Prediction Of Coal Mine Gas Emission Based On SHSCOA-BiLSTM Model

JIA Jinzhang[1,2], Wang Yixing[1,2*], JIA Peng[1,2,3], and Che Defu[4]

[1] College of Safety Science and Engineering, Liaoning Technical University, Huludao 125105, China
[2] Key Laboratory of Mine Thermodynamic disasters and Control of Ministry of Education, Huludao 125105, China
[3] Ordos Institute of Liaoning Technical University, Ordos 017010, China
[4] School of Resources and Civil Engineering, Northeastern University
[*] Corresponding author. E-mail: 545740846@qq.com

Received: Sep. 24, 2025; Accepted: Dec. 10, 2025

Accurate prediction of coal mine gas emission is crucial for disaster prevention, yet challenging due to complex, non-stationary data and traditional models' tendency to converge to local optima. The present study proposes a novel SHSCOA-BiLSTM model, which integrates an enhanced chimpanzee optimisation algorithm to optimise a bidirectional long short-term memory network. The methodology employs data imputation, principal component analysis, and enhanced global search strategies to tune critical hyperparameters. The model has been validated on real-world data, and it has been demonstrated to significantly outperform existing benchmarks, with a reduction in mean absolute percentage error of 57.18 − 74.10% and mean squared error of 80.16 − 92.35%. The findings indicate that the SHSCOA-BiLSTM model offers a highly accurate and robust instrument for gas emission forecasting, providing a reliable scientific foundation for early warning systems that can significantly enhance proactive safety management and prevent gas-related disasters in coal mines.

Keywords: mine gas outflow prediction, bi-directional long and short-term memory network, chimpanzee optimisation algorithm, hyperparameter optimization, principal component analysis

http://dx.doi.org/10.6180/jase.202607_30.039

1. Introduction

Gas is a core factor that threatens the safe production of coal mines and the stable supply of coal [1], and the prediction of gas outflow is the technical basis for preventing and controlling the occurrence of gas disasters [2]. Research and construction of high-precision prediction model of gas outflow, accurate prediction of gas outflow can reduce the probability of gas disaster and economic losses [3], which is of great significance to the safe production of coal mines [4–6].

In recent years, Artificial Intelligence (AI) technology, especially Artificial Neural Networks (ANN), has deeply penetrated into various fields of engineering, becoming a key tool for improving efficiency and accurate prediction. This trend is particularly evident in recent studies [7–9]. Lunarzewski [10] first used geomechanics and gas emission modelling to predict underground gas emissions from coal mines. Wang Dong [11] used the source and statistical methods to predict and analyse the deviation of gas outflows from different production periods during mine production, which improved the accuracy over traditional methods. Qiu Liming [12] proposed a weighted K-nearest neighbour-based method for identifying abnormal gas concentration and established a sudden risk prediction model based on convolutional neural network. C. Özgen Karacan [13] developed a multilayer perceptron (MLP)-type artificial neural network (ANN) model to simulate the generation of vent holes, and the results showed that the exactitude of the predictions of the two-hidden-layer

model exceeded 90%. The development of Tutak's [14] new method for predicting methane concentrations using MLP networks has advanced the field and provided an opportunity to implement modern solutions across the industry. Al-Zwainy [15, 16] successfully utilised intelligent techniques to predict the residual strength of sustainable self-compacting concrete after elevated temperatures, providing an important basis for building safety and sustainability. Dong Dingwen [17] studied the time series analysis method of gas emission based on Gaussian process regression model for gas emission prediction, and the prediction results were accurate and reliable. Bi Shenghao [18] established the RF-IGWO-SVR model and verified the model prediction effect by using the measured gas emission data of a coal mine. Liang Ma [19] proposed a machine learning method based on artificial neural network (ANN) to predict weather monitoring data, and the computational results of the trained artificial neural network matched well with the measurement results. Yang Li [20] introduced a hybrid model CEEMDAN-OVMDTransformer for predicting water inflow in the Heidaigou coal mine, and the prediction results were stable and accurate. Yuqiang Zheng [21] proposed a quadratic decomposition-based and improved TSMixer model for methane outburst volume prediction, a framework that improves the sensitivity and detection robustness to sudden feature changes through a weighted fusion strategy of global trends and residual anomalies. Suren Rathnayake [22] has built the first predictive model for predicting flow pressures in unconventional wells, which has been successfully validated against extensive industry data.

The prediction methods of gas influx and outflow continue to evolve from basic statistical methods to intelligent algorithm optimisation [23]. Scholars have developed models that integrate multivariate theories and advanced algorithms, significantly improving prediction precision and nonlinear processing capabilities [24]. However, gas outflow is a complex phenomenon driven by multiple interdependent factors-including geological endowment, mining intensity, and ventilation conditions-coupled with a significant time-lag effect. These dynamics collectively exhibit strong non-stationary, cross-scale, and nonlinear characteristics [25].Failure to effectively identify and quantify the coupling relationships among these dynamic factors leads to redundant and covariant model inputs [26]. Consequently, timelag mismatches and noise create a highly nonconvex objective function [27]. This, in turn, predisposes gradient-based training processes to fall into local optima [28]. The practical manifestations of this issue include a sluggish response to peaks and sudden changes, unstable generalization under varying conditions, and ultimately, diminished support for on-site decisions regarding gas extraction and ventilation. Therefore, accurately portraying the coupling relationship and constructing a high-precision prediction model for non-stationary sequences [29, 30] is the key to improve the reliability and practical value of gas hazard warning.

In view of these challenges, this paper presents for the first time a gas influx prediction model that combines the hybrid improved chimpanzee optimisation algorithm (SHSCOA) with BiLSTM. Recent studies have demonstrated the potential of hybrid metaheuristic-DL models. However, these models frequently remain susceptible to premature convergence and suboptimal hyperparameter configuration in highly non-stationary environments, such as gas emission. The SHSCOA model has been developed to address these limitations through a unique improvement strategy. This strategy comprises Sobol-sequence population initialization, a dynamic convergence factor, an adaptive weighting mechanism, and a sine mapping-based bounce-out strategy. This set of enhancements is specifically designed to robustly escape local optima and systematically identify a superior hyperparameter set for the BiLSTM, ensuring a more reliable fit to the complex dynamics of gas flow. The model's superior predictive correctness and stability are rigorously verified against SSA-BiLSTM, PSO-BiLSTM, GABiLSTM and DE-BiLSTM baselines, demonstrating its practical value for gas disaster early warning.

2. Theory and formula

Currently, deep learning neural networks are widely used and recognised in various industries and have achieved remarkable results. In this section, a gas influx prediction model is constructed based on bidirectional long and short-term memory neural networks and their optimisation algorithms.

2.1. Bidirectional long- and short-term memory networks

Long Short-Term Memory (LSTM) [31], as a special kind of Recurrent Neural Network (RNN), plays a key role in solving the problem of long-term sequence dependency. However, LSTM suffers from drawbacks such as limited unidirectional information transfer and insufficient context utilisation.

