Journal of Scientific & Industrial Research
Vol. 80, August 2021, pp. 716-725

Probabilistic Approach to Predict Landslide Susceptibility based on Dynamic Parameters for Uttarkashi, Uttarakhand (India)

Poonam Kainthura[1,2]* and Neelam Sharma[1]
[1]Banasthali Vidyapith, Jaipur 304 022, India
[2]University of Petroleum and Energy Studies, Dehradun 248 007, India
Received 19 January 2021; revised 30 March 2021; accepted 10 July 2021

The changing climate and global warming affect the stability of slopes, resulting in landslides. Landslides are frequent in hilly regions all over the world. The present work compares three GIS-based machine learning techniques to predict the changes in landslide susceptibility patterns classified as low, moderate, and high from observed records. The state-of-the-art methods include Random Forest (RF), Support Vector Machine (SVM), and Multinomial Logistic Regression (MLR). The landslide inventory contains a total of 1239 locations, which are divided into three subsets for training, testing, and validation purposes. A total of seven influencing factors were selected to understand the relationship between selected factors and observed landslides. The models were compared using the Receiver Operating Characteristics (ROC) curve and other statistical measures, including accuracy, precision, recall, sensitivity, and specificity. The RF model outperformed with the highest training ($RF_{Accuracy}$=91%), testing ($RF_{Accuracy}$=88%), and validation ($RF_{Accuracy}$=86%) accuracy. The ROC values computed for the validation dataset for three models are 0.749, 0.734, and 0.874 for the MLR, SVM, and RF models respectively. The outcome of the present study could be instrumental for policy and decision-makers concerning risk planning and mitigation.

Keywords: Decision making, Influencing factors, Machine learning, Receiver operating characteristics, Risk mitigation

Introduction

Landslides are a dangerous hazard that can be natural as well as artificial. These landslides commonly occur in hilly regions, which are highly sensitive. A landslide always induces loss of life and infrastructure in the hilly regions all over the world. In most cases, landslips are due to climatic and anthropogenic conditions. Predicting the change in landslide patterns is a challenging task for risk managers. Therefore, it is important to plan a suitable model to minimize the influence of landslides. Previous studies describe several GIS-based statistical and machine learning methods for dealing with landslide disasters. Different machine learning methods are used to map landslide susceptibility.[1] The susceptibility mapping depends on the influencing factors that, in combination, trigger landslides in hilly regions.[2] The factors that induce landslides can be varied for different study regions. Additionally, there can be different types of landslides that are analyzed individually.[3–5] There are various reasons these landslides occur, such as tectonic activities, natural slope failures, anthropogenic activities, and heavy and intense rainfall.[6–8] Landslide predictions and their causes have captured the attention of researchers throughout the past decades. Various studies in the past show the comparison of different techniques to analyze landslide problems.[9–11] Different comparative analysis exhibits that multivariate analysis is critical in comparison to bivariate analysis.[12] The researchers carried out various studies using a statistical and data-driven approach to understand landslide scenarios.[13] The major challenge today is to understand and predict landslides based on changing conditions. Mostly, geological and geo-morphological factors control changes in landslide patterns over time.[14] Researchers all over the world are conducting studies to analyze how changing factors are inducing landslide conditions. The Machine learning approach is proving highly powerful for understanding landslide scenarios.[15] The identification of landslides in changing environments can be helpful for risk managers to assess risk and plan mitigation.[16] The authors applied machine learning techniques support vector machine and adaptive neuro-fuzzy inference system to perform susceptibility analysis for Icheon township, South Korea. An Artificial neural network

*Author for Correspondence
E-mail: poonamkainthura@gmail.com

technique was used to build rainfall prediction model.[17] Random forest method was applied using sixteen variables to map landslide susceptibility.[18] Multiple logistic models were used to predict landslide disasters integrated with digitized geology, slopes, and geology factors.[19] The authors compared various machine learning techniques to analyze landslide susceptibility predictions, from which random forest outperformed.[20] The objective of this work is to present an analytical and prediction model capable of predicting changes in the landslide susceptibility pattern under dynamic conditions. Various authors conducted several studies related to the study areas to understand and predict landslide susceptibility. A SVM integrated with eight thematic layers was used to analyze and map landslide susceptibility for the Mandakini river basin, Gadhwal Himalaya.[21] The stability analysis for Balia Nalain Nainital, Uttarakhand, has revealed the major reasons behind the unstable slopes.[22] A comparative study was conducted to understand landslide susceptibility in the Himalayan regions using machine learning techniques.[23] The geo-morphological and geological terrain of the Himalayan belts are very complex, and geo-engineering projects such as roads and dams induce slope failures and landslides.[24] The literature shows that very few comparative studies have been conducted to map landslide susceptibility using multiple machine learning algorithms for the study region. Looking at the present landslide conditions in the Uttarakhand region, more comparative studies must be conducted to construct a more enhanced and interactive model that will predict landslide susceptibility on dynamically changing geological conditions. The lack of such models has enabled the authors to conduct this study by applying different machine learning models and GIS (Geographical Information System) techniques. Therefore, to bridge this research gap, the present study aims to identify the most influencing factors that induce landslides in the hilly regions of Uttarakhand. The relationships among the selected factors were used to predict the change in landslide susceptibility patterns under dynamic conditions using Random Forest (RF), Support Vector Machine (SVM), and Multinomial Logistic Regression (MLR) techniques. The patterns were divided into three classes: low, moderate, and high. Machine learning is the emerging technology to explore and analyze available landslide inventory to predict landslide changing patterns. However, achieving a hundred percent accuracy is a challenging task for researchers due to the frequent changing conditions of the study region.

