Research on Carbon Emission Allowance Price Prediction Based on Quadratic Decomposition

Xueying Zhang*
School of Mathematics and Statistics, North China University of Water Resources and Electric Power, Zhengzhou, Henan, China
**Corresponding Author*

Abstract: To address the pronounced non-stationarity and high complexity inherent in carbon emission allowance price series, this study introduces a hybrid forecasting framework grounded in secondary decomposition to enhance predictive precision. Initially, the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) algorithm is employed to disaggregate the original price data into multiple intrinsic mode functions (IMFs). These components are subsequently reconstructed into high-, medium-, and low-frequency clusters via K-means clustering based on fuzzy entropy. The high-frequency constituents undergo further decomposition using Variational Mode Decomposition (VMD) to mitigate volatility and noise. A Long Short-Term Memory (LSTM) network optimized by Particle Swarm Optimization (PSO) is applied to predict the refined high-frequency sub-sequences, while the medium- and low-frequency components are forecast using the XGBoost algorithm for its efficiency in handling structured sequences. The final prediction is derived through the aggregation of all reconstructed subsequences. Empirical validation using carbon market data from Hubei demonstrates that the proposed model achieves a reduction in mean absolute percentage error by up to 1.77 percentage points and improves the coefficient of determination by 15.31%, confirming its superior accuracy and robustness against benchmark models.

Keywords: Secondary Decomposition; Fuzzy Entropy; Carbon Price Forecasting; LSTM Model; Multi-Frequency Combination

1. Introduction

Against the backdrop of increasingly severe global climate change, China formally announced in September 2020 its commitment to achieve a peak in carbon dioxide emissions by 2030 and realize carbon neutrality by 2060. This target demonstrates China's responsibility in global climate governance, injects momentum into the global low-carbon transition, and outlines a direction for the sustainable development of China's economy and society [1]. Driven by the "Dual Carbon" goals, the national carbon emissions trading market was officially launched on July 16, 2021, as a key policy tool. The establishment of China's carbon market not only accelerates the low-carbon transformation of key industries but also provides a new market mechanism paradigm for global climate governance, exerting a profound impact on the synergistic development of the economy-energy-environment system [2].

In recent years, research on carbon emission allowance price prediction has garnered significant attention in the fields of environmental economics and energy finance. Traditional statistical and econometric methods are straightforward and applicable for analyzing quantitative characteristics and changing patterns of economic phenomena. However, most models assume that carbon price series are stationary and linear, which contradicts reality. The primary characteristics of carbon allowance prices are nonlinearity and non-stationarity, limiting the effectiveness of traditional econometric models in predicting their long-term trends [3-5].

Compared with traditional statistical and econometric approaches, artificial intelligence prediction methods possess powerful adaptive learning capabilities. They can accurately capture the nonlinear and non-stationary features of carbon allowance prices, thereby enhancing prediction performance. Although certain single AI prediction methods have outperformed traditional models in empirical analyses, the performance of single prediction models still has

room for improvement, given that carbon trading prices are influenced by complex market factors and exhibit high volatility [6-8].

As research progresses, it has been recognized that traditional single models are inadequate for handling nonlinear time series characteristics. To address issues such as excessive data volatility, non-stationarity, and the identification of complex temporal information, some scholars have begun exploring model coupling and proposed hybrid prediction methods. For instance,Ma Zhonglin conducted a systematic study on Shanghai carbon prices by establishing a decomposition and error correction-based TCN-LSTM-Attention model, which demonstrates considerable potential for practical application [9]. Chen et al. proposed a carbon price prediction model based on CEEMDAN and sample entropy (SE) integrated with a Gated Recurrent Unit (GRU) network, and applied it to the Hubei carbon trading market.[10]. Duan and Yang adopted a CEEMDAN-SSA-GRU model to predict carbon market price data from Hubei Province and Beijing. Their experiments showed that this model can accurately predict carbon price data across different regions and time scales while maintaining significant prediction stability[11].

