

Article

Variable-Weighted Linear Combination Model for Landslide Susceptibility Mapping: Case Study in the Shennongjia Forestry District, China

Wei Chen, Hongxing Han, Bin Huang, Qile Huang and Xudong Fu *

School of Civil Engineering, Wuhan University, Wuhan 430072, China; whucw@whu.edu.cn (W.C.); hxhan@whu.edu.cn (H.H.); cucumberhb@163.com (B.H.); huang-ican@whu.edu.cn (Q.H.)

* Correspondence: xdfu@whu.edu.cn; Tel.: +86-139-8611-1225

Received: 21 August 2017; Accepted: 3 November 2017; Published: 7 November 2017

Abstract: A landslide susceptibility map plays an essential role in urban and rural planning. The main purpose of this study is to establish a variable-weighted linear combination model (VWLC) and assess its potential for landslide susceptibility mapping. Firstly, different objective methods are employed for data processing rather than the frequently-used subjective judgments: K-means clustering is used for classification; binarization is introduced to determine buffer length thresholds for locational elements (road, river, and fault); landslide area density is adopted as the contribution index; and a correlation analysis is conducted for suitable factor selection. Secondly, considering the dimension changes of the preference matrix varying with the different locations of the mapping cells, the variable weights of each optimal factor are determined based on the improved analytic hierarchy process (AHP). On this basis, the VWLC model is established and applied to regional landslide susceptibility mapping for the Shennongjia Forestry District, China, where shallow landslides frequently occur. The obtained map is then compared with a map using the traditional WLC, and the results of the comparison show that VWLC is more reasonable, with a higher accuracy, and can be used anywhere that has the same or similar geological and topographical conditions.

Keywords: geographic information system (GIS); landslide; susceptibility; K-means clustering; binarization; analytic hierarchy process (AHP); variable-weighted linear combination; China

1. Introduction

Landslides are common geological disasters in mountainous areas that cause considerable economic and ecological damage. Many countries and regions suffer from landslides; for example, because of the complex geographical conditions and expanding human activities, more than 20,000 landslides occur in China annually, leading to nearly 1000 deaths or missing people, and a direct economic loss of six billion dollars [1,2].

A landslide susceptibility map plays an essential role in urban and rural planning. A landslide susceptibility map (LSM) that emphasizes static landslide-prone conditions presents information on the spatial distribution and probability of landslides in a certain region. In addition, these maps can assist decision makers in risk mitigation, land use management, space development, and environmental conservation to succeed in optimal development [3,4]. Using a geographic information system (GIS), various methods have been proposed, which mainly include expert evaluation, statistical methods, and mechanical approaches [5]. Expert evaluation depends on the judgment of experts, whereas mechanical approaches assess slope stability using deterministic methods and/or numerical methods with a high accuracy [6,7]. However, as they are limited by model selection and data acquisition, mechanical approaches are only applied to small regions and situations that require high accuracy. Compared with the abovementioned approaches, the statistical methods based on connections between influence

factors and landslides, not only avoid the dependence of mechanical models on high-precision data, but also reduce the subjectivity brought by expert evaluation statistical methods [8,9]. The statistical methods include the weighted liner combination model (WLC) [10,11], the logistic regression model [12–14], fuzzy synthetic evaluation model [15,16], and neural network model [17,18].

Previous studies have shown that data processing and the weight of influence factors are the key components when applying statistical methods to landslide susceptibility mapping [19]. Recently, Wang et al. (2016) [20] took the statistical index curve and Zhang et al. (2014) [21] utilized the cumulative frequency curve to classify influence factors for data processing; however, different X-axis intervals may affect the distribution of graphic features and result in different classifications. Meanwhile, buffers are built to consider the influence of locational elements on landslides in GIS analyses, and there is a need to determine buffer length thresholds before building them. Since there are no universal criteria or guidelines, scholars prefer to utilize subjective methods as a solution to this issue. Various weighting methods can be used in regional landslide susceptibility mapping. One commonly-used method is the analytical hierarchy processes (AHP) that is considered as a decision-aiding method for handling multi-criteria decisions [22–25]. However, the AHP method has some limitations [3,26], one of which is its inability to characterize the importance of each influence factor in different locations. For instance, if a mapping cell is far away from a locational element, the influence of this locational factor on landslides is negligible, leading to dimension changes in the preference matrix, and the constant weights calculated using the AHP method cannot deal with this issue well.

For these reasons, K-means clustering and binarization were introduced for data processing. Considering the different locations of mapping cells, the variable-weighted method was established based on the AHP method. On this basis, the variable-weighted linear combination model was established and applied to regional landslide susceptibility mapping for the Shennongjia Forestry District, China.

2. Study Area and Data Used

2.1. Study Area

The Shennongjia Forestry District is located within latitude $31^{\circ}15'–31^{\circ}75'$ N and longitude $109^{\circ}56'–110^{\circ}58'$ E in Northwest Hubei, China (Figure 1). The total area is about 3250 km², and the population was 79,248 as of 2014. The district is bound by the Ta-pa Mountains, with altitudes ranging from 430 m to 3100 m. Located in the mid-latitude northern subtropical monsoon zone, the district has a slightly cool and rainy climate, an average annual temperature of 12 °C, and precipitation is mainly concentrated between May and October. The district also has a dense network of rivers. The main north–south travel route is China National Highway 209, and Hubei Provincial Route 307 runs through the northeastern part of the district. Houses, tourist attractions, and other buildings are widely distributed along the rivers and roads.

In recent years, with the expansion of human activities, landslides have occurred more and more frequently in this district. Among the existing landslides, shallow landslides have the most extensive distribution (almost 80% of the total landslides over the last decade), which pose serious threats to resident life and also restrict the economy. Therefore, it is necessary for decision makers to take into account susceptibility to shallow landslides during urban and rural planning.

2.2. Data Used

2.2.1. Landslide Inventory Map

Four different datasets were prepared for landslide susceptibility mapping (Table 1). The first of these was a landslide inventory map. A landslide inventory map helps to analyze the connections between influence factors and landslides [27–30].

By means of extensive field surveys and spatial analysis using GIS, a digitized inventory map of the district was completed by the Land Resources Bureau of the Shennongjia Forestry District; from which relevant data regarding 199 shallow landslides were extracted. However, the original inventory map only presented the type, occurrence time, area, and location (central position). To get more attributes of the existing landslides, buffers were built with the landslide points as centers, the sizes of which were equal to the recorded areas of landslides. Then, using the zonal statistics tool of the mapping and analytics platform, ArcGIS 10.2 [31], the averages within the buffers were extracted from the factor layers (elevation, gradient, etc.). On this basis, 127 landslides (64%) were used for modeling and the remaining 72 landslides (36%) were used for model validation.

