Risk Assessment of Debris Flow in Ya'an City Based on BP Neural Network

Xiaoyu Liu, Xiaolong Shu, Xing Liu*, Zhe Duan and Zhiyi Ran

College of Water Resources and Hydropower, Sichuan Agricultural University, Ya'an 625014

*Xing Liu’s e-mail: hydro_lx@sicau.edu.cn

Abstract. Since the occurrence of debris flow hazards is related to many factors, this paper selects five evaluation indicators: rainfall, lithology level, slope gradient, undulation degree and distance from faultages based on the three scopes of geomorphology, meteorology and geology. In this paper, the data extracted from the occurrence point of historical debris flow disaster is used, and BP neural network is used for training, then the weights of evaluation indicators are obtained. Finally, the GIS software is used to superimpose and analyze these indicators, and the risk of debris flow in Ya'an city is divided into five levels, thus the danger distribution map of debris flow in Ya'an city is obtained. And the map is compared with historical debris flow disaster occurrence points, the results show that, three districts of Tiancheng County, Lushan County and Baoxing County have the highest overall risk; the overall risk of Yingjing County, Yucheng District and Shimian County is high; and the overall risk of Hanyuan County and Mingshan District is relatively low. The evaluation results reflect the actual situation of the spatial distribution of debris flow in Ya'an city, which can provide a scientific basis for the development of reasonable flood control and disaster against planning in Ya'an city.

1. Introduction

Debris flow is one of the most common mountain geological disasters. It is a rapid surface runoff formed by saturated dilution of soft soil hills, which is caused by short-term heavy rainfall and other bad weather. Debris flow carries a large amount of solid detrital materials, such as sediment and rocks, flows to low-lying areas and causes loss of life and property, which is characterized by sudden developed, quickly formed, serious harm and difficulty in monitoring [1]. China is a mountainous country with nearly 70% of the country's land area is mountainous. The number of mudslides recorded alone has reached 111,000, which is spread over mountainous areas of 26 cities and is a large number and a wide range. As a whole, the average annual direct economic loss is huge. Among the areas mentioned above, Yunnan, Sichuan and Gansu are serious disaster areas for landslides and debris flows. These areas are mostly located in the seismic belt on the edge of the Qinghai-Tibet Plateau. In addition, the precipitation of the mountains on the surface of the Tibetan Plateau leads to the loosening of the surface structure, which forms a large amount of loose deposit strata. Moreover, human improper activities damage the environment, causing frequent mudslide occurrences.

In 1982, Chinese scholar Wang Lixian proposed a method for dividing the risk of debris flow valleys [2]; in 1986, Tan Bingyan proposed that 15 basic indicators should be selected from the direction of determining the severity of debris flow gully and scored according to the requirements of quantitative theory [3]; in 1999, Tang Jiafa used GIS software to classify the debris flow in the upper
reaches of the Minjiang River [4]; in 2002, Li Huixia evaluated and classified the risk degree of soil and water loss in Ya'an area from two aspects of the status and driving force of soil and water loss, and classified the risk degree of soil and water loss in Ya'an city [5]; in 2004, Liu Xilin selected a number of indicators to the debris flow through a comparative analysis of the vulnerability of mudslides in Sichuan Province in 1990 and 2000. The quantitative calculation of disaster risk shows that Ya'an city is a high-risk area for mudslide disasters in Sichuan Province [6]; in 2011, Wang Xueliang used analytic hierarchy process (AHP) to select 10 index factors to construct a hierarchy index system for debris flow risk evaluation [7]; in 2013, Zhao Xin used multiple regression analysis to establish a single-ditch debris flow risk assessment model [8]; in 14 years, based on GIS spatial modeling, Jin Rong selected four factors combined with analytic hierarchy process to evaluate the risk of debris flow in Shimian County and Hanyuan County of Ya'an city [9]; in 2015, taking the debris flow in Dayyelong gully, Tianquan county, Ya'an city as the research object, Wang Jiazhu calculated the dynamic characteristics of debris flow and made a quantitative evaluation of its risk [10].

