Journal of Innovation and Development
ISSN: 2958-4752 | Vol. 2, No. 3, 2023

Carbon price decomposition ensemble hybrid forecasting model based multi-scale feature extraction

Qiongfang Zhang

School of Management, Shanghai University, Shanghai 200444, China

Abstract: The carbon emission market is the core policy tool to achieve the goal of carbon peaking and carbon neutrality. To fully extract the complex features of carbon price series such as non-stationary, non-linear, and multi-scale etc. This paper constructs an integrated hybrid forecasting model CEEMD-GWO-LSSVR based on the multi-scale decomposition of carbon price decomposition. Firstly, the original carbon price series are decomposed into eigenmodal functions (IMFs) of different scales by complementary ensemble empirical modal decomposition (CEEMD), and the LSSVR model optimized by the grey wolf optimization algorithm (GWO) is used as the prediction model to forecast the obtained IMFs, and finally, the prediction results of all IMFs are linearly integrated. This paper selects the price data of the Shanghai carbon trading market for the empirical study, and the empirical results show that the prediction accuracy of the hybrid model proposed in this paper is significantly better than that of the benchmark model.

Keywords: Carbon price forecasting; Complementary ensemble empirical mode decomposition; Grey wolf optimizer; Least square support vector regression.

1. Introduction

Since the industrial revolution, with the rapid development of industrialization and urbanization, human beings have consumed a large number of fossil fuels such as coal in production and life and emitted a large number of greenhouse gases. The greenhouse effect caused by greenhouse gases has endangered the earth's ecosystem. According to existing studies, various phenomena indicate that climate warming has brought irreversible damage to the earth, and the greenhouse effect has led to a reduction in the production of major crops . Therefore, solving the problem of global warming is a severe challenge facing human society. After many years, the EU Emissions Trading Scheme (EU ETS) initiated by the European Union has proved to be an effective tool to deal with climate change, which provides important reference value and significance for the formulation and development of trading systems in carbon trading markets around the world. As one of the major carbon emitters, China plays a crucial role in global climate protection. At the same time, as an emerging market economy, China has also been fulfilling its responsibilities. Since 2013, China has successively launched eight carbon emission trading pilot markets including Shenzhen, Shanghai, Beijing, and Guangzhou. In March 2021, in the government work report, it was pointed out that it is necessary to do a good job in the work of carbon peaking in 2030 and carbon neutrality in 2060. Although China's carbon trading market has conducted a lot of pilot work, it is still in the early stages of the launch of a unified carbon emissions market. Due to insufficient market stability and inadequate trading mechanisms, trading prices are highly volatile, and carbon prices exhibit characteristics such as nonlinearity, non-stationary nature, and multiple scales. Therefore, forecasting carbon prices is a great challenge. Effective and accurate prediction of carbon emission rights prices can, on the one hand, understand the changing rules of carbon emission rights prices, establish effective and stable carbon pricing mechanisms, and on the other hand, provide practical guidance for production, operation, and decision-making.

Since the establishment of the carbon market, many scholars have studied carbon price forecasting, and the commonly used methods are mainly divided into three categories: statistical econometric model, artificial intelligence model, and mixed model. The traditional statistical measurement model is widely used in carbon price forecasting because of its simple and flexible performance. Statistical econometric models include linear regression model and GARCH etc. For example, Guobrandsdottir and Haraldsson studied the driving factors of carbon price change and made a prediction based on linear multiple regression. Byun and Cho used the GARCH model to predict the volatility of carbon prices, and the results showed that the model was superior to the volatility method and the K-nearest neighbor method. Benz and Truck used the Markov transformation method and the standard AR-GARCH model to predict the volatility of carbon spot prices at different stages. However, carbon price fluctuations do not follow the mean regression process, and traditional econometric models cannot achieve good forecasting results. Artificial intelligence models can effectively improve the shortcomings of traditional measurement models, and artificial intelligence models have also been widely used in carbon price forecasting. The commonly used artificial intelligence models mainly include MLP neural network, BP neural network, LSSVR, etc. For example, Fan et al. established a multi-layer perceptron neural network (MLP) forecasting model for the strong nonlinear characteristics of carbon prices. The results show that the MLP model has good performance in carbon price forecasting. Yi et al. used the Back Propagation Neural Network (BPNN) method to predict carbon prices, which has higher forecasting accuracy than traditional statistical models. Zhu et al. used the unique advantages of the ARIMA model capturing linear models and the LSSVR model capturing nonlinear models to predict the EU carbon emission price, and the results showed that the proposed method has good forecasting accuracy.

