RESEARCH ARTICLE

Forecasting China carbon price using an error-corrected secondary decomposition hybrid model integrated fuzzy dispersion entropy and deep learning paradigm

Po Yun[1] · Yingtong Zhou[1] · Chenghui Liu[1] · Yaqi Wu[2] · Di Pan[1]

Abstract

Forecasting China's carbon price accurately can encourage investors and manufacturing industries to take quantitative invest-ments and emission reduction decisions effectively. The inspiration for this paper is developing an error-corrected carbon price forecasting model integrated fuzzy dispersion entropy and deep learning paradigm, named ICEEMDAN-FDE-VMD-PSO-LSTM-EC. Initially, the improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) is used to primary decompose the original carbon price. Subsequently, the fuzzy dispersion entropy (FDE) is conducted to identify the high-complexity signals. Thirdly, the variational mode decomposition (VMD) and deep learning paradigm of particle swarm optimized long short-term memory (PSO-LSTM) models are employed to secondary decompose the high-complexity signals and perform out-of-sample forecasting. Finally, the error-corrected (EC) method is conducted to re-modify and strengthen the above-predicted accuracy. The results conclude that the forecasting performance of ICEEMDAN-type secondary decomposition models is significantly better than the primary decomposition models, the deep learning PSO-LSTM-type models have superiority in forecasting China carbon price, and the EC method for improving the forecasting accuracy has been proved. Noteworthy, the proposed model presents the best forecasting accuracy, with the forecasting errors RMSE, MAE, MAPE, and Pearson's correlation are 0.0877, 0.0407, 0.0009, and 0.9998, respectively. Especially, the long-term forecasting ability for 750 consecutive trading prices is outstanding. Those conclusions contribute to judging the carbon price characteristics and formulating market regulations.

Keywords Forecasting carbon price · ICEEMDAN · Secondary decomposition · Deep learning paradigm · Error corrected · Fuzzy dispersion entropy

Introduction

Global warming is closely related to carbon emissions, while the increasing emissions are originated from rapid economic development and industrial advancement (Wu et al. 2021; Nazifi and Milunovich 2010). Limiting the rise of global warming within 1.5 °C is the most ambitious goal of the Paris Agreement signed in 2015; however, this goal is almost impossible at the current emission level. According to the International Energy Agency (IEA), the global emissions of energy-related greenhouse gases in 2022 reached 41.3 billion tons, while China's carbon emissions have improved from 7.71 billion tons to 11.477 billion tons between 2009 and 2022. As the global climate issue is getting increas-ingly severe, China, the world's largest energy consumer and greenhouse gas emitter, is facing severe environmental pressure. To solve the environmental constraints on high-quality development, the Chinese government officially implemented the "carbon peak, carbon neutral" strategy in September 2020, and the construction of a carbon market has become a concrete measure to reduce carbon emis-sions. The carbon market plays a crucial role in helping China reduce industrial pollutant emissions (Byun and Cho 2013; Zhang et al. 2019). The higher the market efficiency, the more obvious the reduction of emissions performance

(Chevallier 2009; Schneider et al. 2019); therefore, studying the price formation and driving mechanism of the carbon market and forecasting the carbon price accurately are the keys to achieving the goals of the proposed strategy. However, the mainstream carbon price forecasting models based on empirical mode decomposition (EMD) usually present the problems of large decomposition errors and signal complexity identifying subjectively, which may result in low forecasting accuracy. This paper develops a new secondary decomposition carbon price-forecasting model that integrates the fuzzy dispersion entropy and deep learning paradigm; noteworthy, we design an error-corrected idea to strengthen the deep learning forecasting ability which is commonly ignored in most previous studies. The expected forecasting accuracy can support valuable market decision and contribute to emission reduction at a low cost.

The rest of this paper is organized as follows: the "Literature review" section is the literature review. The "Methodology" section introduces the basic models and describes the logical framework of the proposed model. The "Empirical discussion" section discusses and analyzes the out-of-sample forecasting results and analyzes the robustness in different forecasting periods. The "Conclusions" section is the conclusions.

Literature review

Most studies have found that the carbon price has nonlinear, non-stationary, non-normal, and multi-scale characteristics (Zhu et al. 2015; Pan et al. 2023; Zhang et al. 2023; Yue et al. 2023). Different from the volatility modeling technologies, the EMD technology has become the main method for carbon price forecasting in recent years (Tang et al. 2017; Mao and Zeng 2023). Based on the decomposition differences, existing literature can be divided into primary decomposition and secondary decomposition carbon price forecasting studies.

Primary decomposition carbon price forecasting studies

Conducted the EMD technology to decompose the carbon price, used the least squares support vector machine optimized by the particle swarm optimization algorithm (PSO-LSSVM), and generalized autoregressive conditional heteroskedasticity (GARCH) model to take the out-of-sample forecasting, the results put that the EMD-PSO-LSSVM and EMD-GARCH models have high forecasting accuracy in Europe carbon future price (Jianwei et al. 2021; Zhu et al. 2018; Zhang and Wu 2022). The ensemble empirical mode decomposition (EEMD) and complete EEMD (CEEMD) are developed to primary decompose the carbon price; the

autoregressive integrated moving average (ARIMA), local polynomial prediction (LPP), and the particle swarm optimized gray neural network (PSO-GNN) are employed for forecasting, and it is found that the EEMD-ARIMA-LPP, EEMD-LPP, and CEEMD-PSO-GNN models are efficient carbon price forecasting tools (Qin et al. 2020; Zhang et al. 2018). Furthermore, the modified EEMD (MEEMD) and variational mode decomposition (VMD) technologies can also effectively reduce decomposition errors and improve decomposition efficiency (Yang et al. 2020; Guo et al. 2022). The mode reconstruction (MR) method is introduced to identify the contribution of each component, and long short-term memory (LSTM) and GARCH models are used to forecast; the conclusions suggested that the model of VMD-MR-LSSVM has the highest forecasting accuracy in China's Hubei and Shenzhen carbon markets, while the VMD-LSTM-GARCH model has higher forecasting accuracy in European carbon price (Zhu et al. 2019; Huang et al. 2021). Adopted the multi-resolution singular value decomposition (MRSVD) technology to decompose the original carbon price and utilized the extreme learning machine (ELM) optimized by the adaptive whale optimization algorithm (AWOA) to forecast the carbon price, the findings maintained that the forecasting accuracy of MRSVD-AWOA-ELM model in China and European carbon market is better than other comparative models such as EMD-AWOA-ELM and EMD-AWOA-BP (Sun and Zhang 2018). The model of complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), improved CEEMDAN (ICEEMDAN), and sample entropy (SE) are used to decompose and extract the complex carbon price signals; the studies suggested that the CEEMDAN-SE-LSTM and ICEEMDAN-type models have stronger robustness and generalization ability in China short-term carbon price forecasting (Wang et al. 2021; Yun et al. 2023; Hao et al. 2020).

Secondary decomposition carbon price forecasting studies

Although the EMD-type models can decompose the original carbon price into multi-scale time-frequency signals, they may produce larger decomposition errors and mode mixing problems theoretically (Nguyen and Phan 2022; Junior et al. 2020), while the secondary decomposition technology can reduce the decomposition errors effectively by recognizing the high-complexity signals to some extent (Kong et al. 2022).

For example, employed the EMD, EEMD, CEEMD, CEEMDAN, and VMD technologies to primary decompose and secondary decompose the carbon price; classified high-complexity signals by the sample entropy (SE), multi-scale fuzzy entropy (MFE), and partial autocorrelation function (PACF) (Li et al. 2021; Yang et al. 2023a, 2023b); and

further built the deep learning models such as PSO-LSSVM, backpropagation network (BP), gate recurrent unit (GRU), LSTM, and ELM models for out-of-sample forecasting (Li et al. 2021; Wang et al. 2022), the results convinced that the secondary decomposition EMD-VMD-BP and EMD-VMD-LSTM models have stable forecasting performance in Beijing and Shanghai carbon markets (Sun and Huang 2020), while VMD-EEMD-GRU, CEEMD-VMD-BP, and CEEMDAN-VMD-LSTM models have superior forecasting accuracy in Hubei and Guangdong carbon markets (Wu and Liu 2020; Zhou and Wang 2021; Liu et al. 2023). Theoretically, ICEEMDAN technology has obvious advantages over other EMD-type models (Liu et al. 2023). The research applied the ICEEMDAN and CEEMD technologies to primary and secondary decompose the carbon price and adopted the support vector machine (SVM) and multi-layer perceptron (MLP) for price forecasting, the findings show that the ICEEMDAN-CEEMD-SVM, ICEEMDAN-CEEMD-SVM, and ICEEMDAN-CEEMD-MLP models are significantly superior to other primary decomposition models for forecasting China carbon price (Li and Liu 2023; Yang et al. 2023a; Yang et al. 2023b).

Research gaps and contributions

In summary, the carbon price forecasting models in previous studies have developed from primary decomposition to secondary decomposition. However, there are still some improvements for further study. (1) The EMD-type technologies have theoretical defects of large decomposition errors, mode mixing, endpoint effects, and mode alignment problems, while the developed CEEMDAN technology may lead to false component signals during the initial stage of mode decomposition. (2) The complex signals recognized methods such as sample entropy and fuzzy entropy algorithm may have the problems of discontinuous entropy data, relatively sensitive data length and high sensitivity to noise interference. A new signal complexity identification method needs to be developed. (3) The existing models ignore the forecasting performance of the proposed model under different forecasting periods, and the reliability of the conclusions needs to be strengthened. (4) For convincing the forecasting ability, previous studies mainly focus on comparing the gap between the predicted price and the real price, while the role of predicted errors term in improving the final forecasting results has been ignored in most previous studies.

Based on the special carbon price characteristics, this article carried out a new logic of "primary decomposition - complexity recognized - secondary decomposition - forecasting and summing - error corrected" for guiding the hybrid model construction in forecasting China's carbon price. The contribution is designing a new error-corrected secondary decomposition hybrid model integrated fuzzy

dispersion entropy and deep learning paradigm, named ICEEMDAN-FDE-VMD-PSO-LSTM-EC model. Specifically, initially, the ICEEMDAN technology is used to primary decompose the original carbon price to solve the common problems of larger decomposition errors and mode mixing in traditional EMD-type technologies. Secondly, the fuzzy dispersion entropy (FDE) algorithm is conducted to identify the high-complexity signal to overcome the discontinuous entropy and sensitivity to noise in other entropy algorithms. Then, the VMD technology is used to secondary decompose the high-complexity signals that are recognized by the FDE algorithm. Fourthly, we conduct the deep learning LSTM model that has a time series fitting ability to forecast the acquired low complexity signals and further use the particle swarm optimization (PSO) algorithm to optimize the model parameters and improve the forecasting ability, and then the predicted price can be obtained by summing up the predicted components. Finally, the error-corrected (EC) method is used to forecast the errors of each model again, and the final predicted price with the error corrected has been calculated by adding up the predicted errors and the predicted price.

