中華水土保持學報, 57(1):45-56(2026)
Journal of Chinese Soil and Water Conservation, 57 (1): 45-56 (2026)
DOI: 10.29417/JCSWC.202603_57(1).0004

Comparison of data mining models to assess landslide susceptibility in Karganeh Watershed, Lorestan Province, Iran

Ebrahim Karimi Sangchini [1]* Seyed Hossein Arami [2] Ali Dastranj [3]

ABSTRACT Landslide is the movement of materials on the slope (containing natural rocks, soil, artificial accumulations or a mixture of them) that are moved downward by the force of gravity. Therefore, the preparation of landslide susceptibility maps is very vigorous so that mitigate damaging effects. The aim of the research is to spatially landslide susceptibility model via three methods of random forest (RF), maximum entropy (ME) and support vector machine (SVM) algorithms, and compare the efficiency of these models in landslide susceptibility mapping in Karganeh Watershed, Lorestan Province, Iran. Landslide inventory map was primed via aerial photographs interpretation and general field surveys. In this research, 16 vital landslide causal factors were reflected to search their spatial connotation to landslide, established happening native geomorphological and human effects. Then, landslide susceptibility maps were built via tree models in geographic information system (GIS). The ROC curve via AUC index was usage to evaluate too compare landslide susceptibility models. Results presented that SVM model provided considerably higher prediction accuracy of landslide susceptibility map in the Karganeh Watershed by ROC equal to 0.913. The subsequent landslide susceptibility maps be able to be appropriate in fitting watershed management practices in this watershed.

Key Words: Landslide susceptibility, Forest, Entropy, Vector Machine, Karganeh Watershed, Iran

Introduction

Numerous forms of natural dangers and related tragedies containing earthquakes, volcanic eruptions, tsunamis, cloud burst, floods, soil erosion etc. happening the world and amongst such troubling natural hazards landslides are the awful types of greatest recurrent occurrences all everywhere the world [1-4]. Landslide is reason many monetary costs and lives yearly [5-6]. Every year, landslides have affected huge damages of life and stuff, concluded the damages of forests, fruitful cultivated land, habitation area, and network communication in addition to tourist adverts. Additionally, alteration of the earth surface is also responsible for devastating landslides. Consequently, landslides are accountable for huge flooding and its hill areas, tsunami in the seaside zones; river forms alterations sideways through geomorphic and topographical modifications [7-9]. Consequently, landslide susceptibility mapping be able to be one of the key phases in diminishing these costs. These maps are also vital tools for engineers, geoscientists, planners, and managers to choice appropriate sites for agriculture, construction, and other development actions [10-12]. The analysis of landslides has drawn universal consideration mostly caused by cumulative alertness of the socio-economic consequence of landslides, in addition to, the aggregate burden of development on the highland situation [13]. Iran has confronted numerous categories of natural threats and disasters, for example severe soil erosion concluded gully expansion, vulgar floods, and disturbing landslides. So, because of the numerous occurrences of landslides and huge financial damages have develop national disasters of Iran. The landslide event in Iran has caused about 500 billion financial damages. [14-15]. Essentially, being of the sole natural structures for example physiographic, environmental, climatic condition along by anthropogenic actions and its rising demand on natural resources are very susceptible to landslides action in north part of mountainous areas in Iran [16-18]. Landslide susceptibility assessment is a vital procedure for managing of natural tragedies. It is moreover an important phase for integrated catchment management, hazard mitigation, natural and urban design in government strategies universal. There is no solitary method to recognize and prepare a zoning map to measure the susceptibility caused by the wide series of happenings of landslides [19-22]. By applying logical methods, a set of accurate tools is providing for the preparation and best use of the landslide-zoning map, as well as the use of landslide forecast models, which decreases the problem of landslide susceptibility identification and zoning [12, 15, 18, 23-30]. Consequently, for the perfect modeling of landslide susceptibility assessment forecast result has been mainly depends on the availability of respectable quality information, the working scale and the appropriate approach. In current years, data mining methods have been used due to their great accuracy and strong information processing facility for example: Artificial Neural Network [19, 31-32], Logistics Regression [11, 33-37].

The Forest method is one of machine learning and decision tree modeling that assessment the connection among landslide

〔1〕Assistant Professor, Soil Conservation and Watershed Management Research Department, Lorestan Agricultural and Natural Resources Research and Education Center(AREEO), Khorramabad, Iran.
〔2〕Assistant Professor, Forests and Rangelands Research Department, Khuzestan Agricultural and Natural Resources Research and Education Center, Agricultural Research Education and Extension Organization (AREEO), Ahvaz, Iran
〔3〕Assistant Professor, Soil Conservation and Watershed Management Department, khorasan Razavi Agricultural and Natural Resources Research and Education Center (AREEO), Mashhhad. Iran.
* Corresponding Author. E-mail: E.karimi64@gmail.com

happening and environmental factors via merging the results achieved from unlike trees [19]. Between the researchers who used Forest Algorithm in their studies for landslide susceptibility zoning, we can reference [30, 38-42]. The fallouts of these researchers' studies indicate that the random forest algorithm has suitable accuracy in landslide susceptibility zoning. The Entropy model is another machine learning method that has been usually used in new year [19, 43-49] used the entropy method in their studies to estimate landslide susceptibility. The fallouts of these studies highlight the efficiency of the Entropy model in landslide susceptibility zoning. Additional machine learning method is Support Vector Algorithm (SVM), which was used in this study for landslide susceptibility zoning. In their studies, [8, 19, 33, 50-54] conducted landslide susceptibility zoning via the support vector algorithm. The results of these researchers' studies established the usefulness of the support vector algorithm in landslide hazard zoning. Selecting the correct model that has great accuracy and truth can be extensively and efficiently used in the forecast and management of landslides in the text of land surveying package. In this study, an effort has been finished to use the greatest real factors on the happening of landslides, in addition to new machine learning approaches in the study area.

In this study, tree models of data mining including Forest, Entropy and vector machine algorithms were related for landslide susceptibility assessment in Karganeh Watershed. Consequently, so that perform our investigation aim now, it used 16 suitable landslides for this watershed area. Moreover, historic information of 95 landslide polygon were choice for promote development in our investigation work. Lastly, three models using in research have been validated via statistical examination of receiver working characteristics-area below curve (ROC-AUC), thus, landslide susceptibility assessment be able to assistance the planners for final management of environmental squalor and natural resources from delicate damages and eventually development of economic action of this watershed area. The planned methods use both the capability thoughts and ground fact at the like time. This could be taken as a brand-new methodology toward landslide zoning difficulties.

Materials and Methods

1. Study area

Karganeh Watershed is located between 33° 25′ 12″ to 33° 37′ 12″ latitude and 48° 23′ 59″ to 48° 44′ 24″ longitude, occupying about 294.2 sq km in the Lorestan Province, west of Iran (Fig. 1). This watershed is one of the main sub basins of Karkheh River. Altitude in the study area differs among 1,300 to 2,700 m. Based on Iranian meteorological organization report; the average annual precipitation in the watershed is 469 mm. This watershed is located in the middle of Zagros Mountains. In terms of stratigraphy, Karganeh Basin has Bakhtiari, Gachsaran, Asmari, Kashkan, Amiran, Cretaceous limestones and contemporary sediments. 41% of this watershed is covered via rangelands and residual lands are covered by orchard, forest, agricultural and rocky lands (about 59% from this watershed).

