Vietnam Journal of Earth Sciences, 48(2), 289−308, https://doi.org/10.15625/2615-9783/24214

VAST

Vietnam Journal of Earth Sciences
https://vjs.ac.vn/index.php/jse

Urban expansion trends and their relationship with flood susceptibility during the period 2014–2024 in Hanoi City

Huu Duy Nguyen[1*], Thanh Binh Doan[2], Si Dung Pham[2], Duc Dung Tran[3,4]

[1]*Faculty of Geography, University of Science, Vietnam National University, Hanoi, Vietnam*
[2]*Faculty of Architecture and Planning, Hanoi University of Civil Engineering, Hanoi, Vietnam*
[3]*National Institute of Education (NIE), Earth Observatory of Singapore (EOS) and Asian School of the Environment (ASE), Nanyang Technological University of Singapore, Singapore*
[4]*Center of Water Management and Climate Change (WACC), Institute for Environment and Resources (IER), Vietnam National University Ho Chi Minh City (VNU-HCM), Vietnam*

Received 23 November 2025; Received in revised form 26 January 2026; Accepted 13 February 2026

ABSTRACT

Over the past few decades, urban expansion has accelerated worldwide. This process can increase future flood risks due to local changes in hydrological conditions and the increased exposure and vulnerability of communities in flood-prone areas. Therefore, assessing the impact of urban expansion on flood susceptibility is an important task that can support local authorities in urban planning and in mitigating flood impacts. The objective of this study was to assess the impact of urban expansion on flood susceptibility in Hanoi using machine learning models: Deep Neural Networks (DNN), Adaptive Boosting (ADB), Extreme Gradient Boosting (XGB), and Random Forest (RF). A total of 1058 flood points and 14 conditioning factors corresponding to 2014 and 2024 were used as input to the models. Statistical indices, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Area Under the Curve (AUC), and Coefficient of Determination (R^2) were used to evaluate the performance of the proposed model. The results showed that the DNN model achieved the highest performance in assessing the impact of urban expansion on flood susceptibility (AUC=0.92), followed by XGB (0.91), ADB (0.86), and RF (0.82). During 2014–2024, urban expansion combined with the impacts of climate change has significantly increased the areas susceptible to flooding. In Hanoi, areas in the "high" and "very high" flood-susceptibility categories have been expanding continuously, accounting for about 25% of the total study area.

In contrast, the "medium" group has a slight decreasing trend, while the "low" and "very low" areas have narrowed. This shows that urban expansion is increasing the area prone to flooding. The results of this study provide a solid scientific basis, supporting planners and policymakers in identifying limitations in current flood risk adaptation measures and in developing more appropriate spatial and temporal strategies to minimize flood impacts.

Keywords: Urban expansion, flood susceptibility, Hanoi, machine learning.

1. Introduction

Urban expansion is closely associated with population growth, economic development, and urbanization, exerting significant impacts on the environment, economy, and quality of life (Bhunia et al., 2025). This process occurs both horizontally and vertically, leading to vegetation loss and a substantial increase in impervious surface areas. According to the World Bank (2021), between 2000 and 2020,

*Corresponding author, Email: nguyenhuuduy@hus.edu.vn

the land area of suburban and inner cities worldwide increased by approximately 930,000 km^2 (Nguyen et al., 2022). Many large cities in the world clearly show this trend. For example, Karachi expanded from 62.27 km^2 in 1940 to 89.24 km^2 in 1980; Paris doubled its inner-city area to 27,000 ha within two decades; New York had 12.3 million inhabitants in 1950; Tokyo is currently the largest metropolis, with over 34 million residents (Baig et al., 2024; Taubenböck et al., 2019). It is projected that by 2050, nearly two-thirds of the global population will reside in urban areas, with Africa experiencing the fastest urbanization rate (20%), followed by Asia and Oceania (26%) (Kundu and Pandey, 2020). The world population is expected to reach 9.1 billion by mid-century (Estoque and Murayama, 2017; Kundu and Pandey, 2020).

Numerous studies have demonstrated that rapid urban expansion significantly increases both the frequency and intensity of natural hazards, particularly floods (Güneralp et al., 2015). The expansion of impervious surfaces combined with rising population density strongly alters surface runoff processes and flood dynamics (Luu et al., 2022; Lan and Tien, 2009; Tien et al., 2016). Urban development along riverbanks, often coupled with weak land-use management, reduces permeable surfaces and amplifies surface runoff, thereby increasing flood risk. Moreover, rapid urbanization is often associated with inadequate drainage systems, exacerbating flooding during the rainy season. These processes not only intensify flood risk but also interact with environmental factors, including climate change (Das and Sahoo, 2025). Therefore, assessing the relationship between urban expansion and flood susceptibility is essential to support the development of effective planning strategies.

In recent years, machine learning models have been highly effective at analysing the nonlinear relationships between flood-prone areas and the combined impacts of urban expansion, land-use change, and hydrological factors (Viet et al., 2025). Some famous models, such as SVMs, are applied to classification, regression, and anomaly detection; meanwhile, RF works by having each decision tree predict the target variable, then aggregating these predictions to form a more powerful model; feedforward neural networks (FNNs) are capable of handling complex, nonlinear patterns from raw data. However, machine learning models require large data sets and are susceptible to noise, making them difficult to apply in areas with incomplete data, leading to significant performance degradation. In addition, these models have limitations in quantifying performance: floods are often represented as discrete variables (yes/no), whereas evaluation indicators such as accuracy or RMSE depend on the continuity of the output (Mosavi et al., 2018). Furthermore, each geographic area has distinct natural and socioeconomic characteristics, so it is necessary to develop and validate machine learning models tailored to specific areas. In the context of increasing urban expansion, there is an urgent need for powerful tools that can handle multidimensional, nonlinear data to assess the impact of this expansion on the likelihood of flooding.

Although the number of studies assessing flood risk in Hanoi has increased considerably in recent years, very few have thoroughly analyzed the relationship between urban growth and flood susceptibility using advanced methods, such as machine learning (Van Pham et al., 2025). Many previous studies have assessed flood risks independently of urban dynamics or have been limited to analyses based solely on condition land-use conditions (Madhuri et al., 2021). However, numerous studies have highlighted that assessing the impact of urban growth on flood susceptibility is crucial for supporting policymakers and planners in developing sustainable and resilient land-use strategies (Nguyen et al., 2024a).

We selected the machine learning models Adaptive Boosting (ADB), Random Forest (RF), XGBoost (XGB), and Deep Neural Networks (DNN) for their ability to exploit the complex, nonlinear relationships between human-natural factors and flood phenomena. Many previous studies have demonstrated that integrating multisource remote sensing data, including environmental and socioeconomic indicators, into machine learning models can significantly improve the accuracy. (Asfaw et al., 2025) used 28 environmental and rainfall factors extracted from remote sensing images for the urban area of Addis Ababa as input to the Random Forest model, and the results showed that adding urban infrastructure features significantly improved the accuracy of flood forecasting. Based on this, the objective of this study is to apply machine learning models - specifically ADB, XGB, RF, and DNN - to assess the impact of urban expansion on flood susceptibility in Hanoi. The difference from previous studies is that this study directly considers the spatial-temporal impact of urban expansion on flood susceptibility, rather than analysing the flood phenomenon solely under the influence of terrain or climate hydrology. The research results can help policymakers and planners develop effective urban development strategies, helping minimize the impact of floods.

2. Study area and material

2.1. Study Area

The study area is located between 20°53’–21°23’N, 105°44’–106°02’E (Fig. 1). The terrain is mainly low mountains, hills, and plains, gradually decreasing in altitude from northwest to southeast, following the flow of the Red River; mountainous areas are concentrated in the region's northern and western parts. The average altitude ranges from 5 to 20 m above sea level. The city has a dense river network, comprising the Red, To Lich, Da, and Nhue rivers, as well as other smaller river systems. Hanoi is located in the tropical monsoon climate zone, with high temperatures from April to June; the rainy season lasts from May to October, and the dry season from November to April. The average annual rainfall is approximately 1,760 mm. In 2024 alone, September rainfall totaled 697 mm, 525 mm above the same period in 2023, causing severe flooding in many low-lying areas.

In the period 2020–2024, Hanoi's population grew from 8.25 million to 8.72 million, with an average density of 2,595 people/km². In 2024, the immigration rate reached 5.9, while the emigration rate was only 4.5, reflecting the attraction of Hanoi for workers from the provinces, especially the Northern region. This growth led to high housing demand, driving the rapid conversion of agricultural land to urban and industrial uses. It is expected that by 2030, Hanoi's population will reach 10.5 million, rising to 13 million by 2050, with a scheduled urbanization rate of 80–85%. However, the current urban drainage system has not met demand, increasing the risk of flooding during the rainy season.

