Carbon financial price forecasting based on VMD noise reduction and improved DELM optimized by the WDO algorithm

Jingying Fan[1,#], Jiayin Chen[2,#], Haoxuan Li[3,*,#]

[1] SILC business school, Shanghai University, Shanghai, China, 201899

[2] College of Economic and Management, Shanghai Polytechnic University, Shanghai, China, 201209

[3] College of Economic and Management, Changchun University of Science and Technology, Changchun, China, 130012

*Corresponding author: 1376880822@qq.com

[#] These authors contributed equally

Abstract. Carbon price forecasting can help stabilize the carbon pricing mechanism and reduce carbon market risks. This paper firstly uses the closing price of the Guangzhou carbon exchange to predict carbon prices. Secondly, this paper constructs model based on a wind-driven algorithm (WDO) and deep extreme learning machine (DELM), compared with the results of a backpropagation neural network (BP). The prediction results are reliable, with a 49.81% decrease in mean square error (MSE), which shows that the validity of the hybrid VMD-DELM approach is verified.

Keywords: VMD; WDO; DELM; Carbon price; forecasting.

1. Introduction

China starts to set up several carbon emission trading markets in 2013 in order to achieve China's green technology innovation and green economy goals. The study of carbon price formation mechanism and carbon price fluctuation factors provides a valuable reference basis for understanding and avoiding carbon market risks.

The existing methods of carbon price forecasting research topics can be divided into parametric and non-parametric methods [2]. Some scholars think parametric methods such as Generalized Auto Regressive Conditional Heteroskedasticity(GARCH) models are more effective in fitting the time series of carbon price fluctuations and the trend of yield fluctuations [3], and the difference between predicted and actual values in the carbon market is small [4]. However, China only has a short time to establish the carbon emission trading market, and the small average annual price data of carbon trading without detailed data support. Hence, the data characteristics of the GARCH model are not available and lack corresponding experimental accuracy. On this basis, scholars began to introduce nonlinear methods to deal with the nonlinear fluctuations of carbon market prices. For example, Zhu et al. empirically proposed the GMDH-PSO-LSSVM model to analyze the carbon futures prices at two different expiration times of the EU ETS [5]. The study found that this literature lacks the original noise reduction process for carbon price data and some experimental accuracy. Meanwhile, DELM can capture the mapping relationship between data more comprehensively and improve the accuracy of processing high-latitude input variables to improve the experimental accuracy [6].

Thus, this paper proposes a DELM carbon price prediction method based on VMD decomposition noise reduction and wind-driven optimization for the carbon price prediction problem. The innovative points of this paper are as follows.

1) VMD decomposes the original carbon price time-series data to improve the data quality and obtain the optimal parameters, which paves the way for improving the prediction accuracy [7].

2) Adopt the WDO forecasting to optimize the parameters. The problem of large difference between DELM and real values is effectively solved, and the subsequent addition of the DELM model

makes the training process time-consuming and accurate [8], with strong advantages in learning rate and generalization ability.

3) To improve the validity of the model, this paper conducts a multi-model comparison and uses MSE and other indicators to conduct model indicator error tests.

This paper uses Guangzhou carbon market data and a multi-model comparison. Except for Chapter 1, the remaining chapters of this paper are arranged as follows. Chapter 2 introduces the basic principles of the VMD model, and Chapter 3 carries out the example measurement based on the data processed by VMD after noise reduction and uses WDO-DELM to forecast the data. Based on this, the unprocessed data from the BP model is compared with our VMD-WDO-DELM model. Chapter 4 continues with the conclusion and outlook and provides relevant conclusions and recommendations.

2. Introduction of the Model

2.1 VMD Decomposition Noise Reduction

2.1.1 Variational modal decomposition

VMD is an entirely non-recursive adaptive signal processing method whose entire structure is a variational question [9]. The signal decomposition process is achieved by searching for the optimal solution of the constrained variational model, iteratively updating the frequency center and bandwidth of each eigenmode function in solving the variational model [10].

