



A PRESENTATION OF FLOODS MONITORING SYSTEM

G.RAJKUMAR

B.SURIYA

K.NARAYAN

V.POOVIZHI

MENTORED BY

P.SUJATHA

Prediction and early warning method of inundation process at waterlogging points based on Bayesian model average and data-driven

ARTICLE INFO

Keywords:

Urban flood
Waterlogging point
Bayesian model average
Prediction

ABSTRACT

Study region: Built-up area of Zhengzhou, China.

Study focus: An urban flood forecasting model based on multi-model integrated forecasting is proposed. Based on the data of rainfall process and inundation process, a multi-model ensemble prediction model was established for 27 typical waterlogging points in Zhengzhou by using BMA (Bayesian Model Average) coupling different data-driven methods. Combined with the rainfall forecast data, the BMA was driven to predict and warn the inundation process of urban flood.

New hydrological insights for the region: Using the BMA to predict urban flood can improve the accuracy and stability of using single model to predict urban flood. The results indicate that the prediction accuracy of BMA model is 36–46% higher than that of single model, which demonstrate that BMA makes effective use of the advantages of each model and can provide higher accuracy in prediction and early warning. Additionally, BMA significantly reduces the uncertainty of single model prediction in the prediction of inundation process. The analysis of early warning in two different urban flood events indicates that BMA is more suitable for the prediction of severe waterlogging and illustrates the great potential and prospects of BMA in urban flood early warning.

1. Introduction

Globally, the occurrence of urban floods has been unprecedented resulting in huge economic and social losses (Sundaram et al., 2021). In China, a country with rapid urbanization construction and varying topography and climatic conditions, more than 100 cities suffered from urban floods every year from 2008 to 2019 (Wu et al., 2021), causing direct economic losses of more than 100 billion yuan every year. In July 2021, the "7.20 rainstorm" in Zhengzhou, China caused 380 deaths and a direct economic loss of 40.9 billion yuan. Urban flood disaster has become a crucial problem restricting the healthy development of China's economy and society (Zheng et al., 2016). Therefore, it is necessary to explore the rapid, accurate, and stable prediction and early warning methods of urban flood, in order to reduce the loss of urban flood disasters.

In recent years, hydrologic experts are devoted to study a variety of one-dimensional models, two-dimensional (2D) hydrodynamic models and data-driven models to simulate and predict urban flood disaster (Alfieri et al., 2014; Bates et al., 2010; Dimitriadis et al., 2016; Hou et al., 2021). However, some recent studies have shown that the prediction performance of a single method or model is

* Corresponding author.

E-mail address: wanghuiliang@zzu.edu.cn (H. Wang).

<https://doi.org/10.1016/j.ejrh.2022.101248>

Received 24 February 2022; Received in revised form 6 October 2022; Accepted 20 October 2022

Available online 21 October 2022

2214-5818/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

always limited (Liu and Merwade, 2018; Yan and Moradkhani, 2016). Due to the uncertainty of model structure, parameters and inputs, a specific model that works well in one flood event might not perform well in another (Liu and Merwade, 2019). Therefore, relying on a single model for the flood prediction usually increases the statistical deviation of prediction (Liu and Merwade, 2018). In order to deal with this problem, recent studies used multi-model combination methods to deal with model uncertainty and improve model performance. Among the current existing multi-model combination approaches, including simple model average (SMA), weighted model average (WMA), and multi-model super-ensemble (MMSE) (Ajami et al., 2006; Chowdhury and Sharma, 2009; Najafi and Moradkhani, 2015, 2016), BMA is one of the most common and widely used methods to generate a reliable deterministic model prediction by synthesizing several models (Li et al., 2017; Liu and Merwade, 2019; Madadgar and Moradkhani, 2014). In recent years, the BMA technique has been widely used in hydrology (Basher et al., 2020; Yan and Moradkhani, 2015), meteorology (Baran and Möller, 2015), ecology (Yang et al., 2021) and economy (Fernández et al., 2001). Huo et al. (2019) simulated the flood process of three semi-humid regions by using seven hydrological models and the BMA method. They found that BMA method effectively makes use of the advantages of each model and provides more accurate prediction. Qu et al. (2017) applied the BMA method to predict runoff in Fuhe River Basin, China. The results show that BMA post-processing approach can improve the performance of runoff prediction and obtain a more accurate prediction probability density function over a range of lead times from 24 to 120 h. Yan and Moradkhani (2016) proposed an extreme flood information estimation method considering the uncertainty of distribution and model structure using the BMA method. They found that the main uncertainty in extreme flood forecasting is determined by the model structure. Liu and Merwade (2018) established a robust deterministic water level prediction and prediction distribution by combined the different hydraulic models. The results show that the BMA method provides consistent and reliable basin flood level prediction, although it does not always outperform the best model.

As indicated earlier, the BMA provides a more robust method for the flood prediction. The BMA technique can not only quantify the uncertainty of the model through a priori probability and a posteriori probability, but also weighted average the model according to a posteriori probability. Its core idea is to assign a weight to each model according to the performance of the model during training (Liu and Merwade, 2018). The advantage of this approach is that BMA provides a model weighted average prediction of the interested variable (Rings et al., 2012). However, to the authors' best knowledge, literature about the BMA method to study the urban flood simulation and prediction is quite limited, and no study has applied the BMA and data-driven approach to the prediction of local urban flooding.

In summary, this study attempts to establish a comprehensive urban flood prediction model using BMA and data-driven methods, and explore the application of data-driven model using BMA as a fusion-based method to predict urban flood depth at different waterlogging points. The main objectives that are addressed in this study are to; (i) construct a comprehensive prediction model by

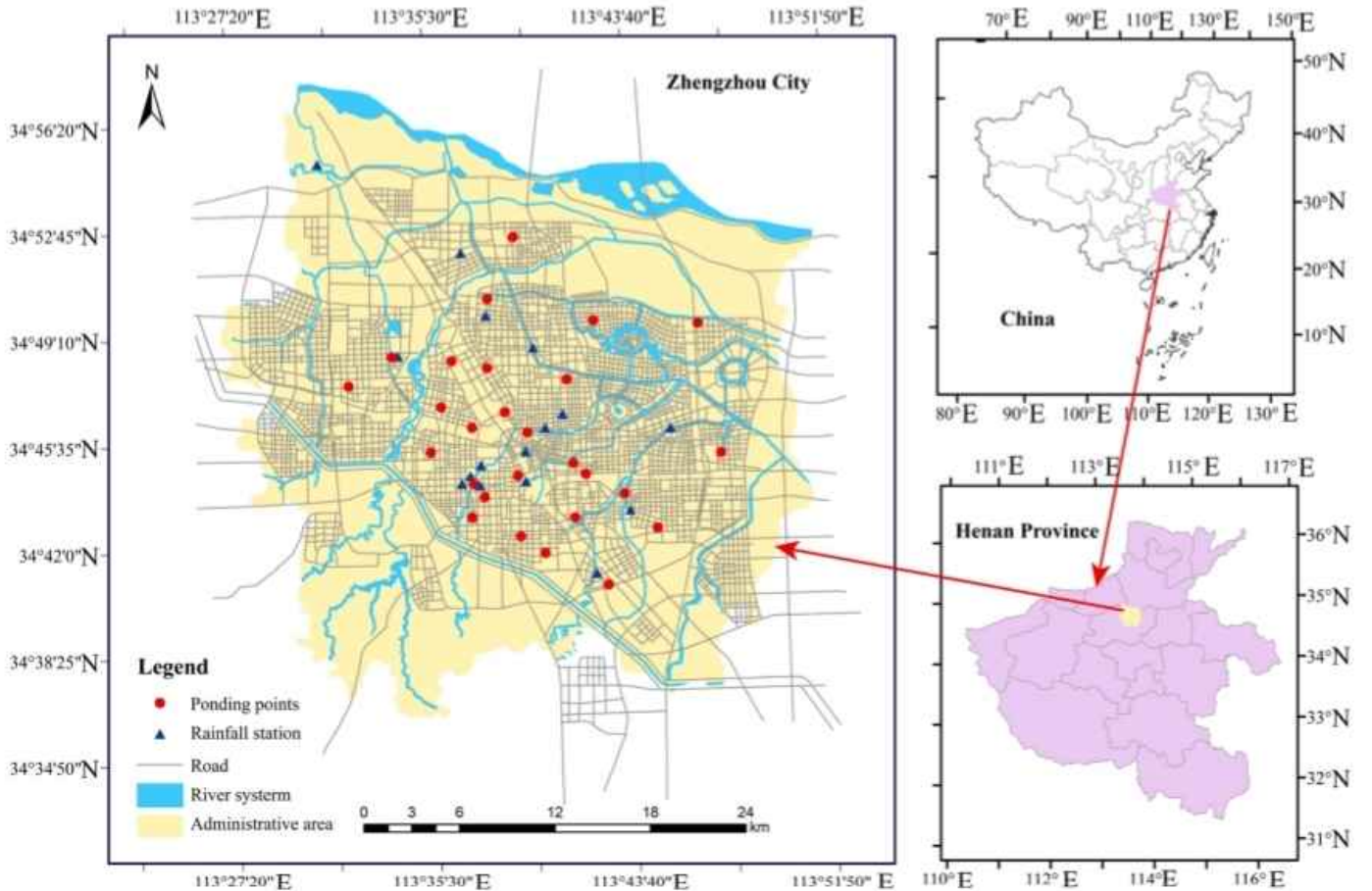


Fig. 1. The location of Zhengzhou City.

BMA and three machine learning methods (support vector machines (SVM), Back Propagation Neural Network (BPNN) and Adaptive Boosting (AdaBoost)); (ii) analyze the performance of BMA and single model in urban flood prediction from the perspectives of model efficiency, accuracy and stability; and (iii) use rainfall forecast data to drive BMA model for fine early warning of urban flood. The innovation of this study is to establish an accurate and stable prediction model of urban flood by using BMA and data-driven model. The research results can provide new ideas for urban flood prediction, as well as provide a scientific basis for urban flood control and early warning.

2. Materials and methods

2.1. study area

Zhengzhou city, the capital of Henan Province in China, was selected for this study (Fig. 1). As of the end of 2020, the permanent population of Zhengzhou exceeded 12 million, and the total Gross Domestic Product (GDP) reached 1.20 trillion yuan. The annual average precipitation in this region is 639.5 mm. However, the distribution of precipitation is very uneven and the interannual variation is large, the precipitation in the flood season (June to September) accounts for about 60% of the annual precipitation. In the year of 2018, 2019 and 2021, serious urban floods occurred successively in this region, especially the urban flood event in July 20, 2021, which caused 380 deaths and a direct economic loss of 40.9 billion yuan. The urban flood disaster has become a prominent problem affecting the sustainable and healthy development of Zhengzhou.

2.2. data

In this paper, the rainfall and inundation data of 18 historical rainfall events in Zhengzhou from 2016 to 2018 were selected for inundation process simulation to evaluate the performance of BMA model. Among them, three randomly selected rainfall events with different rainfall intensity and duration were used as test data, and the other 15 rainfall events were used as training data. And the rainfall forecast data of two different types of rainfall events from 2019 to 2020 were used for early warning. The detailed data were described as follows:

Rainfall data: Rainfall data refer to the rainfall process data with 10 min time resolution observed and recorded by 16 rainfall stations in Zhengzhou, which are: June 4, June 23, August 4, August 25 and September 12 in 2016; May 22, June 22, July 18, August 12, August 25 and August 30 in 2017; May 15, July 4, July 13, July 27, August 10, September 15 and September 25 in 2018. The rainfall process data of each waterlogging point were obtained by interpolating the rainfall process data of the rainfall station by using GIS and Kriging interpolation method. The rainfall data of the rainfall station were collected from the meteorological department of Zhengzhou city.

Waterlogging data: Waterlogging data refer to the water depth of waterlogging points in historical waterlogging events, which mainly from the monitoring equipment of waterlogging at each road intersection. The water depth data of 27 typical waterlogging points with 1 min time resolution were collected from the urban management department of Zhengzhou. In order to be consistent with rainfall data, the time resolution of waterlogging data was set to 10 min. The waterlogging data of 18 rainfall events from 2016 to 2018 correspond to the above rainfall data, are used as the training and verification data of the model. The water depth data of rainfall events on August 1, 2019 and August 7, 2020 were used as the verification data of early warning results.

Rainfall forecast data: Rainfall forecast data is basis of urban flood early warning. The rainfall forecast data of two rainfall events in August 1, 2019 and August 7, 2020 were obtained by calling the API of the meteorological technology company (Caiyun Technology). The rainfall forecast data have a temporal resolution of 1 min and a spatial resolution of 1 km. In order to unify the time resolution, the time resolution of rainfall forecast data was changed to 10 min. And the forecast period of rainfall forecast data is from 0 to 120 min.

