SSRG International Journal of Civil Engineering
ISSN: 2348-8352/ https://doi.org/10.14445/23488352/IJCE-V12I1P106

Volume 12 Issue 1, 46-66, January 2025

Original Article

Assessing Slope Stability Reliability through Visual Exploratory Data Analysis and Machine Learning in Kumaon Region of Uttarakhand in India

Pratul Raj[1], Lal Bahadur Roy[2]

[1,2]*Department of Civil Engineering, NIT Patna, Bihar, India.*

[1]*Corresponding Author : pratul.raj162@gmail.com*

Received: 17 October 2024 Revised: 18 November 2024 Accepted: 04 December 2024 Published: 26 December 2024

Abstract - *Slope stability is essential when planning or building any structure or formation on a soil slope. Accurate slope stability analysis involves considering variability in the properties of soil. Several methods are available within these probabilistic frameworks for determining a slope's safety factor (F.S). Increasing the reliability and accuracy of F.S value calculations can enhance both the stability analysis and stabilization procedures. Researchers have used computational intelligence approaches to obtain high-precision values of F.S. This paper focuses on F.S estimation using various machine learning and computational intelligence methods for comparison. It used six soft computing techniques, namely: Decision Tree (DT), Linear Regression (LR), Support Vector Regression (SVR), Neural Network (NN), k-Nearest Neighbor (KNN), and Random Forest (RF). To improve prediction accuracy, these models accounted for variability in critical soil properties such as slope angle, internal friction, unit weight, cohesion, and slope height. The models were trained with data from field cases, with the safety factor being the output variable. Validations were done using Morgenstern-Price (MP) LEM and Geo-Studio 2018 software. Model performance was carried out in terms of the metrics developed, such as R^2, RMSE, MAE, MSE, and VAF. The LR model resulted in $R^2 = 0.9354$, RMSE = 0.0911, MAE = 0.0703, MSE = 0.0083, and VAF = 93.62%. The graphical analyses applied were ROC curves, actual-versus-predicted plots, and residual graphs, all of which showed that the LR model was effective.*

Keywords - *Machine learning, Slope stability, Factor of safety, Limit equilibrium methods, Predictive model.*

1. Introduction

Slope stability is a worldwide problem that arises not from natural conditions but is often inducted by human activities. It can occur at any scale and with various types of movements. A full understanding of the problem, irrespective of the design approach, is needed for accurate knowledge of likely causes and an understanding of the mechanisms that lead to instability.

Evaluating slope stability requires understanding the geological and geotechnical properties of the soils and rocks that affect resisting forces in the shearing directions [1]. Slope stability analysis forms an important topic in geotechnical engineering. Failures are dependent on specific factors, such as geometrical characteristics of slopes, soil type, stratification, groundwater, and seepage [2]. The standard method to evaluate slope stability is by the calculation of the safety factor. The grounds of instability can vary in any number of ways, and so the phenomenon presents itself under different characteristics depending on the conditions of the ground, [3] which made it crucial for large geotechnical projects such as open-pit mines, embankment dams, highways, canals, tunnels, and railways.

In geotechnical engineering, slope stability analysis plays a vital role, providing a crucial evaluation of the safety and dependability of both natural and artificial slopes. Due to its complexity and difficulty in determining input parameters for geotechnical analysis, [4] it is arduous to estimate soil stabilization precisely. The paper discusses the capability of regression [5] and ANN [6] in predicting slope stability in two-dimensional applications. It is a challenging task, but over the past twenty years, it has gained much more attention due to numerous researchers who have developed new models.

Traditional methods, such as LEMs, are still used in slope stability analysis. Among these methods, the Morgenstern-Price method is highly efficient as it considers a comprehensive approach to the force and moment calculation on the sliding surface. Here, machine learning methods, including both regression models and ANNs, also revolutionized the slope stability analysis. These methods further enhance the prediction of slope behavior through large

datasets and discoveries of complex relationships that exist between the different factors of stability [7, 8]. Some of the common algorithms that are used include support vector machines (SVM), decision trees, k-nearest neighbours (KNN) [9] and neural networks.

Probabilistic approaches are widely accepted. Field data and laboratory tests are used to analyze the variability of the characteristics in soils. For example, [2] used the annealing technique to determine the upper and lower bound values of parameters influencing slope instability, whereas [10] used the response surface approach coupled with FORM for determining the reliability index for the earth dam sections. Recent studies by Roy [11-13] showed examples of carrying out an analysis of reliability in determining the stability of soil slopes under different conditions.

Besides, several researchers have also employed machine learning models. These are multi-layer feed-forward ANNs for forecasting slope stability through parameters like geometric and shear strength parameters [14]. Other similar methods include the integration of ANNs to a predictive model by [15], in which adaptive sine cosine pattern search was successfully applied to find the safety factor for slopes under static or dynamic conditions.

There are methods for reducing dimensions, such as PCA and t-SNE [16], that have been applied in the data analysis for predicting slope stability. For example, models of Support Vector Machine (SVM) [17], Decision Tree (DT) [8], Random Forest (RF) [17], logistic regression (LR) [18], k-nearest neighbor (KNN) [19], and Neural Network (NN) [20] these techniques have been applied to enhance the accuracy of predictions, and different metrics have been used to evaluate the performance of the models.

The Lower Himalayas in Uttarakhand, India, mainly consist of low-grade metamorphic rocks, sedimentary formations, and a few crystalline rock types. These fragile, soft, and brittle sedimentary and low-grade metamorphic rocks are subjected to atmospheric conditions. A slope failure is typically caused by one or more of the following factors: slope profile, geotechnical properties, land use, slope material strength, unplanned urbanization, geo-hydrological conditions, earthquakes, weathering, high rainfall, disturbances, and lithological all contribute to the creation of weak zones. Landslides on hill roads impede the provision of everyday commodities to uphill locations as well as transportation movement to and from them. Furthermore, the problem of slope failures or landslides is ongoing due to its annual occurrence.

A survey of the literature reveals that there has been little research on landslide-prone zones despite their importance as major geological hazards along highway cut slopes that pose serious dangers to infrastructure and safety.

In light of the foregoing, the threat of landslides must be handled more comprehensively with novel yet simple procedures that are not only easily available, economical, and efficient but also easy to apply. Researchers discovered that traditional methods underestimate soil property variability, whereas slope stability reliability analysis using machine learning techniques indicates reduced variability in results. Further testing is needed to determine its applicability in this field.

The paper aimed to carry out the reliability analysis of soil slope stability using the models of DT, RF, KNN, NN, LR, and SVR. Although the traditional methods have been in existence for a long period, the thorough literature review revealed that they do consider the variability of characteristics of the soil, which is not accounted for in those methods. Among the prediction models, Linear Regression emerged as an effective one, and model performance was carried out in terms of the metrics developed, such as R^2, RMSE, MAE, MSE, and VAF also in the graphical analyses applied were ROC curves, actual-versus-predicted plots, and residual graphs, all which showed that the LR model was effective For each of these models, the reliability assessment, with regard to the soil slope stability was carried out.

2. Modelling

Intelligent algorithms change our attitude towards complex problems by processing heavy amounts of data quickly, adapting to changing environments, and giving solutions at unmatched speed with precision. Their primary strength lies not only in the large dataset that they can handle but also in learning through those data and improving over time. These algorithms are essential in machine learning for decision-making, classifications, and predictions across various fields.

2.1. Machine Learning

This subset of artificial intelligence, often referred to as machine learning, refers to designing an algorithm in such a manner that computers may learn from data and make predictions or decide without explicit programming [21]. These algorithms are pattern recognition-based and relationship-based and tend to form trends within data that make them improve over time. The process can either be supervised, where models learn from labeled data, or unsupervised, where the data are unlabeled, and the model discerns hidden patterns. From healthcare and finance to marketing and engineering, machine learning algorithms are being applied in various sectors across the world.

2.1.1. Random Forest

Random Forest is an ensemble learning method in which learning is performed by generating many decision trees during training to improve prediction accuracy. It differs from a single decision tree, which suffers from overfitting- the tendency to memorize training examples instead of

generalizing. Random Forest has minimized that problem by developing many trees and then combining the predictions of all the trees. In classification, it produces the class that appears most frequently in all the predictions, while in regression, it averages the predictions [17]. In a Random Forest, each of the trees is trained on a randomly chosen subset of both the data and the features. This automatically introduces variability, thus increasing the robustness of the model. The approach is specifically good for avoiding overfitting, as no single tree dominates in the prediction process. Random Forest is used very widely for tasks such as:

- Classification: Assigning labels to the objects based on features.
- Regression: Predicting the continuous outcome.
- Feature Importance: It finds which features best predict the values of each response variable and how those important factors contribute to its predictions.

For the experiment, the Random Forest algorithm used all the default parameters, that is, 100 decision trees, based solely on Mean Squared Error (MSE) as the standard for choosing which node splits when building a tree. This flexibility hugely contributes to why Random Forest is such a powerful tool in machine learning and can be used in numerous applications of image recognition to medical diagnoses.

2.1.2. Decision Tree (DT)

A decision tree is a diagram that models decision-making, where each node represents a choice based on a specific feature or attribute, and each branch illustrates a possible outcome or result from that decision. Finally, at the leaves of the tree, final decisions or predictions are presented [22].

- Decision Trees is a widely applied versatile tool in both classification and regression. Its simplicity and interpretability are a reason for its popularity in most applications, such as:
- Classification: In decision trees, the data points are classified by classifying them through nodes and branches and assigning them to distinct categories on the basis of their features.
- Regression: Decision Trees also employ a prediction for continuous values; the value predicted is represented by a leaf node.
- Decision Support Systems: Illustrate and explain complex decision-making processes using representations of what may happen.
- Strategic Decision-Making: Used in business environments to model complex decisions and the consequences that occur as a result.
- Medical Diagnosis: Applied in trying to understand patients' symptoms and test results in order to determine the nature of the medical condition of concern.

We use the Decision Tree Regressor in scikit-learn with all its default parameters; the maximum depth is unlimited, and MSE is the splitting criterion. One of the most significant advantages of Decision Trees is their simplicity, but be careful not to overfit them since it can be done either by pruning the tree or limiting depth.

2.1.3. k-Nearest Neighbour (k-NN)

The k-Nearest Neighbours (k-NN) algorithm is a straightforward and efficient machine learning technique that makes classifications or predictions by evaluating the closeness of nearby data points. The parameter "k" represents the number of neighbours taken into account, serving as a crucial hyperparameter that influences the model's performance [19].

k-NN assumes that it has no model a priori. Also, it does not assume knowledge of any data's underlying distribution, meaning that k-NN is a non-parametric algorithm. Rather, k-NN memorizes the whole training set, and it bases its decisions solely on the similarity-difference (distance) of the new data point to the ones it has memorized. Although k-NN is successful, with large datasets, it becomes computationally costly.The algorithm is applied widely in the following areas:

- Classification: It assigns a class to a new data point based on the majority class of its closest neighbours.
- Regression: Here, the continuous variable is predicted based on averaging the values of the nearest neighbours.
- Anomaly Detection: This also identifies outliers by computing the distance from other points.

