ARTICLE

Uniaxial Compressive Strength Prediction for Rock Material in Deep Mine Using Boosting-Based Machine Learning Methods and Optimization Algorithms

Junjie Zhao, Diyuan Li[*], Jingtai Jiang and Pingkuang Luo

School of Resources and Safety Engineering, Central South University, Changsha, 410083, China

*Corresponding Author: Diyuan Li. Email: diyuan.li@csu.edu.cn

ABSTRACT

Traditional laboratory tests for measuring rock uniaxial compressive strength (UCS) are tedious and time-consuming. There is a pressing need for more effective methods to determine rock UCS, especially in deep mining environments under high *in-situ* stress. Thus, this study aims to develop an advanced model for predicting the UCS of rock material in deep mining environments by combining three boosting-based machine learning methods with four optimization algorithms. For this purpose, the Lead-Zinc mine in Southwest China is considered as the case study. Rock density, P-wave velocity, and point load strength index are used as input variables, and UCS is regarded as the output. Subsequently, twelve hybrid predictive models are obtained. Root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and the proportion of the mean absolute percentage error less than 20% (A-20) are selected as the evaluation metrics. Experimental results showed that the hybrid model consisting of the extreme gradient boosting method and the artificial bee colony algorithm (XGBoost-ABC) achieved satisfactory results on the training dataset and exhibited the best generalization performance on the testing dataset. The values of R^2, A-20, RMSE, and MAE on the training dataset are 0.98, 1.0, 3.11 MPa, and 2.23 MPa, respectively. The highest values of R^2 and A-20 (0.93 and 0.96), and the smallest RMSE and MAE values of 4.78 MPa and 3.76 MPa, are observed on the testing dataset. The proposed hybrid model can be considered a reliable and effective method for predicting rock UCS in deep mines.

KEYWORDS

Uniaxial compression strength; strength prediction; machine learning; optimization algorithm

Nomenclature

ρ	Density
V_P	P-wave velocity
I_{s50}	Point load strength
UCS	Uniaxial compressive strength

1 Introduction

As one of the most critical parameters of rock strength, uniaxial compressive strength (UCS) is widely used in geotechnical engineering, tunneling, and mining engineering. The reliability of

acquiring rock UCS *in situ* directly influences subsequent operations, such as drilling, digging, blasting, and support. Typically, the UCS value of rock materials can be obtained by following the well-established regulations of the International Society for Rock Mechanics (ISRM) [1] and the American Society for Testing Materials (ASTM) [2]. Laboratory testing has strict requirements for specimen preparation. It is challenging to obtain high-quality rock samples from layered sedimentary rocks, highly weathered rocks, and fractured rock masses [3–6]. Moreover, the accuracy of laboratory testing depends on the professionalism of the operators. For this reason, other highly efficient and straightforward methods for rock UCS prediction have been developed in various studies, such as statistical methods (single regression analysis methods and multiple variable regression methods) and soft computing-based methods.

Several studies have employed statistical equations to investigate the relationship between a single variable and UCS [7,8]. Fener et al. [9] conducted several laboratory tests and developed an equation for rock UCS prediction based on point load test results, achieving better performance with an R^2 value of 0.85. Yasar et al. [10] investigated the relationships between hardness (Shore Scleroscope hardness and Schmidt hammer hardness) and UCS, revealing that the hardness property showed a high correlation coefficient with rock UCS. Yilmaz [11] introduced a new testing method to determine the UCS of rock, called the core strangle test (CST). The UCS prediction results obtained from CST were more accurate than those from the point load tests. Basu et al. [12] pointed out that point load strength could be used to predict the UCS of anisotropic rocks, and in their study, the final R^2 result reached 0.86. Khandelwal [13] adopted the linear regression method to fit the relationships between P-wave velocity (Vp) and rock mechanical properties, indicating that Vp was highly correlated with UCS. Amirkiyaei et al. [14] developed a statistical model to predict the UCS of building rock materials after a freeze-thaw operation. The porosity (n) and Vp of the fresh stone were used as input variables, and the model achieved acceptable accuracy. Moreover, the Schmidt hammer rebound value (SHR), another nondestructive test index, was widely accepted for rock UCS prediction due to the data accessibility. Yagiz [15] investigated the correlation between SHR and UCS on nine rock types. The developed equation showed that SHR was strongly correlated with UCS. Nazir et al. [16] proposed a new correlation to predict UCS based on L-type SHR, achieving an R^2 value of up to 0.91, but the correlation was not recommended for highly weathered rock. In another study, Wang et al. [17] also concluded that L-type SHR could be considered an effective parameter for predicting UCS.

However, predicting UCS using a single related factor is not advised because rock strength is determined by a combination of physical and mechanical properties [5]. Multiple regression models have shown better performance in rock UCS prediction than single regression methods. Azimian et al. [5] compared the prediction performance of UCS between the standalone regression model and the multiple regression method. The R^2 values of the single regression equations using only Vp and point load strength index (I_{s50}) were 0.90 and 0.92, respectively. In contrast, the prediction result obtained through the multiple regression model was more accurate ($R^2 = 0.94$). Similarly, Minaeian et al. [18] compared the performance of a multivariate regression model with two single regression models in estimating rock UCS. Vp and SHR were employed in the multiple regression method, and then used in the two single regression models, respectively. The variance accounted for (VAF%) for the multivariate regression model and the two single prediction models were 97.24%, 94.34%, and 94.39%, respectively. Farhadian et al. [19] developed a simple nonlinear multiple regression method for predicting rock UCS based on Vp and SHR, and the results showed that the model exhibited strong prediction capability. In addition, Mishra et al. [20] confirmed that the multiple regression method outperformed the simple regression model in rock UCS prediction. Although the statistical methods presented above showed improved prediction performance, their ability to generalize to new datasets is limited.

In recent years, artificial intelligence (AI) methods have become widely used in engineering to tackle complex nonlinear problems due to their superior capabilities. They are recommended for solving other complex problems, such as earth pressure calculation and rock profile reconstruction [21,22]. Research reviews have revealed that AI-based methods have been successfully applied in areas such as tunnel squeezing analysis [23,24], rock mass failure mode classification [25], rock lithology classification [26,27], and rock burst prediction and assessment [28–30]. The outstanding results achieved by AI-based methods have garnered significant attention in the prediction of rock mechanical parameters. For example, Ghasemi et al. [31] proposed the M5P model tree method to predict the UCS of carbonate rocks. SHR, n, dry unit weight, Vp, and slake durability index were considered as input parameters. Wang et al. [32] employed the random forest algorithm to estimate the UCS value based on L-type SHR and Vp. The proposed method achieved comparable results to the measured values. Cao et al. [4] combined the extreme gradient boosting method with the firefly optimization algorithm to predict rock UCS. Input data included rock dry density (ρ), Vp, and the proportion of crystals in the rocks (such as quartz (Qtz), biotite (Bi), plagioclase (Plg), chlorite (Chl), and alkali feldspar (Kpr)). The model achieved high-precision results. Barzegar et al. [6] developed an ensemble tree-based method to estimate rock UCS using Vp, SHR, n, and I_{s50}. The research results revealed that the ensemble model outperformed both the standalone machine learning (ML) models and multivariate regression models. Jin et al. [33] created a hybrid model combining the extreme learning machine approach with the grey wolf optimization algorithm. Four rock properties (n, Vp, SHR, I_{s50}) were used. The hybrid model achieved better prediction performance ($R^2 = 0.951$) on the testing dataset than the other four models. Saedi et al. [34] investigated six input factors, including n, cylindrical punch index, block punch index, Brazilian tensile strength, I_{s50}, and Vp, and created six prediction models. The findings revealed that the fuzzy inference system and multivariate regression models outperformed the single regression technique. Li et al. [35] adopted six optimization algorithms to enhance the random forest model's performance for rock UCS prediction, employing five parameters: SHR, Vp, I_{s50}, n, and ρ. The proposed model achieved excellent performance with an R^2 value of 0.9753 after optimization. Mahmoodzadeh et al. [36] compared six AI-based methods for rock UCS prediction using four input variables (n, SHR, Vp, and I_{s50}) in the models. Among all predictive models, the Gaussian process regression model performed the best ($R^2 = 0.9955$). Skentou et al. [37] developed an artificial neural network model with three optimization algorithms. Three parameters, such as pulse velocity, SHR, and effective n, were applied in their study, and the best hybrid model obtained outstanding performance ($R^2 = 0.9607$).

AI-based methods offer distinct advantages over traditional empirical formulas for rock UCS prediction. Traditional approaches to determining UCS through laboratory testing face limitations. These include challenges in obtaining high-quality rock samples from certain geological conditions, such as severely fragmented rock masses, in-situ core disking, and lower efficiency in high-quality rock sample collection. Moreover, few studies have been conducted on predicting the UCS of rock materials in deep mines using boosting-based approaches. Therefore, this study aims to develop a simple and robust rock UCS predictive model tailored for deep mining environments. The Lead-Zinc mine in Southwest China serves as the case study. Three easily accessible parameters ρ, Vp, and I_{s50} are used as input parameters, with UCS considered the output. The main contents are as follows: (1) constructing a comprehensive database based on previous studies and new field data obtained from deep mines in southwestern China; (2) developing a straightforward and reliable model for rock UCS prediction, integrating three boosting-based machine learning methods, adaptive boosting (AdaBoost), gradient boosting (GBoost), extreme gradient boosting (XGBoost), and four optimization algorithms: Bayesian optimization algorithm (BOA), artificial bee colony (ABC), grey

wolf optimization (GWO), and whale optimization algorithm (WOA); (3) systematically comparing the performance of all hybrid models in rock UCS prediction.

2 Methods

This study considers boosting-based ML methods such as AdaBoost, GBoost, and XGBoost base models for rock UCS prediction. Boosting, an advanced ensemble method, combines several weaker learners to create a strong learner, ensuring improved overall performance of the final model. In addition, four optimization algorithms, namely BOA, ABC, GWO, and WOA, are employed to obtain the optimal parameters for all the base models. These algorithms are renowned for their robustness and efficiency in solving complex optimization problems. They excel in multidimensional and nonlinear search spaces, providing effective solutions in various fields, from engineering to data science. Detailed descriptions of the Boosting-based approaches and the optimization algorithms are provided in the subsequent sections.

2.1 Boosting-Based ML Algorithms

2.1.1 Adaptive Boosting

AdaBoost, a typical ensemble boosting algorithm, combines multiple weak learners to form a single strong learner. By adequately considering the weights of all learners, it produces a model that is more accurate and less prone to overfitting. The procedure of the algorithm is outlined below:

(1) Determine the regression error rate of the weak learner.

- Obtain the maximum error;

$$E_t = max\,|y_i - h_t(x_i)|,\ \ i = 1, 2, \ldots, N \tag{1}$$

where x_i, y_i indicates the input data; N is the number of the sample; $h_t(x)$ represents the weak learner of the t-th iteration.

- Estimate the relative error for each sample using the linearity loss function;

$$\epsilon_{ti} = \frac{|y_i - h_t(x_i)|}{E_t} \tag{2}$$

- Determine the regression error rate;

$$\epsilon_t = \sum_{i=1}^{n} w_{ti}\epsilon_{ti} \tag{3}$$

where w_{ti} is the weight of the sample x_i.

(2) Determine the weight coefficient α of the weak learner.

$$\alpha_t = \frac{\epsilon_t}{1 - \epsilon_t} \tag{4}$$

(3) Samples weights updation for the $t + 1$ iteration round.

$$W_{t+1}(x_i) = \frac{W_t(x_i)}{Z_t}\alpha_t^{1-\epsilon_{ti}} \tag{5}$$

$$Z_t = \sum_{i=1}^{N} W(x_{ti})\,\alpha_t^{1-\epsilon_{ti}};\ Z_t \text{ is the normalization factor} \tag{6}$$

(4) Build the ultimate learner.

$$H(x) = \sum_{t=1}^{T} \ln\left(\frac{1}{\alpha_t}\right) f(x) . f(x) \text{ is the median value of } \alpha_t h_t(x) (t = 1, 2, \ldots, T) \tag{7}$$

Additionally, the AdaBoost method requires only a few parameter adjustments, such as the decision tree depth, the number of iterations, and the regression loss function. Consequently, it has been extensively used in various fields, including rock mass classification [38], rockburst prediction [39], rock strength estimation [40,41], and tunnel boring machine performance prediction [42].

2.1.2 Gradient Boosting

The GBoost method, another ensemble algorithm, is inspired by gradient descent. It trains a new weak learner based on the negative gradient of the current model loss. This well-trained weak learner is then combined with the existing model. The final model is constructed by repeating these accumulation steps, as shown below:

(1) Initialize the base learner.

$$H_0(x) = \arg min_c \sum_{i=1}^{m} L(y_i, c) \tag{8}$$

$$c = \frac{\sum_{i=1}^{m} y_i}{m} \tag{9}$$

where $L(x)$ is the squared error loss function; c is a constant; y_i is the measured value; m is the number of the sample; $H_0(x)$ is the initial weak learner.

