



OPEN ACCESS

EDITED BY

Wen Nie,
Jiangxi University of Science and
Technology, China

REVIEWED BY

Junlong Sun,
Kunming University of Science and
Technology, China
Mohammad Khajehzadeh,
Islamic Azad University, Anar, Iran

*CORRESPONDENCE

Dan Lu,
✉ 120121555@qq.com

RECEIVED 08 May 2024

ACCEPTED 06 August 2024

PUBLISHED 22 October 2024

CITATION

Huang J, Lu D, Lin W and Yang Q (2024)
Enhancing slope stability prediction through
integrated PCA-SSA-SVM modeling: a case
study of LongLian expressway.
Front. Earth Sci. 12:1429601.
doi: 10.3389/feart.2024.1429601

COPYRIGHT

© 2024 Huang, Lu, Lin and Yang. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

Enhancing slope stability prediction through integrated PCA-SSA-SVM modeling: a case study of LongLian expressway

Jianxin Huang¹, Dan Lu^{2*}, Wei Lin¹ and Qiuyan Yang¹

¹Guangxi Geology and Mineral Construction Group Co., Ltd., Guangxi, China, ²Guangxi Shengfeng Construction Group Co. Ltd., Guangxi, China

China is one of the regions most frequently affected by landslides, which have significant socio-economic impacts. Traditional slope stability analysis methods, such as the limit equilibrium method, limit analysis method, and finite element method, often face limitations due to computational complexity and the need for extensive soil property data. This study proposes a novel approach that combines Principal Component Analysis (PCA), Sparrow Search Algorithm (SSA), and Support Vector Machine (SVM) to improve the accuracy of slope stability prediction. PCA effectively reduces data dimensionality while retaining critical information. SSA optimizes SVM parameters, addressing the limitations of traditional optimization methods. The integrated PCA-SSA-SVM model was applied to a dataset of 257 slope stability samples and validated using five-fold cross-validation to ensure the model's generalization capability. The results show that the model exhibits superior performance in prediction accuracy, precision, recall, and F1-score, with the test set achieving an accuracy of 84.6%, a recall of 84.7%, a precision of 83.1%, and an F1-score of 84.6%. The model's robustness was further validated using slope data from the LongLian Expressway, demonstrating high consistency with the actual stability status. These findings indicate that the PCA-SSA-SVM-based slope stability prediction model has significant potential for practical engineering applications, providing a reliable and efficient tool for slope stability forecasting. Classify the training samples through cross-validation, using the accuracy of cross-validation as the fitness of the sparrow individual. Retain the optimal fitness value and position information.

KEYWORDS

machine learning, side slope stability, PCA, SSA-SVM, prediction

1 Introduction

China is notably one of the regions in Asia, if not the world, that frequently witnesses landslide disasters. According to credible sources (Li et al., 2022; Moayedi et al., 2019; Wei et al., 2021; Xie et al., 2021), between 2011 and 2020, the country experienced over 100,000 geological disasters, of which a staggering 70,000 were landslides, leading to more than 5,000 casualties. The economic impact was profound, resulting in direct losses amounting to 45 billion yuan (Hu et al., 2021). Considering the annual number of geological disasters, it's alarming to

note that landslides consistently constitute over half of these incidents. These landslides, predominantly resulting from slope instability, unleash a multitude of repercussions. They not only cause severe property damage but also result in casualties, traffic disruptions, destruction of homes, hindrances in daily life, and considerable production losses.

To illustrate the scale and implications of such events, let's consider a significant landslide incident from 2019 in Shuicheng County, Liupanshui City, Guizhou Province. This colossal landslide was triggered primarily due to rainfall infiltration causing slope instability. The aftermath was devastating: the surrounding houses crumbled, resulting in numerous casualties. The direct economic ramifications of this single event reached CNY 190 million, with the volume of displaced land estimated to be around 1.8 million m³.

Therefore, accurate analysis and evaluation of slope stability hold significant practical importance. Through effective analysis, treatment, and protection of slopes, casualties and economic losses can be prevented or reduced. Currently, the primary methods used both domestically and internationally include the limit equilibrium method, the limit analysis method, and the finite element method. The limit equilibrium method is one of the first applied to slope stability analysis due to its clear concept and straightforward calculations. [Duncan \(1996\)](#) further analyzed and discussed the influence of different simplified methods and assumptions on the limit equilibrium analysis results of slopes. However, due to its various assumptions, this method has evolved into several different classification methods, typically divided into strict and non-strict segmentation methods. The basic idea of the limit analysis method to solve the slope stability coefficient is to first divide the assumed slip surface into oblique strips, then establish a coordinated velocity field based on the deformation coordination basis, and finally calculate according to the principle that internal energy dissipation equals the external force. [Sloan \(1989\)](#) improved and optimized the lower bound principle finite element method combined with a mathematical programming method to find the lower bound solution of the slope stability safety factor. Although the limit analysis method is widely used in geotechnical engineering, it is challenging to analyze complex shapes and heterogeneous geotechnical engineering cases. Moreover, due to the subjective assumptions made by the limit analysis method, its further application has been significantly limited. As one of the most widely used numerical analysis methods, the finite element method has developed into two main research directions: the finite element strength reduction method and the finite element limit equilibrium method. [Zienkiewicz et al. \(1975\)](#) proposed an alternative analytical approach that eliminates the necessity of presupposing the configuration of the sliding surface. Central to this methodology is the systematic reduction of strength parameters, namely cohesion (c) and internal friction angle (φ). The decrement value of these parameters at the critical juncture is designated as the stability coefficient for the geomaterial mass.

Stability coefficient is one of the important indexes to evaluate whether the slope is unstable. The stability coefficient values obtained by using different slope stability analysis methods under different working conditions need to be systematically verified ([Wang et al., 2022](#); [Zhang W. et al., 2022](#)). For different

types and different conditions of slopes, it is necessary to judge the practicability of these methods. However, it is often computationally intensive, requires specialized software and high-performance computer hardware, and requires detailed soil property data ([Zhang et al., 2023](#)). Recently, machine learning has found efficacious applications across various civil engineering challenges, notably in slope stability evaluation. The stability of mine slopes is influenced by a myriad of factors that exhibit substantial intercorrelation, necessitating heightened precision in predictive modeling ([Xu et al., 2013](#); [Suman et al., 2016](#); [Luan et al., 2023](#); [Lu and Rosenbaum, 2003](#)). Combined with the artificial intelligence algorithm that has emerged in recent years, experts and scholars at home and abroad have proposed many practical models for slope stability prediction research, and have achieved good results. By training a large amount of data, machine learning models can capture the complex relationship between soil and slope characteristics without the need for a clear physical or empirical model. For example, [Gu et al. \(2009\)](#) employed the PCA-GEP algorithm for slope stability prediction and analysis, obtaining favorable outcomes. [Chen et al. \(2014\)](#) utilized PCA in tandem with the BP neural network to anticipate varying types of slope stability, resulting in satisfactory model outcomes. Meanwhile, [Bu et al. \(2009\)](#) introduced a real-coded DE-BP neural network predictive model leveraging the differential evolution algorithm (DE), achieving noteworthy predictive precision. BP neural network models, despite their widespread application in pattern recognition and predictive modeling, exhibit some clear limitations. The primary issue is their propensity to get trapped in local minima, potentially leading to suboptimal model performance. Additionally, BP networks often struggle with overfitting, particularly when dealing with small datasets or a large number of features. They also require significant training time and resources, especially when the network architecture is deep. In the face of these limitations, Support Vector Machine (SVM) models demonstrate their strengths. SVMs enhance classification efficiency by maximizing the margin of the decision boundary, making them more effective in dealing with nonlinear problems and demonstrating superior generalization capabilities when predicting unknown data. The kernel trick in SVMs enables them to efficiently handle high-dimensional data, and they are generally less prone to overfitting, which is particularly valuable for complex pattern recognition challenges. [Zhang et al. \(Bu et al., 2009\)](#), aiming for rapid evaluation of the stability of redbed highway slopes, established a model centered around the SVM algorithm for quick assessment of such slopes, applying it to 16 slopes along the Renmu-Xin Expressway. However, the performance of the SVM model can significantly diminish without proper data preprocessing. [Jin et al. \(Zhang S. et al., 2022\)](#), on the other hand, employed the Sparrow Search Algorithm (SSA) to optimize the Support Vector Machine (SVM), creating an SSA-SVM model for intelligent prediction of slope instability, demonstrating notable advantages in forecasting such events. [Ding et al. \(Jin et al., 2022\)](#) proposed a slope stability prediction model based on Principal Component Analysis (PCA) and Support Vector Machine (SVM), where PCA was used to extract principal components as inputs for SVM training. The results indicated that this method could reduce the dimensionality of input variables and enhance the precision of slope stability prediction in engineering. However,

previous studies, while effective, often faced challenges with high-dimensional datasets and computational demands. Conventional models struggled with multicollinearity among input variables, leading to less accurate predictions and increased computational complexity. Our study introduces an innovative approach that leverages the strengths of three robust techniques: Principal Component Analysis (PCA), the Sparrow Search Algorithm (SSA), and Support Vector Machines (SVM). PCA effectively reduces the number of features in a dataset while retaining crucial information, simplifying subsequent model training and computational demands. SSA, an emerging optimization technique known for its strong global search capabilities and fast convergence, optimizes the parameters of the SVM, addressing limitations of traditional optimization methods used in previous studies. SVM excels in classification and regression tasks, providing superior generalization capabilities compared to other machine learning models. By combining these three approaches, the PCA-SSA-SVM model can effectively process high-dimensional data, improve prediction accuracy through optimal parameter selection, and achieve high-precision predictions. This integrated method is particularly suited for complex data processing tasks that require feature dimensionality reduction, model parameter optimization, and high-precision predictions. Our study introduces a novel hybrid model for slope stability prediction, addressing limitations of previous models. By reducing dimensionality and optimizing model parameters simultaneously, our approach enhances predictive performance and computational efficiency. The application of this model to real-world engineering data, such as the LongLian Expressway slopes, demonstrates its practical utility and effectiveness, achieving a high degree of accuracy and reliability in predictions. In this research, parameters such as rock weight (γ), cohesion (C), internal friction angle (φ), slope height (H), slope angle (β), and pore water pressure (γ_u) are designated as input variables, with the slope safety factor serving as the output variable. PCA is used to reduce the dimensionality of these input variables, selecting fewer and linearly independent factors for data prediction. The SSA-SVM model then trains these new input variables. This methodology presents an innovative avenue for slope stability forecasting, addressing limitations of previous studies and providing a robust, accurate predictive model.

