

Article

Developing Hybrid DMO-XGBoost and DMO-RF Models for Estimating the Elastic Modulus of Rock

Weixing Lin ^{1,2}, Leilei Liu ^{1,*}, Guoyan Zhao ¹  and Zheng Jian ¹

¹ School of Resources and Safety Engineering, Central South University, Changsha 410083, China; gyzhao@csu.edu.cn (G.Z.)

² Changsha Institute of Mining Research Co., Ltd., Changsha 410012, China

* Correspondence: leilei_liu@csu.edu.cn

Abstract: Accurate estimation of the elastic modulus (E) of rock is critical for the design of geotechnical projects such as mining, slopes, and tunnels. However, the determination of rock mechanical parameters usually involves high budget and time requirements. To address this problem, numerous researchers have developed machine learning models to estimate the E of rock. In this study, two novel hybrid ensemble learning models were developed to estimate the E of rock by optimizing the extreme gradient boosting (XGBoost) and random forest (RF) algorithms through the dwarf mongoose optimization (DMO) approach. Firstly, 90 rock samples with porosity, dry density, P -wave velocity, slake durability, and water absorption as input indicators were collected. Subsequently, the hyperparameters of XGBoost and RF were tuned by DMO. Based on the optimal hyperparameters configuration, two novel hybrid ensemble learning models were constructed using the training set (80% of the data). Finally, the performance of the developed models was evaluated by the coefficient of determination (R^2 score), root mean squared error (RMSE), mean absolute error (MAE), and variance accounted for (VAF) on the test set (20% of the data). The results show that the DMO-RF model achieved the best comprehensive performance with an R^2 score of 0.967, RMSE of 0.541, MAE of 0.447, and VAF of 0.969 on the test set. The dry density and slake durability were more influential indicators than others. Moreover, the convergence curves suggested that the DMO-RF model can reduce the generalization error and avoid overfitting. The developed models can be regarded as viable and useful tools in estimating the E of rock.

Keywords: elastic modulus of rock; machine learning; extreme gradient boosting; random forest; dwarf mongoose optimization

MSC: 86-10



Citation: Lin, W.; Liu, L.; Zhao, G.; Jian, Z. Developing Hybrid DMO-XGBoost and DMO-RF Models for Estimating the Elastic Modulus of Rock. *Mathematics* **2023**, *11*, 3886. <https://doi.org/10.3390/math11183886>

Academic Editors: Zhanbo Cheng, Shuai Zhao, Defeng Wang and Lijun Han

Received: 17 August 2023

Revised: 1 September 2023

Accepted: 11 September 2023

Published: 12 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Elastic modulus (E) is one of the key parameters characterizing the properties of rock, and the accurate estimation of the E of rock is essential for the safe construction of geotechnical engineering [1–5]. The E of rock can be determined by situ and laboratory tests, and the test process should comply with a series of operating specifications. However, it is sometimes difficult to extract high-quality core specimens from fragile, weak, and stratified rock masses. Additionally, the actual testing process is often limited by time and cost constraints [6–8]. Therefore, it is essential to develop economical, non-destructive, and indirect methods to estimate the E of rock.

In the early stage of the development of rock mechanics, several scholars attempted to estimate the E of rock by some easily accessible rock indexes, such as the point-load index, slake durability index, Schmidt hammer rebound number, and P -wave velocity [9–12]. As a result, several empirical formulas between E and the above indexes were established. For example, Sachpazis [13] calculated a correlation coefficient of 0.78 between E and the

Schmidt hammer rebound number in carbonate rocks. Lashkaripour [14] analyzed the correlation between the porosity and E of mudrocks. Yasar and Erdogan [15] determined the statistical relations between the hardness and E of rocks. Moradian and Behnia [16] derived fit equations for E and the P -wave velocity of sedimentary rocks by using an ultrasonic test. Armaghani et al. [17] proposed several simple and multiple regression equations to calculate the E of granite rocks. Shan and Di [18] developed formulas to calculate the E of multiple-jointed basalt rocks by laboratory tests. Karakus et al. [19] adopted a multiple regression model to analyze the elastic properties of intact rock. However, the estimation results of these empirical formulas are only applicable to specific rock data and do not reach the generalization purpose.

On the other hand, in recent years, machine learning (ML) approaches have garnered considerable interest in geotechnical engineering because of their powerful nonlinear data processing capabilities [20–22]. With the accumulation of data, some researchers have attempted to develop ML models to estimate the E of rock. For instance, Armaghani et al. [17] proposed an adaptive neuro-fuzzy inference system (ANFIS) for predicting unconfined compressive strength (UCS) and E of granite rocks. Cao et al. [23] hybridized an extreme gradient boosting machine (XGBoost) with the firefly algorithm to estimate the E and UCS of granite rock. Meng and Wu [24] combined the experimental, numerical, and random forest (RF) methods to predict the UCS and the E of frozen fractured rock. Abdi et al. [25] investigated the feasibility of tree-based techniques, including RF, Adaboost, XGBoost, and Catboost models, in predicting the E of weak rock. Acar and Kaya [26] adopted the least square support vector machine (LS-SVM) model to predict the E of weak rock. Several scholars [27–29] adopted different artificial neural network models (ANN) to predict the E of different types of rocks, respectively. More related works on the estimation of E using ML methods and details are tabulated in Table 1.

Table 1. Related works on E estimation using ML approaches.

Year	References	Models	Inputs	Number of Samples
2010	Majdi and Beiki [30]	GA-ANN	UCS, GSI, RQD, ρ , n , NJ	120
2010	Dehghan et al. [29]	ANN	V_p , $I_{S(50)}$, R_n , n	30
2011	Khandelwal and Singh [31]	ANN	UCS, BTS	120
2012	Ocak and Seker [32]	ANN	UCS, γ	195
2015	Armaghani et al. [17]	ANFIS	ρ , V_p , content of Qtz, Kpr, Plg, and Bi	45
2018	Bejarbaneh et al. [33]	ANN	$I_{S(50)}$, V_p , R_n	96
2018	Saedi et al. [34]	ANFIS	BPI, BTS, V_p , $I_{S(50)}$	120
2018	Rezaei [35]	MFIS	H , ρ , n , DI	50
2019	Yang et al. [36]	Bayesian	$I_{S(50)}$, R_n , V_p , n , UCS	71
2020	Acar and Kaya [26]	LS-SVM	V_p , γ , $I_{S(50)}$, BTS	575
2022	Cao et al. [23]	XGBoost-firefly	ρ , V_p , content of Qtz, Kpr, Plg, Bi	45
2022	Pappalardo and Mineo [27]	ANN	n , γ , V_p , E_{dyn} , UCS	/
2023	Meng and Wu [24]	RF	PF, SF, IAF, TF, NF	/
2023	Abdi et al. [25]	RF	n , ρ , V_p , Id_2 , Abs	90

Note: GSI is geological strength index; RQD is rock mass quality designation; GA is genetic algorithm; BTS is Brazilian tensile strength; γ is unit weight; ρ is density; n is porosity; NJ is number of joints per meter; V_p is P -wave velocity; $I_{S(50)}$ is point load index; R_n is the Schmidt hammer rebound number; Qtz is quartz; Kpr is alkali feldspar; Plg is plagioclase; Bi is biotite; BPI is block punch index; H is depth of coring; DI is durability index; E_{dyn} is dynamic elastic modulus; Id_2 is slake durability; Abs is water absorption; PF is persistence factor of ice-filled fractures; SF is spacing between fractures; IAF is inclination angle of fractures; TF is thickness of fractures; NF is number of fractures.

In comparison to other methods, ML approaches can yield dependable results by establishing the nonlinear relationship between input and output variables [37,38]. It is a promising method for estimating the E of rock. Recently, the XGBoost and RF algorithms have shown great potential to improve prediction accuracy and have been successfully applied in many fields, such as electricity consumption forecasting [39,40], infectious disease prediction [41,42], mining maximum subsidence prediction [43,44], and heavy metal contamination prediction [45,46]. XGBoost and RF are two efficient ensemble methods that combine multiple homogeneous weak learners in certain ways to reduce overfitting.

However, the ML algorithms require a proper hyperparameter setting to improve the accuracy. As a novel heuristic optimization algorithm, the dwarf mongoose optimization (DMO) algorithm has proven to have a strong global search capability and high search efficiency in solving optimization problems [47]. Therefore, it may be an efficient approach to optimize the hyperparameters of XGBoost and RF models. The available studies show that the DMO algorithm has not been integrated with the XGBoost and RF models to estimate the E of rock.

This study aims to investigate the feasibility of XGBoost and RF algorithms optimized by the DMO approach to estimate the E of rock. First, a dataset comprising 90 specimens with five indicators was compiled from available rock mechanics data. Next, the DMO-XGBoost and DMO-RF models were developed. The proposed models can be used to estimate the E of rock easily by avoiding the complexity of core specimen preparation in the laboratory.