BiLSTM (Fig. 1) achieves the extraction of forward and backward long-term features of waveform sequences with the help of memory units and bi-directional connections, which improves the correctness and flexibility of prediction, and is able to better explore the intrinsic connection of

individual features, and has better prediction results for gas influx.

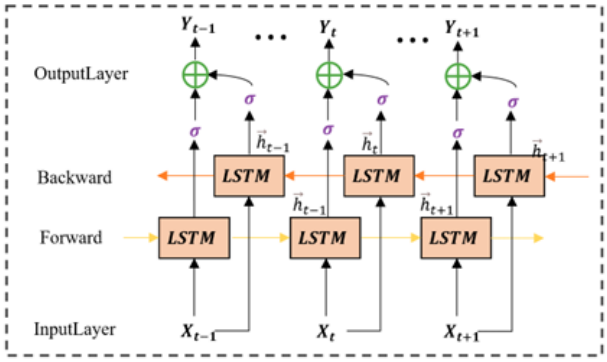


Fig. 1. Structure of bidirectional long short-term memory network

2.2. Basic Chimpanzee Optimisation Algorithm

Chimpanzees with different individual abilities have different hunting tasks, COA algorithm is an intelligent algorithm proposed by simulating the behaviour of chimpanzee group hunting prey [32]. The individuals of the population in the COA algorithm are classified as attackers, evictees, impediments and chasers. The rest of the chimpanzees will assist the attackers when the population hunts, and the chimpanzee group updates its position in the way shown in Fig. 2.

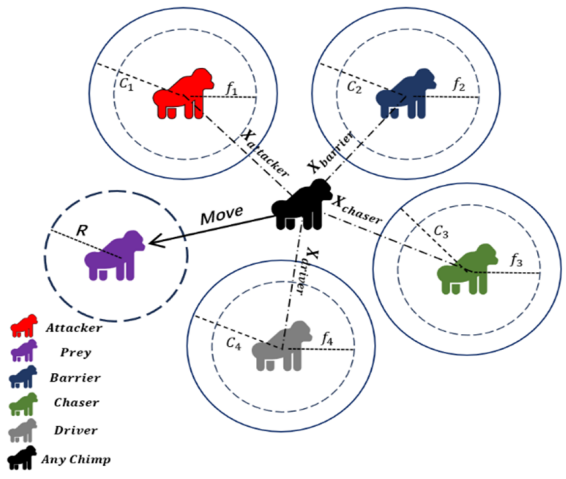


Fig. 2. Update COA Location

Mathematical modelling of chimpanzee expulsion and pursuit of prey is as follows.

$$d = \left| C \cdot X_{\text{prey}}(t) - m \cdot X_{\text{chimp}}(t) \right| \tag{1}$$

$$X_{\text{chimp}}(t+1) = X_{\text{prey}}(t) - a \cdot d \tag{2}$$

Where $X_{\text{prey}}(t)$ is the position of the prey at iteration t; X_{chimp} is the current chimpanzee position vector; d is the distance between the prey and the chimpanzee; and a, m and C are the three position coefficients, which are calculated as follows.

$$a = 4\left(1 - \frac{t}{T_{\max}}\right)(r_1 - 1) \tag{3}$$

$$m = \text{turbulent\_value} \tag{4}$$

$$C = 2r_2 \tag{5}$$

Here, $r_1, r_2 \in [0,1]$ are random vectors; $T_{\max}$ is the maximum number of iterations; f is the convergence factor, which decreases linearly from 2 to 0; $a \in [-f; f]$; m is the chaos factor, which represents the effect of representative motifs on the position of individual chimpanzees; and $C \in [0,2]$ is used to control the effect of the position of the prey on the position of individual chimpanzees. The mathematical model describing how the positions of the chimpanzee individuals of the above four classes are updated after the initialisation of the stochastic population is as follows.

$$X_1 = X_{\text{attacker}} - a_1 \cdot |C_1 \cdot X_{\text{attacker}} - m_1 \cdot X| \tag{6}$$

$$X_2 = X_{\text{barrier}} - a_2 \cdot |C_2 \cdot X_{\text{barrier}} - m_2 \cdot X| \tag{7}$$

$$X_3 = X_{\text{chaser}} - a_3 \cdot |C_3 \cdot X_{\text{chaser}} - m_3 \cdot X| \tag{8}$$

$$X_4 = X_{\text{driver}} - a_4 \cdot |C_4 \cdot X_{\text{driver}} - m_4 \cdot X| \tag{9}$$

$$X(t+1) = (X_1 + X_2 + X_3 + X_4)_{/4} \tag{10}$$

Where $a_{1-4}, m_{1-4}, C_{1-4}$ are the behavioural control coefficients, chaotic mapping inertia factors, and random interference coefficients, respectively, and X_{attacker}, X_{barrier}, X_{chaser} and X_{driver} are the position vectors of the attacker, obstacle, chaser, and repeller, respectively.

2.3. Chimpanzee optimisation algorithm for improved strategies

The novel algorithm presented in this paper is distinct from the conventional COA algorithm in several ways. Primarily, it demonstrates an enhanced capacity to address the deficiencies inherent to the traditional algorithm, such as inadequate population diversity, an absence of directionality in the initial search phase, sluggish convergence, and the stagnation of the search process due to the attainment

of a local optimal solution. The subsequent section will elucidate the enhancement strategy employed by the novel algorithm.

2.4. Sobol Sequence Initialisation Population

The Sobol sequence is a low-discrepancy sequence with short computational period and fast sampling speed. Let the optimal solution take the value range of $[ub, lb]$, and the random number $S_i \subseteq [0,1]$ is generated from the Sobol sequence, therefore, the initial position of the population can be expressed as:

$$X_n = lb + S_i \cdot (ub - lb) \tag{11}$$

As can be seen in Fig. 3, the initial populations generated by the Sobol sequence fall into each range with almost the same amount of individuals, which are more evenly distributed and more widely traversed.

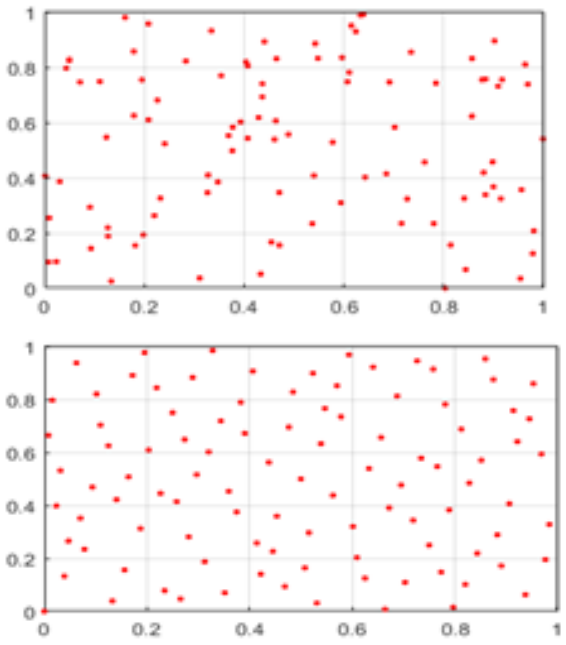


Fig. 3. (a) Stochastic initialisation of the spatial (b) Sobol sequences initialise the spatial Comparison of the spatial distribution of populations with different initializations.

