

Article

Advancing SWAT Model Calibration: A U-NSGA-III-Based Framework for Multi-Objective Optimization

Huihui Mao ¹, Chen Wang ¹, Yan He ¹, Xianfeng Song ^{1,*}, Run Ma ², Runkui Li ¹ and Zheng Duan ³

¹ College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 101408, China; maohuihui16@mails.ucas.ac.cn (H.M.); wangchen223@mails.ucas.ac.cn (C.W.); heyan211@mails.ucas.ac.cn (Y.H.); lirik@ucas.ac.cn (R.L.)

² Artificial Intelligence Key Laboratory of Sichuan Province, Sichuan University of Science & Engineering, Yibin 644005, China; marun@suse.edu.cn

³ Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, 223 62 Lund, Sweden; zheng.duan@nateko.lu.se

* Correspondence: xfsong@ucas.ac.cn

Abstract: In recent years, remote sensing data have revealed considerable potential in unraveling crucial information regarding water balance dynamics due to their unique spatiotemporal distribution characteristics, thereby advancing multi-objective optimization algorithms in hydrological model parameter calibration. However, existing optimization frameworks based on the Soil and Water Assessment Tool (SWAT) primarily focus on single-objective or multiple-objective (i.e., two or three objective functions), lacking an open, efficient, and flexible framework to integrate many-objective (i.e., four or more objective functions) optimization algorithms to satisfy the growing demands of complex hydrological systems. This study addresses this gap by designing and implementing a multi-objective optimization framework, Py-SWAT-U-NSGA-III, which integrates the Unified Non-dominated Sorting Genetic Algorithm III (U-NSGA-III). Built on the SWAT model, this framework supports a broad range of optimization problems, from single- to many-objective. Developed within a Python environment, the SWAT model modules are integrated with the Pymoo library to construct a U-NSGA-III algorithm-based optimization framework. This framework accommodates various calibration schemes, including multi-site, multi-variable, and multi-objective functions. Additionally, it incorporates sensitivity analysis and post-processing modules to shed insights into model behavior and evaluate optimization results. The framework supports multi-core parallel processing to enhance efficiency. The framework was tested in the Meijiang River Basin in southern China, using daily streamflow data and Penman–Monteith–Leuning Version 2 (PML-V2(China)) remote sensing evapotranspiration (ET) data for sensitivity analysis and parallel efficiency evaluation. Three case studies demonstrated its effectiveness in optimizing complex hydrological models, with multi-core processing achieving a speedup of up to 8.95 despite I/O bottlenecks. Py-SWAT-U-NSGA-III provides an open, efficient, and flexible tool for the hydrological community that strives to facilitate the application and advancement of multi-objective optimization in hydrological modeling.

Keywords: SWAT model; multi-objective optimization; parallel processing; U-NSGA-III; parameter calibration; sensitivity analyses



Citation: Mao, H.; Wang, C.; He, Y.; Song, X.; Ma, R.; Li, R.; Duan, Z. Advancing SWAT Model Calibration: A U-NSGA-III-Based Framework for Multi-Objective Optimization. *Water* **2024**, *16*, 3030. <https://doi.org/10.3390/w16213030>

Academic Editor: Avi Ostfeld

Received: 31 August 2024

Revised: 10 October 2024

Accepted: 18 October 2024

Published: 22 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hydrological models serve as essential tools for simulating natural system processes, playing a crucial role in watershed management and decision-making [1]. These models assist researchers in achieving a profound understanding of the hydrological cycle within the Earth system, including simulations of water quantity, water quality, and studies on climate change [2–5]. In recent years, distributed hydrological models have gained increasing attention due to their ability to capture spatial heterogeneity [6,7]. These models are typically characterized by their intricate structures and numerous parameters. While some

parameters can be derived from measurements or observations, many remain unobservable and necessitate calibration for determination [8,9]. In automated model calibration, genetic algorithms (GAs), multi-objective algorithms based on the Pareto concept, and many-objective optimization algorithms proposed in recent years have been experienced in previous studies [10–13]. However, the calibration process typically demands extensive iterative optimization, significantly increasing the computational burden of the model. As a result, calibrating hydrological model parameters remains one of the most challenging tasks in the hydrological community [14]. Traditional calibration methods that rely solely on streamflow observations are susceptible to the problem of equifinality, where different parameter sets yield similar or acceptable model performance [15]. Consequently, the efficient and accurate calibration of hydrological models has emerged as a critical issue that needs addressing.

The SWAT is a continuous, long-term, distributed, and physically based hydrological model [16]. Since its release, it has been adopted due to its open-source nature [4,17]. Various optimization algorithms have been proposed for SWAT calibration. Among them, SWAT-CUP is the most commonly used calibration tool and contains multiple optimization algorithms, primarily single-objective ones [18]. Many classical optimization algorithms have also been applied in SWAT models [19,20], including GAs [21], Shuffled Complex Evolution (SCE-UA) [22,23], and its improved algorithm Shuffled Complex Evolution Metropolis algorithm (SCEM-UA) [24], as well as the Dynamically Dimensioned Search (DDS) algorithm [25]; however, these algorithms were also developed to solve single-objective optimization problems. Single-objective approaches often use aggregation functions when dealing with multi-objective problems, where results depend on the weighting of different objective functions, which introduces subjectivity and can disguise trade-offs among competing criteria [26]. Studies have demonstrated that there are often conflicts and trade-offs between calibration objectives in hydrological models and that single performance measures (e.g., root-mean-square error (RMSE)) are insufficient to characterize the complex behavior of a watershed, necessitating multi-objective optimization (MOO) to identify a set of Pareto optimal solutions [12,26–28]. In practice, MOO facilitates the consideration of multiple, potentially competing objectives, such as streamflow at multiple gauges, multiple variables (such as streamflow, ET, leaf area index (LAI), and soil moisture (SM)), and multiple objective functions (such as Nash-Sutcliffe Efficiency (NSE), coefficient of determination (R^2), Kling-Gupta Efficiency (KGE), and Percentage Bias (PBIAS)) [29–35]. Multi-source data, such as remote sensing data, enrich observations for hydrological models to help mitigate the equifinality problem [36–39]. In addition, MOO algorithms can simultaneously optimize multiple objectives by incorporating one or more hydrological variables from remotely sensed data, thereby improving model accuracy [37]. Evolutionary algorithms, represented by genetic algorithms, are inherently capable of parallel processing, which can accelerate model execution and alleviate the computational burden in SWAT model parameter calibration [40], making them increasingly popular in the hydrological community [41].

Recent advances have seen the extensive application of multi-objective algorithms in calibrating the SWAT model. Researchers have explored various multi-objective algorithms, including the Non-dominated Sorted Genetic Algorithm II (NSGA-II), the multi-algorithm, the genetically adaptive multi-objective method (AMALGAM), Borg, and U-NSGA-III, within the SWAT framework. For example, Bekele and Nicklow [29] calibrated the SWAT model for streamflow and sediment using the Non-dominated Sorted Genetic Algorithm II (NSGA-II) algorithm and found that multi-objective calibration improved model performance. Ercan and Goodall [14] developed a calibration library based on the NSGA-II algorithm and demonstrated its effectiveness for SWAT model calibration. Zhang, et al. [42] introduced the PP-SWAT tool for efficient SWAT model calibration based on the AMALGAM algorithm [43], achieving a speedup of 45–109 times on a Linux computer cluster. Besalatpour, et al. [44] used parallel processing techniques in combination with the AMALGAM multi-objective optimization algorithm for multi-site calibration of SWAT model

parameters, which accelerated the computational speed of the model. Zhang, et al. [45] further compared the performance of the AMALGAM, Strength Pareto Evolutionary Algorithm 2 (SPEA2), and NSGA-II algorithms in the SWAT model and found that AMALGAM was superior. Chilkoti, et al. [33] used hydrological signatures as objective functions in a multi-objective optimization framework and employed the Borg multi-objective evolutionary algorithm (MOEA) to calibrate the SWAT model; their results showed that combining hydrological signatures with conventional statistical functions improved the accuracy of low-flow simulations. Hernandez-Suarez, et al. [13] applied the U-NSGA-III algorithm for multi-objective calibration of ecologically relevant hydrologic indices (ERHIs) in the SWAT model. Herman, et al. [34] used the U-NSGA-III algorithm in the SWAT model, considering streamflow and remote sensing actual evapotranspiration (ETa) simultaneously for multi- and many-objective calibration, and verified the efficacy of the U-NSGA-III algorithm in the SWAT model. To overcome the limitations of parallel processing in the SWAT-CUP software, alternative tools have emerged. For instance, Zamani, et al. [46] developed the LCC-SWAT system, a cloud-based framework that integrates the SUFI-2 and DDS algorithms for parallel runs.

Despite significant progress in the multi-objective calibration of the SWAT model, several limitations persist within existing calibration tools. Firstly, some tools are not publicly available, including optimization algorithms or parallel modules [44,47–51]. Secondly, mainstream calibration tools (e.g., SWAT-CUP) either lack evolutionary multi-objective algorithms [52] or fail to support many-objective calibration [14,18,46,53], limiting their ability to calibrate multiple objectives simultaneously. The increasing incorporation of multiple remote sensing hydrological variables into hydrological models simultaneously has the potential to give rise to complex many-objective optimization challenges [38,54–57]. However, traditional multi-objective algorithms like NSGA-II may struggle to search for many-objective Pareto solutions [58,59]. Finally, some existing tools are not flexible enough to meet the customized needs of users [52]. Consequently, there is an urgent need to develop an open, efficient, and flexible multi-objective calibration framework for the SWAT model that supports many-objective calibration to address the evolving needs of the hydrological community utilizing multi-source remote sensing data.

The NSGA-III algorithm holds promise for many-objective optimization, but previous research has found that its performance may degrade when dealing with one or two objectives [60]. In response, the U-NSGA-III algorithm was developed to be suitable for optimization problems with single objectives, multiple objectives, and many objectives. This study aims to construct a multi-objective calibration framework for the SWAT model that integrates the U-NSGA-III algorithm, named Py-SWAT-U-NSGA-III, which applies from single-objective to many-objective calibration. This framework is developed on Python and is openly accessible. It also supports parallel processing, which can improve the running efficiency of the model. Furthermore, the Python-based platform allows SWAT researchers the flexibility to make customized modifications tailored to their specific needs.

To further broaden the application scenarios of the Py-SWAT-U-NSGA-III framework, this study introduces multiple calibration schemes, including multi-site, multi-variable, and multi-objective function calibrations. Among them, the multi-variable calibration incorporates remotely sensed hydrological variables, such as ET, LAI, and biomass, which can improve the accuracy of SWAT model sub-processes. The framework supports sensitivity analysis, enabling time-varying sensitivity analysis for daily data, and provides post-processing modules (Multi-Criteria Decision Making (MCDM), Result Output/Result Plotting, etc.) to facilitate user-friendly analysis and understanding. In summary, Py-SWAT-U-NSGA-III is expected to provide an effective tool for the SWAT community, promoting broader and deeper applications of multi-objective optimization algorithms in the SWAT model.

This paper is organized as follows: Section 2 introduces the SWAT model, the U-NSGA-III algorithm, and the overall design and specific implementation of the Py-SWAT-U-NSGA-III framework. Section 3 conducts sensitivity analysis, evaluates the parallel

efficiency of the framework, and presents three typical case studies (including multi-site, multi-variable, and multi-objective functions calibration) in the Meijiang River Basin in southern China. Section 4 discusses the results and summarizes the advantages, limitations, and future directions of the framework. Finally, Section 5 presents the conclusions. We have provided the source code for the Py-SWAT-U-NSGA-III framework as an open-source and freely available repository through GitHub: https://github.com/Huihui-Mao/Py-SWAT_U-NSGA-III/tree/master (accessed on 30 August 2024).

2. Methods

2.1. SWAT Model

The SWAT model is a physically based continuous-time long-term semi-distributed watershed-scale hydrological model developed by the Agricultural Research Service of the United States Department of Agriculture (USDA-ARS) [16,61]. The SWAT was developed to predict the impact of land management practices on water, sediment, and agricultural chemical yields in large complex watersheds with varying soils, land use, and management conditions over long periods [61]. Due to its open-access nature, it has been applied widely in a variety of watershed studies around the world, including simulations of both water quality and quantity [4,17]. The SWAT model divides a watershed into subbasins and then further into hydrological response units (HRUs) based on a unique combination of land use, soil type, and land slope. The hydrological cycle simulated by SWAT is divided into two major phases: the land phase and the water or routing phase [61]. In the land phase, the fluxes of water, sediment, nutrients, and pesticide loads are first calculated for each HRU and then calculated by aggregating the fluxes and loads of the HRUs in the subbasin for each subbasin; the results are used as the input for the main channel in the subbasin. The routing phase controls the movement of water, sediments, etc., through the channel network of the watershed to the outlet.

2.2. U-NSGA-III Algorithm

Over the past two decades, multi-objective evolutionary algorithms (MOEAs) have seen widespread use in hydrological modeling applications [13,26,28,29,41]. In particular, the population-based Non-dominated Sorting Genetic Algorithm II (NSGA-II) [62] has received considerable attention for its effectiveness and efficiency in addressing multi-objective optimization problems [14,29,32,63,64]. However, previous research has highlighted the limitations of NSGA-II in solving problems with more than three objective functions [58]. To address these limitations, NSGA-III, an elitist reference-point-based algorithm, has emerged as a viable solution for many objective problems [58]. The main difference between the two algorithms is the niching method, NSGA-II, which uses crowding distances, whereas NSGA-III employs reference directions, which are vectors that uniformly fill the objective space [58]. While NSGA-III's performance diminishes when handling one or two objective functions, the U-NSGA-III algorithm addresses this issue by integrating an explicit selection procedure, allowing it to effectively tackle single-, multi-, and many-objective optimization problems without introducing additional parameters [60]. Moreover, U-NSGA-III can handle unconstrained and constrained problems [65]. Accordingly, the U-NSGA-III algorithm has been selected for implementation in this study.

The U-NSGA-III algorithm has been used in previous studies for hydrological modeling, with applications including multivariate calibration [34] and multi-objective calibration [13,66]. In this study, a maximum of 200 generations was set as the stopping criterion for the multi-objective optimization. The population size was set at 100 for the case studies and 50 for the model efficiency evaluation, with reference directions equal to 100 and 50, respectively. The reference directions were generated using the Riesz s -Energy method [67], included in the Pymoo library. This resulted in a total number of 20,000 model evaluations (5000 model evaluations for the model efficiency evaluation). The U-NSGA-III operators selected were the simulated binary crossover (SBX) and polynomial mutation (PM). The crossover probability and distribution index for the SBX operator and the mutation prob-

ability and distribution index for the PM operator were defined as 0.9, 10, 1/14 (i.e., the reciprocal of the number of calibration parameters), and 20, respectively. Latin Hypercube Sampling (LHS) is selected as the parameter sampling method [68].

2.3. *Py-SWAT-U-NSGA-III: A Framework for Parallel Calibration of the SWAT Model with U-NSGA-III*

2.3.1. Overall Design

Py-SWAT-U-NSGA-III was developed in Python 3.11 and is compatible with multiple platforms, including Windows, Linux, and macOS. It comprises three modules, as illustrated in Figure 1. The first module is the main program of the framework. It involves setting parameters for the U-NSGA-III algorithm, such as population size, sampling method, crossover and mutation operator parameters, number of generations, as well as multi-criteria decision-making (MCDM), convergence analysis, result output and plotting, and sensitivity analysis. The second module comprises various submodules related to the execution of the SWAT model, including tasks such as parameter file reading, model control file modification/project file copying to the parallel running directory, observation data reading, parameter file modification, model running, model simulation data reading, and objective function calculation. The computationally expensive SWAT-related submodules were parallelized using the Python multiprocessing function, except for parameter file reading. These modules were integrated with Pymoo [69] to implement the U-NSGA-III algorithm. Pymoo is a multi-objective optimization framework in Python that offers state-of-the-art single- and multi-objective optimization algorithms. The third module defines the optimization problem within the Pymoo library, including the number of design variables, the number of objective functions, the number of constraints (optional), the upper and lower bounds of the design variables, etc. Moreover, to accommodate a wide range of applications, the second module offers three calibration schemes for multi-site, multi-variable (such as ET, LAI, and biomass), and multi-objective functions, as well as two spatial scales (HRU and subbasin) and three temporal scales (daily, monthly, and yearly).

2.3.2. Implementation of the Py-SWAT-U-NSGA-III Framework

This section provides further elaboration on the implementation of the three modules in the Py-SWAT-U-NSGA-III framework. The usage of the U-NSGA-III algorithm in the first module and the definition of optimization problems in the third module can be found on the Pymoo official website (<https://pymoo.org/index.html>, accessed on 21 October 2024). The implementation details are as follows:

In the first module, as described in the previous section, the first module serves as the main program of the Py-SWAT-U-NSGA-III framework.

