

Article

# Method and Validation of Coal Mine Gas Concentration Prediction by Integrating PSO Algorithm and LSTM Network

Guangyu Yang <sup>1,2</sup>, Quanjie Zhu <sup>3,\*</sup> , Dacang Wang <sup>4</sup>, Yu Feng <sup>5</sup>, Xuexi Chen <sup>4</sup> and Qingsong Li <sup>6</sup><sup>1</sup> Coal Mining Research Institute Co., Ltd. of CCTEG, Beijing 100013, China; ustb\_guangyu@163.com<sup>2</sup> State Key Laboratory of Coal Mining and Clean Utilization, Beijing 100013, China<sup>3</sup> School of Emergency Technology and Management, North China Institute of Science and Technology, Sanhe 065201, China<sup>4</sup> School of Mine Safety, North China Institute of Science and Technology, Sanhe 065201, China; wdc15102587295@163.com (D.W.); xuexichen@ncist.edu.cn (X.C.)<sup>5</sup> School of Civil Engineering, Sun Yat-sen University, Zhuhai 519082, China; fengy253@mail.sysu.edu.cn<sup>6</sup> Guizhou Mine Safety Scientific Research Institute Co., Ltd., Guiyang 550025, China; xili3549@gmail.com

\* Correspondence: zhqj2016@ncist.edu.cn

**Abstract:** Gas concentration monitoring is an effective method for predicting gas disasters in mines. In response to the shortcomings of low efficiency and accuracy in conventional gas concentration prediction, a new method for gas concentration prediction based on Particle Swarm Optimization and Long Short-Term Memory Network (PSO-LSTM) is proposed. First, the principle of the PSO-LSTM fusion model is analyzed, and the PSO-LSTM gas concentration analysis and prediction model is constructed. Second, the gas concentration data are normalized and preprocessed. The PSO algorithm is utilized to optimize the training set of the LSTM model, facilitating the selection of the training data set for the LSTM model. Finally, the MAE, RMSE, and coefficient of determination  $R^2$  evaluation indicators are proposed to verify and analyze the prediction results. Gas concentration prediction comparison and verification research was conducted using gas concentration data measured in a mine as the sample data. The experimental results show that: (1) The maximum RMSE predicted using the PSO-LSTM model is 0.0029, and the minimum RMSE is 0.0010 when the sample size changes. This verifies the reliability of the prediction effect of the PSO-LSTM model. (2) The predictive performance of all models ranks as follows: PSO-LSTM > SVR-LSTM > LSTM > PSO-GRU. Comparative analysis with the LSTM model demonstrates that the PSO-LSTM model is more effective in predicting gas concentration, further confirming the superiority of this model in gas concentration prediction.

**Keywords:** PSO; LSTM; gas concentration prediction; big data utilization

**Citation:** Yang, G.; Zhu, Q.; Wang, D.; Feng, Y.; Chen, X.; Li, Q. Method and Validation of Coal Mine Gas Concentration Prediction by Integrating PSO Algorithm and LSTM Network. *Processes* **2024**, *12*, 898. <https://doi.org/10.3390/pr12050898>

Academic Editor: Qingbang Meng

Received: 28 March 2024

Revised: 24 April 2024

Accepted: 26 April 2024

Published: 28 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Coal mine gas accidents are characterized by their suddenness and unpredictability. For a long time, the prevention and control of gas disasters in coal mines has been a crucial task in the safety management of high-gas coal mines. The prediction of gas concentration is of great significance for the safety and production management of coal mines. On one hand, the prediction of gas concentration holds an extremely important position in the safety production of coal mines. The accurate prediction of gas concentration enables timely implementation of measures such as ventilation and gas drainage to ensure miners' safety and the stable operation of the mine. On the other hand, predicting gas concentrations also helps enhance the production efficiency of coal mines. By monitoring and predicting gas concentrations in real-time, production operations can be scheduled effectively, minimizing production downtime caused by gas levels exceeding the standard. This, in turn, enhances the economic benefits of coal mines. Predicting gas concentrations is a critical technical method for ensuring safe production in coal mines. Accurately predicting gas concentrations enables the timely implementation of measures such as ventilation and gas extraction

to prevent gas explosions, reduce casualties, and is of great significance in safeguarding personnel safety, ensuring production continuity, and enhancing economic benefits.

The prediction of gas concentration in mine roadways is a typical nonlinear problem. This is because the generation, release, transport, and accumulation processes of gas are influenced by a variety of complex factors, such as geological conditions, gas content in coal seams, mining methods, and ventilation conditions. The interactions of these factors make the changes in gas concentration highly nonlinear, thus complicating predictions using traditional methods. To enhance the accuracy and applicability of gas concentration prediction models, experts and scholars from both domestic and international communities have endeavored to tackle this issue through various methods. They have proposed numerous prediction models, including conventional methods and artificial intelligence approaches, to offer academic support for the prevention and control of coal mine gas disasters. Conventional methods include Random Forest [1–3], Regression Trees [4,5], and Multiple Linear Regression [6,7]. These conventional methods for predicting gas concentrations have certain shortcomings. Firstly, conventional methods often rely on linear assumptions which makes it difficult to accurately capture the nonlinear characteristics of gas concentration changes, leading to less accurate prediction results. Secondly, conventional prediction models may rely too heavily on specific datasets, lacking sufficient generalizability. When faced with new or changing conditions, the predictive performance of the model may significantly decline. Furthermore, as mining conditions and environments change, conventional methods struggle to adapt quickly to new changes. Adjusting model parameters and structures can be challenging and time-consuming.

With the development of computer technology, methods for predicting gas concentrations have been continuously improved. In recent years, artificial intelligence technologies, especially machine learning and deep learning methods, have been increasingly applied to predict gas concentrations. These methods can handle highly nonlinear data, learn complex data patterns, and provide more accurate and robust prediction results. For example, techniques such as Artificial Neural Networks (ANN) [8–10], Support Vector Machines (SVM) [11–13], and LSTM [14] can recognize complex nonlinear relationships between gas concentration and various influencing factors through training, thereby enhancing the accuracy and reliability of predictions. Moreover, these methods exhibit good adaptability and flexibility, enabling them to continuously optimize prediction models as data are updated and accumulated. For instance, Li Huan et al. [15] proposed a gas concentration prediction model based on an improved Ant Colony Algorithm—Least Squares Support Vector Machine (ACO-LS-SVM) using Support Vector Machines (SVM). This model addresses issues such as the LS-SVM model's tendency to fall into local optima, slow search efficiency, and premature convergence. To address issues such as the complexity of prediction models and the demand for numerous parameters, Li Shugang et al. [16] constructed a coal mine working face gas concentration prediction model based on re-current neural networks. This model exhibited smaller training errors compared to backpropagation neural network prediction models and bidirectional recurrent neural network prediction models. Zhang Zhen et al. [17] developed a mine gas concentration prediction model using Keras Long Short-Term Memory networks, which exhibited higher prediction accuracy and a longer effective prediction step length. Fu Hua et al. [18] proposed constructing a multi-sensor gas concentration prediction model using deep LSTM networks. Liu Chao et al. [19] proposed a gas concentration prediction model based on CNN-GRU and an LSTM working face gas concentration prediction model using Pearson feature selection. This effectively addresses the challenge that traditional methods face in effectively linking gas concentration with its influencing factors when dealing with a large number of features and big data. Cao et al. categorized methane hazard levels into four categories: normal, attention, warning, and danger. They utilized a coal mine working face methane hazard prediction and early warning model based on Least Squares Support Vector (LS-SVM) multiclassifiers and regressors to forecast changes in methane concentration, achieving commendable predictive

accuracy [12]. Brodny et al. proposed a diagnostic and prognostic development method capable of determining short-term methane sensitivity predictions [20].

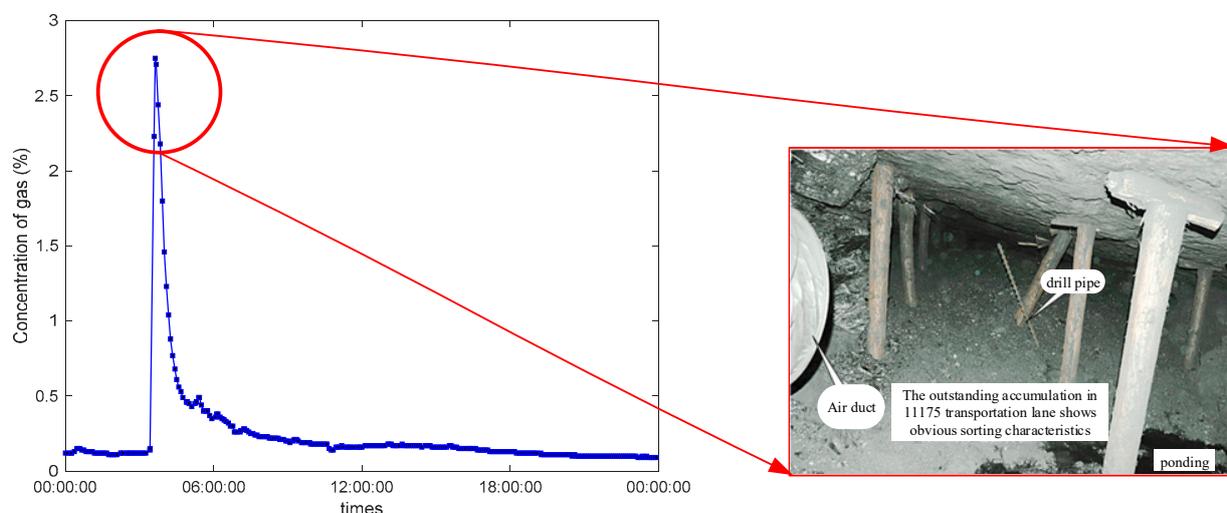
Against the backdrop mentioned above, addressing issues such as the low prediction accuracy of conventional gas concentration prediction models and the complexity of efficient prediction model structures, research on gas concentration prediction using a Particle Swarm Optimization-Long Short-Term Memory network (PSO-LSTM) has been conducted.

## 2. Gas Concentration Data and Its Characteristics

When gas accidents occur, there are often noticeable anomalies in gas concentration. Gas concentration is a type of time-varying data that naturally possesses continuity over time. Exploring its time-varying patterns and predicting its trends are of significant importance for the prevention of coal mine gas accidents. To illustrate this point, let us analyze two instances of gas outburst accidents. During these accidents, gas sensors in critical areas recorded changes in gas concentration within the tunnels, and significant abnormal changes in concentration were observed before the accidents occurred.

