



# Using Machine Learning to Identify and Optimize Sensitive Parameters in Urban Flood Model Considering Subsurface Characteristics

Hengxu Jin<sup>1,2</sup> · Yu Zhao<sup>1,2</sup> · Pengcheng Lu<sup>1,2</sup> · Shuliang Zhang<sup>1,2</sup> · Yiwen Chen<sup>1,2</sup> · Shanghua Zheng<sup>1,2</sup> · Zhizhou Zhu<sup>1,2</sup>

Accepted: 21 January 2024 / Published online: 21 February 2024  
© The Author(s) 2024

## Abstract

This study presents a novel method for optimizing parameters in urban flood models, aiming to address the tedious and complex issues associated with parameter optimization. First, a coupled one-dimensional pipe network runoff model and a two-dimensional surface runoff model were integrated to construct an interpretable urban flood model. Next, a principle for dividing urban hydrological response units was introduced, incorporating surface attribute features. The K-means algorithm was used to explore the clustering patterns of the uncertain parameters in the model, and an artificial neural network (ANN) was employed to identify the sensitive parameters. Finally, a genetic algorithm (GA) was used to calibrate the parameter thresholds of the sub-catchment units in different urban land-use zones within the flood model. The results demonstrate that the parameter optimization method based on K-means-ANN-GA achieved an average Nash-Sutcliffe efficiency coefficient (*NSE*) of 0.81. Compared to the ANN-GA and K-means-deep neural networks (DNN) methods, the proposed method better characterizes the runoff generation and flow processes. This study demonstrates the significant potential of combining machine learning techniques with physical knowledge in parameter optimization research for flood models.

**Keywords** Artificial neural network · Coupled urban flooding model · Genetic algorithm · K-means algorithm · Sub-catchment delineation · Uncertain parameters

## 1 Introduction

In recent years, urban flood disasters caused by sudden heavy rainfall have become increasingly severe, posing a serious threat to urban public infrastructure and the safety of residents' lives and property (Nguyen and Bae 2020; Lu and Sun 2021). Flood forecasting is one of the important non-engineering measures for flood control and disaster reduction, and the construction of a waterlogging model is the cornerstone of urban waterlogging forecasting (Chang et al. 2021; Li et al. 2022; Nandi and Reddy 2022). Urban waterlogging models involve a large number of parameters (Zeng et al. 2020; Liao et al. 2022), and a significant portion

of them cannot be directly obtained from measurable catchment characteristics (Sinnathamby et al. 2017; Guo and Su 2019; Wang et al. 2020). However, the accuracy of runoff depth simulation largely depends on how relevant parameters are defined (Huo and Liu 2019; Feigl et al. 2022a). Sensitivity analysis of the model parameters is fundamental for improving the efficiency and accuracy of model simulations. However, traditional methods for parameter sensitivity analysis involve complex processes, which severely hinder the efficiency of sensitivity analysis for urban waterlogging models.

Therefore, it is crucial to efficiently identify and optimize these sensitive parameters (Wood et al. 2016; Willis et al. 2019). Parameter optimization calibration is a key step in simulating urban waterlogging models and can be performed through manual or automatic calibration methods (Jung et al. 2017; Feigl et al. 2022b; Katipoğlu and Sarıgöl 2023). In the past, manual calibration methods, which are tedious and time consuming, were commonly used. To overcome the difficulties of manual parameter optimization calibration, researchers, both in China and internationally, have

✉ Shuliang Zhang  
zhangshuliang@njnu.edu.cn

<sup>1</sup> Key Laboratory of VGE of Ministry of Education, Nanjing Normal University, Nanjing 210023, China

<sup>2</sup> Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

developed computer-based automatic optimization methods. For instance, Wu et al. (2021) proposed a deep-learning-based method for optimizing the uncertainty parameters in flood processes. Yuan et al. (2021) implemented automatic calibration of rainfall-runoff model parameters using a BP neural network algorithm. Wang et al. (2022) addressed the issue of the inability to automatically calibrate parameters in the optimization module of flood models and proposed an automatic calibration method based on a genetic algorithm (GA) for rainfall-runoff models.

In recent years, there have been rapid developments in the field of deep learning. Compared to traditional hydrodynamic methods, deep learning resembles a “black box” (Adnan et al. 2021). While the internal structure and physical mechanisms of these methods are not fully understood, they have the ability to quickly capture trends and relationships in the data through extensive training (Yan et al. 2021; Ye et al. 2022). Therefore, the application of methods based on artificial neural networks (ANN) and other deep learning techniques for the rapid identification of sensitive parameters in urban waterlogging models is worth exploring. The K-means clustering machine learning algorithm (referred to as K-means) is widely used in flood forecasting research owing to its simple mathematical principles and fast convergence speed (Xu and Peng 2015). Li et al. (2016) used flood similarity to expand real-time corrective information and combined it with a K-means algorithm to achieve flood classification and forecasting in a transitional river basin. Hu et al. (2022) constructed a rapid flood classification forecasting model for the Jingle Basin based on K-means and backpropagation (BP) neural networks. Sun et al. (2022) applied improved sub-watershed division rules combined with the K-means algorithm for parameter calibration in the Storm Water Management Model (SWMM). However, there is a lack of discussion regarding the physical significance of the model parameters and the universal laws that govern the relationship between these parameters and the complex underlying urban surface. The application of the K-means algorithm in flood forecasting research has mostly focused on watershed areas and often classifies floods based on the characteristics of rainfall-flood events, with few studies exploring the sensitive parameters of urban waterlogging models.

In recent years, urbanization has led to increased levels of complexity in underlying surface conditions. Although traditional pipe network routing models perform well in simulating water routing within a network, they cannot provide information on the extent of surface inundation and water depths (Yang et al. 2020). This limitation makes it challenging to simulate the two-dimensional inundation process under complex urban conditions (Rai et al. 2016; Shahed Behrouz et al. 2020). Surface runoff models have the advantage of simulating the extent, depth, and process

of urban inundation based on overflow processes (that is, flow rate versus time relationship) at overflow nodes and urban topography. But such models do not take into account underground space (Zeng et al. 2017; Dao et al. 2022; Yang et al. 2022). Therefore, this study combined a one-dimensional pipe network routing model with a two-dimensional surface runoff model to construct an integrated urban waterlogging model.

Although the efficiency of identifying and optimizing sensitive parameters in urban flood models can be improved with the help of high-performance computers and machine learning methods, current research largely overlooks the physical significance of these model parameters and the universal laws that exist between these parameters and complex urban underlying surfaces of urban areas (Zang et al. 2022). The parameters are often assigned in a simplistic manner, and the sensitivity analysis of these parameters often involves cumbersome and complex processes, such as multiple simulations of the model. Furthermore, these studies have often directly used the simulation results of one-dimensional hydrological models to represent the surface inundation situation without considering the two-dimensional hydrodynamic processes of surface water runoff (Cai et al. 2019).

Investigating the differences in sensitive parameters among different land use functional zones in cities is helpful for reflecting the actual conditions of the study area and improving the efficiency and accuracy of urban flood simulations. Therefore, in this study, we proposed a principle for dividing urban hydrological response units based on the coupling model of the pipe network and surface. This principle incorporates the surface attribute features. Subsequently, we employed K-means clustering to explore the clustering patterns of the uncertain model parameters and identify sensitive parameters using artificial neural networks. Finally, we calibrated the threshold values of the sensitive parameters for the sub-watershed units in different land use functional zones using a genetic algorithm.

## 2 Methodology

In urban flood simulations, there are several issues, such as inadequate consideration of urban underlying surface attributes, unclear principles for dividing urban hydrological response units, and the complex and tedious parameter optimization process. In view of these issues, this study considered the characteristics of underlying urban surfaces and used the K-means-ANN-GA machine learning method to identify and optimize sensitive parameters in urban flood models. The research framework is shown in Fig. 1.

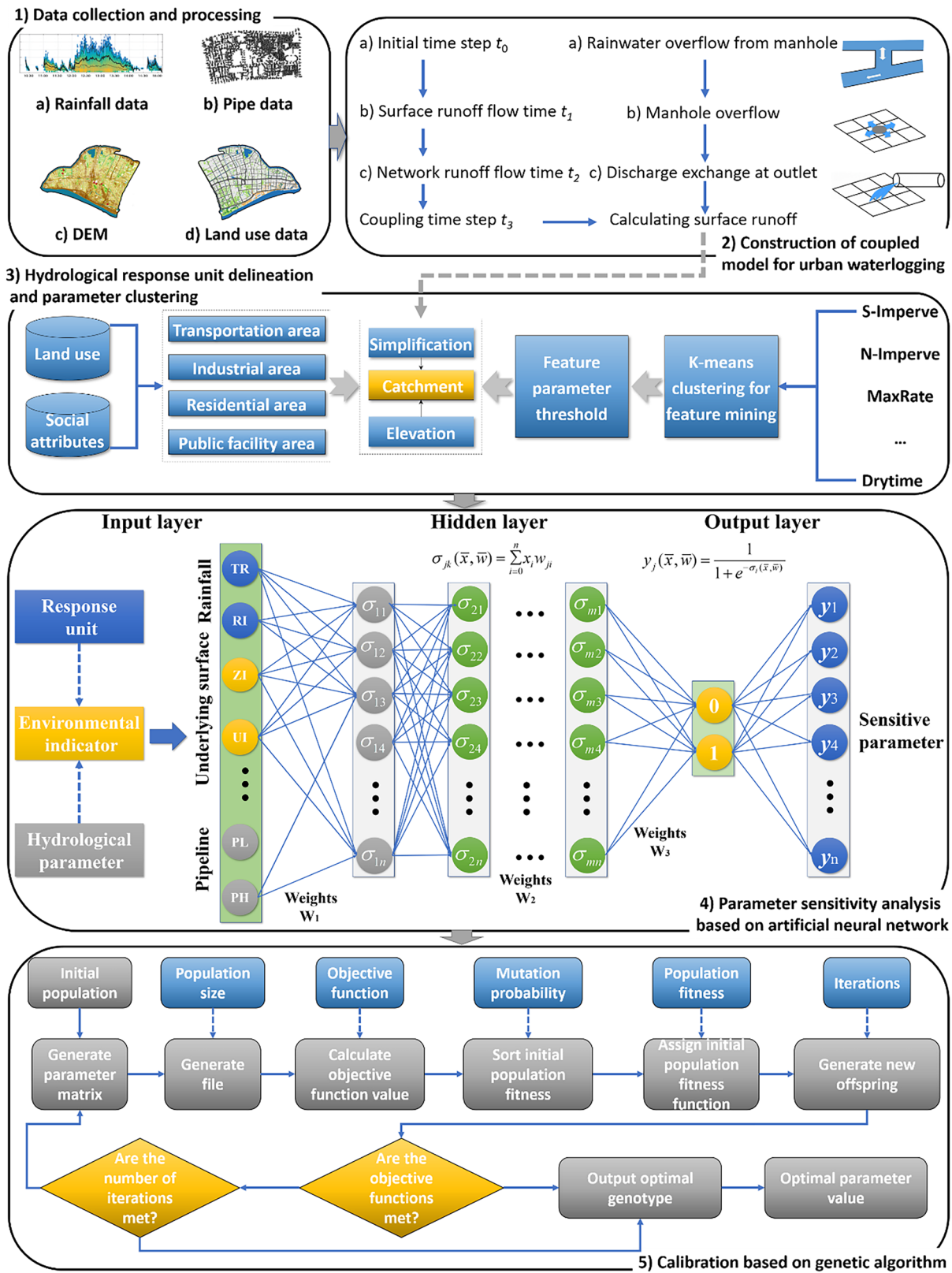


Fig. 1 Research framework

## 2.1 Urban Flood Coupling Model Construction

This study combines a one-dimensional pipe network routing model with a two-dimensional surface runoff model to construct an integrated urban waterlogging model. This integrated model aims to overcome the limitations of individual models and provide a comprehensive understanding of urban inundation processes by considering both surface and underground water routing (Yang et al. 2022).