Methodology

ICEEMDAN model

The improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) model is an improved adaptive decomposition form based on the traditional EMD technology (Colominas et al. 2014). The classical EMD and EEMD technologies may cause problems of larger reconstruction errors and mode mixing (Wu and Huang 2009). In contrast, the core of the CEEMDAN model is adding positive and negative pairs of white noise to the residual of the decomposition process and calculating the average mode information after obtaining the first-order intrinsic mode functions (IMF) components (Zhou et al. 2022). Those special noise-added styles can solve the problem of white noise transfer from high frequency to low frequency effectively so that the reconstruction errors can be reduced greatly. Different from the CEEMDAN model, the significant feature of the ICEEMDAN model is regarding the IMF components as the white noise for adding it to the decomposition process, conducting the adaptive empirical mode decomposition (AEMD) to decompose the multi-scales signals. The decomposition process of the carbon price by ICEEMDAN can be as follows:

Step 1: define the operator $E_k(\cdot)$ as the k-th mode component, $M(\cdot)$ indicate the local mean signal that needs to be decomposed, and define $\langle \rangle$ as the mean process. Then,

add the Gaussian white noise to the original carbon price signal x:

$$x_1^{(i)} = x + \varepsilon_0 E_1\left(w^{(i)}\right) \tag{1}$$

Among them, $w^{(i)}$ means that the i-th white noise needs to be added and ε_0 is the noise standard deviation when the first signal is decomposed. $E_1(\cdot)$ represents the first IMF calculation process. Then, the residual of the first decomposition is expressed as $r_1 = \langle M(x^{(i)}) \rangle$.

Step 2: calculate the first mode component IMF1:

$$\text{IMF1} = x - r_1 \tag{2}$$

Step 3: calculate the second mode component IMF2:

$$\text{IMF2} = r_1 - r_2 \tag{3}$$

Step 4: calculate the k-th mode component IMFk:

$$\text{IMF}k = r_{k-1} - r_k \tag{4}$$

Step 5: iteration termination condition:

$$\sigma = \frac{\|\text{IMF}k - \text{IMF}(k-1)\|_2}{\|\text{IMF}k\|_2} \tag{5}$$

where $r_2 = \langle M(r_1 + \alpha_1 E_2(w^{(i)})) \rangle$, $r_k = \langle M(r_{k-1} + \alpha_{k-1} E_k(w^{(i)})) \rangle$, and $k = 2, 3, \cdots, N$, and σ means the standard deviation between two IMF components. When $k \geq 2$, the first mode decomposition is completed, and the next decomposition is continued until it meets the iteration termination condition. This paper adopts the Cauchy criterion as the iterative convergence condition, that is, when the standard deviation $\sigma < 0.2$, the iteration terminates. The idea of the ICEEM-DAN decomposition process is shown in Fig. 1.

Fuzzy dispersion entropy

Fuzzy dispersion entropy (FDE) is a reliable indicator to measure the time series complexity by calculating the probability of new signal generation (Rostaghi et al. 2021). The FDE algorithm overcomes the defects of discontinuous entropy data, sensitive data length that commonly exists in the calculation of sample entropy, fuzzy entropy, and other methods. For a given carbon price series $x = \{x_1, x_2, \cdots, x_N\}$, the FDE can be calculated as follows:

Firstly, map the carbon price series $x_j(j = 1, 2, \cdots, N)$ to $y_j(j = 1, 2, \cdots, N)$ by the standard normal distribution function $y_j = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_j} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt$.

Secondly, continue to transmit y_j to $c \times y_j + 0.5$ by linear function of $z_j^c = (c \times y_j + 0.5)$, where c represents the number of categories.

Thirdly, calculate the dispersion pattern $\pi_{v_0 v_1 \cdots v_{m-1}}(v = 1, 2, \cdots, c)$, where $v_0 = u_i^c$, $v_1 = u_{i+d}^c$, $v_{m-1} = u_{i+(m-1)d}^c$, m represents the embedding dimensions, d means the time delay, $u_i^{m,c} = \left\{ u_i^c, u_{i+d}^c, \cdots u_{i+(m-1)d}^c \right\}$, $i = 1, 2, \cdots, N - (m-1)d$, and $u_i^c = \text{round}(c \times y_i + 0.5)$. $\text{round}(\cdot)$ indicates the round function, the dispersion pattern consists of c-digit numbers, each number has m values, and there are c^m corresponding patterns.

Fourthly, calculate each dispersion pattern probability $\pi_{v_0 v_1 \cdots v_{m-1}}$.

Finally, based on the Shannon entropy definition, the FDE is defined as

$$\text{FDE} = \text{Fuzzy\_DismEN}(x, m, c, d) = -\sum_{\pi=1}^{c^m} p\left(\pi_{v_0 v_1 \cdots v_{m-1}}\right)$$
$$\times \ln\left(p\left(\pi_{v_0 v_1 \cdots v_{m-1}}\right)\right) \tag{6}$$
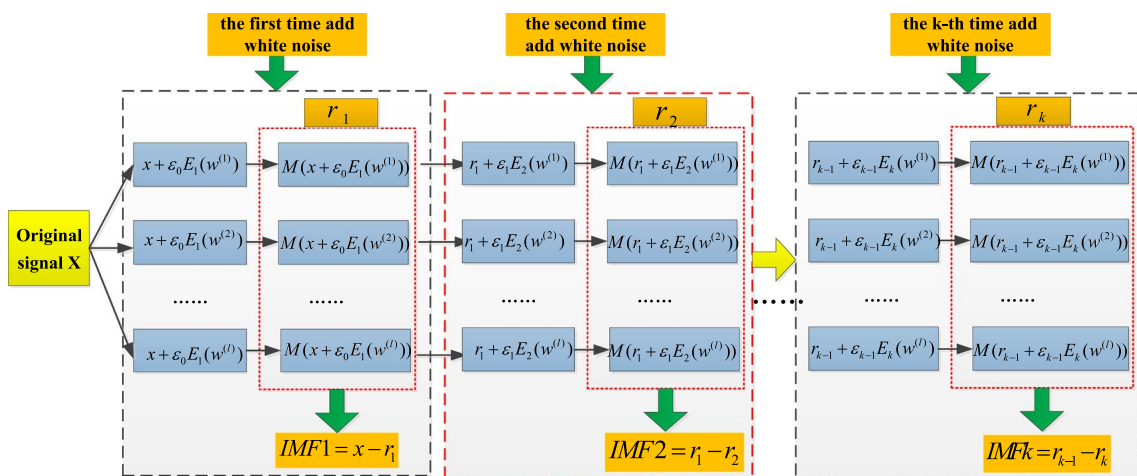

Fig. 1 The mode decomposition framework of the ICEEMDAN model

Theoretically, if the entropy value is greater than 1, it means that the input signal has a strong complexity and the signal noise is relatively large that needs to conduct the secondary decomposition process to reduce the noise complexity (Wang et al. 2022).

VMD model

The typical advantages of VMD are determining the mode number of the original carbon price signals flexibly and effectively avoiding the common problems of endpoint effect and mode mixing that are presented in traditional recursive algorithms (Dragomiretskiy and Zosso 2013). So, the VMD model is suitable for dealing with complex, nonlinear, and non-stationary time series theoretically. Assuming the carbon price is composed of a signal with a specific center frequency and finite width, the VMD model applies the Wiener filter method to adaptively search for the best center and width during the variational search and solution process. The decomposition steps of the VMD are as follows:

Firstly, decompose the carbon price signal S into K IMF components and ensure the decomposed signals are mode components with a specific center frequency and limited width, the signal variation expression is defined as:

$$
\begin{cases}
\min\limits_{\{IMFK\},\{w_K\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times IMFk(t) \right] e^{-jwkt} \right\|_2^2 \right\} \\
s.t. \sum_{k=1}^{K} IMFk = S
\end{cases}
\tag{7}
$$

where $\{IMFk\} = \{IMF1, IMF2, \cdots, IMFk\}$ is the obtained carbon price multi-scale signals, $\{w_k\} = \{w_1, w_2, \cdots, w_k\}$ is the center frequency of each mode signal, and $\delta(t)$ means the pulse signal.

Secondly, introduce the second-order penalty coefficient α and Lagrangian function L to solve the above optimization and variational problems:

$$
\begin{aligned}
L\left[\{IMFk(t)\}, \{w_k(t)\}, \tau(t) \right] &= \alpha \sum_{k=1} \left\| \partial_t \left[(\delta(t) + j/\pi t) \times IMFk(t) \right] e^{-jwkt} \right\|_2^2 \\
&+ \left\| S(t) - \sum_{k=1}^{K} IMFk(t) \right\|_2^2 + \left\langle \tau(t), S(t) - \sum_{k=1}^{K} IMFk(t) \right\rangle
\end{aligned}
\tag{8}
$$

where τ is the Lagrange multiplier.

Finally, update each mode component and its center frequency and calculate the final optimal solution by the following formula:

$$
\hat{I}MFk^{n+1}(w) = \frac{\hat{S}(w) - \sum_{i \neq k} \hat{v}_i(w) + \hat{\tau}_i(w)/2}{1 + 2\alpha(w - w_k)^2}
\tag{9}
$$

In the above formula, w represents the fluctuation frequency and $\hat{I}MFk^{n+1}(w)$, $\hat{v}_i(w)$, and $\hat{\tau}_i(w)$ are the Fourier transforms of $IMFk^{n+1}(w)$, $v_i(w)$, and $\tau_i(w)$ respectively. $IMFk^{n+1}(w)$ is the residual after Wiener filtering of $\hat{S}(w) - \sum_{i \neq k} \hat{v}_i(w)$.

According to the obtained IMF components, the center frequency w of the current mode can be updated by the following transformation:

$$
w_k^{n+1} = \frac{\int_0^\infty w \left| IMFk^{n+1}(w) \right|^2 dw}{\int_0^\infty \left| IMFk^{n+1}(w) \right|^2 dw}
\tag{10}
$$

PSO algorithm

Particle swarm optimization (PSO) is an improved algorithm with low parameter dependence for searching global solutions (Wang et al. 2018). The hidden layers, neurons, and learning rate of a neural network are important factors affecting the training performance. So, we use the PSO algorithm to optimize those parameters. According to the theory, any particle only has two attributes of movement speed and position; each particle conducts a separate optimization movement to search for the local optimal values. During this process, individual extreme value is shared with other particles to determine the global optimal value. Similarly, other particles also adjust their search speed and position accordingly. The optimization process of the PSO algorithm is as follows:

$$
\begin{aligned}
v_{i,t+1} &= w \times v_{i,t} + c_1 \times \text{rand}() \times (pbest_i - x_{i,t}) \\
&+ c_2 \times \text{rand}() \times (gbest_i - x_{i,t})
\end{aligned}
\tag{11}
$$

$$
x_{i,t+1} = x_{i,t} + \lambda v_{i,t+1}
\tag{12}
$$

Among them, $v_{i,t}$ indicates the current movement speed, *rand*() represents a random number between 0 and 1, *pbest* and *gbest* are the local and global optimal solutions, respectively, w represents the inertia coefficient, $c1$ and $c2$ are the learning factors, $v_{i,t+1}$ denotes the maximum value of $v_{i,t}$, and λ is the speed coefficient.

PSO-LSTM forecasting model

Based on the research design, the deep learning LSTM model is used for the nonlinear fitting of each low-complex component after the VMD process. One reason for choosing the LSTM model is that the model can effectively filter, screen, and update input information by a specially designed gate structure and retain the memory characteristics of the sample series (Hochreiter and Schmidhuber 1997). The training process of the LSTM is as follows:

The function of the forget gate is to screen each carbon price mode component to obtain the filtered output.

$$f_t = \sigma\left(W_f \times [h_{t-1}, x_t] + b_f\right) \tag{13}$$

The input gate updates the forgetting output, so as to calculate the updated characteristics of each mode component by the following function:

$$i_t = \sigma\left(W_i \times [h_{t-1}, x_t] + b_i\right) \tag{14}$$

$$\tilde{C}_t = \tanh\left(W_C \times [h_{t-1}, x_t] + b_C\right) \tag{15}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{16}$$

The output gate determines the mode component characteristics that need to be memorized. Thus, the output of this network can be acquired by the activation function.

$$o_t = \sigma\left(W_o \times [h_{t-1}, x_t] + b_o\right) \tag{17}$$

$$h_t = o_t \times \tanh\left(C_t\right) \tag{18}$$

Among them, f_t represents the data output screened by the forget gate, i_t and $\tilde{C}_t$ are the update features of the input gate and candidate cell vector, C_t indicates the updated vector, o_t means the output information of the cell, h_t is the network output, w and b represent the weight and bias, and σ is the sigmoid activation function.

