Journal of Data Science **9**(2011), 65-81

An Assessment of the Use of an Advanced Neural Network Model with Five Different Training Strategies for the Preparation of Landslide Susceptibility Maps

Biswajeet Pradhan
University Putra Malaysia

Abstract: Data collection for landslide susceptibility modelling is often an almost inhibitive activity. This has been the reason for quite sometimes landslide was described and modelled on the basis of spatially distributed values of landslide related attributes. This paper presents landslide susceptibility analysis at Selangor area, Malaysia, using artificial neural network model with the aid of remote sensing data and geographic information system (GIS) tools. To meet the objectives, landslide locations were identified in the study area from interpretation of aerial photographs and supported with extensive field surveys. Then, the landslide inventory was grouped into two categories: (1) training data (2) testing data. Further, topographical, geological data and satellite images were collected, processed, and constructed into a spatial database using GIS tools and image processing techniques. Nine landslide occurrence attributes were selected and analyzed using an artificial neural network model to generate the landslide susceptibility maps. Landslide location data (training data) were used for training the neural network and five training sites were selected randomly in this case. The use of five training sites ensemble to investigate the model reliability, including the role of the thematic variables used to construct the model, and the model sensitivity to changes in the selection of the training sites. By studying the variation of the neural network's susceptibility estimate, the error associated with the model is determined. The results of the neural network analysis are shown on five sets of landslide susceptibility maps. Then the susceptibility maps were validated using "receiver operating characteristics (ROC)" method as a measure for the model verification. Landslide training data which were not used during the training of the neural network was used for the verification of the maps. The results of the analysis were verified using the landslide location data and compared between five different cases. Qualitatively, the model seems to give reasonable results with accuracy observed was 87%, 83%, 85%, 86% and 82% for five different training sites respectively.

Key words: Artificial neural network, GIS, landslide, Malaysia, susceptibility.

1. Introduction

Landslide presents a significant constraint to development in many parts of Malaysia which experiences frequent landslides, with the most recent occurring in 2000, 2001, 2004, 2007 and 2008. Damages and losses are regularly incurred because; historically there has been too little consideration of the potential problems in land use planning and slope stability analysis. Landslides are mostly occurred in Malaysia mainly due to heavy tropical rainfall. In recent years greater awareness of landslide problems has led to significant changes in the control of development on unstable land. So far, few attempts have been made to predict these landslides or preventing the damage caused by them. In last few years, landslide susceptibility analysis using GIS and data mining such as fuzzy logic, and artificial neural network methods have been applied by researchers in different countries (Ercanoglu and Gokceoglu 2002; Gomez and Kavzoglu, 2005; Pistocchi *et al.* 2002; Lee *et al.* 2003a, 2003b, 2004; Pradhan 2010a, 2010b, 2010c, 2010d; Pradhan and Lee, 2010a, 2010b, 2010c; Pradhan, Lee and Buchroithner, 2010a, 2010b; Pradhan, Oh and Buchroithner, 2010; Pradhan and Buchroithner, 2010; Pradhan *et al.*, 2010; Pradhan and Youssef, 2010; Pradhan and Pirasteh, 2010). But their result output can not be directly used in the Malaysian landslide susceptibility analysis. This is due to the changes in the geographical environment set up, litho types and different climatic condition etc. The local geographical settings cause different landslide types based on completely different mechanisms and are absolute incomparable. Through scientific analysis of landslides, we can assess and predict landslide-susceptible areas, and thus decrease landslide damage through proper preparation. To achieve this aim, landslide susceptibility analysis techniques have been applied, and verified in the study area using artificial neural network.

In landslide literature, there have been many studies carried out on landslide hazard evaluation using GIS. There are number of different approaches for the measurement of landslide hazard, including direct and indirect heuristic approaches, and deterministic, probabilistic, statistical and data mining approaches. Recently, there have been studies on landslide hazard evaluation using GIS, and many of these studies have applied probabilistic models (Baeza and Corominas, 2001; Lee and Min, 2001; Lee, Chwae and Min, 2002; Gokceoglu, Sonmez, and Ercanoglu, 2000; Lee and Choi, 2003; Lee, Choi and Min, 2004; Lee and Pradhan, 2006, 2007; Pradhan, Singh and Buchroithner, 2006; Youssef *et al.*, 2009). One of the multivariate models available, the logistic regression models, has also been applied to landslide hazard mapping (Dai and Lee, 2002; Pradhan *et al.*, 2008a; Pradhan and Lee, 2009a; Pradhan and Youssef, 2009). In last few years, a new approach to landslide hazard evaluation using GIS, data mining using fuzzy logic, and artificial neural network models have been applied (Catani *et al.*, 2005; Ercanoglu and Gokceoglu, 2002; Ermini, Catani, and Casagli, 2005; Neaupane and Achet, 2004; Pradhan, Lee and Buchroithner, 2009; Pradhan *et al.*, 2008b; Pradhan and Lee, 2009a, 2009b; Pradhan and Lee, 2007).

