Transportation Infrastructure Geotechnology
https://doi.org/10.1007/s40515-024-00446-y

TECHNICAL PAPER

Application of Soft Computing Techniques for Slope Stability Analysis

Rashid Mustafa[1] · Akash Kumar[1] · Sonu Kumar[1] · Navin Kumar Sah[1] · Abhishek Kumar[1]

Accepted: 18 July 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Past studies have indicated significant uncertainties in determining the safety factor of slope using a deterministic approach. To reduce these uncertainties and enhance the stability of slopes, this study utilized machine learning (ML) techniques. The primary goal was to develop an efficient ML model to predict the factor of safety (FOS) of slope in c-φ soil. Three ML models were evolved: artificial neural network (ANN), Gaussian process regression (GPR) and hybrid ANN model, which combines with the meta-heuristic optimization technique namely particle swarm optimization (PSO) to make ANN-PSO. Five input parameters, i.e. unit weight of soil, cohesion, angle of shear resistance, slope angle and slope height, are used to compute FOS. The efficacy of the ML models is evaluated using a range of performance indicators, such as coefficient of determination (R^2), variance account factor, Legate and McCabe's index, a-10 index, root mean square error (RMSE), RMSE-observations standard deviation ratio, mean absolute error and median absolute deviance in both the training (TR) and testing (TS) stages. Among all the models, ANN-PSO performed better due to its higher value of R^2 (TR = 0.932, TS = 0.833 and Overall = 0.920) and lowest value of RMSE (TR = 0.060, TS = 0.073 and Overall = 0.063) followed by GPR and ANN. The reliability index is calculated using the first-order second moment method for all the models and compared with the observed value. Further tools used to evaluate the model's performance are rank analysis, reliability index, regression curve, William's plot, error matrix and confusion matrix. The overall performance of ANN-PSO is superior to the other two ML models while predicting FOS. The influence of each input parameter on the output is also computed using sensitivity analysis.

Keywords ANN · GPR · PSO · Machine learning · Slope stability · William's plot · Confusion matrix

Extended author information available on the last page of the article

Published online: 05 August 2024

Abbreviations

Symbol	Nomenclature
ML	Machine learning
FOSM	First-order second moment method
MCS	Monte Carlo simulation
FOS	Factor of safety
MDRM	Multiplicative dimensional reduction method
c	Cohesion
φ	Angle of shear resistance
γ	Unit weight of soil
β	Reliability index
P_f	Probability of failure
PEM	Point estimate method
MPEM	Modified point estimate method
RFEM	Random finite element method
RFM	Random finite method
H	Slope height
α	Slope angle
u	Pore pressure
c'	Effective cohesion
LEM	Limit equilibrium method
ANN	Artificial neural network
ANFIS	Adaptive neuro fuzzy inference system
PSO	Particle swarm optimization
NS	Nash Sutcliffe efficiency
LMI	Legate and McCabe's index of agreement
U_{95}	Expended uncertainty
RMSE	Root mean square error
VAF	Variance account factor
R^2	Coefficient of determination
Adj. R^2	Adjusted coefficient of determination
PI	Performance index
BF	Bias factor
RSR	Root mean standard deviation ratio
NMBE	Normalized mean bias error
MAPE	Mean absolute percentage error
RPD	Relative percentage difference
WI	Willmott's index of agreement
MBE	Mean bias error
MAE	Mean absolute error
GPI	Global performance indicator
GPR	Gaussian process regression
SVR	Support vector regression
DT	Decision trees
LSTM	Long-short term memory
DNN	Deep neural network

KNN	K-nearest neighbour
r_u	Pore pressure ratio
SVM	Support vector machine
MARS	Multivariate adaptive regression spline
GRNN	Generalized regression neural network
S_u	Undrain strain parameter
ELM	Extreme learning machine
ABC	Artificial bee colony algorithm
LSSVM	Least square support vector machine
R	Coefficient of correlation
MLP	Multilayer perceptron
RF	Random forest
γ_d	Dry density
GEP	Gene expression programming
ARE	Average relative error
GBM	Gradient boosting machine
ROC	Receiver characteristics operator
AUC	Area under curve
LR	Linear regression
BR	Bayesian ridge
ENR	Elastic net regression
GBR	Gradient boosting regression
SPCE	Sparse polynomial chaos expansion
BN	Bayesian network
HPO	Hyper parameter optimization
GBAEF	Global best artificial electric field algorithm
MCMC	Bayesian Markov chain Monte Carlo method
SCA	Sine cosine algorithm
ASCPS	Adaptive sine cosine algorithm-pattern search
GA	Genetic algorithm
FFA	Firefly algorithm
GWO	Grey wolf optimization
XGBoost	Extreme gradient boosting
ACO	Ant colony optimization
ALO	Artificial lion optimization
ICA	Imperialist competitive algorithm
SCE	Shuffled complex evolution
τ_f	Mobilized shear strength
τ	Shear stress
σ_n	Normal stress
σ_w	Standard deviation
μ_w	Average value
PF	Processing factor
SR	Standardized residuals
L	Leverage

SA Sensitivity analysis
SOR Strength of relation

1 Introduction

The analysis of slope stability is a crucial issue in the field of civil engineering, playing an important role in geotechnical engineering. It depends on many factors such as soil properties, geological structure, environmental conditions and more. By utilizing statistical and probabilistic techniques, machine learning models evaluate soil slope behaviour and stability through probabilistic analysis. This method provides detailed information about uncertainties surrounding slope stability, aiding engineers and geologists in making better decisions.

Many researchers have used the analytical method for the reliability analysis of slope stability. Malkawi et al. (2020) used two analytical methods, namely the FOSM method and Monte Carlo simulation (MCS) to predict the FOS. They took cohesion (c), angle of shear resistance (φ) and unit weight of soil (γ) as input parameters. Model performances were judged using computation of the β. Wang et al. (2022) used the multiplicative dimensional reduction method (MDRM), FOSM and MCS for the probabilistic analysis of slope stability. They have also computed β and probability of failure (P_f) value. Griffiths and Fenton (2004) used two types of analytical methods, namely the random finite element method (RFEM) and MCS to predict FOS. Based on the P_f, value, they concluded that the RFEM is better than other analytical methods. Griffiths et al. (2011) used three analytical methods FOSM, FORM and random finite method (RFM) to predict FOS. They took slope height (H), c, slope angle (α), γ, pore pressure (u), effective cohesion (c') and effective angle of shear resistance (φ') as input parameters. Johari and Javadi (2012) considered the reliability assessment of infinite slope stability using the jointly distributed random variables method. They took three input parameters, namely c, γ and φ. The reliability assessment of the stability of infinite slope is compared with MCS, PEM and FOSM to compute FOS. Al-Karni and Al-Shamrani (2000) employed the study of the effect of soil anisotropy on slope stability using method of slices. They have used analytical method, namely limit equilibrium method (LEM) to compute FOS.

To anticipate the FOS, numerous researchers have employed a numerous machine learning (ML) techniques. Ray and Roy (2021) used soft computing techniques for the reliability analysis of slope stability in soil. They have used two types of soft computing techniques namely ANN and adaptive neuro fuzzy inference system (ANFIS) to predict FOS individually and make one hybrid model (PSO-ANN) by using metaheuristic optimisation technique namely particle swarm optimization (PSO). They took γ, c and φ as input parameters. The model's performance was judged by numerous performance parameters like Nash Sutcliffe efficiency (NS), Legate and McCabe's index (LMI), expended uncertainty (U_{95}), root mean square error (RMSE), variance account factor (VAF), coefficient of determination (R^2), t-statistics, adjusted coefficient of determination (Adj. R^2), performance index (PI), bias factor (BF), root mean standard deviation ratio (RSR), normalized mean bias error (NMBE), mean absolute percentage error (MAPE), relative percentage difference (RPD), Willmott's index of agreement (WI),

mean bias error (MBE), mean absolute error (MAE), global performance indicator (GPI) and reliability index (β). Based on various performance parameters, they concluded that the PSO-ANN model performed better compared to the other two models. Mahmoodzadeh et al. (2022) conducted a comparison of soft computing techniques for predicting of safety factors of slope stability in soil. They used six soft computing ML models, namely GPR, support vector regression (SVR), decision trees (DT), long-short term memory (LSTM), deep neural network (DNN) and k-nearest neighbour (KNN) to predict FOS of the considered slope. They took γ, H, α, φ, c and pore pressure ratio (r_u) as input parameters. The model's performance was assessed by numerous parameters like R^2, MAE, RMSE and MAPE. Based on these performance parameters, the GPR model was identified as the most effective among the five ML models. Zhao (2008) used support vector machine (SVM) as a soft computing technique to predict FOS of finite slope in soil. They have taken three type of input parameter, such as γ, φ and c. Model performances were judged using computation of β. β is computed using FOSM method. Cho (2009) used ANN soft computing technique to predict FOS of slope stability in soil. They have taken two input parameters, such as c and φ. Model performances were judged using computation of β and P_f value. Kumar et al. (2017) used soft computing techniques, namely ANFIS, GPR, multivariate adaptive regression spline (MARS) and generalized regression neural network (GRNN) to predict FOS of infinite slope in soil. They have considered three input parameters namely c, φ and γ. The ML model performance was assessed using computation of β. They concluded that the MARS model is better as compared to three ML models. Gupta et al. (2022) used two types of soft computing techniques, DNN and ANN, to predict FOS of slope stability in soil. They considered input parameters, such as undrain strain parameter (S_u) and φ. The model was assessed using numerous performance parameters such as CC, RMSE, MAE, SI and NSE. Based on performance parameters, they concluded that the DNN model performed better than ANN ML models. Kang et al. (2016) used metaheuristic optimization technique, such as PSO and artificial bee colony algorithm (ABC), individually to predict the FOS of slope stability in soil and made one hybrid model least square support vector machine with particle swarm optimization (PSO-LSSVM) by using metaheuristic optimization technique namely PSO. They took two input parameters, such as c and φ. They used various performance parameters, such as MAE, RMSE and coefficient of correlation (R). Based on various performance parameters, they concluded that the PSO-LSSVM model performed better compared to other ML models. Nanehkaran et al. (2022) used five different ML models, namely multilayer perceptron (MLP), SVM, KNN, DT and random forest (RF) to predict FOS of slope stability in soil. They took five input parameters like H, α, c, φ and dry density (γ_d). They used various performance parameter such as confusion matrix (accuracy, precision, recall and F1-score), MAE, MSE and RMSE. They concluded that the MLP gives best performance as compared to other ML models. Johari et al. (2018) used gene expression programming (GEP) to predict the FOS of unsaturated slope stability in soil. They have taken input parameter such as void ratio, water content, clay content, silt content, c', φ' and γ. They have used numerous performance parameters namely average relative error (ARE), MSE and R^2. Deris et al. (2021) used soft computing technique, namely SVM and DT to predict the FOS of slope stability in soil. They have taken five input parameter such as γ, c, φ, α and H. They used confusion matrix to check the performance of

