

Article

Comparison of Statistical Analysis Models for Susceptibility Assessment of Earthquake-Triggered Landslides: A Case Study from 2015 Earthquake in Lefkada Island

Christos Polykretis ¹, Kleomenis Kalogeropoulos ¹ , Panagiotis Andreopoulos ¹,
Antigoni Faka ¹, Andreas Tsatsaris ^{2,*} and Christos Chalkias ¹

¹ Department of Geography, Harokopio University, El. Venizelou 70, 17671 Athens, Greece

² School of Engineering, Department of Surveying & Geoinformatics Engineering, University of West Attica, Ag. Spyridonos 28, 12243 Athens, Greece

* Correspondence: atsats@uniwa.gr; Tel.: +30-2105385388

Received: 1 July 2019; Accepted: 7 August 2019; Published: 9 August 2019



Abstract: The main purpose of this study is to comparatively assess the susceptibility of earthquake-triggered landslides in the island of Lefkada (Ionian Islands, Greece) using two different statistical analysis models, a bivariate model represented by frequency ratio (FR), and a multivariate model represented by logistic regression (LR). For the implementation of the models, the relationship between geo-environmental factors contributing to landslides and documented events related to the 17th November 2015 earthquake was investigated by geographic information systems (GIS)-based analysis. A landslide inventory with events attributed to the specific earthquake was prepared using satellite imagery interpretation and field surveys. Eight factors: Elevation, slope angle, slope aspect, distance to main road network, distance to faults, land cover, geology, and peak ground acceleration (PGA), were considered and used as thematic data layers. The prediction capability of the models and the accuracy of the resulting susceptibility maps were tested by a standard validation method, the receiver operator characteristic (ROC) analysis. Based on the validation results, the output map with the highest reliability could potentially constitute an ideal basis for use within regional spatial planning as well as for the organization of emergency actions by local authorities.

Keywords: landslide susceptibility; 2015 earthquake; frequency ratio; logistic regression; GIS; Greece

1. Introduction

Globally, landslides are related to a great number of human losses and economic damages (Table 1). One of the major mechanisms by which a landslide can occur is triggering due to earthquakes [1]. Over the last two decades, strong earthquakes have triggered the occurrence of landslides in many islands of the Ionian Sea (2003 in Lefkada, 2014 in Kefhalonia, 2015 in Lefkada, 2018 in Zakynthos). This part of Greece can be characterized as a high-seismicity region, particularly prone and vulnerable to earthquake-triggered landslides.

In order to determine the most fragile landslide areas under the influence of a given earthquake in the future, it is important to indicate these areas through landslide susceptibility assessment and mapping [2]. A landslide susceptibility map gives an indication of “where” future landslides are likely to occur over a region on the basis of local geo-environmental conditions. Susceptibility mapping of earthquake-triggered landslides can be particularly useful for decision makers and planners. The knowledge about the spatial probability of their occurrence is very important for hazard management and mitigation, and safe planning [3].

Today, there are numerous landslide susceptibility assessment models, with most of them being assisted by the advances in Geoinformatics technologies such as geographic information systems (GIS) and remote sensing [2]. These models are divided into two main approaches: Qualitative and quantitative. In the qualitative approaches, used extensively during the 1970s and 1980s, the “expert” (engineering geologist or geomorphologist), based on his/her experience and knowledge of the terrain conditions, determines the degree of susceptibility by geomorphologic analyses, combinations of index maps, and logical analytical models [4–6]. In the quantitative approaches, widely applied in the last few decades, the numerical expressions of the relationship between landslide occurrence and influencing factors are estimated by geotechnical engineering, statistical analysis, and soft computing models [7–10].

Statistical analysis models can be either bivariate [11–13] or multivariate [14–16]. In bivariate models, each individual factor is combined with a landslide inventory map, and weight values based on landslide densities are calculated for its corresponding categories [17]. In multivariate models, the factors are sampled, and for each of the sampling units, the presence or absence of landslides is firstly determined. Following this, weight values for the factors are statistically estimated [18].

The main purpose of this study is to produce a reliable landslide susceptibility map that will present “where” landslides are likely to occur in the island of Lefkada in case of a potential earthquake with magnitude similar to the that witnessed on the 17th November 2015 ($M_w = 6.4$). The epicenter of the earthquake was located in the south-western part of the island between the villages Athani and Agios Petros and, with a focal depth of 10.7 km, was felt in a significant part of western Greece. The main consequences were the loss of two human lives, as well as the occurrence of extensive landslides, which in turn caused serious damage to properties and the road network. These landslide events affected mainly the western part of the island, namely: (a) The villages of Komilio, Dragano, and Athani, (b) the coastal zone from Eggremni to Gialos, and (c) the road axis from Tsoukalades to Agios Nikitas, with an overall length of 6 km [19].

Previous works referring to the aforementioned earthquake have focused on the mapping of triggered landslide events and the evaluation of their effects in the island of Lefkada. Specifically, Kazantzidou-Firtinidou et al. [20] carried out two field surveys in the island after the 17th November 2015 earthquake to observe and record the geotechnical damages represented by the triggered landslides, and the resulting structural damages. Papathanassiou et al. [19] provided detailed information regarding the spatial distribution and characteristics of the earthquake-induced landslides by field mapping. Furthermore, Grendas et al. [21] investigated the engineering geological conditions at areas influenced by landslides, and assessed the characteristics and physical quantities (type, area, volume, etc.) of these slope instability events. Based on the interpretation of aerial imagery and UAV-acquired photos, Valkaniotis et al. [22] identified the characteristics of a landslide triggered by the 17th November 2015 earthquake. A research effort focusing on the susceptibility assessment of earthquake-triggered landslides for the island of Lefkada is found in the research of Papathanassiou et al. [23] who, however, analyzed the previous similar magnitude ($M_w = 6.3$) earthquake in 2003.

In order to create the targeted susceptibility map, two different statistical analysis models, a bivariate like frequency ratio (FR) and a multivariate like logistic regression (LR), were applied. The implementation of these models has been widely preferred for landslide susceptibility assessment in various case studies [24–27]. Based on a landslide inventory, including landslides that were triggered by the 17th November 2015 earthquake, each of the models produced a susceptibility map. Between the two maps, the most reliable was found through a standard validation method, the receiver operator characteristic (ROC) analysis, in terms of accuracy and prediction ability of the models.

Table 1. Human and economic effects due to landslides represented on a global scale (Data provided by EM-DAT: The Emergency Events Database, Universite catholique de Louvain-CRED, D. Guha-Sapir, www.emdat.be, Brussels, Belgium).

Year	Occurrence	Total Deaths	Injured	Affected	Homeless	Total Affected	Total Damage (US\$)
1970–1979	45	7217	1041	94,019	3100	98,160	124,166
1980–1989	78	5647	4250	860,691	2,520,332	3,385,273	1,030,141
1990–1999	93	5104	977	590,027	1,170,860	1,761,864	1,695,190
2000–2009	147	6182	1276	1,119,597	193,539	1,314,412	288,107
2010–2019	126	6579	1304	2,543,562	84,378	2,629,244	1,633,414

2. Study Area

Lefkada is located in the Ionian Sea, between Corfu and Kefalonia, and a narrow strip of sea separates it from mainland Greece (Figure 1). With an extent of about 300 km², it is the fourth largest Ionian Island. Lefkada can be characterized as a mostly mountainous region (about 70% of the total area) with an average altitude of 500 m. Its climate includes mild and rainy winter, and hot summer. Because of these climatic conditions and its geomorphology, the island offers a variety of landscapes, a key element of which is the rich vegetation, narrow plateaus, white beaches, springs, and gorges.

