Engineered Science
DOI: https://dx.doi.org/10.30919/es1083

Machine Learning Approaches for the Prediction of the Seismic Stability of Unsupported Rectangular Excavation

Divesh Ranjan Kumar,[1] Warit Wipulanusat,[1,*] Jirapon Sunkpho,[2] Suraparb Keawsawasvong,[3] Wittaya Jitchaijaroen[3] and Pijush Samui[4]

Abstract

The seismic stability of unsupported rectangular excavations poses significant challenges in geotechnical engineering, especially underground structures. This study addresses the need for accurate prediction methods to assess the vulnerability of such excavations under seismic loading conditions. This study addresses seismic stability in excavations and underground structures using a random forest (RF) model with three distinct optimization algorithms: the whale optimization algorithm (WOA), dragonfly optimization algorithm (DOA), and sparrow search optimization algorithm (SSOA). The method focuses on four dimensionless factors, with the seismic stability number *(N)* serving as the output. The results obtained from the proposed data-driven models indicate that the RF-DOA model has the best predictive performance and highest accuracy. In addition, scatter plots, error plots, line plots, and Taylor diagrams were generated to compare the performances of all the proposed models. Shapley analysis showed that the soil friction angle (ϕ) is the most significant influencing factor, and the horizontal seismic coefficient (k_h) is the least significant influencing factor. This research advances seismic stability prediction for underground structures, providing models for designing earthquake-resistant excavations. The RF-DOA hybrid model is highlighted for its practicality and efficiency in predicting seismic stability, proving essential for geotechnical engineering applications.

Keywords: Stability; Unsupported excavation; Pseudo-static seismic; Random forest.
Received: 24 November 2023; Revised: 04 January 2024; Accepted: 10 January 2024.
Article type: Research article.

1. Introduction

Unsupported excavation, defined as the process of excavating without retaining wall systems, is a cost-effective technique commonly employed in various construction projects, including piers, foundations, oil and water tanks, pipelines, shallow tunnels, and underpasses. Unsupported excavation substantially reduces costs by removing the need for temporary walls and bracing systems, making it a popular choice in underground construction. However, in areas prone to seismic activity, the lack of adequate support in these excavations can result in their collapse during earthquakes. Therefore, evaluating seismic forces is essential for determining the stability of an excavation. A straightforward approach used to determine the seismic stability of underground structures is the pseudo-static approach. This approach considers the simplified horizontal and vertical seismic coefficients as seismic factors by defining them as ratios of gravity acceleration.

Early works by Taylor[1] and Janbu[2] employed the limit equilibrium method to predict planar slope stability. The plastic theorem and the limit analysis approach were employed by Chen[3] and Pastor *et al*.[4] to derive the lower bound (LB) and upper bound (UB) stability solutions for planar unsupported excavations. Over the past three decades, finite element limit analysis (FELA), detailed further in Sloan,[5] has been utilized in various studies, including those by Yu *et al*.[6] and Martin,[7] to obtain LB and UB stability solutions for planar unsupported excavations. However, those

[1] *Research Unit in Data Science and Digital Transformation, Department of Civil Engineering, Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani 12120, Thailand.*
[2] *College of Innovation, Thammasat University, Bangkok 10200, Thailand.*
[3] *Research Unit in Sciences and Innovative Technologies for Civil Engineering Infrastructures, Department of Civil Engineering, Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Pathumthani 12120, Thailand.*
[4] *Department of Civil Engineering, National Institute of Technology Patna 800005, India.*
**Email:* wwarit@engr.tu.ac.th (W. Wipulanusat)

previous studies involved excavations under plane strain conditions. The stability solutions for cylindrical excavations under axisymmetric conditions were derived by Griffiths and Koutsabeloulis,[8] Britto and Kusakabe,[9,10] Bottero *et al*.,[11] Pastor and Turgeman,[12] Khatri and Kumar,[13] Kumar and Chakraborty,[14] and Kumar *et al*.[15] using several numerical and analytical methods.

By employing the 2D FELA, Keawsawasvong and Ukritchon,[16] Ukritchon and Keawsawasvong,[17] Yodsomjai *et al*.,[18–20] and Lai *et al*.[21] provided the static stability solutions of unsupported conical excavations under axisymmetric conditions. Recently, 3D FELA was carried out by Ukritchon *et al*.[22] and Lai *et al*.[23] to derive undrained stability solutions for unsupported rectangular excavations in purely cohesive soils. In addition, for cohesive-frictional soils considering both soil cohesion and friction angle, the 3D FELA was used by Petchkaew *et al*.[24,25] to consider the seismic stability solutions for 3D unsupported excavations by addressing the influence of the horizontal seismic coefficient. Several other researchers have also performed numerical analysis of deep excavation and seismic analysis using finite element analysis.[26–35]

Artificial intelligence (AI) has gained popularity in recent years, leading to the widespread adoption of various machine learning (ML) techniques in the science and engineering domains.[36] A range of ML algorithms, from single to hybrid models, are now commonly used to address complex and nonlinear problems. The random forest (RF) algorithm and other soft computing techniques are widely used in several engineering practices and have yielded promising results.[37–45] Kumar *et al*.[46] used several machine learning methods, including the RF model, to predict the electrical conductivity (EC) of surface water. Markuna *et al*.[47] proposed the RF model and three other machine learning techniques for long-term rainfall prediction. Elbeltagi *et al*.[48] performed a case study in Jaisalmer, India, to predict drought indicators using several data driven ML models, including the RF model. Elbeltagi *et al*.[49] proposed the RF, random tree and Gaussian regression models to predict meteorological drought and precipitation indices. Kumar *et al*. proposed the hybrid extreme learning machine (ELM) and least-square support vector machine (LSSVM) algorithm to predict the bearing capacity of ring foundations.[50] Moreover, the application of various hybrid models has spurred numerous research studies. For example, Liu *et al*.[51] proposed a hybrid model of the whale optimization algorithm and random forest (WOA-RF) model to predict the failure risk level of a backfill pipeline in mines in Sichuan, China. Chen *et al*.[52] proposed a hybrid model of the dragonfly optimization algorithm (DOA) and RF model to predict the thickness of the excavation damage zone. In another study, Chen *et al*.[53] introduced random forest regression (RFR) and an improved sparrow search optimization algorithm to predict the heavy metal contents in soil and found that their RF-SSOA algorithm efficiently predicted these contents. Moreover, these hybrid machine learning models have been frequently employed to address challenging and non-linear issues in civil engineering. The significance of machine learning is evident in the engineering achievements realized through the application of the aforementioned ML algorithms. In this study, to construct hybrid machine learning models, the previous results of Petchkaew *et al*.[25] were employed to estimate the seismic stability number.

Furthermore, an extensive literature survey indicates no prior utilization of hybrid machine learning models for evaluating seismic stability numbers (N) for unsupported rectangular excavations in cohesive-frictional soils under pseudo-static seismic body forces. Thus, considering the above points in this study, a hybrid intelligence approach involving a RF algorithm and three optimization algorithms, namely, the whale optimization algorithm (WOA), dragonfly optimization algorithm (DOA), and sparrow search optimization algorithm (SSOA), was utilized for the prediction of the seismic stability of an unsupported rectangular excavation. Accurate predictions of the seismic stability number enable engineers to design cost-effective solutions by avoiding over-design or unnecessary safety measures. This can lead to significant cost savings in construction projects without compromising safety. On the other hand, seismic stability predictions contribute to ongoing research and development in the field of geotechnical engineering. This helps improve the understanding of soil-structure interaction and leads to the development of more advanced and effective design methods. In addition, the performances of the proposed RF-WOA, RF-DOA, and RF-SSOA, were compared by evaluating several performance metrics. Moreover, scatter plots, error plots, line plots, and Taylor diagrams are also shown to assess the performance of the proposed models. Finally, Shapley analysis is performed to evaluate the weight and impact of each input variable on the seismic stability index (N).

2. Data collection

Figure 1 illustrates the geometry of the 3D unsupported excavation created in Optum G3.[43] The dimensions of the excavation were represented by the variables H for depth, B for width, and L for length. To cohesive-frictional soil, key variables were assigned to describe its characteristics. These included the soil cohesion denoted as 'c', the soil friction angle represented by 'ϕ' and the unit weight indicated as 'γ'. In this study, the pseudo-static approach, a simplified, commonly used technique in earthquake engineering, is employed to create the seismic forces acting on excavations. This approach simplifies the dynamic ground motion induced by earthquakes into a static force that affects the excavation. Within this approach, it is assumed that both the footing and the surrounding soil experience horizontal seismic acceleration. The horizontal force is characterized by the horizontal seismic coefficient, denoted as k_h. By employing this pseudo-static approach, the excavation is subjected to the horizontal seismic body force denoted by $k_h\gamma$. It is worth noting that this study

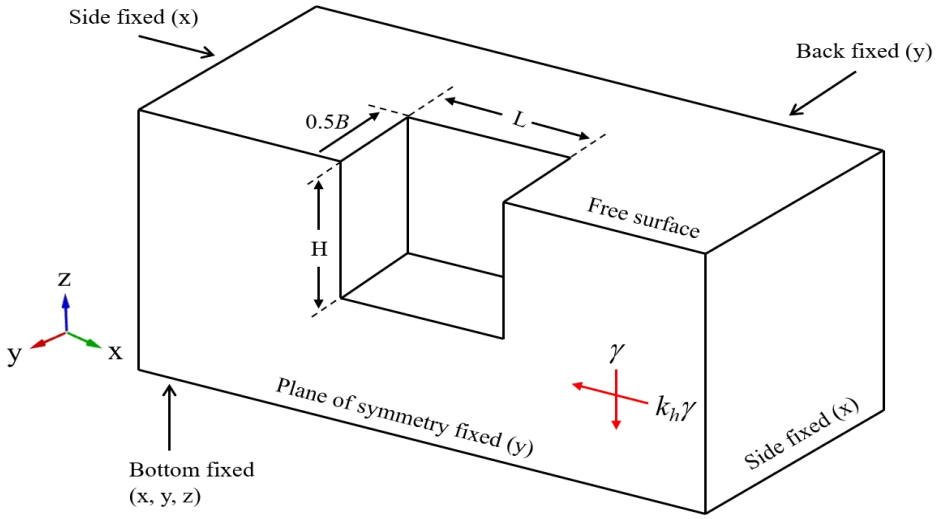

Fig. 1 Problem definition of unsupported rectangular excavation under pseudo-static seismic body force.

does not consider the vertical seismic coefficient (k_v), as it has a negligible influence on the stability of structures and underground constructions.[54–62] Other details regarding the numerical model in Optum G3 can be found in Petchkaew *et al.*[25]

Based on six input parameters (B, L, H, c, ϕ, and k_h) and the use of dimensional analysis,[63] these input parameters are combined into four dimensionless parameters that define a stability number, expressed as follows:

$$N = \frac{\gamma H}{c} = f\left(\phi, k_h, \frac{B}{L}, \frac{H}{B}\right) \quad (1)$$

where N is the seismic stability number, representing the sole output.
ϕ is the soil friction angle;
k_h is the horizontal seismic coefficient;
B/L is the rectangular excavation's aspect ratio;
H/B is the excavated depth ratio.
The selected values of these input parameters are shown in Table 1.

