LaSaRiZ: A Software for Landslide Susceptibility Zonation and Risk Assessment using Neural Network and Fuzzy Set Approaches

M.K. ARORA [1], S. CHAUHAN[2], M. SHARMA[3] and D.P. KANUNGO[4]

[1] *Professor, Department of Civil Engineering, IIT Roorkee, Roorkee-247667*

[2]*Geo-Information Analyst, Shell Technology India, Bengaluru - 560048*

[3]*Assistant Professor, Geology, H.P. University Regional Center, Dharamshala-176215*

[4]*Scientist, CBRI, Roorkee-247667*
(*manojfce@iitr.ernet.in*)

ABSTRACT: Availability of accurate and objective landslide susceptibility maps depicting zones defined on the basis of probability of occurrence of landslides is one of the critical inputs in assessing risk to property and lives in any mountainous region, particularly in the Himalayas. The aim of this study is to portray the usage of an in-house developed indigenous software, acronymmed as **L**andslide **S**usceptibility **a**nd **Ri**sk **Z**onation (LaSaRiZ), for landslide susceptibility zonation and risk assessment. The software incorporates the implementation of three recently proposed objective approaches for Landslide Susceptibility Zonation (LSZ) and two new approaches for Landslide Risk Assessment (LRA). The efficacy of the software was examined through a case study in Western Himalayan region of India. The results from the software indicate that LSZ map based on combined neural network and fuzzy approach performed exceedingly better than those produced from neural network black box approach and fuzzy relation based approaches. The landslide risk assessment maps obtained through the two new approaches implemented in this software were able to clearly depict the resource categories, which were under risk due to landslides in the region.

1   INTRODUCTION

Landslides in the Himalaya are one of the major and widely spread natural hazards that often strike life and property and are a major concern. One of the requirements for effective landslide mitigation and management program is the availability of an accurate Landslide Susceptibility Zonation (LSZ) map. Preparation of LSZ maps requires evaluation of the relationships between various terrain conditions and instances of landslide occurrence. A skilled earth scientist, through his vast experience based on the assessment of the overall terrain conditions usually identifies the causative factors affecting the occurrence of landslides in a region. On their assessment, these factors and their categories are assigned weights and ratings respectively as per their importance in landslide occurrences. The knowledge in the form of weights and ratings is typically input to any LSZ process in several different ways, which can either be manual or in a Geographic Information System (GIS) environment as is evident from a number of studies (e.g., Saha *et al.* 2002, Sarkar and Kanungo 2004, Saha *et al.* 2005, Pareek *et al.*, 2010, Chauhan *et al.*, 2010a). Thus, in these studies, the weights have usually been assigned on the basis of the knowledge domain of the expertise about the subject and the area. This weight assignment strategy, at times, may however be highly subjective and may therefore contain some implicit biasness.

Therefore, in order to reduce the subjectiveness in weight assignment procedure, a number of alternative strategies (e.g., Saha *et al.* 2005) have been attempted in recent years for LSZ mapping in Himalaya and other parts of the world. Most of these studies are based on establishing the relationships between categories of the causative factors and incidences of the existing landslides in a given region through spatial data analyses. These relationships are defined in the form of weights and ratings. Thus, a number of data driven approaches have been proposed, which include logistic regression and multivariate statistical methods (Chauhan *et al.*, 2010b), artificial neural network (Arora *et al.* 2004; Kanungo *et al.* 2006), fuzzy relations (Kanungo *et al.* 2006) and neuro-fuzzy approaches (Kanungo *et al.* 2006).

However, these studies, had to depend on various independent software resources to implement ANN, fuzzy set and statistical data processing due to non-availability of a dedicated software to fulfill the needs of the study. Difficulties are, therefore, faced in importing and exporting of data in different format from one software to another for LSZ mapping. In this study, an exclusive and dedicated GUI based software acronymmed as LaSaRiZ (Landslide Susceptibility and Risk Zonation) has been developed to produce LSZ and LRA maps using a combined neuro-fuzzy approach, as documented in one of earlier papers cited as Kanungo *et al.* (2006).