(2) Calculate the negative gradient. Then, the leaf node region R_{tj} of the t-th regression tree (J is the number of leaf nodes of regression tree t) could be obtained based on (x_i, r_{ti}).

$$r_{ti} = -\frac{\partial L(y_i, h_{t-1}(x_i))}{\partial h_{t-1}(x_i)}, \ i = 1, 2, \ldots, m \tag{10}$$

(3) Obtain the optimal-fitted value c_{tj} for the leaf region $j = 1, 2, \ldots, J$, and update the strong learner.

$$c_{tj} = \arg min_c \sum_{x_i \in R_{tj}} L(y_i, h_{t-1}(x_i) + c) \tag{11}$$

$$h_t(x) = h_{t-1}(x) + \sum_{j=1}^{J} c_{tj}I, \ (x \in R_{tj}) \tag{12}$$

where $h_{t-1}(x)$ is the regression learner obtained after $t-1$ iterations; I is the unit matrix.

(4) Finally, a strong learner could be obtained.

$$f(x) = f_T(x) = \sum_{t=1}^{T} \sum_{j=1}^{J} c_{tj}I \ (x \in R_{tj}) \tag{13}$$

An advantage of GBoost is its flexibility in choosing the loss function, allowing for the use of any continuously differentiable loss function. This characteristic makes the model more resilient to noise. Due to these advantages, GBoost-based methods have achieved considerable success, resulting in many improved versions [43–45].

2.1.3 Extreme Gradient Boosting

XGBoost, an extension of the gradient boosting algorithm, was developed by Chen et al. [46]. This model has robust applications in classification and regression tasks. The primary idea behind the

XGBoost approach is to continually produce new trees, with each decision tree being updated based on the difference between the previous tree's result and the target value, thereby minimizing model bias. Given a dataset $D = \{(x_i, y_i)\}$, where x_i is the sample, the prediction result of the well-trained XGBoost model is calculated as follows:

$$\widetilde{y}_i = \sum_{k=1}^{K} f_k(x_i) \tag{14}$$

where x_i is the sample; $f_k(x_i)$ is the prediction result of the k-th tree for the sample x_i; K is the number of decision trees; $\widetilde{y}_i$ is the sum of the prediction results of all the decision trees.

$$Obj = \sum_{i=1}^{n} l(y_i, \widetilde{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{15}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} (w_j)^2 \tag{16}$$

where $l(y_i, \widetilde{y}_i)$ is the loss function, such as mean square error and cross-entropy loss; y_i is the target value; n is the number of samples; $\Omega(f_k)$ is the complexity of the tree; T is the number of leaves; w_j is the L_2 norm of leaf scores; γ and λ are the factors that aim to change the complexity of the tree.

The XGBoost method combines the loss function and regularization factor into its objective function, resulting in higher generalization than other models, as shown in Eqs. (14), (15). Another distinctive feature of the XGBoost method is its incorporation of a greedy algorithm and an approximation algorithm for searching the split nodes of the tree. The main optimization parameters are the decision tree depth, the number of estimators, and the maximum features. Chang et al. [47] developed an effective model for credit risk assessment using the XGBoost method. Zhang et al. [48] used the XGBoost method to forecast the undrained shear strength of soft clays. In another work, Nguyen-Sy et al. [49] used the XGBoost method to predict the UCS of concrete. The XGBoost model outperforms the artificial neural network (ANN) model, support vector machine (SVM), and other ML methods.

2.2 Optimization Algorithms

Optimization algorithms are usually designed to automatically find the best global solution within the given search space, shortening the model development cycle and ensuring model robustness. Hence, four well-performing optimization algorithms, i.e., BOA, ABC, GWO, and WOA, were implemented in this study to optimize the parameters of the aforementioned boosting-based ML models for the prediction of rock UCS.

2.2.1 Bayesian Optimization Algorithm

Compared to the commonly used grid search and random search algorithms, the BOA method, proposed by Pelikan et al. [50], fully utilizes prior information to find the parameters that maximize the target function globally. The algorithm comprises two parts: (1) Gaussian process regression, which aims to determine the values of the mean and variance of the function at each point, and (2) constructing the acquisition function, which is employed to obtain the search position of the next iteration, as shown in Fig. 1. BOA has the advantages of fewer iterations and a faster processing speed and has been applied in several fields. Díaz et al. [51] studied the UCS prediction of jet grouting columns based on several ML algorithms and the BOA method. The optimized model obtained significant improvement compared to existing works. Li et al. [26] proposed an intelligent model for rockburst prediction using BOA for hyperparameter optimization. Lahmiri et al. [52] used BOA to obtain the optimal parameters of models for house price prediction. Bo et al. [53] developed

an ensemble classifier model to assess tunnel squeezing hazards, with the optimal values of the seventeen parameters obtained utilizing the BOA method. Additionally, Díaz et al. [54] investigated the correlations between activity and clayey soil properties. Thirty-five ML models were introduced in their research to predict the activity using the clayey soil properties, with the BOA method being used to fine-tune the ML models' hyperparameters, producing promising results.
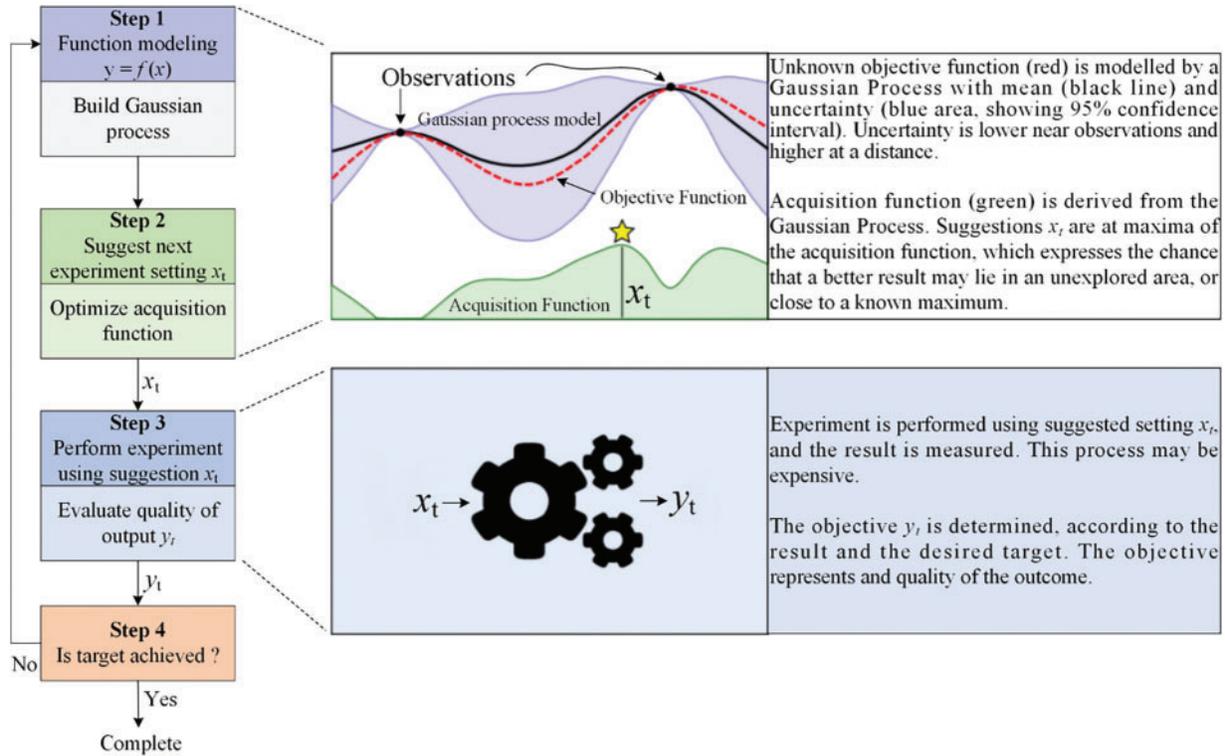

Figure 1: Schematic of the Bayesian optimization algorithm [55]

2.2.2 Artificial Bee Colony

The ABC method, developed for multivariate function optimization problems by Karaboga [56], divides bees in a colony into three groups (employed, onlookers, and scouts) based on task assignment [57], as illustrated in Fig. 2. Employed bees are tasked with finding available food sources and gathering information. In contrast, onlookers collect good food sources based on data transferred from employed bees and perform further searches for food. Scouts are responsible for finding valuable honey sources around the beehive. In mathematical terms, the food source represents the problem's solution, and the nectar level equates to the fitness value of the solution [58]. Parsajoo et al. [59] adopted the ABC method to tune and improve the model performance for rock brittleness index prediction. Zhou et al. [60] recommended an intelligent model for rockburst risk assessment and applied the ABC method to obtain optimal hyperparameters for the model. The results revealed ABC to be a valuable and successful strategy.
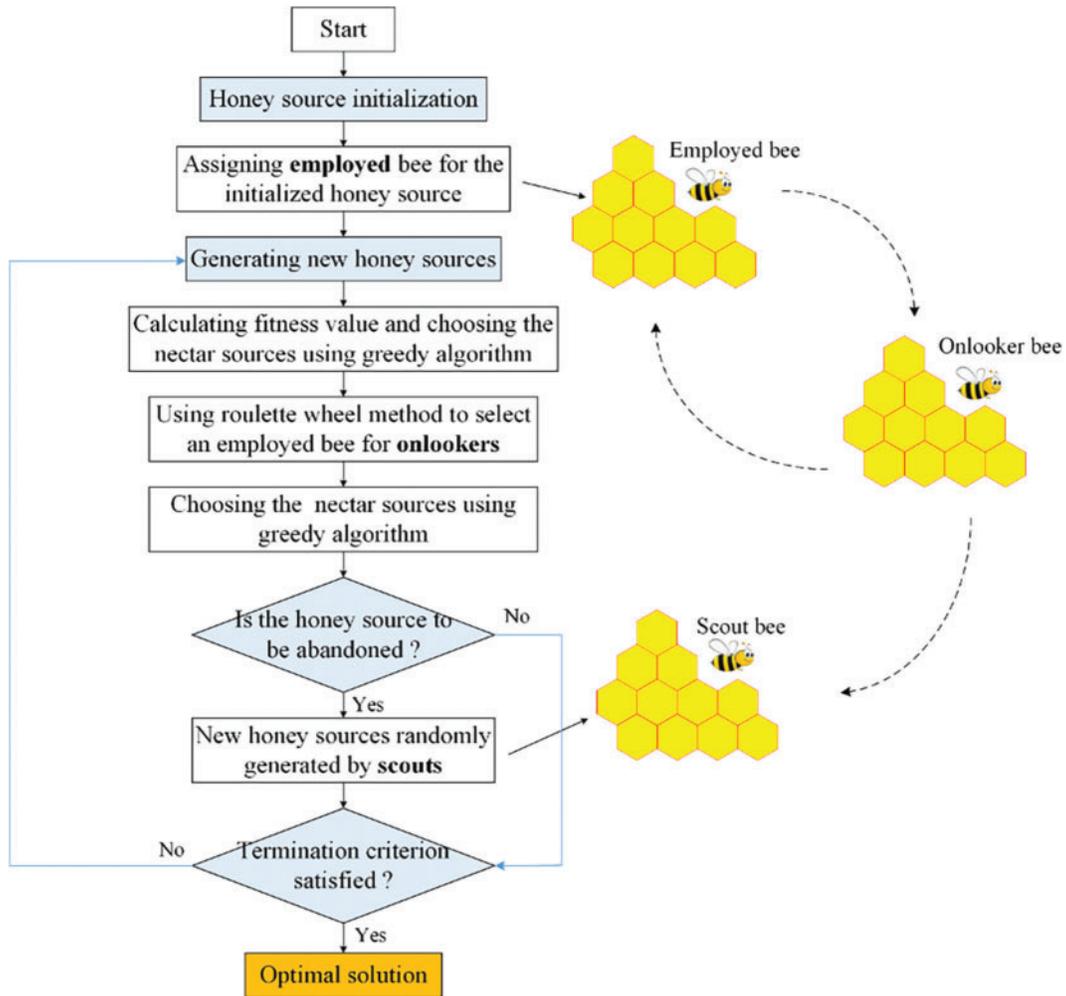
Figure 2: Schematic of the artificial bee colony method

2.2.3 Grey Wolf Optimization Algorithm

GWO is a metaheuristic method developed by Mirjalili et al. [61]. The algorithm simulates grey wolf predation in nature, and wolves are divided into four hierarchies: α, β, δ, w, α, β, δ are the head wolves leading the other wolves (w) moving toward the destination, and the position of wolf w is updated based on α, β, or δ using the following equations:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_p(t) - \vec{X}(t) \right| \tag{17}$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A} \cdot \vec{D} \tag{18}$$

where $\vec{X}_p$ and $\vec{X}(t)$ are the prey and grey wolf position vectors, respectively; vectors $\vec{A}$ and $\vec{C}$ are the constant coefficients; t is the current iteration.