In summary, in the context of carbon allowance price prediction, single statistical forecasting models tend to yield low direct prediction accuracy due to the high complexity of carbon price sequences. Although some scholars have adopted hybrid prediction approaches that first decompose the carbon price series to reduce complexity before forecasting, single decomposition methods are prone to issues such as mode mixing, which can compromise the overall prediction accuracy and reliability. To address these limitations, this paper proposes a hybrid prediction model based on a "divide and conquer" strategy, employing a dual decomposition framework. The model is validated using carbon allowance price data from Hubei Province. Comparative experiments demonstrate that the proposed model effectively enhances the prediction accuracy of carbon allowance prices.

2. Related Technologies and Principles

2.1 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

CEEMDAN represents an enhanced version of Empirical Mode Decomposition (EMD), with its principal advancement residing in a distinctive strategy for noise management. At each stage of the decomposition process, it adaptively adds white noise of a specific amplitude to the current residual signal, generating multiple copies. For each copy, only one round of EMD is performed to extract the candidate Intrinsic Mode Function (IMF) of the current level. The candidate IMFs of the same level from all copies are then ensemble-averaged to obtain the final IMF component. Through this stage-wise noise injection and averaging mechanism, CEEMDAN effectively mitigates the problem of mode mixing and ensures the completeness of signal reconstruction, while also achieving higher computational efficiency compared to its predecessor, EEMD.

2.2 Variational Mode Decomposition (VMD)

VMD is a fully non-recursive, intrinsic signal decomposition technique. Unlike empirical methods, it frames the decomposition process as a variational optimization problem, the objective of which is to minimize the aggregate bandwidth of all modal components under the strict condition that their summation precisely reconstructs the original signal. This optimization is efficiently implemented in the frequency domain through the Alternating Direction Method of Multipliers (ADMM). Notable strengths of VMD lie in its rigorous mathematical foundation, its capability to effectively suppress mode mixing, and its robustness in the presence of noise. The method has found successful applications across diverse domains, including fault diagnosis, financial time-series analysis, and biomedical signal processing.

2.3 XGBoost

XGBoost is an efficient gradient boosting algorithm. Its core idea is to integrate multiple decision trees and optimize the objective function using gradient descent, thereby enhancing the predictive accuracy and generalization capability of the model. During training, it computes a weighted sum of predictions from each decision tree and continuously adjusts the training direction and weight parameters of subsequent trees

based on the gradient information of the loss function. Additionally, regularization terms are introduced to prevent overfitting. Compared to traditional gradient boosting algorithms, XGBoost demonstrates significant improvements in computational efficiency, model performance, and robustness to outliers. It is widely used in various machine learning tasks such as classification and regression.

2.4 Particle Swarm Optimization (PSO)

PSO is a stochastic optimization technique based on population dynamics, drawing inspiration from collective intelligent behavior. In this algorithm, each candidate solution is conceptualized as a particle, defined by two fundamental attributes: a position vector, which represents the current solution in the search space, and a velocity vector, which dictates the direction and step size of the particle's movement in the next iteration. A group of these particles constitutes a swarm that collectively navigates through the solution domain.

The optimization process proceeds as follows:

(1)Initialization:

The swarm is initialized by randomly generating the position and velocity vectors for each particle within the predefined search space.

(2) Fitness Evaluation:

The fitness value of each particle is computed using a predefined fitness function. Based on this evaluation, each particle's personal best position (p_{Best}), representing the best solution it has achieved so far, and the global best position (g_{Best}), denoting the best solution found by any particle in the entire swarm, are recorded and updated.

