

# Chapter 19

## Landslide Hazard Assessment Using Machine Learning and GIS



Amit Jaiswal, A. K. Verma, T. N. Singh, and Jayraj Singh

**Abstract** India's National Disaster Management Authority (NDMA) has highlighted landslides as a complex geological natural hazard. As a result, identifying hazardous zones under varied geometrical and geotechnical conditions has remained a complex problem that has received much attention in the classic era. This work aims to predict landslide hazard zones using spatial analysis by applying multi-criteria analysis methods. For the investigation, Ramban area of Jammu & Kashmir, India, has been considered. The locations of landslides were identified by analysis of Google earth images and field surveys. The major causative factors of landslides such as relative relief, slope, geological structure, lithology, soil thickness, hydrological condition, and land use and land cover were extracted from QGIS software tool, Google earth images, and field survey. The two major triggering factors such as rainfall and seismicity, were also included in the study. The cumulative effect of all these factors was considered for preparing the dataset that was scrutinized to produce a landslide hazard zonation map. It has become very much possible to develop intelligent models calibrated to experimental data to predict landslide zonation with the least potential errors with the advancement of machine learning-based computation technologies. In view of these observations, in this chapter: (i) the relevant spatial data are identified, processed, and analysed (ii) data is used to construct an intelligent machine learning model, namely the back-propagation neural network (BPNN), to predict the hazardous zonation. The study has been validated by comparing the landslide hazard zonation map with the actual occurrence of landslides. The outcomes of this research show that the designed AI-based model is quite promising and may be utilised successfully by practicing professionals to estimate landslide zonation with reasonable accuracy.

---

A. Jaiswal · A. K. Verma (✉)

Department of Mining Engineering, Indian Institute of Technology BHU, Varanasi 221005, India  
e-mail: [amitvermaism@gmail.com](mailto:amitvermaism@gmail.com)

T. N. Singh

Department of Civil and Environmental Engineering, Indian Institute of Technology,  
Patna 801106, Bihar, India

J. Singh

Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi 110016, India

**Keywords** Landslide Hazard · Remote sensing and GIS · Machine learning and landslide

## 19.1 Introduction

Landslides are one of the most prevalent natural disasters that take lives, property, and the environment all over the world. Every year, landslides claim the lives of people and property worldwide, including in India. The Himalaya's dynamic nature has resulted in numerous landslides across the region. Jammu & Kashmir is the westernmost part of the Indian Himalayas and is prone to landslides. The area commonly witnesses extreme weather conditions with excessive seasonal precipitation and snowfall that generally trigger natural hazards such as landslides. National Highway (NH-1A) of Jammu-Srinagar frequently gets blocked at a number of sites during monsoons due to mass movement. The most affected portion is Ramban-Banihal stretch of the national highway (NH-1A). This is one of the most susceptible zones to landslides in the world (Chingkhei et al. 2013). The area passes through some major destructive landslides like the Digdol slide, Panthiyal slide, Battery Chashma slide, and Khoni Nallah slide. Landslide hazard zonation is the most preliminary and important approach to identifying and studying landslide effects in an area. The first approach in landslide hazard zonation mapping was achieved by using factor rating of different parameters causing landslides. With the advancement in hazard zonation mapping approach, analytical hierarchy process, weighted overlay model, rainfall intensity model, frequency ratio model, logistic regression modelling, and different statistical approach have become more important zonation methods. The threshold of different causative factors like slope, elevation, aspect, lithology, hydrological conditions, geological structures, and triggering factors like seismicity and rainfall are the main causes of slope instability. Various statistical and probabilistic approaches for slope instability identification and hazard zonation using GIS and remote sensing have been established as an important approach for studying landslides globally.

The varying characteristics of the material and other uncertainties found in the geological parameters may cause slope instability (Singh et al. 2019). However, finding reliable values of input parameters is cumbersome and becomes a crucial task. The researchers began to gain more attention towards machine learning-based methods with technological advancements (Singh et al. 2016) for fast and efficient predictions. Different Artificial Neural Network (ANN) based models (Pistocchi et al. 2002; Lee et al. 2003a, b; Arora et al. 2004; Lee et al. 2006; Yilmaz 2009) have been employed for landslide susceptibility assessment. Sequentially, a combination of neural and fuzzy approaches has been utilized by Kanungo et al. (2006) in developing landslide susceptibility maps. In a study by Nefeslioglu et al. (2008), the ANNs have achieved a very optimistic result in the estimation of landslide susceptibility than the logistic regression method. Further, Melchiorre et al. (2008) have shown that predictive capability can be improved through cluster analysis in ANN. Furthermore,

Pradhan and Lee (2010a, b, c) have used the logistic regression, frequency ratio, and ANN models for landslide susceptibility analysis.