The constant coefficients $\vec{A}$ and $\vec{C}$ can be calculated as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \tag{19}$$

$$\vec{C} = 2 \cdot \vec{r}_2 \tag{20}$$

where $\vec{r}_1$ and $\vec{r}_2$ are vectors randomly distributed in [0, 1]; $\vec{a}$ is a convergence factor that drops linearly from 2 to 0 as iterations increase.

During the iteration process, the best solution can be obtained by the head wolves α, β, and δ, and the value of $\left|\vec{A}\right| > 1$ means that the candidate solution is far away from the prey. In contrast, when the value of $\left|\vec{A}\right| < 1$ indicates that the candidate solution is close to the prey. The flowchart is shown in Fig. 3. Several studies highlight the GWO method due to its simple structure, few parameters, and easy implementation. Golafshani et al. [62] developed a model to estimate concrete UCS, demonstrating that the GWO-optimized model outperformed the original prediction model. Shariati et al. [63] reported that integrating the GWO approach can greatly improve the model's predictive capability.
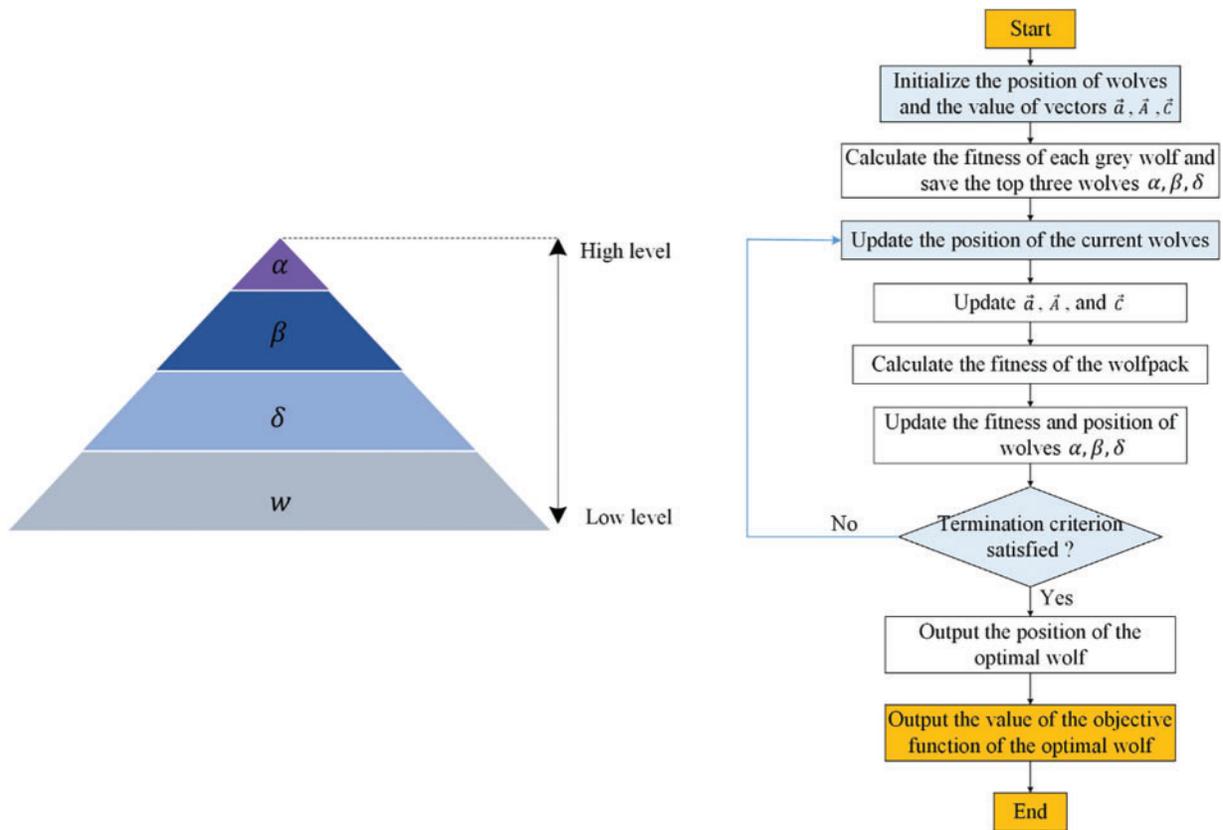

Figure 3: Flowchart of the grey wolf optimization algorithm

2.2.4 Whale Optimization Algorithm

WOA, a unique population intelligence optimization method mimicking whale-feeding behavior, was introduced by Mirjalili et al. [64]. The mathematical modeling process of WOA is comparable to that of the GWO approach. However, a critical distinction between the two algorithms is that humpback whales complete their prey behaviors using either random whale individuals or the ideal individuals, as well as the spiraling bubble-net mode. Zhou et al. [65] applied WOA to obtain optimal parameters for the SVM model for tunnel squeezing classification, achieving higher prediction accuracy. In another study, Tien et al. [66] presented a model for predicting concrete UCS using various optimization techniques, with WOA-based optimization performing the best. Nguyen et al. [67] combined SVM and WOA algorithms to create an intelligent model for predicting fly rock distance, demonstrating that the hybrid WOA-SVM model outperformed standalone models.

3 Database

3.1 Data Preparation

Rock samples collected from a deep lead-zinc ore mine in Yunnan Province, Southwest China, were used as the dataset for developing rock UCS prediction models. The current maximum mining depth of the lead-zinc ore mine has exceeded 1,500 meters, and field rock sample collection operations were conducted at different sublevels, including lower and upper plates surrounding rock and ore, as shown in Fig. 4. All drilling works were done using the KD-100 fully hydraulic pit drilling rig, a small and easy-to-operate machine propelled by compressed air.

Figure 4: Fieldworks. (a)∼(b) indicate borehole sampling at different sublevels, (c)∼(d) are the collected samples in-situ, (e)∼(f) are the corresponding high-quality samples

After that, all the high-quality samples were made into standard specimens with dimensions of $\varnothing 50 \times 100$ mm following the instructions of ISRM and ASTM. Fig. 5 shows the well-processed standard specimens. Then, the rock parameters ρ, V_p, I_{s50}, and UCS were obtained through laboratory tests, as depicted in Fig. 6. The rock density $\rho = m/V$ and the value of the P-wave velocity V_p was measured by the sonic parameter tester. I_{s50} was obtained through the irregular lump test, and the irregular blocks were collected from the same sublevels as the drilling, as shown in Fig. 7. The UCS test was conducted on INSTRON 1346 equipment. Finally, 40 sets of physical and mechanical parameters of the rock were established from the deep mine, as shown in Table 1. In addition to the field dataset, 66 datasets provided in the study [68] were collected to expand the database, and thus, a total of 106 datasets were integrated as the final dataset to develop the models.

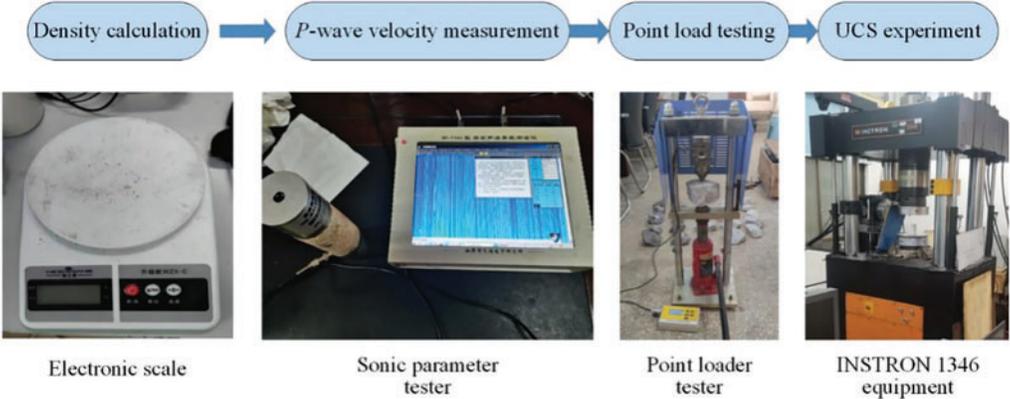**Figure 5:** The well-processed standard specimens

Figure 6: Laboratory tests

Figure 7: *In-situ* irregular rock blocks collection for point load testing

Table 1: The results of laboratory tests

ID	ρ (g/cm³)	V_p (m/s)	I_{s50} (MPa)	UCS (MPa)	ID	ρ (g/cm³)	V_p (m/s)	I_{s50} (MPa)	UCS (MPa)
1	2.83	4401	2.98	72.70	22	2.66	4599	3.68	91.12
2	2.76	3053	4.14	101.11	23	2.72	4207	3.45	91.68
3	2.74	2798	4.03	98.35	24	2.71	3604	2.33	51.89
4	2.76	4208	2.03	49.61	25	2.70	4219	2.75	66.48
5	2.71	4192	2.37	60.25	26	2.86	4205	3.69	92.99
6	2.76	3879	4.08	100.49	27	2.89	3738	3.51	86.90
7	2.96	3883	4.19	102.33	28	2.88	3878	3.78	89.82
8	2.75	4393	2.50	63.26	29	2.88	4043	2.75	67.19
9	2.77	3634	3.28	80.13	30	2.78	4820	4.93	120.31
10	2.70	4393	4.14	100.95	31	2.77	4393	5.34	135.10
11	2.67	4396	4.23	103.11	32	2.68	4821	2.51	63.65
12	2.77	3367	2.42	61.04	33	2.75	4396	3.23	78.90
13	2.73	4595	2.81	71.40	34	2.66	4205	2.84	71.83
14	3.75	3482	2.78	71.24	35	2.66	4211	3.01	73.21
15	4.74	3474	3.11	75.91	36	2.78	4047	2.45	62.12
16	4.85	4823	2.62	63.61	37	2.66	3248	2.38	55.48
17	4.57	3366	3.12	78.63	38	2.78	4206	3.42	85.96
18	4.62	3637	4.89	119.22	39	2.78	3252	4.28	106.99
19	4.93	3599	3.40	82.96	40	2.71	4219	2.78	70.25
20	4.80	4038	3.33	81.18					
21	4.85	4209	3.05	74.50					

Note: The irregular blocks are collected from the same sublevels as the drilling, and yield the I_{s50} as the corresponding point load strength of the standard specimen.

