Flood risk mapping and performance efficiency evaluation of machine learning algorithms: Best practice in northern Iran

Mahdieh Shirmohammadi[1,2], Saied Pirasteh[*2], Weilian Li[3], Davood Mafi-Gholami[4]
[1]School of Geography, Nanjing Normal University, Nanjing, China
[2]Institute of Artificial Intelligence, Shaoxing University, Shaoxing, 508 West Huancheng Road, Yuecheng District, Zhejiang Province, Postal Code 312000, China. Email: mshirmohammadi.khu@gmail.com; sapirasteh1@usx.edu.cn
[3]Southwest Jiaotong University, Faculty of Geosciences and Engineering, China vgewilliam@my.swjtu.edu.cn
[4]Metrology Research Group, Quality Assessment and Management Systems Research Center, Standard Research Institute, PO Box 31585-163, Karaj, Alborz, Iran, d.mafigholami@sku.ac.ir

Keywords: Flood, Machine Learning, GIS, Risk Mapping, Performance Efficiency.

Abstract

Flooding is one of the most devastating natural hazards, and inadequate management can amplify its impacts, leading to severe social, economic, and environmental consequences. Accurate and efficient flood risk mapping is essential for mitigating these effects and supporting effective disaster management strategies. However, challenges remain in optimizing the accuracy and reliability of machine learning (ML) algorithms for flood susceptibility assessment. In this study, we applied several ML algorithms, including Random Forest (RF), XGBoost (Extreme Gradient Boosting), LightGBM, CatBoost, and Support Vector Machine (SVM), to develop flood risk maps for a region in northern Iran. For the analysis, we selected a comprehensive set of environmental and geographical parameters influencing flood susceptibility. These included the Digital Elevation Model (DEM), slope, aspect, Topographic Wetness Index (TWI), Stream Power Index (SPI), river distance, river density, rainfall, lithology, Normalized Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI), soil texture, and land use. Data processing, feature extraction, and model training were conducted using Python, Google Earth Engine, and ArcGIS. Our results demonstrate a strong level of consistency across the models. XGBoost achieved the highest Area Under the Curve (AUC) of 0.87, closely followed by CatBoost at 0.86, Random Forest (RF), and LightGBM, each reaching 0.85. SVM recorded a slightly lower AUC of 0.82. These findings underscore the robust performance of advanced ML algorithms, particularly ensemble methods with tree-based structures, in flood risk mapping, especially within complex environmental contexts.

1. Introduction

Advancing flood susceptibility mapping has been among the top disaster risk research. Flooding is among the most destructive natural disasters, causing widespread economic, environmental, and societal impacts. Accurate flood susceptibility mapping (FSM) is critical for identifying flood-prone areas and enabling effective disaster management, urban planning, and resource allocation. Northern Iran, particularly the Mazandaran, Gilan, and Golestan provinces, is highly vulnerable to recurrent flooding due to its mountainous terrain, proximity to the Caspian Sea, extensive river networks, and frequent heavy rainfall. These vulnerabilities are compounded by climate change, deforestation, and rapid urbanization, which reduce the land's natural water absorption capacity. The catastrophic March 2019 flood in Golestan Province underscored the urgent need for robust flood management strategies.

Machine learning (ML) techniques have emerged as transformative tools for FSM, offering significant advantages over traditional methods (Li et al., 2022; Habibi et al., 2023). These techniques excel in processing large, complex datasets, capturing nonlinear relationships, and improving predictive accuracy. Researchers have increasingly emphasized the integration of diverse ML algorithms and the incorporation of environmental, social, and economic parameters to enhance FSM accuracy and determine algorithm performance (Demissie et al., 2024; Baida et al., 2024). For example, Demissie et al. (2024) highlighted the robustness of Random Forest (RF) and Extreme Gradient Boosting (XGBoost) in capturing spatial and temporal flood patterns. Their study demonstrated superior predictive performance of these models over Logistic Regression (LR), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and AdaBoost. In another study, Kazemi et al. (2024) evaluated Multivariate Adaptive Regression Splines (MARS), Classification and Regression Trees (CART), TreeNet, and RF, emphasizing their capability to handle nonlinear relationships in FSM.