1. **Multi-Criteria Decision Making (MCDM):** This is used to obtain the best tradeoff solution after multi-objective optimization, which includes two methods: compromise programming [70] and pseudo-weights [71]. The compromise programming method uses user-defined distance metrics to determine the Pareto optimal solution closest to the reference point; typically, the reference point is the ideal point, representing the best-expected objective function values [13]. Meanwhile, the pseudo-weights method generates a vector for each Pareto optimal solution, representing the relative importance (or weights) of each objective function [13]. The sum of the different weights in each vector is forced to one. In this framework, the compromise programming method is the default option;
2. **Convergence Analysis:** This employs the hypervolume indicator [72], a commonly used metric for evaluating the approximations of the Pareto front (PF) generated by MOEAs [73]. A higher value of the hypervolume indicator indicates a better solution and the ideal value is one. The definition of the hypervolume indicator, its calculation, and cases in the space of the 2 objective functions are provided in S1 and Figure S1 in the Supplementary Materials;

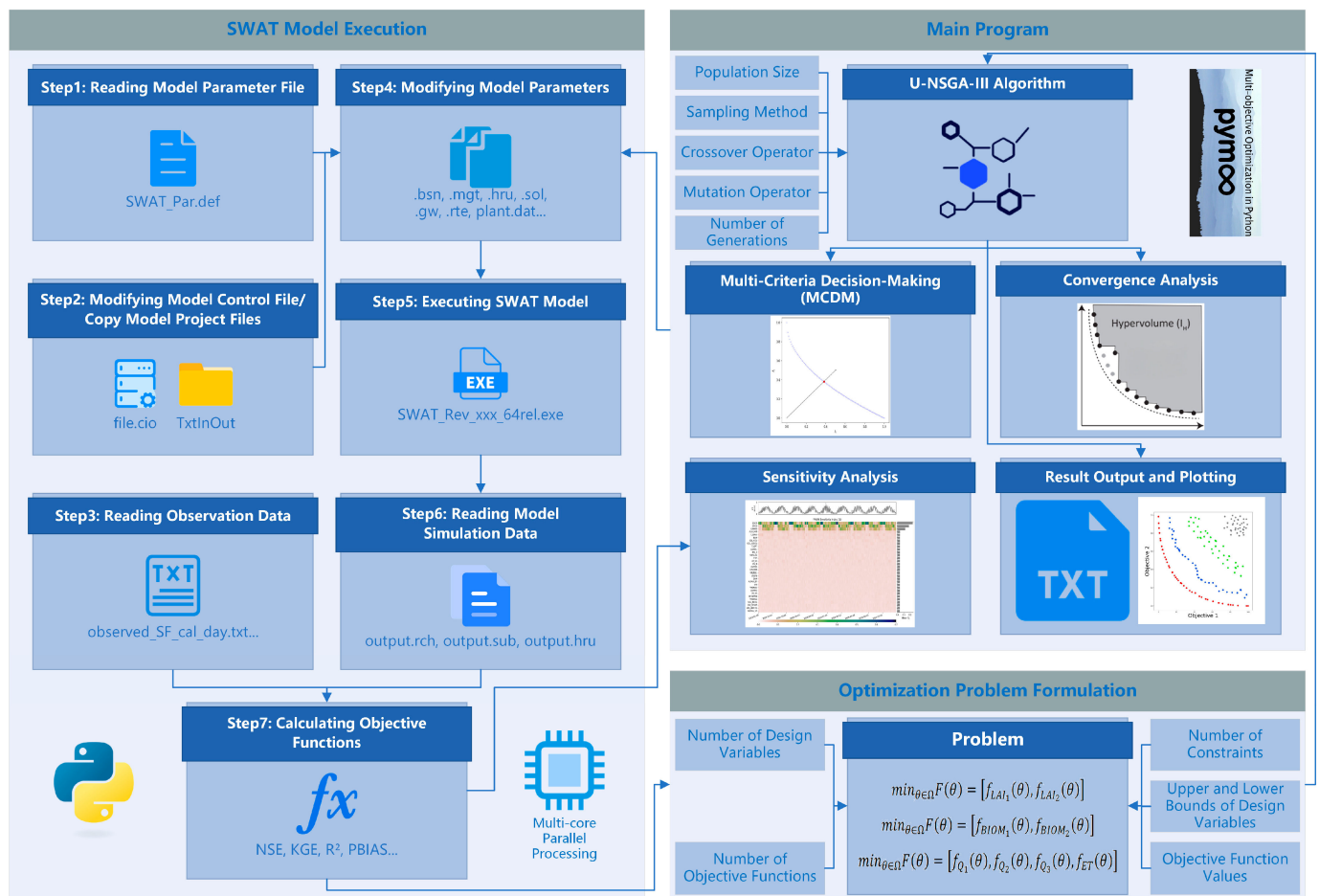


Figure 1. Schematic diagram of Py-SWAT-U-NSGA-III (The MCDM in the figure represents the selection of the trade-off solution from a set of Pareto solutions, see the introduction of Multi-Criteria Decision Making (MCDM) in Section 2.3.2 for details; the Result Output/Result Plotting in the figure give the process of the multi-objective algorithm in the iterative process to find the Pareto solution, which is getting closer and closer to the Pareto solution for the increase of the number of generations; the grey colour in the figure represents the distribution of the initial solution, the green colour and the blue colour represent the distribution of the solution in the iterative process, and the red colour represents the final Pareto solution).

- 3. Result Output/Result Plotting:** The result output section outputs the Pareto and compromise solutions, parameter values during each generation, the number of non-dominated solutions during each generation, and the hypervolume during each generation to a text file. The result plotting section includes the hypervolume plot during each generation, the number of non-dominated solutions plot during each generation, the objective space distribution plot of Pareto and compromise solutions, the parallel coordinate plot of parameter values corresponding to non-Pareto, Pareto, and compromise solutions, time series plots of optimization variables (such as streamflow, leaf area index (LAI), ET, biomass), and flow duration curve plots;
- 4. Sensitivity Analysis:** This includes Sobol' [74] and PAWN [75] methods while time-varying sensitivity analysis (TVSA) is also supported for daily-scale long-time series data. The Sobol' method is one of the most popular variance-based global sensitivity analysis methods [76,77]. It is a global analysis method based on variance decomposition, attributing the variance of model outputs to individual parameters and their interactions. For detailed information on using Sobol's method to calculate hydrological model sensitivity, refer to [76,77]. One major limitation of variance-based sensitivity indices is that they require a large number of output samples to obtain

accurate estimates; another major limitation is that they implicitly assume that output variance is a sensible measure of the output uncertainty, which might not always be the case [75]. To address these issues, a density-based GSA method, PAWN [75,78], has been proposed. PAWN uses the entire model output distribution, rather than just its variance, to quantify the relative influence of parameters on model outputs. By definition, the PAWN method is a moment-independent GSA method. It is very easy to implement, and the analysis of the robustness and convergence of PAWN sensitivity indices is computationally efficient.

In the second module, the SWAT model project files need to be prepared in advance, specifically the 'TxtInOut' directory, which contains various parameter files, model control file, model simulation output files, meteorological data, etc. The second module includes the following steps:

1. **Reading Model Parameter File:** Read the parameter names and corresponding ranges from the model parameter file (SWAT_Par.def). The definition of model parameters in this file is mainly according to the specifications in the SWAT-CUP software [18,52]. Parameters of the same type (determined by the parameter file suffix, e.g., .mgt, .gw, .sol, .rte, .hru, .bsn, .sub, and .plant.dat) are stored in a Python dictionary. The parameterization methods and value ranges of all parameters are also stored;
2. **Modifying Model Control File/Copy Model Project Files (in parallel):** Based on user settings (such as simulation years, simulation start year, simulation output time interval, and warm-up period), modify the model control file (file.cio), then copy the model TxtInOut project folder to the parallel run directory according to the population size of the U-NSGA-III algorithm;
3. **Reading Observation Data (in parallel):** Read the observation data, including the subbasin and HRU scale. For spatial distribution data, a watershed weighted average value is used, and the calculation formula is as follows, taking ET data as an example. For daily-scale spatial distribution data, the framework provides zonal statistics and outputs them to text files.

$$\overline{ET}_j = \frac{1}{A_T} \sum_{i=1}^n A_i ET_{ij}, \quad (1)$$

where $\overline{ET}_j$ is the average ET for month (or day) j ; A_T is the total surface area of the watershed; A_i is the surface area of subbasin (or HRU) i ; ET_{ij} is the ET for subbasin (or HRU) i and month (or day) j ; and n is the number of subbasins (or HRUs);

4. **Modifying Model Parameters (in parallel):** As with SWAT-CUP, it incorporates three parameterization methods (r for percentage change; v for replacement; and a for addition; see the SWAT-CUP manual for further details), which facilitates the alteration of diverse types of parameters, including those about vegetation, soil, management, groundwater, and so on. The initial parameterization values are derived from LHS, and subsequently from the optimized values of the previous generation;
5. **Executing the Model (in parallel):** Call the SWAT.exe executable in each parallel run directory;
6. **Reading Model Simulation Data (in parallel):** Read the data from the model output files (output.rch, output.sub, output.hru). Similar to step 3, a watershed weighted average value for spatial distribution data is established;
7. **Calculating Objective Functions (in parallel):** Calculate the objective functions based on the observed and simulated data. The currently available objective functions include NSE [79], KGE [80], R^2 , PBIAS [81], and Root Mean Square Error (RMSE). They are calculated as the following:

$$NSE = 1 - \frac{\sum_{i=1}^n (S_i - O_i)^2}{\sum_{i=1}^n (O_i - \overline{O_i})^2}, \quad (2)$$

$$KGE = 1 - \sqrt{(1-r)^2 + (1-\alpha)^2 + (1-\beta)^2}, \quad (3)$$

$$R^2 = \frac{[\sum_{i=1}^n (O_i - \bar{O}_i)(S_i - \bar{S}_i)]^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2 \sum_{i=1}^n (S_i - \bar{S}_i)^2}, \quad (4)$$

$$PBIAS = 100 \frac{\sum_{i=1}^n |S_i - O_i|}{\sum_{i=1}^n O_i}, \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (S_i - O_i)^2}{n}}, \quad (6)$$

where S_i and O_i are the simulated and observed data; $\bar{S}_i$ and $\bar{O}_i$ are the means of simulated and observed data; r is the linear regression coefficient between simulated and measured data; $\alpha = \frac{\sigma_s}{\sigma_o}$, $\beta = \frac{\mu_s}{\mu_o}$, μ_s and μ_o are the means of simulated and measured data; σ_s and σ_o are the standard deviation of simulated and measured data; and n is the number of observations. Here, PBIAS uses absolute values for calculation.

The third module involves defining the optimization problem. In this module, besides setting the objective functions, constraint conditions can also be set.

The links between the three modules are as follows:

- (1) When performing a sensitivity analysis, the calculation of the objective function values in step 7 in the second module is passed directly to the sensitivity analysis part in the first module;
- (2) When multi-objective optimization is performed, initially, the number of parameters, the upper and lower bounds of each parameter, the number of objective functions, the number of constraints, the parameterized values for each generation, and the value of the objective functions for the current generation of the population computed in the last step in the second module are passed on to the third module. Subsequently, the optimization problem defined in the third module is passed on to the first module to perform the multi-objective optimization. When the model running reaches the maximum number of generations, the optimization result (approximate Pareto solutions) is subjected to a Multi-Criteria Decision Making (MCDM) process (selection of compromise solution from a set of Pareto solutions), followed by Convergence Analysis and Result Output/Result Plotting; otherwise, the optimized parameter values are passed to step 4 in the second module to continue the multi-objective optimization process.

The framework offers flexibility to determine the number of CPU cores used for parallel running, depending on the specific optimization problem, hardware conditions, and desired computational efficiency. The parallel run is performed using the Process Pools in Python multiprocessing (<https://docs.python.org/3/library/multiprocessing.html#module-multiprocessing.pool>, accessed on 21 October 2024), with each CPU core processing running each SWAT model project file (individual in the U-NSGA-III algorithm population) separately. Details of Py-SWAT-U-NSGA-III are provided on the GitHub site.

3. Case Studies

3.1. Study Area

The study area selected in the case study is the Meijiang River Basin (115°36′–116°38′, 26°00′–27°09′), a tributary of the upper Gan River in Jiangxi Province, southeastern China, which contributes to Poyang Lake and then the Yangtze River (Figure 2), with a drainage area of 6341 km². The watershed is dominated by mid-altitude and low mountains, with elevation ranges from 137 to 1425 m and an average elevation of 351.8 m. The watershed has a subtropical monsoon humid climate. The annual average temperature is 17 °C. The annual precipitation is 1628 mm, and the annual average relative humidity is >80%. The soil types are mainly red earth (63.65%), submergenic paddy soils (26.24%), and yellow red earth (7.13%), with the main texture of loam soil. Evergreen forests (needle leaf and broad leaf) and mixed needle-broad leaf forests (67.02%) are the main vegetation types in the watershed, among which *Pinus massoniana*, *Pinus elliottii*, and *Cunninghamia lanceolata* are

the main needle leaves. Cropland accounts for 13.55% of the total, including paddy fields and dry fields, and paddy fields are mainly planted with rice, which is double-cropping. The rest are shrubs (10.95%) and grasses (2.10%), among other vegetation types.

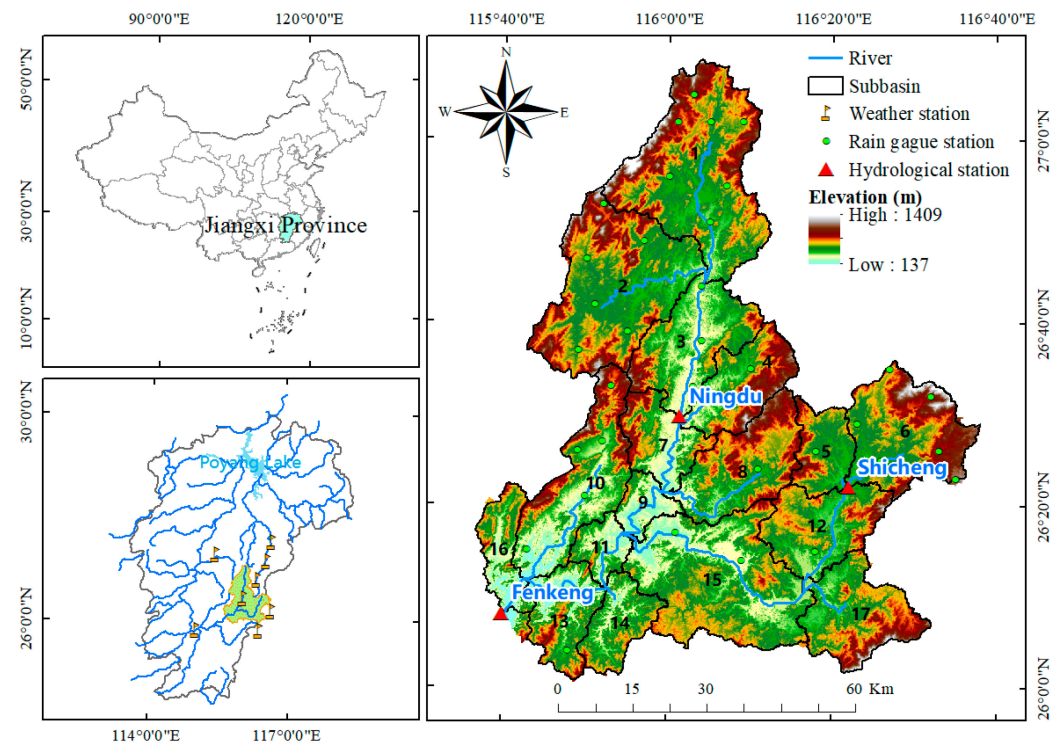


Figure 2. Location of the Meijiang River Basin.

3.2. Input Data and Model Setup

The datasets used in this research are listed in Table 1. The basic datasets required for the SWAT model include topographical data (DEM), land use data, soil data (including soil type and soil attribute data), and meteorological data. There is a difference between the original land use types and the land use types used in the SWAT model. It was necessary to reclassify the land use data of the study area to match the land use type map required by the SWAT model. The soil type data were obtained from the 1:1,000,000 Soil Map of the People's Republic of China, provided by the Resource and Environmental Science Data Platform. Soil attribute data include soil texture, clay content, silt content, sand content, gravel content greater than 2 mm, bulk density, organic carbon content, soil pH, and soil thickness, and vertically include five soil depths: 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, and 60–100 cm. The meteorological data required for the SWAT model include daily precipitation, daily maximum and minimum temperatures, daily solar radiation, daily relative humidity, and daily average wind speed. This study used daily meteorological observation data from 2001 to 2011 from eight meteorological stations (Ganxian, Yongfeng, Nancheng, Nanfeng, Ningdu, Guangchang, Ninghua, and Changting) located within or near the watershed, provided by the China Meteorological Data Service Center. Solar radiation data were calculated using the Ångström-Preseott model based on sunshine duration and empirical parameters [82,83]. Missing values were estimated using the SWAT built-in WXGEN weather generator [61]. Daily observed precipitation data from 37 stations and daily observed streamflow data from three hydrological stations (Ningdu, Shicheng, and Fenkeng) between 2001 to 2011 were also collected from the Annual Hydrological Report (Volume VI: Hydrological Data of Changjiang River Basin). The distribution of meteorological, rainfall, and hydrological stations is shown in Figure 2.

In this study, ArcSWAT was used to configure the SWAT model, with SWAT2012 (revision 692) being employed for the execution of model simulations. The Meijiang River

Basin was divided into 17 subbasins based on DEM data (see Figure 2) and subbasins were further subdivided into 1589 Hydrologic Response Units (HRUs) based on the land use, soil type, and slope of the study area. The slopes were categorized into three classes: 0–18.18%, 18.18–41.66%, and above 41.66%. The initial 2 years (2001–2002) were used as the warm-up period, and the SWAT model was calibrated using daily time steps from 2003 to 2007 and validated from 2008 to 2011. ET data were used at the subbasin scale and synthesized to the watershed scale, as indicated by Equation (1). The hydrological processes in SWAT include surface runoff, ET, percolation, lateral flow, and groundwater. In this study, surface runoff is estimated using the modified version of the Soil Conservation Service curve number (CN) method [84]. The potential ET was estimated using the Penman–Monteith equation. The variable storage routing method is selected for the routing method [85].

Table 1. Datasets used in the SWAT model.

Data Type	Spatial/Temporal Resolution	Source
Digital Elevation Model (DEM)	30 m	ASTER GDEM (https://lpdaac.usgs.gov/products/astgtmv003/ , accessed on 21 October 2024)
Land use	30 m	ChinaCover2010 (Wu, et al. [86], http://www.geodata.cn , accessed on 21 October 2024)
Soil type	30 m	1:1,000,000 Soil Map of the China (https://www.resdc.cn/Default.aspx , accessed on 21 October 2024)
Soil properties	/	Basic soil property dataset of high-resolution China Soil Information Grids (2010–2018) (http://soil.geodata.cn , accessed on 21 October 2024)
Climate	Daily (2001–2011)	CMDC (China Meteorological Data Service Center) (http://data.cma.cn/ , accessed on 21 October 2024)
Rainfall	Daily (2001–2011)	Annual Hydrological Report (Volume VI: Hydrological Data of Changjiang River Basin)
Streamflow	Daily (2001–2011)	Annual Hydrological Report (Volume VI: Hydrological Data of Changjiang River Basin)
ET	Daily (2001–2011)/500 m	PML-V2(China) (He, et al. [87])

3.3. Model Calibration

To evaluate the effectiveness of Py-SWAT-U-NSGA-III, three representative case studies were implemented. In addition, a time-varying sensitivity analysis (TVSA) of daily-scale streamflow and ET was performed using the PAWN method, with a total of 10,000 simulations.

Three calibration schemes are set up as follows:

- (1) Multi-site calibration;
- (2) Multi-variable calibration;
- (3) Multi-objective functions calibration.