3.2 Exploration Data Analysis

For regression prediction problems, correlation analysis between independent and dependent variables is always essential [69]. Fig. 8 shows the analysis results of the correlation between the variables ρ, V_p, I_{s50}, and UCS. There is a positive correlation between the input variables and the output parameter UCS, as can be seen in the last line of Fig. 8, where the values of the Pearson correlation coefficient between ρ, V_p, I_{s50}, and UCS are 0.53, 0.53, and 0.54, respectively. Moreover, the correlation between independent variables shows lower values, which indicates that the input variables used in this paper for the prediction model development are reasonable and feasible. Additionally, the data distribution of each parameter is illustrated through violin plots, as shown in Fig. 9, where the values of ρ, V_p, I_{s50}, and UCS are evenly distributed.
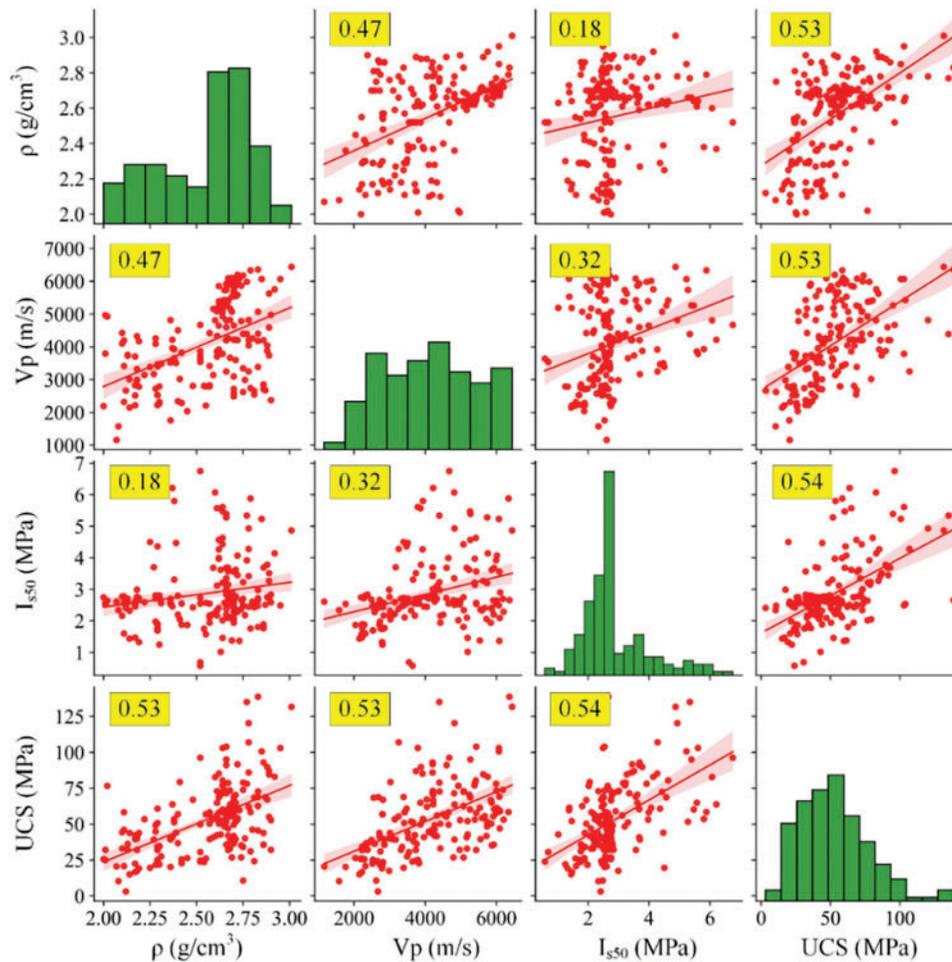
Figure 8: Correlation analysis between input and output variables
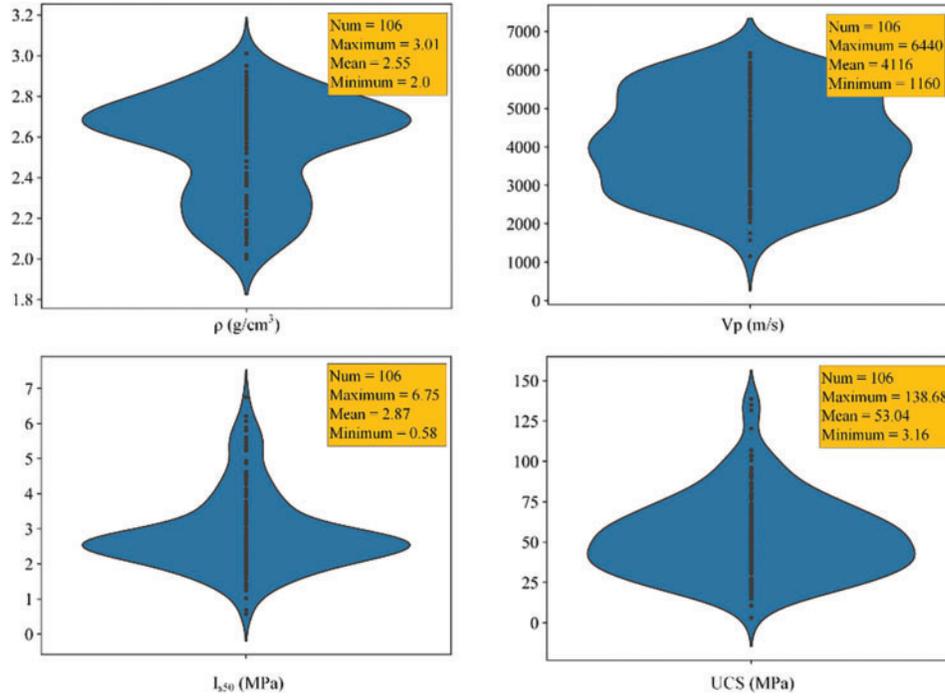
Figure 9: Violin plot of variables

Furthermore, to balance the quantity of training and testing datasets, half of the randomly selected field data and data acquired from the literature were used as training data (86 sets of data), while the remaining field data were used as testing data (20 sets of data). The final ratio between the training and testing data is approximately 8:2.

Simultaneously, all data were normalized prior to model training due to the different magnitudes of the variables. The standardization process was as follows:

$$x^* = \frac{x - \mu}{\sigma} \tag{21}$$

where x is the input variables; μ and σ are the mean value and standard deviation of each variable parameter.

4 Evaluation Indexes

The performance of all hybrid models was evaluated using four evaluation indices: root mean square error (RMSE), mean absolute error (MAE), R^2, and A-20. Typically, lower values of RMSE and MAE indicate a better model, suggesting that the model's predictions are closer to the actual values. Conversely, a larger R^2 value signifies a more robust model, with a maximum value of 1. The value of A-20 equals the proportion of samples where the mean absolute percentage error between the predicted and actual values is less than 20 percent, as shown in Eq. (25). A larger A-20 value indicates more accurate model predictions.

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (y_i - y_i')^2} \tag{22}$$

$$MAE = \frac{1}{M}\sum_{i=1}^{M}\left|y_i - y_i'\right| \tag{23}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{M}\left(y_i - y_i'\right)^2}{\sum_{i=1}^{M}\left(y_i - \overline{y}\right)^2} \tag{24}$$

$$A - 20 = \frac{count}{M}$$

$$count = \sum(errors < 20) \tag{25}$$

$$errors = \left|\frac{y_i - y_i'}{y_i}\right| \times 100$$

where y_i, y_i', and $\overline{y}$ are the target value, prediction result, and the average of all the target values, respectively; M is the number of the dataset. *errors* is the mean absolute percentage error between the predicted value and the actual value.

5 Development of the Prediction Models

This study used boosting-based ML algorithms, including AdaBoost, GBoost, and XGBoost, to predict rock UCS. In addition, four optimization approaches, BOA, ABC, GWO, and WOA, were employed to determine the best parameters for all boosting models, ensuring prediction accuracy. The hybrid models were then tested on testing data, providing the optimal rock UCS prediction model. Fig. 10 depicts the flowchart for hybrid model construction.

5.1 Development of the AdaBoost Model

For the AdaBoost method, the default parameters for regression problems include the base estimator, the number of estimators, the learning rate, and the loss function. The number of estimators and the learning rate were the optimization parameters. However, the base estimator and the loss function were set to their default values (CART decision tree and linear loss function). Additionally, the population size values for the ABC, GWO, and WOA optimization algorithms ranged from 10 to 50, with 5-unit intervals, and the total number of iterations was 100.

To obtain the optimal hyperparameters of the AdaBoost model, the R^2 and RMSE values of different AdaBoost hybrid models during training were summarized, as shown in Figs. 11 and 12. Fig. 11 shows the R^2 values R^2 for four hybrid models trained under different population sizes. The AdaBoost-GWO and AdaBoost-WOA models achieved the highest R^2 values at population sizes of 15 and 40, respectively. The AdaBoost-ABC model performed equally well at population sizes of 40 and 45 but converged faster at 40. The AdaBoost-BOA model was also trained with the same number of iterations, as shown in Fig. 11d. Moreover, the RMSE values for all hybrid models were calculated, and the results are presented in Fig. 12. Based on these results, the optimal hyperparameters of the AdaBoost model using the four optimization algorithms are listed in Table 2. The corresponding prediction performances of the four optimized hybrid models on training and testing datasets are displayed in Fig. 13. The bars with solid color filling indicate the prediction capabilities of the hybrid models on training data, while the bars with slash-filling represent the model prediction performance on testing data. The AdaBoost-ABC and AdaBoost-GWO models exhibited the most robust prediction abilities, with the highest R^2 and A-20 values (0.60 and 0.85), and the lowest RMSE and MAE values (11.45 MPa and 10.30 MPa, respectively) on the testing datasets. Therefore,

the optimal parameters of the AdaBoost method were determined based on the ABC and GWO optimization algorithms, with the best learning rate being 0.197 and the optimal number of estimators rounded to 6.
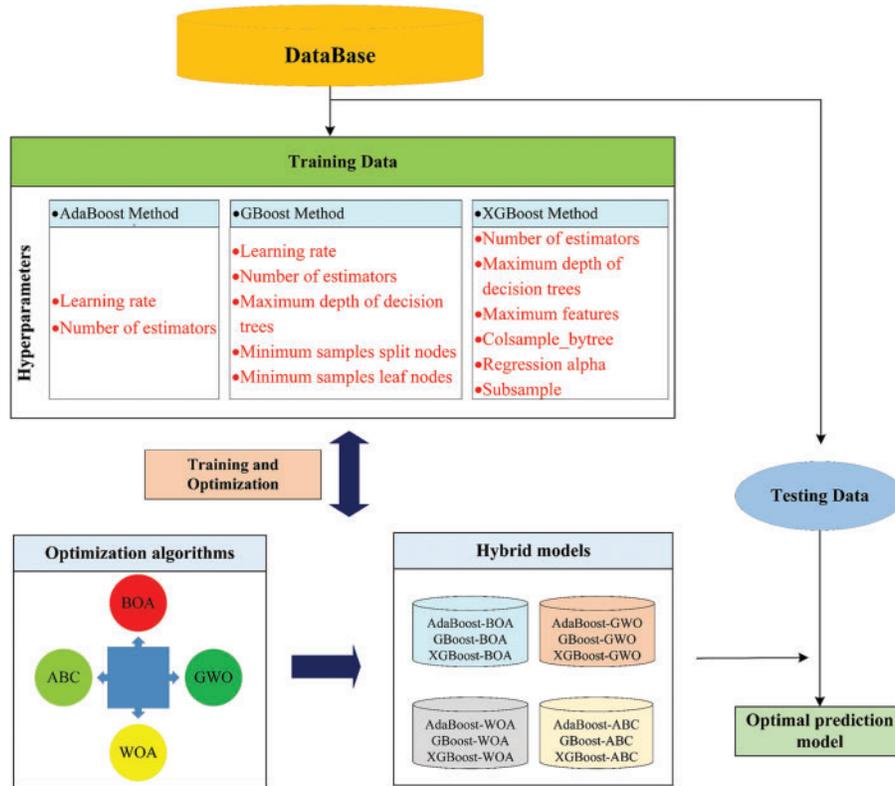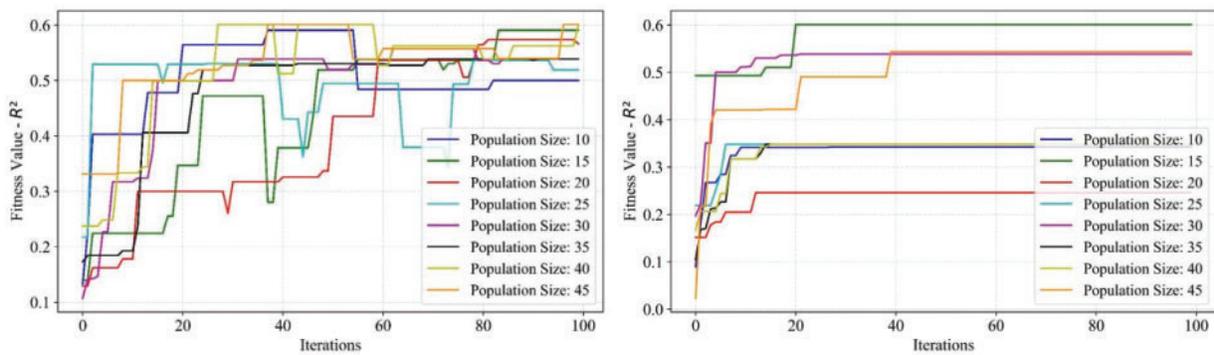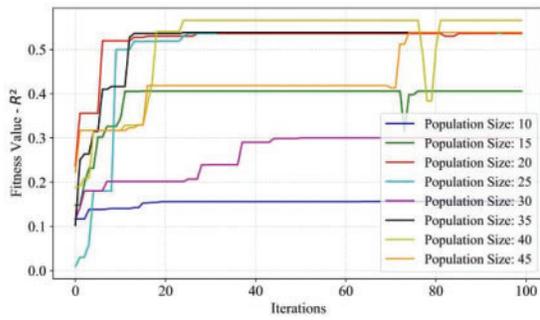

Figure 10: Schematic of the overall framework for rock UCS prediction
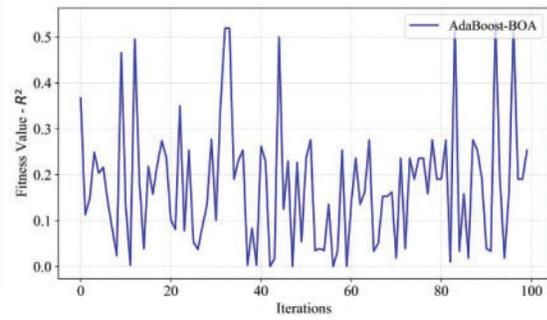

