www.cafetinnova.org

Indexed in
Scopus Compendex and Geobase Elsevier, Chemical Abstract Services-USA, Geo-Ref Information Services-USA, List B of Scientific Journals, Poland, Directory of Research Journals

ISSN 0974-5904, Volume 08, No. 02

International Journal of Earth Sciences and Engineering

April 2015, P.P.18-26

Modeling of Landslide Susceptibility using based Back-Propagation Neural Network and GIS in Kothagiri Region, India

KARTIC KUMAR M[1], ANNADURAI R[1] AND ARUMUGAM K[2]
[1]Department of Civil Engineering, S.R.M University, Kattankulathur, Chennai - 603203 Tamil Nadu, India
[2]Department of Civil Engineering, Kongu Engineering College, Perundurai-638 052, Tamil Nadu, India
***Email:** karthickumar.msm@gmail.com, annadurai.r@ktr.srmuniv.ac.in, geosixface@gmail.com*

Abstract: Landslides are encountered during the monsoon season in the undulated topography of the Western Ghats, resulting in major damage to wealth and human life. Landslide-susceptibility mapping is one of the most critical issues in the Western Ghats. The main intention of the paper is modeling and investigation of landslide susceptibility using Geographical Information System (GIS) - based back-propagation Artificial Neural Networks (ANN). A landslide associated database is constructed from various sources such as: slope, aspect, curvature, geology, precipitation, soil texture, distance from road, distance from lineament, distance from drainage, land use and the vegetation index value from IRS ID satellite images are extracted. For the computation of the relative weight above the factor to a particular landslide occurrence, an artificial neural network (ANN) method is applied. The entire factor's weight is determined by the back-propagation training method. Different training sites are randomly selected to train the neural network, and landslide-susceptibility indexes (LSI) are calculated. Finally, landslide susceptibility map is generated using ArcGIS software. Landslide susceptibility maps are formed from neural networks models and results are compared by existing landslide locations. Among the 84 landslides are mapped in GIS, 70% locations are chosen for modeling purpose and the remaining are used for validation. The results show a high concordance between the landslide inventory and the high susceptibility estimated zone. To achieve the accuracy, landslide susceptibility maps are cross-validated using respective Receiver Operating Characteristic (ROC) curves. The investigation results explain that reasonable acceptance between the landslide susceptibility map and the existing landslide location. The objective of the susceptibility modeling is to reduce the major impact of landslides by determining the areas at vulnerability.

Keywords: *Landslide Susceptibility Modeling - Kothagiri - Back-propagation - Artificial Neural Network.*

1. Introduction

A landslide phenomenon is one of the highest risk factors for people, environment and economic activities. Urbanization and expansion of construction activities in hilly regions has frequency of landslides in recently. Expansion of urban and man-made structures into potentially hazardous areas leads to extensive damage to infrastructure and occasionally results in loss of life every year [1]. Landslide presents a significant constraint to development in many parts of the study area which experience frequent landslides. Steep slope, presence of clay layer in rocks, heavy rain and improper land use practices play a major role in the genesis of landslides in the Kothagiri region [2]. Landslide susceptibility mapping is a valuable tool for assessing current and potential risks that can be used for developing early warning systems, mitigation plans and land use restrictions.

Recently few attempts have been made to predict the landslides and prevent the damages. By prediction, landslide damages could be seriously reduced to certain degree. Recently, statistical methods have applied such as logistic regression analysis and frequency ratio have been applied for landslide susceptibility mapping [3,4]. Other traditional methods such as geotechnical and factor of safety are good tool to assess the landslide hazard analysis and have the ability to develop model [5]. As a new approach to landslide hazard evaluation using GIS, soft computing techniques such as artificial neural networks [6,7], fuzzy approaches [8,9] and fuzzy logic analytical hierarchical process analysis have been used for the landslide susceptibility evaluation [10]. GIS provide functions like geo-statistical analysis and data base processing. In addition, the extension of the analysis to include environmental impact assessment of a slope failure can be easily and effectively performed using GIS and the statistical tools for modeling the natural calamities [11]. The study of landslide hazard also applied these basic tools frequently with intensively use of Digital Elevation Model (DEM) and satellite images.

Different techniques have been used to predict the landslide susceptibility, including heuristic, probabilistic and statistical approaches recently. The method put forward in the research focuses on soft computing techniques derived from Artificial Intelligence viz., Back Propagation-Neural Network model (BP-NN). The landslide inventory database has been identified through the published reports and field surveys in combination with the Global Positioning System (GPS) and GIS. The susceptibility index map is developed by considering the causative factors such as slope, aspect, curvature, elevation, geology, soil texture, distance from drainage, rainfall, Stream Power Index (SPI), distance from lineament, proximity to the road and precipitation are generated through GIS. Landuse / landcover and Normalized Difference Vegetation Index (NDVI) extracted from IRS ID LISS IV satellite image for the investigation. The extracted factors are converted to a 10 × 10 m grid (ArcGIS software package). ANN is applied to find the weightage of individual factor and the weights are applied to all the causative factors for generating landslide susceptibility mapping. The map is verified and validated using the known landslide locations. The investigation presents the mapping of landslide susceptibility for the area of Kothagiri obtained by using applications of the computing technique viz., Artificial Neural Network (ANN).

