Chapter

Evaluation of Landslide Susceptibility of Şavşat District of Artvin Province (Turkey) Using Machine Learning Techniques

Halil Akinci, Mustafa Zeybek and Sedat Dogan

Abstract

The aim of this study is to produce landslide susceptibility maps of Şavşat district of Artvin Province using machine learning (ML) models and to compare the predictive performances of the models used. Tree-based ensemble learning models, including random forest (RF), gradient boosting machines (GBM), and extreme gradient boosting (XGBoost), were used in the study. A landslide inventory map consisting of 85 landslide polygons was used in the study. The inventory map comprises 32,777 landslide pixels at 30 m resolution. Randomly selected 70% of the landslide pixels were used for training the models and the remaining 30% were used for the validation of the models. In susceptibility analysis, altitude, aspect, curvature, distance to drainage network, distance to faults, distance to roads, land cover, lithology, slope, slope length, and topographic wetness index parameters were used. The validation of the models was conducted using success and prediction rate curves. The validation results showed that the success rates for the GBM, RF, and XGBoost models were 91.6%, 98.4%, and 98.6%, respectively, whereas the prediction rate were 91.4%, 97.9%, and 98.1%, respectively. Therefore, it was concluded that landslide susceptibility map produced with XGBoost model can help decision makers in reducing landslide-associated damages in the study area.

Keywords: landslide susceptibility mapping, machine learning, RF, GBM, XGBoost, Şavşat

1. Introduction

Natural disasters cause displacement of people, injuries, loss of life, and damage to infrastructure facilities and cultural heritage, which can directly give rise to extreme economic losses. According to the data from Emergencies Database (EM-DAT), managed by the Center for Research on the Epidemiology of Disasters (CRED), 11,755 people died worldwide due to 396 natural disasters that occurred in 2019; 94.9 million people were affected by these disasters and an economic loss of 103 billion dollars was suffered [1]. On the contrary, according to the report prepared by the AON company, which provides insurance and reinsurance brokerage and risk management consultancy services, the damage caused by natural disasters in 2020 is estimated to be 268 billion dollars [2]. In the AON report prepared in 2020, the value of total economic losses caused by natural disasters in the 2010–2019 period was calculated as 2.98 trillion dollars. In the same report, the economic losses in question were reported to be 1.1 trillion dollars higher than that in the 2000–2009 period [3].

Landslide is generally defined as the downward movement and displacement of the material forming a slope with the effect of gravity [4]. Rabby and Li [5] stated in their study that landslides are a very common phenomenon and account for 9% of disasters in the world. Landslides, especially those caused by rainfall, are the most damaging natural disasters in mountainous and rugged regions, resulting in loss of life, damage to property, and economic loss [6]. Landslide susceptibility maps are one of the important data needed to identify landslide-hazardous areas and to reduce losses due to landslides [7, 8]. Many different approaches and models have been implemented in the production of landslide susceptibility maps. Merghadi et al. [9] and Tang et al. [10] classified the modeling approaches into four categories: the heuristic, physically based, statistical, and machine learning (ML) models. Heuristic and physically based models (also known as deterministic models) have their own characteristics and disadvantages. Heuristic models are highly subjective and rely on experts' opinions and experience on assigning weightage to landslide-conditioning factors [11–14]. In this approach, differences in expert opinions or insufficient information about the study area may cause inconsistent results [15]. Physically based or deterministic models use laws of mechanics to analyze slope stability. The advantages of these models are that they do not require long-term landslide inventory data and are more useful in areas where landslide inventories are missing [15]. However, deterministic models are suitable for small areas where landslide types are simple and ground conditions are fairly uniform [14], but they require detailed geotechnical and hydrogeological data on these areas [13]. To overcome the disadvantages of the above two approaches and to produce reliable landslide susceptibility maps, statistics-based models have been developed [14]. Statistics-based models evaluate the correlation between past landslides and the conditioning factors that had an impact on their occurrence [16] and they need landslide inventory data for this [17].

In recent years, machine learning (ML) techniques such as support vector machine [18, 19], decision tree [20, 21], generalized linear model [22, 23], logistic model tree [13, 16], artificial neural networks [6, 24, 25], and Naïve Bayes [26–28] have been widely applied for landslide susceptibility mapping (LSM). Sahin [29] and Merghadi et al. [9] stated that tree-based ensemble algorithms provide better prediction performance for LSM compared to any single model. In addition, Sahin [30] stated that ensemble learning techniques, such as random forest (RF), gradient boosting machine (GBM), and extreme gradient boosting (XGBoost), are efficient and robust for creating landslide susceptibility maps and that these algorithms would be preferred more frequently in the future for their robustness.

The most common natural disasters in Turkey are landslides and floods. Artvin is one of the provinces in Turkey that experiences the most frequent natural disasters. Landslides occur almost every year in the province of Artvin, especially due to meteorological conditions (extreme rainfall) and anthropogenic activities, such as agricultural activities, excessive irrigation, and road excavations. Şavşat is one of the districts of Artvin where landslides are most common. Şavşat, a Cittaslow city, stands out with its historical and natural beauties and has a high tourism potential. For this reason, it is very important to evaluate the landslide susceptibility to reduce the landslide-associated damages in the district. The aim of this study is to produce landslide susceptibility maps of Şavşat district of Artvin Province using RF, GBM, and XGBoost ML models and to evaluate the performances of the models. Eleven factors commonly used in LSM studies were used in the study. The produced landslide susceptibility maps were validated using the validation dataset.

2. Study area and data used

Şavşat, like other districts of Artvin, is a district with a rugged terrain. Şavşat, spreading on a 1272.27 km^2 land, is located between 41°05′11″ and 41°30′56″ north latitudes and 42°04′30″ and 42°35′47″ east longitudes (**Figure 1**). In the study area, the altitude varies between 590 and 3005 m with the average altitude being 1789.14 m. The average slope of the study area is 21.17°, whereas the maximum slope is 72.53°. The slope is over 20° in ~55% of the study area.

According to the data from the Turkish Statistical Institute (TURKSTAT), the total population of Şavşat district in 2020 is 17,024. Of this population, 6,123 live in the town and 10,901 live in villages [31]. There is a transitional climate between the Black Sea climate and the continental climate in the district. While semi-humid climatic conditions are observed in the low valley floors, cold humid climatic conditions are observed in the higher elevations. In addition, winters are very long in places with high altitudes. According to the data (November 2012–March 2021) from the General Directorate of Meteorology, sum of monthly average rainfall in the study area is 715.60 mm. The monthly average rainfall is minimum in February with 27.8 mm and maximum in May with 111.03 mm. In the study area, the monthly average temperature was maximum at 32.8°C in August and minimum at −7.4°C in December [32].

Şavşat is located in the eastern part of the Eastern Pontides and the southern part of Transcaucasia. In the study area, intrusive, volcanic, and volcano-sedimentary facies have developed due to the magmatic activities that took place in the Dogger, Late Cretaceous, and Eocene ages. In the north and northwest part of the region, units representing the same stratigraphic unity surfaces in a range extending from the Liassic to the Early-Middle Eocene. In the southern part, units representing two separate stratigraphic units are surfaced. The sequence in the west of the southern section is characterized by units of Early-Middle Jurassic and Late

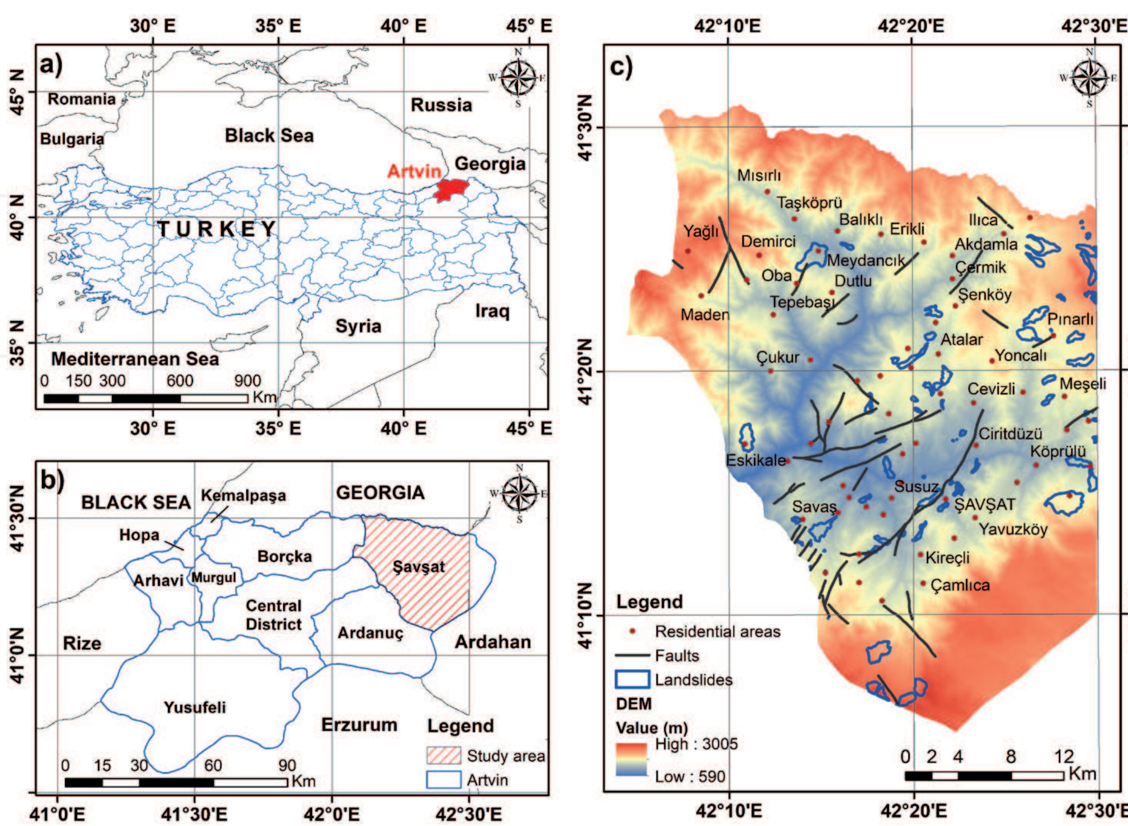

Figure 1.
Study area.