(3) Velocity and Position Update:

Each particle's velocity vector is updated based on its current velocity, its pbest, and the swarm's gbest. The new velocity is then used to update the particle's position. This process drives the swarm toward promising regions of the solution space.

$$v_i^{t+1} = wv_i^t + c_1 r_1 (p_{Best_i} - x_i^t) + c_2 r_2 (g_{Best} - x_i^t) \quad (1)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (2)$$

In the formula,$v_i^{t+1}$is the velocity of particle i in the t generation; w is the inertia weight; r_1 and r_2 are acceleration coefficients, usually used to adjust the speed at which particles approach individual and global optimal solutions; c_1 and c_2 are random numbers between 0 and 1, used to introduce randomness; p_{Best_i} is the individual optimal position of the particle; g_{Best} is the global best position for the group; x_i^{t+1} is the position of the particle i in the $t+1$ generation.

(4) Fitness Evaluation and Update:

The fitness value of each particle is recalculated based on its updated position. The quality of the current solution is evaluated, and both the personal best position (p_{Best}) of each particle and the global best position (g_{Best}) of the entire swarm are updated accordingly if better solutions are found.

(5) Termination Condition Check:

The algorithm terminates if the predefined maximum number of iterations is reached or if a specific precision requirement is satisfied. Otherwise, the process returns to Step (3) to continue the iterative optimization.

2.5 Long Short-Term Memory Network (LSTM)

LSTM is specifically engineered to model both long-term and short-term temporal dependencies in sequential data. Its fundamental innovation resides in three specialized gating mechanisms—the forget gate, input gate, and output gate—which work in concert to regulate and manage the information flow through the memory cell. The functional roles of each gate are defined as follows:

(1) Forget Gate:

This gate controls the extent to which information from the previous cell state is retained or discarded. It determines which information is no longer relevant and should be forgotten. The computation is defined as follows:

$$f_t = \sigma(W_f \times [h_{t-1}, x_t] + b_f) \quad (3)$$

(2) Input Gate:

The input gate selectively updates the cell state by incorporating new information from the current input. It decides which values should be stored in the cell state. The computation is performed as follows:

$$i_t = \sigma(W_f \times [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\widetilde{C}_t = \tan(W_c \times [h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = i_t \times \widetilde{C}_t + f_t \times \widetilde{C}_{t-1} \quad (6)$$

(3) Output Gate:

The output gate controls the amount of information from the current cell state that is

exposed to the hidden state and subsequent layers. It filters the cell state to produce the final output:

$$O_t = \sigma(W_0 \times [h_{t-1}, x_t] + b_0) \quad (7)$$

$$h_t = O_t \times \tan(C_t) \quad (8)$$

Among them: f_t is the forget gate, i_t is the input gate, O_t is the output gate, and C_t is the control gate; σ is the Sigmoid function; W_f. W_c and W_0 are respectively the weight matrices for the forget gate, input gate, and output gate; b_f. b_i and b_0 are the bias vectors for the forget gate, input gate, and output gate, respectively; h_t is the hidden layer output; tanh is the activation function.

2.6 A Dual-Decomposition-Based Carbon Emission Allowance Price Prediction Model

This paper presents a dual-phase decomposition and integration framework for forecasting carbon emission allowance prices. The approach begins by applying Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to disassemble the original price series into a set of intrinsic mode functions (IMFs). The fuzzy entropy of each IMF is computed and subsequently clustered via the K-means algorithm, leading to the reconstruction of the subseries into three distinct categories: high-frequency, medium-frequency, and low-frequency components. Following this, the high-frequency components undergo secondary decomposition using VMD to mitigate their complexity and volatility. The resulting subcomponents are forecast using a LSTM network, the hyperparameters of which are optimized via PSO; the individual predictions are then integrated. For the medium- and low-frequency subseries, the XGBoost algorithm is employed for prediction, leveraging its computational efficiency and modeling robustness. The final forecast is obtained by aggregating all predicted components through a comprehensive integration procedure.