2. Study Area

Kothagiri region is located on an eastern slope of Western Ghats. It has been accepted that Kothagiri is frequently subjected to landslides. The study area is located in the eastern part of the Nilgiri district, Tamil Nadu. The area is a mountainous character. The study area covers 396.65 km^2 and lies between latitudes 11° 10' 00" to 11° 42' 00" North and longitudes 76° 14' 00" to 76° 02' 00" East (Fig. 1) with average altitude of 1400 m above the mean sea level. The study area receives rainfall both in southwest and northeast monsoons with a mean of 1920 mm. The climate of the study area is temperate. The pronounced wet seasons are during the north-east monsoon (October and November). The northeast monsoon is moderate, contributing about 40 percent. The major landslides in the region are either transitional or rotational debris flow and rock fall [12]. The general physiographic trend is NW to SE following the main drainage system. The geology of the study area consists mainly of Charnockite rocks. A few portions in the western side are covered by lateritic soils. Forest types have been described as evergreen and deciduous forest, barren shrubs, and forest plantation.

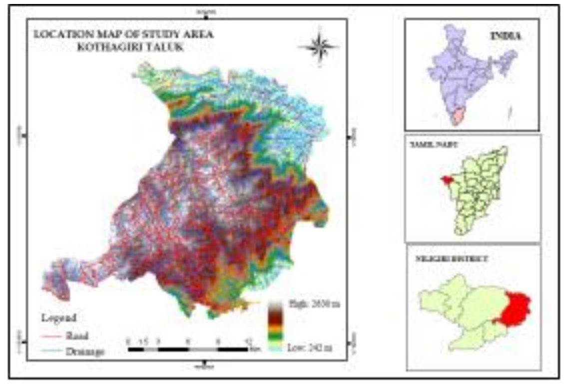

Figure 1 *Location of the study area*

3. Materials and Methods

For modeling the landslide susceptibility, the main stage is collection of data and generation of spatial database and the relevant landslide causative factors are extracted. Data is used employing GIS, for preparing the landslide susceptibility maps for ANN. All the data are utilized in the effort of the selection of the influencing parameters and for their GIS thematic layers. Thirteen factors are considered for the landslide susceptibility analysis, and the factors are extracted from the created spatial database. The factors are transformed into a vector-type spatial database using GIS. For the investigation, the inventory map is processed based on field survey for determining the exact location. Eighty four landslide locations are identified (Fig. 2) for the study; fifty nine landslide locations are used for training set and the rests are used for validation purpose.

Digital Elevation Model (DEM) is generated from the topographic database with resolution of 10 m contour interval. Using DEM, the slope angle, slope aspect, slope curvature and elevation are calculated. In the case of the curvature positive (> 0) values of plan curvatures express convexity in the slope direction, negative (< 0) values of plan curvatures illustrate concavity of slope curvature in the slope direction. Plan curvatures value around zero indicates the flatness of surface. The curvature map is prepared using the avenue routine in Spatial Analyst Tool in ArcGIS 10.0. Additionally, the distance from drainage is calculated using the topographic database. The distance from drainage is calculated in 100 meter intervals using topographic database. Using the geology database, the types of lithology is extracted, and the distance from lineament are calculated. Landuse / Landcover data is classified using IRS 1D LISS IV image employing an unsupervised classification method and it is verified with field survey. Forest evergreen, forest deciduous, forest scrub, forest plantation, barren land, agricultural plantation, fallow and built-up regions are extracted for land cover mapping and NDVI map is obtained from IRS ID LISS IV satellite images. The NDVI value denotes the areas of vegetation in an image. In this study, Survey of India (SOI) topo sheet (scale: 1:25000) is used to detect the landslide locations. Recent landslides are observed during the survey. The landslide-susceptibility modeling is a function of different factors (Table 1). For the landslide-susceptibility analysis, the study area and the spatial correlations between the variables and neural network equations are established. Further, using the detected landslide locations and the constructed spatial database, the weights for all the variables are computed. The calculated and extracted factors are converted to 10 m x 10 m grid (GIS grid) and it is converted to ASCII data to be used further by an artificial neural network program run in MATLAB 8.0. The analysis results are converted to grid data using GIS. The dimensions of the study area grid are 1706 rows and 2325 columns with 3966500 pixels. 2380 pixels denoting areas where landslide not occurred or occurred. The flow chart for constructing ANN is given in Fig. 3.

