water

Article

# Automatic SWMM Parameter Calibration Method Based on the Differential Evolution and Bayesian Optimization Algorithm

**Jiawei Gao [1], Ji Liang [1,*], Yu Lu [1], Ruilong Zhou [1] and Xin Lu [2]**

1 School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China; gaojiawei@hust.edu.cn (J.G.); luyu5250@hust.edu.cn (Y.L.); chou5118@hust.edu.cn (R.Z.)
2 Sichuan Hydraulic Research Institute, Chengdu 610072, China; scsky_lx@126.com
* Correspondence: larkwater1@163.com

**Abstract:** In response to the low accuracy exhibited by the Storm Water Management Model (SWMM), we propose an enhanced Differential Evolution and Bayesian Optimization Algorithm (DE-BOA). This algorithm integrates the global search capability of the differential evolution algorithm with the local search capability of the Bayesian optimization algorithm, which enables a more comprehensive exploration of the vector solution space. A comparative analysis of various types of rainfall events is conducted. For model calibration and validation, a drainage subzone in Jinshazhou, Guangzhou City, is selected as the research subject. In total, 20 specific rainfall events are selected, and the DE-BOA algorithm outperforms the manual calibration, the differential evolution algorithm, and the Bayesian optimization algorithm regarding model calibration accuracy. Furthermore, the DE-BOA algorithm exhibits robust adaptability to rainfall events characterized by multiple peaks and higher precipitation levels, with the Nash–Sutcliffe efficiency coefficient values surpassing 0.90. This study's findings could hold significant reference value for dynamically updating model parameters, thereby enhancing the model simulation performance and improving the accuracy of the urban intelligent water management platform.

**Keywords:** SWMM; differential evolution algorithm; Bayesian optimization algorithm; automatic parameter calibration; urban drainage



**Citation:** Gao, J.; Liang, J.; Lu, Y.; Zhou, R.; Lu, X. Automatic SWMM Parameter Calibration Method Based on the Differential Evolution and Bayesian Optimization Algorithm. *Water* **2023**, *15*, 3582. https://doi.org/10.3390/w15203582

Academic Editor: Adriana Bruggeman

Received: 2 September 2023
Revised: 6 October 2023
Accepted: 10 October 2023
Published: 13 October 2023



## 1. Introduction

The simulation of dynamic rainfall runoff in urban areas using urban stormwater modeling serves as the cornerstone of research in urban rainwater management. Among these models, SWMM—one of the most extensively employed urban stormwater runoff models [1–5]—simulates and predicts urban rainwater runoff and drainage conditions [6]. As a structurally intricate distributed hydrological model, SWMM contains parameters that lack precise physical definitions in the process of model abstraction and simulation. These parameters, such as the impermeable manning coefficient, the depression storage in impermeable areas, and the minimum infiltration rate, are typically determined based on past experiences and may vary with different environmental factors [7–10]. Consequently, parameter calibration and validation are typically conducted during the urban stormwater runoff simulation process, to ensure the precision and applicability of the simulation results [11].

The calibration of SWMM parameters primarily involves two approaches: manual calibration and automatic calibration. Manual calibration, which is more straightforward, relies upon the expertise of experienced model users or hydraulic specialists to manually adjust the model parameters. Calibration is achieved through iterative simulations and parameter value adjustments [12]. This approach demands a high level of expertise, specialized knowledge, a high time consumption, and intense labor, which makes it unsuitable for large-scale parameter optimization.

Meanwhile, automatic calibration utilizes computer algorithms such as particle swarm optimization, simulated annealing, grid search, and random search to automatically optimize the model parameters. These algorithms can efficiently search through many parameter combinations within a short period, which enables the identification of optimal parameter combinations and consequently enhances the accuracy and robustness of the model. The question of how to maximize the accuracy of parameter calibration using optimization algorithms has garnered widespread attention, and several studies have proposed different improved algorithms to enhance calibration performance. For example, Jin et al. (2011) employed an enhanced adaptive genetic algorithm to optimize SWMM calibration problems. Wan and James (2012) introduced a sensitivity-based runoff model automatic calibration method. Gao et al. (2020) simultaneously considered parameter and input uncertainties and proposed a Bayesian-based automatic calibration framework. Perin et al. (2020) utilized the Parameter Estimation (PEST) software package to automatically calibrate model parameters by minimizing the deviation between model outputs and measured data. Shahed Behrouz et al. (2020) developed a new open-source software called OSTRICH-SWMM for single-objective and multi-objective automatic calibration [13–17]. Exploring precise and efficient methods for automatic parameter calibration improves the accuracy of simulation results and significantly reduces the modeling time to enhance the applicability of the SWMM model.

Among the many intelligent algorithms, many have been proposed as applicable solutions to problems in the field of water resources engineering. Armin Azad et al. (2019) investigated the application of a genetic algorithm (GA), ant colony optimization for the continuous domain (ACOR), particle swarm optimization (PSO), and differential evolution (DE) in order to improve the performance of ANFIS models in simulating monthly rainfall magnitudes and discussed how intelligent algorithms can avoid falling into local optima and improve the performance of ANFIS [18]. Saeed Farzin et al. (2021) proposed to hybridize the bat algorithm (BA) with the PSO algorithm and found that the hybrid algorithm (HBP) has a great potential for the optimization of nullahs design [19]. Mariacrocetta Sambito et al. (2021) applied the Bayesian approach to determine the optimal sensor distribution for solving pollution detection and localization problems in urban drainage systems [20].

Among these algorithms, the DE and Bayesian optimization algorithms (BOA) are widely used because of their powerful parameter searching ability. The DE algorithm was developed by Kenneth Price in an attempt to solve the Chebyshev polynomial problem given to him by Rainer Storn [21], and the BOA algorithm is based on the Bayesian formulation and has been developed over many generations. However, both the DE and the BOA have weaknesses that can be improved by hybrid algorithms.

Yukun Bai et al. (2022) proposed an adaptive mutation differential evolution Markov chain (AM-DEMC) algorithm, combining DE and BOA, for the groundwater contamination source identification (GCSI) problem [22]. Tyler Jon Smith et al. (2008) explored the role of Bayesian methods, specifically Markov chain Monte Carlo (MCMC) techniques combined with the DE algorithm, in the assessment of uncertainty and parameter estimation in hydrologic models [23].

Although some studies have applied intelligent algorithms to the automatic rate determination of SWMM, the rate determination performance has not been improved by considering the combination of the global and local search capabilities of the DE and BOA algorithms; therefore, the present study focuses on the Jinshazhou drainage subdistrict in Guangzhou City as the research area. Utilizing the open-source nature of SWMM, Python code is developed to modify and record the engineering input files (.inp) and result output files (.out) [24]. Regarding the precision of parameter calibration, the DE and BOA algorithms are improved in the following two aspects. Firstly, the global search capability of the DE algorithm is combined with the local search capability of the BOA algorithm to prevent parameters from being trapped in local optima during the calibration process, so that the SWMM model parameters are calibrated with high precision. Secondly, an adaptive weight allocation strategy is used; during the calibration process, the proportions of the DE algorithm and BOA algorithm are dynamically allocated by the equilibrium level between constraint conditions and objective function to enhance the efficiency and precision of model calibration. The advantages and disadvantages of using the hybrid algorithm for parameter rate determination are also summarized by comparing and analyzing the performance of the hybrid algorithm in the SWMM model under different rainfall events. This study aims to explore a new approach to improve the accuracy of parameters in urban stormwater runoff models, which enhances the predictive capability of intelligent urban stormwater runoff models.

## 2. Materials and Methods

### *2.1. Differential Evolution Algorithm*

The Differential Evolution (DE) algorithm is a vector-based global optimization algorithm that explores the global optimum of the objective function through operations such as differentiation, mutation, and selection among candidate solution vectors [21,25,26]. The DE algorithm is a metaheuristic algorithm, which implies that it does not rely on strict mathematical proofs and derivations in solving optimization problems; instead, DE conducts searches based on heuristic rules and strategies. The fundamental concept of DE is similar to that of genetic algorithms, which involves continuously updating the population through mutation, crossover, and selection operations until the optimal solution is found.

