Landslide Susceptibility Hazard Mapping Techniques Review

Mutasem Sh. Alkhasawneh, Umi Kalthum Bt Ngah, Tay Lea Tien and Nor Ashidi Bin Mat Isa
School of Electrical and Electronics Engineering, Engineering Campus,
University Sains Malaysia, 14300Nibong Tebal, Malaysia

Abstract: Landslides continue to be one of the worst nature disasters around the world. Worldwide, landslide causes billion of dollars in damage and claim as many as thousands of lives each year. To date, there is no specific methodology for assessment and prediction pattern of landslide to occur because the nature of earth is not the same and the factors triggering the land slide is not consistent. However, qualitative and quantitative methods have been used to detect and predict landslide. The goal of this review paper is to discuss the quantitative methods which involve neural network and fuzzy logic approach that have been proposed, designed and developed by previous researchers in order to overcome the drawbacks of landslide mapping. These techniques may assist geologist, construction companies and the emergency department in detecting the landslide and alerting awareness.

Key words: Landslide, landslide hazard map, neural network, artificial intelligent, fuzzy logic approach

INTRODUCTION

The term of landslide has many definitions, based on Cruden (1991) “any movement of debris or earth down a slope or a mass of rock consider a landslide”. Varnes (1984) considered a probability of movement of the earth downward or outwards under the effect of the gravity, rain and slope as a landslide. For basic landslide hazard description, some information should be included like the landslide location, velocity of landslide, resultant detached material, the area volume and the occurrence possibility in a specified period of time. On the other hand, landslide inventory is an inventory of activity containing date of occurrence of land sliding, location, classification and volume. All these definitions are based on the International Union of Geological Science (IUGS, 1997). Neural networks has many applications, like cancerous diagnostic system (Mat-Isa et al., 2008) GIS image compression and restoration (Al-Bastaki, 2006), object tracking (Bouzemlal et al., 2007), handwritten check words recognition (Noori et al., 2011), evaluating prices of housing (Erkki and Udegunah, 2008), water quality analysis (He et al., 2006), predicting gastric cancer (Amiri et al., 2008) and temperature prediction (Sharma and Agarwal, 2012), aggregate classification (Al-Batah et al., 2009), pattern classification (Hee and Mat-Isa, 2011).

Landslide triggering factors: Landslides are triggered by many causative factors which may be divided into six factors geomorphology, geological, meteorological soil, land use, land cover and hydrologic conditions. Geomorphology, geological and meteorological factors soil includes slope aspect, gradient, relative relief, lithology, degree of weathering, depth, permeability and porosity. Varnes (1984) and Hutchinsson (1995) have given more details about the factors influence the land sliding.

Landslide history of damage: Thousands of landslides occur annually. Based on estimates from the Red Crescent Societies and Red Cross, landslide kill 1500 people average every year (Natural Disaster, 2006). In the summer of 1998 multiple major landslides followed a heavy rain which hit Bangladesh and China. More than 1100 were killed in the former and around 4000 died in the latter. On October 30th, 1998 a major landslide around the volcano of Casitas buried around 2,200 people and caused millions of dollars in property losses. More than 1500 were killed in March of 1998 in Pakistan after landslide and flooding hit the southwestern part of the country. In November 2001, a major landslide left Bab El-Oued, Algeria with more than 1000 people either dead or missing and a quarter of the country sinking on the mud and ditches. In late February 2005 landslide occurred in Bandung Indonesia which killed more than 140 people. On August 10th, 2010 China suffered the worst landslide in decades which killed 702 people and left thousands of people missing.

Related work: To predict the landslide occurrence different methods have been applied and developed. This methods divided to quantitative and qualitative methods.