2.3. Data-driven approach

For the purpose of modeling inundation process prediction model for each waterlogging point, three data-driven models with different structures and complexities, including SVM, BPNN and AdaBoost, were selected. SVM is a supervised machine learning algorithm. The core idea is to use the kernel function to find the optimal hyperplane by mapping data to a high-dimensional feature space (Yu et al., 2008). BPNN is a feedforward neural network, which is widely used in data mining. It repeatedly modifies the weight of each layer through two processes of forward propagation and back propagation to train the model (Chen et al., 2010). AdaBoost is an ensemble learning algorithm, which was usually selected the decision tree as the basic learner. The core idea is to update the sample weight value by using the error of the previous iteration during each iteration (Taherkhani et al., 2020). In order to ensure the consistency of model comparison, the parameters of the above three data-driven models were calibrated by Monte Carlo sampling. Mean Absolute Error (MAE) and Nash-Sutcliffe efficiency coefficient (NSE) were selected to evaluate their accuracy and performance.

The results using single model are the basis of the BMA model. Therefore, the above data-driven models were used to construct the BMA model for inundation process prediction in this study. However, the modeling process of using data-driven model will not be introduced in detail in this study since the purpose and core of this paper is to evaluate the performance of BMA model in the local urban flooding prediction. A complete mathematical description and modeling process using data-driven approach can be found in the previous research (Wu et al., 2020a, 2020b).

2.4. Basic principle of the BMA algorithm

Assuming that y is simulation variable of multi-mode, $[y_1, y_2, \dots, y_T]$ is the measured inundation process data required to calibrate the model, $f = [f_1, f_2, \dots, f_k]$ represents the model space composed of SVM, BPNN and AdaBoost. $p_k(y|f_k, D)$ is the posterior distribution of simulation variable y given measured data D and model. According to the law of total probability, the posterior probability of simulation variable y in BMA can be expressed as:

$$p(y|D) = \sum_{k=1}^K p(f_k|D) \cdot p_k(y|f_k, D) \quad (1)$$

Where $p(f_k|D)$ is the posterior probability of the k -th model f_k given the measured data w_k . (In other words, w_k is the weight of each model in BMA, $\sum_{k=1}^K w_k = 1$).

The predicted value of BMA was obtained by the weighted average of each data-driven model. Assuming that the predicted value and measured value of each model obey the normal distribution, the expectation and variance of the BMA model can be expressed as:

$$E(y|D) = \sum_{k=1}^K p(f_k|D) \cdot E[p_k(y|f_k, D)] = \sum_{k=1}^K w_k \cdot f_k \quad (2)$$

$$Var[y|D] = \sum_{k=1}^K (f_k - \sum_{k=1}^K w_k f_k)^2 + \sum_{k=1}^K w_k \cdot \sigma_k^2 \quad (3)$$

EM (Expectation maximization) algorithm is an effective method to calculate BMA, which requires that the prediction results of data-driven models obey normal distribution (Raftery et al., 2005). In this study, the measured and model simulation sequences were normalized by the Box-Cox function in MATLAB. Based on this, the probability distribution parameters w_k and σ_k^2 of simulation variable y in BMA were calculated using the EM algorithm. Assuming that $\theta = \{w_k, \sigma_k^2, k = 1, 2, \dots, K\}$ is the parameter to be solved of the BMA model, the log likelihood function can be expressed as:

$$l(\theta) = \ln \{p(y|D)\} = \ln \left\{ \sum_{k=1}^K w_k \cdot p_k(y|f_k, D) \right\} \quad (4)$$

The EM algorithm is iterative and alternates between two the E step and the M step. In step E, latent variable is estimated according to the current guess of the parameters. In step M, θ is estimated according to the current value of the latent variable. The EM algorithm iterates repeatedly through expectation and maximization. The iteration stops when the error of likelihood obtained by the iteration is less than the allowable error (10^{-6} in this study). The log-likelihood function will converge to the local maximum of the likelihood since the log-likelihood function increased at each EM iteration (Wu, 1983). A complete mathematical and technical description of BMA and EM algorithm can be found in Raftery et al. (2005), Liu and Merwade (2018).

2.5. The uncertainty analysis of BMA model

Uncertainty of the model reflects the stability of the prediction performance. After the weight (w_k) and variance (σ_k^2) of the BMA model were obtained using the EM algorithm, the uncertainty interval of the prediction value at any time t was deduced using the Monte Carlo Method (Hammersley and Handscomb, 1964). A large number of sampling and gradual approximation are the core idea of Monte Carlo Method. When the sampling times are large enough, the deduced results gradually approach the real results stably. Therefore, the sampling time was determined to 10,000 at every time node of inundation process in this study, and 95% confidence interval of the BMA model was deduced. As shown in Fig. 2, the estimated value (\hat{y}) at any time t was calculated by sampling method,

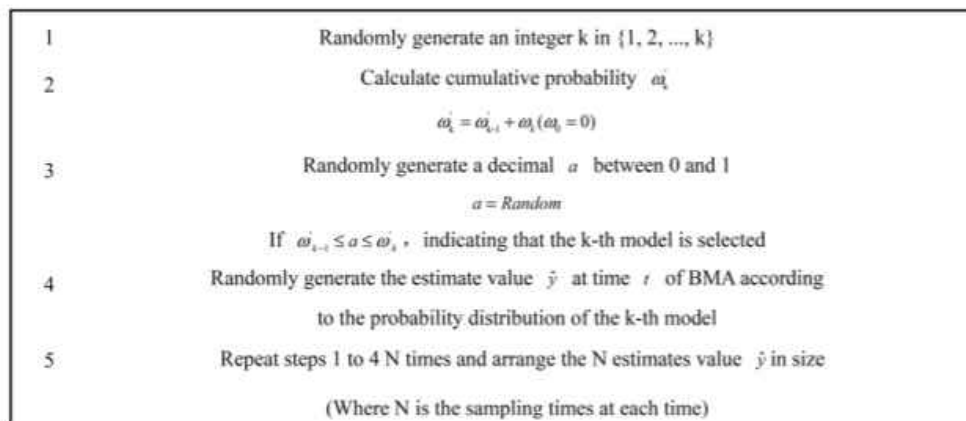


Fig. 2. Steps of deriving uncertainty interval Using the Monte Carlo Method.

and arrange all the estimated values (\hat{y}) in size. The 95% confidence interval of BMA model is the middle part of the quantile of 2.5% and 97.5%.

In addition, the Monte Carlo Method was also used for single models, and the sampling times were also set to 10,000 times. The 95% confidence interval of the predicted value was deduced according to the probability distribution of each model.

2.6. The evaluation statistics of BMA

Three types of indicators are considered to evaluate the performance of BMA algorithm in this study. In the first type, the MAE and NSE were selected to evaluate the performance of the model in validation periods (Eqs. (5), (6)) since the BMA approach produces both deterministic prediction and probabilistic prediction. The smaller the MAE and the closer the NSE is to 1, indicating that the model accuracy and performance is better.

$$MAE = \frac{\sum_{i=1}^n |x_i - y_i|}{n} \quad (5)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2} \quad (6)$$

Where x_i is the predicted value, y_i is the measured value, and \bar{y} is the average of the measured value.

In the second type, Overall Accuracy (OA), Overestimation Rate (OR) and Underestimation Rate (UR) were used to evaluate the overall accuracy of early warning results (Eqs. (7–9)). Specifically, OA is the proportion of correct samples in all samples of early warning results. UR refers to the proportion of samples with early warning level lower than the observed waterlogging level. In contrast, OR refers to the proportion of samples with early warning level higher than the observed waterlogging level ($OA + OR + UR = 1$).

$$OA = \frac{\sum_{n=1}^4 \sum_{m=1}^4 x_{mn}}{N} (m = n) \quad (7)$$

$$UR = \frac{\sum_{n=1}^4 \sum_{m=1}^4 x_{mn}}{N} (m < n) \quad (8)$$

$$OR = \frac{\sum_{n=1}^4 \sum_{m=1}^4 x_{mn}}{N} (m > n) \quad (9)$$

Where m is the predicted waterlogging level, n is the measured waterlogging level, N refers to the total number of samples, and x_{mn} refers to the number of samples that the measured m level is predicted as level n .

In addition, Precision and Recall are selected to evaluate the accuracy of early warning results at each level (Eqs. (10), (11)). Precision is the proportion of true positive samples in the prediction samples, Recall refers to the proportion of true positive samples in all positive samples (Faceli, 2011).

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

Where TP is the number of samples correctly classified as positive, FP is the number of samples incorrectly classified as positive because the right category is negative, and FN is the number of samples incorrectly classified as negative because the right class is positive (Faceli et al., 2011).

When calculating the Precision and Recall of early warning results, TP, FP and FN need to be determined first. On this basis, the Precision and Recall of early warning results at each level were calculated by using Eq. (10) and Eq. (11). Finally, the overall Precision and Recall of early warning results were obtained by weighted averaging (the weight was determined according to the number of waterlogging points under each level) the Precision and Recall of each early warning level.

3. Results and discussion

3.1. The characteristics of rainfall and waterlogging data

3.1.1. The characteristics of rainfall data

In this study, the rainfall data of 18 rainfall events from 2016 to 2018 were selected as the training data of the ponding prediction model. From the perspective of precipitation (Fig. 3), the precipitation of the 18 urban flood events was distributed between 38 mm and 105 mm, of which the rainfall events over 70 mm account for 83.3% of the total rainfall events. The rainfall duration was distributed between 170 min and 290 min, and the location coefficient was distributed between 0.19 and 0.77. Among them, the rainfall events with the location coefficient lower than 0.5 account for 61.1% of the total rainfall events. Therefore, although the rainfall events with small location coefficient (less than 0.5) and large precipitation (more than 70 mm) account for the majority of the total rainfall events, the selected rainfall events basically cover different types of rainfall, which can provide more comprehensive sample data for the construction of ponding prediction model. In addition, the average precipitation of 18 urban flood events reached 79.4 mm, indicating that large precipitation was one of the main reasons for urban flood. Furthermore, it can be easily found that the precipitation of the 1st, 14th and 16th urban flood events did not exceed 45 mm (Fig. 3), and the rainfall duration of three urban flood events were no more than 190 min, indicating that the short-term heavy rainfall with a small total precipitation was also one of the reasons for urban flood.

3.1.2. The characteristics of waterlogging data

Fig. 4(a) reflects the water accumulation characteristics of 27 waterlogging points in 18 rainfall events. The maximum depth and average depth of water accumulation in the 27 waterlogging points fluctuate greatly, the maximum depth is between 0.15 m and 0.68 m, and the average depth is distributed between 0.11 m and 0.55 m. Among them, the water accumulation at the waterlogging point of number 10 is the smallest. The water accumulation at the waterlogging points of number 1, 16, and 26 is very serious, and the maximum depth is more than 0.5 m. Urban flood control should pay more attention to these waterlogging points with serious water accumulation. In contrast, the minimum depth of water accumulation at these waterlogging points fluctuates less, ranging from 0.05 m to 0.19 m. The main reason may be that it is difficult to form serious water accumulation at these waterlogging points under relatively small precipitation, which makes the water accumulation depth show a small difference. In general, the water accumulation depth of the waterlogging points in the 18 rainfall events is between 0.05 m and 0.68 m, covering different degrees of waterlogging events, which can provide more comprehensive waterlogging sample data for the construction of urban flood prediction models.

In order to analyze the water accumulation characteristics of 27 waterlogging points under similar rainfall conditions, the water depth of each waterlogging point under the 2nd, 7th and 18th rainfall events (with similar rainfall, rainfall duration and location coefficient) was analyzed. As shown in Fig. 4(b), under similar rainfall conditions, the average water depth of each waterlogging point was between 0.08 m and 0.44 m, and the number of waterlogging points with more than 0.2 m water depth exceeded 88%. It shows that although most waterlogging points have accumulated water, there are significant differences in the severity of the accumulated water. The main reason is that the topography, land use and drainage conditions of each waterlogging point were quite different. The waterlogging points in sunken terrain such as tunnels and overpasses are more likely to form serious waterlogging. This spatial difference also provides some reference for urban flood control, which should pay more attention to these waterlogging points with serious waterlogging. In addition, these spatial differences of waterlogging data also can provide more comprehensive waterlogging sample data for the construction of urban flood prediction models.

