Application and comparison of logistic regression model and neural network model in earthquake-induced landslides susceptibility mapping at mountainous region, China

Peng Xie\textsuperscript{a,b}, Haijia Wen\textsuperscript{a,b,c}, Chaochao Ma\textsuperscript{a,b}, Laurie G. Baise\textsuperscript{c} and Jialan Zhang\textsuperscript{b}

\textsuperscript{a}Key Laboratory of New Technology For Construction of Cities in Mountain Area (Chongqing University), Ministry of Education, Chongqing, China; \textsuperscript{b}School of Civil Engineering, Chongqing University, Chongqing, China; \textsuperscript{c}Department of Civil and Environmental Engineering, Tufts University, Medford, MA, USA

ABSTRACT

The main objective of this study is to evaluate the performances of different earthquake-induced landslides susceptibility mapping models at mountainous regions in China. At first, 160 earthquake-induced landslide points were identified from field investigations. Concurrently, based on the results of a literature review and the field investigation, 12 influencing factors were considered, and the corresponding thematic layers were generated using geographic information system (GIS) technology. Subsequently, 20 groups with a fixed number of cells were collected as a common training dataset for the two different models, based on a random selection from the entire database (including landslide cells and no-landslide cells). The neural network (NN) model and logistic regression (LR) model were developed with R software. Finally, earthquake-induced landslides susceptibility maps of Wenchuan county were produced, very low, low, medium, high and very high susceptibility zones cover. The validation results indicate that the landslide data from field investigations are in good agreement with the evaluation results, and the LR model has a slightly better prediction than the NN model in this case. In general, the NN model and LR models are satisfactory for susceptibility mapping of earthquake-induced landslides at mountainous regions.

1. Introduction

Landslide is one of the types of natural disasters that cause severe damage, including human and economic losses (Chen and Huang 2013; Wen 2015). Earthquake-induced landslides have been documented from at least as early as 373 or 372 B.C. Earthquakes have been recognized as a major cause of landslides. China has a long history of earthquakes: a neighbouring region was affected by M7.3 event in 1933 (Brown et al. 2012). According to existing statistical data, 301 earthquakes struck China in the past 25 years (from 1990 to 2014), and there has been an average annual incidence rate of 12 times (Liu et al. 2014). The M7.9 Wenchuan earthquake in the Sichuan province has been the most destructive and deadly earthquake for more than 30 years (Brown et al. 2012). After the 12 May 2008 earthquake, the Sichuan Provincial Department of Land and Resources organized and implemented a geological disaster emergency investigation work of 36 geological disaster-prone counties (cities, districts). Numerous members of our team participated in the geological disasters investigation in Wenchuan county, which was undertaken by the Chongqing Shutong Geotechnical...
Engineering Company. The site investigation indicates there are 160 landslide points in the entire territory of Wenchuan county.

Earthquake-induced landslides are the significant composition of geo-hazards and can have considerable and destructive consequences. Correspondingly, these landslides have caused widespread concerns in recent decades. So far, a large number of studies have been conducted addressing many aspects of these phenomena. For a better understanding of the detailed information of earthquake-induced landslides, including the spatial location, the type, the geometrical parameters, traditional site investigation; remote sensing (RS) technology and geographic information system (GIS) technology have been used in practice (Huang 2000; Keefer 2002; Mohammad et al. 2006; Masahiro and Hiroshi 2004; Zhang and Wang 2007; Huang and Li 2009; Dai et al. 2011; Yin et al. 2009; Gorum et al. 2011; Huang et al. 2013; Li et al. 2013a; Xu et al. 2014; Xu et al. 2014a, 2014b; Xu et al. 2015; Guo et al. 2017; Rosser and Carey 2017). Furthermore, based on comprehensive analysis of statistic data and actual conditions, the spatial distribution characteristics, the initiation mechanism and causative factors, have been summarized (Keefer 2000; Masahiro and Hiroshi 2004; Sato et al. 2005; Meunier et al. 2007; Rajabi et al. 2010; Wang et al. 2010; Chen et al. 2011; Lee 2013; Zhang et al. 2013; Zheng et al. 2013; Yin et al. 2014; Xu et al. 2014a, Xu et al. 2014a, 2014b; Xu and Xu 2014c; Tang et al. 2015; Xu et al. 2016; Tian et al. 2017). An important subject is the assessment of the stability or performance of the slope under seismic conditions, whereby the existing analytical methods can be classified into three different categories (Tropeano et al. 2017). Furthermore, the Newmark model and related empirical predictive models have been frequently used for earthquake-induced slope displacement simulation and landslide hazard assessment at a regional scale (Wilson and Keefer 1983; Jibson 1993; Miles and Ho 1999; Paruelo et al. 2004; Sayg and Rathje 2008; Wang et al. 2011; Pareek et al. 2014; Borfecchia et al. 2016; Chousianitis et al. 2016; Du and Wang 2016; Caccavale et al. 2017; Shinoda and Miyata 2017). Landslide susceptibility assessment can assess the spatial distribution of potential landslides in an area, and provide valuable assistance for hazards mitigation. To achieve a scientific landslides susceptibility mapping, multivariate statistical methods and the data-driven methods were used to examine the relative strength and significance of influencing factors, including analytic hierarchy process (AHP) model (Wen et al. 2017), logistic regression (LR) model (Ohlmacher and Davis 2003; Ayalew and Yamagishi 2005; Lee 2005; Nefeslioglu et al. 2006; Bai et al. 2013; Xu et al. 2013; Wang et al. 2016), ANN model (Lee and Evangelista 2006; Yilmaz 2010), frequency ratio (Yilmaz 2009; Pradhan and Lee 2010; Umar et al. 2014), support vector machines (SVM) (Yesilnacar and Topal 2005; Xu et al. 2012) and so on.

According to the available literature, significant achievements have been made, and a series of methods have been induced for earthquake-induced susceptibility assessment. However, no uniformly agreed-upon method is available for predicting the likelihood and spatial extend of seismically-induced landslides (Nowicki et al. 2014). The main purpose of the study is to evaluate the performances of different susceptibility mapping approaches of earthquake-induced landslides at mountainous region. LR and neural network (NN) models are the multivariate statistical and the data-driven methods, respectively, and they are well-suited to analyse the presence/absence of a variable. Therefore, the susceptibility assessments of earthquake-induced landslides using LR and NN models were conducted in Wenchuan county. The receiver operating characteristic (ROC) curves and area under the curves (AUC) were used for validation purpose.

