Applied Mathematics, Modeling and Computer Simulation
C.-H. Chen et al. (Eds.)

doi:10.3233/ATDE231006

A Novel Hybrid Model for Forecasting China Carbon Price Using CEEMDAN and Extreme Learning Machine Optimized by Whale Algorithm

Li NI[a,b,1],Venus Khim-Sen Liew[a]
[a] *Faculty of Economics and Business of Universiti Malaysia Sarawak, Sarawak, 94300, Malaysia*
[b] *School of Finance and Taxation of Anhui Economics & Management College, Hefei, 230601, China*

Abstract. The carbon market can provide economic incentives for manufacturing industry to reduce carbon emissions. This paper follows the idea of "primary decomposition- noise reduction-secondary decomposition- forecasting and integration", the contribution is constructing a hybrid carbon price forecasting model using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and Extreme Learning Machine (ELM) optimized by the Whale Optimization Algorithm (WOA). The results conclude that, the CEEMDAN-type secondary decomposition hybrid models have high forecasting accuracy, the WOAELM-type models can effectively reduce the forecasting errors. Noteworthy, the forecasting errors RMSE, MAE and MAPE of the proposed CEEMDAN-SE-CEEMD-WOAELM model are 2.587, 2.04 and 0.108 respectively, that is the lowest in all the comparative models. The forecasting accuracy and reliability of the proposed model have been convinced. Those findings can provide valuable reference for manufacturing industry to reduce pollutant emissions and take low-carbon investment.

Keywords. Carbon price forecasting, CEEMDAN, extreme learning machine, whale optimization algorithm, sample entropy

1. Introduction

The carbon market is a mean of utilizing resource allocation mechanism to reduce carbon emissions. The price mechanism is the carbon market core, as a result, studying the price formation of the carbon market, forecasting carbon price accurately are the keys to reduce the pollutant emissions in a low-cost manner.

The EMD (Empirical Mode Decomposition, EMD) technology have proved to be an effective model to capture the non-linear and non-stationary characteristic of carbon price [1-3]. Conducted the EMD technology to decompose the carbon price, used the particle swarm optimization least squares support vector machine (PSO-LSSVM) model for out-of-sample forecasting, the results suggest the EMD-PSO-LSSVM model has forecasting superiority in Europe carbon price [4-5].To reduce the decomposition

1 Corresponding Author, Li NI, School of Finance and Taxation of Anhui Economics & Management College, Hefei 230601 China; E-mail:21010257@siswa.unimas.my.

noise, the CEEMDAN and VMD technologies were used to perform primary and secondary decomposition of the original price, LSTM and ELM models were used for forecasting. The findings put that the CEEMDAN-VMD-LSTM and CEEMDAN-VMD-ELM models have substantially forecasting accuracy in China carbon market [6-8].

Generally, the traditional EMD-type technologies have defects such as mode mixing and large reconstruction errors. In particular, the high complexity signals before the secondary decomposition needs to be strictly identified rather than empirically determined. The contribution of this paper is constructing a hybrid carbon price forecasting model using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and Extreme Learning Machine (ELM) model optimized by Whale Optimization Algorithm (WOA). Namely, the model of CEEMDAN-SE-CEEMD-WOAELM, that is, conduct the CEEMDAN technology to primary decompose the carbon price, use Sample Entropy to identify the high complexity signals, employ the CEEMD (Complementary Ensemble Empirical Mode Decomposition,CEEMD) to secondary decompose the high complexity signals, perform the WOAELM model to conduct the out-of-sample forecasting of the obtained signal components, and finally acquire the predicted carbon price.

