





Article

Landslide Susceptibility Mapping Using Different GIS-Based Bivariate Models

Ebrahim Nohani ¹, Meisam Moharrami ², Samira Sharafi ², Khabat Khosravi ³, Biswajeet Pradhan ^{4,5}, Binh Thai Pham ^{6,*}, Saro Lee ^{7,8,*} and Assefa M. Melesse ⁹

¹ Young Researchers and Elite Club, Dezful Branch, Islamic Azad University, Dezful 64616-45169, Iran

² Department of GIS and RS, Faculty of Geography and Planning, University of Tabriz, Tabriz 51666-16471, Iran

³ School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada

⁴ Faculty of Engineering and IT, University of Technology Sydney, Sydney, NSW 2007, Australia

⁵ Department of Energy and Mineral Resources Engineering, Choongmu-gwan, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Korea

⁶ Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam

⁷ Division of Geoscience Research Platform, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124 Gwahang-no, Yuseong-gu, Daejeon 305-350, Korea

⁸ Korea University of Science and Technology, 217 Gajeong-ro Yuseong-gu, Daejeon 34113, Korea

⁹ Department of Earth and environment, Florida International University, Miami, FL 33174, USA

* Correspondence: binhpt@utt.edu.vn (B.T.P.); leesaro@kigam.re.kr (S.L.)

Received: 12 May 2019; Accepted: 3 July 2019; Published: 8 July 2019



Abstract: Landslides are the most frequent phenomenon in the northern part of Iran, which cause considerable financial and life damages every year. One of the most widely used approaches to reduce these damages is preparing a landslide susceptibility map (LSM) using suitable methods and selecting the proper conditioning factors. The current study is aimed at comparing four bivariate models, namely the frequency ratio (FR), Shannon entropy (SE), weights of evidence (WoE), and evidential belief function (EBF), for a LSM of Klijanrestagh Watershed, Iran. Firstly, 109 locations of landslides were obtained from field surveys and interpretation of aerial photographs. Then, the locations were categorized into two groups of 70% (74 locations) and 30% (35 locations), randomly, for modeling and validation processes, respectively. Then, 10 conditioning factors of slope aspect, curvature, elevation, distance from fault, lithology, normalized difference vegetation index (NDVI), distance from the river, distance from the road, the slope angle, and land use were determined to construct the spatial database. From the results of multicollinearity, it was concluded that no collinearity existed between the 10 considered conditioning factors in the occurrence of landslides. The receiver operating characteristic (ROC) curve and the area under the curve (AUC) were used for validation of the four achieved LSMs. The AUC results introduced the success rates of 0.8, 0.86, 0.84, and 0.85 for EBF, WoE, SE, and FR, respectively. Also, they indicated that the rates of prediction were 0.84, 0.83, 0.82, and 0.79 for WoE, FR, SE, and EBF, respectively. Therefore, the WoE model, having the highest AUC, was the most accurate method among the four implemented methods in identifying the regions at risk of future landslides in the study area. The outcomes of this research are useful and essential for the government, planners, decision makers, researchers, and general land-use planners in the study area.

Keywords: landslide; bivariate models; ROC; GIS; Kelijanrestagh; multicollinearity

1. Introduction

A landslide is the movement of a debris, rock, or soil mass down the slope and is one of the most frequently occurring natural disasters in mountain areas that puts lives and properties of the people at risk [1–3]. Landslides occur in smaller scales than other natural disasters, but have higher distribution and are more dangerous in many cases [4]. Landslides lead to the evolution of landforms and are considered as the biggest threat in a great number of regions all around the world [5]. Based on the reports of the Centre for Research on the Epidemiology of Disasters (CRED), landslides are the cause of 17% of all casualties of natural hazards in the world [6]. Some researchers expect this trend to increase in the future with the increase in urbanization, deforestation, and changes in climate conditions [2,7]. The damages due to the occurrence of landslides are also predicted to rise in the subsequent decades with population growth, progression of residential areas and infrastructure in high-risk areas, continuing deforestation, and the increase in regional precipitation [3].

One of the main approaches for developing hazard reduction strategies is creating the landslide susceptibility map (LSM). The LSM can provide spatial distribution of potential slope failures, thus has a significant role in risk mitigation of landslides. Van Den Eeckhaut et al. [8] stated that landslides were more likely occur in the areas with the background of occurrences of landslides. A recent work combined LSM with rainfall threshold to have both spatial and temporal forecasting of landslide occurrences [9].

In the past two decades, substantial research studies on LSM have been conducted worldwide. Many scholars have tried different approaches to creating LSMs. In addition to crisis planning, LSMs are crucial for identifying the areas prone to the risk of landslides as well as managing and reducing the risk [10]. These maps can be provided using an appropriate model by having the landslide data and a set of independent variables [7]. There are three main groups of landslide susceptibility methods, including innovative, deterministic, and statistical [11]. The basis of the innovative models is the opinion of the experts in identifying the weight of each factor. Thus, this type of models has a high potential for error [12,13]. Deterministic models are developed on the basis of mathematical relationships. These models are based on the physical laws, which require calculating the relationship between the resistance forces and drivers of mass movements [14].

Recently, with the advancements in the commercial world, remote sensing, and quick access to geographic information system (GIS) data in natural hazard mapping, landslide modeling has been improved [15]. Many studies on LSM have been conducted using the frequency ratio (FR) [16,17], weight of evidence (WoE) [17], evidential belief function (EBF) [15], artificial neural networks [18–20], neuro-fuzzy systems [21,22], fuzzy logic [23,24], analytical hierarchy processes (AHPs) [25], Shannon entropy (SE) [6], logistic regression [5], and statistical index [3,5] in GIS. Some of the aforementioned techniques have been utilized in other fields of study, such as identification of flood-prone areas [26,27].

Zhang et al. [28] studied landslides in China by combining the statistical index and AHP methods to prepare LSMs. They stated that residential areas and sporadic forests with geological units of red layered moderate soft mixture of clastic rocks placed in altitude class of 0–200 m were quite prone to landslides. In Iran, landslides have been most frequently reported in Mazandaran province, Iran [29]. Klijanrestagh Watershed, located in Mazandaran, is one of the areas with a high risk of landslides due to special physiographic and climatic conditions, geological formations susceptible to the occurrence of landslides, and many villages in the highlands.

As landslides frequently occur in the area under study, this research study was basically aimed at investigating the areas with the potential of the occurrence of future landslides in order to manage and reduce losses. Another objective is to identify the major factors of the occurrence of landslides in the area under study using FR, SE, WoE, and EBF, and finally, evaluate the performance of these models in the identification of landslide-prone areas. The main research gap is the lack of a comprehensive study on the application of bivariate models. Thus, the main motivation of this study was to investigate and determine the prediction power of the most popular and widely used state-of-the-art statistical

bivariate models in a landslide susceptibility assessment in a Klijanrestagh Watershed, a hotspot area for landslide occurrences.

Klijanrestagh basin with an area of 638.9 km² is located between 53°2'0" to 53°21'0" east longitude and 36°11'40" to 36°30'20" north latitude in Mazandaran province in northern Iran (Figure 1). The highest and lowest altitudes of the area are 67 and 1667 m above the sea level. The slope is variable between 0 and 60 degrees. The average annual rainfall is 450 mm. According to Emberger climate classification, the area under study has a humid cold climate and most of the rainfall occurs in the fall and winter seasons. The majority of the area has Cenozoic sediments with formations of Mmsl. The basin consists of 12 different land uses and a major part of the area is allocated to semi-dense forests with an area of 446 km².

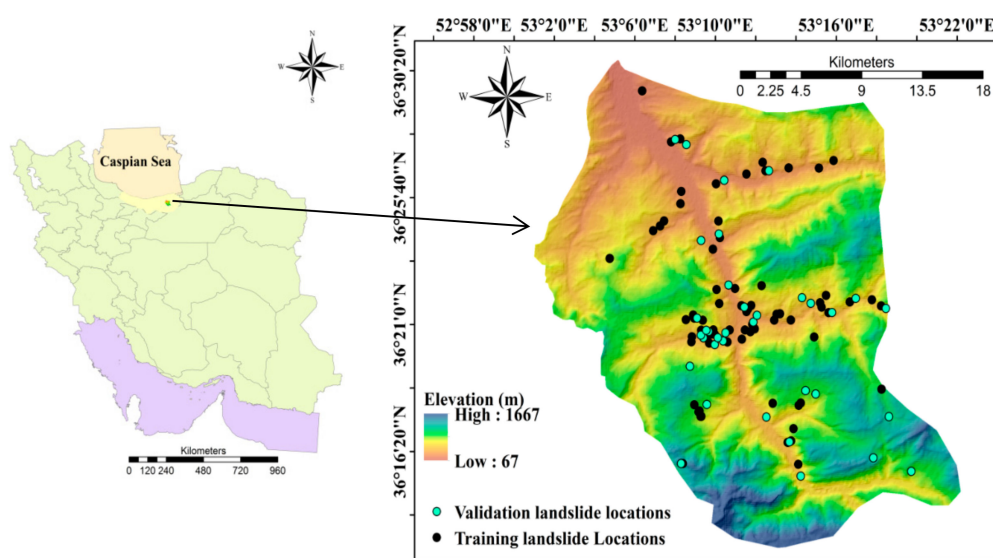


Figure 1. Landslide location map of the study area.

2. Materials and Methods

In this section, we describe the data collected and used and the background theory of the methods used for landslide susceptibility modeling. From these, the data including landslide susceptibility map and landslide conditioning factors were prepared and processed to generate the datasets for constructing and validating the models. The methods include bivariate models (FR, SE, WoE, and EBF) which were used for landslide susceptibility assessment, whereas validation methods (the ROC and AUC) were selected for validating the performance of the bivariate models.