Study Area

The study area, Uttarakhand, shown in (Fig. 1), is well known for its natural beauty and pilgrimage. Uttarakhand is the northern part of India located at 30.0668° N and 79.0193° E coordinates and surrounded by national and international boundaries. The state shares an international border with China in the North and Nepal in the East, and interstate boundaries with Himachal Pradesh in the West and Uttar Pradesh in the South. Uttarakhand consists of thirteen hill districts covering 53,483 km^2 geographical area and has diverse geographical features, ranging from snow-capped peaks in the north to tropical forests in the south. The climate and vegetation in the study region vary with elevation ranging from 190 –7816 m. The slopes in this region are unstable due to ongoing tectonic activities. The temperature varies from sub-zero to 43°C, and the average annual rainfall in the region is 1,550 mm. According to India's state forest record, the total forest area on paper is 34,651 km^2. The state's road network is divided into three types, i.e., National highways (NH), State highways, and major district roads. Many sacred rivers, like the Ganga and Yamuna, originate from the hills of Uttarakhand. However, the increase in anthropogenic activities in the study region is disturbing the natural stability of slopes, resulting in landslide situations. From the records, it is observed that Uttarakhand has always been prone to natural disasters in which landslides are the most frequent danger over this region.[25]

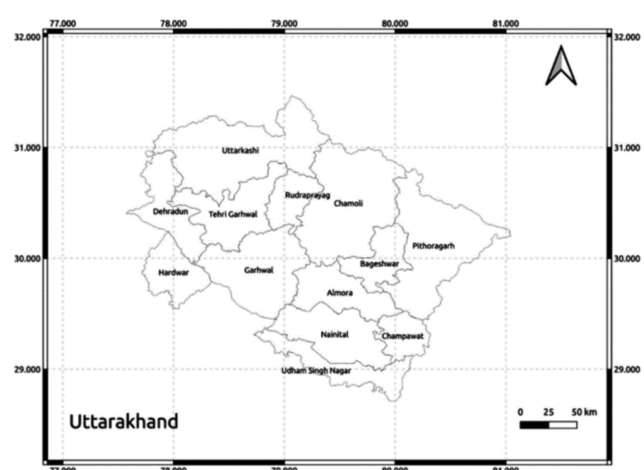

Fig. 1 — Study Area (Uttarakhand)

Materials and Methods

In the present study, the landslide susceptibility pattern prediction is achieved in four steps, as shown in Fig. 2: (1) Landslide data inventory preparation (2) identification of most influencing factors that induce landslides in the study area (3) construction of landslide susceptibility maps and prediction models, and (4) evaluation and comparison of landslide susceptibility prediction models. The steps are discussed as follows:

Landslide Data Inventory Preparation

Landslide inventory can be prepared as a primary source through field investigations or as secondary source directly from government agencies. For the present study, data was collected from reports prepared by the Geological Survey of India. The reports contain landslide locations (x, y coordinates) and some geological factors associated with landslides. The authors have converted the tabular data into shape files using the available landslide locations (coordinates). The landslide inventory also includes features of less importance and some missing values in the dataset. The model construction on such a dataset always results in unreliable predictions. Therefore, it is essential to clean the data before the construction of the prediction model. Data cleaning was achieved by filling in the missing values using the K-nearest neighbor impute method. Later, features of higher importance were selected by applying the correlation-based feature selection method. This method evaluates the value of an attribute by measuring the correlation between it and the target class. The relevance of features can be seen in Table 1.

A total of 554 recorded instances of landslides were finalized for the construction of the model. A sample of the landslide inventory is shown in Table 2.

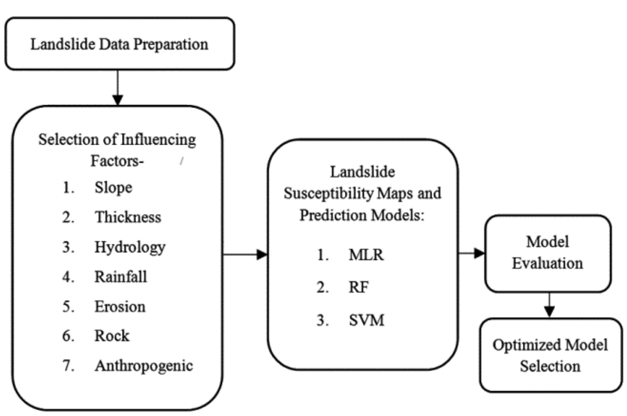

Fig. 2 — Workflow of the Proposed Methodology

Influencing Factors

In landslide susceptibility prediction, the identification of conditioning features is essential to avoid disturbing elements that degrade the predictive capability of models. These parameters are the reason behind landslides in the past. For the current study, seven landslide triggering factors were identified and integrated to predict the landslide changing pattern. The selected influencing factors are discussed below:

Erosion

Erosion is one of the critical causative factors that supports landslide conditions. Due to high-intensity rainfall, soil suction is reduced, which further reduces the strength of shear. Such conditions start surface erosion, bank erosion, and remove the toe that reduces slope support and induces landslides.