(a) AdaBoost-ABC   (b) AdaBoost-GWO
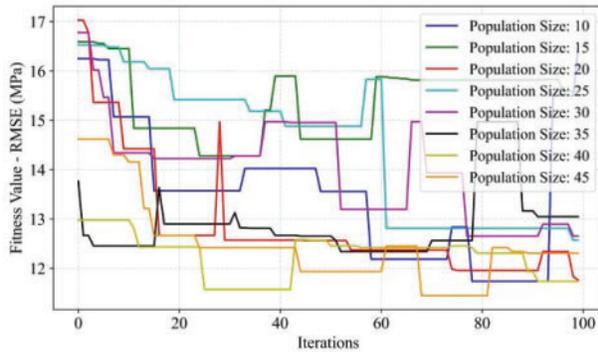
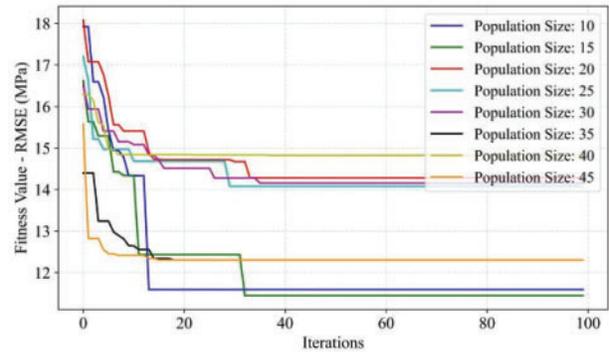**Figure 11:** (Continued)

(c) AdaBoost-WOA   (d) AdaBoost-BOA

Figure 11: Fitness values of R² of the AdaBoost model using four optimization methods, ABC, GWO, WOA, and BOA, respectively

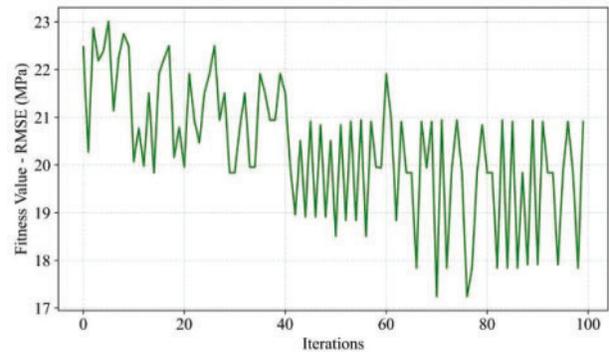(a) AdaBoost-ABC   (b) AdaBoost-GWO
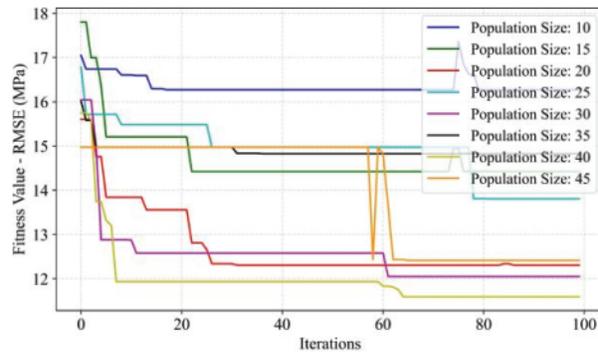
(c) AdaBoost-WOA   (d) AdaBoost-BOA

Figure 12: Fitness values of RMSE of the AdaBoost model using four optimization methods, ABC, GWO, WOA, and BOA, respectively

Table 2: Hyperparameters of the AdaBoost model using four optimization algorithms

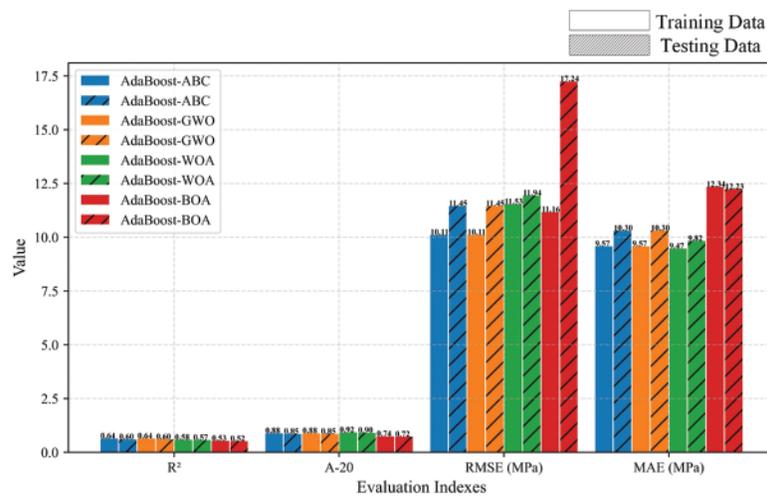| Hybrid models	Hyperparameters	
	Learning rate (1e-4∼1)	Number of estimators (1∼200)
AdaBoost-ABC	0.1973	6.47
AdaBoost-GWO	0.1978	6.15
AdaBoost-WOA	0.1974	11.59
AdaBoost-BOA	0.1978	14.72

Figure 13: Comparison results of the prediction performance of the AdaBoost hybrid models

5.2 Development of the GBoost Model

Compared to the AdaBoost model, the GBoost model required fine-tuning of more hyperparameters, including the learning rate, the number of estimators, maximum decision tree depth, minimum sample split node, and minimum sample leaf node. Similarly, the training process and the corresponding results of R^2 and RMSE for all the hybrid GBoost models were obtained. From Figs. 14 and 15, it was evident that the hybrid models GBoost-ABC, GBoost-GWO, and GBoost-WOA achieved satisfactory results in terms of R^2 and RMSE at the population sizes of 45, 30, and 20, respectively. The GBoost-BOA model also yielded better results with the increase in iterations. Subsequently, four different sets of hyperparameters for the GBoost model were obtained. The prediction capabilities of the four optimized hybrid models were comprehensively analyzed, as shown in Fig. 16. It could be seen that the R^2 results R^2 for all the optimized hybrid models on the training data were excellent, especially for the hybrid models GBoost-GWO and GBoost-WOA, where the values of R^2 were almost equal to 1. However, combined with the results of the other two indicators, RMSE and MAE, it was found that the hybrid models GBoost-GWO and GBoost-WOA showed poor generalization performance on the testing data. The hybrid models GBoost-ABC and GBoost-BOA, on the contrary, performed well on the training datasets and obtained comparable results on the testing datasets. Moreover, as shown in Fig. 16, the overall performance of the hybrid model GBoost-BOA

was superior to that of GBoost-ABC. The RMSE and MAE results for the hybrid model GBoost-BOA were lower than those for GBoost-ABC on the testing datasets. Therefore, the hybrid model GBoost-BOA was considered the best compared to the other GBoost hybrid models. Table 3 shows the optimum hyperparameters for the GBoost model using the BOA method.
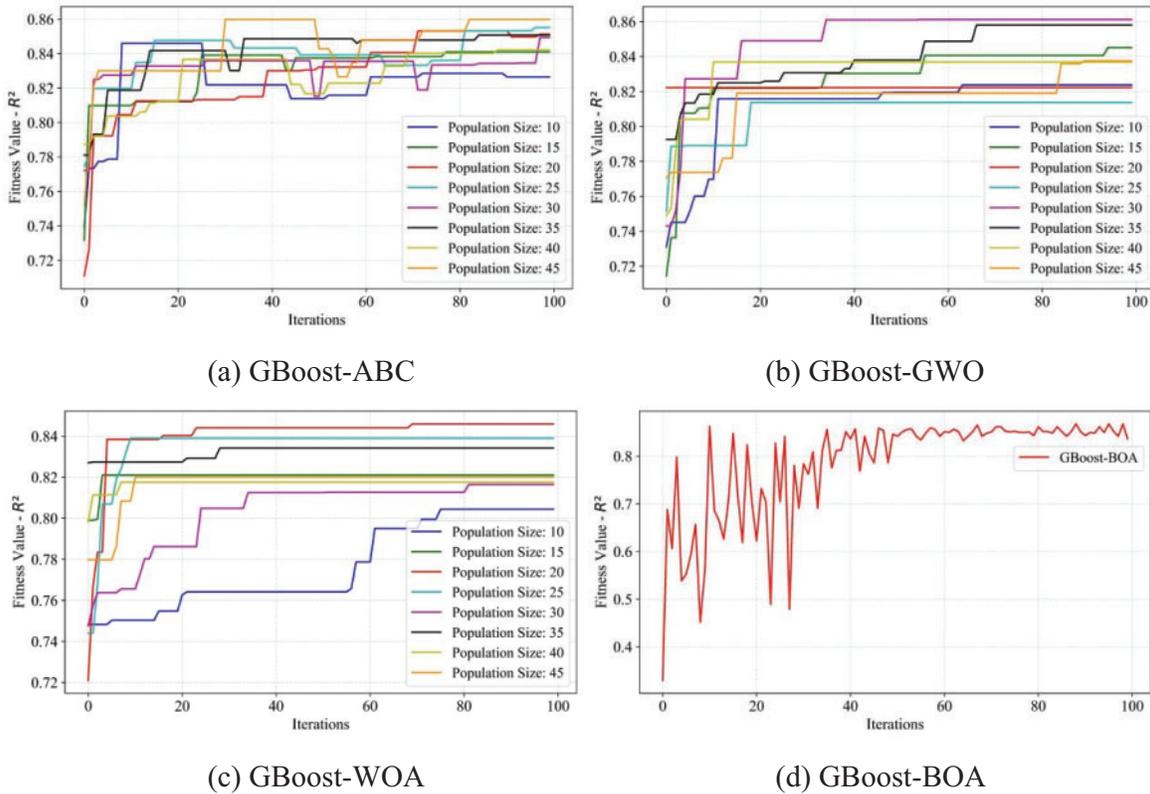

(a) GBoost-ABC

(b) GBoost-GWO

(c) GBoost-WOA

(d) GBoost-BOA

Figure 14: Fitness values of R^2 of the GBoost model using four optimization methods, ABC, GWO, WOA, and BOA, respectively

(a) GBoost-ABC

(b) GBoost-GWO

Figure 15: (Continued)
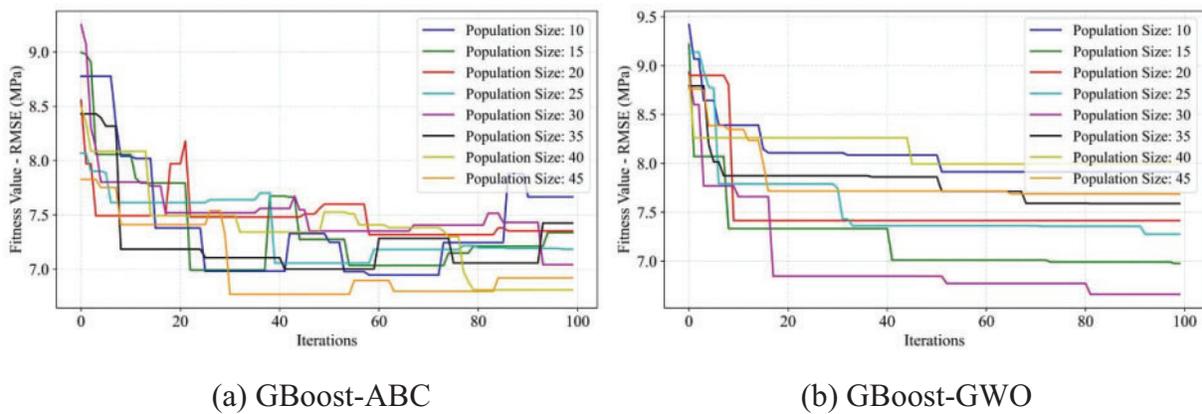
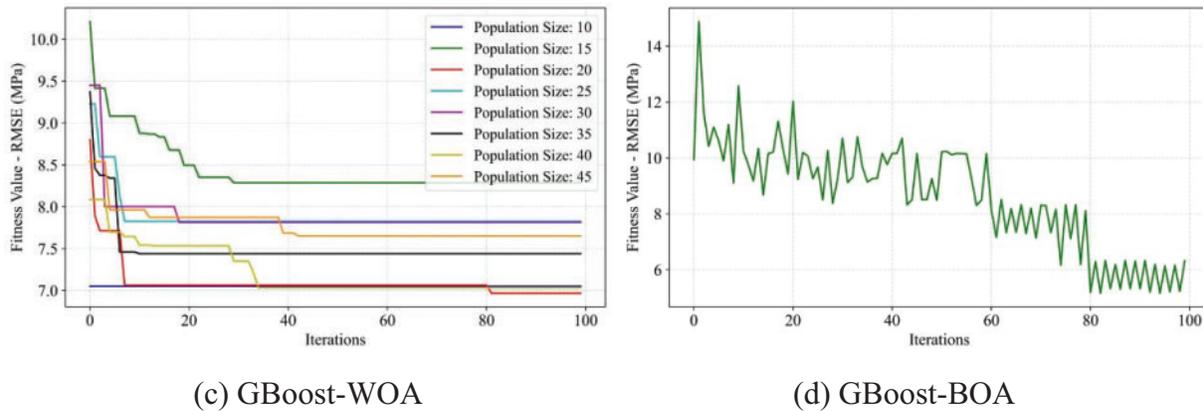(c) GBoost-WOA   (d) GBoost-BOA

