



Research article

Strength prediction and application of cemented paste backfill based on machine learning and strength correction

Bo Zhang ^{a,b}, Keqing Li ^{a,b}, Siqi Zhang ^{a,b}, Yafei Hu ^{a,b}, Bin Han ^{a,b,*}^a School of Civil and Resource Engineering, University of Science and Technology Beijing, Beijing 100083, China^b Key Laboratory of Ministry of Education of China for Efficient Mining and Safety of Metal Mines, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Keywords:

Backfill
Uniaxial compressive strength
Cemented paste backfill
Neural network
Optimization

ABSTRACT

Cemented paste backfill (CPB) is widely used in mines production practices around the world. The strength of CPB is the core of research which is affected by factors such as slurry concentration and cement content. In this paper, a research on the UCS is conducted by means of a combination of laboratory experiments and machine learning. BPNN, RBFNN, GRNN and LSTM are trained and used for UCS prediction based on 180 sets of experimental UCS data. The simulation results show that LSTM is the neural network with the optimal prediction performance (The total rank is 11). The trial-and-error, PSO, GWO and SSA are used to optimize the learning rate and the hidden layer nodes for LSTM. The comparison results show that GWO-LSTM is the optimal model which can effectively express the non-linear relationship between underflow productivity, slurry concentration, cement content and UCS in experiments ($R = 0.9915$, $RMSE = 0.0204$, $VAF = 98.2847$ and $T = 16.37$ s). The correction coefficient (k) is defined to adjust the error between predicted UCS in laboratory (UCS_M) and predicted UCS in actual engineering (UCS_A) based on extensive engineering and experimental experience. Using GWO-LSTM combined with k , the strength of the filling body is successfully predicted for 153 different filled stopes with different stowing gradient at different curing times. This study provides both effective guidance and a new intelligent method for the support of safety mining.

1. Introduction

Filling mining is a green and efficient mining method that can both recycle solid waste in mine and improve the safety of underground mining operations [1, 2, 3, 4]. It is widely used in mines all around the world because it can effectively solve the safety hazards from the mined-out areas and tailings reservoir [5, 6, 7, 8]. Cemented paste backfill (CPB) is an important technology in filling mining methods which uses tailings, waste rock or river sand as aggregate and mixed it with binder to make a filling slurry and then filled it into the mined-out area [9, 10, 11, 12, 13, 14]. CPB forms a filling body with strength and can effectively support the mined-out area. Uniaxial compressive strength (UCS) is the essential strength metric for CPB, which is affected by slurry concentration, binder content, curing time and particle size for tailings [15, 16, 17, 18, 19, 20]. Many scholars have conducted mechanistic modeling of UCS based on different affecting factors and have achieved instructive results [21, 22, 23, 24, 25]. However, the CPB is a complex multiphase mixture and the formation of UCS depends on

the hydration of the binder and the coupling effect of different factors. Therefore, the use of simple mechanistic modeling cannot construct a valid mapping relationship between UCS and different affecting factors.

In the age of artificial intelligence, many scholars have done prospective and innovative studies on predicting UCS in recent years, as detailed in Table 1 [1, 20, 26, 27, 28, 29, 30, 31]. Artificial neural networks (ANN), a popular branch of machine learning technology, have been widely used in various engineering practices and modeling. ANN processes complex information by simulating the nervous system of the human brain and have a unique knowledge representation and intelligent adaptive learning capability. Meanwhile, intelligent optimization algorithms (IOA) as the core of artificial intelligence have also gained popularity in engineering, especially in the field of optimization in engineering design. IOA can rapidly solve optimization problems in multidimensional spaces that are difficult to solve by traditional computational methods. In filling mining, ANN and IOA are combined and applied in the strength design of the CPB. This is because IOA can be used to optimize ANN for parameters such as hidden layer nodes and

* Corresponding author.

E-mail addresses: zszhangbo210@163.com (B. Zhang), hanbin_ustb@163.com (B. Han).<https://doi.org/10.1016/j.heliyon.2022.e10338>

Received 5 February 2022; Received in revised form 30 April 2022; Accepted 12 August 2022

Table 1. Summary of ANN and IOA applications in filling mining.

Authors	Publish time	Main methods	Input	Output
Sivakugan et al. [29]	2005	ANN	Cement content, solids content, curing time and grain size distribution	Strength
Orejarena et al. [31]	2010	ANN	Water-cement ratios, binder composition and binder content	150d UCS
Qi et al. [28]	2018	PSO, ANN	Tailings type, cement-tailings ratio, solids content, and curing time	3d, 7d, 28d UCS
Qi et al. [26]	2018	PSO	Physical and chemical characteristics of tailings, cement-tailings ratio, solids content, and curing time.	7d, 28d UCS
Qi et al. [27]	2018	GA	Physical and chemical characteristics of tailings, the cement-tailings ratio, the solids content, and the curing time	28d UCS
Xiao et al. [30]	2020	ANN	Proportion of OPC, CG, CFA, and YS, solids content, and curing time	1d, 3d, 28d UCS
Yu et al. [20]	2021	ELM, SSA	Cement-to-tailing mass ratio, solid mass concentration, fiber content, fiber length, and curing time	3d, 7d, 28d, and 56d UCS
Li et al. [1]	2021	GA, PSO, SSA	Fiber properties, cement type, curing time, cement-tailings ratio and concentration	3d, 7d, 28d UCS

learning rate. Many researchers have done a lot of research on the application of ANN and IOA in filling mining.

As can be seen in Table 1, ANN and IOA have been shown to have excellent application in filling mining, especially in predicting the UCS of CPB, in previous researches. However, traditional ANNs suffer from gradient disappearance and gradient explosion in predicting UCS [26, 27, 28]. In addition, the neurons of the hidden layer in a traditional ANN are independent and do not affect each other, which makes the ANN lack the memory and storage function of data when processing information [32]. In filling mining, different materials and multiple factors have different degrees of effect on the strength of the filling body. Besides, force majeure factors such as uneven slurry mixing and impurity mixing during the experimental process can also affect UCS. This can lead to outliers in the UCS obtained from laboratory experiments [15, 16, 17, 18, 19, 20]. In the conditions of above problems, the weakness of traditional ANNs without data storage and memory mechanisms is exposed, which makes the accuracy of the constructed UCS prediction models also suffer. This not only affect the accuracy of ANN in predicting UCS, but also limit the in-depth application of ANN in filling mining. In a word, the traditional ANN in predicting UCS needs to be improved to suit the requirements of future mine intelligence development.

Long short-term memory neural network (LSTM), as a new neural network with excellent generalization ability and robustness, is currently less used in fill mining [33, 34]. Its unique gating mechanism changed the structure of the traditional ANN, which gives ANN the function of memorializing and forgetting information [35]. The core of this study is to explore and construct the prediction model for UCS of CPB with LSTM as the fundamental model and IOA as the optimization method.

In this paper, section 2 and section 3 explain that underflow productivity for tailings, slurry concentration, cement content and curing time are used as affecting factors to design the mix proportion experiments of the filling slurry and 180 sets of experimental data are obtained after the UCS test. Section 4 presents four neural networks, four optimization algorithms, and the performance metrics in this study. Section 5 illustrates the process of selecting LSTM as the basic model for this study by comparing the prediction accuracy of different neural networks. Section 6 illustrates the accuracy of the improved LSTM prediction of UCS of CPB using different optimization algorithms and determines the LSTM improved by grey wolf optimization (GWO-LSTM) model as the optimal prediction model. In section 7, GWO-LSTM that is the optimal model, is combined with the correction coefficient (k) to predict the strength of the filling body of 153 approaches in filled stope. This study not only provides guidance for safety mining, but also bring new tools for the intelligent development of mines.

2. Materials

The material used to prepare CPB is generally composed of two parts: aggregate and binder. Aggregate, an inert material, is used to

Table 2. The main chemical composition of the tailings.

Composition	SiO ₂	CaO	MgO	Fe ₂ O ₃	Al ₂ O ₃	Others
Content (%)	69.3	2.47	0.59	2.03	18.84	6.77

form the skeleton structure of CPB to enhance durability and effectively inhibit crack expansion. Commonly used aggregates are tailings, river sand and wind sand, etc. The binder mainly plays the role of gelling effect, which triggers the hydration reaction after mixing with water and generates hydration products to fill the gaps between the aggregates. Commonly used binders are cement, fly ash, slag powder, etc. The combination of the skeletal structure formed by the aggregates and the gelling effect of the binder gives CPB its strength. In this study, tailings and cement are used as aggregate and binder, respectively.

2.1. Tailings

The tailings used in this study are from the Jinfeng Mining Limited (Gold Mine), Southwest of Guizhou province, China. The composition of the tailings by using X-ray fluorescence (XRF) is shown in Table 2. From the Table 2, the main composition of the tailings is SiO₂, which indicates that the tailings is an inert material and can be used as an aggregate to produce the filling slurry. Due to the fine size of the tailings, the hydrocyclone is used to classify the tailings and obtained classified tailings with underflow productivity of 45%, 55%, 65% and 75% respectively. The comparison of particle size for tailings of different underflow productivity is shown in Table 3.

2.2. Binder

The binder is the Portland cement (P.O42.5). The density of cement is 3.1 g/cm³ and the 28d compressive strength is higher than 45 MPa. The cement performance parameter is shown in Table 4.

3. Experiments

UCS is an important indicator used to evaluate whether CPB fulfills engineering standards. The level of UCS affects whether CPB satisfies the requirements of mine safety production. The current methods of measuring the UCS of CPB are mainly mix proportion experiments in the laboratory and the UCS test using a hydraulic testing machine.

3.1. Mix proportion of materials

In this study, proportional tests are conducted with underflow productivity, slurry concentration, cement content and curing time as the main affecting factors of UCS of CPB. The mix proportions are shown in Table 5.

Table 3. The comparison of particle size for tailings of different underflow productivity.

Underflow productivity	Main particle size and cumulative proportion					
	-5 μm	-10 μm	-20 μm	-30 μm	-45 μm	-75 μm
100%	30.07%	44.44%	58.92%	70.73%	79.31%	96.08%
75%	30.53%	41.30%	52.81%	64.44%	74.62%	96.14%
65%	23.92%	33.20%	43.87%	58.65%	72.08%	94.09%
55%	18.76%	26.80%	36.32%	50.86%	65.39%	93.82%
45%	13.24%	20.36%	29.52%	42.55%	57.38%	83.67%

Table 4. The cement performance parameter.

Level	Fineness	Bending strength		Compressive strength		Setting time	
		3d	28d	3d	28d	Initial	Final
P.O42.5	≤6.0%	17 MPa	20 MPa	≥25 MPa	≥45 MPa	2.0h	4.0h

Table 5. The mix proportion of materials.

Affecting factors	Mix proportion				
Underflow productivity	45%	55%	65%	75%	\
Slurry concentration	63%	64%	65%	\	\
Cement content	5%	5.5%	6%	6.5%	7%
Curing time	5d	7d	28d	\	\

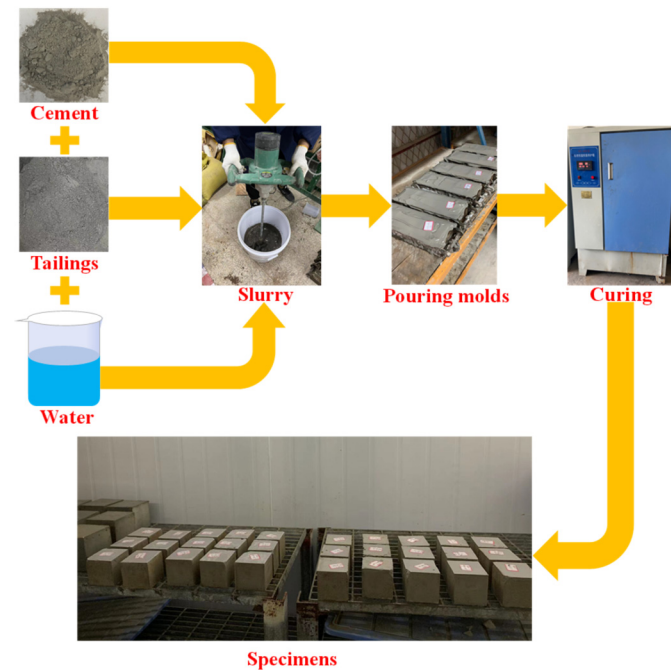


Fig. 1. The flow of the experiments (preparing materials, mixing slurry, making and curing specimen).