Slope Type

The influence of slope type on the distribution of landslides is of high importance. For the selected study region, the slopes are categorized as gentle (3°–5°), moderate (5°–8.5°) and steep (35°–45°). Generally, the slope falls during the rainy season and due to tectonic activities.

Anthropogenic Activities

Due to the increase in the human population and related anthropogenic activities, the complex

Table 1 — Feature Significance

Parameter	Rank
Erosion	0.2339
Slope Type	0.1930
Anthropogenic	0.1030
Hydrology	0.0813
Overburden Thickness	0.0713
Rock Char	0.0710
Rainfall	0.0537
Landslide style	0.0372
Geomorphology	0.0356
Landslide Material	0.0215
Landslide Movement	0.0172

Table 2 — Sample Landslide Inventory

Influencing Factor	Values			
Erosion	Yes	no	Yes	no
Slope Type	Steep	gentle	Moderate	steep
Anthropogenic	Yes	No	Yes	no
Hydrology	Dry	damp	Flowing	damp
Overburden Thickness (m)	0–1	2–5	>5	1–2
Rock Char	Fractured	Jointed	Massive	Sheared
Rainfall	Yes	Yes	Yes	No
Landslide	Yes	no	Yes	Yes

geological settings of the region and its natural state are disrupted. Landslides are the result of anthropogenic activities such as roads, deforestation, industrialization, etc.

Hydrology

Due to heavy and intense rainfall, the water flows through the fractured rocks, increasing the pore pressure and decreasing the shear strength of the slope leading to landslide conditions. Another reason could be internal seepage inside the hill, which activates erosion resulting in excessive surface run-off through drainage.

Overburden Thickness

Overburden is material made up of soil, debris, clay, and rocks. This overburdened material on slopes accumulates and leads to rotational or translational slides. In this work, the depth of this overburden ranges from 0–1 m to >5 m.

Rock Char

The hilly region of Uttarakhand is composed of different types of rocks, such as quartize, phyllite, etc. For multiple reasons, like seismic and anthropogenic activities rock in these regions is fractured and sheared.

Other characteristics of rocks are massive and jointed, and due to external factors like rainfall, earthquakes, weathering, these rocks move down, resulting in landslides.

Rainfall

Rainfall is the most influencing parameter that induces landslides. Uttarakhand state receives intense rainfall during the monsoons, and integrating with other geological factors, brings landslides in hilly regions. Records show that majority of the landslides occurred due to rainfall. The average rainfall it receives is 1069 mm.

Classification Methods

Various machine learning techniques exist to predict landslide susceptibility patterns. In the present study, three methods were selected based on the seven selected conditioning factors of the study area. The selected techniques are Support vector machine, Multinomial logistic regression, and Random forest. A total of 1239 instances were used to construct a landslide susceptibility map. The dataset was split into three subsets to build a susceptibility map, i.e. training, testing, and validation. For training and testing purposes, 554 landslide instances were divided into the proportion of 70:30. The remaining 685 locations were used for validation purposes. Firstly, the learning of the model is done by applying training samples, and then the trained model is tested using testing samples. Secondly, a validation dataset was used to validate the reliability of the model. The landslide susceptibility zones were classified into three classes (low, moderate, and high) using natural break method. The following subsections describe the models in detail:

Support Vector Machine

Support Vector Machine[26] is a learning technique that classifies s samples with n features. This technique is capable of analyzing linearly non-separable and multidimensional data sets. The Support vector machine model computes the optimal decision boundary to group the classes for multidimensional datasets. The best decision boundary can be calculated by:

$$f(X) = sign(\sum_{i=1}^{n} \alpha_i Y\, K(X_i) + b \tag{1}$$

where X_i is the vector of landslide conditioning parameters, $Y \epsilon$ *{+1,-1}* represents the vectors of the target class, α_i are constants, b states the bias value and $K(X_i)$ is a kernel function.

Random Forest

The Random Forest[27] is a widely used ensemble classification technique for landslide risk pattern prediction. This model combines multiple decision trees for classification, applying resampling to the dataset, and randomly changing the rules over the different trees. The Random forest method assumes an unweighted majority of individual decision tree class votes to predict the final class, as shown in Eq. (2).

$$RF(X) = arg_z\left(\sum_{i=1}^{k} I\ DT_i(\,x)\right) = out) \tag{2}$$

where, $RF(X)$ states Random forest model, DT_i is an independent decision tree, I is the indicative function, and out is the output variable.

To construct a random forest model, first individual decision trees DT_i were trained. The ID3 method was used to construct a decision tree. The ID3 uses the theory of information gain, which is expressed as:

$$Gain(A) = info(D) - info_A(D) \tag{3}$$

where, $info(D)$ and $info_A(D)$ are computed as:

$$info_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} info(D_j) \quad (4)$$

where $|D_j|/|D|$ denotes the weight of the j_{th} partition, v denotes the total number of divisions.

$$info(D) = -\sum_{i=1}^{m} p_i \, log_2(p_i), \quad (5)$$

where info (D) is the average information to identify target class, p_i denotes non-zero probability, m represents the number of classes.