In recent years, Lee and Pradhan (2006), Lee and Pradhan (2007), and Pradhan and Lee (2009a) investigated the landslide susceptibility in Malaysia. Pradhan and Lee (2009b) evaluated three models for landslide susceptibility analysis using frequency ratio, logistic regression and artificial neural network model. Pradhan and Lee (2009b) analyzed the rainfall precipitation in the Penang area using back-propagation neural networks. However, they could not have a detail landslide hazard analysis due to lack of rainfall intensity data. Slope stability and rainfall intensity is very important factors causing most of the landslides in Malaysia. Besides these two important factors of rainfall and slope, soil weight and distance to drainage are also important factors in some regions. Pradhan, Lee and Buchroithner (2009), investigated the landslide susceptibility using fuzzy model at Penang Island and they pointed out some important factors, such as topographic slope, topographic aspect, topographic curvature, distance to drainage, lithology, distance to faults, soil texture, landcover, vegetation index and accumulated rainfall intensity.

The objective and motivation of this study is to demonstrate a data mining model for the landslide hazard analysis with the aid of GIS. In order to get a stable and reliable result, in this paper, nine geological and geomorphological factors including, topographic slope, topographic aspect, topographic curvature, distance to drainage, lithology, distance to faults, soil texture, landcover and normalized difference vegetation index (ndvi) to predict landslide susceptible areas. These nine factors constructed an ANN using the back propagation algorithm for landslide susceptibility analysis. To meet the objectives, firstly the ANN model was trained using training sites which can be directly utilized for the landslide susceptibility analysis as long as the recorded nine factors are fed into an ANN model. Five different training samples were selected to train the ANN in order o avoid bias effect in the final results. Finally, the results of the landslide susceptibility maps were validated using the existing landslide location data with the aid of receiver operating characteristics (ROC) approaches.

2. Study Area

The eastern part of Selangor state has suffered much landslide damage following heavy rains and it was selected as a suitable pilot area to evaluate frequency and distribution of landslides. The study area is located approximately between $3°23'53.6^{"}E$ and $3°45'18.05^{"}$ and $101°30'55.33^{"}$ N and $101°3'36.3^{"}$N. The landuse of the state is mainly peat swamp forest, plantation forest, inland forest, scrub, grassland and ex-mining area. The landform of the area ranges from very flat terrain, especially for the peat swamp forest, ex mining, grassland and scrub area, to quite hilly area for the natural forest ranging between 0– 420 m. a.s.l. Based on Malaysian Meteorological Services Department, the temperature of northern part of Selangor is between 29° C to 32° C and mean relative humidity of 65% to 70%. The rainfall from 58 mm to 240 mm per month was recorded in the study area (Tanjung Karang weather station provided by Malaysian Meteorological Services Department).

Tectonically, state of Selangor forms a part of the Sunda Shield. Its fold-mountain system has northern to north-western regional trend which is a southern continuation of eastern Burma through Thailand, and Indonesian Borneo. All the systems, forms the Cambrian to the Quaternary rock formations in Peninsular Malaysia. The pre-Triassic rocks are essentially marine whereas the post-Triassic rocks are characteristically non-marine. The Triassic rocks are of both marine and non-marine origins but in general, the non-marine deposits occur in the Upper Triassic. Within Selangor, it is believed that sedimentation was continuous throughout the Paleozoic and Mesozoic but because of the instability of the basin, the sedimentary record is not complete. Major breaks are apparent between the Paleozoic, Mesozoic and Cenozoic groups of rocks. Granitoids occupy almost half of the study area. These bodies commonly form topographic highs, the largest of which is the main range situated on the eastern flank of the area. Although many of the granite bodies are aligned parallel to the structural trend, they do not always occupy the anticlinal ridges of the sedimentary covers and some of the smaller bodies are found to cut across the structural trends. Regional metamorphism is widespread and most of the Paleozoic and Mesozoic rocks show slight to moderate deformation. In general, the older rocks show a greater degree of metamorphism than the younger rocks. Contact metamorphism is not intense and generally forms narrow aureoles around the igneous bodies. There are at least four major episodes of granite emplacement and it is believed that much of the known mineralization occurred during the later episodes and commonly associated with faulting. At least three sets of faults have been recognized in the study area, the youngest of which is at most post-Early Cretaceous in age.

3. Spatial Database Construction

The data used is shown in Figure 1 and Table 1. Accurate detection of the location of landslides is very important for probabilistic landslide hazard analysis. The application of remote sensing methods, such as aerial photographs and satellite images, are used to obtain significant and cost-effective information on landslides. In this study, 1:25,000–1:50,000-scale aerial photographs were used to detect the landslide locations. These photographs were taken during the period

1981–2000, and the landslide locations were detected by photo interpretation and the locations verified by fieldwork. Recent landslides were observed in aerial photographs from breaks in the forest canopy, bare soil, or other geomorphic characteristics typical of landslide scars, for example, head and side scarps, flow tracks, and soil and debris deposits below a scar. To assemble a database to assess the surface area and number of landslides in each of three study areas, a total of 327 landslides were mapped in a mapped area of 8,179.28 km^2.