the model. Based on performance parameter, they have concluded that SVM model performs better than DT. Yang et al. (2023) used soft computing techniques like SVM, RF, DT, KNN and gradient boosting machine (GBM) to predict the FOS of slope stability in soil. They have taken c, φ, γ, H, r_u and α as input parameters. They have analyzed them using the confusion matrix, receiver characteristics operator (ROC) and area under curve (AUC) value. They have concluded that RF performed better as compared to the other four model. Liu et al. (2021) used ELM as soft computing technique to predict the FOS of slope stability in soil. They took three input parameters, such as γ, c and φ. They analyzed them using the R^2. Lin et al. (2021) used various types of ML models, namely linear regression (LR), Bayesian ridge (BR), elastic net regression (ENR), SVM, DT, RF, adaptive boosting machine, KNN, GBM, bagging, and extra trees, to predict the FOS of finite slope in soil. They used three performance parameters such as R^2, MAE and MSE. Based on performance parameters, they have concluded that gradient boosting regression (GBR), SVM, and bagging are better performed than another model. Xu et al. (2023) used numerous types of ML techniques, namely ANN, ANFIS, DNN, SVM, GPR, DT, RF, sparse polynomial chaos expansion (SPCE), Bayesian network (BN) and hyper parameter optimization (HPO), to predict the FOS of slope stability in soil. They have used performance parameters such as R^2, MSE, RMSE, MAE and MAPE. According to performance parameter, they concluded that ANFIS is better as compared to the other model. Khajehzadeh and Keawsawasvong (2023) used two types of ML model, namely, SVR and global best artificial electric field algorithm (GBAEF), to predict the FOS of finite slope in soil. They took five input parameters γ, c, φ, α, H and r_u. They used four types of performance parameters, such as RMSE, MAE, MSE and R^2. Based on the performance parameter, SVR performs better as compared to the other models. Kang et al. (2015) used GPR computational ML model to predict the FOS of slope stability in soil. Ahmad et al. (2023) used ensemble ML models namely decision tree regression (DTR), multiple linear regression (MLR), k-nearest neighbour regression (KNN), random forest regression (RF), extreme gradient boosting regression (XGB) and support vector regression (SVR) to predict FOS values for railway embankments. Fattahi and Ilghani (2019) used ML technique for slope stability analysis using Bayesian Markov chain Monte Carlo method (MCMC) using the software win BUGS to predict the FOS of slope stability in soil. They have considered six input parameters, such as γ, c, φ, α, H and r_u. The ML model was assessed by using the performance parameter such as R^2 and MSE. Khajehzadeh et al. (2022) used four types of ML model namely ANN, sine cosine algorithm (SCA) and adaptive sine cosine algorithm-pattern search (ASCPS) to predict the FOS of infinite slope in soil. They used various performance parameters such as RMSE and R-value. Based on model performance parameters, they obtained ANN model performance better than the other model. Ahmad et al. (2023) used convolutional neural networks (CNN), deep neural networks (DNN), artificial neural networks (ANN) and multiple linear regression (MLR) to predict FOS of railway embankment. Sabri et al. (2024) used artificial neural networks (ANN), Bayesian neural networks (BNN), convolutional neural networks (CNN) and deep neural networks (DNN) to evaluate the slope stability of a heavy haul freight corridor.

Some researchers have also employed various types of ML technique for probabilistic analysis in the other geotechnical engineering fields. Yousuf et al. (2023) used

ANN and ELM to predict the load sustaining capability of the rectangular footing. Mustafa et al. (2022) utilized four types of meta-heuristic optimization techniques, namely PSO, genetic algorithm (GA), firefly algorithm (FFA) and grey wolf optimization (GWO) to predict FOS of gravity retaining wall individually and make hybrid with soft computing model ANFIS such as ANFIS-PSO, ANFIS-FFA, ANFIS-GA and ANFIS-GWO. Mustafa et al. (2024) utilize three soft computing models, ANFIS, ELM and extreme gradient boosting (XGBoost), to predict the thermal conductivity of soil. Kumar et al. (2024) used soft computing technique ANN and ANN make hybrid with four different type of metaheuristic optimization such as ant colony optimization (ACO), artificial lion optimization (ALO), imperialist competitive algorithm (ICA) and shuffled complex evolution (SCE) to predict the bearing capacity of strip footings subjected to inclined loading.

The probabilistic analysis of stability of slope in c-φ soil is rarely replicated by the application of machine learning (ML) methods. The aim of this study is to develop reliable machine learning models for computing safety factor for the soil slope. The literature review shows that the artificial neural network (ANN), hybrid model of ANN with PSO (ANN-PSO) and Gaussian process regression (GPR) models have not yet been widely used to evaluate the prediction of FOS of c-φ soil slope. In order to eliminate the need for the tedious, conventional and time-consuming calculating procedure, ML models are intended to give a method that is ready for use and that can compute the FOS of slope in c-phi soil. In order to evaluate the probabilistic analysis of slope in c-φ soil, the ML models were trained taking into account cohesion (c), unit weight of soil (γ), angle of shearing resistance (φ), slope angle (α) and height of soil slope (H). The performance of the proposed model was thoroughly assessed using a range of performance parameters, including rank analysis, regression curves, reliability index, William's plot, error matrix and confusion matrix. Geotechnical engineers may find the proposed models to be a helpful tool in rapidly and simply estimating the FOS of slope in c-φ soil, even with limited computer knowledge.

2 Details of the Present Study

Using conventional approach factor of safety (FOS) for an infinite slope in c-φ soil is defined as the ratio of mobilized shear strength $\left(\tau_{\mathrm{f}}\right)$ of the soil to the shear stress (τ) acting on the soil. The FOS can be expressed mathematically as:

$$\mathrm{FOS} = \frac{\tau_{\mathrm{f}}}{\tau} \tag{1}$$

As per Mohr's coulombs theory which is described in Das (1985), the mobilized shear strength $\left(\tau_{\mathrm{f}}\right)$ of soil normally expressed as a function of c and φ. The mathematical expression for the τ_{f} can be written as:

$$\tau_{\mathrm{f}} = c + \sigma_n \tan\varphi \tag{2}$$

where c is the cohesion of the soil, φ is the angle of shear resistance and σ_{n} is the normal stress acting on the soil which can be expressed as follows:

$$\sigma_{\mathrm{n}} = \gamma H \cos^2 \alpha \tag{3}$$

The shear stress acting on the soil (τ) can be expressed as:

$$\tau = \gamma H \cos\alpha \sin\alpha \tag{4}$$

where γ is the unit weight of soil, H is the slope height and α is the slope angle. From Eqs. (1–4), the FOS can be calculated as follows:

$$\mathrm{FOS} = \frac{C + \gamma H \cos^2 \alpha \tan \varphi}{\gamma H \cos \alpha \sin \alpha} \tag{5}$$

Reliability analysis in geotechnical engineering deals with the uncertainty in structural systems and soil mechanics, which can be due to variations in soil properties, or the analytical methods used. The goal is to observe this uncertainty or find a reliable method to predict the FOS for an infinite slope in c-φ soil. The FOSM approach is useful for calculating the β. In this approach, σ_{Y} and μ_{Y} represent the standard deviation and average value of the performance function Y, respectively. τ_{f} and τ are represented as P and Q, respectively, and μ_{P} and μ_{Q} are the mean values and σ_{P} and σ_{Q} are the standard deviation of P and Q, respectively.

$$\beta = \frac{\mu_Y}{\sigma_Y} = \frac{\mu_P - \mu_Q}{\sqrt{\sigma_P^2 + \sigma_Q^2}} \tag{6}$$

Since the standard deviation and mean of the observed FOS have an important impact on the β, the FOS can be calculated as follows:

$$\beta = \frac{\mu_{\mathrm{FOS}} - 1}{\sigma_{\mathrm{FOS}}} \tag{7}$$

where the mean and standard deviation of the observed FOS are represented by μ_{FOS} and σ_{FOS}, respectively.

3 Proposed ML Model

As noted in the previous section, many researchers have employed various statistical parameters to conduct probabilistic analysis of an infinite slope. The primary objective of this study is to provide the most effective ML models, enabling designers and engineers to calculate the FOS for infinite slope in c-φ soil. The context of soft computing techniques briefly explained in the subsequent parts.

3.1 Artificial Neural Network (ANN)

The ANN model relies on weighted simulated neurons and various specialized units, simulated by the genetic neural system, which shown in Fig. 1 (Anand et al., 2021). ANNs

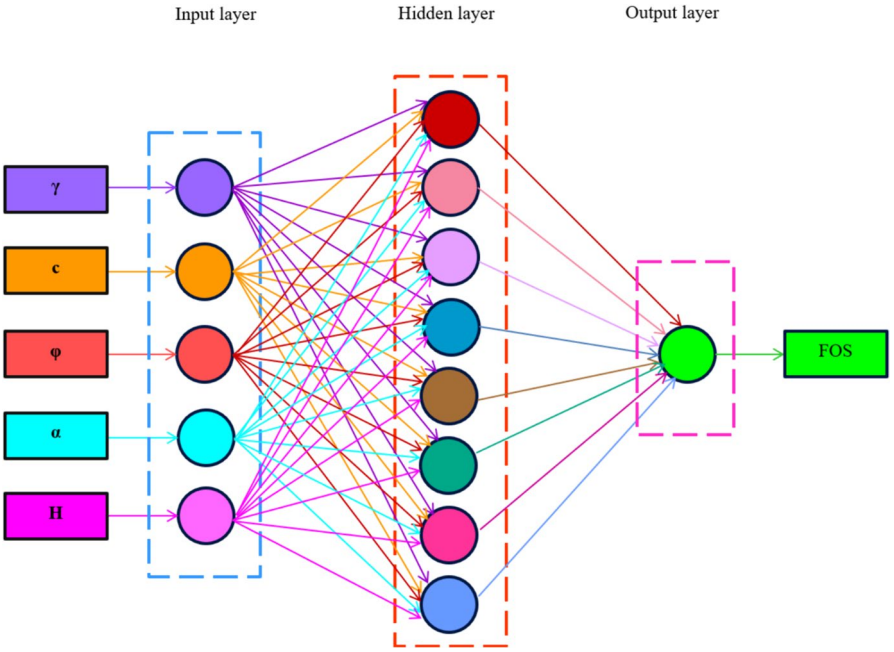


Fig. 1 Schematic representation for an ANN

are also known as a processing factor (PF) that serve as information processors. Each PF has one or more transfer functions, an output and weighted inputs. Generally, a PF is an equation that equates the output and input. As the connection weights put the system's memory, ANNs are identified as connectionist models. An ANN depend on three major sets of layers with neurons. The last and first layers of the ANN are designated as the output and input layers, respectively. They have more neurons representing the output and input variables. The hidden layer is located between these two layers. The hidden layers serve as the predictors, and the decision-making process is show within the output layer. Biases and weights are important parameters of an ANN. Biases determine the degree of freedom, while weights show the interconnection between a layer's neurons. Every node, apart from the input nodes, receives the output by using a nonlinear activation function. The output serves as the input for the next node. This process repeats until a correct solution to the given problem is not obtained. The primary objective of the problem is to calculate the error by comparing the prediction (i.e. the outcome of the network) to the actual result. Backpropagation is utilized for error transmission. Then, the error is transmitted back to a layer back at a time throughout the ANN structure. Adjustments are made to the weights based on the portion of the error. Table 1 depicts the hyperparameters of ANN.

Table 1 Hyperparameters of the model ANN

Parameters	ANN
Batch size	08
Number of epoch	100
Input layer	40 Neurons
Hidden layer	40 Neurons
Output layer	1 Neuron
Optimizer	Adam
Loss function	RMSE

3.2 Particle Swarm Optimization (PSO)

The PSO algorithm simulates the common actions of organisms, mostly bird gathering or fish schooling. It was proposed by the authors Kennedy and Eberhart (1995). The main goal of this algorithm is to find a universal optimal solution in the search space. It initializes the particles with random locations and velocities. After updating their position, each particle determines its individual and global most suitable position in the search space. The best position in each particle, along with the suitable ranking of particles, shows the most suitable state in the universe. The discrete most suitable direction and position in the universe plays an important role in the updating process of a particle. The difference between the global most suitable state and each particle's most suitable position is important for revising the particle's speed. As represent in Fig. 2,

$$v_n^{t+1} = uv_n^t + a_1 r_1 \left(\mathrm{p} - \mathrm{best_n} - \mathrm{y_n^t}\right) + a_2 r_2 \left(\mathrm{g} - \mathrm{best_n} - \mathrm{y_n^t}\right) \tag{8}$$

$$y_n^{t+1} = y_{n+}^t v_n^{t+1} \tag{9}$$

where v and y denote velocity and position, respectively; p-best and g-best shows the best particle position and suitable group position, respectively; r_1 and r_2 are the random number between 0 and 1 and $\mathrm{a_1}$ and $\mathrm{a_2}$ are the cognitive and social coefficients. These parameters are problem solving; thus, their main aim is to decide the level of dependence of a particle on its universal and individual position, which is determined by its inertia weight parameter, denoted as u^{k}. This parameter affects the particle's movement within the universe space which directly depend on time, which expressed as:

$$u\mathrm{k} = u_{\max} - \frac{u_{\max} - u_{\min}}{\mathrm{t_{max}}} t \tag{10}$$

