GIS-based soil planar slide susceptibility mapping using logistic regression and neural networks: a typical red mudstone area in southwest China

Shuai Zhang\(^a\), Can Li\(^a\), Jingyu Peng\(^a\), Dalei Peng\(^{b,d}\), Qiang Xu\(^b\), Qun Zhang\(^c\) and Bate Bate\(^a\)

\(^a\)MOE Key Laboratory of Soft Soils and Geoenvironmental Engineering, Zhejiang University, Hangzhou, China; \(^b\)State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, China; \(^c\)Geo-environmental Monitoring Station of Sichuan Province, Chengdu, China; \(^d\)Faculty of Science and Technology, Technological and Higher Education Institute of Hong Kong, Hong Kong

ABSTRACT

Global warming increases the frequency and intensity of extreme rainfall, putting many areas at risk of landslides. Landslide susceptibility assessment is essential to understand the threats and to predict, prevent, and mitigate landslides. In this study, a soil landslide inventory was constructed based on satellite images, topological maps, and extensive field studies. Subsequently, eight different GIS layers, which were geomorphology, elevation, slope angle, slope aspect, slope structure, slope curvature, antecedent rainfall, and cumulative rainfall on 16 September, were produced as control factors of soil planar slides for the susceptibility mapping. Landslide susceptibility mapping was performed using two different methods, logistic regression model and backpropagation (BP) neural network. Landslide susceptibility in the study area is divided into four levels, which are high, moderate, low, and no susceptibility in both the logistic regression model and the BP neural network model. In both the two models, most of the observed soil planar slides were located in areas with high or moderate susceptibility. For the logistic regression model, total 605 soil planar slides locate in the area with high susceptibility, of which the area is 800.56 km\(^2\), accounting for 40.31% of the total area. Finally, the validation of two models was evaluated. The AUC value of the logistic regression model was 0.878 and the parameters of BP neural network has the correlation coefficient of 0.880, which shows the two models are both reliable and reasonable for predicting the spatial susceptibility of soil planar slides. According to field checks, the BP neural network model is verified to have more accurate spatial prediction performance than the logistic regression model.

ARTICLE HISTORY

Received 30 June 2020
Accepted 3 January 2021

KEYWORDS

Soil landslides; slope stability; spatial susceptibility; logistic regression model; artificial neural network

CONTACT

Bate Bate  
batebate@zju.edu.cn

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
1. Introduction

Southwest China is one of the most earthquake-prone areas. Several powerful earthquakes including, the Yiliang Earthquake on 7 September 2002 (Wang et al. 2013), the Wenchuan Earthquake on 12 May 2008 (Xu et al. 2016a), and the Lushan Earthquake on 20 April 2013, were recorded in Southwest China in just the past two decades. Strong earthquakes like these have further loosened what was already loose soil in the area, making the area prone to failure, especially after prolonged exposure to heavy rainfall and human engineering activities (Zhang et al. 2012, 2016b, 2017; Zhang and Zhang 2017). On 16 September 2011, a total of 1296 soil landslides occurred in Nanjiang County, Sichuan Province, China (Fig. 1). These landslides resulted in huge risks to the population and properties in the nearby downstream area (Fig. 2). Those landslides were triggered by heavy rainfall according to a geological survey conducted by the Nanjiang County Land Resources Bureau. Although many investigations on the triggering factors, geological conditions, or failure mechanisms have been conducted for the red-layered landslides that occurred in Nanjiang (Li et al. 2014; Zhang et al. 2014, 2015, 2016a, 2017), it is still necessary to characterize how soil landslides are distributed in the red mudstone area of the Nanjiang County. It is also important to analyze the susceptibility of this area to landslide, which can provide strong theoretical and technical support for the prevention, management, monitoring, and early warning of landslide hazards.

Figure 1. The study area: (a) The location of the study area; (b) The map of the Nanjiang county; (c) Digital elevation model of the study area. Source: Author.
Landslide susceptibility is the likelihood of a landslide to occur in a particular area based on the local terrain conditions (Brabb 1984). With developments in GIS data-processing techniques, research on the formation mechanisms of landslides, hazard evaluation, and susceptibility assessment of landslides have been conducted over the last two decades (e.g., Van Westen et al. 1997; Guzzetti et al. 1999; Lee et al. 2001, Wooten 2006; Yilmaz 2007; Kawabata and Bandibas 2009; Jadda et al. 2011; Poiraud 2014; Pourghasemi et al. 2014; Xu et al. 2016b; Pham et al. 2017; Singh et al. 2020; Wu et al. 2020). In particular, a lot of research on landslide susceptibility mapping used for susceptibility analysis has been conducted (e.g., Lee et al. 2001; Ohlmacher and Davis 2003; Ayalew et al. 2004; Lee et al., 2006; Yesilnacar and Topal 2005; Van Den Eeckhaut et al. 2006; Lee and Sambath 2006; Dahal et al. 2008; Kamp et al. 2008; Oh et al. 2009; Yilmaz 2009; Bai et al. 2010; Akgun 2012; Felicísimo et al. 2013; Aniya 1985; Su et al. 2015; Fei et al. 2017; Gigović et al. 2019; Bălțeanu et al., 2020). In the authors’ process of susceptibility mapping, many models were used for calculating the value of susceptibility, which can be qualitative or quantitative. In studies using qualitative methods, landslide hazard levels in a study area were illustrated using descriptive terms; on the other hand, studies using quantitative methods evaluated the numerical probabilities of landslide occurrence in a specific hazard zone. However, geotechnical models only appear to be useful to a limited extent.
considering the difficulty in collecting data for necessary physical variables in large regions. Hence, statistical approaches are currently pursued to assess landslide hazards (Yesilnacar and Topal 2005). For example, Ohlmacher and Davis (2003) adopted a multiple logistic regression method to create a landslide hazard map for landslides on hilly terrains along the Kansas and Missouri rivers in northeastern Kansas. The results indicated that slopes are the most important variable for estimating landslide hazards in the study area. A statistical multivariate method, i.e., rare events logistic regression, is evaluated by Van Den Eeckhaut et al. (2006) to create a landslide susceptibility map for a 200 km² study area of the Flemish Ardennes (Belgium). In a study by Yalcin (2008), the analytical hierarchy process (AHP), the statistical index ($W_i$), and weighting factor ($W_j$) methods were used to produce and later compare three susceptibility maps. The results showed that the AHP method gave a more realistic picture of the actual distribution of landslide susceptibility than the $W_i$ and $W_j$ methods. Bai et al. (2010) used a logistic regression based on GIS data to produce a detailed susceptibility map of one of the most landslide-prone areas in China, the Zhongxing-Shizhu segment in the Three Gorges Reservoir region. In recent years, more and more machine learning methods have been used for landslide susceptibility mapping, such as artificial neural networks (Møller, 1993; Lee and Min 2001; Lee et al. 2001; Zeng-Wang 2001; Yesilnacar and Topal 2005; Nawi et al. 2006; Pham et al. 2017; Le et al. 2019a, 2019b; Shariati et al. 2019; and Lv et al. 2020), support vector machine (Guo et al. 2005; Yao and Dai 2006; Basak et al. 2007; and Dou et al. 2015), decision Tree (Chu et al. 2009; Saito et al. 2009; Nefeslioglu et al. 2010; and Tien Bui et al. 2012); random forest (Dou et al. 2019), neuron fuzzy (Pradhan 2013) and so on. Also, comparison between these models have been made to figure out the most accurate model for landslide susceptibility mapping. For example, Chen et al. (2018) assess and compare four advanced machine learning techniques, namely the Bayes’ net (BN), radical basis function (RBF) classifier, logistic model tree (LMT), and random forest (RF) models, for landslide susceptibility modelling in Chongren County, China. Out of the tested models, the RF model was verified to have the highest sensitivity.

