Article

Assessing Landslide Susceptibility by Coupling Spatial Data Analysis and Logistic Model

Antonio Ganga [1], Mario Elia [2,*], Ersilia D'Ambrosio [2], Simona Tripaldi [3], Gian Franco Capra [1], Francesco Gentile [2] and Giovanni Sanesi [2]

[1] Dipartimento di Architettura, Design e Urbanistica, Università degli Studi di Sassari, Viale Piandanna n 4, 07100 Sassari, Italy; anto.ganga@gmail.com (A.G.); pedolnu@uniss.it (G.F.C.)
[2] Department of Agricultural and Environmental Sciences, University of Bari Aldo Moro, 70126 Bari, Italy; dambrosioersilia@libero.it (E.D.); francesco.gentile@uniba.it (F.G.); giovanni.sanesi@uniba.it (G.S.)
[3] Department of Earth and Geo-Environmental Sciences, University of Bari Aldo Moro, 70121 Bari, Italy; simona.tripaldi@uniba.it
[*] Correspondence: mario.elia@uniba.it

Abstract: Landslides represent one of the most critical issues for landscape managers. They can cause injuries and loss of human life and damage properties and infrastructure. The spatial and temporal distribution of these detrimental events makes them almost unpredictable. Studies on landslide susceptibility assessment can significantly contribute to prioritizing critical risk zones. Further, landslide prevention and mitigation and the relative importance of the affecting drivers acquire even more significance in areas characterized by seismicity. This study aimed to investigate the relationship between a set of environmental variables and the occurrence of landslide events in an area of the Apulia Region (Italy). Logistic regression was applied to a landslide-prone area in the Apulia Region (Italy) to identify the main causative factors using a large dataset of environmental predictors (47). The results of this case study show that the logistic regression achieved a good performance, with an AUC (Area Under Curve) >70%. Therefore, the model developed would be a useful tool to define and assess areas for landslide occurrence and contribute to implementing risk mitigation strategy and land use policy.

Keywords: landslide; logistic regression; environmental hazard; risk assessment

1. Introduction

Landslides, one of the most devastating natural/human-induced disasters worldwide, cause injuries and loss of human life as well as damage to properties and infrastructure [1,2]. Haque et al. [3] estimated that from 1995 to 2014 in Europe, there were a total of 476 landslide events, with 1370 deaths and 740 injuries recorded. These alarming numbers illustrate the need to enhance national and regional efforts to prevent or minimize such impacts on human and natural assets.

With 620,808 recorded events, Italy is one of the most affected European countries [4]. Every year landslide events cause the displacement of hundreds of thousands of people, building and infrastructure damages, and conspicuous loss of cultural heritage. To show the magnitude of the recent status, 172 main events were recorded in 2017, 146 in 2016, 311 in 2015, 211 in 2014, and 112 in 2013. Related economic damages have been estimated at c.a. EUR 2.5 billion per year [4] in the period between 2014 and 2020.

Identifying critical zones becomes essential to managing landslide susceptibility, especially in areas where landslides severely threaten human and natural resources and the country's massive cultural heritage. In this regard, the scientific community can help decision-makers by developing ad hoc risk analysis models based on empirical evidence and drivers of landslide occurrence. Geomorphological and topographic conditions, seismicity, intense rainfall, and anthropogenic factors, such as road and railway infrastructures,

urban sprawl, land cover changes, and agricultural practices, represent the main predisposing and triggering factors of landslide events [5–7].

Studies focusing on landslide susceptibility assessment can provide an outstanding contribution to these purposes since they estimate the spatial distribution of landslide occurrence probability in a given area based on predisposing factors [1]. The literature is characterized by inventory-based studies, data-driven (including bivariate and multivariate statistics) analysis, knowledge-driven works, and probabilistic, physically-based, and deterministic approaches [2]. Recently, several machine learning-based techniques have been tested to develop landslide monitoring tools. However, these methods still feature significant drawbacks for modeling complex algorithms, including uncertain model performances, scarce aptitude to detect local variability, and the use of "black boxes", which is not appreciated yet by the operative world. Consequently, most landscape managers still prefer more traditional and sound models, such as logistic regression [8].

In fact, the most widely used method at local and regional scales is logistic regression [2,9,10], which has an advantage over other multivariate statistical methods in that it is independent of data distribution and able to handle continuous, categorical, and binary data [11]. In a logistic regression approach, the binary dependent variables are related to a vast dataset of independent drivers, such as topographic parameters, land uses, and others [12]. This study aims to test logistic regression in a survey area (Apulia, Italy), characterized by frequent seismic events and historically affected by widespread landslides. Specifically, we aimed to (*i*) understand the relationship between selected environmental variables and the occurrence probability of landslide events and (*ii*) produce landslide susceptibility maps for management, mitigation, and prevention purposes.

The proposed analytical framework can be applied to other areas worldwide with similar environmental features. Consequently, results may be helpful for landscape/environmental planners involved in landslide risk mitigation and land-use planning activities.

2. Materials and Methods

2.1. Study Area

The research was carried out across the Dauni Mountains (902.9 km^2; Min. Lat. 41°4′32.01″, Min. Long.14°55′59″, Max. Lat. 41°55′35.27″, Max. Long. 15°31′51.6″) which are located in the western part of the Apulia Region (Italy) (Figure 1). They belong to the Dauno Sub-Apennine dominion, representing the marginal external front of the Southern Apennine chain [13,14]. The chain units, affected by intense tectonic deformations, are thrusted over the Bradanic Trough unit, which in turn overlays the Apulian Foreland units. Folds and thrust structures and eastward verging, along with clay-rich flysch formations, characterize the structural and geological framework of the area [15,16].

As reported by the parametric catalog of Italian earthquakes CPTI15 [17,18], the study area has low seismic activity. However, more intense seismic activity characterizes the nearby Southern Apennines and Gargano Promontory areas, located less than 50 km from the study area [19,20]. Due to the more frequent and more energetic seismic activity of the Southern Apennine chain, the study area may be affected by seismic shaking able to induce landslides, as reported by Del Gaudio et al. (2012), who created a seismic hazard map of an area of the Daunia.