Figure 15: Fitness values of RMSE of the GBoost model using four optimization methods, ABC, GWO, WOA, and BOA, respectively
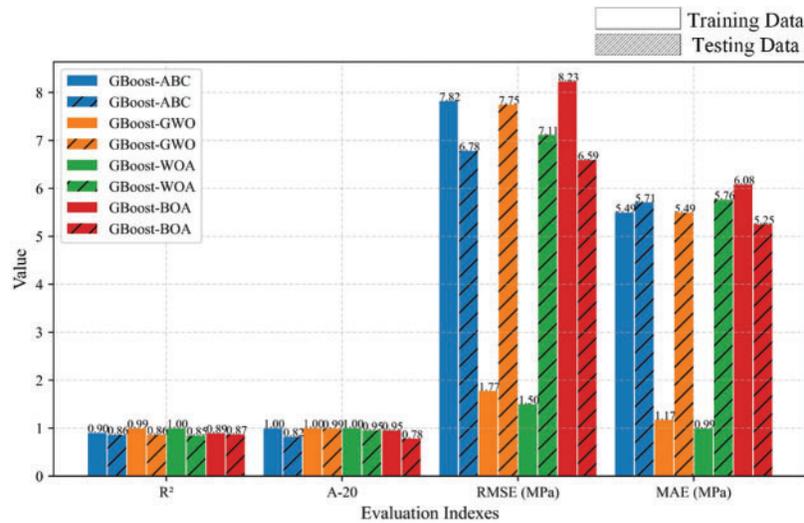

Figure 16: Comparison results of the prediction performance of the AdaBoost hybrid models

Table 3: The optimum parameters of the GBoost model

Model	Hyperparameters	Optimal value	Search space
GBoost	Learning rate	0.49	1e-4~1
	The number of estimators	65	1~200
	Maximum depth of decision trees	15	1~50
	Minimum samples split nodes	20	2~30
	Minimum samples of leaf nodes	24	1~30

5.3 Development of the XGBoost Model

Finally, the hyperparameters of the XGBoost method, such as the number of estimators, maximum decision tree depth, maximum features, tree colsample, regression alpha, and subsample, were confirmed by using the four optimization methods. Figs. 17 and 18 show the corresponding training results for the hybrid models. The hybrid models XGBoost-ABC, XGBoost-GWO, and XGBoost-WOA obtained the best results at 35, 30, and 15 population sizes, respectively. The hybrid model XGBoost-BOA also performed robustly during the training process. Fig. 19 presents the prediction performance of each hybrid model on the training and testing datasets. The hybrid model XGBoost-ABC exhibited the strongest robustness on both training and testing data, with the evaluation index R^2 values of 0.98 and 0.93, respectively, on the training and testing datasets. Meanwhile, the RMSE and MAE values were the lowest on the testing dataset. Thus, the hybrid model XGBoost-ABC was deemed the best prediction model, with the optimal population size of the ABC optimization method being 35, as shown in Fig. 17a. Table 4 summarizes the hyperparameters of the XGBoost method obtained through training with the ABC optimization algorithm.
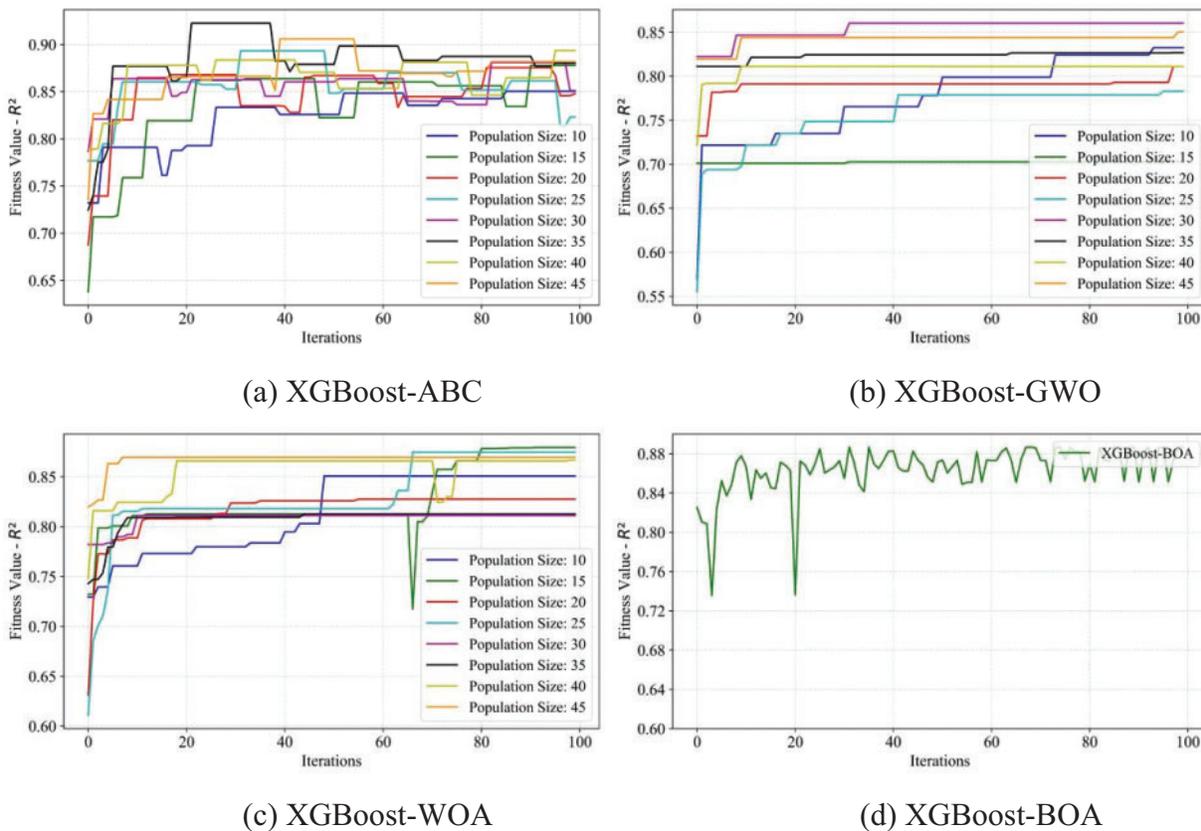

(a) XGBoost-ABC   (b) XGBoost-GWO

(c) XGBoost-WOA   (d) XGBoost-BOA

Figure 17: Fitness values of R^2 of the XGBoost model using four optimization methods, ABC, GWO, WOA, and BOA, respectively

(a) XGBoost-ABC   (b) XGBoost-GWO

(c) XGBoost-WOA   (d) XGBoost-BOA
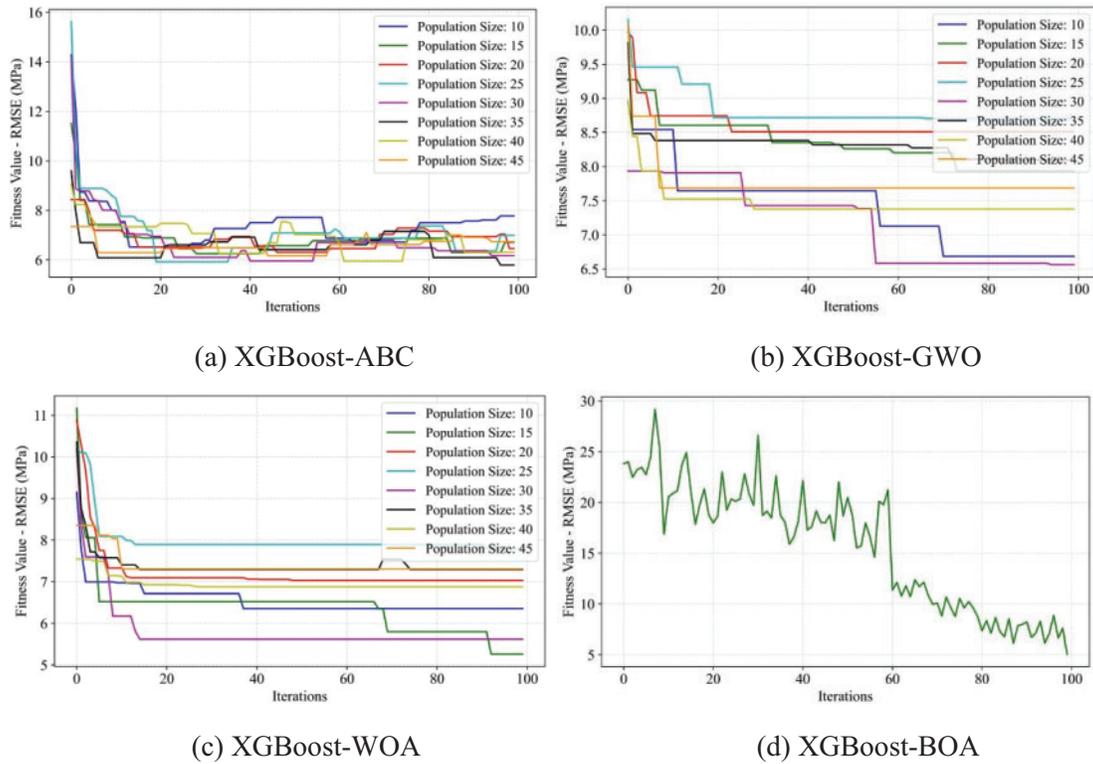
Figure 18: Fitness values of RMSE of the XGBoost model using four optimization methods, ABC, GWO, WOA, and BOA, respectively
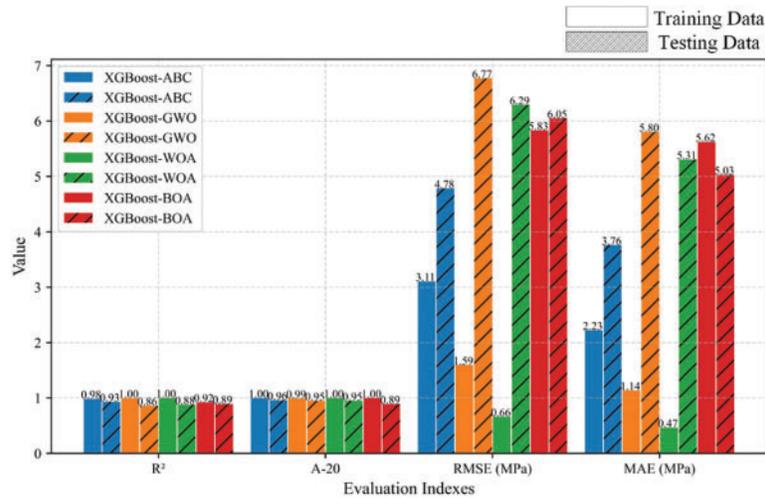

Figure 19: Comparison results of the prediction performance of the AdaBoost hybrid models

Table 4: The optimal parameters of the XGBoost method

Model	Hyperparameters	Optimal value	Search space
XGBoost	The number of estimators	168	1~200
	Maximum depth of decision trees	39	1~50
	Maximum features	7	1~10
	Colsample_bytree	1	0~1
	Regression alpha	0.796	0~1
	Subsample	0.493	0~1

6 Model Prediction Performance Analysis and Discussion

Taylor diagrams [70] were employed to discuss the predictive model's performance on the training and testing datasets. In Fig. 20, the markers indicate different models; the radial direction represents the correlation coefficient; the X-axis indicates the standard deviation, unnormalized; and the green dotted curves reflect the centered RMSE. The reference point with a black Pentastar indicates the actual UCS and other markers closer to this point denote better performance. The hybrid AdaBoost models show poor performance on both training and testing datasets. The hybrid GBoost and XGBoost models achieve acceptable results on the training datasets. However, hybrid XGBoost models perform better in the testing stage than GBoost models. In summary, XGBoost-ABC performs best compared to XGBoost-GWO, XGBoost-WOA, and XGBoost-BOA.
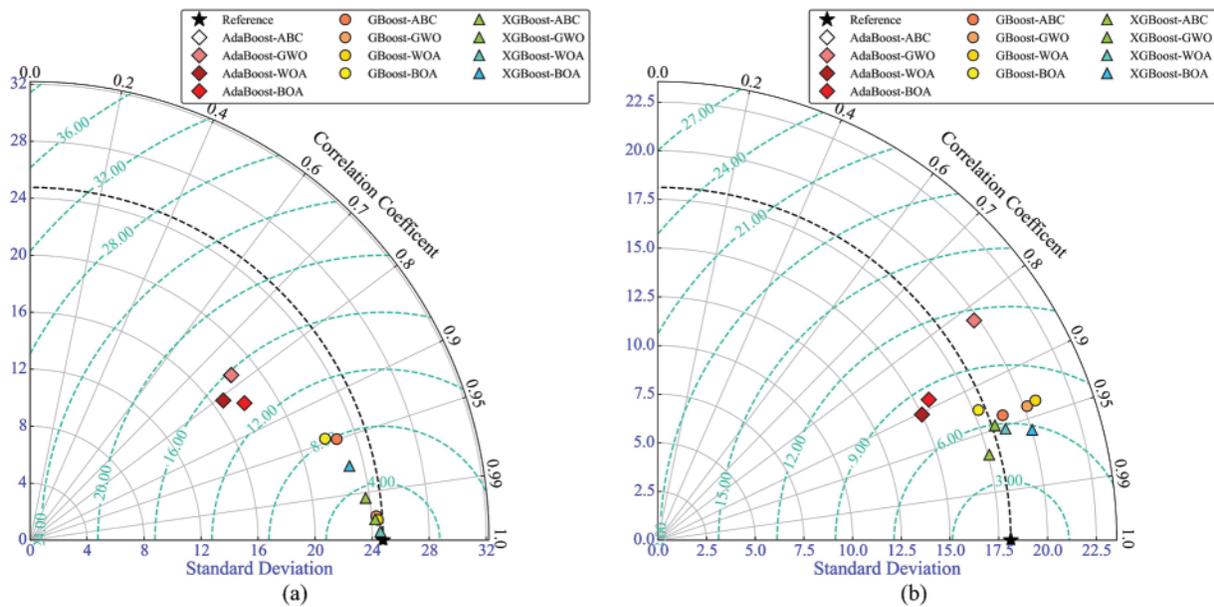