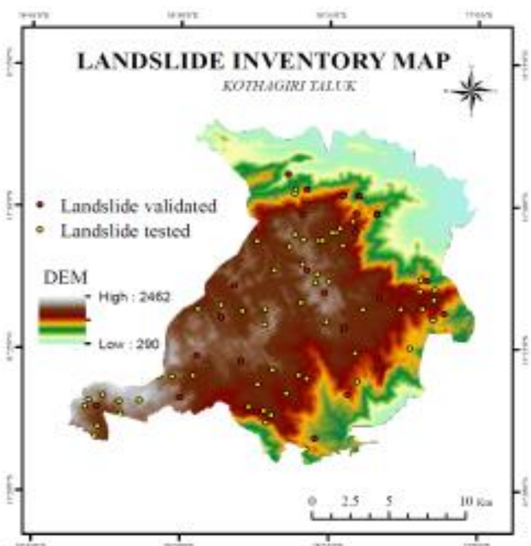

Figure 2 *Landslide inventory map of the study area*

Table 1: *Thematic layer for modeling ANN*

Category	Layer	Data formats	Scale
Landslide	Landslide Inventory	Point	1:25,000
Topographic map	Elevation	Point and line	1:25,000
	Slope angle and aspect	GRID	10 m × 10 m
	Curvature and SPI	GRID	10 m × 10 m
Drainage map	Distance from drainage	Polygon	1:50,000
Geological map	Lithology	Polygon	1:50,000
Lineament map	Distance from lineament	Polygon	1:50,000
Soil map	Soil texture	Polygon	1:50,000
Road map	Distance from Road	Polygon	1:25,000
Precipitation map	Precipitation	GRID	1:25,000
LULC map	Land use and land cover	GRID	5.8 m × 5.8m
NDVI map	NDVI	GRID	5.8 m × 5.8m

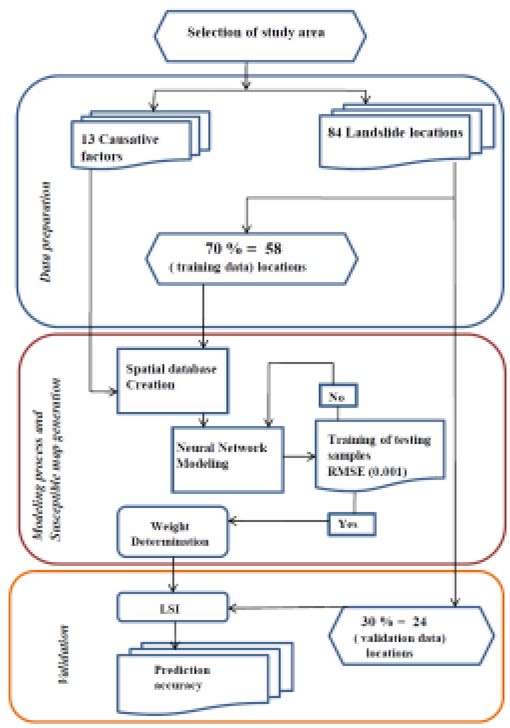

Figure 3 *Methodology adopted for the modeling - ANN.*

3.1. Artificial Neural Network

ANN is a structure which includes densely interconnected adaptive simple processing elements that are capable of performing massively parallel computations for data processing and knowledge representation [13]. Hypothesis behind neural networks is based on an attempt to imitate human learning processes through the creation of artificial neural networks. The model proposed in this work is based on the neural analysis technique which is considered as a quantitative technique and a black-box model [14]. ANN is non-linear program suitable for evaluation of indirect landslide susceptibility assessment. It provides excellent predictions even the data contains noisy and vague. An advantage of ANN method is independent from the statistical distribution of the data and there is no need for specific statistical variables [15]. The individual weights of all factors for landslide occurrences are successively created from the ANN model. A back-propagation network is a multi-layer neural network (MNN). The MNN with back-propagation neural network (BMNN) has been used as a mapping and prediction tool in the earth science studies. The back-propagation technique has extended the types of problems to which ANN can be applied [15, 16] developed a method to integrate ANN to calculate the Landslide Susceptibility Index (LSI). The results are checked by ranking the susceptibility index in classes of equal area and the results indicated a good agreement between the susceptibility map and the landslide inventory.

3.2. Modeling of Artificial Neural Networks

Artificial Neural Networks are computational networks which attempt to simulate the networks of nerve cell (neurons) of the biological central nervous system. ANN is networks of highly interconnected neural computing elements that have the ability to respond to input stimuli and to adapt to the environment. It establishes rules during the learning phase and uses the rules to predict outputs. The input neurons in the neural network intrinsic factors to the landslide susceptibility such as slope, aspect, curvature, elevation, geology, soil texture, landuse, NDVI, distance from drainage, distance from fault, distance from road; data from different sources and precipitation are stored in a GIS and the factors generated an ANN model using the back-propagation technique for landslide prediction. The weights of the attribute layers are computed using an ANN. This method consists of three layers viz., input, hidden and output layers (Fig. 4).

There are thirteen input layer neurons and one hidden layer neurons that will allow complexities to develop in the mapping. The hidden layer neurons and output layer neurons method their inputs by multiply all input by a related weight, summing the product, and processing the sum using a nonlinear transfer function to produce a result [17]. An ANN method learns by adjusting the weights among the layer neurons in response to the errors between the definite output values and the end output values. 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 phases involved in using ANN for multi-layer classification: the training phase, in which the internal weights are attuned, and the classifying phase. Normally, the back-propagation program trains the all multi-layered complex until some targeted minimal error is achieved among the desired and actual output values of the neural network. Once the training is accomplished, the network is used as a feed-forward structure to produce a classification for the entire data [17]. The back-propagation Multi-Layer Perception (MLP) is a commonly used neural network structure in spatial analysis. The modeling of the ANN has been computed in a program developed in MATLAB 8.0. This module uses back propagation (BP), the learning algorithm which allows the modification of connection weights so as to minimize function error. The updating of weights represents the training phase of the neural network. The neural network is prepared by adjusting factors viz., number of hidden nodes, training pixels per category, momentum factor, learning rate, number of training cycles (iterations) and Root Mean Square Error (RMSE). The probability of occurrence of landslides is calculated based on various input parameters and knowledge-based classification. The various input thematic GIS layers are mostly in vector data format and it is converted into raster grid format using ArcGIS software.