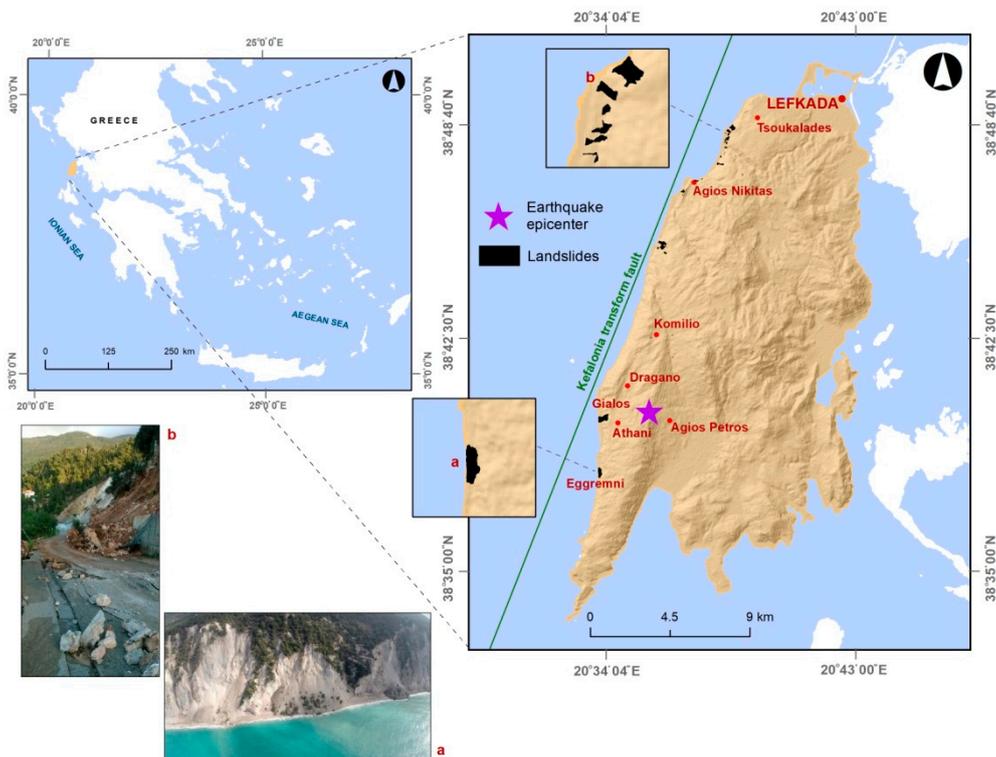


Figure 1. The landslide inventory map, including location map and images illustrating two landslide events triggered by the 17th November 2015 earthquake in the island of Lefkada: (a) Earth slide on the Egremni beach; (b) Rock slide on the road axis from Tsoukalades to Agios Nikitas.

Geologically, most of its terrain belongs to the western margin of the Ionian zone covered by thick limestone formations, while a small part of its south-western territory belongs to the northern extent of the pre-Apulian (Paxos) zone composed of limestones extending to a total thickness of 200 m. Even younger deposits have been detected on the north-east of the island, while the most recent deposits comprising alluvial sediments of loose sand and soft marine clays are found in the homonym capital town [28].

In regard to its tectonic aspect, a large number of faults are situated in Lefkada as a result of its geodynamic complexity. However, the main active tectonic structure accommodating the relative

motion of the African and Eurasian lithospheric plates, and thus highly influencing the seismicity of the region, is the Kefalonia transform fault (KTF) (Figure 1). The fault consists of two segments, the Kefalonia and Lefkada segments. Its length and width for the Kefalonia segment is estimated as 85 and 20 km, respectively, whereas for the Lefkada segment as 40 and 15 km, respectively. Due to the activity of KTF, Lefkada has a rich history of strong earthquakes with magnitudes up to 6.5 (1914, 1948, 1973, and 2003). An earthquake of similar magnitude occurred in the island on 17th November, 2015.

3. Data

Based on the above-described study area, the appropriate data were collected. Specifically, in order to carry out the landslide susceptibility assessment, a spatial database was designed and developed in GIS environment (by ArcGIS ver. 10.2 software package, ESRI, California, CA, USA). This database included two different types of datasets: (a) The landslide inventory dataset, and (b) the datasets with the geo-environmental factors.

3.1. Landslide Inventory

The landslide inventory contains critical information in order to apply quantitative modeling. Following the introduced landslide classification by [29], in the present inventory the term landslide is used to describe earth and rock slides, debris flows, and complex movements (Figure 1). Its creation was based on the use of cloud-free satellite imagery from Google Earth, and field surveys. A landslide database maintained by the Laboratory of Engineering Geology in the Department of Geology at the University of Patras [30] was previously exploited for the detection of landslide events triggered by the 17th November 2015 earthquake in the island of Lefkada. The final landslide inventory contained 43 landslide events. Given that in landslide inventories it is not always possible to differentiate their depletion and accumulation zones [15], these zones were mapped together in an entire area forming a single polygon feature for each landslide event (Figure 1).

3.2. Geo-Environmental Factors

A landslide constitutes the result of interaction between several geo-environmental factors. These factors can be subdivided into two categories: (a) The preconditioning factors that are expected to have an effect on the landslide occurrence, and (b) the triggering factors that trigger it. Due to the “nature” of the examined landslides as earthquake-triggered events, both preconditioning and triggering factors were used in this study. The selection of these factors mainly depends on the scale of the analysis and the availability of data [31]. Seven preconditioning factors representing the conditions in the study area before the occurrence of the earthquake, such as land cover, lithology, elevation, slope angle, slope aspect, distance to main road network, and distance to faults, and one (during the) earthquake-related triggering factor like peak ground acceleration (PGA) were taken into account.

Changes in land cover as a result of human activities like forest logging, deforestation, cultivation on steep slopes, and road construction can have a significant impact on landslide occurrence [32]. Land cover types were defined from the European CORINE (Coordinate of Information on the Environment) program [33] by following the level-3 classification scheme of the CORINE 2018 data. Given that the various lithological formations have different slope stability performance, lithology is also characterized as a significant factor for landslide susceptibility assessment [34]. The lithological units of the study area were derived from the geological map sheets (Leykas and Agios Petros) at 1:50,000 scale, provided by the Institute of Geology and Mineral Exploration (IGME). It is worth mentioning that by grouping the initial categories based on their common characteristics, the final land cover and lithology datasets were created.

The elevation factor is useful to classify the local relief and locate points of maximum and minimum heights within terrains [31]. Furthermore, an increase in slope angle is correlated with an increased likelihood of failure [35]. The slope aspect expressing the azimuthal orientation of a slope constitutes a decisive proxy for slope-influencing natural processes like incoming precipitation,

differential weathering, as well as wind and solar radiation, which may modulate slope stability. The elevation, slope angle, and slope aspect datasets were produced from the ASTER DEM (digital elevation model) product (with spatial resolution of 30 m) using GIS-based spatial analysis tools.

Road openings at the slope bases have negative impacts on slope stability. The dataset of distance to main road network was based on the digitization of the main roads from satellite images. Moreover, since the seismic ground acceleration and intensity diminish with the distance to tectonic elements, it was necessary for the relative factors (distance to faults and PGA) to be included in the present study. The dataset of distance to faults was derived from the digitization of the faults from the 1:50,000 geological map sheets. A PGA dataset was obtained from the earthquake strong-motion record of the Hellenic Strong Motion Network (HSMN) maintained by the Institute of Geodynamics-National Observatory of Athens (IG-NOA). At a national scale, HSMN consists of permanent stations in continuous and triggering mode across the entirety of Greece, an instrument-based observation array, and a network management system [36].

All the factors were organized in the relative GIS data layers and converted into raster grid format with pixel size 30 m. An overview of these factors is presented in Figure 2.