Furthermore, Singha et al. (2024) used Sentinel-1 SAR data and an array of ML algorithms—such as RF, AdaBoost, rFerns, and XGBoost—integrated with interpretability techniques like SHAP and Boruta for assessing feature importance. In addition, other researchers have taken the advanced parameter optimization techniques and combined RF with Bayesian Optimization (BO), Genetic Algorithms (GA), and Grid Search (GS) for enhanced accuracy in urban FSM (Habibi et al., 2023; Baida et al., 2024). However, they integrated Multi-Criteria Decision Analysis (MCDA) and addressed spatial and hydrological complexities in conjunction with the integration of high-resolution data (Hitouri et al., 2024). They leveraged Synthetic Aperture Radar (SAR) data with ML models (e.g., RF, CART, SVM, XGBoost) to improve flood susceptibility predictions or used multi-input Convolutional Neural Networks (CNNs) with Interferometric SAR (InSAR) and optical imagery for precision mapping of flood-affected areas. Moreover, researchers have studied holistic flood risk assessments (Eini et al., 2020), long-term forecasting, and climate impacts (Amiri et al., 2024; Nguyen et al., 2024; Li et al., 2024).

Nevertheless, building on recent advancements, this study focuses on FSM in a mountainous region characterized by steep terrain and rural settlements. The primary objectives include 1. evaluating and comparing the performance of RF, XGBoost, LightGBM, CatBoost, and SVM to identify the most optimized FSM model. 2. Assessing the influence of environmental and geographical factors on flood susceptibility using feature importance analyses, and 3. improving mapping accuracy to inform targeted flood risk mitigation strategies, particularly for vulnerable rural regions. Finally, the study aims to refine FSM processes and provide actionable insights for addressing urban flooding risks, supporting sustainable development goals (SDGs), and enhancing resilience in flood-prone regions.

2. Materials and Methods

2.1. Study Area

The study area is located in the eastern part of Golestan Province, Iran, bounded by longitudes 55°20'E to 56°05'E and latitudes 37°15'N to 37°25'N. This region spans an area of approximately 632.50 km² and exhibits significant topographical variation, with elevations ranging from 154 m to 2,345 m above sea level. The higher altitudes are characterized by densely forested mountainous landscapes, contributing to the area's ecological richness and biodiversity. In contrast, the lower, flatter regions are predominantly occupied by rural settlements and agricultural lands, reflecting the human adaptation to the terrain. The region experiences a semi-humid climate, with an average annual rainfall of approximately 600 mm, which significantly shapes the area's hydrological and geomorphological processes. Due to its varied topography, climatic conditions, and land use patterns, the region is prone to flooding, especially during heavy rainfall events. Figure 1.

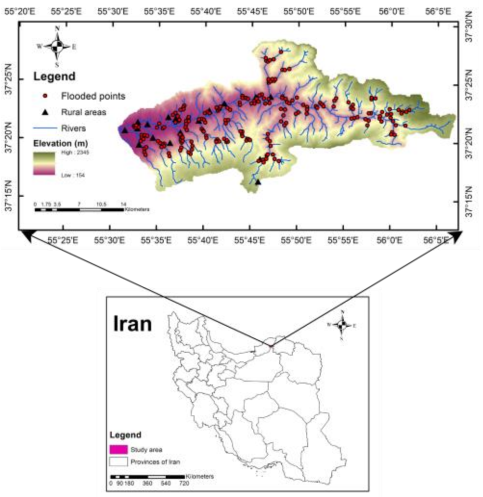

Figure 1. Study area

2.2. Data Collection and Preparation

This study obtained data from historical flood records and field observations collected via global positioning system (GPS) during the catastrophic flood event in Golestan Province in March 2019. A total of 403 data points were gathered and classified into two categories: flooded points (204 points), which represent areas directly affected by the flood, and non-flooded points (199 points), indicating areas that remained unaffected (Figure 1). The dataset was divided into training and testing subsets to develop and validate the flood susceptibility model. Specifically, 80% of the data was used to train the model, enabling it to learn patterns and relationships influencing flood susceptibility, while the remaining 20% was reserved for testing to evaluate the model's accuracy and predictive performance.

