

## Prediction of Permeability Using Random Forest and Genetic Algorithm Model

Junhui Wang<sup>1</sup>, Wanzi Yan<sup>1</sup>, Zhijun Wan<sup>1,\*</sup>, Yi Wang<sup>2,\*</sup>, Jiakun Lv<sup>1</sup> and Aiping Zhou<sup>3</sup>

<sup>1</sup>Key Laboratory of Deep Coal Resource Mining (CUMT), Ministry of Education of China, School of Mines, China University of Mining and Technology, Xuzhou, 221116, China

<sup>2</sup>College of Safety and Emergency Management Engineering, Taiyuan University of Technology, Taiyuan 030024, China

<sup>3</sup>China Energy Investment Group Company, Beijing, 100011, China

\*Corresponding Authors: Zhijun Wan. Email: zhjwan@cumt.edu.cn; Yi Wang. Email: wangyi@tyut.edu.cn

Received: 17 September 2020; Accepted: 26 October 2020

**Abstract:** Precise recovery of Coalbed Methane (CBM) based on transparent reconstruction of geological conditions is a branch of intelligent mining. The process of permeability reconstruction, ranging from data perception to real-time data visualization, is applicable to disaster risk warning and intelligent decision-making on gas drainage. In this study, a machine learning method integrating the Random Forest (RF) and the Genetic Algorithm (GA) was established for permeability prediction in the Xishan Coalfield based on Uniaxial Compressive Strength (UCS), effective stress, temperature and gas pressure. A total of 50 sets of data collected by a self-developed apparatus were used to generate datasets for training and validating models. Statistical measures including the coefficient of determination ( $R^2$ ) and Root Mean Square Error (RMSE) were selected to validate and compare the predictive performances of the single RF model and the hybrid RF–GA model. Furthermore, sensitivity studies were conducted to evaluate the importance of input parameters. The results show that, the proposed RF–GA model is robust in predicting the permeability; UCS is directly correlated to permeability, while all other inputs are inversely related to permeability; the effective stress exerts the greatest impact on permeability based on importance score, followed by the temperature (or gas pressure) and UCS. The partial dependence plots, indicative of marginal utility of each feature in permeability prediction, are in line with experimental results. Thus, the proposed hybrid model (RF–GA) is capable of predicting permeability and thus beneficial to precise CBM recovery.

**Keywords:** Permeability; machine learning; random forest; genetic algorithm; coalbed methane recovery

### Abbreviations

CBM Coalbed Methane;  
RF Random Forest;



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

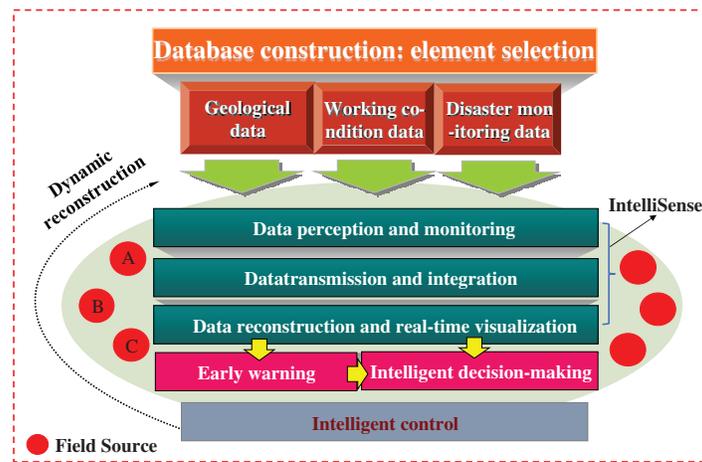
<b>GA</b>	Genetic Algorithm;
<b>UCS</b>	Uniaxial Compressive Strength;
<b>RMSE</b>	Root Mean Square Error;
<b>CNN</b>	Convolutional Neural Networks;
<b>BP</b>	Back Propagation;
<b>APSO</b>	Adaptive Particle Swarm Optimization;
<b>LSSVM</b>	Least Squares Support Vector Machine;
<b>WLS-SVM</b>	Weighted Least Square Support Vector Machine;
<b>MABC</b>	Modified Artificial Bee Colony;
<b>CEA</b>	Cross Entropy Algorithm;
<b>REV</b>	Representative Elementary Volume;
<b>DT</b>	Decision Tree;
<b>OOB</b>	Out-of-Bag;
<b>LWD</b>	Logging-While-Drilling;
<b>IS</b>	Importance Score;
<b>PDP</b>	Partial Dependence Plot.

## 1 Introduction

Coalbed Methane (CBM) is a form of unconventional natural gas found in underground coal mines, where the porous media and fracture networks store and transport CBM [1,2]. CBM is also considered a potential hazard and thus vented out before coal extraction [3]. Coal seam permeability is a measure of gas flowability in the reservoir and governs the production performance and recovery efficiency of CBM [4,5]. Precise CBM recovery based on transparent reconstruction of geological conditions is a branch of intelligent mining, and its flowchart is shown in Fig. 1. The foundation of this process is to construct a database with selected elements including geological data, working condition data and disaster monitoring data [6]. Then, as the existing field sources (e.g., A, B, and C in Fig. 1 denote stress-strain field, temperature field, and seepage field, respectively) would change definitely under any control action, intellisense, early warning (or intelligent decision-making) and intelligent control steps circulate, as a result, the dynamic reconstruction process will make information stereoscopic. From data perception to real-time data visualization, three-dimensional static permeability is reconstructed, which can serve for early warning to hazards; in addition, visualization helps identify favorable areas of CBM recovery based on reservoir permeability assessment [7], followed by gas extraction in the workable zone or in situ modification to enhance gas recovery in the unworkable zone. So the permeability prediction is an essential part of intellisense, serving for decision-making and control action.

The research on multi-field multi-source coupled evolution mechanism and disaster breeding evolution pattern supports theoretically transparent reconstruction of geological conditions for precise CBM extraction [8]. Ye et al. [9] believed that the dynamic change of coal permeability depends on multiple factors. Tao et al. [10] analyzed the patterns of permeability variation in coal seam No. 3 of Qinshui Coalfield and their influences on CBM recovery. To date, researchers have studied experimentally the effect(s) of one or a few variables on permeability. In the initial stage of CBM production, the effective stress level was found to be the controlling factor in permeability reduction [11], and the permeability-strain curve of coal corresponds well to its full stress-strain curve, implying that flowing characteristics of CBM are closely related to damage evolution in the coal during loading [12]. In recent years, the development of triaxial loading devices equipped with high precision sensors and servo control systems has allowed the permeability evolution to be deliberated, stress, as a main controlling factor, and other factors (e.g., moisture [13], temperature [14], geoelectric field [15], type of adsorbed gas [16], etc.) for

simulating the heat rejection, gas displacement, and hydraulic fracturing, etc., have been added to investigate the permeability change tendency [17–21]. With advances in scientific knowledge and field observations, remarkable progresses have been made in analytical permeability models to predict the unique permeability behavior of CBM reservoirs [22–25]. The permeability variation models can be classified into two categories: Porosity-based models where permeability is cubically related to porosity, and stress–strain based models where permeability is exponentially related to the variation in effective stress [26]. Admittedly, empirical formulas for permeability evolution determined by test data fitting are often used in engineering practice, such as the Langmuir-type curve [27].



**Figure 1:** Flowchart of transparent reconstruction of geological conditions

Permeability of gas-saturated coal are affected by numerous factors, making gas seepage process time-variant, nonlinear, and fuzzy, technical bottlenecks of transparent reconstruction of geological conditions are complex procedure, difficulty in nonlinear relationship representation, focusing on univariate analysis and prediction while weakening multivariate impact, and limited model scope, etc., so dynamic geological information transparency fails to satisfy the demand of modern manufacturing. Relying on powerful data processing capability, machine learning simulates spontaneously human learning activity and finds latent rules of the data through analysis so as to enable analytical research and spontaneous decision making on new samples. Recently, Tian et al. [28] used CNN modelling to investigate the inherent relationships between permeability and the microstructural parameters of porous media. Yin et al. [29] and Xie et al. [30] employed BP neural network model to predict the permeability of gas-saturated coal, but this neural network gets easily into local minimum and converges slowly. Some scholars modeled and predicted small-sample data of permeability of gas-saturated coal by using PSO-LSSVM [31], APSO-WLS-SVM [32], MABC-SVM [33], CEA-SVM [34], etc., kernel function and parameters have to be selected in SVM, while in LSSVM, advantages of standard SVM in robustness and sparsity are lost. Habibi et al. [35] predicted the permeability in dual fracture media by multivariate regression analysis. Sharma et al. [36] evaluated the permeability of Indian coal using adaptive neuro-fuzzy inference system technique and compared the outcomes with those obtained by the traditional statistical method of multiple regression analysis. Xie et al. [37] constructed a permeability prediction model using rough set theory based on conditional attributes of the reservoir in Qinshui Basin including reservoir pressure, in situ stress, thickness, and depth, but the prediction result

is not a specific value but a range. As can be seen, machine learning excels in prediction, but selection of input variables limits applicability of the algorithm, in addition, another important challenge is about improving the accuracy and reliability of predictions.

The main objective of this study is to develop a novel hybrid model RF–GA using goodness of individual models for better permeability prediction of gas-saturated raw coal at Xishan Coalfield which is one of the largest CBM reservoirs in China, based on Uniaxial Compressive Strength (UCS), effective stress, temperature, and gas pressure. The study is organized as follows: firstly, the dataset was prepared by a self-developed apparatus to conduct series of seepage tests and UCS tests; then, R2 and RMSE were used to estimate the prediction performance of the RF and RF–GA models; thirdly, correlations and sensitivity analysis were carried out to analyze the change tendency of permeability with input variables.