It can be found that past research has mainly focused on: (1) choosing the parameters used for landslide susceptibility—in order to achieve a more accurate assessment, the evaluation parameters with the highest correlation with landslides would be selected by researchers; (2) modeling landslide susceptibility—a large number of susceptibility evaluation models considering different control factors and mathematical methods have been established; (3) verifying evaluation results. Although many researches have been conducted on the triggering factors, geological conditions, and failure mechanisms for red mudstone landslides in the Nanjiang County, studies on the distribution characteristics of red mudstone landslides and their susceptibility are still limited.

This paper aims to (1) construct a soil landslide inventory map that incorporates the spatial distribution characteristics of soil planar slides in the red mudstone area of Nanjiang; (2) establish an evaluation parameter system of soil landslide susceptibility mapping, based on the correlation between each influencing factor of red mudstone landslides; and (3) determine the appropriate evaluation model and conduct a susceptibility mapping of soil landslides in the study area.
2. Description of the study area

The study area is located in Nanjiang, Sichuan Province of China, which lies between the longitudes of 106°27' to 107°28'E and latitudes of 31°53' to 32°45'N (Figs. 1a and b). In general, according to the topographical map (Fig. 1c), the altitude of the south and middle sections of the study area ranges from 500 to 1000 m and increases to above 1000 m in the north of the area. This area is characterized by widely distributed cuesta (Fig. 3a) with gently inclined rock layers and quasi cuesta, and most of the hillslope area has been explored as cultivated land and paddy field in a ladder shape (Zhang et al. 2017). The slope angle of the rock layers ranged from 10° to 30°. As shown in Fig. 3a, tectonic structure is developed in the northern mountainous area of Nanjiang County, which is not active in the southern area. The main fault structure in the county is the Lanchaiba Fault. The main fold structures are the Zhongzi Mountain Synclinore, the Liang Anticlinorium, Guomatan syncline, the Shatan Anticline, the Xinhua Syncline and the Longfengchang Anticline. On the whole, the tectonic movement in Nanjiang County is mainly manifested as fold tectonic movement. And the fault structure in Nanjiang County is underdeveloped, only some small faults are developed. Lithology is rather complicated in the study area, where the main lithology formations are loose soil, shale, mudstone, dolomite, and diorite (Fig. 3b). Mesozoic Jurassic and Cretaceous strata are the two basic underlying rock layers that consist of the most distributed red mudstone layers in this area. Most of the exposed red mudstone is deeply fractured and highly weathered. During the rainy seasons, these materials are easily weathered and softened. Due to the low strength of the red layer materials, it is easy to collapse under the changing climate
conditions (Zhang et al., 2016). Hence, rainfall-induced landslides occur frequently in the study area. During the extreme rain event on 16 September 2011, over 1700 landslides were triggered in Nanjiang County. Over 90% of soil landslides in Nanjiang were concentrated in the red mudstone area (Fig. 3b). The uneven distribution of soil landslides can contribute to the poor mechanical properties of red mudstone, which is characterized by low strength, strong water sensitivity, and well-developed structural plane. Such soft rock can weather rapidly under the effect of dry-wet cycles. Weak layers between the upper soil and lower bedrock can develop due to water infiltration. Thus, the completely or highly weathered red mudstone can easily become the source of a potential landslide. To perform more accurate modeling, the red mudstone area in the south of Nanjiang has been chosen as the target study area of this research.

For the hydrological conditions, Nanjiang has a subtropical humid monsoon climate, and the average annual precipitation from 2000 to 2011 was approximately 1074 mm, which was slightly less than the annual rainfall of 1149.7 mm in 2011. Rainfall is mainly concentrated between May and September, accounting for nearly 79.2% of the annual rainfall. Figure 4 shows the rainfall contours of the red mudstone area of the Nanjiang County in 2011.

Figure 4. Rainfall distribution in red mudstone area of the Nanjiang County in 2011. Source: Author.
area for Nanjiang in 2011. The maximum and minimum rainfall took place in the northeast and southeast, respectively. The upper bound and lower bound were around 2,000 mm and 1,100 mm. Heavy rainfall occurred in the study area from 6 to 15 September 2011 with an accumulated rainfall of 268.1 mm. From 8:00 am on 16 September to 8:00 am on 18 September 2011, a daily rainfall of 250.4 mm and 179.1 mm were recorded, respectively. The total accumulated rainfall within the 12 days was 698.6 mm (Fig. 5). Compared with the historical monthly average rainfall of 182.6 mm in September, the extreme rain event provided likely prerequisites for rainwater infiltration and softening effects on the sliding surfaces of the potential landslides.

3 Spatial database construction

3.1. Landslide database

This study began with the construction of a soil landslide inventory map based on three basic datasets, including Landsat TM5 satellite images with a spatial resolution of 30 m, a 1:60000-scale topological map, and extensive field studies. According to the landslide classification by Varnes (1978), Hungr et al. (2014), these soil landslides contains 1296 soil planar slides and 455 soil rotational slides. As shown in Fig. 6 and Table 1, based on the field investigation, planar slides are often of small magnitude with a stepped sliding surface in 1–5 m deep from the ground. And they usually occur in a gentle terrain, mainly within 10–30°. On the contrary, rotational slides are of large magnitude with circular sliding surface in the depth of 10–15 m. Planar slides mainly occurred along the gentle dipped bedding surface between the shallow overlying soil and bedrock. However, rotational slides always occur along the circular sliding surface due to gravity or rainfall. Since the most soil slides in the study area is planar slides, rotational slides occur most frequently in fairly homogeneous materials, the failure principle of them is usually simple, and hence easier to be prevented from
causing damage to human lives and properties. So, only the soil planar slides are studied in this research, and 1146 soil planar slides in the red mudstone area in this database have been chosen for this study’s target, and the detailed distribution characteristics of the soil planar slides were studied.

Table 1. Classification of the soil landslides.

| Soil landslides | Soil planar slides | Soil rotational slides |
|-----------------|--------------------|-----------------------|
| Characteristics | 1. Small magnitude | 1. Large magnitude     |
|                 | 2. Depth of sliding surface, mainly in 1–5 m | 2. Depth of sliding surface, mainly in 10–15 m |
|                 | 3. Gentle terrain, mainly within 10–30° | 3. Gentle terrain, mainly within 30–40° |
|                 | 4. Stepped slope surface | 4. Circular slope surface |
| Failure modes   | Failures mainly occurred along the gentle dipped bedding surface between the shallow overlaying soil and bedrock | Failures always occur along the circular sliding surface due to gravity or rainfall |

Figure 6. Failure modes of (a) planar slides (b) rotational slides. Source: Author.
3.2. Landslide control factors

The occurrence of a landslide is controlled by multiple factors. Identification and mapping of a suitable set of parameters in a relationship with slope failure require prior knowledge of the main causes of landslides (Guzzetti et al. 1999). Based on the study of mechanics of soil landslides in the study area and data collected, eight factors are selected as control factors used for landslide susceptibility. Statistical graphs of landslide control factors in the study area are shown in Figs. 7 and 8, to intuitively illustrate the characteristics of landslides. Eight GIS layers, namely geomorphology, elevation, slope angle, slope aspect, slope structure, slope curvature, antecedent rainfall, and cumulative rainfall on 16 September, were then produced as the control factors of soil planar slides for susceptibility mapping (Fig. 9).