The traditional backpropagation ANN is implemented for the determination of weight of causative factors via a connectionist weighting process (Olden *et al.* 2004). A fuzzy relation concept based on cosine amplitude method is implemented to determine ratings (equivalent to fuzzy membership values) of categories of causative factors. The backpropagation ANN, as implemented in LaSaRiZ, can also be used as a black box to

independently produce an LSZ map. Similarly, fuzzy relation concept can also be used to independently produce an LSZ map.

Further, since LSZ maps form key inputs to any landslide risk assessment (LRA) process, the LaSaRiZ also includes two LRA approaches to produce a landslide risk assessment map of a region. Two new semi-quantitative landslide risk assessment approaches, one based on danger pixel concept and the other based on fuzzy set theory, are implemented in LaSaRiZ software to quantify landslide risk at regional level.

The inputs to LaSaRiZ software are a set of raster thematic data layers, each pertaining to a selected causative factor. These thematic data layers were assumed to be created through digitization of past maps or derived from remote sensing images using an image processing and GIS software. The outputs from LaSaRiZ software are a set of LSZ and LRA maps. The LaSaRiZ software was then effectively used to carry out landslide susceptibility and risk assessment study in Garhwal Himalayas. The evaluation of landslide susceptibility maps derived from the implemented approaches was performed through landslide density analysis and interpretation of Receiver Operating Characteristic (ROC) curves.

2 LANDSLIDE SUSCEPTIBILITY APPROACHES

The three LSZ approaches incorporated in the software are; the ANN black box approach, the fuzzy relation concept based approach and the combined neural network and fuzzy approach.

2.1 Artificial Neural Network Black Box Approach for LSZ

ANNs are generic non-linear function approximators extensively used for pattern recognition and classification like problems. The categorization of a terrain into ordinal zones of landslide susceptibility may also be regarded as a classification problem. For example, in a typical neural network architecture for landslide susceptibility zonation, seven input neurons correspond to seven thematic data layers (one for each causative factor affecting landslide) and one output neuron corresponds to presence or absence of landslide.

Three stages involved in ANN data processing for LSZ problem are; the training stage, the weight determination stage and the classification stage. The network output values for the whole dataset are categorized into one of the five landslide susceptible zones to produce the LSZ map from ANN. This approach has been referred to as ANN black box approach, since in this case the weights remain hidden (Arora *et al.* 2004).

2.2 Cosine amplitude based fuzzy relation approach for LSZ

The fuzzy set contains elements that have varying degrees of belongingness (or memberships) in the set (Ross 1995). Fuzzy relations between dependent variable (i.e., LSZ classes) and independent variables (i.e., causative factors) can be characterized based on the similarity in the data play an important part in fuzzy modeling (Dubois and Prade 1980). These methods will lead to determination of strength of relationships or membership values of the categories of each factor, which can be assumed as surrogates of ratings of the categories (Kanungo *et al.* 2006). The ratings can be integrated to produce an LSZ map.

In this software, cosine amplitude similarity method of fuzzy relation is incorporated to produce ratings of each category of a causative factor. The method evaluates the relation between the existing landslide occurrence and the categories of each causative factor considered. The categories of existing landslide distribution layer and categories from each thematic data layer (corresponding to each causative factor), taken one at a time, were considered as two datasets for the computation of ratings or strength of relationship. In the landslide distribution layer, the pixels in the landslide areas were assigned a value of 1, whereas remaining pixels were assigned a value of 0. Similarly, a value of 1 was assigned to pixels belonging to a particular category of a causative factor and a value of 0 to the remaining pixels. The membership values of each category for each fuzzy set (i.e., causative factor) were determined by the strength of relationship between the landslide incidences and causative factors.

2.3 Combined neural network and fuzzy approach for LSZ Mapping

In the ANN black box approach, the weights and ratings remain hidden and are not known. In fuzzy relation based cosine amplitude approach, the LSZ map is produced based only on the ratings of the categories of causative factors. A combined neural network and fuzzy approach takes advantage from both to integrate ANN derived weights and fuzzy set derived ratings to produce an LSZ map. This map is expected to be more accurate than the LSZ maps produced either from ANN or fuzzy relation concept. The combined neural and fuzzy approach involves three steps,

    i)   Determination of weights of causative factors through ANN connection-weight approach
    ii)   Determination of ratings for categories of causative factors using cosine amplitude method

iii)   The integration of weights and ratings to produce the LSZ map

3   LANDSLIDE RISK ASSESSMENT APPROACHES

Landslide Risk Assessment (LRA) is an extremely important aspect in practical applications of landslide studies. The LRA can be regional or site-specific in nature, depending upon the varied applications at different stages of the decision-making process. The landslide risk assessment component of LaSaRiZ software, consists of implementation of two markedly different approaches, named as, danger pixel approach and fuzzy linguistic scale based approach.