In this experiment, the default k-NN Regressor is used with five neighbours. Of course, the main advantage of k-NN is also its greatest weakness-independence from feature engineering, which makes the selection of the distance metric and the value of "k" highly critical to model performance.

2.1.4. Support Vector Regression (SVR)

SVR are robust algorithms used for both classification and regression tasks. SVR identify an optimal hyperplane that divides data into classes, maximizing the margin, which is the distance between the hyperplane and the nearest data points, known as support vectors [8]. This margin maximization enhances the model's accuracy and generalization capability. Support Vector Machine (SVM) is highly useful for high-dimensional space and allows working with the nonlinear relationship only provided by proper utilization of kernels. The applications of kernels like the RBF allow the SVM to transform its given data into a higher dimension in which there is a possibility of the existence of linear separation. The SVR is an application of the principles of SVM but tailored to regression, whereby the intention is to get an appropriate hyperplane that passes through the data points and minimizes the error.

The algorithm is used, for example, in applications such as:

- Image Recognition: In classifying images through distinguishing different classes.
- Text Classification: This simply means sorting documents based on content.
- Bioinformatics: In the classification of biological data, especially in medical research.

In this paper, the default settings used are that of SVR with RBF kernel. These settings produce a flexible and powerful approach to regression tasks.

2.1.5. Linear Regression (LR)

Linear regression is the basic statistical technique widely applied in machine learning as well as in standard statistical analysis. It forecasts the output of the dependent variable by taking into account the linear connection between one or more independent variables and the dependent variable [18].

Logistic regression is a form of linear regression; it is well suited, particularly for binary classification problems. The output is transformed using the logistic function, meaning the resulting predictions will always be between 0 and 1. This makes logistic regression very convenient to use in applications where the outcome variable is dichotomous-for, for example, yes/no or true/false. Some other applications include spam detection, like whether an email is spam or not.

Medical Probability: The chance or likelihood that a patient has some given disease.

In this paper, we apply Linear Regression as an elementary but effective technique in predictive analytics.

2.1.6. Artificial Neural Networks (ANN)

ANN simulates the human brain with interconnected neurons learning data patterns [23]. They can work in unison to identify patterns, make predictions, or solve complex tasks like classification and regression [3]. Each neuron basically processes the input data and passes the result to the next layer so that the network can learn complex patterns based on its training on large datasets.

ANNs are highly flexible and have found use in many different applications, some of which include:

- Image and Speech Recognition: Pattern recognition in images and voice.
- Natural Language Processing (NLP): Comprehension and generation of human language.
- Predictive Modeling: Forecasting future events from historical data.

In this paper, an MLP regressor is implemented using only a single layer consisting of 100 nodes with ReLU as the activation function. The stochastic gradient descent was applied for optimization. This algorithm allows one to control the updating of the network's weights from the correct values based on the network's estimation errors.

The models thus developed using these machine learning algorithms increase the accuracy and efficiency of predictions [24]. It takes care of large datasets and complex relationships, so it has become an essential tool for research in various industries. Every machine learning method has inherent limits. Random Forest (RF) models, while resilient and capable of handling big datasets, are computationally costly and may overfit if not correctly tweaked. Logistic Regression (LR) presupposes a linear relationship between features and output, which limits its ability to capture complicated, nonlinear patterns.

The Support Vector Machine (SVM) is very sensitive to kernel selection and can be inefficient for large datasets because of its high computational complexity. K-Nearest Neighbour (KNN) has high memory utilization and becomes computationally demanding as the dataset grows.

Neural Networks (NN), while very flexible and capable of learning complicated patterns, require a vast amount of data, have long training cycles, and are prone to overfitting if not properly regularized. Decision trees (DT), while simple to read, tend to overfit the data, making them less generalizable unless pruned or utilized in ensemble approaches such as RF.

3. Study Area

The study area is central in the Kumaon Himalaya region of the state of Uttarakhand. This area falls under the Lesser Himalayan terrane. This section geologically holds the two important tectonic elements here: Ramgarh Thrust, RT, and North Almora Thrust, NAT [23, 25]. These thrusts stratify the sediments in the Kumaon section of the Lower Himalayas. A study area is along an elevation of 6 km length that stretches west from the Chhara landslide site in the Almora District to Kakri Ghat.

This region is highly prone to landslides, partly due to unfavorable geographical and geological conditions, which cause heavy rain. The entire Himalayan region falls under seismic zones IV and V, as depicted in the seismic zoning map for India. The specific NH-109 section, which is under study, [26] comes under Zone IV (IS Code 1803 1070, Part I: 2002), which means there is an extreme risk of earthquake.

Climate From Chhara to Kakri Ghat, NH 109, is relatively warm from mid-April to mid-July, with an average annual temperature of 17°C (PWD). June is the hottest month, with an average temperature of 26.1°C, while January is the coldest, with temperatures dropping during the period of mid-December to mid-February. In this region, as per the Indian

Metrological Department (IMD), during the monsoon season, the amount of rainfall is usually recorded at approximately 1130 mm per year, causing recurrent landslides in that area.

In this region, National Highway 109 (NH-109) acts as a critical link, connecting the top half of the Kumaon region, including Almora, Bageshwar, and Pithoragarh (hilly areas), to the lower section of Kumaon, particularly Haldwani city, which is located in the plains and has substantial medical facilities. This route is critical to enabling access to important services and transportation for the region's residents. However, during the rainy season each year, the mountainous terrain becomes more prone to landslides as the soil loses stability.

The loosened debris from the slopes frequently slides down onto the roadway, causing road jams and increasing the danger of accidents. Furthermore, these landslides cause extensive damage, including the loss of agricultural land, human lives, fertile topsoil, and property. This recurring issue underlines the critical need for proper slope stabilization and erosion control measures to protect both the highway and the local population's livelihoods.

Geological layout: It is better visualized by referring to Figure 1, which displays geological maps of the region.

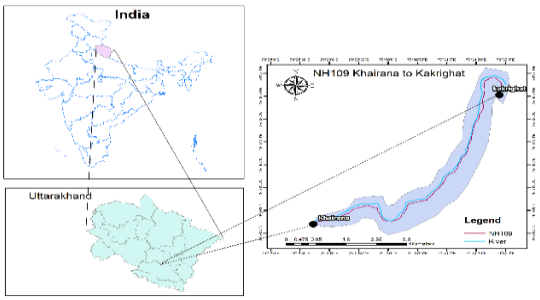

Fig. 1 Map of the study area

4. Slope Field Case Collection and Data Analysis

Slope field case collection means collecting actual data to draw the slope of a differential equation at various points. The data thus collected gets analyzed to determine the trends and find further behaviours of the system that are under consideration while using the differential equation.

4.1. Slope Field Case Collection

For this study, soil samples were collected en route from Chhara village to Kakri Ghat. From the pre-site analysis, it was decided to represent the different types of soils along the road and select 13 locations for investigation. Altogether, 101 pits were excavated down to a one-meter depth, as shown in Figure 2, to gather representative samples and observe conditions below the surface.

Fig. 2 Sample collection for the different Sites

The disturbed samples of soils collected were taken to the laboratory [27] with their unit weight, grain size distribution, and shear strength determined through a direct shear test as indicated in Table 1 performed in the laboratory and the results. These are the input parameters used in the analysis and find out the FOS by using the GeoStudio 2018 Slope/W software carried out in the Limit Equilibrium Method (LEM) Supported by Laboratory Tests End.

The [28] cohesion and internal friction angle are derived from the remolded soil samples using a drained test from the apparatus of the Direct Shear Test. Soil unit weight was determined using a calibrating cylinder, while grain size distribution helped classify soil types.

In other words, each soil sample is a case study in slope engineering incorporating five major slope characteristics as independent factors, while the stability of the slope is evaluated using a Factor of Safety (FOS) as the dependent factor. Table 1 highlights the input parameters used for Geo-Studio software analysis alongside the laboratory test results.

4.1.1. Shear Strength Parameters (c, φ)

Shear strength parameters were determined using the laboratory DST apparatus. These samples were remolded to their original weight and moisture content to give the field conditions. For every sample, four tests were conducted. Shear stress was applied under different vertical loads up to failure. Shear stress was plotted against normal stress by a graph on the Y-axis, with the X-axis showing normal stress [29]. From this graph, cohesion (c) and the friction angle (φ) of each of the 101 representative samples were calculated.

4.1.2. Method of Analysis

Then, the Slope/W program modeled the geometry of the slope, and the [30] Morgenstern-Price (MP) Method determined the FOS calculated as a ratio of the driving forces or moments to the resisting forces or moments based on a two-dimensional equilibrium condition. This limit equilibrium analysis has been performed with GeoStudio 2018 Slope/W software.

$$\text{FOS}=\frac{\textbf{Driving Force (moment)}}{\textbf{Resisting Force (moment)}} \quad (1)$$

The Morgenstern-Price (1965) approach achieves force and moment equilibrium by using an inter-slice force function f(x) and a scaling factor λ. It takes into account both inter-slice forces, allowing for the selection of the force function, and calculates the factor of Safety (FOS) for force and moment equilibrium.

Geo-studio 2018 (Slope/W) helps with slope stability analysis by merging different analyses into a single model, allowing for construction sequence modelling, sensitivity assessments, and complex problem decomposition. Using the limit equilibrium method, slopes are divided into slices to ensure equilibrium per the Mohr-Coulomb failure criterion, hence improving knowledge of slope behaviour.

4.2. Data Pattern Exploration

Preliminary data analysis is essential before developing prediction models. This is where one can figure out very important aspects like the integrity of data, distributional patterns, and relationships among factors, and all these found during this step would be foundations for selecting the correct model.

4.2.1. Data Integrity

Table 1 reports 101 samples, each of which comprises five independent factors or characteristics and one dependent factor or indicator that will be useful for the evaluation of slope stability. The data was complete since all the samples had a valid Factor of Safety. In the MP method, 56 samples had an FOS equal to or more than 1, which means stability, while the remaining samples 45 had an FOS less than 1, and hence indicate slope failures. The ratio between these two groups is found to be nearly 1:1.244; it will indicate a fair, uniform distribution of the values.

To visually inspect the integrity of data, violin plots are illustrated for relevant parameters such as angle of internal friction, cohesion and angle of slope; see Figures 6, 5 and 4. For example, Figure 4 provides two violin plots with stable slopes and slope failures for the various angles of the slope. A white rectangle within each plot indicates the median of the characteristic, while the box represents the lower and upper quartiles [22]. The 95% confidence interval is a thin black line within each of the boxes. This violin reflects the kernel density estimation for each characteristic [22].

The results validated that the data were complete, well-distributed, and normally distributed; hence, it was smooth sailing for further analysis based on this dataset's integrity.

