Natural Hazards (2025) 121:10299–10321
https://doi.org/10.1007/s11069-025-07197-0

ORIGINAL PAPER

Comparative assessment of machine learning models for landslide susceptibility mapping: a focus on validation and accuracy

Mohamed M. Abdelkader[1,2] · **Árpád Csámer**[1,3]

Received: 13 September 2024 / Accepted: 20 February 2025 / Published online: 13 March 2025

Abstract
Accurate landslide susceptibility mapping (LSM) is critical to risk management, especially in areas with significant development. Although the receiver operating characteristic–area under the curve (ROC–AUC) performance metrics are commonly used to measure model effectiveness, showed that these are not enough to check the reliability of the generated maps. In this study, the effectiveness of three machine learning models—logistic regression (LR), random forest (RF), and support vector machine (SVM)—were evaluated and compared in predicting landslide risk in a hilly region east of Cairo, Egypt. A comprehensive dataset was gathered to achieve that, including 183 landslide and 183 non-landslide locations, which were detected through fieldwork and high-resolution satellite imagery. Fourteen conditioning factors from different categories; topographical, geological, hydrological, anthropological, and trigger-related variables, were used as independent factors during the generation of the different LSM. All three models achieved high ROC–AUC values, with RF scoring 0.95, SVM 0.90, and LR 0.88, indicating strong performance. However, further assessment with additional performance metrics like accuracy (ACC), recall, precision, F1 score, and check rationality of the maps revealed key differences. Among the models, only the RF model appeared as the most reliable, with superior across all performance metrics, and fewer misclassifications in critical areas. In contrast, SVM and LR exhibited higher misclassification rates for both landslide-prone and safe locations. These findings show that high ROC–AUC values do not always equate to practical reliability.

Keywords Machine learning · Random forest · Landslide · SHAP value · ROC–AUC · Cairo

✉ Árpád Csámer
csamera@unideb.hu

Mohamed M. Abdelkader
mohamed_abdelkader@sci.asu.edu.eg

[1] Department of Mineralogy and Geology, University of Debrecen, Debrecen 4032, Hungary

[2] Geology Department, Faculty of Science, Ain Shams University, Cairo, Egypt

[3] Cosmochemistry and Cosmic Methods Research Group, University of Debrecen, Debrecen 4032, Hungary

1 Introduction

Landslides are one of the most common hazards against any development plan in hilly regions (Liao et al. 2022). Nowadays, because of global climate change and expanding urbanization, the number of these landslides has increased, which has cost a lot of lives and infrastructure (Youssef et al. 2024). Hence, exploring a reliable approach for evaluating landslide hazards has become very critical to avoid or reduce disasters (Akinci and Zeybek 2021). Therefore, the importance of landslide susceptibility maps (LSM) increased due to their potential to predict landslides and identify areas prone to landslides. Such maps are used as an effective tool in landslide management and mitigating disaster losses within many regions (Guo et al. 2023; Wang et al. 2023). Owing to the continuous development of geographic information systems (GIS), several approaches have been used to generate LSM which depend on physical-based, qualitative, and quantitative methods (Bopche and Rege 2022; Liu et al. 2024).

Qualitative methods are highly subjective and rely on expert judgment to assign weight to landslide conditioning factors (Effat and Hegazy 2014). The problem with this approach is that the difference in user opinions leads to different results for the same areas and conditions (Tyagi et al. 2022). In contrast, the physical approach relies on fundamental laws of physics to understand how slopes move, and the data comes from direct field observations and laboratory experiments (Bednarik et al. 2024; Wang et al. 2024). Although these models are useful in places where landslide inventory maps are not available, they are difficult to use in large areas since they need detailed geotechnical and hydrological data (Chicas et al. 2024).

Quantitative methods, including statistics-based, machine learning-based (ML), and deep learning-based (DL) models, have been developed to overcome the drawbacks of the two approaches mentioned above (Zhang et al. 2021; Yao et al. 2023; Usta et al. 2024). Quantitative methods build numerical relations between landslide events and the landslide-responsible conditional factors in the same area. Among these, ML has attracted a lot of interest due to its capacity to produce predictive models and analyze complex patterns. During the past decade, many studies have used different ML methods, such as logistic regression (LR) (Chowdhury et al. 2024), support vector machines (SVM) (Liu et al. 2021; Akinci and Zeybek 2021), artificial neural network (ANN) (Akinci 2022), random forest (RF) (Shang et al. 2023; Zhang et al. 2023), extreme gradient boosting (XGBoost) (Pradhan and Kim 2020; Zhang et al. 2023) to generate LSM.

Despite numerous studies related to the application of ML for LSM, no universally reliable model performs effectively across different regions (Alkan Akinci et al. 2024). Numerous studies have compared various models and approaches to evaluate their performance and determine the most suitable methods for different regions (Youssef and Pourghasemi 2021; Wang et al. 2023; Khalil et al. 2023). However, the findings consistently indicate that model outcomes vary significantly due to differences in parameters, data quality, geographic characteristics, and environmental factors. These variations highlight the need for further research aimed at developing adaptable models, particularly for regions where a limited number of studies have been conducted (Akinci 2022; Khalil et al. 2023).

Another significant gap in previous LSM studies is the dependence on receiver operating characteristic– area under the curve (ROC–AUC) as the only primary metric of validation. Most of the studies regarded a value of ROC–AUC higher than 0.7 as an indication of the efficiency of the model (Sun et al. 2018; Kamran et al. 2021; Lee and Lee 2024); this threshold gives a limited perspective into the reliability of the model. Therefore, the

dependence on this metric may be misleadingly interpretative and lead to wrong judgments of model performance. However, some studies have attempted to incorporate additional statistical metrics, such as overall accuracy (ACC), recall, and F1-score (Zhao et al. 2022; Wang et al. 2023; Yao et al. 2023), but their use remains limited. Moreover, few research articles have highlighted the importance of spatial validation techniques in ensuring the rationality of generated LSMs. Such analyses are very important in confirming whether the models can correctly classify landslide-prone areas and safe locations within the study region. Without this type of validation, the reliability of the susceptibility maps is not guaranteed, and their practical use is limited (Yavuz Ozalp et al. 2023; Usta et al. 2024).

This research aims to enhance the LSM within arid regions, where studies remain limited. The study's contribution lies in presenting a comprehensive framework that integrates ML techniques with strong evaluation and interpretability techniques, considering the geological and environmental settings of arid areas. To achieve that, logistic regression (LR), random forest (RF), and support vector machine (SVM) are employed and compared. Also, a strong validation framework is proposed, combining traditional metrics such as ROC–AUC, recall, ACC, precision, and F1-score with spatial validation techniques to ensure geographically rational predictions. Moreover, the SHAP (SHapley Additive exPlanations) values are applied to identify the most influential factors related to landslide susceptibility, making the models more transparent and interpretable. The research objective is to propose and validate a comprehensive LSM that improves the reliability of the model and provides valuable insights for land-use planning and disaster risk reduction. This work sets a benchmark for future studies in this field, offering a significant contribution to advancing LSM, especially in arid regions.

2 The study area

Egypt is one of the most highly populated countries in Africa and the Middle East. Especially in Cairo, the high population density (Fig. 1a) leads to significant social and economic challenges. Therefore, it is of critical importance to redistribute this density to uninhabited regions such as hilly areas around Cairo. One of these new areas is our study area. It is part of Cairo Suez District and is located southeast of Cairo and west of the New Administrative Capital of Egypt (Fig. 1b). This area is bounded by longitudes 31°16′27.8″ E and 31°42′21.5″ E and latitudes 29°42′0″ N and 29°59′6″ N and extending over 756 km^2. The region has witnessed major urban development in terms of the construction of many residential areas and the construction of new roads, and the first high-speed electric train line in Egypt will pass through it. Only one attempt has been made so far to deal to study landslide susceptibility in this area; however, it used the qualitative method, which is, as discussed before, a fully subjective assessment (Effat and Hegazy 2014).