2. Study area

2.1. Description of study area

On 12 May 2008, a strong earthquake with a magnitude M7.9 struck the Sichuan province. The occurrence of the Wenchuan earthquake resulted from the collision of the Indian and the Eurasian plates. The Indian plate had been moving northward resulting in the uplift of the Tibetan plateau (Xu et al. 2013). The focal depth ranged from 14 to 19 km. According to official sources on 10 July
2008, there were 69,197 casualties, 374,176 injured and 18,377 reported missing. Approximately
1.8 million people lost their homes and needed at least temporary relocation. Figure 1 is a seismic
map of Wenchuan earthquake, obtained from the USGS website. The epicentre was in the town of
Yingxiu that was located in the southeast of Wenchuan county. Figure 2 presents the town of Ying-
xiu based on a low-altitude aerial photo acquired on 14 May 2008 from the Fengone net. It shows
that the earthquake damage was very serious at a meizoseismal area, and trigged numerous
landslides.

Wenchuan county is one of the most affected areas, has undergone significant destructions and it is at the junction of the Sichuan basin and the Tibetan plateau (Figure 3). Wenchuan county covers the latitudes of 30°45’37” and 31°43’10” and the longitudes of 102°51’46” and 103°44’37”. Ridges, valleys and slopes are the main landforms of this area, with elevations ranging between 572 and 6132 m. It has a northeast orientation and occupies the south-western part of Longmen mountain fault belts, including Maowen fault, Jiudingshan, Yingxiu fault and Erwangmiao faults. The geological time in the Wenchuan county ranges from Cenozoic Quaternary, Mesozoic, Jurassic, Cretaceous, Paleozoic, to Precambrian, and granite, basalt, dolomite, limestone, sandstone, slate, shale, phyllite, etc. are the dominant lithology in these zones (Figure 4).

2.2. Distribution of earthquake-induced landslides in study area

To derive the actual distribution of landslides triggered by the Wenchuan earthquake, the Sichuan Provincial Department of Land and Resources organized and implemented a geological disaster emergency investigation work for 36 geological disaster-prone counties (cities, districts). The investigation work of Wenchuan county was undertaken by the Chongqing Shutong Geotechnical Engineering Company, where one part of the field work was undertaken by our research team. Figure 5 shows the actual distribution of the earthquake-induced landslides in the entire territory of the Wenchuan county, whereby 160 landslide points were identified. Comparing the distribution of earthquake-induced landslides with a map of Wenchuan county, it is apparent that the landslides
are mainly located in the areas most heavily affected by the earthquake, such as the towns of Yingxiu and Longxi.

3. Influencing factors and spatial database

3.1. Influencing factors

To develop a universal applicable mapping model of earthquake-induced landslides, the determination of the influencing factors and the measured parameters is crucial. Some researchers believe that the selection of the influencing factors for the assessment of landslides depends on the assessment scale, the landslide type, the failure mechanisms and the main causes of landslides (Xu et al. 2013; Li et al. 2013a, 2013b). In the meanwhile, a spatial database for landslide-related factors could be constructed easily and efficiently, and is another most important principle. Based on the information of the site investigations and several examples from the literatures, 12 influencing factors have been defined, including peak ground acceleration (PGA), distance from faults, distance from drainage, distance from highways, slope gradient, average annual rainfall, normalized difference vegetation index (NDVI), lithology, slope position, elevation, slope aspect and micro-landform.

3.2. Spatial database

All the 12 influencing factors (PGA, distance from faults, distance from drainage, distance from highways, slope gradient, average annual rainfall, NDVI, lithology, slope position, elevation, slope aspect and micro-landform) have been used to create the spatial database. The spatial database was built using the software ARCGIS 10.1.
For the morphometric factors, the data of slope gradient, slope position, slope aspect and micro-landform were derived from 30-m resolution digital elevation model (DEM), and the data of elevation is extracted from 30-m resolution DEM from Aster satellite data. For the geological factors, the data of lithology and distance from faults were derived from geological maps at a scale of 1:200,000, and ARCGIS 10.1 was used to transform the geologic vector map to a raster image. The data of distance from streams is derived from 30-m resolution DEM. The data of NDVI was downloaded from NASA’s website. For other factors, the data of PGA was downloaded from USGS’s website, and the

Figure 3. Location of study area, Wenchuan county.
distance from highway was from GOOGLE EARTH. The average annual rainfall layer was obtained by accessing the local average rainfall data spanning several years and rainfall contour was generated.

The larger relative elevation could produce larger gravitational potential energy and the amplification effect of high-altitude topographies, which always aggravate the landslide disasters when earthquake occur (Tian et al. 2017). Slope gradient is an important topographic influencing factor, and many landslides shared an obvious common feature of being distributed on steep slopes during field investigation (Qi et al. 2010). PGA and the distance from fault are considered as the seismic factors, and vegetation has the ability to stabilize the slope materials (Zhang et al. 2013). The seepage effect of rainfall could weaken the properties of rock and soil, resulting in brittleness and fragility (Bhandary et al. 2013; Tian et al. 2017). Highway construction increases the instability of slopes, and the distance to highway may reflect the strength of human engineering activities in mountains region. The landslide density decreases gradually with greater distance to highways (Sato et al. 2005;
Qi et al. (2010); Lee (2013); Tang et al. (2015). Based on experts’ experience, field investigation and the results of a literature review, the classification was performed using equal intervals in the value range for the continuous data. Lithology always plays a fundamental role on landscape evaluation and landslide occurrences (Xu et al. 2013; Liu et al. 2014; Wen et al. 2017). According to physical and mechanical features of the lithology, 10 categories were defined. Slope aspect and micro-landform are morphometric factors, and they can be classified into 9 and 10 categories, respectively (Weiss 2001). Slope position can be quantified by the topographic position index (TPI), slope position can be classified into 6 categories (Weiss 2001). Each location in the study area was classified based on the values of all 12 factors at that location, and the classification scheme used for each factor is shown in Table 1. Figure 6 shows the results of the classification.

