Hindawi
Mathematical Problems in Engineering
Volume 2019, Article ID 1934796, 15 pages
https://doi.org/10.1155/2019/1934796

Research Article

Nickel Price Forecast Based on the LSTM Neural Network Optimized by the Improved PSO Algorithm

Bilin Shao,[1] Maolin Li,[1] Yu Zhao,[1] and Genqing Bian[2]

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

Correspondence should be addressed to Maolin Li; 1412426478@qq.com

Received 19 May 2019; Revised 16 July 2019; Accepted 16 August 2019; Published 8 September 2019

Academic Editor: Francesco Lolli

Nickel is a vital strategic metal resource with commodity and financial attributes simultaneously, whose price fluctuation will affect the decision-making of stakeholders. Therefore, an effective trend forecast of nickel price is of great reference for the risk management of the nickel market's participants; yet, traditional forecast methods are defective in prediction accuracy and applicability. Therefore, a prediction model of nickel metal price is proposed based on improved particle swarm optimization algorithm (PSO) combined with long-short-term memory (LSTM) neural networks, for higher reliability. This article introduces a nonlinear decreasing assignment method and sine function to improve the inertia weight and learning factor of PSO, respectively, and then uses the improved PSO algorithm to optimize the parameters of LSTM. Nickel metal's closing prices in London Metal Exchange are sampled for empirical analysis, and the improved PSO-LSTM model is compared with the conventional LSTM and the integrated moving average autoregressive model (ARIMA). The results show that compared with the standard PSO, the improved PSO has a faster convergence rate and can improve the prediction accuracy of the LSTM model effectively. In addition, compared with the conventional LSTM model and the integrated moving average autoregressive (ARIMA) model, the prediction error of the LSTM model optimized by the improved PSO is reduced by 9% and 13%, respectively, which has high reliability and can provide valuable guidance for relevant managers.

1. Introduction

Nickel is a rare metal with outstanding physical and chemical properties. It is known as the "vitamin of the steel industry" and is also the raw material of green batteries. In recent years, nickel metal plays an increasingly important role in the fields of high-tech industry and military industry and is a strategic metal resource for many countries. However, China's nickel resources are relatively small, and its reserves only account for 3.93% of the world [1]. Besides, China's nickel mining and smelting technology are relatively backward, and the annual average nickel output is less than 5% of global nickel production. The contradiction between supply and demand makes China need to import a large number of nickel resources every year. The price of nickel metal is susceptible to many factors such as supply and demand, national policy [2, 3], and the WTO environment. So the nickel price time series presents the characteristics of highly unstable, complex, and unpredictable, making it difficult for nickel market participants such as related companies, investors, and consumers to grasp business opportunities and conduct normal production, operation, and consumption accurately. Therefore, it is of great significance to predict the price of nickel metal effectively.

In recent years, many scholars have conducted a series of studies on the price forecast of commodities. Research objects involve energy [4, 5], metals [6, 7], and agricultural products [8, 9]. The research methods can be divided into two categories: qualitative and quantitative forecasting. The qualitative prediction primarily analyzes the influencing factors and fluctuation characteristics of prices [5, 10], and the quantitative prediction mainly uses statistical methods [11], shallow learning [12], and deep learning methods [13] to construct prediction models. Although the existing literature on price predicting covers a wide range of research subjects, there are few studies on the prediction of nickel prices at home and abroad. And most studies focus on the development of nickel [14] and explore the influencing factors of nickel price fluctuations [15–17], but there is no formal, highly reliable quantitative prediction model of nickel prices. In addition, although there are many quantitative prediction methods, the commonly used statistical methods and the shallow learning methods still have certain defects, and there are certain conditions for use. For example, the integrated moving average autoregressive model (ARIMA) and the gray prediction model are generally suitable for processing linear or stationary sequences, but they are difficult to obtain the nonlinear characteristics of the data. Although artificial neural networks have self-organizing and self-adjusting capabilities to deal with complex nonlinear problems, their ability to predict long-term sequences is limited. The long-short-term memory neural network (LSTM) has memory cells, which can extract deep features from a small number of samples. It is suitable for processing time series and has achieved sound application effects in many fields [18, 19]. However, the parameters of the LSTM model are usually determined by experience, so the subjectivity is strong and will affect the fitting ability of the model. Given the above problems, Feng and Zi-Jun [20] constructed a BP neural network with chaotic PSO optimization to predict carbon price. The results show that the prediction accuracy and stability of the model are better than the traditional BP neural network and the model optimized by standard PSO. Catalao et al. [21] and other scholars proposed a hybrid method based on wavelet transform (WT), particle swarm optimization (PSO), and adaptive network fuzzy inference system (ANFIS) for short-term electricity price forecasting and applied to the short-term electricity price prediction in the Spanish electricity market. By comparing with ARIMA, NN, wavelet-ARIMA, FNN, and other models, it is found that the hybrid model WPA is perfect in both prediction accuracy and calculation time.

To explore a more reasonable and efficient method to predict nickel metal price, the LSTM neural network optimized by the improved particle swarm optimization (PSO) algorithm is proposed to predict the price of nickel metal. Nonlinear decreasing assignment method and sine function are introduced to improve the inertia weight and the learning factors of the PSO algorithm, respectively. Then, the improved PSO algorithm is used to optimize the parameters of the LSTM, and the nickel price of the London Metal Exchange is used as a sample to predict the nickel price. Then, the model is compared with the conventional LSTM model and the time-predictive model ARIMA with high prediction accuracy to verify the validity and reliability of the model.

2. Related Research Theory

The LSTM neural network can learn the complex association between features and tags, but its learning process is highly susceptible to time step, the number of hidden layers, and the number of nodes in each hidden layer. However, these parameters are usually determined by manual adjustment, which not only increases the complexity of the operation process but also may result in lower prediction accuracy [22]. The PSO algorithm is simple in operation and fast in convergence and has a significant effect in solving complex optimization problems. Therefore, the PSO algorithm is adopted in this article to optimize the three parameters of the LSTM model. According to the root mean square error of prediction results of the LSTM model corresponding to different parameters, the PSO algorithm adaptively adjusts the position and velocity of particles to find the optimal parameter combination of the LSTM model.

2.1. LSTM Model. The long-short-term memory (LSTM) model is a variant of the recurrent neural network, which was proposed and improved based on the recurrent neural network [23]. LSTM changes the weight of the self-loop by adding input gate, forgetting gate, and output gate, which can ease the problem of gradient disappearance and gradient explosion in model training, and make up the defects of traditional RNN model. In addition, LSTM has excellent advantages in dealing with nonlinear time-series data with associated relationships for the particular memory function. The basic unit structure of LSTM is shown in Figure 1.

In Figure 1, the LSTM is a four-layer structure with interactions between structures, where h_{t-1} and h_t represent the output of the previous cell and the current cell, x_t represents the input of the current unit, the box represents the neural network layer, the content in the box is the activation function, and the circles represent the arithmetic rules between vectors. C_t represents the state of the neuron at time t and f_t represents the forgetting threshold, which controls the probability of forgetting the state of the last unit neuron through the sigmoid activation function. i_t represents the input threshold, which determines the information that needs to be updated by the sigmoid function, and then uses the tanh activation function to generate a new memory C_t', and ultimately controls how much new information will be added to the neuron state. And o_t represents the output threshold, which determines which parts of the neuron state are output by the sigmoid function, and uses the tanh activation function to process the neuron state to get the final output.

The input layer of the LSTM neural network contains three parameters: sample, time step, and feature dimension, where the time step can be understood as the length of the sliding window. The value of the time step determines how many previous consecutive input data will affect the current input data; that is, how many historical data will be used to predict the data of the next time. The setting of this parameter helps the LSTM neural network to learn long-term dependency information within the time-series data to improve the accuracy of the prediction results. The hidden layer has one parameter: the number of hidden layer nodes, that is, the number of neurons contained in the hidden layer. In an LSTM unit, the physical implementation of each gate is essentially a gate function implemented by several hidden-layer neurons. The hidden-layer neurons are fully connected

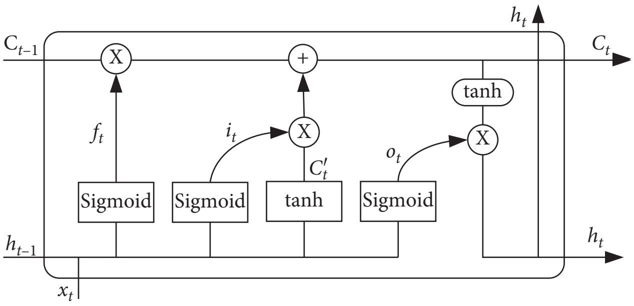


Figure 1: The structure diagram of the LSTM unit.

with the input vector, and the input vectors are weighted and summed by the weight coefficient matrix, and then the offset matrix is added to obtain the output of the hidden layer through the excitation function. The output layer contains two parameters: the number of hidden layer nodes and the output dimension. The calculation process of the LSTM memory unit can be expressed as follows:

$$\begin{aligned} f_t &= \text{sigmoid}\left(w_f\left[x_t, h_{t-1}\right] + b_f\right), \\ i_t &= \text{sigmoid}\left(w_i\left[x_t, h_{t-1}\right] + b_i\right), \\ o_t &= \text{sigmoid}\left(w_o\left[x_t, h_{t-1}\right] + b_o\right), \\ C'_t &= \tanh\left(w_C\left[x_t, h_{t-1}\right] + b_C\right), \end{aligned} \tag{1}$$

where w_f, w_i, w_o, and w_C represent the weight coefficient matrix corresponding to the forgetting gate, the input gate, the output gate, and the neuron state matrix, respectively; and, b_f, b_i, b_o, and b_C represent corresponding offset constants, respectively. According to the above formula, the state C_t and the output h_t of the neuron can be further calculated as shown in the following equations:

$$C_t = f_t C_{t-1} + i_t C'_t, \tag{2}$$

$$h_t = o_t \tanh\left(C_t\right). \tag{3}$$

This is the forward calculation process of the LSTM neural network model. The particular structure enables it to learn long-term dependencies and has been widely used in text analysis [24], time-series prediction [25], and other fields.