According to Petchkaew *et al.*,[25] the FELA technique[5] was used to determine the seismic stability solutions. To increase the precision and dependability of the results, the technique of auto-mesh adaptivity, as explained by Ciria *et al.*[64] and Ali,[65] was utilized to increase the number of elements in areas with significant dissipation of shear forces. Additionally, automatic iterations were executed to further refine the analysis. In this particular study, the initial element count was set at 5,000, while the final element count was increased to 10,000 through three adaptive iterations. The Optum G3 software[66] employed in the study allowed for the activation of the described adaptive technique, which enabled the visualization of the failure mechanism of an unsupported rectangular excavation. This was accomplished by presenting the ultimate adaptive meshes and the progressive shear strain contour, as illustrated in Figs. 2 and 3, respectively. Significantly, the training data utilized in this study originated

Table 1. List of parametric values used in this study.

Parameters	Selected values
k_h	0, 0.1, 0.2, 0.3
H/B	0.5, 1, 2, 3, 4
B/L	1/8, 1/4, 1/2, 1/3, 1
ϕ	10, 20, 30, 40°

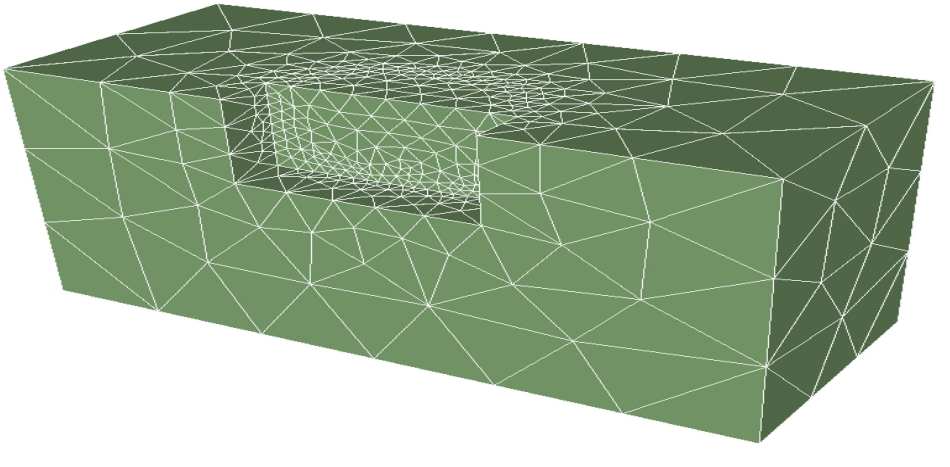

Fig. 2 Example of typical adaptive meshes.

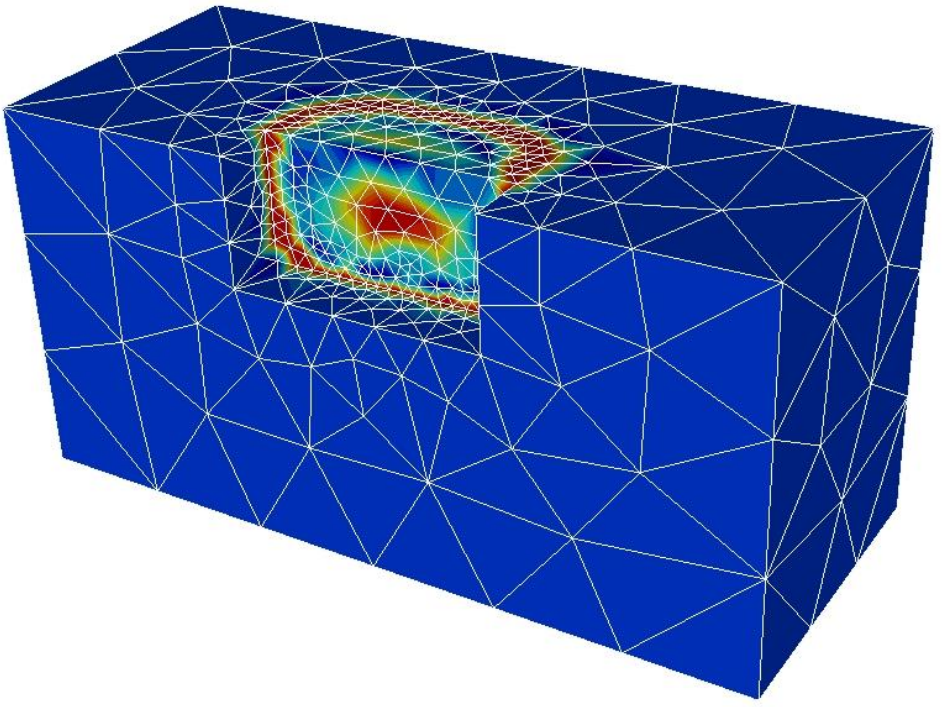

Fig. 3 Example of incremental shear strain contour.

exclusively from the research conducted by Petchkaew *et al*.,[25] which served as the basis for developing machine learning approaches. For more comprehensive information regarding the FELA simulation of this problem, readers can refer to Petchkaew *et al*.[25]

2.1 Data preprocessing

In this study, the proposed models were constructed using 400 three-dimensional finite element limit analysis solutions. To train the machine learning models, the entire dataset was normalized to the range [0,1]. Subsequently, it was randomly divided into two subsets: training and testing. The training subset comprises 80% of the total dataset (*i.e.*, 320 data points) and is used for the hybrid RF models. The testing subset, consisting of the remaining 20% of the data, is employed for validating the constructed models. After the development of the models, the effectiveness of the models was assessed using a variety of performance indicators, such as the coefficient of determination (R^2), mean absolute error (*MAE*), root mean square error (*RMSE*), performance index (*PI*), variance account factor (*VAF*), and weighted mean absolute percentage error (WMAPE). The mathematical expressions of the performance metrics and their ideal values are presented in Table 2.

2.2 Statistical characteristics

Table 3 represents the statistical description of the 400 FELA-generated solutions. The dataset comprises four dimensionless input parameters, namely, the soil friction angle (ϕ), horizontal seismic coefficient (k_h), rectangular excavation's aspect ratio (B/L), and excavated depth ratio (H/B), along with one output parameter, the seismic stability number (N). According to the statistical descriptive results in Table 3, the parameter k_h varies from 0 to 0.3. The other parameters ϕ, B/L, H/B, and N range from 10 to 40, 0.125 to 1, 0.5 to 4, and 3.93 to 36.53, respectively. Additional statistical descriptions of the dataset can also be found in Table 3.

Table 2. Equations for the performance metrics and ideal values.

Performance metrics and acronyms	Equation used	Ideal Value
Coefficient of Determination (R^2)	$R^2 = \frac{\sum_{i=1}^{n}(A_i - A_{avg})^2 - \sum_{i=1}^{n}(A_i - P_i)^2}{\sum_{i=1}^{n}(A_i - A_{avg})^2}$	1
Mean Absolute Error (*MAE*)	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert(P_i - A_i)\rvert$	0
Root Mean Square Error (*RMSE*)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A_i - P_i)^2}$	0
Performance Index (*PI*)	$PI = adj.R^2 + (0.01 \times VAF) - RMSE$	2
Variance Account Factor (*VAF*)	$VAF = \left(1 - \frac{var(A_i - P_i)}{var(Ad_i)}\right) \times 100$	100
Weighted Mean Absolute Percentage Error (WMAPE)	$WMAPE = \frac{\sum_{i=1}^{n}\left\lvert\frac{A_i - P_i}{A_i}\right\rvert \times A_i}{\sum_{i=1}^{n} A_i}$	0

Table 3. Descriptive details of the generated dataset.

Index	k_h	ϕ	B/L	H/B	N
Max.	0.3	40	1	4	36.53
Min.	0	10	0.125	0.5	3.93
Mean	0.15	25.00	0.51	2.10	9.86
Stnd. Dev.	0.11	11.19	0.31	1.28	5.11
Skewness	0.00	0.00	0.33	0.21	1.91
Kurtosis	-1.36	-1.36	-1.16	-1.39	4.44
Counts	400	400	400	400	400

2.3 Multicollinearity analysis of the input features

Multicollinearity is a common issue in regression analysis, particularly when working with multiple independent variables. To assess how these variables correlate with each other, multicollinearity analysis was performed. Highly correlated features can lead to several problems, including unstable coefficient estimates, challenges in interpreting the model, and reduced predictive power. Therefore, identifying and addressing multicollinearity in the dataset is crucial. In this study, the Pearson correlation coefficient and the variance inflation factor (VIF) were used for multicollinearity analysis. The VIF results are summarized in Table 4. The VIF quantifies how much the variance of an estimated regression coefficient is inflated due to multicollinearity, and its calculation is presented in Eq. (2).

$$VIF = \frac{1}{1-R_i^2} \quad (2)$$

where the R_i^2 term represents the coefficient of determination of a regression model where the i-th features are taken as the output variable against all other independent features. A VIF greater than 1 indicates multicollinearity. A commonly used rule of thumb is to consider VIF values above 5 or 10 as problematic. In this analysis, the VIF values for the soil friction angle (ϕ), horizontal seismic coefficient (k_h), rectangular excavation's aspect ratio (B/L), and excavated depth ratio (H/B) were all found to be 1. This indicates the absence of multicollinearity among the selected input parameters. Therefore, all the input parameters should be included in the construction of regression models.

2.4 Hypothesis testing

Hypothesis testing is a statistical approach employed for identifying a population using a sample dataset. This study used parametric tests such as the Analysis of Variance (ANOVA) and the Z-test to evaluate the hypotheses. The study's hypotheses are outlined in the following statements:

- The soil friction angle (ϕ), horizontal seismic coefficient

Table 4. Results of multicollinearity analysis.

Parameters	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	R^2	VIF
Intercept	-3.43	0.46	-7.45	0.00	-4.33	-2.52	-	-
k_h	-5.24	1.13	-4.62	0.00	-7.47	-3.01	0.00	1
ϕ	0.24	0.01	21.22	0.00	0.22	0.26	0.00	1
B/L	7.22	0.41	17.67	0.00	6.42	8.03	0.00	1
H/B	2.09	0.10	21.09	0.00	1.89	2.28	0.00	1

Table 5. ANOVA results for the seismic stability number (N).

Source of Variation	SS	df	MS	F	P-value	F crit
ANOVA for k_h and N						
Between Groups	18849.19	1	18849.19	1440.8	0	3.853
Within Groups	10439.8	798	13.082			
Total	29288.98	799				
ANOVA for ϕ and N						
Between Groups	45855.82	1	45855.82	605.5	0	3.853
Within Groups	60434.8	798	75.732			
Total	106290.6	799				
ANOVA for B/L and N						
Between Groups	17483.38	1	17483.38	1332.2	0	3.853
Within Groups	10473.24	798	13.124			
Total	27956.62	799				
ANOVA for H/B and N						
Between Groups	12037.42	1	12037.42	866.1	0	3.853
Within Groups	11090.8	798	13.898			
Total	23128.22	799				

where *SS* represents the sum of squares, *df* represents the degree of freedom, *MS* represents the mean of squares, and *F* and *F crit* represent the *F* state and critical value, respectively.

(k_h), rectangular excavation's aspect ratio (B/L), and excavated depth ratio (H/B) are significant parameters for determining the seismic stability number (N).

•The considered input parameters have no multicollinearity involved, so the performance of the proposed hybrid ML algorithms is not affected.

2.5 Analysis of Variance (ANOVA)

Analysis of Variance, commonly known as ANOVA, is a statistical method employed to examine the differences in group means within a given sample. ANOVA was used to determine whether there were statistically significant differences between the group means. This section presents the results of an ANOVA conducted using Excel 2021's data analysis features. Table 5 summarizes the findings from this analysis.