Using the Meijiang River Basin as the study area, the first scheme simultaneously calibrates the streamflow of the three hydrological stations (Ningdu, Shicheng, and Fenkeng). The second scheme employs multi-variable calibration using streamflow from the three hydrological stations and PML-V2(China) ET data [87]. The NSE is selected as the objective function for the first and second schemes. By default, U-NSGA-III assigns the same weight to all objectives. As a result, streamflow calibration holds most of the weight in the overall search for this scheme because we have streamflow data from three hydrologic stations. Optimization under these default settings may result in poor performance for ET calibration. Therefore, the weights of streamflow data are adjusted according to previous studies, so that streamflow data and ET data have the same weight [34]. The third scheme involves multi-objective functions calibration using streamflow from the watershed outlet (Fenkeng) and assessing R^2 and PBIAS metrics. Following the multi-objective optimization, MCDM was performed using the compromise programming approach to obtain a compromise so-

lution, and the optimization results were exported and plotted. Each optimization problem was formulated as follows:

$$\min_{\theta \in \Omega} F(\theta) = [f_{Q_1}(\theta), f_{Q_2}(\theta), f_{Q_3}(\theta)], \quad (7)$$

$$\min_{\theta \in \Omega} F(\theta) = [f_{Q_1}(\theta), f_{Q_2}(\theta), f_{Q_3}(\theta), f_{ET}(\theta)], \quad (8)$$

$$\min_{\theta \in \Omega} F(\theta) = [f_{Q_{11}}(\theta), f_{Q_{12}}(\theta)], \quad (9)$$

$$f = 1 - NSE, \quad (10)$$

$$f_{Q_{11}} = 1 - R^2, \quad (11)$$

$$f_{Q_{12}} = PBIAS, \quad (12)$$

where F is a vector composed of multiple objective functions; θ is a vector containing values for p model calibration parameters; Ω is a p -dimensional parameter space limited by the calibration ranges for each model parameter; f_{Q_1} , f_{Q_2} , and f_{Q_3} are the objective functions evaluated for the streamflow of three hydrological stations; f_{ET} is the objective function evaluated for ET; and $f_{Q_{11}}$ and $f_{Q_{12}}$ are the two objective functions for the streamflow of the Fenkeng station.

Based on the results of the sensitivity analysis, 28 sensitive parameters were selected for multi-objective calibration in the model, including the widely used parameters related to surface runoff, groundwater, channel routing, soil water, ET, and biophysical parameters, as shown in Table 2.

Table 2. Parameters used to perform sensitivity analysis in the SWAT model, where the gray area parameters are those with higher sensitivity and are used in the subsequent calibration process (“v_” refers to a replacement by a value from the given parameter range, and “r_” means a relative change to the initial parameter value).

Processes	Parameter	Description	Initial Range
Surface runoff	r_CN2	Initial SCS runoff curve number for moisture condition II	[−0.2, 0.2]
	r_OV_N	Manning’s n value for overland flow	[−0.5, 0.5]
	v_SURLAG	Surface runoff lag coefficient	[0.05, 24]
Groundwater	v_ALPHA_BF *	Baseflow alpha factor (1/days)	[0, 1]
	v_GWQMN *	Threshold depth of water in the shallow aquifer required for return flow to occur (mm H ₂ O).	[0.01, 5000]
	v_GW_REVAP *	Groundwater “revap” coefficient	[0.02, 0.06]
	v_REVAPMN *	Threshold depth of water in the shallow aquifer for “revap” or percolation to the deep aquifer to occur (mm H ₂ O)	[0.01, 500]
	v_RCHRG_DP *	Deep aquifer percolation fraction	[0, 1]
Channel routing	v_CH_N2 *	Manning’s “n” value for the main channel	[0, 0.3]
	v_CH_K2 *	Effective hydraulic conductivity in main channel alluvium (mm/h)	[0, 400]
Soil water	r_SOL_AWC(1)	Available water capacity of the soil layer (mm H ₂ O/mm soil)	[−0.2, 0.2]
	r_SOL_K(1)	Saturated hydraulic conductivity (mm/h)	[−0.2, 0.2]
	r_SOL_BD(1)	Moist bulk density (Mg/m ³ or g/cm ³)	[−0.2, 0.2]
	r_SOL_ALB(1)	Moist soil albedo	[−0.2, 0.2]
Evapotranspiration	v_ESCO *	Soil evaporation compensation factor	[0.01, 1]
	v_EPCO *	Plant uptake compensation factor	[0.01, 1]
	v_CANMX	Maximum canopy storage (mm H ₂ O)	[0, 10]

Table 2. Cont.

Processes	Parameter	Description	Initial Range
Biophysical parameters	v__BIO_E *	Radiation-use efficiency or biomass-energy ratio ((kg/ha)/(MJ/m ²))	[10, 90]
	v__BLAI *	Maximum potential leaf area index	[0.5, 10]
	v__FRGRW1	Fraction of the plant growing season or fraction of total potential heat units corresponding to the 1st point on the optimal leaf area development curve	[0, 0.2]
	v__LAIMX1 *	Fraction of the maximum leaf area index corresponding to the 1st point on the optimal leaf area development curve	[0, 1]
	v__FRGRW2	Fraction of the plant growing season or fraction of total potential heat units corresponding to the 2nd point on the optimal leaf area development curve	[0.25, 1]
	v__LAIMX2 *	Fraction of the maximum leaf area index corresponding to the 2nd point on the optimal leaf area development curve	[0, 1]
	v__DLAI *	Fraction of growing season when leaf area begins to decline	[0.15, 1]
	v__CHTMX *	Maximum canopy height (m)	[0.1, 20]
	v__T_OPT *	Optimal temperature for plant growth (°C)	[11, 38]
	v__T_BASE *	Minimum (base) temperature for plant growth (°C)	[0, 18]
	v__GSI *	Maximum stomatal conductance at high solar radiation and low vapor pressure deficit (m·s ^{−1})	[0, 5]
	v__VPDFR *	Vapor pressure deficit (kPa) corresponding to the second point on the stomatal conductance curve	[1.5, 6]
	v__FRGMAX *	Fraction of maximum stomatal conductance corresponding to the second point on the stomatal conductance curve	[0.001, 1]
	v__ALAI_MIN	Minimum leaf area index for plant during dormant period (m ² /m ²)	[0, 2.5]

Note(s): * The ranges for the factors are based on literature data and the SWAT user manual; Biophysical parameters are biome-specific.

3.4. Efficiency Evaluation Metrics

Speedup [88] was utilized to evaluate the computational efficiency of Py-SWAT-U-NSGA-III. The metric is calculated as follows:

$$S_P = \frac{T_1}{T_P}, \quad (13)$$

where T_1 is the time taken by the program to run on a single processor, and T_P represents the time it takes to run the program on P processors. The speedup metric provides a quantitative measure of the program's acceleration under parallel execution conditions. To perform this evaluation, we conducted several running tests, using a single CPU core as the baseline and a maximum of 60 CPU cores. This comparison can provide insights into identifying optimal computing resource allocation for the framework.

The SWAT model scheme involved was a multi-site streamflow calibration based on three hydrological stations, with 2001–2002 as the warm-up period and 2003–2005 as the calibration period, using daily time steps. For the U-NSGA-III algorithm settings, as described in Section 2.2, the population size was set to 50, and the maximum number of generations was defined to 100, resulting in 5000 model evaluations for each running test. The objective function used was the NSE. Model running speed is inherently influenced by the computer's configuration. In this study, the computational environment for the program is a server running Windows Server 2019 operating system with 80 cores, equipped with two Intel Xeon Silver 4316 @ 2.30GHz processors (Intel Corporation®, Santa Clara, CA, USA), 1024 GB of RAM, and a 480 GB SSD.

3.5. Results

3.5.1. Sensitivity Analyses

To demonstrate the temporal dynamics of SWAT model parameters, a time-varying sensitivity analysis (TVSA) was performed, with the results presented in Figure 3. The parameters were ranked based on their average sensitivity over the entire simulation period. For streamflow data at Ningdu station, the most sensitive parameters included ALPHA_BF, CH_N2, CANMX, CH_K2, RCHRG_DP, ESCO, and EPCO. Similarly, at Shicheng station, ALPHA_BF, CH_K2, CH_N2, CANMX, RCHRG_DP, ESCO, GWQMN, and EPCO exhibited the highest sensitivity. At the watershed outlet, Fenkeng station, ALPHA_BF, CH_N2, CH_K2, CANMX, RCHRG_DP, ESCO, GWQMN, and EPCO were identified as the most sensitive parameters. Regarding ET data, ESCO, EPCO, CANMX, ALAI_MIN, T_BASE, BLAI, and SOL_K(1) were identified as the most sensitive parameters. Notably, the sensitive parameters for streamflow across the three hydrological stations were largely consistent, though their sensitivity rankings varied.

Among the most sensitive parameters, ALPHA_BF, RCHRG_DP, and GWQMN are related to groundwater processes. ALPHA_BF, the baseflow recession constant, significantly impacts baseflow within the model. This parameter demonstrated the highest sensitivity across all three hydrological stations, particularly during periods of low streamflow when baseflow predominates. This finding aligns with the actual physical conditions of the watershed. RCHRG_DP, the aquifer percolation coefficient, modulates water partitioning into deep aquifers. GWQMN, the threshold water level in the shallow aquifer for groundwater contribution to the main channel to occur, affects baseflow.

CH_N2 and CH_K2, both related to the main channel water routing, exhibit higher sensitivity during streamflow peaks. CH_N2, Manning's roughness coefficient, affects the rate and velocity of flow in the channel, whereas CH_K2 influences transmission losses from the channel.

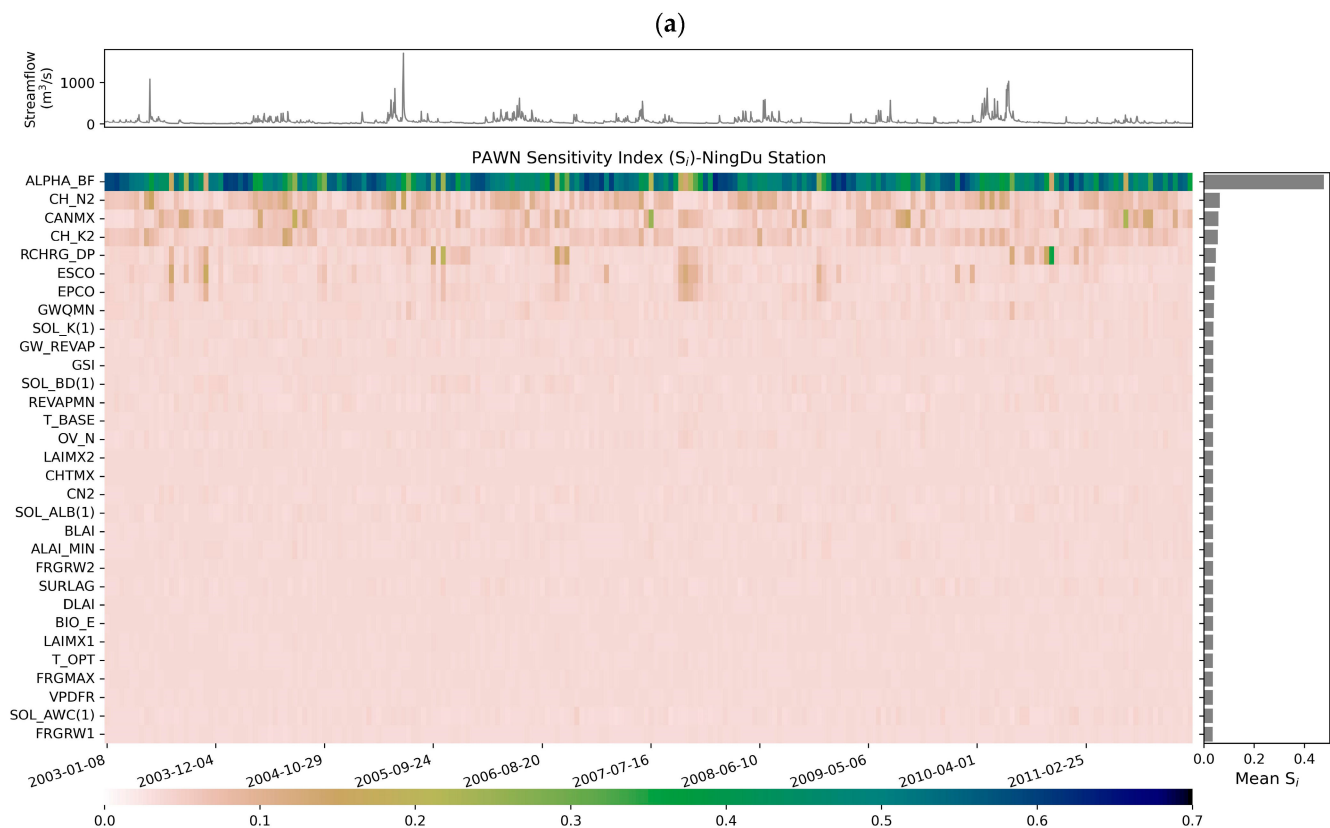


Figure 3. Cont.

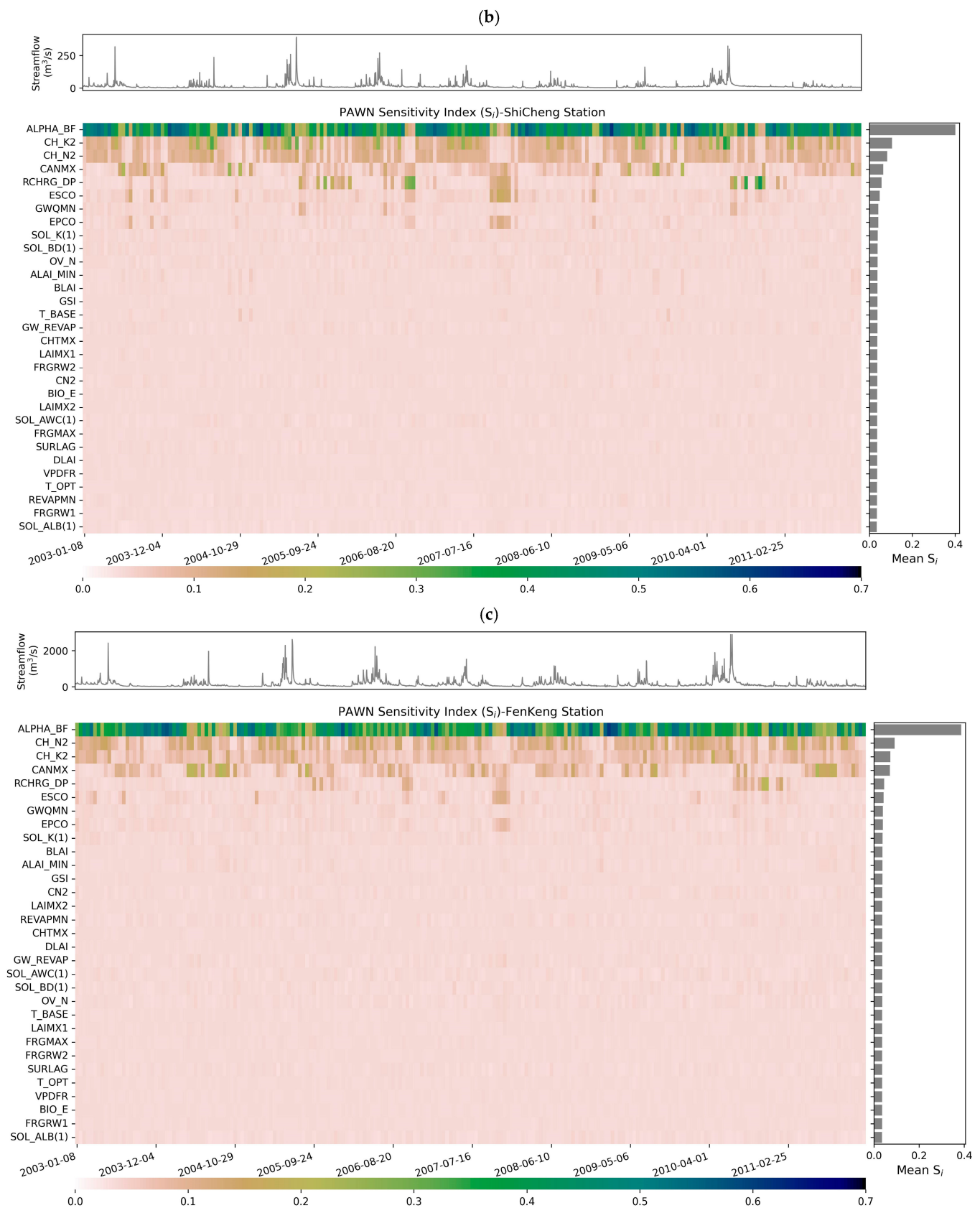


Figure 3. Cont.

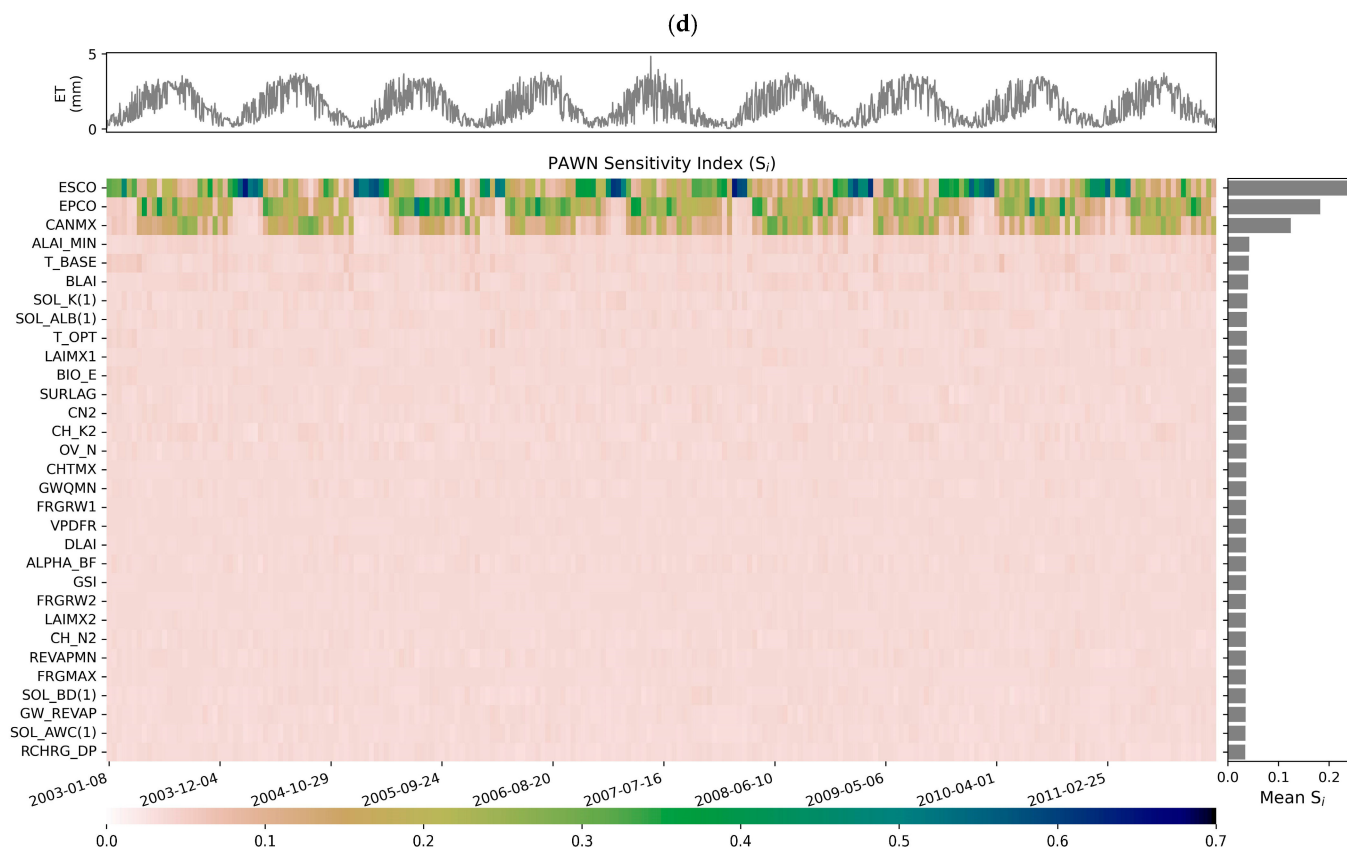


Figure 3. TVSA of each parameter for each model variable: (a–c) streamflow and (d) ET.