2.2. Standard PSO Algorithm. Particle swarm optimization (PSO) is a global optimization algorithm with simple rules and fast convergence [26]. It has been widely used in neural network training and structural optimization design [20].

When solving the optimization problem, PSO updates the velocity and position of the particle by tracking the individual optimal particle and the group optimal particle. It can be described as follows: In the D-dimensional search space, there are m particles forming a group. In the t-th iteration, the position and velocity of the i-th particle are $X_{i,t}$ and $V_{i,t}$, respectively. The particle updates its position and speed by supervising the individual's optimal fitness value $p\text{best}_i$ and the group's current optimal fitness value $g\text{best}_t$. The specific evolution formula is shown in the following equations:

$$\begin{aligned} V_{i,t+1} = {} & w * V_{i,t} + c_1 * \text{rand} * \left(p\text{best}_i - X_{i,t}\right) \\ & + c_2 * \text{rand} * \left(g\text{best}_t - X_{i,t}\right), \end{aligned} \tag{4}$$

$$X_{i,t+1} = X_{i,t} + \lambda * V_{i,t+1}, \tag{5}$$

where w is the inertia weight, c_1 and c_2 are the learning factors, rand is the random number between [0, 1], and λ is the velocity coefficient, $\lambda = 1$. The particle's velocity update formula consists of three parts. The first part is $w * V_{i,t}$, which indicates that the particle maintains its previous speed trend and plays the role of balancing global search and local search. Among them, the inertia weight w will affect the performance of the algorithm to a large extent. If the weight is set too large, it is not conducive to the local search in the later stage of the algorithm. If the setting is too small, it is not conducive to the global search in the early stage of the algorithm, which will slow down the convergence speed of the population. The second part is $c_1 * \text{rand} * (p\text{best}_i - X_{i,t})$, which indicates the ability of the particle to tend to the best position in its history. Among them, the learning factor c_1 reflects the preference degree of the particle learning to the extreme value of the individual, and its value will affect the global searching ability of the algorithm, which will affect the convergence speed of the algorithm. The third part is $c_2 * \text{rand} * (g\text{best}_t - X_{i,t})$, which indicates the ability of the particle to tend to the best position in the history of the population. Among them, the learning factor c_2 reflects the preference degree of the particle learning to the global extremum, and its value will affect the local search ability of the algorithm, which may cause local optimal.

2.3. ARIMA Model. The ARIMA model is called autoregressive integrated moving average model, which consists of autoregressive and moving average. Its expression is ARIMA(p, d, and q), where p is the autoregressive order, d is the difference order, and q is the moving average order. ARIMA is a linear time-series prediction method with high accuracy commonly used in statistics, which can well describe the linear characteristics of time series when dealing with complex time series. Therefore, it is widely used in the field of price prediction research.

The expression of the ARIMA model is as follows:

$$\begin{aligned} X_t = {} & \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \\ & + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \end{aligned} \tag{6}$$

where $\varphi_1, \varphi_2, \ldots, \varphi_p$ are autoregressive coefficients, p is the autoregressive order, $\theta_1, \theta_2, \ldots, \theta_q$ are moving average coefficients, q is the moving average order, and $\{\varepsilon_t\}$ is the white noise sequence.

The process of constructing ARIMA model for time-series prediction is shown in Figure 2.

The premise of constructing the ARIMA model is that the sequence is stationary, so the first step is to verify the smoothness of the data. In this article, the ADF (Augmented

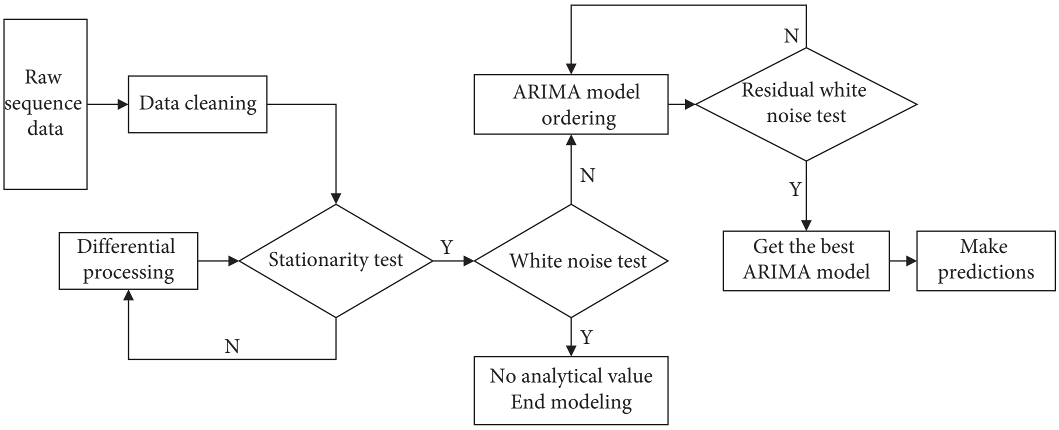


Figure 2: Prediction flowchart of ARIMA model.

Dickey–Fuller test) method is used to test the stability of the sequence. In addition, it is necessary to verify whether the data are a white noise sequence because the white noise sequence is randomly perturbed and cannot be predicted. In the model identification stage, the parameters' range of the ARIMA model can be roughly determined by plotting the autocorrelation graph and the partial autocorrelation graph, and then the parameters of the model are finally determined according to the information criterion. The Akaike Information Criterion (AIC) is a commonly used information criterion. To avoid overfitting, the model with the smallest AIC value is generally preferred. After determining the parameters of the model, it is necessary to verify whether the model is appropriate. In this stage, we mainly test whether the residual sequence of the model is white noise. If the residual sequence is not white noise, it indicates that the currently constructed model does not fit all the valuable information in the sequence, and the parameters of the model need to be redetermined until the residual sequence is classified as white noise.

3. Construction of LSTM Prediction Model Optimized by Improved PSO

Although PSO has a significant effect in solving complex optimization problems, it lacks effective parameter control when dealing with optimization problems. There are problems such as slow convergence, easy to fall into local optimum, and low precision in the later iteration [27]. Through the analysis of particle swarm evolution formula in Section 2.2, this article improves the inertia weight and learning factor of PSO algorithm and uses the improved PSO algorithm to optimize the parameters of LSTM model to reduce the subjective influence of artificially selected parameters.

3.1. Improved PSO Algorithm

3.1.1. Improvement of Inertia Weight. The inertia weight w is mainly used to control the influence of the migration velocity on the current particle velocity, which is manifested as the performance of the PSO. The commonly used inertia weight assignment strategy is a linear decrement assignment; that is, the weight w decreases linearly with the number of iterations. Although this strategy can improve the performance of the PSO algorithm to a certain extent, as the number of iterations decreases linearly, the local search ability of the PSO will be worse. Therefore, in order to improve the overall optimization level of PSO, based on the previous research [28], the nonlinear decrement assignment method is adopted, as shown in the following equation:

$$w = w_{\max} - \left(w_{\max} - w_{\min}\right) * \sqrt{\frac{i}{\text{item\_max}}}, \qquad (7)$$

where $w_{\max}$ and $w_{\min}$ are the maximum inertia weight and the minimum inertia weight, respectively. i is the current iteration number and item_max is the maximum iteration number.

3.1.2. Improvement of Learning Factors. The learning factors c_1 and c_2 are mainly used to adjust the step size of the particle moving to the individual optimal position and the global optimal position. In practical applications, with the advance of the iteration process, it is usually required that the value of c_1 changes from large to small to accelerate the search speed in the early iteration and improve the global search ability. And the value of c_2 is required to be changed from small to large to facilitate the local refinement search at the later stage of the iteration and improve the accuracy, simultaneously. However, the standard PSO usually sets $c_1 = c_2 = 2$, which cannot meet the demands of practical applications. Therefore, the sine function is introduced to improve the learning factor, as shown in the following equation:

$$\begin{cases} c_1 = 2\sqrt{1 - \sin\left(\dfrac{\pi}{2} * \dfrac{i}{\text{item\_max}}\right)}, \\ c_2 = 2\sqrt{\sin\left(\dfrac{\pi}{2} * \dfrac{i}{\text{item\_max}}\right)}. \end{cases} \qquad (8)$$

3.2. Implementation Process of the LSTM Optimized by the Improved PSO. In this article, the LSTM neural network model optimized by the improved PSO algorithm is used to predict the price of nickel. The specific implementation process is shown in Figure 3.

