Scientia Iranica A (2022) 29(3), 1077–1094

Sharif University of Technology
Scientia Iranica
Transactions A: Civil Engineering
http://scientiairanica.sharif.edu

An assessment of data mining and bivariate statistical methods for landslide susceptibility mapping

A. Aram[a], M.R. Dalalian[a,*], S. Saedi[a], O. Rafieyan[b], and S. Darbandi[c]

a. *Department of Soil Science, Tabriz Branch, Islamic Azad University, Tabriz, Iran.*
b. *Department of Environmental Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.*
c. *Department of Water Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.*

Received 4 January 2021; received in revised form 23 April 2021; accepted 14 November 2021

KEYWORDS
Bagging;
Landslide hazard assessment;
Random forest;
Shannon entropy;
West Azerbaijan.

Abstract. Landslide is recognized as one of the environmental challenges that causes land degradation, fertility reduction, and other significant damages to the ecosystem. Therefore, proper identification of landslide-prone areas through modeling is significantly helpful for land development managers and planners by providing them with appropriate management strategies for preventing land degradation. In this research, landslide susceptibility mapping was carried out in West Azerbaijan province, Iran using Frequency Ratio (FR), Shannon Entropy (SE), Random Forest (RF), and an ensemble of Random Forest and Bagging (RF-BA) methods. Based on field surveys, local interviews, and review of similar studies, 12 factors were identified that affected landslide occurrence, namely altitude, slope angle, slope aspect, distance from fault, distance from river, distance from road, drainage density, road density, rainfall, soil, land use, and lithology. In the field surveys, 110 landslides in the area were specified; 70% of the data (77 landslides) were randomly selected and used for modeling and the remaining 30% (33 landslides) for validation. The results of the ROC curve exhibited the accuracy rates of 0.92, 0.91, 0.89, and 0.88 through RF-BA, RF, FR, and SE models, respectively.

1. Introduction

Hillsides in nature are affected by external processes due to geological, geomorphological, and climatic changes, which may lead to the occurrence of slope movements such as landslides. Landslide is a prominent case of land degradation among the seven soil threat factors that change landscape, reduce production resources, and disrupt transportation activities [1]. Although surface landslides apparently do not pose much hazard and displace relatively small volumes of soil down the hillsides, they may cause significant damages to agriculture, structures, infrastructures, and transportation, which are sometimes intense and even fatal [2]. Landslide is smaller in scale, but more frequent than other geological hazards around the world and in many cases, it is more dangerous [3]. Landslide, after flood and earthquake, is recognized as the third most significant natural hazard in the world. According to the World [4], 3.7 million square kilometers of land area and 300 million of the population of the world (about 5%) are at the risk of landslides.

Given the repercussions of landslides on agricultural and natural resources as well as human infrastructure, the significance of research and plans for preventing or reducing their negative impacts remains undisputed [5,6]. One way to reduce the risk of landslide hazards in life and economy is to increase readiness through landslide forecasting systems or

*. *Corresponding author.*
E-mail address: mdalalian@iaut.ac.ir (M.R. Dalalian)

doi: 10.24200/sci.2021.57334.5240

identify the susceptible and endangered areas [7,8]. Therefore, recognizing the unstable regions and causes of landslide events, investigating their geometrical and morphological features, and determining the relationship between the factors affecting landslides and their morphological characteristics have gained much more significance. A better understanding of the interconnection among the environmental factors in the occurrence of landslides is essential to quantitative assessment of landslide sensitivity. Hence, a reliable evaluation depends on the quality of available data and proper utilization of analysis and modeling methods [9]. So far, much research has been carried out on landslide modeling using Geographic Information System (GIS) and remote sensing. Bivariate statistical methods [10], weight of evidence [11], evidential belief function [12], certainty factor [13], Shannon Entropy (SE) [14], machine learning methods (e.g., artificial neural network) [15], Adaptive Neuro-Fuzzy Inference System (ANFIS) [16], support vector machine [17], decision tree [18], Random Forest (RF) [19], logistic regression [20], decision-making methods (e.g., analytic hierarchy process) [21], and analytic network process [22] are some of the approaches that have been adopted in such studies.

The main challenge in landslide modeling is the uncertainty that affects the predictive ability of models as well as the accuracy of landslide susceptibility maps derived from the models [23]. In other words, the major problem is the lack of a comprehensive framework that is acceptable for use in all extended models [12]. Given that a number of factors in nature affect landslides that are not easily identifiable and collectible, the models developed by planners may not be complete and the parameters of different models may be different. Accordingly, the results of no specific model can be accepted with total certainty [24]. Not surprisingly, the current trend of research is mostly directed at combining algorithms to go beyond the existing limitations in order to develop a standard framework for integrating the main concepts of modeling and reducing the uncertainties in studies on landslide. The final objective is to achieve spatial prediction maps with higher reliability. ANFIS methods in combination with genetic algorithm and particle swarm optimization [12], combined Artificial Neural Network (ANN) and particle swarm optimization [25], rotation forest-based support vector machines [26], least-squares support vector classification and bat algorithm [27], an ensemble of Stochastic Gradient Descent (SGD) and AdaBoost meta classifier [28], kernel logistic regression [29], Random Subspace (RS) [30], Best First Decision Trees (BFDT) [31], Logistic Model Tree (LMT) [32], SGD [27], and Nave Bayes Tree (NBT) [26] are samples of such efforts. On the other hand, since there is no algorithm identified with the highest global performance, researchers have always been looking for new and more robust algorithms. The main principle behind the ensemble model is that a group of weak learners come together to form a strong learner, thus increasing the accuracy of the model [33]. When we try to predict the target variable using any machine learning technique, the main causes of the difference in actual and predicted values are noise, variance, and bias. Application of some techniques such as Bagging helps decrease the variance and increase the robustness of the model. Combinations of multiple classifiers decrease variance, especially in the case of unstable classifiers, and may produce a more reliable classification than that of a single classifier [34].

The present study also attempts to achieve high-accuracy landslide susceptibility mapping by combining data mining algorithms (RF and RF-BA). The results were then compared with the those of bivariate statistical methods (FR and SE) to test the reliability of the proposed method for the study area.

The rest of this paper is organized as follows. In Section 2, the study area is generally introduced that comprises the following information on the factors affecting the occurrence of landslides and how to prepare them, statistical methods, and data mining, and evaluation. In Section 3, the results obtained from statistical methods as well as those from combined data mining methods, landslide susceptibility map visualization, and validation of results are examined. In Section 4, the results received by statistical methods are discussed, the importance of these criteria is highlighted, and data mining models are compared. Finally, concluding remarks are given in Section 5.

2. Material and method

2.1. Study area

West Azerbaijan province is located in the northwestern Iran with an area of 37412 km^2 between east longitudes of 44°3′ and 47°24′ and north latitudes of 36°05′ and 39°46′. Topographically, the minimum and maximum elevations in the province are 605 and 3600 m, respectively, and the slope angle ranges from 0 to 62°. The average annual rainfall ranges between 310 and 900 mm, and the maximum annual temperature, minimum annual temperature, and annual mean temperature are 31.7°C, −6.9°C, and 10.9°C, respectively. In terms of land use, most of the province is occupied by agricultural land use. The study area along with the landslide locations is shown in Figure 1.

2.2. Landslide inventory map

Precise determination of the location of landslides and establishment of a landslide database are essential for risk studies. However, determining the exact locations and areas of landslides is challenging and