In 2024, Typhoon Yagi brought heavy rains, raising the flood level on the Red River in Hanoi to 10.76 m, exceeding the second alarm level by 0.26 m; on the Duong River in Thuong Cat, it reached 10.11 m (0.11 m above the second alarm level). 47 hectares of rice and 26.5 hectares of vegetable crops were submerged entirely; 6,144 hectares of rice, 15 hectares of vegetable crops, and approximately 2,500 trees were knocked down; over 2 hectares of fruit trees were damaged. Buildings were also affected, including five households whose roofs were blown off.

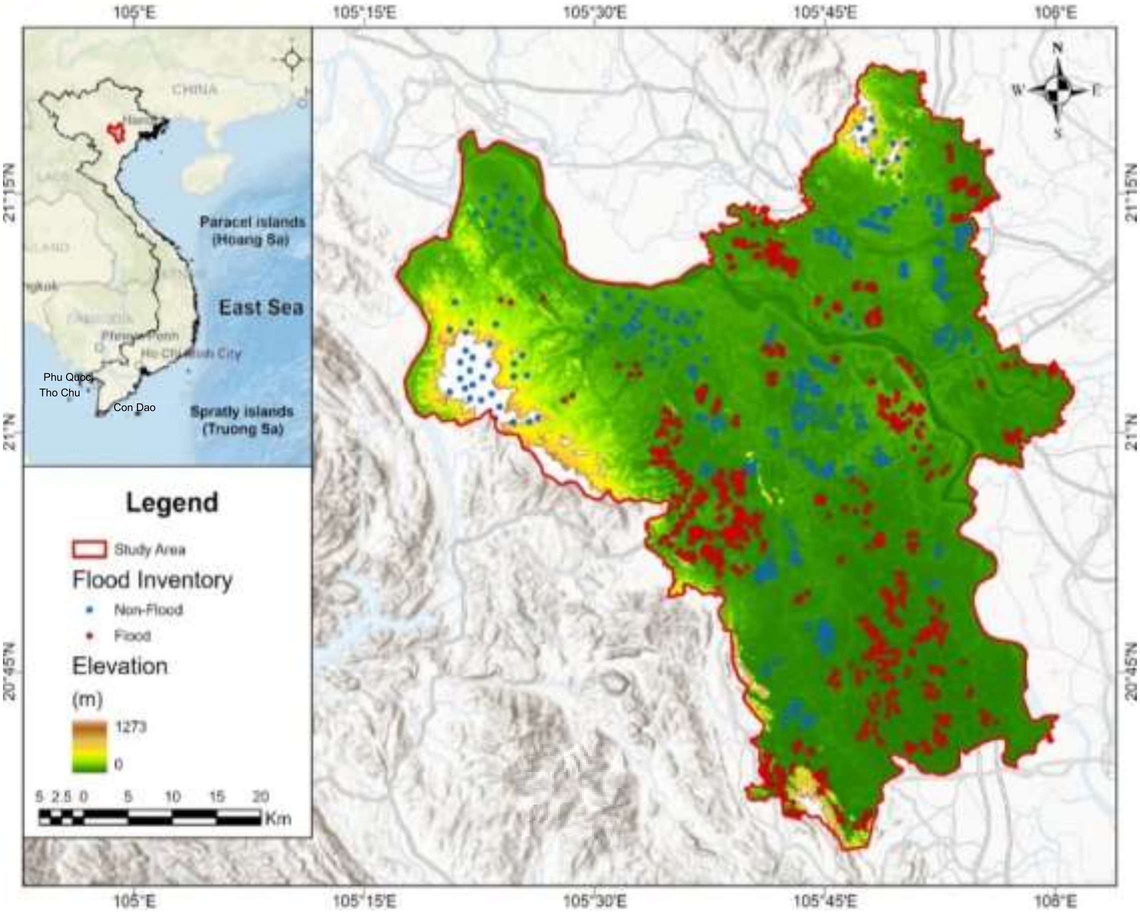

Figure 1. Location of Ha Noi Capital

2.2. Material

Flood Inventory and Conditioning Factors

Flood inventory maps play an important role in machine learning applications, as they reflect the relationship between past flood events and their causes (Long et al., 2026). Flood inventories are compiled from various data sources, including field-measurement data, hydrological-hydraulic models, and satellite images (Demissie et al., 2024). To improve data quality, flood inventory in this study was collected during a major flood event that occurred from 8–13 September 2024, associated with Typhoon Yagi. This Typhoon was selected because it represents a severe large-scale flood affecting in Hanoi City.

To determine the flood zones associated with Typhoon Yagi, this study used Sentinel-1A imagery. More precisely, a single Sentinel-1A image covering the entire administrative territory of Hanoi was used to ensure the spatial consistency of the flood map. The Sentinel-1A image acquired on September 12, 2024, was selected as the pre-flood image, representing the surface conditions just before the flood peak, while the image acquired on September 15, 2024, was used as the post-flood image. The use of SAR data enables reliable flood detection under all weather conditions. Finally, the flood inventory was generated by comparing the pre-flood and post-flood Sentinel-1A images.

In total, 524 flood points were collected to construct the flood susceptibility model for Hanoi City. Furthermore, the model used in this study is binary, therefore, non-flood points also play an important role. Several studies have shown that the number of flood and non-flood points is similar, which improves model performance. Thus, 524 non-flood points were collected in areas that had never been affected by flooding. In total, 1058 flood and non-flood points were collected to build the flood susceptibility model. This data was divided into two parts: 70% for training and 30% for testing.

The proper selection of conditioning factors is indispensable for assessing the reliability of the forecasting model. Selection depends on data availability and the natural and social characteristics of the study area. In this study, 14 conditional factors were used: aspect, curvature, elevation, slope, Normalised Difference Built-up Index (NDBI), Normalised Difference Vegetation Index (NDVI), Water Index (NDWI), and Urban Index (UI), Impervious Built-up Index (IBM), Normalized Difference Impervious Surface Index (NDISI), Land Use/land cover (LULC), Distance to river, Distance to road and Pumping station density. Aspect, curvature, elevation, and slope were extracted from DEMs built based on topographic maps, while the NDBI, NDVI, NDWI, and UI indices for the period 2014–2024 were calculated from Sentinel-2A satellite images.

Elevation, aspect, curvature, and slope were extracted from the DEM (constructed from the 1:50,000 scale topographic map available from the Ministry of Agriculture and the Environment). Distance to river and distance to road were calculated from the 1:50,000 scale topographic map using the Euclidean Distance tool in ArcGIS 10.6. NDBI, NDVI, IBM, NDISI, NDWI, and UI were extracted from the Sentinel-2A imagery recorded on August 2, 2025, and July 1, 2014. The detailed impacts of each conditioning factor on flood susceptibility are presented in Table 1A.

3. Methodology

In this study, the methodology used to assess the effects of urban growth on floods was divided into three main stages: (i) data collection, (ii) construction of machine learning model, and (iii) assessment of the effects of urban growth on flood susceptibility. The details of the methodology are presented in Fig. 2.