The specific implementation steps of the forecast model of nickel price are as follows:

(1) Preprocess the sample data: to better fit the regularity contained in the time series and avoid the influence of the data on the gradient descent method, the nickel price data need to be smoothed and normalized. The specific calculation methods are shown in the following equations:

$$\Delta x_t = x_{t+1} - x_t, \tag{9}$$

$$y = \frac{\Delta x - \Delta x_{\min}}{\Delta x_{\max} - \Delta x_{\min}}, \tag{10}$$

where x_t and x_{t+1} are sample data corresponding to time t and time $t+1$, respectively, and Δx_t is a first-order difference corresponding to sample x_t. Δx is the sample data after the difference, $\Delta x_{\min}$ and $\Delta x_{\max}$ are the minimum and maximum values of the sample data after the difference, respectively, and y is the value obtained by normalizing the Δx.

(2) Initialize the parameters: determine the population size, particle dimension, the number of iterations, learning factor, inertia weight, and the defined interval of the parameter to be optimized.

(3) Initialize the particles: initialize the position and velocity of the particle, and randomly generate a particle $X_{i,0}$ (node1, node2, look_back), where node1 represents the number of neurons in the first hidden layer, node2 represents the number of neurons in the second hidden layer, and look_back represents the time step.

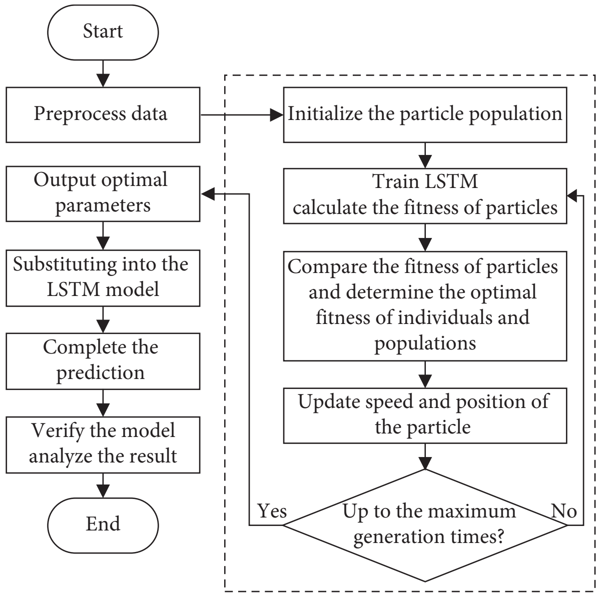

Start
Preprocess data
Output optimal parameters
Substituting into the LSTM model
Complete the prediction
Verify the model analyze the result
End
Initialize the particle population
Train LSTM calculate the fitness of particles
Compare the fitness of particles and determine the optimal fitness of individuals and populations
Update speed and position of the particle
Up to the maximum generation times?
Yes
No

FIGURE 3: The implementation process of the improved PSO-LSTM model.

(4) Set the fitness function of the particle: the sample data are divided into training data, verification data, and test data. In this article, the training data are input to the neural network for training. The root mean square error (RMSE) of the verification data of the LSTM model obtained after reaching the limit of the training times is selected as the individual fitness function, and the minimum fitness value is taken as the iterative target of the PSO algorithm. Then, the improved PSO was used to find the best parameters to be optimized to determine the optimal prediction model for nickel price. The fitness value (fitness) is calculated as follows:

$$\begin{aligned} \text{fitness} &= \text{RMSE}\,(\text{LSTM}\,(\text{node1}, \text{node2}, \text{look\_back})), \\ \text{RMSE}\,(\text{LSTM}\,(\text{node1}, \text{node2}, \text{look\_back})) &= \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(x_t - \widehat{x}_t\right)^2}, \end{aligned} \tag{11}$$

where N is the number of verification samples and x_t and $\widehat{x}_t$ are the real and fitted values of the verification sample, respectively.

(5) Update the speed and position of the particles: calculate the fitness value of each particle and determine the individual optimal fitness value and the group optimal fitness value. In the iterative process, the velocity and position of the particle are continuously updated according to the two optimal values.

(6) Complete the prediction and analyze the results: the parameter values obtained when the end condition of the improved PSO algorithm is satisfied are substituted into the LSTM neural network model, and then the test samples are input into the model for prediction, and finally, the prediction results are analyzed.

4. Case Analysis

In order to verify the effectiveness of the LSTM neural network model with improved PSO optimization, five prediction models are constructed based on the deep learning library Keras. They are as follows: (1) LSTM model with only one hidden layer optimized by standard PSO algorithm (PSO-LSTM11), (2) LSTM model with only one hidden layer optimized by improved PSO algorithm (PSO-LSTM12), (3) LSTM model with two hidden layers optimized by standard PSO algorithm (PSO-LSTM21), (4) LSTM model with two hidden layers optimized by improved PSO algorithm (PSO-LSTM22), and (5) conventional LSTM model. In addition, the autoregressive integrated moving average model (ARIMA) for processing time-series prediction is constructed as a control experiment. For the same sample data, the above six models are used to predict nickel prices.

4.1. Selection of Evaluation Indicators. To show the prediction effect of each model, the three indicators of root mean square error (RMSE), mean absolute deviation (MAE), and mean absolute percentage error (MAPE) are used to measure the performance of each model. RMSE is sensitive to the large deviation between the predicted value and the real value, which can reflect the accuracy of the prediction result well. MAE can avoid the problem that the positive and negative errors cancel each other out. MAPE considers the deviation between the predicted value and the actual value as well as the relationship between the error and the real value, which can better reflect the accuracy of the prediction result. The calculation formulas for the three indicators are as follows:

$$\begin{aligned} \mathrm{RMSE} &= \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(x_t - \hat{x}_t\right)^2}, \\ \mathrm{MAE} &= \frac{1}{N}\sum_{t=1}^{N}\left|x_t - \hat{x}_t\right|, \\ \mathrm{MAPE} &= \frac{100}{N}\sum_{t=1}^{N}\left|\frac{x_t - \hat{x}_t}{x_t}\right|, \end{aligned} \tag{12}$$

where N is the number of samples and x_t and $\hat{x}_t$ are the true and predicted values of the sample, respectively.

4.2. Data Preprocessing. As the largest nonferrous metals exchange in the world, London Metal Exchange's trading price is regarded as the benchmark of the world metal trade, which has an important impact on the production and sales of nonferrous metals in various countries. Therefore, this article uses the monthly average closing price of LME nickel on the London Metal Exchange from June 2006 to July 2018 as the research sample, as shown in Figure 4(a).

4.2.1. Data Preprocessing of the LSTM Model. From Figure 4(a), it can be found that the price fluctuation of nickel metal is frequent and extensive. In order to reduce the influence of other factors on the prediction results, the article makes a smooth treatment of nickel price according to equation (9). Then, this article normalizes the differential data to avoid numerical problems and make the network converge quickly, and the normalized form is shown in equation (10). The final preprocessed data are shown in Figure 4(b).

After data processing, the data set is divided into three parts. The data of 2006.06–2013.07 are used as training data, the data of 2017.08–2015.12 are used as verification data, and the data of 2016.01–2018.07 are used as test data.

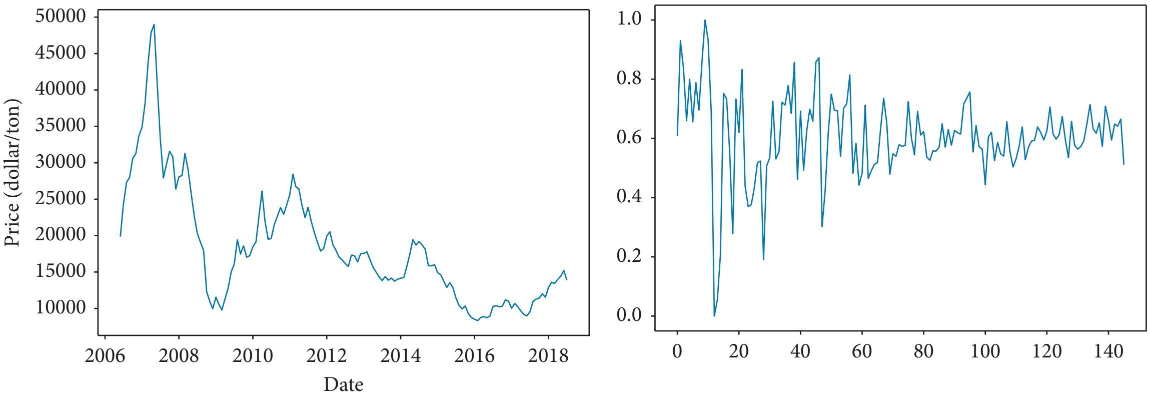
4.2.2. Data Preprocessing of the ARIMA Model. Since the validation set is not needed when constructing an ARIMA model for prediction, the data set is divided into two parts. In order to ensure the fairness of the model performance comparison, the data of 2006.06–2015.12 are used as training data and the data of 2016.01–2018.07 are used as test data.

