We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Geo-spatial Technology for Landslide Hazard Zonation and Prediction

Dericks P. Shukla, Sharad Gupta, Chandra S. Dubey and Manoj Thakur

Additional information is available at the end of the chapter

Abstract

Similar to other geo hazards, landslides cannot be avoided in mountainous terrain. It is the most common natural hazard in the mountain regions and can result in enormous damage to both property and life every year. Better understanding of the hazard will help people to live in harmony with the pristine nature. Since India has 15% of its land area prone to landslides, preparation of landslide susceptibility zonation (LSZ) maps for these areas is of utmost importance. These susceptibility zonation maps will give the areas that are prone to landslides and the safe areas, which in-turn help the administrators for safer planning and future development activities. There are various methods for the preparation of LSZ maps such as based on Fuzzy logic, Artificial Neural Network, Discriminant Analysis, Direct Mapping, Regression Analysis, Neuro-Fuzzy approach and other techniques. These different approaches apply different rating system and the weights, which are area and factors dependent. Therefore, these weights and ratings play a vital role in the preparation of susceptibility maps using any of the approach. However, one technique that gives very high accuracy in certain might not be applicable to other parts of the world due to change in various factors, weights and ratings. Hence, only one method cannot be suggested to be applied in any other terrain. Therefore, an understanding of these approaches, factors and weights needs to be enhanced so that their execution in Geographic Information System (GIS) environment could give better results and yield actual ground like scenarios for landslide susceptibility mapping. Hence, the available and applicable approaches are discussed in this chapter along with detailed account of the literature survey in the areas of LSZ mapping. Also a case study of Garhwal area where Support Vector Machine (SVM) technique is used for preparing LSZ is also given. These LSZ maps will also be an important input for preparing the risk assessment of LSZ.

Keywords: Landslide, LSZ, Remote Sensing and Geographic Information System, Modeling, SVM, Garhwal Himalaya

1. Introduction

According to the information on the International Red Cross, there are roughly 200 major natural disasters that occur each year in the world. These natural disasters cause an annual average loss of nearly 130,000 persons, and more than 140 million normal lives are affected. The frequency of occurrences of these natural disasters has increased many times in the recent past, and its effects are becoming more severe in the coming years. The major attribute is being the population growth, urbanization/industrialization leading to climate change. In general, most of the "natural risks" are accentuated by humans themselves by direct or indirect interference with the nature. Understanding a natural disaster is very difficult as it is a very complex system that involves various controlling and contributing factors. This means that no easy, one-sided solutions can be found, but applying the holistic approach to tackle such problems could yield some beneficial results. Currently, many researches are being carried out to understand the phenomenon acting behind these natural disasters such as floods, tsunamis, cyclones, earthquakes, landslides, etc. So to combat these natural risks, the holistic concepts should be developed and applied, particularly to tackle landslide risk as landslides are one of the major environmental problems in our society.

The adverse impacts of climate change on developing countries have been highly consequential. High-magnitude flash floods and increased rains has been one of the pertinent causes of extensive landslides, which accounts for around 4.89% of the globally occurring natural disasters during the last two decades. The unplanned urbanization and development coupled with continued deforestation may be attributed to this rise in figure. Landslides are quite frequent along the tectonically active Himalayan region. In the year 1984, Varnes defined the term **hazard** as "the probability of occurrence of a potentially damaging phenomenon within a specified period of time and within a given area". When such spatial distributions of hazards are represented on maps into various classes, it gives zonation maps. Thus, landslide hazard zonation refers to the division of area into various classes, which is categorised on the basis of degrees of actual/potential hazard caused by landslides. Hence, hazard zonation forms a critical factor for effective landslide management and is used as a tool for planning mitigation measures. The preparation of the landslide hazard map requires the analysis of most determining factors that leads to soil failure. The preparation of landslide hazard zonation requires detailed landslide inventory, processes involved in slope instability, triggering factors and many other associated studies. Landslides may occur due to a variety of conditional and triggering factors such as change in slope angle, slope aspect, faults, lithology, deforestation, improper drainage system, rainfall, and earthquakes. Thus, this zonation can be carried out at various scales from national (1:1 million) to local (1:5000). Depending on the scale of map, the parameters/factors and their accuracy varies.

With the advent of satellite data and various sensors, the scope of remote sensing has increased widely. The bird's eye view of the area at moderate to fine resolution gives fast and quick information about the terrain. Clubbed with the spectral and temporal characteristics of the satellite, the ability to identify and recognise landslides for the preparation of inventory map has been improved a lot. Both visual as well as automatic processes are well developed for

recognition of landslide features. This preparation of inventory map has been made more effective with recent developments of resolution merging where data from different sensors could be merged to obtain better, sharp and good resolution images. Not only in the identification of landslides but also in the preparation of other contributing and controlling factors, remote sensing plays a crucial role. The elevation data from DEM (Digital Elevation Model) are used for the preparation of slope, aspect, relief, curvature, etc., parameter that controls the behaviour of landslide as well as the slope stability/instability. Not only the optical and multispectral data but the Radar and SAR data are being used for the analysis of landslides. The interferometric SAR technique is capable of distinguishing very minute changes in elevation and slope; hence, it is used for the identification of higher-resolution and correspondingly smaller area. Data from various sensors, i.e. optical, multispectral, thermal and microwave/radar, are being used for landslide studies.

There are various methods for the preparation of Landslide Susceptibility Zonation (LSZ) such as based on Fuzzy logic, Artificial Neural Network, Discriminant Analysis, Direct Mapping, Regression Analysis, Neuro-Fuzzy approach and other techniques. These different approaches apply different rating system and the weights, which are area and factors dependent. Therefore, these weights and ratings play a vital role in the preparation of susceptibility maps using any of the approach. However, one technique that gives very high accuracy in certain might not be applicable to other parts of the world due to change in various factors, weights and ratings. Hence, only one method cannot be suggested to be applied in any other terrain. This chapter discusses the methods being used in the field of LSZ, what are the input parameters being used, what the accuracy is and how best the method map the LSZ. However, it should be kept in mind that most of these methods/analysis are based on landslide inventory of any area, so the first and foremost step for working towards LSZ should be preparation of landslide inventory. Finally, this chapter discusses a case study of application of geo-spatial technology for preparation of LSZ in Garhwal Himalayan region, which is tectonically very active and prone to landsliding.

2. Various Approaches for LSZ Mapping

2.1. Regression Analysis

People are normally interested in finding the relationship between different variables. For example, whether smoking causes lung cancer? Regression analysis is the statistical method of finding relationship between dependent/predicted variable (denoted as y) and independent/predictor variables (denoted as $x_1, x_2, \ldots, x_n$), where n denotes the number of predictor variables [1]. The true relationship between y and $x_1, x_2, \ldots, x_n$ can be approximated by the regression model as indicated in equation 1

$$y = f\left(x_1, x_2, \ldots, x_n\right) + \varepsilon \tag{1}$$

where ε is assumed to be a random error representing the discrepancy in the approximation. It accounts for the failure of the model to fit the data exactly [2]. Typically, regression analysis is used for one of these three purposes [3] viz. (i) Modelling the relationship between x and y, (ii) Prediction of target variable, and (iii) Testing of hypotheses.

There are three types of regression models:

Simple linear regression: It models the linear relationship between two variables; out of which one is dependent variable y, and other is independent variable x. In this model, regression equation is given as below in equation 2

$$y = ax + b + \varepsilon \tag{2}$$

where a = slope of regression line, b = intercept and ε = random error. Simple linear regression is shown in Figure.1.

Multiple linear regression: There are many situations when result depends on one or more predictor variables. In such situations, simple linear regression is not sufficient to model the output, hence it requires a regression equation as given in eq 3, which models the linear relationship between one dependent variable y and more than one independent variables $x_1, x_2, \ldots, x_n$. In this model, regression equation is given as below

$$y = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n + a_0 + \varepsilon \tag{3}$$

where $a_1, a_2, \ldots, a_n$ are regression coefficients, a_0 = intercept and ε = random error

After the determination of regression model, its parameters are estimated based on the collected data. This is called as parameter estimation and model fitting. Most commonly used method of estimation is called the least square method [1, 2, 3].

