ISSN 2616-5902 Vol. 4, Issue 11: 33-38, DOI: 10.25236/AJBM.2022.041106

Carbon Trading Price Forecast Based on LSTM and ARIMA—Take the Shanghai Area for Example

Shurong Dong[1], Xinyue Zhang[2]

[1]College of Textile and Clothing Engineering, Soochow Univeisity, Suzhou, Jiangsu, 215325, China
[2]College of Information Science and Technology, Bohai University, Jinzhou, Liaoning, 121010, China

Abstract: *In order to solve environmental problems such as air pollution, China has tried to carry out carbon emission trading market. For the sake of improving the risk management ability of the carbon emission trading market and make reasonable policies, it is very important to analyze and forecast the price of the carbon emission trading market. Considering that a single model has many self-limitations, in order to ensure that the predicted results can better match the actual situation, this paper adopts the combination model of ARIMA and LSTM to study the carbon emission trading price. Based on the reasonable prediction of carbon emission trading price of Shanghai Stock Exchange in recent years, the results show that ARIMA and LSTM can well combine their mutual advantages, effectively improve the accuracy of prediction, and provide certain guarantee for the sound development of carbon emission trading market.*

Keywords: *Carbon emissions trading, ARIMA, LSTM, CRITIC*

1. Introduction

Considering the late development of China's carbon emission trading market, pilot carbon emission trading markets were first launched in Beijing, Shanghai, Shenzhen and other places. With the rise of the price of carbon emission rights, China's carbon emission trading market is constantly improving, which also means that enterprises need to formulate carbon emission trading strategies scientifically.

We need to find a way to effectively predict the price of carbon emissions trading, because the price of carbon emissions trading is an important basis for formulating the policy of carbon emissions trading market, but also to improve the capacity of carbon emission market risk management.

In recent years, the research methods of market trading prices such as carbon emissions have undergone extensive evolution, from statistical model to intelligent model, from single model to combined model. Statistical models are represented by the differential autoregressive moving average (ARIMA) model [1] and the generalized autoregressive conditional heteroscedasticity (GARCH) model [2], which are widely used. However, the above model needs to check and pre-process the original data, which may lead to information loss when using big data samples. Intelligent models represented by artificial neural network (ANN) are good at dealing with incomplete, fuzzy, uncertain or irregular data and have a good fit for nonlinear relations [3]. Zhang (2003) [4] once discussed a combination model of ARIMA and neural network, and based on this idea, people built a combination model for prediction, and confirmed that the combination model is more advantageous [5].

Considering the volatility and uncertainty of the carbon emission trading market, this paper decided to use ARIMA model to fit the linear part of the carbon emission trading price, and use LSTM model to fit the nonlinear part, and use the combination of CRITIC to obtain the prediction results, in order to combine the advantages of the two models to get better results.

2. Approches

2.1 ARIMA

ARIMA model consists of autoregression (AR), difference (I) and moving average (MA).

ARIMA model has three important data indicators, which are P, Q and D respectively, where P is the number of autoregression items. Q is the number of terms in the moving average, and D is the

ISSN 2616-5902 Vol. 4, Issue 11: 33-38, DOI: 10.25236/AJBM.2022.041106

difference fraction that makes it a stationary series. To obtain accurate prediction results, it is necessary to ensure that the time series used for prediction is stable, and the unstable time series can be changed into stable time series through difference.

ARIMA (P, D, Q) can be obtained by combining autoregressive moving average model (AR), moving average model (MA) and difference method (I). We can express the ARIMA model as:

$$\left(1-\sum_{i=1}^{P}\phi_i L^i\right)(1-L)^d X_t=\left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \tag{1}$$

Where L is the hysteresis operator（Lag operator）, $d \in Z, d > 0$.

2.2 LSTM neural network model

Long Short-Term Memory Networks (LSTM) The purpose is to solve the problems of gradient explosion (too small weight after activation of input information) and gradient disappearance in general recurrent neural network. The concept of gate and Cell state is introduced to realize gradient adjustment.