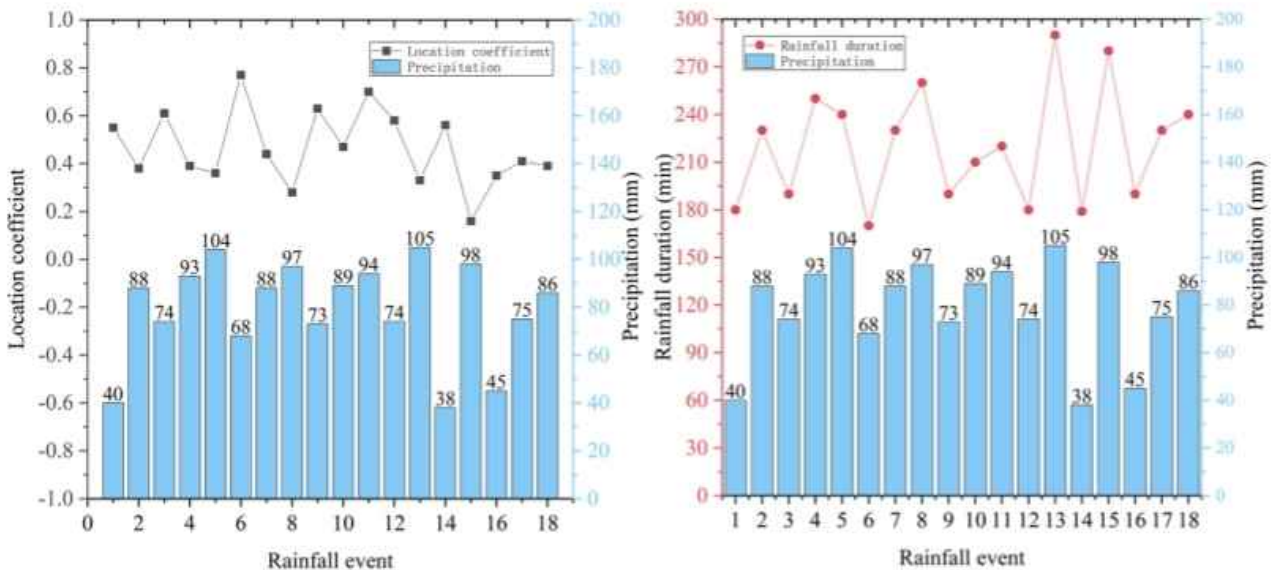


Fig. 3. Characteristics of rainfall data.

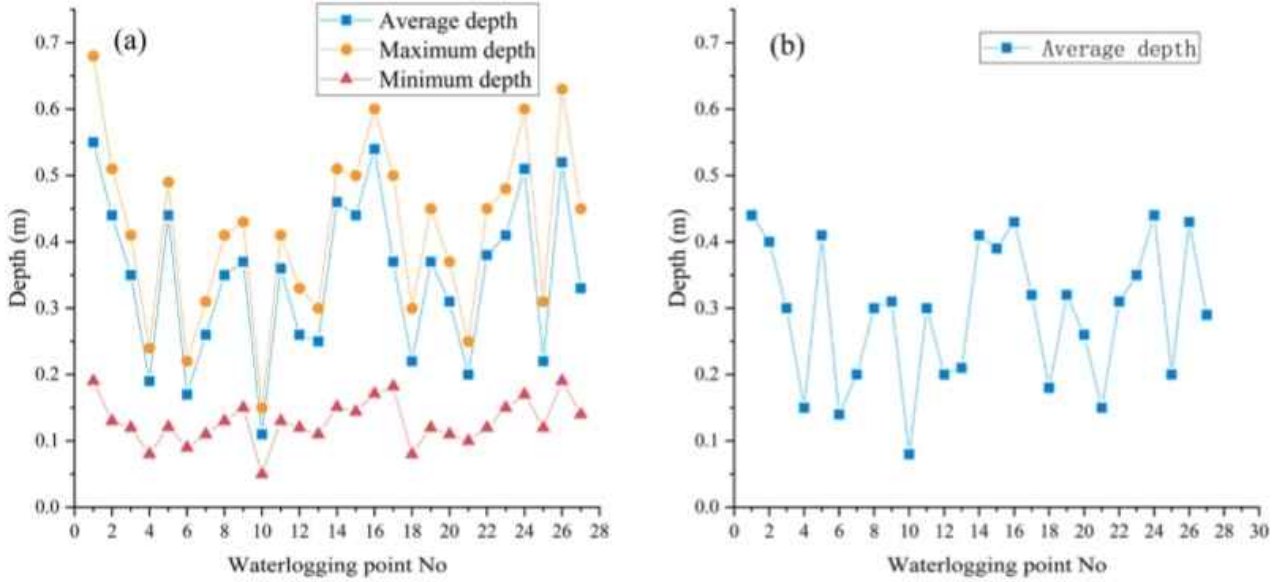


Fig. 4. Characteristics of waterlogging data under (a) 18 rainfall events, (b) the 2nd, 7th and 18th rainfall events.

3.2. Validation of BMA prediction performance

3.2.1. The overall performance of BMA

Three randomly selected rainfall events with different location coefficients and rainfall duration from 2016 to 2018 were selected as verification events to evaluate the performance of BMA model, which were: August 25 in 2016 (with 0.39 location coefficient and 250 min rainfall duration), August 30 in 2017 (with 0.70 location coefficient and 220 min rainfall duration), July 27 in 2018 (with 0.16 location coefficient and 280 min rainfall duration). The Root Mean Square Error (RMSE) was selected as the objective function. The prediction sequence results of BMA model were obtained by EM algorithm. In addition, the performance of BMA model was analyzed from the running time and accuracy of the model. In terms of model running time, the average running time of BMA and three single models were 5.6 s (BMA), 1.6 s (BPNN), 4.2 s (SVM) and 0.6 s (AdaBoost). It can be seen that BMA model has the longest running time. The main reason is that BMA needs to couple multiple models for prediction, so that the running time of BMA model contains the running time of single model. However, although the calculation efficiency of BMA is lower than single model, the running time of 5.6 s will hardly affect the prediction period of flood warning, which is also a significant advantage of data-driven model in calculation efficiency.

Fig. 5 shows the overall accuracy and difference between the BMA model and the single model. The prediction results of SVM, BPNN and AdaBoost were greater than 0.94, indicating that these models have good prediction performance in local urban flooding prediction. Among them, SVM has the highest NSE and the lowest MAE, showing the best prediction effect. In contrast, AdaBoost has a lower NSE and a higher MAE, showing the worst performance. The main reason was that the limited sample data make it difficult for AdaBoost to give full play to its advantages, and SVM also performs well in the limited sample. making the performance of SVM in local urban flooding prediction slightly higher than AdaBoost. In addition, although both the single model and BMA model have achieved good results in local urban flooding prediction, the BMA model has further improved the prediction accuracy. The MAE of the local urban flooding prediction in BMA is 36–46% lower than that of SVM, BPNN and AdaBoost. Specifically, the MAE in BMA is less than

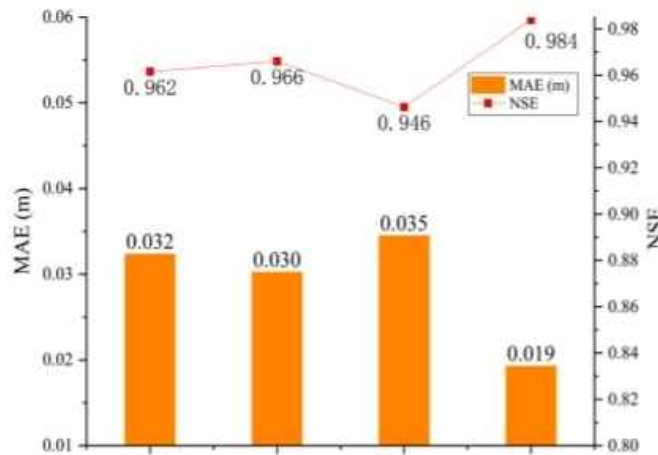


Fig. 5. The overall accuracy comparison of the BMA model and single model (SVM, BPNN and AdaBoost).

0.02 m (0.019 m), but the MAE in SVM, BPNN and AdaBoost are more than 0.03 m. In addition, the NSE of the BMA model is also significantly higher than SVM, BPNN and AdaBoost. Therefore, the above results indicate the overall accuracy of the BMA model is significantly better than the single model (SVM, BPNN and AdaBoost) in the local urban flooding prediction.

Fig. 6 demonstrates the performance between BMA model and single model in different types of rainfall events. BPNN has the highest prediction accuracy in the verification event 1 and 2, but AdaBoost shows higher accuracy in the verification event 3, which indicates that it is difficult for the single model to maintain the highest prediction accuracy in all rainfall event. In contrast, the BMA model always has the lowest MAE and the highest NSE in three different types of rainfall events, which shows that the stability of BMA model is significantly higher than SVM, BPNN and AdaBoost. In addition, Fig. 6 showed that the prediction performance of AdaBoost in different verification events were quite different. Although it has better prediction performance than SVM and BP in the verification event 3, the prediction performance in the first verification event was significantly lower than other models. On the contrary, the prediction performance of SVM and BP was relatively stable. The main reason may still be that the amount of rainfall and waterlogging data was not large enough, making it difficult for AdaBoost to fully explore the potential relationship between rainfall and waterlogging data.

3.2.2. The performance of BMA in each waterlogging point

In order to verify the robustness of BMA model, the prediction performance of the BMA and single model in each waterlogging point was analyzed. It should be pointed out that the characteristics and number of waterlogging points under different rainfall events are uncertain. Therefore, the waterlogging points in the three verification events may have certain differences, which may cause interference to the model performance analysis. Therefore, this study selects 27 identical waterlogging points in three validation events to analyze the prediction performance of the model, in order to eliminate the impact of changes in waterlogging points on the performance analysis.

As shown in Fig. 7, the MAE fluctuation range of BMA model is between 0 and 0.030 m, which is 34.8–68.1% lower than that of SVM (0.009–0.055 m), BPNN (0.010–0.056 m) and AdaBoost (0.012–0.106 m). And the lowest NSE value of BMA model (0.945) is 26.0–33.1% higher than that of SVM (0.74), BPNN (0.75) and AdaBoost (0.71). These results show that the prediction accuracy of BMA model is significantly better than that of single model (SVM, BPNN and AdaBoost) in the first validation event. Similarly, in the second and third validation events, the lowest NSE values of BMA model were 0.87 and 0.922, which were 14.4–72.7% and 15.0–61.8% higher than that of single model. And compared with the single model, the MAE fluctuation range of BMA model also decreased by 29.1–48.4%. These results indicate that the BMA model not only has higher prediction accuracy in different waterlogging points, but also has stable prediction performance. However, the prediction performance of BMA is not the highest in all cases. For example, the NSE of BMA (0.95) is lower than that of BPNN model (0.98) in the waterlogging point of number 21. The main reason is that the results of BMA were obtained by weighted averaging the results of single models. Therefore, when the local error of single model is large, the prediction accuracy of the BMA model may be lower than that of the single model with the highest accuracy. Nevertheless, the prediction performance of the BMA model is significantly better than that of the single model in most cases, which can provide more stable urban flood prediction.

In order to deeply explore the potential relationship between the model performance and different types of waterlogging points, three different types of waterlogging points (#16, #23 and #26) were selected from 27 waterlogging points to compare the prediction performance by reflecting the characteristics of the waterlogging points. Among them, the waterlogging point of number 26 was located near the drainage outlet at the bottom of the tunnel, the waterlogging point of number 16 was located near the road intersection, and the waterlogging point of number 23 was located at the edge of the auxiliary road under the elevated road. As shown in Fig. 8, from the perspective of water accumulation characteristics, the water depth at the waterlogging point of number 26 is the largest, and the water accumulation process line shows the characteristics of steep increase, gentle, and then slow decrease. The main reason is that the waterlogging point of number 26 is located at the underpass tunnel. Due to the influence of the microtopography of

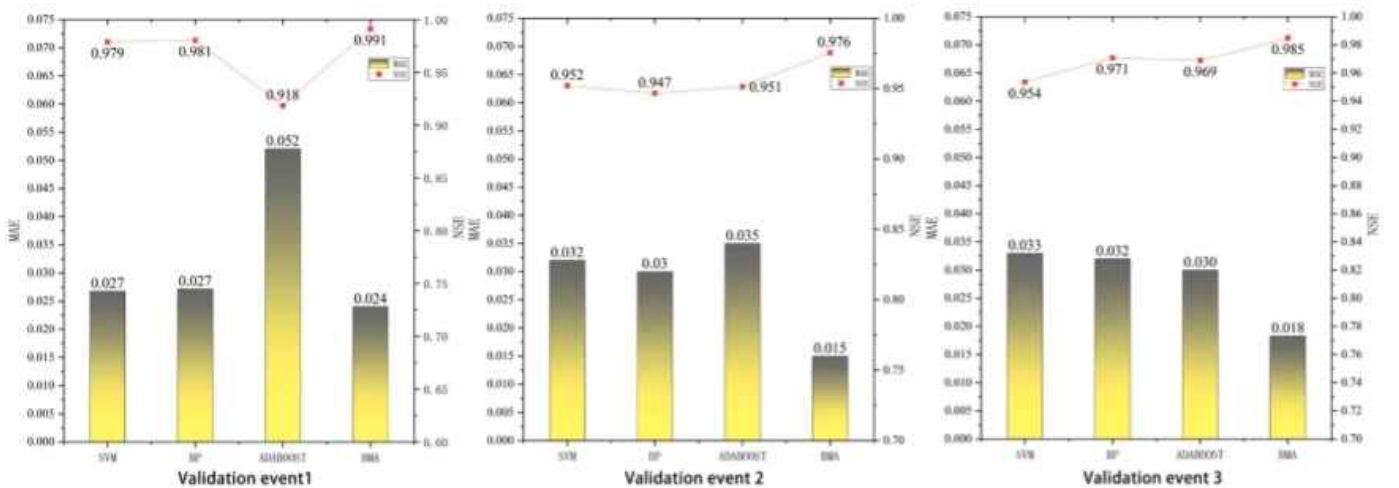


Fig. 6. The performance between BMA model and single model in different types of rainfall events.