3.2. Experimental process

The tailings and cement used in the experiment are weighed according to the different mix proportions in section 3.1 and the backfill slurry is prepared according to Chinese standards GB17671-1999. First, the solid materials are mixed according to above proportions for not less than 5 minutes. Then, put the mixture into a bucket with water and blend for 5 min by using a handheld electric mixer. Lastly, the prepared slurry is poured into a mold with a size of 70.7 mm × 70.7 mm × 70.7 mm. After the specimen is shaped, the mold is dismantled and the specimen is placed in the standard curing room (the temperature is 20 °C and the humidity range is 90-95%) to cure until the predetermined curing time. In addition, 3 specimens are made for each mix proportion and the UCS is taken as their average. Finally, the cured specimens are tested with a press for UCS and 180 sets of UCS data are obtained. The flow of the experiments is shown in Fig. 1

3.3. Experiment results

The experimental results of UCS are shown in Fig. 2. From Fig. 2(a), the mean of UCS obtained from the experiment is 0.3533 and the standard deviation is 0.1252. From Fig. 2(b), the minimum, maximum, first quartile (Q1), third quartile (Q2) and median of UCS are 0.14, 0.96, 0.265, 0.43 and 0.33, respectively. Set the two values of Q3 and Q1 plus or minus 1.5 times the interquartile distance (Q1-1.5 IQR and Q3+1.5 IQR) as the truncation points of the outliers. There are 4 outliers in the experimental UCS and they are all distributed in the upper side of Q3+1.5 IQR, which is the reason for the rightward bias of the distribution of UCS in Fig. 2(a). From Fig. 2, the experimental UCS is concentrated between 0.265 MPa and 0.43 MPa. Although there are a very few outliers in the experimental UCS, this is normal in the UCS test.

4. Methodology

This section introduces 4 neural networks (BPNN, RBFNN, GRNN and LSTM), 3 intelligent optimization algorithms (GWO, PSO and SSA) and 4 performance metrics (RMSE, R, VAF and T). Neural networks are used to construct mapping relationships between inputs (underflow productivity) and output (UCS); intelligent optimization algorithms are used to optimize the neural networks that obtain the best mapping relationships; and performance metrics are used to compare the performance of different models.

4.1. Neural network

Neural network is an algorithmic mathematical model that imitates the behavioral characteristics of animal neural networks and performs distributed parallel information processing. It relies on the complexity of the system and by adjusting the interconnections between internal nodes for the purpose of processing information. This section introduces 4 neural networks for searching and constructing the relationship between inputs and outputs in this study.

For the fairness of the comparison of different neural networks, the hidden layer nodes, the learning rate and the maximum iterations are set to 5, 0.1 and 100 for all neural networks. At the same time, the number of hidden layers is set to 1 for both BPNN and RBFNN; the number of pattern layer and summation layer of GRNN are both set to 1. In addition, the important parameter spread factor of RBF and GRNN is set to 5. The above parameter settings are based on relevant researches and modeling experience [20, 36, 37, 38, 39].

4.1.1. Back propagation neural network

Back Propagation neural network (BPNN) is a multilayer feedforward neural network [40, 41]. It is characterized by the forward transmission of the signal while the error is reverse transmitted. In forward transmission, the input signal is processed layer by layer from the input layer through the hidden layer to the output layer. In this process, the state of neurons in each layer affects only the state of neurons in the

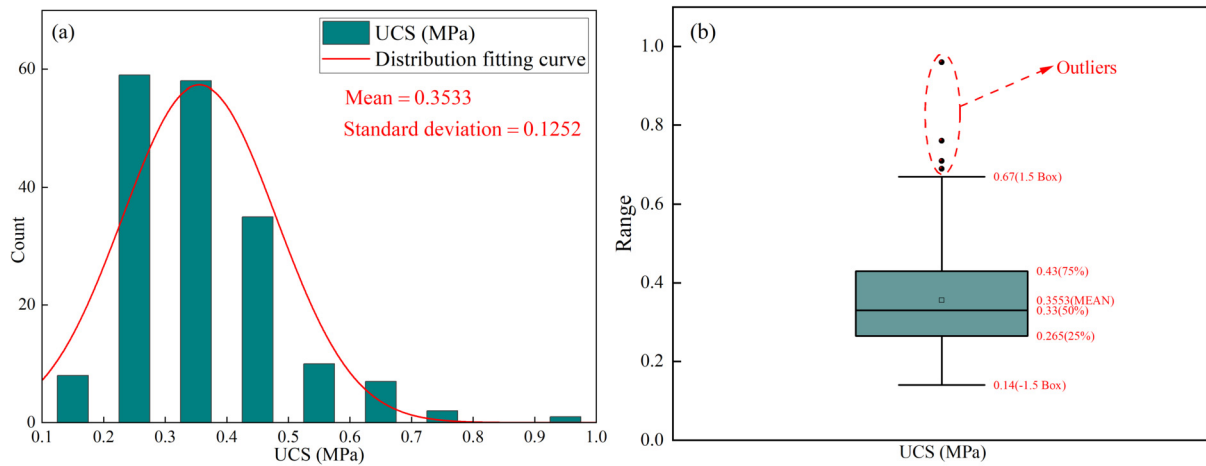


Fig. 2. Distribution of experimental UCS (a) distribution curve of UCS; (b) box line diagram of UCS.

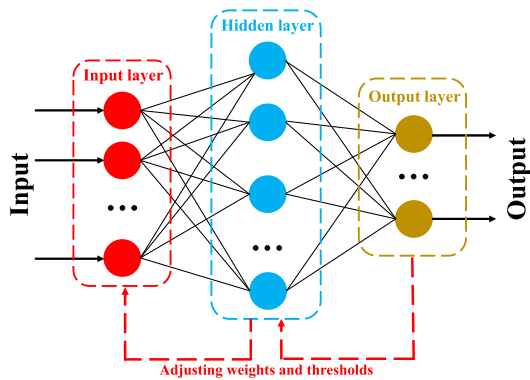


Fig. 3. The network structure of BPNN.

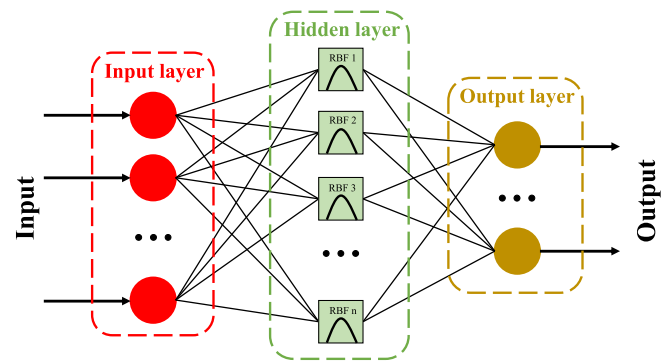


Fig. 4. The network structure of RBFNN.

next layer [42, 43]. If the desired output is not obtained in the output layer, the network is back-propagated and the weights and thresholds are adjusted according to the prediction error, which enables the predicted output of the BPNN to continuously approach the desired output. The network structure of BPNN is shown in Fig. 3. The neuronal structure of the BPNN in this study is set according to the literatures [26, 27, 28] as 4-8-1.

4.1.2. Radial basis function neural network

Similar to BPNN, Radial basis function neural network (RBFNN) is also a forward neural network, which has a network structure consisting of input layer, hidden layer and output layer [44]. RBFNN is characterized by the introduction of RBF as the transform function of neurons in its hidden layer. This enables the hidden layer unit to transform the input vector from low-dimensional and linearly inseparable to high-dimensional and linearly separable to improve generalizability and robustness [45, 46]. The network structure of RBFNN is shown in Fig. 4. To make the comparison fair, the neuronal structure of RBFNN is set to 4-8-1.

4.1.3. General regression neural network

General regression neural network (GRNN) is a modified RBFNN. Unlike RBFNN, the structure of GRNN has 4 layers: input layer, pattern layer, summation layer and output layer [47, 48]. The regression of GRNN on the inputs is different from the least-square superposition of Gaussian weights in RBF, which uses the density function to predict the output by using the calculation of the pattern layer and the summation layer [49, 50, 51]. The network structure of GRNN is shown in Fig. 5. The number of neurons in the pattern layer of GRNN is the number of variables in the input layer. In addition, two types of neurons are used in

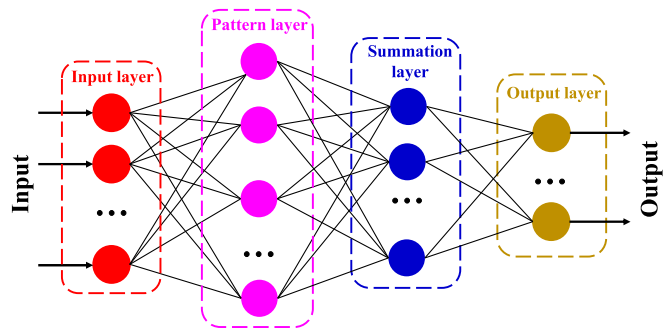


Fig. 5. The network structure of GRNN.

the summation layer for summation. Therefore, the neuronal structure of GRNN is 4-144-2-1. The setting of the neuron structure of GRNN is referred to the literatures [47, 48, 49, 50, 51] and is not described here due to the limitation of space.

4.1.4. Long short-term memory neural network

Long short-term memory neural network (LSTM) is an improved recurrent neural network (RNN). The difference between LSTM and traditional RNN is its addition of memory units to the network structure [52, 53, 54]. The memory cell structure of LSTM is shown in Fig. 6. The LSTM processes information through a four-part gating mechanism consisting of forget gate f , input gate i , output gate o , and memory cell c [55]. The structure of LSTM is as follows.

(1) There are two parts of input at moment t : the output result h_{t-1} at the previous moment and the new input information x_t at this moment.

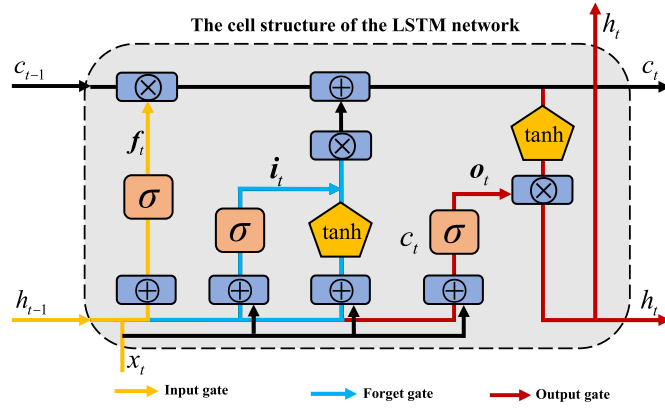


Fig. 6. The memory cell structure of LSTM.

The forget gate f is used to control the amount of input information that would be forgotten. Its calculation equation is as follows.

$$f_t = \sigma(\mathbf{W}_f x_t + \mathbf{U}_f h_{t-1} + \mathbf{b}_f) \quad (1)$$

(2) The input gate i is used to control the amount of input information that would be engaged in the update of the cell. Its calculation equation is as follows.

$$i_t = \sigma(\mathbf{W}_i x_t + \mathbf{U}_i h_{t-1} + \mathbf{b}_i) \quad (2)$$

(3) Calculate the state of the cell \tilde{c}_t at this moment for the current information. Its calculation equation is as follows.

$$\tilde{c}_t = \tanh(\mathbf{W}_c x_t + \mathbf{U}_c h_{t-1} + \mathbf{b}_c) \quad (3)$$

(4) The cell state c is the key to the LSTM model. Updates to the previous cell state c_{t-1} . Its calculation equation is as follows.