## 2 Experimental Program and Data Collection

### 2.1 Materials and Apparatus

The coal samples were collected from coal seam No. 8 at Xishan Coalfield, Shanxi Province, China (Fig. 2a). The coal seam is characterized by high gas content and low permeability. In order to obtain a dataset with multidimensional experimental values, raw coal samples in standard size ( $\varphi 50 \times 100$  mm) were fabricated using a diamond wire cutting machine. Forced drying of the samples was avoided and the average moisture content  $w$  was set to 1.86%.

The dataset was collected through a self-developed triaxial testing apparatus as shown in Fig. 2b. Detailed design of the experimental apparatus can be found in the literature [38]. The testing apparatus enables UCS tests and seepage tests by changing the force head, and mainly consists of a loading frame, a servo hydraulic station, an air path system, a triaxial chamber, a constant-temperature oil bath, and a data acquisition system. The specimen was placed in the chamber, and the paths, rates and target values of loading and heating can be adjusted during the process or set in advance. Seepage test was followed by mechanical test using the same batch of coal samples, yielding values of permeability and UCS at variable temperatures, gas pressures, and stresses.

A dual-porosity seepage model is illustrated in Fig. 2c. The model consists of matrix blocks known as Representative Elementary Volumes (REVs) and fracture network. The REV has pore clusters supported by skeletons, and can be represented by a capillary bundle seepage model. Under the action of loading and heating, the void space or the tortuosity path will change obviously. External loading affects matrix skeleton deformation or tube rearrangement, while heating ( $<100^\circ\text{C}$ ) enables thermal expansion and sorption-induced shrinkage of the blocks. In addition, gas pressure will also induce elastic deformation of skeletons. Thus, volumetric strain of a skeleton in the REV is  $\varepsilon_m = \varepsilon_{es} + \varepsilon_T + \varepsilon_p + \varepsilon_s$  (Fig. 2c) in which,

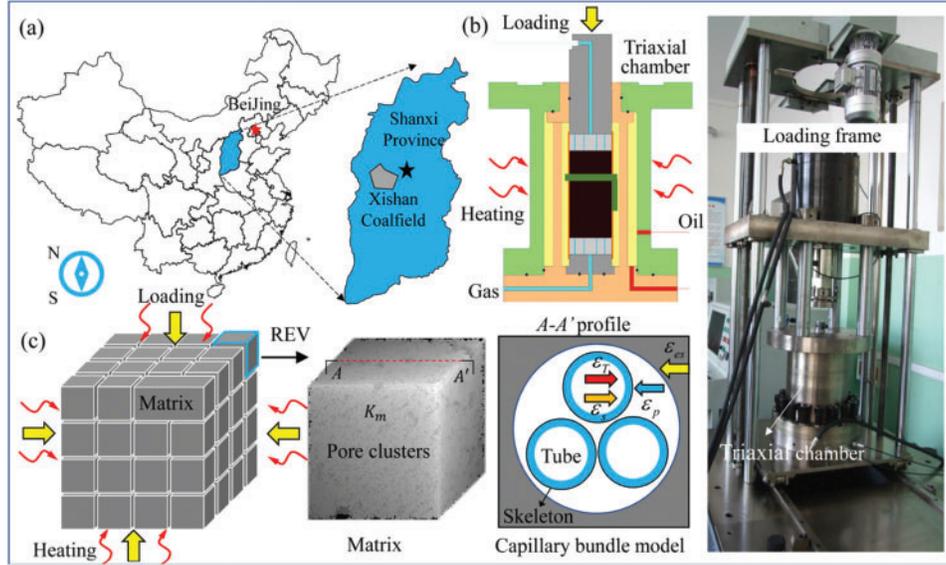
Strain  $\varepsilon_{es}$  due to effective stress can be defined as

$$\varepsilon_{es} = -\frac{1}{K_m} \Delta\sigma_m \quad (1)$$

where  $\Delta\sigma_m$  is the increment of effective stress (MPa) and  $K_m$  is the stiffness of the matrix. The effective stress can be calculated by Eq. (2), without considering the contributions of adsorbed gas and temperature to effective stress.

$$\sigma_m = \frac{1}{3} (\sigma_1 + 2\sigma_2) - \frac{1}{2} (P_1 + P_2) \quad (2)$$

where  $\sigma_1$  and  $\sigma_2$  represent the axial stress and confining pressure (MPa), respectively;  $P_1$  and  $P_2$  denote the inlet and outlet gas pressure, MPa.



**Figure 2:** Materials and apparatus: (a) Sampling site; (b) Testing apparatus; (c) Dual-porosity seepage model

Strain  $\varepsilon_T$  due to thermal expansion can be described as [39]:

$$\varepsilon_T = \alpha_T \Delta T \tag{3}$$

where  $\alpha_T$  is the coefficient of volumetric thermal expansion ( $K^{-1}$ ) and  $\Delta T$  is the increment of temperature (K).

Tube compression by gas pressure yields strain  $\varepsilon_p$ :

$$\varepsilon_p = -C_Y \Delta P \tag{4}$$

where  $C_Y$  is the coefficient of volumetric compression,  $MPa^{-1}$ ;

Strain  $\varepsilon_s$  due to sorption-induced shrinkage can be expressed by the following equation where a temperature correction term is introduced based on the Langmuir formula [40]:

$$\varepsilon_s = \left[ \frac{\varepsilon_L P_m}{P_L + P_m} \exp\left(-\frac{c_2 \Delta T}{1 + c_1 P_m}\right) - \frac{\varepsilon_L P_{m0}}{P_L + P_{m0}} \right] / (1 - \varphi_0) \tag{5}$$

where  $\varepsilon_L$  is the coefficient of gas sorption-induced expansion;  $P_m$  is the current gas pressure, MPa;  $P_L$  is the Langmuir pressure, MPa;  $c_2$  is the coefficient of temperature correction,  $K^{-1}$ ;  $c_1$  is the coefficient of pressure correction,  $MPa^{-1}$ ;  $P_{m0}$  is the initial gas pressure, MPa; and  $\varphi_0$  is the initial absolute porosity, %.

On the assumption that the moisture in the tubes occupies the pores uniformly, porosity  $\varphi_{mw}$  of aqueous matrix is [41]

$$\varphi_{mw} = (1 - w) \left[ 1 - \frac{1 - \varphi_0}{1 + e} (1 + \varepsilon_m) \right] \quad (6)$$

where  $e$  is the measured volumetric strain of REV and  $w$  is the moisture content of pores.

From the definition of the coefficient of volumetric compression, we have [42]

$$1 + e = \exp(-C_Y \Delta \sigma_m) \quad (7)$$

By integrating Eqs. (1)–(7), a porosity evolution model in consideration of temperature, gas pressure, and effective stress can be obtained:

$$\varphi_{mw} = (1 - w) \left[ 1 - \frac{1 - \varphi_0}{\exp(-C_Y \Delta \sigma_m)} (1 + \varepsilon_{es} + \varepsilon_T + \varepsilon_S + \varepsilon_p) \right] \quad (8)$$

The relationship between permeability and porosity is given in the Kozeny–Carman equation established on the basis of the capillary tubes model [43]. By integrating Eqs. (1)–(5) and (8), coal permeability  $k_m$  can be obtained:

$$k_m = \frac{k_{m0} (1 - w)}{\exp(-C_Y \Delta \sigma_m)} \times \left\{ 1 + \frac{e - \left( \alpha_T \Delta T - \frac{1}{K_m} \Delta \sigma_m - C_Y \Delta P \right) (1 - \varphi_0) - \left[ \frac{\varepsilon_L P_m}{P_L + P_m} \exp\left(-\frac{c_2 \Delta T}{1 + c_1 P_m}\right) - \frac{\varepsilon_L P_{m0}}{P_L + P_{m0}} \right]}{\varphi_0} \right\}^3 \quad (9)$$

## 2.2 Dataset Used

### 2.2.1 Input Variables

In dynamic prediction of permeability  $k_m$ , attention should be paid to initial reservoir conditions, seepage path deformation, and sorption state change, as shown in Fig. 3.

Initial reservoir permeability  $k_{m0}$  varies significantly with geological conditions, such as temperature field, moisture content, pore-fracture space, gas pressure, coal metamorphism (or coal rank), etc. Fracturing and heat injection around the surface wells or boreholes always encounter different temperatures and moisture contents. In this paper, the samples are deemed consistent in moisture content, and temperature range is 20–70°C. Coal metamorphism (or coal rank) governs gas content in the coal seam. Prior to permeability test, the samples were saturated. Pore-fracture space of specimen determines the gas flow path, the higher the UCS of a specimen, the less likely the pore-fracture path gets distorted.

The actual change in permeability of coal reservoirs is also dominated by the seepage path deformation. The effective stress, which is an equivalent value in consideration of both external and internal stresses in porous coal, is able to cause unrecoverable structural deformation. In this study, the unidirectional loading path was set, and the loading rate was kept constant (0.1 MPa/s) until the setpoints were reached. The variations in temperature and gas pressure will lead to force rebalancing around the tubes because of the thermal expansion, sorption-induced shrinkage and gas pressure compression (Fig. 2c), and the bulk deformation is recoverable.

Apart from in situ modification of the seepage path, the other method of improving permeability is to apply forcibly physical fields (microwave irradiation, geoelectric field, sound field, etc.) to coal, and its operating principle is normally enhancing thermal effect or reducing sorption potential to promote the transformation from adsorbed gas to free gas.