CANMX, ESCO, and EPCO pertain to ET processes. CANMX, the maximum canopy storage, affects infiltration, surface runoff, and ET, with peak sensitivity corresponding to ET peaks. ESCO, the soil evaporation compensation factor, influences soil evaporation, with higher ESCO values resulting in lower soil evaporation compensation from lower soil layers, leading to higher water yields [61]. The temporal dynamics of ESCO sensitivity showed higher sensitivity during periods of low ET when soil evaporation constitutes a larger proportion of total ET. Conversely, EPCO, the plant uptake compensation factor, impacts transpiration, with higher EPCO values leading to higher plant transpiration compensation from lower soil layers and then lower water yields. EPCO sensitivity peaked during periods of high ET when plant transpiration dominates.

ALAI_MIN, T_BASE, and BLAI are parameters related to vegetation, closely linked to the plant growth process.

SOL_K(1), the saturated hydraulic conductivity of the first soil layer, is a soil-related parameter that affects infiltration, percolation, and lateral flow.

In conclusion, the sensitivity analysis effectively captured the most influential parameters for each variable within the SWAT model, aligning well with their physical processes. The temporal dynamics of parameter sensitivity also corresponded closely with the observed variations in each variable over time, providing confidence in the model's representation of the physical processes within the watershed.

3.5.2. Efficiency Evaluation

Figure 4 presents the variation in execution time and efficiency of the framework as a function of the number of cores used. In this study, speedup was employed as a metric to evaluate the parallel efficiency of the framework. As depicted in Figure 4, when the number of cores is below 25, there is a remarkable reduction in execution time as the core count increases, leading to a rapid increase in efficiency. However, when the core count exceeds 25, there is a slight increase in execution time and a corresponding decrease in

efficiency. Under the conditions of this study, the optimal core allocation was found to be 25, yielding the highest efficiency with a speedup of 8.95.

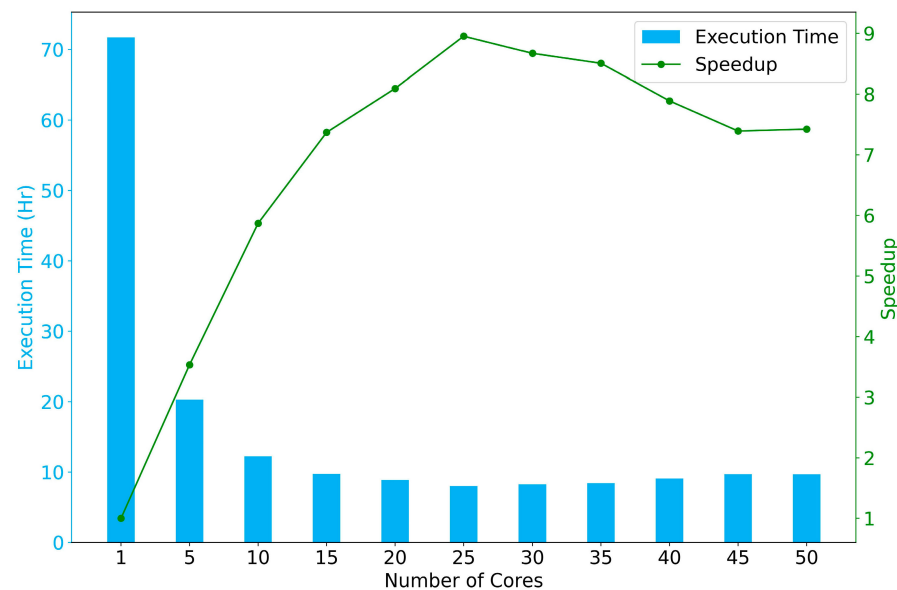


Figure 4. Py-SWAT-U-NSGA-III execution time and parallel speedup as a function of the number of cores used in the run.

3.5.3. Example 1: Multi-Site Calibration

This section evaluates the multi-site calibration capability of the Py-SWAT-U-NSGA-III framework using streamflow data from three hydrological stations, resulting in a multi-objective calibration scheme with three objective functions.

Figure 5 illustrates the distribution of Pareto and compromise solutions within a three-dimensional objective space, as well as the projection of different objective functions in a two-dimensional space. The figure shows a modest trade-off between streamflow at different stations, although the trade-off between Ningdu and Fenkeng stations is relatively weak.

Figure 6 presents the distribution of parameter values corresponding to the Pareto and compromise solutions within the parameter space, depicted using parallel coordinate plots. The results indicate an appreciable narrowing in the parameter ranges for ALPHA_BF, GWQMN, SOL_K(1), ESCO, CANMX, BLAI{8}, and ALAI_MIN{8} across all Pareto-optimal solutions, suggesting a high degree of parameter identifiability for these parameters.

Figure 7 shows the convergence analysis using the hypervolume indicator and the number of non-dominated solutions. After 113 generations, both indicators stabilize (hypervolume indicator around 0.91 and the number of non-dominated solutions at 39), indicating that the model has converged to a near-optimal Pareto front.

Figure 8 describes the time series and mean seasonal dynamics during simulation, including the calibration and validation periods. For streamflow data, flow duration curves for the three hydrological stations are also presented. Figure 8a shows that the optimal compromise solution for streamflow at Ningdu station closely matches the observed data, whereas, at Shicheng station, the optimal solution underestimates the streamflow. At Fenkeng station, streamflow is underestimated during the rising period and overestimated during the recession period. Figure 8b further shows that streamflow at Ningdu station is slightly underestimated under peak and high-flow conditions. It aligns well with the observed data under mid and low-flow conditions. Shicheng station generally shows a larger underestimation trend. At Fenkeng station, streamflow is slightly underestimated under peak and high-low conditions but significantly overestimated under mid- and low-flow conditions.

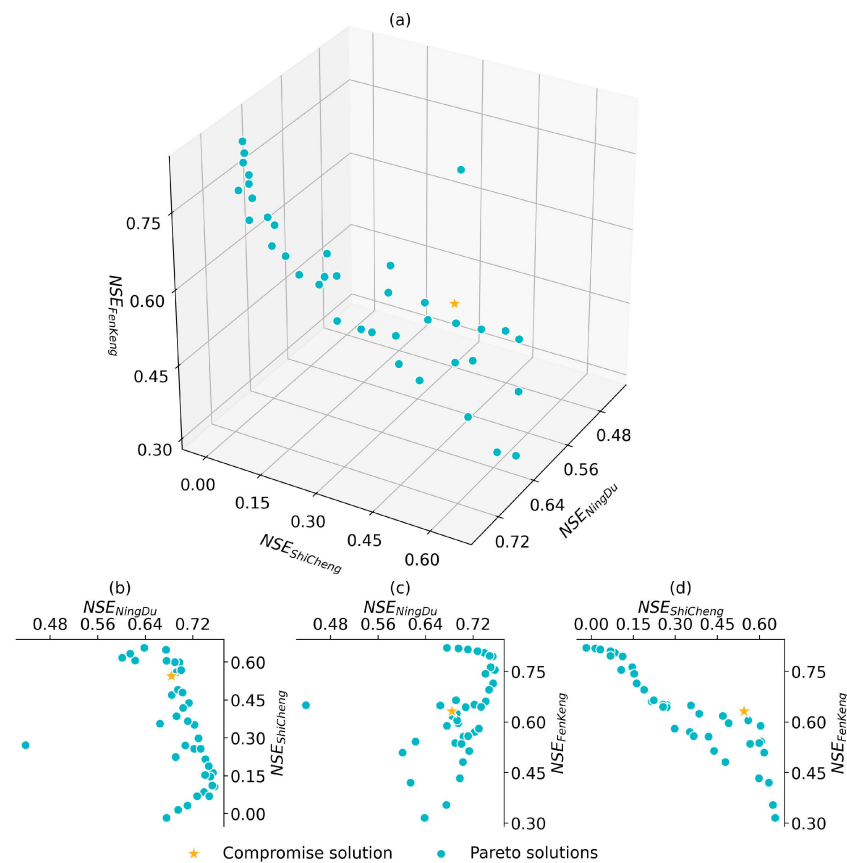


Figure 5. Scatterplot of the distribution of Pareto and compromise solutions in the objective function space: (a) Distribution of Pareto and compromise solutions in three-dimensional objective space, the x, y, and z axes represent the NSE of streamflow at the three hydrological stations, respectively; (b–d) Two-dimensional spatial projections of different objective functions.

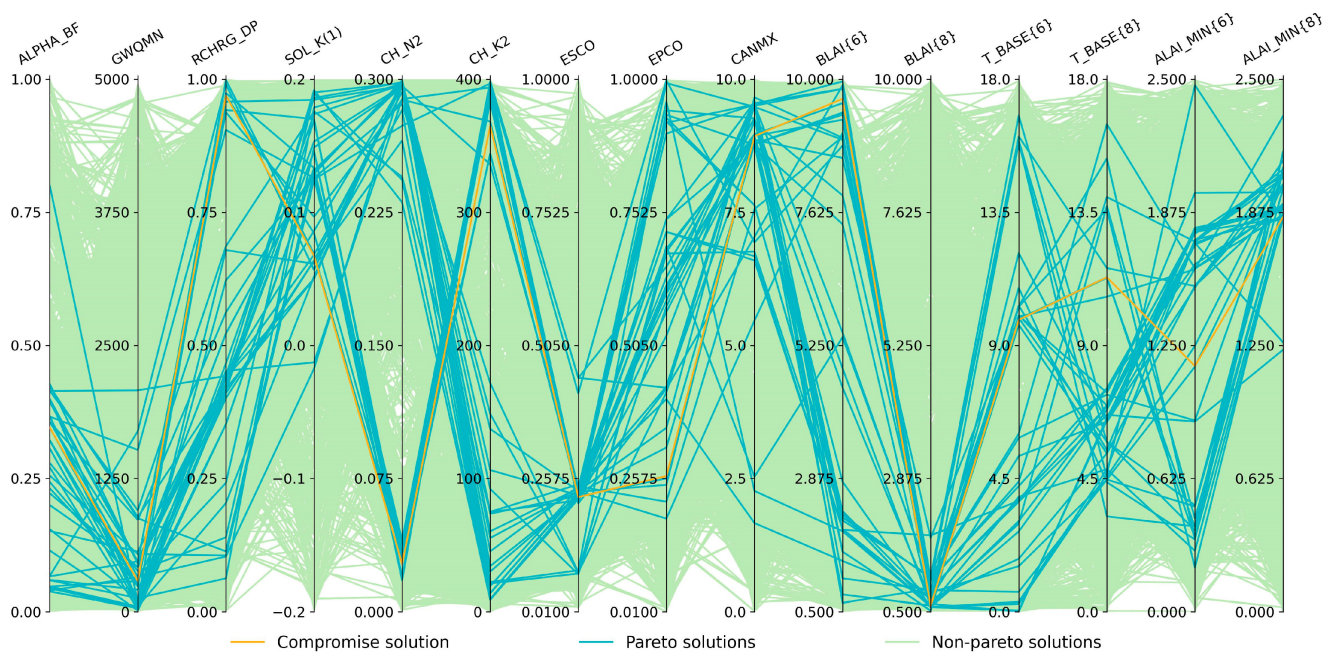


Figure 6. Parallel coordinate plots of Pareto and compromise solutions in parameter space distribution in the multi-site calibration (due to the large number of vegetation-related parameters, only forest-related vegetation parameters are shown).

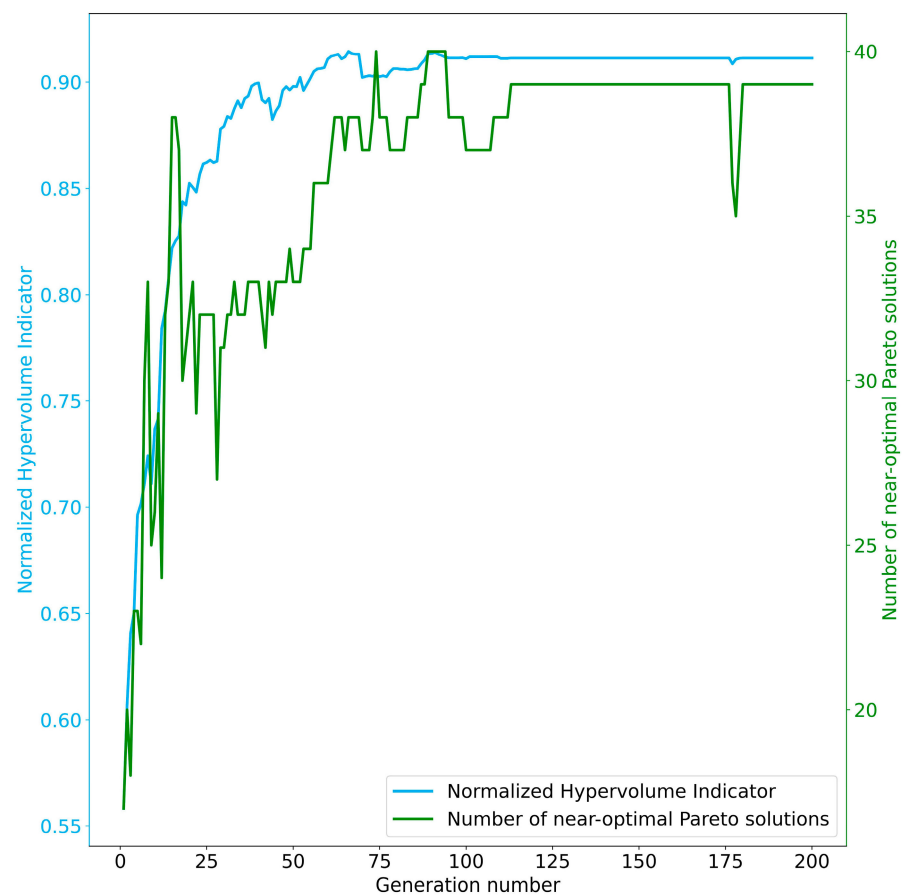


Figure 7. Normalized hypervolume indicator and the number of Pareto solutions as a function of the number of generations in the multi-site calibration.

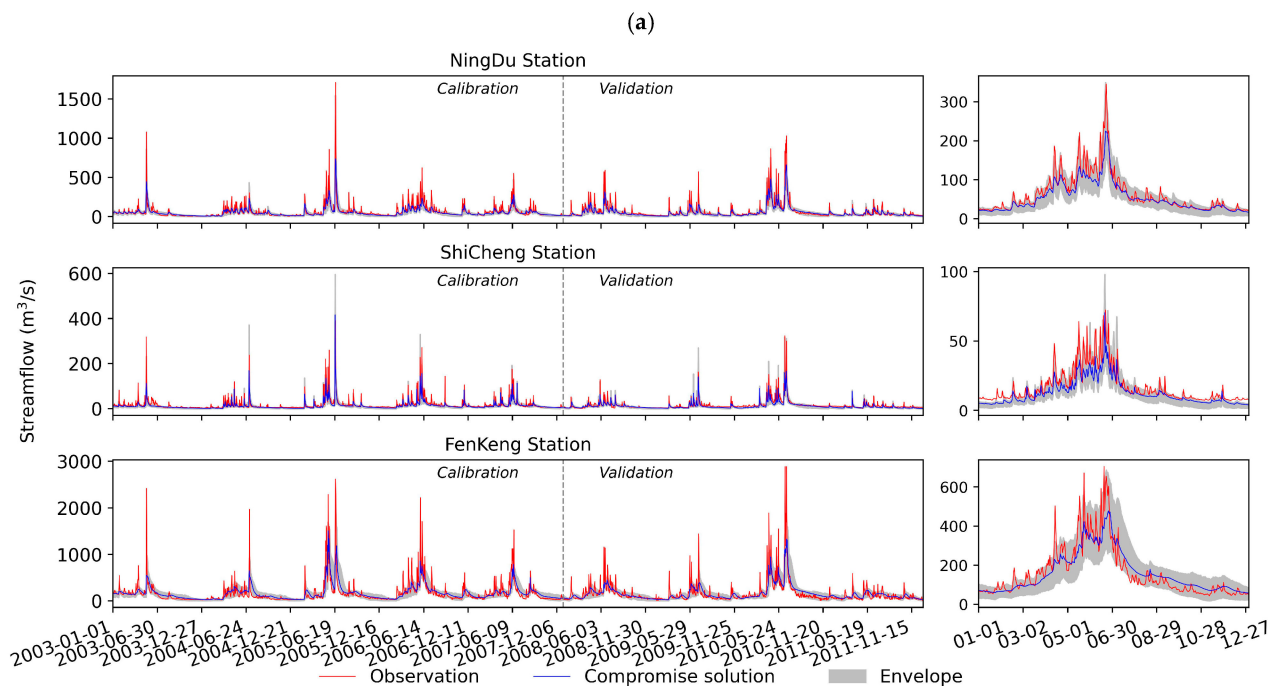


Figure 8. Cont.