This article takes the hidden layers, neurons, and learning rate of the LSTM model as the optimization variables for PSO particles. By updating the speed and position of the particles, the fitness value of carbon price forecasting is minimized, and the optimal model parameters can be obtained. The forecasting steps of the PSO-LSTM are as follows:

Firstly, initialize the particle swarm. We need to set the initial position and speed for each particle (that is the hidden layers, neurons, and learning rate of the LSTM model). Those initial values are usually randomly generated, and the range is determined by the constraint conditions. Secondly is fitness assessment. For each particle, construct an LSTM model using the current parameters, and then use this model to forecast. The forecasting errors are recognized as the fitness value of the particle. Thirdly, update particle position and speed. Update the position and speed of the particles based on their global optimal positions. Fourthly is iterative optimization. Repeat the above steps until the stop condition is met. In each iteration, the position and speed of particles are updated, and the fitness values are re-evaluated. Finally, model forecasting. After the parameters training steps above, use the obtained global optimal parameter to construct the final LSTM model and forecast the test data.

The proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model

To adapt the special carbon price characteristics of nonlinear, non-stationary, and non-normal, this paper designs an error-corrected secondary decomposition hybrid model integrated fuzzy dispersion entropy and deep learning paradigm. That is integrating the advantages of ICEEMDAN, FDE, VMD, PSO-LSTM, and EC methods to improve the forecasting accuracy of carbon prices. The idea of the proposed model is shown in Fig. 2.

Step 1: adopt the ICEEMDAN technology to decompose the original carbon price signals, so as to obtain several IMF signals and a residual term. Step 2: employ the FDE algorithm to calculate the signal complexity of the IMF that is acquired by the ICEEMDAN technology. Step 3: conduct the VMD technology to perform secondary decomposition on the high-complexity signals identified in step 2 to reduce the signal noise and decomposition errors. Step 4: apply the deep learning paradigm of PSO-LSTM to perform one-step-forward forecasting on the low-complexity signals. The predicted price can be obtained by summing up the predicted values of each IMF component. Step 5: re-modify the final predicted results with the error-corrected method. Specifically, (1) subtract the real carbon price from the predicted price obtained in step 4 to acquire the real errors and (2) use the PSO-LSTM model to forecast the errors to obtain the predicted errors, and (3) the final predicted price with the error corrected can be calculated by summing the predicted errors and predicted price.

For estimating the carbon price forecasting performance of the proposed model, this paper also constructed hybrid and single comparative models based on the decomposition technologies of CEEMDAN, EEMD, EMD, and machine learning models such as GRU and BP.

Fig. 2 The logical framework of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model
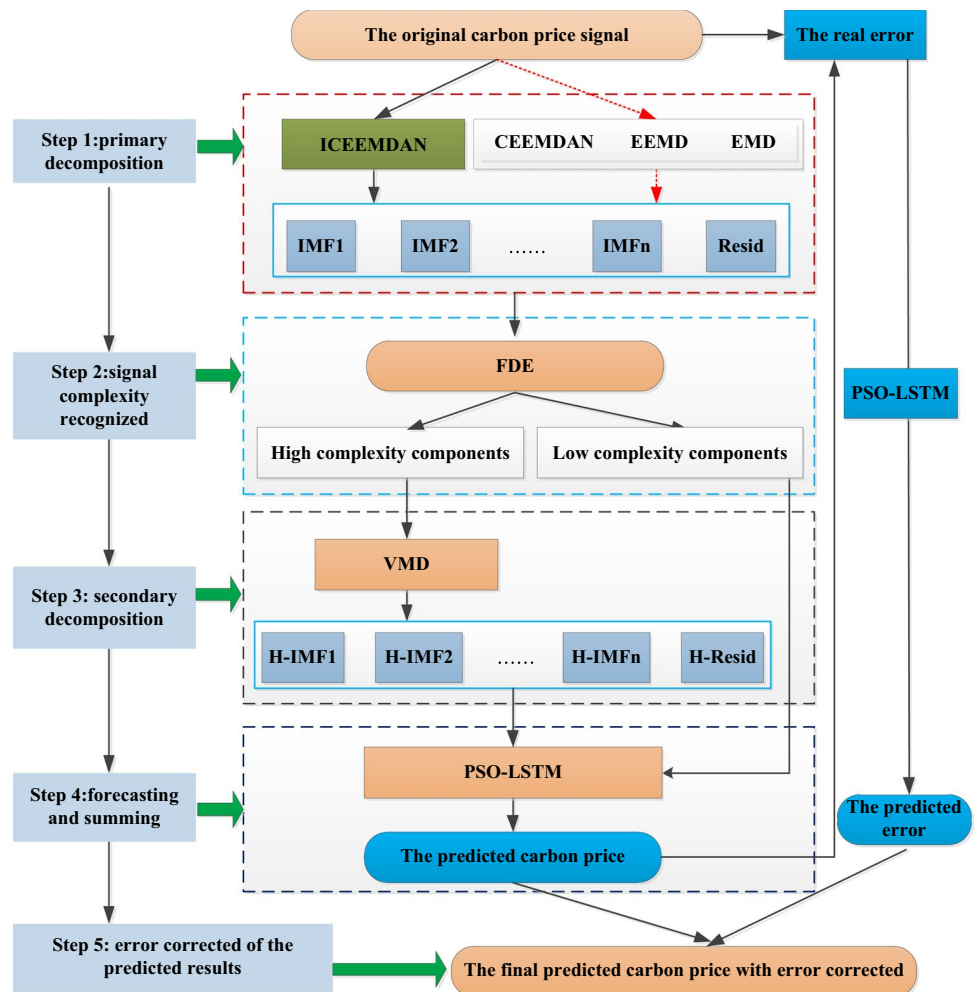
Evaluation criteria

The error evaluation indicators root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and the Pearson correlation (Corr) are conducted as criteria for evaluating the forecasting performance of the proposed model and its comparative models. Among them, the smaller the error indicators, the better the model's superiority. The Corr reflects the correlation between the final predicted price and the real price. The greater the correlation, the model forecasting accuracy is commonly recognized as higher; otherwise, the forecasting ability is poor.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{N}} \tag{19}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{20}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{21}$$

$$Corr_{y,\hat{y}} = \frac{\text{cov}\left(y,\hat{y}\right)}{\sigma_y\sigma_{\hat{y}}} = \frac{E\left[\left(y - \mu_y\right)\left(y - \mu_{\hat{y}}\right)\right]}{\sigma_y\sigma_{\hat{y}}} \tag{22}$$

Among them, y_i and $\hat{y}_i$ are the real carbon price and the predicted price, respectively, σ_y and $\sigma_{\hat{y}}$ represent the variance, μ_y and $\mu_{\hat{y}}$ indicate the mean value, and N represents the sample.

Empirical discussion

Data and basic statistics

Since October 2011, the Chinese government has carried out carbon trading pilots in Beijing, Tianjin, Shanghai, Chongqing, Hubei, Guangdong, Shenzhen, and Fujian. Due

to the large differences in economic development, energy consumption, and market regulation among pilot regions, the trading volume and trading price of each market show differences correspondingly. Especially, the Hubei carbon market has become the most representative and active trading pilot. For example, at the end of 2022, the Hubei carbon market had a total trading volume of 375 million tons, accounting for 44.6% of China's whole share. Furthermore, as the official establishment of China's unified carbon market in 2021, the Hubei carbon market plays a guiding role in market registration and trading settlement of the proposed market (Zhou and Li 2019; Zhang et al. 2020). Therefore, we select the Hubei carbon market as the research object.

The data are sourced from the Hubei carbon emissions trading center (https://www.hbets.cn), and the sample is the daily transaction price from April 28, 2014, to May 31, 2023, and the totaling 2174 samples are obtained. Additionally, after the decomposition of the original price by the EMD-type technologies, the first 80% of the obtained components are used for the parameter optimization of the neural network, and the last 20% are used for the one-step forecasting.

According to Fig. 3, the carbon price volatility is high, the nonlinear characteristics are obvious, and its distribution histogram and probability density do not conform to the normal distribution. Specifically, firstly, the average carbon price is 28.14, and the standard deviation and skewness are 10.76 and 0.624 according to Table 1. A positive right skewed distribution means that the carbon price has a right "outlier"

(Hubert and Vandervieren 2008), and its density function is different from the normal distribution form. Secondly, the ADF statistical value is 1.714, which is not statistically significant, which means that the carbon price rejects the stationary assumption, it has non-stationary characteristics. Thirdly, the critical values of BD statistics are significant at the 1% level, suggesting that carbon price has nonlinear characteristics. Therefore, the original carbon price signal has non-normal, non-stationary, and nonlinear characteristics, which make it suitable for multi-scale decomposition by the mode decomposition technology.

Decomposition of the carbon price signal

Primary decomposition by the ICEEMDAN technology

During the carbon price decomposition process of ICEEMDAN, the ratio of noise deviation to decomposed signal deviation is set to 0.3, the number of noise addition is set to 100, and the maximum iteration is set to 1000.