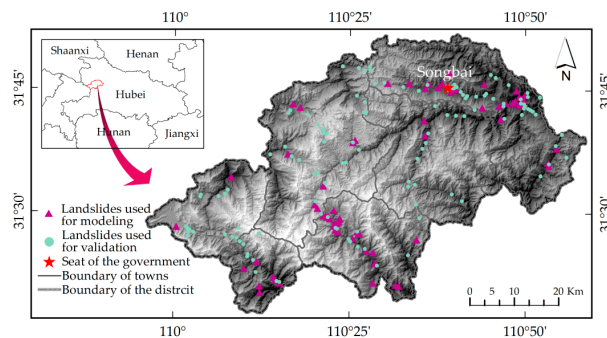


Figure 1. Location map of the Shennongjia Forestry District (provided by the Land Resources Bureau of Shennongjia).

Table 1. List of data sources used in the paper.

	Data	Type	Scale	Data source
1	Landslide inventory map	Point		Land Resources Bureau of Shennongjia Forestry District
2	DEM	Raster	10 m	School of Urban Design of Wuhan University
3	Road	Line		
	River	Line		
	Faults	Line	1:100,000	Geological map of Hubei province
	Lithology	Polygon	1:100,000	
4	Soil moisture content	Polygon	1:200,000	Hydrologic map of Hubei province
	Land use	Polygon		School of Urban Design of Wuhan University
	Precipitation			Meteorological Bureau of Shennongjia Forestry District

2.2.2. Topographical Factors

The second dataset was the digital elevation model (DEM)-based dataset, including topographical factors [32]: elevation, gradient, aspect, relief amplitude (RA), surface roughness (SR), plan curvature (PLC), and profile curvature (PRC). However, having more factors does not necessarily mean a more complete and accurate landslide susceptibility map [13,21,33]. Due to being derived from the same DEM data source, the topographical factors were not independent of each other. To select suitable factors, correlations of topographical factors were analyzed using the Pearson correlation coefficient, which is the covariance of two variables divided by the product of their standard deviations:

$$R = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) \quad (1)$$

where n is the number of samples; x_i , y_i are the single samples indexed with i ; \bar{x} is the sample mean of x_i ; \bar{y} is the sample mean of y_i ; s_x is the sample standard deviation of x_i ; s_y is the sample standard deviation of y_i . The R has a value between -1 and $+1$, where 1 is total positive linear correlation, 0 is no linear correlation, and -1 is total negative linear correlation.

Using the statistical analysis software, SPSS 22, the Pearson correlation coefficients of each topographical factor were calculated. As shown in Table 2, the following pairwise sets of factors have relatively high correlations: Gradient with RA, gradient and SR, RA with SR, PLC with PRC. Therefore, to reduce correlations as much as possible, and to make the selected factors more universal, only elevation and gradients with $R = 0.08$ were considered as influencing factors for landslide susceptibility mapping.

Table 2. Correlation analyses between topographical factors.

Factors	R						
	Elevation	Gradient	Aspect	RA	SR	PLC	PRC
Elevation	1.0	0.08	−0.05	0.05	0.04	0.02	−0.01
Gradient		1.0	−0.07	0.97	0.93	0.08	−0.12
Aspect			1.0	−0.08	−0.08	−0.07	−0.04
RA				1.0	0.96	0.10	−0.17
SR					1.0	0.09	−0.17
PLC						1.0	−0.60
PRC							1.0

2.2.3. Locational Factors

The locational factors refer to the proximity to locational elements (road, river, and fault), and the maps of locational elements constituted the third dataset. Roads affect landslides in the form of slope cuts and barriers to surface flow [34]. The influences of rivers on landslides are mainly manifested through scouring and changes of water level, which lead to a decline in support capacity. Land that is close to faults is considered to be relatively weak, with altered surface material structures and increased permeability [19,35].

2.2.4. Lithology, Aquifer Storage Capacity, Precipitation, and Land Use

The fourth dataset includes lithology, aquifer storage capacity, precipitation, and land use [36,37]. Variations of lithology lead to different resistances to erosion processes, due to varied characteristics, such as different composition, structure, and compactness [38–40]. The high storage capacity of aquifers can increase soil moisture and soften the structural plane of slopes, leading to a reduction in the original stability. For being directly digitized from public thematic maps, maps of lithology, aquifer storage capacity have been classified already: The lithology of the outcrop in the district has been classified into nine categories according to period: Triassic (T), Permian (P), Silurian (S), Ordovician (O), Cambrian (ϵ), Sinian (Z), Qingbaikouan (Q_n), and Middle-Proterozoic (Pt_2) (Figure 2a). Additionally, the aquifers have been classified into high, medium, and low, according to their relative storage capacities (Figure 2b).

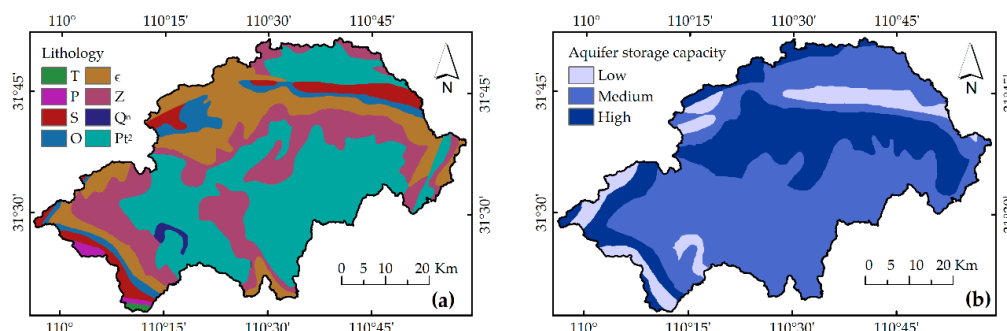


Figure 2. Input layers of influence factors: (a) lithology; (b) aquifer storage capacity.

It is universally acknowledged that precipitation is the dominant exterior factor triggering landslides [41,42]. Using precipitation data from the last ten years, from eight meteorological observation stations in the district, a map of average annual precipitation was produced using the Kriging method, an interpolation method of ArcGIS 10.2. Different land-use types have different influences on the surrounding environment of slopes. The land of the district was classified into three main categories: Agricultural land, construction land, and unused land, based on Article Four of Land Administration Law of China (2004). Agricultural land had the largest proportion (98.68% of the total), which included forest land (92.5%), artificial forest land (2.93%), irrigation land (0.9%), etc., while unused land had the smallest (0.59%). The land use seemed to be uniform throughout the district; therefore, the land use was not selected as the influence factor.