Multinomial Logistic Regression

Generally, the logistic regression model uses binary dependent features equivalent to the presence (1) or absence (0) of landslide risk. Multinomial logistic regression[28] is an extension of binary logistic regression with more than two classes. For the present study, three categories, low, moderate, and high, were taken. MLR considers that the class values of the dependent variable are entirely different. It is a technique applied to predict the probabilities of various possible outcomes known as polychotomous for categorical dependent features based on independent features. This technique assigns records to the class based on the input variables which have the highest probability. The computation of MLR is based on the basic logistic regression formula expressed in Eq.(6).

$$\log\left(\frac{P(0)}{P(1)}\right) = a_0 + b_1 \times x_1 + \cdots b_1 \times x_n \quad (6)$$

where $\log\left(\frac{P(0)}{P(1)}\right)$ is reference point logit class; a_0 is the intercept; b_1 is the coefficient and $x_{1\ldots}x_n$ are independent variables.

Considering low as a reference class, the MLR model has computed the probability of three landslide risk classes as shown in Eqs (7–9).

$$P(Low) = \frac{1}{1+e^{a_0+b_1\times x_1+\cdots b_1\times x_n}+e^{a_0+b_2\times x_1+\cdots b_2\times x_n}} \quad (7)$$

$$P(High) = \frac{e^{a_0+b_1\times x_1+\cdots b_1\times x_n}}{1+e^{a_0+b_1\times x_1+\cdots b_1\times x_n}+e^{a_0+b_2\times x_1+\cdots b_2\times x_n}} \quad (8)$$

$$P(Moderate) = \frac{e^{a_0+b_2\times x_1+\cdots b_2\times x_n}}{1+e^{a_0+b_1\times x_1+\cdots b_1\times x_n}+e^{a_0+b_2\times x_1+\cdots b_2\times x_n}} \quad (9)$$

Model Performance and Evaluation Metrics

The authors applied the Receiver operating characteristics (ROC) curve and other statistical measures to the present study, namely accuracy, precision, recall, sensitivity, and specificity, to evaluate the models. ROC represents the curve that states the true positive (TP) percentage against false positive (FP) to analyze landslide patterns. The area under the curve (AUC) is used to compare the models and their predictive capabilities. The AUC values are negligible (<=50%) when the prediction rate is poor. On the other hand, AUC values are higher (50%–100%) when prediction capability is good. To evaluate and compare the prediction capability of selected models, the authors applied statistical measures expressed in the following Eqs (10–13):

$$ROC = \left(\frac{(\sum TP + \sum TN)}{P+N}\right) \quad (10)$$

$$Accuracy = \left(\frac{TP+TN}{TP+FP+TN+FN}\right) \quad (11)$$

$$Precision = \left(\frac{TP}{TP+FP}\right),\ Recall = \left(\frac{TP}{TP+FN}\right) \quad (12)$$

$$Specificity = \left(\frac{TN}{N}\right),\ Sensitivity = \left(\frac{TP}{P}\right) \quad (13)$$

where P represents the total number of landslides and N represents the total number of non-landslides; TP, FP, TN, and FN, represent true positives, false positives, true negatives, and false negatives. Furthermore, if the values of all considered metrics are higher (maximum 1), the prediction model will be confirmed as accurate and reliable. Accuracy is the total number of correct predictions divided by the total number of instances. Precision is the ratio of correct positive classifications to the whole positive categories. A recall is how many accurate predictions were retrieved. Sensitivity assesses model capability to predict true positive (TP) of each class label, whereas specificity assesses model capability to predict true negatives of each class label.

Results and Discussion

Landslide Susceptibility Map Construction

Three landslide susceptibility maps were generated for this work using GIS-based Support Vector Machine, Random Forest, and Multinomial Logistic Regression methods. The distribution of susceptibility classes (low, moderate, and high) and landslide points on the maps are shown in (Fig. 3). The distribution of these classes is done by applying the natural break[29] classification technique. Changes are highlighted in black in (Fig. 3a), and a combination of red and black in (Fig. 3b), which shows a few changes in the label of landslide susceptibility of MLP and SVM to the proposed best model, random forest. From the results, it is observed that the hilly regions of Uttarakhand are highly susceptible to landslides.

Landslide Model Evaluation

For the present work, three models, Support Vector Machine (SVM), Random Forest (RF), and Multinomial Logistic Regression (MLR), were applied to predict landslides and susceptibility patterns in the Uttarakhand state, India. The models were constructed by employing training together with the testing dataset and validation dataset. The performance of constructed training models, testing models, and validation models are compared using statistical evaluation measures described in Eqs (10–13). The comparison of three models using training and testing and validation datasets is shown in Tables 3–5. The Receiver operating characteristics

