

## Article

# Analysis and Prediction of Grouting Reinforcement Performance of Broken Rock Considering Joint Morphology Characteristics

Guanglin Liang<sup>1</sup>, Linchong Huang<sup>2</sup> and Chengyong Cao<sup>3,\*</sup> <sup>1</sup> School of Aeronautics Astronautics, Shenzhen Campus of Sun Yat-sen University, No. 66 Gongchang Road, Guangming District, Shenzhen 518107, China; liangglin@mail2.sysu.edu.cn<sup>2</sup> School of Civil Engineering, Sun Yat-sen University, Zhuhai 519082, China; hlinch@mail.sysu.edu.cn<sup>3</sup> College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China

\* Correspondence: cy-caoszu.edu.cn

**Abstract:** In tunnel engineering, joint shear slip caused by external disturbances is a key factor contributing to landslides, instability of surrounding rock masses, and related hazards. Therefore, accurately characterizing the macromechanical properties of joints is essential for ensuring engineering safety. Given the significant influence of rock joint morphology on mechanical behavior, this study employs the frequency spectrum fractal dimension (D) and the frequency domain amplitude integral (Rq) as quantitative descriptors of joint morphology. Using Fourier transform techniques, a reconstruction method is developed to model joints with arbitrary shape characteristics. The numerical model is calibrated through 3D printing and direct shear tests. Systematic parameter analysis validates the selected quantitative indices as effective descriptors of joint morphology. Furthermore, multiple machine learning algorithms are employed to construct a robust predictive model. Machine learning, recognized as a rapidly advancing field, plays a pivotal role in data-driven engineering applications due to its powerful analytical capabilities. In this study, six algorithms—Random Forest (RF), Support Vector Regression (SVR), BP Neural Network, GA-BP Neural Network, Genetic Programming (GP), and ANN-based MCD—are evaluated using 300 samples. The performance of each algorithm is assessed through comparative analysis of their predictive accuracy based on correlation coefficients. The results demonstrate that all six algorithms achieve satisfactory predictive performance. Notably, the Random Forest (RF) algorithm excels in rapid and accurate predictions when handling similar training data, while the ANN-based MCD algorithm consistently delivers stable and precise results across diverse datasets.

**Keywords:** quantitative reconstruction of joints; numerical simulation; direct shear test; machine learning; performance prediction

**MSC:** 74M25

Academic Editor: Oleg Sergiyenko

Received: 14 November 2024

Revised: 6 January 2025

Accepted: 10 January 2025

Published: 15 January 2025

**Citation:** Liang, G.; Huang, L.; Cao, C. Analysis and Prediction of Grouting Reinforcement Performance of Broken Rock Considering Joint Morphology Characteristics. *Mathematics* **2025**, *13*, 264. <https://doi.org/10.3390/math13020264>

**Copyright:** © 2025 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

The development of infrastructure is profoundly influenced by terrain, often necessitating the construction of culverts, tunnels, and other mountainous or underground structures tailored to local conditions. When underground projects are situated in unstable and weak surrounding rocks, such as fractured rock masses, engineering challenges—including rock mass instability and collapse—can arise during construction [1]. This issue stems from the significantly lower strength of fractured rock masses compared to intact rock masses. Furthermore, the loading rate and normal stress play a critical role in the mechanical behavior

of rocks under direct shear loading. Bahaaddini et al. [2] investigated the shear behavior and asperity degradation mechanisms of rock joints through experimental and numerical direct shear tests. Their findings revealed that loading rate and normal stress influence the damage processes of rock joints, including sliding, surface wear, and asperity shearing. Similarly, Shen et al. [3] employed PFC2D to simulate sandstone models with standard JRC values, exploring the effect of normal stress on the shear strength of rock joints. They proposed the formula  $SEI = JRC + 2\sigma_n$  to classify shear failure types, determining that  $SEI > 14$  induces negative scale effects due to asperity shearing, with scale effects being more pronounced under low normal stress. This study validated the effectiveness of the BPM approach in simulating rock shear behavior and provided a quantitative framework for analyzing rock mechanical responses. Accurately determining the macroscopic mechanical properties of jointed rock masses is therefore essential for ensuring the safety of underground engineering projects, such as tunnels. Research by Huang et al. [4,5] highlights that the geometric morphology of joints is a key factor influencing joint shear stress. Consequently, accurately quantifying the effects of joint morphology on mechanical properties is a prerequisite for precise assessment of the macroscopic mechanical behavior of rock joint surfaces.

In recent decades, extensive research has focused on the quantitative characterization of joint geometric morphology. Barton et al. [6] introduced the Joint Roughness Coefficient (JRC) method, derived from comprehensive shear tests. This method estimates joint roughness by visually comparing joint profiles against ten standard reference curves. Despite its widespread application, the JRC method is prone to significant errors due to subjective interpretation. To address these limitations, Xie et al. [7] proposed the fractal dimension (D) as a parameter to intuitively represent the intricate fractal geometry of rock joints. Tse et al. [8,9] further quantified joint morphology using the root mean square of the first derivative of the contour line ( $Z_2$ ) and the roughness profile index (RP). G. Grasselli et al. [10,11] highlighted that only the contact area oriented in the shear direction contributes to shear resistance, introducing the inclination angle ( $\theta$ ) and parameter (M) to describe the anisotropic characteristics of joint morphology. However, despite these advancements, the evaluation of joint morphology remains an ongoing challenge, as a unified and standardized evaluation index system has yet to be fully established.

Over the years, researchers have extensively explored various methods for predicting shear stress, often relying on experimental data and semi-empirical models. For example, Li et al. [12,13] proposed a constitutive model to predict shear stress variations in joints under cyclic loading, considering factors such as joint roughness and the refilling of rock debris. Lin et al. [14] developed a prediction formula accounting for scale effects by assuming a random distribution of micro-unit strength in rock materials. This approach yielded results closely aligned with experimental observations. Kulatilake [15] introduced a shear stress model that explicitly incorporates the anisotropic characteristics of rock joints. Additionally, Jiao Feng and colleagues [16] performed extensive shear tests on marble joints filled with sand and gravel of varying friction coefficients using the GCTS rock shear testing system. Through statistical analysis and nonlinear fitting of the experimental data, they derived predictive formulas for shear stress. These studies collectively highlight the ongoing advancements in shear stress modeling, offering valuable insights into understanding and predicting the mechanical behavior of rock joints under varying conditions.

Empirical formulas and constitutive models have generally shown strong agreement in predicting shear stress, yet traditional prediction methods based on these models often face challenges such as low accuracy and efficiency. These limitations primarily stem from small sample sizes and the inherent uncertainty associated with empirical and semi-empirical formulas. In contrast, machine learning methods offer superior nonlinear fitting

capabilities and robust computational power, significantly improving the applicability, accuracy, and efficiency of predictive models. For example, Lin [17] achieved promising results in predicting the shear mechanical properties of joints by establishing a stable neural network topology through cross-validation and optimizing the BP neural network using algorithms like MBGD, Adam, and RMSprop. Netos et al. [18] employed BP neural networks to predict shear displacement curves and shear expansion displacement curves of rock joints under Constant Normal Stiffness (CNS) and Constant Normal Load (CNL) boundary conditions. Babanouri et al. [19] developed an ANFIS-TLBO prediction model by optimizing an adaptive neuro-fuzzy inference system (ANFIS) through algorithmic techniques. They compared the model's predictions with Barton's empirical formula, demonstrating the ANFIS-TLBO model's superior capability in capturing the nonlinear behavior of rock joints. Zhang et al. [20–25] applied machine learning methods in several studies, providing effective solutions to a range of practical engineering challenges. While numerous models have been developed to predict the strength performance of rock joint surfaces using machine learning algorithms, there remains a lack of systematic research on the adaptability and effectiveness of different machine learning models in predicting the mechanical properties of rock joint surfaces.

This paper presents a method for the quantitative evaluation of joint morphology by introducing two parameter indices: the spectrum fractal dimension (D) and the frequency domain amplitude integral (Rq). The study further achieves the quantitative reconstruction of joint geometry by utilizing Fourier transform technology. After calibrating the numerical model through laboratory experiments, the paper systematically investigates the influence of factors such as stress conditions, joint morphology, and bonding strength on the shear performance of rock joint surfaces. Multiple machine learning algorithms are employed to establish a quantitative mapping relationship between these influencing factors and the mechanical properties of rock joint surfaces. Through a comprehensive comparison of predicted results, this paper provides valuable insights and references for research in various areas. Ultimately, the paper proposes an intelligent prediction method for the mechanical properties of fractured rock masses, incorporating joint morphology characteristics.