3 Methodology

3.1 Landslide inventory mapping

Following the principle that 'the past is the key to the future', future landslides will likely to occur under conditions like those of past events (Youssef et al. 2024). Therefore, a landslide inventory map, which illustrates the geographical distribution of

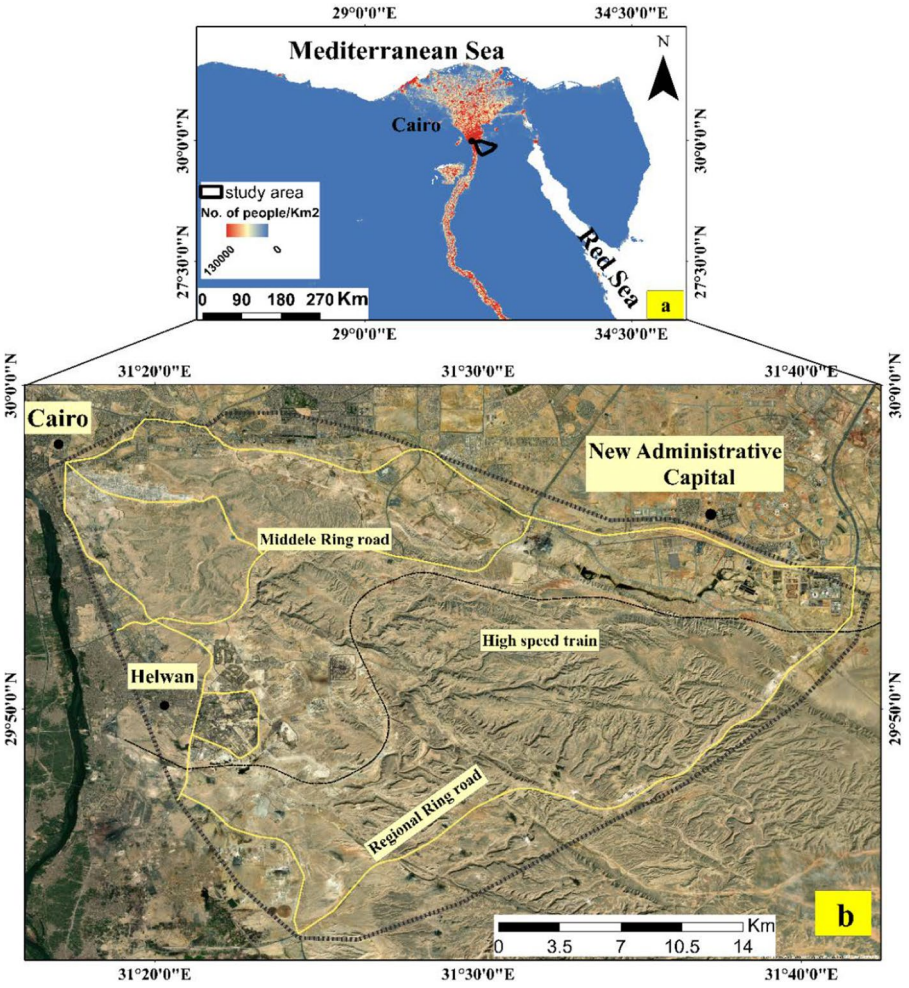

Fig. 1 **a** The distribution of population density in Egypt (after WorldPop and Bondarenko (2020); **b** Location of the study area

existing landslides in specific regions, is crucial for ensuring the accuracy of landslide susceptibility models (Gaidzik and Ramírez-Herrera 2021; Wang et al. 2023). Since there is no record of the landslides in Egypt, a landslide inventory database for the study area was created through extensive fieldwork from February 2023 to August 2023. This work was complemented by analyzing high-resolution time-series images from Google Earth. As a result, a comprehensive database of 183 identified landslide locations, including 110 sites detected during field surveys, was established (Fig. 2). The identified landslides include sliding, falling, and toppling events. To avoid bias, the same number of safe locations was used (Akinci and Zeybek 2021).

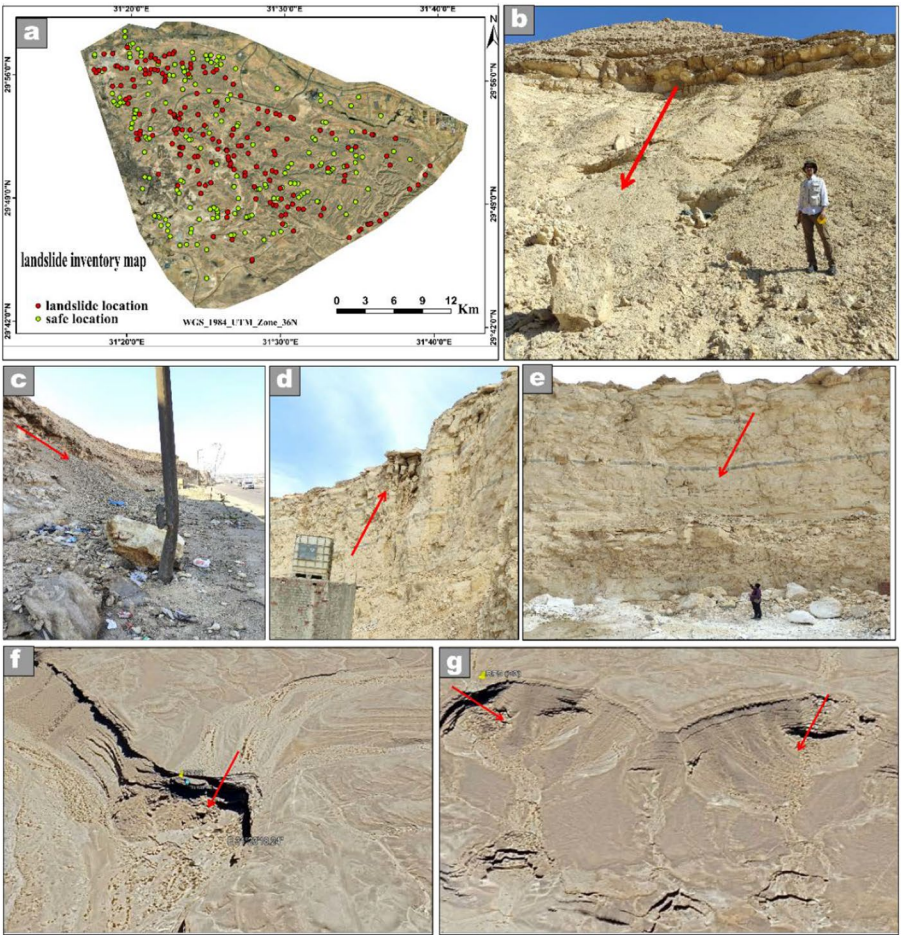

Fig. 2 **a** Landslide inventory map of the study area indicating the recorded landslides and safe locations; **b**–**e** Images from the field showing landslide by the red arrows; **f** and **g** images from Google Earth showing landslides by the red arrows

3.2 Preparation and preprocessing of landslide conditioning factors

In the current study, fourteen landslide conditioning factors, representing five categories (topographic, geologic, hydrological, anthropogenic, and triggering factors), were selected based on field observations and a review of previous studies (Yavuz Ozalp et al. 2023; Liu et al. 2024). These factors were preprocessed as an essential step before being used in the modeling process. Firstly, all predictor datasets were reprojected into a uniform coordinate system and resampled to a consistent resolution of 30 m to ensure accuracy (Wang et al. 2024). Additionally, the predictor variables were normalized to a common scale to enhance the performance of ML algorithms (Liao et al. 2022). Also, the multicollinearity among the predictors has been checked to reduce redundancy and improve model robustness. Then, this dataset was split into training and testing subsets for performance evaluation. These fourteen factors, sourced from various datasets, are

detailed in Table 1. The following sections provide a step-by-step explanation of the preprocessing methods.

3.2.1 Topographic factors

Many studied found that topographic factors have a major role in landslide occurrence (Agboola et al. 2024; Chicas et al. 2024). In this study, key topographic factors included slope, aspect, elevation, profile curvature, and plan curvature. Depending on the relation between the steep slopes and the probability of landslide events, the slope was considered a vital factor in any LSM (Akinci 2022). The slopes in the study area are either natural or artificial slopes and their values are between 0.00° and 62° (Fig. 3a). Aspect is also studied according to its importance to show the direction and grade of terrain via its value reflecting the maximum variation in the downslope direction from a specific cell to its surroundings (Wang et al. 2023). The aspect map has been constructed for the study area, with values ranging from 0° to 360° (Fig. 3b). The elevation map is also a vital topographic parameter in studying landslide occurrence by recognizing the relative relief of the study area (Bopche and Rege 2022). It was found that higher elevations are highly vulnerable to landslides. The elevation map showed that region spans from − 16 to 580 m (Fig. 3c). The last topographic factors in this study are plan curvature (Fig. 3d) and profile curvature (Fig. 3e). Plan curvature describes the concavity or convexity of the slope which affects water flow and erosion patterns and, varies from − 4.62 to 4.67. The profile curvature, ranging from − 6 to 5.51, affects the acceleration of water flow on the slope surface, thereby impacting erosion and deposition processes (Akinci 2022).

3.2.2 Geological factors

The study area is mainly covered by Middle and Upper Eocene sedimentary rocks. The Middle Eocene rocks comprise the Gebel Hof Formation and Observatory Formations, which are characterized mainly by white chalky limestone interbedded with marl (Moustafa and Abd-Allah 1991). While the Upper Eocene rocks comprise the El Qurn, Wadi Garawi, and Wadi Hof Formations, which are composed of limestone, sandy limestone, sand, and marl (Moustafa et al. 1985). Besides the Eocene rock units, Quaternary wadi deposits also occur. The lithological units of the study area were digitized from the composite map made from the previous works (Moustafa et al. 1985; Moustafa and Abd-Allah 1991) (Fig. 3f). These units controlled the chemical, mechanical, and hydraulic properties of rocks which in turn impacted the landslide events (Akinci 2022).