2.1. Landslide Inventory Mapping

The basic phase in LSM is building an inventory map based on the past landslide occurrences [24]. The accuracy of the data on past and present landslides is crucial in predicting future landslides [11,30]. In this research, for collecting the required data, extensive surveys in the field were conducted and aerial photographs were interpreted. A sum of 109 landslides were determined and mapped for preparing the landslide inventory map (Figures 1 and 2). Chung and Fabbri [31] noted the impossibility of validating the functionality of a model without dividing the dataset. Accordingly, the inventory data was randomly categorized into two parts of 70% (74 locations) for building the model, which was called training dataset, and 30% (35 locations), which, not being used in the training step, was applied for validating the model performance, called the test dataset. The ratio of 70/30 for dividing the data for generating training and testing datasets was selected based on the suggestion of Pham et al. [16,30], Pradhan [19] and Pourghasemi et al. [24].



Figure 2. Some photos of landslides occurring in the area under study.

2.2. Landslide Conditioning Factors

To develop an assessment method for the assessment of susceptibility to landslides, identifying the landslide conditioning factors is crucial [32]. Landslide conditioning factors have some characteristics such as easy obtainability, representativeness, and practicality [21]. Ten conditioning factors affecting the occurrence of landslides were determined for LSM in this study based on the literature review and field surveys. The mechanism of landslide occurrence was identified to be related to hydrology, geomorphology, geology, etc. These factors were the slope aspect, curvature, elevation, distance from fault, lithology, normalized difference vegetation index (NDVI), distance from the river, distance from the road, the slope angle, and land use. In subsequent parts of the study, the first step is preparation of the digital elevation model (DEM). Using DEM of 30×30 m resolution (22,819 row and 9321 column), downloaded from ASTER global DEM, the slope angle, the slope aspect, elevation, and curvature were identified in ArcGIS 10.1 software.

2.2.1. Slope Aspect

The slope aspect is among the most significant factors affecting the occurrence of landslides due to various wetness of the aspect [33]. It affects hydrological processes because of evapotranspiration [34] and then, influences soil moisture and vegetation cover [35,36]. The information about the slope aspect was directly taken from DEM and divided into five categories, namely (1) flat, (2) north, (3) east, (4) south, and (5) west (Figure 3a).

2.2.2. Curvature

The morphology of the topography was identified by curvature [5]. This factor controls surface runoff and therefore, has an impact on landslide occurrence [37,38]. The curvature was straight drawn from the DEM and categorized into three classes of negative curvature, i.e., concave (< -0.05), zero curvature, i.e., flat (-0.05 – 0.05), and positive curvature, i.e., convex (> 0.05) (Figure 3b).

2.2.3. Elevation

The influence of elevation is not direct on landslide occurrences, but it affects other parameters such as rainfall, tectonics, etc. [39,40]. Similarly, this factor was taken from the DEM and categorized in six groups based on the landslide frequency analysis: (1) 67–300; (2) 300–600; (3) 600–900; (4) 900–1200; (5) 1200–1500; and (6) > 1500 (Figure 3c).

2.2.4. Distance from Fault

The fault leads to discontinuity of soil and rocks [41]. Generally, tectonic breaks reduce the rock strength and initiate many landslides [42]. The fault layer was derived from the fault map of Iran. Then, the distance from the fault was extracted using the multiple ring buffers command in ArcGIS10.1

with five classes of 0–100, 100–200, 200–300, 300–400, and >400 m, based on literature studies [41,42] (Figure 3d).

2.2.5. Lithology

The different formations have various structures, compositions, and permeability, which influence the formation material strength [43]. Therefore, the confluence of these factors with curvature and slope steepness has a remarkable influence on landslide occurrences [44]. The lithological feature was derived from the geological map of Mazandaran, which consisted of 22 units (Table 1 and Figure 3e).

Table 1. Lithological description of the study area.

No	Code	Lithology
1	Et.1	Andesitic volcanic tuff
2	PLQmc	Conglomerate, Sandston, siltstone, Silty Marl
3	JL2	Lime
4	K11	Hyporite bearing limestone (Senonian)
5	K2l	Thick-bedded to massive limestone (maastrichtian)
6	K2m	Lime and Quinasine
7	l	Marl, shale and detritic limestone
8	K2Pems	Marl, shale, limestone and mixture of conglomerate
9	M23m,s,l	Marl, calcareous sandstone, sandy limestone and minor conglomerate
10	M2msl	Marl and calcareous sandstone
11	Mg	Red conglomerate and sandstone
12	Mm2	Marl and calcareous
13	Mms	Marl, calcareous sandstone and sandy limestone
14	Pel	Medium to thick-bedded limestone
15	Pelm	Medium to thick-bedded limestone
16	Pemls	Dark grey medium-bedded to massive limestone
17	Plc	Polymictic conglomerate and sandstone
18	PLQcs	Conglomerate, Sandston and siltstone
19	Pr	Fusulina limestone, dolomitic limestone
20	Q2	New Alluvial
21	Qal	Loose alluvium in the river channels
22	TR3l,sh	Shale, Lime and Dolomite

2.2.6. NDVI

The NDVI map was generated through Landsat 7/ETM+ by ENVI 5 software (Figure 3f). The greater the NDVI value is, the more the amount of vegetation cover. The root of vegetation leads to stabilization of the hill slope and reduction in landslide occurrences.

2.2.7. Distance from River

The water of the river is considered as one of the main sources of e soil moisture. Binh Thai et al. [45] in their research stated that 65.12% of landslides occurred in the first class of the distance from the river (0–40 m). Accordingly, it is a very important factor in LSM. This factor was extracted for the river layer through buffering and classified into five classes including (1) 0–100, (2) 100–200, (3) 200–300, (4) 300–400, and (5) >400 m (Figure 3g).

2.2.8. Distance from Road

Tuan and Dan [46] stated that landslides were mostly distributed near the road system. By cutting the slope hills for construction of roads in the slopes more than 10 degrees, discontinuity is made in the soil and rock. Accordingly, the area can be prone to landslides [41]. The distances to the roads were obtained and divided into six classes of (1) 0–100, (2) 100–200, (3) 200–300, (4) 300–400, and (5) >400 m (Figure 3h).

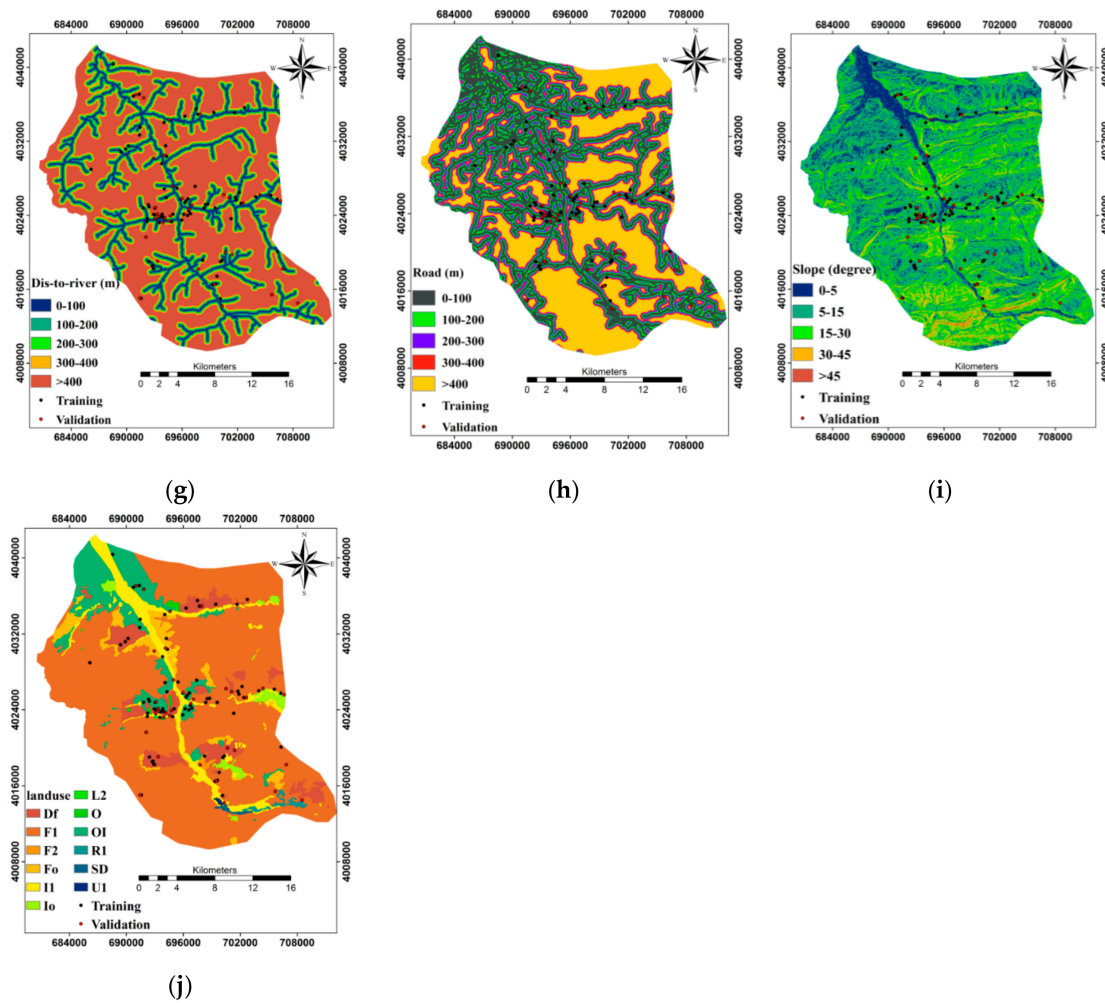


Figure 3. (a) Slope aspect, (b) curvature, (c) elevation, (d) distance from the fault, (e) lithology, (f) NDVI, (g) distance from the river, (h) distance from the road, (i) the slope angle, and (j) land use.