3.1 Danger pixel approach for LRA

In this approach, landslide risk assessment was considered as a function of landslide potential and the resource damage potential. Landslide potential (LP) was equated in terms of danger pixels defined as those pixels that lie in very high and high landslide susceptible zones in the LSZ map (Kanungo *et al.* 2008). Thus, a danger pixel map was prepared. It is a binary map in which a pixel value of 1 corresponds to either very high or high landslide susceptibility zone, and is regarded as the danger pixel whereas rest of the pixels are assigned a value of 0 and are regarded as no danger pixel. For assessing the resource damage potential (RDP), land use and road network in the region were considered as resources as the potential candidates for damage due to occurrence of landslides. Thus, a resource map was produced by integrating the land use land cover map with the road network map of the area. The resource map is a categorical data thematic data layer in which the pixel values depict either the type of land use or the road. The pixel values in resource map and the danger pixel map were multiplied to ascertain the risk so as to prepare a landslide risk assessment (LRA) map. Thus, the LRA map produced from danger pixel approach provides a spatial distribution of resources that appear to be under risk due to landslides.

3.2 Fuzzy set based approach for LRA

Similar to danger pixel approach, the linguistic scale based approach for landslide risk assessment also requires two maps (i.e., LSZ map and resource map). A linguistic scale was designed to assign fuzzy membership values to landslide susceptible zones of LSZ map (Kanungo *et al.* 2008). Similarly, a linguistic scale was designed to allocate fuzzy membership values to the categories of resource map. These fuzzy membership values were based on the opinion of experts. The maps with linguistic scales were named as landslide potential layer and resource damage potential layer, respectively. These two layers were then integrated *via* a multiplicative function to produce landslide risk assessment matrix (LRAM). The elements of LRAM were interpreted and categorized into five risk zones to produce an LRA map. The higher the value of an element in LRAM, higher is the risk.

4   SALIENT FEATURES OF LASARIZ SOFTWARE PACKAGE

A user-friendly software, acronymed as "**La**ndslide **S**usceptibility **a**nd **Ri**sk **Z**onation (LaSaRiZ)", for landslide susceptibility zonation and risk assessment is written in Visual C++ (VC++), an object-oriented programming (OOP) language, which enables the programmers to create modules that do not need to be changed when a new type of application is added. A programmer can simply create a new application that inherits many of its features from existing applications. This introduces flexibility in programming. In addition, VC++ is also very useful for coding engineering and scientific problems.

The LaSaRiZ software has three basic modules; display module, landslide susceptibility zonation module, landslide risk assessment module.

4.1 Display module

In this module, the images (both input and output) in ASCII (i.e. text files) or any image file format (i.e. Erdas Imagine (.img), Bitmap (.bmp), Graphic Interchange File (.gif) and JPEG files (.jpg and .jpeg)) can be opened and displayed.

4.2 Landslide Susceptibility Zonation module

LSZ module is designed to perform landslide susceptibility zonation based on the following approaches, which are implemented as three separate modules; back-propagation neural network black box approach, fuzzy set based cosine amplitude approach, combined neural network and fuzzy relation approach

Back-propagation Neural Network (BPNN) module

The network architectural parameters (i.e., number of units in input layer, hidden layers and output layer) are given as input through a dialog box, as shown in **Fig. 1a**, which is the main dialog box for this module. One additional dialog box has been created for inputting of training parameters. This dialog box pops-up on clicking Training File button and allows the user to key-in the training parameters (**Fig. 1b**). On clicking, the event of button is associated with its definition and starts the execution.

Fuzzy relation module

The main dialog box of this module is shown in **Fig. 2.** Image file and the existing landslide distribution data layer are given as inputs to the module.

Combined Neural Network and Fuzzy module

The main dialog box of this approach is shown in **Fig. 3**. Files containing matrices of connection weights (input-hidden, hidden-hidden and hidden-output layers), as obtained from BPNN module, are directly input here.

