EEET ECOLOGICAL ENGINEERING & ENVIRONMENTAL TECHNOLOGY

Ecological Engineering & Environmental Technology, 2025, 26(1), 201–215
https://doi.org/10.12912/27197050/195845
ISSN 2719–7050, License CC-BY 4.0

Received: 2024.11.01
Accepted: 2024.11.28
Published: 2024.12.01

Developing an explainable and interpretable machine learning model for flood susceptibility mapping

Loubna Khaldi[1*], Ali EL Bilali[2], Alae Elabed[1], Nir Krakauer[3,4], Abdessalam El Khanchoufi[1]

[1] Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[2] River Basin Agency of Bouregreg and Chaouia, Benslimane, Morocco
[3] Department of Civil Engineering, The City College of New York, New York, NY 10031, USA
[4] Earth and Environmental Sciences, City University of New York Graduate Center, New York, NY 10016, USA
* Corresponding author's e-mail: loubna.khaldi@usmba.ac.ma

ABSTRACT

This study evaluates flood susceptibility in the Fez-Meknes region of Morocco by comparing the performance of five machine learning (ML) models using 14 environmental variables. The selected models, including random forest (RF), support vector machine (SVM), K-Nearest neighbor (KNN), recursive partitioning and regression trees (RPART), and logistic regression (LR), were assessed for prediction accuracy and enhanced with partial dependence plots (PDP) and local interpretable model-agnostic explanations (LIME) to increase interpretability. Results indicate that the RF model outperforms other models, achieving a high prediction accuracy with an AUC of 96%, low mean absolute error (MAE) of 0.26, and root mean squared error (RMSE) of 0.31, along with strong Nash Sutcliffe efficiency (NSE) and correlation coefficient (R^2). Through PDP and LIME, the primary factors influencing flood susceptibility were identified as proximity to rivers, drainage density, slope, normalized difference vegetation index (NDVI), terrain roughness index (TRI), and land use and land cover (LULC). These findings highlight the potential of interpretable ML models to enhance flood risk assessment, providing valuable insights for urban planning and flood mitigation strategies in vulnerable regions.

Keywords: flood susceptibility, modeling and mapping, random forest model, Fez-Meknes region.

INTRODUCTION

Flooding represents a serious hydrological hazard worldwide, posing significant risks to infrastructure, ecosystems, and human lives (Dolchinkov, 2024; Sayers et al., 2015). As urbanization expands and climate change exacerbates extreme weather events, understanding flood-prone areas and improving flood susceptibility modeling (FSM) are essential for effective disaster preparedness and mitigation. Traditionally, flood risk management has incorporated a combination of land-use planning, structural flood control measures, and early warning systems. However, recent advances in ML have introduced new opportunities to model flood susceptibility with greater precision, enabling the creation of detailed flood risk maps that aid in identifying and preemptively addressing high-risk zones (Pham et al., 2021; Saha et al., 2021).

While deterministic and statistical approaches, such as the analytic hierarchy process (AHP) integrated with geographic information systems (GIS), have been widely used to assess flood susceptibility (Khaldi et al., 2023), these methods are often constrained by subjectivity and inherent uncertainties in factor weight assignments. Such limitations can affect model robustness and the reliability of susceptibility predictions. For instance, outcomes from AHP-based flood models may vary due to subjective assessments, as noted (Mudashiru et al., 2022), and prior research has advocated for the incorporation of sensitivity and uncertainty analyses to minimize biases in AHP results. Moreover,

recent studies have suggested augmenting AHP with statistical techniques like correlation and regression analyses to validate results and enhance model robustness (Mousavi et al., 2022)

In response to these challenges, ML approaches such as RF, artificial neural networks (ANN), decision trees, and support vector machines (SVM) have gained traction for FSM due to their flexibility in handling complex, nonlinear relationships among variables (Chapi et al., 2017a; Chen et al., 2019), ANN (Bui et al., 2019; Kia et al., 2012).These models have demonstrated strong predictive performance in FSM, yet they exhibit varying degrees of reliability depending on regional and data-specific conditions (Davoudi Moghaddam et al., 2019; Khosravi et al., 2019). Consequently, this underscores the importance of comparing these models' performance and adaptability within different environmental contexts.

A critical limitation of many high-performing ML models, however, is their "black-box" nature, which often obscures the specific influence of input variables on flood predictions. The interpretability of ML models is crucial for FSM, as stakeholders need to understand not only where flood risks are high but also which factors contribute most significantly to these risks. Local interpretable model-agnostic explanations (LIME) (Mishra et al., 2017) and partial dependence plots (PDP) (Friedman, 2001; Holzinger et al., 2022)are two interpretive techniques that have emerged as powerful tools for elucidating ML models. LIME enables localized explanations by examining feature importance for individual predictions, while PDP provides a global view of variable influence by plotting marginal effects across the feature space. Together, these techniques offer a transparent framework for interpreting complex models, thereby enhancing their utility in practical flood risk management applications.

The Fez-Meknes region of Morocco, predominantly characterized by rain-fed agriculture, has witnessed frequent and disruptive flood events. Despite these recurring hazards, FSM studies using advanced ML models remain scarce for this area. The lack of region-specific FSM models presents a significant gap in the literature, particularly as the region faces heightened flood risks associated with climate variability. This study addresses this gap by employing five ML algorithms – SVM, RF, KNN, LR, and Recursive partitioning and regression trees (Rpart)—to model flood susceptibility in Fez-Meknes. Notably, we incorporate LIME and PDP for a comprehensive interpretation of model outputs, thus bridging the gap between predictive accuracy and interpretability.

In this study, we aim to produce a transparent and scientifically robust flood susceptibility map for the Fez-Meknes region. By using LIME and PDP, we provide detailed insights into the role of environmental, hydrological, and topographical variables in flood susceptibility. This approach not only enhances model interpretability but also provides critical information for urban planners and decision-makers to better allocate resources, plan infrastructure, and mitigate flood risks. To the best of our knowledge, this study represents the first application of explainable machine learning models, combined with interpretative techniques, for flood susceptibility mapping in this region. This innovative approach significantly advances the field of flood susceptibility modeling (FSM) and regional flood management.

MATERIALS AND METHODS

Study area

The region of Fez-Meknes is located in the north-central part of Morocco, covering an area of 47.705 square kilometers. It includes the prefectures of Fez and Meknes, as the well as seven provinces, and is located between 33° 02′ 00″ and 35° 02′ 00″ north latitude and 4 ° 00′00′ and 6 ° 00′00′ longitude (Figure 1).