(1) Mutation. The DE algorithm employs the disparity between two individual vectors to induce mutation in the existing individuals, which subsequently utilize this mutation vector to generate new offspring individuals. Upon initializing the population, the individuals undergo mutation operations to form mutation vectors. The equation of the mutation process is as follows:

$$m_i(t) = x_{r1}(t) + F \times (x_{r2}(t) - x_{r3}(t)) \tag{1}$$

where $x_{r1}(t)$ , $x_{r2}(t)$, and $x_{r3}(t)$ are three mutually unequal individuals randomly selected from the initial population, and $F \in (0,1)$ is the mutation factor, which is used to control the variation rate of individuals.

(2) Crossover. The DE algorithm employs a certain probability to select the mutated vector of the offspring, thereby generating experimental individuals to enhance the population diversity and promote structural differentiation.

$$u_{i,j}(t) = \begin{cases} m_{i,j}(t), & rand(0,1) \le CR \text{ } or \text{ } j = rand(j) \\ x_{i,j}(t), & rand(0,1) > CR \text{ } and \text{ } j \ne rand(j) \end{cases} \tag{2}$$

where $CR \in (0,1)$ is the crossover factor, and $rand(j)$ is a randomly generated score vector that ensures that at least one variant individual provides at least a one-dimensional component after the crossover of the test individuals.

(3) Selection. The DE algorithm employs a greedy selection strategy, where the individuals are selected based on the values of the sought-after fitness function; the optimal candidates are selected from both offspring and parent classes.

$$x_i(g+1) = \begin{cases} u_i(t), & f(u_i(t)) \le f(x_i(t)) \\ x_i(t), & f(u_i(t)) > f(x_i(t)) \end{cases} \tag{3}$$

where $f(x)$ is the value of the fitness function.

### *2.2. Bayesian Optimization Algorithm*

The Bayesian Optimization algorithm (BOA), used to optimize black-box functions [27,28], has a primary advantage in its ability to find the global optimum with very few experimental iterations, thus making it particularly suitable for tasks that require many experiments, such as training SWMM models. The BOA algorithm achieves this capability by constructing a Gaussian process model to describe the probability distribution of the function to be optimized. Then, it balances exploration and exploitation by selecting the next experimental point and iteratively approaches the global optimum. The advantages of the BOA algorithm include considering previous parameter information, requiring fewer iterations, and its robustness.

The mathematical principles underlying the BOA algorithm are as follows:

$$P(X \mid Y) = \frac{P(Y \mid X)P(X)}{P(Y)} \tag{4}$$

where $X$ and $Y$ are random events; $P(X)$ is the prior probability of event $X$, i.e., the probability of occurrence of $X$; $P(Y \mid X)$ is the conditional probability of occurrence of event $Y$ given the occurrence of event $X$; $P(Y)$ is the probability of occurrence of event $Y$; and $P(X \mid Y)$ is the posterior probability of occurrence of event $X$ given the occurrence of event $Y$, i.e., the conditional probability.

The algorithm can be divided into two core components: (1) modeling the objective function, which entails calculating the mean and variance of the function values at each point, typically achieved through Gaussian process regression; and (2) constructing an acquisition function that determines the next sampling point. Using Gaussian process regression and the acquisition function, the algorithm determines new search points based on historical search points to enhance the precision and efficiency of the BOA algorithm.

Modeling the objective function,

$$f(x) \approx N\left(\mu(x), \sigma^2(x)\right) \tag{5}$$

where $N$ is the Gaussian distribution, $\mu(x)$ is the mean, and $\sigma^2(x)$ is the variance.

Using Gaussian process regression, the posterior probability of the fitness function values for any individual can be calculated. Subsequently, an acquisition function is constructed based on this posterior probability, which can identify the maximum points of the function as the following search points. By repeating this process, the optimal solution can be attained.

### *2.3. Improvement of the DE-BOA Algorithm*

The DE algorithm has been widely applied in the construction of SWMM models due to its adaptive adjustment of the search space and its ability to globally optimize complex nonlinear functions [29]. However, DE, which exhibits strong global search capabilities in the early stages, tends to slowly converge and possibly become trapped in local optima as it progresses. Meanwhile, the BOA algorithm effectively utilizes the information from the objective function to construct a Gaussian process regression model, which guides the selection of the following sampling point based on the acquisition function. This method enhances the efficiency and accuracy of parameter calibration and prevents the occurrence of local optima [20]. Nevertheless, for large-scale SWMM models, the construction of the Gaussian process model can become excessively complex, resulting in computational challenges and potentially rendering it infeasible.

To address these limitations, this paper proposes the following enhancements to the DE and BOA algorithms.

(1) Building upon the foundation of the DE algorithm, we introduce the Gaussian process regression model from the BOA algorithm. This model captures the posterior distribution of the target function based on existing sample points. Using the collection function to guide the selection of the following sampling points, we combine the global search capability of the DE algorithm with the local search capability of the BOA algorithm. This hybrid optimization approach prevents the parameters from becoming trapped in local optima during the calibration process, so it achieves a

high-precision parameter estimation for the SWMM model. Please refer to Figure 1 for the detailed algorithmic structure.

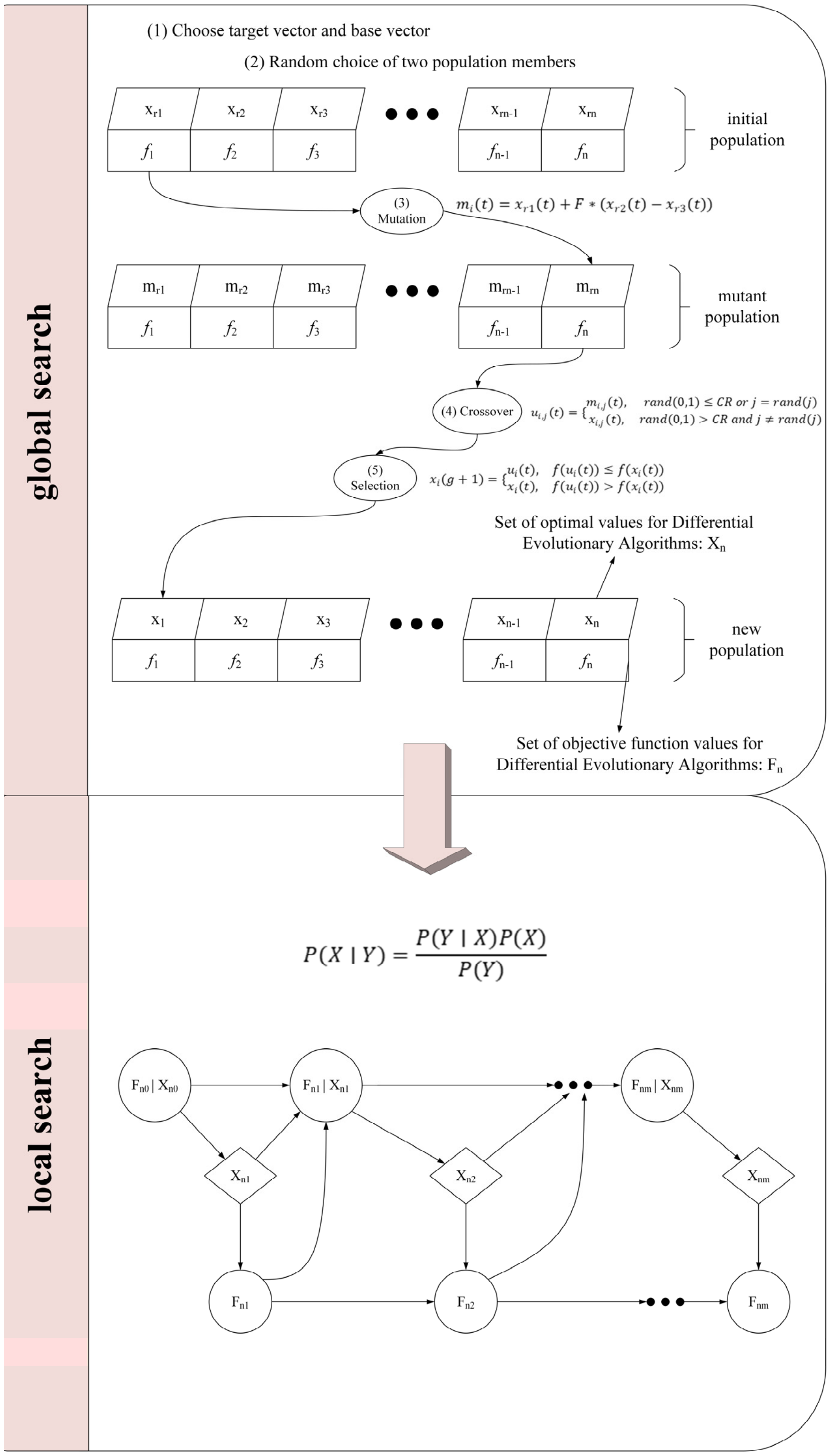


**Figure 1.** Structure of the Differential Evolution and Bayesian Optimization Algorithm.

(2) Using an adaptive weight allocation strategy, we tailor the distribution proportions of the DE algorithm and BOA algorithm, based on the balance between the constraints and the objective function of each specific SWMM model. During the initial stages of

