Spatial Prediction of Earthquake-Induced Secondary Landslide Disaster in Beichuan County Based on GIS

Zhuowei Hu, Lai Wei, Dan Fang, Te Lai, Qing Wang

1College of Resources Environment and Tourism, Capital Normal University,
2Kay Lab of Resources Environment and GIS,
3Key Laboratory of Integrated Disaster Assessment and Risk Governance of the Ministry of Civil Affairs, Beijing 100048, China
4Graduate School of Education,
5Department of Geography, SUNY University at Buffalo, Buffalo 14221, USA

Abstract: In earthquake-stricken area, with the occurrence of aftershocks, heavy rainfall and human activity, the earthquake-induced secondary landslide disaster will threaten people’s life and property in a very long period. So, it makes secondary landslide became a research hotspots that draw much attention. The forecasting of natural disaster is considered as a most effective way to prevention or mitigation disaster and the spatial prediction is the base work of landslide disaster research. The aim of this study is to analyze the landslide prediction, taking the case of Beichuan County. Six factors affecting landslide occurrence have been taken into account, including elevation, slope, lithology, seismic intensity, distance to roads and rivers. The correlations of landslide distribution with these factors is calculated, the multiple regression and neural network model are applied to landslide spatial prediction and mapping. The model calculates result is ultimately categorized into four classes. It shows that the high and very high susceptibility areas most distribute in Qushan, Chenjiaba towns, etc., along the rivers and the roads around the area of Longmenshan fault. The precision accuracy using multiple regression models is about 73.7% and the neural network model can be up to 81.28%. It can be concluded that in this study area, the neural network model appears to be more accurate in landslide spatial prediction.

Keywords: Beichuan, earthquake, GIS, landslide, spatial prediction

INTRODUCTION

China is a large agricultural country and most of China's population live in rural areas. Because the development of rural logistics lags behind, a large number of agricultural products can’t smoothly to enter the market (Li, 2003). This not only makes city residents not obtain adequate supply of agricultural products, greatly reduce the confidence of farmers to produce, but also further exacerbates the contradiction between supply and demand in cities and rural areas and hinders harmonious economic development of China’s society (Huang, 2000). At present, Chinese government is actively promoting new rural construction project, whose goal is continually to develop rural economy for improving farmers’ livelihood, while the construction of rural logistics distribution center has become one of the substantive contents (Guo et al., 2010).

The earthquake landslide is the most common form of secondary geological disasters. It leads widespread concern because of huge hazard force. But sometimes, the harm of earthquake landslide is larger than earthquake itself, especially in mountain area. Wenchuan earthquake is occurred in Longmen Shan region where geological environment is fragile. It has many characteristics, such as high magnitude, long duration and strong ground vibration response. The earthquake can lead to occur lots of secondary geological disasters, such as landslide, avalanche and barrier lake and debris flow. In addition, the impacts of the earthquake on secondary geological disaster also have a certain lag, because it can constitute serials of threat for people in future (Xu et al., 2009). While Wenchuan earthquake was occurring, many scholars conducted emergency survey of secondary geological disaster. They discovered that quantity and coverage in that region are very rare. Some scholars think Beichuan secondary geological disaster is more serious than Wenchuan secondary geological disaster. So the researches of Beichuan secondary geological disaster after the earthquake are very crucial for post-disaster
reconstruction, disaster area planning and economic development (Hu et al., 2007; Peng et al., 2010).

The Evaluation of Landslide Prediction (ELP) has been a hotspot and difficulty in landslide research. ELP includes two parts, time and spatial prediction. Spatial prediction is a spatial distribution probability analysis of existing or potential landslide in a certain area without time factor. It can answer the question which area is prone to occurring landslide. The theoretical basis of prediction is the area where has been occurring landslide is more prone than other area, in similar natural environment condition (Wang and Sassa, 2007). Landslide spatial prediction has been applied to lots of earthquake landslide mitigation researches. There are two ways in landslide prediction, qualitative description and quantitative analysis. Quantitative method can be summed up in two ways. There are mechanism model analysis based on process or physics and mathematical model analysis based on statistics. The former one is suitable for individual landslide, while the latter is suitable for massive landslide. Because it uses factor characteristic of landslide history caused instability to predict future trend of landslide without detailed physical and mechanical characteristics data.