2.5. Weighting factors

The standard COA suffers from a loss of population diversity in later iterations. This occurs because individuals over-converge toward the attacker's position, leading to premature convergence. To address this, we introduce an adaptive weighting factor that provides a more guided update direction. The proposed position update formula is as follows:

$$\begin{cases} x(t+1) = \left(\frac{x_1 f(x_1)}{f(X)} + \frac{x_2 f(x_2)}{f(X)} + \frac{x_1 f(x_1)}{f(X)} + \frac{x_1 f(x_1)}{f(X)}\right)/4 \\ f(X) = f(x_1) + f(x_2) + f(x_3) + f(x_4) \end{cases} \tag{12}$$

2.5.1. *Sine mapping-based automatic bounce mechanism*

In the classical COA algorithm, the algorithm will not be able to find the global optimal solution if the attackers in the population converge too early. In this paper, we design an automatic jump-out mechanism based on Sine mapping, i.e., if the attacker's position is unchanged due to convergence during the successive iterative loops of the algorithm, the automatic jump-out mechanism is activated and the attacker jumps according to Eq. (13).

$$X_{\text{chaos}}(g+1) = \mu_3 \cdot 0.5 \cdot \sin(\pi \cdot X_{\text{attacker}}(g)) \tag{13}$$

where μ_3 is the mapping parameter. After the jump-out mechanism is executed, based on , it is judged whether the attacker's updated position is retained or not, i.e., the regular position is retained when and only when its adaptation is strictly better than the chaotic position, otherwise the chaotic position is adopted.

$$X(g+1) = \begin{cases} X(t+1), & \text{Fitness}(X(t+1)) \\ & < \text{Fitness}(X_{\text{chaos}}(t+1)), \\ X_{\text{chaos}}(g+1), & \text{Fitness}(X(t+1)) \\ & \geq \text{Fitness}(X_{\text{chaos}}(t+1)). \end{cases} \tag{14}$$

2.6. Optimisation algorithm performance testing

In order to verify the global exploration ability and convergence stability of the optimised algorithm, Rastrigin function and Rosenbrock function are selected for the comparison of algorithm optimisation [33], and the function images are shown in Fig. 5.The corresponding mathematical expressions are as follows:

$$f_1(\mathrm{x}) = 20 + x_1^2 + x_2^2 - 10\cos(2\pi x_1) - 10\cos(2\pi x_2) \tag{15}$$

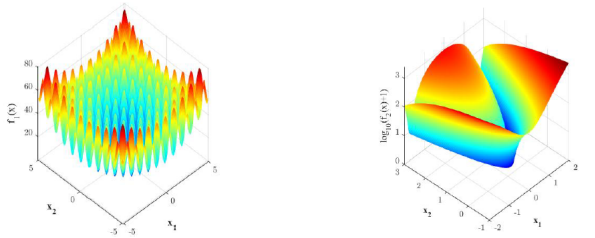
$$f_2(\mathrm{x}) = 100\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2 \tag{16}$$

Two test functions, $f_1(\mathrm{x})$ and $f_2(\mathrm{x})$, are selected to evaluate the optimality seeking performance of SHSCOA, ICOA, COA and WOA. In order to explore their global exploration ability and convergence stability, various cluster sizes are set to 30 , the maximum number of iterations is 500 , and

Algorithm 1 SHSCOA: Strategic Hybrid Sobol-Chimpanzee Optimization Algorithm

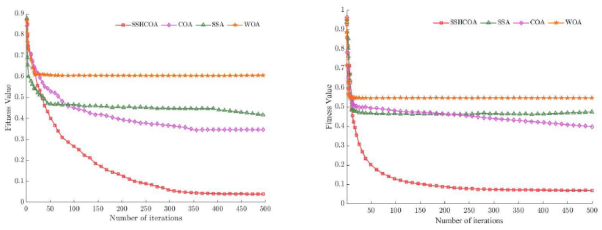
Require: Population size N, max iterations T_{max}, bounds $[lb, ub]$, parameter μ_3
Ensure: Global best X_{best}, best fitness f_{best}
Initialize: $t \leftarrow 0$, $stagnation\_counter \leftarrow 0$, $prev\_best \leftarrow \infty$
Sobol Initialization:
for $i = 1$ to N **do**
  $X_i \leftarrow lb + sobol\_seq[i] \cdot (ub - lb)$ ▷ Eq. (11)
end for
while $t < T_{max}$ **do**
  Evaluate fitness for all individuals
  Update leaders: $X_{attacker}, X_{barrier}, X_{chaser}, X_{driver}$
  for each X_i in *population* **do**
    Calculate a, m, C coefficients ▷ Eqs. (3)-(5)
    Compute candidate positions X_1, X_2, X_3, X_4 ▷ Eqs. (6)-(9)
    Adaptive Weighting:
    $w_k \leftarrow f_k / \sum f_{leaders}$ ▷ Eq. (12)
    $X_{new} \leftarrow \sum w_k \cdot X_k$
    if Fitness$(X_{new}) < fitness[i]$ **then**
      Update X_i and $fitness[i]$
    end if
  end for
  Sine Escape Mechanism:
  if stagnation detected **then**
    $X_{chaos} \leftarrow \mu_3 \cdot 0.5 \cdot \sin(\pi \cdot X_{attacker})$ ▷ Eq. (13)
    if Fitness$(X_{chaos}) < f_{att}$ **then** ▷ Eq. (14)
      $X_{attacker} \leftarrow X_{chaos}$, reset stagnation counter
    end if
  end if
  Update global best if improved
  $t \leftarrow t + 1$
end while
return X_{best}, f_{best}

Fig. 4. SHSCOA algorithm pseudo-code



(a) Rastrigin (b) Rosenbrock

Fig. 5. function $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$



(a) Rastrigin convergence (b) Rosenbrock convergence

Fig. 6. Location of the KVMRT project

the convergence curves of each test function are shown in Fig. 6.

From these two convergence graphs, it can be seen that the SHSCOA algorithm is superior to the other three algorithms in terms of global exploration ability and convergence stability.

2.7. Eigenvalue Screening

An experimental study of mine gas prediction was carried out based on the actual collected daily gas outflow data for 600 days from July 2021 to February 2023 for a mine of Shanxi Huayang New Material Technology Group Co.