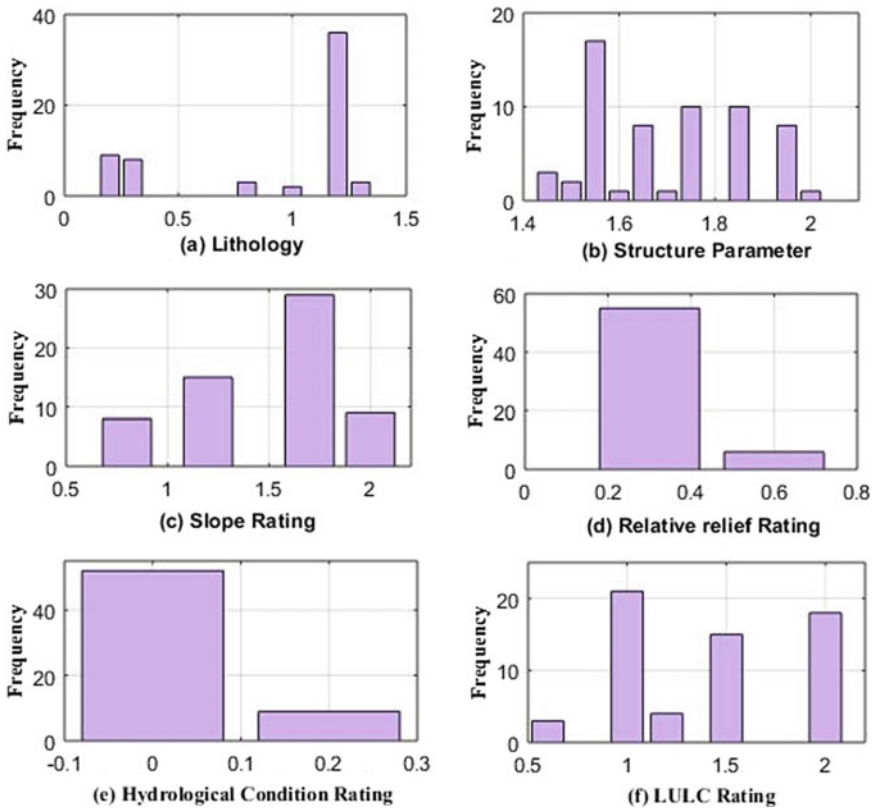
In the present work, the ANN and Logistic Regression (LR) based models have been utilized for landslide hazard assessment, where the considered landslide zones of the study area located along NH-1A from Ramban to Banihal of Ramban district, Jammu & Kashmir, India. Stratigraphically, some part of Ramban to Banihal stretch lies between two major thrust zone of Himalaya of regional importance i.e., Main boundary thrust (MBT) in the south and Main Central Thrust (MCT) in the north, and covers some part of central crystalline to the north of MCT. The area is tectonically very active, due to which it has developed very complex geology (Shanker et al. 1989). Rock type in the study area comprises mainly phyllites, slates, limestone, quartzites, volcanic, and some fine to medium-grained sedimentary rocks near Banihal (Shanker et al. 1989; Bhat et al. 1999). Granitic intrusions are exposed between Digdole and Makerkote. The regional strike of the area varies from NW–SE to WNW–ESE, with a dip varies from moderate to very steep towards N or S. Strata are characterized by three to four joint sets with one joint parallel to the road. Different rating given in the Bureau of Indian Standards (BIS) code for landslide hazard zonation mapping for hilly terrain on the scale (1:50,000) has been utilized (BIS 1998) during this study. The rating for landslide causative factors and triggering factors such as lithology, structure, relative relief, slope angle, land use and land cover, hydrological condition, seismicity, and rainfall (Anbalagan et al. 2008) have been considered for the proposed assessment. The QGIS software tool has been utilized to find the data parameters from the digital elevation model (DEM). The QGIS software is an open-source GIS tool that gained huge popularity among engineers, researchers, and scientists of earth science to compute the geospatial data by incorporating a large number of data access, processing, visualization, and analysis operations. On the contrary, the simulation design of machine learning models for the presented study has been accommodated using ‘WEKA’ an open-source software (Hall et al. 2009) tool. The materials and methods used in this study have been mentioned in the next section.

## 19.2 Materials and Methods

Generally, a machine learning model predicts the outcome(s) through the interpolation of the data set. Data splitting was carried out to train and test the data set on which the performance was measured to build an efficient classifier. The data splitting for the training and testing sets was done with statistical considerations (i.e., mean, standard deviation, and coefficient of variation). The important parameters which influence the instability of the slope are depicted in Fig. 19.1, whereas the statistical aspects (minimum, maximum, mean, and standard deviation) are illustrated in Table 19.1. In the present work, the data set has been divided into a ratio of 70:30 for training and testing, respectively. Generally, 70% data is considered sufficient in training any network, and the rest of the data is reserved for testing and validating the model. To make a predictive classifier, relative relief, slope, lithology, geological

structures, hydrological condition, and land use and land cover were defined as input parameters into the machine learning model, and the type of hazardous zone, i.e., very high hazard zone, high hazard zone, moderate hazard zone and low hazard zone and very low hazard zone state indicates output.

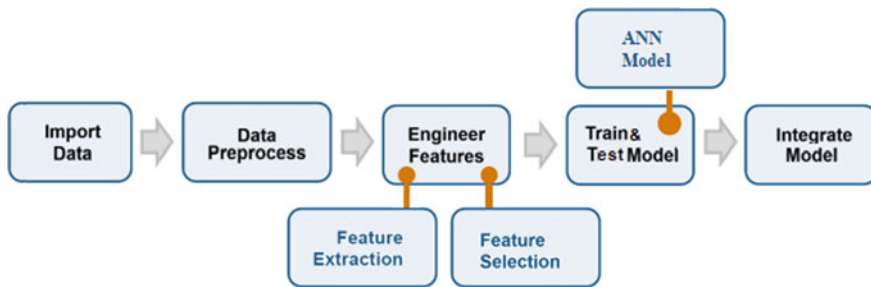
The important attributes of the data based on feature extraction and selection have been processed in the data processing phase. The general performance of the classification models has been achieved using training data sets. On the contrary, testing data was used to check the final model’s predictive capabilities. A complete flowchart of the procedure has been demonstrated in Fig. 19.2. In order to tune the outcome and to prevent the problem of over-fitting the models, the hyper-parameters are defined during training, which is selected based on a trial-and-error method.



