



Article

Urban Flood Risk Assessment through the Integration of Natural and Human Resilience Based on Machine Learning Models

Wenting Zhang ^{1,2,*}, Bin Hu ¹, Yongzhi Liu ^{3,4}, Xingnan Zhang ¹ and Zhixuan Li ¹¹ College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China² The State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China³ Hydrology and Water Resources Department, Nanjing Hydraulic Research Institute, Nanjing 210029, China⁴ The State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, NHRI, Nanjing 210029, China

* Correspondence: zwt@hhu.edu.cn

Abstract: Flood risk assessment and mapping are considered essential tools for the improvement of flood management. This research aims to construct a more comprehensive flood assessment framework by emphasizing factors related to human resilience and integrating them with meteorological and geographical factors. Moreover, two ensemble learning models, namely voting and stacking, which utilize heterogeneous learners, were employed in this study, and their prediction performance was compared with that of traditional machine learning models, including support vector machine, random forest, multilayer perceptron, and gradient boosting decision tree. The six models were trained and tested using a sample database constructed from historical flood events in Hefei, China. The results demonstrated the following findings: (1) the RF model exhibited the highest accuracy, while the SVR model underestimated the extent of extremely high-risk areas. The stacking model underestimated the extent of very-high-risk areas. It should be noted that the prediction results of ensemble learning methods may not be superior to those of the base models upon which they are built. (2) The predicted high-risk and very-high-risk areas within the study area are predominantly clustered in low-lying regions along the rivers, aligning with the distribution of hazardous areas observed in historical inundation events. (3) It is worth noting that the factor of distance to pumping stations has the second most significant driving influence after the DEM (Digital Elevation Model). This underscores the importance of considering human resilience factors. This study expands the empirical evidence for the ability of machine learning methods to be employed in flood risk assessment and deepens our understanding of the potential mechanisms of human resilience in influencing urban flood risk.

Keywords: flood risk assessment; ensemble learning model; urban flood; remote sensing data; machine learning



Citation: Zhang, W.; Hu, B.; Liu, Y.; Zhang, X.; Li, Z. Urban Flood Risk Assessment through the Integration of Natural and Human Resilience Based on Machine Learning Models. *Remote Sens.* **2023**, *15*, 3678. <https://doi.org/10.3390/rs15143678>

Academic Editors: Venkatesh Merwade, Adnan Rajib and Zhu Liu

Received: 31 May 2023

Revised: 13 July 2023

Accepted: 19 July 2023

Published: 23 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As the most common natural disaster, floods cause a large number of casualties and economic losses every year [1]. With urbanization and climate change, an increasing number of cities are affected by flood disasters [2,3]. The large-scale construction of houses and paving of roads in the process of urbanization have led to a significant increase in surface imperviousness, a decrease in infiltration, and a continuous increase in urban runoff, which increases the load on drainage facilities [4,5]. At the same time, climate change has increased the frequency of extreme weather events. In its 2021 report, the Intergovernmental Panel on Climate Change (IPCC) stated that climate change is effecting global weather extremes and extreme climate events, causing an increase in the intensity

and frequency of extreme precipitation in regions such as East Asia, Southeast Asia, and South Asia. The serious situation of flood disasters has prompted people to take technical and non-technical disaster prevention measures to build a flood protection system for cities [6,7].

In recent years, non-engineering measures represented by flood risk assessment have gradually ascended as the dominant approach to urban flood control [8,9]. The primary methods of flood risk assessment are historical disaster mathematical and statistical methods, multi-criteria decision analysis, remote sensing image analysis, scenario simulation analysis, and machine learning methods [10,11]. The historical disaster mathematical and statistical methods refer to the collection of historical flood event disaster data and then the analysis of these data using mathematical and statistical methods [12–15]. The multi-criteria decision analysis method is used to evaluate flood risk in the study area by constructing a system of flood risk assessment indicators and applying methods such as hierarchical analysis and the fuzzy comprehensive evaluation method [16–19]. This method can visually reflect the relationship between each indicator and flood risk. Most of the current indicator weights are calculated based on expert knowledge and experience [20]. The remote sensing image analysis method uses remote sensing technology to obtain information on the inundation extent, inundation duration, and affected bodies in the disaster area and then uses GIS and other tools to spatially analyze this remote sensing information [21–24]. The scenario simulation analysis method, for different scenarios, uses hydrodynamic models to simulate possible disaster events [25–27], and risk assessment is performed based on the simulation results. With the continuous advances in artificial intelligence, remote sensing, and computer technology, machine learning methods have started to be applied in flood risk assessment [28] and provide a superior performance and more cost-effective solution for flood disaster prediction [29].

However, the historical disaster mathematical and statistical method necessitates detailed historical data, which limits its flexibility for risk assessment in rapidly changing urban areas. The remote sensing image analysis method may inaccurately capture flood dynamics due to constraints in temporal and spatial resolution, particularly for small-scale incidents. The multi-criteria decision analysis method heavily relies on expert knowledge, resulting in subjectivity and uncertainty in the evaluation outcomes. The scenario simulation analysis method requires a substantial amount of high-resolution geographical, hydrological, and artificial facility data. The modeling process is complex and entails significant computational resources [30].

Compared to traditional models mentioned above, the machine learning methods exhibit higher performance and less complexity [31]. They offer notable advantages, including: (1) the rapid extraction of features and information from extensive datasets, (2) the utilization of interdisciplinary techniques for processing large amounts of data from multiple sources, and (3) high speed in generating predictions, making them highly promising for real-time flood modeling and risk prediction.

Thus far, numerous attempts have been made to apply machine learning models for the purpose of flood risk assessment and zoning in both watersheds and urban areas. Tehrany used a decision tree model for flood risk assessment in Kelantan, Malaysia [32]. Mojaddadi combined frequency ratios with support vector machines for flood risk analysis in the Baisalot River Basin in Malaysia [33]. Tehrany improved the support vector machine model using weight of evidence (WoE) to improve the accuracy of flood risk assessment [34]. Pham used a method combining a deep learning network and hierarchical analysis to map regional flood risk more accurately [35]. Wang used the random forest model for flood risk assessment and used the support vector machine for comparison [36]. Zhao used a semi-supervised support vector machine model to address the sparse sample size, which led to some improvement in the accuracy of the prediction results [37]. Zhao and Wang used a convolutional neural network for flood risk assessment, considering the influence of the surrounding environment, and achieved superior results compared to traditional machine learning methods [38,39].

The main differences between this study and the previous published works are mainly reflected in the following two aspects: Firstly, the majority of flood risk assessment research has predominantly focused on meteorological, hydrological, and geographical environmental factors from a natural perspective, with some literature also considering the impact of social vulnerability [40,41]. However, the impact of human resilience factors, such as urban flood control measures, has received limited attention, despite their undeniable relevance to urban flood risk. Therefore, this research aims to address this gap by considering factors related to human resilience and integrating them with meteorological and geographical factors, thus constructing a more comprehensive flood assessment framework.

Secondly, ensemble learning models, as machine learning models with superior algorithms, have gradually started to be applied in the assessment of various natural disasters. However, their applicability and generalization ability in urban flood assessment have not been fully explored [42]. The authors will introduce and compare ensemble learning models based on heterogeneous learners with the predictive performance of traditional machine learning models.

Hence, the characteristics of the research domain and the availability of data were taken into consideration in this study. Factors related to urban flooding from three perspectives, namely, natural geography, meteorological hydrology, and human resilience, were selected. The assessment of flood risk in the research area was conducted using multiple single machine learning models and ensemble learning models. The training and testing datasets for these models consisted of historical flood inundation hotspots. After optimizing the hyperparameters of these models, predictions were made regarding the spatial distribution of flood risk within the study area. The applicability of the different models was evaluated by considering their accuracy and their alignment with the historical inundation areas, and the underlying mechanisms between urban flood risk and its driving factors were determined. The research outcomes can provide valuable references for flood management in cities with similar geographical environments and levels of urbanization.

2. Study Area and Materials

2.1. Study Area

The study area is located in the central district of Hefei City, Anhui Province, China, covering a total area of 514.37 km². Geographically, it lies between a longitude of 116°40' to 117°52' east and latitude of 31°30' to 32°32' north. The underlying terrain in this region is predominantly hilly, with higher elevations in the northwest and lower elevations in the southeast. There are plains with a relatively flat topography along rivers and lakes, while certain areas exhibit the presence of hills. The ground elevation ranges from approximately 12 to 45 m, with a few low-lying areas adjacent to the rivers, measuring around 10 to 12 m. Hefei City is intersected by numerous rivers, including the Nanfei River, Shiwuli River, and Tangxi River, flowing from west to southeast and ultimately converging with Chao Lake, as shown in Figure 1. Hefei is in a subtropical monsoon humid climate zone, with an average annual precipitation of 966 mm. Due to its location within a transitional zone between humid and sub-humid regions, its precipitation distribution is uneven and influenced by its topography and water vapor sources. The summer months (June to August) receive the highest precipitation, accounting for 41.3% of the annual total. Historically, Hefei has been susceptible to frequent flood disasters, and in recent years, rapid urban development has further increased the potential for flood and waterlogging incidents in the area.