2.3. Bivariate Models

2.3.1. Frequency Ratio (FR)

As Bonham-Carter [51] stated, *FR* represents the possibility of the occurrence for a given characteristic. The *FR* method is on the basis of the observed relationships between the distribution of landslide occurrence locations and each of the factors influencing landslides and is used to reveal the correlation between them [52]. Greater *FR* values than 1 indicate the great proportion of landslide occurrences to the area and a high correlation. Conversely, the smaller *FR* values than 1 show a weak correlation [53]. Equation (1) shows the *FR* algorithm used in this study [54]:

$$W_{ij} = \frac{FL_{ij}}{FN_{ij}} \tag{1}$$

where W_{ij} shows the *FR* of class i of parameter j ; FL_{ij} is the frequency of occurrence of landslides for class i of parameter j ; and FN_{ij} is the frequency of non-occurrence of landslides for class i of parameter j .

To calculate and make the LSM, the *FR* values of parameters influencing landslides were added together [54]. Finally, the LSM was determined by adding the *FR* values of parameters together using Equation (2).

$$LSM = \sum_{j=1}^n W_{ij} \tag{2}$$

where W_{ij} is the weight of class i of parameter j , and n is the number of conditioning factors.

2.3.2. Shannon Entropy (SE)

Entropy is the measure of abnormality, changeability, unstable behavior, and uncertainty in each system [6]. Generally, the entropy concept is described as the quantity of abnormality between causes and results or in determination on various topics [55]. The entropy index is an estimate of the difference between the average shares of single groups in the whole system [55]. The SE model benefits from an extremely close relationship between the quantity and entropy, called Boltzmann, which is utilized to show the thermodynamic conditions of a system [6]. The Shannon method is adjusted on the basis of the Boltzmann concept and the entropy model is utilized in information theory [6]. The following formulae are used in the calculation of the information coefficient and V_j represents the parameter value to total value ratio [56], determined the equation below:

$$E_{ij} = \frac{FR}{\sum_{j=1}^{M_j} FR} \tag{3}$$

where FR represents the value of the frequency ratio and E_{ij} is the probability density for each class:

$$H_j = -\sum_{i=1}^{M_j} E_{ij} \log_2 E_{ij}, \quad j = 1, \dots, n \tag{4}$$

$$H_{jmax} = \log_2 M_j, \quad M_j\text{-number of classes} \tag{5}$$

$$I_j = (H_{jmax} - H_j / H_{jmax}), \quad I = (0, 1), \quad j = 1, \dots \tag{6}$$

$$V_j = I_j FR \tag{7}$$

where H_j and H_{jmax} are the values of entropy, I_j is information coefficient, M_j is number of classes in each conditioning factor, and V_j is the achieved weight value for the given parameter. The range of variations is between zero and 1. The values close to 1 have higher inconsistency and imbalance [27].

2.3.3. Weights of Evidence (WoE)

Some binary patterns may be joined together to predict another binary pattern for the WoE model. Many benefits of the WoE model over other statistical methods, including being data-driven based on the Bayesian probabilistic method, have made it a fairly popular statistical method [19].

The two basic parameters, which are very crucial to running the model, are negative weight (W^-) and positive weight (W^+) [19,27]. This method estimates the weight of each conditioning factor (A) on the basis of the presence or absence of occurrences (e.g., of landslide) within a specified area as in the following [51].

$$W^+ = \ln \frac{P\{B|A\}}{P\{B|\bar{A}\}} \tag{8}$$

$$W^- = \ln \frac{P\{\bar{B}|A\}}{P\{\bar{B}|\bar{A}\}} \tag{9}$$

where P is probability and \ln is the natural logarithm. B and \bar{B} are the presence or absence of landslide conditioning factors, respectively. Further, A and \bar{A} are the presence or absence of a landslide occurrence, respectively [26,57]. The variation of the parameter W^+ from W^- is called weight contrast [58].

The amount of the weight contrast displays the spatial relationship between factors affecting the occurrence of a landslide. The standard deviation of positive and negative weights, $S(c)$, is computed as follows:

$$S(C) = \sqrt{S^2 W^+ + S^2 W^-} \tag{10}$$

where SW^+ and SW^- are the variances of positive and negative weights, respectively. The variances of positive and negative weights can be estimated through the following equations:

$$S^2W^+ = \frac{1}{N\{B \cap A\}} + \frac{1}{B \cap \bar{A}} \quad (11)$$

$$S^2W^- = \frac{1}{\{\bar{B} \cap A\}} + \frac{1}{\bar{B} \cap \bar{A}} \quad (12)$$

The studentized contrast, i.e., the ratio of the contrast to its standard deviation, is used to calculate confidence:

$$W_{final} = (C/S(C)) \quad (13)$$

2.3.4. Evidential Belief Function (EBF)

The EBF model is fundamentally based on the theory of Dempster-Shafer of evidence [59]. The advantage of this model is that it presents predicted flood zones, landslide zones, etc. with uncertainty [60]. There are four basic parameters in EBF model, including degree of belief (Bel), degree of disbelief (Dis), degree of uncertainty (Unc), and degree of plausibility (Pls) [15,61].

It should be noted that suitable classification of all factors influencing landslides should have the same size of cells and appropriate time resolution. In this phase, L number of the input data layers can be considered as probability (Eij) in order to achieve correct results, in which (ij) represents i number of factors and j is the class characteristics. The total values of the EBF for factors affecting landslides were calculated one after the other by Equations (14) to (19). Equations (14) and (15) show how to achieve the results of Bel.

$$\lambda(Tp)Eij = N/D = [N[L \cap Eij]/N[L]]/[N(Eij) - N(L \cap Eij)/(N(A) N(L))] \quad (14)$$

$$Bel = \lambda(Tp)Eij / \left(\sum \lambda(Tp)Eij \right) \quad (15)$$

where $N(L \cap Eij)$ represents the number of landslide pixels per class, $N(L)$ stands for the sum of the occurrences of landslides, $N(Eij)$ is pixels per class for each factor, $N(A)$ denotes the sum of pixels in the area under study, N is the proportion of areas with landslide occurrences, and D represents the proportion of areas without landslides. These values are obtained by Equations (16) and (17).

$$\lambda(T\bar{p})Eij = \frac{K}{H} = [[N[L] - N[L \cap Eij]]/N[L]]/[N(A) - N(L) - N(Eij) + N(L \cap Eij)/(N(A) - N(L))] \quad (16)$$

$$Dis = \lambda(T\bar{p})Eij / \left(\sum \lambda(T\bar{p})Eij \right) \quad (17)$$

where K is the rate of landslides not occurring and H is the rate of occurrences in non-landslide areas. Equations (18) and (19) were utilized for all classes of factors affecting landslides and then, Equations (13) and (15) were used to prepare the results for Bel and Dis. The values of Pls and Unc were determined by the following equations.

$$Unc = 1 - Dis - Bel \quad (18)$$

$$Pls = 1 - Dis \quad (19)$$

The range of Bel and Pls values is between 0 and 1. The most important limitation of the EBF, according to the findings of Park [60], is that if there is no occurrence of landslide in a specific class, then Bel results are equal to zero. Dis is also be equal to zero and in this case, Unc or uncertainty values are equal to 1.

2.4. Model Verification

To verify the performance of LSM, in this study, the authors have used the receiver operating characteristic (ROC) curve and the area under the curve (AUC). ROC is one of the most popular methods [62,63] and it is a practical technique of showing the quality of deterministic and probabilistic detections [64] used to determine the accuracy of the LSMs [23]. The ROC curve was drawn by plotting specificity on the X axis and sensitivity on the Y axis. Sensitivity is the false positive rate and specificity is the false negative as presented below:

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \quad (20)$$

$$\text{Specificity} = \text{TN}/(\text{FP} + \text{TN}) \quad (21)$$

where true positive (TP) is the number of observed landslides predicted accurately and true negative (TN) is the total non-occurring landslides that have been predicted accurately. On the other hand, false positive (FP) is considered as the number of occurring landslides inaccurately categorized in the non-landslide classes and false negative (FN) is considered as the number of non-occurring landslides inaccurately categorized in the landslide classes.

The AUC can identify the model's accuracy and ability in predicting future landslides. The range of the AUC ranges from 0.5 to 1 with the AUC of 1 representing perfect prediction and the closer the value of the AUC to this number, the better the performance of the model [58].

3. Results

3.1. Multicollinearity Diagnosis

In this study, a multicollinearity test was conducted by variance inflation factors (VIF) and the tolerances method for selection of suitable landslide conditioning factors to identify whether there was correlation between the conditioning factors (Table 2). The values of VIF > 10 or tolerance < 0.1 show a problem of multicollinearity in conditioning factors [6]. According to the multicollinearity test results (Table 2), no collinearity was observed among the 10 conditioning factors affecting the landslide.

Table 2. Indices of multicollinearity diagnosis for landslide conditioning factors.

Model	Collinearity Statistics	
	Tolerance	VIF
Slope aspect	0.353	2.833
Curvature	0.300	3.331
Altitude (m)	0.806	1.240
Distance from fault	0.774	1.293
NDVI	0.332	3.016
Lithology	0.784	1.275
Distance from river (m)	0.714	1.401
Distance from road (m)	0.876	1.142
Slope angle (degree)	0.876	1.142
Land use	0.377	2.656

3.2. FR Model

The results of the FR method showed the weights obtained for each of the classes of factors. As represented in Table 3, the FR was calculated for 10 factors based on their relationships with landslide occurrences. The ratio of the area with landslide occurrences to the whole area was identified. The higher the value of the FR is, the stronger the correlation of the factors with landslide occurrences. The greater values of the FR than 1 show a strong correlation and those below 1 show a weak correlation [65].

Table 3. The correlation between occurring landslides and each conditioning factor using bivariate models.