With the development of artificial intelligence technology, object evaluation based on machine deep learning has become a future development trend. This paper will combine BP neural network and GIS software to analyze various debris occurrences of debris flow disasters, and evaluate the risk of debris flow in Ya'an city, which can provide a scientific basis for scientifically, reasonably predicting and coping with the occurrence of debris flow in this area.

2. Materials and Methods

2.1 Overview of the study area

Ya'an city is located at the western edge of the Sichuan Basin. It spans north latitude 28°51′10″–30°56′40″ and east longitude 101°56′26″–103°23′28″, which is the transition zone from the Sichuan Basin to the Qinghai-Tibet Plateau. The terrain is higher in the north, west and south, and lower in the middle and east. The mountain ranges in the city, the land surface is rugged, and the types of landforms are complex and diverse. There are many mountains, few hills and flat dams. The climate type is subtropical monsoon humid climate, which has large difference in north and south. The annual average temperature is between 14.1 °C to 17.9 °C, the average annual rainfall is about 1800 mm, and the resident population is about 1,537,800.

The causes of debris flow hazards are complex, and there are many factors affecting and controlling the risk of debris flow. Therefore, the selection of evaluation indicators should follow the principles of scientific, operational, hierarchical, objective, comparability, qualitative and quantitative [11]. The establishment of the evaluation index system should meet the principles of applicability, completeness, comparability, comprehensiveness and systematicness [12]. Therefore, according to the previous research results [13], and combined with the actual situation of the study area and GIS spatial analysis tools, the selected evaluation indicators are: lithology level, undulation degree, slope gradient, distance from faultage and rainfall.

2.2 Data processing

The terrain topographic evaluation indicators are extracted from “ASTGTM 30m resolution data”, and the slope gradient and the undulation degree are obtained from research area DEM by the GIS software. The original data of the geological evaluation indicators are derived from the “2.5 million Chinese geological map”, the lithology level of the stratum is obtained from the lithology data in the geological map through GIS software, and then the solidity coefficient is divided into five grades according to the "Rock Classification and Hardness Level", the distance from the faultage is obtained through the fault file to establish a grading buffer; hydrometeorological evaluation factor uses the monthly rainfall data of 0.25°×0.25° for the “TRMM 3B43 satellite” from 1998 to 2016. The average rainfall is calculated by MATLAB software. The measured data of the debris flow originated from “Sichuan Province in 1997”. The spatial distribution data of the 1:10 000 000 debris flow in Ya'an city and the “Handheld Query System for the Hidden Danger Points of Sichuan Province”.
indicators, the slope and rainfall are positive correlation indicators, that is, the larger the index value, the higher the risk; the undulation, lithology level and distance from the fault are negative correlation indicators, that is, the more the index, the smaller the risk.

Because the indicators are not uniform and the range of values varies greatly, which is not conducive to neural network training. It is necessary to standardize the data and eliminate its dimensions, and the formula is as follows [14]:

\[ X = \frac{X_i - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \]  

(1)

where \( X \) is the standard value of the indicator; \( X_i \) is the measured value of a certain indicator; \( X_{\text{max}} \) and \( X_{\text{min}} \) are the maximum and minimum values of the indicator respectively.