According to reports and data from the Water Conservancy Department, two short-duration heavy rainstorms took place on 29 June and 18 July 2010, setting a record for short-duration rainfall intensity and causing more than 30 waterlogged spots in the Hefei urban area.

Furthermore, on 20 August 2012, a heavy rainstorm occurred in the southwest of the urban area, with a maximum hourly rainfall of 90 mm, leading to waterlogging in 68 locations.

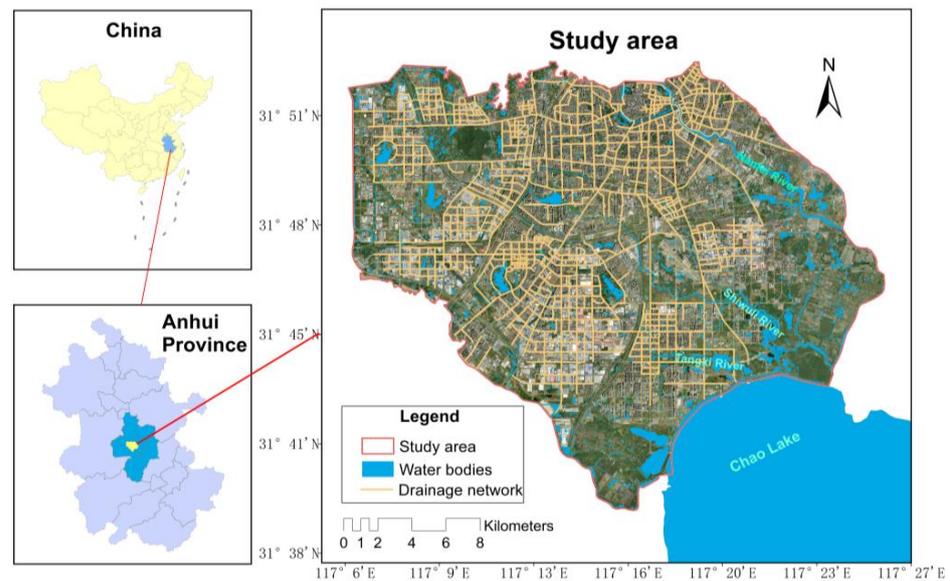


Figure 1. Location of the study area in China.

From 20:00 on 17 July 2020 to 06:00 on 19 July 2020, Hefei experienced a heavy rainfall process. The average rainfall in Hefei was recorded as 187 mm. The water level of Chao Lake exceeded the historical extreme. According to a report released by the Meteorological Bureau, the daily rainfall in Hefei has reached this standard once in 70 to 80 years. This flood disaster affected 805,136 people in Hefei, resulting in a direct economic loss of CNY 5.06 billion.

2.2. Flooding Event Sample Dataset

The construction of historical flood event sample datasets is the key to machine learning model training, which directly affects the rationality of the flood characteristics captured by the model and the flood risk assessment results. In this study, the historical flood inundation locations of Hefei City from 2017 to 2021 were obtained. On this basis, combined with the regional flood risk map in the flood control and drainage planning of Hefei City, 294 flood hotspots and 169 non-flood spots were finally determined as sample points for the training and validation of the machine learning model, as shown in Figure 2.

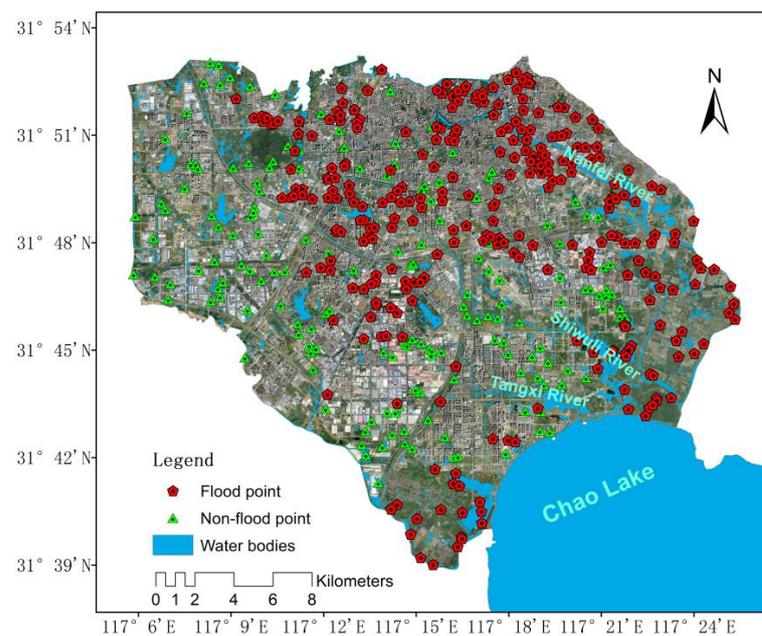


Figure 2. Spatial distribution of the sample points.

After normalizing and standardizing the data, the flood event sample dataset was randomly split into two datasets for training (80% of data, $n = 370$) and testing (20% of data, $n = 93$).

3. Methodology

3.1. Risk Assessment Framework

The risk assessment framework of this study is primarily divided into three aspects, as shown in Figure 3. Firstly, nine indicators relevant to urban flood risk from the perspectives of natural geography, meteorological hydrology, and human resilience are selected, and a sample database required for machine learning is constructed using historical inundation hotspots. Secondly, six machine learning models—SVM, RF, MLP, GBDT, voting, and stacking—are chosen, and the spatial distribution of flood risk in the study area is predicted after optimizing the hyperparameters of the models. Lastly, the predictive accuracy and performance of the models are evaluated by combining the results from different models. Furthermore, an analysis of the driving contributions of the influencing factors is conducted to elucidate the dominant factors causing flooding disasters in the region.

