

Landslide Susceptibility Assessment Using Conditional Analysis and Rare Events Logistics Regression: A Case-Study in the Antrodoco Area (Rieti, Italy)

Vittorio Chiessi^{1*}, Simona Toti², Valerio Vitale³

¹Geological Survey of Italy, Department of Soil, Institute for Environmental Protection and Research, Rome, Italy

²Directorate for Methodology and Statistical Process Design, National Institute of Statistics, Rome, Italy

³Directorate for Environmental and Territorial Statistics, National Institute of Statistics, Rome, Italy

Email: *vittorio.chiessi@isprambiente.it

How to cite this paper: Chiessi, V., Toti, S. and Vitale, V. (2016) Landslide Susceptibility Assessment Using Conditional Analysis and Rare Events Logistics Regression: A Case-Study in the Antrodoco Area (Rieti, Italy). *Journal of Geoscience and Environment Protection*, 4, 1-21.
<http://dx.doi.org/10.4236/gep.2016.412001>

Received: September 9, 2016

Accepted: November 27, 2016

Published: November 30, 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

This paper discusses some methodological aspects for the production of susceptibility maps of slope instability developed within the CARG Project (Geological Cartography of Italy at 1:50,000 scale). It describes an example of a susceptibility map in the presence of low susceptibility, using database having zero or negligible cost, with the aim to test some methodologies that can be easily reproducible to get a first estimate of the landslide susceptibility on a wide area. Two statistical approaches have been applied: the non-parametric conditional analysis and the logistic analysis for rare events. The predictive ability obtained from the two methodologies, was evaluated by the success-prediction curves for the conditional analysis, and by the Receiver Operating Characteristic curve (ROC), for the logistic model. The landslide susceptibility maps have been classified into four classes using both the Natural Breaks algorithm and the method proposed by Chung and Fabbri (2003). The paper considers the influence of these two classification methods on the quality of final results.

Keywords

Landslide Susceptibility, Antrodoco, Conditional Analysis, Rare Events Logistic Regression, Classification Methods

1. Introduction

The assessment of landslide susceptibility of an area can be addressed through determi-

nistic methods, which take into account the physical and mechanical relationships between the known *a priori* geological features, or through statistic methods that identify which key factors are associated with landslide event. In general, the former are based on the safety factor for a given infinite slope model [1], while the latter make use of logistic regression [2] [3] [4], discriminant analysis [5] [6] and more recent neural networks approach [7] [8] [9] [10].

This study was developed within the CARG Project (Geological Cartography of Italy at 1:50,000 scale) of ISPRA (Institute for Environmental Protection and Research), aimed at carrying out geological and geothematic maps from geological data collected at 1:25,000 scale (in some cases at 1:10,000 scale) [11]. The project is realized by the Local Government Agencies (Regioni and Province Autonome), Universities and Research Institutions in Italy.

In this context, ISPRA has ongoing the “landslide susceptibility maps for slope instability at 1:50,000 scale in the Antrodoco area” Project, to test and evaluate different methods useful for the redaction of guidelines [12] [13].

The landslide susceptibility is defined as the likelihood of a landslide occurring in an area on the basis of local terrain conditions [14]. It is the degree to which a terrain can be affected by slope movements. Landslide susceptibility does not include a time frame or frequency of landsliding. It should consider all landsliding which can affect the study area and it includes landslides which are above the study area but may travel onto it, and landslides below the study area which may retrogressively fail up-slope into it [15].

The Antrodoco data were analyzed using two different statistical models: the non-parametric conditional analysis and the logistic analysis for rare events. These methods are easy to apply and do not need costly investigation for data collection. The disadvantages are a probabilistic prediction with a degree of uncertainty; this evaluation cannot be used for the planning of mitigation actions.

In addition, the susceptibility maps have been classified into four classes by the Jenks and Caspall (1971) method (JC) [16] and by the Chung and Fabbri (2003) method (CF) [17]. The influences of these two classification methods on the quality of results have been evaluated.

2. Study Area: Geographical and Geological Aspects

The study area is located in the central Apennines and it extends for 650 km², divided into the Rieti and L’Aquila provinces, characterized by mountain landscape including the Mt. Terminillo, with over 2000 m a.s.l.

The geological interest of the area is due to the presence of sedimentary successions which are deposited in paleoenvironments ranging from the basin to the inner carbonate platform, with all the transitional facies, and with an age ranging from the upper Triassic to the Neogene.

The different sedimentary successions crop out in 5 distinct structural units, by which the sheet can be divided (Figure 1). Four of them are separated by relevant fault

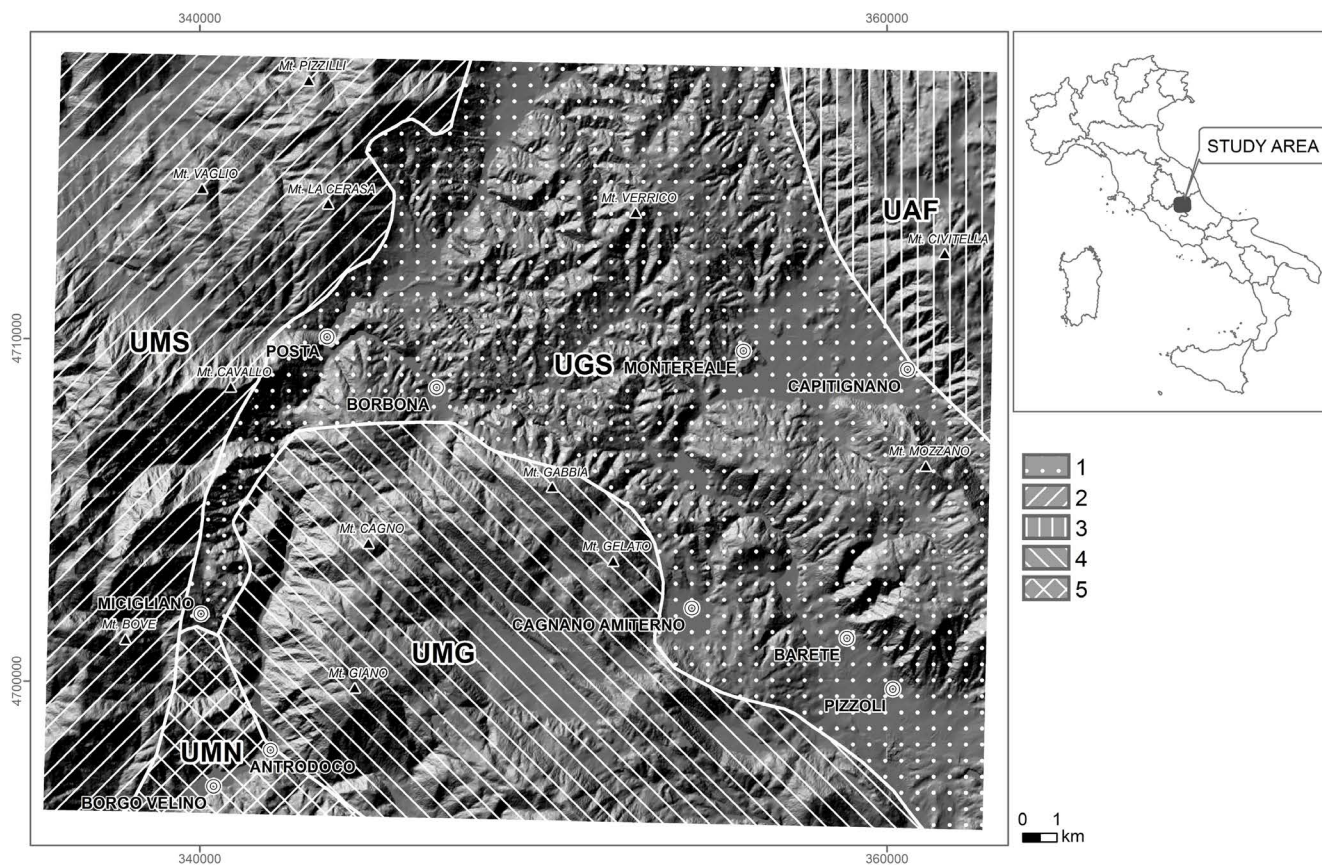


Figure 1. Location and geological sketch of the study area: (1) Gran Sasso-Cittareale Unit (UGS); (2) Mts. Sibillini Unit (UMS); (3) Acquasanta-Mt. dei Fiori Unit (UAF); (4) Mt. Giano-Mt. Gabbia Unit (UMG); (5) Mt. Nuria Unit (UMN). Reference Cartographic System: UTM ED50 Zone 33.

lines. Moreover, some smaller low-flat areas are occupied by Quaternary continental deposits [18] [19]. The first unit (Mts. Sibillini Unit-UMS) in the western part of the sheet, consists of trending meridian ridges with accentuated morphology and altitudes exceeding 2000 m a.s.l. The stratigraphic succession starts with carbonate platform deposits evolving to escarpment/basin ones and it includes all the terms known from the Calcare Massiccio of lower Jurassic up to the Cerrognola marls of Miocene. The second area (Mt. Nuria Unit-UMN) in the southern sector of the sheet, is characterized by ridges descending from S and W to the N and to E, towards the Velino Valley. The third area (Mt. Giano-Mt. Gabbia Unit-UMG) in the southern and central part, is characterized by Apennine trending ridges.