The elevation ranges between 73 and 1135 m a.s.l. (mean 566 m a.s.l.). The morphology is typically hilly-mountainous, strongly shaped by landslides favored by lithological and soil features, seismicity, and the area steepness (*vide infra*). The most diffuse lithotype are clay-limestone (CL), sands and conglomerates (SC), and clayey and clayey-limestone (CCL) units. Soils are strongly influenced by both parent material and the whole environmental features, being more developed in CL and CCL units (Alfisols and Ultisols) [21,22] while less (Inceptisols and Entisols) in SC units and where landslide movements modify their original nature. Land uses are featured by common arable crops (70%, with olives as main tree crops), broad-leaved, coniferous, and mixed forests (21%). Urban areas are very limited (1%). Mean annual rainfall (1918–2007) ranges from 533 to 931 mm. Mean monthly

temperatures range from 5 °C in January to 25 °C in August, negatively correlated with the altitude.
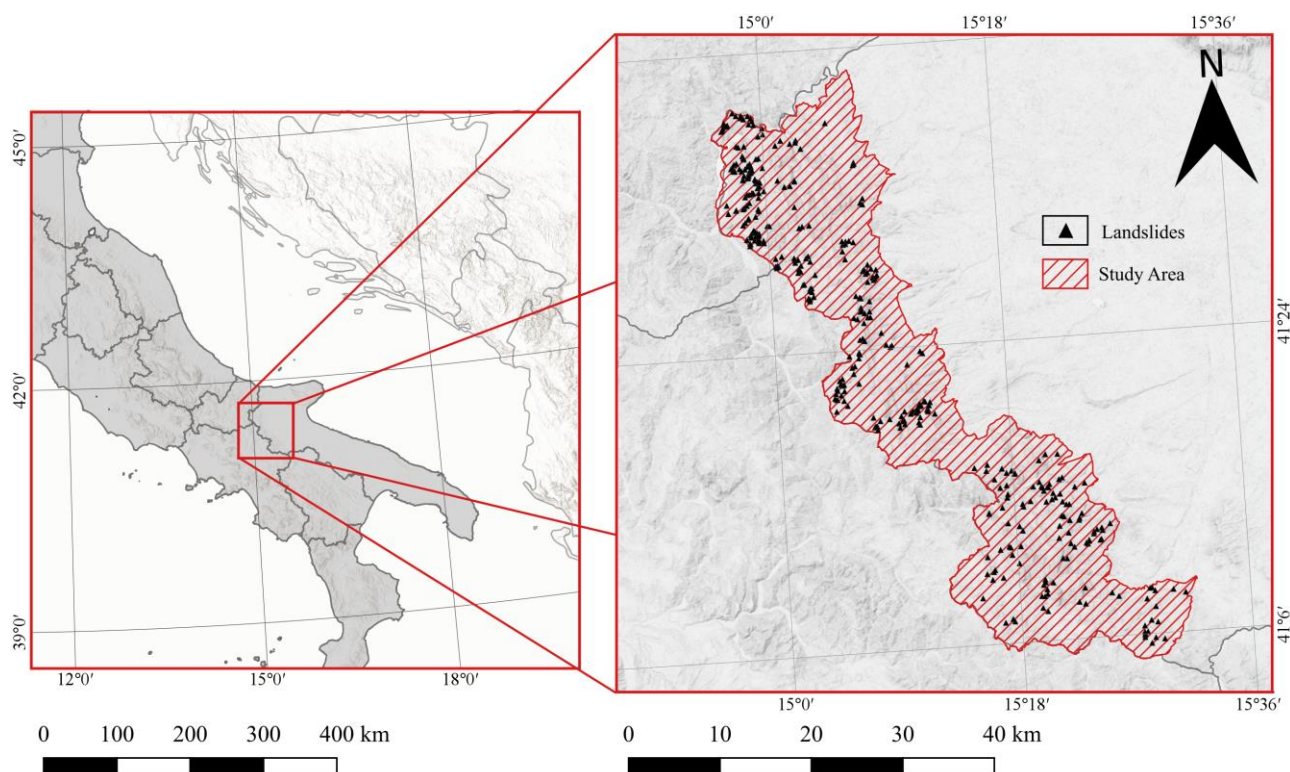
Figure 1. Dauni Mountains study area location with landslide detachment niches identification.

The study area is characterized by a high seismic hazard that, together with the previously reported environmental features, makes it extremely prone to landslide events; between 1918 and 2007, 515 mass movements were registered (Table 1).

Table 1. Number of landslides registered within the study area (1918–2007) with their main features.

Landslide Type	N.
Slow earthflow	228
Rotational/traslational landslide	143
Complex landslide	96
Rapid debris flow	18
Fall/topple	11
Area affected by numerous shallow landslide	11
Unclassified landslide	7
Sinkhole	1
Total number	515

2.2. Data Collection

We extracted data from multisource and multiscale geographic databases (Table S2). A total of 47 predictors were selected according to 7 main macro-class classification/factors: topography, seismic, geolithologic, pedologic, land use, morphologic, and climatic. Some of them are represented in the Figure S1.

The whole investigated area was divided into several grids of 250 square meters each. For each grid, the value of all predictors (vide supra) was calculated using GIS software. The seismic factor was evaluated using data from the national database and additionally evaluating the fault distance by applying the Euclidean distance algorithm.

Concerning the dependent variable, the location of landslide detachment niches was obtained from the landslides regional inventory. Since this database provides information concerning landslides verified between 1918 and 2007 without providing a specific occurrence date, the causative factors' assessment was limited to these years. Thus, climate causative factors were quantified as mean values for the same period. Meanwhile, the study area was reduced to subareas in which land use did not vary throughout this period in order not to consider the effect of land cover variation on slope stability. Available Corine Land Cover Maps (1990, 2000, 2006) were crossed, and only land cover persistence areas (582.2 km^2) were retained for further analysis. Overall, the total number of landslide detachment niches considered within the study area was 328.

Finally, the modified study area was divided into 9316 cells of 250 × 250 m, and, for each of them, the mean values of the causative factors were assessed. In addition, a dichotomic value (i.e., 0 or 1) linked to the absence or presence of landslide detachment niches was assigned to each cell as the dependent variable's value.

2.3. Causative Factor Selection

The first fundamental step was factor selection. Several factors occur in the process of landslide hazard assessment [1,2,23]. In fact, they are the result of interaction between intrinsic and external factors [24]. As a matter of fact, there are no definitive factor selection lists, and the tendency is often to rely on accessible databases [9]. However, the selected model allows for the use of a large number of independent variables, and therefore, a large number of databases have been considered, which are collected in the categories below (Figure 2).
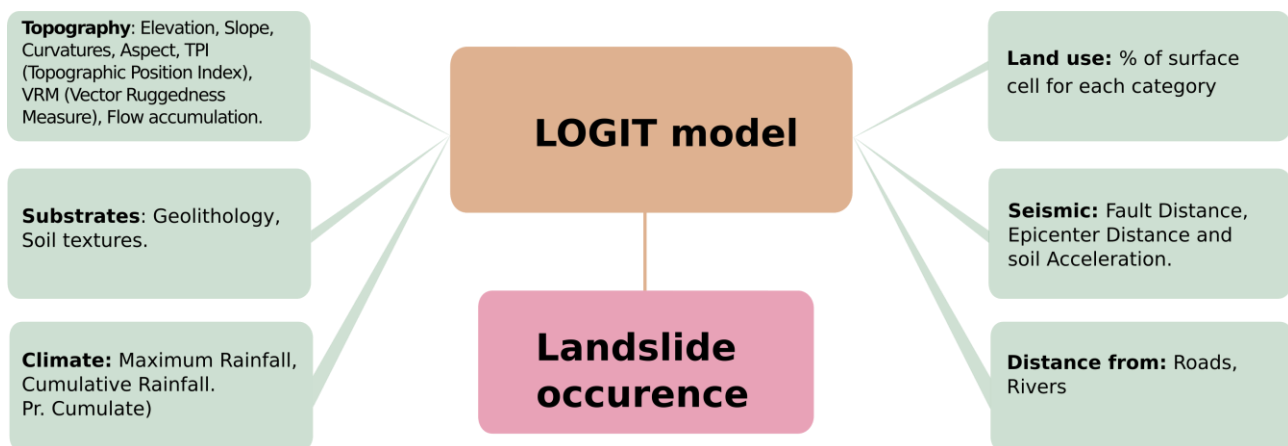
Figure 2. Flow chart of the methodological approach used.

