

Research Article

GIS-Based Landslide Susceptibility Mapping Using Information, Frequency Ratio, and Artificial Neural Network Methods in Qinghai Province, Northwestern China

Bin Li ^{1,2,3} Nianqin Wang,¹ and Jing Chen ⁴

¹College of Geology and Environment, Xi'an University of Science and Technology, Xi'an, Shaanxi 810000, China

²The Third Non-ferrous Geological Exploration Institute of Qinghai Province, Xining, Qinghai 810000, China

³Key Laboratory of Geological Processes and Mineral Resources, Northern Qinghai-Tibet Plateau, Xining, Qinghai 810000, China

⁴Qinghai Normal University, Xining, Qinghai 810000, China

Correspondence should be addressed to Jing Chen; chenjing_nu@126.com

Received 11 April 2021; Revised 16 May 2021; Accepted 7 June 2021; Published 1 July 2021

Academic Editor: Jian Ji

Copyright © 2021 Bin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Landslides are one of the nature hazards causing a lot of casualties and property losses in the world. Over the last decades, many researchers have made contributions in landslide susceptibility maps using qualitative and quantitative methods. Parameters of DEM, geology, etc. are selected to analyze the mechanism of landslides. The quality of data is essential in the landslide studies, and more credible results can be obtained if the data is adequate and accurate from the wide range of parameters. The aim of this study is to evaluate the landslide susceptibility of Huangyuan County of Qinghai. Through field investigations, 100 landslide disaster locations in the study area were selected, and 11 influencing factors including elevation, slope, aspect, plane curvature, profile curvature, road distance, river distance, fault distance, stratum rock property, vegetation coverage index, and terrain humidity index were selected as the influencing factors of landslide disaster based on GIS. In this paper, the information method (IM) model, frequency ratio (FR) model, and artificial neural network (ANN) model are used to evaluate the susceptibility of geological hazards, and the receiver operating characteristic (ROC) curve of disaster points at different levels is used to test the evaluation accuracy of three models. The results show that factors that have great influence on landslides are associated with witness, and the terrain humidity index has the highest weight in the occurrences of landslide. The values of AUC indicate that the ANN model is the best evaluation model suitable for the study area and can be extremely useful for landslide hazard mitigation strategies. Based on the calculation of ANN model, three valley areas are determined with high landslide susceptibility, and necessary reinforcement measures should be taken.

1. Introduction

Landslide is a kind of natural hazard in the mountainous regions, threatening human life and property [1–4]. Over the past decades, the huge catastrophe ability of landslides attracted many researchers to devote themselves to assessing landslide susceptibilities [5–7]. From 2014 to 2020, 47614 geological disasters occurred in China, including 33659 landslides, accounting for 68.27% of the total number of disasters. In the Qinghai region, large loess areas existed, where there are mountainous topographical features and have a high incidence of severe landslides.

The occurrences of landslides are extremely complicated and affected by many factors such as geologic structure, lithological association, topography, rainfall, earthquake, and human activity [8–12]. Based on these factors, various assessing methods have been proposed to analyze the landslide susceptibility, which can be divided into two categories: qualitative analysis and quantitative analysis [13–18]. The qualitative method mainly relies on the judgment of experts, which is can be seen as a kind of subjective method [19–24]. In the process of qualitative analysis, the spatial distribution of unstable slopes is based on experts' understanding of the relationship between the occurrence of

landslides and the assumed antecedent factors, directly determined by the existing landslides or potentially unstable areas. Quantitative analysis, an objective evaluation, is a kind of numerical estimation, such as calculating the probability of landslides. This analysis method is to estimate the potentially unstable area by using the inducing factors related to landslide occurrence.

The quantitative methods mainly include deterministic and statistical ones [25]. A deterministic method is a mathematical model based on the physical and mechanical mechanism to control the slope failure. The most common method is to combine various hydrological models and slope stability evaluation methods to obtain the safety factor of corresponding units. It has the highest accuracy in these models. On the other hand, a statistical model is to analyze the relationship between the landslide catalogue map and factors affecting landslide occurrence and then get the spatial possibility of landslide occurrence, which belongs to the indirect quantitative evaluation method. This method is more practical for landslide evaluation at the mesoscale and is the most widely used method in landslide sensitivity evaluation at present. Each evaluation model has been proved to have different advantages and disadvantages, some of which may be more suitable for solving specific problems in specific areas or implementing specific projects [2].

Although a large number of models and methods have been proposed to produce landslide susceptibility maps (LSM) using geographic information systems (GIS) [13–17], a consensus has not been established regarding which methods are the most suitable, because qualitative techniques can be limited by unconsidered phenomena or incomplete knowledge that the expert decisions are based upon. On the other hand, quantitative methods suffer from inaccurate or low-precision data. Until July 2012, 171 geological disaster points have been investigated. Among them, landslide and debris flow account for 81% [26]. The disasters damaged 20 houses, killed 54 people, and injured thousands of people. Economic losses caused by these events are estimated to be around 70 million RMB (about 10 million U.S. dollars). Therefore, it is necessary to assess and manage areas that are susceptible to landslides and to mitigate any risk associated with them.

Our study aims to find a more suitable landslide susceptibility model for Huangyuan City, Qinghai Province, China. Firstly, a total of 100 landslides were mapped in the study area based on a geological hazard survey (1 : 50,000) of the Qinghai region. Then, according to the geological data, field survey, and landslide information, eleven influencing factors, namely, elevation, slope, aspect, plane curvature, profile curvature, road distance, river distance, fault distance, stratum rock property, vegetation coverage index, and terrain humidity index, were selected for landslide susceptibility mapping. Then, the information method (IM) model, frequency ratio (FR) model, and artificial neural network (ANN) model were adopted to establish landslide susceptibility models. Finally, the receiver operating characteristic (ROC) curve was used to validate and compare the prediction abilities of the landslide susceptibility models and select the optimal one. The results of this study can be

extremely useful for landslide hazard mitigation strategies in the Qinghai region.

2. Study Area

The study area is located in the southeast of Qinghai Province, which is the transition zone between the first and second terrain steps in China (Figure 1), and also the marginal zone of the uplift of the Qinghai Tibet Plateau. The unique geographical location creates unique geological environment conditions, which provides a good disaster pregnant environment for the development of landslide disasters in the study area. The whole study area covers an area of 1545 km², with an altitude of 2470 m–4484 m, the maximum vertical elevation difference of 2014 m, and the terrain inclines from north, West, and south to East. According to the geomorphic types, the aiming area can be divided into four units: tectonic erosion high mountain area, tectonic erosion middle mountain area, tectonic erosion low mountain and hilly area, and valley belt plain area. There are 86 large and small rivers, which belong to the Yellow River system.

The exposed pre-Quaternary strata are Proterozoic, Triassic, Cretaceous, and Paleogene. Tectonically, the study area is located in the first-order tectonics of the Qilian geosynclinal fold system, crossing two second-order tectonic belts of the middle Qilian geosyncline and South Qilian geosyncline. The fault structures are developed, and the folds are mostly in the form of compound structures. The tectonic movements of each stage are reflected in varying degrees. The main faults are distributed along the NW, NW, nearly EW, NE. The tectonic line is NWW. Affected by various external forces, geological disasters easily occur, which is one of the most developed areas in Huangshui River Basin. The climate of the study area belongs to continental semiarid climate, the rainfall is concentrated in May to September every year, and the average annual precipitation is about 405.5 mm. According to statistics, many fatal landslides occurred in this period.

Data of this study includes 100 landslide collapse points, which are provided by Qinghai geological environment monitoring station. Through field investigation, this study makes an in-depth study on the landslide. The study shows that almost all landslides and most soil collapses occur in the loess layer, and the bedrock is the basement stratum of the whole area, which constitutes the sliding bed of the Loess bedrock interface landslide. The losses slopes are a main type in the study, as shown in Figure 2.

3. Methodology

Figure 3 shows the main steps of this study: (a) data preparation, (b) landslide susceptibility modeling, and (c) validation and selection of the optimal models.