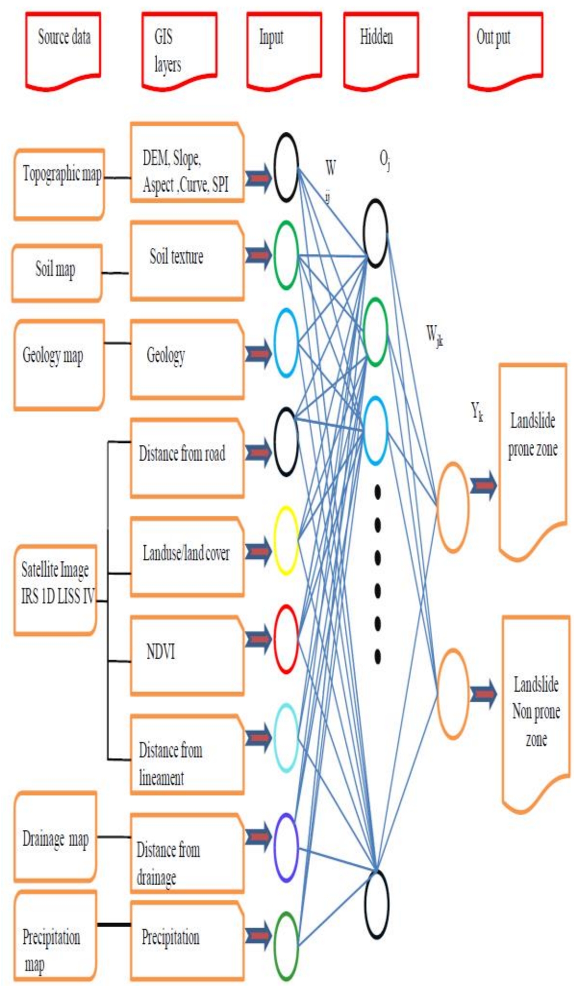

***Figure 4** Three tiered architecture of feed-forward, back-propagation neural network (after Pradhan and Lee. 2009).*

Table 2: *Weights of factors estimated by ANN*

Factors	1	2	3	4	5	6	7	8	9	10	SD	Mean	Weight
Rainfall	0.112	0.125	0.148	0.137	0.139	0.141	0.143	0.142	0.140	0.132	0.011	0.136	0.608
Slope	0.198	0.210	0.205	0.205	0.206	0.209	0.205	0.207	0.204	0.207	0.003	0.206	1.000
Aspect	0.025	0.022	0.020	0.024	0.024	0.023	0.021	0.023	0.024	0.020	0.002	0.023	0.136
Elevation	0.072	0.071	0.058	0.074	0.066	0.067	0.075	0.076	0.075	0.071	0.006	0.071	0.336
Plane Curvature	0.076	0.075	0.071	0.095	0.065	0.077	0.045	0.054	0.056	0.069	0.014	0.068	0.326
Geology	0.124	0.125	0.105	0.116	0.129	0.119	0.128	0.118	0.115	0.108	0.008	0.119	0.536
Soil Texture	0.082	0.092	0.090	0.092	0.089	0.085	0.091	0.095	0.095	0.089	0.004	0.090	0.417
NDVI	0.079	0.062	0.099	0.085	0.064	0.078	0.078	0.085	0.082	0.081	0.011	0.079	0.372
Landuse/Landcover	0.056	0.074	0.071	0.073	0.073	0.075	0.075	0.075	0.073	0.074	0.006	0.072	0.341
Distance from road	0.085	0.084	0.085	0.084	0.087	0.081	0.081	0.085	0.082	0.081	0.002	0.084	0.390
Distance from river	0.098	0.095	0.094	0.092	0..940	0.094	0.096	0.099	0.094	0.097	0.002	0.095	0.439
Distance from fault	0.089	0.084	0.085	0.106	0.089	0.081	0.096	0.101	0.105	0.091	0.009	0.093	0.428
SPI	0.015	0.015	0.019	0.018	0.020	0.018	0.013	0.013	0.018	0.017	0.002	0.017	0.111

3.3. Training of the ANN

Before modeling the ANN program, the data are commonly separated into two sub datasets, such as training dataset and test dataset. The training dataset include entire data belonging to the related problem. The data set is used in the training phase of the model development to update the weights. The test dataset is entirely different from those used in the training phase. The importance of the sub dataset is to check the network performance using untrained data and its accurateness. There is no precise mathematical law to establish the required minimum size of these sub datasets [18]. About 70% of the dataset is usually considered as sufficient for training the network and the rest of it is normally reserved for testing the model [19]. The landslide occurrence area and the landslide-not-occurrence area are selected as training data. Pixels from the two classes are randomly selected as training pixels denoting areas where landslide not occur or occurred. The area where landslides have not occurred and slope is zero; it is classified as areas not occurred to landslide whereas landslides are well-known to exist are assigned as "areas prone to landslide" training set. Training sites have been selected based on landslide location as prone training site and with a varying slope values as non-prone training site and trained back propagation algorithm has been computed. The selected landslide -prone locations are assigned as (0.1, 0.9), and the non-landslide-prone locations are assigned as (0.9, 0.1). To minimize the error among the predicted output and the calculated output values, the back propagation algorithm is used. In this modeling 70% are randomly selected for training the ANN and the remaining are used for the prediction testing.

