

Article

# Comparing Forward Conditional Analysis and Forward Logistic Regression Methods in a Landslide Susceptibility Assessment: A Case Study in Sicily

Dario Costanzo <sup>1,\*</sup>  and Clemente Irigaray <sup>2</sup>

<sup>1</sup> Department of Earth and Sea Sciences, University of Palermo, 90123 Palermo, Italy

<sup>2</sup> Department of Civil Engineering, University of Granada, 18010 Granada, Spain; clemente@ugr.es

\* Correspondence: costanzodario@gmail.com

Received: 15 June 2020; Accepted: 9 July 2020; Published: 10 July 2020



**Abstract:** Forward logistic regression and conditional analysis have been compared to assess landslide susceptibility across the whole territory of the Sicilian region (about 25,000 km<sup>2</sup>) using previously existing data and a nested tiered approach. These approaches were aimed at singling out a statistical correlation between the spatial distribution of landslides that have affected the Sicilian region in the past, and a set of controlling factors: outcropping lithology, rainfall, landform classification, soil use, and steepness. The landslide inventory used the proposal of building the models like the official one obtained in the PAI (hydro geologic asset plan) project, amounting to more than 33,000 events. The 11 types featured in PAI were grouped into 4 macro-typologies, depending on the inherent conditions believed to generate various kinds of failures and their kinematic evolution. The study has confirmed that it is possible to carry out a regional landslide susceptibility assessment based solely on existing data (i.e., factor maps and the landslide archive), saving a considerable amount of time and money. For scarp landslides, where the selected factors (steepness, landform classification, and lithology) are more discriminate, models show excellent performance: areas under receiver operating characteristic (ROC) (AUCs) average > 0.9, while hillslope landslide results are highly satisfactory (average AUCs of about 0.8). The stochastic approach makes it possible to classify the Sicilian territory depending on its propensity to landslides in order to identify those municipalities which are most susceptible at this level of study, and are potentially worthy of more specific studies, as required by European-level protocols.

**Keywords:** landslide susceptibility assessment; forward logistic regression; forward conditional analysis; GIS; Sicily

## 1. Introduction

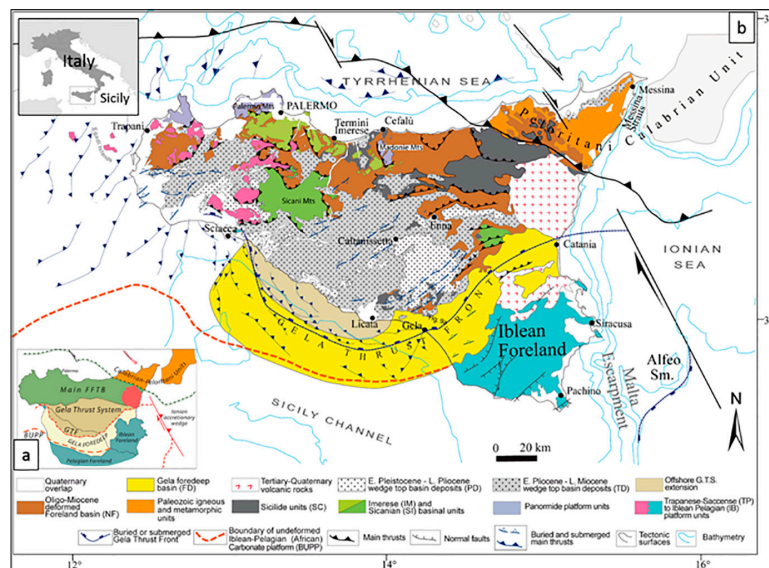
One of the most obvious effects of rapid territorial expansion in recent decades is the growing impact that natural disasters have on man and his activities. Therefore, institutions are committed to investing their resources in both the implementation of structural interventions to mitigate risk as well as early warning systems, and in defining guidelines for land management and civil protection issues [1,2]. Landslides are among the major contributors to the dynamics of the morphological evolution of slopes and occur when slope stability conditions change due to increased stress or decreased resistance along a failure surface. A deformation occurring on a pre-existing failure surface is called re-activation, whilst one along a new fracture plane is referred to as a neo-activation. International literature refers to landslide susceptibility as the spatial probability for gravitational instability conditions within an area, based on its physical-environmental conditions [3,4]. Therefore, depending on the spatial variability of the physical-environmental features of the study area (typically, a river-basin or an administrative

territorial unit), a landslide susceptibility map allows the units into which it is subdivided to be differentiated, according to the higher or lower probability of a landslide occurring. Most approaches and methods were designed to evaluate landslide susceptibility based on the identification and spatial characterization of a set of control factors, and on quantifying the relationships between these and an archive of past landslides. This concept is a reinterpretation of the actualism principle [5] from a geomorphological point of view: the past and the present are the keys to the future [1,6–8]. This principle suggests that the areas affected by landslides in the future will be those that share similar characteristics to those already recorded in past landslides [7,8]. One of the first aspects to be faced in planning research stages aiming to define a set of susceptibility conditions of an area is to establish the scale and approach at which the analysis should be performed. Time- and economic resources-allowing geostatistical approaches are often the most suitable for areas of hundreds or thousands of square kilometers [4,9]. Landslide susceptibility assessment poses specific methodological issues when performed for regional mapping purposes [2]. Indeed, within regional applications, overall forecasting performance is seriously hindered by the lack of data required or the inaccuracy of these data: landslide inventories and thematic maps of the controlling factors. Likewise, no matter what the resolution of the processed data is, some basic issues in model building procedures, such as the modelling approach, mapping units, landslide classification, and representation and validation strategies [1], need to be optimized when applied to regional multi-scale assessment procedures. Thus, an expert European landslide group has recently proposed some criteria for a multi-level method (TIER [10]) in order to define shared approaches to landslide susceptibility mapping. Three susceptibility TIER levels have been proposed and reference data and model building procedures have been recommended for each one. The TIER approach is strictly dependent on the quality of the available data in landslide inventories and on thematic maps, required for the whole European territory. This protocol consists of three nested levels (TIER1, TIER2, and TIER3) with a gradually increasing degree of resolution for the predictive models [10–13]. In a nested tiered approach, when the scale of work changes, the resolution of the factors, the mapping units, as well as the complexity and the type of the techniques used, may change too. The TIERS protocol features a high degree of objectivity and spatial and/or time repeatability, which is why it represents an important safety benchmark for national and regional administrations. For small-scale studies (smaller than 1:100,000), the methods recommended by the Joint Research Centre (JRC research group) are those based on expert-driven approaches such as conditional analysis, heuristics, and/or weighted factors. In this piece of research, we have had the opportunity to verify and compare the use of the binary logistic regression (BLR) statistical technique on extensive areas, in the range of tens of thousands of square kilometers, contrary to suggestions by European Commission experts (Tier-based approach). We have then compared results with those obtained through the conditional analysis (CA) approach and with those obtained by other authors researching the same area [10,13,14]. Susceptibility scenarios described by the maps are also compared and homogenized with those hazards arising from the PAI program [15], working towards a punctual and detailed analysis of all possible discrepancies that may result. Further to validation, and once the robustness of the scientific guidelines testing has been tested, the skills and the experience acquired may be used as a basis for the drafting of a Sicilian municipality susceptibility map, a useful tool for the policy-makers dealing with land management.

## 2. Study Area

The geological setting of Sicily (Figure 1) consists of three main structural elements [16,17]: a range sector, running along the northern strip of the island, from the Peloritan Peaks to the Aegadian Islands, where Triassic to Mesozoic structural–stratigraphic units, mainly characterized by carbonate (at the base) to clayey (at the top) formations, tectonically overlap; a northwest-dipping foredeep area lying in the mid-southern area of the Sicilian territory, consisting of Plio-Pleistocene pelagic marly limestones, silty mudstones, and sandy clays overlying Messinian evaporites separated by the Sikan Peaks; the Hyblean Plateau located in southeastern Sicily, made up of a Triassic-Liassic platform and scarp-basin

carbonates overlaid by Jurassic-Eocene pelagic carbonates and Tertiary open-shelf clastic deposits; southern and central Sicily feature Cretaceous-Lower Pleistocene clastic-terrigenous deposits and Messinian evaporites [18].



**Figure 1.** Schematic structural map of Sicily [17]. (a) The three main geological structural elements of Sicily; (b) distribution of the main stratigraphic–structural units.

The mechanical characteristics of outcropping lithotypes are among the main geo-environmental factors directly influencing the geomorphological stability of Sicilian slopes. Slides and flows are mainly located where the clay continental sequences outcrop. Hard block outcrops (metamorphic and carbonate) are affected by falls, topples, and lateral spreading. The rainfall may cause denudation slope actions able to generate, in the presence of debris on the metamorphic units, debris flows or debris avalanches. The triggering factors which are most likely to influence the activation of landslides in Sicily, are, in order: rainfall, human activity, volcanic eruptions, and earthquakes. Carbonatic rocks outcropping in the foreland sector are almost exclusively affected by falls. Both the ductile clayey formations and the weathered top coverage of the metamorphic units may experience rapid debris avalanche/debris flow phenomena.

The geomorphological features of the island of Sicily have been crafted by the collision of the Eurasian and African plates acting in synergy to shape the current landscape. The topography is directly influenced by the stratigraphic structure of the area and by the surface uplifting and subduction that occurred during the quaternary period, influenced heavily by significant eustatic sea level variations. The changes to Sicily's geology through the epochs have led to a mountainous and hilly terrain.

Furthermore, the influence of medieval human civilization is still apparent in the small Sicilian urban centers. These historic centers are often surrounded by harsh terrain, where slopes with a gradient of 50% or more can often be seen encompassing the centers, features which would have been favorable to the population who first settled there when defending their land. This, however, leads to limited space and opportunity for further urban development.

The flat areas of the island, a total of just 7% of the entire territory, are represented by the alluvial plain of Catania, the coastal plain of Licata and Gela, the coastal area of Trapanese, and that between Syracuse and Scicli, at the foot of the Monti Iblei.

From analysis of the rainfall data of the Sicilian stations, it is possible to highlight how the rainfall is concentrated, especially in the October–March period, though it is somewhat appreciable in the spring (April–May) and of little importance in the summer months. The maximum and minimum values of average annual precipitation are respectively 919 mm in the Monreale station, in the province of Palermo, and 510.6 mm in the Agrigento station. Frequently, at the highest peaks, there is snowfall,

of which there is no significant long-term data, due largely to the lack of a sufficient number of snow stations. A meteoric event of considerable importance in the northern mountain ranges, in particular for the northern slopes of the Madonie due to the presence of fog, which, in addition to integrating normal water supplies through condensation, performs a mitigating and compensating action for extreme climatic events, limiting precipitation and keeping temperatures lower during periods of summer water deficit, as well as decreasing the intensity of weather events harmful to plants, such as late frosts. As regards the thermometric data, there is an inverse trend compared with that of rainfall, as occurs throughout the Mediterranean region. There is, in fact, a gradual increase between March and April, and a more marked increase from May to July–August—a period in which the absolute maximum values are reached—beyond which the temperatures gradually decrease until October, and then drop sharply until December and touch the minimum values in January–February which is the coldest time of the year. The highest average annual temperature is 18.8 °C for the Cefalù (Palermo) station, while the lowest, 13.3 °C, is recorded in Petralia Sottana (Palermo). The lowest average annual minimum temperature value (9.3 °C) is recorded in the Petralia Sottana station, while the highest annual average maximum temperature value (24.1 °C) in the Lentini (Siracusa) and Palermo stations (Castelnuovo Institute).