4. Data Preparation

4.1. Landslide Inventory Mapping. The information of the existing landslide distribution is essential for the

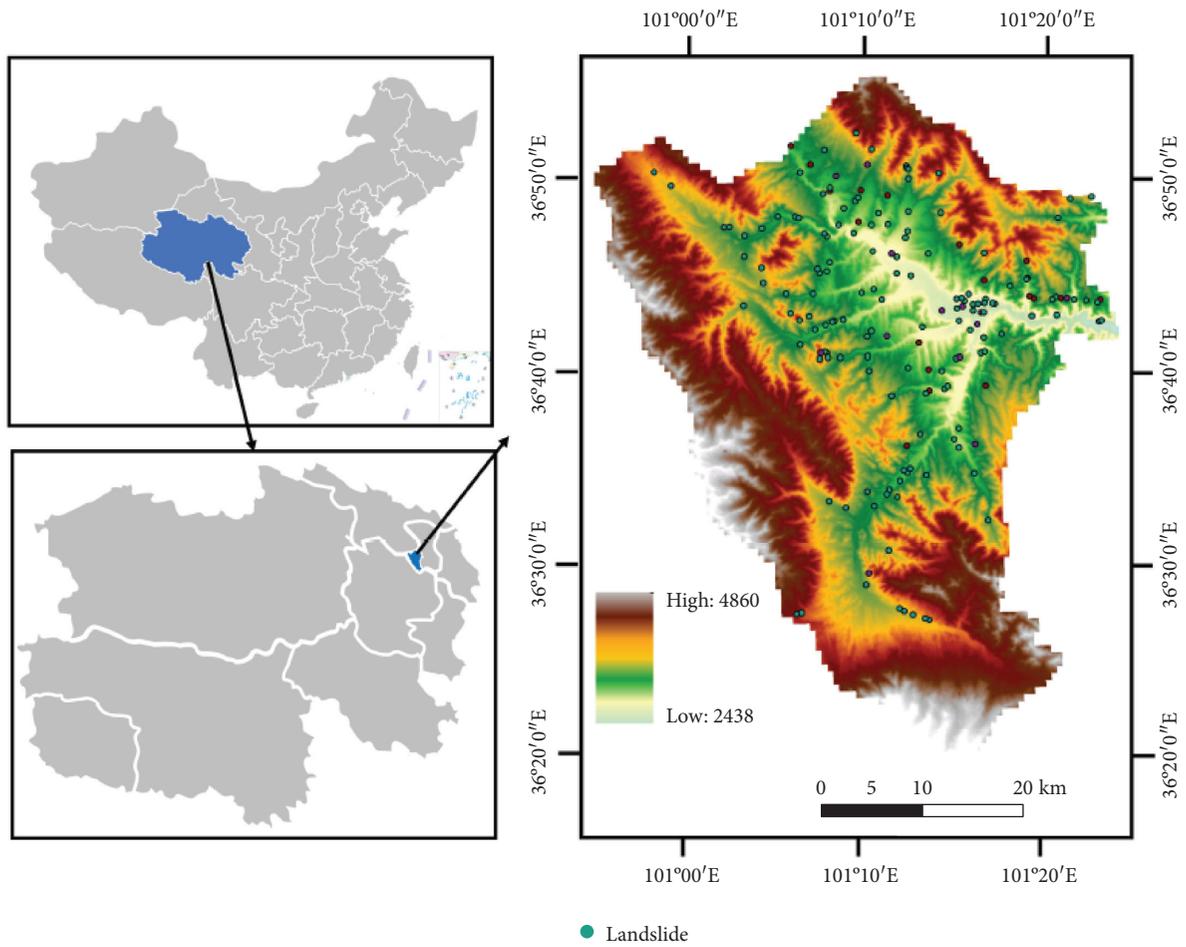


FIGURE 1: Geographical location map of the study area.

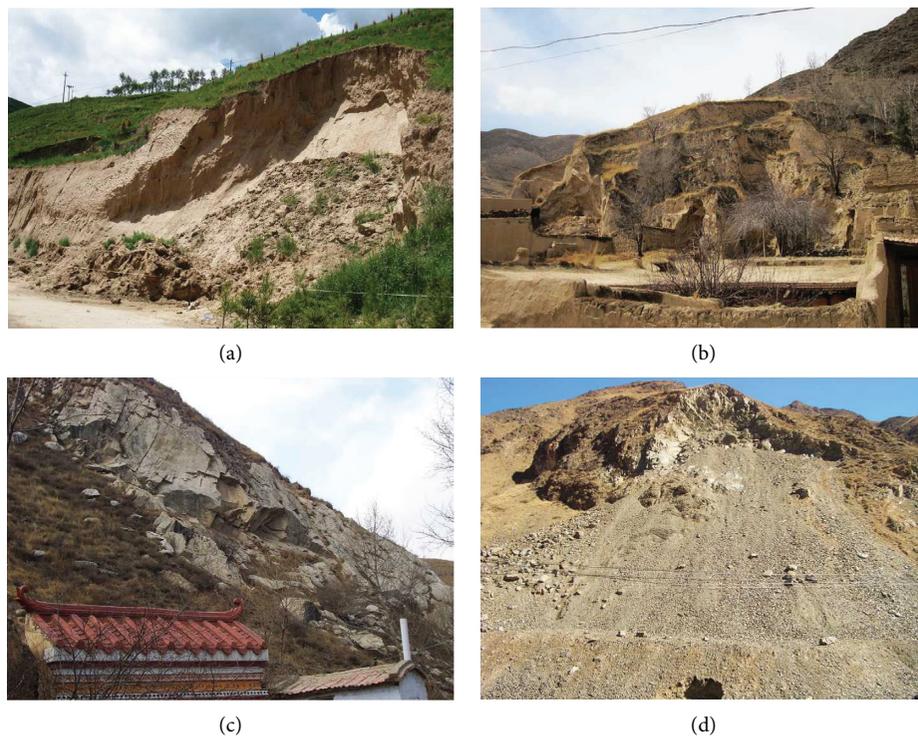


FIGURE 2: Examples of landslides in the study area.

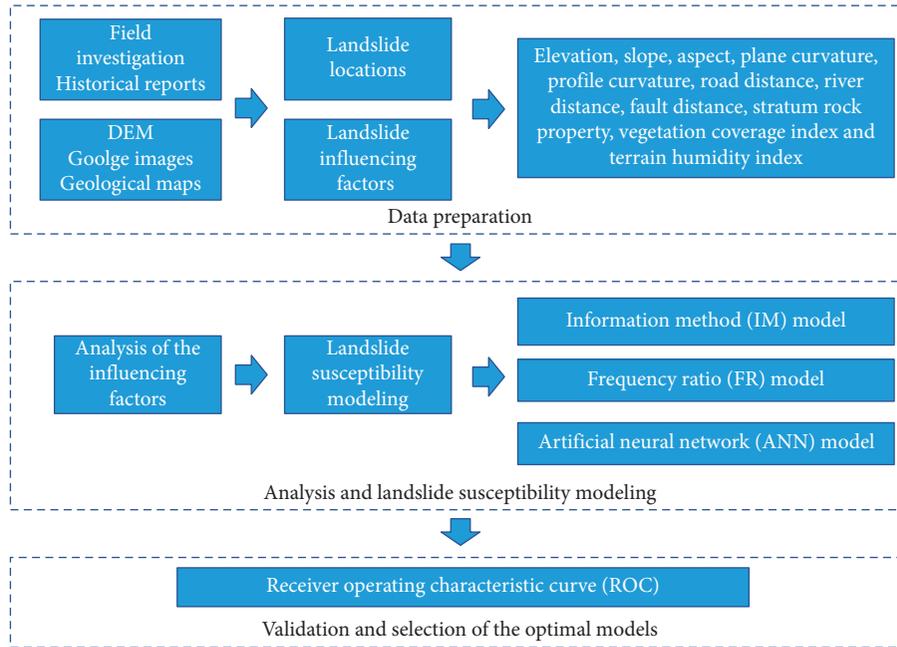


FIGURE 3: Flowchart describing the methodology applied for the analysis.

identification of likelihood of landslides [27]. In this study, landslide locations were determined from the analysis of Google Earth images, the historical records of which are the landslide inventory maps of the Department of Natural Resources of Qinghai Province, and from field investigation (Table 1). The total landslides mapped are 100, which are translated to 572 pixels of 50-meter landslide cell size. The landslide inventories were divided into training data (80% of total landslide cases, 456 pixels) and validation data (20% of total landslide cases, 116 pixels).