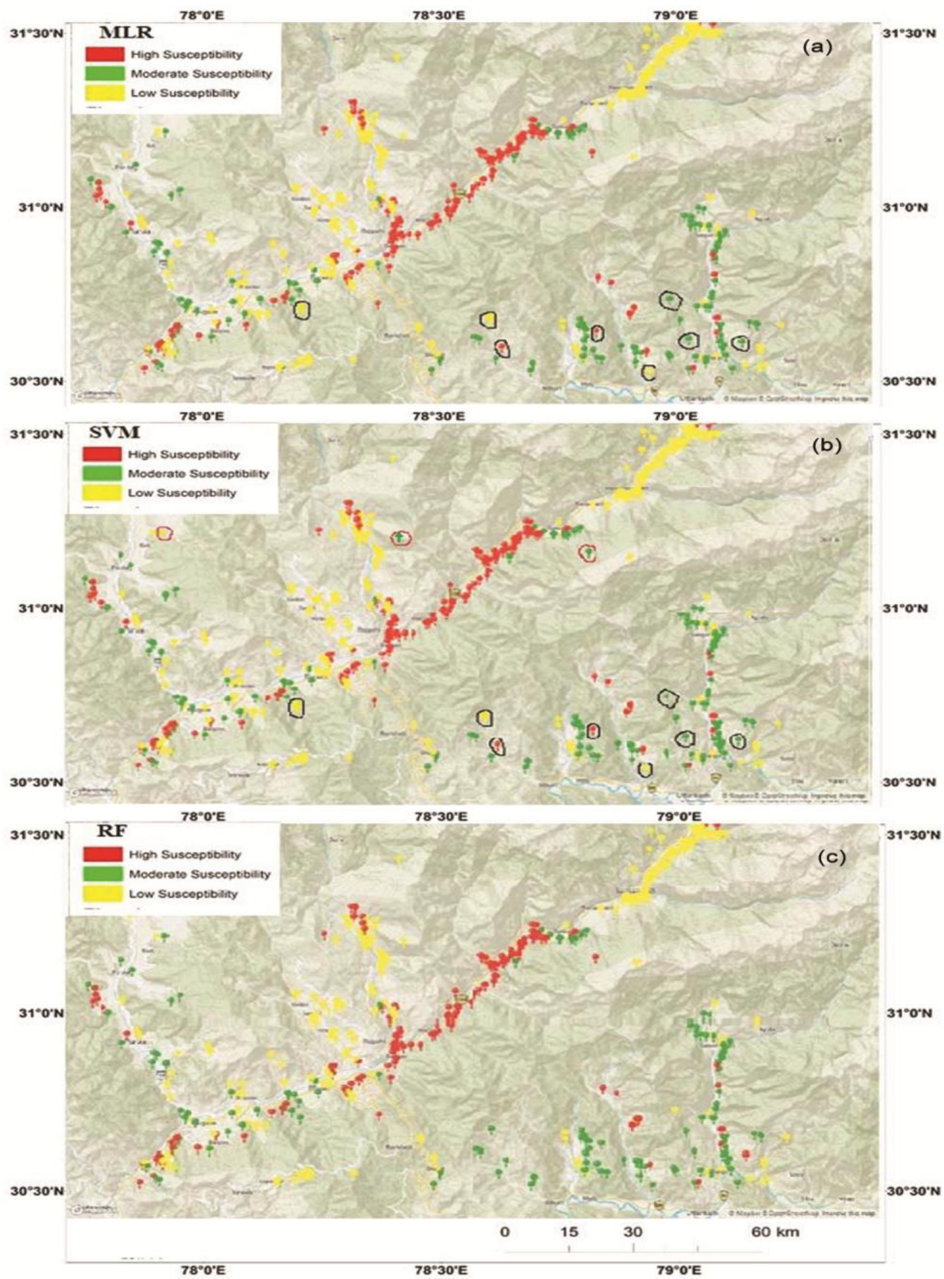

Fig. 3 — Landslide Susceptibility Maps for (a) Multinomial Logistic Regression (MLR), (b) Support Vector Machine (SVM), and (c) Random Forest (RF) classifier

(ROC) curves (Figs. 4–6) compare Sensitivity against Specificity for the three models on the training, testing, and validation datasets.

The results reveal that all three models can predict landslide susceptibility patterns. The models produced are acceptable, but the performance of the Random forest model is higher for training, testing, and validation datasets. For the random forest model, the RF_{ROC} = 0.960 and $RF_{Accuracy}$ = 91% for training, RF_{ROC} = 0.849 and $RF_{Accuracy}$ = 88% for the testing dataset and RF_{ROC} = 0.874 and $RF_{Accuracy}$ = 86% for the validation dataset, which is higher in comparison to other models.

Landslide Susceptibility Prediction

Once the susceptibility maps are prepared, the challenge is to predict changes in susceptibility patterns. Even though various techniques have been discussed and applied for landslide susceptibility predictions it is still difficult to identify the most appropriate technique for different study regions. In

Table 3 — Performance of Multinomial Logistic Regression (MLR), Support Vector Machine (SVM), Random Forest (RF) Classifier on training dataset

Performance Measure	Results		
	RF	SVM	MLR
Sensitivity	0.967	0.801	0.792
Specificity	0.104	0.225	0.206
Precision	0.966	0.800	0.790
Recall	0.966	0.799	0.792
ROC	0.960	0.852	0.853
Accuracy (%)	91%	86%	85%

Table 4 — Performance of Multinomial Logistic Regression (MLR), Support Vector Machine (SVM), Random Forest (RF) Classifier on test dataset

