Gonbad Kavous University

Journal of New Approaches in Water Engineering and Environment

Online ISSN: 2981-0280

Homepage: https://nawee.gonbad.ac.ir

Performance evaluation of the random forest model in flood hazard assessment of the Kashkan Watershed

Atefe Amiri [1], Afshin Honarbaksh [2*], Rafat Zare Bidaki [3], Hossein Zeinivand [4]

[1] Department of Natural Engineering, Faculty of Natural Resources, Shahrekord University, Shahrekord, Iran.
[2] Department of Natural Engineering, Faculty of Natural Resources, Shahrekord University, Shahrekord, Iran.
[3] Department of Natural Engineering, Faculty of Natural Resources, Shahrekord University, Shahrekord, Iran.
[4] Department of Natural Engineering, Faculty of Natural Resources, Lorestan University, Khorramabad, Iran.

Article Info

Article type:

Research Article

Article history:

Received 25 February 2025

Revised form 29 March 2025

Accepted 20 June 2025

Published online 26 June 2025

Keywords:

Flood control

flood map

AUC

RF

ABSTRACT

Objective: Flooding is one of the most dangerous natural events worldwide, caused by a combination of climatic, hydrological, geomorphological, and geological factors. Floods can occur due to heavy rainfall, prolonged rainfall, rapid snowmelt, or dam failure. Regardless of the cause, floods lead to widespread destruction and damage to human societies and infrastructure. Given the severe risks, assessing flood hazards has become essential. Flood sensitivity maps are useful tools to analyze and manage flood-prone areas.

Material and Methods: This study aims to identify flood-sensitive zones using the Random Forest (RF) model in the Kashkan Basin, Lorestan Province. Thirteen flood-related factors and a map of past flood events were used. Of the 58 recorded flood locations, 73% were used for model training and 27% for validation.

Results and Discusion: The analysis revealed that proximity to rivers, elevation, slope, and roughness index are the most influential factors in the region's flooding. The RF model's performance was evaluated using the ROC index, which scored 0.97, indicating excellent model accuracy in generating the flood sensitivity map.

Conclusions: Flooding is driven by various environmental and human factors. Based on the flood risk prediction map, proper management strategies can be implemented to reduce damage and casualties caused by floods.

*Corresponding author, Email: afshin.honarbakhsh@gmail.com

Cite this article: Amiri, Atefe, Honarbaksh, Afshin, Zare Bidaki, Rafat, & Zeinivand, Hossein. (2025). Performance evaluation of the random forest model in flood hazard assessment of the Kashkan Watershed. Journal of New Approaches in Water Engineering and Environment, 4(2), 131-159.

https://doi.org/10.22034/nawee.2025.508845.1143.

DOI: https://doi.org/10.22034/nawee.2025.508845.1143

1. Introduction

Every day, many events occur all over the world, some of which turn into natural disasters. Natural disasters are catastrophes or calamities caused by the occurrence of a dangerous natural phenomenon, such as a flood, drought, earthquake, landslide, storm, or volcano, which causes a lot of financial and human damage to human societies (Fernandez and Lutz, 2010). Among natural disasters, floods are one of the most important and destructive disasters, and due to the lack of information and knowledge about flood casualties, the accurate assessment of their risks is failing on different scales (Grahn and Nyberg, 2017). Factors such as unprincipled changes in land use, destruction of vegetation in an area, reduction of permeable levels, and other human interventions in an area will intensify the risk of flooding (Donyaii et al., 2021). Since various factors are involved in the flooding process, it is difficult to predict it. The phenomenon of flooding in its current form in our country is more a result of the disruption of the natural balance and the geographical conditions of the region than due to heavy rainfall. So much so that even normal rainfall can cause flooding (Khairizadeh et al., 2013). One of the important methods for controlling and reducing surface runoff to delay flooding is watershed management operations. Another method for reducing flood risks is flood zoning, which provides valuable information about the nature of floods, their effects on floodplain areas, and the determination of river boundaries. As a result, it enables the timely issuance of warnings during flood risks and facilitates rescue and relief operations (Khairizadeh et al., 2013). Despite the efforts of experts, decision-makers, stakeholders, and government agencies in recent decades to reduce the impacts of flooding, the number of incidents and related economic and human casualties is increasing globally. This phenomenon is not only prevalent in developing countries but is the most common natural hazard worldwide (Hazarika et al., 2018).

In recent years, investigating flood sensitivity has become an important research topic, and various researchers have investigated the use of different models for the spatial prediction of flood-prone areas.

Avand et al. (2020) used generalized Bayesian and linear random forest models in the Tajan watershed to create a flood sensitivity map. Thirteen factors, including elevation, distance from the river, and slope, were key. The RF model performed better than Bayesian in accuracy. Kiyaniasl et al. (2022) compared RF and SVM algorithms for flood sensitivity maps in the Maroon watershed. Sixteen factors were used, and accuracy was evaluated with ROC. RF achieved 0.997 accuracy, outperforming SVM's 0.947. Satarzade et al. (2022) in the Karkheh watershed used a hybrid RF and weight-of-evidence model to map flood-prone areas, with 11 predictors. ROC assessed accuracy, revealing that 20.49% of the area had above-average flood susceptibility. Mahdizadeh Gharakhanlou and Perez (2023) in Canada used six machine learning models (DT, RF, MLP-NN, AdaBoost, LR, SVM) to predict flood-prone areas with spatial variables. RF had the highest accuracy; lithology and drainage density were key factors. Youssef et al. (2023) in Egypt compared LR, EGB, and RF models for flood sensitivity, using remote sensing, topography, geology, meteorology, and field visits. RF was identified as the best algorithm, with EGB and LR following. Li et al. (2023) in China used deep learning and group learning models (DL, FC-DL, RF-DL, RSS-DL) for flood sensitivity mapping.