Geographic Information System (GIS) technique becomes one of the main support technologies in disaster research because of advantage of spatial data processing and model building. In recent year, many researchers attempt to research in landslide disaster, mostly in rainfall landslide, which combined with mathematical statistics model and supported GIS. Seismic landslide is more complexity and uncertainty. We need to find a scientific and rational seismic landslide susceptibility assessment and space prediction method (Wang et al., 2001). Some scholars use expert rating, field investigation and expert prior knowledge to determine influence factor of landslide and adopt weight analysis to predict landslide (Xin et al., 1999). Other scholars use semi-quantitative or quantitative model, such as fuzzy comprehensive evaluation, certainty coefficient method and information model, to predict landslide. Statistical regression model is popular because of simple operation and easy to understand (Chen and Ran, 2008). Other pattern of recognition, such as neural network and SVM, are also increasingly being used for disaster research because of nonlinear capacity and adaptive capacity (Zhuang et al., 2010). However, it is still rare for seismic landslide, since Landslide is a complex physical process. It is a nonlinear system affected by lots of factors. The nonlinear method is more suitable for landslide prediction than deterministic model and general linear statistical method (Li et al., 2008). Multilayer perceptron is suitable for spatial prediction of local landslide. It has a high nonlinear mapping capability to meet the complexity of landslide. In this study, Beichuan County is chosen as research area, because there were serious earthquakes there. The author extracts the landslide affecting factor and analyze relationship between landslide distribution and factor by GIS technique. CF multiple regression model and Multilayer Neural Network model have been applied to predict secondary landslide disaster spatially, to analyze and compare with the results. And then we examined a suitable way for predicting landslide in research area. The way will help us make decision for the future researches.

**Study area and data:** The location of Beichuan County is E103°44′-104°44′, N31°41′-32°14′ in northwest of Sichuan province. This district is located in Longmen Mountain, which is transition zone from Sichuan Basin to Qinghai-Tibet Plateau with lots of mountains and complex geological environment. The causative fault of Wenchuan earthquake crosses the southeast of study area. There are much phyllite, sandstone of Silurian and carbonatite of Carboniferous. Quaternary loose stacking layers dew on both sides of the river and gully area. Strongly rock weathering provides material source for geological disaster. Furthermore, there belongs to famous Deer Mountain Storm district. Therefore geographic location, topography and geological background make extremely frequent geological disasters in there.

**Data:** In this research, the data include collapse landslide point of Beichuan County, which come from China Geological Survey Bureau, 1:200,000 local geological map, 1:250,000 basic geographic data, 30 m SRTM DEM, intensity distribution map of Wenchuan earthquake, remote sensing image, such as UAV, SPOT, Quickbird and other album information.

**METHODOLOGY**

**Technological process:** Firstly, influencing factor of landslide growth is analyzed according to priori knowledge. Then indicator of forecast and analysis is converted to raster data of united coordinate, format and size by GIS platform based on gathering fact. After that, multivariate regression model and neural network prediction model are constructed based on GIS spatial analysis technique and sample set is built according to certain rules. Sensitivity index of each pixel is calculated and classified. Finally, spatial prediction of landslide is analyzed by comparison of accuracy between the two kinds of model.

**Selection of sample set:** The sample set, which is used for spatial prediction by regression model, is including landslide sample and non-landslide sample. The non-landslide sample from each other should have low correlation according to certain rule. It can reflect
environment characteristics of non-landslide area. In this study, non-landslide point is chosen by certain rule, which is similar to landslide point. The rule is as follows. Constraint range is surrounded by landslide point. The random point is made outside the constraint range. The data grid in this study is 30 m. So the constraint range also is 30 m. There are 1364 random points have made outside of each landslide point by 30 m.

**Extraction of landslide affecting factor:** Landslide is a non-linear system because of complexity, uncertainty and existence of multiple factor. Seismic landslide is controlled by two main factors. One is impact of slope affected by the earthquake, such as seismic intensity and focal depth. Another is suitable disaster environment of pregnancy for landslide, such as terrain, topography, lithology and human activity. In this study, the following parameters are index of spatial prediction of landslide factor and analyze relationship between distribution of landslide and it.