Due to the complex impact of factors such as trading mechanisms, energy prices, heterogeneous environments, and government policies on the carbon emissions trading market, its price series exhibits complex characteristics such as non-stationary, nonlinear, and multiscale characteristics. Using a single forecasting model cannot fully capture the volatility of carbon prices, so hybrid models are widely used. Wang et al. innovatively proposed a TEI@I Complex system methodology that emphasizes the idea of "decomposition and ensemble", and combines data preprocessing and decomposition technology with artificial intelligence models, significantly improving the overall forecasting accuracy of the model. A general hybrid forecasting model can be divided into three stages: decomposition, forecasting, and ensemble; First, the decomposition method is used to decompose the original sequence to obtain different subsequences, and then the forecasting model is applied to the modeling and forecasting subsequences respectively. Finally, the forecasting results of each subsequence are integrated to obtain the final forecasting results. In the studies of Yang et al. , Chai et al. , Zhu et al. and Wang et al., mixed prediction model was used.

The above studies have shown that the effect of decomposing and then predicting subsequences is better than that of non-decomposition models, reflecting the effectiveness of the decomposition method. Some commonly used decomposition methods have certain drawbacks, such as modal aliasing and endpoint effects in EMD. Wavelet analysis (WT) is easily affected by sampling frequency and noise. The CEEMD decomposition method can reduce reconstruction errors based on avoiding modal aliasing. The prediction model LSSVR is an improvement on SVR, which overcomes the shortcomings of standard SVR by converting the solution of a quadratic programming problem into a solution of a linear equation system, and has a high solution speed and prediction accuracy. The LSSVR model has been applied to traffic flow forecasting, wind power forecasting, and other fields. However, the regularization parameter C and kernel function parameters σ of the LSSVR model need to be set empirically. Therefore, given the shortcomings of the above single decomposition method, incomplete feature extraction, and the need to manually set the parameters of the prediction method, this paper proposes a decomposition ensemble hybrid forecasting model based on multiscale decomposition.

The remainder of the paper is organized as follows: Section 2 describes the fundamental methods and framework design used in this paper; Section 3 is empirical research, which applies the forecasting framework proposed in this paper to the carbon market price in Shanghai; Section 4 summarizes the conclusions of this study.

2. Fundamental methods and the proposed model

2.1. Complementary ensemble empirical mode decomposition

Wu and Huang et al. proposed an adaptive signal decomposition method, namely the EMD algorithm. Because this method can adaptively decompose signals, it is widely used in various fields, including engineering applications, finance, digital signals, etc. However, EMD has problems such as modal aliasing and endpoint effects. In 2010, Yeh et al. proposed a new method, CEEMD, which adds a set of positive and negative standard white noise with the same amplitude to the signal decomposition. This method can effectively eliminate the added white noise, reduce the reconstructed signal error caused by residual white noise, and improve the efficiency of signal decomposition calculation.

The specific steps of CEEMD are as follows:

(1) A pair of standard white noise signals $[n_i(\tau), -n_i(\tau)]$ with the same amplitude and phase angle of 180 ° are added to the signal $X(\tau)$ to obtain two new signals:

$$\begin{cases} X_i^+(\tau) = X(\tau) + n_i(\tau) \\ X_i^-(\tau) = X(\tau) - n_i(\tau) \end{cases} \tag{1}$$

Where n_i is the i th added white noise and $X_i^+(\tau)$ $X_i^-(\tau)$ are the i th positive noise and negative noise signals, respectively, $(i = 1, 2 \cdots, M)$.

(2) EMD processing is performed on the new signal $\left[X_i^+(\tau), X_i^-(\tau)\right]$, and s set of IMF components is achieved for each signal:

$$\begin{cases} X_i^+(\tau) = \sum_{j=1}^{m} s_{ij}^+(\tau) \\ X_i^-(\tau) = \sum_{j=1}^{m} s_{ij}^-(\tau) \end{cases} \tag{2}$$

Where $s_{ij}^+(\tau)$ and $s_{ij}^-(\tau)$ are the j th IMFs obtained in the i th positive trial and negative trial, respectively, $(j = 1, 2 \cdots, m)$.