**Fig. 19.1** Causative factor rating according to BIS code with minimum, maximum, mean, and standard deviation **a** Lithology rating, **b** Relative relief rating, **c** Structural parameter rating, **d** Slope rating, **e** Hydrological condition rating

**Table 19.1** Different statistical aspects of the data parameters

Factors rating	Minimum	Maximum	Mean	Standard Deviation	Coefficient of variation
Lithology	0.2	1.2	0.800	0.460	0.576
Relative relief	0.3	0.6	0.450	0.212	0.472
Structural parameter	1.45	2	1.700	0.187	0.110
Slope	0.8	2	1.425	0.532	0.373
Hydrological condition	0	0.2	0.100	0.141	1.414
LULC	0.6	2	0.260	0.527	2.028

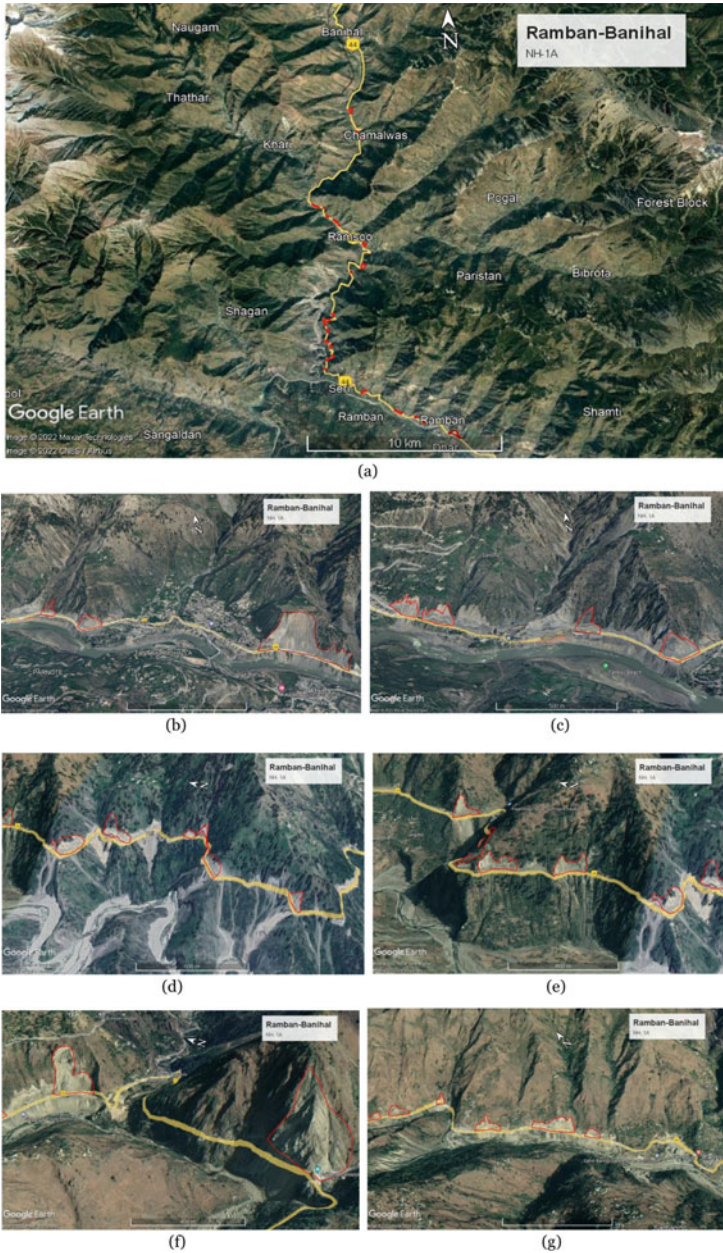


**Fig. 19.2** Flow diagram of prediction classification model using BPNN

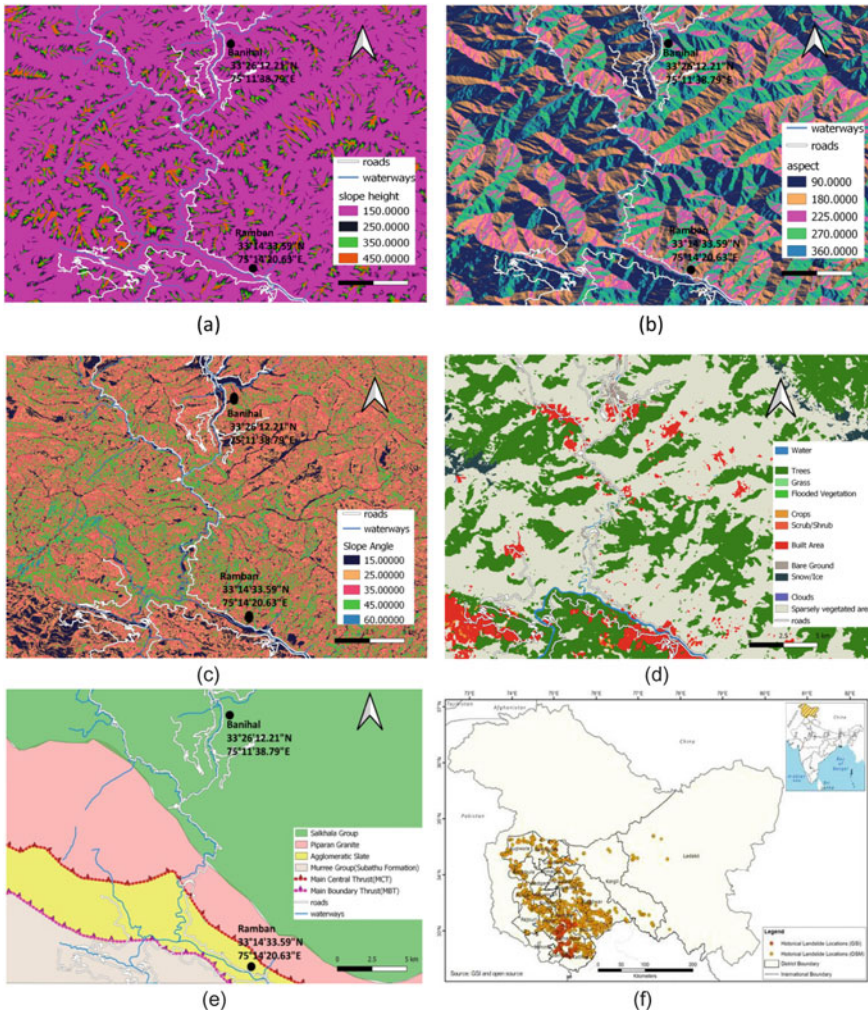
### 19.2.1 Dataset and Its Preliminary Analysis