	Landslide Number	Percent of Landslide	Number of Pixels	Percent of Pixels	FR	SE	WoE	BEL	DIS	UNC	PLS
Factors											
Aspect											
Flat	0	0.00	32625	2.04	0.00		None	0.00	0.20	0.795	0.79
North	33	44.00	448162	28.05	1.57		3.01	0.39	0.15	0.448	0.84
East	12	16.00	346063	21.66	0.74	0.14	−1.18	0.18	0.21	0.598	0.78
South	14	18.67	387209	24.24	0.77		−1.12	0.19	0.21	0.590	0.78
West	16	21.33	383399	24.00	0.89		−0.54	0.22	0.20	0.568	0.79
Curvature											
Concave	8	10.67	198991	12.46	0.86		−0.47	0.21	0.20	0.579	0.79
Flat	51	68.00	1091866	68.35	0.98	0.005	−0.07	0.25	0.20	0.546	0.79
Convex	16	21.33	306601	19.19	1.11		0.47	0.28	0.19	0.524	0.80
Elevation											
67–300	18	24.00	241935	15.14	1.58		2.11	0.40	0.18	0.421	0.82
300–600	47	62.67	682822	42.74	1.47		3.39	0.37	0.13	0.499	0.86
600–900	9	12.00	490349	30.70	0.39	1.804	−3.32	0.09	0.25	0.646	0.74
900–1200	1	1.33	156384	9.79	0.14		−2.07	0.03	0.22	0.746	0.78
1200–1500	0	0.00	24115	1.51	0.00		None	0.00	0.20	0.796	0.79
>1500	0	0.00	1853	0.12	0.00		None	0.00	0.20	0.799	0.79
Distance to Fault											
0–100	14	18.67	225014	14.09	1.33		1.13	0.33	0.19	0.476	0.81
100–200	4	5.33	212463	13.30	0.40		−1.95	0.10	0.21	0.680	0.78
200–300	8	10.67	190586	11.93	0.89	0.035	−0.34	0.22	0.204	0.571	0.79
300–400	7	9.33	165573	10.37	0.90		−0.29	0.22	0.203	0.570	0.79
>400	42	56.00	803733	50.32	1.11		0.98	0.28	0.178	0.542	0.82
Lithology											
EtI	0	0.00	7587	0.47	0.00		None	0.00	0.202	0.798	0.79
PLQmc	0	0.00	1586	0.10	0.00		None	0.00	0.201	0.799	0.79
DAM	0	0.00	1620	0.10	0.00		None	0.00	0.201	0.799	0.79
JL2	0	0.00	2113	0.13	0.00		None	0.00	0.201	0.799	0.79
K11	0	0.00	61922	3.88	0.00		None	0.00	0.209	0.791	0.79
K21	0	0.00	21591	1.35	0.00		None	0.00	0.204	0.796	0.79
K2m	0	0.00	48897	3.06	0.00		None	0.00	0.207	0.793	0.79
K2ml	0	0.00	17979	1.13	0.00		None	0.00	0.203	0.797	0.79
K2pem, s	0	0.00	8257	0.52	0.00		None	0.00	0.202	0.798	0.79
M2,3m, s, l	0	0.00	935299	58.55	0.00		None	0.00	0.484	0.516	0.51
M2,3msl	39	52.00	857	0.05	969.2		32.75	256	0.096	−255	0.90
Mg	0	0.00	76790	4.81	0.00	40.7	None	0.00	0.211	0.789	0.78
Mm2	1	1.33	57006	3.57	0.37		−1.00	0.09	0.205	0.700	0.79
Mms	2	2.67	61271	3.84	0.70		−0.52	0.17	0.203	0.621	0.79
Pel	0	0.00	30284	1.90	0.00		None	0.00	0.205	0.795	0.79
Pelm	1	1.33	21613	1.35	0.99		−0.01	0.24	0.201	0.551	0.79
Pem,l,s	0	0.00	4655	0.29	0.00		None	0.000	0.201	0.799	0.79
PLC	0	0.00	109490	6.85	0.00		None	0.00	0.216	0.784	0.78
PLQCs	17	22.67	11766	0.74	30.77		13.33	7.77	0.156	−6.927	0.84
Pr	0	0.00	2445	0.15	0.00		None	0.00	0.201	0.799	0.79
Q2	1	1.33	85342	5.34	0.50		−1.42	0.06	0.209	0.728	0.79
Qal	13	17.33	19138	1.20	14.47		9.34	3.65	0.168	−2.818	0.83
TR31, sh	1	1.33	9909	0.62	4.30		0.77	0.54	0.199	0.259	0.80
NDVI											
<0.15	24	32.00	104780	6.56	4.88		7.69	1.23	0.146	−0.376	0.85
0.15–0.3	22	29.33	116634	7.30	4.02		6.55	1.01	0.153	−0.166	0.84
0.3–0.45	10	13.33	210087	13.15	1.01	0.57	0.05	0.25	0.200	0.544	0.80
0.45–0.55	13	17.33	543431	34.02	0.51		−2.95	0.12	0.252	0.620	0.74
>0.55	6	8.00	622515	38.97	0.21		−4.68	0.05	0.303	0.646	0.69

Table 3. Cont.

	Landslide Number	Percent of Landslide	Number of Pixels	Percent of Pixels	FR	SE	WoE	BEL	DIS	UNC	PLS
Distance to River											
0–100	13	17.33	157133	9.84	1.76		2.14	0.44	0.184	0.372	0.81
100–200	15	20.00	168234	10.53	1.90		2.61	0.47	0.180	0.342	0.82
200–300	9	12.00	182699	11.44	1.05	0.06	0.15	0.26	0.200	0.536	0.80
300–400	15	20.00	196794	12.32	1.62		2.00	0.40	0.183	0.407	0.81
>400	23	30.67	892509	55.87	0.55		−4.20	0.13	0.315	0.546	0.68
Distance to Road											
0–100	36	48.00	422355	26.44	1.82		4.08	0.45	0.142	0.400	0.85
100–200	18	24.00	291134	18.23	1.32		1.29	0.33	0.187	0.481	0.81
200–300	6	8.00	210762	13.19	0.61	0.11	−1.31	0.15	0.213	0.634	0.78
300–400	8	10.67	155134	9.71	1.10		0.28	0.27	0.199	0.524	0.80
>400	7	9.33	517984	32.43	0.29		−3.88	0.07	0.269	0.658	0.73
Slope (degree)											
0–5	7	9.33	231304	14.48	0.64		−1.25	0.16	0.213	0.625	0.78
5–15	39	52.00	804153	50.34	1.03		0.29	0.26	0.194	0.545	0.80
15–30	26	34.67	477816	29.91	1.16	0.11	0.90	0.29	0.187	0.521	0.81
30–45	3	4.00	78286	4.90	0.82		−0.36	0.20	0.203	0.591	0.79
>45	0	0.00	5899	0.37	0.00		None	0.00	0.202	0.798	0.79
Land use											
Df	18	24.00	111868	7.00	3.43		5.30	0.86	0.164	−0.028	0.83
F1	25	33.33	1115519	69.82	0.48		−6.25	0.12	0.444	0.436	0.55
F2	1	1.33	3747	0.23	5.69		1.74	1.43	0.199	−0.632	0.80
Fo	3	4.00	65227	4.08	0.98		−0.04	0.24	0.201	0.552	0.79
I1	4	5.33	111053	6.95	0.77		−0.55	0.19	0.204	0.602	0.79
Io	1	1.33	23016	1.44	0.93	0.62	−0.08	0.23	0.201	0.566	0.79
L2	0	0.00	238	0.01	0.00		None	0.00	0.201	0.799	0.79
O	2	2.67	5088	0.32	8.37		3.00	2.11	0.196	−1.308	0.80
OI	21	28.00	150690	9.43	2.97		5.12	0.74	0.160	0.092	0.84
R1	0	0.00	7197	0.45	0.00		None	0.00	0.202	0.798	0.79
SD	0	0.00	3046	0.19	0.00		None	0.00	0.201	0.799	0.79
U1	0	0.00	1015	0.06	0.00		None	0.00	0.201	0.799	0.79

The results showed that among the geographical directions, north weighing 1.57 had the most significant impact on the landslide occurrences in Klijanrestagh Watershed in Mazandaran province. The lowest weight was for the flat areas where no landslide had occurred.

Regarding the conditioning factor of curvature, convex curvature with the value of 1.1 had the greatest impact and concave and flat curvatures respectively weighing 0.86 and 0.98 had very small impact on the landslides in the area under study. The other important factor that may affect landslide is elevation. The lowest elevation (67–300 m) had the greatest impact on the landslide (1.58) followed by the second class (300–600), which had a weight of 1.47. Other elevation classes had a low impact on landslides. To study the influence of regional faults on landslides, the impact of different distances from the fault on landslides was calculated and the outcomes revealed that the first class (distance of 0 to 100 m) weighing 1.33 had the highest impact on landslide occurrences in the area under study. The units of PLQCs, Qal, and TR31, sh with the respective weights of 30, 14, and 4 had a high impact on landslides in the area under study.

The lowest NDVI value had the highest impact on landslides. Another factor affecting landslides is the distance from the river. In the results, the first 4 classes (from 0 to 400 m) had positive effects on landslides and the effects were smaller for classes of more than 400 m. Among various distances in the road network, the first class (0–100 m) weighing 1.8 had the greatest impact on landslides and then was the second class (100–200 m) with a weight of 1.3. The analysis of the FR method showed that the

third class (15–30) and second slope (5–15) with FR values of 1.16 and 1.03 exerted the greatest effect on landslides occurring in the area under study.