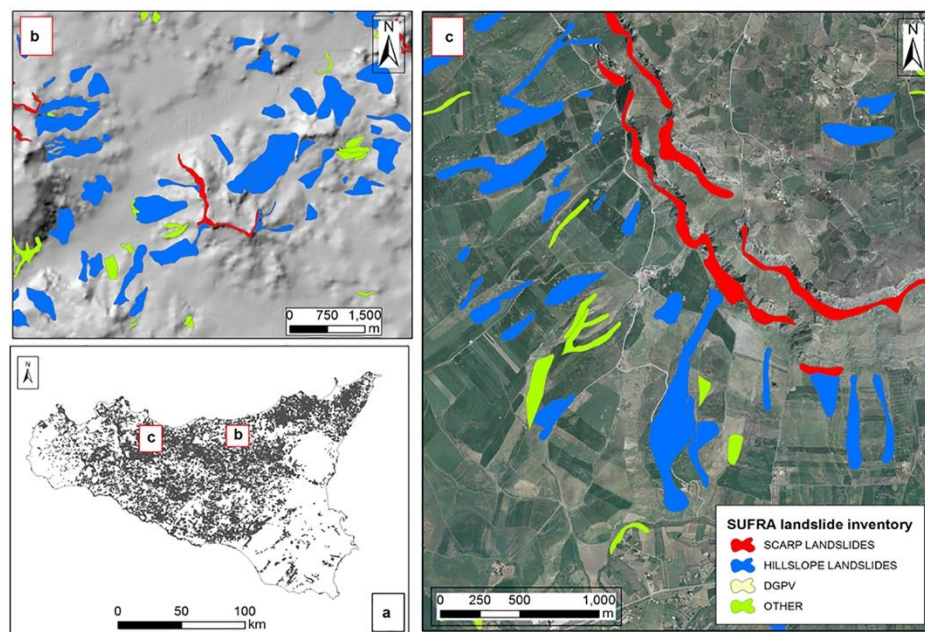
### 3. Data Collection and Processing

#### 3.1. Landslide Inventory

Landslides are natural events in the evolution of a slope. They are a problem and become a hazard and/or risk when they interact with man and the man-made environment. Landslides can be classified according to their movement types and the nature of the displaced material type, as well as the state, distribution, and style of their activity [8,19–21]. As raised by [22], there is a conceptual ambiguity concerning landslides stemming from the use of the very same term (i.e., landslide) for both the landslide deposit (displacement volume) and the movement of material on a slope or a pre-existing landslide body [20,23]. This is in addition to general confusion originating from the variable and complex nature of the phenomenon itself [24], due to profoundly different morphological characteristics, behavior, activity states, and their evolution. The construction of the landslide inventory is a fundamental and critical step towards the application of statistical models designed to estimate the probability of new activations affecting previously uninvestigated areas. A landslide inventory commonly represents the sum of all the events that have occurred in an area. Alterations to a slope profile, pointing to ongoing landslides, tend, over time, to become less evident because of erosion, new landslides, human activity, and vegetation, making the “in landslide/not in landslide” border hard to detect with the passage of time. Generally, “newer” phenomena generated by recent heavy rainfall or earthquakes are more easily identifiable and interpretable than more remote ones, where diagnostic elements begin to dissolve. Certainly, there is a critical issue concerning the updating of the landslide inventory, as the public administration lacks economic and human resources. This shows, even more clearly, the limitations of PAI methodology in mapping areas at risk, as it leads to a zoning system, which is not closely related to all new activations. No systematic archive of slope instability exists for the research area of this paper. The latest archive of slope failures is the one belonging to the PAI project, which now holds 33,094 landslides (latest update May 2016) classified into 11 different types. The landslides that were mapped within the PAI project affect an area of approximately 1300 square kilometers, approximately 5.1% of the size of the region (Table 1; Figure 2). The landslide archive derives from a historical inventory and territorial analysis, completed with the execution of inspections, which began in 2003. The archive is continuously updated by reports from the following: literature and scientific publications; studies to support urban development projects in municipalities; regional civil protection archives; and reports from local authorities to regional and national bodies (Civil Protection, Territory and Environment Departments) of geomorphological phenomena that have occurred (from 1998 to today).

**Table 1.** Landslide inventory, statistics, and their reclassification.

PAI TYPOLOGY	Number of Cases	Area [m <sup>2</sup> ] for a Single Landslide				Total (km <sup>2</sup> )	Percentages		SUFRA		
		max	min	mean	Std. Dev.		Landslide Area	Total Area	Number of Cases	Area (km <sup>2</sup> )	TPOLOGY
1. Falls/Topples	5460	1,630,259	25	15,310	1,138,586	84	6.4	0.3	5460	84	1. Scarp Landslide
2. Rapid flows	853	1,152,779	79	15,403	48,938	13	1.0	0.1	10,202	435	2. Hillslope landslide
4. Slides	2835	45,106	530	11,350	11,987	81	6.2	0.3			
5. Complex	3076	735,524	766	50,290	114,782	215	16.5	0.8			
7. Slow flow	3438	194,381	869	34,263	41,727	125	9.6	0.5			
6. DPGV/Spreads	28	7,891,513	212	509,802	509,802	14	1.1	0.1	28	14	3. DPGV and Spreads
3. Sinkhole	43	1,380,425	25	51,918	219,976	2	0.2	0.0	17,404	772	4. Others
8. Areas with diffused landslides	2877	225,203	1482	53,560	225,203	288	22.1	1.1			
9. Slowly surface deformation	3512	124,610	605	23,534	29,059	155	11.9	0.6			
10. Badlands	1266	1,625,318	3471	42,755	302,626	54	4.1	0.2			
11. Caused by accel. erosion	9706	254,775	9358	30,171	57,136	273	20.9	1.1	33,094	1305	Total
Total	33,094				2,699,822	1305	100.0	5.1			
Area of Sicily 25.832 Km <sup>2</sup>											



**Figure 2.** PAI landslide inventory. (a) Regional distribution of landslides; (b) detail of a portion of the territory of landslides on the hillshade grid; (c) PAI landslides on orthophotos.

The PAI archive focuses more on urban and populated areas. In this piece of research, various models were created to record landslides affecting scarps (scarp landslides; SCR\_LSN) and landslides lying over slopes (hillslope landslides; HILL\_LSN).

### 3.2. Modelling Approach

A predictive model must represent the response of a natural system to the trigger conditions described by its environmental features; the response may lie in the spatial distribution of new landslides or in the so-called prediction image. The effectiveness of the model can be measured by comparing the final expected results (the susceptibility map) with the actual results that are directly observed (the new landslide map). When studying the natural environment, the use of models is essential, as they allow for a simplification of the infinite natural variables, as well as operating within an acceptable processing time with conventional computers. Thanks to a greater exchange of information, the development of the hardware and software component for the acquisition and processing of data, and the interaction between different research groups, the methods for assessing susceptibility from landslides have evolved rapidly [3,4,24–27]. In the last decade, several applications have been carried out with the aim of comparing results from different statistical approaches, as applied to the implementation of models of landslide susceptibility on the same area, using the same landslide inventory and the same control factors [28–31]. Although full agreement does not exist within the scientific community as to what the best approach to follow is, the experience gained through our previous studies tells us which statistical techniques may be adopted from among the many available (discriminant analysis, conditional analysis, binary logistic regression, classification, and regression trees) and which one leads to greater performance in terms of predictive fitting, and robustness [27,32–39].

### 3.3. Conditional Analysis (CA)

The CA statistical approach has been widely adopted in landslide susceptibility assessment [24,26,33,40–44]. The CA concept is simple and easily manageable in a GIS environment. That is why it is the single most recommended method when investigating the landslide susceptibility conditions of larger areas and at scales greater than 1:100,000.

The CA method is based on the concept that landslide density, computed on homogenous domains, represents the relational function between the environmental variables and the framework of a landslide area [2,26,45].

According to the CA concept, we must compute the density of cells in an unstable area for unique condition units (UCUs), defined by overlapping and combining a set of selected control factors and intersecting the UCU and landslide layers. From a statistical point of view, the landslide density corresponds to the susceptibility value of an area in a landslide, linked to a particular combination of control factors. Mathematically, this concept can be expressed as

$$P = \delta UCU(i) = \frac{UCU_{unst}(i)}{UCU(i)} = \frac{\left(\frac{UCU_{unst}(i)}{UCU_{unst}}\right) * \left(\frac{UCU_{unst}}{UCU_{all}}\right)}{\left(\frac{UCU(i)}{UCU_{all}}\right)} \quad (1)$$

where probability (P) can be computed as the ratio between unstable (UCU<sub>unst</sub>) and total counts (UCU<sub>all</sub>) of cells for the cells which have a value.

Landslide density is the relational function between the environmental conditions and the landslides in a given area, meaning that a density ranking order corresponds to a scale of landslide susceptibility [32,33,46].

#### 3.4. Binary Logistic Regression (BLR)

One of the key points in determining the susceptibility conditions of an area with multivariate statistical techniques is the selection of an appropriate number of factors that can justify the spatial distribution of past and future forms of instability. In fact, many of these techniques provide an estimate of the importance of each factor in relation to the others, or its specific contribution in generating a particular type of landslide in the area under investigation. Many of these techniques rank, through hierarchization, the contribution of each factor in determining the landslide-specificity of an area by identifying the minimum and maximum number of factors needed, beyond which the performance variation of the model may be defined as insignificant or even negative [47,48].

According to Hosmer and Lemeshow [49], the aim of an approach based on binary logistic regression (BLR) is to enable the singling out of the best linear relationship between a dichotomous dependent variable (such as 1 or 0 representing the “presence”/“absence” of landslides, respectively) and a set of independent variables representing control geo-environmental factors. In the logistic regression equation, the expected dependent variable  $f(x)$  may be expressed as

$$\text{logit}(x) = \ln(\text{odds}) = \ln\left[\frac{\pi}{1-\pi}\right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (2)$$

where  $\text{logit}(x)$  corresponds to a natural logarithm of odds  $[\pi/(1-\pi)]$ , expressed as a ratio between the likelihood of the presence of landslides ( $\pi$ ) over the likelihood of their absence ( $1-\pi$ );  $\alpha$  is the intercept of the model; and  $\beta_1, \beta_2$  up to  $\beta_n$  are the coefficients, which measure the contribution of each independent input variable [50–52].

In other words, the BLR allows us to estimate the contribution of each input variable using its coefficient in the probability equation by adopting the maximum likelihood classifier concept. Comparing the maximum likelihood computed for every  $\beta$ -value with each estimated error, the significance of the coefficients is tested using the Wald test [31,49,53]. The probability of occurrence can be estimated by multiplying the  $-2 \log$ -likelihood ratio; thus, the negative log-likelihood ( $-2LL$ ) statistic is obtained, which has a chi-square ( $\chi^2$ ) distribution. According to this approach, BLR is executed through a stepwise procedure that allows the differentiation of only those predictor variables with a significant impact on the performance of the multivariate model [36,54,55].

A problem arising when using BLR is that which is related to choosing the appropriate sample size for an unstable area (positive cases) as opposed to stable ones (negative cases). Indeed, the number

of positive cases is significantly lower than negative ones, thus generating an imbalance. Therefore, it is one of the most important choices concerning the random selection technique likely to lead to a balanced dataset of landslide and non-landslide pixels [56]. In this piece of research, the TANAGRA open-source software was used to apply forward stepwise logistic regression [57].

### 3.5. Variables Selection and Factor Class Definition

Landslides are directly connected to the many environmental condition types, mainly depending on slope morphology (slope angle, orientation slope, curvature, elevation, roughness, etc.) and other intrinsic characteristics (lithology, soil use, tectonic condition, landform classification, etc.), while the activation of new landslides depends on trigger factors, such as intensive rainfall or earthquakes [58–60].

All kinds of geo-environmental factors may be considered as potential predictive factors and can be introduced into the analysis. All the factors should be drawn at a specific measurement for classification (categorical, continuous). Acquiring each factor requires time and money. In addition, we must also consider the computing time needed to process huge volumes of data. This is why only those variables thought to be directly and/or indirectly capable of conditioning the established of a slope, thanks to knowledge acquired from previous studies, may enter the analysis. The use and reclassification of a limited number of factors also prevents the generation of a large number of not very widespread mapping units and thus the overestimation of the density for untrained mapping units.

The CA method, requires preselecting the factors to be entered into the analysis, and then defining the UCUs representing the mapping unit in this type of analysis. The univariate approach was followed for each acquired variable and their density calculated for the different types of landslide analyzed (scarp and hillslope landslide). The univariate approach was followed for each acquired variable and their density calculated for the different types of landslide analyzed (scarp and hillslope landslide). Univariate analysis allows to distinguish the significant from insignificant variables, and then combine only those variables among them which are likely to be the most influential in the production of the mapping units GRID layer. The following continuous and categorical variables have been acquired and analyzed (Table 2).