The climate exhibits a Mediterranean character in the northern region, characterized by hot summers and cool winters, while it tends to be cooler in the southern mountainous areas. Rainfall varies across the region, with wetter areas in the south receiving an average of 800 mm and drier areas in the north and northeast receiving an average of 479 mm between 1988 and 2017. The region has several natural rivers and lakes, including the Fez River, Guigou River, Boufekrane River, Tizguit River, Agay River, and DayetIfrah Lake, which have experienced severe flooding in the past. In 2020, it is estimated that the useful agricultural area is about 13356.39 km², or about 15% of the total useful agricultural area of Morocco. The region, including the notable cedar forests of the Middle Atlas and other areas like

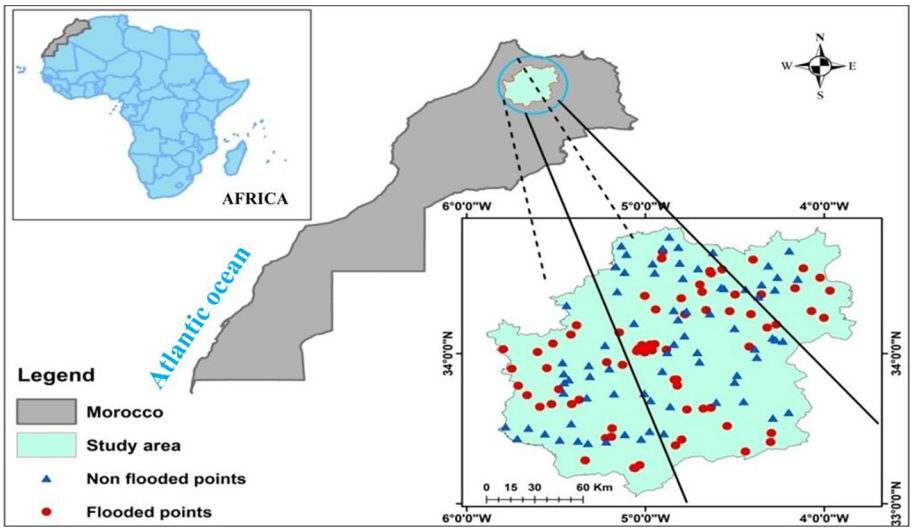

Figure 1. Study area location

Ifrane-Azrou, Ain Alleuh, Taza, and the Rif, holds significant national importance due to its forests' role in landscape aesthetics and economic potential. With an estimated population of around 4 million and an urbanization rate expected to reach 67% by 2036, the region faces growing challenges.

Methods

The evaluation of flood susceptibility in this study employed various ML algorithms, such as RF, SVM, KNN, Rpart, and LR. To elucidate the methodology used in this research (Figure 2), serves as an illustrative framework.

Preparation of data set

Table 1 summarizes the datasets employed in this study. The process commenced with the use of a 1:10,000 scale topographic map to develop a high-resolution digital elevation model (DEM) with a 30-meter spatial resolution. This DEM served as the foundation for calculating several topographic attributes, including slope, flow accumulation, curvature, elevation, the topographic wetness index (TWI), terrain roughness index (TRI), and stream power index (SPI), which were derived using ArcGIS software. LULC and NDVI data were acquired

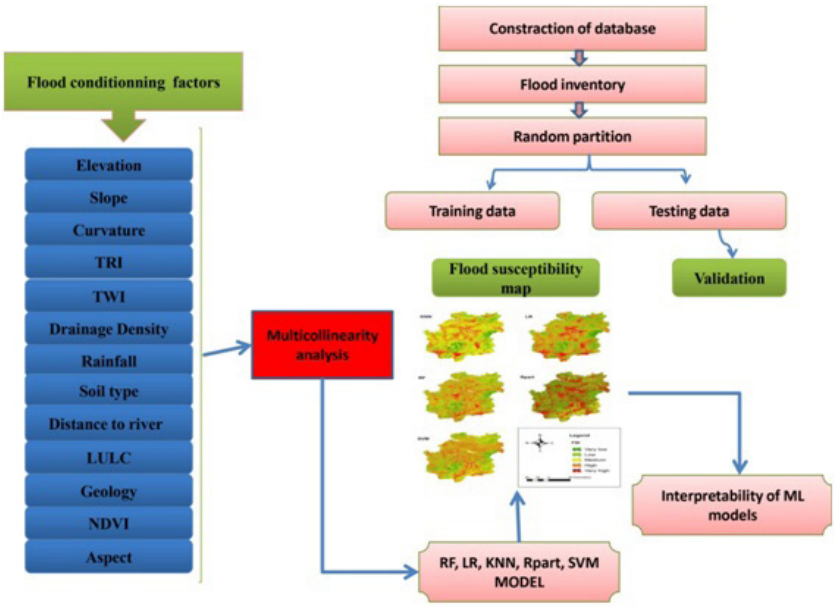

Figure 2. Methodological flow chart

Table 1. Sourcesfor databases

Category	Data source	Resolution	Recording time	Data output type
DEM (elevation)	USGS website	30 × 30 m	2019	Raster
Slope	Extracted from DEM	30 × 30 m	2019	Raster
Curvature	Extracted from DEM	30 × 30 m	2019	Raster
Ruggedness of the terrain index (TRI)	Extracted from DEM	30 × 30 m		Raster
Topographic weightiness index (TWI)	Extracted from DEM	30 × 30 m	2019	Raster
Drainage density	Extracted from DEM	30 × 30 m	2019	Raster
Rainfall	Krigin method	30 × 30 m	2019	Raster
Soil	From the geology map	30 × 30 m	1985	Raster
Distance from the river	Extracted from DEM	30 × 30 m	2019	Raster
Land use and land cover (LULC)	Extracted from Satellite image	30 × 30 m	2019	Raster
Aspect	Extracted from DEM	30 × 30 m	2019	Raster
Normalized difference vegetation index (NDVI)	Landsat		2019	Raster
Geology	Geology map of Morocco	1.100,000	1985	Raster
Stream power index (SPI)	Extracted from DEM	30 × 30 m	2019	Raster
Flood extent	Field survey and historical flood data	Randomly	1980-2019	vector

from Landsat OLI imagery from 2020, available through the EarthExplorer platform (https://earthexplorer.usgs.gov). Additionally, to construct the precipitation map, annual precipitation data covering the period from 1965 to 2020 were collected from 15 monitoring stations managed by the Watershed Agency and the Directorate of Water and Forestry.

Flood inventory

For this study, an inventory map spanning the period from 2000 to 2019 was developed by compiling GPS data from field surveys, annual flood reports, and satellite imagery from Google Earth. This data enabled the identification and precise mapping of locations impacted by flooding.Some of the data were obtained from the Sebou River Basin Agency (ABHS) and the Fez Water and Forests and Meteorology administration. For binary analysis, an equivalent number of non-flooded sites were chosen through reference to historical records and on-site assessments. These 150 flood-prone and 150 non-flood locations were randomly partitioned into two groups: (70%) for training and (30%) for validation in order to construct and assess the flood susceptibility map. Both the training and validation datasets were transformed into raster format, with flood-prone and non-flood locations represented as 1 and 0, respectively.

Multi-collinearity analysis

The selected conditioning factors (FCFs) underwent a multicollinearity assessment by analyzing the tolerance (TOL) and variance inflation factor (VIF) values. This assessment helps detect high correlations among FCFs, which could compromise the model's accuracy. A TOL value below 0.1 combined with a VIF value exceeding 5 indicates potential multicollinearity issues, as noted by (Khosravi et al., 2019).These metrics were calculated using Equations 1 and 2.

$$TOL = 1 - R_j^2 \quad (1)$$

$$VIF = \frac{1}{TOL} \quad (2)$$

where: R^2_j denotes the regression value associated with the j variable when considered alongside the other explanatory variables.