## 2. Quantitative Indicators and Reconstruction Methods for Joint Morphology

To deeply explore the geometric characteristics of rock joints, this study employs Fourier transform techniques to perform frequency domain analysis on joint surfaces. Fourier transform [26] allows the complex geometric morphology of joint surfaces to be converted into amplitude and frequency information in the frequency domain, thereby quantifying the roughness and irregularity of the joints. By applying the Fourier transform to the joint surfaces, we capture key features of joint morphology and identify a linear relationship between amplitude and frequency in the logarithmic coordinate system. This relationship, as shown in Figure 1a, is crucial for understanding the mechanical behavior of joints. It allows us to quantify the morphological characteristics of the joints through the relationship between spectral amplitude (PSD) and frequency ( $q$ ), where the spectral amplitude PSD is inversely proportional to the frequency raised to the power of  $q$ , and the proportionality constant is the spectral fractal dimension,  $D$ . This finding provides a theoretical foundation for the quantitative reconstruction of joint morphology. As depicted in Figure 1b, the processing flowchart outlines the steps of our analysis, from data collection to the final reconstruction of joint morphology.

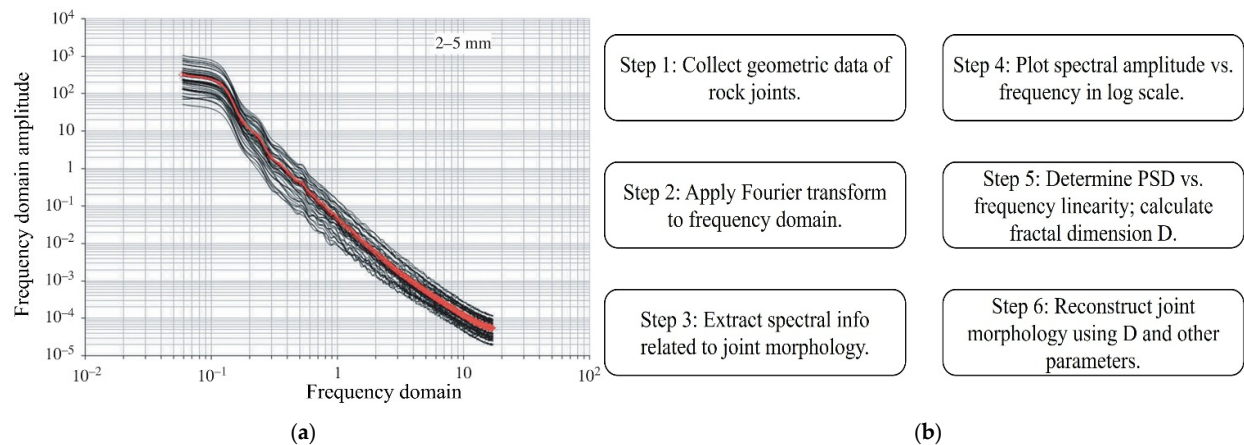


**Figure 1.** (a) Fourier spectrum analysis results. (b): Workflow of rock joint analysis.

Specifically, the Fourier amplitude at different frequencies exhibited the following relationship:

$$PSD(q \geq q_c) = C_0 \left( \frac{q}{q_c} \right)^\alpha \quad (1)$$

$$D = 1 + \alpha/2 \quad (2)$$

In the above formula,  $c_0$  represents a constant,  $\alpha$  represents the power spectral index,  $D$  denotes the spectrum fractal dimension, and  $PSD$  refers to the spectrum amplitude. The accurate formula for  $PSD$  can also be expressed as:

$$PSD(q_x) = \frac{1}{2\pi} \int A(x) e^{-i(xq_x)} dx \quad (3)$$

Here,  $A(x)$  is the autocorrelation function of one-dimensional joint elevation  $h(x)$ , and  $q_x$  represents different frequencies. The weighted integral sum of  $PSD$  in all frequency domains can be expressed as:

$$R_q = 2\pi \int PSD(q) q dq \quad (4)$$

This study selects the frequency spectrum fractal dimension ( $D$ ) and the frequency domain amplitude integral ( $R_q$ ) as the primary parameters for characterizing the roughness of rock joint surfaces. This choice is based on the application of fractal geometry theory in surface roughness characterization. As noted by Yang et al. in the Proceedings of the Royal Society A, the fractal dimension quantifies the roughness of soil particle surfaces and is closely related to the mechanical behavior of soil [27]. The fractal dimension ( $D$ ) reflects the complexity of surface roughness and captures features across multiple scales, while the frequency domain amplitude integral ( $R_q$ ) provides a quantitative assessment of the overall impact of surface roughness. By associating these parameters with the shear strength of rock joints, we can gain a more comprehensive understanding and prediction of the mechanical behavior of rock joints. Furthermore, previous studies support the selection of these parameters, indicating that fractal dimensions and frequency domain parameters effectively characterize the surface properties of rock joints and correlate with their shear performance [28–31].

Building upon the work of Yang (2016) [27], who proposed a quantitative construction method for three-dimensional joint surfaces with varying morphological features through the systematic construction of dual frequency domain amplitudes, we extend this approach by incorporating the definition of two-dimensional morphological features. This enables the development of a two-dimensional joint surface construction method capable of specifying

quantitative morphological characteristics. The detailed construction method is illustrated in Figure 2.

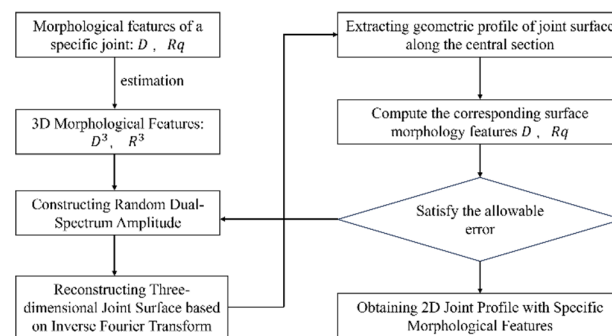


**Figure 2.** Flow chart for joint surface reconstruction with specified morphological features.

The results of the calculations reveal a strong positive correlation between the two-dimensional and three-dimensional parameters, as demonstrated in [27]. Consequently, the parameters  $D$  and  $Rq$  in the two-dimensional space can be extended to the three-dimensional space for estimation. This process involves determining  $D$  and  $Rq$  for the two-dimensional joint morphology and estimating the corresponding  $D^3$  and  $R^3$  values. A random three-dimensional joint surface is then constructed using the inverse Fourier transform, with the middle section's profile extracted for morphological analysis, as shown in Figure 3a. If the quantitative parameter error between the reconstructed joint morphology and the preset morphology meets the specified criteria, the joint contour with the desired morphological features is output, as illustrated in Figure 3b.

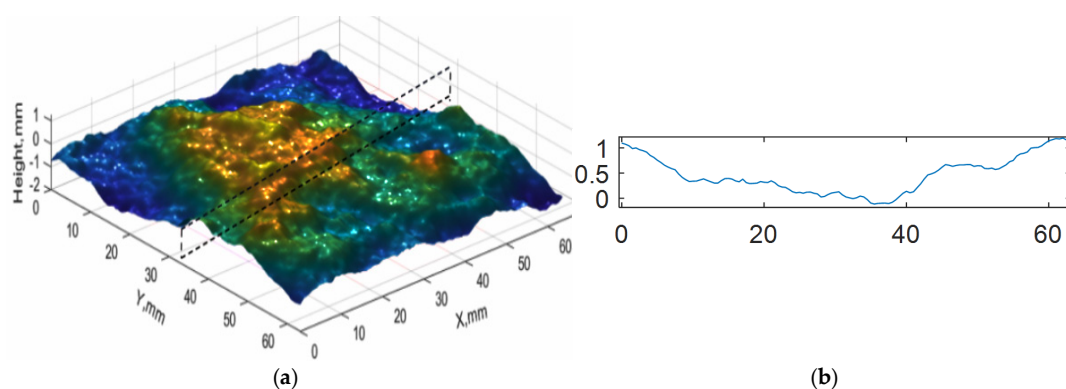


**Figure 3.** Two-dimensional joint structure with designated morphological features. (a) Interception of the axial plane of three-dimensional random joints. (b) Output of 2D joint contours with specified morphological features.

### 3. Direct Shear Test of Joint Surface

To investigate the influence of joint morphology on shear stress in rock joints and to establish initial calibration parameters for numerical simulation models, 3D printing technology was employed to perform indoor direct shear tests on joints exhibiting varying morphological characteristics.

#### 3.1. Experimental Plan

The model fabrication process and the direct shear test procedure for the rock joint surface are illustrated in Figure 4.

