Article

Comparison of Support Vector Machine, Bayesian Logistic Regression, and Alternating Decision Tree Algorithms for Shallow Landslide Susceptibility Mapping along a Mountainous Road in the West of Iran

Viet-Ha Nhu [1,2] , **Danesh Zandi** [3] , **Himan Shahabi** [3,4,*] , **Kamran Chapi** [5] ,
Ataollah Shirzadi [5] , **Nadhir Al-Ansari** [6,*] , **Sushant K. Singh** [7] , **Jie Dou** [8]
and Hoang Nguyen [9]

1   Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam; nhuvietha@tdtu.edu.vn
2   Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam
3   Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran; daneshjan@gmail.com
4   Board Member of Department of Zrebar Lake Environmental Research, Kurdistan Studies Institute, University of Kurdistan, Sanandaj 66177-15175, Iran
5   Department of Rangeland and Watershed Management, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran; k.chapi@uok.ac.ir (K.C.); a.shirzadi@uok.ac.ir (A.S.)
6   Department of Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 97187 Lulea, Sweden
7   Virtusa Corporation, 10 Marshall Street, Irvington, NJ 07111, USA; sushantorama@gmail.com
8   Department of Civil and Environmental Engineering, Nagaoka University of Technology, 1603-1, Kami-Tomioka, Nagaoka, Niigata 940-2188, Japan; douj888@gmail.com
9   Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; nguyenhoang23@duytan.edu.vn
*   Correspondence: h.shahabi@uok.ac.ir (H.S.); nadhir.alansari@ltu.se (N.A.-A.)

check for updates

Abstract: This paper aims to apply and compare the performance of the three machine learning algorithms–support vector machine (SVM), bayesian logistic regression (BLR), and alternating decision tree (ADTree)–to map landslide susceptibility along the mountainous road of the Salavat Abad saddle, Kurdistan province, Iran. We identified 66 shallow landslide locations, based on field surveys, by recording the locations of the landslides by a global position System (GPS), Google Earth imagery and black-and-white aerial photographs (scale 1: 20,000) and 19 landslide conditioning factors, then tested these factors using the information gain ratio (IGR) technique. We checked the validity of the models using statistical metrics, including sensitivity, specificity, accuracy, kappa, root mean square error (RMSE), and area under the receiver operating characteristic curve (AUC). We found that, although all three machine learning algorithms yielded excellent performance, the SVM algorithm (AUC = 0.984) slightly outperformed the BLR (AUC = 0.980), and ADTree (AUC = 0.977) algorithms. We observed that not only all three algorithms are useful and effective tools for identifying shallow landslide-prone areas but also the BLR algorithm can be used such as the SVM algorithm as a soft computing benchmark algorithm to check the performance of the models in future.

1. Introduction

A landslide is defined as the movement of the slope covers, including soil, rock, and organic materials, under the influence of a gravitational force down the slope [1]. Among natural hazards (e.g., pollution, flooding, earthquakes, and landslides), landslides, the topic of this paper, rank seventh globally in terms of death and economic impact [2], including damage to roads rail lines, power lines, and touristic and historical [3–5]. Landslides can significantly affect the geomorphic evolution of the landscape that create some geological disasters throughout the world [6].

Iran is one such country; nearly 4900 destructive landslides recorded in the country up to the end of September 2007, causing approximately USD 12.7 billion (126,893 billion Iranian Rials) damage [7,8]. Between 2500 and 4000, people have died in landslides between AD 763 and 2016 [9]. Slumps are the most common mass movements in the region [9]. Debris and rock avalanches are relatively uncommon but are responsible for much loss of life. Landslides are particularly frequent in the Alborz and Zagros Mountains [4,9]. The risk posed by landslides in these areas is amplified by the inadequate scientific knowledge and resources aimed at dealing with the problem [5,10]. The mountainous road of Salvat Abad saddle is one of the vital links of the strategic road network in Kurdistan province of Iran. This road corridor was severely affected by several landslides every year that caused damages to tens of thousands of dollars every year [11].

Considering the damage from the landslides particularly along the roads, monitoring the areas prone to landslide occurrence by preparing landslide susceptibility mapping is a mandatory for preventing and mitigating the most vulnerable areas [4,12,13]. According to the definition, landslide susceptibility refers to preventing and mitigating risk posed to the most vulnerable areas [14–16]. This map can specify the landslide prone areas with degrees of susceptibility so that land managers, governments, environmental planners, and policy makers can manage these areas before a catastrophic landslide. Indeed, the areas along the road, Salavat Abad, are classified into the same susceptibility classes to introduce these areas to implementing organizations, in order to manage them through control of the landslide before it occurs.

Many methods have been developed for landslide susceptibility mapping, for example (1) expert knowledge-based models such as analytical hierarchy process (AHP) [17], PROMETHEE II, and fuzzy AHP [18]; (2) bivariate and multivariate statistical models such as frequency ratio (FR) [19–21]; index of entropy (IOE) [22–24], weighted linear combination (WLC) [25], certainty factors (CF) [26,27], and logistic regression (LR) [15,28]; (3) deterministic (physical-based) models such as Stability Index Mapping (SINMAP) [29], Shallow Landsliding Stability Model (SHALSTAB) [30], Self-Organized Slope (SOSlope) [31], PRobabilistIc MUltidimensional shallow Landslide Analysis (PRIMULA) [32], SHETRAN [33], Transient Rainfall Infiltration and Grid-based Regional Slope-Stability Model (TRIGRS) [34], and (4) machine learning models. Literatures show that AHP, LR, and SVM are the most commonly used methods [35]. Expert knowledge-based models are typically based on questionnaires and expert opinions that may differ from one expert to another and also may suffer from cognitive limitations centered around uncertainty and subjectivity [36]. Bivariate statistical methods, although robust and flexible, lack sensitivity in their analysis of conditioning factors and also oversimplify input data [37]. Multivariate methods allow users to order parametric importance before using modeling, but require a more profound knowledge of mathematics, statistics, and software [38].

Recently, machine learning methods have gained popularity over expert knowledge-based and bivariate and multivariate methods in studies of natural hazards because of their objectivity and high accuracy. A wide variety of machine learning models are now in use in natural hazard research, including: artificial neural network (ANN) [39,40], adaptive neuro-fuzzy inference system (ANFIS) [41–44], support vector machine (SVM) [15,45,46], K-nearest neighbor (KNN) [47], logistic model tree (LMT) [48,49], alternating decision tree (ADTree) [50–53], random subspace (RS) [54], credal decision tree (CDT) [14], quantile regression (QR) [55], radial basis function (RBF) [56], stochastic gradient descent (SGD) [57]. classification and regression tree (CART) [58], J48 decision tree [59], reduced error pruning tree (REPT) [60–62], reduced error pruning tree (REPTree) [61,63], random

forest (RF) [64], naïve bayes tree (NBT) [11,15], baysian logistic regression (BLR) [65], Fisher's linear discriminant function (FLDA) [66], bayes net (BN) [67], grey wolf optimizer (GWO) [68,69], naïve bayes (NB) [70]; naïve Bayes tree (NBTree) [71–73], evidential belief function (EBF) [74], and kernel logistic regression (KLR) [75].

　　All machine learning algorithms must be tested and validated in landslide-prone areas to select those with the highest performance and prediction accuracy. Therefore, the main aim of this study is to compare the efficiency of BLR, SVM, and ADTree algorithms to landslide susceptibility along a road section using in Kurdistan province, Iran. BLR is an algorithm that is a combination of a base-based theory algorithm and a logistic regression function. However, the ADTree is a decision tree algorithm. Its performance on landslide modeling and susceptibility modeling has been earlier confirmed and suggested [51,53,76–79]. Therefore, in this study, we aim to compare the performance of a functional-based algorithm, SVM, a bayes-based theory algorithm, BLR, with a decision tree-based algorithm, ADTree, for shallow landslide susceptibility modeling in the study area. SVM, in particular, can handle complex and non-linear datasets [35], and thus, is a robust benchmark model that has been successfully used in landslide susceptibility mapping. This study is a pioneering step in the application of advanced predictive machine learning algorithms in landslide susceptibility research in the study area Another objective is to check the ability of the BLR and ADTree algorithm as the benchmark models, such as the SVM in landslide susceptibility mapping.

2. Study Area

　　The Salavat Abad saddle is located in southwest Kurdistan province, Iran (Figure 1). The study area covers about 18.7 km^2 and ranges in elevation from 1699 to 2500 m above sea level [19]. A road through the saddle, which connects Sanandaj City to Tehran, the capital of Iran, has strategic, economic, and socio-cultural importance. Much of Kurdistan province is located in the Zagros Mountains, a tectonically active range dominated by sedimentary and volcanic rocks [80].
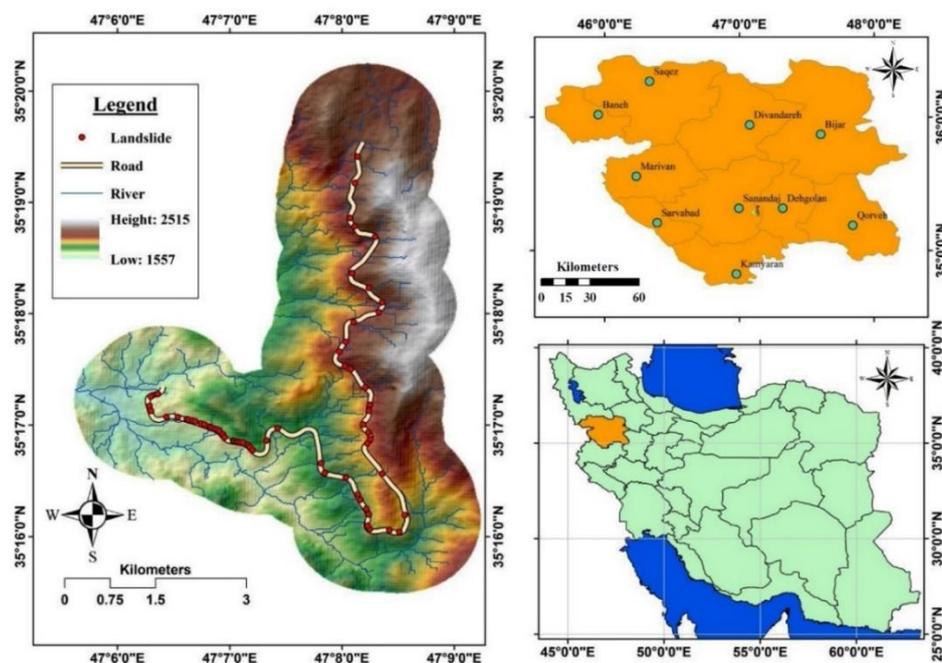

Figure 1. The study area and its location in Kurdistan province (upper right) in northwest Iran (lower right).

　　The climate of the study area is influenced by warm Mediterranean air masses, resulting in rainfall and snowfall in winter, with an average precipitation of about 470 mm [19]. Many costly and fatal mass movements occur in the winter season.