The results present that IMF9 components and 1 residual term can be obtained, as depicted in Fig. 4, and the signal frequency, average period, and time-frequency characteristics of the proposed IMF signals are significantly different. Specifically, the variance contribution, average period, and correlation of IMF1 are all the smallest according to Table 2, with values of 0.002, 2.812, and 0.059, respectively. Those findings mean that the IMF1 signal has a higher fluctuation
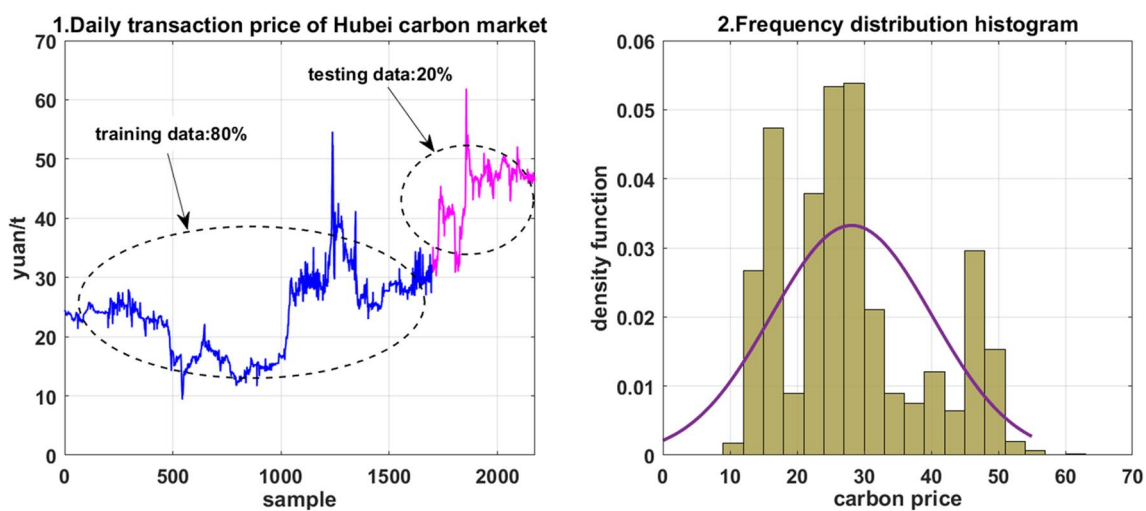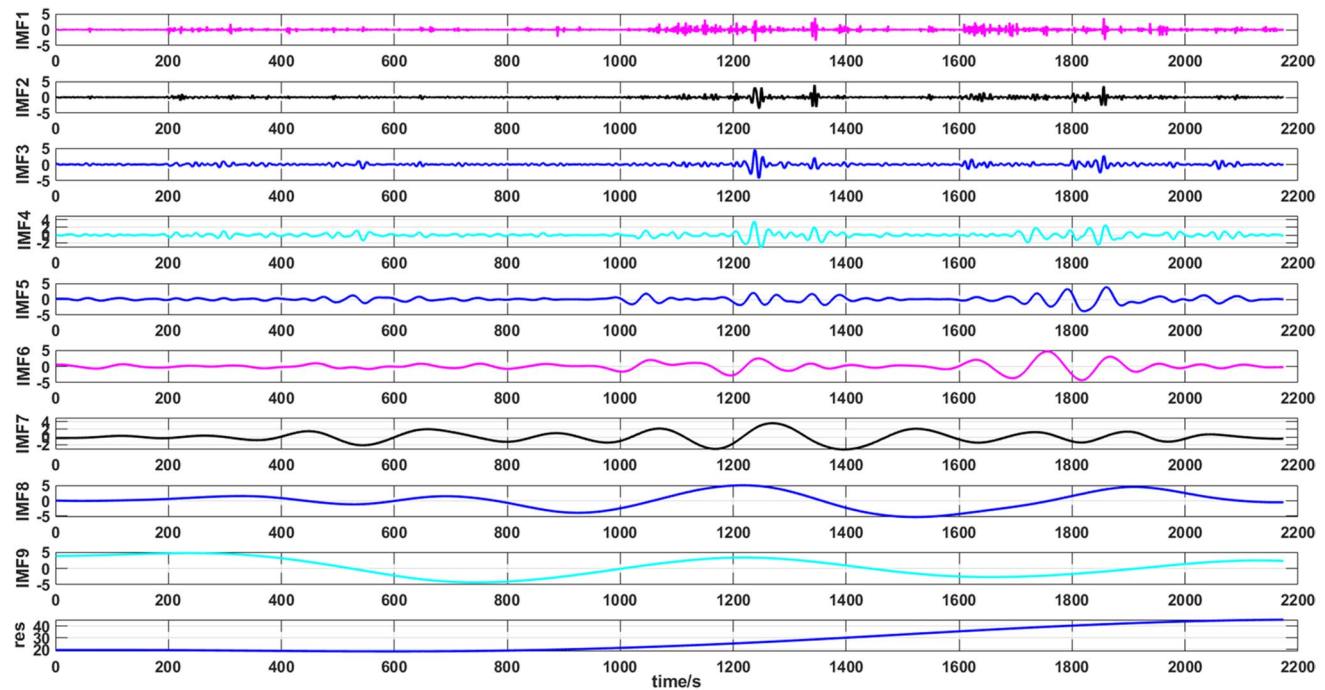

Fig. 3 The carbon price and its probability density

Table 1 Descriptive statistics (2014.4.28–2023.5.31)

Mean	Std. dev.	Skewness	Kurtosis	ADF	JB-Stat	BDS(10)-Stat
28.14	10.76	0.624	2.478	1.714	168.74***	357.29***

*** indicates the significance level under the 1%

Fig. 4 The primary decomposition of carbon price based on the ICEEMDAN model

Table 2 The mode statistic of carbon price signal after the ICEEM-DAN primary decomposition

IMFs	VarR	Average period	Pearson cor-relation	FDE
IMF1	0.002	2.812	0.059	1.486
IMF2	0.002	5.721	0.082	0.882
IMF3	0.002	11.323	0.080	0.614
IMF4	0.003	23.890	0.085	0.534
IMF5	0.006	43.480	0.079	0.358
IMF6	0.012	94.522	0.156	0.223
IMF7	0.015	197.636	0.150	0.200
IMF8	0.059	543.500	0.417	0.052
IMF9	0.065	724.667	0.249	0.027
res	0.796	2174.000	0.873	0.002

frequency and a shorter fluctuation period that may hide more decomposition noise, while the variance contribution of the residual is the highest at 0.796, with a sample period of 2174 days, and the highest correlation at 0.873. Those indicate that the decomposition noise of the residual term is low, as a result, we regard it as a long-term trend with low complexity. Furthermore, for other mode signals from IMF1 to IMF9, the variance contribution, average period and correlation gradually improve, which indicates a decrease in signal noise and an improved ability to explain the carbon premium.

Additionally, as a comparison, this article also employs CEEMDAN, EEMD, and EMD to decompose the original

carbon price (as shown in Appendix 1 of Figs. 13, 14, and 15). Especially, the variance contribution, average period, and correlation of the obtained IMF signals gradually increase, that means the volatility of the signal components gradually decreases (as shown in Appendix 2 of Tables 9, 10, and 11). Those findings are completely consistent with the decomposition results of the ICEEMDAN technology mentioned above.

Secondary decomposition by the ICEEMDAN-FDE-VMD technology

The obvious advantage of FDE is focusing on the threshold determination and mapping probability of complex fuzzy signals (Rostaghi et al. 2021). This article conducts the FDE to measure the entropy values of obtained IMF signals. The results show (the last column of Table 2) that only the entropy value of the IMF1 is greater than 1, which is 1.486, while other IMF signal entropy is less than 1. Therefore, we conclude that the IMF1 series hides more chaotic and complex information compared with other IMF signals.

Based on this, we further adopt the VMD model to secondary decompose the IFM1, empirically setting the number of decomposed signals to 9, as a result, 8 mode components and 1 residual term can be obtained (as shown in Fig. 5). Although the mode decomposition numbers of the ICEEMDAN model have increased after secondary decomposition (as shown in Fig. 6), the FDE value of the high complexity IMF1 has been greatly reduced. The entropy values of the secondary decomposition signals are all less than 1, and the whole sample

Fig. 5 The VMD secondary mode decomposition after the ICEEMDAN primary decomposition (IMF0 means the high-complexity mode signal that needs to be decomposed)
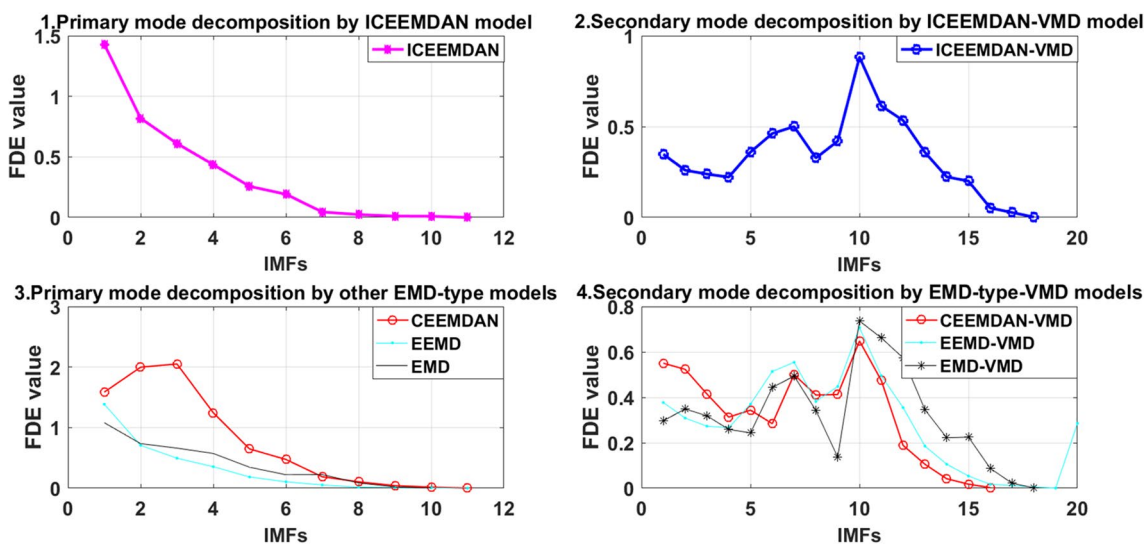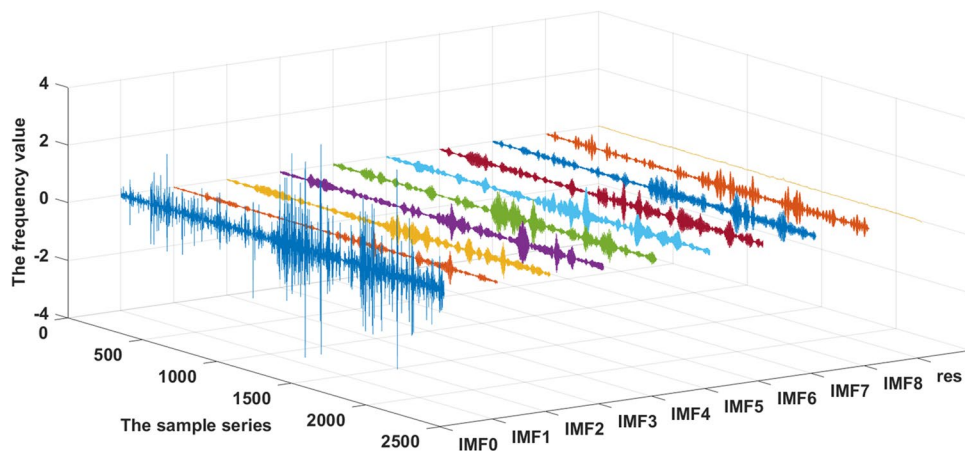
Fig. 6 The curve of fuzzy dispersion entropy based on the ICEEMDAN and other EMD-type models

Table 3 Comparison of the FDE value before and after the secondary decomposition

	Decomposition model	IMF numbers	Average FDE	Maximum value of FDE
Primary mode decomposition	ICEEMDAN	10	0.4378	1.4855
	CEEMDAN	11	0.7600	2.0514
	CEEMD	11	0.3485	1.4299
	EEMD	11	0.3022	1.3870
	EMD	10	0.3957	1.0814
Secondary mode decomposition	ICEEMDAN-VMD	18	0.3350	0.8822
	CEEMDAN-VMD	16	0.3270	0.6481
	CEEMD-VMD	19	0.2896	0.8186
	EEMD-VMD	19	0.2859	0.7074
	EMD-VMD	18	0.3200	0.7362

complexity has been reduced greatly. For example, according to the results in Table 3, the IMF number of the ICEEMDAN primary decomposition is 10, with a maximum FDE value of 1.4855 and an average FDE value of 0.4378. After secondary decomposition, the maximum and average FDE values are only 0.8822 and 0.3350. Thus, the secondary decomposition process effectively reduces the signal complexity.

Similarly, the mode number of other mode decomposition technologies also significantly increases after secondary decomposition, while the signal complexity gradually decreases. The corresponding results are shown in Appendix 3 of Figs. 16, 17, and 18.

Out-of-sample carbon price forecasting based on the proposed model

Parameter optimization

The PSO algorithm is designed to optimize the network structure of the LSTM-type hybrid carbon price forecasting models. Based on experience, we set the initial particle swarm to 20, the initial population to 2, the iterations number to 100, the learning factors c_1 and c_2 to 2, and the maximum and minimum weight to 1.2 and 0.8, respectively.