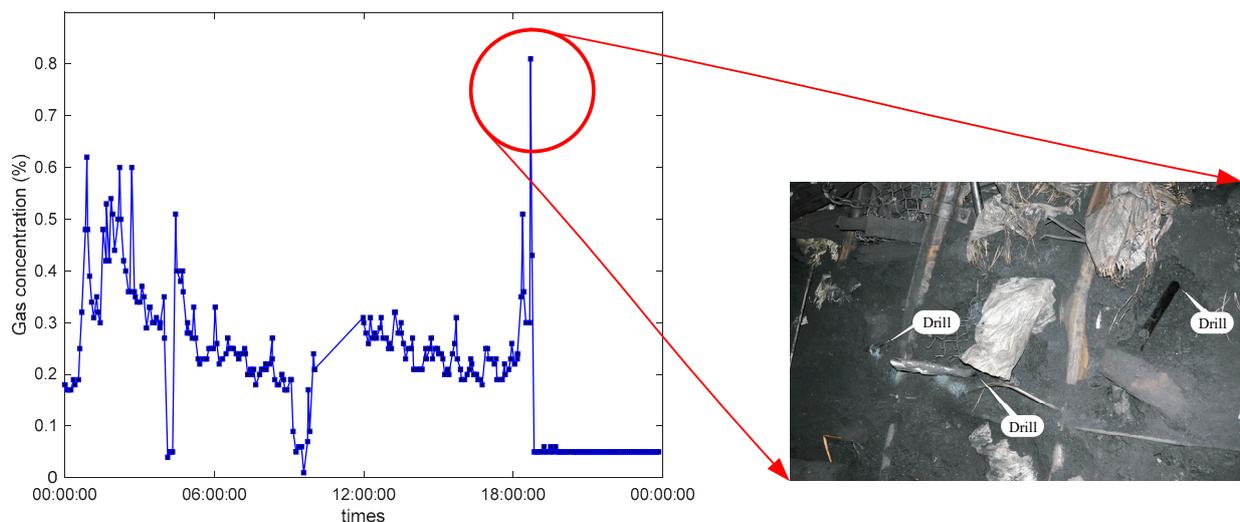
### 2.1. Abnormal of Gas Concentration at the Time of Coal and Gas Outburst

- (1) **Case Study One.** A mine in Guizhou Province was developed with an inclined shaft. Both the main and auxiliary inclined shafts connect to the main and auxiliary mine doors at the elevation of +1410 m, respectively. They are linked to the belt conveyance downhill and track conveyance downhill. The belt conveyance downhill, track conveyance downhill, and the return air inclined shaft were leveled and interconnected at the elevation of +1358 m. The coal and gas outburst accident occurred at the 11,175 open-off cut (drift) development face of the mine. At the time of the accident, the excavation of the open-off cut was progressing upward from the 11,175 transport tunnel, with an inclined length of about 5 m. The coal seam at the 11,175 open-off cut had an inclination angle of approximately six degrees and a coal thickness of around 1.7 m. The accident site was at an elevation of about +1375 m, corresponding to a ground elevation of about +1885 m, with a burial depth of about 510 m. On the day of the accident, the mine utilized a gas concentration monitoring system to gather data on gas concentration at intervals of every 5 min. For this case, 288 gas concentration data points collected from 0:00 on 29 November 2012 to 0:00 the following day were selected. The first 200 data points were used as the training set, and the remaining 88 data points were used as the test set. The PSO-SLTM model was used to predict changes in gas concentration. A typical gas concentration dataset is shown in Figure 1.



**Figure 1.** Typical gas concentration data one.

- (2) Case Study 2: A mine in Guizhou Province was developed with inclined shafts, including a main inclined shaft, an auxiliary inclined shaft, and a return air inclined shaft. The elevation of the main inclined shaft's entrance is +1234.5 m, the auxiliary inclined shaft's entrance is +1236 m, and the return air inclined shaft's entrance is +1240.963 m. A significant carbon monoxide poisoning incident occurred in this mine on 29 March 2012 at the 2902 transport tunnel heading face in the second mining area. On the day of the accident, the mine's gas concentration monitoring system collected data on gas concentration at 5 min intervals. For this analysis, 288 gas concentration data points collected from 00:00 29 March 2012 to the following day at 00:00 were selected. The first 200 data points were used as the training set, and the remaining 88 were used as the test set for the predictive analysis of gas concentration changes using the PSO-SLTM model. A typical gas concentration dataset is shown in Figure 2.



**Figure 2.** Typical gas concentration data two.

### 2.2. Characteristics of Gas Concentration Data and the Difficulty of Prediction

Predicting gas concentration in mine roadways is a typical nonlinear problem. The generation, release, migration, and accumulation processes of gas are influenced by a variety of complex factors, such as geological conditions, gas content in coal seams, mining methods, and ventilation conditions. These factors interact with each other, making the variation in gas concentration highly nonlinear and challenging to predict using conventional methods.

The difficulties mainly lie in the following aspects: (1) Multivariate influence: Roadway gas concentration is affected by numerous variables that have complex interrelations, making it difficult to describe using simple linear models. (2) Data nonlinearity: The relationship between gas concentration and its influencing factors is often nonlinear and may vary over time and spatial location. (3) Dynamic changes occur in the gas concentration in roadways as mining activities progress, which adds complexity to predicting it because of temporal and spatial variations. (4) Data quality and availability: High-quality, high-frequency monitoring data are essential for accurate predictions. However, in practice, there may be difficulties in data collection due to missing data or noise interference.

The complexity of the coal mining environment and operational processes suggests that the variation in gas concentration at the working face exhibits the following characteristics [21]:

- (1) Periodicity: The gas concentration underground is influenced by various factors, leading to periodic variations with similar magnitudes across most periods. By learning the periodic variations in gas concentration characteristics through machine learning, it can be used to predict its changes in the next period.
- (2) Trend: Gas concentration data are a time series that change over time and exhibit trend characteristics, such as level, rising, and falling trends. Trend changes are one of the important foundations for making predictions.

- (3) Volatility and high nonlinearity: The complex mining conditions underground lead to significant fluctuations in gas concentration monitoring data, resulting in high nonlinearity.

These characteristics indicate that models constructed using traditional statistical methods or basic machine learning techniques may not effectively capture the underlying patterns of gas concentration. Therefore, it is necessary to construct appropriate deep learning models to thoroughly analyze and understand the variation characteristics of gas concentration data. This will effectively predict gas concentration changes over a future period, enhance the predictive capability of coal mine gas safety situations, and offer decision support for routine gas monitoring and supervision at the working face.

To this end, a combined prediction model based on PSO-LSTM is proposed for precise research on coal mine gas concentration prediction, using the gas concentration data collected from a mine in Guizhou as an example. This study initially models typical time-series data using the Long Short-Term Memory Recurrent Neural Network model (LSTM). It then calculates the key parameters of the combined model using the Particle Swarm Optimization (PSO) algorithm and finally compares this model with conventional prediction models. The results show that the combined LSTM-PSO model proposed in this paper outperforms conventional models in prediction accuracy and performance.

### 3. Introduction to Prediction Theory and Method

#### 3.1. PSO Algorithm

Particle Swarm Optimization (PSO) is an algorithm inspired by the study of bird flocking behavior, with its core idea being the search for optimal solutions through collaboration and information sharing among individuals within the group [22]. The algorithm uses mass-less particles to simulate individuals within a bird flock, each possessing only velocity and position attributes. Velocity represents the direction and distance of the next iteration's movement, while position represents a potential solution to the problem being addressed. PSO initiates with all individuals foraging within their own space; when an individual finds food (the optimal solution to the problem), that solution becomes the individual's best solution ( $P_{id,pbest}$ ). The best solutions of all individuals are shared within the group, and by evaluating the quality of solutions with a fitness function, the group's best solution ( $P_{d,gbest}$ ) is determined. In the next iteration, each individual updates its velocity and position based on the group's best solution, continuing the foraging process in a cycle until a unique optimal solution is obtained.

Assuming there are  $N$  particles in a  $D$ -dimensional search space, where each particle represents a potential solution, then the position of the  $i^{th}$  particle can be denoted as  $X_{id} = \{x_{i1}, x_{i2}, \dots, x_{iD}\}$ , and the velocity of the  $i^{th}$  particle can be denoted as  $V_{id} = \{v_{i1}, v_{i2}, \dots, v_{iD}\}$ . The best position found using the  $i^{th}$  particle (the personal best solution) is represented as  $P_{id,pbest} = \{p_{i1}, p_{i2}, \dots, p_{iD}\}$ , and the best position found using the entire swarm (the global best solution) is represented as  $P_{d,gbest} = \{p_{1,gbest}, p_{2,gbest}, \dots, p_{D,gbest}\}$ . The individual historical best fitness value is denoted by  $f_p$ , and the swarm's historical best fitness value is denoted by  $f_g$ . The PSO algorithm is characterized by its straightforward structure and ease of implementation, boasting advantages such as high precision and short computation time [23,24].

PSO is initialized with a group of random particles (random solutions). Then, it iteratively finds the optimal solution. In each iteration, particles update themselves by tracking two "extremes" (pbest and gbest). After finding these two optimal values, particles update their velocity and position using the following formula:

$$v_{id}^{k+1} = wv_{id}^k + c_1r_1(P_{id,pbest}^k - x_{id}^k) + c_2r_2(P_{d,gbest}^k - x_{id}^k) \quad (1)$$

Location update formula:

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \quad (2)$$

In the formula:  $i$  represents the particle index;  $d$  represents the dimension index of the particle;  $k$  is the iteration number;  $w$  is the inertia weight;  $c_1$  is the cognitive (individual) learning factor;  $c_2$  is the social (group) learning factor;  $r_1$  and  $r_2$  are random numbers between 0 and 1;  $v_{id}^k$  is the velocity vector of particle  $i$  in dimension  $d$  during iteration  $k$ ;  $P_{id,pbest}^k$  is the historical best position of particle  $i$  in dimension  $d$  during iteration  $k$ ;  $P_{d,gbest}^k$  is the historical best position of the group in dimension  $d$  during iteration  $k$ .

### 3.2. LSTM Model

The LSTM network model integrates short-term memory with long-term memory through gated controls, exhibiting excellent capabilities in processing long sequential data [25]. LSTM is composed of multiple repeating structural modules, each containing three gates: the forget gate, the input gate, and the output gate. The LSTM network controls these three gates through activation functions, thereby managing the retention and forgetting of historical information. The LSTM prediction model primarily includes three steps:

The forgetting gate  $f_t$  determined using the Sigmoid function layer is calculated by the following formula:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

where,  $\sigma$  is the Sigmoid activation function;  $W_f$  is the forgetting gate weight matrix;  $h$  is neuron output;  $x$  is neuron input;  $t$  is the current moment;  $t - 1$  is the previous time;  $b$  is the offset item.

Add the input gate  $i_t$  of the state quantity, and optionally update. The calculation formula is shown in Equations (4)–(6):

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (4)$$

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

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (6)$$

where,  $\tilde{C}_t$  is the alternate update information at  $t$  times;  $\tanh(\cdot)$  is the activation function;  $W_c$  is the weight matrix of memory cells;  $C_t$  and  $C_{t-1}$  were the state of memory cells at  $t$  and  $t - 1$  moments, respectively.

Output gate  $O_t$ , the calculation formula is as follows:

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (7)$$

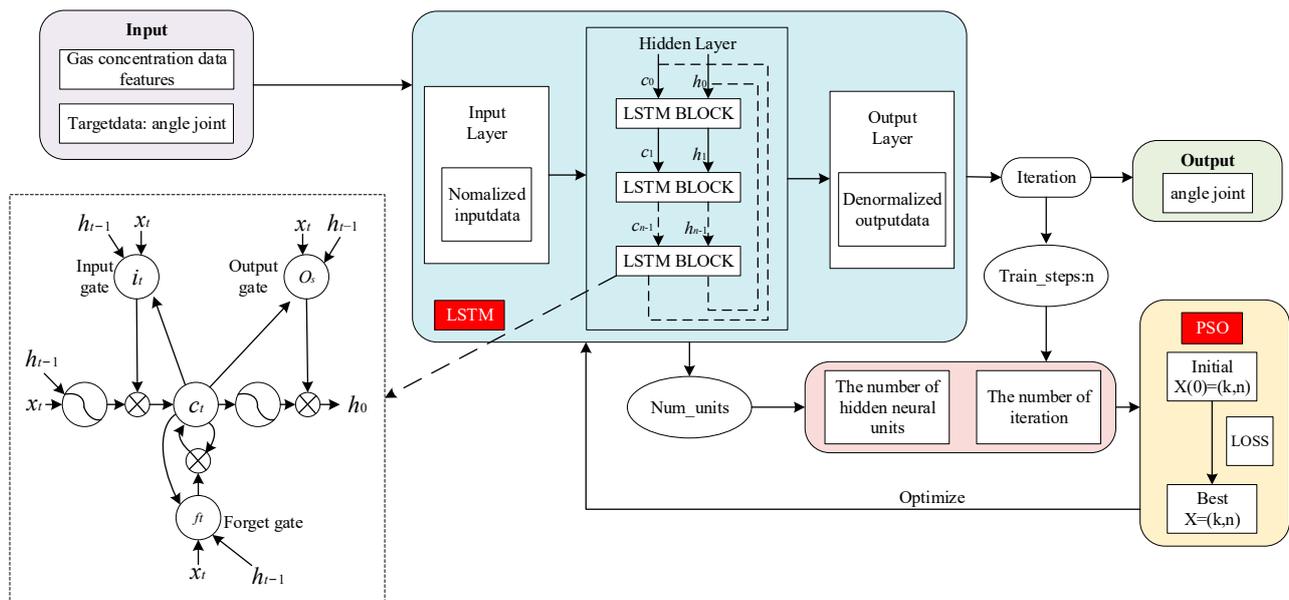
$$h_t = O_t \tanh C_t \quad (8)$$

where,  $W_o$  is the output gate weight matrix.