3. Data Acquisition

3.1. Landslide Inventory Map

The dataset for this study comprises 66 landslides previously mapped by the Forest, Rangeland, and Watershed Management Organization of Iran [19]. We examined the landslides by reviewing aerial photographs (1:40,000 scale) and Google Earth image, and by inspection in the field. Most of the landslides are the result of the slope modification of slopes due to road construction (Figure 2). In this study, landslide bodies were converted into the points (central points) and each polygon of landslides was considered as one landslide location that was applied for the modeling procedure.
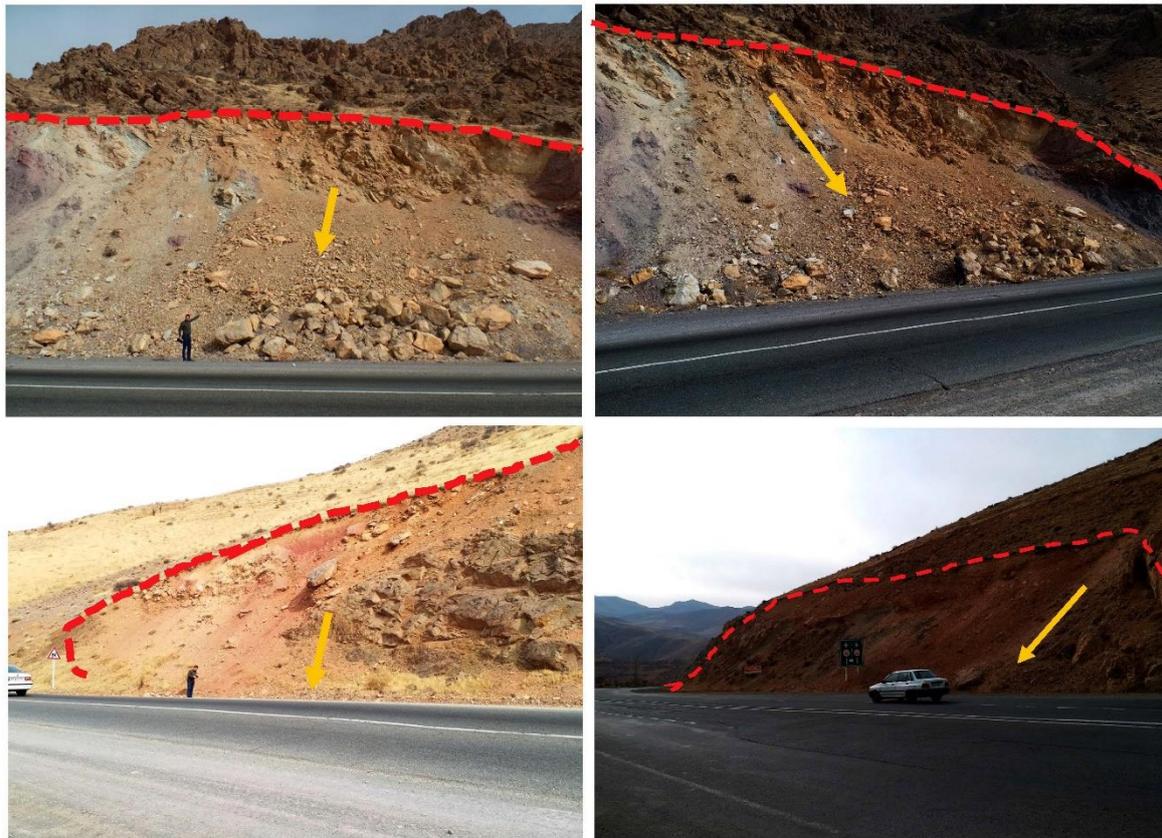
Figure 2. Examples of shallow landslides in the study area.

3.2. Landslide Conditioning Factors

Based on the literature, data availability, and our experience, we selected 18 landslide conditioning factors for modelling: Slope angle, slope aspect, elevation, distance to road, topographic wetness index (TWI), normalized difference vegetation index (NDVI), lithology, land use/land cover, rainfall, distance to fault, plan curvature, profile curvature, slope length-angle index (LS), solar radiation, stream power index (SPI), distance to the river, river density, and fault density. The factors are described briefly in the following subsections:

3.2.1. Slope Angle

A map of slope angles was extracted from DEM, and values were then grouped into five categories (0–30, 31–46, 47–56, 59–66, and >66°) using a natural break classification method. Most landslides occurred on slopes steeper than 47° (Figure 3a).
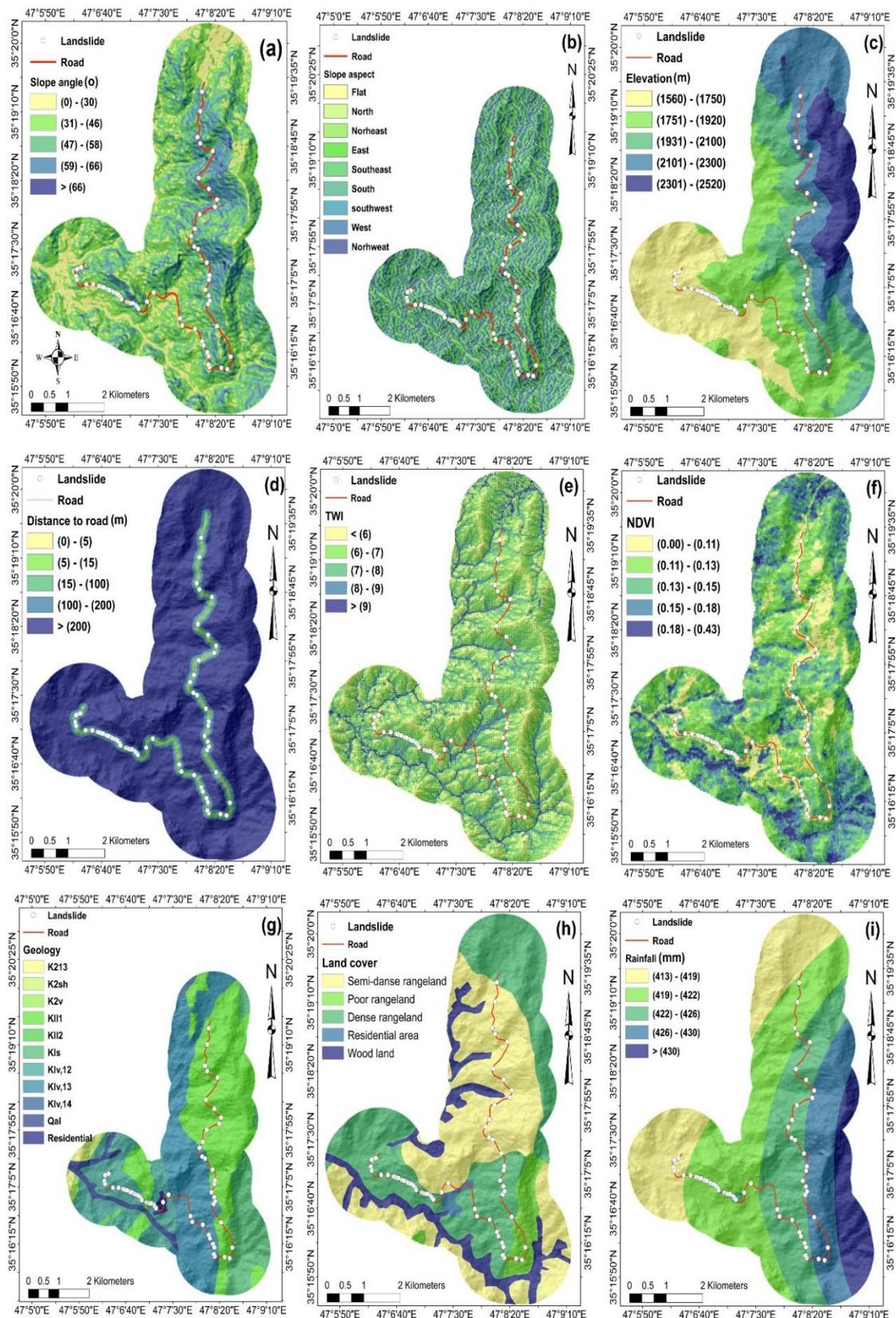
Figure 3. *Cont*.
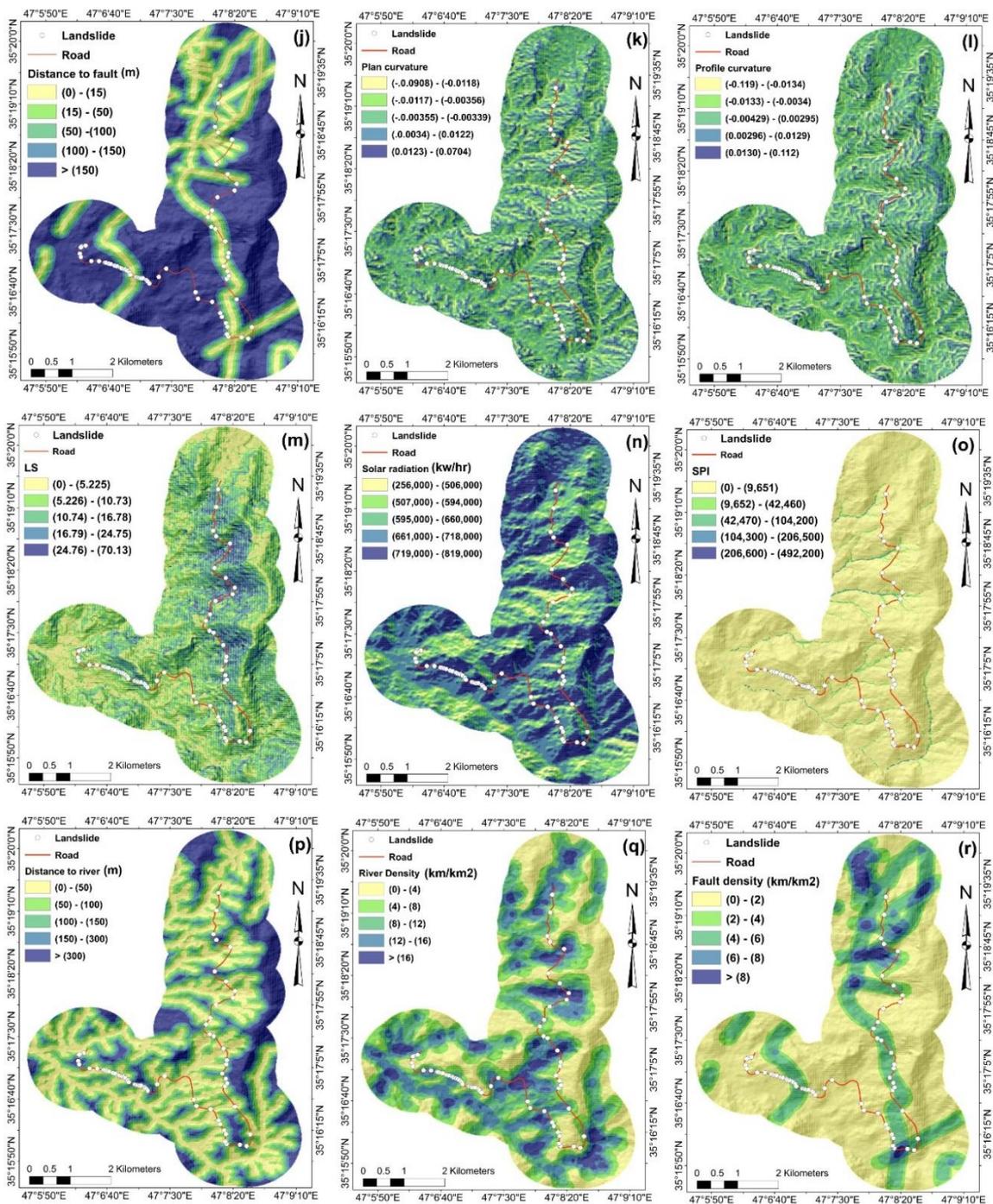
Figure 3. Spatial database for landslide susceptibility analysis: (**a**) slope, (**b**) aspect, (**c**) elevation, (**d**) distance to road, (**e**) TWI, (**f**) NDVI, (**g**) lithology, (**h**) land cover, (**i**) rainfall, (**j**) distance to fault, (**k**) plan curvature, (**l**) profile curvature, (**m**) LS, (**n**) solar radiation, (**o**) SPI, (**p**) distance to river, (**q**) river density, (**r**) fault density.

3.2.2. Slope Aspect

The slope aspect is defined as the cardinal direction of the maximum slope [81]. We extracted nine slope aspect classes from the DEM with a resolution of 12.5 m, obtained from Advanced Land Observing Satellite (ALOS) Phased Array L-type Synthetic Aperture Radar (PALSAR): North, northeast, northwest, east, southeast, southwest, south, west, and flat. Most landslides are located on the southwest- and east-facing slopes (Figure 3b).

3.2.3. Elevation

An elevation map was extracted from DEM, and values placed in five categories using the natural break classification method: 1557–1751, 1751–1917, 1917–2096, 2096–2300, and 2300–2515 m asl. Nearly half of the landslides are located in the lowest elevation class (Figure 3c).

3.2.4. Distance to Road

Distances to the arterial road were separated into five categories using the "Euclidean distance tool" in ArcGIS 10.2: 0–50, 50–100, 100–150, 150–200, and >200 m (Figure 3d).

3.2.5. Topographic Wetness Index

The TWI introduced by Beven and Kirkby [82] in rainfall-runoff modeling to identify the impact of topography and wetness on rates of runoff. It can be computed as follows,

$$TWI = Ln(\frac{\chi}{\tan \gamma}) \tag{1}$$

where χ is the specific catchment area and γ is the slope angle (in degree). We created a TWI layer and defined five categories using the natural break classification method: <6, 6–7, 7–8, 8–9, and >9. Most landslides are within the TWI > 9 class (Figure 3e).

3.2.6. Normalized Difference Vegetation Index

Normalized difference vegetation index (NDVI) provides a measure of vegetation within an area [57]. NDVI can be formulated as follows,

$$NDVI = \frac{(NIR(Band4) - Red(Band3))}{(NIR(Band4) + Red(Band3))} \tag{2}$$

where Red and NIR are the red and near-infrared bands, respectively. The NDVI map was generated using Landsat 8 OLI from 2017. Our NDVI map is shown in Figure 3f.

3.2.7. Lithology

A geology map of the study area at a scale of 1:100,000, obtained from Geological Surveys of Iran (GSI). The description of the lithological units are shown in Table 1, and its categories are shown in Figure 3g.