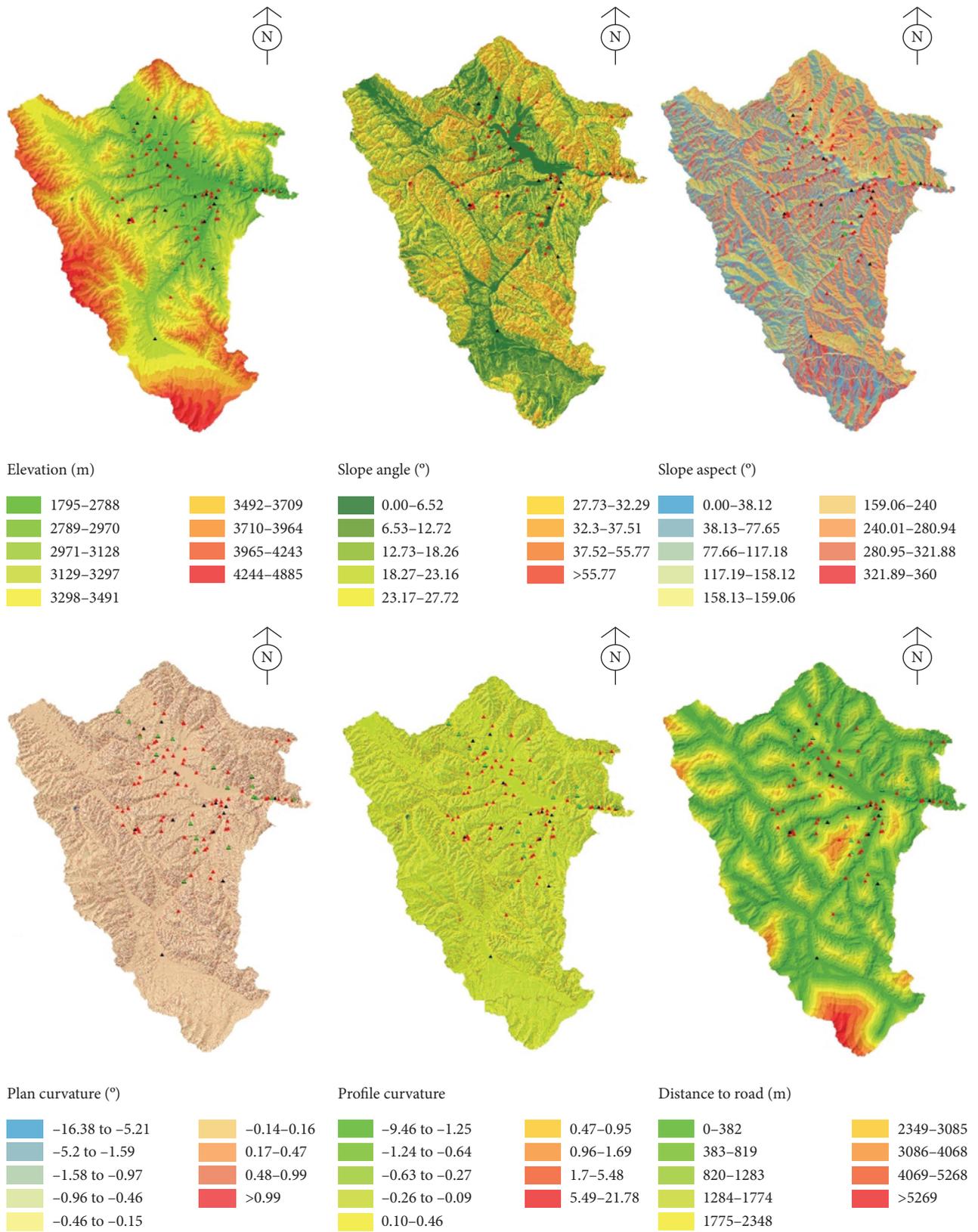
4.2. Influencing Factors. The occurrences of landslide are comprehensive effects of influencing factors. Through the analysis of literature and data, combined with field investigation, this paper selects 11 landslide-prone factors for modeling, which are elevation, slope angle, slope aspect, plane curvature, profile curvature, distance to road, distance to river, distance to fault, stratum lithology, vegetation coverage index, and terrain humidity index, as shown in Figure 4. Elevation has connection with landslide occurrence [29], especially in plateau area. Slope angle can affect the slope stability, and the slope aspect controls the hours of sunshine and the effects of rainfall, moisture, and wind conditions over the study area [30]. Plane curvature and profile curvature are morphological factors, which control the water flow on Earth surface affecting landslide occurrences [31]. The traffic and road construction can produce vibration, which is an inducement of landslide, so the distance to road is also an important influencing factor [29]. The erosion by the river to the bank can reduce the strength of the soil and make the slope less stable, which is a positive correlation with the distance to river [30]. The existence of fault makes the rock fragmented and increase the probability of slope instability [30, 32]. Stratum lithology is the material

TABLE 1: Data sources.

Data	Primary format	Data source
DEM	Grid	ASTER GDEM
Geology	Shapefile (polygon)	Qinghai digital basemap [28]
River and roads	Shapefile (line)	Qinghai digital basemap [28]
Landslide inventory	Shapefile (point)	Field survey

basis of landslides, which is also an influencing factor of slope stability [32]. Vegetation coverage index and terrain humidity index are important environmental factors and are associated with the structure of soil, which are frequently used in mapping the landslide susceptibility [30]. The Jenks natural breaks method is a one-dimensional clustering algorithm. This method considers that there are discontinuities in the data itself, so that the variance of different categories is the largest, and the variance of the same category is the smallest to optimize the classification. The Jenks natural breaks method can be used to classify the landslide impact factors, which can well characterize the distribution of the impact factors. Therefore, the Jenks natural breaks method was applied in the process of classification [29].

4.2.1. Elevation. The 1 : 50000 topographic map of the study area is processed by ArcGIS, and an elevation model with an accuracy of 30 m × 30 m is drawn. They can be divided into 9 categories: 1795–2788 m, 2788–2970 m, 2970–3128 m, 3128–3297 m, 3297–3491 m, 3491–3709 m, 3709–3964 m, 3964–4243 m, and 4243–4885 m by using Jenks natural breaks method. In the background of elevation, the landslide geological disasters in the study area are mainly



(a)

FIGURE 4: Continued.