For the network design of the LSTM, more hidden layers and neurons can improve the training and forecasting performance but may lead to over-fitting or high training costs, while fewer hidden layers and neurons may lead to defects of weak predictive ability (Bengio et al. 2013; Shen et al. 2015). Therefore, we use the step-by-step experimental method to select the optimal parameters calculated by the PSO-LSTM experiment and empirically preset the number of hidden layers to 2, that is, test the validation errors of each neuron with 2, 4, 8, 16, 32, 64, 128, and 256, respectively. The parameter optimization is operated on the training set data that is classified above, with the first 90% of the training data for parameter optimization of the PSO-LSTM model, and the last 10% used to verify the forecasting performance. The results suggest (Table 4) that, when the hidden layer neurons are 2 and 256, the loss errors MAE and RMSE are the smallest, with 1.043 and 1.512, respectively. The correlation coefficient is 0.895, which is the biggest of the whole sample. The learning rate is 0.15. Therefore, we select the neurons of the PSO-LSTM model to 2 and 256 for out-of-sample forecasting.

Analysis on the superiority of ICEEMDAN-type secondary decomposition models

Secondary decomposition is re-decomposing the high-complexity signals, and its purpose is to reduce the signal complexity and improve the potential forecasting ability. The findings regarding the ICEEMDAN secondary decomposition are as follows:

(1) The forecasting errors of the secondary decomposition hybrid carbon price forecasting models are generally lower than those of the primary decomposition, and the correlation is also high. Therefore, the forecasting accuracy and reliability of the secondary decomposition models are relatively satisfactory, and the reduction of signal complexity and decomposition errors can improve the model's out-of-sample forecasting performance; these findings have been proved in the research of Li and Liu (2023) and Yang et al. (2023a, 2023b). For example, the forecasting errors RMSE, MAE, and MAPE of the secondary decomposition model ICEEM-DAN-FDE-VMD-LSTM are 0.2431, 0.1692, and 0.0038, respectively, which are significantly lower than the errors of the corresponding primary decomposition model ICEEMDAN-LSTM. The forecasting errors of the secondary decomposition model CEEMDAN-FDE-VMD-LSTM are 0.3361, 0.2476, and 0.0058, respectively, which are also lower than the errors of the corresponding primary decomposition model CEEMDAN-LSTM. Similarly, the forecasting errors of EEMD-FDE-VMD-LSTM and EMD-FDE-VMD-LSTM models are also lower than their corresponding primary decomposition model. Furthermore, according to Table 5, the forecasting ability of the secondary decomposition carbon price-forecasting model based on GRU and BP is also significantly better than their corresponding primary decomposition models.

(2) The carbon price forecasting models based on ICEEM-DAN technology have better forecasting performance, especially the secondary decomposition ICEEMDAN-type models present higher forecasting accuracy. The superiority of the ICEEMDAN-type hybrid models has been proved in previous studies of forecasting China's carbon prices (Li et al. 2022; Zhu et al. 2023). For example, for the secondary decomposition hybrid forecasting models based on LSTM, GRU, and BP, the forecasting errors of the ICEEMDAN-type models are significantly lower than those of other CEEMDAN-type, EEMD-type, and EMD-type models. As shown in Fig. 7, the deviations between the predicted price and the real price of the ICEEMDAN-FDE-VMD-LSTM, ICEEMDAN-FDE-VMD-GRU, and ICEEMDAN-FDE-VMD-BP models are relatively low. The predicted price is basically consistent with the actual one, and the difference is small. So, the forecasting performance of ICEEM-DAN-type models is superior to other CEEMDAN-type, EEMD-type, and EMD-type forecasting models.

(3) It is interesting to note that although the forecasting performance of the secondary decomposition hybrid models is better than that of the primary decomposition models,

Table 4 The parameter training of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM model

Hidden layer1	Hidden layer2	MAE	RMSE	Corr	Learning rate
2	2	3.1870	3.7947	0.8309	0.1132
	4	5.2242	5.7620	0.6924	0.1166
	8	9.1572	10.4020	0.0007	0.0886
	16	2.7895	3.1606	0.7323	0.1253
	32	7.7908	8.7560	0.2222	0.1101
	64	5.4095	6.1841	0.2685	0.1128
	128	9.9165	11.0546	0.0002	0.1080
	256	**1.0434**	**1.5121**	**0.8945**	**0.1500**
4	2	4.8227	5.3985	0.6451	0.0812
	4	2.6987	3.1431	0.7948	0.1189
	8	6.4924	7.1835	0.5467	0.0559
	16	3.0601	3.4667	0.7222	0.0779
	32	5.5428	6.2877	0.2311	0.0863
	64	1.9079	2.5335	0.7108	0.0537
	128	2.7740	3.5905	0.6799	0.1164
	256	5.8056	6.5174	0.5583	0.1482
8	2	12.5457	14.6143	0.0084	0.1500
	4	5.1572	6.4172	0.2103	0.1017
	8	4.9779	5.4816	0.7368	0.1148
	16	2.9772	3.3113	0.8504	0.0770
	32	6.9300	7.8116	0.2717	0.1246
	64	10.9621	11.9490	0.0563	0.0829
	128	7.3115	8.3651	0.2576	0.0305
	256	4.8301	5.7070	0.5469	0.0372
16	2	6.3904	7.1249	0.5156	0.0829
	4	2.0481	2.4756	0.8184	0.1203
	8	6.6193	6.9793	0.7720	0.1111
	16	3.9331	4.6887	0.5117	0.0874
	32	6.8521	7.6048	0.4719	0.0539
	64	3.6241	4.0166	0.7376	0.0819
	128	4.7044	5.4078	0.6077	0.0769
	256	7.7709	8.3073	0.5961	0.1500
32	2	1.8292	2.0845	0.9194	0.0737
	4	14.5265	16.3173	0.0002	0.1219
	8	1.3816	2.0578	0.8691	0.0794
	16	4.6863	5.2192	0.6973	0.1114
	32	4.9282	6.2662	0.4439	0.1126
	64	4.1376	4.4019	0.8718	0.1039
	128	4.7220	5.8850	0.8829	0.0383
	256	3.0483	3.5860	0.7334	0.0222
64	2	6.9378	7.1383	0.8716	0.1037
	4	8.3434	9.5272	0.0971	0.0305
	8	6.5729	7.2587	0.5059	0.0649
	16	2.5246	5.5927	0.5794	0.0750
	32	10.4071	12.0477	0.1712	0.1269
	64	7.2255	8.0758	0.3681	0.0388
	128	9.4055	10.2049	0.3152	0.0805
	256	7.6821	8.2508	0.5710	0.0486

Table 4 (continued)

Hidden layer1	Hidden layer2	MAE	RMSE	Corr	Learning rate
128	2	3.9435	4.1816	0.9009	0.0223
	4	4.7129	5.1602	0.6922	0.1130
	8	4.9992	5.2850	0.8605	0.0870
	16	8.1082	8.6965	0.5819	0.0726
	32	5.0178	5.5292	0.7049	0.1183
	64	1.6790	3.4544	0.6967	0.0294
	128	13.2428	16.2507	0.3994	0.0290
	256	5.5271	6.1242	0.6590	0.0540
256	2	4.0970	4.4851	0.8536	0.0673
	4	8.2356	8.8939	0.4763	0.0255
	8	3.2214	4.5801	0.5730	0.1372
	16	4.0611	4.3460	0.8431	0.0877
	32	10.8633	11.2055	0.7007	0.0290
	64	9.9711	10.1360	0.8460	0.0100
	128	8.7590	9.2300	0.6085	0.0220
	256	6.2200	7.1273	0.0914	0.1086

Bold numbers mean the optimal neural network parameter results

the evidence does not show the forecasting results of the primary decomposition models are completely better than the single models. For example, the forecasting errors RMSE, MAE, and MAPE of the EEMD-BP model are 3.4916, 2.9202, and 0.0618, respectively, and the forecasting errors of the EMD-BP model are 4.2496, 3.4595, and 0.0730, respectively. In contrast, the forecasting performance of the single model LSTM, GRU, and BP are significantly better than the EEMD-BP and EMD-BP primary decomposition models. Especially, the forecasting errors of the single model LSTM are lower than most primary decomposition models. One reason is that the primary decomposition process may lead to high-complexity signals and easily result in lower out-of-sample forecasting accuracy.

Analysis on the superiority of PSO-LSTM hybrid models

(1) Compared with GRU-type and BP-type hybrid and single models, the LSTM-type hybrid models and single models have relatively low forecasting errors, which suggests high forecasting accuracy and high correlation with the real price. For example, the forecasting errors RMSE, MAE, and MAPE of the ICEEMDAN-FDE-VMD-LSTM model are 0.2431, 0.1692, and 0.0038, respectively, which are significantly lower than the errors of ICEEMDAN-FDE-VMD-GRU and

Table 5 The carbon price forecasting errors and Pearson correlation of the ICEEMDAN-FDE-VMD-PSO-LSTM-EC hybrid model and its comparative models

	Model	Abbre.	RMSE	MAE	MAPE	Corr
Hybrid forecasting models	Panel A: secondary decomposition models for carbon price forecasting					
	ICEEMDAN-FDE-VMD-PSO-LSTM-EC	M1	**0.0877**	**0.0407**	**0.0009**	**0.9998**
	CEEMDAN-FDE-VMD-PSO-LSTM-EC	M2	0.1337	0.0616	0.0013	0.9996
	EEMD-FDE-VMD-PSO-LSTM-EC	M3	0.1394	0.0645	0.0014	0.9996
	EMD-FDE-VMD-PSO-LSTM-EC	M4	0.1624	0.0678	0.0014	0.9994
	ICEEMDAN-FDE-VMD-PSO-LSTM	M5	0.2167	0.1529	0.0034	0.9989
	CEEMDAN-FDE-VMD-PSO-LSTM	M6	0.3111	0.2229	0.0050	0.9977
	EEMD-FDE-VMD-PSO-LSTM	M7	0.3084	0.2365	0.0053	0.9977
	EMD-FDE-VMD-PSO-LSTM	M8	0.3321	0.2604	0.0059	0.9967
	ICEEMDAN-FDE-VMD-LSTM	M9	0.2431	0.1692	0.0038	0.9988
	CEEMDAN-FDE-VMD-LSTM	M10	0.3361	0.2476	0.0058	0.9976
	EEMD-FDE-VMD-LSTM	M11	0.3498	0.2674	0.0060	0.9974
	EMD-FDE-VMD-LSTM	M12	0.3981	0.2835	0.0064	0.9963
	ICEEMDAN-FDE-VMD-GRU	M13	0.3677	0.2568	0.0057	0.9967
	CEEMDAN-FDE-VMD-GRU	M14	0.4474	0.3502	0.0078	0.9976
	EEMD-FDE-VMD-GRU	M15	0.3869	0.2764	0.0062	0.9964
	EMD-FDE-VMD-GRU	M16	0.4495	0.3165	0.0071	0.9952
	ICEEMDAN-FDE-VMD-BP	M17	0.4294	0.2813	0.0061	0.9961
	CEEMDAN-FDE-VMD-BP	M18	0.7184	0.5233	0.0116	0.9881
	EEMD-FDE-VMD-BP	M19	0.5487	0.3705	0.0082	0.9927
	EMD-FDE-VMD-BP	M20	0.8121	0.3167	0.0119	0.9840
	Panel B: primary decomposition models for carbon price forecasting					
	ICEEMDAN-LSTM	M21	0.4557	0.3256	0.0161	0.9964
	CEEMDAN-LSTM	M22	0.7112	0.5229	0.0115	0.9879
	EEMD-LSTM	M23	0.5699	0.3676	0.0081	0.9923
	EMD-LSTM	M24	0.7645	0.5410	0.0122	0.9839
	ICEEMDAN-GRU	M25	0.4592	0.3865	0.0083	0.9990
	CEEMDAN-GRU	M26	1.1338	0.9947	0.0211	0.9965
	EEMD-GRU	M27	1.0623	0.9349	0.0198	0.9970
	EMD-GRU	M28	0.7669	0.5863	0.0125	0.9946
	ICEEMDAN-BP	M29	1.5222	1.3265	0.0287	0.9875
	CEEMDAN-BP	M30	1.8872	1.4476	0.0315	0.9328
	EEMD-BP	M31	3.4916	2.9202	0.0618	0.9129
	EMD-BP	M32	4.2496	3.4595	0.0730	0.8295
Single forecasting models	LSTM	M33	0.7032	0.5842	0.0125	0.9957
	GRU	M34	1.2518	0.8172	0.0183	0.9643
	BP	M35	1.2554	0.7933	0.0178	0.9630

Bold numbers mean the optimal forecasting results with the smallest errors and the biggest correlation

ICEEMDAN-FDE-VMD-BP models. The forecasting errors of the ICEEMDAN-LSTM model are also lower than the corresponding errors of the ICEEMDAN-GRU and ICEEMDAN-BP models. Similarly, the forecasting errors of LSTM-type models based on other CEEMDAN, EEMD, and EMD technologies are also lower than those of GRU-type and BP-type models. Those findings proved that the LSTM-type models are more suitable for out-of-sample price forecasting in China's

carbon market, which is generally consistent with the conclusion of Zhou et al. (2022).