### 2.1.1 Pipe Network Runoff Model

Surface water in urban areas flows naturally into low-lying areas and enters the stormwater pipe networks through rainwater inlets. The hydraulic characteristics of the pipe network and stormwater nodes after the rainwater entered the pipe network were calculated using the dynamic wave method. Once rainwater enters the pipe network, the flow state within the pipes continuously switches between an open-channel flow and a pressurized pipe flow (Ye et al. 2021). The Pressimann virtual slit method and unsteady Saint-Venant equations were employed to model stormwater runoff in a pipe network. The specific equations used for the calculations are as follows:

$$\frac{\partial M}{\partial t} + \frac{1}{N} \frac{\partial Q}{\partial x} = q \quad (1)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left( \alpha \frac{Q^2}{M} \right) + gM \left( \frac{\partial y}{\partial x} \right) + gMS_f - uq = 0 \quad (2)$$

where  $M$  represents the cross-sectional area of the pipe,  $N$  represents the width of the virtual slit,  $Q$  represents the flow rate in the pipe cross-section,  $u$  denotes the lateral boundary inflow velocity along the pipe,  $q$  represents the lateral boundary flow rate,  $x$  represents the distance along the pipe,  $\alpha$  is the momentum correction coefficient,  $g$  represents the acceleration due to gravity,  $y$  represents the water head position, and  $S_f$  denotes the friction slope of the pipe. To solve the aforementioned stormwater pipe network runoff model, an explicit numerical algorithm was utilized, in which the hydraulic parameters and geometric characteristics of the pipe network were obtained from the data of the stormwater pipe network model (Schilling and Tränckner 2022).

### 2.1.2 Surface Runoff Model

The surface runoff model incorporates clear physical mechanisms, enabling the calculation of a surface water depth distribution consistent with the precision of the terrain data through water exchange between grid cells. Owing to the similarity between representing urban surfaces using regular

grids and the modeling approach that employs raster data to describe surface attributes in geographic information system (GIS) (Sosa et al. 2019; O'Loughlin et al. 2020; Shustikova et al. 2020), this study adopted a grid-based hydraulic computation model to simulate the surface runoff process in urban areas.

According to the soil permeability, the underlying surface can be classified as impermeable, semi-permeable, permeable, and highly permeable surfaces. Modeling the surface runoff process using regular grid data involves employing hydraulic methods to calculate the water exchange between grid cells, simulating the movement of water under the influence of gravity and structures, and outputting water depth distribution results that are consistent with the topographic grid. The main control equation is shown as follows:

$$g \frac{n^2 u \sqrt{u^2 + v^2}}{H^{\frac{1}{3}}} + \frac{\partial J}{\partial t} + \frac{\partial(uJ)}{\partial x} + \frac{\partial(vJ)}{\partial y} + gH \frac{\partial z}{\partial x} = 0 \quad (3)$$

$$g \frac{n^2 v \sqrt{u^2 + v^2}}{H^{\frac{1}{3}}} + \frac{\partial K}{\partial t} + \frac{\partial(uK)}{\partial x} + \frac{\partial(vK)}{\partial y} + gH \frac{\partial z}{\partial y} = 0 \quad (4)$$

$$\partial J / \partial x + \partial K / \partial y + \partial H / \partial t = 0 \quad (5)$$

where  $x$  and  $y$  represent the distances in the X and Y directions in the Cartesian coordinate system, respectively.  $H$  represents the depth of the surface water,  $t$  represents time, and  $J$  and  $K$  represent the discharge per unit width in the X and Y directions, respectively.  $g$  represents the acceleration due to gravity, and  $z$  represents the water level on the surface, which is the cumulative quantity of the water depth and surface elevation.  $u$  and  $v$  represent the components of the velocity vector in the X and Y directions (vertical direction), respectively, and  $n$  represents the Manning roughness coefficient.

The implicit finite difference method was used to solve the dynamic model of surface runoff. This allows the calculation of the magnitude and direction of the flow between adjacent grid cells. Subsequently, the water depth in the grid cells was updated based on the flow in different directions.

### 2.1.3 Coupled Model of Surface and Subsurface Drainage Networks

Compared to the one-dimensional pipe network routing model, the two-dimensional surface runoff model excels in simulating flow in uncertain directions. However, it lacks consideration for underground space routing (Zeng et al. 2022). Therefore, in this study, a coupled approach was employed to integrate a one-dimensional pipe network runoff model with a two-dimensional surface runoff model. The pipe network provides net overflow rates at the nodes,

whereas the surface runoff model simulates the extent, depth, and inundation process. By coupling these models, it is possible to leverage their respective strengths to simulate urban flooding processes effectively. The coupling process consists of the following steps.

- Step 1 Construct a pipe network runoff model for the study area with minimal or no generalization of susceptible flooding nodes.
- Step 2 Run the one-dimensional pipe network runoff model to extract the overflow process at the overflow nodes of the pipelines.
- Step 3 Calculate the overflow flow rates for the overflowing pipelines.
- Step 4 Use the overflow processes and overflow rates as point-source boundary conditions to drive the two-dimensional surface runoff model.
- Step 5 Input the elevation grid data and configuration files such as partial node reflux sequences into the two-dimensional surface runoff model to compute the extent and depth of surface inundation.

## 2.2 Identification and Optimization of Sensitive Parameters in the Coupled Urban Flood Model

The parameters of urban flood models can be classified into two categories: deterministic and uncertain parameters. Deterministic parameters are obtained through field measurements or software analyses, generally refers to Area, Width, Imperv, Slope, Pipe shape, Pipe length, and Node elevation (Liu et al. 2023). Uncertain parameters are derived from expert experience or parameter calibrations. The cumbersome and complex calibration optimization process of uncertain parameters is a core issue that affects the efficiency of model simulations and predictions.

The sensitivity of uncertain parameters in urban flood models is widely acknowledged to be closely related to underlying surface conditions. For example, surface depression storage reflects the depth of water storage in hydrological response units, the surface Manning's coefficient represents the resistance encountered by precipitation during runoff in hydrological response units, the pipe network roughness coefficient determines the velocity of underground runoff, while infiltration rate and decay coefficient reflect the subsurface infiltration capacity in the Horton overland flow model (Padiyedath Gopalan et al. 2019; Hu et al. 2020). Thus, based on this understanding, this study proposed a rule for dividing urban hydrological response

units, which incorporates the characteristics of underlying surface attributes. Subsequently, the differential parameter thresholds obtained through K-means clustering analysis were assigned to each watershed unit. Finally, a genetic algorithm was used to calibrate the parameters of the sub-watershed units in different urban land use functional zones within the coupled urban flood model. The main steps are as follows:

- Step 1 Division of hydrological response units considering urban surface characteristics. This ensures that each sub-catchment corresponds to an independent urban land use functional zone.
- Step 2 Outputting feature parameter values for different urban land use functional zones using the K-means clustering algorithm. A crowdsourced dataset of uncertain parameters in urban flood models was established. First, the relevant literature and historical experiential data are searched to obtain a set of prior sample parameters with explicit values for nine uncertain parameters in urban flood models: S-Imperv, S-perv, N-Imperv, N-perv, MaxRate, MinRate, Decay, Drytime, and Roughness. Then, the prior sample parameters are input into the K-means clustering model for analysis, resulting in the output of the feature parameter values assigned to different urban land use functional zones.
- Step 3 Proposal of a sensitive parameter identification mechanism based on an ANN model. This involves using environmental indicators that affect the sensitivity of parameters in different hydrological response units as inputs and constructing a binary classification ANN model with sensitive parameters in the urban flood model as outputs. By adjusting the number of hidden layers and the maximum iteration count of the neural network model, the sensitive parameters of the urban flood model can be quickly obtained.
- Step 4 Calibration of sensitive parameters using a genetic algorithm. Based on the K-means clustering analysis and the sensitive parameter identification mechanism using the ANN model, the thresholds of the sensitive parameters in the urban flood model are distributed to each sub-watershed unit according to the distribution pattern of the urban land use functional zones. Multiple rainfall-runoff events are selected, and based on the mapping relationship between two-dimensional surface runoff generation and one-dimensional pipe network runoff

mechanisms, the optimal parameter values are determined using genetic algorithm.

### 2.2.1 Hydrological Response Unit Division Considering Subsurface Characteristics

Hydrological response units serve as both spatial discretization units and modeling entities for urban flood models. Due to human activities, the natural drainage network within urban areas undergoes continuous fragmentation and merging, transforming from a traditional dendritic structure to a complex network structure (Shen et al. 2019; Zhang et al. 2022). Currently, most hydrological response unit delineation methods in flood simulations are time-consuming and do not consider the influence of topography and human infrastructure. In this study, we not only relied on empirical partitioning methods but also imposed constraints on the number and fundamental spatial scales of hydrological response units. Therefore, this study proposed a set of rules for urban hydrological response unit delineation that incorporates both social and natural characteristics of land surfaces. The specific steps are as follows.

- (1) Based on the natural features of the terrain, conduct preliminary analysis of the hydrological characteristics of the digital elevation model (DEM) using ArcMap software, and delineate primary watershed units based on the flow direction of the pipelines.
- (2) The distribution of major drainage pipelines and roads is used as a basic principle for further refinement and division of watershed areas. At the same time, control the number of sub-watersheds to ensure a balance between the number of pipe segments and nodes with the number of sub-watershed areas (Sun et al. 2022).
- (3) Since each watershed unit possesses both natural and social attributes, overlay these two types of attributes to perform urban functional zoning.

The specific delineation principles are as follows.