3.2.8. Land Cover/Land Use

Our field survey indicated that most of the landslides in the study area have happened near the road where the vegetation has been removed. In this study, we extracted a land cover/land use map from the Kurdistan province land cover map printed at a scale of 1:100,000 (Figure 3h).

3.2.9. Rainfall

A mean annual rainfall map was prepared from data acquired from eight meteorological stations in and around the study area using the IDW (Inverse distance weighted) interpolation method. We defined five categories: 413–419, 419–422, 422–426, 426–430, and >430 mm (Figure 3i).

3.2.10. Distance to Fault

A map of fault distances was extracted from the geology map. We defined five categories based on the manual classification method: 0–50, 50–100, 100–150, 150–200, and >200 m (Figure 3j).

Table 1. Lithological units and its description of the study area.

	Lithological Unit	Description	Age Era	Age Period
1	Kll1	Gray and light gray, thick-bedded to massive, fetid orbitolina bearing limestone.	MESOZOIC	Early Cretaceous
2	Kll2	Thick-bedded, gray to dark gray, rudist and orbitolina bearing limestone.	MESOZOIC	Early Cretaceous
3	Klv,12	Basaltic and andesitic volcanics, tuff, volcanic breccia and calcareous shale with intercalations of limestone.	MESOZOIC	Early Cretaceous
4	K213	Gray and light gray micritic limestone and radioarian limestone with calcschiste structure.	MESOZOIC	Late Cretaceous
5	K2v	Basalt, andesite, spilitic basalte and pyroclastic rocks with developed layering and climbing.	MESOZOIC	Late Cretaceous
6	Klv,13	Spilitic basalt, basalt and andesite lava with intercalation of red, blue, gray limestone, red shale and sandstone.	MESOZOIC	Early Cretaceous
7	Klv,14	Tuff, green tuff, andesite and andesitic dacite, shale, limestone and sandstone.	MESOZOIC	Early Cretaceous
8	Kls	Purpel to red medium to thick-bedded sandstone with intercalation of polymictic conglomerate and silty shale.	MESOZOIC	Early Cretaceous
9	K2sh	Black, dark gray, yellow shale, silty shale and phillitic shale with minor sandstone and micritic limestone intercalations. (Sanandaj shale)	MESOZOIC	Late Cretaceous
10	Qal	Recent alluvium (alluvial channel deposits).	CENOZOIC	Quaternary
11	Residential area	Salavat Saddle		

3.2.11. Plan Curvature

Plan curvature provides a measure of convergence or divergence of runoff on slopes [83]. Values can be positive (concave curvature), negative (convex curvature), or zero (flat slopes. A plan curvature layer was extracted from the DEM and divided into five categories using the natural break classification method: [(−0.0908)–(−0.0118)], [(−0.0118)–(−0.00355)], [(−0.00355)–(−0.0034)], [0.0034–0.0122], and [0.0122–0.0704] (Figure 3k).

3.2.12. Profile Curvature

Profile curvature can affect the velocity of runoff and thus erosion [48]. We extracted the profile curvature from the DEM and created five categories using the natural break classification method: [(−0.119)–(−0.0134)], [(−0.0134)–(−0.0034)], [(−0.00429)–0.00296], [0.00296–0.0129], and [0.0129–0.112] (Figure 3l).

3.2.13. Slope Length-Angle Index (LS)

The LS index was obtained by summing the slope length (L) and slope angle (S). It was extracted from the DEM using SGAG software. We defined five LS categories using the natural breaks classification method: 0–5.225, 5.226–10.73, 10.74–16.78, 16.79–24.75, and 24.76–70.13 (Figure 3m).

3.2.14. Solar Radiation

Solar radiation was extracted from DEM in ArcGIS using the "Area solar radiation" tool and then grouped into five categories: 256,000–506,000, 507,000–594,000, 595,000–660,000, 661,000–718,000, and 719,000–819,000 kw/hr (Figure 3n).

3.2.15. Stream Power Index (SPI)

SPI can be formulated as follows [84],

$$SPI = A_r \tan \gamma \tag{3}$$

where A_r is the specific catchment area and γ is the slope angle. We created an SPI map from the DEM in SGAG software and then defined five categories based on the natural breaks classification method: 0–9651, 9652–42,460, 42,470–104,200, 104,300–206,500, and 206,600–492,200 (Figure 3o).

3.2.16. Distance to the River

A layer of river distance was created based on mapped rivers in the study area. We defined five categories based on the manual classification method: 0–50, 50–100, 100–150, 150–200, and >200 m (Figure 3p).

3.2.17. River Density

Our river density map shows the total length of the river per km². Five categories were established based on the natural break classification method: 0–4, 4–8, 8–12, 12–16, and >16 km/km² (Figure 3q).

3.2.18. Fault Density

The fault density matrix is defined as the total length of the faults within a standard area of 1 km² [85]. We prepared this map based on mapped faults and created five categories using the natural break classification method: 0–2, 2–4, 4–6, 6–8, and >8 km/km² (Figure 3r).

4. Machine Learning Algorithms

4.1. Support Vector Machine

Support vector machine (SVM) is one of the machine learning methods used for classification and regression [86]. The main objective of the algorithm is to classify data with the highest confidence margin using linear data sorting. It maps input data to a much higher level using the Phi function on a training dataset. A linear equation called the 'surface separator' separates the data into two classes (in the case of this study, landslide, and non-landslide). SVM minimizes error by classifying and separating data with the help of a separator-hyperplane. Training points near the line of separation are termed 'surface vectors' [87].

Consider set X_i, which includes linear training data i = (0, 2, 3, ... , n), referred to as training vectors. The training vectors contain two classes denoted by $y_i = \pm 1$. The support vector machine maximizes the two datasets by finding an n-dimensional hyperplane (Figure 4), expressed as follows,

$$\frac{1}{2}\|w\|^2 \tag{4}$$

with the following condition,

$$y_i((w.x_i) + b) \geq 1 \tag{5}$$

where w is the normal separator hyperplane, b is a scalable datum, and (.) signifies a multiplication operation. The following is obtained using Lagrangian coefficients of cost,

$$L = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \lambda_i(y_i((w.x_i) + b) - 1) \tag{6}$$

where λ_i is the Lagrangian multiplier. Equation (7) can be minimized by using the w and b ratios as a standard. For cases that are noisy and indistinguishable (Figure 4), a variable ξ_i can be used as a weak meaning (slack variables ξ_i), in which case Equation (8) becomes:

$$y_i((w.x_i) + b) \geq 1 - \xi_i \tag{7}$$

$$L = \frac{1}{2}\|W\|^2 - \frac{1}{vn}\sum_{i=1}^{n} \xi_i. \tag{8}$$
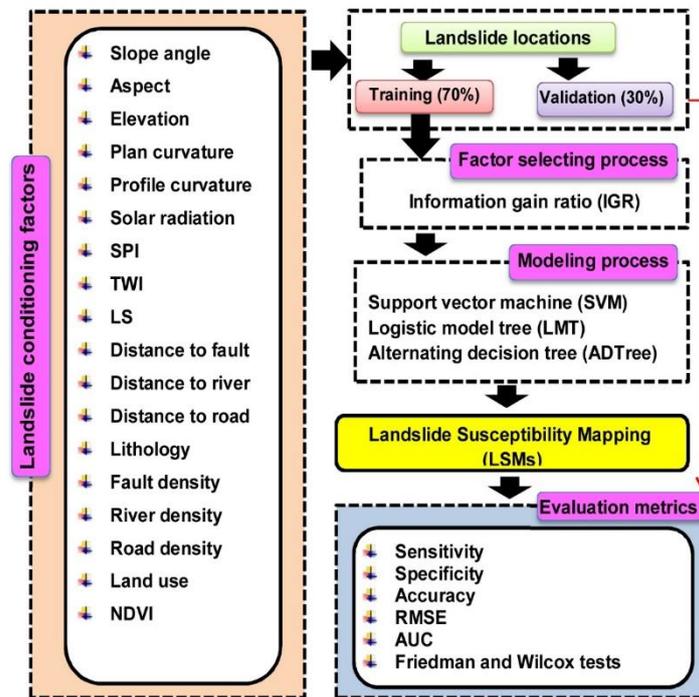
Figure 4. Flowchart of the methodology used in this study.

4.2. Bayesian Logistic Regression

Bayesian logistic regression (BLR) has been used with a two-state dependent variable about effective factors of landslides [88]. With this method, a logistic regression model is created based on the relations between dependent and independent variables. Then, a Bayesian function is applied based on the behavior and response of the effective factors using a prior probability function [88]. A Bayesian function is created in three stages, as follows [88]:

(1) Determine the prior probability of parameters
(2) Determine the likelihood function for data
(3) Create a posterior distribution function for parameters

If x is a training dataset and $x = (x_1, x_2, \ldots, x_n)$, landslide conditioning factors, and y = (y1, y2) is a dependent variable (landslides and non-landslides), a logistic function obtains the posterior probability function for samples belonging to a specific class,

$$P(\text{Class}|x_1, x_2, \ldots, x_n) = \frac{1}{\left(1 + \exp^{(b + w_0 * c + \sum_{i=1}^{n} w_i * f(x_i))}\right)} \tag{9}$$

where x_i are the effective factors, c is the prior log odds ratio ($c = \log \frac{P(\text{class}=0)}{P(\text{class}=1)}$), and b is the bias. w_0 and w_i are the weights trained by training data, and i_{th} factors of xi are used to calculate the $f(x_i)$ function using $\log \frac{P(x_i|\text{class}=0)}{P(x_i|\text{class}=1)}$ (for binary variables). A prior univariate Gaussian function is used to calculate weights in Bayesian-logistic regression model,

$$P(W|\sigma_i) = N(0, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp^{\frac{-(w^2)}{2\sigma_I}} \tag{10}$$

where '0' and 'σ_i' are the data average, and variance, respectively [89].

4.3. Alternating Decision Tree

The alternating decision tree (ADTree) algorithm uses the rules of a tree algorithm for classification by combining tree and boosting algorithms [90]. ADTree identifies and eliminates gaps among the tree and boosting algorithms. The algorithm includes decision and prediction nodes. The decision node expresses a situation, and the prediction node includes a numerical value [91]. ADTree first searches the best constant prediction coefficient for the training data in the stem of the tree. The tree is then grown based on the repetition of data, using the boosting algorithm, and a new rule is added. Next, a decision node and two prediction nodes are created [90]. Then, the algorithm allocates weight to each prediction node so that its predictability can be calculated by summing all weights [92].

(x_1, y_1), ..., (x_m, y_m), are pixels in the training data, $x_i \in R^d$, and y_i is the equivalence of occurrence and non-occurrence of landslides. The boosting algorithm grows the tree, with each repetition (t) supporting two conditions–a precondition (Pt) and a group of rules (Rt). A group of major conditions, C, is created by the weak algorithm in each repetition of the boosting algorithm. The algorithm works as follows:

1- Initialization

Let *Rt* be correct for creating major rules, assuming a precondition (related to the selection of a prediction node for entering the algorithm) and the condition (related to the decision node in the stem of the tree). The first predicted amount is obtained by the following equation,

$$a = \frac{1}{2} \ln \frac{W_+(T)}{W_-(T)} \tag{11}$$

where $W_+(T)$ and $W_-(T)$ are the sums of positive, and negative weights, respectively, and they justify the C condition in the training data.

2- Pre-adjustment

The test samples are weighted again by $W_{i,1} = W_{i,0}e^{-\alpha y_i}$
where t = 1,2, ... , T

- Create a C group of rules by the weak algorithm using the weight-related to each training sample $W_{i,t}$.

- For each main precondition $c_1 \in P_t$ and each condition $c_1 \in P_t$ calculate:

$$Z_t(c_1, c_2) = 2\left(\sqrt{W + (c_1 \hat{c}_2) W - (c_1 \hat{c}_2)} + \sqrt{W + (c_1 \widehat{c}_2) W - (c_1 \widehat{c}_2)} \right) + W(\widehat{c}_1) \quad (12)$$

- Select c_1, c_2 with minima $Z_t(c_1, c_2)$ and run R_{t+1} and R_t through the adding R_t rule so that the precondition and condition are equal to, respectively, c_1 and c_2. Then predict the two prediction amounts:

$$\alpha = \frac{1}{2} \ln \frac{W_+(c_1 \hat{c}_2) + \varepsilon}{W_-(c_1 \hat{c}_2) + \varepsilon} \quad (13)$$

$$b = \frac{1}{2} \ln \frac{W_+(c_1 \widehat{c}_2) + \varepsilon}{W_-(c_1 \widehat{c}_2) + \varepsilon} \quad (14)$$

- Establish P_{t+1} and P_t by adding $c_1 \hat{c}_2$ and $c_1 \widehat{c}_2$
- Update the weights based on the following equation for each repeat:

$$W_{i,t+1} = W_{i,te^{-r_t(x_i)yi}} \quad (15)$$

3- Output

Sum all weights and all major rules $Rt + 1$:

$$\text{class}(x) = \text{sign}\left(\sum_{t=1}^{T} r_t(x) \right) \quad (16)$$

4.4. Multicollinearity Tests

The correlation between the factors increases the redundancy affecting the landslide modelling and the accuracy of the results. Therefore, the multi-collinearity test of the conditioning factors is necessary to analysis when evaluating landslide modeling and susceptibility. For this analysis, two measures, including the tolerance (TOL) (TOL=1 − R^2) and variance inflation factor (VIF) (VIF = 1/TOL), have been used in the multicollinearity test [93,94]. If TOL > 0.1 and VIF < 10, there is a correlation among the factors and the factor with having such information should be removed from the modeling process [48,75].

