

Article

Landslide Geo-Hazard Risk Mapping Using Logistic Regression Modeling in Guixi, Jiangxi, China

Wenchao Huangfu¹, Weicheng Wu^{1,*} , Xiaoting Zhou¹, Ziyu Lin¹, Guiliang Zhang², Renxiang Chen², Yong Song², Tao Lang², Yaozu Qin¹, Penghui Ou¹, Yang Zhang¹ , Lifeng Xie¹, Xiaolan Huang¹ , Xiao Fu¹, Jie Li¹, Jingheng Jiang¹, Ming Zhang¹, Yixuan Liu¹, Shanling Peng¹, Chongjian Shao¹, Yonghui Bai¹, Xiaofeng Zhang³, Xiangtong Liu⁴ and Wenheng Liu¹

- ¹ Key Laboratory of Digital Lands and Resources and Faculty of Earth Sciences, East China University of Technology, Nanchang 330013, China; 201810818004@ecut.edu.cn (W.H.); 201900818004@ecut.edu.cn (X.Z.); zylin@ecut.edu.cn (Z.L.); qyz60010@ecut.edu.cn (Y.Q.); 201810818013@ecut.edu.cn (P.O.); 201810818002@ecut.edu.cn (Y.Z.); 201810705007@ecut.edu.cn (L.X.); 201810705009@ecut.edu.cn (X.H.); 201910705015@ecut.edu.cn (X.F.); 201910705016@ecut.edu.cn (J.L.); 201910818023@ecut.edu.cn (J.J.); 201910818003@ecut.edu.cn (M.Z.); 201910853001@ecut.edu.cn (Y.L.); pshanling@ecut.edu.cn (S.P.); scj350936@ecut.edu.cn (C.S.); 201960070@ecut.edu.cn (Y.B.); 201460013@ecut.edu.com (W.L.)
- ² 264 Geological Team of Jiangxi Nuclear Industry, Ganzhou 341000, China; zgl-63@163.com (G.Z.); crxkcy@163.com (R.C.); songy_6611@163.com (Y.S.); ecitlangtao@163.com (T.L.)
- ³ School of Geophysics and Measurement-Control Technology, East China University of Technology, Nanchang 330013, China; 201160023@ecut.edu.cn
- ⁴ Faculty of Geomatics, East China University of Technology, Nanchang 330013, China; 200760007@ecut.edu.cn
- * Correspondence: Wuwc030903@sina.com or wuwch@ecut.edu.cn; Tel.: +86-18970847278



Citation: Huangfu, W.; Wu, W.; Zhou, X.; Lin, Z.; Zhang, G.; Chen, R.; Song, Y.; Lang, T.; Qin, Y.; Ou, P.; et al. Landslide Geo-Hazard Risk Mapping Using Logistic Regression Modeling in Guixi, Jiangxi, China. *Sustainability* **2021**, *13*, 4830. <https://doi.org/10.3390/su13094830>

Academic Editor: Stefano Morelli

Received: 19 March 2021

Accepted: 15 April 2021

Published: 25 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Reliable prediction of landslide occurrence is important for hazard risk reduction and prevention. Taking Guixi in northeast Jiangxi as an example, this research aimed to conduct such a landslide risk assessment using a multiple logistic regression (MLR) algorithm. Field-investigated landslides and non-landslide sites were converted into polygons. We randomly generated 50,000 sampling points to intersect these polygons and the intersected points were divided into two parts, a training set (TS) and a validation set (VT) in a ratio of 7 to 3. Thirteen geo-environmental factors, including elevation, slope, and distance from roads were employed as hazard-causative factors, which were intersected by the TS to create the random point (RP)-based dataset. The next step was to compute the certainty factor (CF) of each factor to constitute a CF-based dataset. MLR was applied to the two datasets for landslide risk modeling. The probability of landslides was then calculated in each pixel, and risk maps were produced. The overall accuracy of these two models versus VS was 91.5% and 90.4% with a Kappa coefficient of 0.814 and 0.782, respectively. The RP-based MLR modeling achieved more reliable predictions and its risk map seems more plausible for providing technical support for implementing disaster prevention measures in Guixi.

Keywords: multiple logistic regression (MLR); certainty factor (CF); landslide hazard; risk prediction and mapping

1. Introduction

Landslide hazard is one of the frequent geo-hazards, especially in south China, and it has brought severe damages and losses to human society, as well as to the environment [1–3]. According to incomplete statistics, from the beginning of the 21st century to present time, landslide hazards in China have caused as many as 1200 casualties and the estimated economic loss is as high as 5–10 billion yuan, which further restricts the socioeconomic development in the hazard-affected regions [2]. It is therefore essential to conduct an accurate and reliable risk prediction for landslide prevention and early warning.

For this purpose, it is necessary to understand the physical process, or rather, the mechanism of landslides. In general, studied landslides are presumably triggered by

precipitation, which may cause landslides when the rainfall is intense, soil resistance is relatively small, and the slope inclination is relatively high. The landslides cover a larger area if the slip length is longer [3–7]. However, the works of Zhang et al. [8], Ou et al. [9], and Zhou et al. [10] also revealed the important role of road construction and housing development in landslide events. Hence, it is critical to consider geological conditions, environmental conditions, and human activities in landslide risk assessment.

With the advent of the big data era, traditional geological methods have gradually become unable to meet current needs to deal with the more and more abundant big geodata. Artificial intelligence (AI) has long been used in scientific research but it was further promoted by machine learning (ML) and especially by deep learning (DL) techniques because of the capacity for processing hyper-dimensional data and for accurate prediction [11]. AI has not yet been widely applied in geoscience; it may therefore be of critical importance to develop and apply relevant algorithms or adapt the available algorithms to tackle big geodata for ore mineralization prediction or geo-hazard risk assessment to serve our society [12,13]. This will clearly be a meaningful exploratory experiment [13]. Landslide risk prediction is one of the important tasks for ML application in environmental geoscience.

As early as the 1970s, developed countries started to study the risk of geo-hazards, mainly landslides, in the world. In 1999, the United Nations General Assembly proposed an initiative on “the international strategy for disaster reduction” to reduce social vulnerability and the risk of natural disasters. In response to this initiative and based on previous research, countries in North America and Europe carried out studies on geo-hazard risk and its relationship with land use, making great progress in the domain of hazard mitigation and risk reduction [14]. For example, Lumb [15] analyzed the relationship between rock-soil types of slope mass and the occurrence of landslide. Collison et al. [16] suggested that the incline stability is associated with the vegetation coverage in tropical regions. Evans [17] revealed the fact that rainfall distribution in mountainous areas was affected to a certain extent by the elevation of the incline. Mark et al. [18] used 1500 landslide data to study the relationship between the shallow landslide frequency and terrain, and concluded that the steeper the incline, the greater the probability of a landslide.

Studies have also been undertaken to determine the concrete landslide causative factors. For example, Jiang and Eastman [19] used genetic algorithms to optimize the hazard-causative factors in landslide susceptibility mapping; and in terms of information entropy, Wan et al. [20] harnessed a data mining technique to determine the key controlling factors and their thresholds to provoke landslide hazards. At present, there are in fact a number of methods for landslide risk zoning, e.g., analytic hierarchy process (AHP), Fuzzy comprehensive evaluation, information quantity, certainty factor (CF) approach, and multiple logistic regression (MLR) analysis [21–24]. However, there are problems with the single evaluation model, e.g., its limitation in objectively determining the weight of the causative factors, subjective disturbances to the processing of variables in the modeling process, and negligence of the auto-correlation of the causative factors. It is thus a challenge to objectively, quantitatively, and accurately conduct hazard risk assessment at a regional scale [25,26]. Machine learning algorithms have also been utilized for landslide research [8,10], and a series of models are available, such as support vector machines (SVMs) [27], artificial neural networks (ANNs) [28], random forests (RFs) [29], boosted regression trees (BRTs) [30,31], and limit learning machine (LLM) [32]. These algorithms have been reported to have high accuracy for landslide risk assessment, though they have clear disadvantages such as a complex modeling process, unstable model performance, and weak explanatory capability [33,34]. For example, it is difficult to select the penalty coefficient c and kernel function g for SVM models, and to determine the number of neuron nodes in the hidden layer of backpropagation (BP) neural networks, etc. [35]. At present, the selection of these parameters lacks a systematic method and largely relies on the experience of researchers, which brings great uncertainty and difficulty for machine learning modeling. The combination of the MLR method based on GIS and some quantitative evaluation methods may have great potential in geo-hazard risk zoning and assessment. For example,

Xu et al. [36] evaluated the landslide susceptibility provoked by an earthquake based on GIS and certainty coefficient analysis, and Miao et al. [37] conducted the susceptibility analysis of Lanzhou landslide hazard based on GIS and a MLR model. Qian et al. [38] and Ou et al. [9] used a combination of binary MLR and an information value (IV) approach to study the geo-hazard risk of Renhe watershed and Jiangxi, respectively. It is seen that MLR has a high potential for this kind of risk assessment. For this reason, this paper intends to utilize MLR to establish the functional relation between the landslide hazard (probability) and its geo-environmental factors, and to produce a risk map to provide technical support for disaster reduction and prevention action taking Guixi as an example.