The geomorphology factor is shown in Fig. 9a, which is derived from the interpretation of the satellite images and then checked on the field. The soil planar slides in the red mudstone area of Nanjiang are mainly distributed in the Quasi-Cuesta area (515 landslides, with a density of 0.56 landslides/km²), the Mesa area (471 landslides, density 0.53 landslides/km²), and the Cuesta (160 landslides, density 0.77 landslides/km²). In the above three different geomorphic units, one landslide occurred almost every 2 km². It is found that the special terrain in the study area, i.e., monoclinic structure, is one of the control factors for the occurrence of soil planar slide.

The topological attributes, such as elevation, slope angle, slope aspect, slope structure, and curvature, are derived from the digital elevation method (DEM) with a resolution of 30 m. These are generated from a triangulated irregular network (TIN) model. As shown in Fig. 8b, the potential soil planar slides in the study area are mainly located between elevations of 500 m and 1000 m. The soil planar slides in the red layer area of Nanjiang are mainly distributed within the elevation range of 230–1500 m, and especially between 500–1000 m. According to the statistics, the elevation of the soil planar slides follows a normal distribution. The landslide densities in elevations of 230–500 m, 500–1000 m and greater than 1000 m are 0.63 landslides/km², 0.68 landslides/km², and 0.20 landslide/km², respectively (Fig. 9b). It can be found that not only does the slope structure at the elevation of 500–1000 m is mainly monoclinic, making it easy for a landslide to occur, but it is also due to intense human activities, including cropland irrigation and engineering excavations. These human activities are concentrated mainly within the same range of elevation, which further reduces the stability of the slope. As shown in Fig. 8c, the density points of landslides with a slope angle smaller than 10°, between 10°–30°, and larger than 30° are 0.542, 0.664, and 0.346 landslides/km², respectively. Soil planar slides are normally distributed in the monoclinic region within the range of 10°–30°. Rainwater can gather and penetrate the surface of the gentle slopes easily. Moreover, the dip-direction of the bedding slope is consistent with the inclination direction, which leads to the occurrence of bedding failures along with the soil-bedrock interface. Local human activities and land use within the affected area can accelerate the occurrence of landslides. Furthermore, the soil planar slides are evenly distributed at different slope aspects in the red mudstone area of Nanjiang, where landslides in the directions of 90°–180° and 270°–360° are the most accounted for (Fig. 9d). The majority of the soil planar slides are distributed in dip slopes with a density of 0.5926 landslides/km²,
Figure 7. Distribution of the red-layered soil planar slides with (a) geomorphology; (b) elevation; (c) slope angle; (d) slope aspect; (e) structure. Source: Author.
followed by oblique slopes with a density of 0.5811 landslides/km², and anti-dip slopes of 0.5515 landslides/km² (Fig. 9e). The dip slopes are prone to landslides due to the underlying dipping strata. Large sheets of rock tend to slide down the dip slopes, whereas for anti-dip slopes, the effect is the opposite.

Rainfall is a major triggering factor of soil planar slides in Nanjiang. As shown in Fig. 8a, according to statistical analyses, the 14-day antecedent rainfall before the extreme rainfall event on 16 September is mainly between 303–723 mm, following a normal distribution. The distribution curve peaks around 600 mm with the highest frequency of landslide. The density points (per square kilometer) of landslides for the 14-days of rainfall of 303–400 mm, 400–500 mm, and 500–723 mm are 0.35, 0.40, and 0.61 landslides/km², respectively (Fig. 9b). The 48-hour cumulative rainfall recorded is 204–735 mm, also with a normal distribution (Fig. 9c). The density point of landslides for the 48-hours of cumulative rainfall for 204–300 mm, 300–400 mm, 400–500 mm, and 500–735 mm is 0.49, 0.55, 0.69, and 0.99 landslides/km².
respectively (Fig. 9d). Like the antecedent rainfall, the density of soil planar slides increases with the increasing cumulative rainfall.

4. Landslide susceptibility mapping based on GIS platform

4.1. Logistic regression model

The logistic regression model can form a regression relationship between a dependent variable and one or more independent variables, in which the results of the dependent variable is 0 or 1 (Menard 1995; Atkinson and Massari 1998). A good nonlinear logistic regression model has the advantage that the variables may be either discrete or continuous and may not follow a certain statistical distribution. Therefore, the logistic regression model is an efficient method to establish a relationship between the occurrence of landslide and the multiple control factors (Ayalew and Yamagish 2005; Bai et al. 2010; Wang et al. 2015).

In the logistic regression model, the independent variables are the control factors, and the dependent variable is the occurrence of landslides, where 0 means landslides do not occur, and 1 represents landslides do occur. The formulas of the logistic regression model are as follows:

\[
P = \frac{1}{1 + e^{-z}}
\]

\[
z = a_0 + a_{1j} \times X_{1j} + a_{2j} \times X_{2j} + \cdots + X_{nj}
\]

in which, \(P\) is the probability of occurrence of a landslide, its value range is (0, 1); \(n\) is the number of independent variables; \(a_1, a_2, \ldots, a_n\) is a logistic regression coefficient; and \(a_0\) is the intercept of this model which equals logarithmic change value of the probability of occurrence, divided by the probability of not occurring when a factor changes by one unit.

According to data from DEM and interpretation of satellite images, eight different GIS layers have been produced for the control factors for soil planar slides, including geomorphology, elevation, slope angle, slope aspect, slope structure, slope curvature, antecedent rainfall, and 48-hour cumulative rainfall (Fig. 9). However, some of those variables are nominal variables that should be transformed into numeric variables when used in a logistic regression model. It is significant to find an appropriate approach to quantify those nominal variables. An approach using landslide density is applied in this study where control factors are quantified by the ratios landslide points to the area occupied by each secondary control factor. This ratio may not only reflect the distribution of the disaster per unit area, but also the frequency of occurrence. The equation is as follows:

\[
X_{ij} = \frac{N_{ij}}{S_{ij}}
\]

where \(i = 1, 2, \ldots, n\) is the number of the primary control factor; \(j = 1, 2, \ldots, m\) is the number of the secondary control factor; \(N_{ij}\) is the number of disaster points; \(S_{ij}\) is the
Figure 9. Control factors in producing the landslide susceptibility map (a) geomorphology; (b) elevation; (c) slope angle; (d) slope aspect; (e) slope structure; (f) slope curvature; (g) antecedent rainfall; (h) 48-hour cumulative rainfall of the 9.16 rain event. Source: Author.
area of each secondary factor occupied; $X_{ij}$ is the weight of the control factor, which can be normalized to $I_{ij}$ and summarized in Table 2.