Nonlinear regression: When the relationship between dependent and independent variable cannot be modelled using straight line, nonlinear regression is used. For example, nonlinear regression model for growth of a particular organism (y) as a function of time (t) can be written as

$$y = \frac{\alpha}{1 + e^{\beta t}} + \varepsilon \tag{4}$$

where α and β are model parameters and ε = random error. All nonlinear functions that can be transformed into linear functions are called linearizable functions [2, 3].
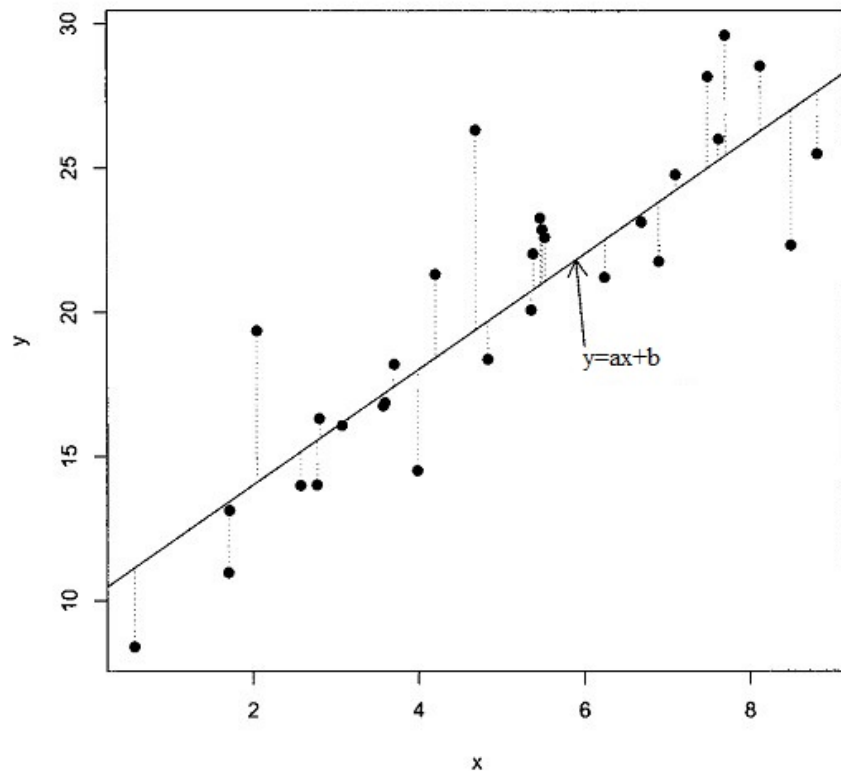
Figure 1. Simple linear regression model, solid line corresponds to true regression line and the dotted line corresponds to random error ε [3].

2.1.1. Estimation Using Least Square

The least square method for linear regression finds regression coefficients $a_0, a_1, a_2, \ldots, a_n$ such that sum of squared distance from actual value y_i and fitted value $\hat{y}_i$ reaches minimum for all possible choices of regression coefficients $a_0, a_1, a_2, \ldots, a_n$, [1, 4] using the given eq 5.

$$\sum_{i=1}^{n}\left[y_i - \left(a_0 + a_1x_1 + a_2x_2 + \ldots + a_ix_i\right)\right]^2 \tag{5}$$

For any choice of observed coefficients $\hat{a}$, the estimated/fitted value given for the observed values is

$$\widehat{y_i} = \widehat{a_0} + \widehat{a_1}x_1 + \widehat{a_2}x_2 + \ldots + \widehat{a_i}x_i \tag{6}$$

The difference between observed value y_i and fitted value $\hat{y}_i$ is called residual.

When dealing with regression analysis, if there is only one response variable, regression analysis is called univariate regression, and in case of two or more response variables, the regression is called multivariate regression. The difference between simple and multiple regressions is determined by the number of predictor variables (i.e. simple means one predictor variable and multiple means two or more predictor variables), whereas the difference between univariate and multivariate regressions is determined by the number of response variables. A brief summary of various classifications is given in Table-1. Out of all these regression types, logistic regression method is used a lot since most variables in hazard zonation mapping tends to be qualitative rather than quantitative.

Types of Regression	Conditions
Univariate	Only one quantitative response variable
Multivariate	Two or more quantitative response variables
Simple	Only one predictor variable
Multiple	Two or more predictor variables
Linear	All parameters enter the equation linearly, possibly after transformation of the data
Nonlinear	The relationship between the response and some of the predictors is nonlinear or some of the parameters appear nonlinearly, but no transformation is possible to make the parameters appear linearly
Analysis of Variance	All predictors are qualitative variables
Analysis of Covariance	Some predictors are quantitative variables and others are qualitative variables
Logistic	The response variable is qualitative

Table 1. Various Classifications of Regression Analysis [2].

2.1.2. Logistic Regression

Logistic regression model is a general linear model, which models the data with binary responses [1], i.e. it predicts the presence or absence of an outcome based on the values of a set of predictor variables [5]. The dependent variable in logistic regression is binary (i.e. 0 or 1, true or false), whereas the independent variable can be categorical, dichotomous or interval [6]. For landslide study, dependent variable is binary, showing either the presence or the absence of landslide.

Example: For determining risk factor for cancer, health data of several people were collected on several variables such as age, sex, smoking, diet, and the family's medical history. The response variable "y" is the person having cancer (y=1) or not having cancer (y=0) [2].

Coefficients of logistic regression can be used to calculate ratios for each independent variable in the model. Logistic regression model can be represented in simplest form as shown in equation 7

$$p = \frac{1}{1 + e^{-y}} \tag{7}$$

where p is the probability of occurrence of an event (varies between 0 and 1 on S-shaped curve), and y is dependent variable and calculated using the logistic regression equation 8

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \tag{8}$$

where a_1, a_2, ..., a_n are logistic regression coefficients and a_0 = intercept, x_1, x_2, ..., x_n are independent variables [7].

2.1.3. Applications [2, 4]

i.   Agricultural sciences (e.g. analysis of data of milk production).

ii.   Management, industrial and labour relations (e.g. Do chief executive officers (CEOs) and their top managers always agree on the goals of the company?).

iii.   Environmental sciences (e.g. exploration of relationship between water quality and land use).

iv.   Psychology (e.g. What are the factors that impact the likelihood of a moonlighting worker becoming aggressive toward his or her supervisor?).

v.   Geography (Can the population of an urban area be estimated without taking a census?).

2.1.4. Landslide Hazard Zonation using Regression Analysis

Regression analysis is one of the most widely used statistical tool as it provides simple methods for establishing a functional relationship among variables. Logistic regression has been used widely for preparation of landslide hazard zonation maps [5, 6, 8, 9]. Slope, aspect, curvature, distance from drainage, lithology, distance from lineaments, land cover, vegetation index, and precipitation are considered as landslide-causing factors in many literatures. In logistic regression model, LHI is calculated by solving the regression equation. Correlation between landslide event and landslide affecting factors is estimated, and then, equation predicting the landslide is obtained.

2.2. Analytic Hierarchy Process

AHP, developed by Thomas L. Saaty in 1975, is an effective tool for decision making. It helps the decision makers in setting priorities and making best decision on complex decisive problems. It distributes the problems in hierarchy of criteria and options (alternatives), i.e. it reduces complex decisions to pairwise comparisons and then synthesizes the result. The AHP considers both the rational and the intuitive to select the best from a number of alternatives

evaluated with respect to several criteria. It checks for consistencies in decision maker's evaluation and also allows limited inconsistencies in judgements.

2.2.1. Working of AHP

The AHP uses a set of evaluation criteria and a set of alternative options among which the best decision is to be made. It generates a weight for each evaluation criteria according to pairwise comparisons of criteria. The criteria with higher weight are selected since it is most important of all the criteria. Further, for fixed criteria, it assigns a score to each alternative option according to pairwise comparisons of options based on those criteria. Higher the score for an option, better the performance of that option w.r.t. considered criteria. Information is then arranged in a hierarchical tree. Finally, the AHP generates global score for each option using the combinations of the criteria weights and options scores and determines relative ranking of alternatives. A simple hierarchy with three levels is shown in Figure.2.
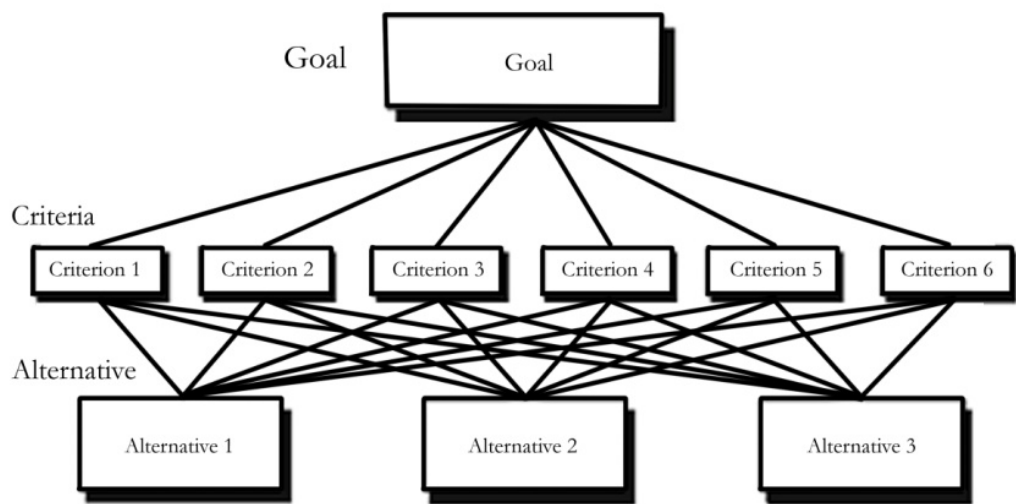

Figure 2. A three level hierarchy [10].

