Research Article

Wind Speed Forecast Based on the LSTM Neural Network Optimized by the Firework Algorithm

Bilin Shao [iD],[1] **Dan Song** [iD],[1] **Genqing Bian** [iD],[2] **and Yu Zhao** [iD][1]

[1]*School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, China*
[2]*School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China*

Correspondence should be addressed to Dan Song; 674355101@qq.com

Wind energy is a renewable energy source with great development potential, and a reliable and accurate prediction of wind speed is the basis for the effective utilization of wind energy. Aiming at hyperparameter optimization in a combined forecasting method, a wind speed prediction model based on the long short-term memory (LSTM) neural network optimized by the firework algorithm (FWA) is proposed. Focusing on the real-time sudden change and dependence of wind speed data, a wind speed prediction model based on LSTM is established, and FWA is used to optimize the hyperparameters of the model so that the model can set parameters adaptively. Then, the optimized model is compared with the wind speed prediction based on other deep neural architectures and regression models in experiments, and the results show that the wind speed model based on FWA-improved LSTM reduces the prediction error when compared with other wind speed prediction-based regression methods and obtains higher prediction accuracy than other deep neural architectures.

1. Introduction

As a green renewable energy source, wind power has an immeasurable commercial development prospect, and the research on related forecasting technologies is also more important. However, the randomness, volatility, and intermittency of wind resources have brought great troubles and challenges to the stable operation of the power system. Traditional wind power forecasting technologies are no longer sufficient to solve the above problems. For this reason, it is urgent to introduce cutting-edge artificial intelligence technology. Artificial intelligence is a branch of computer science dedicated to the research and development of theories, methods, technologies, and application systems for simulating, extending, and expanding human intelligence. In recent years, the rapid development of artificial intelligence-related machine learning, deep learning, and other technologies has provided new ideas for the research and implementation of high-precision wind power prediction technology and brought new development opportunities.

Wind power prediction relies on wind speed estimation. Due to the cyclical, daily pattern, and high stochastic variability, accurate prediction of wind power is too complicated. Therefore, it is clear that efficient transformation and application of the wind energy resources require exact and complete information on the wind features of the region, and local and regional climates, topography, and impediments include buildings; all affect wind energy. In the last decades, scholars have proposed different prediction methods based on the time series of historical wind speed and in general, these models can be divided generally into four types: physical, statistical, intelligence learning model, and hybrid model.

Physical approaches, which are based on a detailed physical description of the atmosphere, used meteorological data such as air temperature, topography, and pressure to predict wind speed, thus leading to intricate calculations and high costs [1]. Statistical methods, such as Autoregressive Integrated Moving Average (ARIMA) model, Seasonal ARIMA (SARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, and Monte

Carlo Simulation [2], predict wind speed on the premise of linear assumption and are more accurate than physical methods [3, 4]. However, the variation of wind speed contains significant nonlinear and chaotic characteristics, and it is usually difficult to accurately and effectively predict the future wind speed simply by applying these methods or models. In addition, statistical methods require a large amount of data for learning and modeling and are more suitable for ultrashort-term wind power prediction. Intelligent learning methods, such as Support Vector Regressor (SVR), Decision Tree Regressor (DTR), Multivariate Linear Regression (MLR), Artificial Neural Network (ANN), train and predict the wind speed data with better performance in the fitting of the nonlinear changes of wind speed [5–7]. SVR, MLR, and DTR have advantages in sparsity and generalization and solving nonlinearity prediction problems, but its key parameters mainly rely on manual selection [8–11]. ANN [12, 13] has the advantages of good nonlinear fitting and strong self-learning ability, but it is unstable, slow convergence rate is easy to fall into the local optimal value, and it is difficult to obtain its network structure including the number of hidden layers. The wind speed prediction model based on Convolutional Neural Networks (CNN) can consider the temporal and spatial correlation of wind speed, to make the ultrashort-term prediction of the spatial distribution of wind speed [14–16].

Wind speed is affected by many factors, and a single prediction model cannot fully include all these factors. Particularly in extreme weather cases, a single model does not have sufficient learning, which may lead to a large deviation in the prediction. The combined prediction method takes into account the respective advantages of different models at the same time, optimally combining a variety of single models and giving play to the advantages of each model can significantly improve the accuracy of prediction [17–20]. Combination prediction methods mainly include weighted combination prediction and fusion combination prediction [21]. The key lies in the determination of the weight coefficient. The combination method of the fixed weight coefficient [22] is simple and easy to realize. The combination method of the variable weight coefficient [17] is strong adaptability and high accuracy. The fusion combination is optimized by other prediction methods in different prediction stages, including input data stabilization, model parameter optimization, and output error correction. Based on empirical mode decomposition (EMD) [22–25], variational mode decomposition (VMD) [26–29], analytical mode decomposition (AMD) [30, 31], the wavelet decomposition [14, 25, 32], and so on, the wind speed sequence data was preprocessed to make the data stable. Better prediction results are achieved. In addition, the Hilbert–Huang transform (HHT) [33], fast correlation filter [34], principal component analysis (PCA) [35], and so on extracted the input features of wind speed data and obtained good prediction results by optimizing the short-term wind speed prediction model combined with other prediction methods. It is an important way to optimize the parameters of the model by using an intelligent algorithm. According to the characteristics of the wind speed data, the intelligent

algorithm is used to determine the parameters adaptively during the training process to improve the learning ability and generalization ability of the model. Genetic algorithm [36], particle swarm optimization algorithm [27], and cuckoo algorithm [37] are used to optimize the hybrid model combining the parameters and threshold values of BPNN, LSTM, SVM, and other intelligent learning models, which can overcome the problem of low prediction accuracy of a single model and improve the accuracy of wind speed prediction. The prediction results of the traditional method are substituted into the error model to correspond to the superposition and correct the error, which has strong universality and is not limited to the specific prediction process [38–41].

In this paper, based on the measured data of a wind turbine in a power plant and the analysis of wind power time series, the combined prediction method is proposed. Firstly, a wind speed prediction model based on LSTM is established. Then, from the perspective of model hyperparameter optimization, the fireworks algorithm (FWA) is used to automatically search for the best hyperparameter combination suitable for wind speed data. Finally, the optimized FWA-LSTM is used to predict and analyse the wind speed data, and its feasibility and effectiveness are verified.

This paper is organized as follows: in Section 2, we constructed a wind speed prediction model based on LSTM; in Section 3, we studied the firework algorithm, hyperparameters optimization of LSTM by the firework algorithm, and optimized LSTM wind speed prediction algorithm based on firework algorithm. In addition, experimental environment configuration and parameter settings, wind speed prediction results based on the proposed method, and the comparison are discussed; and finally, the main conclusions are drawn in Section 5.

2. Wind Speed Prediction Model Based on LSTM

2.1. LSTM Neural Network Model. The traditional neural network model will lose the remote information, and it is difficult to learn the long-distance dependent information. LSTM is an improvement of the recurrent neural network, which aims to overcome the defects of the recurrent neural network in processing long-term memory. The LSTM introduced the concept of cellular states, which determine which states should be preserved and which should be forgotten. The basic principle of LSTM is shown in Figure 1.

As shown in Figure 1, X_t is the input at time t, h_{t-1} is the output of the hidden layer at time $t-1$, and C_{t-1} is the output of the historical information at time $t-1$; f, i, and o are, respectively, the forgetting gate, input gate, and output gate at time t, and e is the internal hidden state, namely, the transformed new information. LSTM conducts parameter learning for them in the training. C_t is the updated historical information at time t, and h_t is the output of the hidden layer at time t.