2. Methodology

2.1. CEEMDAN Model

CEEMDAN is an improved technology for decomposing the non-linear and non-stationary data into different multi-scale signals. The obvious feature of CEEMDAN is that the noise addition is the IMF (Intrinsic Mode Function,IMF) components decomposed by EMD process, rather than directly adding Gaussian white noise to the original signals. This style of white noise addition can solve the problems of mode mixing and large reconstruction errors [9]. The idea of CEEMDAN is as follows:

Step 1: Add K times (K>1) white noise series $v_i(t)$ (i=1,2,3...N) with a mean value of 0 and a constant standard deviation to the original signal x(t), then get a new signal $x'(t)$:

$$x_i'(t)=x_i(t)+\varepsilon v_i(t) \tag{1}$$

Where, ε represents the weight coefficient of Gaussian white noise, $v_i(t)$ is the Gaussian white noise added in the i-th time.

Step 2: Perform the EMD process on the noise-added signal $x'(t)$ to obtain the first mode components, then take the mean of mode components as the first IMF of CEEMDAN decomposition.

$$IMF_1(t)=\frac{1}{K}\sum_{i=1}^{K} IMF_1^i(t) \tag{2}$$

$$r_1(t)=x(t)-IMF_1(t) \tag{3}$$

Where, $IMF_1(t)$ represents the first IMF signal obtained by CEEMDAN decomposition, $IMF_1^i(t)$ means the first IMF signal obtained by EMD decomposition, $r_1(t)$ is the residual .

Step 3: Continue to add white noise to the j-th residual, then perform EMD process to acquire the following components:

$$IMF_j(t)=\frac{1}{K}\sum_{i=1}^{K}E_1(r_{j-1}(t)+\varepsilon_{j-1}E_{j-1}(v_i(t))) \tag{4}$$

$$r_j(t)=r_{j-1}(t)-IMF_j(t) \tag{5}$$

Among them, $IMF_j(t)$ indicates the component obtained after the j-th decomposition, $E_j(\cdot)$ means the j-th IMF component, ε_j represents the white noise weight coefficient added to the j-th order residual, $r_j(t)$ is the jth-order residual.

2.2. Sample Entropy

Sample Entropy is a classic method of time series complexity. Its essence is to calculate the probability of input signal generating new patterns, the higher the probability, the higher the signals complexity. For a time series $\{y_m\}\{y_1, y_2, \cdots, y_M\}$,the sample entropy is calculated as follows:

$$SE = SampEn(n,s,M) = -\ln\left[\frac{A^{(n)}(s)}{B^{(n)}(s)}\right] \tag{6}$$

Among them, A and B represent the number of j whose distance between series y_{mi} and y_{mj} is less than or equal to s. n means the dimension of input vector. In this paper, if the sample entropy is greater than 1, we believe the signal complexity is high and require secondary decomposition process [10].

2.3. WOAELM

Extreme Learning Machine (ELM) is a feed-forward neural network (SLFN) with superior generalization and fast learning ability. The training speed of the ELM model is faster, and it is not easy to fall into local optimal solution. In this paper, the ELM model is used to forecast each IMF component, and the network is optimized by the Whale Optimization Algorithm (WOA) . The WOA is originated from the special bubble net foraging method of humpback whales, the algorithm has the advantages of simple operation and less parameter adjustment. The process of WOA is as follows:

Step 1: Encircling prey. The WOA assumes the current optimal alternative solution is prey. After determining the location of the prey, the remaining search agents will move towards the optimal search agent direction and update the position:

$$D = \left| C \bullet X^*(t) - X(t) \right| \tag{7}$$

$$X(t+1) = X^*(t) - A \bullet D \tag{8}$$

Among them, t represents the iteration numbers; $X^*(t)$ is the current optimal solution, $X(t)$ is the position vector, $|\ |$ means the absolute process, and $\bullet$ represents the elements multiplication. A and C are the coefficient vectors. D is the distance between the whale and its prey (optimal solution). During the iteration process, a decreases from 2 to 0.

Step 2: Bubble net trapping. There are two main processes for humpback whales to prey on. When using a bubble net for predation, the position update between the humpback whale and its prey is expressed as:

$$X(t+1) = D' \times e^{bc} \times \cos(2\pi c) + X^*(t) \tag{9}$$

$$D' = \left| X^*(t) - X(t) \right| \tag{10}$$

Where D' indicates the distance between the current search agents and the optimal solution; b represents the spiral shape parameter; c is a random number that follows a mean distribution of [-1,1].