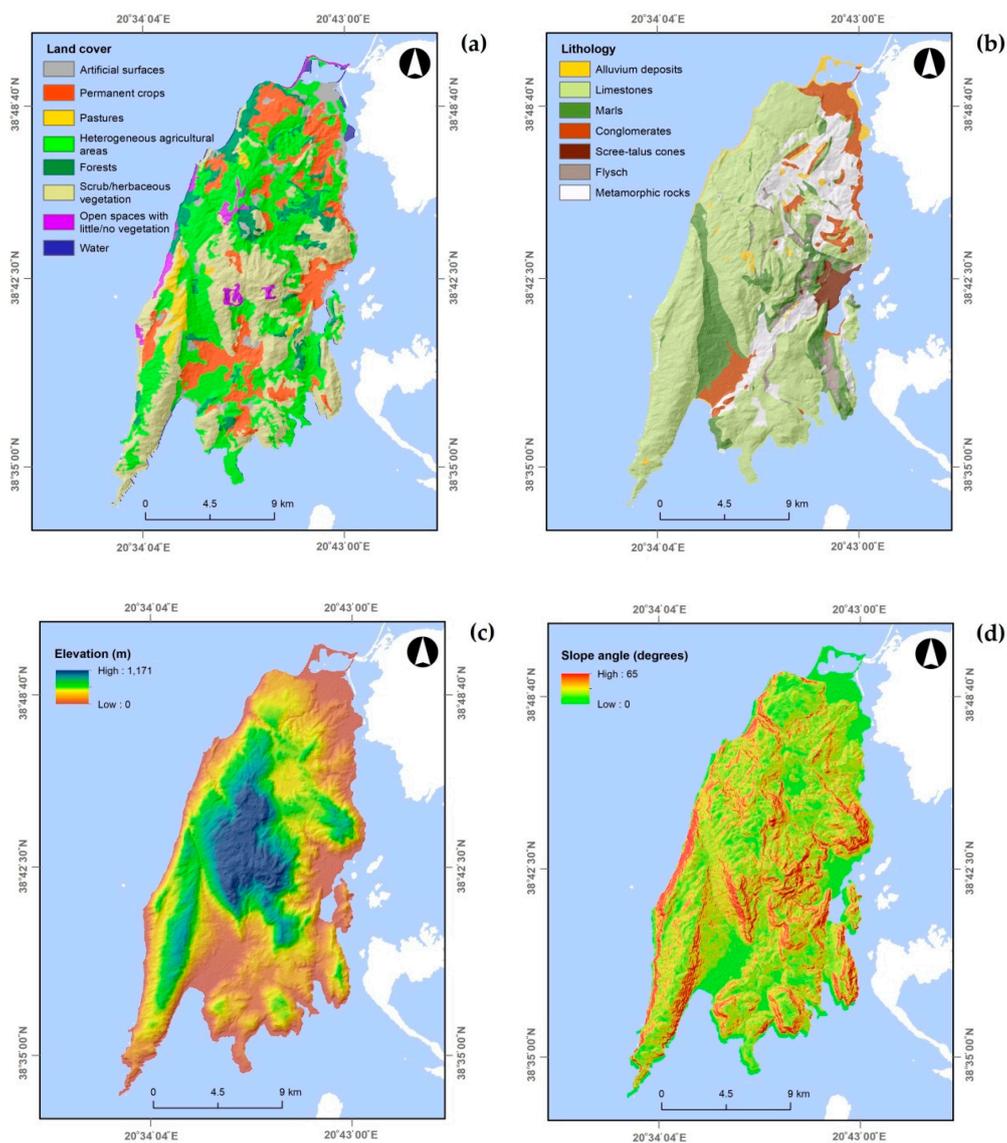


Figure 2. Cont.

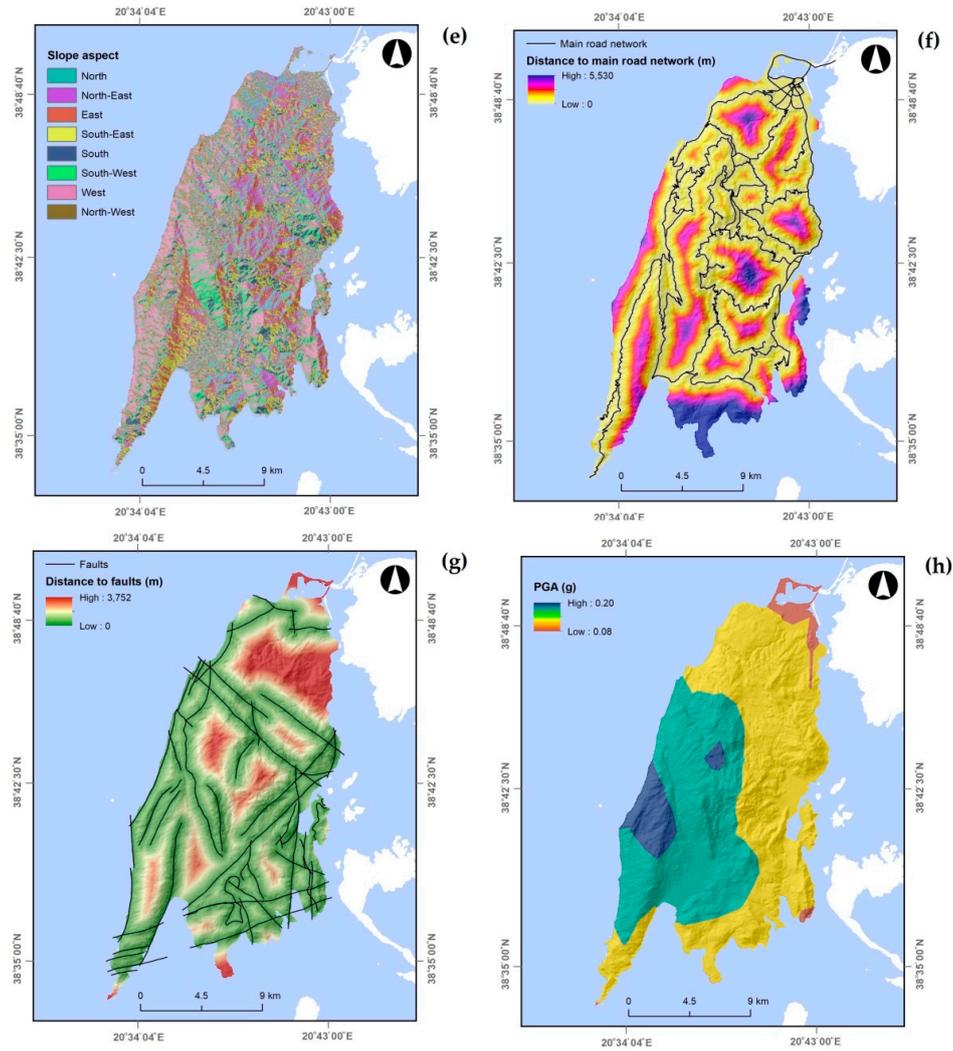


Figure 2. Geo-Environmental factors: (a) Land cover; (b) lithology; (c) elevation; (d) slope angle; (e) slope aspect; (f) distance to main road network; (g) distance to faults; (h) peak ground acceleration (PGA).

4. Methodology

4.1. Frequency Ratio (FR) Model

Frequency ratio is a simple and understandable bivariate statistical model based on the spatial associations between distribution of landslide events and each of the landslide-influencing factors, to expose the level of their correlation. It is calculated for each category of the factors by dividing the landslide occurrence ratio by the area ratio:

$$FR = \frac{\frac{N_{pix}(S_j)}{\sum N_{pix}(S_j)}}{\frac{N_{pix}(N_j)}{\sum N_{pix}(N_j)}} \quad (1)$$

where $N_{pix}(S_j)$ is the number of landslide pixels in factor category j , and $N_{pix}(N_j)$ is the number of pixels in the same factor category. A FR value of 1 (average value) means that the density of landslides in the category is proportional to the size of the category. If the value is greater than 1, then there is a high correlation, whereas a value of less than 1 means a lower correlation [37].

The overall landslide susceptibility (LS) value for each pixel can be obtained by summing the FR values of different landslide-influencing factors:

$$LS = \sum_{i=1}^n FR_{i,j} \quad (2)$$

where $FR_{i,j}$ is the frequency ratio value for the category j of the factor i , and n is the total number of the factors.

