Seismic Landslide Hazard Identification and Assessment Based on BP Neural Network

Jinsheng Fan[1,2], Weidong Li[1,*] and Xinjian Shan[2]
[1]College Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China
[2]State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing 100029, China
[*]Corresponding author

***Abstract*—Based on geographic information systems and remote sensing technology, this article used BP neural network method and choose slope, aspect, intensity, faults, water, elevation, DEM, hardness 8 earthquake landslide factors as influencing factors in the study area (E103° ~ E105°, N30.8° ~ N32°) to identify the earthquake and landslide-prone evaluation studies. The results show: BP neural network landslide recognition correct rate reached 85.3%, and 70% of the *landslide* occurred in the predicted high-risk areas, and the evaluation of seismic landslide convex curve showing a steep trend, the using of BP neural network is feasible to evaluate the seismic landslide.**

Keywords-BP neural network; seismic landslide; geographic information systems; identification and assessment

I. INTRODUCTION

With an increased awareness of the landslide mechanism and with the rapid development of the modern theory of mathematical mechanics and computer technology, various landslide prediction methods continue to emerge[1]. For example, slices method and numerical analysis method based on the traditional limit equilibrium theory, comprehensive analysis based on AHP[4-7], the weight of evidence[8] and multiple logistic regression based on the method of statistical theory[10], modern artificial neural networks based on intelligent data mining algorithms[2,9,16], fuzzy comprehensive evaluation method based on fuzzy theory[14]. In the study area (E103° ~ E105°, N30.8° ~ N32°), We established BP[11-13] neural network model and choose slope, aspect, intensity, faults, water, elevation, DEM, hardness 8 earthquake landslide factors as influencing factors for the earthquake and landslide-prone evaluation studies.

II. DATA PREPARATION AND PROCESSING

This paper used the following data: 1) the heavy disaster area of Wenchuan earthquake triggered slip slope layer data, 2) DEM layer of the study area, 3) Wenchuan earthquake intensity distribution layer, 4) Water layer of the study area, 5) Active fault Layer of the study area, 6) Wenchuan geological distribution Layer, 7) National 1:400 000 basic information layer.

We made the conjoint analysis with the landslide and the influence factor layers under Arcgis platform. We got the landslide attribute tables which contained not only the attributes of area, perimeter, but also 8 attributes of corresponding slope, slope direction, intensity, stream buffer distance, fault buffer distance, rock and soil hardness, DEM, and height, all of the above were prepared for choosing landslide samples. We took the gravity of each slide surface as a point, and converted the landslide surface into landslide point layers[3,15]. Based on 8946 landslide surfaces which are the representative of the heavy disaster area, we selected sample points and generated a number of random no-slipping samples in the region that no landslide occurred. We combine the samples of the landslide and the non slip samples to form the layer of final training samples, and exported their attribute tables to EXCEL, which are used as the training dataset. the landslide surface layer of study area(figure I), Landslide sample group layer(figure II), Non slip sample group layer(figure III), joint layer(figure IV).