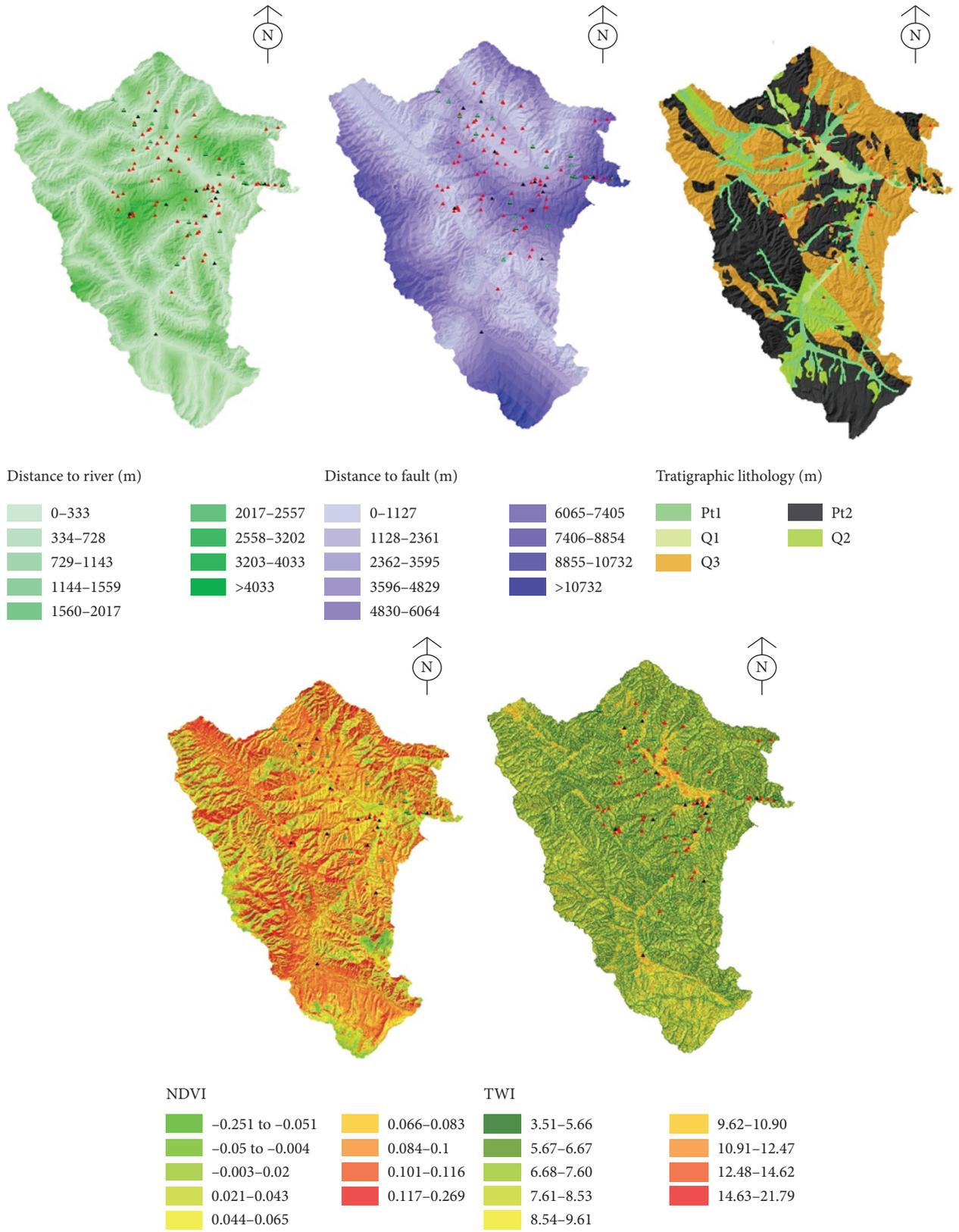


FIGURE 4: The spatial database of landslide susceptibility analysis.

concentrated in the elevation of 1795–3297 m, and the landslide with the elevation of 2788–2970 m accounts for the largest proportion, 42%. However, there is no record of landslide disaster above 3297 m, which is related to the area where people live.

4.2.2. Slope Angle. When the slope increases, the shear stress at the foot of the slope will also increase; that is, the slope with a larger slope is more prone to deformation and failure. For digital elevation model (DEM), otherwise, it can be directly changed to slope data. According to the Jenks natural breaks method, it can be divided into 9 categories: 0° – 6.52° , 6.53° – 12.72° , 12.73 – 18.26° , 18.27 – 23.16° , 23.17 – 27.72° , 27.73 – 32.29° , 32.3 – 37.51° , 37.52 – 55.77° , and 55.78 – 83.17° . In the slope background, the landslide geological disasters in the study area are mainly concentrated in the slope range of 6° to 32° , and the landslide disasters with the slope range of 18.27° to 23.16° account for the largest proportion (23%).

4.2.3. Slope Aspect. The surface receives different solar radiation for different aspect of slope, so it will affect the degree of vegetation coverage, surface weathering, and surface evaporation, thus affecting the occurrence of landslides. The DEM data can be divided into 9 types according to the slope aspect: 0° to 40° , 40° to 80° and 80° – 120° , 120 – 160° , 160 – 200° , 200 – 240° , 240 – 280° , 280 – 320° , and 320 – 360° . Under the slope aspect background, the landslides in the study area mainly occur between 160° and 200° of the slope aspect, and the landslides with the slope aspect of 160° – 200° account for the largest proportion for 21%.

4.2.4. Plan Curvature. The plane curvature reflects the terrain. Positive value means that the terrain surface is convex, negative value means that the terrain surface is concave, and 0 means that the ground is flat. According to the plane curvature, DEM data can be divided into nine categories: -16.38 – -5.21 , -5.21 – -1.59 , -1.58 – -0.97 , -0.96 – -0.46 , -0.45 – -0.15 , -0.14 – 0.16 , 0.17 – 0.47 , 0.48 – 0.99 , and 1 – 9.99 . Under the background of plane curvature, the landslide disasters in the study area are mainly concentrated in the plane curvature range of -0.45 – 0.99 , of which -0.14 – 0.16 is the largest, accounting for 38%.

4.2.5. Profile Curvature. The curvature of the profile reflects the change rate of the ground slope. A positive value indicates that the terrain surface is convex, a negative value indicates that the terrain surface is concave, and 0 indicates that the ground is flat. The DEM data are classified into 9 types according to the slope curvature by the Jenks natural breaks method: -9.46 – -1.25 , -1.25 – -0.62 , -0.63 – -0.27 , -0.26 – 0.09 , 0.1 – 0.46 , 0.47 – 0.95 , 0.96 – 1.69 , 1.7 – 5.48 , and 5.49 – 21.78 . Under the background of section curvature, the landslide disasters in the study area are mainly concentrated

in the section curvature of -0.26 – 0.95 , of which the landslide disasters between 0.1 and 0.46 account for the largest proportion, 39%.

4.2.6. Distance to Road. When people build roads, they will carry out a series of activities, such as manual excavation of slope toe and blasting. These activities will change the original rock and soil structure, affect the stability of rock and soil, and leave hidden dangers for the occurrence of landslide disasters. According to the Jenks natural breaks method, it can be divided into 9 categories: 0 – 382 m, 382 – 819 m, 819 – 283 m, 1283 – 1774 m, 1774 – 2348 m, 2348 – 3085 m, 3085 – 4068 m, 4068 – 5269 m, and > 5269 m. Under the background of road distance, the landslides in the study area are mainly concentrated in the range of 0 – 819 m, and the landslides with road distance of 0 – 382 m account for the largest proportion (59%).

4.2.7. Distance to River. The occurrence of landslide disasters is closely related to surface water. The river constantly erodes and hollows out the slope toe, which will lead to slope instability. Based on the Euclidean distance calculation of the river data in the study area by ArcGIS, the river data can be divided into nine categories: 0 – 333 m, 333 – 728 m, 728 – 1143 m, 1143 – 1559 m, 1559 – 2017 m, 2017 – 2557 m, 2557 – 3202 m, 3202 – 4033 m, and >4033 m by using the Jenks natural breaks method. Under the background of river distance, landslides in the study area are mainly concentrated in the river distance of 0 – 728 m, and the landslide disaster between 0 and 333 m accounts for the largest proportion, 48%.

4.2.8. Distance to Fault. In the fault zone, the structure of rock and soil is broken and easy to be weathered, which has a certain impact on the occurrence of landslide disaster. The Euclidean distance of fault data in the study area is calculated by ArcGIS, and then the Jenks natural breaks method is used to divide it into 9 categories: 0 – 1127 m, 1127 – 2361 m, 2361 – 3595 m, 3595 – 4829 m, 4829 – 6064 m, 6064 – 7405 m, 7405 – 8854 m, 8854 – 10732 m, and >10732 m. Under the background of fault distance, the landslide geological disasters in this study area are mainly concentrated in the fault distance of 0 – 2361 m, and the landslide disasters with fault distance of 0 – 1127 m account for the largest proportion, accounting for 29%.

4.2.9. Stratigraphic Lithology. Stratigraphic lithology is the internal control factor that affects the occurrence of landslide geological disasters, and its type and composition affect slope stability. There are 5 types of lithology in the study area: hard thick medium thick-bedded metamorphic rock group (Pt1), hard block intrusive rock group (Pt2), single structure Aeolian loess group (Q1), double structure alluvial proluvial sand and sand gravel pebble soil group (Q2),

multilayer structure clay, argillaceous gravel pebble, and broken stone soil group (Q3). According to ArcGIS data reclassification, under the background of stratum lithology, the landslide disaster in the study area is mainly in the single structure Aeolian loess, accounting for 34%.

4.2.10. Normalized Difference Vegetation Index. Vegetation can play a role in soil and water conservation and slope protection, which has a certain impact on the occurrence of landslide geological disasters. Normalized Difference Vegetation Index (NDVI) is used to represent the state of plant growth. The negative value indicates that the ground is covered by clouds or snow, 0 indicates that the ground is covered by bedrock or bare soil, and the positive value indicates that the ground is covered by vegetation. Band image data are obtained from the geospatial data of Huangyuan county: Landsat 8 LC 8132035201 6028IGN00. The image data was generated on January 28, 2016, and the cloud cover was 1.89, which can observe the vegetation coverage in the study area. NDVI was calculated by env5.3, and then the data were classified into 9 categories by ArcGIS Jenks natural breaks method: $-0.251-0.051$, $-0.05-0.004$, $-0.003-0.020$, $0.021-0.043$, $0.044-0.065$, $0.066-0.083$, $0.084-0.1$, $0.101-0.116$, and $0.117-0.269$. Under the background of NVDI, the landslide disasters in the study area are mainly concentrated in the range of NVDI $0.004 \sim 0.116$, and the landslide disasters with NVDI $0.084 \sim 0.1$ account for the largest proportion, 33%.