Implementation of AHP

AHP can be implemented in three simple steps

i.   Computation of weight vector for all criteria

ii.   Computation of score matrix for all options

iii.   Ranking of options based on final score

Once the goal has been set, then for all the alternatives, different ranks are given based on the criterion fixed to reach that goal. In this way, the priorities are set, and these factors are compared pairwise. For example, in case of landslide zonation, the goal could be to identify the areas that are prone to landsliding and the factors/parameters, such as slope, elevation, soil type, rock type, distance to drainage, etc., controlling it would become the alternatives. And

to select the areas prone to landsliding, the criteria could be fixed such as slope should be more than 45º, soil type should be clayey, rock type should be other than granite/gneiss (hard rock), etc. Hence, the area fulfilling these criteria will be selected. This way of preparing the landslide susceptibility map is area specific, and the criteria applicable to one location may not be true for other location. Hence, a different approach is needed where the system adjust itself with the given conditions and scenarios.

2.2.2. The Fundamental Scale

The AHP is a general theory of measurement and is used to derive relative priorities of different criteria on absolute scales. Pairwise comparison judgments in the AHP are applied to pairs of homogeneous elements. The fundamental scale represents the intensities of judgments. In many cases, the elements to be compared are almost equal in measurements. In this situation, comparison must be made not on what fraction it is larger than the other [10]. Pairwise comparisons of criteria and/or options are performed based on the scale given in Table-2.

Intensity of Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
2	Weak	
3	Moderate importance	Experience and judgment slightly favour one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgment strongly favours one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance is demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favouring one activity over another is of the highest possible order of affirmation

Table 2. The fundamental scale by T. L. Saaty [10, 11].

2.2.3. Applications of AHP [10, 12, 13].

i.   Evaluation of cities for livelihood and planning

ii.   Ranking of countries

iii.   Customers adoption of mobile devices and mobile services

iv.   Human organ transplants

v.   Prediction of winners in chess matches

vi.   Natural resource management

2.2.4. Landslide Hazard Zonation using AHP

Various authors [14, 15, 16, 17, 18, 19] have used AHP for giving weights to various factors of landslide occurrence. The effect of each factor and factor classes, on landslide occurrence, is determined using pairwise comparison, and an equation is modelled for landside susceptible index (LSI), as given below in equation 9

$$LSI_{AHP} = \sum_{i=1}^{n} Factor_i * W_{AHP_i} \qquad (9)$$

where $Factor_i$ = landslide conditioning factor such as slope, aspect, lithology, etc. W_{AHP_i} = Weightage for each causative factor. Pixel (LSI) values derived from above equation are classified into various susceptibility classes (low, moderate, high, and very high) based on natural break.

2.3. Artificial Neural Network

Artificial neural network attempts to model the information processing capabilities of the brain. The operation of the brain is based on simple basic elements called as neurons. Neurons are connected to each other with transmission lines called as axons and receptive lines called as dendrites. Information is stored at synapses. Each neuron has an activation level that ranges between some minimum and maximum value [20, 21]. A neural network is a massively parallel distributed processor made from simple processing units, which can store knowledge gained from experiments and can utilize it later. It replicates the processing of the brain in two respects [22].

i.   Knowledge is acquired by the network from its environment through a learning process.

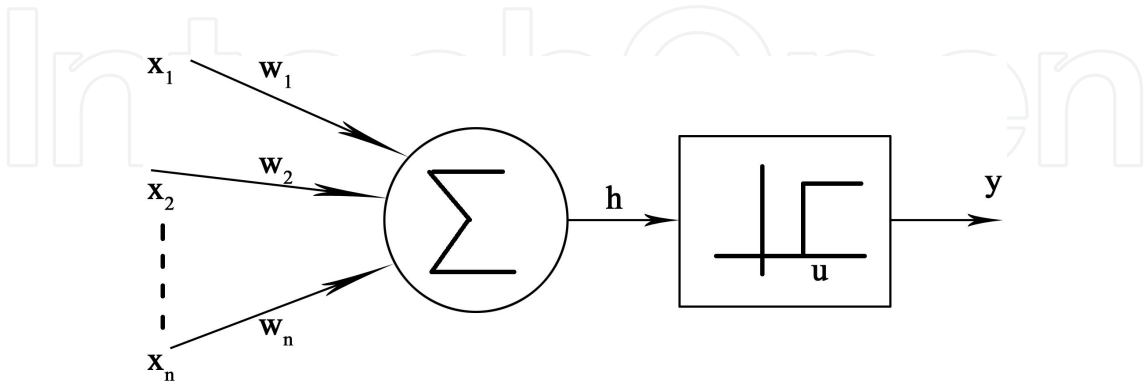ii.   Synaptic weights are used to store the acquired knowledge.

Figure 3. McCulloch and Pitts model of artificial neuron [20].

In 1943, McCulloch and Pitts proposed a computational model for artificial neuron, based on binary threshold [23]. This neuron calculates a weighted sum of 'n' input signals, x_j where j =

1, 2, 3…...n, and generates an output of 1 if this sum is above a certain threshold 'u', else output 0. The model [24] is shown in Figure. 3 and given by equation 10.

$$y = 1, if \ \sum_{j=1}^{n} w_j x_j > u,$$ (10)

$$y = 0, otherwise$$

ANN is a weighted directed graph, in which artificial neurons are nodes and directed edges with weights are connections between neuron outputs and neuron inputs. ANN can be grouped in two categories [20, 22].

i.   Feed-forward network, where graph has no loops, as shown in Figure. 4. Here, all the nodes in each layers are connected to every other node in forward layer, hence it is called fully connected network. If some of the links are missing, then it is called partially connected network. Example: single-layer perceptron, multilayer percep-tron, radial basis function, etc.

ii.   Recurrent or feedback network, where graph has loops because of feedback connec-tions, as shown in Figure. 5. Here, output from all the neurons is applied to input using feedback connection. Example: self-organizing map, adaptive resonance theory model, Hopfield network, etc.
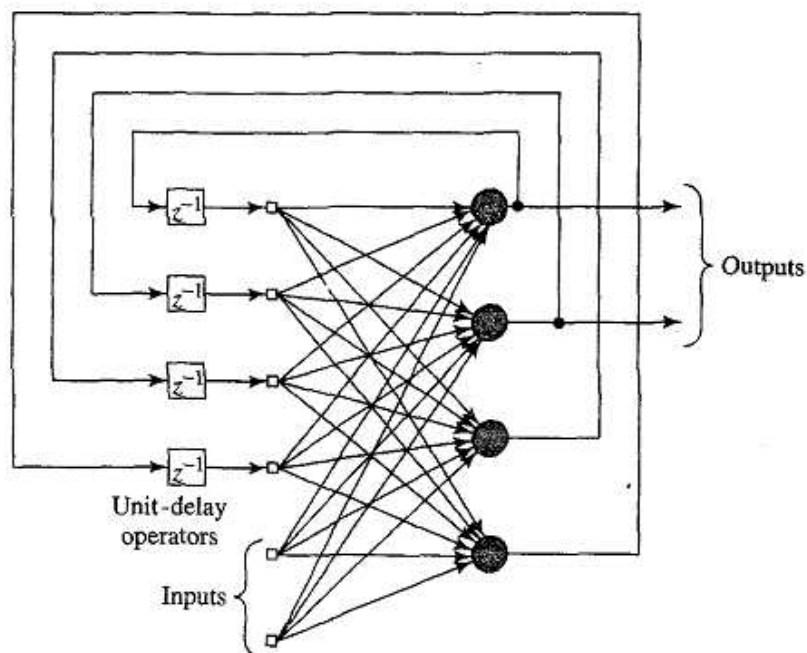

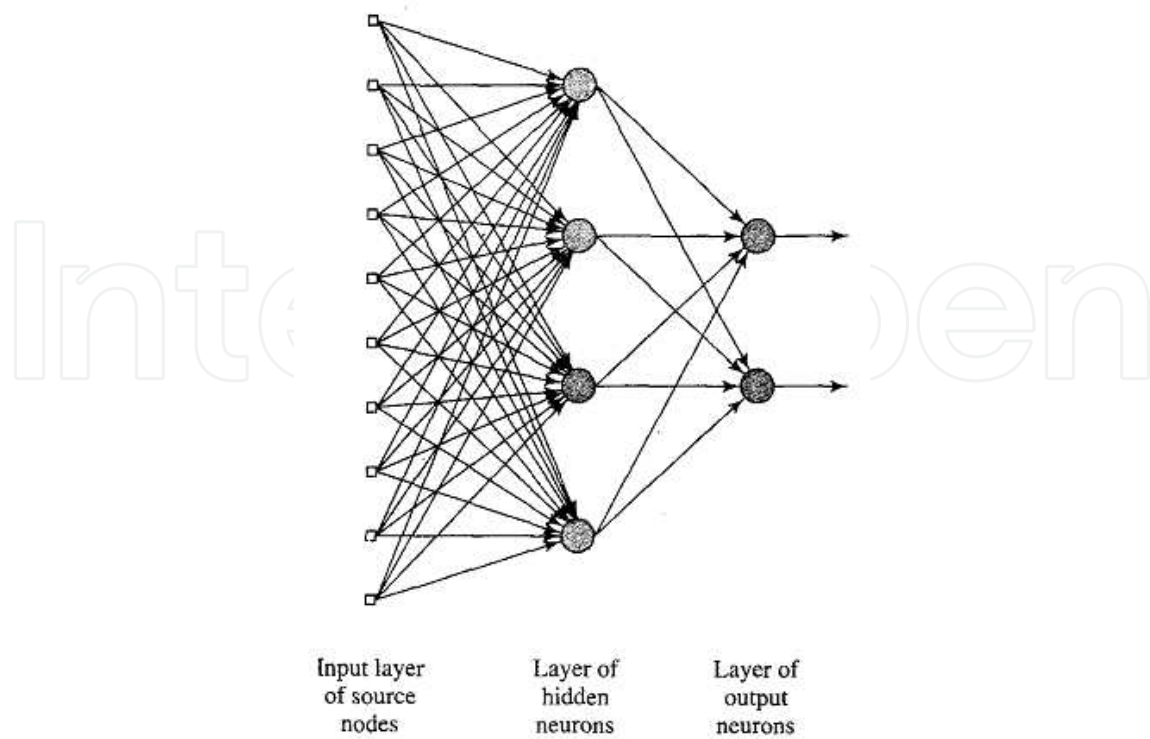**Figure 4.** Example of feed forward network with one hidden layer & one output Layer