2. Methodology

2.1. Study Area

Situated in the northeast of Jiangxi, Guixi is a county covering a surface area of 2292 km², extending geographically between 116°57'43" E and 117°28'06" E and between 27° 50'53" N and 28°37'33" N in space (Figure 1). Dominated by the humid subtropical monsoon climate, Guixi is characterized by a hot summer and mild winter, plentiful precipitation, and four distinct seasons. The mean annual temperature is 18.2 °C and the average annual precipitation is 1790 mm with a maximum of 2761.2 mm and a minimum of 1056.8 mm. The rainfall period is mainly concentrated in March to July, with an accumulated average of 1229 mm, accounting for 68.7% of the annual precipitation. However, it is less humid in the period from September to January. The terrain is complex in the study area, with an elevation ranging from 20 to 1504 m and a slope from 0 to 73°. Guixi is generally high in altitude in the south as a part of the Wuyishan Mountains (Mts), which is composed of granitic massif, and low in the north as a plain, with Red Basin and hills in the center. Xinjiang is the main river in the central part of Guixi flowing from east to west, 60 km in length of which passes through the study area. The basin encompasses more than 50 km² in surface and includes 11 tributaries, e.g., Baita and Luotang Rivers from the south, and Yingshi, Taqiao, and Sili Rivers from the north.

Landslides are the most widespread and harmful geo-hazard in Guixi. According to field investigation, this hazard occurs frequently in spring and summer, especially in June and July; they are mostly small in scale and distributed on both sides of the roads. Most of the landslides contain a clear sliding surface, i.e., soil-rock interface, and are shallow slides (<10 m, mostly <3 m in thickness).

2.2. Data

The selection of landslide causative factors based on field observation is the first key step in risk zoning and assessment. In combination with the results of other authors [8–10,39–44], the following geo-environmental factors were obtained and prepared in order to achieve our objectives: elevation, slope, aspect, proximity to roads, proximity to rivers, proximity to faults, lithology of strata, land cover, average annual precipitation, accumulated average precipitations of March–July, May–July, and March–June, and the normalized difference vegetation index (NDVI) standing for the condition of vegetation development. The sources of these data are listed in Table 1 and details of their quantification are presented in the following sections.

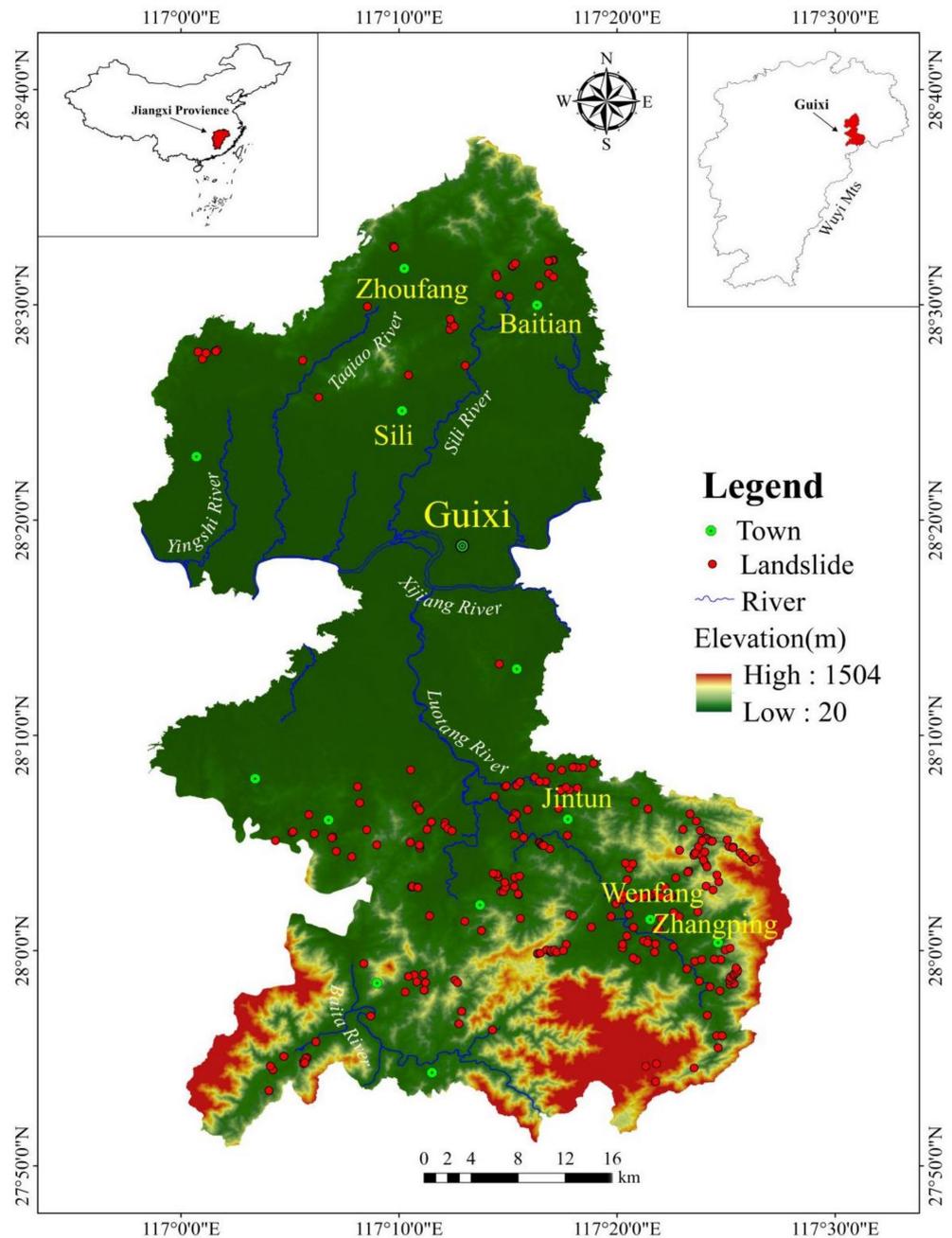


Figure 1. Location map of the study area, Guixi, Jiangxi, China.

Table 1. Data types and sources.

Data Type	Source	Type
Geological map	1/50,000 scale geological map	shapefile
Rivers and roads	Google Earth	shapefile
Meteorological data	Ground stations	shapefile and grid
Topographical data	ASTGTMV003 30m from NASA (www.earthdata.nasa.gov , accessed on 11 August 2020)	grid (30 m)
Land cover types and NDVI	Landsat data from USGS data server (https://glovis.usgs.gov , accessed on 8 May 2020)	grid (30 m)

In the study area, 273 landslides were recorded based on the field survey, and they were divided randomly into two parts in a ratio of 7:3: training set (TS, 191 landslides) and validation set (VS, 82 landslides). Additionally, 514 non-landslide points (i.e., no risk) were identified and selected in the areas where slope is less than 3° , e.g., urban areas or cropland where there is low probability of landslide occurrence. A total of 360 non-landslide points were added to the TS for the successive risk modeling and 154 points were incorporated into the VS for verification and validation purposes.

2.3. Geo-Environmental Factors

The geological and environmental factors involved in the study are presented in the following.

2.3.1. Elevation

Elevation is a fundamental characteristic of landform. A number of studies have revealed that there is an obvious relationship between geo-hazards and elevation [45–50], as the latter is an important indicator of the potential energy of a slope. The higher the incline in elevation, the higher the potential of energy.

2.3.2. Slope

Slope is a critical environmental factor that affects the occurrence of geo-hazards as it determines the possibility of slope mass displacement, or rather, soils in the weathered crust that have vertically different layers. The slope indirectly affects the relative stability between the layers and this determines, to a large extent, the mechanism of the incline deformation and formation of tension joints or fissures that make the slope massif unstable [51,52]. Slope may also influence rainwater permeation and water erosion. However, landslides may occur only when the slope reaches certain thresholds, e.g., $28\text{--}38^\circ$ [53], in natural conditions. Generally, bank slopes of large rivers, lakes (reservoirs), seas, and ditches, open front hillsides, railways, highways, and slope-cutting for building, etc., are all prone to landslides (Figure 2b).

2.3.3. Aspect

Different aspects result in different exposure to sunlight, which brings about differences in vegetation coverage, evapotranspiration, soil erosion, and especially, a different weathering process. These will, in turn, influence the slope runoff, rainfall penetration, the physical and mechanical properties of slope mass, and finally, the stability of the incline [8,54]. Aspects are presented in Figure 2c.

2.3.4. Lithology

Stratigraphic lithology is an important factor for all geo-hazards. On the one hand, lithology determines the development of joints and fractures of rocks under stress or tension. On the other hand, if the lithological characteristics of a given slope are different from those of the underlying rocks, there will be a high potential to form an interbedding or unconformity surface favorable for landslide. Given the same other condition, different lithologies may have different resistance to landslide events, e.g., rigid granitic massif would be more resistant than shale and mudstone. Instead of using AHP as in Roccati et al. [24], we followed Zhang et al. [8] in assigning a proneness weight to each category of lithology shown in Table 2. The assignment rule is that the higher the value, the higher the tendency of the lithology to landslides.