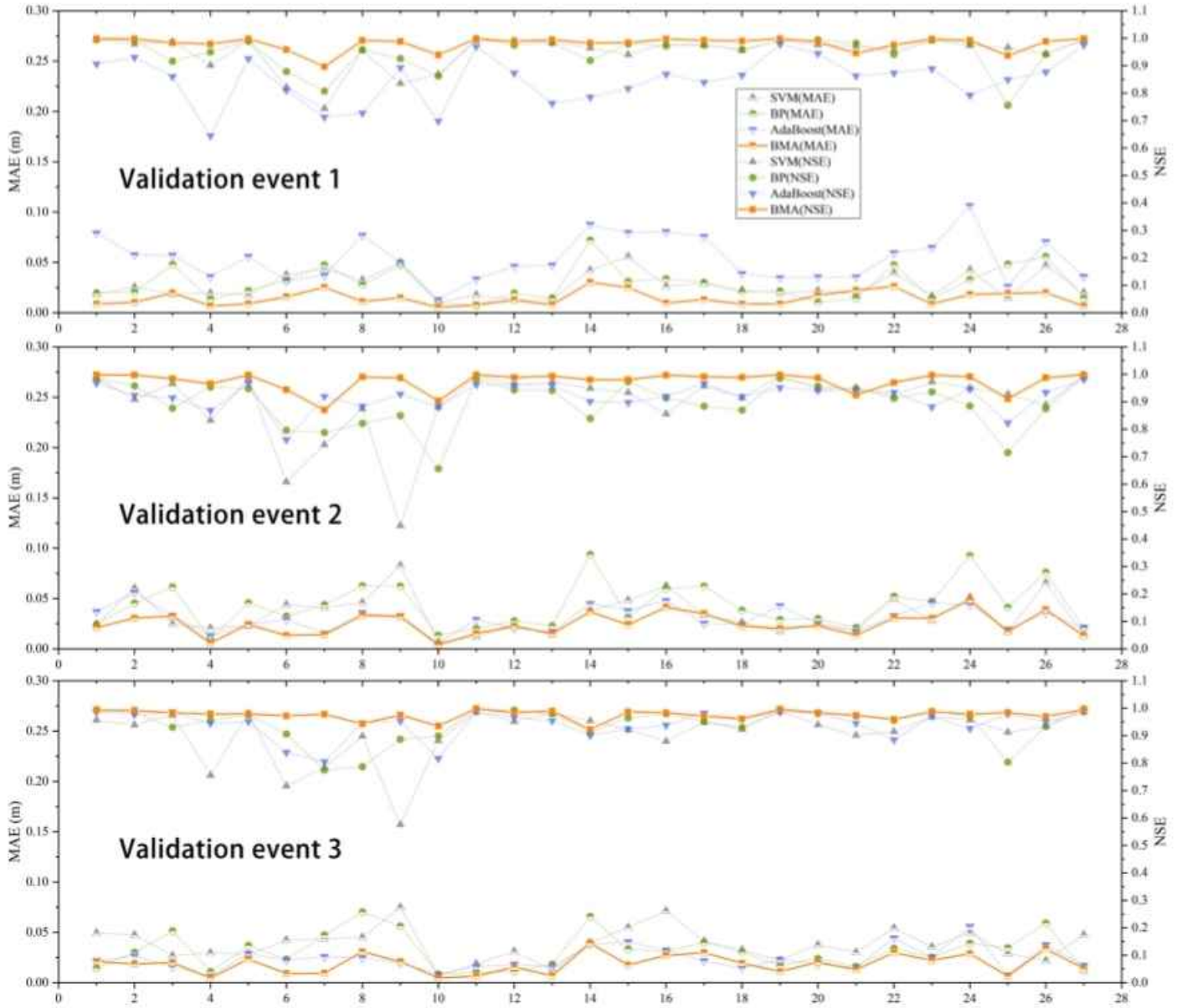


Fig. 7. The prediction performance of the BMA and single model in each waterlogging point.

the underpass tunnel, the rainfall converges rapidly after reaching the surface, resulting in the rapid increase of the water depth. And the water depth at the waterlogging point of number 16 is the smallest. The reason may be that the terrain at the road intersection is relatively flat. The lower surface confluence speed and the flat terrain make the water depth smaller. In addition, it is precisely because of this terrain that the water accumulation curve is relatively gentle. From the perspective of prediction performance, the prediction performance of single model shows great uncertainty. For example, the NSE of AdaBoost model at #16 and #23 waterlogging point is greater than 0.96, but at #23 waterlogging point is only 0.943. Similarly, the NSE of BPNN at #26 and #23 waterlogging point is greater than 0.96, but at No. 16 waterlogging point is only 0.924. In contrast, the prediction results of BMA model in three different types of waterlogging points are 0.984 (# 26), 0.971 (# 16) and 0.982 (# 32), showing high accuracy and stable performance. Therefore, it can be concluded that BMA model has stable prediction performance in the prediction of different types of waterlogging points, and can effectively improve the prediction effect of single model.

3.3. The inundation characteristics based on the BMA model

The rainfall data under different rainfall return periods with 180 min rainfall duration (once in half a year, once a year, once in two years, once in five years, once in ten years, once in twenty years and once in fifty years) (Wu et al., 2020a) were obtained using the classic Chicago rainstorm method and the rainstorm intensity formula in Zhengzhou City (Shao and Liu, 2018) (Eq. (12)), then the local urban flooding prediction results of BMA model under different rainfall return periods were obtained by inputting rainfall data into the model. The Polynomial fitting and cluster method were used to analyze the trend and characteristics of water depth under different rainfall return periods. Specifically, the increasing trend of water depth under different rainfall return periods was determined by using the Polynomial Fitting method, and the trend of 27 waterlogging points was clustered into four categories by cluster

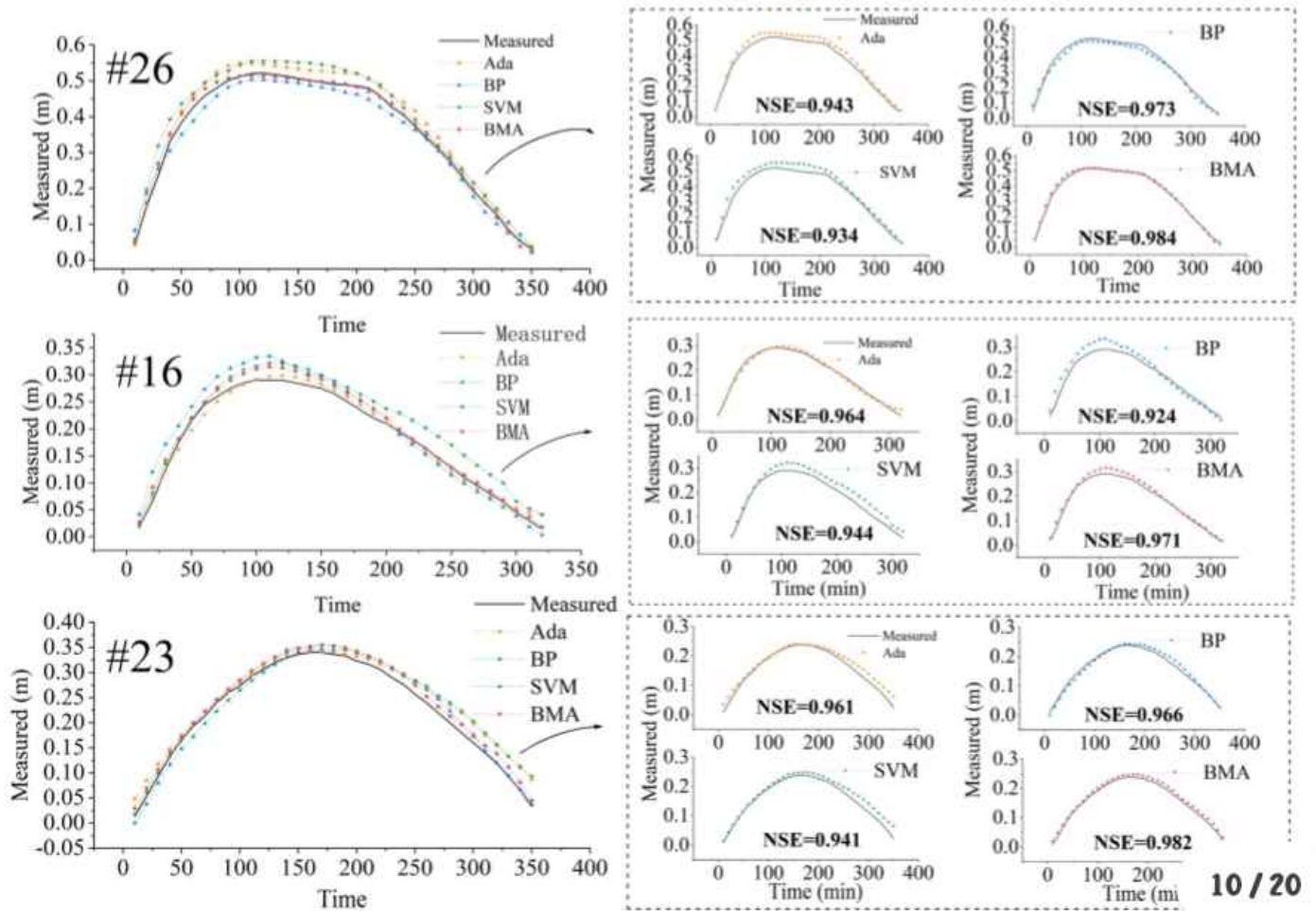


Fig. 8. The prediction performance in different types of waterlogging points.

analysis.

$$i = \frac{40.1(1 + 0.794)\lg P}{(t + 25.8)^{0.948}} \quad (12)$$

Where i is the intensity of rain, P refers to the rainfall return period, t is the duration of rain.

As shown in Table 1 and Fig. 9, including 12 waterlogging points in Category 2, accounting for 44.4% of the waterlogging points in this study. The inundation characteristics of these waterlogging points are that the water depth is approximately linear with rainfall, and there are 7 waterlogging points with the water depth of more than 0.3 m and 11 waterlogging points with the water depth of more than 0.2 m under the rainfall return period of 10a. In Category 3, including 5 waterlogging points, showing growth trend of gradually decrease in the water depth. The water depth exceeds 0.3 m in 10 years rainfall return period and 0.4 m in 50 years rainfall return period. However, the growth trend of water depth in Category 4 (including 4 waterlogging points) is gradually increasing, which is opposite to the inundation characteristics of Category 3. Under the rainfall return period of 50 years, the water depth is more than 0.4 m in all waterlogging points of Category 4 and the water depth is more than 0.6 m in the waterlogging point of No. 1 and No. 26, which is higher than other waterlogging points. On the contrary, the water depth of the waterlogging points in Category 1 is generally low. To be specific, there is almost no waterlogging in these waterlogging points under the rainfall return period of 0.5 year, 1 year and 2 years. The water depth increases gradually when the rainfall return period over 5 years. However, water depth still does not exceed

Table 1
The characteristics of four categories by cluster analysis.