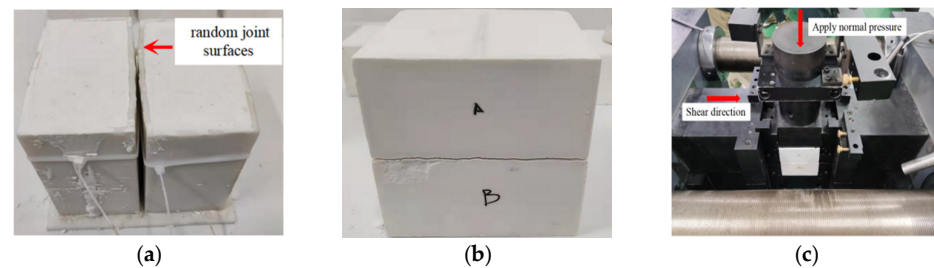


**Figure 4.** Preparation of rock sample model in occlusion state, demolding and shearing process. (a) 3D printing molds. (b) Sample in occlusal state. (c) Direct shear test.

Initially, a mold featuring the joint surface was fabricated using 3D printing (as shown in Figure 4a). Gypsum was then poured into the mold to form a cast, resulting in a sample with complete upper and lower occlusion upon demolding (see Figure 4b). Finally, the direct shear test was performed, as illustrated in Figure 4c.

In this study, we utilized Future R4600 white resin as the material for 3D-printed molds, which were employed to fabricate experimental samples. This resin is known for its precision and durability, making it ideal for producing components that require high levels of accuracy and toughness. Additionally, it offers excellent geometric adaptability, facilitating customized designs that address the limitations of traditional sampling methods. The high efficiency and convenience of Future R4600 resin enable the rapid production of a large number of consistent samples in a short time. With a heat deflection temperature of 44 °C (as per ASTM D 648 standards, measured under a 66 PSI load), the resin ensures mold stability across a range of environmental conditions. Furthermore, the resin demonstrates a tensile strength of 38 MPa and an elongation at break of 8%, with these key parameters measured according to ASTM D 638 standards, ensuring the mold material's reliability and precision when simulating rock joint structures.

In this research, plaster was selected as the rock simulation material for detailed shear tests. With a density of 2300 kg/m<sup>3</sup>, it closely mirrors the density characteristics of rock, while the Poisson's ratio of 0.25 accurately represents the relationship between lateral and axial deformation under load. Regarding shear performance, plaster exhibits a cohesion of 1 MPa and an internal friction angle of 30°, parameters that effectively characterize its cohesion and friction behavior during shearing. The friction coefficient is set at 0.3, describing the material's frictional properties during shear. Additionally, the plaster's elastic modulus is measured at 2.4 GPa, reflecting its elastic response under compression. These material properties provide a reliable foundation for the experiment and ensure accurate simulation of rock mechanical behavior.

The direct shear test was performed with fixed constraints applied to the lower half of the rock model, while normal pressure was exerted on the top. Subsequently, transverse shear was applied to the upper half of the rock model. Detailed experimental conditions are provided in Table 1, and the corresponding results of the direct shear test are presented in Figure 5.

**Table 1.** Experimental conditions.

| Quantitative Characterization Indicators of Joint Geometry |                                        | Cutting Parameters    |            |
|------------------------------------------------------------|----------------------------------------|-----------------------|------------|
| Fractal dimension $D$                                      | Frequency domain amplitude integral Rq | Normal pressure       | Shear rate |
| 1.4                                                        | 1.4                                    | 0.1 MPa; 4 MPa; 6 MPa | 0.8 mm/min |

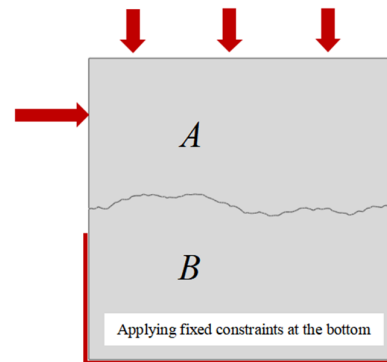


**Figure 5.** Shear scheme.

### 3.2. Numerical Calculation Model

The numerical model consists of a  $150 \times 150 \times 150$  mm square rock sample, where a global cohesive element insertion method is utilized to capture changes in contact during the simulation process. The contact settings are as follows: normal hard contact, tangential frictional contact, and triangular mesh discretization. The Mohr–Coulomb yield criterion is applied, and the simulation is solved explicitly. Boundary conditions are illustrated in Figure 5, with the bottom unit nodes fixed and a uniform load applied to the top surface, while a constant lateral displacement is imposed to simulate the shearing process. The specific parameters of the model are provided in Table 2.

**Table 2.** Numerical model calculation parameters.

| Elastic Modulus/GPa | Density/kg/m <sup>3</sup> | Poisson Ratio        |
|---------------------|---------------------------|----------------------|
| 2                   | 2450                      | 0.2                  |
| Cohesion force/MPa  | Friction angle/°          | Friction coefficient |
| 1                   | 32                        | 0.2                  |

### 3.3. Experimental Results

Figure 6 presents a comparison between the results of the direct shear tests and those obtained from the numerical simulation under varying pressures. The analysis of the graph reveals a marked increase in shear stress for both methods as vertical pressure rises, accompanied by a pronounced strain softening effect. The pressure–displacement curves from the numerical simulation closely match the experimental results across different pressures, with only minor local deviations. This consistency indicates that the numerical model accurately captures the influence of joint morphology and pressure on the shear stress behavior of rocks.

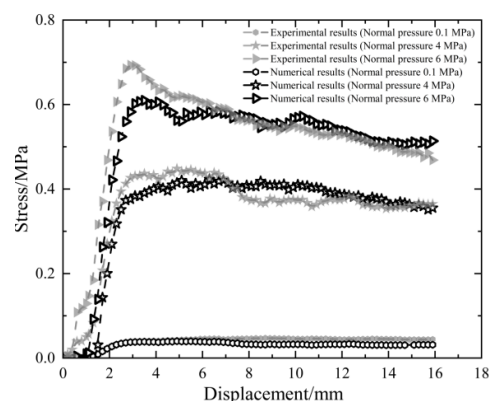


**Figure 6.** Comparison of direct shear test and numerical simulation results.

#### 4. Analysis of Factors Influencing Shear Strength of Rock Containing Joint Surfaces

Building on the validated numerical model, a further numerical model for fractured rocks was developed, incorporating joints with various morphological characteristics. Since the fractal dimension  $D$  and frequency-domain amplitude integral  $R_q$  quantitatively represent joint morphology across different scales, this study applied the aforementioned joint generation method to create 30 sets of joint profiles with varying morphological features. The specific combinations and joint profiles, with different parameters, are presented in Table 3 and illustrated in Figure 7. Considering the significant influence of the friction coefficient between the upper and lower interfaces of the joint on shear stress, three different friction coefficients were used in the calculations to examine the impact of interface friction properties on the resulting shear stress.

**Table 3.** Morphological parameters of random joint structures.

|                                                     |                              |
|-----------------------------------------------------|------------------------------|
| Fractal dimension $D$ (5 groups)                    | 1; 1.2; 1.4; 1.6; 1.8        |
| Frequency domain amplitude product $R_q$ (6 groups) | 0.4; 0.9; 1.4; 1.9; 2.4; 2.9 |
| Interface friction coefficient $\mu$ (3 groups)     | 0.2; 0.5; 0.8                |

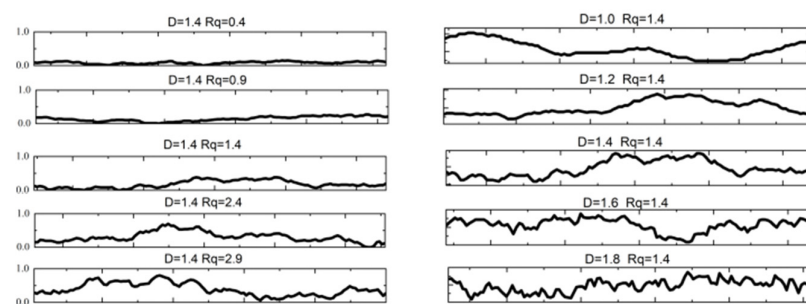


**Figure 7.** Two-dimensional joint profiles with different parameters.

Simultaneously, to investigate the influence of the friction coefficient and normal pressure between the upper and lower interfaces of the joint on shear stress, three different friction coefficients and normal pressures were applied. The effects of these four factors on shear stress were examined, with the results presented in Figures 8–11.

