Research on the Application of Artificial Neural Network to the Flood Risk Assessment

Qiong Li [a] Donghan Liu [b]

[a] *School of Mathematics and Physics, Hubei Polytechnic University, Huangshi, Hubei, China*

[b] *Mechanical and electrical school, Hubei Polytechnic University, Huangshi, Hubei, China*

Abstract: Flood is a most serious hazard to life and property. The traditional probability statistical method is acceptable in analyzing the flood risk but requires a large sample size of hydrological data. This paper puts forward a method based on artificial neural network (ANN) for flood analysis. An artificial neural network model-BP neural network is used to map multi-dimensional space of disaster situation to one-dimensional disaster situation and to raise the grade resolution of flood disaster loss. This technique contributes to a reasonable prediction of natural disasters risk. As an example, its application is verified in the flood risk analysis in China, and the risks of different flood grades are obtained. Our model yield very good results and suggests that the methodology is effective and practical so that it has the potentiality to be used to forecast the flood risk in flood risk management.

Keywords: artificial neural network · flood · risk analysis · assessment

1 Introduction

Natural disasters are increasing alarmingly worldwide. Flooding is a common natural disaster which very often causes property and human losses. Recent flooding disasters have shown the vulnerability of the so called developed and developing countries to such events. In China, flood disasters occur frequently, and about two-thirds of its area are facing the threat of different types and degrees of floods which is the result of natural and unnatural reasons such as social, economic factors. Given this, natural disasters present a great challenge to society today. And flood risk assessment of an area is important for flood disaster managers so they could implement a compensation and disaster-reduction plan. As severe floods occurring frequently, flood risk assessment and management play an important role in guiding the government take timely and correct decision for disaster rescue and relief.

Risk management for the operation of an existing flood protection system is the sum of actions for a rational approach to flood disaster mitigation. Its purpose is the control of flood disasters, in the sense of being prepared for a flood, and to minimize its impact. It includes the process of risk analysis, which forms the basis for decisions on maintaining and improving the system.

Risk analysis, one of the main subjects of flood management is a challenging task at the present. However, assessing flood risk is difficult because of the lack of objective measures of acceptable risk, scarcity of data, and abundance of unknown probability distributions. The flood risk analysis methods have shown a progress from direct integral method, Monte Carlo method, and mean first-order-second-moment method, to advanced first-order-second-moment method, second-order-method and JC method. The theories and methods of flood risk analysis were established according to the studied by the authors (Ang and Tang, 1984, Ashkar and Rousselle, 1981, Diaz-Granados et al. , 1984, Kuczera, 1982, Stedinger and Taylor, 1982, Todorovic and Rousselle, 1971, Todorovic and Zelenhasic, 1970, Wood and Rodríguez-Iturbe, 1975). Recently, many risk analysis approaches have been based on using linguistic assessments instead of numerical values. Using fuzzy sets theory (Zadeh 1965), data may be defined on vague, linguistic terms such as low probability, serious impact, or high risk.

In traditional flood risk assessment, probability statistics method is usually used to estimate hydrological variables' exceedance probability because of its mature basic theory and easy application. But in the case of practical issues, problems exist in the feasibility and reliability. Especially in small sample issues, results based on the classical statistical methods are usually unreliable. In fact, it is rather difficult to collect long sequence flood data and the sample is usually small.

Till now, scholars have made a deep research on the flood's random characteristics in risk analysis, but the study on some aspects such as its fuzziness (Chen 1998), gray characteristic (Xia 2000), unascertained characteristic (Liu et al. 1994), fractal dimension characteristic and chaos characteristic of the flood is relatively weak, so the researches of the risk analysis on such aspects need to be developed further. And neural network is data driven, and it can be described as mapping an input space to an output space. Many problems exist for which there is no underlying knowledge of the process that converts the measured inputs into the observed outputs. Artificial neural networks are well suited to this class of problem because they are excellent data mappers in that they map inputs to outputs. It's suggested that some neural net algorithm might provide a solution. Therefore, an artificial neural network model-BP neural network is used in this paper for evaluating the degree of flood disaster, where the disaster loss degree is a more reasonable continuous real number.