4.5. Selecting the Most Important Conditioning Factors by IGR

Several methods have been used to determine the importance of different factors for landslide occurrence, notably fuzzy-rough theory [95], relief algorithm [96], information gain, and information gain ratio [97]. Information gain specifies the amount of information that a factor can provide about the class. It selects factors with high levels of probability, and does not consider factors with low entropy level. This result is achieved using the IGR index, which was introduced by Quinlan in 1996 [98]. Effective factors for prediction have high IGR values. In this study, we evaluated the importance of our 18 conditioning factors using the Average merit (AM) of the IGR technique [62]. Average merit (AM) quantitatively determines the importance and ranking of factors [62]. The AM is the weight computed by the IGR feature selection technique.

Assume S is the training dataset with n input samples, and also that n (Li, S) is the number of training data in S belonging to Li class (landslide and non-landslide). Then:

$$Info(S) = \sum_{i=1}^{2} \frac{n(L_i, S)}{|S|} \log_2 \frac{n(L_i, S)}{|S|} \quad (17)$$

If we consider the factors impacting landslides, the needed information gain for dividing S into $(S_1, S_2, \ldots, S_m)$ is as follows:

$$Info(S,A) = -\sum_{j=1}^{m} \frac{S_j}{|S|} Info(S) \tag{18}$$

The following equation is used to calculate the information gain for each effective factor, for example, slope angle (A):

$$IGR(S,A) = \frac{Info(S) - Info(S,A)}{SplitInfo(S,A)} \tag{19}$$

SplitInfo is the information gained by the ratio of S training data to a subset with m items using the following equation:

$$SplitInfo(S,A) = -\sum_{j=1}^{m} \frac{S_j}{|S|} \log_2 \frac{S_j}{|S|} \tag{20}$$

4.6. Validation and Comparison of the Models

In this study, we evaluated model accuracy using the following metrics: Sensitivity, specificity, accuracy, kappa, root-mean-square deviation (RMSE), and area under the curve (AUC). There are four types of possible significance, i.e., true positive (TP), false positive (FP), true negative (TN), and false-negative (FN). TP is the number of expected landslides that are truly landslides. FP is the number of expected landslides that are non-landslides. TN denotes the number of expected non-landslides that are truly non-landslide, whereas FN is the number of non-landslides. Better predictive ability is indicated by higher values of sensitivity, specificity, AUC, and accuracy and the lower values of RMSE [21]. A kappa index value of 1 indicates an ideal model, whereas a value of −1 signifies a non-reliable model. The mentioned metrics are expressed as follows:

$$Sensitivity = \frac{TN}{TN + FP} \tag{21}$$

$$Specificity = \frac{TN}{TP + FN} \tag{22}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{23}$$

$$Kappa = \frac{O - E}{1 - E} \tag{24}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(X_{\text{predicted}} - X_{\text{actual}}\right)^2} \tag{25}$$

where O and E are the observed and expected agreement, respectively, $X_{\text{predicted}}$ and X_{actual} are the predicted and observed values of the ith instance from models, and n is the number of instances.

The receiver operating characteristic (ROC) curve has been used to test the overall performance of LSM methods [93]. The area under the ROC curve is the statistical summary of the overall performance of models [72]. The x, and y axes of the AUC are, respectively, the sensitivity and 100-specificity. The values of AUC range from 0 to 1, with values closer to 1 indicating a better predictive ability; an AUC value of 1 indicates perfect model performance [94]. The schematic diagram of methodology is illustrated in Figure 4.

5. Results and Analysis

5.1. Correlation between the Conditioning Factors

Table 2 shows the correlation between 18 conditioning factors. The results conclude that there is no correlated problem among the models, and all of them can be selected as inputs to modeling procedure by the machine learning algorithms.

Table 2. The multicollinearity tests of the factors.

Factors	TOL	VIF
Slope angle	0.520	1.587
Aspect	0.322	2.184
Elevation	0.212	1.135
Profile curvature	0.825	2.381
Plan curvature	0.705	1.843
Distance to road	0.553	2.529
NDVI	0.498	1.814
Land use	0.340	2.321
Lithology	0.263	1.311
LS	0.541	1.849
Rainfall	0.887	1.552
Solar radiation	0.670	2.698
TWI	0.776	1.541
SPI	0.732	1.873
Distance to river	0.820	2.987
River density	0.922	1.784
Distance to fault	0.712	2.835
Fault density	0.825	2.781

5.2. The Most Important Landslide Conditioning Factors in the Study Area

Each conditioning factor contributes differently to landslide occurrence. Therefore, their predictive power must be assessed. According to the values of the AM of IGR (AMIGR) with 10-fold cross-validation, only 10 factors of the 18 conditioning factors have predictive power (Table 3). We, therefore, used only those 10 factors when modeling. Distance to road (AMIGR = 0.1434) is the most significant factor, followed by NDVI (AMIGR = 0.0725), land use (AMIGR = 0.0187), slope aspect (AMIGR = 0.0139), lithology (AMIGR = 0.0097), slope angle (AMIGR = 0.0091), rainfall (AMIGR = 0.0090), distance to fault (AMIGR = 0.0087), elevation (AMIGR = 0.0078), and TWI (AMIGR = 0.0040) (Table 3).

Table 3. Ranks of significant landslide conditioning factors based on the IGR technique and the training dataset.

Conditioning Factor	Rank	AMIGR
Distance to road	1	0.1434
NDVI	2	0.0725
Land use	3	0.0187
Aspect	4	0.0139
Lithology	5	0.0097
Slope angle	6	0.0091
Rainfall	7	0.0090
Distance to fault	8	0.0087
Elevation	9	0.0078
TWI	10	0.0040

5.3. Landslide Modeling and Evaluation Process

After selecting significant conditioning factors, we performed the modeling process on the training dataset using SVM, BLR, and ADTree and then tested the results. The goodness-of-fit analysis indicates that all landslide models predict the spatial distribution of landslides well (Table 4), although the SVM model performed best with values of sensitivity, specificity, accuracy, kappa, RMSE, and AUC of, respectively, 0.943, 0.981, 0.962, 0.924, 0.192, and 0.986. The sensitivity and specificity values show that 94.3% of landslides were correctly classified as landslides and 98.1% of non-landslide locations were correctly classified as non-landslide locations. In comparison, the BLR model results yielded values of sensitivity, specificity, accuracy, kappa, RMSE, and AUC of, respectively, 0.906, 0.981, 0.943, 0.886, 0.237, and 0.943, and corresponding values for ADTree are 0.8887, 0.943, 0.915, 0.846, 0.245, and 0.912.

Table 4. Performances of the SVM, ADTree, and BLR algorithms on the training dataset.

Parameter	SVM	BLR	ADTree
TP	50	48	47
TN	52	52	50
FP	1	1	3
FN	3	5	6
Sensitivity	0.943	0.906	0.887
Specificity	0.981	0.981	0.943
Accuracy	0.962	0.943	0.915
Kappa	0.924	0.886	0.846
RMSE	0.192	0.237	0.245
AUC	0.986	0.943	0.912

Next, we assessed the predictive power of the models using the validation dataset (Table 5). The SVM model yielded the highest sensitivity (85.7%), followed by the BLR (78.6%) and ADTree (71.4%) models. The SVM model also provided the highest specificity value (91.7%), followed by the BLR (83.3%) and ADTree (81.8%) models. Accuracy values for the three models are 88.5% (SVM model), 80.8% (BLR), and 76.0% (ADTree). Kappa (0.869) and AUC (0.976) for the SVM model are greater than corresponding values for the BLR and ADTree models. Finally, the SVM model has the lowest RMSE (0.251), followed by the BLR (0.277) and ADTree (0.343) modes. Overall, the results from both the training and validation datasets show that the SVM model outclassed the BLR and ADTree models in predicting the locations of landslides in the study area.

Table 5. Performances of the SVM, ADTree, and BLR algorithms on the validation dataset.

Parameter	SVM	BLR	ADTree
TP	12	11	11
TN	11	10	9
FP	1	2	2
FN	2	3	4
Sensitivity	0.857	0.786	0.714
Specificity	0.917	0.833	0.818
Accuracy	0.885	0.808	0.760
Kappa	0.869	0.846	0.751
RMSE	0.251	0.277	0.343
AUC	0.976	0.923	0.910

In addition to the comparison of the performance of the models, based on the statistical-indexed base metrics, we assessed the efficiency of the three algorithms based on the CPU time during the modeling implementation. We concluded that in the SVM algorithm the CPU time to process by the training and validation datasets were 0.03 s; however, in the BLR this time for training dataset was 0.05 s and for validation dataset was 0.03 s. Moreover, the ADTree that had the lowest goodness-of-fit and prediction accuracy had 0.09 s and 0.06 s based on the training and the validation datasets, respectively.

5.4. Development of Landslide Susceptibility Maps

After training the SVM, BLR, and ADTree machine learning models with the training dataset and validating them with the validation dataset, we ran the models and obtained outputs as weights (landslide susceptibility indexes, LSIs). LSIs were assigned to each pixel of the study area to construct the landslide susceptibility maps. There are a variety of classification methods in ArcGIS, including manual, equal interval, natural break, quantile, geometric interval, and standard deviation used [16,35,98–100]. We selected the most classification used methods to create the landslide susceptibility maps such as natural breaks, quintile, and geometric intervals for reclassifying the LSIs. In the natural breaks method, no jump is detected in the values [101]. However, the quintile and geometric interval methods essentially split the distribution of susceptibility values into equal divisions, with similar proportions of the total area attributed to each class [102].

As selecting the method used to reclassify the LSIs depends on the LSI histogram [99,103,104], we prepare the histogram of the three mentioned methods based on the landslide pixel against susceptibility classes. A histogram is a better that most of landslide pixels have been placed in high susceptibility (HS) and very high susceptibility (VH) classes. Then, we chose the quintile classification method for the landslide susceptibility map derived using the SVM and ADTree models, and the geometrical interval method for the model based on the BLR (Figure 5). Each map has five susceptibility classes: very low susceptibility (VLS), low susceptibility (LS), moderate susceptibility (MS), high susceptibility (HS), and very high susceptibility (VHS) (Figure 6).
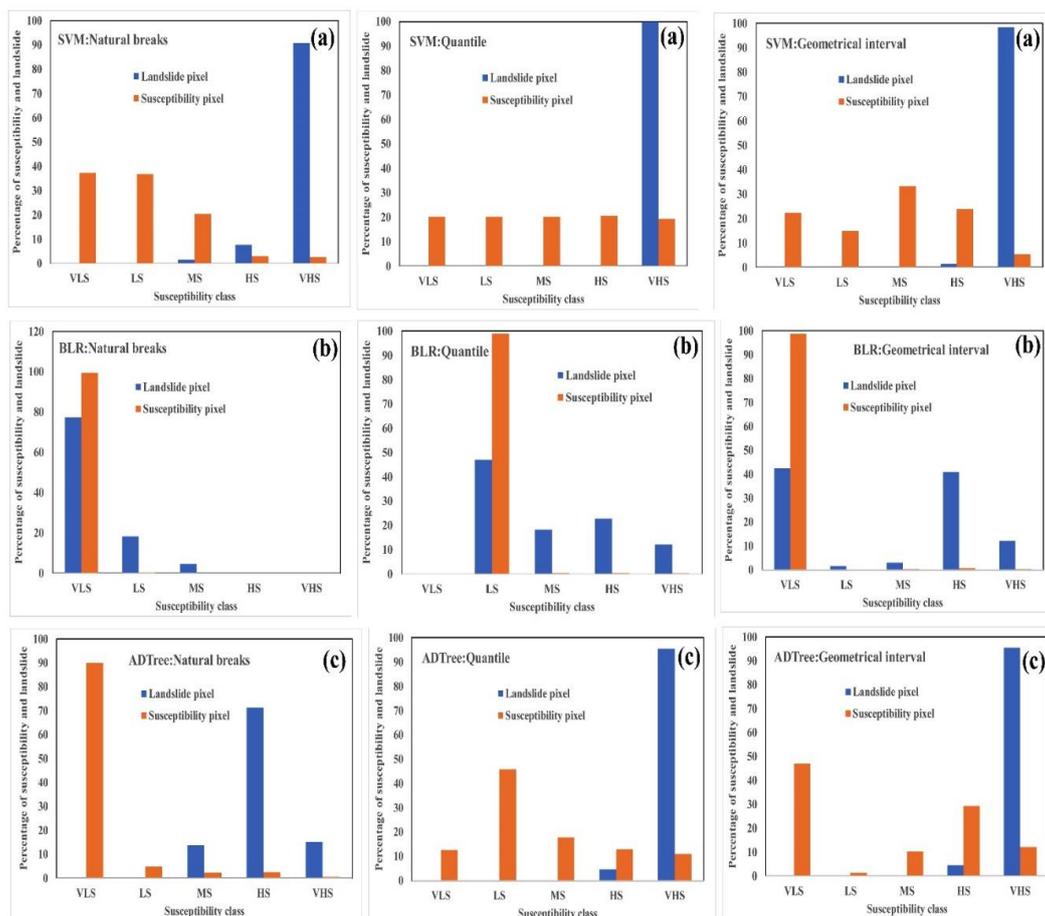