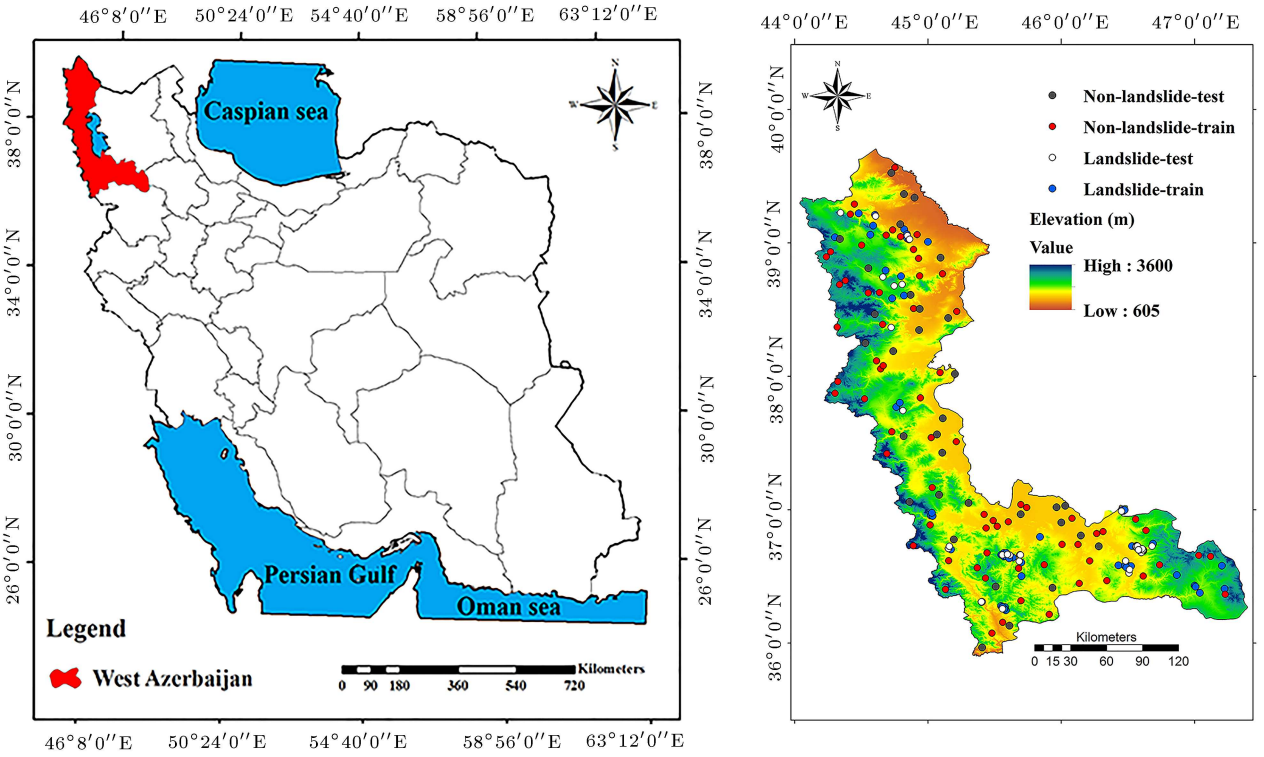

Figure 1. Study area and landslide location.

time-consuming. It requires surveying the terrain and checking the recorded areas through aerial photographs and satellite images. In this study, first, the landslide location information was obtained from the Iranian Forest, Range, and Watershed Management Organization (FRWO). Then, 110 landslides were identified through field surveys, aerial photo interpretation, and Google Earth satellite imagery (Figure 1).

2.3. Criteria for landslide modeling

This study employed altitude, slope angle, slope aspect, distance from fault, lithology, distance from river, distance from road, drainage density, road density, land use, soil, and rainfall for landslide hazard modeling. Factors affecting the occurrence of landslides were selected based on previous studies, expert opinion, and data access [12,16]. A digital elevation model with a spatial resolution of 30 m was prepared using ASTER satellite images, and the altitude, slope angle, and slope aspect layers were prepared in ArcGIS 10.3. Altitude indirectly affects vegetation and soil moisture, thus determining hillside stability status [35]. The altitude map is divided into five classes of 610–1200 m, 1300–1600 m, 1700–1900 m, 2000–2400 m, and 2500–3600 m (Figure 2(a)). Slope angle is also a key factor that affects the stability status of hillside. Upon increasing the slope angle, shear force and, consequently, the likelihood of landslide occurrence would increase [36]. The slope angle map is also divided into five classes of 0°–4.6°, 4.7°–10°, 11°–17°, 18°–25°, and more than 26°–62°(Figure 2(b)). The slope aspect affects the hydrological processes of evapotranspiration, weathering, and vegetation [18]. The slope aspect map is divided into nine classes of flat, north, northeast, east, southeast, south, southwest, west, northwest, and north (Figure 2(c)). Rainfall is regarded as one of the landslide driving factors that reduces the hillside confidence factor and consequently, increases the probability of landslide occurrence. Rainfall data were obtained from the Meteorological Organization of Iran. Rainfall maps were then prepared using station locations in the area under study through the kriging interpolation method (30-year stats). To select the best interpolation method, the Root Mean Square Error (RMSE) index was used, which was the lowest in the kriging method among the other methods for interpolating rainfall. Semivariograms related to the interpolation of rainfall through the kriging method are shown in Figure 3. The rainfall map was divided into five classes of 310–380 mm, 390–440 mm, 450–540 mm, 550–700 mm, and 710–900 mm (Figure 2(d)). In most cases, discontinuous lithological structures such as faults, seams, and crevices can accelerate the occurrence of landslides and provide the conditions for weathering. The geological map was obtained from the Geological Survey of Iran. The distance-from-fault map was derived from the geological map of West Azarbaijan province on a scale of 1:100,000, hence divided into six classes of 0–500 m, 500–1000 m, 1000–2000 m, 2000–5000 m, 5000–10000 m, and

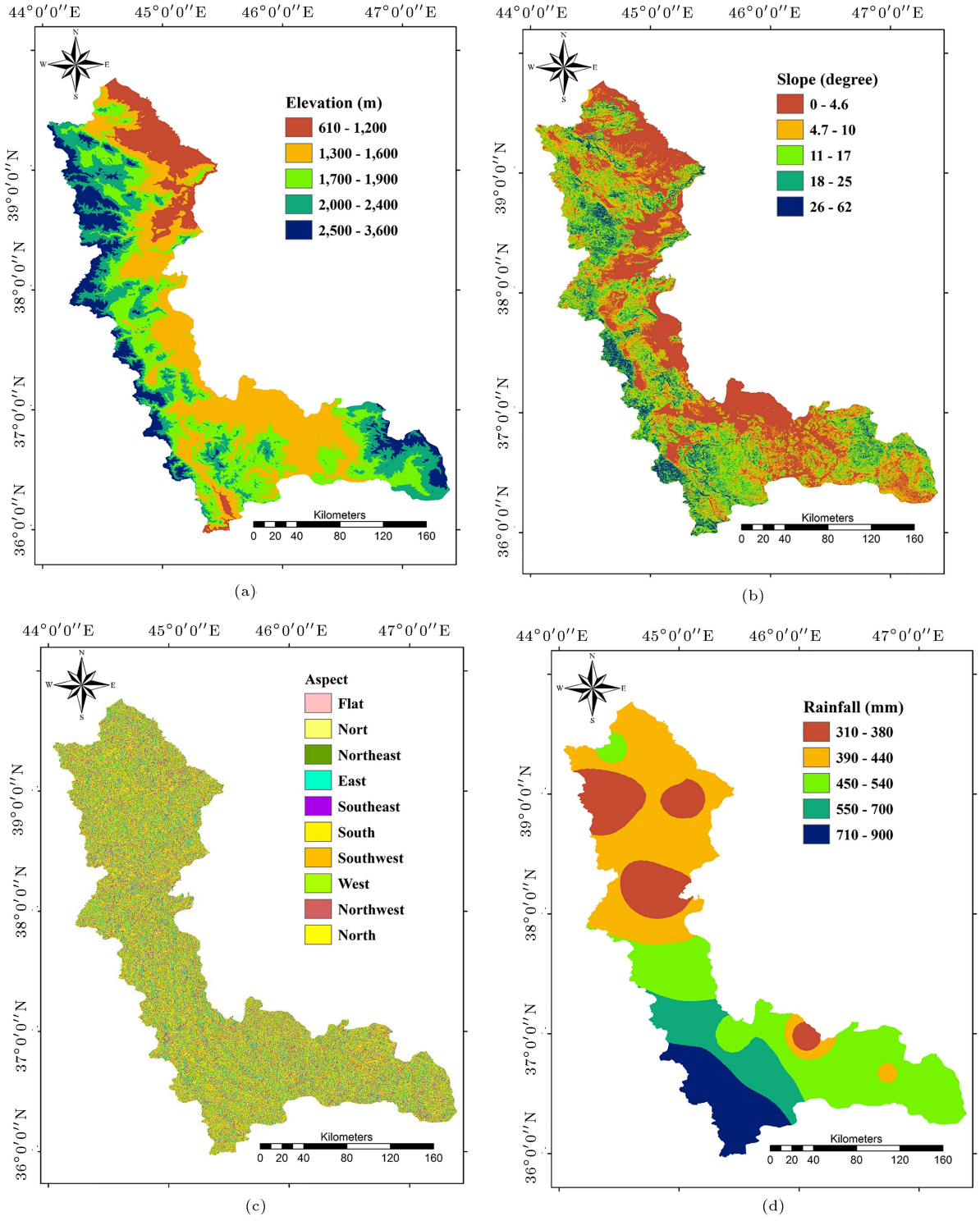

Figure 2. Landslide conditioning factors.

