RESEARCH ARTICLE

Research on Gas Outburst Prediction Model Based on Multiple Strategy Fusion Improved Snake Optimization Algorithm With Temporal Convolutional Network

HUA FU[1], HAOFAN SHI[1], YAOSONG XU[1], AND JINGYU SHAO[2], (Member, IEEE)
[1]Faculty of Electrical and Control Engineering, Liaoning Technical University, Huludao 125105, China
[2]Baidu Inc., Beijing 100000, China

Corresponding author: Haofan Shi (1458336228@qq.com)

ABSTRACT A gas outburst prediction model based on multiple strategy fusion and improved snake optimization algorithm (MFISO) and temporal convolutional network (TCN) is proposed to address the problems of low accuracy of deep learning prediction models for gas outburst in underground mines. By adopting the phase space reconstruction method, the time series of multiple complex influencing factors related to gas outburst were reconstructed and used as model inputs. Sine chaos mapping, spiral search strategy and snake dynamic adaptive weight are introduced to improve the snake optimization algorithm (SO), which enhances the local optimal escape capability and global search capability of the algorithm. The tangent-based rectified linear unit (ThLU) was used to improve the rectified linear unit (ReLU) of the standard TCN to strengthen the generalization capability of the model. The MFISO algorithm was used to optimize the relevant hyperparameters of the improved TCN model to optimize the accuracy of gas outburst prediction. The TCN, GRU, LSTM, SO-TCN, WOA-TCN, and PSO-TCN prediction models were selected to compare the prediction performance with the MFISO-TCN gas outburst prediction model, and the results showed that the mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) of the MFISO-TCN model were 3.11%, 0.47% and 3.31% are lower than those of other models, which verifies that the method of this paper effectively intensifies the performance of gas outburst prediction model in underground mines.

INDEX TERMS Gas outburst, phase space reconstruction method, multiple strategy fusion, snake optimization algorithm, tangent-based rectified linear unit, temporal convolutional network.

I. INTRODUCTION

A gas outburst prediction model is structured by multiple influencing factors related to gas outburst can effectively prevent gas outburst disasters in underground mines [1].

Numerous prediction methods combining deep learning and machine learning have been proposed by domestic and foreign scholars to address the outburst problem of gas. Rong Liang proposed a gas concentration forecasting

The associate editor coordinating the review of this manuscript and approving it for publication was Nuno Garcia[ID].

model with a bidirectional gated recurrent unit neural network (Adamax-BiGRU) using an adaptive moment estimation maximum (Adamax) optimization algorithm [2]. To solve the problem of low prediction accuracy of gas concentration regression prediction algorithms, Yonghui Xu proposesd a gas concentration prediction algorithm based on a stacking model [3]. Ningke Xu proposed an IWOA-LSTM-CEEMDAN model based on an improved whale optimization algorithm (IWOA) and the complete ensemble empirical model decomposition with adaptive noise (CEEMDAN) method is used for gas concentration

prediction [4]. Yaoyong Xu proposed an improved gravitational search algorithm (IGSA) to train a model that optimizes the initial weights and thresholds of BP neural networks [5]. Pengtao Jia proposed a gas concentration prediction model based on convolutional neural network (CNN) optimized by particle swarm algorithm and bi-directional gated unit neural network (aBiGRU) optimized by adaptive moment estimation maximum [6]. Qingwei Xu proposed a new accident prevention technique using a novel safety assessment method based on fault tree basic event importance, grey relational analysis and the bow tie model [7]. Yadong Cai classifies gas concentration data into warning and non-warning classes based on concentration thresholds, and proposes a probability density machine (PDM) algorithm with good adaptability to unbalanced data distributions [8].

Due to the serious nonlinearity in the complex and variable influencing factors related to gas outburst, we propose a gas outburst prediction model is structured by multiple strategy fusion improved snake optimization algorithm (MFISO) optimizes temporal convolutional network (TCN). Selected phase space reconstruction technique to reconstruct the multiple time series of influence factors related to gas outburst. Then combined Sine chaos mapping, spiral search strategy and snake dynamic adaptive weights to improve the snake optimization algorithm. The tangent-based rectified linear unit was used to improve the standard TCN, and the MFISO-TCN gas outburst prediction model was established. Finally, the real gas outburst data from a coal mine in Shanxi was selected for experimental analysis and comparison, which verified that the method of this paper can effectively improve the accuracy of gas outburst prediction.

II. PHASE SPACE RECONSTRUCT TIME SERIES

The influencing factors related to gas protrusion are coal seam thickness, coal seam spacing, advancement speed, etc. The factors often present complex non-linear relationships, which will directly lead to a decrease in the precision of the gas outburst prediction model. The phase space reconstruction method [9] is used to reconstruct the characteristic time series of multiple influencing factors related to gas outburst, each time series is mapped into a higher dimensional space by the delay time and expanding dimensionality to remove invalid data features and restore the underlying principles of the original time series related to gas outburst in the chaotic state. Set a single time series of gas outburst $[Z_1, Z_2, \cdots, Z_n]$, the specific equation is as (1).

$$Z_i(n) = \left[Z_i, Z_{i+\tau}, \cdots, Z_{n+(m-1)\tau} \right]$$

$$= \begin{bmatrix} z_1 & z_{1+\tau} & \cdots & z_{1+(m-1)\tau} \\ z_2 & z_{2+\tau} & \cdots & z_{2+(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ & & \vdots & \\ z_n & z_{n+\tau} & \vdots & z_{n+(m-1)\tau} \end{bmatrix} \quad (1)$$

In (1), $Z_i(n)$ is the phase space of the ith feature variable, τ represents the delay time, m is the embedding dimension, z_n denotes the phase point in the phase space.

A. DETERMINING DELAY TIME

The delay time τ and the embedding dimension m determine the extension and randomness of each influencing factor related to gas outburst in the high-dimensional space [10], and the interactive information method was chosen to determine the parameter τ.

Assume that there are two time series of influencing factors related to gas outburst $[q_1, q_2, \cdots, q_m]$, $[s_1, s_2, \cdots, s_n]$.

$$F(S) = -\sum_{i=1}^{n} H_S(s_i) \log_2 H_S(s_i) \quad (2)$$

In (2), $F(Q)$ and $F(S)$ are the entropies of the two time series, respectively. $H_Q(q_i)$ and $H_S(s_j)$ are the probability of the two time series of gas outburst, respectively.

$$F(Q, S) = -\sum_{i=1}^{m} \sum_{j=1}^{n} H_{Q,S}(q_i, s_j) \log_2 H_{Q,S}(q_i, s_j) \quad (3)$$

$$I(Q, S) = F(Q) + F(S) - F(Q, S) \quad (4)$$

In (3) and (4), $H_{Q,S}(q_i, s_j)$ is the probability of the joint distribution of $H_Q(q_i)$ and $H_S(s_j)$. $I(Q, S)$ is a correlation function with respect to the delay time τ.