Figure 5. Histograms used to prepare the landslide susceptibility maps for three classification methods (natural break, quantile, and geometrical interval): (**a**) SVM, (**b**) BLR, (**c**) ADTree.
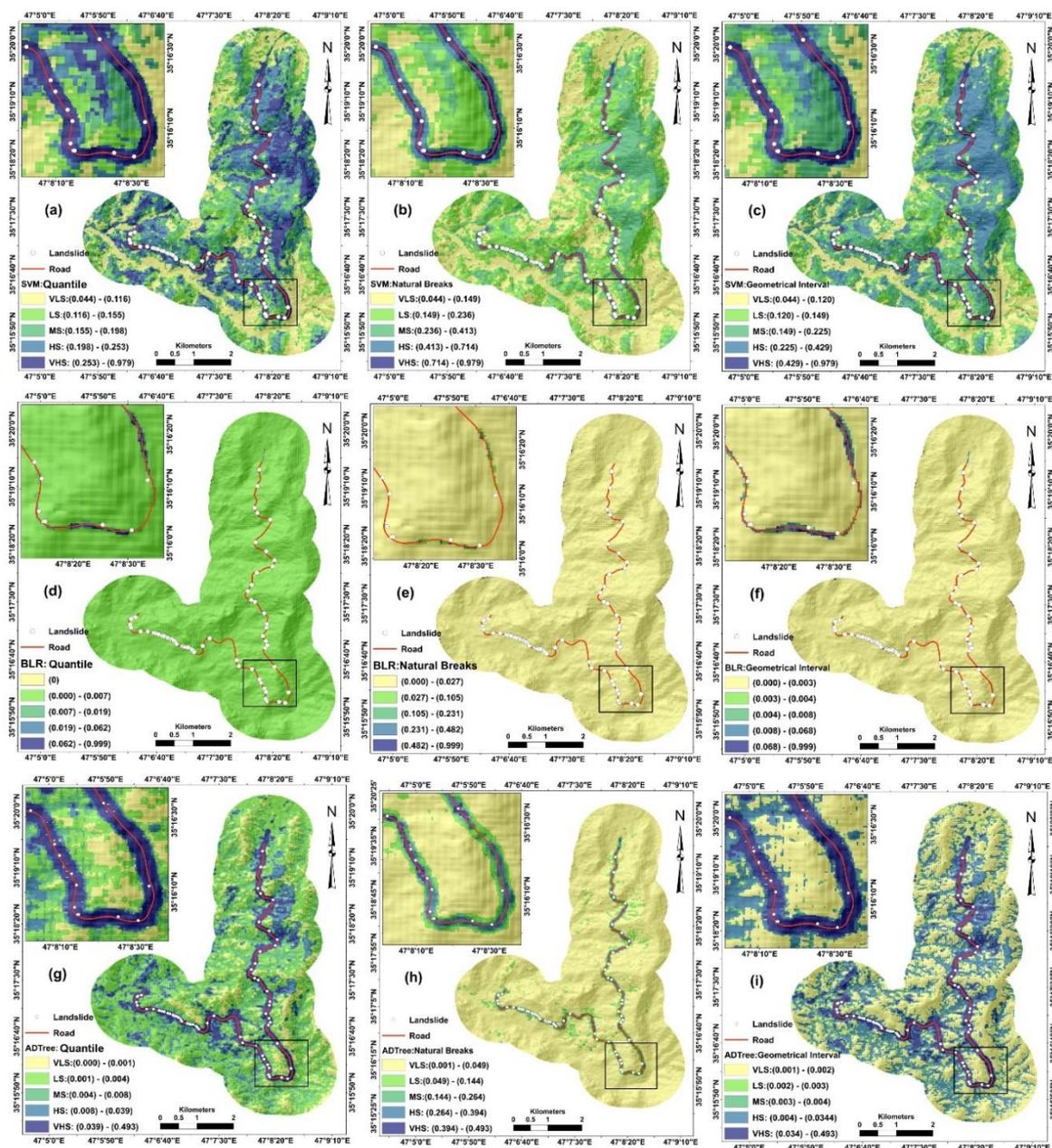
Figure 6. Landslide susceptibility maps generated by (**a**) SVM with the quantile method, (**b**) SVM with the natural break method, (**c**) SVM classified with the geometrical interval method, (**d**) BLR with the quantile method, (**e**) BLR with the natural break method, (**f**) BLR with the geometrical interval method, (**g**) ADTree with the quantile method, (**h**) ADTree with the natural break method, and (**i**) ADTree with the geometrical interval method.

5.5. Evaluation of Landslide Susceptibility Maps

We used the ROC curves for the training and validation datasets to evaluate the machine learning models. The probabilities of landslides calculated for the training and validation datasets provide measures of the performance, and prediction accuracy of the models, respectively [11]. The *x*-axis, and *y*-axis of the ROC curves are, respectively, the sensitivity and 100-specificity indices. The performance and prediction accuracy of the three models are shown in Figure 7a,b, respectively. The performance of the SVM model is slightly higher (AUC = 0.988) than that of the BLR (AUC = 0.985), and ADTree (AUC = 0.977) models. The prediction accuracy of SVM is also slightly higher (AUC = 0.984) than that of the BLR (AUC = 0.980) and ADTree (AUC = 0.977).
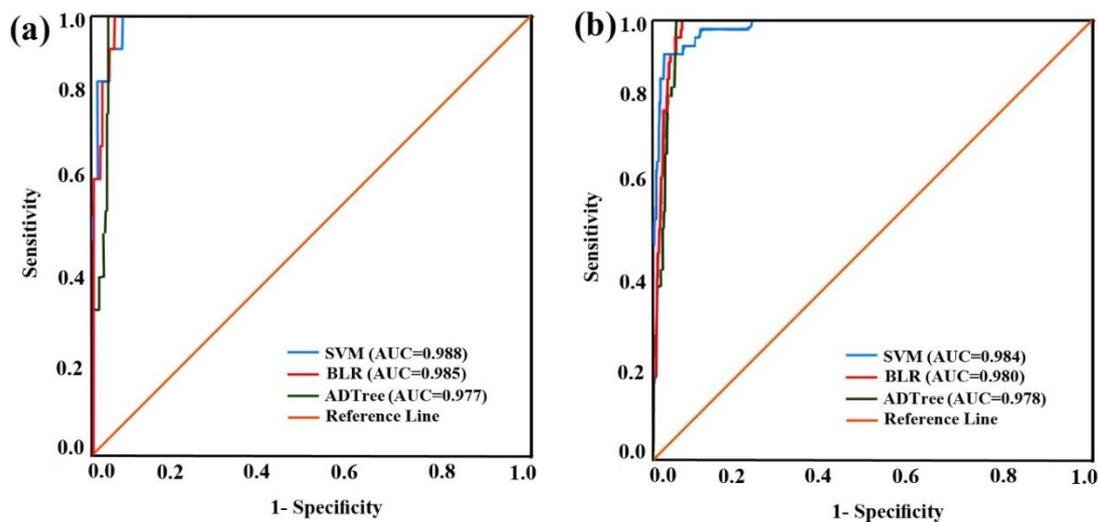
Figure 7. ROC curve and AUC for the SVM, BLR, and ADTree models: (**a**) training and (**b**) validation datasets.

6. Discussion

The susceptibility of an area to landslides is a function of different possible conditioning factors. As all of the factors might have no predictive capability, the most important must be objectively chosen to strengthen the performance and accuracy of the learning algorithms in the training phase. In this study, we used the IGR technique to identify factors with high predictive powers. The weight of each factor in the training phase was calculated using an entropy index. In this study, tested 18 conditioning factors with a raster resolution of 10 m and found that 10 factors were significant: Distance from the road, normalized difference vegetation index, land use, slope aspect, lithology, slope angle, precipitation, distance from faults, elevation, and topographic wetness index. Eight factors were removed from final modeling because they had AM values of 0: e distance from stream, slope-length, annual solar radiation, profile curvature, plan curvature, fault density, drainage density, and stream power index.

Researchers have used three main methods to display classes on landslide susceptibility maps: the natural break, geometrical interval, and quantile methods. We statistically tested the three methods for producing maps using the three machine learning algorithms investigated in this study. The natural breaks classification method was selected for the SVM and ADTree models, and the quantile method for the BLR model. Most of the researchers in landslide susceptibility mapping confirmed the capability of the natural break method to classify the LSIs [105–109]. The quantile method among the classification methods is generally the most effective and commonly used method [104,110,111]. Nhu et al. [16] applied the natural break, geometrical interval, and quantile and based on their histogram of landslide probability values selected the natural break classification method for the random forest algorithm. However, the geometrical interval method for the three ensemble models of rotation forest-based random forest (RF-RAF), bagging based random forest (BA-RAF), and random subspace-based random forest (RS-RAF) to produce shallow landslide susceptibility maps.

Distance to the road is the factor most closely related to landslides in the study area. All the susceptibility maps showed that most landslides are less than 100 m from the road through the study area. The road is located at high elevation in wet areas (high topographic wetness index), which are other significant landslide conditioning factors. In recent years, the road through the study area has been widened, and new bridges have been constructed, changing the landscape and initiating instability along the road. The road is also trafficked by trucks and other heavy vehicles. Therefore, they should be more considered during road widening and other engineering construction in future.

The second and third most crucial landslide conditioning factors are the normalized difference vegetation index land use. Most landslides in the study area have happened in unvegetated or sparsely vegetated areas, including rangeland.

Slope aspect is another critical factor as landslides tend to occur on slopes oriented toward the northwest and west because these aspects experience more precipitation and runoff than other factors. Precipitation in the study area is higher than the average for the country. By considering landslides susceptibility maps and its histograms with three models, it can be concluded that the susceptible areas of landslides belong to a very high susceptible class.

After selecting appropriate landslide conditioning factors, we prepared landslide susceptibility maps using three machine learning models and the natural break, geometrical intervals, and quantile classification methods. Our finding concluded that natural break and quantile had most concordance and consistency with the reality of the study area. We have shown that the SVM algorithm has the highest goodness-of-fit and prediction accuracy of the three machine learning algorithms tested in this study based on both training and validation datasets. This result is consistent with the findings of other landslide researchers [22,112–117]. For example, Kalantar et al. [115], compared the performance of SVM, LR, and ANN for landslide assessment in a catchment in the Dodangeh watershed, Mazandaran province, Iran. They concluded that SVM outperformed the other models, and therefore, it was potentially known as the most powerful algorithm for landslide modeling in their study area. Abedini et al. [116] compared the performance of the SVM and LMT models for landslide susceptibility mapping in Kamyaran county, also in Kurdistan province, and confirmed the superiority of the SVM model. SVM has also been successfully used in landslide susceptibility mapping in the Cameron Highlands, Malaysia [19]. In contrast, Nhu et al. [15] compared LMT, LR, NBT, ANN, and SVM models for landslide susceptibility mapping in Bjar city, Kurdistan province, and found that LMT had the highest, and SVM the lowest, goodness-of-fit and prediction accuracy.