To sum up, all the above 12 independent parameters are all prepared in ARCGIS format with 90 × 90 m cell size, the total number of cell is 556,884. In the meanwhile, the classes and rating values of 12 independent parameters could be assigned to each cell unit.
| Serial number | Factors | Classifications | Rating value | Serial number | Factors | Classifications | Rating value |
|---------------|---------|-----------------|--------------|---------------|---------|-----------------|--------------|
| 1             | PGA (Pg/g) | >0.8            | 1            | 8             | Lithology (Li) Q |          | 10             |
| 0.7–0.8       | 2        | D_{1p}P_{2}     | 9            |               | S_{max3}T_{1b}T_{3x3} | 8          |
| 0.6–0.7       | 3        | S_{max1}S_{max3}D_{2+3} | 7          |               | D_{wsl1}D_{wsp2}Ds1 | 8          |
| 0.5–0.6       | 4        | S_{max5} | 6          |               | O.S_{max5}P_{1y}T_{3av}C_{2yv} | 6          |
| 0.4–0.5       | 5        | D_{2}Z_{bd}    | 7            |               | P_{thn1}D_{1y}D_{2y}D_{3z} | 8          |
| 0.3–0.4       | 6        | P_{thn1}P_{thn2} | 3          |               | Z_{av}P_{wh},\delta | 1          |
| 0.2–0.3       | 7        | <0.2            | 8            |               | P_{thn1}P_{thn2}C_{04} | 2          |
| 0.1–0.0       | 8        | 200–300         | 4            |               | Elevation (El/m) <1000 | 1          |
| 2             | Distance from faults (Df/m) | <2.0 | 7            |               | 1000–1500 | 2          |
|               |          | 2.0–4.0         | 6            |               | 1500–2000 | 3          |
|               |          | 4.0–6.0         | 5            |               | 2000–2500 | 4          |
|               |          | 6.0–8.0         | 4            |               | Valleys | 5          |
|               |          | 8.0–10.0        | 3            |               | Ridge | 6          |
|               |          | 10.0–12.0       | 2            |               | Midslope | 1          |
|               |          | >12.0           | 1            |               | Flat | 2          |
| 3             | Distance from streams (Ds/m) | <100 | 6            |               | LowSlope | 3          |
|               |          | 100–200         | 5            |               | UplSlope | 4          |
|               |          | 200–300         | 4            |               | Valleys | 5          |
|               |          | 300–400         | 3            | 10 | Elevation (El/m) <1000 | 1          |
|               |          | 400–500         | 2            |               | 1000–1500 | 2          |
|               |          | >500            | 1            |               | 1500–2000 | 3          |
|               |          | >1000           | 8            |               | 2000–2500 | 4          |
| 4             | Distance from highways (Dh/m) | <100 | 7            |               | Valleys | 5          |
|               |          | 100–200         | 7            |               | Ridge | 6          |
|               |          | 200–300         | 6            |               | Valleys | 5          |
|               |          | 300–400         | 5            |               | Ridge | 6          |
|               |          | 400–500         | 4            |               | Valleys | 5          |
|               |          | 500–600         | 3            |               | Ridge | 6          |
|               |          | 600–700         | 2            |               | Valleys | 5          |
|               |          | >700            | 1            |               | Ridge | 6          |
| 5             | Slope gradient (Sl/°) | <10 | 11 | Slope aspect (As/°) | -1/Flat | 1          |
|               |          | 10–20           | 2            |               | 0–22.5 or 337.5–360/N | 2          |
|               |          | 20–30           | 3            |               | 22.5–67.5/NE | 3          |
|               |          | 30–40           | 4            |               | 67.5–112.5/E | 4          |
|               |          | 40–50           | 5            |               | 112.5–157.5/SE | 5          |
|               |          | 50–60           | 6            |               | 157.5–202.5/S | 6          |
|               |          | >60             | 7            |               | 202.5–247.5/SW | 7          |
| 6             | Average annual rainfall (Ra/mm) | <525 | 1             |               | 247.5–292.5/SW | 8          |
|               |          | 525–625         | 2            |               | 292.5–337.5/NW | 9          |
|               |          | 625–725         | 3             | Micro-landform (Lf) | Canyons, deeply incised streams | 1          |
|               |          | 725–825         | 4             |               | Midslope drainage, shallow valleys | 2          |
|               |          | 825–925         | 5             |               | Upland drainage, headwaters | 3          |
|               |          | 925–1025        | 6             |               | Upland drainage, headwaters | 3          |
| 7             | NDVI (Nd) | >1025           | 7            |               | U-shape valleys | 4          |
|               |          | <0              | 4            |               | Plains | 5          |
|               |          | 0–0.1           | 1            |               | Open slopes | 6          |
|               |          | 0.1–0.2         | 2            |               | Upper slopes, mesas | 7          |
|               |          | 0.2–0.3         | 3            |               | Local ridges in valleys | 8          |
|               |          | 0.3–0.4         | 4            |               | Midslope ridges, small hills in plains | 9          |
|               |          | 0.4–0.5         | 5            |               | Mountain tops, high ridges | 10         |
|               |          | 0.5–0.6         | 6            |               | Plain | 5          |
|               |          | 0.6–0.7         | 7            |               | Open slopes | 6          |
|               |          | >0.7            | 8            |               | Upper slopes, mesas | 7          |

**Table 1.** Classification of each influencing factors.
Figure 6. The influencing parameters of landslide: a. PGA (Pg), b. Distance from faults (Df), c. Distance from Highways (Dh), d. Distance from streams (Ds), e. Slope angle (Sl), f. Average annual rainfall (Ra), g. NDVI (Nd), h. Lithology (Li), i. Slope position (Sp), j. Elevation, k. Slope aspect (As), l. Micro-landform (Lf).
Figure 6. (Continued)
4. Susceptibility mapping model

4.1. Neural network (NN) model

NN is a non-linear soft computing method with strong non-linear mapping capability. To develop the model, three steps can be identified as follows: (1) data selection for the network training, (2) multi-layer perceptron (MLP) design and (3) network training (Arnone et al. 2014).