3.4. Determination of weights using the ANN model

The back-propagation algorithm is applied to calculate the weights between the input and the hidden neuron as well as the hidden and the output neuron. The weights between layers acquired by neural network training are calculated in reverse; the contribution of each factor is determined. All the causative factor layers are used as input dataset and the landslide locations are used as training datasets. In the ANN model, the back-propagation method is applied. A three-layered feed-forward network is applied using the MATLAB 8.0 software package based on the framework [20]. Feed-forward denotes that the interconnections between the layers propagate forward to the second layer. The number of hidden layers and the number of nodes in a hidden layer required for a particular classification problem are difficult. The flowchart of neural network training for weight determination is depicted in Fig.5. It consisting of an input layer (13 input neurons) × one hidden layer (30 hidden neuron) × one output layer (2 output neurons) is selected for the network model of 13-30-2, with input data normalized within the range between 0.1–0.9. In order to training the network, the landslide inventory map is reclassified by assigning a value of 1 to the landslide pixels and a value of 0 to the not-landslide pixels. From the two classes (landslide and non-landslide), 1666 pixels per class are selected as training pixels at random, as proved that the number of training locations did not influence the analysis. The nominal and interval class group data are converted to continuous values ranging between 0.1 and 0.9. The continuous values are not ordinal data, however nominal data, and the numbers denote the classification of the input data. The learning rate is set to 0.01 and the initial weights are randomly selected 0.1 to 0.9. The computation is repeated ten times to predict if the randomly extracted sample represented each class. The result shows that there are no much differences in the ten classes. First, the initial weights are assigned using the random values, and the results showed different values in each iteration. Hence, in this study, the calculation is repeated 10 times and the results revealed similar values. The selected landslide locations are assigned (0.1, 0.9) and the non-landslide - prone locations are assigned (0.9, 0.1). The weights calculated from ten tests are compared to determine whether the variation in the final weights which is dependent on the selection of the initial weights. The back-propagation computation algorithm is used to minimize the error between the predicted output values and the calculated output values. The ANN algorithm propagated the error backwards and iteratively adjusted the weights. The number of epoch is set up to 10,000, and Root Mean Square Error (RMSE) value used for the stopping criterion is set to 0.01. Most of the training datasets met the 0.011 RMSE goals. However, the maximum number of iterations is set to 10,000 epochs default. After the training dataset, the weights of all factors have been computed by neural network. RMSE is utilized to interrupt the training phase when it is lower than the specified value (0.019 selected for this study). The final weights between the layers are obtained during training and the contribution of all the factors is used to predict susceptibility of landslide. After training, the weights are determined.

The result reveals that the initial weight did not influence the final result. The two classes (landslide and non-landslide), 1666 pixels per class are randomly selected as training pixels. The calculation is repeated ten times by randomly assigning the initial weights to determine how the extracted sample represented each class, whereas the final results are not same. The standard deviation is distributed from 0.002 to 0.011 so as the random sampling does not have a large effect on the result. For easy understanding, the average values are calculated and the values are normalized by the average of the weights of the SPI which is used at the minimum value 0.111 and the slope is used at the maximum value 1.000 (Table. 2). The final weights between layers acquired during training of the neural network and the contribution or importance of all the factors used to predict landslide susceptibility index.

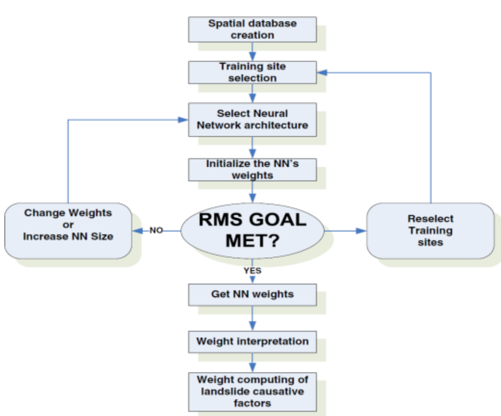

Figure 5*Flow chart for weight determination using ANN model (after Pradhan 2010)*

3.5. Preparation of Landslide Susceptibility Map

In the ANN model, the networks are successfully trained and the weights are computed by both back-propagation and the spatial database. After the network is well trained, the study area is fed into the network in order to estimate the landslide susceptibility index. The landslide susceptibility index contains the minimum grid value of 0.146 and maximum value of 0.928 for each pixel. Using the values, the Landslide Susceptibility Indices (LSI) are determined and used to create the landslide susceptibility maps. The set of susceptibility values derived in each grid are converted to a raster file and a landslide susceptibility map is produced using ArcGIS software (Fig. 6). The values of susceptibility are reclassified into five classes (very low, low, moderate, high, and very high) by means of the natural-breaks method [21]. This method identifies break points by picking the class that breaks the best group in similar values, maximizing the differences between classes [22]. After the tabulation of the result, the area and percentage distribution of the susceptibility classes are determined.