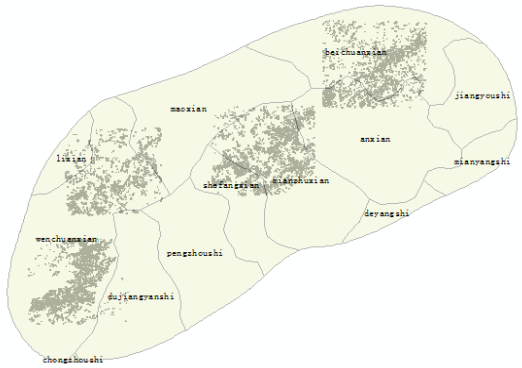

FIGURE I. THE LANDSLIDE SURFACE LAYER OF STUDY AREA

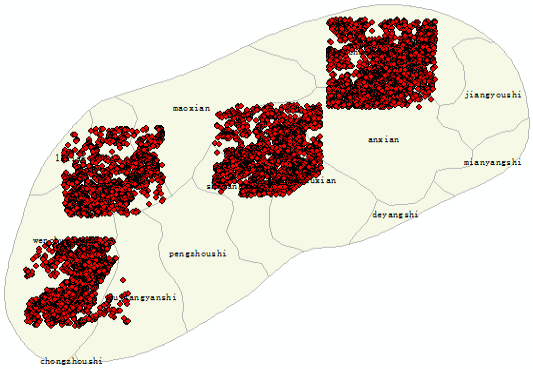

FIGURE II. LANDSLIDE SAMPLE GROUP LAYER

© 2015. The authors - Published by Atlantis Press

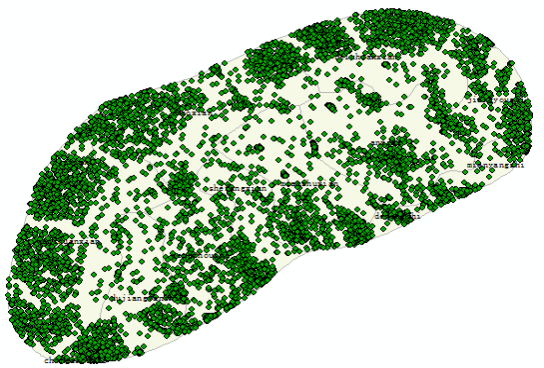

FIGURE III. NON SLIP SAMPLE GROUP LAYER

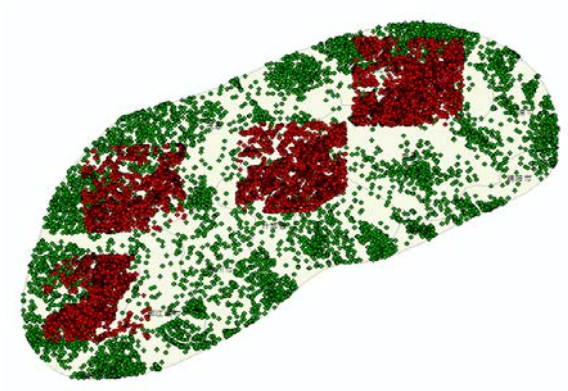

FIGURE IV. JOINT LAYER

In order to evaluate the landslide susceptibility, we used raster data format to carry out weight overlay analysis (because vector data is limited by the scale, it is very difficult to be very precise), and transformed the vector layer into grid layer, the size was 7095 (columns) *4254 (rows).

III. MATLAB IMPLEMENTATION OF THE MODEL AND ITS APPLICATION OF LANDSLIDE IDENTIFICATION

There are totally 17226 samples(including 8946 landslide samples and 8340 no-slipping landslide samples),which is prepared in the previous text. But taking into account the convergence model itself, this paper finally randomly selected samples and non-slip landslide 1000 samples respectively for training, with the left 15,226 samples for landslide recognition and test. Of course, the paper also were carried out 4000, 6000, 8000, 10000 samples of training, but the final results show that the 2000 samples in the model established in this paper is the fastest convergence, and the correct rate of identifying is relatively high, reaching 85.32%. A different number of samples to identify the correct rate comparison table as shown in Table I.

TABLE I. CORRECT RATE COMPARISON TABLE

number of samples	2000	4000	6000	8000	10000
Correct rate	85.32 %	80.82%	80.72%	82.04%	81.45 %