Evaluation metrics (AUC, kappa, ACC, F1) showed all models performed acceptably. Tich Vu et al. (2023) in Vietnam used SVM and RF with remote sensing to study land use's effect on flood susceptibility. RF achieved an AUC of 0.98, outperforming SVM's 0.97 in predicting land use and climate change impacts on floods.

The current research aims to determine the flood potential of the Kashkan watershed using a random forest algorithm. The preparation of flood sensitivity maps using new methods has been a focus of recent research. Creating a flood sensitivity map for the Kashkan watershed is particularly important due to its vulnerability to flooding. Due to the large size and lack of data, and sometimes the unavailability of some areas of this basin, it is not possible to use hydrological models, or it requires a lot of time and money.

Therefore, in this research, the RF model, which is one of the best machine learning algorithms, has been used in Python to prepare a flood sensitivity map. Also, various factors are effective in flooding a watershed, and the use of a wide range of these data will increase the accuracy of the results. In this regard, the participation of some environmental factors in the flood phenomenon and the preparation of a flood zoning map are other innovative aspects of this research.

2. Materials and methods

2.1. The Study area

Kashkan watershed is located in Lorestan province, with coordinates 47° 12″ to 48° 58″ and 32° 55″ to 34° 02″ in the western region of Iran. The area of the studied basin is 9492 square kilometers, and its elevation is between 498 and 3627 meters above sea level (Figure 1). The average annual rainfall and temperature of the basin are 965 mm and 33.3 degrees Celsius, respectively. A large part of land use in the region is rainfed agriculture. Geologically, a huge part of the region belongs to the Quaternary Formation. Inceptisols are the most abundant type of soil in the region, with a relatively recent origin and poor appearance of the horizons due to low accumulation of clay and organic matter. In terms of climate, the studied area has a cold and dry climate. The four main cities of Aleshtar, Kohdasht, Khorram Abad, and Poldokhtar are located in the four corners of the area. Among the permanent rivers of Lorestan province, the Keshkan River is the most flood-prone river in the province, with 17 floods over 1000 cubic meters per second in the last 68 years. The biggest flood recorded in the province was related to the flood on April 12, 2018, with a peak discharge of 6500 cubic meters per second, which caused significant damage (Based on the regional water information of Lorestan province.).

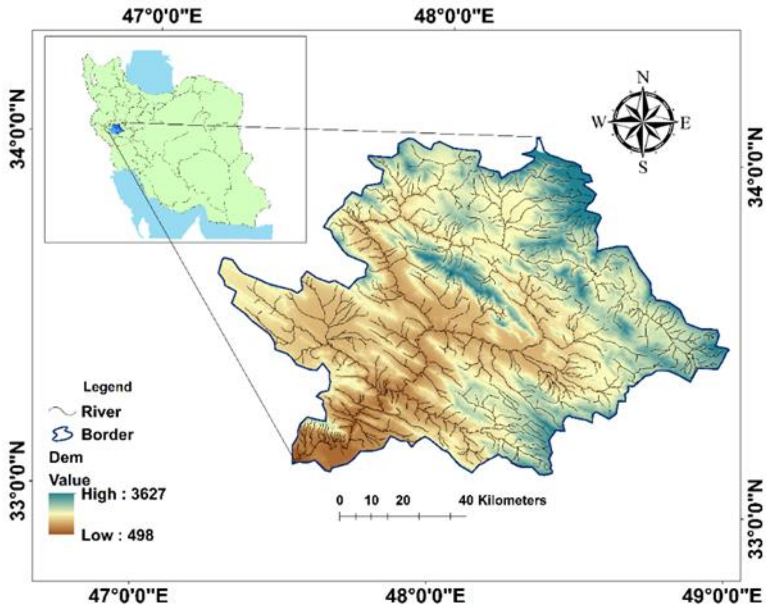

Figure 1- Location of the Kashkan basin in Lorestan province

2.2. Map of past flood locations

Flood locations are an important part of the relationship between the flood event and the factors that caused it. Historical flood events are considered a basis for predicting the occurrence of floods in the future, so that areas close to past events are highly sensitive to flooding. The Kashkan watershed is one of the basins with high flood potential. Thus, 58 flood locations in the study area have been recorded by the regional water organization of Lorestan province (Figure 2), 73% of them were used for model training and 27% for model validation.58 non-flood locations were also selected by using a topographic map and Google Earth software according to areas such as hills and mountains where the flood cannot progress.