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (4)$$

(5) Calculate the output gate o and calculate how the amount of information under the output gate control would be used to generate h_t . Its calculation equation is as follows.

$$o_t = \sigma(\mathbf{W}_o x_t + \mathbf{U}_o h_{t-1} + \mathbf{b}_o) \quad (5)$$

$$h_t = o_t \otimes \tanh(c_t) \quad (6)$$

(6) Transfer h_t to the output layer and output y . The equation is calculated as follows.

$$y = \mathbf{W}_y h_t + \mathbf{b}_y \quad (7)$$

In equations (1)–(7), σ is the sigmoid function; \tanh is the hyperbolic cosine function; \mathbf{W} and \mathbf{U} are the weight matrices; \mathbf{b} is the bias vector; \otimes is the sign of the convolution operation; y is the output at time t .

4.2. Optimization algorithm

The optimization algorithm is a heuristic optimizer, which is an algorithm that imitates the laws of the biological world to solve the problem. It is used for solving the optimal solution of the problem under complex conditions. This section presents 3 bionic algorithms to solve the parameters of a neural network with optimal mapping relationship.

4.2.1. Grey wolf optimization

Grey wolf optimization (GWO) is a new intelligent algorithm that imitates the hunting behavior of gray wolves in nature to optimize objectives [56, 57, 58]. Wolves are classified into 4 classes of α , β , δ and ω according to their fitness degree during hunting. Their ranks are ordered from highest to lowest: $\alpha > \beta > \delta > \omega$. In addition, α is called the

head wolf which leads the other wolves in the hunting process and ω is the bottom wolf which is dominated by the other three classes of wolves in the hunting process. The hunting process is as follows.

(1) Search for and surround prey. Its calculation equation is as follows.

$$D = |C \cdot X_p(t) - X(t)| \quad (8)$$

$$X(t+1) = X_p(t) - A \cdot D \quad (9)$$

In equations (8)–(9), t is the current number of iterations; A and C are coefficient vectors; X_p is the position vector of the prey; X is the position vector of the grey wolf; D is the distance vector between the grey wolf and the prey. A and C are calculated as follows.

$$A = 2a \cdot r_1 - a \quad (10)$$

$$C = 2 \cdot r_2 \quad (11)$$

In equations (10)–(11), a decreases from 2 to 0 as the number of iterations increases; r_1 and R are vectors of random numbers in the range [0, 1].

(2) Attacking the prey. Suppose that α is the gray wolf closest to the prey and β and δ are the next closest. Use equation (12) to calculate their distance to the prey and updating the positions of α , β and δ by equation (13). Finally, other gray wolves gradually approach prey through equation (14). Its calculation equation is as follows.

$$D_\alpha = |C_1 \cdot X_\alpha - X|, \quad D_\beta = |C_2 \cdot X_\beta - X|, \quad D_\delta = |C_3 \cdot X_\delta - X| \quad (12)$$

$$X_1 = X_\alpha - A_1 D_\alpha, \quad X_2 = X_\beta - A_2 D_\beta, \quad X_3 = X_\delta - A_3 D_\delta \quad (13)$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3}, \quad (14)$$

In equations (12)–(14), X_α , X_β and X_δ are the positions of α , β and δ at the t -th iteration.

4.2.2. Particle swarm optimization

Particle swarm optimization (PSO) is an intelligent algorithm designed by simulating the hunting behavior of a group of birds [59, 60, 61]. It uses the sharing of information by individuals in a population to make the entire movement of the group in the problem solution space from disorder to order evolutionary process to obtain the optimal solution. PSO is mainly implemented by the following 2 core equations.

$$V_{id}(t+1) = \omega(t)V_i(t) + c_1 r_1 (P_{ibest}(t) - X_{ibest}(t)) + c_2 r_2 (G_{best}(t) - X_{ibest}(t)) \quad (15)$$

$$X_{id}(t+1) = X_{id}(t) + V_{id}(t+1) \quad (16)$$

In equations (15)–(16), r_1 and R are random numbers uniformly distributed within [0, 1]; c_1 and c_2 are learning factors; ω is inertia weight; $P_{ibest}(t)$ is the i -th particle optimal position; $G_{best}(t)$ is the population optimal position; $X_{id}(t)$ is the current position of the particle; $V_{id}(t)$ is the current speed of the particle.

4.2.3. Sparrow search algorithm

The sparrow search algorithm (SSA) was proposed as a new algorithm to solve optimization problems in 2020 [62, 63]. The core principle of SSA is from the group predatory behavior of sparrows in nature [63, 64, 65]. SSA divides sparrow group into finder and follower. The function of the finder is to guide the entire sparrow group to search and hunt. Its position equation is as follows.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot iter_{max}}\right), & R_2 < ST \\ X_{i,j}^t + Q \cdot L, & R_2 \geq ST \end{cases} \quad (17)$$

In equations (17), i is iteration; j is the individual dimension; $X_{i,j}^t$ is the position of the sparrow; $iter_{max}$ is the maximum iteration; $\alpha \in [0, 1]$ is a random number; $R_2 \in [0, 1]$ and $ST \in [0.5, 1]$ are the alarm value and the security threshold respectively. Q is random number.

The role of a follower is to follow the finder in order to acquire better adaptive ability. The equation for its position is as follows.

$$X_{i,j}^{i+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst}^t - X_{i,j}^t}{t^2}\right), & i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L, & i \leq \frac{n}{2} \end{cases} \quad (18)$$

In equations (18), X_{worst}^t is the worst position of finder; X_p^{t+1} is the best position the producer; A is a matrix with each element being -1 or 1 ; n is the number of individual sparrows. In the above equation, $A^+ = A^T(AA^T)^{-1}$.

Some finders can act as vigilantes to help the finder with foraging. When danger is present, the vigilantes would counter-trap or approach other sparrows. The equation for its position is as follows.

$$X_{i,j}^{i+1} = \begin{cases} X_{best}^t + \lambda \cdot |X_{i,j}^t - X_{best}^t|, & f_i > f_g \\ X_{i,j}^t + J \cdot \left[\frac{X_{i,j}^t - X_{worst}^t}{(f_i - f_w) + \epsilon}\right], & f_i = f_g \end{cases} \quad (19)$$

In equation (19), X_{best}^t is the global best position; $\lambda \sim N(0, 1)$ is the parameter used to control the step size; $J \in [-1, 1]$ is a random number; f_i is adaptation degree of current sparrow; f_g and f_w are the global best adaptation degree value and the global worst adaptation degree value, respectively; ϵ is a constant close to 0 to avoid the denominator being 0.

4.3. Performance metrics

To construct the optimal LSTM, the prediction results of different models are compared in this paper. The root mean square error (RMSE), correlation coefficient (R), variance account for (VAF) and running time of model (T/s) are used as evaluation metrics [66, 67]. The equations of the first three metrics are as follows.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [y_i - f(x_i)]^2} \quad (20)$$

$$R = \frac{\sum_{i=1}^n [(y_i - \bar{y})(f(x_i) - \bar{f}(x_i))]}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \cdot \sqrt{\sum_{i=1}^n (f(x_i) - \bar{f}(x_i))^2}} \quad (21)$$

$$VAF = \left[1 - \frac{VAR(y_i - f(x_i))}{VAR(y_i)}\right] \times 100 \quad (22)$$

In equation (20)–(22), y_i and $f(x_i)$ are the measured values and predicted values; \bar{y} is the mean of the measured values; VAF is used to indicate the degree of fit between the measured and predicted values. In addition, VAR is the variance of the dataset and the range is $[0, 1]$. The closer the $RMSE$ is to 0, the closer the R is to 1 and the closer the VAF is to 100, the higher the prediction accuracy of the model.

Moreover, statistical evaluation of measured and predicted values using 95% confidence interval in this paper. The narrower the confidence interval and the more sample points included, the more significant the accuracy of the prediction results in the regression fitting fig.

5. Comparison and selection of networks

This section constructs 4 neural networks to characterize the mapping relationships between inputs and output based on data set in section 4 in conjunction with section 2. Also, the LSTM with the best mapping relationship is optimized using 4 optimized methods and the model with the best prediction effect applicable to this study is obtained.

In this paper, 4 neural networks are trained based on the experimental data in section 3. 80% of the experimental data is used as the training set to train networks and 20% is used as the validation set to test the reliability of the output of the neural networks.

The comparison of the results for validation set of the 4 networks is shown in Fig. 7. From Fig. 7, the RMSE of BPNN, RBFNN, GRNN and LSTM are 0.0703, 0.0719, 0.0883 and 0.0871, respectively. The R of BPNN, RBFNN, GRNN and LSTM are 0.8707, 0.8882, 0.8812 and 0.8974, respectively. The VAF of BPNN, RBFNN, GRNN and LSTM are 75.7899, 76.8267, 61.3377 and 99.9874, respectively. The T of BPNN, RBFNN, GRNN and LSTM are 9.28 s, 8.74 s, 5.46 s and 13.11 s, respectively. The LSTM has a better fit of the predicted values to the measured values but the longest running time, which is because the complex structure of LSTM increases the amount of computation while enhancing the understanding of the data. Moreover, a ranking method for the performance of different models proposed by Zorlu is utilized as shown in Table 6 [68]. From Fig. 7 and Table 6, LSTM and RBFNN have relatively better prediction performance (the total rank are both 11) for UCS of CPB compared with BPNN and GRNN (the total rank are both 10). For the prediction of UCS of CPB, the main focus in practical application is on prediction accuracy and error. Although the running time of LSTM is longest, it is still the optimal model in this study. Therefore, LSTM is selected as the basic model and the next step of the research is conducted.

6. Optimization of LSTM and results

Two parameters, the hidden layer nodes and the learning rate, have important effects on the performance of LSTM [69, 70, 71]. Generally, the higher the hidden layer codes, the higher the computational accuracy and complexity of the LSTM. Therefore, selecting the appropriate the hidden layer codes is essential for LSTM. Similarly, the learning rate has a significant effect on the output of the LSTM. The learning rate is used to control the step size during gradient descent in order to optimize the output of the LSTM. So, selecting an appropriate learning rate can prevent the LSTM from missing the optimal solution and improve its robustness. For LSTM, the values of the above two parameters were usually determined using trial-and-error or some experimental equations, which leads to an unstable performance of LSTM in practical applications. To improve the performance of LSTM and apply it to this study, the hidden layer nodes and the learning rate are optimized using trial-and-error, GWO, PSO and SSA. The optimization and modeling process is shown in Fig. 8. From Fig. 8, experimental data is randomly classified into training set and validation set. The training set is used to train the LSTM improved by different optimization methods to obtain the optimal hyperparameters and to construct the mapping relationship between input and output. The validation set is used to validate whether the trained LSTM has generalizability and prediction accuracy. Moreover, the maximum iterations of LSTM and each optimization algorithm are set to 100 and the population size of each optimization algorithm is set to 20 according to related literatures to ensure fairness in the comparison.

6.1. Results of trial-and-error-LSTM

Some researches have shown that the optimal performance can be achieved when the hidden layer nodes do not exceed 2 times the input plus 1. Similarly, the learning rate is set to a range of 0.1 to 1 (in steps of 0.1) according to previous researches [35, 72, 73]. Therefore, various LSTM with the hidden layer nodes is from 1 to 9 and the range of learning rate is (0.1, 1) be constructed to select the optimal parameters for LSTM. In this process, the 10-fold cross-validation method is used to evaluate the generalization performance of the LSTM and to guide the selection of the parameters of the optimal LSTM. Using the mean absolute error (MAE) as an evaluation metric, the performance of the LSTM with different combinations of the above parameters is shown in Fig. 9.