According to the best of our knowledge of the literature on landslide susceptibility mapping, SVM can be successfully used as a benchmark computing machine learning algorithm in new ensemble models [118,119]. For example, Pham et al. [119] proposed a new ensemble model consisting of random subspace base classification and regression tree (RSSCART) for landslide modeling and assessment in the Luc Yen district of Yen Bai province, Viet Nam. They compared their new ensemble model, the SVM benchmark model. SVM offers several advantages over other machine learning models: (i) It is free from feature selection techniques that are required by other models such as decision trees; (ii) it can handle complex and non-linear problems with large datasets; and (iii) it solves the convex quadratic programming optimization problem of separating the hyper-plane and thus is a suitable replacement for artificial neural networks [35,113,120].

Our results indicate that the BLR algorithm outperformed ADTree in the landslide modeling process and susceptibility map assessment. BLR is an LR method, within the Bayesian paradigm, that includes a posterior distribution function for evaluating each landslide conditioning factor. BLR also offers several advantages that make it as a robust algorithm for modeling: (i) it can estimate probability intervals of landslide occurrence; (ii) it can be used with small samples, as it does not rely on large-sample approximations; (iii) available prior information about regression coefficients can be incorporated in the Bayesian model; and (iv) multi-level data or models are particularly suited to the hierarchical structure of Bayesian modeling [121,122]. The performance and prediction accuracy of the BLR model has been confirmed and reported, not only for landslide modeling [65,88], but also for flood [123] and land subsidence [92] susceptibility mapping.

ADTree has been suggested and used by some environmental researchers [53,92,124]. An advantage of using ADTree is that it has the fastest induction time for domain problems with few discriminative features [125]. Moreover, it has been successfully used as a base classifier in coupled ensemble models, including multiboot (MB), bagging (BA), random subspace (RS), and rotation forest (RF) for landslide susceptibility mapping [51,78]. Our results will be useful to landslide hazard managers, decision-makers, and researchers when selecting the most appropriate models for landslide

susceptibility mapping. However, we acknowledge the limitations of the present study, largely uncertainties in input data. For example, results can differ depending on the sample size and raster resolution. Shirzadi et al. [78] studied these uncertainties and suggested a raster resolution of 10 m for training/validation sample sizes 60/40% and 70/30%; and a resolution of 20 m for sample sizes of 80/20% and 90/10%. Another limitation of the current study is related to model selection. Each algorithm has a specific probability distribution function or rule, not all of which fit a given training dataset. Therefore, it is necessary to test the models and select the best one for a given study. This process is mainly done using a trial-and-error technique and is time-consuming.

A limitation of landslide susceptibility mapping, in general, is that maps generated with machine learning techniques can accurately show where landslides are likely to occur based on geo-environmental factors, but the important physical, mechanical, and elastic properties of soil such as porosity, permeability, cohesion, and pore water pressure are not considered. These soil-related properties strongly control landslide occurrence at the site scale, yet preparing maps showing their distributions is costly and time-consuming. We recommend that researchers consider these factors and specially to use them in conjunction with slope stability models and deterministic numerical models that address the factor of safety (FOS). For example, Shallow Landsliding Stability (SHALSTAB) and SINMAP (Stability Index MAPping). These models couple a hydrologic model with an infinite slope form of the Mohr-Coulomb failure law to spatially predict slope failures. Therefore, one of the ways to enhance the accuracy of the susceptibility maps is to use the soil-related factors in future.

Additionally, the application of high-resolution data such as airborne laser scanning of Light Detection and Ranging (LiDAR) not only could enhance the quality of the conditioning factors but also the prediction accuracy of the models. The ability of high-resolution data has been confirmed and evaluated by some landslide researchers [99,126–128]. For example, Jebur et al. [128] by very high-resolution data, LiDAR, optimized the used landslide conditioning factors, and they concluded that a high-quality, informative database, is essential and classification of landslide types prior to landslide susceptibility assessment is necessary to help improve model performance.

7. Conclusions

Accurate landslide susceptibility maps provide land-use managers and government officials with a valuable tool for managing landslide hazard and risk. In this paper, we evaluate the performance and prediction accuracy of three well-known machine learning models (SVM, BLR, and ADTree) for landslide susceptibility mapping in the Salavat Abad saddle, Kurdistan province, Iran. The saddle is an important area that connects Kurdistan to other provinces of Iran, and thus, a priority for landslide management and remediation. We determine the most critical geo-environmental factors using the IGR technique to delineate better, visualize, and interpret landslide-prone areas. In our study area, the essential factors for landslide modeling are distance to road, NDVI, and land use. Our models show that the area bordering the arterial road in the Salavat Abad saddle is most susceptible to landsliding. We also show the SVM algorithm has a high goodness-of-fit and prediction accuracy of landslides in the study area, and that BLR and ADTree are suitable alternatives in the study area. Therefore, we suggest the SVM and BLR as soft computing benchmark models in similar areas in terms of topographic, climate, and lithology features.

Author Contributions: V.-H.N., D.Z., H.S., K.C., A.S., N.A.-A., S.K.S., J.D., and H.N contributed equally to the work. D.Z., H.S., and A.S. collected field data and conducted the landslide mapping and analysis. D.Z., H.S., A.S., S.K.S., and J.D. wrote the manuscript. V.-H.N., H.S., K.C., A.S., N.A.-A., and H.N. provided critical comments in planning this paper and edited the manuscript. All the authors discussed the results and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Varnes, D.J. The international association of engineering geology commission on landslides and other mass movements on slopes. 1984. Landslide hazard zonation: A review of principles and practice. *Nat. Hazards* **1984**, *3*, 63–79.
2. Nadim, F.; Kjekstad, O.; Peduzzi, P.; Herold, C.; Jaedicke, C. Global landslide and avalanche hotspots. *Landslides* **2006**, *3*, 159–173. [CrossRef]
3. Assilzadeh, H.; Levy, J.K.; Wang, X. Landslide catastrophes and disaster risk reduction: A gis framework for landslide prevention and management. *Remote Sens.* **2010**, *2*, 2259–2273. [CrossRef]
4. Rezaei, S.; Shooshpasha, I.; Rezaei, H. Reconstruction of landslide model from ert, geotechnical, and field data, nargeschal landslide, iran. *Bull. Eng. Geol. Environ.* **2019**, *78*, 3223–3237. [CrossRef]
5. Arabameri, A.; Saha, S.; Roy, J.; Chen, W.; Blaschke, T.; Tien Bui, D. Landslide susceptibility evaluation and management using different machine learning methods in the gallicash river watershed, iran. *Remote Sens.* **2020**, *12*, 475. [CrossRef]
6. Pourghasemi, H.R.; 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. [CrossRef]
7. Kornejady, A.; Pourghasemi, H.R.; Afzali, S.F. Presentation of rffr new ensemble model for landslide susceptibility assessment in iran. In *Landslides: Theory, Practice and Modelling*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 123–143.
8. Pourghasemi, H.R.; Pradhan, B.; Gokceoglu, C.; Mohammadi, M.; Moradi, H.R. Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at haraz watershed, iran. *Arab. J. Geosci.* **2013**, *6*, 2351–2365. [CrossRef]
9. Ehteshami-Moinabadi, M. On the historical landslide fatalities in the iranian plateau. In *NHAQ97*; Geographical Organization of Iran: Tehran, Iran, 2019.
10. Aghda, S.F.; Bagheri, V.; Razifard, M. Landslide susceptibility mapping using fuzzy logic system and its influences on mainlines in lashgarak region, tehran, iran. *Geotech. Geol. Eng.* **2018**, *36*, 915–937.
11. Shirzadi, A.; Bui, D.T.; Pham, B.T.; Solaimani, K.; Chapi, K.; Kavian, A.; Shahabi, H.; Revhaug, I. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* **2017**, *76*, 60. [CrossRef]
12. Party, I.L.W. *Iranian Landslides List*; Forest, Rangeland and Watershed Association: Tehran, Iran, 2007; Volume 60.
13. Duncan, C.; Norman, I. *Stabilization of Rock Slopes*; Landslides investigations and mitigation, special report 247; Transportation Research Board, National Research Council: Washington, DC, USA, 1996; pp. 474–506.
14. Wang, G.; Lei, X.; Chen, W.; Shahabi, H.; Shirzadi, A. Hybrid computational intelligence methods for landslide susceptibility mapping. *Symmetry* **2020**, *12*, 325. [CrossRef]
15. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Singh, S.K.; Al-Ansari, N.; Clague, J.J.; Jaafari, A.; Chen, W.; Miraki, S.; Dou, J. Shallow landslide susceptibility mapping: A comparison between logistic model tree, logistic regression, naïve bayes tree, artificial neural network, and support vector machine algorithms. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2749. [CrossRef] [PubMed]
16. Nhu, V.H.; Mohammadi, A.; Shahabi, H.; Ahmad, B.B.; Al-Ansari, N.; Shirzadi, A.; Clague, J.J.; Jaafari, A.; Chen, W.; Nguyen, H. Landslide Susceptibility Mapping Using Machine Learning Algorithms and Remote Sensing Data in a Tropical Environment. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4933. [CrossRef] [PubMed]
17. Kumar, R.; Anbalagan, R. Landslide susceptibility mapping using analytical hierarchy process (ahp) in tehri reservoir rim region, uttarakhand. *J. Geol. Soc. India* **2016**, *87*, 271–286. [CrossRef]
18. Roodposhti, M.S.; Rahimi, S.; Beglou, M.J. Promethee ii and fuzzy ahp: An enhanced gis-based landslide susceptibility mapping. *Nat. Hazards* **2014**, *73*, 77–95. [CrossRef]
19. Shirzadi, A.; Chapi, K.; Shahabi, H.; Solaimani, K.; Kavian, A.; Ahmad, B.B. Rock fall susceptibility assessment along a mountainous road: An evaluation of bivariate statistic, analytical hierarchy process and frequency ratio. *Environ. Earth Sci.* **2017**, *76*, 152. [CrossRef]
20. Shahabi, H.; Hashim, M.; Ahmad, B.B. Remote sensing and gis-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central zab basin, iran. *Environ. Earth Sci.* **2015**, *73*, 8647–8668. [CrossRef]

21. Shahabi, H.; Khezri, S.; Ahmad, B.B.; Hashim, M. Landslide susceptibility mapping at central zab basin, iran: A comparison between analytical hierarchy process, frequency ratio and logistic regression models. *Catena* **2014**, *115*, 55–70. [CrossRef]

22. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Alizadeh, M.; Chen, W.; Mohammadi, A.; Ahmad, B.B.; Panahi, M.; Hong, H. Landslide detection and susceptibility mapping by airsar data using support vector machine and index of entropy models in cameron highlands, malaysia. *Remote Sens.* **2018**, *10*, 1527. [CrossRef]

23. Hong, H.; Shahabi, H.; Shirzadi, A.; Chen, W.; Chapi, K.; Ahmad, B.B.; Roodposhti, M.S.; Hesar, A.Y.; Tian, Y.; Bui, D.T. Landslide susceptibility assessment at the wuning area, china: A comparison between multi-criteria decision making, bivariate statistical and machine learning methods. *Nat. Hazards* **2019**, *96*, 173–212. [CrossRef]

24. Chen, W.; Xie, X.; Peng, J.; Shahabi, H.; Hong, H.; Bui, D.T.; Duan, Z.; Li, S.; Zhu, A.-X. Gis-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. *Catena* **2018**, *164*, 135–149. [CrossRef]

25. Shahabi, H.; Hashim, M. Landslide susceptibility mapping using gis-based statistical models and remote sensing data in tropical environment. *Sci. Rep.* **2015**, *5*, 1–15. [CrossRef]

26. Nhu, V.-H.; Rahmati, O.; Falah, F.; Shojaei, S.; Al-Ansari, N.; Shahabi, H.; Shirzadi, A.; Górski, K.; Nguyen, H.; Ahmad, B.B. Mapping of groundwater spring potential in karst aquifer system using novel ensemble bivariate and multivariate models. *Water* **2020**, *12*, 985. [CrossRef]

27. Azareh, A.; Rahmati, O.; Rafiei-Sardooi, E.; Sankey, J.B.; Lee, S.; Shahabi, H.; Ahmad, B.B. Modelling gully-erosion susceptibility in a semi-arid region, iran: Investigation of applicability of certainty factor and maximum entropy models. *Sci. Total Environ.* **2019**, *655*, 684–696. [CrossRef]

28. Shirzadi, A.; Saro, L.; Joo, O.H.; Chapi, K. A gis-based logistic regression model in rock-fall susceptibility mapping along a mountainous road: Salavat abad case study, kurdistan, iran. *Nat. Hazards* **2012**, *64*, 1639–1656. [CrossRef]

29. Pack, R.T.; Tarboton, D.; Goodwin, C. *Sinmap 2.0-a Stability Index Approach to Terrain Stability Hazard Mapping, User's Manual*; U.S. Forest Service, Rocky Mountain Research Station: Logan, UT, USA, 1999.