Table 1: Various attribute data layers used in the analysis

Classification	Sub-Classification	GIS Data Type	Scale
Geological Hazard	Landslide	Point coverage	1:25,000
Basic Map	Topographic Map	Line and Point coverage	1:25,000
	Geological Map	Polygon coverage	1:63,300
	Drainage	Line coverage	1: 25,000
	Land Cover	GRID	30 m × 30 m
	Soil Map	GRID	10 m × 10 m
	Vegetation Index (ndvi)	GRID	10 m × 10 m

There were nine factors considered, and the factors were extracted from the constructed spatial database. The factors were transformed into a vector-type spatial database using the GIS, and landslide-related factors were extracted using the database. A digital elevation model (DEM) was created first from the topographic database. Contour and survey base points that had elevation values from the 1:25,000-scale topographic maps were extracted, and a DEM was constructed with a resolution of 10 m. Using this DEM, the slope angle, slope aspect, and slope curvature were calculated. In the case of the curvature, negative curvatures represent concave, zero curvature represent flat and positive curvatures represents convex. The curvature map was prepared using the avenue routine in ArcView 3.2. In addition, the distance from drainage was calculated using the topographic database. The drainage buffer was calculated in 100 m intervals. Using the geology database, the lithology was extracted, and the distance from lineament were calculated. The lithology map was obtained from a 1:63,300-scale geological map, and the distance from lineament map was calculated in 100 m intervals. Land cover data was classified using a LANDSAT TM image employing an unsupervised classification method and field survey. The nine classes identified, such as urban, water, forest, agricultural area, tin mines, rubber and palm oil plantation were extracted for land cover mapping. Finally, the Normalized Difference Vegetation Index (NDVI) map was obtained from SPOT satellite images. The NDVI value was calculated using the formula $NDVI = (IR - R)/(IR + R)$, where IR value is the infrared portion of the electromagnetic spectrum, and R-value is the red portion of the electromagnetic spectrum. The NDVI value denotes areas of vegetation in an image.

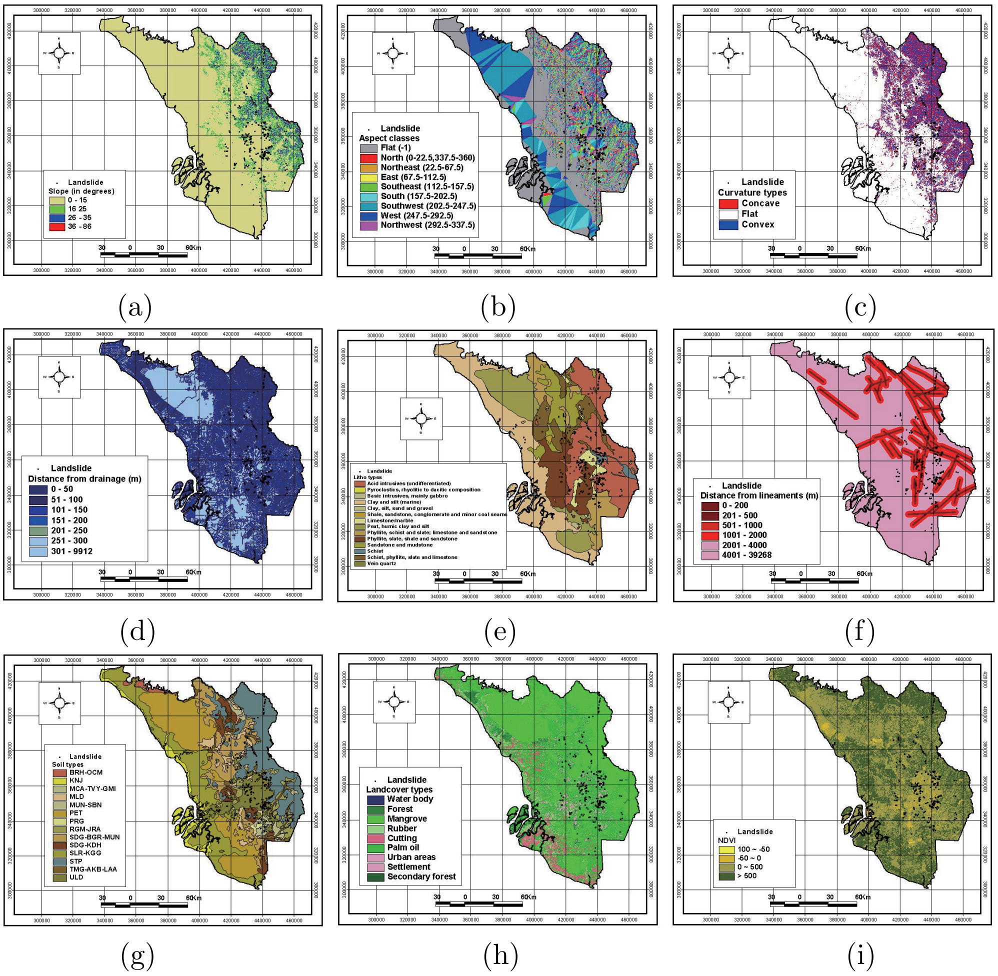


Figure 1: TInput data layers (a) Slope; (b) Aspect; (c) Curvature; (d) Distance to drainage; (e) Lithology; (f) Distance to fault; (g) Soil; (h) Land cover; and (i) Normalised difference vegetation index (ndvi)

The factors were converted to a raster grid with 10 m $\times$ 10 m pixels for application of the artificial neural network. The area grid was 14,140 rows by 12,900 columns and 327 pixels had landslide occurrences.