4.2. Logistic Regression (LR) Model

A logistic regression model forms a multivariate relation between a dependent variable and several independent variables [38]. The dependent variable is binary (i.e. it can take only the value 1 and 0), while its predictors, the independent variables, can be either continuous, discrete, or any combination of both types. In terms of a landslide susceptibility assessment, the goal of LR is to find the best fitting model to describe the relationship between the absence or presence (value of 0 or 1) of landslides (dependent variable) and a set of landslide-influencing factors (independent variables). The model can be expressed in its simplest form as:

$$P = \frac{1}{1 + e^{-z}} \quad (3)$$

where P is the probability of landslide occurrence, which ranges from 0 to 1, and z is a linear sum of a constant and the product of the independent variables with their respective coefficients. The value of z varies from $-\infty$ to $+\infty$ and is calculated from the equation:

$$z = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (4)$$

where n is the number of independent variables, x_i ($i = 1, 2, \dots, n$) are the independent variables, b_0 is the constant of the model, and b_i ($i = 1, 2, \dots, n$) are the coefficients.

Another way to write the LR model is the logit transformation using an equation in the form of:

$$\text{Logit}(p) = \ln\left(\frac{p(Y = 1)}{1 - p(Y = 1)}\right) = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (5)$$

where $p(Y = 1)/1 - p(Y = 1)$ is the so-called odds or likelihood ratio. The result represents the probability that a landslide event will occur divided by the probability that it fails to do so. If a coefficient is positive, its transformed log value will be greater than 1, meaning that the event is more likely to occur. If a coefficient is negative, the relative value will be less than 1 and the odds of the event occurring decreases. A coefficient of 0 has a transformed log value of 1, and it does not change the odds one way or the other [31].

Generally, the model estimates the coefficients and statistics based on the values of independent variables and the status of the dependent variable in a sampling dataset, using a maximum likelihood method [39]. Using the estimates derived from the implementation of the model on the selected sample, the probability of a landslide can be calculated on a pixel by pixel basis (Equation (3)).

4.3. Data Processing

The data processing included two different processes, such as the landslide data sampling and preparation of factor data. For the landslide data sampling, the landslide dataset derived from the inventory should be separated into two parts: (a) A training dataset, for the implementation of the models, and (b) a validation dataset, for the evaluation of their outputs. The total of 43 landslide polygons were converted into raster grid format with pixel size 30 m, resulting to 542 pixels. Among these pixels, 80% (434 pixels) was randomly selected as the training dataset, and the remaining 20%

(108 pixels) made up the validation dataset [2,3]. Furthermore, an equal number of pixels from the landslide-not-occurrence area were randomly selected for both the training and validation datasets. The FR model handled only a landslide dataset, whereas the LR model handled both the landslide and non-landslide datasets. Thus, the training dataset for FR model contained 434 pixels, and for the LR model 868 pixels. The validation dataset contained a total of 216 pixels. In both training and validation datasets, the target value of 1 was assigned to the landslide pixels, while the target value of 0 to the non-landslide pixels.

For the preparation of factor data used in FR model, the GIS-based “Natural Breaks” categorization was implemented for most of the factors with a continuous numerical scale (elevation, slope angle, distance to main road network, and distance to faults), except for PGA factor whose categorization was executed in a manually according to its presented values. In “Natural Breaks” categorization, category breaks identify best group similar values and maximize the differences between categories according to the deviations about the median [40]. For the preparation of factor data used in the LR model, the factors with continuous numerical scale, such as elevation, slope angle, distance to main road network, distance to faults, and PGA were handled in their original format, in order not to alter the state and information presented in their data layers. On the contrary, the factors with discrete categorical scale, such as land cover, lithology, and slope aspect were re-scaled in the range of 0.1 to 0.9 by coding and ranking their categories based on the relative landslide densities.

4.4. Implementation of Models

For the FR model, by crossing the landslide training dataset (434 pixels) with each factor layer, the relative landslide density in each factor category was calculated. The FR value for each of these categories was then estimated with Equation (1) (Table 2). Finally, all factor layers were overlaid by summation (Equation (2)) and a resultant the landslide susceptibility map was obtained. This map was categorized into five categories (“Very Low,” “Low,” “Moderate,” “High,” and “Very High” susceptibility) through the “Natural Breaks” method (Figure 3).

Table 2. Values derived from the frequency ratio (FR) model for all factor categories, and multicollinearity checking indexes (tolerance (TOL) and variance inflation factor (VIF)) and coefficients derived from logistic regression (LR) model for all factors.

Factors and Their Categories	FR Model			LR Model		
	Number of Total Pixels	Number of Landslide Pixels	FR Value	TOL	VIF	Coefficients
Elevation (m)				0.615	1.627	−0.015
(1) 0–134	96,907	315	2.78			
(2) 135–306	86,741	111	1.09			
(3) 307–501	95,228	4	0.04			
(4) 502–754	58,042	4	0.06			
(5) 755–1171	33,779	0	0			
Slope Angle (Degrees)				0.898	1.113	0.052
(1) 0–8	97,552	56	0.49			
(2) 9–16	100,441	102	0.87			
(3) 17–25	85,991	95	0.94			
(4) 26–35	58,647	93	1.35			
(5) 35–65	28,066	88	2.68			

Table 2. Cont.

Factors and Their Categories	FR Model			LR Model		
	Number of Total Pixels	Number of Landslide Pixels	FR Value	TOL	VIF	Coefficients
Slope Aspect				0.670	1.493	4.709
(1) North	34,034	9	0.22			
(2) North-East	40,218	6	0.12			
(3) East	56,505	0	0			
(4) South-East	49,155	3	0.05			
(5) South	34,312	4	0.10			
(6) South-West	41,160	19	0.39			
(7) West	60,759	243	3.35			
(8) North-West	46,861	150	2.68			
Distance to Main Road network (m)				0.850	1.177	−0.004
(1) 0–347	179,992	166	0.79			
(2) 348–846	111,966	218	1.66			
(3) 847–1,648	63,214	50	0.68			
(4) 1,649–3,231	12,889	0	0			
(5) 3,232–5,530	2611	0	0			
Distance to Faults (m)				0.627	1.595	−0.002
(1) 0–324	162,088	425	2.24			
(2) 325–765	101,455	6	0.05			
(3) 766–1,354	62,190	1	0.01			
(4) 1,355–2,163	29,937	2	0.06			
(5) 2,164–3,752	15,002	0	0			
PGA (g)				0.704	1.420	60.946
(1) 0,08	10,895	0	0			
(2) 0.09–0.12	238,830	35	0.13			
(3) 0.13–0.16	98,023	398	3.47			
(4) 0.17–0.20	22,924	1	0.04			
Land Cover				0.816	1.225	2.948
(1) Artificial Surfaces	14,182	9	0.54			
(2) Permanent Crops	60,462	7	0.10			
(3) Pastures	8997	0	0			
(4) Heterogeneous Agricultural Areas	114,705	45	0.34			
(5) Forests	48,515	83	1.46			
(6) Scrub/Herbaceous Vegetation	107,438	241	1.92			
(7) Open Spaces with Little/No Vegetation	12,907	49	3.24			
(8) Water	3466	0	0			
Lithology				0.822	1.216	2.274
(1) Alluvium Deposits	12,059	41	2.89			
(2) Limestones	209,228	381	1.55			
(3) Marls	47,455	6	0.11			
(4) Conglomerates	25,303	0	0			
(5) Scree-Talus Cones	6466	0	0			
(6) Flysch	12,736	1	0.07			
(7) Metamorphic Rocks	55,804	5	0.08			

