International Journal of the Physical Sciences Vol. 6(6), pp. 1374-1378, 18 March, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS10.648
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

Neural network model and geographic grouping for risk assessment of debris flow

Jeng-Wen Lin[1,2]

[1]Department of Civil Engineering, Feng Chia University, Taichung, Taiwan 407, R.O.C., Taiwan.
[2]Du Lin Construction Co. Ltd., Taichung, Taiwan 407, R.O.C., Taiwan. E-mail: jwlin@fcu.edu.tw. Tel: +886-4-24517250#3150. Fax: +886-4-24516982.

Accepted 14 March, 2011

There were 1500 rivers and streams in Taiwan with potential sources of debris flows. Debris flows often occurred after earthquakes and heavy rains, leading to severe damage to human lives and properties. This study aimed to develop an accurate neural network-based risk assessment model of debris flows. The back propagation network was adopted because of its supervised training nature and of its ability to solve complex pattern-matching problems. Real cases of debris flows that occurred in Hualien area of Taiwan from 2007 to 2008 were taken as the database. Such updated data were necessary for accurate predictions of debris flows since the hydrological and geologic data could vary as time proceeds. According to related documentation, this study selected 6 influential factors, including average gradient, catchment area, effective catchment area, accumulated rainfall, rainfall intensity, and geologic condition, as input variables. The results proved that the established model was quite suitable for debris flows risk assessment; the obtained normalized relative error was 8.56% and reduced to 7.04% if geographically close sites were collected as groups.

Key words: Back propagation, debris flow, geographic grouping, neural network, risk assessment.

INTRODUCTION

Debris flows often occurred suddenly in mountain areas with steep slopes, characterizing as high particle concentrations and high flow speeds, coming with full violence, strong physical force and high destructive power. After an earthquake, loose earth became the source of debris and thus the frequency of the occurrence of debris flows was increased (Chang et al., 2009; Liu et al., 2009). The amount of critical rainfall required to trigger debris flows reduced significantly after earthquakes (Chang et al., 2009). Debris flows became products of natural disasters, such as earthquakes and typhoons. However, some conditions of the occurrence of debris flows were difficult to investigate and forecast. A lot of experts and scholars were therefore dedicated to researches in the mechanism and modeling of debris flows.

Among the prediction models for warning of potential debris flows, neural network was the most popular one in the literature although this methodology had some deficiencies. It was well recognized that artificial neural network approaches had some crucial defects such as the convergence of the estimation to local minima, and it required a considerable amount of time to properly train the network (Masri et al., 1999). However, this method was particularly suitable for real life and system unknown applications, solving optimization problems in a very large parameter space (Lin, 2001). For example, Wang et al. (2003) presented a three-layer feed-forward neural network model to forecast the surge of debris flows according to the time series data collected in the Jiangjia Ravine, situated in north part of Yunnan Province of China. Liu et al. (2006) presented a neural network (NN) based model to assess the regional hazard degree of debris flows in Lake Qionghai Watershed, China; from which debris flow management and strategies were proposed for each sub-watershed. The NN model was used in order to effectively handle the nonlinearity and uncertainty inherent in the debris flow hazard analysis. This NN model determined the number of nodes in the middle layer by varying the number of nodes until an optimal performance was achieved.

In order to establish a stable and reliable analytical model for the occurrence prediction of debris flows, Chang and Chao (2006) applied an artificial neural network model constructed by seven significant factors and used back-propagation algorithm. To improve the effectiveness of the debris flow warning system, Chang et al. (2007) developed a NN (near neighbours) framework to connect all the clusters produced by the shared near neighbours (SNN) method, whereas the connected weights of the network were adjusted through a supervised learning method. By using the SNN, the hydro-geological data set could be meaningfully clustered into categories, which were characterized as “occurrence” or “no-occurrence” of debris flows.

In order to evaluate the impact of rainfall on debris flows, Chen et al. (2005) implemented a rainfall-induced debris flow warning system employing real-time rain gauge data. The effects by which the regional rainfall patterns (antecedent rainfall, duration, intensity, cumulative rainfall) and geological settings combined together to trigger a debris flow were analyzed for real-time monitoring. Lu et al. (2007) developed a GIS-based decision support system, which incorporated local topographic and rainfall effects on debris flow vulnerability. Rainfall at a scale compatible with the digital elevation model resolution was obtained using a NN with a wind-induced topographic effect and rainfall derived from satellite rain estimates, along with an adaptive inverse distance weight method. A critical element of the work by Winter et al. (2010) was the continuing development of a rainfall threshold to indicate conditions likely to produce debris flow activity, and the development of a tentative threshold was provided. Any change in rainfall patterns had the potential to affect the frequency and intensity of debris flow and thus the effectiveness of the associated management strategy for such events.