4. Artificial Neural Network Model

The artificial neural network approach has many advantages compared with other statistical methods. Firstly, the artificial neural network method is inde-

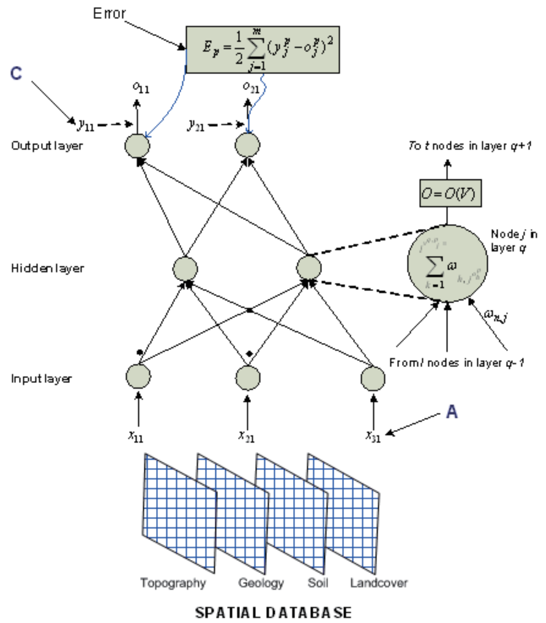


Figure 2: Three tiered architecture of feed-forward, back-propagation neural network (multilayer perception). (A): The presentation of training input data layer pattern 1with the values of and . (B) The output values of the network ( ). (C): The desired output pattern for the first samples of the training data (Modified after Moody and Katz, 2003)

pendent of the statistical distribution of the data and there is no need for specific statistical variables. Neural networks allow the target classes to be defined in relation to their distribution in the corresponding domain of each data source (Zhou, 1999), and therefore integration of remote sensing data or GIS data is convenient. An artificial neural network is a “computational mechanism able to acquire, represent, and compute a mapping from one multivariate space of information to another, given a set of data representing that mapping” (Atkinson and Tatnall, 1997). Most ANN models share a number of characteristics. These will be identified before proceeding to describe particular models (Moody and Katz, 2003). First, unlike expert systems, ANNs are not initialized with any external rule base. Rather the goal of the ANN is to internally identify a set of rules for matching input data to output conclusions. An ANN is composed of a set of nodes and a number of interconnected processing elements. ANN uses learning algorithms to model knowledge and save this knowledge in weighted connections, mimicking the function of a human brain (Turban and Aronson, 2001). One of the most commonly used ANN models is the feed-forward back-propagation ANN. This is a supervised, pattern recognition model that needs to be trained using a data set for which both the input values (x) for a set of predictors and the correct output values (y) are known for a set of examples. The architecture of this ANN is based on a structure known as the Multi-Layer Perceptron (MLP). The MLP, as the name implies, consists of a set of layers, each of which is composed of a set of nodes (alternatively referred to as "processing elements", "units", "processing units", or "neurons"). The MLP with the back-propagation algorithm is trained using a set of examples of associated input and output values. The purpose of an artificial neural network is to build a model of the data-generating process, so that the network can generalize and predict outputs from inputs that it has not previously seen. This learning MLP algorithm is trained with the "Back-Propagation algorithm", which consists of an input layer, hidden layer, and an output layer.

The first layer of the network, or input layer, contains a node for each of l input variables (Figure 2). The l input variables are analogous to the independent variables in multiple regressions. When a given set of l input values for one of the n samples in the training data set is presented to the input nodes, we say that the network is presented with an input pattern ($x^p_{il} = (x_{i1,\ldots,x_{i3}}$, where $i = 1$ to n). The superscript p indicates terms that consists of or refer to a given pattern of values (Moody and Katz, 2003).

The last layer of the network, or output layer, contains t nodes, one for each output type (Figure 2). In this case, there are nine input nodes (one each for slope, aspect, curvature, distance from drainage, distance from lineaments, lithology, landuse, vegetation index and soil texture). Sandwiched between the input and output layers is one "hidden" layer which will allow complexities to develop in the mapping functions. In this case, a three tired ANN architecture model is used. The hidden layer, like the input and output layer, consists of nodes.

The hidden and output layer neurons process their inputs by multiplying each input by a corresponding weight, summing the product, and then processing the sum using a nonlinear transfer function to produce a result. An artificial neural network "learns" by adjusting the weights between the neurons in response to the errors between the actual output values and the target output values. At the end of this training phase, the neural network provides a model that should be able to predict a target value from a given input value.

There are two stages involved in using neural networks for multi-source classification: the training stage, in which the internal weights are adjusted; and the classifying stage. Typically, the back-propagation algorithm trains the multi-layered until some targeted minimal error is achieved between the desired and actual output values of the network. Once the training is complete, the network is used as a feed-forward structure to produce a classification for the entire data

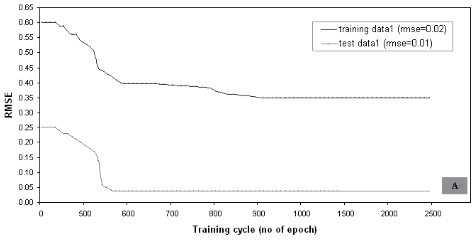
(Paola and Schwengerdt, 1995). For this study, the neural networks were simulated in the neural network module of Mathworks MATLAB (The MathWorks Inc. 1999). The back propagation multilayer perceptron (MLP) is a commonly used and widely available neural network structure in geospatial analysis and was used in this study.