10000–50000 m (Figure 2(e)). The distance from river was introduced into the model to evaluate the role of runoff and effect of erosion (heel) in the waterway in the occurrence of landslide [13]. The streams layer was obtained from the Natural Resources Organization of West Azerbaijan Province. The layer for the distance from river was prepared on a scale of 1:50000 and divided into six classes of 0–500 m, 500–1000 m, 1000–2000 m, 2000–5000 m, 5000–10000 m, and more than 10,000 m (Figure 2(f)). The distance from road is considered as one of the factors of human interference in a number of studies on landslide risks [37]. Road construction near the hillside may cause changes in the natural conditions of areas, which may have been in

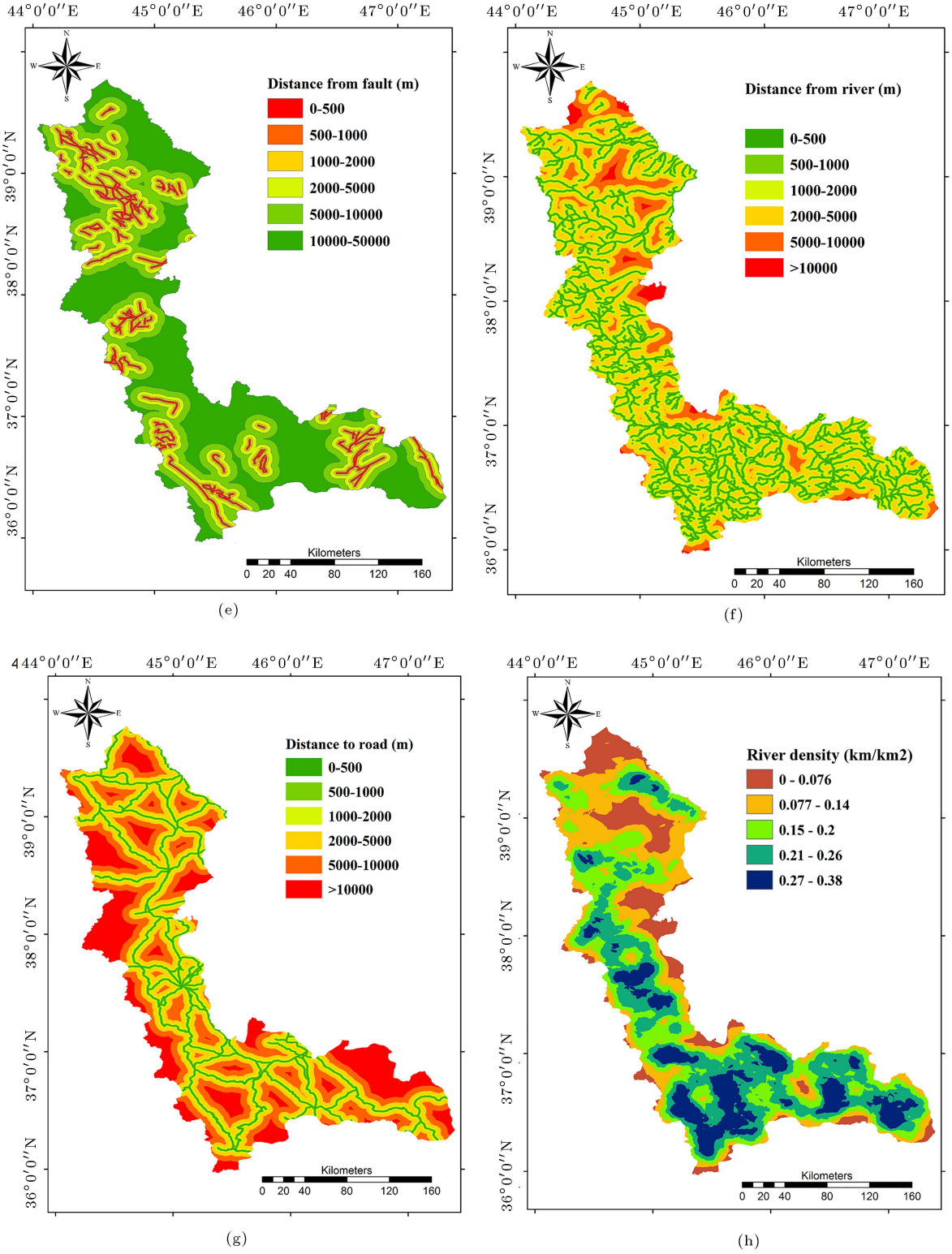

Figure 2. Landslide conditioning factors (continued).

full equilibrium before such constructions. The layer of distance from road was created on a scale of 1:50000 and divided into six classes of 0–500 m, 500–1000 m, 1000–2000 m, 2000–5000 m, 5000–10000 m, and more than 10,000 m (Figure 2(g)). With an increase in the drainage density, both rock permeability and surface flow velocity would decrease. As a result, followed by saturation of the surface layers of the earth, landslides may occur [38]. Drainage density layer was divided into five classes of 0–0.076, 0.077–0.14, 0.15–0.2, 0.21–0.26,

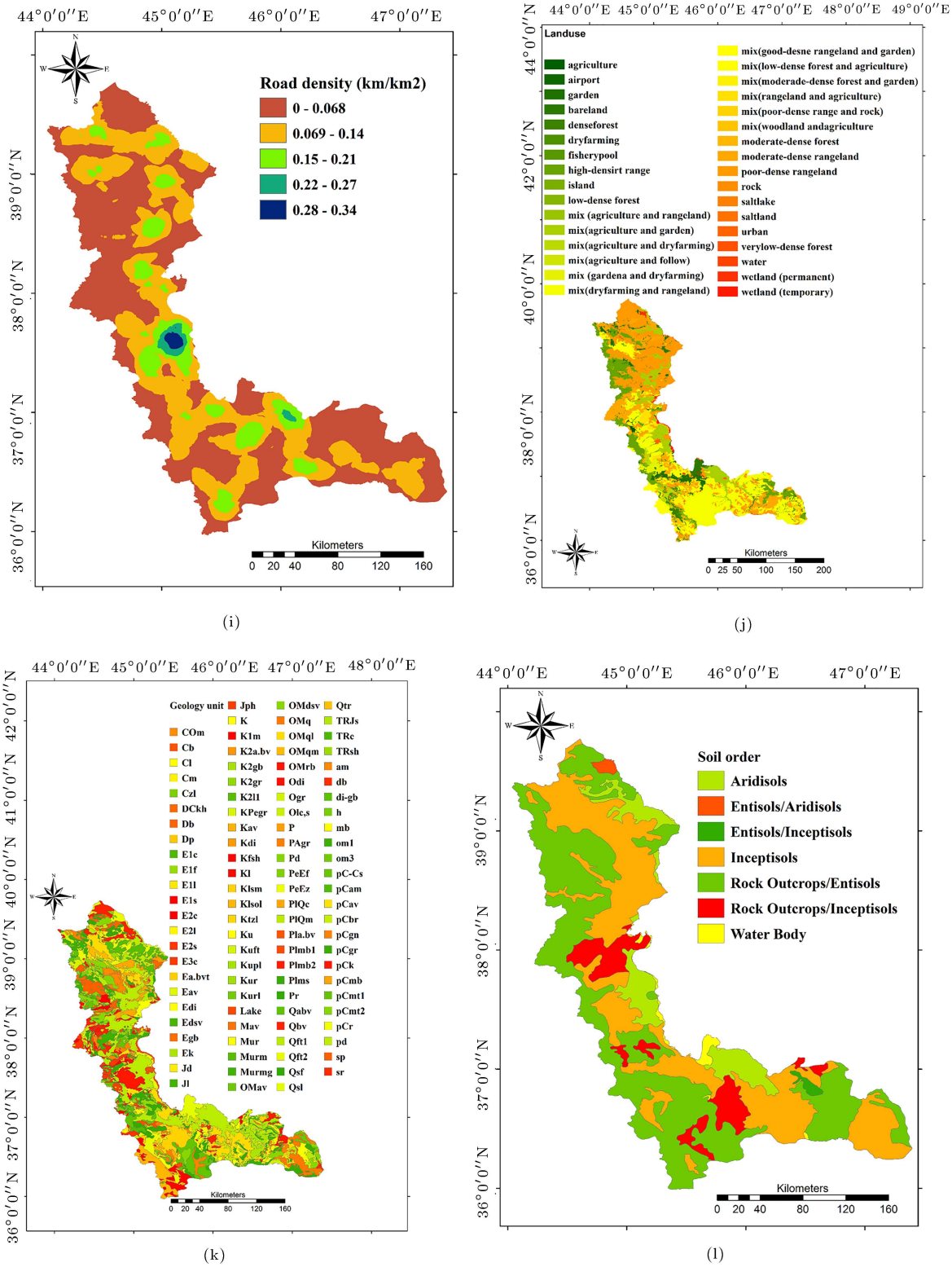

Figure 2. Landslide conditioning factors (continued).

and 0.27–0.38 (Figure 2(h)). An increase in the road density is indicative of the higher influence of human activities on the location as well as a rise in the possibility of disturbing natural conditions and equilibrium of the slopes. In addition, soil adhesion and internal friction decreased, thus decreasing the reliability factor of the hillside, especially in mountainous areas. The road layer was taken from the Open Street Map (OSM) website. Road density layer was divided into five classes of 0–0.086, 0.069–0.14, 0.15–0.21, 0.22–0.27,

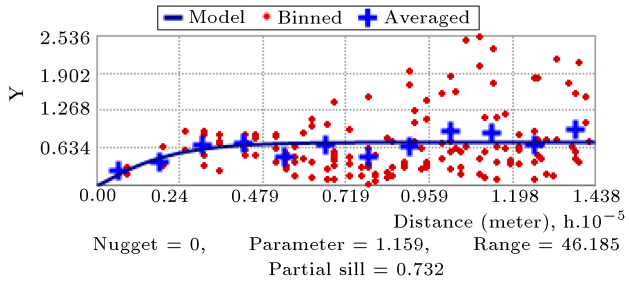

Figure 3. Toxic results of variogram.

and 0.28–0.34 (Figure 2(i)). Land use and vegetation play an important role in hillside stability, and several studies have emphasized their significance in landslide risk assessment. The land use map of the area under study was prepared using images from Landsat 8. The maximum likelihood supervised classification algorithm in ENVI 4.5 software was employed to classify surface coverage, and 33 types of land use were identified in the study area (Figure 2(j)). Since the resistance of different geological units against landslide varies, the properties of the constituents of a domain such as resistance and relative permeability play an important role in their stability. Therefore, lithology is an important and influential factor in the likelihood of landslides. The lithology map was prepared on a scale of 1:100,000 (Figure 2(k)). The soil map was prepared based on the information from the west Azerbaijan natural resources organization.