Liu et al. (2009) found that the Chi-Chi earthquake induced many landslides in the region of central Taiwan, especially for landslides turned into debris flows when typhoon Toraji struck Taiwan in 2001. The topography, hydrogeology, and rainfall factors were introduced into a probabilistic NN to build a prediction model for the probability of the occurrence of debris flows. The results revealed that the susceptibility of debris flows was elevated after the Chi-Chi earthquake. Chang et al. (2009) also investigated the impact of the Chi-Chi earthquake on the occurrence of debris flows using artificial neural networks (ANNs), taking hydrological and geomorphologic influences into account. Significant differences before and after the earthquake were found. In particular, the size of landslide area was proportioned to the occurrence of debris flows. Pradhan and Lee (2009) used remote sensing data, GIS tools and ANNs with different training sites for landslide hazard and risk analysis at Penang Island in Malaysia. Landslide locations were identified through interpretation of aerial photographs and through field surveys.

Other approaches for the risk assessment of debris flows included the work by Chang et al. (2010), who proposed using the genetic algorithm to weigh seven variables according to a principle similar to natural selection. The genetic algorithm described a global optimization technique according to the principles of organic evolution: Selection, crossover, and mutation, involving randomness in the search procedure without requiring any derivative information for the performance criterion (Masri et al., 1999). The weighted variables were then fed into a NN model to predict debris flow occurrences based on relevant factors.

Although, there were many successful models in forecasting the debris flow, the designs were contingent on the systems settings. A general and easy to use approach was seldom proposed. In addition, the recently updated data must be used to assess the hazard degree of debris flows because the hydrological and geologic data could vary as time progresses. Hence, the work of model construction for the prediction of debris flows was still an open arena, requiring continuous and further studies. Based on these arguments, this study investigated the risk of debris flows in the Hualien County, Taiwan. This terrain consisted of mountains, rivers, and plains, in which mountains cover most of the whole area: about 87%, referring to an area with potential debris flows. Related data were retrieved from 25 observation stations in this area. Hydrological and geologic conditions were included in the data as influential factors for analyses (Hsieh et al., 1992; Lin et al., 2009). The data from real cases provided by the Soil and Water Conservation Bureau, Council of Agriculture, Executive Yuan were used for validation. The back propagation network (BPN) decision system was adopted to determine the risk of debris flows. Normalized relative errors were computed to verify the accuracy of the established model.

NEURAL NETWORK DECISION SYSTEMS

An artificial neural network (ANN) was created as a methodology for systems analysis. An ANN is not only useful in addressing problems (Bayram et al. 2011) requiring recognition of complex patterns and performing nontrivial mapping functions but is also capable of “learning”; that is, it can be trained to improve its performance by either supervised or unsupervised learning (Chen, 2004; Chang and Chao, 2006).

The back propagation network (BPN) is one kind of supervised training systems including gradient-based techniques (Lin et al., 2011). Such a network is designed to operate as a multilayer and feed-forward network, using the supervised mode of learning as shown in Figure 1 (Chen, 2004). It has basic elements including (1) the nodes (also called neurons or units), (2) the layers, (3) the links, denoted by arrows, each of which represents a numerical weight, and (4) the activation functions. The network learns a predefined set of input-output sample pairs through a cycle of the feed-forward and back propagation weight training.