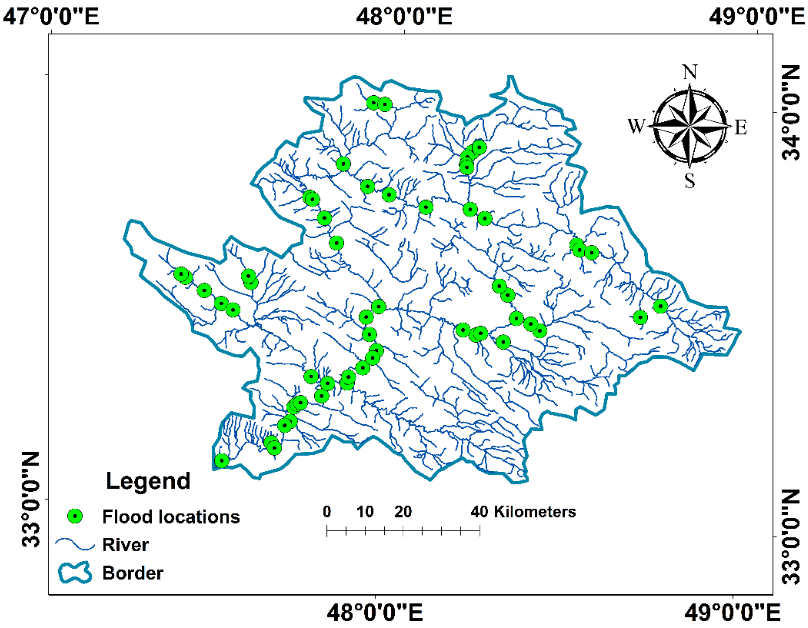

Figure 2- Flood locations in the Kashkan basin

2.3. Conditioning factors in the preparation of the flood sensitivity map

To prepare a flood sustainability map or generally to produce a model to assess the damage of natural disasters, a set of conditioning factors must be defined. In this research, 13 factors affecting the occurrence of floods, including elevation, slope, slope aspect, distance from the river, drainage density, topographic wetness index (TWI), topographic position index (TPI), slope curvature, terrain roughness index (TRI), geology, precipitation, soil, and land use have been used. For the preparation of elevation maps and slope direction, a Digital Elevation Model (DEM) with a resolution of 30 meters and ArcGIS 10.8 software were used. The earth curvature map was also prepared based on the digital elevation model and using the Curvature package of ArcGIS 10.8 software. A digital elevation model and the Slope add-on package were used in ArcGIS 10.8 software to prepare the slope map. The map of distance from neighboring rivers was prepared based on the digital layer of the stream network of the Koshkan watershed and using the Euclidean Distance function in ArcGIS 10.8 software. According to the role of drainage density in water transfer and its infiltration rate in the ground, the drainage density map was also prepared based on the digital layer of the waterway network in ArcGIS 10.8 software. Due to the effect of soil characteristics on the permeability of the land and its effect on the production of runoff, the soil map of the region was also used in this research. Land use is one of the effective factors in determining the runoff and flood potential of a watershed.

The land use map was prepared based on digitalization and modification of the land use map of the General Department of Natural Resources of Lorestan province. In addition to the characteristics of the geological structure, the geological condition also indicates the type of lithology of the basin. The lithology map of the studied area was prepared using the geological map of the country (scale 1:100,000). Precipitation is one of the most important factors affecting the occurrence of floods in a region. In this research, to prepare the precipitation map of the target area, the 20-year precipitation statistics of 11 meteorological and synoptic stations (Sarab Said Ali, Kakareza, Dehno, Cham Anjir, Doab Visian, Afrineh, Kohdasht, Noorabad, Aleshatar, Silakhor, Poldokhtar) in and around the basin have been used. The Kriging interpolation method was used to prepare the precipitation map. Among other factors influencing the determination of flood potential in the watershed, we can mention the topographic wetness index (TWI), topographic position index (TPI), and terrain roughness index (TRI). TWI is defined based on Equation 1.

$$TWI = {A_s}/{tan_\beta} \qquad (1)$$

Where A_sIs the specific area of the watershed and β is the land slope (expressed in percentage).TPI is an index of the landscape's topographic status, calculated by the elevation difference of each cell compared to the average elevation of its neighboring cells. Negative TPI values indicate cells that have a lower elevation than the surrounding cells, while positive values indicate cells that are higher in elevation compared to the cells that surround them (Wehrhan and Sommer, 2021). The Terrain Ruggedness Index (TRI) is used to express the elevation difference between adjacent cells of a DEM. This raster function template is used to generate a visual representation of the TRI with elevation data. The results are interpreted as follows: 0-80m is considered to represent a level terrain surface 81-116m represents a nearly level surface 117-161m represents a slightly rugged surface 162-239m represents an intermediately rugged surface 240-497m represents a moderately rugged surface 498-958m represents a highly rugged surface 959-4367m represents an extremely rugged surface. The digital height model and SAGA GIS software were used to prepare the map of the three desired indicators.

2.4. Random Forest Model (RF)

Tree-based methods are non-parametric (model-free) statistical methods for performing classification analysis and regression analysis using the algorithm of recursive extraction (Breiman, 1984). Random forests are a modern type of tree-based method that includes a multitude of classification and regression trees (Breiman, 2001). The most important feature of random forests is their high performance in measuring the importance of variables to determine what role each variable plays in predicting the response. The random forest produces many decision trees. To classify a new object, the input vector is placed at the end of each random forest tree, each tree produces a classification and it is said that this tree gives a class (vote) to that tree. The forest resulting from the classification that has the most votes is selected (among all the trees in the forest). The Random Forest (RF) algorithm is based on a group of decision trees and is currently one of the best learning algorithms (Lee et al., 2017). In this research, Python programming language was used to implement the RF model, and the model was executed in the PyCharm 2020.3 software environment. For the implementation of the model, the training and validation values were initially set manually (with a 73 to 27 ratio). Then, the RF model will randomly select the defined data through a random search.