2 Method

2.1 Principal component analysis

Principal Component Analysis (PCA) is a statistical method used to simplify the dimension of the data set while retaining as much variability as possible in the original data. It is widely used in data compression, feature extraction and data visualization.

Step.1 Centralize the data (reduce the mean value of each dimension to 0), as shown in Equation 1.

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}} \tag{1}$$

where \bar{x}_j is the arithmetic mean of x_{ij} , σ_{x_j} is the standard deviation of x_{ij} , $i = 1, 2, \dots, m, j = 1, 2, \dots, n$.

Step.2 Calculate the covariance matrix of the data as shown in Equation 2.

$$\text{cov}(z_i, z_j) = \frac{\sum_{k=1}^n (z_{ik} - z_j)(z_{jk} - z_j)}{n - 1} \tag{2}$$

Step.3 Calculate the eigenvalues and eigenvectors of the covariance matrix, as shown in Equation 3.

$$\| \lambda E - R \| = 0 \tag{3}$$

Step.4 Eigenvalues are arranged in descending order, and the eigenvectors associated with the first k largest eigenvalues are chosen to form a projection matrix. When the cumulative contribution of the current q principal components exceeds 85%, it indicates that they capture a predominant portion of the overall information.

Step.5 This projection matrix is used to transform the original data into a new k -dimensional space. After principal component analysis, the initial variables x_1, x_2, \dots, x_n are transformed into the relationship of n comprehensive index factors y_1, y_2, \dots, y_n , as shown in Equation 4.

$$\begin{cases} y_1 = c_{11}x_1 + c_{12}x_2 + \dots + c_{1n}x_n \\ y_2 = c_{21}x_1 + c_{22}x_2 + \dots + c_{2n}x_n \\ \vdots \\ y_n = c_{n1}x_1 + c_{n2}x_2 + \dots + c_{nn}x_n \end{cases} \tag{4}$$

in the formula, c_{ij} and y_i are not related to each other, and c_{ij} satisfies $c_{i1}^2 + c_{i2}^2 + \dots + c_{in}^2 = 1$. Therefore, the number of initial variables is reduced to achieve the purpose of dimensionality reduction.

2.2 Sparrow search algorithm model

Sparrow Search Algorithm (SSA) is a novel swarm intelligence optimization algorithm proposed in 2020, inspired by the foraging and antipredator behavior of sparrows (Ding et al., 2011; Zhang S. et al., 2022; Jin et al., 2022). SSA is not constrained by the differentiability, derivability, and continuity of the objective function. It boasts strong global search capability, excellent stability, and fast convergence. As a novel and well-organized metaheuristic algorithm, SSA can be employed to solve optimization problems across various fields.

Assuming there are N sparrows in a D -dimensional search space, and the position of the i th sparrow in the D -dimensional search space is denoted as $X_i = [x_{i1}, x_{i2}, x_{id}, \dots, x_{iD}]$. The position of the population X is composed of N sparrows, detailed, as shown in Equation 5.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1D} \\ x_{21} & x_{22} & \dots & x_{2D} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{ND} \end{bmatrix} \tag{5}$$

in the equation, x_{id} represents the position of the i th sparrow in dimension D . Here, the accuracy of slope stability is employed as the fitness function, continuously updating the optimum value to achieve the best recognition rate. The fitness values F_X for all sparrows can be represented as shown in Equation 6.

$$F_X = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_i \end{bmatrix} \quad (6)$$

In SSA, the fitness value F_X represents energy reserves, and f denotes the fitness function. During the search process, producers with higher energy reserves obtain food first. Generally, 10%–20% of the population are producers responsible for finding food, and their foraging search range is larger than that of the predators. At the same time, producers should update through Equation 3.

$$X_{id,t+1} = \begin{cases} X_{id,t} \times \exp\left(\frac{-i}{\alpha T}\right), R_2 < S_T \\ X_{id,t} + Q \times L, R_2 \geq S_T \end{cases} \quad (7)$$

In the formula, t is the current iteration number; $X_{id,t+1}$ is the position of the i th sparrow in dimension d during the $(t+1)$ iteration; T is the maximum number of iterations; α is a uniformly distributed random number, $\alpha \in (0,1]$; Q is a random number following the normal distribution; L is a $1 \times d$ matrix where all elements are 1; R_2 is the alert value, $R_2 \in [0,1]$; S_T is the safety value, $S_T \in [0.5,1]$.

When $R_2 < S_T$, there are no predators around the foraging area, and producers can perform extensive search operations. When $R_2 \geq S_T$, the scout sparrows in the swarm have identified a predator and immediately alert the other sparrows. The sparrows in the swarm then begin antipredatory behaviors, adjusting their search

strategy and quickly moving to a safe area. During the foraging process, apart from the producers, all sparrows act as seekers looking for the best foraging area. The seekers update their position according to Equation 7.

When $i > n/2$, the i th seeker gets no food and is in a state of starvation, with low adaptability. Such a sparrow is likely to fly to another place to forage and gain higher energy. When $i \leq n/2$, the i th seeker finds a random position near the current best position x_b to forage, as shown in Equation 8.

$$X_{id',t'+1} = \begin{cases} Q \times \exp\left(\frac{x_{wd,t} - x_{id,t}}{\alpha T}\right), i > \frac{n}{2} \\ x_{bd,t+1} + |x_{id,t} - x_{bd,t+1}| A^+ L, i \leq \frac{n}{2} \end{cases} \quad (8)$$

In the equation, $x_{wd,t}$ is the worst position of the sparrow in the d^{th} dimension during the t th iteration; $x_{bd,t+1}$ is the best position in the d^{th} dimension during the $(t+1)$ iteration; A is a $1 \times d$ matrix, each of its elements is randomly assigned to 1 or -1 , ensuring $A^+ = AT$.

When danger is detected, the sparrows at the edge of the population will quickly move to a safe area to obtain a better position, while the sparrows in the middle of the population will move randomly to approach other sparrows. The mathematical expression for the movement is, as shown in Equation 9.

$$X_{dd,t+1} = \begin{cases} x_{bdt} + \beta(x_{id,t} - x_{bdt}), f_t \neq f_b \\ x_{id,t} + K \left(\frac{x_{dtd} - x_{wdt}}{|f_i - f_w| + \epsilon} \right), f_t = f_b \end{cases} \quad (9)$$

in the formula, x_{bdt} is the best global position of the alert sparrows during the t th iteration; β is a step length control parameter, a random number from the normal distribution with a mean of 0 and a variance of 1; K is a random number indicating the direction of sparrow's movement, also a step length control parameter, and