The raw data contains 10 characteristic variables related to the gas outflow from the mine, namely, the total air return volume of the mine X_1, the daily output of the mine X_2, the mining depth of the backing face X_3, the backing rate of the backing face X_4, the total thickness of the neighbouring layer X_5, the average spacing of the neighbouring layer from the mined layer X_6, the average depth of the neighbouring layer X_7, the length of the excavated roadway X_8, and the daily excavation advance at the excavation face X_9. Mine gas extraction X_{10}. absolute gas influx Y is the output index. Table 1 shows some of the data of the mine, there are missing values due to sensor failure, communication abnormality, and omission of shift report.

2.7.1. *Multiple filling*

The sample data table in this experiment contains some missing values. Simply deleting the samples with missing data is not a reasonable approach. Instead, these missing values can be estimated using interpolation based on adjacent observations.

Multiple Imputation Methods (MI) is a technique for handling missing values based on model estimation and repeated simulations. When dealing with a large number of missing values, MI is a commonly chosen method. First, 3-10 complete datasets are created by estimating the model. Statistical analyses are then performed on each imputed dataset, and the results are merged using Rubin's rule to obtain the final estimate.

Using SPSS software to perform multiple imputation on the sample data, the analysis revealed that the missing rate of the sample data was 5.3%. The number of imputation iterations was set to 5, with a nonlinear model (Nonlinear Models) as the estimation model. The median values from the 5 imputation iterations were selected as the imputed values to obtain the complete data sample, as shown in Table 2.

2.7.2. *Data Dimension Reduction*

To improve the efficiency of the model in processing data, the interpolated data samples were subjected to dimensionality reduction using SPSS software, resulting in the principal component analysis scatter plot and component matrix. As shown in Fig. 7, the cumulative variance contribution rate of the first four principal components reached

Table 1. Partial data

No.	X_1' ($m^3\ min^{-1}$)	X_2 (t/d)	X_3 (m)	X_4' (%)	X_5 (m)	X_6' (m)	X_7' (m)	X_8 (m)	X_9 (m/d)	X_{10}' ($m^3\ min^{-1}$)	Y' ($m^3\ min^{-1}$)
1	5380	4899	1370	88.67	12.14	100	385	588	5	9.22	2.32
2	5380	5146	1370	–	12.14	100	385	643	5	9.61	2.21
3	5380	4676	1370	93.07	–	100	–	672	4	–	–
4	5380	4270	1370	89.93	12.31	100	420	–	4	–	2.43
5	5380	3661	–	–	12.31	100	–	712	5	9.22	2.53
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
598	5290	5633	1450	87.32	15.43	95	–	831	8	–	7.38
599	5290	6375	1450	–	–	95	680	–	11	10.05	7.19
600	5290	5348	1450	90.72	15.43	95	680	955	6	9.66	7.66

Table 2. Partial data after filling

No.	X_1' ($m^3\ min^{-1}$)	X_2 (t/d)	X_3 (m)	X_4' (%)	X_5 (m)	X_6' (m)	X_7' (m)	X_8 (m)	X_9' (m/d)	X_{10}' ($m^3\ min^{-1}$)	Y' ($m^3\ min^{-1}$)
1	5380	4899	1370	88.67	12.14	100	385	588	5	9.22	2.32
2	5380	5146	1370	86.98	12.89	100	385	643	5	9.61	2.21
3	5380	4676	1370	93.07	13.01	100	414	672	4	8.46	2.37
4	5380	4270	1370	89.93	13.31	100	420	699	4	8.46	2.43
5	5380	3661	1370	91.26	14.18	100	425	712	5	9.22	2.53
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
598	5290	5633	1450	87.32	17.22	95	674	831	8	8.51	7.38
599	5290	6375	1450	88.12	15.43	95	680	890	11	10.05	7.19
600	5290	5348	1450	90.72	16.93	95	680	955	6	9.66	7.66

86.5% (with contribution rates of 42.2%, 18.4%, 15.2%, and 10.7%, respectively). Based on the selection principles of principal component analysis, when the cumulative contribution rate exceeds 85%, the main features of the sample data can be maximally retained. Therefore, the first four sample features were selected. Based on the component matrix load analysis, principal components 1, 2, 3, and 4 are primarily influenced by X_2, X_9, X_{10} and X_3 respectively. Replacing the original data with the four sample data, Table 3 and Table 4 present the component matrix and the reduceddimension sample data, respectively.

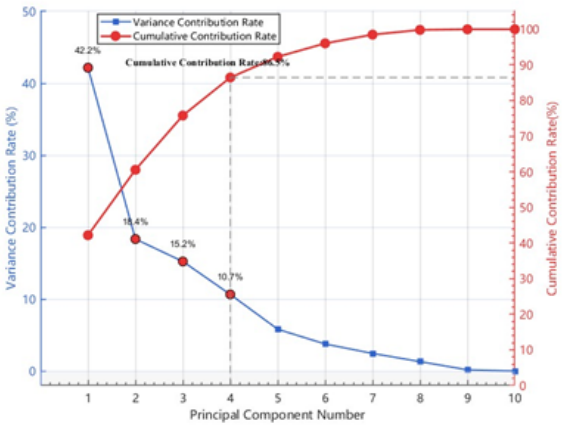


Fig. 7. Principal component analysis gravel plot

2.8. Model evaluation indicators

To accurately analyse the predictive fidelity of the constructed SHSCOA-BiLSTM model, the mean absolute error (MAE), root mean square error (RMSE), coefficient of determination (R^2), mean absolute percentage error (MAPE), and mean square error (MSE) were selected as model evaluation metrics. The calculation process is as follows:

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|\hat{y}_i - y_i| \tag{17}$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2} \tag{18}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{m}(\hat{y} - y_i)^2} \tag{19}$$

$$MAPE = \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{20}$$

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 \tag{21}$$

m is the sample size; y_i is the true value of the i-th sample; $\hat{y}_i$ is the predicted value of the i-th sample; the smaller the values of MAE, RMSE, and MSE, the better the performance.