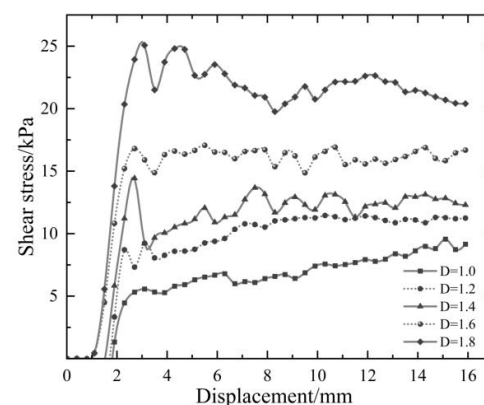


**Figure 8.** Calculation results of shear stress with different  $D$  values ( $R_q = 0.4$ ,  $\mu = 0.2$ ).

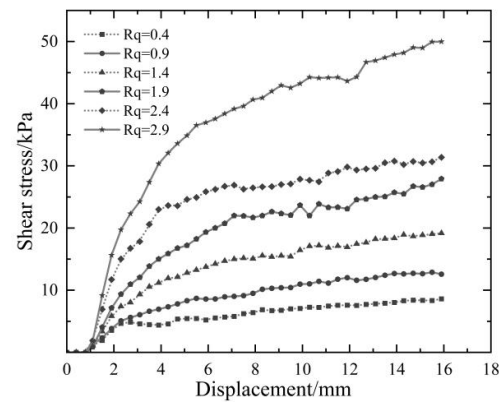


**Figure 9.** Calculation results of shear stress with different Rq values ( $R_q = 0.4$ ,  $\mu = 0.2$ ).

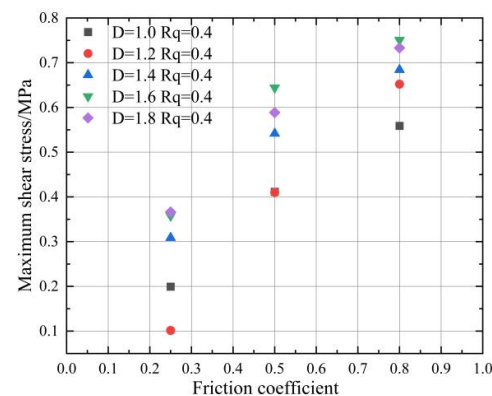


**Figure 10.** Calculation results of peak strength with different friction coefficients.

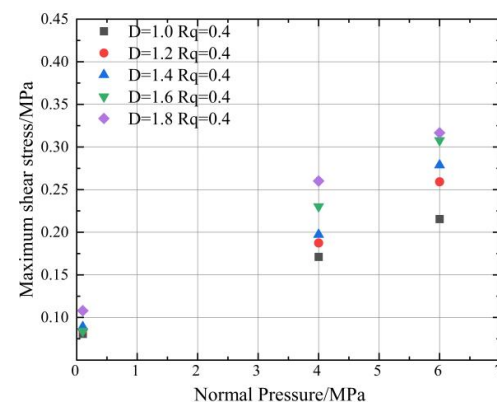


**Figure 11.** Calculation results of peak strength with different normal pressure conditions.

The shear behavior of joint surfaces is generally positively correlated with geometric morphology, interface friction coefficient, and surrounding rock pressure. As these parameters increase, the shear behavior exhibits more pronounced nonlinear characteristics. For example, when the fractal dimension exceeds 1.4, as shown in Figure 8, the pressure-displacement curve demonstrates a clear strain-softening effect. In Figure 9, the peak strength reaches a higher value shortly after the initial shear failure, indicating that post-damage, the upper and lower parts of the joint achieve a more significant occlusion. As normal pressure increases, the shear stress, as depicted in Figure 11, becomes more sensitive to changes in normal pressure. In conclusion, the mechanical properties of rock joints are influenced by both linear and nonlinear factors, which must be carefully considered for accurate prediction of shear stress in rocks containing joint planes.

## 5. Machine Learning-Based Method for Predicting Shear Strength on Joint Surfaces

Traditional statistical methods encounter several limitations in addressing nonlinear problems, such as slow computation speed, low prediction accuracy, and poor generalization capability. To overcome these challenges, machine learning algorithms have emerged as a promising alternative. Each machine learning algorithm has unique characteristics and advantages, making them suitable for different applications. Therefore, this study performs a comparative analysis of the predictive performance of various machine learning algorithms in estimating joint shear strength, based on case studies. The goal of this analysis is to provide valuable insights for selecting the most appropriate intelligent prediction algorithms for joint shear strength assessments.

### 5.1. BP Neural Network

The backpropagation (BP) neural network, a widely used type of feedback neural network, operates through error backpropagation and continuous updates to weights and thresholds [31]. As shown in Figure 12a, the BP neural network is inspired by biological neural structures and consists of an input layer, multiple hidden layers, and an output layer. Multiple initial input signals are mapped through various weight-threshold combinations, generating a new input signal for the subsequent layer.

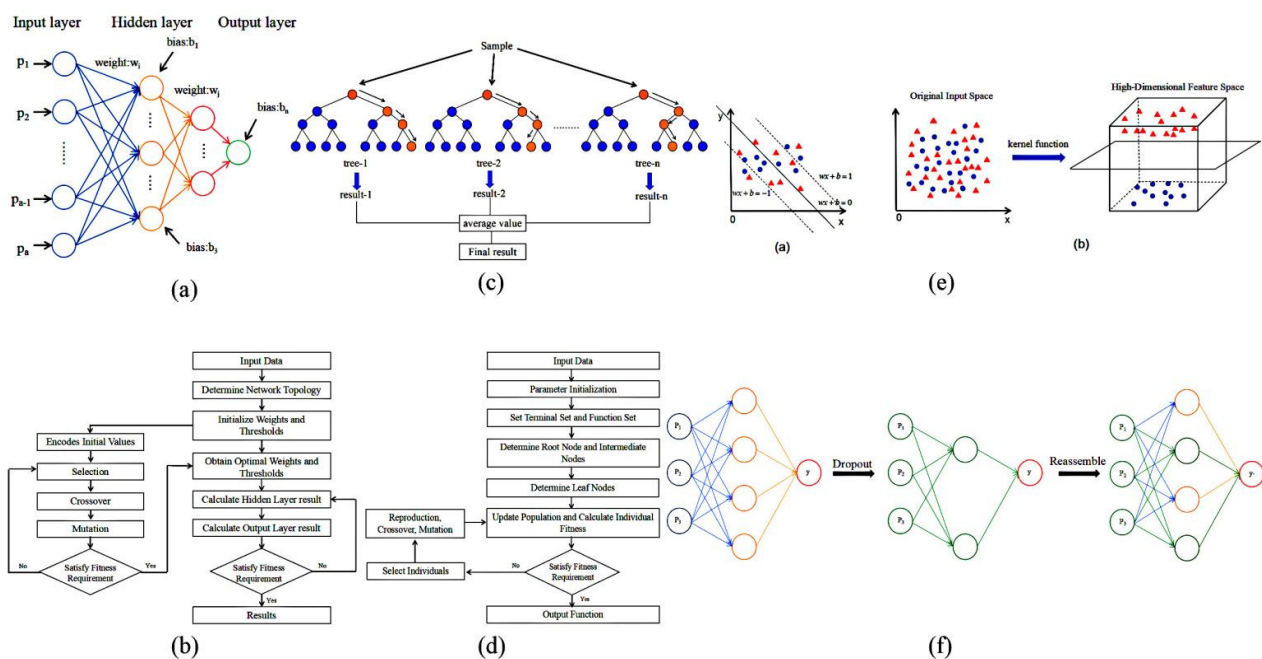


**Figure 12.** Schematic diagram of six different machine learning models. (a) Neural Network; (b) Genetic Programming; (c) Random Forest; (d) Genetic Algorithm; (e) Support Vector Regression; (f) Monte Carlo Dropout.

The output is unidirectionally transmitted to the next layer after being activated by the activation function, continuing this process until the final output is obtained. If the output fails to meet the desired accuracy, error backpropagation is employed to update the weight  $w$  and threshold  $b$ . Specifically, the gradient descent method (as shown in Formula 5) is utilized to calculate the gradient vector of the error with respect to the weights and thresholds for each layer. The weights and thresholds are then adjusted in the opposite

direction of the gradient vector [32,33], and this iterative process continues until the output reaches the desired accuracy or the predefined maximum number of iterations is met.

$$\text{grad}f(w, b) = \left( \frac{\partial E}{\partial w}, \frac{\partial E}{\partial b} \right)^T \quad (5)$$

The complexity and nonlinearity of BP neural networks increase with the addition of more hidden layers and nodes. Furthermore, the gradient descent method may lead to slow convergence and cause the final output to be trapped in local minima, which can impair the model's overall performance and accuracy.