The LSTM network is a type of Recurrent Neural Network (RNN) that processes input data by traversing time steps and updating the RNN state. The RNN state contains information remembered from all previous time steps. LSTM neural networks can predict subsequent values of a time series or sequence by using previous time steps as input.

### 3.3. PSO-LSTM Algorithm

Time series data contain a large amount of uncertain information and predictions made with a single model often are not ideal. Therefore, to improve prediction accuracy, combining the LSTM and PSO models and applying the PSO algorithm to solve for the combination coefficients can speed up the solution process and fully leverage the advantages of the combined model [26]. Compared to individual LSTM and other prediction methods, PSO-LSTM has stronger global search capabilities, adaptability, the ability to overcome local optima, and suitability for large-scale data, enhancing prediction accuracy and stability. Based on this, the paper proposes a mine gas concentration prediction model based on PSO-LSTM (as shown in Figure 3).



**Figure 3.** Structure diagram of PSO-LSTM model.

For illustrative purposes, we formulate the weight update formula for the LSTM algorithm part of the PSO-LSTM model as:

$$W_k = W_{k-1} + \Delta W \quad (9)$$

where  $W_t$  is the weight of the  $t$ th iteration;  $W_{t-1}$  is the weight of the  $t-1$ st iteration; and  $\Delta W$  is the increment of the weight.

In this paper, we have chosen to use a LSTM model for the prediction of gas concentration in the roadway. Generally, the LSTM model is highly influenced by its parameters; therefore, the LSTM needs to be optimized with the help of other methods. Here the PSO particle swarm algorithm is used to optimize the selection of this parameter, and the process is to first define the problem, i.e., the key parameters of the LSTM model need to be defined, and the objective function that needs to be optimized. In this example, we choose to optimize the accuracy of the LSTM. Then, the key parameters of the LSTM are encoded and a one-dimensional vector of the parameters is constructed, as in Section 3.1: the variable  $P$ . Next, the particle swarm is initialized, a random velocity is assigned to each particle, and the initial position of each particle is set to a random combination of parameters, and the accuracy and fitness of each type of parameter are calculated. The best position of each individual particle is compared with the global best position by constantly adjusting it. According to the formula of particle swarm optimization, the speed and position of each particle are updated until the iterative requirements are met, i.e., the optimal position is reached. The parameters selected at this time are the optimal parameters, as shown in Equation (9).

Compared to individual models, the PSO-LSTM combined model, which integrates the PSO algorithm and LSTM neural network, effectively solves the challenge of establishing key LSTM parameters and has the following advantages: a strong global search capability, high adaptability, the ability to overcome local optima, and suitability for large-scale data. This makes PSO-LSTM highly practical and scalable for real-world engineering applications.

## 4. Prediction Model Construction and Evaluation Method

### 4.1. Data Preprocessing

Based on the research presented earlier, this chapter will conduct studies from three aspects: data preprocessing, model construction steps, and model performance evaluation

methods. Due to the influence of the on-site environment and gas monitoring equipment, the monitoring data often have issues such as missing values and a low signal-to-noise ratio. Data cleaning is required before prediction. The data cleaning process can be divided into two steps: missing value detection, and missing value imputation. Missing values are filled using linear interpolation, which involves setting a sliding window to calculate the average value over a certain time period and replacing the missing values with this average. Specifically, the positions of missing values in the data are identified, and the average of historical and recent data is used to fill in these gaps, obtaining complete original input data.

Due to complex underground environmental conditions and issues such as equipment failures and communication breakdowns, gas concentration data often have missing values, which can make predictions difficult. Based on the superiority of spline interpolation in filling missing values in point source time series data, spline interpolation can be used to supplement missing values in gas concentration time series data.

The Spline Interpolation Method (SIM) is a special type of piecewise cubic polynomial interpolation method. Compared to ordinary polynomial interpolation, spline interpolation typically offers a smoother fit to the dataset and results in smaller interpolation errors. Given  $n + 1$  distinct observation times  $t_i$ , satisfying  $t_0 < t_1 < \dots < t_{n-1} < t_n$ , and  $n + 1$  observed values  $Y(t_i)$ , the essence of spline interpolation is to construct an  $n$ -th degree spline function  $Y(t_i)$  that approximates the observed dataset.

$$Y(t) = \begin{cases} Y_0(t) & t \in (t_0, t_1) \\ Y_1(t) & t \in (t_1, t_2) \\ \dots & \dots \\ Y_{n-1}(t) & t \in (t_{n-1}, t_n) \end{cases} \quad (10)$$

$Y(t)$  represents a set of gas concentration data with missing values, as shown in Figure 4. After filling in the missing values using spline interpolation, the result is as shown in Figure 5. From Figure 4, it can be seen that there are multiple gaps in the data, which are related to the on-site environment and monitoring equipment. These gaps are often caused by coal mine tunnel power outages or construction activities that result in sensor disconnections. Missing data can impact the extraction of time series data features and the accuracy of predictions.

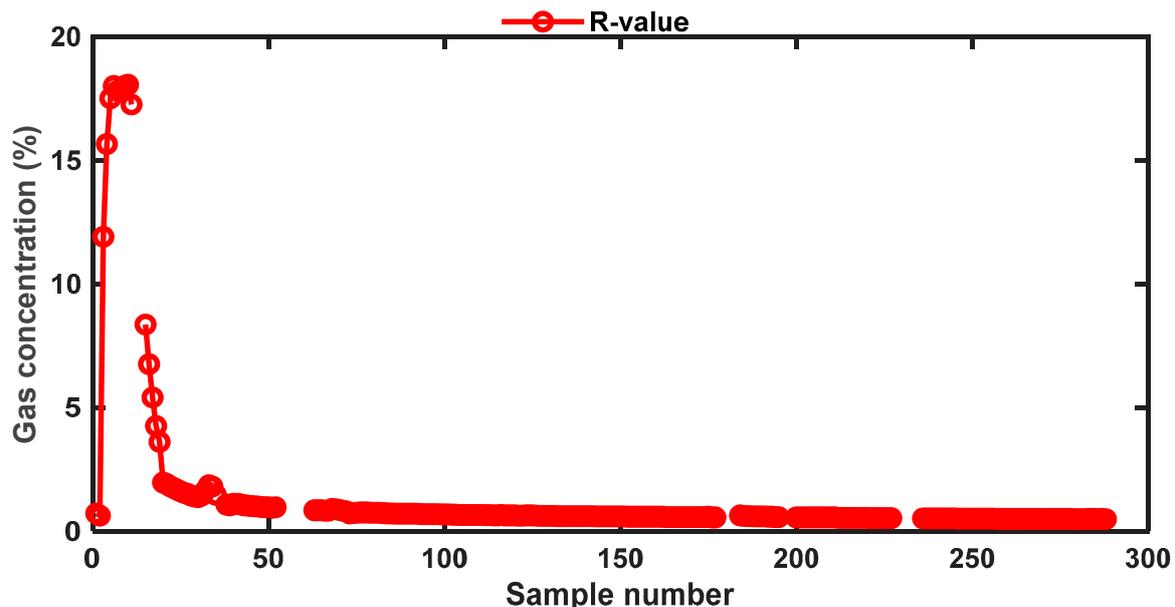


Figure 4. Raw time series with missing values.

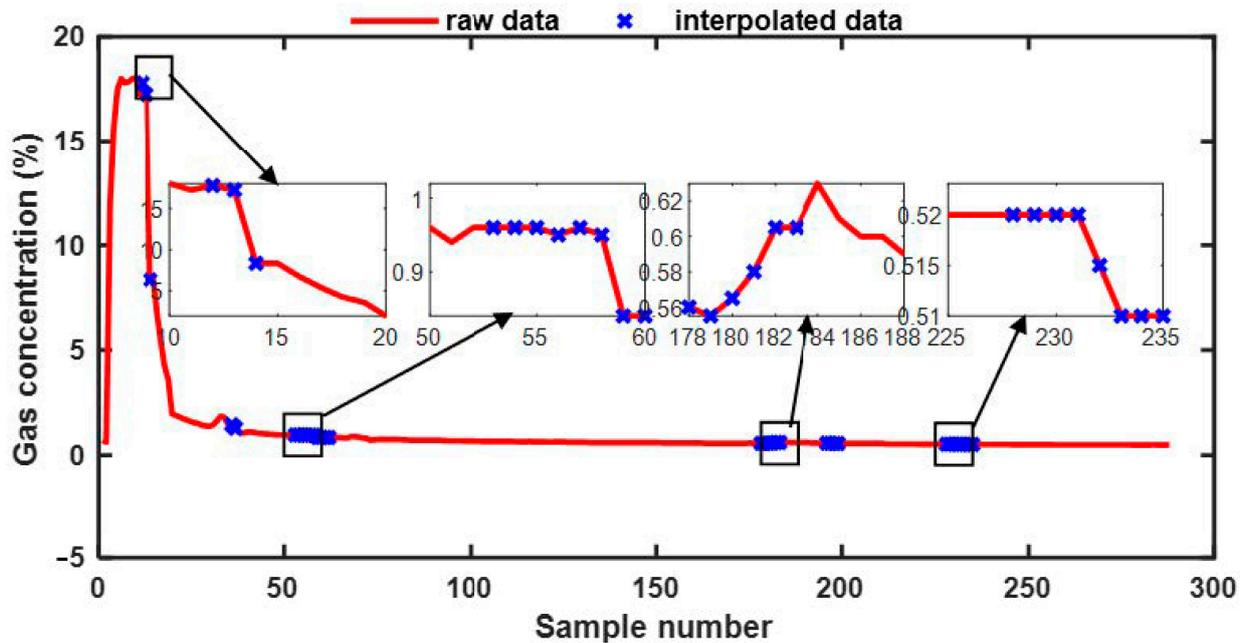


Figure 5. Gap-filled data series.

The steps for implementing the method may be briefly described as follows: Step 1, prepare relevant data, such as missing data caused by equipment failure, communication interruption, or other factors. Step 2 involves selecting relevant spline interpolation parameters, segmenting the data, and interpolating using information from neighboring data points to approximate the actual data.

To reduce the impact of this factor, preprocessing was performed using the method described above, as shown in Figure 5. This method effectively completes the supplementation of missing data (indicated by blue “×” symbols) and maintains a good fit with the original curve (red line), preserving the characteristics and trends of the original data. However, this method is limited by the sliding window parameters of the missing value filling method and further optimization is needed. The advantages of this method include the relatively smooth curve, the simplicity of the algorithm, and the accuracy of estimating missing data points.