2.9. Model Prediction Process

Improving the predictive reliability of the BiLSTM model hinges on the reasonable configuration of its core parameters. Traditional empirical selection methods are highly random and lack guidance, often leading to fluctuations in

Table 3. Component matrix

	PC_1	PC_2	PC_3	PC_4
X_1	-0.12	0.91	0.15	0.03
X_2	0.93	0.08	-0.10	0.05
X_3	0.05	0.02	-0.08	0.94
X_4	0.45	0.15	0.48	0.12
X_5	0.34	0.23	0.25	-0.13
X_6	0.25	0.18	0.30	0.08
X_7	0.38	0.12	0.33	0.42
X_8	-0.21	-0.15	0.13	0.10
X_9	0.20	0.28	0.64	-0.08
X_{10}	-0.31	0.20	0.84	0.18

Table 4. Partial sample data after dimensionality reduction

No.	PC_1	PC_2	PC_3	PC_4	Y
1	-1.32	1.58	0.29	-0.16	2.32
2	-1.60	1.23	0.29	-0.19	2.21
3	-1.41	1.64	0.08	0.38	2.37
4	-1.18	1.62	0.34	0.62	2.43
⋮	⋮	⋮	⋮	⋮	⋮
599	-1.15	1.35	0.50	0.95	7.19
600	-0.96	0.92	0.32	0.03	7.66

model performance. While grid search can systematically screen for optimal parameter combinations, it is limited by its significant computational overhead and inefficiency. To address this, this study introduces an improved chimpanzee optimisation algorithm (SHS) to optimise five key parameters, including the number of neurons in the two BiLSTM layers, the Dropout rate, and the batch size. The number of neurons is related to model complexity and memory capacity, the Dropout mechanism enhances robustness, and the batch size affects training efficiency and stability. The specific algorithm flow is shown in Fig. 8.

3. Results and discussion

3.1. Experimental parameter settings

The SHSCOA-BiLSTM model established in this paper uses the Adam gradient descent algorithm for parameter updating. To effectively prevent model overfitting, the learning rate decay factor is adjusted to 0.05 . SHSCOA is used to optimise the parameters in the BiLSTM model, and the optimisation parameters are set as shown in Table 5.

The fitness convergence curve is shown in Fig. 9. The SHSCOA algorithm reached the optimal fitness value at the 7th iteration, with a minimum root mean square error of 0.017051. The optimal values for Neurons1, Dropout1, Neurons2, Dropout2, and Batch size were 11, 0.052, 157, 0.032, and 14 , respectively.

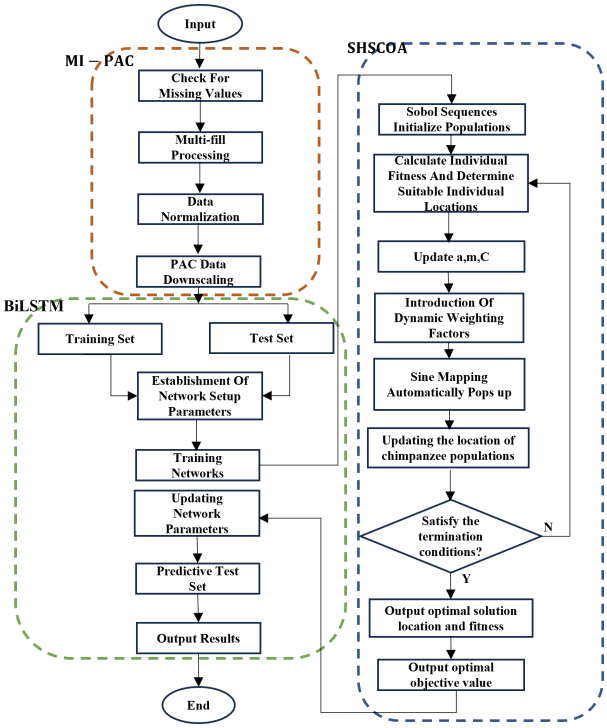


Fig. 8. SHSCOA -BiLSTM model prediction process

3.2. Prediction Results and Analysis

After debugging, the pre-processed gas outflow data set was divided into a 7 : 3 ratio, with 420 data points in the training set and 180 data points in the test set. To verify

Table 5. SHSCOA initial parameter setting

Population size	Number of iterations	Neuron s1	Drop out1	Neuro ns2	Dro pout 2	Batch size
55	100	[10, 200]	[0, 1]	[10, 200]	[0, 1]	[1, 100]

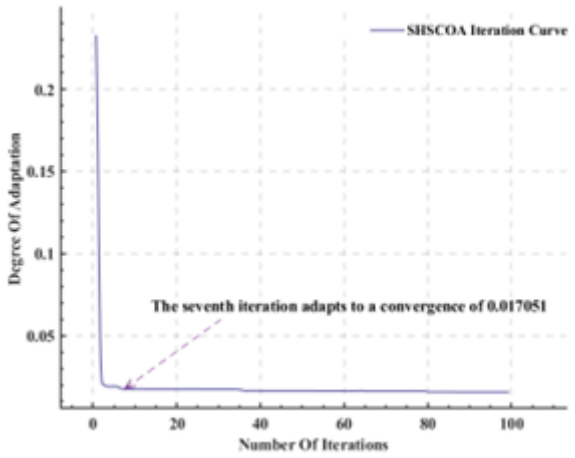


Fig. 9. Adaptation convergence curve

the rigor of the SHSCOA-BiLSTM model's prediction results, four control models, SSA-BiLSTM, PSO-BiLSTM,GA-BiLSTMand DE-BiLSTM, were introduced for experimentation. The hyperparameter selection is shown in Table 6. The prediction results of each model are shown in Fig. 10.

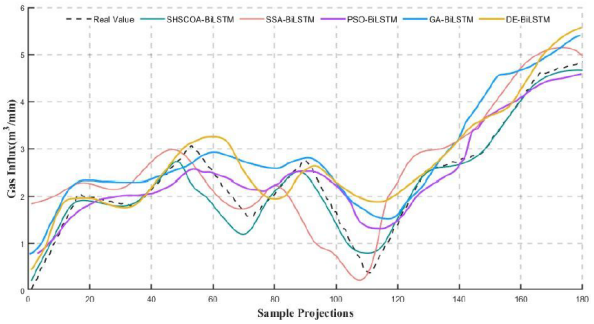


Fig. 10. Comparison of predicted and true values of each model(The black dashed line

4. Represents the true values; colored lines show predictions from different models.)

As demonstrated in Fig. 10, the SSA-BiLSTM model exhibits suboptimal performance in capturing the long-term dependence of sequences. However, this deficiency is effectively mitigated by the global exploration capability of the sparrow search algorithm, which alleviates the gradient decay problem. Nevertheless, the model still experiences a slight lag in handling extreme mutation points.The PSO-BiLSTM model is susceptible to attaining a local optimum due to the particle swarm optimization, and its fitness to non-smooth sequences is weaker, resulting in evident deviations and oscillations of the predictions at the peaks and valleys. The GA-BiLSTM model is limited by the precocious convergence property of the genetic algorithm. This results in a degradation of its ability to fit complex fluctuation intervals, as evidenced by a sluggish response to mutation characteristics and excessive smoothing.The DE-BiLSTM model exhibits the lowest fitting ability due to the lack of efficiency of the differential evolution algorithm in exploring high-dimensional parameter space. This leads to significant prediction deviation, particularly in the presence of noise disturbances.SHSCOA-BiLSTM employs a hybrid improvement strategy that significantly enhances global exploration and local mining ability in hyper-parameter space. This is achieved through the synergistic effect of Sobol sequence initialisation, dynamic convergence factor and sinusoidal mapping jump-out mechanism. Consequently, the network structure, learning rate and other parameters are better fitted to the characteristics of non-smooth sequences. The model demonstrates the highest degree of matching precision and the smallest hysteresis effect at both peaks and valleys. Furthermore, it exhibits enhanced noise resistance, and its prediction curve exhibits the closest proximity to the measured value. The evaluation indexes of the prediction effect of each model are shown in Fig. 11.