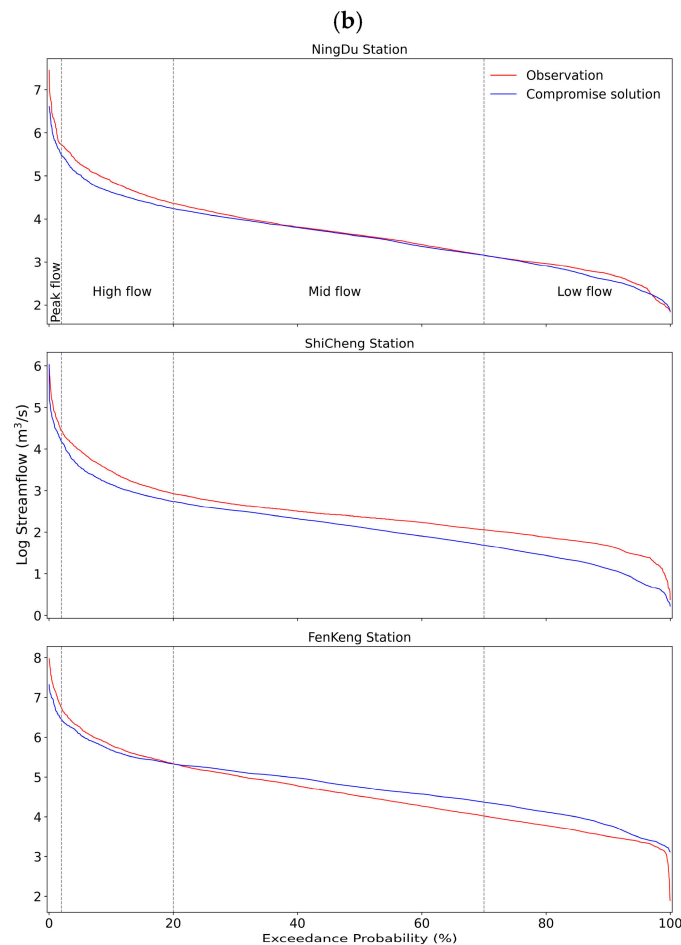


Figure 8. Plotting of calibration variables: (a) Time series curves of streamflow at three hydrological stations, panels on the right show the mean seasonal cycle (the envelope of the grey area represents the uncertainty band encompassing all simulations derived from the Pareto optimal front); (b) Flow duration curves of three hydrological stations.

3.5.4. Example 2: Multi-Variable Calibration

This section evaluates the multivariable calibration capability of the Py-SWAT-UNSGA-III framework, leveraging streamflow data from three hydrological stations and ET data from the watershed scale to formulate four objective functions, constituting a many-objective calibration scenario.

Figure 9 is similar to Figure 5, but with an additional objective function, the fourth objective function is the NSE of the ET, indicated using a gradient color bar. In this three-dimensional space, Figure 9 reveals a good trade-off between streamflow and ET. However, the trade-offs between the streamflow at different stations are less pronounced.

Figure 10 is similar to Figure 6. It highlights a substantial reduction in the parameter ranges for ALPHA_BF, GWQMN, RCHRG_DP, CH_N2, and CH_K2 across all Pareto-optimal solutions, indicating a high level of identifiability for these parameters. Notably, these parameters are also identified as highly sensitive to streamflow, as discussed in Section 3.5.1. This finding is consistent with the study by Wu, et al. [89], which highlighted that identifiability largely depends on the sensitivity or significance of calibration parameters to response variables.

Figure 11 is similar to Figure 7. After 118 generations, both indicators stabilize (the hypervolume indicator at approximately 0.54 and the number of non-dominated solutions at 27), indicating that the model has converged to a near-optimal Pareto front.

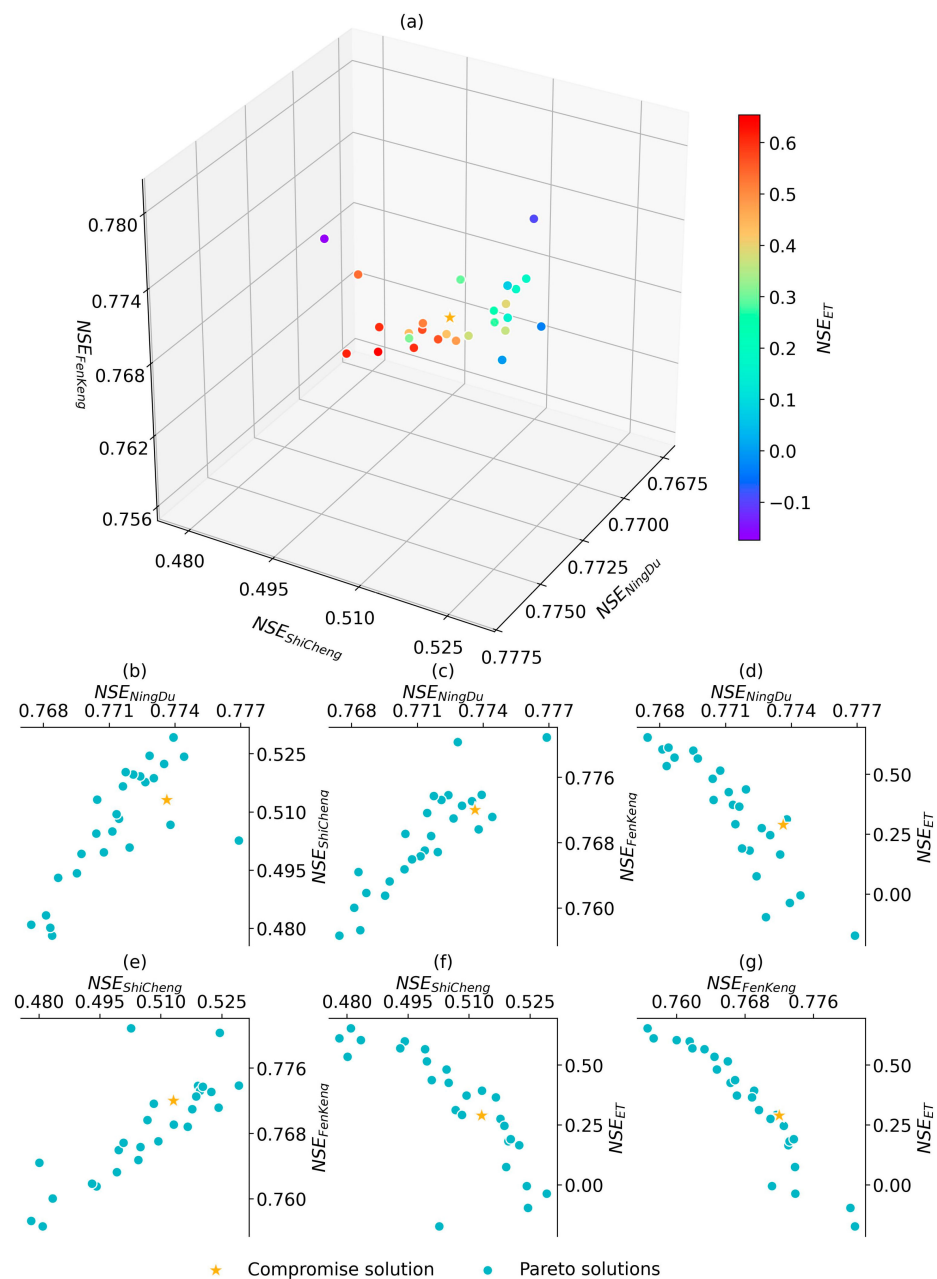


Figure 9. Scatterplot of the distribution of Pareto and compromise solutions in the objective function space: (a) Distribution of Pareto and compromise solutions in three-dimensional objective space, the x, y, and z axes represent NSE of streamflow at the three hydrological stations, respectively, and the NSE of ET is represented using a gradient color bar. (b–g) Two-dimensional spatial projections of different objective functions.

Figure 12 is similar to Figure 8. Figure 12a indicates that the optimal compromise solution for streamflow at three stations and ET across the watershed aligns well with the observed data. Figure 12b further reveals that the streamflow at Ningdu station aligns well with observed data across peak, high, and mid-flow conditions but is underestimated in low-flow conditions. Shicheng Station exhibits similar trends but with slightly more underestimation in the mid-flow range. At Fenken station, streamflow is slightly underestimated at peak flow, overestimated at high and mid-flow, and shows both overestimation and underestimation at low flow.

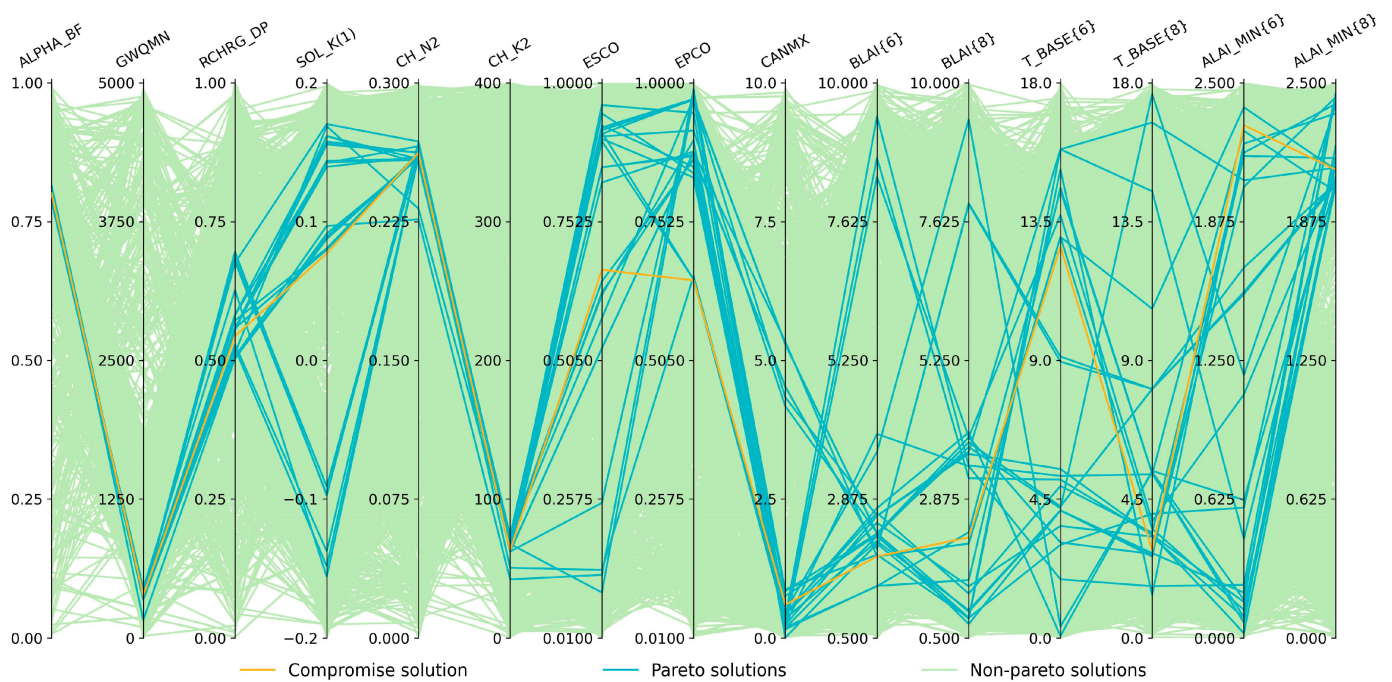


Figure 10. Parallel coordinate plots of Pareto and compromise solutions in parameter space distribution in the multi-variable calibration (due to the large number of vegetation-related parameters, only forest-related vegetation parameters are shown).

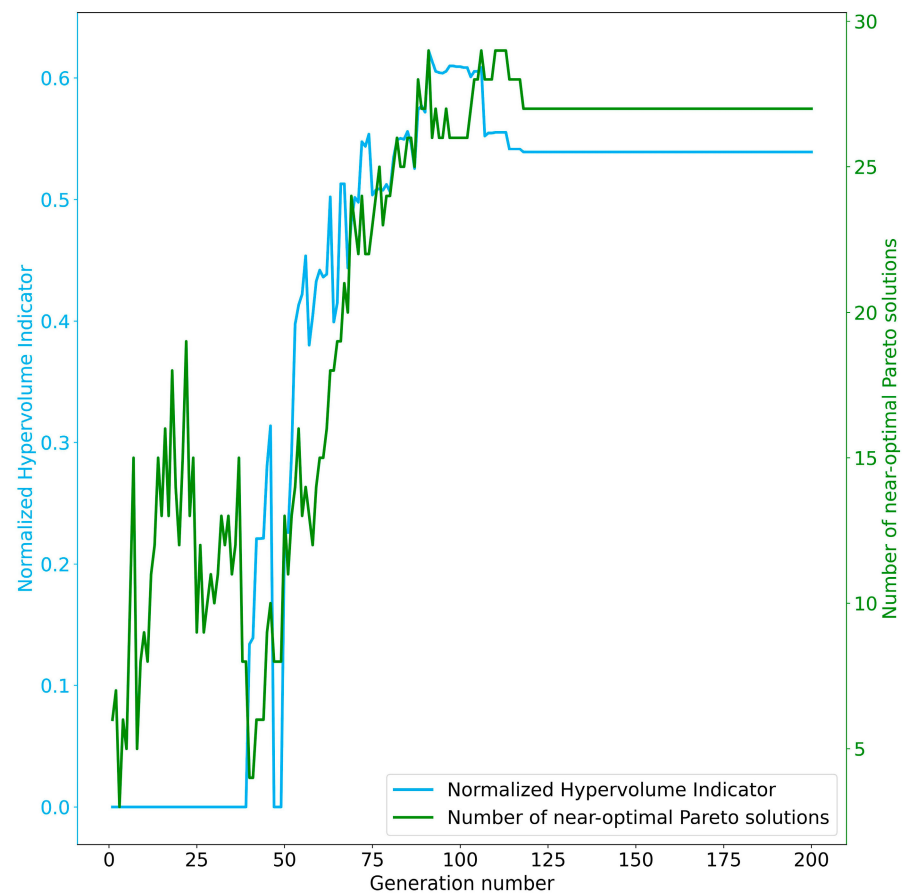


Figure 11. Normalized hypervolume indicator and the number of Pareto solutions as a function of the number of generations in the multi-variable calibration.

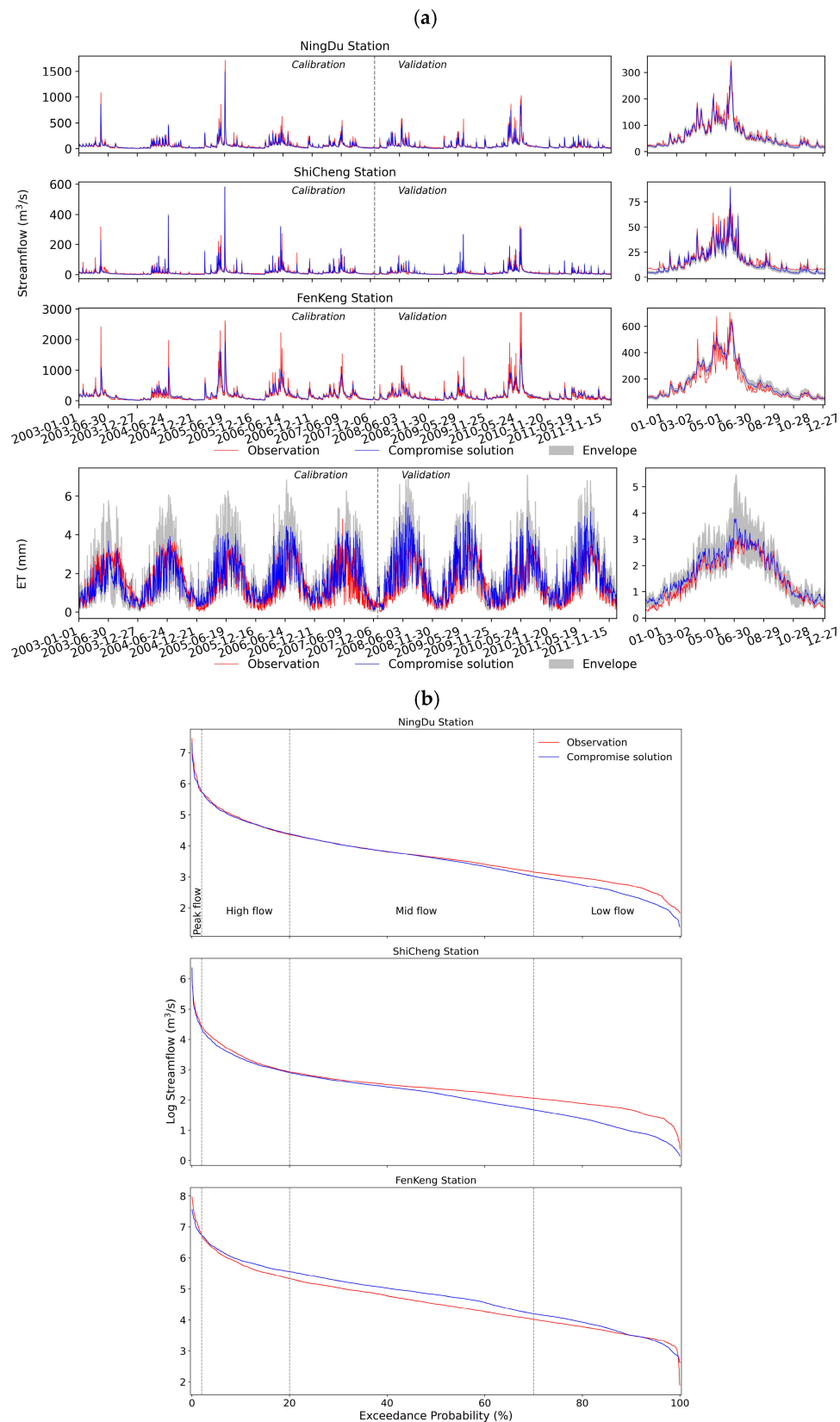


Figure 12. Plotting of calibration variables. **(a)** Time series curves of streamflow at three hydrological stations and ET, panels on the right show the mean seasonal cycle (the envelope of the grey area represents the uncertainty band encompassing all simulations derived from the Pareto optimal front); **(b)** Flow duration curves of three hydrological stations.

3.5.5. Example 3: Multi-Objective Functions Calibration

This section evaluates the multi-objective calibration capability of the Py-SWAT-U-NSGA-III framework using streamflow data from the Fenkeng station, with R^2 and PBIAS as the two objective functions, forming a dual-objective calibration scheme.

Figure 13 illustrates the distribution of Pareto and compromise solutions within a two-dimensional objective space. The figure shows a significant trade-off between the R^2 and PBIAS objectives for streamflow at Fenkeng station. When the R^2 indicator is optimized, the PBIAS indicator is at its worst, and vice versa.

Figure 14 is similar to Figures 6 and 10. The results indicate that, except for GWQMN, ESCO, and EPCO, the parameter ranges for the remaining parameters are significantly narrowed across all Pareto-optimal solutions, suggesting a high level of parameter identifiability.

Figure 15 is similar to Figures 7 and 11. After 98 generations, both indicators stabilize (hypervolume indicator around 0.96 and the number of non-dominated solutions at 31), indicating that the model has converged to a near-optimal Pareto front.