2.5. Evaluation of the model

The relative performance detection curve (ROC) was used to evaluate the performance of the prepared model. The ROC curve is a graphical display of the trade-off between negative and positive error rates for each possible value of the cutoffs. The area under the ROC curve (AUC) represents the prediction value of the system by describing its ability to correctly estimate the events that have occurred (flood occurrence) and the non-occurrence of the event (non-flood occurrence). The area under this curve quantitatively shows the accuracy of flood forecasting maps. If the area under the curve is between 0.5 and 0.6, it indicates the randomness of the results, If the area under the curve is between 0.7 and 0.8, the results have good validity, if the area under the curve is between 0.8 and 0.9, the results have very good validity, and if the area under the curve is more than 0.9, the validity of the obtained results is at an excellent level (Tehrany et al., 2015). Also, in this research, the accuracy and precision criteria were used to evaluate the results of the random forest model. The Accuracy criterion is equal to the number of cases that we predicted correctly, which we call True Positive, divided by the total number of predictions that have been made. TruePositive represents the number of correct predictions (Relation 2).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad (2)$$

The accuracy evaluation criterion for measuring the ratio of correct predictions is positive. In other words, accuracy provides a measure of the model's ability to identify correct positive samples (Relation 3).

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (3)$$

True Positive (TP): Correctly identified as positive. False Negative (FN): The positive sample was incorrectly identified as negative. True Negative (TN): Correctly identified as negative. False Positive (FP): Negative sample incorrectly identified as positive.

3. Results and discussion

Multiple factors in a watershed affect surface water flow, each with a different potential impact on water flow performance and the occurrence of floods. In this study, 13 factors influencing flood occurrence were used, as shown in Figure 3. Determining the most important and effective factors impacting flood occurrence is essential in modeling. In this research, the prioritization of factors affecting floods in the Kashkan watershed area is presented in Figure 4. The results obtained from the machine learning model indicated that the factors of distance from the river (17.28), elevation (31.16), slope (44.12), and TRI (10.53) have the greatest influence on the flood sensitivity map in the studied region, respectively. The factor of distance from the river and elevation has an inverse relationship with the flooding phenomenon, meaning that the greater the distance from the river and elevation, the lower the likelihood of flooding and its associated damage. The slope factor also has an inverse relationship with flooding occurrence; the steeper the area, the lower the likelihood of flooding. The reason for this might be attributed to the possibility of more focused water flow in low slopes, where the flood potential increases as we move towards the lower part of the watershed due to the area's gentle slope. The terrain ruggedness index in highland areas has a positive value, while flat or low-gradient areas have a ruggedness index close to zero. The higher the terrain ruggedness index, the lower the likelihood of flooding and flood-related risks. As shown in the terrain ruggedness index map, areas with a ruggedness index approaching zero have the highest flood sensitivity in the flood sensitivity map.