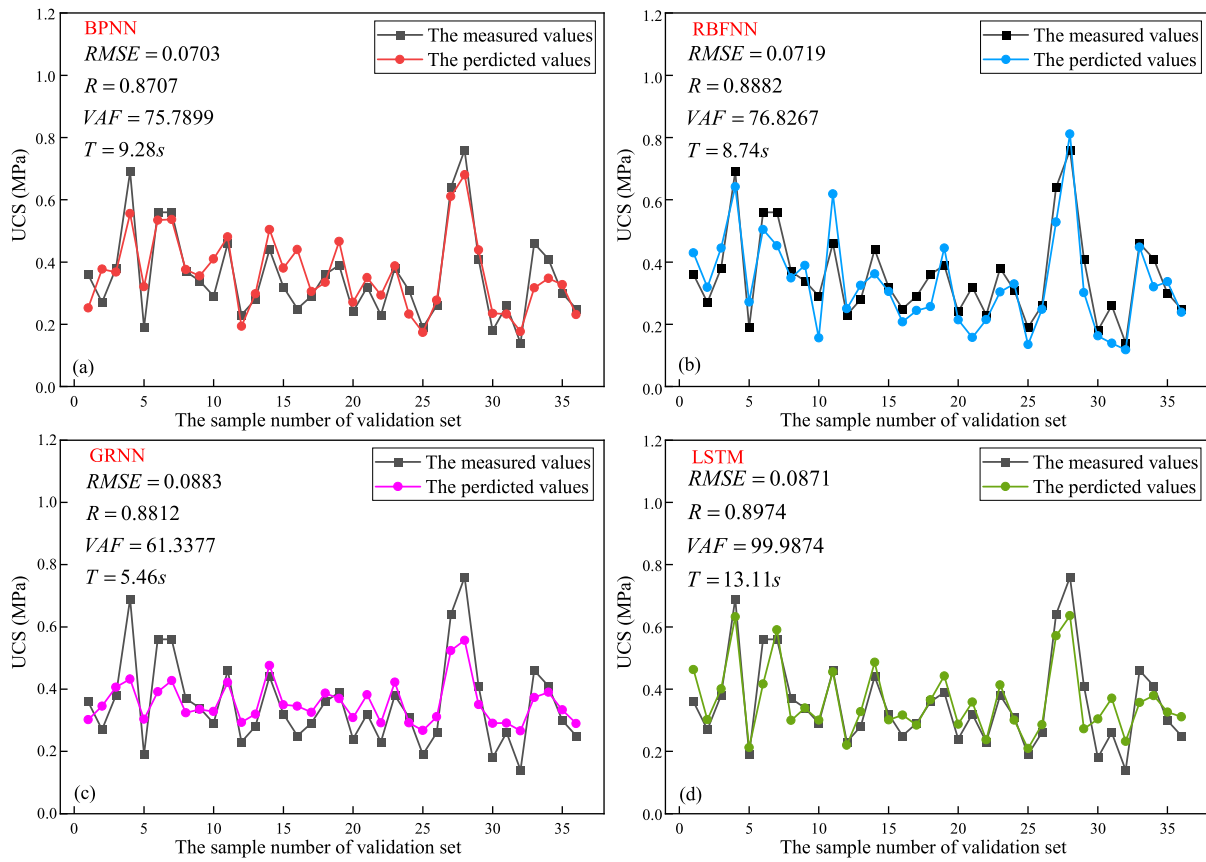


Fig. 7. The prediction results of networks for validation set (a) BPNN; (b) RBFNN; (c) GRNN; (d) LSTM.

Table 6. The performance indices of different networks.

Network	Results				Rank values				Total rank
	R	RMSE	VAF	T	R	RMSE	VAF	T	
BPNN	0.8707	0.0703	75.7899	9.28 s	1	4	2	1	10
RBFNN	0.8882	0.0719	76.8267	8.74 s	3	3	3	2	11
GRNN	0.8812	0.0883	61.3377	5.46 s	2	1	1	4	10
LSTM	0.8974	0.0871	99.9874	13.11 s	4	2	4	1	11

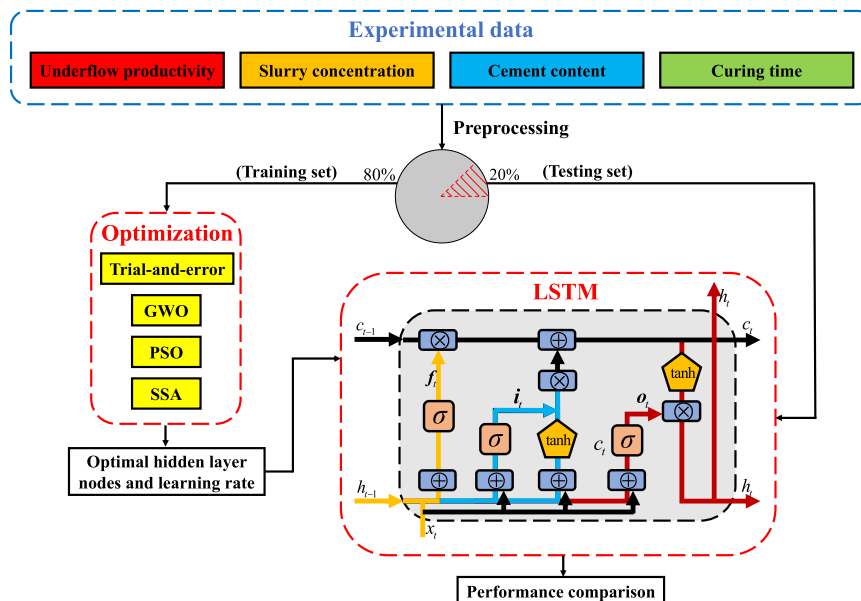


Fig. 8. The optimization and modeling process.

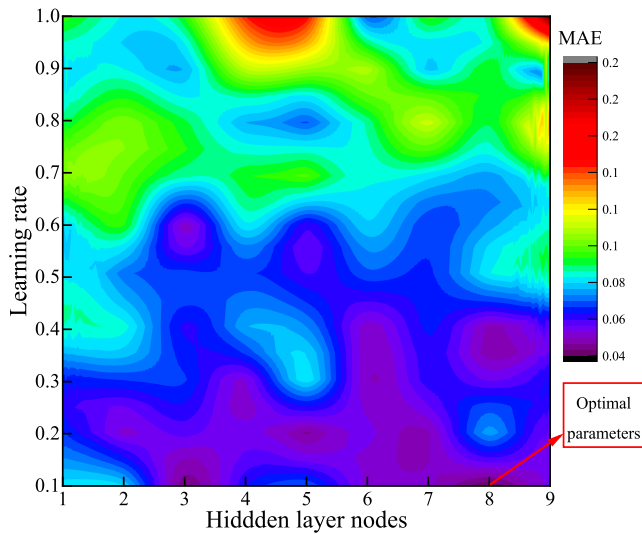


Fig. 9. The performance of the LSTM using trial-and-error.

From the Fig. 9, it can be seen that the effect of learning rate on MAE is more significant than that of hidden layer nodes. When the learning rate is lower than 0.65, MAE is low and basically in the interval (0.04, 0.08); when the learning rate is higher than 0.65, MAE increases significantly and is basically higher than 0.1. When the learning rate is between 0.4 and 0.65, the overall trend of MAE increases and then decreases with the increase of hidden layer nodes and shows a fluctuating change. Also, when the learning rate is between 0.9 and 1, the change of MAE is characterized by fluctuations. MAE obtains minimum when the hidden layer nodes is 8 and the learning rate is 0.1.

The results of the prediction of the validation set using trial-and-error-LSTM are shown in Fig. 10. From the Fig. 10(a), the LSTM yield a prediction performance of $R = 0.9004$, $RMSE = 0.0657$, $VAF = 79.5911$ and $T = 18.61$ s for validation set. From the Fig. 10(b), the narrowest and widest confidence intervals are (0.3521, 0.3856) and (0.6028, 0.7042), respectively; the narrowest and widest prediction intervals are (0.2699, 0.4738) and (0.5409, 0.7661), respectively. This shows that the width of the prediction band is far wider than the confidence band, which indicates that the error between the predicted and measured values is high using trial-and-error-LSTM for the validation set.

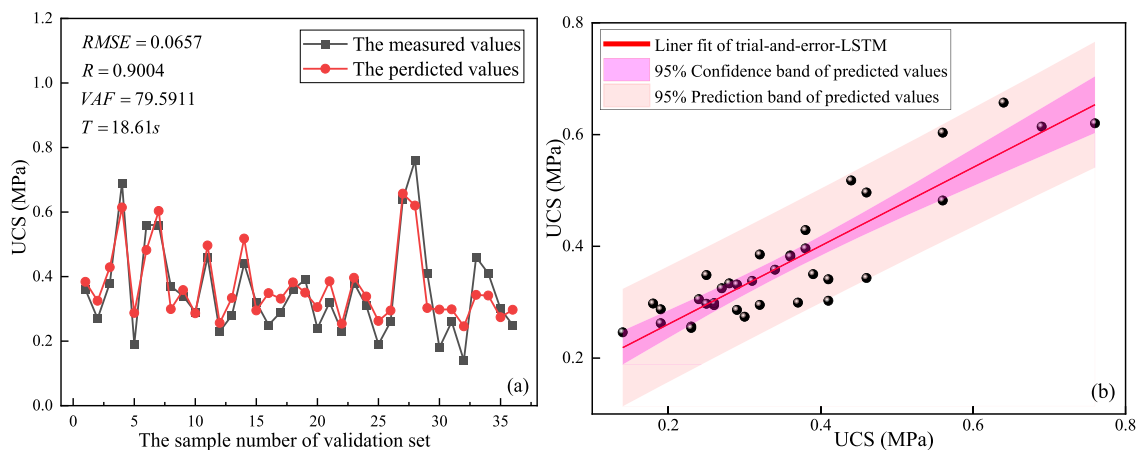


Fig. 10. The prediction performance of the trial-and-error-LSTM (a) the measured and predicted results of trial-and-error-LSTM; (b) the confidence band and prediction band of trial-and-error-LSTM.

6.2. Results of GWO-LSTM

To search for the optimal parameters of the LSTM, the search space of the hidden layer nodes and the learning rate are expanded to (1, 100) and (0.01, 1), respectively. Same settings in both PSO-LSTM and SSA-LSTM as above. The hidden layer nodes obtained using GWO is 25 and the learning rate is 0.0785.

The results of the prediction of the validation set using GWO-LSTM are shown in Fig. 11. From the Fig. 11(a), the GWO-LSTM yield a prediction performance of $R = 0.9915$, $RMSE = 0.0204$, $VAF = 98.2847$ and $T = 16.37$ s for validation set. From the Fig. 11(b), the narrowest and widest confidence intervals are (0.3563, 0.3689) and (0.7378, 0.7761), respectively; the narrowest and widest prediction intervals are (0.3214, 0.4011) and (0.7145, 0.7994), respectively. This shows that both the prediction and confidence bands are narrow and overlap highly, which indicates the high prediction accuracy of GWO-LSTM for the validation set.

6.3. Results of PSO-LSTM

The hidden layer nodes obtained using PSO is 29 and the learning rate is 0.1475. The results of the prediction of the validation set using PSO-LSTM are shown in Fig. 12. From the Fig. 12(a), the PSO-LSTM yield a prediction performance of $R = 0.9403$, $RMSE = 0.0507$, $VAF = 88.3365$ and $T = 21.31$ s for validation set. From the Fig. 12(b), the narrowest and widest confidence intervals are (0.3541, 0.3867) and (0.6901, 0.7888), respectively; the narrowest and widest prediction intervals are (0.2719, 0.4701) and (0.6299, 0.8489), respectively. This shows that the width of the prediction band is wider than the confidence band, which indicates that the prediction error using PSO-LSTM for the validation set is higher than GWO-LSTM. In addition, the presence of misfit points outside the prediction band also demonstrates that the prediction accuracy of PSO-LSTM is lower than that of GWO-LSTM. In PSO, the learning factors and inertia weights of PSO are referred to some related research results to obtain the optimal optimization performance [63, 74, 75].