From the results presented in Table 5, it can be observed that the P-values for k_h, ϕ, B/L, and H/B are less than the significant values (*i.e.*, p<0.05), and the *F*-statistic values are greater than the critical *F* values for each of the input parameters. The null hypothesis (H_0) in ANOVA posits no significant differences between the group means, implying that all group population means are equal. Therefore, these ANOVA results reject the null hypothesis (H_0) of the current study.

2.6 Z- test

A Z-test is a statistical method employed to ascertain whether there is a notable difference between the means of two data sets, given that the population standard deviation is known. This parametric test, which relies on the normal distribution, is typically employed for comparing a sample mean to a known population mean. The results of the Z-test are summarized in Table 6. The obtained z-statistic value (z) is greater than both the critical one-tail and two-tail z-values. Furthermore, the critical one-tail z-value is less than the two-tail z-value, and the p-values for both the one-tail and two-tail is below the significance p value (*i.e.*, p<0.05). Therefore, it can be concluded that the Z-test supports the research hypothesis of this study.

3. Machine learning models and optimization algorithm

3.1 Random forest (RF)

Breiman[67] proposed the RF algorithm, a sophisticated ensemble learning algorithm that utilizes multiple decision trees for basic predictions. The advantages of the RF algorithm include not requiring feature scaling, efficient handling of non-linear parameters, and a reduction in the number of model parameters. These advantages contribute to the RF model's very low risk of overfitting during construction. The foundation of the RF regression algorithm lies in decision trees. By using various training datasets and feature spaces, the algorithm produces several decision tree models. As demonstrated in Fig. 4, the final result is obtained by amalgamating these individual decision tree outputs through techniques such as voting or averaging. The mathematical expression for the RF regression algorithm is detailed in Eq. (3) as follows:

$$Y(x) = \frac{1}{n}\sum_{i=1}^{k} h_i(x) \quad (3)$$

Here, $Y(x)$ represents the final output of the RF regression model, and $h_i(x)$ represents the output of the ith DT.

3.2 Whale Optimization Algorithm (WOA)

The whale optimization algorithm (WOA), a nature-inspired optimization algorithm, was proposed by Mirjalili and Lewis[68] The algorithm is inspired by the social behavior of humpback whales, particularly their hunting strategies. The WOA addresses optimization challenges by emulating the social interactions and hunting behaviors of these marine mammals. These algorithms are categorized as metaheuristic algorithms, signifying their role as versatile problem-solving methods suitable for various optimization challenges. The WOA algorithm is constructed in the following stages: encircling prey, creating spiral bubbles, and searching for prey.

• Encircling prey

First, after initializing a population of potential solutions, the WOA randomly determines the best search agent by finding the prey location within a search space. The objective is to relocate the prey's location. New positions are calculated using mathematical Eqs. (4) and (5), which are dependent on the current positions and the positions of the global best whale.

$$D = \left|\vec{C} \times \vec{X}^*(t) - \vec{X}(t)\right| \quad (4)$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \times \vec{D} \quad (5)$$

where $\vec{D}$ denotes the position vector between the whale and prey, $\vec{X}^*(t)$ represents the best position vector, and $\vec{X}(t)$

Table 6. Results of Z-test analysis for all variables.

Index	k_h	ϕ	B/L	H/B	N
Mean	0.15	25	0.5083	2.1	9.858
Known Variance	0.0125	125	0.0961	1.64	26.152
Observations	400	400	400	400	400
Hypothesized Mean Difference	0	0	0	0	-
z	37.9582	24.632	36.4988	29.4321	-
P(Z<=z) one-tail	0	0	0	0	-
z Critical one-tail	1.6448	1.6448	1.6448	1.6448	-
P(Z<=z) two-tail	0	0	0	0	-
z Critical two-tail	1.9599	1.9599	1.9599	1.9599	-

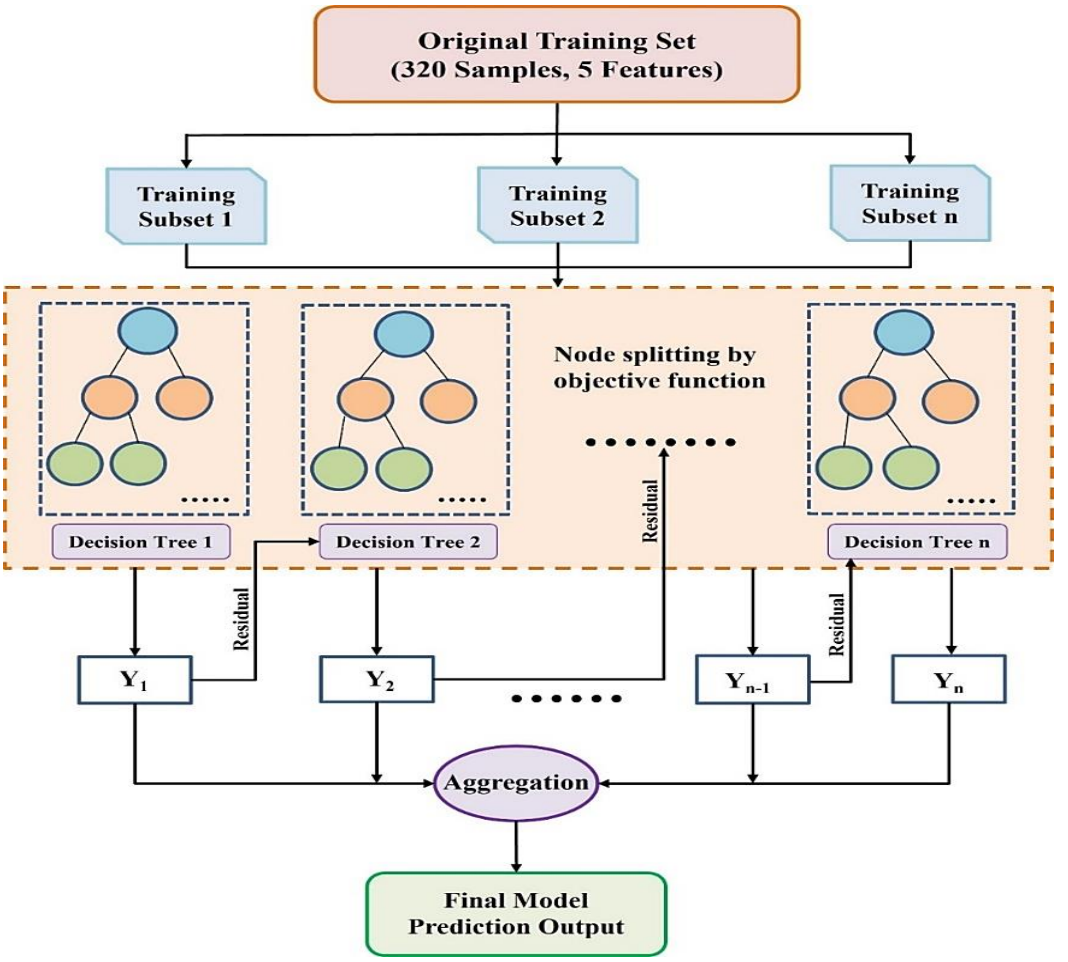

Fig. 4 Flowchart of the RF model.

represents the position vector. The best position vector is determined using the coefficient vectors $\vec{A}$ and $\vec{C}$, which are calculated using Eqs. (6) and (7), respectively.

$$\vec{A} = 2\vec{a} \times \vec{r} - \vec{a} \quad (6)$$

$$\vec{C} = 2\vec{r} \quad (7)$$

where $\vec{r}$ represents the linearly varying random vector from 2 to 0, and $\vec{a}$ represents a random number varying between 0 and 1.

- Creating Spiral Bubbles

During this phase, the whale employs a spiral formation strategy to capture its prey. This approach is characterized by a spiral equation that defines the distance between the whale and its prey, as outlined in Eq. (8).

$$\vec{X}(t+1) = \overrightarrow{D^t} \times e^{bt} \times \cos(2\pi l) + \vec{X}^*(t) \quad (8)$$

where the shape of the logarithmic helix is defined by a constant b, and l represents the random vector with a uniform distribution in the range from -1 to $+1$.

The final solution is the position of the best whale found during the optimization process by selecting bubble net predation based on the probability function p defined using Eq. (9).

$$X_w(t+1) = \begin{cases} x_p(t) - A \times D & p \le 0.5 \\ D \times e^{bt} \times \cos(2\pi l) + \vec{X}^*(t) & p \ge 0.5 \end{cases} \quad (9)$$

where the probability p lies in the range of [0,1] and denotes the probability of the predation mechanism.

- Searching for prey

In this phase, the algorithm imitates the exploitation behavior of whales when they locate prey.

3.3 Dragonfly optimization algorithm (DOA)

The dragonfly algorithm (DA), inspired by the dynamic and static swarming behaviors of dragonflies, such as their hunting and migration patterns, is a swarm intelligence optimization algorithm. Introduced by Mirjalili,[69] this algorithm draws from the natural tendencies of dragonflies to form smaller groups for foraging and larger groups for migration. The DA operates through a series of steps as follows:

Step 1: Initialize a population of dragonflies. Each dragonfly is represented by a solution to the optimization problem.

Step 2: Evaluate the fitness of each dragonfly.

Step 3: For each dragonfly, perform the following steps: a) The dragonfly is moved to a new position by adding a weighted combination of several vectors to its current position, which includes i) the position of the best dragonfly in the population, ii) the position of the dragonfly's nearest neighbor, and iii) the dragonfly's position. b) If the new position offers better fitness, retain it; otherwise, revert to the old position.

Step 4: Repeat steps 2 and 3 until the algorithm achieves convergence or hits the predetermined maximum iteration

count.

This research does not provide a detailed exploration of the underlying mechanisms of DA operation because Mirjalili[69] has made significant contributions to this area. For a comprehensive explanation of DA, readers are encouraged to consult Mirjalili's seminal 2016 work.[69]

3.4 Sparrow Search Optimization Algorithm (SSOA)

The sparrow search algorithm (SSA), proposed by Xue and Shen,[70] is a swarm intelligence optimization algorithm inspired by the foraging and anti-predation behaviors of sparrows. The SSA is simple to implement and understand due to its fast search and convergence behavior. The working procedures of the SSA are presented as follows:

Step 1: Initialize a population of sparrows, classified as either discoverers or followers, with each sparrow representing a solution to the optimization problem.

Step 2: Evaluate the fitness of each sparrow.

Step 3: For each sparrow, perform the following actions: a) move the sparrow to a new position by adding a random vector to its current position. b) maintain the new position if it offers better fitness; otherwise, revert to the original position.

Step 4: Keep repeating steps 2 and 3 until the algorithm either converges or the set maximum number of iterations is met.

It is important to note that the study does not fully disclose the specific functioning principles of SSA. More details about the SSA can be found in the original publications of Xue and Shen.[70]

3.5 The development procedure of the hybrid RF and OA models

In this paper, the RF model is used to predict the seismic stability number (N) and three optimization algorithms, namely, the WOA, DOA, and SSOA, are used to optimize the hyperparameters (n_{tree} and m_{try}) of the RF model. The use of multiple OAs to improve the efficiency of random forest models has been the subject of a great deal of research over the past few years. In the RF model, to obtain a more robust aggregate model, a greater number of estimators is used, which requires more computational power. OAs can efficiently achieve an optimal result and a more robust hybrid model by optimizing the hyperparameters (such as N_estimators (n_{tree}), max_depth, max_features (m_{try}), and bootstrap, *etc.*) of RF models through exploration and exploitation methods. Ultimately, the predictive efficacy of the proposed hybrid models was assessed through the computation of various performance metrics (R^2, VAF, RMSE, MAE, PI, and WMAPE) for the training and testing sets. The following steps are involved in building the hybrid RF models:

(1) The data was preprocessed and separated into training and testing phases.