The second significant geological factor is the distance to the fault. Near faults, rock strength is considerably weak, which results in frequent landslide disasters. Therefore, the distance from faults is a major factor in studying LSM (Wang et al. 2023). The fault map was generated from the previous work of Moustafa et al. (1985) and Moustafa and Abd-Allah (1991). After that, the distance to faults map was created by using the Euclidean distance tool in ArcMap software. The values read from the map are between 0 and 6457 m (Fig. 3g).

3.2.3 Hydrological factors

Topographic wetness index (TWI) is also a vital factor in landslide investigation that identifies the water saturation zones within regions with a large catchment area and a

Table 1 Source of data

Conditional factor	Source	Scale
Slope, Aspect, Elevation, Plan curvature, Profile curvature, Topographic wetness index (TWI)	Extract from SRTM DEM	30 × 30 m
Lithology, Distance to faults	Composite geologic map after (Moustafa et al. 1985; Moustafa and Abd-Allah 1991)	1:50,000
Land use / land cover (LULC)	Generated from Landsat 8 Oli images	30 × 30 m
Distance to roads	Generated from OpenStreetMap server: (https://www.openstreetmap.org/)	Vectors
Rainfall	Constructed from version 4 of the Climatic Research Unit (CRU) time series (TS) monthly high-resolution gridded multivariate climate dataset	0.5 × 0. 5 degree grids
Magnitude and depth of earthquake	ISC-GEM global instrumental earthquake catalogue	Stations

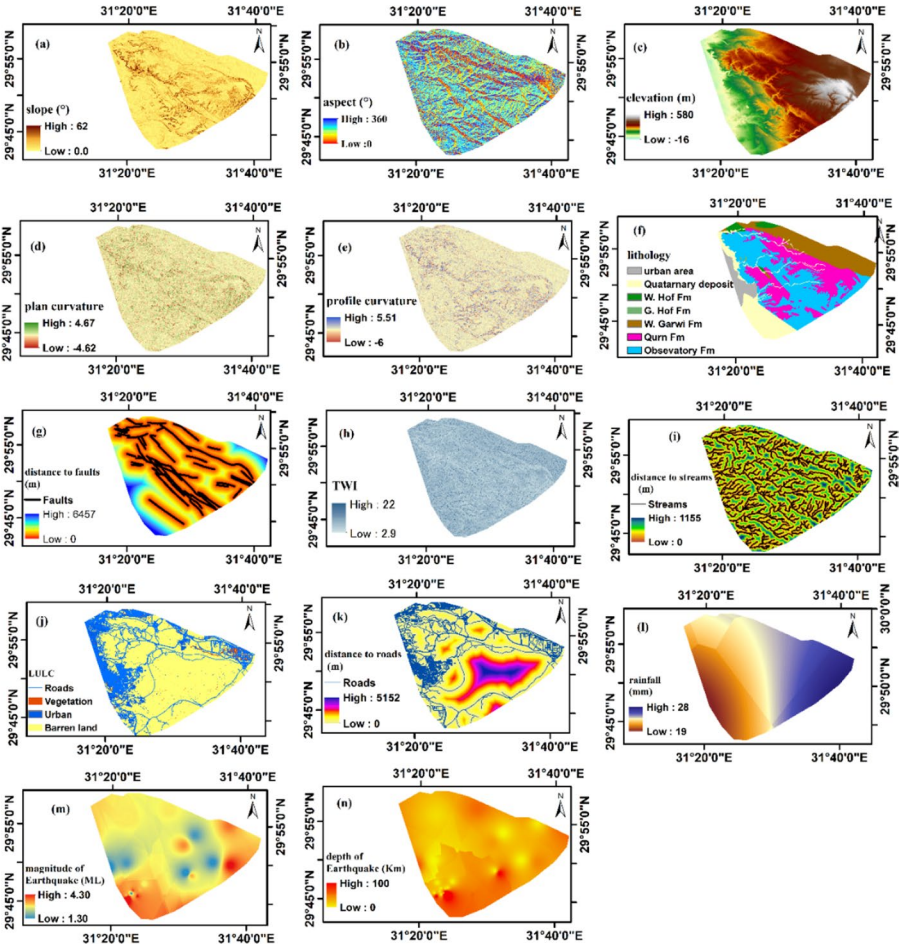

Fig. 3 Landslide conditional factors; **a** Slope, **b** Aspect, **c** Elevation, **d** Plan curvature, **e** Profile curvature, **f** Lithology, **g** Distance to faults, **h** TWI, **i** Distance to streams, **j** LULC, **k** Distance to roads, **l** Rainfall, **m** Magnitude of earthquakes, **n** Depth of earthquakes

gentle slope gradient (Singh et al. 2021). The values of TWI of the study area are calculated using Eq. 1, ranging from 2.9 to 22, as shown in (Fig. 3h).

$$\text{Twi} = \ln\left(\frac{\text{flow accumulation}}{\tan(slope)}\right) \tag{1}$$

Moreover, being close to streams and wadis can reduce slope shear resistance by removing sediments from the slope base, which in turn increases the chance of landslide events (Farooq and Akram 2021). Therefore, the distance to streams map for the study area was created through the hydrological tool and Euclidean distance tool in ArcMap software, as displayed in Fig. 3i, with values ranging from 0 to 1155 m.

3.2.4 Anthropogenic factors

Human activity is one of the main common reasons for the occurrence and reactivation of landslides. In this study, human activity is represented by LULC and distance to road maps. The LULC plays an essential role by showing human-induced modifications that affect vegetation cover and slope design (Chen et al. 2019). Three LULC categories have been recognized from the Landsat 8 satellite imagery (Fig. 3j): Barren land, Vegetation, and Urban area.

Distance to road map for the study area was generated through the Euclidean distance tool in ArcMap by using road data from the OpenStreetMap server, as shown in (Fig. 3k) the distances range from 0 to 5152 m. The construction of roads and railroads in hilly regions can cause landslides due to the extensive excavation of natural slopes and modification of the slope design (Tanyaş et al. 2022).

3.2.5 Trigger factors

Numerous studies have shown that heavy rainfall and earthquakes are responsible for the initiation of landslides (Fan et al. 2019; Kamran et al. 2021). Heavy rainfall can penetrate the rock mass through the discontinuities or the primary porosity of the rocks, leading to shear strength reduction and producing a landslide (Peruccacci et al. 2017). The rainfall distribution of the study area is obtained from data from version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset (Harris et al. 2020). The rainfall map was created by using the inverse distance weight (IDW) interpolation technique in ArcMap, and values range from 19 to 28 mm (Fig. 3l).

Earthquake shaking can also weaken slopes across the landscape, making them more prone to failures and an increased rate of occurrence (Bai et al. 2020). The data on the magnitude and depth of earthquakes are obtained from the ISC-GEM Global Instrumental Earthquake Catalogue (International Seismological Centre 2023)**.** The earthquake magnitude and depth distribution maps (Fig. 3m, Fig. 3n, respectively) were generated by the IDW technique in ArcMap. Earthquakes' magnitude ranged from 1.30 to 4.30 ML, and the depths of earthquakes range between 0 to 100 km.

3.3 Preparation of training and testing data

Although there is no standard accepted guideline for creating training and validation subsets, many researchers prefer to use a 70:30 ratio for selecting landslide samples in LSM studies (Khalil et al. 2023; Usta et al. 2024; Youssef et al. 2024). In this study, 128 landslide locations (70%) were allocated for training, while 55 landslide locations (30%) were used for validation. Furthermore, during both the training and validation stages, the number of positive and negative samples in the datasets was kept equal, maintaining a 1:1 ratio (Akinci and Zeybek 2021).

3.4 Multicollinearity test

One of the initial steps in LSM is a multicollinearity analysis, which checks the independence of the input variables of the model. This analysis highlights the selected influence factors that are highly intercorrelated, thus affecting the accuracy of LSMs (Lee et al. 2020; Wang et al. 2023; Usta et al. 2024). In LSM studies, the indicators used most frequently

for assessing multicollinearity are tolerance (TOL) and the variance inflation factor (VIF), which can be calculated using Eqs. 2 and 3 (Pradhan and Kim 2020; Akinci 2022). VIF represents the inflation in the variance of a regression coefficient that is caused by multicollinearity, while TOL, as the reverse of VIF, is also an indicator of multicollinearity among variables (Yavuz Ozalp et al. 2023). A variable is considered having high multicollinearity and should be excluded from the model if its VIF is greater than 10 or if its TOL is less than 0.1 (Akinci 2022; Khalil et al. 2023).

$$VIF = \frac{1}{1 - Rj^2} \tag{2}$$

$$\mathrm{TOL} = \frac{1}{\mathrm{VIF}} \tag{3}$$

where R_j is the correlation coefficient of a particular conditioning factor on the remaining factors.