Corresponding Author: Mutasem Sh. Alkhasawneh, School of Electrical and Electronics Engineering, Engineering Campus, University Sains Malaysia, 14300Nibong Tebal, Malaysia Tel: +60 (0) 10 436 6063
Table 1: Comparison among the techniques employed in landslide mapping system

<table>
<thead>
<tr>
<th>Research aim</th>
<th>Study area location</th>
<th>Study area (km²)</th>
<th>No. of factors</th>
<th>No. of slides</th>
<th>Technique</th>
<th>Accuracy</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>landslide map and factors weight</td>
<td>Korea</td>
<td>N/A</td>
<td>7</td>
<td>280</td>
<td>Back propagation</td>
<td>Satisfactory results</td>
<td>Lee et al. (2001)</td>
</tr>
<tr>
<td>Classify the susceptibility</td>
<td>Turkey</td>
<td>275.4</td>
<td>7</td>
<td>266</td>
<td>Fuzzy logic</td>
<td>Satisfactory results</td>
<td>Ercanoglu and Gokceoglu (2004)</td>
</tr>
<tr>
<td>landslide hazard map</td>
<td>Riomaggiore, Italy</td>
<td>47</td>
<td>17</td>
<td>5</td>
<td>MLP</td>
<td>73%</td>
<td>Emmini et al. (2005)</td>
</tr>
<tr>
<td>Factors weight</td>
<td>Potenza, Italy</td>
<td>47</td>
<td>7</td>
<td>920</td>
<td>Back propagation algorithm</td>
<td>69%</td>
<td>Canini et al. (2007)</td>
</tr>
<tr>
<td>Comparison of FR, NN and LR</td>
<td>Turkey</td>
<td>8.13</td>
<td>5</td>
<td>600</td>
<td>FR, ANN and LR</td>
<td>NN is the best</td>
<td>Yilmaz (2009)</td>
</tr>
<tr>
<td>Find the weight factors</td>
<td>Cameron highland, Malaysia</td>
<td>12</td>
<td>10</td>
<td>324</td>
<td>LR</td>
<td>83.3%</td>
<td>Pradhan et al. (2008)</td>
</tr>
<tr>
<td>landslide map</td>
<td>Penang, Malaysia</td>
<td>8.06</td>
<td>7</td>
<td>48</td>
<td>Neuron fuzzy</td>
<td>84.30%</td>
<td>Oh and Pradhan (2011)</td>
</tr>
</tbody>
</table>

These two methods vary with respect to methodology have been used in drawing the landslide mapping. The direct field mapping, geomorphological analysis and any methods based on the human judgments are some examples of the qualitative methods while deterministic analyses, artificial intelligence, probabilistic approaches and statistical methods represent the quantitative methods and base on mathematical models. As a result, much less human judgment and experience is needed to produce the models. Table 1 illustrates the comparison among the techniques that have been employed in landslide mapping system. Moreover, the general agreement about the ideal method for producing landslide susceptibility map has not been reached yet (Guzzetti et al., 2009). Initiation of landslide mapping begins in the 1970’s (Fell et al., 2008). Then, particularly in 1980’s, in line with the achievements in computer technology and GIS (Geographic Information Systems), there was a boom related to the landslide mapping in the scientific literature. At the beginning of 1990’s, GIS application for landslide mapping started with simple applications for few cases. In some cases GIS package demonstrated the ability to achieve major analysis on landslide mapping whereas the usage of GIS was partial in other cases. However, in the 1990’s, utilization of GIS provided indispensable tool for mapping and evaluating landslides, particularly for regional or medium scale studies. Giving a general idea about the works which have been done before, in the landslide susceptibility mapping by using the quantitative methods is the aim of this paper.