The fourth area (Gran Sasso-Cittareale Unit-UGS), central-eastern and central-northern part of the sheet, passes toward NE to the Acquasanta-Montagna dei Fiori Unit without particular tectonic disruptions. The area is characterized by a long ridge developed with an Apennine trend; it represents the western most part of the Gran Sasso-Mt. S. Franco chain. The succession is characterized initially by carbonate platform sediments (Calcare Massiccio) and successively by slope/basin deposits as in the Monti Sibillini Unit. The fifth area (Acquasanta-Mt. dei Fiori Unit-UAF), north eastern

part of the sheet, is geologically the direct prosecution of the Gran Sasso-Cittareale Unit. Within this unit the terrigenous Tortonian-Messinian lithotypes outcrop (upper part of marne con Cerrognna, marne a Pteropodi and flysch della Laga). These deposits represent the foredeep evolution of the slope-basin succession of the Gran Sasso-Cittareale Unit.

3. Definition of Variables

Most of the considered explanatory variables are part of the public Italian databases, characterized by an easy availability and a low cost of acquisition, to make the elaboration the most general and repeatable as possible. The information not present in the databases, are acquired *ex novo*.

All information layers were included in a geographic database with coordinate reference system UTM ED50 zone 33.

The Italian Landslides Inventory, IFFI Project [20] was considered the starting dataset, which has been validated and integrated through field surveys, interpretation of aerial photographs and historical examination of the archives, obtaining a landslide distribution map at the 1:10,000 scale (Figure 2).

A total of 366 landslides were acquired, also including elements coded as high angle scarps subject to fall, Deep-Seated Gravitational-slope Deformations (DSGD), falls and debris flows. These areal elements were excluded from processing thus obtaining a dataset consisting of 159 landslides with a total area of 3.9 km². The landslides involve mainly the carbonate platform, slope and basin succession (58%) and the flysch della Laga formation (29%). They involve marginally the marly-arenaceous formation of Miocene (5%) and debris deposits (8%). The most frequent types of movement are translational slides (30%) and complex landslides (27%) and the most extended landslide areas are flows (44%) and complex landslides (29%).

Lithological features were derived from the Geological Map of Italy at 1:100,000 scale, that is currently the only official geological source of the national territory, pending publication of the New Geological Map 1:50,000 (CARG project).

The low quality of data is due to the low precision of the scale used for the geological survey and also to the lack of cartography update, dated 1955 [21]. Lithotypes with similar physico-mechanical properties were collapsed, e.g. the carbonate rock-like formation, defining 8 modalities: Quaternary deposits, consolidated Quaternary deposits, flysch deposits, marly limestones, marls and clay marls, limestone marls and clays alternation, stratified limestone and dolostones. Geographical distribution of the information layer is given in Figure 2 and Table 1 shows the geological description of the lithotypes.

The study area is one of higher seismic risk area of Italy. The seismic acceleration describes partially the real local seismic response, which depends for a completely description on many other parameters, not observed here.

The national dataset of the seismic hazard (in terms of PGA-Peak Ground Acceleration) is provided by the National Institute of Geophysics and Volcanology (INGV),

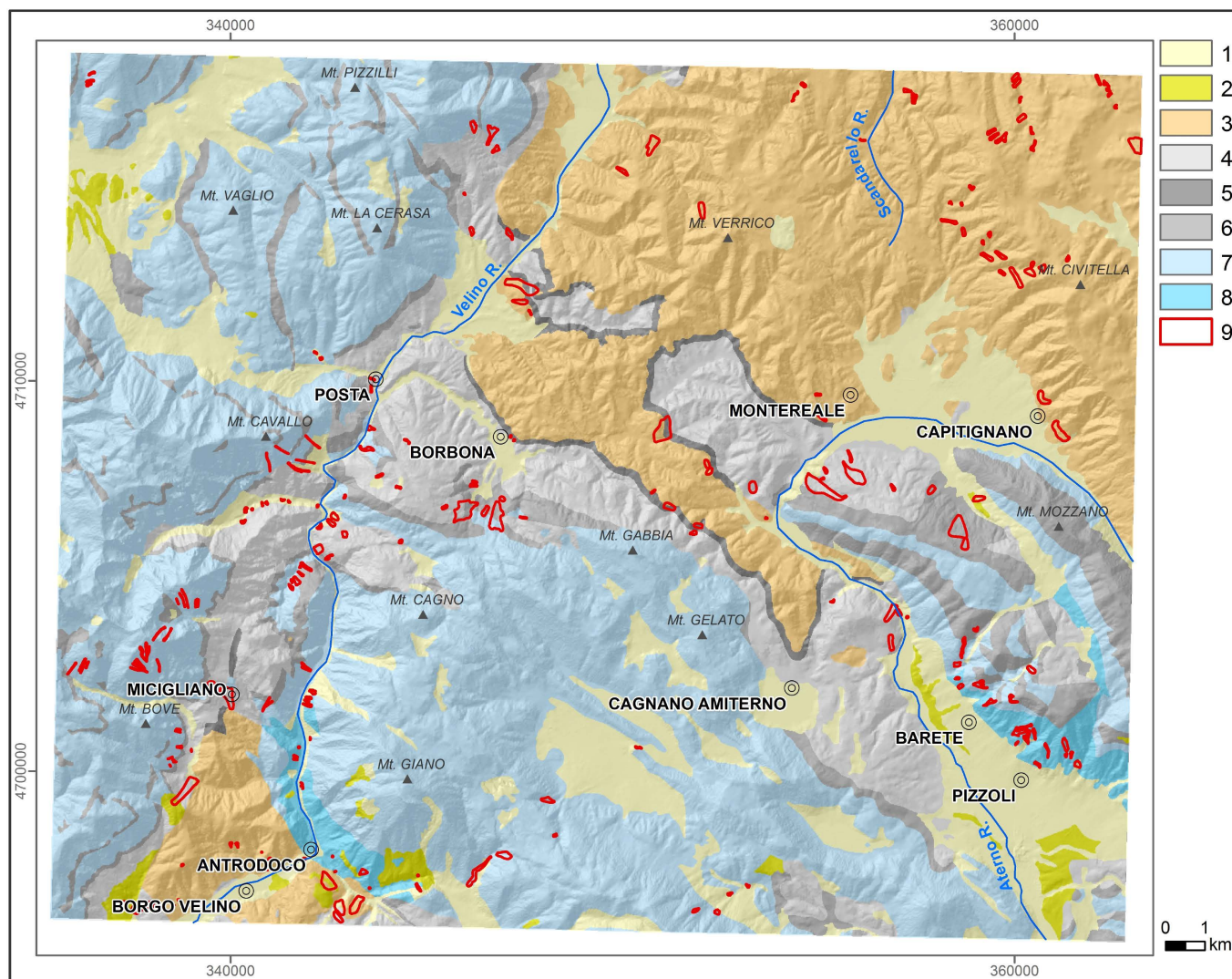


Figure 2. Lithological and landslide inventory map: (1) Quaternary deposits; (2) Consolidated Quaternary deposits; (3) Flysch deposits; (4) Marly limestones; (5) Marls and clay marls; (6) Limestones, marls and clays alternations; (7) Stratified limestone; (8) Dolostones; (9) Landslide inventory.

published on the website <http://esse1.mi.ingv.it/>. The seismic hazard was evaluated for grid points with 10 km spacing and reference interval of 30 years [22]. The values are uniformly distributed (ranging from 22 g to 26 g \times 100), growing from SE to NE (**Figure 3(a)**).

The national digital terrain model (DTM) realized by the Military Geographic Institute (IGM) was obtained by interpolating the orographic data in order to produce a matrix with 20 m spacing. The continuous variable was discretized into 6 categories ranging between min = 449 m a.s.l. and max = 2150 m a.s.l. (**Figure 3(b)**). From this information layer the slope, the aspect and the curvature of the study area was obtained, using the analysis tools in Arc Gis spatial analyst extension [23] [24].

In **Figure 3(c)** the distribution of the slope is shown. The values ranging from 0 to 75 degrees and the highest values are observed in the portion SW of the territory. Consi-

Table 1. Geological description of the lithotypes in the lithological map of **Figure 2**.