B. DETERMINING EMBEDDING DIMENSION

The Cao method was chosen to determine the parameter d.

$$E(m) = (n - m\tau)^{-1} \cdot \sum_{i=1}^{n-d\tau} a(i, m) \quad (5)$$

$$E^*(m) = (n - m\tau)^{-1} \cdot \sum_{i=1}^{n-m\tau} |x(i + m\tau) - x(n(i, m) + m\tau)| \quad (6)$$

$$E1(m) = E(m + 1)/E(m) \quad (7)$$

$$E2(m) = E^*(m + 1)/E^*(m) \quad (8)$$

In (5), (6), (7), and (8), $E(m)$ is the evaluation criterion for reconstructing the time series based on known τ and unknown m. $E^*(m)$ is the correlation evaluation criterion for the two sets of time series of the influencing factors. The value of the embedding dimension m is determined by $E1(m)$. $E2(m)$ evaluates the randomness of the time series of influencing factors related to gas outburst, which works best when it is constant equal to 1.

III. SO ALGORITHM

The standard snake optimization algorithm [11] achieves the optimization of model hyperparameters by simulating various behaviors of snakes in the mating process. The snake optimization algorithm consists of several stages and modes, and the initialization process of the population in the early stage is relatively simple resulting in poor population stochasticity. As the number of iterations increases, the global search capability of the algorithm in the development stage gradually decreases, and the local optimal escape capability of the algorithm decreases largely. The population convergence speed is

significantly improved in the later stage, but the population individuals are prone to aggregation phenomenon resulting in the degradation of the optimal search performance of algorithm. The population position update equations of the snake optimization algorithm during the development phase are as follows.

$$X_{i,j}(t+1) = X_{food} \pm c_2 \cdot Temp \cdot r \cdot (X_{food} - X_{i,j}(t)),$$
$$Q \in (0.25, 1] \ and \ Temp \in (0.6, 1] \qquad (9)$$

$$Q = c_1 \cdot \exp(\frac{t-T}{T}) \qquad (10)$$

$$Temp = \exp(-\frac{t}{T}) \qquad (11)$$

In (9), (10), and (11), $X_{i,j}(t+1)$ represents the optimal position of the snake in $t+1$ iterations. X_{food} is the optimal position of the target food. c_1 and c_2 are constants. r is a random number within $[0, 1]$. $X_{i,j}$ represents the current individual position of the snake. t is the number of current iterations. T is the maximum number of iterations.

The equations for updating the population position in the mating mode of the snake optimization algorithm during the development phase are as follows.

$$X_{i,m}(t+1)$$
$$= X_{i,m}(t) + c_2 \cdot A_m \cdot r \cdot (Q \cdot X_{i,f}(t) - X_{i,m}(t)),$$
$$Q \in (0.25, 1] \ and \ Temp \in (0, 0.6) \qquad (12)$$

$$A_m = \exp(-\frac{f_{i,f}}{f_{i,m}}) \qquad (13)$$

$$X_{i,f}(t+1)$$
$$= X_{i,f}(t) + c_2 \cdot A_f \cdot r \cdot (Q \cdot X_{i,m}(t) - X_{i,f}(t)),$$
$$Q \in (0.25, 1] \ and \ Temp \in (0, 0.6) \qquad (14)$$

$$A_f = \exp(-\frac{f_{i,m}}{f_{i,f}}) \qquad (15)$$

In (12), (13), (14), and (15), $X_{i,m}(t+1)$ and $X_{i,m}(t+1)$ represent the new generation of individual positions of the ith male and female snakes, respectively. r is a random number within $[0, 1]$. $X_{i,f}(t)$ and $X_{i,m}(t)$ are the current positions of the ith male and female snakes, respectively. A_m and A_f are the mating ability of the male and female snakes, respectively. $f_{i,m}$ and $f_{i,f}$ represent the fitness values of the positions of the ith male and female snakes, respectively.

A. MFISO ALGORITHM
To solve the problems of the above SO algorithm, Sine chaos mapping, spiral search strategy and snake dynamic adaptive weight are used to improve the SO algorithm and boost optimal performance of algorithm.

1) SINE CHAOS MAPPING INITIALIZES SNAKE POPULATIONS
The traditional SO algorithm generates the population individuals between $[0, 1]$ irregularly, which is prone to the problems of low quality of population individuals and uneven distribution of population, resulting in the degradation of
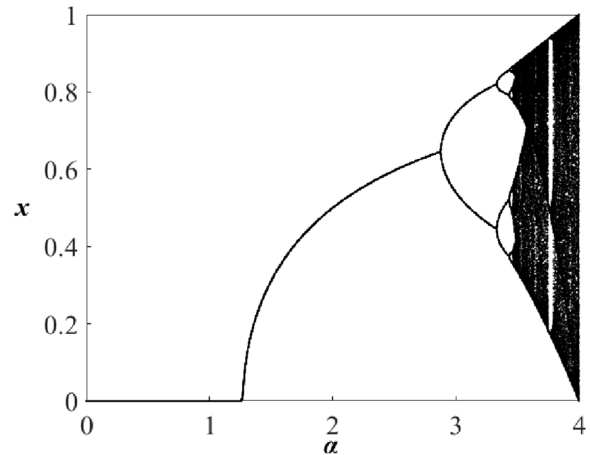

FIGURE 1. Population sequence distribution under different bifurcation parameters.

the algorithm itself in terms of the performance of the optimization search. Sine chaos mapping [12] is introduced to initialize the snake population, improve the overall quality of the population, make the population distribution uniform and diverse, and led to an improvement in the convergence speed and accuracy of the algorithm. The specific equation is as follows.

$$x_{k+1} = \frac{\alpha}{4} \times \sin(\pi x_k) \qquad (16)$$

In (16), k represents the number of iterations of the algorithm. α is the bifurcation parameter. $\alpha \in (0, 4]$ and $x \in (0, 1)$.

Set the initial value x_0, the distribution of the Sine chaotic sequence x as α varies is shown in Fig. 1.

As shown in Fig. 1, when $\alpha \in (0, 3.436]$, the sequence x presents a finite distribution, and the sequence x gradually presents a quasi-divergence state as α increases. When $\alpha \in (3.436, 4]$, the values of x diverge to 0 and 1, and chaos appears. When $\alpha = 4$, the sequence x has the best distribution and the highest randomness of the population. Therefore, the Sine chaotic sequence at $\alpha = 4$ is selected to initialize the population, and the spatial distribution of the Sine chaos mapping in Fig. 2 is obtained by iterating 300 times.

(16) is used to map the values of the variables generated after Sine chaos mapping to individual snakes to achieve the initialization of the snake population. The specific equation is as (17).

$$X_i = X_{\min} + x_{k+1} \cdot (X_{\max} - X_{\min}) \qquad (17)$$

In (17), X_i is the position of the ith snake after mapping. $X_{\max}$ and $X_{\min}$ are the upper and lower bounds of each snake and each dimension.

2) SPIRAL SEARCH STRATEGY IMPROVES SO ALGORITHM
Snake population of the standard SO algorithm is prone to a narrow global search range when updating its position in the
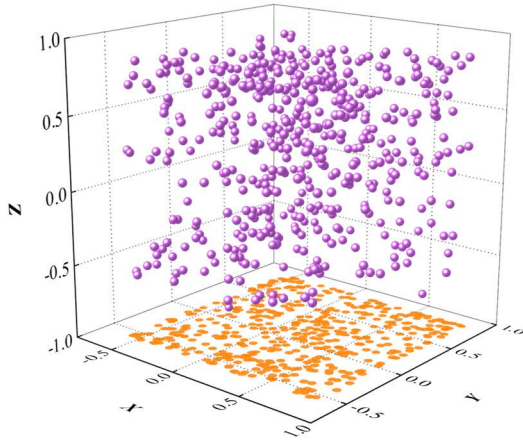
FIGURE 2. Spatial distribution of Sine chaos mapping.