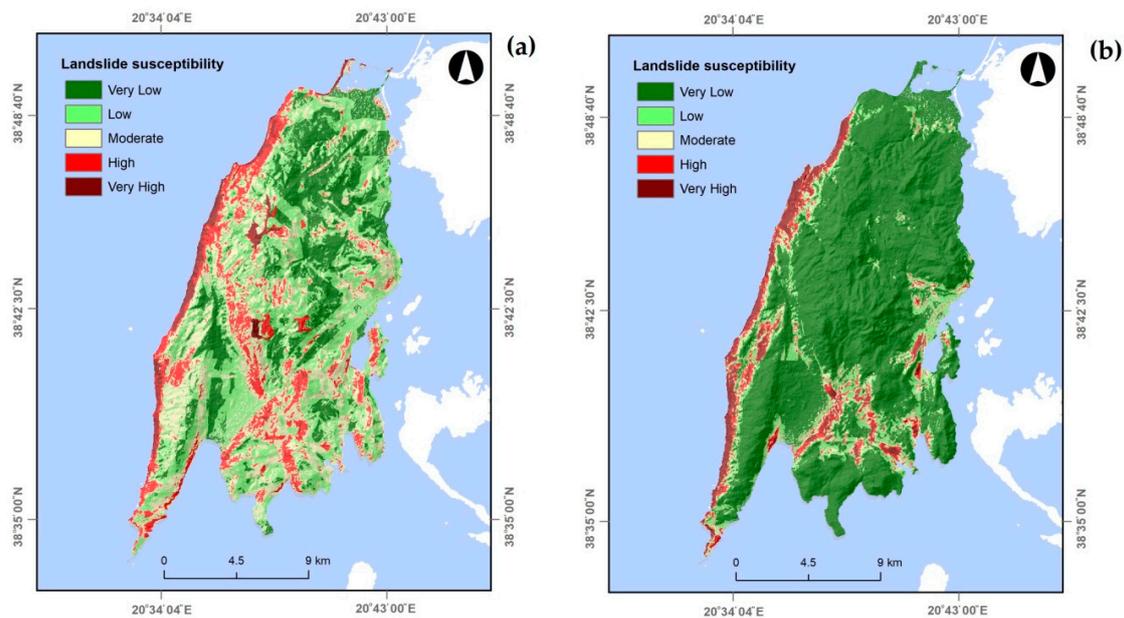


Figure 3. Landslide susceptibility map produced by: (a) FR model; (b) LR model.

In case of the LR model, the training dataset (868 landslide and non-landslide pixels) were matched with each factor layer in order to create a database. This database with the eight factors as independent variables, and the presence and absence of landslide (binary value of 0 and 1) as the dependent variable was imported into the SPSS (ver. 23, Company, City, Country IBM, New York, N.Y., USA) software package. The tolerance (TOL) and variance inflation factor (VIF) indexes were then estimated to check the multicollinearity of independent variables. TOL estimation is based on the determination of a regression of each explanatory variable on all the other explanatory variables. VIF is calculated by $1/\text{TOL}$ [41]. These indexes were found to be greater than 0.2 for TOL and less than 10 for VIF revealing that there is no multicollinearity between any of the factors (Table 2). Following the multicollinearity checking, the binary LR algorithm was carried out to calculate the correlation of landslides to each factor. This correlation is expressed by the coefficients (Table 2). After the assignment of coefficients to all the factors, a GIS-based weighted overlay was applied using Equation (4). By inserting the output of overlay into Equation (3), a resultant landslide susceptibility map was produced and finally categorized into five categories (Figure 3).

5. Results

The results of the FR and LR models are summarized in Table 2. Based on the FR model, the factor categories that presented the highest correlation with the occurrence of landslides triggered by the 17th November 2015 earthquake were “0.12 to 0.16 g” (FR = 3.47) for PGA, “west” facing (FR = 3.35) for slope aspect, “open spaces with little/no vegetation” (FR = 3.24) for land cover, “alluvium deposits” (FR = 2.89) for lithology, “0 to 134 m” (FR = 2.78) for elevation, “35 to 65 degrees” (FR = 2.68) for slope angle, and “0 to 324 m” (FR = 2.24) for distance to faults. There is no mention for the factor of distance to main road network because its highest value (FR = 1.66) was very close to 1 (average value). According to the LR model, the factors of PGA, slope aspect, land cover, lithology, and slope angle had positive coefficients, whereas the factors of elevation, distance to main road network, and distance to faults had negative coefficients. Among the factors with positive coefficients, the highest value was presented by PGA (coefficient of 60.946) followed by slope aspect (coefficient of 4.709).

The landslide susceptibility maps produced by the two models are presented in Figure 3. In the map from the FR model, the “High” and “Very High” susceptibility categories were mainly located in the western coastal zone of Lefkada Island, including large pockets of “High” susceptibility in its central and southern parts. These two categories cover 13% and 3%, respectively, of the study area

(Figure 4). In the map from the LR model, the “Very High” susceptibility category was detected in the western coastal zone of the island, with some limited pockets of “High” and “Very High” susceptibility in its southern part. The coverage percentages for the two susceptibility categories were found to be 4% and 5%, respectively (Figure 4). Moreover, the cross-comparison of the two output maps in terms of coverage differences between their categories (Table 3) indicated that the highest difference percentages (29% and 19%, respectively) were shown between the “Low” and “Moderate” categories of the FR model, and the “Very Low” category of LR model. It is also worth mentioning that no coverage relation was observed between the “High” and “Very High” categories of the LR model, and the “Very Low” and “Low” categories of the FR model. On the contrary, 5% and 1% of the “High” and “Very High”, categories of the FR model were characterized with “Very Low” susceptibility in the map from the LR model. Concerning the coverage similarities, it is indicated that 2% of the study area was characterized by similar susceptibility for each of two, “High” and “Very High”, categories. The corresponding percentages for the other three, “Moderate”, “Low”, and “Very Low”, categories were 3%, 4% and 21%, respectively. The distribution of these similarities are presented in Figure 5.

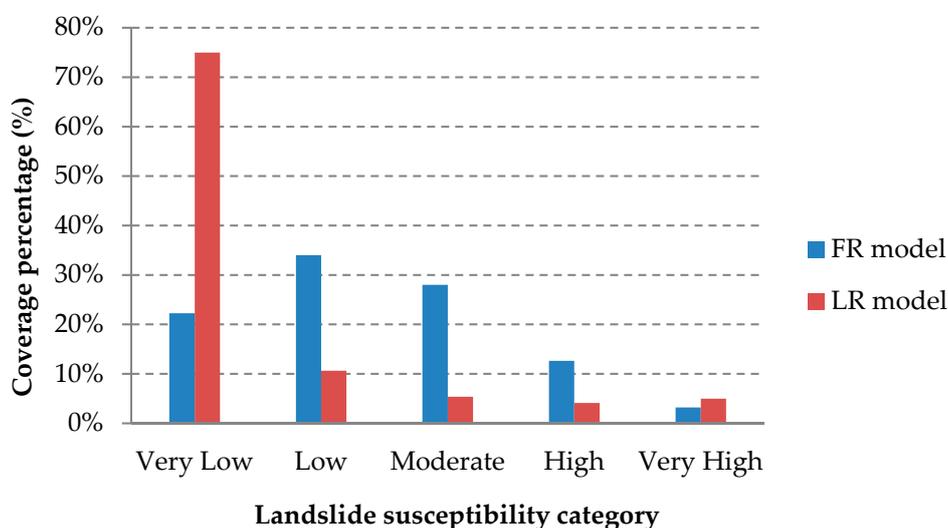


Figure 4. Coverage percentages of landslide susceptibility categories.

Table 3. Coverage cross-comparison for the landslide susceptibility categories between both the FR and LR models.