If the indicator is negatively correlated with risk, the reverse indicator is normalized to:

\[ X = 1 - \frac{X_i - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \]  

(2)

After normalizing, the indicators are divided into five grades according to the risk according to the degree of influence on the risk of debris flow. The higher the grade, the greater the risk (Table 1):

| Level            | I     | II    | III   | IV    | V     |
|------------------|-------|-------|-------|-------|-------|
| Lithology level  | 0     | 0.247 | 0.498 | 0.749 | 1     |
| Undulation degree| 0.353 | 0.512 | 0.641 | 0.786 | 1     |
| Slope gradient   | 0.164 | 0.297 | 0.420 | 0.582 | 0.800 |
| Distance from faultage | 0.200 | 0.400 | 0.600 | 0.800 | 1     |
| Rainfall         | 0.165 | 0.293 | 0.455 | 0.672 | 0.672 |

### 2.3 Research methods

Artificial Neural Networks (ANN) is a nonlinear dynamic system composed of a large number of simple processing units (neurons), which has learning, memory, calculation, recognition, prediction and many other intelligent processing functions. It is different from the modern electronic computers called "computers" and the artificial intelligence systems based on symbolic deduction, but another artificial intelligence system that mimics the human brain nervous system at different levels and degrees [15].

BP neural network is an error back-transfer learning algorithm. The principle is to transform an input vector into an output vector through an implicit layer to realize mapping from input space to output space. The forward mapping is implemented by the weight, and the output of the network under the current weight is compared with the expected output of the mapping requirement that is desired to be implemented [16]. The BP algorithm has not only input layer nodes, output layer nodes, but also one or more hidden layer nodes. For the input signal, it must be propagated forward to the hidden layer node. After the action function, the output signal of the hidden node is propagated to the output node, and finally the output result is given.

This paper chooses a 3-layer neural network model as a debris flow risk assessment model. The specific steps are shown as follow[17]:

1. Operation initialization: the three-layer structure weight \( W_{ij}, W_{jk} \) and a threshold value \( Q_j \) of the network model to be established are respectively assigned a small enough random value.

2. Each model pair (input value and expected output value) of the sample set (known sample) is input to the network.

3. Calculate the output values of each layer node through forward propagation between networks:

\[ X_j^{(i)} = f \left( \sum W_{ij}^{(i)} X_j^{(i-1)} + Q_j \right) \]  

(3)
where: $s$ is the input and output known sample number; $i, j$ is the serial number of each corresponding layer; $X_{j}^{(s)}$ is the connection weight of the $j_{th}$ neuron on the $s_{th}$ layer; $W_{ij}^{(s)}$ is the connection weight of the $i_{th}$ neuron to the $j_{th}$ neuron on $s_{th}$ layer; $f$ is the action function, which generally is the sigmoid function, as $f = 1/(1 + \exp(-x))$.

(4) Calculate the error $\delta_{k}^{(s)}$ of the output layer node $k$ and the error $\sigma_{j}$ of the intermediate layer node $j$ based on the actual output $o_{k}$ of the output layer and the expected output $t_{k}$ and the output $y_{j}$ of the intermediate layer node $j$:

$$\sigma_{k} = (o_{k} - t_{k})\alpha_{k}(1 - o_{k})$$  \hspace{1cm} (4)$$
$$\sigma_{j} = \sum_{s} \delta_{k}^{(s)}W_{j}^{(s)}(1 - y_{j})$$  \hspace{1cm} (5)$$

where: $s$ is the total number of learning modes.

(5) Adjust the connection weight of each layer according to the error of each layer

$$W_{j}^{(s)} = W_{j}^{(s)} + \alpha \delta_{k}^{(s)} y_{j}$$  \hspace{1cm} (6)$$
$$W_{ij}^{(s)} = W_{ij}^{(s)} + \alpha \sigma_{j} I_{i}$$  \hspace{1cm} (7)$$

where: $\alpha$ is the learning indicator and $I_{i}$ is the output value of the input layer node $i$.

(6) Repeat step (4) to step (7). If the output layer cannot obtain the desired output value, the error signal between the actual output value and the expected output value will be circulated along the original link path until the accuracy requirement is reached, that is, the actual output. The difference between the value and the expected value is within a certain acceptable range, and the network training ends. After the training, the network found the influence of each evaluation indicator on the risk of debris flow from the sample example, and fixed it in the form of network connection weight and other structural parameters, and remained stable. At this point, the trained neural network can be used to input the unknown sample parameters to be predicted, and the corresponding prediction results can be directly obtained.

