

Assessment of Landslide Susceptibility Based on Weighted Information Value Model in Pingshan County

Yuanji Yang^{1,2}, Wenjun Zhang^{1,2}

¹School of Environment and Resources, Southwest University of Science and Technology, Mianyang, China ²Mianyang S&T City Division National Remote Sensing Center of China, Mianyang, China Email: 1414295352@qq.com

How to cite this paper: Yang, Y.J. and Zhang, W.J. (2021) Assessment of Landslide Susceptibility Based on Weighted Information Value Model in Pingshan County. Open Access Library Journal, 8: e8199. https://doi.org/10.4236/oalib.1108199

Received: November 17, 2021 Accepted: December 27, 2021 Published: December 30, 2021

Copyright © 2021 by author(s) and Open Access Library Inc. This work is licensed under the Creative **Commons Attribution International** License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/ ۲

Open Access

Abstract

This article takes Pingshan County of Yibin City as an example. After collecting various basic data and reviewing the literature, the slope, aspect, Normalized Difference Vegetation Index (NDVI), distance from water system, distance from road, stratum, distance from fault, rainfall 8 kinds of influencing factors are used as evaluation factors. The index of entropy model (index of entropy, IOE) and information values (IV) were used to evaluate the susceptibility of landslide disasters in the study area. On this basis, a weighted information values model (WIV) was proposed. Using the AUC area under the ROC curve to test and compare the three models, the results show that the success rates of the three models of IOE, IV and WIV are 76.2%, 75.6%, and 79.1%, respectively. Finally, according to the natural discontinuity method, the landslide susceptibility index calculated by ArcGIS software is divided into extremely high, high, medium, and low susceptible areas, and the landslide susceptibility zoning map of Pingshan County is obtained, which can provide reference for relevant departments for disaster prevention and mitigation.

Subject Areas

Environmental Sciences

Keywords

Landslide, Susceptibility Assessment, Index of Entropy Model, Weighted Information Values Model, ROC Curve

1. Introduction

Landslide disasters are very serious all over the world, and our country is one of

the countries with the most landslide disasters. Every year, the casualties and economic losses caused by landslide disasters are very large. In recent years, with the acceleration of my country's economic development, the development of mountainous areas has also increased, and the number and scale of landslides have also increased sharply. Pingshan County, Yibin City is a mountainous area. The increasing number of landslide disasters has seriously affected social stability and economic development, and threatened people's lives. Therefore, the evaluation of landslide susceptibility in this area can provide a scientific basis for local relevant departments to carry out disaster prevention and mitigation and carry out various engineering activities.

Landslide susceptibility evaluation research lays the foundation for subsequent prediction, prevention and management of landslide disasters, and it is also an important scientific basis for the orderly progress of various engineering activities. At present, Chinese and foreign scholars have done a lot of research on the evaluation of landslide susceptibility, and put forward many different evaluation methods. These methods can be roughly distinguished based on qualitative and quantitative, and quantitatively divided into semi-quantitative evaluation and quantitative evaluation [1]. AHP, logistic regression method [2], information method [3], weight of evidence method, frequency ratio method, support vector machine model [4], random forest model, artificial neural network model [5], etc. are all relatively common evaluation method. Using different methods to evaluate the same research area often results in different evaluation results, and whether the advantages of different models can be combined to make the evaluation method more scientific and reasonable has become the focus of research by many scholars. Based on this, based on the use of entropy index model and information method to evaluate the landslide susceptibility of the study area, a weighted information model is constructed, and the first-level index weight of the entropy index model is multiplied by the second-level index weight of the information method. In this way, the comprehensive weight of the study area is obtained, which is closer to the true weight of the data, and the result calculated by it is more scientific and reasonable.

2. Overview of the Study Area

Pingshan County is located in the west of Yibin City, Yibin County to the east, Suijiang County of Zhaotong City, Yunnan Province to the south across the Jinsha River, Leibo County and Mabian Yi Autonomous Prefecture to the west, and Muchuan County to the north. Pingshan County has jurisdiction over 8 towns and 3 townships (including 2 Yi townships). The geographic location is the southern margin of the Sichuan Basin, on the north bank of the lower reaches of the Jinsha River, at 103°36'31" - 104°23'36" east longitude and 28°28'26" -28°53'51" north latitude [6]. The whole area is narrow from north to south and wide from east to west, about 47 KM wide from north to south, and 78 KM long from east to west. The total area of the district is about 1504 KM². The county's annual average temperature is 14.9°C, and the annual average rainfall is 1066 mm. It belongs to the mid-subtropical humid monsoon climate zone, with large vertical climate differences. There are many rivers in the county area, mainly including the Jinsha River and the Minjiang River, as well as larger rivers such as Xining River, Zhongdu River and Longxi River.

