Hazard Assessment of Debris Flow on Yakang Highway Based on GA-BP Neural Network

Yang Yang and Huaqiang Luo

Southwest Petroleum University School of civil engineering and architecture, Sichuan Chengdu 610500.China

Abstract

The hazard assessment of debris flow along highway is one of the most important subjects on predictions of debris flow. It is of great significance for highway construction and maintenance later. BP neural network has good non-linear information process ability, but the network has such defects as slow convergence rate, poor generalization ability and local minimum trapping in practical application. Network optimization with genetic algorithm as the study object, this paper divides hazard assessment units based on watershed and models GA-BP neural network for hazard assessment on the 19 debris flow gullies.

Keywords

Debris Flow, Hazard Assessment, Genetic Algorithm, BP Neural Network, Yakang Highway.

1. Introduction

Yakang highway is a main part of the national highway network and also the only highway from mainland to Ganzi Tibetan prefecture. It not only provides convenient way of going but also undertakes a great role in national defense transportation. However, frequent geological disasters occur in the area due to complicated mountainous terrain, changeable climate and intensified human activities. Thereinto, debris flow is the most destructive. Since highway construction is a linear project, the destruction of a certain part will cause highway disruption and even traffic paralysis, and will also lead to heavy losses of life and property. Therefore, the hazard assessment of debris flow on Yakang highway is practically significant.

The hazard assessment of debris flow has always been an emphasis in geologic hazard research at home and abroad. The constantly updated and perfected assessment method has improved the speed and accuracy of hazard assessment. In 1977, the Japanese scholar Katiji Adati calculated the probability of debris-flow risk degree for the first time considering the three factors of topography, debris flow form, and atmospheric precipitation [1].In 1981, Hollingsworth and Kovacs built the basic framework on hazard assessment of debris flow along highway with the method of scoring, and completed the hazard assessment of landslide with the method of factor overlapping [2]. Arora et al. analyzed the landslide hazard in the Bhagirathi River drainage area in the Himalayas [3]. In 1993, Tang Chuan et al. combined two-dimensional unsteady debris flow theory with mathematical theory to construct hazard assessment model and put it into practice [4]. In 2016, Sun Zhengchao et al. applied the research of threshold effect of drainage network extraction based on DEM in different levels to the division of debris-flow watershed. The research has improved the accuracy of debris-flow hazard assessment [5][6].

BP neural network is one of the most widely used models of artificial neural network. Its function approximation ability is very suitable for solving the problems of multi-target complexity and uncertainty in risk assessment [7].In this paper, the optimization of initial weight and threshold value of BP neural network with genetic algorithm have further improved the accuracy, fitting precision and efficiency of the network.

2. The profile of study area

Yakang highway passes through the transition sections of western edge of Sichuan Basin and eastern edge of Qinghai-Tibet Plateau, ascending rapidly from east to west with complicated and varied topography, big swells, ravines and gullies. Thereinto, Luding-Kangding segment with the length of about 46 km tilts west towards the east, so the west side is higher than the east with the biggest altitude difference of 1400 m and more than half of slops over 35°. The highway crosses the entire Longmen Mountain tectonic belt where there are Dadu River breakage and Hongfeng breakage with intense neotectonic movements. Formation lithology along the highway is complicated. Most of them are exposed metamorphic intrusive rocks. The Luding-Kangding segment is located at the transition zone from the basin to the plateau. The great altitude difference leads to various climates such as subtropical monsoon climate and continental monsoon plateau climate. The mountains in the area of highway are north-south oriented so that they cut off vapor transportation from the east and west. Foehn effect is caused by the sink of dry air making the river valley warmer, thus form the tendency of temperature increment and water diminishment from east to west and from south to north. In the religions the highway passes through, the precipitation has no big change and there are clear dry and wet seasons. The annual average rainfall is 880 to 930 mm. Xianshuihe seismic belt is the main seismic belt in Lukang segment of Yakang highway and the most active one ever in Sichuan. There have been thousands of earthquakes in recent 400 years including 9 quakes over magnitude 7.

3. Hazard assessment of debris flow based on GA-BP neural network

3.1 BP neural network

BP neural network is a multilayer feedforward network with one-way transmission. It consists of the input layer, the hidden layer and the output layer. The hidden layer can have one or more layers and there is no coupling in the same-layer nodes. The input signal is successively passed from input nodes to hidden-layer nodes and then to output nodes. The output of nodes in a certain layer only affects that of the next layer, as shown in figure 1.