- (1) Class I sub-catchments correspond to transportation areas where land use is predominantly composed of paved surfaces. The surface is relatively flat and compact, with minimum values for ponding storage capacity, Manning's coefficient, infiltration rate, and attenuation coefficient.
- (2) Class II sub-catchments correspond to commercial and industrial areas, where land use consists primarily of buildings and paved surfaces. The surface is relatively flat, with relatively small values for ponding storage capacity, Manning's coefficient, infiltration rate, and attenuation coefficient.

- (3) Class III sub-catchments correspond to relatively dispersed residential areas characterized by a mixture of paved roads, roofs, and limited green spaces. Compared to commercial and industrial areas, residential areas have greater surface roughness and variability as well as slightly better permeability. The values for ponding storage capacity, Manning's coefficient, infiltration rate, and attenuation coefficient are moderate.
- (4) Class IV sub-catchments correspond to public land areas dominated by gardens and green spaces. The land use consists primarily of grasslands and forests, with the highest surface roughness, good permeability, and water storage capacity. The values for ponding storage capacity, Manning's coefficient, infiltration rate, and attenuation coefficient are the highest.

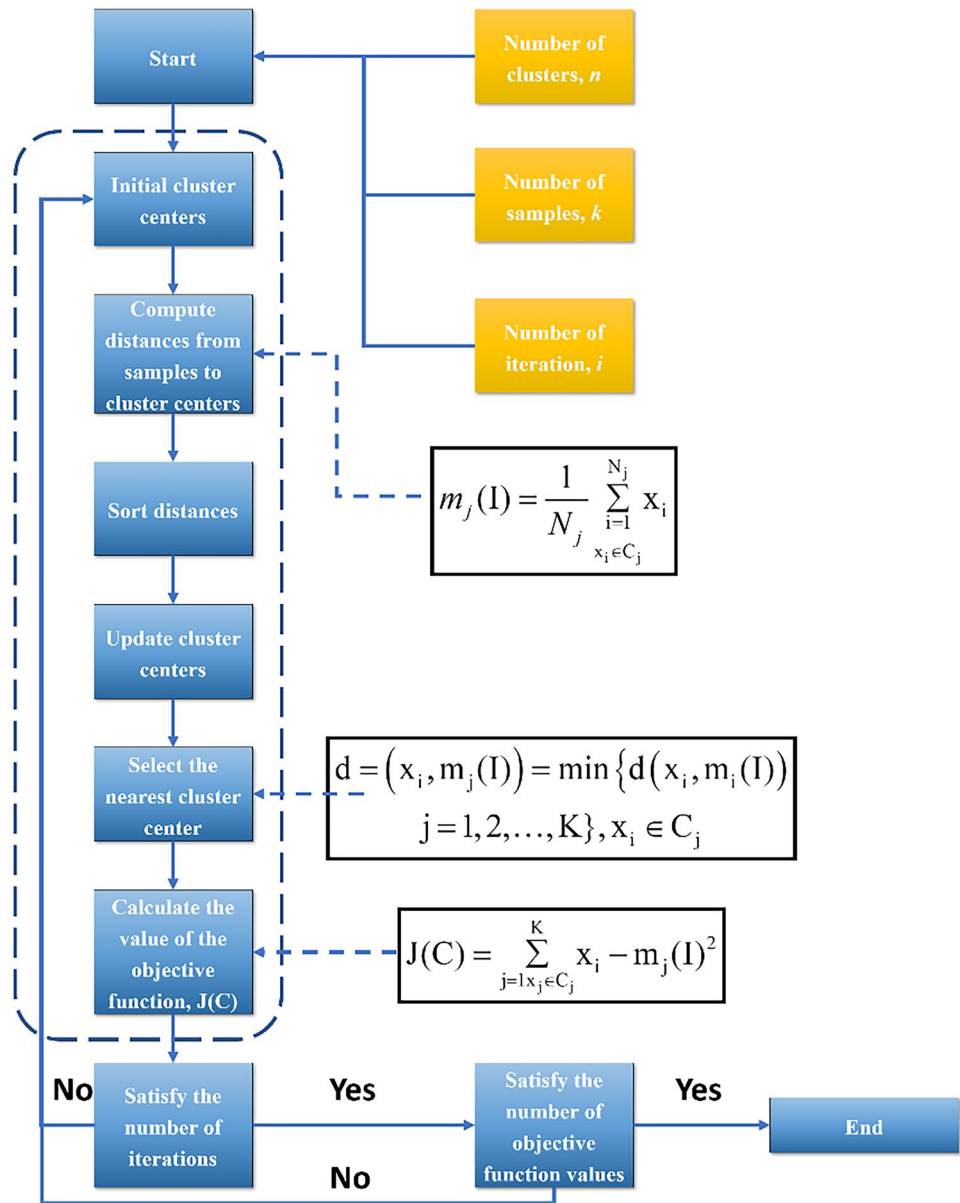
### 2.2.2 Parameter Clustering Based on K-Means

The K-means clustering algorithm is an iterative machine learning analysis algorithm that partitions the sample data into multiple distinct clusters based on the similarity of their feature characteristics. It accomplishes this by randomly selecting  $k$  initial sample points as the initial cluster centers (Liu et al. 2015). In this study, we searched the relevant literature and historical experiential data to obtain a prior sample parameter set that includes nine uncertainty parameters: S-Imperv, S-perv, N-Imperv, N-perv, MaxRate, MinRate, decay, dry time, and roughness. Subsequently, the K-means clustering algorithm was applied to cluster the uncertainty parameters of different functional land use areas in an urban context. The number of clusters  $k$  was set, and the prior parameter samples were inputted into the K-means model for analysis. The algorithm outputs characteristic parameter values (that is, cluster centroids) under  $k$  different clustering conditions. These values were assigned to transportation, commercial and industrial, residential, and public facility areas, representing four distinct land use functional areas in the urban context. A flowchart of the clustering algorithm is shown in Fig. 2.

### 2.2.3 Sensitive Parameter Identification Based on Artificial Neural Network (ANN) Model

Artificial neural networks are powerful tools for processing deep-learning algorithms and are widely applied in regression and classification tasks. In general, ANN can fit any nonlinear function with a well-designed network structure, making them suitable for handling nonlinear systems or black-box models with complex internal representations. However, ANN models require substantial data training to achieve stability, and the training process can be time-consuming. Once trained, though, the model could be rapidly applied to new datasets.

**Fig. 2** Flowchart of clustering analysis based on K-means algorithm



To expedite the identification of sensitive parameters in urban flood models using ANN, the input and output layer data must be prepared. Because this study focused on a coupled model of surface and pipe networks, the determinants of parameter sensitivity were attributed to rainfall, underlying surface, and pipe networks. A quantitative representation of these three factors was achieved by utilizing 11 measurable representative environmental indicators (Table 1), forming part of the ANN model input. The output-layer data (sensitive parameters) were prepared using the Morris method. Specifically, within the parameter range obtained from K-means clustering, each parameter  $x_i$  was randomly altered, and the model was run to generate different outputs corresponding to the

varied  $x_i$  values. The sensitivity  $e_i$  of parameter  $i$  can be expressed as:

$$e_i = \frac{y^* - y}{x_i - x} \quad (6)$$

where  $x$  represents the values of the model parameters, and the corresponding output is  $y$ . When this parameter is changed to  $x_i$ , the corresponding output is  $y^*$ .

The sensitivity identification of parameter  $i$  through K-means clustering is treated as a binary classification problem, where the output is 1 if the parameter is sensitive and 0 otherwise. The performance and prediction accuracy of a neural network are influenced by hyperparameters. By

**Table 1** Environmental indicators affecting parameter sensitivity

| Influencing factor | Environmental indicator                     | Unit      |
|--------------------|---|-----------|
| Rainfall           | Total rainfall                              | mm        |
|                    | Mean rainfall intensity                     | mm/10 min |
| Land surface       | Hydrological response unit area             | ha        |
|                    | Hydrological response unit slope            | %         |
|                    | Proportion of impervious surface            | %         |
|                    | Proportion of completely impervious surface | %         |
| Pipe               | Pipe burial depth                           | m         |
|                    | Pipe length                                 | m         |
|                    | Pipe slope                                  | %         |
|                    | Pipe segment type                           | –         |
|                    | Pipe segment radius                         | m         |

adjusting the hyperparameters of the ANN model, a sensitive parameter identification model with higher accuracy can be obtained. The structure of the model is illustrated in Fig. 3.

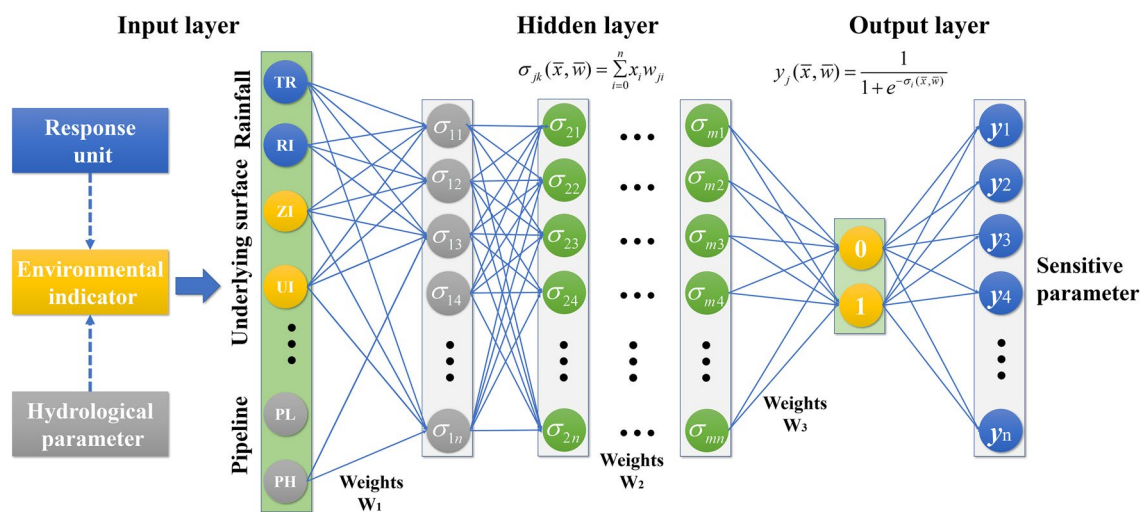
**2.2.4 Parameter Calibration Based on Genetic Algorithm**

Genetic algorithms are primarily used for optimizing problem-solving in machine learning. The basic framework of this algorithm consists of four components: solution vector encoding, solution vector population, fitness function evaluation, and genetic operations (Song et al. 2009). In this study, we adopted the clustering results (that is, specific parameter values) in conjunction with the sensitive parameter clustering results. This approach aims to prevent the singularization of sensitive parameters in each sub-catchment area and enhance the adaptability of sub-catchment areas in different urban land use functional zones. The clustering results are used as thresholds for calibrating the parameters of the urban flood model in sub-catchment areas. The threshold values of

the sensitive parameter clustering features are distributed to individual catchment units based on the distribution patterns of the urban land use functional zones. A genetic algorithm is utilized to determine the optimal values based on the mapping relationship between two-dimensional surface runoff generation and one-dimensional pipe network runoff mechanisms (to reduce uncertainty in the simulation process, the calibration results of this method can only serve as a reference and still require further refinement in conjunction with manual calibration by experts). The specific procedure of the algorithm can be found in Fig. 1, Step 5.