In terms of geomorphology, the eastern part of Pingshan County is the hills of southern Sichuan, the west and north are connected with the mountains of southwest Sichuan, and the south is connected with the eastern Yunnan plateau [7]. Most areas in the county are 300 - 1400 meters above sea level. The central and western regions of the county are mountainous landforms, and the hilly landforms are mainly distributed in the northeast of the county. Geologically, the outcrop of Pingshan County is dominated by the Jurassic, accounting for about 65% of the county's area. The Triassic, Cretaceous, and Permian systems are also exposed. The geological structure system is composed of the Maoshui anticline, the Wuzhishan compound anticline and the Zhongduhe syncline. Stratigraphic faults are mainly distributed in the northeast edge, northwest edge and southwest edge area, and there are few faults in general. There are 143 historical landslides in the area, including 59 small landslides, 73 medium landslides and 11 large landslides. Landslides mainly occurred in Xin'an Town, Jinping Town, Fuyan Town, etc. The vast majority of landslide locations contain groundwater as fissure water or pore water.

3. Selection of Evaluation Model

3.1. Information Model

The information model is based on the actual occurrence of historical landslide disasters in the study area and the information provided by various geological surveying and mapping data, and transforms the factors that have a greater impact on the stability of the landslide into the information value reflecting the stability of the landslide in the study area. The greater the information value, the greater the degree of correlation between the influencing factor and the occurrence of landslide disasters [8]. The information value is usually calculated by conditional probability, but in actual calculations, due to the limitation of the number of landslide disasters, the statistical frequency of landslide sample data is generally used to approximate the conditional probability.

$$I(X_i, Y) = \ln \frac{N_i/N}{S_i/S}$$
(2)

In formula (2): N_{i}/N is the factor corresponding to the proportion of graded landslides, and S_{i}/S is the factor corresponding to the proportion of graded area.

Generally speaking, landslide disasters occur under the combined action of multiple influencing factors. Therefore, the sum of the information values of various influencing factors in the same study area is the total information value of the study area.

$$I_i = \sum_{i=1}^{p} \ln\left(\frac{N_i/N}{S_i/S}\right)$$
(3)

3.2. Entropy Index Model

The entropy index model is a classification prediction model based on probability theory and mathematical statistical analysis. It can also be used as the input data of a mixed model. The model can calculate the weight of the variable in the event through the attribute value of each input variable, and the weight of each variable is represented by the entropy index. The value of the entropy index is positively correlated with the effect of the influence factor on the landslide. The weight value is calculated by the following formula:

$$P_{ij} = \frac{b}{a} \tag{4}$$

$$\left(P_{ij}\right) = \frac{P_{ij}}{\sum_{j=1}^{s} P_{ij}}$$
(5)

$$H_{i} = -\sum_{j=1}^{s} \left(P_{ij} \right) \log_{2} \left(P_{ij} \right), \quad i = 1, 2, \cdots, n$$
(6)

$$H_{i\max} = \log_2 s \tag{7}$$

$$I_i = \frac{H_{i\max} - H_i}{H_{i\max}}, \ i = 1, 2, \cdots, n$$
 (8)

$$P_{i} = \frac{1}{S_{i}} \sum_{j=1}^{s} P_{ij}$$
(9)

$$W_i = I_i P_i \tag{10}$$

In the above formula: *i* is the evaluation factor; *j* is the grade of each factor; *a* and *b* are the area and the percentage of the number of landslides; P_{ij} is the frequency ratio of the factor grade; (P_{ij}) is the probability density; H_i is the entropy value; H_{imax} is the maximum of the corresponding factor Entropy; *s* is the number of factor levels; I_i is the information rate; W_i is the entropy index value.