2 Basis of Artificial Neural Network

The essential of the risk analysis is to estimate the probability density of an index. Because of the incompleteness of the data, the application of traditional statistical methods can not guarantee a high precision. So we use the neural network with the observed sample and get their degree values, and then get the risk estimations by risk analysis. This paper uses artificial neural network and gets continuous degree index values of the samples, then it turns the degree values of observed sample into the continuous real degree number and then gets the risk values. It is tested by a case showing that the method is superior to traditional statistic model , so as to improve the result of traditional estimation.

Artificial Neural Network (ANN) are massively parallel interconnected networks of simple (usually adaptive) nodes which are intended to interact with objects of the real world in the same way as biological nervous systems do(Simon Haykin, 2009). It was proposed based on modern biology research concerning human brain tissue, and can be used to simulate neural activity in the human brain (Markopoulos1, Manolakos, & Vaxevanidis, 2008). ANN has the topological structures of information processing, distributing parallel. The mappings of input and output estimation responses are obtained via combinations of nonlinear functions.

In terms of their structures, neural networks can be divided into two types: feedforward networks and recurrent networks. In a feedforward network, the neurons are generally grouped into layers. Signals flow from the input layer through to the output layer via unidirectional connections, the neurons being connected from one layer to the next, but not within the same layer. The multi-layer perceptron (MLP) is perhaps the best known type of feedforward networks. For the typical multi-layer perceptron of the feed-forward mode neural network, it consists of the input layer, output layer, and hidden layer. Neurons in the input layer only act as buffers for distributing the input signals x_j to neurons in the hidden layer. Each neuron j in the hidden layer sums up its input signals x_j after weighting them with the strengths of the respective connections w_{ji} from the input layer and computes its output y_j as a function f of the sum,viz.

$$y_j = f(\sum w_{ji} x_i) \tag{1}$$

In which f can be a simple threshold function or sigmoidal, hyperbolic tangent or radial basis function. The output of neurons in the output layer is computed similarly.

The backpropagation (BP) algorithm, a gradient descent algorithm, is the most commonly adopted MLP training algorithm. It gives the change Δw_{ji} the weight of a connection between neurons i and j as follows:

$$\Delta w_{ji} = \eta \delta_j x_i \tag{2}$$

Where η is a parameter called the learning rate and δ_j is a factor depending on whether neuron j is an output neuron or a hidden neuron. For output neurons,

$$\delta_j = \left(\frac{\partial f}{\partial net_j}\right)(y_j^{(t)} - y_j) \tag{3}$$

And for hidden neurons,

$$\delta_j = \left(\frac{\partial f}{\partial net_j}\right)\sum_q w_{qj}\delta_q \tag{4}$$

In Equation (3), net j is the total weighted sum of input signals to neuron j and $y_j^{(t)}$ is the target output for neuron j.

The neural cell of each layer only affects the status of the next neural cell. If the expected output signals cannot be obtained in the output layer, the weight values of each layer of the neural cells must be modified. Erroneous output signals will be backward from the source. Finally, the signal error will arrive in certain areas with repeated propagation. After the neural networks' training procedure is complete we can start to analyze the forecast information with weight values and thresholds.

3 Flood Disaster Risk Assessment

According to the above theory, we can calculated the probabilities of each degree of flood disasters in China based on the historical data from 1950 to 2009 collected by the Ministry of Water Resources of the People's Republic of China(see Table 1). We select the set of 60 records as the large sample, and then 30 records are randomly chosen to form a small sample in order to compare the results of them by the method. Damage area, inundated area, dead population, and collapsed houses have been chosen as the disaster indicator in flood risk analysis. And by frequency analysis we classify it into four levels: small, medium, large and extreme (see Table 2).