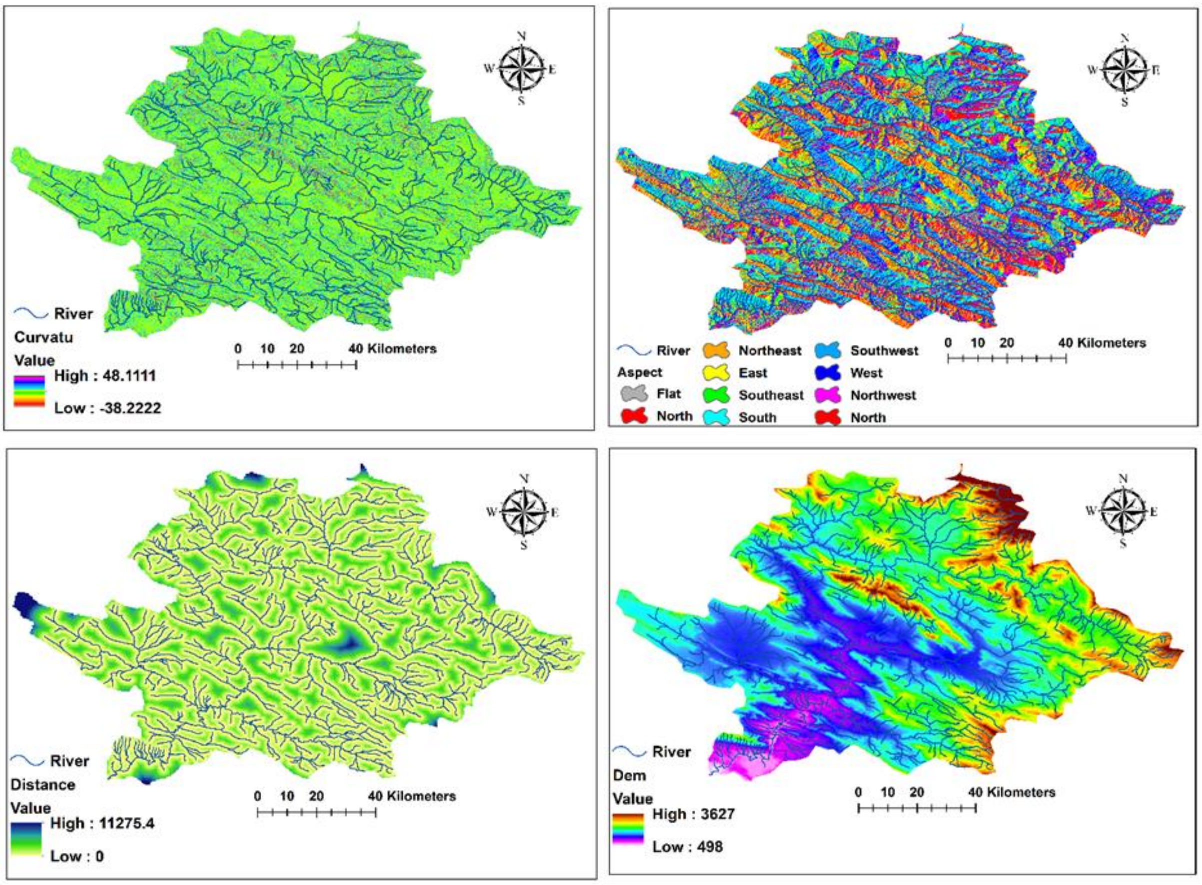

Figure 3- Factors affecting the preparation of a flood susceptibility map

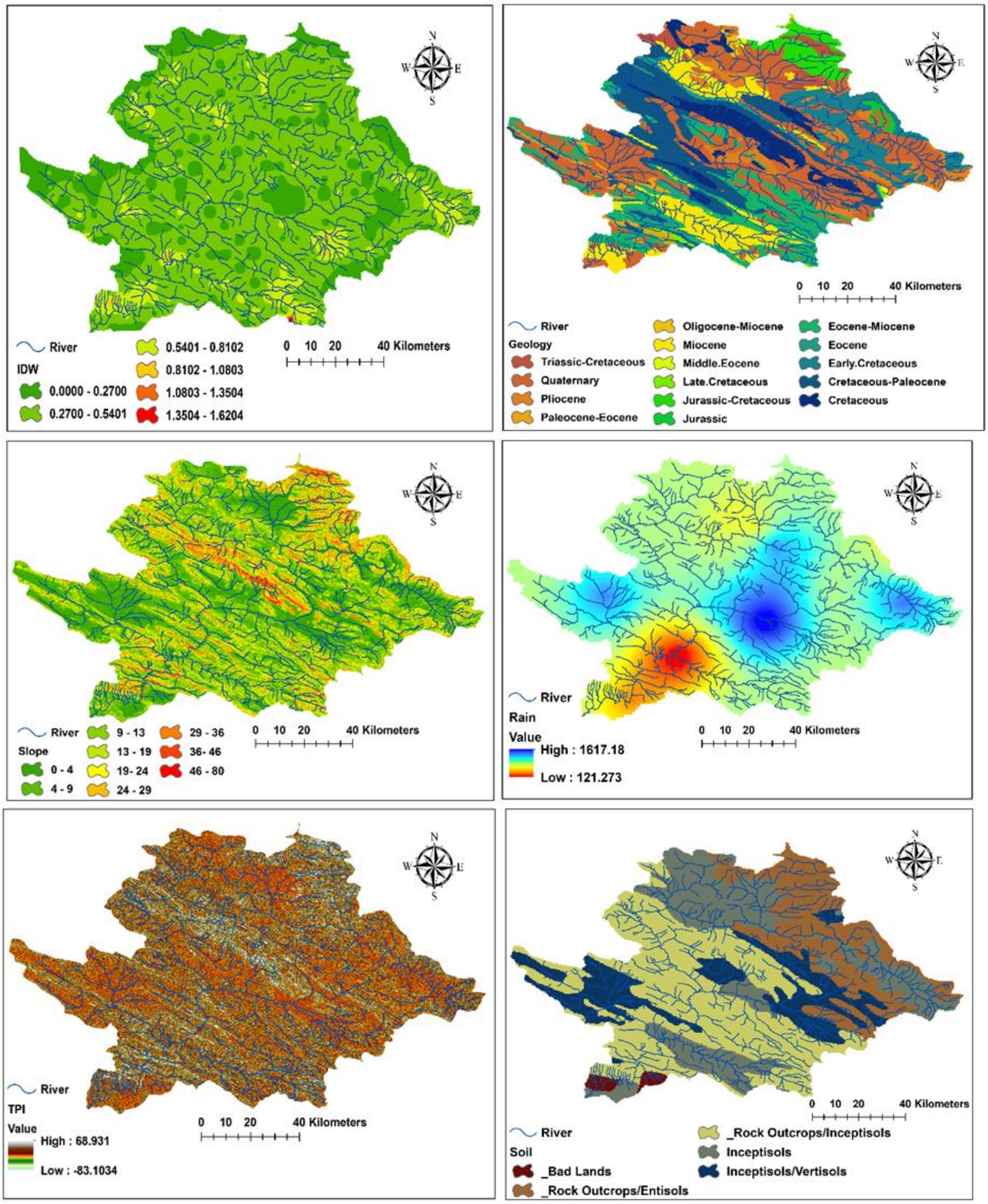

Figure 3- Factors affecting the preparation of flood susceptibility map

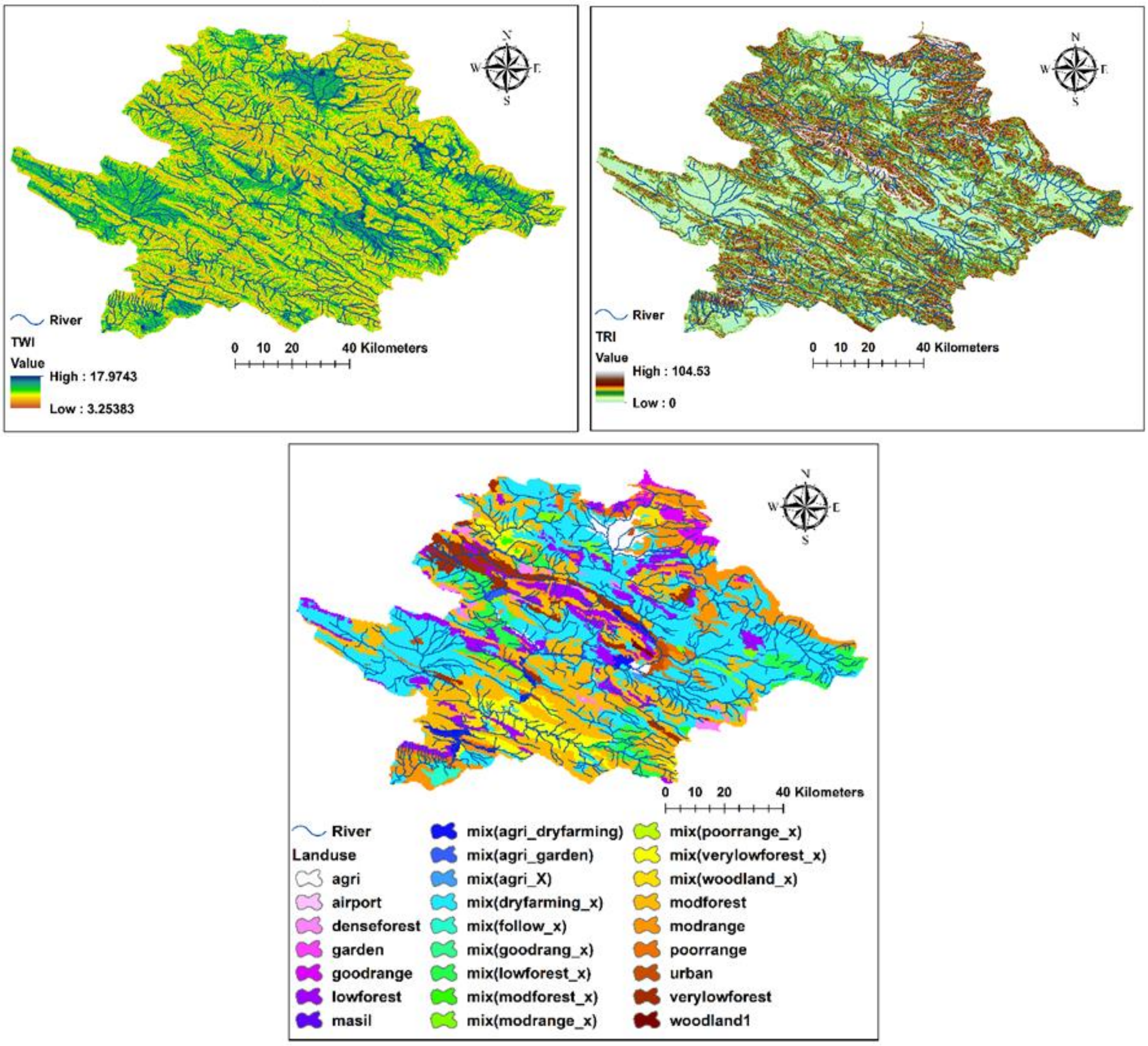

Figure 3- Factors affecting the preparation of flood susceptibility map

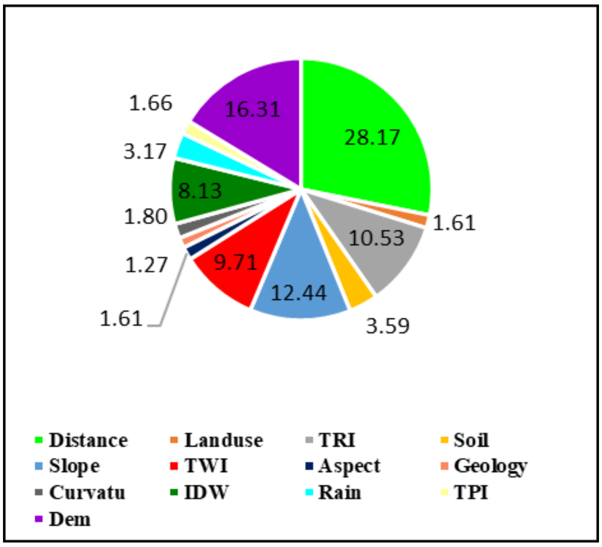

Figure 4- Importance of factors used in the preparation of flood susceptibility map

The flood sensitivity map was prepared using the RF model, and ultimately, based on various sources and using the Natural Breaks