Cretaceous age, and the sequence in the east of the southern section is characterized by units of Late Cretaceous and Middle Eocene age. Tertiary units surfacing in the eastern and southeastern parts of the region are considered as common units [33]. According to the earthquake zone map of Turkey, Şavşat district is located in the third degree earthquake zone. However, the most common natural disaster in the district is landslide [34]. The landslides occurring in the study area are mostly of complex type. Landslides are observed in larger areas with respect to Quaternary alluvium and slope debris [33].

2.1 Landslide inventory map

To reliably predict future landslides, reliable landslide inventory maps containing information about past landslides are needed [16]. As stated by Parise [35], landslide inventory maps represent the spatial distribution of landslides and provide information about the location, typology, and activity status of landslides. In this study, the landslide inventory map produced by Artvin Provincial Directorate of Disaster and Emergency was used. The landslide inventory map contains 85 landslide polygons. The area of the smallest landslide polygon in the study area is 0.01 ha (99.34 m^2), and the area of the largest landslide polygon is 325.97 ha. The average area of the landslide polygons is 34.75 ha. Landslides cover ~3% of the study area. The lengths of the landslides in the region vary between 13 and 3100 m and their widths vary between 10 and 2780 m. According to their activities, 28 of these landslides are active, 32 are stalled, and 25 are inactive landslides. According to Varnes [4] classification of mass movements, 6 of the landslides were classified as slide, 2 as lateral spread, 20 as flow, and the remaining 57 as complex.

2.2 Landslide-conditioning factors

Evaluation of landslide susceptibility in a region depends on determining the factors that are effective in the formation of landslides in that region and on collecting spatial data related to these factors [36]. Yi et al. [8] stated that there is no widely accepted procedure for the selection of factors used in LSM. Yanar et al. [37], on the contrary, stated that the main limitation in determining the factors to be used to create landslide susceptibility maps is the availability of data. In this study, 11 factors including altitude, aspect, curvature, distance to drainage network, distance to faults, distance to roads, land cover (CORINE 2018), lithology, slope, slope length, and topographic wetness index (TWI) were used based on the availability of data, geo-environmental conditions of the study area, and literature survey. Spatial data on these factors are collected from different sources (**Table 1**). Landslide-conditioning factor maps were generated using ESRI ArcGIS 10.5 and SAGA GIS 7.9.0 software and were converted into raster format with 30 m spatial resolution.

2.2.1 Altitude

Altitude is associated with various geomorphological and meteorological factors such as weathering, weather conditions, wind effect, and precipitation, which are effective in the formation of landslides [6]. For this reason, it has been used in almost all LSM studies. The digital elevation model (DEM) of the study area was created using 10-m-interval contours on the topographic maps and it was converted to raster format with 30-m spatial resolution. The altitude map of the study area was generated from this DEM. The altitude in the study area varies between 590 and 3005 m. DEM was reclassified into 10 classes at 240 m intervals (**Figure 2a**).

Original data	Factors	Data type	Scale	Data provider
Landslide inventory	Landslide locations	Polygon	1/25,000	Artvin Provincial Directorate of Disaster and Emergency
Geological map	Lithology	Polygon	1/100,000	General Directorate of Mining Research and Exploration (GDMRE)
	Distance to fault lines	Polyline	1/100,000	
Topographical map	Altitude	GRID	1/25,000	General Directorate of Mapping
	Slope	GRID	1/25,000	
	Slope length	GRID	1/25,000	
	Aspect	GRID	1/25,000	
	Curvature	GRID	1/25,000	
	TWI	GRID	1/25,000	
	Distance to drainage network	GRID	1/25,000	
Road network	Distance to roads	Polyline	1/25,000	Basarsoft Information Technologies Inc.
CORINE 2018	Land cover	Polygon	1/100,000	European Union Copernicus Land Monitoring Service

Table 1.
Data and data sources.

2.2.2 Aspect

Aspect has an important role in landslide formation as it affects factors such as exposure to sunlight and the intensity of solar radiation, wind, rainfall and, soil moisture [38, 39]. For this reason, aspect is widely used in LSM studies [6, 26, 36, 40]. The aspect map used in this study was produced from DEM and divided into nine classes (flat, north, northeast, east, southeast, south, southwest, west, and northwest) (**Figure 2b**).

2.2.3 Curvature

Curvature, which is widely used in geomorphometric analysis, is one of the basic terrain parameters and reflects the shape of the land surface [23, 41]. In curvature map, positive curvature values indicate that the surface is convex, negative curvature indicates that the surface is concave, and zero indicates that the surface is flat [42]. In this study, curvature map was derived from DEM using ArcGIS 10.5 software and divided into three subclasses, i.e., concave, flat, and convex (**Figure 2c**).

2.2.4 Distance to drainage network

The distance to the drainage networks is one of the important conditioning factors used in landslide susceptibility studies, since the pore water pressure that causes the formation of landslides increases in areas close to the drainage networks [23]. Drainage networks in the study area were generated from DEM using functions in ArcHydro toolbox in ArcGIS 10.5 software. The distance to the drainage networks was calculated using the Euclidean distance tool in ArcGIS 10.5.

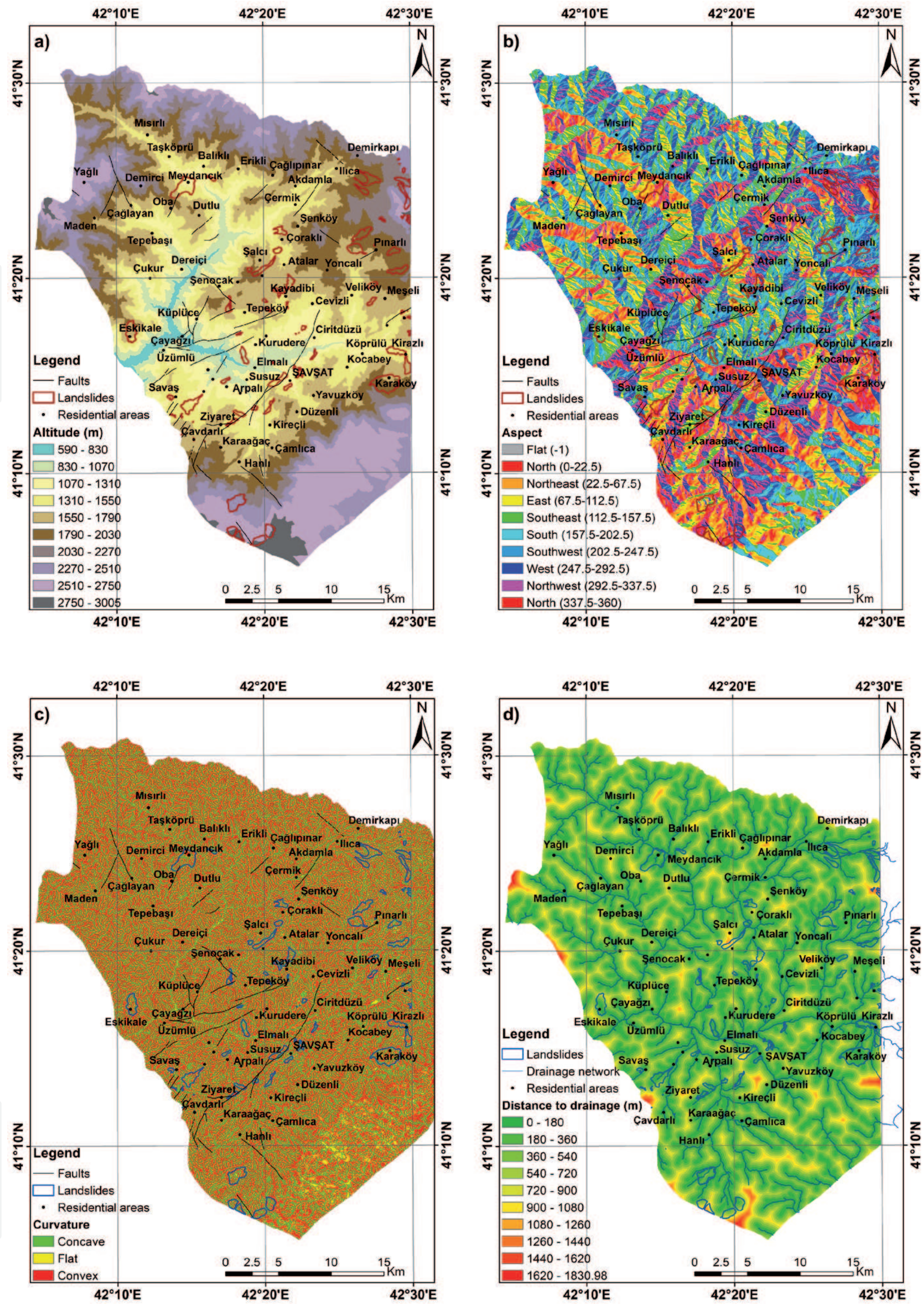

Figure 2.
The landslide conditioning factor maps: a) altitude b) aspect c) curvature d) distance to drainage network.

The maximum distance to the drainage networks in the study area has been calculated as 1830.98 m. The distance to the drainage networks is reclassified into 10 subclasses with equal intervals of 180 m (**Figure 2d**).

