land

MDPI

Article

Integrating Multivariate (GeoDetector) and Bivariate (IV) Statistics for Hybrid Landslide Susceptibility Modeling: A Case of the Vicinity of Pinios Artificial Lake, Ilia, Greece

Christos Polykretis *, Manolis G. Grillakis, Athanasios V. Argyriou, Nikos Papadopoulos and Dimitrios D. Alexakis

Laboratory of Geophysical—Satellite Remote Sensing and Archaeo-Environment (GeoSat ReSeArch Lab), Institute for Mediterranean Studies (IMS), Foundation for Research and Technology—Hellas (FORTH), 74100 Rethymno, Greece; grillakis@hydrogaia.gr (M.G.G.); nasos@ims.forth.gr (A.V.A.); nikos@ims.forth.gr (N.P.); dalexakis@ims.forth.gr (D.D.A.)

* Correspondence: polykretis@ims.forth.gr; Tel.: +30-283-110-6021

Abstract: Over the last few years, landslides have occurred more and more frequently worldwide, causing severe effects on both natural and human environments. Given that landslide susceptibility (LS) assessments and mapping can spatially determine the potential for landslides in a region, it constitutes a basic step in effective risk management and disaster response. Nowadays, several LS models are available, with each one having its advantages and disadvantages. In order to enhance the benefits and overcome the weaknesses of individual modeling, the present study proposes a hybrid LS model based on the integration of two different statistical analysis models, the multivariate Geographical Detector (GeoDetector) and the bivariate information value (IV). In a GIS-based framework, the hybrid model named GeoDIV was tested to generate a reliable LS map for the vicinity of the Pinios artificial lake (Ilia, Greece), a Greek wetland. A landslide inventory of 60 past landslides and 14 conditioning (morphological, hydro-lithological and anthropogenic) factors was prepared to compose the spatial database. An LS map was derived from the GeoDIV model, presenting the different zones of potential landslides (probability) for the study area. This map was then validated by success and prediction rates—which translate to the accuracy and prediction ability of the model, respectively. The findings confirmed that hybrid modeling can outperform individual modeling, as the proposed GeoDIV model presented better validation results than the IV model.

Keywords: landslides; susceptibility; hybrid modeling; Geographical Detector; information value; Greece

Citation: Polykretis, C.; Grillakis, M.G.; Argyriou, A.V.; Papadopoulos, N.; Alexakis, D.D. Integrating Multivariate (GeoDetector) and Bivariate (IV) Statistics for Hybrid Landslide Susceptibility Modeling: A Case of the Vicinity of Pinios Artificial Lake, Ilia, Greece. *Land* **2021**, *10*, 973. https://doi.org/10.3390/land10090973

Academic Editors: Enrico Miccadei, Cristiano Carabella and Giorgio Paglia

Received: 20 August 2021
Accepted: 13 September 2021
Published: 15 September 2021

1. Introduction

A landslide is a gravity-driven environmental process which involves the movement of rocks, debris, earth, or a combination of them down a slope [1]. According to official data, landslides constituted the third (after floods and storms, and before earthquakes) most frequent natural disaster worldwide in 2020 [2]. Generally, the extreme weather events due to climate change, and the high seismic activity in combination with the poorly planned expansion of human activities (deforestation of slopes, uncontrolled irrigation, etc.), have contributed to a global upward tendency in landslide occurrence in the recent years [3].

Due to their occurring without warning and seriously threatening both natural and human environments, landslides are a major problem. Due to severe damage, or even destruction, of infrastructure and properties, they generate larger annual economic losses (billions euro) than any other natural disaster in many countries. In addition, a considerable number of people each year are injured and, in some cases, killed by them. It is indicative that during 1998–2017, totally 4.8 million people were affected by landslides worldwide, with 18,414 of them being killed [4]. In addition, the environmental effects of landslides are mainly changes in terrain morphology, and increased sediment loads in rivers and subsequent transport to dams.

The increased frequency of landslides and the severity of their effects have led to growing interest from international scientific community. Since predictions of occurrence and intensity remain challenging, most of the attention has been given to the determination of potential spatial locations. The acquisition of this spatial information can be achieved through landslide susceptibility (LS) assessments and mapping. LS refers to the potential landslide activity as a result of terrain conditions [5]. An assessment depends on the spatial distribution of past landslides in an area and their relation to its terrain conditions, in order to generate spatial predictions for areas that are not landslide-affected but have similar conditions. The output is a map presenting the region of interest divided into homogeneous zones of susceptibility [6]. LS maps with high levels of accuracy and reliability are considered crucial tools that can then be used as inputs for disaster management plans.

The advancements in the geospatial tools of geographic information systems (GIS) and remote sensing (RS), assisted by improvements in computer processing power, have improved LS modeling over the last few decades. Based on the literature, a considerable number of models are currently available for assessing LS at different spatial scales. In terms of degree of objectivity and necessity for landslide occurrence data, all these models can be separated into two different groups, the qualitative and quantitative models. The qualitative (or semi-quantitative) models estimate a susceptibility score on the basis of weights assigned to landslide conditioning factors from one or more expert(s). They suffer from low objectivity associated with the experts' subjective judgements [7]. On the other hand, the quantitative models decrease bias in the weight assignments, since they depend on fixed mathematical rules, regardless of any expert judgement [8]. Particularly, the impacts of different conditioning factors on past occurrences are quantitatively determined, resulting in high objectivity.

The current capability for acquiring multi-temporal landslide occurrence data through RS-based approaches has led to wide use of the data-driven quantitative models. These models range from complicated geotechnical and advanced machine learning models to more conventional statistical analysis models. Based on mechanical laws for the calculation of a safety factor, the geotechnical models [9,10] examine the slope stability from the perspective of the mechanical properties of the slope. Being based on human learning procedures, machine learning models are used to solve problems characterized by nonlinear functions and data. Commonly applied machine learning models are artificial neural networks (ANN), support vector machines (SVM), random forests (RF) and decision trees (DT) [11–13].

Regarding statistical analysis models, their fundamental principle is to estimate the probability of a landslide under the existence of spatial associations between the conditioning factors and past landslides [14]. Depending on the examination of factors individually or cumulatively, they can be either bivariate or multivariate. In bivariate modeling, weights are calculated for the classes of each individual factor by their levels of association with landslides in a historic dataset. Frequency ratio (FR), information value (IV) and weights of evidence (WoE) constitute the main representatives of bivariate models [15,16]. Conversely, in multivariate modeling, all the factors are sampled, and the presence or absence of landslide is determined for each of the sampling units [17]. Then, weights are calculated for the factors via statistical means. Among the multivariate models, logistic regression (LR) is doubtless the most used [18,19]. However, models such as LR consider the factors as explanatory variables without taking into account the spatial information contained in them and exploring their impacts on landslide occurrence (dependent variable) from a spatial perspective. In order to overcome this limitation, new spatially-based multivariate models have been put forward recently. These models can address the specificities of each space and consider that spatial variations in landslides may cause different responses to variations in the factor variables. Such a model is the Geographical Detector (GeoDetector). Although GeoDetector has been tested in various studies of health, social and environmental sciences [20–22] over the last few years, its use in landslide-related research has been quite limited. Since it provides an effective way to identify and eliminate redundant

variables, GeoDetector has been used in a few relevant studies [14,23,24] for factor selection purposes.

In general, all the quantitative models have been proven beneficial for identifying locations that are prone to landslides; however, some shortcomings still characterize them. The geotechnical models require detailed mechanical data of soil or rock, and as a result they are only suitable for studying small regions or single slopes. Although the statistical models are easy to understand and perform well in most cases, they find it difficult to solve situations with large amounts of data. Moreover, despite their ability to handle large amounts of nonlinear data, machine learning models are not significantly better than the statistical ones, and cannot perform well under different conditions and in different areas [25]. In order to produce the most reliable LS map for a region of interest, one possible solution is to compare different models and select the optimum in terms of accuracy and prediction ability. Several studies have compared two or more different models to recognize the most suitable for a specific region [26–28].

The aforementioned shortcomings tend to increase the uncertainty and reduce the efficiency of models when applied individually. Thus, another solution has gained popularity recently, which is the development of hybrid models. Hybrid modeling can resolve the shortcomings of individual models and improve performance. This type of modeling has been gradually applied in LS assessment studies over the last decade. For instance, in the work of Arabameri et al. [8], the efficiency of the integration of statistical (FR) and machine learning (RF) models was explored for LS mapping in northern Iran. For assessing the LS in a region of India, Saha et al. [29] integrated a statistical and a machine learning model to improve on their individual accuracies. Chen et al. [30] applied a combination of bivariate (WoE) and multivariate (LR) statistical models with a machine learning model (RF) for LS mapping of a mountainous region of China. Roy et al. [31] delineated LS zones in districts of India by integrating bivariate statistical (WoE) and machine learning (SVM) models. Chowdhuri et al. [32] introduced hybrid models from statistical and machine learning model integrations for predicting spatially the landslide occurrence in a basin of India. In addition, some studies have improved the performances of machine learning models by combining them with optimization or meta-heuristic algorithms [33,34].

In Greece, landslide activity has been highly facilitated by the frequent occurrence of intense rainfall and seismic events. Along with them, its complex geo-morphological settings (strained geological formations and steep slopes) and the uncontrolled land-use in landslide-prone areas have contributed. As a result, the interest in and awareness of the importance of LS assessments for regions of Greece have increased, particularly over the last decade. However, the majority of relevant studies has focused on the implementation of individual statistical and machine learning models [35–37], rather than integrated approaches. It could be mentioned that the work of Chalkias et al. [38] constitutes an exception.

The region of Peloponnese has experienced severe natural disasters, including floods, earthquakes, landslides and wildfires. Specifically, landslides have highly damaged settlements within its boundaries (mainly in its northern and western parts), resulting in partial destruction and necessary re-locations to nearby geologically stable lands. Considering all the above, the present study aimed to assess the LS and create a reliable map of a wetland in northwestern Peloponnese. Therefore, a hybrid LS modeling is proposed based on the integration of two different statistical models, the multivariate GeoDetector and bivariate IV. Past landslide occurrence and conditioning factor datasets were incorporated into the hybrid model, named GeoDIV, and analyzed in a GIS environment to determine the spatial distribution of susceptibility. In order to confirm the targeted reliability of LS map, the performance of proposed GeoDIV model was compared with that of the individual IV model in a validation procedure.

2. Study Area

The surrounding area of the Pinios artificial lake was selected for investigation in this study. It is located in the western part of Greece and the northwestern part of geographical area of Peloponnese (Figure 1), covering a total extent of approximately 239 km^2. It belongs administratively to the Prefecture of Ilia and hydrologically to the drainage basin of Pinios River. The boundaries of the study area are defined in the north and south by the basin's boundaries, and in the west and east by altitude contours of 100 and 200 m, respectively. The Pinios artificial lake was created in 1960, after the construction of a dam on the homonym river, and is the largest in Peloponnese (with a total extent of approximately 20 km^2). Its water is used for the irrigation of the plain of Ilia, and hence it is considered one of the most important land improvement projects in the entire prefecture. The total quantity of water withdrawn from the lake annual for irrigation and water supply purposes amounts to 126 million m^3.

