A Hybrid Artificial Intelligence Model for Predicting the Strength of Foam-Cemented Paste Backfill

JINGPING QIU\(^1,2\), ZHENBANG GUO\(^1,2\), LONG LI\(^1,2\), SHIYU ZHANG\(^1,2\), YINGLIANG ZHAO\(^1,2\), AND ZHENGYU MA\(^1,2\)

\(^1\)Key Laboratory of Ministry of Education on Safe Mining of Deep Metal Mines, Northeastern University, Shenyang 110819, China
\(^2\)Science and Technology Innovation Center of Smart Water and Resource Environment, Northeastern University, Shenyang 110819, China

Corresponding authors: Long Li (long\_li@hotmail.com) and Shiyu Zhang (1710369@stu.neu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51774066, in part by the National Science and Technology Planning Project under Grant 2017YFC1503105, in part by the National Science and Technology Planning Project under Grant 2018YFC0604604, in part by the Research and Development Project, Liaoning, under Grant 2019JH2/10300051, in part by the Innovation Program for College Students, Northeastern University, under Grant 200060, and in part by the Fundamental Research Funds for the Central Universities under Grant N2001024.

ABSTRACT Foam-cemented paste backfill (FCPB) has become a trend to solve the problem of roof-contacted filling. In order to solve the time-consuming and labor-intensive disadvantages of laboratory uniaxial compressive strength (UCS) tests, a hybrid artificial intelligence model which combines random forest (RF) algorithm and grid search optimizer (GSO) was proposed for FCPB strength prediction. Moreover, the effects of foaming agent on cement hydration and pore structure characteristics were studied. The results showed that GSO can effectively tune the hyper-parameters of the proposed GSO-RF model and the developed model is an efficient and accurate tool to predict the UCS for FCPB. Though the foaming agent will not change the influence trend of cement-tailings ratio, solid content and curing time, the influence degree will be weakened by the foaming agent. The ranking of the relative importance of influencing variables is: cement-tailings ratio $>$ curing time $>$ foaming agent dosage $>$ solid content. In addition, the foaming agent has no significant effect on the hydration of cement. The foaming agent mainly changes the strength by changing the pore structure characteristics (especially the large pore volume). This research can provided some guidance for developing UCS prediction model for FCPB.

INDEX TERMS Foam-cemented paste backfill, uniaxial compressive strength, random forest, grid search optimizer, variable importance.

I. INTRODUCTION

Cemented paste backfill (CPB) is a mining filling material prepared from dewatered processing tailings, cementitious materials (such as cement) and water \cite{1,2,3}. It has many advantages such as the disposal of tailings waste, reduced ore loss and dilution rate, controlled surface subsidence and improved working environment\cite{4,5,6,7,8,9,10}. A detailed review for CPB has been provided in Ref \cite{11}, where the current status and future perspectives were discussed. Although CPB is favored by mining workers due to the above advantages, there are still many engineering and technical problems. One of the outstanding problems is that due to the dewatering and consolidation of CPB \cite{12}, it is difficult for the hardened CPB to contact the roof (see Fig. 1), which not only affect mineral resources recycling, underground mining operations security, but can also trigger a series of engineering geological hazards \cite{13,14,15,16}. Therefore, it is necessary to study the backfill that can achieve the purpose of roof contacting.

In view of the problem of roof-contacted filling, many mines in China (such as the Aoniu Iron mine, Jinchuan Nickel mine) have adopted foam-cemented paste backfill (FCPB) and achieved good results \cite{17}. Therefore, FCPB is likely to become an effective method to solve the problem of roof-contacted filling. FCPB is a mixture of CPB and foaming agent, which reduces the gap area between roof and backfill by foaming (foaming agent generates gas-phase void) \cite{18}. Moreover, compared with CPB, FCPB shows...
lower cement consumption and higher capability to absorb blasting energy [19], [20]. Like CPB, the advantages of FCPB depend on its mechanical stability [21]–[23]. The mechanical stability of CPB or FCPB is usually characterized by the uniaxial compressive strength (UCS). It can be said that the UCS is an important factor to be considered in the design of CPB or FCPB [23]–[26]. Based on this, numerous researchers [19], [26]–[31] have investigated the UCS of CPB or FCPB. Yin et al. [28] examined the effect of solid content (SC) on strength of CPB. It was found that an increase in the water content has a negative influence on the mechanical strength. Cao et al. [27] investigated the effect of cement-to-tailings (c/t) ratio on the mechanical properties of CPB. They found that the UCS performance of CPBs increases with increasing cement-to-tailings ratio and a linear equation can quantify the relationship between UCS and cement-to-tailings ratio. Fall et al. [29], Fall and Pokharel [30], and Yilmaz et al. [26], [31] also investigated experimentally the influence of curing time (T) on strength of CPB samples. One can observe from those works that CPB samples tend to get hardened with increasing curing time mainly due to the progress of cement hydration. Xu et al. [19] found that the foaming agent dosage had an important effect on the UCS of FCPB.