### 5.2. GA-BP Neural Network

The GA-BP neural network improves upon the traditional BP neural network by optimizing the initial weights and thresholds using genetic algorithms, [34,35] thereby enhancing convergence speed and prediction accuracy. In contrast to the BP network, which randomly initializes weights and thresholds, the GA-BP network utilizes genetic operations such as selection, replication, crossover, and mutation to evolve better individuals, discarding those with lower fitness, as illustrated in Figure 12b. The fitness probability of individuals is given by Formula (6).

$$p_i = \frac{fit(x)}{\sum_{x=1}^i fit(x)} \quad (6)$$

The fitness function,  $fit(x)$ , is inversely related to the error function and determines the probability  $P_i$  of an individual being selected for genetic operations. Individuals with higher fitness and lower error are more likely to be chosen for the next generation, ensuring the overall improvement of the population's quality.

### 5.3. Random Forest

The random forest algorithm is a robust machine learning technique that improves the prediction accuracy of weak learners through ensemble decision-making [36]. A random forest consists primarily of regression trees and random subsets. Each regression tree is composed of a root node, intermediate nodes, and leaf nodes. At each node, input samples are recursively partitioned into sub-samples, with the final output corresponding to the results derived from the associated features, as shown in Figure 12c. A key challenge in constructing a regression tree is selecting the optimal feature distribution, which involves determining the best split points and output values for the intervals created by these splits. The binary method can be used to iteratively identify these optimal split points and intervals. Formulas (7) and (8) are applied to calculate the total minimum error and the output following segmentation, respectively [37].

$$Error = \min_{m.s} \left[ \min_{c_1} \sum_{x_i \in R_1(m,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(m,s)} (y_i - c_2)^2 \right] \quad (7)$$

$$c_1 = \frac{1}{n_1} \sum_{x_i \in R_1(m,s)} y_i, c_2 = \frac{1}{n_2} \sum_{x_i \in R_2(m,s)} y_i \quad (8)$$

The randomization process in the random forest algorithm incorporates both feature and sample randomization. Given a training sample set with M features and N samples, random samples are selected using the bootstrap sampling method, while random features are chosen through partial sampling. This dual randomization process ensures that the

trees within the random forest share some similarities but are not identical, effectively addressing both underfitting and overfitting issues. As a result, this approach improves the reliability and robustness of collective decision-making.

#### 5.4. Genetic Programming

Genetic programming (GP) follows similar design principles to genetic algorithms and exhibits structural similarities with regression trees in random forests [38]. The fitting function is generated by randomly selecting combinations from a predefined set of functions and symbols. The error between the fitting function and the true mapping relationship is then computed to establish a threshold. The prediction accuracy of the fitting function is considered reliable when this error is below a predefined value [39].

The structure and algorithm flow of genetic programming (GP) are depicted in Figure 12d. By setting parameters such as the maximum depth and selection criteria, only the root and intermediate nodes are selected from the feature set, while both the feature and terminal sets can be chosen simultaneously. The selection of sub-nodes is always restricted to the terminal set, resulting in a tree structure with either distinct or identical sub-trees. The function and symbol sets can be either fixed or customized, with the richness of the random fitting functions positively correlating with the diversity of the function and symbol sets.

#### 5.5. Support Vector Regression

Support Vector Machine (SVM) regression is a supervised learning algorithm designed to improve the generalization ability of a learning model by optimizing its structure and minimizing risk, ultimately reducing both empirical risk and the confidence range [40]. Unlike traditional linear regression, SVM regression introduces a threshold  $\epsilon$  around the target, creating a region within which errors of sample points are considered negligible, as shown in Figure 12e. Outside this region, the loss is calculated. As a result, the regression problem is reformulated as an optimization problem [41]. The optimization objective function is given in Formula (9), and the specific simplification and solution process will not be revisited here.

$$\begin{aligned} \min_{w,b} &= \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n l_{\epsilon}(f(x_i), y_i) \\ \text{s.t.} & 0 \leq y_i(w^T x_i + b) \leq 1 \end{aligned} \quad (9)$$

In this context,  $\omega$  represents the  $n$ -dimensional normal vector of the hyperplane, defining its direction, while  $b$  is the offset term indicating the distance from the hyperplane to the origin of the coordinate system. The variable  $c$  serves as a penalty term, and  $l$  denotes the loss function. To simplify the regression equation, a kernel function is employed to map the low-dimensional sample space into a high-dimensional feature space, enabling more efficient identification of the hyperplane [42], as illustrated in Figure 12e. When dealing with small sample sizes and complex data structures, Support Vector Regression (SVR) models can effectively balance computational accuracy and efficiency, as the calculation dimensionality often scales to  $n^2$ ,  $n^3$ , or higher for problems involving a sample size of  $n$ .

#### 5.6. Artificial Neural Network with Monte Carlo Dropout

Based on previous studies [43], the weight and threshold allocation in artificial neural networks (ANNs) may prioritize features that have a significant impact on the output results. As the learning process progresses, these features tend to exert a greater influence on the overall behavior of the neural network.

The ANN-MCD algorithm incorporates a node random failure mechanism, which randomly removes nodes with a specified probability during both the training and testing phases. This mechanism updates and adjusts the weights and thresholds of the remaining

nodes, resulting in a modified network structure (as illustrated in Figure 12c). By repeating this process, multiple sub-artificial neural networks are generated and treated as an ensemble. The Monte Carlo method is then employed to assess the overall error of each sub-network, with the goal of minimizing the output error [44–46].

The Monte Carlo method and the node random failure mechanism provide significant advantages by disrupting the cooperative adaptability among neurons. This disruption enables weight and threshold updates that are no longer dependent on the joint action of hidden nodes, leading to clearer and more balanced feature extraction within the modified network. As a result, the model's generalization ability and robustness are improved. By incorporating ensemble learning techniques and replacing single-precision models with multiple sub-artificial neural networks, the model becomes better equipped to handle more complex sample prediction tasks.

### 5.7. Prediction Results and Analysis

The sample set consists of four characteristic factors: fractal dimension, friction coefficient, normal pressure, and roughness, with a single output representing joint peak strength. All 300 datasets are used for training the model, and after the training process, 30 datasets are randomly selected to assess the model's predictive performance.

To comprehensively assess the applicability of each model to different samples, this paper evaluates the machine learning models based on six criteria: the training set correlation coefficient ( $R^2$ ), Mean Relative Error, Root Mean Square Error, runtime, complexity of model topology and parameter quantity, and the ability to generalize to new data.

To ensure the accuracy and robustness of our models, we meticulously preprocessed the data before presenting the prediction results. This included applying normalization to address the impact of differing scales, achieved through the following formula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (10)$$

where  $x$  represents the original data,  $x'$  represents the normalized data, and  $\min(x)$  and  $\max(x)$  are the minimum and maximum values of the data, respectively. Additionally, during the feature selection process, we based our choices on adjusted input parameters from the parameter table, including the sampling rate of independent variables and the horizontal sampling rate. These adjustments improve the model's generalization capacity and predictive accuracy. To mitigate the influence of potential outliers, we averaged the results of multiple trials, reducing the impact of random errors and enhancing the stability and reliability of the data. These preprocessing steps laid a strong foundation for our predictive models, ensuring the effectiveness of the subsequent analyses.

The correlation coefficient, which reflects the degree of correlation between the true value and the predicted value, is shown in Figure 13.

The correlation coefficient  $R^2$  for the training set reflects the effectiveness of the machine learning model in capturing the mapping relationship between the input and output. The closer  $R^2$  is to 1, the more accurate the predicted results, indicating superior prediction performance. As shown in Figure 13, the correlation coefficients for all machine learning models during the training process exceed 0.9, with no significant differences between them. This suggests that the structure of each machine learning model is scientifically sound, and the prediction results are reliable. Such high  $R^2$  values imply that the models can precisely capture the relationship between input parameters and the shear strength of rock joints, which holds significant practical implications for engineering design and construction. It not only enhances the credibility of model predictions but also provides

solid data support for engineering decisions, ensuring the safety and economy of structural design.