5. Landslide Susceptibility Analysis Using the Artificial Neural Network Model

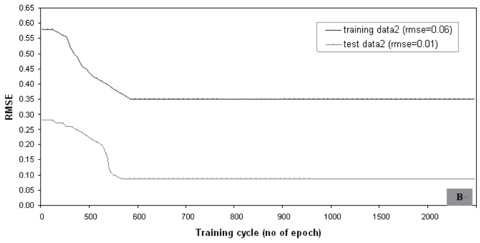
The probabilities of occurrence of landslides were calculated based on (a) the various input attributes that have been listed in table 1 and their cumulative influence (weightage values were derived from ground-based information) and (b) knowledge based classification. Before running the artificial neural network program, the training site should be selected. So, the landslide-prone (occurrence) area and the landslide-not-prone area were selected as training sites. Pixels from each of the two classes were randomly selected as training pixels, with 327 pixels denoting areas where landslide not occurred or occurred. First, areas where the landslide was not occurred were classified as "areas not prone to landslide" and areas where landslide was known to exist were assigned to an "areas prone to landslide" training set. Training sites were selected based on landslide location as prone training site and with a varying slope values as non-prone training site and then the MLP trained back propagation algorithm was computed. Five different training sites were selected randomly to produce five susceptibility maps.

The MLP trained with the Back-Propagation algorithm was then applied to the input attribute layers by modifying the number of hidden nodes and the learning rate. Hidden layers were selected two times of input attribute layers. So obviously, the output will have both "existing" and "non-existing" landslide areas. Some of the input attributes layers are continuous and others categorical in nature. Therefore, these data were converted to raster grid in order to apply the ANN model. Three-layered feed-forward network was implemented using the MATLAB software package. Here, "feed-forward" denotes that the interconnections between the layers propagate forward to the next layer. The number of hidden layers and the number of nodes in a hidden layer required for a particular classification problem are not easy to deduce. In this study, a $9 \times 19 \times 2$ structure was selected for the network, with input data normalized in the range 0.1-0.9. The nominal and interval class group data were converted to continuous values ranging between 0.1 and 0.9. Therefore, all the layers were normalized in the range 0.1- 0.9. The categorical data and their interval class group were converted to a continuous values ranging 0.1 - 0.9. In this way, the continuous values became nominal for back propagation modeling.

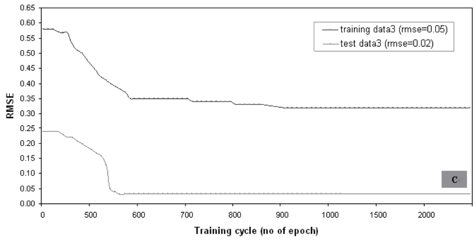
The learning rate was set to 0.01, and the initial weights were randomly



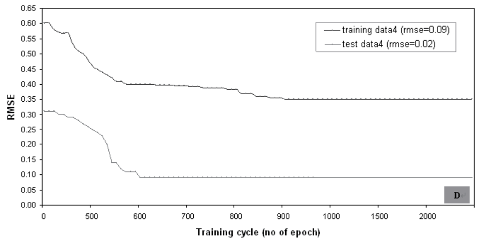
(a)



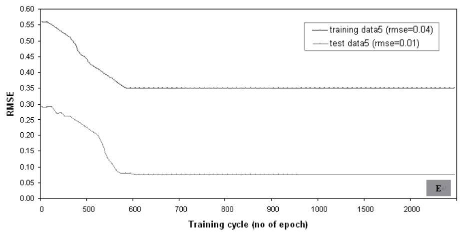
(b)



(c)



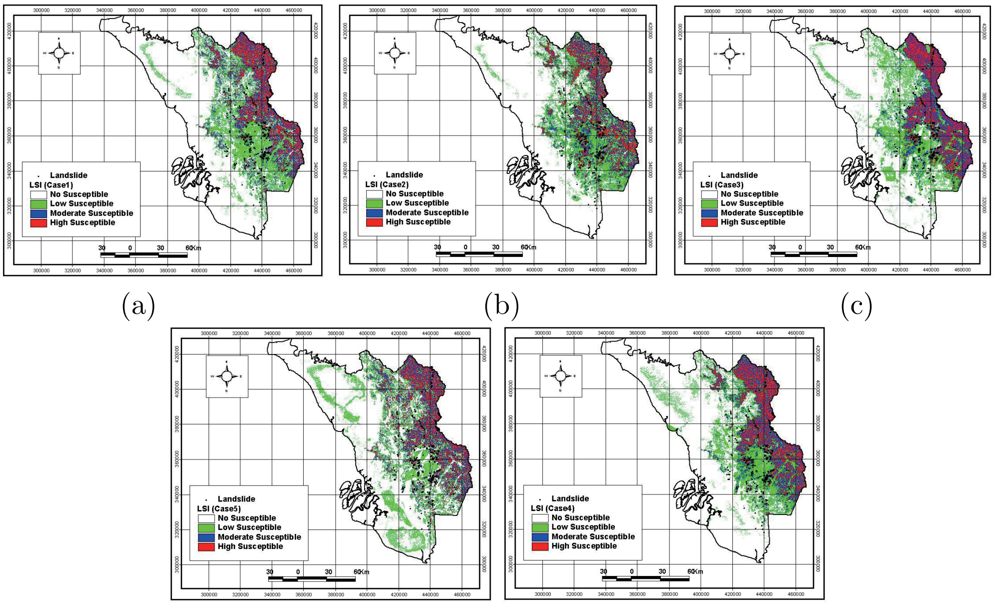
(d)