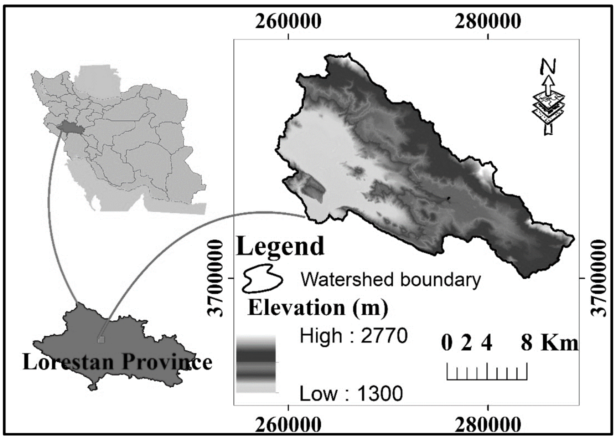


Fig.1 Location map of the study area

2. Landslide inventory map

One of the most important phases of landslide susceptibility assessment is to identify and prepare a distribution map of current landslides in the watershed. This map was prepared via assembly the information associated with landslides or via analyzing the data from remote sensing and GIS techniques. In current research, a landslide inventory map was ready by field surveys, local data, and aerial photographs interpretation. To make this map for susceptibility assessment, one point was located each 100 hectares. As numerous points by landslides, the similar number of areas without landslides were taken (Figure 2).

3. Selection of effective factors

From literature review and studying conditions of Karganeh Watershed, in total, 19 factors were chosen as effective factors on landslide occurrence. The lithology map was ready from the geological map (at a scale of 1:100,000) of this watershed (http://www.gsi.ir/). The land use map prepared by the General Department of Natural Resources of Lorestan was used and modified through field visits and Google Earth satellite images (https://lorestan.frw.ir/). The roads system maps were ready from the topographic map that was composed from National Geographic Organization of Iran (www.ngo-org.ir) at a scale of 1:1:50,000. Curving commonly shows the rate of slope/gradient variations in an exact direction (Lombardo et al. 2014). In which the curvature of the plan and profile represents the topographic features of a watershed and defines a different landscape through quantitative indicators [36, 55]. Therefore, curvature and its forms are key factors for modelling of landslide susceptibility modelling. Numerous key parameters i.e. distance to rivers, roads and faults are deliberated most vigorous landslide effective factors for assessment landslide hazard. Generally, in the great mountainous area, nearer the streams regions are additional susceptible to landslide events. This is caused by dampening volume of the shallow region from rivers water and variability features of the nearby areas. Nearby this, the constancy of the slope is also accountable for river erosion at the slope and therefore this factor has a benefit on landslide susceptibility modelling [56-57]. Hilly roads are extremely effects on the runoff form and activating the landslides events in a disturbing way. Nearby this, during the building of roads in the mountainous regions, it has been extremely exaggerated land constancy and as costs, land actions happen. The road map of the current research area has been composed from Mapping Organization of Iran at a scale of 1:25.000 [2]. Distance to fault is too extremely activating factor adversary the land movements and occupied it as a vigorous landslide effective factors. Fault are principally a wrecked and unbalanced rock surface region, which is importantly pretentious variability of the surface region, and disposed to land movements descending of the slope [58]. All of these linear factors have been dignified via using Euclidean distance buffering in GIS platform.

Altitude is extremely associated with landslides and one of the key factor used in assessment landslide hazard. It is generally influence on vegetation cover, topographic features, human action etc. on a region. Contingent on this effect factors altitude too effects constancy of the slope [59]. The design of precipitation-runoff is connected with precipitation. The strength and period of the precipitation effects the infiltration volume of the whole spaces and rises the gravity of the rock constituents owed to the soaking volume. It is fact that heavy precipitation through lengthy period reasons great frequency of landslide events. Slope is the most vigorous factor for landslide effective factors owing to the features of its constancy. Upper slope value is greatly prone to landslide events and on the contrary. TPI is the amount of slope situation of a topographic feature. Generally, TPI is used to separate among the central point elevation and the mediocre altitude about the central point [60]. Positive and negative TPI scores show the sites that are greater and lesser than the mediocre nearby regions correspondingly. However, zero TPI score shows that, the region is flat or an endless slope. The TPI has been calculated via the next equation [61].

$$TPI = M_0 - \sum_{n-1} M_n / n \tag{1}$$

Where, M_0 is the elevation of the middle point, M_n is the elevation of the grid and n is the total number of neighborhood region.

Generally, the erosional volume of a river is calculated through SPI. In the case of landslide susceptibility modelling, it is vital to distinguish the erosion capacity of the river, as it is extremely accountable for the erosion of the riverside and the influence of the slope by subsiding or departures. Therefore, SPI is taking into thought for landslide susceptibility modelling. If the SPI score is great then erosional competence of a slope surface is too great and on the contrary. The score of SPI was strongminded via next equation [62].

$$SPI = As * tan\beta \tag{2}$$

Wherever, AS shows upslope causal region and β shows the slope angle. Soil moisture of a region is recognized concluded TWI examination [63-64]. So, TWI is vital for landslide susceptibility modelling by way of it is mainly be contingent on moisture form of land surface area wherever landslides happening. In this research, TWI was calculated by next equation [33].

$$TWI = In\left(\frac{A_s}{tan\beta}\right) \tag{3}$$

Where, A_s characterize the area of catchment in m^2 and β characterize the gradient of the slope in radians. Generally, vector dispersal is the diversity of location of the earth's surface, which is used to characterize this factor via vectors vertical to the earth's surface [65]. Through changing the vector delivery on a level, the size of the subsequent vector that is got from the summation of vectors too changes, which stretches the idea of vector ruggedness. Vector ruggedness measurement is condensed as VRM. Soil texture too strongminded the form of landslide and its frequency by way of it is responses on precipitation, absorbency, rigidity, vegetation cover etc. In the current research four kinds of soil texture were documented i.e. clay, clay loam, loam and sandy loam [66]. The features of rock mass be contingent on numerous lithological elements that are hugely activating the landslides happenings [67]. The lithological map of the current research area was composed using the Geological Society of Iran (GSI).

In the next phase, Tolerance and VIF tests were used in SPSS software version 20 to determine the most effective factors. In the investigation of collinearity between the factors. Tolerance scores fewer than 0.1 and VIF scores bigger than 10 show collinearity between factors. The attendance of collinearity between the factors reduces the accuracy of the landslide susceptibility map [68]. Some works researches indicate that there are current numerous methods to assess the multi collinearity, between them greatest general are Pearson’s correlation coefficients, variance inflation factors (VIF), tolerances (TOL) etc. [69-70]. In the current research VIF and TOL methods was accustomed assessment multi collinearity between the different landslide effective factors. The multi collinearity difficulties happen while the threshold value of VIF is >five and TOL is <0.1. To compute the VIF and TOL of multi collinearity the next equation has been used [71].

$$TOL = 1 - R_j^2 \tag{4}$$

$$VIF = \frac{1}{TOL} \tag{5}$$

Where, R_j^2 characterize coefficient of multiple determination of j on the forecaster variables.

The results of the collinearity test using the Tolerance and VIF indices are shown in Table 1. After investigation, slope, slope direction, elevation classes, geology, distance from the river, distance from the road, distance from the fault, river power index (SPI), topographic moisture index (TWI) and slope length index (LS), topographic position index (TPI), topographic roughness index (TRI) and vector roughness measurement index (VRM), land use, distance from the village, and rainfall were selected as the most effective factors of landslide occurrence in the Karganeh Watershed.