3. Methodology

3.1. Weighted Linear Combination Model

The weighted linear combination model (WLC) is one of the most effective methods for landslide susceptibility mapping. WLC can transform multi-factor evaluation results into comprehensive ones through a linear superposition, according to the importance of each factor. The model is based on the following equation:

$$LSI = \sum_{i=1}^n w_i \cdot u_i \quad (2)$$

where LSI is the landslide susceptibility index, n is the total number of influence factors, w_i represents the weight of factor i , and u_i represents the contribution index of factor i .

The procedure of susceptibility mapping based on WLC can be divided into three steps: data processing, weighting, and mapping and model comparison (Figure 3).

3.2. Classification

The classification of influence factors refers to a process by which a single factor is classified into secondary conditions. In this paper, lithology and aquifer storage capacity have been already classified as they were directly extracted from the original thematic maps; K-means clustering and binarization were introduced for data processing of the remaining selected factors.

3.2.1. K-Means Clustering

K-means clustering is a dynamic clustering method based on distance measurements, which is widely used in the fields of data exploration [36,43,44]. K-means clustering takes the Euclidean distance as the similarity criterion and calculates the distance from sample x_i to each centroid, u_k , for a given database, $X = \{x_1, x_2, \dots, x_n\}$, and a given number (k) of clusters $C = \{c_1, c_2, \dots, c_k\}$:

$$J(c_k) = \sum_{x_i \in c_k} \|x_i - \mu_k\|^2 \quad (3)$$

K-means uses an iterative algorithm to minimize the sum of point-to-centroid distances and sums over all k clusters:

$$J(C) = \sum_{k=1}^k J(c_k) \quad (4)$$

To verify the accuracy of the classification, the Tabular Accuracy Index (TAI) [45] was employed. TAI refers to the ratio between the classification error and the “potential error” inherent to the unclassified data, which is a standardized measure of the data classification [46]:

$$TAI = 1 - \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} |a_{ij} - \bar{a}_j|}{\sum_{i=1}^N |a_i - \bar{a}|} \quad (5)$$

where N is the total number of samples, k is the number of clusters, \bar{a} is the average of all the samples, \bar{a}_j is the average of class j , and n_j is the total number of classes j . The TAI values range from 0 to 1, and the higher the TAI value is, the greater the accuracy of the classification.

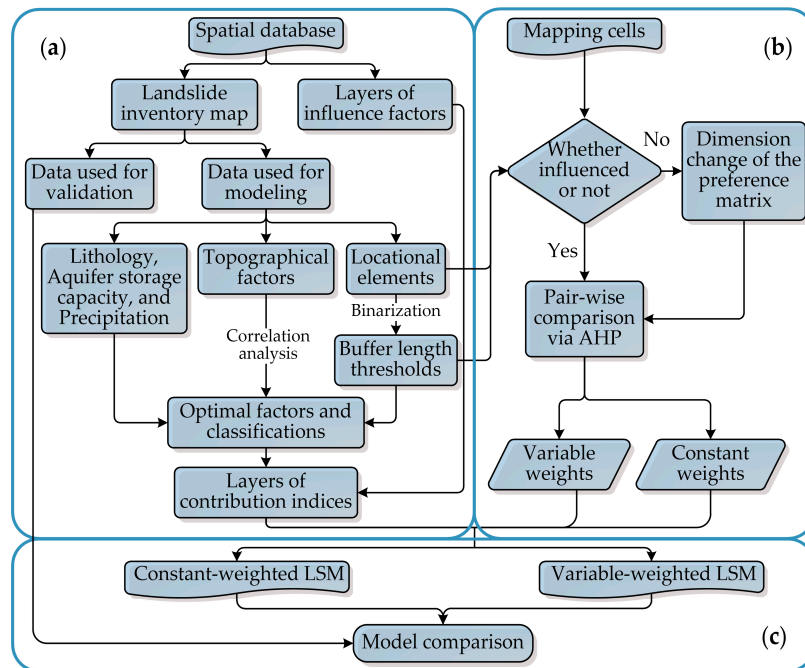


Figure 3. Procedural flow chart adopted in this paper: (a) data processing, which includes classification, correlation analysis, and the determination of the contribution index; (b) weighting according to the different locations of mapping cells based on the improved AHP; and (c) landslide susceptibility mapping and model comparison.

3.2.2. Binarization

The distance data between the locational elements and the existing landslides are the basis for determining buffer length thresholds. However, the extreme data of distance, which reveals that a landslide is located far away from the elements, may create a larger threshold, leading to an increasing cost in the prevention and management of landslides. To eliminate these extreme data, binarization, commonly-used in field of image segmentation [47–49], was introduced:

$$u_i = \begin{cases} 0 & , \quad d_i \geq L \\ 1 & , \quad d_i < L \end{cases} \quad (6)$$

where u_i represents the binary value, d represents the distance between the locational elements and existing landslide i , and L represents the maximum threshold.

Assuming that $L^{(i)}$ (i is the number of the iterations) divides the existing landslides (C) into a part that is influenced by the locational elements ($c_1^{(i)}$) and a part that is not ($c_0^{(i)}$), the sum of point-to-centroid distances was calculated based on K-means clustering:

$$J(C) = \sum_{k=1}^2 J(c_k) = \sum_{c \in c_1^{(i)}} \|d_c - u_1^{(i)}\|^2 + \sum_{c \in c_0^{(i)}} \|d_c - u_2^{(i)}\|^2 \quad (7)$$

where $u_1^{(i)}$ is the centroid of part $c_1^{(i)}$; $u_0^{(i)}$ is the centroid of part $c_0^{(i)}$.

A minimum value of $J(C)$ identified the final centroid:

$$[u_1, u_0] = \arg \min J(C) \quad (8)$$

The average of the centroids $((u_1 + u_0)/2)$ was taken as the maximum buffer length threshold L . Then, the remaining distance data were classified by K-means to obtain other thresholds.

3.3. The Contribution Index

Instead of the commonly-used expert scores, landslide density u_{num} , and landslide area density u_{area} were adopted as contribution indices of influence factors to reduce subjectivities [50]:

$$(u_{num})_{ij} = N_{ij} / \sum_{j=1}^m N_{ij} \quad (9)$$

$$(u_{area})_{ij} = s_{ij} / \sum_{j=1}^m s_{ij} \quad (10)$$

where N_{ij} is the number of landslides in class j of factor i , and is analogous to s_{ij} .