2.2.5 Distance to faults

Areas close to faults are highly susceptible to landslides as the strength decreases due to tectonic fractures [28]. Ba et al. [43] stated that landslides tend to occur around faults due to fractures in the rock mass. For this reason, the distance to the faults is taken into account in the landslide susceptibility analysis [14, 40, 44]. In this study, the distance to the faults was obtained using the Euclidean distance tool of ArcGIS 10.5 software. The maximum distance to the faults in the study area has been calculated as 13,016.61 m. The distance to the faults was classified into 10 subclasses with 1200 m intervals and used in the landslide susceptibility analysis (**Figure 3a**).

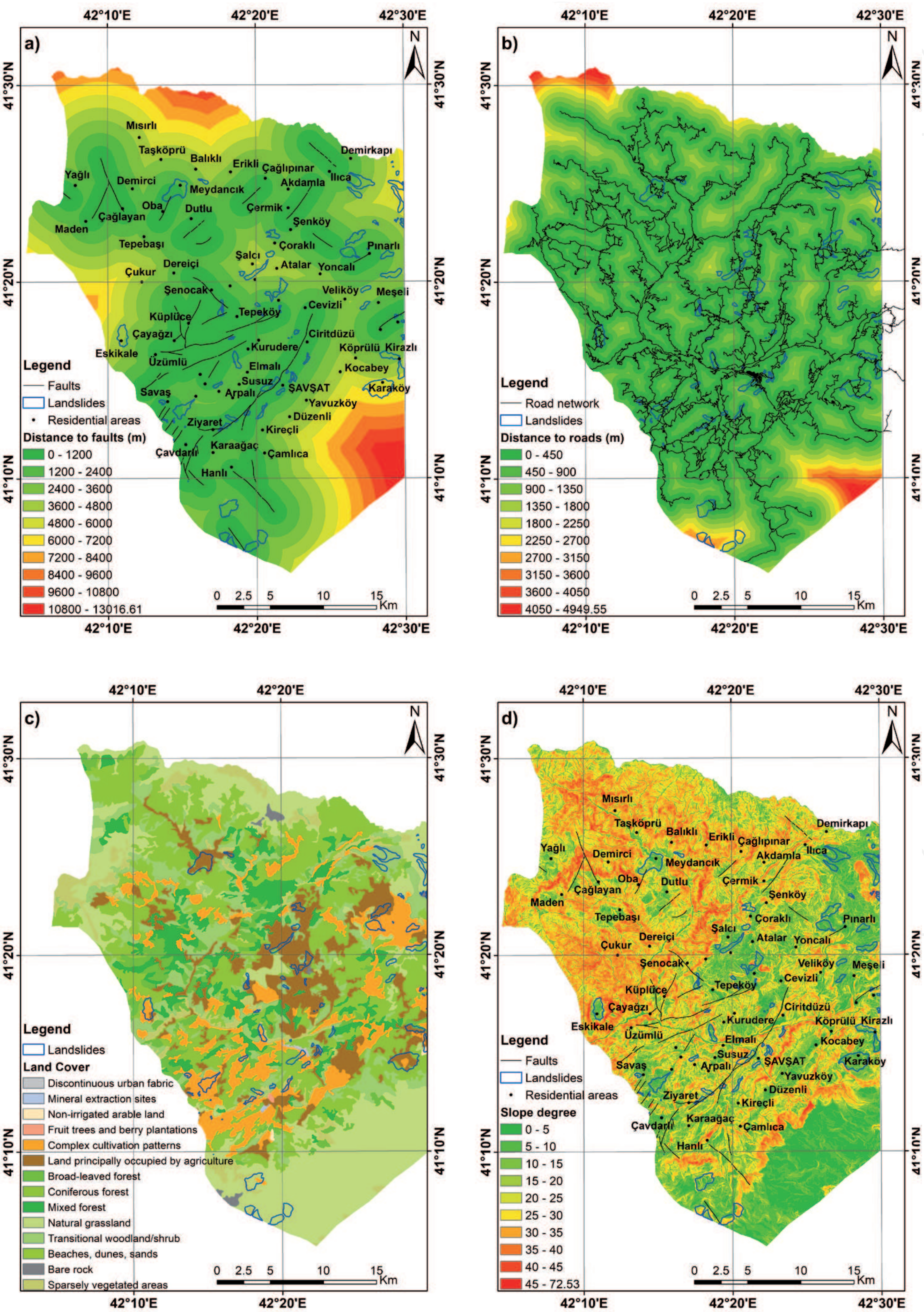

Figure 3.
The landslide conditioning factor maps: a) distance to faults b) distance to roads c) land cover d) slope.

2.2.6 Distance to roads

Road construction, which is considered to be one of the most important anthropogenic factors, destabilizes the slopes, so the probability of landslides along a road increases [43]. Roads built on slopes in areas with rough topography cause loss of toe support, change in topography, increase in tension behind the slope, and development of tension cracks [45, 46]. For this reason, distance to the road has been considered as one of the important conditioning factors in many studies [14, 17, 47]. The road network in the study area was supplied in digital format from Başarsoft Information Technologies Inc., which collects road data for the production of navigation maps in Turkey. Distance to roads was calculated using the Euclidean distance tool in ArcGIS 10.5 and reclassified into 10 subclasses at 450 m intervals (**Figure 3b**).

2.2.7 Land cover

Land cover maps, in general, represent what physical classes or materials (e.g., forest, pasture, field, lake, and wetland) the Earth's surface is spatially covered with. Land use or land cover maps are usually used in LSM studies for taking into consideration the effects of anthropogenic activities on rugged slopes on landslide formation [5]. In this study, CORINE 2018 land cover (CLC 2018) data provided by Copernicus Land Monitoring Service, one of the European Union's Earth Observation Programme services, were used. According to this dataset, the study area includes 14 different land cover classes (**Figure 3c**).

2.2.8 Slope

The slope angle, one of the most important factors governing the stability of slopes, is closely related to the shear forces acting on the slopes. As the angle of

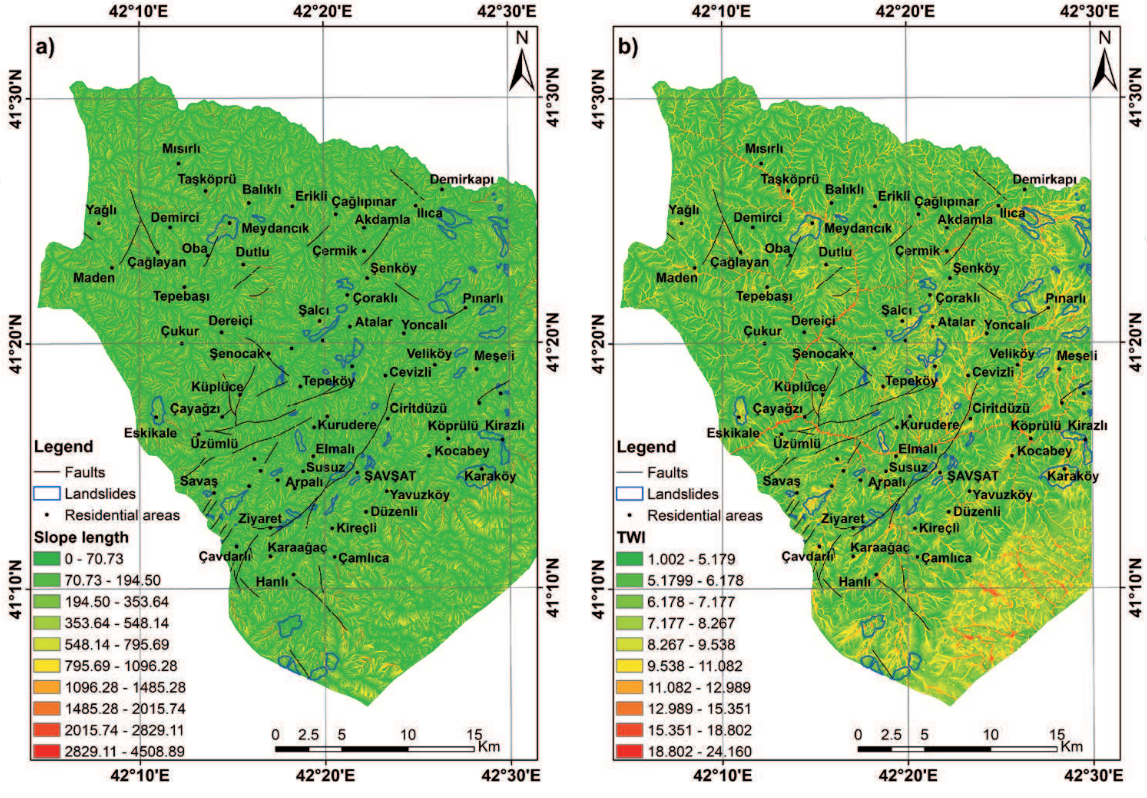

Figure 4.
The landslide conditioning factor maps: a) slope length b) TWI.

inclination increases, the shear stress in the materials forming the slope generally increases [48]. For this reason, slope angle has been used in all LSM studies, as is the case for the lithology parameter [18, 40, 49–51]. The slope in the study area varies between 0° and 72.53°. In this study, the slope was divided into 10 classes with 5° spacing, and a slope map of the study area was produced (**Figure 3d**).

2.2.9 Slope length

Slope length is one of the important topographic factors that affect the formation of landslides [6]. Kavzoglu et al. [18] defines the slope length as "the distance along a slope subject to uninterrupted over land flow." Slope length affects hydrological processes and soil loss, especially in mountainous areas [23]. This factor is closely related to the formation of landslides, because the potential for the materials forming the slopes to be carried downhill also increases with the increase of the slope length [52]. In this study, slope length was produced from DEM using SAGA GIS software and it was reclassified into 10 classes using the natural break classification method (**Figure 4a**).