(1) Training samples selection

In the study, the entire database consisted of landslide cells (1) and no landslide cells (0). Landslide cells consisted of 160 cells and each landslide point became a single cell. In consideration of the geometric effect, all landslide cells were excluded by setting up a 300-m buffer zone for all the 160 landslide points, and the remaining area is the no-landslide cells. To develop an accurate susceptibility mapping model of earthquake-induced landslide, it mainly depends on effective training samples. There were different opinions among researchers for the number of training samples (Ayalew and Yamagishi 2005). Some researchers use the total or only part landslide pixels/points and randomly selected an equal number of non-landslide pixels from landslide free area. Besides, it is recommended to use equal proportions of 1 (landslide) and 0 (no-landslide) pixels in a statistical model (Ohlmacher and Davis 2003; Ayalew and Yamagishi 2005; Xu et al. 2013; Li et al. 2013a, 2013b). Due to the relatively small number of landslides, from each of the two classes (landslide cells and no-landslide cells), 20 groups with a fixed number of cells were randomly selected as a common training sample database in the study, and the ratio between landslide and no-landslide cells is near to 1:5 (149 landslide cells and 751 non-landslide cells for each group).

(2) MLP design

The definition of the MLP structure requires the definition of input, hidden and output layers and the number of nodes for each layer. There were 12 input layers in this study, which corresponds to the number of influencing factors. In regard to the number of hidden layers, the entire process was performed with the use of NNET package of R software. First, three different hidden layers were used, and the number of hidden layers were 1, 2, and 5, respectively. Concurrently, different weight decays were defined, and the corresponding parameters were 0, 0.1, and 0.001. During the software self-optimizing, root mean squared value (RMSE) was selected as the evaluation criteria. After a 760-step iterative calculation, a single hidden layer is determined to be used in landslide analysis applications. In addition, there were 20 nodes for the hidden layer. One node was used in the output layer. Therefore, the structure of the NN was 12-20-1, as shown in Figure 7.

(3) Network training

During the training process, two groups with a fixed number of cells were selected randomly from the training sample dataset, whereby one group was used to train and the other one was used for validation. The whole process was performed using NNET package of R software. Output values ranged from 0 to 1, corresponding to the output value of susceptibility at each cell. To quantify the validation of training, the success rate of the validation sample can be used on the basis of statistical analysis. As shown in Table 2, 102 landslide cells and 711 non-landslide cells among the validation sample were predicted correctly. The overall accuracy of validation dataset is 90.33%.

4.2. Logistic regression (LR) model

LR model is one of the multivariate statistical analysis models, and it is useful for predicting the presence or absence of a characteristic based on the values of a set of controlling factors. The
The association between the predictor variables and the presence/absence of the landslides is tested using the maximum likelihood model. The relationship between the probability of landslide occurrence and the independent variables can be expressed as follows:

$$P = \frac{1}{1 + e^{-Y}} \quad (1)$$

where $Y$ can be defined as follows:

$$Y = C_0 + \sum_{i=1}^{n} C_i X_i \quad (2)$$

where $P$ is the probability of landslide occurrence, $Y$ is the weighted linear combination of the independent variables, $C_0$ is the intercept of the model, $C_i$ ($i = 1, 2, ..., n$) is the coefficient estimated from the sample data, $n$ is the number of independent variables and $X_i$ ($i = 1, 2, ..., n$) is the independent variable, and $n$ is equal to 12 in this study. The output probability values range from 0 to 1, with 0 indicating no-landslide occurrences and 1 indicating 100% (Dai et al. 2004).

For the LR model, the same training sample database was used, which had been determined by NN model. Eleven groups with a fixed number of cells were selected randomly from the training sample database, 10 groups with a fixed number of cells were used to train and the remaining 1 group with a fixed number of cells was used for validation. Regression analysis was performed using the glm() function of R software, and the average of 10 groups with a fixed number of cells regression analysis results was determined as the correlation coefficients, as listed in Table 3.
In order to quantify the quality of the regression, the success rate of the validation sample was carried out. Statistical analysis outcomes are shown in Table 4; 105 landslide cells and 714 non-landslide cells among the training sample were predicted correctly. The overall accuracy of the sampling dataset was 90.56%. This showed that the LR model was slightly more accurate than the NN model.

5. Susceptibility mapping

5.1. Susceptibility mapping using neural network (NN) model

Based on the designed MLP network model, P values (ranging from 0 to 1) for all 556,884 cells were calculated using ARCGIS 10.1 software. The study area was classified in five categories (<0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and >0.8) using four break points (0.2, 0.4, 0.6, and 0.8), corresponding to very low, low, medium, high, and very high susceptibility zones cover. An earthquake-induced landslides susceptibility map of Wenchuan county is shown in Figure 8.

5.2. Susceptibility mapping using logistic regression (LR) model

Based on Equation (1), P values of all 556,884 cells were calculated using ARCGIS 10.1 software. The study area was classified in five categories (<0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and >0.8) using four break points (0.2, 0.4, 0.6, and 0.8), corresponding to very low, low, medium, high, and very high susceptibility zones cover. An earthquake-induced landslides susceptibility map of Wenchuan county is shown in Figure 9.

6. The validity of susceptibility mapping models

6.1. The validity of neural network (NN) model

Figure 8 is an earthquake-landslide susceptibility map (produced based on NN model) of Wenchuan county, which is overlaid by the earthquake-induced landslide points (obtained from field investigations). The high and very high susceptibility class areas match well with the actual earthquake-induced landslide points.
Furthermore, statistical analysis was used to evaluate the effectiveness of the model quantitatively, including the cells numbers and coverage percentages of each susceptibility classification, landslide point mounts, and percentages and density of each susceptibility class area. All the statistical data are shown in Table 5. According to the statistical results, Figure 10 plots the rating area percentage versus rating landslide point percentage based on the NN model. For the NN model, more than 67.5% of field landslides points were located in approximately 6.7% coverage (high and very high susceptibility classes), but only about 14.3% of the field landslide points were located in approximately 87.8% coverage (low and very low susceptibility classes). The landslide point density increased from $0.02428 \times 10^{-3}$ to $4.29288 \times 10^{-3}$, reflecting an approximate increase of 177 times, when the susceptibility class varied from very low to very high.

The AUC value (area under the receiver operating characteristic ROC curve) is usually used to test the validation of the model (Xu et al. 2013; Li et al. 2013a, 2013b). As shown in Figure 11, the AUC value is 0.930 for the NN model, which indicates that the model possesses high predictive power.