The larger coefficient of variation (CV), the more diversified the data and the more relevant the correlation, therefore CV was used to compare the dispersion degrees between the two indices for the suitable index selection [51]. Additionally, the unit area index (UAI) was introduced to characterize the average area of landslides in a certain class:

$$(UAI)_{ij} = (u_{area})_{ij} / (u_{num})_{ij} \quad (11)$$

3.4. Variable-Weighted Method

AHP involves pair-wise comparisons of decision factors and provides a relative dominance value, ranging from one to nine, to calibrate the qualitative and quantitative performances of priorities [52]. The consistency index is used to improve the consistency of judgments:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (12)$$

where n is the number of factors and λ_{\max} is the maximum preference matrix eigenvalue.

The consistency ratio is used to verify whether the preference matrix is randomly generated:

$$CR = \frac{CI}{RI} \quad (13)$$

where RI is the random consistency index. A CR value of less than 0.1 means a satisfactory consistency.

However, the influences of the factors on landslides differ with the locations of the mapping cells. For this reason, the variable-weighted method was established to determine the weight of each

factor according to the locations of the mapping cells. The preference matrix can be established as the following equation:

$$(a_{ij})_{(n-k) \times (n-k)} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,m-k} & \cdots & a_{1,n-k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m-k,1} & \cdots & a_{m-k,m-k} & \cdots & a_{m-k,n-k} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n-k,1} & \cdots & a_{n-k,m-k} & \cdots & a_{n-k,n-k} \end{pmatrix} \quad (14)$$

where n is the number of factors; k is the number of locational factors, $n < k$; $m = 1, 2, \dots, n$; $(a_{ij})_{N \times N}$ is the result of pair-wise comparison, and $a_{ij} = 1/a_{ji}$.

If the mapping cell is located in a region where k locational factors are negligible, the combination of factors changes, leading to a dimension change of the preference matrix. It can be seen that the number of combination changes is $\sum_{i=0}^n C_n^i$.

4. Results

4.1. The Classification of Elevation, Gradient and Precipitation Using K-Means

Using K-means clustering, elevation, gradient, and precipitation were classified. As shown in Table 3, $TAIs$ using K-means are generally higher than Jenks (an internal algorithm of ArcGIS 10.2), especially the TAI of precipitation using K-means, which is 0.1 higher than the TAI using Jenks. Hence, the input layers of elevation, gradient, and precipitation were produced using the relevant class intervals in ArcGIS 10.2. (Figure 4a–c).

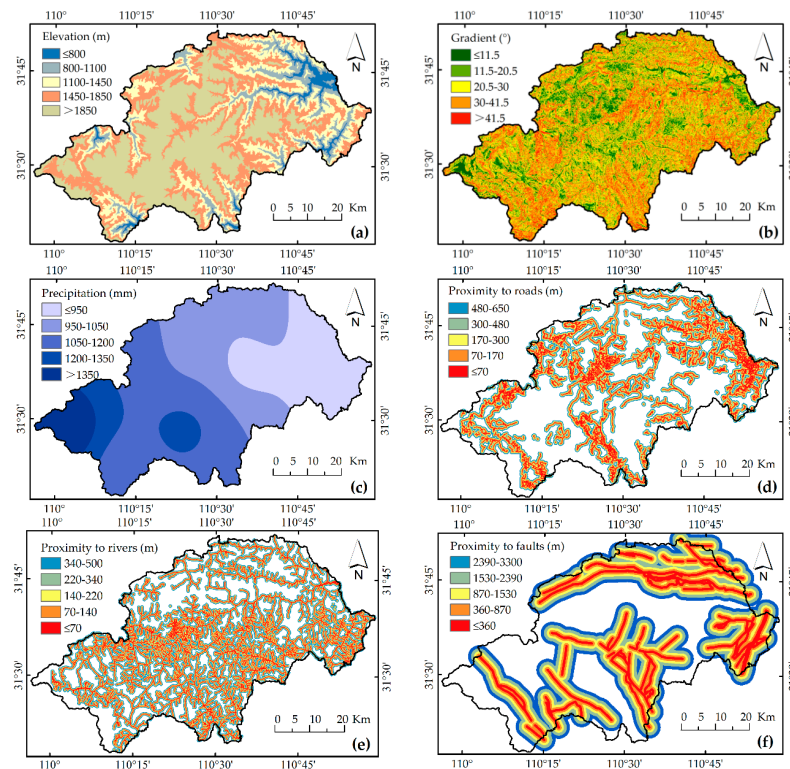


Figure 4. Input layers of influence factors: (a) elevation; (b) gradient; (c) precipitation; (d) proximity to roads; (e) proximity to rivers; and (f) proximity to faults (produced using ArcGIS 10.2).

4.2. The Building of Buffers by Binarization

Using binarization, the buffer length thresholds of locational elements were determined. The computation shows that the data of 13 landslides were eliminated using binarization, leading to a sharp decrease in the maximum threshold, from 19,650 m to 650 m, and also showed an increase in *TAI* from 0.78 to 0.81 (Table 4). The situations are the same for rivers and faults. Subsequently, the buffers were built with the relevant thresholds to consider the influence of locational factors. The obtained vector layers were converted into raster ones, with a cell size of 10 m, to be compatible with the spatial resolution (Figure 4e–f).

Table 3. Classification of elevation, gradient, and precipitation.

Factors	Algorithms	Class Intervals					TAI
		I	II	III	IV	V	
Elevation (m)	K-means	≤800	800–1100	1100–1450	1450–1850	>1850	0.78
	Jenks	≤1100	1100–1500	1500–1850	1850–2250	>2250	0.78
Gradient (°)	K-means	≤11.5	11.5–20.5	20.5–30	30–41.5	>41.5	0.73
	Jenks	≤14.5	14.5–24	24–32.5	32.5–42.5	>42.5	0.71
Precipitation (mm)	K-means	≤950	950–1050	1050–1200	1200–1350	>1350	0.79
	Jenks	≤950	950–1050	1050–1300	1200–1500	>1500	0.69

Table 4. Classification of locational factors.

Factors	Algorithms	Data Size	Class Intervals					TAI
			I	II	III	IV	V	
Proximity to roads (m)	Jenks	127	≤150	150–400	400–700	700–1150	1150–19,650	0.78
	Binarization	114/127	≤70	70–170	170–300	300–480	480–650	0.81
Proximity to rivers (m)	Jenks	127	≤110	110–250	250–500	500–1000	1000–16,700	0.77
	Binarization	104/127	≤70	70–140	140–220	220–340	340–500	0.77
Proximity to faults (m)	Jenks	127	≤850	850–2050	2050–3500	3500–5350	5350–76,300	0.83
	Binarization	95/127	≤360	360–870	870–1530	1530–2390	2390–3300	0.82

4.3. Contrastive Analysis among the Contribution Indices

Landslide density (u_{num}), and landslide area density (u_{area}) of each factor were calculated using Equations (9) and (10). As shown in Figure 5a, u_{num} is evenly distributed along the elevations, and the greatest u_{area} (59.34%) corresponded to an elevation range of 800–1100 m. The CV of u_{area} is far greater than that of u_{num} , which reveals that the contribution of the elevation range of 800–1100 m can be better-characterized by the greater dispersion of u_{area} . Meanwhile, the distribution of the *UAI* indicated that the unit area of landslides in the elevation range of 800–1100 m was greater than that of other elevations.