Performance Measure	Results		
	RF	SVM	MLR
Sensitivity	0.880	0.765	0.765
Specificity	0.245	0.444	0.459
Precision	0.880	0.754	0.754
Recall	0.880	0.765	0.765
ROC	0.890	0.783	0.769
Accuracy (%)	88%	76%	77%

Table 5 — Performance of Multinomial Logistic Regression (MLR), Support Vector Machine (SVM), Random Forest (RF) Classifier on validation dataset

Performance Measure	Results		
	RF	SVM	MLR
Sensitivity	0.859	0.739	0.741
Specificity	0.231	0.429	0.419
Precision	0.825	0.654	0.659
Recall	0.854	0.675	0.677
ROC	0.874	0.734	0.749
Accuracy (%)	86%	73%	75%

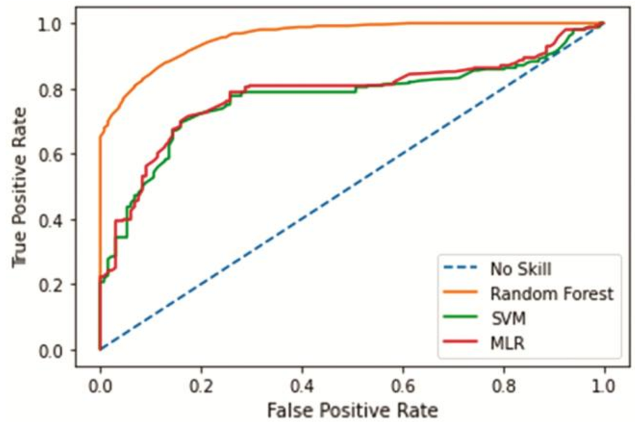

Fig. 4 —Performance comparison of Trained Random Forest (RF), Support Vector Machine (SVM), Multinomial Logistic Regression (MLR), classifier using ROC curve

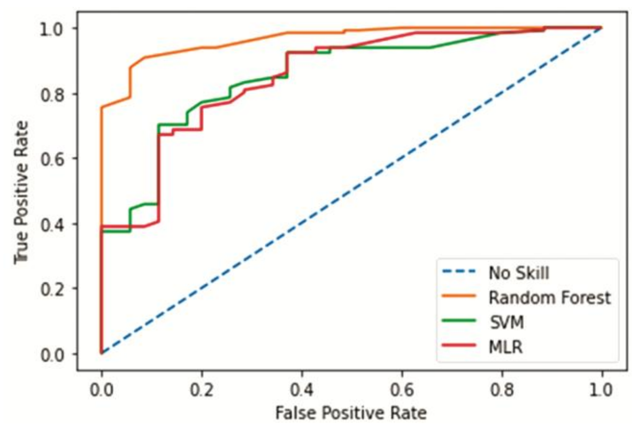

Fig. 5 — Performance comparison of Tested Random Forest (RF), Support Vector Machine (SVM), Multinomial Logistic Regression (MLR), classifier using ROC curve

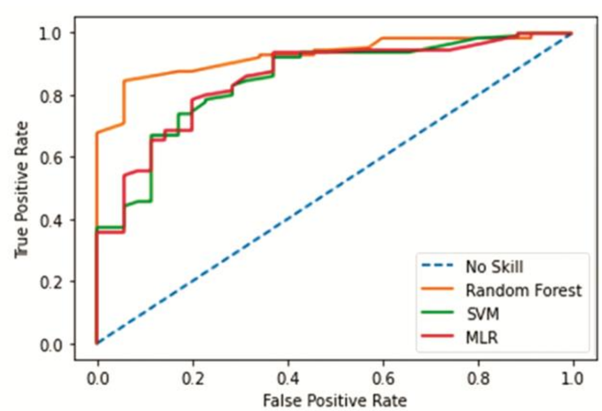

Fig. 6 — Performance comparison of Validated Random Forest (RF), Support Vector Machine (SVM), Multinomial Logistic Regression (MLR), classifier using ROC curve

