Computer Modeling in Engineering & Sciences

Tech Science Press

Doi:10.32604/cmes.2025.064179

ARTICLE

SSA-LSTM-Multi-Head Attention Modelling Approach for Prediction of Coal Dust Maximum Explosion Pressure Based on the Synergistic Effect of Particle Size and Concentration

Yongli Liu[1,2], Weihao Li[1,2,*], Haitao Wang[1,2,3] and Taoren Du[4]

[1]Institute of Interdisciplinary Research on Intelligent Mines, Heilongjiang University of Science and Technology, Harbin, 150022, China
[2]School of Mining Engineering, Heilongjiang University of Science and Technology, Harbin, 150022, China
[3]School of Resources and Engineering Department, Heilongjiang University of Technology, Jixi, 158100, China
[4]Heilongjiang Longmei Jixi Mining Co., Ltd., Xinfa Coal Mine, Jixi, 158199, China

*Corresponding Author: Weihao Li. Email: lwh158619@126.com

Received: 07 February 2025; Accepted: 17 April 2025; Published: 30 May 2025

ABSTRACT: Coal dust explosions are severe safety accidents in coal mine production, posing significant threats to life and property. Predicting the maximum explosion pressure (P_m) of coal dust using deep learning models can effectively assess potential risks and provide a scientific basis for preventing coal dust explosions. In this study, a 20-L explosion sphere apparatus was used to test the maximum explosion pressure of coal dust under seven different particle sizes and ten mass concentrations (C_{dust}), resulting in a dataset of 70 experimental groups. Through Spearman correlation analysis and random forest feature selection methods, particle size (D_{10}, D_{20}, D_{50}) and mass concentration (C_{dust}) were identified as critical feature parameters from the ten initial parameters of the coal dust samples. Based on this, a hybrid Long Short-Term Memory (LSTM) network model incorporating a Multi-Head Attention Mechanism and the Sparrow Search Algorithm (SSA) was proposed to predict the maximum explosion pressure of coal dust. The results demonstrate that the SSA-LSTM-Multi-Head Attention model excels in predicting the maximum explosion pressure of coal dust. The four evaluation metrics indicate that the model achieved a coefficient of determination (R^2), root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) of 0.9841, 0.0030, 0.0074, and 0.0049, respectively, in the training set. In the testing set, these values were 0.9743, 0.0087, 0.0108, and 0.0069, respectively. Compared to artificial neural networks (ANN), random forest (RF), support vector machines (SVM), particle swarm optimized-SVM (PSO-SVM) neural networks, and the traditional single-model LSTM, the SSA-LSTM-Multi-Head Attention model demonstrated superior generalization capability and prediction accuracy. The findings of this study not only advance the application of deep learning in coal dust explosion prediction but also provide robust technical support for the prevention and risk assessment of coal dust explosions.

KEYWORDS: Coal dust explosion; deep learning; maximum explosion pressure; predictive model; SSA-LSTM; multi-head attention mechanism

1 Introduction

In the coal production process, coal dust, as one of the primary hazard sources in underground mines, poses a significant threat to both the safety of coal mining operations and economic efficiency. Hertzberg [1] demonstrated through experimental research that the combustion and explosion of coal dust are caused by the release of flammable gases (volatiles) when coal dust is heated. These gases mix with air to form a

combustible gas mixture, which is then ignited by a high-temperature heat source. A common hazard source of coal dust disasters is the dust cloud formed by coal dust particles suspended in the air. When exposed to high-temperature heat sources such as blasting flames, mechanical friction sparks, or electrical sparks, coal dust explosion accidents may occur [2]. Yuan et al. [3] analyzed hundreds of dust explosion incidents and found that the number of dust explosion accidents worldwide has decreased over time; however, this trend does not apply to China. On 12 January 2019, a roof collapse at the Baiji Coal Mine in Shaanxi Province, China triggered a coal dust explosion, resulting in 21 fatalities. On 27 September 2021, a ventilation system failure at the Songzao Coal Mine in Chongqing, China caused a dust explosion, leading to 10 deaths. On 16 July 2022, an equipment malfunction at the Tashan Coal Mine of the Datong Coal Mine Group in Shanxi Province, China caused a coal dust cloud explosion, resulting in 8 fatalities. Therefore, studying the influence of coal dust's intrinsic factors on explosion characteristics and conducting effective risk assessments are crucial for the safe production of coal mines.

One of the key parameters in the coal dust explosion process is the maximum explosion pressure (P_m), which serves as a critical indicator of the severity of dust explosions. By measuring and analyzing P_m, a better understanding of the characteristics of dust explosions can be achieved, facilitating the development of effective explosion prevention measures. P_m is primarily influenced by factors such as humidity, moisture content, ambient temperature, particle size, and dust concentration. In this study, dust concentration and particle size were selected as independent variables to investigate their effects on P_m. Previous research has extensively explored the influence of these factors on dust explosion characteristics. Tan et al. [4] conducted experiments showing that, at the same concentration, the explosion pressure increases as the coal dust particle size decreases. Azam et al. [5] found that finer coal dust, due to its larger specific surface area and more significant inter-particle heat conduction, requires the addition of more rock dust to suppress explosions. Additionally, Yang et al. [6] discovered that the explosion intensity P_m value of coal dust increases and then decreases with increasing mass concentration. This phenomenon may be due to the restricted flame propagation at low concentrations, resulting in a lower explosion intensity. As the concentration increases, the explosion intensity increases. However, at high concentrations, the limited availability of oxygen leads to a decrease in explosion intensity.

Prevention and suppression of dust explosions are two core measures to ensure the safety of coal mine production. To effectively prevent dust explosions, researchers employ various measurement methods to determine key safety parameters and assess the potential risks of dust explosions based on these parameters. To this end, researchers have conducted extensive experimental studies and utilized artificial neural network technology to develop models for predicting the characteristics of combustible materials in fire and explosion environments. Qi et al. [7] developed a machine learning model to predict the spontaneous combustion temperature of boreholes by combining Random Forest (RF) with the Hunger Games Search optimization algorithm (HGS). The results demonstrated that the HGS-RF hybrid model exhibited the best performance. Lei et al. [8] compared the accuracy of RF and Support Vector Machine (SVM) models in predicting coal spontaneous combustion and found that RF provided accurate predictions without requiring specific parameter settings, making it more suitable for practical applications. Shankar et al. [9] employed the Extreme Gradient Boosting (XGB) model for predicting the susceptibility of coal seams to spontaneous combustion, and the results indicated that this model outperformed the other four methods evaluated. Prasanjit et al. [10] developed a hybrid framework model that utilizes t-distributed Stochastic Neighbor Embedding (t-SNE) for dimensionality reduction and Variational Autoencoder (VAE) for gas feature extraction, combined with a Bidirectional Long Short-Term Memory (bi-LSTM) network for prediction. This model demonstrated lower Mean Squared Error (MSE). Borhani et al. [11] predicted the auto-ignition temperatures of 813 hydrocarbons using a Genetic Algorithm-optimized Multiple Linear Regression (GA-MLR) model and an Artificial Neural

Network (ANN) model. The results indicated that the ANN model provided more accurate predictions and was more convenient for practical applications. Liu et al. [12] employed Principal Component Analysis (PCA) combined with a Backpropagation (BP) Neural Network to predict the flame propagation characteristics of coal dust explosions. The experimental results demonstrated that PCA effectively enhanced the prediction accuracy of the BP neural network. Qi et al. [13] measured the explosion characteristic parameters of four different coal dust and gas mixtures in a standard 20-L explosion vessel and established an effective prediction method. The results indicated that the Bartknecht model exhibited certain applicability for coal dust and gas mixtures. Khan et al. [14] conducted explosibility tests on coal samples from three different regions of Khyber-Pakhtunkhwa, Pakistan, using a 1.2-L Hartmann apparatus. They investigated the factors triggering coal dust explosions and utilized the RF model to predict the maximum explosion pressure. The model's prediction accuracy was relatively low, with experimental results showing an R^2 value of 0.89.

The aforementioned studies indicate that prediction models based on artificial neural networks possess strong theoretical support and practical relevance in the prevention and control of fires involving combustible materials and dust explosions. However, conducting P_m testing not only requires significant time and cost but also entails safety risks. Therefore, establishing an accurate prediction model for coal dust P_m is of great significance for the prevention of coal dust explosions and risk assessment in coal mines. In this study, a 20-L explosion spherical vessel was used to test the P_m values of 7 coal dust samples (C1–C7) at 10 different mass concentrations. Meanwhile, based on the coal dust explosion experimental data, a hybrid modeling method combining Long Short-Term Memory (LSTM) and Multi-Head Attention Mechanism was proposed. By integrating the advantages of both approaches, a P_m prediction model for coal dust is constructed. In this model, LSTM controls the flow of input data through gating mechanisms and outputs hidden states containing historical information. The Multi-Head Attention module receives the hidden state sequence output by the LSTM and utilizes the global attention mechanism and multi-head parallel computing capabilities to capture the dependencies between different positions of the input sequence. To further improve the model's performance, the Sparrow Search Algorithm (SSA) is used to optimize the hyperparameters, and the model's predictive performance is compared with existing machine learning models.

2 Experimental Method and Materials

2.1 Samples of Coal Dust

The coal samples used in this study were obtained from the Shuangyang Coal Mine in Shuangyashan City, China. The coal samples were ground into seven different particle size distributions, and the particle size of the sieved coal dust was measured using a laser particle size analyzer. Each coal sample was classified into ten different mass concentrations (2.5 g, 5 g, 7.5 g, 10 g, 12.5 g, 15 g, 17.5 g, 20 g, 22.5 g, and 25 g) for experimental investigation of the P_m of coal dust. Prior to the experiments, the samples were dried at 50°C for at least 12 h, while the laboratory ambient temperature was maintained at approximately 27°C. The particle size distribution of the seven coal dust samples (C1–C7) is shown in Table 1. The mass concentration of each dust sample was calculated according to Eq. (1) [15].

$$C_{dust=}\frac{m}{V} \tag{1}$$

where, C_{dust} represents the mass concentration of coal dust, m is the mass of the coal dust, and V is the volume of the 20-L spherical explosion chamber (0.02 m^3). The calculated coal dust mass concentration values are 125.0, 250.0, 375.0, 500.0, 625.0, 750.0, 875.0, 1000.0, 1125.0, and 1250.0 g/m^3.