**2.3 Evaluation Metrics**

In this study, the surface water depth at monitoring stations was evaluated using the Nash-Sutcliffe efficiency coefficient (*NSE*), root mean square error (*RMSE*), and peak time difference (*PTD*) (Ichiba et al. 2018). The calculation methods for these evaluations are:



**Fig. 3** Artificial neural networks (ANN) model structure



$$NSE = 1 - \frac{\sum_{t=1}^T (D_{obs}^t - D_{sim}^t)^2}{\sum_{t=1}^T (D_{obs}^t - \overline{D_{obs}})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (D_{obs}^t - D_{sim}^t)^2} \quad (8)$$

$$PTD = |t_{sim}^i - t_{obs}^i| \quad (9)$$

where  $D_{obs}$  and  $D_{sim}$  represent the observed and simulated water depth at time  $t$ , respectively.  $NSE$  serves as an important indicator for evaluating the quality of the model simulation results. The closer the value of  $NSE$  is to 1, the more plausible it is that the model simulates the evolution of the inundation.  $t_{sim}^i$  and  $t_{obs}^i$  represent the occurrence times of the simulated and observed flood peaks, respectively, and this study selects whole-hour intervals.

### 3 The Study Area

This section first presents an overview of the research area and its scientific rationale as a typical case study. It then elaborates on the sources of rainfall, water depth monitoring, pipeline, and geographical data.

#### 3.1 Overview of the Study Area

The study area is located in the western part of Nanjing City (Fig. 4). The total land area is approximately 28  $\text{km}^2$ , with surface elevations ranging from 0 to 19.5 m. The drainage pipe network within the area is relatively independent and consists of various artificial channels and lakes as storage units. Therefore, this area is an urban watershed with clear boundary conditions. Furthermore, the study area has a high level of urbanization, resulting in high surface temperatures during summer. Automobile exhaust emissions and air-conditioning heating contribute to the high temperature of the lower atmospheric layer, whereas densely built structures impede atmospheric circulation, resulting in a relatively stable air mass in the region. These conditions create favorable circumstances for the occurrence of heavy rainfall and urban flooding disasters (Zhu et al. 2018).

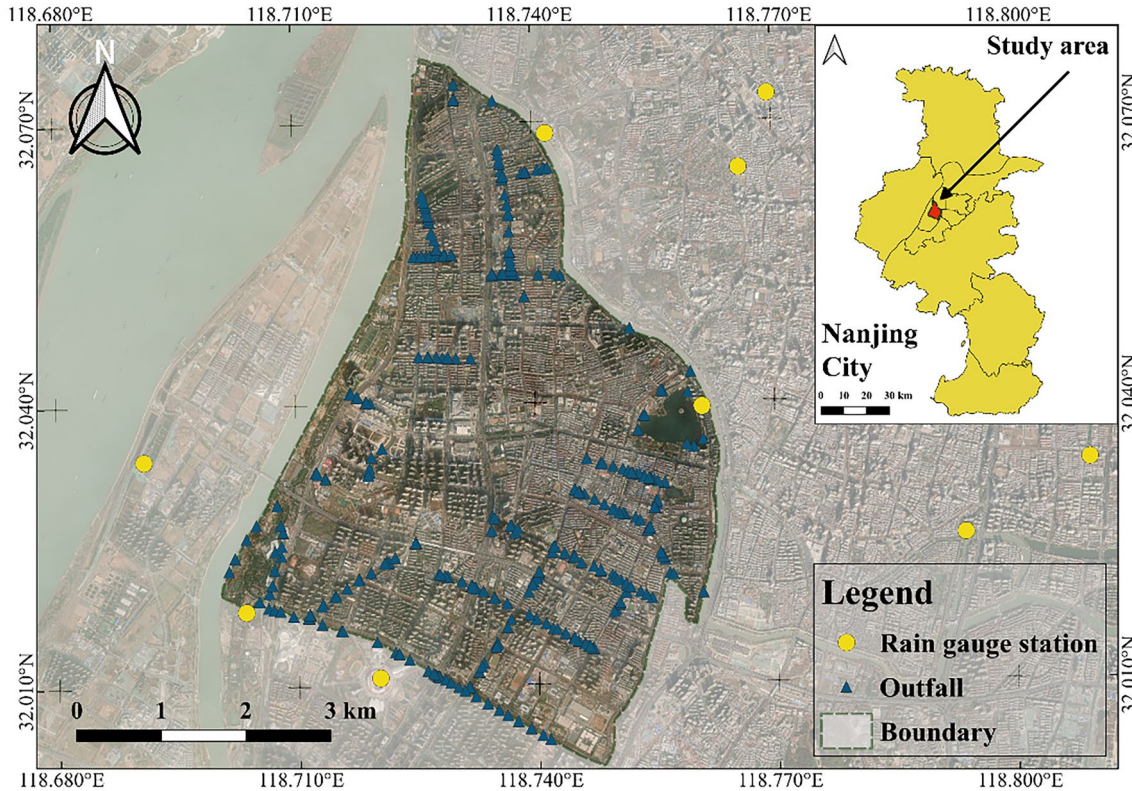


Fig. 4 Overview of the study area

### 3.2 Data Sources

This study collected hourly precipitation data from 1 January 2016, to 14 August 2019. Two rainfall stations are located within the study area in Nanjing City, along with three rainfall stations located on the outskirts. The data from these rainfall stations are useful for determining the timing of heavy rainfall events and optimizing parameters in the urban stormwater flooding model. Surface water depth monitoring data were also collected from six flood-prone sites in the study area. The monitoring information included the time of observation and water depth. To facilitate the modeling of the urban surface in the study area, geographical data, such as DEM, building data, water system distribution, and land use types, were obtained from the Nanjing Planning and Natural Resources Bureau. All the data were based on the WGS-84 coordinate system, and the UTM zone 50N projection was uniformly adopted for the map projection when using the planar coordinate system. In addition, the required pipeline network data for this experimental area was provided by the Nanjing Survey and Design Research Institute Co., Ltd., which include rainwater sewers, inspection wells (rainwater inspection wells and sewage inspection wells), pipelines (rainwater pipelines, sewage pipelines, and a small number of combined rain and sewage pipelines), and drainage outlets (rainwater drainage outlets and sewage drainage outlets).

## 4 Results and Discussion

The experiment first divided the study area into different urban land use functional zones based on the hydrological response unit partitioning rules. Subsequently, a parameter sensitivity analysis and optimization were conducted using the K-means-ANN-GA machine learning method. Finally, the constructed urban inundation model was tested through simulations of three observed rainfall events, and a comparative analysis was performed to discuss the validity of the methodology.

### 4.1 Hydrological Response Unit Delineation

According to the partition rules described in Sect. 2.2.1, the study area was divided into distinct urban land use functional zones, and specific hydrological response units were delineated for each zone.

Land use types provide insights into the natural characteristics of urban surfaces. In this study, land use type and geographical data were obtained from the Nanjing Planning and Natural Resources Bureau. ArcGIS software, in combination with manual identification, was used to classify the land surface of the study area (Fig. 5).

Land use planning reflects the social attributes of urban surfaces. Based on the overall urban land use planning of Nanjing City, the study area was divided into several zones, including residential, school, administrative office, commercial, financial and industrial areas, public green spaces, and water bodies.

To delineate urban land use functional zones, we combined the natural and social attributes of the land surface following the classification of land use and land use planning in the study area. This approach aimed to avoid the issue of having too few categories to differentiate the hydrological response units or too many categories to determine the threshold values for each subcategory. Therefore, the urban land use in the research area was divided into four categories: transportation areas (TA), commercial and industrial areas (CA), residential areas (RA), and public facility areas (PA). The urban functional zoning of the study area is shown in Fig. 6.

Based on the current status and planning conditions of the drainage network in the area, this study modified and generalized the drainage network based on spatial topological relationships (with minimal or no generalization at points vulnerable to waterlogging). Watershed units were delineated based on the principles of hydrological response unit division and hydrodynamic knowledge.

### 4.2 Sensitivity Analysis and Optimization of Parameters in the Study Area

Using parameter values obtained from relevant literature as samples, the K-means clustering algorithm was employed to calculate the parameter thresholds for different urban land use functional areas. The number of clusters  $K$  was set to four, and the resulting parameter thresholds for different urban land use functional areas are presented in Table 2.

Table 2 presents the uncertain parameters required for the model. As multiple parameters have similar effects, selecting a few sensitive parameters for calibration is sufficient. Therefore, based on the method of identifying sensitive parameters described in Sect. 2.2.3, an artificial neural network was employed to rapidly identify the sensitive parameters of the coupled urban waterlogging model. The neural network used environmental indicators as inputs and sensitive parameters as outputs.