Figure 5d shows a different view: A greater gradient is not necessarily more prone to landslide occurrences; the landslides were distributed in the gradient range of 0–41.5°. The distribution of the *UAI* indicated that landslides with a greater area were concentrated in the gradient range of 11.5° to 30°, with the greatest u_{area} being 64.28%. Aquifer storage capacity and precipitation had similar distributions (Figure 5b,e). This is because regions with flat terrain and appropriate humidity are rare in the district, where the human activities are concentrated.

Figure 5c shows that Ordovician (O) rock had the greatest u_{area} (35.25%), and a relatively high u_{num} (16.54%), where landslides have a relatively greater unit area.

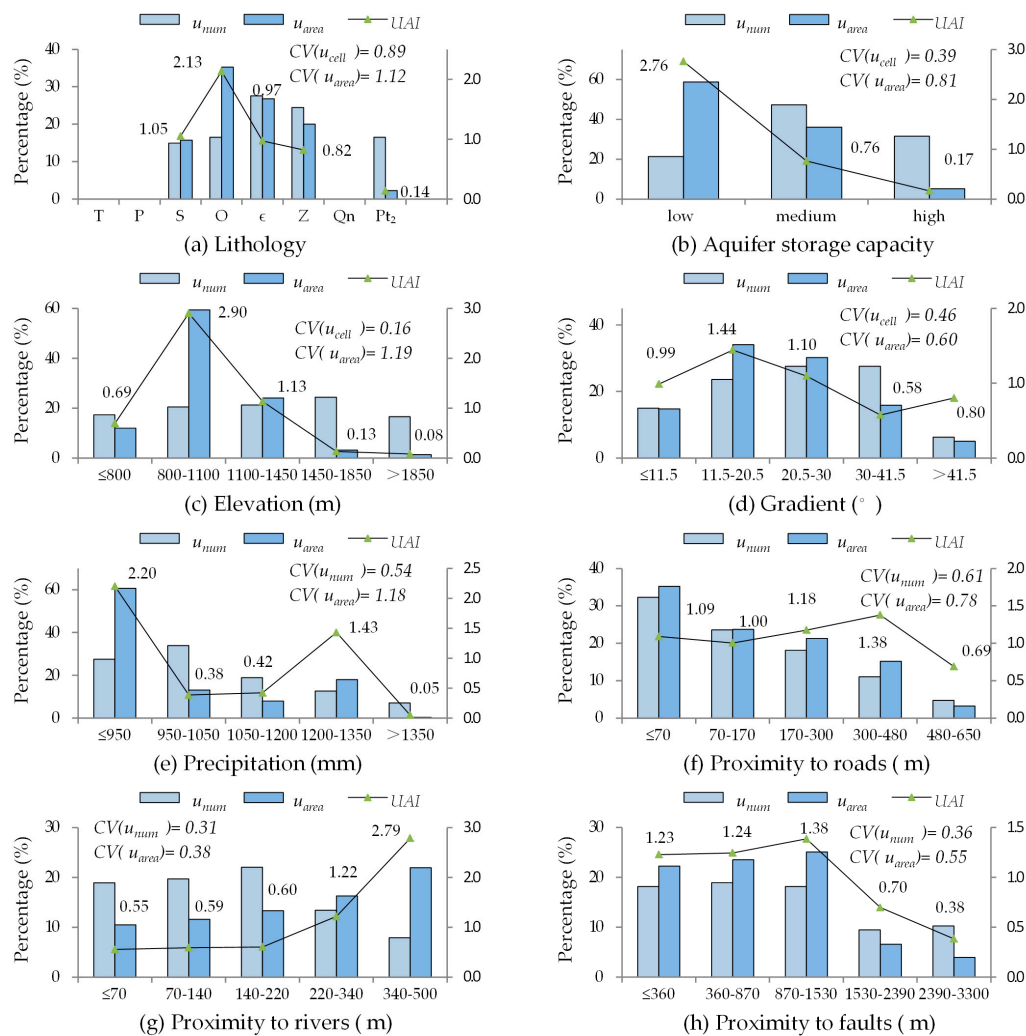


Figure 5. u_{num} and u_{area} of each influence factor: (a) lithology; (b) aquifer storage capacity; (c) elevation; (d) gradient; (e) precipitation; (f) proximity to roads; (g) proximity to rivers; and (h) proximity to faults. CV is the coefficient of variation; UAI is the unit area index.

As shown in Figure 5f, in the regions influenced by roads, with increasing distance from a road, the number of landslides decreases, while area increases. Although the CVs of u_{num} and u_{area} show little difference, both the indices showed that landslides are prone to occur in regions near roads. Figure 5g shows that the u_{num} of proximity to rivers changes minimally within 0–220 m and then slowly decreases, while the u_{area} continually increases with the increase in distance to a river. Regarding faults, the u_{num} and u_{area} change minimally within the distance range of 0–1530 m and then slowly decrease (Figure 5h).

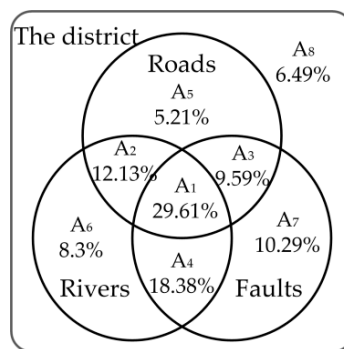
It can be seen above that landslide area density u_{area} can characterize the contribution of a certain class for having the greater CVs. Therefore, landslide area density u_{area} is more efficient for reflecting the connections between influence factors and landslides, and it was taken as the contribution index (Table 5).

Table 5. Contribution indices of influence factors.