3.3. Weighted Information Model

The entropy index model only considers the first-level index weight of the impact factor, and the information model only considers the relative weight of the second-level index of the impact factor, which cannot reflect the true weight of each factor. Based on this, the model multiplies the weight calculated by the entropy index model with the weight calculated by the information volume model to obtain the comprehensive weight of the factor, which is closer to its true weight. Then the weight values of each factor are superimposed and summed to obtain the comprehensive weight value of the entire study area, that is, the landslide hazard susceptibility index LSI. The specific calculation formula is as shown in formula (11):

$$\mathrm{LSI}_{\mathrm{WIV}} = \sum_{i=1}^{F} W_i I V_{ij} \tag{11}$$

F is the number of impact factors.

4. Selection and Analysis of Evaluation Factors

4.1. Data Preparation

All data mainly include the Pingshan County landslide catalog database, 30 m \times 30 m Digital Elevation Model (DEM), road network map, water system map, geological map with a scale of 1:50,000, remote sensing image data of Sentinel No. 2 with a resolution of 10 m, and rainfall. Among them, DEM is downloaded from the cloud of geospatial data, geological maps are downloaded from 91 Satellite Map Assistant, remote sensing images are downloaded from the ESA website, and rainfall maps are obtained from rainfall site data downloaded by China Meteorological Data Network through Kriging interpolation. And randomly select non-landslide points equal to the number of landslide points.

4.2. Selection of Evaluation Factors

The selection of evaluation factors is a very critical step in the evaluation of landslide hazard susceptibility. The incorrect selection of evaluation factors will affect the accuracy of the evaluation results. Too few evaluation factors will cause the evaluation model to be under-fitted and affect the evaluation accuracy. Too many numbers will cause over-fitting of the model, reduce the calculation speed of the entire model, and also affect the accuracy of the model.

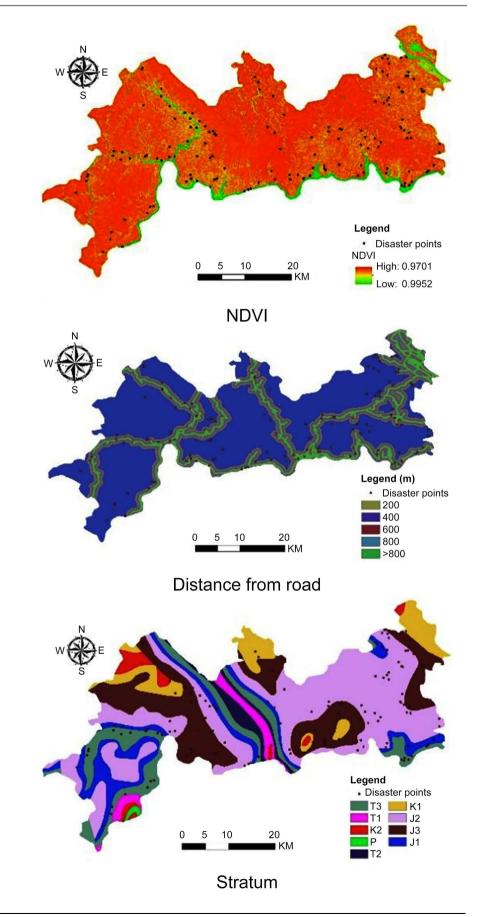
Based on existing data and field surveys, combined with predecessors' experience in selecting factors affecting landslide hazards [9] [10] [11] [12], 9 susceptibility assessment indicators were selected comprehensively, namely slope, aspect, elevation, and normalization. Vegetation Index (NDVI), stratum, distance from fault, distance from water system, distance from road, rainfall.