FIGURE 3. Curve of snake dynamic adaptive weight.

development phase. Although algorithm can achieve convergence quickly, the population diversity plummets, resulting in a significant reduction in the global search capability of algorithm and ability to jump away from the local optimal solution [13]. The Whale Optimization Algorithm [14] (WOA) has excellent local optimal escape ability and performs well in accomplishing the global search task, in which the spiral search strategy of the whale in the hunting phase is that the whale swims in a spiral shape toward the prey and shrinks the encirclement. This not only ensures the convergence of the algorithm, but also expands the search range of the population, which is a good balance between global search ability and local optimal escape ability. The position update equations of the spiral search strategy are as follows.

$$X(t+1) = \begin{cases} X^*(t) - A \cdot D, & \text{if } p < 0.5 \\ X^*(t) + D' \cdot e^{bl} \cdot \cos(2\pi l), & \text{if } p \geq 0.5 \end{cases}$$ (18)

$$D = |C \cdot X^*(t) - X(t)|$$ (19)

$$D' = |X^*(t) - X(t)|$$ (20)

In (18), (19), (20), A takes a random between $[-1, 1]$, D' denotes the distance of the ith whale from its prey, b is a logarithmic spiral constant, l is a random number between $[-1, 1]$.

The spiral search strategy was introduced into the snake location update process in the development phase of the snake optimization algorithm to make full use of the spiral search mechanism to obtain population location information and probabilistic changes. The improved equations of updating the snake position in the development phase are as follows.

$$X_{i,j}(t+1)$$
$$= \begin{cases} X_{food}(t) - A \cdot D, & \text{if } p < 0.5 \\ X_{food}(t) + D' \cdot c_3 \cdot e^{bl} \cdot \cos(2\pi l), & \text{if } p \geq 0.5 \end{cases}$$ (21)
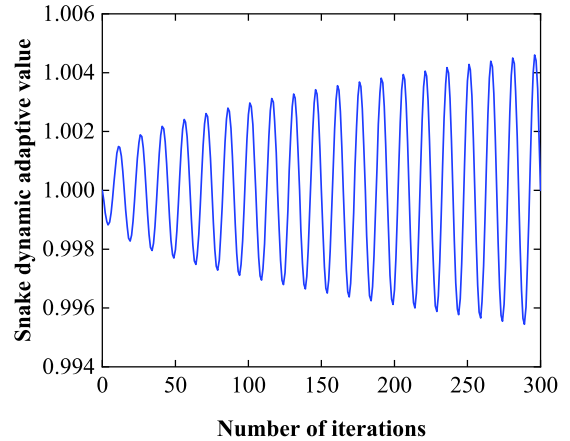
$$l = (Temp - 1) \times rand + 1$$ (22)

$$Temp = \exp(-\frac{t}{T})$$ (23)

$$D' = |X_{food} - X_{i,j}(t)|$$ (24)

In (21), (22), (23), and (24), the logarithmic spiral constant b is taken as 1, c_3 is a constant taken as 2, D' is the distance between the current location of the individual snake and the best location of the target food.

3) SNAKE DYNAMIC ADAPTIVE WEIGHTS

Since the mating population of the SO algorithm aggregates faster, leading to the algorithm prone to local optimal stagnation. Snake dynamic adaptive weight is used to diversify the optimal positions of the snake population and enhance the ability to reproduce optimal next generation individuals. After combining the snake dynamic adaptive weight, the snake position update equations in mating mode are as follows.

$$X_{i,m}(t+1)$$
$$= X_{i,m}(t) + \omega \cdot A_m \cdot r \cdot (Q \cdot X_{i,f}(t) - X_{i,m}(t))$$ (25)

$$X_{i,f}(t+1)$$
$$= X_{i,f}(t) + \omega \cdot A_f \cdot r \cdot (Q \cdot X_{i,m}(t) - X_{i,f}(t))$$ (26)

where the snake dynamic adaptive weight is formulated as (27).

$$\omega = 1 + \frac{e^{t^\beta}}{10T} \cdot \cos(\frac{\pi t}{\gamma T} + \frac{\pi}{2})$$ (27)

In (27), t denotes the current iteration of the algorithm, T is the maximum number of iterations of the algorithm. β is the adaptive range coefficient, which determines the location update range of the snake population, $\beta \in [0.1, 0.2]$, β is taken as 0.17. γ is the adaptive density coefficient, which determines the individual update density of the snake population, $\gamma \in [0.015, 0.035]$, γ is taken as 0.025.

Fig. 3 shows the curve of the adaptation value of 300 iterations of the snake dynamic adaptive weight. As can be seen from the figure, with the increase of iterations, the trend of the adaptive curve presents an expansive spiral state, and
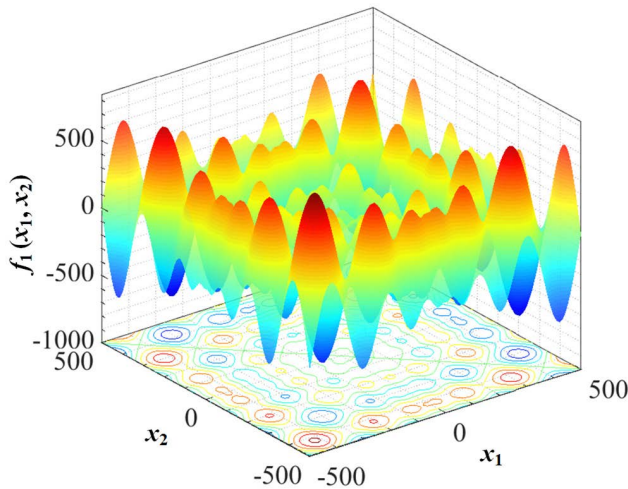
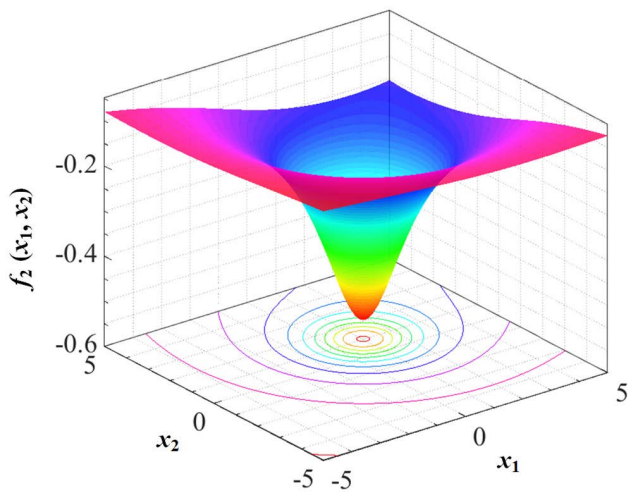**FIGURE 4.** Testing function $f_1(x)$.