30. Dietrich, W.E.; de Asua, R.R.; Coyle, J.; Orr, B.; Trso, M. A validation study of the shallow slope stability model, shalstab, in forested lands of northern california. *Stillwater Ecosyst. Watershed Riverine Sci. Berkeley CA* **1998**, *11*, 16–60.

31. Cohen, D.; Schwarz, M. Effects of tree roots on shallow landslides distribution and frequency in the european alps using a new physically-based discrete element model. *EGUGA* **2017**, *19*, 6154.

32. Cislaghi, A.; Bischetti, G.B. Source areas, connectivity, and delivery rate of sediments in mountainous-forested hillslopes: A probabilistic approach. *Sci. Total Environ.* **2019**, *652*, 1168–1186. [CrossRef] [PubMed]

33. Ewen, J.; Parkin, G.; O'Connell, P.E. Shetran: Distributed river basin flow and transport modeling system. *J. Hydrol. Eng.* **2000**, *5*, 250–258. [CrossRef]

34. Baum, R.L.; Savage, W.Z.; Godt, J.W. Trigrs—A fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis. *US Geol. Surv. Open File Rep.* **2002**, *424*, 38.

35. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using gis-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2014**, *11*, 425–439. [CrossRef]

36. Pourghasemi, H.; Moradi, H.; Aghda, S.F. Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. *Nat. Hazards* **2013**, *69*, 749–779. [CrossRef]

37. Thiery, Y.; Malet, J.-P.; Sterlacchini, S.; Puissant, A.; Maquaire, O. Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. *Geomorphology* **2007**, *92*, 38–59. [CrossRef]

38. Hong, H.; Pourghasemi, H.R.; Pourtaghi, Z.S. Landslide susceptibility assessment in lianhua county (china): A comparison between a random forest data mining technique and bivariate and multivariate statistical models. *Geomorphology* **2016**, *259*, 105–118. [CrossRef]

39. Shirzadi, A.; Shahabi, H.; Chapi, K.; Bui, D.T.; Pham, B.T.; Shahedi, K.; Ahmad, B.B. A comparative study between popular statistical and machine learning methods for simulating volume of landslides. *Catena* **2017**, *157*, 213–226. [CrossRef]

40. Tian, Y.; Xu, C.; Hong, H.; Zhou, Q.; Wang, D. Mapping earthquake-triggered landslide susceptibility by use of artificial neural network (ann) models: An example of the 2013 minxian (china) mw 5.9 event. *Geomat. Nat. Hazards Risk* **2019**, *10*, 1–25. [CrossRef]

41. Rahmati, O.; Panahi, M.; Ghiasi, S.S.; Deo, R.C.; Tiefenbacher, J.P.; Pradhan, B.; Jahani, A.; Goshtasb, H.; Kornejady, A.; Shahabi, H.; et al. Hybridized neural fuzzy ensembles for dust source modeling and prediction. *Atmos. Environ.* **2020**, *224*, 117320. [CrossRef]

42. Wang, Y.; Hong, H.; Chen, W.; Li, S.; Panahi, M.; Khosravi, K.; Shirzadi, A.; Shahabi, H.; Panahi, S.; Costache, R. Flood susceptibility mapping in dingnan county (china) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm. *J. Environ. Manag.* **2019**, *247*, 712–729. [CrossRef]

43. Chen, W.; Panahi, M.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Panahi, S.; Li, S.; Jaafari, A.; Ahmad, B.B. Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility. *Catena* **2019**, *172*, 212–231. [CrossRef]

44. Tien Bui, D.; Khosravi, K.; Li, S.; Shahabi, H.; Panahi, M.; Singh, V.P.; Chapi, K.; Shirzadi, A.; Panahi, S.; Chen, W. New hybrids of anfis with several optimization algorithms for flood susceptibility modeling. *Water* **2018**, *10*, 1210. [CrossRef]

45. Pham, B.T.; Prakash, I.; Dou, J.; Singh, S.K.; Trinh, P.T.; Tran, H.T.; Le, T.M.; Van Phong, T.; Khoi, D.K.; Shirzadi, A. A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. *Geocarto Int.* **2019**, 1–25. [CrossRef]

46. Bui, D.T.; Shirzadi, A.; Shahabi, H.; Geertsema, M.; Omidvar, E.; Clague, J.J.; Pham, B.T.; Dou, J.; Asl, D.T.; Ahmad, B.B. New ensemble models for shallow landslide susceptibility modeling in a semi-arid watershed. *Forests* **2019**, *10*, 743.

47. Shahabi, H.; Shirzadi, A.; Ghaderi, K.; Omidvar, E.; Al-Ansari, N.; Clague, J.J.; Geertsema, M.; Khosravi, K.; Amini, A.; Bahrami, S. Flood detection and susceptibility mapping using sentinel-1 remote sensing data and a machine learning approach: Hybrid intelligence of bagging ensemble based on k-nearest neighbor classifier. *Remote Sens.* **2020**, *12*, 266. [CrossRef]

48. Chen, W.; Zhao, X.; Shahabi, H.; Shirzadi, A.; Khosravi, K.; Chai, H.; Zhang, S.; Zhang, L.; Ma, J.; Chen, Y. Spatial prediction of landslide susceptibility by combining evidential belief function, logistic regression and logistic model tree. *Geocarto Int.* **2019**, *34*, 1177–1201. [CrossRef]

49. Rahmati, O.; Naghibi, S.A.; Shahabi, H.; Bui, D.T.; Pradhan, B.; Azareh, A.; Rafiei-Sardooi, E.; Samani, A.N.; Melesse, A.M. Groundwater spring potential modelling: Comprising the capability and robustness of three different modeling approaches. *J. Hydrol.* **2018**, *565*, 248–261. [CrossRef]

50. Bui, D.T.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Omidavr, E.; Pham, B.T.; Asl, D.T.; Khaledian, H.; Pradhan, B.; Panahi, M. A novel ensemble artificial intelligence approach for gully erosion mapping in a semi-arid watershed (iran). *Sensors* **2019**, *19*, 2444.

51. Pham, B.T.; Shirzadi, A.; Shahabi, H.; Omidvar, E.; Singh, S.K.; Sahana, M.; Asl, D.T.; Ahmad, B.B.; Kim Quoc, N.; Lee, S. Landslide susceptibility assessment by novel hybrid machine learning algorithms. *Sustainability* **2019**, *11*, 4386. [CrossRef]

52. Bui, D.T.; Shirzadi, A.; Chapi, K.; Shahabi, H.; Pradhan, B.; Pham, B.T.; Singh, V.P.; Chen, W.; Khosravi, K.; Ahmad, B.B. A hybrid computational intelligence approach to groundwater spring potential mapping. *Water* **2019**, *11*, 2013.

53. Shirzadi, A.; Solaimani, K.; Roshan, M.H.; Kavian, A.; Chapi, K.; Shahabi, H.; Keesstra, S.; Ahmad, B.B.; Bui, D.T. Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. *Catena* **2019**, *178*, 172–188. [CrossRef]

54. Chen, W.; Zhao, X.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Wang, X.; Ahmad, B.B. Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping. *J. Hydrol.* **2020**, *583*, 124602. [CrossRef]

55. Rahmati, O.; Choubin, B.; Fathabadi, A.; Coulon, F.; Soltani, E.; Shahabi, H.; Mollaefar, E.; Tiefenbacher, J.; Cipullo, S.; Ahmad, B.B. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and uneec methods. *Sci. Total Environ.* **2019**, *688*, 855–866. [CrossRef]

56. Chen, W.; Peng, J.; Hong, H.; Shahabi, H.; Pradhan, B.; Liu, J.; Zhu, A.-X.; Pei, X.; Duan, Z. Landslide susceptibility modelling using gis-based machine learning techniques for chongren county, jiangxi province, china. *Sci. Total Environ.* **2018**, *626*, 1121–1135. [CrossRef]

57. Tien Bui, D.; Shahabi, H.; Omidvar, E.; Shirzadi, A.; Geertsema, M.; Clague, J.J.; Khosravi, K.; Pradhan, B.; Pham, B.T.; Chapi, K. Shallow landslide prediction using a novel hybrid functional machine learning algorithm. *Remote Sens.* **2019**, *11*, 931. [CrossRef]

58. Felicísimo, Á.M.; Cuartero, A.; Remondo, J.; Quirós, E. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. *Landslides* **2013**, *10*, 175–189. [CrossRef]

59. Hong, H.; Liu, J.; Bui, D.T.; Pradhan, B.; Acharya, T.D.; Pham, B.T.; Zhu, A.-X.; Chen, W.; Ahmad, B.B. Landslide susceptibility mapping using j48 decision tree with adaboost, bagging and rotation forest ensembles in the guangchang area (china). *Catena* **2018**, *163*, 399–413. [CrossRef]

60. Pham, B.T.; Prakash, I.; Singh, S.K.; Shirzadi, A.; Shahabi, H.; Bui, D.T. Landslide susceptibility modeling using reduced error pruning trees and different ensemble techniques: Hybrid machine learning approaches. *Catena* **2019**, *175*, 203–218. [CrossRef]

61. Bui, D.T.; Shirzadi, A.; Amini, A.; Shahabi, H.; Al-Ansari, N.; Hamidi, S.; Singh, S.K.; Pham, B.T.; Ahmad, B.B.; Ghazvinei, P.T. A hybrid intelligence approach to enhance the prediction accuracy of local scour depth at complex bridge piers. *Sustainability* **2020**, *12*, 1063.

62. Nhu, V.-H.; Janizadeh, S.; Avand, M.; Chen, W.; Farzin, M.; Omidvar, E.; Shirzadi, A.; Shahabi, H.; Clague, J.J.; Jaafari, A. Gis-based gully erosion susceptibility mapping: A comparison of computational ensemble data mining models. *Appl. Sci.* **2020**, *10*, 2039. [CrossRef]

63. Chen, W.; Hong, H.; Li, S.; Shahabi, H.; Wang, Y.; Wang, X.; Ahmad, B.B. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles. *J. Hydrol.* **2019**, *575*, 864–873. [CrossRef]

64. Nhu, V.-H.; Shirzadi, A.; Shahabi, H.; Chen, W.; Clague, J.J.; Geertsema, M.; Jaafari, A.; Avand, M.; Miraki, S.; Asl, D.T. Shallow landslide susceptibility mapping by random forest base classifier and its ensembles in a semi-arid region of iran. *Forests* **2020**, *11*, 421. [CrossRef]

65. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Shahabi, H.; Chapi, K.; Pham, B.T.; Bin Ahmad, B.; Tien Bui, D. A novel hybrid approach of bayesian logistic regression and its ensembles for landslide susceptibility assessment. *Geocarto Int.* **2019**, *34*, 1427–1457. [CrossRef]

66. Chen, W.; Pradhan, B.; Li, S.; Shahabi, H.; Rizeei, H.M.; Hou, E.; Wang, S. Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. *Nat. Resour. Res.* **2019**, *28*, 1239–1258. [CrossRef]

67. Taheri, K.; Shahabi, H.; Chapi, K.; Shirzadi, A.; Gutiérrez, F.; Khosravi, K. Sinkhole susceptibility mapping: A comparison between bayes-based machine learning algorithms. *Land Degrad. Dev.* **2019**, *30*, 730–745. [CrossRef]

68. Jaafari, A.; Panahi, M.; Pham, B.T.; Shahabi, H.; Bui, D.T.; Rezaie, F.; Lee, S. Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility. *Catena* **2019**, *175*, 430–445. [CrossRef]

69. Chen, W.; Hong, H.; Panahi, M.; Shahabi, H.; Wang, Y.; Shirzadi, A.; Pirasteh, S.; Alesheikh, A.A.; Khosravi, K.; Panahi, S. Spatial prediction of landslide susceptibility using gis-based data mining techniques of anfis with whale optimization algorithm (woa) and grey wolf optimizer (gwo). *Appl. Sci.* **2019**, *9*, 3755. [CrossRef]

70. He, Q.; Shahabi, H.; Shirzadi, A.; Li, S.; Chen, W.; Wang, N.; Chai, H.; Bian, H.; Ma, J.; Chen, Y. Landslide spatial modelling using novel bivariate statistical based naïve bayes, rbf classifier, and rbf network machine learning algorithms. *Sci. Total Environ.* **2019**, *663*, 1–15. [CrossRef]

71. Chen, W.; Li, Y.; Xue, W.; Shahabi, H.; Li, S.; Hong, H.; Wang, X.; Bian, H.; Zhang, S.; Pradhan, B. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Sci. Total Environ.* **2020**, *701*, 134979. [CrossRef]

72. Nguyen, P.T.; Tuyen, T.T.; Shirzadi, A.; Pham, B.T.; Shahabi, H.; Omidvar, E.; Amini, A.; Entezami, H.; Prakash, I.; Phong, T.V. Development of a novel hybrid intelligence approach for landslide spatial prediction. *Appl. Sci.* **2019**, *9*, 2824. [CrossRef]