The proposed work considers 61 landslide locations of hazard zones along NH 1-A of Ramban district, Jammu, and Kashmir for preparing the study dataset. According to BIS code, the study area falls under two hazards zone, namely moderate hazard and high hazard, in the classification (IS 14496 Part 2 1998). The google earth image of the area shown in Fig. 19.3 shows the aforesaid landslide presence.

For the above landslides area, 11-inputs factors were considered in the dataset, including six causative factors, two triggering, and three geological structure parameters’ relationship between slope and rock strata. The causative factors include relative relief, slope, lithology, hydrological condition, soil thickness, land use, and land cover while triggering factors include rainfall and seismicity. The rainfall data were taken from IMD, Ramban and according to seismic zonation of the country area lies in seismic zone IV. Based on the severity of these factors in initiating landslide, different rating has been described in the Bureau of Indian Standards (BIS) code. The important features were extracted from QGIS, Google earth images, and field photographs, while the physiographical and geological studies were carried out in the field itself. The thematic maps of causative factors are depicted using Fig. 19.4.



**Fig. 19.3** Google earth image depicts **a** Ramban to Banihal stretch, **b** landslide located near Ramban along NH-1A, **c** landslide located between Ramban and Banihal along NH-1A, **d** landslide located between Ramban and Banihal along NH-1A, **e** landslide located between Ramban and Banihal along NH-1A, **f** landslide located between Ramban and Banihal along NH-1A, **g** landslide located near Banihal along NH-1A



**Fig. 19.4** Causative factors of landslide thematic maps **a** Relative relief, **b** Slope aspect, **c** Slope angle, **d** LULC, **e** Lithological map, **f** Historical landslide map

### 19.3 Selection of Machine Learning Models

Many researchers have used computational intelligence methodologies with the progress of high computational processing power (Singh. et al. 2016; Al-Najjar and Pradhan 2021; Dou et al. 2015). Artificial intelligence and machine learning approaches have enhanced the accuracy of predicting sensitive areas by landslide evaluation. A Back-Propagation Neural Network (BPNN) was used for the current study to predict landslide zones with large accuracy. The approach is used to build

a predictive classifier by examining multivariate data. The BPNN method was compared to logistic regression and SVM in order to validate and compare the method's performance. The aforementioned approaches have become quite popular because of its ease of use, efficacy, and ability to generate complicated non-linear relationship models. Table 19.2 gives a quick overview of these methods, however, the readers are encouraged to consult the necessary references for more detail (Pradhan and Lee 2010b, c; Dou et al. 2015; Li et al. 2012).

The BPNN algorithm is found to perform reasonably well when applied to the supplied dataset. In comparison to logistic regression and SVM, the BPNN has a high prediction accuracy. Furthermore, because of its simplicity, diversity, and capacity to handle big datasets, missing data, and the risk of over-fitting problems, the BPNN technique has been widely utilised. Besides, it has previously demonstrated remarkable performance for such nonlinear problems (Khandelwal and Singh 2009; Bashiri and Geranmayeh 2011; Khaw et al. 1995; Tortum et al. 2007). Because of these following facts, the BPNN regression model was used for the study.

**Table 19.2** Overview of machine learning algorithms employed for landslide assessments

ML models	Specifications
Back Propagation Neural Network (BPNN)	BPNN is a multi-layered feed-forward neural network component. It is made up of three layers: the input layer, the output layer, and the hidden layer. The input layer receives the input signal to be processed. The output layer is responsible for tasks such as prediction and classification (Dou et al. 2015)
Support Vector Machine (SVM)	SVMs revolve around the notion of a 'margin' on either side of a hyperplane that separates two data classes. Maximizing the margin distance provides some reinforcement to classify future data points with more confidence (Singh et al. 2016; Cortes et al. 1995)
M5-rule Algorithm	M5 rule-Generates a decision list for regression problems using separate and conquer. It builds a model tree in each iteration using the M5 algorithm and makes the "best" leaf into a rule (Holmes et al. 1999; Witten and Wang 1997)
M5-P tree	M5P tree algorithm combines a conventional decision tree with the possibility of linear regression functions at the terminal (leaf) nodes that can predict continuous numerical attributes. It is a decision tree learner for a regression problem that is used to predict the values of a numerical response variable (Singh et al. 2017; Quinlan 1992)



### **19.3.1 Back-Propagation Neural Networks (BPNN)**

The BPNN is a machine learning algorithm developed to act like a human memory to perform any specific work (Haykin 1999). It was invented in the 1970s but has gained popularity in recent times because of its capability to be used in every field. As the neurons work in the human nervous system, ANN also works in the same fashion, learns from the data, and prepares itself to respond either as classification or prediction. It involves many interconnected processing units which process the information. In the configuration of the ANN model, the neurons help in processing the stimulation with the ability to receive and transmit the signal. The signals are transmitted or received from one neuron to another in a well-defined network connected with synaptic weight which represents the bond strength of each neuron. The synaptic weight is positive relationships among the contributing nodes (Fausett 1994). ANN model is able to predict on the basis of input data even if the relationship between the input factors is not very clear or their physical significance is hard to describe. Due to this characteristic of ANN model, it is more preferred and advantageous than that statistical and empirical method in which a well-defined relationship between the input data is needed in the form of linear, nonlinear, etc. The main difference between conventional processing of information and neural network is that it has the ability of proper data manipulation if appropriate weights between layers are selected. It also has weight changing mechanism.