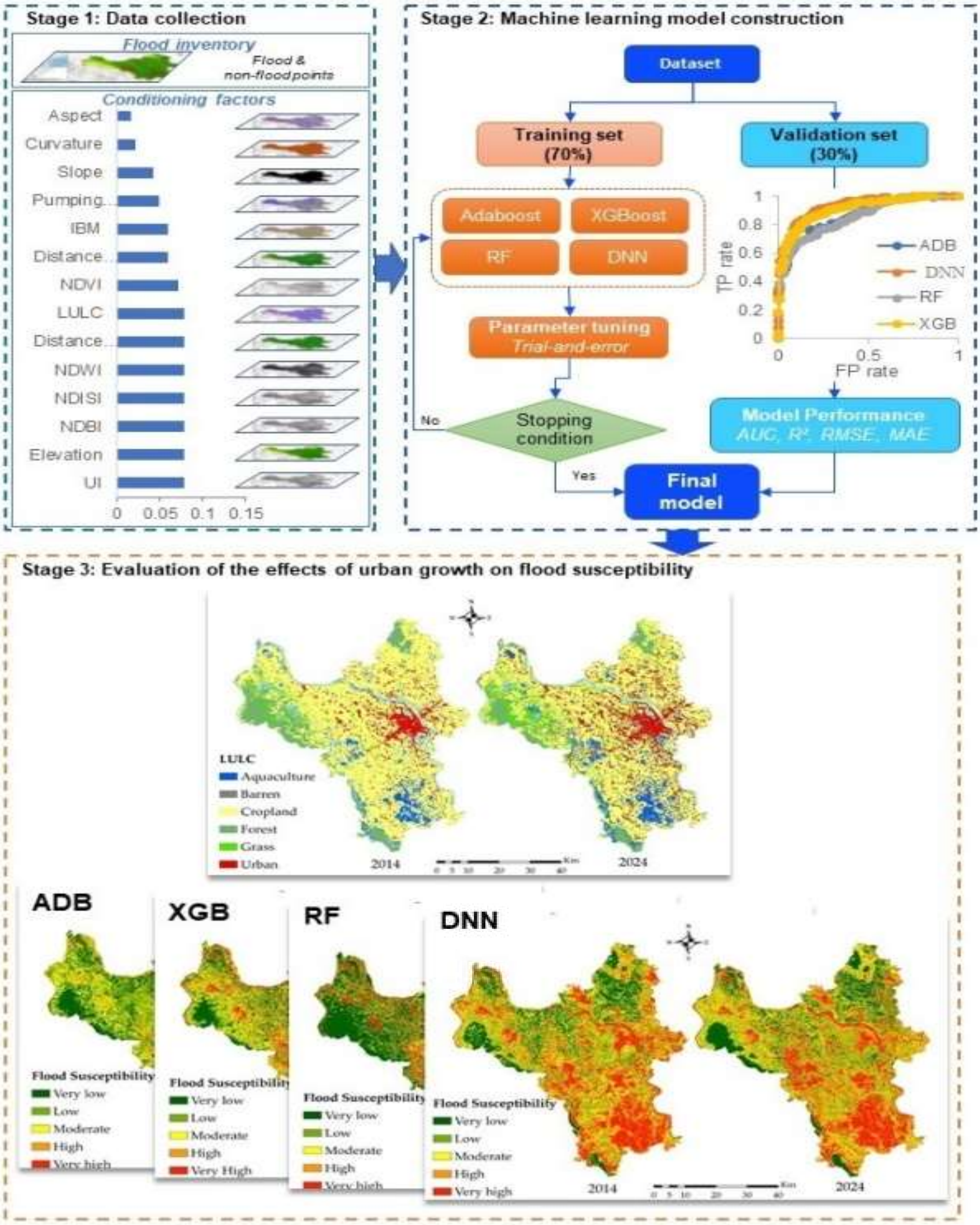

Figure 2. The methodology used for this study

3.1. Adaptive Boosting

ADB is a machine learning algorithm developed by and applied in many flood-related fields, such as flood detection, flood mapping, and sensitivity analysis. This algorithm works on the principle of combining multiple weak learners to form a strong learner with greater forecasting accuracy (Jahanbani et al., 2024). The core mechanism of ADB is to assign higher weights to training samples that were incorrectly predicted in previous iterations, thereby forcing the model to focus more on difficult-to-classify cases (Liu et al., 2017). Over many iterations, ADB builds a set of submodels, each assigned a weight proportional to its forecasting accuracy, thereby gradually improving overall performance (Li, 2025).

The efficiency of ADB depends heavily on hyperparameter tuning. The parameter n_estimators determines the number of weak models to be combined; a larger number can improve accuracy but also increase computational costs. The learning_rate parameter controls how much each weak model contributes to the overall model, with low values making the model learn slowly but more consistently, while high values can reduce generalization ability. The base_estimator parameter specifies the type of weak model used, from simple decision trees to more complex models. Finally, the algorithm parameter allows choosing between SAMME (suitable for multiclass classification) and SAMME.R (which uses predicted probabilities, which is more effective in many cases). With these features, ADB has proven to be effective and flexible in improving the accuracy of flood forecasting and flood mapping.

3.2. Extreme Gradient Boosting

XGB is part of the Gradient Boosting family of algorithms, designed to optimise performance, speed, and scalability. It was developed by the Distributed Machine Learning Community (DMLC) (Ma et al., 2021). Unlike traditional machine learning methods, it operates on an iterative principle, in which weak models (weak learners) are continuously added and refined to overcome the errors of previous models, thus creating a strong model (strong learner) with high accuracy (Ren et al., 2024). In particular, XGB utilises both the first and second derivatives of the loss function to determine the optimal direction and level of adjustment and integrates regularization techniques to reduce model complexity and limit overfitting, thereby improving the reliability of flood forecasting (Linh et al., 2022).

The performance of XGB depends heavily on hyperparameter tuning. The eta (learning rate) parameter controls the influence of each new tree added to the model; the max_depth parameter specifies the maximum depth of the decision tree, where a larger depth can help the model learn better but also increases the risk of overfitting. The subsample parameter determines the proportion of randomly selected data samples used to train each tree, while colsample_bytree controls the proportion of features sampled for each tree, thereby increasing diversity and reducing correlation between trees. In addition, a set of parameters related to the learning task also plays an important role, including objective (determining the problem type, such as binary classification, multiclass, or regression) and eval_metric (a metric for evaluating model performance, such as AUC, RMSE, or logloss). The appropriate selection and tuning of these parameters is key to XGB's high performance in flood risk forecasting and analysis.

3.3. Random Forest

The RF algorithm works on the principle of bagging, in which multiple decision trees are built independently and in parallel, and the results are then combined through a voting mechanism (for classification) or averaging (for regression) (Wang et al., 2015). To increase diversity among trees, RF uses random subsets of data and features during training. The model's performance in flood forecasting depends heavily on the selection and tuning of hyperparameters (Chen et al., 2020). The parameter n_estimators determines the number of trees in the forest, with higher values generally improving accuracy but increasing computational costs. The max_depth parameter controls the maximum depth of each tree: too small a depth can lead to underfitting, while too large a depth can easily lead to overfitting. The max_features parameter specifies the number of features considered at each split; smaller values increase diversity among trees and reduce the risk of overfitting. Additionally, the min_samples_split and min_samples_leaf parameters determine the minimum node split and the minimum leaf size, thereby limiting the complexity of the tree. Finally, the bootstrap parameter enables random sampling with replacement, which contributes to maintaining the randomness and stability of the model.

3.4. Deep Neural Networks

DNNs exploit multilayer neural networks to analyze complex datasets from raw data, thereby significantly improving the accuracy of flood forecasting and risk assessment. This method is especially suitable for multidimensional and nonlinear data processing problems, such as satellite image analysis for flood identification, time-series-based flood forecasting, or modeling the impact of environmental variables (Anbarasan et al., 2020). Input data often includes satellite images, digital elevation models (DEMs), and information about past flood events. After collection and preprocessing, DNNs can be used to identify flood areas from remote sensing images, while sequential regression architectures such as RNN or LSTM support flood forecasting based on rainfall time series. Image segmentation models such as U-Net or SegNet are also used to identify detailed flood boundaries, thereby improving the accuracy of risk mapping (Panahi et al., 2021). During training, DNNs rely on the backpropagation mechanism to optimise and reduce prediction errors. Unlike tree-based models such as XGB, the performance of DNNs depends heavily on the network architecture and training process. Hyperparameters play an important role, such as the learning rate, which adjusts the speed at which the weights are updated (too high can cause the model not to converge, while too low slows down the training process) (Hawamdeh et al., 2025). The number of layers and neurons determines the network's performance, but an overly complex architecture can lead to overfitting, especially with limited flood data (Shao et al., 2024). Nonlinear activation functions, such as ReLU and Sigmoid, enable the model to learn complex relationships in weather and terrain data. Optimisation algorithms such as Adam or SGD adjust the approach to minimising the loss function, while batch size directly affects the stability of the training process. Additionally, regularisation techniques such as dropout and batch normalisation help reduce overfitting by randomly inactivating neurones or normalizing inputs across layers, thereby improving the model's generalisation ability (Johri et al., 2024).

4. Results

4.1. Change in land use

Table 2A and Figure 3 show the changes in land use in Hanoi during the period 2010-2024. There was a significant change in land-use structure, especially the rapid expansion of urban areas. The forested area decreased dramatically from 319.56 km² (9.56%) to 204.53 km² (6.12%), reflecting the levels of urbanization and conversion to other land uses. The water surface area also decreased rapidly, from 285.77 km² (8.55%) to 187.38 km² (5.60%), largely due to the filling of lakes and ponds for urban development. Agricultural land, although still representing a large proportion of the study area, also decreased sharply, from 2137.99 km^2 (63.96%) to 1902.20 km^2 (56.90%), as a direct result of urbanization. Barren land decreased slightly from 48.24 km^2 (1.44%) to 42.18 km^2 (1.26%), while aquaculture increased dramatically, from 173.53 km^2 (5.19%) to 288.03 km^2 (8.62%), and grassland more than doubled, from 122.64 km^2 (3.67%) to 300.32 km^2 (8.98%), reflecting the trend of expanding livestock farming and food demand. Urban land increased rapidly at a rate of 47.6%, expanding 12.52% from 255.63 km^2 in 2010 to 418.65 km^2 in 2024.