FIGURE 5. Testing function $f_2(x)$.

the global search ability of the algorithm steadily increases. The high-density curve makes the population diversification ability greatly improved, thus strengthening the local search capability of the algorithm.

B. MFISO ALGORITHM TESTING

In order to confirm the optimization validity and stability of the MFISO algorithm, the Schwefel2.26 function and the Shekel function were selected for the function optimization comparison experiments and compared with the standard SO algorithm, the Particle Swarm Optimization (PSO) algorithm [15] and the WOA algorithm for the analysis, and the number of iterations of all algorithms was 300 times. Fig. 4 and Fig. 5 are the testing function diagrams, and the specific equations are as follows.

$$f_1(x) = -\sum_{i=1}^{n} -x_i \sin(\sqrt{|x_i|}) \quad (28)$$

$$f_2(x) = -\cos(x_1)\cos(x_2)\exp(-(x_1 - \pi)^2 - (x_2 - \pi)^2) \quad (29)$$

TABLE 1. Test results of $f_1(x)$ for four optimization algorithms.

	MFISO	SO	WOA	PSO
Optimum value	-4189.8289	-4182.2076	-4179.5349	-2274.8226
Average value	-4138.3968	-4126.5823	-3992.0641	-2075.1941
Standard deviation	232.3281	251.3214	198.4602	504.1946

TABLE 2. Test results of $f_2(x)$ for four optimization algorithms.

	MFISO	SO	WOA	PSO
Optimum value	-10.5464	-10.1286	-5.1285	-5.1757
Average value	-10.2962	-9.9259	-5.1112	-5.0663
Standard deviation	0.5065	0.7551	0.6323	0.3218

From Table 1, for the testing function $f_1(x)$, both the SO algorithm and the MFISO algorithm can find their theoretical optimal value, and the results are better than those of the WOA algorithm and the PSO algorithm, but the SO algorithm is worse than the MFISO algorithm in the calculation of the mean value and standard deviation.

From Table 2, for the testing function $f_2(x)$, only the MFISO algorithm finds the optimal solution, which is a significant improvement over the SO algorithm in terms of the accuracy of the search for the optimal value.

Regardless of which testing function evaluation index, the MFISO algorithm performs better than the other three algorithms, indicating that the MFISO algorithm is significantly better than the other algorithms in terms of stability and accuracy of the optimization search. The improved snake optimization algorithm through multiple strategy fusion is able to have both strong global search capability and local optimal escape capability when searching for optimal solutions sufficiently and efficiently.

From Fig. 6 and Fig. 7, it is concluded that for the testing function $f_1(x)$, the MFISO algorithm reaches the convergence state in 24 iterations, which is faster than the other three algorithms and the optimal fitness value of the search is better than the other three algorithms. For the testing function $f_2(x)$, the convergence speed of the four algorithms is not much different, but the optimal fitness value of the MFISO algorithm is better than the other three algorithms in all cases.

The iterative search performance test of the four algorithms by testing functions $f_1(x)$ and $f_2(x)$ illustrates the comprehensive strength of the improved snake optimization algorithm is stronger than the other three algorithms in the three aspects of convergence speed, search accuracy and local escape capability.
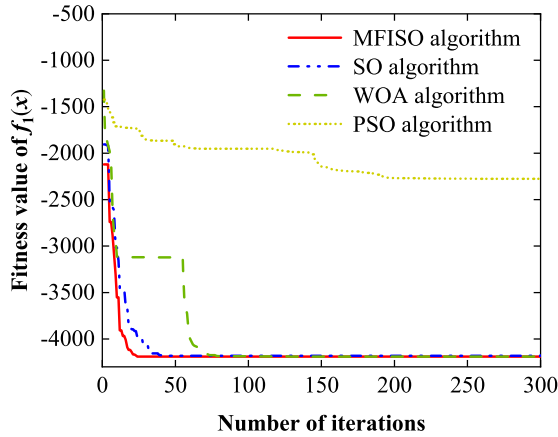
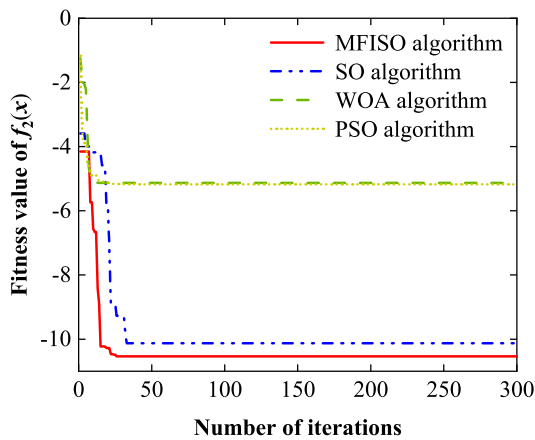**FIGURE 6.** Comparison of the optimization search for $f_1(x)$ for each optimization algorithm.

FIGURE 7. Comparison of the optimization search for $f_2(x)$ for each optimization algorithm.

IV. TCN

TCN is a new network structure for processing long time series obtained by improving, combining and optimizing on the basis of traditional convolutional neural network and recurrent neural network [16]. Due to the complexity and large time span of time series of gas outburst, TCN is used to accurately capture more historical data and accurately predict future data.

A. BUILDING THE TCN NETWORK MODEL

1) CAUSAL CONVOLUTIONS AND DILATED CONVOLUTIONS

Causal convolution ensures that the output of a complex gas prominence time series at moment t is only correlated with the constituents at moment $(t-1)$ in the previous layer of the network, and the length of the gas prominence sequence is guaranteed to be the same in the preceding and following network layers by the method of complementary zeros [17]. Compared with the conventional convolution, the introduction of the dilation factor in the dilated convolution greatly enhances the perceptual field of the convolution
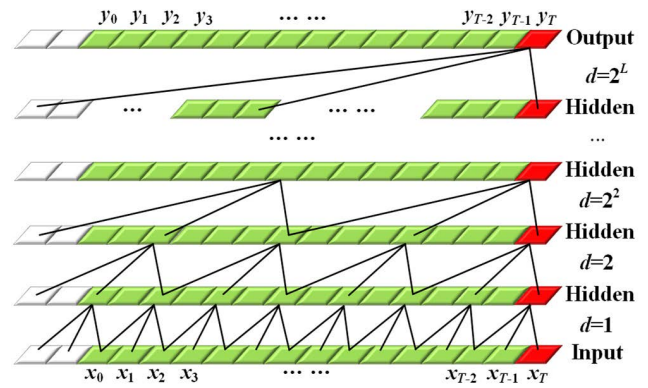

FIGURE 8. The structure diagram of causal convolution and dilated convolution of TCN. Comparison of activation functions.

kernel and ensures that input sequences of arbitrary length can be captured [18].

Fig. 8 shows the structure of causal convolution and dilated convolution of TCN. It is assumed that there are L convolutional layers, where the equation of dilated convolution equation is as (30).

$$F(s) = (x * df)(s) = \sum_{i=0}^{k-1} f(i) X_{x-d \cdot i} \tag{30}$$

In (30), the number of convolution layers L determines the dilation factor size, and the dilation factor $d = [1, 2^1, \cdots, 2^L]$. The convolution kernel size k, also called the model filter coefficient.