2. Methodology

2.1. XGBoost Algorithm

XGBoost is a boosting ensemble algorithm with a decision tree (DT) as the base learner that improves the classical gradient boosting decision tree (GBDT) algorithm [48]. It is designed using an efficient, flexible, and portable distributed gradient boosting framework. Compared with GBDT, the calculation speed is faster while retaining excellent performance. The flowchart of this algorithm is indicated in Figure 1. The core principle of XGBoost is to build DTs one after another and train the subsequent DT with the residuals of the previous ones. The values computed by all the constructed DTs are integrated to achieve a better result. It has been regarded as an advanced evaluator with ultra-high performance in both classification and regression [21]. The XGBoost algorithm is explained as follows:

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{1}$$

where $\hat{y}_i^{(t)}$ denotes the final model, $\hat{y}_i^{(t-1)}$ is the previous model, x_i represents the features corresponding to the sample i , $f_t(x_i)$ is the newly generated DT model, and t is the total number of DT models.

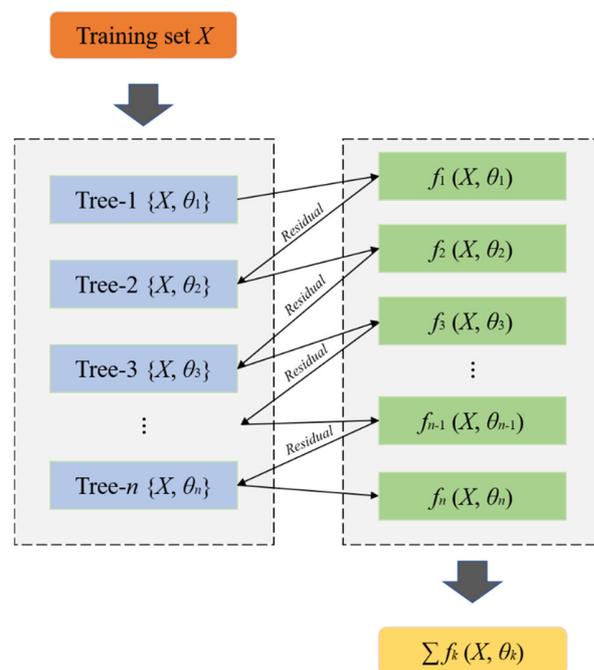


Figure 1. Principle diagram of XGBoost model.

The objective function of the ensemble algorithm is an alternative. In order to avoid overfitting and improve iteration efficiency, XGBoost introduces model complexity to measure the computational efficiency [49]. Accordingly, the objective function of XGBoost can be given by:

$$Obj^{(t)} = \sum_{i=1}^t l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^t \Omega(f_i) \tag{2}$$

where y_i is the actual value, $l(y_i, \hat{y}_i^{(t)})$ is a convex loss function describing how well the model fits with training data, and $\Omega(f_i)$ is the penalty term for regularization to avoid overfitting.

Gradient statistics are applied to the loss function, which effectively eliminates constant parameters. The simplified objective function is obtained as follows:

$$Obj^{(t)} = \sum_{i=1}^t \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{3}$$

where $g_i = \partial \hat{y}_i^{(t-1)} l(y_i, \hat{y}_i^{(t)})$ and $h_i = \partial^2 \hat{y}_i^{(t-1)} l(y_i, \hat{y}_i^{(t)})$ are the first and second order gradient statistics of the loss function, respectively. The penalty term $\Omega(f_t)$ is evaluated by:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \alpha |w| + \frac{1}{2} \lambda \|w\|^2 \tag{4}$$

where α and λ are regularization parameters, whose default values are 1 and 0, respectively, T is the number of leaf nodes controlled by the parameter γ , and w denotes the corresponding weight of the leaf nodes.

2.2. Random Forest Algorithm

RF is a bagging ensemble algorithm with DT as the base learner [50]. Its principle diagram is shown in Figure 2. Different DTs are constructed to set up a forest by extracting sample and feature subsets from the original data. It is a stochastic process that ensures the independence of each DT, which improves the generalization capacity and avoids overfitting [51]. Moreover, all constructed DTs tend to grow freely without pruning, which makes the construction process of each DT fast and the computational efficiency of RF improved. By integrating the results generated by all constructed DTs, the voting (for classification problems) or averaging (for regression problems) approach is used to determine the final results [21]. According to the type of output variable in this paper, the RF algorithm is explained as follows:

$$\bar{y}(x) = \frac{1}{T} \sum_{i=1}^T f_i(x) \tag{5}$$

where $\bar{y}(x)$ represents the predicted value of the RF model, $f_i(x)$ is the predicted value of each DT, and T represents the number of constructed DTs.

2.3. Dwarf Mongoose Optimization

The DMO algorithm is a swarm intelligence optimization algorithm that simulates the seminomadic life of dwarf mongoose in nature. The dwarf mongooses are known for foraging and scouting as a unit, and the DMO algorithm finds the optimal solution by simulating their social behavior. The DMO algorithm divides the swarm into three groups: the alpha group, the scout group, and the babysitter group [47,52,53].

Firstly, the alpha group leads the scout group to forage and scout for a new sleeping mound, leaving the babysitters group to take care of the young dwarf mongooses in the old sleeping mound. Secondly, when a food source is not found, the alpha and scout groups will go back to the old sleeping mound to exchange members with the babysitter group. Thirdly, the alpha group continues to go out foraging and scouting for new sleeping

mounds. The family returns intermittently to exchange babysitters and repeats the cycle until a food source is found.

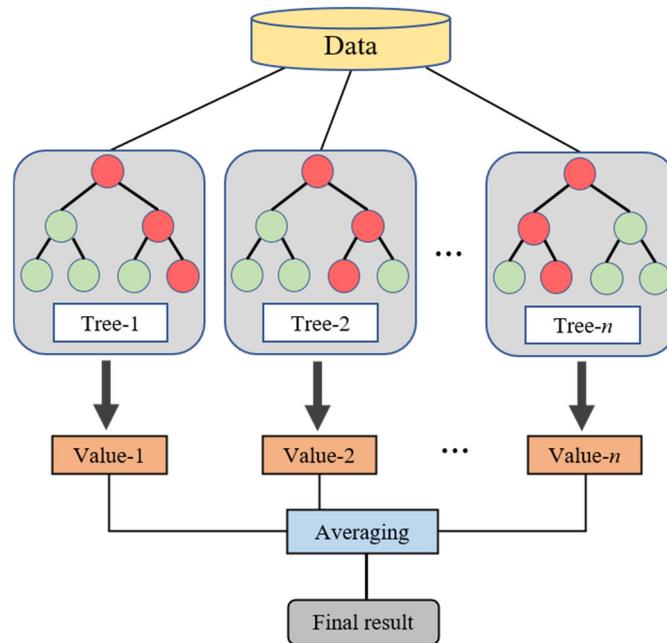


Figure 2. Principle diagram of RF model.

The process of DMO for hyperparameter tuning can be broken down into the following steps [47]:

(1) Initialization: The DMO algorithm starts by initializing the population of solutions as follows:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,d-1} & x_{1,d} \\ x_{2,1} & x_{2,2} & \dots & x_{2,d-1} & x_{2,d} \\ \vdots & \vdots & x_{i,j} & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \dots & x_{n,d-1} & x_{n,d} \end{bmatrix} \tag{6}$$

where X is the set of current candidate populations, n denotes the population size, and d is the dimension of the problem. These solutions represent different sets of hyperparameters for the XGBoost and RF models.

(2) Objective function: The DMO algorithm requires an objective function that quantifies the performance of each solution. In this study, the average root mean squared error (RMSE) values of five-fold cross-validation (CV) were employed as the fitness function. The foraging behavior of the dwarf mongoose is mimicked by dividing the population into alpha, scout, and babysitter groups. The probability of each individual becoming alpha is computed by:

$$\alpha = \frac{fit_i}{\sum_{i=1}^n fit_i} \tag{7}$$

where fit_i denotes the fitness value of the i -th individual.

(3) Behavior of dwarf mongooses: In the optimization of the DMO process, these dwarf mongooses work together to find food and a new sleeping mound. The movement of dwarf mongooses is determined by the leader of its subgroup. Their movement is based on the position of the alpha and the current position of the solution, which is calculated as follows:

$$M = \sum_{i=1}^n \frac{X_i \times sm_i}{X_i} \tag{8}$$

where X_i is the position of the alpha group, and sm_i denotes the current position of the solution.

(4) Cooperative search: DMO encourages a cooperative search among the solutions in the population. This means that solutions collaborate and share information to improve the overall performance of the population. In the context of hyperparameter optimization, it implies that solutions with promising hyperparameters influence or guide other solutions towards better hyperparameter configurations.

(5) Communication and information sharing: In DMO, solutions share information about their performance and hyperparameters with others. This communication helps the population converge toward better solutions.

(6) Update hyperparameters: Based on the cooperative search and shared information, the hyperparameters of each solution are updated.

(7) Fitness evaluation: After updating the hyperparameters, the fitness of each solution is re-evaluated using the objective function. Solutions that perform better are favored and contribute more to the cooperative search process.

(8) Termination: The algorithm terminates when a stopping criterion is met.

(9) Final solution selection: Once the optimization process concludes, the solution with the best-performing hyperparameters, according to the objective function, is selected as the final model configuration.