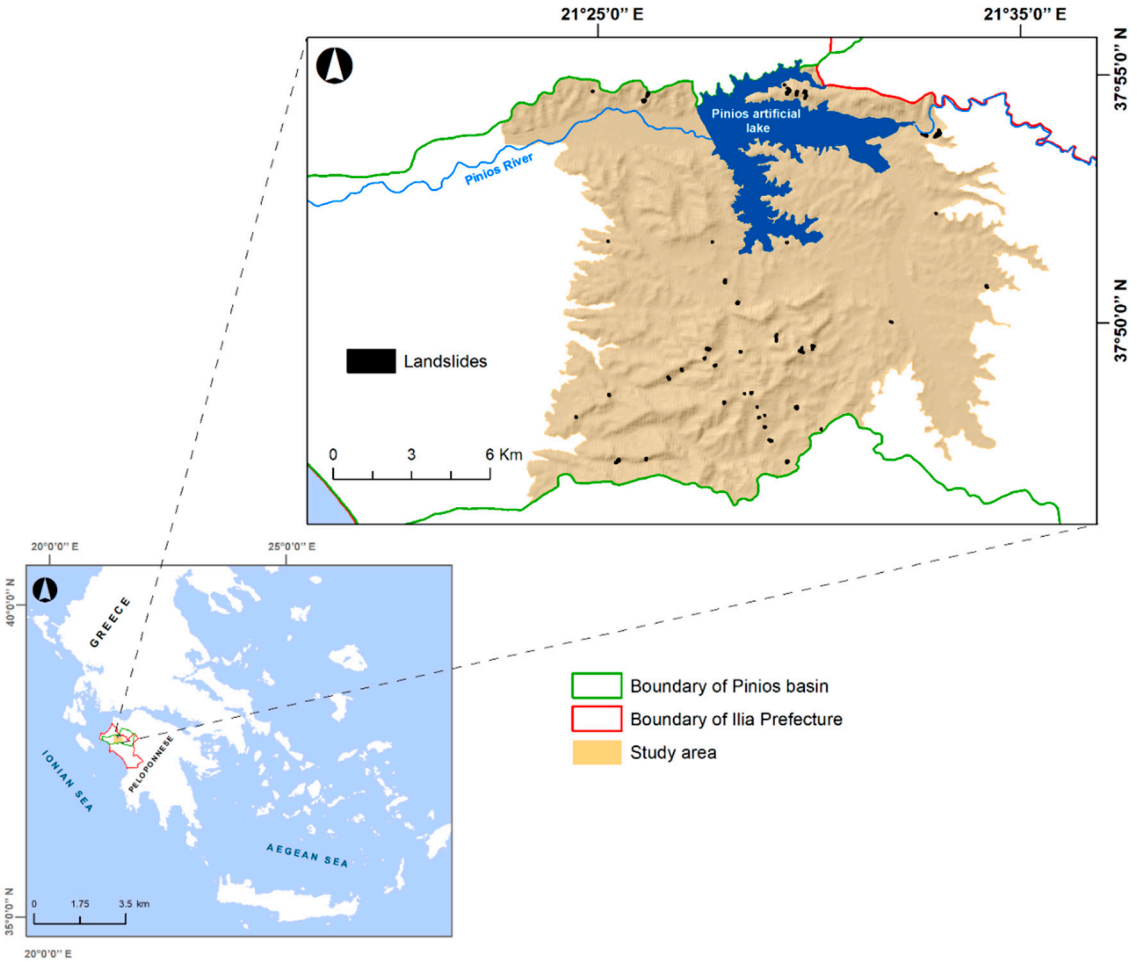


Figure 1. The study area and the locations of events from the landslide inventory.

Following the typical landscape of Ilia Prefecture, the study area can be characterized as an agricultural region at a low altitude (mean altitude at 154 m above sea level). Heterogeneous croplands or fields mixed with natural vegetation represent the predominant agricultural lands. More than 30 settlements are situated within its boundaries, containing 5400 inhabitants according to the official 2011 census [39].

The climate is Mediterranean mild with a mean temperature ranging from 20 to 25 °C in the summer months, and from 4 to 10 °C in the winter months [40]. Long-term rainfall records including the period of the last two decades show a mean annual value reaching approximately 500 mm. From a geological perspective, the study area is mainly covered by

Neogene and Quaternary loose deposits varying in thickness and consistency. Confined granular aquifer systems have been formed inside alluvial deposits, and unconfined aquifers have been developed in Quaternary deposits where groundwater flows to the direction of the sea [41].

3. Data and Methods

In this study, a hybrid model was developed for LS assessment based on the integration of two different individual models, the GeoDetector and IV. A spatial database was created in GIS to be used in hybrid modeling, including: (a) the landslide inventory dataset and (b) the conditioning factor datasets.

3.1. Landslide Inventory

Such a dataset provides information about the landslide events that occurred in the past in a given region. Hence, this information is crucial for any quantitative LS modeling effort. A database maintained by the Laboratory of Engineering Geology at the Department of Geology at University of Patras referring to landslides that occurred between 2000 and 2015 [42], and field surveys, were initially exploited for the spatial locations of past landslides in the study area. Then, multi-temporal Google Earth satellite imagery (Figure 2) was used for their delimitation. Based on the classification proposed by Varnes [1], for this study, the term landslide included shallow debris flows and earth rotational slides, varying in extent from some hundreds to several thousands of meters squared (Table 1). Since it is not always possible to differentiate the depletion and accumulation zones of these landslide types in an inventory map [18], these zones were mapped together in an entire area forming a single polygon feature for each landslide. Therefore, 60 landslide polygons were eventually represented in the relevant inventory map (Figure 1).

Table 1. Types and basic morphometrical parameters of the landslides in the study area.

Landslide Type	Amount of Events	% of Total Landslide Events	Area (sq. m)			Altitude (m)			Slope Angle (Degrees)		
			Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
Debris flows	40	67	4187	103	1268	348	110	222	60	15	26
Earth rotational slides	20	33	18,000	240	4188	332	89	199	45	15	30

3.2. Conditioning Factors

Landslide occurrence is considered to be affected by a variety of natural and anthropogenic factors representing the conditions of a given region. These conditioning factors can be separated into two main categories: (a) the preparatory factors which create suitable conditions for a landslide by changing the state of a slope from stable to marginally stable, and (b) the triggering factors which initiate a landslide by changing the state of a slope from marginally stable to unstable [43]. Morphological and hydro-lithological conditions of the region of interest are represented by natural preparatory factors, whereas the human interventions on it are represented by anthropogenic preparatory factors. The triggering factors mainly represent climatic and seismic conditions related to rainstorms and earthquakes, respectively.

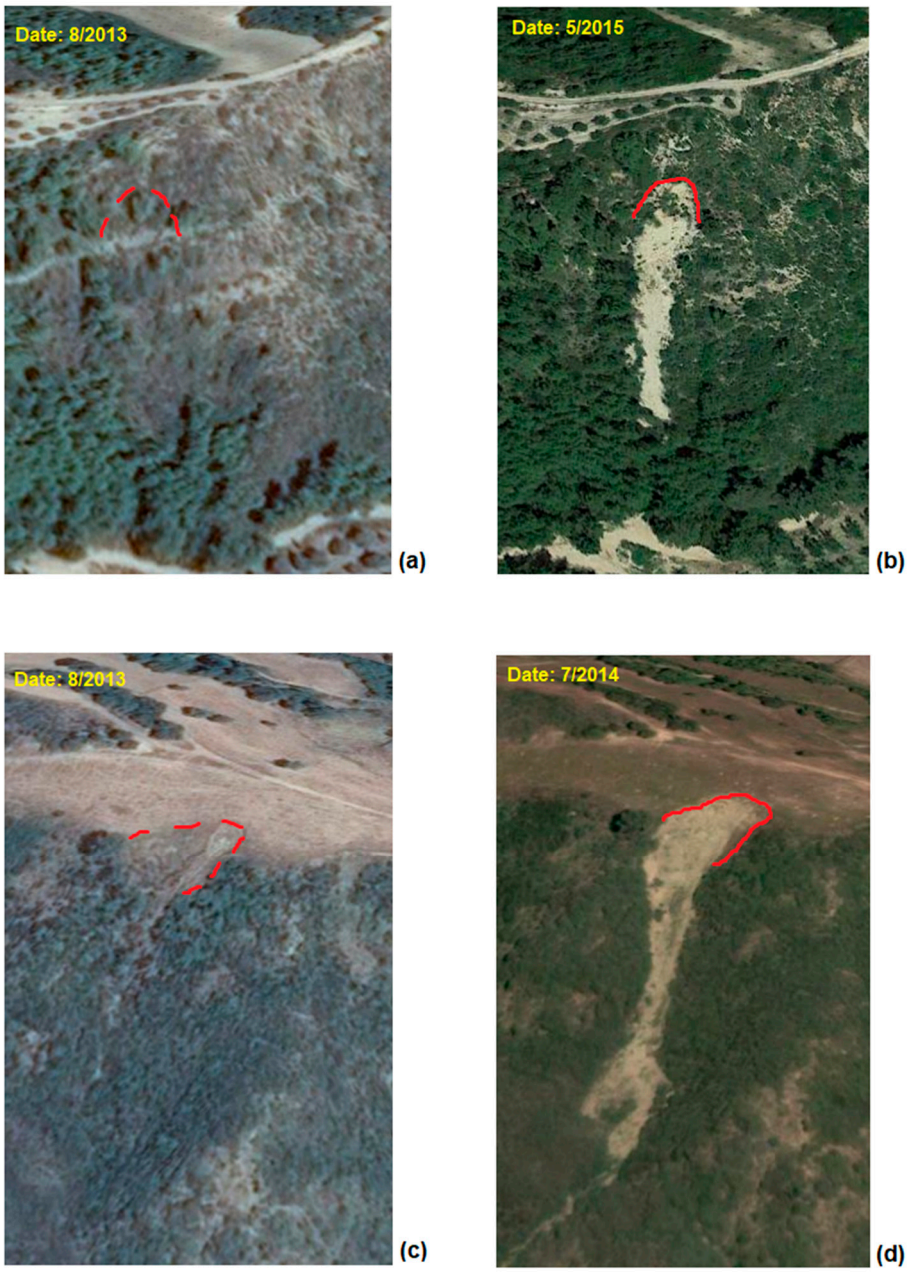


Figure 2. Multi-temporal Google Earth images: (**a**,**c**) before the landslides; (**b**,**d**) after the landslides. The red dashed lines indicate the location of the landslide before it happened, and the red solid line shows the scar of the landslide after it occurred.

Since no official guidelines are used by the scientific community for the selection of factors, the characteristics of the study area, data availability and a literature review [29,30] were taken into account for this study. In total, fourteen conditioning factors were selected, including both preparatory and triggering factors. In particular, the altitude, slope angle, slope aspect, profile curvature, plan curvature, stream density, stream power index (SPI), topographical wetness index (TWI), lithology, proximity to faults and soil type were used as natural preparatory factors; the land use/cover and proximity to roads were used as anthropogenic preparatory factors, and the mean annual rainfall was used as a triggering factor.