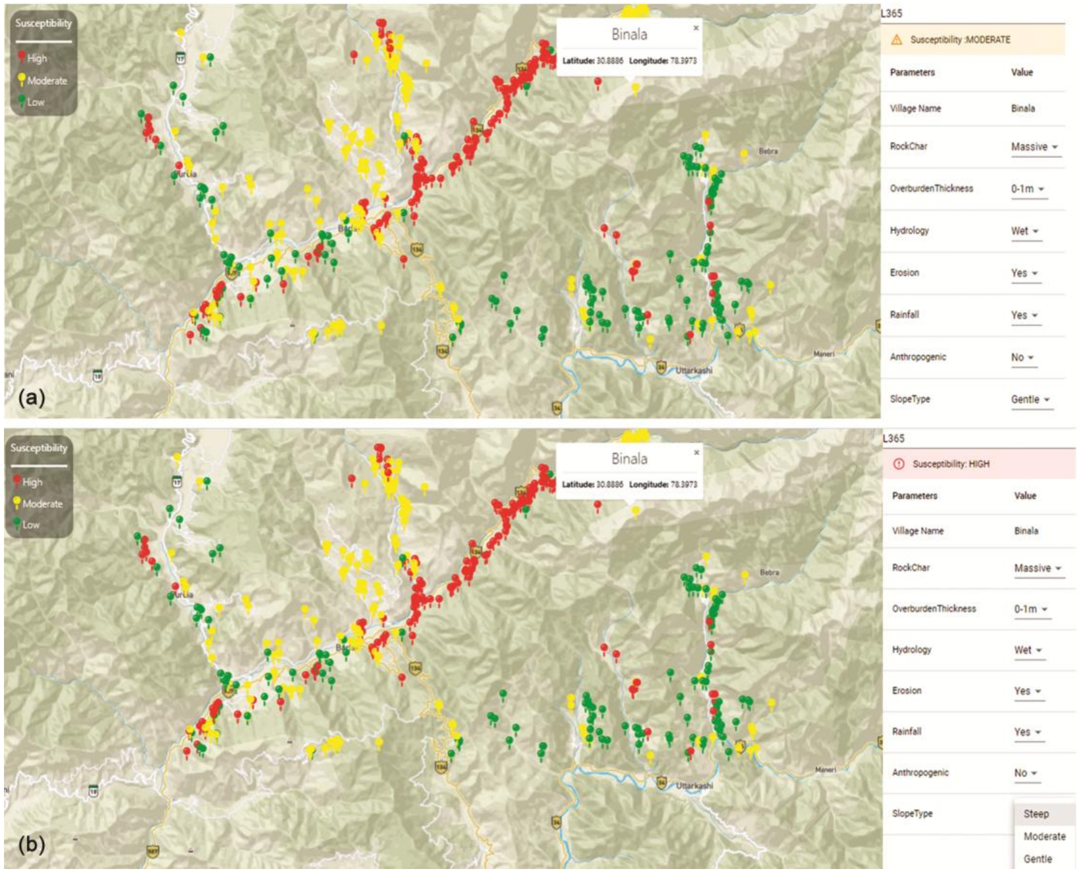

Fig. 7 — Landslide Susceptibility Prediction Using GIS-Based Random Forest (RF) model (a) Actual susceptibility and (b) Susceptibility with changed values of the parameter

the present work, three models, MLR, RF, and SVM, were applied, and the outcomes were compared to assess the best model to predict changes in landslide patterns. It is also observed that for better performance of the model, the quality of data matters. Modeling outcomes show that all three models performed well, but the RF model's predictive capability was best followed by SVM and MLR. The results of testing different models applying statistical measures show that the RF model has achieved the highest values. Considering the RF model as a final prediction model for the study area, a machine Learning and GIS-based system are designed to predict the landslide susceptibility based on changing values of the parameters. The authors used Open Street maps to predict the change in landslide susceptibility patterns to achieve such a system. In the present work, seven landslide influencing factors were considered for prediction. Each factor is correlated with other factors in some way, leading to landslides. A random forest model was finally applied to predict the change in susceptibility pattern. Using this interactive and dynamic system, the analyst can change the values of the parameters, and the system will indicate the new susceptibility class. The visualization of the changing susceptibility patterns can be seen in (Fig 7). On changing the 'Slope Type' factor from gentle to steep, the system alters the susceptibility from moderate to high. Using this system, analysts will easily find the hidden combinations of factors that together induce landslides to assist policymakers, risk analysts, risk management, and mitigation teams.[30,31]

Conclusions

From the results, it can be confirmed that the Random Forest method produced the highest accuracy rate in comparison to other models. Overall, the model proved a promising model for landslide susceptibility predictions. The results confirm that the Uttarakhand region is susceptible to landslides due to the geological and geo-morphological settings of the area. Other external factors, such as rainfall and anthropogenic activities, also play a major role in landslides. The impact of these factors is hidden, and identifying these hidden rules requires time. Therefore, the authors have constructed Machine learning and GIS-based interactive maps to predict the change in susceptibility of the Uttarakhand region. The benefit of such a system is for risk analysts who can interact with these systems. On feeding new input from the environment, the system will predict the probability of landslides. The study region is prone to landslides due to changing conditions in the background. The recent study contributes good knowledge to analyze landslide conditions for the selected region. Therefore, the system will be helpful to policymakers, risk analysts, risk management, and mitigation teams.

Acknowledgments

The authors are grateful to the Geological Survey of India, Dehradun, India, to provide the landslide dataset.

References

1 Shahabi H, Khezri S, Ahmad B B & Hashim M, Landslide susceptibility mapping at central Zab basin, Iran: A comparison between analytical hierarchy process, frequency ratio and logistic regression models, *CATENA*, **115** (2014) 55–70.

2 Jebur M N, Pradhan B & Tehrany M S,Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale, *Remote Sens Environ*, **152** (2014), 150–165.

3 Xie J, Uchimura T, Wang G, Selvarajah H, Maqsood Z, Shen Q, Mei G & Qiao S, Predicting the sliding behavior of rotational landslides based on the tilting measurement of the slope surface, *Eng Geol*, **269** (2020) 105554.

4 Schilirò L, Montrasio L & Mugnozza G S, Prediction of shallow landslide occurrence: Validation of a physically-based approach through a real case study, *Sci Total Environ*, **569–570** (2016) 134–144.