At time T, the formula defined by LSTM neural network is as follows:

$$f_t = sigmoid\left(W_f \bullet \left[h_{t-1}, x_t\right] + b_f\right) \tag{2}$$

$$i_t = sigmoid\left(W_i \bullet \left[h_{t-1}, x_t\right] + b_i\right) \tag{3}$$

$$o_t = sigmoid\left(W_o \bullet \left[h_{t-1}, x_t\right] + b_o\right) \tag{4}$$

$$\tilde{c}_t = \tanh\left(W_c \bullet \left[h_{t-1}, x_t\right] + b_c\right) \tag{5}$$

$$c_t = f_t \times c_{t-1} + i_t * \tilde{c}_t \tag{6}$$

$$h_t = o_t \times \tanh\left(c_t\right) \tag{7}$$

Sigmoid and TANH are two activation functions except that i_t, f_t, o_t and c_t, W_* mentioned above respectively represent the recursive connection weights of their corresponding thresholds. Figure 1 shows the cell structure of the hidden layer. When training the LSTM neural network, the model will first input the data features to the input layer. At this time, the time of the input data is t, and then the model outputs the result through the excitation function. Then the nodes of the LSTM structure will receive the result of Output, the Output of the hidden layer at t-1, and the information stored in the cell at t-1, and then output the data to the next hidden layer or Output layer through model processing. Finally, the result of the LSTM structure node is output to the neurons of the output layer, the back-propagation error is calculated, and each weight is updated. The training order of the whole model is input gate, output gate, forget gate and unit cell.

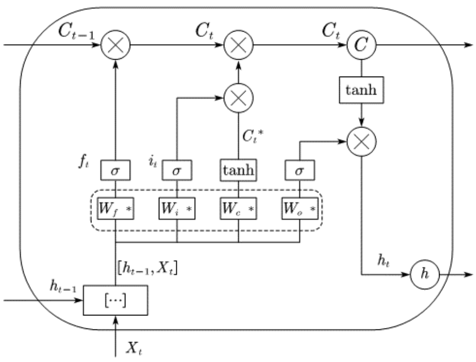

Figure 1: LSTM detail diagram

ISSN 2616-5902 Vol. 4, Issue 11: 33-38, DOI: 10.25236/AJBM.2022.041106

2.3 CRITIC combination model

The algorithm flow chart of ARIMA-LSTM combined model constructed in this paper is shown in Figure 2.

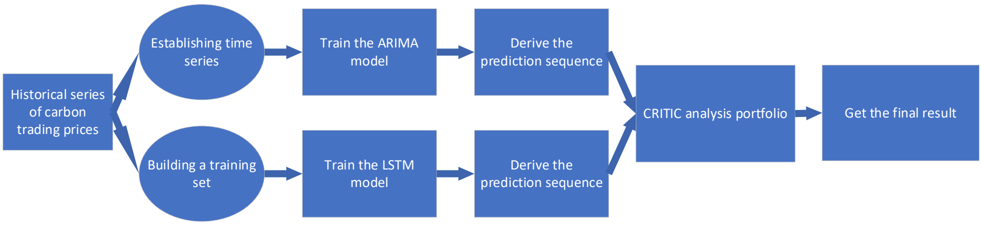

Figure 2: Composite model flow chart

Critical method is an objective weight method. It uses two basic methods to determine the objective weight of the sequence. One is comparative intensity, and the other is the conflict between evaluation indicators. The first method uses standard deviation to express the value difference between evaluation schemes of the same index. If the standard deviation is larger, the value difference between schemes is larger. The second is to use the correlation between indicators to evaluate the conflict between indicators. If the conflict between the two indicators is low, there is a strong conflict between them. The quantitative indicators that jth indicators conflict with other indicators is: $\sum_{t=1}^{n}\left(1-r_{tj}\right)$, r_{tj} evaluates the correlation coefficient between t and J.

Let Cj represent the amount of information, then Cj can be expressed as:

$$C_j = \delta_j \sum_{t=1}^{n}\left(1-r_{tj}\right) \tag{8}$$

The objective weight of the JTH index is as follows:

$$\theta_j = C_j / \sum_{j=1}^{m} C_j, j = 1, 2, \ldots, m \tag{9}$$

3. The example analysis

3.1 Data acquisition

Considering that Shanghai is a priority pilot city of carbon trading, the data credibility is relatively high. Therefore, the data from January 2017 to March 2021 is selected as the data set, excluding the carbon trading price data of 1000 days in total during holidays when there is no trading, and the data set is made in time sequence.