2.4. Model Evaluation Metrics

In this study, four metrics were employed to assess the performance of the proposed models, including the coefficient of determination (R^2 score), the RMSE, the mean absolute error (MAE), and the variance accounted for (VAF). RMSE and VAF are maximization performance metrics, while MAE and VAF are minimization performance metrics [54].

R^2 score represents the proportion of the squared correlation between the predicted and actual values of the target variable, which can be calculated by:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2} \tag{9}$$

where y_i indicates the actual value, \hat{y}_i indicates the predicted value of the model, and \bar{y}_i represents the average of the actual values.

RMSE represents the standard deviation of the fitted error between the predicted value and the actual value, which can be calculated by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \tag{10}$$

where n indicates the total number of samples.

VAF characterizes the prediction performance by comparing the standard deviation of the fitted error with the standard deviation of the actual value, which is defined as:

$$VAF = \left[1 - \frac{\text{var}(y_i - \hat{y}_i)}{\text{var}(y_i)} \right] \times 100 \tag{11}$$

where var indicates the variance.

MAE represents the average error between predicted value and actual value, which can be calculated by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \tag{12}$$

2.5. Proposed Approach

In this study, two novel hybrid ensemble learning models were proposed to estimate the E of rock by optimizing the XGBoost and RF models through the DMO algorithm. The structure of the proposed approach is indicated in Figure 3. Firstly, a database including 90 rock samples with five indicators was established. Secondly, 80% of the samples were

utilized for training, while the remaining 20% were reserved for testing [55]. Thirdly, to validate the superiority of the proposed hybrid ensemble learning models, a comparison against default XGBoost and RF models was performed. Finally, the R^2 score, RMSE, MAE, and VAF were adopted to evaluate the performance of models on the test set.

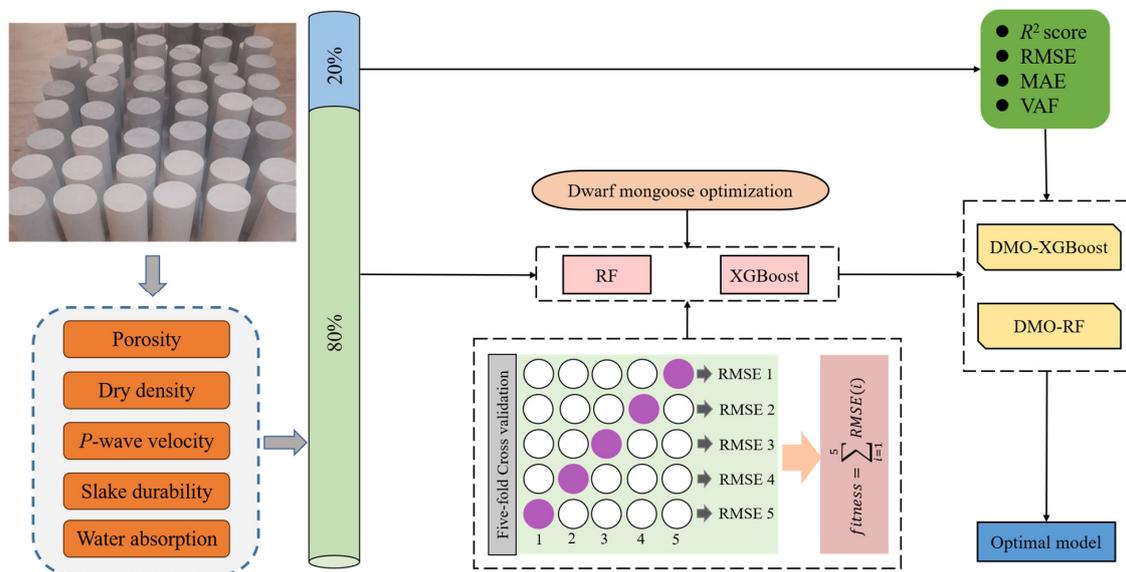


Figure 3. Flowchart of the proposed approach.

The XGBoost and RF algorithms were implemented on the Python library “sci-kit-learn” [56], and the DMO algorithm was implemented on the Python library “MELPY” [57,58]. All experiments were processed using a Windows 10, 64-bit computer with 8 Gb of RAM running an Intel® Core™ i7-9700F CPU @ 3.00 GHz × 2.

3. Data and Variables

According to the work of Abdi et al. [25], four types of rocks, including marl, siltstone, claystone, and limestone, were collected and cored in the laboratory to obtain standard core samples. A total of 90 rock samples were obtained for conducting physical tests and developing E -estimation models. Among the 90 samples, each set of data contains five indicators, including porosity (A_1), dry density (A_2), P -wave velocity (A_3), slake durability (A_4), and water absorption (A_5). Among them, porosity (A_1) is an important factor in determining the strength and deformation behavior of rock, and water absorption (A_5) is related to porosity. Dry density (A_2) and P -wave velocity (A_3) are two common petrophysical properties related to the E of rock. Yagiz et al. [28] found that slake durability (A_4) cycles have a significant effect on the prediction of UCS and modulus of elasticity for carbonate rocks. It is important that these five indicators can be conveniently measured in the laboratory, and therefore, they were selected as input indicators. The detailed descriptions of these indicators are displayed in Table 2, and the statistics of the dataset are illustrated in Table 3.

Table 2. Descriptions of input indicators.

Indicator	Description
Porosity (%)	The ratio of void space to the total volume of rock.
Dry density (g/cm^3)	The mass per unit volume of a rock sample without any water content.
P -wave velocity (m/s)	The speed at which compressional waves travel through a rock sample.
Slake durability (%)	A measure of the resistance of a rock sample to disintegration when exposed to wetting and drying cycles.
Water absorption (%)	The amount of water that a rock sample can absorb when it is immersed in water for a specified period of time.

Table 3. Statistical values of the dataset.

Indicators	Minimum	Median	Maximum	Mean	Standard
A_1 (%)	5.44	21.32	56.55	22.69	9.59
A_2 (g/cm ³)	1.61	2.06	2.98	2.17	0.24
A_3 (m/s)	1011.53	2075.05	3250.45	1989.66	541.78
A_4 (%)	22.75	80.52	96.54	74.87	16.40
A_5 (%)	2.51	10.59	26.54	11.62	5.24
E (GPa)	1.12	3.59	13.23	4.50	2.92

The violin plots of the five indicators and E are shown in Figure 4. These plots are a combination of box plots and density plots, offering insights into the overall distribution of the dataset. In each violin plot, the white dot at the center represents the median of the samples, the upper and lower extents of the thick black line denote the third and first quartiles of the samples, the top and bottom of the thin black line indicate the upper and lower adjacent values, and the black dots are outliers. From Figure 4, it can be seen that the distribution of all the data is relatively balanced, but there are some outliers, which indicates the complexity of estimating the E of rock.

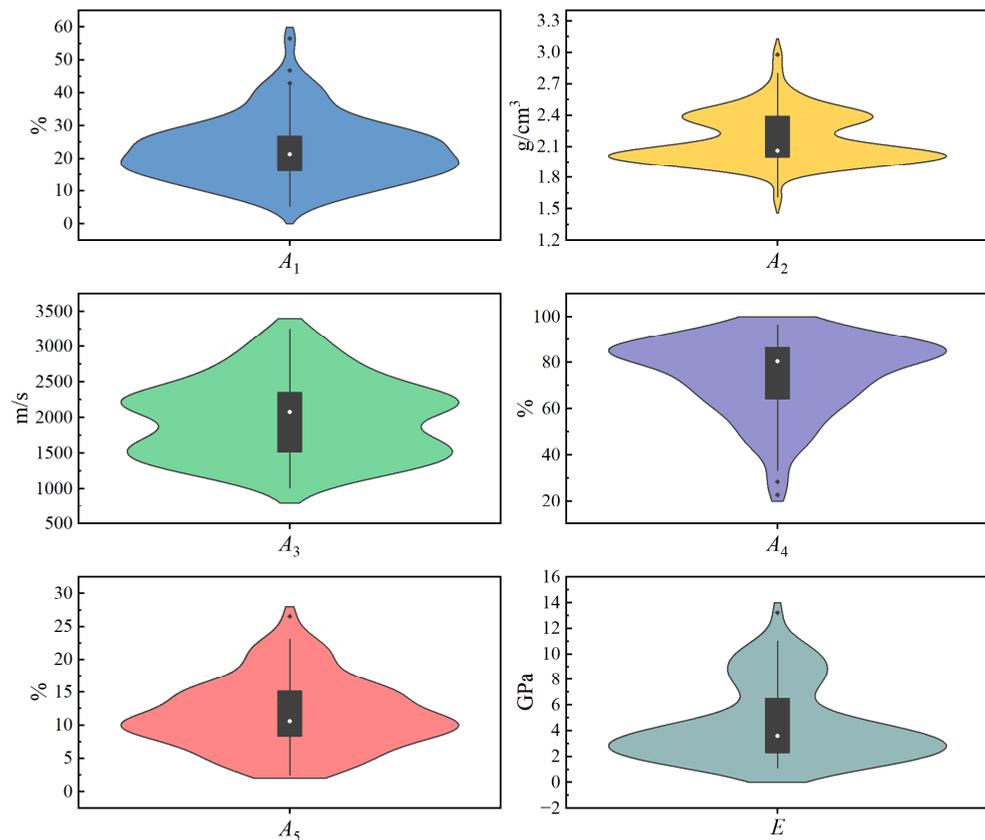


Figure 4. Violin plots of the dataset.