Table 1: Measurement results of particle size parameters of coal dust samples

Particle size parameters	C1	C2	C3	C4	C5	C6	C7
D_{10} (μm)	6.78	4.62	4.29	3.97	2.12	1.74	1.33
D_{20} (μm)	15.99	11.13	9.29	7.75	4.50	3.75	3.13
D_{30} (μm)	26.95	22.98	15.99	15.99	7.75	5.39	4.50
D_{40} (μm)	36.58	33.02	22.98	19.17	11.13	7.47	6.46
D_{50} (μm)	44.59	40.34	33.49	30.33	17.44	11.47	8.70
D_{60} (μm)	51.87	47.44	39.58	39.58	22.98	15.99	13.34
D_{70} (μm)	59.31	52.15	47.44	47.44	27.55	22.98	19.17
D_{80} (μm)	65.17	56.87	55.41	56.87	39.58	27.55	27.55
D_{90} (μm)	77.93	67.43	63.12	64.16	45.73	36.60	33.81

2.2 Experimental Conditions and Methods

The experiment was conducted using a 20-L dust explosion experimental apparatus (HY16426D) manufactured by Jilin Hongyuan Scientific Instruments Co., Ltd. (Jilin, China). The P_m values were measured in accordance with the GB/T 16426-1996 standard [16]. The experimental system configuration of the apparatus is shown in Fig. 1. The 20-L explosion experimental apparatus primarily consists of an explosion tank, a control system, a dust dispersion system, an ignition system, and a data acquisition system. The control system is responsible for automatically initiating dust dispersion and ignition, as well as setting the ignition delay time. The nozzle of the dust dispersion system adopts an outwardly flared blade design. During the dust dispersion preparation process, the tank is first evacuated to an absolute pressure of 0.04 MPa, and then the dust storage is pressurized to an absolute pressure of 2.1 MPa. The control system then releases compressed air to disperse the dust within the dust storage, ensuring that the coal dust is ignited under normal atmospheric pressure. The ignition system consists of an electrode spark rod and a chemical ignition head. The electrode current ignites the chemical ignition head, generating 5 kJ of ignition energy to ignite the dust cloud. The chemical ignition head is composed of metallic zirconium, barium nitrate, and barium peroxide in a mass ratio of 4:3:3, with a total mass of 0.48 g. To ensure experimental reproducibility, all coal sample tests are conducted using the same ignition head and following the same procedure. The ignition delay time is set to 60 ms to ensure sufficient turbulence within the tank, allowing the dust to be uniformly dispersed at the time of ignition. The data acquisition system primarily relies on pressure transducers installed on the tank wall. When an explosion occurs, the system software collects the time-pressure change curve within 0.2 ms and transmits the data to the computer.

To ensure the reliability of the experimental data, coal dust samples C1–C7 were tested at least three times, and the average values were taken. The experimental results are shown in Table 2. The pressure variation within the spherical tank during the experiment is shown in Fig. 2. After the coal dust is released from the dust storage, the pressure in the tank gradually rises to the normal atmospheric pressure range. After the ignition delay, the dust explodes, and the pressure curve sharply increases to the maximum value. As the coal dust combustion in the explosion tank completes, the reaction terminates, and the pressure curve shows a decreasing trend. The explosion test data includes the maximum pressure (P_m) reached within the tank during a single deflagration process, as well as the ignition delay time (t_d).

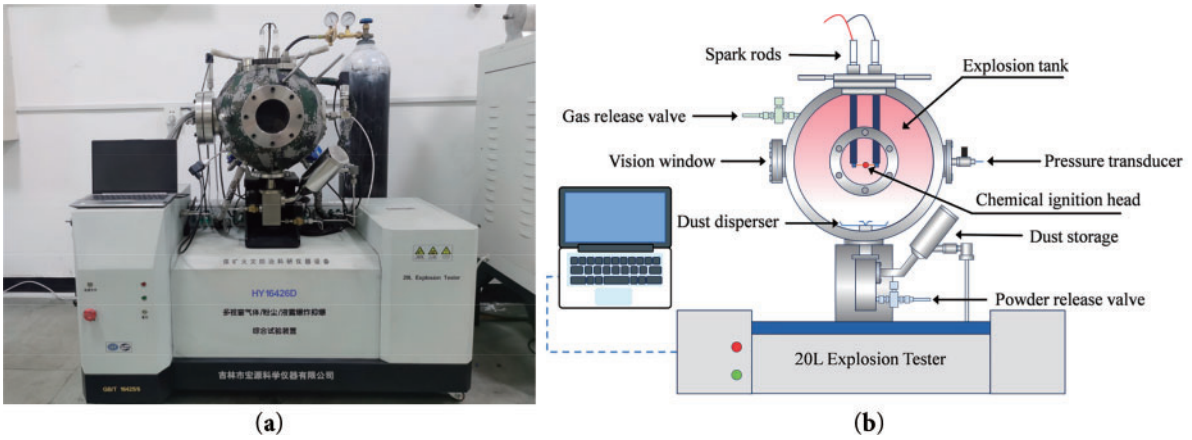

Figure 1: 20-L dust explosion experimental apparatus: (**a**) Physical diagram of the experimental apparatus; (**b**) Structure of the experimental apparatus

Table 2: Explosion experiment results of 7 kinds of coal dust samples

No.	Coal powder sample	Concentration	P_m	No.	Coal powder sample	Concentration	P_m
		g/m^3	MPa			g/m^3	MPa
1	C1	125	0.483	36	C1	750	0.661
2	C2	125	0.512	37	C2	750	0.673
3	C3	125	0.526	38	C3	750	0.683
4	C4	125	0.659	39	C4	750	0.702
5	C5	125	0.678	40	C5	750	0.716
6	C6	125	0.683	41	C6	750	0.739
7	C7	125	0.693	42	C7	750	0.754
8	C1	250	0.538	43	C1	875	0.657
9	C2	250	0.568	44	C2	875	0.669
10	C3	250	0.598	45	C3	875	0.675
11	C4	250	0.669	46	C4	875	0.694
12	C5	250	0.683	47	C5	875	0.707
13	C6	250	0.696	48	C6	875	0.724
14	C7	250	0.709	49	C7	875	0.739
15	C1	375	0.583	50	C1	1000	0.646
16	C2	375	0.631	51	C2	1000	0.652
17	C3	375	0.651	52	C3	1000	0.668
18	C4	375	0.679	53	C4	1000	0.689
19	C5	375	0.687	54	C5	1000	0.696
20	C6	375	0.704	55	C6	1000	0.717
21	C7	375	0.714	56	C7	1000	0.727
22	C1	500	0.617	57	C1	1125	0.633
23	C2	500	0.659	58	C2	1125	0.647

(Continued)

Table 2 (continued)

No.	Coal powder sample	Concentration	P_m	No.	Coal powder sample	Concentration	P_m
		g/m^3	MPa			g/m^3	MPa
24	C3	500	0.669	59	C3	1125	0.666
25	C4	500	0.683	60	C4	1125	0.679
26	C5	500	0.691	61	C5	1125	0.689
27	C6	500	0.711	62	C6	1125	0.699
28	C7	500	0.719	63	C7	1125	0.719
29	C1	625	0.675	64	C1	1250	0.627
30	C2	625	0.683	65	C2	1250	0.635
31	C3	625	0.688	66	C3	1250	0.653
32	C4	625	0.691	67	C4	1250	0.669
33	C5	625	0.701	68	C5	1250	0.682
34	C6	625	0.716	69	C6	1250	0.695
35	C7	625	0.724	70	C7	1250	0.699

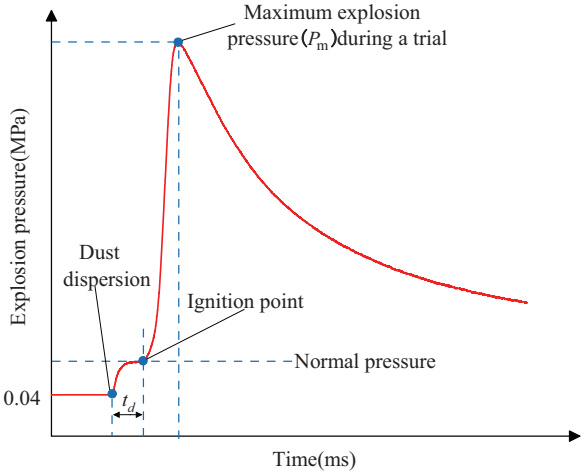

Figure 2: Classical pressure-time curve for 20-L dust explosion experiment

3 Model Theoretical Basis and Establishment

3.1 Long Short-Term Memory Network

Deep learning models are a type of deep neural network with multiple non-linear mapping layers, capable of extracting features from input signals layer by layer and uncovering deeper potential patterns [17]. Among many deep learning models, Long Short-Term Memory (LSTM) is a special architecture of Recurrent Neural Networks (RNNs) that alleviates the vanishing or exploding gradient problem in RNNs by adding a gating mechanism to regulate information flow, making it more adept at capturing semantic dependencies in long sequences. LSTM was created by Hochreiter and Schmidhuber in 1997 [18]. Its core structure consists

of four components: the forget gate, input gate, memory cell state, and output gate, as shown in Fig. 3. The LSTM network structure is calculated as shown in Eqs. (2)–(7) [19,20]:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{2}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{3}$$

$$o_t = \sigma\left(W_o \left[h_{t-1}, x_t\right] + b_o\right) \tag{4}$$

$$\tilde{C}_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{5}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{6}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{7}$$

where, σ denotes the sigmoid activation function, f_t represents the forget gate, i_t denotes the input gate, and o_t represents the output gate. W_f, W_i, and W_o correspond to the weights for each gate, while b_f, b_i, and b_o are the bias terms. h_t and C_t denote the hidden state and the memory cell state at the moment of t, respectively, $\widetilde{C}_t$ is the candidate memory cell state at moment t obtained by the tanh activation function, and x_t is the input sequence at moment t.

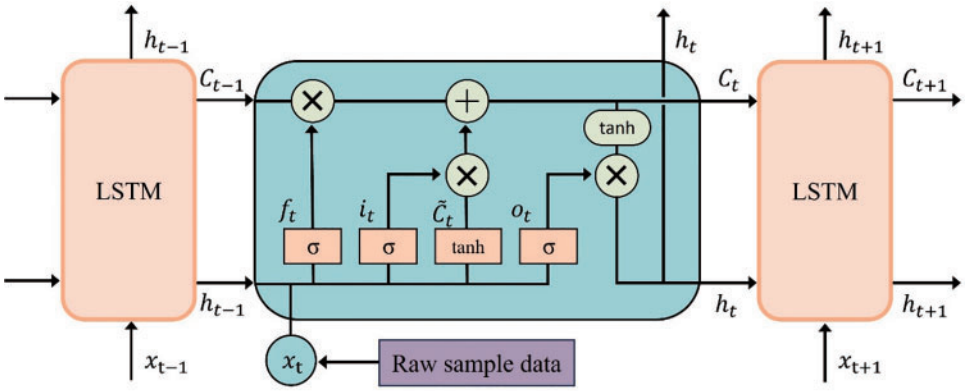

Figure 3: Internal structure of LSTM network

3.2 Multi-Head Attention Mechanism

The Multi-Head Attention Mechanism [21] is a technique that enhances the self-attention mechanism. By decomposing the attention computation into multiple parallel self-attention heads, it allows the model to simultaneously capture various relationships and features from different parts of the input sequence. Each head independently calculates the self-attention scores for the query (Q), key (K), and value (V) in its subspace and outputs the result, thereby enhancing the model's ability to capture complex dependencies and improving its effectiveness in processing sequential data. Multi-Head Attention calculation process is as shown in Eqs. (8)–(11) [22,23]:

$$Q = HW^Q \tag{8}$$

$$K = HW^K \tag{9}$$

$$V = HW^V \tag{10}$$

$$\text{Attention}\left(Q, K, V\right) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{11}$$

where, H is the output vector after LSTM computation, W^Q, W^K, and W^V are the three parameter matrices. The softmax function serves to normalize the computed attentional weights to values of attentional weights between (0, 1) and all weights summing to 1. d_k represents the dimensionality of matrix K, and $\sqrt{d_k}$ serves as a scaling factor, which is used to balance the dot product results and avoid the issue of gradient vanishing due to excessively large dot product values of large vectors. In the Multi-Head Attention mechanism, W^Q, W^K, and W^V perform multiple linear transformations on the vector H, mapping it into different Q, K, and V spaces. Each self-attention head calculates the attention weights using the respective Q and K matrices and performs a weighted summation on the value vector V. Finally, the outputs of all self-attention heads are concatenated into a matrix, and a linear transformation matrix W^o is applied to obtain the final output, as shown in Fig. 4. The output expression of each self-attention head is shown in Eq. (12), and the multi-head attention calculation expression is shown in Eq. (13) [24].

$$\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right) \tag{12}$$

$$\text{MultiHead}\left(Q, K, V\right) = \text{Concat}\left(\text{head}_1, \ldots, \text{head}_h\right) W^o \tag{13}$$

where, head_i represents the attention output of the i-th head, Concat refers to the concatenation operation, and MultiHead is the output result of the multi-head attention layer.