The attached table (Table S1) shows all the calculated factors, the original source, and the spatial/temporal resolution. These data were acquired from public data infrastructures managed by individual regions, Italian Environmental Agencies (such as ISPRA), and the European Environmental Agency (EEA). According to [23], we classified the factors into (*i*) predisposing factors (e.g., topography and geological substrate), (*ii*) triggering factors (e.g., precipitations and earthquakes), and (*iii*) accelerating factors (e.g., human-altered systems).

Regarding predisposing factors, the following groups of predictors were considered:

- *Morphology.* The landform is one of the most important factors influencing landslide occurrence [25]. The slope angle is considered one of the main triggering factors for landslides [26,27]. The aspect is decisive for triggering landslide processes. It is strictly connected to climatic conditions [28] and the factor impacting slope stability [29]. The topographical indices are useful for describing and quantifying the intensity and

characteristics of hydrological processes [25]. The main topographical indices are reported in Table S1.

- The topographic factors were calculated starting with a high-resolution digital terrain model (DTM) with 8-m resolution.
- *Geolitology.* The lithological features also define the resistance capacity of the substrate to external forces. Indeed, this is related to material strength because they have varied compositions and structures for different rock types [9,30,31].
- *Pedology* (texture). Soil texture is one of the soil's main physical features. Previous work has shown how soils of different textures can have different levels of susceptibility to the triggering of landslides [24]. Hamza et al. (2017) indicated that limestone has more probability of landslide occurrence, thus providing a hazard index value of 1.10 (Table 2). In contrast, the probability of occurrence of landslide for gypsum and sandstone is comparatively lower, with a hazard index ranging from 0.80 to 0.77.
- *Distance from river*. This parameter has a strong link to the erosion process in hilly regions [9]. Rivers and streams play an important role in landscape modification by modeling the landforms and shaping the lithological substrate [32]. The distance from the river is calculated as the Euclidean distance by the river stream extracted from a high resolution (1:10000) land use map.

Table 2. List of causative factors.

Class of Factors	Factors Selected
Topography	Elevation; Flow accumulation; Planform curvature; Profile curvature; Standard curvature; Slope; Aspect; Topographic indexes (Slope roughness, Terrain roughness, Topographic wetness index, Vector Ruggedness Measure, Topographic Position Index, Topographic factor).
Seismic Factors	Peak ground acceleration; Distance to active fault; Distance to earthquake epicenter.
Geolithology	Percentage of cover area of Terrace alluvium; Sandstones and clays; Clays and marls; Clays; Debris; Alluvium and river-lake deposits; Beaches; Lakes and glaciers; Sands and conglomerates; Sandy and sand-marly units; Clayey and clayey-limestone units; Marly limestone units.
Pedology	Percentage of cover area of clay; Percentage of sand; Percentage of silt.
Land Cover	Percentage of cover area of Urban area; Agricultural area; Grassland; Forest; Natural area, Rivers and lakes.
Morphology	Distance to rivers; Distance to roads.
Climate	Mean annual maximum rainfall observed in 1 day; Mean annual rainfall.

Considered triggering factors were:

- *Seismic factor*. The dynamics of triggering landslides following an earthquake are well documented [33–35]. Earthquake-induced landslides are one of the worst natural hazards [36]. Specifically, there is a strong relationship between distance from faults and epicenters [37]. Furthermore, the factors triggering earthquake-induced landslides are also related to earthquake characteristics, such as ground acceleration [11].
- *Climate*. The climate affects landslides both directly and indirectly. For instance, intense rainfall is the most common triggering agent of landslides [38].

Finally, the following accelerating factors were considered:

- *Land cover/land use*. It influences the ability to eventually prevent and/or limit the extent and distribution of landslides. Forested and natural systems usually showed a statistically lower landslide occurrence, and they have an important role in prevent-

ing it. This is especially true when located in areas with critical topographical and lithological conditions [39].

- *Distance from road*. Roads are an important susceptibility factor to the triggering of landslides as their construction could modify the land topography and shape. As a matter of fact, roads represent an important driver of the soil profile and hillslope alteration [2].

Variable Selection

Predictors were selected starting from the original 47×9316 observation matrix dataset. In a multivariate statistical analysis, the selection process is one of the most important tasks [40]. In this work, the elimination of highly correlated predictors was carried out in the first variables selection according to previous works [41,42]. In this step, the predictors with Pearson's correlation coefficient >0.75 were removed. The final correlation matrix is reported in Table S1.

2.4. Logistic Regression

Logistic regression analyzes the relationship between multiple independent variables and a categorical dependent variable while estimating the occurrence probability of an event by fitting data to a logistic curve [43]. The logistic function is described as [31]:

$$P = \frac{1}{1 + (exp^{-Z})} \tag{1}$$

where P is the probability of landslide occurrence and Z a linear combination of casual X_i factors:

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n \tag{2}$$

where β_i are the landslide casual factor's coefficients.

The models were performed using R statistical software. In order to operate a predictor's selection, the regressions were performed using a stepwise model. Stepwise regression is a set of iterative search and model comparison procedures that identify which independent variables have the strongest association with the dependent one (Draper and Smith, 1981; Hauser, 1974). Using this approach, the model was performed for each type of landslide and for the total landslide.

For the four models, the Area under Curve (AUC), the Accuracy, and the Akaike information criterion (AIC) were calculated. In order to evaluate the predictive capacity of the logistic model, the McFadden pseudo R^2 area was calculated. This is intended as a logistic regression analog of R^2 and uses ordinary least-squares (OLS) regression [44].

3. Results

3.1. Main Factors Affecting Slope Stability

Table 3 reports the results of significant drivers for each model. In particular, stepwise logistic regression was performed using four different sub-datasets, from the whole landslide dataset, according to the following typology: (*i*) total landslides (TL); (*ii*) rotational/translational landslide (RTL); (*iii*) slow earth flow (SEF), and (*iv*) complex landslide (CL).

Table 3. List of coefficient drivers. Table S2 reports a complete list of factors with their significant metadata. The value of factors removed by stepwise logistic regression was not reported.