Firstly, the input x_t at time t and the output h_{t-1} of the hidden layer are copied into four copies, and different weights are randomly initialized for them, so as to calculate the forgetting gate f, input gate i, and output gate o, as well as the internal hidden state e. Their calculation methods are
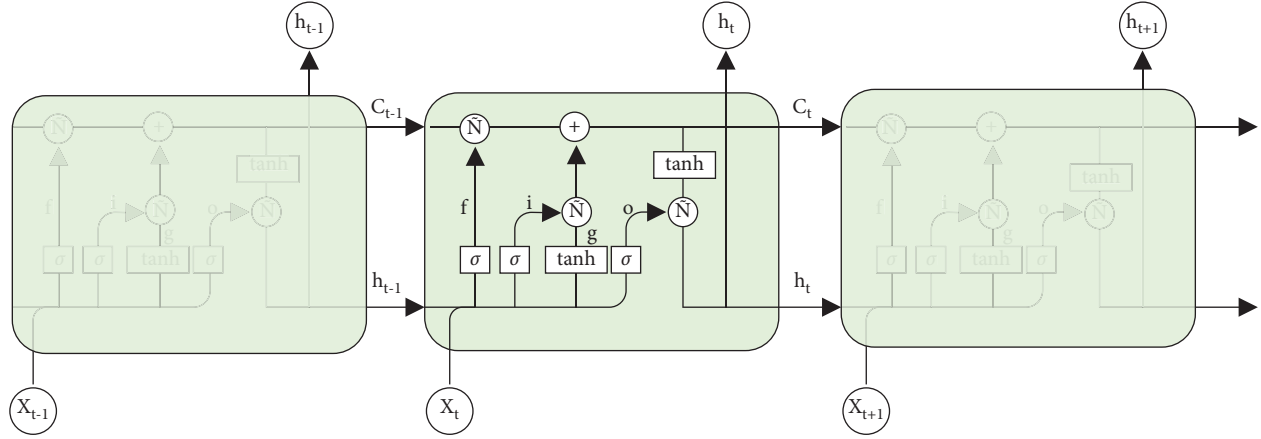
FIGURE 1: Basic principle of LSTM.

shown in formulas (1)–(4), where W is the parameter matrix from the input layer to the hidden layer, U is the self-recurrent parameter matrix from the hidden layer to the hidden layer, r is the bias parameter matrix, and σ is the sigmoid function so that the output of the three gates remains between 0 and 1:

$$f = \sigma\left(W_f x_t + U_f h_{t-1} + r_f\right), \qquad (1)$$

$$i = \sigma\left(W_i x_t + U_i h_{t-1} + r_i\right), \qquad (2)$$

$$o = \sigma\left(W_o x_t + U_o h_{t-1} + r_o\right), \qquad (3)$$

$$g = \sigma\left(W_g x_t + U_g h_{t-1} + r_g\right). \qquad (4)$$

Secondly, forgetting gate f and input gate i are used to control how much historical information C_{t-1} is forgotten and how much new information e is saved, to update the internal memory cell state C_t. The calculation method is as follows:

$$C_t = f_t \otimes C_{t-1} \oplus i \otimes e. \qquad (5)$$

Finally, output gate o is used to control how much C_t information of the internal memory unit is output to the implicit state h_t, and its calculation method is shown as follows:

$$h_t = o \otimes \tanh\left(C_t\right). \qquad (6)$$

2.2. Wind Speed Prediction Model Based on LSTM. The process of using LSTM to predict wind speed data is shown in Figure 2. It mainly includes wind speed data preparation and preprocessing (data resampling and null filling), data normalization, data division, prediction model establishment and evaluation, and data prediction.

First, the wind speed data is modeled as a nonnegative matrix X of an $N \times T$, where N represents the number of wind speed monitoring points, T represents the number of time slots sampled, and each column in the wind speed data matrix represents the wind speed value at different points in a specific time interval.

Wind speed prediction can obtain the predicted value of the future time through the historical time series, $X(i, j)$ represents the scale of $N \times T$ flow matrix, $x_{n,t}$ represents the wind speed value of row n and column t. Wind speed prediction is defined by a series of historical wind speed data $(x_{n,t-1}, x_{n,t-2}, x_{n,t-3}, \ldots, x_{n,t-1})$ to predict the wind speed at time t in the future. In the wind speed prediction model based on LSTM (Figure 2), it is assumed that the wind speed at a certain point in the t-slot is predicted, the input of the model is $(x_{n,t-1}, x_{n,t-2}, x_{n,t-3}, \ldots, x_{n,t-1})$, and the output is the predicted value $\hat{x}_t$ of the wind speed at the t-slot at this point.

(1) Wind speed data preparation and preprocessing: to meet the time-frequency (seconds, minutes, hours, days, etc.) requirements of wind speed data prediction, it is necessary to resample the original data, that is, to convert the time series from one frequency to another through downsampling or upsampling. In addition, if there are null values in the resampled data sequence, the null values need to be filled. Here, we use the machine learning method—the K-Nearest Neighbours (KNN)—to fill with null values of wind speed data.

(2) Data normalization: the range standardization method is used to process the wind speed data so that the sample data value is between 0 and 1. The calculation method of the range standardization method is shown as follows:

$$X_{nor} = \frac{X - X_{min}}{X_{max} - X_{min}}. \qquad (7)$$

In formula (7), X_{max} represents the maximum value of wind speed data and X_{min} represents the minimum value of wind speed data.

(3) Data division: the wind speed data after preprocessing and normalization is divided into a training set and a test set according to a simple cross-validation method. While keeping the wind speed data sequence unchanged, fivefold cross-validation is used to divide into the training set and the test set, which are used for the training and prediction of the LSTM wind speed prediction model, respectively.
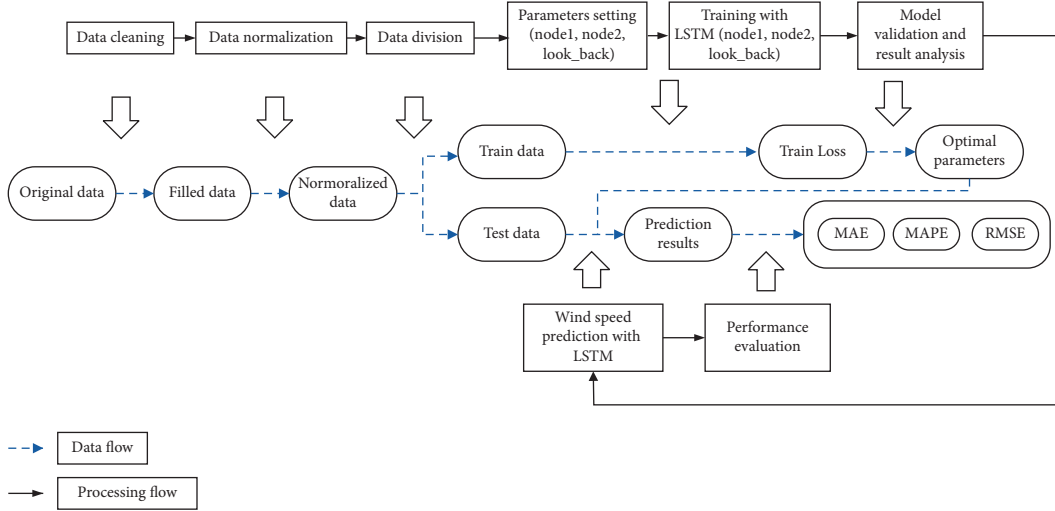
FIGURE 2: General schematic diagram for LSTM-based wind speed prediction.