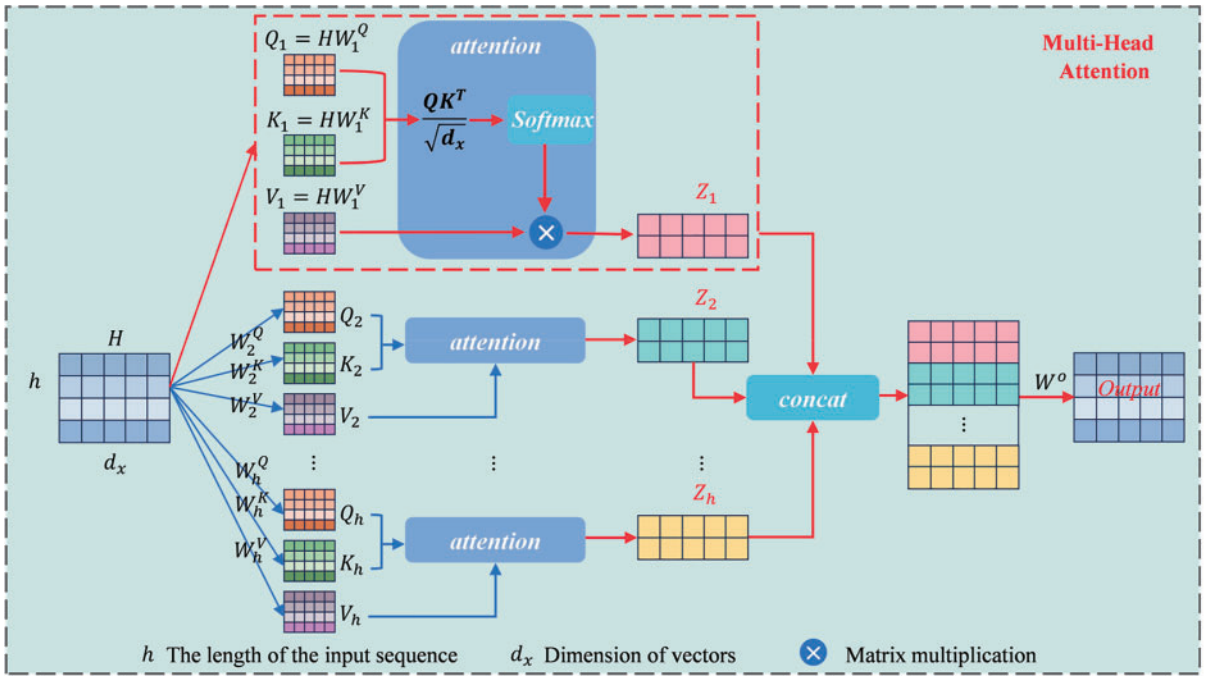

Figure 4: Calculation structure of the multi-head attention mechanism

3.3 Sparrow Search Algorithm

The Sparrow Search Algorithm (SSA) [25] is a swarm intelligence optimization algorithm inspired by the foraging and anti-predation behaviors of sparrows. Due to its excellent stability and powerful search capability, SSA has been widely applied in deep learning modeling. The sparrow population is randomly initialized into discoverers and joiners. Discoverers are responsible for guiding the movement direction of the population and iteratively searching for the global optimal solution. Joiners perform local searches based on the optimal solutions generated by the discoverers. The anti-predation behavior refers to the ability of sparrows to detect danger and promptly issue warning signals, prompting the population to rapidly move to a safe area and update their positions, thereby preventing the algorithm from being trapped in local optima. The population updating process of SSA algorithm is as shown in Eqs. (14)–(16) [26,27]:

Discoverers location update:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} \cdot \exp\left(-\frac{i}{\alpha \cdot iter_{max}}\right), & R_2 < ST \\ X_{i,j}^{t} + Q \cdot \boldsymbol{L}, & R_2 \geq ST \end{cases} \tag{14}$$

where, t denotes the current iteration number, $X_{i,j}^{t}$ represents the position information of the i-th sparrow in the j-th dimension, $iter_{max}$ denotes the maximum number of iterations, $\alpha \in (0,1]$ is a random number, $R_2(R_2 \in [0,1])$ and $ST(ST \in [0.5,1])$ represent the warning threshold and safety threshold, respectively. Q is a random number following a normal distribution, and L is a $1 \times d$ matrix with all elements equal to 1.

Joiners location update:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{worst} - X_{i,j}^{t}}{i^2}\right), & i > n/2 \\ X_{p}^{t+1} + \left| X_{i,j}^{t} - X_{p}^{t+1} \right| \cdot A^{+} \cdot \boldsymbol{L}, & \text{otherwise} \end{cases} \tag{15}$$

where, i represents the number of Joiners, n denotes the total number of sparrows, X_p is the current optimal position occupied by the discoverer, X_{worst} denotes the current global worst position, and A is a $1 \times d$ matrix with each element randomly assigned as either 1 or −1, and $A^{+} = A^{T}(AA^{T})^{-1}$.

Sparrows in Danger location update:

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^{t} + \beta \left| X_{i,j}^{t} - X_{best}^{t} \right| & , f_i > f_g \\ X_{i,j}^{t} + K\left(\frac{|X_{i,j}^{t} - X_{worst}^{t}|}{(f_i - f_w) + \varepsilon}\right) & , f_i = f_g \end{cases} \tag{16}$$

where, X_{best} represents the current global best position, β denotes a random number following a normal distribution with a mean of 0 and a variance of 1, $K \in [-1,1]$ is a random number indicating the movement direction of the sparrow, f_i is the fitness value of the current sparrow, while f_g and f_w represent the current global maximum and minimum fitness values, respectively. ε is a small constant to avoid division by zero in the denominator.

3.4 SSA-LSTM-Multi-Head Attention Prediction Model Construction

In this study, an LSTM-Multi-Head Attention model is employed as the main framework for predicting the maximum explosion pressure of coal dust, and the SSA optimization algorithm is utilized to identify the optimal hyperparameters of the model to enhance prediction accuracy. The original data samples are obtained through coal dust explosion experiments, where the input features and labels of the model are represented as $\{x_{1,}x_{2,}\cdots,x_{n,}\}$ and $\{p_{1,}p_{2,}\cdots,p_{n,}\}$, respectively. Here, x_n denotes the n-th sample data, and p_n represents the corresponding explosion pressure. Fig. 5 illustrates the architecture of the model, and the specific construction process is as follows:

(1) Coal dust explosion pressure experiments were conducted in a 20 L explosion tank to investigate the explosion behaviors under varying particle sizes and concentrations, resulting in the collection of 70 raw data samples. Data normalization was applied to eliminate dimensional differences among the features. In the dataset, particle size and concentration were utilized as input features, while the P_m was designated as the output label. Additionally, 56 data points were randomly selected as the training set, and the remaining 14 data points were used as the test set to evaluate and compare the predictive performance of the developed model.

(2) The LSTM module processes input sequential data using a gating mechanism, effectively extracting the dynamic features of particle size and concentration variations over time during the coal dust explosion process. By storing and updating historical information through memory cells, LSTM captures long-term dependencies within the sequential data, ensuring that the hidden state at each time step not only retains current feature information but also incorporates historical information. These hidden states are then fed into the Multi-Head Attention module to further enhance the model's ability to focus on key features.

(3) The Multi-Head Attention module, through its global attention mechanism, is capable of understanding the dependencies between different time steps throughout the entire explosion process. It utilizes multi-head parallel computation to obtain attention weights, thereby adaptively focusing on the important features in the input sequence. This module not only further enhances the LSTM's ability to handle long sequence data but also improves the model's understanding of the complex nonlinear relationships in the explosion process by capturing the global dependencies among input features, thereby enhancing the predictive performance.

(4) During the model training and optimization process, the SSA algorithm improves the training performance by adjusting hyperparameters. Specifically, SSA uses the model's mean squared error (MSE) as the fitness value to evaluate the optimization effect, ensuring that the model obtains the best hyperparameters and achieves optimal performance during training.

(5) To validate the accuracy and applicability of the SSA-LSTM-Multi-Head Attention model, the prediction results were compared with those of existing models, including PSO-SVM, LSTM, SVM, RF, and ANN, using the test set. The predictive performance of each model was thoroughly evaluated.

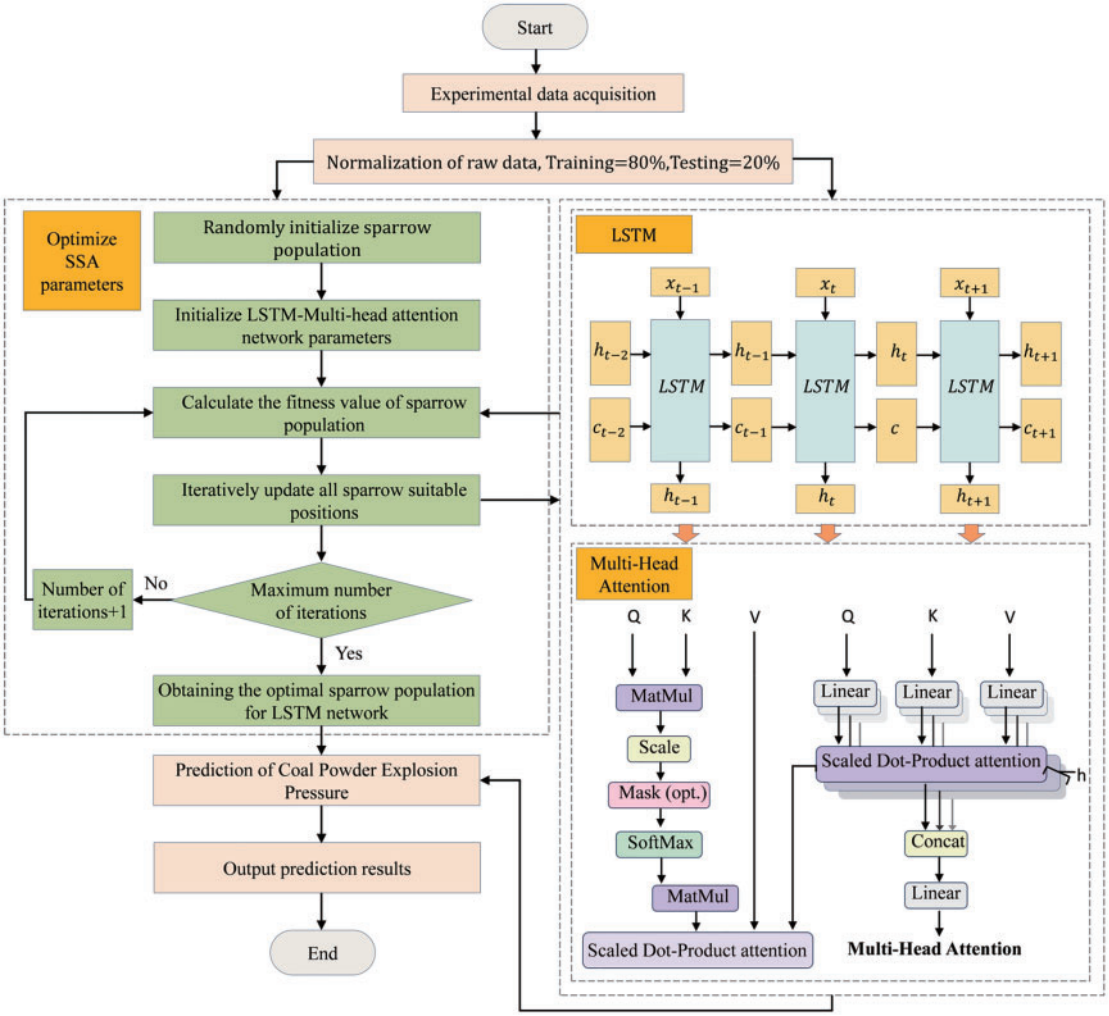

Figure 5: SSA-LSTM-multi-head attention prediction model architecture

3.5 Model Performance Evaluation Metrics

In this study, four evaluation metrics were used to assess the prediction performance of each model for coal dust maximum explosion pressure on the test set:

Root mean square error (RMSE): The smaller the value, the less the divergence between results predicted by the model and the real value, and the better the forecast is [28]. The expression is shown in Eq. (17).