4.2.11. Topographic Wetness Index. Terrain Wetness Index (TWI), which can quantitatively display the control of terrain on the spatial distribution of soil moisture, is a widely used terrain attribute. According to ArcGIS, the data were classified into 9 categories: $3.51-5.66$, $5.67-6.67$, $6.68-7.6$, $7.61-8.53$, $8.54-9.61$, $9.62-10.9$, $10.91-12.47$, $12.48-14.62$, and $14.63-21.79$. Landslides in the study area are mainly concentrated in the range of TWI $3.51-8.53$, and the

landslides with TWI $5.67-6.67$ account for the largest proportion, accounting for 36%.

5. Evaluation Methodologies and Results

5.1. Variance Inflation Factor (VIF). The variance inflation factor is a method to judge multicollinearity by examining the degree to which a given explanatory variable is explained by all other explanatory variables in the equation. Any influencing factor with a VIF value of greater than 10 should be excluded from the landslide susceptibility model. The VIFs in Table 1 show that the values of the influence factors are all below 10, so no factor needed to be excluded from the landslide susceptibility model.

5.2. Information Model. The information method (IM) [16, 19] is a kind of bivariate statistical analysis method. By analyzing the actual situation and the information provided in the deformed or occurring geological disaster areas, it studies the quantity and quality of the information that has an impact on their stability and quantifies the degree of its impact through the information. The information is calculated as follows:

$$I_i = \sum_1^n I(M, x_i) \quad (1)$$

$$= \ln \frac{N_i/N}{S_i/S},$$

where N_i is the total number of geological hazards in the class i evaluation factor of the study area; N is the total number of units with geological hazards in the study area; S_i is the number of units with class i evaluation factors in the study area; S is the total number of evaluation factor units in the study area; I_i is the total information value of the study area; n is the number of evaluation factors.

The landslide susceptibility index (LSI) can be calculated as follows:

$$LSI_{ICM} = IM_{\text{elevation}} + IM_{\text{slope angle}} + IM_{\text{slope aspect}} + IM_{\text{plane curvature}} \\ + IM_{\text{profile curvature}} + IM_{\text{distance to road}} + IM_{\text{distance to river}} + IM_{\text{distance to fault}} \\ + IM_{\text{stratum lithology}} + IM_{\text{vegetarian cover age index}} + IM_{\text{terrain humidity index}}, \quad (2)$$

where IM indicates the influencing factor maps that have been reclassified as per their information content values.

5.3. Frequency Ratio Method. The frequency ratio method (FR) [31, 33] is a more traditional statistical analysis model. A landslide (M) is affected by many factors (n) (such as elevation, aspect, and lithology), so the frequency ratio method divides M into n classes or n grades according to certain rules. The simplified calculation formula is as follows:

$$FR_{ij} = \frac{HA_{ij} / \sum HA_{ij}}{HA_{ij} / \sum HA_{ij}} * 100\%, \quad (3)$$

$$LSI_{FR} = \sum FR,$$

where FR_{ij} is the frequency ratio of i th factor, j th subfactor; HA_{ij} is landslide area of i th factor, j th subfactor; XA_{ij} is the area of i th factor, j th subfactor; LSI_{FR} is the Landslide sensitivity index.

5.4. Artificial Neural Network. The artificial neural network model (ANN) [31, 34] consists of the input layer, one or more hidden layers and an output layer with the different number of neurons. The adjacent layers are fully connected, and each connection is assigned weight. According to Kolmogorov theorem [35], an artificial neural network model with a hidden layer can simulate any nonlinear mapping from n -dimension to m -dimension on a closed set with any precision. Therefore, this study adopts a three-layer network structure, including an input layer, output layer, and a hidden layer. In the input layer, the number of neurons is the number of landslide risk factors; in the output layer, the softmax function is used to output two nodes, which represent the landslide-prone and not prone. The landslide-prone is marked as (1, 0), and the not prone is marked as (0, 1); in the hidden layer, the tanh function is used. The model structure used in this paper is shown in Figure 5. In the process of calculation, each neuron receives the output data of the upper layer neuron, processes the data according to the corresponding connection weight, and outputs the calculation results to the next layer neuron. The learning process of the artificial neural network model is the process of continuously adjusting the network parameters. BP artificial neural network model, based on the cumulative error between the real value and the output value, adopts a gradient descent algorithm to optimize the parameters. The determination process of disaster areas can be divided into three steps. Firstly, the nonlandslide units and landslide units selected from the information partition are used as the test set and training set, 80% of which are trained, and the remaining 20% are used as the test set. Then, 1826315 grids of Huangyuan county are input into the trained neural network model to get the grid probability. Finally, the Jenks natural breaks method is used to reclassify in probability.

6. Results

6.1. IM Method. Based on the information model (IM), the IM values of the influencing factors were calculated, and the results are presented in Table 2. If the value of IM is bigger than 0, it means that the tendency of the landslide is high; on the other hand, it means that the tendency is low. The IM values of Elevation show that the occurrences of landslide are mainly concentrated in the low elevation area. For the slope angle, the classes 12.73–18.26° and 18.27–23.16° have the largest IM values. The higher the slope angle, the greater the effect on the landslide. For the slope aspect, the IM values of the classes 160–200° and 200–240° are 0.391 and 0.105, respectively. For the plane curvature and profile curvature, the highest IM values are 1.532 for class -0.14–0.16 and 0.938 for class 0.47–0.95, respectively. The IM values of distance to a river, road, and fault show that too small distance can also cause landslides. Especially, the vibration from passing vehicles on the road, the erosion of the river, and the weakness of fault are the main influencing factors. The values in Table 3 show that the IM of classes Q1 and Q2 is 1.395 and 1.322, respectively. It indicates that Q1 and Q2 have a high probability of landslide occurrences in this region. The lithology of Q1 is mainly loess, and the lithology

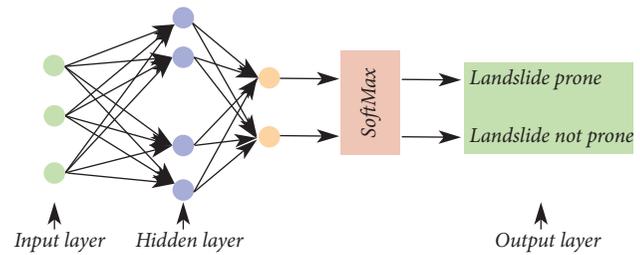


FIGURE 5: Structure of artificial neural network model.

TABLE 2: The VIF values for the influencing factors.

Influencing factor	VIF
Elevation	1.902
Slope angle	1.835
Slope aspect	1.213
Plan curvature	1.078
Profile curvature	1.153
Distance to road	1.233
Distance to river	1.058
Distance to fault	1.008
Stratigraphic lithology	1.621
Normalized difference vegetation index	1.746
Topographic wetness index	1.621

of Q2 is mainly argillaceous gravel pebble and broken stone soil. The soil of Q1 and Q2 is extremely weak of the property and easy to lose stability. Thus, landslides are easy to occur in these two strata. The vegetation can effectively reduce the landslide, but the wetness of soil can promote the landslide based on the IM of Normalized Difference, Vegetation Index, and Topographic Wetness Index. The landslide susceptibility map produced by the IM method is shown in Figure 6(a).

6.2. FR Method. FR values of each class of the eleven influencing factors are summarized in Table 2. The FR values in Table 2 show a similar tendency with the results of IM. The biggest FR values of slope angle and aspect also are in the classes 18.27–23.16° and 160–200°, respectively. The elevation class generated a high FR value in the 2788–2970 m subclass. The distance to the river, road, and fault shows the highest FR value in the subclasses of 0–333 m, 0–382 m, and 0–1127 m, respectively. For the lithology, the subclass of Q1 and Q2 had the highest FR value. The highest FR values of other influencing factors are focused on the same subclass with IM values. The landslide susceptibility maps were reclassified using Jenks natural breaks method. The outcome was an interpretable map showing increasing spatial possibility of future landslide incidence ranging from very low to very high susceptibility to a landslide (Figure 6(b)).