Moreover, floods are intricate and unpredictable phenomena, making selecting conditioning factors a complex and region-specific challenge. The absence of a universal framework for choosing suitable factors is attributed to the diverse environmental, climatic, hydrological, and anthropogenic characteristics that vary across regions (Nachappa et al., 2020). Identifying appropriate factors requires careful consideration of these unique attributes and the specific dynamics influencing flood events. In this study, 13 key factors influencing flooding were selected based on insights from previous research, data availability, and the region's topographic and environmental characteristics. These factors include elevation, aspect, slope, distance to rivers, river density, land use, rainfall, topographic wetness index (TWI), soil texture, normalized difference vegetation index (NDVI), normalized difference moisture index (NDMI), stream power index (SPI), and lithology Figure 2.

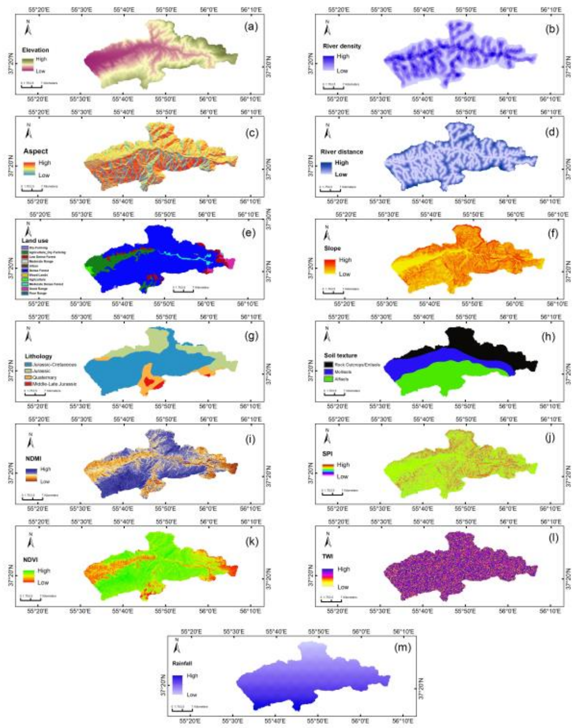

Figure 2. Flood conditioning factors: (a) Elevation, (b) River density, (c) Aspect, (d) River distance, (e) Land use, (f) Slope, (g) Lithology, (h) Soil texture, (i) NDMI, (j) SPI, (k) NDVI, (i) TWI, (m) Rainfall.

2.3. Multicollinearity Analysis

To assess multicollinearity, we examined the correlation matrix of the predictor variables to identify pairs with high correlation coefficients. This preliminary analysis helped pinpoint variables that might contribute to multicollinearity issues. To further quantify the degree of multicollinearity, we calculated the Variance Inflation Factor (VIF) for each predictor variable. VIF values provided a measure of how much the estimated coefficients were influenced by multicollinearity. Variables with excessively high VIF values were flagged as potential candidates for removal to enhance the model's accuracy and reliability.

2.4. Machine Learning Algorithms and Model Execution

Selecting appropriate ML algorithms is essential for accurate flood susceptibility mapping, particularly in regions with small datasets and complex, localized geographic features such as mountainous terrains. The algorithms chosen for this study—Random Forest (RF), XGBoost, LightGBM, CatBoost, and Support Vector Machine (SVM)—were selected for their proven ability to handle such challenges effectively. Random Forest is a robust ensemble method that reduces overfitting and improves prediction accuracy by averaging results from multiple decision trees. Its inherent feature selection and ability to model non-linear relationships make it ideal for analyzing diverse flood susceptibility factors, such as slope, rainfall, and soil type, in data-limited regions. Gradient-boosting algorithms, including XGBoost, LightGBM, and CatBoost, offer high accuracy and efficiency in uncovering complex patterns. XGBoost excels in

https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-1347-2025 |

handling missing data and identifying critical predictors, while LightGBM's histogram-based approach enhances generalization in spatially variable areas. CatBoost stands out for managing categorical variables like land use and vegetation, providing strong performance and reducing overfitting in small datasets. Support Vector Machine is particularly effective for small datasets due to its ability to model complex non-linear relationships and its resistance to overfitting. Its kernel functions allow it to separate classes effectively, making it well-suited for flood susceptibility mapping in regions with high-dimensional and diverse predictors.