Defined as the height above a reference point (typically above the mean sea level), altitude is an important conditioning factor due to its gravitational potential energy. In general, the higher the slope angle is, the higher the likelihood of failure. Therefore,

steep slopes are more prone to failures. The slope aspect is defined as the azimuth-based orientation of terrain and is highly related to exposure to sunlight; evapotranspiration; and rainfall's effects on weathering, soil, vegetation cover and root development [44]. Expressed by different types, such as plan and profile, the curvature indicates the runoff and erosion factors of water. The plan curvature is perpendicular to the maximum slope direction, whereas the profile curvature is parallel to the same direction [45]. By retaining more rainfall water and erosion-induced sediment than convex slopes, concave slopes are correlated with higher likelihoods of failure.

Considering its effects on groundwater recharge, stream density constitutes another important factor for landslide activity. This factor determines the ratio of the total length of streams to the extent of the study area. A high stream density is linked to low surface water infiltration and thus mass movements with high velocity [46]. SPI is another hydrological factor that measures the erosive power of the streams. On the other hand, TWI quantifies the moisture content of the surface [32].

Lithology is one of the most crucial factors for LS assessments, since different lithological formations have different slope instability performances in terms of strength and permeability. In a tectonically active country such as Greece, the faults seem to be associated with extensive fractured zones and steep relief anomalies presenting favorable conditions for landslides [35]. Hence, landslides are usually found in proximity to faults. Additionally, different soil types can have different impacts on surface infiltration and groundwater flow, depending on their particular physical and mechanical properties [47].

Changes in land use/cover as a result of human activities such as cultivation, deforestation and forest logging can significantly affect the occurrence of landslides. Proximity to roads can also reflect the human impact on landslides, as road construction at the base of a slope tends to degrade its stability.

Rainfall—causing an increase in the pore water pressure and a reduction in the shear strength of the soil [48]—is a basic triggering mechanism for not only the development of new landslides but also the re-activation of old ones. Particularly in Greece, rainfall-triggered landslides are among the most frequent and devastating disasters [38]. It is worth mentioning that since the majority of earthquakes that occurred in the study area during the last two decades were characterized by relatively low magnitudes (with M_w between 3.0 and 3.5) and great depth (greater than 15 km) [49], seismic factor was not included in the analysis.

As is shown in Table 2, all the above conditioning factors were represented by GIS-supported data formats. Most of them were in raster format (grids), but others were converted from vector (point, line, or polygon features) to a raster format with 25 m spatial resolution.

3.3. Geographical Detector (GeoDetector)

GeoDetector is a spatially-based multivariate statistical model which was developed in 2010 by Wang et al. [50]. It can detect the spatially stratified heterogeneity of a given phenomenon according to the basic principle that if a determinant is associated with the phenomenon, then there may be some similarities between their spatial distributions. Furthermore, it can reveal the driving forces behind the phenomenon by quantifying the impacts of individual determinants and of their pairwise interactions. The phenomenon under investigation as a dependent variable can be represented by either numerical continuous or discrete classified (stratified) data, and the determinants as explanatory variables exclusively by classified data.

Table 2. Summary of the datasets representing the conditioning factors.

Factor	Dataset	Data Source	Spatial/Temporal Scale/Resolution	Primary Format
Altitude	EU-DEM (v1.1)	"Copernicus" Land Monitoring Service	25 m/2011	Raster (grid)
Slope angle	DEM derivative		25 m/2011	Raster (grid)
Slope aspect	DEM derivative		25 m/2011	Raster (grid)
Plan & Profile curvatures	DEM derivative		25 m/2011	Raster (grid)
Stream density	Rivers and streams	General Use Map of Greece (Hellenic Military Geographical Service)	1:50,000/1989	Vector (line)
SPI	DEM-based hydrological analysis		25 m/2011	Raster (grid)
TWI				
Lithology	Lithological formations	Geological Map of Greece (Institute of Geology and Mineral Exploration)	1:50,000/1993	Vector (polygon)
Proximity to faults	Faults			
Soil type	Soil types	Hellenic Ministry of Environment and Energy	1:50,000/1997	Vector (polygon)
		Soil Map of Greece (Aristotle University of Thessaloniki)	1:500,000/2015	
Land use/cover	"CORINE" features	"Copernicus" Land Monitoring Service	1:100,000/2018	Vector (polygon)
Proximity to roads	Main roads	"OpenStreetMap"	–/2020	Vector (line)
Mean annual rainfall	"E-OBS" daily precipitation	"Copernicus" Climate Change Service	0.1 degrees/2000–2015	Raster (grid)

In the case of LS, GeoDetector can detect whether a conditioning factor (explanatory variable) causes the spatial stratified heterogeneity of landslide occurrence (presence or absence of a landslide, dependent variable) or not. In particular, it can quantify the degree of impact of each factor on the landslide occurrence using a q-statistic calculated as follows [51]:

$$q = 1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N\sigma^2} \quad (1)$$

where $h = 1, 2, \ldots, L$ is a given class (stratum) of an explanatory variable; L is the number of classes; N_h and N are the numbers of samples in class h and entire study area, respectively; and σ_h and σ are the variance of dependent variable in class h and entire study area, respectively. Ranging from 0 to 1, the higher the q value is, the more this explanatory variable contributes to the dependent variable. A p-statistic, an indicator of statistical significance for each explanatory variable, is also calculated by a non-central F-distribution:

$$p(q < x) = p\left(F < \frac{N-L}{L-1}\frac{x}{1-x}\right) = 1 - a \quad (2)$$

where a is the probability of q being higher than or equal to x. In a 95% confidence interval, an explanatory variables with a p value greater than 0.05 is considered to have a statistically insignificant relationship with the dependent variable and could be eliminated from the model.

By estimating the value of q-statistic corresponding to the interaction of two explanatory variables, GeoDetector can also quantify the degree of the interactive impact of each pair of conditioning factors on landslide occurrence. As is shown in Table 3, based on the

comparison of this value with the individually estimated values, the type of interaction can be then determined.

Table 3. Types of interaction between two explanatory variables (X1 and X2).

Interaction Type	Description
Nonlinear-weaken	q(X1∩X2) < Min(q(X1), q(X2))
Univariate-weaken	Min(q(X1), q(X2)) < q(X1∩X2) < Max(q(X1), q(X2))
Bivariate-enhanced	q(X1∩X2) > Max(q(X1), q(X2))
Independent	q(X1∩X2) = q(X1) + q(X2)
Nonlinear-enhanced	q(X1∩X2) > q(X1) + q(X2)

3.4. Information Value (IV)

IV is a bivariate statistical model which was initially proposed by Yin and Yan [52] and later modified by van Westen [53]. It includes class-level estimations of weight values based on the spatial associations between the landslide occurrence and each class of each conditioning factor. The IV for a given factor class is derived from a mathematical formula of the ratio of landslide density in this class to the landslide density in entire study area (or factor):

$$IV = ln\left(\frac{Npix(Si)/Npix(Ni)}{\sum Npix(Si)/\sum Npix(Ni)}\right) \quad (3)$$

where *Npix(Si)* is the number of landslide pixels within the factor class *i*, and *Npix(Ni)* is the number of all pixels in the same class. The calculated value can be either positive or negative, and the higher (or lower) it is, the more (or less) significant the contribution of the relevant factor class to landslide occurrence.

4. LS Assessment by Hybrid Modeling

Considering the functionalities and data requirements of the two models composing the GeoDIV hybrid model, two GIS-based data processing procedures initially took place under the general methodological framework (Figure 3). These procedures were the (non)landslide sampling and the factor preparation. For sampling, the landslide inventory dataset was divided into two subsets used as inputs in the model's training (training dataset) and validation (validation dataset), respectively. Among the amount of 60 landslides contained in the inventory, 80% of them (48 in amount) were randomly selected for the training dataset in this study. The remaining 20% (12 in number) constituted the validation dataset. Based on the sizes of mapped landslides and the spatial resolution of obtained factor data, the entire study area was then tiled into grid pixels of 25 × 25 m as the basic analysis unit, resulting in 188 training and 41 validation landslide pixels. The IV model required only a landslide dataset, whereas the GeoDetector model required both landslide and non-landslide datasets. Hence, in order to construct the dependent variable for GeoDetector, an equal number of pixels from the not landslide-affected part of study area were also selected in a random way for the training dataset (totally 376 pixels). The target values of 0 and 1 were assigned to the non-landslide and landslide pixels, respectively, making the dependent variable a binary classified dataset.

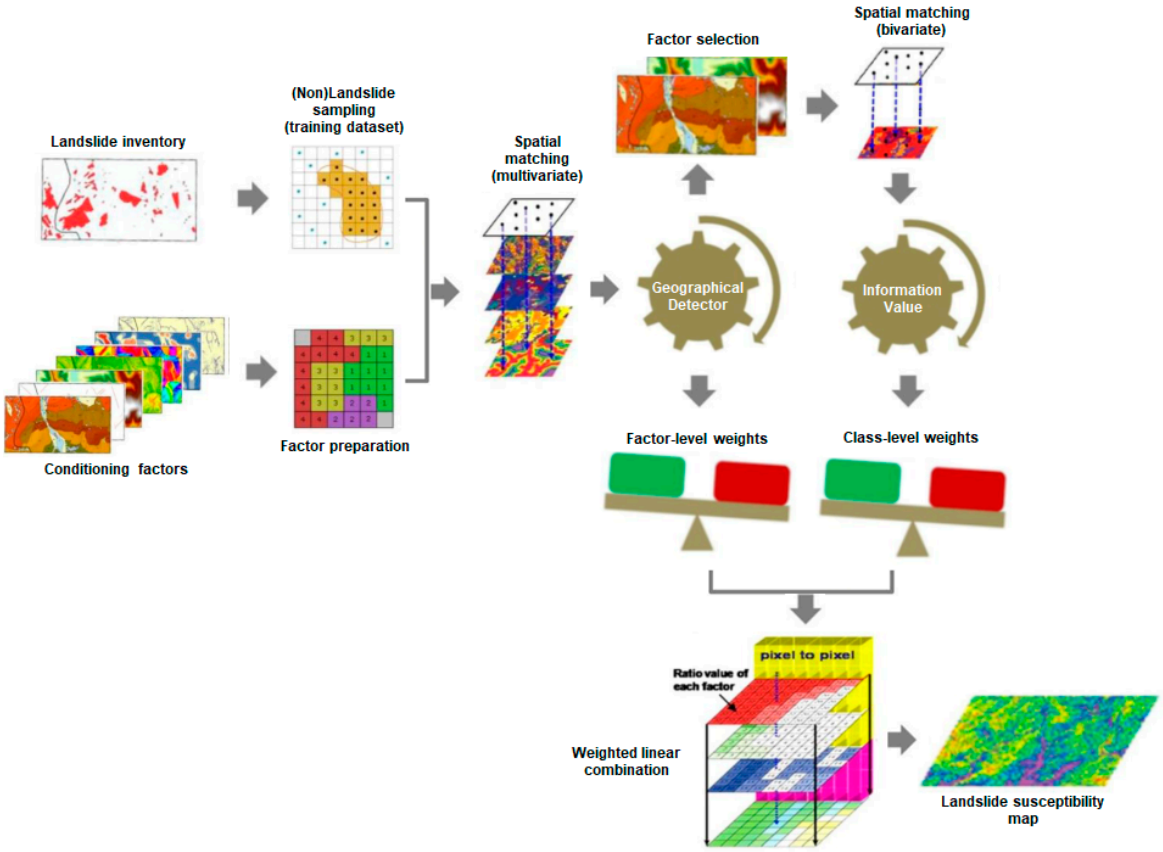


Figure 3. Methodological framework for the development of the hybrid GeoDIV model.