(e)

Figure 3: RMSE variations observed for the training and test datasets with respect to each five cases of training sites; (a) training site 1; (b) training site 2; (c) training site 3; (d) training site 4; and (e) training site 5

selected between 0.1 and 0.3. The MLP trained back-propagation algorithm was used to minimize the error between the predicted output values and the calculated output values. The algorithm propagated the error backwards iteratively by adjusting the weights. The number of epochs was set to 2,500, and the root mean square error (RMSE) value used for the stopping criterion was set to 0.01. Most of the training datasets met the 0.01 RMSE goal. The results of the learning rate for the training datasets are shown in Figure 3. However, if the RMSE value was not achieved, then the maximum number of iterations was terminated at 2,000 epochs. When the latter case occurred, then the maximum RMSE value was 0.213. Finally, the landslide susceptibility maps were generated using the five training sites (Figure 4). The values were classified by equal areas and grouped into four classes (highest 10%, second 10%, third 20% and reminding 60%) based on equal area classification for visual interpretation.



(a) (b) (c)

(d) (e)

Figure 4: RMSE variations observed for the training and test datasets with respect to each five cases of training sites; (a) training site 1; (b) training site 2; (c) training site 3; (d) training site 4; and (e) training site 5

6. ROC Curve Evaluations of the Landslide Susceptibility Maps

The outputs of the neural network model after their spatialization are generally presented in the form of maps expressed qualitatively or quantitatively. In this case, both physical and Receiver Operating Characteristic (ROC) model validations of these outputs have been done in the five cases of susceptibility maps. The landslide susceptibility analysis results were verified using known landslide test locations. The ROC curves were created and their areas under curve were calculated for all five cases. The ROC curve explains how well the model and attributes predict the landslide. So, the area under curve (AUC) can assess the prediction accuracy qualitatively. To obtain the relative ranks for each prediction pattern, the calculated index values of all pixels in the study area were sorted in descending order. Then the ordered pixels values were set on the y-axis with accumulated intervals on the x-axis. .The rate verification results appear as a line in Figure 5. For example, in the case of all attributes used, 90 to 100% (10%) class of the study area where the landslide susceptibility index had a higher rank could explain 35% of all the landslides. In addition, the 80 to 100% (20%) class of the study area where the landslide susceptibility index had a higher rank could explain 58% of the landslides. To compare the result quantitatively, the areas under the curve (AUC) were re-calculated based on the total area value 1 which means perfect prediction accuracy. So, the area under a curve can be used to assess the prediction accuracy qualitatively. Verification results show that in the randomly selected training site 1 (case 1), the area ratio was 0.9259 and the prediction accuracy was 92.59%. In the training site 2 (case 2), the area ratio was 0.8374 and the prediction accuracy was 83.74. In the training site 3 (case 3), the area ratio was 0.8507 and the prediction accuracy was 85.07%. In the training site 4 (case 4), the area ratio was 0.8604 and the prediction accuracy was 86.04%. In the training site 5 (case 5), the area ratio was 0.8292 and the prediction accuracy was 82.92%. So from the prediction accuracy graphs (Figure 5), it is quite evident that, training site 1 where slope equal to "zero" used for susceptibility map shows the best prediction accuracy of 92.59%, where as training site 5 shows the least prediction accuracy of 82.92% with difference is about 10%.

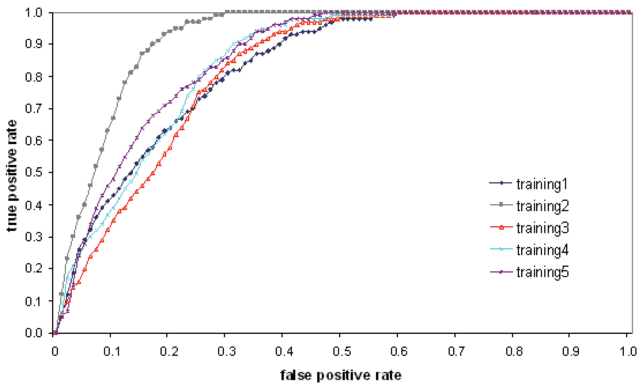


Figure 5: ROC curve evaluations for the susceptibility maps constructed using the five different training cases.

6. Discussions and conclusion

Landslides possess a significant constraint to development in Malaysia, notably through the sudden reactivation of ancient inland landslides. A series of Government funded research projects has provided much background information and identified suitable methods for the use of landslide susceptibility information in land use planning. However, a number of significant problems remain over the use of this information. In this study, an artificial neural network model was used to estimate the landslide susceptible areas using remote sensing data and GIS tools.

In this neural network model, it is difficult to follow the internal processes of the procedures, and the method entails a long execution time with a heavy computing load. There is a need to convert the database to another format, such as ASCII; the method requires that data be converted to ASCII for use in the artificial neural program and later reconverted to incorporate it into a GIS layer. Moreover, the large amount of data in the numerous layers in the study area cannot be processed in artificial neural network program quickly and easily. Using the attribute data, landslide occurrence potential can be assessed, but the landslide events cannot be predicted. However, landslide susceptibility can be analyzed qualitatively. While dealing with continuous and discrete data in an artificial neural network model is an intriguing task. To optimize the model, the artificial neural network methods have to be improved by applying in different areas with more case studies.