3. Empirical Analysis

3.1 Data Selection and Error Evaluation Metrics

The daily carbon emission allowance price data from the Hubei Emission Trading Scheme, covering the period from April 2, 2014, to April 22, 2025, were selected for this study. After excluding non-trading days and missing values, a total of 2,546 valid observations were retained. The dataset was divided into training and testing sets in an 8:2 ratio.

To assess the predictive performance of the proposed model, four evaluation metrics were utilized: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the Coefficient of Determination (R^2). Among these, a higher R^2 value indicates superior explanatory power and model fit, whereas lower values of MAE, RMSE, and MAPE correspond to greater forecasting accuracy. The computational formulas for these metrics are provided below:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (10)$$

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (12)$$

Among them: y_i is the true value; $\hat{y}_i$ is the predicted value; $\bar{y}$ is the average value; N is the sample size.

3.2 Analysis of Decomposition Experiment Results

3.2.1 CEEMDAN decomposition

To mitigate the pronounced volatility and inherent complexity of the carbon emission allowance price series, and to improve the forecasting capability of downstream models, the original price sequence is decomposed using the CEEMDAN method. The initial carbon price series is characterized by strong non-stationarity and significant fluctuation. CEEMDAN is employed to break down the target series into a collection of Intrinsic Mode Functions (IMFs), each representing a distinct frequency component. As shown in Figure 1, the complexity of each IMF component decreases progressively after decomposition. In the figure, the horizontal axis represents the timeline in daily units, and the vertical axis indicates the oscillation frequency of each IMF component.

3.2.2 Fuzzy entropy calculation and clustering

Fuzzy entropy is employed in this study as an indicator to evaluate the amount of information

contained in each subseries. The core idea involves reconstructing the phase space and introducing a fuzzy measure to fuzzify similarity comparisons, thereby obtaining a stable entropy value that reflects the complexity of the time series. A higher fuzzy entropy value indicates greater complexity in the time series. Based on the fuzzy entropy values of individual components, the subseries are clustered into three categories: high-frequency, low-frequency, and trend components. This clustering process reduces the number of components used in modeling, thereby minimizing error accumulation and computational burden, and effectively lowering the complexity of the forecasting model.

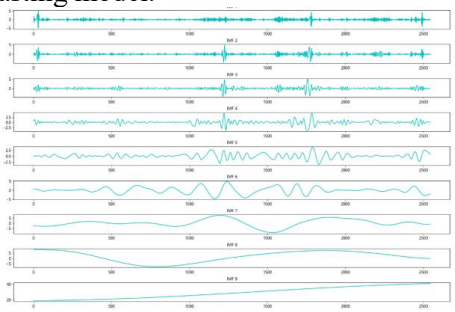

Figure 1. CEEMDAN Decomposition Result Diagram

3.2.3 VMD decomposition

The reconstructed high-frequency subseries remains complex; therefore, Variational Mode Decomposition (VMD) is applied for secondary decomposition to further reduce its complexity. The parameters of VMD are set as follows: noise tolerance is set to 0, convergence accuracy is set to e^{-7} , and the penalty parameter is set to 2000. The selection of the parameter K (number of modes) directly affects the number of modal components. This study uses the center frequency observation method to determine the optimal value of K. Different modal components exhibit distinct center frequencies, making it possible to select K by analyzing the changes in center frequencies under different K values.

Tests were conducted with K values ranging from 3 to 8. When K was increased from 3 to 5, each increment led to a significant increase in the center frequency of the last modal component, indicating that the decomposition was incomplete. However, when K was set to 7 or 8, the increase in the center frequency of the last component became negligible compared to the previous K value. This suggests that the information within the original sequence was almost fully decomposed at K=6. Therefore, the number of decomposition modes K in VMD was set to 6.

3.3 Experimental Results and Comparative Analysis

3.3.1 Comparative experiments

To rigorously evaluate the performance of the proposed hybrid forecasting framework, a comparative analysis was carried out against several established benchmark models, including standalone XGBoost, LSTM, as well as hybrid approaches such as CEEMDAN-LSTM and VMD-LSTM. The predictive outcomes of each model are visually summarized in Figures 2 through 6.