The objective of permeability prediction in this study is to serve the industrial field, database construction requires selected input variables that (1) have high grey relation to permeability (the output), 2) have strong timeliness (e.g.,  $\mu$  CT imaging is technically difficult and not timely), and (3) are easily monitored. Therefore, UCS, effective stress, temperature, and gas pressure in Fig. 3 were selected as input variables.

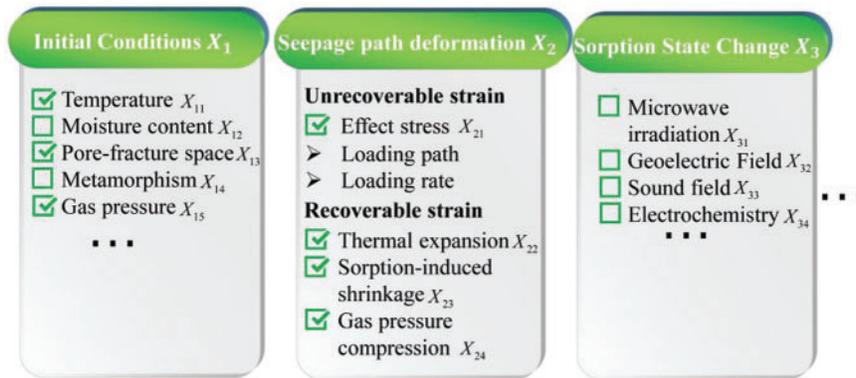


Figure 3: Input variables in the dataset used

### 2.2.2 Output

The output of this study is the permeability of gas-saturated raw coal under different experimental conditions. Regardless of the starting pressure gradient, the permeability of a coal sample is calculated according to the Darcy's law by Eq. (10) [44]:

$$k = \frac{2QP_a\mu L}{A(P_1^2 - P_2^2)} \quad (10)$$

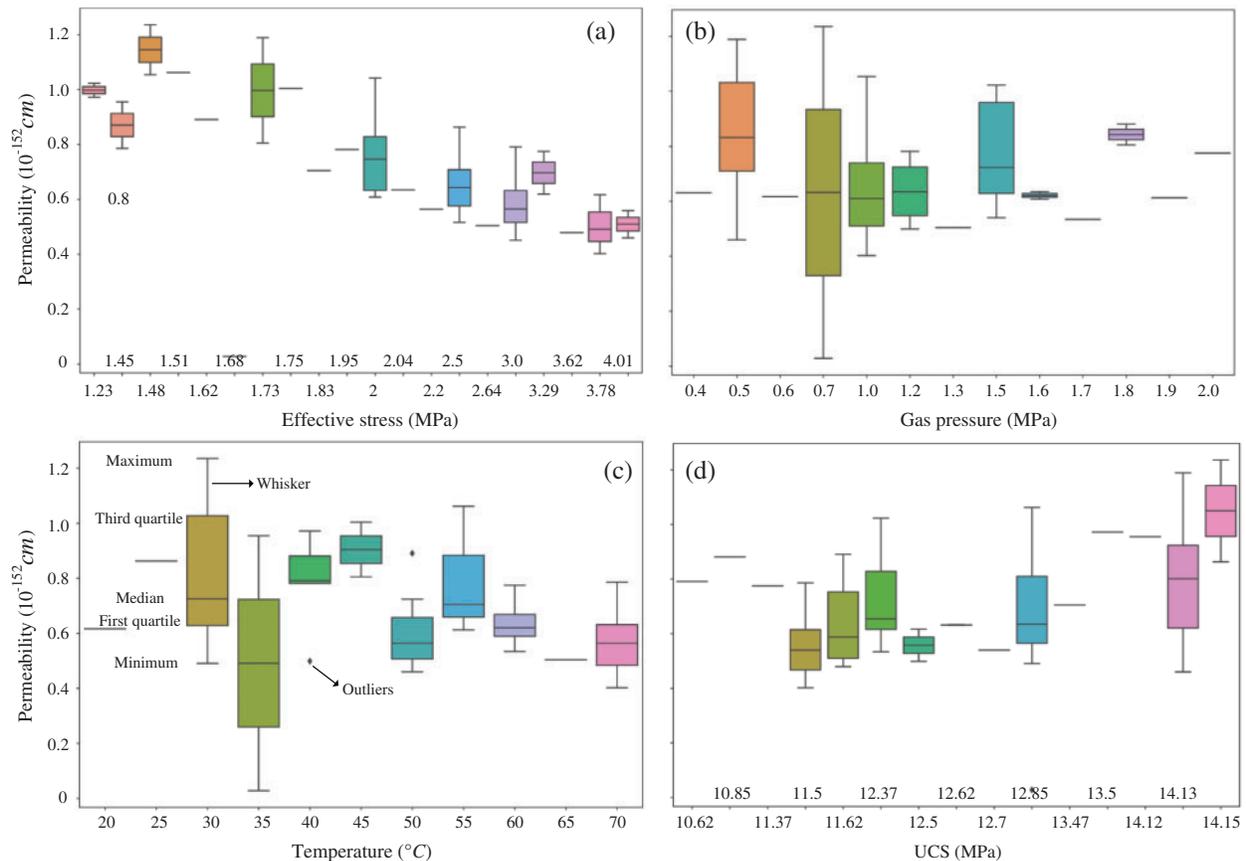
where  $k$  is permeability of the coal sample,  $10^{-3} \mu\text{m}^2$ ;  $Q$  is the gas flow rate at the outlet,  $\text{cm}^3/\text{s}$ ;  $P_a$  is the atmospheric pressure, 0.1 MPa;  $\mu$  is the dynamic viscosity coefficient of gas,  $\text{Pa} \cdot \text{s}$ ,  $\mu = 1.36 \times 10^{-4} T^{0.77}$ ;  $L$  is the length of deformed coal sample, cm; and  $A$  is the area of deformed coal sample,  $\text{cm}^2$ ;

$$L = L' - l_1 \quad (11)$$

$$A = \frac{1}{4}\pi \left( d + \frac{l_2}{\pi} \right)^2 \quad (12)$$

where  $L'$  and  $d$  are the initial height and diameter of the coal sample, mm;  $l_1$  is the axial deformation of the coal sample after the force loading, mm;  $l_2$  is the radial deformation of the coal sample after the force loading, mm.  $l_1$  and  $l_2$  are measured by the axial displacement sensor and circumferential extensometer.

In this study, 50 sets of permeability data of gas-saturated coals varying in compressive strength, gas pressure, effective stress, and temperature were acquired. For such 50 sets of permeability data, Min value was  $0.028 \times 10^{-15} \text{ m}^2$ , Max value was  $1.236 \times 10^{-15} \text{ m}^2$ , Mean value was  $0.705 \times 10^{-15} \text{ m}^2$ , and Standard deviation was 0.225. Boxplots are excellent representation of distribution frequencies and features of the permeability data within the range from the minimum to the maximum in the cases of four input variables (Fig. 4), a short interquartile range, i.e., the distance between the first quartile and the third quartile, indicates many values distributed within a very small range, and a long whisker indicates that the data have very high standard deviation and variance [45]. There are very few outliers in Fig. 4, indicating that the test data are reliable for permeability prediction. Longer interquartile ranges are observed in Figs. 4a–4d, indicating that, under univariate condition, three other variables have significant impact on the output, too, thus conventional analytical methods (e.g., permeability models, or empirical formulas) have limitations in presence of multiple variables having complex coupling relationships, whereas machine learning is advantageous in handling highly nonlinear problems and multivariate analysis.



**Figure 4:** Boxplots of experimental results

### 3 Machine Learning Modelling

#### 3.1 Fundamentals of RF and GA

##### 3.1.1 Random Forest (RF)

Random Forest (RF) is a powerful integrated learning algorithm that was first introduced by Breiman for solving regression, unsupervised learning, and classification problems [46]. RF consists of a committee of Decision Trees (DT), and each individual tree is a fairly simple model that has root nodes, split nodes and leaves. The randomness of node selection is the core of RF.

As shown in Fig. 5 random forest is built by DT evaluation using Bootstrap/Bagging algorithm, and the flow of the RF model is described as follows:

- Step 1: Bootstrap sampling. From original training dataset ( $N$  samples,  $M$ -dimensional features),  $k$  new independent subsets are randomly selected with replacement by Bootstrap.
- Step 2: Model training. For subsets,  $k$  decision trees are constructed. For one DT,  $m$  ( $m < M$ ) features are randomly selected from  $M$  features at each node of the tree and split in the principle of node impurity minimization. The samples not selected in each sampling constitute  $k$  Out-of-Bag (OOB) data for model estimation.
- Step 3: Model predicting. These learners end up making predictions.
- Step 4: Result aggregating. For example, in the case of classification, this can be a majority voting; and in the case of regression, this can be averaging upon the predicted values.