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{p}_i - p_i)^2} \tag{17}$$

Mean absolute error (MAE): It represents the mean absolute error between the predicted values and the actual values. The smaller this value, the smaller the deviation between the predicted and actual values, indicating higher prediction accuracy of the model [29]. The expression is shown in Eq. (18).

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}|\hat{p}_i - p_i| \tag{18}$$

Mean absolute percentage error (MAPE): It measures the relative error between the predicted values and the actual values by calculating the percentage of the prediction error relative to the actual value [30]. The expression is shown in Eq. (19).

$$\mathrm{MAPE} = \frac{1}{m}\sum_{i=1}^{m}\frac{|\hat{p}_i - p_i|}{p_i} \tag{19}$$

Coefficient of determination (R^2): The value ranges from 0 to 1 and is used to interpret the degree of fit between the model's predictions and the actual values [31]. The expression is shown in Eq. (20).

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(\hat{p}_i - p_i)^2}{\sum_{i=1}^{m}(p_i - \overline{p})^2} \tag{20}$$

where, m represents the number of samples, p_i is the actual value of the i-th sample, $\hat{p}_i$ is the predicted value of the i-th sample, and $\overline{p}$ is the mean of the actual values in the sample set.

4 Results and Discussion

4.1 Feature Correlation Analysis

The maximum explosion pressure (P_m) of coal dust is influenced by multiple factors, and changes in these factors can trigger chain reactions in other conditions. Additionally, interactions exist among various factors in the experimental results. Therefore, analyzing the correlation between different particle size parameters, mass concentration, and P_m, identifying key influencing factors, and eliminating redundant features can not only intuitively reflect the impact of different factors on the maximum explosion pressure but also effectively reduce model complexity, shorten training time, enhance generalization ability, and prevent overfitting. In this study, the Spearman correlation coefficient and random forest feature selection methods were employed to screen the most influential feature parameters from nine particle size parameters and coal dust concentration, aiming to analyze the influence patterns of key factors and optimize model performance.

Correlation coefficient is a statistical method used to measure the degree of relationship between two sets of variables. In practical applications, common correlation coefficients include Pearson, Spearman, and Kendall [32,33]. In this study, since the experimental parameters are continuous variables and have a nonlinear relationship with the maximum explosion pressure, the Spearman correlation coefficient is chosen for analysis.

The Spearman correlation coefficient evaluates the monotonic relationship between two variables through ranking, avoiding reliance on specific values. Fig. 6 presents the results of the Spearman correlation coefficient in the form of a heatmap, where darker colors indicate stronger correlations. The correlation coefficients of different parameters with P_m are calculated as shown in Eqs. (21) and (22):

$$D = \sum_{i=1}^{n}\left(X_i^L - P_i\right)^2 \tag{21}$$

$$\rho = 1 - \frac{6D}{n\left(n^2 - 1\right)} \tag{22}$$

where, n represents the number of samples, X_i^L denotes the rank of the L-th feature in the original sample x_i after sorting, P_i represents the rank of the true value p_i of the i-th sample after sorting, and D is the sum of squared differences after ranking. ρ represents the Spearman correlation coefficient, with a range of $[-1, 1]$. A

value of 1 indicates a perfect positive correlation, −1 indicates a perfect negative correlation, and 0 indicates no monotonic relationship.

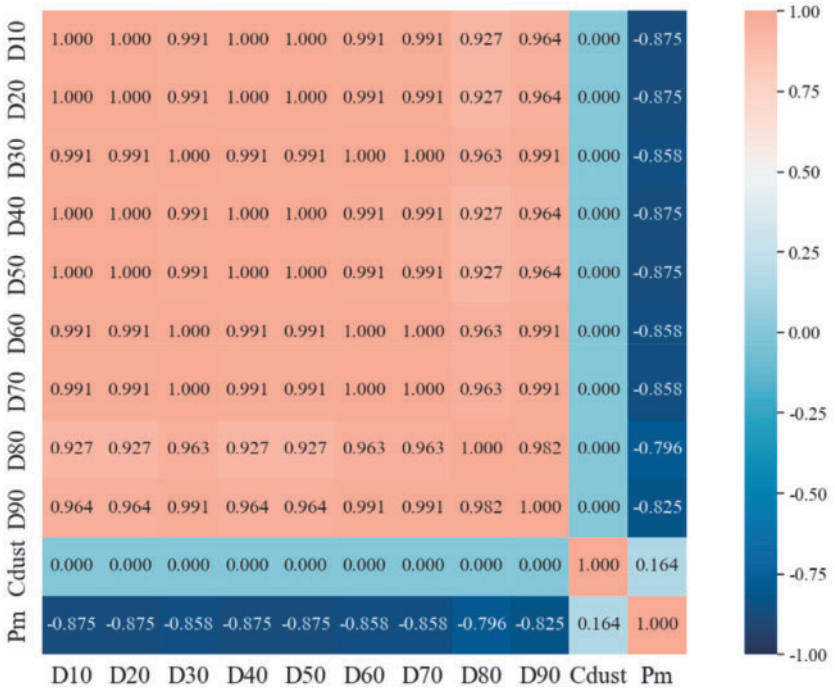

Figure 6: Heat map of Spearman correlation coefficient

The Spearman correlation coefficient can effectively determine the monotonic relationship between features and the target variable. However, due to the complex interactions between features and the intricate linear relationships between features and the target variable, Spearman analysis alone is insufficient to comprehensively assess the impact of features on the target variable. Random Forest is an ensemble learning method that constructs multiple decision trees and uses the voting results of these trees to obtain the final prediction value [34]. It not only excels in classification and regression tasks but has also been applied to feature selection by Genuer et al. [35] and Marie et al. [36]. Research has shown that random forests can effectively capture nonlinear relationships and interactions between features and the target variable, demonstrating excellent stability in feature selection. In this study, the impact of each parameter on the explosion pressure in coal dust explosion experiments is ranked based on the mean squared error change in the random forest. The parameter rankings after feature selection are illustrated in Fig. 7. The random forest feature importance calculation process is as shown in Eqs. (23)–(25):

$$\mathrm{MSE}_{\mathrm{original}} = \frac{1}{n}\sum_{i=1}^{n}\left(p_i - \hat{p}_i\right)^2 \tag{23}$$

$$\mathrm{MSE}^{L}_{\mathrm{perturbed}} = \frac{1}{n}\sum_{i=1}^{n}\left(p_i - \hat{p}_i^{L}\right)^2 \tag{24}$$

$$\mathrm{Importance}\left(L\right) = \mathrm{MSE}^{L}_{\mathrm{perturbed}} - \mathrm{MSE}_{\mathrm{original}} \tag{25}$$

where, n represents the number of samples, p_i denotes the true value of the i-th sample, $\hat{p}_i$ represents the predicted value of the i-th sample, and $\hat{p}_i^{L}$ denotes the predicted value of the i-th sample after perturbing

feature L. $MSE_{original}$ is the mean squared error of the original data, $MSE^{L}_{perturbed}$ is the mean squared error after perturbing feature L, and Importance(L) represents the importance coefficient of feature L.

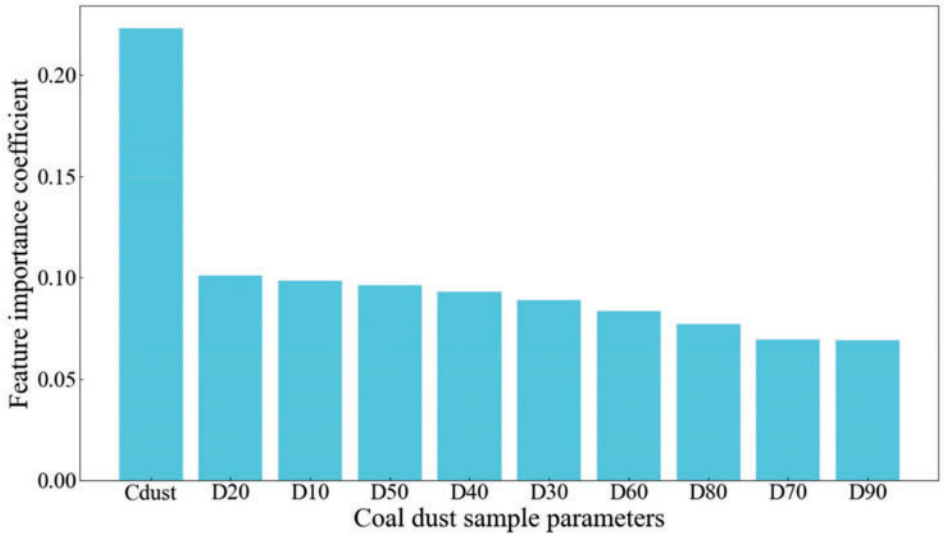

Figure 7: Random forest feature selection results

The relationship between the 10 parameters of coal dust and P_m, as well as the rankings of feature importance, were determined by combining Spearman correlation analysis with the Random Forest feature selection method. Table 3 displays the importance and correlation coefficients for each parameter.

Table 3: Feature importance coefficients and Spearman correlation coefficients with *t*-test *p*-values

Coal dust sample parameters	P_M		
	Random forest importance coefficients	Spearman correlation coefficients	*p* values of the *t*-test
D_{10}	0.098	−0.875	0.000
D_{20}	0.100	−0.875	0.000
D_{30}	0.088	−0.858	0.000
D_{40}	0.093	−0.875	0.000
D_{50}	0.096	−0.875	0.000
D_{60}	0.084	−0.858	0.000
D_{70}	0.069	−0.858	0.000
D_{80}	0.077	−0.796	0.000
D_{90}	0.069	−0.825	0.000
C_{dust}	0.223	0.164	–

In dust explosion research, the median particle size (D_{50}) is commonly used to represent the average particle size of dust samples. However, Castellanos et al. [37] proposed that using D_{50} alone may not accurately describe the impact of particle size on explosion severity. This is further confirmed by the Spearman correlation analysis shown in Fig. 6. Overall, all particle size parameters of coal dust exhibit

a negative correlation with P_m, indicating that P_m decreases as particle size increases. In contrast, the monotonic correlation between dust mass concentration and P_m is relatively weak. Among the nine particle size percentile parameters, the particle size components at the 10th, 20th, 40th, and 50th percentiles exhibit the strongest monotonic correlation with the maximum explosion pressure, followed by the 30th percentile. This indicates that, in addition to D_{50}, D_{10}, D_{20}, and D_{40} are also suitable parameters for characterizing the influence of coal dust particle size on explosion pressure. Therefore, relying solely on the distribution of the D_{50} parameter cannot accurately assess the impact of particle size on explosion pressure and may lead to the erroneous assumption that the hazard of smaller particle sizes applies to the entire dust cloud. This is because, during the reaction process, smaller particles are typically ignited first, acting as new ignition sources to form flame fronts that subsequently ignite surrounding coal dust particles, thereby promoting the progression of the explosion reaction. Table 3 presents the p-values of the particle size parameters after statistical significance testing. The results indicate that the Spearman correlation analysis between the independent variable (particle size) and the dependent variable (P_m) is statistically reliable.