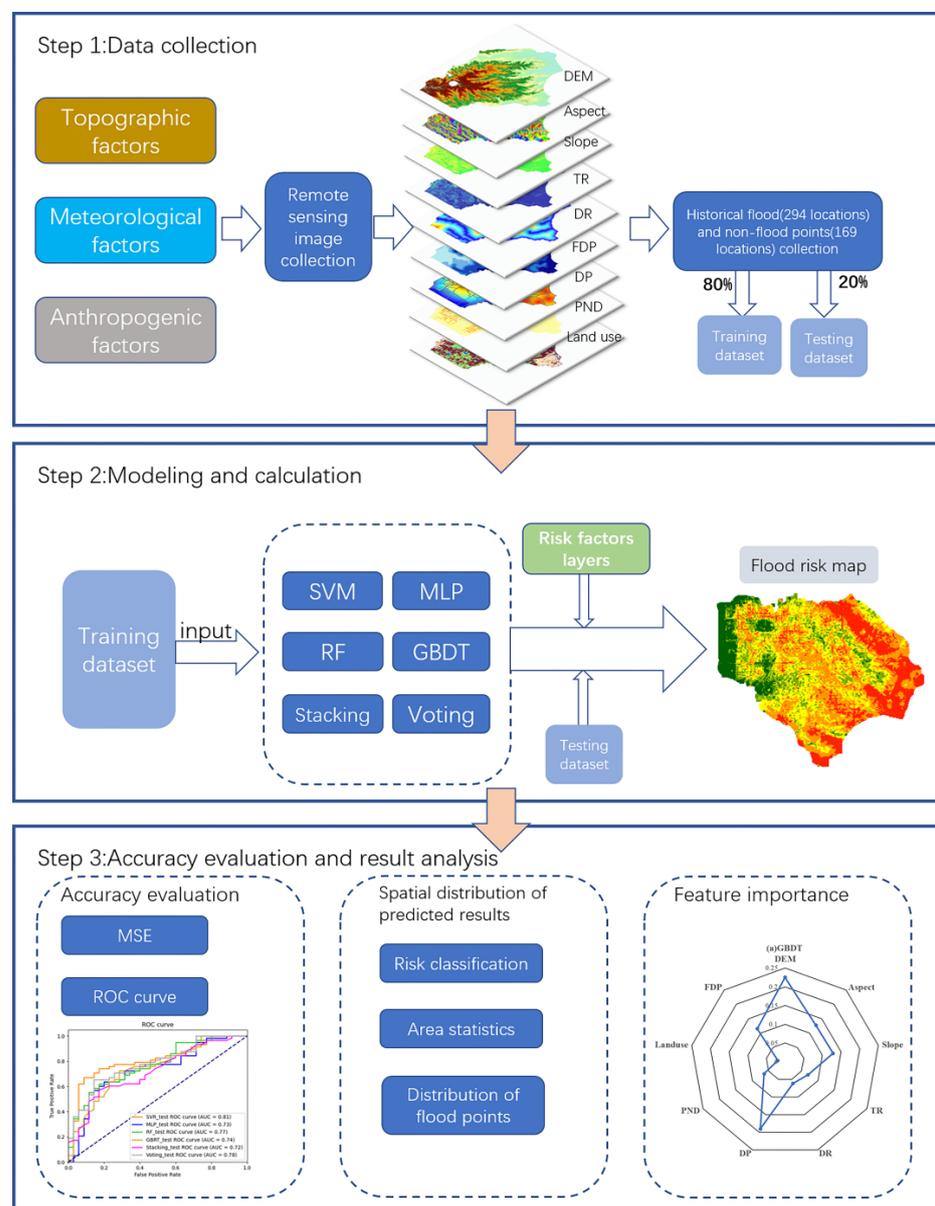


Figure 3. Flood risk assessment framework based on machine learning models.

3.2. Factors Affecting Urban Flooding

The selection of appropriate impact factors is a crucial step in risk assessment. Urban flooding is influenced by a variety of natural and social factors, and there are no universally prescribed selection criteria. In this study, taking into account the local characteristics of the study area and referencing the relevant literature, nine primary factors were determined to influence flooding in terms of meteorological factors (daily precipitation during the flood season), geographical environment factors (DEM, aspect, slope, topographic relief, distance to rivers, land use), and human resilience factors (distance to pumping stations, pipe network density).

The increasing availability of remote sensing technology has allowed for the development of increasingly reliable data collection methods, and the source data for most of the factors listed above can be obtained by processing the corresponding satellite remote sensing imagery, as described in Table 1.

Table 1. Data and data sources.

Data	Data Source
DEM	Geospatial Data Cloud (https://www.gscloud.cn , accessed on 10 December 2022)
Land use	Star Cloud Data Service Platform (http://data.starcloud.pcl.ac.cn/zh , accessed on 10 December 2022)
Pump station, pipe network	Flood control and drainage planning of Hefei City
Precipitation	HRLT dataset (https://doi.org/10.1594/PANGAEA.941329 , accessed on 6 March 2023)

(1) Digital Elevation Models (DEMs)

Elevation is the fundamental form of representation of terrain features [43,44]. In many studies on flood risk assessment, Digital Elevation Models (DEMs) have been employed as essential evaluation parameters [18]. A DEM with a 30 m spatial resolution was extracted from geospatial data clouds. The elevation ranges from 5.01 m to 262.89 m, as shown in Figure 4a.

(2) Slope, Aspect, and Topographic Relief (TR)

Slope and aspect have emerged as commonly selected evaluation factors owing to their significant influences on water flow velocity and direction [45,46]. In this study, slope and aspect data were derived from the DEM using ArcGIS, as shown in Figure 4b,c. As a macroscopic indicator for describing regional terrain features, topographic relief was calculated using ArcGIS based on the DEM, as shown in Figure 4d.

(3) Distance to Rivers (DR)

Many cities are located near mountains and rivers, and these areas tend to have relatively low elevations. Riverbanks and flood-prone zones are more vulnerable to flood impacts [47,48]. Distance to water bodies is an important factor in the analysis of waterlogging risk. Utilizing ArcGIS, the Euclidean distance from each point in the research area to water bodies was calculated, as shown in Figure 4e.

(4) Distance to pumping stations (DP), pipe network density (PND)

Urban drainage relies mainly on underground stormwater pipe networks, and the drainage capacity of a region depends on the distance to the pumping station and the density of the pipe network. In general, the denser the pipe network is, the stronger the drainage capacity of the area closer to the pump station will be, and it will be less susceptible to waterlogging. The pipe network density and distance to pumping station layers were obtained through ArcGIS editing, as shown in Figure 4f,g.

(5) Land use

Runoff conditions vary widely between different land use and land cover patterns [49,50]. Land use data with a 30 m spatial resolution were obtained by downloading remote sensing

images from the Star Cloud Data Service Platform. As shown in Figure 4h, there are a large number of impervious water surfaces in the study area, and the cultivated land is mainly distributed along the banks of the Nanfei River, Pai River, and Chao Lake.

(6) Daily Precipitation during the Flood Season (FDP)

Urban flooding is predominantly induced by heavy precipitation [51,52]. The occurrence of heavy precipitation is concentrated during the flood season. To capture the precipitation characteristics unique to this season, the average daily precipitation was computed for the period from June to September between 2009 and 2019 using the HRLT rainfall dataset [53]. Figure 4 illustrates the computed average daily precipitation for the flood season.

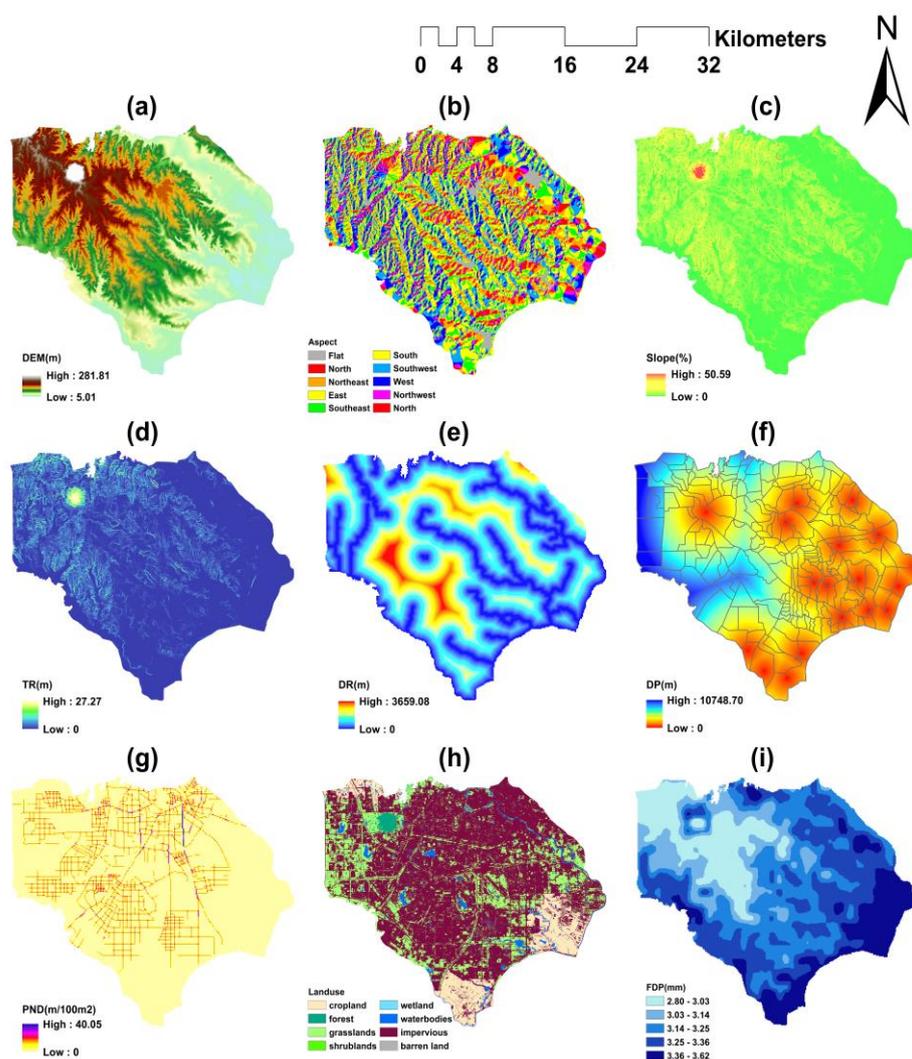


Figure 4. Spatial distribution of flood-influencing factors: (a) DEM, (b) aspect, (c) slope, (d) topographic relief, (e) distance to rivers, (f) distance to pump stations, (g) pipe network density, (h) land use, (i) daily precipitation during flood season.