(4) Construct an LSTM wind speed prediction model: define an LSTM neural network and set the parameters, including time step, network layer number, number of neurons in each layer, dropout, activation function, return value type and number, hidden layer dimension size, learning rate, batch size, and values for the number of iterations.

(5) Compile the network: set the optimizer, error measurement indicators, and training record parameters and compile the constructed LSTM wind speed prediction model.

(6) Evaluate the network: the training set data is substituted into the model for training, the error of the established prediction model is evaluated, and the parameter settings of the model are fine-tuned according to the result to obtain a better prediction effect.

(7) Forecast and evaluation: use the optimized wind speed prediction model to make predictions, compare the prediction results with the real data, and calculate the error.

3. The LSTM Wind Speed Prediction Model Optimized by the Firework Algorithm

3.1. The Firework Algorithm. The fireworks algorithm (FWA) [42–44] is a simple-rule, fast-convergence-speed swarm intelligence optimization algorithm. It searches the solution space mainly by the sparks generated by the firework explosions, and the fireworks and the sparks from the explosion formed the whole crowd. In this algorithm, the firework is seen as a feasible solution in the solution space of the optimization problem, and the process of firework explosion to generate sparks is the way of searching the neighbourhood. FWA includes the following steps: initialization, calculating the fitness, generating sparks by firework explosions, and calculating the optimal solution.

Firstly, FWA sets a series of initial parameter values including the number of fireworks population N, the explosion range control parameter $\widehat{A}$, the maximum number of sparks m, the number of variant sparks $\widehat{m}$, the parameters a and b that limit the number of sparks produced by the explosion, The minimum normal value ε of zero, and the solution space boundaries B_u and B_i, where B_u is the upper boundary and B_i is the lower boundary. The firework algorithm mainly uses random initialization to generate N initial fireworks in the solution space.

Secondly, calculate the fitness value of each firework, and generate sparks based on the fitness value. The calculation for generating the number of sparks in FWA is shown as follows:

$$S_i = m \frac{Y_{\max} - f(x_i) + \varepsilon}{\sum_{i=1}^{N}(Y_{\max} - f(x_i)) + \varepsilon}. \tag{8}$$

In formula (8), S_i is the number of sparks produced by the ith firework, m is a constant which limits the total number of sparks produced, $Y_{\max}$ is the objective function value of the firework with the worst fitness of the current population, $f(x_i)$ is the fitness function of the firework x_i, and ε is the minimum number of the machine.

The calculation of FWA explosion amplitude is shown as follows:

$$A_i = \widehat{A} \frac{f(x_i) - Y_{\min} + \varepsilon}{\sum_{i=1}^{N}(f(x_i) - Y_{\min}) + \varepsilon}. \tag{9}$$

In formula (9), A_i is the explosion amplitude of the ith firework, that is, the explosion radius, $\widehat{A}$ is a constant which represents the maximum explosion amplitude, and $Y_{\min}$ is the fitness value of the firework with the best current population fitness value.

Thirdly, according to the actual firework attributes and the actual situation of the search problem, sparks are generated in the radiation space of the firework. To ensure the diversity of the population, the fireworks need to be appropriately mutated, such as Gaussian mutation.

The calculation of the Gaussian mutation algorithm in FWA is shown as follows:

$$x_i^k = x_i^k \times g. \tag{10}$$

In formula (10), x_i^k is the position of the ith individual on the kth dimension and g is the value of the Gaussian distribution function where $g \sim (1, 1)$.

Finally, calculate the optimal solution of the population and decide whether the termination condition is met. If it satisfies the requirements, stop the search; else, continue iterating.

In the entire population, the spark with the best fitness value is selected and retained as the next-generation fireworks, and the remaining sparks are selected by roulette. The probability of each spark being selected is calculated as follows:

$$P(x_i) = \frac{R(x_i)}{\sum_{j \in k} R(x_i)}. \tag{11}$$

In formula (11), $P(x_i)$ is the probability of the ith spark and $R(x_i)$ is the sum of the distance between the x_i and the candidate fireworks except for x_i.

Compared with particle swarm optimization (PSO) and genetic algorithm (GA), the fireworks algorithm has higher convergence and solving accuracy and has been applied to solve many practical optimization problems, of which parameter optimization is an important aspect [45–47].

3.2. Hyperparameter Optimization of LSTM by the Firework Algorithm.

The hyperparameter selection of the LSTM model has an important influence on the prediction accuracy of the model. The existing hyperparameter selection generally adopts the empirical method. The empirical method is arbitrary and blind in the choice of parameters without universality. Therefore, combining multiple hyperparameters into a multidimensional solution space and traversing the solution space to obtain the optimal parameter combination can reduce the randomness and blindness of parameter selection. The selection of multiple hyperparameters is often carried out in a larger solution space, and a better performance optimization algorithm is needed to quickly obtain the global optimal solution. Therefore, the firework algorithm with global optimization and fast convergence speed is adopted to optimize the LSTM model's hyperparameters to improve the scientificity of model parameter selection and thus improve the prediction accuracy of the model.

Suppose that n hyperparameters of the LSTM wind speed prediction model need to be optimized, and each firework represents a set of hyperparameters in the solution space. Assuming that there are q sets of hyperparameter combinations in the n-dimensional continuous search space, for the ith hyperparameters $i(i = 1, 2, \ldots, q)$ in the spark, the n-dimensional current position vector $x^i(k) = [x_1^i \ x_2^i \ \ldots x_n^i]^T$ represents the current value of the ith group of hyperparameters in the solution space, and the n-dimensional velocity vector $v^i(k) = [v_1^i \ v_2^i \ \& \ v_n^i]^T$ represents the search direction of the group of hyperparameters.

The goal of wind speed prediction is to make the predicted value close to the actual value, that is, the error between the predicted value and the actual value is as small as possible, so the Root Mean Square Error (RMSE) of the training data in the wind speed prediction model is selected as the objective function. Let fitness = RMSE; then, the objective function is to minimize RMSE. The calculation method of RMSE is as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}. \tag{12}$$

In formula (12), $\hat{y}_i$ is the predicted value, $\hat{y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_i\}$, y is the true value, $y = \{y_1, y_2, \ldots, y_i\}$.

According to the firework algorithm, two important hyperparameters of the LSTM wind speed prediction model are optimized: the time step and the number of neurons in each layer. Two LSTM models, single-layer and double-layer LSTM, are used as the research objects to optimize the hyperparameters. Use node to represent the number of neurons and look_back to represent the time step. For a single-layer LSTM model, fitness = RMSE (node, look_back); for a two-layer LSTM model, fitness = RMSE (node 1, node 2, look_back).

According to the FWA process (as shown in Figure 3), the process of the hyperparameters optimization of the LSTM wind speed prediction model mainly includes six steps:

Step 1: initialize the parameters of FWA: set the initial firework population size, namely, the number of hyperparameter combinations N, the explosion range control parameter $\widehat{A}$, the maximum number of sparks m, and the number of variant sparks $\widehat{m}$ and limit the number of sparks produced by the explosion parameters a and b, the minimum normal value ε that tends to zero, and the solution space boundaries B_u and B_i, where B_u is the upper boundary and B_i is the lower boundary. Using random initialization, N initial fireworks are generated in the solution space. Set the maximum number of iterations $item\_max$ and the preset error Pre_error.

Step 2: calculate the fitness of each firework; that is, calculate the fitness value of the objective function of each group of hyperparameters. According to the fitness value, the explosion operator, the number of sparks, the explosion amplitude, and the offset value are calculated. Each firework explosion generates sparks of the hyperparametric group, and the sparks beyond the boundary are mapped according to the rules. At the same time, a certain number of Gaussian variation sparks of the hyperparametric group are generated by using Gaussian variation.