**Table 2.** Geo-environmental factors and their spatial distribution.

Categorical Variables				
Variable	References	Description	Code	Percentage Distribution (%)
Soil Use	(Corine Land Cover project, 2006)	Continuous urban fabric	USE_111	1.94
		Discontinuous urban fabric	USE_112	2.42
		Transitional areas	USE_13	0.27
		Green urban areas	USE_14	0.06
		Arable land	USE_21	32.54
		Permanent crops	USE_22	21.61
		Heterogeneous agricultural areas	USE_23	14.97
		Forest	USE_31	7.79
		Shrub and/or herbaceous associations	USE_32	17.31
		Open spaces with little or no vegetation	USE_33	0.65
		Water bodies	USE_51	0.44



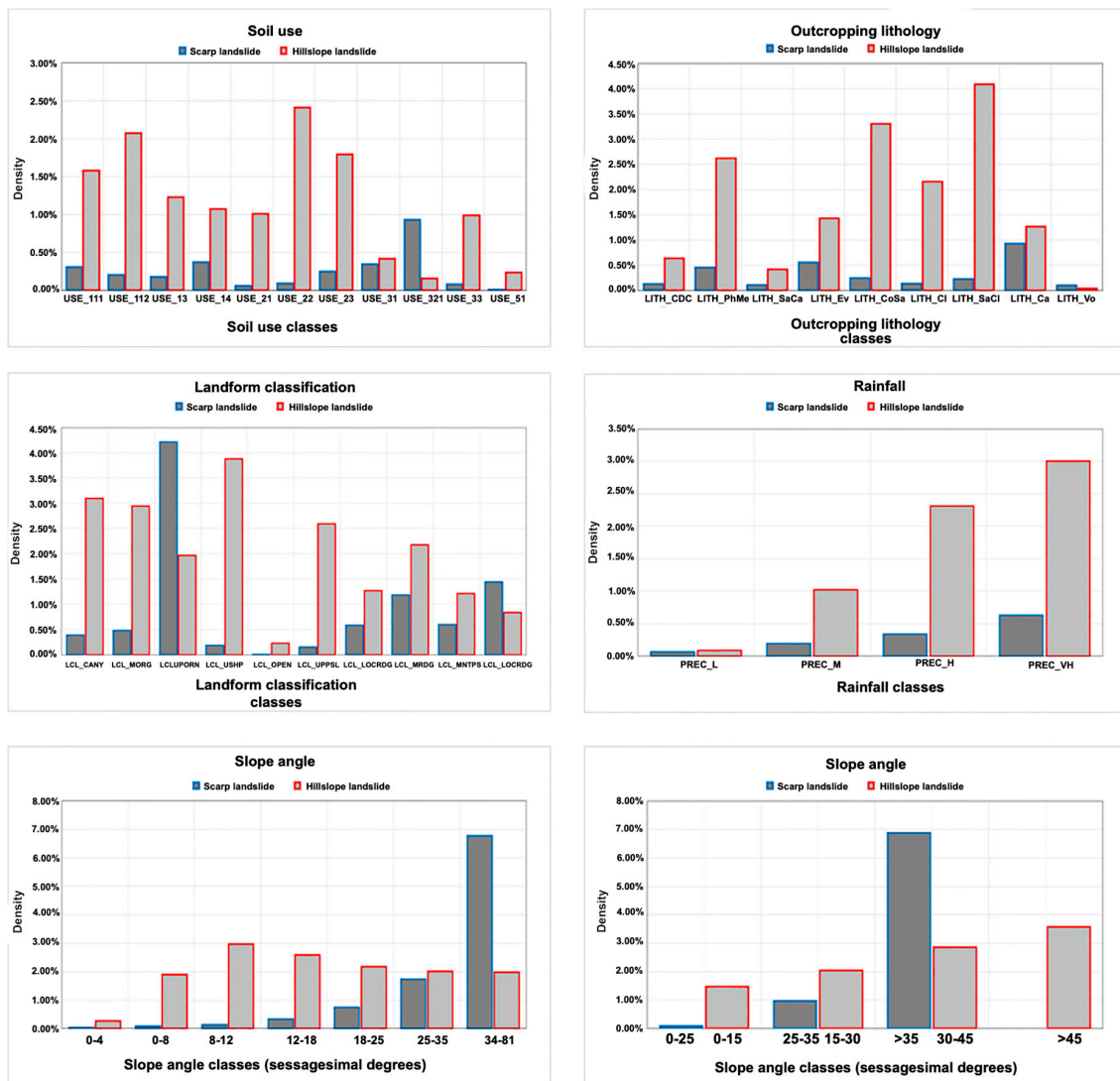
Table 2. Cont.

Categorical Variables				
Variable	References	Description	Code	Percentage Distribution (%)
<b>Outcropping lithology</b>	Lithological Complex	Continental clastic deposition complex	LITH_CDC	12.94
		Phyllitic and metamorphic complex	LITH_PhMe	3.51
		Sandy-calcarenitic complex	LITH_SaCa	13.22
		Evapotitic complex	LITH_Ev	4.86
		Conglomerate-sandstone	LITH_CoSa	2.74
		Clay complex	LITH_CI	34.13
		Sandstone and clay	LITH_SaCl	8.66
		Carbonatic complex	LITH_Ca	13.41
		Volcanic complex	LITH_Vo	6.53
<b>Landform classification</b>	Landform Classification (Weiss, 2001)	Canyons	LCL_CANY	7.51
		Midslope drainage	LCL_MDRG	2.27
		Upland drainage	LCL_UPDRN	2.32
		U-shaped valleys	LCL_USHP	2.92
		Plains	LCL_PLAINS	32.58
		Open slopes	LCL_OPEN	39.51
		Upper slope	LCL_UPPSL	2.80
		Local ridge	LCL_LOCRDG	0.00
		Midslope ridge	LCL_MRDG	2.00
		Mountain tops	LCL_MNTPS	8.07
<b>Rainfall (mm)</b>	SIAS, 2015	0–450	RAIN_L	1.58
		450–600	RAIN_M	63.32
		600–800	RAIN_H	20.75
		>800	RAIN_VH	14.34
<b>Slope Angle (Scarp landslide)</b>	$\Theta = \text{TAN } \Delta y / \Delta x$	Canyons	SLO_L	87.81
		Midslope drainage	SLO_M	10.72
		Upland drainage	SLO_H	1.47
<b>Slope Angle (Hillslope landslide)</b>	$\Theta = \text{TAN } \Delta y / \Delta x$	U-shaped valleys	SLO_L	78.09
		Plains	SLO_M	15.28
		Open slopes	SLO_H	6.43
		Upper slope	SLO_VH	0.20

### 3.5.1. Continuous Variables

The 2-m ARTA-DEM was used and resampled to generate the digital elevation model (DEM). Slope angle, slope aspect, and landform classification were calculated using the DEM.

- Slope angle (SLO) is usually considered as one of the main controlling factors in landslide modelling. At first, SLO was classified into 5 natural break intervals [14], expressed in sexagesimal degrees ( $0^{\circ}$ – $5^{\circ}$ ;  $5^{\circ}$ – $12^{\circ}$ ;  $12^{\circ}$ – $18^{\circ}$ – $18^{\circ}$ – $32^{\circ}$ ;  $> 32^{\circ}$ ). The raster-file of the slope angle was obtained by resampling the 2-meter resolution ARTA-DTM flight ATA (2007/2009) to 100 m per side. As shown in Figure 3, the proposed reclassification for the slope angle for the hillslope landslide does not reveal the theoretical concept for the slope increase, which corresponds to an increase in the likelihood of landslides occurring. This does not happen with the scarp landslide, where increasing the slope angle leads to an increase in the percentage of landslides.



**Figure 3.** Univariate analysis for the continuous variables. **Soil use.** USE\_111: continuous urban fabric; USE\_112: discontinuous urban fabric; USE\_13: transitional areas; USE\_14: green urban areas; USE\_21: arable land; USE\_22: permanent crops; USE\_23: heterogeneous agricultural areas; USE\_31: USE\_321: shrub and/or herbaceous associations; USE\_33: open spaces with little or no vegetation; USE\_51: water bodies. **Landform classification.** LCL\_CANY: canyons; LCL\_MDRG: midslope drainage; LCL\_UPDRN: upland drainage; LCL\_USHP: U-shaped valleys; LCL\_PLAINS: plains; LCL\_OPEN: open slopes; LCL\_UPPSL: upper slope; LCL\_LOCRDG: local ridge; LCL\_MRDG: midslope ridge; LCL\_MNTPS: mountain tops. **Rainfall (mm).** PREC\_L: 0–450; PREC\_M: 450–600; PREC\_H: 600–800; PREC\_VH: >800. **Slope Angle (scarp landslide).** SLO\_L: 0°–25°; SLO\_M: 25°–35°; SLO\_H: >35°; **Slope Angle (hillslope landslide).** SLO\_L: 0°–15°; SLO\_M: 15°–30°; SLO\_H: 30°–45°; SLO\_VH: >45°.

### 3.5.2. Category Variables

- Landform classification (LCL). Using an ArcMap open source tool, the LCL variable was derived directly from the DEM. LCL provides a simple and repeatable method to classify the landscape into slope position and landform category comparison. The different landform category classes can be determined by classifying the combination of a small and large neighborhood topographic position index (TPI) computed for each cell from different scales. The TPI is simply the difference between a cell elevation value and the average elevation of the neighborhood around that cell. Positive values mean the cell is higher than its surroundings, while negative values mean it is

lower [61]. To compute the LCL, the small and the large neighborhood areas were set to 500 and 100 m, respectively. Ten landform classes were thus obtained (Table 2);

- Outcropping lithology (LITH). Together with the slope itself, the lithological conditions of an area are the most important factors influencing the geomorphological processes on the slope. The lithology controls the response of the slope in terms of the trigger-time of the collapse because of rainfall or seismic forces and evolution of the process. The lithotypes cropping out in the map of the Sicilian region were used in this research and grouped into 9 different “lithological complexes”, according to their geotechnical characteristics. The output lithological complexes were named as shown in Table 2. The clay complex is the most widespread one in the Sicilian territory, as it crops out in almost 35% of the area (more than 50,000 hectares);
- Soil use (USE). In this test, we used a soil use map derived from the 1:100,000 Corine Land Cover project (2006) based on a revised version of the Corine Land Cover 2000 dataset with the results of Landsat 1988 and photointerpretation of aerial photos. Table 2 shows land cover characteristics in 11 different classes, for terrain units larger than 0.25 km<sup>2</sup>. The Corine 2006 map was converted into a soil cover digital map provided by the Sicilian region, using the second level of the Corine legend, except for the “urban areas” class, which has been divided into continuous (USE\_111) and discontinuous (USE\_112) urban fabric, corresponding to level III of the Corine Land Cover classification. Arable land (USE\_21) covers more than 30% of the research area. Forest crops cover about 7% of the area and mainly appear in the northeastern sectors. Areas covered by shrubby and herbaceous vegetation associations are dispersed around the study area: they cover 17%. Urban area cover is only 4.5%;
- Rainfall (RAIN). For rainfall, 280 rainfall stations were used to create the GRID rainfall map using the inverse distance weighted method. The database from the Sicilian regional administration office (<http://www.osservatoriodelleacque.it>) was used to extract the mean annual precipitation (for the period 1921–2009). Regarding precipitation, Sicily can be divided into three main sectors with three different pluviometric regimes: the northern sector: includes all the Tyrrhenian coast of the island. Rainfall here is characterized by a rainy season (autumn–winter) and a dry spring and summer. Eastern Sicily: in this area, rainfall is also greater in winter. Precipitation is often concentrated into short spells and is sometimes very violent. This is because the precipitation depression bearers come from Africa and are very hot and humid, favoring strong thermal contrasts. Southern Sicily: includes all the area bordered by the Mediterranean Sea. As in the rest of the island, winter is the rainy season. The number of rainy days is less than in the northern area (<60 days per year). In some areas, rainfall is sparse, especially in the coastal zone. The areas with the highest rainfall are the Madonie, Nebrodi, and Peloritani peaks, Etna, and the area south of Palermo. The driest areas are the Plain of Catania and the southern coast, in particular, Gela city.

Generally, there are two reasons that lead us to choose the smallest possible number of geo-environmental variables for the construction of the forecasting model, which allow the realization of what is called the “best model” for a specific area, capable of providing an acceptable performance forecast. On the one hand, achieving or obtaining each parameter requires spending a considerable amount of time and money, on the other, a large number of environmental variables results in a large number of possible combinations characterizing each of the territorial units chosen in an excessively specific manner, as the basis for statistical analysis. A high number of combinations bring a progressive decrease in the distribution of each combination class. The consequence is an unexpected decrease in the performance of the susceptibility model caused by the inclusion of variables which are closely related to a small number of cells, but poorly correlated to the global distribution of the remaining part, thus affecting the choice of the most predictive variables. Even the selection of factors is an essential step in landslide susceptibility assessment procedures, in which the nature of geomorphological criteria takes priority.

An expert-driven univariate analysis procedure for the following 5 processing control variables was carried out (Figure 3). Depending on the type of landslides, a maximum moderation criterion in

the number of factors used is necessary to identify a first set of control factors that can be justified based on morphodynamic models, defined as heuristic at first approximation, for the distribution of the observed phenomena. Then, regression techniques may be applied to highlight the actual role played by each of the geo-environmental variables considered. However, it is also common practice, and recommended by applied statistics handbooks, to maintain certain diagnostic values of the variables (i.e., slope), even when stepwise regression procedures have greatly reduced their influence. A type of approach which, in a certain way, is the opposite of the analytical-geomorphological way, relying on deterministic or physically based methods.

That is why, for example, for “Slope angle”, values were reclassified into 4 different classes (for hillslope landslide) and 3 classes (landslide scarp). At this point, a new univariate analysis was carried out and the results are shown in Figure 3.