optimization, when the initial population exhibits good dispersion, we leverage the randomness of the initial population by assigning higher weights to the DE algorithm. By invoking the DE algorithm multiple times in the early phases, we enhance the exploration for the global optimum. However, as the optimization progresses, to avoid falling into local optima, we dynamically adjust the weights and increase the frequency of utilizing the BOA algorithm. Leveraging the characteristics of the BOA algorithm, which consider a more significant amount of historical data, we enhance the precision of SWMM model calibration and more thoroughly explore the global optimum. Please refer to Figure 2 for the detailed algorithmic flow.

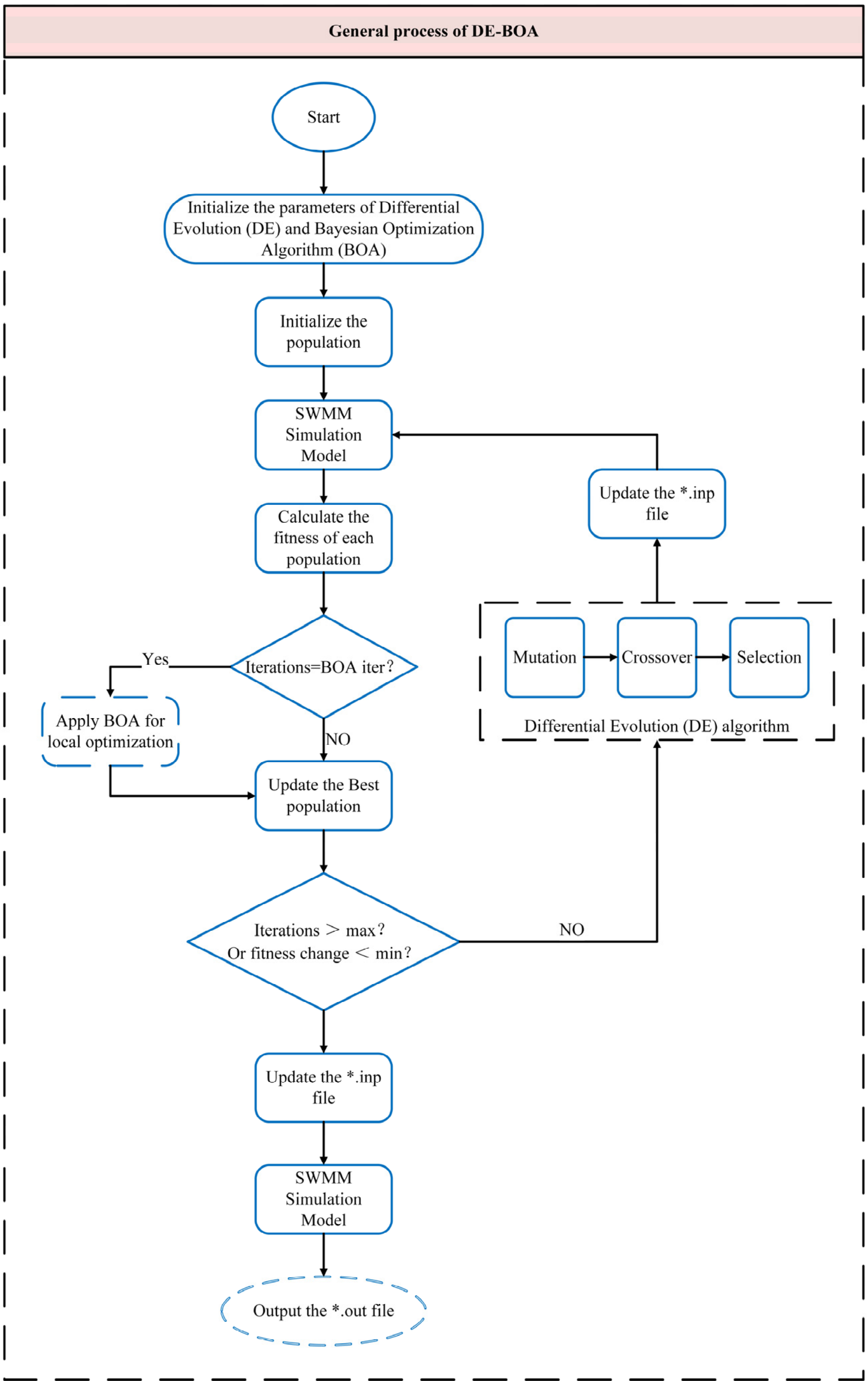


**Figure 2.** Flowchart of the Differential Evolution and Bayesian Optimization Algorithm.

### *2.4. Case Study*

#### 2.4.1. Overview of the Study Area

The present study selects an urban drainage sub-basin in Jinshazhou, Guangzhou, China, as the research case. The area of this sub-basin is 0.82 km$^2$, and its location is illustrated in Figure 3a. Water systems, green spaces, and buildings intertwine in a mosaic-like distribution within the region, while primary and secondary roads crisscross, which creates a diverse urban land structure. Among them, impervious surfaces cover an area of 0.3 km$^2$, which accounts for 36.6%, while pervious surfaces cover an area of 0.52 km$^2$, which accounts for 63.4% (the land use types are depicted in Figure 4).