Figure 16 is similar to Figures 8 and 12. The flow duration curve for the Fenkeng station is also presented. Figure 16a shows that streamflow at Fenkeng station is significantly overestimated under low flow conditions, where baseflow dominates. Figure 16b further reveals that streamflow at Fenkeng station is underestimated under peak and high-flow conditions but significantly overestimated under mid and low-flow conditions.

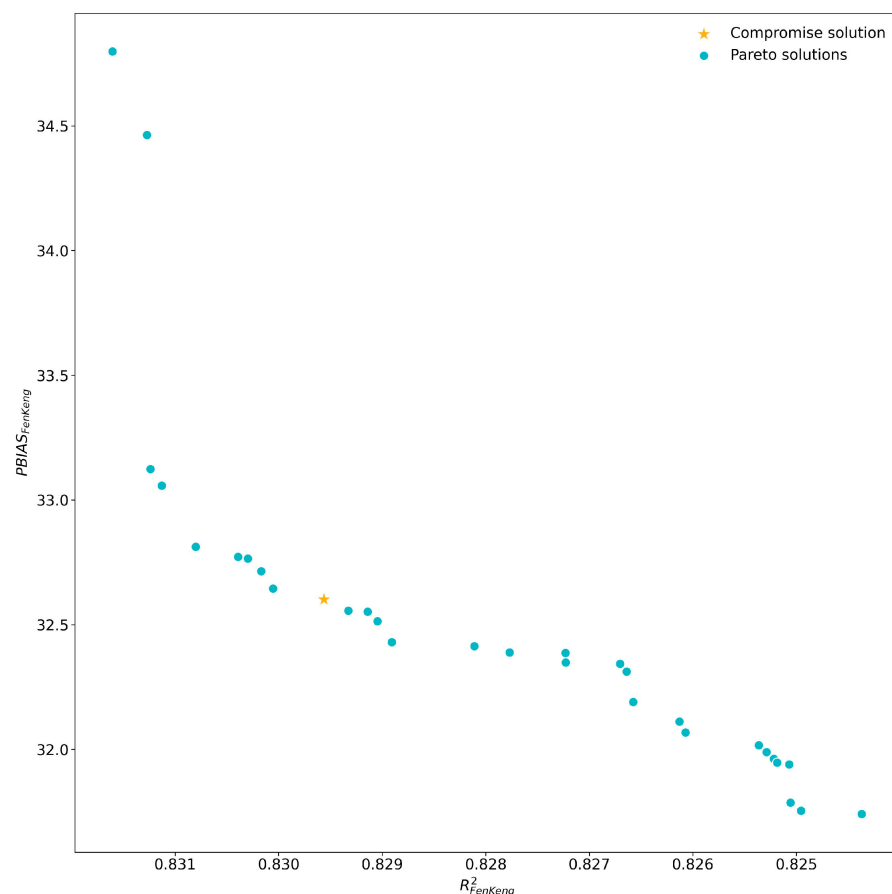


Figure 13. Scatterplot of the distribution of Pareto and compromise solutions in the objective function space.

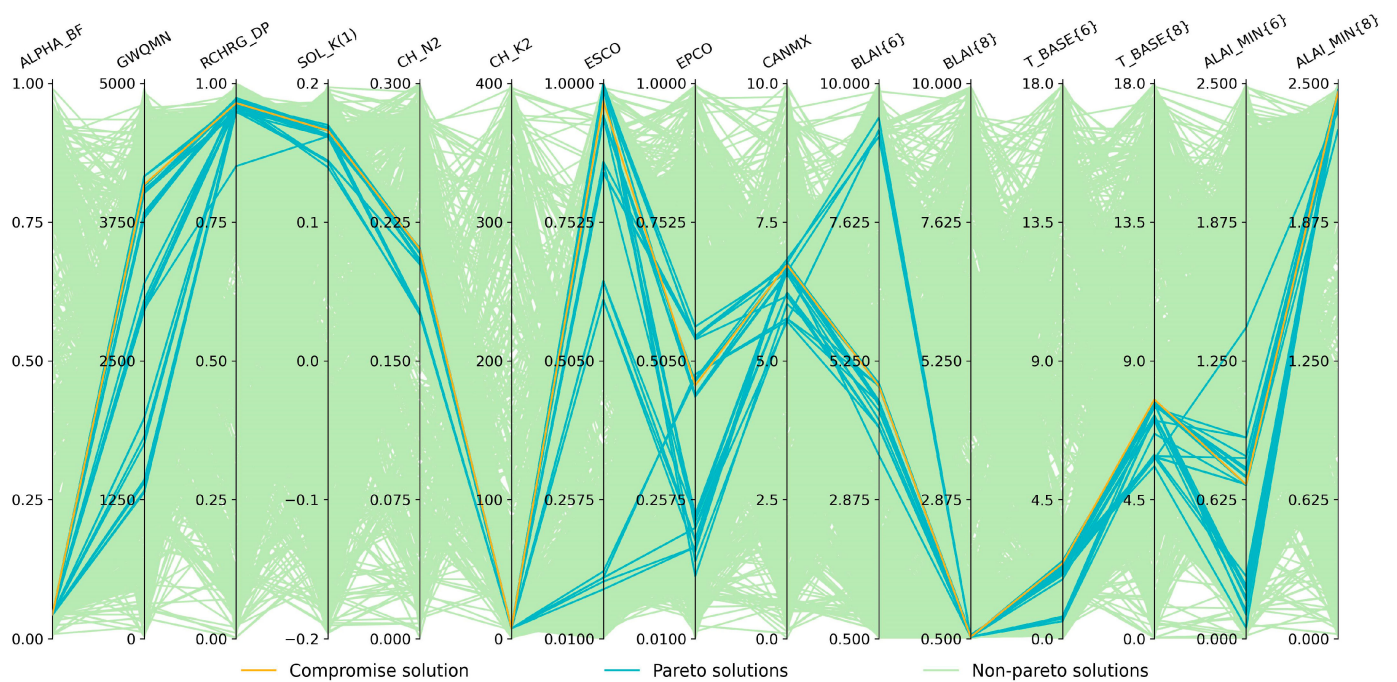


Figure 14. Parallel coordinate plots of Pareto and compromise solutions in parameter space distribution in the multi-objective functions calibration (due to the large number of vegetation-related parameters, only forest-related vegetation parameters are shown).

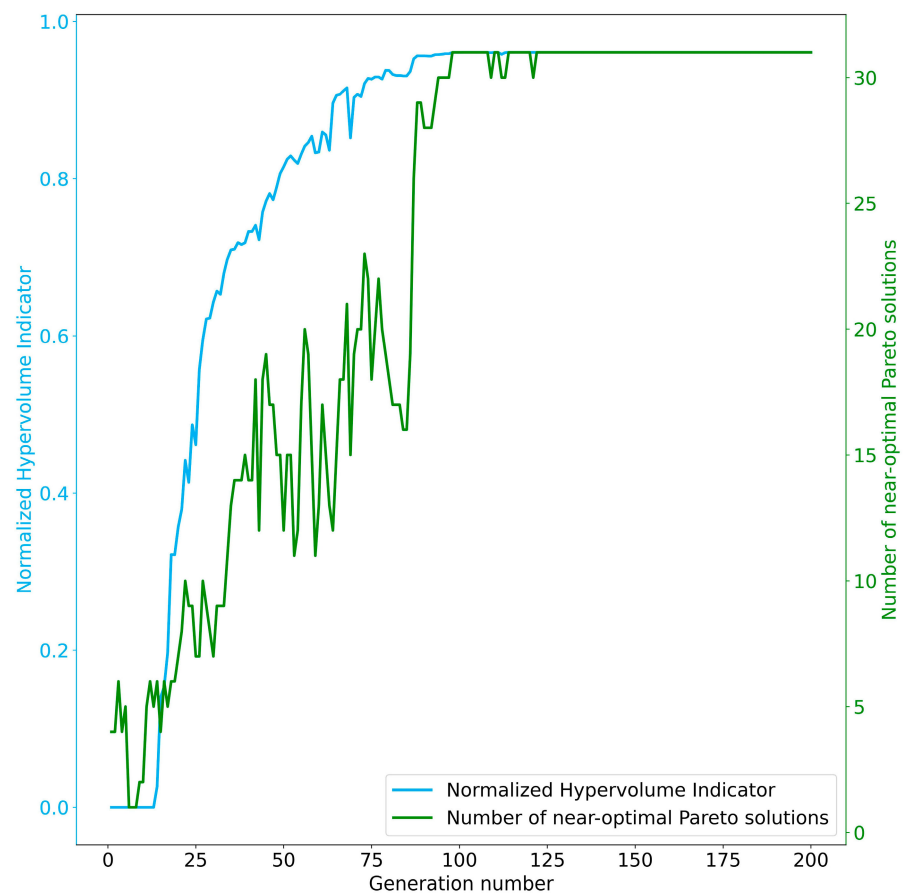


Figure 15. Normalized hypervolume indicator and the number of Pareto solutions as a function of the number of generations in the multi-objective functions calibration.

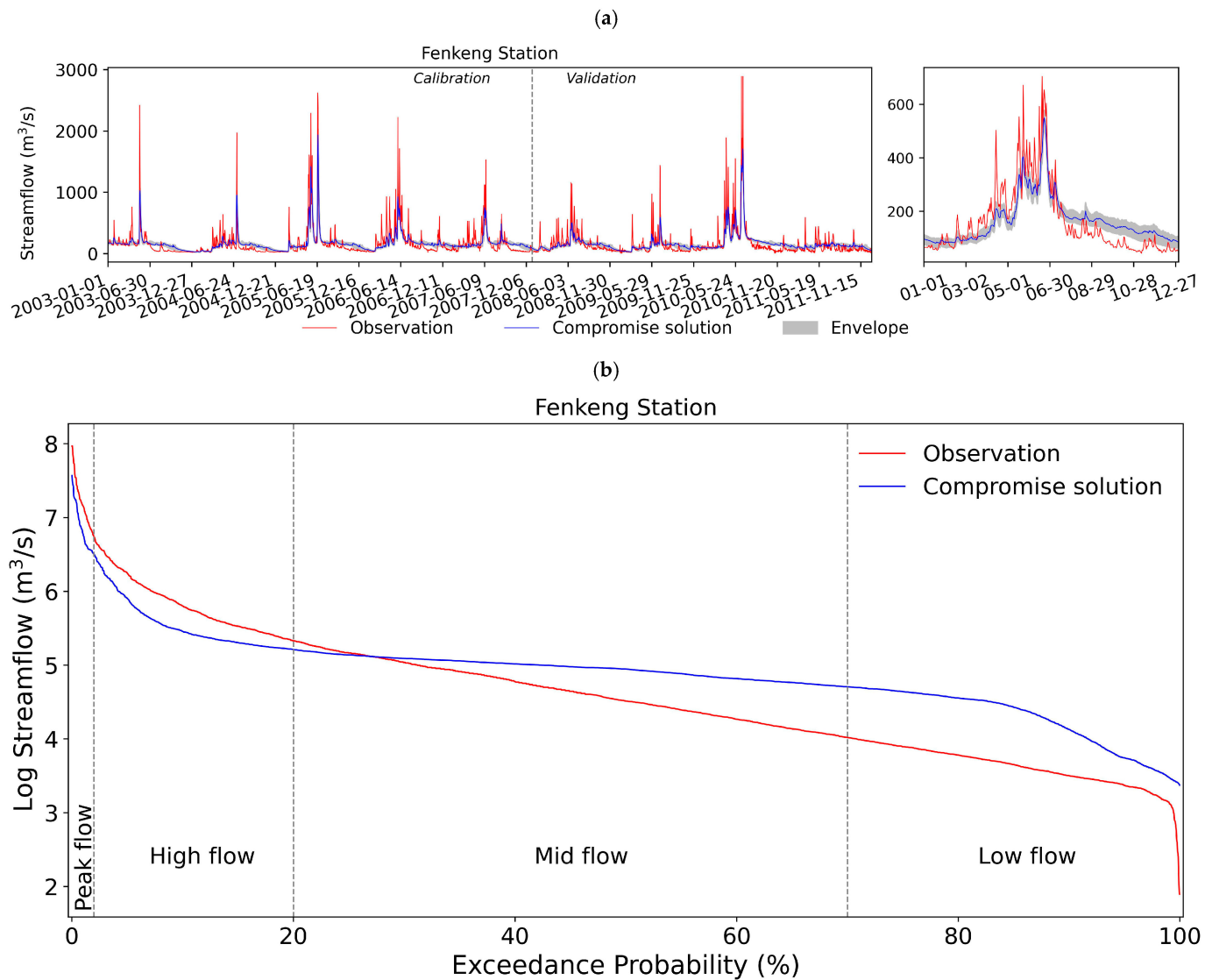


Figure 16. Plotting of calibration variables. (a) Time series curves of streamflow at Fenkeng hydrological station; the panel on the right show the mean seasonal cycle (the envelope of the grey area represents the uncertainty band encompassing all simulations derived from the Pareto optimal front); (b) Flow duration curves of Fenkeng hydrological station.

4. Discussion

4.1. Performance Analysis

Py-SWAT-U-NSGA-III integrates two sensitivity analysis methods: Sobol and PAWN. Sobol, a classic variance-based global sensitivity analysis method, is widely used but computationally demanding. The number of model evaluations increases significantly with the number of parameters, often necessitating tens to hundreds of thousands of model evaluations [90], which imposes a considerable computational burden. In contrast, PAWN, a density-based sensitivity analysis method, requires fewer model evaluations—studies suggest 5000 evaluations are sufficient to rank parameter sensitivities [78]. Given the relatively high model complexity (1589 HRUs) and the large number of parameters in this study, we conducted sensitivity analyses using 10,000 model evaluations. Additionally, we employed time-varying sensitivity analysis (TVSA) to assess temporal changes in parameter sensitivities, providing deeper insights into the interactions between model parameters, hydrological variables, and physical processes. Previous studies on SWAT model parameter sensitivity have primarily focused on overall parameter sensitivity rankings [76,91], poten-

tially leading to information loss and limiting the understanding of the model's internal hydrological processes [92]. A thorough understanding of parameter sensitivity forms a robust foundation for subsequent model calibration, aiming to enhance the predictive accuracy and robustness of SWAT simulations.

The highly parameterized nature of the SWAT model necessitates extensive input/output (I/O) operations due to the large number of parameter files involved [76]. Population-based genetic algorithms, such as U-NSGA-III, naturally lend themselves to parallel processing [40], enhancing computational efficiency. The framework's parallel efficiency is influenced by model complexity, available computing resources, and disk I/O performance. Preliminary tests revealed that model execution on a hard disk drive (HDD) was time-consuming, whereas a solid-state drive (SSD) significantly improved efficiency. In our study environment, the highest parallel efficiency was achieved with 25 cores, resulting in a speedup of 8.95. After 25 cores, increasing the number of cores did not necessarily improve efficiency. The basic model in this study comprised 1589 HRUs and over 10,000 parameter files, making parallel processing particularly challenging and placing a significant strain on disk I/O. Given that the optimal parallel efficiency was achieved with fewer cores than the available cores, it can be inferred that the bottleneck was due to factors other than CPU resources, primarily model complexity and disk performance, particularly I/O saturation caused by extensive parameter file modifications. This phenomenon has been reported in previous studies as well [42,46,93]. Therefore, in practical applications, it is advisable to adjust model complexity appropriately (e.g., the number of HRUs and simulation period) based on specific needs, prioritize SSDs, and select an optimal number of CPU cores to maximize parallel processing efficiency.

To evaluate the multi-objective optimization performance of Py-SWAT-U-NSGA-III, three typical application cases were presented: multi-site, multi-variable, and multi-objective functions calibration, and a visual analysis of the results. In multi-site calibration, the NSE values for streamflow at three hydrological stations ranged from 0.55 to 0.68, which, according to Moriasi, et al. [94], falls within the satisfactory to good range. In multi-variable calibration, the NSE values for streamflow at three hydrological stations ranged from 0.51 to 0.77, while the NSE for evapotranspiration was 0.29. According to Moriasi, et al. [94], the streamflow results range from satisfactory to very good, while the ET results are unsatisfactory. In multi-objective functions calibration, the R^2 and PBIAS for streamflow at the Fenkeng station were 0.83 and 32.60, respectively, which are considered unsatisfactory based on the criteria of Moriasi, et al. [94]. It is important to note that Moriasi's performance ratings apply to monthly scale data, whereas the results in this study are based on daily scale simulations. The higher variability and fluctuations in daily scale data compared to monthly data can potentially degrade objective function performance, particularly for ET data (see Figure 12a for ET time series). From this perspective, the results of this study are acceptable. Comparative analysis of the multi-site and multi-variable calibration revealed that the multi-variable calibration streamflow accuracy was better than the multi-site streamflow accuracy, except for the Shicheng station. This may be due to the inclusion of ET data in the multi-variable calibration, which improves the overall accuracy of the hydrologic sub-process of the SWAT model. Regarding the comparison of the convergence results of the three case studies, the case of the multi-objective functions with two objectives converged after 98 generations with a hypervolume indicator of 0.96. The multi-site case with three objectives converged after 113 generations with a hypervolume indicator of 0.91, and the multi-variable case with four objectives converged after 118 generations with a hypervolume indicator of 0.54. These findings indicate that as the number of objective functions increases, the U-NSGA-III algorithm faces greater challenges in finding approximate Pareto solutions, resulting in more difficulties with convergence, which is similar to the results of Reed, et al. [41]. Previous research has indicated that hydrological model calibration problems can cause MOEAs to fail in obtaining high-quality Pareto approximations within limited model evaluations [41,95]. In this study, the population size of the U-NSGA-III algorithm was set to 100, with a maximum of 200 generations.

Reed, et al. [41] emphasized that population size has a significant impact on MOEAs. Considering the complexity of many-objective hydrological model calibration, increasing the population size and maximum generations may lead to better Pareto approximations, thereby further improving the calibration accuracy of the SWAT model. Secondly, if there is a difference in the number of different types of objective functions (e.g., streamflow from three hydrological stations and 1 watershed-scale ET in this paper), it is also necessary to consider adjusting the weights between various kinds of objective functions to prevent the optimization algorithms from skewing the search results toward one type of objective function (see Section 3.3).