Table 1 Values of flood indexes during 60 years

Table 2 Flood disaster rating standard

3.1 Artificial Neural Network Procession

In order to map multi-dimentional space of disaster situation to one-dimensional disaster situation, a relationship between the disaster degree and the degree indexes is needed. But it is impossible to describe the relationship using a related function. Therefore, we adopt the "simulation" and "memory" of the neural networks in flood degree evaluation. This is because the advantages of neural networks can be used to simulate and record the relationship of the input variables and output variables in the complex "function" through training and learning without any mathematical models.

We take damage area, inundated area, dead population, and collapsed houses as input variables and disaster grading value as an output variable, and then we set the nodes of the input as 4 and of the output layers as 1. It follows on from Kolmogorov's theorem(Hecht-Nielsen, 1987) that the number of nodes in the hidden layer is at least 2n + 1, where n is the number of nodes in the input layer. Since n = 4, the number of nodes in the hidden layer is at least 9. Considering the accuracy, we determine that the number of nodes in the hidden layer is 10. Thus, we can obtain the topology structure (4, 10, and 1) of the neural networks for flood degree forecasting.

The four flood grades are small, medium, large and extreme flood, whose degree value are in the interval [0,1]、[1,2]、[2,3]、[3,4]; We use the disaster grading standard boundary values (table 1) as 5 two-dimensional training samples for training and learning in the BP neural network. Meanwhile initial parameters of BP model weights and biases are randomly assigned before the commencement of training. With 100,000 cycles of training and learning in the training samples, the global error of the networks was set $E=10^{-6}$. Learning rate and impulse parameter of the network are changed adaptive, and function trainlm is used for fast training.

The calculated output values are compared with the expected values where the mean square error is 5.49809*exp(-8), indicating a good fitting. Thus the BP neural network has completed the training procedure. So we can use the BP network to forecast disaster degrees of all the samples with the weighting coefficients and the thresholds modified. The flood degree estimations of the 60 samples can be calculated out in Table 3 with BP neural network.

Table3 Disaster degree estimations based on the BP network evaluation

3.2 Flood risk assessment

In this way, the disaster degree values of all the 60 samples are obtained as shown in

Table 3. The relationship between the recurrence interval N and probability p can be expressed as $N = 1/p$, and then the exceedance probability curve of flood to disaster degree value is shown as Figure 1 using piecewise cubic hermite interpolating polynomial.