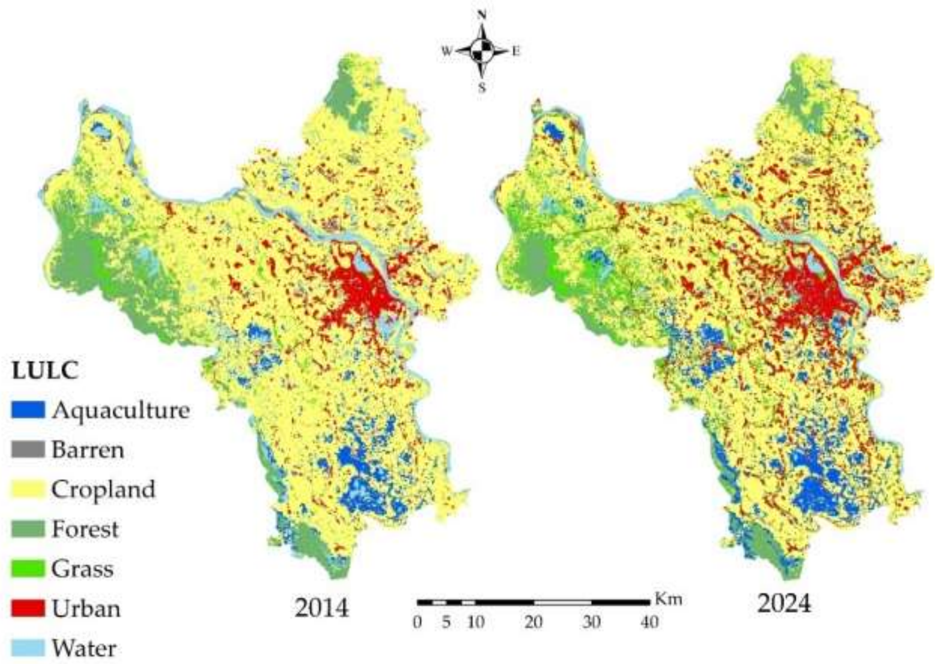

Figure 3. Land use/land cover in Hanoi in 2014 and 2024

4.2. Conditioning factor selection

Evaluating the importance of conditioning factors is an essential step when using machine learning to build flood susceptibility maps, because it eliminates unnecessary or unimportant factors that affect the accuracy of prediction models. We used RF, a popular feature selection method, to assess the importance of the factors. The results showed that UI, elevation, NDBI, NDISI, NDWI, and distance to river are the most important factors for flooding in Hanoi; the UI, NDBI, and NDISI indices directly influence the waterproof capacity. In recent years, the city's urban growth has accelerated, leading to increased concrete, asphalt, and roofing. Areas with high UI, NDBI, or NDISI values represent regions of concrete expansion. These surfaces prevent rainwater from seeping into the soil, thereby increasing surface runoff. This causes water to stagnate or flow rapidly over the ground during rain. Elevation determines the capacity for rainwater accumulation. Specifically, low-lying areas near rivers are more susceptible to flooding. Hanoi is located in the Red River Delta. Most of the city is less than 20 m above sea level; therefore, it is more susceptible to flooding during heavy rains. Several studies have shown that areas below 8 m above sea level have a very high risk of flooding, while areas above 31 m have a lower risk.

NDWI was ranked fifth in importance because a higher NDWI value reflects a high volume of surface water or highly saturated soil. An increase in NDWI indicates that the soil contains more water, which predicts an elevated risk of flooding during heavy rainfall. Although distance to the river is considered an important factor, its influence can be masked by topographic and land-use factors, which exert greater control over surface runoff and water accumulation in the urban area. Distance to the river is the sixth most important factor because areas near rivers are always directly affected by rising water levels. In Hanoi, history shows that when the Bui and Tich rivers rise, the areas along them are affected by flooding. It should be noted that the models used in this study are statistical models; therefore, the importance of the factors depends strongly on the statistical relationships between the conditioning factors and the flood points.

LULC, NDVI, distance to road, IBM, and pumping station density factors have an average influence on the probability of flooding in Ha Noi City. In Ha Noi City, urban growth leads to an increase in the area of construction, such as buildings and roads, and thus a reduction in the area of vegetation. This increases impervious surfaces and the probability of flooding. For example, the Hong Mai and Hai Ba Trung regions are mainly built-up areas with few trees. Therefore, these areas are often affected by flooding during heavy rains. In suburban areas such as Thanh Tri and Gia Lam, many agricultural areas still contribute to relatively efficient water drainage. Areas with high NDVI often reduce surface runoff due to their water-retention capacity and reduced water velocity. In Ha Noi, low NDVI is concentrated in residential areas and the city centre, while high NDVI is found in suburban areas. For example, the Ha Dong and Thanh Tri areas (with high NDVI indices) are less affected by flooding than urban areas such as Hoang Mai and Long Bien. Road construction increases impermeable surfaces and can hinder water flow. The density of the pumping station represents the number of water-pumping facilities per unit area. In Ha Noi, the high pumping station density allows rapid water drainage, but its impact is only moderate because, during heavy rains, pumping capacity is limited. More specifically, during heavy rains, the Thanh Tri and Hoang Mai regions remain flooded, even when the pumping station is operating at full capacity. The factors slope, curvature, and aspect are less influential on the probability of flooding because Hanoi has a flat relief, so their values are almost identical throughout the territory. Therefore, they are not as relevant (Figure 1A).

4.3. Model Performance and Comparison

Figure 2A presents the AUC value of all the models proposed to evaluate the effects of urban growth on flood susceptibility in Hanoi. The results showed that all proposed models performed well during both training and validation. Among them, during training, the XGB model performed best, with an AUC of 0.98, followed by DNN (0.97), RF (0.88), and ADB (0.87). In the validation set, the DNN model outperformed the other models, with an AUC of 0.92, followed by XGB (0.91), ADB (0.86), and RF (0.85).

This study also used RMSE, MAE, and R^2 to evaluate model performance. In terms of the training process, for XGB, the value of RMSE (0.18) and MAE (0.14) was lower than that of other models; therefore, its performance was better than that of other models, followed by DNN (RMSE = 0.22, MAE = 0.17), RF (RMSE = 0.33 and MAE = 0.25), and ADB (RMSE = 0.35 and MAE = 0.25). In the validation process, the RMSE and MAE values of the XGB model remained lower than those of other models (RMSE = 0.26, MAE = 0.2), followed by DNN (RMSE = 0.28, MAE = 0.21), RF (RMSE = 0.36, MAE = 0.27), and ADB (RMSE = 0.38, MAE = 0.29).

The R^2 value of the XGB model was higher than that of other models in both the training and validation processes (R^2 = 0.97 for the training process and R^2 = 0.88 for the validation process), followed by DNN (R^2 = 0.93, R^2 = 0.87), RF (R^2 = 0.85, R^2=0.82), and ADB (R^2 = 0.84, R^2 = 0.81) (Table 1).

Table 1. Performance of the models using RMSE, MAE, AUC, and R^2

	Training dataset				Validation dataset			
	RMSE	MAE	AUC	R2	RMSE	MAE	AUC	R2
ADB	0.35	0.25	0.87	0.84	0.38	0.29	0.86	0.81
DNN	0.22	0.17	0.97	0.93	0.28	0.21	0.92	0.87
RF	0.33	0.25	0.88	0.85	0.36	0.27	0.85	0.82
XGB	0.18	0.14	0.98	0.97	0.26	0.2	0.91	0.88

4.4. Effect of Urban Expansion on flood susceptibility

Figure 3 presents the flood susceptibility map in Hanoi using XGB. While Fig. 3A presents the flood susceptibility map produced by DNN, RF, and ADB. The map shows that the distribution of flood risk in Hanoi is clearly differentiated in space and time. The areas along the Red River, especially in the south and southwest of the city, have a high to very high flood risk. These low-lying areas near the river are often directly affected by flooding. In contrast, the urban centre is located mainly in a medium- to high-risk zone, reflecting high urbanization and a limited drainage network. The northern and northwest areas have low to very low risk, consistent with the characteristics of high terrain and dense vegetation. The distribution of flood susceptibility classes in Ha Noi City for each model is presented in Table 3A.