6.4. Results of SSA-LSTM

The hidden layer nodes obtained using SSA is 6 and the learning rate is 0.0378. The results of the prediction of the validation set using SSA-LSTM are shown in Fig. 13. From the Fig. 13(a), the SSA-LSTM yield a prediction performance of $R = 0.9117$, $RMSE = 0.0586$, $VAF = 83.0121$ and $T = 15.19$ s for validation data. From the Fig. 13(b), the narrowest and widest confidence intervals are (0.3401, 0.3758) and

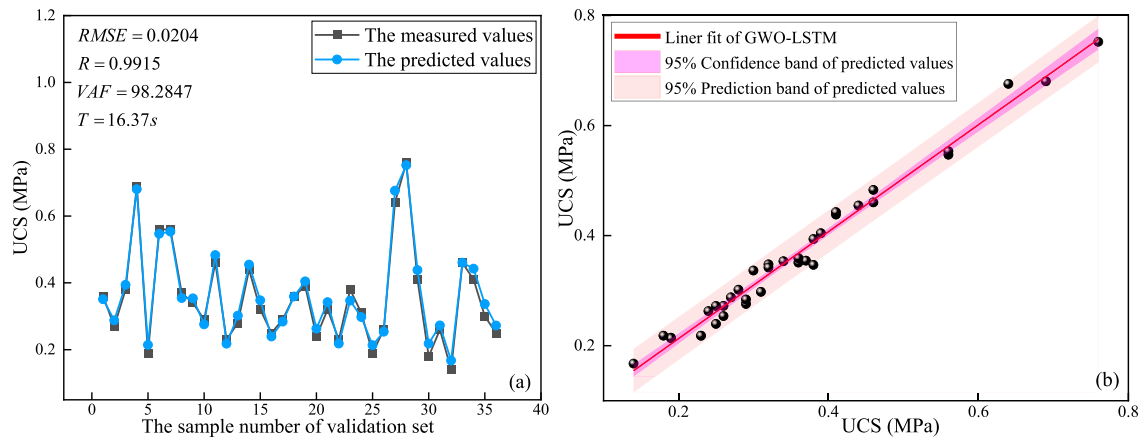


Fig. 11. The prediction performance of the GWO-LSTM (a) the measured and predicted results of GWO-LSTM; (b) the confidence band and prediction band of GWO-LSTM.

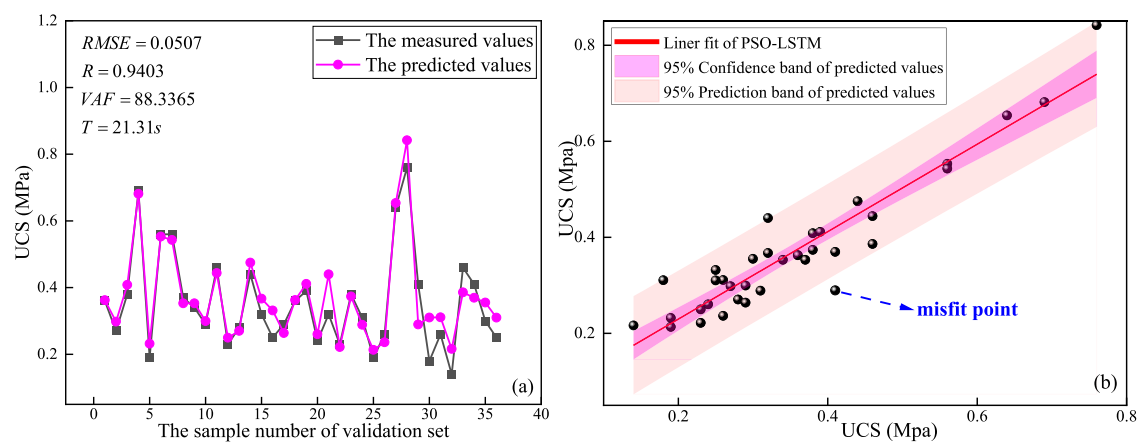


Fig. 12. The prediction performance of the PSO-LSTM (a) the measured and predicted results of PSO-LSTM; (b) the confidence band and prediction band of PSO-LSTM.

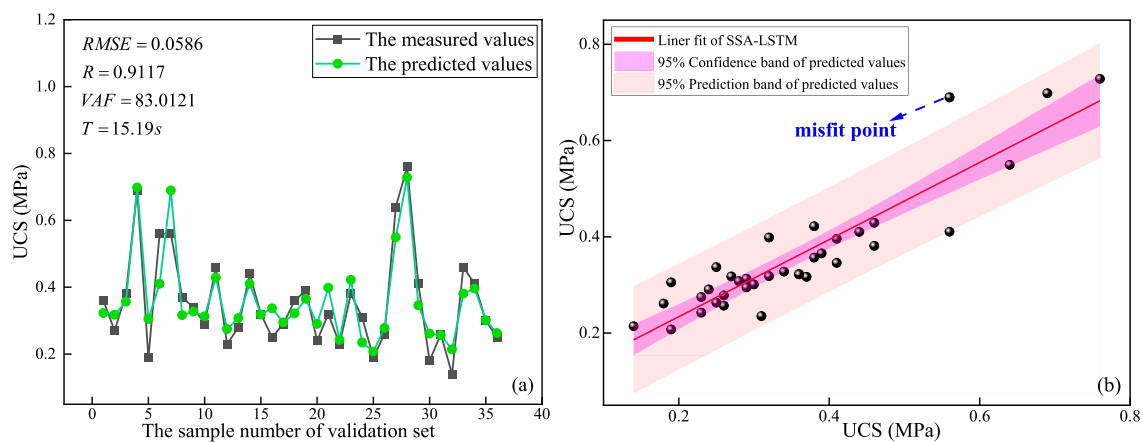


Fig. 13. The prediction performance of the SSA-LSTM (a) the measured and predicted results of SSA-LSTM; (b) the confidence band and prediction band of SSA-LSTM.

(0.6286, 0.7366), respectively; the narrowest and widest prediction intervals are (0.2494, 0.4665) and (0.5627, 0.8026), respectively. This shows that the width of the prediction band is wider than the confidence band, which indicates that the prediction error using SSA-LSTM for the validation set is higher than GWO-LSTM. In addition, the presence of misfit points outside the prediction band also demonstrates that the prediction accuracy of SSA-LSTM is lower than that of GWO-LSTM. In SSA,

the values of ST and PD in SSA are taken with reference to the relevant literature [76].

6.5. Analysis of results

The running time of the 4 improved LSTMs and neural networks in section 5 are shown in Table 7. From Table 7, it can be seen that the

Table 7. The running time of different models.

Model	T	Model	T
BPNN	9.28 s	trial-and-error-LSTM	18.61 s
RBFNN	8.74 s	GWO-LSTM	16.37 s
GRNN	5.46 s	PSO-LSTM	21.31 s
LSTM	13.11 s	SSA-LSTM	22.19 s

running time of LSTM is significantly higher than that of other neural networks and the running time of the improved LSTM is higher than that of the original LSTM. The reasons for this phenomenon are analyzed as follows. First, the core of LSTM is a gating mechanism, which means that the network structure of LSTM is complex and requires many parameters to be calculated. Second, LSTM is a time series model, which means that when processing data the results of the later moment are dependent on the information of the previous moment and cannot be executed in parallel. Lastly, the use of optimization algorithms to improve the LSTM increases the amount of operations, which can increase the running time. Although running time is an important metric in machine learning, the prediction accuracy and error of the model are of more important in this study and in practical engineering applications.

An analysis of the errors distribution between the measured and predicted values of the 4 models is shown in Fig. 14. From the Fig. 14(a), the errors of trial-and-error-LSTM are concentrated between 0.25 and 0.75 with mean and standard deviation of 0.0111 and 0.0644, respectively. From the Fig. 14(b), the errors of GWO-LSTM are concentrated between -0.02 and 0.03 with mean and standard deviation of 0.0084 and 0.0019, respectively. From the Fig. 14(c), it can be seen that the errors of PSO-LSTM are concentrated between -0.025 and 0.075 with mean and standard deviation of 0.0157 and 0.0491, respectively. From the Fig. 14(d), it can be seen that the errors of SSA-LSTM are concentrated between -0.075 and 0.075 with mean and standard deviation of

0.0032 and 0.0593, respectively. The absolute value of the center error of GWO-LSTM is almost 0 (mean = 0.0084) and the minimum value of width value (standard deviation = 0.0019). All of above analysis indicate that the GWO-LSTM has the lowest error between the predicted and measured values and the highest prediction accuracy.

In addition, Taylor diagrams with standard deviation, *R* and RMSE as the core are plotted in Fig. 15. In the Taylor diagram, each model corresponds to a measured point according to the 3 metrics mentioned above. The radial line represents *R*, the solid arc represents standard deviation and the red dashed line represents RMSE in Fig. 15. From the Fig. 15, GWO-LSTM has not only the lowest RMSE and standard deviation but also the highest *R*, which indicates that GWO-LSTM has the highest prediction accuracy among the 4 optimized LSTMs.

According to the no free lunch theorem, the performance of the optimization algorithm is equivalent due to the mutual compensation of all possible functions. This means that one algorithm (Algorithm A) outperforms another algorithm (Algorithm B) on a specific dataset and must be associated with Algorithm A outperforming Algorithm B on another specific dataset. Without certain assumptions about the prior distribution of training set in the feature space, there are as many cases of better performance as there are cases of worse performance, which means that if the algorithm wants to get a positive performance improvement on problems, a negative performance degradation on problems must be paid. In this study, the prediction accuracy of LSTM for UCS is higher than other neural networks but the running time is long. In addition, the predictions of UCS using trial-and-error, GWO, PSO, and SSA optimized LSTM get higher accuracy but they have longer running time. Although GWO-LSTM is the optimal model in this study, this optimal effect exists at the sacrifice of running time. Therefore, it can only be said that GWO-LSTM has excellent performance in this research and it still has much room for improvement.