To visualize the distribution of this dataset, the correlation pair plots and heatmap of the Pearson correlation coefficient between indicators and E are displayed in Figure 5. The results show that A_2 , A_3 , and A_4 are positively correlated with E , while others are negatively correlated. Among them, the strongest correlation was found between A_2 and E with an absolute value of the correlation coefficient of 0.78, and the weakest correlation was found between A_1 and E with an absolute value of the correlation coefficient of 0.28.

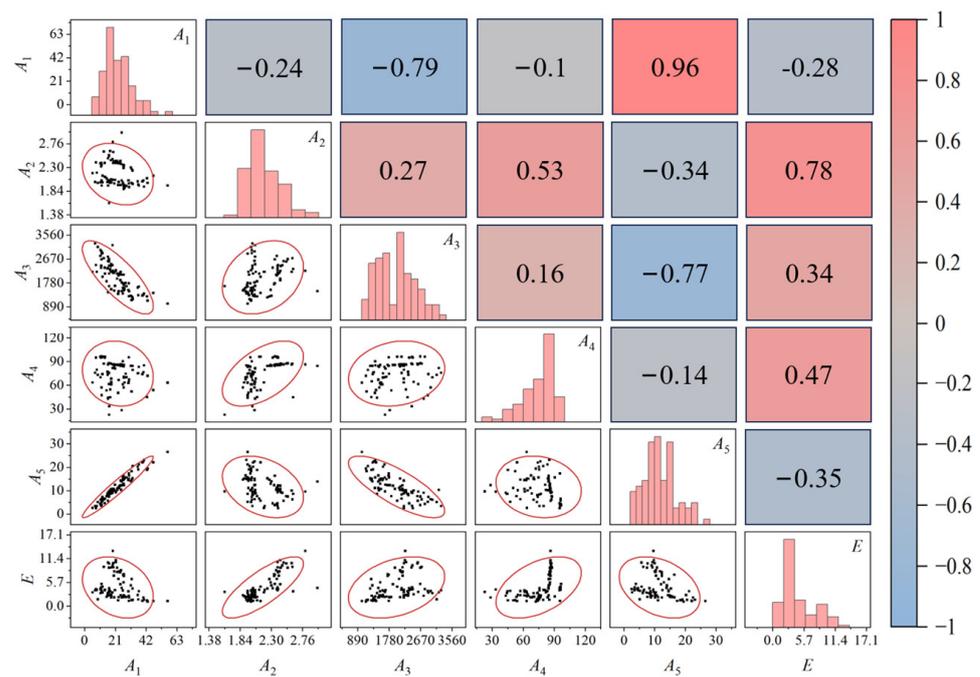


Figure 5. Correlation pair plots and heatmap.

4. Results and Analysis

4.1. Results of Hyperparameters Tuning

In contrast to other optimization methods, swarm intelligence algorithms need to determine the population size and the number of iterations, which directly affect the running time of the models and the ability to tune the optimal hyperparameters [59]. In this study, each model performed 100 iterations at population sizes of 10, 20, 30, 40, 50, 75, and 100, and the average RMSE values of five-fold CV were employed as the fitness function. The fitness curves and the running time of all models are indicated in Figure 6. With an increase in the number of iterations, the average fitness value gradually reaches a stable point. The minimum fitness value corresponds to the optimal population size of the models. However, an excessively large population size will significantly prolong the running time, which is not conducive to the practical application of these models in engineering [60,61]. Therefore, the best trade-off between performance and efficiency was achieved when the population sizes of the DMO-RF and DMO-XGBoost models were 75 and 20, respectively. After tuning the population size, the hyperparameters of the RF and XGBoost models were tuned based on the optimization algorithm. The scope, interval values, and ultimate optimization results of the hyperparameter values are presented in Table 4.

Table 4. Scope and interval values of DMO-RF and DMO-XGBoost models.

ML Algorithms	Hyperparameters	Scope of Values	Interval of Values	Optimal Values
DMO-RF	n_estimators	(10, 300)	10	10
	max_depth	(1, 20)	1	4
	min_samples_split	(2, 10)	0.001	5
	n_estimators	(10, 300)	10	12
DMO-XGBoost	max_depth	(1, 15)	1	5
	learning_rate	(0.001, 1)	0.001	1
	reg_alpha	(0.001, 1)	0.001	1
	reg_lambda	(0.001, 1)	0.001	0.636

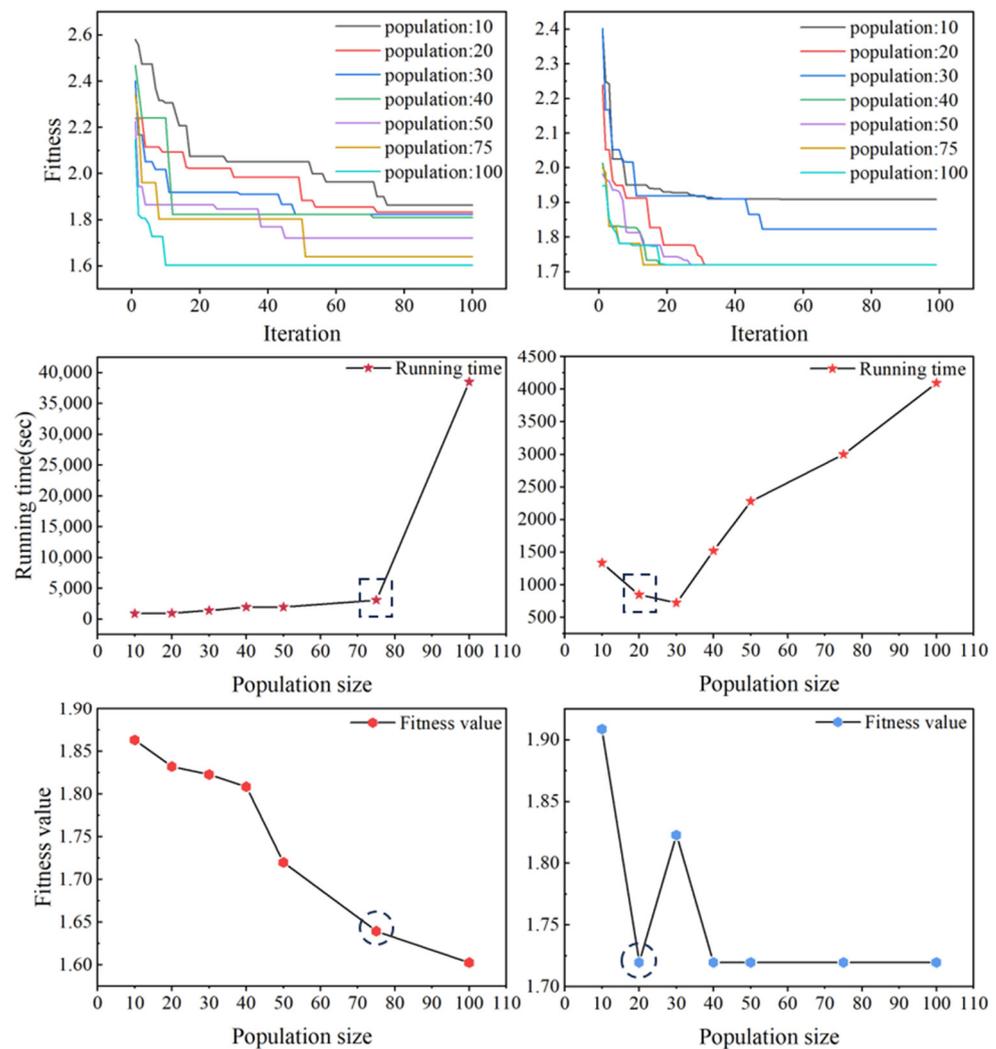


Figure 6. Results of development with DMO-RF and DMO-XGBoost models.

4.2. Model Comparison and Evaluation

To further compare the estimation performances of the two proposed hybrid ensemble learning models with the default XGBoost and RF models, the scatter plots of the four models in the training and testing phases are presented in Figure 7. The vertical and horizontal axes represent predicted and actual values of E , respectively. The diagonal dashed line represents the ideal regression line, representing that the predicted and actual values are equal. Models with more data points lying on this line exhibit higher predictive performance. The solid blue lines are the 10% boundaries, which were set to better observe the distribution of points and compare models. The green and red scatters represent the training data points and test data points, respectively. At the same time, the values of the evaluation metrics (R^2 , RMSE, MAE, and VAF) of each model in the training and testing phases are recorded in Table 5. Some observations can be obtained from Figure 7 and Table 5. First, in the training phase, the XGBoost and DMO-XGBoost models performed better than the RF and DMO-RF models without the phenomenon of the green scatters outside the 10% boundaries, and the default XGBoost model performed the best in all evaluation metrics. Second, in contrast to the training phase, the predictive performance of each model decreased during the testing phase, but the DMO approach improved the performance of the default XGBoost and RF models in the testing phase. Third, in the testing phase, the DMO-RF and DMO-XGBoost models had fewer red scatters outside the 10% boundaries than the default RF and XGBoost models, and the DMO-RF model performed best in all evaluation metrics.