Figure 20: Taylor diagram. (a) Training data, (b) Testing data

Fig. 21 shows the prediction performance comparison of the optimal hybrid model for each boosting method. The results of the three evaluation indices for AdaBoost-ABC and GBoost-BOA are 0.60, 11.45, and 10.30 MPa; 0.87, 6.59, and 5.25 MPa, respectively. The XGBoost-ABC hybrid model

achieved the highest $R^2 = 0.93$ and the smallest RMSE and MAE (4.78 MPa and 3.76 MPa). The UCS prediction results of the hybrid model XGBoost-ABC on the testing dataset are shown in Fig. 22, where the red solid line represents the model prediction results, and the blue dotted line indicates the measured values.
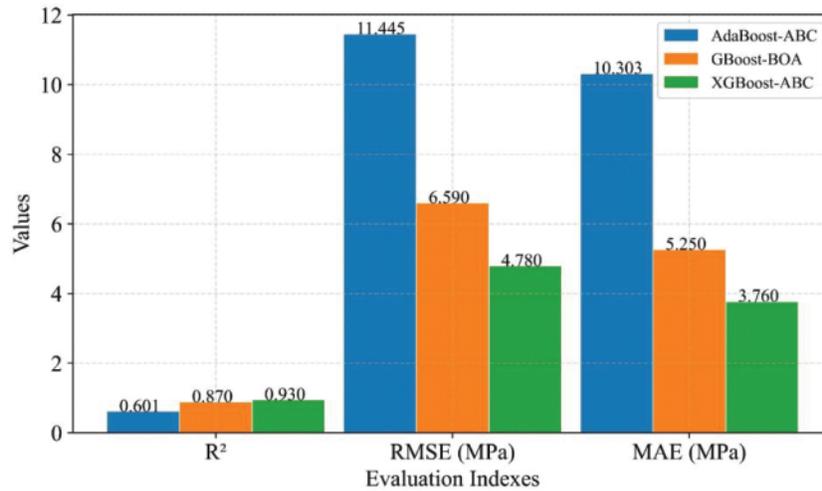

Figure 21: The performance comparison of the optimal hybrid model for each boosting method
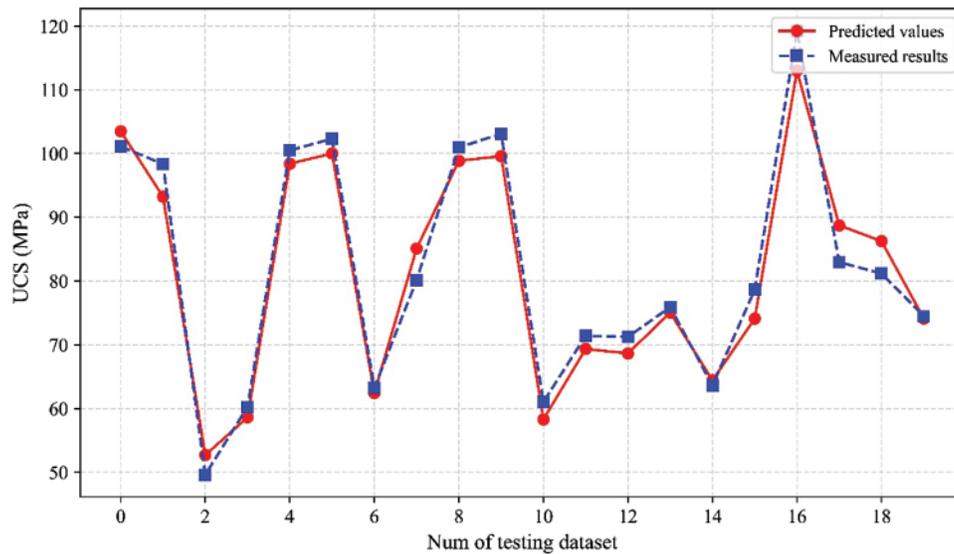

Figure 22: The prediction results of the hybrid XGBoost-ABC model on the testing dataset

Sensitivity analysis was employed to better understand the intrinsic relationships between the selected independent variables and rock UCS. The relevancy factor, a commonly used method to illustrate the sensitivity scale [71,72], was applied in this paper to assess the effect of each variable on UCS. The greater the absolute value of the relevancy factor between the independent and dependent variables, the stronger the influence. The calculation process of the sensitivity relevancy factor (SRF) is as follows:

$$SRF = \left| \frac{\sum_{i=1}^{n} (x_{li} - \overline{x_l}) \times (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_{li} - \overline{x_l})^2 \sum_{i=1}^{n} (y_i - \overline{y})^2}} \right| \tag{26}$$

where $\overline{x_l}$ is the mean value of all data for variable l (l includes S/B, H/B, B/D, T/B, Pf, X_B, and E); x_{li} is the i-th value of variable l; n is the number of the variable data; y_i and $\overline{y}$ are the i-th measured value of variable l and the average value of the prediction results, respectively.

The results showed that the most influential parameter on the UCS is the point load strength index (I_{s50}) of 1.0, after that is the P-wave velocity (V_P) of 0.3, the density (ρ) of 0.03. However, it should be noted that this importance ranking of all input variables is only for the data used in this study and cannot be used as a general criterion.

On the other hand, other intelligent algorithms, including random forest and artificial neural network (ANN), were trained with the same datasets to verify the superiority of the XGBoost-ABC model. The ANN model structure was 3-7-4-1, i.e., two hidden layers with 7 and 4 neurons, respectively. The prediction results of the two models are shown in Fig. 23, and the comparison results between the hybrid model XGBoost-ABC are presented in Table 5. The results demonstrate that the hybrid model XGBoost-ABC proposed in this paper performs better.
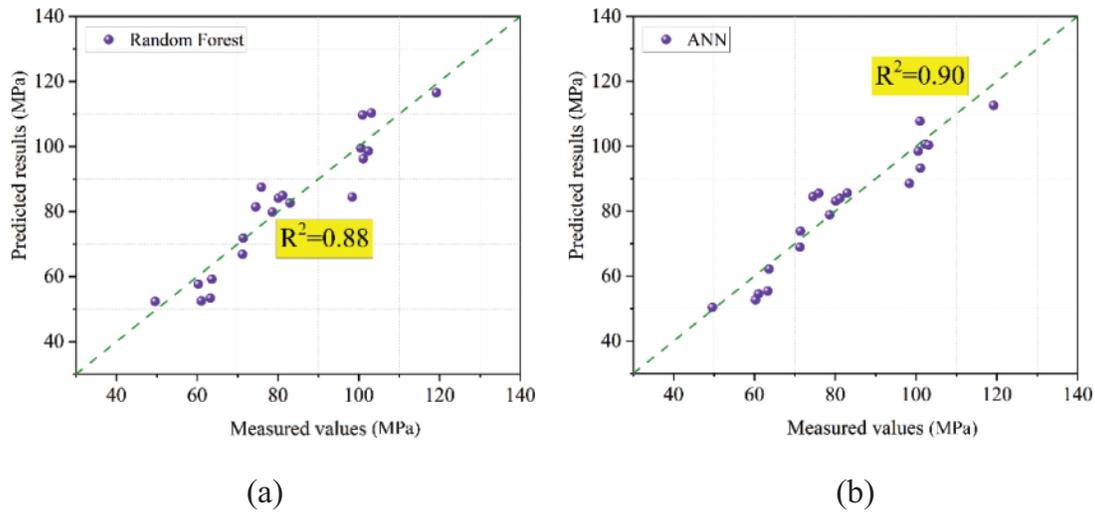

(a)   (b)

Figure 23: Performance of the random forest and ANN models using testing datasets. (a) is the prediction result of the random forest model, (b) is the prediction result of the ANN model

Table 5: Performance indicators of the predictive models

Predictive models	R^2	RMSE (MPa)	MAE (MPa)
XGBoost-ABC	0.93	4.78	3.76
Random forest	0.88	6.33	5.14
ANN (3-7-4-1)	0.90	5.71	4.73

Based on the findings mentioned above, the proposed model XGBoost-ABC achieved an acceptable UCS prediction result. Fig. 24 presents the developed Graphical User Interface (GUI), which

engineers can use as a portable tool to estimate the UCS of rock materials in deep mines. Nevertheless, it is essential to note that the developed model in this study is designed to address UCS prediction of rock in deep mining environments with three parameters: rock density, P-wave velocity, and point load strength.

Figure 24: GUI for the prediction of rock UCS

7 Conclusions and Limitations

In this research, a total of 106 samples are employed to investigate the mechanical properties of rocks in underground mines. Among them, 40 sets of data are taken from a deep lead-zinc mine in Southwest China, which can be regarded as a valuable database for investigating the mechanical properties of rocks in deep underground engineering. Three boosting-based models and four optimization algorithms are implemented to develop intelligent models for rock UCS prediction based on the established dataset. Based on the comparison results, it was found that the proposed hybrid model XGBoost-ABC exhibited superior prediction performance compared to the other models, with the highest R^2 values of 0.98 and 0.93, smallest RMSE values of 3.11 MPa and 4.78 MPa, and the smallest MAE values of 2.23 MPa and 3.76 MPa on the training and testing datasets, respectively.

Overall, the proposed hybrid model achieves promising prediction accuracy on the data presented in this study. However, it is suggested that the model be fine-tuned on other datasets to ensure model prediction accuracy. In addition, more real-world data can be supplemented to enhance the robustness of the model. Finally, other physical and mechanical parameters can also be considered to develop rock strength prediction models in the future.

Author Contributions: Junjie Zhao: Writing–original draft, coding, model training. Diyuan Li: Supervision, writing review. Jingtai Jiang, Pingkuang Luo: Data collection and process. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The datasets are available from the corresponding author upon reasonable request. https://github.com/cs-heibao/UCS_Prediction_GUI.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. ISRM. (1981). *Rock characterization testing and monitoring. ISRM suggested methods.* Oxford: Pergamon Press.

2. ASTM (1986). Standard test method of unconfined compressive strength of intact rock core specimens. ASTM Standard. 04.08 (D 2938).

3. Jamshidi, A. (2022). A comparative study of point load index test procedures in predicting the uniaxial compressive strength of sandstones. *Rock Mechanics and Rock Engineering, 55(7),* 4507–4516.

4. Cao, J., Gao, J., Nikafshan Rad, H., Mohammed, A. S., Hasanipanah, M. et al. (2022). A novel systematic and evolved approach based on xgboost-firefly algorithm to predict young's modulus and unconfined compressive strength of rock. *Engineering with Computers, 38(5),* 3829–3845.

5. Azimian, A., Ajalloeian, R., Fatehi, L. (2014). An empirical correlation of uniaxial compressive strength with P-wave velocity and point load strength index on marly rocks using statistical method. *Geotechnical and Geological Engineering, 32,* 205–214.

6. Barzegar, R., Sattarpour, M., Deo, R., Fijani, E., Adamowski, J. (2020). An ensemble tree-based machine learning model for predicting the uniaxial compressive strength of travertine rocks. *Neural Computing and Applications, 32,* 9065–9080.

7. Yılmaz, I., Sendır, H. (2002). Correlation of schmidt hardness with unconfined compressive strength and Young's modulus in gypsum from Sivas (Turkey). *Engineering Geology, 66(3–4),* 211–219.

8. Kahraman, S. A. I. R., Gunaydin, O., Fener, M. (2005). The effect of porosity on the relation between uniaxial compressive strength and point load index. *International Journal of Rock Mechanics and Mining Sciences, 42(4),* 584–589.

9. Fener, M. U. S. T. A. F. A., Kahraman, S. A. İ. R., Bilgil, A., Gunaydin, O. (2005). A comparative evaluation of indirect methods to estimate the compressive strength of rocks. *Rock Mechanics and Rock Engineering, 38,* 329–343.

10. Yaşar, E., Erdoğan, Y. (2004). Estimation of rock physicomechanical properties using hardness methods. *Engineering Geology, 71(3–4),* 281–288.