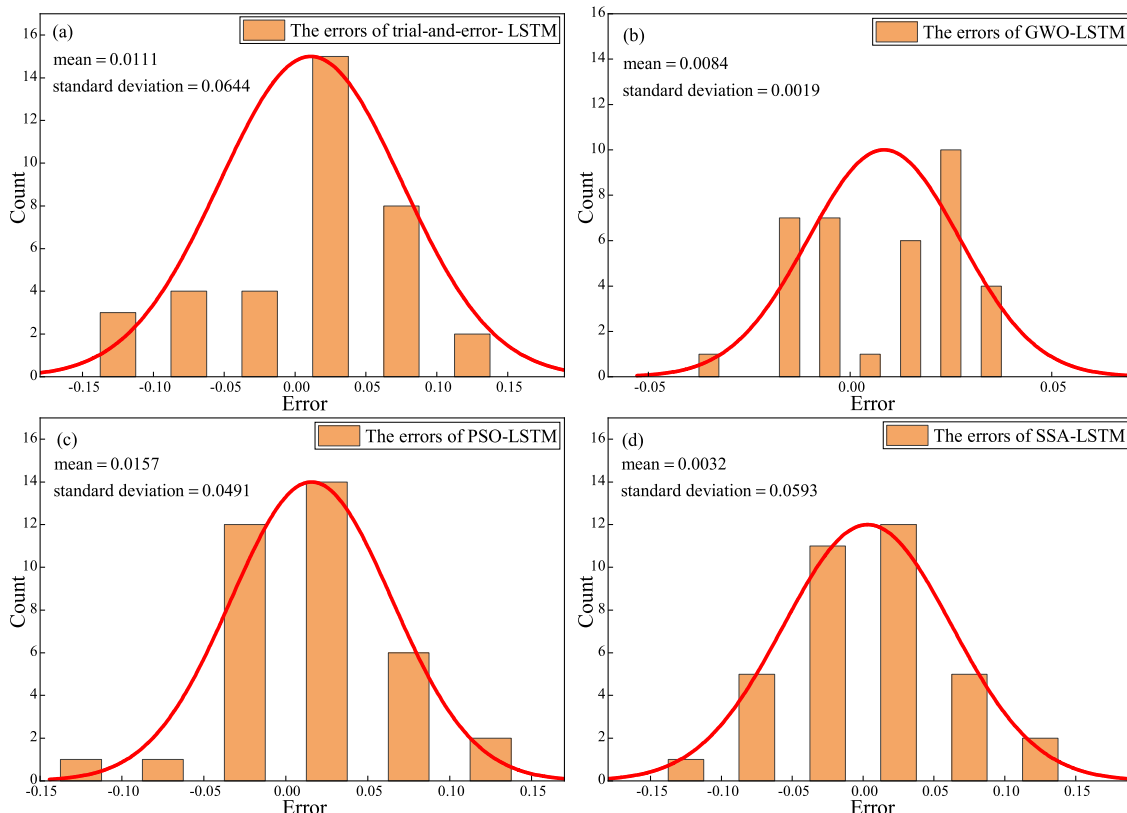


Fig. 14. The comparison of errors distribution from different models (a) trial-and-error-LSTM; (b) GWO-LSTM; (c) PSO-LSTM; (d) SSA-LSTM.

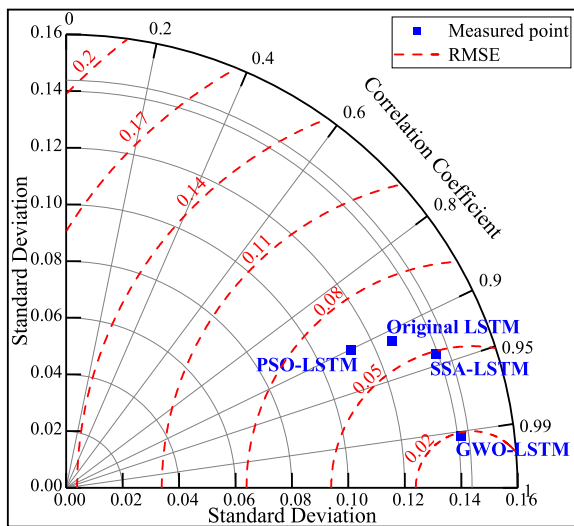


Fig. 15. The Taylor Diagram for performance from different model.

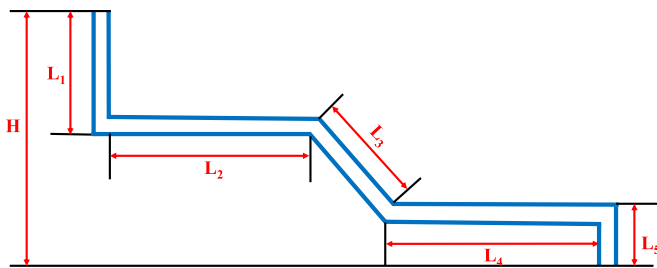


Fig. 16. The schematic representation for stowing gradient.

7. Engineering applications

In this section, the GWO-LSTM constructed in section 6 is used as the core model and combined with filling body of 153 underground approaches for 3 filled stopes to predict the UCS of CPB.

7.1. Engineering background

The transportation with gravity by pipelines is generally used to transport the filling slurry to the mining stope. The stowing gradient is applied to describe the filling range that the gravity system can reach. The stowing gradient characterizes the engineering characteristics of the filling pipeline, which is one of the critical factors for the transportation with gravity of the filling slurry and is essential for the design and production management of the transportation system for the filling slurry. The stowing gradient is calculated and represented as shown in equation (23) and Fig. 16.

$$L = N/H \tag{23}$$

In equation (23), N is the stowing gradient; H is the height difference from start and finish points of filling pipeline; L is the total length of the filling pipeline including the converted length of fittings such as elbows and joints ($L = L_1 + L_2 + L_3 + L_4 + L_5$).

Due to the different stopes and depths of underground mining, the stowing gradients of different filled stopes are variant. If the filling slurry concentration is adjusted blindly under the above conditions, the strength of the filling body would be unstable. The gold mine studied in this paper is mined using the drift fill stopping and cemented filling method. The filling slurry is transported by gravity to the mined-out area and the stowing gradient of each filled stope ranges from 3 to 7.

Filled stope with the range of stowing gradient of 3 to 4.5 are set up as filled stope A; the range of 4.5 to 6 are set up as filled stope B; the range of 6 to 7 are set up as filled stope C. The different filled stopes are zoned according to the stowing gradient as shown in Fig. 17.

7.2. Correction coefficient

Related researches have shown that the actual values of strength for the filling body in the filled stope with the same slurry mix proportion is significantly lower than the laboratory measured values. Although many of UCS prediction studies have been conducted by researchers, the predicted UCS measured in laboratory (UCS_M) rather than the predicted UCS in actual engineering (UCS_A) with engineering guidance. This means that the above studies have limitations in engineering applications. To accurately obtain the UCS of the filled stope and to provide guidance for safe production, a core machine is used to drill the filling body in filled stope and process it into standard specimens. The final UCS of filling body is obtained for different filled stopes and different curing time.

For the error between UCS_M and UCS_A , a correction coefficient (k) is introduced to adjust for this error and is used to guide production. The correction coefficient is defined and calculated by the following equation.

$$k = UCS_M/UCS_A \tag{24}$$

The correction coefficients for the strength of different filled stopes are obtained by comparing UCS_M and UCS_A , which is as shown in Fig. 18. From the Fig. 18(a)-(c), k for 5d, 7d and 28d in filled stope A are 1.63, 1.51 and 1.48, respectively; k for 5d, 7d and 28d in filled stope B are 1.57, 1.46 and 1.41, respectively; k for 5d, 7d and 28d in filled stope C are 1.48, 1.37 and 1.33, respectively. The $k = 1.63$ is the maximum value, which is found in stope A with a curing time of 5d; $k = 1.33$ is the minimum value, which is found in stope C with a curing time of 28d. This indicates that there is still a certain error between UCS_M and UCS_A when using GWO-LSTM for prediction but the distribution of error is regular, which can be seen from the range of distribution of k ($1.33 \leq k \leq 1.63$). On the other hand, this can also demonstrate the rationality of using k to correct error. From the Fig. 18(d), the minimum value of k is obtained in the filled stope C regardless of the curing time, which indicates that the error between UCS_M and UCS_A in filled stope C is lowest. In addition, k of different filled stopes show a decreasing trend when the curing time is extended. This means that with the stowing gradient increases, the k increases gradually for different curing times. The reasons for the phenomenon above are analyzed as follows. The filling slurry in filled stope with high stowing gradient (e.g. filled stope C) is less fluid during gravity transport. To ensure the filling slurry can be transported to the filling stope by gravity, water is added to the stirred tank for the slurry to dilute in engineering practice. This leads to a decrease in filling slurry concentration and causes a decrease in UCS. Therefore, the larger the stowing gradient for filled stope, the larger the error between UCS_M and UCS_A and the larger the k .

7.3. Strength prediction of CPB in filled stope

In the production of mines using the filling mining, the strength of the filling body in the underground is directly determined by the safety of the mining work. Based on the high-precision UCS prediction model constructed in this paper, the strength of filling body in 153 filling approaches in filled stopes filled with a horizontal height from 330 m to 430 m is predicted to provide guidance for the safe mining. The slurry mix proportions for the 153 filling approaches (A:48, B:42 and C:63) are input into GWO-LSTM to obtain the UCS_M . Then UCS_M is divided by k and the UCS_A is obtained. According to safety standard of the mine, 5d UCS must not be lower than 0.3 MPa, 7d UCS must not be than 0.35 MPa and 28d UCS must not be lower than 0.5 MPa. The statistical distribution of the UCS_A is shown in Fig. 19.

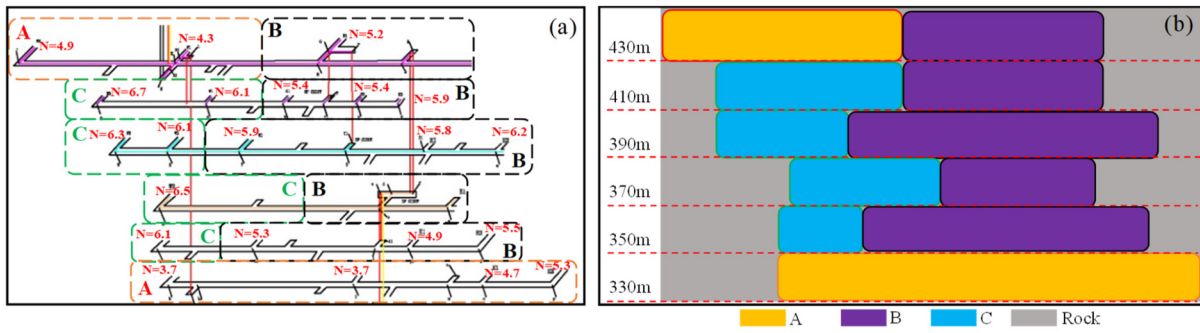


Fig. 17. Filled stope zoning based on stowing gradient (a) the diagram of stowing gradient; (b) the filled stopes divided by different stowing gradient.

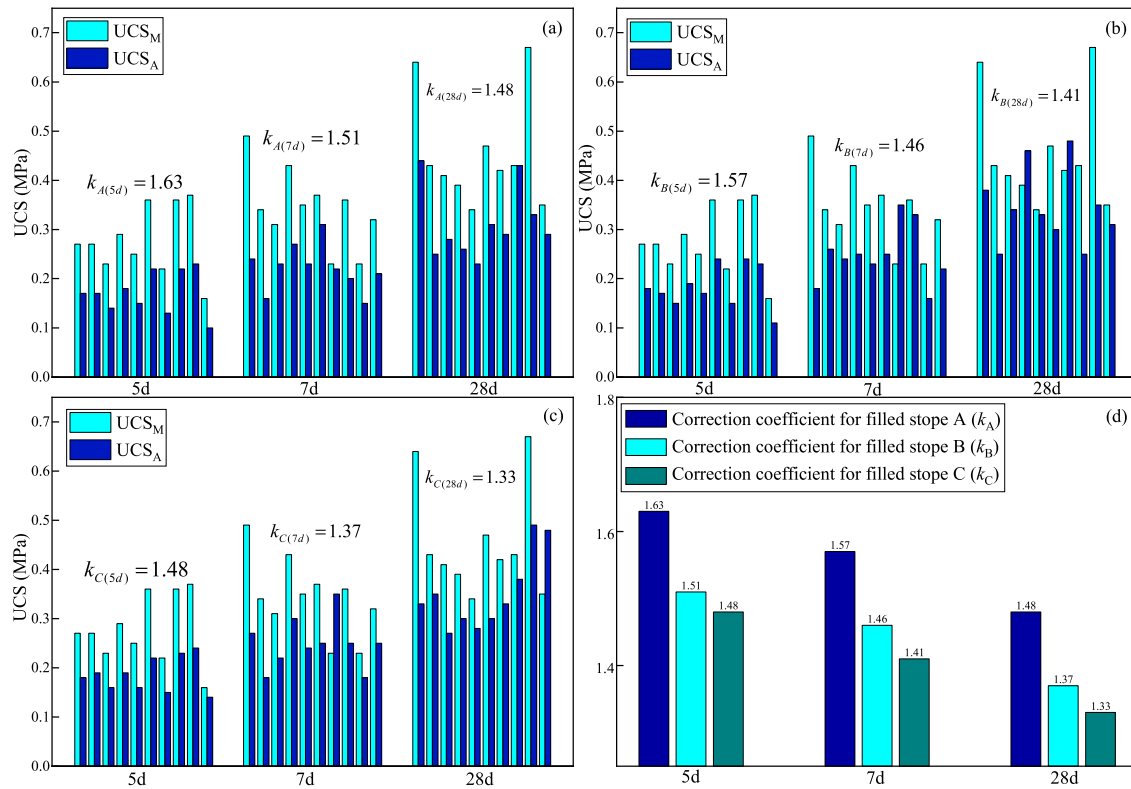


Fig. 18. The correction coefficients in different filled stopes (a) the results of UCS_M and UCS_A in stope A; (b) the results of UCS_M and UCS_A in stope B; (c) the results of UCS_M and UCS_A in stope C; (d) the correction coefficient of different stopes.