Furthermore, the search agents select surrounding predation and bubble net predation based on probability p, and the position updates are as follows:

$$X(t+1) = \begin{cases} X^*(t) - A \bullet D & p \le 0.5 \\ D' \times e^{bc} \times \cos(2\pi c) + X^*(t) & p > 0.5 \end{cases} \tag{11}$$

Where, p is the probability of the predator mechanism with a range of random numbers between 0 and 1. As the iterations t increases, the parameter A gradually decrease. Each whale gradually approaches the current optimal solution if $|A| < 1$.

Step 3: Search for prey. To ensure all the whales can fully search in the solution space, WOA updates the position based on the distance among the whales population. Therefore, when $|A| \ge 1$, the search agents will swim towards other random whales.

$$D'' = \left| C \bullet X_{rand}(t) - X(t) \right| \tag{12}$$

$$X(t+1) = X_{rand}(t) - A \bullet D \tag{13}$$

In above formula, D'' is the distance between the current searched agents and the random agents, Xrand (t) indicates the position of the current random agents.

2.4. The proposed CEEMDAN-SE-CEEMD-WOAELM model

The specific process of the proposed model is as follows:firstly, CEEMDAN technology is used to decompose the carbon price. Secondly, measures the sample entropy of each IMF component to determine the high complexity signals. Thirdly, conducts the CEEMD technology for secondary decomposition the high complexity signals to reduce decomposition noise. Fourthly, the ELM model optimized by whale algorithm, namely WOAELM model is used for out-of-sample forecasting. Fifthly, the final predicted carbon price can be obtained by summing the predicted IMF signals.

For evaluating the forecasting performance, the comparative models based on EEMD (Ensemble Empirical Mode Decomposition,EEMD) and EMD technologies are constructed. The indicators such as RMSE (Root Mean Squared Error, RMSE), MAE (Mean Absolute Error,MAE) ,MAPE (Mean Absolute Percentage Error,MAPE) and Pearson correlation are used as the evaluation standard.

3. Empirical Analysis and Discussion

3.1. Research Data and Basic Statistics

As the first pilot carbon market in China, the obvious advantage of Shenzhen carbon market lies in its high liquidity. This article selects the Shenzhen carbon price as the research object, with a time range of January 3, 2018 to May 31, 2023, a total of 1146 data are obtained. The data sourced from the Shenzhen Carbon Emission Rights Exchange (https://www.szets.com). In addition, we take the first 70% of the sample as the training set, and the remaining data as the test set.

Based on Table 1, it is evident that the mean and variance of the test set are both the highest, with the values of 36.287 and 20.178 respectively. The mean and variance of the training set are the smallest, with the values of 19.09 and 9.759 respectively. Those findings indicate the volatility of the training set data is stable, while the test set may has high price risk. Furthermore, the results of the JB statistic and ADF statistic also show the research sample have non-linear and non-stationary characteristic.

Table 1. Basic Statistics of the sample

	Meam	Std.Dev.	Skewness	Kurtosis	Jarque-Bera Statistic	ADF	Observations
The whole sample	24.761	16.392	0.865	2.706	146.954***	1.258	1146
Training set	19.090	9.759	0.369	1.966	53.549***	3.273	796
Test set	36.287	20.178	-0.271	1.415	38.917***	0.438	350

3.2. Decomposition of the Carbon Price Signals

3.2.1. Primary Decomposition Process of the Carbon Price by CEEMDAN Model

This article conducts CEEMDAN technology to primary decompose the carbon price. After the decomposition, 10 IMF components and 1 residual are obtained. It can be found that the frequency and period of each mode are quite different as shown in Table 2. Specifically, the signal periods of IMF1-IMF4 are shorter, the Pearson correlation with the original carbon price is weak, and the variance explanation for carbon price is the worst, those evidence demonstrate that these signals fluctuate more frequently, and the signal noise is larger, make it difficult to provide effective explanations for carbon premiums. Additionally, the sample entropy of IMF1-IMF4 are greater than 1, that means these signals have high complexity and require further secondary decomposition to reduce decomposition noise.

