Article

Prediction of Uniaxial Compressive Strength in Rocks Based on Extreme Learning Machine Improved with Metaheuristic Algorithm

Junbo Qiu [1], Xin Yin [1,*], Yucong Pan [1], Xinyu Wang [2] and Min Zhang [3]

1   School of Civil Engineering, Wuhan University, Wuhan 430072, China
2   Yellow River Engineering Consulting Co., Ltd., Zhengzhou 450003, China
3   Beijing Aidi Geological Engineering Technology Co., Ltd., Beijing 100144, China
*   Correspondence: yinxin_engineering@163.com

Abstract: Uniaxial compressive strength (UCS) is a critical parameter in the disaster prevention of engineering projects, requiring a large budget and a long time to estimate in different rocks or the early stage of a project. If predicted accurately, the UCS of rocks significantly affects geotechnical applications. This paper develops a dataset of 734 samples from previous studies on different countries' magmatic, sedimentary, and metamorphic rocks. Within the study context, three main factors, point load index, P-wave velocity, and Schmidt hammer rebound number, are utilized to estimate UCS. Moreover, it applies extreme learning machines (ELM) to map the nonlinear relationship between the UCS and the influential factors. Five metaheuristic algorithms, particle swarm optimization (PSO), grey wolf optimization (GWO), whale optimization algorithm (WOA), butterfly optimization algorithm (BOA), and sparrow search algorithm (SSA), are used to optimize the bias and weight of ELM and thus enhance its predictability. Indeed, several performance parameters are utilized to verify the proposed models' generalization capability and predictive performance. The minimum, maximum, and average relative errors of ELM achieved by the whale optimization algorithm (WOA-ELM) are smaller than the other models, with values of 0.22%, 72.05%, and 11.48%, respectively. In contrast, the minimum and mean residual error produced by WOA-ELM are less than the other models, with values of 0.02 and 2.64 MPa, respectively. The results show that the UCS values derived from WOA-ELM are superior to those from other models. The performance indices (coefficient of determination (R^2): 0.861, mean squared error (MSE): 17.61, root mean squared error (RMSE): 4.20, and value account for (VAF): 91% obtained using the WOA-ELM model indicates high accuracy and reliability, which means that it has broad application potential for estimating UCS of different rocks.

Keywords: uniaxial compressive strength; prediction model; extreme learning machine; metaheuristic algorithm

MSC: 68T99

1. Introduction

Uniaxial compressive strength (UCS) plays a vital role in rock engineering projects from design to construction and operation. In general, the UCS can be obtained by conducting laboratory tests using the approaches provided by the International Society of Rock Mechanics (ISRM2007) and the American Society for Testing Materials (ASTM2001a) [1]. However, the uniaxial compression test requires high quality and strict specimen size. Therefore, obtaining a core sample from soft, weak, highly weathered, or fragile rocks is almost impossible. In addition, direct estimation of UCS in the laboratory is costly, complicated, and time-consuming [2]. Therefore, the precise prediction of UCS is a challenge.

As a result, proposing a method for obtaining UCS conveniently and quickly to overcome associated problems and save time and cost is vital.

There are three methods to determine the UCS, including empirical formulation, multiple regression analysis, and soft computing modeling. Some empirical models with non-destructive test results to estimate UCS were proposed to overcome the difficulty in preparing core specimens. Several researchers investigated the relationship between UCS and other physical properties of rock mass, such as Brazilian tensile strength [3], point load strength index [4–6], slake durability index [7], Schmidt hammer rebound number [8,9], and P wave velocity [9–11]. The empirical formulas derived using these techniques are often applied to the sampling area or the same rock type. Empirical formulas include multiple fitting forms but usually consider a single factor, ignoring the effects of multiple factors. Other researchers proposed fuzzy and multiple regression analysis [12–15] to obtain UCS, hence controlling the aforementioned issues. However, these methods cannot solve the nonlinear relationship between UCS and other rock parameters; consequently, the soft computing method was presented to address this issue. Sarkar et al. [16] proposed an artificial neural network (ANN) model to estimate the UCS using slake durability index, dynamic wave velocity, density, and point load index. Yagiz et al. [17] predicted UCS using an ANN model and nonlinear technique. They discovered that ANN models are more accurate in determining UCS than regression techniques. In addition, Yesiloglu et al. [18] developed an adaptive neuro-fuzzy inference system (ANFIS) and an ANN to predict UCS, considering tensile strength, point load index, block punch index, and P-wave velocity as input parameters. They indicated that the performance evaluation of the ANFIS model was more precise than others. Gene expression programming (GEP) [19] and Multilayer Perceptron Neural Network (MLPNN) [20] were utilized to estimate UCS. Li and Tan [21] suggested a least squares vector machine for the UCS prediction model. Nevertheless, Mahmoodzadeh et al. [22] utilized machine learning methods to predict UCS, proving that Gaussian process regression (GPR) performed best. Gupta and Natarajan [23] assessed the ability of density-weighted least squares support vector machine, extreme learning machine (ELM), and random forest (RF) to estimate UCS of rocks and concluded that an improved unique machine learning model has a better predictive capability than other normal models. Recently, a comprehensive model has been combined with the ANN model and particle swarm algorithm (PSO) to predict UCS [24]. Fang et al. [25] also put effort into developing two comprehensive predictive models using hybrid ANN with a imperialism competitive algorithm (ICA) and artificial bee algorithm (ABC).

These methods are valuable for determining UCS with rock physical properties obtained by non-destructive tests. However, empirical formula measures performed unique effects with different factors. Multiple regression models cannot map the nonlinear relationship between UCS and influence factors. Machine learning models have a better predictive capability to estimate UCS than traditional models. Support vector machine (SVM) and Radial basic neural network (RBF) had good performance with small data [26,27]. However, the weight and bias of ANN and the hyper-parameters of the machine learning model demand optimization and have some constraints, such as falling into local minimum and including a low learning rate [28]. The ELM is a single hidden layer feedforward neural network introduced by Huang [29]. Some previous studies revealed that ELM is better than ANN and SVM in overcoming low learning rates and local minimum problems of regression analysis [30]. Therefore, the ELM is used to map the nonlinear relationship between UCS and the influential factors. Meanwhile, the ELM requires optimization algorithms to achieve improved performance. The metaheuristic algorithms inspired by the natural behavior of animals have good performance [31]. Additionally, the datasets for these associated measures for obtaining UCS come from the same area or rock type and are short datasets. The simple data mining methods normally do not provide the required efficiency for small data [32,33]. Hence, a bigger dataset must be established to estimate UCS. Accordingly, the present study aims to develop a new forecast model that estimates

UCS using a dataset of various rocks collected from previous research based on an ELM coupled metaheuristic algorithm.

The main contribution of this paper can be summarized as follows:

1. Collecting a dataset of 734 samples from previous studies of magmatic rocks, sedimentary rocks, and metamorphic rocks in different countries to overcome the problem of requiring a large budget and a long time to estimate UCS in different rocks or at the early stage of a project.
2. Optimizing the hidden neurons and activation function between ELM to map the nonlinear relationship between the UCS and the non-destructive test indices.
3. Utilizing five metaheuristic algorithms (PSO, GWO, WOA, BOA, and SSA) to estimate UCS.
4. Comparing the optimized model to other techniques to prove efficiency.

The remainder of this paper is organized as follows: Section 2 contains the characteristics and visualization of the dataset; Section 3 describes the mathematical relationships of the ELM and metaheuristic algorithm; Section 4 describes the optimization procedures of the ELM optimized by PSO, GWO, WOA, BOA, and SSA; Section 5 contains the statistical evaluation indices of the models; Section 6 summarizes the results of this work and compares the proposed models' effectiveness with other approaches; Section 7 contains the conclusions and recommendations for future research.

2. Dataset

One of the drawbacks of previous studies is that they mainly focused on datasets that are based on a single rock type. Accordingly, this study collects 734 magmatic, sedimentary, and metamorphic rock samples in a single dataset (see Supplementary Materials) to develop the prediction models. Some of these data points are rocks from quarries and natural outcrops in Turkey [14,34–38] and Iran [1,39], while others are natural outcrops and tunnels in India [15,40], Malaysia [24,41,42], and China [43]. Previously, the tensile strength, point load index (I_s), block punch index, density, porosity, and P-wave velocity were utilized as inputs to the numerical models for estimating the UCS. Some studies used the point load index, P-wave velocity (V_p), and Schmidt hammer rebound number (SR_n) to estimate UCS. Accordingly, this study collects these non-destructive test results when developing the UCS dataset. Considered ranges of UCS and influence factors are provided in the Table 1 and Figure 1. Table 1 shows brief descriptive statistics of the dataset used in this research. The SR_n ranges from 10 to 72, the maximum value of V_p is 4675 m/s, and its min value is 375 m/s. In addition, the I_s value ranges from 0.53 MPa to 23.10 MPa. Moreover, the UCS ranges from 2.03 MPa to 239 MPa.