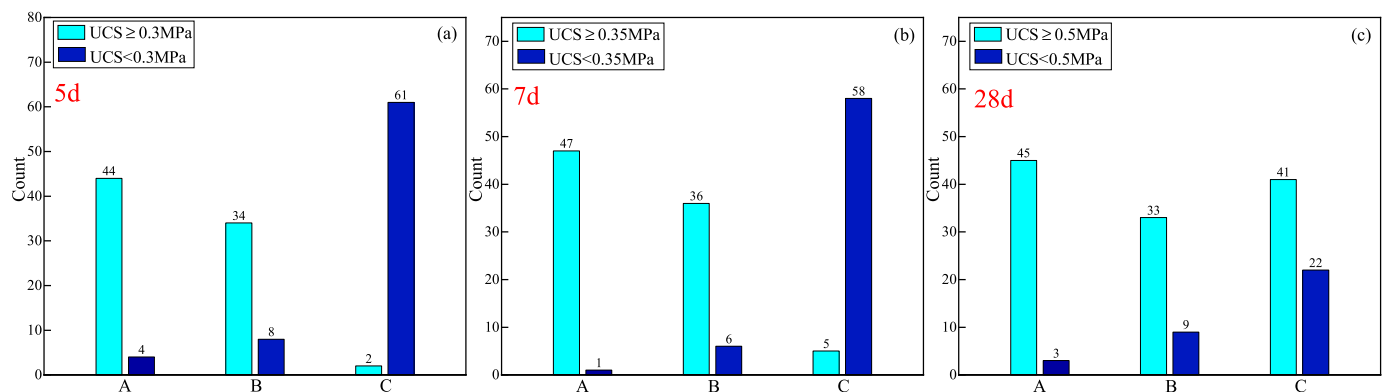


Fig. 19. The statistical distribution of the UCS_A (a) 5d; (b) 7d; (c) 28d.

From the Fig. 19(a), there are 44, 34 and 2 UCS_A at 5d higher than 0.3 MPa in filled stope A, B and C respectively. From the Fig. 19(b), there are 47, 36 and 5 UCS_A at 7d higher than 0.35 MPa in filled stope A, B and C respectively. From the Fig. 19(c), there are 45, 33 and 41 UCS_A at 28d higher than 0.5 MPa in filled stope A, B and C respectively. This shows that the percent of UCS_A compliant with safety standards obtained using GWO-LSTM in combination with k are 97.3%, 80.9% and 65.1% at 5d, 7d and 28d in three filled stope, respectively. In addition, the strength of 3, 9 and 22 approaches in filled stope A, B and C respectively cannot reach the safety standard and need to conduct some safety measures.

Although this study has made some progress in engineering applications, it still has some limitations. First, the dataset composed of UCS of CPB is usually homogeneous and heterogeneous, which makes it more difficult to perceive the working conditions in the engineering applications of the constructed models. Second, the mechanism of hydration reaction and mechanical properties of filling materials is extremely complex and have not been understood and recognized uniformly, which makes the modeling process lack of theoretical knowledge to guide. Lastly, it is difficult for the model to balance real-time and dynamic with limited data sets, which can affect the reliability and robustness of the constructed model. Therefore, there is still room for progress in this study.

8. Conclusion

(1) The UCS database include 180 sets of mix proportion experiments for filling slurry is constructed using tailings as aggregate and cement (PO.42.5) as binder from a gold mine in Guizhou, China. In the above database, the underflow productivity for tailings, slurry concentration, cement content and curing time are set as input and UCS as output. The experimental values of UCS were concentrated between 0.265 MPa and 0.43 MPa (maximum value of 0.96 MPa and minimum value of 0.14 MPa).

(2) Four neural networks, BPNN, RBFNN, GRNN and LSTM, are trained to construct mapping relationships between input and output variables and are used to predict the UCS based on the constructed database. Although RMSE is not the lowest and the running time is higher than the other three neural networks, the LSTM is the optimal model in the prediction of UCS of CPB based on engineering practice. ($R = 0.8974$, $RMSE = 0.0871$, $VAF = 99.9874$, $T = 13.11$ s and Total rank = 11).

(3) The learning rate and the hidden layer nodes in the LSTM are optimized using trial-and-error, PSO, GWO, and SSA to improve the prediction performance of the LSTM. The simulation results show that the prediction performance of the GWO-LSTM (learning rate = 0.0785 and hidden layer nodes = 25) is optimal ($R = 0.9915$, $RMSE = 0.0204$, $VAF = 98.2847$ and $T = 16.37$ s).

(4) The correction coefficient (k) is defined to correct for the errors of UCS_M and UCS_A. The k and GWO-LSTM are combined and successfully predict the strength of the filling body in 153 approaches for 3 filled stopes with different stowing gradient. The results in section 7 demonstrate the rationality of k and the high prediction accuracy of GWO-LSTM for UCS of CPB. This provides guidance for the safety mining of filled stopes.

(5) Future studies will focus on the processing of multi-source heterogeneous data sets in machine learning modeling and the mechanistic modeling of hydration reactions of filling materials.

Declarations

Author contribution statement

Bo Zhang: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Keqing Li, Siqi Zhang, Bin Han: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Yafei Hu: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

Professor Bin Han was supported by the National Key Research and Development Program of China [2018YFC1900603]. Keqing Li was supported by the National Key Research and Development Program of China [2018YFC0604604].

Data availability statement

The authors do not have permission to share data.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

- [1] E. Li, J. Zhou, X. Shi, D. Jahed Armaghani, Z. Yu, X. Chen, P. Huang, Developing a hybrid model of salp swarm algorithm-based support vector machine to predict the strength of fiber-reinforced cemented paste backfill, *Eng. Comput.* 37 (4) (2021) 3519–3540.
- [2] C. Qi, A. Fourie, Cemented paste backfill for mineral tailings management: review and future perspectives, *Miner. Eng.* 144 (2019) 106025.
- [3] S. Zhang, T. Shi, W. Ni, K. Li, W. Gao, K. Wang, Y. Zhang, The mechanism of hydrating and solidifying green mine fill materials using circulating fluidized bed fly ash-slag-based agent, *J. Hazard. Mater.* 415 (2021) 125625.
- [4] W. Sun, H. Wang, K. Hou, Control of waste rock-tailings paste backfill for active mining subsidence areas, *J. Clean. Prod.* 171 (2018) 567–579.
- [5] J. Tian, C. Qi, Y. Sun, Z.M. Yaseen, B.T. Pham, Permeability prediction of porous media using a combination of computational fluid dynamics and hybrid machine learning methods, *Eng. Comput.* 37 (4) (2021) 3455–3471.
- [6] A. Fourie, Preventing catastrophic failures and mitigating environmental impacts of tailings storage facilities, *Proc. Earth Planet. Sci.* 1 (1) (2009) 1067–1071.
- [7] H. Lu, C. Qi, Q. Chen, D. Gan, Z. Xue, Y. Hu, A new procedure for recycling waste tailings as cemented paste backfill to underground stopes and open pits, *J. Clean. Prod.* 188 (2018) 601–612.
- [8] W. Sun, D. Wu, H. Liu, C. Qu, Thermal, mechanical and ultrasonic properties of cemented tailings backfill subjected to microwave radiation, *Constr. Build. Mater.* 313 (2021) 125535.
- [9] J. Qiu, Z. Guo, L. Yang, H. Jiang, Y. Zhao, Effects of packing density and water film thickness on the fluidity behaviour of cemented paste backfill, *Powder Technol.* 359 (2020) 27–35.
- [10] C. Qi, Q. Chen, X. Dong, Q. Zhang, Z.M. Yaseen, Pressure drops of fresh cemented paste backfills through coupled test loop experiments and machine learning techniques, *Powder Technol.* 361 (2020) 748–758.
- [11] X. Zhao, H. Zhang, W. Zhu, Fracture evolution around pre-existing cylindrical cavities in brittle rocks under uniaxial compression, *Trans. Nonferr. Met. Soc. China* 24 (3) (2014) 806–815.
- [12] C. Wang, X. Hou, Z. Liao, Z. Chen, Z. Lu, Experimental investigation of predicting coal failure using acoustic emission energy and load-unload response ratio theory, *J. Appl. Geophys.* 161 (2019) 76–83.
- [13] Y. He, Q. Zhang, Q. Chen, J. Bian, C. Qi, Q. Kang, Y. Feng, Mechanical and environmental characteristics of cemented paste backfill containing lithium slag-blended binder, *Constr. Build. Mater.* 271 (2021) 121567.
- [14] D. Wu, W. Sun, S. Liu, C. Qu, Effect of microwave heating on thermo-mechanical behavior of cemented tailings backfill, *Constr. Build. Mater.* 266 (2021) 121180.
- [15] H.A.C.K. Hettiarachchi, W.K. Mamppearachchi, Effect of vibration frequency, size ratio and large particle volume fraction on packing density of binary spherical mixtures, *Powder Technol.* 336 (2018) 150–160.
- [16] Y. Zhang, W. Gao, W. Ni, S. Zhang, Y. Li, K. Wang, X. Huang, P. Fu, W. Hu, Influence of calcium hydroxide addition on arsenic leaching and solidification/stabilisation behaviour of metallurgical-slag-based green mining fill, *J. Hazard. Mater.* 390 (2020) 122161.