#### 4. Model-Building Strategy

The five reclassified geo-environmental variables were combined in unique conditions units (UCUs) aiming to get all the possible combinations of the classes of the different factors in order to assess the landslide susceptibility. The UCU layers were laid over with the landslide ones (SCR\_LSN and HILL\_LSN, respectively). The landslide densities were computed for each UCU class combination, as the ratio between landslide area counts and total pixel counts: according to Bayes’ theorem [26,45], these values express the conditional probability of landslide occurrence, given a factor condition. The variables were combined and a UCU layer was derived from these. The combination produced a large number of combination classes (8355 for the HILL\_LSN; 4173 for the SCR\_LSN). The CA method requires that the combination of output factors has total areas which are large enough to guarantee the statistical significance of the “observed” sample. Table 3 shows the combination classes of the most susceptible UCUs. The UCUs found to have the highest susceptibility values are those characterized by predominantly consistent lithologies in SCR\_LSN, (LITH\_Ca), LCL\_MNTPS (as landform classification), and high and very high slope angle values. Though uncommon, these UCU combinations have a susceptibility value of 100%. For hillslope landslides, clayey lithologies and open space LCLs are the most widespread combinations capable of generating slope failures (Table 3).

**Table 3.** Most susceptible unique conditions units (UCUs) in the target area.

Most diffused UCUs for SCR_LSN							
UCU Code	Area (Ha)	LCL	LITH	USE	SLO	RAIN	$\delta$
1502	2	LCL_MNTPS	LITH_Ca	USE_112	SLO_VH	RAIN_VH	100.00%
1515	3	LCL_UPDRN	LITH_Ca	USE_23	SLO_VH	RAIN_VH	100.00%
1204	8	LCL_USHP	LITH_SaCa	USE_13	SLO_H	RAIN_H	100.00%
1517	3	LCL_MNTPS	LITH_Ca	USE_13	SLO_H	RAIN_L	100.00%
1217	3	LCL_MNTPS	LITH_Ca	USE_13	SLO_VH	RAIN_VH	100.00%
1001	2	LCL_UPDRN	LITH_SaCa	USE_13	SLO_H	RAIN_VH	100.00%
4095	1	LCL_USHP	LITH_CoSa	USE_13	SLO_M	RAIN_H	100.00%
1075	2	LCL_UPDRN	LITH_Ca	USE_13	SLO_VH	RAIN_M	100.00%
3403	1	LCL_USHP	LITH_Ca	USE_13	SLO_VH	RAIN_M	100.00%
3850	3	LCL_UPDRN	LITH_CoSa	USE_112	SLO_H	RAIN_VH	100.00%
Most diffused UCUs for HILL_LSN							
UCU Code	Area (Ha)	LCL	LITH	USE	SLO	RAIN	$\delta$
2778	14	LCL_OPEN	LITH_CI	USE_13	SLO_VH	RAIN_VH	100.00%
3230	11	LCL_UPDRN	LITH_CI	USE_21	SLO_H	RAIN_H	89.00%
2759	34	LCL_OPEN	LITH_CI	USE_21	SLO_L	RAIN_M	62.00%
2777	17	LCL_OPEN	LITH_CI	USE_13	SLO_VH	RAIN_M	57.00%
3370	13	LCL_UPDRN	LITH_CI	USE_32	SLO_VH	RAIN_VH	36.00%
2711	45	LCL_OPEN	LITH_CI	USE_13	SLO_H	RAIN_VH	31.00%
4376	14	LCL_OPEN	LITH_SaCa	USE_21	SLO_H	RAIN_H	25.00%
2735	32	LCL_OPEN	LITH_CI	USE_13	SLO_M	RAIN_L	24.00%
1272	14	LCL_UPDRN	LITH_SaCa	USE_21	SLO_M	RAIN_VH	24.00%
3381	7	LCL_USHP	LITH_CI	USE_32	SLO_VH	RAIN_VH	23.00%

## 5. Results and Validation

Regardless of the statistical approach, models and susceptibility maps should always be validated to test the forecasting performance of the proposed protocol. Generally, a statistical validation procedure correlates the performance forecast (the expected landslides) that represents the space distribution of the model created by using a part of the observed phenomena (training dataset) with the distribution of a set of landslides which are not used in model-building (test dataset). This test dataset should therefore be temporally or spatially different from the training dataset used for model building, but the regional administration does not provide a systematic and continuous mapping of new landslides which occur reasonably often and in all Italian regions and worldwide. Many validation approaches have been presented in the last two decades in the relevant literature. In this paper, the random partition strategy was used, based on the random partition of the slope failures in two balanced populated training and test subsets to compare and estimate the robustness and reliability of the susceptibility models proposed.

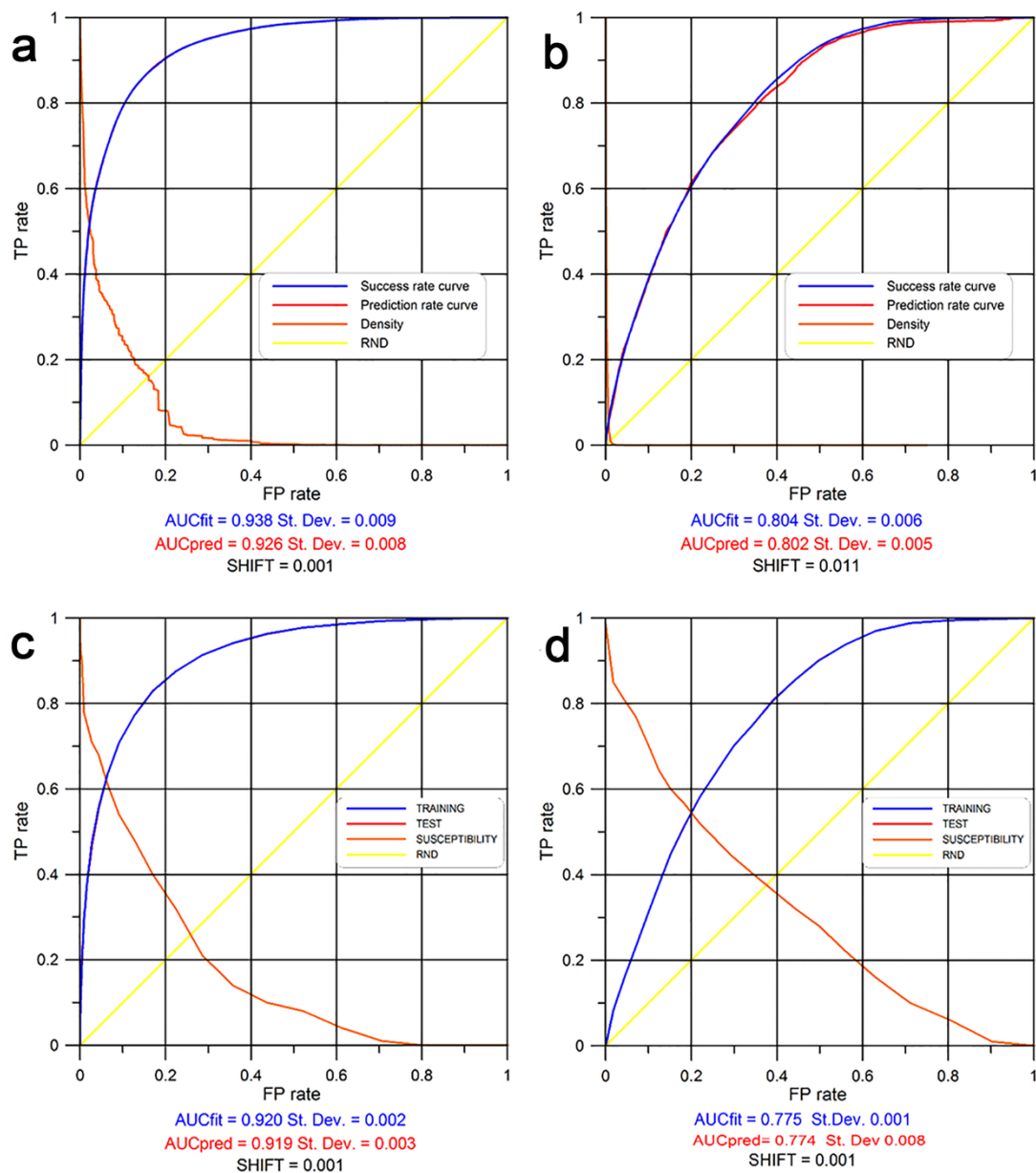
The models presented in this research were derived using the CA and BLR statistical approaches, as the result of the average of 100 different replicas obtained by splitting the dependent variable, in 75% TRN\_Subset and 25% TST\_Subset randomly 100 times.

A reliable approach to test the performance of susceptibility models is the receiver operating characteristic (ROC) curve, allowing the comparison of the predictive performance models. The ROC curve is a plot of the probability expressing the sensitivity (TP rate) that represents the area classified correctly as susceptible (x-axis) versus the 1-specificity (FP rate), representing the probability of false prediction in response to an event for all the cutoff probability values (y-axis). ROC curve analysis allows the differentiation between two classes of events: unstable and stable cells [62]. The quantitative measure of model performance can be tested by computing the area under the curve (AUC) ranging from 0 to 1 [63]. The closer the AUC values are to 1, the higher the predictive performance of the model will be, while the closer the values are to 0.5 (random performance) the higher the inaccuracy of the model will be [50,64,65]. A value equal to 1 denotes a perfect discrimination between positive and negative cases. ROC curves were created for each of the two different statistical approaches (CA and BLR). In Figure 4, the ROC curves for both CA and BLR models were drawn for the training and test subsets with the aim of evaluating the model fitting for both approaches.

The BLR approach has been applied for forward stepwise selection independent predictive variables [28]. As shown in Table 4, average results for the 100 different splits of dependent variables in terms of SUCU. The model suite produced for SCR\_LSN is characterized by a mean error rate of 0.17 (St.dev. 0.004) and AUC > 0.9 (outstanding). The graphs (Figure 4) show the average AUC values are close to excellent for HILL\_LSN (>0.77) according to Hosmer and Lemeshow [49].

For HILL\_LSN, the mean error rate is higher (about 0.29) but still very stable (St. dev. 0.001): a ranking of predictor variables derived from exploiting the forward stepwise statistical procedure.

Using BLR grid-cell-based models, the positive cases (unstable cells) are dramatically less than the negative cases (stable cells) and a suite of 10 different models (both for SCR\_LSN and HILL\_LSN) were prepared in order to estimate the model fitting, prediction skill, and robustness of the proposed approach [22,28,30,36,66]. Each suite model is made up of a subset of the unstable/positive cells and by an ever-changing subset which contains the stable/negative cells. The SCR\_LSN grid-cell models were prepared by merging 6674 positive cells (5% of the total area) with 10 different subset, randomly selected negative 100 × 100 m cells. As for the HILL\_LSN, suite models were created by merging 101,860 unstable cells (about 80% of the total area) and an equal number of stable cases. The negative cases in the subset were randomly selected in order to prevent stable cells from overlapping.



**Figure 4.** Receiver operating characteristic (ROC) curves: (a) conditional analysis (CA) for scarp landslide (SCR\_LSN); (b) CA for hillslope landslide (HILL\_LSN); (c) binary logistic regression (BLR) analysis for SCR\_LSN; (d) BLR analysis for HILL\_LSN.

Categorical variables were binarized and BLR was applied. Each model underwent the BLR procedure 10 times. An open source statistical package (TANAGRA) was used to generate the contingency tables (Table 4) and automatically extract the true positive (TP) and true negative (TN), and single estimates of the sensitivity or hit rate ( $TP/(TP + FN)$ ) and 1-specificity ( $FP/(TP + FN)$ ); [66–70].

For SCR\_LSN, drawing from 18 predictors, 13 were always selected in all 100 model repeats of the models. SLO, LCL\_MNTPS, and USE\_321 were systematically extracted as the first three significant factors, and with positive coefficients. The regression coefficients, for each selected predictor, were marked by conceptually coherent signs. A scan was easily imagined, and LCL\_PLAINS and LCL\_OPEN had a negative coefficient instead (Table 5).

**Table 4.** Contingency table for the 10-model suite. a for the SCR\_LSN; b for HILL\_SLN. TP = true positives, TN = true negatives, FP = false positives, FN = false negatives.