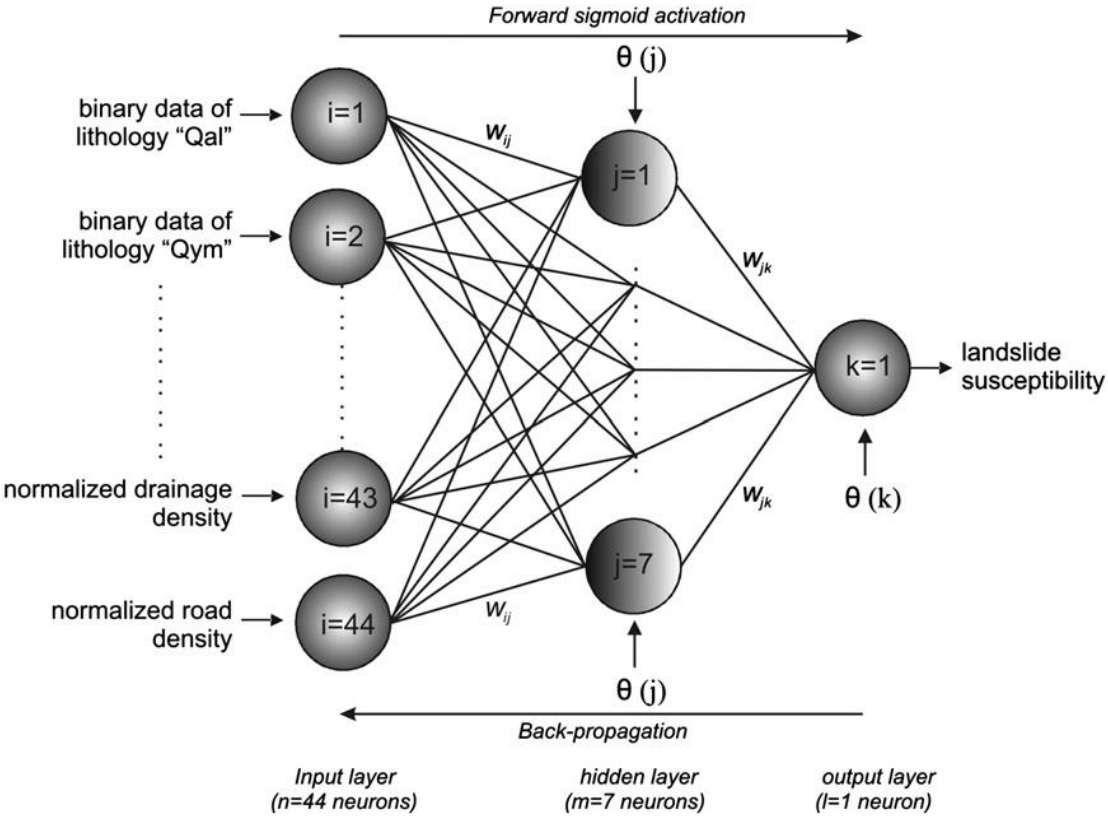

Figure 1. Topological structure chart of BP neural network

When the structure of BP neural network is determined, network training should begin after the setup of target error and raining times. That is to say, the network weights and threshold values are studied and revised in order to make the error function decrease toward negative gradient direction until it reaches the final target value. In the following formula, *h*, *m* and *n* represent the quantity of input layer, the hidden layer and the output layer respectively. w_{ij} represents the weight between input-layer node *i* and hidden-layer node j. θ_j refers to the threshold value of hidden-layer node j, w_{jk} to the weight between hidden-layer node j and output-layer node K. θ_k is the threshold value of output-layer node k.O_k refers to the output of the output layer. T_k stands for the expected output with the weight and threshold value $\in (-1,\ 1)$. BP neural network training can be divided into the following two steps:

Training is started by input known samples of learning to the network. The outputs of each neuron are calculated step by step from the first layer based on the pre-set of network structure, the initial weight and threshold value. *f* is the activation function of sigmoid function. Here are the formulas of network training.

$$S_j = \sum_{i=0}^{m-1} w_{ij} O_i + \theta_j \tag{1}$$

$$O_j = f(S_j) \tag{2}$$

(2)The influence of weights and threshold values on overall error is calculated from the last layer step by step. Then changes are made to the weights and threshold value. BP neural network uses δ learning rule (also called the rule of error correction) to modify weights and threshold value according to gradient descent method.