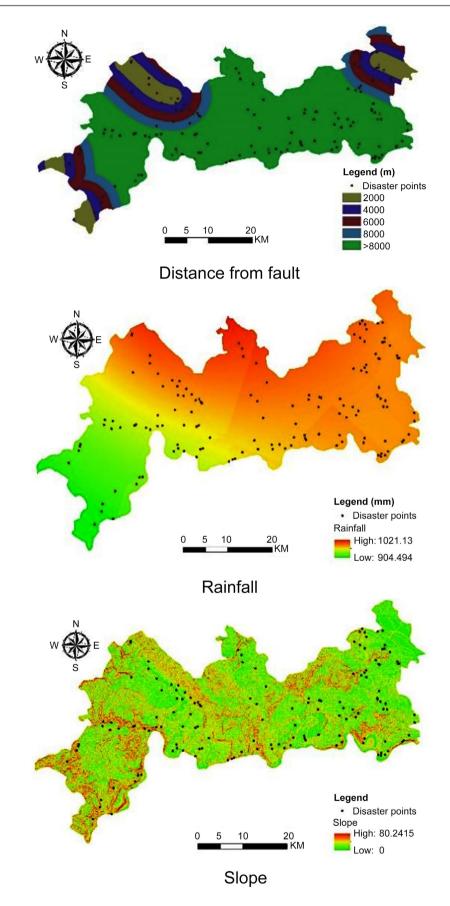
4.3. Evaluation Factor Classification

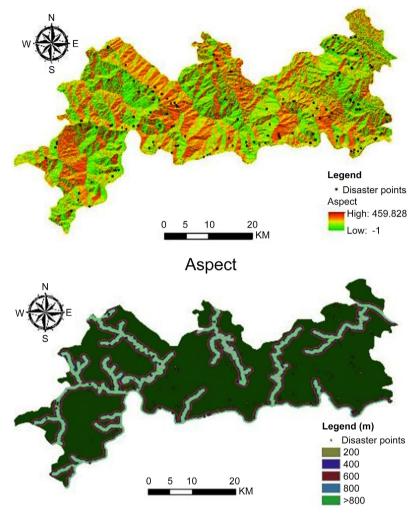
With reference to previous experience [13] [14], combined with the actual situation of the study area, the distance from the water system and the distance from the road are classified according to 200 m, 400 m, 600 m, 800 m and >800 m; the distance from the fault is divided into 5 levels according to the interval of 2000 m; Use the equal interval method to divide the slope into 8 parts; divide the stratum into 9 levels according to the actual stratigraphic conditions; other influencing factors such as rainfall and NDVI are divided into 5 levels according to the natural discontinuity method. The specific classification is shown in **Figure** 1.

4.4. Correlation Analysis of Evaluation Factors

There are many factors that affect the occurrence of landslides. Some of the influencing factors may have a high correlation between the two, and the relatively







Distance from water system

Figure 1. Influence factor extraction graph.

high correlation factors may reduce the operating speed of the model and make the model more complicated. Reduce the accuracy of the entire model. Therefore, when evaluating the susceptibility of landslide hazards, the selected evaluation factors must have low correlation or irrelevance. This paper uses Kendall's correlation evaluation method, and the correlation coefficient is represented by *R*. When the absolute value of *R* is less than |R| < 0.3, it indicates that the correlation between factors is weak, and the factor can be used for landslide susceptibility evaluation. In addition, the correlation between factors is required to have significant reference value, that is, the significance level of correlation between factors sig < 0.05. The data of each factor obtained after ArcGIS software processing is imported into SPSS software for correlation analysis to obtain the correlation coefficient between factors (**Table 1**). It can be seen from **Table 1** that the correlation coefficient between elevation and NDVI, and the correlation coefficient between elevation and distance from the road are all too high, so the elevation factor is removed.
 Table 1. Factor correlation coefficient.

	DEM	slope	aspect	NDVI	road	stratum	fault	water system	rainfall
DEM	1								
slope	0.161	1							
aspect	0.064	0.070	1						
NDVI	0.364	0.191	0.049	1					
road	0.368	0.102	0.046	0.261	1				
stratum	-0.037	-0.051	0.103	-0.050	0.094	1			
fault	0.110	0.047	-0.011	0.059	-0.061	-0.103	1		
water system	0.282	-0.017	0.091	0.199	0.273	0.057	0.121	1	
rainfall	-0.058	-0.118	0.062	0.260	0.056	0.040	-0.100	0.039	1

5. Evaluation Results of Landslide Hazard Susceptibility 5.1. Division of Evaluation Unit

Before evaluating the landslide susceptibility of the study area, in order to make subsequent calculations simple and convenient and ensure the scientificity of the results, it is necessary to re-divide the entire study area and divide all evaluation factors into a unified measurement. Because the grid unit has the advantages of being simple and easy to superimpose, it is widely used when dividing the research area. Use the empirical formula obtained from the DEM test to calculate the appropriate grid size [15]:

$$G_{\rm s} = 7.49 + 0.0006S - 2.0 \times 10^{-9} S^2 + 2.9 \times 10^{-15} C^3$$
(12)

In formula (12): *C* is the scale of the topographic map.