Code	Litotype	Description
1	Quaternary deposits	Quaternary sediments loose or weakly cohesive consisting of edge deposits as actual alluvion and river and lake terraces.
2	Consolidated Quaternary deposits	Quaternary age deposits cemented behavior lithoid or semi lithoid.
3	Flysch deposits	Structurally complex formations with arenaceous-silty lithology.
4	Marly limestones	Marly limestone, calcareous stratified rocks with the predominance of carbonate lithologies. Layers with of a few decimetres thickness.
5	Marls and clay marls	Marls and clay marls, stratified rocks with the prevalent presence of argillaceous lithologies. Layers with decimeter thickness.
6	Limestones, marls and clays alternations	Limestones, marls and clays alternations structurally complex formations with alternating lithologies clay-marl and limestone.
7	Stratified limestone	Stratified limestone, Limestone and calcareous marl layered thickly, Limestone in layers of a few decimeter thick or benches.
8	Dolostones	Dolomite, Dolomitic and dolomitic limestone with detrital behavior often appear brecciate and powdery for tectonic causes.

dering a morphological point of view, slope contributes to determining the extent of the outflows, in particular their sliding velocity. The continuous variable was discretized into 7 classes: 0 - 5, 5 - 10, 10 - 20, 20 - 30, 30 - 40, 40 - 50, 50 - 75.

Aspect identifies the downslope direction of the maximum rate of change in value from each cell to its neighbors. It is measured clockwise in degrees with continuous values from 0° to 360° (**Figure 3(d)**).

The subdivision adopted involves the use of 4 classes each having an amplitude of 90° and therefore the subdivision is: 0 - 90, 90 - 180, 180 - 270, 270 - 360.

Curvature is the second derivative of the input surface on a cell-by-cell basis. Positive values indicate that the surface is upwardly convex at that cell; negative values indicate that the surface is upwardly concave; 0 indicates that the surface is flat. From an applied viewpoint, curvature can be used to describe the physical characteristics of a drainage basin (erosion and runoff processes): it affects the acceleration and deceleration or the convergence and divergence of flow. In the present study the curvature is subdivided in 3 parts to identify the slope as concave, flat or convex: $-81/-1.50$, $-1.50/1.50$, $1.50/66$, respectively (**Figure 3(e)**).

The rainfall data, provided by many rain gauge stations included in the Hydrological Yearbooks (available at the website: <http://www.acq.isprambiente.it/annali.pdf>), were used to describe the rainfall height (h) as a function of their duration (t), by the following equation: $h \text{ (mm)} = a t^n$, where a is a variable depending on the return period and n is a constant for a given value of t.

The analysis was performed using the Gumbel probability distribution [25]. The reference period is from 1970 to 2002. The rainfall duration are 1, 3, 6, 12 and 24 hours and the return period adopted is 30 years. The data were interpolated using the Ordinary Kriging geostatistic method, obtaining the map in **Figure 3(f)**, where the values ranging from 93 to 157 mm of rain, increasing proceeding from O to E. The JC classification algorithm was used to defined 4 classes: 93 - 102, 102 - 115, 115 - 128, 128 - 157.

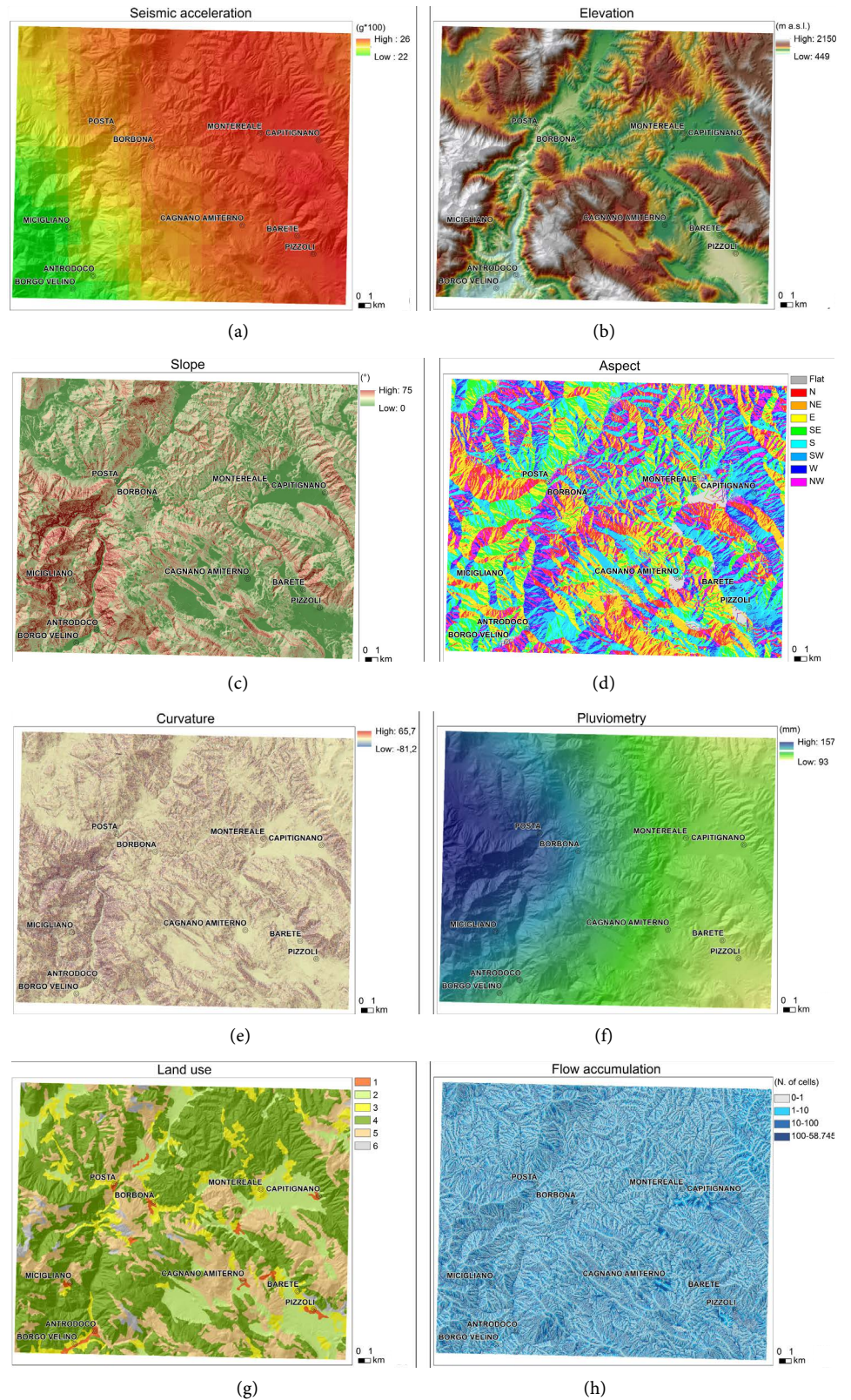


Figure 3. (a) Seismic acceleration map; (b) DTM; (c) Slope map; (d) Aspect map; (e) Curvature map; (f) Pluviometric map; (g) Land use map; (h) Flow accumulation map.

The land use was derived from the Corine Land Cover classes (www.eea.europa.eu/data-and-maps/explore-interactive-maps/corine-landcover-2006), aggregated at the second level, combining all the artificial surfaces as the Arable land and Permanent crops. Then, the variable is defined with 6 modalities: Artificial surfaces, Permanent crops, Agricultural areas, Forests, Scrub and/or herbaceous vegetation, Open spaces with little or no vegetation (**Figure 3(g)**).

The flow direction and flow accumulation functions, including in the toolset Hydrology for ArcGIS, are applied to the DTM. The first function determines the direction of flow from every cell in the raster. The second calculates the number of upslope cells that flow into each cell. Finally 4 classes were identified: 0 - 1, 1 - 10, 10 - 100, 100 - 58,745 (**Figure 3(h)**).

4. Mapping Units

The analysis was achieved by converting all the vector layer in raster layer with cell size equal to 20×20 m. The territorial units used are: cells (pixels) and UCU (Unique Condition Units) [26] [27] [28].

The explanatory variables considered in the analysis are 9: the DTM, the aspect, the curvature, the flow accumulation, the land use, the pluviometry, the slope, the seismic acceleration and finally the lithology. **Table 2** shows the absolute frequencies and percentages of the total, landslides and non-landslides pixels.

Unique Condition Units (UCU) are generated by combinations of the modalities of the 9 explanatory variables considered, by the map algebra functions of ArcGIS Spatial Analyst (ESRI), using the instrument “raster calculator”. They are homogeneous domains of the territory, that is aggregation of pixels with the same values of the parameter set chosen as variables characterizing the territory itself [29]. To reduce the number of UCU, the Cramer’s V statistics [30] was performed between each explanatory variables and the presence/absence of landslide.

The analysis showed that the variables lithology, DTM, slope, land use and aspect are more associated with the variable presence/absence of landslides, assuming the highest value of the index Cramer’s V. Then, the UCU were redefined considering only these five factors, and so the unique code used was $d_1 d_2 d_3 d_4 d_5$, where “ d_i ” identifies the modality of the variable i , then, respectively: the aspect (d_1), the land use (d_2), the DTM (d_3), the slope (d_4) and the lithology (d_5). The final number of UCU was 3644.