2.1.2 Basic principles of variational modal decomposition

VMD decomposes an initial signal f into K discrete IMFs (eigenmodal functions) u_k (k ∈ 1, 2, ..., K), u_k with a specific sparsity of bandwidth in the frequency domain.

For each mode, the associated resolved signal is calculated by Hibert Transform to obtain a one-sided spectrum. Moreover, the spectrum of each mode is modulated onto the corresponding baseband using the predicted center frequency. The bandwidth is estimated from the Gaussian Smoothing of the demodulated signal above, enabling the corresponding constrained variational model expression to be obtained as:

$$\min_{\{u_k\},\{\omega_k\}}\left\{\sum_k\left\|\partial_t\left[\left(\partial(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\} \tag{1}$$

$$\sum_k u_k = f \tag{2}$$

Where f is the original signal, u_k is the mode function. ω_k is the center frequency of each mode.

The Lagrange function is introduced using the two-form factorization and the Lagrange multiplier methods. Then, alternating updates find the saddle point of the Lagrange expression. The frequency domain is then obtained by Fourier transforming.

$$\hat{u}_k^{n+1} = \text{argmina}\left\{\alpha\left\|j\omega\left[\left(1+\text{sgn}(\omega+\omega_k)\right)\hat{u}_k(\omega+\omega_k)\right]\right\|_2^2+\left\|\hat{f}(\omega)-\sum_i\hat{u}_i(\omega)+\frac{\hat{\lambda}(\omega)}{2}\right\|_2^2\right\} \tag{3}$$

Where α denotes the penalty parameter; λ denotes the Lagrangian factor; ω denotes the random frequency; S denotes the set of u_k.

The updated expression for the final solution u_k^{n+1}、 ω_k^{n+1} is shown in equations (4) and (5) by substituting the $\omega=\omega-\omega_k$ variable.

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum_{i\neq k}\hat{u}_i(\omega)+\frac{\hat{\lambda}(\omega)}{2}}{1+2\alpha(\omega-\omega_k)^2} \tag{4}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 \, d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 \, d\omega} \tag{5}$$

2.1.3 Algorithmic flow of the variational modal decomposition

(1) Initialize u_k^1、ω_k^1、λ_k^1、n to 0.

(2) $n = n + 1$, executing the entire loop of the algorithm.

(3) According to equation (6), equation (7) u_k、ω_k to the K number of modal decompositions is needed.

(4) Based on.

$$\lambda^{n+1} = \lambda^n + \tau(f - \sum_k \hat{u}_k^{n+1}) \tag{6}$$

Update λ.

(5) Repeat the cycle until the following conditions are met.

$$\sum_k \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \left\| \hat{u}_k^{n+1} \right\|_2^2 < e \tag{7}$$

The VMD algorithm performs noise reduction decomposition of daily average transaction price data of carbon emission allowances. Then, the signal can be adaptively divided in the frequency domain while describing the local characteristics of the signal, and the number of modal decompositions can be adaptively determined according to the sequence to be decomposed, thus matching the optimal center frequency and finite bandwidth of each mode [11].

2.2 Model Prediction Framework

2.2.1 DELM

In 2004, Guangbin Huang proposed that ELM can improve Backward Propagation (BP) to improve learning efficiency and simplify the setting of learning parameters. The ELM is the basic unit of the Deep Extreme Learning Machine DELM. The idea of DELM is to learn high-level features of the original data after each layer of training by minimizing the reconstruction error.