Table 2. Basic statistical characteristics of IMFs signals decomposed by CEEMDAN model

IMFs	Periods	Correlation with carbon price	VAR	Sample Entropy
IMF1	2.788	0.195	0.033	1.456
IMF2	4.182	0.131	0.002	1.792
IMF3	4.064	0.118	0.007	2.442
IMF4	7.253	0.121	0.008	1.653
IMF5	14.883	0.144	0.007	0.836
IMF6	32.743	0.195	0.018	0.530
IMF7	71.625	0.245	0.017	0.444
IMF8	191	0.262	0.043	0.205
IMF9	286	0.784	0.218	0.046
IMF10	382	0.826	0.045	0.030
res	1146	0.679	0.161	0.010

3.2.2. Secondary Decomposition Process of the carbon price by CEEMD Model

The CEEMD technology is used to secondary decompose the high complexity signals of IMF1-IMF4, as a result (table 3), 18 IMF components can be obtained with an average sample entropy of 0.259 and a maximum sample entropy of 0.836. While only 11 IMF components are acquired with an average sample entropy of 0.859 and a maximum value of 2.442 in the primary decomposition process. Therefore, the sample entropy of each mode after the secondary decomposition are greatly reduced. Those conclude that although the secondary decomposition increase the mode components, the signal decomposition errors and signal complexity are greatly reduced.

Table 3. Comparison of the Sample Entropy value before and after the secondary decomposition

	Model	IMF numbers	Average Sample Entropy	Maximum value of Sample Entropy
Secondary decomposition	CEEMDAN-SE-CEEMD	18	0.259	0.836
	EEMD-SE-CEEMD	19	0.264	0.705
	EMD-SE-CEEMD	19	0.318	0.906
Primary decomposition	CEEMDAN	11	0.859	2.442
	EEMD	10	0.489	1.524
	EMD	9	0.532	1.464

3.3. Out-of-sample forecasting of carbon price by CEEMDAN-SE-CEEMD-WOAELM

3.3.1. Findings on the Superiority of ELM-Type Models Optimized by Whale Algorithm

The WOAELM-type hybrid models can reduce carbon price forecasting errors and improve the correlation between predicted price and the actual carbon price. For example, according to Table 4, as for the secondary decomposition forecasting models, the average forecasting errors RMSE, MAE and MAPE of the WOAELM-type models are 6.762, 4.765 and 0.128 respectively, with a correlation of 0.984. While the other models that without optimized by the whale algorithm have high forecasting errors, with average forecasting errors of RMSE, MAE and MAPE are 10.413, 8.217 and 0.197 respectively, and the correlation is 0.983. Those findings reveal that the whale algorithm used in this paper enhance the forecasting performance and accuracy of the ELM-type models greatly. In addition, similar conclusions can be obtained in the primary decomposition carbon price forecasting models.

3.3.2. Findings on the Superiority of the Secondary Decomposition Hybrid Models

According to the Table 4, the secondary decomposition carbon price forecasting models have low forecasting errors and superior forecasting performance, the forecasting accuracy and reliability have been convinced.

For example, the average forecasting errors RMSE, MAE and MAPE of the secondary decomposition hybrid models are 8.588, 6.491 and 0.162 respectively, with a correlation of 0.98. However, as for the primary decomposition hybrid models, the forecasting errors RMSE, MAE and MAPE are relatively high, with 14.134, 11.648 and 0.316 respectively, the correlation is 0.958. Therefore, we can prove that the secondary decomposition hybrid models have superior forecasting performance than other primary decomposition models. One reason is that the IMF components obtained after the secondary decomposition are all components with lower decomposition errors. Based on this, the out-of-sample forecasting errors can be reduced.