#### **19.3.1.1 Layers of Back Propagation Neural Network**

BPNN is catalogued into three layers. The layers comprise connected nodes which have a 'stimulation function'.

(i) *Input Layer*

The input layer acts as a receiver for the input of raw information. Generally, input nodes are equal to a number of independent variables present in the data set that needs to be fed into the network. Each input node is connected to one or more hidden layers. The nodes of the input layer do not change the input value.

(ii) *Hidden Layer*

This layer basically transforms the raw value in the network according to the desired activity and acts as a link between input data and output result. There can be one or more hidden layers in any network. The values given to hidden nodes are multiplied by hidden layers weight and produce a single number as an output value.

(iii) *Output Layer*

The output layer is connected either from a hidden layer or from the input layer. It produces output value based on the activity of the hidden layer and the weights defined among the output and hidden. In the case of prediction, the output value

corresponds to the response of the input variable. In the case of a classification problem, there is only a single output node.

### 19.3.1.2 Training Methods

During the training process, neural networks acquire information as humans do, and the information or the signal is stored in the network system (Haykin 1999). The BPNN model is trained in such a way that it can combine the thinking process of the human brain with mathematical intellect (McClelland et al. 1986). The training method has three steps:

- (i) In this step, supervised network training is done with the help of a set of training patterns related to the target output.
- (ii) Unsupervised training is done in this step, where the network is trained by a group of a similar set of training patterns as that of a supervised training pattern.
- (iii) In the last step, the reinforcement training is done with the help of the interaction of the network with the environment.

The training phase output and hidden layer neuron of the network processes the input and multiple each input with the corresponding weightage and then process the sum with a non-linear transformation function to yield the final result. The ANN can also adjust weight among the neurons in feedback to the error between the target output and actual output values. The outcome of training the network is the development of a model that can predict the target value from a given input value.

### 19.3.1.3 Features of BPNN

The back-propagation neural network comprises three different attributes: network structure, training algorithm, and function for activation (Fausett 1994; Benardos and Benardos 2005). The network structure is also termed network architecture, and it is the design or the relation by which different neurons are connected. A network with only two units, i.e., input and output layer, is the simplest structure of the network, and when one more layer, i.e., a hidden layer, is added to the network, it becomes capable of predicting desired output values. Design of the network is the most critical part of the network on the basis of understanding of the problem (Benardos and Kaliampakos 2004). In the training algorithm, part weightage value is given to the connection of neurons. The activation function is used in the last stage of the development of the model. In the feed-forward network, the signal flows towards the output direction from the input direction. The BPNN is a multi-layer, feed-forward neural network component. It is made up of three layers: the input layer, the output layer, and the hidden layer. The input signal to be processed is received by the input layer. The output layer is responsible for tasks such as prediction and classification. The real computational engine of the BPNN is an arbitrary number of hidden layers inserted between the input and output layers. In a BPNN, data flows forward from the input to

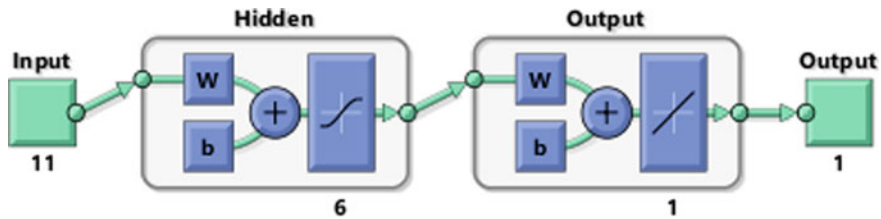


Fig. 19.5 Architecture of back propagation multi-layer neural network (BPNN)

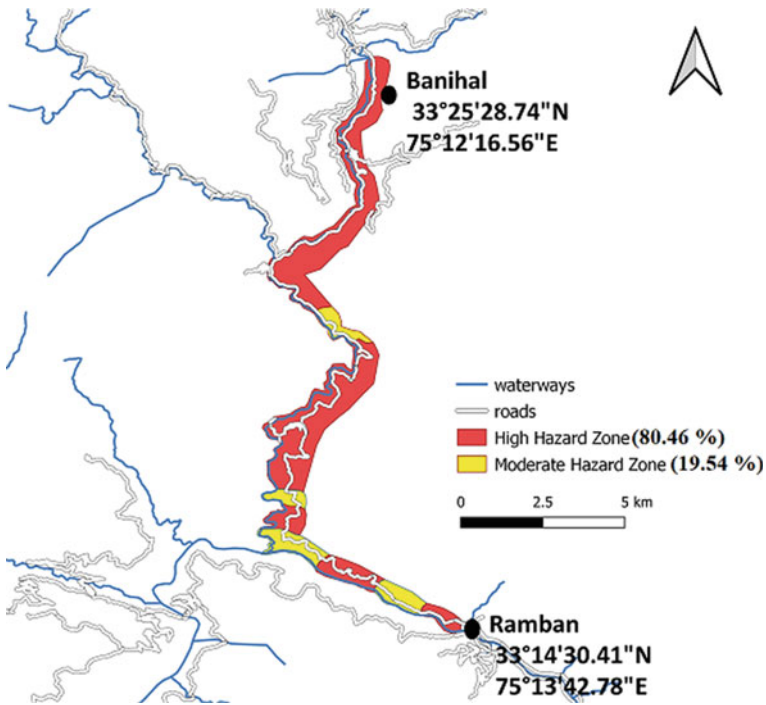
the output layer, similar to a feed-forward network. The hidden layer has nodes due to which complexities develop in the model. The architecture design of the employed BPMNN is shown in Fig. 19.5.