Machine learning models

Support vector machine

The support vector machine (SVM) is a widely used nonparametric training technique in remote sensing (Mountrakis et al., 2011). The algorithm operates by discerning the optimal hyperplane that maximizes the separation between two data classes, which are separated by the hyperplane using a set of support vectors (Cracknell and Reading, 2014). SVM aims to classify items into

specific classes based on their location relative to the optimal hyperplane (Maxwell et al., 2018). There are four main types of kernel function used in SVM algorithms, including linear, polynomial, radial basis function (RBF) and sigmoid kernels, with each having its own advantages and disadvantages (Kavzoglu and Colkesen, 2009).

Random forest (RF)

The RF is a popular ensemble learning algorithm in remote sensing research due to its high classification accuracy (Belgiu and Dra, 2016). In RF, multiple decision trees (ntrees) are created using a bagging procedure with randomly selected training data subsets. Each tree grows independently to its maximum size, dividing each node using the best input variable from a random subset of size mtry without any adjustments (Breiman, 2001).The predicted classification is then determined by the class receiving the most votes from each tree predictor. Therefore, ntree and mtry are important parameters that need to be defined. For this, tuning these parameters through cross-validation and experimentation is crucial to finding the right balance between model performance and computational efficiency.

K-Nearest neighbor (KNN)

The KNN is a nonparametric approach that assumes that items in the same category are clustered together. The distance between the target point and its k nearest neighbors is usually calculated using the Euclidean distance method. After finding the k neighbors of each considered point, it is assigned the class with the most neighbors. KNN is considered a "lazy learner" because it needs to calculate the distances between every point in its neighborhood. Compared to other ML models, it takes longer to train (Guo et al., 2003).

Logistic regression (LR)

LR stands as a widely adopted methodology for examining various variables that have the potential to impact the probability of flooding. This method was originally introduced by (Mc Fadden, 1974) when he devised a formula to construct conditioning factors for evaluating the likelihood of a disaster occurring within a specific geographical area. A notable advantage of this approach is its lack of reliance on data adhering to a normal distribution. Additionally, these conditioning factors can take on various forms, including continuous, discrete, or a combination of both, offering a high degree of flexibility in the analysis (Lee and Sambath, 2006). This technique empowers researchers to scrutinize the relationship between binary dependent variables, utilizing scalar and nominal values as conditioning factors.

Recursive partitioning and regression trees (Rpart)

The Rpart model is a popular decision tree technique in image-based classification due to its non-parametric nature and ease of interpretation. It is capable of handling both regression and classification trees through recursive partitioning. The model can perform a regression tree if the response variable is numeric or a classification tree if it is a categorical value or factor.

Evaluation of model performance

In this study, we evaluated the effectiveness of the model using several statistical indicators, including specificity, sensitivity, Kappa index, and accuracy, following the methodology proposed by (Pham et al., 2021). These metrics were calculated using Equations 3, 4, 5, and 6.

$$\text{Sensitivity} = \frac{TP}{TN + FN} \tag{3}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4}$$

$$\text{Accuracy} \frac{TP + TN}{TP+TN+FP+FN} \tag{5}$$

$$\text{Kappa} = \frac{L_0 - L_e}{L_t - L_e} \tag{6}$$

True positives (TP), signifies the count of pixels accurately identified as positive predictions (indicating flooding), while true negatives (TN) denotes the count of pixels correctly identified as negative predictions (indicating non-flooding). Meanwhile, false positives (FP) and false negatives (FN) refer to the number of pixels incorrectly classified as positive (indicating flooding) or negative (indicating non-flooding), respectively. To gauge the agreement between observed and predicted flooding locations, the Kappa coefficient or Kappa index is employed, where L_0 represents observed agreement, L_e represents projected agreement, and L_t represents total agreement.

Furthermore, we employed the receiver operating characteristic (ROC) curve to evaluate flood susceptibility models, a method commonly used in similar studies (Chen et al., 2020; Gudiyangada Nachappa et al., 2020; Tehrany et al., 2013). The AUC derived from the ROC curve, ranging from 0 to 1, offers a key performance metric, indicating the model's ability to distinguish flooded and non-flooded areas. The AUC was calculated using Equation 7.

$$AUC = \frac{(TP + TN)}{(P + N)} \quad (7)$$

where: P represents total number of flooding locations, N total number of non-flooding locations.

Additionally, mean absolute error (MAE), root mean squared error (RMSE), percent bias (PBIAS), Nash-Sutcliffe efficiency (NSE), correlation coefficient (r), and coefficient of determination (R^2) were utilized to provide insights into the accuracy, reliability, and overall performance of each machine learning method in predicting flood susceptibility.

Interpretability of machine learning

Partial dependency plots

Partial dependency plots (PDPs) illustrate the incremental impact of a specific feature. The essence of PDPs lies in their ability to alter the feature under scrutiny within a ML model while keeping all other features constant. This allows for a comprehensive assessment of how the chosen feature influences the model's predictions. One notable advantage of PDPs compared to certain other interpretability methods is their capacity to reveal the functional relationship between one or two features and the predicted outcomes.

Local interpretable model-agnostic explanations

Local interpretable model-agnostic explanations (LIME) is a specific technique in the realm of ML interpretability. LIME is designed to provide explanations for individual predictions made by ML models. It focuses on explaining the predictions of a model for a particular data point rather than providing a global explanation for the entire model. The key idea behind LIME is to approximate the behavior of the complex ML model being explained by training a simpler, interpretable model on a locally generated dataset.

RESULTS

Multicollinearity test

Based on a previous study, 15 factors were selected to study flood susceptibility in the Fez Meknes region using AHP (Khaldi et al., 2023). Table 2 presents the results of the multicollinearity assessment among the 15 flood conditioning

Table 2. The evolution of the influencing factors using VIF and tolerance

Parameter	15 factors		14 factors	
Factor	Tolerance	VIF	Tolerance	VIF
NDVI	0.687	1.502	0.667	1.500
LULC	0.792	1.295	0.776	1.289
Curvature	0.756	1.338	0.749	1.336
Elevation	0.592	1.673	0.598	1.672
Aspect	0.919	1.095	0.913	1.095
Density	0.843	1.201	0.847	1.181
Distance	0.922	1.097	0.912	1.096
Geology	0.596	1.669	0.600	1.668
Sol	0.823	1.213	0.825	1.212
TRI	0.876	1.237	0.814	1.228
TWI	0.563	1.629	0.383	2.609
Rainfall	0.804	1.3	0.777	1.287
Slope	0.559	2.005	0.499	2.004
SPI	0.003	334.397	0.622	1.608
Accumulation	0.003	338.690		–

factors. The observed TOL and VIF values reveal significant intercorrelations between two factors, suggesting possible multicollinearity issues within the dataset. TheToL and VIF values showed that 13 of the flood factors ranged from 1.08 to 1.79 and 0.559 and 0.919 respectively. However, the calculation of VIF for SPI and accumulation resulted in values 334 and 338, respectively. Since the VIF values for SPI and accumulation exceeded 5 and the ToL values were less than 0.1, a variable should be removed from the final model. When the accumulation factor was excluded, the findings revealed that there is no collinearity among the 14 selected variables, as the VIF and TOL values remained within acceptable thresholds.