FR	LR				
	VL (%)	L (%)	M (%)	H (%)	VH (%)
VL	21	1	–	–	–
L	29	4	1	–	–
M	19	4	3	2	1
H	5	2	1	2	2
VH	1	–	–	–	2

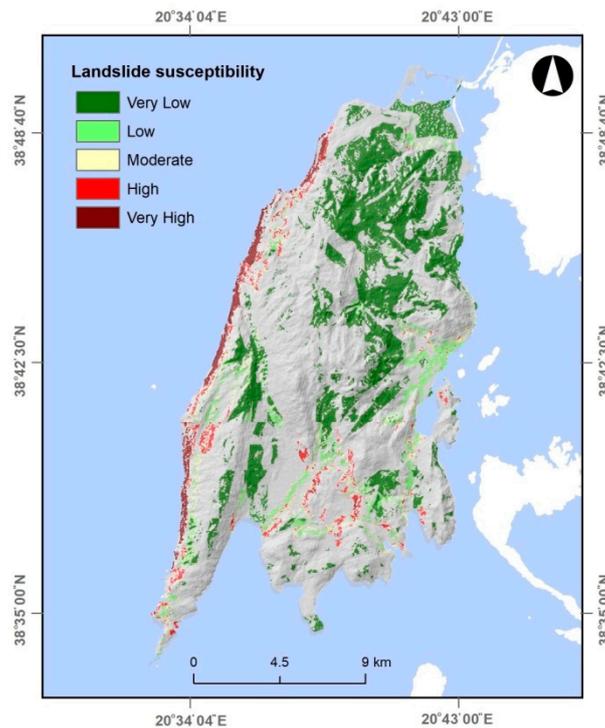


Figure 5. Map with coverage similarities between the landslide susceptibility categories of the two models.

Validation of Results

Validation is an essential process to obtain knowledge about the accuracy and prediction ability of the models [42]. The overlay of produced susceptibility maps with a landslide dataset is considered as a standard validation method. In the present study, the results of this overlay with the landslide validation dataset (Figure 6) showed that, for the FR model, 69% and 28% of the landslide pixels were found within the “Very High” and “High”, susceptibility categories and none of these pixels were found within the “Very Low” and “Low” susceptibility categories. Regarding the LR model, the percentages were 87% and 8% for the “Very High” and “High”, susceptibility categories, and only 1% for the “Very Low” and “Low” susceptibility categories.

A more advanced validation method, which has been widely applied, is the receiver operating characteristics (ROC) analysis [8,15,43,44]. Based on the overlay of landslide susceptibility maps with an “independent” validation dataset (with landslide and non-landslide data), a ROC graph was created. In this graph, the sensitivity of a model, which is determined as the percentage of the correctly predicted landslide pixels by the model, was plotted against specificity, which is the percentage of predicted landslide pixels over the total study area [45]. The value of area under the ROC curve (AUC) indicates the prediction ability of the model. With a range from 0.5 to 1.0, the higher this value is, the better is the prediction ability of the model. More details about ROC analysis can be found in [46]. The results of ROC analysis are presented in Figure 7 and Table 4. The AUC value of the LR model was equal to 0.984, followed by the FR model with a value of 0.976. The accuracy percentages were equal to 93.1% and 78.2%, respectively.

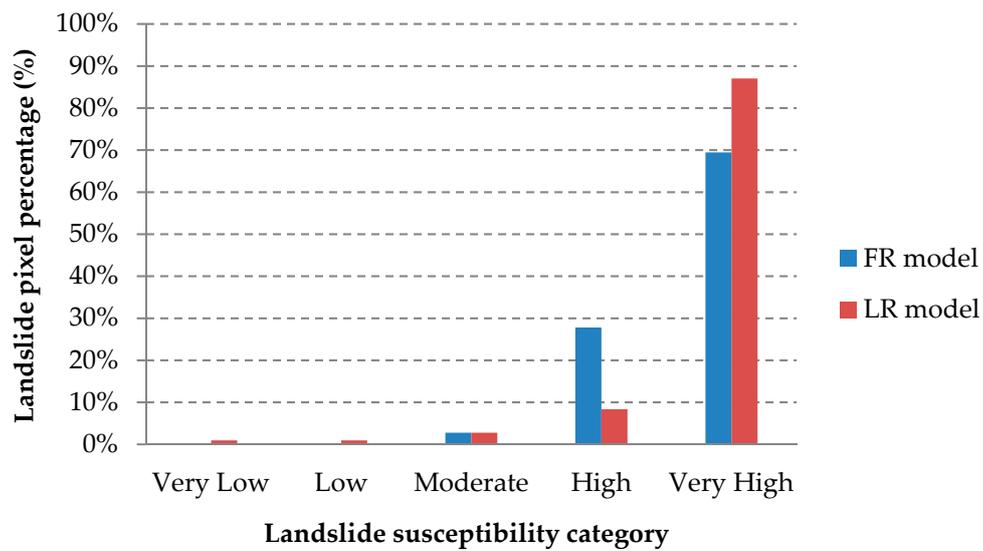


Figure 6. Landslide pixel percentages of landslide susceptibility categories.

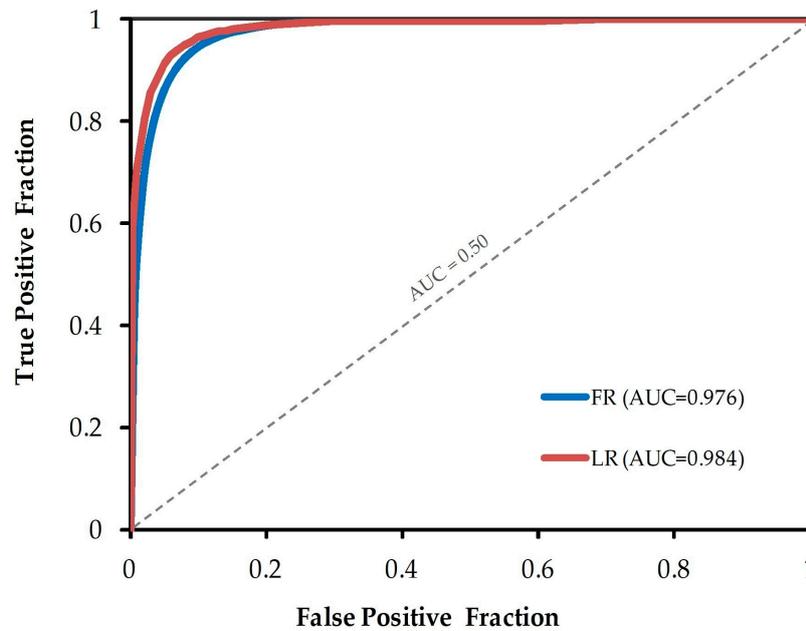


Figure 7. Receiver operating characteristics (ROC) curves of the models.

Table 4. ROC analysis results of the models.

ROC Analysis Results	FR	LR
Number of Cases	216	216
Number Correct	169	201
Positive Cases Missed	0	2
Negative Cases Missed	47	13
Accuracy (%)	78.2	93.1
Sensitivity (%)	100	98.1
Specificity (%)	56.5	88

6. Discussion

For the seismically active region of Ionian Islands, and specifically the island of Lefkada, the occurrence of strong earthquakes acting as a triggering mechanism for landslides, is a frequent

phenomenon. Therefore, the acquisition of knowledge about the spatial definition of potential earthquake-triggered landslides is considered as an essential need for the region. In terms of acquisition of this knowledge, the present study focused on the creation of a reliable landslide susceptibility map that will spatially define the landslide occurrence for the island of Lefkada. This was designed in case an earthquake with a magnitude similar to the one witnessed on the 17th November 2015 occurs in the future.

In order to obtain the desired map, two different statistical models, such as bivariate FR and multivariate LR were applied based on the integration of eight geo-environmental factors and a landslide inventory including landslides triggered by the specific earthquake. As a bivariate model, FR uses the observed densities of landslide pixels within factor datasets to estimate susceptibility in a “sub-factor”, category level. Concerning the category results of the FR model in this study, the areas of island with low elevation, very high slope angle and western facing, being very close to faults, covered by little or no vegetation and alluvium deposits, and presenting high PGA during the earthquake, were highly correlated with the occurrence of the resultant landslides.