Table 4. The out-of -sample forecasting errors of proposed model and its comparative models

Model types	Models	RMSE	MAE	MAPE	Correlation
Secondary decomposition forecasting models	CEEMDAN-SE-CEEMD-WOAELM	2.587	2.04	0.108	0.993
	EEMD-SE-CEEMD-WOAELM	9.577	6.435	0.138	0.975
	EMD-SE-CEEMD-WOAELM	8.123	5.819	0.138	0.983
	Mean errors 1	6.762	4.765	0.128	0.987
	CEEMDAN-SE-CEEMD-ELM	8.236	6.453	0.166	0.984
	EEMD-SE-CEEMD-ELM	14.196	11.325	0.265	0.973
	EMD-SE-CEEMD-ELM	8.806	6.872	0.159	0.989
	Mean errors 2	10.413	8.217	0.197	0.983
	Mean errors of secondary decomposition models	8.588	6.491	0.162	0.980
Primary decomposition forecasting models	CEEMDAN-WOAELM	8.403	6.905	0.171	0.991
	EEMD-WOAELM	10.247	7.723	0.179	0.967
	EMD-WOAELM	17.228	15.083	0.459	0.968
	Mean errors 3	11.959	9.904	0.270	0.975
	CEEMDAN-ELM	9.34	6.504	0.135	0.971
	EEMD-ELM	14.452	12.075	0.318	0.976
	EMD-ELM	25.133	21.595	0.634	0.873
	Mean errors 4	16.308	13.391	0.362	0.940
	Mean errors of primary decomposition models	14.134	11.648	0.316	0.958

3.3.3. Findings on the Superiority of the CEEMDAN-Type Hybrid Models

The forecasting errors of the CEEMDAN-type models are obviously smaller than that of other decomposition technologies hybrid models. Those evidence suggest that the CEEMDAN can improve the signal decomposition efficiency and reduce the decomposition errors.

As for the secondary decomposition hybrid forecasting models, the forecasting errors RMSE, MAE and MAPE of the CEEMDAN-SE-CEEMD-WOAELM model constructed in this paper are 2.587, 2.04 and 0.108 respectively, which are obviously lower than the forecasting errors of the EEMD-SE-CEEMD-WOAELM and EMD-SE-CEEMD-WOAELM models. Meanwhile, the forecasting errors RMSE MAE and MAPE of the CEEMDAN-SE-CEEMD-ELM model are 8.236, 6.463 and 0.1668 respectively, which are also obviously lower than the errors of the EEMD-SE-CEEMD-ELM and EMD-SE-CEEMD-ELM models.

3.3.4. Findings on the Superiority of the CEEMDAN-SE-CEEMD-WOAELM Model

As depicted in above analysis, the forecasting errors of the proposed CEEMDAN-SE-CEEMD-WOAELM model is the lowest compared with other comparative models, and the correlation is the maximum value in all of the models. Those findings demonstrate that the proposed model can decompose the original carbon price into multi-scales signals accurately, and can also make reliable out-of-sample forecasting. Furthermore, the error deviation between the predicted value and the true value is negligibly small (as shown in Figure 1 and Figure 2). Therefore, the proposed model has proved to be an excellent model for forecasting China carbon price, that can provide references for investors to analyse market conditions and conduct price analysis.

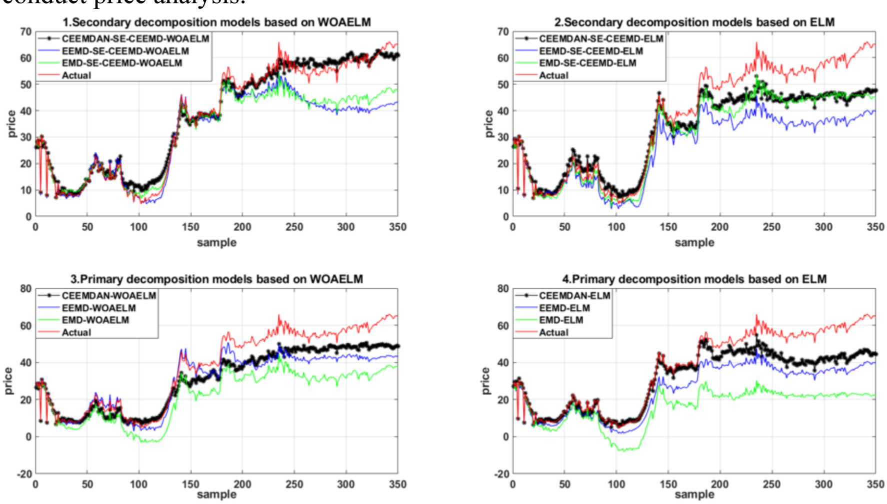

Figure 1. The out-of-sample fitting performance of the proposed model and its comparative models