4.3 Landslide Risk Assessment Module

Two approaches for landslide risk assessment as implemented in LaSaRiZ software are; danger pixel approach, fuzzy set based approach.

Danger pixel module for LRA

The main dialog box for this module is shown in **Fig. 4**. After input files are provided to this module, the number of LSZ maps and their categories and the number of resource category are input through dialog box shown in **Fig. 5**.

Fuzzy linguistic scale module for LRA

The main dialog box for this module is shown in **Fig. 6**. On clicking the Input file button, two more dialog boxes are popped up, as shown in **Fig. 7**. Using these dialog boxes, the user can input the membership values to be assigned to each susceptibility zone and each resource category.

5 AN EXAMPLE ON USAGE OF LaSaRiZ FOR LSZ AND LRA IN HIMALAYAS

The study area belongs to parts of Chamoli and Rudraprayag districts of the State of Uttrakhand, India, in the Himalayan region and covers about 600 sq km. Spatial data pertaining to various causative factors were collected from satellite remote sensing images (IRS-1C and P6 satellite sensors - PAN, LISS-III and LISS IV), Survey of India (SOI) topographic maps, Valdiya's geological map and the field campaigns. **Table 1** provides the details of these data along with their usage in the study.

The spatial data provided in **Table 1** were appropriately processed and analyzed in an image processing and GIS software to prepare seven thematic data layers. Further details on the study area, causative factors and their justification, and the process of preparation of thematic data layers can be found in Chauhan *et al.* 2010ab. The seven thematic data layers were named as slope, aspect, relative relief, structural buffer, lithology, drainage density and landuse landcover. These layers were stacked as a single database in ASCII format to be input to LaSaRiZ for various operations.

The identification and mapping of existing landslides is a pre-requisite to develop any data driven model for LSZ. Therefore, existing landslide locations were interpreted visually and mapped from high-resolution LISS IV (MX) image and PAN sharpened multi-spectral image. A total of 154 landslides of varying dimensions were mapped, which were subsequently digitized and rasterized to create a landslide distribution data layer. From this landslide distribution data layer, the pixels of landslide and no-landslide attributes were extracted, which were then used for training and testing of the three approaches for LSZ mapping implemented in LaSaRiZ software.

6 RESULTS AND DISCUSSION

The LSZ maps produced from the three approaches through LaSaRiZ software were evaluated with each other in respect of the distribution of existing landslides in the area, i.e. on the basis of landslide density, and then through ROC curves. Landslide density was defined as the ratio of the existing landslide area (in percent) obtained from landslide distribution layer to the area of each landslide susceptibility zone (in percent) obtained from an LSZ map. The distribution of landslide susceptibility zones and the landslide densities for all the three approaches are given in **Table 2**. It can be seen that in case of back-propagation neural network approach, 72% of the observed landslides fall in 39% of the total area categorized into very high and high susceptibility zone. Also, a very large area of about 29% is obtained as very high susceptibility zone in neural network approach, which does not show any defined pattern and is found to be distributed overall in the map. In case of the fuzzy relation based approach, 64.65% of observed landslides fall in 24.9% of identified very high and high

susceptibility zones. However, in case of combined neural network and fuzzy approach, 74.5% of observed landslides fall in 29.0% of identified very high and high susceptibility zones, which in fact should be the case (i.e., areas belonging to very high and high susceptibility zones have been further narrowed down). This outcome can also be corroborated from the study of landslide density values. Usually, an ideal LSZ map should have the highest landslide density for VHS zone, as compared to other zones and there ought to be a decreasing trend of landslide density values successively from VHS to VLS zone.

To assess the performance of all the three LSZ approaches, the ROC curves were generated. The ROC curves for the LSZ maps from the ANN black box (AUC of 0.84), fuzzy relation based (AUC of 0.86) and combined neural network and fuzzy (AUC of 0.92) are shown in **Fig. 8**. These curves clearly depict that the combined neural network and fuzzy based approach model is the most successful one in predicting the probability of landslide susceptibility for the study area, since the AUC value for the LSZ map derived from this approach is higher than that obtained from other two approaches. These results, thus, sufficiently demonstrate that the combined neural network and fuzzy approach portrays actual scenario of landslide occurrences in a given region.