Lee et al. (2001) have proposed a study to draw the landslide hazard mapping by using neural network and the new technique was applied in the Yongin in Korea as a case study area. Seven landslide causing factors were collected and extracted from a special data base. These factors include the curvature, slope, soil effective thickness and texture, drainage and timber age and diameters. The back propagation algorithm was used twice in this study, first to create the landslide map, second to determine the weights of each factor in the landslide map. The verification results between the susceptibility index and existing landslide location data shows a good agreements and satisfactory output results. The calcification map is divided into five classes of risk i.e. very low, low, medium, high and very high. These maps were made with regards to the landslide map location. To obtain susceptibility based on the ratio values in order to determine each of the factor’s weight, a neural network is used to implement three-layer feed-forward. Topographic slope had the highest value while the lowest is topographic curvature (Ercanoglu and Gokceoglu, 2004). Carried out fuzzy logic as a new methodology to create the landslide map to the West Black Sea Region in Turkey. Details such landslide inventory, air photograph and survey field, for 275.4 km² the study area was collected. The study area were contained two hundred and 66 landslides topographical parameters such as slope shape, slope angles, slope aspect, topographic elevation and distances to network have been considered. Geological parameters such as closeness to structural elements, relationship between the discontinuities and the slopes and environmental parameter have been used in the landslide inventory. Ercanoglu and Gokceoglu (2004) have used computer program to utilize the fuzzy relations to produce the landslide susceptibility map automatically. The map of the case study area location was classified to very high; high, moderate, low and very low or no susceptibility area and the results came up with 9.6, 10.3, 8.9, 27.5 and 43.8%, respectively. From the previous results, the fuzzy logic showed good performance in producing the landslide susceptibility map (Ercanoglu and Gokceoglu, 2004). In addition, the approach was considered as a useful tool because its result was obtained automatically from the actual data of landslide. Probabilistic Neural Network (PNN) and Multi Layered Perception (MLP) were used by (Emmini et al., 2005) to predict landslide hazard map. Five factors were considered in this study. Namely lithology, profile curvature, slope angle, land cover and upslope. The size of the case study area is 17 km², located in Riomaggiore Italy which is considered as an ideal space for performing tests on landslide hazard analysis. The five
factors were used in the analysis are the classic controlling variable which control the landslide hazard. All the input factors have been converted to the binary variable string consist of 19 positions. These factors were used as input to the ANN. Satisfactory results with preference of MLP was shown by comparison with the recent landslide inventory of the study area and ANN has the ability to predict the hazard mapping with satisfactory results.

The back propagation learning algorithm was used in the landslide susceptibility mapping by Caniani et al. (2007) with Potenza, Italy chosen as the study area. Nine hundred and twenty landslides were recognized, represent the earth flow, rotational slide and rotational slides, spreading over 46 km² which represents around 26% of the entire area of Potenza. Three layers of neural network input layer, hidden layer and output layer were connected to each other, respectively. The famous back propagation algorithm with the three layers input, hidden and output layer was used as learning algorithm. Landslide triggering factors like geomorphology, geological, meterorological and hydrologic conditions which includes lithology, slope aspect and angle, elevation, topographic index and topographical shape and land use. All the morphometric parameters were derived from the digital elevation model (DEM) of the area of Potenza with a resolution of 20 m. The work was divided into two phases i.e., the training phase and the validation phase. In this study, 32% of the landslide site was selected for training phase and the rest of the landslide site was used for the validation phase. The weights of each factor on the seven factors were calculated. Slope aspect, slope angle, slope gradient, lithology and elevation had the highest weight. The verification step found those factors were the most effective factors which could lead to landslide susceptibility. ANN showed good performance by classifying 80% of the landslide pixels correctly.

Pradhan et al. (2010) partially applied the 10 factors mining for landslide hazard mapping by using ANN to calculate the weight factors. Field survey and aerial photographs were used to identify the landslide location of the part Cameron highland in Malaysia. Three hundred and Twenty four landslide were found in the case study area, the database of study area was divided in to three part to assemble access to the map of the database again multi-layer neural network with one input one output and one hidden layer was used, the weight of each factor between the layers was calculated by applying the back propagation algorithm. Sigmoid function was applied as a transfer function to some of the input weight between the layers. Back propagation was the training algorithm. The aim of the study was to calculate the weight of each factor, once the weight of each factor calculated, it can be used in the classification step by passing the new data which never been used in the neural network before. The rate curve were created by finding the error value between the actual output value and the neural network output value. The area under curve detect to the importance of the factor, buy using the MATLAB software to implement the fed forward neural network the relative importance of factors between weights showed the slope factor has the highest value among the all factors which is 2.05, then the distance from drainage is 1.4 and followed by geology whose value is 1.1 the network accuracy prediction for the landslide still in the average accuracy which 83.45.