		TEST SUBSET-PREDICTION SKILL									
(a)	MODELS	PREDICTED YES		PREDICTED NO		RECALL		FALL-OUT		ERROR RATE	AUC
		YES	NO	YES	NO	YES	NO	YES	NO		
		TP/oP	TN/oN	FP/pP	FN/pN						
	1	13,648	2850	3011	13,861	0.819	0.829	0.173	0.178	0.176	0.916
	2	13,672	2853	3008	13,837	0.820	0.829	0.173	0.179	0.176	0.916
	3	13,849	2686	2831	14,004	0.830	0.839	0.162	0.168	0.165	0.921
	4	13,802	2634	2878	14,056	0.827	0.842	0.160	0.170	0.165	0.923
	5	13,686	2814	2994	13,876	0.821	0.831	0.171	0.177	0.174	0.917
	6	13,735	2743	2945	13,947	0.823	0.836	0.166	0.174	0.170	0.919
	7	13,732	2817	2948	13,873	0.823	0.831	0.170	0.175	0.173	0.917
	8	13,794	2755	2896	13,935	0.826	0.835	0.166	0.172	0.169	0.918
	9	13,790	2622	2890	14,068	0.827	0.843	0.160	0.170	0.165	0.923
	10	13,781	2688	2902	14,002	0.826	0.839	0.163	0.172	0.168	0.917
	ALL	13,748.9	2746.2	2930.3	13,945.9	0.824	0.835	0.166	0.174	0.170	0.919
		65.1	86.2	60.9	83.6	0.004	0.005	0.005	0.004	0.004	0.003
		TEST SUBSET-PREDICTION SKILL									
(b)	MODELS	PREDICTED YES		PREDICTED NO		RECALL		FALL-OUT		ERROR RATE	AUC
		YES	NO	YES	NO	YES	NO	YES	NO		
		TP/oP	TN/oN	FP/pP	FN/pN						
	1	206,596	100,179	48,054	154,471	0.811	0.607	0.327	0.237	0.291	0.776
	2	206,596	100,179	48,054	154,471	0.811	0.607	0.327	0.237	0.291	0.775
	3	207,100	100,614	47,550	154,036	0.813	0.605	0.327	0.236	0.291	0.776
	4	206,960	101,894	47,550	154,036	0.813	0.602	0.330	0.236	0.293	0.772
	5	206,275	99,734	48,375	154,916	0.810	0.608	0.326	0.238	0.291	0.776
	6	206,665	100,608	47,659	153,889	0.813	0.605	0.327	0.236	0.291	0.766
	7	206,991	100,761	47,659	153,889	0.813	0.604	0.327	0.236	0.291	0.775
	8	206,455	99,808	48,195	154,842	0.811	0.608	0.326	0.237	0.291	0.775
	9	206,755	101,015	47,895	153,635	0.812	0.603	0.328	0.238	0.292	0.773
	10	206,880	101,576	47,770	153,074	0.812	0.601	0.329	0.238	0.293	0.775
	ALL	206,727	100,636	47,876	154,125.9	0.811	0.604	0.327	0.237	0.292	0.774
		259.0	710.5	285.3	562.8	0.001	0.002	0.001	0.001	0.001	0.003

The main significant controlling factors for SCR\_LSN in the study area which showed a slope angle for LCL\_MNTPS was landform classification (positive coefficient), and LCL\_PLAINS with a negative coefficient.

On to the HILL\_LSN, the predictors extracted by BLR, are more than those extracted for SCR LSN, and above all, this slope angle variable does not appear to be as highly significant as may have been imagined. This is probably due to the fact that the diagnostic areas used here (whole landslide body area) are not so discriminant and able to determine the preparatory conditions for landslides. Additionally, it may owe something to the reasonably large cell size. The PAI landslide inventory is not characterized by high accuracy and some landslide typologies, like for “Areas with diffused landslide”, do not allow a true discrimination of the geo-environmental variables which influence the slope stability conditions. Among the variables, LCL PLAINS ranks as the first predictor extracted in the analysis with a negative coefficient. Among all the variables, 20 were systematically extracted in all 100 different repeats, lithological conditions being those selected most frequently, and with a higher weight. Among them, LITH\_Ca is characterized by having a negative coefficient value (Table 6).

Table 5. Factors analysis for SCR\_LSN.

FACTORS ANALYSIS							
SCR_LSN							
Attribute	Avg. ( $\beta$ )	Avg. (Std-dev)	Avg. Wald	Avg. Signif	Avg. OR	Avg. RK	n
SLO	0.111	0.003	1046.925	0.0000	1.118	1.0	100
LCL_MNTPS	0.822	0.084	95.372	0.0000	2.281	2.0	100
USE_321	0.977	0.070	195.690	0.0000	2.664	3.0	100
USE_111	2.237	0.207	116.737	0.0000	9.419	4.7	100
LCL_PLAINS	-2.071	0.164	159.161	0.0000	0.127	5.1	100
LITH_Ca	0.624	0.091	53.558	0.0000	1.886	8.3	100
LITH_Ev	0.934	0.120	63.700	0.0000	2.596	8.5	100
USE_23	0.599	0.084	51.154	0.0000	1.827	9.4	100
USE_112	1.571	0.220	51.222	0.0000	4.901	9.7	100
LCL_OPEN	-0.562	0.072	61.954	0.0000	0.570	10.0	100
LITH_SaCl	-0.644	0.111	37.426	0.0062	0.531	10.1	100
LITH_PhMe	-0.898	0.143	42.396	0.0002	0.413	10.3	100
LCL_USHP	-0.743	0.178	18.055	0.0005	0.480	13.8	100
LITH_CDC	0.650	0.130	25.732	0.0003	1.947	12.0	97
USE_22	0.308	0.095	10.624	0.0024	1.362	15.2	52
LITH_CI	-0.028	0.106	5.054	0.2675	1.008	9.3	38
LITH_CoSa	0.576	0.179	10.373	0.0030	1.802	16.3	25
LITH_SaCa	0.329	0.156	7.230	0.1161	1.427	13.3	14

Table 6. Factors analysis for HILL\_LSN.

FACTORS ANALYSIS							
HILL_LSN							
Attribute	Avg. ( $\beta$ )	Avg. (Std-dev)	Avg. Wald	Avg. Signif	Avg. OR	Avg. RK	n
LCL_PLAINS	-1.7649	0.0561	5788.5652	0.0050	0.1807	1.0	100
RAIN_H	0.0042	0.0001	3906.3562	0.0000	1.0042	2.0	100
LITH_CI	4.2016	0.0893	2217.0432	0.0000	66.9131	3.0	100
LITH_SaCl	4.0756	0.0898	2062.8770	0.0000	58.9854	4.1	100
LITH_CoSa	4.3585	0.0930	2200.4857	0.0000	78.2727	4.9	100
LITH_Ev	4.0339	0.0927	1895.9701	0.0000	56.5819	6.0	100
LCL_MNTPS	-0.5599	0.0548	775.7442	0.0334	0.6012	7.2	100
USE_22	-0.5557	0.0242	930.3611	0.0031	0.5768	8.0	100
USE_31	-0.7374	0.0296	930.1456	0.0000	0.4809	8.8	100
LITH_PhMe	3.6490	0.0927	1552.2501	0.0000	38.5212	10.4	100
LITH_CDC	3.6064	0.0919	1543.0805	0.0000	36.8982	11.4	100
LITH_Ca	-3.3053	0.0903	1343.2796	0.0000	27.3051	12.4	100
LITH_SaCa	3.2414	0.0923	1234.5050	0.0000	25.6089	13.4	100
LCL_UPPSL	-0.5569	0.0665	311.2864	0.0315	0.6040	14.4	100
USE_22	-0.6382	0.0565	133.4693	0.0000	1.9036	16.7	100
SLO	0.0130	0.0009	196.2342	0.0000	1.0130	16.8	100
LCL_USHP	0.4612	0.0612	171.5003	0.0001	1.6692	17.4	100
LCL_MRDG	-0.3298	0.0699	106.7273	0.0154	0.7555	18.7	100
USE_51	-1.2700	0.1325	94.3962	0.0000	0.2827	20.1	100
USE_32	-0.1319	0.0225	85.1833	0.0019	0.8815	20.7	100
USE_14	-11.0917	79.2952	0.0235	0.8809	0.0000	22.9	94
LCL_CANY	0.3067	0.0563	114.4982	0.0005	1.4338	19.4	90
USE_33	-0.4630	0.1721	7.4578	0.0137	0.6318	12.3	47
USE_112	0.2241	0.0632	11.5918	0.0035	1.2653	23.0	43
USE_13	-0.3685	0.1254	8.7318	0.0044	0.6923	23.6	31
USE_21	0.0654	0.0419	10.9851	0.0024	1.0881	23.9	24
LSC_OPEN	0.2898	0.1851	63.6125	0.0693	1.5858	17.4	20
LCL_MDRG	0.2982	0.2121	25.2307	0.0123	1.6133	21.1	18
USE_23	0.3639	0.0903	15.6449	0.0003	1.4481	23.9	8



## 6. Discussion

These analyses allow the generation of the susceptibility maps for SCR\_LSN and HILL\_LSN for both CA and for BLR.

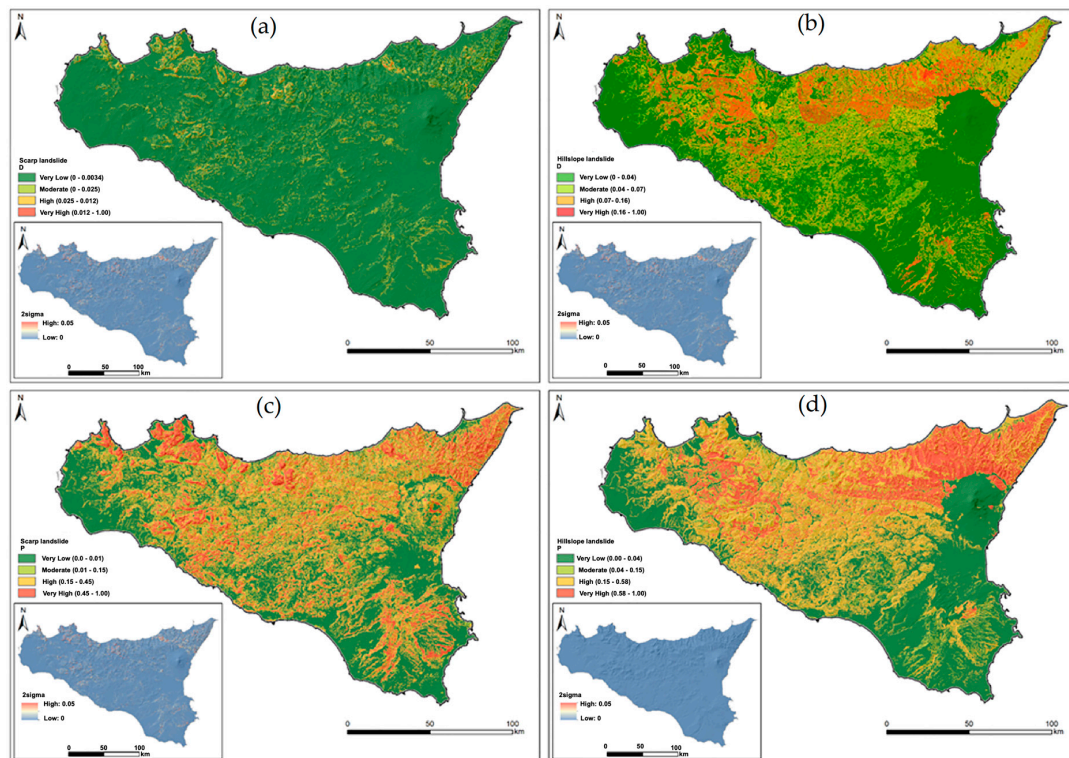
The GRID UCU layers were intersected with 100 different TRNsubset landslide layers (SCR\_LSN and HILL\_SCR) and the mean density was calculated for each mapping unit. The susceptibility model derived from the CA approach was obtained by assigning each UCU to its corresponding class, according to its computed density. For each landslide type, the computed density corresponds to the susceptibility function ( $S_{UCU}$ ) equivalent to the conditional probability of a new landslide, given the selected predictive variables [44].

Although numerous studies on landslide susceptibility zoning have been published, no global approach is yet shared by the scientific community to classify susceptibility maps. In the studies, where the classification of territories in accordance with their level of landslide susceptibility is the aim, it is appropriate to determine the optimal cutoff classification values which are capable of dividing the mapping units mainly into two large domains: stable areas (susceptibility values are less than the cutoff) to the left of the cutoff and unstable terrain (susceptibility greater than the cutoff) to the right of the cutoff [66]. When statistical approaches such as CA and BLR are used, the statistical software sets the significant cutoff ( $mcutoff$ ) as equal to 0.5 by default [54,71–74] in order to correctly classify the predicted stable or unstable cells. This research area is quite extensive (>25,000 km<sup>2</sup>), therefore, it is useful to divide the territory into a few different classes, depending on the value of susceptibility, and define some useful class range boundaries for the susceptibility (or density) index. The aim is to be able to divide the entire Sicilian territory into four classes of susceptibility: very low, moderate, high, and very high. To such an end, three cutoff values are required: the low cutoff ( $lcutoff$ ), the central cutoff ( $mcutoff$ ), and the high cutoff ( $hcutoff$ ). Operationally, the  $mcutoff$  is first identified graphically, on the ROC curves, as the maximum value of the difference between the FP and TP rate. Subsequently, with the same method, were identified the other limits of the classes for the areas to the right of the  $mcutoff$  (high) and for the areas of the left of the  $mcutoff$  (low).

Table 7 shows the identification of  $mcutoff$  for the model obtained by following the CA for SCR\_LSN. The susceptibility maps based on both statistical approaches have been derived and with test inventory subsets verified. Therefore, the maps presented in Figure 5 classify the Sicilian territory with low, moderate, high and very high susceptibility values according to their degree of propensity to instability and second cutoff probability values identified objectively.