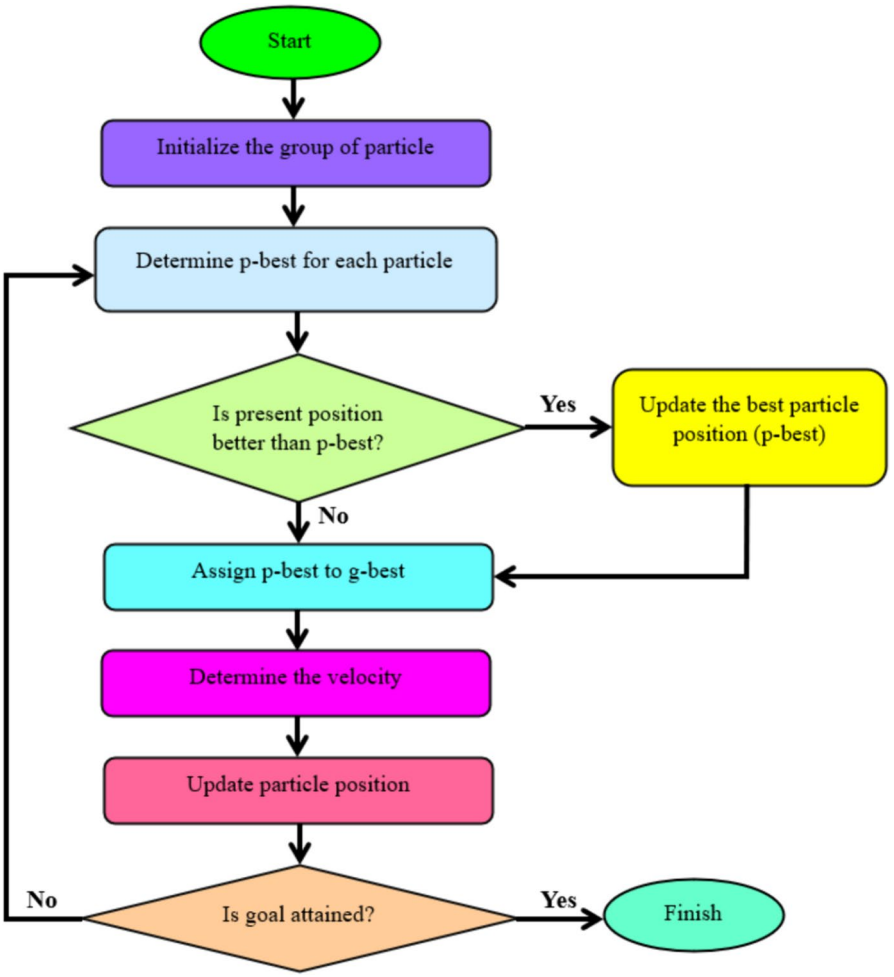


Fig. 2 Steps presenting the process involved in PSO

Here, u_{max} and u_{min} show the maximum and minimum inertia weights, respectively; t_{max} is defined based on the highest number of iterations and k represents the current iteration number. The value of u^k effects the particle's behaviour during optimization. A low value of u^k defines the particle's individual best position and higher value of u^k gives more weight to the global best position, inspiring study of the search space. Figure 2 indicates the flowchart of PSO. Table 2 shows the hyperparameters used for PSO.

Table 2 Hyperparameters used for PSO

Parameters	PSO
Population size	50
Number of iterations	100
Cognitive coefficient	0.50
Social coefficient	0.30
Inertia weight	0.90
Number of epochs	100
Batch size	08
Input layer	64 Neurons
Hidden layer	40 Neurons
Output layer	01 Neuron
Optimizer	Adam
Loss function	RMSE

3.3 Gaussian Process Regression (GPR)

The GPR is an informal Bayesian is an informal Bayesian approach to controlled learning that models the link between input and output parameters, as discussed by Mahmoodzadeh et al. (2021) and Zhu et al. (2021). It is applied to issues involving classification as well as regression.

The Gaussian process is a generalization of the Gaussian distribution and is used to define the variation in functions, whereas the Gaussian distribution can be used to describe the distribution of random variables. The covariance function $n\left(z,\ z'\right)$ and the mean function j (z) in the function space can be used to create a Gaussian process.

$$j(z) = E(g(z)) \tag{11}$$

$$n\left(z, z'\right) = E\left(\left(g(z) - j\left(z'\right)\right)\left(g\left(z'\right) - j\left(z'\right)\right)\right) \tag{12}$$

The Gaussian process may be expressed as:

$$g(z) = GP \sim \left(j(z),\ n\ (z, z')\right) \tag{13}$$

In general, at this point, we take the mean function as zero for the purpose of simplicity in notation, as discussed by Wang et al. (2019). Considering a dataset L containing n observations ($L = \{(z_i, o_i) \mid i = 1, 2, 3, \ldots, n\}$), where the output scalar is denoted by o_i and z_i is the input vector, which is M-dimensional.

$$o = g(z) + \varepsilon \tag{14}$$

where g(z) is the random regression function and Ɛ is the Gaussian noise with a variance of ${\sigma_n}^2(\text{Ɛ} \sim N(0,\ {\sigma_n}^2))$ and an independent, identically distributed Gaussian distribution. Two matrices, $Z = [z_1, z_2, \ldots, z_n]$ and $O = [o_1, o_2, \ldots, o_n]$, represent the input and output data, respectively. The group of functions $g = [g(z_1),\ g(z_2),\ \ldots.,\ g(z_n)]^P$ follows the Gaussian process is to determine $(q(g|T) = B(0,\ K))$, where K is the covariance function matrix $n(z, z')$.

Table 3 Hyperparameters of GPR model

Parameters	GPR
Input layer	64 Neurons
Hidden layer	40 Neurons
Learning rate	0.005
Batch size	08
Adam decay	0.000001
Output layer	01 Neuron
Number of hidden layers	06

$$K(Z,Z)=\begin{bmatrix} n(z_1,z_1) & \cdots & n(z_1,z_n) \\ \vdots & \ddots & \vdots \\ n(z_n,z_1) & \cdots & n(z_n,z_n) \end{bmatrix} \tag{15}$$

The multivariate normal distribution is equally distributed as the prediction output, which includes the training outputs (o) and testing outputs (o^*). When testing at a predetermined position, the joint distribution of the actual target value (o) and the predicted value (o^*) is expressed as follows.

$$\begin{bmatrix} o \\ o^* \end{bmatrix} \backsim N\left(0, \begin{bmatrix} K(Z,Z) & K(Z,Z^*) \\ K(Z^*,Z) & K(Z^*,Z^*) \end{bmatrix}\right) \tag{16}$$

Then, the predictive distribution of the function values o^* at test points $Z^*=[z_1^*, z_2^*, \ldots., z_n^*]$ is calculated using GPR.

$$q(o^*|Z^*,Z,O) \sim N(\overline{g^*}, cov(g^*)) \tag{17}$$

$$\overline{g^*} = K(Z^*,\ Z)K(Z,\ Z)^{-1}O \tag{18}$$

$$\mathrm{COV}(g^*)=K(Z^*,Z^*)-K(Z^*,Z)^{-1}K(Z,Z^*) \tag{19}$$

One of the most popular uses of Gaussian process regression is Bayesian optimization, as discussed by Sonek et al. (2012). By utilizing a Gaussian process model of the objective function to estimate the next evaluation point, Bayesian optimization techniques effectively identify the best value when evaluating the objective function which is highly costly. Table 3 shows the hyperparameters of GPR model.

4 Dataset Preparation

The uncertainty of the FOS for an infinite slope in c-φ soil is directly correlated with the analytical method used. For the reliability analysis, we have generated total 200 datasets to predict the FOS. We have generated the datasets randomly using the command NORM.INV (RAND (), Mean, Standard deviation) in MS-Excel. Statistical input parameters such as γ, c, α, φ and H were prepared, which follow the normal distribution function. For this study, datasets were generated using mean and standard deviation values taken from previous studies by Nanehkaran et al. (2023). Table 4 shows the descriptive statistics of the collected dataset, and Fig. 3 represents

Table 4 Statistical overview of input and output data

Parameters	Input parameters					Output
	Unit weight of soil (γ) (in kN/m^3)	Cohesion (c) (in kN/m^2)	Angle of shearing resistance (φ) (in Degree)	Slope angle (α) (in Degree)	Slope height (H) (in m)	Factor of safety (FOS)
Minimum	14.539	7.824	26.623	11.953	4.203	0.442
Maximum	21.788	300.927	39.185	88.558	46.957	6.746
Mean	18.000	139.000	33.500	56.000	18.500	2.243
Standard deviation	1.4142	63.053	2.1213	22.627	9.1923	1.412
Standard error	0.099	4.458	0.149	1.599	0.649	0.099
First quartile (Q1)	16.910	96.831	32.051	40.028	13.000	1.259
Second quartile (Q2)	17.989	136.230	33.556	56.021	20.165	1.752
Third quartile (Q3)	18.890	186.027	35.004	67.074	25.225	2.767
Range	7.249	293.102	12.562	76.605	42.753	6.304
Kurtosis	−0.265	−0.610	0.193	−0.680	−0.289	1.479
Skewness	0.137	0.084	−0.057	−0.306	0.348	1.424

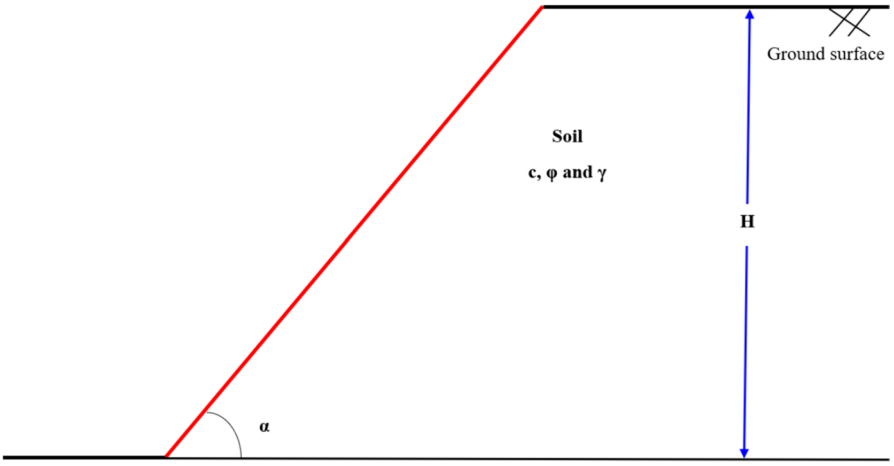


Fig. 3 Geometry of considered slope for this study

considered slope for this study which is taken from the previous study of Feng et al. (2018).

The pre-processing steps involved in preparing a dataset for a machine learning model include normalization, a processing technique used to remove the dimensional influence of the variables. In this case, the value of both input and output variables were normalized between 0 and 1. The normalization of the dataset can be done as follows:

$$F_{Normalized} = \left(\frac{F - F_{min}}{F_{max} - F_{min}} \right) \tag{20}$$

where $F_{\max}$ and $F_{\min}$ are the maximum and minimum values of the parameter (F), respectively. After normalization has been performed, the total datasets are divided into two phases namely training and testing phases. For this, 80% of the entire dataset (160 datasets) is used for training the machine learning model, and the remaining 20% (40 datasets) is used for testing the performance. Figure 4 indicates the methodology flowchart.