Step 3: set the optimal objective function value F_i of each group of hyperparameters. For the ith group of hyperparameters, compare its current objective function value current_fitness with F_i. If it is less than F_i, use current_fitness as the best objective function value F_i of the ith group of hyperparameters; that is, let F_i = current_fitness.
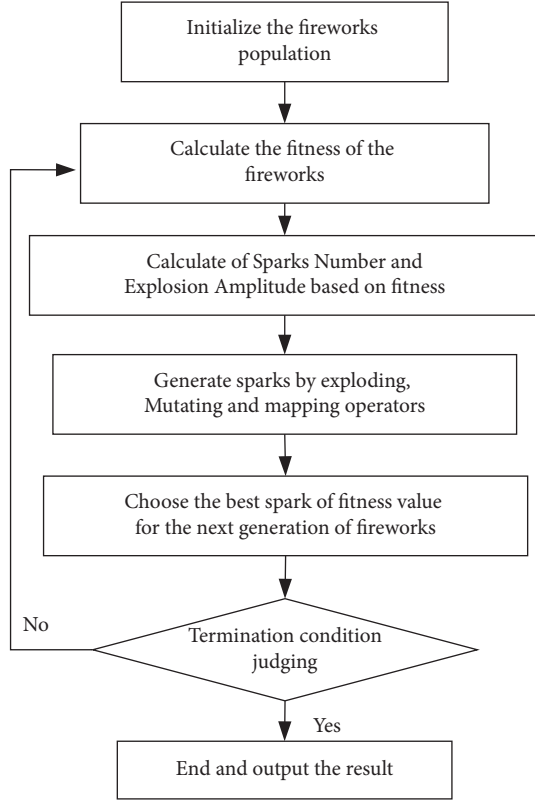
Figure 3: Process of FWA.

Step 4: set the global optimal value F_g. For the ith group of hyperparameters, compare F_i with F_g. If it is less than F_g, use F_i as the optimal value F_g of the current population; that is, let $F_g = F_i$.

Step 5: update the explosion range and spark number of each group of hyperparameters according to formulas (8) and (9).

Step 6: check the termination conditions. If the set conditions (preset error or maximum number of iterations) are not reached, return to Step 2 to continue execution.

3.3. Optimized LSTM Wind Speed Prediction Algorithm Based on the Firework Algorithm.
According to the wind speed prediction steps based on LSTM and the process of the FWA hyperparameter optimization, the call relationship between them can be obtained as in Figure 4.

It is obtained that the wind speed prediction algorithm based on LSTM optimized by FWA—the FWA-LSTM wind speed prediction algorithm—is derived. The pseudocode of the algorithm is shown in Algorithm 1.

Algorithm 1 firstly preprocesses the wind speed data, normalizes and divides the data to obtain a training set and a test set, then establishes the LSTM wind speed prediction model, and uses FWA to optimize the LSTM hyperparameters to obtain the optimal parameter combination; finally, the parameters are substituted into the model to complete the prediction and error calculation of wind speed data.
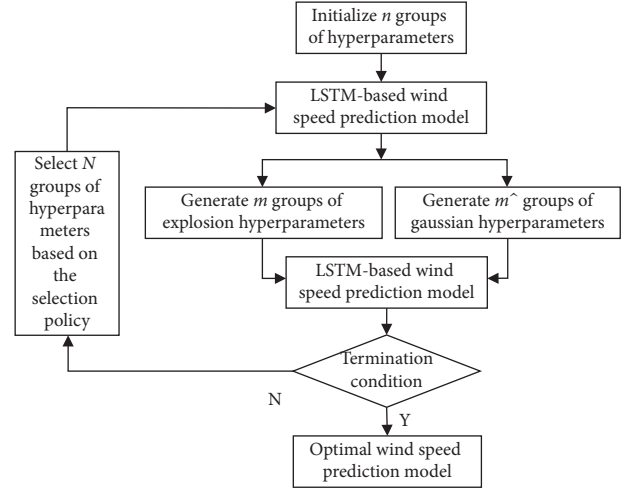

Figure 4: The specific calling relationship between the FWA hyperparameter optimization and the LSTM-based wind speed prediction model.

4. Experimental Evaluations

4.1. Experimental Environment Configuration and Parameter Settings.
This study selects the measured wind speed data of a wind farm in 2015, starting from January 1, 2015, to December 31, 2015, with an interval of 1 hour, each containing 8759 data packets. This paper selects some data segments for model analysis.

For the prediction results of the network model, three error analysis indicators are used to verify the prediction accuracy, namely, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Error (Mean Absolute Percentage Error, MAPE). The calculation methods of MAE and MAPE are shown in equations (13) and (14).

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|, \tag{13}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%. \tag{14}$$

It can be seen from formula (12) that the smaller the value of RMSE, the smaller the average error between the prediction result and the actual data, the higher the prediction accuracy of the model, and the better the prediction performance of the model. Similarly, it can be seen from formulas (13) and (14) that the more MAE and MAPE values tend to 0, the better the prediction effect of the model is and the more perfect the model is. on the contrary, the larger the values, the greater the error and the worse the prediction effect of the model.

To fully verify the prediction effect of the LSTM wind speed prediction model on the wind speed data after FWA optimization, the optimized model prediction results were compared with the typical LSTM prediction results, other neural network models, and regression prediction methods. As shown in the firework algorithm [45, 46], the initial

(1) Wind speed data preparation and preprocessing;
(2) Normalize the raw data;
(3) Divide training set and test set;
(4) Construct LSTM wind speed prediction model. Set partial parameters and fix the number n of optimized parameter;
(5) FWA parameter initialization (fireworks population size P, solving space dimension d, maximum number of iterations $iter\_max$, explosion amplitude range control parameter $\hat{A}$, the maximum number of sparks m, the number of variation sparks $\hat{m}$, the parameters a and b that limit the number of sparks produced by the explosion, the minimum normal value ε that tends to zero, the solution space boundaries B_u and B_i);
(6) Initialize the values of n-dimensional parameter combinations of P groups randomly in the solution space;
(7) Initialize the global optimal parameter combination gbest_parameters, the partial optimal parameter combination pbest_parameters, and the best fitness function value Pg;
(8) While the end condition is false:
(9)   Apply the n-dimensional parameter combinations of P groups, respectively, to the LSTM network flow prediction model for training, and calculate the current fitness function value;
(10)   Get the current best fitness value P_i and the corresponding parameter combination pbest_parameters;
(11)   if $F_i < F_g$
(12) $F_g = F_i$;
(13)   gbest_parameters = pbest_parameters;
(14)   end if
(15)   for each parameter combination
(16)   Calculate the search direction and position of the new parameter combination according to formulas (8) and (9);
(17)   Fix the updated parameter in the selected values;
(18)   end for
(19)   The number of iterations+1;
(20) end while
(21) Return the gbest_parameters;
(22) gbest_parameters is introduced into LSTM wind speed prediction model to predict test data and calculate prediction error.

ALGORITHM 1: FWA-LSTM wind speed prediction algorithm.

number of fireworks $N = 5$, the size of the fireworks population $P = 50$, the preset maximum explosion amplitude $\hat{A} = 40$, the maximum number of sparks $m = 20$, the number of mutant sparks $\hat{m} = 5$, the constants $a = 0.04$, $b = 0.8$, and the maximum number of generations is set to 100.

4.2. Wind Speed Prediction Results Based on FWA-LSTM

4.2.1. Data Processing

(1) Data resampling: Figure 5 shows wind speed data after null filling by the KNN algorithm.