It is noteworthy that almost all the above studies obtained strength through laboratory UCS tests. Although the UCS test is simple, it is tedious and time-consuming, especially when a large number of UCS tests are needed [32], [33]. Some studies [34]–[38] conducted numerous tests and the relationship between UCS and ultrasonic pulse velocity, UCS and electrical resistivity, and UCS and microstructural characteristics were established to predict the strength of CPB. However, the above methods are often restricted to one type of tailings and require specific experimental equipment [34]. In addition, hydrogen peroxide is extremely oxidizing and may affect the hydration degree of binder materials [39]. Moreover, the gas produced by the decomposition of hydrogen peroxide will introduce many extra-large pores (close to the millimeter level), which makes FCPB more loose and porous than CPB. Li et al. [40] concluded that extra-large pores greatly affect the strength characteristics of materials. On the other hand, when the environment in which the hydrogen peroxide is located changes (such as cement-tailings ratio, solid content, etc.), the foaming effect also changes, that is, the bubbles have different properties (i.e., number, size, and stability) in fresh backfill [41]. It has been reported that the strength of backfill materials depends on the pore structure and hydration degree [42]. Therefore, the existing strength prediction model of CPB is no longer applicable to FCPB, and the importance degree of influencing factors of FCPB has not been studied before. It is of great significance for FCPB design to find a convenient and accurate method to predict the UCS of FCPB.

Artificial intelligence (AI) technique has been widely used in engineering areas due to its high accuracy in modeling the relationship between input variables and output variables [43]–[46]. At present, some literatures have used AI technique to study the performance of CPB [32], [34], [47]–[50]. For example, Orejarena and Fall [47], [48] employed artificial neural network (ANN) to predict the UCS of CPB, considering the effect of sulphate attack. Qi et al. [32] forecasted the compressive strength of CPB by associating the boosted regression trees (BRT) model with the particle swarm optimization (PSO). To predict the performance (for example, the slump, the bleeding rate, the compressive strength) and get the optimal mix proportion of the cemented filling material, Xue-jie et al. [49] established an improved BP Neural Network model. Qi et al. [50] introduced a genetic programming (GP) model to predict the compressive strength of CPB, where the characteristics of...
tailings, cement-tailings ratio, solid content and curing time were selected as inputs. Lu et al. [34] utilized the state-of-the-art ensemble learning method to improve the effectiveness of UCS prediction models of CPB.

Although these previous studies have made significant progress in predicting the strength of CPB by using AI technique, all of these models are developed for CPB and have limited guidance for FCPB. Therefore, the purpose of this paper is to propose a new hybrid model to accurately predict the strength of FCPB and study the degree of effect of influencing variables. Random forest (RF) is widely used in regression tasks [51], [52], mainly due to its high prediction accuracy, strong robustness, and ability to process high-dimensional data [53]. RF can also rank the importance of input variables, and when faced with high-dimensional data, it can reduce overfitting problems caused by data redundancy [54], [55]. In addition, the successful application of RF requires the combination of hyper-parameter optimization algorithms. Grid search algorithm (GSO) algorithm is widely used in hyper-parameter tuning and its stability in hyper-parameter tuning has been proved [56]–[58]. Therefore, this paper utilized the GSO to optimize the hyper-parameters of RF. To the author’s knowledge, there are no relevant reports about using AI technology to predict the strength of FCPB. In addition, it is the first time in the literature to study the combined effects of cement-tailings ratio, solid content, curing time and foaming agent dosage (HP) and quantify the relative importance of these variables of FCPB.

II. METHODS

This section is divided into three parts, including the methodological backgrounds of RF and GSO, the descriptions of k-fold cross-validation and model performance measures.

A. RF MODEL

RF algorithms are widely used in classification and regression problems, which mainly profits from good robustness, high accuracy, and returnable characteristic measures [59]–[61]. Using bootstrap aggregating technology to extract multiple samples from the original sample and model each bootstrap sample with a decision tree (the data set size of each sample is the same) [62], [63]. A RF regression predictor [64] can be expressed as:

\[
\hat{f}_{RF}^n(x) = \frac{1}{n} \sum_{i=1}^{n} S_i(x)
\]  

(1)

where \(x\) and \(n\) represent the vectored input variable and the number of trees, respectively; \(S_i(x)\) refers to a single regression tree constructed from a subset of input variables and the bootstrapped samples.

The main implementation steps of RF (Fig. 2) are summarized as follows:

1) Using bootstrap to generate \(n\) samples from the original sample set to form a subset, these subsets are used to grow the tree [65];

2) When establishing a decision tree, a certain number of features are randomly selected, and the most suitable feature is selected as the split node according to the mean square error until it no longer splits [66]. Part of the data in the original sample (1/3 of the total data) will not be used for training and is called Out-of-Bag (OOB) data, which is used to estimate the model’s generalization error [67], [68];

3) In the end, the output of the RF is determined by \(n\) decision trees, where the output of the regression problem is the average of the output of the \(n\) decision trees, and the output of the classification problem is determined by the voting of the \(n\) decision trees [66].

B. GRID SEARCH OPTIMIZER (GSO)

Grid search algorithm, also known as exhaustive attack method, is often used to solve the problem of constrained nonlinear extremes. The principle of the algorithm is to divide the hyper-parameters into grids at a certain interval within a certain spatial search range, and each intersection in the grid corresponds to an objective function value [69]. By traversing each point in the grid regularly to find the intersection of the optimal objective function, the optimal combination of hyper-parameters can be obtained [69], [70]. In the grid search algorithm, the hyper-parameters of each group are independent of each other, and the multi-solution problem caused by possible coupling between hyper-parameters is avoided in the optimization process [71], [72].