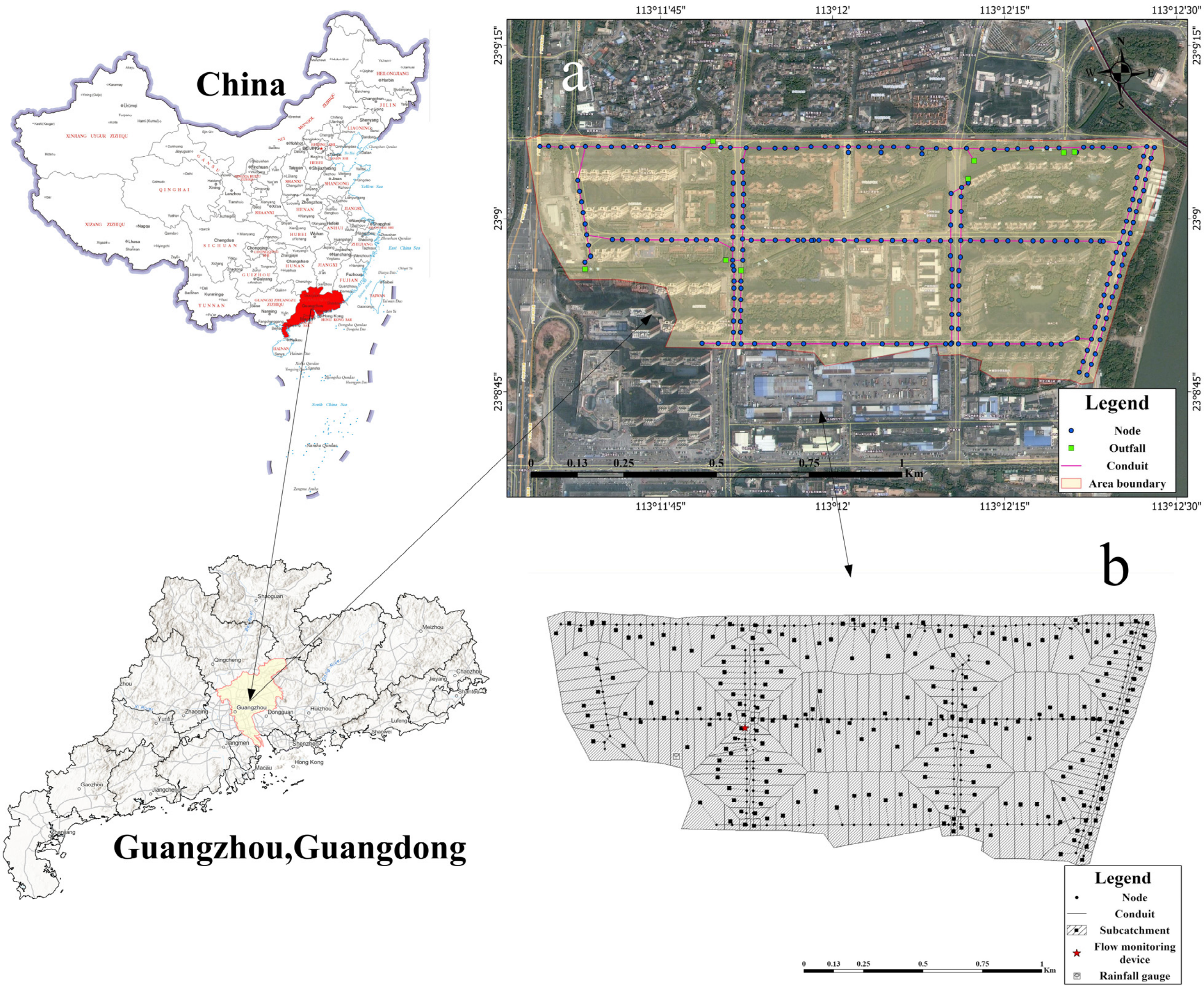


**Figure 3.** Location and model structure of the study area: (**a**) geographical location of the study area; (**b**) SWMM pipe network mode.

An urban stormwater runoff model was constructed based on the network data, land use types within the study area, and observed hydro-meteorological data. This SWMM model comprises 226 sub-catchments, 237 conduits, 226 nodes, and 8 outfalls, as depicted in Figure 3b. The rainfall and runoff data recorded by the rainfall gauge and flow monitoring device in Jinshazhou were utilized as the foundation for the precipitation and runoff analysis in this study.

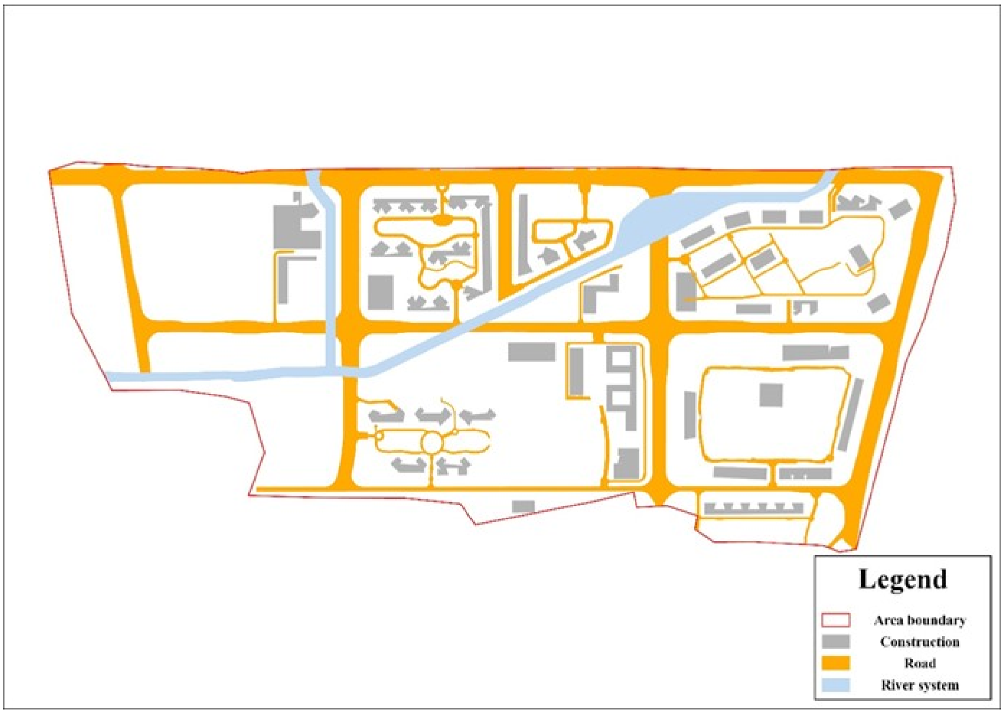


**Figure 4.** Land use types.

### 2.4.2. Analysis of Rainfall Events

This study collected independent rainfall events from April to August 2022. To conduct statistical analysis on these events, a daily interval was used to distinguish between them. Since rainfall events with small total amounts and longer durations have minimal impact on the stormwater network facilities in the study area, as input conditions, the SWMM model for the study area only utilized rainfall events with total precipitation exceeding 10 mm within 24 h. Table 1 lists the rainfall events.

**Table 1.** The statistics of rainfall events in the study area.

| ID | Start Time | End Time | Total Rainfall (mm) | Rainfall Duration (h) | Rainfall Scale |
|---|---|---|---|---|---|
| R1 | 30 April 2022 16:00 | 1 May 2022 0:00 | 14.02 | 8 h | Moderate rain |
| R2 | 1 May 2022 4:10 | 1 May 2022 20:20 | 36.02 | 16.17 h | Heavy rain |
| R3 | 7 May 2022 7:40 | 7 May 2022 18:20 | 19.87 | 10.67 h | Moderate rain |
| R4 | 10 May 2022 1:50 | 11 May 2022 0:00 | 48.47 | 22.17 h | Heavy rain |
| R5 | 11 May 2022 0:00 | 12 May 2022 0:00 | 81.26 | 24 h | Rainstorm |
| R6 | 12 May 2022 10:30 | 13 May 2022 20:20 | 24.65 | 9.84 h | Moderate rain |
| R7 | 13 May 2022 0:00 | 14 May 2022 0:00 | 80.75 | 24 h | Rainstorm |
| R8 | 21 May 2022 13:40 | 22 May 2022 0:00 | 26.91 | 11.67 h | Heavy rain |
| R9 | 6 June 2022 10:10 | 6 June 2022 22:50 | 26.73 | 12.67 h | Heavy rain |
| R10 | 7 June 2022 9:30 | 7 June 2022 19:10 | 43.70 | 9.67 h | Heavy rain |
| R11 | 9 June 2022 7:20 | 10 June 2022 12:30 | 32.78 | 5.17 h | Heavy rain |
| R12 | 12 June 2022 8:00 | 12 June 2022 18:30 | 25.18 | 10.5 h | Heavy rain |
| R13 | 13 June 2022 1:40 | 13 June 2022 16:50 | 35.09 | 15.17 h | Heavy rain |
| R14 | 15 June 2022 8:20 | 15 June 2022 20:50 | 34.64 | 12.5 h | Heavy rain |
| R15 | 17 June 2022 4:30 | 17 June 2022 17:10 | 32.30 | 12.67 h | Heavy rain |
| R16 | 19 June 2022 1:20 | 19 June 2022 14:20 | 23.78 | 13 h | Moderate rain |
| R17 | 2 July 2022 0:00 | 2 July 2022 22:40 | 46.73 | 22.67 h | Heavy rain |
| R18 | 6 July 2022 11:40 | 6 July 2022 19:40 | 30.94 | 8 h | Heavy rain |
| R19 | 3 August 2022 15:10 | 3 August 2022 18:40 | 23.00 | 3.5 h | Moderate rain |
| R20 | 10 August 2022 8:00 | 10 August 2022 16:30 | 27.09 | 8.5 h | Heavy rain |

By conducting statistical analysis on 20 rainfall events, we classified the rainfall intensity levels of these events based on the classification standards issued by the China Meteorological Administration, i.e., rainfall between 10 and 24.9 mm in 24 h is regarded as moderate rainfall, rainfall between 25.0 and 49.9 mm in 24 h is regarded as heavy rainfall, and rainfall between 50.0 and 99.9 mm in 24 h is regarded as torrential rainfall. Among them, moderate, heavy, and torrential rain accounted for 25, 65, and 10% of the total rainfall events, respectively.