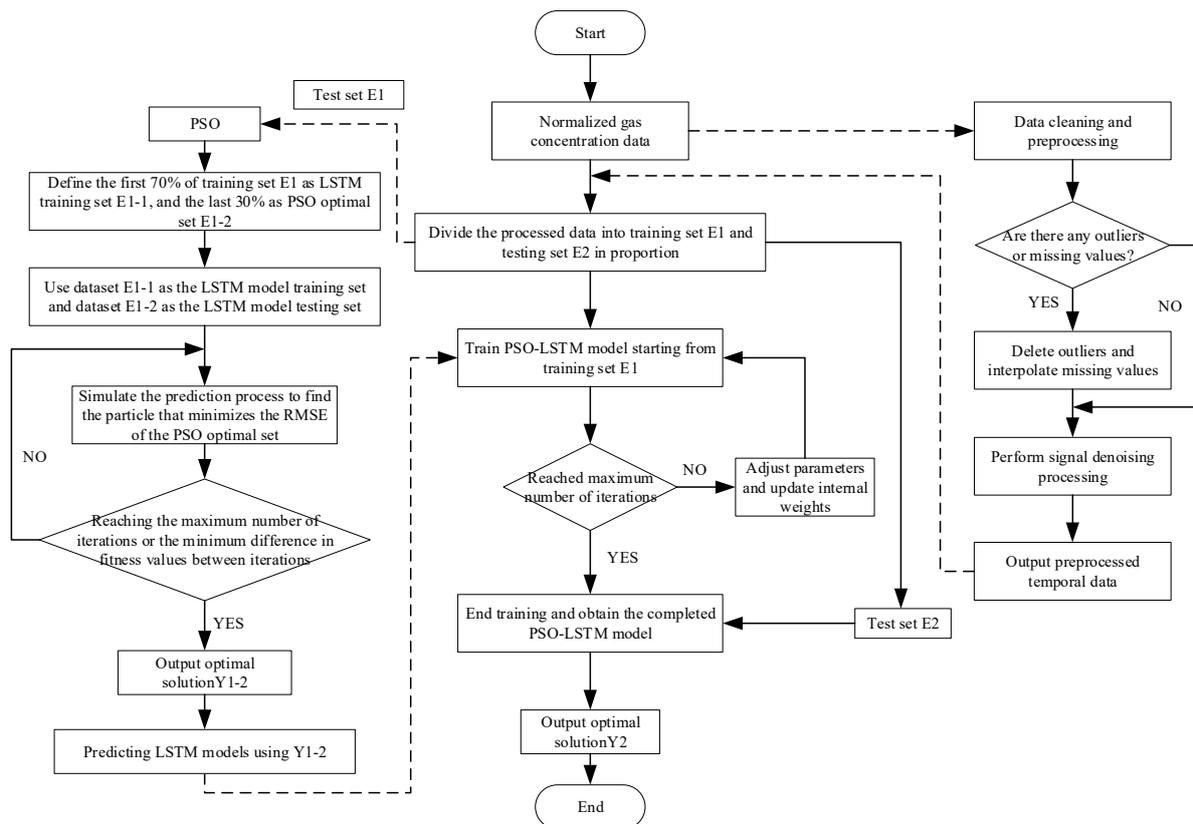
#### 4.2. Composite Model Building Steps

The principle of the PSO-LSTM model is to enhance the training of gas concentration data by combining the PSO algorithm and LSTM algorithm. The PSO algorithm optimizes the data, which are then used as the test set for the LSTM algorithm. This process ensures that the training results of the LSTM algorithm are more accurate. The flowchart is shown in Figure 6.

The specific process is as follows:

- (1) Step 1: The sample data processed using normalization are divided into training set  $E_1$  and test set  $E_2$  of the PSO-LSTM model in a ratio of 7:3.
- (2) Step 2: Start to optimize the training LSTM model with the training set  $E_1$  and PSO models. The specific process is as follows: ① initialize the PSO algorithm and form the optimization community; ② the first 70% of training set  $E_1$  is defined as the LSTM model training set  $E_{1-1}$  and the last 30% is defined as PSO model optimization set  $E_{1-2}$ . ③ Data set  $E_{1-1}$  is used as the training set of LSTM model and the optimal solution of the data set  $E_{1-2}$  is used as the test set of LSTM model; ④ start the simulation prediction process to find the particle (training result) that makes the PSO optimal selection RMSE minimum. At the same time, the LSTM model is trained using data set  $E_{1-1}$ . The Root Mean Square Error (RMSE) was selected as an evaluation index to detect whether the PSO model training reached the maximum number of iterations or

- the minimum difference in the adaptation values between iterations. ⑤ The optimal solution of the output data set  $E_{1-2}$ ,  $Y_{1-2} = \{Y_{1-2}^1, Y_{1-2}^2, \dots, Y_{1-2}^i\}$ ,  $i = \{1, 2, \dots, n\}$ .
- (3) Step 3: Use the optimal solution  $Y_{1-2}$  of data set  $E_{1-2}$  to perform a predictive analysis on the LSTM model that has been trained.
  - (4) Step 4: Adjust the model parameters according to the prediction results. Predict whether the training result reaches the maximum number of iterations. If not, adjust the parameters, update the internal weights, and return to the PSO-LSTM model again for training until the training result reaches the maximum number of iterations.
  - (5) Step 5: After the training of the PSO-LSTM model, the test set  $E_2$  was used to conduct predictive analysis of the trained PSO-LSTM model.
  - (6) Step 6: Output the optimal solution of the PSO-LSTM model,  $Y_2 = \{Y_2^1, Y_2^2, \dots, Y_2^i\}$ ,  $i = \{1, 2, \dots, n\}$ .



**Figure 6.** Flow chart of PSO-LSTM model.

In short, the PSO-LSTM time series prediction algorithm has advantages in global search, parameter adjustment, adaptive ability, scalability and easy implementation, which makes it have higher performance and efficiency when dealing with time series prediction problems.

#### 4.3. Performance Evaluation Method

Evaluating the performance of prediction models typically requires considering multiple indicators. The choice of suitable evaluation metrics depends on the specific circumstances. Generally, RMSE and MAE are used to analyze the average magnitude of errors, with RMSE being more sensitive to outliers [27]. Additionally, the Mean Error (ME) focuses more on whether the predictions are consistently too high or too low but does not reflect the magnitude of the errors. Based on this, the performance testing of the prediction model uses MAE, RMSE, ME, and R2 as criteria to judge the quality of time series prediction, providing a comprehensive set of information to assess the performance of the prediction

model. The mathematical expressions are presented in Equations (11)–(14). MAE can be expressed as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y - f(X)| \quad (11)$$

where  $Y$  represents the sequence of actual values,  $f(X)$  represents the sequence of predicted values, and  $N$  is the length of the sequence.

The Root Mean Squared Error (RMSE) can be expressed as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y - f(X))^2} \quad (12)$$

The Mean Error (ME) can be expressed as:

$$ME = \frac{1}{N} \sum_{i=1}^N (Y - f(X)) \quad (13)$$

The coefficient of determination  $R^2$  can be described as:

$$R^2 = 1 - \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (14)$$

In the formula,  $y_i$  stands for the values to be fitted, with their mean being  $\bar{y}$  and the fitted values being  $\hat{y}_i$ .

Furthermore, the introduction of  $R^2$  is an important metric used to evaluate the extent to which a regression model fits the data. It ranges from 0 to 1, indicating the proportion of the variance in the dependent variable that is predictable from the independent variable(s). The closer  $R^2$  is to 1, the better the model fits the data. Conversely, the closer  $R^2$  is to 0, the worse the model fits the data.

## 5. Parameter Selection and Model Evaluation

### 5.1. Evaluation of Model Key Parameters

When constructing a PSO-LSTM model, it is crucial to consider how to establish key parameters and the impact of the data itself on prediction accuracy, in order to enhance the model's performance [28]. Parameters within the LSTM neural network include the number of LSTM layers, the number of units, and the learning rate. Parameters for the PSO algorithm include the size of the swarm, maximum number of iterations (epochs), and inertia weight. Additionally, model training and validation parameters must be considered, including the division of training and validation sets, and the size of the time window (LSTM time steps). Establishing these parameters typically requires experimentation and tuning, with methods such as cross-validation used to assess the performance of the model under different parameter combinations, selecting the combination that performs best on the validation set. It is also important to be mindful of avoiding overfitting and underfitting to improve the model's generalization ability.

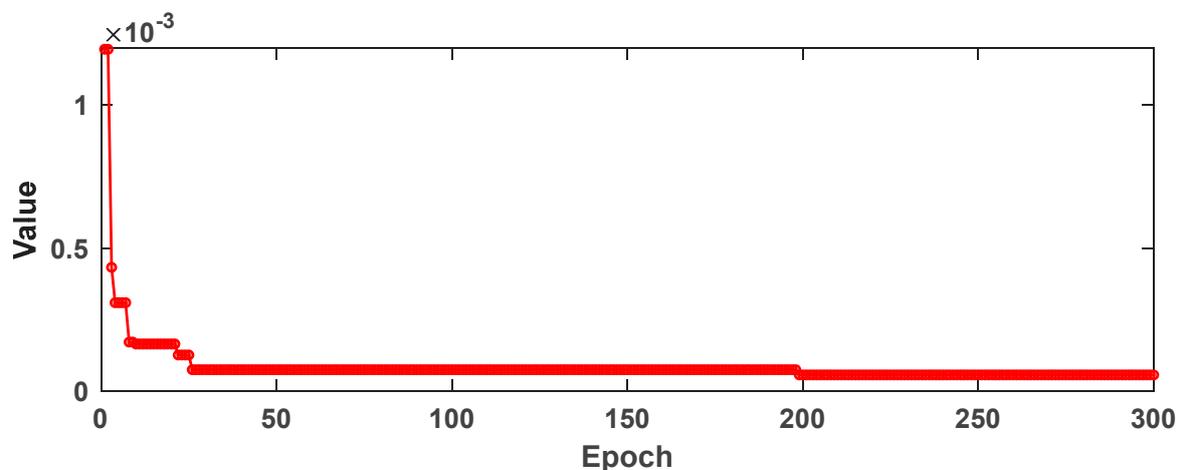
Identifying the key parameters of the LSTM model to achieve optimal performance is an iterative process, typically involving methods such as Grid Search, Random Search, and Bayesian Optimization for evaluation and tuning. This paper considers the superior characteristics of PSO and chooses to perform parameter tuning using the PSO algorithm, thereby obtaining the optimal parameter combination for the LSTM model. Initially, a PSO-LSTM model was constructed using MATLAB, establishing key parameters for the PSO algorithm. As shown in Table 1, the initial swarm size is set to 5, the dimensionality

of the swarm to 2, and the maximum number of iterations to 10, with learning factors  $c_1 = c_2 = 2$ , the maximum inertia weight at 1.2, and the minimum inertia weight at 0.8.

**Table 1.** PSO algorithm parameter.

Parameter	Parameter Value	Parameter	Parameter Value
Example initialize the number of communities	5	Learning factors $c_2$	2
Initialize the population dimension	2	Maximum inertia weight	1.2
Initialize the maximum number of group iterations	10	Minimum inertial weight	0.8
learning factors $c_1$	2	tolerance	$10^{-8}$

The key parameters of the LSTM network were obtained using PSO optimization calculation, and the convergence curve of the optimal fitness of the process was shown in Figure 7. The best fitness convergence curve of PSO is an important index to reflect the optimization process of the PSO algorithm, and the convergence curve can show the change trend of the fitness value of the PSO algorithm in the optimization process. As can be seen from the figure, when the epoch of the PSO algorithm belongs to the range [0, 25], the fitness of convergence curve decreases rapidly, indicating that the convergence speed is fast and the algorithm efficiency is high in the initial stage. It becomes stable after epoch = 25 with a value of  $7.49 \times 10^{-5}$ , and remains stable in the range of 25,200 for a long period of time. Then, there will be a slight decrease at epoch = 200 and then it will enter a stable state again, when the value reaches  $5.66 \times 10^{-5}$ . Considering the convergence rate and other factors, epoch = 50 is chosen as the global optimal solution. It can be seen from the figure that the fitness convergence curve of PSO tends to be stable and the fitness value is low, so it can be judged that the PSO algorithm has found the global optimal solution.