Factors	Contribution Indices							
Elevation (m)	≤ 800 0.12		800–1100 0.593		1100–1450 0.241		1450–1850 0.032	> 1850 0.013
Gradient ($^{\circ}$)	≤ 11.5 0.148		11.5–20.5 0.341		20.5–30 0.302		30–41.5 0.159	> 41.5 0.051
Lithology	T 0.352	P 0	S 0.268	O 0	ϵ 0	Z 0.2	Q_n 0.157	Pt_2 0.022
Aquifer storage capacity		Low 0.587			Medium 0.361			High 0.052
Precipitation (mm)	≤ 950 0.607		950–1050 0.13		1050–1200 0.08		1200–1350 0.18	> 1350 0.003
Proximity to roads (m)	≤ 70 0.357		70–170 0.24		170–300 0.216		300–480 0.154	480–650 0.033
Proximity to rivers (m)	≤ 70 0.142		70–140 0.158		140–220 0.181		220–340 0.221	340–500 0.298
Proximity to faults (m)	≤ 360 0.273		360–870 0.289		870–1530 0.308		1530–2390 0.081	2390–3300 0.048

4.4. Weighting Using the Variable-Weighted Method

As shown in Figure 6, the district was divided into eight different regions, according to whether the mapping cells were influenced by single or multiple locational elements. The regions where the traditional constant weights are applicable account for only 29.61% of the district. The weights of each factor were calculated using the variable-Weighted Method (Table 6).

**Figure 6.** Regions divided by the locational elements.**Table 6.** Weights of each factor in different regions.

Factors	Weights/Regions							
	w_1/A_1	w_2/A_2	w_3/A_3	w_4/A_4	w_5/A_5	w_6/A_6	w_7/A_7	w_8/A_8
Elevation	0.149	0.161	0.165	0.179	0.181	0.196	0.204	0.227
Gradient	0.271	0.287	0.296	0.357	0.314	0.392	0.407	0.454
Precipitation	0.052	0.055	0.057	0.068	0.059	0.074	0.077	0.085
Aquifer storage capacity	0.077	0.087	0.09	0.1	0.105	0.118	0.123	0.151
Lithology	0.048	0.051	0.052	0.063	0.056	0.07	0.073	0.082
Proximity to roads	0.213	0.24	0.247	-	0.285	-	-	-
Proximity to rivers	0.109	0.119	-	0.133	-	0.15	-	-
Proximity to faults	0.081	-	0.092	0.1	-	-	0.117	-

4.5. Landslide Susceptibility Mapping

For landslide susceptibility mapping, the landslide susceptibility index was classified into five categories using K-means clustering: Very low (I), low (II), medium (III), high (IV), and very high (V). Subsequently, two landslide susceptibility maps were generated: (i) using variable weights (LSM1, Figure 7a); and (ii) using constant weights (LSM2, Figure 7b).

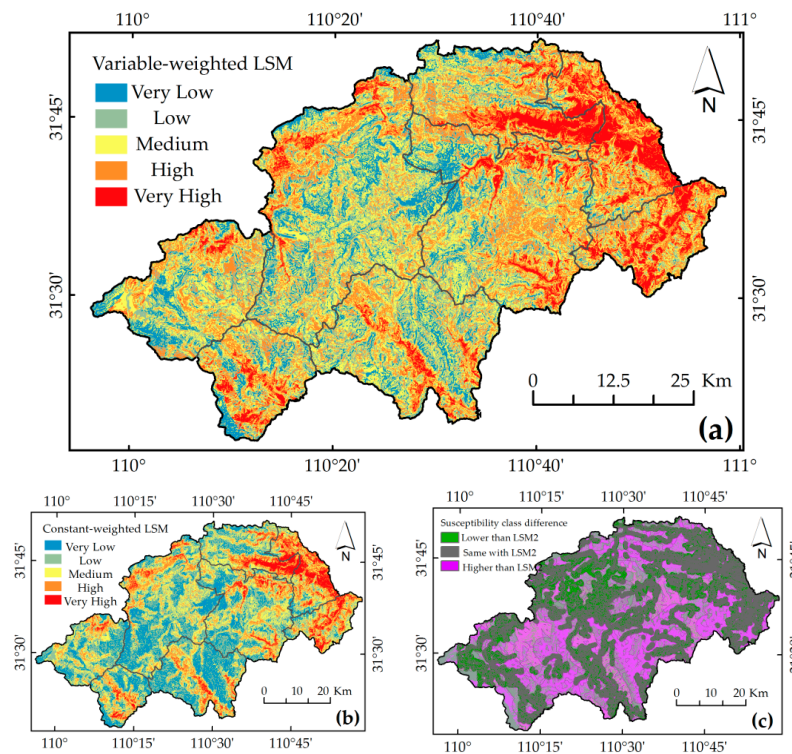


Figure 7. Landslide susceptibility maps: (a) LSM1 (using variable weights); and (b) LSM2 (using constant weights); and (c) susceptibility class difference (produced using ArcGIS 10.2).

5. Discussion

There are two common features between the two different susceptibility maps: high and very high susceptible regions are located mainly in the northeast, where the gradient is gradual and the elevation is relatively low; and the high and very high susceptible regions are mainly distributed along roads, where the slope cuts and overloads may accelerate the instability of slopes.

It can be observed in Figure 7a,b that more landslides occur in regions with the high and very high susceptibility class in LSM1: 59% of total number and 95.5% of the total area of the existing landslides are located in regions with high and very high susceptibility classes in LSM1, while 49.6% of total number and 93.5% of the total area of existing landslides are located in regions with high and very high susceptibility classes in LSM2.

As shown in Figure 7c, 62.3% of the mapping cells have the same susceptibility class and are located in region A₁ where the influences of the three locational factors are considered. Meanwhile, 27% of the mapping cells have one susceptibility class higher than that of using the constant weights, and 0.56% of the mapping cells have two susceptibility classes higher than that of using the constant weights. Cells with higher susceptibility classes are mainly located in regions A₄, A₆, A₇, and A₈, where the influence of the roads is negligible. This is a fairly good result, which illustrates that the mapping cells may have high landslide susceptibility, even in the regions where the influence of the locational factor on landslides is negligible.

Hence, it is concluded that the landslide susceptibility map using the VWLC model has a higher accuracy and a more reasonable distribution of landslides.

6. Conclusions

Urban and rural planning is complex work that needs to take into account social, economic, and ecological factors. As for landslide-vulnerable regions, landslide susceptibility needs to be considered as well. An LSM could contribute to urban and rural planning, such as the movement and merger of villages, which can be planned by combining LSM with population distribution; based on LSM, in conjunction with the locations of infrastructure, property risk can be assessed.

In this paper, the variable-weighted linear combination model was established for the susceptibility mapping of shallow landslides, which mainly included the introduction of binarization for determining buffer length thresholds of the locational elements, and the establishment of the variable-weighted method for weighting, which takes different locations of mapping cells into consideration.

The introduction of binarization not only avoids the subjectivity brought by empirical judgments, but also reduces the effect of extreme values on buffer length thresholds. The maximum threshold of the proximity to roads in the district was 650 m, the maximum threshold of the proximity to rivers was 500 m, and the maximum threshold of the proximity to faults was 3300 m. The variable-weighted method considers the dimensional changes of the preference matrix varying with different locations of the mapping cells, and rationally evaluates susceptibility in regions where the influence of single or multiple locational factors is negligible.