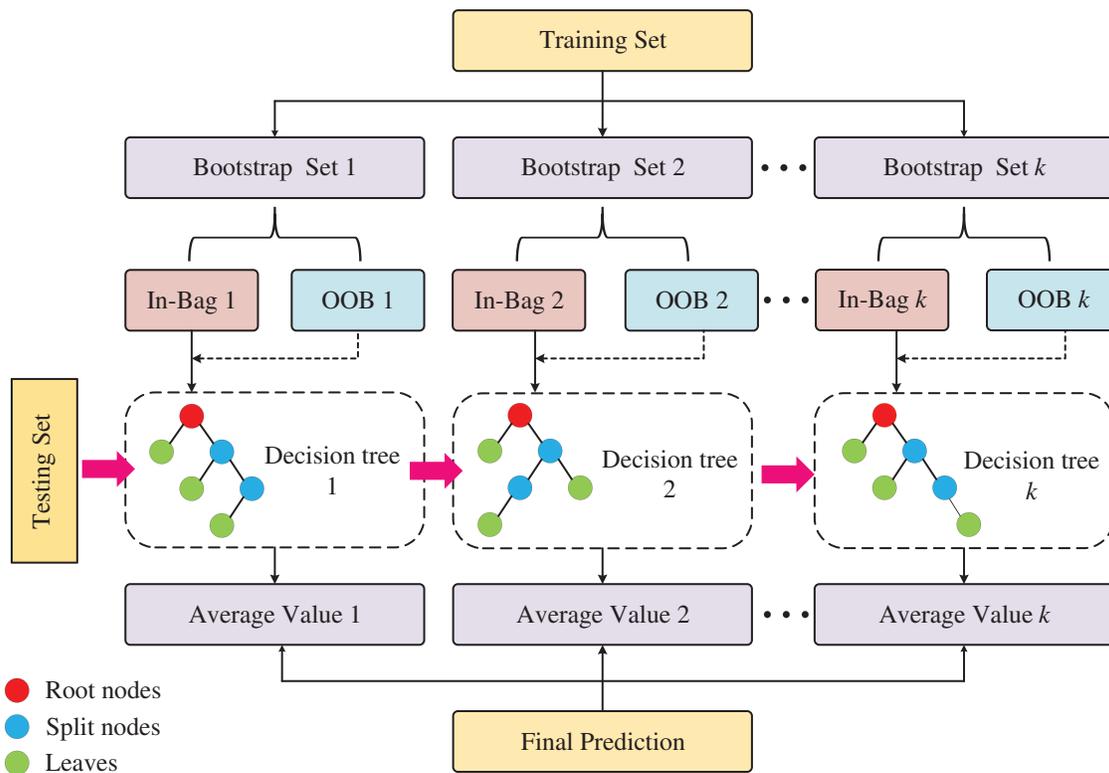


Figure 5: A typical RF process

### 3.1.2 Genetic Algorithm (GA)

Genetic Algorithm (GA), proposed by John Holland in 1970s, is a method of searching for the optimal solution by simulating natural evolutionary process [47], and is used to tune the architecture of the RF for expert performance. It is mainly characterized by direct manipulation of the structural object without limitations in taking the derivative and function continuity, intrinsic implicit parallelism and better global optimization capability, and the use of probabilistic optimization method to obtain automatically and guide the optimal search space without established rules and to tune adaptively the search direction. Since its proposal, GA has matured as one of the most popular evolutionary algorithms for optimization.

GA simulates reproduction, crossover, and genetic mutation phenomena during natural selection and natural inheritance, and can better handle problems with local optima [48]. Core elements of this algorithm include parameter encoding, initial population setting, individual fitness assessment, and GA operators and controlling parameters; a typical GA process is shown in Fig. 6. Firstly, phenotype-to-genotype mapping has to be done, which is usually simplified as binary codes to initialize a population. Then according to the principle of survival of the fittest, the population evolves generation by generation to yield better and better approximate solutions. In each generation, the fitness value of each chromosome is determined by the fitness function. The greater the fitness value, the higher the probability that the chromosome is selected into the next population. With genetic operators of natural genetics, the genes undergo crossover and mutation to generate a population representative of new solution set. Like natural evolution, this process will make the next generation population more adapted to environment than its previous generation, and the optimal individual among the last generation population can be decoded as the approximate optimal solution to the problem.

### 3.2 Modeling and Hyperparameters Tuning

A dataset consisting of 50 samples were split into training set and testing set. The RF model was trained on the training set, while its generalization ability was tested on the testing set. As RF architecture influences its prediction performance, GA was used to optimize the architecture of the RF model. Optimization parameters, and their definitions and valuing ranges are detailed in Tab. 1.

In the current study, a maximum of 100 generations was allowed. In each generation, a total of 10 chromosomes were built. The tournament selection was used for chromosome selection. Crossover probability was set to 0.60, and mutation probability was set to 0.02. All parameter settings in GA were determined by testing.

In the process of hyperparameters tuning, the training performance from 5-fold CV was selected as the fitness function of GA. Every set of hyperparameters were expressed by one chromosome in GA. Gene sequences were updated by chromosomal evolution so as to maximize fitness value. After optimization of GA, the optimal hyperparameters were selected. Finally, the RF model with the optimum hyperparameters was experimentally validated.

### 3.3 RF-GA Model Assessment

In this study, the RF model was assessed as per the criteria of coefficient of determination (R<sup>2</sup>) and Root Mean Square Error (RMSE), and both parameters are most common in validation and comparison of machine learning models [49]. The closer to 1 the value of R<sup>2</sup>, the better the prediction accuracy. RMSE represents the sample standard deviation between the predicted value and the observed value. The smaller the RMSE value, the better the model performance. These

parameters are expressed using the following formulas, respectively:

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} = 1 - \frac{\sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)})^2}{\sum_{i=1}^m (\bar{y} - y^{(i)})^2} \tag{13}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y^{(i)} - \hat{y}^{(i)})^2} \tag{14}$$

where  $SS_{residual}$  denotes the residual sum of squares of the differences between the predicted value and the actual value of sample  $i$ ,  $SS_{total}$  denotes the total sum of squares of the differences between the mean value and the actual value of the sample,  $m$  denotes the number of samples,  $\hat{y}^{(i)}$  denotes the predicted value of sample  $i$ ,  $y^{(i)}$  denotes the actual value of sample  $i$ , and  $\bar{y}$  denotes the sample mean.

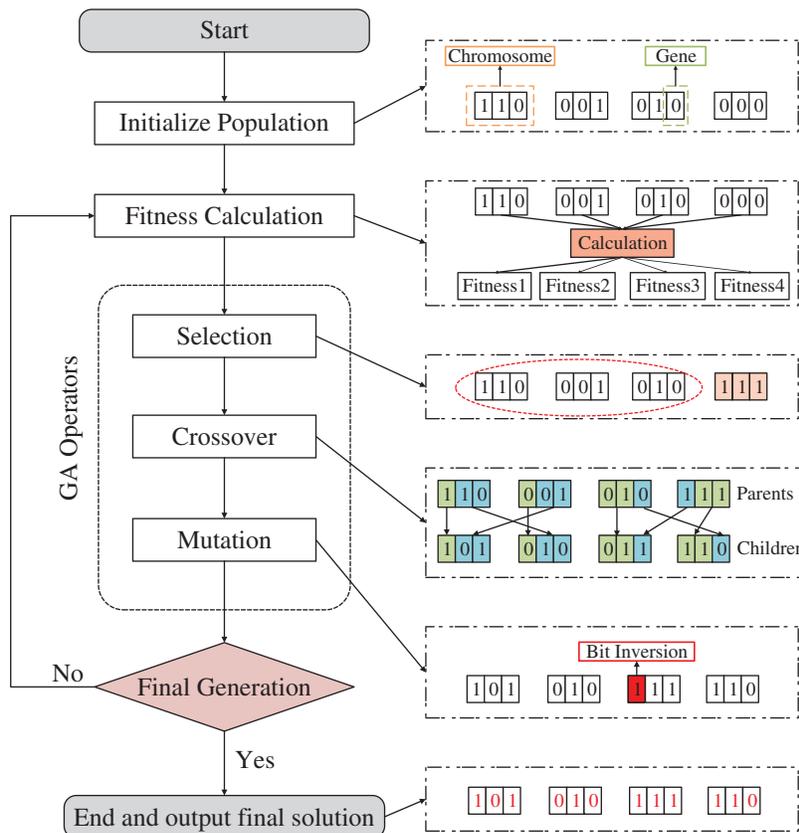
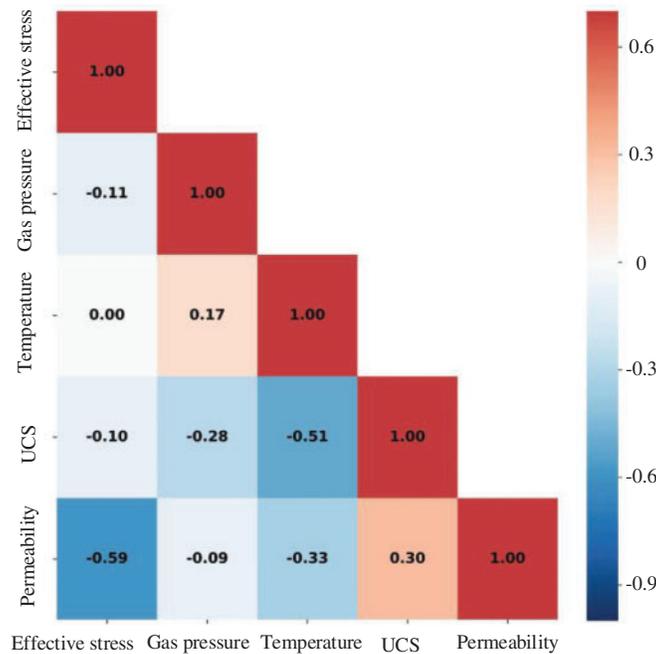


Figure 6: A typical GA process

**Table 1:** Description of hyperparameters and their tuning ranges

No.	Hyperparameter	Explanation	Range
1	Max_depth	The maximum depth of DTs	1–20
2	Min_samples_split	The minimum number of samples for the split	2–10
3	Min_samples_leaf	The minimum number of samples at the leaf node	1–10
4	Max_DT	The maximum number of DT models in the ensemble	1–200
5	Max_features	The number of features considered during the selection of the best splitting	0.4–1

**Figure 7:** Correlation map of selected variables

## 4 Results and Discussion

### 4.1 Input-to-Output Correlations

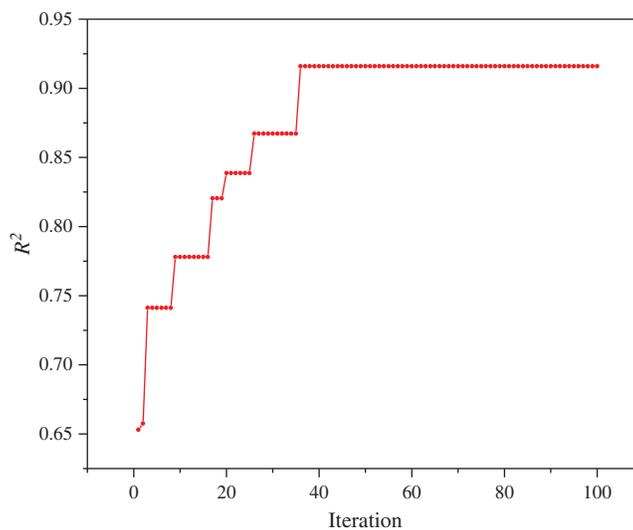
Based on the training and testing of the RF–GA, the relationships between four input variables and the output have been investigated using the Pearson correlation coefficient. Fig. 7 shows the distribution map of correlation coefficients between variables.