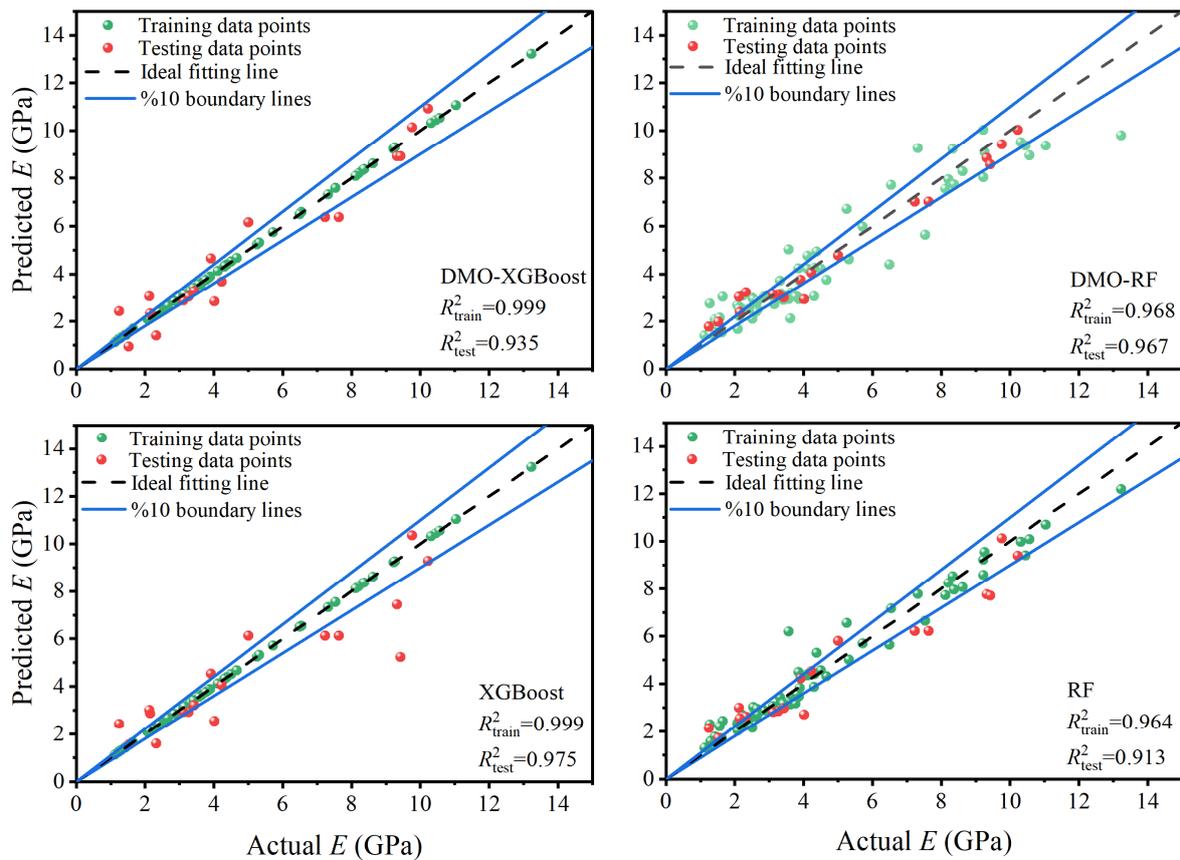


Figure 7. Regression diagrams of all models in the training and testing phases.

Table 5. Performance comparison of different models.

Models	R^2	RMSE	MAE	VAF
Training				
RF	0.964	0.541	0.368	0.965
XGBoost	0.999	0.001	0	0.999
DMO-RF	0.968	0.513	0.353	0.968
DMO-XGBoost	0.999	0.016	0.012	0.999
Testing				
RF	0.913	0.886	0.748	0.920
XGBoost	0.795	1.357	0.994	0.816
DMO-RF	0.967	0.541	0.447	0.969
DMO-XGBoost	0.935	0.763	0.674	0.936

The results presented in bold denote the best values.

5. Discussion

5.1. Comparison with Different Classical Hybrid Models

To verify the effectiveness of the proposed DMO algorithm, the classical simulated annealing (SA) and Bayesian optimization (BO) algorithms were introduced to optimize the RF and XGBoost models for comparison [62–65]. The evaluation metrics of the SA-RF, SA-XGBoost, BO-RF, and BO-XGBoost models in the training and testing phases are shown in Table 6.

To compare the developed models more visually, the Taylor diagram was plotted, as seen in Figure 8. A complete Taylor diagram consists of three components: standard deviation (SD, black and green dashed line), root mean square deviation (RMSD, red dashed line), and correlation coefficient (CC, blue dashed line) [66]. The red pentacle in Figure 8 was the reference point, which represents the actual E values, and the scatters represent

the different prediction models. The scatter with the closest distance to the reference point (with the lowest RMSD value) corresponds to the optimal prediction model. The SD, RMSE, and CC values of all scatters are calculated in Table 7. It can be seen that the DMO-RF model was the best model that was closest to the reference point with an RMSD of 0.530, followed by the BO-RF, DMO-XGBoost, BO-XGBoost, SA-RF, and SA-XGBoost models.

Table 6. Performance comparison of different hybrid models.

Models	R ²	RMSE	MAE	VAF
Training				
SA-RF	0.949	0.646	0.491	0.949
SA-XGBoost	0.997	0.029	0.022	0.997
BO-RF	0.950	0.644	0.478	0.950
BO-XGBoost	0.996	0.183	0.138	0.996
Testing				
SA-RF	0.933	0.774	0.641	0.935
SA-XGBoost	0.918	0.858	0.681	0.921
BO-RF	0.933	0.773	0.654	0.936
BO-XGBoost	0.953	0.648	0.559	0.954

Table 7. Taylor diagram values of all models.

	Reference	DMO-RF	DMO-XGBoost	SA-RF	SA-XGBoost	BO-RF	BO-XGBoost
SD	2.998	2.7	2.999	2.519	2.513	2.671	2.562
CC	0	0.988	0.968	0.977	0.968	0.981	0.975
RMSD	1	0.530	0.757	0.764	0.843	0.641	0.758

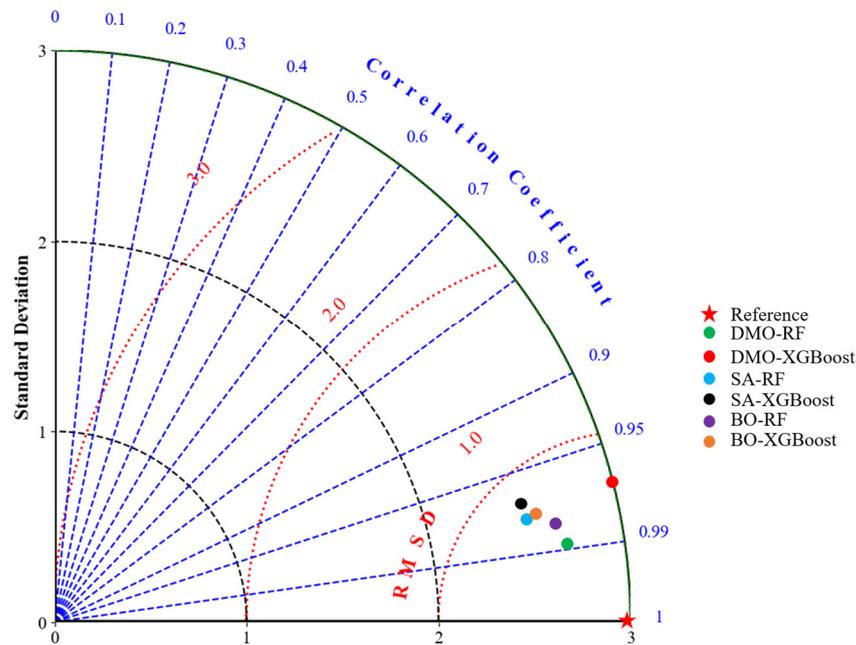


Figure 8. Comparison of models in Taylor diagram.

5.2. Comparison with Other ML Models

To confirm the performance of the developed models, three other ML models were introduced for comparison, namely the SVM, decision tree (DT), and multilayer perceptron neural network (MLPNN). The hyperparameters of these models were default. In addition, Adaptive boosting machine (Adaboost) and Category gradient boosting machine (CatBoost) methods were developed by Abdi et al. [25] on the same dataset. Furthermore, to visually compare the seven ML models, a scoring system was implemented to give corresponding

scores to each model [67]. The values and scores of all models on four evaluation metrics are calculated in Table 8. It can be seen that the overall score ranking for all models is DMO-RF (28) > DMO-XGBoost (24) > CatBoost (20) > Adaboost (16) > SVM (12) > MLPNN (8) > DT (4).

Table 8. Performance comparison of different ML models on test set.