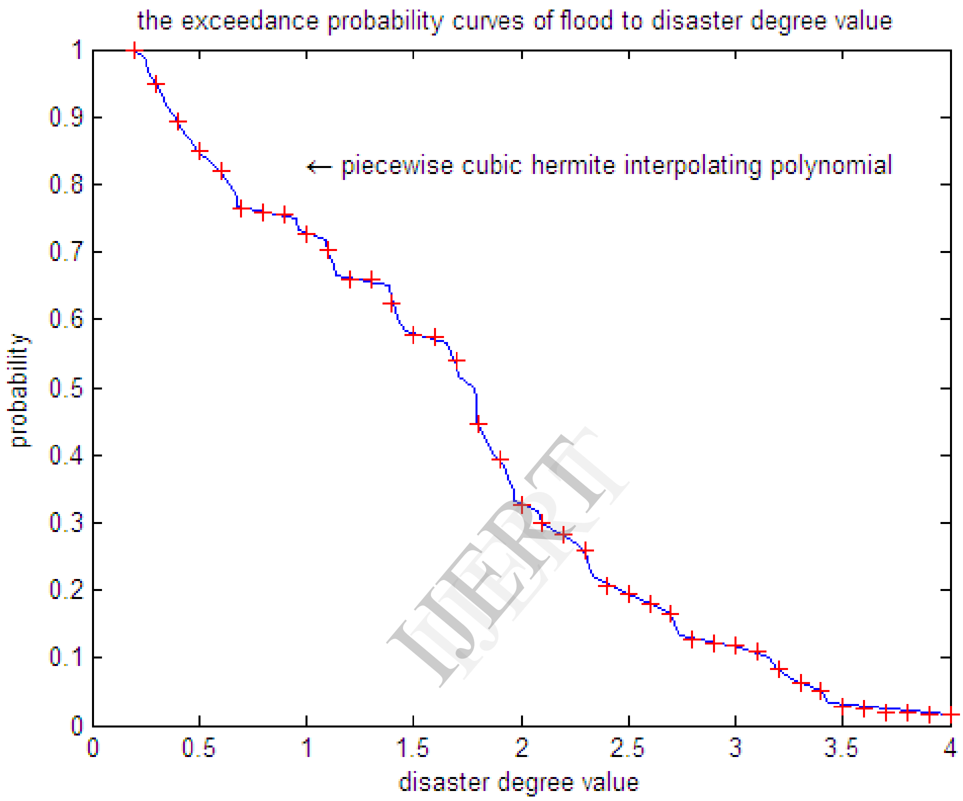

Figure 1: The exceedance probability curves of flood to disaster degree value based on neural network and piecewise cubic hermite interpolating polynomial

Due to the standard of four grades, so we have:

(a) If $0 \le x_i \le 1$, then flood degree belongs to small.

(b) If $1 < x_i \le 2$, then it belongs to medium.

(c) If $2 < x_i \le 3$, then it belongs to large.

(d) If $3 < x_i \le 4$, then it belongs to extreme.

The result in Figure 1 illustrates the risk estimation i.e. the probability of exceeding the disaster degree value. From Figure 1 we know the risk estimation is 0.1180 when the disaster index is 3, in other words, in China, floods exceeding 3 degree value (extreme floods) occur every 8.4757 years. Similarly, the probability of floods exceeding 2 degree (large floods) is 0.3246, namely China suffers the floods exceeding that intensity every 3.0807 years. This indicates the serious situation of floods in China whether on the aspect of frequency or intensity. It also means that BP neural network is useful to analyze probability risk of flood disaster. The frequency and the recurrence interval of the floods of the four grades are shown in Table 4 . These indicate the serious situation of floods in China. The frequency and the recurrence interval of the floods of the four grades are shown in Table 4.

Table 4 Flood disaster risk assessment values in China

Then we calculate the mean error between the results with large sample and small sample by neural network method and traditional statistics. From Table 5, it can be seen that errors given by neural network method are much smaller than that by statistical method, so neural network method is more efficient in solving this problem.

Compared with traditional probabilistic method, the risk values obtained by this neural network method can provide more characteristics of risk system when we analyze the risk of system. The result could help in strategic decision making to manage flood disasters.

Table 5 Comparison of two methods

4 Conclusion

Floods occur frequently in China and cause significant property losses and casualties. In order to implement a compensation and disaster reduction plan, the losses caused by flood disasters are among critically important information to flood disaster managers. This study develops a method of flood risk assessment disasters based on artificial neural network, and some preliminary findings of the analysis of the disaster flood of the state of china have been presented. The approach has been the application of the technique and it has been tested that the method is reliable and that the results are reasonable and stable.

Moreover, the analysis has shown that the method has the potentiality to be used to identify the risks of natural disasters in some area. In view of the facts that the theoretic system of flood risk assessment has been developed enough so far, and the observed series of disasters are quite short or even unavailable, the method based on BP neural network adopted in the paper is indisputably an effective and practical method. This is a new attempt that the model is applied to the case of flood disaster, and more work is needed in order to draw some final lessons from the flood disaster.

5 Acknowledgments

This work is supported by a grant from the National Basic Research Program of China (Project No.2007CB714107), a grant from the Key Projects in the National Science and Technology Pillar Program (Project No. 2008BAB29B08), and a grant from the Special Research Foundation for the Public Welfare Industry of the Ministry of Science and Technology and the Ministry of Water Resources (Project No. 201001080).