	Total landslide (TL)			Rotational/Traslational Landslide (RTL)			Slow Earth Flow (SEF)			Complex Landslide (CL)		
Factor	**Coeff.**	**SE**	*p*	**Coeff.**	**SE**	*p*	**Coeff.**	**SE**	*p*	**Coeff.**	**SE**	*p*
	−0.381	0.030	***	−0.704	0.037	***	−0.694	0.036	***	−0.388	0.031	***
T_TWI	−0.099	0.025	***	/	/	/	/	/	/	−0.207	0.023	***
T_VRM	0.086	0.023	***	0.076	0.027	*	/	/	/	0.050	0.026	.

Table 3. *Cont.*

Factor	Total landslide (TL)			Rotational/Traslational Landslide (RTL)			Slow Earth Flow (SEF)			Complex Landslide (CL)		
	Coeff.	SE	*p*	Coeff.	SE	*p*	Coeff.	SE	*p*	Coeff.	SE	*p*
T_DTM	/	/	/	0.112	0.026	***	−0.081	0.027	**	−0.040	0.026	.
T_F_acc	−0.042	0.025	.	−0.244	0.042	***	/	/	/	−0.223	0.040	***
T_Curv_pl	−0.053	0.022	**	0.100	0.025	***	−0.039	0.025	.	−0.064	0.023	**
T_Curv_st	0.106	0.022	***	0.146	0.024	***	−0.043	0.026	.	−0.037	0.020	.
T_N	/	/	/	0.054	0.025	*	/	/	/	−0.112	0.024	**
T_NE	/	/	/	−0.196	0.030	***	/	/	/	/	/	/
T_E	/	/	/	−0.172	0.029	***	/	/	/	/	/	/
T_SE	/	/	/	−0.040	0.025	.	/	/	/	/	/	/
T_S	0.117	0.021	***	/	/	/	0.061	0.025	.	/	/	/
T_SW	/	/	/	0.038	0.023	.	−0.052	0.025	.	0.060	0.021	**
T_W	−0.038	0.023	.	0.141	0.027	***	−0.088	0.025	***	−0.043	0.024	.
T_NW	0.067	0.022	**	/	/	/	/	/	/	0.089	0.021	.
T_FLAT	−0.060	0.035	.	−0.235	0.041	***	−0.089	0.034	**	/	/	/
T_Ls	0.182	0.024	***	0.115	0.029	***	0.175	0.027	***	0.132	0.022	***
T_TPI	0.112	0.022	***	0.060	0.022	**	0.253	0.022	***	−0.036	0.021	.
S_Epic	0.056	0.027	*	−0.252	0.028	***	0.286	0.031	***	−0.112	0.024	*
G_TA	−0.071	0.043	.	−0.098	0.046	*	−0.081	0.039	*	−0.141	0.041	***
G_SSC	0.137	0.018	***	−0.113	0.024	***	0.206	0.017	***	0.030	0.020	.
G_CM	−0.042	0.022	.	−0.132	0.025	***	−0.224	0.039	***	0.034	0.019	.
G_C	−0.236	0.031	***	−0.333	0.040	***	−0.260	0.031	***	−0.319	0.035	***
G_DDB	−0.262	0.037	***	−0.140	0.043	***	−0.244	0.038	***	−0.262	0.035	***
G_LG	/0.134	0.044	**	−0.077	0.038	*	−0.069	0.037	.	−0.068	0.037	.
G_SC	0.042	0.023	.	0.061	0.023	**	−0.280	0.033	***	/	/	/
G_SSM	0.045	0.019	*	−0.136	0.034	**	−0.119	0.036	**	0.101	0.010	***
G_CCL	0.099	0.022	***	−0.092	0.028	**	0.083	0.023	***	0.103	0.022	*
G_ML	/	/	/	0.157	0.027	**	/	/	/	−0.077	0.024	**
P_sand	−0.110	0.026	***	0.095	0.028	***	−0.241	0.030	***	/	/	/
P_silt	−0.169	0.023	***	−0.311	0.029	***	−0.161	0.024	***	0.106	0.028	***
LU_urb	0.068	0.013	***	0.130	0.011	***	−0.160	0.042	***	0.066	0.013	**
LU_grs	−0.072	0.024	**	−0.073	0.022	***	−0.257	0.037	***	−0.137	0.030	**
LU_for	/	/	/	0.079	0.023	***	−0.089	0.026	***	0.148	0.021	***
LU_nat	0.149	0.016	***	0.111	0.018	***	0.064	0.017	***	/	/	/
D_riv	−0.159	0.024	***	−0.041	0.022	.	−0.213	0.031	***	0.073	0.025	**
D_road	−0.338	0.026	**	−0.389	0.028	***	−0.303	0.027	***	−0.380	0.028	***
C_p_max	−0.039	0.023	.	/	/	/	−0.164	0.025	***	0.168	0.025	***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$.

3.2. Model Performance

In Table 4, performances of the model applied on four sub-datasets (*vide supra*) were reported.

Table 4. Models performance.

Landslide	AIC *	AUC **	Accuracy	McFadden's Pseudo R^2
All	8004.13	0.76	0.71	0.12
Rotational/Translational	7407.88	0.68	0.73	0.19
Slow Earth Flow	7200.94	0.80	0.77	0.21
Complex Landslide	7828.44	0.69	0.70	0.13

* AIC = Akaike Information Criterion. ** AUC = Area Under Curve.

Slow earth flow showed the best performance, with an accuracy >0.75, meaning that this can be considered an accurate model. Overall, all implemented models showed values > 0.70, thus considered suitable. The area under curve (AUC) values are comparable with previous similar studies [45]. To evaluate the model's performance, McFadden's

Pseudo R squared, instead of R squared, was applied. A McFadden's Pseudo R squared between 0.2–0.4 indicates a very good fit. The RTL and SEF landslide models showed the best fit with 0.19 and 0.21, respectively.

3.3. Susceptibility Mapping

According to the method performed by Elia et al. (2019), four susceptibility maps were created; each recorded landslide's detachment niche was georeferenced.