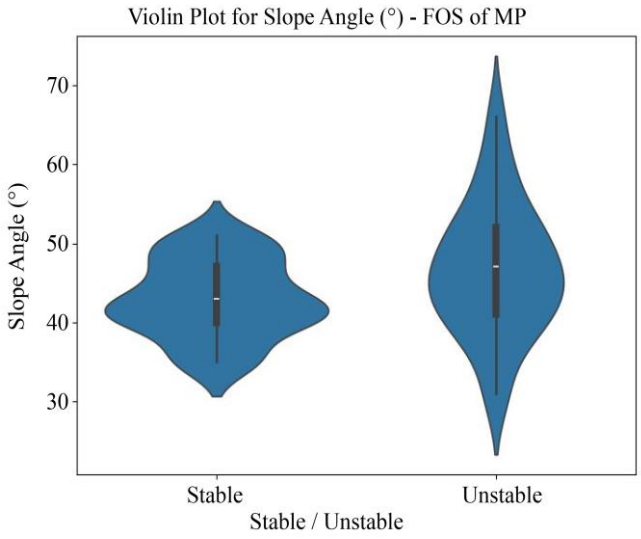

Fig. 3 Violin plot for slope angle

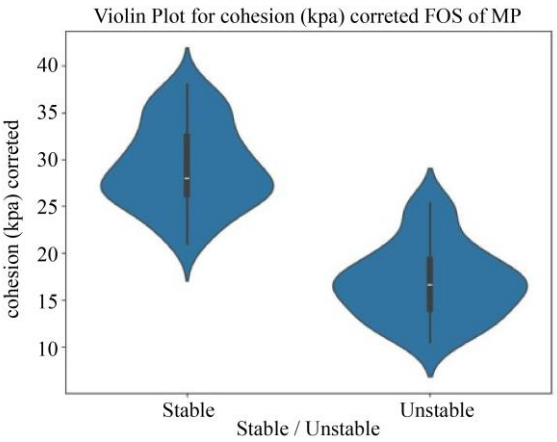

Fig. 4 Violin plot for cohesion

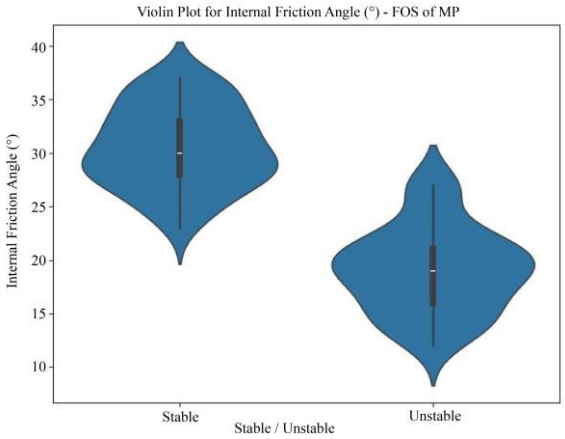

Fig. 5 Violin plot for internal friction

Table 1. Field case of slope engineering

S. No.	Unit Weight (kN/m^3)	Cohesion (kpa)	Angle of Internal friction (◦)	Slope Angle (◦)	Slope Height (m)	MP FOS	Soil Types
1	15.21	21.87	24	39	28	1.191	SP
2	15.5	21	23	39	28	1.13	SW
3	16.07	24.5	26	39	28	1.289	SP
4	16.41	26.25	28	43	39	1.118	SP
5	15.4	15.75	17	43	39	0.67	SP
6	16	27.12	29	43	39	1.173	SP
7	14.56	31.5	32	43	39	1.388	SP
8	13.72	29.75	31	51	53	1.045	SP
9	15.22	25.37	27	51	53	0.854	SP
10	15.75	30.62	32	51	53	1.027	SP
11	13.95	24.5	27	51	53	0.871	SP
12	17.1	23.6	26	43	43	0.979	SW
13	16.06	28	30	43	43	1.187	SW
14	18.3	28.87	29	43	43	1.115	SW
15	16.43	21	22	53	60	0.619	SW
16	16.7	18.37	19	53	60	0.53	SW
17	17.19	19.25	22	53	60	0.594	SW
18	18.4	28	30	44	45	1.136	SW
19	18.55	17.5	20	53	60	0.525	SW
20	15.65	26.25	28	44	45	1.111	SW
21	18.25	13.13	14	44	45	0.507	SW
22	18.03	17.5	21	44	45	0.745	SW
23	16.99	33.25	34	50	48	1.13	SP
24	13.51	36.75	37	50	48	1.359	SW
25	14.02	30.62	33	50	48	1.142	SW
26	14.6	35.87	36	50	48	1.28	SW
27	14.63	31.5	33	50	48	1.137	SP/SU
28	13.74	34.12	36	50	48	1.285	SW
29	16.7	35	35	50	48	1.222	SW
30	17.08	35.87	36	50	48	1.216	SW

31	14.79	20.13	22	43	45	0.85	SP
32	14.9	26.25	27	43	45	1.081	SW
33	14.22	36.75	37	43	45	1.595	SP
34	14.56	24.5	27	43	45	1.065	SP
35	15.2	15.75	18	43	45	0.672	SW
36	13.61	10.5	13	43	45	0.485	SP/SU
37	16.08	17.5	20	45	43	0.716	SP/SU
38	13.8	20.12	21	45	43	0.818	SP/SU
39	15.09	15.75	18	45	43	0.654	SP/SU
40	14.55	25.75	27	45	43	1.052	SP
41	15.44	14	16	39	43	0.362	SP/SU
42	15.7	13.12	17	39	70	0.583	SP/SU
43	15.02	10.5	12	39	70	0.423	SP
44	15.79	12.25	14	39	70	0.489	SP
45	14.95	16.62	20	37	60	0.78	SP/SU
46	14.8	28	30	37	60	1.261	SP/SU
47	15.6	27.12	29	37	60	1.197	SP/SU
48	16.38	21	23	37	60	0.907	SP/SU
49	16.06	23.62	25	41	40	1.028	SP
50	16.7	30.62	32	41	40	1.342	SP/SU
51	14.37	25.37	27	41	40	1.163	SP/SU
52	14.89	13.12	15	41	40	0.6	SP/SU
53	16.45	12.25	14	31	30	0.753	SP
54	16.15	14.87	17	31	30	0.926	SP
55	14.95	14	16	31	30	0.893	SP
56	15.31	23.18	26	35	32	1.327	SP
57	15.2	22.75	25	35	32	1.284	SW
58	16.1	26.25	28	35	32	1.44	SP
59	16.31	27.12	29	35	32	1.49	SP
60	15.34	24.5	26	35	32	1.353	SP
61	15.92	29.75	31	47	45	1.133	SP
62	14.61	34.12	34	47	45	1.322	SP
63	13.62	31.5	33	47	45	1.286	SP
64	15.11	28	30	47	45	1.101	SP
65	15.87	32.37	33	47	45	1.229	SP
66	14	26.25	28	47	45	1.053	SP
67	16.82	25.37	27	47	45	0.945	SW
68	16.2	30.62	32	47	45	1.168	SW
69	16.4	28	30	47	45	1.07	SW
70	18.1	24.5	27	47	45	0.912	SP
71	16.9	21	23	47	56	0.747	SW
72	17.2	30.62	33	40	45	1.352	SW
73	18.24	28.87	31	40	45	1.237	SW
74	18.4	19.25	22	40	45	0.828	SW
75	18.75	27.12	29	40	45	1.138	SW

76	18.32	14	19	66	70	0.336	SP
77	18.17	17.5	20	66	70	0.378	SW
78	17.18	35	35	41	38	1.668	SW
79	13.4	27.12	29	41	38	1.33	SW
80	14.21	15.75	19	41	38	0.788	SW
81	14.8	32.37	33	41	38	1.512	SW
82	13.98	35	35	41	38	1.668	SW
83	16.97	35.87	36	41	38	1.6	SP
84	17.2	37.96	37	41	38	1.663	SW
85	14.62	27.12	29	41	38	1.288	SW
86	14.01	37.62	37	41	38	1.793	SP
87	14.62	26.25	28	41	38	1.244	SP
88	13.5	11.37	13	52	55	0.397	SP/SU
89	16.41	18.35	21	52	55	0.607	SP/SU
90	13.94	16.6	19	52	55	0.58	SP/SU
91	14.92	21.87	24	35	35	1.213	SP/SU
92	14.46	26.25	28	35	35	1.469	SP
93	16.54	14	16	59	60	0.389	SP
94	16.34	16.62	19	59	60	0.453	SP
95	15.12	15.75	16	59	60	0.406	SP
96	16.38	16.62	19	59	60	0.453	SP
97	16.23	18.31	21	48	40	0.718	SP
98	15.02	17.5	20	48	40	0.702	SP
99	16.6	11.81	14	47	51	0.438	SW
100	16.01	12.25	14	47	51	0.45	SP
101	14.51	16.62	19	48	40	0.672	SP

4.2.2. Data Distribution Characteristics

To understand the general data distribution, we computed minimum and maximum standard deviation and mean values for all factors.

Cohesion

Minimum values are 10.5, maximum values 37.96, standard deviation 7.46, and mean 23.75. The histogram with the distribution plot shows a double-peak pattern, which implies that the values for cohesion are uniformly spread in the range, as displayed in Figure 6.

Slope Angle

Minimum, maximum, std dev, and mean for the slope angle are 31.0, 66.0, 6.82, and 44.58, respectively. The distribution plot for this factor is slightly right-skewed, meaning that higher slope angles are relatively less but do occur, as shown in Figure 9.

Unit Weight

The minimum, maximum, standard deviation, and mean for unit weight are 13.40, 18.75, 1.36, and 15.72, respectively. The distribution plot is nearly uniform, with values distributed fairly evenly throughout the range, as shown in Figure 8.

Internal Friction Angle

The minimum, maximum, standard deviation and mean of the internal friction angle are 12.0, 37.0, 6.92, and 25.46, respectively. The distribution for this factor is right-skewed, exhibiting a double-peak pattern, which means more variability in the higher values, as displayed in Figure 7.

Slope Height

The minimum, maximum, standard deviation, and mean for slope height are 28.0, 70.0, 10.0, and 45.87, respectively. A histogram indicates a right-skewed distribution of slope height, i.e., smaller slope heights are more common and larger slope heights less commonly encountered, as shown in Figure 10.

These distributions, as indicated in Table 2, outline the characteristics that underlie the dataset and enable the process of making sense of how each factor influences the stability and behaviour of the slope.