This study utilized a 30-meter resolution Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) to derive topographical factors. Sentinel-2 imagery was processed in Google Earth Engine (GEE) to compute the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Moisture Index (NDMI) (Parto Dezfooli et al., 2025). Precipitation data were sourced from the CHIRPS (Climate Hazards Group InfraRed Precipitation with Station data) dataset, accessed through Google Earth Engine (GEE). The rainfall data were then interpolated using the Inverse Distance Weighting (IDW) method in ArcGIS to generate continuous precipitation surfaces for the study area. Additional variables, including slope, aspect, Stream Power Index (SPI), Topographic Wetness Index (TWI), river density, and river distance, were extracted from the DEM using ArcGIS. Data for soil texture, lithology and land use were sourced from national datasets in Iran. All conditioning factors were converted to raster format at a 30 × 30-meter resolution to ensure compatibility with ML models. Raster values corresponding to each flood-conditioning factor were extracted using the "Extract Values to Points" tool in the Spatial Analyst Toolbox of ArcGIS, and the resulting dataset was exported as a CSV file. Multicollinearity among predictor variables was assessed by examining the correlation matrix to identify pairs with high correlation coefficients. The Variance Inflation Factor (VIF) was calculated for each predictor variable, and variables with a VIF exceeding five were excluded from further analysis. Then, in Python, the StandardScaler from the sklearn.preprocessing library was applied to normalize the dataset. This scaling method transformed the conditioning factors to a standard normal distribution (mean = 0, standard deviation = 1), optimizing the dataset for use with ML algorithms. Next, the flood prediction dataset was constructed by integrating historical flood occurrence points with the conditioning factors. This dataset comprised 14 columns: 13 representing the conditioning factors and the final column indicating flood inventory (where '1' denoted flooded areas and '0' denoted non-flooded areas). The dataset was split into training and validation subsets in an 80:20 ratio, following best practices to ensure model reliability and robustness. Finally, model performance was evaluated using the Area Under the Receiver Operating Characteristic Curve (ROC-AUC). Feature importance was analyzed for each model. All model training and validation were performed in Python, while the preparation and analysis of the flood-prediction dataset were conducted and predicted flood susceptibility maps were classified into three categories of susceptibility: low, moderate, and high within the GIS environment. The flowchart of the methodology framework is depicted in Figure 3.

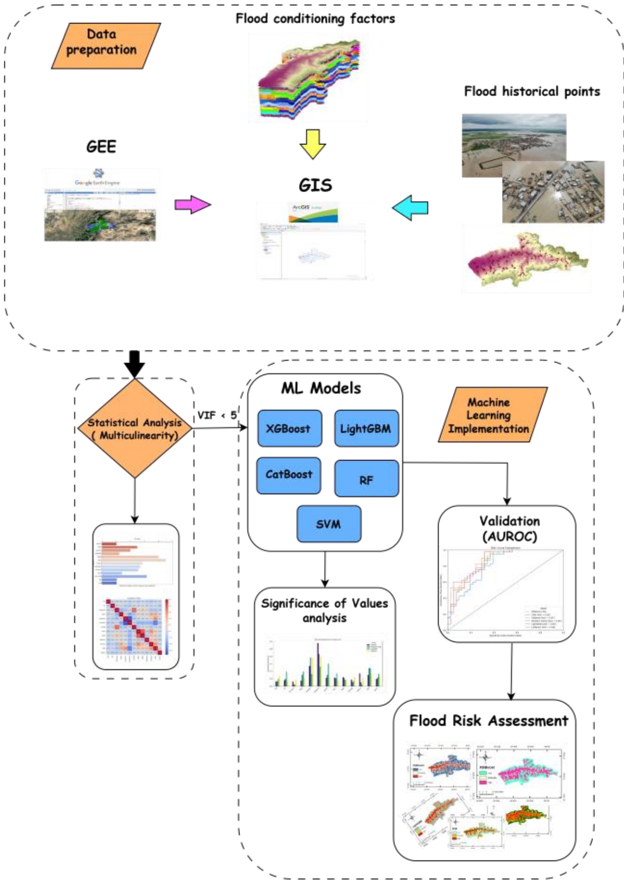

Figure 3. Workflow methodology of flood hazard mapping.