**Figure 7.** Convergence curve of the optimal fitness process.

Therefore, the relevant parameters of the LSTM model are established. The final optimization calculation results are shown in Table 2: the dimension of the input layer is 2, the dimension of the output layer is 1, the number of neurons of the hidden layer is 10–200, the learning rate is 0.01–0.15, the iteration number is 50, and a random gradient descent algorithm is adopted.

**Table 2.** PSO-LSTM algorithm parameter.

Parameter	Parameter Value	Parameter	Parameter Value
Input layer dimension	2	Solver	Adam
Output layer dimension	1	Learning rate	Self-adaptation
Number of hidden layer neurons	10~200	Number of iterations	50
LSTM layer	PSO optimization		

### 5.2. Influence Analysis of Data Length and Sampling Frequency

To verify the reliability of the PSO-LSTM model, an analysis was conducted using actual measured data from a mine, comparing different sample data volumes and frequencies of sample data changes. The analysis from the perspectives of data sample volume and sampling frequency reveals whether these factors impact the prediction model discussed earlier.

#### (1) Analysis of different sample data volumes

The gas monitoring system of the working face collected gas concentration data at 5 min intervals. For this study, 288 gas concentration data points collected from 0:00 28 November 2012 to 0:00 the next day were selected as the sample data to establish a mine gas concentration time series. To study the impact of different sample data volumes on model prediction accuracy, four gas concentration datasets of 6 h, 12 h, 24 h, and 48 h in duration were selected for training and prediction from different data length perspectives. The results are shown in Figure 8a–d; the localized area on the plot is zoomed in on the top of each of these plots, so that the detailed effect of that part can be seen. As can be seen, the predicted and actual values agree well, but there are still minor differences.

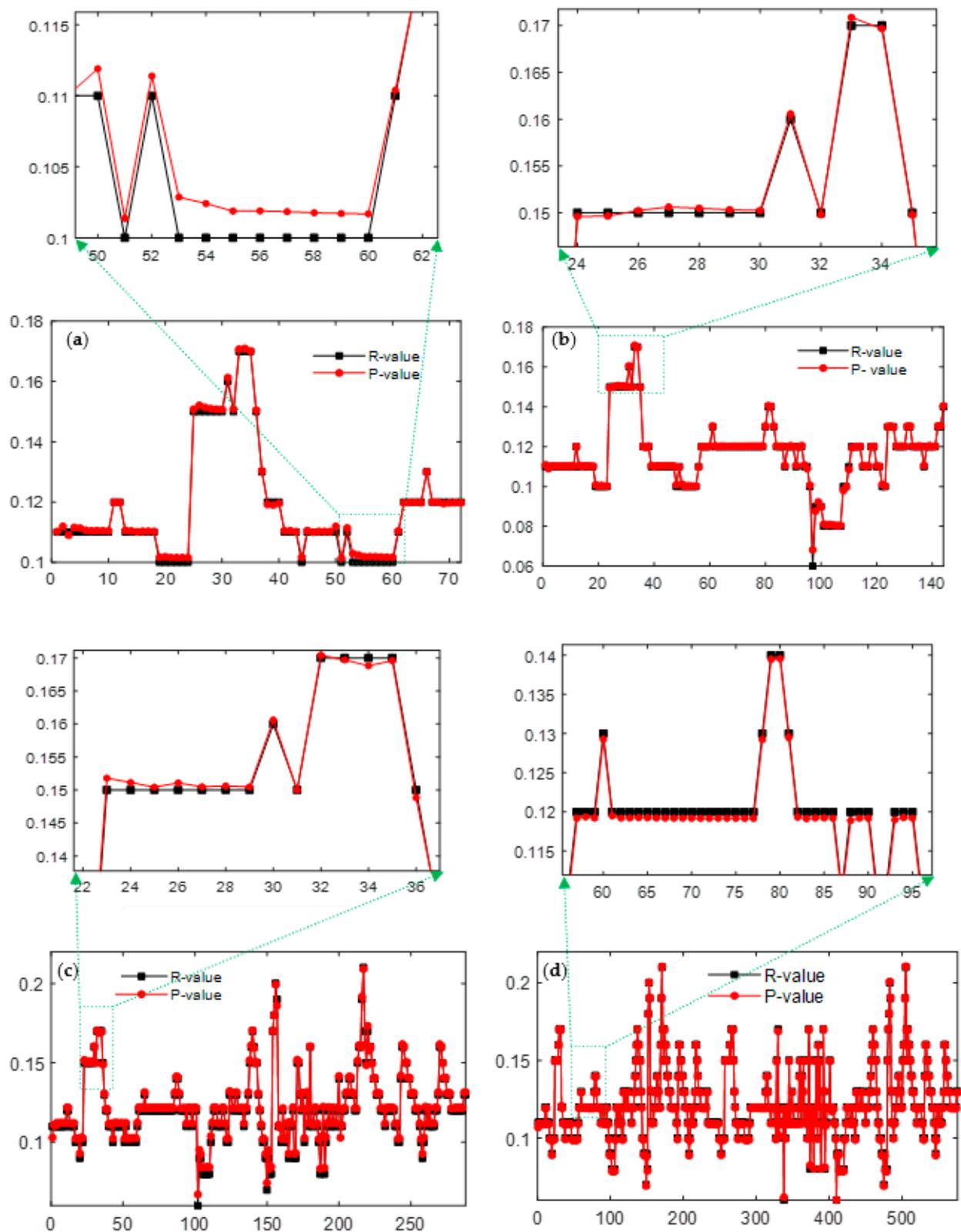
The prediction performance of the four different datasets was calculated, with the Root Mean Square Error (RMSE) values being 0.0013, 0.0010, 0.0010, and 0.0029 (with four significant figures retained), respectively. According to Figure 8, the 12 h and 24 h datasets yielded the smallest RMSE values at 0.0010, while the 48 h dataset had the largest RMSE value at 0.0029. Overall, the fluctuation range of RMSE values is not significant, indicating that the model does not require a high volume of gas concentration data and can adapt to data volume changes within 6 to 48 h.

#### (2) Prediction analysis of sample data volume with different amplitude of change

To verify the impact of the amplitude of data variation on the prediction results of the PSO-LSTM model, gas concentration data from a mine with a smaller frequency of variation were selected as a typical sample (Sample Two) for prediction using the PSO-LSTM model. The prediction's results are shown in Figure 9. Comparing the prediction results of Figure 9 (Sample Two) with those of Figure 8 (Sample One), it was found that the amplitude of gas concentration changes affects the accuracy of predictions. Specifically, the Root Mean Square Error (RMSE) for Sample Two, which had smaller amplitude changes, was lower than that for Sample One, with larger amplitude changes, indicating that Sample Two's prediction error was greater than Sample One's. In summary, when the amplitude of sample data variation is small, it is advisable to keep the total data quantity to no more than 288, meaning that for every 24 h (288 data points) as a data set the PSO-LSTM model yields higher prediction accuracy.

### 5.3. Predictive Performance Evaluation (Data10)

To verify the feasibility and effectiveness of the model, typical gas concentration data from a mine in Guizhou from August 2012 to November 2015 were selected for study. The original input data were first converted to a format compatible with the model's input dimensions. The transformed dataset was then divided into training, testing, and validation sets to verify the model's fitting effectiveness. The first 200 data points were chosen as the training set, the following 88 as the test set, and an additional 40 data points were selected as the validation set. The PSO-LSTM model was used for the gas concentration prediction study and the model's training set loss rate was calculated. The prediction results for the training, test, and validation sets are shown in Figure 10. As shown in Figure 10, on the left side of the figure the prediction results are shown and on the right side the corresponding local zoomed-in image is shown.



**Figure 8.** Analysis of the effect of training sets with different monitoring data lengths on the prediction results: (a) 6 h; (b) 12 h; (c) 24 h; (d) 48 h.

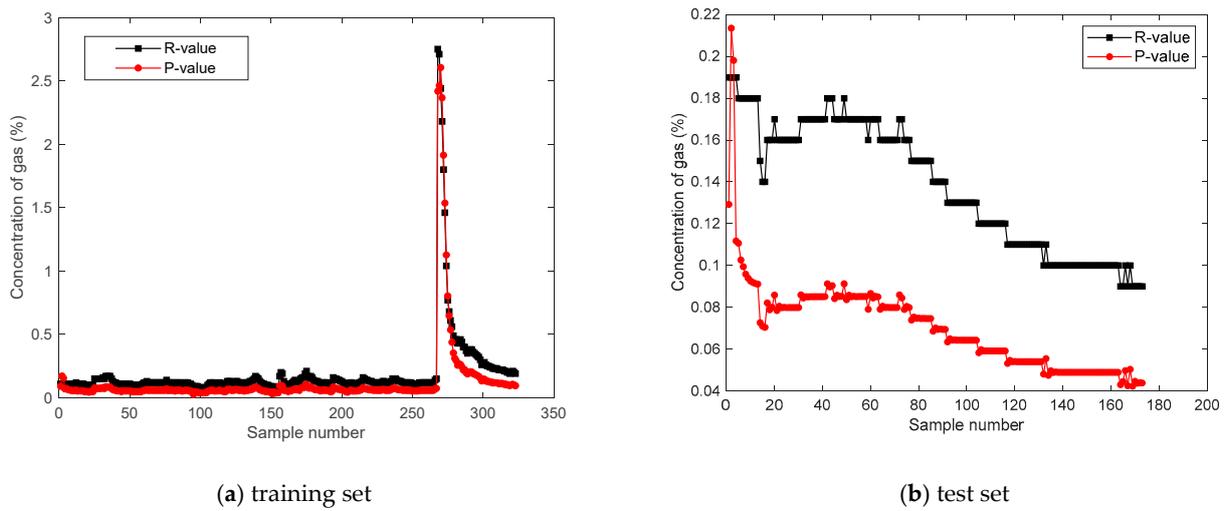


Figure 9. Prediction results of sample data of gas concentration in Sample Two.

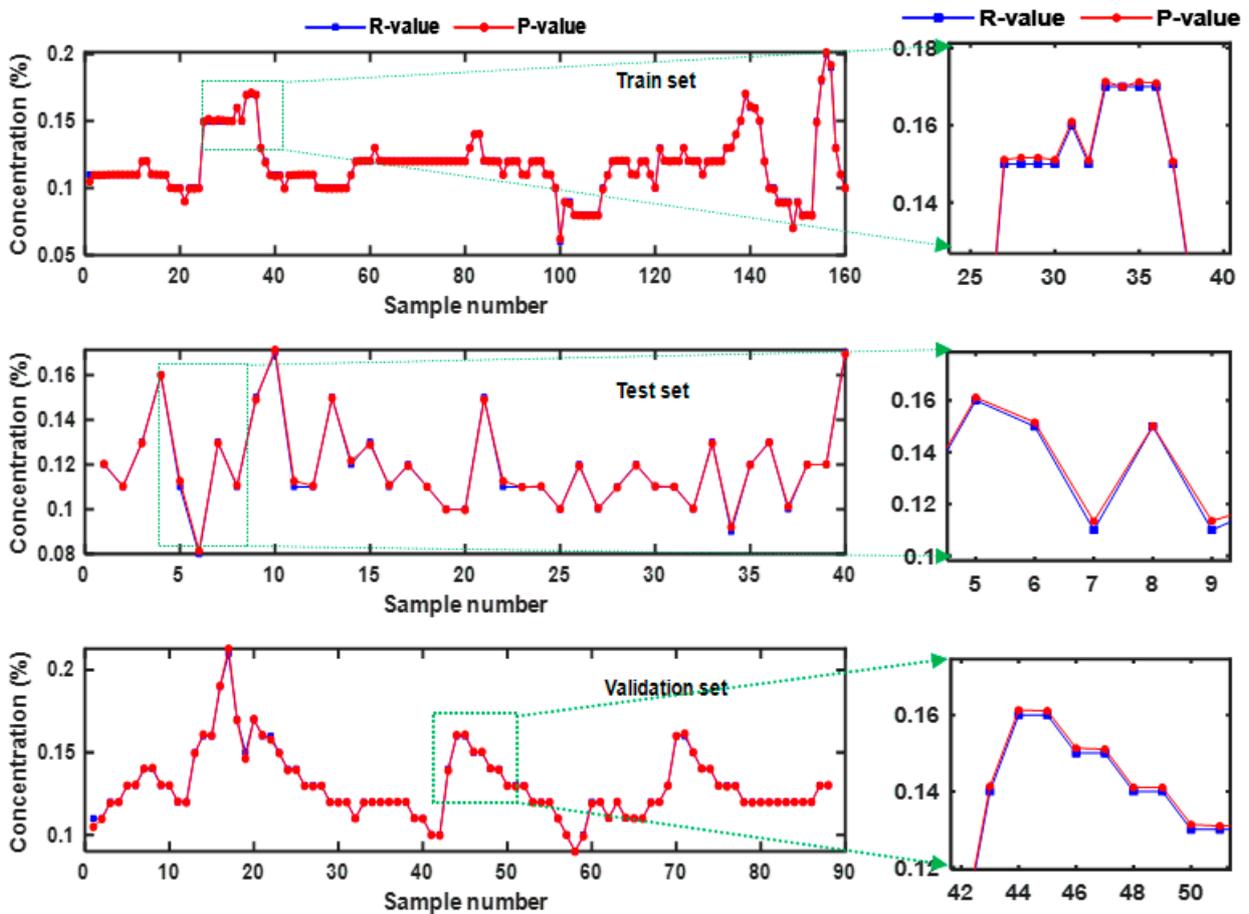


Figure 10. Training set, test set, validation set gas concentration prediction results.