**Table 7.** Cutoff range for (a) HILL\_LSN and (b) SCR\_LSN.

(a)		CA	
		Landslide Tipology	
		SCR_LSN	HILL_LSN
Classes	Very Low	0–0.34	0–0.04
	Moderate	0.34–2.5	0.04–0.07
	High	2.5–12	0.07–0.16
	Very High	0.78–1.00	0.16–1.00
(b)		BLR	
		Landslide Tipology	
		SCR_LSN	HILL_LSN
Classes	Very Low	0–0.10	0–0.10
	Moderate	0.1–0.15	0.1–0.28
	High	0.15–0.48	0.28–0.48
	Very High	0.48–1.00	0.48–1.00



**Figure 5.** Susceptibility distribution for HILL\_LSN using CA (a) and BLR analysis (b); susceptibility distribution for SCR\_LSN using CA (c) and BLR analysis (d).

The presented research is focused on verifying two main concepts: the first, to explore the possibility of using existing data (landslide inventories, thematic maps of predictor variables) to create regional landslide susceptibility models; the second, to verify the exploitability of CA, comparing it with models based on BLR, in order to create landslide susceptibility models for areas and studies bigger than 1:100,000.

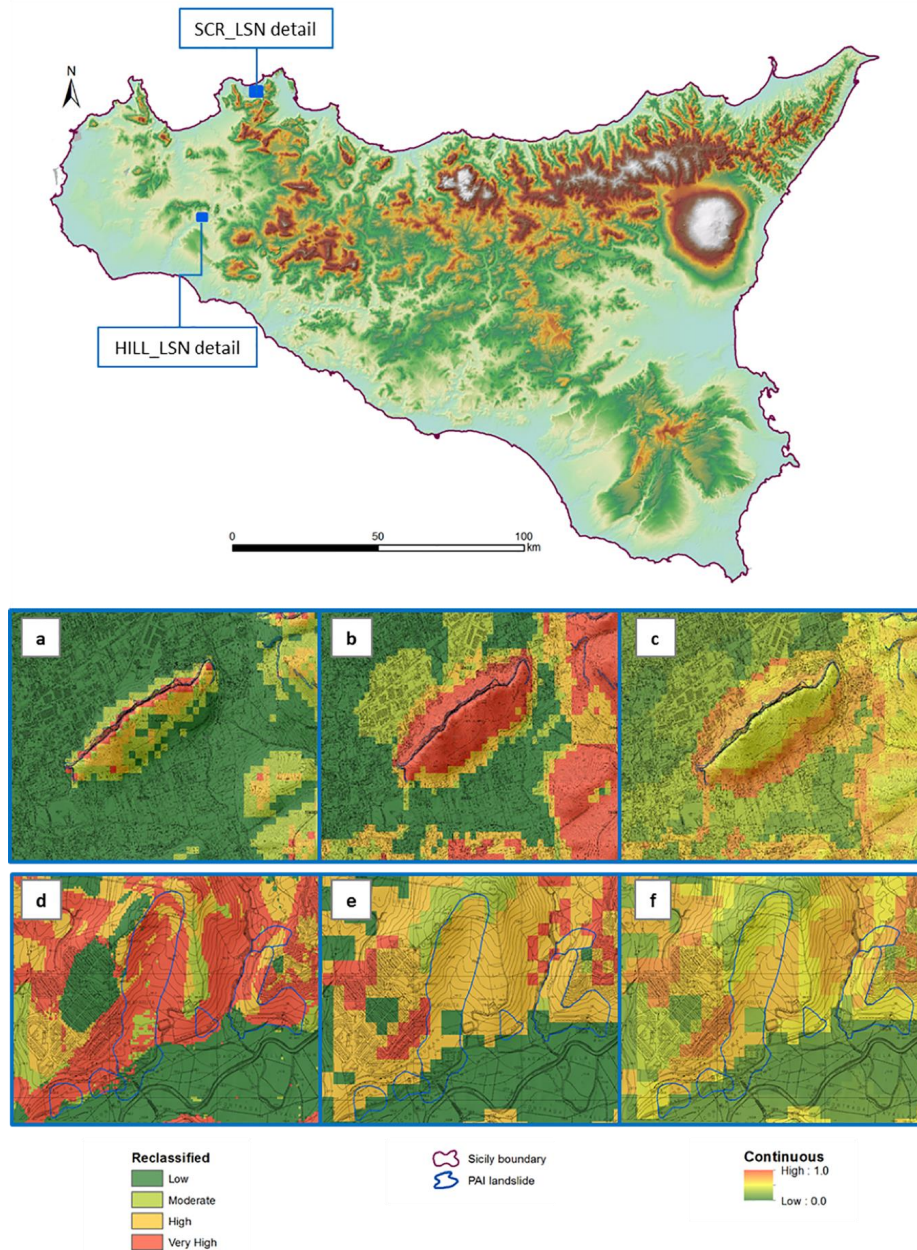
When analyzing the maps presented in Figure 5, it is noteworthy that those created by BLR (b, d) have a chromatic gradation indicating susceptibility conditions for Sicily which are generally higher than those created by CA (a, c). Therefore, this remains an open question: Which statistical approach is more representative of the conditions of susceptibility?

To answer this question, we analyzed the various sectors of the research area in detail and the different maps were compared and analyzed (Figure 6).

In order to single out the targets in the present study, 75% of instability phenomena (training subset) surveyed in the regional inventory of landslides, produced by the Environment and Territory Department of the Sicilian region (ARTA) was used to create two different landslide susceptibility maps: one based on the conditional analysis approach, and another on the binary logistic regression approach. The slope failure archive was simplified into two main types based on expert judgment of which preparatory variables they shared: the scarp landslide and hillslope landslide. The results obtained by CA showed an outstanding predictive ability for models based on a small number of predictive parameters combined in UCUs which were verified through spatial validation using a test subset randomly extracted from the Sicilian regional landslide inventory that covers the entire territory. ROC curve validation of the 100 different models showed the unquestionable excellence and stability of the forecasting performance for both SCR\_LSN and HILL\_LSN. A quality control test on four susceptibility maps was applied according to the degree of fit approach (DF) [24,27,73–75]. DF represents the percentage of an area subject to landslides for each range of classes of susceptibility and is determined by cross-tabulation of the landslide test subset with the susceptibility class of maps. DF can be expressed with the following formula:

$$DF = \frac{LSNi/Si}{\sum LSNi/Si} \quad (3)$$

where  $LSN$  is the area occupied in the  $i$ -class of susceptibility and  $S_i$  is the area of the  $i$  susceptibility class [74,75].



**Figure 6.** Different classifications of susceptibility values for HILL\_LSN (d,e,f) and SCR\_LSN (a,b,c) using BLR (a,d) and CA (b,e) reclassified. Figures (c) and (f) show the susceptibility values obtained with the BLR technique and are shown with continuous values.

The percentage of area subject to landslides falling into the null or moderate susceptibility class may be considered as a false negative error. The histograms shown in Figure 7 highlight an excellent predictive ability of new landslides not seen in the construction of maps. The percentages of relative accuracy are understood as the sum of the degrees of fit of the high and very high susceptibility classes [27] and the value of  $R^2$  of the trend line confirms the goodness of the susceptibility performance models created.

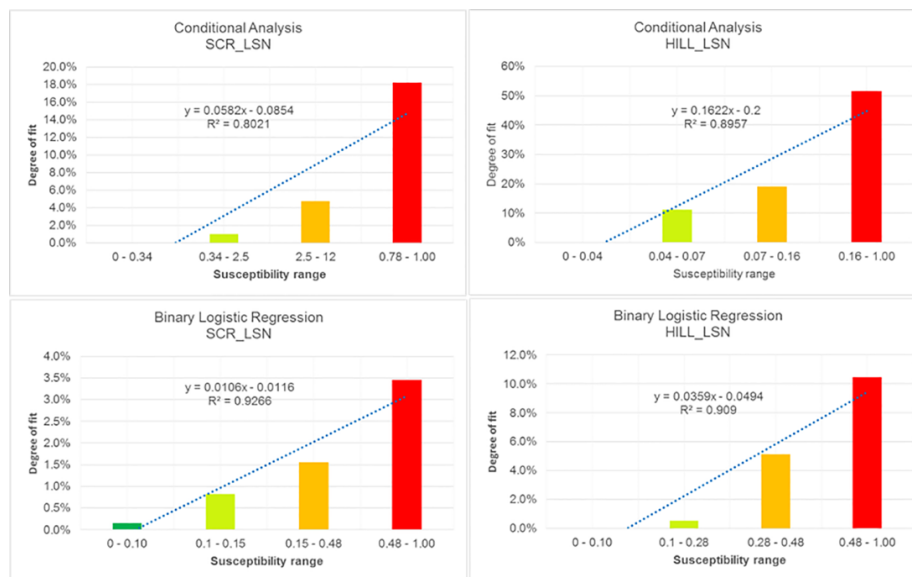


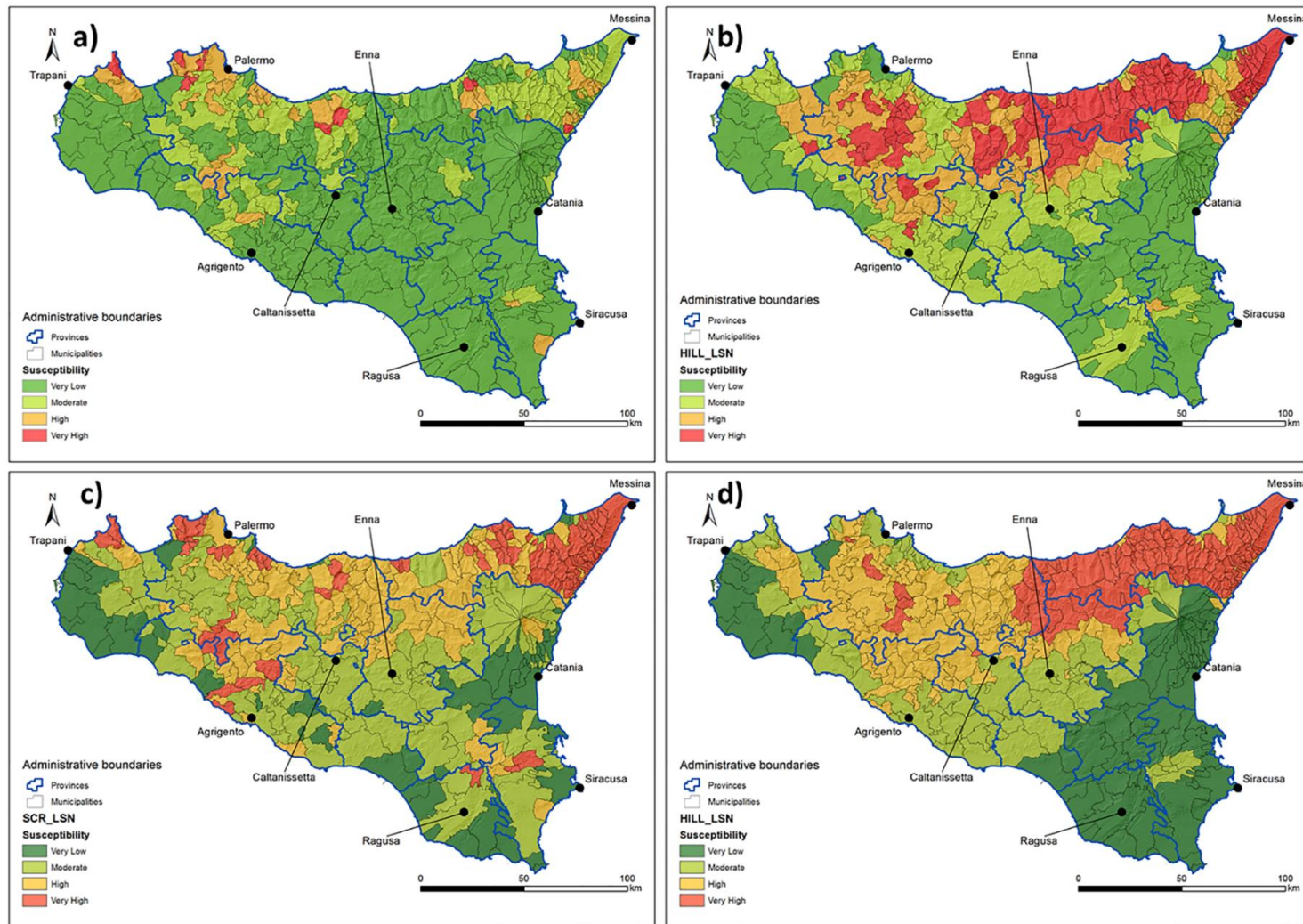
Figure 7. Distribution of landslides by susceptibility classes.

The propensity of a territory to be affected by new landslides and the degree of hazard or risk that characterizes it are usually expressed with the help of a map in which the area is divided into different zones according to the different values that qualify it. In this mapping, the territory is zoned or divided into homogeneous zones or user-defined fields/areas, the ranking of which is defined according to their real or potential degree of landslide susceptibility [8].