Category	Number of waterlogging points	Ponding characteristics	Water depth range (m)		
			10a (79 mm)	20a (90 mm)	50a (104 mm)
1	6	Low water depth; Growth trend is gradually increasing	[0.02,0.15]	[0.04,0.20]	[0.06,0.27]
2	12	Water depth is approximately linear with rainfall	[0.19,0.41]	[0.22,0.50]	[0.27,0.61]
3	6	Growth trend of gradually decrease	[0.30,0.40]	[0.39,0.44]	[0.41,0.51]
4	4	High water depth; Growth trend is gradually increasing	[0.23,0.38]	[0.31,0.50]	[0.41,0.65]

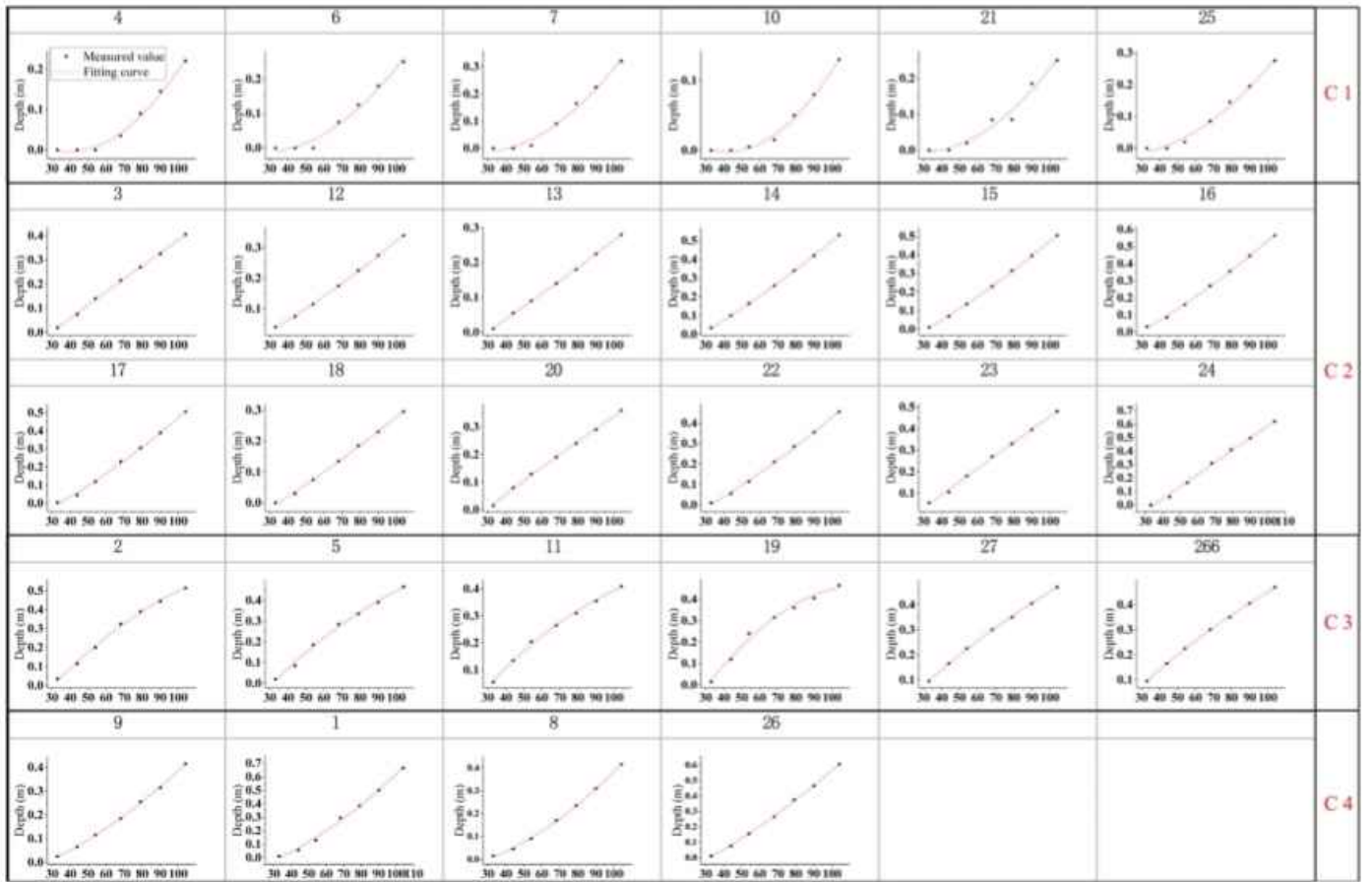


Fig. 9. Cluster analysis of inundation characteristics.

0.2 m under the rainfall return period of 10 years, and there is also only one waterlogging point with the water depth of over 0.3 m under the rainfall return period of 50 A.

On the whole, there are 74% waterlogging points have a water depth of more than 0.25 m when the rainfall return period over (or equal to) 10 years. However, there are only 37% waterlogging points have a water depth of more than 0.25 m when the rainfall return period did not exceed 5 years. In addition, no matter in waterlogging points of Category 2, Category 3 or Category 4, the water depth generally exceeds 0.25 m when the rainfall return period over 10 years, especially in the waterlogging points of Category 4, the growth trend of water depth is increasing with the increase of rainfall, which will threaten the traffic safety of vehicles and pedestrians. Therefore, in urban flood control, urban management departments should pay more attention to the rainfall events with a return period of more than (or equal to) 10 years and the increasing trend of water depth these waterlogging points. Specifically, many effective measures, including dredge rainwater wells before the flood season, using pumping pumps to improve drainage capacity and cutting off the road in case of serious waterlogging, should be taken to ensure the traffic safety of vehicles, the safety of pedestrians and reduce the losses caused by flood disasters.

3.4. Uncertainty analysis of results

Coverage rate (CR) and Bandwidth (B) were used to evaluate the uncertainty of BMA and single model. The CR represents the reliability of the model and it is the proportion of observed results falling within 95% confidence interval. The B (space between upper interval and lower interval) reflects uncertainty of the prediction results and it is the mean width across the inundation process (Liu and Merwade, 2018). The narrower the B, the higher the CR, indicating that the uncertainty of the model is lower. In addition, the

Table 2
Uncertainty analysis results of the BMA and single model.

Model	Validation event 1		Validation event 2		Validation event 3		Total	
	CR (%)	B (m)	CR (%)	B (m)	CR (%)	B (m)	CR (%)	B (m)
SVM	93.2	0.093	95.6	0.084	94.3	0.073	94.4	0.083
BPNN	95.9	0.077	94.7	0.081	96.3	0.082	95.6	0.080
AdaBoost	96.7	0.072	96.4	0.083	95.7	0.067	96.3	0.074
BMA	96.9	0.042	96.4	0.045	96.7	0.041	96.7	0.043

uncertainty of each stage may be significantly different in the prediction of inundation process. Therefore, in order to deeply analyze the uncertainty and accurately identify this difference, four typical waterlogging points were selected by stratified sampling method, which were #7 (C1), #16 (C2), #11 (C3) and #1 (C4).

As shown in Table 2, on the premise that the CR of BMA is not less than 96%, the B of BMA is reduced by 41.9–48.1% compared with the single model, indicating that BMA model has lower uncertainty. In addition, in the three different validation events, BMA model always has the highest CR and the narrowest B, and the CR of BMA is also more than 96%. These results demonstrate that BMA significantly reduced the uncertainty of the single model, which is consistent with the analysis results in Section 3.1.

Figs. 10–12 shows uncertainty of the SVM, BPNN, AdaBoost and BMA under different validation events. As shown in Fig. 10, the B of BMA model is not more than 0.05 m in the first validation event, which is 9.8–61.6% less than that of SVM, BPNN and AdaBoost, indicating that BMA reduces the uncertainty of the prediction results of single model. However, in the middle period of inundation process, although the CR of BMA is 100%, the B of BMA is higher than that in the beginning and end stages of inundation process, indicating that the BMA shows high uncertainty in the middle period of inundation process. In addition, it can be found that the B of BMA model (0.046 m) at the waterlogging point of Category 4 is 7.0–24.3% higher than that of other types of waterlogging points, indicating that the prediction results of BMA model show higher uncertainty at the waterlogging point of Category 4.

In the second validation event (Fig. 11), the B of BMA is between 0.042 m and 0.058 m, which is 25.9–56.7% less than that of single model, indicating that BMA show lower uncertainty than SVM, BPNN and AdaBoost under the 95% confidence interval. To be specific, the prediction results of BMA always have the lowest B in all types of waterlogging points, and the B of BMA across the inundation process is also lower than that of SVM, BPNN and AdaBoost. In addition, the uncertainty of BMA has the same characteristics as the verification event 1 in different stages of inundation process. In the beginning and end stages of inundation process, BMA has a narrow interval width, while in the middle period of inundation process, BMA has a wide interval width, indicating that BMA also shows higher uncertainty in the middle period of inundation process.

Similarly, the B of BMA model (between 0.038 m and 0.059 m) is 20.3–61.8% less than that of single model (Fig. 12), indicating that BMA significantly reduces the uncertainty of the single model prediction. From the characteristics of the interval width under the

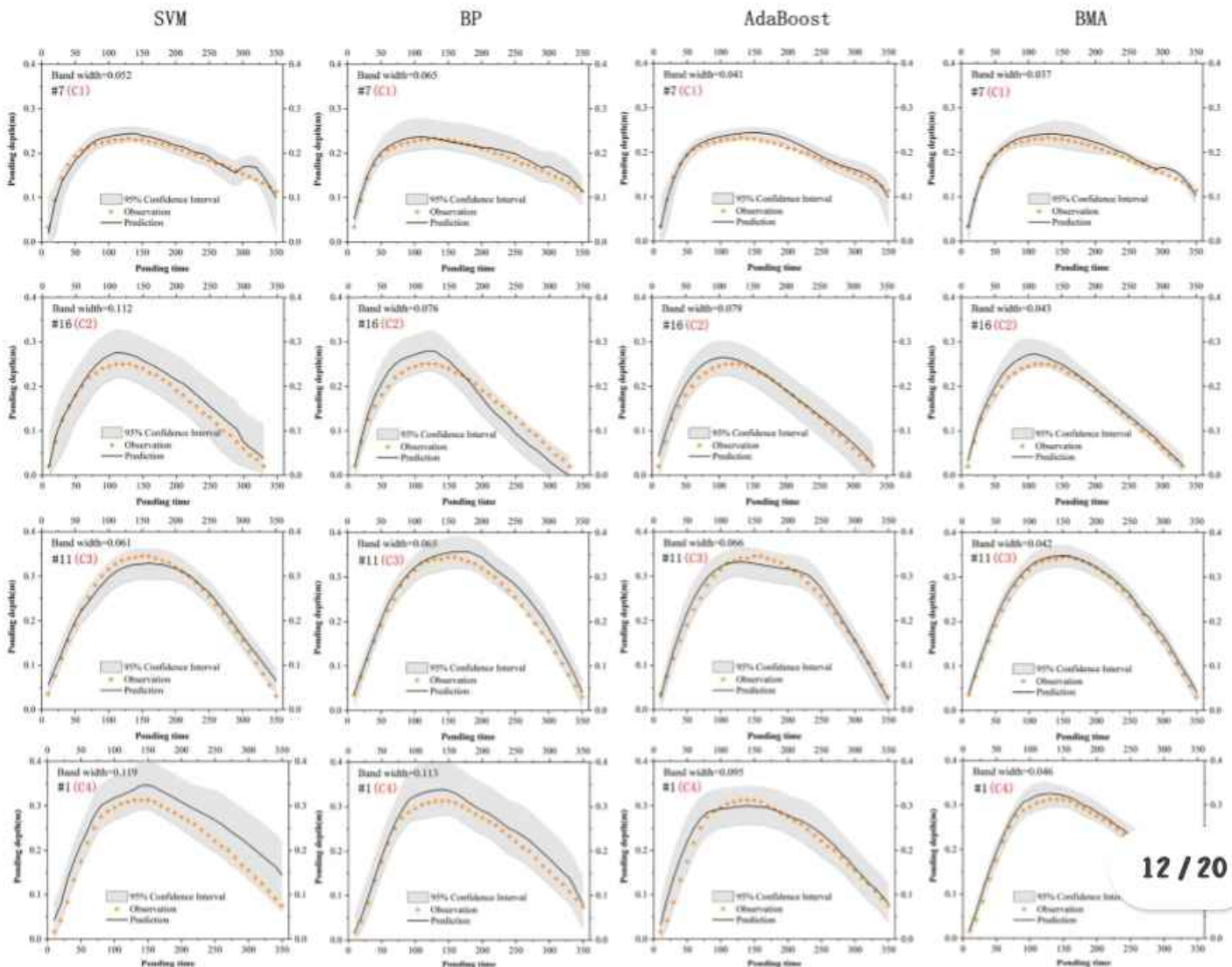


Fig. 10. The uncertainty results under 95% confidence interval in the first validation event (Part).