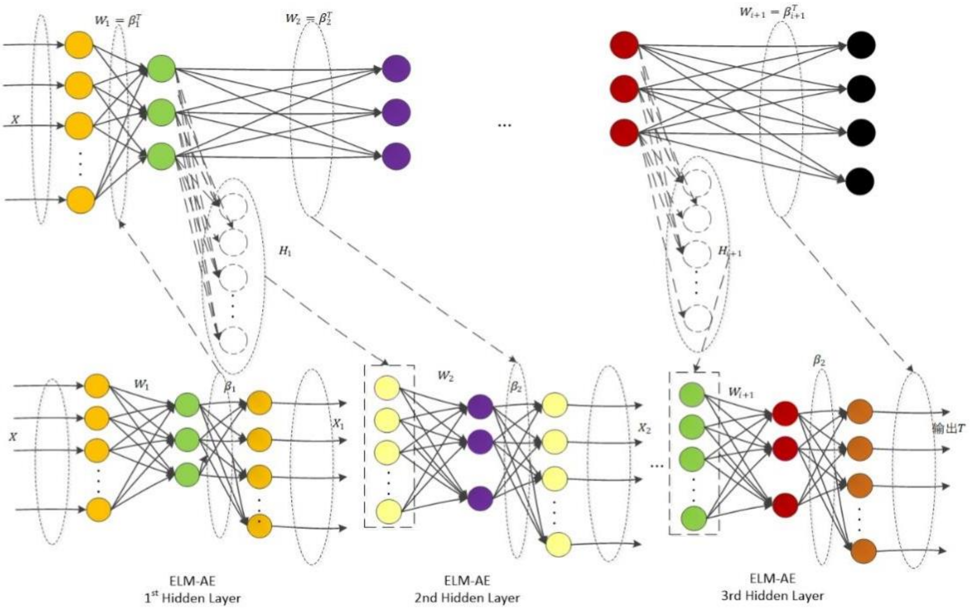

Figure 1. Training process of DELM

Fig.1 depicts how to train the DELM.

In the first step, the input data sample X is taken as the target output of the 1st ELM-AE($X_1 = X$), which in turn is used to obtain the output weight β_1.

In the second step, the matrix H_1 is the output matrix of the 1st hidden layer of DELM, which is taken as the input and target output of the next ELM-AE($X_2 = X$), and is trained layer by layer, while the last layer is trained with ELM to solve the output weight $\beta_{i+1}$of the last hidden layer of DELM. The input weight matrix of each hidden layer of H_{i+1} is $W_{i+1} = \beta_{i+1}^T$.

2.2.2 WDO-DELM

The Wind Driven Optimization (WDO) algorithm is a novel optimization method based on atmospheric motion. It is based on Newton's second law and the ideal gas equation of state. The velocity update equation is thus derived as:

$$u_{new} = (1-\alpha)u_{cur} - gx_{cur} + \left(RT\left|\frac{1}{t}-1\right|\left(x_{opt}-x_{cur}\right)\right) + \left(\frac{cu_{cur}^{other\ dim}}{i}\right) \tag{8}$$

$$c = -2|\Omega|R + T \tag{9}$$

From the above principle, the weights in the original DELM are initialized by random initialization, so the initial weights of the DELM are optimized. The fitness function is designed as follows.

$$Fitness = 2 - Accuracy(Train) - Accuracy(Test) \tag{10}$$

2.3 VMD-WDO-DELM Prediction Model

Carbon financial prices are nonlinear and non-stationary. This paper proposes a prediction model combining variational modal decomposition and an improved WDO-DELM model, and the steps of the prediction model are shown in Fig.2.

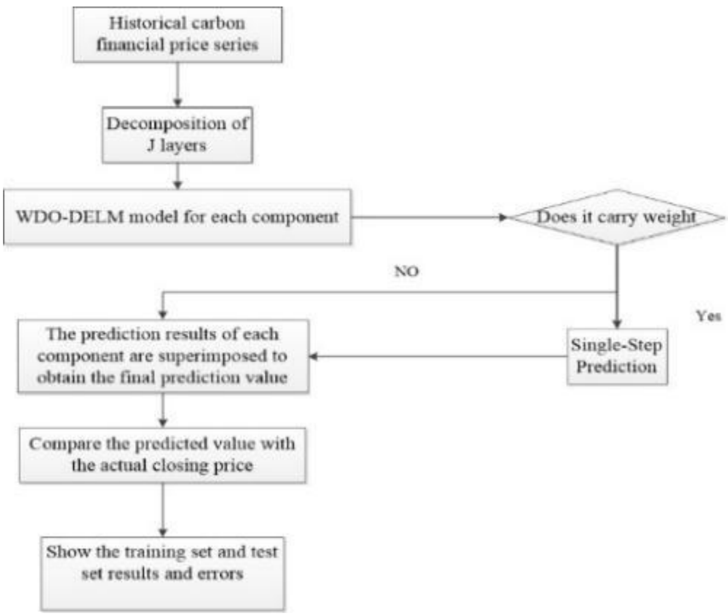

Figure 2. Overview of the Model

3. Results

3.1 Original Data

The experimental data in this paper is the historical carbon emission rights trading price information from the Guangdong Carbon Emission Exchange. A total of 1310 data from March 3, 2015, to March 23, 2021, were selected [12].