Implemented models and resulting maps show the landslide occurrence probability within the study area (Figure 3). Based on probability values, areas were classified in terms of susceptibility as: (i) high (>0.65, red areas); (ii) moderate (0.50–0.65, orange); (iii) moderate-low (0.30–0.50, yellow); (iv) low (20–30, light green); and (v) very low (0.00–0.20, green). In the total landslide map, areas affected by high-level susceptibility are distributed mainly in the northwest area. Conversely, in RTL maps, areas featured by high-level susceptibility are localized in the southern part of the region. In the slow heart flow model maps, the susceptibility distribution is closer to that observed for the total model; however, areas with high susceptibility are more concentrated in the extreme northwestern part. In the complex landslides map, high-level susceptibility spots were identified in several parts of the investigated area. Overall, in every implemented map, the low susceptibility is concentrated in the central–east part of the study area, in correspondence with the less hilly areas. In fact, the study area shows an elevation gradient from west to east, and the landslides distribution follows this gradient, especially in RTL and CL.
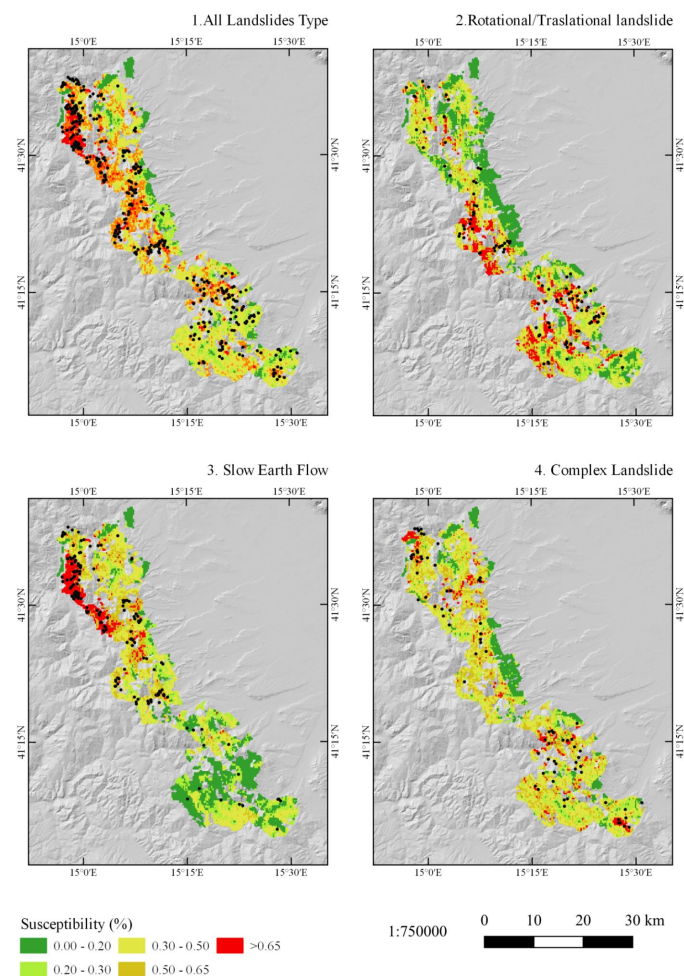
Figure 3. Landslide triggering probability distribution maps. The black dots represent the landslide positions.

4. Discussion

4.1. Logistic Regression Metrics

The assessment of areas likely to produce land movement phenomena was carried out by adopting a logistic regression model, one of the most used approaches in landslide susceptibility analysis. In fact, it allows predicting where landslides are more likely to occur based on the relationship between past events and ad hoc selected causative factors. The logistic regression was demonstrated to be highly suitable in assessing what the most correlated and the most predictive factors in estimating the landslide occurrence susceptibility are. This method provided consistent and significant results in the investigated area. By carrying out a cutoff of the probability of occurrence equal to 0.5, it appears that about 80% of the territory has been classified correctly, i.e., that for 80% of cases where no event has occurred, the probability is lower than 0.5; on the opposite, it is higher where events occurred.

According to the literature, AUC values of 0.5–0.7 are considered low accuracy, 0.7–0.9 suggest valuable applications, and values around 0.9 indicate high accuracy (Swets 1988). In our study, the AUC indicates 0.76, 0.68, 0.80, and 0.69 for TL, RTL, SEF, and CL, respectively, between predicted probabilities and observed outcomes. The findings are consistent with other models developed to estimate landslide susceptibility worldwide, especially given the high rate of human and natural variability across the study area [34,46,47].

4.2. Correlation Analysis

As shown in Table 2, the correlation analysis revealed that only 37 of the 47 causative factors were uncorrelated and were thus included in the subsequent logistic regression. These factors are related to topography (T_E, T_F_acc, T_FLAT, T_NW, T_TPI, T_S, T_TWI, T_VRM, T_W), seismicity (S_Epic), geolithology (G_TA, G_SSC, G_CM, G_C, G_DDB, G_LG, G_SC, G_SSM, G_CCL, G_ML), land cover (LU_urb, LU_grs, LU_for, LU_nat), morphology (D_riv, D_road), and climate (C_p_max). It is well known [48–51] that the slope length and steepness factor (T_Ls) is one of the main landslide predictors, being a parameter used to characterize the effects of topography and hydrology on soil loss [52,53]. In our study, the T_Ls showed the highest coefficient in predicting the totality of recorded events and one of the most predicting variables for the other landslide types, evidencing its influence in the evolution of landslide types. Similarly, Huangfu et al. (2021) [54] found that slope is one of the three most important predictors affecting landslide geo-hazard risk in China.

Generally speaking, landslides occurring on steep slopes are triggered by a reduction of the apparent cohesion of colluvium (soil and debris accumulated upon an impermeable bedrock), resulting from water infiltration into the soil. Concave shape also affects the development of rotational/translational mass movement, reflecting the importance of the hydrological regime in slope stability. In these areas, a convergence of surface and subsurface water streams saturates rapidly the soil, which becomes more prone to movements [51].

As expected, most of the other topographic variables showed a high predicting power in landslide occurrence. For example, T_TPI showed an elevated ability to predict the landslide, particularly the SEF. This finding is consistent with other studies [55,56], showing how the TPI is essential for slope stability analysis aiming at classifying priority zones in landslide occurrence [57]. However, the configuration of such a set of predictors is quite different among the four landslide types. Table 2 suggests that topographic and land cover variables are the most significant predictors for RTL type; seismic and geolithological variables had the highest predicting power for SEF, while the presence of forest and annual maximum rainfall significantly affects CL. These findings demonstrated that different landslide types are regulated by different predisposing or preparatory factors [58]. For example, distance to earthquake epicenter (S_Epic) had only a positive and significant relationship with the SEF landslide type. This is probably because mountainous landscapes, susceptible to landslides, are often characterized by moderate rates of earthquake events. Many authors similarly found this relationship [59–61], arguing that the presence of a fault damage

zone is the primary control on the distribution of earth flows; this was primarily due to ground quaking and the subsequent expansion of superficial debris favoring accelerated water infiltration.

On the contrary, all types of landslides showed a negative relationship with the distance from the roads, while only two types with the distance from rivers. The more the distance from the road increases, the more the probability of a landslide decreases. This negative trend was observed by other authors worldwide [62–65], highlighting the importance of anthropogenic impact on slope stability. Additionally, other authors [26,66] reported that the landslide susceptibility gradually increases up to a proximity zone of 200 m from a road network.

4.3. Perspectives

Our findings in the Apulia region are of primary importance for a more broad assessment of landslide and seismic risk management. The proposed approach can be applied to a larger scale of analysis which constitutes one of the future perspectives. Indeed, we recognize the limits of our study. Firstly, to achieve a significant landslide susceptibility assessment, we used the most detailed available datasets on landslide occurrence and environmental predictors. However, although we employed the most harmonized and available dataset, we are aware that there are several limitations in its accuracy. Numerous factors such as the phenomenon's complexity, the high diversity of geomorphic landslide features, and the use of different mapping and sampling procedures undoubtedly affected the accuracy and precision of the applied landslide dataset [67–69]. Additionally, our model results from data collected in dynamic landscapes and ecosystems constantly subjected to changes. Therefore, starting from these findings, new and innovative approaches must be developed as new knowledge is obtained from a geological, hydrological, and geotechnical point of view.