In regard to factor preparation, the raster layers of conditioning factors on a continuous numerical scale (altitude, slope angle, profile curvature, plan curvature, stream density, SPI, TWI, proximity to faults, proximity to roads and mean annual rainfall) were divided into a number of discrete classes (Figure 4). In this study, the number of categories and their relative break values were mainly determined by the "natural breaks (Jenks)" classification method [54]. In this method, class breaks identify the most similar within-group values and maximize the differences between classes according to the deviations about the median [55]. Additionally, the raster layers of factors originally on a discrete classified scale (slope aspect, lithology, soil type, and land use/cover) were prepared by grouping them into more or less common initial classes (Figure 4).

After the data processing procedures, the GeoDIV model was implemented. A database was firstly created as the result of the matching of the sample of 376 training data with each factor layer. Including the fourteen classified factors as independent variables and the landslide presence or absence (binary target value of 0 and 1) as the dependent variable were determined in the GeoDetector software, developed by Xu and Wang [56], to determine the impacts of the factors and their pairwise interactions on the spatial stratified heterogeneity of landslide occurrence represented by the training sample. This determination included the calculation of q values for the factors and their pairwise interactions (Tables 4 and 5). To incorporate in the model only the factors with statistically significant relationships with landslide occurrence, the estimated p values (Table 4) of the factors were also exploited for factor selection. Despite the requirement for p values less than 0.05 in the 95% confidence interval, factors such as altitude, slope angle, plan curvature, stream density, TWI, proximity to faults, proximity to roads, lithology, soil type and land use/cover remained in the model. Conversely, slope aspect, profile curvature, SPI and mean annual rainfall were not qualified to be further analyzed by the model, indicating that there were statistically insignificant relationships (i.e., p values greater than 0.05) between them and landslide occurrence in the same confidence interval.

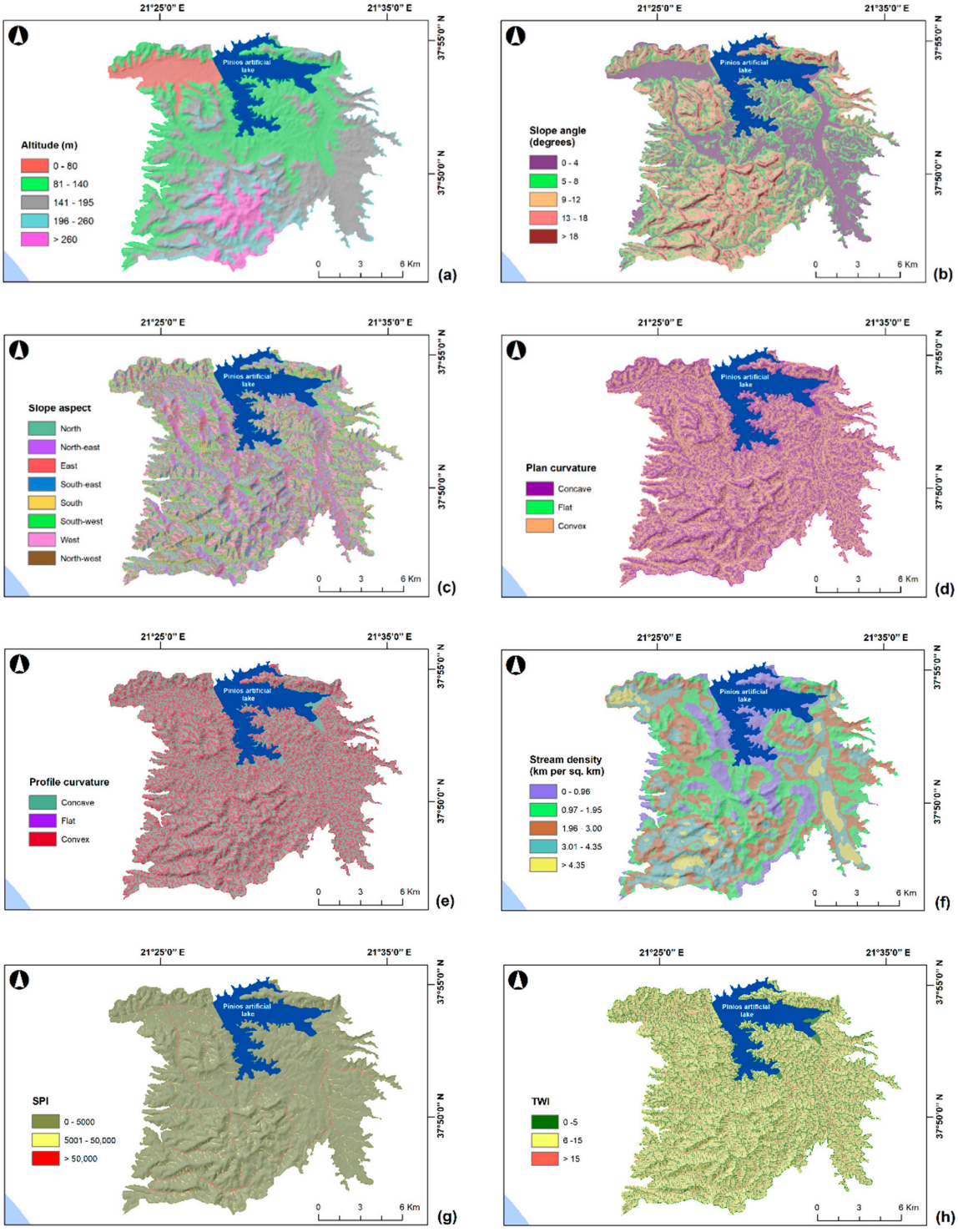


Figure 4. *Cont.*

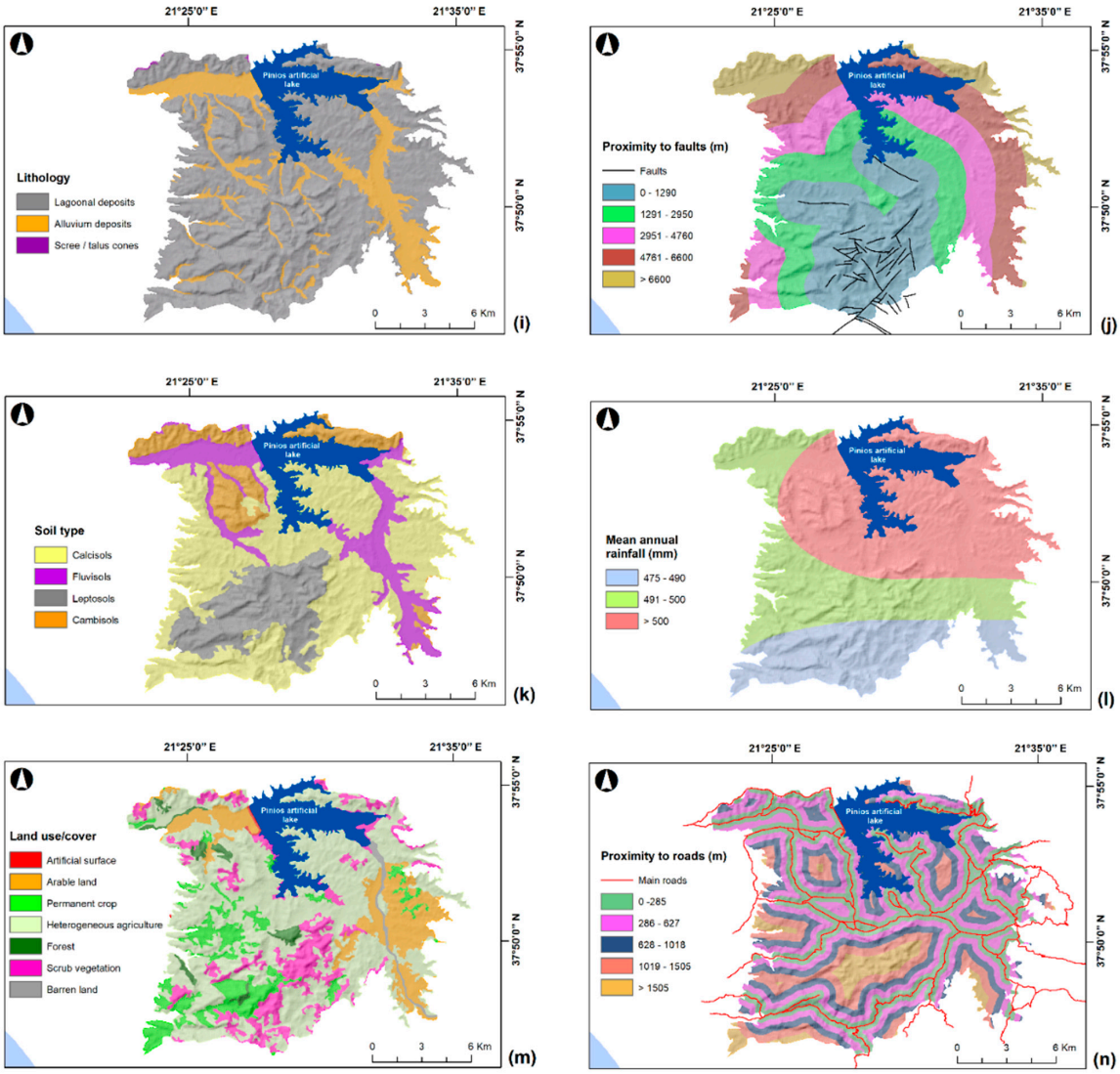


Figure 4. Conditioning factors: (**a**) altitude; (**b**) slope angle; (**c**) slope aspect; (**d**) plan curvature; (**e**) profile curvature; (**f**) stream density; (**g**) SPI; (**h**) TWI; (**i**) lithology; (**j**) proximity to faults; (**k**) soil type; (**l**) mean annual rainfall; (**m**) land use/cover; (**n**) proximity to roads.

Table 4. The *q* and *p*-statistic values for the conditioning factors, calculated using GeoDetector.

Factor	*q* Value	*p* Value
Altitude	0.078	0.00
Slope angle	0.264	0.00
Slope aspect	0.038	0.06 *
Plan curvature	0.016	0.01
Profile curvature	0.021	0.06 *
Stream density	0.065	0.00
SPI	0.003	0.30 *
TWI	0.019	0.04
Lithology	0.053	0.00
Proximity to faults	0.147	0.00
Soil type	0.072	0.00
Land use/cover	0.151	0.00
Proximity to roads	0.174	0.00
Mean annual rainfall	0.001	0.85 *

* indicate the factors eliminated from GeoDetector according to the *p* values.

Table 5. The q-statistic values for the pairwise interactions between the conditioning factors, calculated using GeoDetector.