(3) Repeat step (1) and step (2) with different white noise components every time to obtain two sets of IMF components.

(4) Compute the average value of all the corresponding IMFs:

$$IMFj = \frac{1}{2M} \sum_{i=1}^{M} \left(s_{ij}^+(\tau) + s_{ij}^-(\tau)\right)s \tag{3}$$

Where IMF j is the j th IMF derived by the CEEMD algorithm.

(5) Finally, the original signal $X(\tau)$ is decomposed into m IMF components and a residual component $R_m(\tau)$, that is:

$$X(\tau) = \sum_{j=1}^{m} IMF_j + R_m(\tau) \tag{4}$$

2.2. Least square support vector regression

Traditional support vector regression formulates training processes through quadratic programming, which requires a long training process and often leads to overfitting problems. This requires a long training program and often leads to overfitting problems. To solve the above problems, Suykens and Vandewalle proposed the least squares support vector machine, a hyperplane-based model that has been successfully applied to regression. LSSVR optimizes the constraint conditions and loss functions of the model based on SVR, replacing nonequality constraints in SVR with equality constraints. An error variable is introduced, and the sum of squares of errors is used as a loss function. Reduce the computational dimension and skillfully transform nonlinear problems into linear problems through kernel functions, making them easy to solve. Compared with SVR, the LSSVR model has significant improvement in computational efficiency and good generalization ability.

Assume $\{(x_i, y_i)\}, i = 1, \cdots, N$ is the training dataset for output y_i , LSSVR can be expressed as the following constrained optimization problem based on structural risk minimization:

$$\min J(w,\varepsilon)=\frac{1}{2}w^T w+\frac{C}{2}\sum_{i=1}^{N}\varepsilon_i^2 \; s.t. y_i = w,\psi(x_i)+b+\varepsilon_i \quad (5)$$

Where $i=1,\cdots,N$ C represents a regularization parameter, ε_i represents an approximate error, w and b represent weight vectors, and bias terms, respectively, $\psi(\cdot)$ represents a nonlinear mapping function.

Using the Lagrange multiplier method to solve the above problem, the following vector expression can be obtained:

$$\begin{pmatrix} b \\ \alpha \end{pmatrix}=\begin{pmatrix} 0 & 1^T \\ 1 & K+C^{-1}I \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ y \end{pmatrix} \quad (6)$$

Where $\alpha=(\alpha_1,\cdots,\alpha_N)^T$, $y=(y_1,\cdots,y_N)^T$, $\mathbf{1}=(1,1\cdots1)^T$, matrix K, $k_{ij}=\psi(x_i),\psi(x_j)$, I is an identity matrix. Finally, the LSSVR model can obtain:

$$f(x)=\sum_{i=1}^{N}\alpha_i K(x,x_i)+b \quad (7)$$

Where $K(x,x_i)=\psi(x)\,\psi(x_i)$ is a Gaussian kernel function, $K(x,x_i)=\exp\left(-\frac{(x-x_i)^2}{2\sigma^2}\right)$, σ is the kernel width.

2.3. Grey wolf optimizer

GWO algorithm is a swarm intelligence optimization algorithm proposed by Mirjalili et al. . This algorithm achieves the goal of optimization by simulating hunting behavior in natural wolf packs based on the mechanism of wolf group collaboration. GWO algorithm has the characteristics of a simple structure, fewer parameters to adjust, and easy implementation. The principle of the GWO algorithm is as follows: To mathematically model the social level of grey wolves in GWO, the first three best wolves (optimal solution) are respectively defined as α , β , δ , which guides other wolves to search towards the target. The remaining wolves (candidate solution) are defined as ω , those which update their positions around α , β , δ . The first layer: The α wolf pack is the leader in the population, responsible for hunting prey throughout the wolf pack, that is, optimizing the optimal solution in the algorithm; The second layer: the β wolf pack is responsible for assisting the wolf pack, which is the suboptimal solution in the optimization algorithm; The third layer: the δ wolf packs, following the commands and decisions of α and β , will reduce the fitness of α and β to δ ; The fourth layer: the ω wolf pack, which updates its position around α , β or δ . The hunting of grey wolves involves three main steps: encircle, hunt, and attack their prey.