(2) Model development: In this step, following data preparation, the parameters of the RF model are initialized. Subsequently, various optimization algorithms are sequentially initialized to optimize the RF model's hyperparameters, including N_estimators (n_{tree}), max_depth, max_features (m_{try}), bootstrap, and max_samples. In this study, three optimization algorithms, namely, the WOA, DOA, and SSOA, were utilized, leading to the development of the RF-WOA, RF-DOA, and RF-SSOA hybrid models. After repeatedly searching for the global optimum position, the suggested OAs output the final parameters of the RF model. This involves initializing the populations, determining the fitness of each search agent, and altering the current search agent's position, and, if necessary, checking and correcting the search space border. Fig. 5 shows the flowchart of the

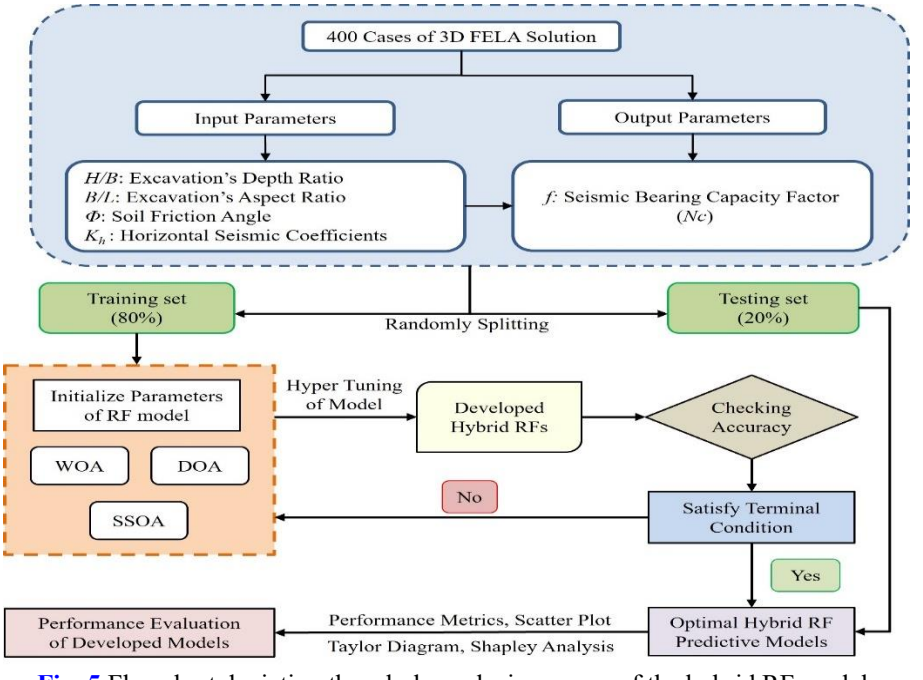

Fig. 5 Flowchart depicting the whole analysis process of the hybrid RF model.

methodology used in constructing the hybrid RF model.

(3) Optimum parametric configuration of the proposed hybrid prediction model: It is vital to look at the hybrid models that offer the best performance and best hyperparameter values. It is essential to examine a reliable hybrid model with optimum values of hyperparameters such as N_estimators (n_{tree}), max_depth, max_features (m_{try}), bootstrap, and max_samples. In this study, to obtain the optimal hyperparameters, a trial-and-error approach was used during the model construction process. The optimum values of the hyperparameters for the developed models are summarized in Table 7.

Table 7. Optimum hyperparameters for the proposed hybrid RF model.

Models	Optimal hyperparameters
RF-WOA	Population Size = 100, the number of trees = 12, max_features (m_{try})=5, Maximum iteration= 400, Bootstrap =True.
RF-DOA	Population Size=150, the number of trees = 20, max_features (m_{try})=7, Maximum iteration= 400, Bootstrap =True.
RF-SSOA	Population Size=120, the number of trees = 15, max_features (m_{try})=5, Maximum iteration= 400, Bootstrap =True.

The convergence curves of all the models are shown in Fig. 6; the proposed RF-DOA converges more quickly than do the other hybrid RF models. The computational costs at 400 iterations were 578.02 s, 664.39 s, and 485.26 s for the RF-WOA, RF-DOA, and RF-SSOA, respectively.

3.6 Cross-validation accuracy

In this study, to assess the accuracy and generalizability of the proposed predictive models, a 10-fold cross-validation technique was employed. By employing 10-fold cross-validation techniques, models are less prone to overfitting. In this process, the whole dataset is divided into 10 equal-size folds, with one fold used as the test set and the remaining nine folds used as the training set to evaluate the model's performance. This process was repeated 10 times by changing the different folds to form a test set, after which 10 estimates of the proposed model's performance were obtained. Finally, the accuracy of the proposed hybrid RF algorithm was reported as an average accuracy. A schematic diagram of the 10-fold cross-validation process is shown in Fig. 7. The results of 10-fold cross-validation are presented in Fig. 8(a) for the coefficient of determination, Fig. 8(b) for the RMSE value corresponding to each fold, Fig. 8(c) for the MAE value corresponding to each fold, and Fig. 8(d) for the overall statistical measure presented as the mean value of R^2, RMSE and MAE of the developed models.

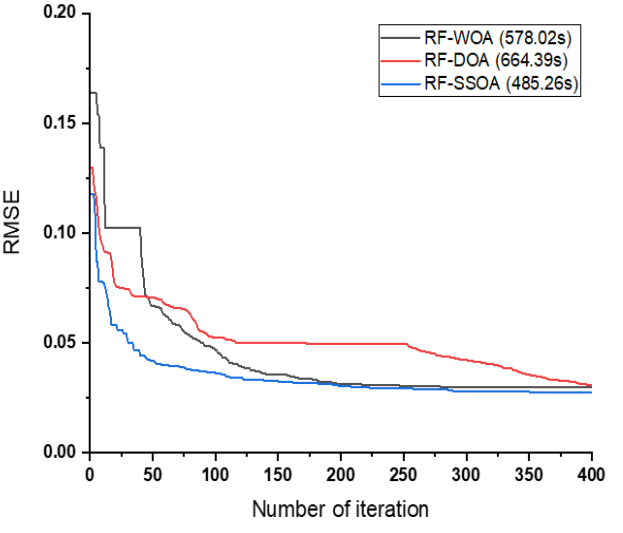

Fig. 6 Convergence Curve for all proposed model.

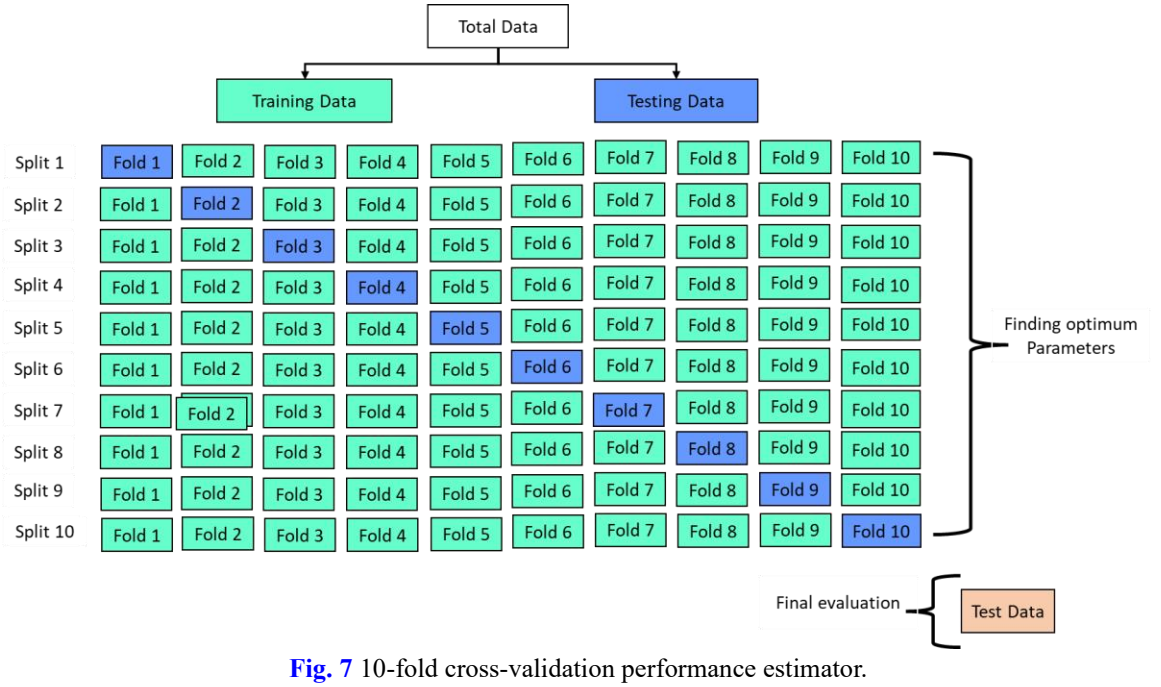

Fig. 7 10-fold cross-validation performance estimator.

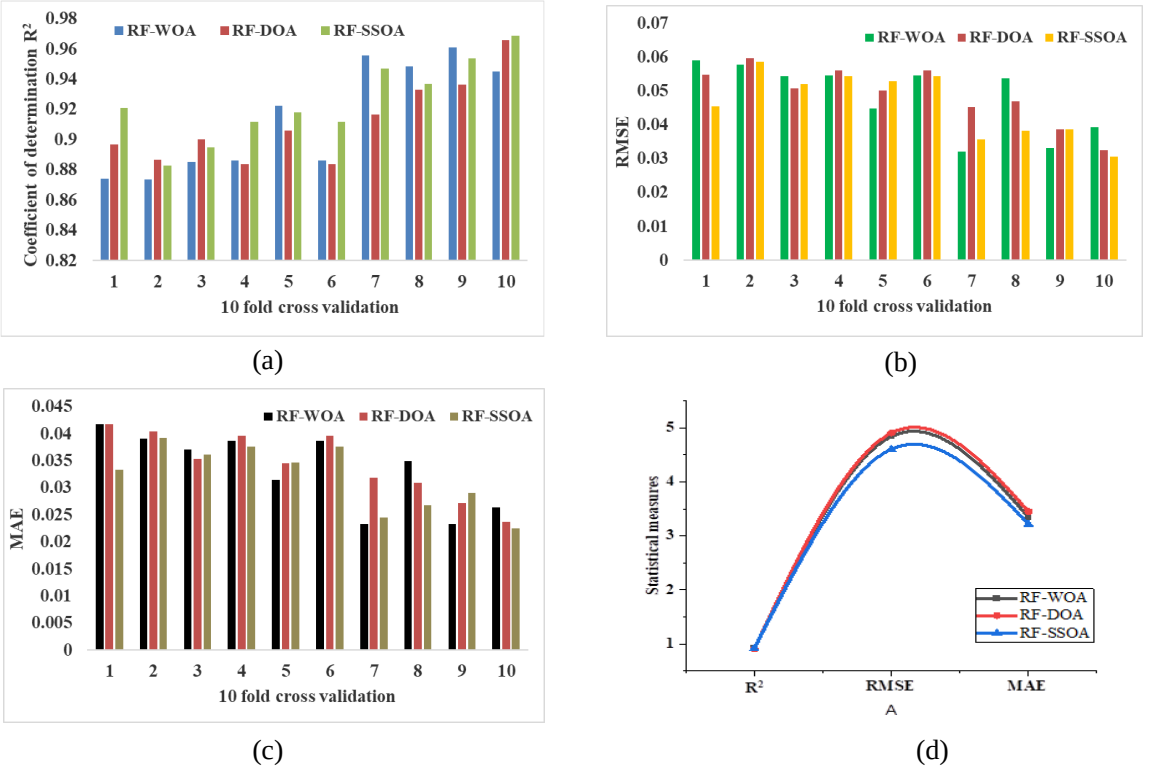

Fig. 8 Comparison of 10-fold cross-validation; (a) R^2; (b) RMSE; (c) MAE; (d) Overall comparison of means values of metrics.