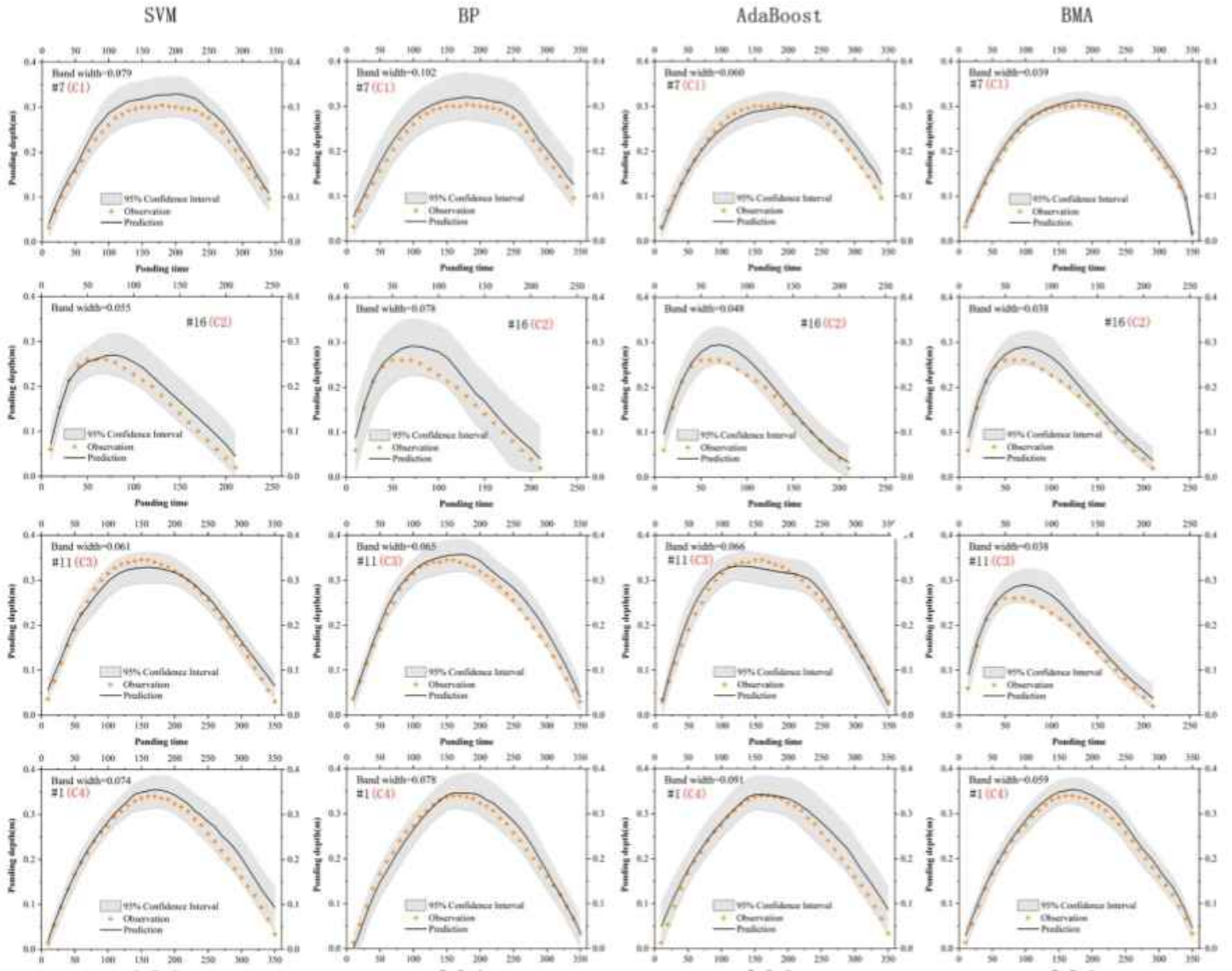


Fig. 12. The uncertainty results under 95% confidence interval in the third validation event (Part).

Table 3
Accuracy analysis of rainfall forecast data.

Lead forecast time (min)	August 1, 2019		August 7, 2020		Total MRE (%)
	MRE of rainfall process (%)	MRE of Precipitation (%)	MRE of rainfall process (%)	MRE of Precipitation (%)	
10	20.5	14.3	19.4	14.6	17.2
20	21.9	18.6	20.1	17.4	19.5
30	23.4	21.5	22.1	20.6	21.9
40	24.5	22.2	24.1	21.2	23.0
50	25.8	23.1	25.4	22.5	24.2
60	26.9	25.4	26.6	24.3	25.8
70	29.4	27.1	28.8	26.3	27.9
80	28.1	26.3	28.7	26.1	27.3
90	30.9	28.5	29.6	27.8	29.2
100	34.3	33.1	34.4	32.7	33.5
110	40.5	39.4	39.9	38.6	39.6
120	42.7	40.8	41.8	40.1	41.3

are large uncertainties in the accuracy of rainfall forecast data, the accuracy of rainfall forecast data gradually improves with the shortening of the forecast period. When the forecast period is shortened to 90 min, the Mean Relative Error (MRE) of rainfall forecast data is reduced to 29.2%, which can basically meet the accuracy requirements of input data of BMA model.

3.5.2. The early warning results based on BMA

By inputting the rainfall forecast data (with different lead forecast time) into the BMA model using Python 3.7, the inundation process results of each waterlogging point were obtained. The water depth was used to evaluate the urban flood disaster risk of each waterlogging point, which is divided into four classes: 0–3 cm, 3–10 cm, 10–25 cm, > 25 cm. On this basis, the early warning results of two rainfall events were obtained using GIS. For example, by inputting the rainfall forecast data with 100 min lead forecast time into the model, combined with the classification standard and GIS, the early warning results with 100 min lead forecast time were obtained. Similarly, with the shortening of lead forecast time, the rainfall forecast data with 90 min, 80 min, 70 min, 60 min, 50 min, 40 min, 30 min, 20 min, 10 min lead forecast time were continuously input into the model, the early warning results under different lead forecast time were obtained. It should be noted that because the running time of BMA model (5.6 s) hardly affects the forecast period of early warning, this study didn't consider the running time in the discussion of forecast time. As shown in Table 4, With the shortening of lead forecast time, the accuracy of early warning results is gradually improved. The accuracy of early warning results has increased from 74.0% (100 min lead forecast time) to 90.7% (10 min lead forecast time). The main reason for this phenomenon is that the accuracy of rainfall forecast data gradually improves with the shortening of lead forecast time, and the improvement of the accuracy of input data will improve the accuracy of the output result. Then the accuracy of early warning results will be improved. It is worth noting that when the lead forecast time is less than (or equal to) 80 min, the Precision and Recall of the BMA model are both exceeding 80%, which basically meets the accuracy requirements of early warning. What's more, the lead forecast time of 80 min is basically enough for relevant managers to make full response. Therefore, these demonstrate that this method has great application potential in the early warning and response of waterlogging.

In addition, Table 4 and Fig. 13 also shows the accuracy of BMA model at each early warning level. The performance of the early warning results gradually improved with the increase of early warning level. To be specific, the BMA model has low Precision in the early warning of level 1 (68.6%), but the Precision in the early warning of other levels is exceeding 88%. Similarly, the Recall of the BMA is from 80.1% to 91.4% with the improvement of early warning level. In addition, it is obvious that the prediction performance of the BMA model for the early warning level of level 4 is higher than the early warning level of other levels, indicating that BMA model is more suitable for the prediction of serious waterlogging.

Fig. 14 reflects the spatial characteristics of waterlogging. It is obvious that the areas with serious waterlogging were mainly concentrated in the middle and south of Zhengzhou City. In the middle area of Zhengzhou City, the waterlogging degree is obviously high since these areas are old urban areas with low pipe network design standards and aging pipe networks. The south area of Zhengzhou City is in the stage of rapid development and construction. The drainage system of these areas is still under construction. In addition, the new project will change the layout of the original drainage pipe network. Therefore, these areas are easier to produce serious waterlogging in heavy rainfall events. In addition, in the rainfall event on August 1, 2019, there were 15 waterlogging points with water accumulation degree of more than grade 3 (including grade 3). However, 11 of the 15 waterlogging points in the rainfall event on August 7, 2020 still have a water accumulation degree of more than grade 3, which indicates that although the distribution and inundation of the waterlogging points may be affected by the rainfall event, the underlying surface characteristics determine that the risk of waterlogging at these waterlogging points is greater. Therefore, urban flood control should pay more attention to these waterlogging points.

3.5.3. Comparison of early warning results between the BMA and single model

The OA, UR and OR were used to compare the early warning performance of SVM, BPNN, AdaBoost and the BMA model. As shown in Fig. 15, the OA of BMA was 90.7% OA, but the OA of three single models were no more than 82%, indicating that the BMA has the higher accuracy than the single model in early warning of local urban flooding. In addition, it can be found that the early warning results in SVM and BPNN have high OR (16.7% in SVM and 18.5% in BPNN) and low UR (1.9% in SVM and 5.6% in BPNN), indicating that the early warning results were mainly overestimated. In contrast, the early warning results in AdaBoost have high UR (13.0%) and low OR (9.3%), which shows that the early warning results were mainly underestimated. However, the UR (3.7%) and OR (5.6%) of the BMA model was not only close, but also relatively low. Therefore, the above results and analysis demonstrate that the BMA method can be used to improve the accuracy of urban flood prediction and early warning. Although the early warning results may cause slight overestimation.

Table 4
Accuracy analysis of early warning results by BMA.

Level	Accuracy	Lead forecast time (min)									
		10	20	30	40	50	60	70	80	90	100
1	Precision (%)	78.5	78.1	74.2	71.9	68.6	66.3	65.2	64.4	62.2	60.8
	Recall (%)	84.3	84.1	83.6	83.2	82.4	82.1	80.5	77.3	73.8	69.4
2	Precision (%)	95.1	94.2	94.8	93.1	94.4	92.9	88.7	83.4	77.9	72.8
	Recall (%)	92.5	91.6	90.9	90.2	89.4	88.2	86.1	84.7	81.9	79.3
3	Precision (%)	94.3	94.2	93.7	93.1	92.8	91.6	88.9	85.3	82.1	79.9
	Recall (%)	94.5	94.3	93.6	93.1	92.8	92.86	89.2	86.9	84.4	81.4
4	Precision (%)	94.9	94.3	94.6	94.2	93.7	93.75	91.4	87.2	85.1	82.5
	Recall (%)	95.1	94.2	94.5	94.1	93.7	93.7	90.4	88.9	85.7	83.8
Total	Precision (%)	90.7	90.2	89.3	88.1	87.4	86.1	83.5	80.1	76.8	74.0
	Recall (%)	91.6	91.05	90.65	90.15	89.6	89.2	86.6	84.5	81.5	78.5

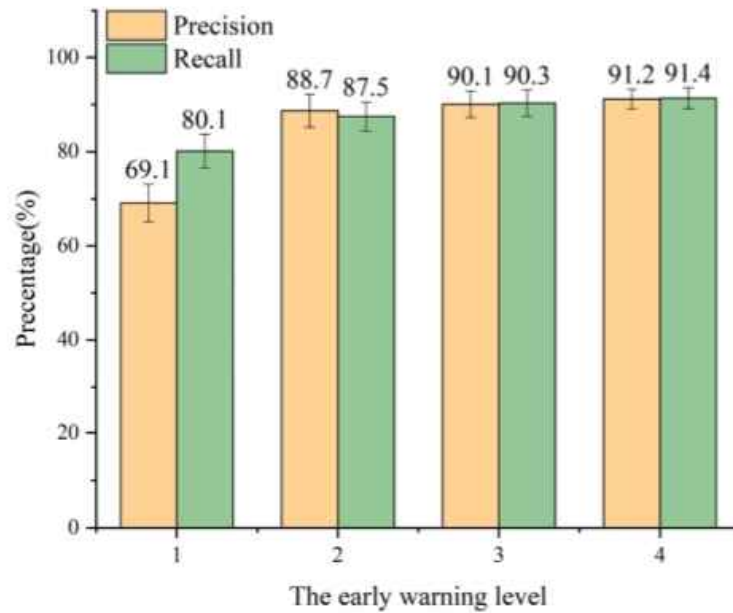


Fig. 13. The characteristics of early warning accuracy under different early warning level.