Figure 8. Earthquake-landslide susceptibility map (produced based on neural network model) of Wenchuan county is overlaid by earthquake-induced landslide points (obtained from field investigation).
6.2. Validity of logistic regression (LR) model

Figure 9 is the earthquake-landslide susceptibility map (produced by logistic regression model) of Wenchuan county overlaid by earthquake-induced landslide points (obtained from field investigation), and shows that high and very high susceptibility class areas match well with the location of the actual earthquake-induced landslide points.

Furthermore, statistical analysis was conducted, all the statistical data are listed in Table 6. According to the statistical results, more than 61.2% field landslide points are located in approximately 4.2% coverage (high and very high susceptibility classes) and approximately 21.2% field landslide points are located in larger 92.6% coverage (low and very low susceptibility classes). The
landslide point density increased by approximately 186 times, from $0.02703 \times 10^{-3}$ to $5.02128 \times 10^{-3}$, when the susceptibility class varied from very low to very high (Figure 12).

As shown in Figure 13, the AUC value is 0.941 for the LR model, which indicates that the model possesses high predictive power.
In this study, the earthquake-induced landslide analyses were performed using the LR and NN models. Subsequently, the analysis results were validated based on success rate of the validation sample. For NN model and LR model, the success rates of the sampling database are 90.33% and 90.56%, respectively (Table 2, Table 4). In addition, the area under the ROC curve (AUC value) characterizes the predict accuracy of the predefined models. The reduction of the AUC values for LR model (94.1%) and NN model (93%) is about 1.1% (comparing Figure 13 with Figure 11). The validation results shows that the LR model has slightly better prediction than the NN model in this case.

Figure 12. Relationship between rating area percentage and rating landslide (produced by the logistic regression model).

Figure 13. ROC curve (produced by the logistic regression model).

7. Comparison between neural network (NN) model and logistic regression (LR) model

In this study, the earthquake-induced landslide analyses were performed using the LR and NN models. Subsequently, the analysis results were validated based on success rate of the validation sample. For NN model and LR model, the success rates of the sampling database are 90.33% and 90.56%, respectively (Table 2, Table 4). In addition, the area under the ROC curve (AUC value) characterizes the predict accuracy of the predefined models. The reduction of the AUC values for LR model (94.1%) and NN model (93%) is about 1.1% (comparing Figure 13 with Figure 11). The validation results shows that the LR model has slightly better prediction than the NN model in this case.
8. Discussion and conclusion

8.1. Discussion

8.1.1. Identification of influencing factors

Based on the integrated results of a literature review and the field investigations, 12 influencing factors were defined, including PGA, distance from faults, distance from drainage, distance from highways, slope gradient, average annual rainfall, NDVI, lithology, slope position, elevation, slope aspect and micro-landform.

The intense ground motion could cause a short-lived distribution in the balance of forces within natural slopes, which is the main reason for landslides. The Arias intensity is the most suitable for characterizing earthquake impacts (Hsieh et al. 2011; Chousianitis et al. 2014), and it is estimated by the integration of the squared acceleration over time. However, owing to the fact that the Arias intensity incorporates more information content, a single parameter, the most commonly used parameter to describe earthquake ground motion is PGA (Xu et al. 2013; Li et al. 2013a; Hsieh et al. 2011; Chousianitis et al. 2014; Li et al. 2013b; Tian et al. 2017).

It is well known that the fault not only increases PGA values, but also reduces the mechanics strength of rock mass in the landslide-prone areas. Thus, almost all researchers used distance from faults as an indispensable parameter in landslide susceptibility mapping studies (Huang and Li 2009; Yin et al. 2009; Xu et al. 2013; Li et al. 2013a, 2013b).

Highway network is the significant composition of transportation system in Chinese mountainous region, and the main engineering activities are commonly adjacent to highway, so the distance from highways can be used to evaluate the effect of engineering activities. Both sides of highway are particularly prone to landslides driven by the collapsing force of an earthquake (Tian et al. 2017).

Many researchers believe that the surface topography such as elevation, lithology, slope gradient, slope aspect and slope position plays an important role in landslide distribution (Xu et al. 2013; Li et al. 2013a, 2013b). Lithology is a representative geological data, its effect can be measured by the physical and mechanical features of the lithology. The slope aspect refers to the compass direction of the slope surface, which would affect the characteristics of natural environment indirectly (vegetation growth, the distribution of rainfall and so on), and slope position can be quantified by the topographic position index (TPI). As slope gradient increases, the degree of gravity-induced shear stress in the slope also increases.

The distribution of rivers has a wide span in China, and it is not only a significant component of the geological environment, but also is indispensable natural resources. The water content of the soil and rock may be different near to rivers. In addition, the flow of river may change the geometric characteristics of the slope, and the free face forms. Therefore, many researchers believe that distance from streams plays an important role in landslide distribution (Xu et al. 2013; Li et al. 2013a, 2013b).

When earthquakes occur, the seepage effect of water could weaken the properties of rock and soil, thereby resulting in brittleness and fragility. In addition, the different initial conditions of soil water content may produce different geometrical shapes of landslides. Thus, water is one major influencing factor of earthquake-induced landslides.

Vegetation can improve slope stability in mountainous region by altering the mechanical and hydrological properties of the soil and reducing surficial erosion and mass wasting. The significant impact has been understood and documented (Zhang et al. 2015). Normalized difference vegetation index (NDVI) could be seen as the parameter to measure the coverage of vegetation (Xu et al. 2013).

To develop a universal applicable mapping model of earthquake-induced landslides at mountainous region, it is necessary to meet the requirement of abundance and accessibility. The 12 influencing factors could reflect the complexity of earthquake-induced landslide. In addition, the corresponding thematic layers were obtained in GIS technology, it is beneficial to construct geospatial database accurately and quickly.
8.1.2. The validation of the two developed models

The logistic regression model (Hosmer & Lemeshow) is well suited to analyse the presence/absence of a dependent variable (Cuzzetti et al. 1999; Lee 2005; Pradhan and Lee 2010), and the relationship between independent variables is specific. The variables may be either continuous or discrete, or any combination of both types, and they do not necessarily have to be normally distributed (Lee 2005). In addition, the multivariate statistical method offers the advantages of significantly reducing the number of factors to be analysed, allowing the identification of additional variables that more influence the dependent variable. Moreover, logistic model is based on the basic assumption that all the variables are independent, and the relationship is linear. For LR model, the original values are used as the input for the continuous data, and the categorical values are used for the remaining data (lithology, slope position, slope aspect and micro-landform). This operation is beneficial to reduce human interference in distinction among original data.