For data cleaning, linear interpolation is used to fill in the missing data to ensure the quality and reliability of deep learning. The isolation Forest method is used for outlier detection, which is used to detect sparsely distributed and far from the high-densified population. The algorithm iForest consists of t iTrees (Isolation Tree), each of which is binary, and the result obtained by cyclic training is shown in Fig.3. For each outlier, the filtering is done with the current day's situation.

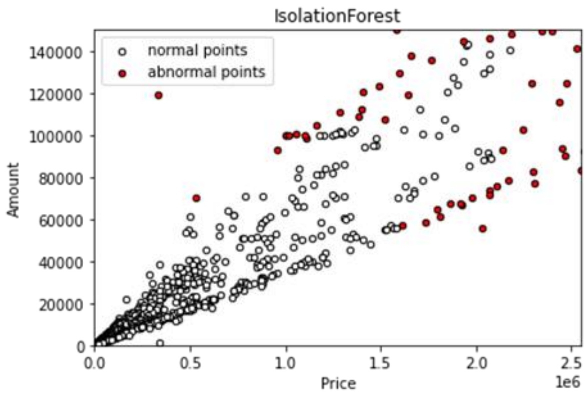

Figure 3. Isolated Forest Test Results

3.2 VMD

The VMD requires determining the number of modes K. The main distinction between modes is the difference in center frequencies. This paper determines K by observing the central frequency, and modes with similar central frequencies are over-decomposed from K of 7. Therefore, the number of modes is chosen to be 6. Then, the default value of 2000 for α is used for VMD. The image of its original data is shown in Figure 1, and the decomposition results are shown in Figs. (4,5).

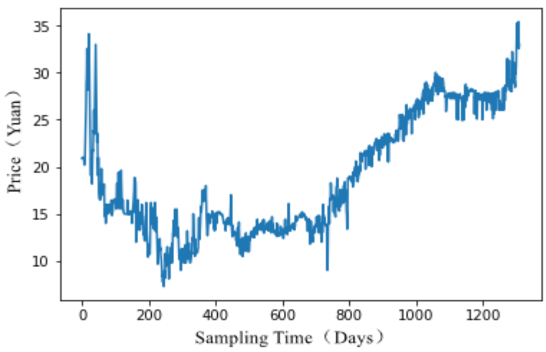

Figure 4. Original Input Signal

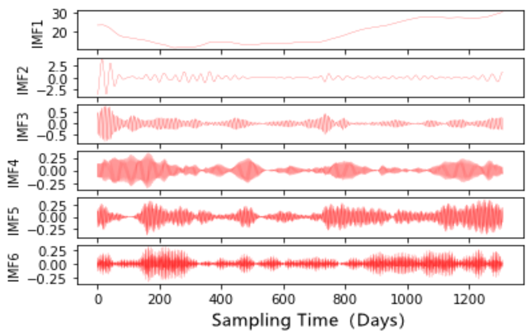

Figure 5. Decomposition Results of VMD

3.3 Comparative analysis of prediction models

In order to illustrate the effectiveness of the WDO-DELM model, the WDO-DELM optimized by VMD noise reduction is compared with the DELM and BP model, and the six components V1-V6 are used to predict the average carbon closing price. This paper uses three error evaluation indicators: mean absolute percentage error (MAPE), mean square error (MSE), and coefficient of determination (R^2).

