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

doi:10.3233/ATDE221078

Carbon Price Prediction Using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise Analysis and Convolutional Neural Network

YuanLing WANG[1], Zhengjie YAN, Yun BAI

School of Management Science and Engineering, Chongqing Technology and Business University, Chongqing 400067, China

Abstract. China has launched national carbon trading marked in 2021, and up to now, Hubei has the largest proportion of carbon trading volume, it is totally important to research the carbon trading price in Hubei. In this paper, we propose a new model for carbon price in Hubei, which is combine complete ensemble empirical mode decomposition with adaptive noise analysis (CEEMDAN) with convolutional neural network (CNN). Firstly, carbon price is decomposed by CEEMDAN into various intrinsic mode function (IMF) which are combined using sample entropy approach. Then, CNN is used to establish a point prediction model. Finally, we calculate the mean square (RMSE), mean absolute error (MAE), and coefficient of determination (R^2) of the model are 0.4893, 0.6809, and 0.9754, respectively. Compare with other two models, the hybrid model proposed in this paper exhibits the best performance.

Keywords. Carbon price; Complete ensemble empirical mode decomposition with adaptive noise analysis; Sample entropy; Convolutional neural network

1. Introduction

Environmental problems caused by carbon emissions have always been the focus [1] of attention from countries around the world. China has launched the carbon trading market in 2013, and the national carbon trading market had opened in 2021 [2]. Through the way of carbon trading, the enterprises' carbon emissions can be reduced, in order to achieve a better emission reduction effect [3]. The Hubei carbon trading market has been opened since 2014, and as of March 2021, 1583 carbon trading data points have been generated.

Due to the carbon trading prices are affected by energy prices, exchange rates, and policies [4-5], carbon price data are nonstationary and nonlinear, so traditional single forecasting models and machine models have poor forecasting effects [6-8]. In recent years, deep learning algorithms have become more and more mature and are widely used in the field of data analysis [9]. More and more scholars have applied deep learning algorithms to forecast carbon trading price [10-12]

This study will completely combine complete ensemble empirical mode decomposition with adaptive noise analysis and convolution neural network. The

[1] Corresponding Author, Yuanling WANG, School of Management Science and Engineering, Chongqing Technology and Business University, Chongqing 400067, China; E-mail:1983308234@qq.com

CEEMDAN solve the EMD decomposition and the remaining white noise problem [13], in recent years, CEEMDAN is widely used in the carbon price prediction. CNN is the most typical deep learning intelligent algorithm [14], it has very good effect on processing of large amounts of data. The combination of CEEMDAN and CNN not only solves the problem of nonstationary and nonlinear of carbon price data, but also improves the prediction accuracy.

2. Methods

2.1 Convolutional Neural Network (CNN)

CNN is really common in the field of image, however, in recent years, it has become more and more common in the field of data analysis. CNN consists of an input layer, a convolution layer, a pooling layer, a fully connected layer and an output layer. This paper uses a one-dimensional convolution layer. The pooling layer connects the convolution layer and we applicate the largest pooling method to extract effective feature information for the carbon price in Hubei. The fully connects layer is connected to the pooling layer, which integrates data features.

2.2 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise Analysis (CEEMDAN)

Compare with the empirical mode decomposition (EMD), CEEMDAN adds white noise sequence, and the error is smaller than EEMD [15]. In this paper, we use CEEMDAN to decompose the series of carbon price, the main steps are as follows:

1. Introduce the original series: $x(t)+p_0 n_j$. The original series is decomposed to obtain the first modal component.

$$IMF_1(t)=\frac{1}{M}\sum_{j=1}^{M} IMF_1^j(t) \tag{1}$$

where p_0 controls the signal-to-noise ratio (SNR) of the additional noise to the original signal, and the residual signal can be expressed as follows.