2.2.10 Topographic wetness index (TWI)

TWI is an index generally used to characterize the spatial distribution of soil moisture [53] and is considered as an important factor contributing to the occurrence of landslides. Yanar et al. [37] stated that TWI indicates the locations and size of the water-saturated regions. For this reason, TWI has been used in many landslide susceptibility studies [26, 54, 55]. The following equation is used to calculate TWI:

$$TWI = \ln\left(\frac{A_s}{tan\beta}\right) \tag{1}$$

In the Eq. (1), As is the specific basin area and β is the slope in degrees. TWI index in the study area, varying between 1.002 and 24.160, was produced using SAGA GIS software. TWI index values were divided into 10 subclasses using the natural break classification method and used in sensitivity analysis (**Figure 4b**).

2.2.11 Lithology

Kavzoglu et al. [18] stated that lithology is one of the main factors that have a direct impact on the formation of landslides, as lithological and structural variations lead to changes in the strength and permeability of rocks and soils. For this reason, lithology has been one of the most important conditioning factors used in all landslide susceptibility evaluation studies. In this study, 1/100,000 scaled digital geological map obtained from General Directorate of Mineral Research and Exploration (GDMRE) was used to produce the lithological map of the study area. The geological map of the study area includes 16 lithological units (**Figure 5**).

3. Methodology

3.1 Random forest

First proposed by Breiman [56], RF is an ensemble learning method that creates multiple decision trees from the training dataset and combines the results of

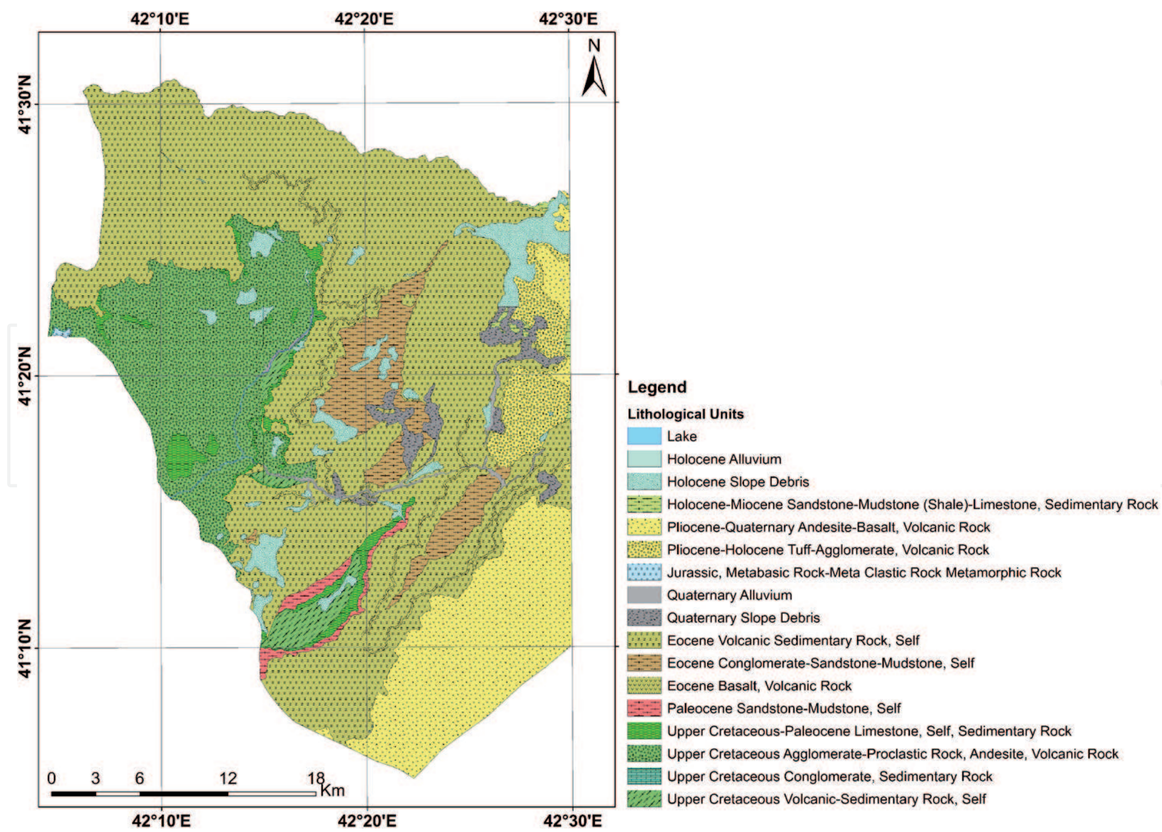

Figure 5.
Lithological map of the study area.

the decision trees to improve the predictive ability of the model [57]. According to Arabameri et al. [44] and Merghadi et al. [9], one of the most important advantages of RF is that it avoids the risk of overfitting, which is a common problem in other decision tree models. In the study conducted by Sahin [29], it is stated that requiring less hyperparameter tuning, compared to gradient boosting algorithms, was RF's main advantage. To create a classification model in RF, two parameters must be defined: *ntree* parameter, which refers to the number of decision trees generated by RF, and *mtry* parameter, which refers to the number of factors or variables used in each node of the decision tree. In this study, "*rf*" method of the "*caret*" package [58] was used in R 3.6.3 to apply the RF model. In the study, the *ntree* parameter was set to 100 and the *mtry* parameter to 8, and a 10-fold cross validation approach was used to reduce the variability of the model results.

3.2 Gradient boosting machine (GBM)

GBM [59] is a ML technique that combines multiple different models through boosting and regression trees to increase prediction precision [60]. The main feature of GBM is that it combines multiple weak learners to improve their performances. GBM, an ensemble learning method, combines multiple decision trees to create a more powerful model that can be used for classification or regression. In GBM, unlike RF, each tree tries to correct the error of the previous tree [61]. For this purpose, the residual errors calculated as a result of the prediction of the previous tree are minimized and the next tree is obtained, and these processes continue until the prediction results are stable or until the maximum number of trees is reached. In practice, the number of trees is chosen to be 100 or greater. There are four parameters that must be set by the user during the execution of the GBM, namely number of trees (*n.trees*), *shrinkage*, number of levels of trees (*interaction.depth*), and the minimum number of observations in trees' terminal nodes (*n.minobsinnode*). For low-variance and accurate predictions, the learning rate is chosen so that it converges to the optimum value with small steps in the right direction. The number of levels of trees is chosen between 8 and 32. In this study, GBM was performed using the "*gbm*" method in R 3.6.3.

3.3 Extreme gradient boosting (XGBoost)

XGBoost, developed by Chen and Guestrin [62], is based on the gradient boosting approach. XGBoost is based on the efficient and effective implementation of the gradient boosting algorithm. For this purpose, it interprets the approximate greedy algorithm with the Newton–Rapson method. XGBoost uses several classification and regression trees and integrates them using gradient boosting [63]. It produces fast and accurate solutions with univocal regression trees, weighed quantile approach, and sparsity aware split finding. It is trained very quickly, and since it is suitable for parallel learning technique, XGBoost increases the overall accuracy (performance) of the model by avoiding the overfitting problem during the training process [64]. XGBoost uses two additional techniques called shrinkage and column (feature) subsampling to avoid overfitting [62]. Wang et al. [61] noted that the computational speed and accuracy of XGBoost has been significantly improved compared to GBM. In this study, the XGBoost model is implemented in R 3.6.3 using the "*xgbTree*" method of the "*caret*" package.

3.4 Preparation of training and validation dataset

"Landslide (or positive)" and "non-landslide (or negative)" samples are needed in the study area during the training and validation of the models used to create landslide susceptibility maps. The ratio of 70:30 has been commonly used in the literature to produce training and validation datasets [6, 8, 65, 66]. In particular, 70% of the landslide inventory data is used for training the models and the remaining 30% is used for the validation of the models. Huang and Zhao [67], on the contrary, stressed that the number of positive and negative samples in the training and validation datasets should be equal, i.e., having a ratio of 1:1. For this reason, as many negative samples as the number of positive samples are selected in the study area. In this study, 85 landslide polygons on the inventory map were converted to 30 m × 30 m resolution raster format and 32,777 landslide pixels were obtained. A value of "1" was assigned to positive or landslide pixels in the study area. Then, 32,777 non-landslide pixels were randomly selected in the study area in the R program and the value of "0" was assigned to these pixels. Randomly selected 70% of the landslide and non-landslide pixels (45,888 pixels in total) were used for training the models and the remaining 30% (19,666 pixels) were used for the validation of the models.

3.5 Multicollinearity analysis for landslide-conditioning factors

One of the important steps of LSM is to control the multicollinearity between landslide-conditioning factors [8]. Multicollinearity is an important analysis used to determine the conditional independence between the factors during the selection of the conditioning factors to be used in susceptibility models, and thus, to prevent the models from producing erroneous predictions [9, 68]. Commonly used indicators for multicollinearity analysis are tolerance (TOL) and variance inflation factor (VIF). A TOL value less than 0.1 or a VIF value greater than 10 indicates multicollinearity [8, 16, 44]. TOL and VIF values calculated using the training dataset for this study are shown in **Table 2**. The results show that there is no multicollinearity

Landslide conditioning factors	Statistics	
	TOL	VIF
Altitude	0.4713	2.1217
Aspect	0.9770	1.0235
Curvature	0.7879	1.2692
Distance to drainage network	0.7916	1.2633
Distance to faults	0.7786	1.2844
Distance to roads	0.5552	1.8011
Land cover	0.7206	1.3877
Lithology	0.8763	1.1412
Slope	0.5373	1.8610
Slope length	0.7345	1.3615
Topographic Wetness Index	0.4595	2.1761

Table 2.
Multicollinearity analysis of landslide-conditioning factors.

among the landslide-conditioning factors used in the study. Therefore, all selected factors were used to produce landslide susceptibility map of the study area.

4. Results and discussion

4.1 Landslide susceptibility mapping

In this study, RF, GBM, and XGBoost models were successfully applied and landslide susceptibility index (LSI) maps were produced via R 3.6.3 using the training data set for each model. Then, landslide susceptibility maps were obtained by reclassifying the LSI maps into five classes: very low, low, medium, high, and very high, using the natural breaks (Jenks) classification method in ArcGIS 10.5 software (**Figure 6**).