Table 1 Collinearity test between effective factors in landslides

Factors	*Tolerance*	VIF	Factors	*Tolerance*	VIF
Slope	0.62	1.6	Topographic Position Index (TPI)	0.8	1.06
Aspect	0.58	1.8	Topographic Roughness Index (TRI)	0.73	1.53
Elevation	0.54	2.3	Vector Roughness Measure (VRM)	0.32	2.6
Geology	0.67	2.3	Soil texture	0.012	13.1
Distance from river	0.78	1.02	Distance from village	0.82	1.61
Distance from road	0.39	3.1	NDVI index	0.035	16.1
Distance from fault	0.26	3.9	Curvature Index	0.063	18.4
River Power Index (SPI)	0.44	2.7	Landuse	0.29	4.1
Topographic Wetness Index (TWI)	0.51	3.5	Precipitation	0.48	2.1
slope length (LS)	0.26	2.5			

4. Landslide susceptibility mapping via Forest (RF)

The algorithm of Forest was principally presented through Breiman. It is an ensemble-based machine learning classifier. This algorithm is constructed on non-parametric multivariate statistical technique. The mechanism of RF is constructed on the building of plentiful choice trees at the preliminary or training period finished the joint act of catching and arbitrarily choice of factors [72]. The catching methodology i.e. from the training dataset numerous trees are supernumerary all the method finished subset in a Forest algorithm. Consequently, it shows that the similar example became likelihoods numerous periods and others could perhaps not be became likelihoods at all [73]. The application of Forest has wide viewpoint in classification, regression lengthways by in unverified learning. It is vital of two factors to produce the Forest classifiers trees and these factors are Ntree i.e. number of choice trees and Mtry i.e. number of variables. Any caring of supposition does not require for founding the association between descriptive and reply variables in a RF algorithm. After determining the factors affecting the occurrence of landslides, a landslide susceptibility map was prepared using the Forest algorithm. Forest model is one of the machine learning methods for decision tree modeling. In this study, R software and Forest package were used in order to apply the Forest model in the assessment of landslide susceptibility [40]. Finally, the final landslide susceptibility map was zoned based on equal sizes in five susceptibility classes (very low, low, medium, high and very high) [11].

5. Landslide susceptibility mapping using Entropy

The Entropy method is one of the machine learning methods that has a high spatial prediction ability in various environmental fields [74]. The Entropy model first tries to find the relationships between independent and dependent variables in order to provide reality-based predictions. To use this model, to predict the behavior of a species or phenomenon, there is no need for points of absence of that species or phenomenon; Rather, it uses a series of influential factors (factors affecting the occurrence of landslides) as well as the presence points of that phenomenon (sliding points for modeling). In this research, the Entropy model in MaxEnt software was used in order to zone the sensitivity of landslide occurrence. In order to use this model, first the independent variables (factors affecting the occurrence of landslides) and the dependent variables (points of occurrence of landslides) were converted to the desired format and entered into the MaxEnt software environment. Based on the principle of Entropy, this model forms a network of communication between independent and dependent variables [47]. Finally, the final landslide susceptibility map was zoned based on equal sizes in five susceptibility classes (very low, low, medium, high and very high) [11].

6. Landslide susceptibility mapping using Vector Machine

This model is one of the supervised machine learning models used to classify and separate data. In other words, the Vector Machine model divides the data into distinct groups after analysis between independent and dependent variables. In the Vector Machine algorithm, each data sample is drawn as a point in the n-dimensional space on the data scatter diagram (n is the number of features of a data sample) and the value of each feature related to the data determines one of the components of the coordinates of the point on the diagram. The main idea of this algorithm is a binary classification using training points, which transforms the original input space into a space with higher dimensions to find a desirable hyperplane. The training points that are close to the desired plane are called support vectors. Once the decision level is obtained, it can be used to estimate new data [50]. In this study, ModEco software and Vector Machine algorithm have been used in order to procedure the Vector Machine model.

7. Evaluation of landslide susceptibility models

The efficiency of the landslide susceptibility models can be evaluated using ROC curve (operating characteristic). The ROC curve is a diagram in which the pixels' relation that is correctly forecast the happening or non-happening of landslides (True Positive) is plotted in contrast to the increase quantity that is the pixels' relation that is incorrectly forecast. As previously stated, the susceptibility model, calculates the revolution in probability in each pixel in a incessant fury of zero and one. By decisive a threshold (e.g. 0.5) the model's production can be rehabilitated to a separate scale of zero and one e.g. the pixels in which that the alteration probability is extra than their threshold, take 1 and pixels in which the change probability is fewer than their threshold takings 0 and the output is obtainable as a map. By comparing this with the landslide inventory, the pixels' ratio can be planned in ROC diagram. The ROC index equivalents to the region below the curve [50].

Results and Discussion

1. Landslide inventory map

Landslide inventory map indicated that there are 37 scattered landslides in the Karganeh Watershed. Exaggerated total area through landslide is 635 ha (2.23% of the watershed area).

2. Explain of effective factors

The map of factors inducing the happening of landslides in Karganeh Watershed is displayed in figure (3). The most important factors affecting landslides in the region include slope, slope direction, elevation classes, geology, drainage network (distance from the river), road (distance from the road), distance from the village, fault (distance from the fault), topographic indicators (river power index (SPI), topographic moisture index (TWI) and slope length index (LS)), geomorphological indices (topographic position index (TPI), topographic roughness index (TRI) and vector strength index (VRM)) of land use and precipitation were identified and analyzed. Nandi and Shakur [66], 7 factors of slope degree, soil type, soil erodibility, soil index, vegetation pattern, rainfall and proximity to waterway. They chose the studied basins as factors affecting landslides. Shano et al. [36] from lithology factors, groundwater conditions, distance from the fault, morphometric parameters (slope, direction and curvature), land use and precipitation, Tyagi et al. [64] from seismic data parameters, precipitation, DEM, SPI, TWI, Drainage network, slope, direction and water reservoirs. Based on the sensitivity results, slope length factors (LS, slope and topographic roughness index (TRI) were found to be the most sensitive factors in the occurrence of landslides in the studied area.