IV. EVALUATION OF LANDSLIDE-PRONE

Landslide-prone evaluation is on the basis of trained BP neural network. We turn the eight factor layers into a raster grid, which can be expressed in terms of classification grid by property[5-8]. For BP neural network which has been trained, because the entire training process is programmed, so we overlay the raster expression of each factor in accordance with the training process to get the desired forecast results. In accordance with the overall neural network architecture diagram, we overlay eight factors three times according to different weights and plus three results to a final result. The weights between the input layer and the hidden layer are shown in table II, the weight of the hidden layer and the output layer are shown in table III. Final superimposed result is shown in Figure V, Its range is -58.90 ~ 89.41, and I did not put this range to map it to a range of 0 to 1 with logsig function so that a wider range of values can be some more conducive to natural classification. According to the natural classification[7,8], the area ranging from -58.90 to -31.57 was divided into extremely low-prone areas, the area ranging from -31.57 to 4.67 was divided into low-prone areas, the area ranging from -4.67 to 17.08 was divided into the middle-prone area, the area range from -17.08 to 43.11 was divided into high-prone areas, the zoning ranging from 43.11 to 89.406 was divided into extremely prone areas. It can be seen that lots of landslides occurred in the landslide predicted high-risk areas. To test the reliability of the results predicted, we adopt a method called the percentage cumulative landslide area - prone area percentage of cumulative curve[6-11]. This method is a quantitative evaluation method of landslide-prone model, we divided the susceptibility results of the evaluation in accordance with one percent of the area into 100 equal parts descending intervals and respectively calculated the percentage of landslide occurring, we finally draw cumulative susceptibility percentage and landslide cumulative percentage curves, namely the evaluation curve shown in Figure VI, and represent the correct rate the model predicted by the area under the curve (AreaUnderCurve).The results show the correct rate reached 83.94%, indicating that the prediction is feasible.

TABLE II. THE WEIGHT OF EACH NODE FROM THE INPUT LAYER TO THE HIDDEN LAYER

slope	aspect	intensity	faults	water	elevation	DEM	hardness	offset
-1.5361	0.5375	-15.2437	12.8972	73.7433	3.5880	-13.4643	-49.7637	-7.8857
3.6591	1.3898	40.6752	-35.2324	-34.3924	30.1944	-7.4514	-20.6409	-3.6987
-20.4957	-0.5745	-61.4237	-17.2035	-1.7283	13.7603	-11.6756	21.9136	-31.5312

TABLE III. THE WEIGHT OF EACH NODE FROM THE HIDDEN LAYER TO THE INPUT LAYER

The first node of hidden layer	The second node of hidden layer	The third node of hidden layer	offset
-43.8870	61.0510	-45.8267	12.4661