The risk due to landslides in the Himalaya was ascertained from the LRA maps. The LRA map based on danger pixel approach and obtained from the LaSaRiZ software can be used to assess the category-wise risk. This map shows spatial distribution those resource categories, which are under risk. The figure suggests that the categories settlement, agriculture, sparse vegetation and road are under risk due to landslides. It can further be observed that settlement areas around Chamoli, a portion of road from Nandprayag to Chamoli, the agricultural land and sparsely vegetated areas surrounding Chamoli town are also under risk due to landslides in this region.

The LaSaRiZ derived LRA map based on fuzzy linguistic scale approach, when superimposed on resource category map, depicts spatial distribution of different risk zones under various resource categories. This LRA map thus provides additional information on further categorizing a resource category into very high to very low risk zones. It can be observed from that 0.54% of total area is under very high risk which comprises of settlement and road categories, 0.73% lies in high risk zone, MR zone occupies 1.83%, an area of 18.67% is represented by LR zone and a large area of 78.23% is represented by VLR zone. Thus, this approach is able to identify that there is a small percentage of areas, which are in very high and high risk zones due to landslides.

7   CONCLUSION

In this paper, a study on landslide susceptibility zonation and risk assessment through an indigenously developed software acronymmed as LaSaRiZ was presented. The GUI based software is quite user-friendly and has been developed exclusively for the stated purpose. The inputs to this software are a set of thematic data layers corresponding to the factors responsible for the occurrence of the landslides in a region and the resource categories under risk due to landslides. These data layers, created in any available image processing and GIS software, were assumed to be available with the user beforehand in the specified data format as per the requirements of LaSaRiZ. As an example, the working of LaSaRiZ was demonstrated through a case study on LSZ and LRA in a landslide prone area in the Himalayan region. The LSZ map produced by the combined neural network and fuzzy approach showed systematic and a decreasing trend of variation in landslide density values from VHS to VLS zones in the region. Thus, for the study area considered, the combined neural network and fuzzy approach provided an accurate representation of the actual scenario of landslide occurrences in the region. Further, the two LRA approaches were found to be effective in explicitly identifying the resource categories under risk and also defining zones of risks due to landslides in those categories in the region. It is expected that LaSaRiZ software will be quite useful to managers and planners involved in landslide mitigation programs for various infrastructural development activities in a region.

Table 1: Data sources and specific use

Data types	Description			Specific Use
Remote Sensing Data	Satellite/Sensor	Spatial Resolution	Year of Acquisition	Land use/ Land cover, Structural Features Landslide distribution
	IRS PAN	5.8m	1999	
	IRS LISS-III	23.5m	2001	
	IRS LISS-IV	5.8m	2008	
Google Earth Data	Various images of study area Google Earth Data			Landslide distribution
Topographical maps	Scale: 1:50000			DEM: Slope, Aspect, Relative Relief, Drainage network
Geological map	Scale: 1:326000			Lithology, Structural Features
Field Data	GPS Surveys			Landslide distribution, Land use/Land cover

Table 2: Landslide distribution in various landslide susceptible zones

Landslide Susceptible Zones	ANN black box approach			Fuzzy set based approach			Combined neural network and fuzzy approach		
	% area of identified zones (a)	% area of observed and slides per class (b)	Landslide density (b/a)	% area of identified zones (a)	% area of observed landslides per class (b)	Landslide density (b/a)	% area of identified zones (a)	% area of observed landslides per class (b)	Landslide density (b/a)
VHS	11.60	29.3	2.52	4.0	24.33	6.08	5.1	33.1	6.49
HS	27.40	42.7	1.56	20.9	40.32	1.93	23.9	41.4	1.73
MS	46.15	27.4	0.59	41.1	27.62	0.67	41.4	21.9	0.53
LS	8.80	0.26	0.29	29.3	7.13	0.24	25.9	3.6	0.14
VLS	6.05	0.34	0.05	4.7	0.6	0.13	3.7	0	0