2) RESIDUAL CONNECTIONS

TCN uses residual connections to deepen the network depth and overcome a series of problems such as gradient disappearance and gradient explosion of the network, which is a good solution to the problem of slow convergence speed and neuron degradation of deep neural networks.

Fig. 9 shows the structure of residual connections of TCN, which encapsulates causal convolutions, dilated convolutions, and residual connections. Dropout will stop the neuron work with certain probability, thus enhancing generalization ability of the model. ReLU is the rectified linear unit, which acts as the activation function; WeightNorm can normalize the weight values.

3) IMPROVEMENT OF TCN NETWORK

The standard TCN uses the rectified linear units (ReLU) as the activation function, and although it performs better in accelerating the convergence of the model, the negative half-axis of the ReLU function is all zero, which is prone to problems such as necrosis of some neurons during the model training [19]. The rectified linear units based on tanh function (ThLU) [20] was selected instead of the ReLU activation function to overcome the problems of TCN model in the network training process such as the decrease of the prediction
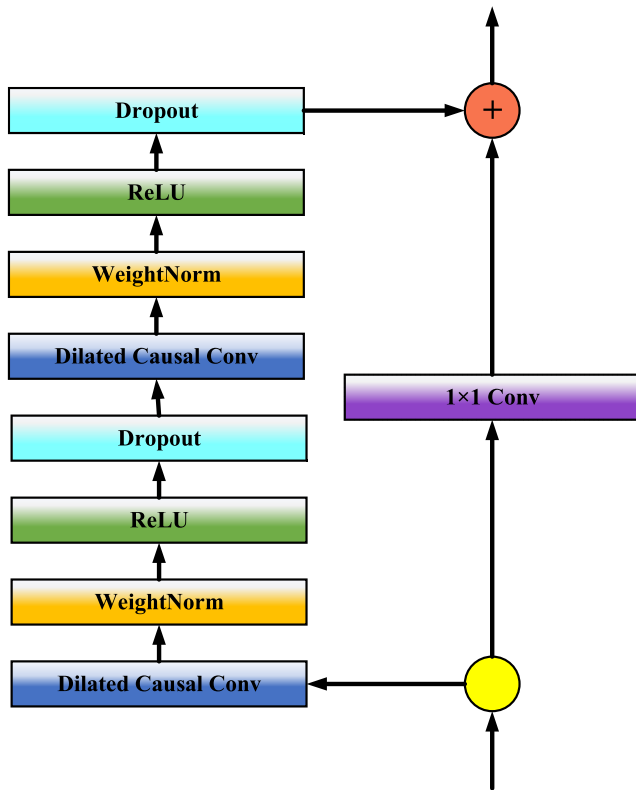
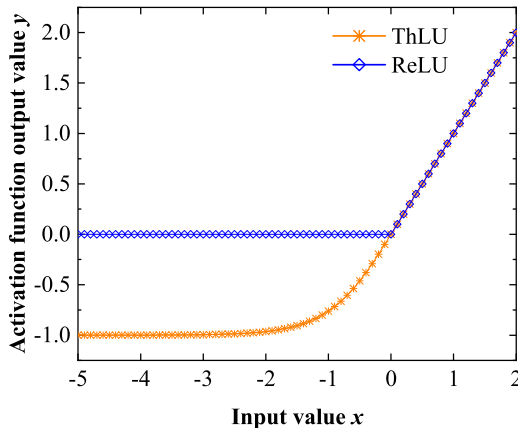**FIGURE 9.** The structure diagram of residual connections of TCN.

FIGURE 10. Comparison of activation functions.

TABLE 3. Optimization intervals of TCN relevant hyperparameters.

hyperparameters	optimization intervals
The number of convolutional layers L	$[0-50]$
The filter coefficients k	$[0-10]$

function are all non-zero and less than zero. The output values of the ThLU function are better than those of the ReLU function, and the improved TCN model is more effective in gas outburst prediction.

B. ESTABLISHING MFISO-TCN GAS PROTRUSION PREDICTION MODEL

In order to make the TCN model have better historical data capturing ability and wider data perception field in gas outburst prediction, the MFISO algorithm was used to optimize the number of convolutional layers L and filter coefficients k of the TCN model to build the MFISO-TCN gas outburst prediction model. The optimization intervals of TCN relevant hyperparameters are shown in Table 3.

The process of establishing MFISO-TCN gas outburst prediction model is shown in Fig. 11.

(1) Set the population number N and divide the population into two groups equally, choose the maximum number of iterations T, and initialize the snake population using (17) to obtain a uniformly distributed snake population.

(2) Initialize the TCN relevant hyperparameters and optimization intervals.

(3) The optimal individual and optimal fitness values for male and female snakes under the current hyperparameters are calculated, and the food quantity Q and temperature $Temp$ are calculated according to (10) and (11).

(4) In the snake development phase, the spiral search strategy was used to update the male and female individual locations according to (21), (22), (23), and (24).

(5) In the mating mode of snakes, the fitness values of male and female snake positions are calculated according to (25), (26), and (27), and if the next generation is successfully hatched, the new individual position is recorded and the worst snake position is replaced.

(6) Determine whether the maximum number of iterations is satisfied, if the condition is satisfied, the optimal hyperparameters are assigned to TCN, otherwise return to (3) to continue iteration.

(7) Use the optimal hyperparameters optimized by MFISO algorithm to construct MFISO-TCN gas outburst prediction model and output the prediction results.

V. EXPERIMENTAL ANALYSIS

Since there are many influencing factors related to gas outburst in underground mines, 8184 sets of data containing ten influencing factors from a coal mine in Shanxi, including
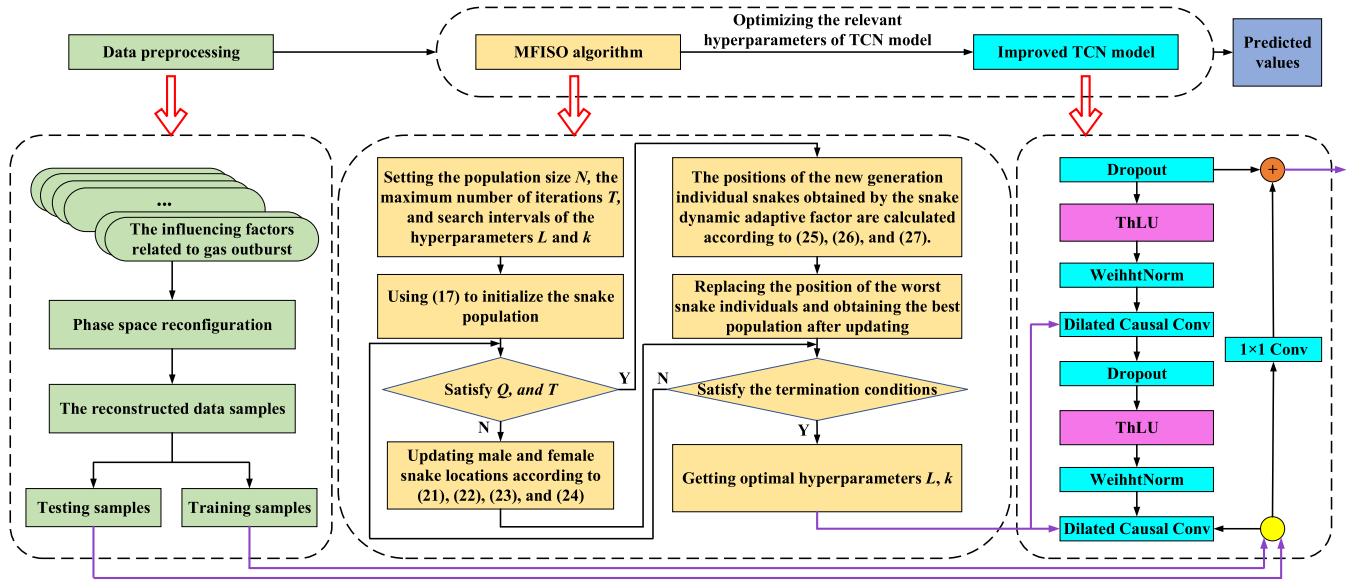
accuracy of gas outburst due to the necrosis of some neurons. The ThLU function is shown in (31).