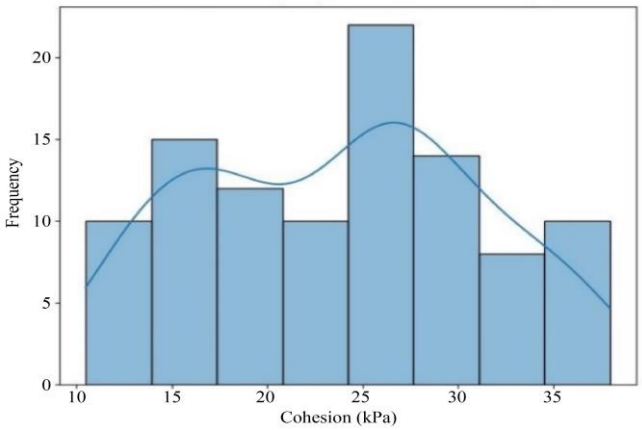

Fig. 6 Distribution histogram of cohesion (kPa) indexes

Table 2 Statistical characteristics of data

Index's	Cohesion (kPa)	Internal friction angle (◦)	Unit weight of soil (kN/m^{-3})	Slope angle (◦)	Slope height (m)
Median	24.50	27.00	15.65	43.00	45.00
Minimum	10.50	12.00	13.40	31.00	28.00
Maximum	37.96	37.00	18.75	66.00	70.00
Mean	23.75	25.46	15.72	44.58	45.87
Standard Deviation	7.46	6.92	1.36	6.82	10.00
Quantile 25%	17.50	20.00	14.62	41.00	39.00
Quantile 75%	28.87	31.00	16.54	48.00	51.00

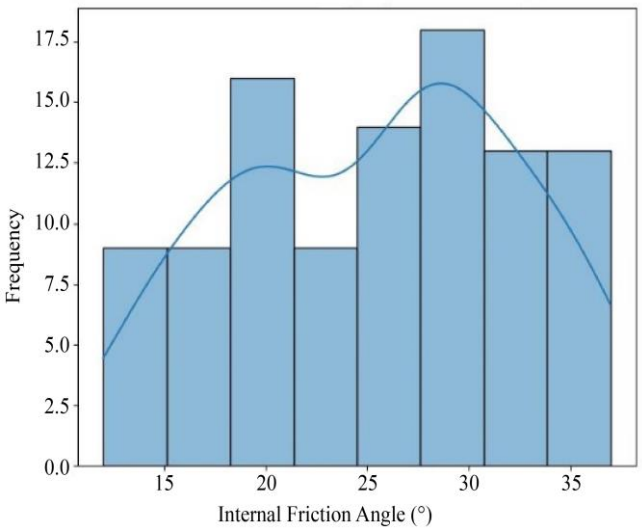

Fig. 7 Distribution histogram of internal friction angle (ϕ) indexes

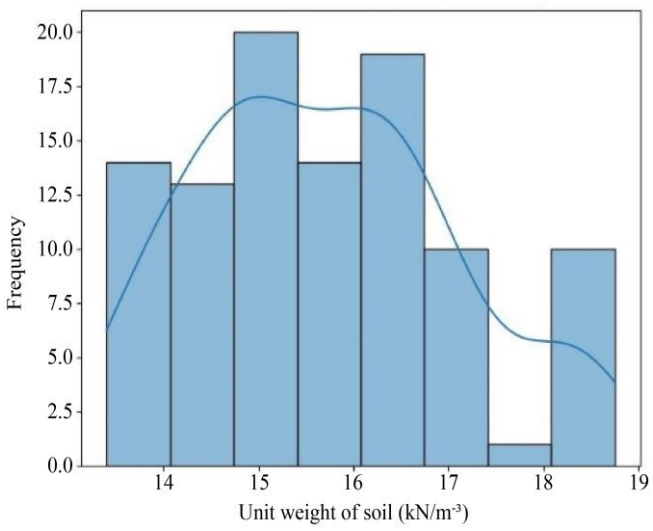

Fig. 8 Distribution histogram of the unit weight of soil (KN/m^3) indexes

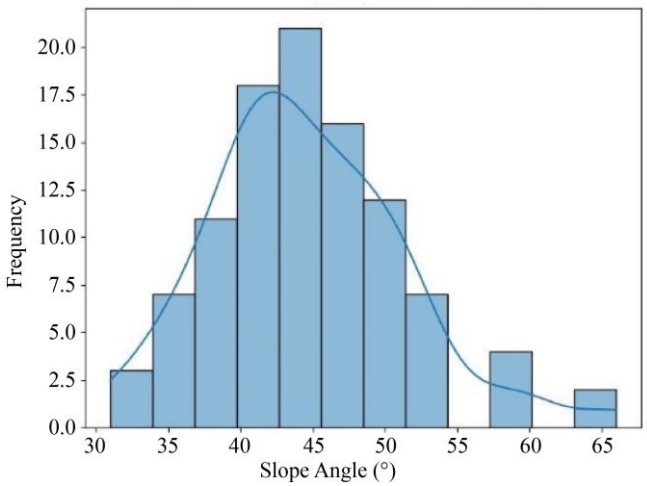

Fig. 9 Distribution histogram of slope angle (°) indexes

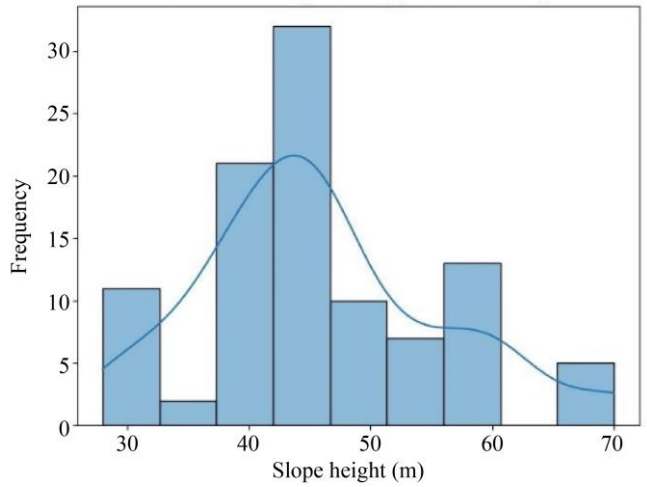

Fig. 10 Distribution histogram of Slope height (m) indexes

4.2.3. Correlation Analysis

Before finalizing the predicting models, the correlation between the five key characteristics needs to be investigated. Strong correlations among these factors can compromise the accuracy of the models and arrive at wrong conclusions against the actual data. This relationship is measured using Pearson's correlation coefficient, denoted by R [22]; correlation coefficients describe the strength and direction of a linear relationship between variables. The formula to calculate the correlation coefficient is as follows:

$$R(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}} \quad (2)$$

Where:

X and Y are independent variables.

R(X, Y) is the correlation coefficient, showing the linear relationship between X and Y.

Cov(X, Y) is the covariance, indicating how X and Y change together.

Var[X] and Var[Y] are the variances of X and Y, showing their spread.

Table 3 displays a correlation coefficient matrix for all five characteristics. A heat map (figure 11) depicts the correlation matrix: where warmer colours-in, in this case, red-will indicate a positive correlation and cooler colors, like blue, will denote a negative correlation. The correlation coefficient varies from -1 to 1, indicating a linear relationship.

Positive correlation: A coefficient between 0 and 1 denotes that the two factors go directly together, and increasing one factor increases another.

Negative correlation: A coefficient between -1 and 0 implies an inverse relationship in which an increase in one element reduces the other.

Negligible or almost no correlation: At an almost negligible linear relationship between the factors, the coefficient becomes 0.

The analysis shows that two pairs have a good positive correlation; these include the cohesion and internal friction angle (ϕ). On the other hand, the two pairs had good negative correlations; these include the height of the soil and cohesion, the height of the soil and the internal friction angle (ϕ). Unit weight indicates a very poor correlation with either the slope angle or the height of the soil; this means that these factors are relatively independent of each other.

Table 3. Heat map Correlation Matrix of variables for both MP Methods of Slope Stability

	Cohesion (c)	**Phie (internal friction Angle)**	**Unit weight**	**Slope Angle**	**Height of soil (Slope Height)**
Cohesion (c)	1	.99	-0.09	-0.05	-0.28
Phie (internal friction Angle)	0.99	1	-0.08	-0.05	-0.27
Unit weight	-0.09	-0.08	1	0.13	0.17
Slope Angle	-0.05	-0.05	0.13	1	0.62
Height of soil (Slope Height)	-0.28	-0.27	0.17	0.62	1

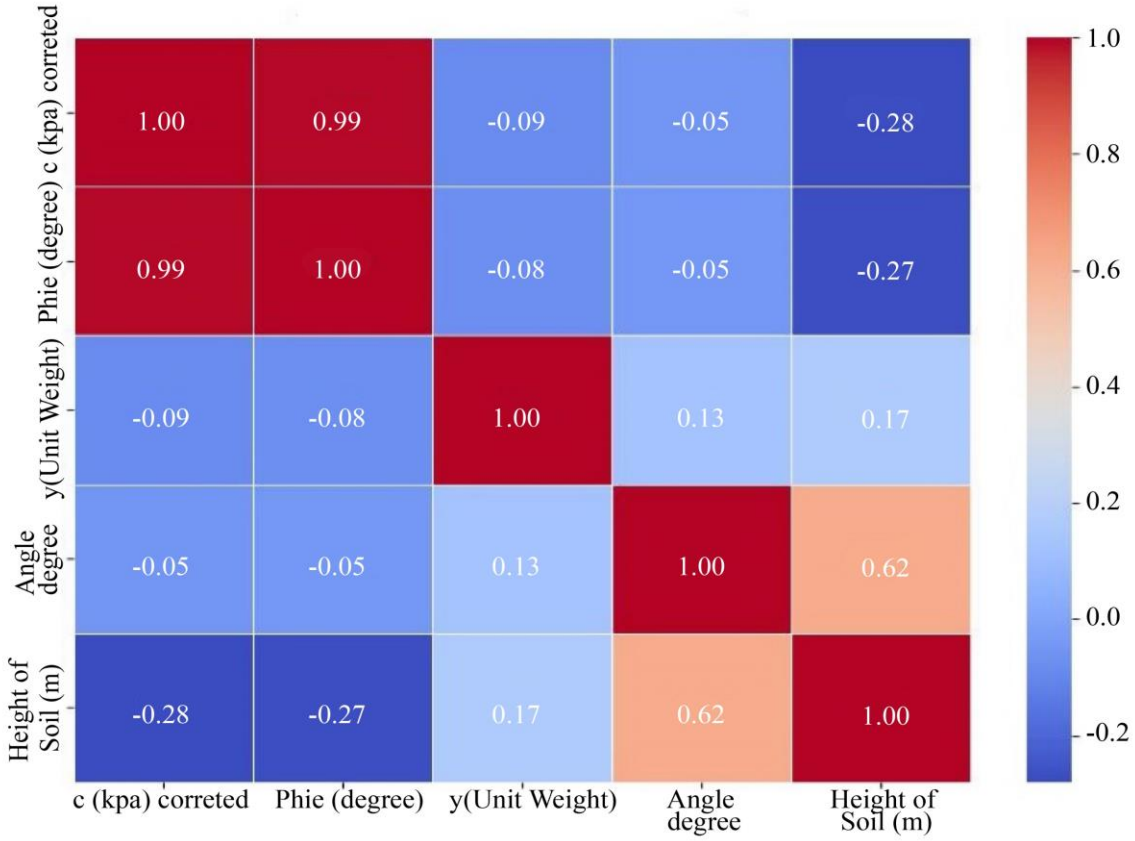

Fig. 11 Correlation matrix

4.2.4. Data of the Current Study

From the table shown in Table 1, a total of 101 samples were recorded. Each sample has five independent characteristics (factors) and two dependent indicators, one of which is the Factor of Safety (FOS). The safety factor was determined using the Morgenstern-Price (MP) method. The GeoStudio 2018 software is utilized in computing for the FOS [31]. The combined data show that the characteristics of each sample were significant, and every sample contained a valid indicator. These datasets were actually used for training and testing models in Python. The data used as input values in soft computing models and the corresponding prediction outputs were obtained through training and testing of these models. The actual FOS values and values predicted by DT, RF, KNN, NN [32], LR, and SVR models were used for comparison with the output of the developed model. Different performance parameters and statistical tests have been applied to determine the most reliable one.