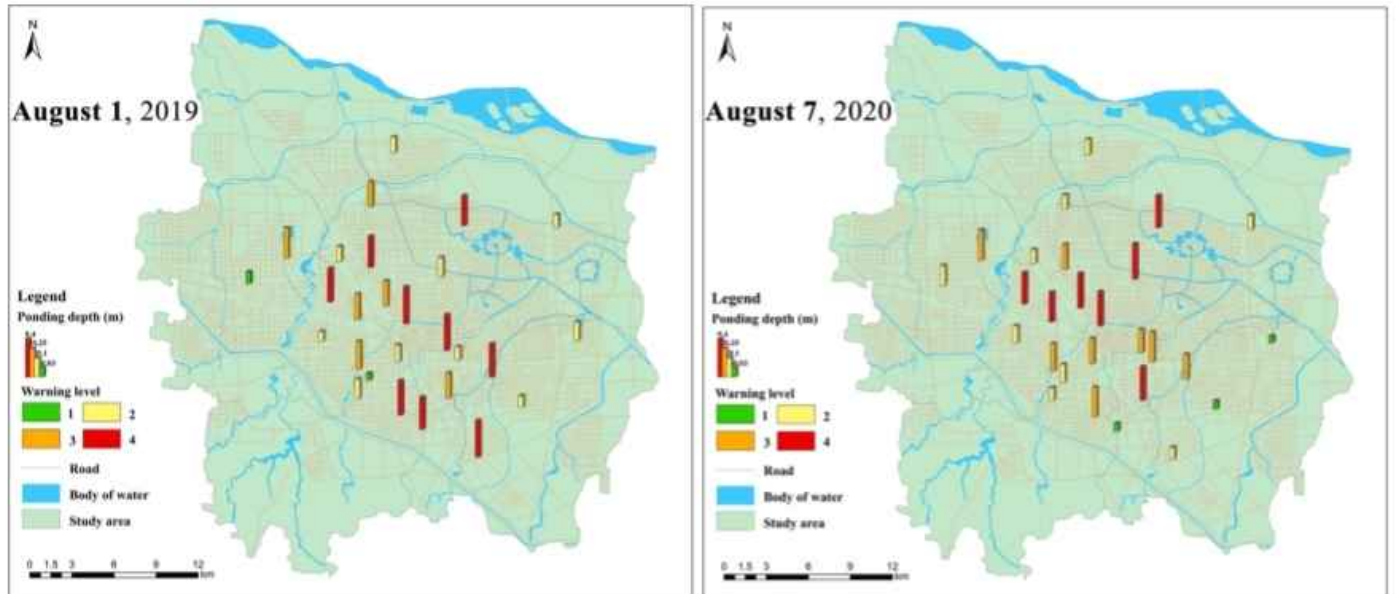


Fig. 14. Early warning results of two rainfall events using the BMA approach (10 min lead forecast).

4. Discussion

4.1. Influence of rainfall forecast data on early warning accuracy

It is generally accepted that the accuracy of rainfall forecast data decreases gradually with the extension of lead forecast time, which directly affect the early warning results of the BMA model under different lead forecast time. As shown in Table 4 and Fig. 13, under different lead forecast time, the accuracy of early warning gradually decreases with the extension of the forecast time, which is mainly due to the influence of the accuracy of rainfall forecast data. Nevertheless, the influence of rainfall forecast data on early warning accuracy is not clear. Therefore, in order to analyze the impact of rainfall forecast data on early warning accuracy, the quantitative impact of rainfall forecast data in different lead forecast time on early warning accuracy was analyzed by using the control variable method. Specifically, by inputting the measured rainfall data into the BMA model, the prediction results of inundation process were obtained. And combined with the early warning classification standard, the early warning results under actual rainfall data were obtained. These results were used as a control group to analyze the impact of the rainfall forecast data in different lead forecast time on the accuracy of early warning. As shown in Fig. 16, on the premise that the measured rainfall data was used as the input data (control group), the error of BMA model is 6.9%. And due to the influence of rainfall forecast data, the error of BMA model increased by 9.3–26%. In addition, the error caused by rainfall forecast data increased significantly after the lead forecast time exceeds 50 min.

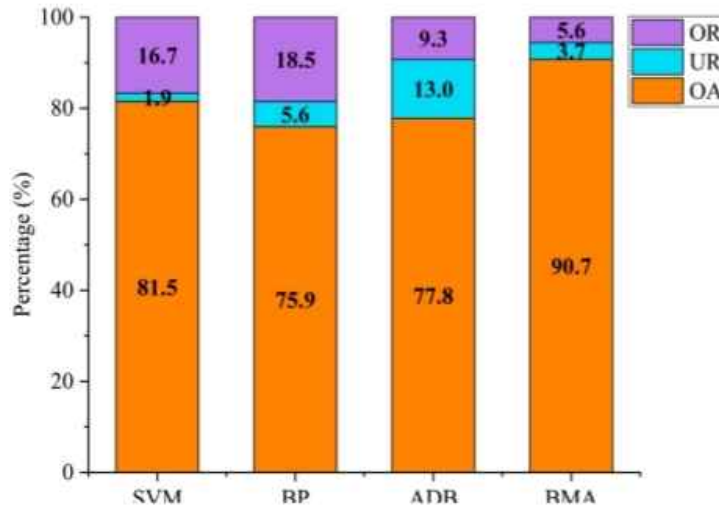


Fig. 15. Comparison of early warning accuracy between BMA and the single model.

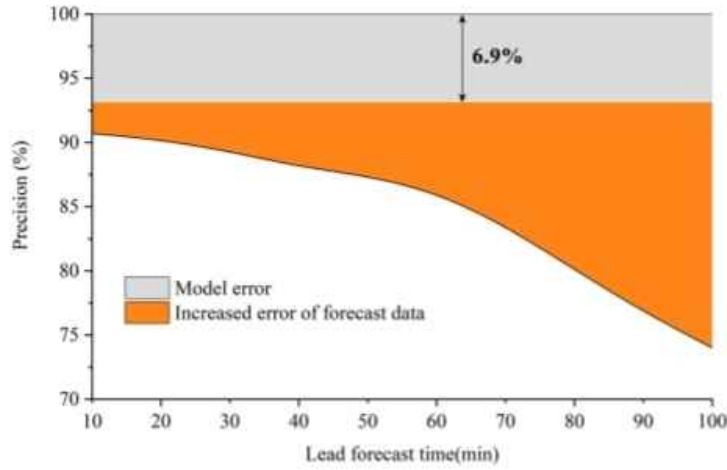


Fig. 16. The impact of rainfall forecast data on early warning accuracy.

Therefore, although the accuracy of the early warning results within 80 min lead forecast time can be accepted (Precision>80%), the early warning within 50 min lead forecast time is more reliable from the changing trend of early warning results.

4.2. Comparative analysis with physical model

Based on three classical machine learning model, this study uses BMA approach to couple these machine learning models to construct an integrated prediction model. In terms of prediction accuracy, the NSE of the prediction results of the BMA and machine learning models are more than 0.94. In terms of calculation efficiency, BMA model needs the average running time of 5.6 s. And in the early warning under different lead forecast time, the Precision of BMA model is more than 80% when the lead forecast time is no more than 80 min

Similarly, the physical model can also predict the inundation process. Our previous research has used SWMM model to simulate the flood events in Zhengzhou (Wang et al., 2022). The results demonstrate that the prediction results of SWMM model have an error of 16.7%. In addition, Qi et al. (2022) established an urban flood simulation model in Haikou City by using PCSWMM. The results show that the NSE of the model is 0.844. Therefore, the physical model can also achieve accurate urban flood prediction, and there is no significant difference in accuracy compared with the data-driven model. In terms of the calculation efficiency, the average calculation time of SWMM model under three rainfall events (20170706, 20160605 and 20160805) was 1.8 min, and the calculation time under the rainfall events on July 20, 2021 was 10.1 min. Recent approaches show that the one-dimensional surface and drainage pipe network model usually needs computation time, ranging from 5 min to 1 h (Bulti and Abebe, 2020), and accurate two-dimensional hydraulic model usually needs the running time of more than one hour (Bulti and Abebe, 2020; Chang et al., 2015). The main reason for this difference in calculation efficiency is the complexity of physical model and the duration of rainfall event. Therefore, it can be concluded that the approach proposed in this study is superior to the physical model in terms of calculation efficiency. However, it should be pointed out that the approach proposed in this study is only applicable to the area with sufficient historical urban flood data. In other words, the data-driven method is difficult to be directly applied in the area lacking data. In the area lacking urban flood

data, this approach can be used as a proxy model of physical model to improve the calculation efficiency. Recent researches (Kwon and Kim, 2021; Mignot and Dewals, 2022) demonstrate that data-driven method can significantly improve the computational efficiency of physical model on the premise of ensuring the accuracy.

In addition, the physical model can also be used as the benchmark model of BMA. Introducing physical model into BMA model for inundation process prediction and early warning needs to meet two basic conditions. One condition is that the physical model can accurately identify the location of waterlogging points and accurately simulate the inundation process of urban flood. Another condition is that the calculation time of physical model should not be too long. Therefore, future research can focus on how to introduce physical model into BMA model on the premise of ensuring accuracy and computational efficiency.

4.3. The generalization ability of BMA model

On July 20, 2021, an extreme rainstorm occurred in Zhengzhou, with a total rainfall of 624.1 mm within 24 h. By inputting the rainfall data into the BMA model, the inundation process was obtained, which aims to verify the applicability of the BMA model under the extreme rainfall event. As shown in Fig. 17, the trend of inundation process predicted by BMA model is consistent with the measured process. However, the maximum depth of inundation process is underestimated, which is 26.1% lower than the actual value. And the predicted inundation process is 1 h earlier than the measured inundation process, the peak time is 1 h later than the actual inundation process, the end time is 1 h earlier than the actual inundation process. The one reason for this error is that the BMA model does not contain sufficient sample data. Specifically, although the return period of rainfall events in the sample data set has reached once in a hundred years, the return period of the rainfall event on July 20, 2021 has reached once in a thousand years, which seriously exceeds the intensity of all rainfall events in the sample data set. Therefore, due to the limitation of sample data set, the accuracy of BMA model in predicting extreme rainfall events has decreased. An important way to solve this problem is to continuously enrich the sample data set to improve the generalization ability of the model. Based on this, this study collected the data of the extreme rainfall event in Zhengzhou on July 20, 2021 and imported it into the sample data set, which aims to improve the generalization ability of BMA model under extreme rainfall events. In addition, another reason for the decrease in the model prediction accuracy is that the hydraulic connection at the waterlogging point has changed significantly. To be specific, the increase of rainfall may increase the catchment area greatly, and the basin flood will also effects on urban flood. Therefore, in the extreme rainfall event, basin flood and urban flood should be considered comprehensively, which is also one of the important development directions of urban flood research in the future.

5. Conclusion

BMA is a multi-model weighted average simulation method integrating the results of various models. In this paper, the BMA was applied to the prediction and early warning of urban flood. The main conclusions are as follows:

(1) In the prediction of local urban flooding, BMA can improve the accuracy and stability of the prediction results. Compare with three single models, the BMA algorithm has the lowest MAE (0.019) and the highest NSE (0.984) in the prediction of local urban flooding. In addition, although BMA does not always have the highest accuracy in all rainfall events and all waterlogging points, the overall predictive performance and stability of the BMA model is better than the single model in different rainfall events and waterlogging points, which demonstrates that BMA can provide a more accurate and stable method for urban flood prediction.

(2) Analysis indicates that urban management departments should pay more attention to the rainfall event with the return period of more than (or equal to) 10 years and the waterlogging points of Category 2, Category 3 or Category 4. There are 74% waterlogging points have a depth of more than 0.25 m when the rainfall return period over (or equal to) 10 years. Especially in the waterlogging points of Category 4, the average depth is 0.34 m under the rainfall return period of 10 years, and the water depth shows an exponential growth trend with the increase of rainfall.

(3) Among the three verification events of different types, BMA always has the highest CR and the narrowest B. In different stages of

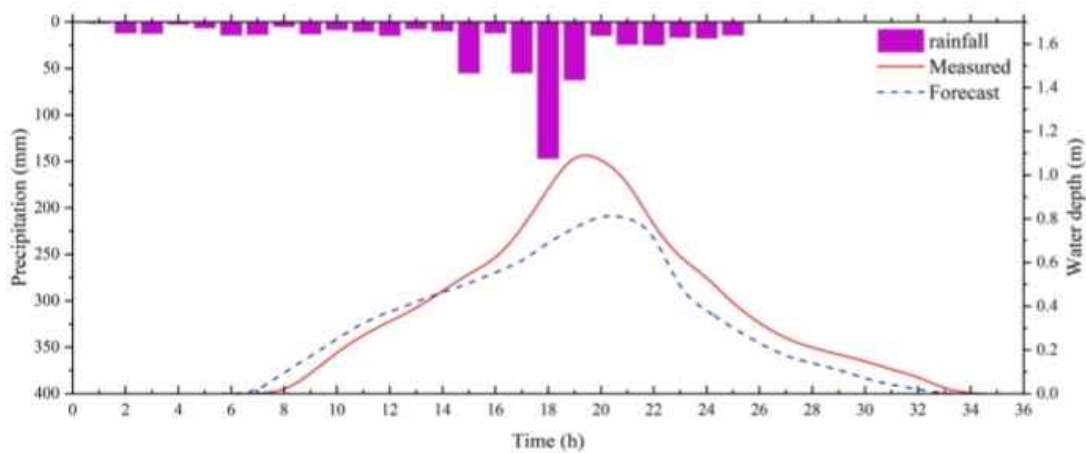


Fig. 17. The verification results of rainfall event on July 20, 2021.

local urban flooding, BMA has lower uncertainty in the beginning and end stages, but in the middle period BMA has higher uncertainty. Nevertheless, the BMA method has reduced the uncertainty of local urban flooding prediction to a certain extent.