Using the classification criteria proposed by Yin et al. [30], which are suitable for Guangzhou City, we classified the rainfall characteristics of the 20 rainfall events based on seven types of heavy rainfall patterns. Table 2 presents the results. The analysis reveals that the selected independent rainfall events of this study encompass all commonly observed rainfall patterns in Guangzhou City, which are predominately single-peak heavy rainfall.

**Table 2.** Rainfall event rain type.

| Rainfall Pattern | Frequency (Times) | Proportion |
|---|---|---|
| Single peak ahead (type I) | 1 | 5% |
| Single peak back (type II) | 5 | 25% |
| Single peak centered (type III) | 5 | 25% |
| Uniform distribution (type IV) | 1 | 5% |
| Distribution before and after bimodal (type V) | 1 | 5% |
| Bimodal distribution (type VI) | 4 | 20% |
| Bimodal mid-post distribution (type VII) | 3 | 15% |

2.4.3. SWMM Model Parameters

In the SWMM model, parameters are typically classified into two types: those that can be directly measured through field surveys, and those that lack a precise physical definition, which makes them challenging to directly obtain. The latter type is usually estimated based on empirical ranges, which require the identification of sensitive parameters that significantly impact the model to improve the efficiency of calibration [31].

In this paper, the improved Morris screening method [31–33] is selected for sensitivity analysis, i.e., the idea of the control variable method is adopted to change the value of a parameter in a fixed step to affect the simulation results, under the condition of ensuring that the other parameters remain unchanged, and the average rate of change of the model results is taken as the sensitivity value of the parameter, which is calculated using the following formulas:

$$S = \sum_{i=1}^{n} \frac{(Y_{i+1} - Y_i)/Y_0}{(P_{i+1} - P_i)/100} / n \quad (6)$$

where $S$ is the sensitivity value; $Y_\mathrm{i}$ is the output value of the ith run of the model; $Y_{\mathrm{i}+1}$ is the output value of the $i$+1 th run of the model; $Y_0$ is the baseline output value of the model; $P_i$ is the percentage change in the parameter value of the ith model run relative to the calibrated parameter value; $P_{\mathrm{i}+1}$ is the percentage change in the parameter value of the $i$+1 st model run relative to the calibrated parameter value; and $n$ is the number of model runs.

When the value of $S$ is less than 0.05, the parameter is insensitive. If the value of $S$ is between 0.05 and 0.2, the parameter is moderately sensitive. If the value of $S$ is between 0.2 and 1.0, the parameter is considered sensitive. A value of $S$ greater than 1.0 indicates that the parameter is very sensitive. In this paper, the initial parameter value is changed in fixed steps of 10%, and the results of increasing its value by −25%, −15%, −5%, 5%, 15%, and 25% are obtained to calculate the sensitivity value of the parameter to total runoff, flood flow, and the peak present time, and the average value of $S$ is taken as the main index for sensitivity analysis.

The results of the parameter sensitivity analysis are shown in Table 3, and the range of parameter values is derived from a combination of local engineering experience values and the SWMM User's Manual.

**Table 3.** Parameters to be determined and value ranges.

| Calibration Parameters | Parameter Meaning | $S$ | Ranges |
|---|---|---|---|
| N-Imperv | Manning coefficient of impermeable area | 0.069567 | 0.01~0.05 |
| N-Perv | Manning coefficient of permeable area | 0.054767 | 0.05~0.4 |
| S-Imperv/mm | Depression volume in impermeable area | 0.054133 | 0.2~10 |
| S-Perv/mm | Depression storage volume in permeable area | 0.0879 | 2~20 |
| MinRate/(mm·h$^{-1}$) | Minimum infiltration rate | 0.051633 | 1~20 |
| MaxRate/(mm·h$^{-1}$) | Maximum infiltration rate | 0.053167 | 20~100 |
| Decay/h$^{-1}$ | Permeation attenuation coefficient | 0.072767 | 2~7 |

#### 2.4.4. SWMM Optimization Calculation

1. Formulate the objective function

Given the significance of the peak flow values and peak occurrence time in urban flood disasters, this study is founded upon minimizing disparities in peak values and peak occurrence times. By integrating the Nash–Sutcliffe Efficiency (NSE), a multi-objective function is established to conduct a fitting analysis of the observed rainfall data and simulated results.

$$num0 = \sum_{i=1}^{n} I\left(\mid Q_{o,t} - Q_{p,t} \mid > P_{dt}\right) \tag{7}$$

where $n$ is the number of peaks, $Q_{\mathrm{o},t}$ is the observed flow at the '$t$'-th time step, $Q_{\mathrm{p},t}$ is the simulated flow at the '$t$'-th time step, and $I$ is the indicator function. The indicator function takes the value of 1 when the absolute difference between $Q_{\mathrm{o},t}$ and $Q_{\mathrm{p},t}$ exceeds the tolerance for peak differences $P_{dt}$, and 0 otherwise.

$$num1 = \sum_{i=1}^{n} I\left(\mid T_{o,t} - T_{p,t} \mid > T_{dt}\right) \tag{8}$$

where $T_{o,t}$ is the peak occurrence time of the observed flow at the '$t$'-th time step, $T_{p,t}$ is the peak occurrence time of the simulated flow at the '$t$'-th time step, and $I$ is the indicator function. The indicator function takes the value of 1 when the absolute difference between $T_{o,t}$ and $T_{p,t}$ exceeds the tolerance for peak differences $T_{dt}$, and 0 otherwise.

$$NSE = 1 - \frac{\sum_{t=1}^{T}\left(Q_{\mathrm{o},t} - Q_{\mathrm{p},t}\right)^2}{\sum_{t=1}^{T}\left(Q_{\mathrm{o},t} - Q_{\mathrm{o,m}}\right)^2} \tag{9}$$

where $T$ is the length of the time series,$Q_{\mathrm{o,m}}$ is the mean of the observed flow, and an $NSE$ value closer to 1 indicates a more favorable calibration outcome of the model.

By combining Equations (7)–(9), we formulate a multi-objective function:

$$F = [num0 + num1] \times P + NSE \tag{10}$$

where $P$ is a penalty factor to penalize the larger variance values that occur in the peak and peak present time.

2. Simulation procedure
    (1) The SWMM model of the study area was simulated using the Horton infiltration model and the dynamic wave method. Four rainfall events (R3, R4, R7, and R13) were set aside for subsequent model validation, while the remaining 16 rainfall events were utilized as input conditions to calibrate the model parameters.
    (2) The population was initialized and the fundamental parameters of the DE-BOA algorithm were set, based on the experience of multiple iterations of trial and error in the previous sensitivity analysis: population size = 20; mutation factor = 0.8; crossover factor = 0.7; maximum iteration count = 100. Bayesian

iteration factors were chosen as 5 and 2, respectively. Batch import and export of SWMM parameters was implemented using the Python language.

(3) Three solutions were randomly selected from the initial population and trial individuals were generated using the mutation and crossover formulas. The Python language was utilized to invoke the SWMM calculation engine for simulation computation and to obtain simulated water depths at the monitoring locations.

(4) The *.out file generated by the SWMM simulation was retrieved; the objective function values were calculated after simulation; a greedy selection strategy was employed; and the optimal value was selected as the calibrated parameter value. The *.inp file was modified to serve as the input file for the next round of SWMM simulation.

(5) At the early stage of the algorithm, the BOA algorithm was applied every five generations based on the optimal value to locally optimize it. As the algorithm progressed, an adaptive strategy to increase the proportion of the BOA algorithm was used and the influence factor of the optimal historical value on the computation was adjusted. Multiple rounds of local optimization were performed based on the optimal value and a greedy selection strategy was applied to the results.

(6) The calibrated *.inp file was generated as output.

3. Evaluation criteria

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Q_{o,t} - Q_{p,t}\right)^2} \tag{11}$$

where $RMSE$ denotes the Root Mean Square Error, which is used to gauge the deviation between simulated flow and measured flow. A small value signifies a minor simulation error.

$$R^2 = \frac{\left[\sum_{t=1}^{T}\left(Q_{p,\,t} - Q_{p,\,m}\right)\left(Q_{o,\,t} - Q_{o,\,m}\right)\right]^2}{\sum_{t=1}^{T}\left(Q_{p,\,t} - Q_{p,\,m}\right)^2\sum_{t=1}^{T}\left(Q_{o,\,t} - Q_{o,\,m}\right)^2} \tag{12}$$

where $R^2$ is the coefficient of determination to assess the degree of fit and modeling capability of the algorithm. A higher value indicates a greater explanatory power of the independent variable on the dependent variable. $Q_{p,t}$ is the mean of the measured flow.

$$PE = \frac{Q_p^p - Q_o^p}{Q_o^p} \tag{13}$$

where $PE$ is the peak relative error, which serves as a measure of the algorithm's deviation in extreme situations. A smaller value signifies better simulation performance. $Q_p^p$ and $Q_o^p$ are the peak values of the simulated flow and measured flow, respectively.