**3. Results and discussion**

This paper takes 270 point position data of actual debris flow disasters in Ya'an as input data, 200 data as training samples and 70 data as test samples. The results are shown in Table 2. Then use the grid calculator of GIS spatial analysis to superimpose each index factor according to formula 8 to obtain the comprehensive distribution map of debris flow hazard in Ya'an city, as shown in Figure 1.

| Factor                | Lithology level | Undulation degree | Slope gradient | Distance from faultage | Rainfall |
|-----------------------|-----------------|-------------------|----------------|------------------------|----------|
| Weight                | 0.140           | 0.094             | 0.450          | 0.197                  | 0.118    |

$$P = L \times \alpha + U \times \beta + S \times \chi + D \times \delta + R \times \epsilon$$  \hspace{1cm} (8)$$

In the formula, $P$ is the flood hazard index; $L$ is the lithology index; $U$ is the undulation index; $S$ is the slope index; $D$ is the distance from the faultage; $R$ is the rainfall index; $\alpha, \beta, \chi, \delta, \epsilon$ are the weight of each indicator.
It can be seen from the figure that the risk of debris flow in the northern part of Ya'an city is significantly higher. Through the analysis of the test samples, 70% is located in the medium and above areas of debris flow hazard, and the results reflect the actual situation of the spatial distribution of debris flow in Ya'an. The extremely high-risk area accounts for 15.3% of the total area. The central and southern parts of Baoxing County, the north-central part of Lushan County, the western part of Tianquan County, and the central part of the central part of Shimian County are extremely dangerous. These areas belong to the area with large and strong slope fluctuation, which are distributed on the fault zone, and are prone to debris flow disaster; the high-risk areas account for 25.2%, including southern Baoxing County, Lushan County, Tianquan County, middle of Yingjing County, and north-central Shimian County, these areas have relatively large undulations and less vegetation cover.
If heavy rains or flash floods occur, debris flows are prone to occur; the moderately dangerous areas account for 29.7%, mainly distributed in the central and southern parts of Ya’an city. Most of these areas are not distributed on the fault zone and the geological conditions are good, the possibility of debris flow is small; the low-risk, extremely low-occupancy areas accounts for 29.8%, mainly located in the southeast of Yucheng District, north of Baoxing County, west of Shimian County and Mingshan District. Most of these areas have good vegetation coverage and flat terrain, and are not prone to debris flow disaster.

Statistical analysis was performed using GIS software, and the results as shown in Table 3 were obtained. It can be seen from the table that the V-level areas of Tianquan County, Lushan County and Baoxing County account for the largest proportion, and the proportion of Grade IV areas is above 25%. It is necessary to be wary that these IV-level areas may break through the critical value. It is converted into a V-class dangerous zone, so these three administrative districts need to do scientific disaster prevention work; Yingjing County, Yucheng District and Shimian County III and above account for more than 50%, and it is also necessary to be alert to the possibility of mudslides. The areas of Grade IV and Grade V in Hanyuan County and Mingshan District are relatively small, and the overall risk of debris flow is the lowest, but it is also necessary to pay attention to the prevention of debris flow.

4. Conclusion
(1) On the basis of summarizing the research results of predecessors, select rainfall, lithology level, distance from faultage, undulation degree and slope gradient as evaluation indicators, construct an evaluation index system, and each index factor is superimposed with its respective weights by GIS software. Then the risk distribution map of the debris flow disaster in Ya’an city is obtained. Finally, the analysis results are compared with the location of the debris flow disaster at the checkpoint, and the research results show good consistency.

(2) Using BP neural network and combining with the actual occurrence location point of debris flow disaster, the evaluation indicators are extracted and standardized for learning and training, then the weight of each evaluation index is obtained, and the accuracy of the research results is improved.