From among the Italian regions, Sicily is one of areas most affected by geomorphological instability. Landslide activity is a clear threat to the territory, facilities, and people present there. From this point of view, each part of the territory is characterized by a landslide vulnerability value. Generally, it depends on the territorial level of exposure to the threat, determined by the socio-economic value of the assets as well as by their resistance to the stresses expected. Interaction between humans and the natural environment is a very complex and diverse issue, not often approached in a systematic way, as resources are primarily invested in risk-mitigating measures, while being severely limited when it comes to the medium- and long-term research needed to understand the environment better and more effectively. The current PAI archive version is highly dependent on the past instability scenario. By looking at those that were counted and catalogued using a matrix system of evaluation, it is possible to derive the conditions of the associated geomorphological risk. This last concept represents a big step forward as it is now necessary to consider strongly the concept of landslide susceptibility which would represent the adoption of a spatial analysis tool with predictive power. On this scale, output cuts will be based on administrative boundaries (municipalities).

To meet this goal, the mean values of susceptibility (CA) and probability (BLR) were assigned to each municipality of the Sicily region (the calculation was only performed for the 382 municipalities included within the main island) in a GIS environment.

The output susceptibility values obtained by means of the two methods (CA and BLR) for the municipal boundaries are represented as maps in Figure 8. The picture shows the susceptibility maps relating to both conditional methods (a, b) and BLR (c, d) for the SCR\_LSN and HILL\_LSN typologies, respectively. In particular, it is possible to observe that for the SCR\_LSN types, maps (a, c) are very different with regard to the distribution of the susceptibility values among the municipal boundaries. On the other hand, the comparison between the maps concerning the HILL\_LSN type shows a convergence among the results and suitable differentiation for the urban units. Both statistical approaches (CA and BLR) show higher susceptibility landslide values in the northeastern portion of the island, in correspondence with the mountainous chains of Nebrodi and Peloritani in the province of Messina.



**Figure 8.** Regional susceptibility map with municipalities boundary. (a,b) Using the CA approach for SCR\_LSN; (c,d) showing the susceptibility classification using the BLR approach.

## 7. Final Remarks and Management Implications

The research results presented here have demonstrated the ability to leverage a set of existing data already in the public administration to generate regional landslide susceptibility maps. A total of 100 performances executed for both the CA and BLR approaches have allowed susceptibility patterns characterized by excellent performance to be obtained, with remarkable stability, robustness, and reliability, even if CA generated more powerful models in terms of forecasted performance (AUC curves). With the goal of creating, for the whole Sicilian territory, susceptibility maps that could be useful for the purposes of effective land management and spatial urban planning, this study proposes a reclassification method into four susceptibility map classes. The reclassified susceptibility maps, determined by the identified cutoff boundaries, were validated quantitatively with the degree of fit technique. Validations demonstrated the reliability of the maps created and the explorability of the proposed protocol. However, some differences may be highlighted when comparing the final maps generated with two different approaches. In fact, the maps generated with CA seem to better represent the conditions leading to the failure of a territory than those created by BLR-generated ones. This is more realistic for the SCR\_LSN than for HILL\_LSN. This is probably due to the fact that the BLR tries to find a correlation equation between landslides and predictor training models, using a diagnostic area (the entire area subject to landslides) not only representative of the trigger conditions, involving areas only passively affected by the landslide and, in some cases generating an overestimation of the susceptibility value.

Of the 382 of Sicilian municipalities examined, 244 (nearly 64%) were correctly classified as unstable (i.e., prone to landslide instability) as they fall in the classes of high and/or very high susceptibility, both for scarp and for hillslope landslides. We particularly want to highlight that susceptibility is predicted as being very high in 90 of 108 municipalities in the province of Messina (more than 80%). The municipalities featuring lower susceptibility values were those within the provinces of Syracuse and Ragusa. This approach is based on the propensity for gravitational instability, unlike PAI, and allows the concept of landslide density or density index to be overcome. Table 8 lists the top 10 municipalities in order, compared to the size of an area that may be affected by new landslides in the future.

For each municipality surface area, the extension area has been reported, the density (D) which corresponds to the landslide density for the current phenomena surveyed in the PAI and the P (or the area which is located to the right of  $m_{cutoff}$ ) for each municipality, and finally, the difference between the current framework of the landslide and that provided (Odd).

The table shows that the towns of Isola delle Femmine and Roccafronza, which despite having a higher SCR\_LSN density (respectively 7.04% and 6.04%) are the last two municipalities in the table, because of the difference, in terms of Odd, between the density and the future probability, which is lower than that seen in the municipalities of San Vito Lo Capo and Frazzanò. Almost 4 km of new territory in the municipality of San Vito Lo Capo, for example, could be affected by new SCR\_LSN activations in the future. Similarly, we can say that linked to HILL\_LSN, the territory of Alcara Li Fusi is the Sicilian municipality with the highest HILL\_LSN density value, but the municipalities of Ficarra (with 8 km<sup>2</sup>) and Sinagra (with 6.54 km<sup>2</sup>), both in the province of Messina, will be most affected by new activations of hillslope landslides.

The large and widespread use of known geostatistical methods has gone through at least three decades of landslide hazard studies, but still does not eliminate some of the conceptual and operational bottlenecks, only sporadically resulting in the safety enforcement schemes by the authorities involved in studying landslide risk in Italy.

**Table 8.** The top 10 municipalities ordered according to the size of the area that may be affected by new landslides in the future ((a) SCR\_LSN in Table; (b) HILL\_LSN).

(a) Municipality	Area (km <sup>2</sup> )	P	D	odd	km <sup>2</sup>
SAN VITO LO CAPO	3.6	4.42%	1.98%	2.44%	3.91
FRAZZANO'	1.2	5.90%	3.02%	2.88%	1.46
GIARDINELLO	50.8	5.76%	2.33%	3.42%	0.89
ISNELLO	6.9	6.76%	3.70%	3.05%	0.82
TORRETTA	26.0	4.28%	1.61%	2.67%	0.69
SAN MARCO D'ALUNZIO	12.8	5.85%	2.58%	3.27%	0.23
BORGETTO	26.0	5.12%	2.97%	2.16%	0.14
CINISI	33.0	4.81%	2.57%	2.24%	0.14
ISOLA DELLE FEMMINE	60.2	10.70%	7.04%	3.66%	0.13
ROCCAFIORITA	25.5	9.49%	6.04%	3.45%	0.04
(b) Municipality	Area (km <sup>2</sup> )	P	D	odd	km <sup>2</sup>
FICARRA	215.6	88.85%	3.72%	85.13%	8.03
SINAGRA	70.0	93.26%	9.34%	83.91%	6.54
ALCARA LI FUSI	31.1	91.98%	16.74%	75.24%	5.21
CASTELL'UMBERTO	30.2	92.05%	7.31%	84.74%	2.21
SAN PIERO PATTI	14.4	89.30%	9.77%	79.52%	1.40
MONTAGNAREALE	11.4	89.73%	11.07%	78.66%	1.26
RACCUJA	15.8	88.78%	5.60%	83.19%	0.88
UCRIA	26.1	94.39%	3.14%	91.25%	0.82
SANT'ANGELO DI BROLO	18.5	89.90%	4.34%	85.56%	0.80
TORTORICI	23.9	89.21%	2.76%	86.46%	0.66

Since economic problems, which are common to all countries, do not allow either investment in research projects on a medium- and long-term scale, the concept of landslide susceptibility should represent, for all political and administrative actors dealing with environmental and territorial policies, a new approach to the problem associated with mass movements. For this reason, the scientific community is engaged in a continuous search for methods and techniques to estimate the degree of real and potential instability, using the minimum amount of equipment, data, and economic resources possible. Generally, substantial difficulty exists in identifying the most reliable procedures, allowing this matter to be approached in a non-traditional manner based on modelling and investigative techniques built on the exchange of experiences between experts, studies, and experiments on every continent, and showing different strategies and possible technical combinations, depending on the type and/or the number and complexity of the investigation, producing susceptibility, hazard, and risk maps, used as the basis for decision-making processes in land management. In this framework, further effort is needed in trying to make the different methods more objective and shared by all, in order to be simple and repeatable and, most of all, in transferring the knowledge gained to laws that underpin territorial planning, building regulations, and in civil defense plans [22].

Over the decades, many research groups and national and international commissions have tried to provide precise definitions, in order to generate maps indicating the different urban planning vocation of an area. For this reason, the scientific community is engaged in a continuous search for methods and techniques to estimate the degree of real and potential instability, using the minimum amount of equipment and possible economic resources.

**Author Contributions:** Writing—original draft, D.C.; review & editing, C.I. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

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

## References

1. 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](#)]
2. Guzzetti, F.; Reichenbach, P.; Cardinali, M.; Galli, M.; Ardizzone, F. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology* **2005**, *72*, 272–299. [[CrossRef](#)]
3. Brabb, E.E. Innovative approach to landslide hazard and risk mapping. In Proceedings of the 4th International Symposium on Landslides, Toronto, ON, Canada, 16–21 September 1984; Volume 1, pp. 307–324.
4. Soeters, R.; van Westen, C.J. Slope instability recognition, analysis, and zonation. In *Landslides Investigation and Mitigation*; Special Report 247; Transportation Research Board: Washington, DC, USA, 1996; pp. 129–177.
5. Lyell, C. *Principles of Geology, Being an Attempt to Explain the Former Changes of the Earth's Surface, by Reference to Causes Now in Operation*; John Murray: London, UK, 1833; Volume 3.
6. Aleotti, P.; Chowdhury, R. Landslide hazard assessment: Summary review and new perspectives. *Bull. Eng. Geol. Environ.* **1999**, *58*, 21–44. [[CrossRef](#)]
7. Carrara, A.; Cardinali, M.; Detti, R.; Guzzetti, F.; Reichenbach, P. Gis techniques and statistical models in evaluating landslide-hazard. *Earth Surf. Process. Landf.* **1991**, *16*, 427–445. [[CrossRef](#)]
8. Varnes, D.J. Slope movements, type and processes. In *Landslides Analysis and Control*; Schuster, R.L., Krizek, R.J., Eds.; Special Report 176; Transportation Research Board: Washington, DC, USA, 1984.
9. Cascini, L. Applicability of landslide susceptibility and hazard zoning at different scales. *Eng. Geol.* **2008**, *102*, 164–177. [[CrossRef](#)]
10. Hervás, J.; Günther, A.; Reichenbach, P.; Chacón, J.; Pasuto, A.; Malet, J.-P.; Trigila, A.; Hobbs, P.; Maquaire, O.; Tagliavini, F.; et al. Recommendations on a common approach for mapping areas at risk of landslides in Europe. In *Guidelines for Mapping Areas at Risk of Landslides in Europe*; Hervás, J., Ed.; JRC Report EUR 23093 EN; Office for Official Publications of the European Communities: Luxembourg, 2007; 53p.
11. Eckelmann, W.; Baritz, R.; Bialousz, S.; Carre, F.; Jones, B.; Kibblewhite, M.; Kozak, J.; Le Bas, C.; Tóth, G.; Várallyay, G.; et al. *Common Criteria for Risk Area Identification according to Soil Threats*; European Soil Bureau Research Report No.20, EUR 22185 EN; Office for Official Publications of the European Communities: Luxembourg, 2006; 94p.
12. Glade, T.; Reichenbach, P.; Resources, N. Preface “Landslide hazard and risk assessment at different scales”. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2169–2171.
13. Günther, A.; Reichenbach, P.; Malet, J.; Van Den Eeckhaut, M.; Hervás, J.; Dashwood, C.; Guzzetti, F. Tier-based approaches for landslide susceptibility assessment in Europe. *Landslides* **2012**, *10*, 529–546. [[CrossRef](#)]
14. Manzo, G.; Tofani, V.; Segoni, S.; Battistini, A.; Catani, F. GIS techniques for regional-scale landslide susceptibility assessment: The Sicily (Italy) case study. *Int. J. Geogr. Inf. Sci.* **2013**, *27*, 1433–1452. [[CrossRef](#)]
15. ARTA\_Sicilia. *Piano Stralcio di bacino per l'Assetto Idrogeologico Della Regione Siciliana*; Assessorato Regionale Territorio e Ambiente della Regione Sicilia: Sicilia, Italy, 2004; 165p.
16. Agnesi, V.; Di Maggio, C.; Macaluso, T. Morphostructural setting and geomorphological evolution in the Madonie Mountains (Northern Sicily, Italy). *Suppl. Geogr. Fis. Din. Quat.* **1997**, *3*, 43–44.
17. Catalano, R.; Agate, M.; Albanese, C.; Avellone, G.; Basilone, L.; Morticelli, G.; Gugliotta, C.; Sulli, A.; Valenti, V.; Gibilaro, C.; et al. *Walking along a Crustal Profile Across the Sicily Fold and Thrust Belt*; Geological Field Trips: Palermo, IT, USA, 2013; Volume 5.
18. Di Maggio, C.; Madonia, G.; Parise, M.; Vattano, M. Karst of Sicily and its conservation. *J. Cave Karst Stud.* **2012**, *74*, 157–172. [[CrossRef](#)]
19. Dikau, R.; Brunnsden, D.; Schrott, L.; Ibsen, M. *Landslide Recognition. Identification, Movement and Courses*; Report No. 1 of the European Commission Environment Programme; Wiley and Sons: Chichester, UK, 1996; 251p.
20. Cruden, D.M. A simple definition of a landslide. *Bull. Int. Assoc. Eng. Geol.* **1991**, *43*, 27–29. [[CrossRef](#)]
21. Cruden, D.M.; Varnes, D.J. Landslides Types and Processes. In *Landslides: Investigation and Mitigation*; Special Report 247; Transportation Research Board: Washington, DC, USA, 1996; pp. 36–75.
22. Guzzetti, F.; Reichenbach, P.; Ardizzone, F.; Cardinali, M.; Galli, M. Estimating the quality of landslide susceptibility models. *Geomorphology* **2006**, *81*, 166–184. [[CrossRef](#)]