(2) Data normalization: the range standardization method (equation (7)) is used to process the wind speed data so that the sample data value is between 0 and 1. The processing result is shown in Figure 6.

(3) Data division: the normalized data is divided into train set and test set according to the simple cross-validation method. The first 80% of the data is used as training data for the training of the LSTM network model. The remaining 20% of the data is used as prediction data to verify the efficiency of the model.

4.2.2. Wind Speed Prediction Based on Basic LSTM

(1) Define the network: this prediction uses a four-layer LSTM model with one input layer, two hidden layers, and one output layer.

The specific connection method of the three-layer LSTM is as follows: the first layer of LSTM receives input with time steps of 1, data_dim = 3, and the number of neurons is 64; the second layer uses the results of the first input layer as input for training and passes its output to the next hidden layer. The number of neurons is the same as that of the first layer, and the third layer of the hidden layer (Dense) uses the first layer of LSTM. The output of the third layer is the input; the output layer of the third layer takes the output of the second hidden layer as the input and is connected to a fully connected layer. A one-dimensional vector with a length of 200 output from the fully connected layer is the final output result, which represents the predicted value of 200 data points in the future. To prevent LSTM from overfitting, a dropout layer is added between the first layer and the hidden layer for regularization. After repeated testing, it is found that the accuracy of the training set is the highest when dropout = 0.25.

(2) Compile the network: LSTM network compilation uses adaptive moment estimation (Adam) algorithm as the optimizer and a mean square error loss function as the objective function.

(3) Fitting the network: the LSTM network is trained on 1600 pieces of train data and 200 pieces of valid data are used for validation. The number of iterations epochs = 50, look_back = {1, 5, 10}, and batch_size = 128.

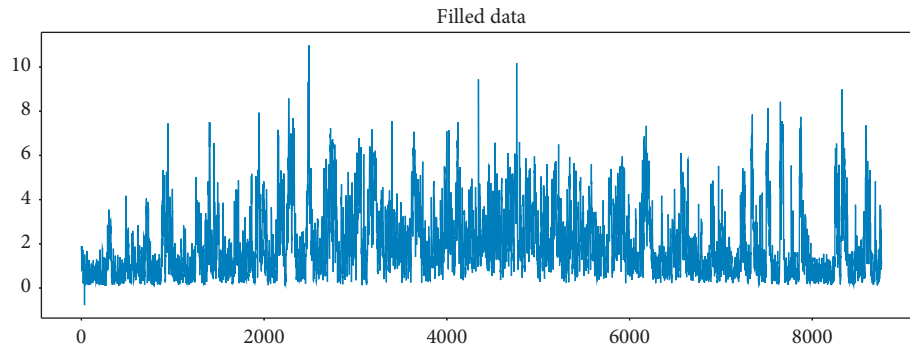Filled data

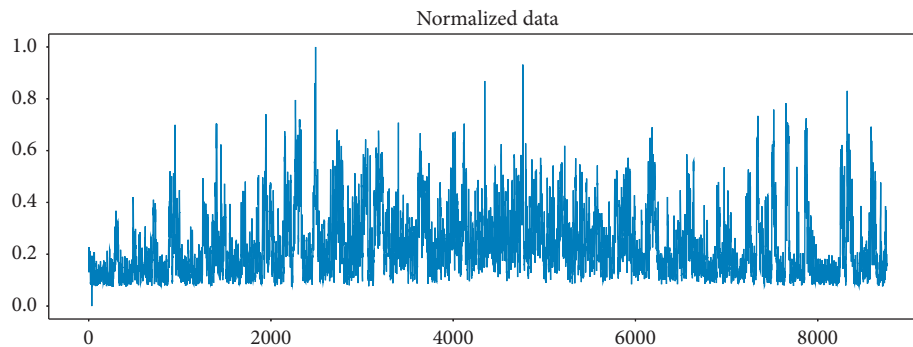FIGURE 5: Wind speed data after null filling.

Normalized data

FIGURE 6: Wind speed data after normalization.

(4) Network evaluation: when look_back takes 1, 5, and 10, respectively, and the number of hidden layers (LN) is 1 and 2, respectively, the loss data of the model training process are shown in Figure 7.

According to Figure 7, the loss of each training shows a downward trend, indicating that the model is effective.

(5) Wind speed forecasting: 200 pieces of test data are predicted and the results are shown in Figure 8. TestOriginal_result represents the original data; testPredict_result_101, testPredict_result_105, and testpredict_110, respectively, represent the prediction results when LN = 1, look_back takes 1, 5, and 10, respectively. TestPredict_result_201, testPredict_result_205, and testpredict_210, respectively, represent the prediction results when LN = 2, and look_back takes 1, 5, and 10, respectively.

(6) Error of the prediction mode: LSTM models corresponding to different parameter combinations were used for wind speed prediction, and errors (RMSE, MAE, MAPE) of each model validation set were compared (see Table 1).

It can be concluded that, for the wind speed data, the prediction effect of the parameter combination set by the empirical method is unstable and cannot achieve the optimal prediction performance. Therefore, the fireworks algorithm (FWA) is adopted to optimize the model; that is, an intelligent algorithm is used to efficiently obtain the parameter combination with the optimal prediction effect.

4.2.3. Hyperparameter Optimization according to FWA. To show the process of the optimal parameter value of the LSTM wind speed model determined by the firework algorithm, Figure 9 shows the changes in the number of nodes and the time step during the optimization process of the FWA-LSTM12 model.

Figure 9 shows that, for the prediction of wind speed data, the fitness value tends to be stable starting from the 9th iteration; that is, the FWA to optimize the LSTM-based wind speed prediction model converges easily.

The changes in the number of nodes and time steps in the optimization process of the FWA-LSTM23 model are shown in Figure 10.

It can be seen from Figure 10 that the optimal parameters of the FWA-LSTM model are set to node 1 = 5, node 2 = 2, and look_back = 5. Therefore, in the prediction of wind speed data used in this article, the best configuration of the LSTM model is that the number of neurons in the first layer is set to 5, the number of neurons in the second layer is set to 2, and the time step is set to 5.

4.3. Result Analysis. To evaluate the prediction performance of the LSTM model after parameter optimization by FWA, wind speed data samples at 200-time points are used for verification. Firstly, the basic LSTM and PSO-LSTM (LSTM optimized by the particle swarm optimization) are tested for comparison. Figure 11 shows the prediction results of these three methods.