Solaizani et al. (2009) have used Anbalagan method to cover 62.07 km² the space of the study area which is located in South-Western part of Sari, Iran. The hazard area was divided to three regions medium, high and very high hazard. Eight factors were used in this study namely the slope, slope aspect, relative relief, lithology, fault, land use and land cover, water ground and soil. The importance of each factor was calculated individually. Result showed that lithology and soil are more importance in landslide occurrence, then, roughness, land use, slope, ground water condition and structure are effective, respectively. The Anbalagan method has showed a good result and performance for the case study comparing with other methods.

Marjanovic et al. (2009) focused on using support vector machine (SVM), Neighbor (k-NN) algorithms and Analytical Hierarchy Process (AHP) for weighting influences of different input parameters. Seven factors have been used to extract the landslide map. These factors include elevation, slope angle, aspect and distance from flows, vegetation cover, lithology and rainfall to represent the natural factors of the slope stability. The study area was the North West slopes of the Fruška Gora Mountain, in the vicinity of Novi Sad, North West Serbia which represent 40 km² of hilly landscape. The research was divided into two parts i.e., the expert's opinion in multi-criteria analysis and the machine learning feature SVM and K-NN algorithm. Multi-criteria analysis is a widespread tool for various types of assessments, especially for spatial implications. It implements a procedure where several inputs fuse a single outcome of the modeled phenomenon. However, these input geo-parameters have got different importance for the phenomenon, requiring to be leveled up in some fashion, which brings us to the Analytical Hierarchy Process (AHP). AHP is a decision making tool, pioneered in 1980's by Saaty (1980), being broadly applied ever since. In context of the research i.e., the research predating this one (Marjanovic, 2009), a standard first-level AHP was
performed over input raster sets. Inputs represent geological, morphometrical and environmental parameters fairly significant for the problem and yet appropriate for raster modeling in coarser scale (Komac, 2006; Esmali and Ahmadi, 2003; Van Westen et al., 2006; Vozniček, 2000). However, those initial input sets were refashioned and adapted to meet the purpose in chosen machine learning approaches. In addition, the final outcome of multi-criteria stage (model of landslide susceptibility) together with the spatial coordinates of ground elements (corresponding pixels, equally sized and geo-referenced in all raster sets) had been appended to complete the input set for the training mode of the algorithm. The K-NN is a simple algorithm usually used for n-dimensional input space. However, this requires sorting and pondering distances per each element, resulting in hardware-demanding and time-consuming procedures. This method turns extremely inefficient with larger number of parameters and bigger percentage of training data and the 57% of accuracy in the average, while SVM method deals with the binary classification model and reached the highest accuracy (88%).

Oh and Pradhan (2011) have used the new technique Adaptive Neuron-Fuzzy Inference System (ANFIS) for the first time in the landslide susceptibility mapping. 8.06 km² of Penang Island was taken as a case study. Forty eight landslides were compiled from various data sources. Data were collected and extracted from the aerial photographs and extensive field surveys, in addition to different recourse like the news records, historical landslide report and any archive data. Seven measure factors were considered in this work namely the Stream Power Index (SPI), slope angle, plane curvature, soil texture, altitude, distance to the drainage and road. Preparing the data for the training and testing is very important step. Many methods once can use to choose the data for training and testing. Oh and Pradhan (2011) have considered 80% of the data is enough to train the neural network and the rest of the data for training. Randomly cells have been selected for training from the landslide occurrence area and the non occurrence area in this study 38 landslides out 48 were used for training and 10 landslides were used for verification. Frequency ratio technique was used to find the relation between the landslide and each one of the seven factors. Two assumptions were considered in the verification stage, first assumption that the landslide has a link with spatial information like topographic and soil type, the second assumption that the rain or an earthquake can trigger the landslide in the future (Chung and Fabbri, 2003). Receiver Operating Characteristics (ROC) was used to check the effectiveness of the susceptibility map. A higher accuracy 84.39% of susceptibility mapping was achieved in this study.