References:

Ang AHS, Tang WH (1984) Probability concepts in engineering planning and design. Decision, Risk, and Reliability. Vol. II. New York: JohnWiley & Sons, Inc.

Anselmo V, Galeati G, Palmirei S, Rossi U, Todini E (1996) Flood risk assessment using an integrated hydrological and hydraulic modeling approach: a case study. Journal of Hydrology 175: 533 - 554.

Archer DR (1981) Seasonality of flooding and the assessment of flood risk. Proceedings- Institution of Civil Engineers 70: 1023 - 1035.

Ashkar, Rousselle J (1981) Design discharge as a random variable: a risk study. Water Resources Research 17 (3): 577 - 591.

Cengiz Kahraman (2009) Risk analysis and crisis response 23 (4): 413-414

Clarke RT (1999) Uncertainty in estimation of mean annual flood due to rating-curve indefinition. Journal of Hydrology 222: 185 - 190.

Chen SY (1998) Engineering Fuzzy Set Theory and Application. Beijing: National Defence Industrial Press, 1 - 221.

Chen,SY (2002) Fuzzy recognition theory and application for complex water resources system optimization, Changchun. China: Jilin University Press.

Chen, SY (2005) Theory and model of engineering variable fuzzy sets - mathematical basis for fuzzy hydrology and water resources. Journal of Dalian University of Technology, 45(2), 308–312.

Chen, SY (2009) Theory and model of variable fuzzy sets and its application. Dalian,China: Dalian University of Technology Press.

Diaz – Granados MA, Valdes JB, Bras RL (1984) A Physically based flood frequency distribution. Water Resources Research 20 (7): 995 -1002.

Ettrick TM, Mawdllsey JA, Metcalfe A V (1987) The influence of antecedent catchment conditions on seasonal flood risk. Water Resources Research 23 (3): 481 - 488.

Fernandez B, Sales JD (1999) Return period and risk of hydrologic events Ⅰ: Mathematical Formulation. Journal of. Hydrological Engineering 4 (4): 297 - 307.

Fernandez B, Sales JD (1999) Return period and risk of hydrologic events Ⅱ: Applications. Journal of. Hydrological Engineering 4 (4): 308 - 316.

Futter MR, Mawdsley JA, Metcalfe AV (1991) Short-term flood risk prediction: a comparison of the cox regression model and a distribution model. Water Resources Research 27 (7): 1649 - 1656.

Germano Resconi (2009) Geometry of risk analysis (morphogenetic system). Stochastic Environmental Research and Risk Assessment 23 (4): 425-432

Hecht-Nielsen R (1987) Kolmogorov's mapping neural network existence theorem. IEEE Press, 11–13.

Jin JL, Zhang XL, Ding J (2002) Projection Pursuit Model for Evaluating Grade of

Flood Disaster Loss, Systems Engineering-theory & Practice,22(2), 140-144.

Jin JL, Jin BM, Yang XH, Ding J (2000) A practical scheme for establishing grade model of flood disaster loss, Journal of Catastrophology, 15(2), 1-6.

Kuczera G (1982) On the relationship between the reliability of parameter estimates and hydrologic time series data used in calibration. Water Resources Research, 18 (1): 146 - 154.

Lei XY, He CM (2004) Study and application of evaluation model of flood risk based on information diffusion theory-take storm-snowmelts flood in the Xindahe river of Akshu river basin as an example, Hydrology 24 (4) 5–8.

Liu SY, Pang YJ, Yuan LR, et al. (1994) Unascertained Mathematics and its application. Baoding: Hebei University Press, 1 - 8.

Markopoulos Ap, Manolakos De And Vaxevanidis Nm (2008) Artificial neural network models for the prediction of surface roughness in electrical discharge machining. *J. Intell. Manuf.* **19** 283–292.