## 19.4 Results and Discussion

The landslide hazard zonation map has been developed for the study area according to the rating laid in BIS code as discussed in the previous sections. The rating has divided the NH 1-A Ramban to Banihal stretch into two categories, “Moderate hazard zone and High hazard zone,” which is represented by 1 or 2, illustrated in Fig. 19.6. The result has revealed that most of the region in the study area falls under the high hazard category, and some part falls in the moderate hazard zone. The percentages of these categories are 80.46 and 19.54%, respectively as shown in Fig. 19.6. The redundant data that was located beyond the mean range was removed, and the remaining data samples were used in a 70–30% ratio for training, testing, and validation. This regression model’s performance was compared across the training, testing, and validation sets. The regression plot shows the prediction responses (output) for the training, testing, and validation sets. The data should fall along a 45-degree line for a perfect fit, with the network outputs equal to the responses. As shown in Fig. 19.7, the result indicates that the data fit is reasonable, with a high  $R$ -value for all data sets.

A plot of training record error values against the number of training epochs has also been discussed using Fig. 19.8. Generally, the error reduces after more epochs of training but might start to increase on the validation data set as the network starts overfitting the training data. The training stops after 1,000 epochs consecutive increases in validation error in the default setup, and the best performance is taken from the epoch with the lowest validation error. The result is found reasonable due to the minimum mean-square error. The test and validation set errors have also shown similar characteristics. Besides, no significant overfitting has occurred by epoch 1,000.

The investigation outcome revealed that the BPNN model is the most appropriate tool for predicting landslide hazard zonation as compared to other discussed machine learning models. The actual vs. predicted results plot is depicted in Fig. 19.9.



**Fig. 19.6** Landslide hazard zonation map of National Highway 1-A along Ramban to Banihal

In comparison to other ML-methods (M5-P, M5-rule, and SVM approaches), the investigation outcomes demonstrated that the BPNN with 6 hidden layers gives the least values of correlation coefficient value as 0.8715, the mean absolute error value is 0.192, and the root means squared error (RMES) value is 0.278. Comparison plots between mean absolute error (MAE) and root mean square error (RMES) over machine learning methods are depicted in Fig. 19.10.

## 19.5 Conclusions

Landslides often occur in the hazardous zone of the Himalayan region. It became cumbersome and challenging due to uncertain geological and geotechnical parameters. The factors causing frequent landslide includes highly weathered and structurally deformed rocks, continuous tectonic activity along regional faults and thrusts, high rainfall, the extension of human settlement in a hilly area, and infrastructure & road construction. In recent times, various engineers and scientists gained attention in identifying these influential factors. This study has done landslide hazard zonation mapping along the road section NH 1-A of Ramban Himalaya according to the rating

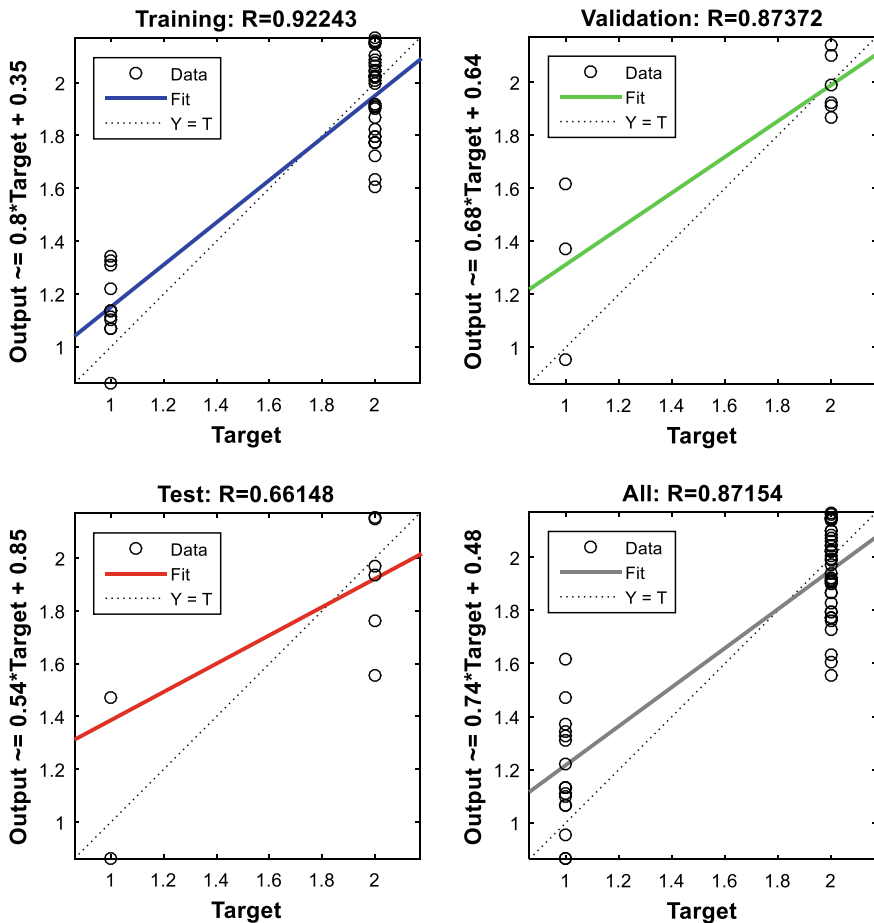


Fig. 19.7 A performance comparison for susceptible area prediction using BPNN model