4.2.5. Performance Parameters

Various statistical methods have been used in order to evaluate the fitness and adequacy of models, among which are:

Root Mean Squared Error (RMSE)[21], Willmott's Index for Agreement (WI)[12], Reliability Index (β)[11], Median Absolute Error, t-Statistic, Performance Index (PI), Bias Factor, RS Ratio (RSR), Maximum Error, Normalized Mean Bias Error (NMBE), Coefficient of Determination (R^2), Expanded Uncertainty (U95)[12], Relative Percentage Difference (RPD), Mean Bias Error (MBE), Global Performance Indicator (GPI), Nash-Sutcliffe Efficiency (NSE) [11], Mean Absolute Error (MAE), Explained Variance, Mean Squared Error (MSE)[12], Variance Account Factor (VAF), Legate and McCabe's Index (LMI), Adjusted R^2[11], Mean Absolute Percentage Error (MAPE)[21] are Key Performance Metrics. These parameters help predict the error of the different models used in this paper.

Soft computing models and approaches are versatile computer strategies for solving complex, uncertain, or imprecise situations. Fuzzy logic manages uncertainty, neural networks to recognize patterns, and evolutionary algorithms to optimize. These methods prioritize approximation and adaptability, making them appropriate for real-world issue resolution.

Mean Squared Error (MSE)

This measure assesses soft computing models' prediction error. A value nearer to zero indicates more accuracy [12].

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2 \quad (3)$$

Root Mean Squared Error (RMSE)

Also indicates prediction error. A lower RMSE means higher prediction accuracy [33].

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2} \quad (4)$$

Mean Bias Error (MBE) and Mean Absolute Error (MAE)

Metrics for determining the accuracy of expected foundation settling. The optimum value for both of these errors is 0 [34], which represents perfect prediction accuracy.

$$\text{MBE} = \frac{1}{n}\sum_{i=1}^{n}(o_i - p_i) \quad (5)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|(o_i - p_i)| \quad (6)$$

Explained Variance

Indicates the error in the predictions by the soft computing models. The ideal value is 1.

$$\text{Explained Variance} = exlained\ variance\ score\ (o_i, p_i) \quad (7)$$

Median Absolute Error

Reflects the error in the predictions. A value of 0 is ideal.

$$\text{Median Absolut Error} = median\ abs\ (o_i, p_i) \quad (8)$$

Maximum Error

Shows the largest error in predictions. The ideal value is 0.

$$\text{Max. Error} = max.(o_i, p_i) \quad (9)$$

Nash-Sutcliffe Efficiency (NSE)

Assesses the forecasting capability of the models. Values nearer to 1 suggest superior performance [35].

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(o_i - p_i)^2}{\sum_{i=1}^{n}(o_i - o_{mean})^2} \quad (10)$$

Legate and McCabe's Index (LMI)

It indicates a variation in prediction accuracy, with values spanning from -∞ to 1 [36]."

$$\text{LMI} = 1 - \frac{\sum_{i=1}^{n}|(o_i - p_i)|}{\sum_{i=1}^{n}|(o_i - o_{mean})|} \quad (11)$$

Expanded uncertainty (U95) represents the model's performance in predicting short-term foundation settlement. A smaller U95 value indicates better model performance [11].

$$\text{U95} = 1.96 * \text{std.}(o_i - p_i) \quad (12)$$

Variance Account Factor (VAF)

Assesses model performance, where values approaching 100% signify superior results [37].

$$\text{VAF} = 1 - \frac{var(p_i - o_i)}{var(o_i - o_{mean})} \quad (13)$$

R^2 and Adj. R^2

These numbers demonstrate the model's capacity to capture variation in soil properties. Values nearer to 1 and closer together indicate higher model accuracy [38].

$$R^2 = \frac{\sum_{i=1}^{n}(o_i - o_{mean})^2 - \sum_{i=1}^{n}(o_i - p_i)^2}{\sum_{i=1}^{n}(o_i - o_{mean})^2} \quad (14)$$

$$\text{Adj } R^2 = 1 - \frac{(1-R^2)*(n-1)}{(n-f-1)} \quad (15)$$

Each performance parameter provides insight into the accuracy and reliability of the models for predicting foundation settlements based on the input characteristics.

t-statistic A lower value indicates better predictive performance of the model [39].

$$\text{tStatistic} = \frac{mean\frac{residuals}{std.(residuals)}}{sqrt(\ln(o_i))} \quad (16)$$

Performance Index (PI)

The value reflects the effectiveness of the model's performance [11].

$$\text{PI} = \frac{\sum_{i=1}^{n}(abs(p_i - o_i)}{\sum_{i=1}^{n}(abs(o_i))} \quad (17)$$

Bias Factor

The value reflects estimation deviation: greater than 1 indicates overestimation, less than 1 underestimation, and exactly 1 signifies an unbiased model [40].

$$\text{Bias Factor} = \frac{1}{n}\sum_{i=1}^{n}\frac{p_i}{o_i} \quad (18)$$

Root Mean Standard (RSR) [50]

Indicates the error index, where values closer to 0 signify higher predictive accuracy.

$$\text{RSR} = \frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n}(o_i - o_{mean})^2} \quad (19)$$

Normalized Mean Bias Error (NMBE)

The model's prediction accuracy for values that differ from the mean. The model performs best when NMBE is equal to zero [11].

$$\text{NMBE} = \frac{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)}{\frac{1}{n}\sum_{i=1}^{n}(o_i)} \quad (20)$$

Relative Percentage Difference (RPD)

Represents the model's performance as computed by Eq. 23, with a greater RPD value indicating stronger predictive capabilities. [11].

$$\text{RPD} = \sqrt{\frac{\frac{1}{n}\sum_{i=1}^{n}(o_i - p_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(o_i - o_{mean})^2}} \quad (21)$$

Willmott's Index (WI)

Willmott's Index (WI) is the prediction error level for shallow foundation settlement using soft computing models. A number of 1 implies optimal prediction accuracy with minimal error on an index scale of 0 to 1 [41].

$$\mathrm{WI} = 1 - \frac{\sum_{i=1}^{n}(o_i - p_i)^2}{\sum_{i=1}^{n}|(o_i - o_{mean})^2 + (p_i - o_{mean})|^2} \quad (22)$$

Global Performance Indicator (GPI)

Global Performance Indicator (GPI) evaluates the model with many assessment parameters in one value [23]. A higher GPI suggests better model accuracy [12].

$$\mathrm{GPI} = \frac{\sum_{i=1}^{n}(abs(p_i - o_i))}{\sum_{i=1}^{n}(abs(p_i - o_{mean}))} \quad (23)$$

Reliability Index (β)

A higher dependability index value corresponds to greater model performance [42].

$$\beta = \frac{mean\ (residual)}{std.(residual)} \quad (24)$$

5. Methodology

Here, o_i is the i_{th} observed value, p_i is the i_{th} prediction value, o_{mean} is the average of observed value, SD represents the standard deviation, and f represents the feature.

5.1. Data Collection and Description

A total of 101 soil samples were gathered from various sites along the stretch between Chhara Village and Karki Ghat in Almora District. Tests were performed to analyze unit weight, grain size distribution, and shear strength.

The data obtained revealed information about soil types φ, c, unit weight, and the height and angle of the soil mass in the field. Five of the most important parameters were used in this study to check the slope stability.

Pore-water pressure is not definite in field measurements and differs from one standard to another; hence, it is impossible to add to prediction models. Thus, the models were deleted for accuracy and reliability. The FOS was determined by Geo-Studio software (2018). In this study, the MP approach is utilized for the slope stability assessment method.

5.2. Model Training and Development

There are only 101 samples because real-world field data are usually hard to collect. To overcome this issue, appropriate machine learning algorithms for small datasets were applied, namely LR, RF, DT, SVR, KNN, and NN. These models improve the reliability of the predictions.

5.3. Testing and Evaluation

The collected samples were used for slope stability prediction. The data size was very small, and k-fold cross-validation was utilized to validate the performance of the prediction models [43].

Six machine learning methods, including DT, RF, KNN, LR, SVR, and NN, have been applied to develop the prediction models. Traditional cross-validation techniques like 21-fold cross-validation suffer from inherent randomness. Rather, a random cross-validation method was used [43].

In this regard, samples 21 were used as a test set through random selection, and the rest were used for training. This type of training was carried out five times on the samples, and the average of all the results mentioned above was taken as the final prediction.

This case study examines the performance of the models using a number of performance metrics such as Performance Index (PI), Uncertainty (U95), R^2 (Coefficient of Determination), NSE, VAF, RS-Ratio (RSR), GPI, t-Statistic, Bias Factor, NMBE, Mean Bias Error (MBE), Reliability Index (Beta), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error, Explained Variance, MAE, Maximum Error, Adjusted R^2, WI, MAPE, Relative Percentage Difference (RPD), LMI. Each metric was utilized in the evaluation of the performance of the models [44].

5.4. Selection of the Model

A procedure that implemented k-fold cross-validation for six different machine learning models resulted in the tabular comparison of the performance metrics of each model, as indicated in Table 4. A careful consideration of these was made, and the most performing model was determined based on parameters such as MSE, RMSE, U95, VAF, and MAPE.

5.5. Prediction

The model was then used to predict new values, and the model was deployed for future predictions.

5.6. Residual Plot Analysis

Tested the residual plots of each model using the MP method of slope stability in this section. Residual plots display differences between observed and predicted values, highlighting model performance and inconsistencies, as shown in Figures (12-17). One gets a scatterplot where the closer the points align with the vertical line, the better the model fits with the data.

- LR fitted the data perfectly
- k-NN demonstrated good fitting after LR
- RF and SVR were approximately average in fitting the data
- DT and NN performed poorly in fitting the data

The residual plot analysis helps to affirm which of the models is appropriate for the prediction of slope stability.