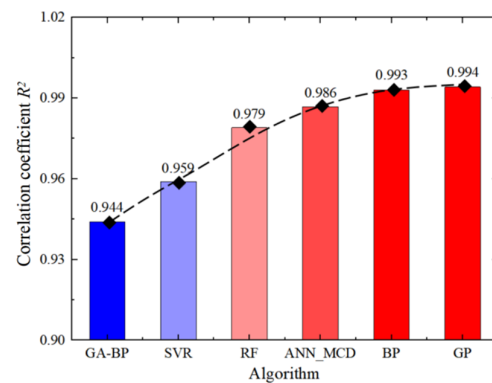


**Figure 13.** Comparison of  $R^2$  during the training process.

The Mean Relative Error (MRE) and Root Mean Square Error (RMSE) are two widely adopted metrics for evaluating regression model performance. Their mathematical expressions are provided in Equations (11) and (12):

$$MRE = \frac{\sum_{i=1}^n \frac{|y - y_i|}{y}}{n} \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y - y_i)^2}{n}} \quad (12)$$

Both the Mean Relative Error (MRE) and Root Mean Square Error (RMSE) quantify the deviation between predicted and actual values; however, they are susceptible to the influence of outliers. To address this sensitivity, outliers were identified and excluded prior to error computation. The refined results are illustrated in Figure 14.

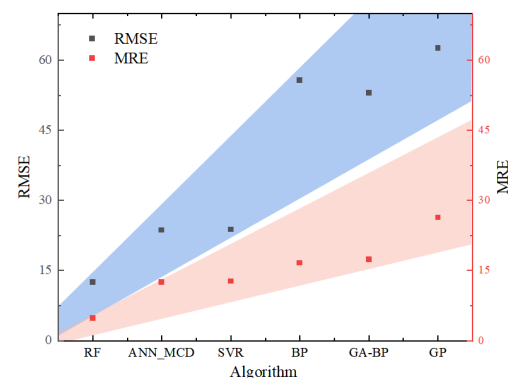


**Figure 14.** Comparison of prediction errors among different machine learning algorithms.

From the analysis presented in Figure 14, the random forest model demonstrates superior performance in error evaluation, showcasing notable advantages as a classification-based regression model. Similarly, the support vector regression (SVR) model, with its complex structure, and the ANN-MCD model exhibit commendable predictive accuracy. In contrast, the GA-BP, BP, and GP models display relatively larger errors. Notably, the GA-BP model, enhanced by genetic algorithm optimization, slightly outperforms the standard BP neural network. However, the GP model records the highest error and exhibits poor stability during training.

The runtime complexity, model topology, and parameter count are intrinsically linked to the underlying algorithmic principles and structural intricacies of machine learning models. These differences become evident during experimental execution. In this predictive experiment, the runtime performance and key hyperparameter configurations of the machine learning models are presented in Figure 15 and Table 4, respectively.

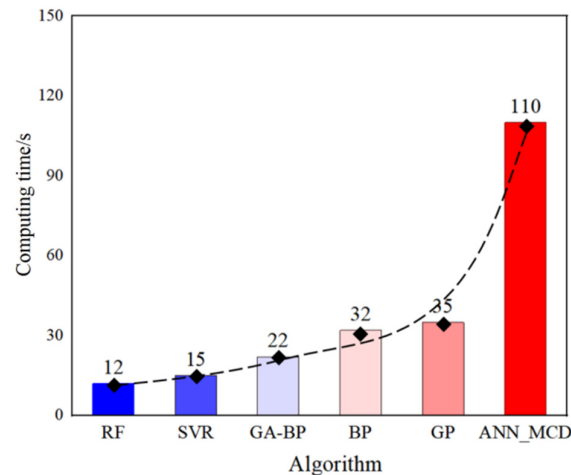


**Figure 15.** Comparison of computing time of different machine learning algorithms.

**Table 4.** Main hyperparameters of each model.

|         |                                                                                                                                                                                                                                                        |
|---------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| BP      | Number of hidden layers: 1, Number of hidden layer nodes: 7, Number of iterations: 100, Maximum error: 0.001<br>Hidden layer activation function: sigmoid, Output layer activation function: linear                                                    |
| GA-BP   | Number of hidden layers: 1, Number of hidden layer nodes: 7, Number of iterations: 100, Maximum error: 0.001, Activation function: sigmoid<br>Population size: 50, Replication probability: 0.2, Crossover probability: 0.1, Mutation probability: 0.3 |
| RF      | Population: 50, Maximum depth: 20, Feature interval: 1                                                                                                                                                                                                 |
| GP      | Population size: 200, Maximum depth: 10<br>Precision: 0.001, Population size: 50, Copy probability: 0.2, Cross probability: 0.1, Variation probability: 0.3                                                                                            |
| SVR     | Kernel function: Gaussian kernel function, Maximum coefficient: 3, Minimum coefficient: −15, Maximum regularization parameter: 15, Minimum regularization parameter: −5                                                                                |
| ANN_MCD | Maximum number of network layers: 4, Maximum number of nodes: 30, Weight threshold, Initial allocation number: 10, Random calculation times: 10, Activation function: relu<br>optimization function adamax, Iterations times: 500, Dropout rate: 0.1   |

The results demonstrate that all machine learning models can efficiently perform predictions on small sample datasets. Among them, the random forest model, leveraging independent decision trees capable of parallel computation, achieves the shortest runtime. Under identical neural network parameter settings, the GA-BP neural network surpasses the BP neural network, underscoring the genetic algorithm's efficacy in mitigating the BP network's slow convergence issue. In contrast, the ANN-MCD model, incorporating a node random failure mechanism and Monte Carlo simulations, requires the longest computation time. Overall, the computational efficiency of each model in this experiment remains significant, largely attributed to the small sample size and well-defined internal relationships among the variables. For larger and more complex datasets, the random forest or support vector regression models are recommended for balancing computational

efficiency and prediction accuracy. Conversely, for smaller datasets with less distinct factor relationships, the ANN-MCD model proves most effective in capturing true values.

The generalization ability and robustness of machine learning models are critical performance metrics, reflecting their capacity to handle unseen data and outliers, respectively—both essential for practical applications. In this study, 270 samples were randomly selected from a dataset of 300 as training data, while the remaining 30 samples were reserved for testing the models' predictive performance on new data. The prediction outcomes (Figure 16) were compared against results obtained from training on the full dataset and predictions based on known data (Figure 17).

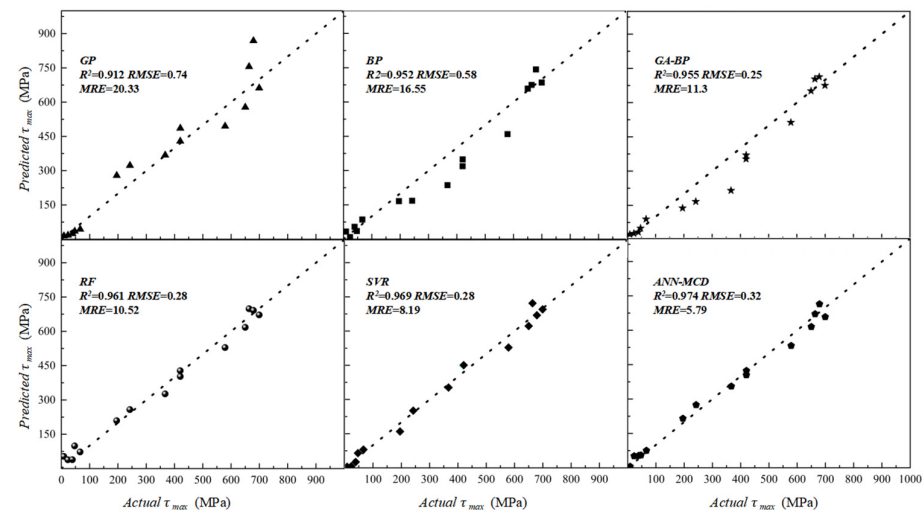


Figure 16. Prediction results of fresh data.

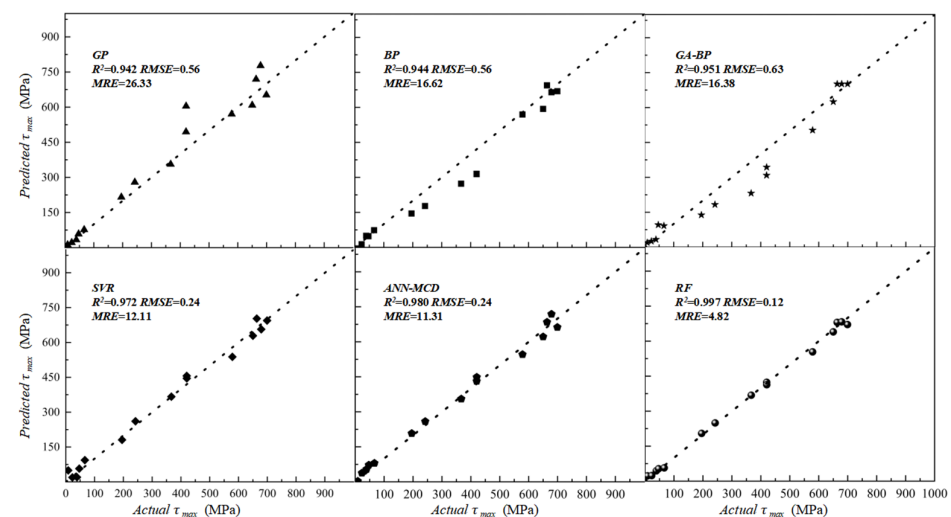


Figure 17. Prediction results of known data.