Factor	Altitude	Slope Angle	Slope Aspect	Plan Curvature	Profile Curvature	SPI	TWI	Proximity to Roads	Proximity to Faults	Stream Density	Mean Annual Rainfall	Lithology	Soil Type	Land Use/Cover
Altitude														
Slope angle	0.370													
Slope aspect	0.268	0.423												
Plan curvature	0.106	0.294	0.098											
Profile curvature	0.101	0.300	0.074	0.034										
SPI	0.093	0.274	0.061	0.022	0.027									
TWI	0.104	0.298	0.110	0.032	0.033	0.021								
Proximity to roads	0.290	0.488	0.347	0.196	0.211	0.190	0.204							
Proximity to faults	0.299	0.398	0.388	0.184	0.188	0.153	0.197	0.319						
Stream density	0.289	0.317	0.237	0.092	0.121	0.083	0.097	0.310	0.377					
Mean annual rainfall	0.155	0.301	0.137	0.030	0.041	0.012	0.027	0.224	0.243	0.227				
Lithology	0.123	0.278	0.133	0.067	0.078	0.058	0.070	0.242	0.224	0.134	0.072			
Soil type	0.195	0.341	0.275	0.090	0.101	0.085	0.099	0.299	0.319	0.274	0.163	0.107		
Land use/cover	0.215	0.350	0.285	0.175	0.177	0.168	0.180	0.326	0.273	0.278	0.170	0.177	0.286	

Subsequently, by matching only the 188 landslide training data with each layer of statistically significant factors, the landslide density for each of their classes was estimated. The IVs were then calculated by Equation (2) to determine the impact of each class on landslide occurrence (Figure 5).

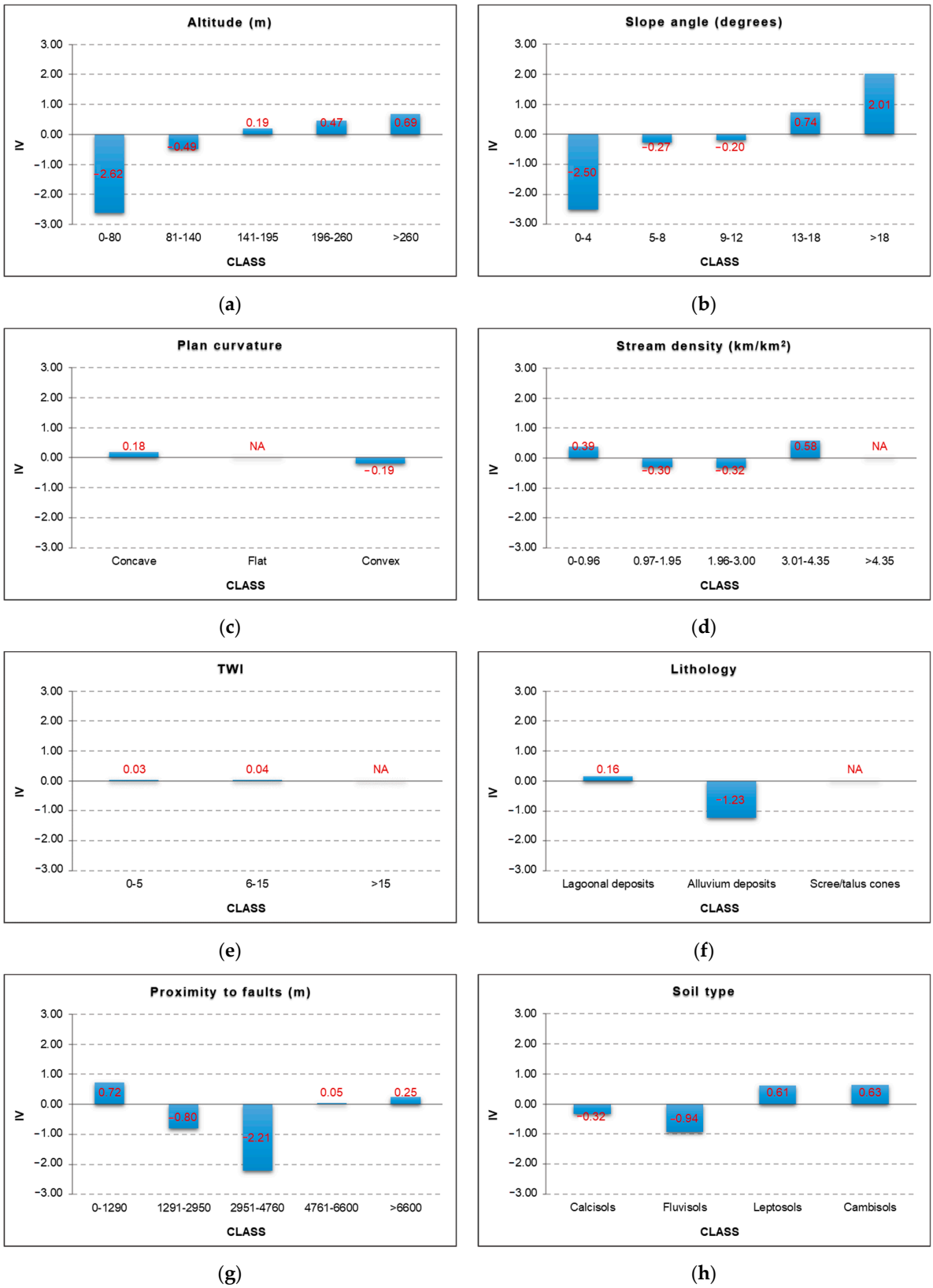
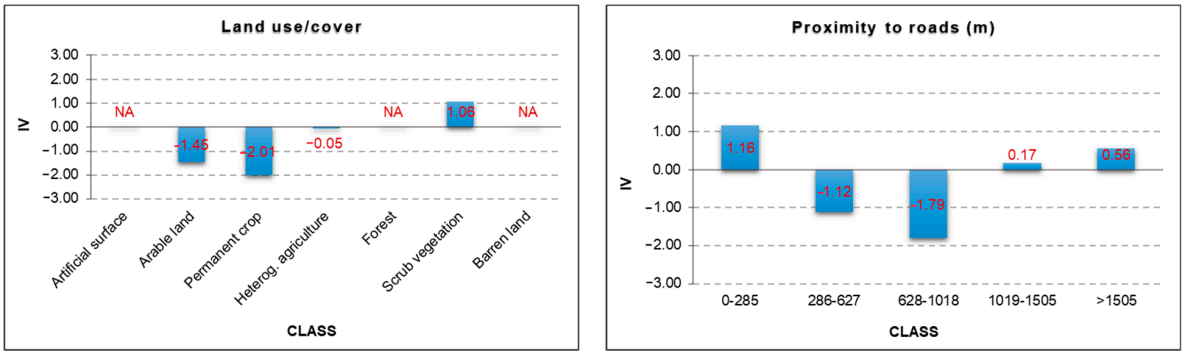


Figure 5. *Cont.*



(**i**) (**j**)

Figure 5. The estimated IVs for the classes of conditioning factors qualified from the factor selection: (**a**) altitude; (**b**) slope angle; (**c**) plan curvature; (**d**) stream density; (**e**) TWI; (**f**) lithology; (**g**) proximity to faults; (**h**) soil type; (**i**) land use/cover; (**j**) proximity to roads. NA values (or no bars) indicate "not applicable" for these classes.

By using the q values from GeoDetector as factor-level weights and IVs as class-level weights, the overall landslide susceptibility (LS) score was estimated through a GIS-based weighted linear combination of statistically significant factors:

$$LS = \sum_{j=1}^{n} W_J \times s_{i,j} \quad (4)$$

where W_j is the weight of a given factor j, $s_{i,j}$ is the weight for a given class i of factor j and n is the number of factors. The spatial distribution of the estimated overall score was visualized by a LS map divided into five classes ("very low", "low", "moderate", "high" and "very high" susceptibility) according to the "natural breaks (Jenks)" method (Figure 6).

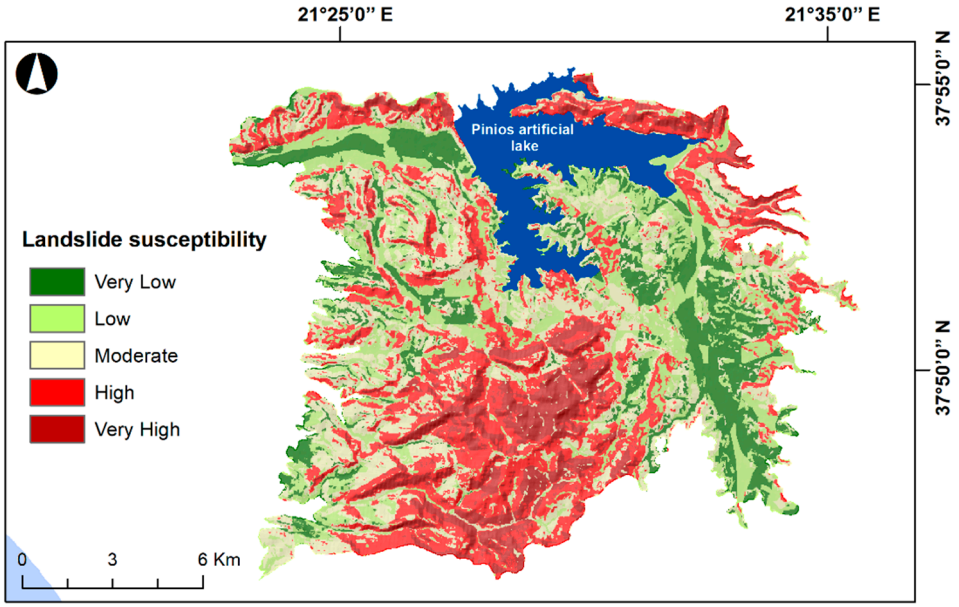


Figure 6. The landslide susceptibility map produced by the hybrid GeoDIV model.

5. Results

The weights from the GeoDIV model are summarized in Table 4 and Figure 5. Among the conditioning factors that eventually remained in the model, the highest factor-level weight was obtained from slope angle (q value of 0.264). It was followed by proximity to roads, land use/cover and proximity to faults (q values of 0.174, 0.151 and 0.147, respectively). For these factors, the classes with the highest class-level weights were the

"greater than 18 degrees" (IV = 2.01) for slope angle, "0 to 285 m" (IV = 1.16) for proximity to roads, "scrub vegetation" (IV = 1.06) for land use/cover and "0–1290 m" (IV = 0.72) for proximity to faults. The rest of conditioning factors were found to have much lower factor-level weights (q values below 0.10). Plan curvature was the factor with the lowest weight (q value of 0.016).

According to the correlations between Tables 3 and 5, the impact degree and types of the different pairwise interactions of factors were determined. The interaction between slope angle and proximity to roads presented the highest weight value (q value of 0.488). This value was greater than the sum of their individual values, indicating that their interaction type was nonlinearly enhanced. Generally, the weights of all the factors (even the lowest of plan curvature) were significantly increased by slope angle, achieving either nonlinear enhancement or bivariate enhancement.

The LS map from GeoDIV model is illustrated in Figure 6. It shows that the "high" and "very high" susceptibility zones are mainly in the southern and northern parts of the vicinity of Pinios artificial lake, with some large pockets of "high" susceptibility in the western part. These two zones cover 25% and 12%of the lake's vicinity, respectively.