$$f(x) = \begin{cases} x, & x \geq 0 \\ \dfrac{1 - e^{-2x}}{1 + e^{-2x}}, & x < 0 \end{cases} \quad (31)$$

Fig. 10 shows the comparison of the activation functions of ReLU and ThLU. When $x \geq 0$, the output values of the positive half-axis of the ThLU function are the same as the output values of the ReLU function, which are both equal to the input values. When $x < 0$, the output values of the ReLU function are all zero, but the output values of the ThLU

FIGURE 11. MFISO-TCN gas outburst prediction model.

TABLE 4. Partial data on influencing factors related to gas outburst.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1	563.00	5.66	11.00	3.68	12.00	3.53	0.38	2410.00	3.74	6.8390
2	590.00	2.52	8.00	4.21	18.00	2.85	0.48	3139.00	3.30	6.8353
3	604.00	2.38	9.00	4.03	16.00	2.64	0.52	3354.00	3.28	6.8533
4	607.00	3.66	9.00	4.34	17.00	2.77	0.12	3087.00	3.14	6.8466
5	634.00	3.45	12.00	4.80	15.00	2.92	1.0	3620.00	5.10	6.8476
6	640.00	3.15	11.00	4.67	15.00	2.75	0.82	3412.00	3.10	6.8548
7	450.00	2.57	12.00	2.43	16.00	4.32	0.82	1996.00	3.58	6.8509
8	544.0	4.02	11.00	3.16	13.00	3.81	0.51	2207.00	3.90	6.8525
9	629.00	2.12	13.00	4.62	19.00	2.80	0.13	3456.00	3.68	6.8383
10	401.00	6.09	10.00	1.87	25.00	4.52	0.53	1855.00	5.64	6.8517

coal mining depth X_1 (m), coal seam thickness X_2 (m), coal seam inclination X_3 (deg), original gas content of mining seam X_4 (m^3/t), coal seam spacing X_5 (m), advancement speed X_6 (m/d), extraction rate X_7, daily production X_8 (m^3), adjacent seam gas content X_9 (m^3/t), and gas outburst X_{10} (m^3/min), were selected. The first 8000 sets of measured data were selected as the research object, the first 6000 sets of data were used as the training samples of the gas outburst prediction model, and the remaining 2000 sets were used as the testing samples of the model. The data

of ten groups of influencing factors are listed, as shown in Table 4.

A. BUILDING THE TCN NETWORK MODEL
The selected time series of ten influencing factors related to gas outburst are reconstructed in phase space, and the delay time τ and embedding dimension d are solved using the interactive information method and Cao method.

From Fig. 12, the first extreme point of $I(Q, S)$ is at the delay time of 3, so $\tau = 1$. From Fig. 13, when $d_0 = 14$, $E1(d)$
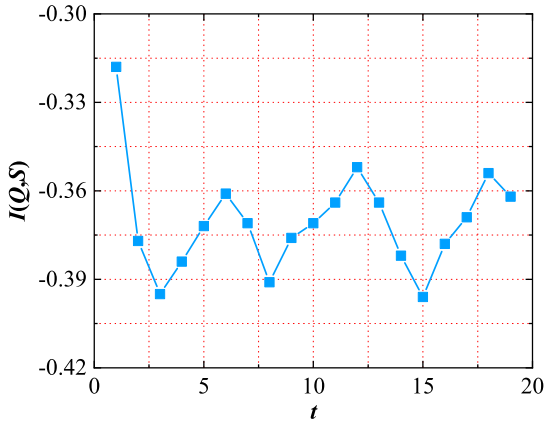
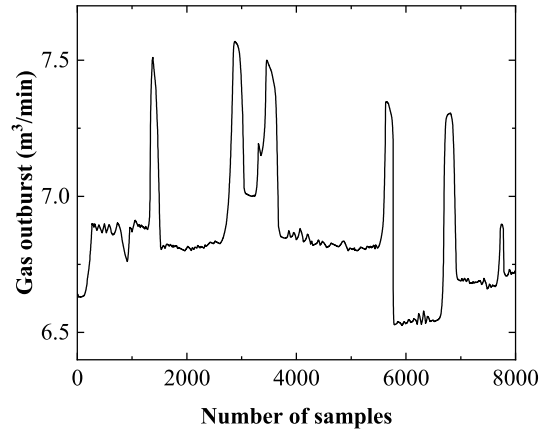**FIGURE 12.** The curve of delay time.

FIGURE 13. The curve of embedding dimensional.

FIGURE 14. The time series of absolute gas outburst.

FIGURE 15. The time series of reconstructed absolute gas outburst.

FIGURE 16. The curves of fitness values.

From Fig. 14 and Fig. 15, as the number of samples gradually increases, the sequence of gas outburst gradually appears noise points and has strong nonlinearity, and the distribution of the reconstructed sequence is regular and the extension performance is good.

B. COMPARISON OF ALGORITHM OPTIMIZATION
The MFISO algorithm is used to search for the optimization of the relevant hyperparameters of the TCN model, and the results are compared and analyzed with those of the SO algorithm, WOA algorithm and PSO algorithm. Fig. 14 shows the iterative comparison of the fitness values of the optimized TCN model for each algorithm.

From Fig. 16, the MFISO algorithm with Sine chaos mapping, spiral search strategy, and snake dynamic adaptive factor reached the convergence state and searched the optimal fitness value after only 22 iterations of the four algorithms in the process of searching for the optimal hyperparameters. The feasibility of using the MFISO algorithm to optimize the relevant hyperparameters of TCN and build the gas outburst prediction model is demonstrated.

tends to 1 and gradually stabilizes, and $E2(d)$ is constant equal to 1. At this time, $(d_0 + 1)$ is the optimal embedding dimension, so $d = 15$. Reconstruct the time series of relevant influence factors according to (1), and remove the invalid data features in the high-dimensional space. Fig. 14 and Fig. 15 show the comparison of the original time series of gas outburst and the reconstructed time series.

TABLE 5. Comparison of model error.