An analysis of the prediction patterns reveals that certain machine learning models, such as random forest and genetic programming (GP), experience a decline in accuracy when handling new data. The random forest model demonstrates a degree of cross-identification with unknown data, while the GP model exhibits limitations in both nonlinear mapping and generalization capabilities. In contrast, the ANN-MCD model, leveraging principles of uncertainty quantification and a reliability evaluation mechanism, maintains high prediction accuracy and stability. As sample size and complexity increase, the deficiencies in generalization and robustness of models like random forest, BP neural networks, and GP are expected to become more pronounced. Conversely, the support

vector regression (SVR) model and ANN-MCD model are less affected by variations in sample data, demonstrating superior adaptability to complex prediction tasks.

Given the inherent complexity and variability of underground engineering, the generalization ability of joint strength prediction models plays a critical role. To comprehensively assess the predictive performance of the six aforementioned machine learning models, a weighted analysis method is employed. This approach integrates multiple evaluation criteria, including the training set correlation coefficient, relative error of the validation set, root mean square error, runtime, model topology and parameter complexity, and generalization performance on new sample data. The specific calculation formula is expressed as follows:

$$R = \sum_{i=1}^n r_i \times w_i \quad (13)$$

$$s.t. w_1 + w_2 + \dots + w_n = 1$$

In the formula,  $R$  denotes the total ranking score,  $r$  represents the ranking position of an individual performance indicator, and  $w$  indicates the corresponding weight. Referring to Zorlu [40], the weights  $w_1$  to  $w_6$  are assigned as 0.1, 0.15, 0.15, 0.05, 0.1, and 0.45, respectively. The comprehensive evaluation results are illustrated in Figure 18.

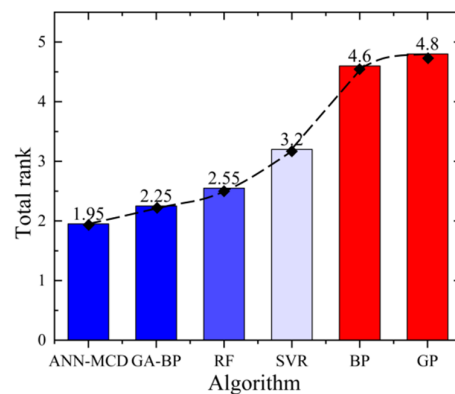


**Figure 18.** Ranking calculation results.

Based on the results of the weighted analysis, the six machine learning algorithms are ranked in terms of generalization ability as follows: ANN-MCD, SVR, GA-BP, RF, BP, and GP. This weighted evaluation ensures that key performance indicators are appropriately prioritized, providing a balanced assessment of the models. It should be noted that calculation time and prediction accuracy may vary with different datasets and parameter configurations. The experiments in this study are specific to the selected samples, and the observed model behaviors reflect these conditions. Nevertheless, these findings offer valuable guidance for selecting suitable machine learning models in future applications.

## 6. Conclusions

This study focuses on predicting joint shear stress by analyzing joint morphology. A quantitative evaluation index for joint morphology is proposed, forming the basis for a two-dimensional quantitative reconstruction method of joint morphology. Through numerical analysis and direct shear tests, the effects of various morphological characteristics and stress levels on interface shear stress are investigated. Additionally, the predictive performance of multiple machine learning algorithms is compared, culminating in the development of a robust shear stress prediction method based on machine learning. The primary conclusions are as follows:

The fractal dimension (D) and roughness (Rq) are effective indicators for characterizing the two-dimensional contours of joints. The proposed quantitative reconstruction method

for joint morphology enables the precise construction of two-dimensional joint profiles with varying morphological characteristics based on different combinations of  $D$  and  $R_q$ .

The fractal dimension ( $D$ ), roughness ( $R_q$ ), interface roughness, and normal pressure exhibit an overall positive correlation with joint shear stress. However, due to the randomness of joint profiles and the occlusion and embedment effects during the shear process, this positive correlation demonstrates localized nonlinear development characteristics.

The ANN-MCD model demonstrates the highest accuracy and robustness in predicting unknown data, making it the most reliable choice for handling unforeseen conditions in engineering applications. This capability is particularly critical for safety assessment and risk management in rock engineering projects, such as tunnels, dams, and mines, where precise predictions play a pivotal role in ensuring structural stability and operational safety.

The random forest model achieves the smallest RMSE and MRE values and the highest  $R^2$  value across both training and test sets, highlighting its superior prediction accuracy and computational efficiency. In engineering applications, this translates to the model's ability to deliver rapid and precise predictions, facilitating faster design and construction decision-making while minimizing risks arising from inaccurate forecasts.

However, distortions observed in certain sample points during fresh data prediction highlight the necessity for ongoing model validation and refinement in practical applications. The SVR and GA-BP models also demonstrate strong capabilities in capturing the complex relationships between dependent variables, achieving prediction performance comparable to the random forest model. This suggests that the optimal choice of a predictive model may vary depending on specific engineering scenarios to ensure the best predictive accuracy and reliability.

It is important to acknowledge that the performance of machine learning models can vary with different datasets and parameter configurations, and the conclusions drawn in this study are specific to the experimental samples used. Nevertheless, these findings provide valuable guidance for selecting suitable machine learning models for predicting joint shear stress. In practical engineering applications, this implies that the optimal model can be selected based on specific geological conditions and project requirements, enabling more accurate predictions of rock joint shear strength and facilitating the optimization of engineering design and construction strategies.

**Author Contributions:** G.L. Writing—original draft, Methodology, Conceptualization. L.H. Writing—review & editing, Methodology, Data curation, Funding acquisition. C.C. Writing—review & editing, Software, Funding acquisition, Formal analysis. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the National Natural Science Foundation of China (No. 52208400, 52208381), Guangdong Basic and Applied Basic Research (Grant No2024A1515012179). This financial support is gratefully acknowledged.

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

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Chen, N.; Zhang, X.; Jiang, Q.; Feng, X.; Wei, W.; Yi, B. Shear Behavior of Rough Rock Joints Reinforced by Bolts. *Int. J. Geomech.* **2017**, *18*, 04017130. [[CrossRef](#)]
2. Bahaaddini, M.; Hagan, P.C.; Mitra, R.; Khosravi, M.H. Experimental and numerical study of asperity degradation in the direct shear test. *Eng. Geol.* **2016**, *204*, 41–52. [[CrossRef](#)]
3. Shen, J.; Sun, C.; Huang, H.; Chen, J.; Wu, C. Scale Effects on Shear Strength of Rough Rock Joints Caused by Normal Stress Conditions. *Sustainability* **2023**, *15*, 7520. [[CrossRef](#)]