method, it was divided into five classes: very low sensitivity, low sensitivity, moderate sensitivity, high sensitivity, and very high sensitivity. Figure 5 shows the flood sensitivity map of the Kashkan basin.

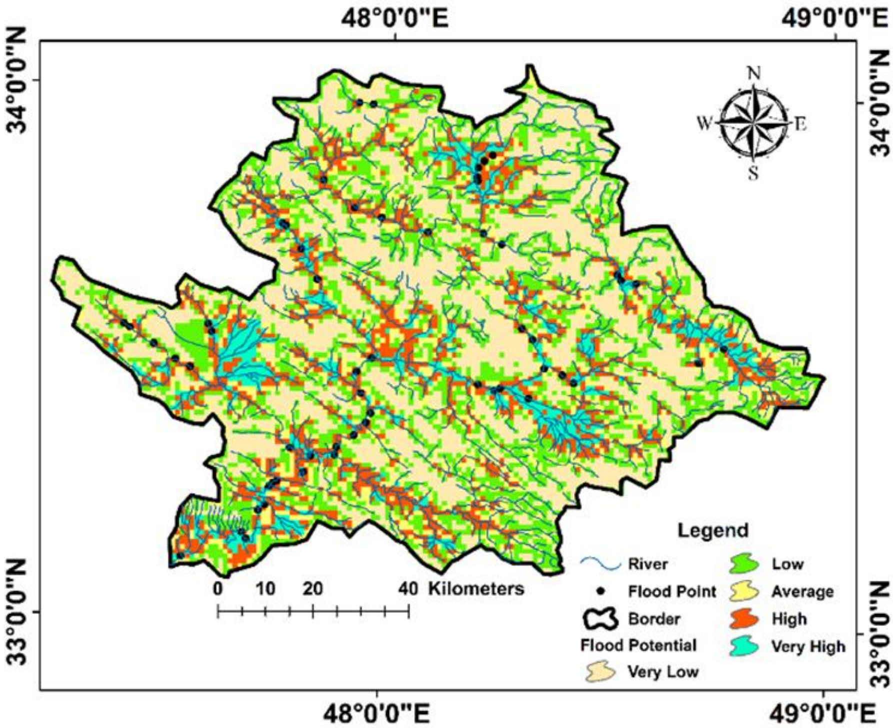

Figure5- Flood risk potential zoning map

Figure 6 shows the area of each flood-sensitive floor in the RF model. The results showed that in the RF model, the highest level of flood sensitivity is related to the very low floor with an area of 3011 square kilometers and the lowest level is related to the very high floor with an area of 749 square kilometers.