Tian et al. (2010) has proposed mixed strategy with evolutionary algorithm and neural network. Using an objective optimization algorithm based on ANN and multi-objective evolution method, to get the essential affecting factors and their weight. Acceptable result was achieved. Miyi county located in southwest of Sichuan, province, China was selected to be the study area. Geomorphology, topography, geology, environmental landslide occurrence and other factors was used in this research. The research methodology in this work can be presented in five steps: (a) initial and coding of susceptibility factors, (b) fitness assignment, (c) training historical data by neural network, (d) sorting by selection strategy (makes evolution toward directional search for Pareto set, keep diversity of the elitist) and (e) calculating individual’s crowd distance, using mix strategy of tournament selection operator and crowd Compare operator. Eighty percent of the data was used for training and the rest of the data for testing. After the calculation by the ANN, The five factors showed the highest weight value were used essential factors to susceptibility zoning, while the factors with smallest weight were neglected.

Yilmaz (2009) have done comparison between Frequency Ratio (FR), Logistic Regression (LR) and Artificial Neural Network (ANN). A small study area was chosen in the Republic of Korea. With a size of 8.13 km², the landslide occurred in 300 locations. Five landslide factors were used they are the topographical factor, hydrological factor, soil factor, forest factor, land cover factor. Area Under Curve (AUC) analysis was built with each model to assets the performance of FR, ANN and LR. Analysis result showed that there was a high correlation between the maps using LR and ANN methods exhibiting the highest correlation coefficient (0.829). The lowest coefficient (0.619) was found between LR and FR methods. Each model has some advantages and disadvantages. FR can be simply applied, whereas an LR method needs data conversion to be read by the statistical software program. And LR method has a limitation in calculation on the program when the data is massive. ANN method is time consuming but attains high accuracy.

Ermini et al. (2005) proposed a comparison of neural networks and logistic regression methods in a medium scale. In this study neural network has proved again that it is more realistic than any other techniques for landslide mapping and landslide susceptibility hazard mapping. The goal of this study was to produce a landslide mapping for natural gas pipeline in the study area which covers 290 km² representing the area around the gas pipeline, which located in Marmara and Black Sea regions of Turkey. Collecting and preparing the data is one of the major steps in landslide susceptibility mapping. In this Study the landslide inventory map was prepared based on
the previous inventory map and extensive field work. Logistic Regression (LR) and Neural Networks (NN) (feed-forward and back-propagation) methods were used to analyze some probable and slide causing factors like slope and slope length, topographic wetness watershed basins index, surface area ratio, curvature plane and profile, distance from road drainage and fault line, elevation, density of drainage and fault land cover and use and the stream power index. In this study, two landslide susceptibility maps were produced by LR and NN. Validation data set and the field check were used to evaluate the two maps. On 1:25000 scale map 112 landslides were found, 33 bodies were extracted and mapped. The preprocessing in this study, were done by putting all the independent variable on witheld and in some subsequent step when those variable determined to be significant will be added to the system and the other will be withheld. In this study, the training step were done by dividing the data randomly to six sets and the chi-square of Hosmer and Leme were used to evaluate the training set. SPSS software package were used to calculate the average accuracy of each training set. Software package observed that there a little bit difference was noticed between the logistic regression training set and neural network training set. The difference came from the training set sample itself. Hence, the random training samples were taking from the whole dataset which is common in the all training procedure and doesn’t make that much difference. This study did not include the slope and geology factor to the model. The data set nature played a highly rule in the accuracy of this comparison. Three of the accuracy indicators were used in this study Percentage of Total Area (PTA) percentage of relevant susceptibility level in whole area, Percentage of Landslide Body (PLB) percentage of relevant susceptibility level in landslide bodies and percentage of relevant susceptibility (PSC) level in seed cells. The ratio of PTA/PSC and relative operating characteristics (ROC) curves are calculated. Now the ration between PTA/PSC and ROC relative operating characteristics curves were calculated, the value of PTA, PSC and the ratio of PTA/PSC should be below the value of PLB and PSC. The result of theses indicator showed once again that the neural fed forward network with back propagation algorithm perform better than the logistic regression model (Swets, 1988). ROC curve is a well known tool of reflecting the accuracy of probabilistic detection and the forecast systems. The ROC estimation for logistic regression technique was lower than fed forward back propagation neural network learning algorithm. The estimation of ROC was increasing slowly comparing with neural back propagation learning algorithm, especially in the low and very low risk zones and in the high and very high risk zone. The logistic regression method showed a lower percentage of landslide prediction comparing with neural network. Therefore using neural network for prediction landslide map is more realistic.