6.3. ANN Method. The ANN method was performed with a mean square error of 0.02 in the training process. The landslide susceptibility value can be obtained by

TABLE 3: Calculations of IM and FR of various evaluation factors.

Factors	Subclass	IM	FR
Elevation	1795–2788 m	0.478	1.321
	2788–2970 m	0.673	1.433
	2970–3128 m	0.195	1.115
	3128–3297 m	-0.237	0.876
	3297–3491 m	-0.398	0.775
	3491–3709 m	-0.786	0.306
	3709–3964 m	-1.231	0.205
	3964–4243 m	-1.432	0.114
	4243–4885 m	-2.075	0.101
Slope angle	0–6.52°	-0.121	0.506
	6.53–12.72°	0.265	1.121
	12.73–18.26°	0.431	1.045
	18.27–23.16°	0.591	1.231
	23.17–27.72°	0.375	1.033
	27.73–32.29°	0.115	0.897
	32.3–37.51°	-0.153	0.555
	37.52–55.77°	-0.271	0.246
	55.78–83.17°	-0.489	0.387
Slope aspect	0–40°	-0.793	0.831
	40–80°	-0.323	1.031
	80–120°	0.051	1.056
	120–160°	-0.115	0.836
	160–200°	0.391	1.154
	200–240°	0.105	1.038
	240–280°	-0.203	1.016
	280–320°	-0.315	1.114
	320–360°	-0.256	0.983
Plane curvature	-16.38--5.21	0.135	1.178
	-5.21--1.59	0.003	1.045
	-1.58--0.97	0.106	1.036
	-0.96--0.46	0.117	1.187
	-0.45--0.15	0.786	0.709
	-0.14-0.16	1.532	0.778
	0.17-0.47	0.975	0.613
	0.48-0.99	0.398	1.077
	1-9.99	0.449	0.832
Profile curvature	-9.46--1.25	-0.451	0.561
	-1.25--0.62	-0.232	0.876
	-0.63--0.27	0.124	1.139
	-0.26-0.09	0.108	0.936
	0.1-0.46	0.549	0.787
	0.47-0.95	0.938	1.259
	0.96-1.69	0.225	1.015
	1.7-5.48	-0.331	0.897
	5.49-21.78	-0.261	0.732
Stratigraphic lithology	Pt1	-2.121	0.325
	Pt2	-1.589	0.446
	Q1	1.395	1.035
	Q2	1.322	1.153
	Q3	0.876	1.087
Distance to river	0-333 m	1.231	2.357
	333-728 m	1.053	1.753
	728-1143 m	-0.113	1.156
	1143-1559 m	-1.452	0.797
	1559-2017 m	-1.121	0.568
	2017-2557 m	-1.643	0.321
	2557-3202 m	-2.152	0.253
	3202-4033 m	-1.988	0.118
	>4033 m	-2.223	0.074

TABLE 3: Continued.

Factors	Subclass	IM	FR
Distance to fault	0-1127 m	2.156	3.754
	1127-2361 m	1.986	2.531
	2361-3595 m	0.993	1.674
	3595-4829 m	0.635	1.152
	4829-6064 m	0.141	1.041
	6064-7405 m	-0.241	0.93
	7405-8854 m	-0.764	0.87
	8854-10732 m	-1.653	0.551
	>10732 m	-2.229	0.324
NDVI	-0.251--0.051	1.532	1.986
	-0.05--0.004	2.11	2.321
	-0.003-0.020	2.563	3.267
	0.021-0.043	0.648	1.323
	0.044-0.065	-0.613	0.976
	0.066-0.083	-0.787	0.815
	0.084-0.1	-0.218	0.693
	0.101-0.116	-1.131	0.556
	0.117-0.269	-2.153	0.323
TWI	3.51-5.66	-0.881	0.653
	5.67-6.67	0.112	1.132
	6.68-7.6	0.308	1.512
	7.61-8.53	0.235	1.073
	8.54-9.61	1.568	3.151
	9.62-10.9	1.332	2.312
	10.91-12.47	0.863	1.457
	12.48-14.62	0.151	1.032
	14.63-21.79	0	0
Distance to road	0-382 m	2.132	4.273
	382-819 m	0.876	1.182
	8191-283 m	0.231	0.998
	1283-1774 m	-0.115	0.693
	1774-2348 m	-0.289	0.671
	2348-3085 m	-0.787	0.773
	3085-4068 m	-0.563	0.712
	4068-5269 m	-1.132	0.683
	>5269 m	-2.225	0.312

$$LSI_{ANN} = \sum_{i=1}^n fw_i Xw_{ij}, \quad (4)$$

where fw_i is the weight of each factor, and w_{ij} is the normalized weight for the category of the factor i .

The values of importance and normalized importance are presented in Table 4. It can be seen that the important values of TWI, slope aspect, and NDVI are 0.16, 0.14, and 0.12, respectively. The relevant normalized importance values are 100%, 87.5%, and 75%, respectively. The values of elevation, profile curvature, and plane curvature are much smaller, 0.06, 0.05, and 0.03, respectively. The results show that the factors associated with the wetness are the main causes of the landslide. The loess soil is the main component in the study area, which is extremely easy to lose stability under the influence of water.

7. Validation and Discussion

7.1. Validation. Based on the grid data map of susceptibility index obtained from the three models, the random point tool

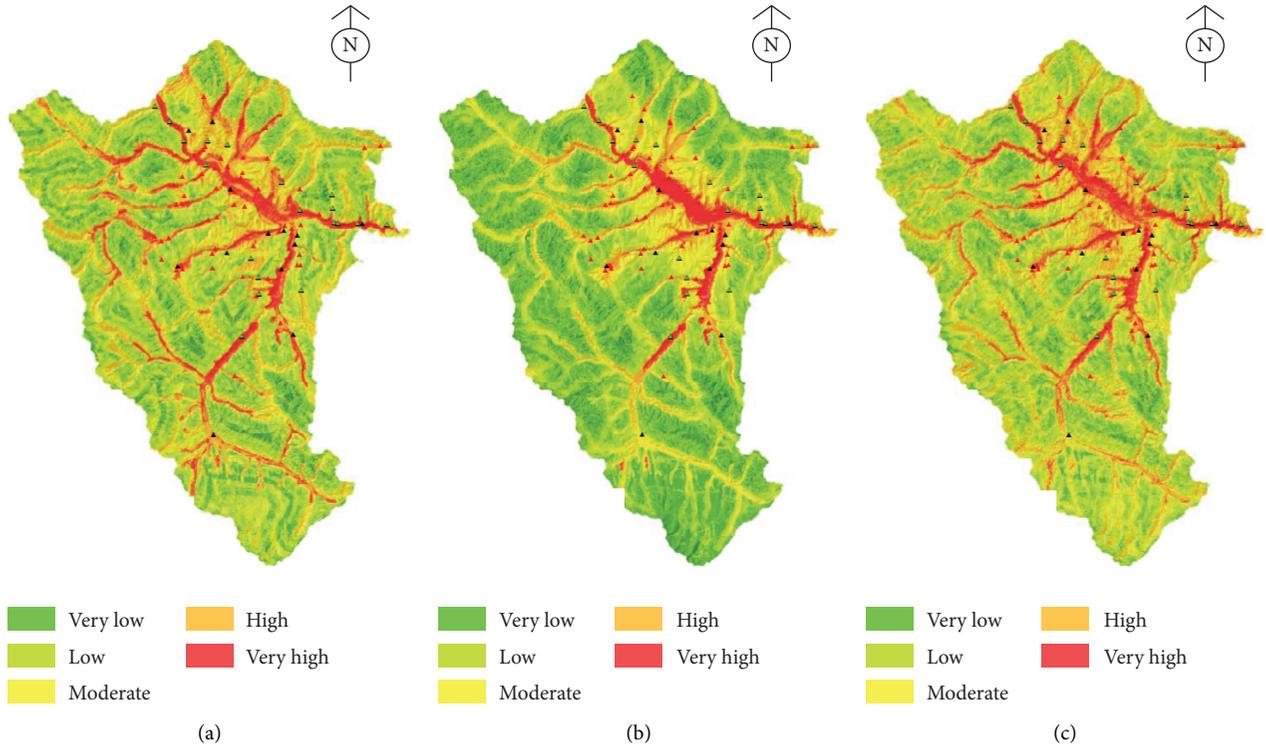


FIGURE 6: Landslide susceptibility maps. (a) IM model. (b) FR model. (c) ANN model.