Models	R^2	Score	RMSE	Score	MAE	Score	VAF	Score	Total Score
Adaboost	0.841	5	1.188	5	0.920	5	0.849	5	20
CatBoost	0.784	4	1.382	4	1.075	4	0.788	4	16
SVM	0.737	3	1.410	3	1.132	3	0.745	3	12
DT	0.688	1	1.679	1	1.321	1	0.701	1	4
MLPNN	0.721	2	1.473	2	1.163	2	0.722	2	8
DMO-RF	0.967	7	0.541	7	0.447	7	0.969	7	28
DMO-XGBoost	0.935	6	0.763	6	0.674	6	0.936	6	24

5.3. Relative Importance of Indicators

In this study, the relative importance of each indicator was determined by combining the DMO-RF and DMO-XGBoost models with the permutation feature importance technique, which is indicated in Figure 9 [68,69]. Based on the DMO-RF model, the rank of the indicators' importance was $A_2 > A_4 > A_3 > A_5 > A_1$. According to the DMO-XGBoost model, the rank of the indicators' importance was $A_2 > A_4 > A_5 > A_3 > A_1$. The results were consistent with the calculations of the feature importance analysis module built into the XGBoost and RF models. Apparently, two indicators were the most important: A_2 (dry density) and A_4 (slake durability). The results can be used as a reference for developing more reliable E estimation models of rock in the future.

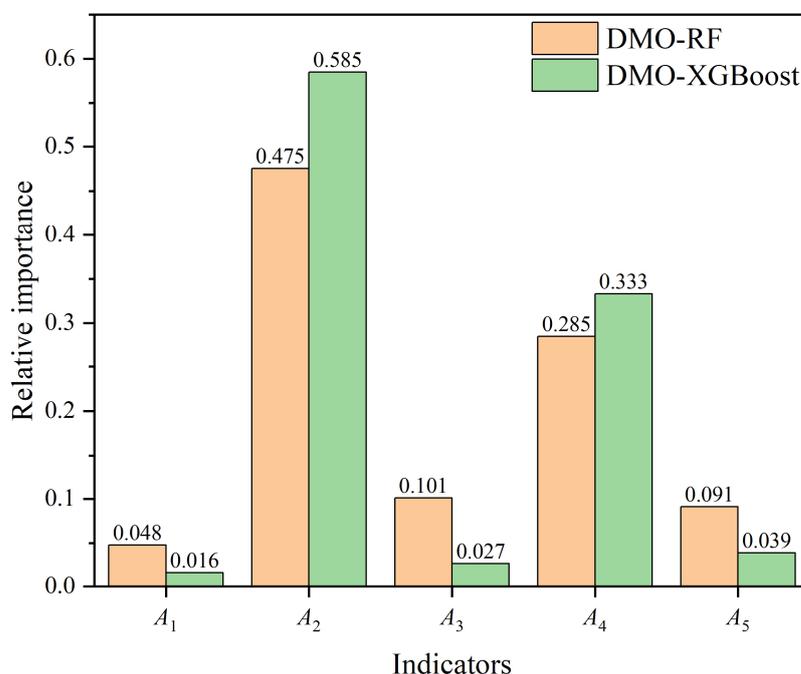


Figure 9. Relative importance of indicators.

5.4. Overfitting Validation

Furthermore, to further analyze whether the proposed DMO-RF and DMO-XGBoost models suffered from overfitting or underfitting, the convergence curves are plotted in Figure 10. The horizontal axis represents the number of samples, while the vertical axis represents the R^2 score obtained through a five-fold CV. It can be seen that the DMO-RF and DMO-XGBoost models tended to converge as the sample size increased. However, the DMO-RF model performed with less error between the training and test sets than

the DMO-XGBoost model, which indicates that the proposed DMO-RF model effectively mitigated the issue of generalization error and exhibited a certain degree of resistance against overfitting.

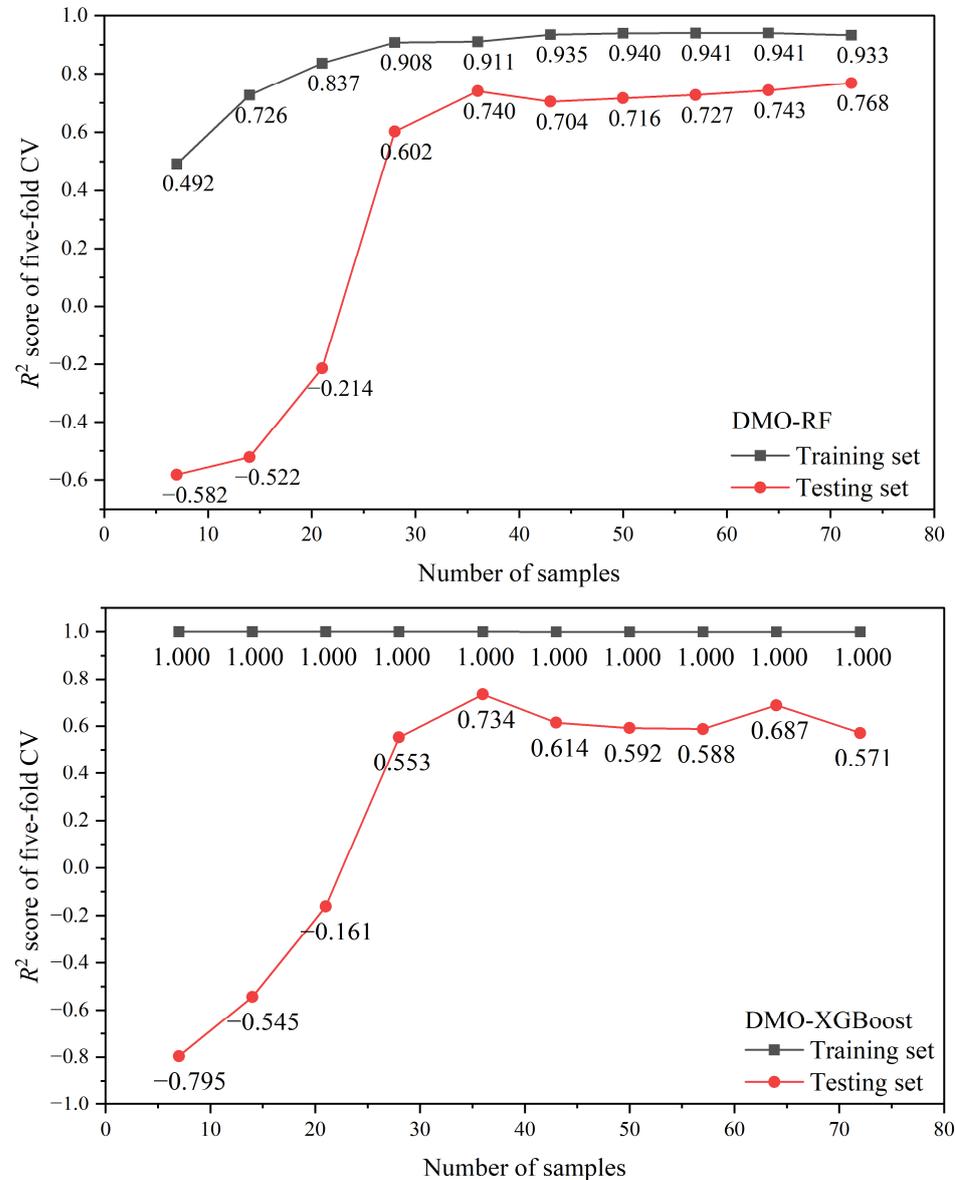


Figure 10. Convergence curves of DMO-RF and DMO-XGBoost models.

5.5. Limitations

Although the proposed hybrid ensemble approaches obtained excellent results in estimating the E of rock, there are still some limitations to consider:

(1) The dataset used in this study is relatively small, and the specimens in the original dataset were composed of four types of weak rocks, including marl, siltstone, claystone, and limestone. There is no consideration of other types of rock, such as granite, basalt, and other hard rocks. This situation might lead to the presence of a sampling bias that might affect the generalizability of the proposed ML models. Therefore, it is crucial to expand the database by collecting various types of rock cases to increase the quantity and quality of the dataset.

(2) More indicators should be considered. Although the five indicators in this study affect the E of rock significantly, other factors such as depth of coring, mineral composition,

grain size distribution, and RQD index also have an effect on the E of rock. It is valuable to explore the influences of these indicators on the E of rock estimation.

Although the aforementioned limitations exist, the developed DMO-XGBoost and DMO-RF models in this study could be considered a feasible and practical tool for estimating the E of rock. Compared to traditional expensive and time-consuming laboratory testing methods for rock mechanical parameters, the proposed models have potential advantages in terms of cost savings and time efficiency. Geotechnical engineers can significantly reduce the need for costly laboratory testing, providing project budget savings while enabling faster and more effective assessment of geologic and geotechnical risks. In addition, the developed models could be extended to estimate other geotechnical parameters, such as UCS, rock shear strength, and rock brittleness, et al.