## 3. Results and Discussion

### *3.1. Parameter Calibration Results*

The model underwent parameter calibration using a selection of 16 rainfall events from Table 1. Taking R5 as an example, the calibrated flow process was compared to the observed flow process for the manual calibration, Bayesian optimization algorithm, differential evolution algorithm, and DE-BOA algorithm (Figure 5). With equal numbers of iterations, the objective function values were recorded to be 0.8846, 0.8704, 0.8795, and 0.9402. Compared to the other methods, the enhanced algorithm exhibited improvements of 6.29%, 8.02%, and 6.91%. Figure 5 shows that the DE-BOA algorithm outperformed the other approaches in flow prediction, since its peak flow and curve trends are more closely aligned with the observed flow.

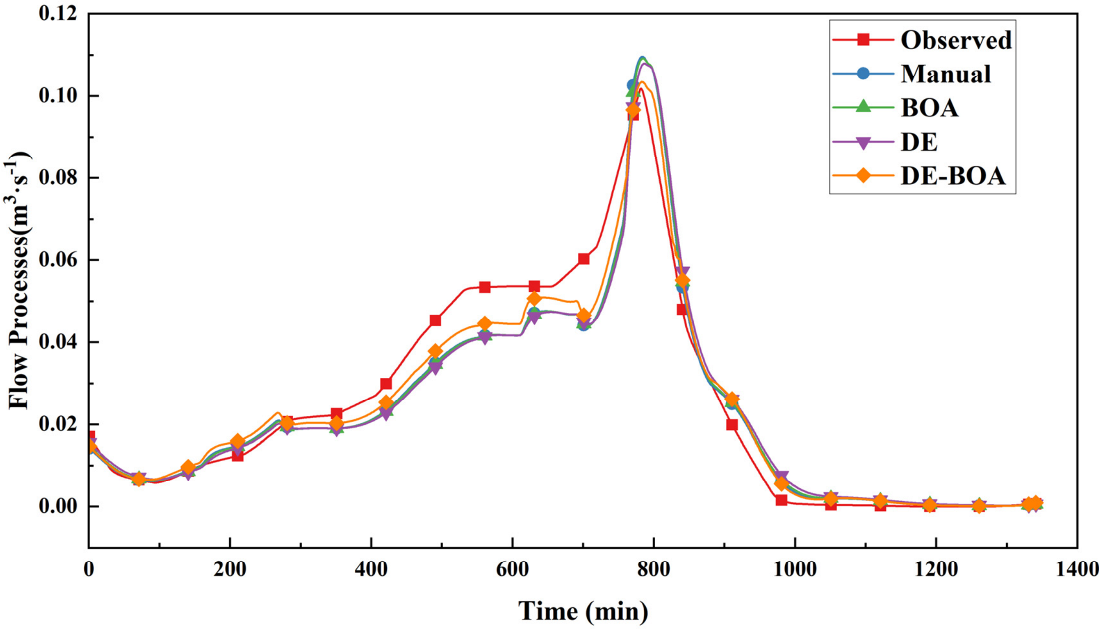


**Figure 5.** Comparison of flow processes for different rate determination methods.

Furthermore, different calibration methods were compared based on the evaluation metrics, as outlined in Table 4. The DE-BOA algorithm exhibits better RMSE, $R^2$, and PE than the other methods to various degrees. The RMSE and $R^2$ values indicate that, when calibrated using the DE-BOA algorithm, there is a good fit between observed and simulated flows. The PE value indicates that even in extreme circumstances, the DE-BOA algorithm maintains a commendable level of simulation performance.

**Table 4.** Evaluation metrics for simulating results using different calibration methods.

| Method | BOA | DE | Manual | DE-BOA |
|---|---|---|---|---|
| RMSE | 0.0071 | 0.0075 | 0.0070 | 0.0052 |
| $R^2$ | 0.9300 | 0.9216 | 0.9339 | 0.9644 |
| PE | 0.0727 | 0.0605 | 0.0768 | 0.0161 |

The parameter rate setting speed of the algorithms is also one of the decisive factors affecting the application of intelligent algorithms. In this study, when comparing the speeds of DE, BOA, and DE-BOA algorithms to be the first to find the optimal objective function in 100 iterations (refer to Figure 6), the BOA algorithm achieved the optimal objective function value of 0.8704 in the 20th iteration, the DE algorithm achieved the optimal objective function value of 0.8795 in the 26th iteration, and the DE-BOA algorithm achieved the optimal objective function value of 0.9402 in the 13th iteration. This shows that the DE-BOA algorithm is not only faster but also more accurate in SWMM parameterization.

Based on the computed objective function values from the 16 rainfall instances, distinct weights were assigned. These weights were employed to calculate the calibrated model parameters (refer to Table 5). In comparison to the parameters obtained through manual calibration, the integrated approach demonstrated enhanced precision and substantial variations in specific parameter values. The primary underlying factor is that manual calibration typically relies on a trial-and-error method, which focuses on single-parameter adjustments. Hence, it fails to account for the intricate interplay among multiple parameters and is subjected to pronounced subjective inclinations. In contrast, the DE-BOA algorithm leverages its algorithmic traits to perform comprehensive global exploration and meticulous local refinement, thus enabling the search for optimal values within the vector solution space, which encompasses all calibrated parameters.

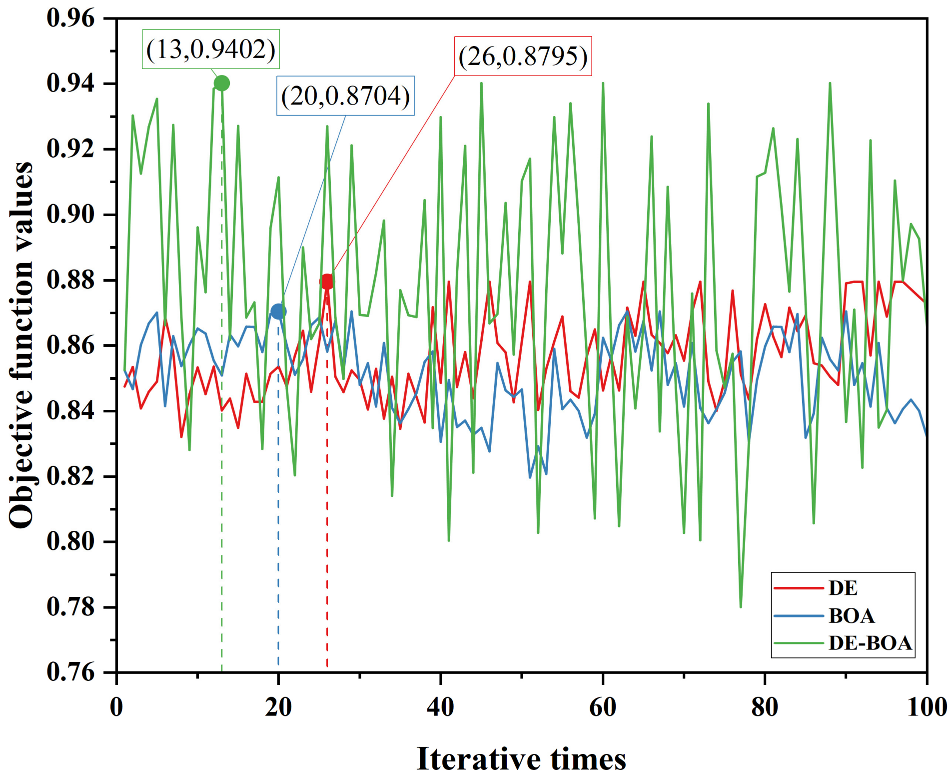


**Figure 6.** Comparison of rate speeds for different rate determination methods.

**Table 5.** SWMM calibration results.

| Parameters | Initial Empirical Value | Rate Value | |
|---|---|---|---|
| | | Manual Calibration | DE-BOA Calibration |
| N-Imperv | 0.012 | 0.01 | 0.01 |
| N-Perv | 0.2 | 0.4 | 0.3528 |
| S-Imperv/mm | 0.45 | 0.2 | 0.2 |
| S-Perv/mm | 2 | 20 | 17.3671 |
| MinRate/(mm·h$^{-1}$) | 12.7 | 18 | 11.0311 |
| MaxRate/(mm·h$^{-1}$) | 200 | 90 | 49.8109 |
| Decay/h$^{-1}$ | 4 | 3 | 2.9728 |

### *3.2. Model Verification*

To assess the accuracy and credibility of the calibrated model, we input the four rainfall events selected for validation (refer to Table 6) into the SWMM model. These events (R3, R4, R7, and R13) include various rainfall intensities and types within the study area, which enables an exploration of the impact of different rainfall event types on the model's simulation.