(1) Encircling: grey wolves encircle prey during the hunt. To mathematically model encircling behavior the following equations are proposed:

$$D=\left|C\cdot X_P(t)-X(t)\right| \quad (8)$$

$$X(t+1)=X_P(t)-A\cdot D \quad (9)$$

$$A=2a\cdot r_1-a \qquad C=2r_2 \quad (10)$$

Where t indicates the current iteration, A and C are coefficient vectors, X_p is the position vector of the prey, X indicates the position vector of a grey wolf, the components of a are linearly decreased from 2 to 0 throughout iterations and r_1 , r_2 are random vectors in [0,1].

(2) Hunting: In the GWO algorithm, the optimal grey wolf is considered to be α , the suboptimal grey wolf β , the third best grey wolf δ , and the rest of the grey wolves are ω . The model is established according to the characteristic that α , β , δ has more knowledge about the position of the prey, and α , β , δ is used to guide the movement ω in the iterative process, to achieve global optimization. Update the positions of all grey wolves using the following equation:

$$D_\alpha=\left|C_1\cdot X_\alpha(t)-X(t)\right|, X_1=X_\alpha-A_1\cdot D_\alpha \quad (11)$$

$$D_\beta=\left|C_2\cdot X_\beta(t)-X(t)\right|, X_2=X_\beta-A_2\cdot D_\beta \quad (12)$$

$$D_\delta=\left|C_3\cdot X_\delta(t)-X(t)\right|, X_3=X_\delta-A_3\cdot D_\delta \quad (13)$$

Where D_α , D_β and D_δ respectively represent the distance between grey Wolf ω and layer α , β and δ wolves; X_α , X_β , X_δ respectively represent the current positions of α , β and δ ; C_1 , C_2 , C_3 are random vectors; X_1 , X_2 , X_3 respectively represent the positions that need to be adjusted by individual ω grey wolves affected by wolves in the layer α , β and δ ; Take the average value ,i.e.

$$X(t+1)=(X_1+X_2+X_3)/3 \quad (14)$$

(3) attacking prey: When the prey stops moving, the wolf starts to attack and the hunting process is terminated.

2.4. Grey wolf optimizer optimized Least square support vector regression

Least Squares Support Vector Regression is an extension of SVR that simplifies the computation and speeds up the solution while ensuring the original advantages. However, when using LSSVR to model the data for forecasting, the prediction accuracy depends on the selection of the regularization parameter C and the kernel function parameter σ . The larger C is, the easier it is to overfit and generalize poorly; the smaller C is, the easier it is to underfit; The larger σ is, the more likely a smoothing effect will occur, resulting in poor training; the smaller σ is, the model will only act near the support vector, resulting in poor prediction. The GWO algorithm has the advantages of a simple structure, few parameters to be adjusted, and easy implementation. Therefore, in this paper, the GWO algorithm is used to optimize the values of two parameters C and σ LSSVR, the mean square error is used as the fitness function, and the minimization of the MSE value is used as the optimization-seeking objective. The main steps of GWO optimization of LSSVR are as follows:

Step 1: Set the range of values of LSSVR parameters C and σ ;

Step 2: Initialize the relevant parameters in the GWO algorithm, and set the position of the grey wolf to [C , σ];

Step 3: The location parameters of each grey wolf in the grey wolf population are used as the values of C and σ . The LSSVR is trained and fitted on the training set, and the MSE values of the fitting results corresponding to all parameter combinations are derived, and the minimum MSE value is used as the fitness value under the parameter combinations C and σ ;

Step 4: Comparison of fitness values under different combinations of parameters C and σ represented by grey wolves, grading of grey wolf populations, and updating of grey wolf locations;

Step 5: Calculate the fitness value of each grey wolf individual after the location update and compare it with the optimal fitness value of the previous generation, and update it to obtain the new optimal fitness value;

Step 6: Repeat steps 3-5 until the maximum number of iterations is reached, then the optimization process ends; the grey wolf position corresponding to the optimal fitness value is the combination of C and σ optimal parameters.

2.5. The framework of our proposed approach

Based on the above methodological principles, this paper proposes an integrated prediction model for carbon price decomposition based on CEEMD-GWO-LSSVR. The prediction process proposed in this paper is shown in Fig 1, with the following steps:

Step 1: Data pre-processing; for complex features such as the non-smooth and non-linearity of carbon price time series, to better separate and extract the features of the original carbon price series, the CEEMD method is used to decompose the original carbon price series adaptively to obtain n sub-series.