Input layer   Layer of   Layer of
of source   hidden   output
nodes   neurons   neurons

Figure 5. Example of recurrent network with hidden layer

2.3.1. Learning Algorithms

To be able to learn is the fundamental trait of intelligence. Although it is difficult to formulate a precise definition of learning, the process of learning in the context of ANN can be defined as the problem of updating network architecture and connection weights so that a network can efficiently perform a specific task [20]. Artificial neural network tries to learn input–output relationships from the given collection of representative examples, instead of following a set of rules specified by human experts. This is one of the major advantages of neural networks over traditional expert systems. A learning algorithm refers to a procedure in which learning rules are used for adjusting the weights. Some examples of learning algorithms are (i) Error correction learning, (ii) Memory-based learning, (iii) Hebbian learning, (iv) Competitive learning, (v) Boltzmann learning, etc. [23, 25].

2.3.2. Feed-Forward Back-Propagation Network (Based on error correction learning)

It is basically a feed-forward multilayer perceptron with back-propagation as learning/training algorithm. In order to train a neural network to perform desired task, the weight of each input has to be adjusted, such that the error between the desired and actual output is minimal (Figure. 6 after [21]) i.e.

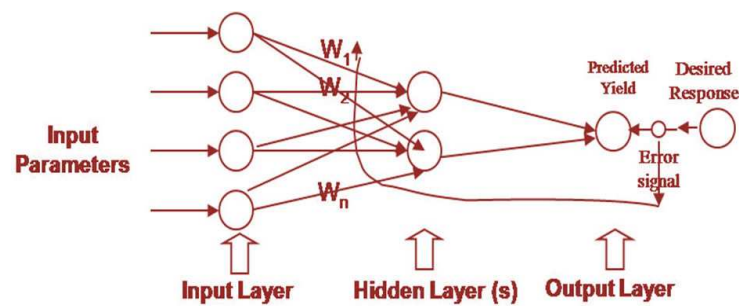$$\text{Error Signal}(e) = \text{Desired Response}(d) - \text{Actual Output}(y)$$

Figure 6. Back-Propagation Neural Network [21].

2.3.3. Applications of ANN

1. Image processing, classification of satellite data, compression of large images, etc.

2. Medical signal processing and developing medical decision support system [26, 27].

3. Paper making industry for prediction of curl in paper reel [28].

4. Calculation of nonlinear interpolation algorithm [29].

5. Detection and classification of vehicles in traffic management [30].

6. Optical and handwritten character recognition [31].

7. Operations research [32].

8. Application in Mineral Potential Mapping [33].

9. Landslide Susceptibility Mapping [34-37].

2.3.4. Application of ANN in Landslide Hazard Zonation

ANN has been used widely in the preparation of LHZ maps [34–37]. People have used variations of ANN with one input layer, two hidden layers, and one output layer for various factors controlling landslide occurrence. ANN connection weights are used to provide weights or rankings to the input data source (landslide-causative factors). Weights of factors and rankings of categories are integrated to provide LSZ map.

2.4. Support Vector Machine

Support Vector Machine is a data classification technique, developed by Vapnik in 1990. Classification process involves separating data into training and testing sets. Each element in the training set contains a corresponding target value (i.e. the class labels) and several attribute (i.e. the features of elements). The ultimate goal of SVM is to predict the target value for the test data, with only attributes of the test data given [38, 39]. Support vector machines are based on the concept of decision planes that define decision boundaries [40]. SVM finds the best hyperplane (n-dimensional plane) that separates all data points of one class from those of other

class. It uses kernel method to project linearly non-separable data to a higher dimension. The kernel can separate classes even if mean values are near to each other. A simple illustration of the method is shown in Figure.7. The data points shown are linearly separable. The maximum margin hyper plane is shown in red, and the margin between the support vectors is shown by the parallel light blue lines. The two classes do not overlap. The support vectors (patterns that are on the margin) are shown [41] as yellow circles for class 1 and triangles for class 2.
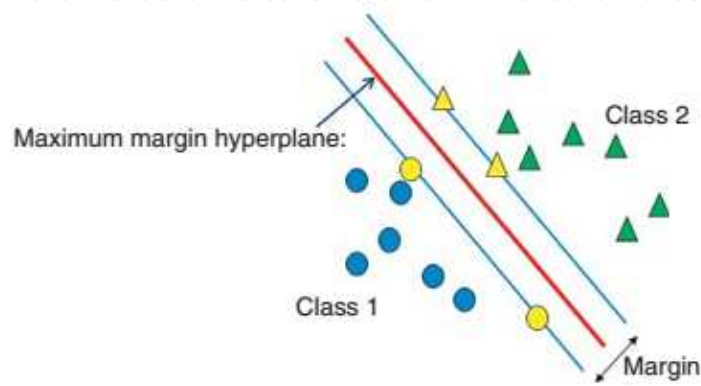

Figure 7. Illustration of the support vector [41].

Let m-dimensional training inputs x_i (i=1,...,M) belong to Class 1 or 2 and the associated labels be y_i = 1 for Class 1 and −1 for Class 2. If these data are linearly separable, we can determine the decision function, which is represented by equation 11 [42]

$$D(x) = w^T x + b \tag{11}$$

where **w** and b are weight and bias, respectively, to map the input into a higher dimensional space. The optimal separating hyper plane (i.e., $w^T x + b$ = 0) is located where the margin between the two classes is maximized, and the misclassification is minimized. The optimal hyper plane satisfies the following constrained minimization as given by equations 12–13

$$Min: \frac{1}{2} w^T w \tag{12}$$

$$w^T x_i + b \begin{cases} > 0 \ for \ y_i = 1, \\ < 0 \ for \ y_i = -1 \end{cases} \tag{13}$$

They can be obtained by solving the following constrained optimization problem by the method of Lagrange multipliers and maximizing the equation 14 as given below

$$L(w,b) = \frac{1}{2}(w.w) - \sum_{i=1}^{m} \alpha_i \left(y_i \left(w.x_i + b \right) - 1 \right) \tag{14}$$

where α_i = Lagrange's multiplier and $\alpha_i \geq 0$. SVM can perform only binary classification; however, classifying data in more than two classes can be performed using pairwise classification [42, 43].

2.4.1. Applications of Support Vector Machine

i.   Image processing for classification of satellite images [44].

ii.   Modelling of Seismic Liquefaction Potential [45].

iii.   Financial Literacy Modelling [46].

iv.   Text charactization [47].

v.   Face detection [48].

vi.   Texture classification [49].

3. Advantages and Disadvantages

All these methods mentioned above have certain advantages as well as disadvantage over the other, hence a detailed comparative Table 3 showing their advantages and disadvantages are given below.

Method	Advantages	Disadvantages
Regression Analysis	Model developer has full knowledge of variables.	It requires the data to be independent.
	It is most strongly predictive of an outcome.	It is sensitive to outliers.
	It runs faster than neural network/support vector machine-based models.	
	It is not "black box" as ANN.	
Analytic Hierarchy Process (AHP)	It is simple, flexible and powerful.	It requires a large number of comparisons.
	All the calculations are driven by decision maker's experience.	Limitation of the use of 9 point T. L. Saaty's scale.
	It does not require an expert system with the decision maker's knowledge embedded in it.	It adds extra burden on decision maker for complex problem.
		Rank reversal

Method	Advantages	Disadvantages
Artificial Neural Network (ANN)	It requires less formal statistical training to develop the network.	Neural network are "black box".
	It can implicitly detect complex nonlinear relationships.	Single-layer perceptron work only on linearly separable classification problems.
	Availability of multiple training algorithms.	It requires greater computational resources.
		It is prone to over fitting.
		Can trap in local minima.
Support Vector Machine (SVM)	It has high prediction accuracy and good mathematical foundation.	The biggest limitation of the support vector approach is the choice of the kernel.
	Overfitting does not occur.	It requires long training time.
	It does not trap in local minima, i.e. it finds the global solution.	Problem has to be formulated as two-class problem.
	It works well with fewer training samples (i.e. number of support vectors do not matter much).	
	It requires fewer parameters (kernel, error cost).	