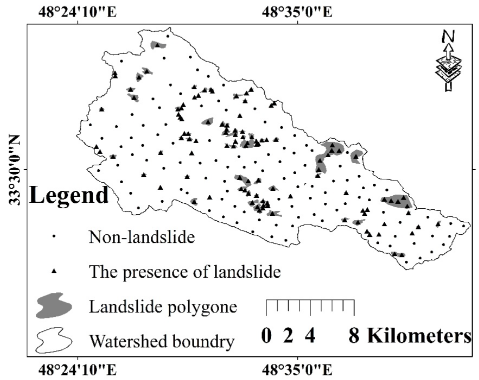
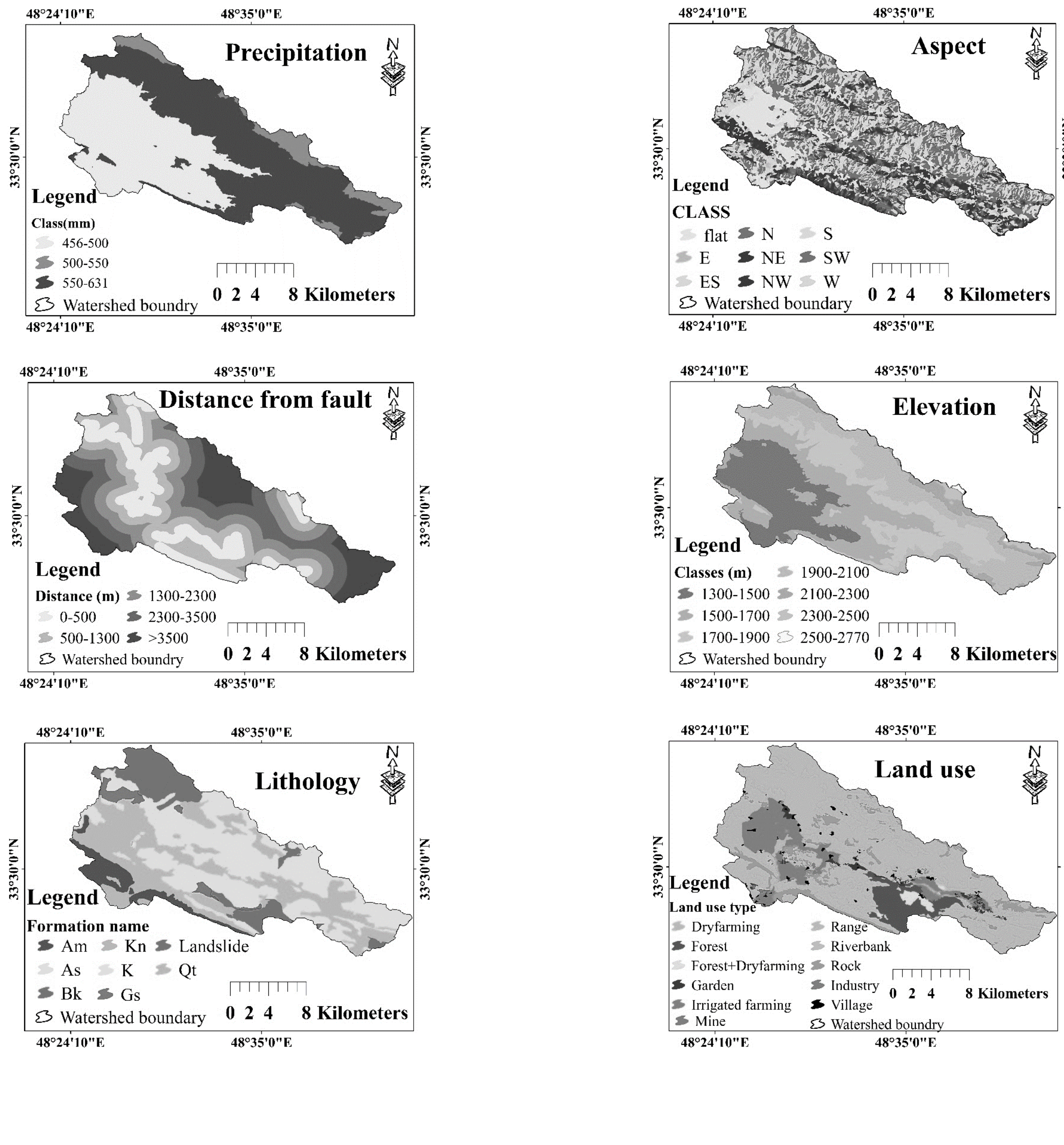


Fig.2 Landslide distribution map in the Karganeh Watershed



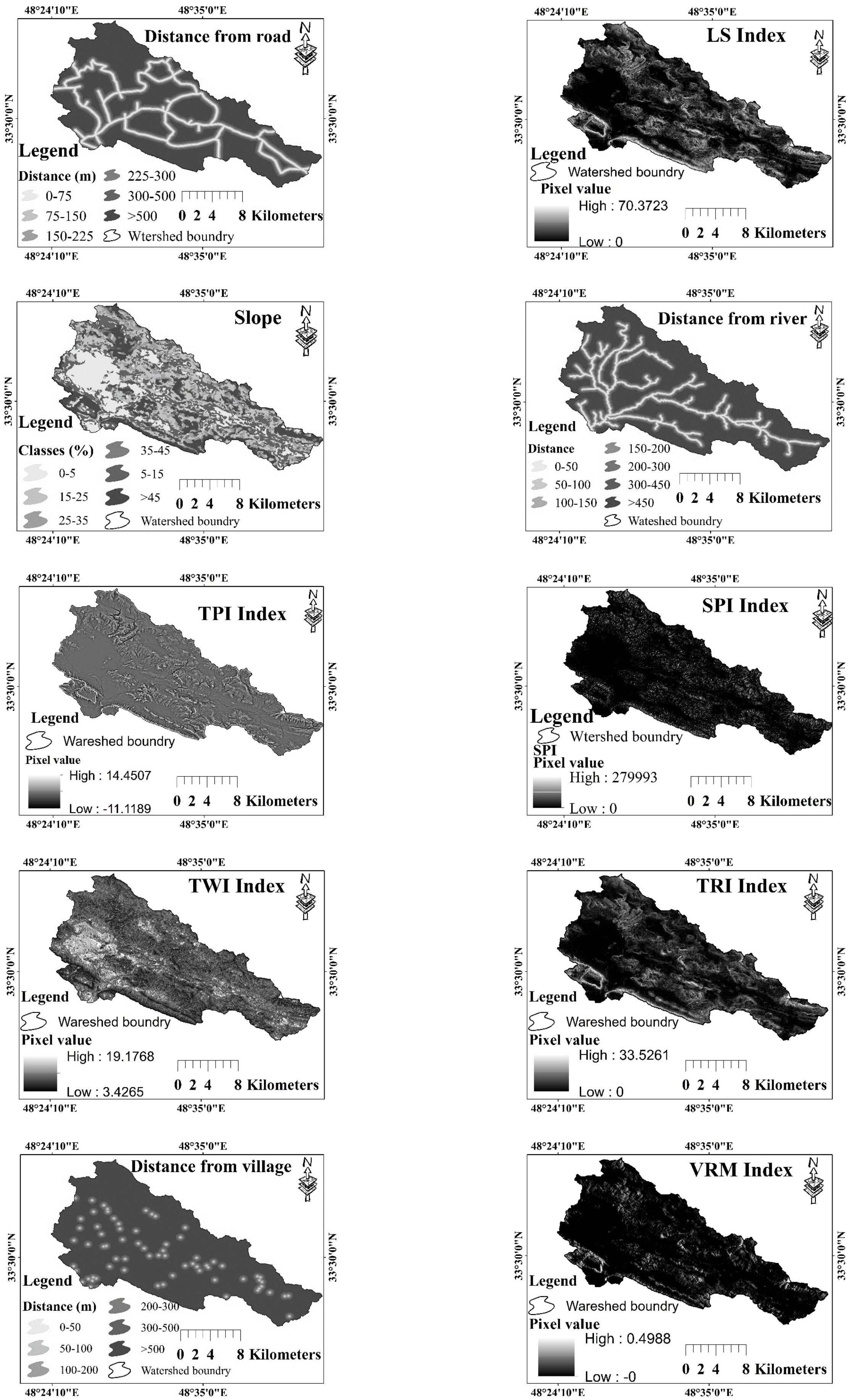


Fig.3 Map of factors affecting the occurrence of landslides in Karganeh Watershed.

3. Landslide susceptibility zonation

Using the resulting models, the landslide susceptibility maps was produced and classified in very low, low, medium, high, and very high classes (Table 2 and Fig 3).