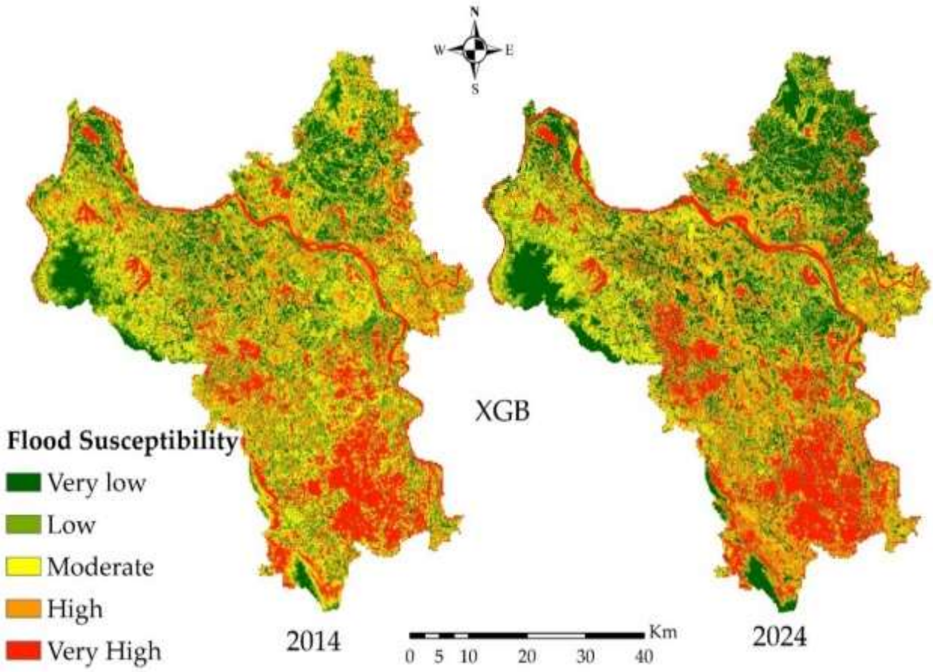

Figure 3. Flood susceptibility in 2014 and 2024 by XGB in Hanoi city

In general, between 2014 and 2024, and 2024, flood-prone areas have tended to increase in several regions due to urban growth. Several studies have highlighted that increased precipitation and temperatures have led to increased intensity and frequency of flooding. Specifically in Hanoi, areas classified as "high" and "very high" show an increasing trend. These two categories represent approximately 25% of the study area and continue to expand. In contrast, the "moderate" category has slightly decreased, while areas classified as "low" and "very low" tend to gradually diminish.

5. Discussions

Floods are dangerous natural hazards that cause significant damage to people and property and hinder a country's development (Alfieri et al., 2017). They have tended to increase in both number and intensity in the context of climate change and urban expansion. As a result, many urban areas, especially in developing countries, are becoming increasingly vulnerable to flooding (Onchi-Ramos et al., 2024). In that context, assessing the impact of urban expansion on flood susceptibility is important for supporting policymakers in developing effective adaptation strategies, reducing risk, and promoting sustainable development (Demissie et al., 2024). The objective of this study was to assess the impact of urban expansion on flood susceptibility using ADB, RF, DNN, ADB machine learning models in Hanoi, Vietnam. The results of this study showed that urban expansion leads to significant changes in land use, increasing impervious area and, consequently, the probability of flooding. By accurately identifying 14 key factors affecting the flood process in Hanoi, this study helps clarify the impact of urban expansion on flood vulnerability. This study has many similarities with the results recorded of many other regions in the world. Since the Doi Moi reforms in 1986, the urban land area in Vietnam's delta regions has nearly doubled, with the rate of expansion accelerating since the 2000s. The pattern of urban expansion in these cities, characterized by fringes and leapfrog, is increasingly common, similar to the trend observed in many global megacities. This process disrupts natural flow and contributes to increased flood risks, identical to the expansion of Jakarta or Sao Paulo, where rapid urbanization combined with heavy rains led to severe floods. In Hanoi, urban expansion has been considerable, reflected in the conversion of agricultural land to construction land and the formation of high-density residential areas.

An increase in built-up land surface leads to more impervious surfaces, reducing the role of natural vegetation in regulating water flow. This increases the study area's vulnerability to flooding (Pham et al., 2015). Historically, the Red River Delta has been highly vulnerable to flooding due to alluvial deposition and low-lying terrain (Luo et al., 2018). Since the country's Doi Moi process in 1986, the pace of urban expansion in Hanoi has accelerated, driven by economic development and rural-urban migration. Furthermore, the merger of Ha Tay into Hanoi in 2008 not only expanded administrative boundaries but also created conditions for foreign investment, thereby promoting economic growth and expanding the capital's urban area (Smith and Scarpaci, 2000a). In addition, economic reforms, especially the gradual reform of land-use rights management and transfer mechanisms and the emergence of the real estate market, have made land an important resource for speculation and capital accumulation. In that context, the development of new urban areas not only plays an essential role in promoting economic growth and attracting participation from local authorities, foreign investors, and residents, but also contributes to rapid urban expansion (Petrişor et al., 2020). This development process has led to a population explosion: Hanoi's population increased from 6.9 million in 2010 to 8.7 million in 2024. This explosion of population and urban space expansion emphasises the direct impact of the transition from a centrally planned economy to a market economy, while opening up to foreign investment and international integration. However, this development leads to increased impervious surfaces and the breakdown of natural structures, thereby increasing the likelihood of flooding. Specifically, the results of the study show that the area of areas at medium, high, and very high risk of flooding increased from 2014 to 2024. This has been demonstrated in many previous studies. In planning strategies, the urbanization rate of Hanoi in 2030 is expected to reach 65-75% (Van Pham et al., 2025). In Hanoi, the urban development structure is expected to be organised around five spaces, five corridors, five dynamic axes, five socioeconomic regions, and five urban areas. These orientations are anticipated to strongly influence land-use transformation, particularly the conversion of agricultural land into urban residential land. This trend is most evident in suburban districts such as Đông Anh and Gia Lâm, where urbanisation is the primary driver of development.

The application of machine learning models is an effective way to improve forecast accuracy and integrate diverse data sources. Previous studies have confirmed that machine learning models can significantly improve the accuracy of flood risk maps and play an important role in flood risk management and mitigation. These models exploit historical data, geospatial information, and environmental factors to identify areas at risk of flooding and support the development of disaster response and preparedness strategies. Incorporating geospatial datasets, such as elevation and land cover, has been shown to improve model accuracy, with many studies reporting AUC values exceeding 0.80 (Al-Kindi and Alabri, 2024; Bui et al., 2023). Several studies have also shown that machine learning models play an important role in flood assessment by simultaneously analysing multiple factors affecting flood risk, using algorithms such as k-nearest neighbours (KNN), Decision Trees (DTs), and Support Vector Machines (SVMs) to improve prediction accuracy, optimise flood management strategies and improve understanding of risk mechanisms in urban areas (Yuwono et al., 2024). Specifically, a study in Jakarta applied support vector regression (SVR), achieving an R^2 of 0.977 and an RMSE of 0.112, demonstrating the potential of SVM for predicting flood levels (Azi et al., 2024). SVM has also been deployed in real-time flood warning systems, helping to provide immediate responses in vulnerable areas. In Malaysia, SVMs are highly effective for flood prediction using 29 months of rainfall and river water level data (Azi et al., 2024), with advantages in classifying flood and non-flood areas, handling complex datasets, and providing reliable results (Mosavi et al., 2018). This evidence confirms that machine learning algorithms, especially variants of SVM, are powerful tools for predicting river water levels and flood occurrence, thereby helping minimise the socioeconomic impact of natural disasters. In addition to SVM, other algorithms such as RF and Gradient Boosting have been applied to analyze hydrological variables (rainfall, streamflow, and flood frequency) and have performed well in flood forecasting, including in Jeddah, Saudi Arabia (Al-Areeq et al., 2022). In addition, ADB has demonstrated the ability to automatically map flood inundation from remote sensing data, such as MODIS imagery, without manual intervention, enabling rapid response in emergencies (Ahamed and Bolten, 2017). Later studies have further improved this method by combining AdaBoost (ADB) with spatial context learning to extract flood maps from optical satellite images (Liu et al., 2017), or by integrating ADB with DTs to analyse factors such as drainage density and rainfall to build reliable flood inventory maps (Coltin et al., 2016). In addition to traditional models, deep learning (DL) models have been widely applied in recent years to assess the impact of land-use change, in general, and urban expansion, in particular, on areas prone to flooding. Neural network architectures have demonstrated superior capabilities for processing satellite imagery to detect and model nonlinear relationships in floods. Additionally, sequential neural regression models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks have proven effective in processing time-series patterns, such as rainfall and streamflow, thereby improving the reliability of flood forecasting (Liu et al., 2023).

Although the study successfully assessed the impact of urban expansion on flood susceptibility, it still had limitations related to data use. The study did not consider the impact of urban shape on flood susceptibility. Although the study used some urban indicators directly related to impervious surfaces and built-up land area, vertical morphological factors that can affect surface runoff, such as bridge characteristics and building shapes, were not integrated. Future studies will incorporate these factors better to assess the impact of urban expansion on floods. In addition, considering the effect of urban drainage systems is quite simple when only indicators related to pumping station density are used, rather than indicators that fully reflect the drainage system's functions, such as pipe density.

6. Conclusions

Flooding regularly impacts people and the economy in Vietnam, and flood risk tends to increase with urban growth. Therefore, assessing the effects of this relationship is essential to support decision-makers in sustainable land use planning. The objective of this study was to assess the effects of urban growth on flood susceptibility using machine learning, namely the DNN, ADB, RF, and XGB models, in the city of Hanoi. The conclusion was as follows:

(i) The models proposed in this study can be generated for application in other parts of in Vietnam and the wider world. ii) Among the proposed models, the DNN model performed better than the others, with an AUC value of 0.92, followed by XGB (AUC=0.91), ADB (AUC=0.86), and RF (AUC=0.82). These results highlight the superiority of the DNN and recommend its use for constructing flood-susceptibility maps and assessing the effects of urban growth on flood vulnerability.