Amendment of weights and threshold values from the hidden layer to the output layer:

$$w_{jk} = w_{jk} - \eta_1 \frac{\partial E(w,\ \theta)}{\partial w_{jk}} = w_{jk} - \eta_1 \delta_{jk} O_j \tag{3}$$

$$\theta_k = \theta_k - \eta_2 \frac{\partial E(w,\ \theta)}{\partial \theta_k} = \theta_k - \eta_2 \delta_{jk} \tag{4}$$

Amendment of weights and threshold values from the input layer to the hidden layer:

$$w_{ij} = w_{ij} - \eta_1 \frac{\partial E(w,\ \theta)}{\partial w_{ij}} = w_{ij} - \eta_1 \delta_{ij} O_i \tag{5}$$

$$\theta_j = \theta_j - \eta_2 \frac{\partial E(w,\ \theta)}{\partial \theta_j} = \theta_j - \eta_2 \delta_{ij} \tag{6}$$

Thereinto, ηrepresents study step, and $E(w,\ \theta)$is the error function in rule learning.

3.2 GA-BP neural network

GA-BP neural network is the integration result of genetic algorithm and BP neural network. It is characterized by the global searching ability with excellent genetic algorithm to search the optimal solution range of weights and threshold values of the BP neural network. Then optimal solution is acquired after fine gradient search on local area with BP algorithm. This method not only increases the work efficiency of network but also improves its accuracy.

3.2.1 Realization of GA-BP neural network

(1) Coding scheme selection. Network weights and threshold values are the coding objects. There are two schemes to choose: binary coding and real coding.

(2) Parameter settings. The parameters of genetic algorithm are set to determine the initial population scale and evolving algebra.

(3) Fitness function. The leading indicators to evaluate the performance of BP neural network are square error between network actual output value and expected value, as well as E_p. The smaller the E_P value is, the greater the network performance. Here is the error formula:

$$E_p = \frac{1}{2}\sum_{n=1}^{h}\left(T_n^p - O_n^p\right)^2 \tag{7}$$

Finally, fitness function is set as $f = C - E_p$, in which C is the biggest constant.

(4) Genetic manipulation. Genetic manipulation including selection, crossover and mutation is the key process of population evolution. A better new generation of population is generated by the mentioned three operations.

(5) After the calculation of genetic algorithm, the fittest individual decoding is selected, and the optimal weight and threshold value are acquired.

(6) Prediction is implemented by putting the prediction samples into good BP neural network.

4. Establishment of GA-BP neural network model

4.1 Determination of BP neural network structure

According to site investigation and literature review, there are 9 indexes selected to assess the debris flow hazard in Lu-kang segment of Yakang highway. They are relief, average gradient, fault density, watershed cutting density, seismic intensity, mean annual precipitation, annual precipitation variation coefficient, formation lithology and NDVI value. Thus, the input layer is set to 9 and the output layer is set to 1. The number of hidden-layer nodes can be calculated according to the empirical formula: $m = (h + n)/2$ or $m = \sqrt{h \times n}$. It is the result of calculation and comparison that m=5 should be the most appropriate.

4.2 Genetic algorithm coding and parameter setting

Both genetic algorithm and BP neural network calculation are completed based on Matlab platform. At first, a set of randomly generated samples are coded to the binary system. The iteration number is set to 10; initial population to 2; crossover probability P_c=0.2 and mutation probability P_m=0.05. Finally, the average fitness value of chromosome tends to be stable when it evolutes to the 10th generation.After the fittest chromosome is acquired, it is decoded to get the optimal initial weight and threshold value appropriate for the BP network.

Table 1. Network weights after optimization

Point	1	2	3	4	5	6	7	8	9
W_{ij}	1.514	2.369	2.051	-2.215	-1.865	-2.078	-2.827	-2.945	0.579
W_{jk}	-2.954	0.272	0.055	-1.519	-2.728				

Table 2. Network threshold values after optimization

Point	1	2	3	4	5
θ_j	-2.382	2.981	-0.846	0.751	-0.640
θ_k	2.050				

4.3 Training of GA-BP neural network

Training models can be established after acquiring the optimized network weights and threshold values. In this paper, 88 debris flow gullies at Kangding county and Luding county are selected as samples including 66 training samples and 22 ones for verification. The data of 66 sets of samples used for training are input to GA-BP neural network after normalization processing. It is resulted that the network performance reaches the optimum when it trains 10000 times, as shown in figure 2. After network completes the training, samples need to be verified to test the network performance and to further judge whether the network available. The data of 22 sets of samples used for verification are input to the network after normalization processing. It is resulted that the error between the actual output and the expected output is within reasonable range. And the actual output with the expected output curve in the function output graph fit very well.Therefore, it is declared that the network could be used to predict debris flow gullies along the highway.