C. K-FOLD CROSS-VALIDATION AND MODEL PERFORMANCE MEASURES

In this study, a k-fold cross-validation method was adopted. This method randomly divides the training set (UCS data) into \(k\) folds, the first fold is used as the validation set and the remaining \(k-1\) folds are used as the training set [43].
It is worth noting that every fold of data will be trained and validated, and the model overfitting can be reduced in the k training-validating rounds [20], [21]. The 10-fold cross-validation is adopted to optimize the process of hyper-parameters tuning (Fig. 3).

The model performance is evaluated by the mean square error (MSE), the mean absolute error (MAE), and the coefficient of multiple determinations ($R^2$). The calculations are as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$  \hspace{1cm} (2a)

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$  \hspace{1cm} (2b)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$  \hspace{1cm} (2c)

where $n$ is the sample size; $y_i$ is the predicted value; $\hat{y}_i$ is the observed value; and $\bar{y}$ is the mean of $\hat{y}_i$.

III. MATERIALS AND EXPERIMENTS

The tailings tested were obtained from the ore processing plant of Aoniu Iron Mine (Liaoning province, China). The particle size distribution of tailings was determined using a Malvern laser Mastersizer 2000, as shown in Fig. 4(a). It can be seen that the fine (minus 20 µm) content of tailings is 15.31%, which can be classified as a coarse size tailings material [21]. The mineral characteristics of tailings were obtained by X-ray diffraction (XRD), as shown in Fig. 4(b). As can be seen from the figure, the crystalline components in tailings are mainly quartz, calcite and dolomite. In addition, X-ray fluorescence (XRF) analysis results showed that the main chemical components of tailings are SiO$_2$ (64.4%), Fe$_2$O$_3$ (18.04%) and Al$_2$O$_3$ (5.65%), as detailed in Table 1. The binder used in this study was ordinary Portland cement (OPC) type 32.5R according to Chinese National Standard GB 175-2007 [73]. The main chemical composition of the cement is shown in Table 1. Tap water was used to mix the materials to achieve the desired consistency of the slurry.
In this experiment, hydrogen peroxide (H$_2$O$_2$, Fig. 5 (a)) with a concentration of 30% was selected as the foaming agent. In alkaline environment, hydrogen peroxide is decomposed to produce oxygen (see formula (3)), leading to the volume expansion of the slurry. In addition, calcium stearate (Fig. 5 (b)) was added as foam stabilizer to improve the bubble distribution in the slurry and subsequent to improve the mechanical strength of FCPB [74]. It is worth noting that, after a lot of preliminary experiments, the optimal dosage of foam stabilizer is 0.7% of the total mass of tailings and OPC.

$$2H_2O_2 \rightarrow 2H_2O + O_2 \uparrow$$  \hspace{1cm} (3)

After the experimental materials were prepared according to the experimental scheme in Table 2, the FCPB slurry was manufactured according to the experimental steps shown in Fig. 6, and then the slurry was poured into a plastic mould with a diameter of 50 mm and a height of 100 mm. After sealing, the mould was cured in a curing box. The curing conditions are consistent with those described in Ref [22].

Following a predetermined curing time of 3, 7, 14 and 28 days, the FCPB specimens (Fig. 5 (c)) were demoulded and subjected UCS test (Fig. 5(d)). The UCS tests were performed at a loading rate of 0.5 mm/min using a computer-controlled mechanical press (Humboldt HM-5030, USA), and the ratio of the maximum stress value ($F$) in the stress-strain curve to the compression area of the specimen ($A$) was regarded as UCS (see formula (4)).

$$UCS = \frac{F}{A}$$  \hspace{1cm} (4)
TABLE 3. Statistical description of inputs and outputs.

| Variables                  | No. | Min  | Max  | Mean | Std  | Skewness | Kurtosis |
|----------------------------|-----|------|------|------|------|----------|----------|
| Cement-tailings ratio      | x1  | 0.125| 0.25 | 0.181| 0.05 | 0.356    | -1.51    |
| Solid content              | x2  | 0.72 | 0.76 | 0.74 | 0.016| 0        | -1.51    |
| Foaming agent dosage       | x3  | 0.015| 0.03 | 0.023| 0.006| 0        | -1.37    |
| Curing time                | x4  | 3    | 28   | 13   | 9.546| 0.634    | -1.11    |
| UCS                        | y   | 0.04 | 1.342| 0.415| 0.292| 1.064    | 0.495    |

ratio of the cement paste sample is 1:1, and the curing time is 4 hours [75]. FCPB samples with fixed variable values (solid content: 72%, cement-tailings ratio: 1:4 and curing time: 28 days) were prepared at foaming agent dosage varying from 0% to 3%. It is worth noting that due to the measurement range of the mercury intrusion porosimetry (MIP), this paper uses image processing methods to measure the size of the large pores (>200 µm). The specific operation steps can refer to the Ref [40].