11. Yilmaz, I. (2009). A new testing method for indirect determination of the unconfined compressive strength of rocks. *International Journal of Rock Mechanics and Mining Sciences, 46(8),* 1349–1357.

12. Basu, A., Kamran, M. (2010). Point load test on schistose rocks and its applicability in predicting uniaxial compressive strength. *International Journal of Rock Mechanics and Mining Sciences, 47(5),* 823–828.

13. Khandelwal, M. (2013). Correlating P-wave velocity with the physico-mechanical properties of different rocks. *Pure and Applied Geophysics, 170,* 507–514.

14. Amirkiyaei, V., Ghasemi, E., Faramarzi, L. (2021). Estimating uniaxial compressive strength of carbonate building stones based on some intact stone properties after deterioration by freeze–thaw. *Environmental Earth Sciences, 80(9),* 352.

15. Yagiz, S. (2009). Predicting uniaxial compressive strength, modulus of elasticity and index properties of rocks using the Schmidt hammer. *Bulletin of Engineering Geology and the Environment, 68,* 55–63.

16. Nazir, R., Momeni, E., Armaghani, D. J., Amin, M. F. M. (2013). Prediction of unconfined compressive strength of limestone rock samples using L-type Schmidt hammer. *Electronic Journal of Geotechnical Engineering, 18,* 1767–1775.

17. Wang, M., Wan, W. (2019). A new empirical formula for evaluating uniaxial compressive strength using the schmidt hammer test. *International Journal of Rock Mechanics and Mining Sciences, 123,* 104094.

18. Minaeian, B., Ahangari, K. (2013). Estimation of uniaxial compressive strength based on P-wave and Schmidt hammer rebound using statistical method. *Arabian Journal of Geosciences, 6,* 1925–1931.

19. Farhadian, A., Ghasemi, E., Hoseinie, S. H., Bagherpour, R. (2022). Prediction of rock abrasivity index (RAI) and uniaxial compressive strength (UCS) of granite building stones using nondestructive tests. *Geotechnical and Geological Engineering, 40(6),* 3343–3356.

20. Mishra, D. A., Basu, A. (2013). Estimation of uniaxial compressive strength of rock materials by index tests using regression analysis and fuzzy inference system. *Engineering Geology, 160,* 54–68.

21. Zhu, J., Chang, X., Zhang, X., Su, Y., Long, X. (2022). A novel method for the reconstruction of road profiles from measured vehicle responses based on the Kalman filter method. *Computer Modeling in Engineering & Sciences, 130(3),* 1719–1735. https://doi.org/10.32604/cmes.2022.019140

22. Chen, Q., Xu, C., Zou, B., Luo, Z., Xu, C. et al. (2023). Earth pressure of the trapdoor problem using three-dimensional discrete element method. *Computer Modeling in Engineering & Sciences, 135(2),* 1503–1520. https://doi.org/10.32604/cmes.2022.022823

23. Ghasemi, E., Gholizadeh, H. (2019). Prediction of squeezing potential in tunneling projects using data mining-based techniques. *Geotechnical and Geological Engineering, 37,* 1523–1532.

24. Kadkhodaei, M. H., Ghasemi, E., Mahdavi, S. (2023). Modelling tunnel squeezing using gene expression programming: A case study. *Proceedings of the Institution of Civil Engineers-Geotechnical Engineering, 176(6),* 567–581.

25. Liu, Z. D., Li, D. Y. (2023). Intelligent hybrid model to classify failure modes of overstressed rock masses in deep engineering. *Journal of Central South University, 30(1),* 156–174.

26. Li, D., Zhao, J., Liu, Z. (2022). A novel method of multitype hybrid rock lithology classification based on convolutional neural networks. *Sensors, 22(4),* 1574.

27. Li, D., Zhao, J., Ma, J. (2022). Experimental studies on rock thin-section image classification by deep learning-based approaches. *Mathematics, 10(13),* 2317.

28. Kadkhodaei, M. H., Ghasemi, E., Sari, M. (2022). Stochastic assessment of rockburst potential in underground spaces using Monte Carlo simulation. *Environmental Earth Sciences, 81(18),* 447.

29. Kadkhodaei, M. H., Ghasemi, E. (2022). Development of a semi-quantitative framework to assess rockburst risk using risk matrix and logistic model tree. *Geotechnical and Geological Engineering, 40(7),* 3669–3685.

30. Ghasemi, E., Gholizadeh, H., Adoko, A. C. (2020). Evaluation of rockburst occurrence and intensity in underground structures using decision tree approach. *Engineering with Computers, 36,* 213–225.

31. Ghasemi, E., Kalhori, H., Bagherpour, R., Yagiz, S. (2018). Model tree approach for predicting uniaxial compressive strength and Young's modulus of carbonate rocks. *Bulletin of Engineering Geology and the Environment, 77,* 331–343.

32. Wang, M., Wan, W., Zhao, Y. (2020). Prediction of uniaxial compressive strength of rocks from simple index tests using random forest predictive model. *Comptes Rendus Mecanique, 348(1),* 3–32.

33. Jin, X., Zhao, R., Ma, Y. (2022). Application of a hybrid machine learning model for the prediction of compressive strength and elastic modulus of rocks. *Minerals, 12(12),* 1506.

34. Saedi, B., Mohammadi, S. D., Shahbazi, H. (2019). Application of fuzzy inference system to predict uniaxial compressive strength and elastic modulus of migmatites. *Environmental Earth Sciences, 78,* 1–14.

35. Li, J., Li, C., Zhang, S. (2022). Application of six metaheuristic optimization algorithms and random forest in the uniaxial compressive strength of rock prediction. *Applied Soft Computing, 131,* 109729.

36. Mahmoodzadeh, A., Mohammadi, M., Ibrahim, H. H., Abdulhamid, S. N., Salim, S. G. et al. (2021). Artificial intelligence forecasting models of uniaxial compressive strength. *Transportation Geotechnics, 27,* 100499.

37. Skentou, A. D., Bardhan, A., Mamou, A., Lemonis, M. E., Kumar, G. et al. (2023). Closed-form equation for estimating unconfined compressive strength of granite from three nondestructive tests using soft computing models. *Rock Mechanics and Rock Engineering, 56(1),* 487–514.

38. Liu, Q., Wang, X., Huang, X., Yin, X. (2020). Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data. *Tunnelling and Underground Space Technology, 106,* 103595.

39. Wang, S. M., Zhou, J., Li, C. Q., Armaghani, D. J., Li, X. B. et al. (2021). Rockburst prediction in hard rock mines developing bagging and boosting tree-based ensemble techniques. *Journal of Central South University, 28(2),* 527–542.

40. Liu, Z., Armaghani, D. J., Fakharian, P., Li, D., Ulrikh, D. V. et al. (2022). Rock strength estimation using several tree-based ML techniques. *Computer Modeling in Engineering & Sciences, 133(3),* 799–824. https://doi.org/10.32604/cmes.2022.021165

41. Liu, Z., Li, D., Liu, Y., Yang, B., Zhang, Z. X. (2023). Prediction of uniaxial compressive strength of rock based on lithology using stacking models. *Rock Mechanics Bulletin, 2(4),* 100081.

42. Zhang, Q., Hu, W., Liu, Z., Tan, J. (2020). TBM performance prediction with Bayesian optimization and automated machine learning. *Tunnelling and Underground Space Technology, 103,* 103493.

43. Chen, T., Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining*, pp. 785–794. San Francisco, CA, USA.

44. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W. et al. (2017). Lightgbm: A highly efficient gradient boosting decision tree. *Proceedings of the Advances in Neural Information Processing Systems*, pp. 3147–3155. Long Beach, CA, USA.

45. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, vol. 31, pp. 6638–6648. Montréal, QC, Canada.

46. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y. et al. (2015). xgboost: eXtreme gradient boosting. *R Package Version 0.4-2, 1(4),* 1–4.

47. Chang, Y. C., Chang, K. H., Wu, G. J. (2018). Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions. *Applied Soft Computing, 73,* 914–920.

48. Zhang, W., Wu, C., Zhong, H., Li, Y., Wang, L. (2021). Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geoscience Frontiers, 12(1),* 469–477.

49. Nguyen-Sy, T., Wakim, J., To, Q. D., Vu, M. N., Nguyen, T. D. et al. (2020). Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method. *Construction and Building Materials, 260,* 119757.

50. Pelikan, M., Goldberg, D. E., Cantú-Paz, E. (1999). BOA: The Bayesian optimization algorithm. *Proceedings of the Genetic and Evolutionary Computation Conference GECCO-99*, vol. 1, pp. 525–532. Orlando, FL, USA.

51. Díaz, E., Salamanca-Medina, E. L., Tomás, R. (2023). Assessment of compressive strength of jet grouting by machine learning. *Journal of Rock Mechanics and Geotechnical Engineering, 16,* 102–111.

52. Lahmiri, S., Bekiros, S., Avdoulas, C. (2023). A comparative assessment of machine learning methods for predicting housing prices using Bayesian optimization. *Decision Analytics Journal, 6,* 100166.

53. Bo, Y., Huang, X., Pan, Y., Feng, Y., Deng, P. et al. (2023). Robust model for tunnel squeezing using Bayesian optimized classifiers with partially missing database. *Underground Space, 10,* 91–117.

54. Díaz, E., Spagnoli, G. (2023). Gradient boosting trees with Bayesian optimization to predict activity from other geotechnical parameters. *Marine Georesources & Geotechnology,* 1–11.

55. Greenhill, S., Rana, S., Gupta, S., Vellanki, P., Venkatesh, S. (2020). Bayesian optimization for adaptive experimental design: A review. *IEEE Access, 8,* 13937–13948.

56. Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization. In: *Technical report-TR06*, Erciyes University, Engineering Faculty, Computer Engineering Department.

57. Bharti, K. K., Singh, P. K. (2016). Chaotic gradient artificial bee colony for text clustering. *Soft Computing, 20,* 1113–1126.

58. Asteris, P. G., Nikoo, M. (2019). Artificial bee colony-based neural network for the prediction of the fundamental period of infilled frame structures. *Neural Computing and Applications, 31(9),* 4837–4847.

59. Parsajoo, M., Armaghani, D. J., Asteris, P. G. (2022). A precise neuro-fuzzy model enhanced by artificial bee colony techniques for assessment of rock brittleness index. *Neural Computing and Applications, 34,* 3263–3281.

60. Zhou, J., Koopialipoor, M., Li, E., Armaghani, D. J. (2020). Prediction of rockburst risk in underground projects developing a neuro-bee intelligent system. *Bulletin of Engineering Geology and the Environment, 79,* 4265–4279.

61. Mirjalili, S., Mirjalili, S. M., Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software, 69,* 46–61.

62. Golafshani, E. M., Behnood, A., Arashpour, M. (2020). Predicting the compressive strength of normal and high-performance concretes using ANN and ANFIS hybridized with grey wolf optimizer. *Construction and Building Materials, 232,* 117266.

63. Shariati, M., Mafipour, M. S., Ghahremani, B., Azarhomayun, F., Ahmadi, M. et al. (2022). A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement. *Engineering with Computers, 38,* 757–779.

64. Mirjalili, S., Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software, 95,* 51–67.

65. Zhou, J., Zhu, S., Qiu, Y., Armaghani, D. J., Zhou, A. et al. (2022). Predicting tunnel squeezing using support vector machine optimized by whale optimization algorithm. *Acta Geotechnica, 17(4),* 1343–1366.

66. Tien Bui, D., Abdullahi, M. A. M., Ghareh, S., Moayedi, H., Nguyen, H. (2021). Fine-tuning of neural computing using whale optimization algorithm for predicting compressive strength of concrete. *Engineering with Computers, 37,* 701–712.

67. Nguyen, H., Bui, X. N., Choi, Y., Lee, C. W., Armaghani, D. J. (2021). A novel combination of whale optimization algorithm and support vector machine with different kernel functions for prediction of blasting-induced fly-rock in quarry mines. *Natural Resources Research, 30,* 191–207.

68. Momeni, E., Armaghani, D. J., Hajihassani, M., Amin, M. F. M. (2015). Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement, 60,* 50–63.

69. Lei, Y., Zhou, S., Luo, X., Niu, S., Jiang, N. (2022). A comparative study of six hybrid prediction models for uniaxial compressive strength of rock based on swarm intelligence optimization algorithms. *Frontiers in Earth Science, 10,* 930130.

70. Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres, 106,* 7183–7192.

71. Chen, G., Fu, K., Liang, Z., Sema, T., Li, C. et al. (2014). The genetic algorithm based back propagation neural network for MMP prediction in CO_2-EOR process. *Fuel, 126,* 202–212.

72. Bayat, P., Monjezi, M., Mehrdanesh, A., Khandelwal, M. (2021). Blasting pattern optimization using gene expression programming and grasshopper optimization algorithm to minimise blast-induced ground vibrations. *Engineering with Computers, 38,* 3341–3350.