This is considered a binary classification problem, where a parameter is assigned a value of 1 if it is sensitive and 0 otherwise. The data were divided into training and testing sets, where TP, TN, FP, and FN represent the numbers of true positives, true negatives, false positives, and false negatives, respectively. These values are presented as percentages in the confusion matrix, as shown in Fig. 7. It can be observed from the figure that, in the testing set, the accuracy of identifying the sensitivity of the nine parameters is

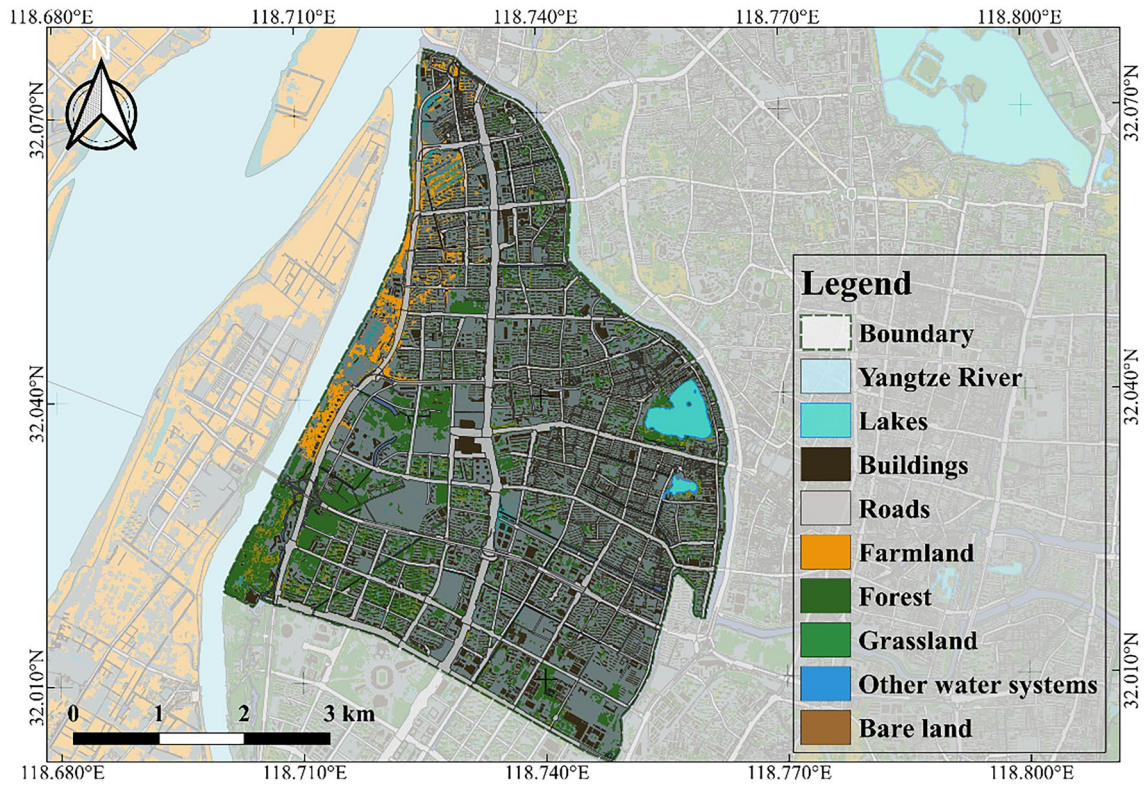


Fig. 5 Land use types of the study area

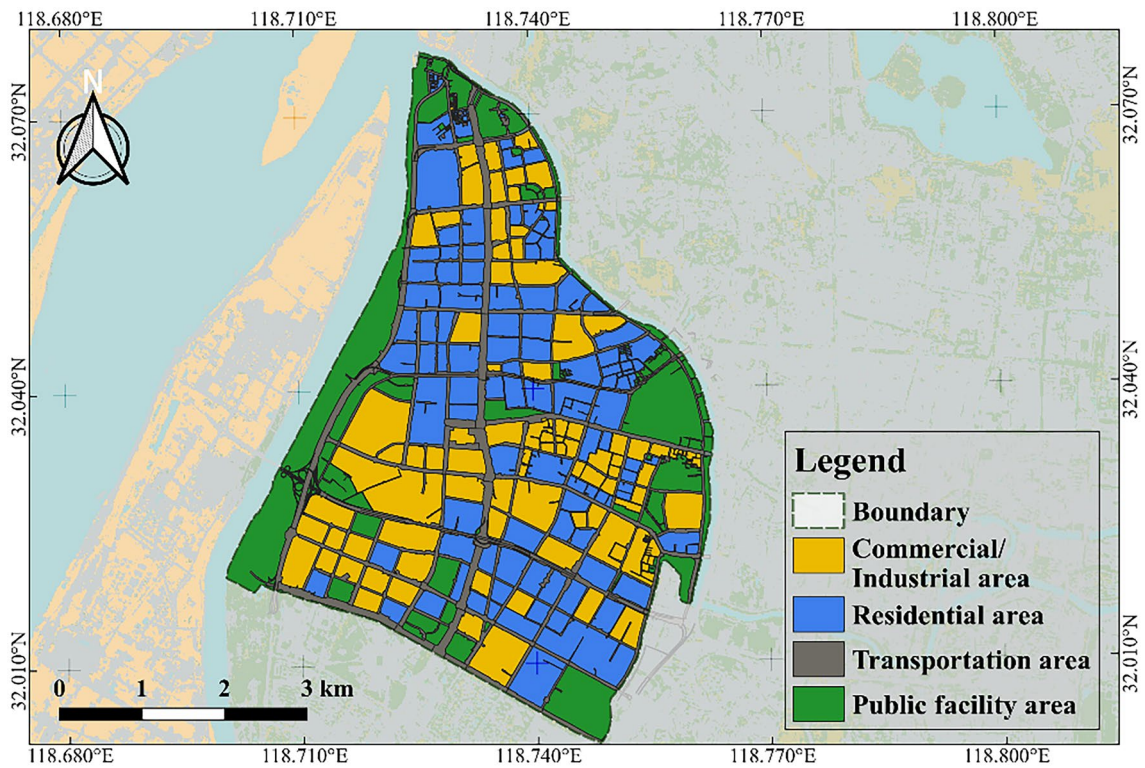
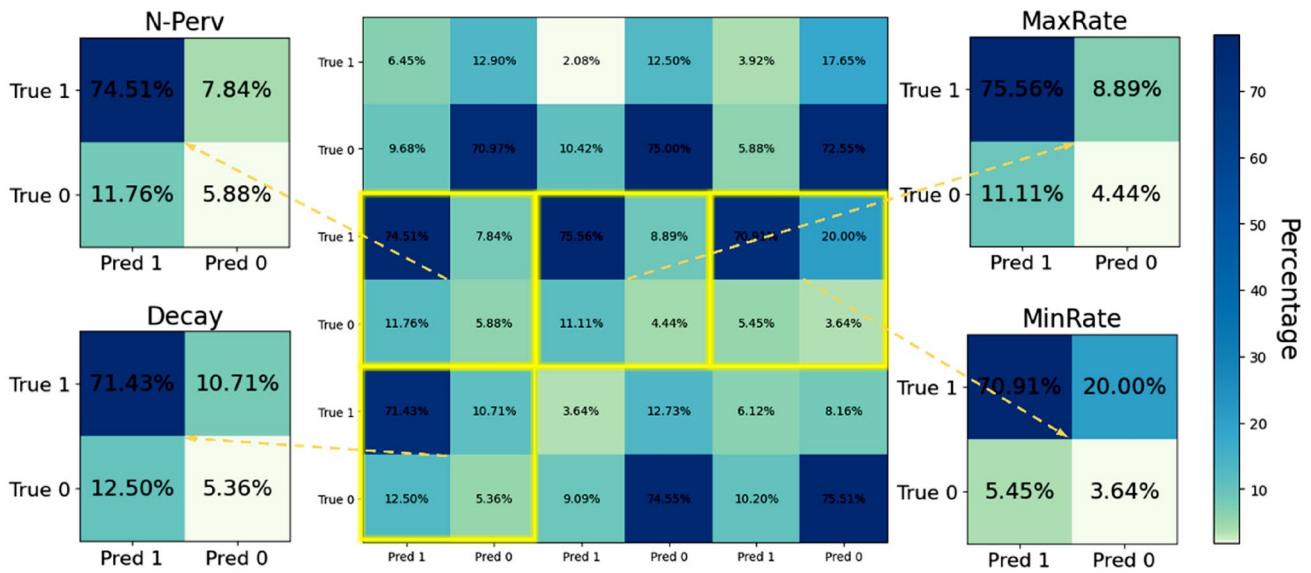


Fig. 6 Urban functional zones of the study area

**Table 2** Uncertainty parameter thresholds

| Parameter                      | S-Imperv (mm) | S-perv (mm) | N-Imperv –  | N-perv –  | MaxRate (mm h <sup>-1</sup> ) | MinRate (mm h <sup>-1</sup> ) | Decay – | Drytime (d) |
|--------------------------------|---------------|-------------|-------------|-----------|-------------------------------|-------------------------------|---------|-------------|
| Transportation (TA)            | 0.2–3.2       | 2.7–5.9     | 0.01–0.018  | 0.05–0.18 | 31.6–40.5                     | 1.3–5.5                       | 2.1–3.2 | 2.9–4.7     |
| Commercial and industrial (CA) | 3.2–5.7       | 5.9–8.2     | 0.018–0.024 | 0.18–0.25 | 40.5–60.2                     | 5.5–9.1                       | 3.2–4.3 | 4.7–8.2     |
| Residential (RA)               | 5.7–7.9       | 8.2–10.4    | 0.024–0.039 | 0.25–0.31 | 60.2–77.8                     | 9.1–15.7                      | 4.3–5.6 | 8.2–10.1    |
| Public facility (PA)           | 7.9–9.2       | 10.4–12.9   | 0.039–0.042 | 0.31–0.39 | 77.8–94.1                     | 15.7–19.3                     | 5.6–6.7 | 10.1–13.5   |



**Fig. 7** Parameter sensitivity identification accuracy confusion matrix

generally above 70%. This indicates the applicability of the method for identifying sensitive parameters using artificial neural networks.

According to the sensitivity analysis results, only N-Perv, MinRate, Decay, and MaxRate were identified as sensitive parameters in a certain region. Using Table 2 as a reference, we assigned range values to these four sensitive parameters for the sub-watersheds within the study area according to the distribution patterns of the urban land use functional areas. A genetic algorithm was then employed in conjunction with the observed urban stormwater flooding events to determine the optimal values of these sensitive parameters. In a certain hydrological response unit within the study area, the optimal values of the four sensitive parameters for different functional areas of urban land use are presented in Table 3.

### 4.3 Model Verification

Referring to the parameter table in Sect. 4.2, we substituted the uncertain parameter values into the model to verify its performance across different sub-watershed categories. Three observed rainfall events were used to simulate

**Table 3** Uncertainty parameters

| Parameter                      | N-perv | MinRate               | MaxRate               | Decay |
|--------------------------------|--------|-----------------------|-----------------------|-------|
| Unit                           | –      | (mm h <sup>-1</sup> ) | (mm h <sup>-1</sup> ) | –     |
| Transportation (TA)            | 0.09   | 3.8                   | 33.2                  | 2.8   |
| Commercial and industrial (CA) | 0.19   | 7.2                   | 46.9                  | 3.9   |
| Residential (RA)               | 0.27   | 12.4                  | 67.7                  | 5.0   |
| Public facility (PA)           | 0.33   | 17.5                  | 85.1                  | 6.5   |

and test the developed urban flooding model. To validate the feasibility of the proposed method, we employed two additional methods for comparison: the ANN-GA method, which does not consider the delineation rules of urban hydrological response units, and the K-means-deep neural networks (K-means-DNN) method, which does not consider the genetic algorithm. The evaluation metrics selected for assessing the simulated results at monitoring stations S1 and S2 were the Nash-Sutcliffe efficiency coefficient (*NSE*), root mean square error (*RMSE*), and peak time difference (*PTD*,  $\Delta t$ , to an integer). Statistical

analyses of the evaluation metrics for the three methods are presented in Table 4.