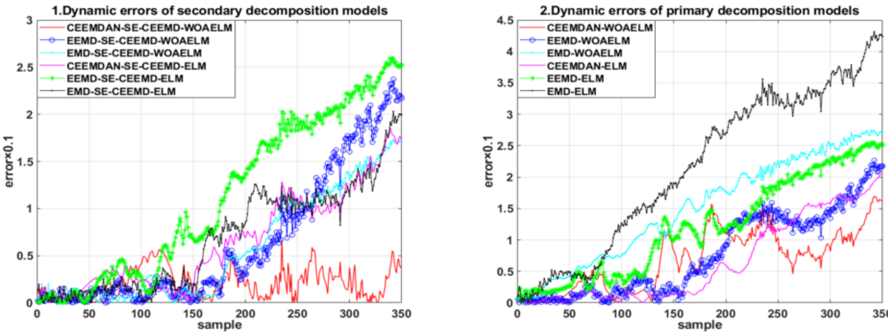

Figure 2. The dynamic forecasting errors of the proposed model and its comparative models

4. Conclusion

This paper focuses on the price forecasting of the emerging carbon market in China. The contribution of this paper is constructing a novel hybrid model for forecasting China carbon price using CEEMDAN and extreme learning machine optimized by whale algorithm. The results conclude that the secondary decomposition carbon price forecasting models, especially the hybrid models based on CEEMDAN technology have high forecasting accuracy, the WOAELM-type models can reduce the forecasting errors effectively. Noteworthy, the forecasting accuracy of the proposed CEEMDAN-SE-CEEMD-WOAELM model is substantially better than other comparative models.

Acknowledgements

This work has been supported by the University Humanities and Social Sciences Research Project of Anhui Province of China (Grant No. SK2021A1105), the scientific projects of Anhui Finance & Trade Vocational College (Grant No. tzpyxj123). The authors also acknowledge UNIMAS for its research facilities provided.

References

[1] Yun P, Huang X, Wu Y, et al. Forecasting carbon dioxide emission price using a novel mode decomposition machine learning hybrid model of CEEMDAN-LSTM[J]. Energy Science & Engineering, 2023,11(1): 79-96.

[2] Li Y L, Zhang J J, Fang Y.The driving factors of China's carbon prices: Evidence from using ICEEMDAN-HC method and quantile regression[J]. Finance Research Letters, 2023,54, 103756.

[3] Li J, Liu D. Carbon price forecasting based on secondary decomposition and feature screening[J]. Energy, 2023,278, 127783.

[4] 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.

[5] Zhang W, Wu Z. Optimal hybrid framework for carbon price forecasting using time series analysis and least squares support vector machine[J]. Journal of Forecasting,2022,41(3): 615-632.

[6] 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.
[7] Zhou J, Wang Q. Forecasting Carbon Price with Secondary Decomposition Algorithm and Optimized Extreme Learning Machine[J]. Sustainability,2021, 13:8413.
[8] Zhou F, Huang Z, Zhang C. Carbon price forecasting based on CEEMDAN and LSTM[J]. Applied Energy,2022,311:118601.
[9] Yang H, Yang X, Li G.Forecasting carbon price in China using a novel hybrid model based on secondary decomposition, multi-complexity and error correction[J]. Journal of Cleaner Production, 2023,401:136701.
[10] Huang G B, Wang D H, Lan Y. Extreme learning machines: a survey[J]. International Journal of Machine Learning and Cybernetics, 2011,2:107-122.