3.5 Machine learning methods for landslide susceptibility assessment

Receny, ML approaches and soft computing techniques have become essential tools for analyzing and modeling different geotechnical hazards owing to their ability to deal with the complex behavior of the geotechnical system (Zhang et al. 2020; Phoon and Zhang 2023). During this work, three ML methods; logistic regression (LR), random forest (RF), and support vector machine (SVM) were applied to assess their effectiveness in mapping the accurate LSM.

3.5.1 Logistic regression (LR)

LR is one of the most widely used ML methods because it requires minimum preparation regarding the quality and size of samples (Rai et al. 2022; Chowdhury et al. 2024). It can deal with various data types, whether continuous or discrete (Sun et al. 2018). During this work, LR was applied through the "glm" package in R software to identify risky zones within the study area (Akinci and Zeybek 2021). The LSM has been generated by the LR model using Eq. 4.

$$\mathrm{P} = \frac{1}{1 + e^{-z}} \tag{4}$$

where P describes the possibility of a landslide event, varying from 0 to 1, Z is expressed as an independent variable, and its coefficient, as shown in Eq. 5.

$$\mathrm{Z} = \beta_0 + \beta_1 \mathrm{x}_1 + \beta_2 \mathrm{x}_2 + \cdots + \beta_\mathrm{n} \mathrm{x}_\mathrm{n} \tag{5}$$

where β_0, β_1, β_2,…, β_n are the model parameters that reflect the contributions of each factor, while x_1, x_2,.., x_n are the independent variables that influence the occurrence of landslides.

3.5.2 Random forest (RF)

RF is considered one of the most effective ML methods introduced by Breiman (2001). It takes advantage of decision trees and uses the bagging method to perform powerful

classifying (Zhang et al. 2023; Usta et al. 2024). RF classifier generates many independent decision trees using different bootstrapped samples of the training data (Wang et al. 2020). This approach helps the algorithm overcome the limitations of individual trees and enhance predictive performance, making it a popular tool in the field of ML (Akinci 2022). This study employed the RF method from the 'caret' package (Kuhn 2008) in the R platform to implement the RF algorithm. During the training process, hyperparameter optimization was performed to enhance the model's predictive performance. After multiple iterations, the best-tuned hyperparameters for the RF model were identified as 500 trees for ntree and 2 for mtry. To ensure the robustness and reliability of the results, the RF model was evaluated using a threefold cross-validation with 10 repetitions, leading to a total of 30 iterations (Akinci 2022; Yavuz Ozalp et al. 2023).

After that, the significance of each factor influencing landslide in the RF model was assessed using the Mean Decrease Gini (MDG) values. These values measure the extent to which a given predictor variable decreases the Gini impurity across all decision trees in the forest, which reflects the contribution of each factor during the model's training process (Chowdhuri et al. 2021).

3.5.3 Support vector machine (SVM)

SVM showed extensive application in LSM (Hong et al. 2017; Kamran et al. 2021). The SVM aims to construct an optimal hyperplane that separates the classes with a decision surface that maximizes the margin between them (Zhao et al. 2022). The points that lie closest to the hyperplane on either side are called support vectors, and the classification boundary is determined only by these support vectors, rather than by the other data points (Cortes and Vapnik 1995). In many classification problems, datasets are not linearly separable. In such cases, kernel functions map the data into a higher-dimensional space, where it becomes separable. There are several types of kernel functions, including linear, polynomial, radial basis function (RBF), and sigmoid kernels (Liu et al. 2021). RBF kernels are commonly used in the generation of LSMs due to their robustness and ability to provide more accurate predictions (Akinci and Zeybek 2021). During this research, SVM was applied via the package of "e1071" in the R platform to construct the LSM (Meyer et al. 2024). To optimize model performance, we conducted tuning using hyperparameters; cost = 1 and gamma = 0.1, based on maximizing accuracy.

After that, the predictor's influence is determined to estimate the relative importance of each predictor in shaping the decisions of the SVM model and aiding in the interpretation and understanding of the model's predictive capabilities. This influence was determined by identifying the support vectors and their associated coefficients representing each support vector's weight in defining the decision boundary.

3.6 SHapley additive exPlanation (SHAP)

LSMs generated by ML models are becoming more advanced, complex, and able to provide accurate predictions. However, as these models became powerful, it became harder to understand how they reached those predictions. This has led to the development of new tools to improve the interpretability of ML models (Usta et al. 2024). SHAP values are one of the most recent methods used to enhance the interpretability of ML methods used in LSM by quantifying the contribution of each predictor to the model's prediction (Lee and Lee 2024). The SHAP value is based on the Shapley values of cooperative game theory,

which offers a fair method for attributing a contribution to each feature in a model. These values are calculated by evaluating all possible feature combinations, calculating each feature's marginal contribution to the prediction across these combinations, and then averaging these contributions (Lee and Lee 2024). This provides one value per feature that represents how much that feature contributed to the individual prediction. The advantage of this approach is that it ensures fair and consistent attribution of feature importance, avoiding biases that may arise from other methods. As a result, SHAP has become a powerful tool for analyzing and understanding complex models within the field of ML (Usta et al. 2024).

Mathematically, the SHAP value for a feature i in a model with n features is defined as:

$$\phi_i(f) = \sum_{s \subseteq N \setminus \{i\}} \frac{|S|!(n - |s| - 1!}{n!} [f(S \cup \{i\}) - f(s)] \tag{6}$$

where (f) is the SHAP value for feature i, N is the set of all features, S is a subset of features excluding i, |S| is the number of features in subset S, f(S) is the model output using the subset S of features, and f(S ∪ {i}) is the model output when feature i is included in the subset S.

3.7 Evaluation of the models

Recognizing the most accurate model for the specific study area is essential in using ML techniques for LSM. This step is mainly done by using the testing part of the database. Three different ways to measure the accuracy of each model were applied during this work. Initially, four statistical indexes, namely accuracy (ACC), precision, recall, and F1 score were calculated based on the confusion matrix, which compares the predicted and actual values for each case using the Eqs. 7–10. The confusion matrix involves four primary components: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The TP and TN show the count of samples correctly identified as positive and negative, respectively. On the other hand, FP and FN give the count of samples incorrectly classified as positive and negative, respectively (Alkan Akinci et al. 2024).

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{7}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

In addition, the study calculated the ROC–AUC, a commonly used method to evaluate the overall model performance of LSMs, where a larger ROC–AUC indicates that the model performs better (Yu et al. 2021). This curve shows the relation between a true positive rate (sensitivity) and a false positive rate (1-specificity), which both can be computed by using Eqs. 11, 12.

$$\text{Sensitivity} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \tag{11}$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True negative} + \text{False positive}} \tag{12}$$

Finally, these susceptibility maps were evaluated by comparing the spatial distribution of the recorded landslide events and safe points against the identified high-risk and low-risk zones across each generated map.

4 Results

4.1 Multicollinearity analysis of the conditional factors

The result of multicollinearity analysis for the conditional factors is presented in Table 2. The values of VIF are between 1.12 and 4.28, while the TOL values range from 0.23 to 0.89. Based on the results, these 14 factors are suitable for application in landslide susceptibility analysis.

4.2 Landslide susceptibility maps

In this study, three LSMs were generated by using LR, RF, and SVM models. These maps were classified into five risk categories (very low, low, medium, high, and very high) by using the natural breaks (Jenks) method in ArcGIS software (Fig. 4).

These resultant maps revealed significant variations in risk levels across the study area depending on the model used (Fig. 4d). For example, the very high susceptibility and very low susceptibility zones covered 16.72% and 22.69% of the study area, respectively, in the LR model; 21.87% and 9.97% in the RF model; and 18.83% and 6.92% in the SVM model.

Moreover, distinctions among the three models are obvious not only in the proportions of categorized areas but also in the location of risk zones within the area. In RF and SVM models, risky zones are mainly concentrated in the northwest, where the steep slopes are prevalent, and also in the south near the roads (Fig. 4b, c). In contrast, the LR model showed high-risk areas extending into the eastern and central regions of the study area (Fig. 4a).