Since the ARIMA model requires the sequence to be stable, the article firstly performs the ADF test on the nickel metal price series. As shown in Table 1, it is found that the p-value is much larger than the significance level of 0.05, so the sequence is considered to have a unit root and is nonstationary. The original sequence needs to be smoothed. In addition, the fluctuation of nickel metal price is large, so in order to alleviate the influence of heteroscedasticity on the model, this article performs the logarithmic operation on the original sequence and performs first-order differential processing on the logarithmic sequence. The ADF test is then performed on the first-order difference sequence, as shown in Table 1. From Table 1, it is found that the p-value of the first-order difference sequence is much smaller than the significance level of 0.05, so the first-order difference sequence is considered to be a stationary sequence. Therefore, the first-order difference sequence of the LME nickel metal price is selected for analysis.

To verify the applicability of the model, the Ljung–Box test is performed on the first-order difference sequence. The Ljung–Box test is a test of pure randomness to verify whether the sequence is a white noise sequence. The results of the Ljung–Box test are shown in Table 2.

From Table 2, it can be found that the p-value of the Q statistic are far less than the significance level of 0.05, so the first-order difference sequence is a nonwhite noise sequence. In summary, the ARIMA model can be constructed to analyze the sequence.

4.3. Prediction of Nickel Prices Based on LSTM. In this article, the adaptive moment estimation (Adam) algorithm is used as the optimization algorithm to train the internal parameters of LSTM. To prevent the overfitting phenomenon in the LSTM training process, use dropout for regularization and set dropout = 0.1. The number of LSTM training is 50. The parameters of the PSO are set as follows: the population number is 20, the number of iterations is 50, the minimum value of the parameter to be optimized is 1, and the maximum value is 15. And the learning factor of the standard



(a)

(b)

Figure 4: The (a) original sequence and (b) processed sequence of nickel metal monthly price.

Table 1: The results of two ADF tests.

Test item	Test result of original sequence	Test result of processed sequence
Test statistic value	−2.2601036330998725	−8.7048239142842423
p-value	0.18515824547409387	0.00000000000003725
Critical value (1%)	−3.4786478891750301	−3.4765979175374011
Critical value (5%)	−2.8827217656441682	−2.8818291230495543
Critical value (10%)	−2.5780653266120561	−2.5775887982253085

Table 2: The Ljung–Box test results of first-order difference sequences.

Lag	AC	Q	p-value
1.0	0.335056	13.137734	0.000289
2.0	0.137655	15.375058	0.000459
3.0	−0.047252	15.641057	0.001343
4.0	0.094434	16.713127	0.002197
5.0	0.134068	18.893789	0.002012
6.0	0.033009	19.027203	0.004118
7.0	−0.080867	19.835415	0.005936
8.0	−0.115688	21.505096	0.005920
9.0	−0.074676	22.207423	0.008244
10.0	−0.000595	22.207468	0.014082

Table 3: The display of partial LSTM prediction models.

Model	Structure	RMSE verification
LSTM1	(look_back, node, layer) = (10, 3, 1)	788.60
LSTM2	(look_back, node, layer) = (15, 3, 1)	617.53
LSTM3	(look_back, node, layer) = (15, 11, 1)	973.40
LSTM4	(look_back, node1, node2, layer) = (15, 4, 11, 2)	624.12

PSO is set to $c_1 = c_2 = 2$, and the inertia weight is set to $w = 0.6$; the maximum inertia weight of the improved PSO is set to 0.8, and the minimum inertia weight is set to 0.2.

4.3.1. Prediction of Nickel Prices Based on Conventional LSTM. In order to determine the parameters of the model when constructing the LSTM model for prediction, 1000 groups of integers in the range of [1, 15] are randomly generated as the value of the time step and the number of nodes in each hidden layer for the LSTM model with only one LSTM layer and two LSTM layers. Then, the LSTM models corresponding to different parameter combinations are used for nickel metal price prediction, respectively, and the RMSE of the verification set of each model is compared. Some results are shown in Table 3.

In Table 3, look_back is the time step, node is the number of nodes contained in the hidden layer in the LSTM model with a single hidden layer, node1 and node2 are, respectively, the numbers of nodes included in each hidden layer in the LSTM model with two hidden layers, and the layer is the number of hidden layers of the LSTM model. Observing the data in Table 3, it can be found that the setting of the numbers of hidden layers, the number of nodes included in the hidden layer, and the time step have a deep influence on the fitting effect of the LSTM. The model LSTM2 (look_back = 15, node = 3, layer = 1) with the smallest RMSE of the verification set is selected as the representative of the traditional LSTM model in the comparative experiment.

4.3.2. Prediction of Nickel Prices Based on the LSTM Model Optimized by Improved PSO. The standard PSO algorithm and the improved PSO algorithm are used to optimize the LSTM model with one hidden layer and two hidden layers, respectively. In the process of optimizing the LSTM by PSO, the change of the fitness value is shown in Figure 5.

From Figure 5, fitness11, fitness21, fitness12, and fitness22 are the fitness values corresponding to the models

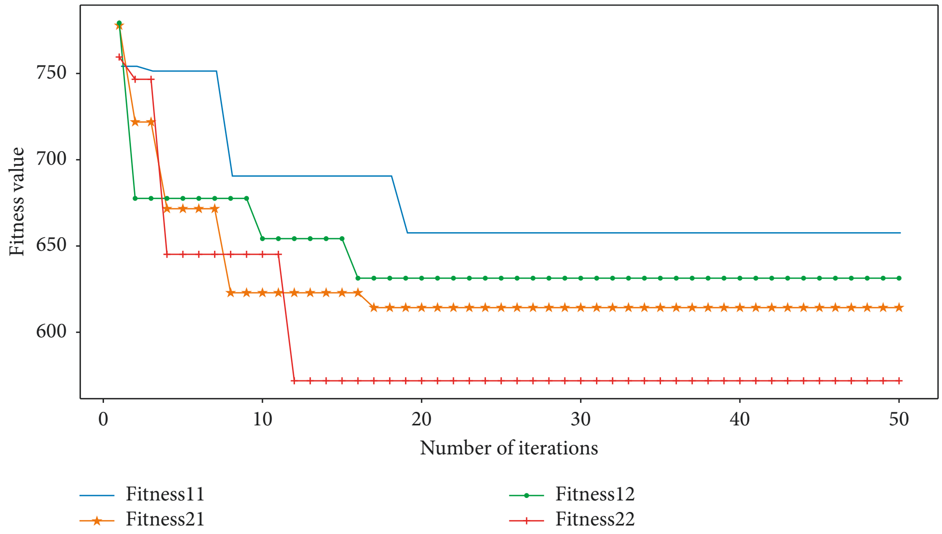


Figure 5: The change of fitness value-1.

PSO-LSTM11, PSO-LSTM21, PSO-LSTM12, and PSO-LSTM22, respectively. The following can be seen from Figure 5:

(1) The final convergence value of fitness21 is smaller than that of fitness11, and its convergence speed is faster than that of fitness11. Fitness22's final convergence value is smaller than that of fitness12, and its convergence speed is faster than that of fitness12. It indicates that the fitness value of the LSTM with two hidden layers optimized by PSO is smaller than that of the LSTM with one hidden layer optimized by PSO, and the convergence speed is faster.

(2) The final convergence value of fitness12 is smaller than that of fitness11, and the convergence speed is faster than that of fitness11. The final convergence value of fitness22 is smaller than that of fitness21, and the convergence speed is faster than that of fitness21. It shows that the fitness value of LSTM optimized by improved PSO is smaller than that of LSTM optimized by standard PSO, and the convergence speed of improved PSO is faster than that of standard PSO.

From the above, it can be seen that the improved PSO algorithm can effectively improve the prediction accuracy of the LSTM neural network and the convergence speed of the PSO algorithm. The prediction effect of the LSTM model optimized by the improved PSO algorithm is significantly better than that of the LSTM model optimized by standard PSO, and the prediction effect of the PSO-LSTM22 is the best.

To verify the rationality and effectiveness of the LSTM optimized by the improved PSO algorithm in this article, the following control experiments were set up: (1) the LSTM model with two hidden layers optimized by PSO with improved inertia weight w according to equation (7) (PSO-LSTM-W) and (2) the LSTM model with two hidden layers optimized by PSO with improved inertia weight w according to equation (7) and learning factors according to the literature [28] (PSO-LSTM-WC). The changes in fitness values for each model during the iterative process are shown in Figure 6.

Fitness21, fitness22, fitnessW, and fitnessWC are fitness values corresponding to the models PSO-LSTM21, PSO-LSTM22, PSO-LSTM-W, and PSO-LSTM-WC, respectively. The following can be seen from Figure 6:

(1) The convergence speed of fitnessW is faster than that of fitness21, but the final convergence values of the two are close. It shows that the improved inertia weight method shown in equation (7) can accelerate the convergence speed of PSO, but the accuracy is not improved.