2.4. FR model

In the Frequency Ratio (FR) model, the set of landslides is introduced as the dependent variables, and the parameters affecting landslide are presented as the independent variables [39]. FR model calculates the probability of a phenomenon with specific characteristics. In this regard, the probability of the occurrence of landslides in each class for all parameters is computed through this method. To determine the effect of each class on each independent variable, Eq. (1) is used:

$$FR = \frac{Fi}{P_i}, \tag{1}$$

where FR is the impact of each class on each parameter, F_i the percentage of the points of landslide in class i, and P_i the percentage of the pixels of class i in the entire study area.

2.5. SE model

Entropy is a measure of disorder, instability, behavioral imbalance, energy distribution, and uncertainty in a system [40]. It determines how the most important factors can be estimated based on the effective factors of a goal and specifies the variables with the highest impact on an event [41]. Therefore, this theory, as a managerial approach, can play an important role in identifying the effective factors and their impacts [40]. Eqs. (2)–(6) are employed to evaluate the criteria using the entropy model:

$$E_{ij} = \frac{FR}{\sum_{j=1}^{M} FR}, \tag{2}$$

$$H_j = -\sum_{i=1}^{s} (E_{ij}) log_2(E_{ij}), \tag{3}$$

$$H_{j\,\max} = log_2 M_j, \tag{4}$$

$$I_j = \frac{H_{j\,\max} - H_j}{H_{j\,\max}}, \tag{5}$$

$$W_j = I_j * FR, \tag{6}$$

where FR is the frequency ratio, E_{ij} is density probability, H_j and $H_{j\,\max}$ are entropy value and maximum entropy, respectively, I_j is information factor, M_j is the number of classes, and W_j is the final weight of each criterion.

2.6. RF model

RF is a modern type of tree-based methods that incorporates a variety of classification and regression trees. Since it is nonparametric, it is suitable for modeling continuous and discrete decision tree data [42]. It is constructed using a set of trees with n independent observational data. This method is a combination of several decision trees in which different Bootstrap instances of data are involved. Moreover, in random construction of each tree, many input variables are involved. Through the Bootstrap method, a large number of n samples from the initial observation dataset are generated. During the sampling process, almost one-third of the data are not sampled, hence regarded as out-of-process sampling. Followed by constructing the whole tree, the test data are introduced to the tree, and the number of trees is calculated for each input. Finally, the final output is calculated by averaging the outputs [39].

2.7. Bagging model

Bagging is one of the simplest, yet most successful, group methods for improving the classification problem [43]. This method is specifically beneficial for dealing with high-volume and high-dimensional data since, in such situations, it is not possible to achieve a model in one step due to the high complexity of the problem [18]. Bagging was first introduced in [44] to reduce the variance of a predictor. In this method, several copies with equal volumes to the initial training dataset are randomly selected. Given that sampling is done by replacement, some data may not appear multiple times

or even in one training sequence. Each of the training sequences is used to train a weak cluster and build a model. The output of the models is combined using bagging technique to achieve the final output.

2.8. Validation

Receiver Operating Characteristic (ROC) curve is used to evaluate the models. It consists of two axes, namely sensitivity (x-axis) and transparency (y-axis) [45]. The x and y axes of the ROC curve are calculated using Eqs. (7) and (8), which are obtained from the confusion matrix by defining the threshold limit between zero and one.

$$X = 1 - \left[\frac{TN}{TN + FP}\right], \tag{7}$$

$$Y = \left[\frac{TP}{TP + FN}\right]. \tag{8}$$

The area under the ROC curve, called Area Under the ROC Curve (AUC), represents the value of the prediction of a system by describing its ability to accurately predict the occurrence of an event (landslide) and its non-occurrence (no landslide).

3. Result

3.1. Results of FR model

The results from the FR model for each of the classes of effective factors are shown in Table 1. According to these results, the middle slopes have the greatest impact on landslide occurrence. In low slopes, due to small gravity, the probability of landslide occurrence is reduced and the steep slopes are the mountainous areas covered with rocks and fragile soil, which is not a suitable condition for the occurrence of landslides. According to the results related to the slope, the south direction with a weight of 1.49 has the highest impact on landslide occurrence and the lowest weight is for the flat areas, where no landslides have occurred. In terms of altitude, more landslides often occur in the middle classes, as pointed out by other researchers [39]. In the study area, as the altitude increases, the probability of landslide occurrence decreases. The results of the rainfall factor show that overall, with an increase in rainfall amounts, the likelihood of landslides increases, which is consistent with the results from other studies (e.g. [43]). According to the results related to distance from fault, the class of 0–500 m with a weight of 5.07 has the greatest impact on landslide occurrence in the study area. In the case of distance from the river, the 0–500 m class has the highest influence on landslide occurrence mainly because permanent rivers are the main source of moisture for landslide occurrence. Although the role of non-standard road construction in landslide events has been confirmed by other researchers [46], it did not have a significant impact in the area under study. The results for drainage density show that the class of 0.19–0.34 has the most significant impact on landslide occurrence ($FR = 3.10$). In terms of road density, the class of 0.07–0.13 has the greatest effect on landslide occurrence in the study area ($FR = 1.28$). Lithology criterion is characterized by the highest FR value for Eva unit ($FR = 23.93$). It has been proven that upon increasing the density and amount of vegetation, the probability of landslide occurrence would decrease due to the role of the roots of plants in preventing this phenomenon. The results related to land use in this study also indicate that low forest areas are most prone to the occurrence of landslides ($FR = 9.28$). Finally, the results of soil criterion show that Entisols/Inceptisols have the greatest impact on landslide occurrence ($FR = 12.33$). The weights obtained by the FR model were applied to the effective criteria. Figure 4 illustrates the landslide susceptibility map.

3.2. Result of SE model

The results achieved by the SE model are summarized in Table 1. The order of the effective factors in landslide incidents is as follows: soil (1.33), land use (0.41), distance from fault (0.36), river density (0.26), road density (0.2), distance from river (0.17), altitude (0.13), lithology (0.13), rainfall (0.11), slope angle (0.08), slope aspect (0.07), and distance from river (0.05). The landslide susceptibility map prepared by applying the weights from the SE model to the effective criteria is presented in Figure 5.

3.3. Results of RF and RF-BA models

To prepare a landslide sensitivity map using the combined data mining models, the combination of bagging model with RF model in WEKA data mining software was used. To this end, the weights of the FR model were used to input the hybrid models. To implement the data mining models, a spatial database provided the points of presence of landslide (Value 1) and absence of landslide (Value 0). Non-occurrence points were randomly created as equal to the training and validation points. Here, 70% of the data were used as training points and 30% as evaluation points. The performance of the hybrid data mining algorithms depends on the optimal selection of the parameters used by this algorithm. This requires performing modeling in many different iterations based on the training data and number of the seeds (to divide the training data). Therefore, by changing each of these parameters, modeling is carried out again with the new conditions and the results will change. Since the two parameters in this study depended on each other, a constant number was assumed for iterations. Then, the number of seeds varied and the percentage for

Table 1. Spatial relationship between each effective factor and landslide locations based on the FR and SE models.

Class	No. of pixels in domain	No. of landslides	FR	E_{ij}	H_j	$H_{j\,\max}$	I_j	W_j
Altitude (m)					1.94	2.32	0.16	0.13
605–1179	358553	3	0.40	0.10				
1179–1562	1225254	24	0.95	0.24				
1562–1942	1006641	38	1.82	0.45				
1942–2381	724626	11	0.73	0.18				
2381–3600	406076	1	0.12	0.03				
Slope angle					2.14	2.32	0.08	0.08
0–4.5	1348663	12	0.43	0.08				
4.5–10.5	940113	23	1.18	0.23				
10.5–17.17	757493	27	1.72	0.34				
17.17–25.1	478448	13	1.31	0.26				
25.1–61.7	196433	2	0.49	0.10				
Slope aspect					2.92	3.17	0.08	0.07
Flat	281	0	0.00	0.00				
North	460586	11	1.15	0.14				
North East	488194	7	0.69	0.09				
East	470580	6	0.62	0.08				
South East	454555	14	1.49	0.19				
South	462122	5	0.52	0.07				
South West	473843	10	1.02	0.13				
West	460561	13	1.36	0.17				
North West	450427	11	1.18	0.15				
Rainfall (m)					2.07	2.32	0.11	0.11
308–383	555800	3	0.26	0.05				
383–443	1132762	12	0.51	0.10				
443–543	1215138	37	1.48	0.29				
543–700	425717	16	1.82	0.35				
700–900	407377	9	1.07	0.21				
Distance from fault					2.13	2.58	0.18	0.36
0–500	162718	17	5.07	0.41				
500–1000	153675	9	2.84	0.23				
1000–2000	274248	13	2.30	0.19				
2000–5000	665573	9	0.66	0.05				
5000–10000	875384	20	1.11	0.09				
10000–50000	1605196	9	0.27	0.02				
Distance from river					2.11	2.58	0.18	0.17
0–500	669330	26	26	0.34				
500–1000	598882	20	20	0.30				
1000–2000	935863	18	18	0.17				
2000–5000	1116796	8	8	0.06				
5000–10000	353574	5	5	0.13				
10000–50000	62345	0	0	0.00				

Table 1. Spatial relationship between each effective factor and landslide locations based on the FR and SE models (continued).