4. Results

4.1 Performance evaluation

The performance assessment of the proposed hybrid machine learning models is a crucial step in the development and deployment of machine learning techniques for predicting the seismic stability number (N). The proposed models perform well on a given 3D LB and UB FELA dataset. Several performance metrics and techniques are used to assess the performance of machine learning models. In this study, as mentioned above, six performance metrics, namely, R^2, the WMAPE, the RMSE, the VAF, the PI, and the MAE, were evaluated to assess the performance of the proposed models. These metrics are widely utilized for evaluating the performance of various machine-learning models. Among these metrics, the RMSE, MAE, and WMAPE are common error parameters, while R^2, VAF and PI are the accuracy and performance metrics, respectively. R^2 quantifies the fraction of variance in the dependent variable that is predictable from the independent variables in the model. A higher R-squared indicates a better fit. Ideally, in a perfect and robust model, these performance metrics should align with their ideal values, as detailed in Table 2. The results of the performance assessment of the hybrid RF models for the training and testing phases are presented in Table 8.

As per the results presented in Table 8, the proposed RF-DOA model achieved the target precision ($R^2 = 0.971$ and RMSE = 0.027) in the training phase, followed by the RF-WOA ($R^2 = 0.966$ and RMSE = 0.030) and RF-SSOA ($R^2 = 0.963$ and RMSE = 0.031). Similarly, in the testing phase, the RF-DOA model again achieved the desired precision ($R^2 = 0.975$ and RMSE = 0.025), followed by the RF-WOA ($R^2 = 0.973$ and RMSE = 0.025) and RF-SSOA ($R^2 = 0.963$ and RMSE = 0.030). The performance of the developed RF-DOA is quite satisfactory in both the training and testing phases compared to that of the RF-WOA and RF-SSOA models. However, among the developed hybrid RF models, RF-DOA was found to be the most effective, and RF-SSOA was the least effective in both the training and testing phases of seismic stability number (N) determination.

Table 8. Performance of the proposed model in the training and testing phases.

Parameters	RF-WOA		RF-DOA		RF-SSOA	
	Train	Test	Train	Test	Train	Test
R^2	0.966	0.973	0.971	0.975	0.963	0.963
WMAPE	0.109	0.104	0.102	0.102	0.108	0.118
RMSE	0.030	0.025	0.027	0.025	0.031	0.030
VAF	96.59	97.22	97.12	97.24	96.27	95.93
PI	1.902	1.920	1.915	1.922	1.894	1.891
MAE	0.020	0.018	0.019	0.017	0.020	0.020

Scatterplots, error plots, and line plots are often used in machine learning to visualize relationships between the actual and estimated seismic stability number (N). When working with machine learning models, scatterplots, error plots, and line plots can be useful for various purposes, such as understanding the data, assessing model performance, and visualizing model predictions. The scatterplots, error plots, and line plots for the RF-WOA, RF-DOA, and RF-SSOA models are presented in Figs. 9, 10, and 11, respectively.

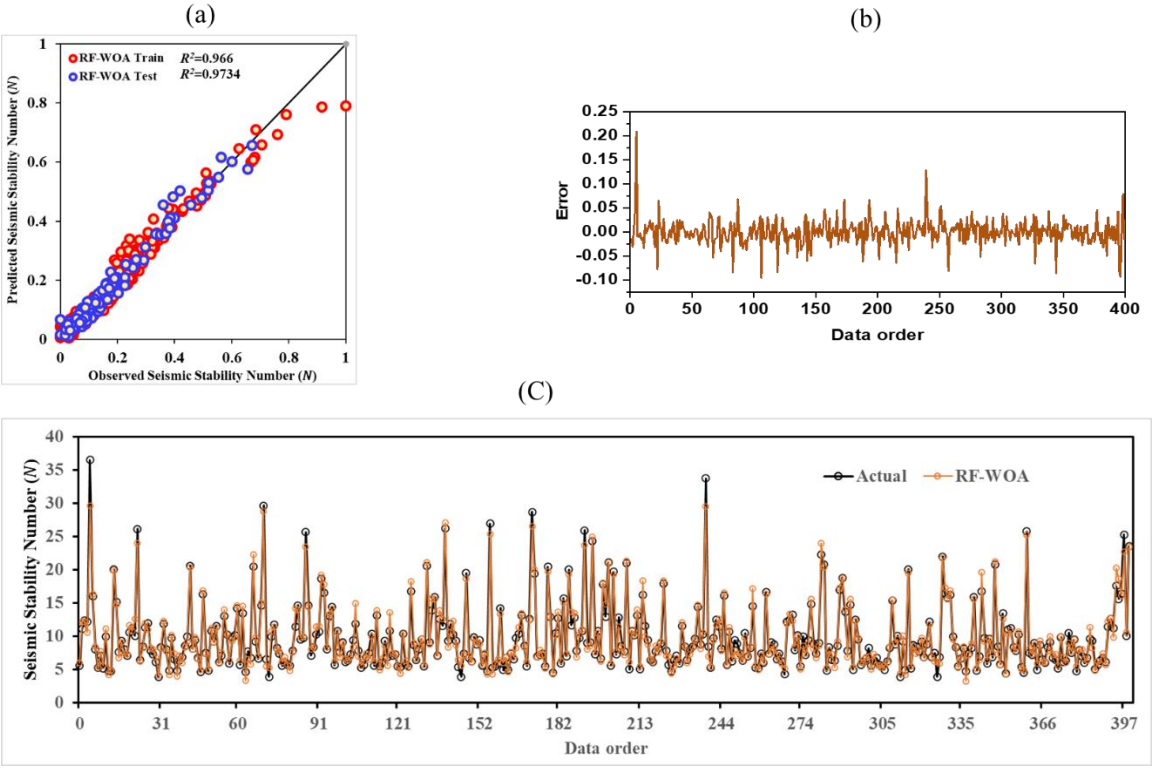

Fig. 9 Performance of RF-WOA model (a) scatterplot, (b) error plot, (c) line plot.

A Taylor diagram or Taylor plot, which is a graphical representation used to compare the performance of machine learning models, was used to analyze and compare the performance of all the proposed models on both the training and testing datasets. Taylor[71] developed this approach as a way to visualize and compare the skills of different machine learning models. The results of the Taylor chart presented in Figs.12 and 13 demonstrate that the RF-DOA model outperformed the other models in the training and testing phases.

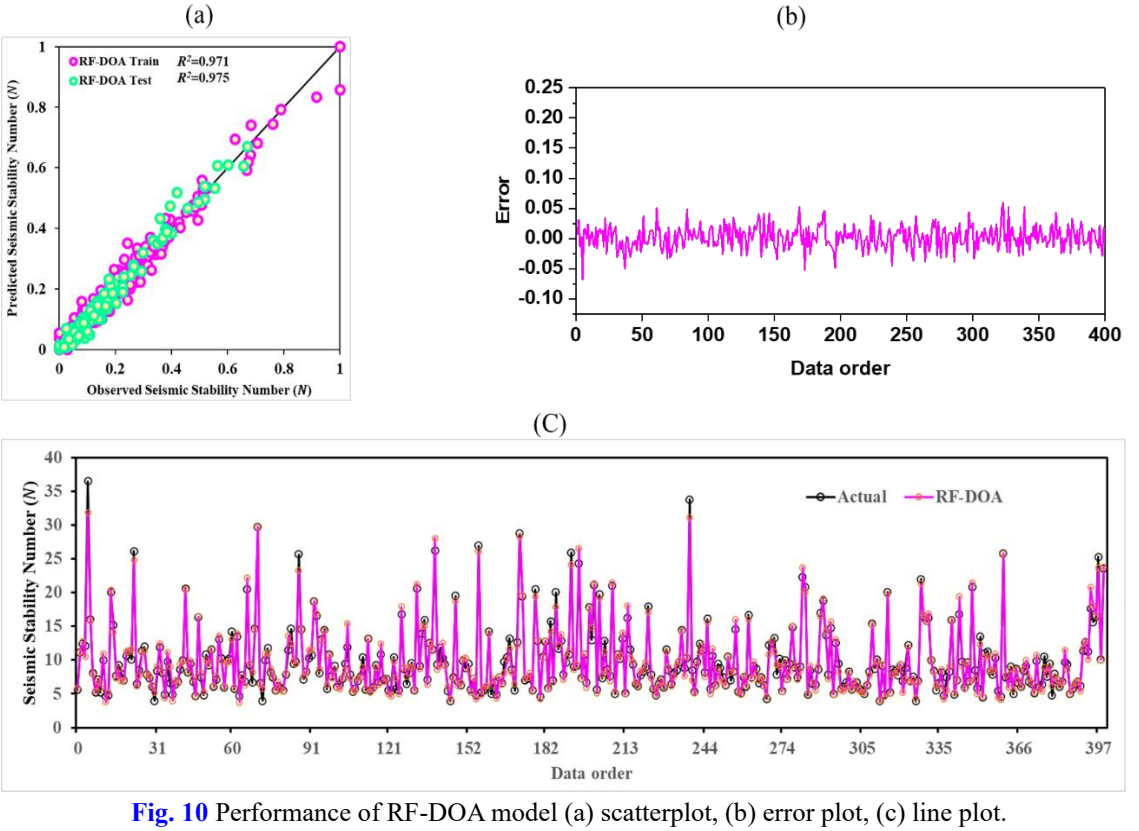

Fig. 10 Performance of RF-DOA model (a) scatterplot, (b) error plot, (c) line plot.