UCS is directly correlated to permeability, whereas, effective stress, gas pressure, and temperature are all inversely related to permeability. The absolute value of inverse correlation coefficient between effective stress and permeability is the maximum, indicating that the effect of structural

deformation on permeability is greater than that of the bulk deformation of tube bundle, and as effective stress increases, coal pores and fractures are compressed to become smaller so that gas flow path shrinks; the effect of temperature on coal permeability is embodied in change of gas state equation, change of gas viscosity, thermal expansion of coal matrix, and gas sorption-desorption re-equilibrium; gas pressure is weakly inversely correlated to permeability, because under low osmotic pressure in this study, the gas was prone to “slip flow effect” [50] and the inflection point was postponed, then the permeability–gas pressure curve tended to be flat with increasing gas pressure.

#### 4.2 Hyperparameters Tuning

Fig. 8 shows the iteration process of GA finding the maximum  $R^2$  value. It can be seen that the value of  $R^2$  increases gradually as GA iterates, indicating that the GA is efficient in tuning RF architecture. At the first iteration, the highest  $R^2$  was 0.65, while at the 36th iteration,  $R^2$  increased to 0.916, when optimal hyperparameters of RF were  $\text{max\_depth} = 13$ ,  $\text{min\_sample\_split} = 3$ ,  $\text{min\_samples\_leaf} = 1$ ,  $\text{Max\_DT} = 20$ , and  $\text{max\_features} = 0.821$ .



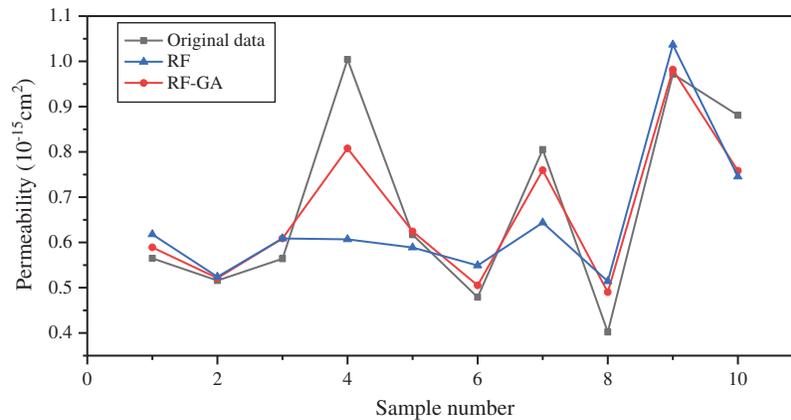
**Figure 8:** Hyperparameters tuning using GA model

#### 4.3 Predictive Capability of the Models

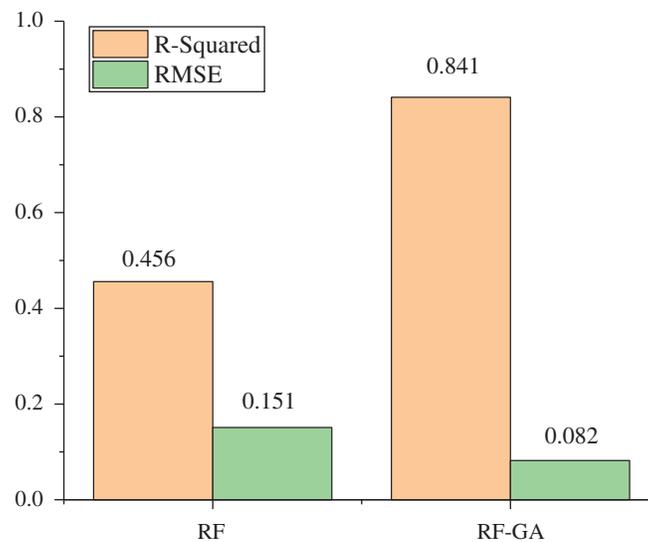
Fig. 9 compares visually prediction performance of typical RF model with that of RF–GA model on the testing set. The sample data consist of 50 points, 80% of which served as the training set while 20% of which were used as the testing set. In prediction with RF model,  $R^2$  value was 0.456, and RMSE value was 0.151. After hyperparameters of the RF model were optimized by GA,  $R^2$  value was 0.841, and RMSE value was 0.082. Fig. 10 compares errors of RF model with those of RF–GA model, demonstrating the feasibility of GA in improvement of RF modeling performance.

As shown by prediction performance, both the hybrid RF–GA model and the single RF model are capable of permeability prediction, though the prediction effect of RF–GA model is superior to that of RF model. GA is an intelligent optimization algorithm that can be efficiently

implemented at ease, thus it is able to solve complex optimization problems effectively. Therefore, GA can be effectively used to improve performance of RF model. In general, the RF-GA model proposed in this paper are able to predict permeability quickly and efficiently, though performance of this model may have to be modified somewhat depending on data source and feature distribution.



**Figure 9:** Comparison between RF model and RF-GA model in permeability prediction performance

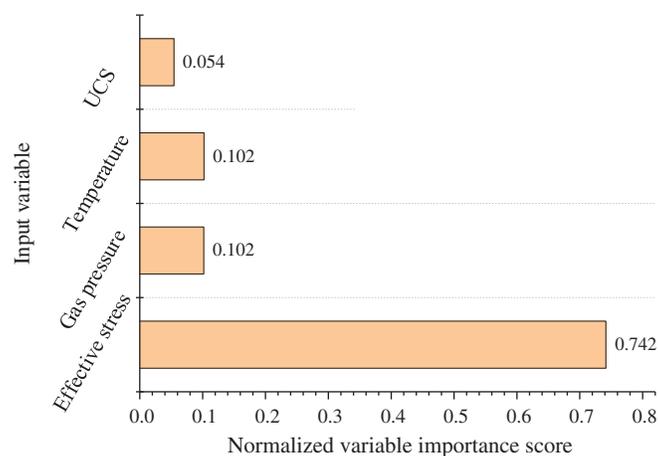


**Figure 10:** Comparison between RF model and RF-GA model in error

#### 4.4 Relative Importance of Input Variables

A sensitivity study of input variables was undertaken for a better understanding of permeability change tendency. Normalized variable Importance Score (IS) and Partial Dependence Plot (PDP) were chosen as methods for interpreting the importance of input variables [51,52].

The RF algorithm was employed to calculate IS of data on the basis of depth of the tree for a variable and mean decrease impurity of nodes, as shown in Fig. 11. The results manifest that the effective stress (IS 0.742) exerts the maximum impact on permeability prediction, followed by temperature or gas pressure (IS 0.102), and then UCS (IS 0.054). The effective stress level was found to be the controlling factor for permeability reduction either in the experiment [53–55] or in CBM production [11]. In Fig. 3a, the interquartile ranges in the effective stress–permeability boxplots were generally short, implying that the data are relatively centered, while the longer interquartile ranges in Figs. 3b–3c prove data discretion due to effective stress. The score of UCS is the lowest because all the samples came from coal seam No. 8 and evidently fractured samples were rejected in advance, leading to narrow range of UCS (10.62–14.15 MPa) which does not impact permeability much.

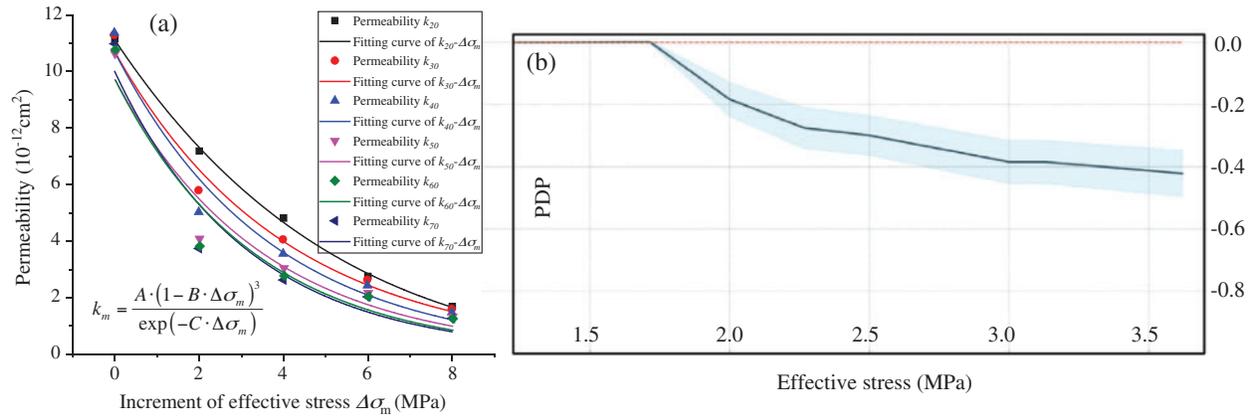


**Figure 11:** Importance scores of input variables

PDPs were introduced by Friedman for understanding the marginal effect of one feature on prediction performance of a machine learning model [56–58], in other words, they are used to study how features influence the prediction. PDPs are obtained by choosing a number of values of the influencing variables, predicting the output using each of those values for all cases of other influencing variables, and then calculating the mean output across the cases.