(2) The final convergence value of fitnessWC is smaller than that of fitness21 and fitnessW, and the convergence speed is faster than that of fitness21 and fitnessW. It shows that the improved method of PSO-LSTM-WC model can effectively improve the search accuracy and speed of the PSO algorithm.

(3) The final convergence value of fitness22 is significantly smaller than that of fitnessWC, and the convergence speed is slightly faster than that of fitnessWC. It indicates that the improved method of PSO-LSTM22 model can significantly improve the convergence speed of the PSO algorithm and the prediction accuracy of the LSTM model, and its performance is better than that of PSO-LSTM-WC model.

In conclusion, the PSO-LSTM22 nickel price prediction model proposed in this article is reasonable and efficient.

In addition, to present the optimal parameter values of the LSTM model determined by the improved PSO algorithm, Figures 7 and 8, respectively, show the changes in the number of the nodes and the time step in the PSO-LSTM22 model's optimization process.

It can be seen from Figures 7 and 8 that the optimal parameter of the PSO-LSTM22 model is set to node1 = 4, node2 = 2, and look_kack = 15.

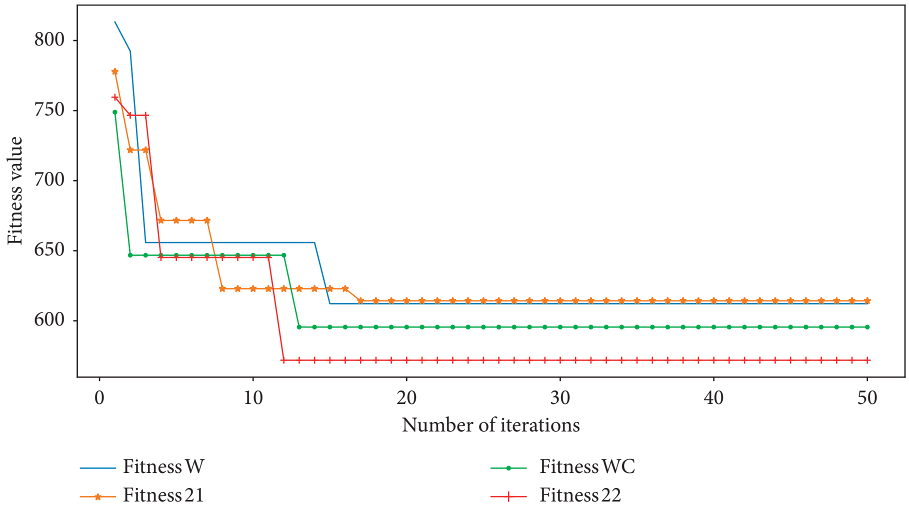


Figure 6: The change of fitness value-2.

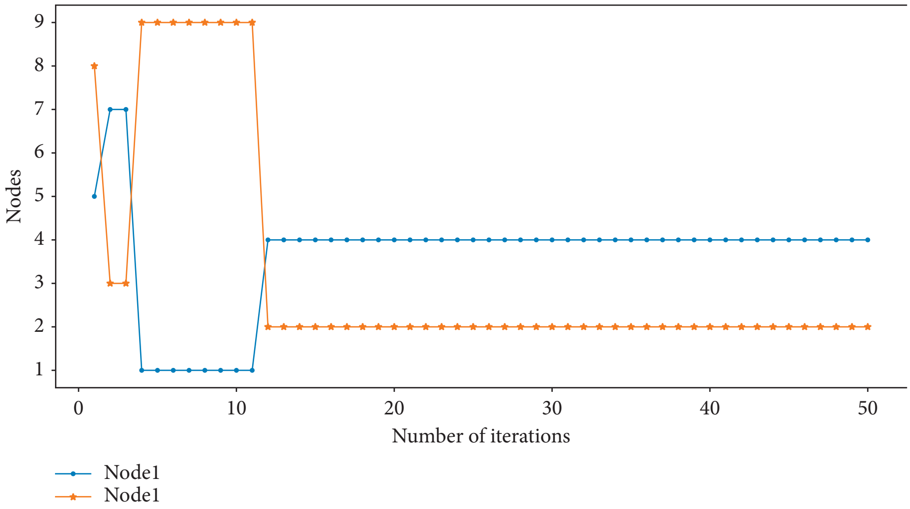


Figure 7: The change in the node number in the hidden layer.

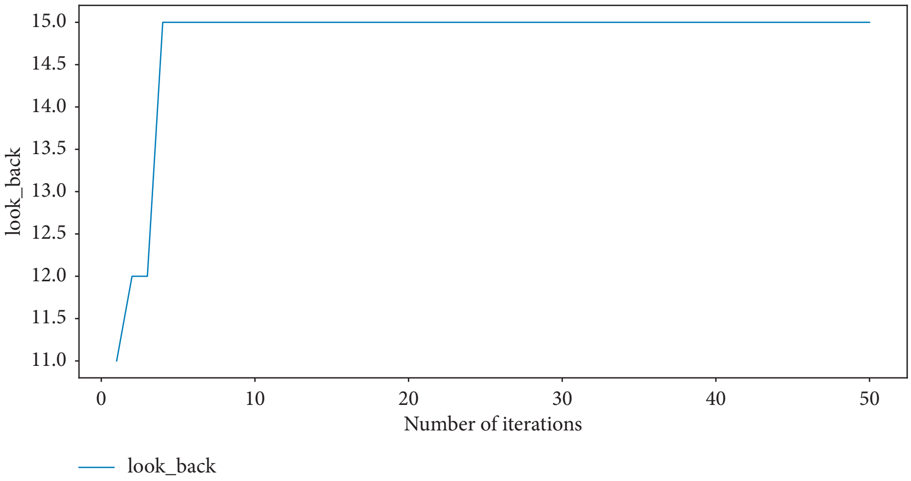


Figure 8: The change in time step.

4.4. Prediction of Nickel Prices Based on ARIMA. In order to determine the order of the ARIMA model, this article first draws the first-order difference sequence and its autocorrelation graph and partial autocorrelation graph, as shown in Figure 9. Then, the order range of the model was identified by observing the autocorrelation graph and the partial autocorrelation graph, and then the p and q of the model were determined according to the AIC criterion.

When determining the order of the model according to the AIC criterion, the order range of p and q are selected as [0, 5], and the heat map corresponding to the AIC value of each model in the range is drawn, as shown in Figure 10.

Five models were selected by observing the heat map and the autocorrelation and partial autocorrelation plot of the first-order difference sequence in the original manuscript: ARIMA(3, 1, 2), ARIMA(3, 1, 3), ARIMA(0, 1, 2), ARIMA(1, 1, 2), and ARIMA(1, 1, 1) to predict nickel metal prices, respectively.

After determining the order of the model, in order to verify the rationality of the model, this article takes the ARIMA(1, 1, 2) model as an example and carries out the Ljung–Box test on the residual sequence. The results are shown in Table 4. Meanwhile, the autocorrelation graph, partial autocorrelation graph, and QQ graph of the residual sequence are drawn, as shown in Figure 11.

From Table 4, it can be found that the p-value of the Q statistic is far greater than the significance level of 0.05, so the residual sequence is considered to be a white noise sequence. It can be seen from Figure 11 that the autocorrelation coefficient and the partial autocorrelation coefficient are mostly within the confidence interval, so the residual is considered to not correlate. In addition, the residuals in the QQ graph are basically concentrated on a straight line, indicating that the residuals obey the normal distribution. It can be considered that the model has a good fitting effect.

Then, in order to test whether the residual is heteroscedastic, an autocorrelation test on the residual square sequence was performed. The results are shown in Table 5.

From Table 5, it can be found that the p-value of lag 1–12 are all greater than the significance level of 0.05, so it can be considered that the residual sequence does not have autocorrelation; that is, the residual sequence does not have the ARCH effect. In summary, it is reasonable to construct an ARIMA(1, 1, 2) model to analyze the sequence.

The verification processes of the other 4 models are the same as above. Finally, the errors corresponding to the test sets of the five models are compared, as shown in Table 6.

From Table 6, it can be found that the RMSE, MAE, and MAPE of the model ARIMA(1, 1, 2) are smaller than those of the other four models, and the prediction effect is better. So the ARIMA(1, 1, 2) model is finally constructed to predict the price of nickel metal.