(3) According to the risk distribution map, the areas with high risk of debris flow disasters are mainly concentrated in the central part of Baoxing County and the western part of Tianquan County and most areas of Lushan County. The population of these three administrative areas is relatively densely distributed, which will aggravate the impact of mudslide disasters. Therefore, appropriate
disaster prevention work should be done.

(4) This paper uses the 0.25° × 0.25° TRMM satellite multi-year average rainfall index, which has a large spatial scale, but is still insufficient spatial differentiation of mountain rainfall. And there is also a lack of description of local short-term heavy rainfall. The surface rainfall on the hourly and empty scales is the average maximum three-day rainfall for many years, and is greater than the average annual rainstorm days of the 50 mm flood season. The above shortcomings are expected to be improved in future research.

Reference:
[1] Wang J Z, Ren G M, Yu T B, Gao B. (2015) Development characteristics and risk assessment of debris flow in Dayelonggou, Lushan earthquake area, Sichuan. J. Chinese Journal of Geological Hazard and Control, 26(04): 1-5.
[2] Zhao X, Cheng Z L, Liu J K, Liu D X, Shi L. (2013) Study on Risk Assessment of Single Gully Debris Flow in Dongchuan District of Yunnan Province. J. Journal of Catastrophology, 28(01): 102-106.
[3] Wang X L, Li J Y. (2011) Research on debris flow risk assessment system based on analytic hierarchy process. J. China Mining, 20(10): 113-117.
[4] Han Y S, Cui P, Liu H J, Yang M Q, Li C K. (2008) Risk assessment method and application of debris flow disaster. J. Chinese Journal of Safety Science, 18(12): 140-147+26.
[5] Liu X L, Su P C. (2004) Risk Assessment of Debris Flow in Sichuan Province. J. Disaster Science, (02): 25-30.
[6] Li H X. (2002) Analysis on Risk Assessment, Zoning and Development Trend of Soil and Water Loss in Ya'an Area. J. Journal of Soil and Water Conservation, (06): 17-19.
[7] Tang J F, Xie H. (1999) Study on debris flow risk zoning under the support of gis technology—Taking the upper reaches of Minjiang River as an example. J. Sichuan Surveying and Mapping, (03): 120-122.
[8] Tan B Y. (1986) Quantitative comprehensive evaluation of the severity of debris flow ditch. J. Bulletin of Soil and Water Conservation, (01): 51-57+44.
[9] Yu J R, Yan L H, Li X. (2014) Evaluation of debris flow risk based on gis space modeling. J. Science Chinese, (4):30-30.
[10] Wang L X. (1982) About the classification of wild rivers. J. Journal of Beijing Forestry College, (03): 94-107.
[11] Fu X G. (2008) Research on Benefit Evaluation System of Slope Runoff Distribution Project in Earth and Rock Mountain Area. D. Chang'an University.
[12] Cui P, Yang K, Wei F Q, Chen X Q. (2001) Index system for debris flow disaster assessment. J. Journal of Natural Disasters, (04): 36-41.
[13] Li X Y, Liu Q S, Bai S Y. (2017) Risk Assessment of Debris Flow Disaster in Ya'an City, Sichuan Province. J. Bulletin of Soil and Water Conservation, 37(02): 278-283+288.
[14] Lai C G, Wang Z L, Song H J. (2011) Risk assessment of flood disaster in Beijiang River Basin based on bp neural network. J. Hydropower Energy Science, 29(03): 57-59+161.
[15] Mao J, Zhao H D, Yao W. (2011) Development and application of artificial neural network. J. Electronic Design Engineering, 19(24): 62-65.
[16] Wang L. (2014) Principle, classification and application of artificial neural network. J. Science and Technology Information, (03): 240-241.
[17] Liu Y J, Hu H T, Bai Z Y. (2001) Neural network method for risk assessment of debris flow. J. Geology and Exploration, (02): 84-87.