Validation and Comparison

In order to evaluate the performance of a model applied for LS assessment and mapping, a validation step is required. Since it can provide information about the accuracy and prediction ability of the model, and thus the reliability of its LS output, this step is crucial for any relevant research effort. A standard validation procedure is one based on success and prediction rates [28,45,48]. This specific procedure depends on the creation of two rate curves explaining the percentages of landslides that fall into defined LS ranks. These curves are graphically presented in cumulative frequency diagrams, with respect to the two different datasets of landslide inventory. For the success rate curve, the landslide training dataset was used to indicate how well the model fits to the training data. On the contrary, for the prediction rate curve, the "independent" landslide validation dataset was used to show how well the model can predict the distribution of future landslides [57].

To obtain the success and prediction rate curves in this study, the overall LS score (Equation (3)) was initially sorted in descending order (from high to low). Then, the ordered LS score was divided into 100 classes with 1% cumulative intervals. The resultant LS ranks (0–100%, where a higher rank means a lower LS score) were plotted on the x-axis, whereas their cumulative percentages of training and validation landslide data are on the y-axis. An area under curve (AUC) value was eventually calculated for each of the two rate curves indicating the accuracy and prediction ability of GeoDIV model, respectively. With a range of 0.5–1.0, this value reflects the model's performance.

Aiming to confirm the potential "superiority" of the hybrid modeling against the individual modeling and explore the impact of GeoDetector-based factor selection on LS assessment, the individual IV model was also applied, and its validation results were compared with those of GeoDIV model. In this context, IVs were additionally calculated for the classes of statistically insignificant factors (not included in GeoDIV model). The overall LS score (presented also by classes, in Figures 7 and 8) was then obtained by the summation of all the fourteen IV-weighted factors as follows:

$$LS = \sum_{j=1}^{n} IV_j \quad (5)$$

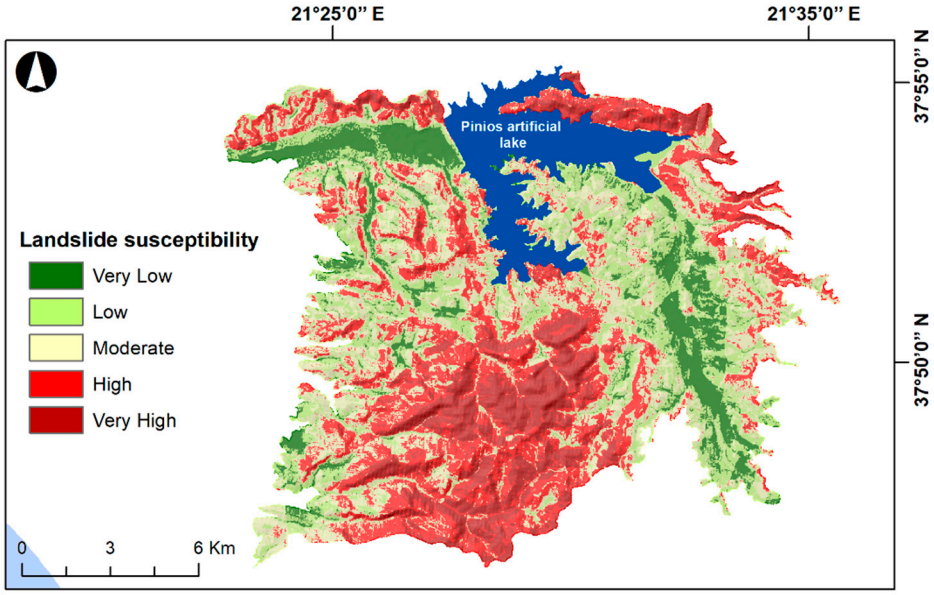
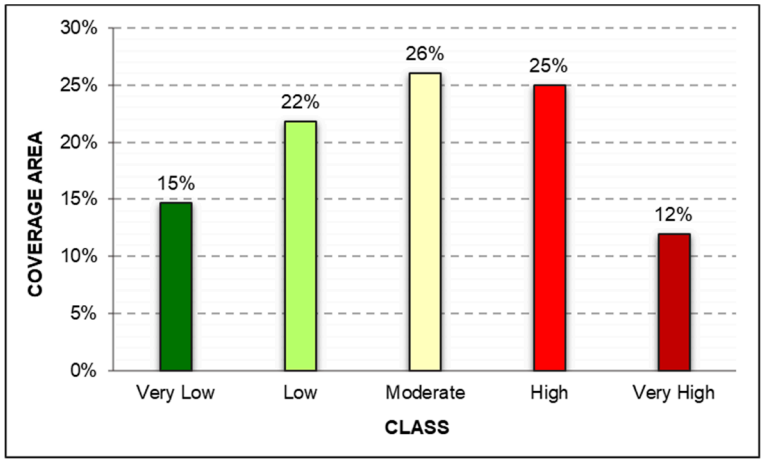
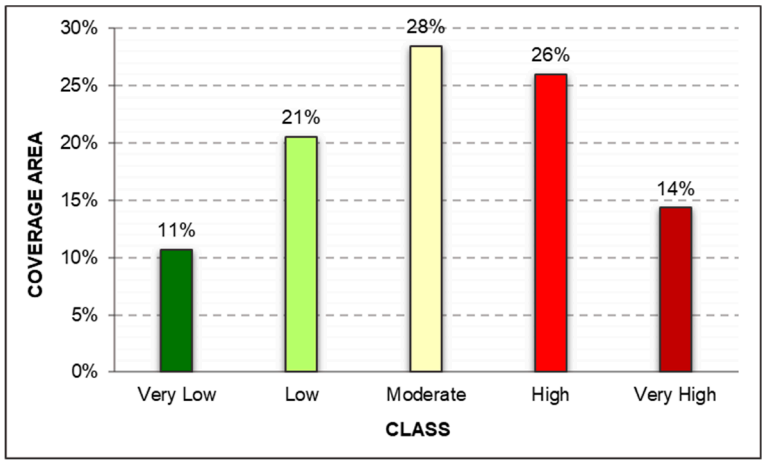


Figure 7. The landslide susceptibility map produced by the individual IV model.



(**a**)

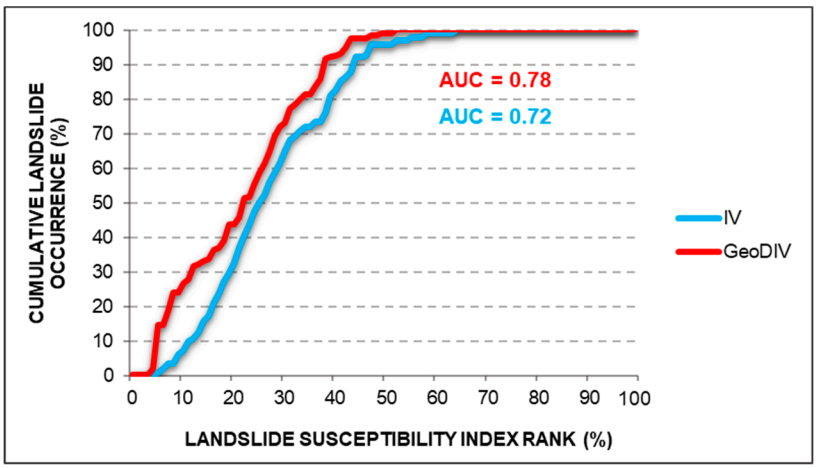


(**b**)

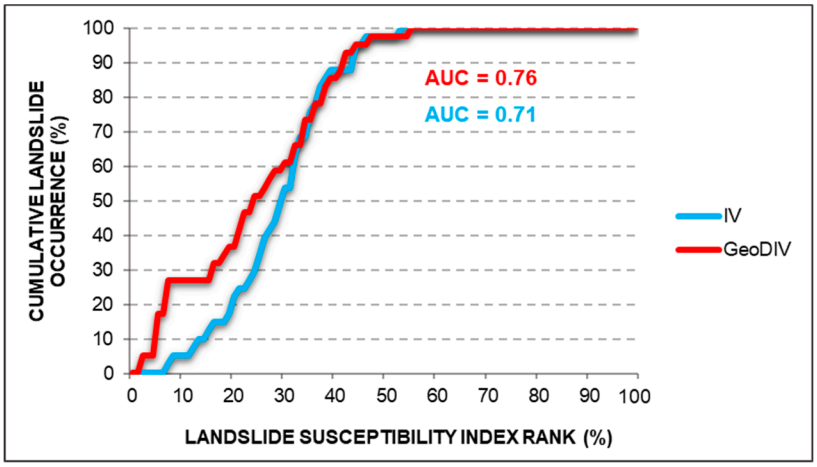
Figure 8. Diagrams with the coverage area percentages of the landslide susceptibility classes for the models: (**a**) hybrid GeoDIV model; (**b**) individual IV model.

Based on the *LS* score rank, the success and prediction rate curves were created, and the relative AUC values were calculated for individual *IV* model as well.

The results of validation procedure for both models are presented in Figure 9. The success and prediction rate curves indicate that the first 30% of the LS ranks derived from GeoDIV model can explain about 70% of landslide training data and 60% of landslide validation data, respectively. Moreover, the relevant AUC values of 0.78 and 0.76 revealed remarkable accuracy (data fitting) and prediction ability of the model. All these results were found to be worse for individual IV model, with explanation percentages of 60% and 50%, respectively, and AUC values of 0.72 and 0.71, respectively.



(**a**)



(**b**)

Figure 9. The results of the validation procedure: (**a**) success rate curves; (**b**) prediction rate curves.

6. Discussion

Due to the observed upward tendency in landslide occurrence, authorities of all the administration levels (national, regional and local) are called on to collaborate with the scientific community to spatially determine potential landslide instances and mitigate, or even prevent, the damage and losses that they may cause. LS assessment and mapping is the first and most basic step for effective risk management and disaster response [58]. Several LS assessment models have been developed and applied, with their own advantages and disadvantages [59]. A current tendency is the integration of these individual models to enhance their benefits and overcome their weaknesses. The consequent hybrid models are expected to reduce the uncertainty and improve the reliability of the output LS maps [60]. In order to address this statement, in the present study, a hybrid model based

on the integration of two different statistical analysis models, multivariate GeoDetector and bivariate IV, was proposed for LS assessment and mapping. In general, GeoDetector, as a new spatial model, has been rarely used in LS studies compared with other models. Hence, its integrated applications are even more limited. To the best of our knowledge and without ignoring the research works of Luo and Liu [7] and Yang et al. [61], the proposed integration had not been tested hitherto for the development of hybrid LS modeling.

A variety of natural and anthropogenic conditioning factors and a landslide inventory for a Greek wetland around the Pinios artificial lake, were analyzed as inputs in the hybrid model named GeoDIV. It can be stated that the advantages (or disadvantages) of GeoDIV model are "inherited" from the two individual models which it was based on. Under strict, prior defined data assumptions, the IV model is capable of evaluating the impact of each class of many conditioning factors due to the occurrence of past landslides; however, the mutual relationship between the factors is mostly neglected [62]. Without any assumptions on the distribution of data, the GeoDetector model is capable of exploring this relationship but not evaluating individually the impact of each factor class.