IV. RF WITH GSO

This section will introduce how to combine the GSO with RF to predict the FCPB strength. The modelling process of GSO-RF mainly includes three parts: datasets, feature correlation analysis, and hyper-parameter optimization.

A. DATASET

In this paper, a total of 432 UCS tests were performed. Three tests were performed for each mix proportion of FCTB and the mean value was used for modelling. Therefore, 144 data were eventually used for model training and performance evaluation.

There are four input variables of GSO-RF model, namely, cement-tailings ratio, solid content, foaming agent dosage and curing time. The output of GSO-RF is UCS of FCPB. Table 3 shows the statistical results of input and output. In the supervised task, the original data is divided into training set and testing set. The training set is selected for hyper-parameter optimization, while the testing set is utilized for evaluating the generalization ability of the model. The training set is obtained by randomly extracting 70% of the original dataset, and the remaining 30% is used as testing set. The specific method is to gradually increase the size of the training set until a steady model performance is reached [43]. Therefore, the UCS data of the training set and the testing set are 100 and 44, respectively.

B. FEATURE CORRELATION ANALYSIS

Before the variables are introduced into the model, it is necessary to perform a correlation analysis to avoid the decrease of the model prediction accuracy caused by the redundancy between the data [76]. Pearson correlation coefficient is widely used in correlation analysis between data [77]–[79]. Therefore, this paper chooses the Pearson correlation coefficient to analyse the four input variables. From Fig. 7 it can be seen that low correlation exists between the four input variables, which can be used as input to the model.

C. HYPER-PARAMETERS TUNING

For different datasets, the optimal performance of RF depends on the optimal combination of different hyper-parameters. Based on the results of previous studies and data characteristics, this paper focuses on analysing four hyper-parameters of RF, including: the number of decision tree ($N_{\text{tree}}$), the maximum depth ($\text{Max_depth}$), the minimum samples required at a leaf node ($\text{Min_samples_leaf}$) and the minimum number of samples required to split an internal node ($\text{Min_samples_split}$). The 10-fold cross-validation and grid search algorithm were used to optimize different combinations of hyper-parameters, and the mean square error was
TABLE 4. Hyper-parameters of RF tuned by GSO.

| Hyper-parameters | Explanations                                      | Tuning range | Interval |
|------------------|--------------------------------------------------|--------------|----------|
| ntree            | The number of decision trees                     | 50-300       | 25       |
| max_depth        | The maximum depth of RF                          | 1-10         | 1        |
| min_samples_leaf | The minimum samples at the leaf node             | 1-10         | 1        |
| min_samples_split| The minimum samples to split an internal node    | 2-10         | 1        |

FIGURE 9. The influence of a) C/t ratio (SC:76%, HP:3%, T:14 d), b) SC (c/t:1:4, HP:1.5%, T:14 d), c) HP (c/t:1:4, SC:76%, T:14 d) and d) T (c/t:1:4, SC:76%, HP:1.5%) on the UCS of FCPB/CPB.

utilized to evaluate the model performance under different combinations of hyper-parameters. Although the grid search algorithm takes more time to optimize hyper-parameters, this method has been proved to be feasible and stable in hyper-parameter optimization (Table 4) [80], [81]. Fig. 8 illustrates the procedures for predicting the UCS of FCPB by applying the proposed GSO-RF model.

V. RESULTS AND DISCUSSION

A. EXPERIMENTAL RESULTS

1) RESULTS OF UCS TESTS

The cement-tailings ratio, solid content, foaming agent dosage, and curing time are the four major factors that affect the strength of the FCPB. To better study the influence of these factors on the UCS of FCPB, the control variable method was used for further analysis [23]. As can be seen from Figure 9, the cement-tailings ratio, curing time and solid content are positively correlated with the UCS of FCPB. This indicates that the UCS increases with the increase of these factors. The overall positive relationship between the above factors and the UCS of FCPB is consistent with findings in CPB samples [4], [23], [35], indicating that the foaming agent will not change the influence trend of these factors.

It is worth noting that although the influence trend of these factors on the strength of FCPB is generally unchanged, it can be clearly seen that the degree of influence is weakened...
FIGURE 10. Portlandite content for 4 h cement paste with various HP.

compared to CPB. For example, when FCPBs at the age of 3, 7, 14 and 28 days (cement-tailings ratio, solid content and foaming agent dosage are fixed at 1: 4, 76% and 1.5%, respectively), the UCSs are 0.47, 0.92, 1.16 and 1.34 Mpa, the corresponding growth percentages are 96.6%, 146.6% and 185.5%, while the corresponding values of CPB are 101.2%, 189% and 250.8%. As for foaming agent dosage, it can be seen that the UCS of FCPB has a significant negative correlation with the foaming agent dosage, that is, with the increase of foaming agent dosage, the strength UCS continuously decreases. This is to be expected because the increase in foaming agent dosage leads to increased porosity in the FCPB, which reduces the compactness and makes FCPB more prone to failure during loading.