5 Statistical Performance Indicators

Using numerous types of statistical performance parameters and graphical judgments such as radar diagram and R-curve, the prediction accuracy of soft computing models ANN, GPR and ANN combined with one metaheuristic optimization

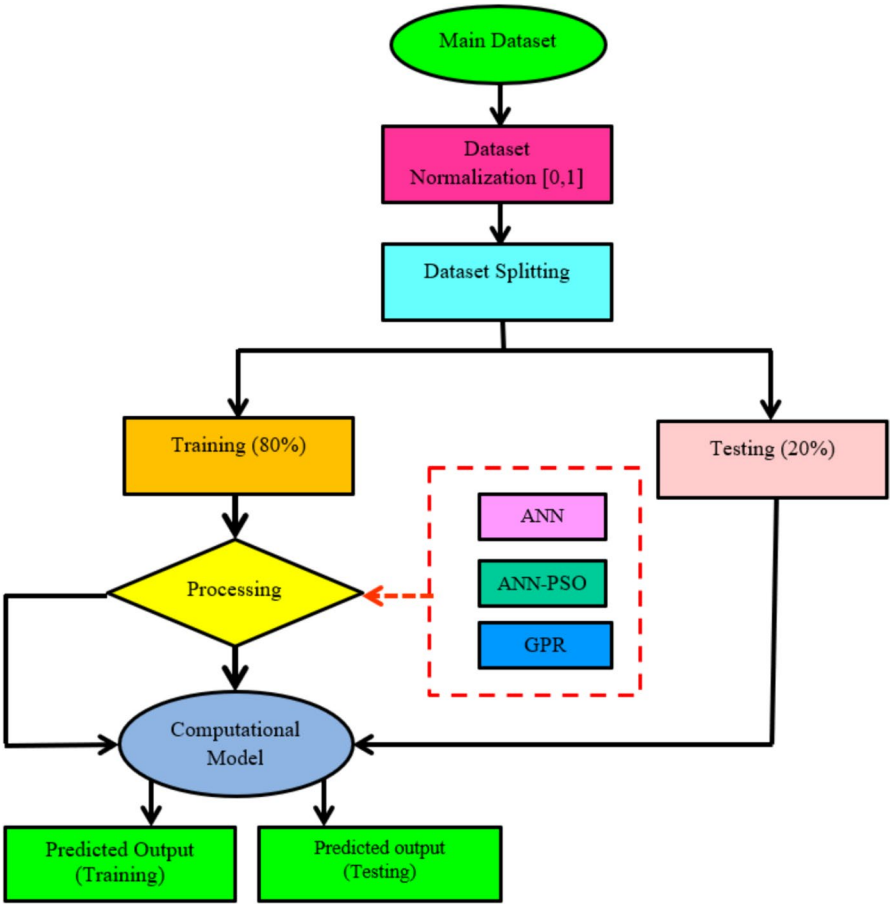


Fig. 4 Methodology flowchart

technique as PSO to create a hybrid model known as ANN-PSO was investigated. Various statistical performance parameters are used, including trained measuring parameters like R^2, VAF, LMI and a-10 index and error measuring parameters like RMSE, RSR, MAE and MAD.

$$R^2 = \frac{\sum_{i=1}^{n}\left(\mathrm{FOS}_{\mathrm{obs,i}} - \overline{\mathrm{FOS}}\right)^2 - \sum_{i=1}^{n}\left(\mathrm{FOS}_{\mathrm{obs,i}} - \mathrm{FOS}_{\mathrm{pred,i}}\right)^2}{\sum_{i=1}^{n}\left(\mathrm{FOS}_{\mathrm{obs,i}} - \overline{\mathrm{FOS}}\right)^2} \tag{21}$$

$$\mathrm{VAF} = \left(1 - \frac{\mathrm{var}\left(\mathrm{FOS}_{\mathrm{obs,i}} - \mathrm{FOS}_{\mathrm{pred,i}}\right)}{\mathrm{var}\left(\mathrm{FOS}_{\mathrm{obs,i}}\right)}\right) \times 100 \tag{22}$$

$$\mathrm{LMI} = 1 - \left[\frac{\sum_{i=1}^{n}\left|\mathrm{FOS}_{\mathrm{obs,i}} - \mathrm{FOS}_{\mathrm{pred,i}}\right|}{\sum_{i=1}^{n}\left|\mathrm{FOS}_{\mathrm{obs,i}} - \overline{\mathrm{FOS}}\right|}\right] \tag{23}$$

$$\mathrm{a} - 10\,\mathrm{Index} = \frac{\mathrm{K10}}{\mathrm{n}} \tag{24}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\left(\mathrm{FOS}_{\mathrm{obs,i}} - \mathrm{FOS}_{\mathrm{pred,i}}\right)^2} \tag{25}$$

$$\mathrm{RSR} = \frac{\mathrm{RMSE}}{\sqrt{\frac{1}{\mathrm{N}}\sum_{i=1}^{n}\left(\mathrm{FOS}_{\mathrm{obs,i}} - \overline{\mathrm{FOS}}\right)}} \tag{26}$$

$$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n}\left|\left(\mathrm{FOS}_{\mathrm{pred,i}} - \mathrm{FOS}_{\mathrm{obs,i}}\right)\right| \tag{27}$$

$$\mathrm{MAD} = \mathrm{Median}\left(\left|\mathrm{FOS}_{\mathrm{pred,1}} - \mathrm{FOS}_{\mathrm{obs,1}}\right|, \left|\mathrm{FOS}_{\mathrm{pred,2}}\right| - \mathrm{FOS}_{\mathrm{obs,2}}|, \ldots\ldots\ldots\ldots\ldots, |\,\mathrm{FOS}_{\mathrm{pred,n}} - \mathrm{FOS}_{\mathrm{obs,n}}\right|\Big) \tag{28}$$

where the real and predicted ith values in this scenario are denoted by $\mathrm{FOS}_{\mathrm{obs,i}}$ and $\mathrm{FOS}_{\mathrm{pred,}}$i, respectively. The mean of actual value is represented by $\overline{\mathrm{FOS}}$, the number of training or testing sample is denoted by *n*, the model input capacity is denoted by *k* and the quantity of data with observed/predicted value ratio between 0.90 and 1.10 is found in k10. Equations (21–28) are used to determine that the model has a minimum value of RMSE, RSR, MAE and MAD and a higher value of R^2, VAF, LMI and a-10 index.

6 Result and Discussion

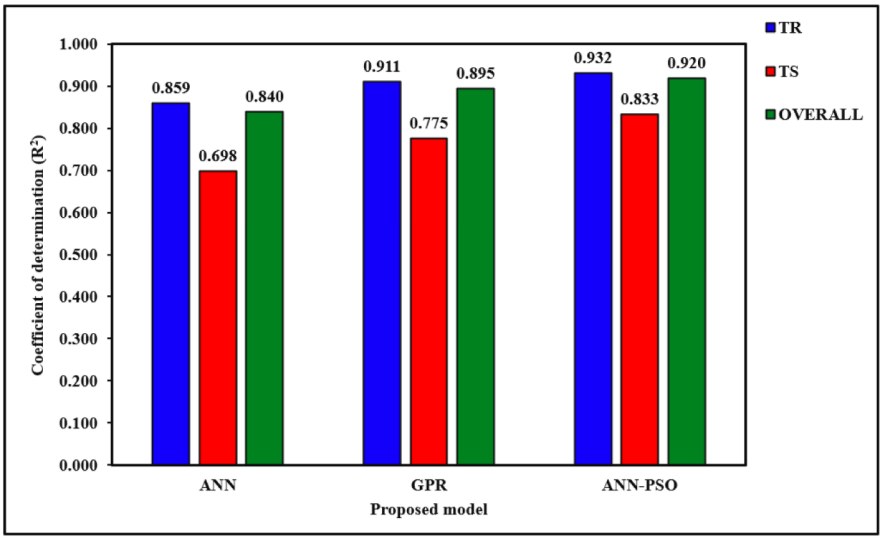
6.1 Prediction Power of the Proposed Model

In this study, the factor of safety (FOS) of an infinite slope is predicted using three ML techniques, namely ANN, GPR and a hybrid model of ANN (ANN-PSO). The performance of these models are evaluated by calculating statistical parameters, namely R^2, VAF, LMI, a-10 index, RMSE, MAE, RSR and MAD. The results of calculation statistics for the training (TR) and testing (TS) stages are provided in Table 5. The comparison of two soft computing techniques and the hybrid model was based on statistical indices. According to the result, it has been noted that ANN-PSO achieved better prediction performance with higher value of $R^2 = 0.931$, VAF = 93.569, LMI = 0.754 and a-10 index = 0.381 and a lower value of RMSE = 0.060, RSR = 0.260, MAE = 0.436 and MAD = 0.031 in training stage, whereas the slightly decrease in the testing stage ($R^2 = 0.833$, VAF = 87.240, LMI = 0.667 and a-10 index = 0.350, RMSE = 0.073, RSR = 0.408, MAE = 0.043 and MAD = 0.026. Among the three ML models, it can be concluded that the ANN-PSO model provided better predictions than the other two models.

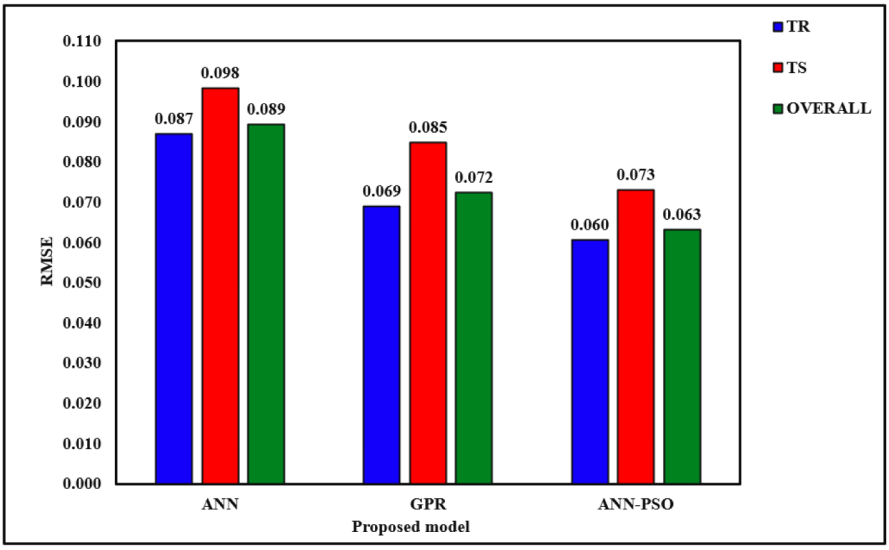
Figure 5a indicates the comparison of the R^2 values for the training (TR) and testing (TS) as well as the overall datasets of the proposed models. The model with the highest R^2 value is considered the best model. In Fig. 5a, the ANN-PSO model has the highest R^2 value in the TR and TS datasets, as well as for the overall datasets, indicating its superiority over the other two models. Similarly, Fig. 5b shows a comparison of the RMSE values for the TR, TS and overall datasets of the proposed models. The model with the lowest RMSE value is deemed the best predicting model. In Fig. 5b, the ANN-PSO model demonstrates the lowest RMSE value across the TR, TS and overall datasets, thereby surpassing the other two models.