Table 3. Advantages and Disadvantages of these methods [10, 12, 13, 50, 51]

4. Literature Survey

The literature survey of some of the available research works carried out for Landslide Susceptibility Zonation is shown in Table 4 below:

S. No.	Techniques used	Accuracy (%)	References
1.	Discriminant Analysis	83.8	Carrara et al [52]
2.	Regression Analysis	70	Jade & Sarkar [53]
3.	Logistic Regression	74.8	Guzzetti et al. [54]
4.	Multilayer Perceptron	73	Ermini et al. [55]
5.	Neuro-Fuzzy approach	97	Pradhan et al. [36]
6.	Combined Neural Network and Fuzzy	74.5	Kanungo et al. [56]

Table 4. A comparative table for various techniques uesd with their accuracy.

The results obtained showed that the Artificial Neuro Fuzzy (ANF) modeling is a very useful and powerful tool for the regional landslide susceptibility risk assessments. Various member-ship functions should be selected and a number of training sets should be carefully and

optimally selected to prevent over learning of the model. Therefore, the results that are to be obtained from the ANF modeling should be assessed carefully because the over learning may cause misleading results [35]. As a final recommendation, the results obtained from various papers showed that the methods followed in the study based on Neuro-Fuzzy approach exhibits a high performance. However, it is not forgotten that the performance of such type maps depends not only on the methodology followed but also on the quality of the available data and the factors considered for preparing LSZ. These input factors can be natural factors (like rainfall, lithology, slope, etc.) and anthropogenic factors (like road construction, mining, etc.). For this reason, if the quality of the data increases, the performance of the maps produced by these methods could increase. The detailed literature survey where various different models have been used for landslide hazard zonation is given below:

Lee and Pradhan[5] used frequency ratio and logistic regression model for mapping the landslide susceptible areas by considering slope, aspect, curvature, distance from drainage, lithology, distance from lineaments, land cover, vegetation index, and precipitation as landslide stimulating factors. They calculated the Landslide Hazard Index (LHI) by summation of frequency ratios for all the factors and solving the regression equation, respectively, for both methods and concluded that the frequency ratio model has 2.7% (93.04–90.34%) better predication accuracy than the logistic regression model.

Pradhan et al [57] combined frequency ratio and fuzzy algorithm for generating landslide hazard maps. Fuzzy membership values were calculated using frequency ratio and detected landslides. Fuzzy algebraic operators (such as fuzzy and, or, product, sum) and fuzzy gamma operators were applied on fuzzy membership values for landslide hazard mapping. Value of fuzzy gamma operator was set to 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, and 0.975 for detecting its effect on landslide hazard maps. After verification, they found that out of 17 cases tested, the gamma operator with value 0.8 performed best (prediction accuracy 80.26%), while 'Fuzzy algebraic sum' and 'fuzzy or' showed worst accuracy of 64.77% and 56.86%, respectively.

Pourghasemi et al[14] showed the applicability of fuzzy logic and analytic hierarchy process in the mapping and zonation of landslide susceptible areas. A total of 12 data layers, which correspond to 12 landslide conditioning factors, were exploited to detect the most susceptible areas. Fuzzy membership values to all pixels were assigned based on the frequency ratio model. Landslide susceptibility was then identified using fuzzy if then else rules. Using the AHP model, weightage of each contributing factor was identified using pairwise comparisons and an equation was modelled for landside susceptible index. Validation of the maps created using both the methods was performed using ROC curve. They concluded that the model with fuzzy logic has the highest area under the curve (AUC) value 0.9194, whereas AHP has 0.8887.

Devkota et al[6] compared certainty factor, index of entropy and logistic regression methods for landslide susceptibility mapping. Slope gradient, slope aspect, altitude, plan curvature, lithology, land use, distance from faults, rivers and roads, topographic wetness index, stream power index and sediment transport index were considered as prominent factors for landslide susceptibility study. The value of the certainty factor ranges between −1 and +1. A positive value means an increasing certainty in landslide occurrence, while a negative value corre-

sponds to a decreasing certainty in landslide occurrence. CF values of the landslide conditioning factors were combined pairwise to generate landslide susceptibility index. Natural breaks were used to classify LSI value to Landslide Hazard Zones. The performance of landslide susceptibility models was assessed using ROC curves. They found that the hazard map prepared using the index of the entropy model has the highest prediction accuracy (90.16%), followed by the logistic regression model (86.29%) and the certainty factor model (83.57%).

Nourani et al[8] prepared landslide hazard zonation maps using genetic programming and compared it with frequency ratio, logistic regression, artificial neural network. Seven factors, i.e. lithology, slope, aspect, elevation, land cover, distance to stream, and distance to road, were considered prominent for landslide hazard zonation study. In the frequency ratio model, landslide hazard index was calculated by summation of frequency ratios for all the factors. In the logistic regression model, LHI was calculated by solving the regression equation. Correlation between landslide event and landslide affecting factors was estimated, and then, equation predicting the landslide was obtained. Three layered feed-forward neural network with back-propagation as training algorithm was used for calculation of LHI. Two different criteria were used to measure the efficiency of the ANN method, i.e. the root mean square error (RMSE) and the determination coefficient (DC). For producing the best landslide susceptibility maps, sensitivity analysis was also implemented in ANN. For verification of LSM, produced by FR, LR, ANN, and GP methods, landslide testing data were compared with these maps. The assessment of AUCs showed that the prediction accuracy of FR, LR, ANN, and GP methods were 89.42%, 87.57%, 92.37%, and 93.27%, respectively.

Bui et al[37] compared the accuracy of landslide prediction, using support vector machine, multilayer perceptron neural network, radial basis function neural network, kernel logistic regression and logistic model tree. Slope, aspect, altitude, relief amplitude, topographic wetness index, stream power index, sediment transport index, lithology, fault density, land use, and rainfall were studied as landslide conditioning factors. For choosing the best subset of conditioning factors, predictive ability of the factors was assessed using the information gain ratio with 10-fold cross-validation technique. The analysis of landslide inventory map showed that landslides mainly occurred during and after the heavy rainfall. The performance of landslide susceptibility models was assessed using receiver operating characteristics (ROC) curves, and reliability was assessed using kappa index. They found that the MLP neural net model has the highest prediction capability of 90.2%, followed by the SVM model 88.7%, the KLR model 87.9%, the RBF neural net model 87.1%, and the LMT model 86.1%.

Youssef et al[9] combined logistic regression and frequency ratio for removing their weaknesses and producing landslide susceptibility maps with better accuracy. Altitude, curvature, distance from wadis, distance from road, distance from fault, stream power index, topographic wetness index, soil type, geology, slope, and aspect were used as contributing factors in landslide occurrences. Frequency ratio was calculated by analyzing the relationship between 11 conditioning factors and landslide occurrence. Landslide hazard index was calculated by summation of frequency ratios for all the factors and solving the regression equation, respectively, for the frequency ratio and logistic regression methods. After this, the probability index

for ensemble of FR and LR was calculated and normalized to be between 0 and 1. For calcu-lating the landslide susceptibility map from ensemble method, the probability index value was classified in five categories using quantile classifier. Probability index value represents the predicted probability of landslide for each pixel in the presence of given set of conditioning factor. Validation of all three models was performed using ROC curves, and they observed that the prediction accuracy of ensemble of FR and LR was higher (82%) than that of FR (58%) and LR (77%) separately.

5. Case Study

The landslide susceptibility mapping is carried out in the Mandakini River basin of Uttarak-hand, which covers an area of about 2439 sq. km and is situated between 30°19'00"N to 30°49'00"N latitude and 78°49'00"E to 79°20'00"E longitude (Figure. 8a) falling in Survey of India toposheet Nos. 53J and 53N.

5.1. Geological setting of the Study Area

The lithological mapping of the area (Figure. 8b) shows the presence of Vaikrita formation in the north, forming most of the Greater/Higher Himalaya in Garhwal. South of this formation, the Munsiyari formation is present in the Lesser Himalaya. South of the Munsiyari formation, the Ramgarh group is present. The southernmost area of the basin is comprised of Berinag Formation. Vaikrita, Munsiyari, Ramgarh, and Berinag formations are, respectively, separated by Main Central Thrust (MCT-I), which is equivalent to Vaikrita Thrust; Main Central Thrust (MCT-II), which is equivalent to Munsiyari/Jutogh Thrust and Main Central Thrust (MCT-III), which is equivalent to Ramgarh/Chail Thrust [58, 59] (Figure. 8b). The presence of MCT Thrust zone causes high shearing and fractures in this area, which makes the rocks weak and highly prone to landslides and other natural hazards.