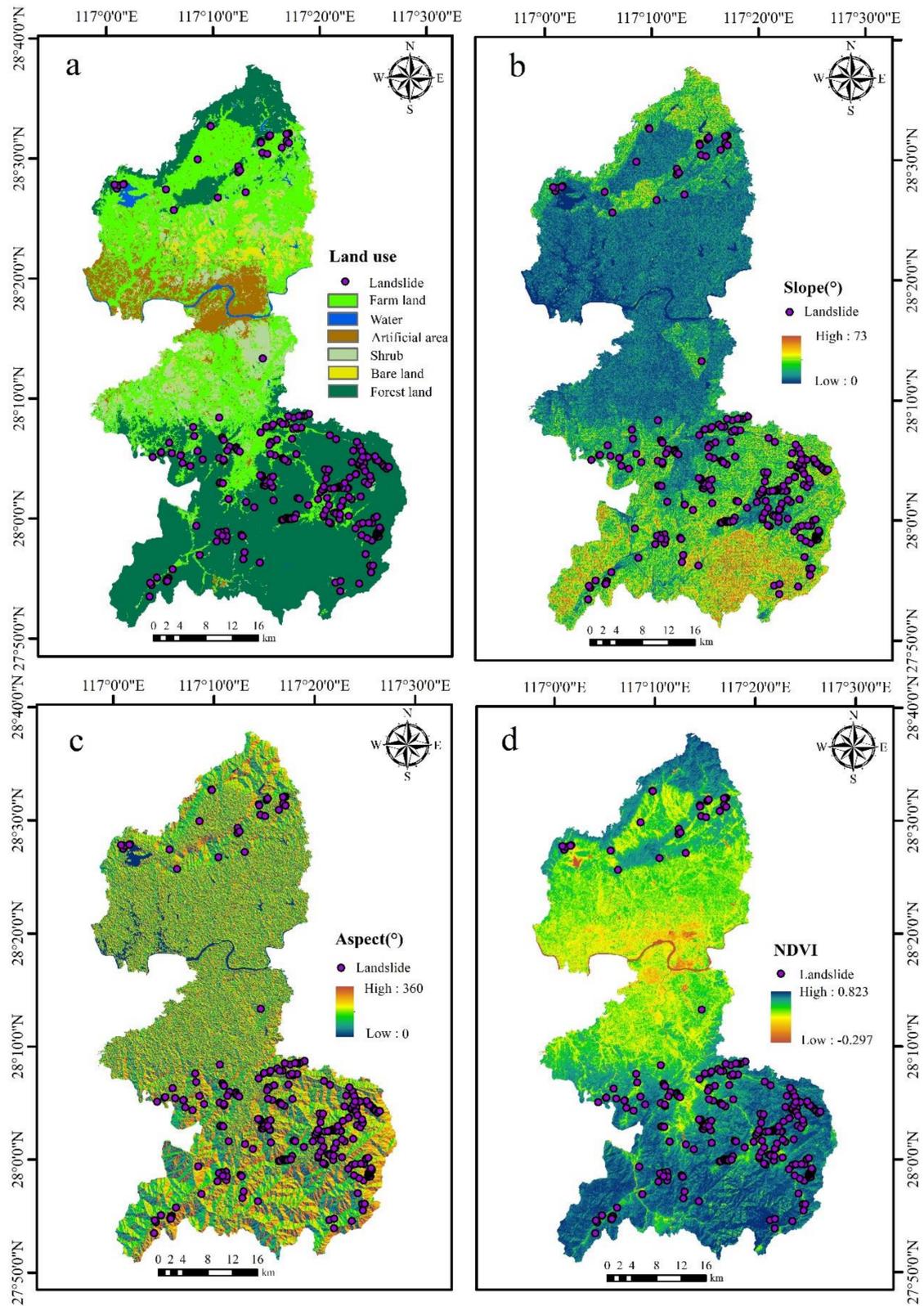


Figure 2. Geo-environmental factors used in the study: (a) land cover, (b) slope, (c) aspect, (d) NDVI.

Table 2. Proneness weight of the different lithological formations.

Lithology	Granitic and Volcanic Rocks	Sandstone	Metamorphic Rocks	Limestone and Carbonatite	Mudstone and Shale
Weight	1	5	7	8	10

2.3.5. Proximity to Faults

The influence of the fault structure on the landslide mainly lies in that whether it is a normal or a reverse fault, a certain width of the fracture zone is generated on both sides of it. In areas where fractures or joint fissures are well developed, dense, weak, and vulnerable zones may be formed [37]. Thus, the nearer to the faults, the higher susceptibility to landslide; and the bigger the scale of the faults, the greater the impacts they exert. In this study, the distance to faults was calculated by a buffer operation within GIS, and the fault buffers were defined as 0–30 m, 30–60 m, 60–90 m, 90–120 m, and >120 m for small scale faults (<10 km in length), which were assigned a weight value of 13, 10, 7, 4, and 1, respectively, and 0–60 m, 60–120 m, 120–180 m, 180–240 m, and >240 m for bigger faults (≥ 10 km in length), which were assigned a weight value of 20, 15, 10, 5, and 1, respectively [8,10] (Figure 3c).

2.3.6. Proximity to Roads

Slope cutting and excavation for road construction and housing development destroyed the stability of the incline, provoking landslides [8,10,24]. In terms of our understanding about the impacts of road construction on landslides in the field, it was noted that distance plays a key role in such hazard events, i.e., the closer to the road, the higher the propensity of landslide. Thus, the distance to roads was regarded as a critical factor which was calculated by a buffering operation. The road buffers were defined with intervals of 0–30 m, 30–60 m, 60–90 m, and 90–120 m, and assigned, respectively, a weight value of 10, 8, 4, and 1 [8,10] (Figure 3a).

2.3.7. Proximity to Rivers

River development is also an unignorable factor for geo-hazards because it erodes and liquidizes its banks or bottom slopes on the two sides, causing collapses and landslides [8]. The development of a river system favors the occurrence of the geo-hazard. Similar to other linear factors, the closer to the rivers, the higher the propensity of landslides. Hence, the distance to rivers was also used as an environmental factor. It was buffered into different zones, e.g., 0–30 m, 30–60 m, 60–90 m, 90–120 m, and 120–150 m for small rivers and streams, and 0–60 m, 60–120 m, 120–180 m, 180–240 m, and 240–300 m for the main river, i.e., Xinjiang River. A weight was allocated to each buffer zone based on its proclivity to landslide, e.g., zones 0–30 m, 30–60 m, 60–90 m, 90–120 m, and 120–150 m were assigned, respectively, 12, 10, 6, 3, and 1, and zones 0–60 m, 60–120 m, 120–180 m, 180–240 m, and 240–300 m, respectively, 20, 16, 10, 5, and 1 [8,10] (Figure 3b).

2.3.8. Precipitation

A large number of landslides show that precipitation is the first important cause of their occurrence. There are generally two cases: either sufficient precipitation duration or intensity [37]. The rainy season in southern China goes from March to July. During this period, there is usually a lot of rainfall and heavy periods of rainfall occur in June–July of each year, and landslide hazards frequently take place in these two months. Therefore, the annual mean precipitation, the accumulated average precipitation of March–July, March–June, and May–July of the past 20 years were obtained from 40 ground weather stations, and then calculated and interpolated with the inverse distance weighting (IDW) method into a grid. These precipitation layers were selected for our risk modeling to investigate which months' combination of rainfall would best reveal its role in the landslide events.

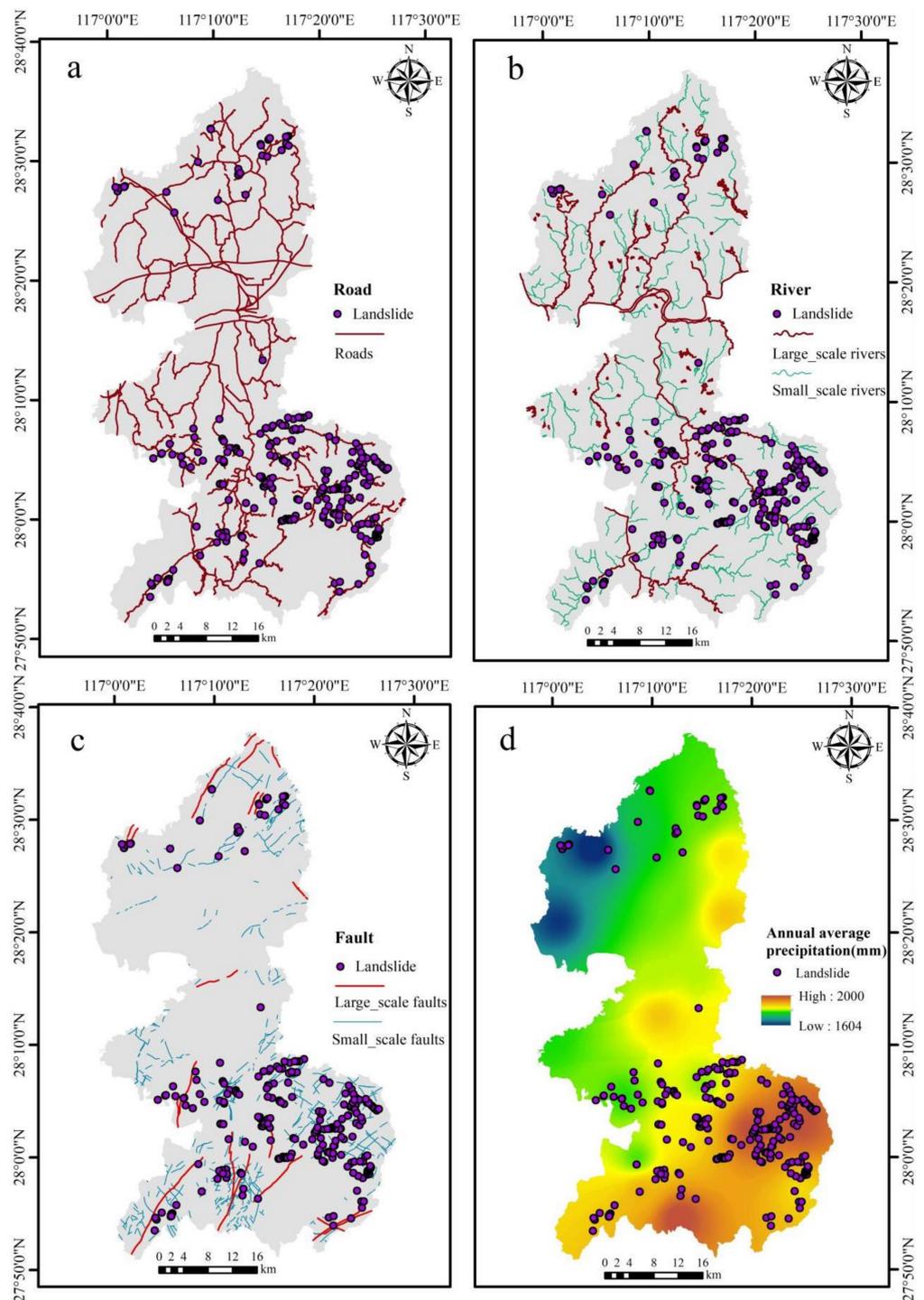


Figure 3. Geo-environmental factors used in the study: (a) roads, (b) rivers, (c) faults, (d) annual average precipitation.