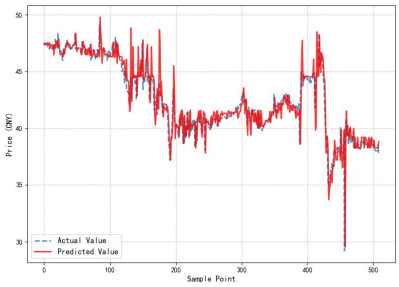

Figure 2. Forecasting Results of the Hybrid Model

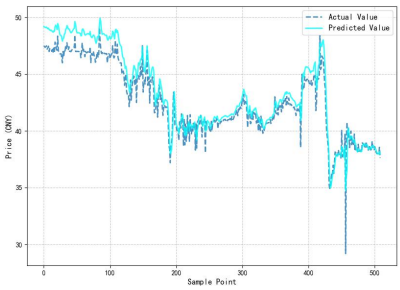

Figure 3. Forecasting Results of XGBoost

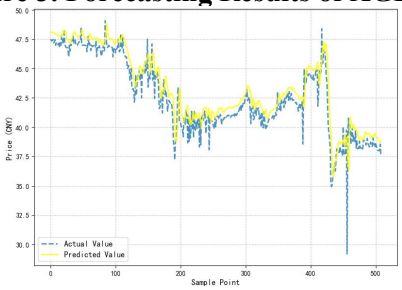

Figure 4. Forecasting Results of LSTM

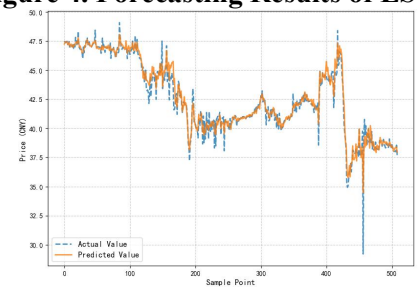

Figure 5. Forecasting Results of CEEMDAN-LSTM

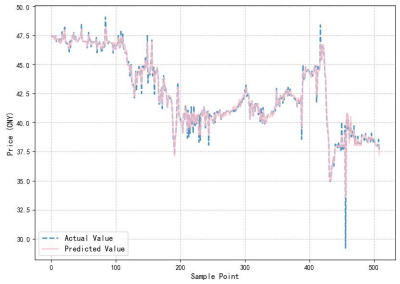

Figure 6. Forecasting Results of VMD-LSTM

Based on the experimental results presented in Figures 2–6 and Table 1, it can be observed that the proposed dual-decomposition hybrid forecasting model achieves smaller errors and demonstrates superior predictive performance and accuracy compared to other models.

When compared with single prediction models—specifically, XGBoost and LSTM—the proposed model reduces RMSE by 71.76% and 70.16%, respectively. MAE is reduced by 75% and 79.57%, and MAPE is lowered by 1.44 and 1.77 percentage points, respectively. These results indicate a significant improvement in prediction accuracy over classical single models.

In comparisons with single-decomposition hybrid models, namely CEEMDAN-LSTM and VMD-LSTM, the proposed model reduces RMSE by 67.83% and 36.21%, MAE by 79.12% and 48.65%, and MAPE by 1.61 and 0.44 percentage points, respectively.

The experimental analysis confirms that the dual-decomposition hybrid forecasting model offers substantial advantages in terms of accuracy. The proposed model significantly outperforms all single models across all evaluation metrics. Moreover, it also exhibits clear improvements over single-decomposition hybrid models, further demonstrating the benefits of the secondary decomposition strategy in enhancing both prediction accuracy and generalization capability. These findings underscore the effectiveness of the proposed dual-decomposition model in handling the nonlinear and complex characteristics of carbon emission allowance price series.