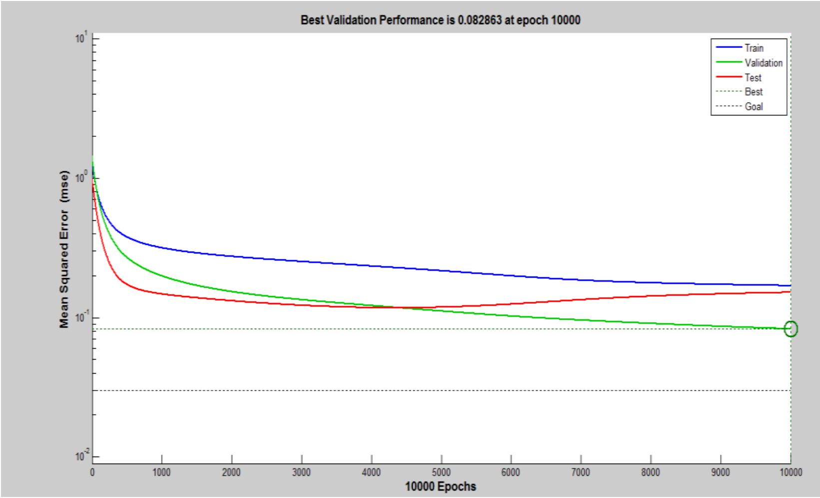

Figure 2. The training result of GA-BP neural network

5. Hazard assessment of debris flow along highway

5.1 Division of assessment unit

The research data in this paper based on DEM data come from ASTER GDEM V2 jointly distributed by NASA and METI with high precision and wide coverage. Hazard assessment units of debris flow can be divided as soon as DEM data are acquired. The main processes are: 1. establish buffer zone 10 km around the highway; 2. calculate the depression in buffer zone and repeatedly fill the depression until it reaches the target value; 3. calculate the flow direction; 4. summarize the calculation and extract river network; 5. extract sub-catchment; 6. determine assessment unit of debris flow hazard in buffer zone. Assessment unit of debris flow hazard along highway is obtained through the above steps, as shown in the figure. According to the figure of watershed unit division and literature research results, there are 19 units of debris flow that can make impact on the highway. Therefore, the 19 debris flow gullies will be evaluated and predicted in this paper.

5.2 Extraction of assessment index data

All the data of assessment on the 19 debris flow gullies along the highway are based on the DEM data extracted from Arcgis10.2 platform. After acquired, these raw data need to be further processed with such formula as shown in formula (8):

$$X'_{IJ} = \frac{X_{IJ} - X_{Imin}}{X_{Imax} - X_{Imin}} \quad (8)$$

Thereinto, X_{IJ} represents the value of index (I) of debris flow (J). X_{Imax} and X_{Imin} respectively refers to the maximum and minimum values of index (I) in all of the debris flow gullies.

5.3 Hazard assessment of debris flow

After the steps above are completed, the assessment of the 19 debris flow gullies is implemented by putting the already processed data into GA-BP neural network model. Then their hazard degrees are obtained respectively combined with the partition table of debris flow hazard degree. The results are shown in table 3.