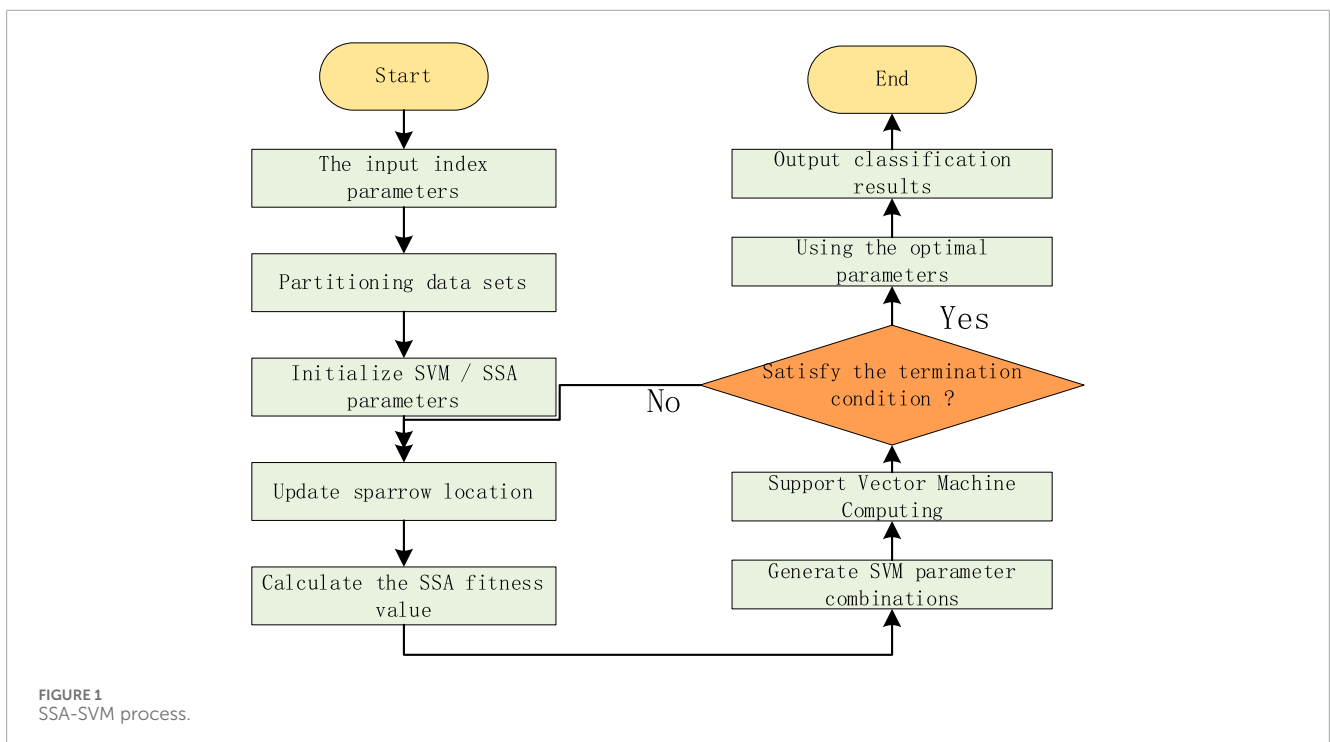


FIGURE 1 SSA-SVM process.

$K \in [-1, 1]$; ϵ is a small constant to avoid division by zero; f_i is the fitness value of the i th sparrow; f_b is the best fitness value of the current sparrow swarm; f_w is the worst fitness value of the current sparrow swarm.

When $f_i = f_b$, the sparrow is at the edge of the population, easily targeted by predators; when $f_i \neq f_b$, the sparrow is in the middle of the population. Once the sparrow is aware of the threat from a predator, it will move closer to other sparrows and adjust its search strategy to avoid being attacked.

2.3 SSA-SVM model

The protagonist of the algorithm is the sparrow, each individual sparrow having only one attribute, which is its position, representing the location of the food it has found. Each sparrow may undergo three states of change: 1) acting as a discoverer, leading the population to forage; 2) being a follower, chasing the discoverer to find food; 3) having a vigilance mechanism, abandoning foraging upon detecting danger. The optimization parameters in the sparrow algorithm are the penalty parameter c and kernel function parameter g in SVM (Support Vector Machine). The fitness function is the SVM's prediction accuracy on the test set.

The selection of SVM penalty factor C and kernel function parameter g greatly influences the classification results. At the same time, the SSA algorithm has strong global search capability and is suitable for optimizing SVM's penalty factor C and kernel function parameter g to obtain a better parameter combination. By using a certain number of sparrows for global optimization, the

optimal parameter combination can be obtained. Then, the optimal C and g obtained by the SSA optimization algorithm are used to establish the SVM identification model and obtain the diagnostic results. The process of optimizing SVM using the SSA algorithm is shown in Figure 1, with specific steps as follows:

First, determine the input and output of the fault diagnosis model. Extract fault features as the input of the diagnostic model and determine the target output values. Establish training and testing sample sets. Specifically, the data was divided into training and testing sets with a ratio of 9:1.

Initialize the related parameters of the sparrow search algorithm, including population size, maximum number of iterations, C , and g . Update the position according to Equations 7–9

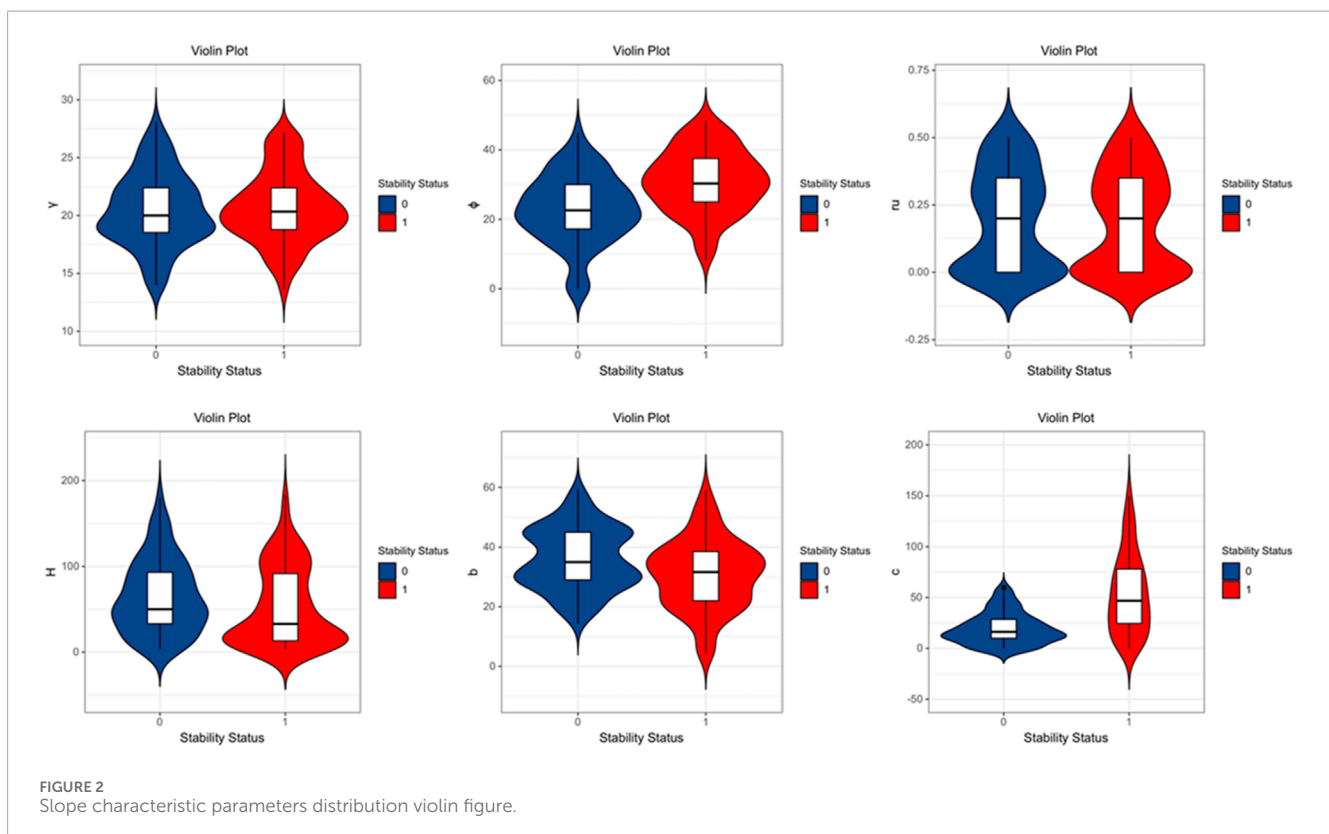
Calculate the fitness value of the sparrow individual's new position, compare the updated fitness value with the original optimal value, and update the global optimal information.

Determine whether the number of iterations meets the termination condition. If not, repeat step (3); otherwise, stop, output the optimal parameters, input the test set samples into the optimal SVM model, and output the diagnostic results.

3 Slope stability prediction model based on PCA-SSA-SVM

3.1 Method principle

For the assessment of slope stability, identifying the most prominent influencing factors is imperative. In this research, slopes with a Factor of Safety (FOS) exceeding 1.3 were classified as stable.