Class	No. of pixels in domain	No. of landslides	FR	E_{ij}	H_j	$H_{j\,\max}$	I_j	W_j
Distance from road					2.44	2.58	0.06	0.05
0–500	263840	2	0.37	0.07				
500–1000	244182	5	0.99	0.18				
1000–2000	437796	10	1.11	0.20				
2000–5000	1004482	22	1.06	0.19				
5000–10000	1006340	31	1.49	0.27				
10000–50000	780150	7	0.44	0.08				
Drainage density					1.78	2.32	0.23	0.26
0–0.03	478741	4	0.41	0.07				
0.03–0.07	810623	8	0.48	0.09				
0.07–0.12	926513	7	0.37	0.07				
0.12–0.19	988682	24	1.18	0.21				
0.19–0.34	532223	34	3.10	0.56				
Road density					1.47	2.32	0.37	0.2
0–0.07	2025379	39	0.93	0.35				
0.07–0.13	1329621	35	1.28	0.48				
0.13–0.2	327766	3	0.44	0.17				
0.2–0.27	37710	0	0.00	0.00				
0.27–0.34	16318	0	0.00	0.00				
Lithology					3.4	6.64	0.49	0.13
am	41533	3	3.45	0.04				
Cb	33661	1	1.42	0.02				
Cl	18377	2	5.20	0.06				
Cm	13230	0	0.00	0.00				
COm	15779	0	0.00	0.00				
Czl	8129	0	0.00	0.00				
db	10917	0	0.00	0.00				
Db	4655	0	0.00	0.00				
DCkh	28756	0	0.00	0.00				
di-gb	13877	0	0.00	0.00				
Dp	497	0	0.00	0.00				
E1c	8324	0	0.00	0.00				
E1f	2425	0	0.00	0.00				
E1l	8348	0	0.00	0.00				
E1s	27632	0	0.00	0.00				
E2c	15940	0	0.00	0.00				
E2l	26756	0	0.00	0.00				
E2s	37308	0	0.00	0.00				
E3c	796	0	0.00	0.00				
Ea.bvt	6077	0	0.00	0.00				
Eav	5996	3	23.93	0.26				
Edi	3385	0	0.00	0.00				
Edsv	2197	0	0.00	0.00				
Egb	1828	0	0.00	0.00				
Ek	53801	9	8.00	0.09				

Table 1. Spatial relationship between each effective factor and landslide locations based on the FR and SE models (continued).

Class	No. of pixels in domain	No. of landslides	FR	E_{ij}	H_j	$H_{j\,\max}$	I_j	W_j
Lithology					3.4	6.64	0.49	0.13
h	8911	0	0.00	0.00				
Jd	177	0	0.00	0.00				
Jl	10266	5	23.29	0.25				
Jph	22777	0	0.00	0.00				
K	5824	0	0.00	0.00				
K1m	3913	0	0.00	0.00				
K2a.bv	3809	0	0.00	0.00				
K2gb	771	0	0.00	0.00				
K2gr	22989	0	0.00	0.00				
K2l1	11985	0	0.00	0.00				
Kav	22416	0	0.00	0.00				
Kdi	1423	0	0.00	0.00				
Kfsh	67657	2	1.41	0.02				
Kl	739	0	0.00	0.00				
Klsm	113364	9	3.80	0.04				
Klsol	89896	3	1.60	0.02				
KPegr	6621	0	0.00	0.00				
Ktzl	11601	0	0.00	0.00				
Ku	37252	0	0.00	0.00				
Kuft	34076	0	0.00	0.00				
Kupl	21822	0	0.00	0.00				
Kur	1570	0	0.00	0.00				
Kurl	17131	0	0.00	0.00				
Lake	22405	0	0.00	0.00				
Mav	9640	0	0.00	0.00				
mb	9788	0	0.00	0.00				
Mur	46757	0	0.00	0.00				
Murm	72510	4	2.64	0.03				
Murmg	2601	0	0.00	0.00				
Odi	813	0	0.00	0.00				
Ogr	7044	0	0.00	0.00				
Olc,s	583	0	0.00	0.00				
om1	111695	0	0.00	0.00				
om3	3805	0	0.00	0.00				
OMav	18437	0	0.00	0.00				
OMdsv	3797	0	0.00	0.00				
OMq	71732	3	2.00	0.02				
OMql	269830	9	1.60	0.02				
OMqm	18873	0	0.00	0.00				
OMrb	97679	0	0.00	0.00				
P	17164	0	0.00	0.00				
PAgr	21779	0	0.00	0.00				
pC-Cs	966	0	0.00	0.00				
pCam	10171	0	0.00	0.00				
pCav	297	0	0.00	0.00				
pCbr	6675	0	0.00	0.00				
pCgn	92759	2	1.03	0.01				

Table 1. Spatial relationship between each effective factor and landslide locations based on the FR and SE models (continued).

Class	No. of pixels in domain	No. of landslides	FR	E_{ij}	H_j	$H_{j\,\max}$	I_j	W_j
Lithology					3.4	6.64	0.49	0.13
pCgr	45744	1	1.05	0.01				
pCk	38036	1	1.26	0.01				
pCmb	1000	0	0.00	0.00				
pCmt1	153162	4	1.25	0.01				
pCmt2	123508	1	0.39	0.00				
pCr	28120	0	0.00	0.00				
pd	3829	0	0.00	0.00				
Pd	35632	0	0.00	0.00				
PeEf	7927	0	0.00	0.00				
PeEz	1679	0	0.00	0.00				
Pla.bv	2521	0	0.00	0.00				
Plmb1	5852	0	0.00	0.00				
Plmb2	9328	0	0.00	0.00				
Plms	31082	1	1.54	0.02				
PlQc	62974	0	0.00	0.00				
PlQm	15018	0	0.00	0.00				
Pr	192755	0	0.00	0.00				
Qabv	9981	0	0.00	0.00				
Qbv	125207	1	0.38	0.00				
Qft1	613205	8	0.62	0.01				
Qft2	158016	2	0.61	0.01				
Qsf	8842	0	0.00	0.00				
Qsl	71960	0	0.00	0.00				
Qtr	2919	0	0.00	0.00				
sp	68322	0	0.00	0.00				
sr	2698	0	0.00	0.00				
TRe	7594	0	0.00	0.00				
TRJs	22752	3	6.31	0.07				
TRsh	7609	0	0.00	0.00				
Land use/ cover					2.41	5.04	0.52	0.41
agri	12170	0	0.00	0.00				
airport	217	0	0.00	0.00				
bagh	159115	6	1.83	0.07				
bareland	42267	0	0.00	0.00				
denseforest	16477	3	8.81	0.34				
dryfarming	40261	0	0.00	0.00				
fisherypool	2032	0	0.00	0.00				
goodrange	363469	2	0.27	0.01				
island	7	0	0.00	0.00				
lowforest	15644	3	9.28	0.36				
mix(agri_bagh)	308717	0	0.00	0.00				
mix(agri_dryfarming)	50912	1	0.95	0.04				
mix(agri_follow)	1643	0	0.00	0.00				
mix(agri_X)	5282	0	0.00	0.00				
mix(bagh_X)	50874	0	0.00	0.00				
mix(dryfarming_x)	926712	37	1.93	0.07				

Table 1. Spatial relationship between each effective factor and landslide locations based on the FR and SE models (continued).