Table 1. Brief descriptive statistics of the dataset.

	SR_n	V_p (m/s)	I_s (MPa)	UCS (MPa)
Minimum	10	375	0.53	2.03
Maximum	72	7943	23.10	239.00
Average	42	4675	4.33	75.05
Standard deviation	11.83	1383.14	3.01	44.70

Figure 1 shows a visualization of the collected dataset. There is a wide distribution of attributes and UCS in the dataset, which means that the collected data include a wide range of rock types. Meanwhile, it can be noticed that the Pearson's correlation coefficients between the UCS and SR_n, V_p exceed 0.64, indicating strong correlations. The correlation between the UCS and I_s is 0.42. Moreover, the correlations between the input parameters range between 0.22 and 0.51, showing slight interactions.
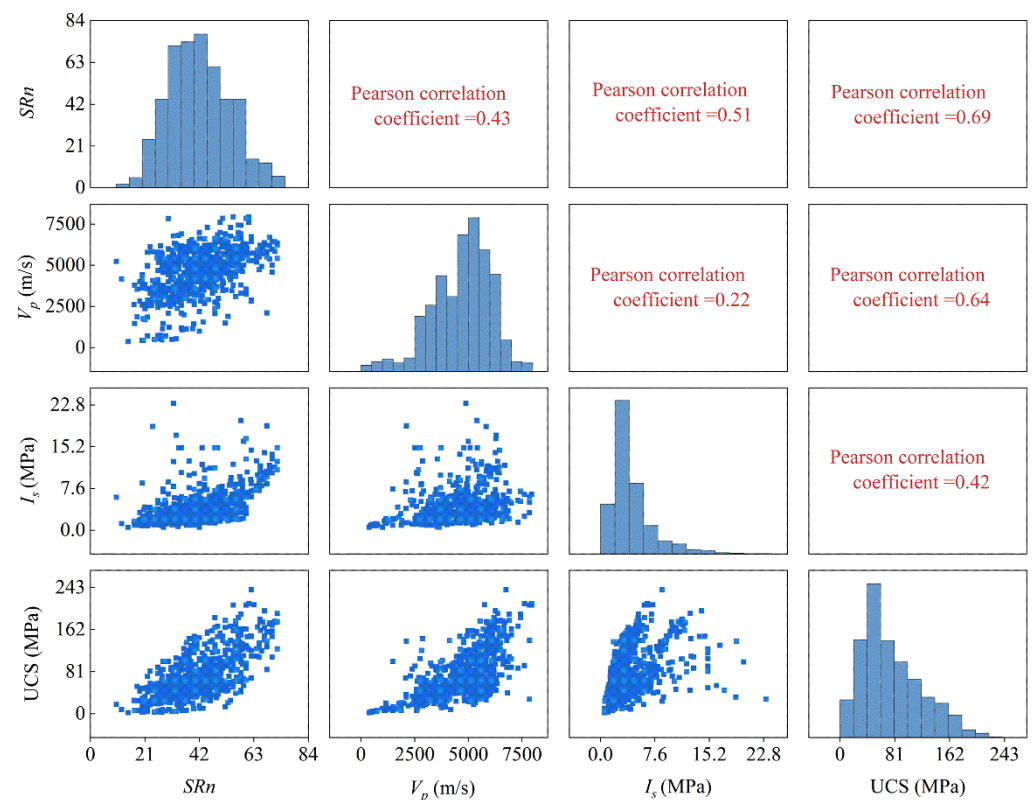
Figure 1. Visual illustration of the collected dataset.

3. Methods

UCS is a critical parameter for rock-mechanic-related investigations in civil, mining, and petroleum projects. However, experimentally evaluating this parameter is rather expensive, complicated, and time-consuming. As a result, previous investigations tended to develop soft-computing models for rapid UCS estimation. Indeed, this study aims to propose a generalized numerical model based on the wide-range dataset present to overcome the complexity of the test procedures. Within the study context, an ELM is used for mapping the nonlinear relationship between the UCS and the non-destructive test indices, and a metaheuristic algorithm is utilized to enhance the prediction ability of the ELM.

3.1. Extreme Learning Machine

The ELM is a single hidden layer feedforward neural network introduced by Huang [29]. The ELM was proposed to solve the time-consuming training problem in feedforward backpropagation neural networks. Similar to other feedforward neural networks, ELM has an input, a hidden, and an output layer, as depicted in Figure 2.

For a data set R of D arbitrary distinct training samples $R = \{(x_i, t_i)|i = 1, 2, 3, \ldots, D\}$, where $x_i = [x_{i1}, x_{i2}, \ldots x_{iD}]^T$ and $t_i = [t_{i1}, t_{i2}, \ldots t_{iD}]^T$ are the inputs and output, ELM mathematic model is defined by Equation (1).

$$o_i = \sum_{i=1}^{L} \beta_i g(x_i) = \sum_{i=1}^{L} \beta_i g(m_i x_i + n_i) \tag{1}$$

where o_i is the output vectors, $g(x)$ is the active function, typically defined as a sigmoid, sine, or hardlim function as shown in Figure 3, m_i is the connection weights between the hidden layer and the input layer node, n_i is the threshold between the hidden layer and the input layer node, β_i is the weight vector between the hidden layer and output layer nodes, and L is the hidden modes.
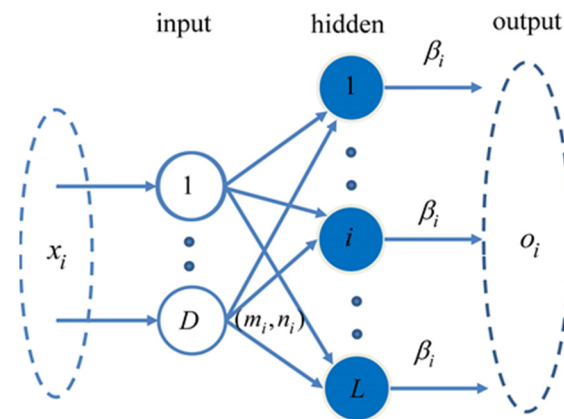
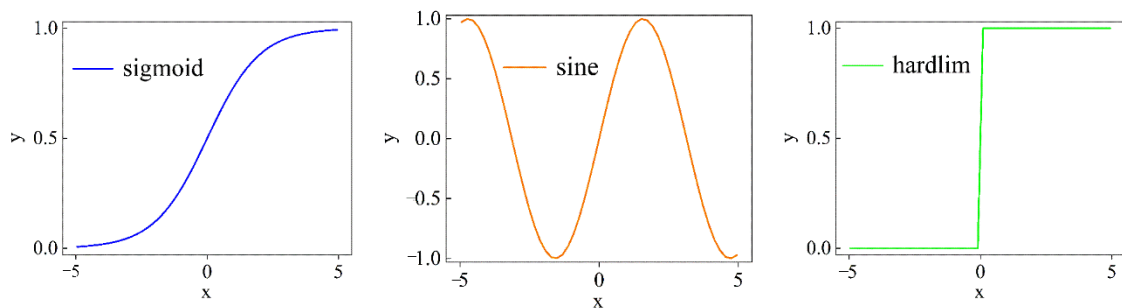**Figure 2.** Flowchart of the ELM architecture.

Figure 3. Various types of activation functions of ELM.

According to the two theorems proposed by Huang, when $g(x)$ is infinitely differentiable, the ELM with L hidden nodes and activation function can be fit to achieve a zero-error approximation of any D samples. Hence, Equation (2) is established.

$$\sum_{j=1}^{L}||o_i - t_i|| = 0 \tag{2}$$

According to Equation (2), there exists specific m_i, n_i, and β_i to make the formula (3) hold.

$$\sum_{i=1}^{L}\beta_i g(m_i x_i + n_i) = t_i \tag{3}$$

The Equation (3) can be simplified in the form of $H\beta = T$, where H is the output matrix of the hidden layer of ELM, T is the target matrix. Unlike traditional gradient-based learning algorithms with fixed input weights and hidden layer bias, the ELM theories claim that the parameters m_i and n_i can be assigned randomly. Then, the issue for training the ELM is transformed into finding a least-square solution. The solution of Equation (3) in the matrix form is defined in Equation (4). According to the two theorems proposed by Huang, when the number of samples and hidden modes is the same, Equation (2) can be established. Therefore, the sample size of the dataset is normally much larger than the number of hidden neurons, and the pseudo-inverse of matrix H is required.

$$\hat{\beta} = H^+ T = (H^T H)^{-1} H^T T \tag{4}$$

where H^+ is the Moore-Penrose generalized inverse of H.