The high susceptibility to landslides in the Mandakini River basin is mainly due to complex geological settings, varying slopes and relief, heavy rainfall, along with ever-increasing human interference in the ecosystem. Extreme climatic events increase the instability of the terrain, which results in landslides, example includes the Kedarnath disaster [60]. Some of the major landslides occurred in the past are near Okhimath in 1997, 1998, 2010, 2012, 2013; in Phata Byung area in 2001, 2005, 2013; in Madhyamaheshwar area in 1998, 2005, 2013, etc., which are dependent on various factors such as geology, structure, land use, old slides, slope, slope aspect, and drainage in the area [61, 62, 63].

5.2. Data Used

The Survey of India (SOI) toposheet Nos. 53N and 53J were used to create the base map of the study area. Landsat satellite image of October 2008 with 30-m spatial resolution was taken to finalize the tectonic and geologic map of the study area (after) [59]. Elevation data were taken from ASTER-GDEM (Advance Spaceborne Thermal Emission and Reflection Radiometer, Global Digital Elevation Model) having spatial resolution of 30 m with an accuracy of ±10 m.

These data sets were analyzed, preprocessed and then categorized using Arc GIS 9.3, ERDAS Imagine 9.1 software to generate various thematic layers such as elevation, slope, aspect, drainages, geology/lithology, soil, buffer of thrusts/faults, and buffer of streams in the study area (Figure 8 a-h).
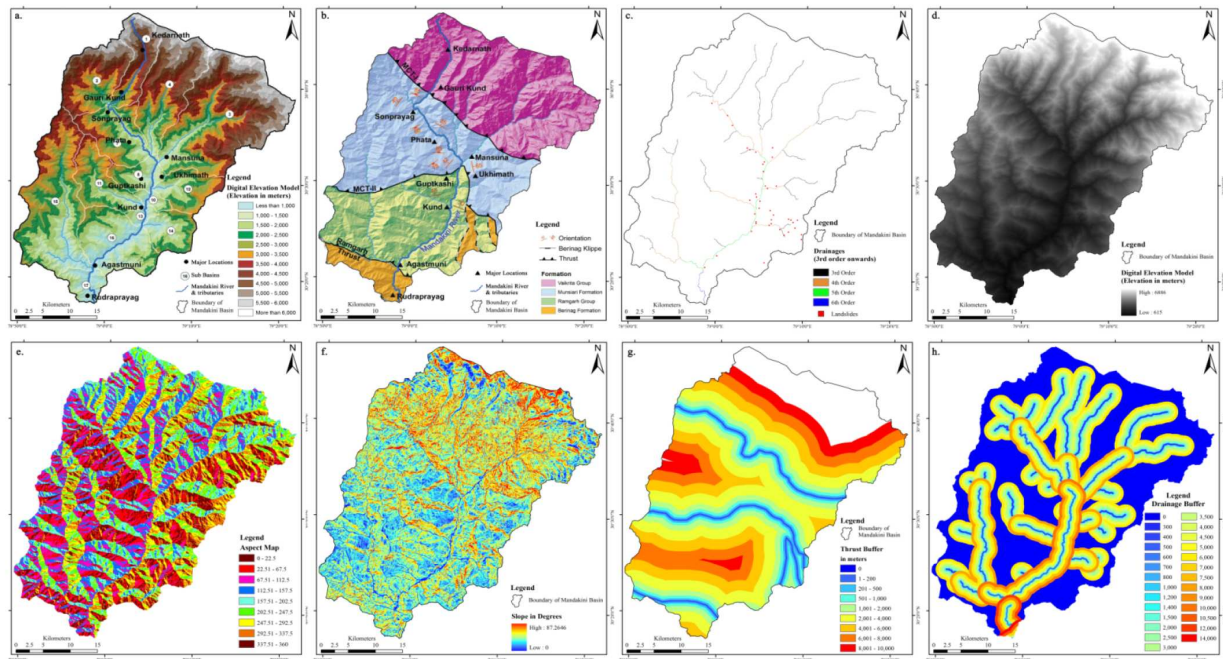

Figure 8. Various thematic layers used in landslide susceptibility prediction using PSVM model. a) Classified elevation map of the study area prepared from ASTER-GDEM showing major locations of Mandakini River basin. b) Geological map showing various formations and structures mainly MCT-I, MCT-II and Ramgarh Thrust (after Shukla et al. [59]). c) Drainage map derived from DEM showing third-order onwards and the presence of landslides in the study area. d) DEM map. e) Aspect map showing variation in the hill facets. f) Slope map showing comparatively higher slopes in northern sides as compared to southern side because of the presence of glacial features. g) Buffer map of the thrusts present in the study area created at specified intervals and reclassified in nine classes. h) Buffer map of the drainages third-order onwards. For the simplicity of the model, first- and second-order streams were not taken.

5.3. Model Selection and Results

All the data sets were generated in Geographic Information System (GIS) environment at 30 × 30 m pixel resolution, the vector layers were converted to raster format with other raster data sets. These raster data sets were converted to ASCII format to be read in MATLAB for using Support Vector Machine (SVM) for prediction of Landslide susceptibility. The landslide data for Okhimath River basin, procured from Geological Survey of India (GSI), were considered to test the SVM model and generate the predictive susceptibility map. The study area contains 1,805,548 pixels, while 2207 pixels are present as landslides. Thus, the pixels representing the landslides are mere 0.125% of the whole study area. The purpose of this study is to predict the landslide, so 1 denotes that pixel involved in landslide and −1 represents pixels that are not involved in landslide. In the whole study area, 2207 pixels were mapped as landslide based on the past data from GSI and other published reports. The whole set of data were divided into 60% as training data and 40% as testing data.
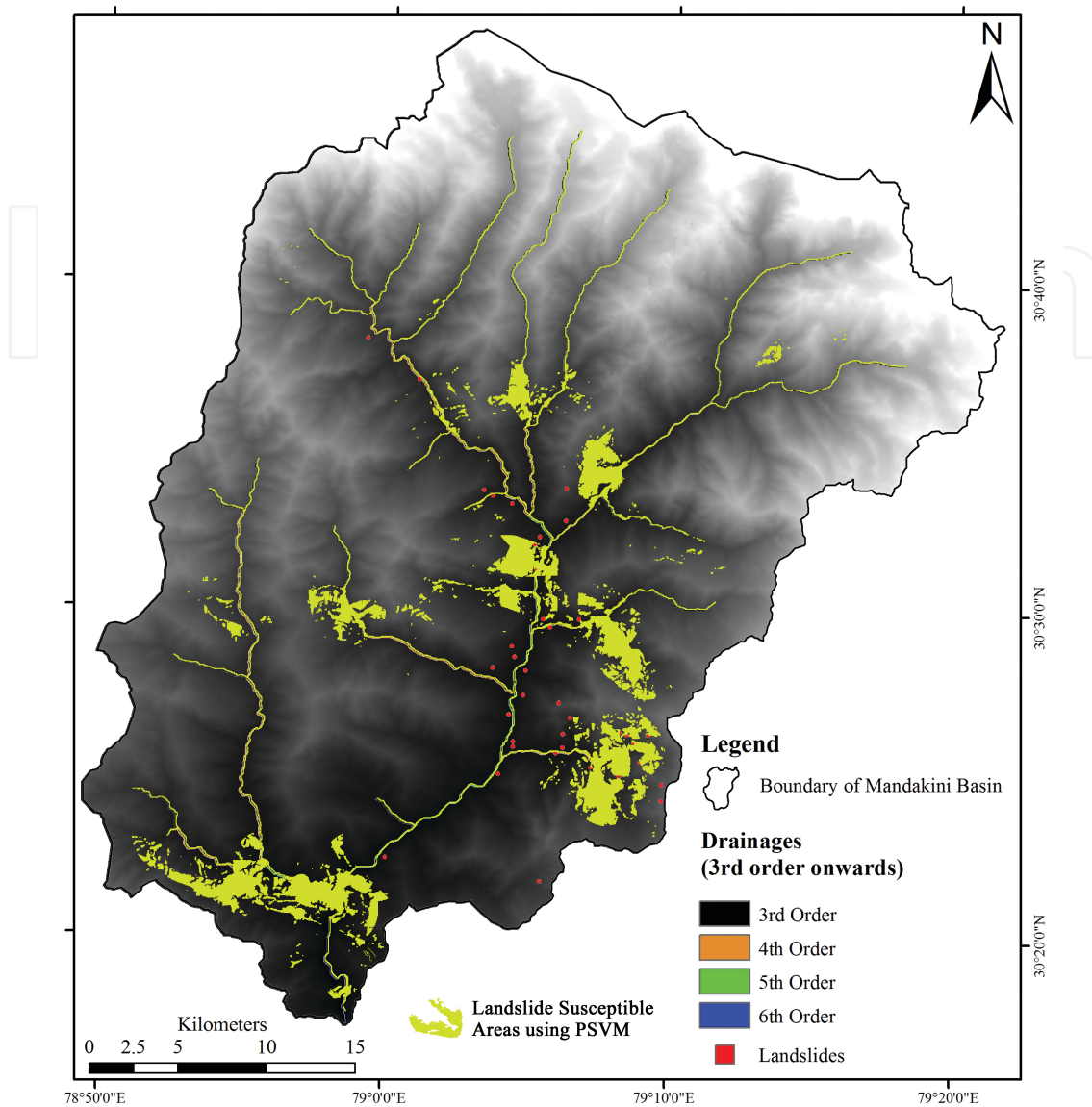
Figure 9. Landslide Susceptibility Map prepared using PSVM model shows areas susceptible to landslides on the DEM and drainage map of the study area with the actual past landslides.