After producing a comprehensive evaluation layer superimposed by all the control factor layers in ArcGIS, the extracted data of each unit in this layer were input into SPSS to calculate the correlation of landslides to each factor, during which the “LR” algorithm is selected. In this process, when selecting an area where no landslide occurred, a buffer zone with a radius of 100 m was created for each landslide point to make the selected area more representative and convincing, based on the study of Suzen and Doyuran (2004). The layer contains circles of the same number as landslide points. Subsequently, the landslide point buffer layer and the comprehensive evaluation layer are analyzed to determine the units in the comprehensive evaluation layer, which do not intersect with any landslide point buffer circles. The value of the non-intersecting units is 0, indicating that there no landslide disaster occurred in the unit. Then the stepwise calculation is carried out within SPSS to eliminate variables that have no or little contribution to the occurrence of landslides. The significance level is set as 0.05. The results indicate that the slope aspect, geological structure, and pattern of the slope’s surface contributed little to the occurrence of soil planar slides, however, the slope angle, the 14-day antecedent rainfall, elevation, and the 48-hour cumulative rainfall had a greater contribution.

Based on the results of the logistic regression, the following logistic regression model is applied in this study to map the study area’s susceptibility to landslides in SPSS:

### Table 2. Parameters in the susceptibility assessment for the red-layered soil planar slides.

| Primary control factors | No. | Secondary control factors | $N_{ij}$ | $S_{ij}$ | $X_{ij}$ | $I_{ij}$ |
|-------------------------|-----|----------------------------|---------|---------|---------|---------|
| Elevation (m)           | 1   | 0–500                      | 88      | 138.78  | 0.63    | 0.55    |
|                         | 2   | 500–1000                   | 975     | 1434.95 | 0.68    | 0.61    |
|                         | 3   | >1000                      | 83      | 411.92  | 0.20    | 0       |
| Slope angle (°)         | 1   | 0–20                       | 727     | 1105.86 | 0.66    | 0.58    |
|                         | 2   | 20–40                      | 404     | 793.73  | 0.51    | 0.39    |
|                         | 3   | 40–60                      | 15      | 53.81   | 0.28    | 0.10    |
| Slope aspect (°)        | 1   | 0–135                      | 438     | 692.01  | 0.63    | 0.55    |
|                         | 2   | 135–180                    | 165     | 235.61  | 0.70    | 0.63    |
|                         | 3   | 180–225                    | 157     | 259.98  | 0.60    | 0.51    |
|                         | 4   | 225–360                    | 386     | 779.82  | 0.49    | 0.37    |
| Geomorphology           | 1   | Table mountain             | 471     | 881.88  | 0.53    | 0.42    |
|                         | 2   | Quasi-cuesta               | 515     | 913     | 0.56    | 0.46    |
|                         | 3   | Cuesta                     | 160     | 207.38  | 0.77    | 0.73    |
| Slope structure         | 1   | Dip slope                  | 240     | 413.01  | 0.58    | 0.48    |
|                         | 2   | Oblique slope              | 380     | 641.25  | 0.59    | 0.50    |
|                         | 3   | Cross-dip slope            | 405     | 712.57  | 0.57    | 0.47    |
|                         | 4   | Anti-dip slope             | 121     | 219.39  | 0.55    | 0.45    |
| Slope curvature         | 1   | Concave                    | 137     | 296.52  | 0.46    | 0.33    |
|                         | 2   | Straight                   | 876     | 1413.67 | 0.62    | 0.53    |
|                         | 3   | Convex                     | 133     | 276.75  | 0.48    | 0.36    |
| Antecedent rainfall (mm)| 1   | 303–400                    | 24      | 67.74   | 0.35    | 0.19    |
|                         | 2   | 400–500                    | 104     | 257.38  | 0.40    | 0.26    |
|                         | 3   | 500–723                    | 1018    | 1675.33 | 0.61    | 0.52    |
| Cumulative rainfall (mm)| 1   | 204–300                    | 65      | 133.01  | 0.49    | 0.37    |
|                         | 2   | 300–400                    | 881     | 1593.82 | 0.55    | 0.45    |
|                         | 3   | 400–500                    | 160     | 231.96  | 0.69    | 0.62    |
|                         | 4   | 500–735                    | 40      | 40.52   | 0.99    | 1       |
where $X_1j$, $X_2j$, $X_3j$ and $X_4j$ represent the weight of the slope angle, the 14-day antecedent rainfall, the elevation, and the 48-hour cumulative rainfall, respectively.

The slope angle, the 14-day antecedent rainfall, elevation, and the 48-hour cumulative rainfall have logistic regression coefficients of 5.93, 4.29, 3.48, 1.66, receptively. According to the value of the coefficients, the slope angle has the greatest impact on landslides, followed by the 14-day antecedent rainfall, elevation, and finally, the 48-hour cumulative rainfall.

Using the factor weight $X_{ij}$ and the established regression model, the weight value of a soil planar slide to occur within each unit can be obtained. Weight values of the resulting susceptibility map are treated as relative and not absolute values (Blahut et al., 2010). Based on the relative probability value obtained by overlapping the layers of each causing factor, the units are ordered from the highest ones to the lowest and then reclassified into classes. A landslide susceptibility map of soil planar slides in the red mudstone area can therefore be developed (Fig. 10).

Four susceptibility levels, i.e., high, moderate, low, and none were classified. The area with high susceptibility is 800.56 km$^2$, accounting for 40.31% of the total area and containing 605 soil planar slides distributed between the town of Ganchang in
the north of Nanjiang, and the towns of Shahe, Changchi, and Xialiang in the south of Nanjiang. The area with moderate susceptibility is 808.44 km$^2$, accounting for 40.71% of the total area and containing 451 soil planar slides distributed mainly in the towns of Zhengzhi and Nanjiang. The low susceptibility regions were mainly found in Dahe town, which accounts for 16.8% of the total area with an affected area of 333.67 km$^2$. A total of 83 soil planar slides were observed. The area with no susceptibility is 43.10 km$^2$, accounting for 2.17% of the total area and is also mainly distributed in Dahe town in the red mudstone area of eastern Nanjiang.

4.2. Backpropagation neural network

Landslide control factors constitute a dynamic system with strong nonlinearity and dimensionality. For nonlinear dynamic problems like this, an artificial neural network is suitable, whereas traditional regression techniques often fail to produce accurate approximation as the dimensionality and/or nonlinearity of the problem increases. A three-layer backpropagation (BP) descent algorithm was used for assessing the susceptibility of soil planar slides in this paper, which is a well-recognized procedure for training a neural network (Multi-Layer Perceptrons—MLPs topology) (Yesilnacar and Topal 2005).

The principle of the BP neural network is using gradient descent for point(s) with minimum error to search for a performance surface (error as a function of neural network weights) (Yesilnacar and Topal 2005). The training of the BP neural network model is mainly composed of two processes, forward and backward propagation. First, the information is input from the input layer and passed to the hidden layer. After being processed layer by layer within the hidden layers, the processed information is finally transmitted to the output layer. This process is called forward propagation. If results are not as expected, an error signal will be issued so that weights of neurons in each layer are modified until the smallest error signal appears; this is the backpropagation process. When the weights are modified to the optimal, the network training is completed. The BP neural network model used in this paper is shown in Fig. 11.