Through calculation, a grid unit with a size of 30 m \times 30 m is selected as the grid in this study area, with a total of 2,166,304 grid units.

5.2. Evaluation Results of the Information Volume Model

The information value IV_{ij} of each factor calculated by the above formula is shown in **Table 3**. It can be seen from **Table 3** that the NDVI is 0.41 - 0.65, and the study area with a distance of less than 200 meters from the road has the largest information value corresponding to the classification. Under this condition, it is most likely to cause landslide disasters.

In ArcGIS software, the information value of different evaluation factors in the study area is superimposed to obtain the total information value of the entire study area, that is, the landslide hazard susceptibility index LSI, which ranges from -4.85 to 4.22. Use the natural discontinuity method to classify the LSI, which are extremely high, high, medium, and low prone areas, and draw a zon-ing map of the proneness level, as shown in **Figure 2**:

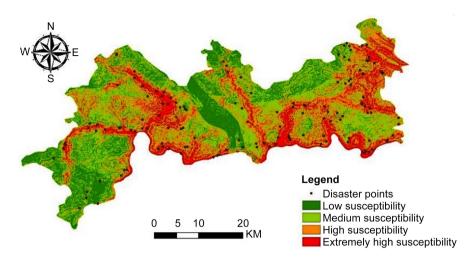


Figure 2. Landslide susceptibility based on IV model.

5.3. Evaluation Results of the Entropy Index Model

The relative weight (W_i) of the first-level index of each evaluation factor obtained by using the above-mentioned entropy index model calculation formula is shown in **Table 2**. It can be seen from **Table 2** that the main factors affecting the occurrence of landslide disasters in Pingshan County are slope, NDVI, strata, and distance from the water system. The LSI in the study area can be calculated by formula (13):

$$LSI_{IOE} = (NDVI \times 0.11) + (slope \times 0.12) + (rainfall \times 0.04) + (aspect \times 0.06) + (water system \times 0.07) (13) + (fault \times 0.01) + (stratum \times 0.08) + (road \times 0.06)$$

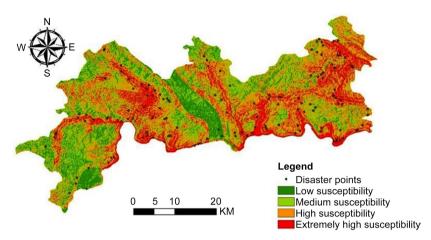
The value range of Pingshan County LSI calculated by the entropy index model is 0.72 - 3.22. The LSI is classified using the natural discontinuity method, which are extremely high, high, medium, and low prone areas, and the landslide proneness level is drawn. Figure, as shown in **Figure 3**.

5.4. Evaluation Results of Weighted Information Model

The weighted information value under the classification state of each evaluation factor calculated by formula (11) is shown in **Table 3**. The calculated value of LSI ranges from -0.41 to 0.34. The natural discontinuity method is used to classify LSI into extremely high, high, medium, and low prone areas, and a zoning map of the landslide proneness level is drawn, as shown in **Figure 4**.

6. Inspection and Comparison of Evaluation Models

This article uses the ROC curve (success rate curve) method to evaluate the model [16]. Use the ROC curve analysis module in SPSS software to analyze the accuracy of the research results. Arrange the susceptibility index LSI of each evaluation unit in the entire study area in descending order to obtain the cumulative area percentage as the horizontal axis, and use the cumulative percentage of the number of landslides in the corresponding range of the landslide suscepti-



bility index as the vertical axis to obtain the landslide proneness of the study area The success rate curve of the sex zoning diagram is shown in **Figure 5**.

Figure 3. Landslide susceptibility based on IOE model.

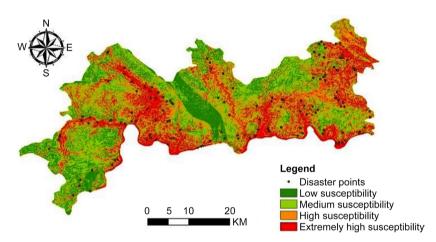


Figure 4. Landslide susceptibility based on WIV model.

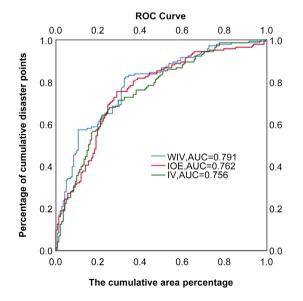


Figure 5. The success rate curve.