Model comparison

Table 3 shows the performance evaluation of five FSMs in both the validation and training datasets based on TP, TN, FP, FN, kappa and AUC. AUC values exceeding 80% are considered high-performing models, while an AUC of 80–90% is considered very good, and an AUC above 90% is considered excellent. Based on the results presented in Figure 3 and Table 3, the RF, SVM, and Rpart models showed high prediction accuracy (AUC) of 96%, 89%, and 87%, respectively, indicating that they are high-performance models.

The RF model displays the highest sensitivity, reaching 0.97 during training and 0.79 in testing, which highlights its effectiveness in accurately detecting positive cases. Conversely, the KNN model exhibits the lowest sensitivity, with values of 0.60 in training and 0.29 in testing. The specificity metric, which measures the ability to correctly identify negative cases, demonstrates that the SVM model has the highest values of 0.87 in training and 0.95 in testing. The KNN model shows the lowest specificity values in both training and testing. When

Table 3. Performance metrics of different classification models

Parameter	Training					Testing				
Metrics	KNN	RF	Rpart	LR	SVM	KNN	RF	Rpart	LR	SVM
TP	45	73	64	60	62	7	19	19	16	16
TN	48	60	56	40	65	20	20	17	14	20
FP	27	15	19	35	10	1	1	4	7	1
FN	30	2	11	15	13	17	5	5	8	8
Sensitivity	0.60	0.97	0.85	0.80	0.83	0.29	0.79	0.79	0.67	0.67
Specificity	0.64	0.80	0.75	0.53	0.87	0.95	0.95	0.81	0.67	0.95
PPV ()	62.50	82.95	77.11	63.16	86.11	87.50	95.00	82.61	69.57	94.12
NPV ()	61.54	96.77	83.58	72.73	83.33	54.05	80.00	77.27	63.64	71.43
Kappa	0.240	0.773	0.600	0.49	0.69	0.23	0.73	0.59	0.33	0.60
AUC	0.69	0.96	0.87	0.75	0.89	0.73	0.97	0.86	0.76	0.90

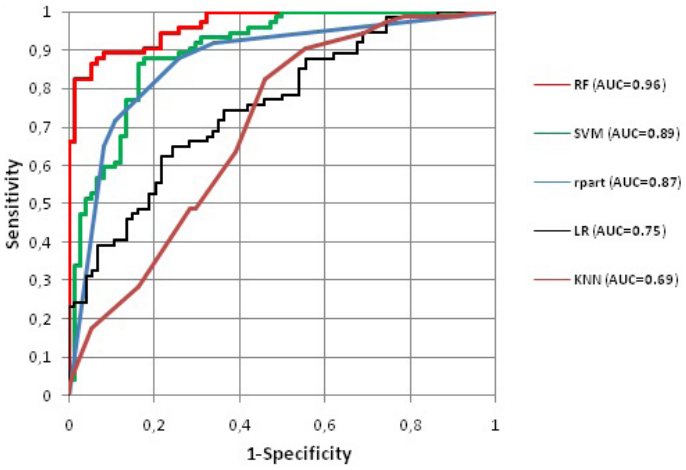

Figure 3. Area under curve of different classification models

considering the positive predictive value (PPV), the RF model performs well with 82.95% in training and 95.00% in testing. The negative predictive value (NPV) highlights the RF model's strength again, achieving high values of 96.77% in training and 80.00% in testing. The Kappa coefficient, which measures agreement between predicted and actual values, indicates that the RF model has the highest value of 0.773 in training and 0.735 in testing, suggesting substantial agreement. The RF model demonstrated robust performance, as confirmed by the area under the ROC curve (AUC) values of 0.96 in the training dataset and 0.976 in the testing dataset, signifying its exceptional discriminatory capability.

Analyzing the uncertainty of ML models based on the provided Table 4 involves understanding the variability and reliability of model predictions across different metrics.

The results show, RF exhibits relatively lower MAE (0.26) and RMSE (0.31) values compared to other models, suggesting higher precision and accuracy in predicting flood susceptibility levels. Conversely, LR shows higher error rates, implying greater uncertainty in its predictions. The PBIAS values reveal insights into potential biases in the models' predictions, with RF and recursive partitioning and regression trees (RPART) demonstrate negative biases, while KNN and SVM models exhibit positive biases, indicating potential overestimation issues.

Furthermore, the Nash-Sutcliffe efficiency, correlation coefficient, and coefficient of determination metrics provide insights into the models' ability to capture and reproduce observed flood susceptibility patterns. RF consistently demonstrates higher NSE (0.63), r (0.83), and R^2 (0.69) values compared to other models, indicating better agreement between predicted and observed values and lower uncertainty in its predictions.

Moreover, the outcomes of both the Friedman and Wilcoxon rank tests indicated the presence of statistically significant distinctions among all the models, as demonstrated in Table 5 and Table 6. The mean rank values spanned from 2.55 to 3.37. Additionally, the Friedman test results yielded a p-value below 0.05, signifying the rejection of the null hypothesis and the acceptance of the alternative hypothesis with a 95% confidence level.

Table 4. Model performances and uncertainties during the validation process

Parameter	LR	RPART	RF	KNN	SVM
MAE	0.39	0.2	0.26	0.45	0.29
RMSE	0.45	0.38	0.31	0.47	0.37
PBIAS	-3.1	-0.6	-2.1	15	16.3
NSE	0.2	0.42	0.63	0.11	0.44
r	0.45	0.68	0.83	0.38	0.69
R^2	0.2	0.46	0.69	0.15	0.47

Table 5. Models comparison utilizing the Wilcoxon signed-rank test (two-tailed)

Test statistics				
	LR - RF	SVM - RF	RPART - RF	KNN – RF
Z	-1.964	-2.862	-.803	-2.299
P-value	0.05	0.004	0.422	0.022
Significance	Yes	Yes	no	Yes

Table 6. The Friedman's test average rank of flood susceptibility ML models

Parameter	Mean rank	Chi square	Asymp. Sig.
LR	2.55	22.313	0.000
SVM	3.37		
RPART	2.92		
RF	2.99		
KNN	3.17		

The z-scores for the KNN-RF, SVM-RF and LR-RF models surpassed the critical threshold of -1.96 and +1.96. Hence, it is evident that there was significant performance variations among the five models employed in flood susceptibility mapping.

Figure 4 instead the five flood susceptibility maps generated from the SVM, RF, Rpart, LR, and KNN models for the purpose of comparative analysis. Each model classified the area into five susceptibility classes using the natural breaks method: very low, low, moderate, high, and very high. The proportion of each class for each model was then calculated and depicted in Figure 5.

It is evident from Figure 4 that the use of ML models significantly alters the distribution of susceptibility zones in the FSM. The models' highest sensitivity in the high and very high classes was observed in urban areas near rivers.The models exhibited the highest sensitivity in urban zones situated near rivers, specifically within areas classified as high and very high risk.

The distribution of flood-prone area classes predicted by different ML models is illustrated in Figure 5. The outcomes highlight the models' diverse tendencies in categorizing areas into various levels of flood-prone classifications.

SVM and RF models showcase relatively balanced outcomes, capturing the low, medium, high and very high flood-prone classes.

Conversely, the greatest disparity between models was observed in estimating the very high-risk class, with the highest and lowest estimates obtained with KNN (42.16%) and Rpart (2.76%), respectively.