4. Huang, L.; Ma, J.; Lei, M.; Liu, L.; Lin, Y.; Zhang, Z. Soil-water inrush induced shield tunnel lining damage and its stabilization: A case study. *Tunn. Undergr. Space Technol.* **2020**, *97*, 103290. [[CrossRef](#)]
5. Lin, Y.; Lai, Z.; Ma, J.; Huang, L. A Combined weighted Voronoi tessellation and random field approach for modeling heterogeneous rocks with correlated grain structure. *Constr. Build. Mater.* **2024**, *416*, 135228. [[CrossRef](#)]
6. Barton, N. Review of a new shear strength criterion for rock joints. *Eng. Geol.* **1973**, *7*, 287–332. [[CrossRef](#)]
7. Xie, H.; Wang, J.-A.; Xie, W.-H. Fractal effects of surface roughness on the mechanical behavior of rock joints. *Chaos Solitons Fractals* **1997**, *8*, 221–252. [[CrossRef](#)]
8. Tse, R.; Cruden, D.M. Estimating joint roughness coefficients. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1979**, *16*, 303–307. [[CrossRef](#)]
9. Tatone, B.S.A.; Grasselli, G. A new 2D discontinuity roughness parameter and its correlation with JRC. *Int. J. Rock Mech. Min. Sci.* **2010**, *47*, 1391–1400. [[CrossRef](#)]
10. Bao, H.; Xu, X.; Lan, H.; Zhang, G.; Yin, P.; Yan, C.; Xu, J. A new joint morphology parameter considering the effects of micro—slope distribution of joint surface. *Eng. Geol.* **2020**, *275*, 105734. [[CrossRef](#)]
11. Yin, P.; Ma, J.; Yan, C.; Huang, L. Estimation of the shear strength of fractured Gosford sandstone based on fractal theory and numerical modelling. *J. Pet. Sci. Eng.* **2019**, *182*, 106278. [[CrossRef](#)]
12. Li, Y.; Du, X.; Ji, Y. Prediction of the transitional normal stress of rock joints under shear. *Int. J. Rock Mech. Min. Sci.* **2022**, *159*, 105203. [[CrossRef](#)]
13. Lin, Y.; Lai, Z.; Ma, J.; Huang, L.; Lei, M. A FDEM approach to study mechanical and fracturing responses of geo-materials with high inclusion contents using a novel reconstruction strategy. *Eng. Fract. Mech.* **2023**, *282*, 109171. [[CrossRef](#)]
14. Lin, H.; Xie, S.; Yong, R.; Chen, Y.; Du, S. An empirical statistical constitutive relationship for rock joint shearing considering scale effect. *Comptes Rendus Mécanique* **2019**, *347*, 561–575. [[CrossRef](#)]
15. Kulatilake, P.H.S.W.; Shou, G.; Huang, T.H.; Morgan, R.M. New maximum shear stress criteria for anisotropic rock joints. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1995**, *32*, 673–697. [[CrossRef](#)]
16. Feng, J.; Baohua, G.; Minglei, Z. Empirical formula for shear strength of sand filled marble joints. *Geotech. Mech.* **2018**, *39*, 7.
17. Lin, Y.X.; Wang, X.; Ma, J.; Huang, L. A finite-discrete element based approach for modelling the hydraulic fracturing of rocks with irregular inclusions. *Eng. Fract. Mech.* **2022**, *261*, 108209. [[CrossRef](#)]
18. Neto, S.; Indraratna, B.; Oliveira, D. Modelling the shear behaviour of clean rock discontinuities using artificial neural networks. *Rock Mech. Rock Eng.* **2017**, *50*, 1817–1831. [[CrossRef](#)]
19. Babanouri, N.; Fattahi, H. An ANFIS-TLBO criterion for shear failure of rock joints. *Soft Comput.* **2020**, *24*, 4759–4773. [[CrossRef](#)]
20. Zhang, P.; Yin, Z.Y.; Jin, Y.F. Machine Learning-Based Modelling of Soil Properties for Geotechnical Design: Review, Tool Development and Comparison. *Arch. Computat. Methods Eng.* **2022**, *29*, 1229–1245. [[CrossRef](#)]
21. Lin, Y.X.; Li, C.H.; Ma, J.; Lei, M.; Huang, L. Effects of void morphology on fracturing characteristics of porous rock through a finite-discrete element method. *J. Nat. Gas Sci. Eng.* **2022**, *104*, 104684. [[CrossRef](#)]
22. Zhang, P.; Yin, Z.-Y.; Jin, Y.-F. State-of-the-Art Review of Machine Learning Applications in Constitutive Modeling of Soils. *Arch. Comput. Methods Eng.* **2021**, *28*, 1–26.
23. Zhang, P.; Yin, Z.-Y.; Jin, Y.-F.; Chan, T.H.T. A novel hybrid surrogate intelligent model for creep index prediction based on particle swarm optimization and random forest. *Eng. Geol.* **2020**, *265*, 105328. [[CrossRef](#)]
24. Ma, J.; Zhao, J.; Lin, Y.; Liang, J.; Chen, J.; Chen, W.; Huang, L. Study on Tamped Spherical Detonation-Induced Dynamic Responses of Rock and PMMA Through Mini-chemical Explosion Tests and a Four-Dimensional Lattice Spring Model. *Rock Mech. Rock Eng.* **2023**, *56*, 7357–7375. [[CrossRef](#)]
25. Chen, R.; Zhang, P.; Wu, H.; Wang, Z.; Zhong, Z. Prediction of shield tunneling-induced ground settlement using machine learning techniques. *Front. Struct. Civ. Eng.* **2019**, *13*, 1363–1378. [[CrossRef](#)]
26. Zhao, L.; Chenghao, Y.; Huang, D.; Wang, X.; Zhang, S.; Jiao, K. Fourier-based reconstruction of rock joint and the statistical analysis of its morphology. *J. Railw. Sci. Eng.* **2017**, *14*, 1606–1614.
27. Yang, H.; Béatrice, A.B.; Yao, T. Characterization of the surface roughness of sand particles using an advanced fractal approach. *Proceedings. Math. Phys. Eng. Sci.* **2016**, *472*, 20160524. [[CrossRef](#)] [[PubMed](#)]
28. Santamarina, C.; Cascante, G. Effect of surface roughness on wave propagation parameters. *Géotechnique* **1998**, *48*, 129–136. [[CrossRef](#)]
29. Yimsiri, S.; Soga, K. Effect of surface roughness on small strain stiffness of soils—Micromechanical approach. In *Proceedings of the Second International Symposium on Pre-Failure Deformation Characteristics of Geomaterials: IS Torino 99*; Jamiolkowski, M., Lancellotta, R., Presti, D.L., Eds.; AA Balkema: Rotterdam, The Netherlands, 1999; Volume 1, pp. 597–602.
30. Otsubo, M.; O’Sullivan, C.; Sim, W.W.; Ibraim, E. Quantitative assessment of the influence of surface roughness on soil stiffness. *Géotechnique* **2015**, *65*, 694–700. [[CrossRef](#)]
31. Wang, Z.; Chen, Q.; Wang, Z.; Xiong, J. The investigation into the failure criteria of concrete based on the BP neural network. *Eng. Fract. Mech.* **2022**, *275*, 108835. [[CrossRef](#)]

32. Jin, B.; Lu, G.; Wang, W. A BP neural network rainfall runoff prediction model based on elastic gradient descent algorithm. *J. Shandong Univ.* **2020**, *50*, 117–124.
33. Ren, J.; Song, R.; Chen, X. Optimization of heat treatment process for 7003 aluminum alloy based on BP neural network gradient descent algorithm. *Aerosp. Mater. Technol.* **2009**, *39*, 6–9+57.
34. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
35. Wu, S.; Zhao, G.; Wu, B. Real-time prediction of the mechanical behavior of suction caisson during installation process using GA-BP neural network. *Eng. Appl. Artif. Intell.* **2022**, *116*, 0952–1976. [[CrossRef](#)]
36. Cheng, Z.-L.; Zhou, W.-H.; Garg, A. Genetic programming model for estimating soil suction in shallow soil layers in the vicinity of a tree. *Eng. Geol.* **2020**, *268*, 105506. [[CrossRef](#)]
37. Koza, J.R. *Genetic Programming: On the Programming of Computers by Natural Selection*; MIT Press: Cambridge, MA, USA, 1992.
38. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
39. Padarian, J.; Minasny, B.; McBratney, A.B. Assessing the uncertainty of deep learning soil spectral models using Monte Carlo dropout. *Geoderma* **2022**, *425*, 0016–7061. [[CrossRef](#)]
40. Gal, Y.; Ghahramani, Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016.
41. Valero-Carreras, D.; Alcaraz, J.; Landete, M. Comparing two SVM models through different metrics based on the confusion matrix. *Comput. Oper. Res.* **2022**, *152*, 305–548. [[CrossRef](#)]
42. Padierna, L.C.; Carpio, M.; Rojas-Domínguez, A.; Puga, H.; Fraire, H. A novel formulation of orthogonal polynomial kernel functions for SVM classifiers: The Gegenbauer family. *Pattern Recognit.* **2018**, *84*, 211–225. [[CrossRef](#)]
43. Yang, K.J.; Li, Z.; Huang, P.K.; Lin, Y.X.; Huang, L.C. Mechanical and failure characteristics of novel tailorable architected metamaterials against crash impact. *Engineering Failure Analysis.* **2024**, *159*, 108141. [[CrossRef](#)]
44. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
45. Lee, J.; Mitici, M. Deep reinforcement learning for predictive aircraft maintenance using probabilistic Remaining-Useful-Life prognostics. *Reliab. Eng. Syst. Saf.* **2023**, *230*, 0951–8320. [[CrossRef](#)]
46. Zorlu, K.; Gokceoglu, C.; Ocakoglu, F.; Nefeslioglu, H.A.; Acikalin, S. Prediction of uniaxial compressive strength of sandstones using petrography-based models. *Eng. Geol.* **2007**, *96*, 141–158. [[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.