Initially, $X = (x_1, x_2, ..., x_l)^T$ is the input vector, $X_j$ is the output value of the hidden layer, and $Y_k$ is the output value of the output layer. The input value of the neuron at the hidden layer is equal to the output value at the input layer. Supposing $l$ is the number of neurons in the input layer, $m$ is the number of neurons in the hidden layer, and $n$ is the number in the output layer, so $X_j$ can be calculated by the following formula:

$$X_j = f_1 \left( \sum_{i=1}^{l} \omega_{ij} x_i - Q_j \right), \quad 1 < j < m$$

$$Y_k = f_2 \left( \sum_{j=1}^{m} \omega_{jk} X_j - Q_k \right), \quad 1 < k < n$$

where $\omega_{ij}$ is the connection weight between node $i$ (input layer) and node $j$ (hidden layer), $\omega_{jk}$ is the connection weight between node $j$ (hidden layer) and node $k$ (output
layer), $Q_j$ is the threshold of node $j$, and $Q_k$ is the threshold of node $k$. Using the lowering direction of Newton’s method to modify the weight. The weight is adjusted as the follows:

$$\Delta \omega = (J^T J + uI)^{-1} J^T E$$  \hspace{1cm} (7)

In which, $J$ represents the Jacobian matrix obtained by differentiating the error with the weight, $I$ represent the initial iteration matrix, and $E$ represents the error vector. When the value of $u$ is large enough, the error correction formula can be further expressed as:

$$\omega_{ij}(t + 1) = \omega_{ij}(t) + \Delta \omega$$  \hspace{1cm} (8)

Using a recursive algorithm can back-propagate the error from the output layer and adjust the weight as shown in the following formula:

$$E = \frac{1}{2} \sum_{K=1}^{n} (d_k - y_k)^2$$  \hspace{1cm} (9)

where $d_k$ is the actual value. The error $E$ of the output will eventually be less than or equal to the expected error $\varepsilon$. At this point, the training process of the network will have been completed.
It is necessary to evaluate the predictive capability of landslide control factors in the BP neural network model to acquire more accurate landslide susceptibility modeling, because some factors may have a negative effect on the generated models (Tien Bui et al. 2016). Thus, in this study, three different groups of parameters are tested in MATLAB to obtain the predictive capability of each group of parameters. For each group of parameters, five hundred typical soil planar slide points and 500 non-slide points in the red layer area of Nanjiang County are extracted using ArcGIS. The 500 slides were randomly divided into two parts, including 350 slides (70%) used for training the BP neural network model and the rest 150 slides (30%) used for validating the built models. The non-slide points were also randomly split into training (70%, 350 non-slides) and validation (30%, 150 non-slides). Then, a non-linear model is established between the parameters and the slope stability, and eventually, the group of evaluation parameters with the most accurate spatial prediction result is selected as control factors of the occurrence of a soil planar slide and used to predict the landslide susceptibility.

The first group of parameters consists of all the eight influencing factors, which are geomorphology, elevation, slope angle, slope aspect, slope structure, slope curvature, antecedent rainfall, and the 48-hour cumulative rainfall. Since there is a large difference in the order of magnitude between the input data and the output data, it will lead to a large error in the prediction result of the BP neural network, which is called ‘overfitting’. In order to avoid the error caused by the magnitude difference, the values of each parameter are normalized before being inputted. The softmax normalization method is used in this study to convert the input data into numbers between 0 and 1. This can eliminate the difference between the orders of magnitude. Examples of some of the input data are shown in Table 3. It can be seen that the normalized data samples are all distributed in intervals of 0 and 1, and there is no abnormal phenomenon that any of the data reaches the two ends, which achieves the expected processing effect. Moreover, for the data of slope stability in Table 3, 1 indicates that the slope is unstable and a landslide occurs, while 0 indicates that the slope is in a stable state.

The MATLAB software is used to complete the calculation of the BP neural network method. The fitted curve of the predicted value and the actual value is shown in Fig. 12(a). The correlation coefficient \( R \) between the predicted value and the actual value is 0.862, indicating the prediction result is highly accurate and that the BP network model has good reliability and can be applied to the spatial prediction of soil planar slide disasters in Nanjiang.

Table 4 shows the calculation results of the weights of the eight influencing factors. Here we can see the parameters of the weights in descending order: slope angle, antecedent rainfall, elevation, the 48-hour cumulative rainfall, slope structure, slope curvature, geomorphology, and slope aspect. The results show that the soil planar slides in the red layer area in the Nanjiang County are greatly affected by slope angle, rainfall, elevation, and geomorphology, and that the slope aspect does not contribute much to the occurrence of the landslide.

The geomorphology and slope aspect with the smallest weight values are removed in the second group of parameters, and the slope structure and slope curvature are
further excluded in the third group of parameters. The fitting curves between the predicted value and the actual value of the stability coefficient of the second and third groups of parameters are shown in Figs. 12b and c, respectively. The calculated weights of the three groups of parameters using the BP neural network model are arranged in the same order. Furthermore, the correlation coefficients (R) between the predicted and actual values of the BP network model, considering there was three different groups of evaluating parameters: 0.862, 0.880, and 0.839, respectively. The prediction results are all good but the correlation coefficient of the second group is the highest, which illustrates that the evaluation parameters selected in the second group are the most reliable and reasonable. Six parameters, which are elevation, slope angle, slope structure, slope curvature, and the 48-hour cumulative rainfall are found to be the key factors that affect the stability of slopes. Using the most reliable and reasonable evaluation parameters, the superposition of layers in ArcGIS is carried out to assess the landslide susceptibility. According to the probability results of a landslide to occur, the units in ArcGIS are also ordered from the highest ones to the lowest ones and then reclassified into classes. A landslide susceptibility map of soil planar slides in the red mudstone area can be developed (Fig. 13). The susceptibility of soil planar slide disasters in the red layer area of the Nanjiang County can be also divided into 4 levels (high susceptibility, moderate susceptibility, low susceptibility, and no susceptibility).

### Table 3. Examples of input data in the BP neural network.

| No. | Control factors | Elevation | Slope angle | Slope aspect | Geomorphology | Slope structure | Slope curvature | Antecedent rainfall | Cumulative rainfall | Stability |
|-----|-----------------|-----------|-------------|-------------|---------------|----------------|-----------------|---------------------|---------------------|-----------|
| 1   |                 | 0.6591    | 0.2393      | 0.4647      | 0.7143        | 0.6667         | 0.5000          | 0.2951              | 0.0812              | 1.0000    |
| 2   |                 | 0.3449    | 0.2910      | 0.4120      | 0.7143        | 0.3333         | 0.5000          | 0.3171              | 0.0876              | 1.0000    |
| 3   |                 | 0.2514    | 0.1861      | 0.5373      | 0.7143        | 0.3333         | 0.5000          | 0.3279              | 0.0705              | 1.0000    |
| 4   |                 | 0.3700    | 0.3687      | 0.6462      | 0.7143        | 0.6667         | 0.5000          | 0.3336              | 0.0876              | 1.0000    |
| 5   |                 | 0.4886    | 0.2807      | 0.6015      | 0.7143        | 0.6667         | 0.5000          | 0.3370              | 0.1026              | 1.0000    |
| 6   |                 | 0.4006    | 0.2836      | 0.6407      | 0.7143        | 0.6667         | 0.5000          | 0.3352              | 0.0962              | 1.0000    |
| 7   |                 | 0.6096    | 0.3858      | 0.2620      | 0.7143        | 0.3333         | 1.0000          | 0.3425              | 0.1303              | 1.0000    |
| 8   |                 | 0.8865    | 0.2801      | 0.5177      | 0.5000        | 0.0000         | 0.5000          | 0.6041              | 0.5569              | 0.0000    |
| 9   |                 | 0.7574    | 0.2158      | 0.4132      | 0.5000        | 0.3333         | 0.5000          | 0.6541              | 0.5529              | 0.0000    |
| 10  |                 | 0.6897    | 0.6649      | 0.0989      | 0.5000        | 0.3333         | 0.5000          | 0.4926              | 0.5020              | 0.0000    |