Step 2: Subsequence forecasting; least squares support vector regression model is used for modeling and , and the optimal C and σ parameters of the model are obtained by the grey wolf optimization algorithm search, and the GWO-LSSVR model is applied to all subsequences obtained in step 1.

Step 3: Reconstruction of sub-series predicted values; the predicted values of each sub-series are superimposed to obtain the final prediction results.

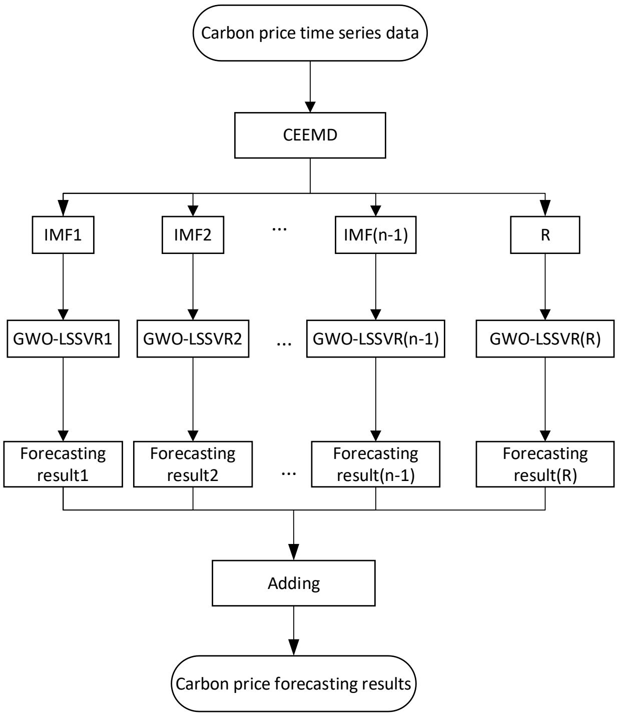

Fig 1. The forecasting process for carbon price with CEEMD-GWO-LSSVR

3. Empirical analysis

3.1. Data sources

In 2011, the National Development and Reform Commission approved seven provinces and cities, including Beijing, Tianjin, Shanghai, Chongqing, Hubei, Guangdong, and Shenzhen, to carry out carbon emission trading pilot projects. In July 2021, the national carbon emission trading market is launched and a trading center is set up in Shanghai. In this paper, the actual data of the Shanghai carbon market is used as the sample for empirical analysis, and the data of the Shanghai carbon market from 2017 to 2021 is selected as the sample, as shown in Fig 2. The dataset was obtained from the Wind database. The sample interval of the Shanghai carbon market is from January 1, 2017, to December 31, 2021, with a total of 737 observations, and the sample data are divided into two parts: a training set and a test set, with the proportion of sample data being about 80% and 20%, respectively. The Shanghai carbon market uses 2017-2020 as the training set with 598 observations and 2021 as the test set with 139 observations. The training set is used to train the GWO-LSSVR model, and the test set is used to test the prediction performance of the model.

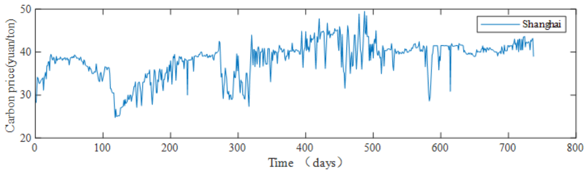

Fig 2. The original carbon price of Shanghai

3.2. Model accuracy assessment

To judge the prediction effect of each model more intuitively, three commonly used error evaluation indicators are selected in this paper to measure the performance of the prediction models and the accuracy of the prediction results, and the three evaluation indicators used are mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE), as follows:

$$MAE = \frac{\sum_{t=1}^{N} \left| y_t - \hat{y}_t \right|}{N} \tag{15}$$

$$MAPE = \frac{\sum_{t=1}^{N} \left| y_t - \hat{y}_t \right| / y_t}{N} \tag{16}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(y_t - \hat{y}_t \right)^2} \tag{17}$$

where y_t and $\hat{y}_t$ respectively represent the real value and predicted values in the period t. Smaller MAE, MAPE, and RMSE indicate better prediction performance and higher prediction accuracy.