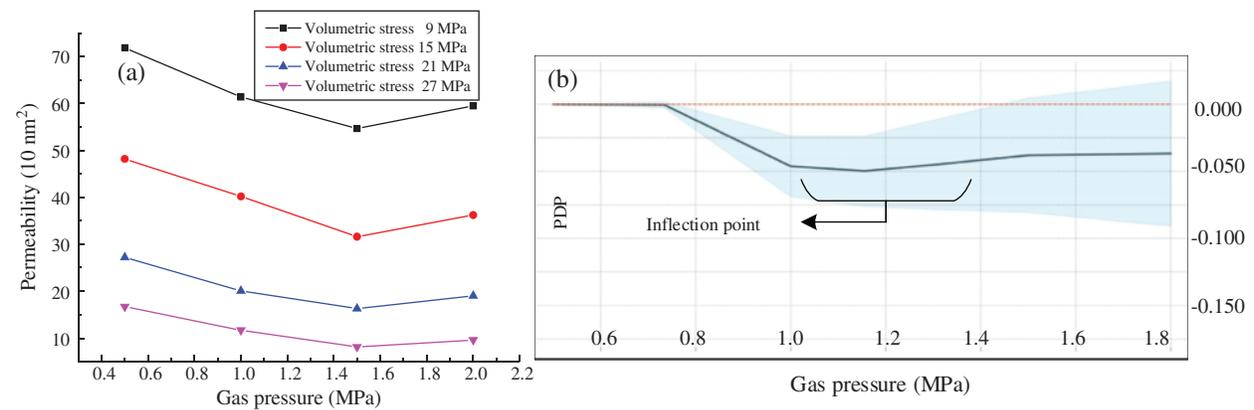
Fig. 12a shows change tendencies of permeability with the increment of effective stress, where  $k_{20}$  is the permeability at 20°C. Fig. 12b is the PDP of the effective stress on permeability prediction, showing that the permeability of coal samples decreases nonlinearly at a reduced rate with the increase of effective stress, which concurs with the tendencies in Fig. 12a. The relationship between permeability and the increment of effective stress can be expressed by the function  $k_m = A \cdot (1 - B \cdot \Delta\sigma_m)^3 / \exp(-C \cdot \Delta\sigma_m)$  which agrees with Eq. (9), where A, B, and C are fitting coefficients.

The mechanism of effective stress influencing permeability is described as follows: (1) As effective stress increases, the coal is further compacted, the pores and fractures wherein are compressed, resulting in narrower gas seepage paths and lower permeability; and (2) After gas seepage paths are compressed to some extent, the compressive effect of effective stress on the seepage paths dwindles, and then the size of seepage paths will tend to be stable.



**Figure 12:** Change in permeability with effective stress: (a) Curve of permeability–increment of mean effective stress; (b) PDP of the effective stress on permeability prediction (Note: The data in Fig. 12a are derived from the samples from coal seam No. 8, panel 18506, Malan Coal Mine, Xishan Coalfield)

Fig. 13a shows change tendencies of permeability with gas pressure. Fig. 13b is a PDP of effective stress on permeability prediction, showing that the change of permeability with gas pressure is not simple monotonic change, but there is an inflection point halfway at gas pressures of 1.0–1.4 MPa, basically in agreement with the test data in Fig. 13a where there are too few data points to determine the range of inflection point very well.

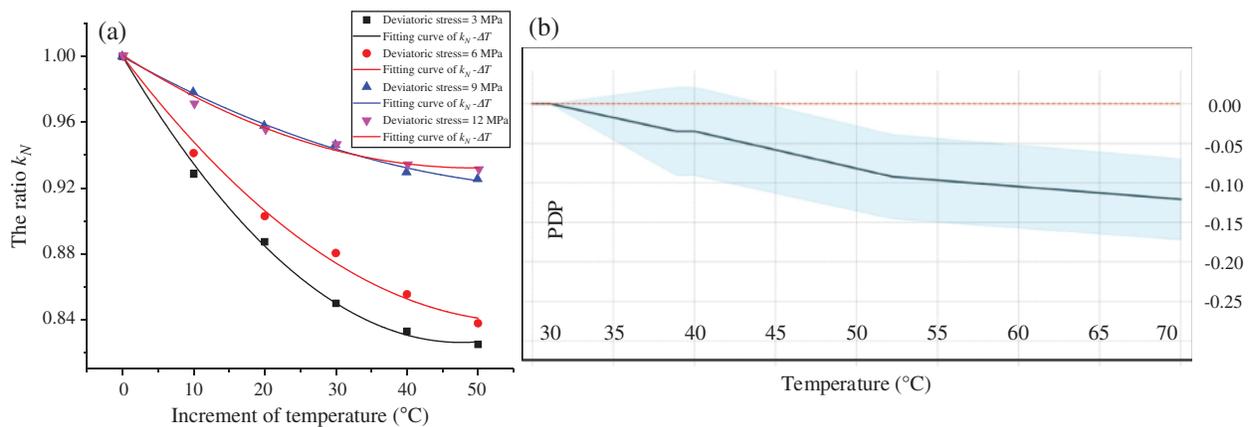


**Figure 13:** Change in permeability with gas pressure: (a) Curves of permeability–gas pressure; (b) PDP of gas pressure on permeability prediction (Note: The data in Fig. 13a are derived from the samples from coal seam No. 8, panel 18506, Malan Coal Mine, Xishan Coalfield)

The reason for occurrence of the above phenomenon in the change of gas pressure versus permeability is Klinbenberg effect that slip flow occurs between gas molecules and solid walls at a low osmotic pressure [50,59,60]. Gas molecules collide each other and with pore-walls when traveling through the pore medium. When the pore radius approaches to the mean free path of gas molecules, the frequency of collision between gas molecules and solid walls increases. Therefore, this additional flux due to the gas flow at the wall surface, which is called slip

flow, becomes effective to enhance the flow rate. When gas pressure further increases, the slip flow phenomenon diminishes gradually while the sorption-induced expansion of skeletons makes seepage paths shrink, then the adsorbed gas will occupy the effective pore path area, resulting in reduced effective pore path section and lower coal permeability. The permeability decreasing tendency will become gradually gentle and there will be an inflection point due to limited space of sorption-induced expansion. Furthermore, the postponement of inflection point enabled very small coefficient of inverse correlation between gas pressure and permeability in Fig. 7.

Fig. 14a shows change tendencies of permeability with temperature. Fig. 14b is a PDP of temperature on permeability prediction, showing that the permeability decreased nonlinearly with increasing temperature. The mechanism of temperature influencing permeability is described as follows: (1) Thermal expansion outperforms sorption-induced shrinkage in leading to seepage path compression; and (2) As temperature rises, gas viscosity increases accordingly, and gas flow in the seepage path becomes slower, resulting in reduction in coal permeability.



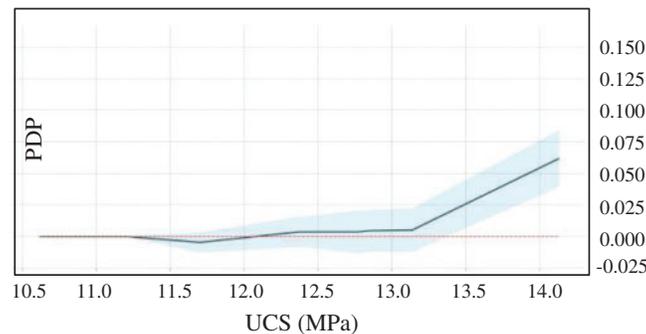
**Figure 14:** Change in permeability with temperature: (a) curves of  $k_N$ —increment of temperature, where  $k_N$  is the ratio of current permeability to initial permeability and the initial temperature is 20°C; (b) PDP of temperature on permeability prediction (Note: The data in Fig. 14a are derived from the samples from coal seam No. 8, panel 18506, Malan Coal Mine, Xishan Coalfield)

Fig. 15 is a PDP of UCS on permeability prediction, showing that compressive strength was weakly directly correlated to permeability. Because coals varying in coal rank and coal quality characteristics differ much in compressive strength, in other words, under equal pressures, they will produce deformation and fractures to variable extents, the higher the compressive strength of coal, the better the primary pores and fractures are protected, and the higher the permeability [61].