To scientifically assess the model’s performance, the overall Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were calculated for each model to verify the feasibility and effectiveness of the method. The final results showed that the PSO-LSTM model’s gas concentration prediction was validated and analyzed using the evaluation formulas for MAE, RMSE, and the coefficient of determination  $R^2$ , yielding an MAE of 0.0008, an RMSE of 0.0010, and an  $R^2$  of 0.9987. The small MAE and RMSE indicate that the

model has a small measurement error, and the  $R^2$  value close to 1 suggests a high degree of model fit, demonstrating the model's practical value.

#### 5.4. Cross Validation Analysis (Data11)

The core idea of cross-validation is to divide the original dataset into groups, using some as the training set to train the model and others as the test set to evaluate the model. The benefits of cross-validation include evaluating the predictive performance of a model, particularly how well a trained model performs on new data, which can help to reduce overfitting to some extent. It also maximizes the extraction of useful information from limited data and aids in selecting the appropriate model. This paper opts for  $k$ -fold cross-validation, which reduces variance by averaging the results of training on  $k$  different groupings, thus decreasing the model's performance sensitivity to data division. Given the modest size of the dataset, a  $k$  value of 5 is chosen. Based on this, the MATLAB cross-validation function (such as `crossvalid`) is used to divide the data. A loop structure is then utilized to sequentially use each subset as the validation set and the remaining subsets as the training set for multiple cross-validations. In each cross-validation, the training set is used to train the PSO-LSTM model, and the validation set to obtain prediction results. After each validation, evaluation metrics (such as RMSE, MAE, ME, etc.) are calculated and recorded. The variation in the prediction model's loss rate (LOSS) during this cross-validation process is illustrated in Figure 11. The LOSS is a performance metric that reflects the model's fit to the training data. Observing the PSO-LSTM model's loss rate reveals that, although initially high, it generally stabilizes after about 20 iterations, indicating the model's gradual convergence to an optimal solution. The fourth cross-validation shows the lowest LOSS, suggesting the best predictive capability in this instance.

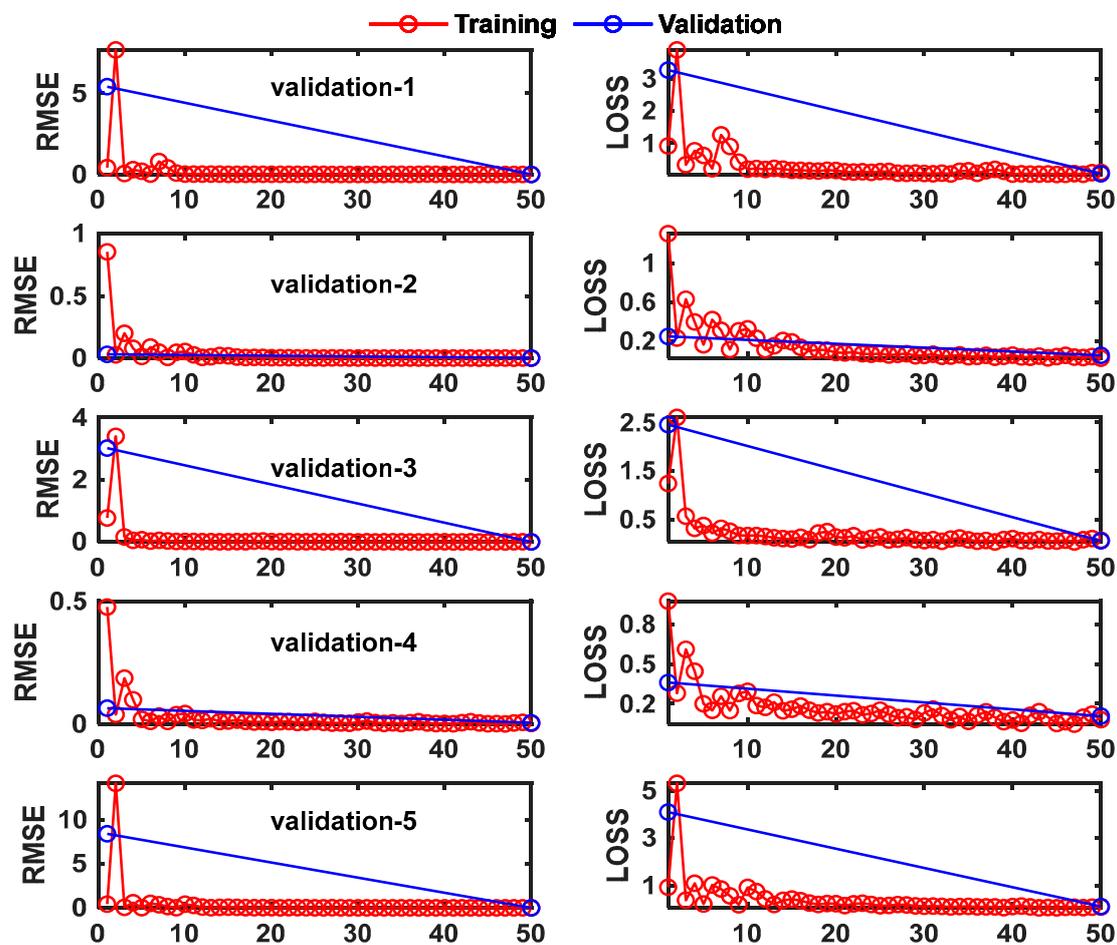


Figure 11. Training set, validation set loss rate obtained from cross-validation.

From the above the LOSS rate loss of the PSO-LSTM prediction model, it can be seen that the training process and performance of the model are good. However, it is important to note that the loss rate is only one indicator of model performance, and other factors should be considered for a comprehensive evaluation. This analysis includes evaluation metrics from the training process such as RMSE, MAE, ME, etc. By synthesizing the prediction results from each cross-validation, the model's generalization ability and stability are assessed. The average evaluation metrics from the cross-validations can be calculated and comparison graphs of prediction results versus observed values can be drawn for visual analysis. The variation in evaluation metrics during the cross-validation process is shown in Figure 12.

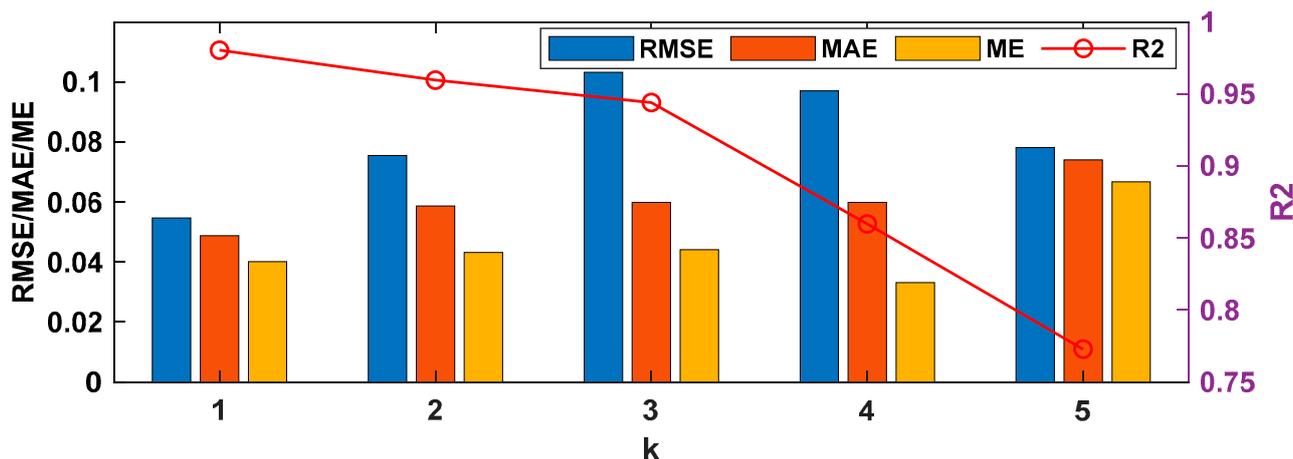


Figure 12. Changes in evaluation indicators during cross-validation.

As can be seen from Figure 12,  $R^2$  is generally higher than 90% when cross-validation is performed for 1–4 times, and in the first cross-validation validation, it reaches 0.9804. With the progress of cross-validation, the low value of 0.7728 appears in the fifth cross-validation, which is consistent with the results in Figure 11, and a high loss rate appears in this cross-validation. From the results of the RMSE, MAE and ME evaluation indicators, the maximum value of RMSE is 0.1032 and the minimum value of 0.0547, while the maximum value of MAE is 0.0740 and the minimum value of MAE is 0.0487. It can be seen that the value of RMSE will be greater than the value of MAE. When the error is small, they may not differ much, but as the error increases, the value of the RMSE will increase more quickly because it is more affected by the large error. Considering these indexes comprehensively, the performance of the prediction model can be evaluated more comprehensively, and the most suitable evaluation indexes can be selected according to the specific needs.

## 6. Case Study and Discussion

In order to verify the feasibility and validity of the model, two typical gas concentration data of a mine in Guizhou province from August 2012 to November 2015 were selected as the research object (outburst accidents occurred during the corresponding period).

### 6.1. Typical Case Analysis

After preprocessing the data, 70% of the data were as selected for training the model, and the remaining 30% were as used as a test set. For the typical gas concentration data of one (Figure 1) and two (Figure 2), a comparative prediction study was conducted using the PSO-LSTM model and the LSTM model. The prediction result for typical gas concentration data one is shown in Figure 13, and the prediction result for case two is shown in Figure 14. From Figures 13 and 14, it can be observed that the prediction curve of the PSO-LSTM model becomes stable more quickly than that of the LSTM model, and its curve is closer to the curve of the actual values.