From the evaluation metrics, the proposed method in this study achieved a Nash-Sutcliffe efficiency coefficient (*NSE*) above 0.73, a root mean square error (*RMSE*) within the range of 2–7, and an average peak time difference (*PTD*,  $\Delta t$ ) of approximately 20 min. Compared to the ANN-GA method, the proposed method shows improvements of 0.29 in *NSE*, 3.74 in *RMSE*, and 0.5 h in *PTD*. Compared with the K-means-DNN method, the proposed method shows improvements of 0.19 in *NSE*, 2.76 in *RMSE*, and 0.17 hours in *PTD*. Therefore, it can be concluded that the proposed method performs better at simulating urban flooding by effectively capturing the distribution patterns of uncertain parameters in different urban functional areas and aligning the model with the actual underlying surface conditions.

Scatter plots of the simulated results and observed values for the two monitoring stations during the flood events on 10 June 2017, and 8 August 2017 are shown in Fig. 8. These scatter plots demonstrate a close alignment between the simulated results based on the optimized parameters in this study and the observed values, following a 1:1 relationship without significant nonlinearity or heteroscedasticity patterns. In addition, we fitted the simulated results of the other two methods to the observed values. Although the fitting line of the K-means-DNN method approaches a 1:1 relationship in some cases (for example, the  $S_1$  monitoring station during the 10 June 2017 event), these fitting lines do not pass the significance test when considering all the events in the testing dataset.

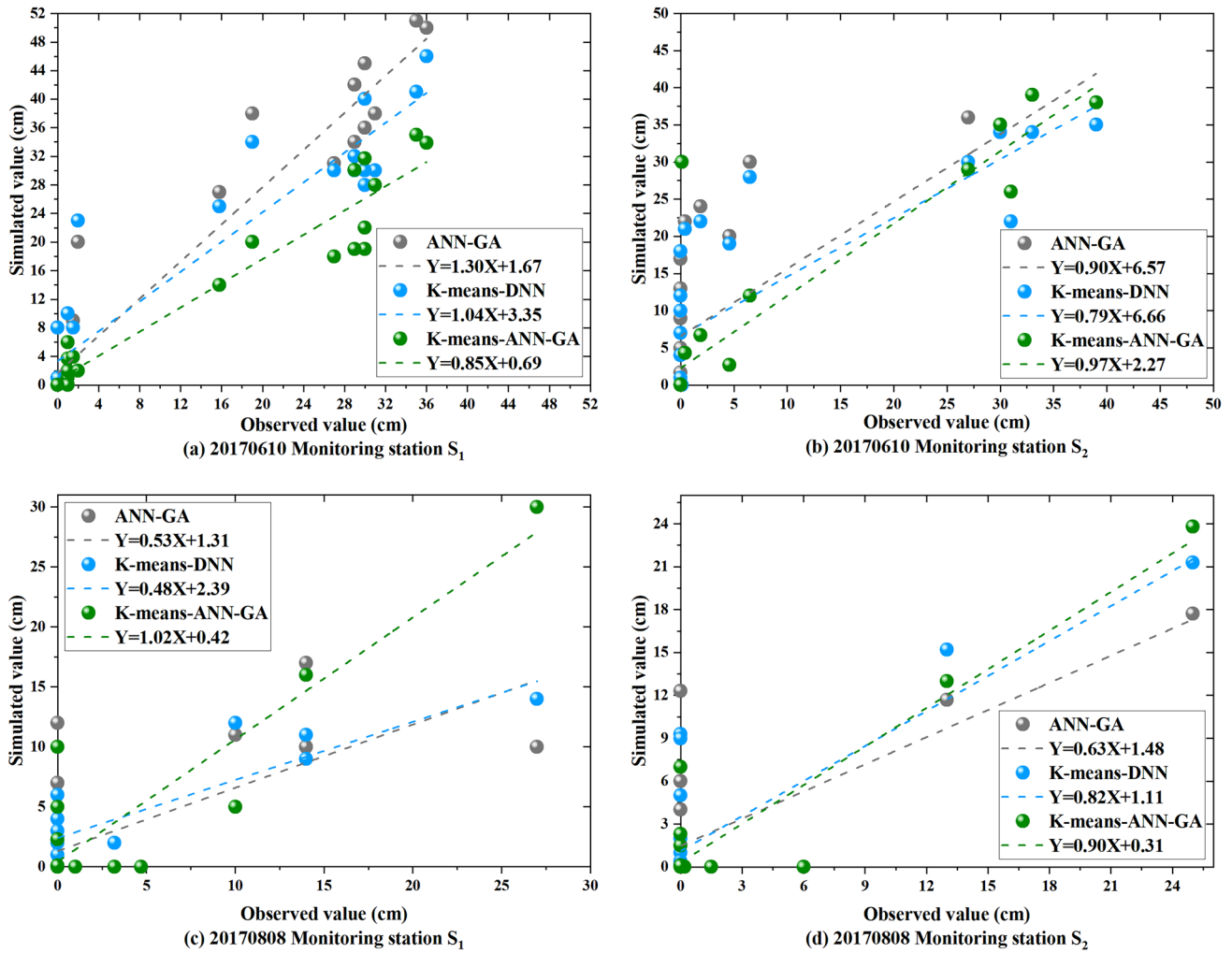
The simulation process of two monitoring stations during the flood events on 10 June 2017 and 8 August 2017 is shown in Fig. 9 (the curve has been smoothed). From the evolution of flood simulation in the study area, it is evident that the K-means-ANN-GA parameter optimization method demonstrates a closer fit between the simulated water depths and observed water depths, as well as a closer match between the simulated and observed peak values, compared to the parameter optimization methods of ANN-GA and K-means-DNN. The ANN-GA method, which neglects the consideration of urban functional areas, exhibits a pronounced attenuation

and significant errors in the computed flood hydrographs. The results further demonstrate that exploring parameter patterns during the modeling process can better reflect urban surface features (Sun et al. 2022). The variation in uncertainty parameter thresholds is also linked to urban functional zones, providing new insights into rapidly obtaining parameters for urban flooding models, consistent with Liu et al.'s research findings (Liu et al. 2023). On the other hand, the K-means-DNN method, which considers the characteristics of urban underlying surfaces based on the division rules of urban hydrological response units, does not employ genetic algorithms for precise threshold determination of sensitive parameters but assigns a single fixed parameter value to each sub-catchment unit. As a result, it fails to adequately account for parameter uncertainty and exhibits poor fitting to the observed values. The results confirm the necessity of conducting parameter precision calibration (Song et al. 2009; Li 2020) and address the challenge of accurately simulating urban flooding depths in the presence of continuously changing urban surfaces, which is difficult to achieve with fixed-parameter models (Kim et al. 2022). The K-means-ANN-GA method proposed in this study effectively captured the complex underlying surface characteristics of the study area. It not only better explores the sensitive parameters of the urban flooding model, but also considers parameter uncertainty. Therefore, in the model simulation process, the parameter optimization method used in this study can more accurately portray the production and convergence processes, which is consistent with the above analyzed results.

Further analysis of the flooding situation during the two flood events on 10 June 2017 and 8 August 2017 in the study area was conducted, and the distribution of the flooded areas at a certain moment is shown in Fig. 10. When combined with the schematic diagram of urban land use functional zones in Fig. 6, it becomes evident that simulated waterlogging is prone to occur in urban blocks, especially in areas with a high concentration of industrial and commercial land use. This is due to the high building density, extensive surface hardening, strong impermeability, low surface elevation, and rapid surface runoff in these areas, making them highly susceptible to waterlogging. In contrast, simulated

**Table 4** Statistical analyses of flood simulation result evaluation indicators

| Method            | Evaluation indicators | ANN-GA     |             |                | K-Means-DNN |             |                | K-Means-ANN-GA |             |                |
|-------------------|-----------------------|------------|-------------|----------------|-------------|-------------|----------------|----------------|-------------|----------------|
|                   |                       | <i>NSE</i> | <i>RMSE</i> | $\Delta t$ (h) | <i>NSE</i>  | <i>RMSE</i> | $\Delta t$ (h) | <i>NSE</i>     | <i>RMSE</i> | $\Delta t$ (h) |
| 10 June 2017      | $S_1$                 | 0.61       | 8.86        | 1              | 0.76        | 7.02        | 0              | 0.91           | 4.22        | 1              |
|                   | $S_2$                 | 0.40       | 10.22       | 1              | 0.44        | 9.83        | 0              | 0.74           | 6.76        | 1              |
| 8 August 2017     | $S_1$                 | 0.47       | 4.79        | 1              | 0.56        | 4.34        | 0              | 0.80           | 2.95        | 0              |
|                   | $S_2$                 | 0.52       | 4.44        | 0              | 0.68        | 3.60        | 0              | 0.87           | 2.32        | 0              |
| 25 September 2017 | $S_1$                 | 0.60       | 7.82        | 1              | 0.64        | 7.02        | 1              | 0.73           | 2.18        | 0              |
|                   | $S_2$                 | 0.51       | 8.51        | 1              | 0.65        | 6.94        | 2              | 0.79           | 3.79        | 0              |
| Average value     |                       | 0.52       | 7.44        | 0.83           | 0.62        | 6.46        | 0.5            | 0.81           | 3.70        | 0.33           |