Fig. 11 shows that all error indicators of the SHSCOA-BiLSTM model are significantly reduced and the prediction accuracy is significantly improved compared with the SSABiLSTM, PSO-BiLSTM, GA-BiLSTM and DE-BiLSTM models. Specifically, the mean square error (MSE) of SHSCOA-BiLSTM is reduced by 56.65%, 79.84%, 82.09%, and 84.62% compared with the above four types of models, respectively; the mean absolute error (MAE) is reduced by 33.68%, 55.93%, 62.92%, and 67.71%, respectively; and the mean absolute percentage error (MAPE) is reduced by 43.72%, 54.67%, 65.27% and 67.00%. Meanwhile, SHSCOA-BiLSTM reaches 0.9925 on the R^2 metric, which is improved by 3.03%, 4.32%, 4.96% and 6.53% compared with the comparison model, demonstrating the optimal goodness-of-fit. In addition, the model also outperforms other compara-

Table 6. Hyperparameter selection for different models

Model	Learning rate	Number of hidden layer nodes	Regularisation parameter
SSA -BiLSTM	0.05	50	0.0351
PSO-BiLSTM	0.02	40	0.0531
GA-BiLSTM	0.04	50	0.0441
DE-BiLSTM	0.06	70	0.0342

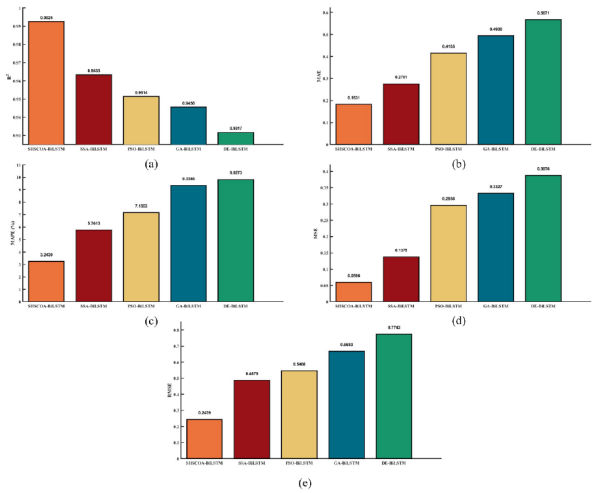


Fig. 11. Summary of integrated evaluation indicators for each model(Except for the coefficient of determination ($\mathbf{R^2}$), lower values indicate better model performance

tive algorithms in terms of convergence speed and anti-interference ability, which fully verifies its high reliability and resilience in predicting gas outflow in a complex multifactor mine environment.

4.1. Statistical Significance Analysis

In order to statistically verify whether the performance improvement of the SHSCOABiLSTM model is significant and not merely a random occurrence, we performed paired t-tests on the prediction errors (MAE, MSE) between this model and each benchmark model (SSABiLSTM, PSO-BiLSTM, GA-BiLSTM, DE-BiLSTM). Each model was executed 30 times, employing distinct random seeds to obtain the distribution of the error metrics.The null hypothesis (H_0) posits that there is no statistically significant difference between the means of the two comparison groups, while the alternative hypothesis (H_1) suggests that there is a significant difference. The results of the t -test (significance level $\alpha = 0.05$) are summarised in Table 7.

As demonstrated in Table 7, all p-values are significantly less than 0.05, indicating that the performance enhancements exhibited by SHSCOA-BiLSTM are statistically substantial. This validates the hypothesis that the observed reductions in MAE and MSE are not attributable to random variations, but rather are a consequence of the efficacy of the proposed model.

4.2. Application analysis of engineering examples

4.2.1. Overview of the mine

In order to verify the universality and stability of the SHSCOA-BiLSTM gas outflow early warning model, an analysis is conducted on a mine of Shanxi Lu'an Group Luning Coal Company Limited. The mine field extends for 3.8 km from north to south and 2.1 km from east to west, with a maximum reserve calculation area of 9.32 km^2. The mine is categorised as a high-gas mine, exhibiting an absolute gas outflow of 38.42 m^3/min and a relative gas outflow of 11.73 m^3/min. The initial gas content was determined to be approximately 6.9 m^3/t, while the mean coal seam thickness was recorded as 2.67 m . The mean dip of the seam was established as 5.21°, and the mean depth of excavation was determined to be 336.15 m . The mean length of the working face was found to be 142.62 m , with the mean extraction rate recorded as 84.9%. The mean gas content of the neighbouring seams was determined to be 4.43 m^3/t, and the mean thickness of the neighbouring seams was determined to be 3.62 m . The mean lithology of the seams was found to be 4.33 . The mean thickness of the adjacent seams was found to be 3.62 m , and the mean lithology of the inter-seams was determined to be 4.33.

4.2.2. Modelling analysis

In summary, there is a danger of gas leakage in the mining process, so it is necessary to predict the degree of danger to ensure the safe production of mines. Of the 400 sets of sample data collected, the first 280 sets were used as the training set and the last 120 sets were used as the prediction set, and the four principal components were extracted by MI-PCA processing, and the SSA-BiLSTM, PSO-BiLSTM, GA-BiLSTM, DE-BiLSTMBiLSTM, GA-BiLSTM, DE-BiLSTM, and SHSCOA- BiLSTM models were compared and analysed with the same model parameter settings as above, and the results of the evaluation metrics of the obtained prediction models are shown in Table 8. From Table 8, it is clear that: the MAE of the SHSCOABiLSTM model based on the gas outflow is the same in the test set as in the test

Table 7. SHSCOA-BiLSTM paired t-test results with the remaining models

Comparison Model	Metric	t-value	p -value	Significant?
SHSCOA-BiLSTM	MAE	8.32	<0.001	Yes
vs SSA-BiLSTM	MSE	9.14	<0.001	Yes
SHSCOA-BiLSTM	MAE	8.43	<0.001	Yes
vs PSO-BiLSTM	MSE	9.27	<0.001	Yes
SHSCOA-BiLSTM	MAE	9.33	<0.001	Yes
vs GA-BiLSTM	MSE	9.73	<0.001	Yes
SHSCOA-BiLSTM	MAE	10.22	<0.001	Yes
vs DE-BiLSTM	MSE	11.35	<0.001	Yes