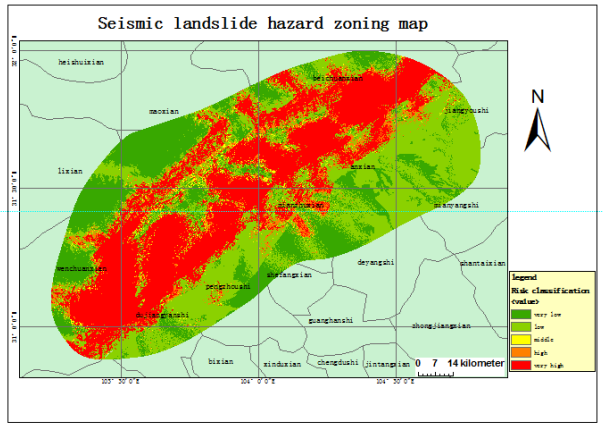

FIGURE V. SUPERPOSITION RESULTS TO PREDICT CLASSIFICATION MAP

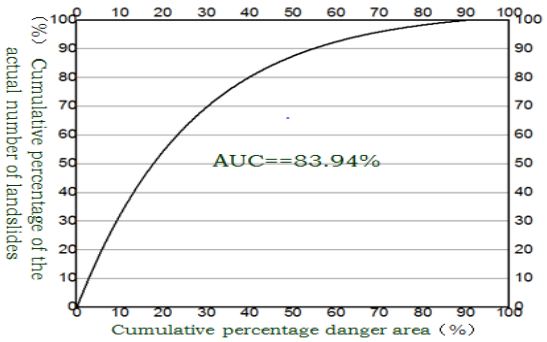

FIGURE VI. EVALUATION OF LANDSLIDE PREDICTION CURVE

V. CONCLUSION

This paper relies on ARCGIS platform to process data to obtain a sample of large quantities, and compared to the use of manually selecting the sample, the efficiency is extremely high[3,15]. In MATLAB platform we carried out the BP neural network model by programming for the training sample, which be used on landslide hazard identification and assessment. The correct rate of identification reached 85.30%, and Risk Assessment presents a steep convex curve trend, and 30% of in the region predicted to nearly 70 percent of the landslide, which indicated that this method is very reliable[11]. But paper also needs to be improved as follows：1) How to select the sample to make it more representative. 2) BP neural network convergence is slow. As to over-fitting problems, we need avoiding over-fitting algorithm to train samples, so that more samples can be used for training the network, which can make result-prediction more accurate.

ACKNOWLEDGMENT

This research was conducted during the collaborative work in State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration.

REFERENCES

[1] Zhang duo, wu zhonghai, li jiachun, jiaangyao. “Review of seismic landslide researche,” Journal of Geomechanics, 2013,19(3), pp.225-241.

[2] Xia youyuan, xiong haifeng. “Artificial Neural Networks factor sensitivity analysis of slope stability,” Chinese Journal of Rock Mechanics and Engineering, 2004,23(16), pp.2703-2707.

[3] Tang guoan, yangxin. “ArcGIS Geographic Information Systems Spatial Analysis Experimental Course,”Science Press, 2014.

[4] Yu lu, shan xinjian, chen xiaoli. “the rapid division of Lushan earthquake landslide hazard zone rating based on Comprehensive Index,” Seismology and Geology,2014,36(4), pp.1106-1115.

[5] liu lina, xu chong. Landslide Hazard Assessment of 2013 Lushan earthquake zone Supported by GIS based on AHP Method,” Journal of Catastrophology,2014,29(4), pp.183-191.

[6] xu chong. “Certainty analysis factor of earthquake landslide disaster basedon GIS,” Chinese Journal of Rock Mechanics and Engineering, 2010,29, pp.2972-2981.

[7] xu chong. “Earthquake landslide-prone evaluation of wencuan based on GIS and certainty factor analysis,” Journal of Engineering Geology,2010,18(1), pp.15-26.

[8] xu chong. “Yushu earthquake landslide hazard assessment based on weight of evidence,” Seismology and Geology,2013,35(1), pp.151-164.

[9] xu chong. “Earthquake landslide susceptibility zoning based on GIS and ANN model,” Geological Science and Technology Information, 2012,31(3), pp.116-121.

[10] xu chong. “Earthquake landslide hazard assessment and testing of wencuan based on Logistic regression model,” hydrogeology & engineering geology,2013,40(3), pp.98-104.

[11] xu chong. “the research of 2010 Yushu Earthquake Landslide Prediction ModelBased on different kernel functions,” Chinese Journal of Geophysics, 2012,55(9), pp.2994-3005.

[12] chen xiaoli, ran hongliu, wang mingming. “Potential Seismic landslide hazard zone division method,” Chinese Journal of Geophysics, 2012, 55(4), pp.1269-1277.

[13] chen xiaoli, zhao jian, ye hong. “Application of radial basis probabilistic neural network on landslide earthquake,” Seismology and Geology, 2006, 28(3), pp.0430-0440.

[14] chen xiaoli, qi shengwe, ye hong. “FuzzyComprehensive Evaluation ofseismiclandsliderisk based on GIS,” the journal of Peking University, 2008,44(3), pp.434-438.

[15] shan xinjian. “Remote sensing technology and geographic information systems integration in Geological Environmental Assessment,” 1999, beijing, Institute of Geology, China Earthquake Administration.

[16] chen xiaoli. “Application of Artificial Intelligence in seismic hazard assessment of landslide,” 2007, beijing, Institute of Geology, China Earthquake Administration .