The NN is a type of data-driven method. Compared with the logistic regression model, it overcomes most of the aforementioned limits. This model is suitable to solve a black-box problem, and has the ability to handle a large amount of information and to learn complex model functions from examples, i.e. by ‘training’, or by using sets of input-output data (Giustolisi & Savic). The main disadvantage is that the output results elicit increased randomness, and that the function cannot be expressed clearly. For NN model, the normalized categorical values are used as the input. The operation is beneficial to avoid the randomness of output result.

8.1.3. Samples selection and training

In this study, 20 groups with a fixed number of cells were randomly selected as a common training dataset, which provide the precondition for the two different susceptibility models.

For any group with a fixed number of cells, a set ratio between landslide and no-landslide cells near to 1:5 (149 landslide cells and 751 non-landslide cells for each group), the more samples is beneficial in improving the accuracy of established models.

For the logistic model, 10 groups with a fixed number of cells were used to train, and the average of regression analysis results from the 10 groups with a fixed number of cells was determined as the correlation coefficients. The method has been proven beneficial in reducing the uncertainty in the calculation of the model coefficients.

8.2. Conclusion

In general, earthquake-induced landslides are sudden serious and dangerous events, so a rapid and accurate susceptibility assessment is significant. In this study, the NN and the LR models were applied in Wenchuan county to produce accurate maps of earthquake-induced landslide susceptibility. Therefore, combining these models with GIS technology provides a relatively flexible and easy-to-use framework towards the spatial prediction of earthquake-induced landslides at mountainous regions.

For the NN model, more than 67.5% of field landslides points were located in approximately 6.7% coverage (high and very high susceptibility classes), but only approximately 14.3% field landslides points were located in approximately 87.8% coverage (low and very low susceptibility classes). The landslide point density increased by 177 times when the susceptibility class changed from very low to very high. For the LR model, more than 61.2% of the field landslides points were located in approximately 4.2% coverage (high and very high susceptibility classes), but approximately 21.2% field landslides points were located in larger 92.6% coverage (low and very low susceptibility classes). The landslide point density increased by 186 times when the susceptibility class varied from very low to very high. The validation results indicate that the NN model and the LR model are satisfactory for susceptibility mapping of earthquake-induced landslides at mountainous region, and the LR model (the AUC value is 0.941) has slightly better prediction than the NN model (the AUC value is 0.930) in this case. County is the basic unit of geo-hazard management in China, and bridges, valleys and slopes are the main landforms of Wenchuan county; these results can provide scientific basis for disaster prevention and mitigation at mountainous region in practice.
Acknowledgement

Authors mainly completed this research at Geo-hazards Engineering Research of Tufts University. We acknowledge China Scholarship Council and Chongqing University fund the corresponding author as a visiting scholar at Tufts University. We also acknowledge Chongqing Shutong Geotechnical Engineering Company offered partly basic data. Meanwhile, we thank Zhu PhD candidate of Geo-hazards Research at Tufts, for her research suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The National ‘Twelfth Five Year’ Plan of Science and Technology Support Project [grant number 2012BAJ22B06]; Venture & Innovation Support Program for Chongqing Overseas Returnees [grant number CX2017125]; the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry [grant number [2015]1098]; High Technology and Industrial Technology Development Special of Chongqing Development and Reform Committee in 2016 [grant number [2016]1270].

References

Arnone E, Francipane A, Noto LV, Scarbaci A, Loggia LG. 2014. Strategies investigation in using artificial neural network for landslide susceptibility mapping: application to a Sicilian catchment. J Hydroinform. 16(2):502–515.
Ayalew L, Yamagishi H. 2005. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology. 65:15–31.
Bai SB, Wang J, Thiebes B, Cheng C, Chang ZY. 2013. Susceptibility assessments of the Wenchuan earthquake-triggered landslides in Longnan using logistic regression. Environ Earth Sci. 71:731–743.
Bhandary NP, Dahal RK, Timilsina M, Yatabe M. 2013. Rainfall event-based landslide susceptibility zonation mapping. Nat Hazards. 69(1):365–388.
Borfecchia F, Canio GD, Cecco LD, Giocoli A, Grauso S, La Porta LL, Martin S, Pollino M, Roselli I, Zini A. 2016. Mapping the earthquake-induced landslide hazard around the main oil pipeline network of the Agri Valley (Basilicata, southern Italy) by means of two GIS-based modelling approaches. Nat Hazards. 81:759–777.
Brown D, Saito K, Liu M, Spence R, So E, Ramage M. 2012. The use of remotely sensed data and ground survey tools to assess damage and monitor early recovery following the 12.5.2008 Wenchuan earthquake in China. Bull Earthquake Eng. 10:741–764.
Caccavale M, Matano F, Sacchi M. 2017. An integrated approach to earthquake-induced landslide hazard zoning based on probabilistic seismic scenario for Phlegraean Islands (Ischia, Procida and Vivara), Italy. Geomorphology. 295:235–259.
Chen CY, Huang WL. 2013. Land use change and landslide characteristics analysis for community-based disaster mitigation. Environ Monit Assess. 185:1–15.
Chen H, Lin GW, Lu MH, Shih TY, Hong MJ, Wu SJ, Chuang B. 2011. Effects of topography, lithology, rainfall and earthquake on landslide and sediment discharge in mountain catchments of southeastern Taiwan. Geomorphology. 133:132–142.
Chousianitis K, Del Gaudio V, Sabatakakis N, Kavoura K, Drakatos G, Bathrellos G, Skilodimou H. 2016. Assessment of earthquake-induced landslide hazard in Greece: From Arias intensity to spatial distribution of slope resistance demand. Bull Seismol Soc Am. 106(1):174–188.
Dai FC, Lee CF, Tham LG, Ng KC, Shum WL. 2004. Logistic regression modelling of storm-induced shallow landsliding in time and space on natural terrain of lantau Island, Hong Kong. Bull Eng Geol Environ. 63(4):315–327.
Chousianitis K, Gaudio DV, Kelogeris I, Ganas A. 2014. Predistive model of Arias intensity and Newmark displacement for regional scale evaluation of earthquake-induced landslide hazard in Greece. Soil Dyn Earthquake Eng. 65:11–29.
Cuzzetti F, Carrara A, Cardinalli M, Reichenbach P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology. 31:181–216.
Dai FC, Xu C, Yao X, Xu L, Tu XB, Gong QM. 2011. Spatial distribution of landslides triggered by the 2008 Ms 8.0 Wenchuan earthquake, China. J Asian Earth Sci. 40(4):883–895.
Du WQ, Wang G. 2016. A one-step Newmark displacement model for probabilistic seismic slope displacement hazard analysis. Eng Geol. 205:12–23.
Gorum T, Fan XM, van Westen CJ, Huang BQ, Xu Q, Tang C, Wang GH. 2011. Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake. Geomorphology. 133:152–167.
Guo CW, Huang YD, Yao LK, Alradi H. 2017. Size and spatial distribution of landslides induced by the 2015 Gorkha earthquake in the Bhone Koshi river watershed. J Mt Sci. 14(10):1938–1950.