Table 8. Comparison of evaluation indicators for different forecasting models

Model	R^2		RMSE		MSE		MAE		MAPE/%	
	Train	Tset	Train	Tset	Train	Tset	Train	Tset	Train	Tset
DE-BiLSTM	0.951	0.936	0.439	0.539	0.332	0.329	0.441	0.423	5.762	8.269
GA-BiLSTM	0.957	0.943	0.397	0.473	0.268	0.317	0.398	0.396	5.213	7.834
PSO-BiLSTM	0.965	0.951	0.351	0.392	0.186	0.295	0.237	0.341	4.221	6.742
SSA-BiLSTM	0.973	0.963	0.241	0.331	0.132	0.231	0.216	0.251	3.914	6.131
SHSCOA-BiLSTM	0.983	0.978	0.224	0.251	0.067	0.193	0.176	0.187	3.117	5.196

Table 9. Comparison of real and predicted values of different prediction models

No.	Real value	SHSCOA-BiLSTM		SSA-BiLSTM		PSO-BiLSTM		GA-BiLSTM		DE-BiLSTM	
		Predicted value	Relative error (%)	Predicted value	Relative error (%)	Predicted value	Relative error (%)	Predicted value	Relative error (%)	Predicted value	Relative error (%)
1	3.52	3.711	5.43	3.809	8.22	3.206	8.93	3.839	9.07	3.865	9.79
2	3.62	3.698	2.16	3.807	5.16	3.346	7.58	4.066	12.31	3.130	13.53
3	4.13	4.431	7.32	4.525	9.38	4.551	10.21	3.861	6.43	3.923	5.18
4	4.60	4.786	4.05	4.769	3.67	4.423	3.85	4.713	2.46	4.431	3.67
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
117	4.94	5.417	9.66	5.452	10.36	5.424	9.79	4.563	7.64	5.398	9.28
118	5.25	5.634	7.31	5.359	2.08	5.463	4.06	5.477	4.32	6.060	15.43
119	7.26	7.349	1.22	7.392	1.82	6.802	6.31	7.873	8.45	8.474	16.72
120	7.80	7.929	1.65	7.974	2.23	7.903	1.32	8.786	12.64	8.334	6.85

set. In the test set, the R^2 values of the predicted models are improved by 1.56%, 2.84%, 3.71%, and 4.49%, the RMSEs are reduced by 24.17%, 35.97%, 46.93%, and 53.43%, the MSEs are reduced by 16.45%, 34.58%, 39.12%, and 41.34%, and the MAEs are reduced by 25.50%, 45.16%, 52.78%, and 55.79%, and MAPE by 15.25%, 22.93%, 33.67%, and 37.16%, respectively, indicating that the prediction accuracy of this model is higher than that of SSA-BiLSTM, PSOBiLSTM, GA-BiLSTM, and DE-BiLSTM models, respectively.

The results of the comparison between the true values and the predicted values of different prediction models are shown in Table 9, from which it can be seen that the SHSCOA-BiLSTM prediction model has the smallest error, with its relative error ranging from $1.22\% - 9.66\%$, and the DE-BiLSTM prediction model has the largest error, with its relative error ranging from 3.67% to 1672%. The results show that the predicted values of SHSCOA-BiLSTM model are closer to the real values, while the rest of the models have fitting phenomenon in the prediction process, which leads to large errors. This shows that the SHSCOA-BiLSTM model has better High degree of conformity and generalisation than the rest of the models in practical applications.

5. Conclusion

1. The integration of Multiple Imputation and Principal Component Analysis efficiently handled missing data and reduced dimensionality. This process reconstructed four principal control factors, which enhanced input information density, minimized overfitting risk, and improved model training efficiency.

2. The proposed SHSCOA algorithm, enhanced with Sobol sequence initialization, a dynamic convergence factor, and a sine mapping bounce-out mechanism, demonstrated stronger global exploration and convergence stability than standard algorithms. This effectively prevented premature convergence to local optima during the hyperparameter tuning of the BiLSTM network.

3. Rigorous validation on real-world mine data demonstrated that the SHSCOA-BiLSTM model significantly outperforms benchmark models (SSA-BiLSTM, PSO-BiLSTM, GABiLSTM, DE-BiLSTM). It achieved marked reductions in MAE, MSE, and MAPE, and a higher R^2, showing superior accuracy in tracking peaks and sudden changes with minimal lag. Statisti-

cal tests confirmed the significance of these improvements.

4. The model's performance depends on data quality and its computational complexity is non-trivial. Future work will focus on developing lightweight versions for real-time application and exploring online learning for dynamic adaptation to new data.

6. Declaration of competing interest

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

7. Data availability

Data will be made available on request.

8. Acknowledgement

This research is conducted with financial support from the National Natural Science Foundation of China (No. 52174183 and 52374203).

References

[1] Y. Wu, M. Chen, K. Wang, and G. Fu, (2019) *"A dynamic information platform for underground coal mine safety based on internet of things"* **Safety Science *113***: 9–18. DOI: 10.1016/j.ssci.2018.11.003.

[2] F. Wang, P. Zhang, B. Cui, Z. Sun, and K. Zhang, (2021) *"Research progress of disaster factors and a prevention alarm index of coal and gas outbursts"* **Arabian Journal of Geosciences *14***: 2042. DOI: 10.1007/s12517-021-07540-2.

[3] M. Petkovic, Y. Chen, I. Gamrath, U. Gotzes, N. S. Hadjidimitrou, J. Zittel, X. Xu, and T. Koch, (2022) *"A hybrid approach for high precision prediction of gas flows"* **Energy Systems *13***: 383–408. DOI: 10.1007/s12667-021-00466-4.

[4] S. Li, M. You, D. Li, and J. Liu, (2022) *"Identifying coal mine safety production risk factors by employing text mining and Bayesian network techniques"* **Process safety and environmental protection *162***: 1067–1081. DOI: 10.1016/j.psep.2022.04.054.

[5] X. Li, Z. Cao, and Y. Xu, (2025) *"Characteristics and trends of coal mine safety development"* **Energy Sources, Part A: Recovery, Utilization, and Environmental Effects *47***: 2316–2334. DOI: 10.1080/15567036.2020.1852339.

[6] C. Zhang, P. Wang, E. Wang, D. Chen, and C. Li, (2023) *"Characteristics of coal resources in China and statistical analysis and preventive measures for coal mine accidents"* **International journal of coal science & technology *10***: 22. DOI: 10.1007/s40789-023-00582-9.

[7] M. S. K. Al-Marsomia and F. M. S. Al-Zwainya, (2023) *"Journal of Project Management"* **Journal of Project Management *8***: 119–132. DOI: 10.5267/J.JPM.2022.11.002.

[8] S. H. R. Aldhamad, R. Maya, S. F. M. Alazawy, and F. M. Alzwainy, (2024) *"Forecasting models for time and cost performance predicting of infrastructural projects"* **Civil and Environmental Engineering *20***: 1024–1039. DOI: 10.2478/cee-2024-0074.