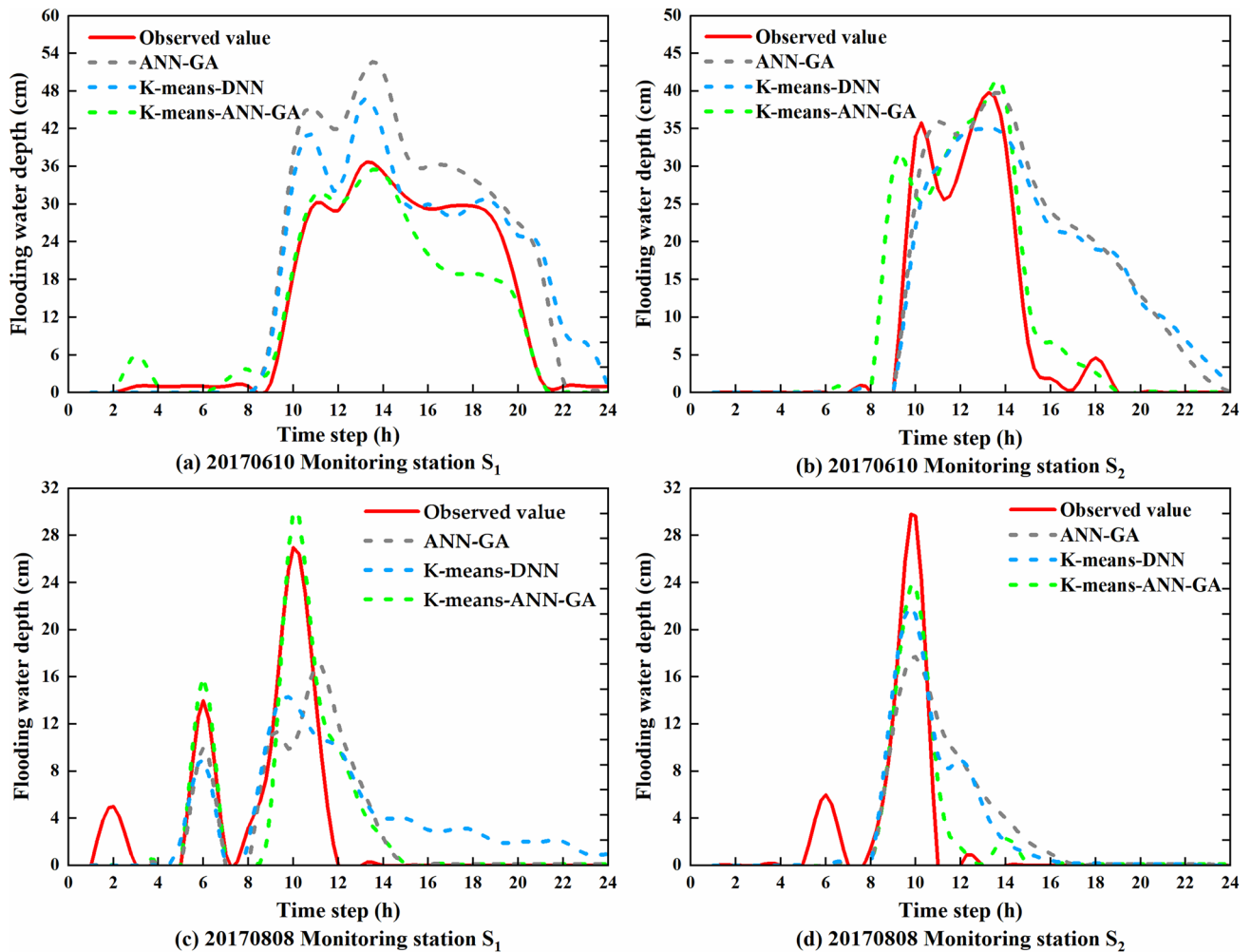


**Fig. 8** Scatter plots of the simulated versus observed values. **a** Monitoring station  $S_1$  on 10 June 2017; **b** Monitoring station  $S_2$  on 10 June 2017; **c** Monitoring station  $S_1$  on 8 August 2017; **d** Monitoring station  $S_2$  on 8 August 2017.

waterlogging is less likely to occur in the urban outskirts because of the abundance of public land, extensive coverage of green spaces, sparse buildings, lush vegetation, weak impermeability, and low susceptibility to waterlogging. The analysis results are consistent with the conclusions of Chen et al. (2022), Peng et al. (2021), Liao et al. (2023), among others. These results further demonstrate the feasibility of the parameter optimization method that considers the social and natural characteristics of underlying surfaces. Moreover, compared to traditional parameter optimization methods, our parameter sensitivity analysis process reduces the tedious and complex processes such as multiple simulations of the model, effectively reducing the time required for parameter optimization (Wu et al. 2021). This significantly improves the modeling efficiency of urban rainfall-flood models and highlights the great potential of combining machine learning with physical knowledge in parameter optimization research for urban flood models (Snieder and Khan 2023).

### 5 Conclusion

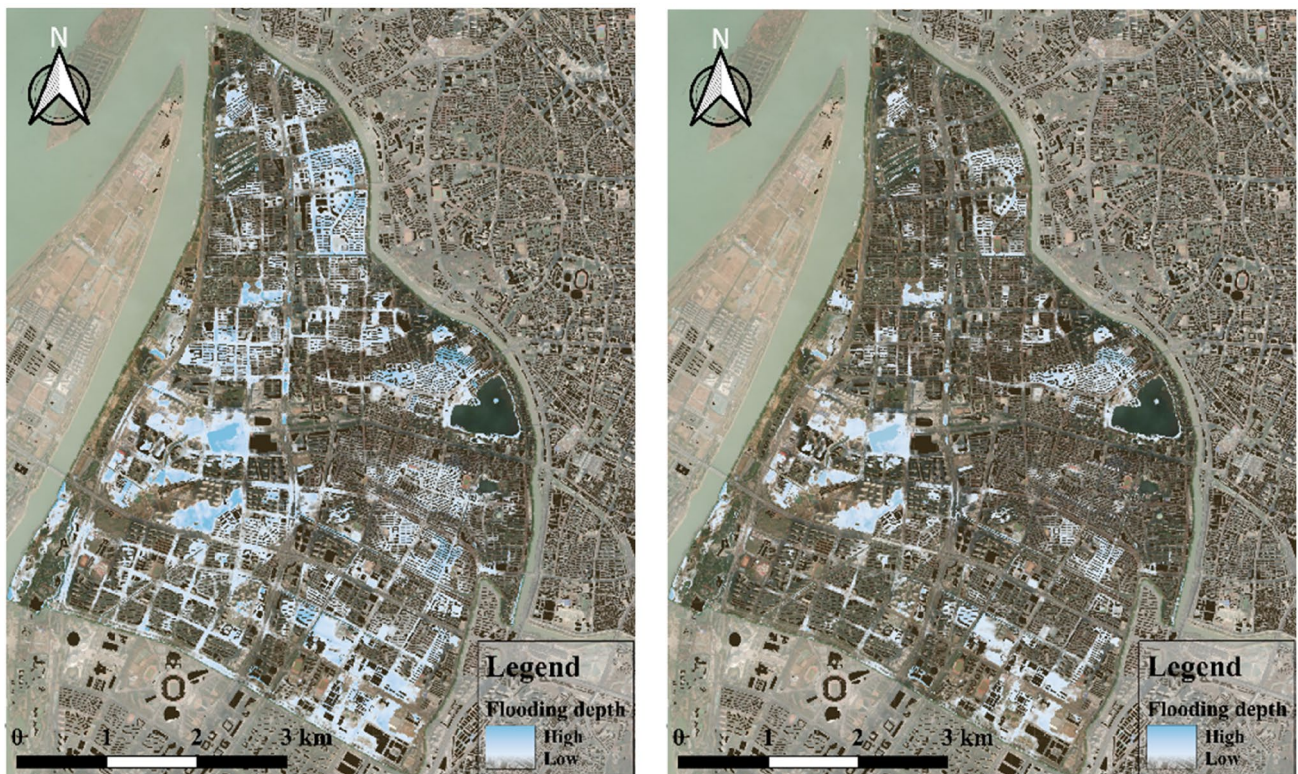
This study employed a coupled approach to integrate a one-dimensional pipe network model and a two-dimensional surface runoff model, harnessing their respective strengths to simulate urban flooding processes in a more detailed manner. To address the challenges related to unclear sub-catchment delineation and complex parameter optimization in urban flooding models, a principle for partitioning urban hydrological response units was proposed. Furthermore, the parameters of the urban flooding model were optimized using the K-means-ANN-GA method. The results indicate that the average Nash-Sutcliffe efficiency coefficient of simulated water depth in the three rainfall events reached 0.81, demonstrating a closer fit to the observed water depths compared to the ANN-GA and K-means-DNN parameter optimization methods. This study explored the general patterns of nine uncertainty



**Fig. 9** Flood process simulation results. **a** Monitoring station  $S_1$  on 10 June 2017; **b** Monitoring station  $S_2$  on 10 June 2017; **c** Monitoring station  $S_1$  on 8 August 2017; **d** Monitoring station  $S_2$  on 8 August 2017.

parameters, including S-Imperv, S-perv, and N-Imperv, and different urban land use functional zones, starting from the relationship between the uncertain model parameters and the complex underlying surface structure of the urban areas. This approach offers a new perspective for the rapid acquisition of parameters for urban flooding models. This article is based on the parameter optimization method of K-means-ANN-GA, which is not entirely automatic.

In the modeling process, a significant amount of manual intervention is still required for the quantification of reflux sequences, as the interpretability of reflux patterns with spatiotemporal heterogeneity is insufficient. Additionally, based on modeling experience, the efficiency of the parameter optimization method using K-means-ANN-GA is high, but it has not been validated in other different types of regions. The scale dependence and transferability of the method require further research.



**Fig. 10** Flood simulation results of the study area. **a** Simulation results for the event on 10 June 2017; **b** Simulation results for the event on 8 August 2017.

**Acknowledgments** This study was supported by the National Natural Science Foundation of China (Grant Nos. 42271483, 42071364) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant No. KYCX23\_1696). We would like to thank Editage for English language editing.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adnan, R.M., A. Petroselli, S. Heddam, C.A.G. Santos, and O. Kisi. 2021. Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stochastic Environmental Research and Risk Assessment* 35(3): 597–616.
- Cai, T., X. Li, X. Ding, J. Wang, and J. Zhan. 2019. Flood risk assessment based on hydrodynamic model and fuzzy comprehensive evaluation with GIS technique. *International Journal of Disaster Risk Reduction* 35: Article 101077.
- Chang, T.-J., H.-L. Yu, C.-H. Wang, and A.S. Chen. 2021. Overland-gully-sewer (2D-1D-1D) urban inundation modeling based on cellular automata framework. *Journal of Hydrology* 603(B): Article 127001.
- Chen, W., H. Wu, J.S. Kimball, L. Alfieri, N. Nanding, X. Li, L. Jiang, W. Wu, et al. 2022. A coupled river basin-urban hydrological model (DRIVE-Urban) for real-time urban flood modeling. *Water Resources Research* 58(11): Article e2021WR031709.
- Dao, D.A., D. Kim, and D.H.H. Tran. 2022. Estimation of rainfall threshold for flood warning for small urban watersheds based on the 1D–2D drainage model simulation. *Stochastic Environmental Research and Risk Assessment* 36(3): 735–752.
- Feigl, M., S. Thober, R. Schweppe, M. Herrnegger, L. Samaniego, and K. Schulz. 2022a. Automatic regionalization of model parameters for hydrological models. *Water Resources Research* 58(12): Article e2022WR031966.
- Feigl, M., S. Thober, R. Schweppe, M. Herrnegger, L. Samaniego, and K. Schulz. 2022b. Automatic regionalization of model parameters for hydrological models. *Water Resources Research* 58(12): Article e2022WR031966.
- Guo, J., and X. Su. 2019. Parameter sensitivity analysis of SWAT model for streamflow simulation with multisource precipitation datasets. *Hydrology Research* 50(3): 861–877.
- Hu, C., X. Zhang, C. Li, C. Liu, J. Wang, and S. Jian. 2022. Real-time flood classification forecasting based on K-means++ clustering and neural network. *Water Resources Management* 36(1): 103–117.
- Hu, J., Q. Zhou, A. McKeand, T. Xie, and S.-K. Choi. 2020. A model validation framework based on parameter calibration under