3.3. Selection of Machine Learning Models

Traditional single machine learning models have been extensively employed in flood risk assessment; however, they are prone to overfitting. In this study, we tried to incorporate ensemble learning models for flood risk assessment. Six machine learning models were selected for evaluation, including two traditional single machine learning models, support vector machine and multilayer perceptron; two ensemble learning models based on homogeneous learners, random forest and gradient boosting decision tree; and two

ensemble learning models based on heterogeneous learners, voting and stacking ensemble learning. The specific principles of the various methods are as follows:

(1) Support Vector Machine (SVM)

Support vector machine (SVM) is a machine learning technique developed based on statistical learning theory. Its basic principle is to identify the best separation hyperplane in the feature space to maximize the interval between positive and negative samples in the training set [54,55]. By learning the two types of samples of flood occurrence and non-flood occurrence, the optimal classification hyperplane is found in the high-dimensional feature space, and the two types of data are correctly separated [56].

(2) Multi-layer sensor (MLP)

MLP is an artificial neural network (ANN) with a feedforward structure, mapping a set of input vectors to a set of output vectors [57]. It is composed of an input layer, hidden layer and output layer. The input layer receives the impact factors of the flood; the hidden layers process the input and transform it into the output, while the output layer predicts the flood risk value [58].

(3) Random Forest (RF)

The random forest algorithm is a bagging algorithm with a decision tree as the estimator. It connects multiple tree models in parallel. The dataset of each tree is randomly selected, and some features are randomly selected as inputs. Finally, all the trees' results are integrated as the final result [59]. For each tree, the flood risk value is finally obtained through the binary tree classification, moving from top to bottom, of the selected flood element index [60,61].

(4) Gradient Lifting Decision Tree (GBDT)

GBDT differs from random forest in that it adopts a boosting strategy as an ensemble learning algorithm [62]. By iterating multiple regression trees to make joint decisions, a learning device is constructed at each iteration step to reduce the loss of flood risk predictions along the steepest gradient direction so as to compensate for the shortcomings of the last iteration.

(5) Stacking ensemble learning

This is a heterogeneous learning technique that combines diverse base learners by training a model, unlike the homogeneous bagging and boosting methods, which directly aggregate the outputs of several learners to obtain the final prediction [63]. Generally, stacking consists of several base learners (level 0) and a meta-learner (level 1), in which the outputs of the base learners serve as the inputs of the meta-learner. Both the precision and variety of base learners affect the performance of a stacking algorithm.

(6) Voting ensemble learning

The voting ensemble method involves the construction of several heterogeneous classifiers, such as SVM, decision tree, logistic regression, and k-nearest neighbors (KNN). These classifiers are then combined using majority voting or weighted averaging to achieve more accurate classification results, significantly reducing model variance and improving overall performance. In this study, a voting regression model was utilized, where the average of the flood risk predictions from multiple base regressors served as the final prediction.

3.4. Model Construction and Hyperparameter Optimization

The optimization of hyperparameters in machine learning models aims to discover the most optimal set of hyperparameters that yield a superior performance on the test dataset. The choice of hyperparameters significantly affects the learning outcome of the model. In this study, we performed hyperparameter optimization for the SVM model's epsilon, the number of trees and maximum tree depth for the RF model, and the number of trees, maximum tree depth, and learning rate for the GBDT model. To carry out the

hyperparameter optimization process, a rigorous 5-fold cross-validation strategy was employed. The training dataset was partitioned into five subsets, and each subset was iteratively employed as a validation set, while the remaining four subsets were utilized for training and parameter tuning. This approach enables an unbiased evaluation of the hyperparameters' performance on data that are not used for training.

In the stacking ensemble model, we assembled the GBDT, RF, and SVM as the fundamental base regressors. The predictions from these three base models were further combined using the RF algorithm. On the other hand, in the voting ensemble model, the GBDT, RF, and SVM were also chosen as the base regressors, and the final prediction was obtained by averaging the outputs of the three base models.

All machine learning models were implemented using the Scikit learn library in Python. The hyperparameter optimization results of each model are shown in Table 2.

Table 2. Hyperparameters of machine learning models.

Model	Hyperparameters
SVM	epsilon = 0.3
MLP	Default
RF	n_estimators = 42, max_depth = 5
GBDT	n_estimators = 53, learning_rate = 0.03, max_depth = 3
Stacking	Base regressors: SVM, RF, GBDT, Final regressor: RF
Voting	Base regressors: SVM, RF, GBDT, Finally, take the average

Note: The hyperparameters not indicated in the table are set to their default values.

4. Results

4.1. Evaluation of Model Performance

(1) Mean Squared Error

The Mean Squared Error (MSE) serves as a metric for assessing the deviation between predicted and true values by measuring the square root of the ratio between the squared deviations and the number of observations (n) [64]. It is known for its sensitivity to outliers within a dataset. Table 3 presents the comparative results of the MSE for the various models used in this study. The results indicate that the RF, stacking, and voting models have the best performance in terms of the MSE on the training set, while on the testing set, the voting, RF, and SVM models exhibit the lowest MSE values. In particular, the stacking model exhibits the highest MSE on the testing set, suggesting its inferior generalization capability.

Table 3. MSE of the training and testing datasets for each model.

Model	SVM	MLP	RF	GBDT	Stacking	Voting
Training	0.396	0.387	0.310	0.375	0.332	0.355
Testing	0.432	0.451	0.429	0.440	0.459	0.428

(2) ROC curve

Receiver Operating Characteristic curve (ROC) is a tool used to evaluate the performance of classification models. The ROC curve is a two-dimensional graph. The horizontal coordinate is the false positive rate (FPR) and the vertical coordinate is the true positive rate (TPR). The closer the ROC curve is to the upper left corner, the better the model performance is. The closer the curve is to the diagonal, the worse the model performance is [65]. As shown in Figures 5 and 6, the RF model results show a good performance for both the test set and training set. Its ROC curve is closest to the upper right corner, and its AUC value is very high (training set 0.97, test set 0.77). The results of the SVM model show that the ROC curves of the test set and the training set are similar, and the AUC values are also similar, indicating that the SVM model has a low degree of overfitting and strong generalization

ability. Compared with the other models, the ROC curve of the MLP model is close to the diagonal, and the model performance is unsatisfactory. Stacking, as an ensemble learning model that integrates multiple models, does not exhibit a better performance.