Table 2. Index of entropy value.

Factor	slope	NDVI	rainfall	aspect	Water system	fault	stratum	road
IOE W _i	0.12	0.11	0.04	0.06	0.07	0.01	0.08	0.06

Table 3. Information value and weighted information value.

Factor	Grade	IVij	WiIVij	Factor	Grade	IV_{ij}	WiIVij
	0 - 11.92	0.15	0.02	Water system	200	0.47	0.03
Slope	11.92 - 21.00	0.35	0.04		400	0.41	0.03
	21.00 - 30.40	-0.21	-0.03		600	-0.57	-0.04
	30.40 - 41.69			system	800	0.69	0.05
	41.69 - 80.25	-1.02	-0.12		>800	-0.21	-0.01
	<-0.10 0.82 0.09		2000	0.02	0.00023		
	-0.10 - 0.41	0.02	0.002		4000	0.03	0.00034
NDVI	0.41 - 0.65	1.00	0.11	Fault	6000	0.25	0.0026
	0.65 - 0.79	0.70	0.08		8000	-0.33	-0.003
	0.79 - 0.98	-0.59	-0.07		>8000	-0.01	0.0001
	<935.02	-0.72	-0.03		T_3	0.32	0.03
	935.02 - 957.80	-0.30	-0.01		T_1	-1.4	-0.12
Rainfall	957.80 - 978.30	0.12	0.004		K_2	-0.56	-0.05
	978.30 - 993.33	0.26	0.009		T_2	-1.42	-0.12
	>993.33	-0.57	-0.02	Stratum	K_1	-0.73	-0.06
	North	0.02	0.001		J_2	0.03	0.003
	North east	0.34	0.02		J_3	0.24	0.02
	East	-0.17	-0.01		J_1	-0.15	-0.01
	Southeast	0.18	0.01		Р	0	0
Aspect	South	0.02	0.001		200	0.87	0.05
	South West	South West -0.19 -0.01			400	0.19	0.01
	West	0.73	0.04	Road	600	0.27	0.02
	Northwest	-0.04	-0.0002		800	0.50	0.03
					>800	-0.39	-0.02

AUC (area under the curve) is used to evaluate the accuracy of the model. The larger the AUC, the closer the simulated value is to the true value of the sample, and the more accurate the model evaluation result is. It can be seen from **Figure 5** that the AUC values of the three models are all higher, and the weighted information model > entropy index model > information method. By calculating the number of landslide hazards that fall in high and extremely high-prone areas, it is found that the weighted information model accounts for 86.71%, which is larger than the information model's 76.22% and the entropy index model's 78.32%. It can be seen from **Table 3** that in the weighted information

model, the main influencing factor of the landslide is NDVI, the distance from the road and the slope, which are consistent with the results of the information model and the entropy index model. This shows that the weighted information model has better accuracy.

7. Conclusions

1) This paper uses three models of IOE, IV and WIV to evaluate the susceptibility of landslides in Pingshan County. The susceptibility zoning map of the corresponding model is obtained by the ArcGIS raster overlay, and the accuracy of the three models is 76.2%, 75.6%, and 79.1% after ROC curve inspection. Among them, the susceptibility zoning map generated by the WIV model has the highest accuracy, indicating that this model is more suitable for landslide hazard susceptibility evaluation research in this type of study area.

2) Based on the topography and geomorphology characteristics of the study area, this paper selects 9 evaluation factors, and after correlation analysis, the slope, aspect, NDVI, rainfall, distance to road, distance to fault, distance to water system, and strata are selected. Evaluation factors establish a landslide susceptibility evaluation system. According to the calculation of the model, the factors that account for the higher weight of the evaluation factors are NDVI, distance from the road, and slope. It can be concluded that landslide hazards are mainly concentrated in areas with medium vegetation coverage, short distance from the road, and gentle slope.