$$r_1(t)=x(t)-IMF_1(t) \tag{2}$$

2. Before the next step of decomposition, introduce the first-order component of self-decomposition of white noise signal and combine with the residual signal. Then, the signal to be decomposed can be expressed as, $x(t)+p_1 emd_1(n_j(t))$ where $emd_i()$ represents the i-th modal component generated by the EMD algorithm. Continue to use the EEMD algorithm to integrate and average the second modal component.

$$IMF_2(t)=\frac{1}{M}\sum_{j=1}^{M} emd_1(r_1(t)+p_1 emd_1(n_j(t))) \tag{3}$$

3. Continue to use the method of step 2 to calculate the residual signal, and the (n+1)-th model component can be obtained as:

$$IMF_{(n+1)}(t)=\frac{1}{M}\sum_{j=1}^{M} emd_1(r_n(t)+p_n emd_n(n_j(t))) \tag{4}$$

4. We can terminate the decomposition algorithm, while the final residual signal satisfies the iterative termination condition. If the number of modal components in the final decomposition is N, the original carbon price sequence can be expressed as:

$$x(t)=\sum_{i=1}^{N} IMF_i(t)+r(t) \tag{5}$$

3. Data Analysis and Preprocessing

Since the launch of carbon trading pilot in 2013, China has developed eight carbon emission trading markets in Hubei, Guangdong, Beijing, Shanghai, Shenzhen, Tianjin, Fujian and Chongqing. By June 2021, Hubei accounted for the largest proportion of carbon trading, accounting for 32.4% (figure 1).

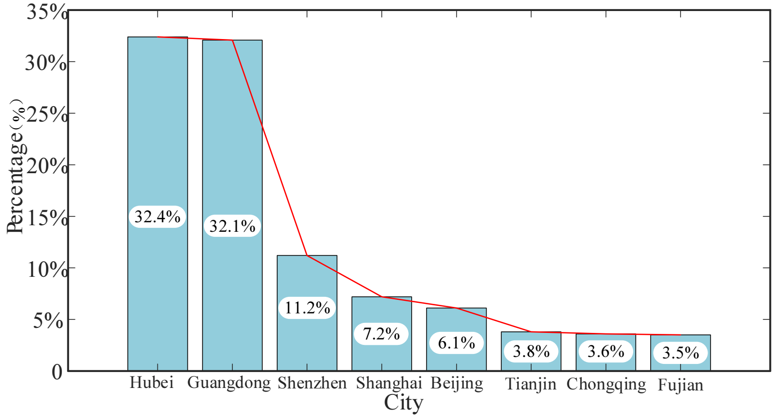


Figure 1. Ratio of carbon trading volume in China's carbon trading market

In this study, we selected carbon trading price data of Hubei during 2013 to 2021 from the China Carbon Trading Network (www.tanpaifang.com). A total of 1583 observation values were divided into training datasets (70%) and testing datasets (30%). We conducted statistical analysis of Hubei carbon price, and the results are shown in table 1.

Table 1. Statistical analysis for carbon price in Hubei (RMB: yuan/ton)

Data type	Data period	Number of data	Mean	Var	Max	Min	Median	Skew	Kurt
Carbon price	4/2014-3/2021	1583	23.41	44.21	53.85	10.07	24.23	0.34	0.20

According to the descriptive statistics of the Hubei carbon trading price series, the variance in the Hubei carbon price data was 44.21, indicating that the original data had a high degree of dispersion. The skewness equals 0, indicating that carbon trading price values of Hubei from 2014 to 2021 are evenly distributed around the mean value of 23.41. The Kurtosis equals 0.34, indicating that the data of carbon price deviates less from the mean value.

4. Experimental Algorithm Analysis

4.1 The Experiment Design

The experiment process is shown in figure 2, it can be described as follows:

Step1, the original data was decomposed into different IMFs by using CEEMADN algorithm.

Step2, we calculate the sample entropy of each IMF, and select the best IMF integration according to the value of sample entropy.

Step3, the prediction model is used to predict the integrated signal separately, then the prediction result is compared with the undecomposed data.