3.3. Prediction process

3.3.1. Analysis of decomposition

From Fig 2, we can see that the Shanghai carbon market price curve shows irregular fluctuations and has obvious nonlinear characteristics, so CEEMD is used to perform adaptive decomposition of the original Shanghai carbon market price series. Eight subsequences and one residual sequence were obtained, noted as IMF1, IMF2, IMF3, IMF4, IMF5, IMF6, IMF7, IMF8, and R. The results are shown in Fig 3. The frequency of IMF1 is the highest, and the frequencies of IMF2, IMF3, IMF4, and IMF5 gradually decrease. It is generally believed that the noise is mainly concentrated in the high-frequency IMF component, and the low-frequency component is less affected by the noise. However, removing the noise would significantly reduce the accuracy of the predictions, so all IMF components were retained. Compared with the original carbon price series, the IMF obtained by decomposing all has a simpler structure, smoother volatility, and stronger regularity, which is conducive to modeling analysis.

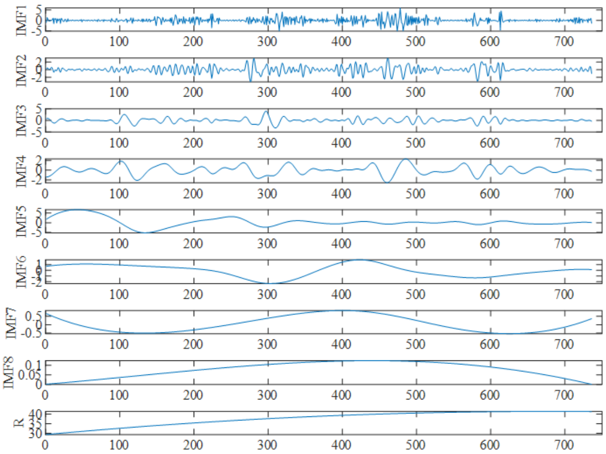

Fig 3. Extracted IMFs of Shanghai carbon prices by CEEMD

3.3.2. Analysis of empirical results

Based on the existing "decomposition-forecasting-ensemble" modeling idea, the proposed hybrid carbon price forecasting model can be represented as CEEMD-GWO-LSSVR, where "CEEMD" indicates that the method decomposes the original carbon price series; The "GWO-LSSVR" indicates that the LSSVR model is optimized using the grey wolf optimization algorithm as the forecasting model. Three sets of comparative experiments were designed to fully validate the validity and feasibility of the proposed model:

(1) Individual model comparison. To verify the effectiveness of the proposed prediction model using the decomposition method, five single models, ARIMA, BP, SVR, LSSVR, and GWO-LSSVR, are used to compare with the model proposed in this paper. These five single models do not perform data preprocessing on the original series and directly predict the original series.

(2) Comparison of algorithm-optimized prediction models. To highlight the advancement of the prediction model LSSVR using the GWO algorithm optimization, the model CEEMD-LSSVR and the model CEEMD-PSO-LSSVR are used to compare with the model proposed in this paper. These two models keep the decomposition method and prediction model unchanged, and only the optimization strategies for the prediction model are different. Model CEEMD-LSSVR does not use the optimization algorithm to optimize LSSVR, and model CEEMD-PSO-LSSVR uses the PSO algorithm to optimize LSSVR.

(3) Comparison of different decomposition methods. To demonstrate the effectiveness of using the CEEMD method in the model proposed in this paper, the model EMD-GWO-LSSVR and the model EEMD-GWO-LSSVR are used to compare with the model proposed in this paper. These two models keep the prediction model unchanged, and use EMD and EEMD as decomposition methods for the original series of carbon prices, respectively.