Table 5 performance parameters of the proposed models

Parameter	ANN		GPR		ANN-PSO		Ideal value
	Training	Testing	Training	Testing	Training	Testing	
R^2	0.858	0.698	0.911	0.775	0.931	0.833	1
VAF	85.919	70.868	91.135	77.850	93.569	87.240	100
LMI	0.666	0.495	0.720	0.550	0.754	0.667	1
a-10 Index	0.306	0.250	0.356	0.225	0.381	0.350	1
RMSE	0.087	0.098	0.069	0.084	0.060	0.073	0
RSR	0.375	0.549	0.297	0.473	0.260	0.408	0
MAE	0.059	0.065	0.049	0.058	0.043	0.043	0
MAD	0.042	0.042	0.037	0.045	0.031	0.026	0



(a)



(b)

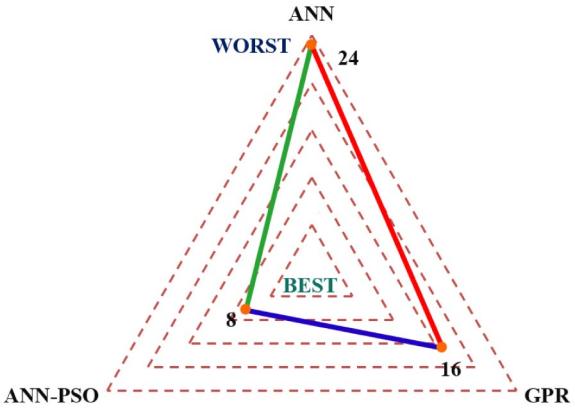
Fig. 5 Comparison of performance parameters for training and testing as well as overall datasets. **a** Coefficient of determination ($R.^2$) and **b** root mean square error (RMSE)

6.2 Rank Study

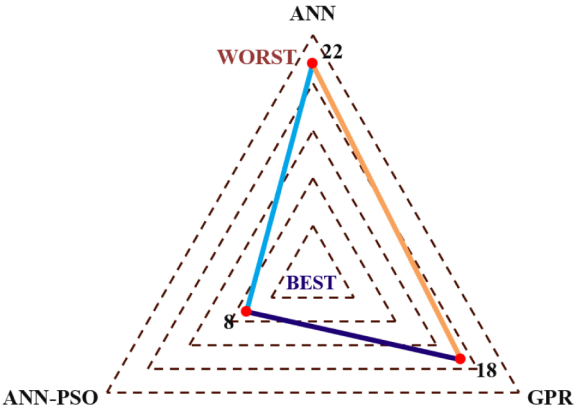
In this section, a ranking study is performed, as shown in Table 6. After computing all the statistical parameters for both the training (TR) and testing (TS) stages, the models are simultaneously ranked. The model with the best performance is assigned rank 1, while the model with the worst performance is assigned rank 3 (as three models namely ANN, GPR and ANN-PSO are used in this study). After this, the sum of all ranks is calculated to obtain the total rank and final rank, which is also determined in this study. The model with the lowest rank is considered the best performing model, while the model with the highest rank is considered the worst performing model. From Table 6, we find that the total ranks for both training and testing datasets are as follows: the ANN-PSO model has $(\text{Rank})_{\text{TR}} = 8$, $(\text{Rank})_{\text{TS}} = 8$ and a final rank = 16; the GPR model has $(\text{Rank})_{\text{TR}} = 16$, $(\text{Rank})_{\text{TS}} = 18$ and a final rank = 34; and the ANN model has $(\text{Rank})_{\text{TR}} = 24$, $(\text{Rank})_{\text{TS}} = 22$ and final rank = 46. This gives a complete calculation of the prediction power and performance of the model. Hence, based on the ranking, we conclude that the ANN-PSO model demonstrates superior predictive ability compared to the other two models. Rank analysis is also represent in the form of radar diagram which shown in Fig. 6a–c.

Table 6 Rank analysis of the proposed model

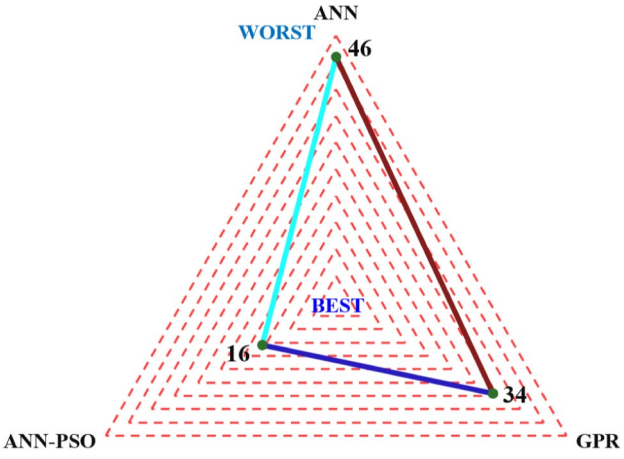
Parameters	ANN		GPR		ANN-PSO	
	Training	Testing	Training	Testing	Training	Testing
R^2	3	3	2	2	1	1
VAF	3	3	2	2	1	1
LMI	3	3	2	2	1	1
a-10 Index	3	2	2	3	1	1
RMSE	3	3	2	2	1	1
RSR	3	3	2	2	1	1
MAE	3	3	2	2	1	1
MAD	3	2	2	3	1	1
Total rank	24	22	16	18	8	8
Final rank	46		34		16	



(a)



(b)



(c)

Fig. 6 The rank analysis is represented in the form of radar diagram. **a** TR stage. **b** TS stage. **c** Final rank

6.3 Reliability Analysis

Reliability analysis aims to observe uncertainty and obtain a reliable approach to predict the FOS for an infinite slope in c-φ soil. The FOSM approach is applied to calculate the reliability index (β). The reliability index value is obtained by using Eq. (7). By comparing the reliability index of proposed models with the observed β, the model whose β is very close to the observed β is considered to have better performance than other models. Table 7 shows that the ANN-PSO model performs better than the other two models and is assigned rank 1, while the ANN model has worst performance and is assigned ranked 3. The bar graph in Fig. 7 shows a comparison of the observe β to the model's β values. Among the proposed models, the model whose β is closer to the observe β is considered the best. The ANN-PSO model's β is closer to the observe β. Hence, it is considered to be best model to predict the FOS of an infinite slope compared to the than other two models.

Table 7 Comparison of the proposed model based on β

Models	Observed β	Model's β	Rank
ANN	0.879	0.952	3
GPR		0.948	2
ANN-PSO		0.895	1

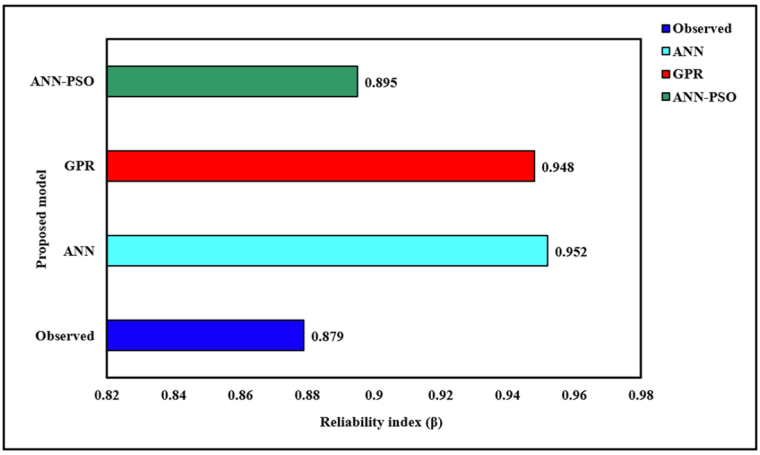


Fig. 7 Comparison of reliability index (β) between observed β and model's β

6.4 Regression Curve

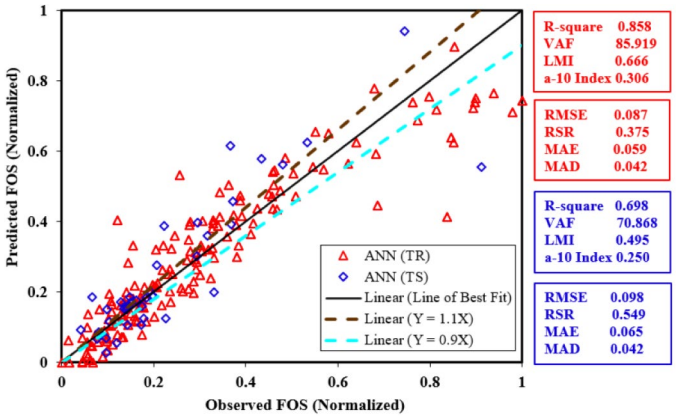
The regression curve is a graphical representation of the observed and predicted values of FOS (normalized), with the observed FOS (normalized) represented on the abscissa and the predicted FOS (normalized) on the ordinate. This curve is also known as R-curve or scatter plot. This *R*-value (coefficient of correlation), which is calculated and presented in Table 5, is derived from this curve.

Using training and testing datasets, Fig. 8a–c represents the observed FOS (normalized) and predicted FOS (normalized) for an infinite slope. From the regression curve, we can observe that all the three models overlap (observed FOS with predicted FOS) each other and follow the almost same trend. The ANN-PSO model shows slight deviation, whereas significant deviations are observed in the ANN model for both testing and training datasets. From the other measures, it is clear that ANN-PSO model has more reliable predictability than other two models.

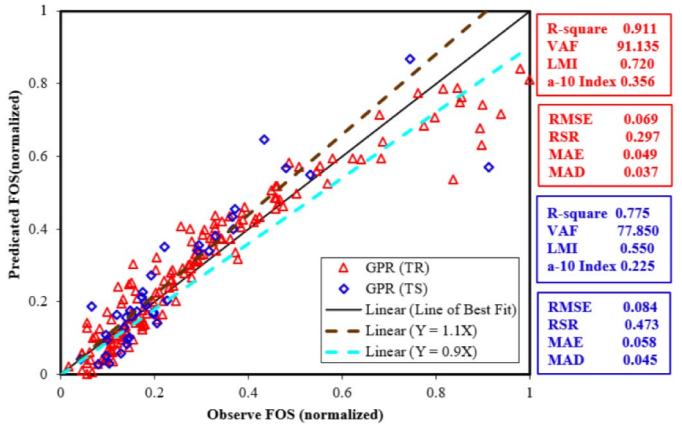
6.5 William's Plot

The William plot is a graphical representation between the standardized residuals versus leverage (*L*), which helps visualize the applicability domain. In this plot, the leverage (*L*) values are represented on the abscissa, while the standardized residuals are represented on the ordinate, as shown in Fig. 9a–c. Evaluating the applicability domain of the three different proposed models is essential for determining whether the model is reliable in prediction. By assessing the leverage (*L*) values for both testing (TS) and training (TR) datasets, the applicability domains for the three distinct proposed models were determined. These graphs show that the applicability domain is enclosed by the boundary region PQRS within a leverage threshold L^* ($L^* = 3(I+1)/w$, where I is the input parameters and w is the number of training datasets) and ± 3 standard deviations. Elements falling within the boundary region PQRS and having a leverage value $L < L^*$ are reliably predicted by the model.

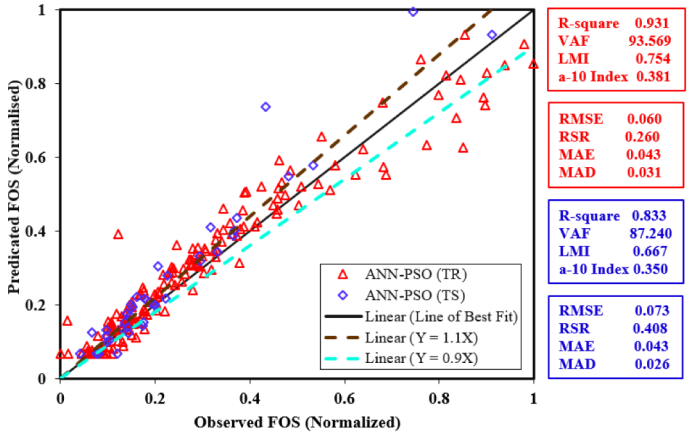
The William plot for both the training and testing datasets is used to evaluate the applicability domain of the ANN, GPR and ANN-PSO models, within the standardized residual ($\pm 3\sigma$) and a leverage threshold $L^* = 0.1125$. From Fig. 9a–c, it can be observe that each element in the training dataset has an $L < L^*$, but two elements in the testing datasets exceed the L^* threshold and five training and one testing dataset elements lie outside the boundary region ($\pm 3\sigma$) in the ANN model. In the GPR model, all training dataset elements have $L < L^*$, but two elements in the testing dataset exceed the L^* threshold and four training and one testing dataset elements have are outside the boundary region ($\pm 3\sigma$). In the ANN-PSO, training dataset elements have $L < L^*$, but three elements in the testing datasets exceed the L* threshold, and two training dataset elements



(a)



(b)



(c)

Fig. 8 Regression curve of observed and predicted FOS for both training (TR) and (TS) datasets for the models. **a** ANN. **b** GPR. **c** ANN-PSO

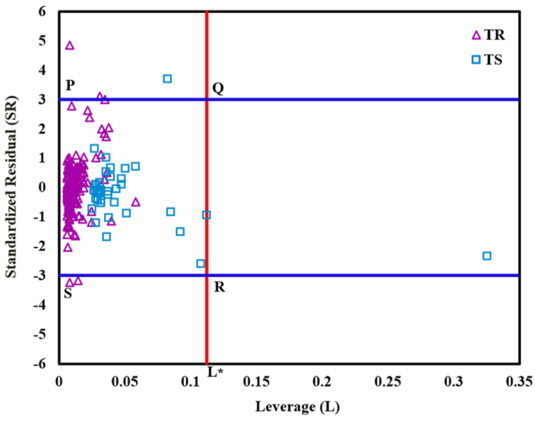


Fig. 9 Williams plot for **a** ANN, **b** GPR and **c** ANN-PSO

lie outside the boundary region ($\pm 3\sigma$). Among the three models, the ANN-PSO model shows less deviation from PQRS region. From the William plot, we conclude that ANN-PSO model predicts better than other two models.