After an input pattern has been applied to the first layer of network units, it is propagated through each upper layer until an output is generated. Assume that the BPN consists of three layers:

The input, hidden, and output layer, such that the input-output relation can be expressed as (Masri et al., 1999):

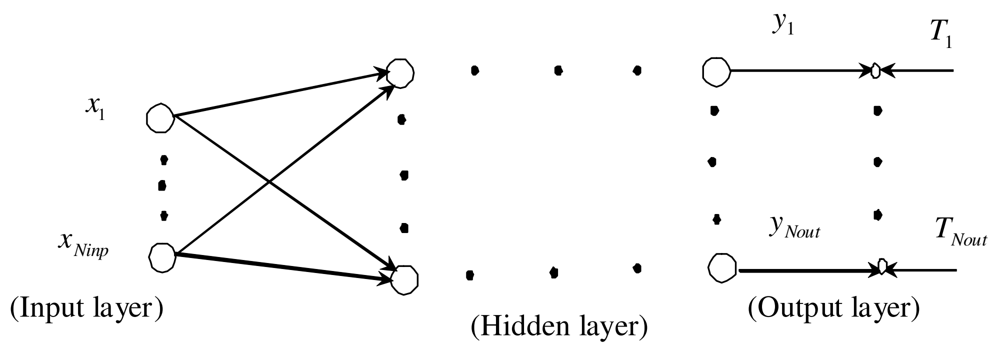

Figure 1. Framework of neural network.

$$y_i = f\left\{\sum_{j=1}^{Nout}\left[w_{ij} f\left(\sum_{k=1}^{Ninp} v_{jk} x_k + b_{vj}\right) + b_{wi}\right]\right\} \tag{1}$$

where w_{ij} and v_{jk} are interconnection weights, b_{wi} and b_{vj} are bias or threshold terms, while $f()$ is the activation function used in both hidden and output layers. The activation function introduces the nonlinear effect to the network and maps the result of computation to the output values between 0 and 1, in order to determine whether there will be a debris flow (Chang and Chao, 2006). The activation function often uses tan-sigmoid function or log-sigmoid function in a multilayer feed-forward network. In this study, the log-sigmoid function is used as the transformation function in both hidden and output layers.

The network output must be compared with the desired output so that the network will learn the desired behavior. In gradient-based techniques, a performance criterion is defined as the error between the desired and actual network response (Masri et al., 1999):

$$J = \left\|\mathbf{y} - \mathbf{y}_d\right\|^2 = \left\|\mathbf{e}_y\right\|^2 \tag{2}$$

where $\mathbf{y}$ denotes the network output vector and $\mathbf{y}_d$ the desired output vector.

Since $\mathbf{y}$ depends on the network parameters (weights), the value of J can be reduced by adjusting the network weights in the direction of the negative gradient of the performance function:

$$\Delta w_{ij} = -\frac{\eta}{2}\frac{\partial J}{\partial w_{ij}} \tag{3}$$

where $\eta > 0$ is the learning rate of the algorithm.

The back-propagation algorithm takes advantage of the parallelism inherent in the network to calculate the error in a forward pass and the adjustment Δw_{ij} in a backward pass (Masri et al., 1999). The error signals are transmitted backward from the output layer to each node in the intermediate layers until each node in the network has received an error signal that describes its relative contribution to the total error. According to the error signal received, the connection weights are updated by each node so that the network converges toward a state with all the training patterns encoded (Chen, 2004).

To perform training for a BPN, the weight and bias values should be initialized first. The Input vector and target output vector are entered using sample datasets. The number of neurons in the hidden layer depends on complexity of the problem. For a general problem, an empirical equation is considered:

$$numberof\ neurons = numberof\ inputelements + numberof\ outputelements \tag{4}$$

The best number of neurons can be found by trial and error. Then the batch training is conducted. The training process is to apply the optimized multi-dimensional algorithm with multiple inputs and outputs to the relationships between the input and the output vectors of the network.

ANALYSIS AND DISCUSSION

In this study, the neural network toolbox of MATLAB R2009a version was used to construct the BPN model for the debris flow prediction, with one hidden layer. Six influential factors, including the average gradient, catchment area, effective catchment area, accumulated rainfall, rainfall intensity, and geologic condition were selected as input variables. The input variables of the BPN model were normalized so that all the values fell in the range of 0 to 1. The objective was then to make sure that all output values of the model were between 0 and 1. For the debris flow prediction, this study defined “likely to occur” as having an output value higher than or equal to 0.5, while “not likely to occur” as having an output value lower than 0.5. After referencing the empirical Equation (4) together with trial and error, the “best” number of neurons for the hidden layer was selected as 6.

There were 25 sites being studied, including 20 sites (80%) being randomly selected for training and the remaining 5 sites (20%) for testing to show predictive accuracy of the BPN (back propagation network). The BPN training and testing output values are listed in Tables 1 and 2, respectively. These values are compared with the actual number of occurrences of debris flows in Hualien area. The comparison uses the normalized

Table 1. The BPN training results of debris flow assessment.