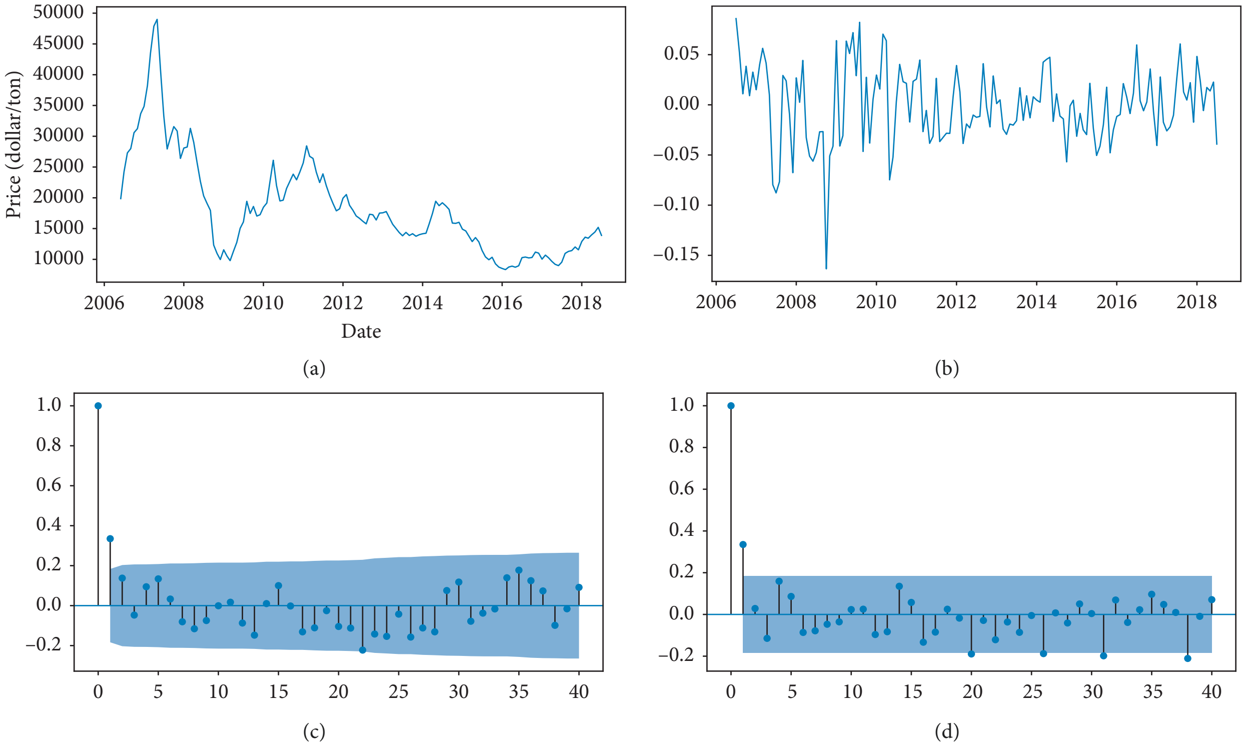
5. Result Analysis

To evaluate the prediction performance of the improved LSTM model optimized by the PSO algorithm, the test samples are used for verification. In this article, the LSTM model optimized by the standard PSO algorithm, the conventional LSTM model, and the ARIMA model are compared with the model. Since the time steps of the five LSTM models all have a value of 15, Figure 12 shows the relative error of each model for the prediction results of the last 15 test data. In addition, the maximum and minimum values of the relative errors are listed in Table 7.

As can be seen from Table 7, the maximum and minimum relative errors of the PSO-LSTM12 model are reduced by 1% and 28%, respectively, compared with the PSO-LSTM11 model. Moreover, the maximum and minimum relative errors of the PSO-LSTM22 model are reduced by 1% and 44%, respectively, compared with the PSO-LSTM21 model. It shows that, compared with the standard PSO algorithm, the improved PSO algorithm can effectively improve the prediction accuracy of the LSTM model. The maximum and minimum relative errors of the LSTM model optimized by PSO are lower than those of the conventional LSTM model and the ARIMA model, indicating that using the PSO algorithm to optimize the LSTM model can effectively improve the prediction accuracy of the LSTM model.

In order to more intuitively display the prediction effects of each model, the predicted performance evaluation index values of each model are calculated, as shown in Table 8.

From Table 8, it can be found that compared with the PSO-LSTM21 model, the RMSE, MAE, and MAPE of the PSO-LSTM22 model are reduced by 3%, 7%, and 6%, respectively. And the RMSE, MAE, and MAPE of the PSO-LSTM12 model are all smaller than those of PSO-LSTM11. It shows that compared with the standard PSO algorithm, using the improved PSO algorithm to optimize the LSTM model can improve the accuracy of the optimization results. Comparing the prediction error of PSO-LSTM11 and PSO-LSTM21, PSO-LSTM12, and PSO-LSTM22, it can be found that the prediction effect of the LSTM with two hidden layers is better than that of the LSTM with one hidden layer. Comparing the prediction error of the PSO-LSTM series model and the standard LSTM model, it can be found that the evaluation index values of the PSO-LSTM model are smaller than those of the LSTM model. In particular, compared with the LSTM model, the RMSE, MAE, and MAPE of the PSO-LSTM22 model are reduced by 9%, 9%, and 10%, respectively. It shows that the PSO algorithm can effectively improve the prediction performance of the LSTM model. The evaluation index values of the PSO-LSTM model and the LSTM model are smaller than those of the ARIMA model, which indicates that the prediction effect of the LSTM model is better than that of the shallow structure model, and the LSTM model is more suitable for dealing with complex nonlinear problems. Also, compared with the ARIMA model, the accuracy of the PSO-LSTM22 model is improved by 13%. In conclusion, the LSTM optimized by the improved PSO constructed in this article has better prediction effect and higher reliability.



(a)

(b)

(c)

(d)

Figure 9: The (a) original data and (b) processed data nickel price and (c) autocorrelation graph and (d) partial autocorrelation graph of the processed sequence.

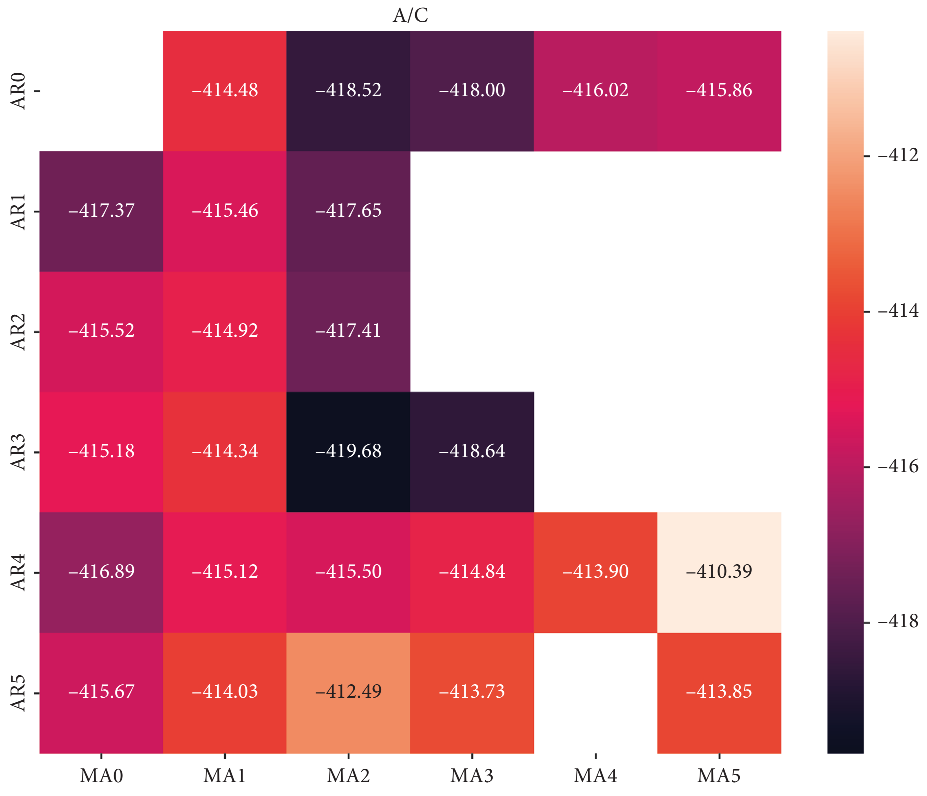
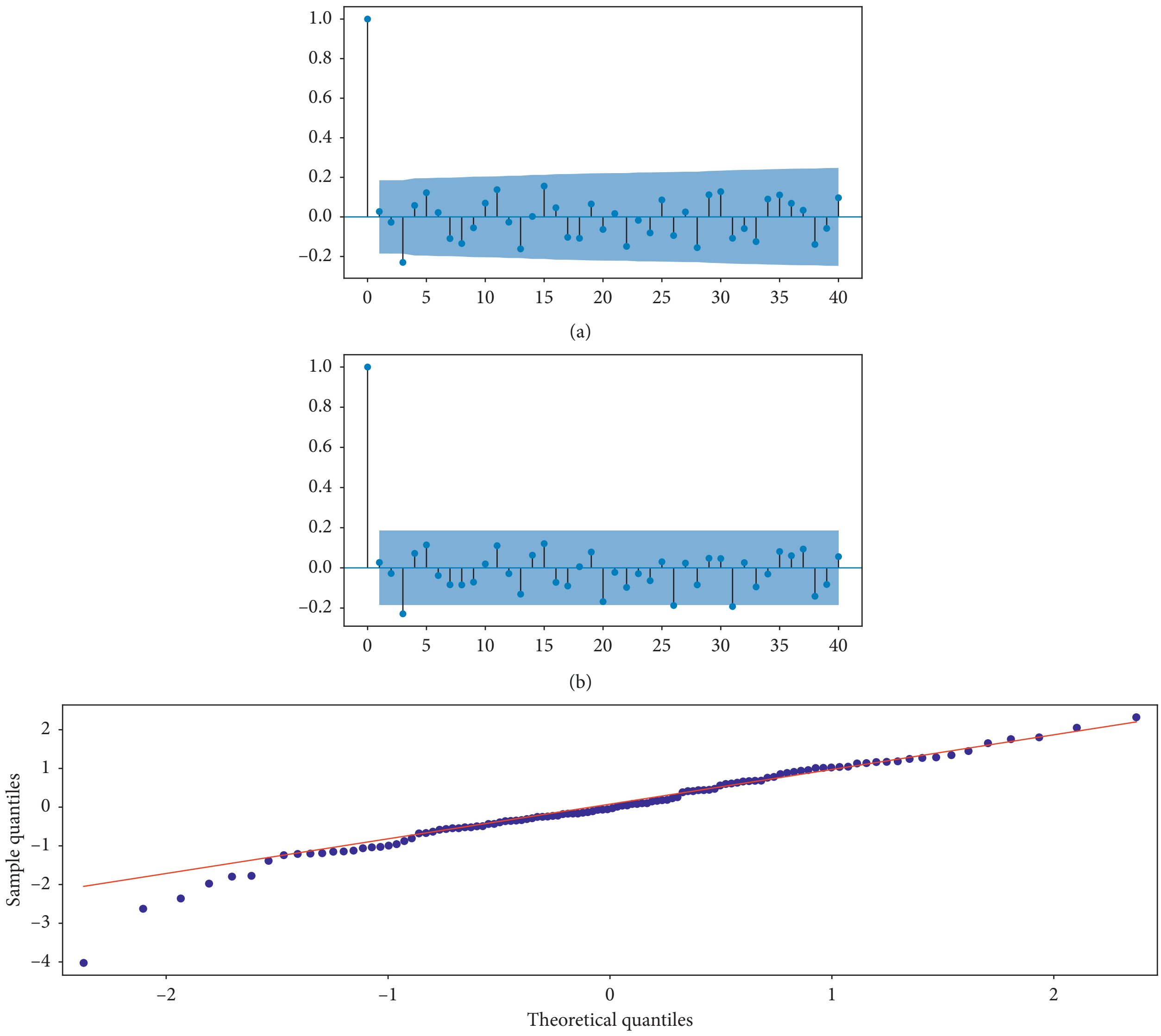