The spatial distributions (in percentages) of the susceptibility classes for each model are given in **Figure 7**. It has been determined that the study area is highly or very highly susceptible to landslides by 27.27%, 11.13%, and 16.89% according to the GBM, RF, and XGBoost models, respectively (**Figure 7**).

The significance degrees of the landslide-conditioning factors used in the study are presented in **Figure 8**. It has been observed in all models that the lithology is the most important parameter. After lithology, the most important or most effective parameters in the study area were determined to be altitude, distance to faults, slope, and land cover parameters. Slope length and curvature were the least significant parameters in all models (**Figure 8**). The findings related to the parameters found to be effective in terms of landslide are explained in the following sections.

When **Table 3** is examined, ~76% of the landslides in the study area can be seen to have occurred at altitudes between 1070 and 2030 m. In respect of altitude, 1070–1310, 1310–1550, 1550–1790, and 1790–2030 m altitude classes were found to be susceptible to landslides (**Table 3**). The main reason why these altitude classes are susceptible to landslides is that more than 90% of the village settlements in the study area are located between these altitudes. Uncontrolled excavations and uncontrolled agricultural activities in villages are the most important factors that trigger landslides. In the study by Erener et al. [34], conducted in Şavşat district and

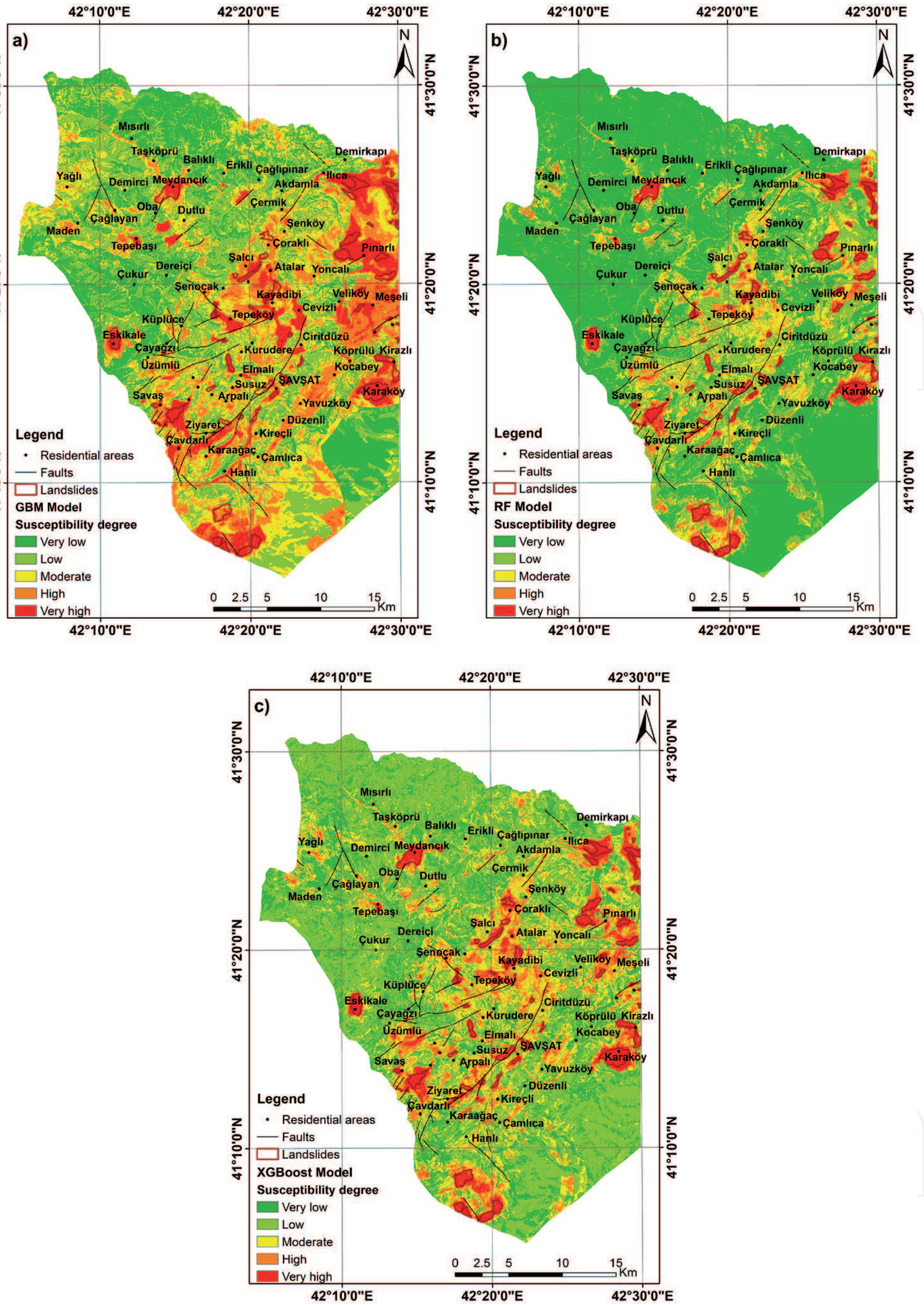

Figure 6.
Landslide susceptibility maps produced using a) GBM b) RF c) XGBoost.

covering a more limited (small) region compared to this study, the altitude class between 1500 and 2000 m was found to be susceptible to landslides.

When the study area is examined in terms of slope, it is seen that 0°–5°, 5°–10°, 10°–15°, and 15°–20° slope classes are more susceptible to landslides (**Table 3**). In these slope classes, 82.31% of the landslides occurred in the study area. The fact that complex mass movements (creeping and spreading) in the study area are generally seen in areas with low slope degrees (approximately in the range of 7°–12°) have provided these results in terms of slope.

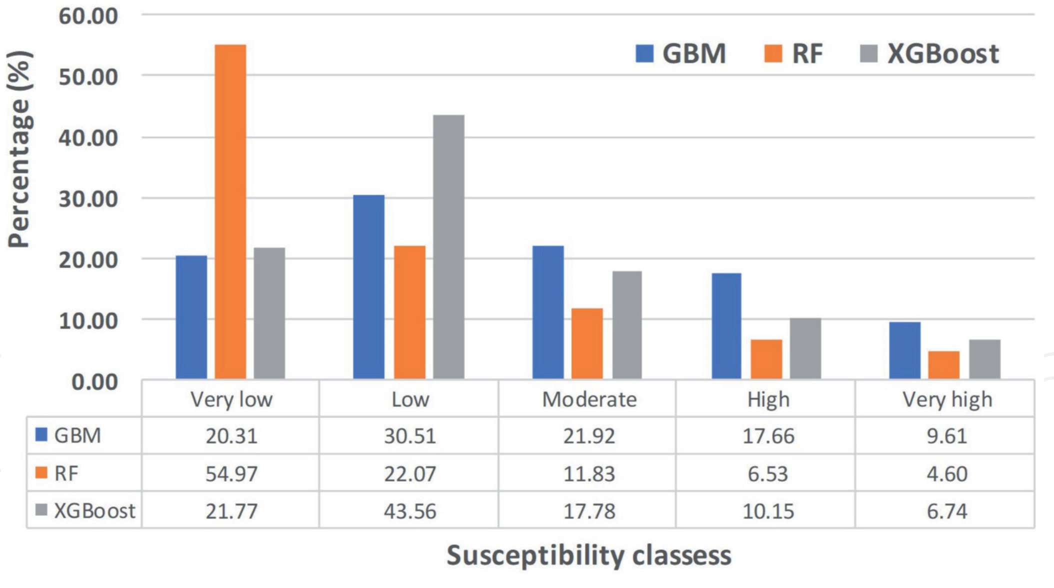

Figure 7.
Percentage distributions of susceptibility classes.

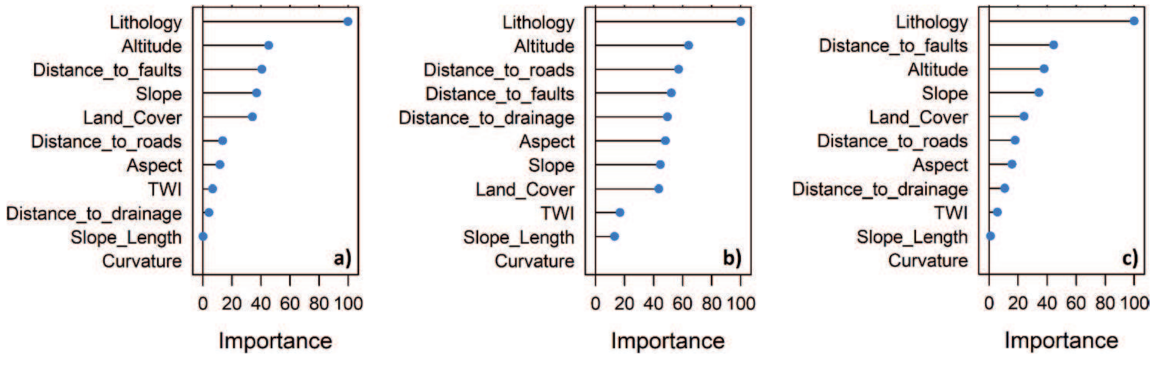

Figure 8.
Importance of landslide-conditioning factors for a) GBM b) RF c) XGBoost.

When **Table 3** is examined, it is seen that ~55% of the landslides in the study area occur on slopes with north, northeast, and northwest aspects. When the frequency ratios in **Table 3** are examined, it is clearly seen that the slopes with these aspects have the highest frequency ratio value, and therefore, they are more susceptible to landslides. In the study conducted by Akıncı and Zeybek [69], in the Ardanuç district, which is adjacent to the Şavşat district and has similar topographical and geomorphological characteristics with the study area, the slopes with north, northwest, and northeast aspects were determined to be more susceptible to landslides.