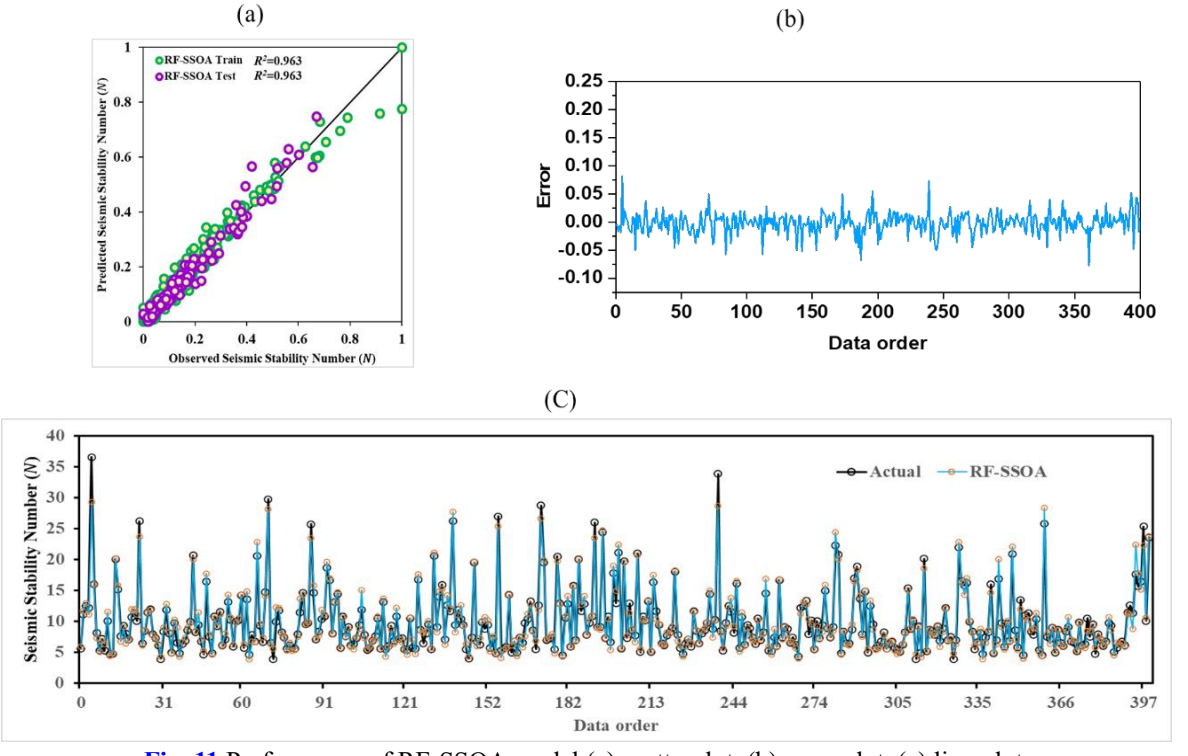

Fig. 11 Performance of RF-SSOA model (a) scatterplot, (b) error plot, (c) line plot.

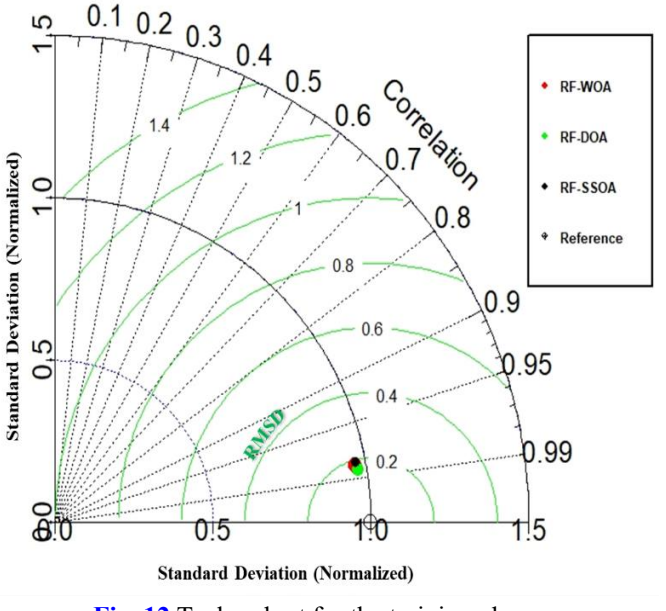

Fig. 12 Taylor chart for the training phase.

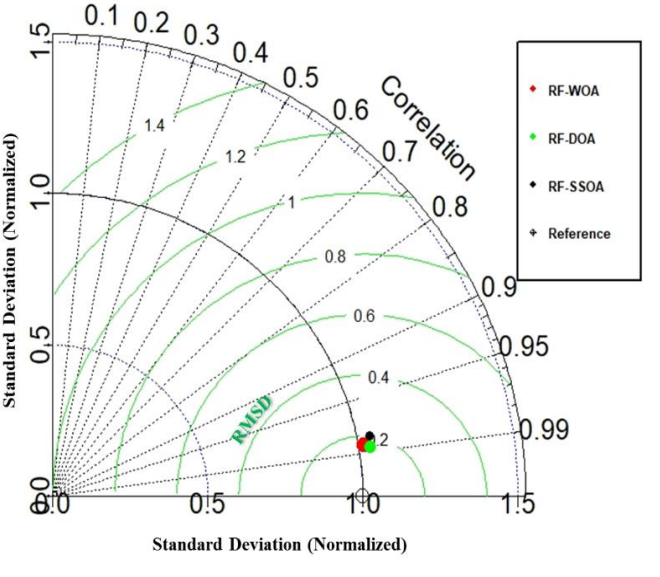

Fig. 13 Taylor chart for testing phase.

4.2 Shapley analysis

As mentioned in the Results and discussion section, the proposed hybrid ML models can accurately predict the seismic stability number (N) by varying the excavation's aspect ratio, excavated depth ratio, horizontal seismic coefficient, and soil material's friction angle. However, accurately estimating the exact value of the seismic stability number (N) is challenging due to the uncertainty associated with the soil material's friction angle and horizontal seismic coefficient. As a result, it is crucial to investigate the effects of all the inputs on the seismic stability index (N).

This section applies the Shapley-Additive Explanation approach to identify the weight and impact of each input variable on the seismic stability index (N). The Shap analysis revealed that the soil friction angle (ϕ), excavated depth ratio (H/B), rectangular excavation's aspect ratio (B/L), and horizontal seismic coefficient (k_h) significantly impact the seismic stability number (N). According to a previous study, ϕ, H/B, and B/L are positively correlated, but k_h is negatively correlated. They also concluded that excavation failure is slightly affected by a small ϕ, but a large ϕ makes it more prone to excavation failure. Furthermore, Fig. 14 shows that the relative shapely significance of each input variable is

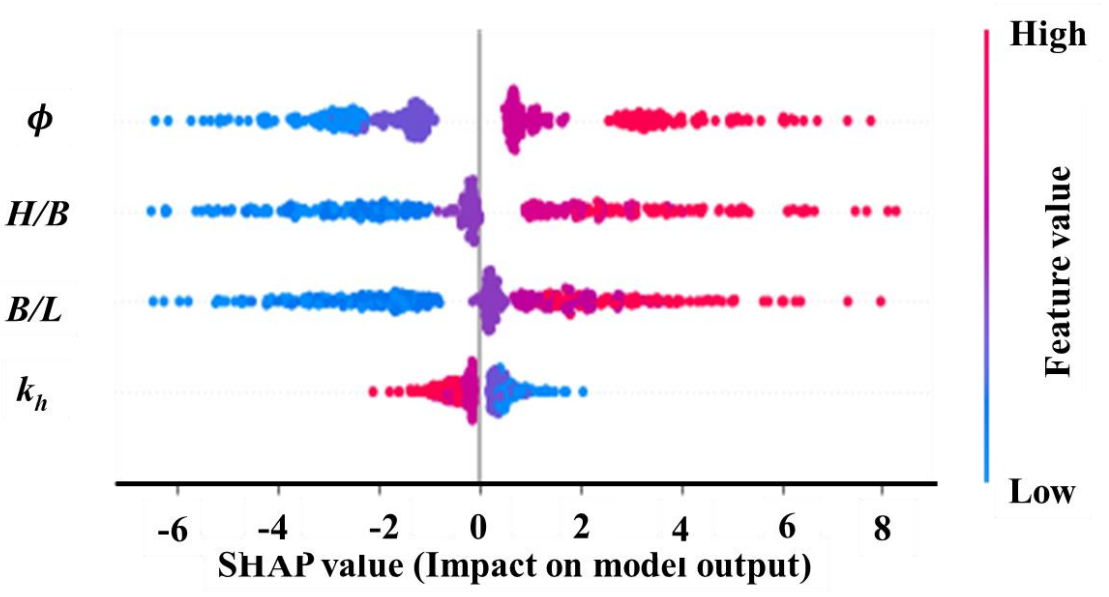

Fig. 14 Shapley analysis of features.

related to the seismic stability number (N). Additionally, Figs. 14 and 15 present the SHAP graph's results for all the dimensionless input parameters concerning the seismic stability number (N). In Fig. 14, the x-axis represents the SHAP value, which represents the impact of each input parameter, and the contribution of each variable is depicted as dot points in distinct colors. The soil friction angle (ϕ), excavated depth ratio (H/B), and rectangular excavation's aspect ratio (B/L) have strong positive correlations with the seismic stability number (N).

5. Discussion

The preceding sections offer a thorough evaluation of the predictive capabilities of the proposed hybrid random forest algorithm in forecasting seismic stability numbers (N) for unsupported rectangular excavations. Various performance metrics were scrutinized and analyzed to compare the predictive efficacy of the proposed models. A dataset comprising 320 three-dimensional finite element limit analysis solutions was employed for model development, and an additional 80 test data points were utilized to validate the performance of the developed models. However, a thorough examination of the results revealed that the RF-DOA outperforms the other proposed hybrid random forest algorithms in both the training and testing phases. The convergence behavior also shows that RF-DOA outperforms RF-WOA and RF-DOA after approximately 350 iterations. On the other hand, shapely analysis reveals that the horizontal seismic coefficient (k_h) has the least significant impact, while the soil friction angle (ϕ) has the most significant influence on predicting the seismic stability number (N).

5.1 Theoretical implications

This research pioneered the use of hybrid machine learning algorithms to assess seismic stability numbers (N) in unsupported rectangular excavations within cohesive-frictional soils subjected to pseudo-static seismic forces. Theoretical implications involve understanding the inner

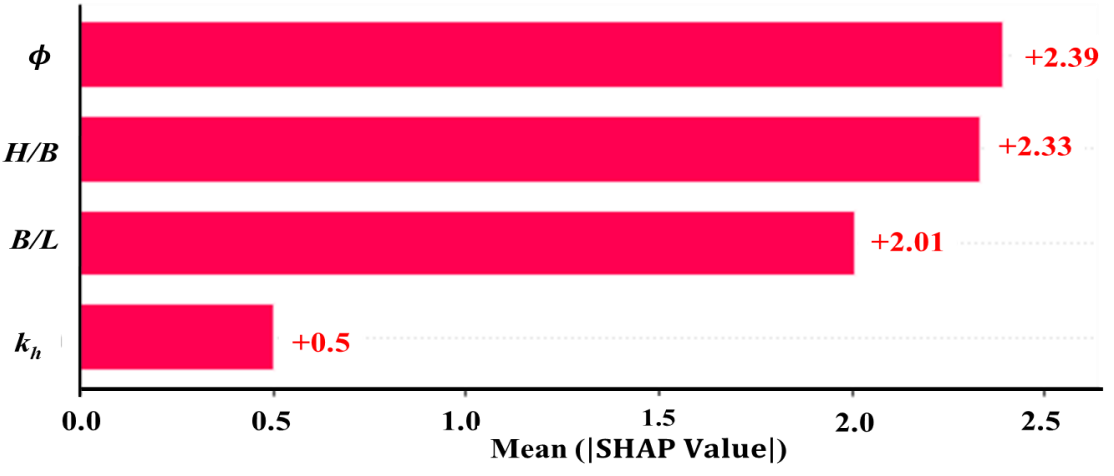

Fig. 15 Features importance.

workings of these models and their relationship to seismic stability numbers (*N*). Machine learning aids in pinpointing crucial factors impacting stability numbers (*N*); furthermore, scrutinizing feature importance enhances the theoretical understanding of parameters influencing unsupported rectangular excavations during seismic events. Considering the quality and representativeness of the training data is imperative in theory.