3.2 ARIMA model construction

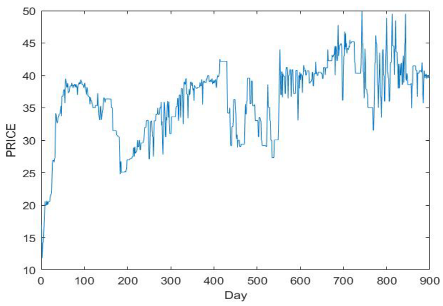

Figure 3: Shanghai carbon trading volume curve

The trading price of 900 days was selected as the training set to construct the time series, and the

ISSN 2616-5902 Vol. 4, Issue 11: 33-38, DOI: 10.25236/AJBM.2022.041106

trading price of 100 days was selected as the verification set. Firstly, according to the actual trading volume, we draw the trading volume - time curve of Shanghai in recent years and observe its trend.

It can be seen from Figure 3 that the trend is very unstable, which does not meet the requirements of ARIMA model for stationary sequence, so first-order difference is carried out for this sequence.

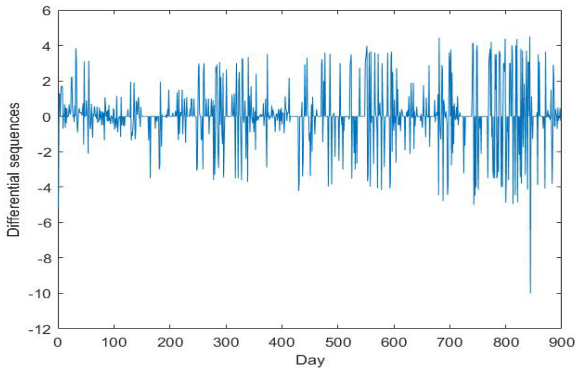

Figure 4: Stationary sequence diagram after difference

It can be clearly seen from the Figure 4 that the time series has turned into a stationary series, so it can be determined that the value of parameter D is 1, and then only the values of P and q need to be determined. We can use autocorrelation coefficient (AC) and partial autocorrelation coefficient (PAC) to identify ARIMA model. This model will draw conclusions by observing the sequence of autocorrelation coefficients generated by ACF(Autocorrelation Function) and PACF(partial autocorrelation Function) or partial autocorrelation coefficients generated by delay number. Based on the principle of parameter determination in Table 1, the autocorrelation coefficient or partial autocorrelation coefficient of the sequence generated by the carbon trading price to the delay number is calculated as shown in Figure 5. Thus, p and q values can be determined to be 6,6 respectively. Finally, a carbon trading price prediction model can be constructed based on ARIMA (1,6,6).

Table 1: Principle of parameter determination

The sequence	AR（p）	MA（p）	ARMA（p, q）
Autocorrelation function	trailing	Q tail cut	trailing
Partial autocorrelation function	The PTH tail is truncated	trailing	trailing

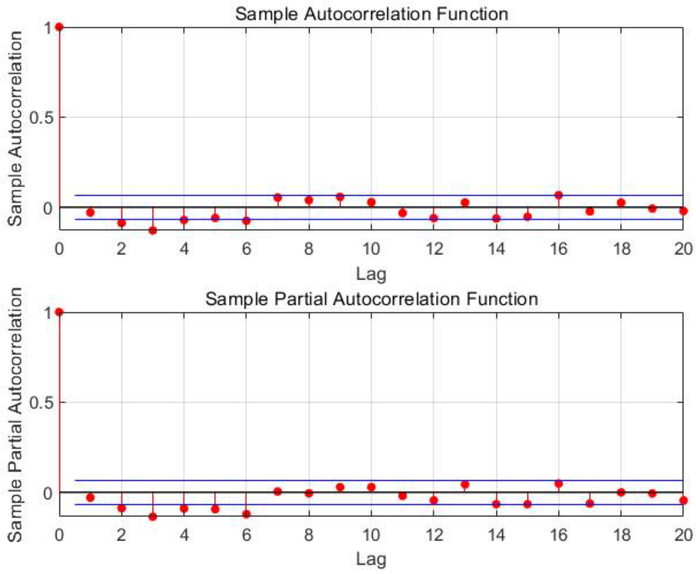

Figure 5: Sample ACF and PACF diagram

ISSN 2616-5902 Vol. 4, Issue 11: 33-38, DOI: 10.25236/AJBM.2022.041106

3.3 Construction of LSTM neural network

Data normalization is to map data points to the interval [0,1]. This model uses the MinMaxScaler method to perform data normalization on the input historical price data. Through data normalization, the network model can be fitted more quickly and the accuracy of model prediction can be improved.