Class	No. of pixels in domain	No. of landslides	FR	E_{ij}	H_j	$H_{j\,max}$	I_j	W_j
Land use/ cover					2.41	5.04	0.52	0.41
mix(goodrang_x)	166383	2	0.58	0.02				
mix(lowforest_x)	34965	0	0.00	0.00				
mix(modforest_x)	17490	0	0.00	0.00				
mix(modrange_x)	73449	1	0.66	0.03				
mix(poorrange_x)	17775	0	0.00	0.00				
mix(woodland_x)	8660	0	0.00	0.00				
modforest	15128	0	0.00	0.00				
modrange	448181	6	0.65	0.03				
poorrange	895493	16	0.86	0.03				
rock	1156	0	0.00	0.00				
saltlake	1333	0	0.00	0.00				
saltland	360	0	0.00	0.00				
urban	19632	0	0.00	0.00				
verylowforest	1192	0	0.00	0.00				
water	8044	0	0.00	0.00				
wetland1	9097	0	0.00	0.00				
wetland2	12555	0	0.00	0.00				
Soil					1.18	2.81	0.58	1.33
Rock outcrops/entisols	1632502	37	1.09	0.07				
Rock outcrops/inceptisols	307784	11	1.72	0.11				
Water body	43606	0	0.00	0.00				
Aridisols	322585	1	0.15	0.01				
Entisols/aridisols	18265	0	0.00	0.00				
Entisols/inceptisols	19576	5	12.33	0.77				
Inceptisols	1371498	23	0.81	0.05				

the Area Under the Receiver Operating Characteristic (AUROC) was recorded. For the iteration parameter and seed parameter, numbers from 10 to 20 and from 1 to 10 were used, respectively. The results of RF and RF-BA modeling are shown in Table 2. According to Figure 4, the sensitivity, transparency, accuracy, and ROC values of the hybrid model (RF-BA) are higher than those of the single model (RF), indicating the better performance of the hybrid model than that of the other. After training the hybrid models, the modeling process was generalized to the whole study area and then, to the landslide susceptibility map using RF-BA and RF models in ArcGIS 10.3 software. Upon use of the natural breaks classification method, the landslide susceptibility map was divided into five categories: very low susceptibility, low susceptibility, medium susceptibility, high susceptibility, and very high susceptibility. The landslide susceptibility maps prepared by RF and RF-BA modeling are shown in Figures 6 and 7, respectively.

Table 2. Result of the RF and RF-BA models.

Criterion	RF-BA		RF	
	Train	**Test**	**Train**	**Test**
TP	75	30	73	28
TN	76	31	74	29
FP	2	3	4	6
FN	1	2	3	5
Sensitivity	0.962	0.797	0.948	0.778
Transparency	1	0.933	0.906	0.917
Accuracy	0.981	0.897	0.926	0.833
AUC	0.993	0.923	0.984	0.897

3.4. Validation of the results

Based on the success rate and prediction rate methods, the results of the landslide susceptibility maps were evaluated by comparing them with the existing landslides [8]. Success rate results were obtained based

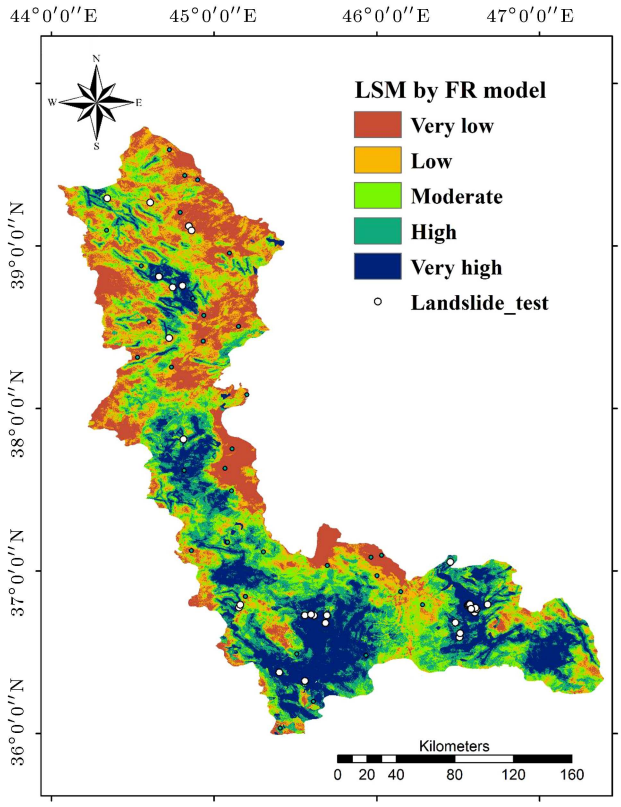

Figure 4. Landslide susceptibility map prepared by the FR model.

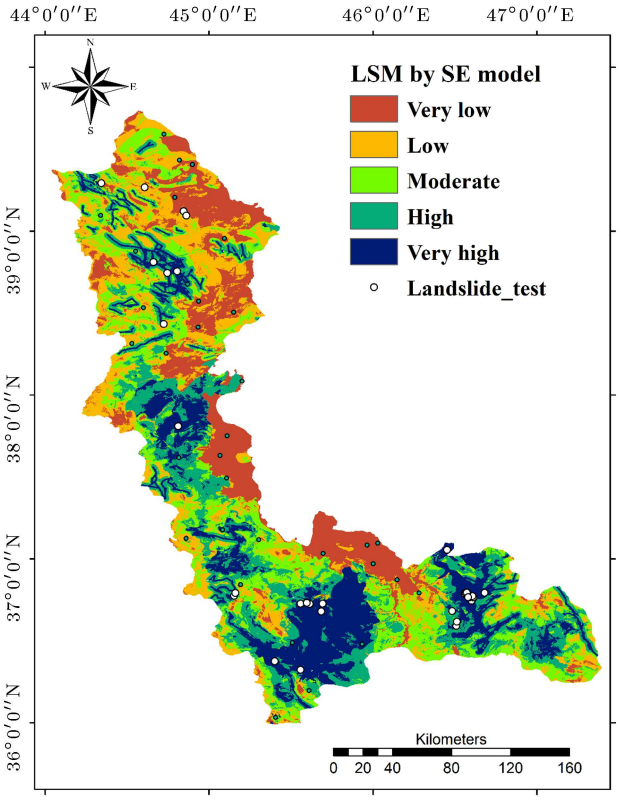

Figure 5. Landslide susceptibility map prepared by the SE model.

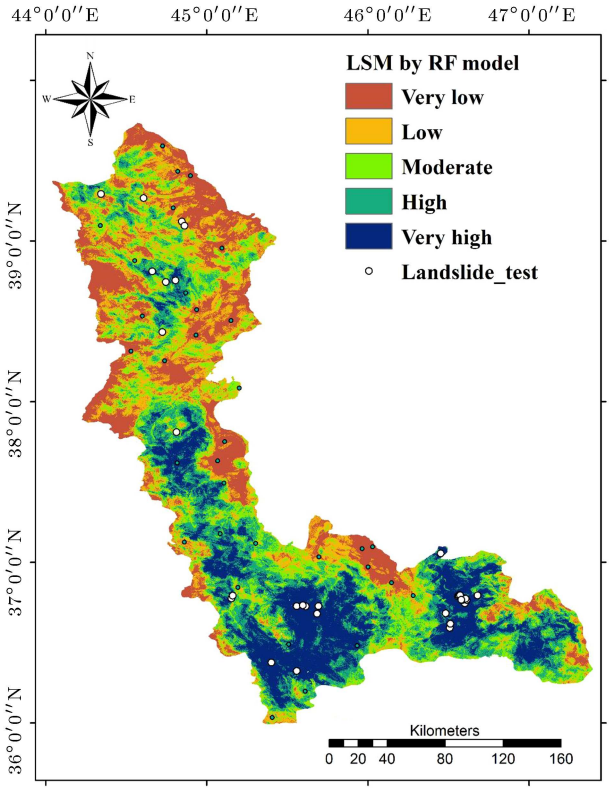

Figure 6. Landslide susceptibility map prepared by the RF model.

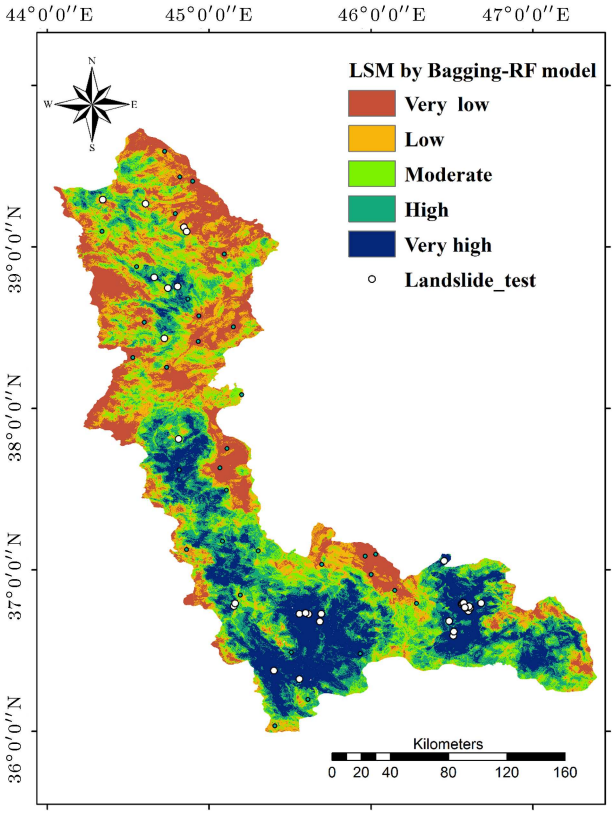

Figure 7. Landslide susceptibility map prepared by the RF-BA model.