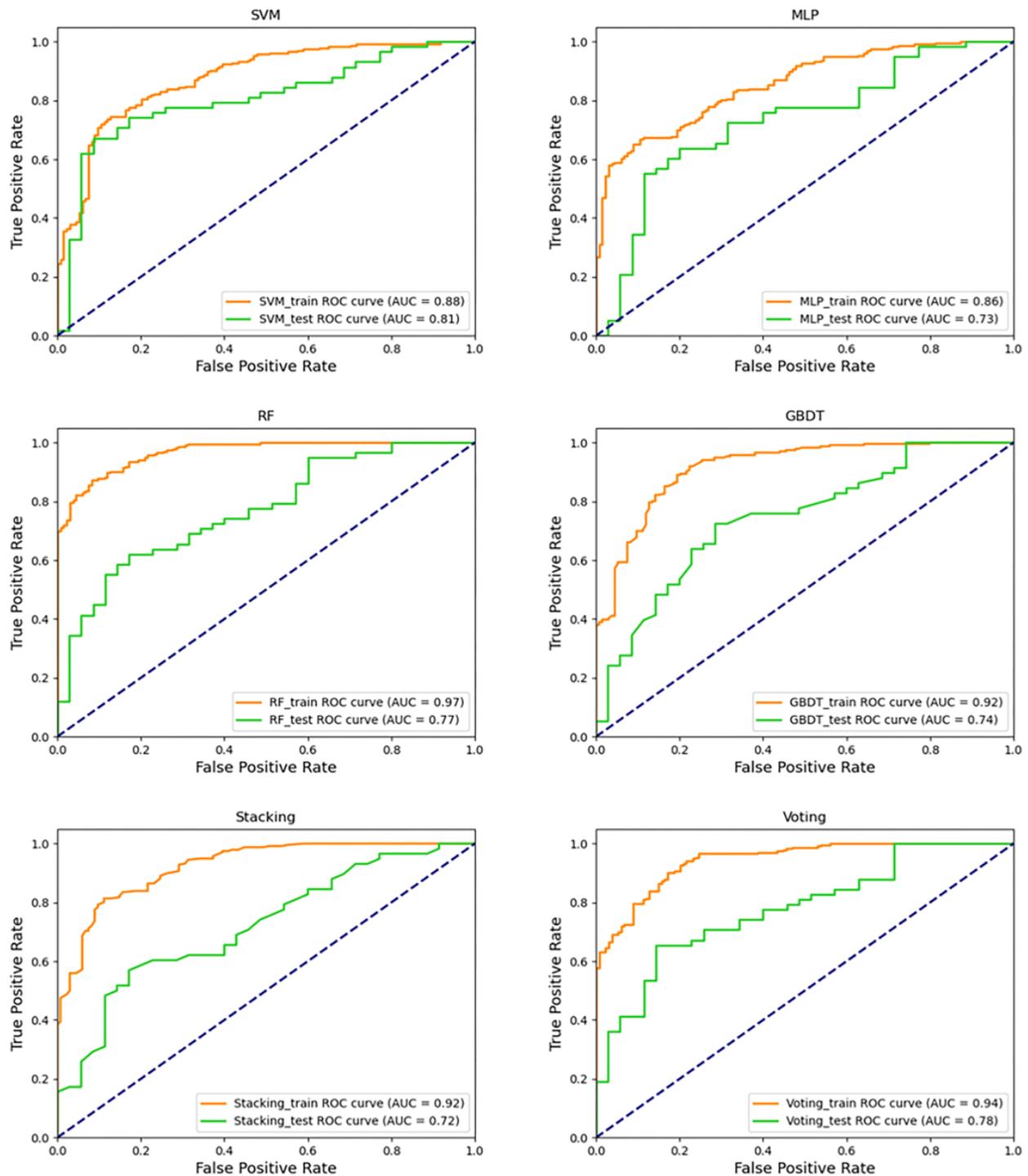


Figure 5. Training and testing ROC curves for each model.

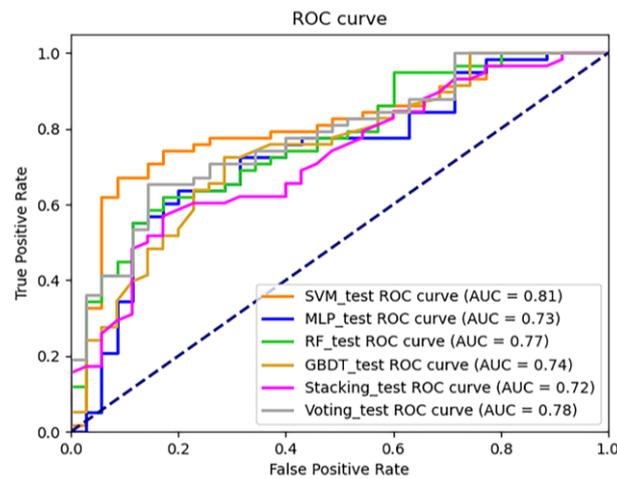


Figure 6. ROC curves of testing from the six models.

(3) F-score, precision, accuracy, and recall

The evaluation metrics, such as the F-score, precision, accuracy, and recall, were calculated for each model. The results are shown in Table 4.

Table 4. F-score, precision, accuracy, and recall for models.

Model	Accuracy		Precision		Recall		F-Score	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
GBDT	0.835	0.677	0.911	0.792	0.822	0.655	0.864	0.717
MLP	0.757	0.699	0.838	0.778	0.767	0.724	0.801	0.750
RF	0.886	0.688	0.949	0.774	0.869	0.707	0.907	0.739
SVM	0.792	0.710	0.835	0.754	0.839	0.793	0.837	0.773
STACKING	0.841	0.634	0.928	0.731	0.814	0.655	0.867	0.691
VOTING	0.859	0.699	0.911	0.768	0.864	0.741	0.887	0.754

From the results, it is evident that the training set results demonstrate that the RF model achieves optimal values across all four indicators. This finding aligns with the MSE conclusion, indicating that the RF model yields the best prediction performance. In the test set results, the SVM model demonstrates a strong performance on the three indicators, which aligns with the ROC results and suggests that SVM possesses robust generalization capabilities.

4.2. Spatial Distribution of Risk Prediction Results

The spatial distributions of flood risk predicted by the different machine learning models are shown in Figure 7. It can be seen that the majority of high-risk and very-high-risk areas are concentrated in low-lying areas, mainly along the banks of the Nanfei, Shiwuli, and Tangxi Rivers and near Chao Lake. In addition, scattered high-risk areas can be observed in the central urban zone. In particular, the result of the RF model closely aligns with the high-risk distribution map of historical floods mentioned in the “Comprehensive Planning of Urban Drainage (Rainwater) and Waterlogging Prevention in Hefei City” report.

In Figure 8, it can be observed that the risk values obtained using each model are primarily concentrated in the moderate- and high-risk categories, while the area classified as low-risk is relatively small across all the models. Specifically, the SVM model predicts a significantly smaller area in the very-high-risk category compared to the other models, while the area classified as high-risk is noticeably larger. This indicates that the predictive performance of the SVM model for the very-high-risk category is unsatisfactory, with a tendency to underestimate the risk.

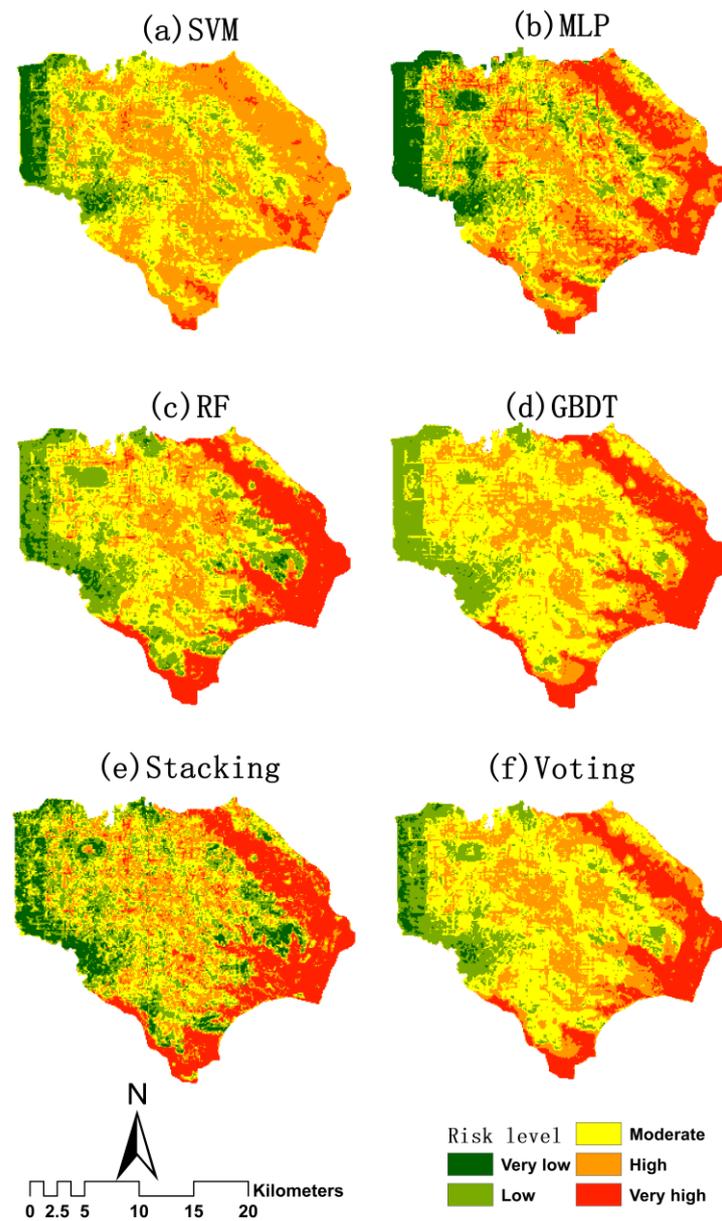


Figure 7. Flood risk maps predicted by the six models.

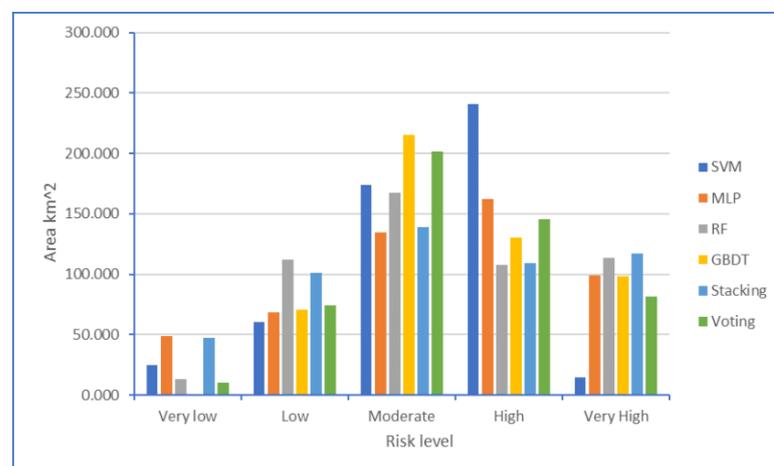


Figure 8. Area statistics of different risk categories for the six models.