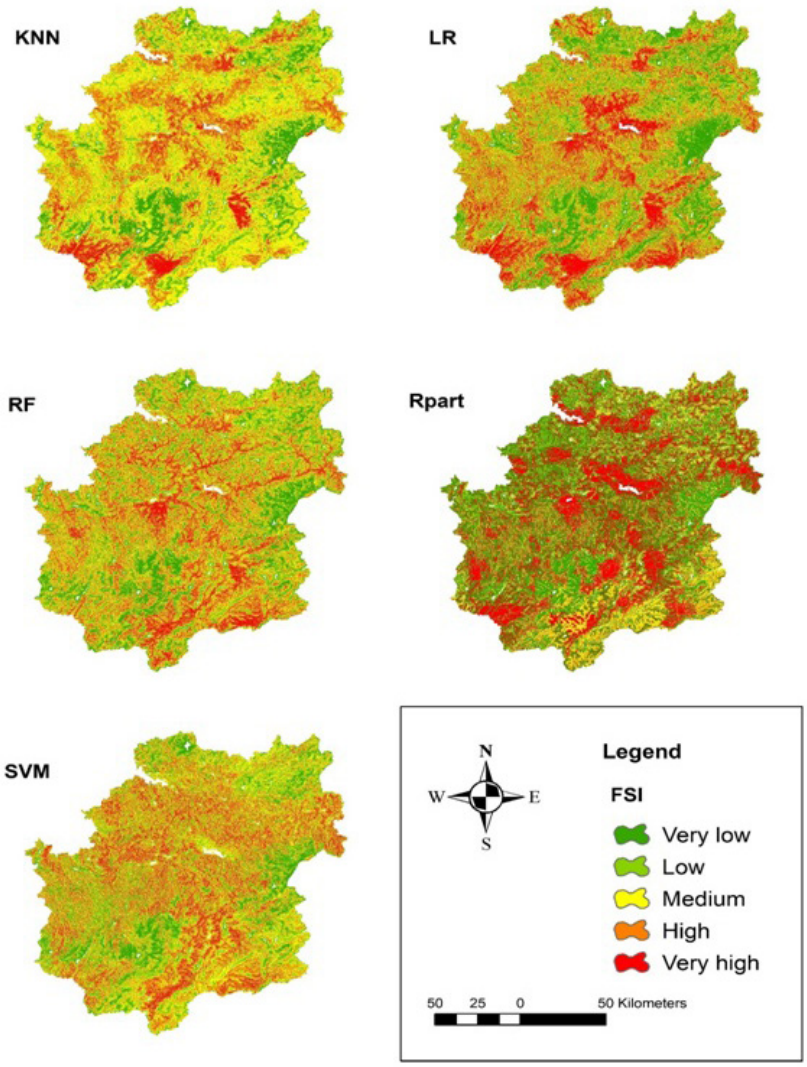

Figure 4. Flood-prone map for each model

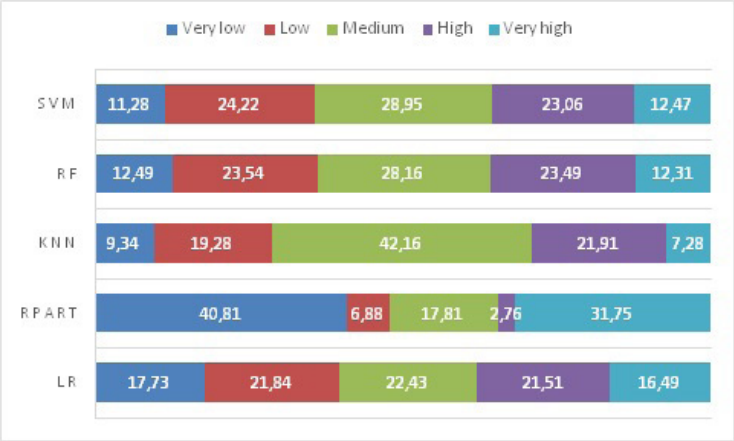

Figure 5. The distribution of the flood prone areas class for each model (%)

Interpretability of the RF model

The flood susceptibility map generated using the Random Forest model was analyzed using the PDP and LIME approaches to investigate the global explanations of ML models. Figure 6 illustrates the PDP results, where the x-axis represents the value of the input conditioning factor, and the y-axis shows the change in the predicted value.

Figure 6 indicates that all factors significantly impact flood susceptibility prediction. Specifically, the flood probability increases with the NDVI value up to a certain point, and it decreases significantly with an increase in TWI. The flood prediction probability is positively correlated with land use/land cover LULC classes and geology, while it is negatively correlated with drainage density and TRI.

Additionally, the LIME bar chart presented in (Figure 7) was used to select and partially interpret the influence of each factor on the flood prediction result. This analysis provided a basic understanding of why the models predict specific areas as flooded or non-flooded.

Hence, it can be concluded that the key factors influencing the flood susceptibility predictions in the Random Forest model are distance to river, drainage density, slope, NDVI, TWI, and LULC. These variables play a significant role in shaping the model's output.