Fig. 1: (a) Main dialog box

Fig. 1(b) Sub dialog box for training parameters
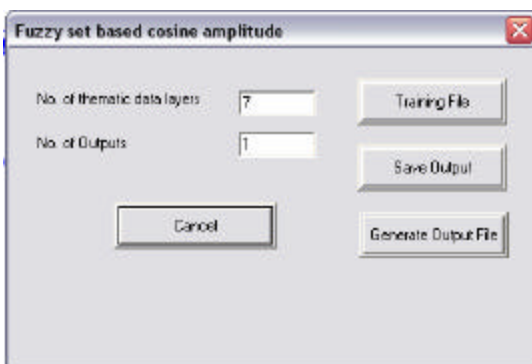
Fig. 2: Main dialog box of fuzzy set based cosine amplitude module
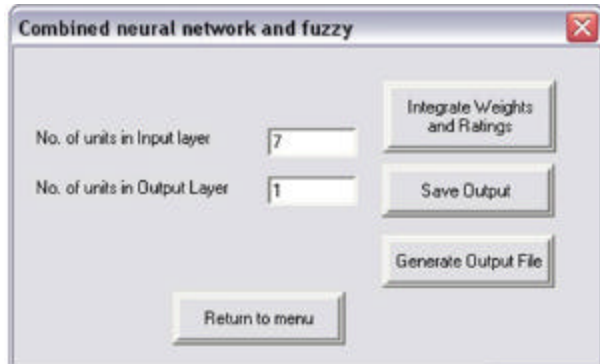

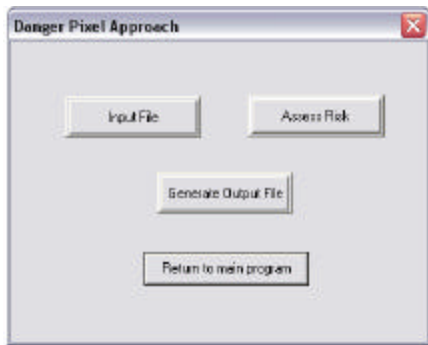**Fig. 3:** Main dialog box of combined neural network and fuzzy module
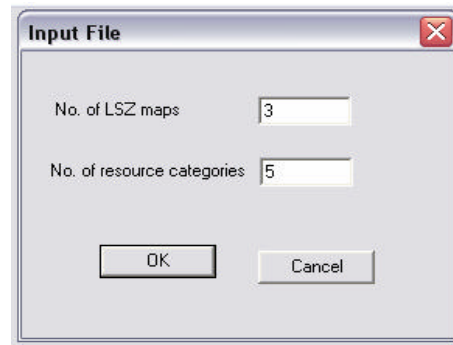
Fig. 4 Main dialog box of danger pixel module

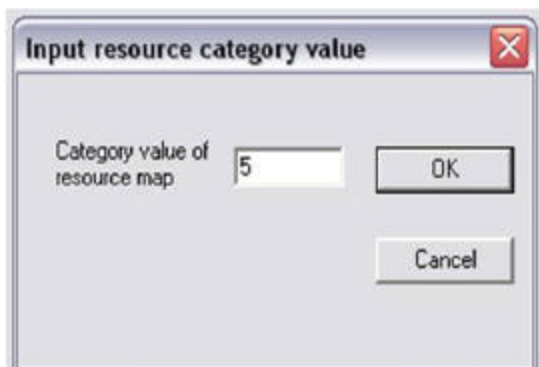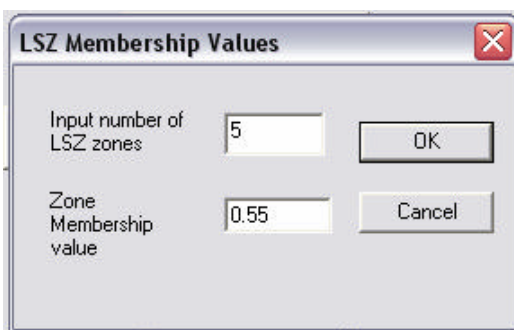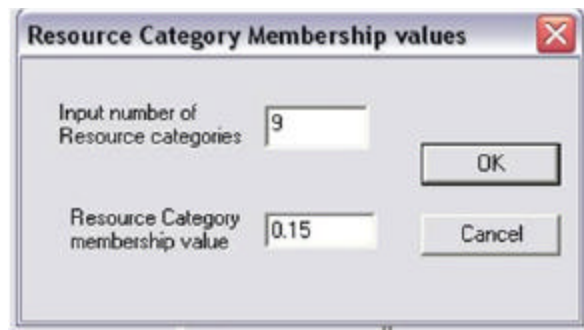**Fig. 5(a)** Sub dialog boxes of danger pixel module