Huang JJ. 2000. Chi-Chi earthquake induced landslides in Tai wan. Earthquake Eng Eng Seismol. 2(2):25–33.

Huang RQ, Li WL. 2009. Analysis of the geo-hazards triggered by the 12 May 2008 Wenchuan earthquake, China. Bull Eng Geol Environ. 68(3):363–371.

Huang RQ, Wang YS, Pei XJ, Li YS, Li WL, Luo YH. 2013. Characteristics of co-seismic landslides triggered by the Lushan Ms7.0 earthquake on the 20th of April, Sichuan Province, China. J Southwest Jiaotong Univ. 48(4):581–589.

Hsieh SY, Lee CT. 2011. Empirical estimation of the Newmark displacement from the Arias intensity and critical acceleration. Eng Geol. 122:34–42.

Jibson RW. 1993. Predicting earthquake-induced landslide displacements using Newmark’s sliding block analysis. Transp Res Board Rec. 1411:9–17.

Keffer DK. 2000. Statistical analysis of an earthquake-induced landslide distribution - the 1989 Loma Prieta, California event. Eng Geol. 58(3-4):231–249.

Keffer DK. 2002. Investigating landslides caused by earthquake - a history review. Surv Geophy. 23:473–510.

Lee CT. 2013. Re-evaluation of factors controlling landslides triggered by the 1999 Chi-Chi earthquake. In: Ugai K, Yagi H, Wakai A, editors. Earthquake-induced landslides: Proceedings of the International Symposium on Earthquake-Induced Landslides, Kiryu, Japan, 2012. Berlin, Heidelberg: Springer Berlin Heidelberg; p. 213–224.

Lee S, Evangelista DG. 2006. Earthquake-induced landslide susceptibility mapping using an artificial neural network. Nat Hazards Earth Syst Sci. 6(5):687–695.

Lee S. 2005. Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. Int J Remote Sens. 26:1477–1491.

Lee S. 2005. Application of logistic model and its validation for landslide susceptibility mapping using GIS and remote sensing data. Int J Remote Sens. 26:1477–1491.

Li WL, Huang RQ, Tang C, Xu Q, van Westen C. 2013a. Co-seismic landslide inventory and susceptibility mapping in the 2008 Wenchuan earthquake disaster area, China. J Mt Sci. 10(3):339–354.

Li WL, Huang RQ, Xu Q, Tang C. 2013b. Rapid susceptibility mapping of co-seismic landslides triggered by the 2013 Lushan earthquake using the regression model developed for the 2008 Wenchuan earthquake. J Mt Sci. 10(5):699–715.

Liu YY, Fang JL, Chen XH, Chen YY. 2014. Evaluation of landslide susceptibility in Zigui Country based on certainty factor method. J Nat Disasters. 23(6):209–217.

Masahiro C, Hiroshi Y. 2004. Geological and geomorphological characteristics of landslides triggered by the 2004 Mid Niigata prefecture earthquake in Japan. Eng Geol. 82:202–221.

Meunier P, Hovius N, Haines AJ. 2007. Regional patterns of earthquake-triggered landslides and their relation to ground motion. Geophys Res Lett. 34(20):L20408.

Miles SB, Ho CL. 1999. Rigorous landslide hazard zonation using Newmark’s method and stochastic ground motion simulation. Soil Dynam Earthquake Eng. 18(4):305–323.

Mohammad RM, Shahryar S, Mohammad KJ. 2006. Landslides triggered by the Avaj, Iran earthquake of June 22, 2002. Eng Geol. 86:166–182.

Nefeslioegl U, Duman TY, Durmaz S. 2006. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). Geomorphology, 94(3-4):401–418.

Nowicki MA, Wald DJ, Hamburger MW, Hearne M, Thompson EM. 2014. Development of a globally applicable model for near real-time prediction of seismically induced landslides. Eng Geol. 173:54–65.

Ohlmacher GC, Davis JC. 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng Geol. 69:331–343.

Pardhan B, Lee S. 2010. Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. Environ Modle Softw. 25(6):747–759.

Pareek N, Pal S, Kaynia AM, Sharma ML. 2014. Empirical-based seismically induced slope displacements in a geographic information system environment: a case study. Georisk: Assess Manage Risk Eng Sys Geohazards. 8(4):258–268.

Paruelo JM, Garbulsky MF, Guerschman JP, Jobbagy EG. 2004. Two decades of normalized difference vegetation index changes in South America: identifying the imprint of global change. Int J Remote Sensing. 25(14):2793–2806.

Pradhan B, Lee S. 2010. Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. Environ Model Softw. 25:747–759.

Qi SW, Xu Q, Lan HX, Zhang B, Liu JY. 2010. Spatial distribution analysis of landslides triggered by 2008 Ms 10 Wenchuan earthquake, China. Eng Geol. 116:95–108.

Rajabi AM, Khamechhiyan M, Mahdavifar MR, DelGaudio V. 2010. Attenuation relation of Arias intensity for Zagros mountains region (Iran) soil dynamics and earthquake engineering, 30:110–118.