laid in BIS code. The map has depicted that most of the region (~80.46%) in the study area falls under the high hazard category while the remaining region (~19.54%) lies under the moderate hazard category. The development of machine learning-based computation technologies has aided the researcher in developing intelligent models to predict landslide hazard zones. This study also covers a machine learning approach called back-propagation neural network (BPNN) for landslide hazard assessment. The cumulative effect, including causative and triggering factors of landslides, were considered for preparing the dataset of the study. The BPNN model’s performance has been investigated over the data set using training and testing. The validation of the performance has been examined based on three other machine learning models such as SVM, M5-rule, and M5P. The investigation revealed that the BPNN model shows a better performance in terms of high prediction accuracy and minimum error

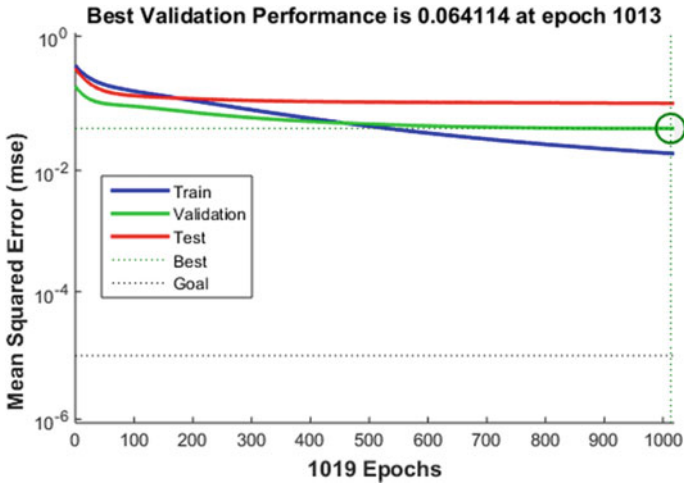


Fig. 19.8 A performance plot of the training errors, validation errors, and test errors

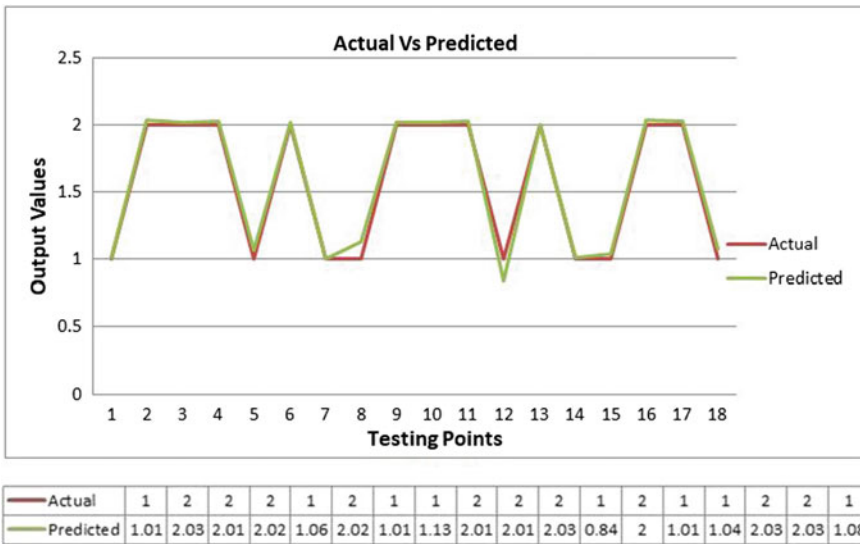
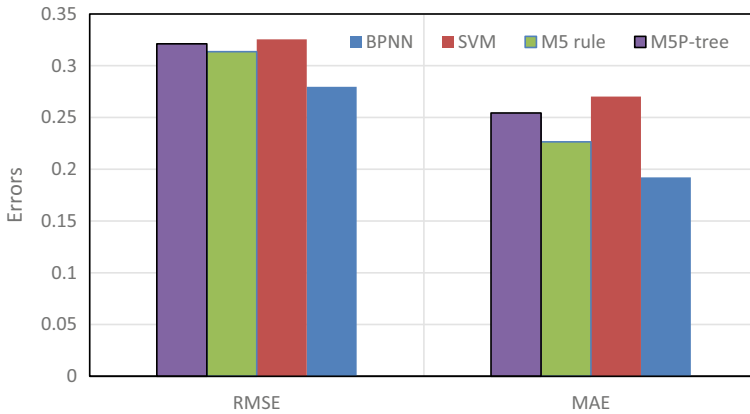


Fig. 19.9 Actual vs predicted landslide plot for the testing dataset

as compared to other discussed machine learning models. Based on the above observations, it is concluded that the proposed method can be used as a significant tool for the identification of landslide hazard zones and provides its mitigation measures. The implication of the study provides theoretical guidance to the researchers and engineers working in the Ramban to Banihal area to assist in developing infrastructure and construction & widening of national highways.