Regarding land use, the garden with the weight of 8.3 and dry farming with the weight of 3.4 had the most significant effects on landslide occurrences. Equation (22) was used in GIS and finally, LSM was prepared using the FR model.

$$LSM_{FR} = FR_{Slope} + FR_{Curvature} + FR_{Dis-River} + FR_{elevation} + FR_{Landuse} + FR_{Dis-Road} + FR_{Dis-fault} + FR_{aspect} + FR_{Lithology} + FR_{NDVI} \quad (22)$$

3.3. SE Model

The results of SE for the relationship of the effective factors with the occurrence of landslides are presented in Table 3. Lithology factors, elevation classes, land use, and NDVI respectively weighing 40.7, 1.8, 0.62, and 0.57 are the most important factors affecting landslides in Klijanrestagh Watershed. Other factors in order of importance and the effects on landslides are aspect, slope, distance from the road, distance from the river, distance from the faults, and terrestrial curvature weighing 0.14, 0.11, 0.09, 0.06, 0.03, and 0.005, respectively. Using Equation (23), the LSM was prepared using the SE model.

$$LSM_{SE} = 0.14 * Aspect_{FR} + 0.005 * curvature_{FR} + 1.804 * elevation_{FR} + 0.35 * fault_{FR} + 40.7 * lito_{FR} + 0.57 * NDVI_{FR} + 0.065 * Dis - River_{FR} + 0.094 * Dis - road + 0.112 * Slope_{FR} + 0.62 * LULC_{FR} \quad (23)$$

3.4. WoE Model

Like the FR model, this model is utilized to attain the weights of factors affecting landslides and the relationship of landslide occurrences with the related classes. The results of this model in this study are shown in Table 3. The WoE model is based on W^+ and W^- . W^+ indicates that predictable variables are present in landslide locations and its magnitude represents the effect magnitude of the condition factors on landslides [66]. W^- indicates the absence of conditioning factors in landslide occurrences and indicates a negative correlation between factors and the occurrence of landslides [3]. Therefore, C for the positive spatial association is positive and for the negative spatial association is negative.

S^2W^+ and S^2W^- are variances of W^+ and W^- , respectively. $S(C)$ is the standard deviation of contrast and $C/S(C)$ is studentized values of contrast (Table 3). $C/S(C)$ is the final weight of the WoE method. The map of effective weighted factors was prepared by assigning this weight to each factor. Finally, the LSM was produced implementing the following equation.

$$LSM_{WoE} = Aspect_{WoE} + curvature_{WoE} + elevation_{WoE} + fault_{WoE} + litjo_{WoE} + NDVI_{WoE} + Dis - river_{WoE} + Dis - road_{WoE} + Slope_{WoE} + LULC_{WoE} \quad (24)$$

Considering the values of $C/S(C)$ in Table 3 for the aspect, the northern class of the aspect had the most significant effect (3.05) on landslides. Other classes, having negative weights, showed no significant correlation with the occurrence of landslides. The results showed that the convex curvature had a significant positive effect (0.47) and concave and flat curvatures had negative effects on landslides. Like in the FR method, the first 2 classes of the elevation conditioning factor had the largest impact and the rest had lower effects on landslide occurrences. The elevation of 300–600 m weighing 3.39 had the greatest effect on landslides. Class 0–100 m of the distance from the fault weighing 1.13 exerted the most remarkable effect on landslides, and classes 100–200 and 200–300 m had the lowest effects on landslides in the area under study.

Regarding the lithology factor, the M2msl unit weighing 32 had the highest impact on landslide occurrences. Another effective factor affecting landslides was NDVI. The lowest NDVI values had the highest impact on landslides. Therefore, the landslide occurrence probability was reduced with the increase in the NDVI value. The risk of landslides decreased with the distance from the river, because moisture, as a major cause of landslide occurrences, decreased. The first class of the distance from the river (100–200 m) weighing 2.61 had the greatest effect on landslides. As with the distance from

the river, the risk of landslides was reduced with the distance from the road. In this study, the first class of the distance from the road (0–100 m) with a weight of 4 had the greatest effect on landslides. The probability of the occurrence of landslides increased with the slope up to 30 degrees and then, decreased. The highest effect on landslides was related to the slope class of 15–30 degrees. Based on this method, dry farming had the greatest effect on landslides, while the FR method introduced garden use for the greatest effect. After dry farming weighing 5.3, the combination of garden and dry farming weighing 5.1, and garden weighing 3 as well as the forest with the average density had the highest effects on landslides, respectively.

3.5. EBF Model

The EBF method was utilized to explore the relationship between locations of the occurring landslides and the landslide conditioning factors. The results depicted the weight of each class of factors in landslides (Table 3).

In the map of the aspect, the northern direction with 44% events had the highest probability of the occurrence of landslides weighting 0.39, as it was less exposed to sunlight and had higher moisture than other directions. The results of the curvature map analysis showed that the convex curvature had the greatest impact on landslides.

As in the FR method, the first class of elevation weighing 4 had the greatest effect on landslides in this method. Class 0–100 m of the distance from the fault had the greatest effect on landslides as 19% of landslide events were observed in this distance, which consisted in only 14% of the basin area. In the case of the lithology factor, the M2,3msl and PLQC units weighing 256 and 7, respectively, had the greatest effects on the occurrence of landslides in Klijanrestagh Basin.

Another factor that would affect landslides and was considered in this study was NDVI. The lowest NDVI value, which was less than 0.15 and weighed 1.2, had the highest impact on the occurrence of landslides. The results of the distance from the river and locations of the occurring landslides showed that the risk of landslides decreased with the distance from the river. Therefore, the second class (100–200) and first class (0–100 m) had the highest effect on landslides occurring in the study area.

The landslide risk decreased with the increase in the distance from the road. The first class (0–100 m) had the greatest effect (BEL = 0.45) on landslides. Considering the slope factor, the slope of 15–30 degrees with BEL of 0.29 had the greatest effect and the slopes more than 45% (BEL = 0) had the lowest effect on landslides. The highest weights were for garden (BEL = 2.1) and forest (BEL = 1.4).

Figure 4 presents the EBF results. The map of bel (Figure 4a) shows the distribution of landslides. A comparison was drawn between belief and disbelief maps (Figure 4b), and the results showed that areas with higher bel values had lower dis and vice versa, that is, the areas more prone to landslides were those with higher degrees of belief and lower degrees of disbelief about the occurrence of landslides. According to Unc map, areas with the highest sensitivity to landslides had the minimum values of Unc. Pls map is almost similar to bel and represents great values for areas with high belief values and low uncertainty. The results comply well with Carranza and Hale [61], Bui et al. [67] and Nampak et al. [68].

Finally, the LSM was utilized through the relationships provided above for the FR, SE, WoE, and EBF models and then, four maps were prepared based on the quantile method, which is one of the most popular classification methods [26], which are shown in Figure 5.

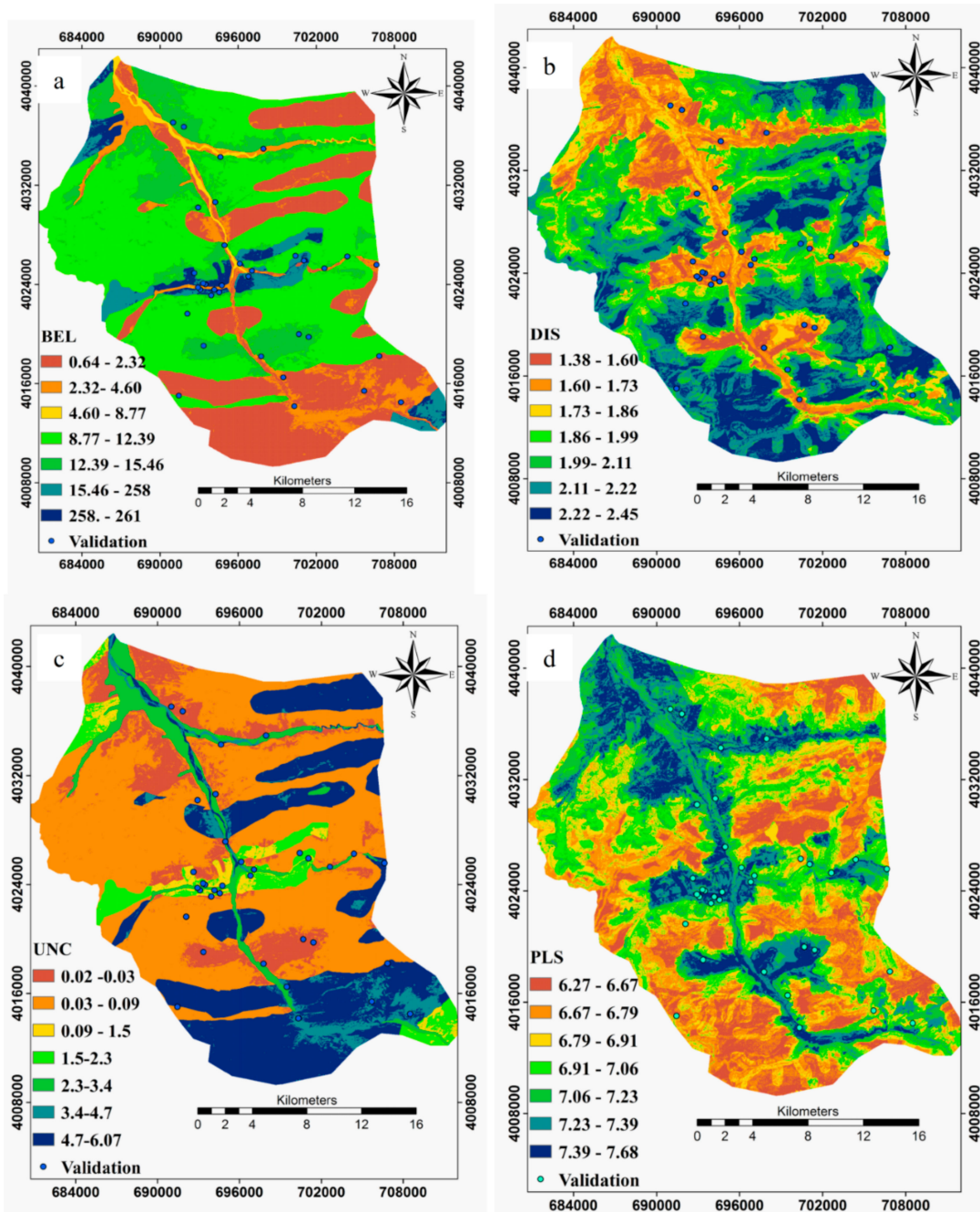


Figure 4. The combined results of the EBF model: (a) belief, (b) disbelief, (c) uncertainty, and (d) plausibility.