(ii) Urban growth has a significant effect on flood susceptibility in Hanoi. More specifically, the very high and high flood susceptibility areas are increasing rapidly.

References

Ahamed A., Bolten J.D., 2017. A MODIS-based automated flood monitoring system for Southeast Asia. International Journal of Applied Earth Observation and Geoinformation, 61, 104−117. https://doi.org/10.1016/j.jag.2017.05.006.

Ahmed K.R., Akter S., 2017. Analysis of landcover change in southwest Bengal delta due to floods by NDVI, NDWI and K-means cluster with landsat multi-spectral surface reflectance satellite data. Remote Sensing Applications: Society and Environment, 8, 168−181. https://doi.org/10.1016/j.rsase.2017.08.010.

Al-Areeq A.M., Abba S.I., Yassin M.A., Benaafi M., Ghaleb M., Aljundi I.H., 2022. Computational machine learning approach for flood susceptibility assessment integrated with remote sensing and GIS techniques from Jeddah, Saudi Arabia. Remote Sensing, 14, 5515. https://doi.org/10.3390/rs14215515.

Al-Juaidi A.E., 2023. The interaction of topographic slope with various geo-environmental flood-causing factors on flood prediction and susceptibility mapping. Environmental Science and Pollution Research, 30, 59327−59348. https://doi.org/10.1007/s11356-023-26616-y.

Al-Kindi K.M., Alabri Z., 2024. Investigating the role of the key conditioning factors in flood susceptibility mapping through machine learning approaches. Earth Systems and Environment, 8, 63−81. https://doi.org/10.1007/s41748-023-00369-7.

Albano R, Adamowski J (2025) Use of digital elevation models for flood susceptibility assessment via a hydrogeomorphic approach: A case study of the Basento River in Italy. Natural Hazards, 121, 9021−9042. https://doi.org/10.1007/s11069-025-07144-z.

Alfieri L., Bisselink B., Dottori F., Naumann G., de Roo A., Salamon P., Wyser K., Feyen L., 2017. Global projections of river flood risk in a warmer world. Earth's Future, 5, 171−182. Doi: 10.1002/2016EF000485.

Anbarasan M., Muthu B., Sivaparthipan C., Sundarasekar R., Kadry S., Krishnamoorthy S., Dasel A.A., 2020. Detection of flood disaster system based on IoT, big data and convolutional deep neural network. Computer Communications, 150, 150−157. https://doi.org/10.1016/j.comcom.2019.11.022.

Asfaw W., Rientjes T., Bekele T.W., Haile A.T., 2025. Estimating elements susceptible to urban flooding using multisource data and machine learning. International Journal of Disaster Risk Reduction, 116, 105169. https://doi.org/10.1016/j.ijdrr.2024.105169.

Azi A., Saleh R.F., Ardana W.M., Kusrini K., 2024. Flood Prediction Using Support Vector Regression (Case Study of Floodgates in Jakarta). Journal of Computer Networks, Architecture and High Performance Computing, 6, 1333−1344. https://doi.org/10.47709/cnahpc.v6i3.4360.

Baig A., Atif S., Tahir A., 2024. Urban development and the loss of natural streams leads to increased flooding. Discover Cities, 1, 9. https://doi.org/10.1007/s44327-024-00010-w.

Bhunia A., Chatterjee U., Gupta J., Gupta K., 2025. Urbanisation and Quality of Life: A Comparative Assessment between Developed and Developing Countries. Sustainability and Urban Quality of Life. Routledge India, 3−21. Doi: 10.4324/9781003604358-2.

Bui Q.D., Luu C., Mai S.H., Ha H.T., Ta H.T., Pham B.T., 2023. Flood risk mapping and analysis using an integrated framework of machine learning models and analytic hierarchy process. Risk Analysis, 43, 1478−1495. https://doi.org/10.1111/risa.14018.

Chen W., Li Y., Xue W., Shahabi H., Li S., Hong H., Wang X., Bian H., Zhang S., Pradhan B., 2020. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. Science of The Total Environment, 701, 134979. https://doi.org/10.1016/j.scitotenv.2019.134979.

Coltin B., McMichael S., Smith T., Fong T., 2016. Automatic boosted flood mapping from satellite data. International Journal of Remote Sensing, 37, 993−1015. Doi: 10.1080/01431161.2016.1145366.

Damayanti H.N., Wikan P.A., Annurhutami F., 2024. GIS-based flood susceptibility mapping in Wawar watershed, Purworejo Regency. IOP Conference Series: Earth and Environmental Science. IOP Publishing, 012051. Doi: 10.1088/1755-1315/1314/1/012051.

Das A., Sahoo S.N., 2025 Impact of land use and climate change on urban flooding: a case study of Bhubaneswar city in India. Natural Hazards, 121, 8655−8674. https://doi.org/10.1007/s11069-025-07130-5.

Demissie Z., Rimal P., Seyoum W.M., Dutta A., Rimmington G., 2024. Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment. Applied Computing and Geosciences, 23, 100183. https://doi.org/10.1016/j.acags.2024.100183.

Estoque R.C., Murayama Y., 2017 Trends and spatial patterns of urbanization in Asia and Africa: a comparative analysis. Urban Development in Asia and Africa: Geospatial Analysis of Metropolises. Springer, 393−414. https://doi.org/10.1007/978-981-10-3241-7_19.

Freund Y., Schapire R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55, 119−139. https://doi.org/10.1006/jcss.1997.1504.

Güneralp B., Güneralp İ., Liu Y., 2015. Changing global patterns of urban exposure to flood and drought hazards. Global Environmental Change, 31, 217−225. https://doi.org/10.1016/j.gloenvcha.2015.01.002.

Hawamdeh A., Tarawneh A., Sharrab Y., Al-Fraihat D., 2025. Deep neural networks hydrologic and hydraulic modeling in flood hazard analysis. Water Resources Management, 1−17. https://doi.org/10.1007/s11269-025-04243-1.

Hoang D.-V., Liou Y.-A., 2024. Assessing the influence of human activities on flash flood susceptibility in mountainous regions of Vietnam. Ecological Indicators, 158, 111417. https://doi.org/10.1016/j.ecolind.2023.111417.

Jahanbani M., Vahidnia M.H., Aghamohammadi H., Azizi Z., 2024. Flood susceptibility mapping through geoinformatics and ensemble learning methods, with an emphasis on the AdaBoost-Decision Tree algorithm, in Mazandaran, Iran. Earth Science Informatics, 17, 1433−1457. https://doi.org/10.1007/s12145-023-01213-2.

Johri S., Pratap S., Narayan J., Dwivedy S.K., 2024. Sensory data classification for gait assessment using deep neural networks: A comparative study with SGD and Adam optimizer. 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). IEEE, 1−6. Doi: 10.1109/IATMSI60426.2024.10503416.

Khosravi K., Shahabi H., Pham B.T., Adamowski J., Shirzadi A., Pradhan B., Dou J., Ly H.-B., Gróf G., Ho H.L., 2019. A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. Journal of Hydrology, 573, 311−323. https://doi.org/10.1016/j.jhydrol.2019.03.073.

Kundu D., Pandey A.K., 2020. World urbanisation: trends and patterns. Developing national urban policies: Ways forward to green and smart cities. Springer, 13−49. https://doi.org/10.1007/978-981-15-3738-7.

Lan N.H., Tien N.N., 2009. GIS based loss estimation modeling for flooding in the case of long wave as tsunami. Vietnam Journal of Marine Science and Technology, 9. https://doi.org/10.15625/1859-3097/9/1/6250.

Li Z., 2025. Flood Hazard Prediction Model Based on K-means Clustering Analysis and AdaBoost Regression Model. 2025 International Conference on Electrical Drives, Power Electronics & Engineering (EDPEE). IEEE, 182−186. Doi: 10.1109/EDPEE65754.2025.00036.

Linh N.T.T., Pandey M., Janizadeh S., Bhunia G.S., Norouzi A., Ali S., Pham Q.B., Anh D.T., Ahmadi K., 2022. Flood susceptibility modeling based on new hybrid intelligence model: Optimization of XGboost model using GA metaheuristic algorithm. Advances in Space Research, 69, 3301−3318. https://doi.org/10.1016/j.asr.2022.02.027.

Liu X., Sahli H., Meng Y., Huang Q., Lin L., 2017. Flood inundation mapping from optical satellite images using spatiotemporal context learning and modest AdaBoost. Remote Sensing, 9, 617. https://doi.org/10.3390/rs9060617.

Liu Y., Yang Y., Chin R.J., Wang C., Wang C., 2023. Long short-term memory (LSTM) based model for Flood forecasting in Xiangjiang River. KSCE Journal of Civil Engineering, 27, 5030−5040. https://doi.org/10.1007/s12205-023-2469-7.