First understand the calculation formula:

$$X_{std}=\frac{X-X.\min\left(axis=0\right)}{X.\max\left(axis=0\right)-X.\min\left(axis=0\right)} \tag{10}$$

$$X_{scaled}=X_{std}*\left(\max-\min\right)+\min \tag{11}$$

In Formula (10) and (11), X: the data to be normalized is usually two-dimensional moment; X. Min (axis=0) : a row vector composed of the smallest values in each column; X. Max (axis=0) : the row vector composed of the maximum value in each column; Max: the maximum value of the interval to map to; Min: minimum value of the interval to be mapped to; X_{std} : standardization results; X_{scaled} : Normalized results.

The training data were normalized and stored for later use, and the LSTM regression network was created to specify the number of hidden units of LSTM layer 128. Since prediction is sequential prediction, input one dimension, output one dimension, training options specified, solver set to Adam, 375 training rounds. The gradient threshold is set to 1. Specify an initial learning rate of 0.005 and reduce it by multiplying it by a factor of 0.1 after 125 rounds of training.

3.4 Data output

MAPE refers to the mean absolute percentage error, which is a relative measure that effectively identifies the MAD scale as a percentage unit rather than a variable unit. Average absolute percentage error is a measure of relative error, which uses absolute value to avoid positive and negative errors cancelling each other out. Relative error can be used to compare the accuracy of prediction of various time series models. In this paper, critic is used to form a combined model of the two models, and MAPE is used to compare ARIMA model, LSTM model and combined model. (Table 2)

Table 2: MAPE values of the three models

MAPE values	7Days	30Days	100Days
LSTM	0.0074	0.0179	0.0591
ARIMA	0.0098	0.0126	0.0459
Portfolio model	0.0087	0.0117	0.0485

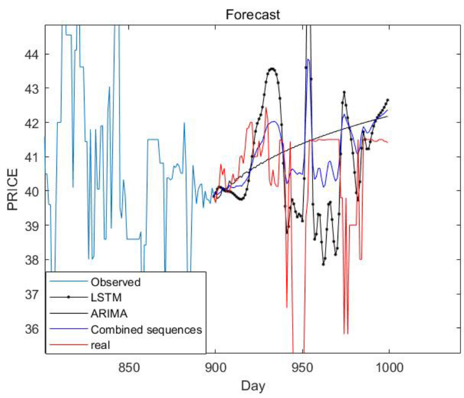

Figure 6: Model output diagram

ISSN 2616-5902 Vol. 4, Issue 11: 33-38, DOI: 10.25236/AJBM.2022.041106

As can be seen from Figure 6 and the value of MAPE, the combination model well integrates the advantages of the two models. It is the curve closest to the real data and has the most advantage in predicting the price in the short and medium term. However, it has no reference for long-term price prediction and can provide reference for the next decision.

4. Conclusion

This paper aims to improve the accuracy of carbon trading price prediction through combination model and provide effective reference for trading. By using ARIMA model's good linear fitting ability and LSTM's powerful nonlinear relationship mapping ability, the price time series is regarded as composed of linear autocorrelation structure and nonlinear structure, and the combination of ARIMA model and LSTM model is used to predict. At the same time, taking the carbon market trading in Shanghai in recent years as an example, the results show that the combination model of ARIMA and LSTM can effectively improve the accuracy of carbon price forecasting in the short and medium term. The combination method gives full play to the advantages of the two models and is an effective method to predict short-term market price changes. At the same time, this model also has good generalization ability, which can be applied to other time series prediction. It can also be used for long-term sequence change analysis. It can be considered to add the wavelet function before the LSTM model to separately predict the sequence decomposition, which can achieve higher accuracy.

References

[1] Rahimi Z H, Khashei M. A Least Squares-based Parallel Hybridization of Statistical and Intelligent Models for Time Series Forecasting [J]. Computers & Industrial Engineering, 2018, (11).

[2] Kim H Y, Won C H. Forecasting the Volatility of Stock Price Index: A Hybrid Model Integrating LSTM With Multiple GARCH-type Models [J]. Expert Systems With Applications, 2018, (103).

[3] Khashei M, Bijari M. Fuzzy Artificial Neural Network (p, d, q) Model for Incomplete Financial Time Series Forecasting [J]. Journal of Intelligent and Fuzzy Systems, 2014, 26(2).

[4] Zhang G P. Time Series Forecasting Using a Hybrid ARIMA and Neu-ral Network Model [J]. Neurocomputing, 2003, (50).

[5] Sun Y, Zhao Z, Ma X, et al. Short-time Scale Gravitational Microlens-ing Events Prediction With ARIMA- LSTM and ARIMA- GRU Hy-brid Model [J]. Lecture Notes in Computer Science, 2019, (11473).