Table 1 shows the comparison of the prediction results between the proposed model and the comparison model in this paper

Table 1. The errors of different forecasting models

Models	MAE	MAPE	RMSE
ARIMA	1.7560	0.0420	2.1147
BP	1.1706	0.0288	1.7018
SVR	1.2960	0.0319	1.7350
LSSVR	1.1363	0.0282	1.8140
GWO-LSSVR	0.7984	0.0213	1.3932
CEEMD-LSSVR	0.4960	0.0132	0.9389
CEEMD-PSO-LSSVR	0.4295	0.0119	0.7463
EMD-GWO-LSSVR	0.4312	0.0127	0.7945
EEMD-GWO-LSSVR	0.4232	0.0114	0.7194
CEEMD-GWO-LSSVR	0.4185	0.0109	0.6795

By analyzing the three prediction evaluation indicators of MAE, MAPE, and RMSE in Table 1, it is seen that the three indicators of MAE, MAPE, and RMSE of the model CEEMD-GWO-LSSVR proposed in this paper outperform all the comparative models in the Shanghai carbon market, which reflects the effectiveness and advancement of the model proposed in this paper.

(1) In the single model comparison experiment, the comparison between five single models ARIMA, BP, SVR, LSSVR, and GWO-LSSVR can be found that the prediction errors of models BP, SVR, LSSVR, and GWO-LSSVR as artificial intelligence models are smaller than those of the econometric model ARIMA, and the prediction effect of the traditional econometric model ARIMA This is because the econometric model mainly portrays the linear characteristics of the series, while the original series of the carbon price has non-linear, non-smooth and other complex characteristics making it difficult for the traditional econometric model to capture the characteristics of the original series and fail to achieve a better prediction effect. Among the remaining four single AI models, the model GWO-LSSVR has better prediction results, which indicates that the GWO algorithm finds more suitable parameters for LSSVR. The prediction results of the above five single models also illustrate from the side that for non-stationary and non-linear series, AI models have more advantages over traditional econometric models.

(2) In the comparison experiments of the algorithm optimization prediction model, by model CEEMD-LSSVR, model CEEMD-PSO-LSSVR and the model proposed in this paper, the three evaluation indexes MAE, MAPE, and RMSE of the model proposed in this paper compared the two models, the prediction results show that GWO-LSSVR prediction model is better than the randomly selected parameters of LSSVR model, and the algorithm GWO optimization prediction model LSSVR is better than the algorithm PSO optimization prediction model LSSVR, which shows that the use of GWO algorithm optimization prediction model LSSVR in this paper can effectively avoid the blindness of human subjective selection of parameters.

(3) In the comparison experiments of different decomposition methods, the model EMD-GWO-LSSVR and the model EEMD-GWO-LSSVR are compared with the model proposed in this paper, and the model proposed in this paper outperforms the two compared models in three evaluation indexes MAE, MAPE, and RMSE, indicating that the CEEMD decomposition method, compared with the EMD and EEMD decomposition methods, can better It shows that the CEEMD decomposition method can better extract the multi-scale characteristics of the Shanghai carbon market

price than the EMD and EEMD decomposition methods.

4. Conclusions

Carbon trading is the most cost-effective tool to reduce carbon emissions, and the accurate prediction of carbon trading prices can provide a theoretical reference for the development of carbon market construction. To fully extract the characteristics of non-stationary and non-linear carbon prices as well as multi-scale, this paper takes the daily closing price of the Shanghai carbon market as the research object and constructs an integrated hybrid prediction model based on the multi-scale decomposition of carbon prices. The research findings of this paper are as follows:

(1) The CEEMD decomposition method is very effective in improving the performance of the hybrid model, which can overcome the modal confounding problem of some decomposition methods, and also can fully extract the multiscale features of the original carbon price series and decompose the data into more stable components. The prediction results show that the prediction performance of the model with the decomposition method is significantly better than that of the model without the decomposition method.

(2) To address the problem that the LSSVR model requires human-set parameters, this paper uses the GWO algorithm to optimize the parameters of LSSVR to obtain the optimal parameter combinations, which substantially reduces the influence of subjectivity of traditional manual assignment and has superior compared with the PSO algorithm.

(3) The decomposition integrated hybrid prediction model based on CEEMD-GWO-LSSVR proposed in this paper performs well in prediction, and the prediction accuracy is substantially improved compared with that of the control model, indicating the effectiveness and superiority of the proposed model in carbon price prediction problems.

Although the decomposition ensemble hybrid prediction model proposed in this paper has high prediction accuracy, the model only takes the historical carbon price as input and does not consider the influence of external factors such as national policies, regional economic development, and energy prices on the carbon price. In future research, we can also consider adding the influence of urea to the prediction to explore the influence of data other than historical prices on the prediction results and further improve the model's prediction performance.