Using the VWLC model and considering eight influence factors, elevation, gradient, precipitation, aquifer storage capacity, lithology, proximity to roads, proximity to rivers, and proximity to faults, a landslide susceptibility map for the Shennongjia Forestry District was produced. Compared with a susceptibility map using constant weights, the map based on VWLC was validated to have a higher accuracy and a more reasonable distribution of landslides.

Therefore, the VWLC model is more beneficial to urban and rural planning, and can be used for any location that has the same, or similar, geological and topographical conditions.

Acknowledgments: The study is funded by National Science and Technology Support Program of China (No. 2014BAL05B07).

Author Contributions: The developed model was conceived and built by Wei Chen and Xudong Fu; programming was done by Bin Huang and Qile Huang; the manuscript was written by Wei Chen; and the manuscript was revised by Hongxing Han and Xudong Fu.

Conflicts of Interest: The authors declare no competing financial interests.

References

1. Bulletin of Land Resources of China. 2016. Available online: <http://data.mlr.gov.cn/> (accessed on 28 April 2017).
2. Bulletin of Nationwide Geologic Hazards of China. 2016. Available online: <http://www.cigem.gov.cn/> (accessed on 22 February 2017).
3. Bathrellos, G.D.; Skilodimou, H.D.; Chousianitis, K.; Youssef, A.M.; Pradhan, B. Suitability estimation for urban development using multi-hazard assessment map. *Sci. Total Environ.* **2017**, *575*, 119–134. [[CrossRef](#)] [[PubMed](#)]
4. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65*, 15–31. [[CrossRef](#)]
5. Fall, M.; Azzam, R.; Noubactep, C. A multi-method approach to study the stability of natural slopes and landslide susceptibility mapping. *Eng. Geol.* **2006**, *82*, 241–263. [[CrossRef](#)]
6. Pastor, M.; Blanc, T.; Haddad, B.; Petrone, S.; Morles, M.S.; Drempetic, V.; Issler, D.; Crosta, G.B.; Cascini, L.; Sorbino, G.; et al. Application of a SPH depth-integrated model to landslide run-out analysis. *Landslides* **2014**, *11*, 793–812. [[CrossRef](#)]

7. Chalkias, C.; Ferentinou, M.; Polykretis, C. GIS Supported Landslide Susceptibility Modeling at Regional Scale: An Expert-Based Fuzzy Weighting Method. *ISPRS Int. J. Geo-Inf.* **2014**, *3*, 523–539. [[CrossRef](#)]
8. Canavesi, V.; Alvares, R.C.D. Changes in Vegetation Cover in Reforested Areas in the State of Sao Paulo, Brazil and the Implication for Landslide Processes. *ISPRS Int. J. Geo-Inf.* **2012**, *1*, 209–227. [[CrossRef](#)]
9. Lin, Y.L.; Xia, K.W.; Jiang, X.; Bai, J.; Wu, P. Landslide Susceptibility Mapping Based on Particle Swarm Optimization of Multiple Kernel Relevance Vector Machines: Case of a Low Hill Area in Sichuan Province, China. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 191. [[CrossRef](#)]
10. Ayalew, L.; Yamagishi, H.; Ugawa, N. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. *Landslides* **2004**, *1*, 73–81. [[CrossRef](#)]
11. Lyu, M.R.; Nikora, A. A heuristic approach for software reliability prediction: The equally-weighted linear combination model. In Proceedings of the 1991 International Symposium on Software Reliability Engineering, Austin, TX, USA, 17–18 May 1991; IEEE: New York, NY, USA, 1991; pp. 172–181. [[CrossRef](#)]
12. Pradhan, B. Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia. *Adv. Space Res.* **2010**, *45*, 1244–1256. [[CrossRef](#)]
13. Ghorbani Nejad, S.; Falah, F.; Daneshfar, M.; Haghizadeh, A.; Rahmati, O. Delineation of groundwater potential zones using remote sensing and GIS-based data-driven models. *Geocarto Int.* **2016**, *32*, 1–21. [[CrossRef](#)]
14. Ahmed, B.; Dewan, A. Application of Bivariate and Multivariate Statistical Techniques in Landslide Susceptibility Modeling in Chittagong City Corporation, Bangladesh. *Remote Sens.* **2017**, *9*, 304. [[CrossRef](#)]
15. Feng, L.; Zhu, X.; Sun, X. Assessing coastal reclamation suitability based on a fuzzy-AHP comprehensive evaluation framework: A case study of Lianyungang, China. *Mar. Pollut. Bull.* **2014**, *89*, 102–111. [[CrossRef](#)] [[PubMed](#)]
16. Park, H.-J.; Jang, J.-Y.; Lee, J.-H. Physically Based Susceptibility Assessment of Rainfall-Induced Shallow Landslides Using a Fuzzy Point Estimate Method. *Remote Sens.* **2017**, *9*, 487. [[CrossRef](#)]
17. Yesilnacar, E.; Topal, T. Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng. Geol.* **2005**, *79*, 251–266. [[CrossRef](#)]
18. Yilmaz, I. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environ. Earth Sci.* **2009**, *61*, 821–836. [[CrossRef](#)]
19. Kanungo, D.P.; Arora, M.K.; Sarkar, S.; Gupta, R.P. A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Eng. Geol.* **2006**, *85*, 347–366. [[CrossRef](#)]
20. Wang, J.; Yin, K.; Xiao, L. Landslide susceptibility assessment based on GIS and weighted information value: a case study of Wanzhou district, Three gorges reservoir. *Chin. J. Rock Mech. Eng.* **2014**, *33*, 797–808.
21. Zhang, J.; Yin, K.; Wang, J.; Liu, L.; Huang, F. Evaluation of landslide susceptibility for Wanzhou district of Three Gorges Reservoir. *Chin. J. Rock Mech. Eng.* **2016**, *35*, 284–296. [[CrossRef](#)]
22. Bathrellos, G.D.; Gaki-Papanastassiou, K.; Skilodimou, H.D.; Papanastassiou, D.; Chousianitis, K.G. Potential suitability for urban planning and industry development using natural hazard maps and geological–geomorphological parameters. *Environ. Earth Sci.* **2011**, *66*, 537–548. [[CrossRef](#)]
23. Pourghasemi, H.R.; Mohammady, M.; Pradhan, B. Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. *Catena* **2012**, *97*, 71–84. [[CrossRef](#)]
24. Kayastha, P.; Dhital, M.R.; De Smedt, F. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: A case study from the Tinau watershed, west Nepal. *Comput. Geosci.-UK* **2013**, *52*, 398–408. [[CrossRef](#)]
25. Ramanathan, R. A note on the use of the analytic hierarchy process for environmental impact assessment. *J. Environ. Manag.* **2001**, *63*, 27–35. [[CrossRef](#)] [[PubMed](#)]
26. Bathrellos, G.D.; Karymbalis, E.; Skilodimou, H.D.; Gaki-Papanastassiou, K.; Baltas, E.A. Urban flood hazard assessment in the basin of Athens Metropolitan city, Greece. *Environ. Earth Sci.* **2016**, *75*. [[CrossRef](#)]
27. Ercanoglu, M.; Gokceoglu, C.; Van Asch, Th.W.J. Landslide susceptibility zoning north of Yenice (NW Turkey) by multivariate statistical techniques. *Nat. Hazards* **2004**, *32*, 1–23. [[CrossRef](#)]