2) RESULTS OF MICROSTRUCTURE TESTS

Figure 10 shows the amounts of portlandite (Ca(OH)$_2$) formed in cement pastes with different dosages of foaming agent. The amount of Ca(OH)$_2$ is an important indicator reflecting the degree of hydration [75]. It can be clearly seen from the figure that the amounts of Ca(OH)$_2$ corresponding to the samples with different dosages of foaming agent are almost the same, which shows that although hydrogen peroxide has strong oxidizing, it has no significant effect on the hydration of cement. The reason for this phenomenon may be that the dosage of hydrogen peroxide is too low, and the relatively high water content further weakens the effect of hydrogen peroxide. Figure 11 shows the pores of FCPB or CPB with various foaming agent dosages. It can be seen from Fig. 11(a) that the porosity of FCPB (49.8%, 60.25%) is much larger than that of CPB (35.2%). Moreover, the more the foaming agent is added, the larger the porosity. This is caused by the decomposition of the foaming agent into gases. The large pore volume is shown in Figure 11(b). It is clear that the large pore volume of FCPB is much larger than that of CPB. At the same time, the large pore volume is positively correlated with the foaming agent dosage.

B. RESULTS OF THE HYPER-PARAMETERS TUNING

This study used a grid search algorithm to optimize the hyper-parameters of the UCS prediction model. Therefore, the four hyper-parameters need to be divided according to a certain distance. The specific division is as follows:

In the RF hyper-parameters, the number of decision tree interval is 25, and the intervals of max_depth, min_samples_leaf and min_samples_split are all kept constant at 1. The 10-fold cross-validation was performed on the data in the training set through a grid search algorithm combined with the MSE evaluation function, and finally the optimal hyper-parameters were obtained. It should be noted that when the mean value of MSE reaches the maximum under 10-fold cross-validation, the optimal hyper-parameters of the UCS prediction model are optimal (Table 5). The GSO-RF model equipped with the optimum hyper-parameters is trained by utilizing the training dataset and then the performance on the training set can be obtained [43].

C. RESULTS OF THE GSO-RF MODEL

The performance of the GSO-RF model was verified on the training set and the testing set using the $R^2$, MSE and MAE, which are determined by experimental and predicted UCS values. In addition, the ratio of experimental UCS to predicted UCS was analysed.

The comparison between the predicted UCS values obtained from the RF model with the optimum hyper-parameters and the experimental UCS data in the training dataset as shown in Fig. 12. It can be seen from Fig. 12(a) that most of the predicted UCS values are very close to the corresponding experimental results. Consequently, it could be said that the proposed GSO-RF model can reliably describe the nonlinear relationship between UCS and the influencing variables. In other words, the proposed model has great potential for the UCS prediction of FCPB. Moreover, as shown in Fig. 12(b), the MSE and MAE of the GSO-RF model are 0.0004 and 0.0158, respectively. The smaller MSE and MAE, the lower the deviation between the experimental and predicted UCS values [43]. The $R^2$ between the predicted and tested UCS values reached up to 0.99, which demonstrates that a good positive correlation is achieved for the UCS training set [32].

The testing set is used to verify the generalization capability of the proposed model [33]. Thus, the testing set was used to evaluate the trained GSO-RF model, and the predicted

| Hyper-parameters | Type   | The optimal parameters |
|------------------|--------|------------------------|
| $N_{tree}$       | Integer| 100                    |
| Max_depth        | Integer| 10                     |
| Min_samples_leaf | Integer| 1                      |
| Min_samples_split| Integer| 2                      |
results are illustrated in Fig. 13. Similar to the results shown in Fig. 12, most of the predicted UCS values were in good agreement with the experimental UCS values, with almost no significant outliers. This can be attributed to the fact that training samples are relatively sufficient. It is worth noting that the $R^2$ value of GSO-RF model on the testing set is 0.9586 (Fig. 13(b)), although lower than the $R^2$ value on the training set, the goodness of fit of the model can meet the experimental prediction needs.

Fig. 14(a) and (b) show the histogram plots of the ratio of experimental/predicted UCS values obtained by the training and testing sets using the GSO-RF model, respectively. It can be clearly seen that the median and mean of density curves for both training and testing sets are close to one. Specifically, the median mean and mean value of the training set are 0.9852 and 0.9821, respectively, while corresponding values of the testing set are 1.028 and 1.058, respectively. This indicates that in the training set, the UCS predicted by the GSO-RF model was slightly less than the actual value, while in the testing set, the result was the opposite. Nevertheless, the results show that the proposed GSO-RF model can reliably predict the UCS of FCPB.

D. RELATIVE IMPORTANCE OF INFLUENCING VARIABLES

To better understand the influence of variables on UCS prediction, partial dependence plots [82] and feature importance scores [83] were introduced to perform variable sensitivity analysis. Partial dependency plots can be utilized to explain the relationship between the output target and each set of input features, and can intuitively reflect in which direction a certain feature has a significant impact on the target output [32]. The feature importance scores provides an importance measure that can rank the importance of input variables [84]. When calculating the variable
importance, the given variable is randomly permuted in OOB and then obtain an error estimate [66]. The difference between this error estimate and the OOB error without permutation named variable importance [54], [66]. The greater the variable importance, the greater its impact on UCS.