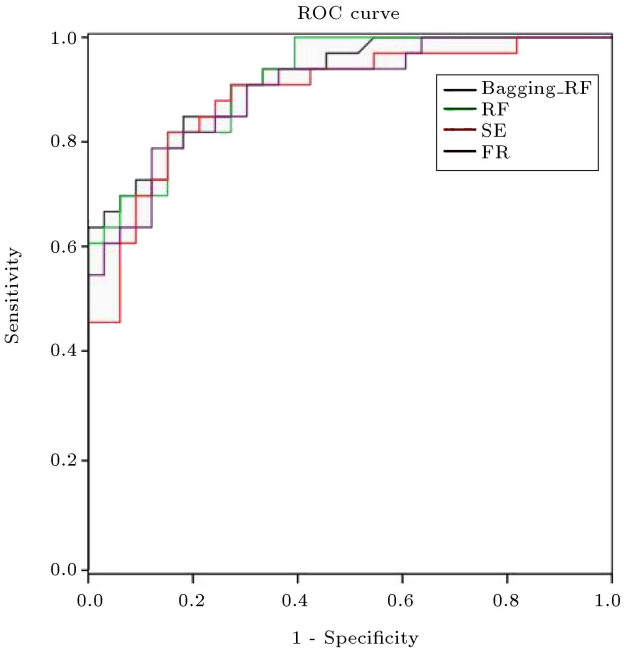

Figure 8. Results for the ROC curve.

on the training data, and the prediction rates were obtained from validation datasets. The areas under the prediction rate curve are 0.88, 0.89, 0.91, and 0.92 for FR, SE, RF, and RF-BA models, respectively (Figure 8). The results indicate that the hybrid data mining model is more efficient than other models, even the single data mining model. In addition, data mining models, in general, are characterized by higher accuracy than bivariate statistical models. The findings of the present study are in agreement with the results obtained in [12,47].

4. Discussion

Landslides have complex mechanisms and many factors are involved in their occurrence. Identification of the locations with the highest potential for this type of mass movement helps managers and planners with land use planning and land development programs. To this end, in this study, 12 parameters were first identified as the primary factors affecting landslides. Then, they were analyzed and used for landslide susceptibility mapping through FR, SE, RF, and RF-BA models. The results from the ROC curve indicated that RF-BA had the highest accuracy of modeling in landslide susceptibility mapping, followed by RF, SE, and FR. In addition, it was observed that data mining models were more accurate than statistical methods. The fundamental difference between the statistical and data mining methods lies in their assumptions or the nature of data being processed [45]. Generally, it is assumed in statistical techniques that data distribution is clear and normal; further, the accuracy or inaccuracy of the final results depends on the accuracy of the initial assumption. On the contrary, data mining methods do not make any assumptions about the data. Data mining methods operate much better when the data is incomplete or contradictory because in such cases, the missing data is somehow retrieved based on the pattern of the data. On the other hand, in statistical methods, the problem of the lack of some data yields an incomplete outcome. While statistical methods are not capable of detecting complex nonlinear patterns, data mining methods due to their exploratory properties begin to model behavior of the data without any initial assumptions, hence adaptable over time [45]. Through data mining, the algorithm becomes more and more vibrant. In addition, the nonlinear, robust structure of these models facilitates simulating the behavior of both social and real environments. The obtained results in this study revealed that the RF-BA model was more accurate than the RF model alone, mainly because in the bagging method, clusters from different sets of initial datasets guaranteed diversity condition and the accuracy increased by averaging sample classifications of different datasets. The findings of the present study are in agreement with the results obtained in [34,48]. Furthermore, according to the results, the entropy model exhibited higher accuracy than the FR model since it employed the results of the FR model. The results from the present study are consistent with those obtained by Hong et al. [49].

5. Conclusion

The main objective of this study was to compare Random Forest (RF) and RF-BA data mining models with Frequency Ratio (FR) and SE bivariate F-statistical models for landslide susceptibility mapping. The following results were achieved:

1. According to the Receiver Operating Characteristic (ROC) curve, the RF-BA, RF, FR, and SE models obtained the accuracy rates of 0.92, 0.91, 0.89, and 0.88, respectively;
2. According to the ROC curve and AUC, RF and RF-BA data mining models exhibited higher accuracy than that of statistical methods in landslide susceptibility mapping;
3. Hybrid models were more accurate than single models in dual data mining and statistical models;
4. Based on the SE model, soil, land use, and distance from fault were the most important factors in landslide occurrence in the study area;
5. According to the results of the FR model, the probability of landslide occurrence in the study area was higher with altitudes between 1562–1942 m,

slope angles between 10.5°–17.7°, south slope aspect, distances from fault less than 500 m, distances from river less than 500 m, distances from road between 5000 – 10000 m, rainfalls between 543 – 700 mm, drainage densities between 0.19 – 0.34, road densities between 0.07 – 0.13, Eva lithology, low forest land use, and Entisols/Inceptisols soil order.

References

1. Keesstra, S.D., Quinton, J.N., van der Putten, W.H., et al. "The significance of soils and soil science towards realization of the United Nations", *SOIL*, **2**(3), pp. 111–128 (2016). https://doi.org/10.5194/soil-2-111-2016
2. Bordoni, M., Meisina, C., Valentino, R., et al. "Site-specific to local-scale shallow landslides triggering zones assessment using TRIGRS", *Nat Hazards Earth Syst. Sci.*, **15**(5), pp. 1025–1050 (2015).
3. Tsangaratos, P. and Benardos, A. "Estimating landslide susceptibility through a artificial neural network classifier", *Natural Hazards*, **74**(3), pp. 1489–1516 (2014).
4. Dilley, M., Chen, R.S., Deichmann, U., et al., *Natural Disaster Hotspots: a Global Risk Analysis*, The World Bank Hazard Management Unit", Washington (2005).
5. Lee, S. and Choi, J. "Landslide susceptibility mapping using GIS and the weight-of-evidence model", *International Journal of Geographical Information Science*, **18**(8), pp. 789–814 (2004).
6. Li, Z., He, Y., Li, H., et al. "Antecedent rainfall induced shallow landslide-A case study of Yunnan landslide, China", *Scientia Iranica*, **26**(1), pp. 202–212 (2019).
7. Lee, S. and Sambath, T. "Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models", *Environmental Geology*, **50**(6), pp. 847–855 (2006).
8. Pak, A. and Sarfaraz, M. "Lattice boltzmann method for simulating impulsive water waves generated by landslides", *Scientia Iranica*, **21**(2), pp. 318–328 (2014).
9. Pandey, V.K., Pourghasemi, H.R., and Sharma, M.C. "Landslide susceptibility mapping using maximum entropy and support vector machine models along the Highway Corridor, Garhwal Himalaya", *Geocarto International*, **35**(2), pp. 168–187 (2020).
10. Yilmaz, I. "Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat-Turkey)", *Computers & Geosciences*, **35**(6), pp. 1125–1138 (2009).
11. Gadtaula, A. and Dhakal, S. "Landslide susceptibility mapping using Weight of Evidence Method in Haku, Rasuwa District, Nepal", *Journal of Nepal Geological Society*, **58**, pp. 163–171 (2019).
12. Chen, Z., Liang, S., Ke, Y., et al. "Landslide susceptibility assessment using evidential belief function, certainty factor and frequency ratio model at Baxie River basin, NW China", *Geocarto International*, **34**(4), pp. 348–367 (2019).
13. Devkota, K.C., Regmi, A.D., Pourghasemi, H.R., et al. "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", *Natural Hazards*, **65**(1), pp. 135–165 (2013).
14. Nohani, E., Moharrami, M., Sharafi, S., et al. "Landslide susceptibility mapping using different GIS-based bivariate models", *Water*, **11**(7), p. 1402 (2019).
15. Pascale, S., Parisi, S., Mancini, A., et al. "Landslide susceptibility mapping using artificial neural network in the urban area of senise and san costantino albanese (basilicata, southern Italy)", In *International Conference on Computational Science and Its Applications*, Springer, Berlin, Heidelberg, pp. 473–488 (2013).
16. Aghdam, I.N., Varzandeh, M.H.M., and Pradhan, B. "Landslide susceptibility mapping using an ensemble statistical index (Wi) and adaptive neuro-fuzzy inference system (ANFIS) model at Alborz Mountains (Iran)", *Environmental Earth Sciences*, **75**(7), p. 553 (2016).
17. Pourghasemi, H.R., Jirandeh, A.G., Pradhan, B., et al. "Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran", *Journal of Earth System Science*, **122**(2), pp. 349–369 (2013).
18. Hong, H., Kornejady, A., Soltani, A., et al. "Landslide susceptibility assessment in the Anfu County, China: comparing different statistical and probabilistic models considering the new topo-hydrological factor (HAND)", *Earth Science Informatics*, **11**(4), pp. 605–622 (2018).
19. Kim, J.C., Lee, S., Jung, H.S., et al. "Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea", *Geocarto International*, **33**(9), pp. 1000–1015 (2018).
20. Oliveira, A., Fernandes, J., Bateira, C., et al. "Influence of digital elevation MODELS ON landslide susceptibility with logistic regression model", *Revista do Departamento de Geografia*, **36**, pp. 33–47 (2018).
21. Park, S., Choi, C., Kim, B., et al. "Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea", *Environmental Earth Sciences*, **68**(5), pp. 1443–1464 (2013).
22. Dano, U.L., Balogun, A.L., Matori, A.N., et al. "Flood susceptibility mapping using GIS-based analytic network process: A case study of Perlis, Malaysia", *Water*, **11**(3), p. 615 (2019).