### 5 Limitations and Outlooks

It has shown that the permeability prediction using RF-GA model is quite promising, but challenges still remain. First, the dataset was collected from Xishan Coalfield in China, resulting in a high possibility that the trained model cannot be generalized to other coalfields. This challenge is mainly due to small dimensionality and volume of the dataset that comprises only four input variables in this study. Another important challenge is about improving the accuracy and reliability of predictions. A prerequisite for solving the above problems is to establish a

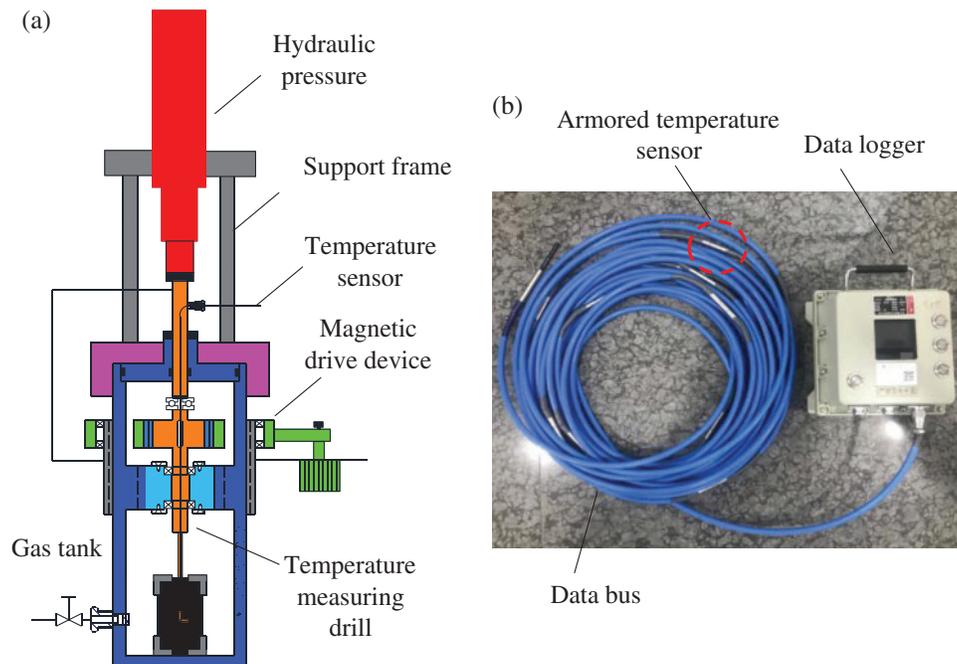
big data information base of static geological elements that includes in situ stress data, basic gas parameters, geo-temperature parameters, and parameters of other physical fields, serving for permeability prediction, dynamic intelligence, intelligent decision-making, and intelligent control of CBM recovery (Fig. 1).



**Figure 15:** PDP of UCS on permeability prediction

To note, input variables in the dataset should be the data obtained from the engineering field and facilitate prediction of permeability dynamics, though laboratory study is indispensable during database establishment. UCS used in this study was measured with a time delay, as laboratory testing is required to characterize pore–fracture characteristic  $X_{13}$ , so we have to establish the relationship of Logging-While-Drilling (LWD) parameter  $X_{13}'$  vs. UCS in the future. Metamorphism degree  $X_{14}$  determines normally gas content  $X_{14}'$  of the coal seam which consists of volume of desorbed gas on site, volume of lost gas, and volume of residual gas, and these are basic gas parameters to be frequently determined in engineering practice. However, the content of gas lost due to gas desorption induced by high temperature of friction during the drilling often relies on empirical formulas that cause errors. To solve two problems above, we developed a gas sorption-desorption apparatus matching Fig. 2b, for which a sorption tank was specially designed, including a temperature-measuring drill and a magnetic drive device (Fig. 16a). The apparatus enables the study on sorption and desorption of gas in lump coal and acquisition of LWD parameters and real-time drilling temperature, thereby the relationship of UCS–LWD parameters–drilling line temperature–volume of lost gas can be established to characterize  $X_{13}$  and  $X_{14}$  with  $X_{13}'$  and  $X_{14}'$ , respectively. In this study, temperature range was 20–70° C, but in fact, the temperature at heat injection site might be higher than 200°C, hence, we designed a geo-temperature tester including a data bus fitted with multi-point temperature sensors and a data logger (Fig. 16b), acquired geo-temperature in the heat injection affected region and conducted permeability evolution test at a temperature above 70°C.

Finally, the expected form of the input dataset is  $k = f(X_{11}, X_{12}, X_{13}', X_{14}', X_{15}, X_{21}, X_{22}, X_{23}, \dots)$ , where  $f$  is a function representing the best robustness machine learning algorithm.



**Figure 16:** Apparatus setup: (a) Gas tank; (b) Multi-point geo-temperature tester

## 6 Conclusions

An intelligent modelling framework for the prediction of raw coal permeability based on Random Forest and Genetic Algorithm was proposed in this study. The coal samples were collected from coal seam No. 8 at Xishan Coalfield in China. Four input variables, including UCS, effective stress, temperature, and gas pressure, were selected based on the established dual-porosity seepage model. A total of 50 seepage and UCS tests were performed using the self-developed triaxial testing apparatus. During hyper-parameters tuning, 5-fold CV was used. Statistical measures including  $R^2$  and RMSE were used to validate and compare the prediction performance of the RF and RF-GA models. Based on the results and discussion, the following conclusions were drawn:

- (1) The RF-GA model was found to be more suitable for permeability prediction with  $R^2 = 0.841$  and  $RMSE = 0.082$ , indicating that GA is efficient in tuning RF architecture. The optimal RF hyperparameters were  $max\_depth = 13$ ,  $min\_sample\_split = 3$ ,  $min\_samples\_leaf = 1$ ,  $Max\_DT = 20$ , and  $max\_features = 0.821$ .
- (2) UCS is directly correlated to permeability, whereas, effective stress, gas pressure, and temperature are all inversely related to permeability.
- (3) A sensitivity study of input variables was undertaken using the methods of Importance Scores and Partial Dependence Plots. The results manifest that the effect stress (IS 0.742) exerts the maximum impact on permeability prediction, followed by temperature or gas pressure (IS 0.102), and then UCS (IS 0.054). PDPs showed the marginal effect of one feature on prediction result, and the results are in agreement with the test results.

In the future, an expanded dataset can be collected to improve the generalization capability of the modelling. We have developed relevant equipment and highlight the practicality of inputs in field engineering.

**Authors' Contributions:** Junhui Wang and Wanzi Yan contributed equally to this work.

**Funding Statement:** This work has been supported by the Fundamental Research Funds for the Central Universities [2017XKZD06].

**Conflicts of Interest:** The authors declare that there is no conflict of interest regarding the publication of this paper.

## References

1. Xue, Y., Dang, F., Li, R., Liu, F. (2018). Numerical investigation of the effect of sorption time on coal permeability and gas pressure. *Computer Modeling in Engineering & Sciences*, 115(3), 345–358.
2. Qin, Y., Moore, T. A., Shen, J., Yang, Z. B. Shen, Y. L. et al. (2018). Resources and geology of coalbed methane in China: A review. *International Geology Review*, 60(5–6), 777–812.
3. Karimpouli, S., Tahmesbi, P., Ramandi, H. L. (2020). A review of experimental and numerical modeling of digital coalbed methane: Imaging, segmentation, fracture modeling and permeability prediction. *International Journal of Coal Geology*, 228, 103552. DOI 10.1016/j.coal.2020.103552.
4. Aminian, K., Ameri, S. (2009). Predicting production performance of CBM reservoirs. *Journal of Natural Gas Science and Engineering*, 1(1–2), 25–30.
5. Lin, J., Ren, T., Cheng, Y. P., Nemcik, J., Wang, G. D. (2019). Cyclic N<sub>2</sub> injection for enhanced coal seam gas recovery: A laboratory study. *Energy*, 188, 116115.
6. Yuan, L., Zhang, P. S. (2020). Framework and thinking of transparent geological conditions for precise mining of coal. *Journal of China Coal Society*, 45(7), 2346–2356.
7. Zhang, Z., Qin, Y., Bai, J., Fu, X., Liu, D. (2016). Evaluation of favorable regions for multi-seam coalbed methane joint exploitation based on a fuzzy model: A case study in southern Qinshui Basin. *China Energy Exploration & Exploitation*, 34(3), 400–417.
8. Xie, H. P., Ju, Y., Ren, S. H., Gao, F., Liu, J. Z. et al. (2019). Theoretical and technological exploration of deep in situ fluidized coal mining. *Frontiers in Energy*, 13(4), 603–611.
9. Ye, J., Shi, B., Zhang, C. (1999). Coal reservoir permeability and its controlled factors in China. *Journal of China Coal Society*, 24(2), 118–122.
10. Tao, S., Wang, Y. B., Tang, D. Z., Xu, H., Lv, Y. M. et al. (2012). Dynamic variation effects of coal permeability during the coalbed methane development process in the Qinshui Basin, China. *International Journal of Coal Geology*, 93, 16–22.
11. Somerton, W. H., Söylemezoğlu, I., Dudley, R. (1975). Effect of stress on permeability of coal. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 12(5–6), 129–145.
12. Wang, D. K., Lv, R. H., Wei, J. P., Fu, Q. C., Wang, Y. T. et al. (2019). An experimental study of seepage properties of gas-saturated coal under different loading conditions. *Energy Science & Engineering*, 7(3), 799–808.
13. Pan, Z. J., Connell, L. D., Camilleri, M., Connelly, L. (2010). Effects of matrix moisture on gas diffusion and flow in coal. *Fuel*, 89(11), 3207–3217.
14. Feng, Z. J., Wan, Z. J., Zhao, Y. S., Li, G. W., Zhang, Y. et al. (2010). Experimental study of permeability of anthracite and gas coal masses under high temperature and triaxial stress. *Chinese Journal of Rock Mechanics & Engineering*, 29(4), 689–696.
15. Wang, E. Y., Zhang, L., He, X. Q., Liu, Z. T. (2004). Electric field response of gas permeability of coal. *Journal of China University of Mining & Technology*, 33(1), 62–65.