Table 3. Assessment indexes and results

Number of debris flow gullies	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	Assessment result	Hazard degree
1	-0.482	-0.147	-0.070	-0.975	0.453	-1	0.402	-0.993	0.870	0.560	moderate
2	-0.634	0.388	0.487	-0.987	0.032	-1	0.407	-0.993	0.723	0.061	low
3	-0.521	0.804	-1	-0.976	-0.198	-1	0.084	-0.993	0.260	0.123	low
4	-0.461	0.572	-1	-0.983	0.260	-1	-0.038	-0.993	0.967	0.159	low
5	-0.498	0.358	-1	-0.972	-0.416	-1	-0.723	-0.993	0.580	0.169	low
6	-0.408	0.691	-1	-0.981	-0.172	-0.466	-0.636	-0.993	0.262	0.180	low
7	0.099	0.614	-1	-0.922	0.025	-0.474	-0.612	-0.993	0.220	0.060	low
8	0.521	0.482	-1	-0.710	0.475	0.748	-0.478	-0.993	-0.060	0.732	high
9	-0.003	0.794	-1	-0.923	0.193	-1	-0.585	-0.993	0.726	0.240	low
10	0.425	0.338	-1	-0.845	0.064	-1	0.049	-0.993	0.777	0.721	high
11	-0.158	0.042	-1	-0.918	-0.417	-1	0.161	-1.000	0.503	0.653	high
12	0.732	0.299	-1	-0.770	-0.018	-1	0.156	-0.993	0.796	0.763	high
13	-0.264	-0.128	-0.127	-0.841	0.006	-1	0.246	-0.958	0.060	0.553	moderate
14	0.189	-0.054	-0.505	-0.637	-0.063	-0.478	-0.748	0.075	-0.033	0.073	low
15	0.831	0.456	-1	-0.665	0.527	-0.930	-0.196	-0.993	-0.171	0.054	low
16	1.049	0.459	-1	-0.481	0.264	-0.703	-0.019	-0.993	0.550	0.673	high
17	-0.575	-0.556	-1	-0.970	0.147	-1	0.224	-0.993	0.220	0.030	low
18	0.548	0.107	-1	-0.705	-0.170	-1	0.199	-0.993	0.764	0.983	extreme-high
19	1.335	0.541	-1	0.252	0.442	-0.771	0.052	-0.993	-0.096	0.663	high

Note: all the assessment indexes in the table are obtained after normalization processing.

6. Conclusion

1. The hazard assessment of debris flow on Luding-Kangding segment of Yakang highway is completed based on GA-BP neural network. The result shows that the debris flow hazard is serious in this segment, including 1 extreme-high hazardous debris flow gully, 6 high hazardous gullies, 2 in moderate hazard degree and 10 in low hazard degree. The debris flow gullies in extreme-high and high hazard degree account for 36% of the total number. The assessment result agrees with the truth, so it can be used as the basis for the subsequent risk assessment.

2. The GA-BP network not only makes full use of function approximation ability of BP algorithm, but also combines it with global search ability of genetic algorithm. Weights and threshold values that can optimize the network are acquired through contrastive analysis of genetic algorithm. They have greatly improved the network efficiency and accuracy, and have ultimately enhanced the reliability of the hazard assessment results of debris flow along Yakang highway.

References

[1] Katsuharu Adachi, Hisao Tokuyama, Chief Junior middle man, et al. Judgment on the risk of debris flow occurrence [J]., 1977, 30:7-16.

[2] Hollingsworth R, Kovacs G S. Soil Slumps and Debris Flows: Prediction and Protection[J]. 1981, 18(1):17-28.

[3] M. K. Arora Corresponding author, A. S. Das Gupta†, R. P. Gupta. An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas[J]. International Journal of Remote Sensing, 2004, 25(3):559-572.

[4] Tang Chuan, Zhou Juqian, Liu Xilin. Numerical simulation of two dimensional unsteady flow of debris flow and its application in risk assessment [J]. Yunnan science and Technology Press, 1993

[5] Sun Zhenchao,Wang Zegen,Su Pengcheng,Hu Depeng,Liao Qiangcheng.Extraction and Precision Analysis of Small Watershed River Network Based on Different Data Sources of DEM: A Case Study of Jiahe Watershed[J]. Electronic Journal of Geotechnical Engineering, 2015, 20(25):12623-12634.

[6] Wang Zegen, Sun Zhengchao, Su Pengcheng, Liao Qiangcheng, Hu Depeng.Research on Threshold Effect of Extracted River Networks Based on DEM with Different Horizontal [J].Electronic Journal of Geotechnical Engineering,2015,20(25):12635-12645.

[7] Zhao Yuan, Liu Xilin. Application of artificial neural network in debris flow risk assessment [J]., 2005, 16(2):135-138.

[8] Zhao Jianbo, Zhao Haiwei, Yang Yuan. Application of artificial neural network in prediction of flood level in tidal reach [J]., 2007.

[9] Wang Min. Study on the extraction of hydrological elements and the establishment of hydrological topological structure based on [D]. DEM Northwestern University, 2012.

[10] Turcotte R, Fortin J P, Rousseau A N, et al. Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network[J]. Journal of Hydrology, 2001, 240(3–4):225-242.

[11] Graymore M L M. Sustaining human carrying capacity: a tool for regional sustainability assessment.[J]. Ecological Economics, 2010, 69(3):459-468.

[12] Liu Xilin. Risk assessment of debris flow [M]. Science Press, 1995