23. Chen, W., Hong, H., Panahi, M., et al. "Spatial prediction of landslide susceptibility using GIS-based data mining techniques of anfis with whale optimization algorithm (WOA) and grey wolf optimizer (GWO)", *Applied Sciences*, **9**(18), p. 3755 (2019).

24. Huang, F., Zhang, J., Zhou, C., et al. "A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction", *Landslides*, **17**(1), pp. 217–229 (2020).

25. Nguyen, V.V., Pham, B.T., Vu, B.T., et al. "Hybrid machine learning approaches for landslide susceptibility modeling", *Forests*, **10**(2), p. 157 (2019).

26. Pham, B.T., Prakash, I., Dou, J., et al. "A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers", *Geocarto International*, **35**(12), pp. 1267–1292 (2020).

27. Bui, D.T., Hoang, N.D., Nguyen, H., et al. "Spatial prediction of shallow landslide using Bat algorithm optimized machine learning approach: A case study in Lang Son Province, Vietnam", *Advanced Engineering Informatics*, **42**, p. 100978 (2019).

28. Tien Bui, D., Shahabi, H., Omidvar, E., et al. "Shallow landslide prediction using a novel hybrid functional machine learning algorithm", *Remote Sensing*, **11**(8), p. 931 (2019).

29. Chen, W., Shahabi, H., Zhang, S., et al. "Landslide susceptibility modeling based on GIS and novel bagging-based kernel logistic regression", *Applied Sciences*, **8**(12), p. 2540 (2018).

30. Shirzadi, A., Soliamani, K., Habibnejhad, M., et al. "Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping", *Sensors*, **18**(11), p. 3777 (2018).

31. Pham, B.T., Khosravi, K., and Prakash, I. "Application and comparison of decision tree-based machine learning methods in landside susceptibility assessment at Pauri Garhwal Area, Uttarakhand, India", *Environmental Processes*, **4**(3), pp. 711–730 (2017).

32. Stocking, M.A. and Murnaghan, N., *A Handbook for the Field Assessment of Land Degradation*, 1st Ed., Routledge (2002). https://doi.org/10.4324/9781849776219

33. Hong, H., Liu, J., and Zhu, A.X. "Modeling landslide susceptibility using LogitBoost alternating decision trees and forest by penalizing attributes with the bagging ensemble", *Science of the Total Environment*, **718**, p. 137231 (2020).

34. Wu, Y., Ke, Y., Chen, Z., et al. "Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping", *Catena*, **187**, p. 104396 (2020).

35. He, S., Pan, P., Dai, L., et al. "Application of kernel-based fisher discriminant analysis to map landslide susceptibility in the Qinggan River delta, Three Gorges, China", *Geomorphology*, **171**, pp. 30–41 (2012).

36. Xu, C., Dai, F., Xu, X., et al. "GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China", *Geomorphology*, **145**, pp. 70–80 (2012).

37. Li, X., Yang, H., Zhang, J., et al. "Time-domain analysis of tamper displacement during dynamic compaction based on automatic control. coatings", **11**(9), 1092 (2021). DOI: 10.3390/coatings11091092

38. Yang, W., Chen, X., Xiong, Z., et al. "A privacy-preserving aggregation scheme based on negative survey for vehicle fuel consumption data", *Information Sciences*, **570**, pp. 526–544 (2021). DOI: 10.1016/j.ins.2021.05.009

39. Li, B., Feng, Y., Xiong, Z., et al. "Research on AI security enhanced encryption algorithm of autonomous IoT systems", *Information Sciences*, **575**, pp. 379–398 (2021). DOI: 10.1016/j.ins.2021.06.016

40. Li, B., Yang, J., Yang, Y., et al. "Sign language/gesture recognition based on cumulative distribution density features using UWB radar", *IEEE Transactions on Instrumentation and Measurement*, **70**, pp. 1–13 (2021). DOI: 10.1109/TIM.2021.3092072

41. Weng, L., He, Y., Peng, J., et al. "Deep cascading network architecture for robust automatic modulation classification", *Neurocomputing (Amsterdam)*, **455**, pp. 308–324 (2021). DOI: 10.1016/j.neucom.2021.05.010

42. Chao, L., Zhang, K., Wang, J., et al. "A comprehensive evaluation of five evapotranspiration datasets based on ground and GRACE satellite observations: Implications for improvement of evapotranspiration retrieval algorithm", *Remote Sensing* (Basel, Switzerland), **13**(12), p. 2414 (2021). DOI: 10.3390/rs13122414

43. Zhang, K., Wang, S., Bao, H., et al. "Characteristics and influencing factors of rainfall-induced landslide and debris flow hazards in Shaanxi Province, China", *Natural Hazards and Earth System Sciences*, **19**(1), pp. 93–105 (2019). DOI: 10.5194/nhess-19-93-2019

44. Zuo, Y., Jiang, S., Wu, S., et al. "Terrestrial heat flow and lithospheric thermal structure in the Chagan Depression of the Yingen-Ejinaqi Basin, north central China", *Basin Research*, **32**(6), pp. 1328–1346 (2020). DOI: 10.1111/bre.12430

45. Xu, J., Wu, Z., Chen, H., et al. "Study on strength behavior of basalt fiber-reinforced loess by digital image technology (DIT) and Scanning Electron Microscope (SEM)", *Arabian Journal for Science and Engineering* (2021). DOI: 10.1007/s13369-021-05787-1

46. Zhang, K., Chao, L., Wang, Q., et al. "Using multi-satellite microwave remote sensing observations for retrieval of daily surface soil moisture across China", *Water Science and Engineering*, **12**(2), pp. 85–97 (2019). DOI: 10.1016/j.wse.2019.06.001

47. Kordestani, H., Zhang, C., Masri, S.F., et al. "An empirical time-domain trend line-based bridge signal decomposing algorithm using Savitzky-Golay filter", *Structural Control and Health Monitoring*, **28**(7), n/a-n/a (2012). DOI: 10.1002/stc.2750

48. Zhou, W., Liu, J., Lei, J., et al. "GMNet: Graded-feature multilabel-learning network for RGB-thermal urban scene semantic segmentation," in *IEEE Transactions on Image Processing*, **30**, pp. 7790–7802 (2021). DOI: 10.1109/TIP.2021.3109518

49. Hong, H., Chen, W., Xu, C., et al. "Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy", *Geocarto International*, **32**(2), pp. 139–154 (2017).

Biographies

Azad Aram received his BSc degree in Water Engineering and an MSc degree in Civil Engineering with a focus on water from the Islamic Azad University, Mahabad Branch. His master thesis is entitled "Hydraulic simulation of river flow to determine the maximum values of flood zone using software Mike 11".

Mohammad Reza Dalalian holds a PhD degree in Physics and Soil Conservation. He conducted several research on wind erosion, modeling the uptake of soil contaminants, and soil improvement. Since 2004, he has been an Assistant Professor at the Department of Soil Science at the Islamic Azad University, Tabriz Branch. In addition to teaching, research activities in the form of master and doctoral students' theses and research projects, he is the Director of the Department of Soil and Environmental Engineering.

Siamak Saedi holds a PhD in Soil Science. Since 2001, he has been an Assistant Professor at the Department of Soil Science at the Islamic Azad University, Tabriz Branch. In addition to teaching, research activities in the form of Master and doctoral students' theses and research projects, he was titled as a distinguished professor in 2011. In fact, he is known as the founder of Soil Science major at the Azad University of Tabriz as well as the Director of the Department of Soil Science for a few years. In addition, he was the executive manager of a private soil, plant and water laboratory for four years. Since 2009, he has worked as an official expert of justice and has studied land reform processes in rural societies in North-Western of Iran.

Omid Rafieyan holds a PhD degree in Natural Resources Engineering. He has been interested in Remote Sensing and GIS since 2000 and has specialized in his master's and doctoral dissertations on the application of RS&GIS in natural resources. His other research interests include multi-criteria decision-making, urban green space, and interpretation of aerial photographs. From 2002 to 2008, he was the Director of scientific-executive projects entitled "Updating vegetation maps using satellite images" and "Preparing the forest type and density map using aerial photographs" in the Yekom Consulting Engineers Company. Since 2008, he has been an Assistant Professor at the Department of Environmental Engineering at the Islamic Azad University, Tabriz Branch. In addition to teaching, research activities in the form of master and doctoral students' theses and research projects, he is the Head of the Scientometrics Office of the university.

Samad Darbandi holds a PhD degree in Irrigation and Drainage. Since 2000, he has been an Assistant Professor at the Department of Water Engineering at the Islamic Azad University, Tabriz Branch. He is interested in modeling and simulation in the fields of water and soil. In addition to teaching, he conducts research activities in the form of master and doctoral students' theses and research projects.