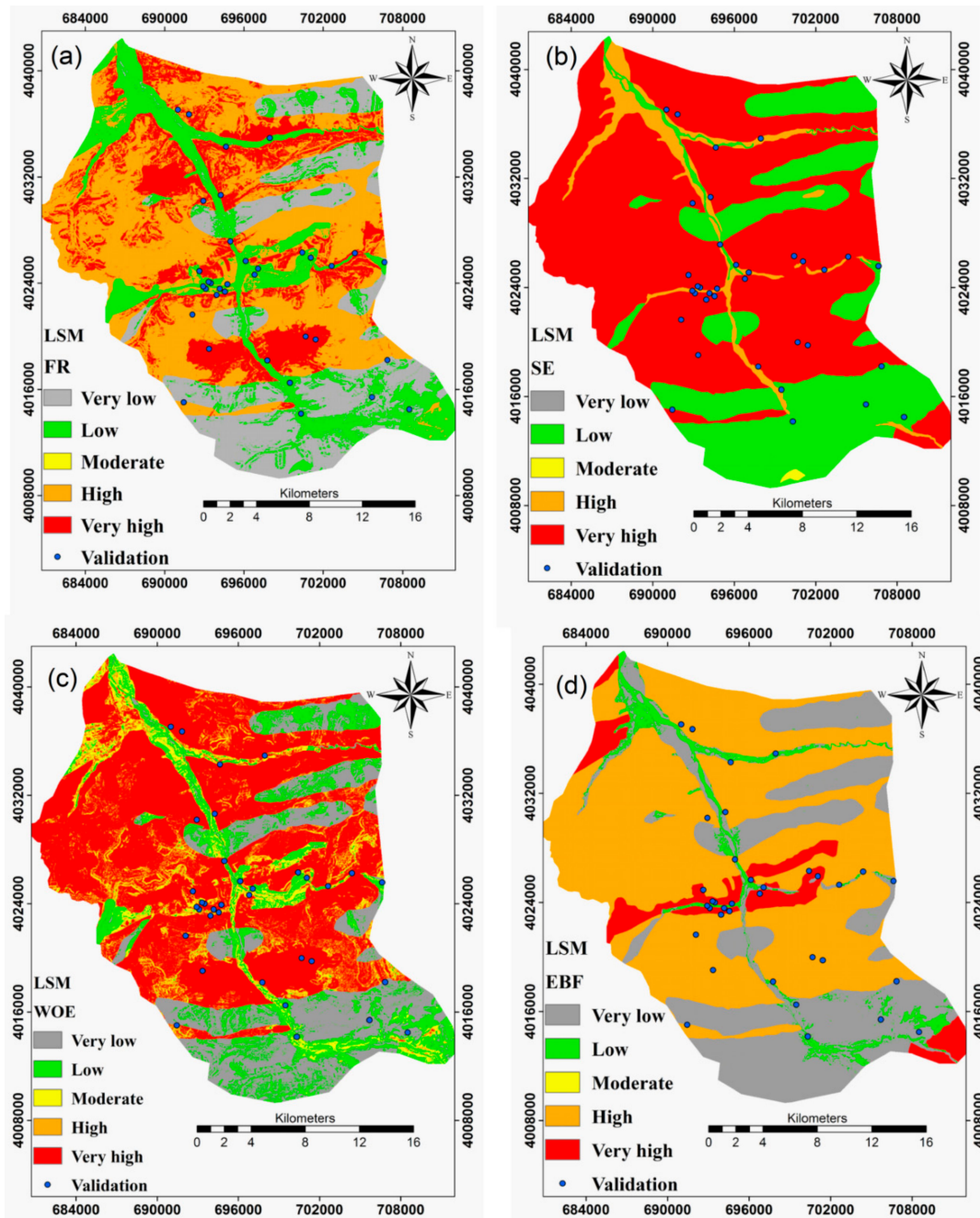


Figure 5. Landslide susceptibility maps produced by (a) FR, (b) SE, (c) WoE, and (d) EBF models.

3.6. Assessment and Comparison of the Models

In this research study, the well-known ROC and area under the curve (AUC) techniques were applied to assess and compare the models. There are two main steps in model evaluation by AUC, namely, the success rate evaluation and the prediction rate evaluation. The results for the success rate were achieved on the basis of training data and the prediction rates were attained by a set of validation data. The results of the success rate represented the fitness of the model for the training data and because the training points on which the model was built were used in the success rate. The success rate method was not useful in assessing the predicting power of the model [68].

The area under the curve of the success rate for the FR, SE, WoE, and EBF models was 0.85, 0.84, 0.86, and 0.8, respectively (Figure 6a). Further, the area under the curve of the prediction rate for the FR, SE, WoE, and EBF models was 0.83, 0.82, 0.84, and 0.79, respectively (Figure 6b). The AUC

amount of 0.84 for the WoE shows 84% prediction accuracy of the model. Considering the results of the evaluation by the AUC, the most accurate prediction in the identification of LSM is for the WoE model and the least accurate prediction is for the EBF. However, generally, prediction accuracies of these four models are reasonable.

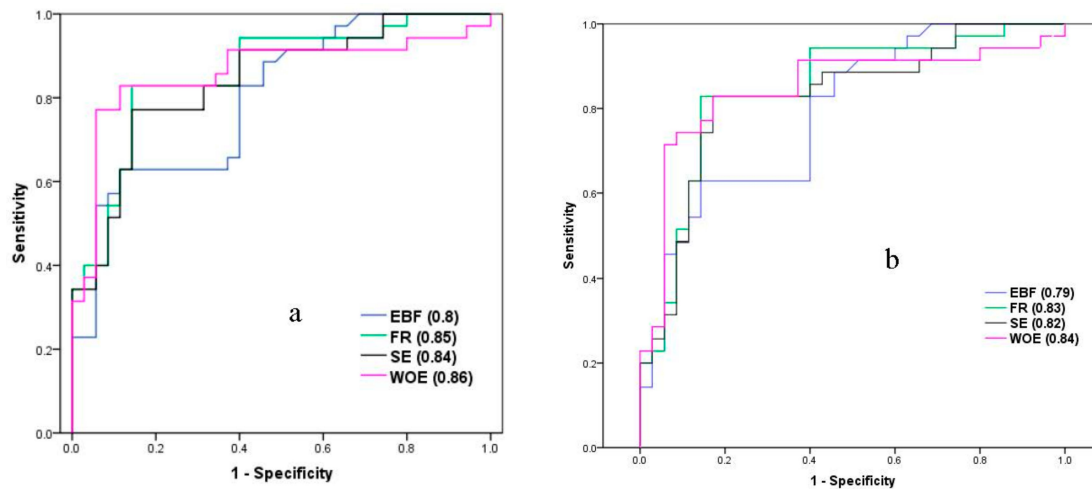


Figure 6. Model validation for Klijanrestagh Watershed: (a) success rate curve and (b) prediction rate curve.

4. Discussion

The preparation of landslide susceptibility is one of the most practical approaches for landslide hazard assessment and proper management tasks [69]. Although, there are many models developed for the LSM, there is no a universal guideline for model selection to model better LSMs. In this study, the four most popular and widely used state-of-the art statistical bivariate models (i.e., FR, SE, WOE and EBF) were applied for the LSM in Klijanrestagh Watershed, Iran.

Modeling results mostly depend on both the model's structure and quality of data [70]. The most effective data were selected and their quality was investigated through a multicollinearity test and the SE model. The results of these two tests showed that all applied conditioning factors are effective and independent. According to the SE model, the aspect, the slope and the distance from the road were the most effective factors and curvature was the least effective factor on landslide occurrences. This result is in accordance with the results of other research reports [70–75]. The improper road construction in the study area can be revealed as the one of the most effective criteria on landslide occurrences.

The results showed that the north aspect had the highest effect on landslide occurrences, and the main reason may be that the north aspect has the most wetness among other aspects resulting from less solar radiation [70]. Regarding the curvature, convex has the most effect on landslide occurrences, especially when combined with improper road construction that cause failure slopes. The risk of landslides decreased with the elevation in the area under study as the depth of soil reduced sharply. The results show that there is a linear relationship between landslide occurrences and faults, as faults are a disconnection in the earth and thus, can cause landslides more easily. The lowest NDVI value had the highest impact on landslides. This is because a higher NDVI value means higher vegetation cover, which can help to prevent landslides. The distances of 100 to 200 m from the river and road exerted the greatest impact on landslides occurring in the area under study since the closer to the river the higher the wetness, therefore, the higher the landslide occurrences probability. These findings are consistent with results of other literature studies [75–77]. These results are approximately consistent among the FR, WoE and EBF models.

This study showed that WOE has the highest prediction power, followed by FR, SE and EBF, respectively. According to Park [60], the main limitation of the EBF method is that if a landslide event does not happen in a given class, the Bel results will be equal to zero. In addition, when the Unc

or uncertainty values are equal to 1, Dis will be zero. But EBF model despite its lower capability for prediction of landslide-prone areas benefited from belief, disbelief, plausibility, and uncertainty maps, which were very practical. Thus, it can be concluded that the results of the predictions about the landslide-prone areas by modeling the uncertainty degree (Unc) show the relationships between landslide occurrences and multiple spatial data layers quantitatively [60]. The power of prediction of the FR was highly identical to that of SE method, which complies well with Youssef et al. [17].

In general, WOE model was found the best model in this study for identification of landslide-prone areas as this data-driven method has some advantages, which may lead to its high accuracy in comparison with other methods. They are (i) WoE model benefits from the Bayesian-based theorem in a log-linear form using prior and posterior probabilities and (ii) WoE model is effectively utilized when sufficient dataset is in hand for predicting the relative significance of evidential themes using statistical tools.