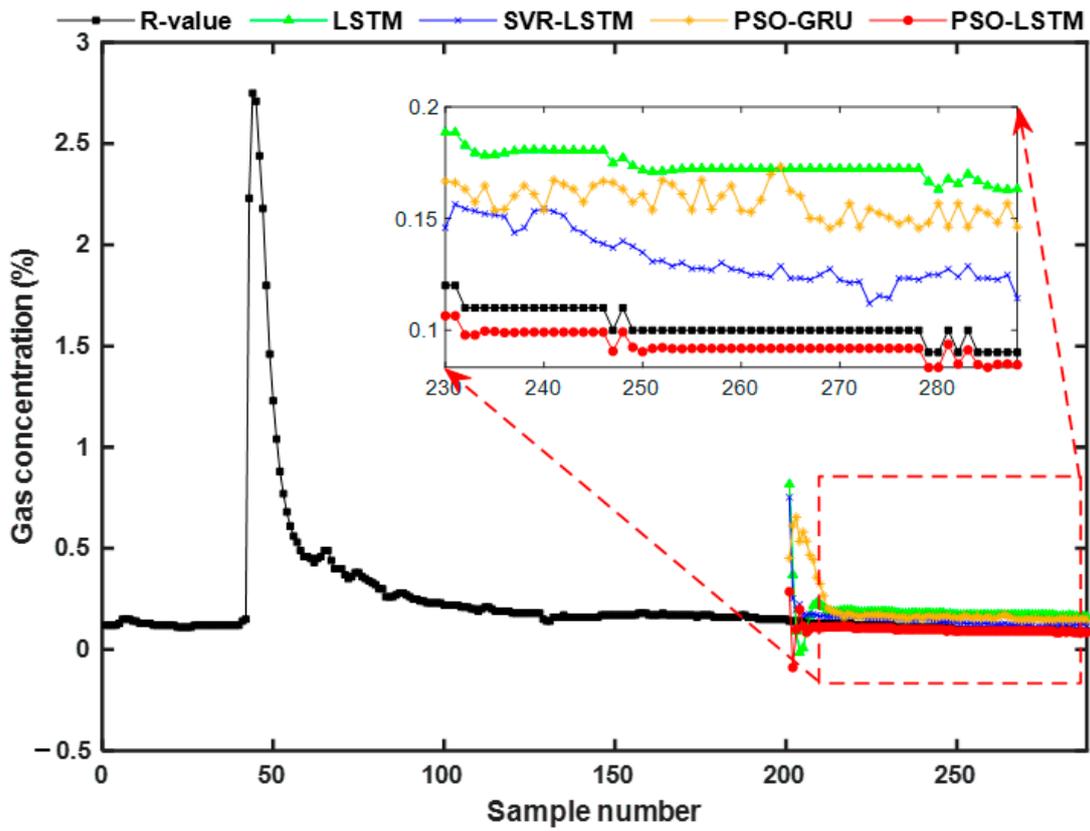


Figure 13. Gas concentration prediction results of Case 1.

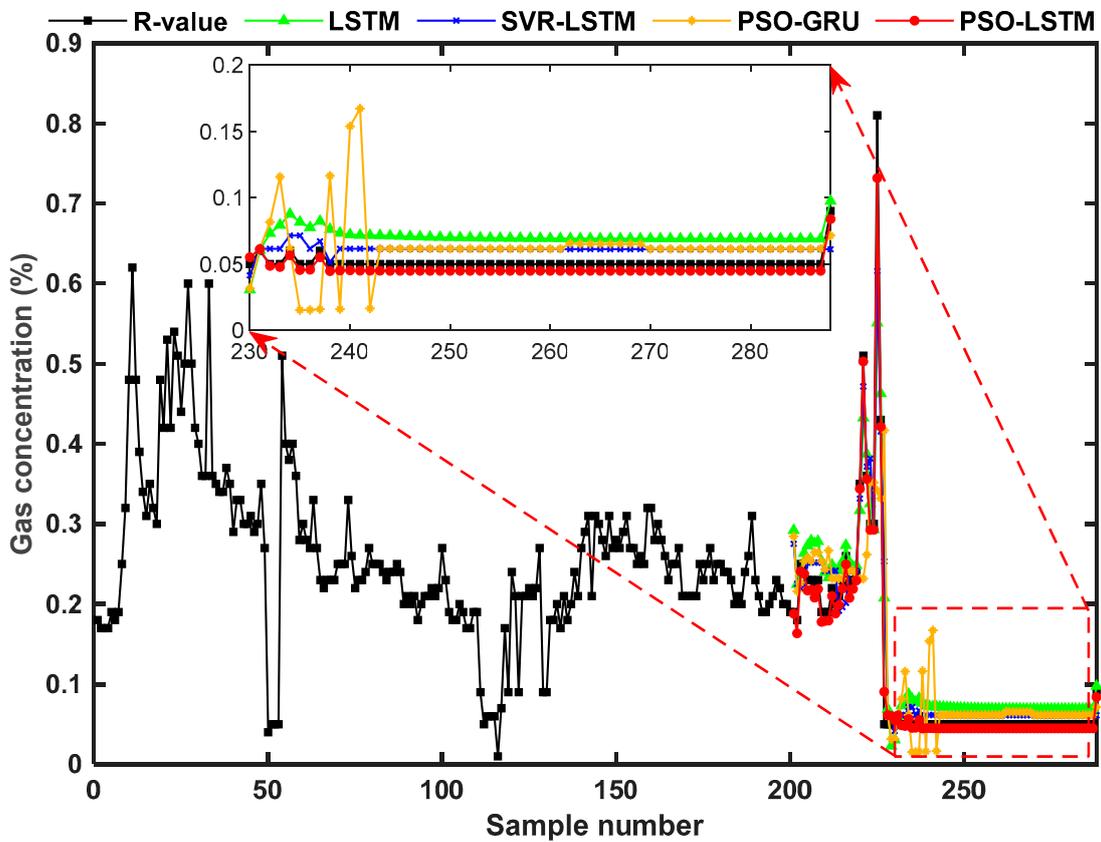


Figure 14. Gas concentration prediction results of Case 2.

In summary, this paper proposes a PSO-LSTM model for predicting the time series data of gas concentration in coal mines. Experimental results demonstrate that the PSO-LSTM model offers better prediction accuracy and performance. It can more effectively capture the trends and patterns in time series data, thereby enhancing the accuracy of predictions. This research provides a new optimization method for time series data prediction, which is expected to play a significant role in practical applications.

## 6.2. Discussion Section

To evaluate the effectiveness and accuracy of the predictions mentioned above, the cross-validation method is employed to divide the dataset into multiple subsets. Each subset is used as the validation set in turn, with the remaining subsets serving as the training set to perform multiple rounds of model training and validation. This approach aims to comprehensively assess the generalization ability and stability of the model. The use of this method facilitates a comprehensive evaluation of the PSO-LSTM prediction model's predictive power and accuracy, thereby providing guidance for model improvements and optimization. Under the same dataset conditions, a quantitative comparative analysis of the predictive efficacy of the combined model with common models (LSTM, SVR-LSTM, PSO-GRU) is conducted. The results, as shown in Table 3, indicate that the PSO-LSTM combined model outperforms the other models, with smaller error values between the predicted and actual values.

**Table 3.** Evaluation and comparison of gas concentration prediction results of different models.

Prediction Model	Case 1			Case 2		
	MAE	RMSE	R <sup>2</sup>	MAE	RMSE	R <sup>2</sup>
LSTM	0.0807	0.1050	0.1945	0.0292	0.0432	0.8605
SVR-LSTM	0.0404	0.0745	0.3473	0.0218	0.0379	0.8906
PSO-GRU	0.0899	0.1366	0.1410	0.0378	0.0790	0.4841
PSO-LSTM	0.0166	0.0327	0.9059	0.0077	0.0118	0.9912

To observe the prediction results more intuitively, the actual values and the predicted values of the four models are visualized for comparison, as shown in Figures 13–15. From these figures, it can be observed that the value of MAE is consistently lower than that of RMSE, and the trend of the R<sup>2</sup> curve aligns with the overall trend of MAE and RMSE. Lower values of RMSE and MAE indicate a higher degree of concordance between predicted and actual values, while higher values indicate greater prediction errors. R<sup>2</sup> evaluates the degree of fit of the curve; the closer it is to 1, the better the fit and thus the better the prediction effect. Specifically, the comparative performance of all models from highest to lowest is: PSO-LSTM > SVR-LSTM > LSTM > PSO-GRU. Among them, the PSO-LSTM combination model exhibits exceptional predictive performance, followed by the SVR-LSTM model. The LSTM model, despite experiencing significant fluctuations, performs better overall than PSO-GRU.

In summary, compared to the other three algorithms, PSO-LSTM surpasses them in prediction accuracy and maintains a more stable prediction rate, aligning with the general trend of time series data (as shown in Figure 15). Especially when facing data with more frequent fluctuations, its predictive efficiency is superior. For instance, in Case 1, despite a significant change, the data overall appears smooth (with smaller anomalies), but its prediction results are less accurate than in the more fluctuating Case 2. This indirectly suggests that the PSO-LSTM prediction model also possesses superior performance for non-stationary data.

Through the comparative analysis mentioned above, compared to individual models, the PSO-LSTM combined model, which integrates the PSO algorithm and the LSTM neural network, has the following advantages: (1) A strong global search capability: The PSO algorithm simulates the movement of particles in the search space, enabling a better

exploration of the problem's global optimal solution. In predicting nonlinear complex issues like gas concentration, PSO-LSTM can better search for potential patterns and trends in the data, enhancing prediction accuracy. (2) A high adaptability: The PSO algorithm is adaptive, allowing it to adjust the search strategy based on the complexity of the problem and the characteristics of the data. Combined with the memory and learning capabilities of the LSTM neural network, PSO-LSTM can flexibly adapt to different data features and prediction requirements, improving the model's generalization ability. (3) Overcoming local optima: The PSO algorithm, through the collaborative search of multiple particles, helps to overcome the problem of local optima that neural network models like LSTM are prone to. This can more effectively uncover potential laws and trends in the data, increasing the accuracy and stability of predictions. (4) Applicability to large-scale data: PSO-LSTM is suitable for handling large-scale data and can effectively process and learn from large amounts of data in nonlinear complex problems like gas concentration, making PSO-LSTM highly practical and scalable in real-world engineering applications.