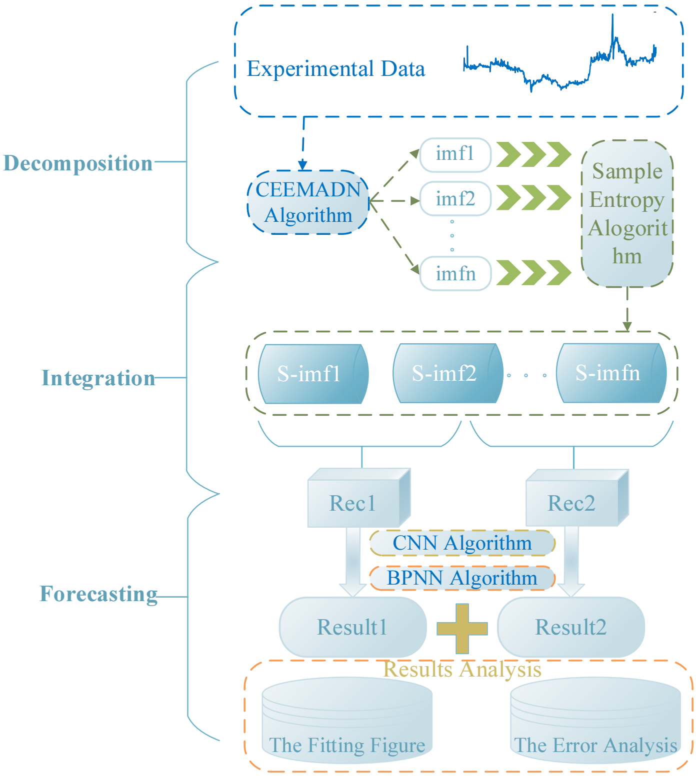


Figure 2. The modeling flowchart

4.2 The Setting of Experiment Models

After many times attempts, the parameters of CEEMDAN, CNN, BP and SE are selected. It can be found in the table 2, the Nstd, NR and MaxIter values in CEEMDAN are set to o.5, 100 and 5000 respectively, the Batch_size, Epoch and SW values in CNN are set to 32, 100 and 5 respectively, the Eta, Error, D, L and Q values in BP are set to 5, 0.5, 0.0002, 5, 1, 5 respectively, the m and r in SE are set to 5, 0.24 repectively.

Table 2. The parameters of models

Method	Symbol	Meaning	Value
	Nstd	Standard deviation of Gaussian white noise	0.5
CEEMDAN	NR	The number of times noise is added	100

	MaxIter	Maximum iteration	5000
CNN	Batch_size	The number of samples selected for a training	32
	Epoch	The number of iterations	100
	SW	The input Number of nodes	5
BP	Eta	Learning rate	0.5
	Error	Precision	0.0002
	D	The input Number of nodes	5
	L	The output Number of nodes	1
	Q	Number of hidden layers	5
SE	m	The refactoring dimension	5
	r	The threshold size, r = 0.1~0.25std(data)	0.24

4.3 The Results

After the experiments, we got the results of prediction. From the figure 3, CEEMADN algorithm was used to decompose the original sequence into 11 IMF components and 1 residual component. Table 3 shows the sample entropy of 12 IMF sequences, after combining the value of SE and several experiments, the most effective method is to combine IMF1 and IMF2 to form a new sequence Rec1, and IMF3-IMF12 to form a new sequence Rec2.