(4) In the fine early warning of local urban flooding using rainfall forecast data, the early warning result of BMA model has the overall accuracy of 90.7%, which is significantly higher than that of SVM (81.5%), BPNN (75.9%) and AdaBoost (77.8%). It shows that BMA can provide a more accurate and stable method for fine early warning of urban flood.

However, there are still some directions for further exploration in this study. For instance, although BMA provides a more stable and accurate local urban flooding prediction and early warning method, this approach can only predict urban flood in the waterlogging points with detailed observation data. Therefore, future research can combine hydrological model, data-driven methods and the BMA to explore comprehensive and refined urban flood prediction and early warning methods.

CRediT authorship contribution statement

Yihong Zhou: Data curation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Zening Wu:** Methodology, Validation, Writing – review & editing, Supervision, Funding acquisition. **Hongshi Xu:** Methodology. **Huiliang Wang:** Methodology, Supervision, Funding acquisition.

Declaration of Competing Interest

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

Data Availability

The data that has been used is confidential.

Acknowledgements

The research was funded by the Key Project of National Natural Science Foundation of China (No: 51739009), Excellent Youth Fund of Henan Province of China (212300410088), Science and Technology Innovation Talents Project of Henan Education Department of China (21HASTT011), Young backbone Teachers Training Fund of Henan Education Department of China (2020GGJS005), National Natural Science Foundation of China (52109040) and China Postdoctoral Science Foundation (2021M702950). We thank the meteorological department and administration bureau of Zhengzhou city, for providing rainfall and waterlogging data. We also thank the anonymous reviewers for their valuable comments.

References

- Ajami, N.K., Qingyun, D., Xiaogang, G., et al., 2006. Multimodel combination techniques for analysis of hydrological simulations: application to distributed model intercomparison project results. *J. Hydrometeorol.* 7 (4), 755–768. <https://doi.org/10.1175/JHM519.1>.
- Alfieri, L., Salamon, P., Bianchi, A., et al., 2014. Advances in pan-European flood hazard mapping. *Hydrol. Process.* 28 (13), 4067–4077. <https://doi.org/10.1002/hyp.9947>.
- Baran, S., Möller, A., 2015. Joint probabilistic forecasting of wind speed and temperature using Bayesian model averaging. *Environmetrics* 26 (2), 120–132. <https://doi.org/10.1002/env.2316>.
- Basher, A., Islam, A.K.M.S., Stiller-Reeve, M.A., et al., 2020. Changes in future rainfall extremes over Northeast Bangladesh: a Bayesian model averaging approach. *Int. J. Climatol.* 40 (6), 3232–3249. <https://doi.org/10.1002/joc.6394>.
- Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *J. Hydrol.* 387 (1/2), 33–45. <https://doi.org/10.1016/j.jhydrol.2010.03.027>.
- Bult, D.T., Abebe, B.G., 2020. A review of flood modeling methods for urban pluvial flood application. *Model. Earth Syst. Environ.* 6 (3), 1293–1302. <https://doi.org/10.1007/s40808-020-00803-z>.
- Chang, T.-J., Wang, C.-H., Chen, A.S., 2015. A novel approach to model dynamic flow interactions between storm sewer system and overland surface for different land covers in urban areas. *J. Hydrol.* 524, 662–679. <https://doi.org/10.1016/j.jhydrol.2015.03.014>.
- Chen, C.-S., Chen, B.P.-T., Chou, F.N.-F., et al., 2010. Development and application of a decision group Back-Propagation Neural Network for flood forecasting. *J. Hydrol.* 385 (1–4), 173–182. <https://doi.org/10.1016/j.jhydrol.2010.02.019>.
- Chowdhury, S., Sharma, A., 2009. Multisite seasonal forecast of arid river flows using a dynamic model combination approach. *Water Resour. Res.* 45 (10) <https://doi.org/10.1029/2008WR007510>.
- Dimitriadis, P., Tegos, A., Oikonomou, A., et al., 2016. Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. *J. Hydrol.* 534, 478–492. <https://doi.org/10.1016/j.jhydrol.2016.01.020>.
- Faceli, K., Lorena, A.C., Gama, J. a., Carvalho, A., 2011. Inteligência Artificial: Uma abordagem de aprendizado de máquina. 2: 192. DOI:Merwade, Venkatesh/0000-0001-5518-2890; Olivera, Francisco/0000-0003-0919-6815.
- Fernández, C., Ley, E., Steel, M.F.J., 2001. Benchmark priors for Bayesian model averaging. *J. Econ.* 100 (2), 381–427. [https://doi.org/10.1016/S0304-4076\(00\)00076-2](https://doi.org/10.1016/S0304-4076(00)00076-2).
- Hammersley, J.M., Handscomb, D.C., 1964. *Monte Carlo Methods*. Methuen, London, pp. 76–84.
- Hou, J., Zhou, N., Chen, G., et al., 2021. Rapid forecasting of urban flood inundation using multiple machine learning models. *Nat. Hazards* 108 (2), 2335–2356. <https://doi.org/10.1007/s11069-021-04782-x>.
- Huo, W., Li, Z., Wang, J., et al., 2019. Multiple hydrological models comparison and an improved Bayesian model averaging approach for ensemble prediction over semi-humid regions. *Stoch. Environ. Res. Risk Assess.* 33 (1), 217–238. <https://doi.org/10.1007/s00477-018-1600-7>.
- Kwon, S.H., Kim, J.H., 2021. Machine learning and urban drainage systems: State-of-the-art review. *Water (Switz.)* 13 (24). <https://doi.org/10.3390/w13243545>.
- Li, W., Zhou, J.Z., Sun, H.W., et al., 2017. Impact of distribution type in bayes probability flood forecasting. *Water Resour. Manag.* 31 (3), 961–977. <https://doi.org/10.1007/s11269-016-1557-6>.

- Liu, Z., Merwade, V., 2018. Accounting for model structure, parameter and input forcing uncertainty in flood inundation modeling using Bayesian model averaging. *J. Hydrol.* 565, 138–149. <https://doi.org/10.1016/j.jhydrol.2018.08.009>.
- Liu, Z., Merwade, V., 2019. Separation and prioritization of uncertainty sources in a raster based flood inundation model using hierarchical Bayesian model averaging. *J. Hydrol.* 578, 124100 <https://doi.org/10.1016/j.jhydrol.2019.124100>.
- Madadgar, S., Moradkhani, H., 2014. Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging. *Water Resour. Res.* 50 (12), 9586–9603. <https://doi.org/10.1002/2014WR015965>.
- Mignot, E., Dewals, B., 2022. Hydraulic modelling of inland urban flooding: recent advances. *J. Hydrol.*, 127763 <https://doi.org/10.1016/j.jhydrol.2022.127763>.
- Najafi, M.R., Moradkhani, H., 2015. Multi-model ensemble analysis of runoff extremes for climate change impact assessments. *J. Hydrol.* 525, 352–361. <https://doi.org/10.1016/j.jhydrol.2015.03.045>.
- Najafi, M.R., Moradkhani, H., 2016. Ensemble combination of seasonal streamflow forecasts. *J. Hydrol. Eng.* 21 (1), 1–13. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001250](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001250).
- Qi, W., Ma, C., Xu, H., et al., 2022. Urban flood response analysis for designed rainstorms with different characteristics based on a tracer-aided modeling simulation. *J. Clean. Prod.* 355, 131797 <https://doi.org/10.1016/j.jclepro.2022.131797>.
- Qu, B., Zhang, X., Pappenberger, F., et al., 2017. Multi-model grand ensemble hydrologic forecasting in the fu river basin using bayesian model averaging. *Water* 9 (2). <https://doi.org/10.3390/w9020074>.
- Raftery, A.E., Gneiting, T., Balabdaoui, F., et al., 2005. Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* 133 (5), 1155–1174. <https://doi.org/10.1175/MWR2906.1>.
- Rings, J., Vrugt, J.A., Schoups, G., et al., 2012. Bayesian model averaging using particle filtering and Gaussian mixture modeling: theory, concepts, and simulation experiments. *Water Resour. Res.* 48 (5) <https://doi.org/10.1029/2011WR011607>.
- Shao, D., Liu, G.-s., 2018. Up-to-date urban rainstorm intensity formulas considering spatial diversity in China. *Environ. Earth Sci.* 77 (14), 541. <https://doi.org/10.1007/s12665-018-7718-6>.
- Sundaram, S., Devaraj, S., Yarrakula, K., 2021. Modeling, mapping and analysis of urban floods in India—a review on geospatial methodologies. *Environ. Sci. Pollut. Res.* 28 (48), 67940–67956. <https://doi.org/10.1007/s11356-021-16747-5>.
- Taherkhani, A., Cosma, G., McGinnity, T.M., 2020. AdaBoost-CNN: an adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning. *Neurocomputing* 404, 351–366. <https://doi.org/10.1016/j.neucom.2020.03.064>.
- Wang, H., Hu, Y., Guo, Y., et al., 2022. Urban flood forecasting based on the coupling of numerical weather model and stormwater model: a case study of Zhengzhou city. *J. Hydrol. Reg. Stud.* 39, 100985 <https://doi.org/10.1016/j.ejrh.2021.100985>.
- Wu, C.F.J., 1983. On the convergence properties of the EM Algorithm. *Ann. Stat.* 11 (1), 95–103. <https://doi.org/10.1214/aos/1176346060>.
- Wu, M., Wu, Z., Ge, W., et al., 2021. Identification of sensitivity indicators of urban rainstorm flood disasters: a case study in China. *J. Hydrol.* 599, 126393 <https://doi.org/10.1016/j.jhydrol.2021.126393>.
- Wu, Z., Zhou, Y., Wang, H., et al., 2020a. Depth prediction of urban flood under different rainfall return periods based on deep learning and data warehouse. *Sci. Total Environ.* 716, 137077 <https://doi.org/10.1016/j.scitotenv.2020.137077>.
- Wu, Z., Zhou, Y., Wang, H., 2020b. Real-time prediction of the water accumulation process of urban stormy accumulation points based on deep learning. *IEEE Access* 8, 151938–151951. <https://doi.org/10.1109/ACCESS.2020.3017277>.
- Yan, H., Moradkhani, H., 2016. Toward more robust extreme flood prediction by Bayesian hierarchical and multimodeling. *Nat. Hazards* 81 (1), 203–225. <https://doi.org/10.1007/s11069-015-2070-6>.
- Yan, H.X., Moradkhani, H., 2015. A regional Bayesian hierarchical model for flood frequency analysis. *Stoch. Environ. Res Risk Assess.* 29 (3), 1019–1036. <https://doi.org/10.1007/s00477-014-0975-3>.
- Yang, Y., Sun, H., Xue, J., et al., 2021. Estimating evapotranspiration by coupling Bayesian model averaging methods with machine learning algorithms. *Environ. Monit. Assess.* 193 (3), 156. <https://doi.org/10.1007/s10661-021-08934-1>.
- Yu, J., Lee, H., Kim, M.-S., et al., 2008. Traffic flooding attack detection with SNMP MIB using SVM. *Comput. Commun.* 31 (17), 4212–4219. <https://doi.org/10.1016/j.comcom.2008.09.018>.
- Zheng, Z., Gao, J., Ma, Z., et al., 2016. Urban flooding in China: main causes and policy recommendations. *Hydrol. Process.* 30 (7), 1149–1152. <https://doi.org/10.1002/hyp.10717>.

A large, teal-colored cloud shape with a white outline, centered on the page. The words "THANK YOU" are written in bold black capital letters inside the cloud.

THANK YOU

Several realistic water droplets of various sizes are scattered around the central cloud. Some are in the top left, some in the top right, and a cluster of larger ones is in the bottom right. The background is a light gray gradient.