5. Conclusions

Landslide susceptibility modeling for the future is one of the vital and difficult tasks in the assessment of the risk of landslides. Many methods have been proposed for spatial prediction of natural hazards including landslides, e.g., bivariate statistical, multivariate statistical, data mining, etc. However, the accuracy of the predictions of these models is still under debate as their accuracy varies from one watershed to another. Kelijanrestagh Watershed in Mazandaran province, Iran, is a hot spot of landslide occurrences with many landslides occurring in each year. In this research, four models, namely FR, SE, WoE, and EBF, were utilized in landslide susceptibility mapping of Kelijanrestagh Watershed in Mazandaran province, Iran.

A total of 109 locations of landslides were determined by extensive surveys in the field and mapped in a landslide inventory map. The locations were categorized into two main datasets for training and validation. In the next step, 10 landslide conditioning factors were selected by a review of the literature and on the basis of the characteristics of the area under study and availability of data. Finally, the LSMs were prepared and mapped with five classes (very low, low, moderate, high and very high) using four models.

Next, the four produced maps were evaluated using training and testing datasets and ROC and AUC (success rate and prediction rate) methods. The AUC results indicated the success rates of 0.8, 0.86, 0.84, and 0.85 for EBF, WoE, SE, and FR, respectively. Further, they presented the prediction rates of 0.84, 0.83, 0.82, and 0.79 for WoE, FR, SE, and EBF models, respectively. Therefore, the WoE model, having the greatest AUC, was the most accurate method among the four implemented methods in identifying areas at risk of future landslides in the area under study.

The outcomes of this research can be helpful for decision makers and future development planners to avoid the zones with high susceptibility during implementing a project. However, it is recommended that a comparative study needs to be conducted (i.e., comparison between data mining algorithms, data intelligence, statistical bivariate models, multivariate models, multi criteria decision making models) applied for solving this challenge of which model should be applied, which precision and accuracy is reasonable, and so on. The field survey is a difficult approach for the preparation of the landslide inventory map, especially for elevation which often affects landslide distribution. Landslides that occur in high altitude areas are often lost, because of the difficulty of accurate field surveys. Thus, it is recommended that identification of landslide locations should be based on high-resolution satellite images.

Author Contributions: E.N., M.S. and S.S. collected the data, K.K. and B.T.P. performed modeling and analysis. E.N., M.S., S.S., K.K. and B.T.P. wrote the manuscript. B.P., S.L. and A.M.M. provided critical comments in planning this paper and edited the manuscript. All the authors discussed the results and edited the manuscript.

Funding: This research was conducted by the Basic Research Project of the Korea Institute of Geoscience and Mineral Resources (KIGAM) funded by the Ministry of Science and ICT.

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

References

- Chakraborty, S.; Pradhan, R. Development of GIS based landslide information system for the region of East Sikkim. *Int. J. Comput. Appl.* **2012**, *49*, 5–9. [[CrossRef](#)]
- Goetz, J.N.; Guthrie, R.H.; Brenning, A. Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology* **2011**, *129*, 376–386. [[CrossRef](#)]
- Regmi, A.D.; Devkota, K.C.; Yoshida, K.; Pradhan, B.; Pourghasemi, H.R.; Kumamoto, T.; Akgun, A. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab. J. Geosci.* **2014**, *7*, 725–742. [[CrossRef](#)]
- Trigila, A.; Iadanza, C.; Esposito, C.; Scarascia-Mugnozza, G. Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampileri (NE Sicily, Italy). *Geomorphology* **2015**, *249*, 119–136. [[CrossRef](#)]
- 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](#)]
- 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](#)]
- Kanungo, D.; Arora, M.; Sarkar, S.; Gupta, R. Landslide Susceptibility Zonation (LSZ) Mapping—A Review. *J. South Asia Disaster Stud.* **2009**, *2*, 81–105.
- Van Den Eeckhaut, M.; Vanwalleghem, T.; Poesen, J.; Govers, G.; Verstraeten, G.; Vandekerckhove, L. Prediction of landslide susceptibility using rare events logistic regression: A case-study in the Flemish Ardennes (Belgium). *Geomorphology* **2006**, *76*, 392–410. [[CrossRef](#)]
- Monsieurs, E.; Dewitte, O.; Demoulin, A. A susceptibility-based rainfall threshold approach for landslide occurrence. *Nat. Hazard Earth Syst. Sci.* **2019**, *19*, 775–789. [[CrossRef](#)]
- Holec, J.; Bednarik, M.; Šabo, M.; Minár, J.; Yilmaz, I.; Marschalko, M. A small-scale landslide susceptibility assessment for the territory of Western Carpathians. *Nat. Hazards* **2013**, *69*, 1081–1107. [[CrossRef](#)]
- Reichenbach, P.; Rossi, M.; Malamud, B.; Mihri, M.; Guzzetti, F. A review of statistically-based landslide susceptibility models. *Earth Sci. Rev.* **2018**, *180*, 60–91. [[CrossRef](#)]
- Dahal, R.K.; Hasegawa, S.; Nonomura, A.; Yamanaka, M.; Masuda, T.; Nishino, K. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environ. Geol.* **2008**, *54*, 311–324. [[CrossRef](#)]
- Hjort, J.; Luoto, M. 2.6 Statistical Methods for Geomorphic Distribution Modeling. In *Treatise on Geomorphology*; Academic Press: San Diego, CA, USA, 2013; pp. 59–73.
- Tsangaratos, P.; Ilia, I.; Rozos, D. Case Event System for Landslide Susceptibility Analysis. In *Landslide Science and Practice*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 585–593.
- Althuwaynee, O.F.; Pradhan, B.; Lee, S. Application of an evidential belief function model in landslide susceptibility mapping. *Comput. Geosci.* **2012**, *44*, 120–135. [[CrossRef](#)]
- Pham, B.T.; Tien Bui, D.; Indra, P.; Dholakia, M. Landslide susceptibility assessment at a part of Uttarakhand Himalaya, India using GIS-based statistical approach of frequency ratio method. *Int. J. Eng. Res. Technol.* **2015**, *4*, 338–344.
- Youssef, A.M.; Pradhan, B.; Jebur, M.N.; El-Harbi, H.M. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. *Environ. Earth Sci.* **2015**, *73*, 3745–3761. [[CrossRef](#)]
- Lee, S.; Ryu, J.-H.; Lee, M.-J.; Won, J.-S. Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. *Environ. Geol.* **2003**, *44*, 820–833. [[CrossRef](#)]
- Pradhan, B. Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches. *J. Indian Soc. Remote Sens.* **2010**, *38*, 301–320. [[CrossRef](#)]
- Yilmaz, I. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environ. Earth Sci.* **2010**, *61*, 821–836. [[CrossRef](#)]
- Oh, H.-J.; Pradhan, B. Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput. Geosci.* **2011**, *37*, 1264–1276. [[CrossRef](#)]

22. Sezer, E.A.; Pradhan, B.; Gokceoglu, C. Manifestation of an adaptive neuro-fuzzy model on landslide susceptibility mapping: Klang valley, Malaysia. *Expert Syst. Appl.* **2011**, *38*, 8208–8219. [[CrossRef](#)]
23. Akgun, A.; Kincal, C.; Pradhan, B. Application of remote sensing data and GIS for landslide risk assessment as an environmental threat to Izmir city (west Turkey). *Environ. Monit. Assess.* **2012**, *184*, 5453–5470. [[CrossRef](#)] [[PubMed](#)]
24. Pourghasemi, H.R.; Pradhan, B.; Gokceoglu, C. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat. Hazards* **2012**, *63*, 965–996. [[CrossRef](#)]
25. Althuwaynee, O.F.; Pradhan, B.; Park, H.-J.; Lee, J.H. A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. *Catena* **2014**, *114*, 21–36. [[CrossRef](#)]
26. Khosravi, K.; Nohani, E.; Maroufinia, E.; Pourghasemi, H.R. A GIS-based flood susceptibility assessment and its mapping in Iran: A comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. *Nat. Hazards* **2016**, *83*, 947–987. [[CrossRef](#)]
27. Khosravi, K.; Pourghasemi, H.R.; Chapi, K.; Bahri, M. Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: A comparison between Shannon's entropy, statistical index, and weighting factor models. *Environ. Monit. Assess.* **2016**, *188*, 656. [[CrossRef](#)] [[PubMed](#)]
28. Zhang, G.; Cai, Y.; Zheng, Z.; Zhen, J.; Liu, Y.; Huang, K. Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in Huizhou, China. *Catena* **2016**, *142*, 233–244. [[CrossRef](#)]
29. Pourghasemi, H.R.; Kerle, N. Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environ. Earth Sci.* **2016**, *75*, 185. [[CrossRef](#)]
30. Thai Pham, B.; Bui, D.T.; Prakash, I. Landslide susceptibility modelling using different advanced decision trees methods. *Civ. Eng. Environ. Syst.* **2018**, *35*, 139–157. [[CrossRef](#)]
31. Chung, C.-J.F.; Fabbri, A.G. Validation of spatial prediction models for landslide hazard mapping. *Nat. Hazards* **2003**, *30*, 451–472. [[CrossRef](#)]
32. 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.
33. Sadr, M.P.; Maghsoudi, A.; Saljoughi, B.S. Landslide susceptibility mapping of Komroud sub-basin using fuzzy logic approach. *Geodyn. Res. Int. Bull.* **2014**, *2*, XVI–XXVIII.
34. Sidle, R.C.; Ochiai, H. *Landslides: Processes, Prediction, and Land Use*; Water Resources Monograph Series; American Geophysical Union: Washington, DC, USA, 2006.
35. 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](#)]
36. Qingfeng, H.; Zhihao, X.; Shaojun, L.; Renwei, L.; Shuai, Z.; Nianqin, W.; Pham, B.T.; Wei, C. Novel Entropy and Rotation Forest-Based Credal Decision Tree Classifier for Landslide Susceptibility Modeling. *Entropy* **2019**, *21*(2), 106.
37. 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](#)] [[PubMed](#)]
38. Pham, B.T.; Prakash, I.; Khosravi, K.; Chapi, K.; Trinh, P.T.; Nog, T.Q.; Hossini, S.V.; Bui, D.T. A comparison of Support Vector Machines and Bayesian algorithms for landslide susceptibility modelling. *Geocarto Int.* **2018**. [[CrossRef](#)]
39. Shirzadi, A.; Soliamani, K.; Habibnejhad, M.; Kavian, A.; Chapi, K.; Shahabi, H.; Chen, W.; Khosravi, K.; Thai Pham, B.; Pradhan, B.; et al. Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping. *Sensors* **2018**, *18*, 3777. [[CrossRef](#)]
40. Thai Pham, B.; Prakash, I.; Dou, J.; Singh, S.K.; Trinh, P.T.; Tran, H.T.; Le, T.M.; Phong, T.V.; Khoi, D.K.; Shirzadi, A.; et al. A Novel Hybrid Approach of Landslide Susceptibility Modeling Using Rotation Forest Ensemble and Different Base Classifiers. *Geocarto Int.* **2018**, 1–38.
41. 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](#)]