Table 3. The values of SE

Series	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
SE	1.24	1.70	1.68	0.98	0.56	0.25	0.17	0.10	0.03	0.01	0.01	0.02

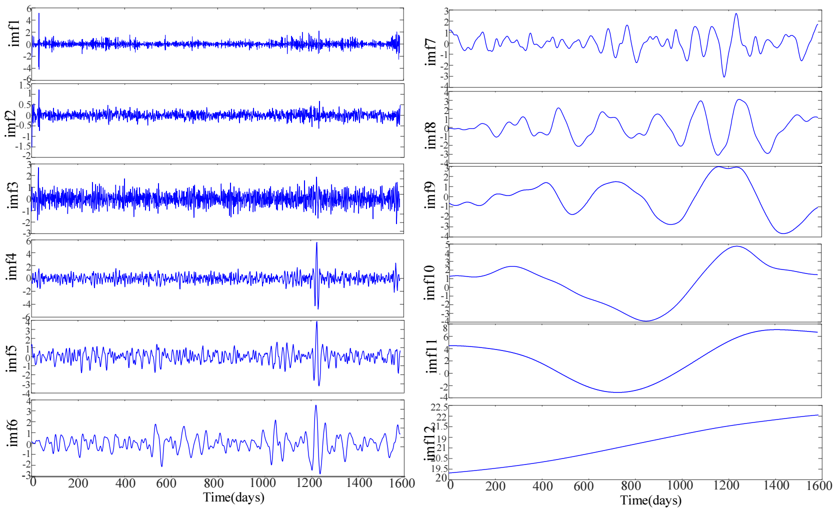


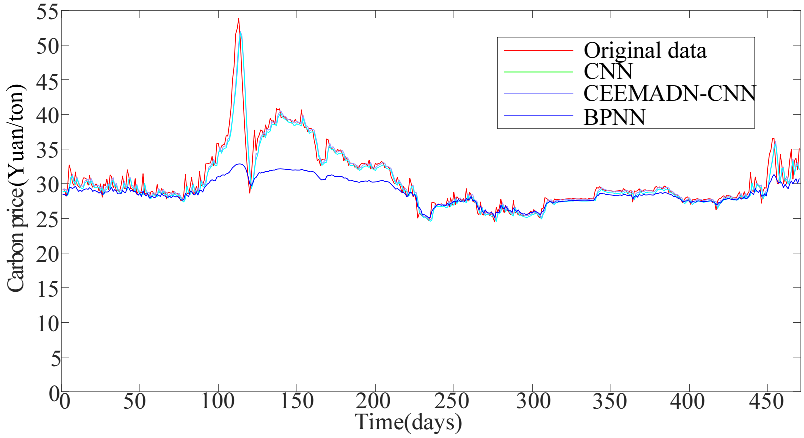
Figure 3. The IMFs after CEEMADN

From table 4, the R^2, MAE, RMSE of CEEMADN-CNN are 0.9754, 0.6809 and 0.4893, respectively. Obviously, CEEMADN-CNN model has better prediction effect than CNN and BP. Compare CNN to BP, the former prediction errors is much better than the latter.

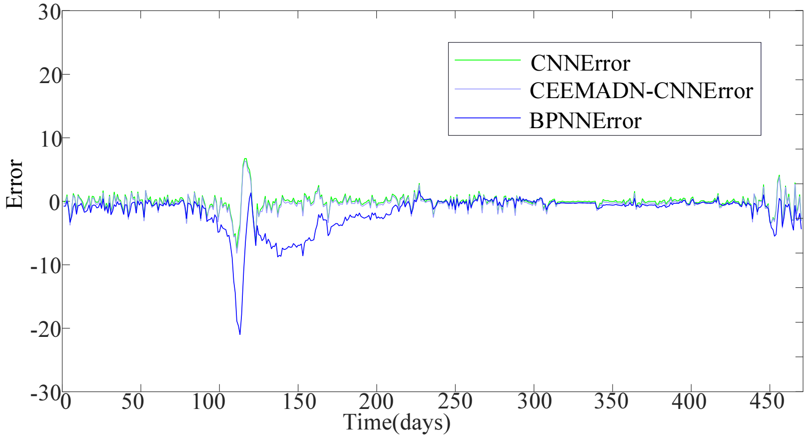
Table 4. The results of prediction

Model	R^2	MAE(yuan/ton)	RMSE(yuan/ton)
CNN	0.9606	0.6618	0.8614
CEEMADN-CNN	0.9754	0.6809	0.4893
BP	0.4231	1.8327	3.2998