Compared to the traditional intelligent algorithm, ELM can be rapidly trained by determining the number of hidden layers. Indeed, the ELM solves the shortcomings of the backpropagation gradient descent method represented by easily falling into local minima.

When solving the weights of the hidden and output layers, a mathematical method with uniqueness and global optimality is used. Therefore, the ELM has superior performance, yet shows some shortcomings. The selection of the number of hidden neurons and the activation function for an ELM model typically follows an iterative approach without a theoretical basis. For practical problems, the network topology and functions optimization require an experienced designer or lots of repeated trials, which increases its application difficulties. In the ELM calculation process, the weights and thresholds of the hidden layer and output layer are calculated using straightforward mathematical methods. However, the weights and biases of an ELM model's hidden and input layer are randomly initialized between 0 and 1. The random weights and thresholds restrict the mapping performance of the ELM. Therefore, the enhancement in the ELM is represented by improving the threshold of random weights generated through the original network, improving the stability of the network, and fully distilling the nonlinear relationship between input and output. In this study, various intelligent algorithms are proposed to improve the shortcomings of the ELM model.

3.2. Particle Swarm Algorithm (PSO)

The PSO algorithm is a bionic intelligent optimization algorithm proposed by Kennedy and Eberhart [44] in the 1990s. In this algorithm, each solution of the optimization problem is simplified to a particle, i.e., a bird swarm individual. The algorithm-solving process aims to find food for each bird swarm individual through group collaboration. The mathematical model of the PSO algorithm is as follows: based on the problem's type, the initial population is set in the D-dimensional search space, and the position and velocity of particles are determined by $pbest_{id}^{t}$ and $gbest_{d}^{t}$. The selection of $pbest_{id}^{t}$ and $gbest_{d}^{t}$ is intended to move each particle to a different point in the solution area. Finally, the optimal solution is obtained through a continuous change of velocity and position. Equations (5) and (6) are used to update the velocity and position, respectively.

$$v_{id}^{t+1} = v_{id}^{t} + c_1 r_1 (pbest_{id}^{t} - x_{id}^{t}) + c_2 r_2 (gbest_{d}^{t} - x_{id}^{t+1}) \tag{5}$$

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t} \tag{6}$$

where v_{id}^{t+1} is the particle velocity in the $t + 1$ generation, c_1 and c_2 are constants between $(0, 2)$, r_1 and r_2 are constants between $(0, 1)$, t is the iteration time, and x_{id}^{t+1} is the position of a particle in the $t + 1$ generation.

3.3. Grey Wolf Optimization (GWO)

Grey wolf optimization is a meta-inspired algorithm that simulates the hunting behaviors of grey wolves. It was proposed by Mirjalili [45] in 2014. In this method, the wolves are divided into α wolf, β wolf, γ wolf, and ω wolf according to their fitness from high to low. The wolf of α, β, and γ leadership search and locate the prey. With the wolf group evolution, the distance of the prey is reduced, and the ω wolf is guided to track and capture the prey. The implementation of the grey wolf algorithm is shown in the following steps:

Step 1, surround the prey. Identify and surround the prey before preying. The following three equations show the distance and updating formulas between the wolf and prey in each grade of the grey wolf group.

$$D = |E \cdot X_p(t) - X(t)| \tag{7}$$

$$X(t + 1) = X_p(t) - A \cdot E \tag{8}$$

$$\begin{aligned} A &= 2 \cdot a \cdot r_2 - a \\ E &= 2r_1 \end{aligned} \tag{9}$$

where X_P and $X(t + 1)$ are the location of prey when the number of iterations is t and $t + 1$, respectively, $X(t)$ is the position of a grey wolf when the number of iterations is t, A and E

are the convergence vector and coefficient vector, respectively, *a* linearly decreases from 2 to 0, respectively, and r_1 and r_2 are constants between (0, 1).

Step 2, hunt for prey. Once the prey is surrounded, the wolves begin to hunt. The optimal, sub-optimal, and third-optimal solutions are α, β, and γ wolves according to the fitness ranking. Their positions are updated as shown in Equation (8). The first three grades of wolves guide the other wolves, and Equation (9) is the update mode.

$$
\begin{aligned}
D_\alpha &= |C \cdot X_\alpha(t) - X(t)|, X_1 = X_\alpha(t) - A_1 \cdot D_\alpha \\
D_\beta &= |C \cdot X_\beta(t) - X(t)|, X_2 = X_\beta(t) - A_2 \cdot D_\beta \\
D_\gamma &= |C \cdot X_\gamma(t) - X(t)|, X_3 = X_\gamma(t) - A_3 \cdot D_\gamma
\end{aligned}
\tag{10}
$$

$$
X(t+1) = \frac{X_1 + X_2 + X_3}{3}
\tag{11}
$$

where $X(t+1)$ is the position update of ω wolves, D_α, D_β and D_γ are the distance update of α, β, and γ wolves and prey, respectively, and X_1, X_2 and X_3 are the position update of α, β, and γ wolves, respectively.

Step 3, attack prey. Similar to the last two steps, the wolf attacks when the prey is exhausted. The mathematical model can be expressed as follows (10) and (11), where A is a random number in the range of $[-2a, 2a]$. When A is outside $[-1, 1]$, it enhances ergodicity. In order to approach prey and reduce the value of *a*, *A* will decrease, and when *A* is within $[-1, 1]$, the grey wolf group attacks.

3.4. Whale Optimization Algorithm (WOA)

The whale optimization algorithm is a nature-inspired algorithm mimicking the motion of whales when hunting their prey. It was first developed by Mirjalili and Lewis [46] to solve optimization problems. The algorithm simulates the actions of the humpback whale in searching the prey and the bubble-net feeding method of encircling prey. The mathematical model of a whale's unique action is following:

3.4.1. Encircling Prey

The humpback whale can recognize the location of prey when they enter the target area or perception space. WOA assumes that the best position (solution) is the target prey. Once the best search agent is proposed, the rest agents try to update their location toward the best position (solution) as described in (12)–(14).

$$
DI = |C \cdot W_p(t) - W(t)|
\tag{12}
$$

$$
W(t+1) = W_p(t) - K \cdot DI
\tag{13}
$$

$$
\begin{aligned}
K &= 2 \cdot k \cdot r_2 - k \\
C &= 2r_1
\end{aligned}
\tag{14}
$$

where *t* is the current iteration, $W(t)$ indicates the position of prey, *K* and *C* are coefficient vectors, W_p is the position of the optimal solution, *k* is a variable linearly decreasing from 2 to 0, and r_1 and r_2 are constants between (0, 1).

3.4.2. Bubble-Net Attack Method

This section mainly introduces the shrinking encircling mechanism and spiral update position. First, the value of *K* is changed with *k* decreases by Equation (14) to achieve the shrinking encircling mechanism, and the whales' positions are updated according to Equations (15) and (16). As the whales are close to the prey (best solution), the distance between the whales and the prey can be calculated. A spiral update equation is then created to mimic the helix-shaped movement of the whales as follows:

$$
W(t+1) = DSe^{bt} \cos(2\pi t) + W_p(t)
\tag{15}
$$

$$DS = W_p(t) - W(t) \tag{16}$$

where DS is the distance between the whale and the prey (current best solution), b is a constant which defines the shape of the logarithmic spiral, and t is a constant between (0, 1).

According to the previous equation, the whale can have two strategies to move close to the prey. The mathematical equation is as follows:

$$W(t+1) = \begin{cases} W_p(t) - K \cdot DI & if \ p < 0.5 \\ DSe^{bt}\cos(2\pi t) + W_p(t) & if \ p > 0.5 \end{cases} \tag{17}$$

where p is a random number in (0, 1).

3.4.3. Exploration Phase

Humpback whales randomly search the prey based on the constant variation to obtain the best solution. This process is mathematically described as follows:

$$DI = |C \cdot W_{rand}(t) - W(t)| \tag{18}$$

$$W(t+1) = W_{rand}(t) - K \cdot DI \tag{19}$$

where $W_{rand}(t)$ represents the random whale in the current population.