(2) The forecasting performance of the LSTM model optimized by the PSO algorithm is obviously better than other models that have not been optimized. This indicates that the optimization of network parameters has played a positive role in improving the carbon price forecasting accuracy, that is, consistent with the optimization results proposed in the studies of Jianwei
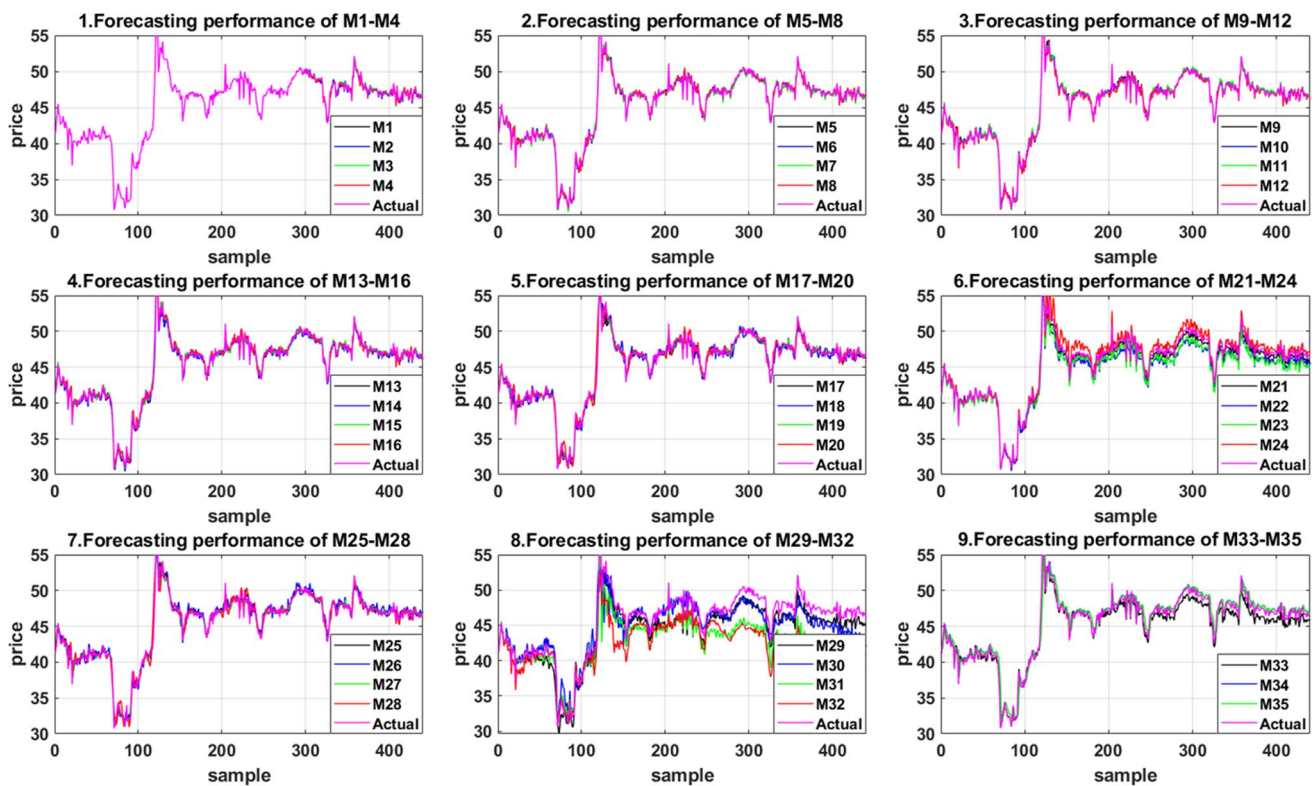
Fig. 7 The carbon price forecasting performance of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model and its comparative models

et al. ([2021](#)) and Zhang and Wu ([2022](#)). For example, the secondary decomposition model ICEEMDAN-FDE-VMD-PSO-LSTM, CEEMDAN-FDE-VMD-PSO-LSTM, EEMD-FDE-VMD-PSO-LSTM, and EMD-FDE-VMD-PSO-LSTM have lower forecasting errors than other comparative models that do not be optimized. Especially the forecasting errors RMSE, MAE, and MAPE of the ICEEMDAN-FDE-VMD-PSO-LSTM model are 0.2176, 0.1529, and 0.0034, respectively, which are the lowest among all non-error corrected hybrid models.

Discussions on the advantages of the error-corrected secondary decomposition models

The error-corrected process improves the final forecasting performance by playing the role of the predicted error term. The results in Table [5](#) show that the forecasting errors of the error-corrected secondary decomposition models are relatively lower compared with other models, and the correlation is also high, consequentially, the effectiveness of the error-corrected carbon price forecasting models can be proven. For example, the forecasting errors RMSE, MAE, and MAPE of the EMD-FDE-VMD-PSO-LSTM-EC model are 0.1624, 0.0678, and 0.0014, respectively; the forecasting errors of the EEMD-FDE-VMD-PSO-LSTM-EC model are

0.1394, 0.0645, and 0.0014, respectively; and the forecasting errors of the CEEMDAN-FDE-VMD-PSO-LSTM-EC model are 0.1337, 0.0616, and 0.0013, respectively, which are significantly lower than the forecasting performance of other comparative models without error corrected.

Discussions on the forecasting results of the ICEEMDAN-FDE-VMD-PSO-LSTM-EC model

Based on the above discussion, we find that, firstly, the hybrid carbon price forecasting models based on secondary decomposition technology have better forecasting accuracy and price correlation than other primary decomposition hybrid models. Especially, the ICEEMDAN-type hybrid models have higher forecasting ability, demonstrating the necessity of selecting ICEEMDAN and VMD technologies for the primary and secondary decomposition of original carbon price in this paper. Secondly, the forecasting performance of the deep learning LSTM-type secondary decomposition models is relatively better than the GRU-type and BP-type models, especially the LSTM-type secondary decomposition models optimized by the PSO algorithm have higher accuracy. Thirdly, the forecasting errors of the error-corrected hybrid forecasting models are significantly lower than other comparative models. Noteworthy, the forecasting errors of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model

constructed in this article are the smallest among all the comparative models, and the correlation is also the maximum value. As shown in Fig. 8, the proposed model has relatively fewer forecasting errors, and the error deviation is relatively stable.

Retesting of the proposed model based on forecasting term differences

To test the stability of the proposed model on different forecasting periods, this paper readjusts the test set and retests

the out-of-sample forecasting performance of the proposed model. That is, the data from the last 750, 500, and 250 consecutive trading days of the carbon price are intercepted as the test set, and the rest samples are used for training. Several models with better forecasting performance are conducted to test the long-term, medium-term, and short-term carbon price forecasting effects. The forecasting results are shown in Tables 6, 7, and 8, respectively.

The results show that, firstly, as for the secondary decomposition hybrid models, the long-term forecasting performance has better forecasting performance, while
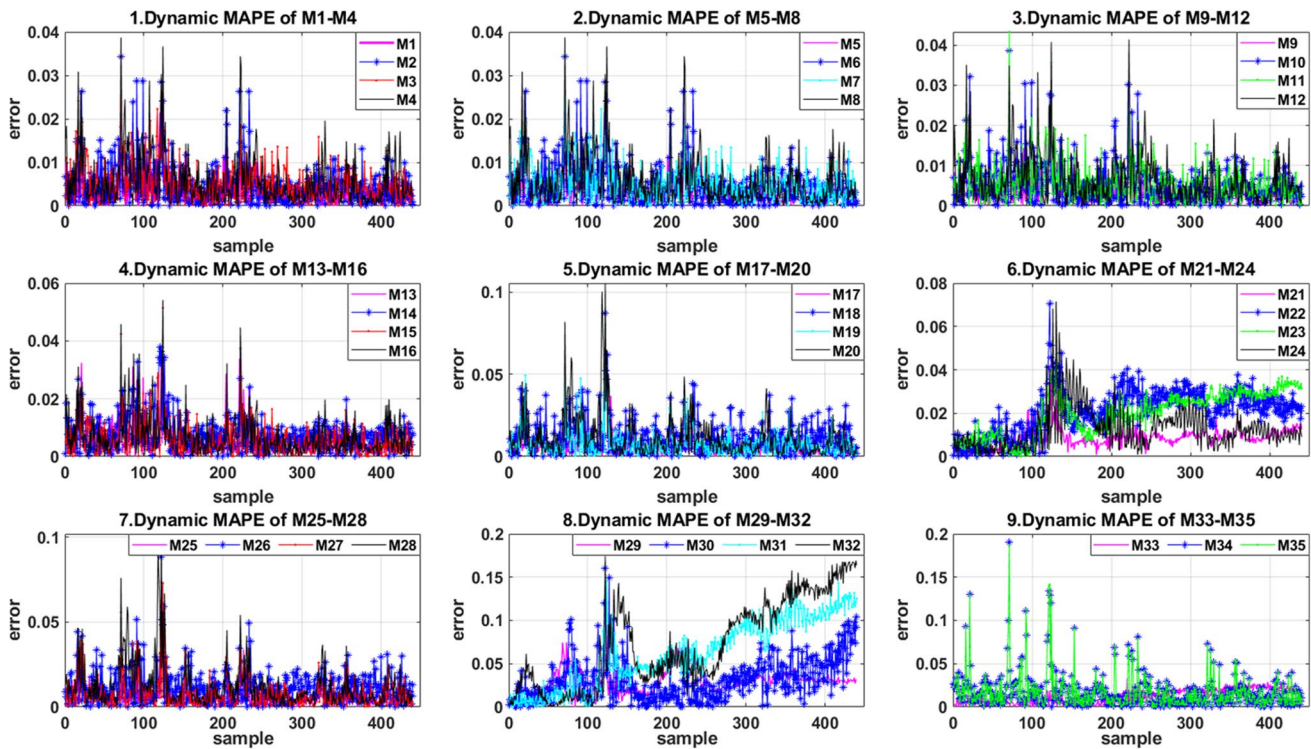
Fig. 8 The dynamic MAPE carbon price forecasting errors of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model and its comparative models

Table 6 The carbon price forecasting errors and Pearson correlation of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model and its comparative models in the long-term periods (750 daily trading data)