6.6 Error Matrix

An error matrix, also known as a confusion matrix, is a table used to describe the performance of different models. It allows visualization of an algorithm's performance by comparing the predicted values of selected variables with their ideal values. The matrix calculates and provides a graphical representation of the amount of error in the predicted models, judged based on the ideal values of each performance parameter. It also provides an understanding of the highest and lowest values of error in a predictive model. Hence, the error values (E%) of a predictive model can be calculated using the two terms listed below for trend-measuring parameters (R^2, VAF, LMI and a-10 Index) and error measuring parameters (RMSE, RSR, MAE and MAD), respectively.

$$E_{\mathrm{TMP}} = \frac{I_{\mathrm{TMP}} - \left|P_{\mathrm{TMP}}\right|}{\left|P_{\mathrm{TMP}}\right|} \times 100 \tag{29}$$

$$E_{\mathrm{EMP}} = \left|(I_{\mathrm{EMP}} - \left|P_{\mathrm{EMP}}\right|)\right| \times 100 \tag{30}$$

where E_{TMP} and E_{EMP} denote the error for trend measuring parameters and error measuring parameters, respectively; I_{TMP} and I_{EMP} denote the ideal values for trend measuring and error measuring parameters, respectively; P_{TMP} and P_{EMP} denote the performance indices estimated for trend measuring and error measuring parameters, respectively. The amount of error is calculated for trend measuring and error measuring parameters using Eqs. (29 and 30), which are located in Table 8.

Figure 10 shows the error matrix for trend measuring and error measuring parameters. The quantity of error is computed for both training and testing datasets by considering trend measuring parameters, namely R^2, VAF, LMI and a-10 index and error measuring parameters, namely RMSE, RSR, MAE and MAD as shown in Fig. 10. ANN-PSO has the lowest error for both trend measuring and error measuring parameters out of all three models. The lowest error range (0–27) is shown in dark green, moderate error range (27–45) in light green, the semi-moderate error range (45–60) in pink and the highest error range in red. The ANN model performs the worst among all three models due to the highest error. Overall, we can conclude that the ANN-PSO model has better prediction for both training and testing datasets because it has the lowest error.

Table 8 Calculation of error for both trend and error measuring parameters (TR and TS dataset)

Parameters	ANN		GPR		ANN-PSO	
	Error in %		Error in %		Error in %	
	TR	TS	TR	TS	TR	TS
R^2	16.550	43.266	9.769	29.032	7.410	20.048
VAF	16.388	41.107	9.727	28.452	6.873	14.626
LMI	50.150	102.020	38.888	81.818	32.625	49.925
a-10 index	226.797	300.000	180.898	344.444	161.780	185.714
RMSE	8.700	9.800	6.900	8.400	6.000	7.300
RSR	37.500	54.900	29.700	47.300	26.000	40.800
MAE	5.900	6.500	7.900	5.800	4.360	4.300
MAD	4.220	4.240	3.700	4.500	3.100	2.600

Fig. 10 Error matrix for trend and error measuring parameters (training and testing datasets)

6.7 Comparative Analysis

Comparative analysis is a method used to examine two or more proposed models based on their performance parameters to identify their similarities and differences. Lin et al. (2021) used the SVM model to predict FOS of slopes. The highest performing model achieved an overall R^2 value of 0.864. In terms of performance, ANN slightly lagged behind SVM, while GPR performed better compared to SVM. However, the ANN-PSO model outperformed SVM. Ray and Roy (2021) used a hybrid model, PSO-ANN, which combines ANN with meta-heuristic optimization such as PSO to predict the FOS of slopes. The highest performing model achieved an overall R^2 value of 0.904. Based on the performance, ANN and GPR lag marginally behind PSO-ANN hybrid model. However, ANN-PSO performed better compared to the same hybrid model. Mahmoodzadeh et al. (2022) used GPR model to predict the FOS of slope stability. The highest performing model achieved an overall R^2 value of 0.813. Based on performance parameters, ANN, GPR and ANN-PSO performed better compared to the GPR model. Ahmad et al. (2024) used the hybrid model ANFIS-PSO to predict the FOS of soil slopes. The highest performing model achieved an R^2 value of 0.901 in training and 0.896 in testing. In terms of performance, ANN slightly lagged the ANFIS-PSO hybrid model, while GPR showed similar performance. However, the ANN-PSO model gives superior result compared to the ANFIS-PSO hybrid model. Table 9 shows the comparative analysis of the proposed models based on coefficient of determination (R^2).

Table 9 Comparative analysis of the proposed models based on coefficient of determination

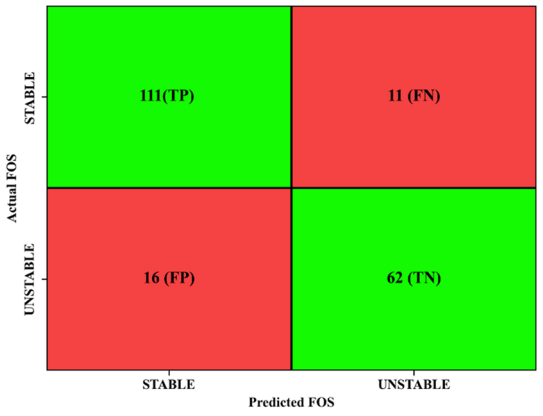
Performance parameters/ Phase	Present study			Past research			
				Lin et al. (2021)	Ray and Roy (2021)	Mahmoodzadeh et al. (2022)	Ahmad et al. (2024)
	ANN	GPR	ANN-PSO	SVM	PSO-ANN	GPR	ANFIS-PSO
R^2 (Training)	0.858	0.911	0.931	-	-	-	0.901
R^2 (Testing)	0.698	0.775	0.833	-	-	-	0.896
R^2 (Overall)	0.840	0.895	0.920	0.864	0.904	0.813	-

6.8 Confusion Matrix

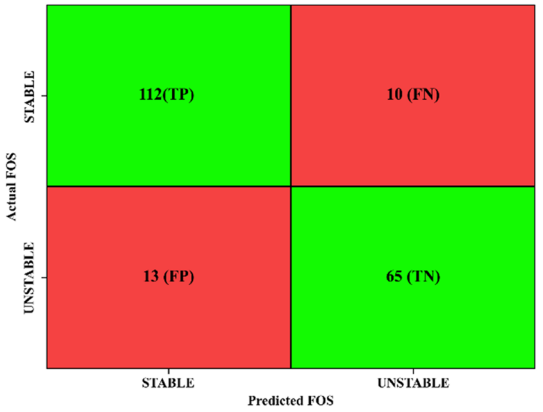
A confusion matrix, a specific type of table layout, is used to visualize the performance of machine learning algorithms. To assess the performance of predictive models, various metrics such as accuracy, F1 score, Matthew's correlation coefficient (MCC), precision and recall are used (Zhou et al. 2021, Wang et al. 2023, Liu et al. 2023 and Liu et al. 2024). It is commonly used in supervised learning and in unsupervised learning; it is referred to as a matching matrix. The rows of the matrix represent instances of a predictive class, while the columns represent instances of an actual class. In predictive analytics, a confusion matrix is a two-row, two-column table that records the counts of true positives, false negatives, false positives and true negatives. This provides a more comprehensive analysis than merely looking at the ratio of correct classification (accuracy). If the dataset is unbalanced, meaning the number of observations in different classes varies significantly, accuracy can be misleading. However, the performance matrix, which includes accuracy, F1 score, Matthew's correlation coefficient (MCC), precision and recall parameters, is a specific table that depicts a prediction algorithm's performance based on its predicted values (also known as evaluation indexes). True positive (TP), true negatives (TN), false positives (FP) and false negatives (FN) are used in classification tasks to compare the classifier's results with reliable external evaluations. Therefore, each confusion matrix offers evaluation indexes used to analyze the capabilities and performances of machine learning classifiers. Precision (also known as positive predictive value) is the ratio of relevant instances among the retrieved instances, while recall is the ratio of relevant instances that have been retrieved. Table 10 shows the mathematical expression and their ideal values of performance parameters. Figure 11 indicates the confusion matrix as classified from the model (a) ANN, (b) GPR and (c) ANN-PSO. Table 11 indicates the performance indices for all the proposed models. For assessing the best model, a ranking is assigned to the model based on its performance indicators. The model with the higher performance indicators receives a lowest rank. In the overall ranking, the ascending sequence of performance is ANN, GPR and ANN-PSO. The best performing model among all is the ANN-PSO model, as it gets the lowest rank of 5, while the ANN model is the lowest performing model with an overall rank of 15. Figure 12 indicates the radar diagram of performance indices for ANN, GPR and ANN-PSO models. From Fig. 12, we can observe that ANN-PSO performed better followed by GPR and ANN.

Table 10 Mathematical expression and their ideal values of performance parameters

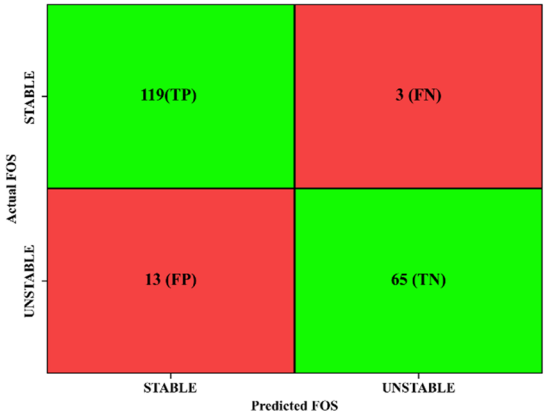
Parameters	Mathematical expression	Ideal value
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	1
F1 score	$\frac{2\times TP}{2\times TP+FP+FN}$	1
Matthew's correlation coefficient (MCC)	$\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)\times(FP+TN)\times(TP+FN)\times(FN+TN)}}$	1
Precision	$\frac{TP}{TP+FP}$	1
Recall	$\frac{TP}{TP+FN}$	1



(a)



(b)



(c)

Fig. 11 Confusion matrix for the model. **a** ANN. **b** GPR. **c** ANN-PSO

Table 11 Computed values and their rank analysis of confusion matrix

ML models	Accuracy	F1 score	MCC	Precision	Recall	Total rank
ANN	0.865	0.891	0.713	0.874	0.909	15
Rank	3	3	3	3	3	
GPR	0.885	0.906	0.756	0.896	0.918	10
Rank	2	2	2	2	2	
ANN-PSO	0.920	0.937	0.832	0.901	0.975	5
Rank	1	1	1	1	1	

6.9 Rate Analysis

In this section, both observed and predicted factors of safety are analyzed on the basis of stable (FOS > 1) or unstable slope (FOS < 1). Analysis is done on the basis of success rate (If FOS > 1) and failure rate (If FOS < 1). The success rate is the percentage of stable slopes whose FOS > 1, while the failure rate is the percentage of unstable slopes whose FOS < 1 among all slope conditions to judge the performance of the proposed models in both TR and TS phase. Figure 13 shows the success and failure rates for both TR and TS phase. It is observed from Fig. 13 that the success rate of the ANN-PSO model has nearly closet to the success rate of observe FOS in the TR phase and equal in the TS phase while predicting FOS. Hence, the best reliable model is ANN-PSO among all three models.