Table 2 Multicollinearity analysis results

Conditional factor	VIF	TOL	Conditional factor	VIF	TOL
Slope	1.51	0.66	Distance to streams	1.28	0.78
Aspect	1.22	0.82	TWI	2.11	0.47
Elevation	4.28	0.23	LULC	1.12	0.89
Profile curvature	1.39	0.72	Distance to roads	1.66	0.60
Plan curvature	2.05	0.49	rainfall	4.06	0.25
Lithology	1.51	0.66	Magnitude of earthquake	1.50	0.67
Distance to faults	1.44	0.69	Depth of earthquake	1.39	0.72

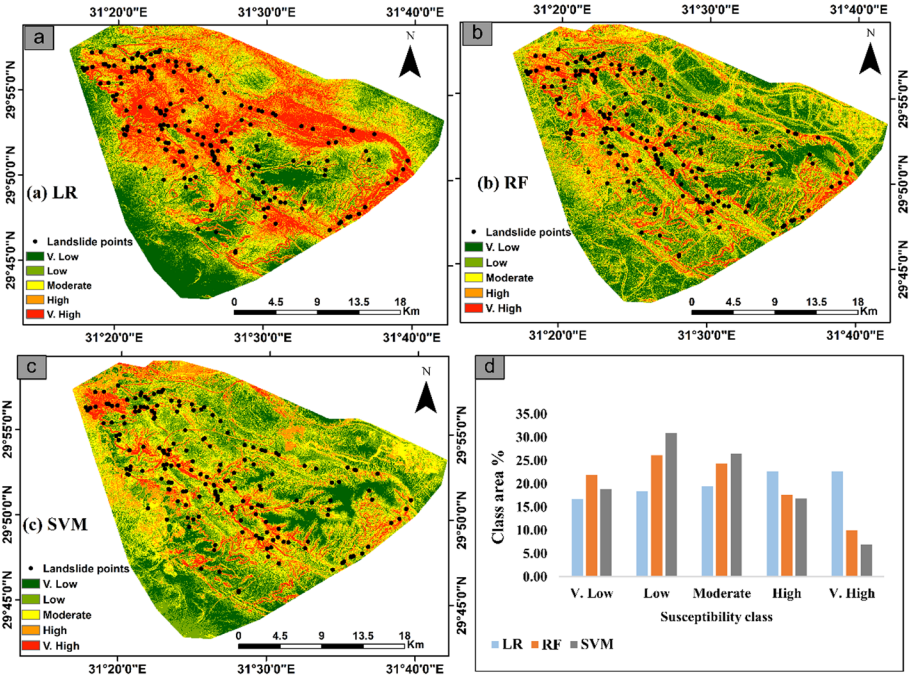

Fig. 4 **a**, **b**, and **c** LSMs generated using LR, RF, and SVM, respectively; **d** bar chart showing the percentage of susceptibility zones for LR, RF, and SVM models

4.3 The contribution of the conditional factors

Since the models have different methodologies, they showed the relationship between predictor variables and landslide occurrence in different ways (Fig. 5). Therefore, we analyzed this relation to understand the contribution of each factor. In the LR model, we assessed the contribution of each predictor using the absolute values of the coefficients (Fig. 5a). The factors with the highest absolute coefficients, indicating the strongest influence on landslide occurrence, were slope (4.53), magnitude of earthquakes (2.56), TWI (2.53), profile curvature (2.48), and distance to roads (2.47). Other significant factors included plan curvature (2.22), elevation (2.07), and distance to faults (1.51), all of which had notable contributions. On contrary, factors such as rainfall (1.05), land use/land cover (LULC) (0.51), depth to earthquakes (0.51), aspect (0.18), lithology (0.18), and distance to streams (0.08) had smaller contributions. The ranking of these coefficients helps us understand the relative importance of each factor in predicting landslide susceptibility.

For the RF model, the contribution of conditional factors was assessed using the values of MDG (Fig. 5b). According to these values, the slope is the most significant factor with predictive capabilities in landslide susceptibility, followed by TWI, profile curvature, aspect, and distance to road, which can be considered moderately important factors. The other factors are considered of low importance.

In the SVM model, the degree of influence of each predictor was calculated (Fig. 5c). The slope was found to be the most influential factor, followed by TWI, distance to faults, profile curvature, and distance to roads, which showed also significant effects on landslide

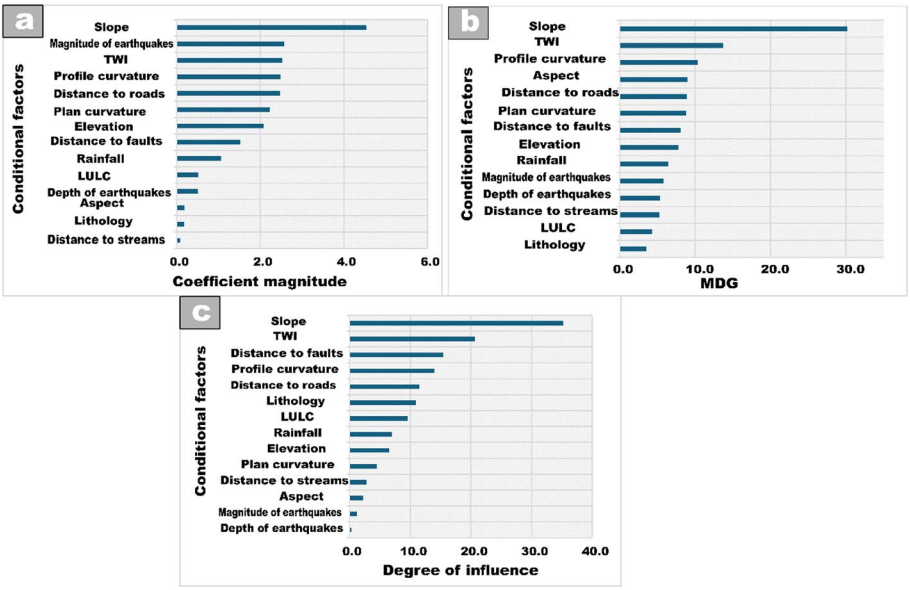

Fig. 5 Contribution of the conditional factors; **a** bar chart showing the values of coefficient magnitude for LR; **b** bar chart showing MDG of the RF model; **c** bar chart showing degrees of influence of predictors in the SVM model

susceptibility. In contrast, the depth and magnitude of earthquakes exhibited minimal influence, suggesting their limited role in shaping the decision-making process of the model.

4.4 Interpretation of models with SHAP

The impacts of conditional factors on model outputs are also evaluated by using SHAP values. SHAP values provided not only the ranking of the importance of each factor but also showed if the factor has a positive or negative impact on the final model (Chen and Fan 2024). In SHAP plots (Fig. 6), the importance of factors and their contribution to landslides were ranked vertically by significance, while the horizontal distribution of SHAP values showed the range and direction of their effects. Also, the color gradient (from red for high values to blue for low) highlights how the factor value contributes to the model output. Analyzing the predictions of the three models using SHAP values reveals that the slope is consistently the most influential factor across the different models, despite the differing importance of other conditional factors.

In the LR model (Fig. 6a), slope and distance to roads appeared as the two most influential factors in shaping the predictions, showing their strong relation to the final prediction. Elevation, TWI, and earthquake magnitude also played important roles, but their effects were not as pronounced. It was shown that topographic factors, represented by slope and profile curvature, along with moisture-related factors, indicated by TWI, were crucial in driving the RF's outcomes (Fig. 6b). Whereas distance to roads and aspect showed less importance. Although topographic factors, particularly slope, and elevation, remained dominant in SVM (Fig. 6c), highlighting the importance of terrain characteristics in shaping the model's predictions, LULC, distance to

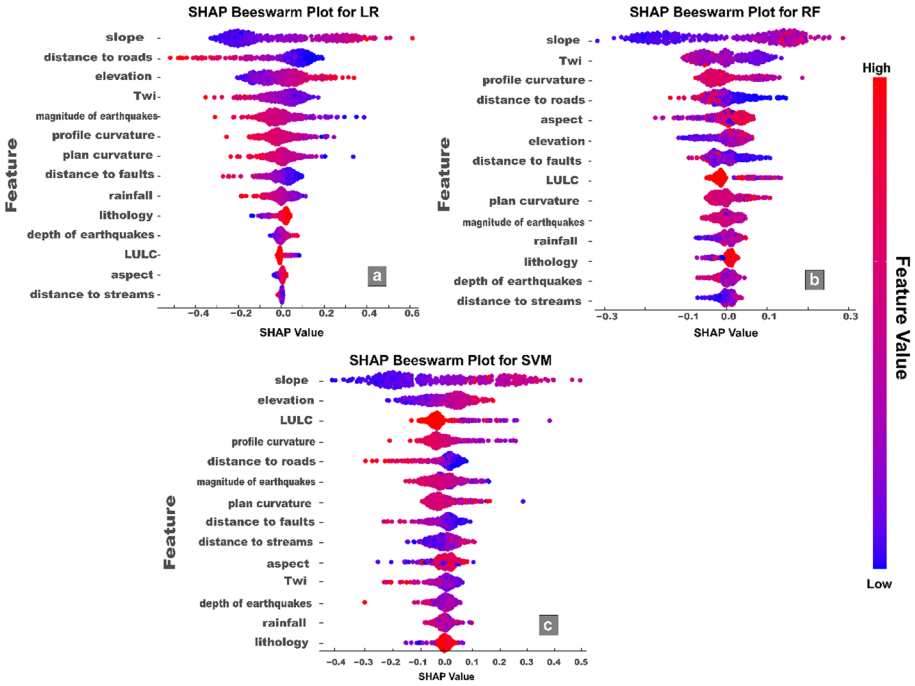

Fig. 6 SHAP beeswarm plots; **a** LR, **b** RF, and **c** SVM

roads, and earthquake magnitude played crucial roles, indicating that human activity and seismic factors also significantly influenced the model's behavior.