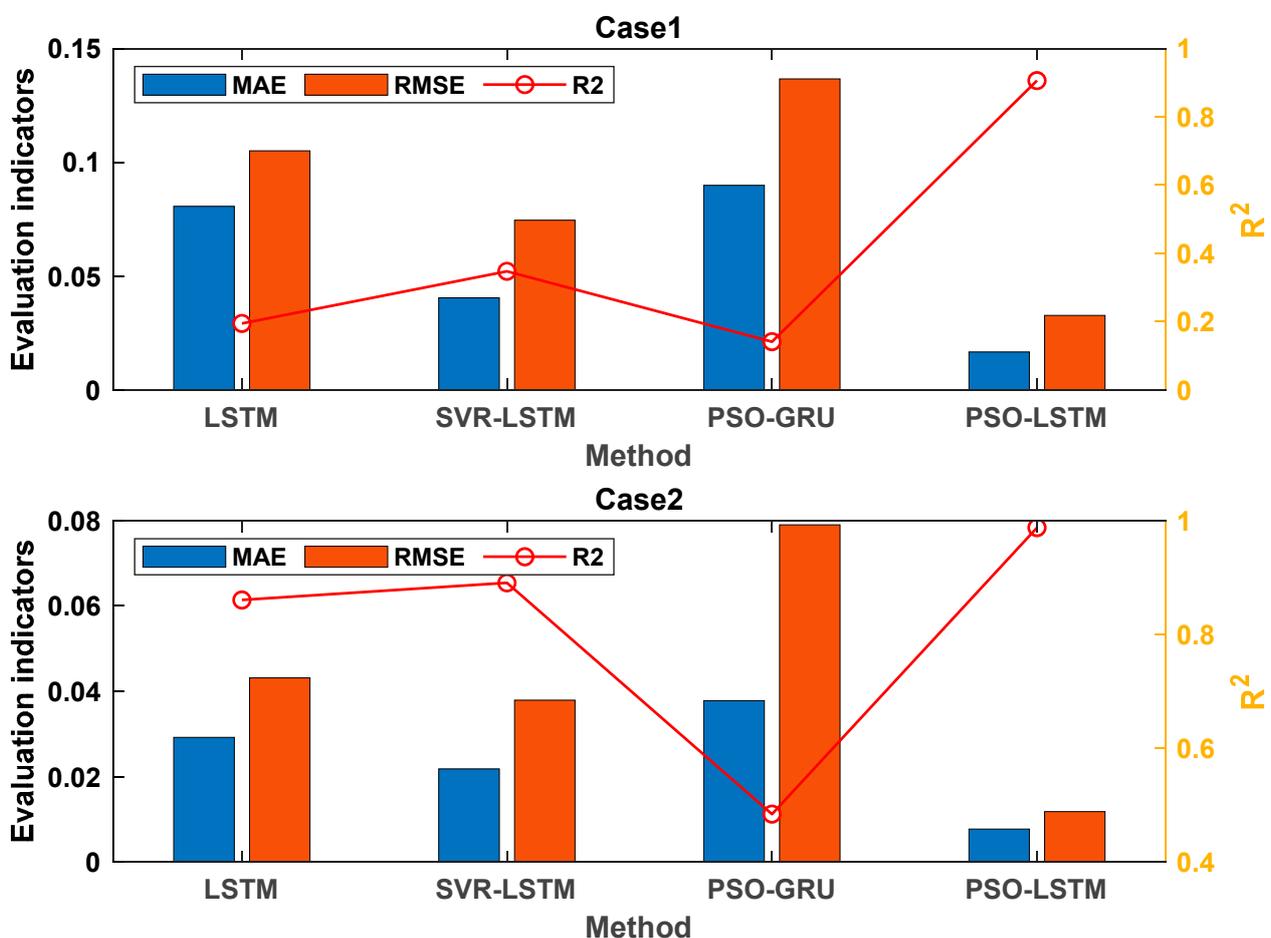


Figure 15. Comparison of the prediction performance of several typical algorithms.

## 7. Conclusions

This paper proposes a gas concentration prediction method based on the PSO-LSTM combined model, tailored to the complex nature of coal mine gas concentration. This method leverages the adaptive PSO to automatically search for the initial input weights, recursive weights, and biases of the LSTM network, as well as the weights and biases of the FCL, for predicting gas concentration time series. Research was conducted on the principles of the prediction model, model implementation steps, data preprocessing, and prediction performance evaluation, and a comparative analysis of typical gas concentration data was carried out alongside conventional methods.

- (1) Based on the characteristic patterns of typical gas concentration time series data, a combined model prediction method utilizing the PSO model and LSTM network is proposed. This model capitalizes on the high accuracy and convergence properties of the PSO algorithm to optimize key parameters of the LSTM model, determining the weight coefficients within the combined model, thereby enhancing the prediction accuracy of the LSTM model. The introduction of this method further improves the accuracy of coal mine gas concentration predictions, offering significant reference value for gas disaster monitoring, early warning, and prevention.
- (2) The prediction accuracy of the PSO-LSTM model was evaluated using 24 h gas concentration data from a mine as a sample. The results show that the PSO-LSTM model performs even better, and this method successfully combines the global search potential of adaptive PSO with the local search capabilities of the Adam optimizer, enhancing model accuracy, reducing the likelihood of falling into local minima, and overcoming underfitting/overfitting issues.
- (3) To verify the superiority of the PSO-LSTM model, two sets of field-measured gas concentration data were used as sample data and compared with the predictive effects of conventional models such as LSTM, SVR-LSTM, and PSO-GRU. The results indicate that the predictive performance of all models ranks as follows: PSO-LSTM > SVR-LSTM > LSTM > PSO-GRU. Compared to the other three algorithms, PSO-LSTM exhibits higher predictive accuracy and maintains a more stable prediction accuracy rate, aligning with the general trends of time series data.

**Author Contributions:** Conceptualization, G.Y. and Q.Z.; data curation, D.W. and Y.F.; formal analysis, D.W., Y.F. and Q.L.; funding acquisition, Q.Z. and X.C.; investigation, Y.F. and X.C.; methodology, G.Y., Q.Z., D.W. and Q.L.; project administration, G.Y.; resources, Q.L.; software, Q.Z. and D.W.; supervision, Q.Z. and Q.L.; validation, G.Y., Y.F. and X.C.; visualization, X.C.; writing—original draft, G.Y. and D.W.; writing—review and editing, Q.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Hebei Natural Science Foundation, grant number E2023508021, Central Guidance on Local Science and Technology Development Fund of Hebei Province, grant number 216Z5401G, the Fundamental Research Funds for the Central Universities, grant number 3142021002, S&T Program of Hebei, grant number 22375401D, and General Program of National Natural Science Foundation of China, grant number 52074121. The APC was funded by Sun Yat-Sen University.

**Data Availability Statement:** The data used to support the findings of this study were supplied by Qianjie Zhu under license and so cannot be made freely available. Requests for access to these data should be made to Qian-jie Zhu (zhqj2016@ncist.edu.cn).

**Conflicts of Interest:** Author Guangyu Yang was employed by the company Coal Mining Research Institute Co., Ltd. of CCTEG. Author Qingsong Li was employed by the company Guizhou Mine Safety Scientific Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The [companies in affiliation and Funding] had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Lu, B.; Hardin, J. A Unified Framework for Random Forest Prediction Error Estimation. *J. Mach. Learn. Res.* **2021**, *22*, 386–426.
2. Wang, H.; Wang, G. Improving random forest algorithm by Lasso method. *J. Stat. Comput. Simul.* **2021**, *91*, 353–367. [[CrossRef](#)]
3. Zhang, H.; Zimmerman, J.; Nettleton, D.; Nordman, D.J. Random forest prediction intervals. *Am. Stat.* **2019**, *74*, 392–406. [[CrossRef](#)]
4. Czajkowski, M.; Czerwonka, M.; Kretowski, M. Cost-sensitive global model trees applied to loan charge-off forecasting. *Decis. Support Syst.* **2015**, *74*, 57–66. [[CrossRef](#)]
5. Blanquero, R.; Carrizosa, E.; Molero-Río, C.; Morales, D.R. On optimal regression trees to detect critical intervals for multivariate functional data. *Comput. Oper. Res.* **2023**, *152*, 106152. [[CrossRef](#)]
6. Zhao, Y.; Kang, N. Prediction of Coal Face Gas Emission Based on Multiple Linear Regression. *Coal Technol.* **2021**, *40*, 118–121.

7. Li, N.; Yang, M.; Hu, X.; Zhu, Z. Research on Influencing Factors of Coal Mine Gas Emissions Based on Multiple Stepwise Linear Regression. *Coal Technol.* **2022**, *41*, 131–134.
8. Tutak, M.; Brodny, J. Predicting methane concentration in longwall regions using artificial neural net-works. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1406. [[CrossRef](#)] [[PubMed](#)]
9. Bukhari, S.M.S.; Moosavi, S.K.R.; Zafar, M.H.; Mansoor, M.; Mohyuddin, H.; Ullah, S.S.; Alroobaea, R.; Sanfilippo, F. Federated transfer learning with orchard-optimized Conv-SGRU: A novel approach to secure and accurate photovoltaic power forecasting. *Renew. Energy Focus* **2024**, *48*, 100520. [[CrossRef](#)]
10. Khan, U.A.; Khan, N.M.; Zafar, M.H. Resource efficient PV power forecasting: Transductive transfer learning based hybrid deep learning model for smart grid in Industry 5.0. *Energy Convers. Manag. X* **2023**, *20*, 100486. [[CrossRef](#)]
11. Olatomiwa, L.; Mekhilef, S.; Shamshirband, S.; Mohammadi, K.; Petković, D.; Sudheer, C. A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Sol. Energy* **2015**, *115*, 632–644. [[CrossRef](#)]
12. Cao, S.G.; Liu, Y.B.; Wang, Y.P. A forecasting and forewarning model for methane hazard in working face of coal mine based on LS-SVM. *J. China Univ. Min. Technol.* **2008**, *18*, 172–176. [[CrossRef](#)]
13. Zhou, J.; Yang, P.; Peng, P.; Khandelwal, M.; Qiu, Y. Performance Evaluation of Rockburst Prediction Based on PSO-SVM, HHO-SVM, and MFO-SVM Hybrid Models. *Min. Metall. Explor.* **2023**, *40*, 617–635. [[CrossRef](#)]
14. Liu, L.; Yang, C. LSTM gas concentration prediction model based on multiple factors. *J. Saf. Sci. Technol.* **2022**, *18*, 108–113.
15. Li, H.; Jia, J.; Yang, X.; Song, C. Gas concentration prediction model for fully mechanize coal mining face. *Ind. Mine Autom.* **2018**, *44*, 48–53.
16. Li, S.; Ma, L.; Pan, S.; Shi, X. Research on prediction model of gas concentration based on RNN in coal mining face. *Coal Sci. Technol.* **2020**, *48*, 33–38.
17. Zhang, Z.; Zhu, Q.; Li, Q. Prediction of Mine Gas Concentration in Heading Face Based on Keras Long Short Time Memory network. *Saf. Environ. Eng.* **2021**, *28*, 61–67+78.
18. Fu, H.; Liu, Y.; Xv, N.; Zhang, J. Research on Gas Concentration Prediction Based on Multi-Sensor-Deep Long Short-Term Memory Network Fusion. *Chin. J. Sens. Actuators* **2021**, *34*, 784–790.
19. Liu, C.; Zhang, A.; Xue, J.; Lei, C.; Zeng, X. LSTM-Pearson Gas Concentration Prediction Model Feature Selection and Its Application. *Coal Sci. Technol.* **2023**, *16*, 2318. [[CrossRef](#)]
20. Brodny, J.; Felka, D.; Tutak, M. The use of the neuro-fuzzy model to predict the methane hazard during the underground coal mining production process. *J. Clean. Prod.* **2022**, *368*, 133258. [[CrossRef](#)]
21. Liang, Y.; Li, X.; Li, Q.; Mao, S.; Zheng, M.; Li, J. Research on intelligent prediction of gas concentration in working face based on CS-LSTM. *Min. Saf. Environ. Prot.* **2022**, *49*, 80–86.
22. Wen, H.; Yan, L.; Jin, Y.; Wang, Z.; Guo, J.; Deng, J. Coalbed methane concentration prediction and early-warning in fully mechanized mining face based on deep learning. *Energy* **2023**, *264*, 126208. [[CrossRef](#)]
23. Al-Hajj, R.; Assi, A.; Fouad, M.; Mabrouk, E. A hybrid LSTM-based genetic programming approach for short-term prediction of global solar radiation using weather data. *Processes* **2021**, *9*, 1187. [[CrossRef](#)]
24. Tang, G.; Sheng, J.; Wang, D.; Men, S. Continuous Estimation of Human Upper Limb Joint Angles by Using PSO-LSTM Model. *IEEE Access* **2021**, *9*, 17986–17997. [[CrossRef](#)]
25. Cao, X.H.; Stojkovic, I.; Obradovic, Z. A robust data scaling algorithm to improve classification accuracies in biomedical data. *BMC Bioinform.* **2016**, *17*, 359. [[CrossRef](#)] [[PubMed](#)]
26. Jiang, Z.; Zhang, S.; Li, W. Gas Concentration Prediction Model Based on Improved Neural Network. *IEEE Access* **2019**, *7*, 100364–100370.
27. Ruma, J.F.; Adnan, M.S.G.; Dewan, A.; Rahman, R.M. Particle swarm optimization based LSTM networks for water level forecasting: A case study on Bangladesh river network. *Results Eng.* **2023**, *17*, 100951. [[CrossRef](#)]
28. Kumar, G.; Singh, U.P.; Jain, S. An adaptive particle swarm optimization-based hybrid long short-term memory model for stock price time series forecasting. *Soft Comput.* **2022**, *26*, 12115–12135. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.