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 handing 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°15′–31°75′ N and longitude 109°56′–110°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.

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

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Scale</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landslide inventory map</td>
<td>Point</td>
<td>Land Resources Bureau of Shennongjia Forestry District</td>
</tr>
<tr>
<td>DEM</td>
<td>Raster</td>
<td>10 m</td>
</tr>
<tr>
<td>Road</td>
<td>Line</td>
<td>School of Urban Design of Wuhan University</td>
</tr>
<tr>
<td>River</td>
<td>Line</td>
<td>Geological map of Hubei province</td>
</tr>
<tr>
<td>Faults</td>
<td>Line</td>
<td>Geological map of Hubei province</td>
</tr>
<tr>
<td>Lithology</td>
<td>Polygon</td>
<td>1:100,000</td>
</tr>
<tr>
<td>Soil moisture content</td>
<td>Polygon</td>
<td>1:200,000</td>
</tr>
<tr>
<td>Land use</td>
<td>Polygon</td>
<td>Hydrologic map of Hubei province</td>
</tr>
<tr>
<td>Precipitation</td>
<td></td>
<td>Meteorological Bureau of Shennongjia Forestry District</td>
</tr>
</tbody>
</table>

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:

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

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 \(x\); \(\bar{y}\) is the sample mean of \(y\); \(s_x\) is the sample standard deviation of \(x\); \(s_y\) is the sample standard deviation of \(y\). 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.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Elevation</th>
<th>Gradient</th>
<th>Aspect</th>
<th>RA</th>
<th>SR</th>
<th>PLC</th>
<th>PRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td>1.0</td>
<td>0.08</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>Gradient</td>
<td>1.0</td>
<td>-0.07</td>
<td>0.97</td>
<td>0.93</td>
<td>0.08</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>Aspect</td>
<td>1.0</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.07</td>
<td>-0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA</td>
<td>1.0</td>
<td>0.96</td>
<td>0.10</td>
<td>-0.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>1.0</td>
<td>0.09</td>
<td>-0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLC</td>
<td>1.0</td>
<td>-0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRC</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 (e), Sinian (Z), Qingbaikouan (Qs), and Middle-Proterozoic (Ptz) (Figure 3c). Additionally, the aquifers have been classified into high, medium, and low, according to their relative storage capacities (Figure 3d).

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:

\[ LSI = \sum_{i=1}^{n} w_i \cdot u_i \]  

(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 2).

Figure 2. 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. 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 already been 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 \( x \) to each centroid, \( w \), for a given database, \( X = \{x_1, x_2, ..., x_n\} \), and a given number \( (k) \) of clusters \( C = \{c_1, c_2, ..., c_k\} \):
\begin{equation}
J(c_i) = \sum_{x \in c_i} ||x - \mu_i||^2
\end{equation}

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

\begin{equation}
J(C) = \sum_{i=1}^{k} J(c_i)
\end{equation}

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]:

\begin{equation}
TAI = 1 - \frac{\sum_{j=1}^{N} \sum_{i \in i} |d_{i} - \bar{d}_{j}|}{\sum_{i=1}^{N} |d_{i} - \bar{d}|}
\end{equation}

where \( N \) is the total number of samples, \( k \) is the number of clusters, \( \bar{d} \) is the average of all the samples, \( \bar{d}_{j} \) is the average of class \( j \), and \( n_{i} \) 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.

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:

\[ u_i = \begin{cases} 0 & d_i \geq L \\ 1 & d_i < L \end{cases} \]

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 \( C^{(i)}_1 \) and a part that is not \( C^{(i)}_0 \), the sum of point-to-centroid distances was calculated based on K-means clustering:

\begin{equation}
J(C) = \sum_{i=1}^{k} J(c_i) = \sum_{c_i \in C^{(i)}_1} \|d_i - u^{(i)}_1\|^2 + \sum_{c_i \in C^{(i)}_0} \|d_i - u^{(i)}_0\|^2
\end{equation}

where \( u^{(i)}_1 \) is the centroid of part \( C^{(i)}_1 \); \( u^{(i)}_0 \) is the centroid of part \( C^{(i)}_0 \).