4. Results and Discussion

Validation

Validation technique is performed by comparing of existing landslide data and landslide susceptibility model results. 30% of the landslide locations are used in the training phase. Global Positioning System (GPS) has been used to collect the landslide locations in the study area. There are 25 active landslides has been identified and added to the inventory for the validation of the result of ANN model. The locations are superimposed on the landslide susceptibility classification obtained by applying the weights derived from the ANN. For validation, two basic assumptions are required to validate the landslide susceptibility computation. The first one is landslide related to spatial data such as elevation, geology, NDVI, landuse, distance from lineaments, distance from drainage and rainfall, and the next is about the future landslides which will be predicted by specific impact factors such as slope and rainfall [23]. For this, the two basic assumptions are satisfied as the landslides and related to the spatial database, slope and rainfall. On the classified susceptibility map, more than 25.53 % of the study area fall under the class of high to very high susceptibility and 56% of the active landslides occur in this class. Landslide susceptibility is very high where the gradient of the slope exceeds 35° which is in the north and southwest part of the study area. The result of the susceptibility map, 52.03% of the total area illustrate that very low and low landslide susceptibility, occurs where the gradient of slope ranges between 0° and 10°, rainfall values are high (> 250 mm). In the landslide susceptibility maps, the landslide pixels generally coincided with the sites falling in the very high, high and moderate susceptibility classes. A relationship between the susceptibility map (Fig.6) and the landslide inventory map, conforms the appropriate classification system 59% of the landslides data set which have been falling in the classes of high and very high susceptibility classes. It is observed that 95% of landslide pixels are appropriately classified. The result illustrates that there is a fair agreement between the prediction accuracy and the occurrence of landslide.

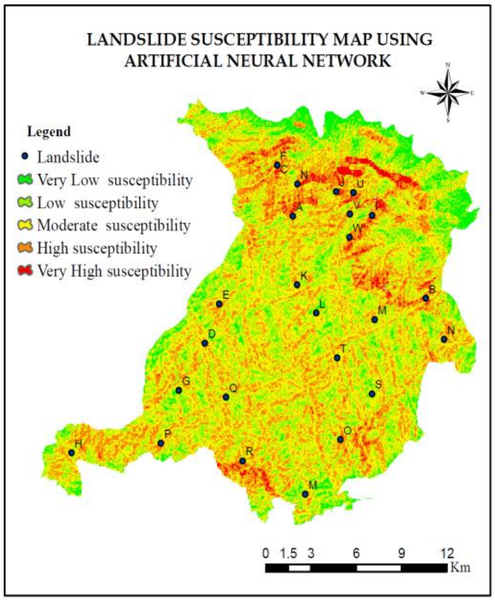

Figure 6 *Susceptibility map of BPNN methodology based on the ANN model*

The total area with their percentage occupied by various landslide susceptibility zones in the image is determined. Relative frequencies of areas influenced by different landslide susceptibility zones are calculated from the frequency ratio. Ideally the frequency ratio value should increase from a very low susceptible zone to a very high-susceptible zone, since the highest landslide susceptible zones are generally more prone to landslides than other zones. Table 3 illustrates the relative frequency ratio of landslide susceptibility classes of ANN model. The landslide susceptibility model obtained is further validated with the goal of checking how well the model matched the occurrences of landslides; it is obtained by means of Receiver Operating Characteristic (ROC) curves (Fig. 7). Accuracy of the model is evaluated using the ROC curves. The result of the neural network model and its predictive capability the ROC curve has been developed [24]. For comparing the result quantitatively, the area under curve is recalculated as a fraction of the total area of the chart.

Table 3: *Frequency ratio associated with the landslide susceptibility classification*

Susceptibility classification	Number of pixels	Area (%)	Landslide (%)	Frequency ratio
Very low	1269280	32.00	0	0
Low	794489	20.03	16	0.798
Moderate	890082	22.44	28	1.247
High	611634	15.42	44	2.853
Very high	401013	10.11	12	1.186

ROC curves plot the sensitivity versus specificity where the sensitivity is the proportion of correctly classified known landslide grid cells as unstable, and the specificity is the proportion of grid cells outside a mapped landslide that is correctly classified as stable [25]. The ROC curve is derived by plotting all grouping of the sensitivities and proportions of false negatives. Sensitivity is the fraction of positive occurrences of landslide that is correctly predicted, while 1-specificity is the fraction of incorrectly predicted cases that did not occur [26]. It confirms the swap between the two successive rates [27]. The ROC curves can be summarized quantitatively with the help of the area under the AUC which gives the accuracy of the developed model for predicting the landslide susceptibility. The AUC value ranges between 0 and 1; a higher value point out a higher prediction rate and the value near 0.5 means the prediction is no better than a random guess [25]. AUC values less than 0.7 reveals a poor performance of the model; values between 0.7 and 0.8 reflect as fair performance of the model; values between 0.8 and 0.9 can be considered as a good and values above 0.9 can be considered as excellent. In this study, all the landslide free cells are considered to be negative for the ROC curve computation. For the purpose of validation, 25 landslide locations are superimposed on the LSI classification. The rate curves for ANN landslide susceptibility map is illustrated (Fig. 7). The ROC curve is a graphical representation of the trade-off between the false-negative and false-positive rates for every possible cutoff value. By tradition, the plot shows the false-positive rate (FPR) on the X axis and the true-positive rate (TPR) on the Y axis. The area under the ROC curve (AUC) characterizes the quality of a forecast system by describing the system's ability to anticipate the correct occurrence or non-occurrence of pre-defined ''events.'' The best method has a curve with the largest AUC; the AUC varies from 0.5 to 1.0. If the model does not predict the occurrence of the landslide any better than chance, the AUC would equal 0.5. A ROC curve of 1 represents perfect prediction. In this investigation reveals the landslide susceptibility model using ANN, the AUC is 0.875 and the prediction accuracy is 87.50 %. Consequently, the ANN derived susceptibility map can be considered as good prediction. The model with ANN has the highest area under the curve value of 0.875. The predicted rate explains how well the model and predictor variable predicts the landslide [28, 29].