3) According to the zoning map of landslide hazard susceptibility, most areas of Taiping Township, Pingshan Town, Fuyan Town, Jinping Town, and Xin'an Town are located in areas with high landslide hazard and extremely high proneness, which is roughly consistent with the actual situation. The proportion of high-density landslide disaster areas is almost the same as that of high-risk areas.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- Zhou, P.H., Li, Y.C., Bai, J., *et al.* (2017) Susceptibility Assessment of Landslide Hazard to Catchment Unit in Xinyuan County. *Science of Surveying and Mapping*, 42, 82-87.
- [2] Wang, N.-Q., Guo, Y.-J., Liu, T.-M., et al. (2017) Assessment of Landslide Susceptibility Based on SVM-LR Model: A Case Study of Lintong District. Science Technology and Engineering, 42, 82-87.
- [3] Li, Y.R., Adnan, A., Xiang, X.Q., Ju, N.P., Zhao, J.J. and Ahmet, O. (2012) Landslide Susceptibility Mapping and Evaluation along a River Valley in China. *Acta Geologica Sinica (English Edition)*, 86, 1022-1030. https://doi.org/10.1111/j.1755-6724.2012.00726.x
- [4] Zhang, J., Yin, K.L., Wang, J.J., *et al.* (2016) Evaluation of Landslide Susceptibility for Wanzhou District of Three Gorges Reservoir. *Chinese Journal of Rock Mechanics and Engineering*, 35, 284-296.

- [5] Chen, L.-H., Ye, J., Wei, C.-J., et al. (2016) Application of ArcGIS and Information Method to Landslide Susceptibility Evaluation. *Journal of Guangxi University* (*Natural Science Edition*), 41, 141-148.
- [6] Wang, W.D., Liu, P. and Gong, L. (2019) Landslide Susceptibility Mapping of Sichuan Province Based on Support Vector Machine. *Journal of Railway Science and Engineering*, 16, 1194-1200.
- [7] Abudikeyinu, X. and He, S. (2019) Spatial Prediction on Landslide Vulnerability Based on MIV-BP Neural Network. *Yangtze River*, 50, 140-144.
- [8] Wei, L.-S., Kong, D.-Y., Luo, Y.-F., *et al.* (2018) Groundwater Occurrence Environment in the Red Beds of Southern Sichuan and the Hydro-Chemical Characteristic: A Case Study in Pingshan County, Yibin City. *Exploration Engineering (Rock & Soil Drilling and Tunneling)*, **45**, 145-150+115.
- [9] Zhang, X.Y., Zhang, C.S., Meng, H.J., et al. (2018) Landslide Susceptibility Assessment of New Jing-Zhang High-Speed Railway Based on GIS and Information Value Model. Journal of Geomechanics, 24, 96-105.
- [10] Wu, S.Y. (2019) Susceptibility Evaluation of Geological Hazards of Collapse and Landslides in Xiamen City Based on Information Model. *Safety and Environmental Engineering*, 26, 22-27.
- [11] Tian, S.J., Zhang, S.S., Tang, Q.S., *et al.* (2019) Comparative Study of Landslide Susceptibility Assessment Based on Different Evaluation Units. *Journal of Natural Disasters*, 28, 137-145.
- [12] Ilia, I. and Tsangaratos, P. (2016) Applying Weight of Evidence Method and Sensitivity Analysis to Produce a Landslide Susceptibility Map. *Landslides*, 13, 379-397. <u>https://doi.org/10.1007/s10346-015-0576-3</u>
- [13] Huang, M.-W., Chen, C.-Y., Wu, T.-H., Liu, S.-Y. and Kao, C.-Y. (2012) GIS-Based Evaluation on the Fault Motion-Induced Coseismic Landslides. *Journal of Mountain Science*, 9, 601-612. <u>https://doi.org/10.1007/s11629-012-2434-2</u>
- [14] Liao, L.P., Yu, M., Wen, H.T., et al. (2019) Evaluation on the Susceptibility of Collapse and Landslide in Rongxian County, Southeastern Guangxi. Earth and Environment, 47, 518-526.
- [15] Guo, T.S., Zhang, J.Q., Han, Y., *et al.* (2019) Evaluation of Landslide Susceptibility in Yanchang County Based on Particle Swarm Optimization-Based Support Vector Machine. *Bulletin of Geological Science and Technology*, **38**, 236-243.
- [16] Thanh Thuy Le, T. and Kawagoe, S. (2019) Evaluation of Landslide Susceptibility in Cau River Basin Using a Physical-Based Model under Impact of Climate Change. *Open Journal of Modern Hydrology*, 9, 1-19. <u>https://doi.org/10.4236/ojmh.2019.91001</u>