3. Results and Discussion

3.1. Multicollinearity analysis

To identify suitable independent variables for mapping flood-prone areas, a correlation matrix was examined to detect pairs with high correlation coefficients Figure 4, and a multicollinearity analysis was performed using the Variance Inflation Factor (VIF) method (Talukdar et al., 2022b). As shown in Figure X, the results indicated that among the 13 independent variables analyzed, NDVI (VIF = 4.51) and NDMI (VIF = 4) exhibited the highest collinearity. However, all variables had VIF values below 5, suggesting no significant multicollinearity. Consequently, all 13 independent variables were considered appropriate for modeling and mapping flood-prone areas in the study region.

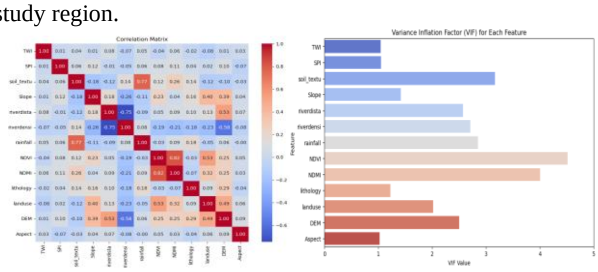

Figure 4. The results of multicollinearity analysis for the independent variables.

3.2. Flood Hazard Mapping

Flood hazard maps were developed using five advanced ML algorithms: RF, XGBoost, LightGBM, CatBoost, and SVM Figure 5. These models provided valuable insights into the spatial distribution of flood risks across the study area, revealing critical patterns of vulnerability and enabling targeted mitigation strategies. One of the most significant findings was that regions located near rivers were consistently identified as high-risk zones.

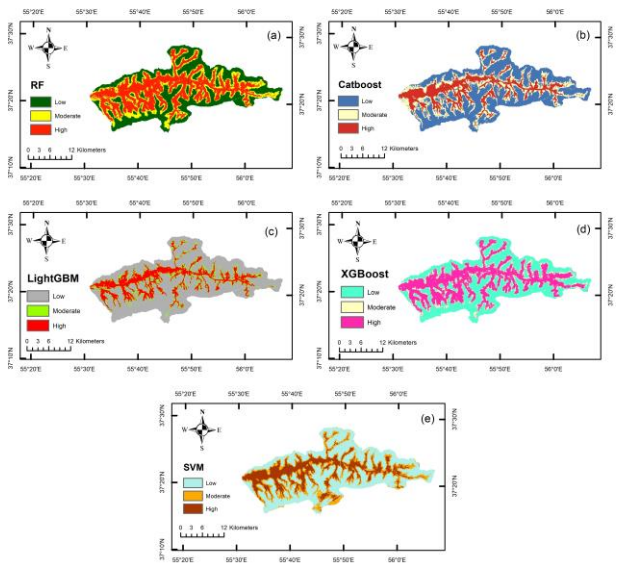

Figure 5. Flood risk assessment using (a) RF, (b) CatBoost, (c) LightGBM, (d) XGBoost, (e) SVM algorithms.

This elevated risk is primarily due to the region's humid climate, characterized by consistent rainfall throughout the year, which saturates the soil and increases runoff potential. Additionally, the area's mountainous and densely forested terrain, marked by steep slopes, exacerbates surface runoff during heavy rainfall events. These conditions amplify the likelihood of flash floods that funnel directly into rivers, increasing flood hazards along riverbanks and downstream areas. Furthermore, rural areas in the west part of the region were classified as low-lying zones highly susceptible to water accumulation. These areas are geographically positioned to collect runoff from higher elevations, effectively acting as natural basins. Their proximity to rivers further intensifies their flood vulnerability, as water from upstream and surrounding terrain converges within these zones. This combination of geographic and hydrological factors poses substantial risks to local communities, endangering infrastructure, disrupting livelihoods, and causing significant economic losses. These findings highlight the importance of prioritizing flood management efforts in river-adjacent and low-lying rural regions. Implementing measures such as improved drainage systems, reforestation with flood-resistant species, and community-based flood preparedness initiatives can help mitigate the adverse impacts of flooding in these high-risk areas.