Results of partial dependence plots of the influencing variables for UCS of FCPB are illustrated in Fig. 15. It is clear that the cement-tailings ratio, solid content, curing time and foaming agent dosage are of great importance in responding to the UCS of the prepared FCPB. As expected, increasing the value of the cement-tailings ratio, solid content and curing time used in the FCPB mix gives rise to the higher UCS. An increase in the foaming agent dosage reduces the UCS of FCPBs. These results are consistent with the experimental results in Section V.A. This further verifies the accuracy of the proposed GSO-RF model. Qi et al. [32] studied the effect of solid content on CPB strength and found that 70% was the critical content for strength growth, and after the solids content greater than 70%, the strength accelerated. It is worth noting that similar phenomena also appeared in the strength growth of FCPB, but the critical content value was 74%. This indicates that both CPB and FCPB have the critical solid content to accelerate the strength growth, and the variation of this value may be related to many factors, such as the characteristics of tailings, the type and dosage of admixtures, which need further experimental research. It is worth noting that, by comparing the strength of FCPB in Ref [19], the UCSs obtained in this paper are significantly smaller when the four main factors, chemical properties of tailings and binder type are similar. Comparing the tailings grading used in the respective experiments, it is obvious that the particle size of the tailings (D$_{50} = 23 \mu m$) used in Ref [19] is significantly finer than the particle size of the tailings (D$_{50} = 106 \mu m$) in this paper. In general, under otherwise equal conditions the finer the tailings will lead to higher strength [85]. It can therefore be inferred that the tailings gradation can affect the foaming effect of the foaming
Fig. 16 shows the relative importance score of the cement-tailings ratio, curing time, foaming agent dosage and solid content, respectively. It can be seen that the magnitude of relative importance is in a descending order as: cement-tailings ratio > curing time > foaming agent dosage > solid content. This order (except the foaming agent dosage) is consistent with some findings in the literatures [23], [32], [50]. The importance score of the cement-tailings ratio is 0.5201, which overweighted the sum of the importance scores of the other three factors. Therefore, the cement-tailings ratio is still the primary consideration in the ratio design of FCPB. The importance score of curing time is 0.3256, ranking second only to the curing time. The early strength of backfill is particularly important to improve mining efficiency and production, and the strength does not change significantly after a period of time [28]. Therefore, the UCS at 3 days, 7 days, 14 days and 28 days are generally studied. The importance scores of foaming agent dosage and solid content were 0.1194 and 0.0349, respectively. The importance of foaming agent dosage is self-evident: when the content is insufficient, the expansion effect of FCPB is poor, thus can’t play the purpose of roof connection; when the dosage is excessive, the strength of backfill is not enough to control of surface subsidence. In this study, when the foaming agent dosage increased from 1.5% to 3%, the strength of FCPB decreased by 35 to 55%. The consistency of slurry plays an important role in the formation and distribution of bubbles [86]. Other things being equal, the higher the solid content, the greater the viscous force of slurry, which is not conducive to the formation of bubbles, and finally shows the characteristics of low porosity of FCPB. However, the slurry with low solid content will also have an adverse effect on the formation of FCPB. Too low consistency will lead to the escape of bubbles and the merger of small bubbles to form large bubbles [87]. Therefore, reasonable solid content is beneficial to the formation of bubbles and better pore size distribution.
VI. CONCLUSIONS

In this study, a hybrid artificial intelligence model (GSO-RF) was proposed to predict the UCS of FCPB. A total of 432 FCPB samples with different combinations of influencing variables (the cement-tailings ratio, the solid content, the forming agent dosage and curing time) were performed for the construction of GSO-RF model. The optimum hyper-parameters were obtained by GSO. A 10-fold cross-validation was adopted to optimize the process of hyper-parameters tuning. In addition, the hybrid model performance is evaluated by MSE, MAE, and R^2. At last, partial dependence plots and feature importance scores were introduced to perform variable sensitivity analysis.

In terms of predicting FCPB strength, the GSO-RF model shows great potential. The GSO can effectively tune the hyper-parameters of the proposed GSO-RF model. In both the training set and testing set, the hybrid model showed a good prediction effect, which indicates that the model can effectively and accurately predict the UCS of FCPB. Through influencing variables sensitivity analysis, it can be found that the foaming agent will not change the influence trend of cement-tailings, solid content and curing time, but the influence degree will be weakened by the foaming agent. The ranking of the relative importance of influencing variables is: cement-tailings ratio > curing time > foaming agent dosage > solid content. In addition, there is a critical solid content (74%) value that accelerates the strength growth of FCPB. The Foaming agent has no significant effect on the hydration of cement. The foaming agent mainly changes the strength by changing the pore characteristics (especially the large pore volume).

It is worth noting that the dataset splitting to the training and testing set was only performed only once in the current study. In this case, the performance analysis will suffer from some randomness issue \[88\], \[89\]. Moreover, the input variables in this study are relatively few, and factors such as the physical and chemical characteristics of tailings and cementing material types are not taken into account. Therefore, correcting randomness issue and incorporating more variables into the GSO-RF model will improve the prediction accuracy and applicability of the model.