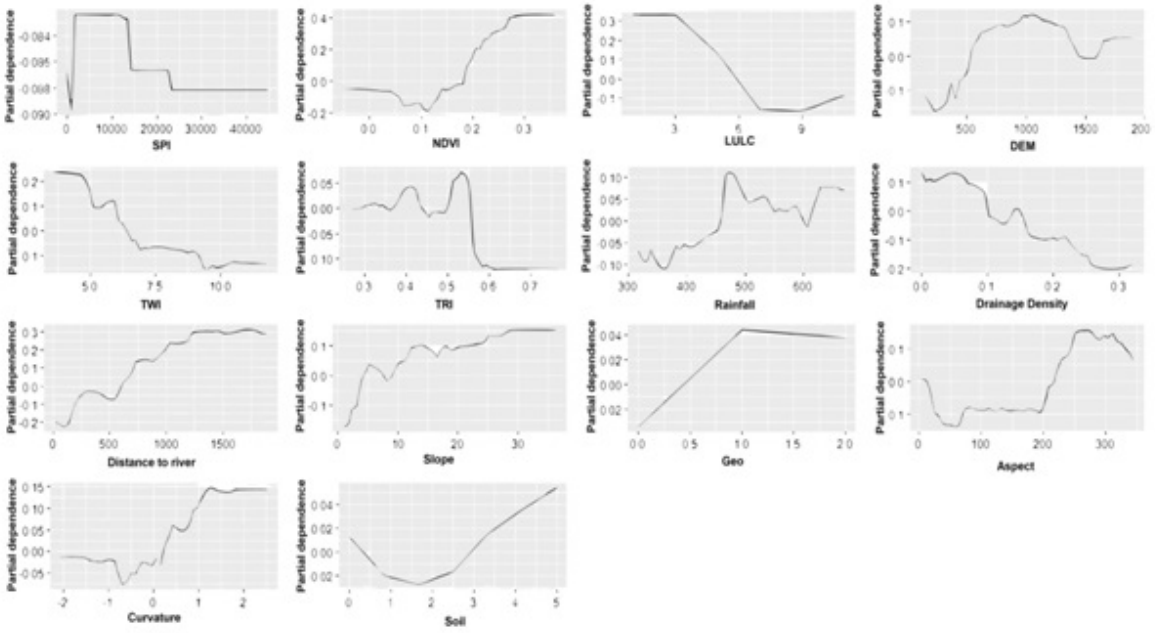

Figure 6. PDP analysis for RF model

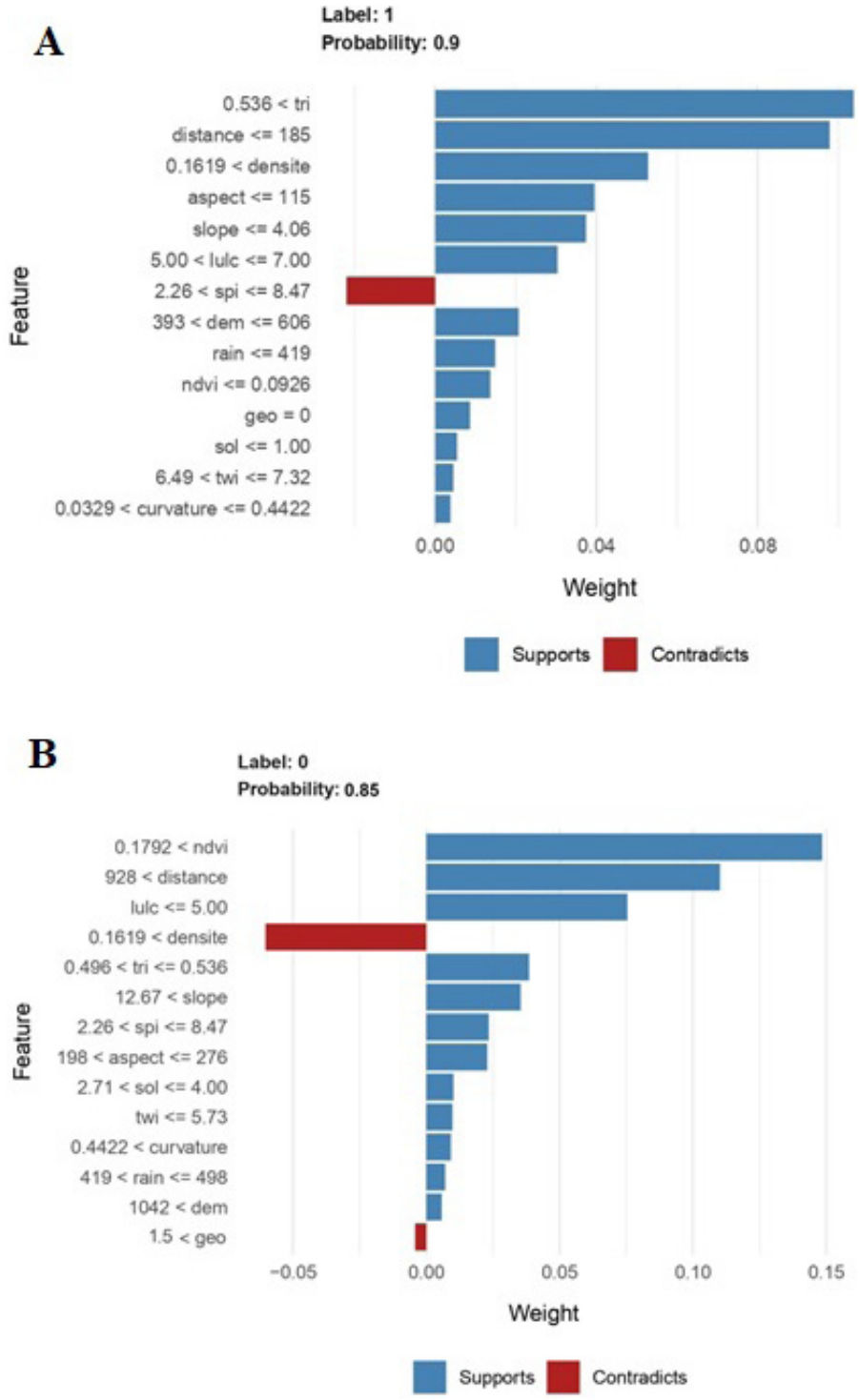

Figure 7. LIME analysis for RF model

DISCUSSION

The use of ML methods for flood susceptibility mapping has gained popularity in recent years (Munawar et al., 2021). However, interpreting the models and understanding the factors contributing to flood susceptibility has remained challenging (Mudashiru et al., 2021). Integrating partial dependence plots (PDP) and local interpretable model-agnostic explanations (LIME) in flood susceptibility mapping is a novel approach that enhances the interpretation of Random Forest models and factors influencing flood susceptibility, a technique not previously used in flood modeling studies. These approaches enables visualizing the relationship between input variables

and model output, explaining the contribution of each variable to predictions (Friedman, 2001; Mishra et al., 2017). Hence, incorporating PDP and LIME in flood susceptibility mapping signifies a significant progress in interpreting and explaining ML models, resulting in more accurate and reliable flood susceptibility maps.

The results of the multicollinearity analysis presented in Table 2 indicate a high intercorrelation between two flood conditioning factors, namely SPI and flow accumulation. To mitigate this issue, one of the variables was removed from the final model, resulting in reduced collinearity among the remaining 14 variables. This step is crucial for developing an accurate model to predict flood susceptibility (Bastiaan et al., 2022; Mahdizadeh Gharakhanlou and Perez, 2022). Indeed, previous studies emphasized the significance of addressing multicollinearity issues in flood susceptibility modeling (Chapi et al., 2017b; Dormann et al., 2013).

Notable differences in performance were observed among the various ML models used for flood susceptibility mapping, consistent with previous research comparing ML model performances (Althuwaynee et al., 2020). The RF model demonstrated superior performance in flood susceptibility mapping in the Fez-Meknes region, as evidenced by its high AUC and other metrics. This finding is consistent with previous studies that have emphasized the effectiveness of RF in flood susceptibility mapping across various geographic areas (Gocic and Trajkovic, 2013).

The uncertainty in flood susceptibility mapping associated with machine learning arises from various factors. Firstly, the complexity and non-linearity of flood dynamics introduce inherent uncertainty into the modeling process. Additionally, environmental processes like rainfall patterns, land cover changes, and topographic characteristics contribute to uncertainty in flood susceptibility predictions. Furthermore, individual characteristics and behaviors of each ML algorithm, along with the complexities of feature selection and interpretation, play a significant role. ML algorithm operates on unique assumptions, mechanisms, and sensitivities to dataset features, resulting in variability in predictions. Differences in learning algorithms, optimization techniques, and decision boundaries further contribute to uncertainty by influencing how models analyze and interpret input data. Varying priorities and interpretations of features among ML algorithm lead to divergent predictions and levels of uncertainty.

The superior reliability of RF model results suggests that ML models can proficiently conduct flood susceptibility mapping in the Fez-Meknes region. Generally, Q-values from geodetector tools offer a statistical measure of the relative importance of factors in flooding (Zhang et al., 2022), they do not offer a comprehensive and interpretable analysis of the contribution of each factor. Therefore, the use of PDP and LIME techniques in interpreting RF model results in this study provides deeper insights into flood prediction.