[9] F. Al-Zwainy, M. G. Al-khazrajy, N. M. Hussein, S. Mohamed, M. M. Sarhan, T. J. Al-Musawi, and G. Hayder, (2024) *"Utilizing Artificial Neural Networks for Predictive KPI Analysis in Bridge Projects."* **Journal of Computational Analysis & Applications *33***:

[10] L. L. W. Lunarzewski, (1998) *"Gas emission prediction and recovery in underground coal mines"* **International Journal of Coal Geology *35***: 117–145. DOI: 10.1016/S0166-5162(97)00007-4.

[11] W. Dong and D. Hong. "Analysis the Influence Factors and Deviation of Gas Outflow in Resumed Mine". In: *IOP Conference Series: Earth and Environmental Science*. ***474***. IOP Publishing. 2020, 042031. DOI: 10.1088/1755-1315/474/4/042031.

[12] L. Qiu, Y. Peng, and D. Song, (2022) *"Risk prediction of coal and gas outburst based on abnormal gas concentration in blasting driving face"* **Geofluids *2022***: 3917846. DOI: 10.1155/2022/3917846.

[13] C. Ö. Karacan, (2009) *"Forecasting gob gas venthole production performances using intelligent computing methods for optimum methane control in longwall coal mines"* **International Journal of Coal Geology *79***: 131–144. DOI: 10.1016/j.coal.2009.07.005.

[14] M. Tutak and Krenicky, (2024) *"Predicting methane concentrations in underground coal mining using a multilayer perceptron neural network based on mine gas monitoring data"* **Sustainability *16***: 8388. DOI: 10.3390/su16198388.

[15] F. M. Al-Zwainy, S. A. Salih, M. R. Aldikheeli, et al., (2021) *"Prediction of residual strength of sustainable self-consolidating concrete exposed to elevated temperature using artificial intelligent technique"* **International Journal of Applied Science and Engineering *18***: 1–15. DOI: 10.6703/IJASE.202106_18(2).012.

[16] J. A. Al-Somaydaii, A. T. Albadri, and F. M. Al-Zwainy, (2024) *"Hybrid approach for cost estimation of sustainable building projects using artificial neural networks"* **Open Engineering *14***: 20220485. DOI: 10.1515/eng-2022-0485.

[17] D. Dong, (2012) *"Mine gas emission prediction based on Gaussian process model"* **Procedia Engineering *45***: 334–338. DOI: 10.1016/j.proeng.2012.08.167.

[18] S. Bi, L. Shao, Z. Qi, Y. Wang, and W. Lai, (2023) *"Prediction of coal mine gas emission based on hybrid machine learning model"* **Earth Science Informatics *16***: 501–513. DOI: 10.1007/s12145-022-00894-5.

[19] L. Ma, C. Huang, Z.-S. Liu, K. A. Morin, M. Aziz, and C. Meints, (2020) *"Artificial neural network for prediction of full-scale seepage flow rate at the equity silver mine"* **Water, Air *231***: 179. DOI: 10.1007/s11270-020-04541-x.

[20] Y. Li, Q. Wu, and F. Lei, (2025) *"Mine Water Inflow Prediction Using a CEEMDAN-OVMD-Transformer Model"* **Applied Sciences *15***: 9710. DOI: 10.3390/app15179710.

[21] Q. Zheng, C. Li, B. Yang, Z. Yan, and Z. Qin, (2025) *"A Method for Predicting Coal-Mine Methane Outburst Volumes and Detecting Anomalies Based on a Fusion Model of Second-Order Decomposition and ETO-TSMixer"* **Sensors *25***: 3314. DOI: 10.3390/s25113314.

[22] S. Rathnayake, A. Rajora, and M. Firouzi, (2022) *"A machine learning-based predictive model for real-time monitoring of flowing bottom-hole pressure of gas wells"* **Fuel *317***: 123524. DOI: 10.1016/j.fuel.2022.123524.

[23] A. S. Hati, P. Kumar, et al., (2023) *"An adaptive neural fuzzy interface structure optimisation for prediction of energy consumption and airflow of a ventilation system"* **Applied Energy *337***: 120879. DOI: 10.1016/j.apenergy.2023.120879.

[24] Y. Wang, Y. Si, B. Huang, and Z. Lou, (2018) *"Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008-2017"* **The Canadian Journal of Chemical Engineering *96***: 2073–2085. DOI: 10.1002/cjce.23249.

[25] S. Rathnayake, A. Rajora, and M. Firouzi, (2022) *"A machine learning-based predictive model for real-time monitoring of flowing bottom-hole pressure of gas wells"* **Fuel *317***: 123524. DOI: 10.1016/j.fuel.2022.123524.

[26] S.-H. Wu, Y.-C. Zhu, Z.-H. Pan, and H. Di, (2025) *"An efficient Expectation-Maximization algorithm for Bayesian operational modal analysis with physics-data fusion model"* **Mechanical Systems and Signal Processing *237***: DOI: 10.1016/j.ymssp.2025.113144.

[27] Y. Li and S. A. Vorobyov, (2017) *"Fast algorithms for designing unimodular waveform (s) with good correlation properties"* **IEEE Transactions on Signal Processing *66***: 1197–1212. DOI: 10.1109/TSP.2017.2787104.

[28] S. Yang, Y. Tian, C. He, X. Zhang, K. C. Tan, and Y. Jin, (2021) *"A gradient-guided evolutionary approach to training deep neural networks"* **IEEE Transactions on Neural Networks and Learning Systems *33***: 4861–4875. DOI: 10.1109/TNNLS.2021.3061630.

[29] H. K. Risan, F. M. Serhan, and A. A. Al-Azzawi. "Management of a typical experiment in engineering and science". In: *AIP Conference Proceedings*. ***2864***. AIP Publishing LLC. 2024, 050001. DOI: 10.1063/5.0186079.

[30] F. Al-Zwainy, E. K. Abdalkarim, W. K. Majeed, E. S. Huseen, and H. S. Jari, (2024) *"Development Artificial Neural Network (ANN) computing model to analyses men's 100-meter sprint performance trends."* **Fizjoterapia Polska** (2): DOI: 10.56984/8ZG5608M3Q.

[31] A. Sherstinsky, (2020) *"Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network"* **Physica D: Nonlinear Phenomena *404***: 132306. DOI: 10.1016/j.physd.2019.132306.

[32] H. Jia, K. Sun, W. Zhang, and X. Leng, (2022) *"An enhanced chimp optimization algorithm for continuous optimization domains"* **Complex & Intelligent Systems *8***: 65–82. DOI: 10.1007/s40747-021-00346-5.

[33] M. Naser, M. K. Al-Bashiti, A. T. G. Tapeh, A. Naser, V. Kodur, R. Hawileh, J. Abdalla, N. Khodadadi, A. H. Gandomi, and A. D. Eslamlou, (2025) *"A review of benchmark and test functions for global optimization algorithms and metaheuristics"* **Wiley Interdisciplinary Reviews: Computational Statistics *17***: e70028. DOI: 10.1002/wics.70028.