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\% \tag{11}$$

$$\mathrm{MSE} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2} \tag{13}$$

The smaller the value of MAPE and MSE and the larger the value of R^2 means, the better the prediction effect. The comparison of the prediction results of the testing sets is shown in Figs. (6,7).

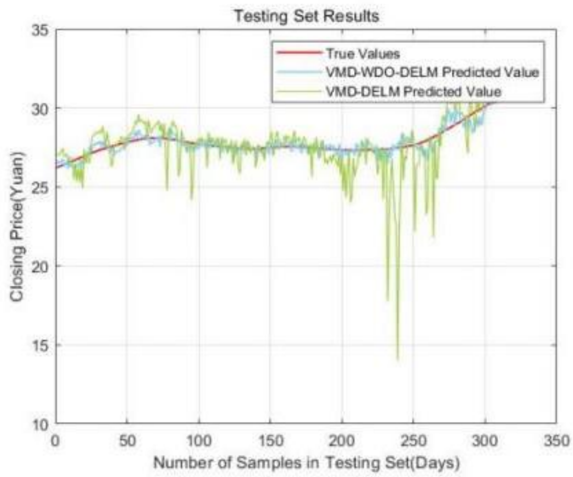

Figure 6. Testing Set Results

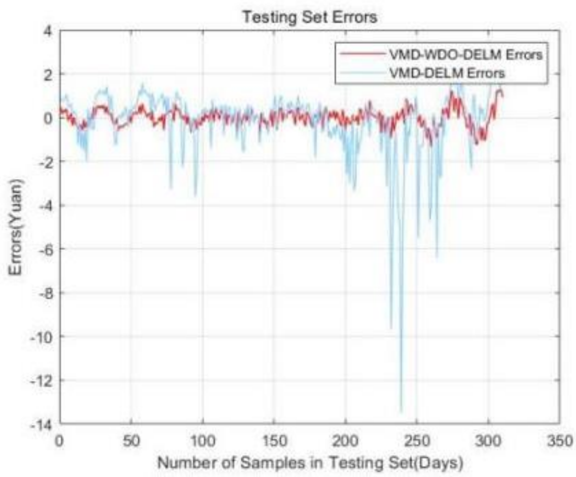

Figure 7. Testing Set Errors

The prediction errors of the testing set for different models are shown in Table 1.

Table 1. Testing set errors for different models

	MSE	MAPE	R^2
BP	1.31010	0.02343	0.00647
VMD-DELM	0.36191	0.01666	0.42848
VMD-WDO-DELM	0.18165	0.01166	0.71314

From the results of the testing set, the VMD-WDO-DELM model reduces MSE by 86.13%, MAPE by 50.23%, and R^2 10,922.26% compared to the BP model VMD-WDO-DELM model reduced MSE by 49.81%, MAPE by 30.01%, and R^2 improved by 44.43%.

The prediction curves of the VMD-WDO-DELM model overlap with the actual curves more than the standard VMD-DELM and BP models. The standard VMD-DELM model has some errors with the actual curve at the mutation points, but the optimized VMD-DELM model by the WDO algorithm can solve this problem, and the overall curve is similar to the actual curve, which achieves better prediction results.

With the development of global carbon trading, this paper adjusts the sample period and selects the five-year data from March 22, 2016, to March 22, 2021, for robustness testing. The results show that compared with the BP model, the MSE of VMD-DELM decreases by 100.00%, MAPE decreases by 99.98%, and R2 increases by 11036.63% in the test set data. Therefore, the model is robust.

4. Conclusion

The launch of the national carbon market has optimized the environment for the development of carbon finance, and policies have encouraged the development of carbon finance activities, such as the Opinions on Further Promoting the Healthy Development of the Capital Market.

This paper proposes a new forecasting framework for the carbon financial price problem. The value of this paper is mainly 3 points.

1) VMD decomposes the original carbon price time-series data to obtain the optimal parameters.

2) The WDO algorithm is used to optimize the parameters, and the optimal parameters are selected as a pop size of 40 and a max iteration of 100.