TABLE 4: Normalized importance value derived from ANN method.

Number	Factors	ANN importance	ANN normalized importance (%)
1	TWI	0.16	100
2	Slope aspect	0.14	87.5
3	NDVI	0.12	75
4	Slope angle	0.1	62.5
5	Distance to road	0.1	62.5
6	Stratigraphic lithology	0.09	56.3
7	Distance to fault	0.08	50
8	Distance to river	0.07	43.8
9	Elevation	0.06	37.5
10	Profile curvature	0.05	31.3
11	Plane curvature	0.03	18.7

is created through ArcGIS 10.2, and 100 nonlandslide points with the same number of geological hazard points to be studied are randomly created and imported into SPSS software. The accuracy of the results of the above three models is evaluated through the ROC curve of the analysis tool in the software. The AUC value can be obtained by equation (5) [36]. The results are shown in Figure 7.

$$AUC = \int_0^1 \frac{\exp(a/(1-b))(x/(1-x))^{(1+b)/(1-b)}}{1 + \exp(a/(1-b))(x/(1-x))^{(1+b)/(1-b)}} dx, \tag{5}$$

where coefficients a and b represent the dependence of the test accuracy on threshold; x is the value of ROC.

It can be seen from Figure 7 that the ROC curve verification results of three different evaluation models are above

the reference line, which is in line with the feasibility expectation of ROC curve verification. The accuracy of the ANN model is the highest, and AUC is 0.907; the accuracy of FR model is the lowest, and AUC is 0.867. In this respect, the ANN model also has the highest predictive power. In terms of methods, it can be seen that the ANN model has a higher prediction ability than the IM model and FR model.

7.2. Discussion. The landslide susceptibilities were calculated by three methods based on the GIS engine, and the areas and percentages distribution of the susceptibility classes were produced. The susceptibility classes were divided into five classes, very low, low, medium, high, and very high risk. To compare the models in detail, the statistical results of the landslide susceptibility maps produced in this study are also

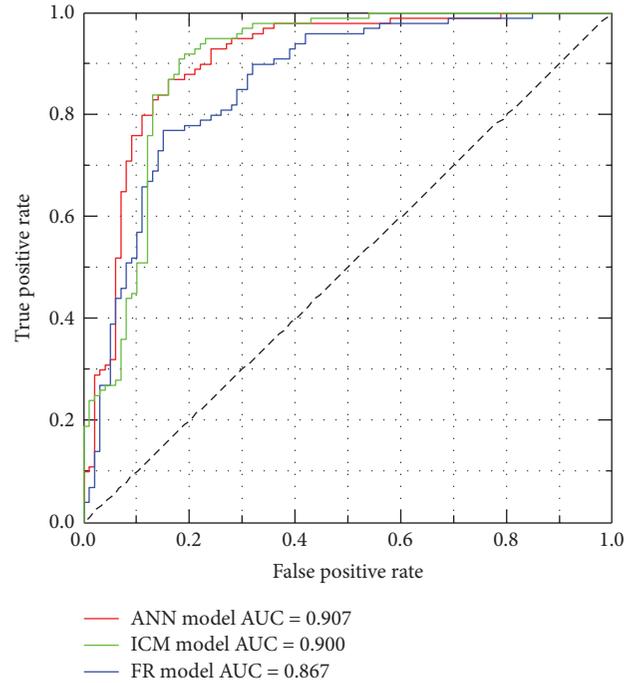


FIGURE 7: ROC curve of the models.

TABLE 5: Statistical results of landslide susceptibility mapping.

Models	Susceptibility	Class area (km ²)	Class ratio (%)	Landslides count	Landslides ratio (%)
IM	Very low	259.95	17.24	1	1
	Low	492.91	32.69	3	3
	Moderate	407.12	27.00	13	13
	High	229.04	15.19	29	29
	Very high	118.82	7.88	54	54
FR	Very low	471.95	31.30	0	0
	Low	529.86	35.14	2	2
	Moderate	309.71	20.54	14	14
	High	130.13	8.63	47	47
	Very high	66.19	4.39	37	37
ANN	Very low	215.95	14.32	0	0
	Low	460.46	30.54	1	1
	Moderate	442.66	29.36	4	4
	High	276.77	18.36	30	30
	Very high	112.00	7.45	65	65

listed in Table 5. The number of landslides included in the high and very high susceptibility classification can also reflect the predictive power of a model. The results show that the areas of five susceptibility classes are 259.95, 492.91, 407.12, 229.04, and 118.82 km², respectively, for the IM model. For the FR model, the areas of five classes are 471.95, 529.86, 309.71, and 130.13 km², respectively. For the ANN model, the landslide susceptibility areas of five classes are 251.95, 460.46, 442.66, 276.77, and 112 km², respectively, sh.

According to the landslide susceptibility map produced by the IM method, the very low and low landslide susceptibility areas are 17.24% and 32.69% of the total study area, respectively. Moderate, high and very high susceptible areas account for 27%, 15.19%, and 7.88% of the total area. The

landslide susceptibility generated by the FR method, which contains 31.3% of the total area, is determined to be very low landslide susceptibility. Low, moderate, and high areas make up 35.14%, 20.54%, and 8.63%, respectively. The very high landslide susceptibility area is 4.39% of the total study area. The landslide susceptibility map, produced through the ANN method, has similar results with IM method, but different with FR method. The very low, low, and moderate area are denoted at values of 14.32%, 30.54%, and 29.36% of the total study area, respectively. The high and very high susceptibility areas are 18.36% and 7.45% of the total area, respectively. The number of landslides divided into the high and very high classes characterizes the accuracy of the model. In this respect, the ANN method has the highest predictive ability.

According to the above analysis, the ANN model is the optimal model. Thus, the ANN model was adopted for landslide susceptibility map analysis in this study. In Figure 6, it can be seen that the very high susceptibility area is mainly distributed along the rivers and valleys. This is consistent with the landslide distribution along the river and valley found in our field investigation. It also can be seen that the very high susceptibility area is mainly distributed in three regions: (a) Yuandong-Dahua; (2) Dongga-Shangshigou; (3) Dongxia-Yuanju. In the study area, collapses and landslides are distributed on both sides of the valley or both sides of the river, and their disaster-causing effect and distribution density are closely related to the erosion and cutting of the valley and river. Generally, in the source and upstream of the gully, vertical erosion is the main factor, and the landslides and collapses occur frequently on both sides of the gully. In the middle and lower reaches of the valley, side erosion is the main cause. The unloading and weathering of the valley slopes on both sides of the valley are strong, and the slope on one side of the river erosion bank is prone to collapse and landslide.

8. Conclusion

In this study, landslide susceptibility for losses slopes in Huangyuan city is proposed. According to geological data, field survey, and landslides information, eleven influencing factors, which include elevation, slope, aspect, plane curvature, profile curvature, road distance, river distance, fault distance, stratum rock property, vegetation coverage index, and terrain humidity index, were selected for the landslide susceptibility of Huangyuan, Qinghai. The IM, FR, and ANN methods were adopted to establish the landslide susceptibility model, which is validated by the ROC curve.

Through the analysis of the 11 factors by three methods, every factor has a specific subclass with high landslide susceptibility. Based on the results of ANN method, the landslides in this study area are mainly influenced by factors associated with the wetness. Therefore, the specific subclass of every factor is conducive to the accumulation of water, which leads to the occurrence of landslide. The ANN model is the most optimal model, with an AUC value of 0.907, followed by the IM model (0.900) and FR model (0.867) in the study area. The high and very high landslide susceptibility classes make up 84%, 84%, and 95% of the total landslide count in the IM model, FR model, and ANN model, which also indicates that the ANN model is better than the others. From the landslide susceptibility map produced by the ANN model, the very high and high susceptibility areas are mainly distributed in valley area. Therefore, it is necessary to prevent and control landslide disasters in these areas by applying measures, such as slope protection, retention, and anchoring.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Applied Basic Research Project of Science and Technology Department of Qinghai Province (2018-ZJ-714) and Special Fund for Key Laboratory of Geological Process and Mineral Resources in Northern Qinghai Tibet Plateau (2019-KY-01).