5. Statistical Analysis

5.1. Conditional Analysis

The conditional analysis was applied to unique condition map units to evaluate the probabilistic relationship between environmental factors and the occurrence of landslides in the considered study area [31]. The probability of landslide occurrence is calculated by the frequency of landslide, within each map units, as follow:

$$P(Y = 1 | d_1, d_2, d_3, d_4, d_5) = \text{landslide area/unit area map}$$

Table 2. Variables considered in the analysis and distribution of the frequencies of related modalities.

Variable	Classes	Variable description	N. of total pixel	N. of pixel with landslide	N. of pixel without landslide
Landslide distribution	0/1	Absence/presence phenomenon	1,520,450 (100.0%)	9840 (0.6%)	1,510,610 (99.4%)
Lithology	1	Quaternary deposits	239,216 (100.0%)	645 (0.3%)	238,571 (99.7%)
	2	Quaternary cemented deposits	19,328 (100.0%)	248 (1.3%)	19,080 (98.7%)
	3	Flysch deposits	370,870 (100.0%)	3272 (0.9%)	367,598 (99.1%)
	4	Marly limestones	200,747 (100.0%)	2998 (1.5%)	197,749 (98.5%)
	5	Marls and clay marls	15,808 (100.0%)	529 (3.3%)	15,279 (96.7%)
	6	Limestones, marls and clays alternations	105,007 (100.0%)	1360 (1.3%)	103,647 (98.7%)
	7	Stratified limestone	544,635 (100.0%)	1630 (0.3%)	543,005 (99.7%)
	8	Dolostones	24,839 (100.0%)	523 (2.1%)	24,316 (97.9%)
Seismic acceleration (g*100)	1	22 - 23	26,619 (100.0%)	256 (1.0%)	26,363 (99.0%)
	2	24	164,621 (100.0%)	1951 (1.2%)	162,670 (98.8%)
	3	25	969,836 (100.0%)	6106 (0.6%)	963,730 (99.4%)
	4	26	359,374 (100.0%)	2892 (0.8%)	356,482 (99.2%)
DTM (m a.s.l.)	1	449-700	60,107 (100.0%)	1333 (2.2%)	58,774 (97.8%)
	2	700-900	355,118 (100.0%)	3147 (0.9%)	351,971 (99.1%)
	3	900-1,100	473416 (100.0%)	4488 (0.9%)	468928 (99.1%)
	4	1,100-1,300	315,147 (100.0%)	1395 (0.4%)	313,752 (99.6%)
	5	1,300-1,500	201,476 (100.0%)	561 (0.3%)	200,915 (99.7%)
	6	1,500-2,150	115,186 (100.0%)	281 (0.2%)	114,905 (99.8%)
Slope (°)	1	0 - 5	268,839 (100.0%)	229 (0.1%)	268,610 (99.9%)
	2	5 - 10	182,601 (100.0%)	855 (0.5%)	181,746 (99.5%)
	3	10 - 20	439,473 (100.0%)	4782 (1.1%)	434,691 (98.9%)
	4	20 - 30	400377 (100.0%)	3152 (0.8%)	397225 (99.2%)
	5	30 - 40	185,410 (100.0%)	1605 (0.9%)	183,805 (99.1%)
	6	40 - 50	38,752 (100.0%)	551 (1.4%)	38,201 (98.6%)
	7	50 - 75	4998 (100.0%)	31 (0.6%)	4967 (99.4%)
Aspect (°)	1	0 - 90	445,836 (100.0%)	3656 (0.8%)	442,180 (99.2%)
	2	90 - 180	350,739 (100.0%)	2243 (0.6%)	348,496 (99.4%)
	3	180 - 270	419353 (100.0%)	2033 (0.5%)	417,320 (99.5%)
	4	270 - 360	304,522 (100.0%)	3273 (1.1%)	301,249 (98.9%)
Terrain curvature	1	-81/-1.50 (concave)	143,594 (100.0%)	1482 (1.0%)	142,112 (99.0%)
	2	-1.50/1.50 (flat)	1,228,190 (100.0%)	8614 (0.7%)	1,219,576 (99.3%)
	3	1.50/66 (convex)	148,666 (100.0%)	1109 (0.7%)	147,557 (99.3%)
Pluviometry (mm)	1	93 - 102	117,924 (100.0%)	948 (0.8%)	116,976 (99.2%)
	2	102 - 115	496,469 (100.0%)	3730 (0.8%)	492,739 (99.3%)
	3	115 - 128	310,874 (100.0%)	2427 (0.8%)	308,447 (99.2%)
	4	128 - 157	595,183 (100.0%)	4100 (0.7%)	591,083 (99.3%)
Land use	1	Artificial surfaces	12,986 (100.0%)	148 (1.1%)	12,838 (98.9%)
	2	Permanent crops	189,538 (100.0%)	332 (0.2%)	189,206 (99.8%)
	3	Agricultural areas	118,367 (100.0%)	1037 (0.9%)	117,330 (99.1%)
	4	Forests	797,509 (100.0%)	6129 (0.8%)	791,380 (99.2%)

continued

	5	Scrub and/or herbaceous vegetation	373,832 (100.0%)	3379 (0.9%)	370,453 (99.1%)
	6	Open spaces with little or no vegetation	28,218 (100.0%)	180 (0.6%)	28,038 (99.4%)
Flow accumulation (N. of cells)	1	0 - 1	698,849 (100.0%)	4315 (0.6%)	694,534 (99.4%)
	2	1 - 10	490,992 (100.0%)	3291 (0.7%)	487,701 (99.3%)
	3	10 - 100	248,217 (100.0%)	2662 (1.1%)	245,555 (98.9%)
	4	100 - 58,745	82,392 (100.0%)	937 (1.1%)	81,455 (98.9%)

where P is the conditional probability of landslide ($Y = 1$) given a certain combination of predisposing factors within the unit of the map ($d_1 d_2 d_3 d_4 d_5$).

The conditional analysis associates, for each UCU, a landslide susceptibility index (SI) calculated as the ratio between the number of pixels with landslides present in a UCU and the number of total pixels constituting the UCU itself, using the ArcGis spatial analysis “Zonal Statistics” tool.

The number of UCU with 0 landslide pixels are 2776 over 3644, corresponding to 76%, with a median size of 37 pixels per UCU (number of pixels for UCU: min = 1, max = 21,450, average = 228.56). There are 868 UCU with at least one pixel with landslides, with a median size equal to 369 pixels per UCU (min = 2, max = 16,413, average = 1021.90).

The conditional analysis results are shown in **Table 3**, where the frequency classes for high susceptibility UCU with $SI > 2.45$ (threshold adopted in CF classification described later) are reported. We emphasize the following aspects:

- 1) Aspect: The modalities of this regressor not affect the phenomenon because the frequency distribution is uniformly distributed;
- 2) Land Use: The modality Forests (4) and Scrub and/or herbaceous vegetation (5) cover themselves almost 70% of areas of high susceptibility UCU, landslides developed in areas covered by vegetation;
- 3) DTM: Altitude up to 1100 meters is characterized by 85% of high susceptibility UCU. It is possible that the landslides occurring in high altitude are missing in the inventory;
- 4) Slope. As usual, the critical slope is between 10° and 40° ;
- 5) Lithology. The frequency of the modalities for the lithology variable are not extremely different, they are comprised between 9% and 18% with the exclusion of the modality Quaternary cemented deposits (6%).

The landslide susceptibility maps shown in **Figure 4(a)** and **Figure 4(b)** were realized using two different classification criteria which divided the territory into four susceptibility classes. The first is the JC classification algorithm, which optimizes the distribution of the classes, minimizing the variance within classes and maximize the variance between them.

The second is the CF method, which defines the limits between classes by constructing the success/prediction rate curves and use the ratio of effectiveness as an indicator of the validity of the prevision.

Table 3. Results of conditional analysis.