Using the artificial neural network, five landslide susceptibility maps were created and verified with the aid of ROC curve method. Five susceptibility maps were prepared using the five randomly selected training sites. The results shows that, training site 1 (case 1) where slope equal to "zero" used for susceptibility map gives higher prediction accuracy than the other training sites (Case 2- 5).

Landslide susceptibility maps are of great help to planners and engineers for choosing suitable locations to implement developments. These results can be used as basic data to assist slope management and land-use planning, but the models used in the study are valid for generalized planning and assessment purposes, although they may be less useful at the site-specific scale where local geological and geographic heterogeneities may prevail. In spite of a number of weaknesses in the database, the ANN modeling approach, combined with the use of remote sensing and GIS spatial data give reasonable accuracy for the landslide prediction. For the model to be more generally applied, more landslide data are needed, as well as application to more regions.

Acknowledgments

Thanks to the Alexander von Humboldt Foundation, Germany for awarding a guest researcher position and adequate fund to carry out research at Dresden University of Technology, Germany. Special thanks to the anonymous reviewer for his valuable comments which helped the manuscript to bring to the current form.

References

Atkinson, P. M. and Tatnall, A. R. L. (1997). Neural networks in remote sensing. *Int. J. Remote Sens.* **18**, 699-709.

Baeza, C. and Corominas, J. (2001). Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surf. Proc. Land* **26**, 251-1263.

Catani, F., Casagli, N., Ermini, L., Righini, G. and Menduni, G. (2005). Landslide hazard and risk mapping at catchment scale in the Arno River Basin. *Landslides* **2**, 329-343.

Dai, F. C. and Lee, C. F. (2002). Landslide characteristics and slope instability modeling using GIS, Lantau Island. Hong Kong. *Geomorphology* **42**, 213-228.

Ercanoglu, M. and Gokceoglu, C. (2002). Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkye) by fuzzy approach. *Environmental Geology* **41**, 720-730.

Ermini, L., Catani, F. and Casagli, N. (2005). Artificial neural networks applied to landslide susceptibility assessment. *Geomorphology* **66**, 327-343.

Gokceoglu, C., Sonmez, H. and Ercanoglu, M. (2000). Discontinuity controlled probabilistic slope failure risk maps of the Altindag (settlement) region in Turkey. *Engineering Geology* **55**, 277-296.

Gomez, H. T. and Kavzoglu, T. (2005). Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. *Engineering Geology* **78**, 11-27.

Lee, S., Choi, J. and Min, K. (2002). Landslide susceptibility analysis and verification using the Bayesian probability model. *Environmental Geology* **43**, 120-131.

Lee, S., Chwae, U. and Min, K. (2002). Landslide susceptibility mapping by correlation between topography and geological structure: the Janghung area, Korea. *Geomorphology* **46**, 149-162.

Lee, S. and Min, K. (2001). Statistical analysis of landslide susceptibility at Yongin, Korea. *Environmental Geology* **40**, 1095-1113.

Lee, S. and Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* **4**, 33-41.

Lee, S. and Pradhan, B. (2006). Probabilistic landslide risk mapping at Penang Island, Malaysia. *Journal of Earth System Science* **115**, 1-12.

Lee, S., Ryu, J. H., Min, K. and Won, J. S. (2003a). Landslide susceptibility analysis using GIS and artificial neural network. *Earth Surface Processes and Landforms* **27**, 1361-1376.

Lee, S., Ryu J. H., Lee, M. J. and Won, J. S. (2003b). Landslide susceptibility analysis using artificial neural network at Boun, Korea. *Environmental Geology* **44**, 820-833.

Lee, S. and Choi, U. (2003) Development of GIS-based geological hazard information system and its application for landslide analysis in Korea. *Geoscience Journal* **7**, 243-252.

Lee, S., Ryu, J. H., Won, J. S. and Park, H. J. (2004). Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. *Engineering Geology* **71**, 289-302.

Lee, S., Choi, J. and Min, K. (2004). Probabilistic landslide hazard mapping using GIS and remote sensing data at Boun, Korea. *International Journal of Remote Sensing* **25**, 2037-2052.

Moody, A. and Katz, D. B. (2003). Artificial intelligence in the study of mountain landscapes. In *Geographic information science and mountain geomorphology* (Edited by Bishop, M. P.; and Shorder, J. F., 219-249). Springer.

Neaupane, K. M. and Achet, S. H. (2004). Use of backpropagation neural network for landslide monitoring: a case study in the higher Himalaya. *Engineering Geology* **74**, 213-226.

Paola, J. D. and Schowengerdt, R. A. (1995). A review and analysis of backpropagation neural networks for classification of remotely sensed multi-spectral imagery. *International Journal of Remote Sensing* **16**, 3033-3058.

Pistocchi, A., Luzi, L. and Napolitano, P. (2002). The use of predictive modeling techniques for optimal exploitation of spatial databases: a case study in landslide hazard mapping with expert system-like methods. *Environmental Geology* **41**, 765-775.