In addition to the models, factor selection also plays a major role in the LS results [14]. Too many redundant factors may lead to less realistic and reliable results. Therefore, the capability of significance statistics-based factor selection provided by GeoDetector makes it an ideal option for selecting the most proper factors and then assigning objective weights to them with regard to their different contributions to past landslide occurrence. By incorporating this property of GeoDetector in hybrid GeoDIV model, among the fourteen conditioning factors initially collected, four of them (slope aspect, profile curvature, SPI, and mean annual rainfall) were identified as statistically insignificant and were not finally included in LS assessment. Similar factors were also eliminated as redundant in [14,23,24].

Focusing on the factors qualified from factor selection, slope angle was highlighted by the factor-level weights (q values) of the GeoDIV model. In GeoDetector's terminology, slope angle can be characterized as the factor which most explains the spatial stratified heterogeneity of landslide occurrence in the study area. In simple words, its weight was found to be much higher than the rest of factors, revealing that slope angle has the greatest impact on landslide activity. This is in line with findings from other studies in Greece which, on the basis of using either qualitative or quantitative models at different scales (national and regional), also indicated slope angle as one of the most important factors [38,63,64].

When slope angle interacted to some degree with proximity to roads, an even greater impact was detected, according to the interaction weights. Generally, the single impacts of all other factors were shown to be significantly improved from their interactions with the slope angle. Except for the particularly influential role of the specific factor, this finding also confirms the "nature" of landslides as a phenomenon that, to a great extent, constitutes the result of interactions between multiple conditioning factors. From a sub-factor perspective, as it was derived from the class-level weights (IVs), the steep parts of study area being very close to roads and covered by scrub vegetation seem to be more prone to landslides.

The output map of the GeoDIV model illustrated the spatial distribution of the estimated LS. It shows that extensive parts, mainly located in south and north, are most likely to have landslides in the future. In comparison with the relevant map from the individual IV model, it can be mentioned that despite the preservation of the general spatial pattern, there was displacement of the pockets from low susceptibility in GeoDIV's map to higher susceptibility in IV's map. This "overestimation" from the IV model may have been due to the inclusion of the additional four conditioning factors, confirming the above statement about the negative impacts of redundant factors on the reliability of LS results.

Regarding the performance of proposed hybrid model, it has to be firstly noted that GeoDIV provided far more than satisfactory validation results in terms of accuracy (success rate) and prediction ability (prediction rate), considering the scale of analysis. Compared to the IV model, although both models seemed to converge to approximate results, the convergence of GeoDIV was found to be faster. This finding proves the expected "superiority" of hybrid against the individual modeling and is in agreement with previous

studies, concluding that the integration of bivariate with multivariate statistical models improved the performance of former ones [30,47].

Some assumptions and limitations of the present study have to be pointed out. The quality of LS assessment and mapping is highly related to both the landslide inventory and conditioning factors. By using Google Earth satellite imagery, the landslides with identifiable signs in the images were mainly mapped in landslide inventory. Hence, the inventory cannot totally represent the landslide-contributing factors of the study area. Moreover, although the preparation of different susceptibility maps for the various types of landslides can provide more realistic predictions [65], the different types of mapped landslides were not considered in this study. On the other hand, the differentiation between landslide source and deposition zones enabled the models to accurately identify source areas and hence to precisely define the factors that contribute to the initiation of a landslide. The lack of this differentiation resulted to a study's assumption concerning the existence of similar terrain conditions within these zones and thus the representation of each landslide by a single polygon feature. Considerable simplification of these polygons had to be then undertaken by converting them to grid pixels. In this way, an underestimation of landslide data may have taken place in some cases. Additionally, the sampling procedure for the creation of the landslide training and validation datasets can affect the model's efficiency. On the basis of appropriate sizes, a sufficient amount of data should be included in the training dataset, and a remaining "independent" amount of data in the validation dataset. From the perspective of conditioning factors, their spatial resolution and classification can affect the precision of the spatial matching between the landslide and factor data. Therefore, the examination of alternatives for these parameters could lead to different results.

7. Conclusions

A hybrid model named GeoDIV was applied to produce a reliable LS map for the vicinity of Pinios artificial lake (Ilia, Greece). Based on the analysis of landslide and factor conditioning data, the GeoDIV framework exploited the multivariate GeoDetector to eliminate redundant factors and objectively quantify the individual and interactive impacts of the remaining ones (factor-level weights) on landslide occurrence. The bivariate IV was used for objectively quantifying the impacts of their classes (class-level weights). In practice, the integration of these two models increased their efficiency. The findings confirmed that hybrid modeling outperform modeling: the GeoDIV model yielded better results than the individual IV model in terms of accuracy and prediction ability. Thus, GeoDIV can be considered as a promising and robust model which can be beneficial not only to the current study area, but also to other regions with similar or even different conditions and settings.

In general, it was revealed that hybrid LS modeling assisted by multiple geospatial tools (RS and GIS) can contribute well to the production of reliable maps. The LS map produced by the GeoDIV model could be important basis for the regional or local authorities in order to develop both general (long-term) and emergency (short-term) strategies centered on "space design" disaster management. Knowledge about the potential for landslides in a region is valuable for policy makers, as it can allow them to select safe locations while planning land use and approving construction projects. Policy makers could also identify threatened settlements and roads, and in response take drastic disaster management measures (including building engineered structures, planning evacuation routes and issuing early warnings).

Future research work will focus on testing the proposed hybrid modeling for LS assessments of other regions characterized by different environmental and/or human settings, with various landslide densities. Comparisons with other advanced models, such as machine learning models, will be also performed.

Author Contributions: Conceptualization, C.P.; methodology, C.P.; software, C.P.; validation, C.P.; formal analysis, C.P., M.G.G. and A.V.A.; data curation, C.P. and M.G.G.; visualization, C.P. and A.V.A.; writing—original draft preparation, C.P., M.G.G., A.V.A., N.P. and D.D.A.; writing—review

and editing, C.P., M.G.G., A.V.A., N.P. and D.D.A.; supervision, N.P. and D.D.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors extremely appreciate the significant contributions of the journal's editor and reviewers to the handling and revision of this article.

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

References

1. Varnes, D.J. Slope movement types and processes. In *Landslides: Analysis and Control*; Schuster, R.L., Krizek, R.J., Eds.; Transportation Research Board Special Report 176: Washington, DC, USA, 1978; pp. 11–33.
2. Centre for Research on the Epidemiology of Disasters—CRED. *Disaster Year in Review 2020: Global Trends and Perspectives*; Cred Crunch; Université Catholique de Louvain: Brussels, Belgium, 2021; Volume 62, p. 2.
3. Psomiadis, E.; Papazachariou, A.; Soulis, K.X.; Alexiou, D.-S.; Charalampopoulos, I. Landslide mapping and susceptibility assessment using geospatial analysis and earth observation data. *Land* **2020**, *9*, 133. [CrossRef]
4. Centre for Research on the Epidemiology of Disasters—CRED; United Nations International Strategy for Disaster Reduction—UNISDR. *Economic Losses, Poverty & Disasters (1998–2017)*; Université Catholique de Louvain: Brussels, Belgium, 2018; p. 33.
5. Brabb, E.E. Innovative approaches to landslide hazard mapping. In *Proceedings of the 4th International Symposium on Landslides*; Canadian Geotechnical Society: Tororno, ON, Canada, 1984; pp. 307–324.
6. Varnes, D.J. *Landslide Hazard Zonation: A Review of Principles and Practice*; UNESCO: Paris, France, 1984; Volume 3.
7. Luo, W.; Liu, C.-C. Innovative landslide susceptibility mapping supported by geomorphon and geographical detector methods. *Landslides* **2018**, *15*, 465–474. [CrossRef]
8. Arabameri, A.; Pradhan, B.; Rezaei, K.; Lee, C.-W. Assessment of landslide susceptibility using statistical- and artificial intelligence-based FR–RF integrated model and multiresolution DEMs. *Remote Sens.* **2019**, *11*, 999. [CrossRef]
9. Ciurleo, M.; Mandaglio, M.C.; Moraci, N. A quantitative approach for debris flow inception and propagation analysis in the lead up to risk management. *Landslides* **2021**, *18*, 2073–2093. [CrossRef]
10. Vieira, B.C.; Fernandes, N.F.; Filho, O.A.; Martins, T.D.; Montgomery, D.R. Assessing shallow landslide hazards using the TRIGRS and SHALSTAB models, Serra do Mar, Brazil. *Environ. Earth Sci.* **2018**, *77*, 260. [CrossRef]
11. Su, C.; Wang, L.; Wang, X.; Huang, Z.; Zhang, X. Mapping of rainfall-induced landslide susceptibility in Wencheng, China, using support vector machine. *Nat. Hazards* **2015**, *76*, 1759–1779. [CrossRef]
12. Taaleb, K.; Cheng, T.; Zhang, Y. Mapping landslide susceptibility and types using Random Forest. *Big Earth Data* **2018**, *2*, 159–178. [CrossRef]
13. Gorsevski, P.V.; Brown, M.K.; Panter, K.; Onasch, C.M.; Simic, A.; Snyder, J. Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: A case study in the Cuyahoga Valley National Park, Ohio. *Landslides* **2016**, *13*, 467–484. [CrossRef]
14. Xie, W.; Li, X.; Jian, W.; Yang, Y.; Liu, H.; Robledo, L.F.; Nie, W. A novel hybrid method for landslide susceptibility mapping-based geodetector and machine learning cluster: A case of Xiaojin county, China. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 93. [CrossRef]
15. Kayastha, P. Landslide susceptibility mapping and factor effect analysis using frequency ratio in a catchment scale: A case study from Garuwa sub-basin, East Nepal. *Arab. J. Geosci.* **2015**, *8*, 8601–8613. [CrossRef]
16. Borrelli, L.; Ciurleo, M.; Gullà, G. Shallow landslide susceptibility assessment in granitic rocks using GIS-based statistical methods: The contribution of the weathering grade map. *Landslides* **2018**, *15*, 1127–1142. [CrossRef]
17. Baeza, C.; Corominas, J. Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf. Process. Landf.* **2001**, *26*, 1251–1263. [CrossRef]
18. Dagdelenler, G.; Nefeslioglu, H.A.; Gokceoglu, C. Modification of seed cell sampling strategy for landslide susceptibility mapping: An application from the Eastern part of the Gallipoli Peninsula (Canakkale, Turkey). *Bull. Eng. Geol. Environ.* **2016**, *75*, 575–590. [CrossRef]
19. Mondal, S.; Mandal, S. RS & GIS-based landslide susceptibility mapping of the Balason River basin, Darjeeling Himalaya, using logistic regression (LR) model. *Georisk* **2018**, *12*, 29–44.
20. Polykretis, C.; Alexakis, D.D. Spatial stratified heterogeneity of fertility and its association with socio-economic determinants using Geographical Detector: The case study of Crete Island, Greece. *Appl. Geogr.* **2021**, *127*, 102384. [CrossRef]
21. Xiong, J.; Pang, Q.; Fan, C.; Cheng, W.; Ye, C.; Zhao, Y.; He, Y.; Cao, Y. Spatiotemporal characteristics and driving force analysis of flash floods in Fujian province. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 133. [CrossRef]
22. Zhou, C.; Zhu, N.; Xu, J.; Yang, D. The contribution rate of driving factors and their interactions to temperature in the yangtze river delta region. *Atmosphere* **2020**, *11*, 32. [CrossRef]
23. Yang, J.; Song, C.; Yang, Y.; Xu, C.; Guo, F.; Xie, L. New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector: A case study of Duwen Highway Basin, Sichuan Province, China. *Geomorphology* **2019**, *324*, 62–71. [CrossRef]