3) The BP model without noise reduction and the VMD-DELM and VMD-WDO-DELM models after noise reduction are compared. The following conclusions are obtained from the error analysis experiments after noise reduction:

i. VMD-WDO-DELM shows a decrease in MSE and increase in MAPE compared with the BP and VMD-DELM models and effectively predicts carbon prices.

ii. After algorithm optimization, the WDO-DELM curve is closer to the actual value.

However, this study also has some limitations. First, the prediction performance of this model is outstanding, but the calculation needs to be simplified. Second, in the integrated forecasting stage, the data treatment of VMD for carbon financial prices is unstable because the data are multifactorial. Therefore, future research will focus on simplifying the models, considering better model selection in multi-factor situations, and using deep learning methods to improve the prediction results.

References

[1] Jiang F, Peng Zijun. Carbon price prediction based on chaotic PSO optimized BP neural network. Statistics and Information Forum, 2018, 33(05): 93-98.

[2] Fan Lili, Liu Lu. Analysis of carbon price yield volatility in China based on ARMA-GARCH model. Economic circles, 2020(05): 25-32.

[3] Paolella M S, Taschini L. An Econometric Analysis of Emission Allowance Prices. Journal of Banking and Finance, 2008, 32(10)

[4] Help Zhu, Yiming Wei. International carbon market price forecasting based on GMDH-PSO-LSSVM. Systems Engineering Theory and Practice, 2011, 31(12): 2264-2271.

[5] Bai Lingwei. Research on volatility spillover effects and price forecasting in China's carbon market[D]. Shanxi University of Finance and Economics,2019.

[6] Qian Lin, Kang Min, Fu Xiuqing, Wang Xingsheng and Fei Xiuguo, Application of VMD-based adaptive morphology in bearing fault diagnosis, Vibration and Shock, Volume 36, Issue 3, Pages 227-233, 2017, doi: 10.13465/j.cnki.jvs.2017.03.036.

[7] Zou NANA, Hu HANGPING, Bai YANPING, Shi NA. A joint denoising method based on optimization of VMD and wavelet packet noise reduction [J]. Practice and understanding of mathematics, 2021, 51(20): 135-142.

[8] Chen, B.B., Cao, C.Q., Yu, S.W.. Optimization of PID parameters based on wind-driven optimization algorithm WDO[J]. Computer Engineering and Applications,2016,52(14):250-253+260.

[9] Tang Guiji and Wang Xiaolong, Application of parametric optimization of variational modal decomposition method in early fault diagnosis of rolling bearings, Journal of Xi'an Jiaotong University, vol. 49, no. 5, pp. 73-81, 2015.

[10] Zhao Xin, Chen Chenpeng, Bi Guihong, Chen Shilong and Xie Xu, 'A multi-model decomposition-based portfolio prediction model for carbon trading prices', Power Science and Engineering, vol. 38, no. 2, pp. 52-59, 2022.

[11] Xie X., Yan S. Ping. Prediction of carbon emission trading price based on ARMA-BP combination model--Fujian Province as an example [J]. Journal of Longyan College, 2021, 39(02): 107-115. DOI: 10.16813/j.cnki.cn35-1286/g4.2021.02.018.

[12] Guangdong Carbon Exchange website http://www.cnemission.com/article/hqxx/

[13] Zhou L, Liu D, Zheng XL. A PSO-DELM-based method for predicting mobile phone Internet traffic [J]. Computer Engineering and Design, 2021, 42(02): 316-323. DOI: 10.16208/j.issn1000-7024.2021.02.003.

[14] JIN Yongchao, ZHUANG Xiaodie, WANG Renfang, CAO Qian. ARIMA-BP neural network-based power coal futures price forecasting [J]. Journal of North China University of Technology (Natural Science Edition), 2022, 44(02): 92-98.

[15] WANG Jinxin, JIANG Ziyan, WANG Qi, LI Dong, YU Furong, LIU Hai, REN Jingyuan. Price prediction in the spot market of electricity in western Mongolia based on SA-BP algorithm[J]. Inner Mongolia Electric Power Technology, 2021, 39(04): 42-46. DOI: 10.19929/j.cnki.nmgdljs.2021.0076.