Table 5 provides the statistics of inundation points in different risk categories for the six models. In comparison to the other models, the SVM model exhibits significantly fewer inundation points in the very-high-risk category, while having a higher number of points in the high-risk category. This aligns with the results shown in Figure 8, further confirming that the SVM model notably underestimated the extent of the very-high-risk range. The voting ensemble learning model showed a notably lower number of inundation points in the very-high-risk category compared to the other models, while the stacking ensemble learning model exhibited a higher number of points in this category. These results indicate that both ensemble learning models have an insufficient matching accuracy in predicting the extent of the very-high-risk area.

Table 5. Spatial distribution of flood points.

Risk Level	Waterlogging Points					
	SVM	MLP	RF	GBDT	Stacking	Voting
Very low	2	3	2	1	3	1
Low	5	11	5	2	18	4
Moderate	61	52	62	69	67	63
High	208	106	100	118	75	134
Very high	18	122	125	104	131	92

4.3. Analysis of Impact Factor Contribution

The GBDT and RF models are both based on tree models. Tree models possess interpretability, allowing for the determination of the contribution values (importance) of various factors to the model’s predictive outcomes, as illustrated in Figure 9. For both models, the top five factors in terms of importance were DEM, DP, slope, aspect, and FDP. The importance of these indicators was roughly similar for both models. However, in the case of the GBDT model, the importance of DP was approximately 4% higher than in the case of the RF model.

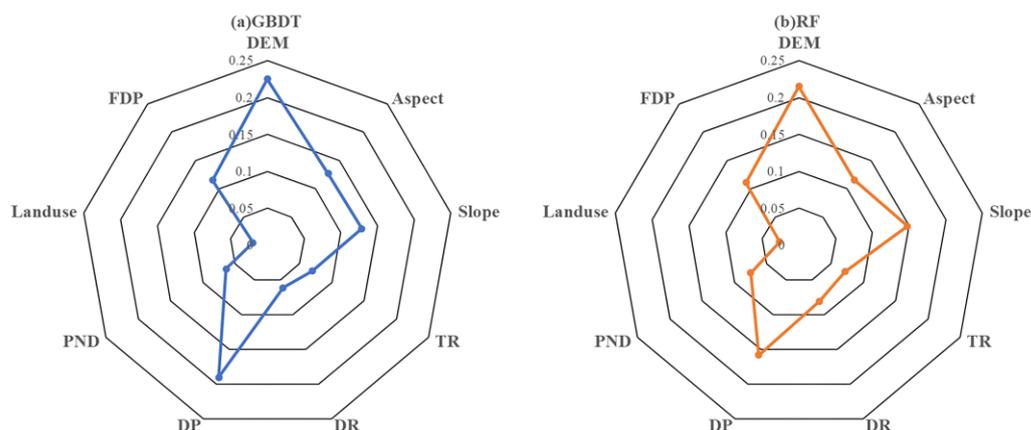


Figure 9. Contributions of the factors to urban flood risk according to GBDT and RF.

5. Discussion

5.1. Application Potential of Machine Learning Models

Theoretically, machine learning models with more advanced algorithms and superior performance should demonstrate better results in capturing flood characteristics and learning capabilities. However, such expectations were not fully realized in the specific application in this study. The results indicate that the RF model not only exhibits a higher accuracy on both the training and testing datasets but also demonstrates a greater congruence between the predicted spatial distribution of flood risk and the historical occurrence of inundation events. It outperforms the voting and stacking ensemble models in terms of prediction accuracy and performance. These outcomes align with those of analogous

investigations in related domains. For instance, Chen et al. [66] found that GBDT outperformed XGBoost in a flood risk assessment of the Pearl River Delta urban agglomeration in China, despite XGBoost generally being considered to have better learning capabilities. Yao et al. [67] discovered that ensemble learning methods were not necessarily superior to their base models in assessing flash flood sensitivity in Jiangxi, China. Stacking did not always outperform SVM or RF in terms of performance. It is worth noting that the applicability and generalizability of machine learning models across different research areas remain uncertain. The two ensemble models used in this paper relied on three base models, and the learning capabilities of the ensemble models were constrained by the abilities of these base models. Future research could explore diverse combinations of alternative foundational models, thus enhancing the performance of the ensemble models.

5.2. Factors Affecting Urban Flood Risk

Existing research on flood risk in watershed and urban areas has revealed that lower-lying regions are more susceptible to inundation disasters [68]. The influence of geographical environment factors, with DEM often taking a prominent position, is widely recognized [69]. Additionally, precipitation, as a factor triggering flooding, offers a substantial contribution [65]. However, this study considered the factor of human society's proactive resilience to flood disasters. The order of influence of various driving factors on the risk outcomes slightly differs from that in other literature. Elevation, distance to pumping stations, and slope emerge as the top three factors governing urban waterlogging risk in the central district of Hefei City. The results further validate the strong correlation between these indicators and waterlogging risk. The highest-risk areas are predominantly concentrated in low-lying areas along the rivers, but there are also a few distributed within the city center. These high-risk zones in the city center exhibit higher elevation than the riverbanks. However, while on one hand, they are relatively far from the coverage of pumping station control, on the other hand, the design standards of drainage networks in these areas are insufficient to meet the needs of urban development, severely hindering water drainage. It is evident that the vulnerability of human resilience measures has led to certain areas in the city center becoming high-risk zones.

5.3. Limitations and Future Directions

There are still some limitations of this study. The rate determination of the hyperparameter values was carried out on the basis of limited sample data in the specific study area examined in this paper and thus cannot be applied to cover various situations. With the continuous development of human activities, such as urbanization, agricultural expansion, and reservoir construction, human activities will have an increasing impact on flood risk. More indicator factors of the social and economic dimensions need to be considered in the assessment system, and the acquisition of this kind of high-resolution data will also be challenging. In addition, the specific mechanisms of human activities with respect to flood risk need to be studied in depth, including the different impacts of human activities on floods of different types and scales, and the mechanisms of interaction between human activities and natural factors. This will contribute to a more comprehensive understanding of the impact of human activities on flood risk and provide more scientific guidance for future flood risk management and decision making.

Currently, research on flood risk is predominantly focused on the urban or watershed scale. Future studies could be conducted on more refined spatial and temporal scales. For example, in-depth research could be conducted using high-resolution remote sensing data and geographic information system (GIS) technology to investigate the details of different land use types within cities and the interaction between cities and the natural environment. This would provide a better understanding of the influence of human activities on flood risk. Moreover, urban flood disaster research involves multiple disciplinary fields, including meteorology, hydrology, urban planning, and social sciences. Future efforts should further advance interdisciplinary research by integrating knowledge and methodologies

from various fields to thoroughly explore the comprehensive mechanisms underlying the combined impacts of climate change and human activities on urban flood disasters.

6. Conclusions

In this paper, nine factors were chosen from three aspects of natural geography, namely, meteorological hydrology, and human resilience. Then, a comprehensive risk assessment factor system and framework were constructed. Combined with the collected historical flood inundation point data, six machine learning models were used to assess the urban flood risk in Hefei City. The prediction results of each model were analyzed, and the potential mechanism of flood risk in these urban areas was revealed. Finally, the following main conclusions were drawn:

1. The results of the MSE analysis reveal that both the RF and voting ensemble models exhibit an excellent performance on both the training and testing datasets. However, it is worth noting that the stacking ensemble model only demonstrates a satisfactory performance on the training dataset, indicating its limited generalization capability. Additionally, based on the ROC curve analysis, the RF model stands out as the top-performing model. These findings, collectively, suggest that the predictive efficacy of ensemble models, which integrate heterogeneous learners, may not necessarily surpass that of their constituent base models.
2. The prediction results of the SVR model underestimate the range of extremely high-risk areas. Relatively speaking, the prediction results of the stacking ensemble model underestimate the range of extremely high-risk areas.
3. The high-risk and very-high-risk areas are mainly concentrated low-lying areas along rivers and near the Chao Lake region. The areas classified as medium- and high-risk outnumber those classified as low-risk. The overall risk level in the study area underscores the daunting challenge of urban flooding facing the city of Hefei.
4. The ranking results of factor importance indicate that geography-related factors constitute the major contributors among the top five contributing factors. It is worth noting that the factor of DP has the second most important driving influence after the DEM. This finding emphasizes the necessity of considering human resilience factors when conducting flood risk analysis in urban areas that are significantly impacted by human activities.