3.5. Butterfly Optimization Algorithm (BOA)

Inspired by the living habits of butterflies in nature, a butterfly optimization algorithm [47] (BOA) was proposed to simulate butterflies' foraging and mating behaviors. Unlike other metaheuristic algorithms, this method's advantage is that each butterfly has its unique odor. The butterfly can perceive and analyze the odor in the air to determine the potential direction of food sources/mating partners. In BOA, the fragrance is formulated as a function of the stimulus's physical strength, as follows:

$$F = cI^a \tag{20}$$

where F is the concentration of aroma emitted by butterflies, c is the sensory mode, I is the stimulus intensity, and a is the power index dependent on the mode, indicating different absorption degrees of aroma among different butterflies.

In most cases, it is possible to define a and c within the range of [0, 1]. When a is 1, the butterfly does not absorb the fragrance. That is, another butterfly perceives the amount of fragrance emitted by a specific butterfly at the same capacity.

The BOA algorithm is divided into three parts, and the detailed steps are as follows:

Initializes the butterfly population by randomly generating the butterfly position in the search space and calculating and storing each butterfly's fragrance and fitness value. The fitness values of randomly generated butterfly populations are sorted to store the butterfly in the best position. Butterflies move toward the best position. The position update equation is as follows:

$$X_i^{t+1} = X_i^t + (r^2 + p_{best}^t - X_i^t) \bullet F(X_i) \tag{21}$$

where r is the random number in (0, 1), indicating the best butterfly for current iterations t, and $F(X_i)$ represents the aroma fitness value of the first butterfly at the current iteration number.

The mathematical model of butterfly population local search stage is as follows

$$X_i^{t+1} = X_i^t + (r^2 + p_{r_1}^t - p_{r_2}^t) \bullet F(X_i) \tag{22}$$

where $p_{r_1}^t$ and $p_{r_2}^t$ represent random two butterfly locations for the tth iteration in the search space, and r_1 and r_2 are random numbers between (0, 1).

3.6. Sparrow Search Algorithm (SSA)

Inspired by the group wisdom, foraging and anti-predation behaviors of the sparrow in nature, Xue [48] proposed the sparrow search algorithm to solve optimization problems. In the SSA, there are two types of sparrows: producer and scrounger. The producers with high levels of energy reserves can search for food sources and guide the movement of the entire population. The position update equation is as follows:

$$
X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot exp \left(\frac{-i}{\alpha \cdot iter_{\max}} \right) & if \ R_2 < ST \\ X_{i,j}^t \cdot Q \cdot L & if \ R_2 \geq ST \end{cases} \tag{23}
$$

where t is the current iteration, $iter_{\max}$ is the maximum number of iterations, $X_{i,j}^t$ and $X_{i,j}^{t+1}$ indicates the position of a sparrow, i is the number of sparrows, j is the dimension of the optimization problem, α is the random number in $(0, 1)$, $R_2 (R_2 \in [0, 1])$ is the alarm value, represents the safety threshold, Q is the random number which obeys normal distribution, and L is a matrix in which each element inside is 1.

When producers expand the search range to find foods without predators threatening and enter the wide search mode, if $R_2 \geq ST$ the sparrows quickly move to safe areas when predators move close to them.

As for the scroungers, if they detect that the producer has found good food, they immediately move the objective position to get food. On the one hand, if scroungers defeat the producer, the update formula is as shown in Equation (23). In contrast, if the producer wins, the scrounger enforces Equation (24).

$$
X_{i,j}^{t+1} = \begin{cases} Q \cdot exp \left(\frac{X_{worst}^t - X_{i,j}^t}{i^2} \right) & if \ i > n/2 \\ X_p^{t+1} + \left| X_{i,j}^t - X_p^{t+1} \right| \cdot A^+ \cdot L & otherwhise \end{cases} \tag{24}
$$

where X_p^{t+1} is the optimal position of the producer, X_{worst}^t is the current global worst position, $A^+ = A^T (AA^T)^{-1}$ and A is a one-dimensional matrix with each element randomly assigned -1 or 1, when $i > n/2$, the scrounger with the worst fitness value cannot find the food.

In SSA, some sparrows, which account for 10% or 20% of the total population, are assumed to be aware of the danger. In such a case, sparrows at the edge of the group quickly move forward to the safety area to get a better position, and other sparrows in the middle group move to others. The mathematical model can be expressed by:

$$
X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \cdot \left| X_{i,j}^t - X_{best}^t \right| & if \ f_i > f_g \\ X_{i,j}^t + \lambda \cdot \left(\frac{\left| X_{i,j}^t - X_{worst}^{t+1} \right|}{(f_i - f_w + \delta)} \right) & if \ f_i = f_g \end{cases} \tag{25}
$$

where X_{best} is the current global best position, β is a step control parameter that obeys the normal distribution of random numbers with a mean value of 0 and a variance of 1, λ is a random number in $(0, 1)$, f_i, f_g and f_w are the fitness value of the current sparrow, current best fitness, and current worst values, respectively, and δ is the smallest constant to avoid zero-division-error.

4. ELM Optimized by PSO, GWO, WOA, BOA, and SSA

This study uses ELM to map the nonlinear relationship between influence factors and UCS. However, the weights and thresholds of the hidden and input layers of the ELM algorithm are random numbers between 0 and 1, which can cause problems. The random weights and thresholds restrict the mapping performance of ELM. To obtain a reliable prediction, it is essential to improve the predictability of ELM. Hence, as an optimizer of weights and thresholds between inputs and hidden layers, PSO, GWO, WOA, BOA, and

SSA are utilized in this study. The development of the optimized ELM to predict UCS has the following steps:

1. To considerably distill the information governing the relationship between the UCS and the input variables, a database including 734 samples was developed in this study and divided into 700 samples for the training and 34 for the testing.
2. Firstly, the training set is used to optimize the ELM model's hidden layer neurons and activation function. After that, the test set is input into the trained ELM, and the obtained results are used to compute the performance metrics, including the root mean squared error (RMSE). The optimized hidden neurons and activation function can be determined when the RMSE is minimized.
3. To enhance the ELM model predictability, the PSO, GWO, WOA, BOA, and SSA are utilized to optimize weights and thresholds between inputs and hidden layers. Figure 4 depicts a process for optimizing ELM using the multi-algorithm.
4. Compare the predicted results and calculate the statistical evaluation indices to select the most precise and reliable model.
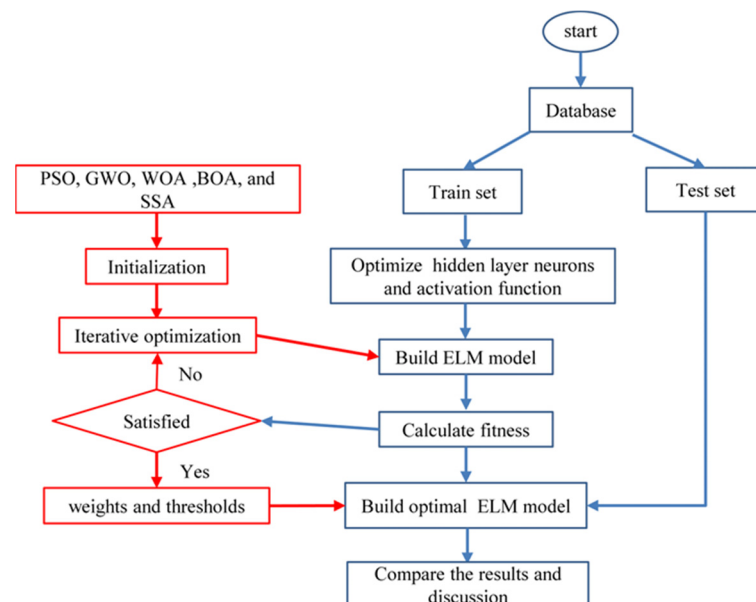

Figure 4. The process for optimizing the ELM using the PSO, GWO, WOA, BOA, and SSA.

5. Statistical Evaluation Indices

In order to evaluate the accuracy of the proposed prediction models, some statistical indices, including root mean squared error (RMSE), coefficient of determination (R^2), amount of value account for (VAF), and mean squared error (MSE), are calculated using Equations (26)–(29).

$$\text{RMSE} = \sqrt{\left(\frac{1}{n}\right)\sum_{k=1}^{n}(y_i - y_i')} \tag{26}$$

$$R^2 = 1 - \frac{sum\ squared\ \text{regression (SSR)}}{sum\ \text{of square total (SST)}} \tag{27}$$

$$\text{VAF} = \left[1 - \frac{var(y_i - y_i')}{var(y_i)}\right] \times 100\% \tag{28}$$

$$\text{MSE} = \left(\frac{1}{n}\right)\sum_{k=1}^{n}(y_i - y_i') \tag{29}$$

where y_i is the measured value, y_i' is the predicted value, and n is the number of observations.