73. Chen, W.; Shirzadi, A.; Shahabi, H.; Ahmad, B.B.; Zhang, S.; Hong, H.; Zhang, N. A novel hybrid artificial intelligence approach based on the rotation forest ensemble and naïve bayes tree classifiers for a landslide susceptibility assessment in langao county, china. *Geomat. Nat. Hazards Risk.* **2017**, *8*, 1955–1977. [CrossRef]

74. Tien Bui, D.; Khosravi, K.; Shahabi, H.; Daggupati, P.; Adamowski, J.F.; Melesse, A.M.; Thai Pham, B.; Pourghasemi, H.R.; Mahmoudi, M.; Bahrami, S. Flood spatial modeling in northern iran using remote sensing and gis: A comparison between evidential belief functions and its ensemble with a multivariate logistic regression model. *Remote Sens.* **2019**, *11*, 1589. [CrossRef]

75. Chen, W.; Shahabi, H.; Shirzadi, A.; Hong, H.; Akgun, A.; Tian, Y.; Liu, J.; Zhu, A.-X.; Li, S. Novel hybrid artificial intelligence approach of bivariate statistical-methods-based kernel logistic regression classifier for landslide susceptibility modeling. *Bull. Eng. Geol. Environ.* **2019**, *78*, 4397–4419. [CrossRef]

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

77. Pham, B.T.; Bui, D.T.; Prakash, I. Landslide susceptibility assessment using bagging ensemble based alternating decision trees, logistic regression and j48 decision trees methods: A comparative study. *Geotech. Geol. Eng.* **2017**, *35*, 2597–2611. [CrossRef]

78. Shirzadi, A.; Soliamani, K.; Habibnejhad, M.; Kavian, A.; Chapi, K.; Shahabi, H.; Chen, W.; Khosravi, K.; Thai Pham, B.; Pradhan, B. Novel gis based machine learning algorithms for shallow landslide susceptibility mapping. *Sensors* **2018**, *18*, 3777. [CrossRef] [PubMed]

79. Gao, H.; Fam, P.S.; Tay, L.; Low, H. An overview and comparison on recent landslide susceptibility mapping methods. *Disaster Adv.* **2019**, *12*, 46–64.

80. Samadian, B.; Fakher, A. Proposing a framework to combine geological and geotechnical information for city planning in sanandaj (iran). *Eng. Geol.* **2016**, *209*, 1–11. [CrossRef]

81. Xu, C.; Dai, F.; Xu, X.; Lee, Y.H. Gis-based support vector machine modeling of earthquake-triggered landslide susceptibility in the jianjiang river watershed, china. *Geomorphology* **2012**, *145*, 70–80. [CrossRef]

82. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology/un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrol. Sci. J.* **1979**, *24*, 43–69. [CrossRef]

83. Ercanoglu, M.; Gokceoglu, C. Assessment of landslide susceptibility for a landslide-prone area (north of yenice, nw turkey) by fuzzy approach. *Environ. Geol.* **2002**, *41*, 720–730.

84. Moore, I.D.; Wilson, J.P. Length-slope factors for the revised universal soil loss equation: Simplified method of estimation. *J. Soil Water Conserv.* **1992**, *47*, 423–428.

85. Bui, D.T.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Hoang, N.-D.; Pham, B.T.; Bui, Q.-T.; Tran, C.-T.; Panahi, M.; Ahmad, B.B. A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. *Remote Sens.* **2018**, *10*, 1538.

86. Hong, H.; Liu, J.; Zhu, A.-X.; Shahabi, H.; Pham, B.T.; Chen, W.; Pradhan, B.; Bui, D.T. A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the wuning area (china). *Environ. Earth Sci.* **2017**, *76*, 652. [CrossRef]

87. Roodposhti, M.S.; Safarrad, T.; Shahabi, H. Drought sensitivity mapping using two one-class support vector machine algorithms. *Atmos. Res.* **2017**, *193*, 73–82. [CrossRef]

88. Das, I.; Stein, A.; Kerle, N.; Dadhwal, V.K. Landslide susceptibility mapping along road corridors in the indian himalayas using bayesian logistic regression models. *Geomorphology* **2012**, *179*, 116–125. [CrossRef]

89. Hosmer, D.W., Jr.; Lemeshow, S.; Sturdivant, R.X. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2013; Volume 398.

90. Freund, Y.; Mason, L. *The Alternating Decision Tree Learning Algorithm*; icml: San Francisco, CA, USA, 1999; pp. 124–133.

91. Chen, W.; Li, Y.; Tsangaratos, P.; Shahabi, H.; Ilia, I.; Xue, W.; Bian, H. Groundwater spring potential mapping using artificial intelligence approach based on kernel logistic regression, random forest, and alternating decision tree models. *Appl. Sci.* **2020**, *10*, 425. [CrossRef]

92. Tien Bui, D.; Shahabi, H.; Shirzadi, A.; Chapi, K.; Pradhan, B.; Chen, W.; Khosravi, K.; Panahi, M.; Bin Ahmad, B.; Saro, L. Land subsidence susceptibility mapping in south korea using machine learning algorithms. *Sensors* **2018**, *18*, 2464. [CrossRef] [PubMed]

93. Jebur, M.N.; Pradhan, B.; Tehrany, M.S. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (lidar) data at catchment scale. *Remote Sens. Environ.* **2014**, *152*, 150–165. [CrossRef]

94. Moosavi, V.; Talebi, A.; Shirmohammadi, B. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by taguchi method. *Geomorphology* **2014**, *204*, 646–656. [CrossRef]

95. Huang, Y.; Zhao, L. Review on landslide susceptibility mapping using support vector machines. *Catena* **2018**, *165*, 520–529. [CrossRef]

96. Sameen, M.I.; Sarkar, R.; Pradhan, B.; Drukpa, D.; Alamri, A.M.; Park, H.-J. Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests. *Comput. Geosci.* **2020**, *134*, 104336. [CrossRef]

97. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.-W.; Khosravi, K.; Yang, Y.; Pham, B.T. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the izu-oshima volcanic island, japan. *Sci. Total Environ.* **2019**, *662*, 332–346. [CrossRef]

98. Quinlan, J.R. Improved use of continuous attributes in c4. 5. *J. Artif. Intell. Res.* **1996**, *4*, 77–90. [CrossRef]

99. Dubois, D.; Prade, H. International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. In *Inference in Possibilistic Hypergraphs*; Springer: Berlin/Heidelberg, Germany, 1990; pp. 249–259.

100. Kononenko, I. European conference on machine learning. In *Estimating Attributes: Analysis and Extensions of Relief*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 171–182.

101. Ayalew, L.; Yamagishi, H. The application of gis-based logistic regression for landslide susceptibility mapping in the kakuda-yahiko mountains, central japan. *Geomorphology* **2005**, *65*, 15–31. [CrossRef]

102. Schicker, R.; Moon, V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology* **2012**, *161*, 40–57. [CrossRef]

103. Meng, Q.; Miao, F.; Zhen, J.; Wang, X.; Wang, A.; Peng, Y.; Fan, Q. Gis-based landslide susceptibility mapping with logistic regression, analytical hierarchy process, and combined fuzzy and support vector machine methods: A case study from wolong giant panda natural reserve, china. *Bull. Eng. Geol. Environ.* **2016**, *75*, 923–944. [CrossRef]

104. 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. *Comput. Geosci.* **2013**, *51*, 350–365. [CrossRef]

105. Falaschi, F.; Giacomelli, F.; Federici, P.; Puccinelli, A.; Avanzi, G.A.; Pochini, A.; Ribolini, A. Logistic regression versus artificial neural networks: Landslide susceptibility evaluation in a sample area of the serchio river valley, italy. *Nat. Hazards* **2009**, *50*, 551–569. [CrossRef]

106. Bednarik, M.; Magulová, B.; Matys, M.; Marschalko, M. Landslide susceptibility assessment of the kraľovany–liptovský mikuláš railway case study. *Physics Chem. Earth Parts A/B/C* **2010**, *35*, 162–171. [CrossRef]

107. Erener, A.; Düzgün, H.S.B. Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of more and romsdal (norway). *Landslides* **2010**, *7*, 55–68. [CrossRef]

108. Constantin, M.; Bednarik, M.; Jurchescu, M.C.; Vlaicu, M. Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the sibiciu basin (romania). *Environ. Earth Sci.* **2011**, *63*, 397–406. [CrossRef]

109. Xu, C.; Xu, X. Controlling parameter analyses and hazard mapping for earthquake-triggered landslides: An example from a square region in beichuan county, sichuan province, china. *Arab. J. Geosci.* **2013**, *6*, 3827–3839. [CrossRef]

110. Bui, D.T.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378.

111. Umar, Z.; Pradhan, B.; Ahmad, A.; Jebur, M.N.; Tehrany, M.S. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in west sumatera province, indonesia. *Catena* **2014**, *118*, 124–135. [CrossRef]

112. Yao, X.; Tham, L.; Dai, F. Landslide susceptibility mapping based on support vector machine: A case study on natural slopes of hong kong, china. *Geomorphology* **2008**, *101*, 572–582. [CrossRef]

113. Marjanović, M.; Kovačević, M.; Bajat, B.; Voženílek, V. Landslide susceptibility assessment using svm machine learning algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [CrossRef]

114. Ballabio, C.; Sterlacchini, S. Support vector machines for landslide susceptibility mapping: The staffora river basin case study, italy. *Math. Geosci.* **2012**, *44*, 47–70. [CrossRef]

115. Kalantar, B.; Pradhan, B.; Naghibi, S.A.; Motevalli, A.; Mansor, S. Assessment of the effects of training data selection on the landslide susceptibility mapping: A comparison between support vector machine (svm), logistic regression (lr) and artificial neural networks (ann). *Geomat. Nat. Hazards Risk* **2018**, *9*, 49–69. [CrossRef]

116. Abedini, M.; Ghasemian, B.; Shirzadi, A.; Bui, D.T. A comparative study of support vector machine and logistic model tree classifiers for shallow landslide susceptibility modeling. *Environ. Earth Sci.* **2019**, *78*, 560. [CrossRef]

117. Pandey, V.K.; Pourghasemi, H.R.; Sharma, M.C. Landslide susceptibility mapping using maximum entropy and support vector machine models along the highway corridor, garhwal himalaya. *Geocarto Int.* **2020**, *35*, 168–187. [CrossRef]

118. Bui, D.T.; Tuan, T.A.; Hoang, N.-D.; Thanh, N.Q.; Nguyen, D.B.; Van Liem, N.; Pradhan, B. Spatial prediction of rainfall-induced landslides for the lao cai area (vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. *Landslides* **2017**, *14*, 447–458.

119. Pham, B.T.; Prakash, I.; Bui, D.T. Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression trees. *Geomorphology* **2018**, *303*, 256–270. [CrossRef]

120. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]

121. Dunson, D.B. Commentary: Practical advantages of bayesian analysis of epidemiologic data. *Am. J. Epidemiol.* **2001**, *153*, 1222–1226. [CrossRef] [PubMed]

122. Rose, C.E.; Pan, Y.; Baughman, A.L. Bayesian logistic regression modeling as a flexible alternative for estimating adjusted risk ratios in studies with common outcomes. *J. Biom. Biostat.* **2015**, *6*, 1–6.

123. Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [CrossRef]

124. Khosravi, K.; Pham, B.T.; Chapi, K.; Shirzadi, A.; Shahabi, H.; Revhaug, I.; Prakash, I.; Bui, D.T. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at haraz watershed, northern iran. *Sci. Total Environ.* **2018**, *627*, 744–755. [CrossRef]

125. Sok, H.K.; Ooi, M.P.-L.; Kuang, Y.C.; Demidenko, S. Multivariate alternating decision trees. *Pattern Recognit.* **2016**, *50*, 195–209. [CrossRef]

126. Mahalingam, R.; Olsen, M.J.; O'Banion, M.S. Evaluation of landslide susceptibility mapping techniques using lidar-derived conditioning factors (oregon case study). *Geomat. Nat. Hazards Risk* **2016**, *7*, 1884–1907. [CrossRef]

127. Gorsevski, P.V.; Brown, M.K.; Panter, K.; Onasch, C.M.; Simic, A.; Snyder, J. Landslide detection and susceptibility mapping using lidar and an artificial neural network approach: A case study in the cuyahoga valley national park, ohio. *Landslides* **2016**, *13*, 467–484. [CrossRef]

128. Jebur, M.N.; Pradhan, B.; Tehrany, M.S. Manifestation of lidar-derived parameters in the spatial prediction of landslides using novel ensemble evidential belief functions and support vector machine models in gis. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *8*, 674–690. [CrossRef]