23. Bosi, C.; Dramis, F.; Gentili, B. Carte geomorfologiche di dettaglio e carte di stabilità: Esempi nel territorio marchigiano. *Geol. Appl. Idrogeol.* **1982**, *20*, 53–62.
24. Chacón, J.; Irigaray, C.; Fernández, T.; El Hamdouni, R. Engineering geology maps: Landslides and geographical information systems. *Bull. Eng. Geol. Environ.* **2006**, *65*, 341–411. [[CrossRef](#)]
25. Carrara, A.; Guzzetti, F.; Cardinali, M.; Reichenbach, P. Use of GIS technology in the prediction and monitoring of landslide hazard. *Nat. Hazards* **1999**, *20*, 117–135. [[CrossRef](#)]
26. Carrara, A.; Cardinali, M.; Guzzetti, F.; Reichenbach, P. GIS technology in mapping landslide hazard. In *Geographical Information Systems in Assessing Natural Hazards*; Carrara, A., Guzzetti, F., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; pp. 135–175.
27. Irigaray, C.; Fernández, T.; El Hamdouni, R.; Chacón, J. Evaluation and validation of landslide-susceptibility maps obtained by a GIS matrix method: Examples from the Betic Cordillera (southern Spain). *Nat. Hazards* **2007**, *41*, 61–79. [[CrossRef](#)]
28. Carrara, A.; Crosta, G.; Frattini, P. Comparing models of debris-flow susceptibility in the alpine environment. *Geomorphology* **2008**, *94*, 353–378. [[CrossRef](#)]
29. Lombardo, L.; Cama, M.; Maerker, M.; Rotigliano, E. A test of transferability for landslides susceptibility models under extreme climatic events: Application to the Messina 2009 disaster. *Nat. Hazards* **2014**, *74*, 1951–1989. [[CrossRef](#)]
30. Rossi, M.; Guzzetti, F.; Reichenbach, P.; Mondini, A.C.; Peruccacci, S. Optimal landslide susceptibility zonation based on multiple forecasts. *Geomorphology* **2010**, *114*, 129–142. [[CrossRef](#)]
31. Van Den Eeckhaut, M.; Hervás, J.; Jaedicke, C.; Malet, J.P.; Montanarella, L.; Nadim, F. Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. *Landslides* **2011**, *9*, 357–369. [[CrossRef](#)]
32. Clerici, A.; Perego, S.; Tellini, C.; Vescovi, P. Landslide failure and runout susceptibility in the upper T. Ceno valley (Northern Apennines, Italy). *Nat. Hazards* **2009**, *52*, 1–29. [[CrossRef](#)]
33. Conoscenti, C.; Di Maggio, C.; Rotigliano, E. GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). *Geomorphology* **2008**, *94*, 325–339. [[CrossRef](#)]
34. Costanzo, D.; Cappadonia, C.; Conoscenti, C.; Rotigliano, E. Exporting a Google Earth™ aided earth-flow susceptibility model: A test in central Sicily. *Nat. Hazards* **2011**, *61*, 103–114. [[CrossRef](#)]
35. Costanzo, D.; Rotigliano, E.; Irigaray, C.; Jiménez-Perálvarez, J.D.; Chacón, J. Factors selection in landslide susceptibility modelling on large scale following the GIS matrix method: Application to the river Beiro basin (Spain). *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 327–340. [[CrossRef](#)]
36. Costanzo, D.; Chacón, J.; Conoscenti, C.; Irigaray, C.; Rotigliano, E. Forward logistic regression for earth-flow landslide susceptibility assessment in the Platani river basin (southern Sicily, Italy). *Landslides* **2014**, *11*, 639–653. [[CrossRef](#)]
37. Lombardo, L.; Cama, M.; Conoscenti, C.; Marker, M.; Rotigliano, E. Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: Application to the 2009 storm event in Messina (Sicily, southern Italy). *Nat. Hazards* **2015**, *79*, 1621–1648. [[CrossRef](#)]
38. Rotigliano, E.; Agnesi, V.; Cappadonia, C.; Conoscenti, C. The role of the diagnostic areas in the assessment of landslide susceptibility models: A test in the Sicilian chain. *Nat. Hazards* **2011**, *58*, 981–999. [[CrossRef](#)]
39. Rotigliano, E.; Cappadonia, C.; Conoscenti, C.; Costanzo, D.; Agnesi, V. Slope units-based flow susceptibility model: Using validation tests to select controlling factors. *Nat. Hazards* **2012**, *61*, 143–153. [[CrossRef](#)]
40. Clerici, A.; Perego, S.; Tellini, C.; Vescovi, P. A GIS-based automated procedure for landslide susceptibility mapping by the Conditional Analysis method: The Baganza valley case study (Italian Northern Apennines). *Environ. Geol.* **2006**, *50*, 941–961. [[CrossRef](#)]
41. Vergari, F.; Della Seta, M.; Del Monte, M.; Fredi, P.; Lupia Palmieri, E. Landslide susceptibility assessment in the Upper Orcia Valley. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 1475–1495. [[CrossRef](#)]
42. Davis, J.C. *Statistics and Data Analysis in Geology*; John Wiley & Sons Inc.: New York, NY, USA, 1973; 564p.
43. Chung, C.F.; Fabbri, A.G. Validation of Spatial Prediction Models for Landslide Hazard Mapping. *Nat. Hazards* **2003**, *30*, 451–472. [[CrossRef](#)]
44. Jiménez-Perálvarez, J.D.; Irigaray, C.; El Hamdouni, R.; Chacón, J. Landslide-susceptibility mapping in a semi-arid mountain environment: An example from the southern slopes of Sierra Nevada (Granada, Spain). *Bull. Eng. Geol. Environ.* **2010**, *70*, 265–277. [[CrossRef](#)]

45. Santacana, N.; Baeza, B.; Corominas, J.; De Paz, A.; Marturiá, J. A GIS-Based Multivariate Statistical Analysis for Shallow Landslide Susceptibility Mapping in La Pobla de Lillet Area (Eastern Pyrenees, Spain). *Nat. Hazards* **2003**, *30*, 281–295. [[CrossRef](#)]
46. Hosmer, D.W.; Lemeshow, S. Interpretation of the fitted logistic regression model. In *Applied Logistic Regression*, 2nd ed.; John Wiley & Sons, Inc.: New York, NY, USA, 2000; pp. 47–90.
47. Akgün, A.; Türk, N. Mapping erosion susceptibility by a multivariate statistical method: A case study from the Ayvalik region, NW Turkey. *Comput. Geosci.* **2011**, *37*, 1515–1524. [[CrossRef](#)]
48. 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](#)]
49. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2013**, *11*, 425–439. [[CrossRef](#)]
50. Van Den Eeckhaut, M.; Hervás, J. State of the art of national landslide databases in Europe and their potential for assessing landslide susceptibility, hazard and risk. *Geomorphology* **2012**, *139–140*, 545–558. [[CrossRef](#)]
51. Dai, F.C.; Lee, C.F.; Li, J.; Xu, Z.W. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ. Geol.* **2001**, *40*, 381–391.
52. Kim, G.; Lee, J.; Lee, K. Application of representative elementary area (REA) to lineament density analysis for groundwater implications. *Water Resour.* **2004**, *8*, 27–42. [[CrossRef](#)]
53. Dai, F.C.; Lee, C.F. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology* **2020**, *42*, 213–228. [[CrossRef](#)]
54. Rakotomalala, R. TANAGRA: A free software for research and academic purposes. In Proceedings of the European Grid Conference, Amsterdam RNTI-E-3, Amsterdam, The Netherlands, 5 February 2005; Volume 2, pp. 697–702.
55. Crozier, M.J. Field assessment of slope instability. In *Slope Instability*; Brunsden, D., Prior, D.B., Eds.; John Wiley & Sons: Chichester, UK, 1984; pp. 103–142.
56. Hansen, A. Landslide hazard analysis. In *Slope Instability*; Brunsden, D., Prior, D.B., Eds.; John Wiley & Sons: Chichester, UK, 1984; pp. 523–602.
57. Palenzuela, J.A.; Jiménez-Perálvarez, J.D.; Chacón, J.; Irigaray, C. Assessing critical rainfall thresholds for landslide triggering by generating additional information from a reduced database: An approach with examples from the Betic Cordillera (Spain). *Nat. Hazards* **2016**, *84*, 185–212. [[CrossRef](#)]
58. Jenness, J. Topographic Position Index (TPI) v. 1.3a. Jenness Enterprises. 2006. Available online: <http://www.jennessent.com/arcview/tpi.htm> (accessed on 8 June 2020).
59. Regmi, N.R.; Giardino, J.R.; McDonald, E.V.; Vitek, J.D. A comparison of logistic regression-based models of susceptibility to landslides in western Colorado, USA. *Landslides* **2013**, *11*, 247–262. [[CrossRef](#)]
60. Hanley, A.J.; McNeil, J.B. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* **1982**, *143*, 29–36. [[CrossRef](#)] [[PubMed](#)]
61. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [[CrossRef](#)]
62. Nandi, A.; Shakoor, A. A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. *Eng. Geol.* **2010**, *110*, 11–20. [[CrossRef](#)]
63. Frattini, P.; Crosta, G.; Carrara, A. Techniques for evaluating the performance of landslide susceptibility models. *Eng. Geol.* **2010**, *111*, 62–72. [[CrossRef](#)]
64. Lasko, T.A.; Bhagwat, J.G.; Zou, K.H.; Ohno-Machado, L. The use of receiver operating characteristic curves in biomedical informatics. *J. Biomed. Inform.* **2005**, *38*, 404–415. [[CrossRef](#)] [[PubMed](#)]
65. Goodenough, D.J.; Rossmann, K.; Lusted, L.B. Radiographic applications of receiver operating characteristic (ROC) curves. *Radiology* **1974**, *110*, 89–95. [[CrossRef](#)]
66. Egan, J.P. *Signal Detection Theory and ROC Analysis*; Academic Press: New York, NY, USA, 1975; 277p.
67. Williams, C.J.; Lee, S.S.; Fisher, R.A.; Dickerman, L.H. A comparison of statistical methods for prenatal screening for Down syndrome. *Appl. Stoch. Models Data Anal.* **1999**, *15*, 89–101. [[CrossRef](#)]
68. Nefeslioglu, H.A.; Gokceoglu, C.; Sonmez, H. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Eng. Geol.* **2008**, *97*, 171–191. [[CrossRef](#)]
69. Ohlmacher, G.C.; Davis, J.C. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Eng. Geol.* **2003**, *69*, 331–343. [[CrossRef](#)]

70. Chung, C.F.; Fabbri, A.G.; Van Westen, C.J. Multivariate regression analysis for landslide hazard zonation. In *Geographical Information Systems in Assessing Natural Hazards*; Carrara, A., Guzzetti, F., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; pp. 107–133.
71. Atkinson, P.M.; Massari, R. Generalised linear modelling of susceptibility to landsliding in the Central Apennines, Italy. *Comput. Geosci.* **1998**, *24*, 373–385. [[CrossRef](#)]
72. Irigaray, C.; Fernández, T.; El Hamdouni, R.; Chacón, J. Verification of landslide susceptibility mapping: A case study. *Earth Surf. Process. Landf.* **1999**, *24*, 537–544.
73. Fernández, T.; Irigaray, C.; El Hamdouni, R.; Chacon, J. Methodology for landslide susceptibility mapping by means of a GIS. Application to the Contraviesa area (Granada, Spain). *Nat. Hazard* **2003**, *30*, 297–308.
74. Remondo, J.; Gonzalez, A.; Diaz De Teran, J.R.; Cendrero, A.; Fabbri, A.; Chung, C.F. Validation of landslide susceptibility maps; examples and applications from a case study in Northern Spain. *Nat. Hazards* **2003**, *30*, 437–449. [[CrossRef](#)]
75. Plattner, T. Modelling public risk evaluation of natural hazards: A conceptual approach. *Nat. Hazards Earth Syst. Sci.* **2005**, *5*, 357–366. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).