**Table 6.** Rainfall events of the validation set.

| ID | Total Rainfall (mm) | Rainfall Duration (h) | Rainfall Scale | Rainfall Pattern |
|---|---|---|---|---|
| R3 | 19.87 | 10.67 h | Moderate rain | IV |
| R4 | 48.47 | 22.17 h | Heavy rain | VII |
| R7 | 80.75 | 24 h | Rainstorm | II |
| R13 | 35.09 | 15.17 h | Heavy rain | V |

In Figure 7, the comparison between calibrated and measured flow processes of the four rainfall events reveals that all NSE values of the model exceed 0.85. This result signifies a remarkable concordance between simulated and observed flow processes, which satisfies the application requirements of the SWMM model. R4 and R13 exhibit the highest NSE values of 0.9317 and 0.9306, respectively, which can be attributed to their correlation with the rainfall dynamics. When multiple rainfall peaks and substantial precipitation occur, the predictive accuracy of the model more closely aligns with the actual observed values. Thus, the DE-BOA algorithm exhibits greater adaptability when handling rainfall data characterized by multiple peaks and a significant rainfall volume.

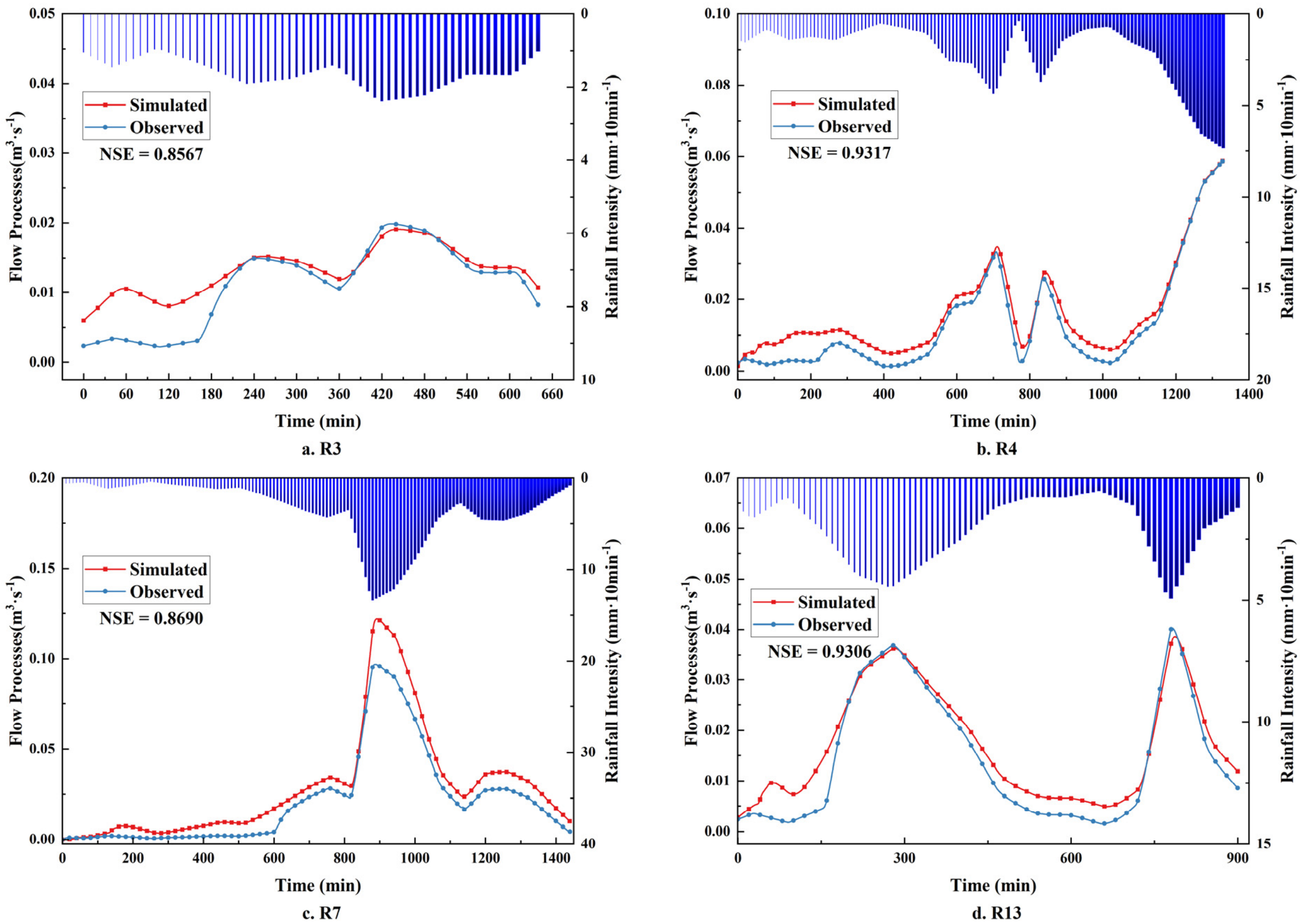


**Figure 7.** Verification of model calibration results.

## 4. Conclusions

The present study proposed a calibration method for the SWMM model based on the DE-BOA algorithm. To demonstrate the feasibility and applicability of the DE-BOA algorithm, a case study was conducted in a drainage sub-catchment in Jinshazhou, Guangzhou. The rainfall events in the area were analyzed for their characteristics, and a multi-objective function was constructed to evaluate the algorithm's performance. The comparison of the simulated flow process of the algorithm with the rest of the rate determination methods shows that the DE-BOA algorithm performs well on the evaluation metrics RMSE, R2, and PE, which reach 0.0052, 0.9644, and 0.0161, respectively, and the objective function value is improved by 6.29%, 8.02%, and 6.91% compared with the rest of the methods, and meanwhile, in terms of the rate setting rate and the rate setting process, the DE-BOA algorithm also shows obvious advantages. The algorithmic model enables the automatic parameter calibration of SWMM with high accuracy and at a faster rate, which effectively improves the modeling efficiency. This result provides valuable insights for the dynamic updating of model parameters and improves the performance and accuracy of modeling on intelligent water management platforms.

The calibration validation of the model demonstrates a significant improvement in the calibration effectiveness of the DE-BOA algorithm for rainfall events with multiple peaks and substantial precipitation. It must be admitted that this study has some limitations, because, due to the confidentiality of the geographic information data, we only obtained a limited number of sample data, which may affect the generalization of the results. Therefore, the question of how to further improve the algorithm rate accuracy, as well as the generalizability, will be considered in future research.

**Author Contributions:** Conceptualization, J.L. and J.G.; methodology, J.L. and J.G.; investigation, Y.L. and R.Z.; data curation, J.L.; writing—original draft preparation, J.G.; writing—review and editing, J.L. and X.L.; supervision, J.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was supported by the National Key R&D Program of China (2022YFC3002704) and the Natural Science Foundation of Hubei Province of China (2022CFB042).