28. Mohammady, M.; Pourghasemi, H.R.; Pardhan, B. Landslide susceptibility mapping at Golestan Province, Iran: A comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models. *J. Asian Earth Sci.* **2012**, *61*, 221–236. [[CrossRef](#)]
29. Galli, M.; Ardizzone, F.; Cardinali, M.; Guzzetti, F.; Reichenbach, P. Comparing landslide inventory maps. *Geomorphology* **2008**, *94*, 268–289. [[CrossRef](#)]
30. Martha, T.R.; van Westen, C.J.; Kerle, N.; Jetten, V.; Kumar, K.V. Landslide hazard and risk assessment using semi-automatically created landslide inventories. *Geomorphology* **2013**, *184*, 139–150. [[CrossRef](#)]
31. Arcgis 10.2. Environmental Systems Research Institute, Inc. Copyright 2007. Available online: <http://desktop.arcgis.com> (accessed on 21 June 2017).
32. Pourghasemi, H.R.; Pradhan, B.; Gokceoglu, C. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Nat. Hazards* **2012**, *63*, 965–996. [[CrossRef](#)]
33. Ayalew, L.; Yamagishi, H. Slope movements in the Blue Nile basin, as seen from landscape evolution perspective. *Geomorphology* **2004**, *57*, 95–116. [[CrossRef](#)]
34. Das, I.; Stein, A.; Kerle, N.; Dadhwal, V.K. Landslide susceptibility mapping along road corridors in the Indian Himalayas using Bayesian logistic regression models. *Geomorphology* **2012**, *179*, 116–125. [[CrossRef](#)]
35. Yalcin, A.; Bulut, F. Landslide susceptibility mapping using GIS and digital photogrammetric techniques: A case study from Ardesen (NE-Turkey). *Nat. Hazards* **2006**, *41*, 201–226. [[CrossRef](#)]
36. Zhang, J.-Y.; Xu, W.; Zhang, W.; Meng, X.; Zhang, Y. A novel compression algorithm for infrared thermal image sequence based on K-means method. *Infrared Phys. Technol.* **2014**, *64*, 18–25. [[CrossRef](#)]
37. Hong, H.; Chen, W.; Xu, C.; Youssef, A.M.; Pradhan, B.; Bui, D.T. Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy. *Geocarto Int.* **2016**, *32*, 1–16. [[CrossRef](#)]
38. Xu, K.; Guo, Q.; Li, Z.; Xiao, J.; Qin, Y.; Chen, D. Landslide susceptibility evaluation based on BPNN and GIS: A case of Guojiaba in the Three Gorges Reservoir Area. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 1111–1124. [[CrossRef](#)]
39. Ilia, I.; Tsangaratos, P. Applying weight of evidence method and sensitivity analysis to produce a landslide susceptibility map. *Landslides* **2015**, *13*, 379–397. [[CrossRef](#)]
40. Dai, K.R.; Li, Z.H.; Tomás, R.; Liu, G.; Yu, B.; Wang, X.; Cheng, H.; Chen, J.; Stockamp, J. Monitoring activity at the Daguangbao mega-landslide (China) using Sentinel-1 TOPS time series interferometry. *Remote Sens. Environ.* **2016**, *186*, 501–513. [[CrossRef](#)]
41. Tomás, R.; Li, Z.; Lopez-Sanchez, J.M.; Liu, P.; Singleton, A. Using wavelet tools to analyse seasonal variations from InSAR time-series data: A case study of the Huangtupo landslide. *Landslides* **2015**, *13*, 437–450. [[CrossRef](#)]
42. Tarolli, P.; Borga, M.; Chang, K.-T.; Chiang, S.-H. Modeling shallow landsliding susceptibility by incorporating heavy rainfall statistical properties. *Geomorphology* **2011**, *133*, 199–211. [[CrossRef](#)]
43. Likas, A.; Vlassis, N.; Verbeek, J.J. The global k-means clustering algorithm. *Pattern Recognit.* **2003**, *36*, 451–461. [[CrossRef](#)]
44. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. C-Appl.* **1979**, *28*, 100–108. [[CrossRef](#)]
45. Jenks, G.F.; Caspall, F.C. Error on Choroplethic Maps—Definition, Measurement, Reduction. *Ann. Assoc. Am. Geogr.* **1971**, *61*, 217–244. [[CrossRef](#)]
46. Carrão, H.; Singleton, A.; Naumann, G.; Barbosa, P.; Vogt, J.V. Optimized System for the Classification of Meteorological Drought Intensity with Applications in Drought Frequency Analysis. *J. Appl. Meteorol. Clim.* **2014**, *53*, 1943–1960. [[CrossRef](#)]
47. Farina, A.; Kovacs-Vajna, Z.M.; Leone, A. Fingerprint minutiae extraction from skeletonized binary images. *Pattern Recognit.* **1999**, *32*, 877–889. [[CrossRef](#)]
48. Fernandez, A.; Lopez, V.; Galar, M.; del Jesus, M.J.; Herrera, F. Analysing the classification of imbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches. *Knowl.-Based Syst.* **2013**, *42*, 97–110. [[CrossRef](#)]
49. Gatos, B.; Pratikakis, I.; Perantonis, S.J. Adaptive degraded document image binarization. *Pattern Recognit.* **2006**, *39*, 317–327. [[CrossRef](#)]

50. Tian, H.; Qiao, J.; Wang, M. A method of early warning on rainfall-induced landslide risk probability based on hazard zoning: A case study on the rainfall-induced landslide of Miyi County, Sichuan, China. *Geol. Bull. China* **2009**, *28*, 1093–1097.
51. Guo, Y.J. *Theory and Application of Comprehensive Evaluation*; Science Press: Beijing, China, 2007.
52. Saaty, T.L. A scaling method for priorities in hierarchical structures. *J. Math Psychol.* **1977**, *15*, 234–281. [[CrossRef](#)]



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).