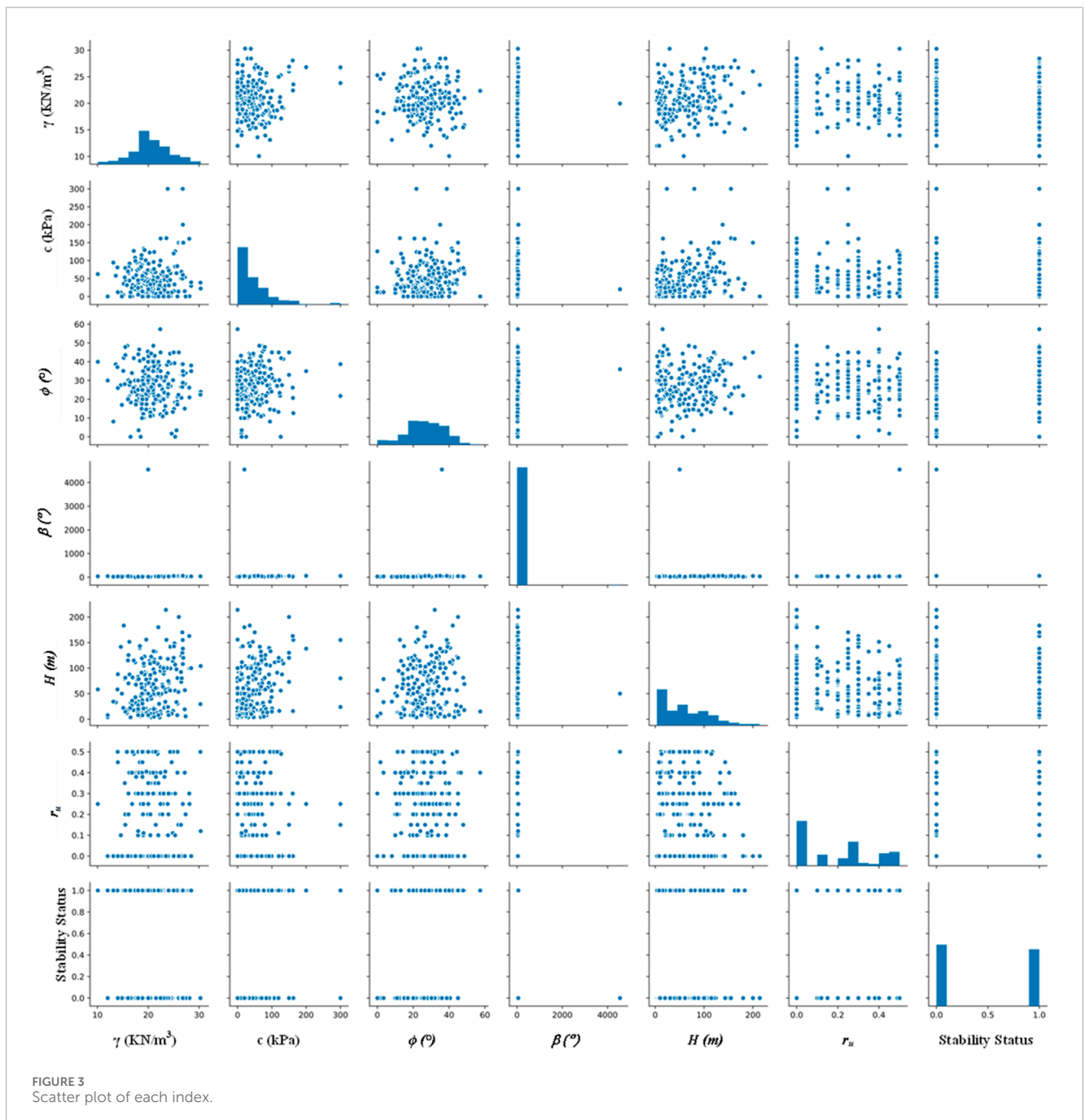
The stability of slopes is categorized into two groups: stable (coded as 1) and unstable (coded as 0). Historically, slope stability prediction efforts have utilized project data for both training and prediction datasets. However, this often results in a limited amount of training data, leading to suboptimal accuracy in the developed models.

To address this limitation, the present study aggregates numerous referenced engineering cases (Emina et al., 2008; Kardani et al., 2021; Zhang R. et al., 2023; Zhang Y. et al., 2023), amassing a total of 257 experimental datasets to evaluate the model's efficacy. The comprehensive dataset is presented in Supplementary Table 1. The slope characteristic parameters vary across different slopes. Visual analysis can

illustrate the characteristic parameter information of the slope in different states and qualitatively analyze the slope's condition to some extent.

In evaluating each characteristic parameter, the violin plot shows the distribution of characteristic parameters under different slope conditions, as shown in Figure 2. Among the different characteristic parameters, the violin shapes of the stable and unstable state data are similar, with few or no outliers, indicating that the dataset is reasonably constructed. The data scatter distribution is delineated in Figure 3.

To further analyze the data, this paper provides descriptive statistics for six indicators, as shown in Table 1. Given that the total



sample size is less than 500, the Shapiro-Wilk (S-W) test is employed to assess the normality of multiple analysis items. The data for the analysis items γ and ϕ follow a normal distribution. Their median and mean values are close, skewness and kurtosis are near zero, and the significance P -values are 0.051 and 0.052, respectively, which are

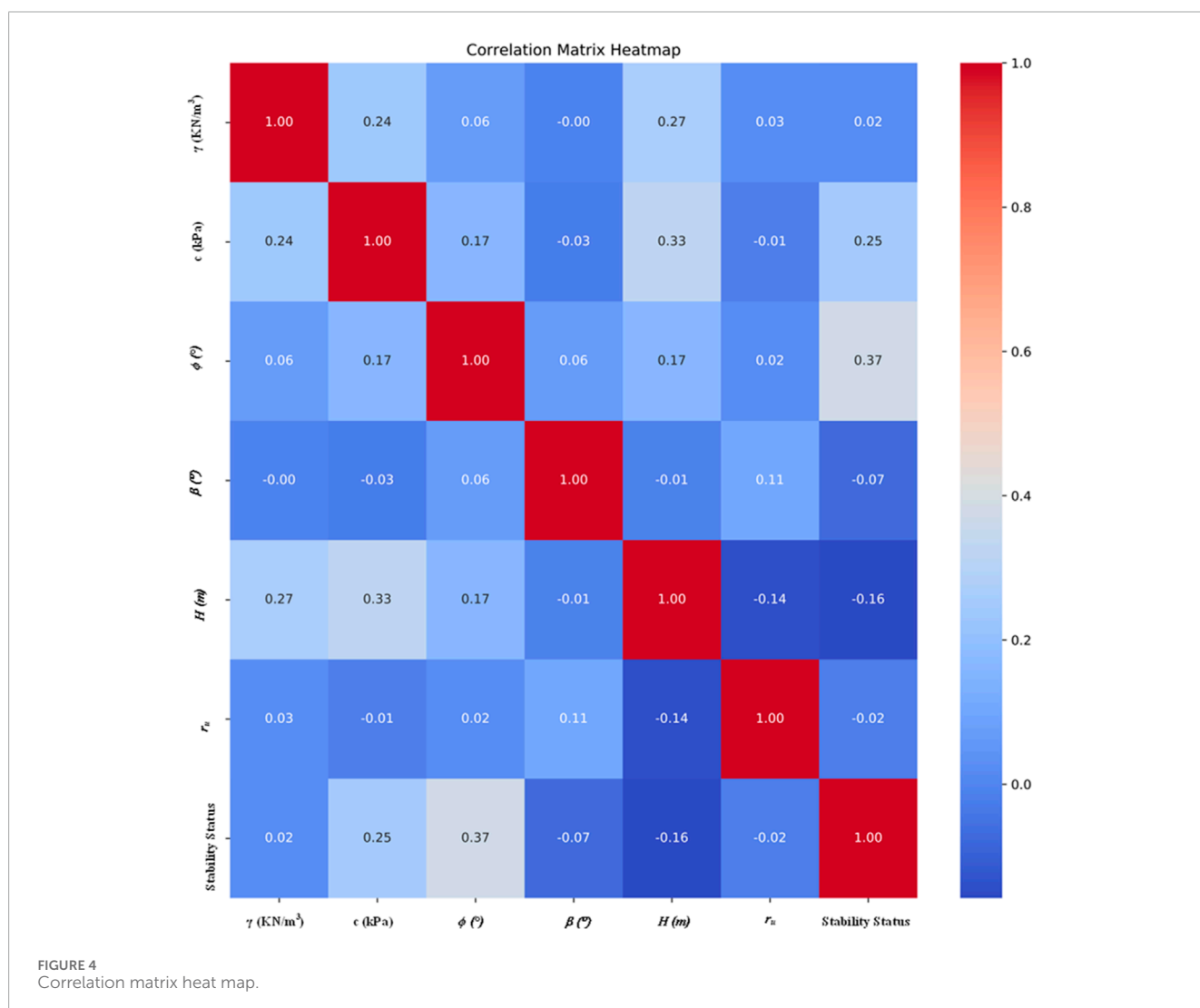
higher than the commonly used 0.05 significance level, so the null hypothesis of normality cannot be rejected (Sah et al., 1994).

However, the data for c , β , H , and γ_u do not follow a normal distribution. Even though the kurtosis and skewness of H are within the common normal distribution range, the significance P -value

TABLE 1 Statistical description of samples.

| Name | Sample size | Median | Average value | Standard deviation | Skewness | Kurtosis | S-W test |
|------------|-------------|--------|---------------|--------------------|----------|----------|-------------------|
| γ | 257 | 20.31 | 20.588 | 3.593 | 0.117 | 0.164 | 0.989 (0.051 *) |
| c | 257 | 30.81 | 46.172 | 48.228 | 2.289 | 7.967 | 0.791 (0.000 ***) |
| ϕ | 257 | 26.81 | 26.732 | 10.565 | -0.218 | 0.085 | 0.989 (0.052 *) |
| β | 257 | 33.33 | 51.059 | 281.957 | 15.993 | 256.174 | 0.052 (0.000 ***) |
| H | 257 | 50 | 60.803 | 46.798 | 0.793 | 0.015 | 0.922 (0.000 ***) |
| γ_u | 257 | 0.2 | 0.194 | 0.181 | 0.261 | -1.395 | 0.852 (0.000 ***) |

Note: *** represents a significance level of 1%. ** represents a significance level of 5%. * represents a significance level of 10%.



remains at 0.000, clearly indicating a departure from normality. Among the six analysis items, two meet the criteria for a normal distribution, while four do not.