Table 2 The distribution of area in different landslide susceptibility classes

	Forest		Entropy		Vector Machine	
Susceptibility class	Area (ha)	% Area	Area (ha)	% Area	Area (ha)	% Area
Very low	19681.29	66.9	7709.38	26.2	15240.48	51.8
Low	1691.32	5.7	7134.48	24.2	3772.63	12.8
Medium	799.74	2.7	6347.89	21.6	1774.75	6.1
High	992.39	3.4	7342.38	25	4501.64	15.3
Very high	6251.5	21.3	882.13	3	4126.74	14
Total	29416.24	100	29416.24	100	29416.24	100

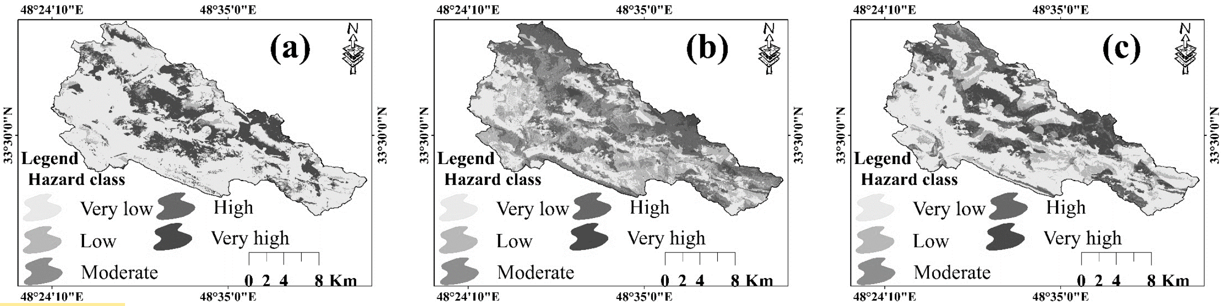


Fig.3 Landslide susceptibility maps based on, a: Random Forest (RF), b: Maximum Entropy (ME), c: Support Vector Machine (SVM)

The results of the validation evaluation of the models used in this research using the ROC curve method are shown in Figure 4. The area under the ROC index diagram for the validation of the Entropy model was 0.787, the area under the ROC index diagram was 0.913 for the validation of the Vector Machine model, the area under the ROC index diagram for validation Forest model measurement was 0.865. This shows that the models used in zoning and determining landslide prone areas in Karganeh Watershed have a very good capability. The Vector Machine model with ROC equal to 0.913 was selected as the best model in landslide risk assessment in Karganeh Watershed. The Entropy model calculates the complex distribution algorithm and by providing diverse results, it helps a lot to understand the phenomenon and process of occurrence and reaction of factors. This method is one of the quantitative methods of determining sensitivity and it is one of the models that has received a lot of attention in the last 10 years and has been used by researchers in different parts of the world. Convertino et al. [44] in Italy and Kim et al. [45] in South Korea have used the Entropy method in their studies. Based on the results obtained from this model, 50.2% of the area of the area is in the very low and low sensitivity class, 21.6% is in the medium sensitivity class, and 28% is in the high and very high sensitivity regionalization level. After evaluating this model with the ROC index, the amount of surface area under the graph in the validation stage was 0.787, this result indicates the average capability of the model in zoning and determining areas prone to landslide susceptibility in Karganeh Watershed. The Vector Machine model is one of the supervised machine learning models used to classify and separate data. The main idea of this algorithm is a binary classification using training points, which transforms the original input space into a space with higher dimensions, in order to find a desirable hyperplane. The training points that are close to the desired plane are called support vectors. Once the decision level is obtained, it can be used to estimate new data. Hang et al. [33], Pandey et al. [47] Lee et al. [51] and Pham et al. [50] and Yang et al. [76] have used the SVM algorithm in their studies for landslide susceptibility zoning. Based on the results obtained from this model, 64.6% of the district surface area is in the very low and low sensitivity class, 6.1% is in the medium sensitivity class, and 29.3% is in the high and very high sensitivity area. The amount of surface area under the ROC index diagram was obtained at 0.913 in the validation stage, which indicates the very good capability of the model in determining areas prone to landslide susceptibility in the Karganeh Watershed. The Forest model forms a cluster of decision trees. During the modeling, it involves the effective underlying factors and the evidence of landslide occurrence. Moreover, with high repetition, it removes the inefficient decision branches in the modeling process and finally continuously improves the predictions. By minimizing the prediction error, it has a very high power in predicting landslide susceptibility. Similar results have been obtained in the researches of [20-21, 73,76]. Based on the results obtained from this model, 72.6% of the area of the district is in the very low and low sensitivity class, 2.7% in the medium sensitivity class, and 24.7% of the area is in the high and very high sensitivity class.

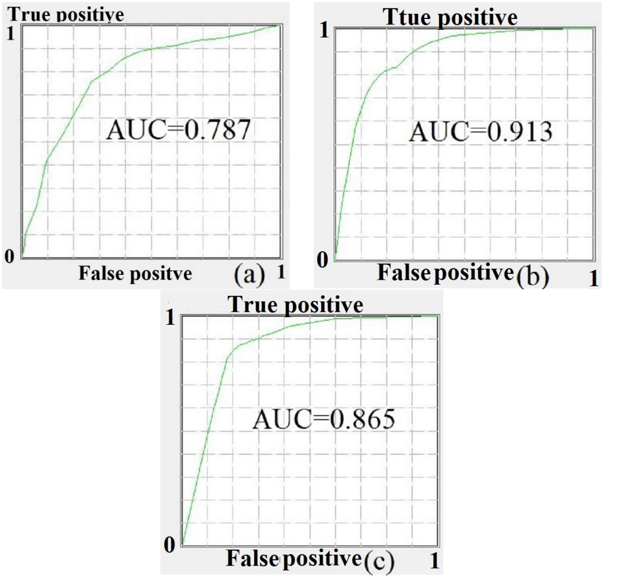


Fig.4 ROC curves a: Maximum Entropy (ME), b: Support Vector Machine (SVM), c: Random Forest (RF)

4. Determining the most important parameters affecting the occurrence of landslides

After selecting the best model, the outputs of this model were used as the base model to determine the most important factor influencing the occurrence of landslides in the Karganeh Watershed. Figure (5) shows the results of the Kappa index diagram to determine the most important influencing parameters and the contribution of each parameter in the prediction of the Vector Machine model. According to the diagram, geological indicators, land use, slope, topographic roughness index (TRI), slope length and slope direction are the most important influencing parameters. Slope is one of the influential parameters in domain instabilities. In factors such, as water penetration in the slopes, the failure angle and soil adhesion. Many researches have emphasized the direct role of slope and its effect on landslides [56-57,77]. Landslide susceptibility increases with increasing slope. It is found that it represents the increase of gravity force in high slopes. Therefore, the slope factor compared to the height is introduced as a better and more complete factor (in terms of available gravity information) for landslide susceptibility modeling. The slope length index (LS) or sediment transport (STI) represents It is the carrying capacity of the waterway. It actually determines the effect of topography on erosion. The longer the slope is, the higher the sediment carrying capacity is, and the condition of landslides on the side of the waterway increases. An increase in the topographic roughness index (TRI) indicates uplift and nontectonic activity. An increase in this index leads to more surface roughness and an increase in slope, which makes landslides more prone to occur.