5 Shanmugam G & Wang Y, The landslide problem, *J Palaeog*, **4(2)** (2015) 109–166.

6 Carlini M, Chelli A, Vescovi P, Artoni A, Clemenzi L, Tellini C & Torelli L, Tectonic control on the development and distribution of large landslides in the Northern Apennines (Italy), *Geomorphology*, **253** (2016) 425–437.

7 Fustos I, Abarca-del-Río R, Mardones M, González L & Araya L R, Rainfall-induced landslide identification using numerical modelling: A southern Chile case, *J S Am Earth Sci*, **101** (2020).

8 Gill J C & Malamud B D, Anthropogenic processes, natural hazards, and interactions in a multi-hazard framework, *Earth-Sci Rev*, **166** (2017) 246–269.

9 Bui D T, Tsangaratos P, Nguyen V T, Liem N V, Trinh P T, Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment, *CATENA*, **188** (2020).

10 Merghadi A, Yunus A P, Dou J, Whiteley J, Pham B T, Bui D T, Avtar R & Abderrahmane B, Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance, *Earth-Sci Rev*, **207** (2020).

11 Huang F, Cao Z, Guo J, Jiang S H, Li S & Guo Z, Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping, *CATENA*, **191** (2020).

12 Schicker R & Moon V, Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale, Geomorphology, **161–162** (2012) 40–57.

13 Tordesillas A, Zhou Z & Batterham R, A data-driven complex systems approach to early prediction of landslides, *Mech Res Commun*, **92** (2018) 137–141.

14 Avanzi G D A, Giannecchini R & Puccinelli A, The influence of the geological and geomorphological settings on shallow landslides. An example in a temperate climate environment: the June 19, 1996 event in northwestern Tuscany (Italy), *Eng Geol*, **73(3–4)** (2004) 215–228.

15 Lu P, Qin Y, Li Z, Mondini A C & Casagli N, Landslide mapping from multi-sensor data through improved change detection-based Markov random field, *Remote Sens Environ*, **231** (2019).

16 Wang H, Zhang L, Yin K, Luo H & Li J, Landslide identification using machine learning, *Geosci Front*, **12(1)** (2021) 351–364.

17 Kashiwao T, Nakayama K, Ando S, Ikeda K, Lee M & Bahadori A, A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency, *Appl Soft Comput*, **56** (2017) 317–330.

18 Sun D, Wen H, Wang D & Xu J, A random forest model of landslide susceptibility mapping based on hyper parameter optimization using Bayes algorithm, *Geomorphology*, **362** (2020).

19 Ohlmacher G C & Davis J C, Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA, *Eng Geol*, **69(3–4)** (2003) 331–343.

20 Chen W, Xie X, Wang J, Pradhan B, Hong H, Bui D T, Duan Z & Ma J, A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility, *CATENA*, **151** (2017) 147–160.

21 Kumar D, Thakur M, Dubey C S & Shukla D P, Landslide susceptibility mapping & prediction using Support Vector Machine for Mandakini River Basin, Garhwal Himalaya, India, *Geomorphology*, **295** (2017) 115–125.

22 Kumar M, Rana S, Pant P D & Patel R C, Slope stability analysis of Balia Nala landslide, Kumaun Lesser Himalaya, Nainital, Uttarakhand, India, *J Rock Mech Geotech*, **9(1)** (2017) 150–158.

23 Peethambaran B, Anbalagan R, Kanungo D P, Goswami A & Shihabudheen K V, A comparative evaluation of supervised machine learning algorithms for township level landslide susceptibility zonation in parts of Indian Himalayas, *CATENA*, **195** (2020).

24 Pradhan S P & Siddique T, Stability assessment of landslide-prone road cut rock slopes in Himalayan terrain: A finite element method based approach, *J Rock Mech Geotech*, **12(1)** (2020) 59–73.

25 Kumar M, Rana S, Pant P D & Patel R C, Slope stability analysis of Balia Nala landslide, Kumaun Lesser Himalaya, Nainital, Uttarakhand, India, *J Rock Mech Geotech*, **9(1)** (2017) 150–158.

26 Huang Y & Zhao L, Review on landslide susceptibility mapping using support vector machines, *CATENA*, **165** (2018) 520–529.
27 Stumpf A & Kerle N, Object-oriented mapping of landslides using Random Forests, *Remote Sens Environ*, **115(10)** (2011) 2564–2577.
28 Du J, Glade T, Woldai T, Chai B & Zeng B, Landslide susceptibility assessment based on an incomplete landslide inventory in the Jilong Valley, Tibet, Chinese Himalayas, *Eng Geol*, **270** (2020).
29 Chen X, Chen H, You Y , Liu J, Susceptibility assessment of debris flows using the analytic hierarchy process method − A case study in Subao river valley, China, *J Rock Mech Geotech*, **7(4)** (2015) 404–410,
30 Ritu, Living with and responding to risk in the Uttarakhand Himalayas: A call for prioritizing lived experiences in research policy praxis, Int J of Disaster Risk Reduction, **48**, 2020.
31 Kala C P, Deluge, disaster and development in Uttarakhand Himalayan region of India: Challenges and lessons for disaster management, *Int J of Disaster Risk Reduction*, **8** (2014) 143–152.