5. Discussion

The susceptibility map can have the potential to predict the spatial characteristics of future landslides. Without some form of validation, the prediction model and images are useless and have little scientific significance (Chung and Fabbri, 2003). According to Can et al. (2005), two rules for a spatially effective landslide susceptibility map were mentioned: i) most landslides should occur on the areas having high susceptibility values, and ii) each susceptibility should be included in the susceptibility map to be produced because the entire area can be classified in subclasses having different susceptibility categories. As the results have shown, most of the observed soil planar slides in this detailed field investigation are located in the areas with high or moderate susceptibility, in both the logistic regression and BP neural network models. In
Figure 12. The fitted curve of the predictive value and the actual value for (a) first, (b) second, and (c) third group of parameters. Source: Author.
The two susceptibility maps produced, the four susceptibility regions, high, moderate, low, and no susceptibility regions, both covered the study area.

Besides, for the logistic regression model, the ROC curve is a useful method to represent the quality of deterministic and probabilistic detections and the forecast systems (Swets 1988). Receiver operating characteristic curve analysis (ROC analysis),
provides tools to differentiate two classes, established through a diagnostic test, in an optimal manner. This analysis is based on the final distribution of a classification method that differentiates between correct and failed predictions according to a confusion matrix composed of four indexes, true positive (TP), false positive (FP), true negative (TN), and false negative (FN). For landslide susceptibility, TP and FN are landslide points which are located successfully and wrongly, respectively. On the contrary, TN and FP are the stable points which are located successfully and wrongly, respectively. Two statistics are calculated, namely TPR (true positive rate) and FPR (false positive rate) for a set of threshold or cut-off values (Vakhshoori and Zare 2018; Cantarino et al. 2019). The quality of a forecast system is characterized by the area under the ROC curve (AUC). The AUC varying from 0.5 to 1.0 shows the system’s ability to correctly predict the occurrence or non-occurrence of the event. The AUC value of the model was calculated to validate the performance of the logistic regression model. The ROC curve of this model is drawn and the AUC value is calculated as 0.878, which shows the model has a high accuracy to predict the spatial susceptibility of landslides (Fig. 14).

For the BP neural network, the predicted results of three different groups of parameters are all good, but the correlation coefficient of the second group is the highest. This illustrates that six parameters, elevation, slope angle, slope structure, slope curvature, and 48-hour cumulative rainfall, are key factors that control the stability of the slopes. The correlation coefficient (R) of this model was 0.880, which shows it is reliable and reasonable for predicting spatial susceptibility of landslide.

Moreover, the performance of the two models was compared to select the optimized model for soil planar slide susceptibility assessment of the study area. Field checks had been conducted and showed that the accurate rate of logistic regression model is 0.821 which is lower than that of BP neural model which is 0.880. The comparison illustrates that the BP neural network have a better performance for predicting the spatial susceptibility of soil planar slide for Nanjiang county.

Figure 14. ROC curve of logistic regression for area with high susceptibility. Source: Author.
6. Conclusions

Landslide susceptibility assessment is essential for understanding the threats and to predict, prevent, and mitigate landslides. On 16 September 2011, heavy rain triggered more than 1000 soil landslides in the red mudstone area in the Nanjiang County, southwest China. In this study, according to statistical analysis of the distribution of 1146 soil planar slides, and mapping of their susceptibilities using ArGIS in the study area, the following conclusions are drawn:

(1) Of the eight influencing factors — geomorphology, elevation, slope angle, slope structure, slope aspects, slope curvature, the 14-day antecedent rainfall, and the 48-hour cumulative rainfall on 16 September — considered to be used for soil planar slide susceptibility mapping. The slope angle, rainfall, and elevation have the greatest impact on the occurrence of soil planar slides, followed by slope structure, geomorphology, slope curvature, and slope aspects in both the logistic regression and the BP neural network models. This shows that the coupling effects posed by these external and internal factors contribute to the occurrence of rainfall-induced soil planar slides in Nanjiang.

(2) Among these factors, the slope angle has the largest weight and contributes the most to the occurrence of soil planar slides. The slope angles are mainly concentrated in the 10–30° range. In natural conditions, the soil slopes in Nanjiang has good stability. However, in extreme cases, such as the heavy rainfall event with rich antecedent rainfall on 16 September, infiltrated water can decrease the soil shear strength, increase the pore water pressure, and soften the interface between the soil and underlying bedrock, meaning landslides may eventually occur. Therefore, higher weights are assigned to the control factors of the 14-day antecedent rainfall and the 48-hour cumulative rainfall.

(3) The landslide susceptibility in the study area is divided into four levels, which are high, moderate, low, and no susceptibility in both the logistic regression model and the BP neural network model, which is shown in the susceptibility maps. In both the two models, most of the observed soil planar slides were located in areas with high or moderate susceptibility. And in the logistic regression model, total 605 soil planar slides locate in the area with high susceptibility, of which the area is 800.56 km², accounting for 40.31% of the total area.

(4) The validation of two models was also evaluated. The AUC value of the logistic regression model was 0.878 and the parameters of BP neural network has the correlation coefficient of 0.880, which shows the two models are both reliable and reasonable for predicting the spatial susceptibility of soil planar slides. According to field checks, the BP neural network model is verified to have more accurate spatial prediction performance than the logistic regression model, since it has a higher accurate rate of 0.880 than the logistic regression model.

Disclosure statement

This research is financially supported by the Youth Program of National Natural Science Foundation of China (Grant No. 41907243), Key Program of National Natural Science
Foundation of China (Grant No. 41630640), the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (Grant No. 41521002).

**Funding**

This research is financially supported by the Youth Program of National Natural Science Foundation of China (Grant No. 41907243), Key Program of National Natural Science Foundation of China (Grant No. 41630640), the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (Grant No. 41521002).

**References**

Akgun A. 2012. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at Izmir, Turkey. Landslides. 9(1):93–106. doi:10.1007/s10346-011-0283-7.

Aniya M. 1985. Landslide-susceptibility mapping in the amahata river basin, Japan. Annals of the Association of American Geographers. 75(1):102–114. doi:10.1111/j.1467-8306.1985.tb00061.x.

Atkinson PM, Massari R. 1998. Generalized linear modeling of susceptibility to landsliding in the central Apennines, Italy. Computers and Geosciences. 24(4):373–385.

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(1-2):15–31.