Variable correlation analysis is an important part of statistics, which is mainly used to explore whether there is a relationship between two or more variables. In order to prevent data redundancy in the model, which affects the prediction accuracy of the model, the correlation coefficient between the two variables in the data set is calculated to determine whether there are redundant characteristic variables and determine the correlation between variables, see Figure 4. According to the correlation analysis table, it can be observed that there is a positive correlation between γ and c and H , among which the correlation with H is the most significant. c mainly has a strong positive correlation with H , and also has a certain degree of positive correlation with ϕ . ϕ shows positive correlation with c and E . The relationship with all other variables is relatively weak, but the positive correlation with ϕ is slightly more obvious. H not only has a significant positive correlation with γ and c , but also has a certain negative correlation with γu . In general, the correlation between H and several other variables is the most prominent, especially the positive correlation with γ , c and the negative correlation with γu ; the correlation with all variables is relatively weak.

3.2 PCA analysis

In constructing the slope stability prediction model, six key factors related to slope stability status are incorporated. Given the multidimensionality of data derived from these factors, model training becomes challenging. To address this, Principal Component Analysis (PCA) is applied to transform this high-dimensional dataset into a reduced-dimensional space while preserving the original data's information content as much as possible. This approach helps isolate the most significant principal components, thereby reducing model complexity and enhancing training efficacy. The procedural steps for PCA analysis, as detailed in Section 2.2, are outlined below:

Step 1. Suitability Assessment: Initially, both the Kaiser-Meyer-Olkin (KMO) and Bartlett's tests are performed to determine the data's suitability for PCA, as shown in Table 2. A KMO value exceeding 0.8 is optimal, while values between 0.7 and 0.8 are considered moderate. KMO values below 0.6 may be deemed unsuitable. The KMO test results revealed a KMO value of 0.604.

TABLE 2 PCA feasibility test.

| KMO test and bartlett test | | |
|----------------------------|------------------------|-----------|
| KMO value | | 0.604 |
| Bartlett sphericity test | Approximate chi-square | 75.447 |
| | df | 15 |
| | P | 0.000 *** |

Note: *** represents a significance level of 1%. ** represents a significance level of 5%. * represents a significance level of 10%.

Additionally, Bartlett's sphericity test results showed a significant P -value of 0.000 ***, which is significant at the respective level. This rejects the null hypothesis, indicating correlations among the variables and affirming the appropriateness of proceeding with PCA.

Step 2. Determining the Number of Principal Components: To ascertain the required number of principal components, examine the variance elucidation table and the scree plot. Table 3 lists the variance contribution rate of each principal component to the original variables, which helps in selecting the number of principal components by considering the descending trajectory of eigenvalues. During PCA, the cumulative variance elucidation rate typically approaches 90%. Based on the information in the variance elucidation table, five principal components are identified as input parameters for the SSA-SVM model.

Step 3. Analysis of principal component load coefficient: Table 4 load coefficient combined to reveal the importance of hidden variables behind each principal component.

Step 4. Spatial Representation of Principal Components: The dimensions of numerous principal components are condensed to two or three principal components, and their spatial orientation is depicted using quadrant diagrams. Extracting two or three principal components allows for a clearer visualization of their spatial interrelations. However, since this study extracts five principal components, their spatial distribution cannot be effectively illustrated using the primary mapping technique.

Step 5. Constructing the Principal Component Equation 10: Based on the component matrix, the component formula for each principal component is formulated and its weight is determined, as shown in Table 5.

$$\begin{cases} y_1 = 0.367 \times \gamma + 0.435 \times c + 0.259 \times \phi - 0.015 \times \beta + 0.456 \times H - 0.091 \times r_u \\ y_2 = 0.079 \times \gamma + 0.029 \times c + 0.305 \times \phi + 0.604 \times \beta - 0.119 \times H + 0.629 \times r_u \\ y_3 = 0.522 \times \gamma + 0.108 \times c - 0.575 \times \phi - 0.363 \times \beta - 0.101 \times H + 0.537 \times r_u \\ y_4 = 0.254 \times \gamma - 0.109 \times c - 0.614 \times \phi + 0.705 \times \beta + 0.199 \times H + 0.368 \times r_u \\ y_5 = -0.73 \times \gamma + 0.770 \times c - 0.413 \times \phi + 0.088 \times \beta + 0.129 \times H + 0.196 \times r_u \\ y = 0.306 \times y_1 + 0.212 \times y_2 + 0.178 \times y_3 + 0.165 \times y_4 + 0.37 \times y_5 \end{cases} \quad (10)$$

Step 6. Comprehensive score output: Finally, calculate and output the comprehensive score based on PCA and the principal component data are shown in Supplementary Table 2.

3.3 PCA-SSA-SVM model construction and training

Utilizing the SSA-SVM approach, an slope stability prediction model is developed. Python software is chosen for scripting the SSA algorithm to enhance the SVM model. After applying principal component analysis (PCA), the initial six evaluation indicators for slope stability are substituted by five principal components. The output variable denotes the safety condition of the slope.

3.4 Model evaluation indicators

The accuracy (ACC), recall rate (TPR), precision rate (PPV) and F1-score were used to evaluate the prediction effect of

TABLE 3 Variance interpretation form.

| Component | Characteristic root | Explanatory rate of variance (%) | Cumulative variance interpretation rate (%) |
|-----------|---------------------|----------------------------------|---|
| 1 | 1.645 | 27.418 | 27.418 |
| 2 | 1.142 | 19.03 | 46.449 |
| 3 | 0.957 | 15.949 | 62.398 |
| 4 | 0.889 | 14.818 | 77.216 |
| 5 | 0.736 | 12.264 | 89.48 |
| 6 | 0.631 | 10.52 | 100 |

TABLE 4 Factor load coefficient table.

| Name | Principal component_1 | Principal component_2 | Principal component_3 | Principal component_4 | Principal component_5 | Common degree |
|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|---------------|
| γ | 0.337 | 0.232 | 0.35 | -0.338 | 0.246 | 0.961 |
| c | -0.253 | 0.144 | -0.078 | 0.65 | 0.381 | 0.855 |
| φ | 0.045 | -0.153 | 0.053 | -0.251 | 0.411 | 0.996 |
| β | 0.305 | 0.287 | 0.108 | 0.528 | 0.483 | 0.994 |
| H | 0.751 | 0.002 | 0.022 | -0.047 | -0.03 | 0.631 |
| γ _u | -0.1 | 0.302 | -0.542 | -0.356 | 0.312 | 0.931 |

TABLE 5 Composition matrix table.

| Name | Principal component_1 | Principal component_2 | Principal component_3 | Principal component_4 | Principal component_5 |
|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| γ | 0.367 | 0.079 | 0.522 | 0.254 | -0.73 |
| c | 0.435 | 0.029 | 0.108 | -0.109 | 0.77 |
| φ | 0.259 | 0.305 | -0.575 | -0.614 | -0.413 |
| β | -0.015 | 0.604 | -0.363 | 0.705 | 0.088 |
| H | 0.456 | -0.119 | -0.101 | 0.199 | 0.129 |
| γ _u | -0.091 | 0.629 | 0.537 | -0.368 | 0.196 |

Consequently, an aggregate score for each principal component can be derived. The novel variable formulations for the five principal components are delineated as.

PCA-SSA-SVM. The calculation formula is, as shown in Equations 11–13 (Khajehzadeh et al., 2022):

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

$$TPR = \frac{TP}{(TP + FN)} \tag{12}$$

$$PPV = \frac{TP}{(TP + FP)} \tag{13}$$

TP: True positives, where actual positive samples are correctly predicted as positive; TN: True negatives, where actual negative samples are correctly predicted as negative; FN: False negatives, where actual positive samples are incorrectly predicted as negative; FP: False positives, where actual negative samples are incorrectly predicted as positive.

The F1-score, also referred to as the F-measure, serves as a comprehensive metric that evaluates the model's precision (PPV) and recall (TPR) as shown in Equation 14. It is computed as the harmonic mean of precision and recall, with values spanning

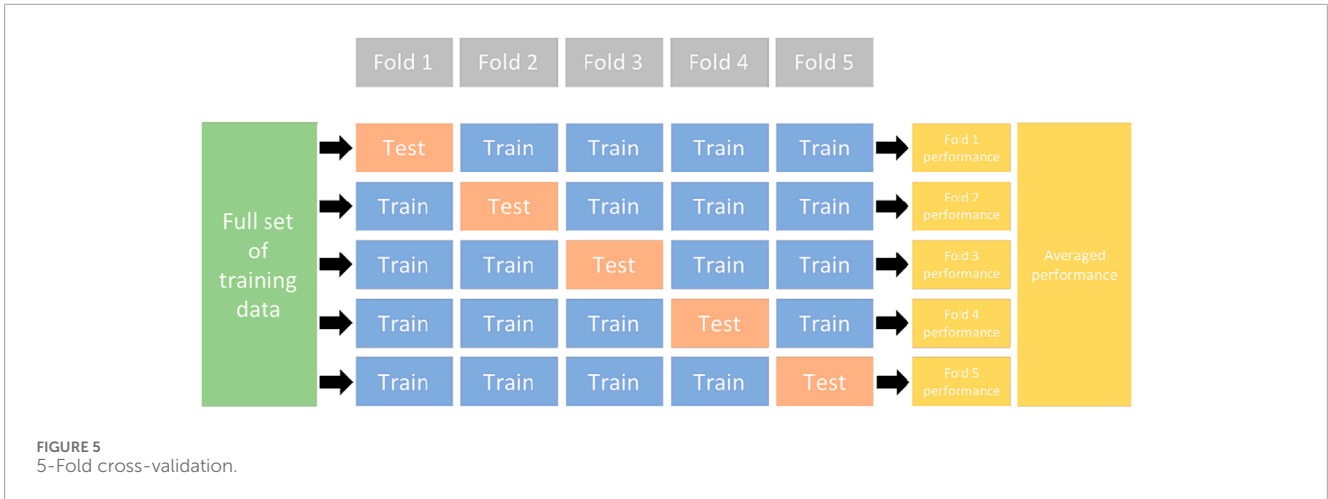


FIGURE 5 5-Fold cross-validation.