From the feature selection results in Fig. 7, it is evident that the parameter with the greatest influence on the target variable P_m is the mass concentration, with an importance coefficient of 0.223, followed by the particle size parameters. This phenomenon arises because, in the 20-L explosion tank experiment, although smaller particles have a larger specific surface area, which increases the contact area with oxygen and promotes the explosion reaction, as the dust concentration increases further, the oxygen content within the tank becomes insufficient to ensure the complete combustion of all coal dust particles, which becomes the primary factor limiting the explosion reaction. As a result, the mass concentration of coal dust has a more significant effect on explosion pressure than particle size. Among the particle size parameters, the top three with the greatest influence are D_{10}, D_{20}, and D_{50}. This further indicates that in explosion experiments, smaller particle sizes are more likely to serve as ignition sources, igniting nearby particles, and therefore have a greater impact on P_m compared to larger particle sizes.

Based on the above conclusions, this study will use mass concentration and the particle size percentages of D_{10}, D_{20}, and D_{50} to analyze their influence on the maximum explosion pressure of coal dust. These five parameters will also be used as input features for the coal dust P_m prediction model. The test set after feature selection is shown in Table 4.

Table 4: Test dataset after feature selection

No.	Coal dust sample	Input				Output
		D_{10} (μm)	D_{20} (μm)	D_{50} (μm)	Concentration (g/m^3)	
1	C3	4.29	9.29	33.49	125	0.526
2	C5	2.12	4.50	17.44	125	0.678
3	C2	4.62	11.13	40.34	250	0.568
4	C4	3.97	7.75	30.33	250	0.669
5	C1	6.78	15.99	44.59	375	0.583
6	C4	3.97	7.75	30.33	375	0.679
7	C1	6.78	15.99	44.59	500	0.617
8	C5	2.12	4.50	17.44	500	0.691
9	C1	6.78	15.99	44.59	625	0.675
10	C4	3.97	7.75	30.33	750	0.702
11	C1	6.78	15.99	44.59	875	0.657
12	C6	1.74	3.75	11.47	1000	0.717

(Continued)

Table 4 (continued)

No.	Coal dust sample	Input				Output
		D_{10} (μm)	D_{20} (μm)	D_{50} (μm)	Concentration (g/m^3)	
13	C2	4.62	11.13	40.34	1125	0.647
14	C6	1.74	3.75	11.47	1250	0.695

4.2 Impact of Coal Dust Concentration on P_m

In this study, the influence of mass concentrations ranging from 125.0 g/m^3 to 1250.0 g/m^3 on the maximum explosion pressure (P_m) of coal dust was investigated for seven different particle sizes, as shown in Fig. 8. Overall, at a constant particle size, P_m initially increases with rising mass concentration but begins to decline after reaching a certain threshold. For coal dust samples C1–C3, P_m shows a significant initial increase with concentration, peaking at 625.0 g/m^3, after which it gradually decreases. Thus, 625.0 g/m^3 is identified as the optimal explosion concentration for samples C1–C3. For samples C4–C7, when the concentration exceeds 750.0 g/m^3, P_m exhibits a notable decreasing trend with further increases in concentration. This trend is primarily influenced by the dispersion characteristics of coal dust and the oxygen content within the explosion tank. When the coal dust concentration is below the optimal explosion concentration (625.0 g/m^3 and 750.0 g/m^3 in this study), the particles are dispersed under turbulent flow, resulting in relatively large inter-particle spacing. Although combustion occurs in an oxygen-rich environment, the large spacing slows down the propagation of the combustion wave, making flame propagation difficult and leading to a lower P_m. As the concentration increases, the inter-particle spacing decreases while oxygen remains sufficient, promoting flame propagation and heat transfer, which raises P_m to its peak. However, when the concentration exceeds the optimal level, the particle density becomes excessively high. While heat transfer between particles is further enhanced, the limited oxygen supply in the tank becomes insufficient to support the complete combustion of all coal dust particles. Consequently, unburned particles are generated, which hinder flame propagation and absorb heat, ultimately causing P_m to decline.

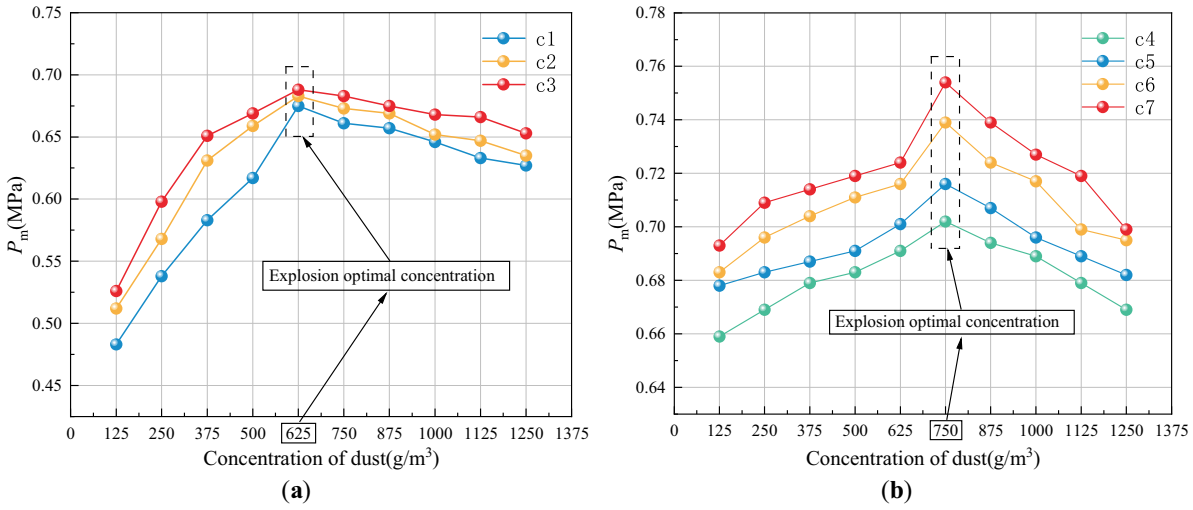

Figure 8: Impact of various concentrations of coal dust samples on P_m: (**a**) The P_m of coal dust samples C1–C3 at different concentrations; (**b**) The P_m of coal dust samples C4–C7 at different concentrations

Additionally, as the particle size of coal dust decreases, the specific surface area increases, providing a larger contact area for oxygen and particles, accelerating the combustion rate, and releasing more energy per unit time, thereby intensifying the explosion reaction. Therefore, under conditions of sufficient oxygen, smaller coal dust particles produce higher peak P_m values compared to larger particles. However, due to the faster combustion rate of smaller particles and the limitation of oxygen content, smaller particles generate more unburned particles, which suppress the progression of the explosion reaction, causing a more pronounced decline in P_m compared to larger particles.

4.3 Impact of Coal Dust Particle Size on P_m

Fig. 9 presents the variation curves of P_m for seven different coal dust samples under dust concentrations of 375, 500, 625, 750, 1000, and 1250 g/m^3. As shown in the figure, at a constant concentration, P_m exhibits an increasing trend with decreasing coal dust particle size. Among the seven coal dust samples, sample C7, which has the smallest mean particle size D_{50}, produces the highest P_m values across all six concentrations, measuring 0.699, 0.714, 0.719, 0.724, 0.727, and 0.754 MPa, respectively. In contrast, samples C1, C2, and C3, which have larger mean particle sizes D_{50}, exhibit maximum explosion pressures below 0.7 MPa compared to the other four samples. This phenomenon can be attributed to the increase in the specific surface area as particle size decreases, which shortens the diffusion time of oxygen to the particle surface, enhances combustion efficiency, and releases more heat. Consequently, the explosion reaction becomes more intense, leading to a higher P_m with decreasing particle size.

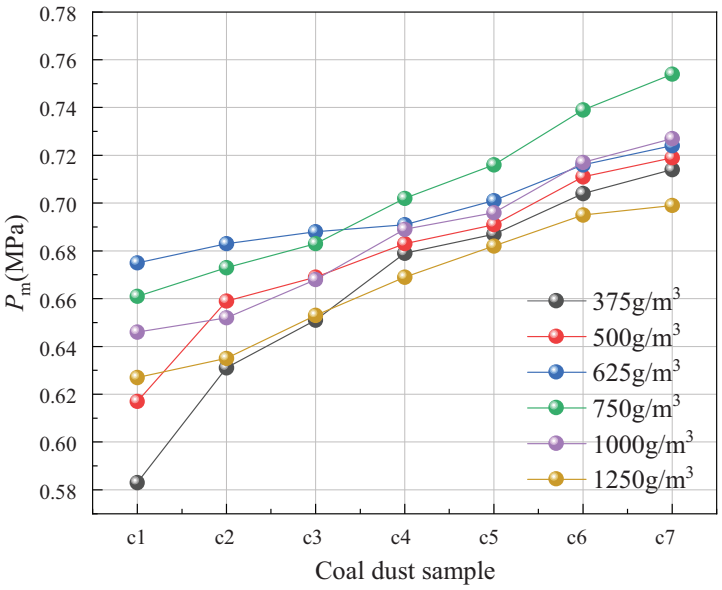

Figure 9: Impact of various particle sizes of coal dust samples on P_m

When the particle size of coal dust is sufficiently fine, the combustion has already developed sufficiently, and particle size is no longer a major limiting factor for the maximum explosion pressure. As shown in Fig. 9, the explosion pressure of samples C1–C3 shows a clear increasing trend, while the increasing trend of samples C4–C7 begins to slow down. This is because the D_{10} and D_{20} values of samples C4–C7 are smaller than those of C1–C3, resulting in a larger specific surface area of the coal dust particles. The larger specific surface area makes the coal dust more easily ignited, quickly absorbing heat and increasing the combustion temperature, thereby promoting the explosion reaction. However, the larger specific surface area also accelerates the

reaction between coal dust and oxygen, leading to a higher rate of oxygen consumption. As the particle size continues to decrease, oxygen becomes insufficient in the sealed explosion tank, limiting the explosion reaction and causing the increase in P_m to gradually plateau.

4.4 Validation of the Prediction Results for P_m of Coal Dust

The model in this study is developed using the TensorFlow framework and written in Python. During the model training phase, the optimization algorithm SSA searches for the model's hyperparameters globally by evaluating the MSE, continuing the process until no further improvement can be made. Fig. 10 illustrates the detailed process of SSA's iterative optimization and the resulting best hyperparameters of the model. The learning rate, number of iterations, number of LSTM hidden layer nodes (for both layers), and batch size found for the model were 0.0002, 204, 52, 81, and 14, respectively. Fig. 11 shows the MSE validation process of the LSTM-Multi-Head Attention model after optimization by SSA, including iterations on both the training and test datasets. At the 212th iteration, the MSE on the training and test sets reached their minimum values of 0.000936 and 0.00103, respectively, indicating that the model's training was optimal and training was stopped at this point.