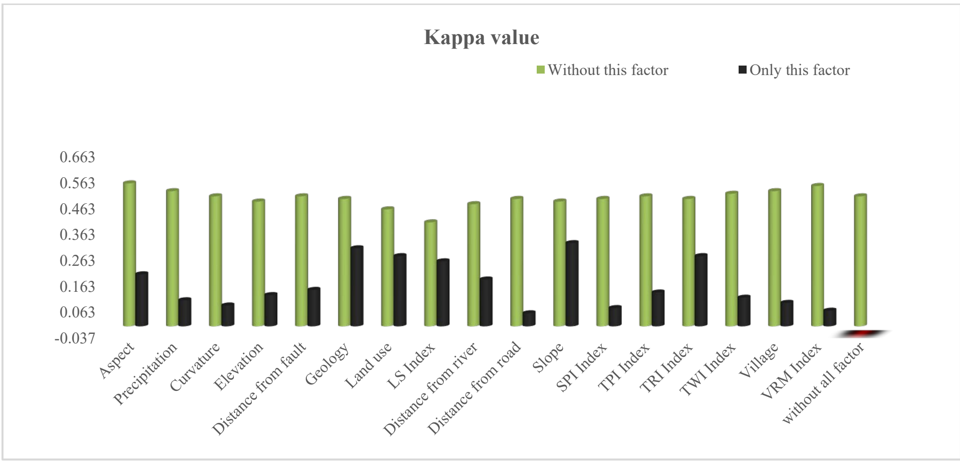


Fig.5 The results of the Kappa index diagram to determine the most important influencing parameters

Conclusion

In this study, it was tried to use all effective factors in order to evaluate landslide susceptibility in Karganeh Watershed. Principal component analysis (PCA), tolerance and VIF were used to determine the relationship between the factors influencing the occurrence of landslides and to determine the most effective factors. In order to determine the best method of landslide sensitivity, three machine learning models including: Entropy, Forest and Vector Machine model were used in Karganeh Watershed. In order to model landslide susceptibility, 70% of landslide points were used for model training and 30% of landslide data were used for model validation. The amount of surface area under the ROC index diagram was obtained at 0.865 in the validation stage, which indicates the good capability of the model in determining landslide susceptibility areas in Karganeh Watershed. According to the results obtained from the evaluation of the models using the area under the ROC curve, the Vector Machine model was selected as the best model in the zoning of landslide susceptibility in Karganeh Watershed. This shows that this model has high accuracy in evaluating the landslide susceptibility in the studied area. By comparing the results obtained with the real conditions through field visits, there is a very high agreement between the results of the landslide susceptibility map using the SVM and the actual evidence in the study area. According to the landslide susceptibility map using the SVM model, the areas with high susceptibility are located on the steep slopes upstream and at high altitudes, which caused the upstream material to fall down due to the high slope and the effect of gravity. In addition, a large part of the studied area is in low slopes with agricultural and orchards, which have low and very low susceptibility in terms of landslide occurrence. The Karganeh Watershed's situations for example climatic condition, geology, roughness, geomorphology and tectonic conditions in addition to human compression factors such as land use and roads changes has formed a suitable background for the landslide that its happening is nearby 95 suitcases with estimated about 635 hectares in watershed. After the zonation using the Vector Machine model in Karganeh Watershed, nearby 24.7% of the watershed area is placed in high and very high susceptibility classes, which it is presentation high susceptibility to landslide for this watershed that must be deliberated in Susceptibility management, landslide damages and land use planning. Changing the rangeland to rain fed farming and road construction is completed severely in Karganeh Watershed throughout recent years and led to awarding great role of human factors on landslide in comparing other factors.

Acknowledgements

This article has been done in the procedure of a research project between the Soil Conservation and Watershed Management Research Institute and the Agricultural and Natural Resources Research Center of Lorestan Province with the number 2-59-29-031-000489. It is necessary to express our sincere gratitude for the collaboration of these centers in all stages of the project. This project was funded under the project entitled "Assessing landslide hazard and risk in Karganeh Watershed, Lorestan Province. We would like to thank the Soil Conservation and Watershed Management Research Office for their support. The authors have no conflict of interests whatsoever.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Oh, H.J., and Lee, S. (2017). "Shallow landslide susceptibility mod-eling using the data mining models artificial neural network and boosted tree." *Applied Sciences*, 7, 1000.

[2] Arabameri, A., Karimi-Sangchini, E., Pal, S.C., Saha, A., Chowdhuri, I., Lee, S., and Bui, D.T. (2020). "Novel Credal Decision Tree-Based Ensemble Approaches for Predicting the Landslide Susceptibility." *Remote Sensing*, 12, 3389.

[3] Karimi, S.E., Salehpour, J.A., and Mosafaii J. (2022). "Flood risk management in Khorramabad wa-tershed using the DPSIR framework. *Natural Hazards*, 122(1), 3101.

[4] Fowdur, T.P., Appadoo, S.S., Cowlessur, M., Gooroochurn, M., Doomah, Z., Jogee, D., and Hurbungs, V. (2025). "A Cloud-Based Landslide and Weather Nowcasting System with Collaborative Machine Learning for the Chitrakoot Region in Mauritius." *Measurement: Digitalization*, 2-3, 100008.

[5] Kelarestaghi, A., and Ahmadi, H. (2009). "Landslide suscep-tibility analysis with a bivariate approach and GIS in North-ern Iran." *Arabian journal of geosciences*, 2, 95.

[6] Ma, Z., and Mei, G. (2025). "Forecasting landslide deformation by integrating domain knowledge into interpretable deep learn-ing considering spatiotemporal correlations." *Journal of Rock Mechanics and Geotechnical Engineering*, 17(2), 960-982.

[7] Pham, B.T., Parkash, I., Singh, S.K., Shirzadi, A., Shahabi, H., Tran, T.T., and Bui, D.T. (2019). "Landslide sus-ceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learn-ing approaches." *Catena*, 175, 203.

[8] Yang, Y., Lu, Y., and Mei, G. (2023). "A federated learning based approach for predicting landslide displacement considering data security." *Future Generation Computer Systems*, 149, 184.

[9] Xiong, H., Ma C., Li M., Tan J., and Wang, Y. "Landslide susceptibility prediction considering land use change and human activity: A case study under rapid urban expansion and afforestation in China." *Science of The Total Environment*, 866, 161430.

[10] Regmi, A.D., Devkota, K., Yoshida, K., Pra-dhan, B., Pourghasemi, H.R., Kumamoto, T., and Akgun, A. (2013). "Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya." *Arabian Journal of Geosciences*, 7(2),725.

[11] Sangchini, K.E., Emami, S.N., Tahmasebipour, N., Pourghasemi, H.R., Naghibi, S.A., Arami, S.A., and Pradhan, B. (2016). "Assessment and compari-son of combined bivariate and AHP models with logistic regression for landslide susceptibility mapping in the Cha-harmahal-e-Bakhtiari Province, Iran." *Arabian journal of ge-osciences*, 9(201).

[12] Sharma, A., Mohana, R., Kukkar, A., and Chodha, V., and Bansal, P. (2023). "An ensemble learning-based experimental framework for smart landslide detection, monitoring, pre-diction, and warning in IoT-cloud environment." Environmental Science and Pollution Research, 30, 122677-12699.

[13] Aleotti, P., and Chowdhury, R. (1999). "Landslide hazard as-sessment: summary review and new perspectives." Bulletin of Engineering Geology and the Environment, 58, 21-44.

[14] Arabameri, A., Saha, S., Roy, J., Chen, W., blaschke, T., and Bui, D.T. (2020). "Landslide Susceptibility Evaluation and Management Using Different Machine Learning Methods in The Gallicash River Watershed, Iran." *Remote Sensing*, 12(3), 475.

[15] He. R., Zhang, W., Dou, J., Jiang, N., Xiao, H., and Zhou, J. (2024). "Application of artificial intelligence in three as-

pects of landslide risk assessment: A comprehensive review." *Rock Mechanics Bulletin*, 3(4), 100144, 2024.