5. Conclusions

Landslides are one of the most devastating natural disasters, causing human and economical losses. In order to mitigate landslide risk, it is fundamental to identify critical risk zones. In the current study, logistic regression was applied to a landslide-prone area in the Apulia Region (Italy) in order to identify the main causative factors and produce a landslide susceptibility map. The model has never been applied before in the study area.

In this case study, it was found that logistic regression achieved a good performance, with an AUC value higher than 70%. Thus, it could constitute a useful tool in identifying critical areas for landslide occurrence and defining a risk mitigation strategy and land use policy. For example, in high–moderate susceptibility areas, mitigation infrastructures, resilience building, and a relocation strategy for families living in those areas could be provided. In addition, as a result of this study, it was found that human infrastructures, such as road networks, are one of the main elements triggering landslides. Thus, peculiar attention must be paid to the choice of the route when designing the transport networks.

However, further detailed studies should be conducted to compare different models, widen the number of causative factors considered (i.e., geotechnical data), and assess landslide susceptibility individually for each type of landslide since the current study combined different types of landslides. Nonetheless, this study gives a comprehensive and preliminary understanding of the likely future insights for researchers and policy makers.

Author Contributions: Conceptualization, A.G., M.E. and E.D.; methodology A.G. and M.E.; software, A.G. and E.D.; validation A.G., M.E., F.G. and G.S.; formal analysis, A.G.; investigation, A.G., M.E. and E.D.; resources, A.G., M.E., F.G. and G.S.; data curation, A.G., M.E. and E.D.; writing—original draft preparation, A.G., M.E., F.G., G.F.C., S.T. and G.S.; writing—review and

References

1. Chen, W.; Sun, Z.; Han, J. Landslide Susceptibility Modeling Using Integrated Ensemble Weights of Evidence with Logistic Regression and Random Forest Models. *Appl. Sci. Switz.* **2019**, *9*, 171. [CrossRef]

2. Nsengiyumva, J.B.; Luo, G.; Amanambu, A.C.; Mind'je, R.; Habiyaremye, G.; Karamage, F.; Ochege, F.U.; Mupenzi, C. Comparing Probabilistic and Statistical Methods in Landslide Susceptibility Modeling in Rwanda/Centre-Eastern Africa. *Sci. Total Environ.* **2019**, *659*, 1457–1472. [CrossRef] [PubMed]

3. Haque, U.; Blum, P.; da Silva, P.F.; Andersen, P.; Pilz, J.; Chalov, S.R.; Malet, J.-P.; Auflič, M.J.; Andres, N.; Poyiadji, E.; et al. Fatal Landslides in Europe. *Landslides* **2016**, *13*, 1545–1554. [CrossRef]

4. Trigila, A.; Iadanza, C.; Bussettin, M.; Lastoria, B. *Dissesto Idrogeologico in Italia: Pericolosità e Indicatori di Rischio*; Rap-porto_Dissesto_Idrogeologico_ISPRA_287_2018_Web.Pdf; ISPRA: Milan, Italy, 2018.

5. Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M.; Ardizzone, F. Probabilistic Landslide Hazard Assessment at the Basin Scale. *Geomorphology* **2005**, *72*, 272–299. [CrossRef]

6. Nepal, N.; Chen, J.; Chen, H.; Wang, X.; Pangali Sharma, T.P. Assessment of Landslide Susceptibility along the Araniko Highway in Poiqu/Bhote Koshi/Sun Koshi Watershed, Nepal Himalaya. *Prog. Disaster Sci.* **2019**, *3*, 100037. [CrossRef]

7. Lafortezza, R.; Tanentzap, A.J.; Elia, M.; John, R.; Sanesi, G.; Chen, J. Prioritizing fuel management in urban interfaces threatened by wildfires. *Ecol. Indic.* **2015**, *48*, 342–347. [CrossRef]

8. Elia, M.; D'Este, M.; Ascoli, D.; Giannico, V.; Spano, G.; Ganga, A.; Colangelo, G.; Lafortezza, R.; Sanesi, G. Estimating the Probability of Wildfire Occurrence in Mediterranean Landscapes Using Artificial Neural Networks. *Environ. Impact Assess. Rev.* **2020**, *85*, 106474. [CrossRef]

9. Soma, A.S.; Kubota, T.; Mizuno, H. Optimization of Causative Factors Using Logistic Regression and Artificial Neural Network Models for Landslide Susceptibility Assessment in Ujung Loe Watershed, South Sulawesi Indonesia. *J. Mt. Sci.* **2019**, *16*, 383–401. [CrossRef]

10. Meinhardt, M.; Fink, M.; Tünschel, H. Landslide Susceptibility Analysis in Central Vietnam Based on an Incomplete Landslide Inventory: Comparison of a New Method to Calculate Weighting Factors by Means of Bivariate Statistics. *Geomorphology* **2015**, *234*, 80–97. [CrossRef]

11. Dou, J.; Bui, D.T.; Yunus, A.P.; Jia, K.; Song, X.; Revhaug, I.; Xia, H.; Zhu, Z. Optimization of Causative Factors for Landslide Susceptibility Evaluation Using Remote Sensing and GIS Data in Parts of Niigata, Japan. *PLoS ONE* **2015**, *10*, e0133262. [CrossRef]

12. Elia, M.; Giannico, V.; Spano, G.; Lafortezza, R.; Sanesi, G. Likelihood and frequency of recurrent fire ignitions in highly urbanised Mediterranean landscapes. *Int. J. Wildland Fire* **2020**, *29*, 120. [CrossRef]

13. Merlini, S.; Mostardini, F. Appennino Centro-Meridionale: Sezioni Geologiche e Proposta Di Modello Strutturale. In Proceedings of the Geologia dell'Italia Centrale, Congresso Nazionale, Rome, Italy, 30 September–4 October 1986; Volume 73, pp. 147–149.

14. Patacca, E.; Scandone, P. Geology of the Southern Apennines. *Boll. Soc. Geol. Ital.* **2007**, *7*, 75–112.

15. Dazzaro, L.; Di Nocera, S.; Pescatore, T.; Rapisardi, L.; Romeo, M.; Russo, B.; Senatore, M.; Torre, M. Geologia Del Margine Della Catena Appenninica Tra Il F. Fortore e Il T. Calaggio (Appennino Meridionale). *Mem. Della Soc. Geol. Ital.* **1988**, *41*, 411–422.

16. Dazzaro, L.; Rapisardi, L. Integration of persistent scatterer interferometry and ground data for landslide monitoring: The Pianello landslide (Bovino, Southern Italy). *Mem. Soc. Geol. It.* **1996**, *16*, 143–147.

17. Rovida, A.; Locati, M.; Camassi, R.; Lolli, B.; Gasperini, P.; Antonucci, A. Italian Parametric Earthquake Catalogue (CPTI15), Version 3.0. Istituto Nazionale Di Geofisica e Vulcanologia (INGV). Available online: https://emidius.mi.ingv.it/CPTI15-DBMI15_v3.0/ (accessed on 1 June 2022).