- aleatory and epistemic uncertainty. *Structural and Multidisciplinary Optimization* 63: 645–660.
- Huo, J., and L. Liu. 2019. Application research of multi-objective artificial bee colony optimization algorithm for parameters calibration of hydrological model. *Neural Computing and Applications* 31(9): 4715–4732.
- Ichiba, A., A. Gires, I. Tchiguirinskaia, D. Schertzer, P. Bompard, and M.-C.T. Veldhuis. 2018. Scale effect challenges in urban hydrology highlighted with a distributed hydrological model. *Hydrology and Earth System Sciences* 22(1): 331–350.
- Jung, D., Y. Choi, and J. Kim. 2017. Multiobjective automatic parameter calibration of a hydrological model. *Water* 9(3): Article 187.
- Katipoğlu, O.M., and M. Sarıgöl. 2023. Prediction of flood routing results in the Central Anatolian region of Türkiye with various machine learning models. *Stochastic Environmental Research and Risk Assessment* 37(6): 2205–2224.
- Kim, S.W., S.H. Kwon, and D. Jung. 2022. Development of a multi-objective automatic parameter-calibration framework for urban drainage systems. *Sustainability* 14(14): Article 8350.
- Li, J. 2020. A data-driven improved fuzzy logic control optimization-simulation tool for reducing flooding volume at downstream urban drainage systems. *Science of The Total Environment* 732: Article 138931.
- Li, D., H. Li, Y. Zhang, M. Guo, and D. Zhang. 2022. Influence of time step synchronization on urban rainfall-runoff simulation in a hybrid CPU/GPU 1D–2D coupled model. *Water Resources Management* 36(10): 3417–3433.
- Li, Y., J. Zhou, R. Jiang, and M. Zhou. 2016. Classified correction of real-time flood forecasting based on K-Cluster analysis. *China Rural Water and Hydropower* 12(160–162): 167 (in Chinese).
- Liao, R., Z. Xu, C. Ye, X. Shu, and D. Yang. 2023. Simulation of rainstorm waterlogging in Dahongmen drainage area based on SWMM and InfoWorks ICM models. *Water Resources Protection* 39(3): 109–117.
- Liao, R., Z. Xu, C. Ye, B. Zuo, D. Xiang, and X. Shu. 2022. Parameter sensitivity analysis methods of storm water management model. *Journal of Hydroelectric Engineering* 41(6): 11–21.
- Liu, K., W. Bao, J. Que, J. Li, and H. Shu. 2015. Application of K-mean cluster method to flood forecasting based on principal component analysis. *Engineering Journal of Wuhan University* 48(4): 447–450; 458 (in Chinese).
- Liu, C., Z. Han, X. Li, Y. Sun, Y. Tang, D. Hou, and C. Hu. 2023. Urban stormwater rapid simulation based on BIC-Kmeans and SWMM. *Water Resources Protection* 39(5): 79–87.
- Lu, P., and Y. Sun. 2021. Scenario-based hydrodynamic simulation of adaptive strategies for urban design to improve flood resilience: A case study of the Mingzhu Bay Region, Guangzhou, Greater Bay Area. *River Research and Applications* 39(7): 1425–1436.
- Nandi, S., and M.J. Reddy. 2022. An integrated approach to streamflow estimation and flood inundation mapping using VIC, RAPID and LISFLOOD-FP. *Journal of Hydrology* 610: Article 127842.
- Nguyen, D.H., and D.-H. Bae. 2020. Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network. *Journal of Hydrology* 584: Article 124710.
- O'Loughlin, F.E., J. Neal, G.J.P. Schumann, E. Beighley, and P.D. Bates. 2020. A LISFLOOD-FP hydraulic model of the middle reach of the Congo. *Journal of Hydrology* 580: Article 124203.
- Padiyedath Gopalan, S., A. Kawamura, H. Amaguchi, T. Takasaki, and G. Azhikodan. 2019. A bootstrap approach for the parameter uncertainty of an urban-specific rainfall-runoff model. *Journal of Hydrology* 579: Article 124195.
- Peng, L., Z. Xu, G. Zhao, B. Zuo, J. Wang, and S. Song. 2021. Simulation of urban rainstorm waterlogging processes based on SWMM and LISFLOOD-FP models: Case study in Jinan City. *South-to-North Water Transfers and Water Science & Technology* 19(6): 1083–1092.
- Rai, P.K., B.R. Chahar, and C.T. Dhanya. 2016. GIS-based SWMM model for simulating the catchment response to flood events. *Hydrology Research* 48(2): 384–394.
- Schilling, J., and J. Tränckner. 2022. Generate\_SWMM\_inp: An open-source QGIS plugin to import and export model input files for SWMM. *Water* 14(14): 2262.
- Shahed Behrouz, M., Z. Zhu, L.S. Matott, and A.J. Rabideau. 2020. A new tool for automatic calibration of the storm water management model (SWMM). *Journal of Hydrology* 581: Article 124436.
- Shen, H., Z. Xu, S. Zhang, and Q. Li. 2019. Analysis and test on rainfall control modes of LID facilities with different confluence types. *Journal of Hydraulic Engineering* 50(5): 578–588.
- Shustikova, I., J.C. Neal, A. Domeneghetti, P.D. Bates, S. Vorogushyn, and A. Castellarin. 2020. Levee breaching: A new extension to the LISFLOOD-FP model. *Water* 12(4): Article 942.
- Sinnathamby, S., K.R. Douglas-Mankin, and C. Craige. 2017. Field-scale calibration of crop-yield parameters in the soil and water assessment tool (SWAT). *Agricultural Water Management* 180(Part A): 61–69.
- Snieder, E., and U.T. Khan. 2023. A novel ensemble algorithm based on hydrological event diversity for urban rainfall-runoff model calibration and validation. *Journal of Hydrology* 619: Article 129193.
- Sosa, J., C. Sampson, A. Smith, J. Neal, and P. Bates. 2019. A toolbox to quickly prepare flood inundation models for LISFLOOD-FP simulations. *Environmental Modelling & Software* 123: Article 104561.
- Song, X., Q. Shu, H. Wang, and W. Liao. 2009. Comparison and application of SCE-UA, genetic algorithm and simplex method. *Engineering Journal of Wuhan University* 42(1): 6–9; 15 (in Chinese).
- Sun, Y., C. Liu, X. Du, F. Yang, Y. Yao, S. Soomro, and C. Hu. 2022. Urban storm flood simulation using improved SWMM based on K-means clustering of parameter samples. *Journal of Flood Risk Management* 15(4): Article e12826.
- Wang, F., X. Qing, S. Yang, and Z. Cui. 2022. Automatic calibration of SWMM parameters based on PySWMM. *China Water & Wastewater* 38(21): 124–130.
- Wang, Y., S. Zhu, L. Yuan, and R. Deng. 2020. An automatic parameter calibration method for the SWAT model in runoff simulation. *River Research and Applications* 36(7): 1321–1333.
- Willis, T., N. Wright, and A. Sleight. 2019. Systematic analysis of uncertainty in 2D flood inundation models. *Environmental Modelling & Software* 122: Article 104520.
- Wood, M., R. Hostache, J. Neal, T. Wagener, L. Giustarini, M. Chini, G. Corato, P. Matgen, and P. Bates. 2016. Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identifiability techniques. *Hydrology and Earth System Sciences* 20(12): Article 4983–4997.
- Wu, Z., B. Ma, H. Wang, C. Hu, H. Lv, and X. Zhang. 2021. Identification of sensitive parameters of urban flood model based on artificial neural network. *Water Resources Management* 35(7): 2115–2128.
- Xu, W., and Y. Peng. 2015. Research on classified real-time flood forecasting framework based on K-means cluster and rough set. *Water Science and Technology* 71(10): 1507–1515.
- Yan, X., K. Xu, W. Feng, and J. Chen. 2021. A rapid prediction model of urban flood inundation in a high-risk area coupling machine learning and numerical simulation approaches. *International Journal of Disaster Risk Science* 12(6): 903–918.
- Yang, Q., Z. Ma, and S. Zhang. 2022. Urban pluvial flood modeling by coupling raster-based two-dimensional hydrodynamic model and SWMM. *Water* 14(11): Article 1760.
- Yang, Y., L. Sun, R. Li, J. Yin, and D. Yu. 2020. Linking a storm water management model to a novel two-dimensional model for urban pluvial flood modeling. *International Journal of Disaster Risk Science* 11(4): 508–518.
- Ye, C., Z. Xu, X. Lei, Y. Chen, C. Ban, and H. Su. 2021. Rapid simulations of storm water runoff in urban community scale. Case study of a community compound in Fuzhou City. *Journal of Hydroelectric Engineering* 40(10): 81–94.

- Ye, C., Z. Xu, X. Lei, W. Liao, X. Ding, and Y. Liang. 2022. Assessment of urban flood risk based on data-driven models: A case study in Fuzhou City, China. *International Journal of Disaster Risk Reduction* 82: Article 103318.
- Yuan, S., D. Li, Y. Chen, Z. He, Q. Cheng, and F. Liu. 2021. Automatic calibration procedure of storm water management model parameters based on back propagation neural network algorithm. *China Water & Wastewater* 37(21): 125–130.
- Zang, Y., Y. Meng, X. Guan, H. Lv, and D. Yan. 2022. Study on urban flood early warning system considering flood loss. *International Journal of Disaster Risk Reduction* 77: Article 103042.
- Zeng, J., Y. Mai, Z. Li, X. Ren, J. Pan, and G. Huang. 2020. Sensitivity analysis of SWMM parameters in Guangzhou Tianhe wisdom city. *Water Resources Protection* 36(3): 15–21.
- Zeng, Z., Z. Wang, and C. Lai. 2022. Simulation performance evaluation and uncertainty analysis on a coupled inundation model combining SWMM and WCA2D. *International Journal of Disaster Risk Science* 13(3): 448–464.
- Zeng, Z., Z. Wang, X. Wu, C. Lai, and X. Chen. 2017. Rainstorm waterlogging simulations based on SWMM and LISFLOOD models. *Journal of Hydroelectric Engineering* 36(5): 68–77.
- Zhang, M., X. Zhang, T. Wang, Y. Xu, and Q. Zhang. 2022. Method of determining hydrological response units for discretization of micro-topography. *Research of Soil and Water Conservation* 29(1): 86–91; 99 (in Chinese).
- Zhu, M., F. Wang, T. Zhang, X. Yin, W. Wu, and H. Zhu. 2018. Research and practice of standard system of city surveying and mapping geographic information—taking Nanjing city as an example. *Science of Surveying and Mapping* 43(12): 151–157.