Model	Abbre.	RMSE	MAE	MAPE	Corr
ICEEMDAN-FDE-VMD-PSO-LSTM-EC	M1	**0.0696**	**0.0286**	**0.0006**	**0.9998**
ICEEMDAN-FDE-VMD-PSO-LSTM	M5	0.2322	0.1615	0.0044	0.9997
ICEEMDAN-FDE-VMD-LSTM	M9	0.3278	0.2311	0.0063	0.9993
ICEEMDAN-FDE-VMD-GRU	M13	0.3769	0.269	0.0074	0.9992
Average errors		0.2516	0.1726	0.0047	0.9996
ICEEMDAN-LSTM	M21	0.6941	0.5236	0.0122	0.9993
ICEEMDAN-GRU	M25	0.4374	0.3131	0.0084	0.9991
LSTM	M33	1.1503	0.9314	0.0217	0.9991
GRU	M34	1.2211	0.755	0.0204	0.9908
Average errors		0.8757	0.6308	0.0157	0.9971

Bold numbers mean the optimal forecasting results with the smallest errors and the biggest correlation

Table 7 The carbon price forecasting errors and Pearson correlation of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model and its comparative models in the medium-term periods (500 daily trading data)

Model	Abbre.	RMSE	MAE	MAPE	Corr
ICEEMDAN-FDE-VMD-PSO-LSTM-EC	M1	**0.072**	**0.0714**	**0.0009**	**0.9992**
ICEEMDAN-FDE-VMD-PSO-LSTM	M5	0.3505	0.2578	0.0062	0.9983
ICEEMDAN-FDE-VMD-LSTM	M9	0.3116	0.2816	0.0081	0.9897
ICEEMDAN-FDE-VMD-GRU	M13	0.3571	0.6911	0.0088	0.9899
Average errors		0.2728	0.3255	0.0060	0.9943
ICEEMDAN-LSTM	M21	0.4862	0.3192	0.0076	0.9969
ICEEMDAN-GRU	M25	0.636	0.4645	0.01	0.996
LSTM	M33	0.5518	0.5781	0.0085	0.9788
GRU	M34	0.2382	0.1657	0.004	0.9992
Average errors		0.4781	0.3819	0.0075	0.9927

Bold numbers mean the optimal forecasting results with the smallest errors and the biggest correlation

Table 8 The carbon price forecasting errors and Pearson correlation of the proposed ICEEMDAN-FDE-VMD-PSO-LSTM-EC model and its comparative models in the short-term periods (250 daily trading data)

Model	Abbre.	RMSE	MAE	MAPE	Corr
ICEEMDAN-FDE-VMD-PSO-LSTM-EC	M1	**0.1797**	**0.1217**	**0.0168**	**0.9659**
ICEEMDAN-FDE-VMD-PSO-LSTM	M5	0.294	0.214	0.0045	0.9801
ICEEMDAN-FDE-VMD-LSTM	M9	1.3084	1.0897	0.0229	0.7529
ICEEMDAN-FDE-VMD-GRU	M13	1.2758	1.237	0.0469	0.9581
Average errors		0.7645	0.6656	0.0228	0.9143
ICEEMDAN-LSTM	M21	0.2987	0.1619	0.0072	0.9959
ICEEMDAN-GRU	M25	0.4655	0.3679	0.0077	0.948
LSTM	M33	0.9137	0.5988	0.0126	0.7946
GRU	M34	0.3797	0.3217	0.0068	0.9659
Average errors		0.5144	0.3626	0.0086	0.9256

Bold numbers mean the optimal forecasting results with the smallest errors and the biggest correlation
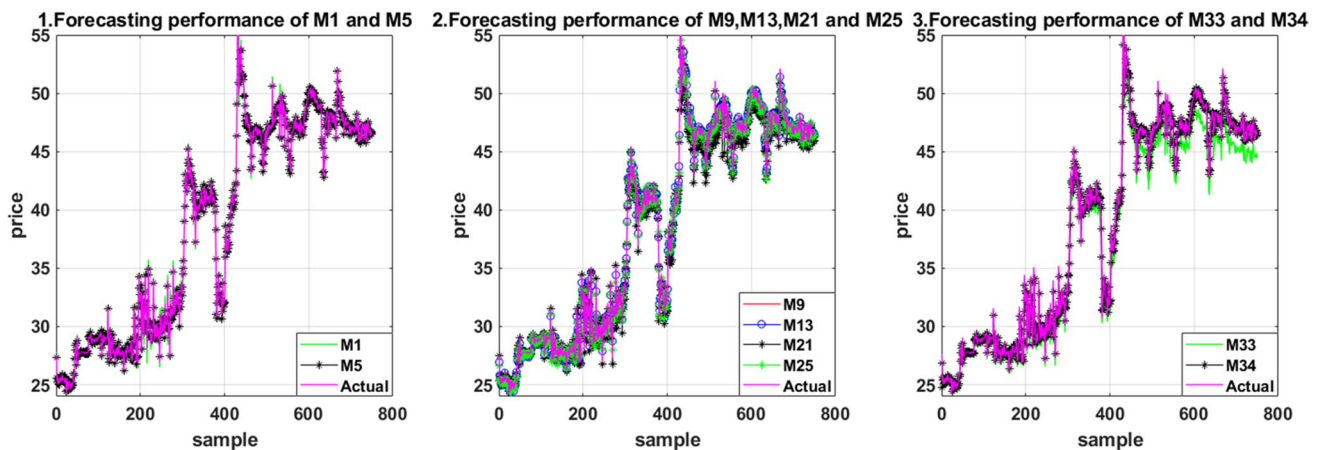
Fig. 9 The one-step-forward out-of-sample forecasting performance of the proposed model in the long-term periods (750 daily trading data)

the short-term forecasting is relatively poor (the forecasting performance of each period is shown in Figs. 9, 10, and 11). For example, the average forecasting errors RMSE, MAE, MAPE, and the correlation of the secondary decomposition models in the long-term period are 0.2516, 0.1726, 0.0047, and 0.9996, respectively (as shown in Table 6); the average errors and correlation of the secondary decomposition models in the medium-term period are 0.2728, 0.3255, 0.0060, and 0.9943, respectively (as shown in Table 7); and the average errors and correlation of the secondary decomposition models in the short-term period are 0.7645, 0.6656, 0.0228, and 0.9143
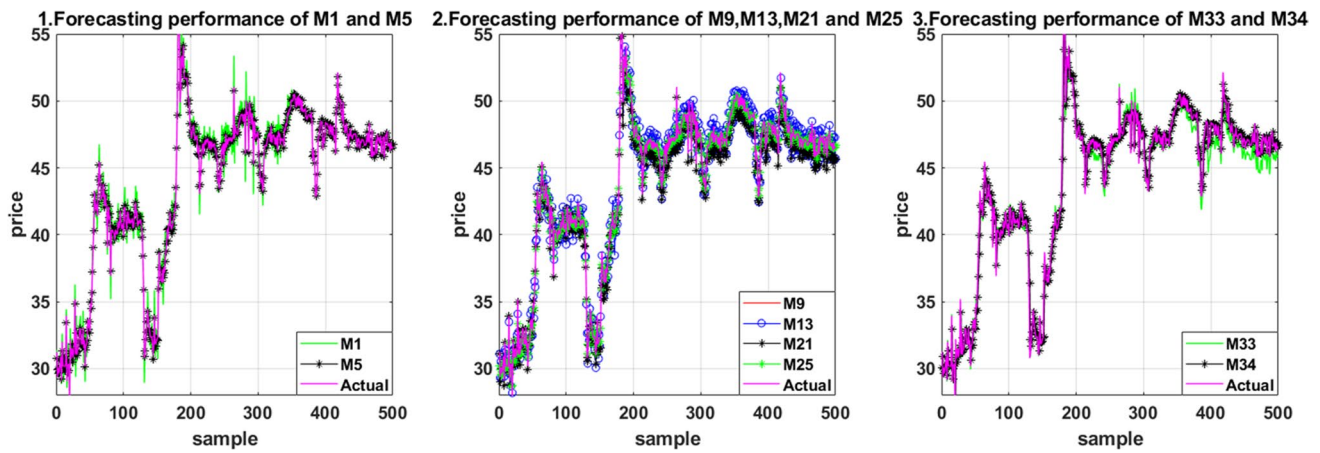
Fig. 10 The one-step-forward out-of-sample forecasting performance of the proposed model in the medium-term periods (500 daily trading data)
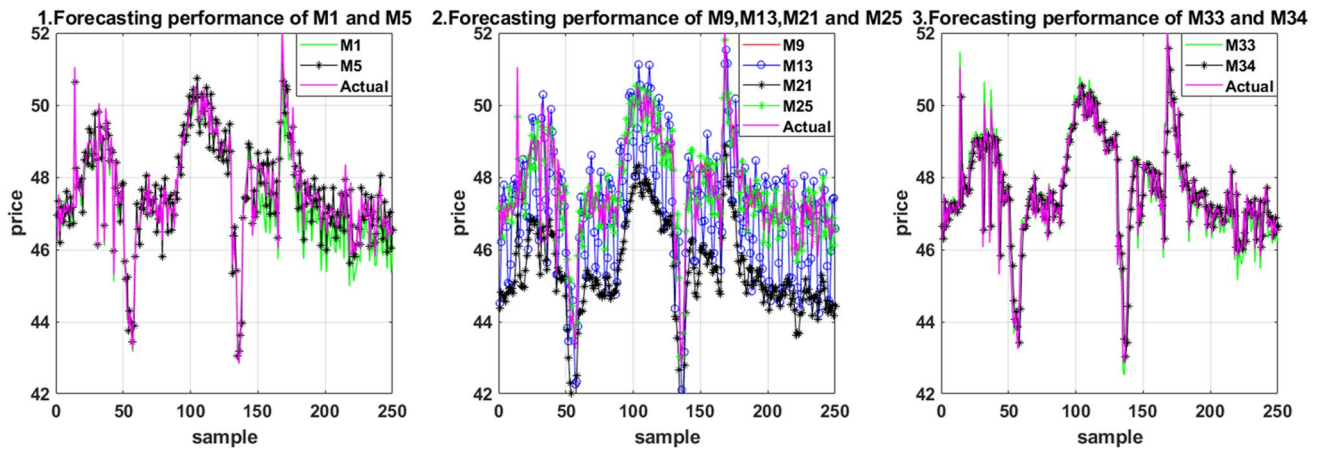

Fig. 11 The one-step-forward out-of-sample forecasting performance of the proposed model in the short-term periods (250 daily trading data)

respectively (as shown in Table 8). This evidence reflects that the secondary decomposition hybrid models are better for performing price forecasting in long-term periods. One possible reason is that the secondary decomposition process can reduce the decomposition errors and signal noise effectively (Li et al. 2022; Sun et al. 2022), especially the large amount of IMF predicted data can maximally offset the forecasting errors during the price integration stage, so as to improve the forecasting accuracy.

Secondly, the ICEEMDAN-FDE-VMD-PSO-LSTM-EC model constructed in this paper has the best forecasting performance and accuracy in all the periods, especially in the long-term period, which is significantly superior to other comparative models. This conclusion is completely consistent with the findings in the "Discussions on the forecasting results of the ICEEMDAN-FDE-VMD-PSO-LSTM-EC model" section above. From the error curve depicted in Fig. 12, it is obvious that the dynamic errors

of the proposed model are relatively small compared with other comparative models; this finding reflects the higher fitting ability between the predicted price and the real price.