6. Calculation Results and Discussion

6.1. ELM Parameters Optimization

As previously stated, the variables had distinct units and wide distribution. The data should be normalized to a value between 0 and 1 before training based on ELM to get good performance, as shown in Equation (30).

$$x_n = \frac{x_a - x_{\min}}{x_{\max} - x_{\min}} \tag{30}$$

where x_n is the normalized value, x_a is the actual value, $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of the dataset.

The hidden layer neurons and activation function must be optimized for the ELM. The RMSE of ELM was utilized to predict a model to tune them. Table 2 and Figure 5 indicate the effects of the number of hidden layer neurons and activation function.

Table 2. Effects of the number of hidden modes on the ELM performance.

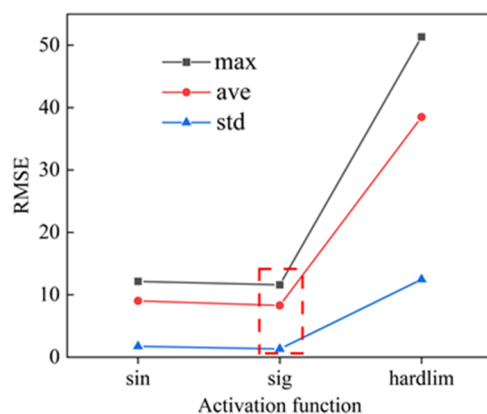| Number of Hidden Modes	RMSE		
	Maximum	Average	Standard Deviation
1	68.475	53.568	9.446
2	50.799	27.179	16.706
3	32.572	12.253	9.216
4	12.072	8.396	1.433
5	11.590	8.297	1.330
6	13.803	8.803	1.904
7	12.338	9.709	1.745
8	13.035	10.368	1.801
9	12.875	11.181	0.853
10	15.286	12.113	1.762

Figure 5. Effects of the activation function on the ELM performance.

Figure 5 depicts the impact of the number of hidden layer modes on the ELM performance. When the number of modes is 5, the standard deviation, maximum, and average error are the smallest. Therefore, there are five hidden modes in the present study. When the activation function for the ELM is a hardlim function, the maximum, average, and standard deviation of error are more significant than the other two functions. The predicted errors in the other two functions are relatively small when the activation function is the sigmoid function; hence, the activation function of the ELM model is the sigmoid function.

6.2. Calculation Results and Performance Comparison

A multi-algorithm is applied to enhance the ELM model predictability after optimizing the parameters of the activation function and the number of hidden neurons. Figure 6

depicts the predicted results using a single ELM model and ELM optimized by PSO, GWO, WOA, BOA, and SSA.
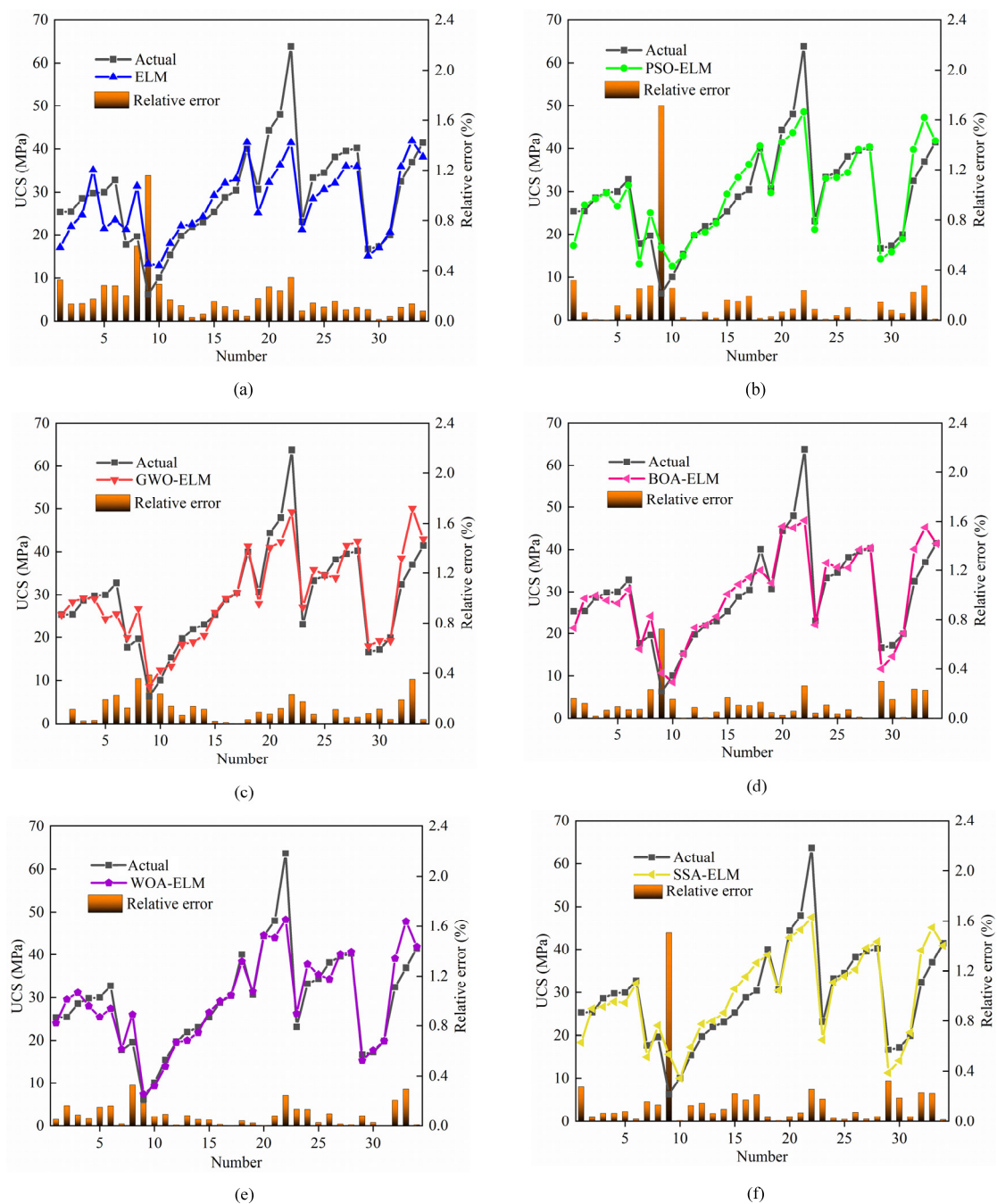

Figure 6. Performance of single ELM models and hybrid ELM models optimized by PSO, GWO, WOA, BOA, and SSA. (**a**) predicted results and relative errors of ELM model; (**b**) predicted results and relative errors of PSO-ELM model; (**c**) predicted results and relative errors of GWO-ELM model; (**d**) predicted results and relative errors of BOA-ELM model; (**e**) predicted results and relative errors of WOA-ELM model; (**f**) predicted results and relative errors of SSA-ELM model.

Figure 6a illustrates that only a few predicted values are close to the actual values; thus, a single ELM model mispredicted UCS. The random generation of the weights and thresholds of the input and hidden layers can limit the performance of ELM. Figure 6b,f show that the actual and predicted curves change together, indicating that the optimized ELM by PSO (PSO-ELM) and SSA (SSA-ELM) can estimate UCS using point load index,

P-wave velocity, and Schmidt hammer rebound. The predicted performance of an ELM optimized by the PSO and SSA model is better than a single ELM model. The minimum and average relative errors of PSO-ELM and SSA-ELM are 0.34% and 0.55%, respectively, which are smaller than the single ELM model. However, the maximum relative errors of PSO-ELM and SSA-ELM are 171.62% and 150.67%, respectively, relatively less than the single ELM model. It indicates that PSO and SSA can relatively enhance ELM predictability, and PSO-ELM and SSA-ELM models are unstable. Compared to the above two algorithms, BOA also improves the predictability of ELM. Figure 6d demonstrates that BOA-ELM predicts better performance than the single ELM, PSO-ELM, and SSA-ELM models. The minimum, maximum, and average relative errors of BOA-ELM are 0.22%, 72.05%, and 11.48%, respectively, smaller than single ELM, PSO-ELM, and SSA-ELM models. However, the maximum relative errors of single ELM, PSO-ELM, SSA-ELM, and BOA-ELM are nearly greater than 50%, indicating that at the lowest value of UCS, its predictive accuracy is almost awful. This can be utilized if data are lacking in ranges with low UCS values or if the ELM parameters require further optimization. As shown in Figure 6c,e, nearly all predicted values are close to actual values, demonstrating that GWO and WOA can further improve the predictability of the ELM model. At low values, the prediction accuracy of GWO-ELM and WOA-ELM models is superior to that of the other three algorithms. The minimum, maximum, and average relative errors of WOA-ELM are 0.22%, 72.05%, and 11.48%, respectively, smaller than the GWO-ELM model and significantly less than the single ELM, PSO-ELM, and SSA-ELM models.