It can be seen from Figure 11 that the prediction effects of FWA-LSTM and PSO-LSTM models are similar and both better than the basic LSTM method. To compare the
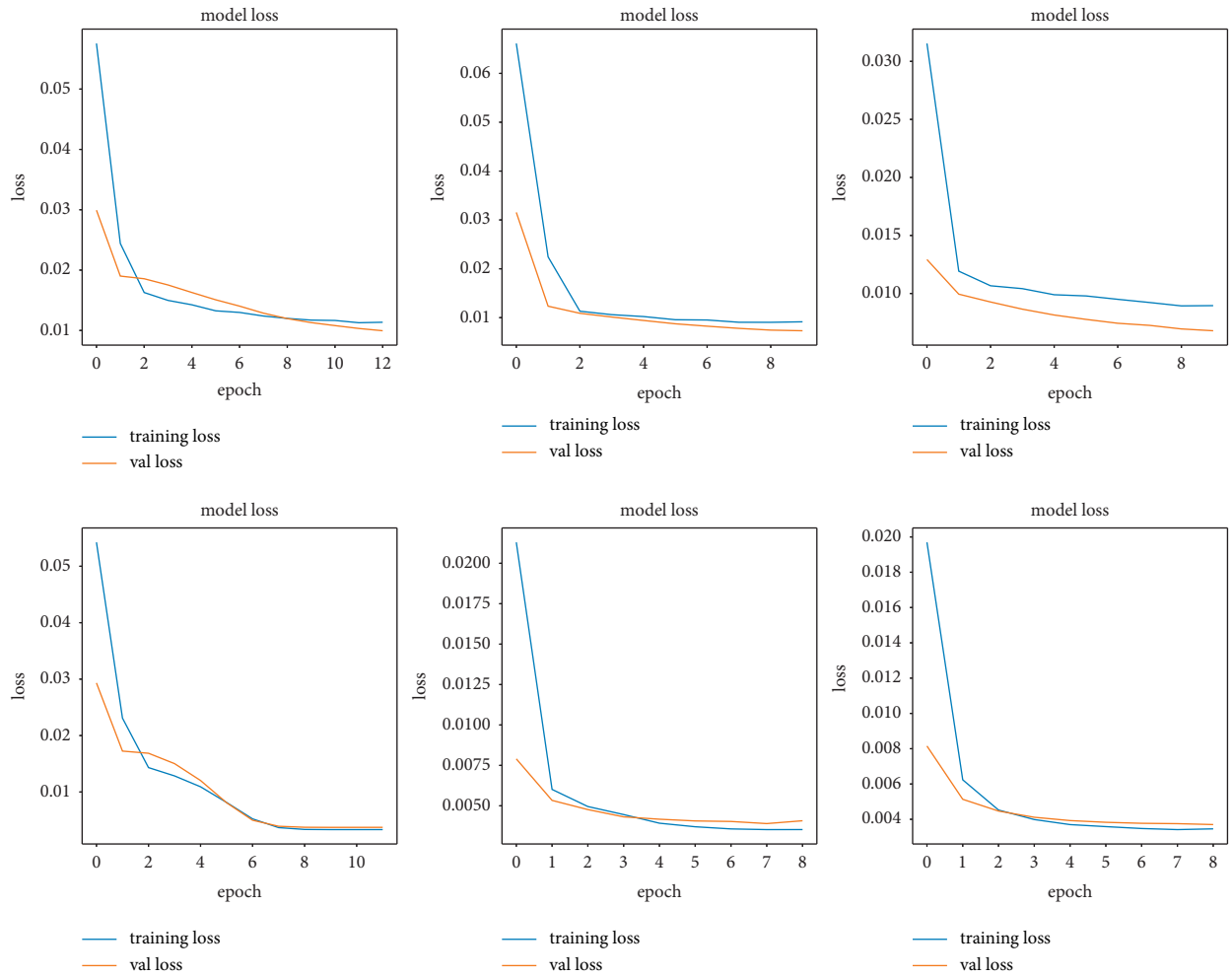
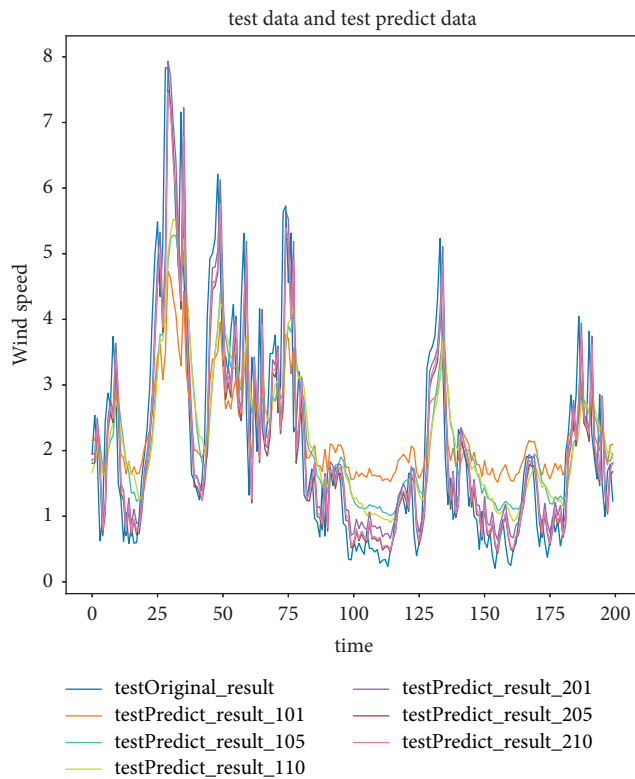FIGURE 7: Model training loss under different parameter combinations.

FIGURE 8: Test results of partial data.

TABLE 1: Errors of LSTM-based wind speed prediction models under different parameter combinations.

Model	Parameters	RMSE	MAE	MAPE (%)
LSTM101	(look_back, LN) = (1,1)	1.18	0.96	94.77
LSTM105	(look_back, LN) = (5,1)	0.98	0.74	66.51
LSTM110	(look_back, LN) = (10,1)	0.96	0.71	63.23
LSTM201	(look_back, LN) = (1,2)	0.73	0.51	42.09
LSTM205	(look_back, LN) = (5,2)	0.73	0.58	33.28
LSTM210	(look_back, LN) = (10,2)	0.72	0.58	35.34

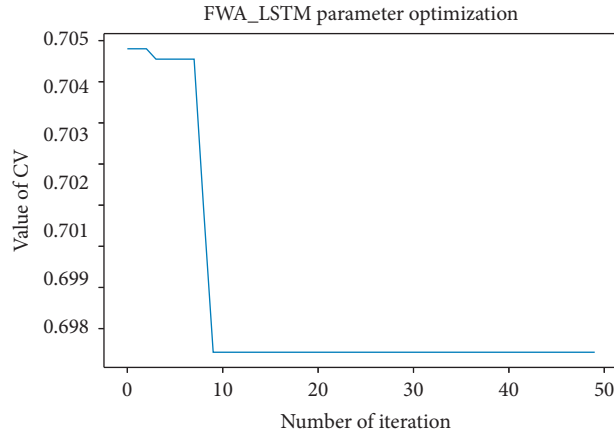FIGURE 9: The change of the fitness value.

(a)   (b)
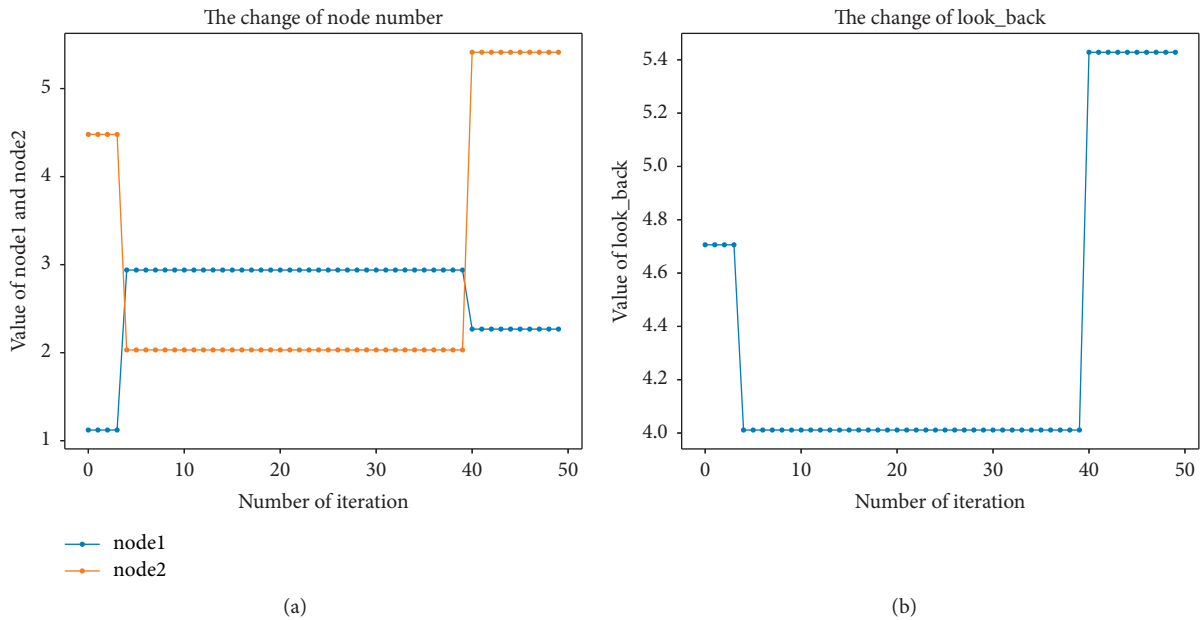
FIGURE 10: The change of parameters during FWA iteration. (a) The change of node number. (b) The change of time step.

performance of the three methods more clearly, their prediction performance evaluation index values are calculated and shown in Table 2.