Liuzzo L., Sammartano V., Freni G., 2019. Comparison between different distributed methods for flood susceptibility mapping. Water Resources Management, 33, 3155−3173. https://doi.org/10.1007/s11269-019-02293-w.

Long G., Tantanee S., Nusit K., Sooraksa P., 2026. Aggregated flood susceptibility mapping in Upper Chao Phraya River Basin using Shannon's Entropy, Machine Learning, and Stacking ensemble methods. Geomatics, Natural Hazards and Risk, 17, 2614731. https://doi.org/10.1080/19475705.2026.2614731.

Luo P., Mu D., Xue H., Ngo-Duc T., Dang-Dinh K., Takara K., Nover D., Schladow G., 2018. Flood inundation assessment for the Hanoi Central Area, Vietnam under historical and extreme rainfall conditions. Scientific reports, 8, 12623. https://doi.org/10.1038/s41598-018-30024-5.

Luu C., Nguyen D.D., Amiri M., Van P.T., Bui Q.D., Prakash I., Pham B.T., 2022. Flood susceptibility modeling using Radial Basis Function Classifier and Fisher's linear discriminant function. Vietnam Journal of Earth Sciences, 44(1), 55−72. https://doi.org/10.1038/s41598-018-30024-5.

Ma M., Zhao G., He B., Li Q., Dong H., Wang S., Wang Z., 2021. XGBoost-based method for flash flood risk assessment. Journal of Hydrology, 598, 126382. https://doi.org/10.1016/j.jhydrol.2021.126382.

Madhuri R., Sistla S., Srinivasa Raju K., 2021. Application of machine learning algorithms for flood susceptibility assessment and risk management. Journal of Water and Climate Change, 12, 2608−2623. https://doi.org/10.2166/wcc.2021.051.

Mosavi A., Ozturk P., Chau K.-W., 2018. Flood prediction using machine learning models: Literature review. Water, 10, 1536. https://doi.org/10.3390/w10111536.

Nguyen H.D., Nguyen Q.-H., Dang D.K., Van C.P., Truong Q.H., Pham S.D., Bui Q.-T., Petrisor A.-I., 2024a. A novel flood risk management approach based on future climate and land use change scenarios. Science of The Total Environment, 921, 171204. https://doi.org/10.1016/j.scitotenv.2024.171204.

Nguyen N.T., Nhon N.T.T., Hai H.T.N., Chi N.D.T., Hien T.T., 2022. Characteristics of microplastics and their affiliated PAHs in surface water in Ho Chi Minh City, Vietnam. Polymers, 14, 2450. https://doi.org/10.3390/polym14122450.

Onchi-Ramos K., Rodríguez-Cuevas C., Couder-Castañeda C., Padilla-Pérez D.-A., 2024. Flood risk assessment of the Garita River in the urban zone of San Luis Potosí City, by hydrodynamic modeling. Scientific Reports, 14, 15891. https://doi.org/10.1038/s41598-024-66743-1.

Panahi M., Jaafari A., Shirzadi A., Shahabi H., Rahmati O., Omidvar E., Lee S., Bui D.T., 2021. Deep learning neural networks for spatially explicit prediction of flash flood probability. Geoscience Frontiers, 12, 101076. https://doi.org/10.1016/j.gsf.2020.09.007.

Petrişor A.-I., Hamma W., Nguyen H.D., Randazzo G., Muzirafuti A., Stan M.-I., Tran V.T., Aştefănoaiei R., Bui Q.-T., Vintilă D.-F., 2020. Degradation of coastlines under the pressure of urbanization and tourism: Evidence on the change of land systems from Europe, Asia and Africa. Land, 9, 275. https://doi.org/10.3390/land9080275.

Pham V.C., Pham T.-T.-H., Tong T.H.A., Nguyen T.T.H., Pham N.H., 2015. The conversion of agricultural land in the peri-urban areas of Hanoi (Vietnam): patterns in space and time. Journal of Land Use Science, 10, 224−242. https://doi.org/10.1080/1747423X.2014.884643.

Rahmati O., Darabi H., Panahi M., Kalantari Z., Naghibi S.A., Ferreira C.S.S., Kornejady A., Karimidastenaei Z., Mohammadi F., Stefanidis S., 2020. Development of novel hybridized models for urban flood susceptibility mapping. Scientific reports, 10, 12937. https://doi.org/10.1038/s41598-020-69703-7.

Ren H., Pang B., Bai P., Zhao G., Liu S., Liu Y., Li M., 2024. Flood susceptibility assessment with random sampling strategy in ensemble learning (RF and XGBoost). Remote Sensing, 16, 320. https://doi.org/10.3390/rs16020320.

Shao Y., Chen J., Zhang T., Yu T., Chu S., 2024. Advancing rapid urban flood prediction: a spatiotemporal deep learning approach with uneven rainfall and attention mechanism. Journal of Hydroinformatics, 26, 1409−1424. https://doi.org/10.2166/hydro.2024.024.

Smith D.W., Scarpaci J.L., 2000b. Urbanization in transitional societies: An overview of Vietnam and Hanoi. Urban Geography, 21, 745−757. Doi: 10.2747/0272-3638.21.8.745.

Taubenböck H., Weigand M., Esch T., Staab J., Wurm M., Mast J., Dech S., 2019. A new ranking of the world's largest cities - Do administrative units obscure morphological realities? Remote Sensing of Environment, 232, 111353. https://doi.org/10.1016/j.rse.2019.111353.

Tien N.N., Thanh N.T., Dang V.H., Vinh V.D., 2016. The initial results of studying the suspended sediment dynamics during the flood season in the Hau River mouth area. Vietnam Journal of Marine Science and Technology, 16, 122−128. https://doi.org/10.15625/1859-3097/16/2/6480.

Van Pham T., Bui D.X., Do T.A.T., Do A.N.T., 2025. Assessing flood susceptibility in Hanoi using machine learning and remote sensing: implications for urban health and resilience. Natural Hazards, 1−22. https://doi.org/10.1007/s11069-025-07211-5.

Versini P.-A., Gaume E., Andrieu H., 2010. Assessment of the susceptibility of roads to flooding based on geographical information-test in a flash flood prone area (the Gard region, France). Natural Hazards and Earth System Sciences, 10, 793−803. https://doi.org/10.5194/nhess-10-793-2010.

Viet H.P., An B.N., Viet H.P., Ngoc A.N., Phuong T.G.T., Cao H.N., Tien B.D., 2025. Enhancing spatial prediction of flash floods with Balancing Composite Motion Optimized Random Forest: A case study in High-Frequency Torrential rainfall area. Vietnam Journal of Earth Sciences, 47(1), 118−135. https://doi.org/10.15625/2615-9783/22189.

Wang Z., Lai C., Chen X., Yang B., Zhao S., Bai X., 2015. Flood hazard risk assessment model based on random forest. Journal of Hydrology, 527, 1130−1141. https://doi.org/10.1016/j.jhydrol.2015.06.008.

Wu Y., She D., Xia J., Song J., Xiao T., Zhou Y., 2023. The quantitative assessment of impact of pumping capacity and LID on urban flood susceptibility based on machine learning. Journal of Hydrology, 617, 129116. https://doi.org/10.1016/j.jhydrol.2023.129116.

Yuwono B.D., Sabri L., Wijaya A., Awaluddin M., 2024. Assessment of Flood Risk Induced by Land Subsidence Using Machine Learning. Indonesian Journal of Geography, 56. https://doi.org/10.22146/ijg.94726.

APPENDIX

Table 1A. Contribution to flood susceptibility

Conditioning factors	**Contribution to flood susceptibility**
Elevation	Elevation is a fundamental factor in assessing flood risk, as it governs the storage and distribution of runoff. Water from high-altitude areas (hills and mountains) often flows into low-lying areas, making them more susceptible to flooding (Albano and Adamowski, 2025).
Slope	Slope is also a key factor as it directly affects the velocity, direction, and volume of runoff. Areas with steep slopes generate fast, strong flows, increasing the risk of flooding in lower areas, while flat or gentle areas are prone to water stagnation during prolonged heavy rains (Damayanti et al., 2024).
Curvature, Aspect	Topographic curvature reflects the surface's convexity, concavity, and flatness, thereby affecting storage capacity. Concave or gently sloping areas near converging streams, where drainage density is high, are often more susceptible to flooding than convex or flat areas (Al-Juaidi, 2023).
NDVI, NDWI	NDVI reflects vegetation density and health, which regulate surface runoff: high NDVI values indicate better infiltration and reduced flood risk, while low NDVI values indicate reduced vegetation cover and increased runoff (Khosravi et al., 2019). NDWI represents surface water content; high values are generally associated with low-lying areas prone to waterlogging and a high risk of flooding, while low values reflect cooler conditions and a lower risk of flooding (Ahmed and Akter, 2017).
NDBI, IBM, NDISI and UI	The NDBI, IBM, NDISI and UI indices represent the degree of urbanisation and the proportion of impervious surfaces. High values reflect increased construction and grey infrastructure, which increases the risk of runoff and flooding. On the contrary, low values indicate the presence of many permeable surfaces, which help reduce flood risk (Hoang and Liou, 2024; Rahmati et al., 2020).
Distance from river	Distance from river plays a key role in assessing the effects of urban growth on flood susceptibility. Areas located close proximity to rivers are more likely to be affected by flooding due to direct exposure to overflows, while this risk gradually decreases as distance from the river increases (Liuzzo et al., 2019).
Distance from road and Pumping station density	Distance from road plays an important role in the effects of urban growth on flood susceptibility, as it directly influences runoff and can create areas of water accumulation. This factor is also associated with high urban density, leading to increased impervious surfaces and, consequently, greater flood risk (Versini et al., 2010). On the contrary, the density of the pumping station helps reduce flood vulnerability, as higher density allows for faster water evacuation; however, its effectiveness is highly dependent on the capacity and distribution of the pumping system (Wu et al., 2023).
Land use (LULC)	Land use (LULC) plays an important role in determining areas highly susceptible to flooding, as changes in land use directly influence surface impermeability and, consequently, soil infiltration capacity (Nguyen et al., 2024a).