		Type of Model			
		MFISO-TCN	TCN	GRU	LSTM
	MAE	3.12%	9.54%	11.35%	12.36%
Type of error	MAPE	0.48%	1.38%	1.58%	1.89%
	RMSE	3.31%	9.95%	11.75%	12.77%

C. COMPARISON OF ALGORITHM OPTIMIZATION

To prove the validity of the research content of this article, two sets of model simulations were compared and mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean squared error (RMSE) were selected to measure the deviation of model prediction values from the true values. The specific equations are as follows.

$$MAE = \frac{1}{H} \cdot \sum_{i=1}^{H} |y_i - y_i'| \times 100\% \tag{32}$$

$$MAPE = \frac{1}{H} \cdot \sum_{i=1}^{H} |\frac{y_i - y_i'}{y_i}| \times 100\% \tag{33}$$

$$RMSE = \sqrt{\frac{1}{H} \sum_{i=1}^{H} (y_i - y_i')^2} \times 100\% \tag{34}$$

In (32), (33), (34), y_i is the real value of the ith sample of data set of gas outburst, y_i' is the predicted value of the model for the ith sample of data set, H denotes the number of samples of data set.

1) PERFORMANCE COMPARISON OF DIFFERENT TYPES OF MODELS

Traditional TCN, gated recurrent unit (GRU) [21] and long short term memory (LSTM) [22] were selected for network training and compared with MFISO-TCN gas outburst prediction model for analysis. All models were trained 200 times.

Fig. 17 shows the comparison of the prediction results of each model for 2000 sets of real data after training. MFISO-TCN model has better prediction and data reduction ability than other models, the difference between traditional TCN and MFISO-TCN is not significant for the data points in the smooth segment, but for the peak values in the segments of 750-850 and 1700-1800, MFISO-TCN performs better. Although GRU and LSTM can predict the trend of the entire gas outburst sequence, their response is worse than MFISO-TCN for complex and unstable gas outburst sequences.

Table 5 shows the comparison of the prediction errors of the above four models, and MAE, MAPE, and RMSE are
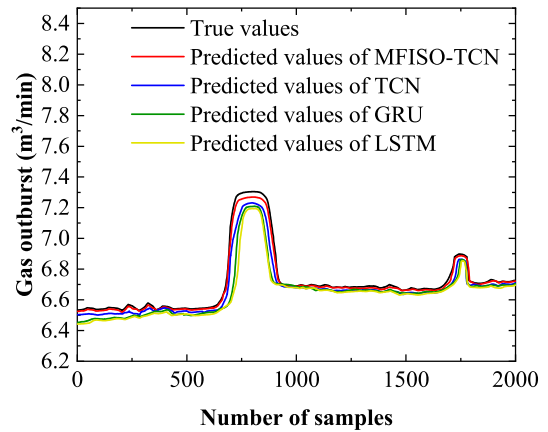

FIGURE 17. Comparison of model prediction results.

used to measure the validity of the models for gas outburst prediction. According to the comparison results, the prediction accuracy of MFISO-TCN is greater than the other three models, and it is effective in gas outburst prediction.

2) PERFORMANCE COMPARISON OF TCN MODELS OPTIMIZED BY DIFFERENT ALGORITHMS

The standard SO algorithm, WOA algorithm, and PSO algorithm were used for hyperparameters searching and network training of TCN models, and finally compared with MFISO-TCN gas outburst prediction model for analysis. All models were trained 200 times.

Fig. 18 shows the comparison of the prediction results of each model after training on 2000 sets of real data. The TCN models optimized by different algorithms all perform well in gas outburst prediction, but the MFISO-TCN model is still more capable of capturing data in different segments than the other models.

Table 6 shows the comparison of the prediction errors of the above four different algorithm optimized models. The comparison shows that the prediction accuracy of the gas outburst prediction model established by the optimized hyperparameters of the SO algorithm has been better than that of the standard TCN, but the effect of improving the accuracy is much less than that of the MFISO-TCN

TABLE 6. Comparison of model error.

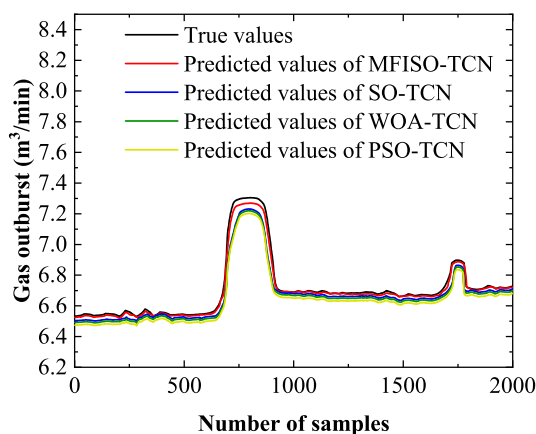|		Type of Model			
		MFISO-TCN	SO-TCN	WOA-TCN	PSO-TCN
	MAE	3.11%	8.72%	9.47%	10.68%
Type of error	MAPE	0.47%	1.31%	1.41%	1.53%
	RMSE	3.31%	8.96%	9.82%	10.97%

FIGURE 18. Comparison of model prediction results.

gas outburst prediction model. The feasibility of using the MFISO algorithm to optimize the relevant hyperparameters of TCN and build the gas outburst prediction model is further demonstrated.

In summary, the gas outburst prediction model established by using the multiple strategy fusion and improved snake optimization algorithm to search for the relevant hyperparameters of TCN can effectively enhance the prediction accuracy of gas outburst in underground mines.

VI. CONCLUSION

1) The complex and variable time series of gas outburst in underground mines are reconstructed by the phase space reconstruction method to analyze the connection between the relevant influencing factors in the high-dimensional space, which removes the invalid data features, accelerates the convergence speed of the model, and intensifies the prediction accuracy of the model.

2) The Sine chaos mapping, spiral search strategy and snake dynamic adaptive factor are used to improve the snake optimization algorithm, which enriches the population diversity, improves the defect of easily falling into local optimum in the process of searching for the best, and strengthens the global search capability of the algorithm.

3) Combining the ThLU function to improve the TCN model effectively overcomes the problem that the neurons of TCN are prone to necrosis during the training process. Optimizing the relevant hyperparameters of the improved TCN using the MFISO algorithm can effectively intensify the generalization ability and prediction accuracy of the TCN.

4) The MFISO-TCN gas outburst prediction model was compared with TCN, GRU, LSTM, SO-TCN, WOA-TCN, PSO-TCN in two sets of experiments, respectively. The proposed method can predict the time series of gas outburst in underground mines more accurately and has certain theoretical research and engineering practical significance.

REFERENCES

[1] W. Nie, S. J. Peng, J. Xu, L. R. Liu, G. Wang, and J. B. Geng, "Experimental analyses of the major parameters affecting the intensity of outbursts of coal and gas," *Sci. World J.*, vol. 2014, Jan. 2014, Art. no. 185608, doi: 10.1155/2014/185608.

[2] R. Liang, X. Chang, P. Jia, and C. Xu, "Mine gas concentration forecasting model based on an optimized BiGRU network," *ACS Omega*, vol. 5, no. 44, pp. 28579–28586, Oct. 2020, doi: 10.1021/acsomega.0c03417.