18. Rovida, A.; Locati, M.; Camassi, R.; Lolli, B.; Gasperini, P. The Italian Earthquake Catalogue CPTI15. *Bull. Earthq. Eng.* **2020**, *18*, 2953–2984. [CrossRef]

19. Valensise, G.; Pantosti, D.; Basili, R. Seismology and Tectonic Setting of the 2002 Molise, Italy, Earthquake. *Earthq. Spectra* **2004**, *20*, 23–37. [CrossRef]

20. Miccolis, S.; Filippucci, M.; de Lorenzo, S.; Frepoli, A.; Pierri, P.; Tallarico, A. Seismogenic Structure Orientation and Stress Field of the Gargano Promontory (Southern Italy) From Microseismicity Analysis. *Front. Earth Sci.* **2021**, *9*, 179. [CrossRef]

21. Malagnini, L.; Scognamiglio, L.; Mercuri, A.; Akinci, A.; Mayeda, K. Strong Evidence for Non-Similar Earthquake Source Scaling in Central Italy. *Geophys. Res. Lett.* **2008**, *35*, 17. [CrossRef]

22. Soil Survey Staff. *Keys to Soil Taxonomy*, 12th ed.; USDA—Natural Resources Conservation Service: Washington, DC, USA, 2014.

23. Santini, M.; Grimaldi, S.; Nardi, F.; Petroselli, A.; Rulli, M.C. Pre-Processing Algorithms and Landslide Modelling on Remotely Sensed DEMs. *Geomorphology* **2009**, *113*, 110–125. [CrossRef]

24. Hamza, T.; Raghuvanshi, T.K. GIS Based Landslide Hazard Evaluation and Zonation – A Case from Jeldu District, Central Ethiopia. *J. King Saud Univ. Sci.* **2017**, *29*, 151–165. [CrossRef]

25. Kadavi, P.R.; Lee, C.-W.; Lee, S. Landslide-Susceptibility Mapping in Gangwon-Do, South Korea, Using Logistic Regression and Decision Tree Models. *Environ. Earth Sci.* **2019**, *78*, 116. [CrossRef]

26. Myronidis, D.; Papageorgiou, C.; Theophanous, S. Landslide Susceptibility Mapping Based on Landslide History and Analytic Hierarchy Process (AHP). *Nat. Hazards* **2016**, *81*, 245–263. [CrossRef]

27. Rozos, D.; Bathrellos, G.D.; Skillodimou, H.D. Comparison of the Implementation of Rock Engineering System and Analytic Hierarchy Process Methods, upon Landslide Susceptibility Mapping, Using GIS: A Case Study from the Eastern Achaia County of Peloponnesus, Greece. *Environ. Earth Sci.* **2011**, *63*, 49–63. [CrossRef]

28. Bednarik, M.; Magulová, B.; Matys, M.; Marschalko, M. Landslide Susceptibility Assessment of the Kraľovany–Liptovský Mikuláš Railway Case Study. *Phys. Chem. Earth Parts ABC* **2010**, *35*, 162–171. [CrossRef]

29. 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]

30. Kanungo, D.P.; Arora, M.K.; Sarkar, S.; Gupta, R.P. A Comparative Study of Conventional, ANN Black Box, Fuzzy and Combined Neural and Fuzzy Weighting Procedures for Landslide Susceptibility Zonation in Darjeeling Himalayas. *Eng. Geol.* **2006**, *85*, 347–366. [CrossRef]

31. Rasyid, A.R.; Bhandary, N.P.; Yatabe, R. Performance of frequency ratio and logistic regression model in creating GIS based landslides susceptibility map at Lompobattang Mountain, Indonesia. *GeoEnviron. Disasters* **2016**, *3*, 19. [CrossRef]

32. Meten, M.; PrakashBhandary, N.; Yatabe, R. Effect of Landslide Factor Combinations on the Prediction Accuracy of Landslide Susceptibility Maps in the Blue Nile Gorge of Central Ethiopia. *Geoenviron. Disasters* **2015**, *2*, 9. [CrossRef]

33. Rollo, F.; Rampello, S. Probabilistic assessment of seismic-induced slope displacements: An application in Italy. *Bull. Earthq. Eng.* **2021**, *19*, 4261–4288. [CrossRef]

34. Cao, J.; Zhang, Z.; Wang, C.; Liu, J.; Zhang, L. Susceptibility Assessment of Landslides Triggered by Earthquakes in the Western Sichuan Plateau. *Catena* **2019**, *175*, 63–76. [CrossRef]

35. Su, L.; Hu, K.; Zhang, W.; Wang, J.; Lei, Y.; Zhang, C.; Cui, P.; Pasuto, A.; Zheng, Q. Characteristics and Triggering Mechanism of Xinmo Landslide on 24 June 2017 in Sichuan, China. *J. Mt. Sci.* **2017**, *14*, 1689–1700. [CrossRef]

36. Tian, Y.; Xu, C.; Hong, H.; Zhou, Q.; Wang, D. Mapping Earthquake-Triggered Landslide Susceptibility by Use of Artificial Neural Network (ANN) Models: An Example of the 2013 Minxian (China) Mw 5.9 Event. *Geomat. Nat. Hazards Risk* **2019**, *10*, 1–25. [CrossRef]

37. Chen, X.L.; Yu, L.; Wang, M.M.; Li, J.Y. Brief Communication: Landslides Triggered by the M_s= 7.0 Lushan Earthquake, China. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1257–1267. [CrossRef]

38. Sidle, R.C. Using Weather and Climate Information for Landslide Prevention and Mitigation. In *Climate and Land Degradation*; Sivakumar, M.V.K., Ndiang'ui, N., Eds.; Environmental Science and Engineering; Springer: Berlin/Heidelberg, Germany, 2007; pp. 285–307. ISBN 978-3-540-72438-4.

39. Persichillo, M.G.; Bordoni, M.; Meisina, C. The Role of Land Use Changes in the Distribution of Shallow Landslides. *Sci. Total Environ.* **2017**, *574*, 924–937. [CrossRef]

40. Zellner, D.; Keller, F.; Zellner, G.E. Variable Selection in Logistic Regression Models. *Commun. Stat. Simul. Comput.* **2004**, *33*, 787–805. [CrossRef]

41. D'Ambrosio, E.; De Girolamo, A.M.; Barca, E.; Ielpo, P.; Rulli, M.C. Characterising the Hydrological Regime of an Ungauged Temporary River System: A Case Study. *Environ. Sci. Pollut. Res.* **2017**, *24*, 13950–13966. [CrossRef]

42. Elia, M.; Giannico, V.; Lafortezza, R.; Sanesi, G. Modeling Fire Ignition Patterns in Mediterranean Urban Interfaces. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 169–181. [CrossRef]

43. Park, H.-A. An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain. *J. Korean Acad. Nurs.* **2013**, *43*, 154. [CrossRef]

44. Smith, T.J.; McKenna, C.M. A Comparison of Logistic Regression Pseudo R2 Indices. *Mult. Linear Regres. Viewp.* **2013**, *39*, 10.