References

- [1] E. Rotigliano, C. Cappadonia, C. Conoscenti, and D. Costanzo, V. Agnesi, Slope units-based flow susceptibility model: using validation tests to select controlling factors," *Natural Hazards*, vol. 61, pp. 143–153, 2012.
- [2] X. Sun, J. Chen, X. Han, Y. Bao, J. Zhan, and W. Peng, "Application of a GIS-based slope unit method for landslide susceptibility mapping along the rapidly uplifting section of the upper Jinsha River, south-western China," *Bulletin of Engineering Geology and the Environment*, vol. 79, no. 1, pp. 533–549, 2020.
- [3] X. J. Yang and L. D. Chen, "Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, pp. 487–495, 2010.
- [4] Y. Bao, S. Zhai, J. Chen et al., "The evolution of the Samaoding paleol and slide river blocking event at the upstream reaches of the Jinsha river, Tibetan plateau," *Geomorphology*, vol. 351, Article ID 106970, 2020.
- [5] G. C. Ohlmacher and J. C. Davis, "Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA," *Engineering Geology*, vol. 69, pp. 331–343, 2003.
- [6] M. S. Roodposhti, S. Rahimi, and M. J. Beglou, "PROM-ETHEE II and fuzzy AHP: an enhanced GIS-based landslide susceptibility mapping," *Natural Hazards*, vol. 73, pp. 77–95, 2014.
- [7] J. Ji, Y. Gao, Q. Lü, Z. Wu, W. Zhang, and C. Zhang, "China's early warning system progress," *Science*, vol. 365, p. 332, 2019.
- [8] Z. Chen, J. Du, J. Yan, P. Sun, K. Li, and Y. Li, "Point estimation method: validation, efficiency improvement, and application to embankment slope stability reliability analysis," *Engineering Geology*, vol. 263, Article ID 105232, 2019.
- [9] E. N. Bromhead, "Reflections on the residual strength of clay soils, with special reference to bedding-controlled landslides," *The Quarterly Journal of Engineering Geology and Hydrogeology*, vol. 46, pp. 132–155, 2013.
- [10] W. C. Li, F. C. Dai, Y. Q. Wei, M. L. Wang, H. Min, and L. M. Lee, "Implication of subsurface flow on rainfall-induced landslide: a case study," *Landslides*, vol. 13, no. 5, pp. 1109–1123, 2016.
- [11] X. H. Guo, Z. P. Lai, Z. Sun, X. L. Li, and T. B. Yang, "Luminescence dating of Suozi landslide in the upper Yellow river of the Qinghai-tibetan plateau, China," *Quaternary International*, vol. 349, pp. 159–166, 2014.
- [12] Y. S. Zhang, C. B. Guo, H. X. Lan, N. J. Zhou, and X. Yao, "Reactivation mechanism of ancient giant landslides in the tectonically active zone: a case study in southwest China," *Environmental Earth Sciences*, vol. 74, no. 2, pp. 1719–1729, 2015.

- [13] C. Yu and J. Chen, "Application of a GIS-based slope unit method for landslide susceptibility mapping in Helong city: comparative assessment of ICM, AHP, and RF model," *Symmetry*, vol. 12, no. 11, p. 1848, 2020.
- [14] L. Nahayo, E. Kalisa, A. Maniragaba, and F. X. Nshimiyimana, "Comparison of analytical hierarchy process and certain factor models in landslide susceptibility mapping in Rwanda," *Modeling Earth Systems and Environment*, vol. 5, pp. 885–895, 2019.
- [15] S. B. Bai, J. Wang, G. N. Lue, P. G. Zhou, and S. S. Hou, S. N. Xu, GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the three Gorges area, China," *Geomorphology*, vol. 115, pp. 23–31, 2010.
- [16] C. Cao, Q. Wang, J. Chen, Y. Ruan, L. Zheng, and S. S. C. Niu, "Landslide susceptibility mapping in vertical distribution law of precipitation area: case of the Xulong hydropower station reservoir, southwestern China," *Water*, vol. 8270 pages, 2016.
- [17] S. Lee, "Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data journals," *International Journal of Remote Sensing*, vol. 26, pp. 1477–1491, 2005.
- [18] A. Yalcin, S. Reis, A. C. Aydinoglu, and T. Yomralioglu, "A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey," *Catena*, vol. 85, pp. 274–287, 2011.
- [19] F. Wang, P. Xu, C. Wang, N. Wang, and N. Jiang, "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*, vol. 6, no. 172, 2017.
- [20] G. Du, Y. Zhang, Z. Yang, C. Guo, X. Yao, and D. Sun, "Landslide susceptibility mapping in the region of eastern Himalayan syntaxis, Tibetan plateau, China: a comparison between analytical hierarchy process information value and logistic regression-information value methods," *Bulletin of Engineering Geology and the Environment*, vol. 78, pp. 4201–4215, 2019.
- [21] W. Chen, M. Panahi, P. Tsangaratos et al., "Applying population-based evolutionary algorithms and a neuro-fuzzy system for modeling landslide susceptibility," *Catena*, vol. 172, pp. 212–231, 2019.
- [22] K. Khosravi, M. Panahi, and D. T. Bui, "Spatial prediction of groundwater spring potential mapping based on an adaptive neuro-fuzzy inference system and metaheuristic optimization," *Hydrology and Earth System Sciences*, vol. 22, pp. 4771–4792, 2018.
- [23] T. Lei, D. Xue, Z. Lv, S. Li, Y. Zhang, and A. K. Nandi, "Unsupervised change detection using fast fuzzy clustering for landslide mapping from very high-resolution images," *Remote Sensing*, vol. 101381 pages, 2018.
- [24] M. Mokarram and A. R. Zarei, "Landslide susceptibility mapping using fuzzy-AHP," *Geotechnical & Geological Engineering*, vol. 36, pp. 3931–3943, 2018.
- [25] P. Aleotti and R. Chowdhury, "Landslide hazard assessment: summary review and new perspectives," *Bulletin of Engineering Geology and the Environment*, vol. 58, no. 1, pp. 21–44, 1999.
- [26] C. Y. Zheng, *Detailed Investigation Report of Geological Hazards in Xining City, Qinghai Province (Huangyuan County)*, Qinghai Hydrogeological Engineering Geological Environment Geological Survey Institute, Xining, China, 2012, in Chinese.
- [27] Y. Achour and R. Hamid Pourghasemi, "How do machine learning techniques help in increasing accuracy of landslide susceptibility maps," *Geoscience Frontiers*, vol. 11, pp. 153–165, 2020.
- [28] Department of Natural Resources of Qinghai Province, China, *Qinghai Digital Basemap*, Department of Natural Resources of Qinghai Province, China, Xining, China, 2019.
- [29] A. Aditian, T. Kubota, and Y. Shinohara, "Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia," *Geomorphology*, vol. 318, pp. 101–111, 2018.
- [30] D. D. Kose and T. Turk, "GIS-based fully automatic landslide susceptibility analysis by weight-of-evidence and frequency ratio methods," *Physical Geography*, vol. 40, pp. 481–501, 2019.
- [31] P. Binh Thai, B. Dieu Tien, I. Prakash, and M. B. Dholakia, "Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS," *Catena*, vol. 149, pp. 52–63, 2017.
- [32] Y. Achour, A. Boumezbeur, R. Hadji, A. Chouabbi, V. Cavaleiro, and E. A. Bendaoud, "Landslide susceptibility mapping using analytic hierarchy process and information value methods along a highway road section in constantine, Algeria," *Arabian Journal of Geosciences*, vol. 10, p. 194, 2017.
- [33] A. Kumar, R. K. Sharma, and V. K. Bansal, "GIS-based comparative study of information value and frequency ratio method for landslide hazard zonation in a part of mid-Himalaya in Himachal Pradesh," *Innovative Infrastructure Solutions*, vol. 4, no. 18, 2019.
- [34] P. Jiang, Z. Zeng, J. Chen, and T. Huang, Zeng, "Generalized regression neural networks with K-Fold cross-validation for displacement of landslide forecasting," in *Advances in Neural Networks—International Symposium on Neural Networks 2014*, Y. Li and I. King, Eds., Springer International Publishing, Berlin, Germany, 2014.
- [35] R. Hecht-Nielsen, "Kolmogorov's mapping neural network existence theorem," vol. 3, pp. 11–14, in *Proceedings of the international conference on Neural Networks*, vol. 3, pp. 11–14, IEEE Press, San-Diego, CA, USA, January 1987.
- [36] S. D. Walter, "Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data," *Statistics in Medicine*, vol. 21, pp. 1237–1256, 2002.