[16] Sangchini, E., Ownegh, M., Sadaldin, A., and Mashayekhan, A. (2011). "Probabilistic Landslide Risk Analysis and Mapping (Case Study: Chehel-Chai watershed, Golestan Province, Iran)." *Journal of Rangeland Science,* 2(1), 425.

[17] Aghda, S.M.F., Baghery, V., and Razifard, M. (2018). "Landslide susceptibility mapping using fuzzy logic system and its influences on mainlines in lashgarak region, Tehran, Iran." *Geotechnical and Geological Engineering*, 36, 915-937.

[18] Li, B., Yongsheng, L., Niu, R., Xue, T., and Duan, H. (2024). "Early warning of landslides based on statistical analysis of landslide mo-tion characteristics and AI Earth Cloud InSAR processing system: a case study of the Zhenxiong landslide in Yunnan Province, China." *Landslides*, 21(11), 1-12.

[19] Chen, W., Poughasemi, H.R., Konejady, A., and Zhang, N. (2017). "Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques." *Geoderma*, 305(1), 314-327.

[20] Rahmati, O., Kornejady, A., Samadi, M., Nobre, A.D., and melesse, A.M. (2018). "Development of an automated GIS tool for reproducing the HAND terrain model." *Environmental modelling and software*, 102, 1-12.

[21] Poughasemi, H.R., and Rahmati, O. (2018). "Prediction of the landslide susceptibility: Which algorithm, which precision?" *Catena*, 162, 177-192.

[22] Ilinca, V., Șangric, I., and Gheuca, I. (2025). "Landslide sus-ceptibility analysis using machine learning: Insights from the bend Subcarpathians-Carpathians, Romania." *Geomorphology*, 486, 109872.

[23] Carrara, A., Giovanni, C., and Frattini, P. (2003). "Geomorphological and historical data in assessing landslide hazard." *Earth Surface Processes and Landforms*, 28(10), 1125-1142.

[24] Dahal, R. K., Hasegawa, S., Nonomura, S., Yamanaka, M., Masuda, T., and Nishino, K. (2008). "GIS-based weights-of-evidence modelling of rainfall-induced land-slides in small catchments for landslide susceptibility mapping." *Environmental Geology*, 54, 311-324.

[25] Bathrellos, G.D., Kalivas, D.P., and Skilodimou, H.D. (2009). "GIS-based landslide susceptibility mapping models applied to natural and urban planning in Trikala, Central Greece." *Estudios Geológicos*, 65(1), 49-65.

[26] Kayastha, P., Dhital, M.R., and De Smedt, F. (2013). "Evaluation and comparison of GIS based landslide susceptibility mapping procedures in Kulekhani watershed, Nepal." *Journal of the Geological Society of India*, 81, 219-231.

[27] Sangchini, E.K., Arami, A., Moghadam, H.R., Khodabakhshi, Z., and Jafari, R. (2014). "Landslide Risk Assessment for Baba Heydar Watershed, Chaharmahal and Bakhtiari Province, Iran." *Iranian journal of earth sciences*, 6(1), 121.

[28] Qian, H, Wang, M., and Liu, K. (2021). "Rapidly assessing earth-quake-induced landslide susceptibility on a global scale us-ing random forest." *Geomorphology*, 391(15), 107889.

[29] Chang, Z., Huang, J., Huang, F., Bhuyan, K., and Raj Meena, S., and Catani, F. (2023). "Uncertainty analysis of non-landslide sample selection in landslide susceptibility predic-tion using slope unit-based machine learning models." *Gondwana Research*, 117, 307-320.

[30] Qian, L., Ou, L., Li, G., Chen, Y., and Qian, B. (2025). "Optimiz-ing the application of machine learning models in predicting landslide susceptibility using the information value model in Junlian County of Sichuan Basin." *Advances in Space Research*, 76(2), 699-717.

[31] Harmouzi, H., Nefeslioglu, H.A., Rouai, M., Sezer, E.A., Dekayir, A., and Gokceoglu, C. (2019). "Land-slide susceptibility mapping of the Mediterranean coastal zone of Morocco between Oued Laou and El Jebha using artificial neural networks (ANN)." *Arabian Journal of Geosciences*, 12, 696.

[32] Yao, J., Qin, S., Qiao, S., Liu, X., Zhang, L., and Chen, J. (2022). "Application of a two-step sampling strategy based on deep neural network for landslide susceptibility mapping." *Bulletin of Engineering Geology and the Environment*, 81, 148.

[33] Hong, H., Pradhan, B., Xu, C., and Bui, D.T. (2015). "Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines." *Catena*, 133, 266-281.

[34] Abedini, M., Ghasemyan, B., and Rezaei Mogaddam, M.H. (2017). "Landslide susceptibility mapping in Bijar city, Kurdistan Province, Iran: a comparative study by logistic regression and AHP models." *Environmental Earth Sciences*, 76, 308.

[35] Hemasinghe, H., Rangali, R.S., Deshapriya, N.L., and Samarakoon, L. (2018). "Landslide susceptibility mapping using logistic regression model (a case study in Badulla District, Sri Lanka)." *Procedia engineering*, 212, 1046-1053.

[36] Shano, L., Raghuvanshi, T.K., and Meten, M. (2022). "Landslide Hazard Zonation using Logistic Regression Model: The Case of Shafe and Baso Catchments, Gamo Highland, Southern Ethiopia." *Geotechnical and Geological Engineering*, 40, 83-101.

[37] Puente, F., Mustafa, A., and Teller, J. (2021). "Landslide Susceptibility Mapping of Urban Areas: Lo-gistic Regression and Sensitivity Analysis applied to Quito, Ecuador." *Geoenviron Disasters*, 8(1).

[38] Zhou, X., Wen, H., Zhang, Y., Xu, J., and Zhang, W. (2021). "Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization." *Geoscience Frontiers*, 12(5), 101211.

[39] Sevgen, E., Kocaman, S., Nefeslioglu, H.A., and Gokceoglu, C. (2019). "A novel performance assessment approach using photogrammetric techniques for landslide sus-ceptibility mapping with logistic regression, ANN and random forest." *Sensors*, 19(18), 3940.

[40] Zhao, L., Wu, X., Niu, R., Wang, Y., and Zhang, K. (2020). "Using the rotation and random forest models of ensemble learning to predict landslide susceptibility." *Geomatics, Nat-ural Hazards and Risk*, 11(1), 1542-1564.

[41] Sun, D., Xu, J., Wen, H., and Wang, D. (2021). "Assessment of landslide susceptibility mapping based on Bayesian hy-perparameter optimization: A comparison between logistic regression and random forest." *Engineering Geology*, 281, 105972.

[42] Burak, F.T., Abbaspour, A., Alimohammadlou, Y., and Tecuci, G. (2021). "Landslide susceptibility analyses using Random Forest, C4.5, and C5.0 with balanced and unbalanced datasets." *Catena*, 203, 105355.

[43] Phillips, S.J., Anderson, R.P., and Schapire, R.E. (2006). "Maximum entropy modeling of species geographic distributions." *Ecological Modeling*, 190(3-4), 231-259.

[44] Convertino, M., Troccoli, A., and Catani, F. (2013). "Detecting fingerprints of landslide drivers: a MaxEnt model." *Journal of Geophysical Research: Earth Surface*, 118(3), 1367-1386.