REFERENCES

[1] M. Fall, M. Benzaazoua, and S. Ouellet, “Experimental characterization of the influence of tailings fineness and density on the quality of cemented paste backfill,” Minerals Eng., vol. 18, no. 1, pp. 41–44, Jan. 2005.
[2] M. Benzaazoua, B. Bussière, I. Demers, M. Aubertin, É. Fried, and A. Blier, “Integrated mine tailings management by combining environmental desulphurization and cemented paste backfill: Application to mine Doyon, Quebec, Canada,” Minerals Eng., vol. 21, no. 4, pp. 330–340, Mar. 2008.
[3] J. Qiu, Z. Guo, L. Yang, H. Jiang, and Y. Zhao, “Effects of packing density and water film thickness on the fluidity behaviour of cemented paste backfill,” Powder Technol., vol. 359, pp. 27–35, Jan. 2020.
[4] A. Kesimul, E. Yilmaz, B. Ercikdi, I. Alp, and H. Deveci, “Effect of properties of tailings and binder on the short-and long-term strength and stability of cemented paste backfill,” Mater. Lett., vol. 59, no. 28, pp. 3703–3709, Dec. 2005.
[5] B. Ercikdi, G. Külekci, and T. Yilmaz, “Utilization of granulated marble wastes and waste bricks as mineral admixture in cemented paste backfill of sulphide-rich tailings,” Construct. Building Mater., vol. 93, pp. 573–583, Sep. 2015.
[6] M. Fall and M. Benzaazoua, “Modeling the effect of sulphate on strength development of paste backfill and binder mixture optimization,” Cement Concrete Res., vol. 35, no. 2, pp. 301–314, Feb. 2005.
[7] Z. Song, J. Mao, X. Tian, Y. Zhang, and J. Wang, “Optimization analysis of controlled blasting for passing through houses at close range in super-large section tunnels,” Shock Vib., vol. 2019, pp. 1–16, Aug. 2019.
[8] X. Zhang, Y. Han, Y. Sun, and Y. Li, “Innovative utilization of refractory iron ore via suspension magnetization roasting: A pilot-scale study,” Powder Technol., vol. 352, pp. 16–24, Jun. 2019.
[9] Z. Song, G. Shi, J. Wang, W. Wei, T. Wang, and G. Zhou, “Research on management and application of tunnel engineering based on BIM technology,” J. Civil Eng. Manage., vol. 25, no. 8, pp. 785–797, Sep. 2019.
[10] X. Zhang, X. Gu, Y. Han, N. Parra-Álvarez, V. Clarenboux, and S. K. Kawatra, “Flotation of iron ores: A review,” Mineral Process. Extractive Metall. Rev., pp. 1–29, Nov. 2019.
[11] C. Qi and A. Fourie, “Cemented paste backfill for mineral tailings management: Review and future perspectives,” Minerals Eng., vol. 144, Dec. 2019, Art. no. 106025.
Q. Huang, J. Mao, and Y. Liu, “An improved grid search algorithm of SVM parameters optimization,” in Proc. IEEE 14th Int. Conf. Commun. Technol., Nov. 2012, pp. 1022–1026.

I. Suyari, A. Prugel-Bennett, and G. Wills, “SVM parameter optimization using grid search and genetic algorithm to improve classification performance,” Telecommun. Comput. Electron. Control, vol. 14, no. 4, p. 1502, 2016.

R. Díaz-Uriarte and S. A. De Andres, “Gene selection and classification of microarray data using random forest,” BMC Bioinf., vol. 7, no. 1, p. 3, 2006.

J. Wei, W. Huang, Z. Li, W. Xue, Y. Peng, L. Sun, and M. Cribb, “Estimating 1-km-resolution PM2.5 concentrations across China using the space-time random forest approach,” Remote Sens. Environ., vol. 231, Sep. 2019, Art. no. 111221.

J. Bialas, T. Oommen, and T. C. Havens, “Optimal segmentation of high spatial resolution images for the classification of buildings using random forests,” Int. J. Appl. Earth Observ. Geoinf., vol. 82, Oct. 2019, Art. no. 101895.

S. Lu, Q. Li, L. Bai, and R. Wang, “Performance predictions of ground source heat pump system based on random forest and back propagation neural network models,” Energy Convers. Manage., vol. 197, Oct. 2019, Art. no. 118664.

C. Beaulac and J. S. Rosenthal, “Predicting university Students’ academic success and major using random forests,” Res. Higher Edu., vol. 60, no. 7, pp. 1048–1064, Nov. 2019.

M. W. Ahmad, J. Reynolds, and Y. Rezgui, “Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees,” J. Cleaner Prod., vol. 203, pp. 810–821, Dec. 2018.

L. Breiman, “Bagging predictors,” Mach. Learn., vol. 24, no. 2, pp. 123–140, Aug. 1996.

L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

F. B. de Santana, W. Borges Neto, and R. J. Poppi, “Random forest as one-class classifier and infrared spectroscopy for food adulteration detection,” Food Chem., vol. 293, pp. 323–332, Sep. 2019.

V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan, and B. P. Feuston, “Random forest: A classification and regression tool for compound classification and QSAR modeling,” J. Chem. Inf. Comput. Sci., vol. 43, no. 6, pp. 1947–1958, Nov. 2003.

H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, “An empirical evaluation of deep architectures on problems with many factors of variation,” in Proc. 24th Int. Conf. Mach. Learn. (ICML), 2007, pp. 473–480.