6.9.1 Sensitivity Analysis (SA)

Sensitivity analysis (SA) is the study of the relative importance of different input parameters on the model output (FOS). Each input (γ, c, φ, α and H) has an impact on the output (FOS), which is determined by the strength of relation (SOR) parameter. A higher SOR value indicates that the input parameter has more influence on the output. The following is an expression for strength of relation (SOR).

$$\mathrm{SOR}_{\mathrm{Z_{s,r}}} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{j}} \mathrm{A_{r,i}} \mathrm{Z_{s,i}}}{\sqrt{\sum_{\mathrm{i=1}}^{\mathrm{j}} \left(\mathrm{A_{r,i}}\right)^2 \left(\mathrm{Z_{s,i}}\right)^2}} \tag{31}$$

where $A_{\mathrm{r,i}}$ implies the ith value of rth independent variable; j and r are the whole observations and total input parameters, respectively; $Z_{\mathrm{s,i}}$ denotes the ith value of sth dependent variable. $\mathrm{SOR}_{\mathrm{Z_{s,r}}}$ is the strength of relation of rth independent variable to sth dependent variable and s is the total dependent variables. In this study, $r = 5$, $s = 1$ and $j = 200$. Figure 14a–d indicates the strength of relation between different input parameters. As shown in Fig. 14a–d, the parameter c has the most influence on the FOS computation in all scenarios (actual case as well as all three proposed models) because it has the highest SOR value out of all five input parameters followed by φ, α, γ and H. Finally, it is clear that the ANN-PSO almost perfectly mirrored the actual output in predicting the FOS.

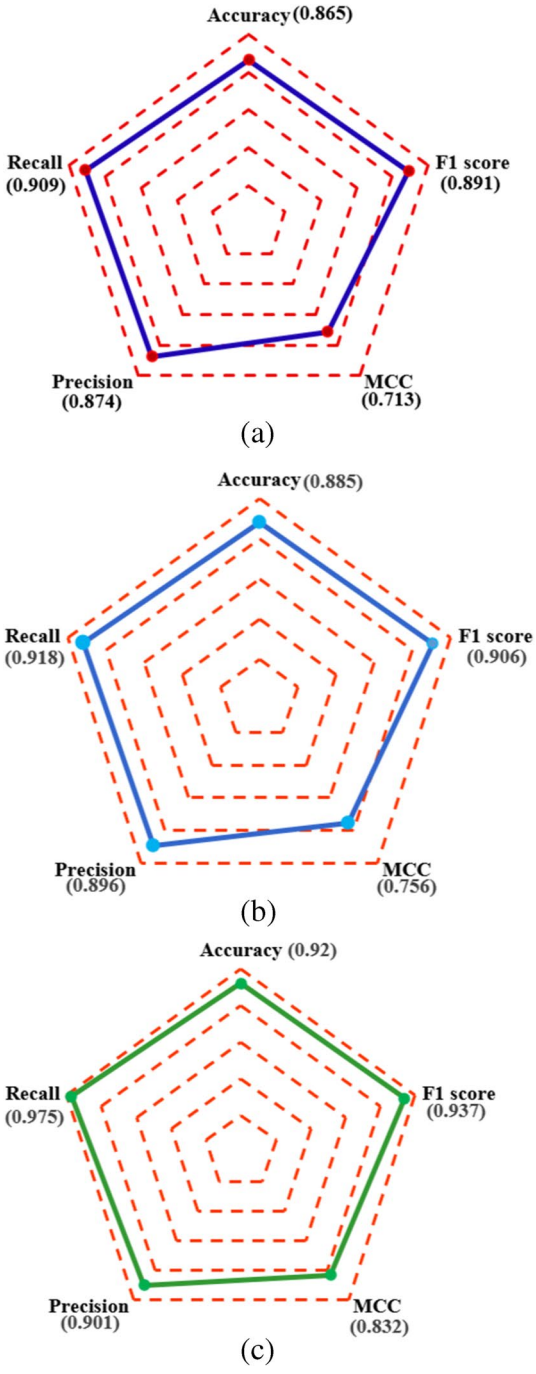
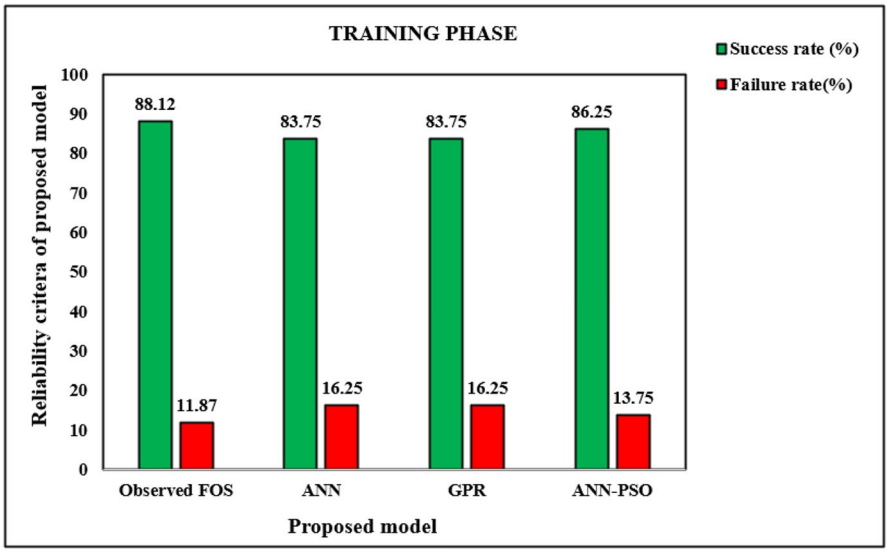
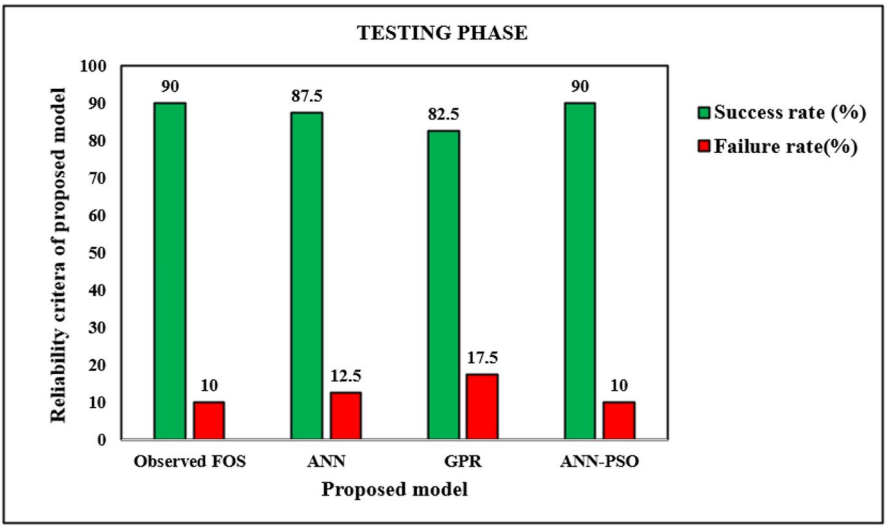


Fig. 12 Radar diagram of performance indices for **a** ANN, **b** GPR and **c** ANN-PSO



(a)



(b)

Fig.13 Reliability criteria of proposed model. **a** TR and **b** TS phase

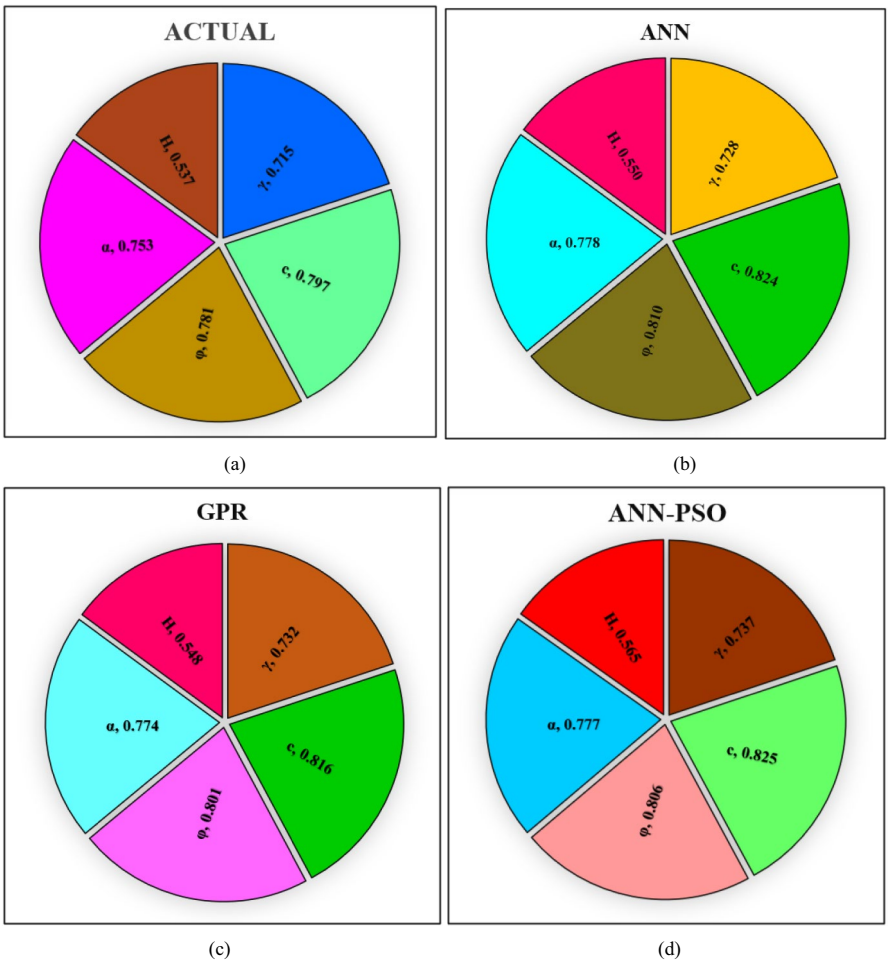


Fig. 14 The relative significance of input parameters for **a** ACTUAL, **b** ANN, **c** GPR and **d** ANN-PSO model

7 Conclusion

This study conducts a reliability analysis of infinite slope stability in c-φ soil, considering five input parameters like γ, c, φ, α and H. we have used two types of soft computing techniques namely artificial neural network (ANN) and Gaussian process regression (GPR) to predict the factor of safety (FOS) individually and make one hybrid model (ANN-PSO) by using metaheuristic optimisation technique namely particle swarm optimization (PSO).

The conclusion can be outlined as follows:

(i) The ANN-PSO hybrid model is better than the other two models to predict the FOS. It achieves superior predictive power with higher values (R^2 = 0.931, VAF = 93.569, LMI = 0.754 and a-10 index = 0.381) and lower values (RMSE = 0.060,

RSR = 0.260, MAE = 0.436 and MAD = 0.031) in the training stage, whereas the slightly decrease in the testing stage (R^2 = 0.833, VAF = 87.240, LMI = 0.667 and a-10 index = 0.350, RMSE = 0.073, RSR = 0.408, MAE = 0.043 and MAD = 0.026).

(ii) On the basis of reliability index (β), the ANN-PSO is superior to the other two models. The FOSM method is used to calculate the reliability index (β). The model's β is closest to the observed β which perform better than the other models. Other criteria, such as rank analysis, R-curve, William's plot, error matrix, confusion matrix and rate analysis, also indicate that ANN-PSO is superior.

(iii) The strength of relation (SOR) value is calculated to analyze the influence of input parameters on FOS for an infinite slope in c-φ soil. Among the five input variables, cohesion (c) has the most significant impact, followed by the angle of shearing resistance (φ), slope angle (α), unit weight of soil (γ) and slope height (H).

(iv) Machine learning techniques offer advantages such as higher accuracy and lowest error, fast decision making, more reliability and time saving. The drawbacks include high cost, inability to think beyond set limits.

Author Contribution Rashid Mustafa: conceptualization, data curation, investigation, methodology, software, supervision, validation; Akash Kumar: writing—original draft; Sonu Kumar: writing—review and editing; Navin Kumar Sah: literature review; Abhishek Kumar: formal analysis.

Availability of Data and Material Data presented in the paper are available with authors.

Declarations

Ethics Approval and Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.