[45] Kim, H.G., Lee, D.K., Park, C., Kil, S., Son, Y., and Park J.H. (2015). "Evaluating landslide hazards using RCP 4.5 and 8.5 scenarios." *Environmental earth sciences*, 73(3),1385-1400.

[46] Kerekes, A., Poszet, S., and Andrea, G. (2018). "Landslide susceptibility assessment using the maximum entropy model in a sector of the Cluj–Napoca Municipality, Romania." *Revista de Geomorfologie*, 20(1), 130-146.

[47] Pandey, V.K., Pourghasemi, H.R., and Sharma, M.C. (2018). "Landslide susceptibility mapping using maximum en-tropy and support vector machine models along the High-way Corridor, Garhwal Himalaya." *Geocarto International*, 35(2), 1-38.

[48] Barman, J., Ali, S.S., Biswas, B., and Das, J. "Applica-tion of index of entropy and Geospatial techniques for land-slide prediction in Lunglei district, Mizoram, India." *Natural Hazards Research*, 3(3), 508-521.

[49] Maashi, M., Alzaben, N., Negm, N., Venkatesan, V., Sabarunisha Begum, S., and Geetha, P. (2025). "Advanced landslide susceptibility mapping and analysis of driving mechanisms using ensemble machine learning models." *Journal of South American Earth Sciences*, 151, 105272.

[50] Pham, B.T., Pradhan, B., Bui, D.T., Prakash, I., and Dholakia, M.B. (2016). "A comparative study of different ma-chine learning methods for landslide susceptibility as-sess-ment: A case study of Uttarakhand area (India)." *Environmental Modelling and Software*, 84, 240-250.

[51] Lee, S., Hong, S.M., and Jung, H.S. (2017). "A support vector ma-chine for landslide susceptibility mapping in Gangwon Province, Korea." *Sustainability*, 9(1), 48.

[52] Kornejady, A., Pourghasemi, H.R., and Afzali, S.F. (2019). "Presentation of RFFR new ensemble model for land-slide susceptibility assessment in Iran." *In Landslides: theory, practice and modeling*, 123,143.

[53] Panahi, M., Gayen, A., Pourghasemi, H.R., Rezaie, F., and Lee, S. (2020). "Spatial prediction of landslide suscep-tibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms." *Science of The Total Environment*, 741, 139937.

[54] Balogun, A.L., Rezaie, F., Pham, Q.B., Gigović, L., Drobnjak, S., Aina, Y.A., Panahi, M., Temitope, S., and Yekeen, S.L. (2021). "Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms." *Geoscience Frontiers*, 12(3), 101104.

[55] Mahalingam, R., Olsen, M.J., and O'Banion, M.S. (2016). "Evaluation of landslide susceptibility mapping techniques using lidar-derived conditioning factors (Oregon case study)." *Geomatics. Natural Hazards and Risk*, 7, 1884-1907.

[56] Lee, S. (2005). "Application of logistic regression model and its val-idation for landslide susceptibility mapping using GIS and remote sensing data." *International Journal of Remote Sensing*, 26(7), 1477-1491.

[57] Devkota, K., Regmi, A., Pourghasemi, H.R, Yoshida, K., Pradhan, B., Ryu, I., Dhital, M., and Althuwaynee, Q. (2013). "Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya." *Nat Hazards*, 65, 135-165.

[58] Kohavi, R. (1995). "A study of cross-validation and bootstrap for accuracy estimation and model selection." In Proceedings of the International Joint Conference on Artifcial Intelligence (IJCAI), Montreal, QC, Canada, 20-25.

[59] Václavík, T., and Meentemeyer, R.K. (2009). "Invasive spe-cies distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions?" *Ecological Modelling*, 220(23), 3248-3258.

[60] Boria, R.A., Olson, L.E., Goodman, S.M., and Anderson, R.P. (2014). "Spatial filtering to reduce sampling bias can improve the performance of ecological niche models." *Ecological Modelling*, 275, 73-77.

[61] Kalantar, B., Ueda, N., Saeidi, V., Ahmadi, K., Halin, A.A., and Shabani, F. (2020). "Landslide Susceptibility Mapping: Machine and Ensemble Learning Based on Remote Sensing Big Data." *Remote Sensing*, 12(11), 1737.

[62] Bui, D.T., Ho, T.C., Pradhan, B., Pham, B.T., Nhu, V.H., and Revhaug, I. (2016). "GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks." *Environmental Earth Sciences*, 75, 1101.

[63] Wilson, J.P., and Gallant, J.C. (2000). *Terrain Analysis: Prin-ciples and Applications*. John Wiley & Sons: Hoboken, NJ, USA; 524.

[64] Tyagi, A., Tiwari, R.K., and James, N. (2021). "GIS-Based Landslide Hazard Zonation and Risk Studies Using MCDM." *In Local Site Effects and Ground Failures, Springer Singa-pore*, 251.

[65] Grohmann, C.H., Smith, M.J., and Riccomini, C. (2011). “Multiscale Analysis of Topo- graphic Surface Roughness in the Midland Valley, Scotland.” *IEEE Transactions on Geoscience and Remote Sensing*, 49(4), 1200-1213.

[66] Nandi, A., and Shakoor, A. (2009). “A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses.” *Engineering Geology*, 110(1-2), 11-20.

[67] Horton, R.E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology, *GSA Bulletin*, 56(3), 275-370.

[68] Chen, W., Yan, X., Zhao, Z., Hong, H., Bui, D.T., and Pradhan, B. (2019). “Spatial prediction of landslide susceptibility using data mining-based kernel logistic regression, naive Bayes and RBFNetwork models for the Long County area (China).” *Bulletin of Engineering Geology and the Environment*, 78, 247-266.

[69] Jebur, M.N., Pradhan, B., and Tehrany, M.S. (2014). “Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale.” *Remote Sensing Environmental*, 152, 150-165.

[70] Mokarram, M., Roshan, G., and Negahban, S. (2015). “Landform classification using topography position index (Case study: Salt dome of Korsia-Darab plain, Iran).” M*odeling Earth Systems and Environment*, 1, 40.

[71] Ozdemir, A. (2011). “Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey).” *Journal of Hydrology*, 405(1-2), 123-136.

[72] Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., and Khosravi, K. (2017). “A novel hybrid artificial intelligence approach for flood susceptibility as-sessment.” *Environmental Modelling and Software*, 95, 229-245.

[73] Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D.T., Duan, Z., and Ma, J. (2017). “A comparative study of logistic model tree, random forest and classification and regression tree models for spatial prediction of landslide susceptibility.” *Catena*, 151, 147-160.

[74] Park, N.W. (2015). “Using maximum entropy modeling for land-slide susceptibility mapping with multiple geo environmental data sets.” *Environmental Earth Sciences*, 73(3), 937-949.

[75] Pontius, R.J., and Schneider, L.C. (2001). “Land-cover change model validation by an ROC method for the Ipswich water-shed, Massachusetts, USA, Agriculture.” *Ecosystems and Environment*, 85(1-3), 239-248.

[76] Yang, S., He, Z., Chai, J., Meng, D., Macek, W., Branco, R., and Zhu, S.P. (2023). “A novel hybrid adaptive frame-work for support vector machine-based reliability analysis: A comparative study.” *Structures*, 58, 105665.

[77] Wang, H., Zhang, L., Yin, K., Luo, H., and Li, J. (2021). “Land-slide identification using machine learning.” *Geoscience Frontiers*, 12(1), 351-364.

2025 年 03 月 12 日　收稿
2026 年 08 月 11 日　修正
2026 年 02 月 20 日　接受