Hence, the landslide susceptibility map for Mandakini River basin was prepared using the Proximal Support Vector Machine (PSVM) model (Figure. 9). It is evident from this figure that the PSVM model classified more areas in landslide susceptible zone as compared to certain landslides have been missed. Hence, various performance metrics such as average prediction accuracy (AA), true positive rate (TPR), true negative rate (TNR) and relative operating characteristic curve (ROC) were computed on testing data to validate the performance of prediction models [64, 65, 66]. The validation results in terms of AUC, and their corresponding testing accuracy showed that the PSVM model has higher AUC values when rainfall data from TRMM were considered with respect to when not considered as shown in Figure 10. The PSVM model with TRMM and without TRMM has an AA of 82.85% and 84.20%, TPR of 79.43% and 72.46%, TNR of 82.85% and 84.22% and an AUC value of 81.15% and 78.34%, respectively

(Table 5). The high value of TNR (82.85% and 84.22%) achieved by the PSVM model in this case is due to the large number of pixels for the study area as compared to pixels forming the landslides. Hence, this model predicted/demarcated the safe areas with 84.22% accuracy when TRMM data were taken into consideration, while it predicted the areas prone to landslide with 79.43% accuracy when TRMM data were taken in consideration because of less number of landslide pixels. Though the AUC values (78.34% and 81.15%) are good, the average accuracy for the PSVM model is quite high between 82.85% and 84.20%. Similar results were also obtained by Pradhan [67] where SVM yielded 81.46% AUC when applied on altitude, slope angle, plan curvature, distance from drainage, distance from road, soil type and NDVI as the input parameters considered for landslide susceptibility mapping for Penang Island in Malaysia.

Model	AA%	TPR%	TNR%	AUC%	C
PSVM (with TRMM)	82.85	**79.43**	82.85	**81.15**	100
PSVM (without TRMM)	**84.2**	72.46	**84.22**	78.34	128

Table 5. Prediction performance for PSVM model.

Best results are shown in bold. AA(%) is the average accuracy, TPR(%) is the true predictive rate, TNR(%) is the true negative rate and AUC(%) is the area under the curve.
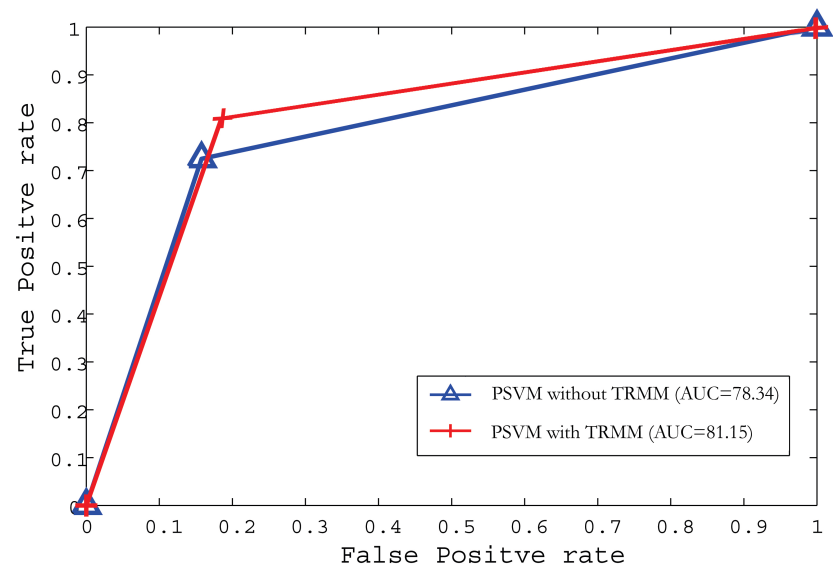

Figure 10. Best Prediction rate and area under the curves (AUC) produced by PSVM model with and without TRMM data consideration.

5.4. Conclusion

In Garhwal Himalaya, Mandakini River basin is highly vulnerable to landslides, especially the town of Okhimath and its nearby villages. In the vicinity of the study area, Mandakini

River crosses various Himalayan thrusts, and due to the presence of these tectonically active MCT zones, the rocks shows high shearing and fracturing and becomes more susceptible for landsliding. The susceptibility to landslide is mainly controlled by valley slopes, attitude of discontinuity of surfaces, soil type, presence of drainage, nature of rocks exposed, and structural and tectonic features present, besides human interaction in the terrain.

Hence, recently developed Support Vector Machine (SVM) learning technique was applied on this area to demarcate the landslide prone and safe areas. The PSVM method has been applied for landslide susceptibility mapping of the study area. The PSVM model showed higher average accuracy (AA) of 82.82%–84.20% for this study area, and the ROC curve indicates that the PSVM model has the prediction accuracy of 81.15%. Nevertheless, this model can be effectively used for landslide susceptibility mapping in this area or similar terrain with these sets of input parameters.

Acknowledgements

Author details

Dericks P. Shukla[1*], Sharad Gupta[1], Chandra S. Dubey[2] and Manoj Thakur[3]

*Address all correspondence to: dericks.82@gmail.com

1 School of Engineering, Indian Institute of Technology, Mandi (HP), India

2 Department of Geology, University of Delhi, Delhi, India

3 School of Basic Sciences, Indian Institute of Technology, Mandi (HP), India

References

[1] Yan, X. & Su, X.G., 2009. Linear Regression Analysis: Theory and Computing, *World Scientific*, pp. 1–4.

[2] Chatterjee, S. & Hadi, A.S., 2006. Regression Analysis by Example. 4th ed., *Wiley Inter-Science*, New Jersey, pp. 12–15.

[3] Chatterjee, S. & Simonoff, J.S., 2013. Handbook of Regression Analysis. *Wiley Inter-Science*, New Jersey, pp. 3–16.

[4] Mendenhall, W. & Sincich, T., 2012. A Second Course in Statistics Regression Analysis. 7th ed., *Prentice Hall*.

[5] Lee, S. & Pradhan, B., 2007. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides*, 4(1), pp. 33–41.

[6] Devkota, K.C., Regmi, A.D., Pourghasemi, H.R., Yoshida, K. et al., 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya. *Natural Hazards*, 65(1), pp.135–165.

[7] Kleinbaum, D.G. & Klein, M., 2010. Logistic Regression: A Self-Learning Text, 3rd ed., *Springer*, pp. 4–10.

[8] Nourani, V., Pradhan, B., Ghaffari, H., & Sharifi, S.S., 2014. Landslide susceptibility mapping at Zonouz Plain, Iran using genetic programming and comparison with frequency ratio, logistic regression, artificial neural network models. *Natural hazards*, 71(1), pp. 523–547.

[9] Youssef, A.M., Pradhan, B., Jebur, M.N., & El-Harbi, H.M., 2014. Landslide susceptibility mapping using ensemble bivariate and multivariate statistical models in Fayfa area, Saudi Arabia. *Environmental Earth Sciences*, 73(7), pp. 3745–3761.

[10] Saaty, T.L. & Vargas, L.G., 2012. Models, Methods, Concepts & Applications of the Analytic Hierarchy Process. *Springer Science & Business Media*, New York, pp. 1–7.

[11] Saaty, T.L. & Kearns, K.P., 1985. Analytical Planning: The Organization of Systems. *Pergamon Press*, pp. 19–40.

[12] Saaty, T.L. & Vargas, L.G., 1982. The Logic of Priorities—Applications in Business, Energy, Health, and transportation. *Springer Science & Business Media*, New York.

[13] Brunelli, M., 2015. Introduction to the Analytic Hierarchy Process. *Springer Briefs in Operations Research*, pp. 1–15.

[14] Pourghasemi, H.R., Pradhan, B. & Gokceoglu, C., 2012. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Natural Hazards*, 63(2), pp. 965–996.

[15] Bhatt, P.B., Awasthi, K.D., Heyojoo, B.P., Silwal, T., & Kafle, G., 2013. Using geographic information system and analytical hierarchy process in landslide hazard zonation. *Applied Ecology and Environmental Sciences*, 1(2), pp. 14–22.

[16] Reza, M. & Daneshvar, M., 2014. Landslide susceptibility zonation using analytical hierarchy process and GIS for the Bojnurd region, northeast of Iran. *Landslides*, 11, pp. 1079–1091.

[17] Tazik, E., Jahantab, Z., Bakhtiari, M., Rezaei, A. & Alavipanah, S.K., 2014. Landslide susceptibility mapping by combining the three methods Fuzzy Logic, Frequency Ratio and Analytical Hierarchy Process in Dozain basin. In *ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 267–272.