- **Terrain factor:** Terrain factor is the most common factor in landslide research. It has been recognized the most important factor on the development of landslide by academics. Some experts even use the terrain discriminant to analyze landslide hazard. The occurrence of landslide requires two conditions in general (Fan et al., 2004). One is material source; the other is suitable sliding surface. The sliding surface is controlled by the terrain.

- **Slope:** The current research shows that slope is a necessary condition in occurrence of landslide. Slope gradient is the main controlling factor in decision of landslide decline power. It not only affected remaining decline power, which is inside the slope along existing or potential slip surface, but also affects the form and mechanism of slope deformation and failure. Furthermore, seismic ground motion amplitude and frequency have changed as the terrain height, slope shape and vary in different part.

  The slope is extracted by 30 m DEM. The study has a statistic of slope distribution at 10º for the interval...
and an overlay statistic between the slope and landslide distribution. In Fig. 1, the slope of the study area is from 0° to 83.50°. It mainly distributes area in the range of 30°-40°, which is about 62% of the total area. Most of the landslide distribution is between 30°-50° in statistic observation. In this study area, there is about 70.7% landslide. 143 landslides are less than 20° in this area, which is about 20% of the total area. More than 50° area is the lowest. There is shown in Fig. 1.

**Elevation:** The relationship between the elevation and landslide has indirect influence on the distribution of landslide. It mainly has following several aspects in former deals:

- Elevation affects the distribution of underground water.
- There are different kinds of vegetables and vegetable coverage.
- Elevation affects the preserves of human being. Human often takes activity in low elevation region, such as land reclamation.
- Seismic wave amplitude and power vary with elevation change (Xiang et al., 2010; Chen et al., 2008).

In this area, the maximum elevation is 4703 m; the minimum elevation is 508 m. Elevation is divided into 5 levels for 1000 m interval. In Fig. 2, the elevation of study area distribute in the range of 1000-2000 m. It is about 60.7% of the total area. Seismic landslide is also located in this range. It is about 56.9% of the total area. However, the area of less than 1000 m elevation is about 11.5% in study area. (38.5%) of landslide falls into this area. As a result, most of the landslide is spread out across the lower elevation place.

**Rock type factor:** Lithology has been proved to have clear control for landslide. Structural characteristic and it are basic condition for slope rock mass strength, stress distribution and failure characteristic. It determines the mechanical strength and resists weathering ability of rock and soil. Different geological age and rock type have different reaction on stability of slope under seismic action. The stratum of study area has been developed from Paleozoic to Cenozoic. In Fig. 3, it mainly includes 11 kinds of rocks, such as phyllite, mudstone, dolostone, limestone, shale, slate, siltstone, quartz sandstone and basalt. Phyllite is about

![Figure 2: Classification of elevation, relation between elevation and landslides frequency](image-url)
Fig. 3: Classification of lithology, relation between lithology and landslides frequency

Fig. 4: Classification of proximity to river system, relation between river system and landslides frequency
Fig. 5: Classification of proximity to river system, relation between proximity to river system and landslides frequency

Fig. 6: Seismic intensity map, relation between seismic intensity and landslides frequency
70.7% of the total area. In this area, there is about 64% landslide of total area.

**Water system factor**: A large number of landslides are developed in both sides of river because of erosion. It makes increase of coastal slope area. Then the upper soil slope exceeds its carrying capacity. Erosion ability of main stream is stronger than tributary. Water system of study area is developed, ravines and gullies criss-cross. Relationship of distance from the water system and landslide is analyzed in buffer of water system within 200 m. There are lots of landslides in buffer; it is about 59.3% of total. Number of landslide has distinct decay with increasing distance from the water system. There is shown in Fig. 4.

**Road factor**: Human activity is an important factor in the landslide. Because cutting slope and excavation of construction project by human change the natural feature. The most important slope of project is the road construction. Road of study area is developed. Sensitivity of different distance from the road to landslides is analyzed in buffer of road within 200 m. There are lots of landslides distribute in the range 600 m from the road. It is about 59.8% of total. Number of landslide has distinct decay with increasing distance from the road. There is shown in Fig. 5.