Figure 7 depicts the residual error results using a single ELM model optimized by PSO, GWO, WOA, BOA, and SSA. The residual error histograms of six models exhibit normal distributions. The range of residual errors of the single ELM model is between 0.25 and 22.21 MPa, with a mean of 5.07 MPa. The mean value of residual errors using PSO-ELM is 3.2 MPa, varying widely from 0.10 to 15.28 MPa. The average residual errors of BOA-ELM and SSA-ELM are 2.91 MPa (0.06–16.78 MPa) and 3.34 MPa (0.06–16.16 MPa). The average value of GWO-ELM residual errors is 3.18, ranging from 0.05 to 14.51 MPa. The minimum, maximum, and average residual errors derived from optimized ELM models are less than the single ELM model. The maximum residual errors of PSO-ELM, GWO-ELM, BOA-ELM, WOA-ELM, and SSA-ELM models are less than 20 MPa, and the smallest of them is 14.51 MPa. The greatest residual error of the WOA-ELM model is 15.41 MPa, which is relatively bigger than the smallest maximum residual error. The minimum and mean values of residual errors (using WOA-ELM) are lower than others. It indicates that a multi-algorithm can improve the ELM model's predictability, and its best performance is WOA.

Figure 8 illustrates the R^2 results produced by a single ELM model and an ELM optimized by PSO, GWO, WOA, BOA, and SSA for UCS.

The ELM model produces an R^2 value for UCS of 0.682, as depicted in Figure 8a. The accuracy of optimized ELM models is more than 0.80 and higher than that of a single ELM model. It is understandable to see that a multi-algorithm can enhance the predictability of the ELM model. Figure 8b, f shows that the R^2 derived from PSO-ELM, SSA-ELM, and GWO-ELM models are 0.812, 0.827, and 0.835, respectively. The R^2 results generated by the above three models fall between 0.80 and 0.85. Accordingly, three algorithms enhance the prediction ability of ELM, but the accuracy must be improved. Figure 8d,e reveals that the R^2 of the BOA-ELM and WOA-ELM models is greater than 0.85, indicating their performance is superior to that of the above three algorithms. Meanwhile, the R^2 of the WOA-ELM model is 0.861, which is higher than the R^2 for the ELM and other optimized ELM models. Therefore, the WOA-ELM model, being a combinatorial approach to the modeling work, performed best compared to ELM and optimized models.
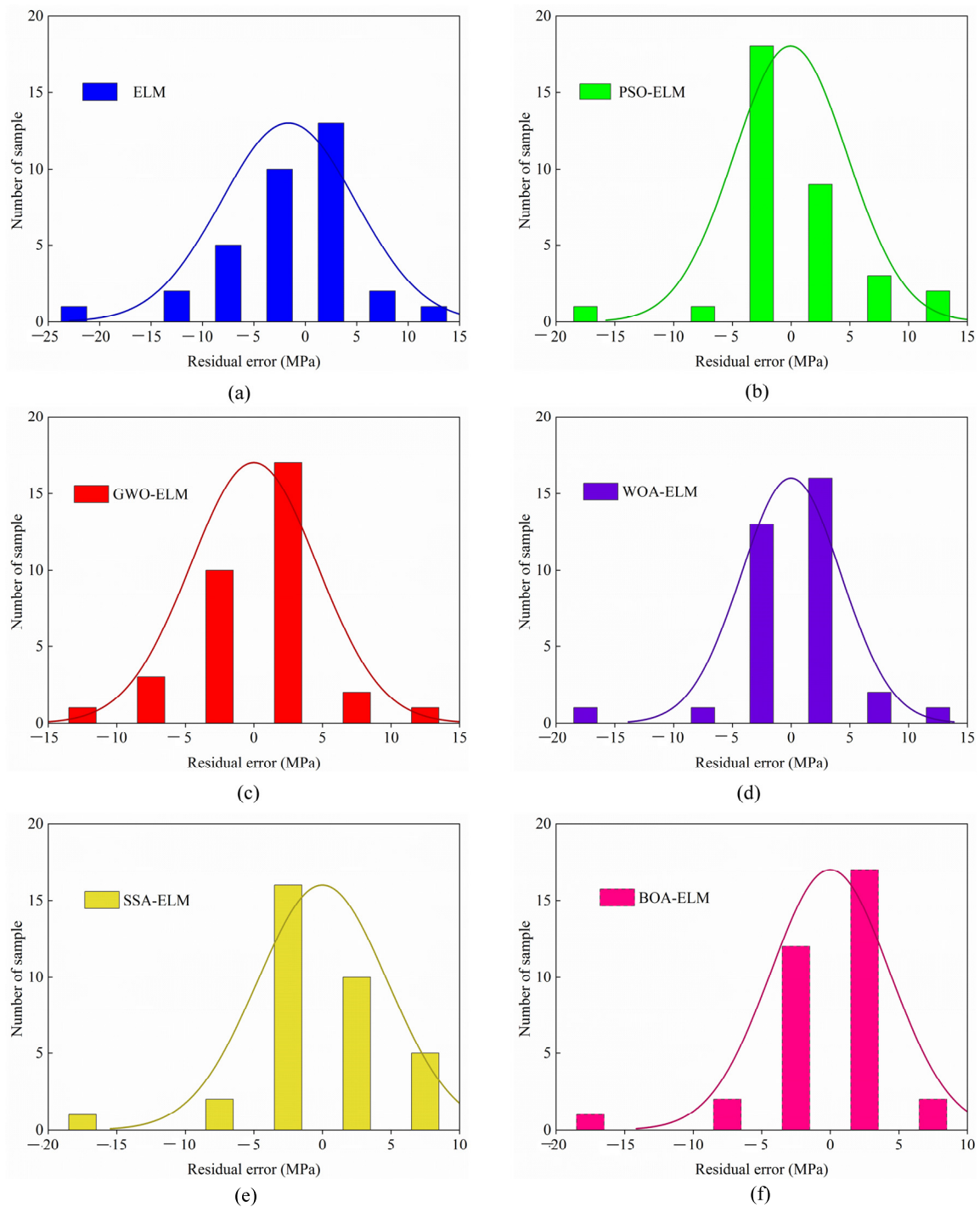
Figure 7. Frequency distributions of residual errors utilizing a single ELM model and hybrid ELM models optimized by PSO, GWO, WOA, BOA, and SSA. (**a**) residual errors based on ELM model; (**b**) residual errors based on PSO-ELM model; (**c**) residual errors based on GWO-ELM model; (**d**) residual errors based on BOA-ELM model; (**e**) residual errors based on WOA-ELM model; (**f**) residual errors based on SSA-ELM model.
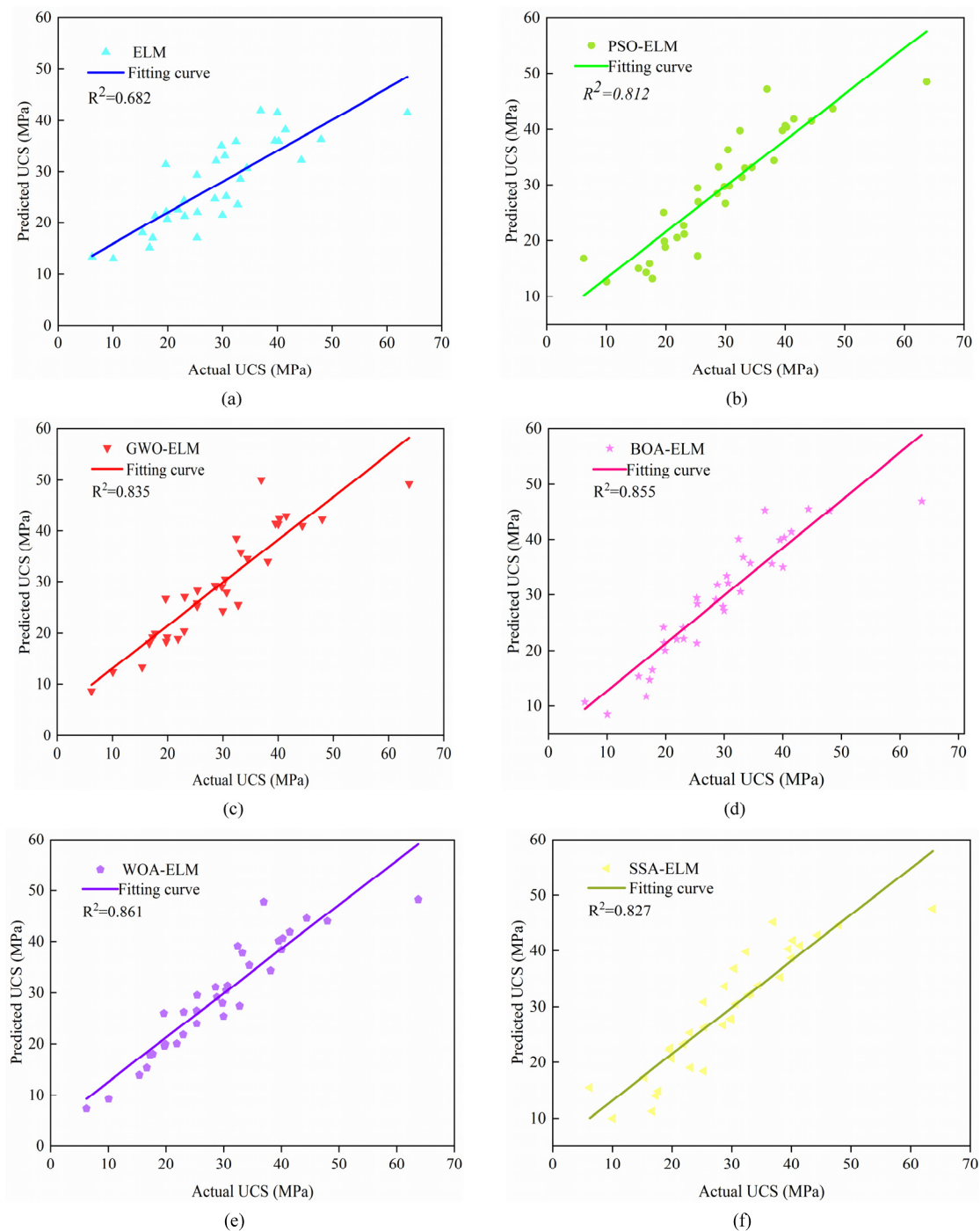
Figure 8. UCS results utilizing a single ELM and hybrid ELM models optimized by PSO, GWO, WOA, BOA, and SSA. (**a**) R^2 of measured and predicted values of UCS using ELM model; (**b**) R^2 of measured and predicted values of UCS using PSO-ELM model; (**c**) R^2 of measured and predicted values of UCS using GWO-ELM model; (**d**) R^2 of measured and predicted values of UCS using BOA-ELM model; (**e**) R^2 of measured and predicted values of UCS using WOA-ELM model; (**f**) R^2 of measured and predicted values of UCS using SSA-ELM model.