5.2 Practical implications

The proposed machine learning algorithms can contribute to the development of early warning systems for seismic events. If the stability of an excavation can be predicted in advance, appropriate safety measures can be implemented, potentially preventing accidents and minimizing damage. The proposed machine learning algorithms can assist in optimizing the design of rectangular excavations to enhance their seismic stability. Engineers can utilize predictive models to progressively enhance designs, making them more resilient to seismic forces. In practical terms, the use of machine learning extends to risk assessment, aiding decision-makers in comprehending potential project-related risks and taking suitable measures for mitigation. The integration of these algorithms with real-time monitoring systems amplifies their practical utility, allowing for continuous data updates that improve prediction accuracy and facilitate dynamic adjustments in response to evolving geological and seismic conditions.

5.3 Limitations

Despite the RF-DOA hybrid model achieving more accurate results in predicting the seismic stability number *N*, several shortcomings of this study need to be addressed in the future. For instance, the study considered only limited criteria for assessing the stability of an unsupported rectangular excavation in seismic-risk areas. Regarding the practical use of excavations in geological engineering, the components that affect excavations are more challenging to quantify and acquire precisely. For such prediction problems, more sophisticated and high-performance models may be developed in the foreseeable future. During this time, the accuracy of the RF model could be vastly enhanced by implementing an appropriate optimization technique. Extending the database size related to the invalidation risk of an unsupported excavation and utilizing a data enhancement algorithm could improve the prediction accuracy of the optimized model. Additionally, these approaches could reduce the generalization error of the hybrid model.

6. Conclusion

This study employs several hybrid machine learning algorithms, such as the RF-WOA, RF-DOA, and RF-SSOA, to determine the seismic stability number (N). For the training and testing of the models, a total of 400 FELA numerical solutions are used. The performance of the proposed models is demonstrated and compared using multiple statistical performance metrics. The results obtained through various analyses confirm the superiority of hybrid machine learning algorithms.

The following major conclusions can be drawn from this analysis.

1) The RF-DOA machine learning algorithm exhibits the best performance compared to the other proposed hybrid algorithms, with a smaller deviation between the observed and model-predicted seismic stability numbers (N). Furthermore, the RF-DOA model demonstrated a greater coefficient of determination ($R^2 = 0.967$ for training and $R^2 = 0.967$ for testing), followed by the RF-WOA and RF-SSOA.

2) According to the statistical analysis in Section 4.1, the RF-DOA model yields lower error metric values of the RMSE, MAE, and WMAPE. This trend toward lower error metrics indicates the superior performance of the proposed RF-DOA model.

3) 10-fold cross-validation was performed to assess the model's accuracy and to prevent overfitting of the models. Thus, the proposed RF-DOA model outperforms the RF-WOA and RF-SSOA models in predicting a more accurate value of the seismic stability number (N).

4) Shapley analysis was also performed to assess the impact of each input variable on the seismic stability index (N). From this analysis, it can be observed that the soil friction angle (ϕ), excavated depth ratio (H/B), rectangular excavation's aspect ratio (B/L), and horizontal seismic coefficient (k_h) significantly impact on the N value. However, among these, the horizontal seismic coefficient (k_h) has the least significant impact, while the soil friction angle (ϕ) has the most significant influence on predicting the N value.

However, certain constraints require further investigation in the future: (a) The proposed hybrid random forest models necessitate comprehensive validation using extensive datasets sourced from diverse domains, including experimental or numerical data. (b) A detailed comparative analysis should be conducted on the performance of the proposed hybrid random forest models with other robust models such as MARS, GP, XGBoost, MT, or SVM. (c) The proposed hybrid random forest models for predicting the seismic stability index (N) are valid for the defined range of input parameters, such as the soil friction angle (ϕ), excavated depth ratio (H/B), rectangular excavation's aspect ratio (B/L), and horizontal seismic coefficient (k_h).

Acknowledgment

This work was supported by Thammasat University Research Unit in Data Science and Digital Transformation.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

References

[1] D. W. Taylor, Fundamentals of soil mechanics, *Soil Science*, 1948, **66**, 161, doi: 10.1097/00010694-194808000-00008.

[2] N. Janbú, Stability Analysis of Slopes with Dimensionless Parameters, Doctoral Thesis. Harvard University, *Division of Engineering and Applied Physics. Cambridge, MA*, 1954.

[3] W.-F. Chen, *Limit Analysis and Soil Plasticity*. Elsevier, 2013; ISBN 0444601066.

[4] J. Pastor, S. Turgeman, Limit analysis in axisymmetrical problems: numerical determination of complete statical solutions, *International Journal of Mechanical Sciences*, 1982, **24**, 95-117, doi: 10.1016/0020-7403(82)90041-8.

[5] S. W. Sloan, Geotechnical stability analysis, *Géotechnique*, 2013, **63**, 531-571, doi: 10.1680/geot.12.rl.001.

[6] H. S. Yu, R. Salgado, S. W. Sloan, J. M. Kim, Limit analysis versus limit equilibrium for slope stability, *Journal of Geotechnical and Geoenvironmental Engineering*, 1998, **124**, 1-11, doi: 10.1061/(asce)1090-0241(1998)124: 1(1).

[7] C. M. Martin, The use of adaptive finite-element limit analysis to reveal slip-line fields, *Géotechnique Letters*, 2011, **1**, 23-29, doi: 10.1680/geolett.11.00018.

[8] D. V. Griffiths, N. Koutsabeloulis, Finite element analysis of vertical excavations, *Computers and Geotechnics*, 1985, **1**, 221-235, doi: 10.1016/0266-352X(85)90025-4.

[9] A. M. Britto, O. Kusakabe, Stability of unsupported axisymmetric excavations in soft clay, *Géotechnique*, 1982, **32**, 261-270, doi: 10.1680/geot.1982.32.3.261.

[10] A. M. Britto, O. Kusakabe, Stability of axisymmetric excavations in normally consolidated clays, Cambridge University Engineering Department Report CUED/D-SOILS, 1982, **109**, 666-681.

[11] A. Bottero, R. Negre, J. Pastor, S. Turgeman, Finite element method and limit analysis theory for soil mechanics problems, *Computer Methods in Applied Mechanics and Engineering*, 1980, **22**, 131-149, doi: 10.1016/0045-7825(80)90055-9.

[12] J. Pastor, T.-H. Thai, P. Francescato, New bounds for the height limit of a vertical slope, *International Journal for Numerical and Analytical Methods in Geomechanics*, 2000, **24**, 165-182, doi: 10.1002/(SICI)1096-9853(200002)24: 2165: AID-NAG62>3.0.CO;2-A.

[13] V. N. Khatri, J. Kumar, Stability of an unsupported vertical circular excavation in clays under undrained condition, *Computers and Geotechnics*, 2010, **37**, 419-424, doi: 10.1016/j.compgeo.2009.11.001.

[14] J. Kumar, D. Chakraborty, Stability numbers for an unsupported vertical circular excavation in c–ϕ soil, *Computers and Geotechnics*, 2012, **39**, 79-84, doi: 10.1016/j.compgeo.2011.08.002.

[15] J. Kumar, M. Chakraborty, J. P. Sahoo, Stability of unsupported vertical circular excavations, *Journal of Geotechnical and Geoenvironmental Engineering*, 2014, **140**, 04014028, doi: 10.1061/(asce)gt.1943-5606.0001118.

[16] S. Keawsawasvong, B. Ukritchon, Stability of unsupported conical excavations in non-homogeneous clays, *Computers and Geotechnics*, 2017, **81**, 125-136, doi: 10.1016/j.compgeo.2016.08.007.

[17] B. Ukritchon, S. Keawsawasvong, A new design equation for drained stability of conical slopes in cohesive-frictional soils, *Journal of Rock Mechanics and Geotechnical Engineering*, 2018, **10**, 358-366, doi: 10.1016/j.jrmge.2017.10.004.

[18] W. Yodsomjai, S. Keawsawasvong, S. Likitlersuang, Stability of unsupported conical slopes in hoek-brown rock masses, *Transportation Infrastructure Geotechnology*, 2021, **8**, 279-295, doi: 10.1007/s40515-020-00137-4.

[19] W. Yodsomjai, S. Keawsawasvong, T. Senjuntichai, Undrained stability of unsupported conical slopes in anisotropic clays based on anisotropic undrained shear failure criterion, *Transportation Infrastructure Geotechnology*, 2021, **8**, 557-568, doi: 10.1007/s40515-021-00153-y.

[20] W. Yodsomjai, S. Keawsawasvong, C. Thongchom, J. Lawongkerd, Undrained stability of unsupported conical slopes in two-layered clays, *Innovative Infrastructure Solutions*, 2020, **6**, 15, doi: 10.1007/s41062-020-00384-x.

[21] V. Q. Lai, Nguyen D.-K., R. Banyong, Keawsawasvong S, Limit analysis solutions for stability factor of unsupported conical slopes in clays with heterogeneity and anisotropy, *International Journal of Computational Materials Science and Engineering*, 2022, **11**, 2150030, doi:10.1142/S2047684121500305.

[22] B. Ukritchon, S. Yoang, S. Keawsawasvong, Undrained stability of unsupported rectangular excavations in non-homogeneous clays, *Computers and Geotechnics*, 2020, **117**, 103281, doi: 10.1016/j.compgeo.2019.103281.

[23] V. Q. Lai, J. Shiau, S. Keawsawasvong, S. Seehavong, L. T. Cabangon, Undrained stability of unsupported rectangular excavations: anisotropy and non-homogeneity in 3D, *Buildings*, 2022, **12**, 1425, doi: 10.3390/buildings12091425.

[24] P. Petchkaew, S. Keawsawasvong, W. Tanapalungkorn, S. Likitlersuang, Seismic stability of unsupported vertical circular excavations in c-φ soil, *Transportation Infrastructure Geotechnology*, 2023, **10**, 165-179, doi: 10.1007/s40515-021-00221-3.

[25] P. Petchkaew, S. Keawsawasvong, W. Tanapalungkorn, S. Likitlersuang, 3D stability analysis of unsupported rectangular excavation under pseudo-static seismic body force, *Geomechanics and Geoengineering*, 2023, **18**, 175-192, doi: 10.1080/17486025.2021.2019321.

[26] S. Likitlersuang, C. Chheng, S. Keawsawasvong, Structural modelling in finite element analysis of deep excavation, *Journal of GeoEngineering*, 2019, **14**, 121-128, doi: 10.6310/jog.201909_14(3).1.

[27] T. S. Nguyen, S. Likitlersuang, Influence of the spatial variability of soil shear strength on deep excavation: a case study of a bangkok underground MRT station, *International Journal of Geomechanics*, 2021, **21**, 04020248, doi: 10.1061/(asce)gm.1943-5622.0001914.