2.3.9. Land Cover Types

Land use refers to all activities that humans impose on land for survival and development. This can be not only production activities such as planting crops, trees, pasture, construction of factories, roads, etc., but also non-production activities such as construction of scenic tourist areas and nature reserves. With increasing demand from human beings, land use has become more and more intensive, resulting in a decline in woodlands/forests and grasslands in some regions or countries [55]. Such land use and change may exert a

significant impact on soil erosion, especially, when there is concentrated rainfall [55,56]. The development of burn-and-slash agriculture makes land prone to landslide hazards. In fact, there are at present six land cover types in the study area, namely farmlands, water bodies (rivers and lakes), built-up areas, forests, shrubs, and finally bare land, and these were allocated a proclivity weight of 0, 0, 0, 1, 4, and 10, respectively [8,10] (Figure 2a). Here, forest cover has the least proclivity while bare land without vegetation protection is the most vulnerable category when the other conditions are the same.

The land cover map was produced from Landsat TM images dated 31 May (Spring) and 7 November 2010 (Autumn) based on the approach proposed by Wu et al. [57].

2.3.10. NDVI

The NDVI (normalized difference vegetation index) reveals information about vegetation cover, e.g., vegetation vigor, density, biomass, and its root system. Roots are able to hold soil and stabilize the hillside mass, and reduce the impact of precipitation on the soil. The level of vegetation coverage, in particular forest cover, reflects the degree of vegetation protection of the land surface. Under the same conditions, areas with greater vegetation coverage are less prone to landslide hazards.

The NDVI of the study area was generated from Landsat 5 TM images acquired in late autumn (late October and early November) in the five-year period 2005–2010. After atmospheric correction using the COST model [55,58], an averaged five-year NDVI was produced to represent the natural vegetation (e.g., forests, woodlands, and shrublands) as croplands were harvested and herbaceous vegetation became withered at that time. This mean NDVI image ranges from -0.297 to 0.823 (Figure 2d).

2.4. Methods

2.4.1. MLR Model

Based on the linear regression in which the dependent variables are binomial, the MLR model applies a logical function to solve the nonlinear problems and this conversion has made MLR highly successful in the field of machine learning. For this reason, it has been increasingly drawing the attention of geographers, biologists, and environmentalists in recent decades thanks to its ability to analyze spatially explicit causes and predict the potential changes and events in the ecosystem [55]. It has hence been widely applied in deforestation [59,60], land use change [61–63], and landslide prediction [64]. Within MLR, independent variables can be continuous or discrete and do not necessarily meet the requirement of normal distribution. Supposing that the probability of occurrence of geo-hazards is P , and the probability of non-occurrence is $1 - P$, then $\ln[P/(1 - P)]$ is taken as the dependent variable, the hazard-causative factors x_i ($i = 1, 2, 3, \dots, n$) as the independent variables, and the MLR model can be expressed as a linear regression in Equation (1).

$$\ln[P/(1 - P)] = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \quad (1)$$

The probability of landslide occurrence can be computed with Equation (2).

$$\begin{cases} Z = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \\ P = \frac{\exp(Z)}{1 + \exp(Z)} \end{cases} \quad (2)$$

where P is the calculated probability of landslide occurrence, x_i ($i = 1, 2, 3, \dots, n$) are the geo-environmental factors as independent predictive variables, α_i ($i = 1, 2, 3, \dots, n$) are the coefficients of these predictive variables, and α_0 is a constant.

After calculation of the probability of landslide occurrence, the landslide risk zoning map can be created.

2.4.2. Modeling Procedures

For landslide risk zoning, it is requisite to determine the potential role of the geo-environmental factors in the landslide event, in other words, to understand which factors

have a greater effect on this hazard occurrence than others. In this research, the prediction and assessment of landslide hazards in Guixi was conducted using two MLR models, i.e., CF-based and RP-based MLR models.

CF Calculation

Proposed by Shortliffe and Buchanan for certainty reasoning in medical diagnosis in 1975 [65], the CF was redefined and improved by Heckerman in 1986 [66]. It is a probability function [67], which is used to examine the tendency of factors governing the occurrence of an event. As the evaluation process is relatively simple and accurate, it is widely used in geo-hazard risk assessment [68–72]. The premise is to assume that the geo-hazards that have occurred and those that will probably take place in the future are under the same geological and environmental conditions. The calculation formula is shown in Equation (3):

$$CF = \begin{cases} \frac{PP_a - PP_s}{PP_a(1 - PP_s)}, PP_a \geq PP_s \\ \frac{PP_a - PP_s}{PP_s(1 - PP_a)}, PP_a < PP_s \end{cases} \quad (3)$$

where PP_a is the conditional probability, or rather, the ratio of the area or number of landslide hazards appearing in subset a to the area of subset a . PP_s is the probability of occurrence of the whole event, and in our case, it can be represented as the ratio of the landslide area or number occurring in the whole study area to the surface area of the study area.

According to Equation (3), the CF varies from -1 to 1 . The closer the CF to 1 , the greater the possibility of a landslide hazard or vice versa. When CF is 0 , it indicates the same probability of landslide occurrence and non-occurrence.

Superimposing the training samples (landslides and non-landslides) on each factor layer that has been divided into subsets, the CF approach was applied to calculate the CF value of each subset of the 13 geo-environmental factors. The CF-based dataset was hence produced and the areas and numbers of the subsets of a part of the factors are presented in Table 3.

Table 3. Division and CF values of the main geo-environmental factors.

Variable	Subset	Area of Subset (km ²)	Area of Landslide (km ²)	CF
Elevation	20–267	1658.736	0.3951	−0.268
	267–515	355.597	0.1134	−0.018
	515–762	160.439	0.0135	−0.742
	762–1009	73.526	0.0063	−0.735
	1009–1257	32.552	0.0018	−0.831
	1257–1504	8.955	0.0000	−1
Slope	<3	271.531	0.0234	−0.735
	3–9	774.307	0.1395	−0.446
	9–15	448.984	0.1800	0.188
	15–21	296.966	0.1062	0.092
	21–27	220.882	0.0567	−0.209
	27–33	144.243	0.0171	−0.634
	33–39	79.566	0.0063	−0.757
	39–45	36.940	0.0009	−0.926
	>45	16.386	0.0000	−1
Aspect	Flat	23.6754	0	−1.000
	North	143.12	0.0261	−0.44
	Northeast	299.695	0.063	−0.354
	East	313.212	0.0576	−0.434
	Southeast	302.942	0.0621	−0.369
	South	281.11	0.0873	−0.043
	Southwest	301.716	0.0909	−0.074
	West	323.207	0.0684	−0.348
	Northwest	311.566	0.0549	−0.459
	North	146.109	0.0198	−0.582

Table 3. Cont.

Variable	Subset	Area of Subset (km ²)	Area of Landslide (km ²)	CF
Proximity to roads	0–30	73.5246	0.306	0.922
	30–60	71.1603	0.1251	0.815
	60–90	69.0336	0.0369	0.393
	90–120	66.8322	0.0117	−0.462
Proximity to rivers	0–30	50.4495	0.054	0.696
	30–60	48.5829	0.045	0.649
	60–90	47.0187	0.0459	0.667
	90–120	45.9873	0.027	0.446
	0–60	69.4548	0.0243	0.071
	60–120	62.4204	0.0126	−0.379
	120–180	57.9708	0.0135	−0.283
180–240	54.7452	0.0243	0.268	
Proximity to faults	0–30	37.4526	0.0198	0.386
	30–60	38.0592	0.0189	0.346
	60–90	38.9853	0.0153	0.171
	90–120	39.2265	0.0189	0.326
	0–60	17.5698	0.0045	−0.212
	60–120	17.9028	0.009	0.354
	120–180	17.7084	0.0054	−0.062
180–240	17.1873	0.0063	0.114	
NDVI	−0.297–−0.073	10.034	0.0000	−1
	−0.073–0.151	30.311	0.0027	−0.726
	0.151–0.375	687.569	0.0954	−0.572
	0.375–0.599	737.238	0.3429	0.301
	0.599–0.823	824.746	0.0891	−0.668
Annual average precipitation	1604.7–1683.8	160.9317	0.0144	−0.723
	1683.8–1762.8	204.5124	0.0081	−0.877
	1762.8–1841.9	947.5398	0.1251	−0.594
	1841.9–1920.9	698.8761	0.1863	−0.179
	1920.9–2000	277.9641	0.1962	0.54

RP Resampling

Within GIS, 50,000 RPs were generated and used to intersect samples (i.e., landslide event polygons for training), and 1858 points were retained in the polygons (0 or 1) in the dependent variable layer, in which 589 points fell in the landslide areas and 1269 points in the no-risk areas. These points were superimposed and intersected with each geo-environmental factor layer to obtain the values corresponding to the landslides and no-risk areas. Thus, an integrated dataset including the training set, i.e., RP-based dataset, was produced.