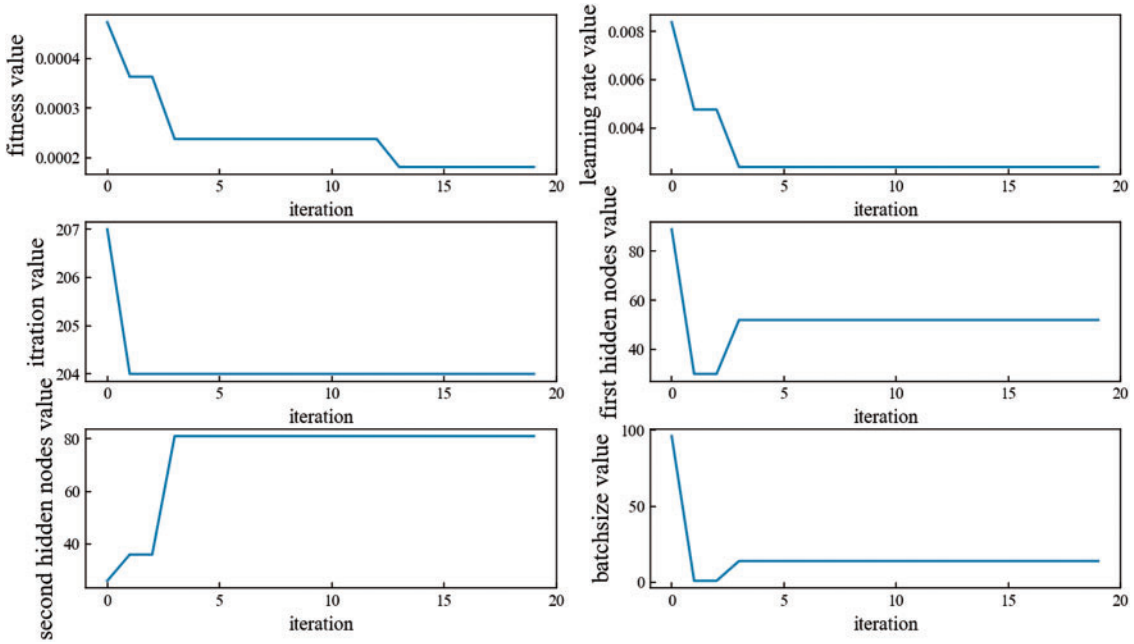

Figure 10: Iteration results of SSA in searching for model hyperparameters

To validate the superiority of the SSA optimization algorithm and the Multi-Head Attention mechanism in improving the LSTM model's prediction of coal dust P_m values, the prediction performance of commonly used models in fire and explosion prediction, including ANN, RF, SVM, PSO-SVM, and LSTM, was compared under the same experimental conditions. Additionally, the prediction performance of each model was quantified using four evaluation metrics. The calculated results of the model's performance metrics are shown in Table 5. From Table 5, it can be observed that the SSA-LSTM-Multi-Head Attention model performs slightly better than the PSO-SVM model, with the R^2 for the test and training sets increasing by 0.0049 and 0.0194, respectively. Compared to the LSTM, SVM, RF, and ANN models, the R^2 values for the test and training sets improved by 0.0049 and 0.0302, 0.0634 and 0.063, 0.0095 and 0.0802, and 0.0062 and

0.1307, respectively. Furthermore, the SSA-LSTM-Multi-Head Attention model exhibited the lowest RMSE, MAPE, and MAE values for both the test and training sets, with R^2 values of 0.9841 and 0.9743, respectively, which are closest to 1. The performance of each model in the training set, after rolling prediction cross-validation, is shown in Figs. 12 and 13. From the figures, it is clear that the SSA-LSTM-Multi-Head Attention model has the lowest MSE and achieves the best values across all four evaluation metrics, demonstrating superior generalization ability.

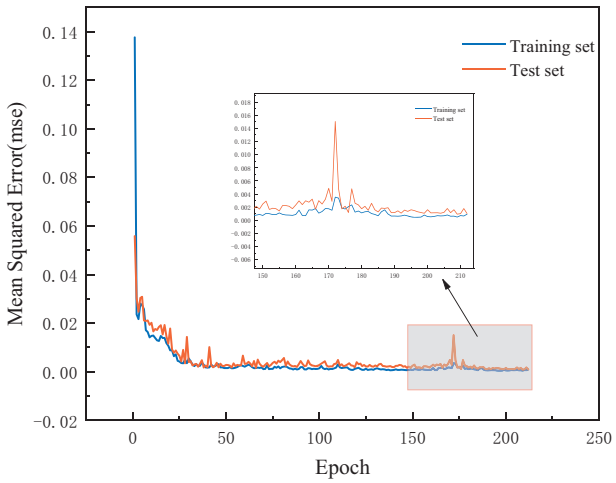

Figure 11: Model iteration and MSE validation

Table 5: Comparison of prediction evaluation metrics for P_m of coal dust in different models

Model	MAPE		RMSE		MAE		R^2	
	Training set	Test set	Training set	Test set	Training set	Test set	Training set	Test set
ANN	0.0070	0.0226	0.0074	0.0215	0.0044	0.0135	0.9779	0.8436
RF	0.0073	0.0210	0.0079	0.0177	0.0044	0.0125	0.9746	0.8941
SVM	0.0186	0.0192	0.0140	0.0162	0.0124	0.0124	0.9207	0.9113
LSTM	0.0077	0.0153	0.0072	0.0128	0.0051	0.0096	0.9792	0.9441
PSO-SVM	0.0093	0.01423	0.0071	0.0115	0.0062	0.0088	0.9792	0.9549
SSA-LSTM-Multi-Head Attention	0.0074	0.0108	0.0030	0.0087	0.0049	0.0069	0.9841	0.9743

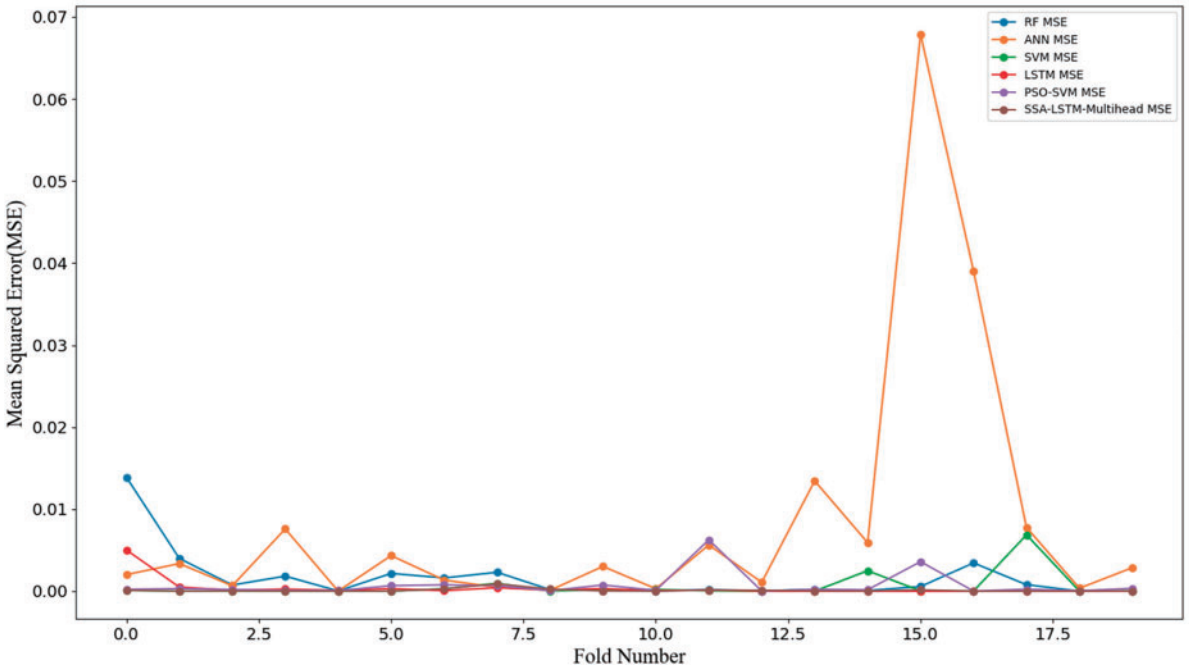

Figure 12: MSE variation of each model in rolling prediction cross-validation

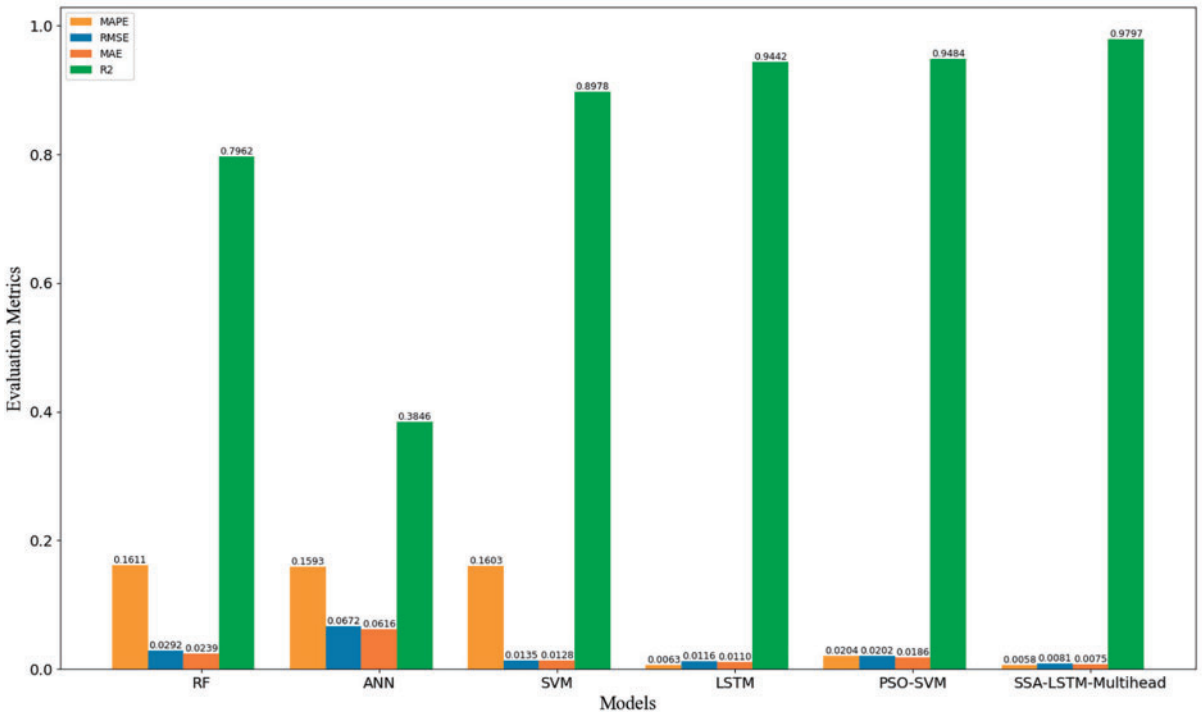

Figure 13: Results of four evaluation metrics for each model after cross-validation

The prediction results of all models on the test set are shown in Fig. 14. From the figure, the coal dust maximum explosion pressure prediction models based on the RF and ANN algorithms exhibit good fitting

performance on the training set, with predicted values closely matching the true values. However, these models are prone to overfitting. When predicting the test set, the prediction accuracy of these two models significantly decreases, and the degree of dispersion increases, leading to a decline in model robustness. In contrast, the SSA-LSTM-Multi-Head Attention, PSO-SVM, and LSTM prediction models show small deviations between predicted and true values on the training set, demonstrating strong generalization ability. Among these, the LSTM-Multi-Head Attention model significantly outperforms the other two models, with smaller errors between predicted and true values. For the test set predictions, all three models maintain relatively high accuracy; however, the LSTM prediction model without SSA optimization exhibits the poorest performance. This further proves that the introduction of the SSA optimization algorithm and the Multi-Head Attention module enhances the model's prediction performance. The SVM prediction model shows slightly larger deviations between predicted and true values on the training set, but its generalization ability on the test set is superior to that of the RF and ANN models. Comparing the MAPE, RMSE, and MAE values of different models, in the test set, the LSTM-Multi-Head Attention model reduces the MAPE by 0.0034, 0.0045, 0.0084, 0.0102, and 0.0118 compared to PSO-SVM, LSTM, SVM, RF, and ANN, respectively. The RMSE is reduced by 0.0028, 0.0041, 0.0075, 0.009, and 0.0128, while the MAE is reduced by 0.0019, 0.0027, 0.0055, 0.0056, and 0.0066.