**Seismic intensity**: Seismic intensity is represent impact and damage extent of surface and engineering building affected by the earthquake. It canms macroscopic reflected influence of some area affected by the earthquake. Regional intensity has relationship with release energy, focal depth and distance from epicenter. It is the most appropriate indicator of seismic parameters. Landslide under different intensity are counted according to “Seismic intensity distribution map of Wenchuan Earthquake”, which is a compilation after the earthquake. The number of landslide has increased with the increasing of intensity. There is shown in Fig. 6.

**PREDICTION ANALYSIS OF LANDSLIDE HAZARD AND VALIDATION**

**Landslide spatial prediction based on CF multivariate regression model**: The problem of argument quantization of different types of data in regression analysis can be resolved in a certain extent, while CF value as argument of regression analysis. CF is an abbreviation of coefficient of certainty. It is a probability function for degree of certainty of event occurrence. The specific formula is:

\[
CF = \begin{cases} 
PP_a - PP_s & PP_a \geq PP_s \\
PP_s(1 - PP_s) & PP_s < PP_a
\end{cases}
\]

where, PP_a is conditional probability of occurrence of seismic landslide in data set a. It is pixel ratio between landslide occurrence and unit in every classification result of each data layer. PP_s is a priori probability of occurrence seismic landslide in study area. It is pixel ratio between landslide and total. The range of CF value is from -1 to 1. A positive value indicates greater certainty of landslide instability and more prone to landslide. A negative value indicates smaller certainty of landslide instability and less prone to landslide. CF value can respond the contribution of landslide instability affected by some factors. Therefore, CF value can be operated as argument of multiple regressions. General multivariate regression model is:

\[
Y = B_0 + \sum_{i=1}^{n} B_i X_i 
\]

where, 
Y : Dependent variable. It represents probability of unit landslide occurrence in this study
X_i : Argument. It represents CF value, which is factor affecting by landslide
B_i : Regression coefficient
B_0 : Constant

The extracted factor has been classified according to previous chapter. Landslide proportions of different levels in each type of data can be obtained by overlay analysis. Then the CF value of each type data and conditional probability can be calculated. The result is imported into statistical analysis software for multivariate regression analysis. The regression coefficient value of impact factor can be calculated (Table 1). Finally, the probability of landslide and earthquake-induced landslide susceptibility index can be calculated according to formula 2.

The range of susceptibility index is form -1.086 to -0.172 according to result. Nature break method is a commonly classification method in statistics. It takes similar status value into a group following the numerical distribution. The difference within the group is the least and difference without the group is the most. It is suitable for swarm event, such as geological hazard. Landslide occurrence probability is divided into very high, high, moderate and mild following the nature break method. The result is shown in Fig. 7.

**Sensitivity analysis based on multilayer perceptron neural network model**: Artificial neural network is a non-liner dynamic system, which simulate and make abstract to human or natural neural networks, rising in 20th century. It, which widely used in computer,
Table 1: List of regression coefficients

<table>
<thead>
<tr>
<th>Constant term</th>
<th>Distance from the road</th>
<th>Elevation</th>
<th>Slope</th>
<th>Earthquake intensity</th>
<th>Distance from river system</th>
<th>Rock type</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_0 )</td>
<td>( B_1 )</td>
<td>( B_2 )</td>
<td>( B_3 )</td>
<td>( B_4 )</td>
<td>( B_5 )</td>
<td>( B_6 )</td>
</tr>
<tr>
<td>0.488</td>
<td>0.08</td>
<td>0.21</td>
<td>0.328</td>
<td>0.144</td>
<td>0.219</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 2: List of MLP train coefficients

<table>
<thead>
<tr>
<th>Correlation parameter</th>
<th>Hidden junction points</th>
<th>Training points</th>
<th>Momentum coefficient</th>
<th>Learning rate</th>
<th>Mean square error</th>
<th>Mean square error</th>
<th>Accuracy rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter values</td>
<td>3</td>
<td>1000</td>
<td>0.5</td>
<td>0.0006</td>
<td>0.000443</td>
<td>0.000444</td>
<td>81.14%</td>
</tr>
</tbody>
</table>

FIGURES

Fig. 7: Landslide spatial prediction map (CF multiple regression model)

Fig. 8: Landslide spatial prediction map (MLP model)