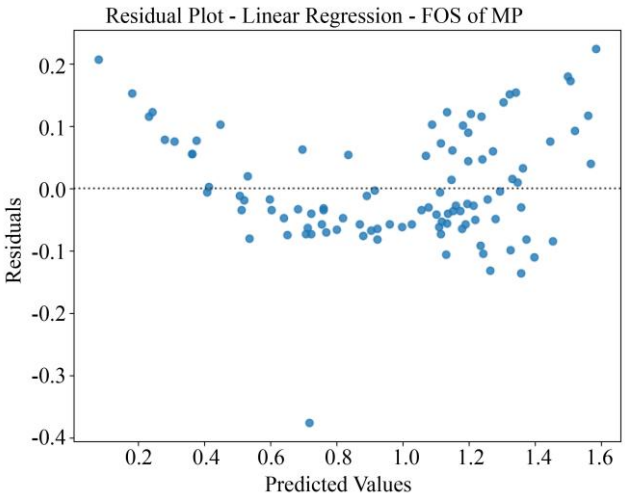

Fig. 12 Residual plot of LR as applied through the MP method

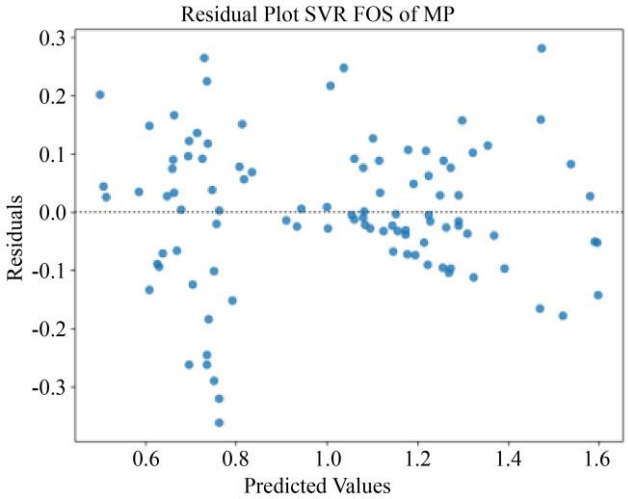

Fig. 15 Residual plot of SVR as applied through the MP method

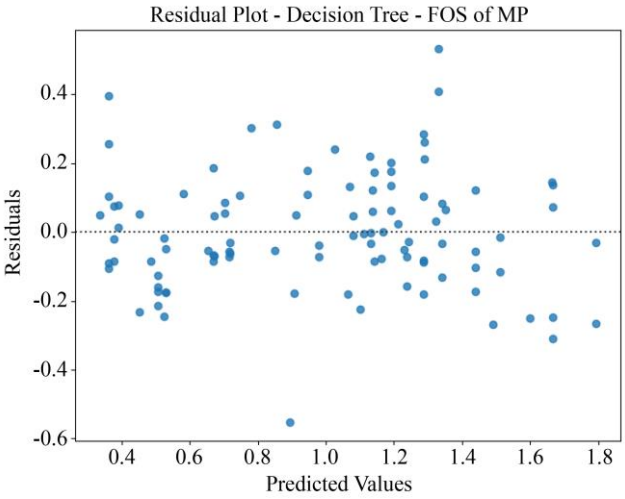

Fig. 13 Residual plot of DT as applied through the MP method

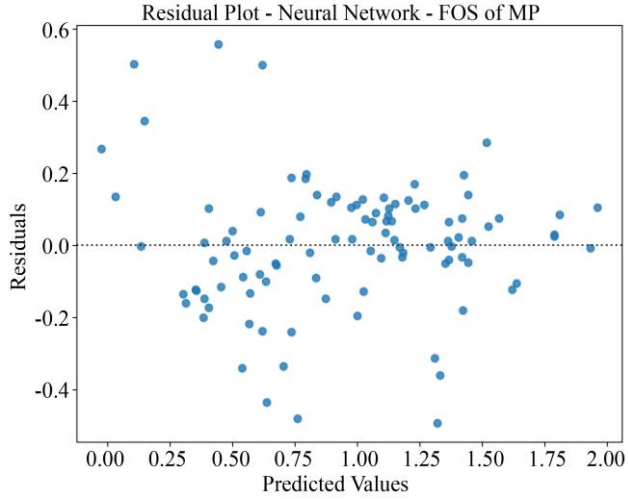

Fig. 16 Residual plot of NN as applied through the MP method

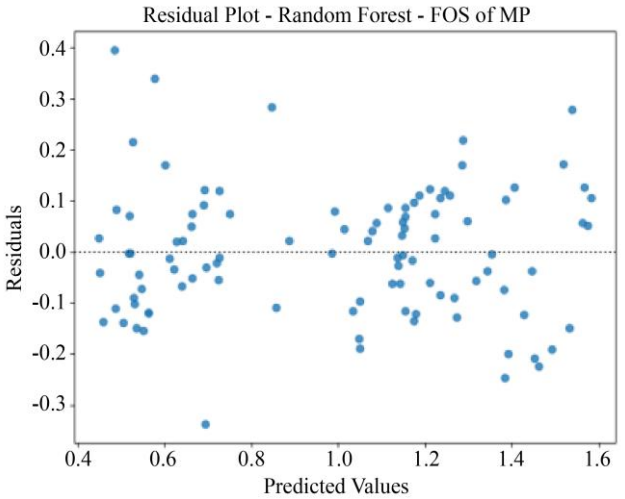

Fig. 14 Residual plot of RF as applied through the MP method

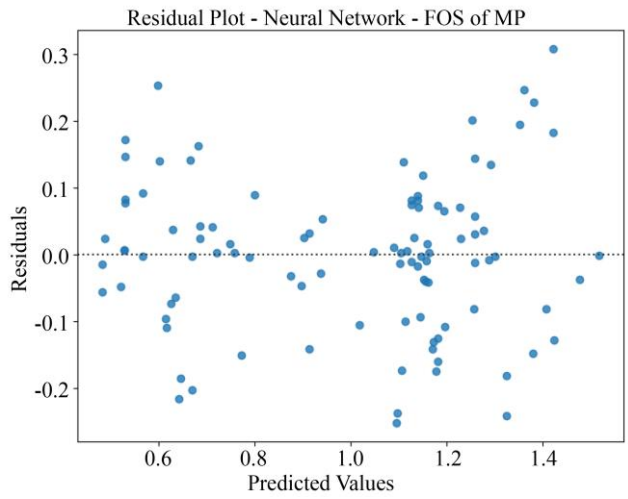

Fig. 17 Residual plot of k-NN as applied through the MP method

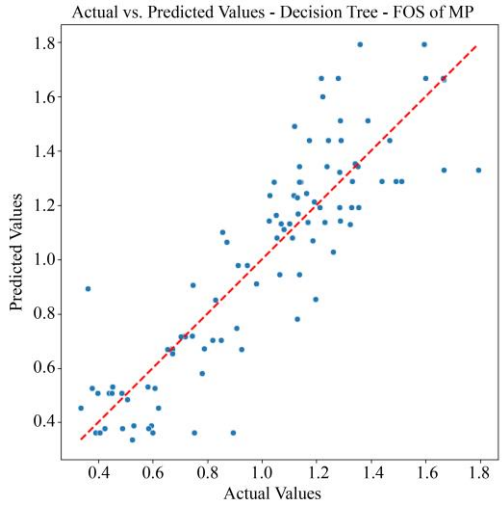

Fig. 18 Performance of Actual vs. Predicated Value of DT model (by MP methods of Slope stability)

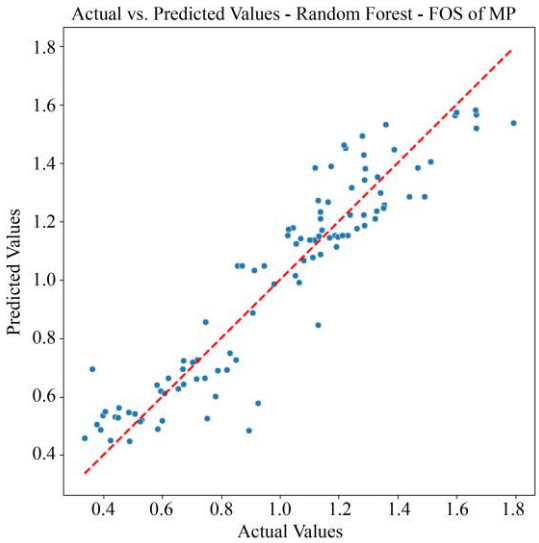

Fig. 19 Performance of Actual vs. Predicated Value of RF model (by MP methods of Slope stability)

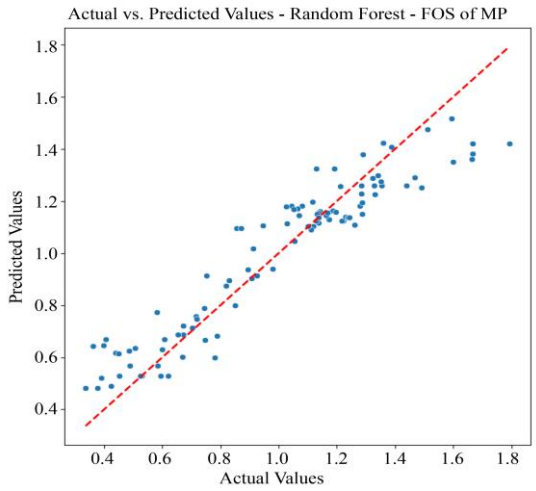

Fig. 20 Performance of Actual vs. Predicated Value of k-NN model (by MP methods of Slope stability)

6. Results and Discussion

FOS was calibrated using the MP method. Six supervised learning models were used on samples of soil. These models include NN, DT, K-NN, SVR RF and LR. Their training, testing, and validation were done using k-fold cross-validation with the FOS as the target variable. Figure (18-23) plots the actual vs. predicted values for all models. We can draw conclusions about the performance of each model from these plots. Between these, we could see that LR provided the best results [45] in that its predicted values have a closer approximation to the actual values. This way, it follows that its model fits better. Below are the ROC curves for all models plotted, as shown in Figure 24. It generates the ROC curve, which depicts the connection between the True Positive and False Positive Rates. True positives correctly predict actual positive values, while false positives incorrectly predict negative values. The ROC curve of a random classifier lies on the diagonal of the graph, and an ideal curve would be placed in the top left corner. From the graphical outcomes of ROC curves, it is clear that the LR model is the best since the curve is closer to the ideal curve. Thus, the LR model is the best based on the ROC analysis.

Tables 4 describe some performance metrics as Willmott's Index (WI), MSE, RMSE, MAE, explained variance, Mean Bias Error (MBE), Nash-Sutcliffe Efficiency (NS), Median Absolute Error, VAF, MAPE, NMBE, RSR, Bias factor, Performance Index (PI), R^2, Adjusted R^2, GPI, LMI, U95, t-statistic, Reliability Index (β), RPD, and maximum error. Among these, the NS of the LR model is closest to 1, which means it has the greatest predictive strength. Besides, comparing the models based on MSE, MAE, RMSE, and VAF, the prediction error appears to be the smallest in LR, which once again reflects its superiority. The R^2 and Adjusted R^2 of the LR models are also closer to 1, and this suggests that the LR model effectively explains the majority of the variability in the soil parameters. Further, the higher the Reliability Index β for the model LR, the better it will be.