J. Wang, H. Du, X. Yao, and Z. Hu, “Using classification structure pharmacokinetic relationship (SCPR) method to predict drug bioavailability based on grid-search support vector machine,” Analytica Chim. Acta, vol. 601, no. 2, pp. 156–163, Oct. 2007.

L. Lin, Z. Xiaolong, Z. Kai, and L. Jun, “Bilinear grid search strategy based support vector machines learning method,” Inf., vol. 38, no. 1, 2014.

Y. Ding, X. Song, and Y. Zen, “Forecasting financial condition of Chinese listed companies based on support vector machine,” Expert Syst. Appl., vol. 34, no. 4, pp. 3081–3089, May 2008.

B. L. Yan, GB/T 175-2007 Common Portland Cement. Beijing, China: Standards Press of China, 2007.

G. Sang, Y. Zhu, G. Yang, and H. Zhang, “Preparation and characterization of high porosity cement-based foam material,” Construct. Building Mater., vol. 91, pp. 133–137, Aug. 2015.

J. Haiqiang, M. Fall, and L. Cui, “Yield stress of cemented paste backfill in sub-zero environments: Experimental results,” Minerals Eng., vol. 92, pp. 141–150, Jun. 2016.

S. Khalid, T. Khalil, and S. Nasreen, “A survey of feature selection and feature extraction techniques in machine learning,” in Proc. Sci. Inf. Conf., Aug. 2014, pp. 372–378.

S. D. Bolboaca and L. Jäntsch, “Pearson versus Spearman, Kendall’s tau correlation analysis on structure-activity relationships of biologic active compounds,” Leonardo J Sci., vol. 5, no. 9, pp. 179–200, Jul. 2006.

F. P. Holgado-Tello, S. Chacón-Moscoso, I. Barbero-García, and E. Vila-Abad, “Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables,” Qual. Quantity, vol. 44, no. 1, pp. 153–166, Jan. 2010.

C. Thurumalai, A. Duba, and R. Reddy, “Decision making system using machine learning and pearson for heart attack,” in Proc. Int. Conf. Electron., Commun. Aeron. Technol. (ICECA), vol. 2, Apr. 2017, pp. 206–210.

Y. Lee and J. Lee, “Binary tree optimization using genetic algorithm for multiclass support vector machine,” Expert Syst. Appl., vol. 42, no. 8, pp. 3843–3851, May 2015.

F. Archetti, I. Giordani, and L. Vanneschi, “Genetic programming for QSAR investigation of docking energy,” Appl. Soft Comput., vol. 10, no. 1, pp. 170–182, Jan. 2010.

E. R. Ziegel, The Elements of Statistical Learning. New York, NY, USA: Taylor & Francis, 2003.

X. Bai, F. Zhang, J. Hou, F. Xia, A. Tolba, and E. Elashkar, “Implicit multi-feature learning for dynamic time series prediction of the impact of institutions,” IEEE Access, vol. 5, pp. 16372–16382, 2017.

B. H. Menze, B. M. Kelin, R. Masuch, U. Himmelreich, P. Bachert, W. Petrich, and F. A. Hamprecht, “A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data,” BMC Bioinf., vol. 10, no. 1, p. 213, 2009.

X. Ke, H. Hou, M. Zhou, Y. Wang, and X. Zhou, “Effect of particle gradation on properties of fresh and hardened cemented paste backfill,” Construct. Building Mater., vol. 96, pp. 378–382, Oct. 2015.

W. Liu and X. Zhang, “Study on volume stability, of chemical foaming cement paste,” KSCE J. Civ. Eng., vol. 21, no. 7, pp. 2790–2797, 2017.

T.-H. Wee, D. S. Babu, T. Tamilselvan, and H.-S. Lim, “Air-void system of foamed concrete and its effect on mechanical properties,” ACI Mater. J., vol. 103, no. 1, p. 45, 2006.

C. Qi, H.-B. Ly, Q. Chen, T.-T. Le, V. M. Le, and B. T. Pham, “Clogging-dewatering prediction of fine mineral tailings using a hybrid machine learning approach,” Chemosphere, vol. 244, Apr. 2020, Art. no. 125450.

C. Qi, Q. Chen, X. Dong, Q. Zhang, and Z. M. Yaseen, “Pressure drops of fresh cemented paste backfills through coupled test loop experiments and machine learning techniques,” Powder Technol., vol. 361, pp. 748–758, Feb. 2020.
SHIYU ZHANG received the B.S. and M.S. degrees from Northeastern University (NEU), Shenyang, China, in 2011 and 2013, respectively, where he is currently pursuing the Ph.D. degree in mining engineering. His research interests include comprehensive utilization of resources and cemented foam backfill.

YINGLIANG ZHAO was born in Yantai, China, in 1987. He is currently pursuing the Ph.D. degree in mining engineering with Northeastern University, Shenyang, China. He is also conducting research on comprehensive utilization of mineral resources and low-carbon cementitious materials.

ZHENGYU MA was born in Shuozhou, Shanxi, China. He is currently pursuing the bachelor’s degree in mining engineering with Northeastern University, Shenyang, China. He has a strong interest in the comprehensive utilization of mineral resources and has carried out research in this area.

* * *