4.2. Advantages, Limitations, and Future Directions

Previous studies have developed various SWAT-based multi-objective calibration tools, some of which are not publicly available [44,47–50], while others have certain limitations. For example, Zhang, et al. [42] developed PP-SWAT using Python 2.6; it may not be compatible with modern Python 3 or directly accessible. Ercan and Goodall [14] designed the NSGA-II algorithm library for SWAT but it did not implement parallel processing for the SWAT model, and NSGA-II may have limitations when dealing with many-objective calibration. Zamani, et al. [46] introduced an LCC-SWAT system that incorporated SUFI-2 and DDS optimization techniques, but SUFI-2 is a single-objective algorithm and DDS is also primarily designed for single-objective optimization [25], which makes it difficult to apply directly to multi-objective optimization. Hernandez-Suarez, et al. [13] presented the U-NSGA-III algorithm tool for SWAT, but it was limited to multi-objective calibration of ecologically relevant hydrologic indices. The classical SCE-UA algorithm proposed by Duan, et al. [22] was initially designed to address single-objective optimization problems.

The Py-SWAT-U-NSGA-III framework offers several advantages, including flexibility, openness, and efficiency. Its integration of advanced multi-objective optimization algorithms, such as U-NSGA-III, allows it to handle a wide range of calibration scenarios, from single-objective to many-objective problems. The framework's support for multi-site, multi-variable, and multi-objective calibration, combined with its ability to integrate remotely sensed data, makes it a versatile tool for hydrological modeling.

Despite its advantages, Py-SWAT-U-NSGA-III has some limitations. For instance, the current framework includes a limited selection of objective function types (see Section 2.3.2). Second, previous studies have utilized hydrological observations such as snow cover area (SCA), snow water equivalent (SWE), or glacier mass balance (GMB) to improve the realism of streamflow simulation in data-scarce glacier watersheds [96–98]. However, the present framework does not consider the integration of such data within the SWAT model. Additionally, while there are numerous MOEAs, different algorithms may yield varying results depending on the specific problem and their respective applicability [41]. This study did not compare with other popular multi-objective optimization algorithms. Finally, the SWAT model typically involves large-scale parameter files, which can create performance bottlenecks and hinder the parallel advantage of multi-core CPUs.

Future research can address these limitations by exploring the following avenues. First, by incorporating additional widely-used objective functions based on specific application needs, such as the RMSE-observations standard deviation ratio (RSR) recommended by Moriasi, et al. [94], as well as commonly used streamflow characteristics (SFCs) indicators like transformed root-mean-square error (TRMSE), runoff coefficient percent error (ROCE), and metrics related to the flow duration curve, such as the slope of the flow duration curve (SFDCE) [13,99]. Second, future studies may consider incorporating additional hydrologic observations, such as snow cover area, snow water equivalent, or glacier mass balance, into the framework, potentially broadening its applicability in data-scarce watersheds. Furthermore, testing other MOEAs, such as Borg and AMALGAM [33,41–43,45,100], with a preference for optimization algorithms already available in the Pymoo library (<https://pymoo.org/index.html>, accessed on 21 October 2024), would require minimal code effort and avoid complex programming. Lastly, considering the computationally intensive nature

of the SWAT model, it is recommended to use SSDs for better read/write performance or to consolidate input data into a structured file (e.g., NetCDF) for block reading, which can alleviate I/O constraints, as suggested by [42]. Moreover, storing SWAT parameter files in databases can improve access speed and parallel efficiency [101]. These proposals collectively expand the potential and significance of Py-SWAT-U-NSGA-III applications within the hydrological community.

Last but not least, although multi-objective optimization has been widely used in hydrological models, researchers should not blindly believe in the effectiveness of multi-objective algorithms when using them in practical applications or need to compare and analyze them from multiple perspectives. Previous studies have also pointed out that multi-objective calibration does not necessarily lead to better model accuracy than single-objective calibration [102]. Therefore, it is necessary to choose the appropriate method according to the specific problem, which also needs to be further compared, analyzed, and verified in future research.

4.3. Prospects for Data-Driven Machine Learning Models (DMLs) and Process-Driven Hydrological Models (PHMs) in Hydrology

Over the past few years, machine learning, including deep learning, has rapidly gained traction in hydrology and water resources management [103–108]. Its unique advantages have been leveraged in various applications, such as drought and flood forecasting, water quality assessment, soil moisture prediction, rainfall-runoff modeling, and predictions of surface and groundwater levels, with streamflow prediction being the most extensively developed application [107]. The advantages and disadvantages of DMLs versus PHMs are summarized below.

Advantages of Data-Driven Machine Learning Models (DMLs) include the following [108]:

- (1) DMLs can handle various input variables flexibly, including non-traditional hydrological measurements;
- (2) DMLs can capture complex nonlinear relationships between input and output data;
- (3) In many hydrological applications, DMLs typically outperform physics-based and conceptual models in terms of predictive accuracy.

However, there are limitations associated with DMLs, as follows [104,105,107]:

- (1) Despite significant successes across numerous fields, including hydrology, DMLs often face criticism for their lack of interpretability and physical consistency, referred to as “black box” models;
- (2) The performance of DMLs is heavily dependent on the quality and quantity of training data;
- (3) DMLs do not fundamentally understand the underlying physical mechanisms, limiting their ability to explain causal relationships between hydrological variables within processes. Additionally, they lack mass balance checks, making it impossible to ascertain how much water is converted from precipitation and how much remains within the watershed;
- (4) DMLs may struggle to capture the effects of non-stationarity, as they assume that the training and testing data distributions are the same;
- (5) Deep learning models typically require substantial computational resources, including high-performance computing (HPC) systems and graphics processing units (GPUs) for training and inference, with computational demands increasing as hydrological datasets become larger and more complex.

Process-driven hydrological models (PHMs) also possess the following strengths:

- (1) PHMs are based on the physical laws governing hydrological processes, thus providing a solid theoretical foundation and interpretability; the parameters in PHMs have clear physical meanings, facilitating calibration and adjustment based on actual hydrological processes;

- (2) PHMs generally exhibit robust predictions across different conditions and regions once parameters have been appropriately calibrated;
- (3) PHMs can be applied to various watersheds and climatic conditions, offering good generalizability.

Nevertheless, PHMs have notable limitations, as follows [107]:

- (1) PHMs often face constraints related to parameterization adequacy, data quality and uncertainty, computational limitations, complexity, and availability;
- (2) PHMs may struggle when simulating complex nonlinear hydrological processes, potentially failing to adequately capture spatial dependencies across different temporal and spatial scales;
- (3) PHMs might introduce higher uncertainties due to differences in datasets, model parameters, and model structures.

Comparative studies have indicated that DMLs can, in certain cases, compete with PHMs, and in some aspects, even outperform them [105,106]. DMLs exhibit superior performance under high flow conditions, making them a valuable tool for real-time flood forecasting to mitigate risks. In contrast, PHMs perform better under low flow conditions and serve as more reliable tools for water resource planning and groundwater management [105,106]. This indicates that both PHMs and DMLs have their respective strengths, and their application should be tailored to specific conditions, such as hydrological, climatic, and topographic factors, as well as data availability and modeling objectives [105].

Future research should explore the combined use of PHMs and DMLs [105,106]. Physics-informed machine learning (PIML) or physics-guided deep learning (PGDL) offer potential avenues for integrating physical mechanisms with data-driven methods, thus enabling complementary advantages, and generating physically consistent and more universally applicable models, thereby enhancing the accuracy of hydrological predictions [104,107].

5. Conclusions

This study introduces Py-SWAT-U-NSGA-III, a multi-objective optimization framework based on the SWAT model, which integrates the U-NSGA-III algorithm to efficiently tackle hydrological model calibration problems across single, multiple, and many-objective scenarios. Developed in the Python environment, the framework is both open and scalable.

The efficacy of Py-SWAT-U-NSGA-III was demonstrated within the Meijiang River Basin in southern China. Sensitivity analysis was conducted using daily streamflow and ET data, highlighting the temporal dynamics of parameter sensitivity with TVSA. Multi-core parallel processing achieved a speedup of up to 8.95, substantially improving the optimization efficiency of the SWAT model. The three case studies involving multi-site, multi-variable, and multi-objective functions calibration confirmed the framework's versatility and effectiveness across different calibration scenarios. Furthermore, the post-processing module facilitated the interpretation of the optimized results, such as the distribution of Pareto and compromise solutions in parameter and objective function spaces, time series curves of calibration variables, and convergence analysis using the hypervolume indicator, thereby deepening insights into model behavior and optimization results.

Overall, Py-SWAT-U-NSGA-III is an open, efficient, and flexible multi-objective optimization framework for the SWAT model that can meet the demands of increasingly complex hydrological systems. Furthermore, we addressed the limitations and potential of Py-SWAT-U-NSGA-III and proposed several directions for improvement. These include the incorporation of streamflow characteristics (SFCs) or ERHIs, the integration of various hydrological observations (SCA, SWE, or GMB), the expansion of multi-objective algorithms, and ongoing efforts to enhance the framework's efficiency to strengthen its potential for continuous application in hydrological modeling.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/w16213030/s1>: S1. Hypervolume Indicator; Figure S1: The

hypervolume indicator in the two-objective case; Table S1: Land use types and their codes in the vegetation database; Table S2: Nash–Sutcliffe efficiency (NSE) for the flow duration curves (FDC) in three case studies; Table S3. Simulation accuracy (NSE) for the three case studies. Refs. [72,109–112] are cited in the Supplementary Materials.

Author Contributions: Conceptualization, H.M. and X.S.; methodology, H.M.; software, H.M.; validation, H.M.; formal analysis, H.M.; investigation, H.M.; data curation, H.M., C.W. and Y.H.; writing—original draft preparation, H.M.; writing—review and editing, H.M., X.S., R.L. and Z.D.; visualization, H.M. and R.M.; supervision, X.S.; project administration, X.S. and R.L.; funding acquisition, X.S., R.M. and Z.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (Grant No. 2020YFC1807103), the National Natural Science Foundation of China (Grant No. 42341206, Grant No. 40771167, and Grant No. 42405145), the Network Security and Informatization Special Application Demonstration Project of Chinese Academy of Sciences (Grant No. CAS-WX2023SF-0403-01), the Fundamental Research Funds for the Central Universities (Grant No. E0E48914X2), and the Natural Science Foundation of Sichuan Province (Grant No. 2024NSFSC0770). In addition, Z.D. is grateful to have received funding from the Crafoord Foundation (Grant No. 20200595 and Grant No. 20210552).

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Acknowledgments: We express our sincere thanks to the scientists of the PML-V2(China) data used in this study, as well as to the developers of the Pymoo open-source library. We would like to acknowledge the data support from “Soil SubCenter, National Earth System Science Data Center, National Science & Technology Infrastructure of China. (<http://soil.geodata.cn>, accessed on 21 October 2024)” and the data support from “National Earth System Science Data Center, National Science & Technology Infrastructure of China. (<http://www.geodata.cn>, accessed on 21 October 2024)”. The authors acknowledge four anonymous reviewers for their very constructive comments and suggestions that substantially improved the quality of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wagener, T.; Boyle, D.P.; Lees, M.J.; Wheeler, H.S.; Gupta, H.V.; Sorooshian, S. A framework for development and application of hydrological models. *Hydrol. Earth Syst. Sci.* **2001**, *5*, 13–26. [CrossRef]
2. Clark, M.P.; Fan, Y.; Lawrence, D.M.; Adam, J.C.; Bolster, D.; Gochis, D.J.; Hooper, R.P.; Kumar, M.; Leung, L.R.; Mackay, D.S.; et al. Improving the representation of hydrologic processes in Earth System Models. *Water Resour. Res.* **2015**, *51*, 5929–5956. [CrossRef]
3. Moges, E.; Demissie, Y.; Larsen, L.; Yassin, F. Review: Sources of Hydrological Model Uncertainties and Advances in Their Analysis. *Water* **2021**, *13*, 28. [CrossRef]
4. Gassman, P.W.; Reyes, M.R.; Green, C.H.; Arnold, J.G. The soil and water assessment tool: Historical development, applications, and future research directions. *Trans. ASABE* **2007**, *50*, 1211–1250. [CrossRef]
5. Keller, A.A.; Garner, K.; Rao, N.; Knipping, E.; Thomas, J. Hydrological models for climate-based assessments at the watershed scale: A critical review of existing hydrologic and water quality models. *Sci. Total Environ.* **2023**, *867*, 161209. [CrossRef]
6. Beven, K. How far can we go in distributed hydrological modelling? *Hydrol. Earth Syst. Sci.* **2001**, *5*, 1–12. [CrossRef]
7. Fatichi, S.; Vivoni, E.R.; Ogden, F.L.; Ivanov, V.Y.; Mirus, B.; Gochis, D.; Downer, C.W.; Camporese, M.; Davison, J.H.; Ebel, B.; et al. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *J. Hydrol.* **2016**, *537*, 45–60. [CrossRef]
8. Beven, K.J. *Rainfall-Runoff Modelling: The Primer*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012.
9. Gao, H.; Birkel, C.; Hrachowitz, M.; Tetzlaff, D.; Soulsby, C.; Savenije, H.H.G. A simple topography-driven and calibration-free runoff generation module. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 787–809. [CrossRef]
10. Babović, V.; Wu, Z.; Larsen, L.C. Calibrating hydrodynamic models by means of simulated evolution. In Proceedings of the First International Conference on Hydroinformatics (Hydroinformatics '94), Rotterdam, The Netherlands, 19–23 September 1994; pp. 193–200.
11. Savic, D.A.; Walters, G.A. Genetic Algorithms for Least-Cost Design of Water Distribution Networks. *J. Water Resour. Plan. Manag.* **1997**, *123*, 67–77. [CrossRef]
12. Khu, S.T.; Madsen, H. Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration. *Water Resour. Res.* **2005**, *41*, W03004. [CrossRef]

13. Hernandez-Suarez, J.S.; Nejadhashemi, A.P.; Deb, K. A novel multi-objective model calibration method for ecohydrological applications. *Environ. Model. Softw.* **2021**, *144*, 105161. [\[CrossRef\]](#)
14. Ercan, M.B.; Goodall, J.L. Design and implementation of a general software library for using NSGA-II with SWAT for multi-objective model calibration. *Environ. Model. Softw.* **2016**, *84*, 112–120. [\[CrossRef\]](#)
15. Beven, K.; Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **2001**, *249*, 11–29. [\[CrossRef\]](#)
16. Arnold, J.G.; Srinivasan, R.; Muttiah, R.S.; Williams, J.R. Large area hydrologic modeling and assessment. Part I: Model development. *JAWRA J. Am. Water Resour. Assoc.* **1998**, *34*, 73–89. [\[CrossRef\]](#)
17. Douglas-Mankin, K.R.; Srinivasan, R.; Arnold, J.G. Soil and Water Assessment Tool (SWAT) Model: Current Developments and Applications. *Trans. ASABE* **2010**, *53*, 1423–1431. [\[CrossRef\]](#)
18. Abbaspour, K.C.; Vajdani, M.; Haghighat, S.; Yang, J. SWAT-CUP calibration and uncertainty programs for SWAT. In Proceedings of the MODSIM 2007 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand, Christchurch, New Zealand, 5 December 2007; pp. 1596–1602.
19. Krishnan, N.; Raj, C.; Chaubey, I.; Sudheer, K.P. Parameter estimation of SWAT and quantification of consequent confidence bands of model simulations. *Environ. Earth Sci.* **2018**, *77*, 470. [\[CrossRef\]](#)
20. Ercan, M.B.; Goodall, J.L.; Castronova, A.M.; Humphrey, M.; Beekwilder, N. Calibration of SWAT models using the cloud. *Environ. Model. Softw.* **2014**, *62*, 188–196. [\[CrossRef\]](#)
21. Zhang, X.; Srinivasan, R.; Zhao, K.; Liew, M.V. Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrol. Process.* **2009**, *23*, 430–441. [\[CrossRef\]](#)
22. Duan, Q.; Sorooshian, S.; Gupta, V. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **1992**, *28*, 1015–1031. [\[CrossRef\]](#)
23. Kavetski, D.; Qin, Y.; Kuczera, G. The Fast and the Robust: Trade-Offs Between Optimization Robustness and Cost in the Calibration of Environmental Models. *Water Resour. Res.* **2018**, *54*, 9432–9455. [\[CrossRef\]](#)
24. Vrugt, J.A.; Gupta, H.V.; Bouten, W.; Sorooshian, S. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.* **2003**, *39*, 1201. [\[CrossRef\]](#)
25. Tolson, B.A.; Shoemaker, C.A. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.* **2007**, *43*, W01413. [\[CrossRef\]](#)
26. Efstratiadis, A.; Koutsoyiannis, D. One decade of multi-objective calibration approaches in hydrological modelling: A review. *Hydrol. Sci. J.* **2010**, *55*, 58–78. [\[CrossRef\]](#)
27. Gupta, H.V.; Sorooshian, S.; Yapo, P.O. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.* **1998**, *34*, 751–763. [\[CrossRef\]](#)
28. Yapo, P.O.; Gupta, H.V.; Sorooshian, S. Multi-objective global optimization for hydrologic models. *J. Hydrol.* **1998**, *204*, 83–97. [\[CrossRef\]](#)
29. Bekele, E.G.; Nicklow, J.W. Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.* **2007**, *341*, 165–176. [\[CrossRef\]](#)
30. Zhang, X.; Srinivasan, R.; Van Liew, M. Multi-Site Calibration of the SWAT Model for Hydrologic Modeling. *Trans. ASABE* **2008**, *51*, 2039–2049. [\[CrossRef\]](#)
31. Rajib, M.A.; Merwade, V.; Yu, Z.Q. Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture. *J. Hydrol.* **2016**, *536*, 192–207. [\[CrossRef\]](#)
32. Herman, M.R.; Nejadhashemi, A.P.; Abouali, M.; Hernandez-Suarez, J.S.; Daneshvar, F.; Zhang, Z.; Anderson, M.C.; Sadeghi, A.M.; Hain, C.R.; Sharifi, A. Evaluating the role of evapotranspiration remote sensing data in improving hydrological modeling predictability. *J. Hydrol.* **2018**, *556*, 39–49. [\[CrossRef\]](#)
33. Chilkoti, V.; Bolisetti, T.; Balachandar, R. Multi-objective autocalibration of SWAT model for improved low flow performance for a small snowfed catchment. *Hydrol. Sci. J.-J. Des Sci. Hydrol.* **2018**, *63*, 1482–1501. [\[CrossRef\]](#)
34. Herman, M.R.; Hernandez-Suarez, J.S.; Nejadhashemi, A.P.; Kropp, I.; Sadeghi, A.M. Evaluation of Multi- and Many-Objective Optimization Techniques to Improve the Performance of a Hydrologic Model Using Evapotranspiration Remote-Sensing Data. *J. Hydrol. Eng.* **2020**, *25*, 04020006. [\[CrossRef\]](#)
35. Mohammadi Igder, O.; Alizadeh, H.; Mojaradi, B.; Bayat, M. Multivariate assimilation of satellite-based leaf area index and ground-based river streamflow for hydrological modelling of irrigated watersheds using SWAT+. *J. Hydrol.* **2022**, *610*, 128012. [\[CrossRef\]](#)
36. Sheffield, J.; Wood, E.F.; Pan, M.; Beck, H.; Coccia, G.; Serrat-Capdevila, A.; Verbist, K. Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions. *Water Resour. Res.* **2018**, *54*, 9724–9758. [\[CrossRef\]](#)
37. Jiang, D.; Wang, K. The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review. *Water* **2019**, *11*, 1615. [\[CrossRef\]](#)
38. Dembélé, M.; Ceperley, N.; Zwart, S.J.; Salvatore, E.; Mariethoz, G.; Schaeffli, B. Potential of satellite and reanalysis evaporation datasets for hydrological modelling under various model calibration strategies. *Adv. Water Resour.* **2020**, *143*, 103667. [\[CrossRef\]](#)
39. Rajib, A.; Evenson, G.R.; Golden, H.E.; Lane, C.R. Hydrologic model predictability improves with spatially explicit calibration using remotely sensed evapotranspiration and biophysical parameters. *J. Hydrol.* **2018**, *567*, 668–683. [\[CrossRef\]](#)