Fig. 5 (b) Main dialog box of danger pixel module

Fig. 6 Main dialog box of Fuzzy Linguistic module

(a)

(b)

Fig. 7 Dialog boxes to enter membership values in Fuzzy Linguistic module
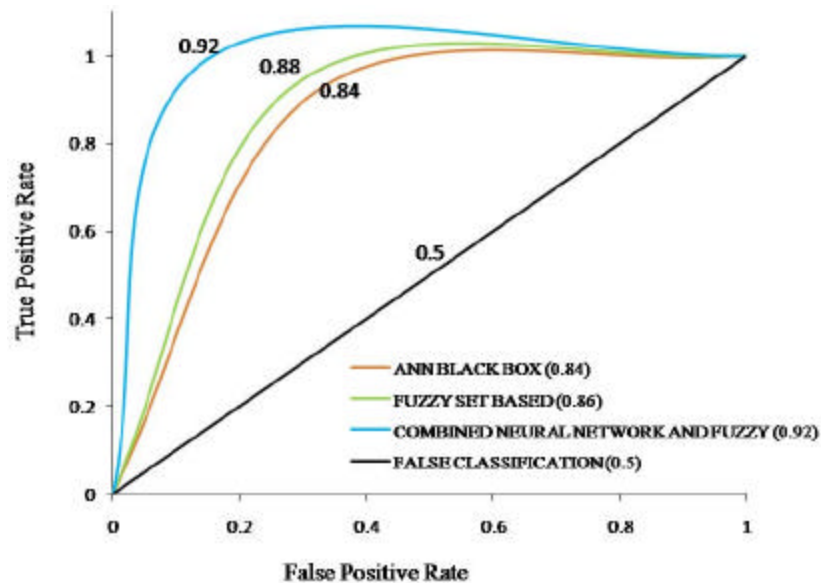
Fig. 8 ROC curves for LSZ maps produced by ANN black box, fuzzy set based and combined neural network and fuzzy approach

REFERENCES

[1]   Arora, M.K., Das Gupta, A.S., Gupta, R.P. (2004) An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. *International Journal of Remote Sensing*, 25, 559–572.

[2]   Chauhan, S., Arora, M. K., and Gupta, N. K. (2010a) Landslide Susceptibility Zonation through ratings derived from Artificial Neural Network, *International Journal of Applied Earth Observation and Geo Information*, 12, 340-350.

[3]   Chauhan, S., Sharma, M and Arora, M. K. (2010b) Landslide Susceptibility Zonation of Chamoli region, Garhwal Himalayas,using Logistic Regression Model, *Landslides*, Published Online.

[4]   Dubois, D., Prade, H. (1980) *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York.

[5]   Kanungo DP, Arora MK, Gupta RP, Sarkar S. (2008) Landslide risk assessment using concepts of danger pixels and fuzzy set theory in Darjeeling Himalayas. *Landslides*, 5, 407–416.

[6]   Kanungo DP, Arora MK, Sarkar S, Gupta RP (2006) A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Engineering Geology*, 85, 347–366

[7]   Olden, J.D., Joy, M.K., Death, R.G. (2004) An accyrate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modeling*, 178, 389–397.

[8]   Ross, T.J., (1995) *Fuzzy Logic with Engineering Applications*. McGraw-Hill, New York.

[9]   Saha, A.K., Gupta, R.P., Arora, M.K. (2002) GIS-based landslide hazard zonation in a part of the Himalayas. *International Journal of Remote Sensing*, 23, 357–369.

[10] Saha, A.K., Gupta, R.P., Sarkar, I., Arora, M.K., Csaplovics, E. (2005) An approach for GIS-based statistical landslide susceptibility zonation—with a case study in the Himalayas. *Landslides* 2, 61–69.

[11] Sarkar, S., Kanungo, D.P. (2004) An integrated approach for landslide susceptibility mapping using remote sensing and GIS. *Photogrammetric Engineering and Remote Sensing*, 70, 617–625.