Pradhan, B. and Lee, S. (2007). Utilization of optical remote sensing data and GIS tools for regional landslide hazard analysis by using an artificial neural network model. *Earth Science Frontier* **14**, 143-152.

Pradhan, B. and Lee, S. (2009a). Landslide risk analysis using artificial neural network model focusing on different training sites. *International Journal of Physical Sciences* **3**, 1-15.

Pradhan, B. and Lee, S. (2010a). Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences* **60**, 1037-1054.

Pradhan, B., Lee, S. and Buchroithner, M. F. (2009). Use of geospatial data for the development of fuzzy algebraic operators to landslide hazard mapping: a case study in Malaysia. *Applied Geomatics* **1**, 3-15.

Pradhan, B., Lee. S., Mansor, S., Buchroithner, M. F. and Jallaluddin, N. (2008). Utilization of optical remote sensing data and geographic information system tools for regional landslide hazard analysis by using binomial logistic regression model. *Journal of Applied Remote Sensing* **2**, 1-11.

Pradhan, B., Mansor, S., Lee, S. and Buchroithner, M. F. (2008b) Application of data mining model for landslide hazard mapping. *Proceedings of ISPRS*, Vol. XXXVII, Part B8, Commission VIII, 187-196.

Pradhan, B., Singh, R. P. and Buchroithner, M. F. (2006). Estimation of stress and its use in evaluation of landslide prone regions using remote sensing data. *Advances in Space Research* **37**, 698-709.

Pradhan, B. and Youssef, A. M. (2009). Manifestation of remote sensing data and GIS for landslide hazard analysis using spatial-based statistical models. *Arabian Journal of Geosciences*, online first, doi http://dx.doi.org/10.1007/s12517-009-0089-2

Pradhan, B. (2010a). Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia. *Advances in Space Research* **45**, 1244-1256.

Pradhan, B. (2010b). Use of GIS-based fuzzy logic relations and its cross application to produce landslide susceptibilitymaps in three test areas in Malaysia. *Environmental Earth Sciences* doi:10.1007/s12665-010-0705-1.

Pradhan, B. (2010c). Manifestation of an advanced fuzzy logic model coupled with Geo-information techniques to landslide susceptibility mapping and their comparison with logistic regression modelling. *Environmental and Ecological Statistics* doi:10.1007/s10651-010-0147-7.

Pradhan, B. (2010d). Application of an advanced fuzzy logic model for landslide susceptibility analysis. *International Journal of Computational Intelligence Systems* **3**, 370- 381.

Pradhan, B. and Buchroithner, M. (2010). Comparison and validation of landslide susceptibility maps using an artificial neural network model for three test areas in Malaysia. *Environmental and Engineering Geoscience* **16**, 107-126.

Pradhan, B. and Lee, S. (2010a). Delineation of landslide hazard areas using frequency ratio, l ogistic regression and artificial neural network model at Penang Island, Malaysia. *Environmental Earth Sciences* **60**, 1037-1054.

Pradhan, B. and Lee, S. (2010b). Regional landslide susceptibility analysis using backpropagation neural network model at Cameron Highland, Malaysia. *Landslides* **7**, 13-30.

Pradhan, B. and Lee, S. (2010c). Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modeling. *Environmental Modelling and Software* **25**, 747-759.

Pradhan, B. and Pirasteh, S. (2010). Comparison between prediction capabilities of neural network and fuzzy logic techniques for landslide susceptibility mapping. *Disaster Advances* **3**, 26-34.

Pradhan, B. and Youssef, A. M. (2010). Manifestation of remote sensing data and GIS on landslide hazard analysis using spatial-based statistical models. *Arabian Journal of Geosciences* **3**, 319-326.

Pradhan, B. Lee, S. and Buchroithner, M. F. (2010a). Remote sensing and GIS-based landslide susceptibility analysis and its crossvalidation in three test areas using a frequency ratio model. *Photogrammetrie, Fernerkundung, Geoinformation* **1**,17-32.

Pradhan, B. Lee, S. and Buchroithner, M. F. (2010b). A GIS-based backpropagation neural network model and its cross application and validation for landslide susceptibility analyses. *Computers Environment and Urban systems* **34**, 216-235.

Pradhan, B. Oh, J.J. and Buchroithner, M. F. (2010c). Weight-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. *Geomatics Natural Hazards and Risk* **1**, 199-223. doi:10.1080/19475705.2010.498151.

Pradhan, B. Sezer, E. Gokceoglu, C. and Buchroithner, M.F. (2010). Landslide susceptibility mapping by neuro-fuzzy approach in a landslide prone area (Cameron Highland, Malaysia). *IEEE Transactions on Geosciences and Remote Sensing* **48**, 4164-4177. doi:10.1109/TGRS.2010.2050328.

Turban, E. and Aronson, J. E. (2001). *Decision Support Systems and Intelligent Systems*. Prentice Hall.

Zhou, W. (1999). Verification of the nonparametric characteristics of backpropagation neural networks for image classification. *IEEE Transactions on Geoscience and Remote Sensing* **37**, 771-779.

Received June 12, 2009; accepted January 18, 2010.

Biswajeet Pradhan
Spatial and Numerical Modelling Laboratory
Institute of Advanced Technology
University Putra Malaysia
43400, UPM, Malaysia
biswajeet@mailcity.com / biswajeet24@gmail.com