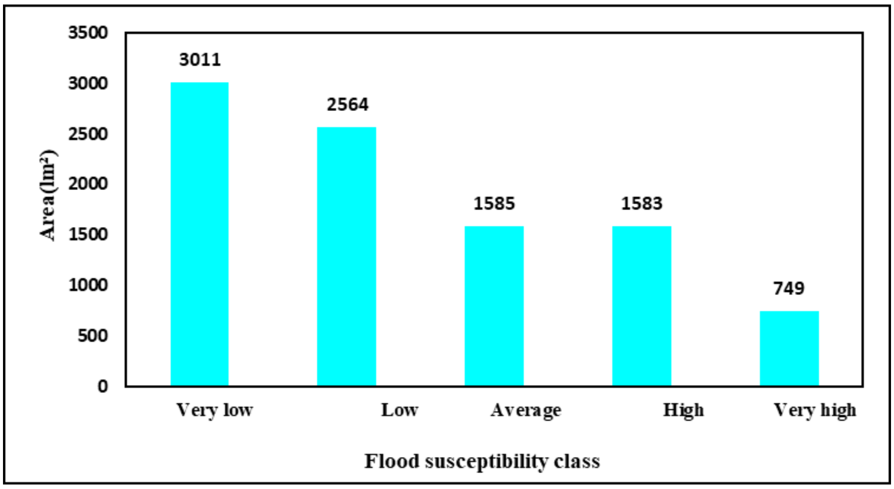

Figure 6- The area of sensitivity classes

The ROC curve method was used to evaluate the accuracy of the flood and inundation potential prediction map. The evaluation results showed that the area under the curve was 0.97 in the validation stage, which indicates the model's excellent performance in assessing the accuracy of the flood susceptibility map of the studied area.

Table 1- Evaluation criteria of the used model

RF	Precision	Accuracy	AUC
Test	0.941	0.971	0.97
Training	0.979	0.986	0.99

Several parameters within a watershed influence water flow at the basin scale, with each factor having a varying potential to impact flow dynamics and the occurrence of floods. Determining the most important and effective factors affecting floods is very necessary for modeling. In this study, to estimate the flood sensitivity map, 13 factors affecting the flood have been used. Among the investigated factors, the parameters of distance from the river, elevation, slope, and TRI are the most important influencing parameters in the flood sensitivity of the studied area. The research results with the studies of Siahkamri et al. (2016), Avand et al. (2020), and Arab Ameri et al. (2017) are consistent. The parameters of distance from the river, elevation, slope, and terrain roughness index have an opposite relationship with flooding in such a way that the greater the distance from the river, elevation, slope, and terrain roughness index, the more the probability of flooding, and as a result, the related damages are reduced. Also, the results of the investigations showed that the highest level of sensitivity was obtained for the class with low sensitivity, and the lowest level was obtained for the class with very high sensitivity. Areas with high sensitivity mostly include areas close to the river with a low slope, and most of these areas are prone and sensitive to floods due to changes in land use and human manipulation, which have caused the increase of agricultural and residential lands. On the other hand, reducing the distance from the main boundary of the river and exposing properties, lands, and people have been other effective factors in increasing damage during floods in this area. According to the soil map of the basin, areas where the soil type is Inceptisols and Vertisols have the highest flood potential. Precipitation is also considered another influential factor in the flood potential of the basin, in such a way that areas with higher rainfall have a greater flood potential. In this study, the results of RF model evaluation criteria showed that the performance of the model for the spatial prediction of the flood capacity of the studied area was very good, which in the study of Chang et al. (2020), Avand et al. (2020) also, the appropriate performance of machine learning models in the field of spatial prediction of flood capacity has been investigated and confirmed.

4. Conclusion

Unfortunately, in Lorestan province, flooding is a dominant and damaging phenomenon that causes great damage to people every year. Through regional analysis and using different methods, it is possible to identify flood-sensitive areas and reduce the damage they cause. The purpose of this research is to determine flood-sensitive areas using a random forest model. The most important results of the current research are as follows: 1) The factors of distance from the river, elevation, slope, and terrain roughness index are the most important parameters affecting the occurrence of floods in this basin. 2) The highest level of the studied basin is related to the class with low sensitivity, and the lowest level is related to the class with very high sensitivity. 3) Since the flood capacity map obtained from modeling in this research was evaluated using real and observed flood event data, its results are reliable to a large extent and can be used in flood management of the studied basin. Executive measures such as the construction of watershed structures, dams to curb the flow, flood walls in rivers and revitalization of vegetation on the watershed level, etc., and management measures such as the correct land use pattern, preventing encroachment on the legal boundaries of rivers, legalizing the harvesting of river materials, etc. can be carried out in each part of the watershed based on the flood capacity map. The results of the study show that flooding is the result of various environmental and human factors, and based on the flood risk prediction map, appropriate management measures can be taken to reduce losses and casualties caused by floods.