TABLE 6 Model prediction evaluation.

| Name | Accuracy rate | Recall rate | Precision rate | F1 |
|--------------|---------------|-------------|----------------|-------|
| Training set | 0.952 | 0.922 | 0.981 | 0.951 |
| Test set | 0.846 | 0.847 | 0.831 | 0.846 |

TABLE 7 Comparison of model predictions using 5-fold cross-validation.

| Name | Accuracy rate | Recall rate | Precision rate | F1 |
|-------------|---------------|-------------|----------------|-------|
| SVM | 0.654 | 0.571 | 0.727 | 0.64 |
| SSA-SVM | 0.731 | 0.801 | 0.75 | 0.772 |
| PSO-SVM | 0.825 | 0.753 | 0.781 | 0.832 |
| GA-SVM | 0.826 | 0.71 | 0.802 | 0.786 |
| PCA-SSA-SVM | 0.846 | 0.847 | 0.831 | 0.846 |

between 0 and 1. A score of 1 indicates impeccable accuracy, while a score of 0 denotes the poorest accuracy.

$$F_{1-score} = 2 \frac{PPV \cdot TPR}{PPV + TPR} \tag{14}$$

3.5 Cross-validation

To further enhance the model's generalization ability and prevent overfitting during the training process, this study employs 5-fold cross-validation, as illustrated in Figure 5. Initially, the original training dataset D is randomly divided into k equally sized subsets: D1, D2, D3, D4, and D5. Let D_j and $D_{-j} = D/D_j$ denote the j-th

test set and the corresponding training set, respectively, for each iteration.

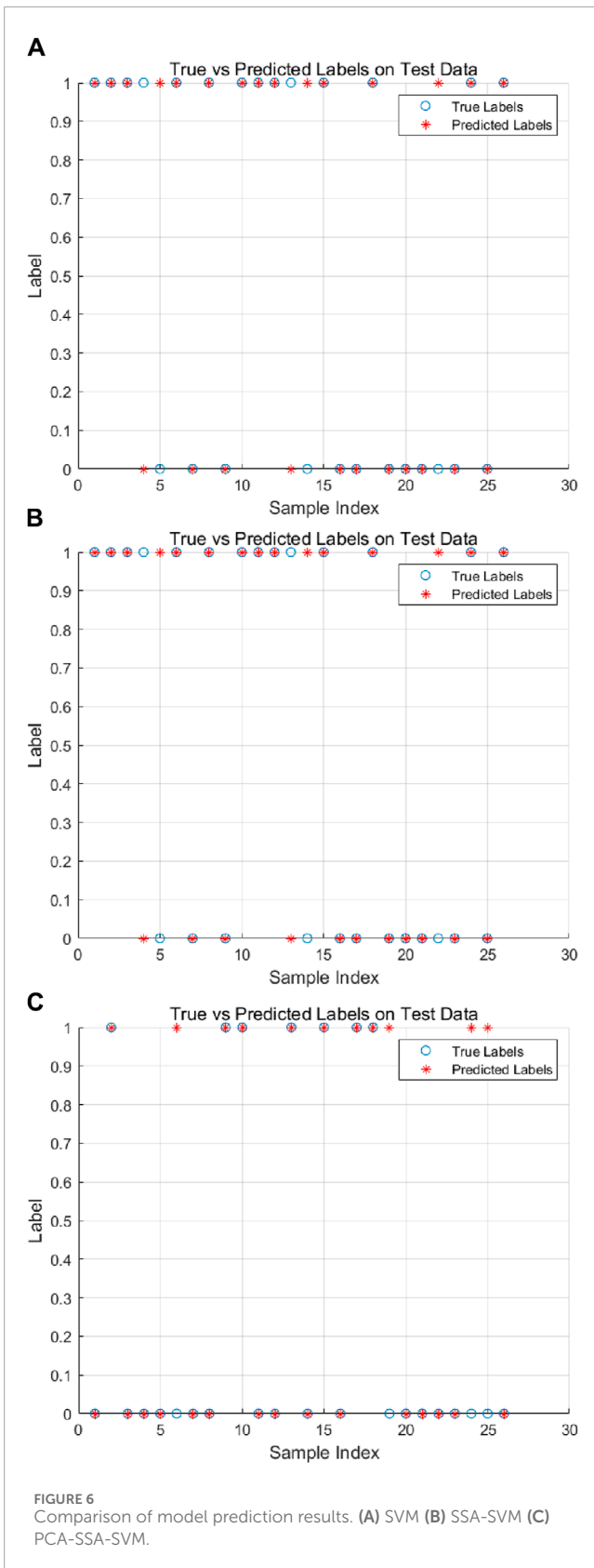
3.6 Prediction results and analysis

During the data preprocessing phase, the initial dataset was randomly permuted to create a modified dataset. To enhance the model's generalizability, 231 data entries were designated as the training set, while the remaining entries were reserved for testing. Utilizing the PCA-SSA-SVM approach, a comprehensive prediction was executed on the 257 slope stability data entries. Combining 5-fold cross-validation with model parameter optimization, the results show that the SSA-SVM model's performance under different parameters indicates that the optimal value for the penalty coefficient (C) in the SVM model is 9.426, and the optimal value for the kernel parameter (gamma) is 0.03.

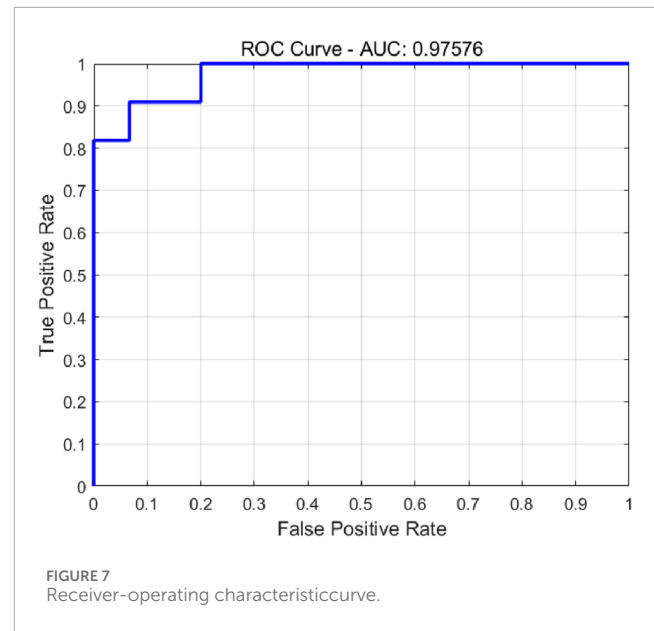
The exhaustive outcomes are presented in Table 6. The prediction outcomes from the test set are compared with the actual results, with a subsequent analysis conducted in conjunction with the predictions for each state, as depicted in Figure 6.

Based on the data presented in Table 7, the PCA-SSA-SVM model exhibits the highest accuracy, recall, precision, and F1 score compared to the other models. Its accuracy rate is notably high at 84.6%, its recall rate is excellent at 84.7%, and its precision rate is impressive at 83.1%. The SSA-SVM model, with an accuracy rate of 73.1%, a recall rate of 80.1%, and a precision rate of 75%, performs better than the simple SVM model across all metrics, suggesting that incorporating the SSA algorithm improves performance. The PSO-SVM model shows a strong performance with an accuracy rate of 82.5%, a recall rate of 75.3%, and a precision rate of 78.1%, while the GA-SVM model has an accuracy rate of 82.6%, a recall rate of 71%, and a precision rate of 80.2%. Both the PSO-SVM and GA-SVM models exhibit better performance than the SVM model but are slightly lower than the PCA-SSA-SVM model in accuracy and F1 score.

Based on the provided metrics, combining PCA with the SSA algorithm yields the best results among all the models. This



integrated approach significantly enhances prediction accuracy and effectiveness, showcasing its superiority in handling high-dimensional data and optimizing model performance. The PCA-SSA-SVM model's robustness and high precision make



it an ideal choice for slope stability prediction in engineering applications.

In order to comprehensively evaluate the performance of the model, this study drew the ROC (receiver operating characteristic) curve and calculated the AUC (area under the curve) value, see Figure 7. The ROC curve is an effective tool for visually demonstrating the relationship between the true positive rate and the false positive rate of the model under various thresholds. The AUC value is the area under the ROC curve, which provides us with a single quantitative indicator of model performance. The AUC value is usually between 0.5 (no discrimination) and 1.0 (perfect discrimination). The model of this study obtained an AUC value of 0.9758, which strongly indicates that the PCA-SSA-SVM model performs well in the slope stability prediction task.