3.3. Significance of Values Analysis

Feature importance was evaluated for the tree-based ML algorithms RF, XGBoost, LightGBM, and CatBoost, which inherently provide mechanisms to assess the contribution of each feature through their tree structures. In contrast, feature importance analysis is not applicable to SVM with an RBF kernel due to its non-linear nature and lack of direct metrics for feature contribution. This study focused exclusively on tree-based models to leverage their ability to quantify feature importance effectively. The importance scores of input variables were evaluated across the four models to assess their relative contributions to predicting the target variable. The results consistently showed that hydrologic features, particularly river distance and river density, were the most significant predictors in all models, underscoring their pivotal role in determining the target variable (Figure 6). Among topographic features, elevation and slope emerged as key predictors, achieving the second-highest importance scores in the Random Forest (RF), LightGBM, and CatBoost models. Land use, NDVI, and NDMI also served as secondary contributors, with these variables primarily influencing the XGBoost model, following hydrologic features. These findings underscore the necessity of carefully selecting input features and models to achieve optimal predictive performance. The results also provide valuable insights into the importance of topographic and hydrologic features in predicting the target variable, regardless of the specific ML model employed.

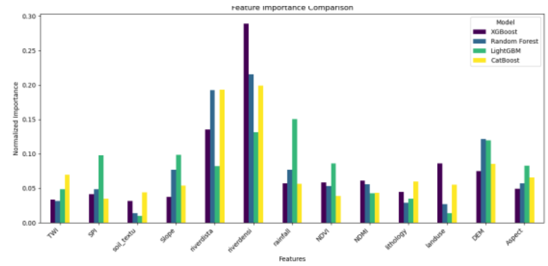

Figure 6. Feature importance RF, XGBoost, LightGBM, and CatBoost algorithms.

3.4. Evaluating Different ML Algorithms Performance

Evaluating flood hazard models is essential for accurate flood risk prediction and management, providing insights for disaster preparedness, urban planning, and resource allocation. Proper evaluation ensures reliable predictions and informed decision-making for stakeholders. The performance of the machine learning algorithms was assessed using the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC), a widely used metric for evaluating the ability of models to distinguish between flood-prone and non-flood-prone areas. The results demonstrate that all models performed well, achieving high AUC values indicating their effectiveness in predicting flood hazards. Our AUC results show strong performance across all models, with XGBoost and CatBoost leading the way. XGBoost achieved the highest AUC of 0.87, indicating excellent discriminatory power in distinguishing flood-prone and non-flood-prone areas. This suggests that XGBoost effectively captures complex relationships and patterns in our data. CatBoost, with an AUC of 0.86, performed similarly, showing that it also handles our data well, particularly with categorical variables, which it processes automatically. The slight difference between the two models may stem from differences in how they handle features and interactions. Random Forest and LightGBM both delivered AUCs of 0.85, which is still very good but slightly behind the boosted models. While Random Forest is robust and reliable, it might not capture intricate patterns as effectively as XGBoost or CatBoost, which use gradient boosting techniques to refine their predictions. Similarly, LightGBM's performance, although competitive, did not surpass Random Forest or the boosting models, likely due to its sensitivity to feature interactions or other model-specific tuning factors. SVM showed the lowest performance with an AUC of 0.82. While this is still decent, SVM may struggle to capture the non-linear relationships in the flood susceptibility data. Given the complexity of the problem, tree-based and boosting models tend to outperform SVM in such cases. To make it short, our study findings indicate that XGBoost and CatBoost stand out as the top performers, offering the best AUC and thus making them the most promising models for our flood susceptibility mapping. Random Forest and LightGBM are still strong contenders but may benefit from further tuning to improve performance. While useful in other contexts, SVM appears less effective for this particular task Figure 7.

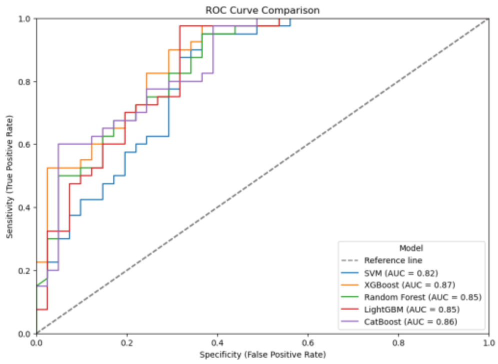

Figure 7. ROC-AUC Curve of all Algorithms.