Variable	Classes	Variable description	Class frequency (UCU with SI > 2.45)
Aspect (°)	1	0 - 90	30.35
	2	90 - 180	22.54
	3	180 - 270	21.97
	4	270 - 360	25.14
Land Use	1	Artificial surfaces	4.34
	2	Permanent crops	4.34
	3	Agricultural areas	15.03
	4	Forests	36.42
	5	Scrub and/or herbaceous vegetation	32.66
	6	Open spaces with little or no vegetation	7.23
DTM (m a.s.l.)	1	449 - 700	25.14
	2	700 - 900	27.75
	3	900 - 1100	31.79
	4	1100 - 1300	10.98
	5	1300 - 1500	4.05
	6	1500 - 2150	0.29
Slope (°)	1	0 - 5	6.07
	2	5 - 10	11.56
	3	10 - 20	22.25
	4	20 - 30	21.97
	5	30 - 40	19.36
	6	40 - 50	16.76
	7	50 - 75	2.02
Lithology	1	Quaternary deposits	12.14
	2	Quaternary cemented deposits	6.07
	3	Flysch deposits	18.79
	4	Marly limestones	17.63
	5	Marls and clay marls	10.69
	6	Limestones, marls and clays alternations	13.58
	7	Stratified limestone	8.67
	8	Dolostones	12.43

A cross-validation methodology was used to evaluate the model [17] [32] [33] [34]. Each UCU was randomly divided in two subdataset, training (50% of the pixels) and testing (the other 50% of the pixels). The landslide susceptibility index was calculated on the training dataset and the UCU are ordered descending, accordingly. Two curves are plotted to assess the adaptation of the model to the data: the success and the prediction curves. The agreement between the two curves is measure of goodness of fit of the model (Figure 5). If the model estimated on the training dataset is correct, the curves present strong gradients in the initial portion and the percentage of area with landslide

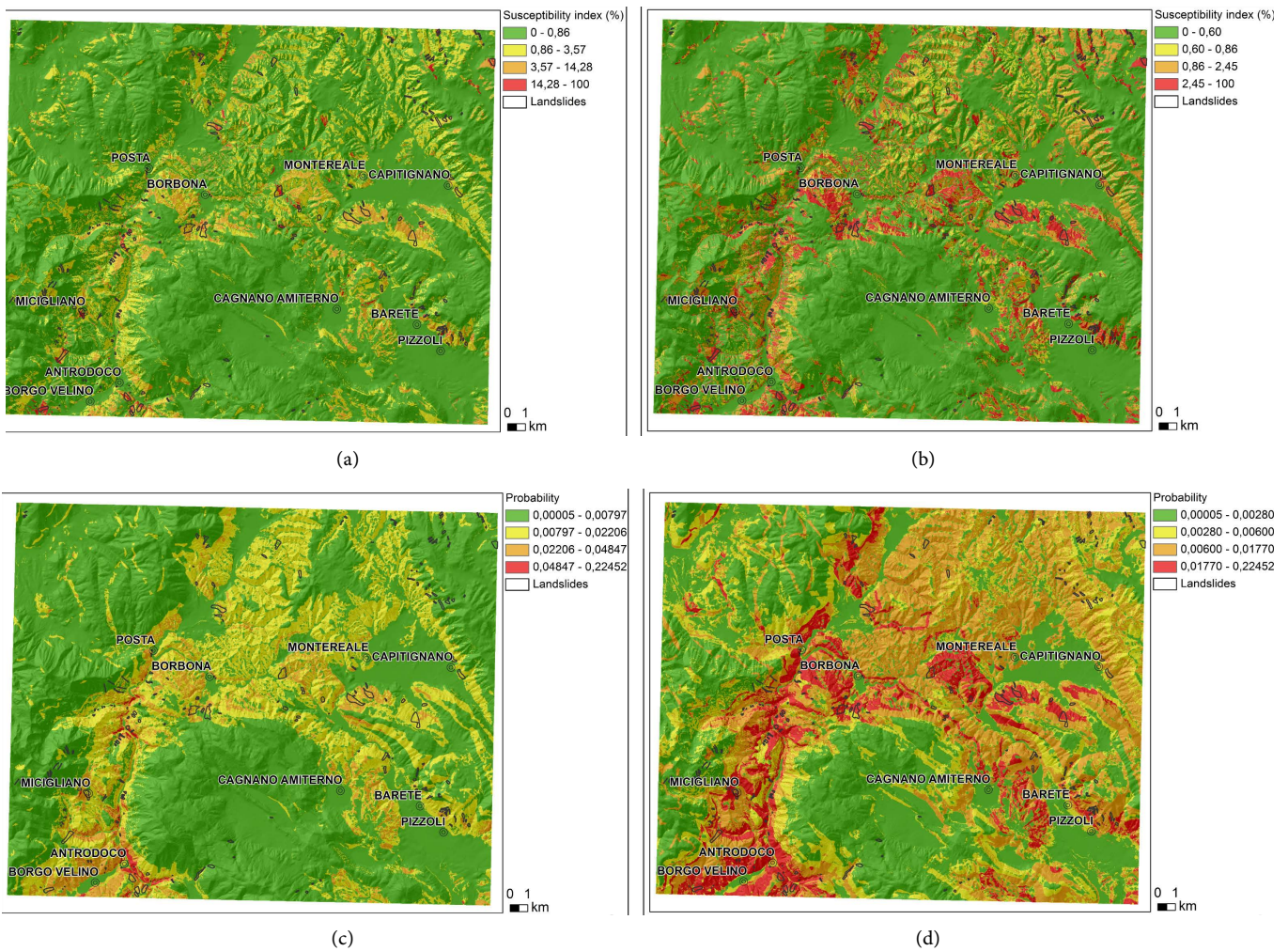


Figure 4. (a) Susceptibility map obtained through the conditional analysis and JC classification algorithm; (b) Susceptibility map obtained through the conditional analysis and CF classification method; (c) Susceptibility map obtained through logistic model and JC classification algorithm; (d) Susceptibility map obtained through logistic model and CF classification method.

predicted will be greater for the classes with greater susceptibility index. A trend according to a 45 degree line corresponds to a completely at random prediction and indicates that the portion of the area with landslide is constant for each area analyzed. The **Figure 5** reveals that 10% of the areas classified as the most susceptible correspond to about 30% of the territory actually affected by landslide.

The CF method defines two high discriminant classes: the higher one corresponding to a value for the ratio of effectiveness greater than 3; the lower one corresponding to a value less than 0.2. Two intermediate classes were defined in order to produce comparable susceptibility maps. The value which separates the two intermediate classes is fixed to 0.86 corresponding to a change in the slope of polyline (point B in the **Figure 5**).

5.2. Rare Events Logistics Regression

The logistic regression model associates the expected value of a dichotomous depen-

dent variable to the values (or modalities) of independent variables, via logit transform. In the present framework, the statistical analysis based on logistic regression model, identifies predictors significantly associated with the presence/absence of landslides and provides the estimates of landslide probability as a function of the geological variables previously defined.

If it is denoted by P the probability of occurrence of a landslide for a pixel with x_1, \dots, x_k geological features, the logistic model can be written as follow:

$$P(Y = 1 | x_1, \dots, x_k) = \frac{1}{1 + e^{-(\alpha + \beta_1 x_1 + \dots + \beta_k x_k)}}$$

The probability of landslide for the Antrodoco region, is underestimated by the logistic model due to the proportion of pixels affected by landslide to respect pixels not in landslide, which is about 6 for 1000 (9840 pixels in landslide for 1,520,450 total pixels). This problem is addressed by a specific approach, named rare events logic regression, in some statistical and geological papers [28] [35] [36].

The rare events logic regression by King and Zeng incorporates three correctives to the standard approach. The first correction concerns the sample: for each pixel affected by landslide, 5 non-landslides pixel have to be randomly sampled. The landslide pixels in the Antrodoco region are 9840, then $5 \times 9840 = 49200$ non-landslides pixels are randomly sampled and the total size of the sample is of 59,040 pixels. The sample is divided into two sub-samples in order to estimate and evaluate the model. The 80% of the pixels randomly sampled are the training sub-sample (47,232 pixels of which 7872 landslide pixels) and the remaining 20% of the sample is the test subsample (11,808 pixels of which 1968 landslide pixels).

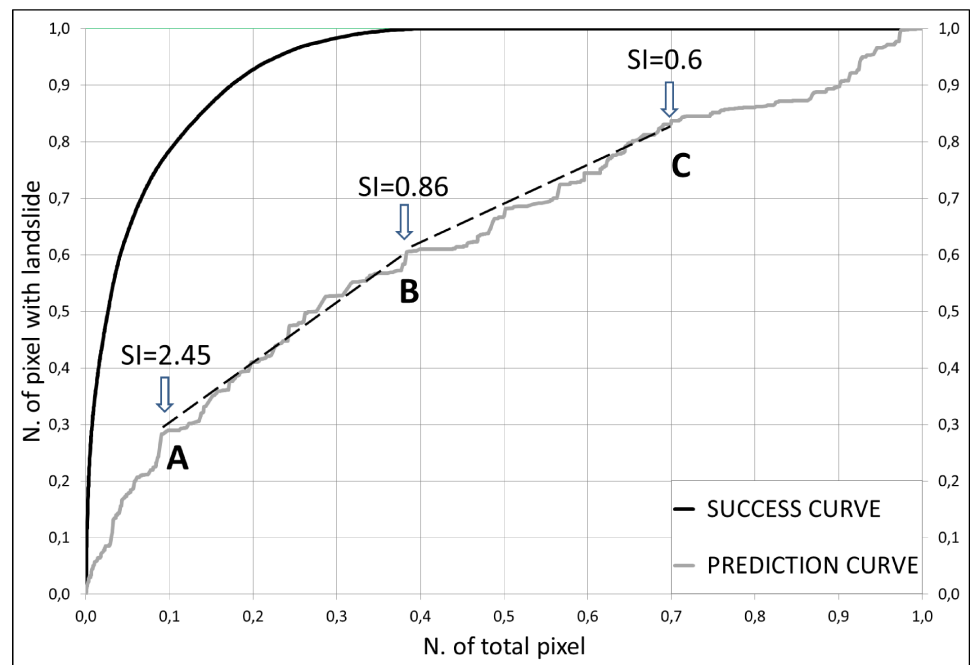


Figure 5. Validation of the results of the conditional analysis.