**Fig. 19.10** Comparison of MAE and RMSE over different machine learning methods

**Conflict of Interest** The authors wish to confirm that there are no conflicts of interest with this research article.

## References

- Anbalagan R, Chakraborty D, Kohli A (2008) Landslide hazard zonation (LHZ) mapping on meso-scale for systematic town planning in mountainous terrain
- Al-Najjar HA, Pradhan B (2021) Spatial landslide susceptibility assessment using machine learning techniques assisted by additional data created with generative adversarial networks. *Geosci Front* 12(2):625–637
- Arora MK, Gupta ASD, Gupta RP (2004) An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. *Int J Remote Sensing* 25(3): 559–572
- Bashiri M, Geranmayeh AF (2011) Tuning the parameters of an artificial neural network using central composite design and genetic algorithm. *Sci Iran* 18(6):1600–1608
- Benardos AG, Benardos AP (2005) Applications of artificial neural networks in geotechnology. *Chron Min Metall* 15(1):65–81 (in Greek)
- Benardos AG, Kaliampakos DC (2004) Modelling TBM performance with artificial neural networks. *Tunn Undergr Space Technol* 19(6):597–605
- Bhat GM, Pandita SK, Singh R, Malik M A, Sarkar S (1999) Northwest Himalayan Successions along Jammu–Srinagar Transect (Field Guide). Indian Association of Sedimentologists, Aligarh, India, 141 pp
- Bureau of Indian Standard (1998) IS: 14496, Preparation of Landslide Hazard Zonation Maps in Mountainous terrains – Guidelines, Part 2 Macro-zonation, BIS, New Delhi
- Chingkhei RK, Shiroyleima A, Singh LR, Kumar A (2013) Landslide hazard zonation in NH-1A in Kashmir Himalaya, India. *Int J Geosci*
- Cortes C, V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
- Dou J, Yamagishi H, Pourghasemi HR, Yunus AP, Song X, Xu Y, Zhu Z (2015) An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island Japan. *Nat Hazards* 78(3):1749–1776
- Fausett L (1994) Fundamentals of neural networks architectures, Algorithms and applications. Prentice-Hall, Upper Saddle River, NJ, p 461

- Holmes G, Hall M, Prank E (1999) Generating rule sets from model trees. In Australasian joint conference on artificial intelligence. Springer, Berlin, Heidelberg, pp 1–12
- Hosmer D, Lemeshow S (2004) Applied logistic regression. Wiley
- Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. *ACM SIGKDD Explorations Newslett* 11(1):10–18
- Haykin S (1999) *Neural networks. A comprehensive foundation*, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ, p 696
- 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. *Eng Geol* 85:347–366
- Khandelwal M, Singh TN (2009) Prediction of blast-induced ground vibration using artificial neural network. *Int J Rock Mech Min Sci* 46(7):1214–1222
- Khaw JF, Lim BS, Lim LE (1995) Optimal design of neural networks using the Taguchi method. *Neurocomputing* 7(3):225–245
- Lee S, Ryu JH, Lee MJ, Won JS (2003a) Landslide susceptibility analysis using artificial neural network at Boun Korea. *Environ Geol* 44:820–833
- Lee S, Ryu JH, Min KD, Won JS (2003b) Landslide susceptibility analysis using GIS and artificial neural network. *Earth Surf Proc Land* 27:1361–1376
- Lee S, Ryu JH, Lee MJ, Won JS (2006) The application of artificial neural networks to landslide susceptibility mapping at Janghung Korea. *Math Geol* 38(2):199–220
- Li Y, Chen G, Tang C, Zhou G, Zheng L (2012) Rainfall and earthquake-induced landslide susceptibility assessment using GIS and Artificial Neural Network. *Nat Hazard Earth Sys Sci* 12(8):2719–2729
- McClelland JL, Rumelhart DE, the PDP Research Group (1986) *Parallel distributed processing: Explorations in the microstructure of cognition*, vol II. MIT Press, Cambridge, MA
- Melchiorre C, Matteucci M, Azzoni A, Zanchi A (2008) Artificial neural networks and cluster analysis in landslide susceptibility zonation. *Geomorphology* 94:379–400
- Nefeslioglu HA, Gokceoglu C, Sonmez H (2008) An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Eng Geol* 97:171–191
- Pistocchi A, Luzi L, Napolitano P (2002) The use of predictive modeling techniques for optimal exploitation of spatial databases: a case study in landslide hazard mapping with expert system-like methods. *Environ Geol* 41:765–775
- Pradhan B, Lee S (2010a) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environ Earth Sci* 60:1037–1054
- Pradhan B, Lee S (2010b) Landslide susceptibility assessment and factor effect analysis: Backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modeling. *Environ Model Softw* 25(6):747–759
- Pradhan B, Lee S (2010c) Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland Malaysia. *Landslides* 7(1):13–30
- Quinlan, J. R. (1992) Learning with continuous classes. In 5th Australian joint conference on artificial intelligence, Vol 92, pp 343–348
- Shanker R, Kumar G, Saxena SP (1989) “Stratigraphy and sedimentation in Himalayas: A Reappraisal” geology and tectonics of Himalaya. *Geo Surv Ind Special Pub* 26:1–26
- Singh J, Verma AK, Haider B, Singh TN, Sachin M (2016) A study of soft computing models for prediction of longitudinal wave velocity. *Arabian J Geosci* 9(3): 224
- Singh B, Sihag P, Singh K (2017) Modelling of impact of water quality on infiltration rate of soil by random forest regression. *Model Earth Syst Environ* 3(3):999–1004
- Singh J, Haider B, Verma AK (2019) A BBO-based algorithm for slope stability analysis by locating critical failure surface. *Neural Comput Appl* 31(10):6401–6418
- Tortum A, Yayla N, Çelik C, Gökdağ M (2007) The investigation of model selection criteria in artificial neural networks by the Taguchi method. *Physica A* 386(1):446–468



- Witten IH, Wang Y (1997) Induction of model trees for predicting continuous classes. In Proc. Poster Papers Europ. Conf. Machine Learning
- Yilmaz I (2009) Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Comput Geosci* 35(6):1125–1138