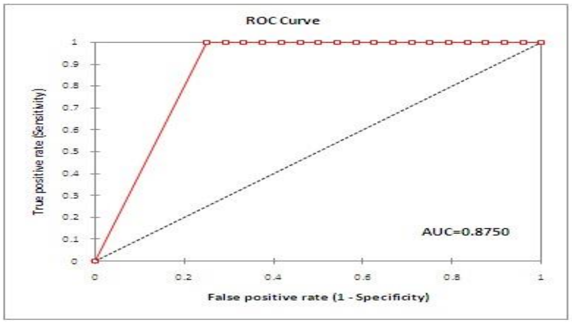

Figure 7 *Receiver operating characteristics plots for the susceptibility maps produced showing false positive rate (x-axis) vs. true positive rate (y-axis).*

5. Conclusions

Landslide susceptibility modeling for Kothagiri region is computed by using an artificial neural network model, employing a back-propagation learning algorithm in the MATLAB 8.0. Slope, aspect, curvature, geology, soil texture, distance from drainage, distance from road, distance from lineaments, SPI, land cover, NDVI, and rainfall parameters are used as the inputs to the ANN model. Through multiple field survey, a landslide inventory map is created. Among the 84 landslides identified from historical records and field survey, the landslides are classified into rotational slide, transitional flow and flow slide. The application of neural network is divided into two stages: first stage is training phase in which random selection of landslide location sites (70 %) are selected as training sites and the weights are calculated, and the validation procedure is the second stage in which the obtained susceptibility map is verified with the remaining 30 % of the landslide inventory map. From this method of an artificial neural network, the relative importance, weights, among the factors are computed. The study reveals that slope is the most important causes of landslide occurrence in the Kothagiri region. The slope illustrates the highest weight index (1.00). Steep slope (> 25^{o}) with concave curvature values and high proximity to stream network areas are found as more susceptible to landslides. The second important causative factor to landslide occurrence is rainfall (0.608) and geology (0.536); areas covering granitic rocks with gentle slope are found as the most susceptible lithology compared to the other rocks. The third important parameter contributing to landslide occurrence is distance from the river (0.439), distance from the fault (0.428) and soil texture (0.417). NDVI (0.372) and land use (0.341) play the next important role on landslide occurrence in the study area. Generally, landslides occurred in fallow land and barren land, and there are few landslides in the dense forest areas. An ANN weight illustrates that the less important factors are curvature (0.326), aspect (0.136), and SPI (0.111). From the results, it is concluded that the slope is the most important factor due to its weight is almost two times higher than the other factors for landslide susceptibility mapping. Slope gradient ranges between 15° to > 35° are found to be more susceptible to slide. Using the ANN weights, the landslide susceptibility Index and map are formed. Furthermore, the landslide susceptibility map is classified into five classes, and frequency ratio analysis is also performed with the landslide location data. The 25 % of total landslide location used for validates the applicability of ANN to landslide susceptibility mapping. The general trend of the relative frequency value illustrated that highest landslide susceptibility zone is generally more prone to landslides than other zones. In order to assess the impact of the derived results, the ROC cross-validation method is applied. The validation reveals that 87.50 % of the total landslide total area is properly classified by the ANN model. Generally, the validation results demonstrated satisfactory agreement between the landslide susceptibility map and the existing landslide location. These results can be used to assist slope management, urban development, and land use planning. However but the models used in the study are valid for generalized planning and assessment purposes.

References

[1] M.H. Vahidnia, A.A. Alesheikh, A. Alimohammadi and F. Hosseinali, "Hospital site selection using fuzzy AHP and its derivatives", Journal of Environmental Management, Volume No.90, Issue No.10, pp 3048-3056, 2009.

[2] G. Manimaran, R.A. Antony, S. Selvam, D. Manimaran and M. Sugan, "Characterization and disaster management of landslides in the Nilgiris mountainous terrain of Tamil Nadu, India", International Journal of Geomatices and Geosciences, Volume No.3, Issue No.1 pp 1-12, 2012, .