24. Rong, G.; Li, K.; Han, L.; Alu, S.; Zhang, J.; Zhang, Y. Hazard mapping of the rainfall–landslides disaster Chain based on GeoDetector and Bayesian Network Models in Shuicheng County, China. *Water* **2020**, *12*, 2572. [CrossRef]
25. Pourghasemi, H.R.; Teimoori Yansari, Z.; Panagos, P.; Pradhan, B. Analysis and evaluation of landslide susceptibility: A review on articles published during 2005–2016 (periods of 2005–2012 and 2013–2016). *Arab. J. Geosci.* **2018**, *11*, 193. [CrossRef]
26. Ciurleo, M.; Cascini, L.; Calvello, M. A comparison of statistical and deterministic methods for shallow landslide susceptibility zoning in clayey soils. *Eng. Geol.* **2017**, *223*, 71–81. [CrossRef]
27. Aditian, A.; Kubota, T.; Shinohara, Y. Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology* **2018**, *318*, 101–111. [CrossRef]
28. Barella, C.F.; Sobreira, F.G.; Zêzere, J.L. A comparative analysis of statistical landslide susceptibility mapping in the southeast region of Minas Gerais state, Brazil. *Bull. Eng. Geol. Environ.* **2019**, *78*, 3205–3221. [CrossRef]
29. Saha, S.; Arabameri, A.; Saha, A.; Blaschke, T.; Ngo, P.T.T.; Nhu, V.H.; Band, S.S. Prediction of landslide susceptibility in Rudraprayag, India using novel ensemble of conditional probability and boosted regression tree-based on cross-validation method. *Sci. Total Environ.* **2021**, *764*, 142928. [CrossRef]
30. Chen, W.; Sun, Z.; Han, J. Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models. *Appl. Sci.* **2019**, *9*, 171. [CrossRef]
31. Roy, J.; Saha, S.; Arabameri, A.; Blaschke, T.; Bui, D.T. A novel ensemble approach for landslide susceptibility mapping (LSM) in Darjeeling and Kalimpong districts, West Bengal, India. *Remote Sens.* **2019**, *11*, 2866. [CrossRef]
32. Chowdhuri, I.; Pal, S.C.; Arabameri, A.; Ngo, P.T.T.; Chakrabortty, R.; Malik, S.; Das, B.; Roy, P. Ensemble approach to develop landslide susceptibility map in landslide dominated Sikkim Himalayan region, India. *Environ. Earth Sci.* **2020**, *79*, 476. [CrossRef]
33. Paryani, S.; Neshat, A.; Javadi, S.; Pradhan, B. Comparative performance of new hybrid ANFIS models in landslide susceptibility mapping. *Nat. Hazards* **2020**, *103*, 1961–1988. [CrossRef]
34. Balogun, A.-L.; Rezaie, F.; Pham, Q.B.; Gigović, L.; Drobnjak, S.; Aina, Y.A.; Panahi, M.; Yekeen, S.T.; Lee, S. Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms. *Geosci. Front.* **2021**, *12*, 101104. [CrossRef]
35. Sabatakakis, N.; Koukis, G.; Vassiliades, E.; Lainas, S. Landslide susceptibility zonation in Greece. *Nat. Hazards* **2013**, *65*, 523–543. [CrossRef]
36. Tsangaratos, P.; Ilia, I. Landslide susceptibility mapping using a modified decision tree classifier in the Xanthi Perfection, Greece. *Landslides* **2016**, *13*, 305–320. [CrossRef]
37. Kavoura, K.; Sabatakakis, N. Investigating landslide susceptibility procedures in Greece. *Landslides* **2020**, *17*, 127–145. [CrossRef]
38. Chalkias, C.; Polykretis, C.; Ferentinou, M.; Karymbalis, E. Integrating expert knowledge with statistical analysis for landslide susceptibility assessment at regional scale. *Geosciences* **2016**, *6*, 14. [CrossRef]
39. Hellenic Statistical Authority (ELSTAT). Population and Housing Census: Resident Population. Available online: https://www.statistics.gr/el/statistics/pop (accessed on 1 June 2021).
40. Lainas, S.; Sabatakakis, N.; Koukis, G. Rainfall thresholds for possible landslide initiation in wildfire-affected areas of western Greece. *Bull. Eng. Geol. Environ.* **2016**, *75*, 883–896. [CrossRef]
41. Dimitriadou, S.; Katsanou, K.; Charalabopoulos, S.; Lambrakis, N. Interpretation of the factors defining groundwater quality of the site subjected to the wildfire of 2007 in Ilia prefecture, South-Western Greece. *Geosciences* **2018**, *8*, 108. [CrossRef]
42. Laboratory of Engineering Geology, Department of Geology, University of Patras. Landslide Management System of Western Greece. Available online: http://landslide.engeolab.gr/ (accessed on 9 March 2021).
43. Popescu, M.E. Landslide causal factors and landslide remediatial options. In Proceedings of the 3rd International Conference on Landslides, Slope Stability and Safety of Infrastructures, Singapore, 11–12 July 2002; pp. 61–81.
44. Shirzadi, A.; Bui, D.T.; Pham, B.T.; Solaimani, K.; Chapi, K.; Kavian, A.; Shahabi, H.; Revhaug, I. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. *Environ. Earth Sci.* **2017**, *76*, 60. [CrossRef]
45. Rabby, Y.W.; Li, Y. Landslide susceptibility mapping using integrated methods: A case study in the Chittagong hilly areas, Bangladesh. *Geosciences* **2020**, *10*, 483. [CrossRef]
46. Pradhan, A.M.S.; Kim, Y.T. Relative effect method of landslide susceptibility zonation in weathered granite soil: A case study in Deokjeok-ri Creek, South Korea. *Nat. Hazards* **2014**, *72*, 1189–1217. [CrossRef]
47. Li, R.; Wang, N. Landslide susceptibility mapping for the Muchuan county (China): A comparison between bivariate statistical models (WoE, EBF, and IoE) and their ensembles with logistic regression. *Symmetry* **2019**, *11*, 762. [CrossRef]
48. Yan, F.; Zhang, Q.; Ye, S.; Ren, B. A novel hybrid approach for landslide susceptibility mapping integrating analytical hierarchy process and normalized frequency ratio methods with the cloud model. *Geomorphology* **2019**, *327*, 170–187. [CrossRef]
49. Institute of Geodynamics—National Observatory of Athens (IG-NOA). Earthquake Inventories and Maps. Available online: http://www.gein.noa.gr/el/seismikotita/ (accessed on 16 April 2021).
50. Wang, J.-F.; Li, X.-H.; Christakos, G.; Liao, Y.-L.; Zhang, T.; Gu, X.; Zheng, X.-Y. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 107–127. [CrossRef]
51. Wang, J.-F.; Zhang, T.-L.; Fu, B.-J. A measure of spatial stratified heterogeneity. *Ecol. Indic.* **2016**, *67*, 250–256. [CrossRef]
52. Yin, K.J.; Yan, T.Z. Statistical prediction model for slope instability of metamorphosed rocks. In Proceedings of the 5th international Symposium on Landslides, Lausanne, Switzerland, 13 September 1988; Volume 2, pp. 1269–1272.

53. Van Westen, C.J. *Application of Geographical Information System to Landslide Hazard Zonation*; ITC-Publication No. 15; ITC-Publication; International Institute for Geo-Information Science and Earth Observation: Enschede, The Netherlands, 1993; p. 245.
54. Calvello, M.; Ciurleo, M. Optimal use of thematic maps for landslide susceptibility assessment by means of statistical analyses: Case study of shallow landslides in fine grained soils. In *Landslides and Engineered Slopes. Experience, Theory and Practice, Proceedings of the 12th International Symposium on Landslides, Napoli, Italy, 12–19 June 2016*; Aversa, S., Cascini, L., Picarelli, L., Scavia, C., Eds.; CRC Press: London, UK, 2016; Volume 2, pp. 537–544.
55. Jenks, G.F. *Optimal Data Classification for Choropleth Maps*; University of Kansas: Lawrence, KS, USA, 1977.
56. Xu, C.-D.; Wang, J.-F. Geodetector: Software for Measure and Attribution of Stratified Heterogeneity (SH). Available online: http://www.geodetector.cn/ (accessed on 19 March 2020).
57. Chung, C.F.; Fabbri, A.G. Probabilistic prediction models for landslide hazard mapping. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 1389–1399.
58. Xing, Y.; Yue, J.; Guo, Z.; Chen, Y.; Hu, J.; Travé, A. Large-scale landslide susceptibility mapping using an integrated machine learning model: A case study in the Lvliang Mountains of China. *Front. Earth Sci.* **2021**, *9*, 722491. [CrossRef]
59. Du, G.-L.; Zhang, Y.-S.; Iqbal, J.; Yang, Z.-H.; Yao, X. Landslide susceptibility mapping using an integrated model of information value method and logistic regression in the Bailongjiang watershed, Gansu Province, China. *J. Mt. Sci.* **2017**, *14*, 249–268. [CrossRef]
60. Luo, X.; Lin, F.; Chen, Y.; Zhu, S.; Xu, Z.; Huo, Z.; Yu, M.; Peng, J. Coupling logistic model tree and random subspace to predict the landslide susceptibility areas with considering the uncertainty of environmental features. *Sci. Rep.* **2019**, *9*, 15369. [CrossRef]
61. Yang, Y.; Yang, J.; Xu, C.; Xu, C.; Song, C. Local-scale landslide susceptibility mapping using the B-GeoSVC model. *Landslides* **2019**, *16*, 1301–1312. [CrossRef]
62. Zhang, G.; Cai, Y.; Zheng, Z.; Zhen, J.; Liu, Y.; Huang, K. Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in Huizhou, China. *Catena* **2016**, *142*, 233–244. [CrossRef]
63. Sakkas, G.; Misailidis, I.; Sakellariou, N.; Kouskouna, V.; Kaviri, G. Modeling landslide susceptibility in Greece: A weighted linear combination approach using analytic hierarchical process, validated with spatial and statistical analysis. *Nat. Hazards* **2016**, *84*, 1873–1904. [CrossRef]
64. Polykretis, C.; Chalkias, C. Comparison and evaluation of landslide susceptibility maps obtained from weight of evidence, logistic regression, and artificial neural network models. *Nat. Hazards* **2018**, *93*, 249–274. [CrossRef]
65. Regmi, A.D.; Devkota, K.C.; Yoshida, K.; Pradhan, B.; Pourghasemi, H.R.; Kumamoto, T.; Akgun, A. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab. J. Geosci.* **2014**, *7*, 725–742. [CrossRef]