Moreover, the SHAP plots showed that lithology, depth of earthquakes, and distance to streams displayed the smallest impact on the three model predictions, as indicated by their lower vertical ranking and narrower distribution of SHAP values. These factors make a very small contribution to the model predictions, with their SHAP values showing limited variation in both direction and magnitude.

4.5 Results of the validation of models

All the examined models yielded ROC–AUC values exceeding 85% (Fig. 7). According to these values, these models provide acceptable outcomes within the study area (Zhang et al. 2023). However, there was considerable variation in ROC–AUC values across all models, with ROC–AUCs of 0.95.4 in RF, 0.90 in SVM, and 0.882 in LR. These models experienced further validation through the calculation of evaluation metrics to ensure a more accurate assessment. The results of these indices are presented in Table 3. The values showed that RF continuously provided higher performance accuracy. For ACC, RF is greater than SVM and LR by 0.07 and 0.13, respectively. Additionally, RF outperforms SVM and LR in terms of precision by 0.16 and 0.29, respectively, and outperforms them in terms of F1 score by 0.09 and 0.17, respectively. Nonetheless, all three models obtained a precision of 0.89.

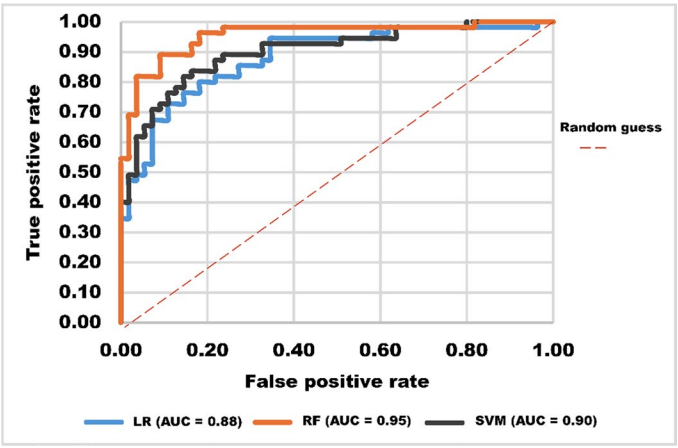

Fig.7 ROC curves and AUC values of the models

Table 3 Performance metrics of the models

Metrics	LR	RF	SVM
ACC	0.77	0.90	0.83
Recall	0.89	0.89	0.89
Precision	0.61	0.90	0.74
F1 score	0.73	0.90	0.81

5 Discussion

The principal role of the LSM is to identify predictable risky zones to offer the decision-maker insights for disaster mitigation (Zhou et al. 2021). Although many studies compared different approaches to generating landslide hazard maps for the same area, until now there has been no standard strategy for building reliable LSM that can be used globally (Sahin et al. 2020). During this study, the LSM generated by different ML approaches (LR, RF, and SVM), and the resultant maps revealed variations in risk zones distribution among the models, particularly in LR. LR identified 45% of the study area as very high and high-risk areas, which was nearly double the coverage seen in RF and SVM, where these zones covered approximately 25% (Fig. 4d). Conversely, RF and SVM classified about 49% of the area as very low and low-risk, significantly higher than 35% of LR. While the moderate-risk zones exhibited slight variation among the models, ranging between 19 and 25%. Furthermore, the spatial distribution of risk zones varied as well. Moreover, RF and SVM identified the risky areas in the northwest, characterized by steep slopes, and in the south, near the roads. In contrast, LR extended risk zones into the eastern and central parts of the study area, emphasizing its different susceptibility predictions. These variations in the LSMs can cause significant confusion, and selecting the wrong model may increase the risk of landslides in the study area rather than mitigate it. Investigation of how each model processes the conditional factors and assessing the reliability of these susceptibility maps are a crucial step.

5.1 Evaluation of the significance of landslide factors

There is no fixed guide for selecting these factors or determining their significance because of the complexity of landslide occurrences and the relative importance of each landslide factor varies depending on the features of the study area, the analysis model used, and the other variables included in the study (Dai et al. 2023). Therefore, studying the influence of each factor on landslide events is vital to our study. ML algorithms use different mechanisms to make decisions. Consequently, their contributions towards model predictions vary even when the same factors are utilized. ML models offer some insights into the relative importance of the factors through feature importance functions; however, this approach does not enough to explain how the models make predictions, hence their characterization as "black box" models. To overcome this limitation, this study used the SHAP value plots for the interpretability of ML-based LSMs (Lu et al., 2023; Pradhan et al., 2023).

The analysis of SHAP plots across the three models (Fig. 6) revealed that topographic factors, particularly slope, elevation, and profile curvature, have a high impact on landslide occurrence predictions. This result aligns with the well-established understanding that topographic factors play a dominant role in driving landslide occurrence due to their influence on slope morphology and natural erosion process (Youssef and Pourghasemi 2021; Agboola et al. 2024). Additionally, the hydrological factor (TWI) and human activities (distance to roads) also demonstrated significant impacts on landslide hazards by affecting slope resistance.

Notably, the relative importance of factors varied across models, except for the slope factor, which consistently ranked as the most influential. This variation highlights the different ways in which each model captured and utilized information from contributing factors. The SHAP values in the LR and SVM showed straightforward patterns, with clear clusters of red and blue points representing positive and negative influences (Fig. 6a, c). These patterns indicate that these models captured the relationships between factors and landslide susceptibility in a linear way. In contrast, the RF model (Fig. 6b) displayed more complex SHAP patterns. The color gradients were less distinct, suggesting that RF captured nonlinear relationships between the factors and landslide susceptibility. This complexity likely explains the higher predictive accuracy of RF, as it was able to account for complex interactions among the variables (Chen and Fan 2024).

5.2 Validation and rationality of susceptibility maps

In this study, the prediction performance of the resultant models was examined firstly by using ROC–AUC, which is use most frequently in LSM studies (Akinci and Zeybek 2021; Wang et al. 2023, 2024). All the models achieved ROC–AUC values exceeding 88%, with ROC–AUCs of 0.95.4 in RF, 0.90 in SVM, and 0.882 in LR (Fig. 7), indicating that these models provide acceptable outcomes align with the previous research projects (Chowdhury et al. 2024; Usta et al. 2024; Lee and Lee 2024). However, while these high ROC–AUC values indicate good overall model performance, they do not necessarily mean that each model is equally reliable for decision-making in LSM, which was cleared by the difference in the risk locations from one map to another for the same area. Thus, in this work, there are two additional methods to validate the generated LSMs. The initial evaluation involved computing the F1 score, recall, ACC, and precision. These metrics are used frequently in recent studies of LSM (Liao et al. 2022; Chen

and Fan 2024; Lee and Lee 2024). The results showed that the RF model consistently performed better than the others whether in predicting the landslide or the safe zone, with values of 0.90 for ACC, 0.89 for recall, 0.90 for precision, and 0.90 for the F1 score (Table 3). These higher performance metrics of the RF model owing to its distinct mechanism to learn complex relationships between the independent factors and the occurrence of landslides, which is compatible with research findings (Khalil et al. 2023; Lee and Lee 2024).

Moreover, the second significant step in the validation stage is investigating the rationality of these maps by comparing the number of safe places in areas of low and very low susceptibility for each vulnerability class and the number of landslides in areas of high and very high susceptibility (Usta et al. 2024; Alkan Akinci et al. 2024). This analysis showed a very effective way of understanding how well each model performs in practical terms. The distribution of landslides across susceptibility classes showed a consistent trend among all models, with an increase in landslide occurrences in high-susceptibility zones (Fig. 8a). The RF model correctly classified 145 landslide locations in the very high risk zone, followed closely by LR with 143 landslides, and SVM with 96 landslides. Additionally, the RF model had only two landslides misclassified in the very low and low susceptibility zones, compared to five misclassified landslides in both the LR and SVM models.

When considering safe locations, the SVM model performed best in correctly classifying them within the very low and low susceptibility categories and had the fewest safe points incorrectly placed in the high and very high-risk zones (Fig. 8b). In contrast, the LR showed the greatest number of misclassifications, with 64 safe locations incorrectly classified in the very high and high-risk areas.