To further compare the proposed models, their performance indices, i.e., RMSE, VAF, and MSE, were calculated as presented in Table 3. Theoretically, a predictive model is better when the RMSE and MSE equal 0, and VAF is 100%. Table 3 indicates that the MSE and RMSE of the ELM model are more significant than those of optimized ELM models. The VAF value of the ELM model is 73% less than that of optimized ELM models.

It should be noted that the ELM model must be improved, and the multi-algorithm can increase the predictability of the ELM model. The RMSE and MSE of the WOA-ELM model are significantly lower than PSO-ELM, GWO-ELM, and SSA-ELM models, and the VAF produced by WOA-ELM is also larger than that of the above three models. Comparatively, the RMSE and MSE of WOA-ELM are smaller than those of the BOA-ELM model. As previously stated, the relative and residual errors of the WOA-ELM model are smaller than those of others, and the R^2 of the present model is closer to 1 than other models. In this study, the WOA-ELM model can predict UCS with a higher degree of accuracy than the ELM and combined ELM models.

Table 3. Performance indices of the proposed predictive models.

Model	ELM	PSO-ELM	GWO-ELM	BOA-ELM	WOA-ELM	SSA-ELM
R^2	0.682	0.812	0.835	0.855	0.861	0.827
MSE	44.37	22.65	20.88	18.36	17.61	21.92
RMSE	6.66	4.76	4.57	4.28	4.20	4.68
VAF (%)	73	88	79	92	91	90

7. Conclusions

Predicting UCS is an interesting and challenging exercise. This study first collects 734 samples to conduct a new dataset that includes magmatic, sedimentary, and metamorphic rocks and rock-like materials from different countries. The ELM was proposed to map the relationship between UCS and point load index, P-wave velocity, and Schmidt hammer rebound number to estimate UCS. In order to further predict UCS, five algorithms (PSO, GWO, WOA, BOA, and SSA) were applied to improve the predictability of ELM. Based on the aforementioned statements, the following conclusions are drawn:

- The optimized ELM model consists of five hidden neurons and a sigmoid activation function.
- Compared to the models proposed above, it can be stated that the predicted performance of the six models for predicting UCS from high to low is as follows: WOA-ELM, BOA-ELM, GWO-ELM, SSA-ELM, PSO-ELM, and ELM. The predicted indices (R^2: 0.861; MSE: 17.61; RMSE: 4.20) produced by WOA-ELM illustrate that it is the more precise model.
- The minimum, maximum, and average relative errors produced by ELM optimized using the whale optimization algorithm (WOA-ELM) are 0.22%, 72.05%, and 11.48% smaller than the other models.
- The minimum and mean residual error produced by WOA-ELM are 0.02 and 2.64 MPa, respectively, smaller than other models.
- The results showed that the WOA-ELM model is the best among other techniques investigated in this study. Its performance indices reveal the high accuracy and reliability of the new model for predicting UCS.

In all, the hybrid models proposed in this study are suitable for different rocks. Thus, the proposed WOA-ELM model in this study has broad application potential in predicting the UCS of various rocks.

The main limitation of this paper is that only one dataset was utilized to evaluate the results of developed models. Meanwhile, this study did not consider that the proposed algorithms have some limitations, such as local minima trapping issues and the inability to exploit local space. To avoid this, additional research will be conducted in the future:

1. The developed model in this study will be applied to other datasets to demonstrate its generalization ability and robustness.
2. We will present strategies to avoid the problem of local minima trapping issues and the inability of metaheuristic algorithms to exploit local space and illustrate their impact on the current model.