Fixing the proportion between landslide/non-landslide pixels requires the introduction of the two remaining correctives on estimating of the intercept and of the probabilities.

In particular, the correct intercept is calculated as:

$$\alpha = \hat{\alpha} - \ln \left[\left(\frac{1-\tau}{\tau} \right) \left(\frac{\bar{y}}{1-\bar{y}} \right) \right]$$

where τ is the fraction of landslide pixels in the area (amounting to 9840 of 1,520,450, approximately 0.00647), \bar{y} the proportion of landslide pixels in the sample and $\hat{\alpha}$ the biased estimate of the intercept.

Instead, the corrected probabilities of landslide are obtained as:

$$P(Y = 1 | x_1, \dots, x_k) = \tilde{p}_i + C_i$$

where the correction factor C_i is added to the to the biased value of standard probability \tilde{p} . The correction factor C_i is a function of the probability \tilde{p} , the vector of x and the covariance matrix of the vector of the estimators of the coefficient β [35].

The rare events logistic regression on the Antrodoco data was computed with the relogit function of the Zelig library [37] in R [38]. The final model is the following:

$$\text{Logit} \left(P(Y = 1 | d_1, \dots, d_5) \right) = \alpha + \beta_1 d_1 + \dots + \beta_5 d_5$$

where the regressors d are: the aspect (d_1), the land use (d_2), the DTM (d_3), the slope (d_4) and the lithology (d_5), indicated above with d_5 . In **Table 4** the estimates of the regression parameters, their standard errors, the corresponding value of the test z and the p -values are reported. The reference modalities are not shown.

All the considered variables are highly significant (p -value < 0.01), with the exception of “rare vegetation” modality (6) for the variable land use. The results concerning the aspect variable suggest a weak north vs. south effect, with negative association for south-facing pixels (2 and 3) and positive association to the landslide probability for the north-facing pixels (1 and 4). The land use variable shows a negative association for any significant modalities respect on the “urbanized areas” modality. Surprising, the results for the DTM variable suggests an increase in negative association between the probability of landslide and the increasing altitude, maybe due to the incompleteness of landslide inventory on the inaccessible sites higher located. The slope is positively associated with the phenomenon, strongly for the gradients between 30° and 75° (5, 6 and 7). Finally, for the lithology, the “Marls e clay marls” (5) is the strongest positive modality respect on the class of “quaternary deposits” (class reference 1), whereas “stratified limestone” (7) is the only negative one.

Receiver Operating Characteristic curve (ROC) for the test sample was used to evaluate the model. To this aim, for the 13,446 test pixels, predicted probabilities were computed and dichotomized using a grid of possible threshold values. For each threshold value, it was counted the number of the true positives (TP), *i.e.* the landslide pixels with a predicted probability greater than the threshold, the positives (P), *i.e.* total pixels with predicted probability greater than the threshold, the false negatives (FN), *i.e.*

Table 4. Results of rare events logistic regression: the estimates of the regression parameters, their standard errors, the corresponding value of the test statistic z and the p -values.

Variable	Classes	Variable description	Estimate (β)	Std. Error	z value	Pr ($> z $)
Intercept (α)			-5.40	0.15	-36.82	<0.01
Aspect ($^\circ$)	2	90 - 180	-0.44	0.03	-12.33	<0.01
	3	180 - 270	-0.66	0.04	-17.97	<0.01
	4	270 - 360	0.10	0.03	3.10	<0.01
Land Use	2	Permanent crops	-1.71	0.14	-12.26	<0.01
	3	Agricultural areas	-0.66	0.13	-5.08	<0.01
	4	Forests	-0.79	0.12	-6.33	<0.01
	5	Scrub and/or herbaceous vegetation	-0.30	0.13	-2.42	0.01
	6	Open spaces with little or no vegetation	0.20	0.16	-1.31	0.19
DTM (m a.s.l.)	2	700 - 900	-0.85	0.05	-16.73	<0.01
	3	900 - 1100	-1.00	0.05	-20.24	<0.01
	4	1100 - 1300	-1.66	0.06	-28.91	<0.01
	5	1300 - 1500	-2.07	0.07	-29.76	<0.01
	6	1500 - 2150	-2.09	0.09	-23.29	<0.01
Slope ($^\circ$)	2	5 - 10	1.51	0.09	17.15	<0.01
	3	10 - 20	2.26	0.08	28.01	<0.01
	4	20 - 30	2.08	0.08	25.33	<0.01
	5	30 - 40	2.30	0.09	26.67	<0.01
	6	40 - 50	2.98	0.10	29.43	<0.01
Lithology	6	50 - 75	2.47	0.24	10.34	<0.01
	2	Quaternary cemented deposits	0.86	0.11	8.14	<0.01
	3	Flysch deposits	0.76	0.06	13.46	<0.01
	4	Marly limestones	0.86	0.06	15.16	<0.01
	5	Marls and clay marls	2.10	0.09	23.62	<0.01
	6	Limestones, marls and clays alternations	1.10	0.06	17.21	<0.01
	7	Stratified limestone	-0.18	0.06	-3.02	<0.01
8	Dolomite	1.18	0.08	14.05	<0.01	

the landslide pixels with a predicted probability less than the threshold, and the negatives (N), *i.e.* total pixels with a predicted probability less than the threshold. ROC curves of **Figure 6** are obtained plotting sensitivity (TP/P) and 1-specificity (1-FN/N) for each threshold value of the grid [39] [40] [41] [42] [43].

The threshold value for predicted probability maximizing sensitivity and minimizing 1-specificity is fixed to 0.008918, which correspond sensitivity = 0.70 and 1-specificity = 0.68.

6. Comparison between Models

In the preceding paragraphs two approaches for the analysis of data of Antrodoco were

presented, one based on the conditional analysis and another on logistic regression model that produced for each UCU, *i.e.* for each combination of modalities of 5 regressors considered, the landslide susceptibility index and the logistic probability of landslide, respectively.

The operational requirement to differentiate four classes of susceptibility, involved the use of two methods of discretization of the continuous value of the estimated probabilities. In particular, the probability is estimated by non-parametric method, divided into four classes of susceptibility, using the JC method and that of CF approach, thus obtaining the two maps of **Figure 4(a)** and **Figure 4(b)**. Instead, the probability is estimated by logistic model for rare events, divided into four classes by the same two methods of classification, providing the two maps shown in **Figure 4(c)** and **Figure 4(d)**.

Table 5 shows the percentages of total pixels and pixels affected by landslide represented by the four maps, classified according to the susceptibility for each of the four analyzes.

It should be noted that the spatial domain has been divided into 4 ordered categories of susceptibility for cartographic purposes only. In the following, the analysis will proceed making particular reference to the extreme classes, capable of characterizing more strongly the obtained results. The threshold that bisects the landslide probability of the logistic model for rare events, calculated from the ROC curve in order to maximize sensitivity and specificity at the same time, falls into medium-low class for both classifications.

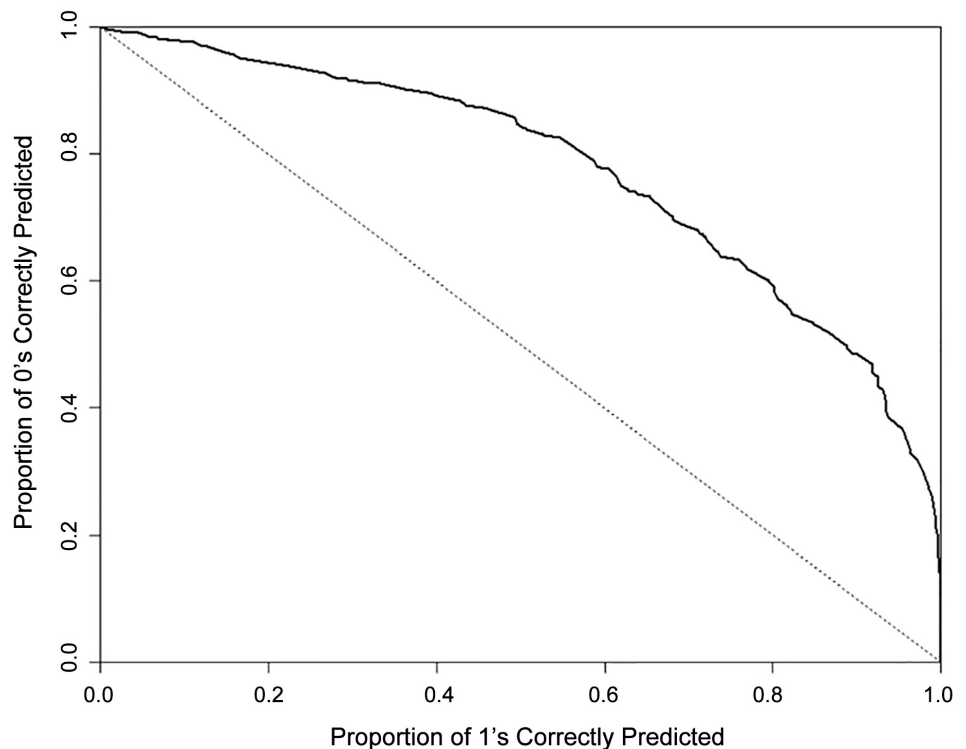


Figure 6. ROC curve for the test dataset.