References

Ahmad, F., Samui, P., Mishra, S.S.: Probabilistic slope stability analysis on a heavy-duty freight corridor using a soft computing technique. Transp. Infrastruct. Geotech. (2023). https://doi.org/10.1007/s40515-023-00365-4

Ahmad, F., Samui, P., Mishra, S.S.: Machine learning-enhanced Monte Carlo and subset simulations for advanced risk assessment in transportation infrastructure. J. Mt. Sci. **21**(2), 690–717 (2024a). https://doi.org/10.1007/s11629-023-8388-8

Ahmad, F., Samui, P., Mishra, S.S.: Probabilistic slope stability analysis using subset simulation enhanced by ensemble machine learning techniques Model. Earth Syst. Environ. **10**, 2133–2158 (2024b). https://doi.org/10.1007/s40808-023-01882-4

Al-karni, A.A., Al-shamrani, M.A.: Study of the effect of soil anisotropy on slope stability using method of slices. Comput. Geotech. **26**(2), 83–103 (2021). https://doi.org/10.1016/S0266-352X(99)00046-4

Anand, M.A.T., Anandakumar, S., Pare, A., Chandrasekar, V., Venkatachalapathy, N.: Modelling of process parameters to predict the efficiency of shallots stem cutting machine using multiple regression and artificial neural network. Journal of Food Process Engineering 45(6) (2021) https://doi.org/10.1111/jfpe.13944

Cho, S.E.: Probabilistic stability analysis of slope using the ANN-based response surface. Comput. Geotech. **36**(5), 787–797 (2009). https://doi.org/10.1016/j.compgeo.2009.01.003

Das, B.M.: Principles of geotechnical engineering. PWS-KENT Publishing Co., Ltd., London (1985)

Deris, A.M., Solemon, B., Omar, R.C.: A Comparative study of supervised machine learning approaches for slope failure production. E3S Web Conf. 325, 01001 (2021) https://doi.org/10.1051/e3sconf/202132501001

Fattahi, H., Ilghani, N.Z.: Slope stability analysis using Bayesion Markov chain Monte Carlo method. Geotechnical and Geological Engineering **38**(3), 1–10 (2019). https://doi.org/10.1007/s10706-019-01172-w

Griffiths, D.V., Fenton, G.A.: Probabilistic slope stability analysis by finite elements. Geo. and Geo. Env. Eng. **130**(5), 507–518 (2004). https://doi.org/10.1061/(ASCE)1090-0241(2004)130:5(507)

Griffiths, D.V., Huang, J., Fenton, G.A.: Probabilistic Infinite Slope Analysis. Com. Geo. **38**(4), 577–584 (2011). https://doi.org/10.1016/j.compgeo.2011.03.006

Gupta, A., Aggarwal, Y., Aggarwal, P.: Deep neural network and ANN ensemble for slope stability prediction. Archives of Materials Science and Engineering **116**(1), 14–27 (2022). https://doi.org/10.5604/01.3001.0016.0975

Johari, A., Nejad, A.H., Mousavi, S.: Probabilistic model of unsaturated slope stability considering the uncertainties of soil-water characteristics curve. Scientia Iranica **25**(4), 2039–2050 (2018). https://doi.org/10.24200/sci.2017.4202

Johari, A., Javadi, A.A.: Reliability assessment of infinite slope stability using the jointly distributed random variables method. Scientia Iranica **19**(3), 423–429 (2012). https://doi.org/10.1016/j.scient.2012.04.006

Kang, F., Han, S., Salgado, R., Li, J.: System probabilistic stability analysis of soil slopes Gaussian process regression with Latin hypercube sampling. Computers and Geotecnics **63**, 13–25 (2015). https://doi.org/10.1016/j.compgeo.2014.08.010

Kang, F., Li, J.S., Li, J.J.: System reliability analysis of slopes using least support vector machines with particle swarm optimization. Neurocomputing **209**, 46–56 (2016). https://doi.org/10.1016/j.neucom.2015.11.122

Kennedy, J., Eberhart, R.: Particle swarm optimization. Proceeding of the ICCN'95-International Conference on Neural Networks, Perth, WA, Australia 4, 1942–1948 (1995) https://doi.org/10.1109/ICNN.1995.488968

Khajehzadeh, M., Taha, M.R., Keawsawasvong, S., Mirzaei, H., Jebeli, M.: An effective artificial intelligence approach for slope stability evaluation. IEEE Access **10**, 5660–5671 (2022). https://doi.org/10.1109/ACCESS.2022.3141432

Khajehzadeh, M., Keawsawasvong, S.: Predicting slope safety using and optimised machine learning model. Heliyon 9(12) (2023) https://doi.org/10.1016/j.heliyon.2023.e23012

Kumar, R., Wipulanusat, W., Kumar, M., Keawsawasvong, S., Samui, P.: Optimized neural network-based state-of-the-art soft computing models for the bearing capacity of strip footings subjected to inclined loading. Intelligent System with Applications **21**, 2667–3053 (2024). https://doi.org/10.1016/j.iswa.2023.200314

Kumar, R., Samui, P., Kumari, S.: Reliability analysis of infinite slope using metamodels. Geotech Geol Eng 35, 1221–1230 (2017) https://link.springer.com/article/https://doi.org/10.1007/s10706-017-0160-9

Lin, S., Zheng, H., Han, C., Han, B., Li, W.: Evaluation and prediction of slope stability using machine learning approaches. Front. Struct. Civ. Eng. **15**, 821–833 (2021). https://doi.org/10.1007/s11709-021-0742-8

Liu, S., Wang, L., Zhang, W., Sun, W., Fu, J., Xiao, T., Dai, Z.: A physics-informed data-driven model for landslide susceptibility assessment in the Three Gorges Reservoir area. Geosci. Front. **14**(5), 101621 (2023). https://doi.org/10.1016/j.gsf.2023.101621

Liu, S., Wang, L., Zhang, W., Sun, W., Wang, Y., Liu, J.: Physics-informed optimization for a data-driven approach in landslide susceptibility evaluation. Journal of Rock Mechanics and Geotechnical Engineering (2024). https://doi.org/10.1016/j.jrmge.2023.11.039

Liu, Z., Wu, D., Sheng, D., Fatahi, B., Khabbaz, H.: Machine learning aided stochastic slope stability analysis. UNCECOMP Proceedings 75–81 (2021) https://doi.org/10.7712/120221.8023.19068

Mahmoodzadeh, A., Mohammadi, M., Ibrahim, H., Rashid, T.A., Aldalwie, A.H.M., Ali, H.F.H., Daraei, A.: Tunnel geomechanical parameters prediction using Gaussian process regression. Mech. Learn. Appl. **3**, 100020 (2021). https://doi.org/10.1016/j.mlwa.2021.100020

Mahmoodzadeh, A., Mohammadi, M., Ali, H.F.H., Ibrahim, H.H., Abdulhamid, S.N., Nejati, H.R.: Prediction of safety factors for slope stability: comparison of machine learning techniques. Nat. Hazards **111**, 1771–1799 (2022). https://doi.org/10.1007/s11069-021-05115-8

Malkawi, A.I.H., Hassan, W.F., Abdulla, F.A.: Uncertainty and reliability analysis applied to slope stability. Struct. Saf. **22**(1), 161–187 (2020). https://doi.org/10.1016/S0167-4730(00)00006-0

Mustafa, R., Samui, P., Kumari, S.: Reliability analysis of gravity retaining wall using hybrid ANFIS. Infrastructures **7**(9), 121 (2022). https://doi.org/10.3390/infrastructures7090121

Mustafa, R., Kumari, K., Kumari, S., Kumar, G., Singh, P.: Probabilistic analysis of thermal conductivity of soil. Arab. J. Geosci. **17**, 22 (2024). https://doi.org/10.1007/s12517-023-11831-1

Nanehkaran, Y.A., Pusatli, T., Chengyong, J., Chen, J., Cemiloglu, A., Azarafza, M., Derakhshani, R.: Application of machine learning techniques for the estimation of the safety factor in slope stability analysis. Water **14**(22), 3743 (2022). https://doi.org/10.3390/w14223743

Nanehkaran, Y.A., Licai, Z., Chengyong, J., Chen, J., Anwar, S., Azarafza, M., Derakhshani, R.: Comparative analysis for slope stability by using machine learning methods. Appl. Sci. **13**, 1555 (2023). https://doi.org/10.3390/app13031555

Feng, X., Li, S., Yuan. C., Zeng, P., Sun, Y.: Prediction of slope stability using naive Bayes classifier. KSCE J. Civ. Eng. **22**, 941–950 (2018). https://doi.org/10.1007/s12205-018-1337-3

Ray, R., Roy, L.B.: Reliability analysis of soil slope stability using ANN, ANFIS, PSO-ANN soft computing techniques. Nat Volatiles Essent Oils **8**(6), 3478–3491 (2021). (https://www.nveo.org/index.php/journal/article/view/4100/3368)

Sabri, M., Ahmad, F., Samui, P.: Slope stability analysis of Heavy-haul freight corridor using novel machine learning approach. Model. Earth Syst. Environ. **10**, 201–219 (2024). https://doi.org/10.1007/s40808-023-01774-7

Sonek, J., Larochelle, H., Adams, R.P.: Practical Bayesian optimization of machine learning algorithms. Adv Neural Inf Process Syst **25**, 2951–2959 (2012). https://doi.org/10.48550/arXiv.1206.2944

Wang, C., Wu, X., Kozlowski, T.: Gaussian process–based inverse uncertainty quantification for trace physical model parameters using steady-state PSBT benchmark. Nucl. Sci. Eng. **193**, 100–114 (2019). https://doi.org/10.1080/00295639.2018.1499279

Wang, Y., Wang, L., Liu, S., Liu, P., Zhu, Z., Zhang, W.: A comparative study of regional landslide susceptibility mapping with multiple machine learning models. Geol. J. (2023). https://doi.org/10.1002/gj.4902

Wang, M., He, Z., Zhao, H.: Dimensional reduction-based moment model for probabilistic slope stability analysis. Appl Sci **12**(9), 4511 (2022). https://doi.org/10.3390/app12094511

Xu, H., He, X., Shan, F., Niu, G., Sheng, D.: Machine learning in the stochastic analysis of slope stability: a state-of-the-art review. Modelling **4**(4), 426–453 (2023). https://doi.org/10.3390/modelling4040025

Yang, Y., Zhou, W., Jiskani, I., Lu, X., Wang, Z., Luan, B.: Slope stability prediction method based on intelligent optimization and machine learning algorithms. Sustainability **15**(2), 1169 (2023). https://doi.org/10.3390/su15021169

Yousuf, S.M., Khan, M.A., Ibrahim, S.M., Sharma, A.K., Ahmad, F.: Response of rectangular footing resting on reinforced silty sand treated with lime using experimental and computational approach. Geomechanics and Geoengineering **19**(2), 139–161 (2023). https://doi.org/10.1080/17486025.2023.2194871

Zhao, H.: Slope reliability analysis using a support vector machine. Comput. Geotech. **35**(3), 459–467 (2008). https://doi.org/10.1016/j.compgeo.2007.08.002

Zhou, X., Wen, H., Zhang, Y., Xu, J., Zhang, W.: Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. Geosci. Front. **12**(5), 101211 (2021). https://doi.org/10.1016/j.gsf.2021.101211

Zhu, B., Hiraishi, T., Pei, H., Yang, Q.: Efficient reliability analysis of slopes integrating the random field method and a Gaussian process regression-based surrogate model. Int. J. Numer. Anal. Methods Geomech **45**, 478–501 (2021). https://doi.org/10.1002/nag.3169

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Rashid Mustafa[1] · Akash Kumar[1] · Sonu Kumar[1] · Navin Kumar Sah[1] · Abhishek Kumar[1]

✉ Rashid Mustafa
talktorashid@keck.ac.in

Akash Kumar
akashkumarjeearvind@gmail.com

Sonu Kumar
853204sonukumar@gmail.com

Navin Kumar Sah
sahnavinkumar3@gmail.com

Abhishek Kumar
abhishekkrnov@gmail.com

[1] Department of Civil Engineering, Katihar Engineering College Katihar, Katihar, Bihar 854109, India