Site	Output by BPN	Assessment result
Bazi stream	0.3643	No occurrence (Success)
Chongde village	0.5252	Occurrence (Success)
Chutian	0.7541	Occurrence (Failure)
Dafong village	0.3741	No occurrence (Failure)
Dafu village	0.9105	Occurrence (Success)
Daquan village	0.9041	Occurrence (Success)
Fahua mountains	0.4873	No occurrence (Success)
Fonglin stream	0.7105	Occurrence (Success)
Fushi village	0.2871	No occurrence (Failure)
Fuxing stream	0.4317	No occurrence (Success)
Hepin village	0.5849	Occurrence (Success)
Hongye stream	0.8774	Occurrence (Failure)
Jiamin villiage	0.6891	Occurrence (Success)
Shapodang stream	0.8147	Occurrence (Success)
Shofeng stream	0.8714	Occurrence (Failure)
Shuiyuan village	0.5771	Occurrence (Success)
Tongmen	0.5426	Occurrence (Success)
Tongshing tribe	0.5385	Occurrence (Failure)
Wanrong village	0.8421	Occurrence (Success)
Xumeiji stream	0.7357	Occurrence (Success)

Table 2. The BPN testing results of debris flow assessment.

Site	Output by BPN	Assessment result
Dashing village	0.9104	Occurrence (Success)
Jingmei village	0.5473	Occurrence (Success)
Qingshui stream	0.5375	Occurrence (Failure)
Rongshu tribe	0.3467	No occurrence (Success)
Xiulin village	0.3587	No occurrence (Failure)

relative error (Lin et al., 2009):

$$E = \frac{\sqrt{\frac{1}{n}\left[(a_1-b_1)^2+(a_2-b_2)^2+\cdots+(a_n-b_n)^2\right]}}{\sqrt{\frac{1}{n}\left[a_1{}^2+a_2{}^2+\cdots+a_n{}^2\right]}} \qquad (5)$$

where n represents the number of terms, a denotes the actual value ratio and b the output value ratio.

It measures the difference between the ratio of real case estimations and the ratio of the BPN outputs. Both ratios are dimensionless, satisfying the consistency of unit. If the numerators of the output values are close to the ones of the actual values and the same for the denominators, the utilization of the normalized relative errors becomes meaningful. The total relative error obtained is 8.56%.

In order to reduce the obtained relative error in assessing debris flows, the factor of the geographic location is considered. Two groups are found, each containing the sites within 20 km of relative distance. Group A consists of 6 sites: Fonglin Stream, Jiamin Villiage, Jingmei Village, Shuiyuan Village, Tongmen, and Xiulin Village. Group B consists of 4 sites: Dafong Village, Dafu Village, Daquan Village, and Dashing Village. Under this condition and using the assessment results in Tables 1 and 2, the total relative error is reduced to 7.04%. The corresponding ratio of success has been promoted. Table 3 displays this summary of successful ratio for debris flow predictions using the BPN for the training and testing phases as well as for the ungrouped and grouped sites by considering the geographic location. It is clear to see the merit of the geographic grouping: The ratio of success for debris flow predictions has been raised from 70 to 86% in the training phase and from 60 to 67% in the testing phase, respectively. The entire effect of the

Table 3. Summary of debris flow predictions using BPN.

	Total	Sample		Ratio (%)	
		Success	Failure	Success	Failure
Training	Ungrouped	14	6	70	30
	Grouped	6	1	86	14
Testing	Ungrouped	3	2	60	40
	Grouped	2	1	67	33

geographic input has increased the accuracy from 68 to 80%. Future studies are required to improve the accuracy of debris flow predictions. Geographic locations can be further explored to find out if it is possible to raise the prediction ratio of success on the one hand. On the other, the success ratio depends on the BPN features, including the inputs, intermediate nodes, layers, the activation function, and the gradient-based algorithm for minimizing the mean-squared error function. Any one of these conditions can influence the accuracy of the risk assessment of debris flows.