45. 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]

46. Chen, W.; Yan, X.; Zhao, Z.; Hong, H.; Bui, D.T.; Pradhan, B. Spatial Prediction of Landslide Susceptibility Using Data Mining-Based Kernel Logistic Regression, Naive Bayes and RBFNetwork Models for the Long County Area (China). *Bull. Eng. Geol. Environ.* **2019**, *78*, 247–266. [CrossRef]

47. Murillo-García, F.G.; Steger, S.; Alcántara-Ayala, I. Landslide Susceptibility: A Statistically-Based Assessment on a Depositional Pyroclastic Ramp. *J. Mt. Sci.* **2019**, *16*, 561–580. [CrossRef]

48. Yang, X.; Yang, X. Digital Mapping of RUSLE Slope Length and Steepness Factor across New South Wales, Australia. *Soil Res.* **2015**, *53*, 216–225. [CrossRef]

49. Chen, W.; Pourghasemi, H.R.; Kornejady, A.; Zhang, N. Landslide Spatial Modeling: Introducing New Ensembles of ANN, MaxEnt, and SVM Machine Learning Techniques. *Geoderma* **2017**, *305*, 314–327. [CrossRef]

50. Çellek, S. Effect of the Slope Angle and Its Classification on Landslide. *Nat. Hazards Earth Syst. Sci. Discuss.* **2020**, 1–23. [CrossRef]

51. Akinci, H.; Zeybek, M.; Dogan, S. *Evaluation of Landslide Susceptibility of Şavşat District of Artvin Province (Turkey) Using Machine Learning Techniques*; IntechOpen: London, UK, 2021; ISBN 978-1-83969-024-2.

52. Hudson, N.W. Soil and Water Conservation (Second Edition). By F. R. Troeh, J. A. Hobbs and R. L. Donahue. Hemel Hempstead, Herts, UK: Prentice Hall, (1991), Pp. 530, £55.65, ISBN 13-832324-X. *Exp. Agric.* **1992**, *28*, 127. [CrossRef]

53. Troeh, F.R.; Hobbs, J.A.; Donahue, R.L.; Troeh, F.R. *Soil and Water Conservation*, 2nd ed.; Prentice Hall: Hoboken, NJ, USA, 1991.

54. Huangfu, W.; Wu, W.; Zhou, X.; Lin, Z.; Zhang, G.; Chen, R.; Song, Y.; Lang, T.; Qin, Y.; Ou, P.; et al. Landslide Geo-Hazard Risk Mapping Using Logistic Regression Modeling in Guixi, Jiangxi, China. *Sustainability* **2021**, *13*, 4830. [CrossRef]

55. Park, H.J.; Jang, J.Y.; Lee, J.H. Assessment of Rainfall-Induced Landslide Susceptibility at the Regional Scale Using a Physically Based Model and Fuzzy-Based Monte Carlo Simulation. *Landslides* **2019**, *16*, 695–713. [CrossRef]

56. Nseka, D.; Kakembo, V.; Bamutaze, Y.; Mugagga, F. Analysis of Topographic Parameters Underpinning Landslide Occurrence in Kigezi Highlands of Southwestern Uganda. *Nat. Hazards* **2019**, *99*, 973–989. [CrossRef]

57. Tağıl, Ş.; Jenness, J.; Tagil, S.; Jenness, J. GIS-Based Automated Landform Classification and Topographic, Landcover and Geologic Attributes of Landforms Around the Yazoren Polje, Turkey. *J. Appl. Sci.* **2008**, *8*, 910–921. [CrossRef]

58. Zêzere, J.L. Landslide Susceptibility Assessment Considering Landslide Typology. A Case Study in the Area North of Lisbon (Portugal). *Nat. Hazards Earth Syst. Sci.* **2002**, *2*, 73–82. [CrossRef]

59. Molnar, P.; Anderson, R.S.; Anderson, S.P. Tectonics, Fracturing of Rock, and Erosion. *J. Geophys. Res. Earth Surf.* **2007**, *112*. [CrossRef]

60. Clarke, B.A.; Burbank, D.W. Bedrock Fracturing, Threshold Hillslopes, and Limits to the Magnitude of Bedrock Landslides. *Earth Planet. Sci. Lett.* **2010**, *297*, 577–586. [CrossRef]

61. Scheingross, J.S.; Minchew, B.M.; Mackey, B.H.; Simons, M.; Lamb, M.P.; Hensley, S. Fault-Zone Controls on the Spatial Distribution of Slow-Moving Landslides. *GSA Bull.* **2013**, *125*, 473–489. [CrossRef]

62. Brenning, A.; Schwinn, M.; Ruiz-Páez, A.P.; Muenchow, J. Landslide Susceptibility near Highways Is Increased by 1 Order of Magnitude in the Andes of Southern Ecuador, Loja Province. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 45–57. [CrossRef]

63. Pasang, S.; Kubíček, P. Landslide Susceptibility Mapping Using Statistical Methods along the Asian Highway, Bhutan. *Geosciences* **2020**, *10*, 430. [CrossRef]

64. Zhou, C.; Cao, Y.; Yin, K.; Wang, Y.; Shi, X.; Catani, F.; Ahmed, B. Landslide Characterization Applying Sentinel-1 Images and InSAR Technique: The Muyubao Landslide in the Three Gorges Reservoir Area, China. *Remote Sens.* **2020**, *12*, 3385. [CrossRef]

65. Ou, P.; Wu, W.; Qin, Y.; Zhou, X.; Huangfu, W.; Zhang, Y.; Xie, L.; Huang, X.; Fu, X.; Li, J.; et al. Assessment of Landslide Hazard in Jiangxi Using Geo-Information. *Front. Earth Sci. China* **2021**, *9*, 648342. [CrossRef]

66. Bathrellos, G.D.; Kalivas, D.P.; Skilodimou, H.D. GIS-Based Landslide Susceptibility Mapping Models Applied to Natural and Urban Planning in Trikala, Central Greece. *Estud. Geológicos* **2009**, *65*, 49–65. [CrossRef]

67. Guzzetti, F.; Carrara, A.; Cardinali, M.; Reichenbach, P. Landslide Hazard Evaluation: A Review of Current Techniques and Their Application in a Multi-Scale Study, Central Italy. *Geomorphology* **1999**, *31*, 181–216. [CrossRef]

68. Wu, X.; Chen, X.; Zhan, F.B.; Hong, S. Global Research Trends in Landslides during 1991–2014: A Bibliometric Analysis. *Landslides* **2015**, *12*, 1215–1226. [CrossRef]

69. Lima, P.; Steger, S.; Glade, T.; Tilch, N.; Schwarz, L.; Kociu, A. Landslide Susceptibility Mapping at National Scale: A First Attempt for Austria. In *Advancing Culture of Living with Landslides*; Mikos, M., Tiwari, B., Yin, Y., Sassa, K., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 943–951.