Mei YD (2002) Tan GM. Risk analysis for flood prevention and safety of dam. Engineering Journal of Wuhan University, 35 (6): 11 – 15

Palm R (2007) Multiple-step-ahead prediction in control systems with Gaussian process models and TS-fuzzy models. Engineering Applications of Artificial Intelligence 20(8):1023–35

Rasmussen PF, Rosbjerg D (1989) Risk estimation in partial duration series. Water Resources Research 25 (11): 2319 - 2330.

Stedinger J R (1983) Design events with specified flood risk. Water Resources Research 19 (2): 511 - 522.

Stedinger J R, Taylor M R (1982) Synthetic Streamflow Generation 1. Model Verification and Validation. Water Resources Research 18 (4): 919 - 924.

Sen Z (1999) Simple risk calculations in dependent hydrological series. Hydrological Science 44 (6): 871 - 878.

Simon H. (2009) Neural Networks and Learning Machines. (3rd ed.) Beijing: China Machine Press, (Chapter 4)

Todorovic P, Rousselle J (1971) Some problems of flood analysis. Water Resources

Research 7 (5):1144 - 1150.

Todorovic P, Zelenhasic E (1970) A stochastic model for flood analysis. Water Resources Research 6 (6): 1641 - 1648.

Wood EF (1975) Bayesian approach to analyzing uncertainty among flood frequency models. Water Resources Research 11 (6): 839 - 843.

Xia J (2000) Grey system hydrology. Wuhan: Huazhong University of Science and Technology Press, 1 - 391.

Xu ZX (1989) Review of nonhomogeneous Poisson model on flood risk probability. Journal of China Hydrology (4):59-63

Xu ZX, Deng YL (1989) Flood risk HSPPB model and its application. Journal of Hydroelectric Engineering, (1): 46 - 55.

Xu ZX, Li JY, ItoK (1998) Clustering stochastic point process model for flood risk analysis. Stochastic Hydrology and Hydraulics 12 (1): 53- 64.

Xu ZX, Ye SZ (1988) Flood risk CSPPC model and its application. Journal of Hydraulic Engineering, (9): 1 - 8.

Xu ZX, Zeng GM (1992) A Flood Frequency HSPPC Model and its Application. Advances in Water Science, 3 (3): 175 - 180.

Tables and Figures:

Table 1: Values of flood indexes during 60 years

year	disaster area (thousand hectares)	inundated area (thousand hectares)	Dead Population (persons)	collapsed houses (ten thousand)
1950	6559.00	4710.00	1982	130.50
1951	4173.00	1476.00	7819	31.80
1952	2794.00	1547.00	4162	14.50
1953	7187.00	3285.00	3308	322.00
1954	16131.00	11305.00	42447	900.90
1955	5247.00	3067.00	2718	49.20
1956	14377.00	10905.00	10676	465.90
1957	8083.00	6032.00	4415	371.20
1958	4279.00	1441.00	3642	77.10
1959	4813.00	1817.00	4540	42.10
1960	10155.00	4975.00	6033	74.70
1961	8910.00	5356.00	5074	146.30
1962	9810.00	6318.00	4350	247.70
1963	14071.00	10479.00	10441	1435.30
1964	14933.00	10038.00	4288	246.50
1965	5587.00	2813.00	1906	95.60
1966	2508.00	950.00	1901	26.80
1967	2599.00	1407.00	1095	10.80
1968	2670.00	1659.00	1159	63.00
1969	5443.00	3265.00	4667	164.60
1970	3129.00	1234.00	2444	25.20
1971	3989.00	1481.00	2323	30.20
1972	4083.00	1259.00	1910	22.80