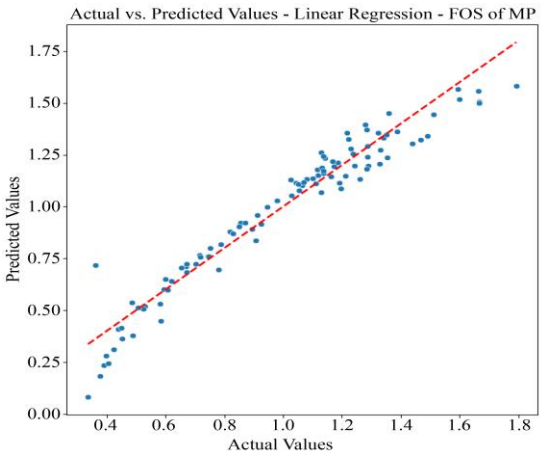

Fig. 21 Performance of Actual vs. Predicated Value of LR model (by MP methods of Slope stability)

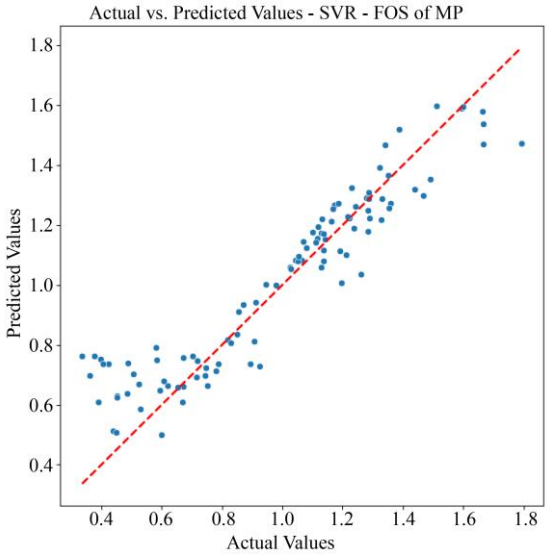

Fig. 22 Performance of Actual vs. Predicated value of SVR model (by MP methods of Slope stability)

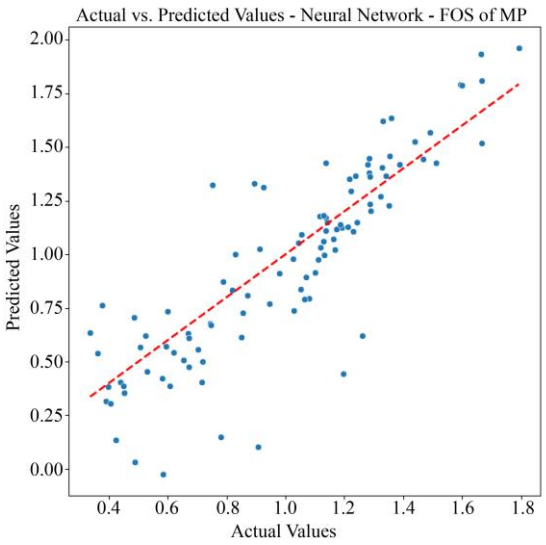

Fig. 23 Performance of Actual vs. Predicated value of NN model (by MP methods of Slope stability)

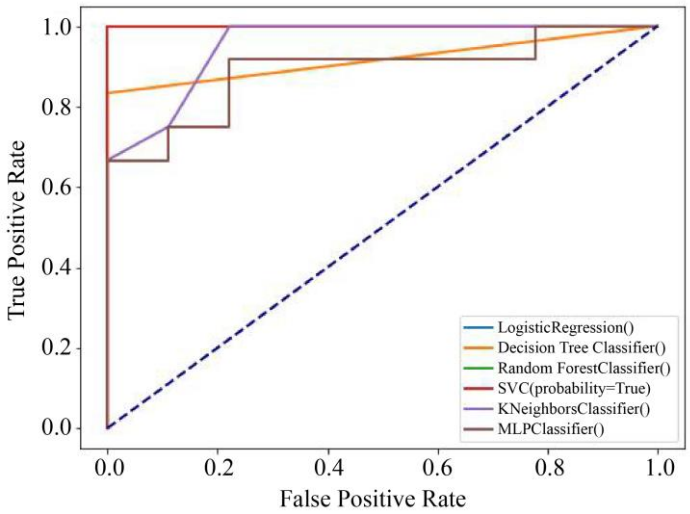

Fig. 24 ROC Curve for FOS of MP

Table 4. Performance parameters of DT, RF, KNN, NN, LR, and SVR model

Model	LR	DT	RF	SVR	k-NN	NN	Ideal Value
MSE	0.00831593	0.03572103	0.016105757	0.017444395	0.01543877	0.05289839	0
RMSE	0.09119174	0.189000079	0.126908461	0.13207723	0.12425284	0.22999651	0
MAE	0.07038892	0.141960396	0.099307624	0.095674387	0.09557822	0.16590351	0
R2	0.93542954	0.722638076	0.874944146	0.864550073	0.8801231	0.58926155	1
Explained Variance	0.93629835	0.722674917	0.875166078	0.870258868	0.88036815	0.60144424	1
Median Absolute Error	0.05173757	0.116	0.08146	0.069654322	0.0784	0.11296109	0
Max Error	0.35496829	0.531	0.40795	0.425592868	0.3724	0.8036963	0
Nash-Sutcliffe Efficiency	0.93542954	0.722638076	0.874944146	0.864550073	0.8801231	0.58926155	1

Legate-McCabe Index	0.99840408	1.135704325	1.01606474	0.854986718	0.85169932	1.23007196	- ∞, 1
Expanded Uncertainty	0.17752927	0.370415551	0.24851977	0.253357332	0.24328652	0.44405747	0
Variance Account Factor	0.93629835	0.722674917	0.875166078	0.870258868	0.88036815	0.60144424	100%
Adj. R2	0.9320311	0.70804008	0.868362259	0.857421129	0.87381379	0.56764374	1
t-Statistic	1.17367452	0.11583208	-0.4237445	-2.10811386	-0.4548484	1.75706448	Smaller value
Performance Index	0.07116397	0.143523524	0.100401101	0.096727859	0.09663063	0.16773028	0
Bias Factor	0.98930559	0.997797798	1.005405105	1.027413638	1.00567968	0.95995337	1
RS Ratio	0.25410718	0.526651615	0.353632371	0.368035226	0.34623244	0.6408888	0
Normalized Mean Bias Error	-0.01069441	-0.0022022	0.005405105	0.027413638	0.00567968	-0.0400466	0
MAPE	9.29622722	16.09777516	11.80823332	14.52444658	11.9440363	20.5802611	0
Relative Percentage Difference	0.25410718	0.526651615	0.353632371	0.368035226	0.34623244	0.6408888	>20%
Willmott's Index	0.83028921	0.840928948	0.855862116	0.874865841	0.88888524	0.8415979	0.0-1.0
Mean Bias Error	-0.01057793	-0.00217822	0.005346238	0.027115073	0.00561782	-0.0396105	0
Global Performance Indicator	0.22961081	0.407095493	0.318313884	0.364443906	0.36548285	0.43925776	Higher Value
Reliability Index	0.11678498	0.011525723	-0.04216415	-0.20976517	-0.0452591	0.17483445	Higher Value

Table 5. Model and parameter

Model	Parameter
DT	Using the Decision Tree Regressor from scikit-learn with default hyperparameters, splitting criterion as MSE, and unlimited maximum depth for the tree.
RF	The RandomForest Regressor is an ensemble of decision trees that uses default settings. These include the number of trees (n_estimators=100) and the MSE criterion for splitting.
KNN	The KNeighbours Regressor is employed with default settings, including the number of neighbours, which is set to 5 by default.
LR	The Linear Regression model uses default settings, with no explicit hyperparameters set in the code.
SVR	The SVR model is used with its default settings, including the Radial Basis Function (RBF) as the default kernel for regression in scikit-learn.
NN	The MLP (Multi-Layer Perceptron) Regressor is used with default settings, including one hidden layer with 100 nodes, the ReLU activation function, and optimization through stochastic gradient descent.

7. Conclusion

In this research, six machine learning techniques- LR, DT, RF, SVR, k-NN, and [46] NN- were applied to analyze the reliability of slope stability. The data set gathered is used to assess soil stability based on the MP method. The data was divided into a training set and a test set; models were trained with k-fold cross-validation. A detailed comparison of the six models, including both performance metrics and graphical analysis, showed that the most effective model is LR. Selection was also made based on parameters such as RMSE, Bias Factor, and Performance Index. Further visual evaluation from ROC curves, precision-recall curves, and plots of actual

vs. predicted values proved the superiority of LR in [47] predicting power, as shown in Figure (18-23) and Figure 24. LR had high predictive capability and the highest accuracy for the stability of soil under prediction.

This research developed prediction models using five key indicators chosen for influencing the quantitative aspects of slope stability. As great as this is for the first endeavor for assessing slope stability, these qualitative factors - pore pressure, rainfall, and existing joints, for example - play a significant role in slope stability [48]. Methods for converting such qualitative factors into measurable indicators are needed in further research. These other methods of slope stability analysis, like Spencer, Bishop and Janbu, must be taken into consideration in further work for more accurate predictions.

The main purpose and challenge for the future in this area of research will be to find even more objective and integral indicators of the assessment of slope stability. This research can be applied to other places with similar geographical, meteorological, or socioeconomic variables, providing significant insights into broader applications. These methods could be used to create real-time predictive models in sectors including hydrology, forest fire risk management, and crop production prediction across various terrains. The study may also look into hybrid techniques, which combine the capabilities of various machine learning models to increase performance across a variety of datasets and applications.

Data availability

All the data generated or analyzed during this study are included in this article**.**

Acknowledgements

The authors would like to express their gratitude to Lalit Dumka, Ravindar Singh Koranga, Rohit Kumar, Dr. Ankur Bisht, Mani, and Bhavna for their kind assistance with the analysis of the results of this research. Additionally, I would like to thank the Head of the Civil Engineering Department at Graphic Era Hill University in Bhimtal for giving me the lab space needed for the research for this paper. I especially want to thank Nitin Harbola for his assistance with the lab work and soil sample collection.