4 Engineering verification

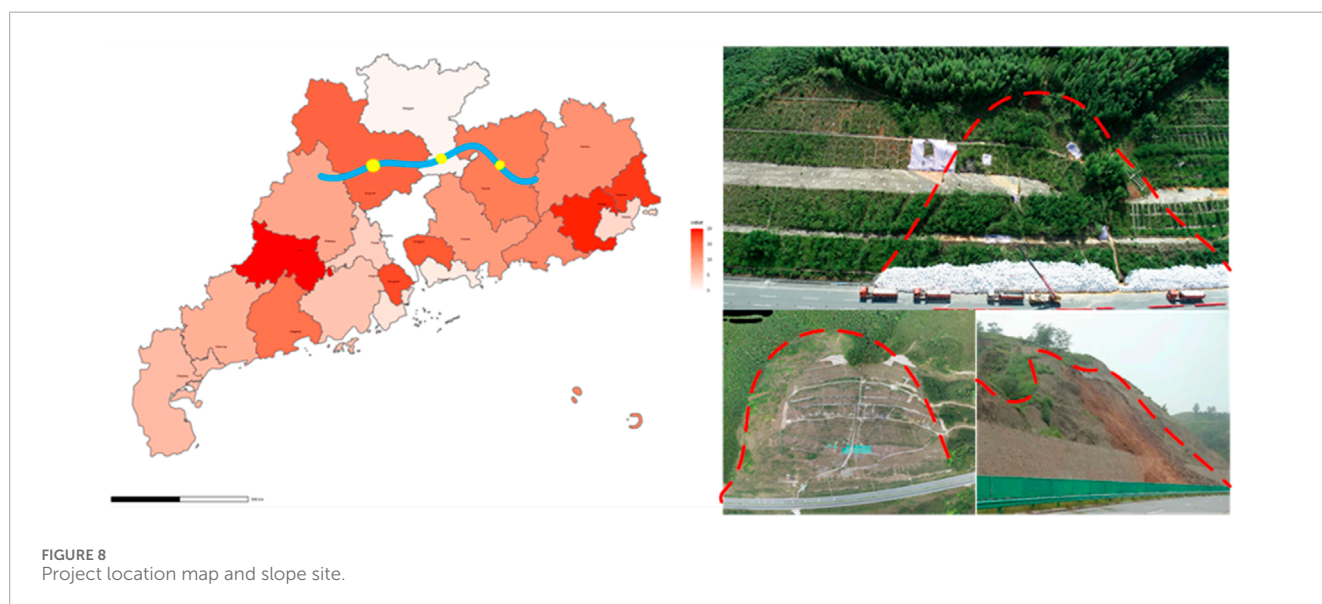
4.1 Engineering background

The Guangdong Province's Longchuan to Huaiji Expressway in China is a part of the original national highway network "7,918" plan, specifically marked as the 17th cross route. A significant portion of this route is the Longchuan to Lianping section of the Guangdong Province Longchuan to Huaiji Expressway, commonly referred to as the "LongLian Expressway". This project is located in the mountainous terrain of northern Guangdong, crossing through Heyuan and Shaoguan cities, encompassing four counties and thirteen towns.

The starting point of the LongLian Expressway is at K206 + 222 in the Old Long Town of Longchuan County, where it connects with the already operational Meihe Expressway. The endpoint is at K334 + 200 in the Tangxia Village, Longxian Town, Wengyuan County, Shaoguan City. The total length of the route is 127.978 km,

TABLE 8 Slope data of LongLian Expressway.

| No. | γ (KN/m ³) | c (kPa) | ϕ (°) | (°) | H (m) | γ_u | Stability status |
|-----|-------------------------------|---------|------------|-------|--------|------------|------------------|
| 1 | 17.93 | 78.20 | 18.49 | 33.42 | 120.79 | 0.0 | 1 |
| 2 | 18.02 | 40.92 | 21.18 | 21.86 | 34.56 | 0.1 | 0 |
| 3 | 25.76 | 64.11 | 21.40 | 15.76 | 30.38 | 0.5 | 0 |
| 4 | 21.03 | 21.23 | 17.72 | 5.79 | 57.31 | 0.0 | 1 |
| 5 | 25.74 | 31.42 | 17.23 | 30.03 | 80.53 | 0.4 | 1 |
| 6 | 18.80 | 14.40 | 25.02 | 19.98 | 30.06 | 0.0 | 0 |
| 7 | 18.77 | 30.01 | 9.99 | 25.02 | 50.00 | 0.1 | 1 |
| 8 | 19.97 | 19.96 | 36.00 | 45.50 | 50.00 | 0.5 | 1 |
| 9 | 26.70 | 50.00 | 26.60 | 50.00 | 34.04 | 0.2 | 1 |
| 10 | 26.80 | 90.00 | 28.80 | 59.00 | 42.37 | 0.3 | 1 |



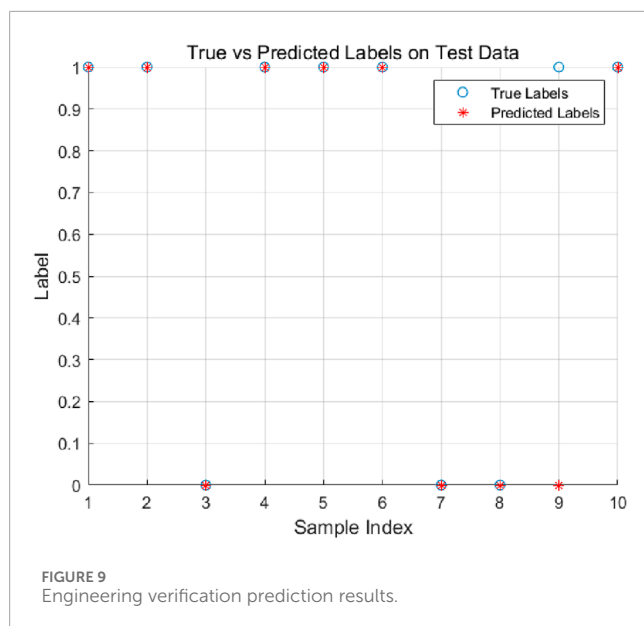
with a designed speed of 100 km/h, adhering to the technical standards of a two-way four-lane expressway, as shown in Figure 8.

4.2 Slope profile

4.2.1 Stratigraphic lithology

The surface layer consists of silty clay and sandy clay soil, with plastic to hard plastic consistency. The underlying strata comprise entirely to strongly weathered andesite porphyry and moderately weathered diorite. The fully weathered andesite porphyry is yellow-brown in color, showing complete weathering with substantial disruption of the original rock structure. The core of the rock manifests as hard soil columns, with a soft rock quality that tends to soften and disintegrate when in contact with water,

locally interspersed with isolated stones. The strongly weathered andesite porphyry appears yellow, yellow-gray, or yellow-brown, with a major portion of the original rock structure destroyed, exhibiting a medium-grain structure and blocky construction. The mineral composition includes quartz, feldspar, and mica, with the rock core primarily displaying fragmentary and blocky forms, block diameters ranging from 2 to 6 cm, and a minor amount appearing as debris. Joints and fissures are well-developed with some fissure surfaces being impregnated with iron-manganese material, displaying uneven weathering, soft rock quality, and fragmented rock mass. The moderately weathered diorite is gray with a medium to coarse-grain structure and blocky construction. Its mineral composition also includes quartz, feldspar, and mica, with well-developed joints and fissures. The rock core manifests as columnar or short columnar structures, with joint lengths ranging from 10 to



70 cm, producing a brittle sound when struck, indicating hard rock quality and a relatively intact rock mass.

4.2.2 Hydrogeological overview

Surface Water: The surface does not exhibit perennial water flow; transient surface runoff only occurs post-rainfall during the rainy season.

Groundwater: The groundwater is composed of upper soil layer pore water and deep bedrock fissure water, with a relatively small water content. The primary source of replenishment relies on the infiltration of atmospheric precipitation, with the discharge base level being the streams at the bottom of the slopes. During the period of this survey, no groundwater levels were detected within the drilled depths.

4.3 Slope data

There are 10 large and small slopes in LongLian Expressway, including 7 unstable slope samples and 3 stable slope samples, as shown in the Table 8.

4.4 Verification based on PCA-SSA-SVM model

Utilizing the well-trained PCA-SSA-SVM model for slope stability prediction, the results are illustrated in Figure 9. From the provided data, it can be observed that the PCA-SSA-SVM model predictions align with the actual stability status data of the LongLian Expressway slopes at most points. Specifically, among the ten data points, the PCA-SSA-SVM model predictions match the actual stability status at nine points. Only at one data point does the model prediction deviate from the actual stability status.

5 Discussion

While the PCA-SSA-SVM model demonstrated high accuracy and robustness, several limitations must be addressed to enhance its applicability and performance in real-world scenarios.