Table 5. Analysis of the percentages of total pixels and pixels with landslide obtained in the four maps.

	Conditional model				Logistic model			
	Jenks & Caspall classes		Chung & Fabbri classes		Jenks & Caspall classes		Chung & Fabbri classes	
	N. of total pixel (%)	N. of pixel with landslides (%)	N. of total pixel (%)	N. of pixel with landslides (%)	N. of total pixel (%)	N. of pixel with landslides (%)	N. of total pixel (%)	N. of pixel with landslides (%)
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Susceptibility								
Low	84.4	13.9	83.0	9.5	69.0	29.8	40.0	6.0
Medium-low	8.0	36.2	1.4	4.3	24.5	40.9	20.3	13.4
Medium-high	4.8	34.5	6.1	29.0	5.7	25.0	29.9	41.2
High	2.8	15.4	9.4	57.1	0.8	4.4	9.9	39.4

Considering the columns (a) of **Table 5**, it can be seen that the susceptibility classes identified by the JC method seem to correctly describe the distribution of landslides in the territory, consisting of a high percentage of area with low landslide probability and a low percentage of area with high landslide probability, both for the conditional model, column (1a), and for logistic model, column (3a). For the latter, the difference in the distribution of the area in the four classes is particularly evident, changing by the JC classification to the CF one: the frequency distribution of the percentages shown in the column (3a) seems more realistic, due to the landslide rarity of the one shown in the column (4a). However, considering the presence of landslides in the area identified by the four susceptibility classes (columns (b) of the same table), it is noted that the number of landslides present in the areas classified in low and medium-low susceptibility, is much higher in the case of JC classes compared to those of CF, for both estimation models. In particular, for the conditional model, there are included in areas classified as low and medium-low susceptibility according to JC, the 50.1% of the landslides (column 1b), while in the case of the classification of CF (column 2b) only the 3.8% of the landslides are included. Even for the classification applied to the estimated probabilities using the logistic model, the percentages ranging from 70.7% of landslides in low and middle-low classes with low susceptibility according to JC algorithm, at 19.4% of landslides with the classification according to the method of CF.

Then, the method of JC seems to produce for the data, a classification less faithful than that obtained by the method of CF, for both methods.

The probability, estimated by the conditional model with CF classification method, attributed low susceptibility to 83% of the pixels, which corresponds to 9.5% of those with landslide, while classifying 57.1% as high susceptibility for 9.4% of the area. For the probability estimated by the rare events logistic model, only 40% of the area is indicated at low susceptibility, among which are included 6% of the total pixels affected by landslide, while in the high susceptibility class are comprised only 39.4% of the pixels with landslide, in an area equal to 9.9% of the total.

Then, the conditional estimation method, especially in conjunction with the CF me-

thod of classification, discriminates better between pixels with landslide and without landslide compared to logistic estimates with both classifications

7. Conclusions

Present paper aims to produce the susceptibility maps of slope instability at 1:50,000 scale of the Antrodoco area. This mapping scale is not a detailed scale on which perform territorial planning; however it reduces the costs of zonation and limits the specific engineering and geotechnical investigations only in high susceptibility areas.

The information available was synthesized in 9 variables at the pixel level: DTM, aspect, curvature, flow accumulation, land use, pluviometry, slope, seismic acceleration and lithology. The continuous variables were appropriately discretized. Using the value of Cramer's V statistic, the 5 variables most associated with landslide presence were identified (lithology, DTM, slope, aspect, land use), determining a manageable number of UCU, *i.e.* group of pixels with same combination of variable modalities.

The probability of landslide, for each UCU, was estimated using two approaches: the conditional model and the rare event logistic regression model. The first compares the number of pixels affected by landslide to the number of pixels of the area (susceptibility index). Using this approach it is not necessary to make assumptions about the phenomenon; however the role of a singular regressor in the probability level cannot be assessed and it is not possible to evaluate standard errors.

The second approach provided estimates and level of significance for the association between each regressor and the landslide presence, using a logistic model adapted to the case of rare events. In this case, it is possible to evaluate the importance of each of them in determining the overall level of landslide susceptibility.

The predictive ability of the landslide probability obtained from the two approaches, was evaluated by the success-prediction curves for the conditional analysis, and by the ROC curve, for the logistic model. For both approaches, an acceptable predictive ability was attained.

The results show that morphological characteristics of the territory, as north-facing aspect and high slope gradient (values between 30° and 75°), and physico-mechanical properties of lithotype (in particular marls and clay marls) are determinant factors for the landslide susceptibility.

The negative association between DTM variable and landslide presence is probably due both to the low quality of the data obtained from national public database (cell size 20 × 20 m), and to the incompleteness of landslide inventory on the inaccessible sites higher located.

The land use variable shows a negative association for any significant modalities respect to the "urbanized areas" modality showing a critical aspect for this category of land use.

The landslide probability was divided into four susceptibility categories (high, medium-high, medium-low and low), by two classification methods: the Natural Breaks algorithm, automatically implemented on GIS systems, and CF method. Four suscepti-

bility maps were produced, applying the two classification methods to the both estimating models.

The map showing higher discriminant ability in explaining the distribution of landslide is obtained by conditional analysis in conjunction with CF classification method. In the case of susceptibility classes constructed by JC the ability of the method to identify the UCU with higher number of landslides, noticeably worsens. The observation remains valid in the case of susceptibility classes built starting from the probability of the logistic model, although the descriptive ability of the model has values less high than the previous one. Thus, the quality of the susceptibility map is dependent not only from the prediction model implemented but also from the classification method adopted.

Acknowledgements

The authors would like to thank Dr Paola Giordano (National Institute of Statistics), for useful discussion and support.

References

- [1] Hammond, C., Hall, D., Miller, S. and Swetik, P. (1992) Level I Stability Analysis (LISA) Documentation for Version 2.0. General Technical Report INT-285, USDA, Forest Service, Intermountain Research Station, Ogden.
- [2] Akbari, A., Yahaya, F., Azamirad, M. and Fanodi, M. (2014) Landslide Susceptibility Mapping Using Logistic Regression Analysis and GIS Tools. *Electronic Journal of Geotechnical Engineering*, **19**, 1687-1696. <http://www.ejge.com/2014/Ppr2014.161mar.pdf>
- [3] Regmi, N.R., Giardino, J.R., McDonald, E.V. and Vitek, J.D. (2014) A Comparison of Logistic Regression-Based Models of Susceptibility to Landslides in Western Colorado, USA. *Landslides*, **11**, 247-262. <https://doi.org/10.1007/s10346-012-0380-2>
- [4] Bai, S.B., Wang, J., Lü, G.N., Zhou, P.G., Hou, S.S. and Xu, S.N. (2010) GIS-Based Logistic Regression for Landslide-Susceptibility Mapping of the Zhongxian Segment in the Three Gorges Area, China. *Geomorphology*, **115**, 23-31. <https://doi.org/10.1016/j.geomorph.2009.09.025>
- [5] Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M. and Ardizzone, F. (2005) Probabilistic Landslide Hazard Assessment at the Basin Scale. *Geomorphology*, **72**, 272-299. <https://doi.org/10.1016/j.geomorph.2005.06.002>
- [6] Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F. and Cardinali, M. (2006) Landslide Hazard Assessment in the Collazzone Area, Umbria, Central Italy. *Natural Hazards and Earth System Sciences*, **6**, 115-131. <https://doi.org/10.5194/nhess-6-115-2006>
- [7] Pradhan, B. (2013) A Comparative Study on the Predictive Ability of the Decision Tree, Support Vector Machine and Neuro-Fuzzy Models in Landslide Susceptibility Mapping Using GIS. *Computers & Geosciences*, **51**, 350-365. <https://doi.org/10.1016/j.cageo.2012.08.023>
- [8] Lee, S., Jeon, S.W., Oh, K.-Y. and Lee, M.-J. (2016) The Spatial Prediction of Landslide Susceptibility Applying Artificial Neural Network and Logistic Regression Models: A Case Study of Inje, Korea. *Open Geosciences*, **8**, 117-132. <https://doi.org/10.1515/geo-2016-0010>
- [9] Falaschi, F., Giacomelli, F., Federici, P.R., Puccinelli, A., D'Amato Avanzi, G., Pochini, A.