6. Conclusions

E is one of the important parameters in rock mechanics. Accurately estimating the E of rock is significant for the design and construction of geotechnical projects. In this study, DMO-XGBoost and DMO-RF models were developed to estimate the E of rock. The effectiveness of the proposed models was verified using a database including 90 rock samples with five indicators. To avoid overfitting or selection bias, the five-fold CV method was combined with the DMO algorithm to tune the hyperparameters on the training set (80% of the dataset). The R^2 score, RMSE, MAE, and VAF were adopted to evaluate the performance of models on the test set (20% of the dataset). In addition, two default ensemble models (XGBoost and RF) were introduced and compared with the proposed two hybrid models. Overall, the DMO algorithm improved the predictive performance of the default XGBoost and RF models, and the DMO-RF model performed best with an R^2 of 0.967, RMSE of 0.541, MAE of 0.447, and VAF of 0.969 on the test set. Furthermore, the classical SA and BO algorithms were introduced to optimize the RF and XGBoost models for comparison, and the Taylor diagram was plotted to determine the comprehensive rank, which was DMO-RF > BO-RF > DMO-XGBoost > BO-XGBoost > SA-RF > SA-XGBoost. The permutation feature importance technique revealed that dry density and slake durability were more influential than other indicators in the evaluation results. Based on the convergence curves, it was verified that the DMO-RF model can reduce the generalization error and avoid overfitting.

In future work, a larger dataset containing higher-quality data should be established to improve the estimation performance. In conclusion, the proposed DMO-RF model in this paper can be considered a viable and useful tool in estimating the E of rock and can be recommended for the application of other geotechnical parameter estimations, such as UCS, rock shear strength, and rock brittleness, amongst others.

Author Contributions: Conceptualization, L.L.; methodology, L.L.; software, Z.J.; validation, W.L. and Z.J.; formal analysis, L.L.; investigation, L.L.; resources, L.L.; data curation, L.L.; writing—original draft preparation, W.L.; writing—review and editing, L.L.; visualization, W.L.; supervision, G.Z.; project administration, G.Z.; funding acquisition, G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (2018YFC0604606).

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bieniawski, Z.T. Estimating the strength of rock materials. *J. South. Afr. Inst. Min.* **1974**, *74*, 312–320. [[CrossRef](#)]
2. Zhao, Y.; Zhao, G.; Xu, L.; Zhou, J.; Huang, X. Mechanical property evolution model of cemented tailings-rock backfill considering strengthening and weakening effects. *Constr. Build. Mater.* **2023**, *377*, 131081. [[CrossRef](#)]
3. Lin, L.; Xu, J.; Yuan, J.; Yu, Y. Compressive strength and elastic modulus of RBAC: An analysis of existing data and an artificial intelligence based prediction. *Case Stud. Constr. Mater.* **2023**, *18*, e02184. [[CrossRef](#)]

4. Liu, J.; Jiang, Q.; Chen, T.; Yan, S.; Ying, J.; Xiong, X.; Zheng, H. Bayesian Estimation for Probability Distribution of Rock's Elastic Modulus Based on Compression Wave Velocity and Deformation Warning for Large Underground Cavern. *Rock Mech. Rock Eng.* **2022**, *55*, 3749–3767. [[CrossRef](#)]
5. Yang, J.P.; Chen, W.Z.; Yang, D.S.; Tian, H.M. Estimation of Elastic Moduli of Non-persistent Fractured Rock Masses. *Rock Mech. Rock Eng.* **2016**, *49*, 1977–1983. [[CrossRef](#)]
6. Thomaz, W.d.A.; Miyaji, D.Y.; Possan, E. Comparative study of dynamic and static Young's modulus of concrete containing basaltic aggregates. *Case Stud. Constr. Mater.* **2021**, *15*, e00645. [[CrossRef](#)]
7. Sonmez, H.; Gokceoglu, C.; Nefeslioglu, H.A.; Kayabasi, A. Estimation of rock modulus: For intact rocks with an artificial neural network and for rock masses with a new empirical equation. *Int. J. Rock Mech. Min. Sci.* **2006**, *43*, 224–235. [[CrossRef](#)]
8. Aksoy, C.O.; Aksoy, G.G.U.; Yaman, H.E. The Importance of deformation modulus on design of rocks with numerical modeling. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2022**, *8*, 103. [[CrossRef](#)]
9. Katz, O.; Reches, Z.; Roegiers, J.C. Evaluation of mechanical rock properties using a Schmidt Hammer. *Int. J. Rock Mech. Min. Sci.* **2000**, *37*, 723–728. [[CrossRef](#)]
10. Cilli, P.A.; Chapman, M. The Power-Law Relation Between Inclusion Aspect Ratio and Porosity: Implications for Electrical and Elastic Modeling. *J. Geophys. Res.-Solid Earth* **2020**, *125*, e2019JB019187. [[CrossRef](#)]
11. Heidari, M.; Khanlari, G.R.; Kaveh, M.T.; Kargarian, S. Predicting the Uniaxial Compressive and Tensile Strengths of Gypsum Rock by Point Load Testing. *Rock Mech. Rock Eng.* **2012**, *45*, 265–273. [[CrossRef](#)]
12. Asem, P.; Gardoni, P. A generalized Bayesian approach for prediction of strength and elastic properties of rock. *Eng. Geol.* **2021**, *289*, 106187. [[CrossRef](#)]
13. Sachpazis, C.I. Correlating schmidt hardness with compressive strength and young's modulus of carbonate rocks. *Bull. Int. Assoc. Eng. Geol.* **1990**, *42*, 75–83. [[CrossRef](#)]
14. Lashkaripour, G.R. Predicting mechanical properties of mudrock from index parameters. *Bull. Eng. Geol. Environ.* **2002**, *61*, 73–77. [[CrossRef](#)]
15. Yasar, E.; Erdogan, Y. Estimation of rock physicochemical properties using hardness methods. *Eng. Geol.* **2004**, *71*, 281–288. [[CrossRef](#)]
16. Moradian, Z.A.; Behnia, M. Predicting the Uniaxial Compressive Strength and Static Young's Modulus of Intact Sedimentary Rocks Using the Ultrasonic Test. *Int. J. Geomech.* **2009**, *9*, 14–19. [[CrossRef](#)]
17. Armaghani, D.J.; Mohamad, E.T.; Momeni, E.; Narayanasamy, M.S.; Amin, M.F.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](#)]
18. Shan, Z.-G.; Di, S.-J. Loading-unloading test analysis of anisotropic columnar jointed basalts. *J. Zhejiang Univ.-Sci. A* **2013**, *14*, 603–614. [[CrossRef](#)]
19. Karakus, M.; Kumral, M.; Kilic, O. Predicting elastic properties of intact rocks from index tests using multiple regression modelling. *Int. J. Rock Mech. Min. Sci.* **2005**, *42*, 323–330. [[CrossRef](#)]
20. Liang, W.; Sari, Y.A.; Zhao, G.; McKinnon, S.D.; Wu, H. Probability Estimates of Short-Term Rockburst Risk with Ensemble Classifiers. *Rock Mech. Rock Eng.* **2021**, *54*, 1799–1814. [[CrossRef](#)]
21. Zhang, W.; Zhang, R.; Wu, C.; Goh, A.T.C.; Lacasse, S.; Liu, Z.; Liu, H. State-of-the-art review of soft computing applications in underground excavations. *Geosci. Front.* **2020**, *11*, 1095–1106. [[CrossRef](#)]
22. Min, C.; Xiong, S.; Shi, Y.; Liu, Z.; Lu, X. Early-age compressive strength prediction of cemented phosphogypsum backfill using lab experiments and ensemble learning models. *Case Stud. Constr. Mater.* **2023**, *18*, e02107. [[CrossRef](#)]
23. Cao, J.; Gao, J.; Nikafshan Rad, H.; Mohammed, A.S.; Hasanipanah, M.; Zhou, J. A novel systematic and evolved approach based on XGBoost-firefly algorithm to predict Young's modulus and unconfined compressive strength of rock. *Eng. Comput.* **2022**, *38*, 3829–3845. [[CrossRef](#)]
24. Meng, W.Z.; Wu, W. Machine Learning-Aided Prediction of the Mechanical Properties of Frozen Fractured Rocks. *Rock Mech. Rock Eng.* **2023**, *56*, 261–273. [[CrossRef](#)]
25. Abdi, Y.; Momeni, E.; Armaghani, D.J. Elastic modulus estimation of weak rock samples using random forest technique. *Bull. Eng. Geol. Environ.* **2023**, *82*, 176. [[CrossRef](#)]
26. Acar, M.C.; Kaya, B. Models to estimate the elastic modulus of weak rocks based on least square support vector machine. *Arab. J. Geosci.* **2020**, *13*, 590. [[CrossRef](#)]
27. Pappalardo, G.; Mineo, S. Static elastic modulus of rocks predicted through regression models and Artificial Neural Network. *Eng. Geol.* **2022**, *308*, 106829. [[CrossRef](#)]
28. 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](#)]
29. Dehghan, S.; Sattari, G.; Chehreh Chelgani, S.; Aliabadi, M.A. 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. [[CrossRef](#)]
30. Majidi, A.; Beiki, M. Evolving neural network using a genetic algorithm for predicting the deformation modulus of rock masses. *Int. J. Rock Mech. Min. Sci.* **2010**, *47*, 246–253. [[CrossRef](#)]