Figure 10: The heat map corresponding to the AIC values of each model.

Table 4: The Ljung–Box test results of residual sequence of ARIMA(1, 1, 2).

Lag	AC	Q	p-value
1.0	0.026972	0.083684	0.772366
2.0	−0.027064	0.168704	0.919108
3.0	−0.229964	6.363354	0.095210
4.0	0.057894	6.759599	0.149151
5.0	0.122480	8.549676	0.128431
6.0	0.022287	8.609507	0.196759
7.0	−0.109559	10.069091	0.184692
8.0	−0.134496	12.289901	0.138730
9.0	−0.055056	12.665653	0.178328
10.0	0.069653	13.272957	0.208808



(a)

(b)

(c)

Figure 11: The (a) autocorrelation graph, (b) partial autocorrelation graph, and (c) QQ graph of the residual sequence.

Table 5: The Ljung–Box test results of the residual square sequence of the ARIMA(1, 1, 2).

Lag	AC	Q	p-value
1.0	0.001880	0.000407	0.983911
2.0	−0.013717	0.022247	0.988938
3.0	0.240741	6.811112	0.078168
4.0	0.060478	7.243522	0.123565
5.0	0.006430	7.248456	0.202808
6.0	0.191994	11.688549	0.069289
7.0	0.073203	12.340164	0.089913
8.0	0.060502	12.78956	0.119298
9.0	0.048683	13.083349	0.158871
10.0	0.184445	17.341826	0.067134
11.0	0.111246	18.906318	0.062796
12.0	−0.011588	18.923464	0.090392

Table 6: Prediction errors of the five models.

	ARIMA(1, 1, 2)	ARIMA(1, 1, 1)	ARIMA(3, 1, 2)	ARIMA(0, 1, 2)	ARIMA(3, 1, 3)
RMSE	633.96	636.84	640.12	640.37	644.39
MAE	508.45	510.32	508.59	511.50	524.97
MAPE (%)	4.61	4.63	4.61	4.63	4.80

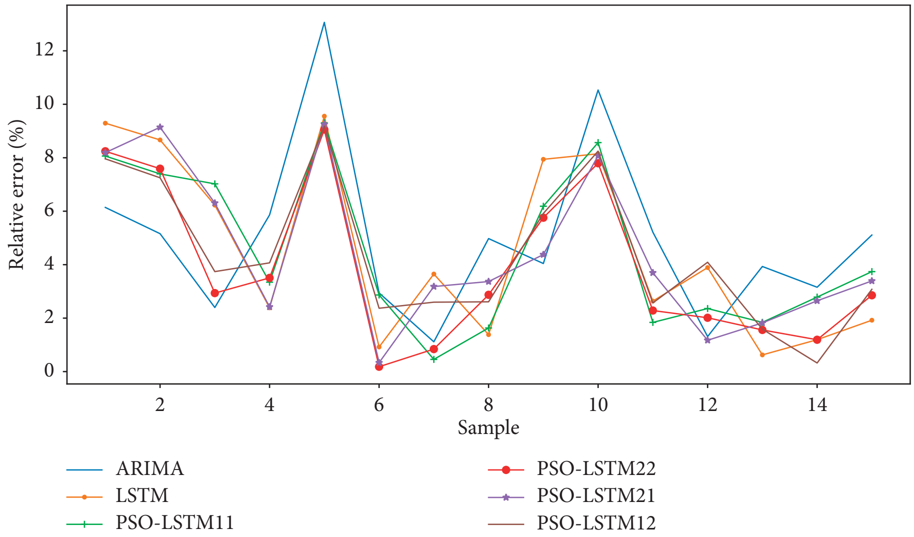


Figure 12: The comparison of prediction errors of the six models.

Table 7: The comparison of relative errors of the six models.

	PSO-LSTM11 (%)	PSO-LSTM12 (%)	PSO-LSTM21 (%)	PSO-LSTM22 (%)	LSTM (%)	ARIMA (%)
Maximum value	9.31	9.25	9.18	9.05	9.55	13.07
Minimum value	0.46	0.33	0.32	0.18	0.62	0.67

Table 8: Comparison of prediction error index values of six models.

	PSO-LSTM11	PSO-LSTM12	PSO-LSTM21	PSO-LSTM22	LSTM	ARIMA
RMSE	588.97	584.72	569.50	553.89	610.32	633.96
MAE	507.76	507.26	495.12	460.03	507.85	508.45
MAPE (%)	4.49	4.49	4.37	4.11	4.56	4.61

6. Conclusions

In this article, PSO is improved by nonlinear decrement assignment and sine function, and the improved PSO algorithm is used to optimize the parameters of LSTM. Then, the LSTM neural network model optimized by the improved PSO algorithm is applied to LME nickel metal price prediction and compared with the standard PSO-optimized

LSTM model, conventional LSTM model, and ARIMA time-series prediction model with higher prediction accuracy. Empirical results show that the LSTM neural network model optimized by the improved PSO algorithm can effectively complement the defects of the standard PSO algorithm with poor robustness and solve the problems that are difficult to be determined by LSTM network structure. Furthermore, it can effectively improve the convergence speed of PSO and the prediction accuracy of the LSTM model.

This article uses the improved LSTM neural network to predict the price of nickel metal, which has excellent prediction performance and strong reliability and can provide technical support for the establishment of a good nickel price prediction mechanism. In addition, effective price prediction facilitates nickel market participants to deeply dig into the law of nickel price fluctuations, timely grasp the trend of nickel market price fluctuations, and make reasonable production, purchase, and investment plans in advance, so as to better deal with the impact of nickel price fluctuations.

This article studies the medium-term forecast of nickel metal price, and it is necessary to further verify whether the model constructed here has the same sound effect on short-term nickel metal price forecast. In addition, based on the existing research, further research is in need about how to comprehensively grasp the influencing factors of nickel metal price and accurately measure the contribution rate of each influencing factor in nickel metal price fluctuation to improve the prediction accuracy.

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 that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Many ideas on the paper are suggested by B. S. to support the work, M. L. contributed to model establishing and paper writing, Y. Z. analyzed the data, and G. B. reviewed the work and modified the article. In general, all authors cooperated as much as possible during all the progress of the research.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61672416).

References

[1] https://max.book118.com/html/2018/0603/170477941.shtm.

[2] Y. Wen-Jing, "Influence factors on recent nickel price increase," *China Titanium Industry*, vol. 14, no. 2, pp. 48-49, 2017.

[3] X. Ai-Dong, "The impact of policy heavy punches on China's nickel market," *China Metal Bulletin*, vol. 19, p. 17, 2005.

[4] S. Yousefi, I. Weinreich, and D. Reinarz, "Wavelet-based prediction of oil prices," *Chaos, Solitons & Fractals*, vol. 25, no. 2, pp. 265–275, 2005.