Table 2A. Land use/land cover type in Hanoi in 2014 and 2024

Land use type	**Area (km²) in 2014**	%	**Area (km²) in 2024**	%
Aquaculture	173.530718	5.19	288.03444	8.62
Forest	319.55843	9.56	204.53017	6.12
Water	285.765851	8.55	187.379918	5.60
Cropland	2137.985683	63.96	1902.20434	56.90
Barren	48.242896	1.44	42.183122	1.26
Urban	255.625298	7.65	418.651382	12.52
Grass	122.643023	3.67	300.318475	8.98

Table 3A. The distribution of flood susceptibility class in Ha Noi city

		Very low (km²)	**Low (km²)**	**Moderate (km²)**	**High (km²)**	**Very high (km²)**
ADB	2014	439.9218	1442.08	1008.671	236.8917	216.1845
	2024	337.7367	1234.267	1336.14	176.7978	274.7871
XGB	2014	453.9843	825.5007	785.9205	678.0762	600.2667
	2024	768.6108	571.4208	620.1423	756.3987	643.1553
DNN	2014	418.2588	683.9658	759.3102	692.9082	789.3054
	2024	475.5591	675.0729	717.9417	794.3094	696.8457
RF	2014	912.7368	738.072	416.4012	320.1966	956.3418
	2024	337.7367	1234.267	1336.14	176.7978	274.7871

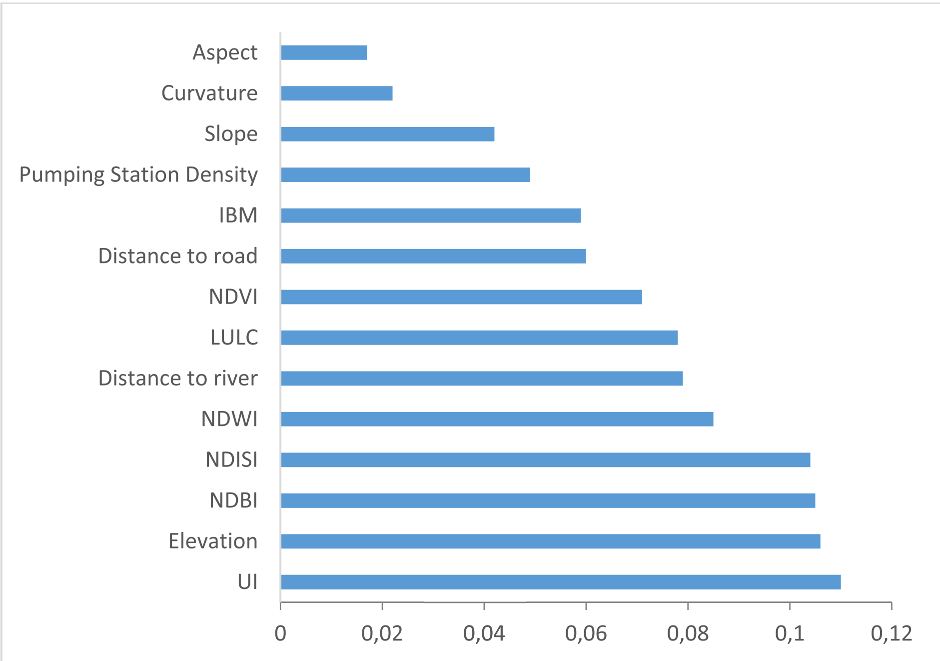

Figure 1A. Importance of factors, using RF

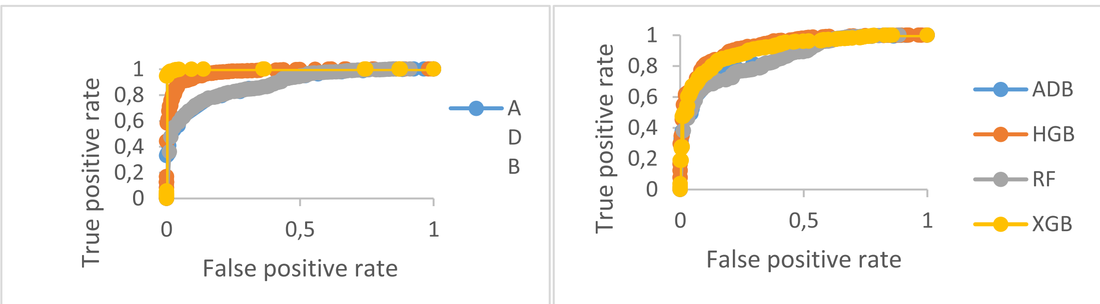

Figure 2A. AUC value for training and validation processus for ADB, DNN, RF, and XGB

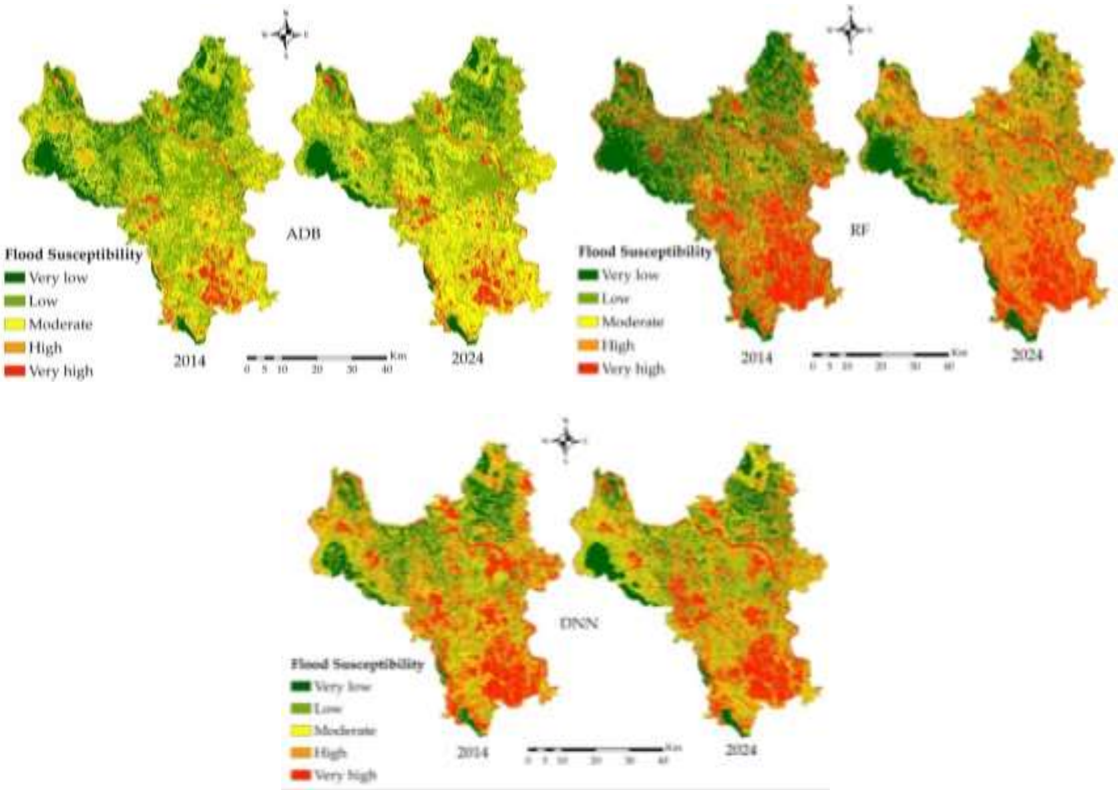

Figure 3A. Flood susceptibility in 2014 and 2024 by ADB, RF, XGB, and DNN in Hanoi city