MLR Modeling

The RP-based and CF-based datasets were input into SPSS20.0 for binary MLR analysis, where landslide events were treated as a dependent variable and other geo-environmental factors as independent (predictive) variables. The first and second landslide risk models, i.e., RP-based MLR-1 and CF-based MLR-2, were thus obtained. The regression coefficient of each factor was considered as its contribution weight to the landslide event.

Risk Mapping

After modeling, the two models were applied back to the predictive variables to calculate the landslide hazard risk of the whole study area. Landslide risk zoning was conducted based on the calculated probability, which was further graded into different risk levels, e.g., when the probability is 0–0.2, the zone is considered stable (no risk), 0.2–0.4 low, 0.4–0.6 moderate, 0.6–0.8 high, and 0.8–1.0 extremely high risk.

2.4.3. Verification and Reliability Analysis

Overall accuracy (OA) and Kappa coefficient [8,11,57,73,74] were the metrics to evaluate the accuracy and reliability of the landslide risk models produced by MLR modeling using RP- and CF-based datasets against the VS. As Landis and Koch [74] interpreted, if KC reaches 0.81–1.00, the prediction reliability is “almost perfect” [8].

2.4.4. Multicollinearity Diagnosis

Collinearity means that two predictive variables are linearly correlated to each other in a regression model, and if there are more than two variables involved, we call it multicollinearity. In such a case, it is not possible to conduct regression modeling as independent variables cannot independently predict the dependent variable. MLR modeling is susceptible to such collinearities among the predictive variables. Tolerance (TOL) and variance inflation factor (VIF) are two criteria for diagnosing multicollinearity [75,76]. If TOL is <0.1 or $VIF > 10$, it indicates a strong multicollinearity among the independent variables [77].

3. Results

The intermediate processing and final results are presented in Tables 4–7, and Figure 4 in this section for further discussion.

Table 4. Results of random points (RP)-based MLR modeling (Model MLR-1).

Variable	β	SE	Wals	df	Sig	exp(β)	Order
Elevation	−0.007	0.001	55.015	1	0.000	0.993	7
Slope	0.135	0.016	75.712	1	0.000	1.145	3
Aspect	0.003	0.001	7.854	1	0.005	1.003	6
Proximity to roads	0.440	0.025	298.674	1	0.000	1.552	1
Proximity to rivers	0.024	0.021	1.259	1	0.262	1.024	4
Proximity to faults	−0.007	0.033	0.039	1	0.843	0.993	7
NDVI	−2.155	1.028	4.389	1	0.036	0.116	12
Lithology	−0.145	0.037	15.269	1	0.000	0.865	10
Land use	0.010	0.078	0.016	1	0.899	1.010	5
Accumulated average precipitation of March–July	0.310	0.061	25.917	1	0.000	1.364	2
Accumulated average precipitation of May–July	−0.119	0.032	14.123	1	0.000	0.888	9
Accumulated average precipitation of March–June	−0.206	0.028	54.259	1	0.000	0.814	11
Annual average precipitation	−0.032	0.016	3.922	1	0.048	0.969	8
Constant	−6.675	3.885	2.952	1	0.086	0.001	

Table 5. Results of certainty factor (CF)-based MLR modeling (Model MLR-2).

Variable	β	SE	Wals	df	Sig	exp(β)	Order
Elevation	1.720	0.728	5.584	1	0.018	5.585	5
Slope	2.645	0.358	54.524	1	0.000	14.083	1
Aspect	0.782	0.600	1.699	1	0.192	2.186	8
Proximity to roads	2.608	0.150	301.480	1	0.000	13.572	2
Proximity to rivers	0.592	0.258	5.266	1	0.022	1.808	10
Proximity to faults	−0.085	0.573	0.022	1	0.882	0.918	12
NDVI	1.403	0.254	30.580	1	0.000	4.067	6
Lithology	2.225	0.313	50.623	1	0.000	9.256	4
Land use	0.729	0.570	1.640	1	0.200	2.074	9
Accumulated average precipitation of March–July	−1.582	0.813	3.787	1	0.052	0.206	13
Accumulated average precipitation of May–July	0.553	0.757	0.533	1	0.465	1.738	11
Accumulated average precipitation of March–June	2.400	0.674	12.681	1	0.000	11.019	3
Annual average precipitation	0.908	0.599	2.294	1	0.130	2.479	7
Constant	2.999	0.389	59.515	1	0.000	20.058	

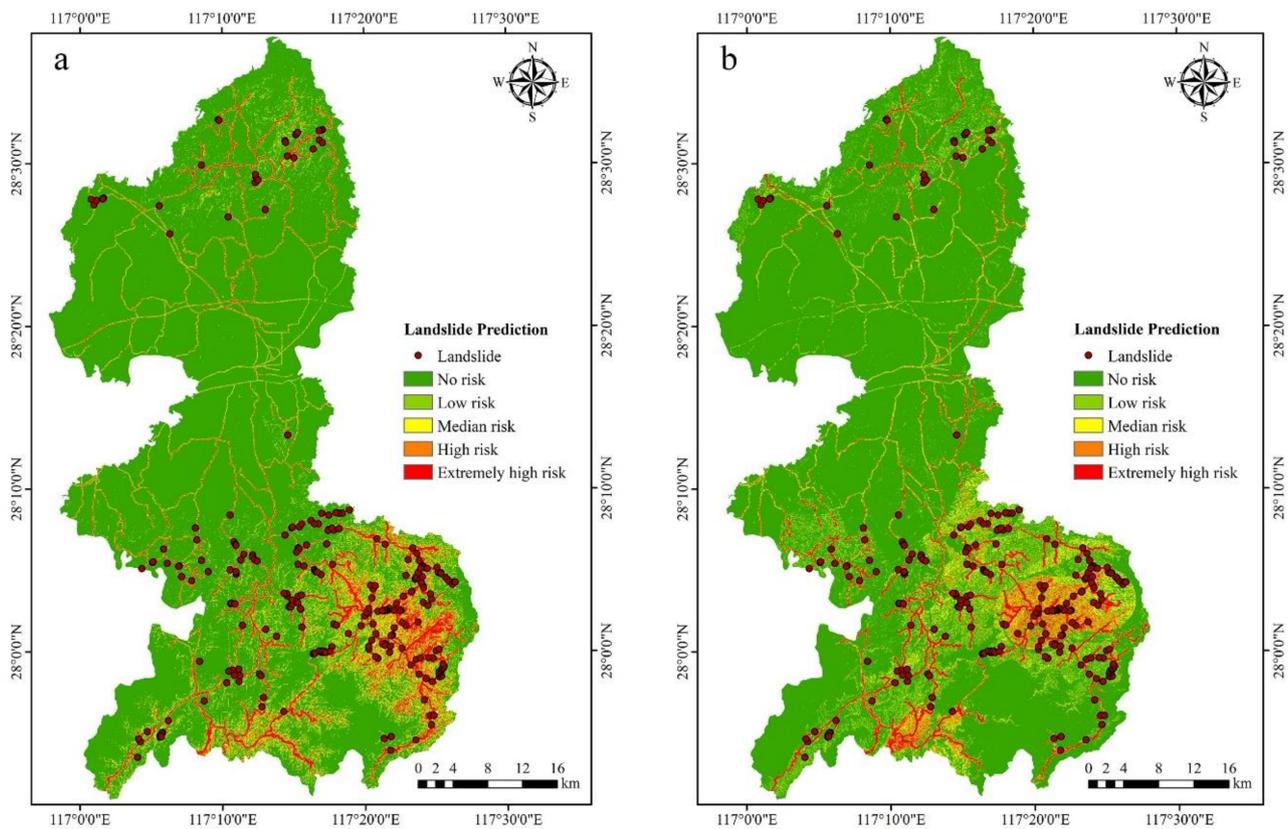
Note: β —regression coefficient of each predictive factor in the model, SE—standard error, Wals—Chi-square value, df—degree of freedom, and Sig—significance.

Table 6. Overall accuracy (OA) and Kappa coefficient (KC) of the two MLR models versus the validation set (VS).

Model	Overall Accuracy (%)	Kappa Coefficient
RP-based MLR-1	91.5%	0.814
CF-based MLR-2	90.4%	0.782

Table 7. Observation accuracy of the predicted risk map versus the observed landslides.

Risk Level	Distribution of the Field Observed Landslides	
	MLR-1	MLR-2
No risk	4.76%	6.60%
Low risk	3.66%	1.83%
Median risk	4.76%	2.93%
High risk	9.89%	6.60%
Extremely high risk	76.92%	82.05%

**Figure 4.** Landslide risk maps: (a) from model MLR-1, and (b) from model MLR-2.

3.1. MLR Models

RP-based and CF-based MLR models are shown in Tables 4 and 5, respectively, where the estimated coefficients of the predictive factors (independent variables) are demonstrated. These models were utilized to compute the probability of landslide risk for the whole study area. The observation accuracy of the two modeling results are illustrated in Table 6. The two obtained models MLR-1 and MLR-2 are expressed in Equations (4) and (5):

$$P = \frac{\exp(-6.675 - 0.007x_1 + 0.135x_2 + \dots + 0.310x_{10})}{1 + \exp(-6.675 - 0.007x_1 + 0.135x_2 + \dots + 0.310x_{10})} \quad (4)$$

$$P = \frac{\exp(2.999 + 1.720x_1 + 2.645x_2 + \dots + 0.908x_{13})}{1 + \exp(2.999 + 1.720x_1 + 2.645x_2 + \dots + 0.908x_{13})} \quad (5)$$

In Equations (4) and (5), x_1 —elevation, x_2 —slope, x_3 —aspect, x_4 —proximity to roads, x_5 —proximity to rivers, x_6 —proximity to faults, x_7 —NDVI, x_8 —lithology, x_9 —land cover, x_{10} —accumulated average precipitation of March–July, x_{11} —accumulated average pre-

precipitation of May–July, x_{12} —accumulated average precipitation of March–June, and x_{13} —average annual rainfall.