References

[1] Zhao C, Liu B, Piao S, et al. Temperature increase reduces global yields of major crops in four independent estimates[J]. Proceedings of the National Academy of Sciences, 2017, 114(35): 9326–9331.

[2] Zhang Y J, Wei Y M. An overview of current research on EU ETS: Evidence from its operating mechanism and economic effect[J]. Applied Energy, 2010, 87(6): 1804-1814.

[3] Guðbrandsdóttir H N, Haraldsson H Ó. Predicting the price of EU ETS carbon credits[J]. Systems Engineering Procedia, 2011, 1: 481-489.

[4] Byun S J, Cho H. Forecasting carbon futures volatility using GARCH models with energy volatilities[J]. Energy Economics, 2013, 40: 207-221.

[5] Benz E, Trück S. Modeling the price dynamics of CO2 emission allowances[J]. Energy Economics, 2009, 31(1): 4-15.

[6] Sheng C, Wang G, Geng Y, et al. The correlation analysis of futures pricing mechanism in China’s carbon financial market[J]. Sustainability, 2020, 12(18): 7317.

[7] Fan X, Li S, Tian L. Chaotic characteristic identification for carbon price and an multi-layer perceptron network prediction model[J]. Expert Systems with Applications, 2015, 42(8): 3945-3952.

[8] Yi L, Li Z, Yang L, et al. The Scenario Simulation Analysis of the EU ETS Carbon Price Trend and the Enlightenment to China [J][J]. Journal of Environmental Economics, 2017, 2(3): 22-35.

[9] Zhu, B., Wei, Y., 2013. Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. Omega-international Journal of Management Science 41 (3), 517–524.

[10] Shouyang W, Lean Y U, Lai K K. Crude oil price forecasting with TEI@ I methodology[J]. Journal of Systems Science and Complexity, 2005, 18(2): 145.

[11] Yang S, Chen D, Li S, et al. Carbon price forecasting based on modified ensemble empirical mode decomposition and long short-term memory optimized by improved whale optimization algorithm[J]. Science of the Total Environment, 2020, 716: 137117.

[12] Chai S, Zhang Z, Zhang Z. Carbon price prediction for China's ETS pilots using variational mode decomposition and optimized extreme learning machine[J]. Annals of Operations Research, 2021: 1-22.

[13] Zhu B, Ye S, Wang P, et al. A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting[J]. Energy Economics, 2018, 70: 143-157.

[14] Wang J, Sun X, Cheng Q, et al. An innovative random forest-based nonlinear ensemble paradigm of improved feature extraction and deep learning for carbon price forecasting[J]. Science of the Total Environment, 2021, 762: 143099.

[15] Wang W B, Fei P S, Yi X M. Chinese stock market forecasting based on EMD and neural network[J]. Systems Engineering Theory and Practice, 2010, 30(6): 1027-1033.

[16] Rayi V K, Mishra S, Naik J, et al. Adaptive VMD based optimized deep learning mixed kernel elm autoencoder for single and multistep wind power forecasting[J]. Energy, 2022, 244: 122585.

[17] Li H, Jin F, Sun S, et al. A new secondary decomposition ensemble learning approach for carbon price forecasting[J]. Knowledge-Based Systems, 2021, 214: 106686.

[18] Wang S, Gu Y, Uchida H, et al. A hybrid short-term traffic flow forecasting method based on EMDW-LSSVM[C]//2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2020: 1-6.

[19] Habib A, Abbassi R, Aristizábal A J, et al. Retracted: Forecasting model for wind power integrating least squares support vector machine, singular spectrum analysis, deep belief network, and locality-sensitive hashing[J]. Wind Energy, 2020, 23(2): 235-257.

[20] Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in adaptive data analysis, 2009, 1(01): 1–41.

[21] Yeh J-R, Shieh J-S, Huang N E. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method[J]. Advances in adaptive data analysis, 2010, 2(02): 135–156.

[22] Suykens J A, Vandewalle J. Least squares support vector machine classifiers[J]. Neural processing letters, 1999, 9(3): 293–300.

[23] Jianwei E, Ye J, He L, et al. A denoising carbon price forecasting method based on the integration of kernel independent component analysis and least squares support vector regression[J]. Neurocomputing, 2021, 434: 67–79.

[24] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. Advances in engineering software, 2014, 69: 46-61.