The findings from the PDP analysis revealed that all 14 flood-conditioning factors significantly influenced flood susceptibility prediction, with some factors exerting a positive effect while others had a negative impact. These results align with previous studies that have identified various factors, including land use/land cover, topography, and soil characteristics, as key variables in flood susceptibility mapping (Al-Areeq et al., 2024; Bouamrane et al., 2022; Menuka et al., 2022; Tehrany et al., 2019). Additionally, the LIME bar chart analysis offered further insights into the influence of each factor on flood prediction outcomes. This method randomly selected factors and partially interpreted their impact on flood prediction, aiding in identifying the primary causative factors behind flood susceptibility maps generated by the Random Forest model in specific areas. Key factors affecting flood susceptibility include proximity to rivers, drainage density, slope, NDVI, TRI, and LULC. Proximity to rivers plays a critical role as areas near rivers are more prone to flooding due to potential overflow and rising water levels (Rahman et al., 2021). Drainage density influences water conveyance, with higher density areas better equipped to handle excess runoff, reducing the risk of water accumulation and localized flooding(Ogden et al., 2011). Slope affects water flow, with steeper slopes contributing to rapid surface runoff and increased flood risk(Mohammadi et al., 2021). TRI provides insights into terrain roughness, influencing flow resistance and potentially prolonging flood duration in areas with rougher surfaces (Avand et al., 2021). Finally, LULC types impact flood risk by affecting surface runoff and infiltration (Tehrany et al., 2019). By considering these factors collectively, flood susceptibility models offer valuable insights for effective flood management and tailored mitigation strategies suited to specific regions and conditions.

PDP and LIME approaches offer more insightful and interpretable analyses of the relationship

between input factors and flooding prediction. While PDP shows the marginal effect of each input variable on predicted outcomes, LIME provides local interpretations of model predictions for specific instances. This information aids decision-makers in understanding the specific drivers of flooding, essential for developing effective mitigation strategies.

CONCLUSIONS

This study explored the utilization of ML methods in flood susceptibility mapping and the interpretation of factors contributing to flood susceptibility. Initially, a multicollinearity test was conducted to select the most relevant flood conditioning parameters. Results suggested that accumulation should be excluded from the final model due to collinearity issues. Subsequently, five ML algorithms were compared based on their performance evaluation, with the RF model demonstrating the highest predictive accuracy. Both Friedman and Wilcoxon rank tests confirmed significant distinctions among the models. Flood susceptibility maps produced by the models showed significant differences in the distribution of susceptibility zones, particularly observing high and very high susceptibility classes in urban areas near rivers. Lastly, interpretability of the RF model was assessed using PDP and LIME approaches, revealing significant impacts of distance to river, drainage density, slope, NDVI, TWI, and LULC on flood susceptibility prediction in the Fes-Meknes region.

Acknowledgements

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

REFERENCES

1. Al-areeq, A.M., Saleh, R.A.A., Ghaleb, M., Abba, S.I., Yaseen, M., (2024). Implication of novel hybrid machine learning model for flood subsidence susceptibility mapping : A representative case study in Saudi Arabia. *J. Hydrol. 630*, 130692. https://doi.org/10.1016/j.jhydrol.2024.130692
2. Althuwaynee, O.F., Balogun, A., Al Madhoun, W., (2020). Air pollution hazard assessment using decision tree algorithms and bivariate probability cluster polar function: evaluating inter-correlation clusters of PM10 and other air pollutants. *GIScience Remote Sens. 57*, 207–226. https://doi.org/10.1080/15481603.2020.1712064
3. Avand, M., Moradi, H., lasboyee, M.R., (2021). Using machine learning models, remote sensing, and GIS to investigate the effects of changing climates and land uses on flood probability. *J. Hydrol. 595*, 125663. https://doi.org/10.1016/j.jhydrol.2020.125663
4. Bastiaan, R.M., Salaki, D.T., Hatidja, D., (2022). Comparing the Performance of Prediction Model of Ridge and Elastic Net in Correlated Dataset. *Oper. Res. Int. Conf. Ser. 3*, 8–13. https://doi.org/10.47194/orics.v3i1.127
5. Belgiu, M., Dra, L., (2016). Random forest in remote sensing : A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing 114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011
6. Bouamrane, A., Derdous, O., Dahri, N., Tachi, S.E., Boutebba, K., Bouziane, M.T., (2022). A comparison of the analytical hierarchy process and the fuzzy logic approach for flood susceptibility mapping in a semi-arid ungauged basin (Biskra basin: Algeria). *Int. J. River Basin Manag. 20*, 203–213. https://doi.org/10.1080/15715124.2020.1830786
7. Breiman, L.E.O., 2001. Random Forests 5–32.
8. Bui, D.T., Ngo, P.T.T., Pham, T.D., Jaafari, A., Minh, N.Q., Hoa, P.V., Samui, P., (2019). A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena 179*, 184–196. https://doi.org/10.1016/j.catena.2019.04.009
9. Bunmi Mudashiru, R., Sabtu, N., Abdullah, R., Saleh, A., Abustan, I., (2022). Optimality of flood influencing factors for flood hazard mapping: An evaluation of two multi-criteria decision-making methods. *J. Hydrol. 612*, 128055. https://doi.org/10.1016/j.jhydrol.2022.128055
10. Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., Khosravi, K., (2017a). A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw. 95*, 229–245. https://doi.org/10.1016/j.envsoft.2017.06.012
11. Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Tien, D., (2017b). Environmental Modelling & Software A novel hybrid arti fi cial intelligence approach for fl ood susceptibility assessment. *Environ. Model. Softw. 95*, 229–245. https://doi.org/10.1016/j.envsoft.2017.06.012
12. Chen, W., Hong, H., Li, S., Shahabi, H., Wang, Y., Wang, X., Bin, B., (2019). Flood susceptibility

modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. *J. Hydrol. 575*, 864–873. https://doi.org/10.1016/j.jhydrol.2019.05.089

13. Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., Wang, X., Bian, H., Zhang, S., Pradhan, B., Ahmad, B. Bin, (2020). Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. *Sci. Total Environ. 701*, 134979. https://doi.org/10.1016/j.scitotenv.2019.134979
14. Cracknell, M.J., Reading, A.M., 2014. Computers & Geosciences Geological mapping using remote sensing data: A comparison of fi ve machine learning algorithms , their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci. 63*, 22–33. https://doi.org/10.1016/j.cageo.2013.10.008
15. Davoudi Moghaddam, D., Pourghasemi, H.R., Rahmati, O., (2019). Assessment of the contribution of geo-environmental factors to flood inundation in a semi-arid region of SW Iran: comparison of different advanced modeling approaches. *Adv. Nat. Technol. Hazards Res. 48*, 59–78. https://doi.org/10.1007/978-3-319-73383-8_3
16. Dolchinkov, N.T., (2024). Natural emergencies and some causes of their occurrence: a review. trends ecol. Indoor Environ. *Eng. 2*, 18–27. https://doi.org/10.62622/teiee.024.2.1.18-27
17. Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., Mcclean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography (Cop.). 36*, 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x
18. Friedman, H., (2001). Greedy function approximation : A gradient boosting machine author (s): Jerome H. Friedman Source : The Annals of Statistics, Oct 29(5) (Oct., 2001), 1189–1232 Published by : Institute of Mathematical Statistics Stable UR 29, 1189–1232.
19. Gocic, M., Trajkovic, S., (2013). Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia. *Glob. Planet. Change 100*, 172–182. https://doi.org/10.1016/j.gloplacha.2012.10.014
20. Gudiyangada Nachappa, T., Tavakkoli Piralilou, S., Gholamnia, K., Ghorbanzadeh, O., Rahmati, O., Blaschke, T., (2020). Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. *J. Hydrol. 590*, 125275. https://doi.org/10.1016/j.jhydrol.2020.125275
21. Guo, G., Wang, H., Bell, D., Bi, Y., Greer, K., (2003). *KNN Model-Based Approach in Classification* 986–996.
22. Holzinger, A., Goebel, R., Fong, R., Moon, T., Müller, Klaus-Robert, Samek, W., (2022). *xxAI-Beyond Explainable AI*.
23. Kavzoglu, T., Colkesen, I., (2009). A kernel functions analysis for support vector machines for land cover classification. Int. J. Appl. Earth Obs. *Geoinf. 11*, 352–359. https://doi.org/10.1016/j.jag.2009.06.002
24. Khaldi, L., Elabed, A., El Khanchoufi, A., (2023). Quantitative assessment of the relative impacts of different factors on flood susceptibility modelling: Case study of Fez-Meknes region in Morocco. *E3S Web Conf. 364*. https://doi.org/10.1051/e3sconf/202336402005
25. Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.B., Gróf, G., Ho, H.L., Hong, H., Chapi, K., Prakash, I., (2019). A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *J. Hydrol. 573*, 311–323. https://doi.org/10.1016/j.jhydrol.2019.03.073
26. Kia, M.B., Pirasteh, S., Pradhan, B., Mahmud, A.R., Nor, W., Sulaiman, A., Moradi, A., (2012). An artificial neural network model for flood simulation using GIS : Johor River Basin , Malaysia. *Environ. Earth Sci*. 251–264. https://doi.org/10.1007/s12665-011-1504-z
27. Lee, S., Sambath, T., (2006). Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environ. Geol. 50*, 847–855. https://doi.org/10.1007/s00254-006-0256-7
28. Mahdizadeh Gharakhanlou, N., Perez, L., (2022). Spatial prediction of current and future flood susceptibility: examining the implications of changing climates on flood susceptibility using machine learning models, *Entropy*. https://doi.org/10.3390/e24111630
29. Maxwell, A.E., Warner, T.A., Fang, F., Maxwell, A.E., Warner, T.A., (2018). Implementation of machine-learning classification in remote sensing : an applied review sensing: an applied review. *Int. J. Remote Sens. 39*, 2784–2817. https://doi.org/10.1080/01431161.2018.1433343
30. Mc fadden, D., 1974. Conditional logit analysis of qualitative choice behavior. *Front. Econom.* https://doi.org/10.1080/07373937.2014.997882
31. Menuka, M., Sachin, T., Santosh, A., Bikram, S., Bikram, M., Amir, S., (2022). Flood susceptibility assessment using machine learning approach in the Mohana- Khutiya River of Nepal. *Nat. Hazard Res.* https://doi.org/10.1016/j.nhres.2024.01.001
32. Mishra, S., Sturm, B.L., Dixon, S., (2017). Local interpretable model-agnostic explanations for music

content analysis. Proc. 18th Int. Soc. *Music Inf. Retr. Conf. ISMIR*, 537–543.

33. Mohammadi, M., Darabi, H., Mirchooli, F., Bakhshaee, A., Torabi Haghighi, A., (2021). Flood risk mapping and crop-water loss modeling using water footprint analysis in agricultural watershed, northern *Iran. Nat. Hazards 105*, 2007–2025. https://doi.org/10.1007/s11069-020-04387-w
34. Mountrakis, G., Im, J., Ogole, C., 2011. ISPRS Journal of photogrammetry and remote sensing support vector machines in remote sensing : A review. *ISPRS J. Photogramm. Remote Sens. 66*, 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001
35. Mousavi, S.M., Ataie-Ashtiani, B., Hosseini, S.M., (2022). Comparison of statistical and MCDM approaches for flood susceptibility mapping in northern Iran. *J. Hydrol. 612*, 128072. https://doi.org/10.1016/j.jhydrol.2022.128072
36. Mudashiru, R.B., Sabtu, N., Abustan, I., (2021). Quantitative and semi-quantitative methods in flood hazard/susceptibility mapping: a review. Arab. *J. Geosci. 14*. https://doi.org/10.1007/s12517-021-07263-4
37. Munawar, H.S., Hammad, A.W.A., Waller, S.T., (2021). A review on flood management technologies related to image processing and machine learning. *Autom. Constr. 132*, 103916. https://doi.org/10.1016/j.autcon.2021.103916
38. Ogden, F.L., Raj Pradhan, N., Downer, C.W., Zahner, J.A., (2011). Relative importance of impervious area, drainage density, width function, and subsurface storm drainage on flood runoff from an urbanized catchment. *Water Resour. Res. 47*. https://doi.org/10.1029/2011WR010550
39. Pham, B.T., Luu, C., Phong, T. Van, Nguyen, H.D., Le, H. Van, Tran, T.Q., Ta, H.T., Prakash, I., (2021). Flood risk assessment using hybrid artificial intelligence models integrated with multi-criteria decision analysis in Quang Nam Province, Vietnam. J. Hydrol. 592, 125815. https://doi.org/10.1016/j.jhydrol.2020.125815
40. Rahman, M., Ningsheng, C., Mahmud, G.I., Islam, M.M., Pourghasemi, H.R., Ahmad, H., Habumugisha, J.M., Washakh, R.M.A., Alam, M., Liu, E., Han, Z., Ni, H., Shufeng, T., Dewan, A., (2021). Flooding and its relationship with land cover change, population growth, and road density. Geosci. *Front. 12*, 101224. https://doi.org/10.1016/j.gsf.2021.101224
41. Saha, S., Arabameri, A., Saha, A., Blaschke, T., Thao, P., Ngo, T., Ha, V., Band, S.S., (2021). Science of the total environment prediction of landslide susceptibility in Rudraprayag , India using novel ensemble of conditional probability and boosted regression tree-based on cross-validation method. *Sci. Total Environ. 764*, 142928. https://doi.org/10.1016/j.scitotenv.2020.142928
42. Sayers, P., Galloway, G., Penning-Rowsell, E., Yuanyuan, L., Fuxin, S., Yiwei, C., Kang, W., Le Quesne, T., Wang, L., Guan, Y., (2015). Strategic flood management: ten 'golden rules' to guide a sound approach. *Int. J. River Basin Manag. 13*, 137–151. https://doi.org/10.1080/15715124.2014.902378
43. Tehrany, M.S., Jones, S., Shabani, F., (2019). Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena 175*, 174–192. https://doi.org/10.1016/j.catena.2018.12.011
44. Tehrany, M.S., Pradhan, B., Jebur, M.N., (2013). Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol. 504*, 69–79. https://doi.org/10.1016/j.jhydrol.2013.09.034
45. Zhang, R., Chen, Y., Zhang, X., Ma, Q., Ren, L., (2022). Mapping homogeneous regions for flash floods using machine learning: A case study in Jiangxi province, China. *Int. J. Appl. Earth Obs. Geoinf. 108*. https://doi.org/10.1016/j.jag.2022.102717