References

1. Heidari, M.; Mohseni, H.; Jalali, S.H. Prediction of uniaxial compressive strength of some sedimentary rocks by fuzzy and regression models. *Geotech. Geol. Eng.* **2017**, *36*, 401–412. [CrossRef]
2. Armaghani, D.J.; Mohamad, E.T.; Momeni, E.; Narayanasamy, M.S.; Amin, M. An adaptive neuro-fuzzy inference system for predicting unconfined compressive strength and Young's modulus: A study on Main Range granite. *Bull. Eng. Geol. Environ.* **2015**, *74*, 1301–1319. [CrossRef]
3. Mishra, D.A.; Basu, A. Use of the block punch test to predict the compressive and tensile strengths of rocks. *Int. J. Rock Mech. Min. Sci.* **2012**, *51*, 119–127. [CrossRef]
4. Şahin, M.; Ulusay, R.; Karakul, H. Point load strength index of half-cut core specimens and correlation with uniaxial compressive strength. *Rock Mech. Rock Eng.* **2020**, *53*, 3745–3760. [CrossRef]
5. Basu, A.; Kamran, M. Point load test on schistose rocks and its applicability in predicting uniaxial compressive strength. *Int. J. Rock Mech. Min. Sci.* **2010**, *47*, 823–828. [CrossRef]
6. Singh, T.N.; Kainthola, A.; Venkatesh, A. Correlation between point load index and uniaxial compressive strength for different rock types. *Rock Mech. Rock Eng.* **2012**, *45*, 259–264. [CrossRef]
7. Kahraman, S.; Fener, M.; Gunaydin, O. Estimating the uniaxial compressive strength of pyroclastic rocks from the slake durability index. *Bull. Eng. Geol. Environ.* **2017**, *76*, 1107–1115. [CrossRef]
8. Zhang, H.; Wu, S.; Zhang, Z. Prediction of uniaxial compressive strength of rock via genetic algorithm—Selective ensemble learning. *Nat. Resour. Res.* **2022**, *31*, 1721–1737. [CrossRef]
9. Yagiz, S. Correlation between slake durability and rock properties for some carbonate rocks. *Bull. Eng. Geol. Environ.* **2011**, *70*, 377–383. [CrossRef]
10. Khandelwal, M. Correlating P-wave velocity with the physico-mechanical properties of different rocks. *Pure Appl. Geophys.* **2013**, *170*, 507–514. [CrossRef]
11. Iyare, U.C.; Blake, O.O.; Ramsook, R. Estimating the uniaxial compressive strength of argillites using brazilian tensile strength, ultrasonic wave velocities, and elastic properties. *Rock Mech. Rock Eng.* **2021**, *54*, 2067–2078. [CrossRef]
12. Wang, S.; Li, X.; Yao, J.; Gong, F.; Du, S. Experimental investigation of rock breakage by a conical pick and its application to non-explosive mechanized mining in deep hard rock. *Int. J. Rock Mech. Min. Sci.* **2019**, *122*, 104063. [CrossRef]
13. Wang, S.; Sun, L.; Li, X.; Zhou, J.; Du, K.; Wang, S.; Khandelwal, M. Experimental investigation and theoretical analysis of indentations on cuboid hard rock using a conical pick under uniaxial lateral stress. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2022**, *8*, 34. [CrossRef]
14. Karakus, M.; Tutmez, B. Fuzzy and multiple regression modelling for evaluation of intact rock strength based on point load, schmidt hammer and sonic velocity. *Rock Mech. Rock Eng.* **2006**, *39*, 45–57. [CrossRef]
15. Mishra, D.A.; Basu, A. Estimation of uniaxial compressive strength of rock materials by index tests using regression analysis and fuzzy inference system. *Eng. Geol.* **2013**, *160*, 54–68. [CrossRef]
16. Sarkar, K.; Singh, A. Estimation of strength parameters of rock using artificial neural networks. *Bull. Eng. Geol. Environ.* **2010**, *69*, 599–606. [CrossRef]
17. Yagiz, S.; Sezer, E.A.; Gokceoglu, C. Artificial neural networks and nonlinear regression techniques to assess the influence of slake durability cycles on the prediction of uniaxial compressive strength and modulus of elasticity for carbonate rocks. *Int. J. Numer. Anal. Methods Geomech.* **2012**, *36*, 1636–1650. [CrossRef]
18. Yesiloglu-Gultekin, N.U.; Gokceoglu, C.; Sezer, E.A. Prediction of uniaxial compressive strength of granitic rocks by various nonlinear tools and comparison of their performances. *Int. J. Rock Mech. Min. Sci.* **2013**, *62*, 113–122. [CrossRef]
19. Dindarloo, S.R.; Siami-Irdemoosa, E. Estimating the unconfined compressive strength of carbonate rocks using gene expression programming. *arXiv* **2016**, arXiv:1602.03854.

20. Gül, E.; Ozdemir, E.; Sarc, D.E. Modeling Uniaxial Compressive Strength of Some Rocks from Turkey Using Soft Computing Techniques. *Measurement* **2020**, *171*, 108781. [CrossRef]

21. Wen, L.; Tan, Z. Research on Rock Strength Prediction Based on Least Squares Support Vector Machine. *Geotech. Geol. Eng.* **2017**, *35*, 385–393.

22. Mahmoodzadeh, A.; Mohammadi, M.; Ibrahim, H.; Abdulhamid, S.N.; Ali, H. Artificial intelligence forecasting models of uniaxial compressive strength. *Transp. Geotech.* **2021**, *27*, 100499. [CrossRef]

23. Gupta, D.; Natarajan, N. Prediction of uniaxial compressive strength of rock samples using density weighted least squares twin support vector regression. *Neural Comput. Appl.* **2021**, *33*, 15843–15850. [CrossRef]

24. Momeni, E.; Armaghani, D.J.; Hajihassani, M.; Amin, M.M. Prediction of uniaxial compressive strength of rock samples using hybrid particle swarm optimization-based artificial neural networks. *Measurement* **2015**, *60*, 50–63. [CrossRef]

25. Fang, Q.; Bejarbaneh, B.Y.; Vatandoust, M.; Armaghani, D.J.; Mohamad, E.T. Strength evaluation of granite block samples with different predictive models. *Eng. Comput.* **2019**, *37*, 891–908. [CrossRef]

26. Izonin, I.; Tkachenko, R.; Shakhovska, N.; Lotoshynska, N. The additive input-doubling method based on the svr with nonlinear kernels: Small data approach. *Symmetry* **2021**, *13*, 612. [CrossRef]

27. Izonin, I.; Tkachenko, R.; Dronyuk, I.; Tkachenko, P.; Rashkevych, M. Predictive modeling based on small data in clinical medicine: Rbf-based additive input-doubling method. *Math. Biosci. Eng. MBE* **2021**, *18*, 2599–2613. [CrossRef]

28. Yin, X.; Liu, Q.; Pan, Y.; Huang, X.; Wang, X. Strength of stacking technique of ensemble learning in rockburst prediction with imbalanced data: Comparison of eight single and ensemble models. *Nat. Resour. Res.* **2021**, *30*, 1795–1815. [CrossRef]

29. Huang, G.B.; Chen, L.; Siew, C.K. Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Trans. Neural Netw.* **2006**, *17*, 879–892. [CrossRef]

30. Kang, F.; Liu, J.; Li, J.; Li, S. Concrete dam deformation prediction model for health monitoring based on extreme learning machine. *Struct. Control. Health Monit.* **2017**, *24*, e1997. [CrossRef]

31. Li, E.; Yang, F.; Ren, M.; Zhang, X.; Zhou, J.; Khandelwal, M. Prediction of blasting mean fragment size using support vector regression combined with five optimization algorithms. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 18. [CrossRef]

32. Yin, X.; Liu, Q.; Huang, X.; Pan, Y. Perception model of surrounding rock geological conditions based on tbm operational big data and combined unsupervised-supervised learning. *Tunn. Undergr. Space Technol.* **2022**, *120*, 104285. [CrossRef]

33. Xin, Y.; Qla, B.; Xing, H.C.; Ypa, B. Real-time prediction of rockburst intensity using an integrated cnn-adam-bo algorithm based on microseismic data and its engineering application. *Tunn. Undergr. Space Technol.* **2021**, *117*, 104133.

34. Tuğrul, A.; Zarif, I.H. Correlation of mineralogical and textural characteristics with engineering properties of selected granitic rocks from Turkey. *Eng. Geol.* **1999**, *51*, 303–317. [CrossRef]

35. Kahraman, S. Evaluation of simple methods for assessing the uniaxial compressive strength of rock. *Int. J. Rock Mech. Min. Sci.* **2001**, *38*, 981–994. [CrossRef]

36. Dinner, I.; Acar, A.; Ural, S. Estimation of strength and deformation properties of Quaternary caliche deposits. *Bull. Eng. Geol. Environ.* **2008**, *67*, 353–366.

37. Kilic, A.; Teymen, A. Determination of mechanical properties of rocks using simple methods. *Bull. Eng. Geol. Environ.* **2008**, *67*, 237. [CrossRef]

38. Çobanoğlu, İ.; Çelik, S.B. Estimation of uniaxial compressive strength from point load strength, Schmidt hardness and P-wave velocity. *Bull. Eng. Geol. Environ.* **2008**, *67*, 491–498. [CrossRef]

39. Aliabadi, S. Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks. *Min. Sci. Technol.* **2010**, *20*, 41–46.

40. Tandon, R.S.; Gupta, V. Estimation of strength characteristics of different Himalayan rocks from Schmidt hammer rebound, point load index, and compressional wave velocity. *Bull. Eng. Geol. Environ.* **2015**, *74*, 521–533. [CrossRef]

41. Jahed Armaghani, D.; Tonnizam Mohamad, E.; Hajihassani, M.; Yagiz, S.; Motaghedi, H. Application of several non-linear prediction tools for estimating uniaxial compressive strength of granitic rocks and comparison of their performances. *Eng. Comput.* **2016**, *31*, 189–206. [CrossRef]

42. Armaghani, D.J.; Mohamad, E.T.; Momeni, E.; Monjezi, M.; Narayanasamy, M.S. Prediction of the strength and elasticity modulus of granite through an expert artificial neural network. *Arab. J. Geosci.* **2016**, *9*, 48. [CrossRef]

43. Ng, I.T.; Yuen, K.V.; Lau, C.H. Predictive model for uniaxial compressive strength for Grade III granitic rocks from Macao. *Eng. Geol.* **2015**, *199*, 28–37. [CrossRef]

44. Kennedy, J. Particle swarm optimization. In Proceedings of the ICNN'95—International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948.

45. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]

46. Mirjalili, S.; Lewis, A.D. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]

47. Arora, S.; Singh, S. Butterfly optimization algorithm: A novel approach for global optimization. *Soft Comput.* **2018**, *23*, 715–734. [CrossRef]

48. Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control. Eng. Open Access J.* **2020**, *8*, 22–34. [CrossRef]