This research has significant practical implications for flood risk management in Lorestan province and similar regions. By accurately identifying flood-sensitive areas, the findings of this study can assist local authorities, urban planners, and environmental managers in implementing targeted interventions. The flood sensitivity map derived from this study can be used to prioritize areas for flood mitigation efforts, such as the construction of flood barriers, water retention structures, or vegetation restoration in vulnerable regions.

In addition, the results can guide land use planning by identifying areas where development should be restricted due to high flood risk. These findings can also inform emergency preparedness strategies, ensuring that at-risk populations are adequately warned and protected during flood events. By integrating the flood risk prediction model into urban and regional planning, authorities can proactively reduce damage and improve resilience against future flood events.

In the end, it can be said that obtaining an accurate and reasonable spatial forecasting map can help city managers and planners identify flood-sensitive areas for the critical management of susceptible regions, contributing to more sustainable and safer development practices.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

Dr. Mehdi Malekian is thanked and appreciated for her valuable and unwavering guidance in conducting this research.

Ethical Considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare no conflict of interest.

References

Arab Ameri, A., Pourghasemi, H.R., Shirani, K. 2017. Flood susceptibility zoning using a new hybrid Bayesian theory-hierarchical analysis process (Case study: Neka Watershed, Mazandaran Province). Ecohydrology, 4(2): 447-462.

Avand, M., Moradi, H.R., Ramazanzadeh, M. 2020. Flood susceptibility mapping using random forest machine learning and generalized Bayesian linear model. Environ. Water Engineering, 6(1): 83–95.

Breiman, L. 1984. Classification and regression trees CA, Wadsworth International Groups. Handbook of data mining and knowledge discovery, 248-276.

Breiman, L. 2001. Random forests. Machine Learning, 45(1): 5-32.

Chang, L.C., Chang, F.J. 2020. IoT-based flood depth sensors in artificial intelligent urban flood warning systems. EGU General Assembly, held online, 4-8 pp 12523

Donyaii, A., Sarraf, A.P., Ahmadi, H. 2021. Optimization of reservoir dam operation using gray wolf, crow search, and whale algorithms based on the solution of the nonlinear programming model. JWSS - Journal of Water and Soil Science, 24(4): 159-175.

Fernandez, D.S., Lutz, M.A. 2010. Urban flood hazard zoning in Tucuman Province, Argentina, using GIS and multicriteria decision analysis. Engineering Geology, 111: 90-98.

Grahn T., Nyberg L. 2017. Assessment of pluvial flood exposure and vulnerability of residential areas. International Journal of Disaster Risk Reduction, 21: 367–375.

Hazarika, N., Barman, D., Das, A.K., Sarma, A.K., Borah, S.B. 2018. Assessing and mapping flood hazard, and risk in the Upper Brahmaputra River valley using stakeholders' knowledge and multicriteria evaluation (MCE). 11(52): 5700-5716.

Kiani, Asl M.A., Moteshaffeh, B., Roshaan, S.H. 2023. Evaluation of Machine Learning Algorithms (RF and SVM) in Producing Flood Susceptibility Mapping Maroon Watershed. Jwmseir, 17 (61):41-51.

Khairizadeh, M., Maleki, J., Amunia, H. 2013. Flood hazard potential zoning in Marduk ChaiCatchment using ANP Model.Quantitative Geomorphology, 1(3): 39-65.

Li, Y., Hong, H. 2023. Modeling flood susceptibility based on deep learning coupling with ensemble learning models. Journal of Environmental Management, Volume 325, Part A, art. no. 116450.

Lee Sunmin Kim, J.C., Jung, H.S., Lee, M.J., Lee, S. 2017 . Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. Geomatics, Natural Hazards and Risk, 8(2): 1185–1203.

Mahdizadeh Gharakhanlou, N., Perez, L. 2023. Flood susceptible prediction through the use of geospatial variables and machine learning methods. Journal of Hydrology, 617 (C).

Siahkamri, S., Zainivand, H. 2016. The potential of flood-prone areas using statistical index model and weight of evidence (Case study: Madarsoo Watershed, Golestan). Remote Sens GIS Nat Resour, 7(4): 116 - 133.

Satarzadeh, E., Sarraf, A.P., Hajikandi, H., Sadeghian, M.S. 2022. Spatial Prediction of Flood Susceptible Areas in Karkheh Watershed of Lorestan Province Using the Combined Random Forest – Weight of Evidence Model. Watershed research, *36*(2): 87-103

Tehrany, M.S., Pradhan, B., Mansor, S., Ahmad, N. 2015. Flood susceptibility assessment using a GIS-based support vector machine model with different kernel types. Catena, 125(1):91–101.

Tich Vu, V., Duy Nguyen, H., Lan Vu, P., Cuong Ha, M., Dong Bui, V., Oanh Nguyen, T., Hiep Hoang, V., Hue Nguyen, T.K. 2023. Predicting land use effects on flood susceptibility using machine learning and remote sensing in coastal Vietnam.WaterPracticeandTechnology,18(6):1543–1555.

Wehrhan, M., Sommer, M. 2021. A Parsimonious Approach to Estimate Soil Organic Carbon Applying Unmanned Aerial System (UAS) Multispectral Imagery and the Topographic Position Index in a Heterogeneous Soil Landscape. Remote Sensing, 13(18):3557.

Youssef, A.M., Mahdi, A.M., Pourghasemi, H.R. 2023. Optimal flood susceptibility model based on performance comparisons of LR, EGB, and RF algorithms. Nat Hazards, 115: 1071–1096.