Ayalew L, Yamagishi H, Ugawa N. 2004. Landslide susceptibility mapping using GIS based weighted linear combination, the case in Tsugawa area of Agano River, Niigata Prefecture, Japan. Landslides. 1(1):73–81.

Bai S, Wang J, Lu G, Zhou P, Hou S, Xu S. 2010. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. Geomorphology. 115(1-2):23–31.

Bălteanu D, Micu M, Jurchescu M, Malet J-P, Sima M, Kucsića G, Dumitrića C, Petrea D, Mărgărint M C, Bilașco Ş, et al. 2020. National-scale landslide susceptibility map of Romania in a European methodological framework. Geomorphology. 371:107432 doi:10.1016/j.geomorph.2020.107432.

Basak D, Pal S, Patranabis DC. 2007. Support vector regression. Neural Inf. Process. Rev. 11: 203–224.

Blahut J, Van Westen C J, Sterlacchini S. 2010. Analysis of landslide inventories for accurate prediction of debris-flow source areas. Geomorphology. 119(1-2):36–51. doi:10.1016/j.geomorph.2010.02.017.

Brabb EE. 1984. Innovative approaches to landslide hazard mapping. In Proceedings 4th International Symposium on Landslides, Toronto, 1, p. 307–324.

Can T, Nefeslioglu HA, Gokceoglu C, Sonmez H, Duman TY. 2005. Susceptibility assessments of shallow earthflows triggered by heavy rainfall at three subcatchments by logistic regression analyses. Geomorphology. 72(1-4):250–271.

Cantarino I, Carrion MA, Goerlich F, Martinez Ibáñez V. 2019. A ROC analysis-based classification method for landslide susceptibility maps. Landslides. 16(2):265–282.

Chen W, Peng J, Hong H, Shahabi H, Pradhan B, Liu J, Zhu A-X, Pei X, Duan Z. 2018. Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Sci. Total Environ. 626:1121–1135. doi:10.1016/j.scitotenv.2018.01.124.

Chu C-M, Tsai B-W, Chang K. 2009. Integrating Decision Tree and Spatial Cluster Analysis for Landslide Susceptibility Zonation. World Acad. Sci. Eng. Technol. 59: 479–483.
Chung CJF, Fabbri AG. 2003. Validation of spatial prediction models for landslide hazard mapping. Natural Hazards. 30(3):451–472.

Cruden, David M, Varnes, David J, Transportation Research Board Special Report, CHAPTER 3 - LANDSLIDE TYPES AND PROCESSES, Schuster, Robert L, Turner, A Keith, National academy Press, Washington, D. C., 1996, 36, 75, 247.

Dahal RK, Hasegawa S, Nonomura A, Yamanaka M, Masuda T, Nishino K. 2008. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Environ Geol. 54(2):311–324.

Dou J, Paudel U, Oguchi T, Uchiyama S, Hayakawa YS. 2015. Shallow and Deep-Seated Landslide Differentiation using support Vector Machines: a Case Study of the Chuetsu Area, Japan. Terr Atmos Ocean Sci. 26(2-2):227.

Dou J, Yunus A P, Tien Bui D, Sahana M, Chen C-W, Zhu Z, Wang W, Pham B T. 2019. Evaluating GIS-Based Multiple Statistical Models and Data Mining for Earthquake and Rainfall-Induced Landslide Susceptibility Using the LiDAR DEM. Remote Sensing. 11(6):638. doi:10.3390/rs11060638.

Fei W, Peihua X, Changming W, Ning W, Nan J. 2017. Application of a gis-based slope unit method for landslide susceptibility mapping along the longzi river, southeastern tibetan plateau, china. International Journal of Geo-Information. 6(6):172.

Felícizimo á. M, Cuartero A, Remondo J, Quirós E. 2013. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides. 10(2):175–189.

Gigovic I, Drobnjak S, Pamucar D. 2019. The application of the hybrid gis spatial multi-criteria decision analysis best-worst methodology for landslide susceptibility mapping. International Journal of Geo-Information. 8(2):79.

Guo Q, Kelly M, Graham CH. 2005. Support vector machines for predicting distribution of Sudden Oak Death in California. Ecol. Model. 182(1):75–90.

Guzzetti F, Carrara A, Cardinali M, Reichenbach P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, central Italy. Geomorphology. 31(1-4):181–216.

Hungr O, Lerouell S, Picarelli L. 2014. The Varnes classification of landslide types, an update. Landslides. 11(2):167–194.

Jadda M, Shafri HZM, Mansor SB. 2011. PFR model and GiT for landslide susceptibility mapping: a case study from Central Alborz, Iran. Nat Hazards. 57(2):395–412.

Kamp U, Growley BJ, Khattak GA, Owen LA. 2008. GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. Geomorphology. 101(4):631–642.

Kawabata D, Bandibas J. 2009. Landslide susceptibility mapping using geological data, a DEM from ASTER Images and an Artificial Neural Network (ANN). Geomorphology. 113(1-2):97–109. doi:10.1016/j.geomorph.2009.06.006.

Lee S, Min K. 2001. Statistical analysis of landslide susceptibility at Yongin, Korea. Environ Geol. 40(9):1095–1113.

Lee S, Ryu J, Min K, Won J. 2001. Development of two artificial neural network methods for landslide susceptibility analysis. In: IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217). vol. 5, p. 2364–2366.

Lee S, Sambath T. 2006. Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. Environ Geol. (Berlin), 50(6):847–855.

Le LT, Nguyen H, Dou J, Zhou J. 2019a. A Comparative Study of PSO-ANN, GAANN, ICA-ANN, and ABC-ANN in Estimating the Heating Load of Buildings’ Energy Efficiency for Smart City Planning. Appl. Sci. 9(13):2630.

Le LT, Nguyen H, Zhou JJ, Moayedi H. 2019b. Estimating the Heating Load of Buildings for Smart City Planning using a Novel Artificial Intelligence Technique PSO-XGBoost. Appl. Sci. 9(13):2714.
Lee S, Ryu J-H, Lee M-J, Won J-S. 2006. The Application of Artificial Neural Networks to Landslide Susceptibility Mapping at Janghung, Korea. Math Geol. 38(2):199–220. doi:10.1007/s11004-005-9012-x.

Li XZ, Kong JM, Cui Y, Tian SJ, Zhong W, Han PF, Zhang Y. 2014. Analysis on distribution law, types and characteristics and development tendency of secondary geohazards induced by Lushan earthquake. Journal of Engineering Geology. 2:272–279. in Chinese.

Lv Y, Le Q, Bui H-B, Bui X, Nguyen H, Nguyen-Thoi TJ, Song X. 2020. A comparative study of different machine learning algorithms in predicting the content of ilmenite in titanium placer. Appl. Sci. 10(2):635.

Menard S. 1995. Applied logistic regression analysis. Sage University Paper Series on Quantitative Applications in Social Sciences. 106Thousand Oaks, California. : pp. 98.

Møller M F. 1993. Efficient Training of Feed-Forward Neural Networks. DPB. 22(464)doi:10.7146/dpb.v22i464.6937.

Nawi NM, Ransing MR, Ransing RS. 2006. An improved learning algorithm based on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method for back propagation neural networks. In: Sixth International Conference on Intelligent Systems Design and Applications, p. 152–157.