Within the first 3600 m margin of the faults, 74% of the landslides occurred in the study area (**Table 3**). In the study area, the landslide susceptibility tends to decrease with distance from the faults. Although the region most susceptible to landslides in terms of distance to faults is 4800–6000 m, it is seen that distance classes of 0–1200, 1200–2400, and 2400–3600 m are also susceptible to landslides (**Table 3**). Althuwaynee et al. [70] stated that the probability of landslide decreases as the distance to the faults increases. Also in the LSM study conducted by Akinci et al. [40] in the area covering Arhavi, Hopa, and Kemalpaşa districts of Artvin Province, the areas within the first 2000 m distance to the faults were determined to be more susceptible to landslides.

Considering the CORINE 2018 land cover data, it was determined that ~56% of the landslides in the study area occurred in agricultural areas (**Table 3**).

DOI: http://dx.doi.org/10.5772/intechopen.99864

Factor	Subclasses	Pixels in domain	Pixels with landslide	Percentage of landslides (%)	Percentage of domain (%)	FR
Altitude (m)	590–830	16659	0	0.00	1.34	0.0000
	830–1070	65715	928	2.83	5.28	0.5363
	1070–1310	129711	3488	10.64	10.42	1.0212
	1310–1550	202140	5893	17.98	16.24	1.1071
	1550–1790	242416	7455	22.74	19.48	1.1679
	1790–2030	188846	8231	25.11	15.17	1.6552
	2030–2270	140243	1174	3.58	11.27	0.3179
	2270–2510	120268	2402	7.33	9.66	0.7585
	2510–2750	121891	1687	5.15	9.79	0.5256
	2750–3005	16840	1519	4.63	1.35	3.4255
Slope (degree)	0–5	86650	3307	10.09	6.96	1.4493
	5–10	156262	10914	33.30	12.55	2.6524
	10–15	160085	8205	25.03	12.86	1.9464
	15–20	160639	4554	13.89	12.91	1.0766
	20–25	174982	2652	8.09	14.06	0.5756
	25–30	194812	1656	5.05	15.65	0.3228
	30–35	176843	1069	3.26	14.21	0.2296
	35–40	96265	302	0.92	7.73	0.1191
	40–45	28481	98	0.30	2.29	0.1307
	45–72.53	9710	20	0.06	0.78	0.0782

Factor	Subclasses	Pixels in domain	Pixels with landslide	Percentage of landslides (%)	Percentage of domain (%)	FR
Aspect	Flat	4416	49	0.15	0.35	0.4214
	North	148077	6845	20.88	11.90	1.7555
	Northeast	151999	5873	17.92	12.21	1.4673
	East	148757	3387	10.33	11.95	0.8647
	Southeast	161166	2816	8.59	12.95	0.6635
	South	162974	2667	8.14	13.09	0.6215
	Southwest	161749	2269	6.92	12.99	0.5327
	West	155008	3680	11.23	12.45	0.9016
	Northwest	150583	5191	15.84	12.10	1.3091
CORINE 2018	112	861	45	0.14	0.07	1.9848
	131	333	0	0.00	0.03	0.0000
	211	927	275	0.84	0.07	11.2657
	222	408	0	0.00	0.03	0.0000
	242	133175	7439	22.70	10.70	2.1213
	243	130192	10747	32.79	10.46	3.1348
	311	46639	241	0.74	3.75	0.1962
	312	340177	3839	11.71	27.33	0.4286
	313	96234	82	0.25	7.73	0.0324
	321	278125	6054	18.47	22.34	0.8266
	324	149283	2678	8.17	11.99	0.6812
	331	1380	0	0.00	0.11	0.0000
	332	4746	181	0.55	0.38	1.4483
	333	62249	1196	3.65	5.00	0.7296

DOI: http://dx.doi.org/10.5772/intechopen.99864

Factor	Subclasses	Pixels in domain	Pixels with landslide	Percentage of landslides (%)	Percentage of domain (%)	FR
Distance to faults (m)	0–1200	353573	10042	30.64	28.41	1.0786
	1200–2400	311578	8822	26.92	25.03	1.0752
	2400–3600	198557	5463	16.67	15.95	1.0448
	3600–4800	132651	2932	8.95	10.66	0.8394
	4800–6000	91279	3955	12.07	7.33	1.6454
	6000–7200	62754	1563	4.77	5.04	0.9459
	7200–8400	43119	0	0.00	3.46	0.0000
	8400–9600	25977	0	0.00	2.09	0.0000
	9600–10800	14205	0	0.00	1.14	0.0000
	10800–13016.61	11036	0	0.00	0.89	0.0000

Factor	Subclasses	Pixels in domain	Pixels with landslide	Percentage of landslides (%)	Percentage of domain (%)	FR
Lithology	Lake	915	0	0.00	0.07	0.0000
	e-10-s	625072	7853	23.96	50.22	0.4771
	e-18-s	64071	1795	5.48	5.15	1.0639
	e-V2	15193	6	0.02	1.22	0.0150
	Jbmclm	524	0	0.00	0.04	0.0000
	k2–10-s	19698	77	0.23	1.58	0.1484
	k2–2-k	256	0	0.00	0.02	0.0000
	k2-pn-8-s	29136	1800	5.49	2.34	2.3461
	k2-V16-V15-V13	189788	942	2.87	15.25	0.1885
	plQ-V13-V2	163657	6297	19.21	13.15	1.4612
	plQ2-V17-V16	41251	1883	5.74	3.31	1.7335
	pn-19-s	10873	383	1.17	0.87	1.3377
	Q-21-k	6144	16	0.05	0.49	0.0989
	Q-23-k	23921	4063	12.40	1.92	6.4502
	Q2–21-k	434	0	0.00	0.03	0.0000
	Q2–23-k	50601	7359	22.45	4.07	5.5229
	Q2m-20-ks	3195	303	0.92	0.26	3.6014

Table 3.
Spatial relationship between landslide-conditioning factors and landslides.

Non-irrigated arable lands (CORINE land cover code 211), agricultural areas within natural vegetation (243), mixed agricultural areas (242), discontinuous urban structure (112), and bare rocks (332) were determined as landslide sensitive areas. The scattered settlements in the villages cause uncontrolled excavations, which in turn triggers landslides. In the landslide susceptibility study conducted by Erener et al. [34] in Şavşat district, it was reported that landslide activity increased in areas where the original vegetation was removed or changed. In the same study, it was determined that farming areas, irrigated or dry, were more susceptible to landslides. Researchers attributed this to the deforestation in agricultural areas.

4.2 Validation and comparison of landslide susceptibility models

Thi Ngo et al. [7] stated that it is important to identify landslide-prone areas with high accuracy and to use an appropriate metric for the performance evaluation to produce a reliable landslide susceptibility map. The performances of the models used in the production of landslide susceptibility maps are mostly evaluated using the receiver-operating characteristics (ROC) curve [28, 38, 45, 60, 71–73]. Therefore, in this study, the receiver-operating characteristic-area under the curve (ROC-AUC) approach was applied to evaluate and measure the performances of ML models. The ROC curve is a graph showing the true positive rate (TPR or sensitivity) on the vertical axis and the false positive rate (FPR or 1-specificity) on the horizontal axis. In the ROC curve, the most important indicator used to evaluate the accuracy or performance of the susceptibility model is the AUC. AUC takes values between 0.5 and 1 [71]. An AUC value close to 1.0 indicates high performance of the model and close to 0.5 indicates low performance of the model. On the contrary, Chen et al. [74] and Wang et al. [17] stated that the AUC value can be classified in five classes: poor (0.5–0.6), moderate (0.6–0.7), good (0.7–0.8), very good (0.8–0.9), and excellent (0.9–1.0).

In the study, success rate and prediction rate curves were created using training and validation data sets, respectively. The success rate curve is used to understand how well the models used to produce landslide susceptibility maps to classify existing landslide areas [74]. In this study, the AUC values of the success rate curves for the GBM, RF, and XGBoost models were calculated as 91.6%, 98.4%, and 98.6%, respectively (**Figure 9a**). Since the success rate curve is produced using the training

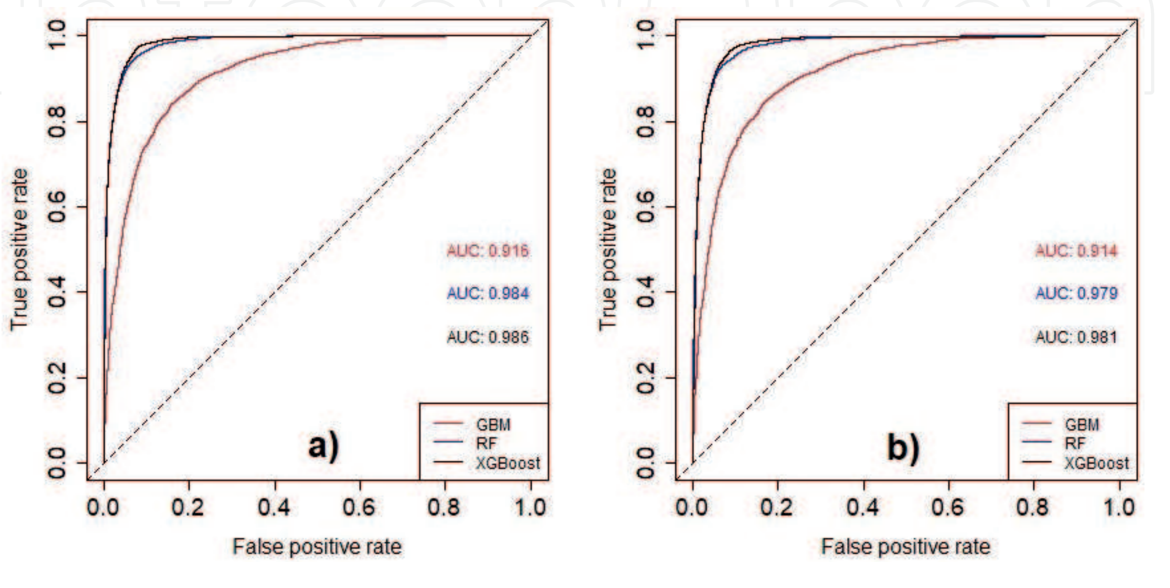

Figure 9.
a) Success rate b) prediction rate curves for ML models.

data set, it is not an appropriate indicator to evaluate the predictive capabilities of the models [21, 42]. The prediction rate curve should be used to evaluate the prediction capabilities of the models [75]. The prediction rate curve shows how well the models predict unknown or probable future landslides [5]. The AUC values of the prediction rate curves produced for the GBM, RF, and XGBoost models were calculated as 91.4%, 97.9% and 98.1%, respectively (**Figure 9b**). AUC value being close to 1.0 in three models show, according to the classification made by Chen et al. [74] and Wang et al. [17], that their performances, i.e., their prediction capacities, are excellent.