From the figure 4, we draw the fitting diagram and error diagram of the test set. As we can see from the figure, the fitting effect of the CEEMADN-CNN model is the best than other two models, and the fitting effect of the BP model is the worst; the error of the CEEMADN-CNN model fluctuates gently, while the error of the BP model fluctuates the most. It indicates that the prediction effect improved after CEEMADN decomposition. What's more, the prediction effect of CNN was better than BP.



a. The curve fitting of test sets



b. The error of test sets

Figure 4. The results of prediction

5. Conclusion

In this study, we proposed a carbon price prediction model combine CEEMADN with CNN. The CEEMADN is used to extract the features of carbon price in Hubei, through decompose the raw data into 12 IMFs. IMFs was reconstructed by means of sample entropy. We tested the performance of the proposed model according to different comparison sets. All the results indicated that the proposed model had the best prediction performance.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (71801044).

References

[1] Yan S., Lv C., Yao L., Hu Z., Wang F. Hybrid dynamic coal blending method to address multiple environmental objectives under a carbon emissions allocation mechanism, Energy, 2022. 254, 124297.
[2] Pan X., Pu C., Yuan S., Xu H., Effect of Chinese pilot's carbon emission trading scheme on enterprises' total factor productivity: The moderating role of government participation and carbon trading market efficiency, Journal of Environmental Management, 2022. 316, 115228
[3] Dong Z., Xia C., Fang K., Zhang W., Effect of the carbon emissions trading policy on the co-benefits of carbon emissions reduction and air pollution control, Energy Policy, 2022. 165, 112998.
[4] Wang Q., Li L. The effects of population aging, life expectancy, unemployment rate, population density, per capita GDP, urbanization on per capita carbon emissions. Sustainable Production and Consumption. 2021. 28, 760-774.
[5] Liu H., Du K., Li J., An improved approach to estimate direct rebound effect by incorporating energy efficiency: A revisit of China's industrial energy demand. Energy Economic, 2019. 80, 720-730.
[6] Zhu B., Wei Y.,. Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. Omega. 2013. 41(3), 517-524.
[7] Huang Y., Dai X., Wang Q., Zhou D., A hybrid model for carbon price forecasting using GARCH and long short-term memory network, 2021. Applied Energy. 285, 116485.
[8] Ma X., Jiang P., Jiang Q., Research and application of association rule algorithm and an optimized grey model in carbon emissions forecasting. Technological Forecasting and Social Change, 2020. 158,120159.
[9] Lu H., Ma X., Huang K., Azimi M.,. Carbon trading volume and price forecasting in China using multiple machine learning models. Journal of Cleaner Production, 2020. 249, 119386.
[10] Liu K., Cheng J., Yi J., Copper price forecasted by hybrid neural network with Bayesian Optimization and wavelet transform, Resources Policy, 2022.75, 102520.
[11] Zhao L., Miao J., Qu S. Chen X., A multi-factor integrated model for carbon price forecasting: Market interaction promoting carbon emission reduction. Science of The Total Environment, 2021.796, 149110.
[12] Sun W., Zhang C., Sun C., Carbon pricing prediction based on wavelet transform and K-ELM optimized by bat optimization algorithm in China ETS: the case of Shanghai and Hubei carbon markets. Carbon Management, 2018. Pages 605-617.
[13] Wang J., Cui Q., Sun X., A novel framework for carbon price prediction using comprehensive feature screening, bidirectional gate recurrent unit and Gaussian process regression. Journal of Cleaner Production, 2021.314,128024.
[14] Ji L., Zou Y., He K., Zhu B., Carbon futures price forecasting based with ARIMA-CNN-LSTM model. Procedia Computer Science, 2019.162, 33-38.
[15] Wang L., Shao Y., Fault feature extraction of rotating machinery using a reweighted complete ensemble empirical mode decomposition with adaptive noise and demodulation analysis, Mechanical Systems and Signal Processing, 2020. 138, 106545.