Pradhan et al. (2008) have presented the use of remote sensing data for landslide hazard analysis around Penang Island area, Malaysia. Linear logistic regression model was used for landslide hazard delineation by using remote sensing and GIS data. Aerial photographs, satellite images with high resolution and field surveys were used to identify the location of landslide. Satellite imageries were used to extract the terrain information such as land cover, topographic slope, aspect and curvature. The other factors could lead to landslide occurrence were chosen like: lithology, soil, distance from both lineament and the drainage, the vegetation index value which is extracted from SPOT 5 image were considered too. The study area this time was Penang island Malaysia with 285 km² covers the area North West coast of Malaysia with big channel separating the Penang island from the main land. A total of 463 landslides were detected in the area of case study. Ten factors like the slope, aspect and curvature were considered in calculating the probability. The logistic regression model with GIS and remote sensing were estimate to predict the landslide hazards in Penang Island. The result of verification of logistic regression model and GIS showed 86.62% prediction accuracy in hazard map.

Furthermore, some researcher have study the factors affect the landslide individually. Bibalani et al. (2007a) studied the link between the vegetation cover factor and soil stability, in northwest of Iran.

Gasim et al. (2010) have came up with study to determine geomorphology and geological features of the Bukit Bukit, Malaysia. The geological and geomorphology factors consider among the important factors cause the landslide hazard. For more details about how those factors affect the landslide review the studies by Bibalani et al. (2007b) and Devkota et al. (2006). Moreover, different methodologies and for different places have been used to predict landslide hazard map like Probabilistic method and logistic regression (Lim et al., 2011). Penang Island, Malaysia has been chosen as a study area landslide hazard maps were produced Probabilistic methods such as frequency ratio, statistical index, certainty factor and landslide susceptibility analysis, logistic regression with 80.05% accuracy represent the best result followed by frequency ratio with 79.68%, landslide susceptibility analysis with 79.6%, statistical index with 79.38% and certainty factor with 79.37%.

**CONCLUSION**

This study gives a general review to the systems and technique that have been proposed and developed for
landslide susceptibility mapping. This review paper is the first one to mention and gather the neural network and fuzzy logic technique which has been used in landslide since 2000. Basically, landslide mapping systems consist of four major steps like, collecting the data, preprocessing, processing and output. All these steps are mentioned in this paper and the methodology and the technique were explained. To date there is no specific method or technique to find the landslide map. Each place has a different map and different considering factors and technique. Neural network and fuzzy logic have shown a good result and they still have huge potential to be expanded and improved in the future.

REFERENCES