42. Devkota, K.C.; Regmi, A.D.; Pourghasemi, H.R.; Yishida, K.; Pradhan, B.; Ryu, I.C.; Dhital, M.R.; Althuwaynee, O.F. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Nat. Hazards* **2013**, *65*, 135–165. [[CrossRef](#)]
43. Hong, H.; Naghibi, S.A.; Pourghasemi, H.R.; Pradhan, B. GIS-based landslide spatial modeling in Ganzhou City, China. *Arab. J. Geosci.* **2016**, *9*, 112. [[CrossRef](#)]
44. Nguyen, V.V.; Pham, B.T.; Vv, B.T.; Prakash, I.; Jha, S.; Shahabi, H.; Shirzadi, A.; Ba, D.N.; Kumar, R.; Chatterjee, J.M.; et al. Hybrid Machine Learning Approaches for Landslide Susceptibility Modeling. *Forests* **2019**, *10*, 157. [[CrossRef](#)]
45. Binh Thai, P.; Dieu, T.B.; Indra, P. Application of Classification and Regression Trees for Spatial Prediction of Rainfall Induced Shallow Landslides in the Uttarakhand Area (India) Using GIS. In *Climate Change, Extreme Events and Disaster Risk Reduction*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 159–170.
46. Tuan, T.; Dan, N. Landslide susceptibility mapping and zoning in the Son La hydropower catchment area using the analytical hierarchy process. *J. Sci. Earth* **2012**, *3*, 223–232.
47. Fayez, L.; Thai Pham, B.; Solanki, H.A.; Pazhman, D.; Dholakia, M.B.; Khalid, M.; Prakash, I. Application of Frequency Ratio Model for the Development of Landslide Susceptibility Mapping at Part of Uttarakhand State, India. *Int. J. Appl. Eng. Res.* **2018**, *13*, 6846–6854.
48. Pham, B.T. A Novel Classifier Based on Composite Hyper-cubes on Iterated Random Projections for Assessment of Landslide Susceptibility. *J. Geol. Soc. India* **2018**, *91*, 355–362. [[CrossRef](#)]
49. Pham, B.T.; Prakash, I.; Jaafari, A.; Bui, D.T. Spatial prediction of rainfall-induced landslides using aggregating one-dependence estimators classifier. *J. Indian Soc. Remote Sens.* **2018**, *46*, 1457–1470. [[CrossRef](#)]
50. Salvatici, T.; Tofani, V.; Rossi, G.; D’Ambrosio, M.; Tacconi Stefanelli, C.; Benedetta Masi, E.; Rosi, A.; Pazzi, V.; Vannocci, P.; Petrolo, M.; et al. Application of a physically based model to forecast shallow. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 1919–1935. [[CrossRef](#)]
51. Bonham-Carter, G.F. *Geographic Information Systems for Geoscientists: Modelling with GIS vol 13*; Elsevier: Amsterdam, The Netherlands, 2014.
52. Lee, S.; Talib, J.A. Probabilistic landslide susceptibility and factor effect analysis. *Environ. Geol.* **2005**, *47*, 982–990. [[CrossRef](#)]
53. Akgun, A.; Dag, S.; Bulut, F. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. *Environ. Geol.* **2008**, *54*, 1127–1143. [[CrossRef](#)]
54. Youssef, A.M.; Pradhan, B.; Pourghasemi, H.R.; Abdullahi, S. Landslide susceptibility assessment at Wadi Jawrah Basin, Jizan region, Saudi Arabia using two bivariate models in GIS. *Geosci. J.* **2015**, *19*, 449–469. [[CrossRef](#)]
55. Wan, S. A spatial decision support system for extracting the core factors and thresholds for landslide susceptibility map. *Eng. Geol.* **2009**, *108*, 237–251. [[CrossRef](#)]
56. Bednarik, M.; Magulová, B.; Matys, M.; Marschalko, M. Landslide susceptibility assessment of the Kraľovany–Liptovský Mikuláš railway case study. *Phys. Chem. Earth Parts A/B/C* **2010**, *35*, 162–171. [[CrossRef](#)]
57. Xu, C.; Xu, X.; Lee, Y.H.; Tan, X.; Yu, G.; Dai, F. The 2010 Yushu earthquake triggered landslide hazard mapping using GIS and weight of evidence modeling. *Environ. Earth Sci.* **2012**, *66*, 1603–1616. [[CrossRef](#)]
58. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol.* **2013**, *504*, 69–79. [[CrossRef](#)]
59. Dempster, A.P. Upper and Lower Probabilities Induced by a Multivalued Mapping. In *Classic Works of the Dempster-Shafer Theory of Belief Functions*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 57–72.
60. Park, N.-W. Application of Dempster-Shafer theory of evidence to GIS-based landslide susceptibility analysis. *Environ. Earth Sci.* **2011**, *62*, 367–376. [[CrossRef](#)]
61. Carranza, E.J.M.; Hale, M. Evidential belief functions for data-driven geologically constrained mapping of gold potential, Baguio district, Philippines. *Ore Geol. Rev.* **2003**, *22*, 117–132. [[CrossRef](#)]
62. Jaafari, A.; Zenner, E.K.; Pham, B.T. Wildfire spatial pattern analysis in the Zagros Mountains, Iran: A comparative study of decision tree based classifiers. *Ecol. Inform.* **2018**, *43*, 200–211. [[CrossRef](#)]

63. Tien Bui, D.; Shahabi, H.; Shiezadi, A.; Chapi, K.; Honang, N.-D.; Pham, B.T.; Bui, Q.-T.; Tran, C.-T.; Panahi, M.; Bin Ahmad, B.; et al. A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. *Remote Sens.* **2018**, *10*, 1538. [[CrossRef](#)]
64. Swets, J.A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285–1293. [[CrossRef](#)]
65. Pradhan, B. Manifestation of an advanced fuzzy logic model coupled with Geo-information techniques to landslide susceptibility mapping and their comparison with logistic regression modelling. *Environ. Ecol. Stat.* **2011**, *18*, 471–493. [[CrossRef](#)]
66. Pradhan, B.; Oh, H.; Buchroithner, M. Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. *Geomat. Nat. Hazards Risk* **2010**, *1*, 199–223. [[CrossRef](#)]
67. Bui, D.T.; Pradhan, B.; Lofman, O.; Revhaug, I.; Dick, O.B. Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): A comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. *Catena* **2012**, *96*, 28–40.
68. Nampak, H.; Pradhan, B.; Manap, M.A. Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *J. Hydrol.* **2014**, *513*, 283–300. [[CrossRef](#)]
69. Guzzetti, F. Landslide Hazard and Risk Assessment. Ph.D. Thesis, University of Bonn, Bonn, Germany, 2006.
70. Pham, B.T.; Khosravi, K.; Prakhsh, I. Application and comparison of decision tree-based machine learning methods in landslide susceptibility assessment at Pauri Garhwal Area, Uttarakhand, India. *Environ. Process.* **2017**, *4*, 711–730. [[CrossRef](#)]
71. Chen, W.; Zhao, X.; Shahabi, H.; Shirzadi, A.; Khosravi, K.; Chai, H.; Zhang, S.; Zhang, L.; Ma, J.; Chen, Y.; et al. Spatial prediction of landslide susceptibility by combining evidential belief function, logistic regression and logistic model tree. *Geocarto Int.* **2019**. [[CrossRef](#)]
72. Gholami, M.; Ghachkanlu, E.; Khosravi, K.; Pirasteh, S. Landslide prediction capability by comparison of frequency ratio, fuzzy gamma and landslide index method. *J. Earth Syst. Sci.* **2019**, *128*, 42. [[CrossRef](#)]
73. Tien Bui, D.; Shahabi, H.; Omidvar, E.; Shirzadi, A.; Geertsema, A.; Clague, J.; Khosravi, K.; Pradhan, B.; Pham, B.; Chapi, K.; et al. Shallow landslide prediction using a novel hybrid functional machine learning algorithm. *Remote Sens.* **2019**, *11*, 931. [[CrossRef](#)]
74. Chen, W.; Shahabi, H.; Zhang, S.; Khosravi, K.; Shirzadi, A.; Chapi, K.; Pham, B.T.; Zhang, T.; Zhang, L.; Chai, H.; et al. Landslide susceptibility modeling based on gis and novel bagging-based kernel logistic regression. *Appl. Sci.* **2018**, *8*, 2540. [[CrossRef](#)]
75. Hung, P.V.; Son, P.Q.; Dung, N.V. The study evaluated arming of risk of landslide in Hoa Binh and Son La reservoir hydropower area on the basis of analyzing high-resolution remote sensing and geographic information systems. *Vietnam J. Earth Sci.* **2015**, *37*, 193–203.
76. Tan, M.T.; Tao, N.V. Studying landslides in Thua Thien—Hue province. *Vietnam J. Earth Sci.* **2014**, *36*, 121–130. [[CrossRef](#)]
77. Thom, B.V.; Son, P.Q.; Hung, P.V.; Anh, N.T.V. Research assessment landslide and sedimentation of Hoa Binh hydropower reservoir. *Vietnam J. Earth Sci.* **2016**, *38*, 131–142.