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

\[ [u^{*}_1, u^{*}_0] = \arg \min J(C) \]

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]:
\[
\left( u_{\text{num}} \right)_j = \frac{N_j}{\sum_{j=1}^{m} N_j} \quad (9)
\]
\[
\left( u_{\text{area}} \right)_j = \frac{s_j}{\sum_{j=1}^{m} s_j} \quad (10)
\]

where \( N_j \) is the number of landslides in class \( j \) of factor \( i \), and is analogous to \( s_j \).

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:

\[
\left( \text{UAI} \right)_j = \left( u_{\text{area}} \right)_j \left( u_{\text{num}} \right)_j \quad (11)
\]

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:

\[
CI = \frac{\lambda_{\text{max}} - n}{n - 1} \quad (12)
\]

where \( n \) is the number of factors and \( \lambda_{\text{max}} \) 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:

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

where \( n \) is the number of factors; \( k \) is the number of locational factors, \( n < k \); \( m = 1, 2, \ldots, n \); \( \left( a_{ij} \right)_{N \times N} \) 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, TAI\(_s\) 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 3a,b,e).

Figure 3. Input layers of influence factors: (a) elevation; (b) gradient; (c) lithology; (d) aquifer storage capacity; (e) precipitation; (f) proximity to roads; (g) proximity to rivers; and (h) 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 3f–h).

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

<table>
<thead>
<tr>
<th>Factors</th>
<th>Algorithms</th>
<th>Class Intervals</th>
<th>TAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation (m)</td>
<td>K-means</td>
<td>≤800 800–1100</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Jenks</td>
<td>1100–1500 1500–1850</td>
<td>0.78</td>
</tr>
<tr>
<td>Gradient (°)</td>
<td>K-means</td>
<td>≤11.5 11.5–20.5 20.5–30</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Jenks</td>
<td>≤14.5 14.5–24 24–32.5</td>
<td>0.71</td>
</tr>
<tr>
<td>Precipitation (mm)</td>
<td>Jenks</td>
<td>≤950 950–1050 1050–1200 1200–1350</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Table 4. Classification of locational factors.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Algorithms</th>
<th>Data Size</th>
<th>Class Intervals</th>
<th>TAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximity to roads (m)</td>
<td>Jenks</td>
<td>127</td>
<td>≤150 150–400</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Binarization</td>
<td>114/127</td>
<td>≤70 70–170</td>
<td>0.81</td>
</tr>
<tr>
<td>Proximity to rivers (m)</td>
<td>Jenks</td>
<td>127</td>
<td>≤110 110–250 250–500</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Binarization</td>
<td>104/127</td>
<td>≤70 70–140</td>
<td>0.77</td>
</tr>
<tr>
<td>Proximity to faults (m)</td>
<td>Jenks</td>
<td>127</td>
<td>≤850 850–2050 2050–3500 3500–5350</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>Binarization</td>
<td>95/127</td>
<td>≤360 360–870</td>
<td>0.82</td>
</tr>
</tbody>
</table>

4.3. Contrastive Analysis among the Contribution Indices

Landslide density ($t_{mean}$), and landslide area density ($t_{area}$) of each factor were calculated using Equations (9) and (10). As shown in Figure 4a, $t_{mean}$ is evenly distributed along the elevations, and the greatest $t_{area}$ (59.34%) corresponded to an elevation range of 800–1100 m. The CV of $t_{area}$ is far greater than that of $t_{mean}$, which reveals that the contribution of the elevation range of 800–1100 m can be better-characterized by the greater dispersion of $t_{area}$. 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 4b 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 $t_{area}$ being 64.28%. Aquifer storage capacity and precipitation had similar distributions (Figure 4d–e). This is because regions with flat terrain and appropriate humidity are rare in the district, where the human activities are concentrated.