Nefeslioglu Ha, Sezer E, Gokceoglu C, Bozkir aS, Duman TY. 2010. Assessment of Landslide Susceptibility by Decision Trees in the Metropolitan Area of Istanbul, Turkey. Math. Probl. Eng. 2010:1–15.

Oh HJ, Lee S, Chotikasathien W, Kim CH, Kwon JH. 2009. Predictive landslide susceptibility mapping using spatial information in the Pechabun area of Thailand. Environ Geol. 57(3):641–651.

Ohlmacher GC, Davis JC. 2003. Using multiple logistic regression and GIS technology to predict landslide in northeast Kansas, USA. Eng Geol. 69(3-4):331–343.

Pham BT, Bui DT, Prakash I, Nguyen LH, Dholakia MB. 2017. A comparative study of sequential minimal optimization-based support vector machines, vote feature intervals, and logistic regression in landslide susceptibility assessment using GIS. Environ Earth Sci. 76(10):371.

Poiraud A. 2014. Landslide susceptibility–certainty mapping by a multi-method approach: A case study in the Tertiary basin of Puy-en-Velay (Massif central, France). Geomorphology. 216:208–224.

Pourghasemi HR, Moradi HR, Fatemi Aghda SM, Gokceoglu C, Pradhan B. 2014. GIS-based landslide susceptibility mapping with probabilistic likelihood ratio and spatial multi-criteria evaluation models (North of Tehran, Iran). Arab J Geosci. 7(5):1857–1878.

Pradhan B. 2013. A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Comput. Geosci. 51:350–365.

Saito H, Nakayama D, Matsuyama H. 2009. Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan. Geomorphology. 109(3-4):108–121.

Shariati M, Mafipour MS, Mehrabi P, Bahadori A, Zandi Y, Salih MNA, Nguyen HJ, Song X, Poi-Ngian S. 2019. Application of a Hybrid Artificial Neural Network-Particle Swarm Optimization (ANN-PSO) Model in Behavior Prediction of Channel Shear Connectors Embedded in Normal and High-Strength Concrete. Appl. Sci. 9(24):5534.

Singh P, Sharma A, Sur U, Rai PK. 2020. Comparative landslide susceptibility assessment using statistical information value and index of entropy model in bhanupali-beri region, himachal pradesh, india. Environment Development and Sustainability. doi: 10.1007/s10668-020-00811-0.

Su C, Wang L, Wang X, Huang Z, Zhang X. 2015. Mapping of rainfall-induced landslide susceptibility in wencheng, china, using support vector machine. Nat Hazards. 76(3):1759–1779.
Suzen ML, Doyuran V. 2004. Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu Catchment. Turkey. Engineering. Geology. 71(3-4):303–321.

Swets JA. 1988. Measuring the accuracy of diagnostic systems. Science. 240(4857):1285–1293.

Tien Bui D, Ho TC, Pradhan B, Pham BT, Nhu VH, Revhaug I. 2016. GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks. Environmental Earth Science. 75:1–22.

Tien Bui D, Pradhan B, Lofman O, Revhaug I. 2012. Landslide Susceptibility Assessment in Vietnam using support Vector Machines, Decision Tree, and Naïve Bayes Models. Math. Probl. Eng. 2012:1–26.

Vakhshoori V, Zare M. 2018. Is the ROC curve a reliable tool to compare the validity of landslide susceptibility maps? Geomatics Nat Hazards Risk. 9(1):249–266.

Van Den Eeckhaut M, Vanwalleghem T, Poesen J, Govers G, Verstraeten G, Vandekerckhove L. 2006. Prediction of landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium). Geomorphology. 76(3-4):392–410.

Van Westen CJ, Rengers N, Terlien MTJ, Soeters R. 1997. Prediction of the occurrence of slope instability phenomena through GIS-based hazard zonation. Geol Rundsch. 86(2):404–414.

Varnes, D. J., Landslides, analysis and control, Slope movement types and processes, Schuster, R. L., Krizek, R. J., National Academy of Sciences, Washington, D. C., 1978, 11, 33, 176.

Wang DP, He SM, Ge SJ, Pan CP, Zhai MG. 2013. Mountain hazards induced by the earthquake of Sep 07, 2012 in Yiliang and the suggestions of disaster reduction. Journal of Mountain Science. 1:101–107. in Chinese).

Wang Q, Li W, Chen W, Bai H. 2015. GIS-based assessment of landslide susceptibility using certainty factor and index of entropy models for the Qianyang County of Baoji city, China. J Earth Syst Sci. 124(7):1399–1415.

Wooten MR. 2006. Landslide hazards and landslide hazard mapping in North Carolina. Geological Society of America Abstracts with Programs. 38 :28.

Wu Y, Ke Y, Chen Z, Liang S, Zhao H, Hong H. 2020. Application of alternating decision tree with adaboost and bagging ensembles for landslide susceptibility mapping. CATENA. 187:104396.

Xu Q, Li YR, Zhang S, Dong XJ. 2016b. Classification of large-scale landslides induced by the 2008 earthquake. Environ Earth Sci. 75(1):22.

Xu C, Shen LL, Wang GL. 2016a. Soft computing in assessment of earthquake-triggered landslide susceptibility. Environ Earth Sci. 75(9):767.

Yalcin A. 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. CATENA. 72(1):1–12.

Yao X, Dai FC. 2006. Support Vector Machine Modeling of Landslide Susceptibility Using a GIS: A Case Study. Geol. Soc. IAEG, London. 793:1–12.

Yesilnacar 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(3-4):251–266.

Yilmaz I. 2007. GIS based susceptibility mapping of karst depression in gypsum: A case study from Sivas basin (Turkey). Eng Geol. 90(1-2):89–103.

Yilmaz I. 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(6):1125–1138.

Zeng-Wang X. 2001. GIS and ANN model for landslide susceptibility mapping. J Geogr Sci. 11(3):374–381.

Zhang S, Xu Q, Hu Z. 2016. Effects of rainwater softening on red mudstone of deep-seated landslide, Southwest China. Eng. Geol. 204:1–13. doi:10.1016/j.enggeo.2016.01.013.
Zhang Q, Xu Q, Ning N. 2014. A study of the stability influence factors and coupling for inclined-shallow soil landslides under the condition of rainfall. Hydrogeology and Engineering Geology. 5:90–94. in Chinese.

Zhang S, Xu Q, Zhang Q. 2017. Failure characteristics of gently inclined shallow landslides in Nanjiang, southwest of China. Eng Geol. 217:1–11.

Zhang M, Yin Y, Huang B. 2015. Mechanism of rainfall-induced landslide in gently inclined red beds in the eastern Sichuan Basin, SW China. Landslides. 12(5):973–983.

Zhang S, Zhang LM. 2017. Impact of the 2008 Wenchuan earthquake in China on subsequent long-term debris flow activities in the epicentral area. Geomorphology. 276:86–103.

Zhang S, Zhang LM, Peng M, Zhang L, Zhao H, Chen H. 2012. Assessment of risks of loose landslide deposits formed by the 2008 Wenchuan earthquake. Nat Hazards Earth Syst Sci. 12(5):1381–1392.

Zhang S, Zhang LM, Lacasse S, Farrokh N. 2016b. Evolution of Mass Movements near Epicenter of Wenchuan Earthquake, the First Seven Years. Scientific Report 6.