3.5. Discussion

Flood susceptibility mapping is essential for understanding the spatial distribution of flood risks and informing decision-making processes for effective flood management. The goal of this study was to evaluate and compare the performance of five widely used machine learning algorithms—Random Forest (RF), XGBoost, LightGBM, CatBoost, and Support Vector Machine (SVM)—in generating flood susceptibility maps for a mountainous region in northern Iran. The results from this study provide important insights into the strengths and limitations of these algorithms for flood risk assessment, particularly in areas characterized by complex topography, hydrology, and land use dynamics. One of the key contributions of ML in flood susceptibility mapping is its ability to provide insights into the relative importance of different predictor variables. This study's feature importance analysis revealed that hydrological features, such as river distance and river density, were consistently identified as the most significant predictors of flood susceptibility across all models. This underscores the critical role that river networks play in flood events, particularly in regions with steep, mountainous terrain, where runoff can quickly accumulate and exacerbate flood hazards. The identification of river-related features as top predictors is consistent with the findings of other studies mentioned earlier in the introduction section, where proximity to rivers and river density have been shown to strongly influence flood risk.

Topographic factors such as elevation and slope also emerged as important predictors in the models. Elevated areas tend to receive more rainfall, and steeper slopes increase runoff, both of which contribute to higher flood susceptibility. These findings are in line with the hydrological principles of runoff generation in mountainous regions, where the terrain’s physical features directly influence water flow and accumulation patterns. The prominence of these topographic variables aligns with the general understanding that floods in mountainous regions are often influenced by steep slopes that accelerate surface runoff. Additionally, land use, vegetation indices (NDVI, NDMI), and soil characteristics were found to play secondary roles in the models, with their relative importance varying across different algorithms. These factors influence the soil's water retention capacity, the ground's permeability, and the presence of vegetation that could either mitigate or exacerbate flooding. The inclusion of vegetation indices, such as NDVI and NDMI, highlights the relevance of land cover in flood susceptibility models (Kumar et al., 2022; Chakraborty et al., 2023). Vegetation can significantly affect soil moisture retention and surface runoff, influencing flood risks in a given area. The ability of XGBoost and CatBoost to automatically handle categorical variables like land use and vegetation further enhanced their performance.

The findings of this study have important implications for flood risk management in northern Iran and similar mountainous regions. The consistently high performance of XGBoost and CatBoost suggests that these models can be reliably used for generating accurate flood susceptibility maps, which are essential for effective flood disaster planning and mitigation. In particular, the identification of river-adjacent areas and low-lying regions as high-risk zones emphasizes the need for targeted flood risk management strategies in these areas. Given the significant role of topographic and hydrological factors in flood susceptibility, the study advocates for integrating these factors into local flood management practices.

4. Conclusion

This study has demonstrated the effectiveness of machine learning algorithms, particularly XGBoost and CatBoost, in generating accurate flood susceptibility maps for a mountainous region in northern Iran. The superior performance of these gradient boosting models highlights their ability to capture complex spatial relationships and interactions between topographic, hydrological, and environmental factors that are critical for flood risk assessment. Additionally, RF and LightGBM also showed strong performance, while SVM proved to be less effective in handling the complexities of the flood susceptibility data. The feature importance analysis revealed that hydrological factors, such as river distance and density, as well as topographic variables like elevation and slope, are the primary determinants of flood susceptibility. These findings are consistent with previous studies and underscore the significant role of river networks and terrain in flood occurrence, particularly in mountainous regions. While secondary, land use, vegetation indices, and soil characteristics also contributed to the flood risk prediction, highlighting the importance of including environmental factors in flood susceptibility models. The results from this study have crucial implications for flood risk management in northern Iran and similar mountainous regions. The successful application of machine learning models, particularly XGBoost and CatBoost, provides valuable tools for flood risk mapping, aiding in the development of effective flood management strategies. These strategies should include improved drainage systems, river channelization, and floodplain restoration, particularly in areas with high river density and proximity to rivers. Additionally, rural areas, which are particularly vulnerable to flood impacts, should be prioritized for flood preparedness measures, including community education and early warning systems. The incorporation of socio-economic factors, such as population density and income levels, into future flood susceptibility models would further enhance the understanding of rural communities' vulnerability and resilience to floods.