It is worth mentioning that in Tables 4 and 5, we can calculate the odds ratio (OR) [55,62] using $\exp(\beta)$ between the geo-environmental factors to understand the likelihood contribution to the hazard event of one factor over the other.

Evaluated versus the VS, the OA and KC of the two MLR models are 91.5% and 0.814 and 90.4% and 0.782, respectively (Table 6). Based on the interpretation of Landis and Koch (1976), the modeling result of Model MLR-1 reaches “almost perfect” while that of MLR-2 is “substantially” good, though the observation accuracy of the latter appears to be slightly higher as more field-observed landslides are located in the predicted high and extremely high zones (Table 7 and Figure 4).

3.2. Collinearity

The results of the collinearity diagnosis are demonstrated in Table 8. For model MLR-1, the TOL of all rainfall factors was <0.1 or $VIF > 10$, which means that they are correlated and it is not necessary to use all of them for MLR modeling. However, for model MLR-2, the TOL was >0.1 with $VIF < 10$, indicating that there was not a serious multicollinearity problem among the independent factors for CF-based MLR modeling.

Table 8. Collinearity of the geo-environmental factors of the two models: MLR-1 and MLR-2.

Variable	MLR-1		MLR-2	
	Collinearity Statistics		Collinearity Statistics	
	Tolerance	VIF	Tolerance	VIF
Elevation	0.403	2.479	0.919	1.088
Slope	0.541	1.847	0.705	1.419
Aspect	0.919	1.088	0.857	1.167
Proximity to roads	0.596	1.678	0.587	1.702
Proximity to rivers	0.863	1.159	0.871	1.148
Proximity to faults	0.937	1.067	0.928	1.078
NDVI	0.389	2.570	0.851	1.176
Lithology	0.652	1.534	0.686	1.458
Land use	0.932	1.073	0.847	1.181
Accumulated average precipitation of March–June	0.014	73.495	0.178	5.623
Accumulated average precipitation of May–July	0.011	92.400	0.129	7.745
Annual average precipitation	0.009	106.564	0.124	8.074
Accumulated average precipitation of March–July	0.002	571.362	0.116	8.649

3.3. Landslide Risk Maps

The calculated risk probability from the two models, MLR-1 (Equation (4)) and MLR-2 (Equation (5)), ranging from 0 to 1.0, was classified into five grades: stable (no risk, 0–0.2), low (0.2–0.4), moderate (0.4–0.6), high (0.6–0.8), and extremely high risk (0.8–1.0). The zonation map of landslide hazard was thus created and presented in Figure 4.

The two maps show that the areas of Guixi susceptible to landslide hazard are largely spread alongside the roads. The high risk and extremely high risk areas are mainly distributed in the towns of Wenfang, Jintun, and Zhangping in the south and Baitian in the north of the study area. The central part of Guixi is a relatively flat plain with a low risk of landslides.

As seen in Tables 4 and 5, the regression coefficients of the geo-environmental factors obtained from MLR analysis denote the order of importance. For example, the importance of roads, slope, and rainfall has been well illustrated in MLR-1 and MLR-2. Road construction and housing development destroy the original stability of the slope massif, triggering the downward slide of the overlying geological strata and the weathered crust along the bedding surface or the downhill slide of the slope massif on the unconformed surface with the underlying rocks. Hence, these human activities form a key part of the landslide

disaster. Precipitation is considered the most important triggering factor of this kind of disaster, especially in south China.

4. Discussion

4.1. Digitization and Weight Assignment

In order to model and predict the landslide risk in Guixi by the MLR algorithm, we employed 13 geological and environmental factors either as continuous or discrete variables, together with field observation data. The non-numeric factors (e.g., geological strata, faults, road, and rivers) were first digitized and then allocated a propensity weight either to different lithology or to buffers of the linear features in line with their proximities to the latter to make these factors digitized. Thence, such processing made data-driven approaches possible and successful in the landslide risk modeling of this study.

4.2. Importance of the Variables in Landslide Events

As revealed in Tables 4 and 5, the first three important geo-environmental factors of Model 1 (MLR-1) are the proximity to roads, the accumulated average precipitation of March–July, and the slope, and of Model 2 (MLR-2) are the slope, the proximity to roads, and the accumulated average precipitation of March–June. Both models have illustrated more or less the similar correlation of the hazard-causative factors such as roads, slope, and accumulated average rainfall of March–June or March–July with landslide events. The accuracy of the RP-based Model 1 seems to be higher than the CF-based Model 2 (Table 6). As already mentioned earlier, most of the observed landslides were distributed along the roads or behind the houses, and resulted directly from human activities, i.e., road construction and housing development. This is the same as what we discovered in Jiangxi [8–10].

Collinearity analysis reveals that, for RP-based MLR modeling, one of the rainfall-related factors is enough, e.g., annual average precipitation or accumulated average precipitation of March–July or May–July. As demonstrated in Table 4, the accumulated average precipitation of March–July is more important than other rainfall factors, and hence can be selected for final risk mapping.

4.3. Lowering the Threshold of Slope

We noted that 234 landslides, taking up 85.9% of the total, took place in zones with a slope ranging from 3° to 21°. Yet, Fan et al. suggested that the sliding threshold be 28°–38° for natural landslides [53]. This means that the threshold has been lowered by the intervention of human activities. This finding is also supported by two other parallel studies [8–10] in the same province.

4.4. Comparison with the Results of Others

Another study [8] employed an RF algorithm to predict the landslide risk in Guixi and achieved the mapping work with an OA of 91.23% and KC of 0.82. This is close to those of Model 1 of this study, indicating that RP-based MLR modeling and prediction can achieve the same satisfactory results as RF algorithm. Ou et al. [9] conducted a comprehensive regional-scale study on landslide hazard mapping in Jiangxi using MLR based on an information value (IV) approach and achieved an accuracy of 86%. This is also comparable to what we have effectuated. In future work, a comparative analysis is to be conducted using SVM and BRT algorithms, as [31] reported that these algorithms may lead to relevant risk prediction and evaluation.

5. Conclusions

The prediction of landslide geo-hazards is a complex nonlinear process. This study has not only indicated ways of processing geo-environmental data, especially non-digital data for quantitative analysis, but also exhibits the fact that even the same modeling algorithm may reach different results if the sampling scheme is different. It is therefore

necessary to select a reasonable one for achieving reliable risk prediction. The RP-based MLR approach seems more reasonable than the CF-based one and will be recommended for extension to similar landslide hazard prediction elsewhere. The prediction and its results may provide technical support and relevant advice to decision-makers for disaster reduction and prevention management in the study area.

This study also illustrates the unfavorable role of anthropogenic activities in causing such geological disasters, confirming the results of Zhang et al., Ou et al. and Zhou et al. [8–10]. Man exploits natural resources, modifies the land surface, and, at the same time, directly breaks the original balance of the landscape. This is the origin of the artificial landslides that are related to the road system development and urbanization. Apart from a better and more comprehensive geological survey before road construction and housing development, man has to think of optimal land use planning to minimize the man-made disasters.

Author Contributions: Conceptualization, W.H. and W.W.; methodology, W.W. and W.H.; software, W.W.; validation, W.H., X.Z. (Xiaoting Zhou), T.L., J.J., M.Z. and Y.L.; formal analysis, W.H., X.Z. (Xiaoting Zhou), P.O., and Y.Z.; investigation, W.H., X.Z. (Xiaofeng Zhang), W.W., Y.S., Z.L., P.O., Y.Z., L.X., Y.L., Y.Q., S.P., C.S., Y.B., X.Z. (Xiaofeng Zhang), X.L., and W.L. and X.H.; resources, W.W., G.Z., R.C., Y.S., Z.L., X.Z. (Xiaofeng Zhang), Y.L. and T.L.; data curation, W.H., X.Z. (Xiaoting Zhou), P.O., and Y.Z.; writing—original draft preparation, W.H. and W.W.; writing—review and editing, W.W.; visualization, W.H., X.F. and J.L.; supervision, W.W.; project administration, W.W. and G.Z.; funding acquisition, W.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Jiangxi Talent Program (Grant No. 900/2120800004) and by the Start-up Fund for Scientific Research from the East China University of Technology (Grant No. DHTP2018001) granted to Weicheng Wu.

Data Availability Statement: The data that support the findings of this study are available the corresponding author, Weicheng Wu, upon reasonable request.

Acknowledgments: A part of the field survey data and the Geological Map of Guixi on scale of 1/50,000 were supplied by the 264 Geological Team of Jiangxi Nuclear Industry. DEM data (AST-GTMV003 30m) were obtained from NASA (www.earthdata.nasa.gov, accessed on 11 August 2020) and Landsat data from USGS data server (<https://glovis.usgs.gov>, accessed on 8 May 2020).

Conflicts of Interest: No potential conflict of interest was reported by the authors.