[3] I. Yilmaz, "Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Condi-tional probability, logistic regression, artificial neural net-works, and support vector machine", Environ. Earth Sci., doi: 10.1007/s12665-009-0394-9, 2010,

[4] M. Mohammady, H.R. Pourghasemi and B. Pradhan, "Landslide susceptibility mapping at Golestan Province Iran: a comparison between frequency ratio, Dempster-Shafer, and weights-of evidence models", J Asian Earth Sci 61, pp 221–236, 2012,

[5] C. Gokceoglu, H. Sonmez and M. Ercanoglu, "Discontinuity controlled probabilistic slope failure

risk maps of the Altindag (settlement) region in Turkey", Eng. Geol, pp 55277–296, 2000

[6] Y. Song, J. Gong, S. Gao, D. Wang, T. Cui, Y. Li, and B, Wei, Susceptibility assessment of earthquake-induced landslides using Bayesian network: a case study in Beichuan, China, Comput Geosci Volume No.42 pp 189–199, 2012.

[7] M. Zare, H.R. Pourghasemi, M. Vafakhah and B. Pradhan, "Landslide susceptibility mapping at Vaz watershed (Iran) using an artificial neural network model: a comparison between multi-layer perceptron (MLP) and radial basic function (RBF) algorithms", Arab J Geosci. Doi: 10.1007/s12517-012-0610-x, 2012.

[8] A. Akgun, "Comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision and likelihood ratio methods: a case study at İzmir, Turkey", Landslides Volume No .9, pp 93–106, 2012.

[9] H.R. Pourghasemi, B. Pradhan, C. Gokceoglu, H.R. Mohammadi, "Application of weights-of-evidence and certainty factor models and their comparison in land-slide susceptibility mapping at Haraz watershed", Iran, Arab. J. Geosci., doi: 10.1007/s12517-012-0532-7, 2012.

[10] P.V. Gorsevski, P. Jankowski and P.E. Paul, Heuristic approach for mapping landslide hazard integrating fuzzy logic with analytic hierarchy process, Control Cybern, Volume No.35 Issue.1, pp 1–26, 2006.

[11] A. Carrara, and R.J. Pike, "GIS technology and models for assessing landslide hazard and risk", Geomorphology, Volume No.94, pp. 257-260, 2008.

[12] M. Kartic Kumar and R. Annadurai, "Mapping of Landslide Susceptibility Using Geospatial Technique - A Case Study in Kothagiri Region, Western Ghats, Tamil Nadu, India", International Journal of Engineering Research & Technology Volume No. 2, pp 1-12, 2012.

[13] R.J. Schalkoff, *Artificial Neural Network*, McGraw-Hill, New York, p 448, 1997.

[14] P. Aleotti and R. Chowdhury, "Landslide hazard assessment: summary review and new perspectives", Bull Eng Geol Environ Vol.58, pp 21–44, 1999.

[15] S. Lee, J.J. Ryu, J.S. Won and H.J. Park, "Determination and application of the weights for landslide susceptibility mapping using an artificial neural network", Eng Geol Volume.71, pp 289–302, 2004.

[16] K.M. Neaupane, and S.H. Achet, "Use of back propagation neural network for landslide monitoring: a case study in the higher Himalaya", Engineering Geology, Volume No.74, pp. 213–236, 2004.

[17] B. Pradhan, "An Assessment of the Use of an Advance Neural Network Model with Five Different Training Strategies for the Preparation of Landslide Susceptibility Maps", Journal of Data Science, Volume No.9, pp 65-81, 2011.

[18] H.A. Nefeslioglu, C. Gokceoglu, and H. Sonmez, "An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps", Engineering Geology Volume No 97, pp 171, 2008.

[19] K. Swingler, *Applying neural networks: A practical guide.* New York: Academic Press. 1996,

[20] J.W. Hines, *Fuzzy and neural approaches in engineering* New York: John Wiley and Sons Inc, 1997.

[21] G.F. Jenks, "Geographic logic in line generalization", Cartographica, Volume No.26, 27–42, 1989.

[22] P.R. Federici, A. Puccinelli, E. Cantarelli, N. Casarosa, G. Amato Avanzi, F. Falaschi, R. Giannecchini, A. Pochini, A. Ribolini, M. Bottai, N. Salvati and C. Testi, "Multidisciplinary investigations in evaluating landslide susceptibility—an example in the River valley (Italy)", Quaternary International Volume No.171, pp 52–63, 2007.

[23] S. Lee, J.H. Ryu, K. Min, and J.S. Won, "Landslide susceptibility analysis using GIS and artificial neural network", Earth Surf Proc Landform, Volume No.28, pp 1361–1376, 2003.

[24] Massimo Conforti , Stefania Pascale , Gaetano Robustelli and Francesco Sdao, "Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy)", Catena, http://dx.doi.org/10.1016/j.catena.2013.08.006 , 2013

[25] S. Beguería, "Validation and evaluation of predictive models in hazard assessment and risk management, Natural Hazards", Volume No. 37 pp 315–329, 2006.

[26] J.A. Swets, "Measuring the accuracy of diagnostic systems, Science", Volume No.240, pp. 1285–1293, 1988.

[27] M.H. Zweig and G. Campbell, "Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine", Clinical Chemistry Volume No.39, pp 561-577, 1993.

[28] D.T. Bui, B. Pradhan, O. Lofman, I. Revhaug and O.B. Dick, "Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models", Catena, Volume No.96, pp 28-40, 2012.