Figure 4c shows that Ordovician (O) rock had the greatest $t_{area}$ (35.25%), and a relatively high $t_{mean}$ (16.54%), where landslides have a relatively greater unit area.
Figure 4. $u_{num}$ and $u_{area}$ of each influence factor: (a) elevation; (b) gradient; (c) lithology; (d) aquifer storage capacity; (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 4f, 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 4g 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 4h).

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.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Contribution Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation (m)</td>
<td>≤800 800–1100 1100–1450 1450–1850 &gt;1850</td>
</tr>
<tr>
<td></td>
<td>0.12 0.593 0.241 0.032 0.013</td>
</tr>
<tr>
<td>Gradient (°)</td>
<td>≤11.5 11.5–20.5 20.5–30 30–41.5 &gt;41.5</td>
</tr>
<tr>
<td></td>
<td>0.148 0.341 0.302 0.159 0.051</td>
</tr>
<tr>
<td>Lithology (°)</td>
<td>T P S O ε Z Q n Pt</td>
</tr>
<tr>
<td></td>
<td>0.352 0.0 0.268 0 0 0.2 0.157 0.022</td>
</tr>
<tr>
<td>Aquifer storage capacity</td>
<td>Low Medium High</td>
</tr>
<tr>
<td></td>
<td>0.587 0.361 0.052</td>
</tr>
<tr>
<td>Precipitation (mm)</td>
<td>≤950 950–1050 1050–1200 1200–1350 &gt;1350</td>
</tr>
<tr>
<td></td>
<td>0.607 0.13 0.08 0.18 0.003</td>
</tr>
<tr>
<td>Proximity to roads (m)</td>
<td>≤70 70–170 170–300 300–480 480–650</td>
</tr>
<tr>
<td></td>
<td>0.357 0.24 0.216 0.154 0.033</td>
</tr>
<tr>
<td>Proximity to rivers (m)</td>
<td>≤70 70–140 140–220 220–340 340–500</td>
</tr>
<tr>
<td></td>
<td>0.142 0.158 0.181 0.221 0.298</td>
</tr>
<tr>
<td>Proximity to faults (m)</td>
<td>≤360 360–870 870–1530 1530–2390 2390–3300</td>
</tr>
<tr>
<td></td>
<td>0.273 0.289 0.308 0.081 0.048</td>
</tr>
</tbody>
</table>

4.4. Weighting Using the Variable-Weighted Method

As shown in Figure 5, 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 5. Regions divided by the locational elements.

Table 6. Weights of each factor in different regions.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Weights/Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevation</td>
<td>w/A1 w/A2 w/A3 w/A4 w/A5 w/A6 w/A7 w/A8</td>
</tr>
<tr>
<td>Gradient</td>
<td>0.149 0.161 0.165 0.179 0.181 0.196 0.204 0.227</td>
</tr>
<tr>
<td>Precipitation</td>
<td>0.271 0.287 0.296 0.357 0.314 0.392 0.407 0.454</td>
</tr>
<tr>
<td>Aquifer storage capacity</td>
<td>0.052 0.055 0.057 0.068 0.059 0.074 0.077 0.085</td>
</tr>
<tr>
<td>Lithology</td>
<td>0.077 0.087 0.09 0.1 0.105 0.118 0.123 0.151</td>
</tr>
<tr>
<td>Proximity to roads</td>
<td>0.048 0.051 0.052 0.063 0.056 0.07 0.073 0.082</td>
</tr>
<tr>
<td>Proximity to rivers</td>
<td>0.213 0.24 0.247 - 0.285 - - -</td>
</tr>
<tr>
<td>Proximity to faults</td>
<td>0.109 0.119 - 0.133 - 0.15 - -</td>
</tr>
<tr>
<td></td>
<td>0.081 - 0.092 0.1 - - 0.117 -</td>
</tr>
</tbody>
</table>
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 6a); and (ii) using constant weights (LSM2, Figure 6b).

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.

Figure 6. Landslide susceptibility maps: (a) LSM1 (using variable weights); and (b) LSM2 (using constant weights); and (c) susceptibility class difference (produced using ArcGIS 10.2).

It can be observed in Figure 6a,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 6c, 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


© 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/).