References

1. Wu, W.; Ai, G. Risk assessment of natural disasters in the course of selection of nuclear waste disposal. *J. East China Geol. Inst.* **1995**, *18*, 260–265. (In Chinese with English abstract)
2. Li, X. Study on Logistic Regression Model Applied to Regional Slope Stability Evaluation Based on GIS. Ph.D. Thesis, China University of Geosciences, Wuhan, China, 2005. (In Chinese with English abstract)
3. Bordoni, M.; Valentino, R.; Meisina, C.; Bittelli, M.; Chersich, S. A Simplified Approach to Assess the Soil Saturation Degree and Stability of a Representative Slope Affected by Shallow Landslides in Oltrepò Pavese (Italy). *Geosciences* **2018**, *8*, 472. [[CrossRef](#)]
4. Crosta, G.B.; Frattini, P. Rainfall thresholds for the triggering of soil slips and debris flows. In Proceedings of the 2nd Plinius Conference on Mediterranean Storms, Siena, Italy, 16–18 October 2000; GNDCI: Perugia, Italy, 2001; pp. 463–488.
5. Muntohar, A.S.; Liao, H.-J. Rainfall infiltration: Infinite slope model for landslides triggering by rainstorm. *Nat. Hazards* **2010**, *54*, 967–984. [[CrossRef](#)]
6. Stamatopoulos, C.A.; Di, B. Analytical and approximate expressions predicting post-failure landslide displacement using the multi-block model and energy methods. *Landslides* **2015**, *12*, 1207–1213. [[CrossRef](#)]
7. Take, W.A.; Bolton, M.D.; Wong, P.C.P.; Yeung, F.J. Evaluation of landslide triggering mechanisms in model fill slopes. *Landslides* **2004**, *1*, 173–184. [[CrossRef](#)]
8. Zhang, Y.; Wu, W.; Qin, Y.; Lin, Z.; Zhang, G.; Chen, R.; Song, Y.; Lang, T.; Zhou, X.; Huangfu, W.; et al. Mapping Landslide Hazard Risk Using Random Forest Algorithm in Guixi, Jiangxi, China. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 695. [[CrossRef](#)]
9. Ou, P.; Wu, W.; Qin, Y.; Zhou, X.; Huangfu, W.; Zhang, Y.; Xie, L.; Huang, X.; Fu, X.; Li, J.; et al. Assessment of landslide hazard in Jiangxi using geo-information technology. *Front. Earth Sci. Geohazards Georisks* **2021**, *9*. [[CrossRef](#)]
10. Zhou, X.; Wu, W.; Lin, Z.; Zhang, G.; Chen, R.; Song, Y.; Wang, Z.; Lang, T.; Qin, Y.; Ou, P.; et al. Zonation of landslide susceptibility in Ruijin, Jiangxi, China. *Int. J. Environ. Res. Public Health* **2020**. under review.
11. Wu, W.; Zucca, C.; Muhaimed, A.S.; Al-Shafie, W.M.; Al-Quraishi, A.M.F.; Nangia, V.; Zhu, M.; Liu, G. Soil salinity prediction and mapping by machine learning regression in Central Mesopotamia, Iraq. *Land Degrad. Dev.* **2018**, *29*, 4005–4014. [[CrossRef](#)]

12. Qin, Y.; Liu, L. Quantitative 3D Association of Geological Factors and Geophysical Fields with Mineralization and Its Significance for Ore Prediction: An Example from Anqing Orefield, China. *Minerals* **2018**, *8*, 300. (In Chinese with English abstract) [[CrossRef](#)]
13. Zhou, Y.; Chen, S.; Zhang, Q.; Xiao, F.; Wang, S.; Liu, Y.; Jiao, S. Advances in Big Data and Mathematical Earth Science Research—Preface to Big Data and Mathematical Earth Science Topics. *Acta Petrol. Sin.* **2018**, *34*, 255–263. (In Chinese with English abstract)
14. Liu, X. Risk zoning of avalanche and slip flow based on logistic regression model and certainty coefficient. *J. Eng. Geol.* **2014**, *22*, 1250–1256. (In Chinese with English abstract)
15. Lumb, P. Slope failures in Hong Kong. *Q. J. Eng. Geol. Hydrogeol.* **1975**, *8*, 31–65. [[CrossRef](#)]
16. Collison, A.J.C.; Anderson, M.G. Using a combined slope hydrology—Stability model to identify suitable conditions for landslide prevention by vegetation in the humid tropics. *Earth Surf. Process. Landf.* **1996**, *21*, 737–747. [[CrossRef](#)]
17. Evans, N.C. *Natural Terrain Landslide Study: Preliminary Assessment of the Influence of Rainfall on Natural Terrain Landslide Initiation*; Geotechnical Engineering Office: Hong Kong, China, 1997; pp. 28–29.
18. Mark, R.K.; Ellen, S.D. *Statistical and Simulation Models for Mapping Debris Flow Hazard*; Carrara, A., Guzetti, F., Eds.; Geographical Information System in Assessing Natural Hazard; Academic Publishers: Amsterdam, The Netherlands, 1995; pp. 93–106.
19. Jiang, H.; Eastman, J.R. Application of fuzzy measures in multi-criteria evaluation in GIS. *Int. J. Geogr. Inf. Sci.* **2000**, *14*, 173–184. [[CrossRef](#)]
20. Wan, E.T.; Mitani, Y.; Xie, M.; Mori, J. Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach. *Eng. Geol.* **2003**, *68*, 373–386.
21. Chau, K.T.; Sze, Y.L.; Fung, M.K.; Wong, W.Y.; Fong, E.L.; Chan, L.C.P. Landslide hazard analysis for Hong Kong using landslide inventory and GIS. *Comput. Geosci.* **2004**, *30*, 429–443. [[CrossRef](#)]
22. Meng, Q.; Sun, W.; Wang, T. Research on geological disasters susceptibility evaluation in Fengxian, Shaanxi. *J. Eng. Geol.* **2011**, *19*, 388–396. (In Chinese with English abstract)
23. Zhang, C.; Chen, Y.; Zhang, Y.; Sun, X. Evaluation of the susceptibility of geological disasters in Zhaotong, Yunnan based on multiple linear regression model. *Hydrogeol. Eng. Geol.* **2016**, *43*, 159–163. (In Chinese with English abstract) [[CrossRef](#)]
24. Roccati, A.; Paliaga, G.; Luino, F.; Faccini, F.; Turconi, L. GIS-Based Landslide Susceptibility Mapping for Land Use Planning and Risk Assessment. *Land* **2021**, *10*, 162. [[CrossRef](#)]
25. Tian, C.; Liu, X.; Wang, J. Evaluation of the susceptibility of geological disasters in Guangdong Province based on CF and Logistic regression models. *Hydrogeol. Eng. Geol.* **2016**, *43*, 154–161. (In Chinese with English abstract) [[CrossRef](#)]
26. Yue, X.; Huang, M.; Xu, Q.; Chen, C.; Gu, X. Susceptibility evaluation of debris flow disaster in Karst area of Guizhou Province. *J. Earth Inf. Sci.* **2015**, *17*, 1396–1403. (In Chinese with English abstract)
27. Huang, F.; Yin, K.; Jiang, S.; Huang, J.; Cao, Z. Assessment of landslide susceptibility based on cluster analysis and support vector machine. *J. Rock Mech. Eng.* **2018**, *37*, 156–167. (In Chinese with English abstract) [[CrossRef](#)]
28. Yang, H.; Wu, B.; Wang, L. Application of PSO-RBF coupling model of chaotic sequence in landslide displacement prediction. *Sci. Technol. Eng.* **2013**, *13*, 9118–9121, 9126. (In Chinese with English abstract)
29. Catani, F.; Lagomarsino, D.; Segoni, S.; Tofani, V. Landslide susceptibility estimation by random forests technique: Sensitivity and scaling issues. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 2815–2831. [[CrossRef](#)]
30. Pourghasemi, H.R.; Kornejady, A.; Kerle, N.; Shabani, F. Investigating the effects of different landslide positioning techniques, landslide partitioning approaches, and presence-absence balances on landslide susceptibility mapping. *Catena* **2020**, *187*, 104364. [[CrossRef](#)]
31. Xiong, K.; Adhikari, B.R.; Stamatopoulos, C.A.; Zhan, Y.; Wu, S.; Dong, Z.; Di, B. Comparison of Different Machine Learning Methods for Debris Flow Susceptibility Mapping: A Case Study in the Sichuan Province, China. *Remote Sens.* **2020**, *12*, 295. [[CrossRef](#)]
32. Huang, F.; Yin, K.; Huang, J.; Gui, L.; Wang, P. Landslide susceptibility mapping based on self-organizing-map network and extreme learning machine. *Eng. Geol.* **2017**, *223*, 11–22. [[CrossRef](#)]
33. Xia, H.; Yin, K.; Liang, X.; Ma, F. Evaluation of landslide susceptibility based on SVM-ANN model—Taking Wushan County in the Three Gorges Reservoir Area as an example. *Chin. J. Geol. Hazard Control* **2018**, *29*, 13–19. (In Chinese with English abstract) [[CrossRef](#)]
34. Li, Y.; Mei, H.; Ren, X.; Hu, X.; Li, M. Evaluation of geological hazard susceptibility based on deterministic coefficient and support vector machine. *J. Geo-Inf. Sci.* **2018**, *20*, 1699–1709. (In Chinese with English abstract)
35. Feng, F.; Wu, X.; Niu, R.; Xu, S.; Yu, X. Landslide sensitivity evaluation of BP neural network based on particle swarm optimization. *Surv. Mapp. Sci.* **2017**, *42*, 170–175. (In Chinese with English abstract) [[CrossRef](#)]
36. Xu, C.; Dai, F.; Yao, X.; Zhao, Z.; Xiao, J. GIS platform and certainty factor analysis method based Wenchuan earthquake-induced landslide susceptibility evaluation. *J. Eng. Geol.* **2010**, *18*, 15–26. (In Chinese with English abstract)
37. Fang, M.; Zhang, J.; Xu, Y. Research on Lanzhou City Landslide Disaster Sensitivity Based on GIS and Logistic Regression Model. *Remote Sens. Technol. Appl.* **2011**, *26*, 845–853. (In Chinese with English abstract)
38. Du, Q.; Fan, W.; Li, K.; Yang, D.; Lv, J. The Application of Binary Logistic Regression and Information Model in Geological Hazard Zoning. *Disaster Sci.* **2017**, *32*, 220–226. (In Chinese with English abstract) [[CrossRef](#)]
39. 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](#)]