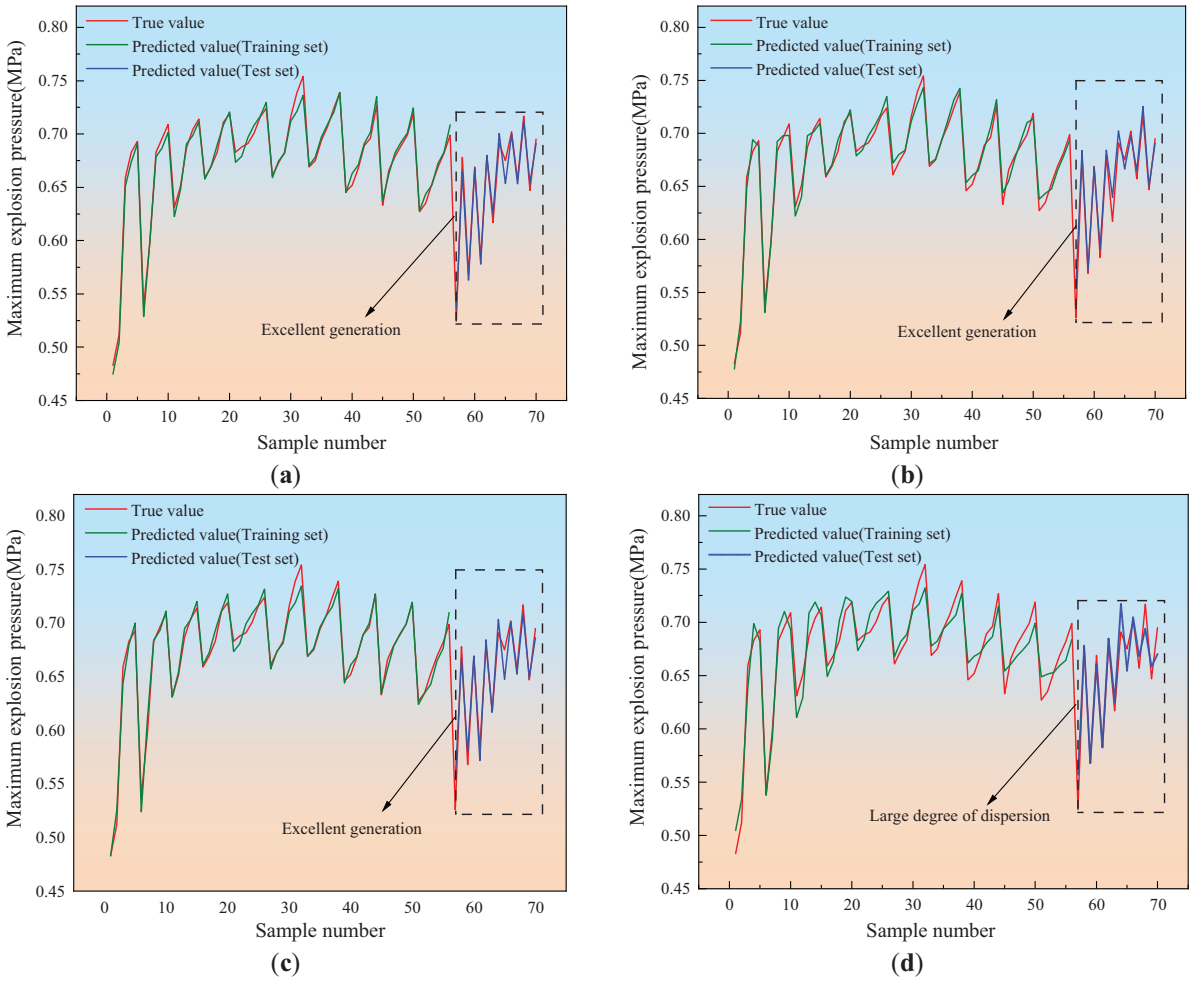

Figure 14: (Continued)

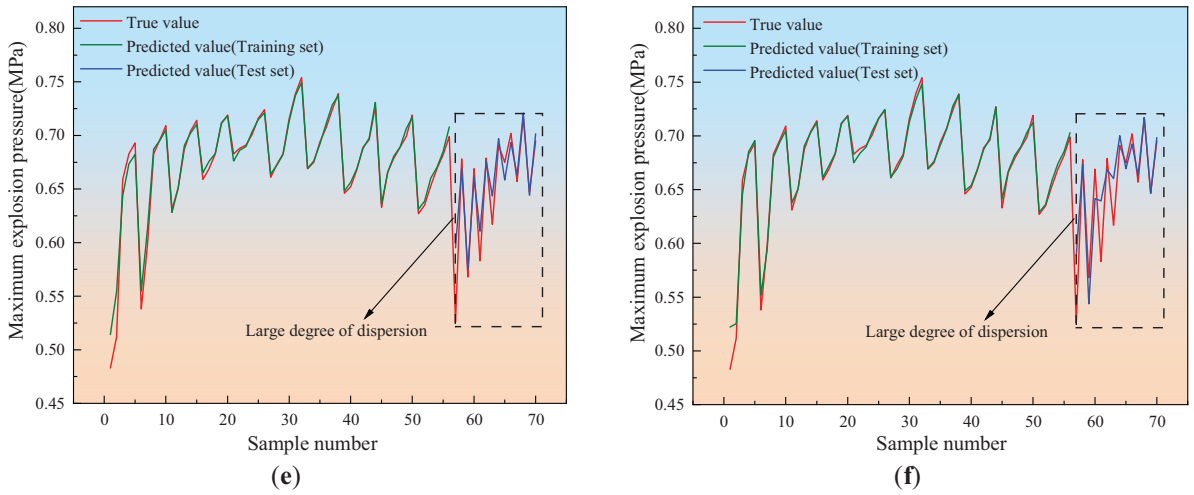

Figure 14: Comparison of model prediction results in test set and training set: (**a**) SSA-LSTM-multi-head attention model coal dust P_m prediction results; (**b**) PSO-SVM model coal dust P_m prediction results; (**c**) LSTM model coal dust P_m prediction results; (**d**) SVM model coal dust P_m prediction results; (**e**) RF model coal dust P_m prediction results; (**f**) ANN model coal dust P_m prediction results

In summary, for the prediction of coal dust P_m, compared to traditional prediction models, the introduction of the Multi-Head Attention mechanism into the LSTM model enhances prediction accuracy by leveraging the strengths of both modules in capturing temporal dependencies in sequence data and global dependencies among features. Additionally, the use of SSA to automatically search for the optimal hyperparameters of the model further improves its stability and generalization capability, making it suitable for predicting the maximum explosion pressure of coal dust.

4.5 Validation of the SSA-LSTM-Multi-Head Attention Model in Predicting the P_m of Coal Dust

Based on the analysis of the above studies, compared to the currently available fire and explosion prediction models. The SSA-LSTM-Multi-Head Attention model constructed with coal dust particle size and concentration as the original dataset has the best prediction accuracy for the value of coal dust P_m, and showed better generalization in the test set. In order to verify the applicability of the model proposed in this study in predicting the P_m of coal dust, coal samples were selected in the "2 Experimental method and materials", and the same equipment and experimental methods for coal dust P_m experiments, and a new set of 10 data points was obtained. According to "Section 3.4 SSA-LSTM-Multi-Head Attention prediction model construction" combined with newly acquired experimental data, SSA-LSTM-Multi-Head Attention model is used to predict the P_m of coal dust. The results are shown in Fig. 15.

From the validation results, the SSA-LSTM-Multi-Head Attention model's predicted values generally show a distribution characteristic greater than the actual values. The average absolute error rate between the actual and predicted maximum explosion pressures of coal dust samples is 0.88%, with the differences in maximum and minimum explosion pressures being 0.0122 MPa and 0.001 MPa, respectively. The average absolute error is 0.0059 MPa. The results indicate that this model can accurately and reliably reflect the maximum explosion pressure generated during coal dust explosions, providing an effective reference for coal mine dust explosion risk assessment.

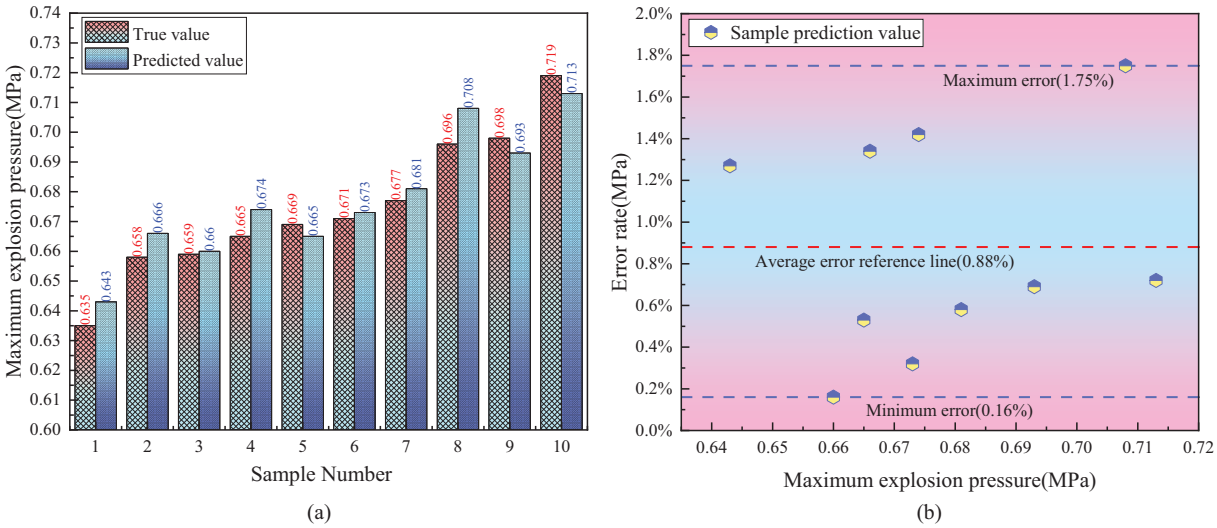

Figure 15: Prediction of P_m of coal dust and error distribution by SSA-LSTM-multi-head attention model: (**a**) Prediction of P_m of coal dust by SSA-LSTM-multi-head attention model; (**b**) Error distribution of SSA-LSTM-multi-head attention model in predicting the P_m of coal dust

5 Conclusions

The Spearman correlation coefficient indicates a positive correlation between coal dust particle size and the maximum explosion pressure. The random forest screening results further reveal that concentration has the greatest impact on P_m. By using mass concentration and particle sizes D_{10}, D_{20} and D_{50} as feature parameters for the coal dust maximum explosion pressure prediction model, the effects of mass concentration and particle size on P_m are analyzed to more accurately assess coal dust explosion pressure.