RESULT ANALYSIS AND VERIFICATION

Both results of models can predict the spatial distribution of landslide in the future. In comparison, the distribution of landslide susceptibility is relatively uniform. The sensitivity of east is substantially higher than west. Very high sensitive area is mainly distributed to both sides of the river and road in southeast fault zone, which are banded and flake. High and moderate sensitive area is mainly distributed to middle area. Mild sensitive area is mainly distributed to western mountain area.
To compare the distribution of seismic landslide and assessment result, predictive validity of 2 models can be further analyzed by quantitative. As shown in Table 3 and 4, very high and high sensitive area of 2 models in calculation result has the most landslide point. There are 57% landslides in 13.14% of total area by CF multivariate regression model. The density is 0.89/km². There are 59% landslides in 13.2% of total area by MLP model. The density is 1.03/km². Both can be summed up more than half of landslide in a small area. It has high assessment accuracy. This study introduces precision estimation method proposed by Sridevi Jadi and further analyzes prediction accuracy in quantitative. Assessment accuracy P expressed by empirical probability (Sridevi et al., 1997; Tao et al., 2010):

\[ P = \frac{KS}{S} \left( 1 - \frac{K - KS}{N - S} \right)^{\frac{1}{2}} \]  

(3)

where,
- \( P \): Prediction accuracy
- \( S \): Total number of landslide unit
- \( KS \): Number of landslide unit in very high, high and moderate sensitive area
- \( K \): Total number in very high, high and moderate sensitive area
- \( N \): Total raster unit

In this study, 4 assessment classes is merged into 2 classes. \( K \) is 1365729 and \( KS \) is 590 in CF multivariate regression model. Then, \( K \) is 1023648 and \( KS \) is 614 in multilayer perceptron neural network model. \( N \) is 3189480, \( S \) is 664. According to the formula 3, the prediction accuracy of CF multivariate regression model and MLP model are 73.7 and 81.28%. Obviously, the prediction accuracy of multilayer perceptron neural network model is more than 80% and higher than CF multivariate regression model.

### CONCLUSION AND DISCUSSION

- In this study, landslide spatial prediction research is adopted 2 methods in Beichuan heavy disaster area. It is a basic work of geological disaster prevention and geological hazard prediction after the earthquake and based on the area, which has been occurred landslide, is prone to occur landslide.
  - By statistical analysis of the relevant study on the relationship between landslide control parameter and landslide distribution, slope of 30°-50° is prone slope of seismic landslide, elevation of 1000-2000 m is prone elevation of seismic landslide, phyllite and siltstone is prone rock of seismic landslide, within 400 m of water and with 600 m of road are prone segment of seismic landslide.
  - Spatial prediction model is constructed; whose argument of multiple regressions is CF value of landslide affect factor. It can solve the consolidation of heterogeneous data. Prediction accuracy is 73.7%. Artificial Neural Network is one of the most common non-linear models. Its prediction accuracy is 81.28%. Landslide system is a non-liner system. So, non-linear neural work method is more suitable for hazard spatial prediction in study area. As shown in result, the linear effect of regression model is more visible.
  - The distribution of high sensitive area by 2 models is similar. High sensitive area of seismic landslide in Beichuan mainly distribute in Qu Mountain, Chen Jiaba Qiang autonomous township and so on in the east. Sensitive of landslide is relatively low in west. However, low sensitive area does not means disaster not occurred. It has relationship with catalog of geological hazards data.
  - The prediction result of 2 methods has large differences without high sensitive area. Multiple regression models are a linear regression model, therefore the result trend to linear distribution. The difference of low, medium and high is not too large because of natural break method. It is a balance distribution. Multilayer perceptron is a non-linear model based on BP algorithm. It has strong adaptive ability. In processing of training, it can continue to adjust the parameters to achieve optimal result. There are more efficient distinguish in high and low sensitive area.
  - The loose debris of seismic landslide can make the debris flow and barrier lake, especially in very high sensitive area. Therefore, reconstruction plan should be avoided very high and high sensitive area of geological hazard.
ACKNOWLEDGMENT

The author thanks the anonymous reviewers for their valuable remarks and comments. This study is supported by 2012 National Science and Technology Pillar Program of China (2012BAH33B03, 2012BAH33B05).

REFERENCES