year	disaster area (thousand hectares)	inundated area (thousand hectares)	Dead Population (persons)	collapsed houses (ten thousand)
1973	6235.00	2577.00	3413	72.30
1974	6431.00	2737.00	1849	120.00
1975	6817.00	3467.00	29653	754.30
1976	4197.00	1329.00	1817	81.90
1977	9095.00	4989.00	3163	50.60
1978	2820.00	924.00	1796	28.00
1979	6775.00	2870.00	3446	48.80
1980	9146.00	5025.00	3705	138.30
1981	8625.00	3973.00	5832	155.10
1982	8361.00	4463.00	5323	341.50
1983	12162.00	5747.00	7238	218.90
1984	10632.00	5361.00	3941	112.10
1985	14197.00	8949.00	3578	142.00
1986	9155.00	5601.00	2761	150.90
1987	8686.00	4104.00	3749	92.10
1988	11949.00	6128.00	4094	91.00
1989	11328.00	5917.00	3270	100.10
1990	11804.00	5605.00	3589	96.60
1991	24596.00	14614.00	5113	497.90
1992	9423.30	4464.00	3012	98.95
1993	16387.30	8610.40	3499	148.91
1994	18858.90	11489.50	5340	349.37
1995	14366.70	8000.80	3852	245.58
1996	20388.10	11823.30	5840	547.70
1997	13134.80	6514.60	2799	101.06
1998	22291.80	13785.00	4150	685.03

year	disaster area (thousand hectares)	inundated area (thousand hectares)	Dead Population (persons)	collapsed houses (ten thousand)
1999	9605.20	5389.12	1896	160.50
2000	9045.01	5396.03	1942	112.61
2001	7137.78	4253.39	1605	63.49
2002	12384.21	7439.01	1819	146.23
2003	20365.70	12999.80	1551	245.42
2004	7781.90	4017.10	1282	93.31
2005	14967.48	8216.68	1660	153.29
2006	10521.86	5592.42	2276	105.82
2007	12548.92	5969.02	1230	102.97
2008	8867.82	4537.58	633	44.70
2009	8748.16	3795.79	538	55.59

Table 2 Flood disaster rating standard

Disaster level	Damage area (thousand hectares)	Inundated area (thousand hectares)	Dead population (persons)	Collapsed houses (ten thousand)	Recurrence interval (years)	Degree value
Small flood	0~9045	0~4989	0~3446	0~112.1	<2	0~1
Medium flood	9045~14197	4989~8216.7	3446~5113	112.1~247.7	2~5	1~2
Large flood	14197~20388	8216.7~13000	5113~10676	247.7~754.3	5~20	2~3
Extreme flood	20388~80000	13000~50000	10676~100000	754.3~5000	>20	3~4

Table 3：disaster degree estimations based on the BP network evaluation during the 60 years in China

Year	Degree value	Year	Degree value
1950	0.4968	1980	1.1020
1951	3.4372	1981	2.6992
1952	1.9236	1982	1.6720
1953	1.6588	1983	3.4072
1954	3.1968	1984	2.0740
1955	1.1336	1985	0.6060
1956	0.2592	1986	0.3456
1957	0.3244	1987	1.9396
1958	2.3124	1988	2.3460
1959	2.4576	1989	1.7848
1960	2.7404	1990	2.7104
1961	0.9520	1991	3.2744
1962	0.3892	1992	1.8760
1963	0.2496	1993	3.0172
1964	0.1996	1994	2.2672
1965	1.0908	1995	2.100
1966	1.7048	1996	3.1532
1967	1.4108	1997	2.5816
1968	1.4584	1998	2.1920
1969	1.4224	1999	0.5588
1970	1.8180	2000	0.2980
1971	1.7900	2001	0.4520
1972	1.8504	2002	0.4144
1973	1.9704	2003	0.6748
1974	1.3876	2004	0.6804
1975	4.0000	2005	1.7912
1976	1.7960	2006	1.1428

1977	0.9656	2007	2.3168
1978	1.7068	2008	0.6428
1979	1.9700	2009	1.3928

Table 4：Flood disaster risk evaluation values

Disasters level	Small flood	Medium flood	Large flood	Extreme flood
Exceedance probability risk	1.0000	0.7269	0.3246	0.1180
Recurrence interval(years)	1.0000	1.3757	3.0807	8.4757

Table 5 Comparison of two methods

Method	BP network	Statistics
Mean error	0.0421	0.0428