[28] B.-C B. Hsiung, S. Likitlersuang, K. H. Phan, P. Pisitsopon, Impacts of the plane strain ratio on excavations in soft alluvium

deposits, *Acta Geotechnica*, 2021, **16**, 1923-1938, doi: 10.1007/s11440-020-01115-3.
[29] W. Tanapalungkorn, W. Yodsomjai, S. Keawsawasvong, T. S. Nguyen, W. Chim-Oye, P. Jongpradist, S. Likitlersuang, Undrained stability of braced excavations in clay considering the nonstationary random field of undrained shear strength, *Scientific Reports*, 2023, **13**, 13358, doi: 10.1038/s41598-023-40608-5.
[30] N. Muenpetch, S. Keawsawasvong, V. Komolvilas, S. Likitlersuang, Numerical investigation on impact of excavations in influence zone of existing MRT tunnels, *Geomechanics and Geoengineering*, 2023, 1-20, doi: 10.1080/17486025.2023.2226107.
[31] L. Z. Mase, S. Likitlersuang, T. Tobita, Analysis of seismic ground response caused during strong earthquake in Northern Thailand, *Soil Dynamics and Earthquake Engineering*, 2018, **114**, 113-126, doi: 10.1016/j.soildyn.2018.07.006.
[32] S. Likitlersuang, P. Plengsiri, L. Z. Mase, W. Tanapalungkorn, Influence of spatial variability of ground on seismic response analysis: a case study of Bangkok subsoils, *Bulletin of Engineering Geology and the Environment*, 2020, **79**, 39-51, doi: 10.1007/s10064-019-01560-9.
[33] L. Z. Mase, S. Likitlersuang, Implementation of seismic ground response analysis in estimating liquefaction potential in northern Thailand, *Indonesian Journal on Geoscience*, 2021, **8**, 371–383, doi: 10.17014/ijog.8.3.371-383.
[34] L. Z. Mase, W. Tanapalungkorn, S. Likitlersuang, K. Ueda, T. Tobita, Liquefaction analysis of Izumio sands under variation of ground motions during strong earthquake in Osaka, Japan, *Soils and Foundations*, 2022, **62**, 101218, doi: 10.1016/j.sandf.2022.101218.
[35] P. Hong-in, S. Keawsawasvong, V. Q. Lai, T. S. Nguyen, W. Tanapalungkorn, S. Likitlersuang, 3D stability and failure mechanism of undrained clay slopes subjected to seismic load, *Geotechnical and Geological Engineering*, 2023, **41**, 3941-3969, doi: 10.1007/s10706-023-02497-3.
[36] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Soft computing techniques for predicting penetration and uplift resistances of dual pipelines in cohesive soils, *Engineered Science*, 2023, **24**, 897, doi: 10.30919/es897.
[37] R. Kumar, A. Kumar, D. Ranjan Kumar, Buckling response of CNT based hybrid FG plates using finite element method and machine learning method, *Composite Structures*, 2023, **319**, 117204, doi: 10.1016/j.compstruct.2023.117204.
[38] D. R. Kumar, P. Samui, A. Burman, Prediction of probability of liquefaction using soft computing techniques, *Journal of the Institution of Engineers (India): Series A*, 2022, **103**, 1195-1208, doi: 10.1007/s40030-022-00683-9.
[39] R. Kumar, B. Rai, P. Samui, A comparative study of prediction of compressive strength of ultra-high performance concrete using soft computing technique, *Structural Concrete*, 2023, **24**, 5538-5555, doi: 10.1002/suco.202200850.
[40] R. Kumar, B. Rai, P. Samui, Machine learning techniques for prediction of failure loads and fracture characteristics of high and ultra-high strength concrete beams, *Innovative Infrastructure Solutions*, 2023, **8**, 219, doi: 10.1007/s41062-023-01191-w.
[41] H. F. Isleem, N. D. K. R. Chukka, A. Bahrami, S. Oyebisi, R. Kumar, T. Qiong, Nonlinear finite element and analytical modelling of reinforced concrete filled steel tube columns under axial compression loading, *Results in Engineering*, 2023, **19**, 101341, doi: 10.1016/j.rineng.2023.101341.
[42] V. Kumar, A. Burman, F. H. M. Portelinha, M. Kumar, G. Das, Influence of variation of soil properties in bearing capacity and settlement analysis of a strip footing using random finite element method, *Civil Engineering Infrastructures Journal*, 2023, doi: 10.22059/CEIJ.2023.360871.1930.
[43] D. R. Kumar, W. Wipulanusat, M. Kumar, S. Keawsawasvong, P. Samui, Optimized neural network-based state-of-the-art soft computing models for the bearing capacity of strip footings subjected to inclined loading, *Intelligent Systems with Applications*, 2024, **21**, 200314, doi: 10.1016/j.iswa.2023.200314.
[44] P. Kumar, P. Samui, Reliability-based load and resistance factor design of an energy pile with CPT data using machine learning techniques, *Arabian Journal for Science and Engineering*, 2023, 1-30, doi: 10.1007/s13369-023-08253-2.
[45] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Bearing capacity of eccentrically loaded footings on rock masses using soft computing techniques, *Engineered Science*, 2023, **24**, 929, doi: 10.30919/es929.
[46] D. Kumar, V. K. Singh, S. Ali Abed, V. K. Tripathi, S. Gupta, N. Al-Ansari, D. K. Vishwakarma, A. Z. Dewidar, A. A. Al-Othman, M. A. Mattar, Multi-ahead electrical conductivity forecasting of surface water based on machine learning algorithms, *Applied Water Science*, 2023, **13**, 192, doi: 10.1007/s13201-023-02005-1.
[47] S. Markuna, P. Kumar, R. Ali, D. K. Vishwkarma, K. S. Kushwaha, R. Kumar, V. K. Singh, S. Chaudhary, A. Kuriqi, Application of innovative machine learning techniques for long-term rainfall prediction, *Pure and Applied Geophysics*, 2023, **180**, 335-363, doi: 10.1007/s00024-022-03189-4.
[48] A. Elbeltagi, M. Kumar, N. L. Kushwaha, C. B. Pande, P. Ditthakit, D. K. Vishwakarma, A. Subeesh, Drought indicator analysis and forecasting using data driven models: case study in Jaisalmer, India, *Stochastic Environmental Research and Risk Assessment*, 2023, **37**, 113-131, doi: 10.1007/s00477-022-02277-0.
[49] A. Elbeltagi, C. B. Pande, M. Kumar, A. D. Tolche, S. K. Singh, A. Kumar, D. K. Vishwakarma, Prediction of meteorological drought and standardized precipitation index based on the random forest (RF), random tree (RT), and Gaussian process regression (GPR) models, *Environmental Science and Pollution Research*, 2023, **30**, 43183-43202, doi: 10.1007/s11356-023-25221-3.
[50] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Machine learning approaches for prediction of the bearing capacity of ring foundations on rock masses, *Earth Science Informatics*, 2023, **16**, 4153-4168, doi: 10.1007/s12145-023-01152-y.

[51] W. Liu, Z. Liu, Z. Liu, S. Xiong, S. Zhang, Random forest and whale optimization algorithm to predict the invalidation risk of backfilling pipeline, *Mathematics*, 2023, **11**, 1636, doi: 10.3390/math11071636.
[52] Y. Chen, W. Yong, C. Li, J. Zhou, Predicting the thickness of an excavation damaged zone around the roadway using the DA-RF hybrid model, *Computer Modeling in Engineering & Sciences*, 2023, **136**, 2507-2526, doi: 10.32604/cmes.2023.025714.
[53] Y. Chen, Z. Liu, C. Xu, Zhao X., L. Pang, K. Li, Y. Shi, Heavy metal content prediction based on random forest and sparrow search algorithm, *The Journal of Chemometrics*, 2022, **36**, e3445, doi: 10.1002/cem.3445.
[54] J. Kumar, V. B. K. Mohan Rao, Seismic bearing capacity of foundations on slopes, *Géotechnique*, 2003, **53**, 347-361, doi: 10.1680/geot.53.3.347.37282.
[55] J. Kumar, P. Ghosh, Seismic bearing capacity for embedded footings on sloping ground, *Géotechnique*, 2006, **56**, 133-140, doi: 10.1680/geot.2006.56.2.133.
[56] D. Chakraborty, J. Kumar, Seismic bearing capacity of shallow embedded foundations on a sloping ground surface, *International Journal of Geomechanics*, 2015, **15**, 04014035, doi: 10.1061/(asce)gm.1943-5622.0000403.
[57] W. Luo, M. Zhao, Y. Xiao, R. Zhang, W. Peng, Seismic bearing capacity of strip footings on cohesive soil slopes by using adaptive finite element limit analysis, *Advances in Civil Engineering*, 2019, **2019**, 4548202, doi: 10.1155/2019/4548202.
[58] M. Beygi, A. Keshavarz, M. Abbaspour, R. Vali, M. Saberian, J. Li, Finite element limit analysis of the seismic bearing capacity of strip footing adjacent to excavation in c-φ soil, *Geomechanics and Geoengineering*, 2022, **17**, 246-259, doi: 10.1080/17486025.2020.1728396.
[59] M. Beygi, R. Vali, A. Keshavarz, Pseudo-static bearing capacity of strip footing with vertical skirts resting on cohesionless slopes by finite element limit analysis, *Geomechanics and Geoengineering*, 2022, **17**, 485-498, doi: 10.1080/17486025.2020.1794058.
[60] R. Ganesh, J. Kumar, Seismic bearing capacity of strip foundations with nonlinear power-law yield criterion using the stress characteristics method, *Journal of Geotechnical and Geoenvironmental Engineering*, 2022, **148**, 04022083, doi: 10.1061/(asce)gt.1943-5606.0002867.
[61] V. Q. Lai, F. Lai, D. Yang, J. Shiau, W. Yodsomjai, S. Keawsawasvong, Determining seismic bearing capacity of footings embedded in cohesive soil slopes using multivariate adaptive regression splines, *International Journal of Geosynthetics and Ground Engineering*, 2022, **8**, 46, doi: 10.1007/s40891-022-00390-2.
[62] D. R. Kumar, P. Samui, W. Wipulanusat, S. Keawsawasvong, K. Sangjinda, W. Jitchaijaroen, Soft-computing techniques for predicting seismic bearing capacity of strip footings in slopes, *Buildings*, 2023, **13**, 1371, doi: 10.3390/buildings13061371.
[63] R. Butterfield, Dimensional analysis for geotechnical engineers, *Géotechnique*, 1999, **49**, 357-366, doi: 10.1680/geot.1999.49.3.357.
[64] H. Ciria, J. Peraire, J. Bonet, Mesh adaptive computation of upper and lower bounds in limit analysis, *International Journal for Numerical Methods in Engineering*, 2008, **75**, 899-944, doi: 10.1002/nme.2275.
[65] A. Ali, A. V. Lyamin, J. Huang, J. H. Li, M. J. Cassidy, S. W. Sloan, Probabilistic stability assessment using adaptive limit analysis and random fields, *Acta Geotechnica*, 2017, **12**, 937-948, doi: 10.1007/s11440-016-0505-1.
[66] O. OptumG, Copenhagen, Denmark: Optum Computational Engineering, 2020, https//optumce.com/.
[67] L. Breiman, Random Forests, *Machine Learning*, 2001, **45**, 5-32, doi: 10.1023/A:1010933404324.
[68] S. Mirjalili, A. Lewis, The whale optimization algorithm, *Advances in Engineering Software*, 2016, **95**, 51-67, doi: 10.1016/j.advengsoft.2016.01.008.
[69] S. Mirjalili, Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems, *Neural Computing and Applications*, 2016, **27**, 1053-1073, doi: 10.1007/s00521-015-1920-1.
[70] J. Xue, B. Shen, A novel swarm intelligence optimization approach: sparrow search algorithm, *Systems Science & Control Engineering*, 2020, **8**, 22-34, doi: 10.1080/21642583.2019.1708830.
[71] K. E. Taylor, Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres*, 2001, **106**, 7183-7192, doi: 10.1029/2000jd900719.