These results highlight an important point about model evaluation: a high ROC–AUC score, while often seen as a sign of good performance, does not automatically mean a model is reliable. This is particularly evident in the case of the LR model, which showed a high ROC–AUC score, and exhibited the highest number of misclassifications when predicting both landslide-prone and safe areas. In contrast, the RF model showed superior performance across multiple evaluation metrics and produced a rational landslide susceptibility distribution, establishing it as the most reliable model for LSM in this study. Finally, these results indicate that choosing the right model for landslide susceptibility mapping comes down from the combination of evaluation metrics and assessment of the rationality of the resultant maps.

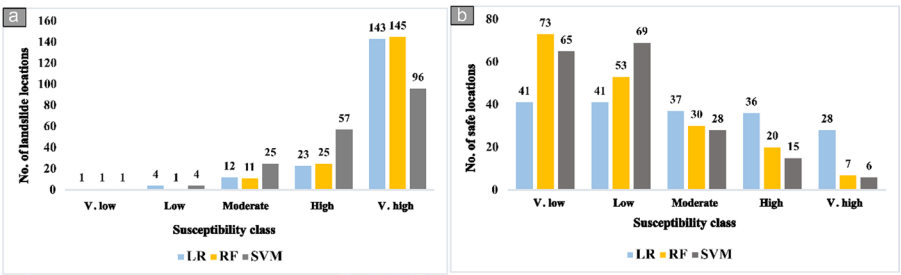

Fig. 8 Rationality of the models; **a** comparison with landslide locations, **b** comparison with safe locations

6 Conclusion

This study compared the accuracy and reliability of three different machine learning methods (LR, RF, and SVM) used in generating LSM maps for an area located southeast of Cairo and east of the New Administrative Capital, Egypt. This research marks the first application of these three models to produce LSM for this region, as previous studies primarily relied on qualitative and subjective methods. 183 landslide locations and an equal number of safe locations were used as a landslide record database for this region that was collected by fieldwork and Google Earth images. Moreover, this study examined the significance of the fourteen factors that could control landslide occurrence by using the SHAP value analysis. The results showed that topographic factors, practically the slope, have the greatest impact on the occurrence of a landslide. Also, human activities and hydrological factors showed an important effect compared to other factors; these can accelerate the rate of landslide. On the contrary, trigger factors such as rainfall and earthquakes exhibited the lowest influence on landslide occurrences.

The findings of this work showed that the RF model exceeded the others, achieving the highest ROC–AUC score of 0.95 and the highest values in ACC, recall, precision, and F1 scores. It demonstrated a consistent ability to correctly classify landslide locations in high-susceptibility zones and minimize misclassification in low- susceptibility areas, making it the most reliable choice for practical LSM. The SVM model, while proficient in classifying safe locations and maintaining a high AUC of 0.90, fell short in accurately predicting landslide-prone areas compared to RF. The LR model, despite achieving a respectable AUC of 0.88, exhibited lower performance metrics and higher misclassification rates for both landslide and safe locations. These outcomes highlight the importance of not relying only on ROC–AUC results when assessing model performance for decision-making purposes. Effective LSM requires models that not only exhibit high predictive accuracy but also demonstrate practical reliability in distinguishing between high-risk and safe zones. The RF model's balanced performance across multiple evaluation metrics highlights its suitability as the most dependable tool for landslide risk management. This study promotes a complete approach to model evaluation that integrates both statistical performance and realistic applicability to better ensure mitigation strategies for vulnerable communities.

Acknowledgements The authors would like to express their sincere gratitude to the anonymized reviewers and the journal editor for their valuable time, insightful comments, and constructive feedback, which significantly improved the quality of this article. Additionally, we thank the joint executive program between Hungary and Egypt (Stipendium Hungaricum scholarship) for funding Mohamed M. Abdelkader.

Funding Open access funding provided by University of Debrecen. Open access funding is provided by the University of Debrecen. Supported by the University of Debrecen Program for Scientific Publication. This research received no external funding.

Declarations

Conflict of interest The authors declare that any known competing financial interests or personal relationships could have influenced none of the work reported in this study.

References

Agboola G, Beni LH, Elbayoumi T, Thompson G (2024) Optimizing landslide susceptibility mapping using machine learning and geospatial techniques. Eco Inform 81:102583. https://doi.org/10.1016/J.ECOINF.2024.102583

Akinci H (2022) Assessment of rainfall-induced landslide susceptibility in Artvin, Turkey using machine learning techniques. J Afr Earth Sci 191:104535. https://doi.org/10.1016/J.JAFREARSCI.2022.104535

Akinci H, Zeybek M (2021) Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey. Nat Hazards 108:1515–1543. https://doi.org/10.1007/S11069-021-04743-4

Alkan Akinci H, Akinci H, Zeybek M (2024) Comparison of diverse machine learning algorithms for forest fire susceptibility mapping in Antalya, Türkiye. Adv Space Res 74:647–667. https://doi.org/10.1016/J.ASR.2024.04.018

Bai S, Lu P, Thiebes B (2020) Comparing characteristics of rainfall- and earthquake-triggered landslides in the Upper Minjiang catchment. China Eng Geol 268:105518. https://doi.org/10.1016/J.ENGGEO.2020.105518

Bednarik M, Yilmaz I, Kralovičová L (2024) Deterministic approach to assess landslide susceptibility and landslide activity in the central-western region of Slovakia. Bull Eng Geol Env 83:1–14. https://doi.org/10.1007/S10064-024-03795-7

Bopche L, Rege PP (2022) Landslide susceptibility mapping: an integrated approach using geographic information value, remote sensing, and weight of evidence method. Geotech Geol Eng 40:2935–2947. https://doi.org/10.1007/S10706-022-02070-4

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Chen L, Guo Z, Yin K et al (2019) The influence of land use and land cover change on landslide susceptibility: a case study in Zhushan Town, Xuan'en County (Hubei, China). Nat Hazard 19:2207–2228. https://doi.org/10.5194/NHESS-19-2207-2019

Chen, C., Fan, L. (2024). Interpretability of statistical, machine learning, and deep learning models for landslide susceptibility mapping in Three Gorges Reservoir Area. ArXiv. https://arxiv.org/abs/2405.11762

Chicas SD, Li H, Mizoue N et al (2024) Landslide susceptibility mapping core-base factors and models' performance variability: a systematic review. Nat Hazards 120:2573–12593. https://doi.org/10.1007/s11069-024-06697-9

Chowdhuri I, Pal SC, Chakrabortty R et al (2021) Torrential rainfall-induced landslide susceptibility assessment using machine learning and statistical methods of eastern Himalaya. Nat Hazards 107:697–722. https://doi.org/10.1007/S11069-021-04601-3

Chowdhury MS, Rahaman MN, Sheikh MS et al (2024) GIS-based landslide susceptibility mapping using logistic regression, random forest and decision and regression tree models in Chattogram District Bangladesh. Heliyon 10:e23424. https://doi.org/10.1016/J.HELIYON.2023.E23424

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. https://doi.org/10.1007/BF00994018

Dai X, Zhu Y, Sun K et al (2023) Examining the spatially varying relationships between landslide susceptibility and conditioning factors using a geographical random forest approach: a case study in Liangshan. China Remote Sens 15:1513. https://doi.org/10.3390/RS15061513

Effat HA, Hegazy MN (2014) Mapping landslide susceptibility using satellite data and spatial multicriteria evaluation: the case of Helwan District, Cairo. Appl Geomat 6:215–228. https://doi.org/10.1007/s12518-014-0137-9

Fan X, Scaringi G, Korup O et al (2019) Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. Rev Geophys 57:421–503. https://doi.org/10.1029/2018RG000626

Farooq S, Akram MS (2021) Landslide susceptibility mapping using information value method in Jhelum Valley of the Himalayas. Arab J Geosci 14:1–16. https://doi.org/10.1007/S12517-021-07147-7

Gaidzik K, Ramírez-Herrera MT (2021) The importance of input data on landslide susceptibility mapping. Sci Rep 11:1–14. https://doi.org/10.1038/s41598-021-98830-y

Guo Z, Tian B, Li G et al (2023) Landslide susceptibility mapping in the Loess Plateau of northwest China using three data-driven techniques—a case study from the middle Yellow River catchment. Front Earth Sci 10:1033085. https://doi.org/10.3389/FEART.2022.1033085

Harris I, Osborn TJ, Jones P, Lister D (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci Data 7:1–18. https://doi.org/10.1038/s41597-020-0453-3

Hong H, Pradhan B, Bui DT et al (2017) Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: a case study at Suichuan area (China). Geomat Nat Haz Risk 8:544–569. https://doi.org/10.1080/19475705.2016.1250112

International seismological centre (2023) ISC-GEM Earthquake catalogue, version 10.0. https://doi.org/10.31905/d808b825