Rosser BJ, Carey JM. 2017. Comparison of landslide inventories from the 1994 Mw 6.8 Arthurs pass and 2015 Mw 6.0 Wilberforce earthquakes, Canterbury, New Zealand. Landslides. 14:1171–1180.
Sato HP, Sekiguchi T, Kojioi R, Suzuki Y, Iida M. 2005. Overlaying landslides distribution on the earthquake source, geological and topographical data: the Mid Niigata prefecture earthquake in 2004, Japan. Landslides. 2:143–152.

Sayg G, Rathje EM. 2008. Empirical predictive models for earthquake-induced sliding displacements of slopes. J Geotech Geoenviro Engr. 134:790–803.

Shinoda M, Miyata Y. 2017. Regional landslide susceptibility following the Mid NIIGATA prefecture earthquake in 2004 with NEWMARK's sliding block analysis. Landslides. 14:1887–1899.

Tang HM, Liu X, Hu XL, Griffiths DV. 2015. Evaluation of landslide mechanisms characterized by high-speed mass ejection and longrun-out based on events following the Wenchuan earthquake. Eng Geol. 26:12–24.

Tian YY, Xu C, Chen J, Zhou Q, Shen LL. 2017. Geometrical characteristics of earthquake-induced landslides and correlations with control factors: a case study of the 2013 Minxian, Gansu, China, Mw 5.9 event. Landslides. 14:1915–1927.

Tropeano G, Silvestri F, Ausilio E. 2017. An uncoupled procedure for performance assessment of slopes in seismic conditions. Bull Earthquake Eng. 15:3611–3637.

Umar Z, Pradhan B, Ahmad A, Jебur MN, Tehrany MS. 2014. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. Catena. 118:124–135.

Wang XY, Nie GZ, Ma MJ. 2011. Evaluation model of landslide hazards induced by the 2008 Wenchuan earthquake using strong motion data. Earthq Sci. 24:311–319.

Wang XY, Nie GZ, Wang DW. 2010. Relationships between ground motion parameters and landslides induced by Wenchuan earthquake. Earthq Sci. 23:233–242.

Wang Y, Song CZ, Lin QG, Li J. 2016. Occurrence probability assessment of earthquake-triggered landslides with Newmark displacement values and logistic regression: The Wenchuan earthquake, China. Geomorphology. 258:108–119.

Weiss A. 2001. Topographic position and landforms analysis. In: Environmental Systems Research Institute. International user Conference; Jul 9–13; San Diego, CA.

Wen HJ, Xie P, Xiao P, Hu DP. 2017. Rapid susceptibility mapping of earthquake-triggered slope geohazards in Lushan County by combining remote sensing with the AHP model developed for the Wenchuan earthquake. Bull Eng Geol Environ. 76:909–921.

Wen HJ. 2015. A susceptibility mapping model of earthquake triggered slope geohazards based on geo-spatial data in mountainous regions. Georisk. 9(1):25–36.

Wilson RC, Keefer DK. 1983. Dynamic analysis of a slope failure from the 6 August 1979 Coyote lake, California, earthquake. Bull Eng Geol Environ. 73(3):863–877.

Xu C, Shyu JBH, Xu XW. 2014. Landslides triggered by the 12 January 2010 Port-au-Prince, Haiti, Mw = 7.0 earthquake: visual interpretation, inventory compiling, and spatial distribution statistical analysis. Nat Hazards Earth Syst Sci. 14:1789–1818.

Xu C, Xu XW. 2014c. Statistical analysis of landslides caused by the Mw 6.9 Yushu, China, earthquake of 2010. Nat Hazards. 72:871–893.

Xu C, Xu XW, Dai FC, Saraf AK. 2012. Comparison of different models for susceptibility mapping of earthquake triggered landslides related with the 2008 Wenchuan earthquake in China. Comput Geosci. 46:317–329.

Xu C, Xu XW, Dai FC, Wu ZD, He HL, Shi F, Wu XY, Xu SN. 2013. Application of an incomplete landslide inventory, logistic regression model and its validation for landslide susceptibility mapping related to the May 12, 2008 Wenchuan earthquake of China. Nat Hazards. 68:883–900.

Xu C, Xu XW, Dai FC, Yao X. 2014a. Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis. Landslides. 11(3):441–461.

Xu C, Xu XW, Shen LL, Yao Q, Tan XB, Kang WJ, Ma SY, Wu XY, Cai JT, Gao MX, et al. 2016. Optimized volume models of earthquake-triggered landslides. Science Report. 6:1–9.

Xu C, Xu XW, Shyu JBH. 2015. Database and spatial distribution of landslides triggered by the Lushan, China Mw 6.6 earthquake of 20 April 2013. Geomorphology. 248:77–92.

Xu C, Xu XW, Shyu JBH, Zheng WJ, Min W. 2014b. Landslides triggered by the 22 July 2013 Minxian–Zhangxian, China, Mw 5.9 earthquake: inventory compiling and spatial distribution analysis. J Asian Earth Sci. 92:125–142.

Yesilinacar E, Topal T. 2005. Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Eng Geol. 79:251–266.

Yılmaz İ. 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat-Turkey). Comput Geosci. 35:1125–1138.

Yılmaz İ. 2010. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci. 61(4):821–836.

Yin Y, Wang F, Sun P. 2009. Landslide hazards triggered by the 2008 Wenchuan earthquake, Sichuan, China. Landslide. 6(2):139–152.
Yin ZQ, Zhao WJ, Qin XG. 2014. Distribution characteristics of geohazards induced by the Lushan earthquake and their comparisons with the Wenchuan earthquake. J Earth Sci. 25(5):912–923.
Zhang DX, Wang GH. 2007. Study of the 1920 Haiyuan earthquake-induced landslides in loess (China). Eng Geol. 94:76–88.
Zhang YS, Dong SW, Hou CT, Guo CB, Yao X, Li B, Du JJ, Zhang JG. 2013. Geohazards induced by the Lushan Ms7.0 earthquake in Sichuan Province, Southwest China: typical examples, types and distributional characteristics. Acta Geol Sin, Engl Ed. 87(3):646–657.
Zheng WJ, Yuan DY, He WG. 2013. Geometric pattern and active tectonics in southeastern Gansu province: discussion on seismogenic mechanism of the Minxian-Zhangxian Ms 6.6 earthquake on July 22, 2013. Chin J Geophys. 56:4058–4071.