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

Funding: This research was funded by the Natural Science Foundation of China (grant No. 42175177, U2240216, 92047203, 42075191), Special Basic Research Key Fund for Central Public Scientific Research Institutes (grant No. Y521002), and the National Key Research and Development Program of China (grant No. 2019YFC1510204).

Data Availability Statement: The DEM dataset can be downloaded through <https://www.gscloud.cn/sources/accessdata/310?pid=302> (accessed on 10 December 2022). The land use data can be downloaded through <http://data.starcloud.pcl.ac.cn/zh/resource/1> (accessed on 10 December 2022). The river network data can be downloaded through <https://download.geofabrik.de/asia/china.html> (accessed on 10 December 2022). The HRLT dataset can be downloaded through <https://doi.pangaea.de/10.1594/PANGAEA.941329?format=html#download> (accessed on 6 March 2023). The pump station, pipe network and flood events sample can be obtained by contacting the corresponding author.

Acknowledgments: We would like to thank the assistant editors, academic editors and reviewers for their efforts to improve the quality of this paper. The authors would like to thank all the experts who contribute to urban flood protection and control.

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

References

1. Zhi, G.; Liao, Z.; Tian, W.; Wu, J. Urban flood risk assessment and analysis with a 3D visualization method coupling the PP-PSO algorithm and building data. *J. Environ. Manag.* **2020**, *268*, 110521. [\[CrossRef\]](#)
2. O'Donnell, E.C.; Thorne, C.R. Drivers of future urban flood risk. *Philos. Trans. R. Soc. A* **2020**, *378*, 20190216. [\[CrossRef\]](#)
3. Miller, J.D.; Hutchins, M. The impacts of urbanisation and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *J. Hydrol. Reg. Stud.* **2017**, *12*, 345–362. [\[CrossRef\]](#)
4. Qian, Y.; Chakraborty, T.C.; Li, J.; Li, D.; He, C.; Sarangi, C.; Chen, F.; Yang, X.; Leung, L.R. Urbanization impact on regional climate and extreme weather: Current understanding, uncertainties, and future research directions. *Adv. Atmos. Sci.* **2022**, *39*, 819–860. [\[CrossRef\]](#)
5. Arnold, C.L., Jr.; Gibbons, C.J. Impervious surface coverage: The emergence of a key environmental indicator. *J. Am. Plan. Assoc.* **1996**, *62*, 243–258. [\[CrossRef\]](#)
6. Ding, X.; Liao, W.; Lei, X.; Wang, H.; Yang, J.; Wang, H. Assessment of the impact of climate change on urban flooding: A case study of Beijing, China. *J. Water Clim. Chang.* **2022**, *13*, 3692–3715. [\[CrossRef\]](#)
7. Kourtis, I.M.; Tsihrintzis, V.A. Update of intensity-duration-frequency (IDF) curves under climate change: A review. *Water Supply* **2022**, *22*, 4951–4974. [\[CrossRef\]](#)
8. Miguez, M.G.; Mascarenhas, F.C.B.; Canedo de Magalhães, L.P.; D'Alterio, C.F.V. Planning and design of urban flood control measures: Assessing effects combination. *J. Urban Plan. Dev.* **2009**, *135*, 100–109. [\[CrossRef\]](#)
9. Waghwal, R.K.; Agnihotri, P.G. Flood risk assessment and resilience strategies for flood risk management: A case study of Surat City. *Int. J. Disaster Risk Reduct.* **2019**, *40*, 101155. [\[CrossRef\]](#)
10. Chen, H.; Liu, Y.; Lin, K.; Lan, T.; Liu, Z.; Li, W. Flood Hazard Assessment Methods: Research Review. *J. Water Resour. Res.* **2020**, *9*, 597. [\[CrossRef\]](#)
11. Li, C.; Sun, N.; Lu, Y.; Guo, B.; Wang, Y.; Sun, X.; Yao, Y. Review on Urban Flood Risk Assessment. *Sustainability* **2023**, *15*, 765. [\[CrossRef\]](#)
12. Sado-Inamura, Y.; Fukushi, K. Empirical analysis of flood risk perception using historical data in Tokyo. *Land Use Policy* **2019**, *82*, 13–29. [\[CrossRef\]](#)
13. Zhao, S.; Zhang, Q. Spatio-temporal risk assessment of crop flood in three provinces of Northeast China. *J. Catastrophol.* **2013**, *28*, 54–60.
14. Qin, N.; Jiang, T. Flood risk zoning and assessment in the middle and lower reaches of the Yangtze River based on GIS. *J. Nat. Disasters* **2005**, *14*, 5–11.
15. Liu, X.; Shi, P. Theory and Practice of Regional Flood Risk Assessment Model. *J. Nat. Disasters* **2001**, *2*, 66–72.
16. Ali, S.A.; Parvin, F.; Pham, Q.B.; Vojtek, M.; Vojteková, J.; Costache, R.; Linh, N.T.T.; Nguyen, H.Q.; Ahmad, A.; Ghorbani, M.A. GIS-based comparative assessment of flood susceptibility mapping using hybrid multi-criteria decision-making approach, naïve Bayes tree, bivariate statistics and logistic regression: A case of Topľa basin, Slovakia. *Ecol. Indic.* **2020**, *117*, 106620. [\[CrossRef\]](#)
17. Paul, P.; Sarkar, R. Flood susceptible surface detection using geospatial multi-criteria framework for management practices. *Nat. Hazards* **2022**, *114*, 3015–3041. [\[CrossRef\]](#)
18. Pham, B.T.; Luu, C.; Van Phong, T.; Nguyen, H.D.; Van Le, H.; Tran, T.Q.; Ta, H.T.; Prakash, I. Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam. *J. Hydrol.* **2021**, *592*, 125815. [\[CrossRef\]](#)
19. Tehrany, M.S.; Kumar, L. The application of a Dempster–Shafer-based evidential belief function in flood susceptibility mapping and comparison with frequency ratio and logistic regression methods. *Environ. Earth Sci.* **2018**, *77*, 490. [\[CrossRef\]](#)
20. Zou, Q.; Zhou, J.; Zhou, C.; Song, L.; Guo, J. Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP. *Stoch. Environ. Res. Risk Assess.* **2013**, *27*, 525–546. [\[CrossRef\]](#)
21. Chubey, M.S.; Hathout, S. Integration of RADARSAT and GIS modelling for estimating future Red River flood risk. *GeoJournal* **2004**, *59*, 237–246. [\[CrossRef\]](#)
22. Barredo, J.I. Major flood disasters in Europe: 1950–2005. *Nat. Hazards* **2007**, *42*, 125–148. [\[CrossRef\]](#)
23. Ding, Z.; Li, J.; Li, L. Method for flood submergence analysis based on GIS grid model. *J. Hydraul. Eng.-ASCE* **2004**, *6*, 56–67.
24. Li, J.; Cao, L.; Pu, R. Progresses on monitoring and assessment of flood disaster in remote sensing. *J. Hydraul. Eng.-ASCE* **2014**, *45*, 253–260.
25. Shrestha, B.B.; Kawasaki, A. Quantitative assessment of flood risk with evaluation of the effectiveness of dam operation for flood control: A case of the Bago River Basin of Myanmar. *Int. J. Disaster Risk Reduct.* **2020**, *50*, 101707. [\[CrossRef\]](#)
26. Su, B.; Huang, H.; Zhang, N. Dynamic risk assessment method for urban waterlogging based on scenario simulation. *J. Tsinghua Univ. Sci. Technol.* **2015**, *55*, 684–690.
27. Patro, S.; Chatterjee, C.; Mohanty, S.; Singh, H.; Raghuvanshi, N.S. Flood inundation modeling using MIKE FLOOD and remote sensing data. *J. Indian Soc. Remote Sens.* **2009**, *37*, 107–118. [\[CrossRef\]](#)
28. Rafiei-Sardooi, E.; Azareh, A.; Choubin, B.; Mosavi, A.H.; Clague, J.J. Evaluating urban flood risk using hybrid method of TOPSIS and machine learning. *Int. J. Disaster Risk Reduct.* **2021**, *66*, 102614. [\[CrossRef\]](#)
29. Mosavi, A.; Ozturk, P.; Chau, K. Flood prediction using machine learning models: Literature review. *Water* **2018**, *10*, 1536. [\[CrossRef\]](#)