5. Conclusions

In this study, RF, GBM, and XGBoost algorithms were used for landslide susceptibility mapping of Şavşat district of Artvin Province. The performances of these models were evaluated using success rate and prediction rate curves. According to the AUC values, the models used in the study showed excellent performance. However, the XGBoost model outperformed the other two models in landslide susceptibility mapping of the study area. Therefore, it was concluded that the susceptibility map produced by the XGBoost model can help decision makers and planners in reducing the risks caused by landslides in the region and in land use planning. In this study, 11 factors—altitude, aspect, curvature, distance to drainage network, distance to faults, distance to roads, land cover, lithology, slope, slope length, and TWI—were used based on the availability of the data, geo-environmental conditions of the study area, and literature survey. As a result of the study, it was concluded that the main factor governing the landslides in the study area in all three models is lithology. The artificial factors that trigger landslides across the province of Artvin, as in Şavşat district, are uncontrolled excavation works (usually road widening), uncontrolled explosive excavations, and uncontrolled agricultural land irrigation. In this respect, providing basic disaster awareness trainings to citizens residing in areas susceptible to landslides in the study area and trainings on the causes, effects, and consequences of landslides will be beneficial in terms of risk reduction. Similarly, taking into account landslide susceptibility maps in selecting dwelling zones in rural areas and in determining the routes through which infrastructure facilities such as drinking water, natural gas, electricity, and sewerage will pass, will be effective in reducing the risks associated with landslides in the study area.

Conflict of interest

The authors declare no conflict of interest.

Author details

Halil Akinci[1]*, Mustafa Zeybek[2] and Sedat Dogan[3]

1 Department of Geomatics Engineering, Artvin Çoruh University, Artvin, Turkey

2 Güneysınır Vocational School, Selcuk University, Konya, Turkey

3 Department of Geomatics Engineering, Ondokuz Mayis University, Samsun, Turkey

*Address all correspondence to: halil.akinci@artvin.edu.tr

References

[1] CRED. Natural Disasters 2019. Centre for Research on the Epidemiology of Disasters (CRED). Brussels: CRED; 2020. Available from: https://emdat.be/sites/default/files/adsr_2019.pdf (Accessed: February 2, 2021)

[2] AON. Weather, Climate & Catastrophe Insight: 2020 Annual Report. AON; Chicago, Illinois. Available from: https://www.aon.com/global-weather-catastrophe-natural-disasters-costs-climate-change-2020-annual-report/index.html (Accessed: February 18, 2021)

[3] AON. Weather, Climate & Catastrophe Insight: 2019 Annual Report. AON; Chicago, Illinois. http://thoughtleadership.aon.com/Documents/20200122-if-natcat2020.pdf (Accessed: February 13, 2021)

[4] Varnes DJ. Slope movement types and processes. In: Schuster RL, Krizek RJ, editors. Landslides, Analysis and Control, Special Report 176: Transportation Research Board. Washington, DC: National Academy of Sciences; 1978. p. 11-33.

[5] Rabby YW, Li Y. Landslide susceptibility mapping using integrated methods: A case study in the Chittagong hilly areas, Bangladesh. Geosciences. 2020;10:483. DOI:10.3390/geosciences10120483

[6] Youssef AM, Pourghasemi HR. Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir region, Saudi Arabia. Geoscience Frontiers. 2021;12:639-655. DOI:10.1016/j.gsf.2020.05.010

[7] Thi Ngo PT, Panahi M, Khosravi K, Ghorbanzadeh O, Kariminejad N, Cerda A, Lee S. Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geoscience Frontiers. 2021;12:505-519. DOI:10.1016/j.gsf.2020.06.013

[8] Yi Y, Zhang Z, Zhang W, Jia H, Zhang J. Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. Catena. 2020;195:104851. DOI:10.1016/j.catena.2020.104851

[9] Merghadi A, Yunus AP, Dou J, Whiteley J, Pham BT, Bui DT, Avtar R, Abderrahmane B. Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews. 2020;207:103225. DOI:10.1016/j.earscirev.2020.103225

[10] Tang Y, Feng F, Guo Z, Feng W, Li Z, Wang J, Sun Q, Ma H, Li Y. Integrating principal component analysis with statistically-based models for analysis of causal factors and landslide susceptibility mapping: A comparative study from the loess plateau area in Shanxi (China). Journal of Cleaner Production. 2020;277:124159. DOI:10.1016/j.jclepro.2020.124159

[11] Trigila A, Iadanza C, Esposito C, Scarascia-Mugnozza G. Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). Geomorphology. 2015;249:119-136. DOI:10.1016/j.geomorph.2015.06.001

[12] Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM. Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir region, Saudi Arabia. Landslides. 2016;13(5):839-856. DOI:10.1007/s10346-015-0614-1

[13] Pham BT, Shirzadi A, Tien Bui D, Prakash I, Dholakia MB. A hybrid machine learning ensemble approach based on a radial basis function neural network and rotation forest for landslide susceptibility modeling: A case study in the Himalayan area, India. International Journal of Sediment Research. 2018;33(2):157-170. DOI:10.1016/j.ijsrc.2017.09.008

[14] Wu Y, Ke Y, Chen Z, Liang S, Zhao H, Hong H. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. Catena. 2020;187:104396. DOI:10.1016/j.catena.2019.104396

[15] Luo W, Liu CC. Innovative landslide susceptibility mapping supported by geomorphon and geographical detector methods. Landslides. 2018;15:465-474. DOI:10.1007/s10346-017-0893-9

[16] Chen W, Xie X, Peng J, Wang J, Duan Z, Hong H. GIS-based landslide susceptibility modelling: A comparative assessment of kernel logistic regression, Naïve-Bayes tree, and alternating decision tree models. Geomatics, Natural Hazards and Risk. 2017;8(2):950-973. DOI:10.1080/19475705.2017.1289250

[17] Wang G, Chen X, Chen W. Spatial prediction of landslide susceptibility based on GIS and discriminant functions. ISPRS International Journal of Geo-Information. 2020;9(3):144. DOI:10.3390/ijgi9030144

[18] Kavzoglu T, Sahin EK, Colkesen I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides. 2014;11(3):425-439. DOI:10.1007/s10346-013-0391-7

[19] Kumar D, Thakur M, Dubey CS, Shukla DP. Landslide susceptibility mapping & prediction using support vector machine for Mandakini River basin, Garhwal Himalaya, India. Geomorphology. 2017;295:115-125. DOI:10.1016/j.geomorph.2017.06.013

[20] Bui DT, Pradhan B, Lofman O, Revhaug I. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naïve Bayes models. Mathematical Problems in Engineering. 2012;2012:974638. DOI:10.1155/2012/974638

[21] Pradhan B. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Computers and Geosciences. 2013;51:350-365. DOI:10.1016/j.cageo.2012.08.023

[22] Goetz JN, Brenning A, Petschko H, Leopold P. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers and Geosciences. 2015;81:1-11. DOI:10.1016/j.cageo.2015.04.007

[23] Pourghasemi HR, Rahmati O. Prediction of the landslide susceptibility: Which algorithm, which precision? Catena. 2018;162:177-192. DOI:10.1016/j.catena.2017.11.022

[24] Aditian A, Kubota T, Shinohara Y. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. Geomorphology. 2018;318:101-111. DOI:10.1016/j.geomorph.2018.06.006

[25] Sevgen E, Kocaman S, Nefeslioglu HA, Gokceoglu C. A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest. Sensors. 2019;19(18):3940. DOI:10.3390/s19183940.

[26] Pourghasemi HR, Gayen A, Park S, Lee C-W, Lee S. Assessment of landslide-prone areas and their zonation using logistic regression, LogitBoost, and NaïveBayes machine-learning algorithms. Sustainability. 2018;10:3697. DOI:10.3390/su10103697

[27] He Q, Shahabi H, Shirzadi A, Li S, Chen W, Wang N, Chai H, Bian H, Ma J, Chen Y, Wang X, Chapi K, Bin Ahmad B. Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF classifier, and RBF network machine learning algorithms. Science of the Total Environment. 2019;663:1-15. DOI:10.1016/j.scitotenv.2019.01.329

[28] Hu Q, Zhou Y, Wang S, Wang F. Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River basin. Geomorphology. 2020;351:106975. DOI:10.1016/j.geomorph.2019.106975

[29] Sahin EK. Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. SN Applied Sciences. 2020;2:1308. DOI:10.1007/s42452-020-3060-1

[30] Sahin EK. Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping. Geocarto International. DOI:10.1080/10106049.2020.1831623.