31. Khandelwal, M.; Singh, T.N. Predicting elastic properties of schistose rocks from unconfined strength using intelligent approach. *Arab. J. Geosci.* **2011**, *4*, 435–442. [[CrossRef](#)]
32. Ocak, I.; Seker, S.E. Estimation of Elastic Modulus of Intact Rocks by Artificial Neural Network. *Rock Mech. Rock Eng.* **2012**, *45*, 1047–1054. [[CrossRef](#)]
33. Bejarbaneh, B.Y.; Bejarbaneh, E.Y.; Amin, M.F.M.; Fahimifar, A.; Armaghani, D.J.; Abd Majid, M.Z. Intelligent modelling of sandstone deformation behaviour using fuzzy logic and neural network systems. *Bull. Eng. Geol. Environ.* **2018**, *77*, 345–361. [[CrossRef](#)]
34. Saedi, B.; Mohammadi, S.D.; Shahbazi, H. Prediction of uniaxial compressive strength and elastic modulus of migmatites using various modeling techniques. *Arab. J. Geosci.* **2018**, *11*, 574. [[CrossRef](#)]
35. Rezaei, M. Indirect measurement of the elastic modulus of intact rocks using the Mamdani fuzzy inference system. *Measurement* **2018**, *129*, 319–331. [[CrossRef](#)]
36. Yang, L.; Feng, X.; Sun, Y. Predicting the Young's Modulus of granites using the Bayesian model selection approach. *Bull. Eng. Geol. Environ.* **2019**, *78*, 3413–3423. [[CrossRef](#)]
37. Liu, L.; Zhao, G.; Liang, W. Slope Stability Prediction Using k-NN-Based Optimum-Path Forest Approach. *Mathematics* **2023**, *11*, 3071. [[CrossRef](#)]
38. Jong, S.C.; Ong, D.E.L.; Oh, E. State-of-the-art review of geotechnical-driven artificial intelligence techniques in underground soil-structure interaction. *Tunn. Undergr. Space Technol.* **2021**, *113*, 103946. [[CrossRef](#)]
39. Pang, X.F.; Luan, C.F.; Liu, L.; Liu, W.; Zhu, Y.C. Data-driven random forest forecasting method of monthly electricity consumption. *Electr. Eng.* **2022**, *104*, 2045–2059. [[CrossRef](#)]
40. Wang, Y.Y.; Sun, S.F.; Chen, X.Q.; Zeng, X.J.; Kong, Y.; Chen, J.; Guo, Y.S.; Wang, T.Y. Short-term load forecasting of industrial customers based on SVM and XGBoost. *Int. J. Electr. Power Energy Syst.* **2021**, *129*, 106830. [[CrossRef](#)]
41. Guo, K.H.; Shen, C.C.; Zhou, X.K.; Ren, S.; Hu, M.; Shen, M.X.; Chen, X.; Guo, H.F. Traffic Data-Empowered XGBoost-LSTM Framework for Infectious Disease Prediction. *IEEE Trans. Intell. Transp. Syst.* **2023**, *2*, 1307–1318. [[CrossRef](#)]
42. Meng, D.L.; Xu, J.; Zhao, J.J. Analysis and prediction of hand, foot and mouth disease incidence in China using Random Forest and XGBoost. *PLoS ONE* **2021**, *16*, e0261629. [[CrossRef](#)] [[PubMed](#)]
43. Gu, Z.Y.; Cao, M.C.; Wang, C.G.; Yu, N.; Qing, H.Y. Research on Mining Maximum Subsidence Prediction Based on Genetic Algorithm Combined with XGBoost Model. *Sustainability* **2022**, *14*, 10421. [[CrossRef](#)]
44. Zhou, X.Z.; Zhao, C.; Bian, X.C. Prediction of maximum ground surface settlement induced by shield tunneling using XGBoost algorithm with golden-sine seagull optimization. *Comput. Geotech.* **2023**, *154*, 105156. [[CrossRef](#)]
45. Bhagat, S.K.; Tung, T.M.; Yaseen, Z.M. Heavy metal contamination prediction using ensemble model: Case study of Bay sedimentation, Australia. *J. Hazard. Mater.* **2021**, *403*, 123492. [[CrossRef](#)]
46. Jia, X.L.; Fu, T.T.; Hu, B.F.; Shi, Z.; Zhou, L.Q.; Zhu, Y.W. Identification of the potential risk areas for soil heavy metal pollution based on the source-sink theory. *J. Hazard. Mater.* **2020**, *393*, 122424. [[CrossRef](#)]
47. Agushaka, J.O.; Ezugwu, A.E.; Abualigah, L. Dwarf Mongoose Optimization Algorithm. *Comput. Methods Appl. Mech. Eng.* **2022**, *391*, 114570. [[CrossRef](#)]
48. Chen, T.Q.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
49. Bentejac, C.; Csorgo, A.; Martinez-Munoz, G. A comparative analysis of gradient boosting algorithms. *Artif. Intell. Rev.* **2021**, *54*, 1937–1967. [[CrossRef](#)]
50. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
51. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *mboxemph27*, 294–300. [[CrossRef](#)]
52. Mehmood, K.; Chaudhary, N.I.; Khan, Z.A.; Cheema, K.M.; Raja, M.A.Z.; Milyani, A.H.; Azhari, A.A. Dwarf Mongoose Optimization Metaheuristics for Autoregressive Exogenous Model Identification. *Mathematics* **2022**, *10*, 3821. [[CrossRef](#)]
53. Agushaka, J.O.; Ezugwu, A.E.; Olaide, O.N.; Akinola, O.; Abu Zitar, R.; Abualigah, L. Improved Dwarf Mongoose Optimization for Constrained Engineering Design Problems. *J. Bionic Eng.* **2023**, *20*, 1263–1295. [[CrossRef](#)] [[PubMed](#)]
54. Yin, X.; Gao, F.; Wu, J.; Huang, X.; Pan, Y.; Liu, Q. Compressive strength prediction of sprayed concrete lining in tunnel engineering using hybrid machine learning techniques. *Undergr. Space.* **2022**, *7*, 928–943. [[CrossRef](#)]
55. Wang, M.; Zhao, G.; Liang, W.; Wang, N. A comparative study on the development of hybrid SSA-RF and PSO-RF models for predicting the uniaxial compressive strength of rocks. *Case Stud. Constr. Mater.* **2023**, *18*, e02191. [[CrossRef](#)]
56. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
57. Thieu, N.V.; Mirjalili, S. MEALPY: An open-source library for latest meta-heuristic algorithms in Python. *J. Syst. Archit.* **2023**, *139*, 102871. [[CrossRef](#)]
58. Thieu, N.V.; Barma, S.D.; Lam, T.V.; Kisi, O.; Mahesha, A. Groundwater level modeling using Augmented Artificial Ecosystem Optimization. *J. Hydrol.* **2023**, *617*, 129034. [[CrossRef](#)]
59. Piotrowski, A.P.; Napiorkowski, J.J.; Piotrowska, A.E. Population size in Particle Swarm Optimization. *Swarm Evol. Comput.* **2020**, *58*, 100718. [[CrossRef](#)]
60. Zhou, J.; Huang, S.; Qiu, Y. Optimization of random forest through the use of MVO, GWO and MFO in evaluating the stability of underground entry-type excavations. *Tunn. Undergr. Space Technol.* **2022**, *124*, 104494. [[CrossRef](#)]

61. Zhao, G.Y.; Wang, M.; Liang, W.Z. A Comparative Study of SSA-BPNN, SSA-E NN, and SSA-SVR Models for Predicting the Thickness of an Excavation Damaged Zone around the Roadway in Rock. *Mathematics* **2022**, *10*, 1351. [[CrossRef](#)]
62. Tsai, C.W.; Hsia, C.H.; Yang, S.J.; Liu, S.J.; Fang, Z.Y. Optimizing hyperparameters of deep learning in predicting bus passengers based on simulated annealing. *Appl. Soft Comput.* **2020**, *88*, 106068. [[CrossRef](#)]
63. Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092. [[CrossRef](#)]
64. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [[CrossRef](#)]
65. Hasanpour, R.; Rostami, J.; Schmitt, J.; Ozcelik, Y.; Sohrabian, B. Prediction of TBM jamming risk in squeezing grounds using Bayesian and artificial neural networks. *J. Rock Mech. Geotech. Eng.* **2020**, *12*, 21–31. [[CrossRef](#)]
66. Zhou, Q.M.; Chen, D.L.; Hu, Z.Y.; Chen, X. Decompositions of Taylor diagram and DISO performance criteria. *Int. J. Climatol.* **2021**, *41*, 5726–5732. [[CrossRef](#)]
67. Zorlu, K.; Gokceoglu, C.; Ocakoglu, F.; Nefeslioglu, H.A.; Acikalin, S. Prediction of uniaxial compressive strength of sandstones using petrography-based models. *Eng. Geol.* **2008**, *96*, 141–158. [[CrossRef](#)]
68. Gregorutti, B.; Michel, B.; Saint-Pierre, P. Correlation and variable importance in random forests. *Stat. Comput.* **2017**, *27*, 659–678. [[CrossRef](#)]
69. Altmann, A.; Tolosi, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340–1347. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.