At the same particle size, P_m shows a trend of initially increasing and then decreasing as the mass concentration increases. Samples C1–C3 reach a peak at a mass concentration of 625.0 g/m^3 and then gradually decrease, while samples C4–C7 continuously decrease after the mass concentration exceeds 750.0 g/m^3. This trend is influenced by the dispersion and flowability of the dust: at low concentrations, the particle spacing is large, hindering flame propagation, resulting in lower pressure; at high concentrations, flame propagation is improved, but oxygen deficiency limits the reaction, leading to a decrease in pressure. At the same time, when the mass concentration remains constant, a decrease in particle size leads to an increase in P_m. Fine coal dust, with a larger specific surface area, accelerates oxygen diffusion and combustion reactions, thereby increasing pm. Although smaller particle sizes enhance oxygen contact, the larger specific surface area also increases oxygen consumption. In a closed explosion chamber, the oxygen content limits the further increase of P_m.

Compared to the PSO-SVM, LSTM, SVM, RF, and ANN models, the SSA-LSTM-Multi-Head Attention model outperforms these models in predicting the maximum explosion pressure of coal dust, demonstrating higher prediction accuracy and generalization ability. Verified through the 20 L explosion tank experiment, the model's prediction results show an average absolute error rate of 0.88%, a maximum explosion pressure difference of 0.0122 MPa, a minimum pressure difference of 0.001 MPa, and an average absolute error of 0.0059 MPa. These results confirm that the model can effectively predict the maximum explosion pressure of coal dust, providing strong support for coal dust explosion prevention and risk assessment in coal mines.

In this study, owing to the inherent safety limitations of explosion experiments, we primarily focus on the influence of coal dust particle size and concentration on explosion pressure P_m. However, other

physicochemical properties of coal dust (e.g., volatile content, moisture content) and different experimental conditions (e.g., oxygen concentration, methane concentration) may also significantly affect P_m. Comprehensive consideration of these factors could further optimize the model's predictive performance. The 70 datasets used in this study are relatively small for deep learning models. Data augmentation and synthetic data generation could be considered to expand the dataset, but it is crucial to ensure that the augmented data aligns with the dynamic principles of dust explosions.

Acknowledgement: The authors acknowledge the support from the Institute of Interdisciplinary Research on Intelligent Mines, Heilongjiang University of Science and Technology, Harbin, China; the School of Resources and Engineering, Heilongjiang University of Technology, Jixi, China; and Heilongjiang Longmei Jixi Mining Co., Ltd., Xinfa Coal Mine, Jixi, China.

Funding Statement: This research was funded by the Research on Intelligent Mining Geological Model and Ventilation Model for Extremely Thin Coal Seam in Heilongjiang Province, China (2021ZXJ02A03), the Demonstration of Intelligent Mining for Comprehensive Mining Face in Extremely Thin Coal Seam in Heilongjiang Province, China (2021ZXJ02A04), and the Natural Science Foundation of Heilongjiang Province, China (LH2024E112).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Haitao Wang; methodology, Weihao Li; software, Weihao Li and Haitao Wang; validation, Weihao Li; formal analysis, Weihao Li; investigation, Haitao Wang; resources, Yongli Liu; data curation, Weihao Li, Haitao Wang and Taoren Du; writing—original draft preparation, Weihao Li; writing—review and editing, Weihao Li and Haitao Wang; visualization, Taoren Du; supervision, Yongli Liu; project administration, Yongli Liu; funding acquisition, Yongli Liu. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the Corresponding Author, Weihao Li, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Hertzberg M, Zlochower IA. Devolatilization rates and intraparticle wave structures during the combustion of pulverized coals and polymethylmethacrylate. Symp Combust. 1991;23(1):1247–55. doi:10.1016/S0082-0784(06)80387-5.
2. Jurca A, Lupu C, Paraian M, Vatavu N, Iacob-Ridzi FT. Analysis of explosivity parameters and environmental safety for combustible dusts. Environ Eng Manag J. 2014;13(6):1433–8. doi:10.30638/eemj.2014.157.
3. Yuan Z, Khakzad N, Khan F, Amyotte P. Dust explosions: a threat to the process industries. Process Saf Environ Prot. 2015;98:57–71. doi:10.1016/j.psep.2015.06.008.
4. Tan Y, Yu C, Gao Y, Yuan H. Effect of particle size of coal dusts on explosion pressure and rise rate of explosion pressure. In: The 2008 International Symposium on Safety Science and Technology; 2008; Beijing, China.
5. Azam S, Mishra PD. Effects of particle size, dust concentration and dust-dispersion-air pressure on rock dust inertant requirement for coal dust explosion suppression in underground coal mines. Process Saf Environ Prot. 2019;126(2):35–43. doi:10.1016/j.psep.2019.03.030.
6. Yang Y, Luo Z, Ding X, Zhang F, Luo C, Zhang H, et al. Effects of dust concentration, particle size, and crude oil concentration on the explosion characteristics of oil-immersed coal dust. Fuel. 2024;356(8):129596. doi:10.1016/j.fuel.2023.129596.
7. Qi Y, Xue K, Wang W, Cui X, Liang R. Prediction model of borehole spontaneous combustion based on machine learning and its application. Fire. 2023;6(9):357. doi:10.3390/fire6090357.

8. Lei C, Deng J, Cao K, Xiao Y, Ma L, Wang W, et al. A comparison of random forest and support vector machine approaches to predict coal spontaneous combustion in gob. Fuel. 2018;239:297–311. doi:10.1016/j.fuel.2018.11.006.
9. Shankar US, Prasad DM, Aishwarya M. Prediction of spontaneous combustion susceptibility of coal seams based on coal intrinsic properties using various machine learning tools. Environ Sci Pollut Res Int. 2023;30(26):69564–79. doi:10.1007/s11356-023-27248-y.
10. Prasanjit D, Saurabh K, Kumar C, Pandit D, Chaulya SK, Ray SK, et al. t-SNE and variational auto-encoder with a bi-LSTM neural network-based model for prediction of gas concentration in a sealed-off area of underground coal mines. Soft Comput. 2021;25(22):14183–207. doi:10.1007/s00500-021-06261-8.
11. Borhani GNT, Afzali A, Bagheri M. QSPR estimation of the auto-ignition temperature for pure hydrocarbons. Process Saf Environ Prot. 2016;103(Pt A):115–25. doi:10.1016/j.psep.2016.07.004.
12. Liu T, Cai Z, Wang N, Jia R, Tian W. Prediction method of coal dust explosion flame propagation characteristics based on principal component analysis and bp neural network. Math Probl Eng. 2022;1:5078134. doi:10.1155/2022/5078134.
13. Qi Y, Gan X, Li Z, Li L, Wang Y, Ji W. Variation and prediction methods of the explosion characteristic parameters of coal dust/gas mixtures. Energies. 2021;14(2):264. doi:10.3390/en14020264.
14. Khan AU, Salman S, Muhammad K, Habib M. Modelling coal dust explosibility of khyber pakhtunkhwa coal using random forest algorithm. Energies. 2022;15(9):3169. doi:10.3390/en15093169.
15. Wang Q, Fang X, Shu CM, Wang Q, Sheng Y, Jiang J, et al. Minimum ignition temperatures and explosion characteristics of micron-sized aluminium powder. J Loss Prev Process Ind. 2020;64:104076. doi:10.1016/j.jlp.2020.104076.
16. GB/T 16426-1996. Determination for maximum explosion pressure and maximum rate of pressure rise of dust cloud. Beijing, China: The State Bureau of Quality and Technical Supervision; 1996.
17. Lecun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. doi:10.1038/nature14539.
18. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–80. doi:10.1162/neco.1997.9.8.1735.
19. Zhang P, Li C, Peng C, Tian J. Ultra-short-term prediction of wind power based on error following forget gate-based long short-term memory. Energies. 2020;13(20):5400. doi:10.3390/en13205400.
20. Sun L, Qin H, Przystupa K, Majka M, Kochan O. Individualized short-term electric load forecasting using data-driven meta-heuristic method based on LSTM network. Sensors. 2022;22(20):7900. doi:10.3390/s22207900.
21. Yu X, Zhang D, Zhu T, Jiang X. Novel hybrid multi-head self-attention and multifractal algorithm for non-stationary time series prediction. Inf Sci. 2022;613(2):541–55. doi:10.1016/j.ins.2022.08.126.
22. Jin Y, Tang C, Liu Q, Wang Y. Multi-head self-attention-based deep clustering for single-channel speech separation. IEEE Access. 2020;8:100013–21. doi:10.1109/ACCESS.2020.2997871.
23. Lin F, Zhang C, Liu S, Ma H. A hierarchical structured multi-head attention network for multi-turn response generation. IEEE Access. 2020;8:46802–10. doi:10.1109/ACCESS.2020.2977471.
24. Guo Q, Huang J, Xiong N, Wang P. MS-pointer network: abstractive text summary based on multi-head self-attention. IEEE Access. 2019;7:138603–13. doi:10.1109/ACCESS.2019.2941964.
25. Xue J, Shen B. A novel swarm intelligence optimization approach: sparrow search algorithm. Syst Sci Control Eng. 2020;8(1):22–34. doi:10.1080/21642583.2019.1708830.
26. Fan Y, Zhang Y, Guo B, Luo X, Peng Q, Jin Z. A hybrid sparrow search algorithm of the hyperparameter optimization in deep learning. Mathematics. 2022;10(16):3019. doi:10.3390/math10163019.
27. Fei T, Wang H, Liu L, Zhang L, Wu K, Guo J. Research on multi-strategy improved sparrow search optimization algorithm. Math Biosci Eng. 2023;20(9):17220–41. doi:10.3934/mbe.2023767.
28. Chai T, Draxler RR. Root mean square error (RMSE) or mean absolute error (MAE)?—arguments against avoiding RMSE in the literature. Geosci Model Dev. 2014;7(3):1525–34. doi:10.5194/gmd-7-1247-2014.
29. Putri TH, Caraka RE, Toharudin T, Kim YH, Chen RC, Gio PU, et al. Fine-tuning of predictive models CNN-LSTM and CONV-LSTM for nowcasting PM2.5 Level. IEEE Access. 2023;12(9):28988–9003. doi:10.1109/ACCESS.2024.3368034.

30. Kim S, Kim H. A new metric of absolute percentage error for intermittent demand forecasts. Int J Forecast. 2016;32(3):66–9. doi:10.1016/j.ijforecast.2015.12.003.
31. Renaud O, Victoria-Feser MP. A robust coefficient of determination for regression. J Stat Plan Inference. 2010;140(7):1852–62. doi:10.1016/j.jspi.2010.01.008.
32. Mukaka MM. Statistics comer: a guide to appropriate use of correlation coefficient in medical research. Malawi Med J. 2012;24(3):69–71. doi:10.2166/wh.2012.000.
33. Kendall MG. A new measure of rank correlation. Biometrika. 1938;30(1–2):81–93. doi:10.1093/biomet/30.1-2.81.
34. Breiman L. Random forest. Mach Learn. 2001;45(1):5–32. doi:10.1023/A:1010933404324.
35. Genuer R, Poggi J, Tuleau-Malot C. Variable selection using random forests. Pattern Recognit Lett. 2010;31(14):2225–36. doi:10.1016/j.patrec.2010.03.014.
36. Marie C, Robin G, Jérôme S. Combining clustering of variables and feature selection using random forests. Commun Stat-Simul Comput. 2021;50(2):426–45. doi:10.1080/03610918.2018.1563145.
37. Castellanos D, Carreto-Vazquez HV, Mashuga VC, Trottier R, Mejia FA, Sam Mannan M. The effect of particle size polydispersity on the explosibility characteristics of aluminum dust. Powder Technol. 2014;254:331–7. doi:10.1016/j.powtec.2013.11.028.