[31] TURKSTAT. Address based population registration system results. Population of Municipalities, Villages and Quarters. Turkish Statistical Institute. Available from: https://biruni.tuik.gov.tr/medas/?kn=95&locale=tr (Accessed: Mach 2, 2021)

[32] GDM. Meteorological Data Information Sales and Presentation System. General Directorate of Meteorology. Available from: https://mevbis.mgm.gov.tr/mevbis/ui/index.html (Accessed: April 14, 2021)

[33] Keskin I. 1:100,000 Scale Geological Map of Turkey, No:178 Artvin-E48 and F48 Map Sheet. General Directorate of Mineral Research and Exploration, Geological Research Department, Ankara, Turkey, 2013 (in Turkish).

[34] Erener A, Mutlu A, Düzgün HS. A comparative study for landslide susceptibility mapping using GIS-based multi-criteria decision analysis (MCDA), logistic regression (LR) and association rule mining (ARM). Engineering Geology. 2016;203:45-55. DOI:10.1016/j.enggeo.2015.09.007

[35] Parise M. Landslide mapping techniques and their use in the assessment of the landslide hazard. Physics and Chemistry of the Earth, Part C: Solar, Terrestrial and Planetary Science. 2001;26(9): 697-703. DOI:10.1016/S1464-1917(01)00069-1

[36] Bera S, Guru B, Ramesh V. Evaluation of landslide susceptibility models: A comparative study on the part of Western Ghat region, India. Remote Sensing Applications: Society and Environment. 2019;13:39-52. DOI:10.1016/j.rsase.2018.10.010

[37] Yanar T, Kocaman S, Gokceoglu C. Use of Mamdani fuzzy algorithm for multi-hazard susceptibility assessment in a developing urban settlement (Mamak, Ankara, Turkey). ISPRS International Journal of Geo-Information. 2020;9(2):114; DOI:10.3390/ijgi9020114

[38] Yan F, Zhang Q, Ye S, Ren B. A novel hybrid approach for landslide susceptibility mapping integrating analytical hierarchy process and normalized frequency ratio methods with the cloud model. Geomorphology. 2019;327:170-187. DOI:10.1016/j.geomorph.2018.10.024

[39] Bahrami S, Rahimzadeh B, Khaleghi S. Analyzing the effects of tectonic and lithology on the occurrence of landslide along Zagros ophiolitic suture: A case study of Sarv-Abad, Kurdistan, Iran. Bulletin of Engineering Geology and the Environment. 2020; 79:1619-1637. DOI:10.1007/s10064-019-01639-3

[40] Akinci H, Kilicoglu C, Dogan S. Random forest-based landslide susceptibility mapping in coastal regions of Artvin, Turkey. ISPRS International Journal of Geo-Information. 2020;9(9):553. DOI:10.3390/ijgi9090553

[41] Soma AS, Kubota T, Mizuno H. Optimization of causative factors using logistic regression and artificial neural network models for landslide susceptibility assessment in Ujung Loe watershed, South Sulawesi Indonesia. Journal of Mountain Science. 2019; 16(2):383-401. DOI:10.1007/s11629-018-4884-7

[42] Pourghasemi HR, Mohammady M, Pradhan B. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. Catena. 2012;97:71-84. DOI:10.1016/j.catena.2012.05.005

[43] Ba Q, Chen Y, Deng S, Wu Q, Yang J, Zhang J. An improved information value model based on gray clustering for landslide susceptibility mapping. ISPRS International Journal of Geo-Information. 2017;6(1):18. DOI:10.3390/ijgi6010018

[44] Arabameri A, Saha S, Roy J, Chen W, Blaschke T, Bui DT. Landslide susceptibility evaluation and management using different machine learning methods in the Gallicash River watershed, Iran. Remote Sensing. 2020;12(3):475. DOI:10.3390/rs12030475

[45] Youssef AM, Al-Kathery M, Pradhan B. Landslide susceptibility mapping at Al-hasher area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models. Geosciences Journal. 2015;19(1):113-134. DOI:10.1007/s12303-014-0032-8

[46] Ding Q, Chen W, Hong H. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. Geocarto International. 2017;32(6):619-639. DOI:10.1080/10106049.2016.1165294

[47] Demir G. GIS-based landslide susceptibility mapping for a part of the north Anatolian fault zone between Reşadiye and Koyulhisar (Turkey). Catena. 2019;183:104211. DOI:10.1016/j.catena.2019.104211

[48] Lee S, Choi J. Landslide susceptibility mapping using GIS and the weight-of-evidence model. International Journal of Geographical Information Science. 2004;18(8):789-814. DOI:10.1080/1365881041 0001702003

[49] Akgun A. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: A case study at İzmir, Turkey. Landslides. 2012;9(1):93-106. DOI:10.1007/s10346-011-0283-7

[50] Colkesen I, Sahin EK, Kavzoglu T. Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. Journal of African Earth Sciences. 2016;118:53-64. DOI:10.1016/j.jafrearsci.2016.02.019

[51] Chen W, Zhang S, Li R, Shahabi H. Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. Science of the Total Environment.

2018;644:1006-1018. DOI:10.1016/j.scitotenv.2018.06.389.

[52] Gómez H, Kavzoglu T. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River basin, Venezuela. Engineering Geology. 2005; 78(1-2):11-27. DOI:10.1016/j.enggeo.2004.10.004

[53] Catani F, Lagomarsino D, Segoni S, Tofani V. Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. Natural Hazards and Earth System Sciences. 2013;13(11):2815-2831. DOI:10.5194/nhess-13-2815-2013

[54] Akgun A, Erkan O. Landslide susceptibility mapping by geographical information system-based multivariate statistical and deterministic models: in an artificial reservoir area at northern Turkey. Arabian Journal of Geoscience. 2016;9:165. DOI:10.1007/s12517-015-2142-7

[55] Huang F, Cao Z, Guo J, Jiang SH, Li S, Guo Z. Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. Catena. 2020;191:104580. DOI:10.1016/j.catena.2020.104580

[56] Breiman L. Random forests. Machine Learning. 2001;45:5-32.

[57] Sun D, Wen H, Wang D, Xu J. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. Geomorphology. 2020;362:107201. DOI:10.1016/j.geomorph.2020.107201

[58] Kuhn M. Building predictive models in R using the caret package. Journal of Statistical Software. 2008; 28(5):1-26.

[59] Friedman JH. Greedy function approximation: A gradient boosting machine. The Annals of Statistics. 2001;29(5):1189-1232.

[60] Pourghasemi HR, Sadhasivam N, Amiri M, Eskandari S, Santosh M. Landslide susceptibility assessment and mapping using state-of-the art machine learning techniques. Natural Hazards. 2021; DOI:10.1007/s11069-021-04732-7

[61] Wang Z, Liu Q, Liu Y. Mapping landslide susceptibility using machine learning algorithms and GIS: A case study in Shexian County, Anhui Province, China. Symmetry. 2020;12:1954. DOI:10.3390/sym12121954

[62] Chen T, Guestrin C. XGBoost. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD'16; New York: ACM Press; 2016. p. 785-794.

[63] Can R, Kocaman S, Gokceoglu C. A comprehensive assessment of XGBoost algorithm for landslide susceptibility mapping in the upper basin of Ataturk dam, Turkey. Applied Sciences. 2021;11:4993. DOI:10.3390/app11114993

[64] Üstüner M, Abdikan S, Bilgin G, Balik Şanli F. Crop classification using light gradient boosting machines. Turkish Journal of Remote Sensing and GIS. 2020;1(2):97-105 (in Turkish).

[65] Achour Y, Pourghasemi HR. How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? Geoscience Frontiers. 2020;11(3): 871-883. DOI:10.1016/j.gsf.2019.10.001

[66] Zhang Yx, Lan Hx, Li Lp, Wu Ym, Chen Jh, Tian Nm. Optimizing the frequency ratio method for landslide susceptibility assessment: A case study of the Caiyuan Basin in the southeast mountainous area of China. Journal of Mountain Science. 2020;17(2):340-357. DOI:10.1007/s11629-019-5702-6

[67] Huang Y, Zhao L. Review on landslide susceptibility mapping using support vector machines. Catena.

2018;165:520-529. DOI:10.1016/j.catena.2018.03.003

[68] Tsangaratos P, Ilia I. Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size. Catena. 2016;145:164-179. DOI:10.1016/j.catena.2016.06.004

[69] Akinci H, Zeybek M. Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey. Natural Hazards. 2021; DOI:10.1007/s11069-021-04743-4

[70] Althuwaynee OF, Pradhan B, Lee S. Application of an evidential belief function model in landslide susceptibility mapping. Computers and Geosciences. 2012;44:120-135. DOI:10.1016/j.cageo.2012.03.003

[71] Akgun A, Sezer EA, Nefeslioglu HA, Gokceoglu C, Pradhan B. An easy-to-use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm. Computers and Geosciences. 2012; 38:23-34. DOI:10.1016/j.cageo.2011.04.012

[72] Chen W, Li W, Hou E, Zhao Z, Deng N, Bai H, Wang D. Landslide susceptibility mapping based on GIS and information value model for the Chencang District of Baoji, China. Arabian Journal of Geosciences. 2014;7(11):4499-4511. DOI:10.1007/s12517-014-1369-z

[73] Zhao Y, Wang R, Jiang Y, Liu H, Wei Z. GIS-based logistic regression for rainfall-induced landslide susceptibility mapping under different grid sizes in Yueqing, Southeastern China. Engineering Geology. 2019; 259: 105147. DOI:10.1016/j.enggeo.2019.105147.

[74] Chen W, Sun Z, Zhao X, Lei X, Shirzadi A, Shahabi H. Performance evaluation and comparison of bivariate statistical-based artificial intelligence algorithms for spatial prediction of landslides. ISPRS International Journal of Geo-Information. 2020; 9:696. DOI:10.3390/ijgi9120696.

[75] Chung CJF, Fabbri AG. Validation of spatial prediction models for landslide hazard mapping. Natural Hazards. 2003;30(3):451-472. DOI:10.1023/B:NHAZ.0000007172.62651.2b