It can be seen from Table 2 that compared with the PSO-LSTM and the basic LSTM, the FWA-LSTM has slightly smaller prediction errors of RMSE and MAPE, while the MAE is close to PSO-LSTM. On the whole, FWA-LSTM is considered to be superior to the PSO-LSTM and the basic LSTM.

To further verify the prediction effect of the improved FWA-LSTM prediction model, it is compared with other neural network prediction methods, such as Gated Recurrent Unit (GRU), Simple Recurrent Neural Network (SimpleRNN), and Bidirectional Recurrent Neural Network (BiRNN), and other predictive models such as SVR and ARIMA were also compared. Similarly, using 200-time points of wind speed data samples for verification, their
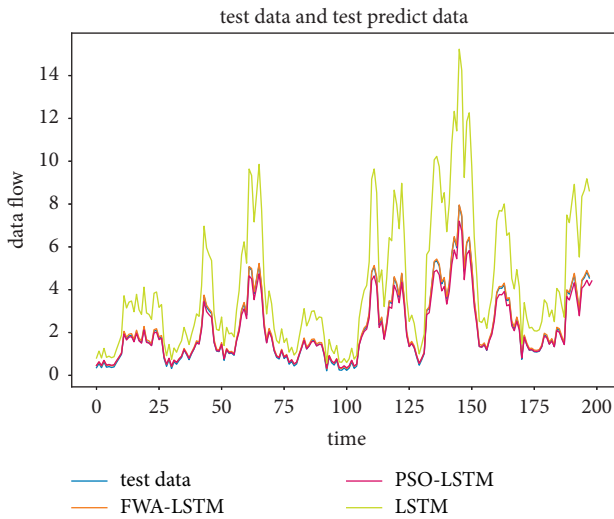
FIGURE 11: Prediction results of FWA-LSTM and other prediction methods.

TABLE 2: The comparison of prediction error index values of the eight models.

Method	RMSE	MAE	MAPE (%)
FWA-LSTM	0.64	0.46	30.05
PSO-LSTM	0.69	0.46	33.07
LSTM	0.76	0.68	38.43

TABLE 3: The comparison of prediction error index values of the six models.

	FWA-LSTM	GRU	SimpleRNN	BiRNN	SVR	ARIMA
RMSE	0.64	0.71	0.74	0.73	0.89	0.83
MAE	0.46	0.51	0.51	0.52	0.68	1.09
MAPE (%)	30.05	30.22	32.13	31.42	38.43	60.16

prediction performance evaluation index values were obtained, respectively, and the results are shown in Table 3.

It can be seen from Table 3 that the prediction errors of RMSE, MAE, and MAPE of the FWA-LSTM model optimized by FWA are all less than those of other tested prediction methods such as GRU, SimpleRNN, BiRNN, SVR, and ARIMA, so it is said that the FWA-LSTM prediction model has a better prediction effect in wind speed prediction. Therefore, it is said that the FWA-LSTM is more suitable for dealing with the real-time sudden change of wind speed data.

In summary, the proposed FWA-LSTM method has a better prediction effect and higher reliability for the future prediction of wind speed.

5. Conclusions

Wind speed prediction can be applied to wind energy optimization and has important reference significance for wind power planning and the stable operation of the power system. This paper first established a wind speed prediction model based on the nonparametric model LSTM neural network, optimized the hyperparameters of the established LSTM prediction model with the firework algorithm, and reduced the prediction Root Mean Square Error compared to the empirical method of obtaining parameters, and the FWA-LSTM is better than the double-layer LSTM in the wind speed data prediction.

The improved model FWA-LSTM is applied to wind speed prediction and compared with the prediction effects of other neural network prediction methods and regression methods. The experimental results show that compared to other prediction models and the traditional LSTM model, the FWA-LSTM method reduces the prediction errors, which significantly reduces the prediction error and improves the accuracy of wind speed prediction. The next step will continue to combine a variety of prediction methods to improve the prediction accuracy of wind speed prediction.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] S. Samadianfard, S. Hashemi, K. Kargar et al., "Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm," *Energy Report*, vol. 6, pp. 1147–1159, 2020.

[2] X. Ma, Y. Chen, W. Yi, and Z. Wang, "Prediction of extreme wind speed for offshore wind farms considering parametrization of surface roughness," *Energies*, vol. 14, no. 4, 2021.

[3] R. Jing, Z. Zhang, Z. Liang et al., "Review of wind speed and wind power prediction methods," *Journal of Applied Sport Manageinent*, vol. 39, no. 6, pp. 1045–1059, 2020.

[4] X. Liu, Z. Lin, and Z. Feng, "Short-term offshore wind speed forecast by seasonal ARIMA - a comparison against GRU and LSTM," *Energy*, vol. 227, 2021.

[5] N. Bokde, A. Feijóo, D. Villanueva, and K. Kulat, "A review on hybrid empirical mode decomposition models for wind speed and wind power prediction," *Energies*, vol. 12, no. 2, 2019.

[6] E. B. Ssekulima, M. B. Anwar, A. A. Hinai, and M. S. E. Moursi, "Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: a review," *IET Renewable Power Generation*, vol. 10, no. 7, pp. 885–989, 2016.

[7] J. Wang, Y. Song, F. Liu, and R. Hou, "Analysis and application of forecasting models in wind power integration: a review of multi-step-ahead wind speed forecasting models," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 960–981, 2016.

[8] L. Xiang and Z. Deng, "Short-term wind speed forecasting based on improved empiricalwavelet transform and least

squares support vector machines," *Acta Energiae Solaris Sinica*, vol. 42, no. 02, pp. 97–103, 2021.

[9] S. Zhang, C. Liu, W. Wang, and B. Chang, "Twin least square support vector regression model based on gauss-laplace mixed noise feature with its application in wind speed prediction," *Entropy*, vol. 22, no. 10, 2020.

[10] A. Khosravi, L. Machado, and R. O. Nunes, "Time-series prediction of wind speed using machine learning algorithms: a case study Osorio wind farm, Brazil," *Applied Energy*, vol. 224, pp. 550–566, 2018.

[11] Y. Dong, J. Niu, Q. Liu et al., "A hybrid prediction model for wind speed using support vector machine and genetic programming in conjunction with error compensation," *Stochastic Environmental Research and Risk Assessment*, vol. 11, pp. 1–14, 2021.

[12] O. Karakuş, E. E. Kuruoğlu, and M. A. Altınkaya, "One-day ahead wind speed/power prediction based on polynomial autoregressive model," *IET Renewable Power Generation*, vol. 11, no. 11, pp. 1430–1439, 2017.