- and Ribolini, A. (2009) Logistic Regression versus Artificial Neural Networks: Landslide Susceptibility Evaluation in a Sample Area of the Serchio River Valley, Italy. *Natural Hazards*, **50**, 551-569. <https://doi.org/10.1007/s11069-009-9356-5>
- [10] Pradhan, B. and Lee, S. (2010) Landslide Susceptibility Assessment and Factor Effect Analysis: Back Propagation Artificial Neural Networks and Their Comparison with Frequency Ratio and Bivariate Logistic Regression Modeling. *Environmental Modelling & Software*, **25**, 747-759. <https://doi.org/10.1016/j.envsoft.2009.10.016>
- [11] D'Ambrogi, C., Pantaloni, M. and Pichezzi, R.M. (2010) I 20 anni del Progetto di cartografia geologica nazionale. *Memorie Descrittive della Carta Geologica d'Italia*, **88**.
- [12] Amanti, A. (2010) Integrazioni geotematiche al rilevamento geologico: Il caso del foglio "Antrodoco". *Memorie Descrittive della Carta Geologica d'Italia*, **88**, 39-46.
- [13] Amanti, A., Chiessi, V., Guarino, P.M. and Serafini, R. (2010) Landslides Cartography in "foglio Antrodoco" Project: Progress Report and Result. *Memorie Descrittive della Carta Geologica d'Italia*, **88**, 139.
- [14] Brabb, E. (1984) Innovative Approaches for Landslide Hazard Evaluation. *IV International Symposium on Landslides*, Toronto, 307-323.
- [15] Corominas, J., Leroi, E. and Savage, W.Z. (2008) Guidelines for Landslide Susceptibility, Hazard and Risk Zoning for Land Use Planning. *Engineering Geology*, **102**, 85-98. <https://doi.org/10.1016/j.enggeo.2008.03.022>
- [16] Jenks, G.F. and Caspall, F.C. (1971) Error on Choroplethic Maps: Definition, Measurement, Reduction. *Annals of the Association of American Geographers*, **61**, 217-244. <https://doi.org/10.1111/j.1467-8306.1971.tb00779.x>
- [17] Chung, C.F. and Fabbri, A.G. (2003) Validation of Spatial Prediction Models for Landslide Hazard Mapping. *Natural Hazards*, **30**, 451-472. <https://doi.org/10.1023/B:NHAZ.0000007172.62651.2b>
- [18] Pantaloni, M., Capotorti, F., D'Ambrogi, C. and Di Stefano, R. (2004) Geological Guide of the 348 Antrodoco Sheet ISPRA. Internal Document, ISPRA, Italy.
- [19] Berti, D., et al. (2009) La Geologia del Foglio n. 348 Antrodoco. *Memorie Descrittive Carta Geologica d'Italia*, **88**, 134.
- [20] ISPRA Institute for Environmental Protection and Research (2008) Landslides in Italy, Special Report 2008, **83**, ISPRA, Italy.
- [21] SGI Servizio Geologico d'Italia (1955) Carta Geologica d'Italia Scala 1:100,000, Foglio n. 139 L'Aquila.
- [22] Meletti, C. and Montaldo, V. (2007) Stime di pericolosità sismica per diverse probabilità di superamento in 50 anni: Valori di ag, Progetto DPC-INGV S1, Deliverable D2. <http://esse1.mi.ingv.it/d2.html>
- [23] Longley, P.A. and Batty, M. (2003) Advanced Spatial Analysis: The CASA Book of GIS. ESRI, Redlands, CA.
- [24] McCoy, J. (2004) ArcGIS 9 Geoprocessing in ArcGIS. ESRI, Redlands, CA.
- [25] Gumbel, E.J. (1958) Statistics of Extremes. Columbia University Press, New York.
- [26] Carrara, A., Cardinali, M., Guzzetti, F. and Reichenbach, P. (1995) GIS Technology in Mapping Landslide Hazard. In: Carrara, A. and Guzzetti, F., Eds., *Geographical Information Systems in Assessing Natural Hazards*, Kluwer Academic Publisher, Dordrecht, 135-176. https://doi.org/10.1007/978-94-015-8404-3_8
- [27] Chung, C.F., Fabbri, A.G. and Van Westen, C.J. (1995) Multivariate Regression Analysis for Landslide Hazard Zonation. In: Carrara, A. and Guzzetti, F., Eds., *Geographical Informa-*

- tion Systems in Assessing Natural Hazards*, Kluwer Academic Publisher, Dordrecht, 107-134. https://doi.org/10.1007/978-94-015-8404-3_7
- [28] Van Den Eeckhaut, M., Reichenbach, P., Guzzetti, F., Rossi, M. and Poesen, J. (2009) Combined Landslide Inventory and Susceptibility Assessment Based on Different Mapping Units: An Example from the Flemish Ardennes, Belgium. *Natural Hazards and Earth System Sciences*, **9**, 507-521. <https://doi.org/10.5194/nhess-9-507-2009>
- [29] Hansen, A. (1984) Landslide Hazard Analysis. In: Brunsen, D. and Prior, D.B., Eds., *Slope Instability*, John Wiley and Sons, New York, 523-602.
- [30] Kendall, M. and Stuart, A. (1979) *The Advanced Theory of Statistics: Inference and Relationship*. Hodder Arnold, London.
- [31] Reichenbach, P., Guzzetti, F. and Carrara, A. (2002) Special Issue on Assessing and Mapping Landslide Hazards and Risk. *Natural Hazards and Earth System Science*, **2**, 1-2.
- [32] Van Westen, C.J., Rengers, N. and Soeters, R. (2003) Use of Geomorphological Information in Indirect Landslide Susceptibility Assessment. *Natural Hazards*, **30**, 399-419. <https://doi.org/10.1023/B:NHAZ.0000007097.42735.9e>
- [33] Brenning, A. (2005) Spatial Prediction Models for Landslide Hazards: Review, Comparison and Evaluation. *Natural Hazards and Earth System Sciences*, **5**, 853-862. <https://doi.org/10.5194/nhess-5-853-2005>
- [34] Chung, C.F. and Fabbri, A.G. (2008) Predicting Landslides for Risk Analysis Spatial Models Tested by a Cross-Validation Technique. *Geomorphology*, **94**, 438-452. <https://doi.org/10.1016/j.geomorph.2006.12.036>
- [35] King, G. and Zeng, L. (2001) Logistic Regression in Rare Events Data. *Political Analysis*, **9**, 137-163. <https://doi.org/10.1093/oxfordjournals.pan.a004868>
- [36] Van Den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten, G. and Vandekerckhove, L. (2006) Prediction of Landslide Susceptibility Using Rare Events Logistic Regression: A Case-Study in the Flemish Ardennes (Belgium). *Geomorphology*, **76**, 392-410. <https://doi.org/10.1016/j.geomorph.2005.12.003>
- [37] Imai, K., King, G. and Lau, O. (2007) Relogit: Rare Events Logistic Regression for Dichotomous Dependent Variables. In: Imai, K. and King, G., Eds., *Zelig: Everyone's Statistical Software*, 493-502. <http://r.iq.harvard.edu/docs/zelig.pdf>
- [38] R Core Team (2012) *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing. Vienna, Austria. <http://wwwR-project.org>
- [39] Lasko, T.A., Bhagwat, J.G., Zou, K.H. and Ohno-Machado, L. (2005) The Use of Receiver Operating Characteristic Curves in Biomedical Informatics. *Journal of Biomedical Informatics*, **38**, 404-415. <https://doi.org/10.1016/j.jbi.2005.02.008>
- [40] Begueria, S. (2006) Validation and Evaluation of Predictive Models in Hazard Assessment and Risk Management. *Natural Hazard*, **37**, 315-329. <https://doi.org/10.1007/s11069-005-5182-6>
- [41] Fawcett, T. (2006) An Introduction to ROC Analysis. *Pattern Recognition Letters*, **27**, 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- [42] Petschko1, H., Brenning, A., Bell, R., Goetz1, J., and Glade, T. (2014) Assessing the Quality of Landslide Susceptibility Maps—Case Study Lower Austria. *Natural Hazards and Earth System Sciences*, **14**, 95-118. <https://doi.org/10.5194/nhess-14-95-2014>
- [43] Chalkias, C., Ferentinou, M. and Polykretis, C. (2014) GIS-Based Landslide Susceptibility Mapping on the Peloponnese Peninsula, Greece. *Geosciences*, **4**, 176-190. <https://doi.org/10.3390/geosciences4030176>

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact gep@scirp.org