40. Feby, B.; Achu, A.; Jimnisha, K.; Ayisha, V.; Reghunath, R. Landslide susceptibility modelling using integrated evidential belief function based logistic regression method: A study from Southern Western Ghats, India. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100411. [[CrossRef](#)]
41. Mandal, K.; Saha, S.; Mandal, S. Applying deep learning and benchmark machine learning algorithms for landslide susceptibility modelling in Rorachu river basin of Sikkim Himalaya, India. *Geosci. Front* **2021**, 101203, in press. [[CrossRef](#)]
42. Shahri, A.A.; Spross, J.; Johansson, F.; Larsson, S. Landslide susceptibility hazard map in southwest Sweden using artificial neural network. *Catena* **2019**, *183*, 104225. [[CrossRef](#)]
43. Othman, A.A.; Gloaguen, R.; Andreani, L.; Rahnama, M. Improving landslide susceptibility mapping using morphometric features in the Mawat area, Kurdistan Region, NE Iraq: Comparison of different statistical models. *Geomorphology* **2018**, *319*, 147–160. [[CrossRef](#)]
44. Nicu, I.C. Frequency ratio and GIS-based evaluation of landslide susceptibility applied to cultural heritage assessment. *J. Cult. Heritage* **2017**, *28*, 172–176. [[CrossRef](#)]
45. Chen, X.; Ran, H. Triggering factors susceptibility of earthquake-induced Landslides in 1976 Longling earthquake. *Acta Sci. Nat. Univ. Pekin.* **2009**, *45*, 104–110. (In Chinese with English abstract)
46. Yang, G.; Xu, P.; Cao, S.; Zhang, W.; Lan, Z.; Chen, J.; Dong, X. Evaluation of regional landslide susceptibility based on certainty factor combination model. *J. Eng. Geol.* **2019**, *27*, 1153–1163. [[CrossRef](#)]
47. Wu, S.; Li, H.; Zhang, K.; Chen, G. Susceptibility analysis of elevation factors to the development of landslide disaster in Pingchang county. *J. Three Gorges Univ.* **2015**, *37*, 62–66. [[CrossRef](#)]
48. Bălțeanu, D.; Chendeș, V.; Sima, M.; Enciu, P. A country-wide spatial assessment of landslide susceptibility in Romania. *Geomorphology* **2010**, *124*, 102–112. [[CrossRef](#)]
49. Kamp, U.; Growley, B.J.; Khattak, G.A.; Owen, L.A. GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. *Geomorphology* **2008**, *101*, 631–642. [[CrossRef](#)]
50. Süzen, M.L.; Doyuran, V. Data driven bivariate landslide susceptibility assessment using geographical information systems: A method and application to Asarsuyu catchment, Turkey. *Eng. Geol.* **2004**, *71*, 303–321. [[CrossRef](#)]
51. Yang, D.; Fan, W. ArcGIS-based zoning assessment of geological hazards susceptibility—Taking Xunyang County as an example. *Chin. J. Geol. Hazards Prev.* **2015**, *26*, 82–86. (In Chinese with English abstract) [[CrossRef](#)]
52. Zhao, Y.; Niu, R. Exploration of landslide risk zoning based on evidence weight method. *Geogr. Geogr. Inf. Sci.* **2010**, *26*, 19–23.
53. Fan, L.; Lehmann, P.; Or, D. Effects of soil spatial variability at the hillslope and catchment scales on characteristics of rainfall-induced landslides. *Water Resour. Res.* **2016**, *52*, 1781–1799. [[CrossRef](#)]
54. Lan, H.; Wu, F.; Zhou, C.; Wang, S. Sensitivity analysis of landslide factors in the Xiao Jiang River Basin in Yunnan based on GIS. *J. Rock Mech. Eng.* **2002**, *10*, 1500–1506. (In Chinese with English abstract)
55. Wu, W. Application de la Géomatique au Suivi de la Dynamique Environnementale en Zones Arides. Ph.D. Thesis, Université de Paris 1-Panthéon-Sorbonne, Paris, France, 2003. (In French).
56. Wu, W.; De Pauw, E.; Helldén, U. Assessing woody biomass in African tropical savannahs by multiscale remote sensing. *Int. J. Remote Sens.* **2013**, *34*, 4525–4549. [[CrossRef](#)]
57. Wu, W.; Zucca, C.; Karam, F.; Liu, G. Enhancing the performance of regional land cover mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 422–432. [[CrossRef](#)]
58. Wu, W.; De Pauw, E.; Zucca, C. Using remote sensing to assess impacts of land management policies in the Ordos rangelands in China. *Int. J. Digit. Earth* **2013**, *6*, 1–22. [[CrossRef](#)]
59. Nualchawee, K.; Miller, L.; Tom, C.; Christenson, J.; Williams, D. *Spatial Inventory and Modelling of Shifting Cultivation and Forest Land Cover of Northern Thailand, with Inputs from Maps, Airphotos and Landsat*; Remote Sensing Center Technical Report, No. 4177; Texas A&M University: College Station, TX, USA, 1981.
60. Lambin, E.F. *Modelling Deforestation Processes: A Review*; Tropical Ecosystem Environment Observations by Satellites; TREES Series B: Research Report No. 1, EUR15744EN; European Commission: Luxembourg, 1994; pp. 45–101.
61. Mertens, B.; Lambin, E.F. Land-Cover-Change Trajectories in Southern Cameroon. *Ann. Assoc. Am. Geogr.* **2000**, *90*, 467–494. [[CrossRef](#)]
62. Serneels, S.; Lambin, E.F. Proximate causes of land-use change in Narok District, Kenya: A spatial statistical model. *Agric. Ecosyst. Environ.* **2001**, *85*, 65–81. [[CrossRef](#)]
63. Verburg, P.H.; Soepboer, W.; Veldkamp, A.; Limpiada, R.; Espaldon, V.; Mastura, S.S.A. Modelling the spatial dynamics of regional land use: The CLUE-S model. *Environ. Manag.* **2002**, *30*, 391–405. [[CrossRef](#)]
64. Wang, X.; Huang, H.; Ban, Y.; Li, J. Drawing of landslide hazard zoning map with GIS assistance—A case study of Panzhuhua City, Sichuan Province. *Bull. Surv. Mapp.* **2015**, *2*, 46–50. (In Chinese with English abstract) [[CrossRef](#)]
65. Shortliffe, E.H.; Buchanan, B.G. A model of inexact reasoning in medicine. *Math. Biosci.* **1975**, *23*, 351–379. [[CrossRef](#)]
66. Heckerman, D. Probabilistic interpretation for MYCIN's certainty factors. *Mach. Intell. Pattern Recognit.* **1986**, *4*, 167–196.
67. Xiao, T. GIS-Based Research on Spatial Simulation of Landslide in Lanzhou. Ph.D. Thesis, College of Earth and Environmental Science, Lanzhou University, Lanzhou, China, 2006. (In Chinese with English abstract).
68. Arabameri, A.; Pradhan, B.; Rezaei, K. Gully erosion zonation mapping using integrated geographically weighted regression with certainty factor and random forest models in GIS. *J. Environ. Manag.* **2019**, *232*, 928–942. [[CrossRef](#)]

69. Azareh, A.; Rahmati, O.; Rafiei-Sardooi, E.; Sankey, J.B.; Lee, S.; Shahabi, H.; Bin Ahmad, B. Modelling gully-erosion susceptibility in a semi-arid region, Iran: Investigation of applicability of certainty factor and maximum entropy models. *Sci. Total Environ.* **2019**, *655*, 684–696. [[CrossRef](#)]
70. Costache, R.; Pham, Q.B.; Avand, M.; Linh, N.T.T.; Vojtek, M.; Vojteková, J.; Lee, S.; Khoi, D.N.; Nhi, P.T.T.; Dung, T.D. Novel hybrid models between bivariate statistics, artificial neural networks and boosting algorithms for flood susceptibility assessment. *J. Environ. Manag.* **2020**, *265*, 110485. [[CrossRef](#)] [[PubMed](#)]
71. Chen, W.; Xie, X.; Peng, J.; Shahabi, H.; Hong, H.; Bui, D.T.; Duan, Z.; Li, S.; Zhu, A. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. *Catena* **2018**, *164*, 135–149. [[CrossRef](#)]
72. Juliev, M.; Mergili, M.; Mondal, I.; Nurtaev, B.; Pulatov, A.; Hübl, J. Comparative analysis of statistical methods for landslide susceptibility mapping in the Bostanlik District, Uzbekistan. *Sci. Total Environ.* **2019**, *653*, 801–814. [[CrossRef](#)] [[PubMed](#)]
73. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [[CrossRef](#)]
74. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [[CrossRef](#)]
75. Zhu, L.; Huang, J.F. GIS-based logistic regression method for landslide susceptibility mapping in regional scale. *J. Zhejiang Univ. Sci. A* **2006**, *7*, 2007–2017. (In Chinese with English abstract) [[CrossRef](#)]
76. Ozdemir, A.; Altural, T. A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *J. Asian Earth Sci.* **2013**, *64*, 180–197. [[CrossRef](#)]
77. Belsley, D.A. *Conditioning Diagnostics: Collinearity and Weak Data in Regression*; Wiley: New York, NY, USA, 1991.