Table 1. Comparison of Evaluation Metrics among Different Models

Comparative experimental models	RMSE	MAE	MAPE/%	R^2/%
XGBoost	1.31	0.76	1.91	83.27
LSTM	1.24	0.93	2.24	84.68
CEEMDAN-LSTM	1.15	0.91	2.08	86.95
VMD-LSTM	0.58	0.37	0.91	96.66
Hybrid model	0.37	0.19	0.47	98.58

4. Conclusion

Carbon emission allowance price time series are characterized by nonlinearity, non-stationarity, high complexity, and multiple influencing factors. When handling such complex sequences, decomposition methods can effectively reduce volatility, while deep learning approaches exhibit superior performance in predicting nonlinear and intricate temporal data due to their network structures. This study takes the carbon price series from the Hubei market as the research object and proposes a hybrid forecasting model based on a "divide and conquer" strategy, which integrates CEEMDAN-VMD dual decomposition with LSTM and XGBoost algorithms. Furthermore, the PSO algorithm is introduced to optimize the hyperparameters of the LSTM model, aiming to enhance overall prediction accuracy and generalization capability. The following conclusions are drawn from comparative experiments:

(1) Traditional machine learning and classical deep learning models exhibit limited prediction accuracy and robustness when directly applied to highly complex carbon allowance price sequences. In contrast, the "decomposition-then-prediction" approach demonstrates improved performance in mitigating these issues.

(2) Single decomposition methods may suffer from mode mixing, excessive noise, and insufficient reduction in data complexity, which can impair prediction performance. The proposed dual-decomposition model combining CEEMDAN and VMD more effectively reduces noise and decreases data complexity, thereby significantly enhancing predictive accuracy.

(3) Empirical analysis based on carbon allowance price data from Hubei Province confirms that the proposed model outperforms conventional forecasting approaches, achieving higher precision in predicting nonlinear, non-stationary, and highly complex carbon price sequences.

References

[1] Li M Y, Wang J, Tian Y. Research on the development of China's economic industries under the carbon peak and carbon neutrality goals. Global Energy Interconnection, 2024,

7(06): 629-639.
[2] Wang K, Lü C. Global and Chinese carbon market review and outlook (2025). Journal of Beijing Institute of Technology (Social Sciences Edition), 2025, 27(02): 19-36.
[3] Huang Y X. Research on the volatility of China's carbon price based on the Skew-t-EGARCH model. China Forestry Economics, 2022, 173(02): 92-96.
[4] Zhou L. Research on the price forecasting of the national carbon emission trading market based on a combined forecasting model. Jinan: Shandong University, 2023.
[5] Huang J K. The impact of economic policy uncertainty on the volatility of Shenzhen's carbon emission trading market: An analysis based on the GARCH-MIDAS model. China Journal of Commerce, 2024, (12): 121-125.
[6] Wen L, Liu Y. A research about Beijing's carbon emissions based on the IPSO-BP model. Environmental Progress & Sustainable Energy, 2017, 36(2): 428-434.
[7] Huang X, He G Y, Sun Y X. Research on carbon trading price forecasting based on RF-PSO-LSTM model. Journal of Hunan City University (Natural Science Edition), 2024, 33(06): 61-66.
[8] Wang Y F, Wang N, Zhang S H. Hybrid forecasting of EU carbon price based on multi-source data features and multi-scale analysis. Journal of Systems Science and Mathematical Sciences, 2024: 1-24.
[9] Ma Z L. Carbon price forecasting based on decomposition and error correction with TCN-LSTM-Attention model. Shanxi University, 2024.
[10] Chen Y, Zhang M, Gao J. Carbon emission allowance trading price prediction based on CEEMDAN-SE-GRU. Information System Engineering, 2024, (11): 123-126.
[11] Duan J T, Yang X Z. Research on carbon price prediction method integrating deep learning and signal decomposition. The World of Survey and Research, 2024.