References

[1] John T. Christian, Charles C. Ladd, and Gregory B. Baecher, "Reliability Applied to Slope Stability Analysis," *Journal of Geotechnical Engineering*, vol. 120, no. 12, pp. 2180-2207, 1994. [CrossRef] [Google Scholar] [Publisher Link]

[2] Y.M. Cheng, "Location of Critical Failure Surface and Some Further Studies on Slope Stability Analysis," *Computers and Geotechnics*, vol. 30, no. 3, pp. 255-267, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[3] M.G. Sakellariou, and M.D. Ferentinou, "A Study of Slope Stability Prediction using Neural Networks," *Geotechnical & Geological Engineering*, vol. 23, pp. 419-445, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[4] Sarat Kumar Das et al., "Classification of Slopes and Prediction of Factor of Safety Using Differential Evolution Neural Networks," *Environmental Earth Sciences*, vol. 64, pp. 201-210, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[5] Thanh Son Nguyen, and Suched Likitlersuang, "Reliability Analysis of Unsaturated Soil Slope Stability under Infiltration Considering Hydraulic and Shear Strength Parameters," *Bulletin of Engineering Geology and the Environment*, vol. 78, pp. 5727-5743, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] I. Yilmaz, and A.G. Yuksek, "An Example of Artificial Neural Network (ANN) Application for Indirect Estimation of Rock Parameters," *Rock Mechanics and Rock Engineering*, vol. 41, no. 5, pp. 781-795, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[7] H. El-Ramly, N.R. Morgenstern, and D.M. Cruden, "Probabilistic Slope Stability Analysis for Practice," *Canadian Geotechnical Journal*, vol. 39, no. 3, pp. 665-683, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[8] Yaser A. Nanehkaran et al., "Comparative Analysis for Slope Stability by Using Machine Learning Methods," *Applied Science*, vol. 13, no. 3, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Prashanth Ragam et al., "Estimation of Slope Stability Using Ensemble-Based Hybrid Machine Learning Approaches," *Frontiers in Materials*, vol. 11, pp. 1-19, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] G.L. Sivakumar Babu, and Amit Srivastava, "Reliability Analysis of Earth Dams," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 136, no. 7, pp. 995-998, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[11] Rahul Ray, and Lal Bahadur Roy, "Reliability Analysis of Soil Slope Stability Using ANN, ANFIS, PSO-ANN Soft Computing Techniques," *NVEO-Natural Volatiles & Essential Oils Journal*, vol. 8, no. 6, pp. 3478-3491, 2021. [Google Scholar] [Publisher Link]

[12] Rahul Ray, Shiva Shankar Choudhary, and Lal Bahadur Roy, "Reliability Analysis of Soil Slope Stability using MARS, GPR and FN Soft Computing Techniques," *Modeling Earth Systems and Environment*, vol. 8, pp. 2347-2357, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Arunava Chakraborty, and Digantab Goswami, "Prediction of Critical Safety Factor of Slopes Using Multiple Regression and Neural Network," *Journal of Geo-Engineering Sciences*, pp. 1-10, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[14] Arunav Chakraborty, and Diganta Goswami, "Prediction of Slope Stability Using Multiple Linear Regression (MLR) and Artificial Neural Network (ANN)," *Arabian Journal of Geosciences*, vol. 10, pp. 1-11, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[15] Mohammad Khajehzadeh et al., "An Effective Artificial Intelligence Approach for Slope Stability Evaluation," *IEEE Access*, vol. 10, pp. 5660-5671, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Xianda Feng et al., "Prediction of Slope Stability Using Naive Bayes Classifier," *KSCE Journal of Civil Engineering*, vol. 22, pp. 941-950, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[17] Shan Lin et al., "Evaluation and Prediction of Slope Stability Using Machine Learning Approaches," *Frontiers of Structural and Civil Engineering*, vol. 15, no. 4, pp. 821-833, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Singh Sudhir Kumar, and Chakravarty Debashish, "Assessment of Slope Stability Using Classification and Regression Algorithms Subjected to Internal and External Factors," *Archives of Mining Sciences*, vol. 68, no. 1, pp. 87-102, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Shuai Huang, Mingming Huang, and Yuejun Lyu, "An Improved KNN-Based Slope Stability Prediction Model," *Advances in Civil Engineering*, vol. 2020, no. 1, pp. 1-16, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Pijush Samui, and Bimlesh Kumar, "Artificial Neural Network Prediction of Stability Numbers for Two-Layered Slopes with Associated Flow Rule," *The Electronic Journal of Geotechnical Engineering*, vol. 11, pp. 1-44, 2006. [Google Scholar] [Publisher Link]

[21] Zaobao Liu et al., "An Extreme Learning Machine Approach for Slope Stability Evaluation and Prediction," *Natural Hazards*, vol. 73, pp. 787-804, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[22] Shuhong Wang, and Muhammad Israr Khan, "Developing Correlations for Advance Prediction of Slope Factor of Safety Using Linear Regression Analysis–Karachi Landslide as a Case Study," *Polish Journal of Environmental Studies*, vol. 30, no. 6, pp. 5849-5862, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] An-Jui Li, Kelvin Lim, and Abdoulie Fatty, "Stability Evaluations of Three-Layered Soil Slopes Based on Extreme Learning Neural Network," *Journal of the Chinese Institute of Engineers*, vol. 43, no. 7, pp. 628-637, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Abidhan Bardhan, and Pijush Samui, "Application of Artificial Intelligence Techniques in Slope Stability Analysis: A Short Review and Future Prospects," *International Journal of Geotechnical Earthquake Engineering*, vol. 13, no. 1, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Rameshwar Bali, A.R. Bhattacharya, and T.N. Singh, "Active Tectonics in the Outer Himalaya: Dating a Landslide Event in the Kumaun Sector," *Earth Science India*, vol. 2, no. 4, pp. 276-288, 2009. [Google Scholar] [Publisher Link]

[26] Indian Standards, "*Criteria for Earthquake Resistant Design of Structures*," IS 1893 (Part 4), Bureau of Indian Standards, pp. 1-23, 2002. [Google Scholar] [Publisher Link]

[27] R.K. Umrao et al., "Stability Analysis of Cut Slopes Using Continuous Slope Mass Rating and Kinematic Analysis in Rudraprayag District, Uttarakhand," *Geomaterials*, vol. 1, no. 3, pp. 1-9, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[28] Sina Javankhoshdel, and Richard J. Bathurst, "Simplified Probabilistic Slope Stability Design Charts for Cohesive and Cohesive-Frictional (C-Φ) Soils," *Canadian Geotechnical Journal*, vol. 51, no. 9, pp. 1033-1045, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[29] L.K. Sharma et al., "Stability Investigation of Hill Cut Soil Slopes along National Highway 222 at Malshej Ghat, Maharashtra," *Journal of the Geological Society of India*, vol. 89, pp. 165-174, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[30] Muhammad Israr Khan, and Shuhong Wang, "Comparing the Various Slope Stability Methods to Find the Optimum Method for Calculating Factor of Slope Safety," *IOP Conference Series: Earth and Environmental Science, International Conference on Energy, Material Science and Environment Engineering*, Taipei, China, vol. 480, pp. 1-13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[31] M. Pandit Vinod, B. Patel Jignesh, K. Thakare Amol, "Comparative Study of Advanced Two Dimensional Methods of Slope Stability Analysis," *International Journal of Advanced Technology in Civil Engineering*, vol. 2, no. 1, pp. 39-42, 2013. [CrossRef] [Publisher Link]

[32] Himanshu Sharma, and Anand Singh Jalal, "Visual Question Answering Model Based on Graph Neural Network and Contextual Attention," *Image and Vision Computing*, vol. 110, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[33] Ozgur Kisi, Jalal Shiri, and Mustafa Tombul, "Modeling Rainfall-Runoff Process Using Soft Computing Techniques," *Computers & Geosciences*, vol. 51, pp. 108-117, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[34] T. Raventos-Duran et al., "Structure-Activity Relationships to Estimate the Effective Henry's Law Constants of Organics of Atmospheric Interest," *Atmospheric Chemistry and Physics*, vol. 10, pp. 7643-7654, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[35] Sharad K. Jain, and K.P. Sudheer, "Fitting of Hydrologic Models: A Close Look at the Nash–Sutcliffe Index," *Journal of Hydrologic Engineering*, vol. 13, no. 10, pp. 981-986, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[36] David R. Legates, and Gregory J. McCabe, "A Refined Index of Model Performance: A Rejoinder," *International Journal of Climatology*, vol. 33, no. 4, pp. 1053-1056, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[37] Candan Gokceoglu, "A Fuzzy Triangular Chart to Predict the Uniaxial Compressive Strength of the Ankara Agglomerates from their Petrographic Composition," *Engineering Geology*, vol. 66, no. 1-2, pp. 39-51, 2002. [CrossRef] [Google Scholar] [Publisher Link]

[38] G.L. Sivakumar Babu, and Amit Srivastava, "Reliability Analysis of Allowable Pressure on Shallow Foundation Using Response Surface Method," *Computers and Geotechnics*, vol. 34, no. 3, pp. 187-194, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[39] R.J. Stone, "Improved Statistical Procedure for the Evaluation of Solar Radiation Estimation Models," *Solar Energy*, vol. 51, no. 4, pp. 289-291, 1993. [CrossRef] [Google Scholar] [Publisher Link]

[40] Sathit Prasomphan, and Shigeru Mase, "Generating Prediction Map for Geostatistical Data Based on an Adaptive Neural Network Using Only Nearest Neighbors," *International Journal of Machine Learning and Computing*, vol. 3, no. 1, pp. 98-102, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[41] Ravinesh C. Deo, Pijush Samui, and Dookie Kim, "Estimation of Monthly Evaporative Loss Using Relevance Vector Machine, Extreme Learning Machine and Multivariate Adaptive Regression Spline Models," *Stochastic Environmental Research and Risk Assessment*, vol. 30, pp. 1769-1784, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[42] USACE, "*Risk-based Analysis in Geotechnical Engineering for Support of Planning Studies*," Technical Letter, Engineering and Design, US Army Corps of Engineers, Washington DC, 1999. [Publisher Link]

[43] Yu Wang et al., "Probabilistic Analysis of Post-Failure Behavior of Soil Slopes Using Random Smoothed Particle Hydrodynamics," *Engineering Geology*, vol. 261, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[44] Tan N. Nguyen et al., "A Novel Analysis-Prediction Approach for Geometrically Nonlinear Problems Using Group Method of Data Handling," *Computer Methods in Applied Mechanics and Engineering*, vol. 354, pp. 506-526, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[45] Behailu G. Habtemariam, Kelifa B. Shirago, and Democracy D. Dirate, "Effects of Soil Properties and Slope Angle on Deformation and Stability of Cut Slopes," *Advances in Civil Engineering*, vol. 2022, no. 1, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[46] D. Leynaud, and N. Sultan, "3-D Slope Stability Analysis: A Probability Approach Applied to the Nice Slope (SE France)," *Marine Geology*, vol. 269, no. 3-4, pp. 89-106, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[47] Muhammad Israr Khan, and Shuhong Wang, "Slope Stability Analysis to Develop Correlations between Different Soil Parameters and Factor of Safety Using Regression Analysis," *Polish Journal of Environmental Studies*, vol. 30, no. 5, pp. 4021-4030, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[48] Rajesh Singh, R.K. Umrao, and T.N. Singh, "Probabilistic Analysis of Slope in Amiyan Landslide Area, Uttarakhand," *Geomatics, Natural Hazards Risk*, vol. 4, no. 1, pp. 13-29, 2013. [CrossRef] [Google Scholar] [Publisher Link]