- Model Limitations:** Although the PCA-SSA-SVM model showed excellent performance on the test set, the study does not deeply explore potential limitations in practical applications. The presence of nonlinear relationships between features or potential outliers in the slope dataset may affect the model's predictive performance. Furthermore, the model's adaptability to slope data under varying geological conditions or climatic environments requires further investigation.
- Feature Selection and Optimization:** While PCA effectively reduces the dimensionality of the data, the specific rationale for selecting the six features used in this study was not elaborated. Future research should explore feature engineering and optimization techniques to further enhance model predictive performance. Introducing additional features related to slope stability could enrich the model inputs and improve accuracy.
- Model Comparison and Validation:** Although comparisons with SVM and SSA-SVM models were presented, a comparative analysis with other advanced machine learning methods was not included. Future research could investigate the performance of models such as Random Forest, Gradient Boosting Trees, or Deep Learning techniques on the same dataset. Such comparisons would help establish the superiority of the PCA-SSA-SVM model and provide a more comprehensive evaluation of its performance.

6 Conclusion

Within this study, the PCA methodology is employed to examine and process the six evaluative indices affecting slope stability across 257 data samples. From these, five principal components are derived, serving as input variables for the SSA-SVM. Subsequently, a PCA-SSA-SVM predictive model for slope stability is formulated. The accuracy of the model is further evaluated by engineering verification. A detailed analysis of the model's predictive outcomes and associated errors yielded the subsequent findings.

- The PCA methodology adeptly addresses the multicollinearity challenge inherent among factors influencing slope stability, streamlining the model's input variables and enhancing simulation efficiency.
- The model's precision, robustness, and adaptability were gauged via metrics such as precision, recall, and the F1-score. The test set outcomes were 84.6%, 84.7%, and 84.6%, respectively. With an AUC value of 0.9758, the model's accuracy and adaptability are deemed commendable.
- The accuracy of the PCA-SSA-SVM model was validated using the slope data from the LongLian Expressway. Among the 10 provided samples, only the prediction for sample 1 was

incorrect. The results indicate that the PCA-SSA-SVM based slope stability prediction model can be effectively applied in engineering practice to achieve slope stability forecasting.

- The model's classification performance has not taken into account factors such as the slope's geometric shape, soil quality of the slope body, and external influencing elements. Further research is required in the future, especially when using machine learning algorithms to estimate slope stability, to continue refining and enhancing feature parameters.

Data availability statement

The original contributions presented in the study are included in the article/[supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

JH: Investigation, Methodology, Software, Writing—original draft. DL: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Resources, Writing—original draft, Writing—review and editing. WL: Writing—original draft. QY: Writing—review and editing.

References

- Bu, N., Zhao, H., and Xie, J. (2009). Slope stability analysis with DE-BP neural network. *Subgr. Eng.* (4), 2.
- Chen, J., Zheng, R., and Chen, H. (2014). Slope stability analysis based on PCA and BP neural network. *China Saf. Sci. J.* 10 (5), 6.
- Ding, H., Zhu, J., and Luo, S. (2011). Research on slope stability prediction model based on PCA-SVM. *Subgr. Eng.* (2), 3.
- Duncan, J. M. (1996). State of the art: limit equilibrium and finite-element analysis of slopes. *J. Geotechnical Eng.* 123 (7), 577–596. doi:10.1061/(asce)0733-9410(1996)122:7(577)
- Emina, E., Torlakovic, Driman, D. K., Parfitt, J. R., Wang, C., Benerjee, T., et al. (2008). Sessile serrated adenoma (SSA) vs. traditional serrated adenoma (TSA). *Am. J. Surg. Pathology* 32, 21–29. doi:10.1097/pas.0b013e318157f002
- Gu, Q., Cai, Z. H., Zhu, L., and Huang, B. (2009). Slope stability prediction based on PCA-GEP algorithm. *Rock Soil Mech.* 30 (3), 757–761 + 768. doi:10.3969/j.issn.1000-7598.2009.03.033
- Hu, J., Cheng, P., and Liu, M. M. (2021). Numerical modeling of 3D slopes with weak zones by random field and finite elements. *Appl. Sci.* 11, 9852. doi:10.3390/app11219852
- Jin, A., Zhang, J., Sun, H., and Wang, B. (2022). An intelligent prediction and early warning model for slope instability based on SSA-SVM. *J. Huazhong Univ. Sci. Technol. Nat. Sci. Ed.* 50 (11), 142–148. doi:10.13245/j.hust.221118
- Kardani, N., Zhou, A., Nazem, M., and Shen, S. L. (2021). Improved prediction of slope stability using a hybrid stacking ensemble method based on finite element analysis and field data. *J. Rock Mech. Geotechnical Eng.* 13 (1), 188–201. doi:10.1016/j.jrmge.2020.05.011
- Khajehzadeh, M., Taha, M. R., Keawsawong, S., Mirzaei, H., and Jebeli, M. (2022). An effective artificial intelligence approach for slope stability evaluation. *IEEE Access* 10, 5660–5671. doi:10.1109/access.2022.3141432
- Li, N., Wang, Y., and Ma, W. (2022). A wind power prediction method based on DE-BP neural network[J]. *Front. energy res.* 10, 844111. doi:10.3389/feart.2022.844111
- Lu, P., and Rosenbaum, M. (2003). Artificial neural networks and grey systems for the prediction of slope stability. *Nat. Hazards* 30 (3), 383–398. doi:10.1023/b:nhaz.0000007168.00673.27
- Luan, B., Zhou, W., Jiskani, I. M., Lu, X., and Wang, Z. (2023). Slope stability prediction method based on intelligent optimization and machine learning algorithms. *Sustainability* 15, 1169. doi:10.3390/su15021169
- Moayedi, H., Bui, D. T., Kalantar, B., and Kok Foong, L. (2019). Machine-learning-based classification approaches toward recognizing slope stability failure. *Appl. Sci.* 9 (21), 4638. doi:10.3390/app9214638
- Sah, N., Sheorey, P., and Upadhyaya, L. (1994). Maximum likelihood estimation of slope stability. *Int. J. Rock Mech. Min. Sci. and Geomechanics Abstr.* 31 (1), 47–53. doi:10.1016/0148-9062(94)92314-0
- Sloan, S. W. (1989). Upper bound limit analysis using finite elements and linear programming. *Int. J. Numer. Anal. Methods Geomechanics* 13 (3), 263–282. doi:10.1002/nag.1610130304
- Suman, S., Khan, S. Z., Das, S. K., and Chand, S. K. (2016). Slope stability analysis using artificial intelligence techniques. *Nat. Hazards* 84 (2), 727–748. doi:10.1007/s11069-016-2454-2
- Wang, G. J., Zhao, B., Wu, B., Zhang, C., and Liu, W. (2022). Intelligent prediction of slope stability based on visual exploratory data analysis of 77 *in situ* cases. *Int. J. Min. Sci. Technol.* doi:10.1016/j.ijmst.2022.07.002
- Wei, W., Li, X., Liu, J., Zhou, Y., and Zhou, J. (2021). Performance evaluation of hybrid WOA-SVR and HHO-svr models with various kernels to predict factor of safety for circular failure slope. *Appl. Sci.* 11 (4), 1922. doi:10.3390/app11041922
- Xie, W., Nie, W., Saffari, P., Robledo, L. F., Descote, P. Y., and Jian, W. (2021). Landslide hazard assessment based on Bayesian optimization–support vector machine in Nanping City, China. *Nat. Hazards* 109 (1), 931–948. doi:10.1007/s11069-021-04862-y
- Xu, C., Xu, X., Dai, F., Wu, Z., He, H., Shi, F., et al. (2013). Application of an incomplete landslide inventory, logistic regression model and its validation for landslide susceptibility mapping related to the may 12, 2008 wenchuan earthquake of China. *Nat. Hazards* 68 (2), 883–900. doi:10.1007/s11069-013-0661-7

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

Authors JH, WL, and QY were employed by Guangxi Geology and Mineral Construction Group Co., Ltd. Author DL was employed by Guangxi Shengfeng Construction Group Co. Ltd.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feart.2024.1429601/full#supplementary-material>

Zhang, H., Wu, W., Zhang, X., Han, L., and Zhang, Z. (2023). Slope stability prediction method based on the margin distance minimization selective ensemble. *Catena* 2023-08-20. doi:10.1016/j.catena.2022.106055

Zhang, R., Su, J., and Feng, J. (2023a). An extreme learning machine model based on adaptive multi-fusion chaotic sparrow search algorithm for regression and classification. *Evol. Intell.* 17, 1567–1586. doi:10.1007/s12065-023-00852-0

Zhang, S., Zheng, D., and Zhang, W. (2022b). A rapid evaluation method for the stability of red-layer highway slopes based on the SVM algorithm. *Northwest Hydropower*, (000-003). doi:10.3969/j.issn.1006-2610.2022.03.001

Zhang, W., Li, H., Han, L., Chen, L., and Wang, L. (2022a). Slope stability prediction using ensemble learning techniques: a case study in yunyang county, chongqing, China. *J. Rock Mech. Geotech. Eng.* 14 (4), 1089–1099. doi:10.1016/j.jrmge.2021.12.011

Zhang, Y., Ding, J., Sun, J., and Zhang, D. (2023b). Prediction and online optimization of strip shape in hot strip rolling process using sparrow search algorithm online sequential deep multilayer extreme learning machine algorithm. *Steel Res. Int.* 94. doi:10.1002/srin.202200832

Zienkiewicz, O. C., Humpheson, C., and Lewis, R. W. (1975). Associated and non-associated visco-plasticity and plasticity in soil mechanics. *Géotechnique* 25 (4), 671–689. doi:10.1680/geot.1975.25.4.671