30. Huang, G.; Luo, H.; Lu, X.; Yang, C.; Wang, Z.; Huang, T.; Ma, J. Study on risk analysis and zoning method of urban flood disaster. *Water Resour. Prot.* **2020**, *36*, 1–6.
31. Mekanik, F.; Imteaz, M.A.; Gato-Trinidad, S.; Elmahdi, A. Multiple regression and artificial neural network for long-term rainfall forecasting using large scale climate modes. *J. Hydrol.* **2013**, *503*, 11–21. [[CrossRef](#)]
32. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol.* **2013**, *504*, 69–79. [[CrossRef](#)]
33. Mojaddadi, H.; Pradhan, B.; Nampak, H.; Ahmad, N.; Ghazali, A.H.B. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1080–1102. [[CrossRef](#)]
34. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J. Hydrol.* **2014**, *512*, 332–343. [[CrossRef](#)]
35. Pham, B.T.; Luu, C.; Van Dao, D.; Van Phong, T.; Nguyen, H.D.; Van Le, H.; Von Meding, J.; Prakash, I. Flood risk assessment using deep learning integrated with multi-criteria decision analysis. *Knowl.-Based Syst.* **2021**, *219*, 106899. [[CrossRef](#)]
36. Wang, Z.; Lai, C.; Chen, X.; Yang, B.; Zhao, S.; Bai, X. Flood hazard risk assessment model based on random forest. *J. Hydrol.* **2015**, *527*, 1130–1141. [[CrossRef](#)]
37. Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Xu, L. Assessment of urban flood susceptibility using semi-supervised machine learning model. *Sci. Total Environ.* **2019**, *659*, 940–949. [[CrossRef](#)]
38. Zhao, G.; Pang, B.; Xu, Z.; Peng, D.; Zuo, D. Urban flood susceptibility assessment based on convolutional neural networks. *J. Hydrol.* **2020**, *590*, 125235. [[CrossRef](#)]
39. Wang, Y.; Fang, Z.; Hong, H.; Peng, L. Flood susceptibility mapping using convolutional neural network frameworks. *J. Hydrol.* **2020**, *582*, 124482. [[CrossRef](#)]
40. Youssef, A.M.; Pradhan, B.; Jebur, M.N.; El-Harbi, H.M. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. *Environ. Earth Sci.* **2015**, *73*, 3745–3761. [[CrossRef](#)]
41. Bui, D.T.; Hoang, N.D.; Martínez-Álvarez, F.; Ngo, P.T.T.; Hoa, P.V.; Pham, T.D.; Samui, P.; Costache, R. A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area. *Sci. Total Environ.* **2020**, *701*, 134413.
42. Wang, Y.; Fang, Z.; Hong, H.; Costache, R.; Tang, X. Flood susceptibility mapping by integrating frequency ratio and index of entropy with multilayer perceptron and classification and regression tree. *J. Environ. Manag.* **2021**, *289*, 112449. [[CrossRef](#)]
43. Daly, C.; Neilson, R.P.; Phillips, D.L. A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteorol. Climatol.* **1994**, *33*, 140–158. [[CrossRef](#)]
44. Gumbo, B.; Munyamba, N.; Sithole, G.; Savenije, H.H.G. Coupling of digital elevation model and rainfall-runoff model in storm drainage network design. *Phys. Chem. Earth Parts A/B/C* **2002**, *27*, 755–764. [[CrossRef](#)]
45. Saravanan, S.; Abijith, D.; Reddy, N.M.; Parthasarathy, K.S.S.; Janardhanam, N.; Sathiyamurthi, S.; Sivakumar, V. Flood susceptibility mapping using machine learning boosting algorithms techniques in Idukki district of Kerala India. *Urban Clim.* **2023**, *49*, 101503. [[CrossRef](#)]
46. Gao, H.; Cai, H.; Duan, Z. Understanding the impacts of catchment characteristics on the shape of the storage capacity curve and its influence on flood flows. *Hydrol. Res.* **2018**, *49*, 90–106. [[CrossRef](#)]
47. Mudashiru, R.B.; Sabtu, N.; Abustan, I.; Balogun, W. Flood hazard mapping methods: A review. *J. Hydrol.* **2021**, *603*, 126846. [[CrossRef](#)]
48. Chan, S.W.; Abid, S.K.; Sulaiman, N.; Nazir, U.; Azam, K. A systematic review of the flood vulnerability using geographic information system. *Heliyon* **2022**, *8*, e09075. [[CrossRef](#)]
49. Shi, P.; Ma, X.; Hou, Y.; Li, Q.; Zhang, Z.; Qu, S.; Chen, C.; Cai, T.; Fang, X. Effects of land-use and climate change on hydrological processes in the upstream of Huai River, China. *Water Resour. Manag.* **2013**, *27*, 1263–1278. [[CrossRef](#)]
50. Khorn, N.; Ismail, M.H.; Nurhidayu, S.; Kamarudin, N.; Sulaiman, M.S. Land use/land cover changes and its impact on runoff using SWAT model in the upper Prek Thnot watershed in Cambodia. *Environ. Earth Sci.* **2022**, *81*, 466. [[CrossRef](#)]
51. Sun, X.; Li, R.; Shan, X.; Xu, H.; Wang, J. Assessment of climate change impacts and urban flood management schemes in central Shanghai. *Int. J. Disaster Risk Reduct.* **2021**, *65*, 102563. [[CrossRef](#)]
52. Roy, S.; Bose, A.; Singha, N.; Basak, D.; Chowdhury, I.R. Urban waterlogging risk as an undervalued environmental challenge: An Integrated MCDA-GIS based modeling approach. *Environ. Chall.* **2021**, *4*, 100194. [[CrossRef](#)]
53. Qin, R.; Zhao, Z.; Xu, J.; Ye, J.; Li, F.; Zhang, F. HRLT: A high-resolution (1 day, 1 km) and long-term (1961–2019) gridded dataset for temperature and precipitation across China. *Earth Syst. Sci. Data* **2022**, *14*, 4793–4810. [[CrossRef](#)]
54. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
55. Vapnik, V.N. An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **1999**, *10*, 988–999. [[CrossRef](#)]
56. Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J.; Sajedi-Hosseini, F.; Mosavi, A. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2019**, *651*, 2087–2096. [[CrossRef](#)]
57. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
58. Firoozishahmirzadi, P.; Rahimi, S.; Esmaeili Seraji, Z. Application of Machine Learning Models for flood risk assessment and producing map to identify flood prone areas: Literature Review. *Int. J. Data Envel. Anal.* **2021**, *9*, 43–88.

59. Lai, C.; Shao, Q.; Chen, X.; Wang, Z.; Zhou, X.; Yang, B.; Zhang, L. Flood risk zoning using a rule mining based on ant colony algorithm. *J. Hydrol.* **2016**, *542*, 268–280. [[CrossRef](#)]
60. Pourghasemi, H.R.; Kariminejad, N.; Amiri, M.; Edalat, M.; Zarafshar, M.; Blaschke, T.; Cerda, A. Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. *Sci. Rep.* **2020**, *10*, 3203. [[CrossRef](#)]
61. Lai, C.; Chen, X.; Zhao, S.; Wang, Z.; Wu, X. A flood risk assessment model based on random forest and its application. *J. Hydraul. Eng.-ASCE* **2015**, *46*, 58–66.
62. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
63. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [[CrossRef](#)]
64. Karunasingha, D.S.K. Root mean square error or mean absolute error? Use their ratio as well. *Inf. Sci.* **2022**, *585*, 609–629. [[CrossRef](#)]
65. Lin, K.; Chen, H.; Xu, C.Y.; Yan, P.; Lan, T.; Liu, Z.; Dong, C. Assessment of flash flood risk based on improved analytic hierarchy process method and integrated maximum likelihood clustering algorithm. *J. Hydrol.* **2020**, *584*, 124696. [[CrossRef](#)]
66. Chen, J.; Huang, G.; Chen, W. Towards better flood risk management: Assessing flood risk and investigating the potential mechanism based on machine learning models. *J. Environ. Manag.* **2021**, *293*, 112810. [[CrossRef](#)]
67. Yao, J.; Zhang, X.; Luo, W.; Liu, C.; Ren, L. Applications of Stacking/Blending ensemble learning approaches for evaluating flash flood susceptibility. *Int. J. Appl. Earth Obs.* **2022**, *112*, 102932. [[CrossRef](#)]
68. Lee, S.; Kim, J.C.; Jung, H.S.; Lee, M.J.; Lee, S. Spatial prediction of flood susceptibility using random-forest and boosted tree models in Seoul metropolitan city. *Korea Geomat. Nat. Hazards Risk* **2017**, *8*, 1185–1203. [[CrossRef](#)]
69. Li, K.; Wu, S.; Dai, E.; Xu, Z. Flood loss analysis and quantitative risk assessment in China. *Nat. Hazards* **2012**, *63*, 737–760. [[CrossRef](#)]

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