40. Deb, K. Multi-objective Optimisation Using Evolutionary Algorithms: An Introduction. In *Multi-Objective Evolutionary Optimisation for Product Design and Manufacturing*; Deb, K., Ed.; Springer: London, UK, 2011; pp. 3–34.
41. Reed, P.M.; Hadka, D.; Herman, J.D.; Kasprzyk, J.R.; Kollat, J.B. Evolutionary multiobjective optimization in water resources: The past, present, and future. *Adv. Water Resour.* **2013**, *51*, 438–456. [\[CrossRef\]](#)
42. Zhang, X.; Beeson, P.; Link, R.; Manowitz, D.; Izaurralde, R.C.; Sadeghi, A.; Thomson, A.M.; Sahajpal, R.; Srinivasan, R.; Arnold, J.G. Efficient multi-objective calibration of a computationally intensive hydrologic model with parallel computing software in Python. *Environ. Model. Softw.* **2013**, *46*, 208–218. [\[CrossRef\]](#)
43. Vrugt, J.A.; Robinson, B.A. Improved evolutionary optimization from genetically adaptive multimethod search. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 708–711. [\[CrossRef\]](#)
44. Besalatpour, A.A.; Pourreza-Bilondi, M.; Aghakhani Afshar, A. Parallelization of AMALGAM algorithm for a multi-objective optimization of a hydrological model. *Appl. Water Sci.* **2023**, *13*, 241. [\[CrossRef\]](#)
45. Zhang, X.; Srinivasan, R.; Liew, M.V. On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site calibration of the SWAT model. *Hydrol. Process.* **2010**, *24*, 955–969. [\[CrossRef\]](#)
46. Zamani, M.; Shrestha, N.K.; Akhtar, T.; Boston, T.; Daggupati, P. Advancing model calibration and uncertainty analysis of SWAT models using cloud computing infrastructure: LCC-SWAT. *J. Hydroinformatics* **2020**, *23*, 1–15. [\[CrossRef\]](#)
47. Confesor, R.B.; Whittaker, G.W. Automatic calibration of hydrologic models with multi-objective evolutionary algorithm and Pareto optimization. *J. Am. Water Resour. Assoc.* **2007**, *43*, 981–989. [\[CrossRef\]](#)
48. Her, Y.; Cibilin, R.; Chaubey, I. Application of Parallel Computing Methods for Improving Efficiency of Optimization in Hydrologic and Water Quality Modeling. *Appl. Eng. Agric.* **2015**, *31*, 455–468. [\[CrossRef\]](#)
49. Kayastha, N.; Lu, S.; Betrie, G.; Zakayo, Z.; van Griensven, A.; Solomatine, D. Dynamic linking of the watershed model SWAT to the multi-objective optimization tool NSGAX. In Proceedings of the Watermatex, 8th IWA Symposium on Systems Analysis and Integrated Assessment, IWA, San Sebastian, Spain, 20–22 June 2011; pp. 10–15.
50. Zhang, X.; Izaurralde, R.C.; Zong, Z.; Zhao, K.; Thomson, A.M. Evaluating the Efficiency of a Multi-core Aware Multi-objective Optimization Tool for Calibrating the SWAT Model. *Trans. ASABE* **2012**, *55*, 1723–1731. [\[CrossRef\]](#)
51. Rouholahnejad, E.; Abbaspour, K.C.; Vejdani, M.; Srinivasan, R.; Schulin, R.; Lehmann, A. A parallelization framework for calibration of hydrological models. *Environ. Model. Softw.* **2012**, *31*, 28–36. [\[CrossRef\]](#)
52. Abbaspour, K.C. *SWAT-CUP: SWAT Calibration and Uncertainty Programs—A User Manual*; Eawag: Dübendorf, Switzerland, 2015; pp. 16–70.
53. Zhang, X.; Srinivasan, R.; Hao, F. Predicting Hydrologic Response to Climate Change in the Luohe River Basin Using the SWAT Model. *Trans. ASABE* **2007**, *50*, 901–910. [\[CrossRef\]](#)
54. Nijzink, R.C.; Almeida, S.; Pechlivanidis, I.G.; Capell, R.; Gustafssons, D.; Arheimer, B.; Parajka, J.; Freer, J.; Han, D.; Wagener, T.; et al. Constraining Conceptual Hydrological Models With Multiple Information Sources. *Water Resour. Res.* **2018**, *54*, 8332–8362. [\[CrossRef\]](#)
55. Pool, S.; Fowler, K.; Peel, M. Benefit of Multivariate Model Calibration for Different Climatic Regions. *Water Resour. Res.* **2024**, *60*, e2023WR036364. [\[CrossRef\]](#)
56. Mahmood, R.; Jia, S. A Comprehensive Approach to Develop a Hydrological Model for the Simulation of All the Important Hydrological Components: The Case of the Three-River Headwater Region, China. *Water* **2022**, *14*, 2778. [\[CrossRef\]](#)
57. Ali, M.H.; Popescu, I.; Jonoski, A.; Solomatine, D.P. Remote Sensed and/or Global Datasets for Distributed Hydrological Modelling: A Review. *Remote Sens.* **2023**, *15*, 1642. [\[CrossRef\]](#)
58. Deb, K.; Jain, H. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints. *IEEE Trans. Evol. Comput.* **2014**, *18*, 577–601. [\[CrossRef\]](#)
59. Köppen, M.; Yoshida, K. Substitute Distance Assignments in NSGA-II for Handling Many-objective Optimization Problems. In Proceedings of the Evolutionary Multi-Criterion Optimization; Springer: Berlin/Heidelberg, Germany, 2007; pp. 727–741.
60. Seada, H.; Deb, K. A Unified Evolutionary Optimization Procedure for Single, Multiple, and Many Objectives. *IEEE Trans. Evol. Comput.* **2016**, *20*, 358–369. [\[CrossRef\]](#)
61. Neitsch, S.L.; Arnold, J.G.; Kiniry, J.R.; Williams, J.R. *Soil and Water Assessment Tool Theoretical Documentation Version 2009*; Texas Water Resources Institute: College Station, TX, USA, 2011.
62. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [\[CrossRef\]](#)
63. Lu, S.; Kayastha, N.; Thodsen, H.; van Griensven, A.; Andersen, H.E. Multiobjective Calibration for Comparing Channel Sediment Routing Models in the Soil and Water Assessment Tool. *J. Environ. Qual.* **2014**, *43*, 110–120. [\[CrossRef\]](#)
64. Maringanti, C.; Chaubey, I.; Popp, J. Development of a multiobjective optimization tool for the selection and placement of best management practices for nonpoint source pollution control. *Water Resour. Res.* **2009**, *45*, W06406. [\[CrossRef\]](#)
65. Jain, H.; Deb, K. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point Based Nondominated Sorting Approach, Part II: Handling Constraints and Extending to an Adaptive Approach. *IEEE Trans. Evol. Comput.* **2014**, *18*, 602–622. [\[CrossRef\]](#)
66. Hernandez-Suarez, J.S.; Nejadhashemi, A.P. Probabilistic Predictions of Ecologically Relevant Hydrologic Indices Using a Hydrological Model. *Water Resour. Res.* **2022**, *58*, e2021WR031104. [\[CrossRef\]](#)

67. Blank, J.; Deb, K.; Dhebar, Y.; Bandaru, S.; Seada, H. Generating Well-Spaced Points on a Unit Simplex for Evolutionary Many-Objective Optimization. *IEEE Trans. Evol. Comput.* **2021**, *25*, 48–60. [\[CrossRef\]](#)
68. McKay, M.D.; Beckman, R.J.; Conover, W.J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. *Technometrics* **2000**, *42*, 55–61. [\[CrossRef\]](#)
69. Blank, J.; Deb, K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* **2020**, *8*, 89497–89509. [\[CrossRef\]](#)
70. Zeleny, M. Multiple Criteria Decision Making (MCDM): From Paradigm Lost to Paradigm Regained? *J. Multi-Criteria Decis. Anal.* **2011**, *18*, 77–89. [\[CrossRef\]](#)
71. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; John Wiley & Sons: Hoboken, NJ, USA, 2001; Volume 16.
72. Zitzler, E.; Thiele, L. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Trans. Evol. Comput.* **1999**, *3*, 257–271. [\[CrossRef\]](#)
73. Singh, H.K. Understanding Hypervolume Behavior Theoretically for Benchmarking in Evolutionary Multi-/Many-Objective Optimization. *IEEE Trans. Evol. Comput.* **2020**, *24*, 603–610. [\[CrossRef\]](#)
74. Sobol', I.Y.M. On sensitivity estimation for nonlinear mathematical models. *Mat. Model.* **1990**, *2*, 112–118.
75. Pianosi, F.; Wagener, T. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environ. Model. Softw.* **2015**, *67*, 1–11. [\[CrossRef\]](#)
76. Nossent, J.; Elsen, P.; Bauwens, W. Sobol' sensitivity analysis of a complex environmental model. *Environ. Model. Softw.* **2011**, *26*, 1515–1525. [\[CrossRef\]](#)
77. Zhang, C.; Chu, J.G.; Fu, G.T. Sobol's sensitivity analysis for a distributed hydrological model of Yichun River Basin, China. *J. Hydrol.* **2013**, *480*, 58–68. [\[CrossRef\]](#)
78. Pianosi, F.; Wagener, T. Distribution-based sensitivity analysis from a generic input-output sample. *Environ. Model. Softw.* **2018**, *108*, 197–207. [\[CrossRef\]](#)
79. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [\[CrossRef\]](#)
80. Gupta, H.V.; Kling, H.; Yilmaz, K.K.; Martinez, G.F. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.* **2009**, *377*, 80–91. [\[CrossRef\]](#)
81. Gupta, H.; Sorooshian, S.; Yapo, P. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *J. Hydrol. Eng.* **1999**, *4*, 135–143. [\[CrossRef\]](#)
82. Liu, X.; Li, Y.; Zhong, X.; Zhao, C.; Jensen, J.R.; Zhao, Y. Towards increasing availability of the Ångström–Prescott radiation parameters across China: Spatial trend and modeling. *Energy Convers. Manag.* **2014**, *87*, 975–989. [\[CrossRef\]](#)
83. Liu, Y.; Tan, Q.; Pan, T. Determining the Parameters of the Ångström–Prescott Model for Estimating Solar Radiation in Different Regions of China: Calibration and Modeling. *Earth Space Sci.* **2019**, *6*, 1976–1986. [\[CrossRef\]](#)
84. USDA, S. *National Engineering Handbook, Section 4: Hydrology*; Soil Conservation Service, USDA: Washington, DC, USA, 1972.
85. Williams, J.R. Flood routing with variable travel time or variable storage coefficients. *Trans. ASAE* **1969**, *12*, 100–103. [\[CrossRef\]](#)
86. Wu, B.; Yuan, Q.; Yan, C.; Wang, Z.; Yu, X.; Li, A.; Ma, R.; Huang, J.; Chen, J.; Chang, C.; et al. Land cover changes of China from 2000 to 2010. *Quat. Sci.* **2014**, *34*, 723–731. [\[CrossRef\]](#)
87. He, S.Y.; Zhang, Y.Q.; Ma, N.; Tian, J.; Kong, D.D.; Liu, C.M. A daily and 500 m coupled evapotranspiration and gross primary production product across China during 2000–2020. *Earth Syst. Sci. Data* **2022**, *14*, 5463–5488. [\[CrossRef\]](#)
88. Eager, D.L.; Zahorjan, J.; Lazowska, E.D. Speedup versus efficiency in parallel systems. *IEEE Trans. Comput.* **1989**, *38*, 408–423. [\[CrossRef\]](#)
89. Wu, X.; Shirvan, K.; Kozłowski, T. Demonstration of the relationship between sensitivity and identifiability for inverse uncertainty quantification. *J. Comput. Phys.* **2019**, *396*, 12–30. [\[CrossRef\]](#)
90. Saltelli, A. Making best use of model evaluations to compute sensitivity indices. *Comput. Phys. Commun.* **2002**, *145*, 280–297. [\[CrossRef\]](#)
91. Khorashadi Zadeh, F.; Nossent, J.; Sarrazin, F.; Pianosi, F.; van Griensven, A.; Wagener, T.; Bauwens, W. Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model. *Environ. Model. Softw.* **2017**, *91*, 210–222. [\[CrossRef\]](#)
92. Pianosi, F.; Wagener, T. Understanding the time-varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis. *Hydrol. Process.* **2016**, *30*, 3991–4003. [\[CrossRef\]](#)
93. Lin, Q.; Zhang, D.; Wu, J.; Chen, X.; Fang, Y.; Lin, B. PASS4SWAT: Orchestration of containerized SWAT for facilitating computational reproducibility of model calibration and uncertainty analysis. *Environ. Model. Softw.* **2024**, *178*, 106085. [\[CrossRef\]](#)
94. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900. [\[CrossRef\]](#)
95. Tang, Y.; Reed, P.; Wagener, T. How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? *Hydrol. Earth Syst. Sci.* **2006**, *10*, 289–307. [\[CrossRef\]](#)
96. Gao, H.; Dong, J.; Chen, X.; Cai, H.; Liu, Z.; Jin, Z.; Mao, D.; Yang, Z.; Duan, Z. Stepwise modeling and the importance of internal variables validation to test model realism in a data scarce glacier basin. *J. Hydrol.* **2020**, *591*, 125457. [\[CrossRef\]](#)
97. Grusson, Y.; Sun, X.; Gascoin, S.; Sauvage, S.; Raghavan, S.; Anctil, F.; Sánchez-Pérez, J.-M. Assessing the capability of the SWAT model to simulate snow, snow melt and streamflow dynamics over an alpine watershed. *J. Hydrol.* **2015**, *531*, 574–588. [\[CrossRef\]](#)

98. Tuo, Y.; Marcolini, G.; Disse, M.; Chiogna, G. A multi-objective approach to improve SWAT model calibration in alpine catchments. *J. Hydrol.* **2018**, *559*, 347–360. [\[CrossRef\]](#)
99. Kollat, J.B.; Reed, P.M.; Wagener, T. When are multiobjective calibration trade-offs in hydrologic models meaningful? *Water Resour. Res.* **2012**, *48*, W03520. [\[CrossRef\]](#)
100. Hadka, D.; Reed, P. Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework. *Evol. Comput.* **2013**, *21*, 231–259. [\[CrossRef\]](#)
101. Zhang, D.J.; Lin, Q.Y.; Yao, H.X.; He, Y.R.; Deng, J.; Zhang, X.X. Accelerating SWAT Simulations Using An In-Memory NoSQL Database. *J. Environ. Inform.* **2021**, *37*, 142–152. [\[CrossRef\]](#)
102. Zhang, R.; Liu, J.; Gao, H.; Mao, G. Can multi-objective calibration of streamflow guarantee better hydrological model accuracy? *J. Hydroinformatics* **2018**, *20*, 687–698. [\[CrossRef\]](#)
103. Chadalawada, J.; Herath, H.M.V.V.; Babovic, V. Hydrologically Informed Machine Learning for Rainfall-Runoff Modeling: A Genetic Programming-Based Toolkit for Automatic Model Induction. *Water Resour. Res.* **2020**, *56*, e2019WR026933. [\[CrossRef\]](#)
104. Herath, H.M.V.V.; Chadalawada, J.; Babovic, V. Hydrologically informed machine learning for rainfall-runoff modelling: Towards distributed modelling. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 4373–4401. [\[CrossRef\]](#)
105. Kim, T.; Yang, T.; Gao, S.; Zhang, L.; Ding, Z.; Wen, X.; Gourley, J.J.; Hong, Y. Can artificial intelligence and data-driven machine learning models match or even replace process-driven hydrologic models for streamflow simulation?: A case study of four watersheds with different hydro-climatic regions across the CONUS. *J. Hydrol.* **2021**, *598*, 126423. [\[CrossRef\]](#)
106. Jin, A.; Wang, Q.; Zhan, H.; Zhou, R. Comparative Performance Assessment of Physical-Based and Data-Driven Machine-Learning Models for Simulating Streamflow: A Case Study in Three Catchments across the US. *J. Hydrol. Eng.* **2024**, *29*, 05024004. [\[CrossRef\]](#)
107. Tripathy, K.P.; Mishra, A.K. Deep learning in hydrology and water resources disciplines: Concepts, methods, applications, and research directions. *J. Hydrol.* **2024**, *628*, 130458. [\[CrossRef\]](#)
108. Nearing, G.S.; Kratzert, F.; Sampson, A.K.; Pelissier, C.S.; Klotz, D.; Frame, J.M.; Prieto, C.; Gupta, H.V. What Role Does Hydrological Science Play in the Age of Machine Learning? *Water Resour. Res.* **2021**, *57*, e2020WR028091. [\[CrossRef\]](#)
109. Shang, K.; Ishibuchi, H.; He, L.J.; Pang, L.M. A Survey on the Hypervolume Indicator in Evolutionary Multiobjective Optimization. *IEEE Trans. Evol. Comput.* **2021**, *25*, 1–20. [\[CrossRef\]](#)
110. Guerreiro, A.P.; Fonseca, C.M.; Paquete, L. The Hypervolume Indicator: Computational Problems and Algorithms. *ACM Comput. Surv.* **2021**, *54*, 119. [\[CrossRef\]](#)
111. Fonseca, C.M.; Paquete, L.; López-Ibáñez, M. An improved dimension-sweep algorithm for the hypervolume indicator. In Proceedings of the IEEE Congress on Evolutionary Computation, Vancouver, Canada, 16–21 July 2006; pp. 1157–1163.
112. Gharari, S.; Hrachowitz, M.; Fenicia, F.; Gao, H.; Savenije, H.H.G. Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4839–4859. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.