Acknowledgements

This study forms part of the Ph.D. thesis of Ms. Mahdieh Shirmohammadi at Nanjing Normal University (NNU) and will be presented at the ISPRS Geospatial Week (GSW) 2025. The authors extend their gratitude to Nanjing Normal University and Shaoxing University for their invaluable support. Additionally, the authors express their appreciation to the Chinese Scholarship Council (CSC) for awarding a scholarship to the first author, enabling her Ph.D. studies at NNU.

References

Amiri, A., Soltani, K., Ebtehaj, I., & Bonakdari, H., 2024: A novel machine learning tool for current and future flood susceptibility mapping by integrating remote sensing and geographic information systems. *Journal of Hydrology*, 632, 130936.

Baida, M., Boushaba, F., Chourak, M., Hosni, M., Sabar, H., & Zahaf, T., 2024: Flood risk decomposed: Optimized machine learning hazard mapping and multi-criteria vulnerability analysis in the city of Zaio, Morocco. *Journal of African Earth Sciences*, 220, 105431.

Chakrabortty, R., Chandra Pal, S., Ruidas, D., Roy, P., Saha, A., & Chowdhuri, I., 2023: Living with floods using state-of-the-art and geospatial techniques: Flood mitigation alternatives, management measures, and policy recommendations. *Water*, 15(3), 558.

Demissie, Z., Rimal, P., Seyoum, W. M., & Dutta, A., 2024: Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment. *Applied Computing and Geosciences*, 23, 100183.

Eini, M., Kaboli, H., Rashidian, M., & Hedayat, H., 2020: Hazard and vulnerability in urban flood risk mapping: Machine learning techniques and considering the role of urban districts. *International Journal of Disaster Risk Reduction*, 50, 101687.

Habibi A., Delavar M.R., Nazari B., Pirasteh S., Sadeghian M.S., Li J., 2023: Novel intelligence models for flood hazard: Development of novel hybridized ensemble machine learning models with SA and IG feature selection algorithms. *International Journal of Applied Earth Observation and Geoinformation*, 122:103443, https://doi.org/10.1016/j.jag.2023.103443.

Hitouri, S., Mohajane, M., Lahsaini, M., Ajim Ali, S., Setargie, T., Tripathi, G., D'Antonio, P., Singh, S., & Varasano, A., 2024: Flood susceptibility mapping using SAR data and machine learning algorithms in a small watershed in Northwestern Morocco. *Remote Sensing*, 16, 858.

Kazemi, M., Mohammadi, F., Hassanzadeh Nafooti, M., Behvar, K., & Kariminejad, N., 2024: Flood susceptibility mapping using machine learning and remote sensing data in the Southern Karun Basin, Iran. *Applied Geomatics*, 16, 731–750.

Kumar, V., Sharma, K., Caloiero, T., Mehta, J., & Singh, K., 2023: Comprehensive overview of flood modeling approaches: A review of recent advances. *Hydrology*, 10(7).

Li W., Zhu J., Pirasteh S., Zhu Q., Fu L., Wu J., 2022: Investigations of disaster information representation from geospatial perspective: progress, challenges, and recommendations. *Transactions in GIS*. 26 (3): 1376-1398. Trans. GIS. DOI: 10.1111/tgis.12922, 00:1–23.

Nachappa, T., Ghorbanzadeh, O., Gholamnia, K., & Blaschke, T., 2020: Multi-hazard exposure mapping using machine learning for the state of Salzburg, Austria. *Remote Sensing*, 12(17), 2757.

Nguyen, H., Nguyen, Q., Dang, D., Van, C., Truong, Q., Pham, S., Bui, Q., & Petrisor, A., 2024: A novel flood risk management approach based on future climate and land use change scenarios. *Science of the Total Environment*, 921, 171204.

Parto Dezfooli F., Valadan Zoej M.J., Mansourian A., Alatab S., Youssefi F., Pirasteh S., 2025: GEE-Based Environmental Monitoring and Phenology Correlation Investigation Using Support Vector Regression. *Remote Sensing Applications: Society and Environment*, 37:101445, https://doi.org/10.1016/j.rsase.2024.101445.

Singha, C., Swain, K., Meliho, M., Abdo, H., Almohamad, H., & Al-Mutiry, M., 2022: Spatial analysis of flood hazard zoning map using novel hybrid machine learning technique in Assam, India. *Remote Sensing*, 14, 6229.