- [17] J. Qiu, Z. Guo, L. Yang, H. Jiang, Y. Zhao, Effect of tailings fineness on flow, strength, ultrasonic and microstructure characteristics of cemented paste backfill, *Constr. Build. Mater.* 263 (2020) 120645.
- [18] Q. Zhang, G. Hu, X. Wang, Hydraulic calculation of gravity transportation pipeline system for backfill slurry, *J. Cent. South Univ. Technol.* 15 (5) (2008) 645–649.
- [19] Q. Chen, Q. Zhang, X. Wang, C. Xiao, Q. Hu, A hydraulic gradient model of paste-like crude tailings backfill slurry transported by a pipeline system, *Environ. Earth Sci.* 75 (14) (2016) 1099.
- [20] Z. Yu, X. Shi, X. Chen, J. Zhou, C. Qi, Q. Chen, D. Rao, Artificial intelligence model for studying unconfined compressive performance of fiber-reinforced cemented paste backfill, *Trans. Nonferrous Met. Soc. China* 31 (4) (2021) 1087–1102.
- [21] B. Yan, W. Zhu, C. Hou, E. Yilmaz, M. Saadat, Characterization of early age behavior of cemented paste backfill through the magnitude and frequency spectrum of ultrasonic P-wave, *Constr. Build. Mater.* 249 (2020) 118733.
- [22] H. Jiang, M. Fall, Y. Li, J. Han, An experimental study on compressive behaviour of cemented rockfill, *Constr. Build. Mater.* 213 (2019) 10–19.
- [23] Y. Shi, L. Cheng, M. Tao, S. Tong, X. Yao, Y. Liu, Using modified quartz sand for phosphate pollution control in cemented phosphogypsum (PG) backfill, *J. Clean. Prod.* 283 (2021) 124652.
- [24] J. Wu, H. Jing, Q. Yin, B. Meng, G. Han, Strength and ultrasonic properties of cemented waste rock backfill considering confining pressure, dosage and particle size effects, *Constr. Build. Mater.* 242 (2020) 118132.
- [25] A. Mohammed, W. Mahmood, K. Ghafor, TGA, rheological properties with maximum shear stress and compressive strength of cement-based grout modified with polycarboxylate polymers, *Constr. Build. Mater.* 235 (2020) 117534.
- [26] C. Qi, A. Fourie, Q. Chen, Q. Zhang, A strength prediction model using artificial intelligence for recycling waste tailings as cemented paste backfill, *J. Clean. Prod.* 183 (2018) 566–578.
- [27] C. Qi, Q. Chen, A. Fourie, Q. Zhang, An intelligent modelling framework for mechanical properties of cemented paste backfill, *Miner. Eng.* 123 (2018) 16–27.
- [28] C. Qi, A. Fourie, Q. Chen, Neural network and particle swarm optimization for predicting the unconfined compressive strength of cemented paste backfill, *Constr. Build. Mater.* 159 (2018) 473–478.
- [29] R.R.N. Sivakugan, Prediction of paste backfill performance using artificial neural networks, in: *Proceedings of the 16th International Conference on Soil Mechanics and Geotechnical Engineering*, 2005, pp. 1107–1110.
- [30] C. Xiao, X. Wang, Q. Chen, F. Bin, Y. Wang, W. Wei, Strength investigation of the silt-based cemented paste backfill using lab experiments and deep neural network, *Adv. Mater. Sci. Eng.* 2020 (2020) 1–12.
- [31] L. Orejarena, M. Fall, The use of artificial neural networks to predict the effect of sulphate attack on the strength of cemented paste backfill, *Bull. Eng. Geol. Environ.* 69 (4) (2010) 659–670.
- [32] A. Altan, S. Karasu, E. Zio, A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer, *Appl. Soft Comput.* 100 (2021) 106996.
- [33] L. Xing, W. Liu, A data fusion powered bi-directional long short term memory model for predicting multi-lane short term traffic flow, *IEEE Trans. Intell. Transp. Syst.* (2021) 1–10.
- [34] W. Zhou, M. Liu, Y. Zheng, Y. Dai, Online error correction method of PMU data based on LSTM model and Kalman filter, in: *2020 12th IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 2020, pp. 1–6.
- [35] K. Kumar, M.T.U. Haider, Enhanced prediction of intra-day stock market using meta-heuristic optimization on RNN-LSTM network, *New Gener. Comput.* 39 (1) (2021) 231–272.
- [36] N. Zhou, J. Zhang, S. Ouyang, X. Deng, C. Dong, E. Du, Feasibility study and performance optimization of sand-based cemented paste backfill materials, *J. Clean. Prod.* 259 (2020) 120798.
- [37] W. Cao, W. Zhu, W. Wang, Y. Demazeau, C. Zhang, A deep coupled LSTM approach for USD/CNY exchange rate forecasting, *IEEE Intell. Syst.* 35 (2) (2020) 43–53.
- [38] C. Qi, Q. Chen, A. Fourie, X. Tang, Q. Zhang, X. Dong, Y. Feng, Constitutive modelling of cemented paste backfill: a data-mining approach, *Constr. Build. Mater.* 197 (2019) 262–270.
- [39] D. Wu, T. Deng, R. Zhao, A coupled THMC modeling application of cemented coal gangue-fly ash backfill, *Constr. Build. Mater.* 158 (2018) 326–336.
- [40] Y. Wang, Y. Cao, L. Cui, Z. Si, H. Wang, Effect of external sulfate attack on the mechanical behavior of cemented paste backfill, *Constr. Build. Mater.* 263 (2020) 120968.
- [41] W. Zhou, G. Sun, Z. Yang, H. Wang, L. Fang, J. Wang, BP neural network based reconstruction method for radiation field applications, *Nucl. Eng. Des.* 380 (2021) 111228.
- [42] H. Xue, H. Cui, Research on image restoration algorithms based on BP neural network, *J. Vis. Commun. Image Represent.* 59 (2019) 204–209.
- [43] W. Liang, G. Wang, X. Ning, J. Zhang, Y. Li, C. Jiang, N. Zhang, Application of BP neural network to the prediction of coal ash melting characteristic temperature, *Fuel* 260 (2020) 116324.
- [44] P. Reiner, B.M. Wilamowski, Efficient incremental construction of RBF networks using quasi-gradient method, *Neurocomputing* 150 (2015) 349–356.
- [45] X. Xu, D. Shan, S. Li, T. Sun, P. Xiao, J. Fan, Multi-label learning method based on ML-RBF and Laplacian ELM, *Neurocomputing* 331 (2019) 213–219.
- [46] N. Zhang, S. Ding, J. Zhang, Multi layer ELM-RBF for multi-label learning, *Appl. Soft Comput.* 43 (2016) 535–545.
- [47] X. Xiao, X. Zhao, Retraction note: evaluation of mangrove wetland potential based on convolutional neural network and development of film and television cultural creative industry, *Arab. J. Geosci.* 14 (21) (2021) 2247.
- [48] H. Rui, C. Gao, Retraction note to: neural network-based urban green vegetation coverage detection and smart home system optimization, *Arab. J. Geosci.* 14 (22) (2021) 2424.
- [49] H. Khatter, A. Kumar Ahlawat, Correction to: an intelligent personalized web blog searching technique using fuzzy-based feedback recurrent neural network, *Soft Comput.* 24 (12) (2020) 9335.
- [50] S. Rohit, S. Chakravarthy, A convolutional neural network model of the neural responses of inferotemporal cortex to complex visual objects, *BMC Neurosci.* 12 (1) (2011) P35.
- [51] Y. Cheng, Retraction note: rainfall trend and supply chain network management in mountainous areas based on dynamic neural network, *Arab. J. Geosci.* 14 (24) (2021) 2779.
- [52] M. Hajiaghayi, E. Vahedi, Code failure prediction and pattern extraction using LSTM networks, in: *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*, 2019, pp. 55–62.
- [53] Y. Guo, Stock price prediction based on LSTM neural network: the effectiveness of news sentiment analysis, in: *2020 2nd International Conference on Economic Management and Model Engineering (ICEMME)*, 2020, pp. 1018–1024.
- [54] S.C. van de Leemput, M. Prokop, B. van Ginneken, R. Manniesing, Stacked bidirectional convolutional LSTMs for deriving 3D non-contrast CT from spatiotemporal 4D CT, *IEEE Trans. Med. Imaging* 39 (4) (2020) 985–996.
- [55] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (8) (1997) 1735–1780.
- [56] P. Siddavaatam, R. Sedaghat, Grey wolf optimizer driven design space exploration: a novel framework for multi-objective trade-off in architectural synthesis, *Swarm Evol. Comput.* 49 (2019) 44–61.
- [57] S. Maroufpoor, E. Maroufpoor, O. Bozorg-Haddad, J. Shiri, Z. Mundher Yaseen, Soil moisture simulation using hybrid artificial intelligent model: hybridization of adaptive neuro fuzzy inference system with grey wolf optimizer algorithm, *J. Hydrol.* 575 (2019) 544–556.
- [58] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey wolf optimizer, *Adv. Eng. Softw.* 69 (2014) 46–61.
- [59] G. Deepa, G.L.R. Mary, A. Karthikeyan, P. Rajalakshmi, K. Hemavathi, M. Dhanisri, Detection of brain tumor using modified particle swarm optimization (MPSO) segmentation via Haralick features extraction and subsequent classification by KNN algorithm, *Mater. Today Proc.* (2021).
- [60] N. Wang, J. Liu, J. Lu, X. Zeng, X. Zhao, Low-delay layout planning based on improved particle swarm optimization algorithm in 5G optical fronthaul network, *Opt. Fiber Technol.* 67 (2021) 102736.
- [61] J. Peng, Y. Li, H. Kang, Y. Shen, X. Sun, Q. Chen, Impact of population topology on particle swarm optimization and its variants: an information propagation perspective, *Swarm Evol. Comput.* (2021) 100990.
- [62] J. Xue, B. Shen, A novel swarm intelligence optimization approach: sparrow search algorithm, *Syst. Sci. Control Eng.* 8 (1) (2020) 22–34.
- [63] H. Zhang, Z. Peng, J. Tang, M. Dong, K. Wang, W. Li, A multi-layer extreme learning machine refined by sparrow search algorithm and weighted mean filter for short-term multi-step wind speed forecasting, *Sustain. Energy Technol. Assess.* 50 (2022) 101698.
- [64] S. Zhou, H. Xie, C. Zhang, Y. Hua, W. Zhang, Q. Chen, G. Gu, X. Sui, Wavefront-shaping focusing based on a modified sparrow search algorithm, *Optik* 244 (2021) 167516.
- [65] P. Kathirolu, K. Selvadurai, Energy efficient cluster head selection using improved sparrow search algorithm in wireless sensor networks, *J. King Saud Univ. Comput. Inf. Sci.* (2021).
- [66] K. Zorlu, C. Gokceoglu, F. Ocakoglu, H.A. Nefeslioglu, S. Acikalin, Prediction of uniaxial compressive strength of sandstones using petrography-based models, *Eng. Geol.* 96 (3–4) (2008) 141–158.
- [67] K. Borna, R. Ghanbari, Hierarchical LSTM network for text classification, *SN Appl. Sci.* 1 (9) (2019) 1124.
- [68] J. Song, S. Tang, J. Xiao, F. Wu, Z.M. Zhang, LSTM-in-LSTM for generating long descriptions of images, *Comput. Vis. Media* 2 (4) (2016) 379–388.
- [69] X. Yuan, C. Chen, X. Lei, Y. Yuan, R. Muhammad Adnan, Monthly runoff forecasting based on LSTM-ALO model, *Stoch. Environ. Res. Risk Assess.* 32 (8) (2018) 2199–2212.
- [70] V. Rybalkin, C. Sudarshan, C. Weis, J. Lappas, N. Wehn, L. Cheng, Correction to: efficient hardware architectures for 1D- and MD-LSTM networks, *J. Signal Process. Syst.* 93 (12) (2021) 1467.
- [71] L. Wang, X. Zhong, S. Wang, Y. Liu, Correction to: ncDLRES: a novel method for non-coding RNAs family prediction based on dynamic LSTM and ResNet, *BMC Bioinform.* 22 (1) (2021) 583.
- [72] S. Rezaei, J. Behnamian, Benders decomposition-based particle swarm optimization for competitive supply networks with a sustainable multi-agent platform and virtual alliances, *Appl. Soft Comput.* 114 (2022) 107985.
- [73] D. Ben Hmamou, M. Elyaqouti, E. Arjdal, J. Chaoufi, D. Saadaoui, S. Lidaighbi, R. Aqel, Particle swarm optimization approach to determine all parameters of the photovoltaic cell, *Mater. Today Proc.* (2021).

- [74] R. Kumar, Fuzzy particle swarm optimization control algorithm implementation in photovoltaic integrated shunt active power filter for power quality improvement using hardware-in-the-loop, *Sustain. Energy Technol. Assess.* 50 (2022) 101820.
- [75] J. Ma, Z. Hao, W. Sun, Enhancing sparrow search algorithm via multi-strategies for continuous optimization problems, *Inf. Process. Manag.* 59 (2) (2022) 102854.
- [76] X.U. Hui, C. Guangbin, Z. Shengxiu, Y. Xiaogang, H. Mingzhe, Hypersonic reentry trajectory optimization by using improved sparrow search algorithm and control parametrization method, *Adv. Space Res.* (2021).