Multivariate LR analyzes variables that are non-symmetrical and show skewed distributions to create a mathematical model that predicts the probability of occurrence in a sampling unit. The coefficients assigned to influencing factors from the LR model are useful to assess their importance on the presence or absence of landslides. In the present study, the factors of PGA, slope aspect, land cover, lithology, and slope angle were found to have positive coefficients indicating that these factors are positively related with the occurrence of earthquake-triggered landslides. However, among them, the factor with the significantly strongest effect on their occurrence was the PGA. It could be said that this finding confirms the “nature” of landslides under investigation as earthquake-triggered events.

The output maps of the two models visualized the spatial distribution of the estimated landslide susceptibility in the study area. Both maps presented the western coastal zone as the part of the island with the highest probability to produce landslides under the influence of a potential earthquake with a magnitude similar to earthquake examined in this paper. According to the study [23], the same part of the island was proved to be riskiest area for earthquake-triggered landslides after the 2003 earthquake.

In regard to reliable results, the susceptibility maps produced by the two models were compared by means of validation using a dataset not originally included in the training process. From the method of overlay, it was derived that the “Very High” and “High” susceptibility categories, contained a total of about 95% of landslide data included in the validation dataset. Although both models present a very good prediction ability (based on their AUC values in Figure 7), the better one is found in the LR model. The same model also shows superiority against the FR model in terms of accuracy, as the last one tended to present a susceptibility overestimation, which is expressed by high sensitivity and simultaneously low specificity values (Table 4). Both models can successfully detect the susceptible areas (“High” and “Very High” categories), but only the LR model can, at the same time, successfully detect the non-susceptible areas (“Very Low” and “Low” categories). Thus, LR can be assumed as more balanced model against FR. In general, among statistical analysis models, multivariate LR has been proved to give better validation results compared to bivariate models for landslide susceptibility assessments [47–49]. Regarding specifically the susceptibility assessment of earthquake-triggered landslides, in the study [2], LR model presented very good validation results from two susceptibility maps produced for a region of China using multi-temporal and earthquake triggered landslide datasets. Moreover, in the study [3], the susceptibility of earthquake-triggered landslides was comparatively evaluated for a watershed of China based on the implementation of six different quantitative models. One of these models was the LR, which provided the best validation results when compared to the other models.

It is worth mentioning that in this study important analysis parameters, such as the number and type of geo-environmental factors, their “pre-model” processing, the size of training and validation datasets, and the categorization of final susceptibility maps, were based on objective

criteria. Considering the data-driven “nature” of the two models, the examination of alternative choices for these parameters could lead to different results.

7. Conclusions and Outlook

Due to its confirmed higher reliability, the landslide susceptibility map produced by the LR model could constitute an important basis for the assessment and management of landslide hazard over the island of Lefkada. This map provides useful information for decision makers and planners to choose suitable locations to implement reconstructions and developments. Furthermore, it could be used for the estimation of the amount of damages and losses that landslide hazards may potentially cause, and the consequent planning of mitigation actions that will reduce the risk of structural damage and loss of life on the island.

Future research work may include testing the LR model in other areas affected by earthquake-triggered landslide events, and its comparison with other advanced soft computing models, such as artificial neural networks (ANNs), support vector machines (SVMs), as well as ensemble random forest and neuro-fuzzy models.

Author Contributions: All the authors contributed substantially to this work. conceptualization, C.P., K.K., P.A. and C.C.; methodology, C.P. and P.A.; validation, C.P. and A.F.; visualization, C.P. and K.K.; writing—original draft preparation, C.P., K.K., P.A., A.F., A.T. and C.C.; writing—review and editing, C.P., A.F., A.T. and C.C.

Acknowledgments: The authors wish to thank the colleagues who politely considered revising this article.

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

References

- Barlow, J.; Barisin, I.; Rosser, N.; Petley, D.; Densmore, A.; Wrigh, T. Seismically-induced mass movements and volumetric fluxes resulting from the 2010 $M_w = 7.2$ earthquake in the Sierra Cucapah, Mexico. *Geomorphology* **2015**, *230*, 138–145. [[CrossRef](#)]
- Bai, S.; Wang, J.; Zhang, Z.; Cheng, C. Combined landslide susceptibility mapping after Wenchuan earthquake at the Zhouqu segment in the Bailongjiang Basin, China. *Catena* **2012**, *99*, 18–25. [[CrossRef](#)]
- Xu, C.; Xu, X.; Dai, F.; Saraf, A.K. Comparison of different models for susceptibility mapping of earthquake triggered landslides related with the 2008 Wenchuan earthquake in China. *Comput. Geosci.* **2012**, *46*, 317–329. [[CrossRef](#)]
- Listo, F.D.L.R.; Carvalho Vieira, B. Mapping of risk and susceptibility of shallow-landslide in the city of Sao Paulo, Brazil. *Geomorphology* **2012**, *169–170*, 30–44. [[CrossRef](#)]
- Avtar, R.; Singh, C.K.; Singh, G.; Verma, R.L.; Mukherjee, S.; Sawada, H. Landslide susceptibility zonation study using remote sensing and GIS technology in the Ken-Betwa River Link area, India. *Bull. Eng. Geol. Environ.* **2011**, *70*, 595–606. [[CrossRef](#)]
- Mandal, B.; Mandal, S. Analytical hierarchy process (AHP) based landslide susceptibility mapping of Lish river basin of eastern Darjeeling Himalaya, India. *Adv. Space Res.* **2018**, *62*, 3114–3132. [[CrossRef](#)]
- Vieira, B.C.; Fernandes, N.F.; Filho, O.A.; Martins, T.D.; Montgomery, D.R. Assessing shallow landslide hazards using the TRIGRS and SHALSTAB models, Serra do Mar, Brazil. *Environ. Earth Sci.* **2018**, *77*, 260. [[CrossRef](#)]
- Chalkias, C.; Ferentinou, M.; Polykretis, C. GIS-Based Landslide Susceptibility Mapping on the Peloponnese Peninsula, Greece. *Geosciences* **2014**, *4*, 176–190. [[CrossRef](#)]
- Hong, H.; Liu, J.; Tien Bui, D.; 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](#)]
- 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](#)]
- Youssef, A.M.; Al-Kathery, M.; Pradhan, B. Landslide susceptibility mapping at Al-Hasher Area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models. *Geosci. J.* **2015**, *19*, 113–134. [[CrossRef](#)]