[13] A. Xie, H. Yang, J. Chen, L. Sheng, and Q. Zhang, "A short-term wind speed forecasting model based on a multi-variable long short-term memory network," *Atmosphere*, vol. 12, no. 5, p. 651, 2021.

[14] Z. Gan, C. Li, J. Zhou et al., "Temporal convolutional networks interval prediction model for wind speed forecasting," *Electric Power Systems Research*, vol. 191, Article ID 106865, 2021.

[15] C. Liang, Y. Liu, J. Zhou et al., "Wind speed prediction at multi-locations based on combination of recurrent and convolutional neural networks," *Power System Technology*, vol. 45, no. 2, pp. 534–542, 2021.

[16] H. Fan, X. Zhang, S. Mei et al., "UItra-short-term wind speed prediction model for wind farms based on spatiotemporal neural network," *Automation of Electric Power Systems*, vol. 45, no. 1, pp. 28–35, 2021.

[17] J. Wang, Y. Wang, Z. Li, H. Li, and H. Yang, "A combined framework based on data preprocessing, neural networks and multi-tracker optimizer for wind speed prediction," *Sustainable Energy Technologies and Assessments*, vol. 40, Article ID 100757, 2020.

[18] H. Liu, X. Mi, and Y. Li, "Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network," *Energy Conversion and Management*, vol. 166, pp. 120–131, 2018.

[19] Z. ZhangL. Ye et al., "Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression," *Applied Energy*, vol. 247, pp. 270–284, 2019.

[20] M.-R. Chen, G.-Q. Zeng, K.-D. Lu, and J. Weng, "A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN, and LSTM," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6997–7010, 2019.

[21] J. Ding, G. Chen, Y. Huang et al., "Short-term wind speed prediction based on CEEMDAN-SE-improved PIO-GRNN model," *Measurement and Control*, vol. 54, no. 1-2, pp. 73–87, 2021.

[22] Y. Jiang, S. Liu, N. Zhao, J. Xin, and B. Wu, "Short-term wind speed prediction using time varying filter-based empirical mode decomposition and group method of data handling-based hybrid model," *Energy Conversion and Management*, vol. 220, Article ID 113076, 2020.

[23] D. Xiong, W. Fu, K. Wang, P. Fang, T. Chen, and F. Zou, "A blended approach incorporating TVFEMD, PSR, NNCT-based multi-model fusion and hierarchy-based merged

optimization algorithm for multi-step wind speed prediction," *Energy Conversion and Management*, vol. 230, Article ID 113680, 2021.

[24] M. D. Liu, L. Ding, and Y. L. Bai, "Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction," *Energy Conversion and Management*, vol. 233, Article ID 113917, 2021.

[25] S. W. Fei, "The hybrid model of empirical wavelet transform and relevance vector regression for monthly wind speed prediction," *International Journal of Green Energy*, vol. 17, no. 10, pp. 583–590, 2020.

[26] Y. Liu, J. Zhou, J. Yan et al., "Wind speed real-time prediction method based on multi-task learning," *Renewable Energy Resources*, vol. 39, no. 04, pp. 481–487, 2021.

[27] Y. Li, X. Chen, C. Li, G. Tang, Z. Gan, and X. An, "A hybrid deep interval prediction model for wind speed forecasting," *IEEE Access*, vol. 9, pp. 7323–7335, 2020.

[28] Y. J. Natarajan and D. S. Nachimuthu, "New SVM kernel soft computing models for wind speed prediction in renewable energy applications," *Soft Computing*, vol. 24, pp. 1–18, 2019.

[29] Y. Xie, C. Li, G. Tang, and F. Liu, "A novel deep interval prediction model with adaptive interval construction strategy and automatic hyperparameter tuning for wind speed forecasting," *Energy*, vol. 216, Article ID 119179, 2021.

[30] M. Li, Z. Zhang, T. Ji, and Q. H. Wu, "Ultra-short term wind speed prediction using mathematical morphology decomposition and long short-term memory," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 4, pp. 890–900, 2020.

[31] S. Zhang, Y. Chen, J. Xiao, W. Zhang, and R. Feng, "Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism," *Renewable Energy*, vol. 174, pp. 688–704, 2021.

[32] L. Lopez, I. Oliveros, L. Torres et al., "Prediction of wind speed using hybrid techniques," *Energies*, vol. 13, no. 23, 2020.

[33] N. Zeng, Z. Huo, X. U. Chang et al., "Minute-scale ultra-short-term wind speed prediction based on improved HHT," *Journal of Chinere Society of Power Engineering*, vol. 41, no. 4, pp. 309–315, 2021.

[34] Y. Zhang, G. Pan, Y. Zhao, Q. Li, and F. Wang, "Short-term wind speed interval prediction based on artificial intelligence methods and error probability distribution," *Energy Conversion and Management*, vol. 224, Article ID 113346, 2020.

[35] W. Xu, P. Liu, L. Cheng et al., "Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy," *Renewable Energy*, vol. 163, pp. 772–782, 2021.

[36] Z. Liu, R. Hara, and H. Kita, "Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting," *Energy Conversion and Management*, vol. 238, Article ID 114136, 2021.

[37] T. Fu and M. Yang, "Mixed wind speed prediction based on ensemble empirical mode decomposition and improved cuckoo algorithm to optimize BPNN," *Journal of Qufu Normal University*, vol. 47, no. 2, pp. 27–34, 2021.

[38] L. Bingchen, Y. Huijun, D. Huaxuan et al., "Short-term wind speed prediction based on ceemd and lstm-arima," *China Measurement and Test*, IEEE, vol. 47, pp. 1–6, 2021.

[39] X. Huang, J. Wang, and B. Huang, "Two novel hybrid linear and nonlinear models for wind speed forecasting," *Energy Conversion and Management*, vol. 238, Article ID 114162, 2021.

[40] X. Chen, C. S. Lai, W. W. Y. Ng et al., "A stochastic sensitivity-based multi-objective optimization method for short-term wind speed interval prediction," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 9, pp. 1–12, 2021.

[41] F. Han, Y. Han, J. Lu et al., "Wind speed prediction based on nested shared weight long short-term memory network," *The Journal of China Universities of Posts and Telecommunications*, vol. 28, no. 1, 2021.

[42] Y. Tan and Y. Zhu, "Fireworks algorithm for optimization," in *Proceedings of the International Conference in Swarm Intelligence*, pp. 355–364, Springer, Beijing, China, June 2010.

[43] Y. Tan and Z. Shaoqiu, "Recent advances in fireworks algorithm," *CAAI Transactions on Intelligent Systems*, vol. 9, no. 05, pp. 515–528, 2014.

[44] X. Han, L. Zheng, L. Wang, Z. Hong, and X. Wang, "Fireworks algorithm based on dynamic search and tournament selection," *International Journal of Computers and Applications*, vol. 43, pp. 1–12, 2019.

[45] J. Zhao, D. Ting, and C. Bo, "AQI prediction based on long short-term memory model with spatial-temporal optimizations and fireworks algorithm," *Journal of Wuhan University (Natural Science Edition)*, vol. 65, no. 3, pp. 250–262, 2019.

[46] Y. Xu, M. Wang, and W. Fan, "Parameter optimization of secondary system fault correlation analysis based on AFWA," *Journal of North China Electric Power University*, vol. 48, no. 4, pp. 1–8, 2021.

[47] X. Li, K. Peng, J. Peng, and D. Hou, "Effect of thermal damage on mechanical behavior of a fine-grained sandstone," *Arabian Journal of Geosciences*, vol. 14, no. 1212, pp. 1–12, 2021.