Conclusion

An efficient and simple approach for the risk assessment and modeling of debris flows is developed in this study. The back-propagation network (BPN) decision system is adopted to determine the risk of debris flows and verified by the real case estimations for Hualien area of Taiwan. Six influential factors are used as input variables of the BPN: The average gradient, catchment area, effective catchment area, accumulated rainfall, rainfall intensity, and geologic condition. Database from 2007 to 2008 are collected for accurate predictions of debris flows because the hydrological and geologic data may vary as time proceeds. A normalized relative error is presented to compare the BPN outputs with the real case estimations. The relative error obtained is 8.56% for the 25 sites in Hualien area of Taiwan. Such an error is reduced to 7.04% if 10 geographically close sites are collected as two groups, with the corresponding prediction ratio of success being promoted from 68 to 80%. For possible future research directions, more data should be collected to demonstrate the efficiency of the geographic grouping in addition to the improvement in element settings of the BPN.

ACKNOWLEDGMENTS

The author would like to thank the National Science Council of the Republic of China, Taiwan, for the financial support of this research under Contract No. NSC 96-2221-E-035-038. Special thanks are due to Mr. Cheng-Yi Peng for his assistance in running related simulations.

REFERENCES

Bayram AF, Gültekin SS, Söğüt AR (2011). Specifications of thermal waters and their classification on the base of neural network method: Examples from Simav geothermal area, Western Turkey. Int. J. Phys. Sci., 6(1): 43-51.

Chang FJ, Tseng KY, Chaves P (2007). Shared near neighbours neural network model: a debris flow warning system. Hydrol. Proc., 21(14): 1968-1976.

Chang FJ, Chiang YM, Lee WS (2009). Investigating the impact of the Chi-Chi earthquake on the occurrence of debris flows using artificial neural networks. Hydrol. Proc., 23(19): 2728-2736.

Chang TC, Chao RJ (2006). Application of back-propagation networks in debris flow prediction. Engr. Geol., 85(3-4): 270-280.

Chang TC, Wang ZY, Chien YH (2010). Hazard assessment model for debris flow prediction. Environ. Earth Sci., 60(8): 1619-1630.

Chen CW (2004). Stability of nonlinear structural control via fuzzy control and neural network. PhD dissertation, National Central University, Taiwan.

Chen CY, Chen TC, Yu FC, Yu WH, Tseng CC (2005). Rainfall duration and debris flow initiated studies for real-time monitoring. Environ. Geol., 47(5): 715-724.

Hsieh CL, Chiang CH, Chen LJ (1992). Field survey and analysis of debris flow in Hualien and Taitung. J. Chin. Soil Water Conserv., 23(2): 109-122. (in Chinese)

Lin JW (2001). Adaptive algorithms for the identification of nonlinear structural systems. PhD dissertation, Columbia University, New York, USA.

Lin JW, Huang CW, Shih CH, Chen CY (2011). Fuzzy Lyapunov stability analysis and NN modeling for tension leg platform systems. J. Vibr. and Cont., 17(1): 151-158.

Lin JW, Peng CY, Chung SH (2009). Occurrence prediction and modeling framework of debris flows. The 15th Conference on Pavement Engineering & Int'l Pavement Expert Summit held at Feng Chia University, Taiwan. No. T008: 1-7.

Liu CN, Dong JJ, Peng YF, Huang HF (2009). Effects of strong ground motion on the susceptibility of gully type debris flows. Engr. Geol., 104(3-4): 241-253.

Liu Y, Guo HC, Zou R, Wang LJ (2006). Neural network modeling for regional hazard assessment of debris flow in Lake Qionghai Watershed, China. Environ. Geol., 49(7): 968-976.

Lu GY, Chiu LS, Wong DW (2007). Vulnerability assessment of rainfall-induced debris flows in Taiwan. Natural Hazards, 43(2): 223-244.

Masri SF, Smyth AW, Chassiakos AG, Nakamura M, Caughey TK (1999). Training neural networks by adaptive random search techniques. J. Engr. Mech., 125(2): 123-132.

Pradhan B, Lee S (2009). Landslide risk analysis using artificial neural network model focusing on different training sites. Int. J. Phys. Sci., 4(1): 1-15.

Wang XK, Huang E, Cui P (2003). Simulation and prediction of debris flow using artificial neural network. Chin. Geog. Sci., 13(3): 262-266.

Winter MG, Dent J, Macgregor F, Dempsey P, Motion A, Shackman L (2010). Debris flow, rainfall and climate change in Scotland. Quart. J. Engr. Geol. Hydrogeol., 43(4): 429-446.