[18] Boroumandi, M., Khamehchiyan, M. & Nikoudel, M.R., 2015. Using of Analytic Hierarchy Process for Landslide Hazard Zonation in Zanjan Province, Iran. *Engineering Geology for Society and Territory*, 2, pp. 951–955.

[19] 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(3), pp. 559–572.

[20] Jain, A.K., Mao, J. & Mohiuddin, K.M., 1996. Artificial neural network: a tutorial. *Computer*, 29(3), pp. 31–44.

[21] Konar, A., 1999. Artificial Intelligence and Soft Computing, Behavioral and Cognitive Modelling of the Human Brain. *CRC Press*.

[22] Haykin, S., 2005. Neural Network: A Comprehensive Foundation. 2nd ed., *Prentice Hall*.

[23] Zurada, J.M., 1992. Introduction to Artificial Neural System. *West Publishing Company*.

[24] Demuth, H.B. & Beale, M., 2002. Neural Network Toolbox. *The MathWorks*, ver. 4.

[25] Alawala, C.R., 2007. Fuzzy Logic and Neural Networks: Basic Concepts and Applications. *New Age International Publisher*, pp. 121–143.

[26] Graupe, D., Liu, R.W. & Moschytz, G.S., 1988. Applications of neural networks to medical signal processing. In *Proceedings of the 27th IEEE Conference on Decision and Control Austin, Texas*, pp. 343–347.

[27] Gorzalczany, M.B., 1996. An idea of the application of fuzzy neural networks to medical decision support systems. In *Proceedings of the IEEE International Symposium on Industrial Electronics*, 1, pp. 398–403.

[28] Edwards, P.J., Murray, A.F., Papadopoulos, G., Wallace, A.R., et al., 1999. The application of neural networks to the papermaking industry. *IEEE Transactions on Neural Networks*, 10(6), pp. 1456–1464.

[29] Sun, Z., 2009. Application of neural network in calculation of nonlinear interpolation algorithm. In *IEEE International Conference on Information Science and Engineering*, pp. 3981–3984.

[30] Daigavane, P.M., Bajaj, P.R. & Daigavane, M.B., 2011. Vehicle Detection and Neural Network Application for Vehicle Classification. In *International Conference on Computational Intelligence and Communication Systems*.

[31] Mani, N. & Srinivasan, B., 1997. Application of artificial neural network model for optical character recognition. In *IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*. pp. 7–10.

[32] Smith, K.A. & Gupta, J.N.D., 2000. Neural networks in business: techniques and applications for the operations researcher. *Computers & Operations Research*, 27, pp. 1023–1044.

[33] *Lee, S. & Oh, H.J., 2011. Application of Artificial Neural Network for Mineral Potential Mapping, Artificial Neural Networks - Application, Dr. Chi Leung Patrick Hui (Ed.). ISBN: 978-953-307-188-6, *InTech*, DOI: 10.5772/16187.

[34] Kanungo, D.P., Arora, M.K., Sarkar, S., & Gupta, R.P., 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(3-4), pp.347–366.

[35] Pradhan, B., Sezer, E. A., Gokceoglu, C., & Buchroithner, M. F. (2010). Landslide susceptibility mapping by neuro-fuzzy approach in a landslide-prone area (Cameron Highlands, Malaysia). *Geoscience and Remote Sensing, IEEE Transactions on*, *48*(12), 4164-4177.

[36] Pradhan, B., Mansor, S. & Pirasteh, S., 2011. Landslide Susceptibility Mapping: an Assessment of the Use of an Advanced Neural Network Model with Five Different Training Strategies, Artificial Neural Networks - Application, Dr. Chi Leung Patrick Hui (Ed.), ISBN: 978-953-307-188-6, *InTech*, DOI: 10.5772/15738.

[37] Bui, D. T., Tuan, T. A., Klempe, H., Pradhan, B., & Revhaug, I. (2015). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, 1-18.

[38] Hsu, C.W., Chang, C.C., & Lin, C.J., 2003. A Practical Guide to Support Vector Classification.

[39] Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*,2(3), 27.

[40] Support Vector Machines (SVM) Introductory Overview, http://www.statsoft.com/Textbook/Support-Vector-Machines

[41] Mather, P.M. & Koch, M., 2011. Computer Processing of Remotely-Sensed Images: An Introduction. 4th ed., *Wiley Blackwell*, pp. 267-268.

[42] Abe, S., 2010. Support Vector Machines for Pattern Classification, 2nd ed., *Springer-Verlag* London, pp. 20-24.

[43] Campbell, C. & Ying, Y., 2011. Learning with Support Vector Machines. *Morgan Claypool Publishers*, pp. 1-5.

[44] Watanachaturaporn, P., Arora, M.K., & Varshney, P.K., 2008. Multisource Classification Using Support Vector Machines: An Empirical Comparison with Decision Tree and Neural Network Classifiers. *Photogrammetric Engineering & Remote Sensing*, 74(2), pp. 239–246.

[45] Samui, P., 2014. Vector machine techniques for modeling of seismic liquefaction data. *Ain Shams Engineering Journal*, 5, pp.355–360.

[46] Huang, R., Samy, M., Tawfik, H., & Nagar, A.K., 2008. Application of Support Vector Machines in Financial Literacy Modelling. *in Second UKSIM European Symposium on Computer Modeling and Simulation*, 2008, pp. 311–316.

[47] Moguerza, J.M. & Munoz, A., 2006. Support Vector Machines with Applications. Statistical Science, 21(3), pp.322–336.

[48] Osuna, E., Freund, R. & Girosi, F., 1997. Training Support Vector Machines: an Application to Face Detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 130-136.

[49] Kim, K.I., Jung, K., Park, S.H., Kim, H.J., 2002. Support Vector Machines for Texture Classification. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 24(11), pp.1542–1550.

[50] Tu, J. V, 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology*, 49(11), pp.1225–1231.

[51] Igelnik, B., 2011. Computational Modeling and Simulation of Intellect: Current State and Future Perspectives. *Information Science Refer*, Chocolate Avenue Hershey PA, pp.226.

[52] Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., & Reichenbach, P. (1991). GIS techniques and statistical models in evaluating landslide hazard.*Earth surface processes and landforms*, 16(5), 427-445.

[53] Jade, S., & Sarkar, S. (1993). Statistical models for slope instability classification. *Engineering Geology*, 36(1), 91-98.

[54] Guzzetti, F., Carrara, A., Cardinali, M., & Reichenbach, P. (1999). Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*, 31(1), 181-216.

[55] Ermini, L., Catani, F., & Casagli, N. (2005). Artificial neural networks applied to landslide susceptibility assessment. *Geomorphology*, 66(1), 327-343.

[56] Kanungo, D. P., Sarkar, S., & Sharma, S. (2011). Combining neural network with fuzzy, certainty factor and likelihood ratio concepts for spatial prediction of landslides. *Natural hazards*, *59*(3), 1491-1512.

[57] Pradhan, B., Lee, S., & Buchroithner, M. F. (2009). Use of geospatial data and fuzzy algebraic operators to landslide-hazard mapping. *Applied Geomatics*, *1*(1-2), 3-15.

[58] Ray, Y., Srivastava, P.: Widespread aggradation in the mountainous catchment of the Alaknanda-Ganga river system: timescales and implications to hinterland-foreland relationships. Quaternary Science Reviews 29(17), 2238-2260 (2010)

[59] Shukla, D., Dubey, C., Ningreichon, A., Singh, R., Mishra, B., Singh, S.: GIS-based morphotectonic studies of Alaknanda river basin: a precursor for hazard zonation. Natural hazards 71(3), 1433-1452 (2014)

[60] Dubey, C., Shukla, D., Ningreichon, A., Usham, A.: Orographic control of the Kedarnath disaster. Current Science 105(11), 1474-1476 (2013)

[61] Rautela, P., Thakur, V.: Landslide hazard zonation in Kaliganga and Madhyamaheshwar valleys of Garhwal Himalaya: a GIS based approach. Himalayan Geol 20:2, 31-44 (1999)

[62] Sati, S., Naithani, A., Rawat, G.: Landslides in the Garhwal Lesser Himalaya, UP, India. Environmentalist 18(3), 149-155 (1998)

[63] Chaudhary, S., Gupta, V., Sundriyal, Y.: Surface and sub-surface characterization of Byung landslide in Mandakini valley, Garhwal Himalaya. Himalayan Geology 31:2, 125-132 (2010)

[64] Bradley, A.P.: The use of the area under the roc curve in the evaluation of machine learning algorithms. Pattern recognition 30(7), 1145-1159 (1997)

[65] Brenning, A.: Spatial prediction models for landslide hazards: review, comparison and evaluation. Natural Hazards and Earth System Science 5(6), 853-862 (2005)

[66] Webb, A.R.: Statistical pattern recognition. John Wiley & Sons (2003)

[67] Pradhan, B.: A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Computers & Geosciences 51, 350-365 (2013)