[5] T. Trainer, "Some factors that would affect the retail price for 100% Australian renewable electricity," *Energy Policy*, vol. 116, pp. 165–169, 2018.

[6] T. Kriechbaumer, A. Angus, D. Parsons, and M. Rivas Casado, "An improved wavelet—ARIMA approach for forecasting metal prices," *Resources Policy*, vol. 39, no. 1, pp. 32–41, 2014.

[7] L. U. Lei-Rong, T. Xian-Zhi, and X. U. Yong-Ge, "Research on prediction of molybdenum price based on cointegration theory," *Rare Metals and Cemented Carbides*, vol. 44, no. 3, pp. 67–72, 2016.

[8] O. Isengildina, S. H. Irwin, and D. L. Good, "Evaluation of USDA interval forecasts of corn and soybean prices," *American Journal of Agricultural Economics*, vol. 86, no. 4, pp. 990–1004, 2004.

[9] A. Trujillo-Barrera, P. Garcia, and M. L. Mallory, "Short-term price density forecasts in the lean HOG futures market," *European Review of Agricultural Economics*, vol. 45, no. 1, pp. 121–142, 2018.

[10] J. Xian-Ling and W. Bing-Nan, "Research on the characteristics and influencing factors of international gold price volatility—analysis from a periodic perspective," *Price Theory and Practice*, vol. 12, pp. 98–101, 2017.

[11] G. W. Crawford and M. C. Fratantoni, "Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices," *Real Estate Economics*, vol. 31, no. 2, pp. 223–243, 2003.

[12] S. Bi-Lin and L. Yuan-Xin, "Research on method of Mo price forecasting based on MLP and RBF neural network," *China Molybdenum Industry*, vol. 40, no. 5, pp. 54–60, 2016.

[13] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models," *Expert Systems with Applications*, vol. 103, pp. 25–37, 2018.

[14] C. Min, "Overview and outlook of the nickel market in September 2014," *Non-ferrous Metal Engineering*, vol. 4, no. 5, pp. 11-12, 2014.

[15] H. Geman and W. O. Smith, "Theory of storage, inventory and volatility in the LME base metals," *Resources Policy*, vol. 38, no. 1, pp. 18–28, 2013.

[16] D. Xiang-Quan, *Research on Nickel Price Forecasting and Purchasing Model and Its Application*, Northeastern University, Boston, MA, USA, 2015.

[17] N. Shan-Qin, H. Song-Yi, W. Hong-Hai, Y. Wen-Jia, C. Qi-Shen, and L. Ting, "Analysis on the main influencing factors and future trend of nickel price," *Resource Science*, vol. 37, no. 5, pp. 961–968, 2015.

[18] X. Li, L. Peng, X. Yao et al., "Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation," *Environmental Pollution*, vol. 231, no. 1, pp. 997–1004, 2017.

[19] M. Cai and J. Liu, "Maxout neurons for deep convolutional and LSTM neural networks in speech recognition," *Speech Communication*, vol. 77, pp. 53–64, 2016.

[20] J. Feng and P. Zi-Jun, "Prediction of carbon price based on chaotic PSO optimized BP neural network," *Statistics and Information Forum*, vol. 33, no. 5, pp. 93–98, 2018.

[21] J. P. S. Catalao, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-PSO-ANFIS approach for short-term

electricity prices forecasting," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 137–144, 2011.

[22] L. Wan, F. Fen-Ling, and J. Qi-Wei, "Prediction of railway passenger traffic volume based on LSTM neural network optimized by improved particle swarm optimization algorithm," *Journal of Railway Science and Engineering*, vol. 15, no. 12, pp. 3274–3280, 2018.

[23] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[24] M. Song, X. Zhao, Y. Liu, and Z. Zhao, "Text sentiment analysis based on convolutional neural network and bidirectional LSTM model," in *Proceedings of the International Conference of Pioneering Computer Scientists, Engineers and Educators*, vol. 902, pp. 55–68, Changsha, China, September 2018.

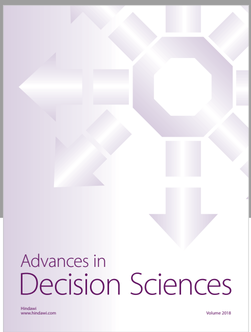
[25] J. Y. Choi and B. Lee, "Combining LSTM network ensemble via adaptive weighting for improved time series forecasting," *Mathematical Problems in Engineering*, vol. 2018, Article ID 2470171, 8 pages, 2018.

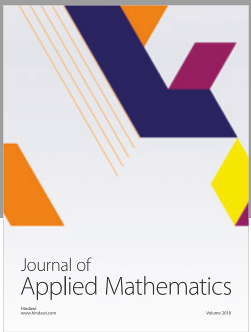
[26] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the International Conference on Neural Networks*, pp. 1942–1948, Anchorage, AK, USA, August 2002.

[27] M. Guo-Qing, L. Rui-Feng, and L. Li, "Particle swarm optimization algorithm of learning factors and time factor adjusting to weights," *Application Research of Computers*, vol. 31, no. 11, pp. 3291–3294, 2014.

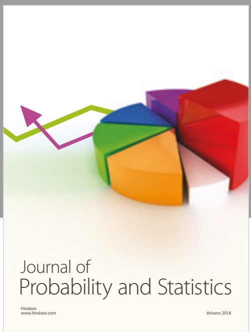
[28] L. Zheng-Shan, W. Wen-Hui, W. Xiao-Wan, and Z. Xin-Sheng, "Soil corrosion prediction of buried pipeline based on the model of RS-PSO-GRNN," *Materials Protection*, vol. 51, no. 8, pp. 47–79, 2018.

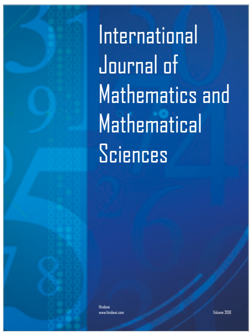
Advances in
Operations Research


Advances in
Decision Sciences

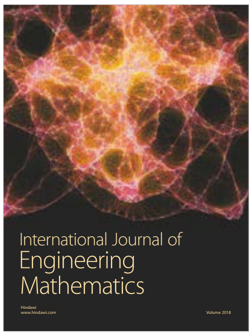

Journal of
Applied Mathematics

The Scientific
World Journal

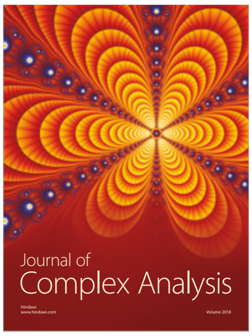

Journal of
Probability and Statistics

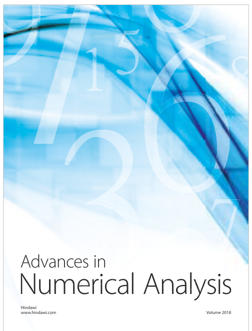

International
Journal of
Mathematics and
Mathematical
Sciences

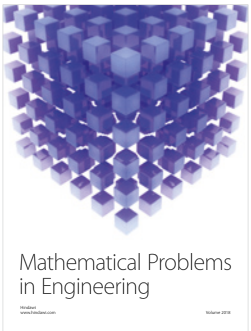
Journal of
Optimization

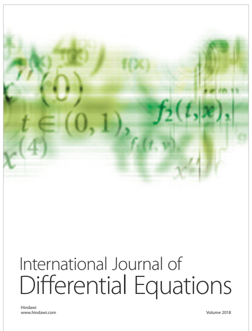

International Journal of
Engineering
Mathematics

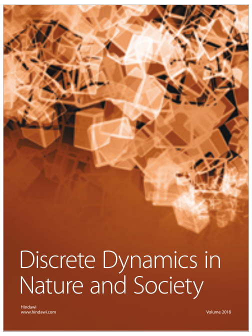
International Journal of
Analysis

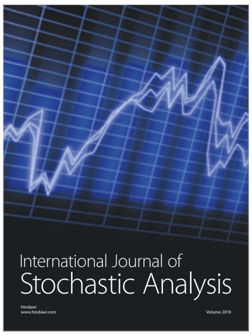

Journal of
Complex Analysis


Advances in
Numerical Analysis


Mathematical Problems
in Engineering

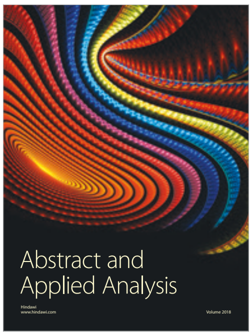

International Journal of
Differential Equations

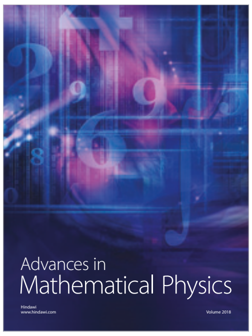

Discrete Dynamics in
Nature and Society


International Journal of
Stochastic Analysis

Journal of
Mathematics

Journal of
Function Spaces


Abstract and
Applied Analysis


Advances in
Mathematical Physics