Kamran KV, Feizizadeh B, Khorrami B, Ebadi Y (2021) A comparative approach of support vector machine kernel functions for GIS-based landslide susceptibility mapping. Appl Geomat 13:837–851. https://doi.org/10.1007/S12518-021-00393-0

Khalil YM, Al-Masnay YA, Al-Areeq NM et al (2023) Estimating landslide hazard distribution based on machine learning and bivariate statistics in Utmah Region, Yemen. Nat Hazards 120:2869–2907. https://doi.org/10.1007/S11069-023-06310-5

Kuhn M (2008) Building predictive models in R Using the caret package. J Stat Softw. https://doi.org/10.18637/jss.v028.i05

Lee SM, Lee SJ (2024) Landslide susceptibility assessment of South Korea using stacking ensemble machine learning. Geoenviron Disasters 11:1–17. https://doi.org/10.1186/S40677-024-00271-Y

Lee DH, Kim YT, Lee SR (2020) Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various non-linear activation functions. Remote Sens 12:1194. https://doi.org/10.3390/RS12071194

Liao M, Wen H, Yang L (2022) Identifying the essential conditioning factors of landslide susceptibility models under different grid resolutions using hybrid machine learning: a case of Wushan and Wuxi counties. China Catena 217:106428. https://doi.org/10.1016/J.CATENA.2022.106428

Liu R, Li L, Pirasteh S et al (2021) The performance quality of LR, SVM, and RF for earthquake-induced landslides susceptibility mapping incorporating remote sensing imagery. Arab J Geosci 14:1–15. https://doi.org/10.1007/S12517-021-06573-X

Liu S, Wang L, Zhang W et al (2024) Physics-informed optimization for a data-driven approach in landslide susceptibility evaluation. J Rock Mech Geotech Eng 16:3192–3205. https://doi.org/10.1016/J.JRMGE.2023.11.039

Lu J, Ren C, Yue W et al (2023) Investigation of landslide susceptibility decision mechanisms in different ensemble-based machine learning models with various types of factor data. Sustain 15:13563. https://doi.org/10.3390/su151813563

Meyer D, Dimitriadou E, Hornik K et al (2024) Misc functions of the department of statistics, probability theory group (formerly: e1071), TU Wien [R package e1071 version 1.7–16]. CRAN: contributed packages. https://doi.org/10.32614/CRAN.PACKAGE.E1071

Moustafa AR, Abd-Allah AM (1991) Structural setting of the central part of the Cairo-Suez District. MERC. Ain Shams Univ Earth Sci Ser 5:133–145

Moustafa AR, Yehia MA, Abdel Tawab S (1985) Structural setting of the area east of Cairo, Maadi, and Helwan. MERC. Ain Shams Univ Earth Sci Ser 5:40–64

Peruccacci S, Brunetti MT, Gariano SL et al (2017) Rainfall thresholds for possible landslide occurrence in Italy. Geomorphology 290:39–57. https://doi.org/10.1016/J.GEOMORPH.2017.03.031

Phoon KK, Zhang W (2023) Future of machine learning in geotechnics. Georisk Assessment Manag Risk Eng Syst Geohazards 17(1):7–22. https://doi.org/10.1080/17499518.2022.2087884

Pradhan AMS, Kim YT (2020) Rainfall-induced shallow landslide susceptibility mapping at two adjacent catchments using advanced machine learning algorithms. ISPRS Int J Geo Inf 9:569. https://doi.org/10.3390/IJGI9100569

Pradhan B, Dikshit A, Lee S, Kim H (2023) An explainable AI (XAI) model for landslide susceptibility modeling. Appl Soft Comput 142:110324. https://doi.org/10.1016/j.asoc.2023.110324

Rai DK, Xiong D, Zhao W et al (2022) An investigation of landslide susceptibility using logistic regression and statistical index methods in Dailekh district. Nepal Chin Geogr Sci 32:834–851. https://doi.org/10.1007/S11769-022-1304-2

Sahin EK, Colkesen I, Acmali SS et al (2020) Developing comprehensive geocomputation tools for landslide susceptibility mapping: LSM tool pack. Comput Geosci 144:104592. https://doi.org/10.1016/J.CAGEO.2020.104592

Shang H, Su L, Chen W et al (2023) Spatial prediction of landslide susceptibility using logistic regression (LR), functional trees (FTs), and random subspace functional trees (RSFTs) for Pengyang County. China Remote Sens 15:4952. https://doi.org/10.3390/rs15204952

Singh P, Sharma A, Sur U, Rai PK (2020) Comparative landslide susceptibility assessment using statistical information value and index of entropy model in Bhanupali-Beri region, Himachal Pradesh, India. Environ Dev Sustain 23(4):5233–5250. https://doi.org/10.1007/s10668-020-00811-0

Sun X, Chen J, Bao Y et al (2018) Landslide susceptibility mapping using logistic regression analysis along the Jinsha River and its tributaries close to Derong and Deqin County, Southwestern China. ISPRS Int J Geo Inf 7:438. https://doi.org/10.3390/IJGI7110438

Tanyaş H, Görüm T, Kirschbaum D, Lombardo L (2022) Could road constructions be more hazardous than an earthquake in terms of mass movement? Nat Hazards 112:639–663. https://doi.org/10.1007/S11069-021-05199-2

Tyagi A, Kamal Tiwari R, James N (2022) A review on spatial, temporal and magnitude prediction of landslide hazard. J Asian Earth Sci X 7:100099. https://doi.org/10.1016/J.JAESX.2022.100099

Usta Z, Akıncı H, Akın AT (2024) Comparison of tree-based ensemble learning algorithms for landslide susceptibility mapping in Murgul (Artvin), Turkey. Earth Sci Inf 17:1459–1481. https://doi.org/10.1007/S12145-024-01259-W

Wang Y, Sun D, Wen H et al (2020) Comparison of random forest model and frequency ratio model for landslide susceptibility mapping (LSM) in Yunyang County (Chongqing, China). Int J Environ Res Public Health 17:4206. https://doi.org/10.3390/ijerph17124206

Wang Y, Wang L, Liu S et al (2023) A comparative study of regional landslide susceptibility mapping with multiple machine learning models. Geol J. https://doi.org/10.1002/GJ.4902

Wang YH, Wang LQ, Zhang WG et al (2024) A physics-informed machine learning solution for landslide susceptibility mapping based on three-dimensional slope stability evaluation. J Cent South Univ. https://doi.org/10.1007/S11771-024-5687-3

WorldPop, Bondarenko M (2020) Individual countries 1km population density (2000–2020). https://doi.org/10.5258/SOTON/WP00674

Yao J, Yao X, Zhao Z, Liu X (2023) Performance comparison of landslide susceptibility mapping under multiple machine-learning based models considering InSAR deformation: a case study of the upper Jinsha River. Geomat Nat Hazards Risk. https://doi.org/10.1080/19475705.2023.2212833

Yavuz Ozalp A, Akinci H, Zeybek M (2023) Comparative analysis of tree-based ensemble learning algorithms for landslide susceptibility mapping: a case study in Rize. Turkey Water 15:2661. https://doi.org/10.3390/W15142661

Youssef AM, Pourghasemi HR (2021) Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. Geosci Front 12:639–655. https://doi.org/10.1016/J.GSF.2020.05.010

Youssef AM, El-Haddad BA, Skilodimou HD et al (2024) Landslide susceptibility, ensemble machine learning, and accuracy methods in the southern Sinai Peninsula, Egypt: assessment and Mapping. Nat Hazards 120:14227–14258. https://doi.org/10.1007/s11069-024-06769-w

Yu X, Zhang K, Song Y et al (2021) Study on landslide susceptibility mapping based on rock–soil characteristic factors. Sci Rep 11:1–27. https://doi.org/10.1038/s41598-021-94936-5

Zhang W, Zhang R, Wu C et al (2020) State-of-the-art review of soft computing applications in underground excavations. Geosci Front 11:1095–1106. https://doi.org/10.1016/J.GSF.2019.12.003

Zhang W, Li H, Li Y et al (2021) Application of deep learning algorithms in geotechnical engineering: a short critical review. Artif Intell Rev 54:5633–5673. https://doi.org/10.1007/S10462-021-09967-1

Zhang W, He Y, Wang L et al (2023) Landslide susceptibility mapping using random forest and extreme gradient boosting: A case study of Fengjie, Chongqing. Geol J 58:2372–2387. https://doi.org/10.1002/GJ.4683

Zhao P, Masoumi Z, Kalantari M et al (2022) A GIS-based landslide susceptibility mapping and variable importance analysis using artificial intelligent training-based methods. Remote Sens 14:211. https://doi.org/10.3390/RS14010211

Zhou X, Wu W, Qin Y, Fu X (2021) Geoinformation-based landslide susceptibility mapping in subtropical area. Sci Rep 11:1–16. https://doi.org/10.1038/s41598-021-03743-5