12. Zhao, C.; Chen, W.; Wang, Q.; Wu, Y.; Yang, B. A comparative study of statistical index and certainty factor models in landslide susceptibility mapping: A case study for the Shangzhou District, Shaanxi Province, China. *Arab. J. Geosci.* **2015**, *8*, 9079–9088. [[CrossRef](#)]
13. Piacentini, D.; Devoto, S.; Mantovani, M.; Pasuto, A.; Prampolini, M.; Soldati, M. Landslide susceptibility modeling assisted by Persistent Scatterers Interferometry (PSI): An Example from the Northwestern Coast of Malta. *Nat. Hazards* **2015**, *78*, 681–697. [[CrossRef](#)]
14. Basu, T.; Pal, S. Identification of landslide susceptibility zones in Gish River basin, West Bengal, India. *Georisk* **2018**, *12*, 14–28. [[CrossRef](#)]
15. Dagdelenler, G.; Nefeslioglu, H.A.; Gokceoglu, C. Modification of seed cell sampling strategy for landslide susceptibility mapping: An application from the Eastern part of the Gallipoli Peninsula (Canakkale, Turkey). *Bull. Eng. Geol. Environ.* **2016**, *75*, 575–590. [[CrossRef](#)]
16. Reichenbach, P.; Busca, C.; Mondini, A.C.; Rossi, M. The Influence of Land Use Change on Landslide Susceptibility Zonation: The Briga Catchment Test Site (Messina, Italy). *Environ. Manag.* **2014**, *54*, 1372–1384. [[CrossRef](#)] [[PubMed](#)]
17. Soeters, R.; van Westen, C.J. Slope Instability Recognition, Analysis and Zonation. In *Landslides, Investigation and Mitigation*; Turner, A.K., Schuster, R.L., Eds.; National Academy Press: Washington, DC, USA, 1996; pp. 129–177.
18. Baeza, C.; Corominas, J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf. Process. Landf.* **2001**, *26*, 1251–1263. [[CrossRef](#)]
19. Papathanassiou, G.; Vakaniotis, S.; Ganas, A.; Grendas, N.; Kollia, E. The November 17th, 2015 Lefkada (Greece) strike-slip earthquake: Field mapping of generated failures and assessment of macroseismic intensity ESI-07. *Eng. Geol.* **2016**, *220*, 13–30. [[CrossRef](#)]
20. Kazantzidou-Firtinidou, D.; Kassaras, I.; Tonna, S.; Ganas, A.; Vintzileou, E.; Chesi, C. The November 2015 M_w 6.4 earthquake effects in Lefkas Island. In Proceedings of the 1st International Conference on Natural Hazards and Infrastructure, Chania, Greece, 28–30 June 2016; pp. 1–10.
21. Grendas, N.; Marinos, V.; Papathanassiou, G.; Ganas, A.; Valkaniotis, S. Engineering geological mapping of earthquake-induced landslides in South Lefkada Island, Greece: Evaluation of the type and characteristics of the slope failures. *Environ. Earth Sci.* **2018**, *77*, 425. [[CrossRef](#)]
22. Valkaniotis, S.; Papathanassiou, G.; Ganas, A. Mapping an earthquake-induced landslide based on UAV imagery; case study of the 2015 Okeanos landslide, Lefkada, Greece. *Eng. Geol.* **2018**, *245*, 141–152. [[CrossRef](#)]
23. Papathanassiou, G.; Valkaniotis, S.; Ganas, A.; Pavlides, S. GIS-based statistical analysis of the spatial distribution of earthquake-induced landslides in the island of Lefkada, Ionian Islands, Greece. *Landslides* **2013**, *10*, 771–783. [[CrossRef](#)]
24. Wang, L.-J.; Guo, M.; Sawada, K.; Lin, J.; Zhang, J. A comparative study of landslide susceptibility maps using logistic regression, frequency ratio, decision tree, weights of evidence and artificial neural network. *Geosci. J.* **2016**, *20*, 117–136. [[CrossRef](#)]
25. Aditian, A.; Kubota, T.; Shinohara, Y. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology* **2018**, *318*, 101–111. [[CrossRef](#)]
26. Demir, G. Landslide susceptibility mapping by using statistical analysis in the North Anatolian Fault Zone (NAFZ) on the northern part of Susehri Town, Turkey. *Nat. Hazards* **2018**, *92*, 133–154. [[CrossRef](#)]
27. Mondal, S.; Mandal, S. Landslide susceptibility and risk: A micro level study from the Balason River basin in Darjeeling Himalaya. *Arab. J. Geosci.* **2018**, *11*, 197. [[CrossRef](#)]
28. Karababa, F.S.; Pomonis, A. Damage data analysis and vulnerability estimation following the August 14, 2003 Lefkada Island, Greece, Earthquake. *Bull. Earthq. Eng.* **2011**, *9*, 1015–1046. [[CrossRef](#)]
29. Cruden, D.M.; Varnes, D.J. Landslide types and processes. In *Landslides: Investigation and Mitigation*; Turner, A.K., Schuster, R.L., Eds.; National Academy Press: Washington, DC, USA, 1996; pp. 36–75.
30. Landslide Inventory Database, Laboratory of Engineering Geology, University of Patras. Available online: <http://www.geoarch.gr/> (accessed on 5 September 2017).
31. 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](#)]
32. Glade, T. Landslide occurrence as a response to land use change: A review of evidence from New Zealand. *Catena* **2003**, *51*, 297–314. [[CrossRef](#)]

33. Heymann, Y.; Steenmans, C.; Croisille, G.; Bossard, M. *CORINE Land Cover Project: Technical Guide*; European Commission, Directorate General Environment, Nuclear Safety and Civil Protection, ECSC-EEC-EAEC: Brussels, Belgium, 1994.
34. Carrara, A.; Cardinali, M.; Detti, R.; Guzzetti, F.; Pasqui, V.; Reichenbach, P. GIS techniques and statistical models in evaluating landslide hazard. *Earth Surf. Process. Landf.* **1991**, *16*, 427–445. [[CrossRef](#)]
35. Dai, F.C.; Lee, C.F. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology* **2002**, *42*, 213–228. [[CrossRef](#)]
36. Margaris, B.; Kalogeras, I.; Papaioannou, C.; Savvaidis, A.; Theodoulidis, N. Evaluation of the national strong motion network in Greece: Deployment, data-processing and site characterization. *Bull. Earthq. Eng.* **2014**, *12*, 237–254. [[CrossRef](#)]
37. Choi, J.; Oh, H.-J.; Lee, H.-J.; Lee, C.; Lee, S. Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. *Eng. Geol.* **2012**, *124*, 12–23. [[CrossRef](#)]
38. Atkinson, P.M.; Massari, R. Generalized linear modeling of susceptibility to landsliding in the central Apennines, Italy. *Comput. Geosci* **1998**, *24*, 373–385. [[CrossRef](#)]
39. Kundu, S.; Saha, A.K.; Sharma, D.C.; Pant, C.C. Remote Sensing and GIS Based Landslide Susceptibility Assessment using Binary Logistic Regression Model: A Case Study in the Ganeshganga Watershed, Himalayas. *J. Indian Soc. Remote Sens.* **2013**, *41*, 697–709. [[CrossRef](#)]
40. Jenks, G.F. *Optimal Data Classification for Choropleth Maps*; University of Kansas: Lawrence, KS, USA, 1977.
41. Pourghasemi, H.R.; Gayen, A.; Park, S.; Lee, C.-W.; Lee, S. Assessment of Landslide-Prone Areas and Their Zonation Using Logistic Regression, LogitBoost, and NaïveBayes Machine-Learning Algorithms. *Sustainability* **2018**, *10*, 3697. [[CrossRef](#)]
42. Remondo, J.; González-Díez, A.; Díaz de Terán, J.R.; Cendrero, A. Landslide susceptibility models utilizing spatial data analysis techniques: A case study from the Lower Deba Valley, Guipúzcoa (Spain). *Nat. Hazards* **2003**, *30*, 267–279. [[CrossRef](#)]
43. Pradhan, A.M.S.; Kim, Y.-T. GIS-based landslide susceptibility model considering effective contributing area for drainage time. *Geocarto Int.* **2018**, *33*, 810–829. [[CrossRef](#)]
44. Zhou, C.; Yin, K.; Cao, Y.; Ahmed, B.; Li, Y.; Catani, F.; Pourghasemi, H.R. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. *Comput. Geosci.* **2018**, *112*, 23–37. [[CrossRef](#)]
45. Romer, C.; Ferentinou, M. Shallow landslide susceptibility assessment in a semiarid environment—A Quaternary catchment of KwaZulu-Natal, South Africa. *Eng. Geol.* **2016**, *201*, 29–44. [[CrossRef](#)]
46. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
47. Mahdadi, F.; Boumezbeur, A.; Hadji, R.; Kanungo, D.P.; Zahri, F. GIS-based landslide susceptibility assessment using statistical models: A case study from Souk Ahras province, N-E Algeria. *Arab. J. Geos.* **2018**, *11*, 476. [[CrossRef](#)]
48. Achour, Y.; Garcia, S.; Cavaleiro, V. GIS-based spatial prediction of debris flows using logistic regression and frequency ratio models for Zêzere River basin and its surrounding area, Northwest Covilhã, Portugal. *Arab. J. Geos.* **2018**, *11*, 550. [[CrossRef](#)]
49. 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](#)]