16. Zhou, J. P., Xian, X. F., Li, X. H., Xu, J. U. (2010). Effect of different adsorptional gases on permeability of coal. *Chinese Journal of Rock Mechanics & Engineering*, 29(11), 2256–2262.
17. Chao, Z. M., Ma, G. T., Wang, M. (2020). Experimental and numerical modelling of the mechanical behaviour of low-permeability sandstone considering hydromechanics. *Mechanics of Materials*, 148, 103454.
18. Wang, M., Wang, F., Zhu, Z. M., Dong, Y. Q., Nezhad, M. M. et al. (2019). Modelling of crack propagation in rocks under SHPB impacts using a damage method. *Fatigue & Fracture of Engineering Materials & Structures*, 42(8), 1699–1710.
19. Wang, M., Zhu, Z., Dong, Y., Zhou, L. (2017). Study of mixed-mode I/II fractures using single cleavage semicircle compression specimens under impacting loads. *Engineering Fracture Mechanics*, 177, 33–44.
20. Yin, Q., Ma, G., Jing, H., Wang, H., Su, H. et al. (2017). Hydraulic properties of 3D rough-walled fractures during shearing: An experimental study. *Journal of Hydrology*, 555, 169–184.
21. Yin, Q., Jing, H., Ma, G., Su, H., Liu, R. (2018). Investigating the roles of included angle and loading condition on the critical hydraulic gradient of real rock fracture networks. *Rock Mechanics and Rock Engineering*, 51(10), 3167–3177.
22. Seidle, J., Huitt, L. (1995). Experimental measurement of coal matrix shrinkage due to gas desorption and implications for cleat permeability increases. *International Meeting on Petroleum Engineering, Society of Petroleum Engineers*, Beijing, China. DOI 10.2118/30010-MS.
23. Palmer, I., Mansoori, J. (1996). How permeability depends on stress and pore pressure in coalbeds: A new model. *SPE Reservoir Evaluation & Engineering*, 1(06), 539–544.
24. Shi, J., Durucan, S. (2004). Drawdown induced changes in permeability of coalbeds: A new interpretation of the reservoir response to primary recovery. *Transport in Porous Media*, 56(1), 1–16.
25. Yin, Q., Liu, R., Jing, H., Su, H., Yu, L. et al. (2019). Experimental study of nonlinear flow behaviors through fractured rock samples after high-temperature exposure. *Rock Mechanics and Rock Engineering*, 52(9), 2963–2983.
26. Zhou, H. W., Rong, T. L., Mou, R. Y., Wang, L. J., Ren, W. G. (2019). Development in modeling approaches to mining-induced permeability of coals. *Journal of the China Coal Society*, 44(1), 221–235.
27. Levine, J. R. (1996). Model study of the influence of matrix shrinkage on absolute permeability of coal bed reservoirs. *Geological Society, London, Special Publications*, 109(1), 197–212.
28. Tian, J. W., Qi, C. C., Sun, Y. F., Yaseen, Z. M. (2020). Surrogate permeability modelling of low-permeable rocks using convolutional neural networks. *Computer Methods in Applied Mechanics and Engineering*, 366, 113103.
29. Yin, G. Z., Li, M. H., Li, W. P., Cao, J., Li, X. (2013). Model of coal gas permeability prediction based on improved BP neural network. *Journal of China Coal Society*, 38(7), 1179–1184.
30. Xie, L. R., Lu, P., Wang, J. R., Gao, L., Niu, Y. Q. et al. (2017). LVQ-CPSO-BP-based prediction technique of coal gas permeability rate. *Journal of Mining and Safety Engineering*, 34(2), 398–404.
31. Shao, L. S., Ma, H. (2015). Model of coal gas permeability prediction based on PSO-LSSVM. *Coal Geology & Exploration*, 38(7), 1179–1184.
32. Mao, Z. Y., Huang, C. J., Lu, S. C., Han, R. Y. (2019). Model of gas-bearing coal permeability prediction based on APSO-WLS-SVM. *Coal Geology & Exploration*, 47(2), 66–71,78.
33. Tang, G. S., Zhang, H. W., Han, J., Song, W. H. (2015). Prediction model on permeability of gas-bearing coal based on MABC-SVM. *Journal of Safety ence and Technology*, (2), 11–16.
34. Li, B., Zhang, X. X., Li, T. T., Che, X. Q. (2016). Prediction method of reservoir permeability of coal bed methane. *Journal of Hlongjiang University of ence and Technology*, 26(5), 480–484.
35. Habibi, M. J., Mokhtari, A. R., Baghbanan, A., Namdari, S. (2014). Prediction of permeability in dual fracture media by multivariate regression analysis. *Journal of Petroleum Science and Engineering*, 120, 194–201.
36. Sharma, L., Vishal, V., Singh, T. (2017). Predicting CO<sub>2</sub> permeability of bituminous coal using statistical and adaptive neuro-fuzzy analysis. *Journal of Natural Gas Science and Engineering*, 42, 216–225.

37. Xie, Y. N. (2011). *Study on rough set theory in the prediction of coal reservoir permeability (Ph.D. Thesis)*. China University of Geoscience, Qingdao.
38. Wang, J., Wan, Z., Wang, Y., Liu, Z., Liu, S. et al. (2020). Effect of stress and moisture content on permeability of gas-saturated raw coal. *Geofluids*, 2020, 8837758.
39. Zhu, W. C., Wei, C. H., Liu, J., Qu, H. Y., Elsworth, D. (2011). A model of coal-gas interaction under variable temperatures. *International Journal of Coal Geology*, 86(2–3), 213–221.
40. Teng, T., Wang, J. G., Gao, F., Ju, Y., Jiang, C. C. (2016). A thermally sensitive permeability model for coal-gas interactions including thermal fracturing and volatilization. *Journal of Natural Gas Science and Engineering*, 32, 319–333.
41. Zhao, Y., Cao, S. G., Li, Y., Zhang, Z. Y., Guo, P. et al. (2018). The occurrence state of moisture in coal and its influence model on pore seepage. *RSC Advances*, 8(10), 5420–5432.
42. Tao, Y. Q., Xu, J., Cheng, M. J., Li, S. C., Peng, S. J. (2009). Theoretical analysis and experimental study on permeability of gas-bearing coal. *Chinese Journal of Rock Mechanics and Engineering*, 28(S2), 3364–3370.
43. Bear, J., Corapcioglu, M. Y. (1983). Mechanics of fluids in porous media. *Eos, Transactions American Geophysical Union*, 64(11), 109–110.
44. Scheidegger, A. E. (1958). The physics of flow through porous media. *Soil Science*, 86(6), 355.
45. Sun, Y., Genton, M. G. (2011). Functional boxplots. *Journal of Computational and Graphical Statistics*, 20(2), 316–334.
46. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
47. Lingaraj, H. (2016). A study on genetic algorithm and its applications. *International Journal of Computerences & Engineering*, 4(10), 139–143.
48. Qi, C. C., Chen, Q. S., Fourie, A., Zhang, Q. L. (2018). An intelligent modelling framework for mechanical properties of cemented paste backfill. *Minerals Engineering*, 123, 16–27.
49. Qasem, S. N., Samadianfard, S., Nahand, H. S., Mosavi, A., Shamshirband, S. et al. (2019). Estimating daily dew point temperature using machine learning algorithms. *Water*, 11(3), 582.
50. Foroozesh, J., Abdalla, A. I. M., Zhang, Z. (2019). Pore network modeling of shale gas reservoirs: Gas desorption and slip flow effects. *Transport in Porous Media*, 126(3), 633–653.
51. Sun, Y., Zhang, J., Li, G., Wang, Y., Sun, J. et al. (2019). Optimized neural network using beetle antennae search for predicting the unconfined compressive strength of jet grouting coalcretes. *International Journal for Numerical and Analytical Methods in Geomechanics*, 43(4), 801–813.
52. Sun, Y., Zhang, J., Li, G., Ma, G., Huang, Y. et al. (2019). Determination of Young's modulus of jet grouted coalcretes using an intelligent model. *Engineering Geology*, 252, 43–53.
53. Liu, R., Huang, N., Jiang, Y., Jing, H., Yu, L. (2020). A numerical study of shear-induced evolutions of geometric and hydraulic properties of self-affine rough-walled rock fractures. *International Journal of Rock Mechanics and Mining Sciences*, 127, 104211.
54. Liu, R., He, M., Huang, N., Jiang, Y., Yu, L. (2020). Three-dimensional double-rough-walled modeling of fluid flow through self-affine shear fractures. *Journal of Rock Mechanics and Geotechnical Engineering*, 12(1), 41–49.
55. Li, Y., Tang, D. Z., Xu, H., Meng, Y. J., Li, J. (2014). Experimental research on coal permeability: The roles of effective stress and gas slippage. *Journal of Natural Gas Science and Engineering*, 21, 481–488.
56. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.
57. Zhang, J., Li, D., Wang, Y. (2020). Predicting tunnel squeezing using a hybrid classifier ensemble with incomplete data. *Bulletin of Engineering Geology and the Environment*, 79(6), 3245–3256.

58. Zhang, J., Wang, Y., Sun, Y., Li, G. (2020). Strength of ensemble learning in multiclass classification of rockburst intensity. *International Journal for Numerical and Analytical Methods in Geomechanics*, 44(13), 1833–1853.
59. Wang, J. J., Yu, L., Yuan, Q. W. (2019). Experimental study on permeability in tight porous media considering gas adsorption and slippage effect. *Fuel*, 253, 561–570.
60. Li, B., Wei, J. P., Wang, K., Jia, Y. N. (2014). Experimental study of nonlinear motion law for gas seepage in coal seams. *Chinese Journal of Rock Mechanics & Engineering*, 33, 3219–3224.
61. Hao, M. (2019). *Study on evolution characteristics of hydraulic fracture and gas seepage law in low permeability coal seam (Ph.D. Thesis)*. Xi'an University of Science and Technology, Xi'an.