[3] Y. Xu, R. Meng, and X. Zhao, "Research on a gas concentration prediction algorithm based on stacking," *Sensors*, vol. 21, no. 5, p. 1597, Feb. 2021, doi: 10.3390/s21051597.

[4] N. Xu, X. Wang, X. Meng, and H. Chang, "Gas concentration prediction based on IWOA-LSTM-CEEMDAN residual correction model," *Sensors*, vol. 22, no. 12, p. 4412, Jun. 2022, doi: 10.3390/s22124412.

[5] Y. S. Xu, C. Y. Qi, and S. C. Feng, "Gas surge prediction model based on IGSA-BP network," *J. Electron. Meas. Instrum.*, vol. 33, no. 5, pp. 111–117, 2019, doi: 10.13382/j.jemi.B1801923.

[6] P. T. Jia, Z. Y. Zhang, R. Liang, H. T. Liu, and Y. F. Miao, "A PSO-CNN-aBiGRU-based gas concentration prediction method," *Mining Res. Develop.*, vol. 41, no. 12, pp. 76–81, 2021, doi: 10.13827/j.cnki.kyyk.2021.12.002.

[7] Q. Xu, K. Xu, L. Li, and X. Yao, "Mine safety assessment based on basic event importance: Grey relational analysis and bow tie model," *Roy. Soc. Open Sci.*, vol. 5, no. 8, Aug. 2018, Art. no. 180397, doi: 10.1098/rsos.180397.

[8] Y. Cai, S. Wu, M. Zhou, S. Gao, and H. Yu, "Early warning of gas concentration in coal mines production based on probability density machine," *Sensors*, vol. 21, no. 17, p. 5730, Aug. 2021, doi: 10.3390/s21175730.

[9] F. Cui, S. F. He, X. P. Lai, J. Q. Chen, B. C. Sun, C. Jia, and Y. J. Gao, "Study on B-value trend of time series of impact ground pressure mine based on phase space reconstruction and deep learning," *J. Coal*, pp. 1–13, Nov. 2022, doi: 10.13225/j.cnki.jccs.2022.0168.

[10] T. Ji, J. Wang, M. Li, and Q. Wu, "Short-term wind power forecast based on chaotic analysis and multivariate phase space reconstruction," *Energy Convers. Manage.*, vol. 254, Feb. 2022, Art. no. 115196.

[11] F. A. Hashim and A. G. Hussien, "Snake optimizer: A novel meta-heuristic optimization algorithm," *Knowl.-Based Syst.*, vol. 242, Apr. 2022, Art. no. 108320.

[12] J. Feng, J. Zhang, X. Zhu, and W. Lian, "A novel chaos optimization algorithm," *Multimedia Tools Appl.*, vol. 76, no. 16, pp. 17405–17436, Aug. 2017.

[13] I. Al-Shourbaji, P. H. Kachare, S. Alshathri, S. Duraibi, B. Elnaim, and M. Abd Elaziz, "An efficient parallel reptile search algorithm and snake optimizer approach for feature selection," *Mathematics*, vol. 10, no. 13, p. 2351, Jul. 2022.

[14] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.

[15] Q. Ge, C. Guo, H. Jiang, Z. Lu, G. Yao, J. Zhang, and Q. Hua, "Industrial power load forecasting method based on reinforcement learning and PSO-LSSVM," *IEEE Trans. Cybern.*, vol. 52, no. 2, pp. 1112–1124, Feb. 2022.

[16] Z. Wang and B. Yao, "Multi-branching temporal convolutional network for sepsis prediction," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 2, pp. 876–887, Feb. 2022.

[17] P. Ardimento, L. Aversano, M. L. Bernardi, M. Cimitile, and M. Iammarino, "Using deep temporal convolutional networks to just-in-time forecast technical debt principal," *J. Syst. Softw.*, vol. 194, Dec. 2022, Art. no. 111481.

[18] W. Li, Y. Wei, D. An, Y. Jiao, and Q. Wei, "LSTM-TCN: Dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network," *Environ. Sci. Pollut. Res.*, vol. 29, no. 26, pp. 39545–39556, Jun. 2022.

[19] K. H. Liu, P. S. Zhong, D. F. Xu, Q. Xia, and M. Liu, "A modified linear cell based on hyperbolic tangent function," *Comput. Integr. Manuf. Syst.*, vol. 26, no. 1, pp. 145–151, 2020, doi: 10.13196/j.cims.2020.01.015.

[20] T. D. Ryck, S. Lanthaler, and S. Mishra, "On the approximation of functions by tanh neural networks," *Neural Netw.*, vol. 143, pp. 732–750, Nov. 2021, doi: 10.1016/j.neunet.2021.08.015.

[21] X. Gao, X. Li, B. Zhao, W. Ji, X. Jing, and Y. He, "Short-term electricity load forecasting model based on EMD-GRU with feature selection," *Energies*, vol. 12, no. 6, p. 1140, Mar. 2019.

[22] W. Zhang, Z. Lin, and X. Liu, "Short-term offshore wind power forecasting—A hybrid model based on discrete wavelet transform (DWT), seasonal autoregressive integrated moving average (SARIMA), and deep-learning-based long short-term memory (LSTM)," *Renew. Energy*, vol. 185, pp. 611–628, Feb. 2022.

HAOFAN SHI was born in Huhehaote, Inner Mongolia, China, in 1999. He received the B.S. degree in measurement and control technology and instrumentation from Liaoning Technical University, Huludao, China, in 2020, where he is currently pursuing the M.S. degree. His current research interests include intelligent detection, data fusion technology, and deep learning.

YAOSONG XU was born in Cangzhou, Hebei, China, in 1979. He received the B.S. degree in detection technology and instrumentation and the M.S. degree in power electronics and power transmission from Liaoning Technical University, Huludao, China, in 2001 and 2004, respectively, and the Ph.D. degree in precision instruments and machinery from Tianjin University, Tianjin, China, in 2012. He has been working as a Teacher with the Department of Measurement and Control Technology and Instrumentation, Liaoning Technical University. He is the author of three books, more than 30 articles, and ten inventions. His current research interests include power quality monitoring, grid load prediction, and electrical equipment fault diagnosis.

HUA FU was born in Fuxin, Liaoning, China, in 1962. She received the B.S. degree in electrical automation, the M.S. degree in power electronics and power transmission, and the Ph.D. degree in mining engineering from Liaoning Technical University, Huludao, China, in 1984, 1996, and 2002, respectively. She was a Senior Visiting Scholar at the University of Michigan, USA, in 2008, and the University of Technology Sydney, Australia, in 2016. She has been a Professor at Liaoning Technical University, Liaoning, China, since 1984. She is the author of six monographs, 21 teaching books, 120 articles, and 34 inventions. Her current research interests include coal mine gas monitoring, safety monitoring and surveillance, and intelligent detection and control.

JINGYU SHAO (Member, IEEE) received the B.S. degree from Beihang University, Beijing, China, the M.S. degree from the University of Technology Sydney, Sydney, Australia, and the Ph.D. degree from Australian National University, Canberra, Australia, in 2021. He is currently a Senior Engineer at Baidu Inc. His research interests include data mining, machine learning, deep learning, and natural language processing.

● ● ●