**Data Availability Statement:** All data, models, and code that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** We are grateful to the editors and anonymous reviewers for their thoughtful comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hossain, S.; Hewa, G.A.; Wella-Hewage, S. A Comparison of Continuous and Event-Based Rainfall–Runoff (RR) Modelling Using EPA-SWMM. *Water* **2019**, *11*, 611. [CrossRef]
2. Tsihrintzis, V.A.; Hamid, R. Runoff Quality Prediction from Small Urban Catchments Using SWMM. *Hydrol. Process.* **1998**, *12*, 311–329. [CrossRef]
3. Jang, S.; Cho, M.; Yoon, J.; Yoon, Y.; Kim, S.; Kim, G.; Kim, L.; Aksoy, H. Using SWMM as a Tool for Hydrologic Impact Assessment. *Desalination* **2007**, *212*, 344–356. [CrossRef]
4. Yao, L.; Chen, L.; Wei, W. Assessing the Effectiveness of Imperviousness on Stormwater Runoff in Micro Urban Catchments by Model Simulation. *Hydrol. Process.* **2016**, *30*, 1836–1848. [CrossRef]
5. Guan, M.; Sillanpää, N.; Koivusalo, H. Storm Runoff Response to Rainfall Pattern, Magnitude and Urbanization in a Developing Urban Catchment. *Hydrol. Process.* **2016**, *30*, 543–557. [CrossRef]
6. Sun, N.; Hong, B.; Hall, M. Assessment of the SWMM Model Uncertainties within the Generalized Likelihood Uncertainty Estimation (GLUE) Framework for a High-Resolution Urban Sewershed. *Hydrol. Process.* **2014**, *28*, 3018–3034. [CrossRef]
7. Zaghloul, N.A. Sensitivity Analysis of the SWMM Runoff-Transport Parameters and the Effects of Catchment Discretisation. *Adv. Water Resour.* **1983**, *6*, 214–223. [CrossRef]
8. Barco, J.; Wong, K.M.; Stenstrom, M.K. Automatic Calibration of the U.S. EPA SWMM Model for a Large Urban Catchment. *J. Hydraul. Eng.* **2008**, *134*, 466–474. [CrossRef]
9. Del Giudice, G.; Padulano, R. Sensitivity Analysis and Calibration of a Rainfall-Runoff Model with the Combined Use of EPA-SWMM and Genetic Algorithm. *Acta Geophys.* **2016**, *64*, 1755–1778. [CrossRef]
10. Kim, S.W.; Kwon, S.H.; Jung, D. Development of a Multiobjective Automatic Parameter-Calibration Framework for Urban Drainage Systems. *Sustainability* **2022**, *14*, 8350. [CrossRef]
11. Tscheikner-Gratl, F.; Zeisl, P.; Kinzel, C.; Leimgruber, J.; Ertl, T.; Rauch, W.; Kleidorfer, M. Lost in Calibration: Why People Still Do Not Calibrate Their Models, and Why They Still Should—A Case Study from Urban Drainage Modelling. *Water Sci. Technol.* **2016**, *74*, 2337–2348. [CrossRef]
12. Gulbaz, S.; Kazezyilmaz-Alhan, C.M. Calibrated Hydrodynamic Model for Sazlidere Watershed in Istanbul and Investigation of Urbanization Effects. *J. Hydrol. Eng.* **2013**, *18*, 75–84. [CrossRef]
13. Jin, X.; Jiang, Y.; Wu, W.; Jin, J. Automatic Calibration of SWMM Model with Adaptive Genetic Algorithm. In Proceedings of the 2011 International Symposium on Water Resource and Environmental Protection, Xi'an, China, 20–22 May 2011; Volume 2, pp. 891–895.
14. Wan, B.; James, W. SWMM Calibration Using Genetic Algorithms. In *Global Solutions for Urban Drainage*; American Society of Civil Engineers: Reston, VA, USA, 2012; pp. 1–14. [CrossRef]
15. Gao, X.; Yang, Z.; Han, D.; Huang, G.; Zhu, Q. A Framework for Automatic Calibration of SWMM Considering Input Uncertainty. *Hydrol. Earth Syst. Sci. Discuss.* **2020**, 1–25. [CrossRef]

16. Perin, R.; Trigatti, M.; Nicolini, M.; Campolo, M.; Goi, D. Automated Calibration of the EPA-SWMM Model for a Small Suburban Catchment Using PEST: A Case Study. *Env. Monit Assess* **2020**, *192*, 374. [CrossRef]
17. Shahed Behrouz, M.; Zhu, Z.; Matott, L.S.; Rabideau, A.J. A New Tool for Automatic Calibration of the Storm Water Management Model (SWMM). *J. Hydrol.* **2020**, *581*, 124436. [CrossRef]
18. Azad, A.; Manoochehri, M.; Kashi, H.; Farzin, S.; Karami, H.; Nourani, V.; Shiri, J. Comparative Evaluation of Intelligent Algorithms to Improve Adaptive Neuro-Fuzzy Inference System Performance in Precipitation Modelling. *J. Hydrol.* **2019**, *571*, 214–224. [CrossRef]
19. Farzin, S.; Valikhan Anaraki, M. Optimal Construction of an Open Channel by Considering Different Conditions and Uncertainty: Application of Evolutionary Methods. *Eng. Optim.* **2021**, *53*, 1173–1191. [CrossRef]
20. Sambito, M.; Freni, G. Strategies for Improving Optimal Positioning of Quality Sensors in Urban Drainage Systems for Non-Conservative Contaminants. *Water* **2021**, *13*, 934. [CrossRef]
21. Storn, R.; Price, K. Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization Over Continuous Spaces. *J. Glob. Optim.* **1995**, *23*, 1–15.
22. Bai, Y.; Lu, W.; Li, J.; Chang, Z.; Wang, H. Groundwater Contamination Source Identification Using Improved Differential Evolution Markov Chain Algorithm. *Env. Sci Pollut Res* **2022**, *29*, 19679–19692. [CrossRef]
23. Smith, T.J.; Marshall, L.A. Bayesian Methods in Hydrologic Modeling: A Study of Recent Advancements in Markov Chain Monte Carlo Techniques. *Water Resour. Res.* **2008**, *44*, 1–9. [CrossRef]
24. McDonnell, B.E.; Ratliff, K.; Tryby, M.E.; Wu, J.J.X.; Mullapudi, A. PySWMM: The Python Interface to Stormwater Management Model (SWMM). *J. Open Source Softw.* **2020**, *5*, 2292. [CrossRef]
25. 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]
26. Brest, J.; Greiner, S.; Boskovic, B.; Mernik, M.; Zumer, V. Self-Adapting Control Parameters in Differential Evolution: A Comparative Study on Numerical Benchmark Problems. *IEEE Trans. Evol. Comput.* **2006**, *10*, 646–657. [CrossRef]
27. Shao, Z.; Xu, L.; Chai, H.; Yost, S.A.; Zheng, Z.; Wu, Z.; He, Q. A Bayesian-SWMM Coupled Stochastic Model Developed to Reconstruct the Complete Profile of an Unknown Discharging Incidence in Sewer Networks. *J. Environ. Manag.* **2021**, *297*, 113211. [CrossRef]
28. Muleta, M.K.; McMillan, J.; Amenu, G.G.; Burian, S.J. Bayesian Approach for Uncertainty Analysis of an Urban Storm Water Model and Its Application to a Heavily Urbanized Watershed. *J. Hydrol. Eng.* **2013**, *18*, 1360–1371. [CrossRef]
29. Yazdi, J.; Yoo, D.G.; Kim, J.H. Comparative Study of Multi-Objective Evolutionary Algorithms for Hydraulic Rehabilitation of Urban Drainage Networks. *Urban Water J.* **2017**, *14*, 483–492. [CrossRef]
30. Yin, L.; Chen, X.; Chen, Z.; Shuai, W.; Zhang, M.; Liu, H. Study on Storm Pattern of Typical Rainfall Station in Guangzhou. *J. Water Resour. Res.* **2013**, *2*, 409. [CrossRef]
31. Zhong, B.; Wang, Z.; Yang, H.; Xu, H.; Gao, M.; Liang, Q. Parameter Optimization of SWMM Model Using Integrated Morris and GLUE Methods. *Water* **2023**, *15*, 149. [CrossRef]
32. Lenhart, T.; Eckhardt, K.; Fohrer, N.; Frede, H.-G. Comparison of Two Different Approaches of Sensitivity Analysis. *Phys. Chem. Earth Parts A/B/C* **2002**, *27*, 645–654. [CrossRef]
33. Morris, M.D. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics* **1991**, *33*, 161–174. [CrossRef]