Thirdly, as for the primary decomposition models and the single models, the long-term forecasting performance is the worst. For example, in the long-term forecasting period, the average errors RMSE, MAE, and MAPE of the ICEEMDAN-LSTM, ICEEMDAN-GRU, LSTM, and GRU are 0.8757, 0.6308, 0.0157, and 0.9971, respectively. Those errors are significantly higher than the medium-term and short-term forecasting performance, and the correlation is also lower than other periods. This evidence maintains that compared with the long-term forecasting advantage of the secondary decomposition models, the primary decomposition carbon price forecasting models are more suitable for medium and short-term price forecasting.
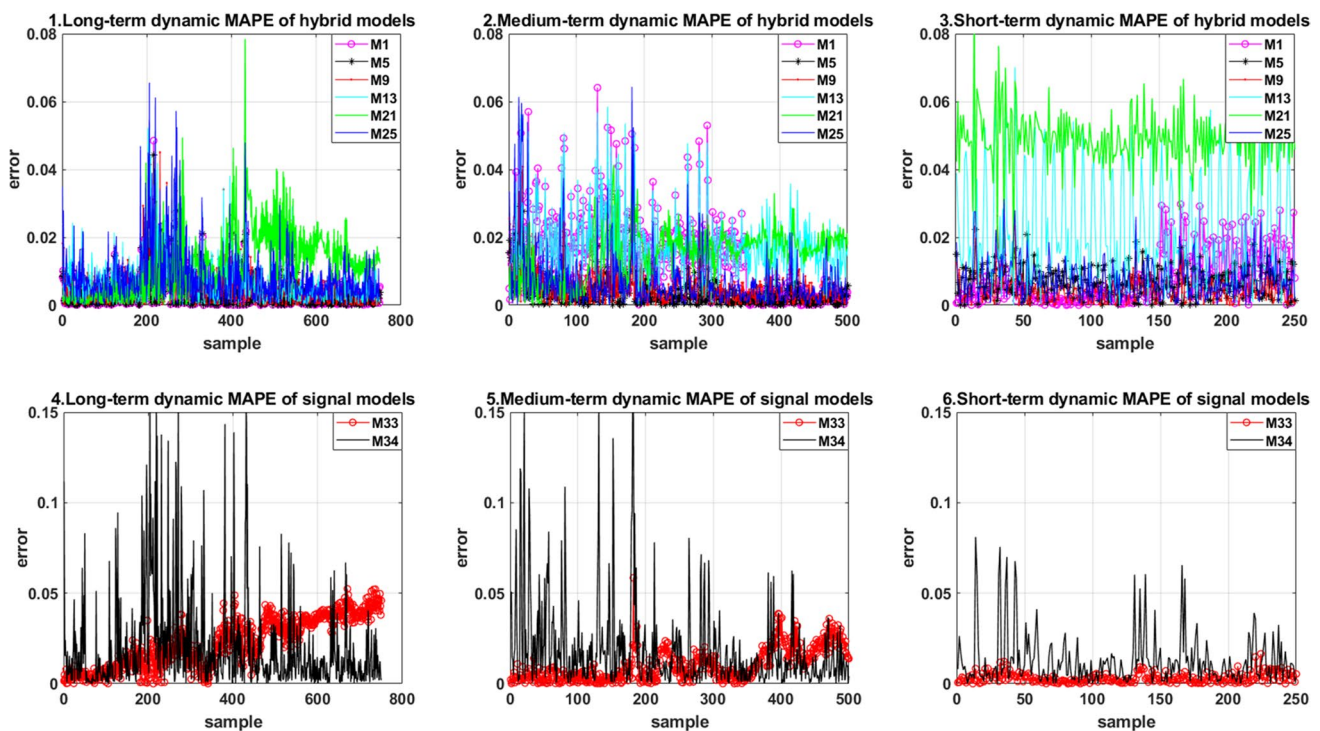
Fig. 12 The dynamic MAPE carbon price forecasting errors of the proposed model in the long-medium short-term periods

Conclusions

The carbon price not only reflects the supply and demand of market allowances but also reveals the nonlinear price formation mechanism. As of July 2023, it has been 2 years since China's power industry was formally included in the national carbon emissions trading system (ETS). With the development of the carbon market, more and more Chinese companies have begun to incorporate the emissions cost into their daily business decision. Therefore, this study focuses on the emerging China carbon market and constructs a novel error-corrected secondary decomposition hybrid model integrated fuzzy dispersion entropy and deep learning paradigm to forecast the carbon price. The main conclusions are as follows:

Firstly, the forecasting performance of the ICEEMDAN-type secondary decomposition hybrid carbon price forecasting models is significantly better than primary decomposition models, and other CEEMDAN, EEND, and EMD-type secondary decomposition hybrid models. These findings show the signal decomposition efficiency of ICEEMDAN technology is relatively high, which can reduce the decomposition errors and improve the forecasting accuracy. Furthermore, the FDE algorithm plays a positive role in identifying high-complexity signals for price forecasting. These conclusions provide modeling ideas for revealing the formation mechanism of complex carbon prices. Secondly, the

deep learning paradigm of LSTM-type models optimized by the PSO algorithm has obvious advantages in fitting and forecasting China's carbon price, the advantage of the LSTM model in dealing with financial time series has been proved. Thirdly, the error-corrected method for improving the forecasting accuracy has achieved satisfactory results. Especially, the ICEEMDAN-FDE-VMD-PSO-LSTM-EC model presents the best forecasting ability, which can provide more accurate modeling technology for investors to take carbon market trading. Finally, the proposed model has obvious superiority in different forecasting periods, particularly, the long-term forecasting for 750 consecutive trading prices is outstanding. This result shows that the proposed model is more suitable for long-term price forecasting in China's carbon market.

The above conclusions provide a valuable reference for judging the price characteristics of China's carbon market and formulating effective market regulations. However, as for the applied aspect, there are some limitations. (1) Due to the limited trading data of China's national carbon emissions trading system, we only use the Hubei carbon price to represent the situation of the whole China's carbon market, and future studies can be developed as the enrichment of the representative price. (2) More additional factors, such as the price and consumption of fossil fuels, the trading volume of carbon allowance, and the air pollution index, can also be included in future price forecasting studies.

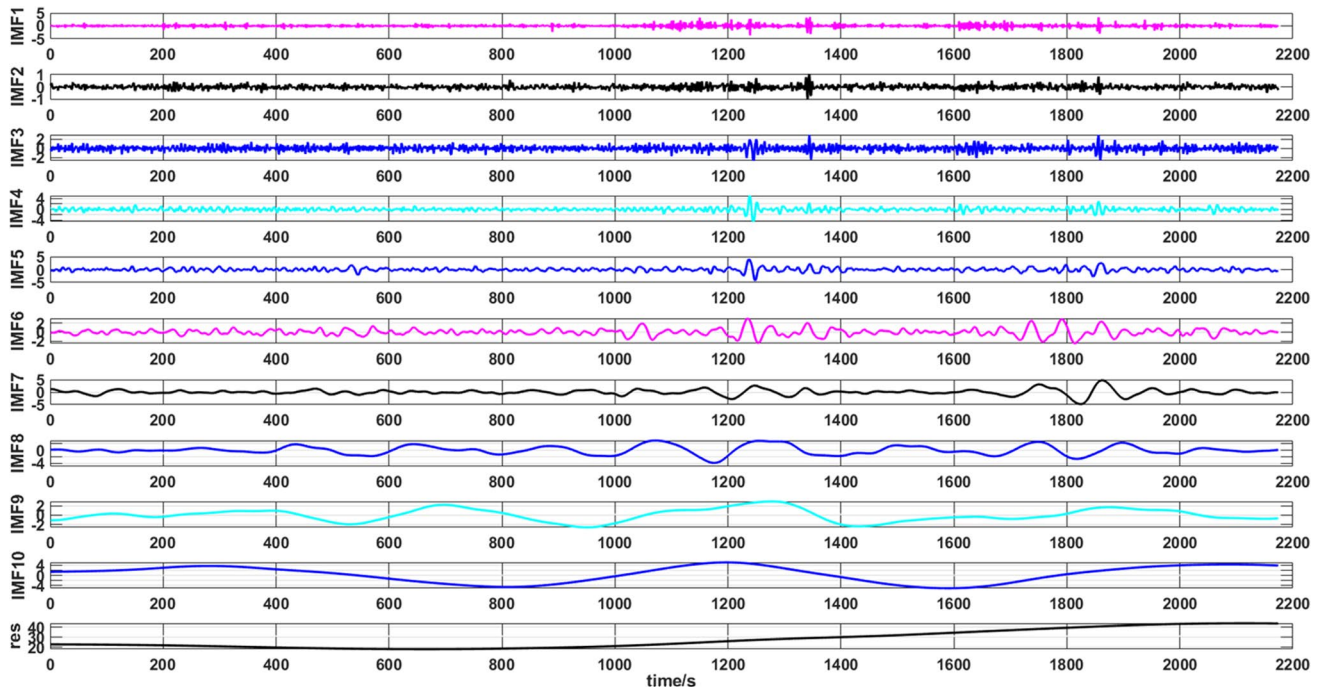# Appendix 1. Primary decomposition of the carbon price signals based on the EMD-type technologies

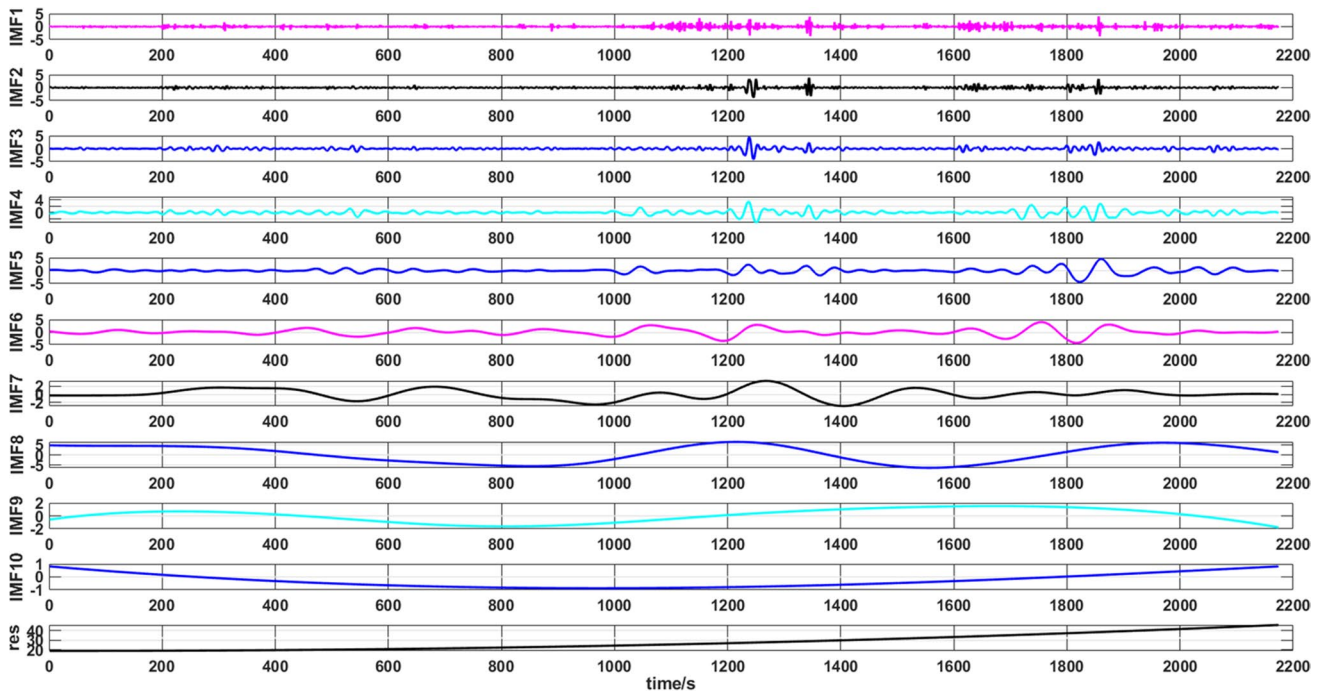**Fig. 13** The primary decomposition of carbon price based on the CEEMDAN model

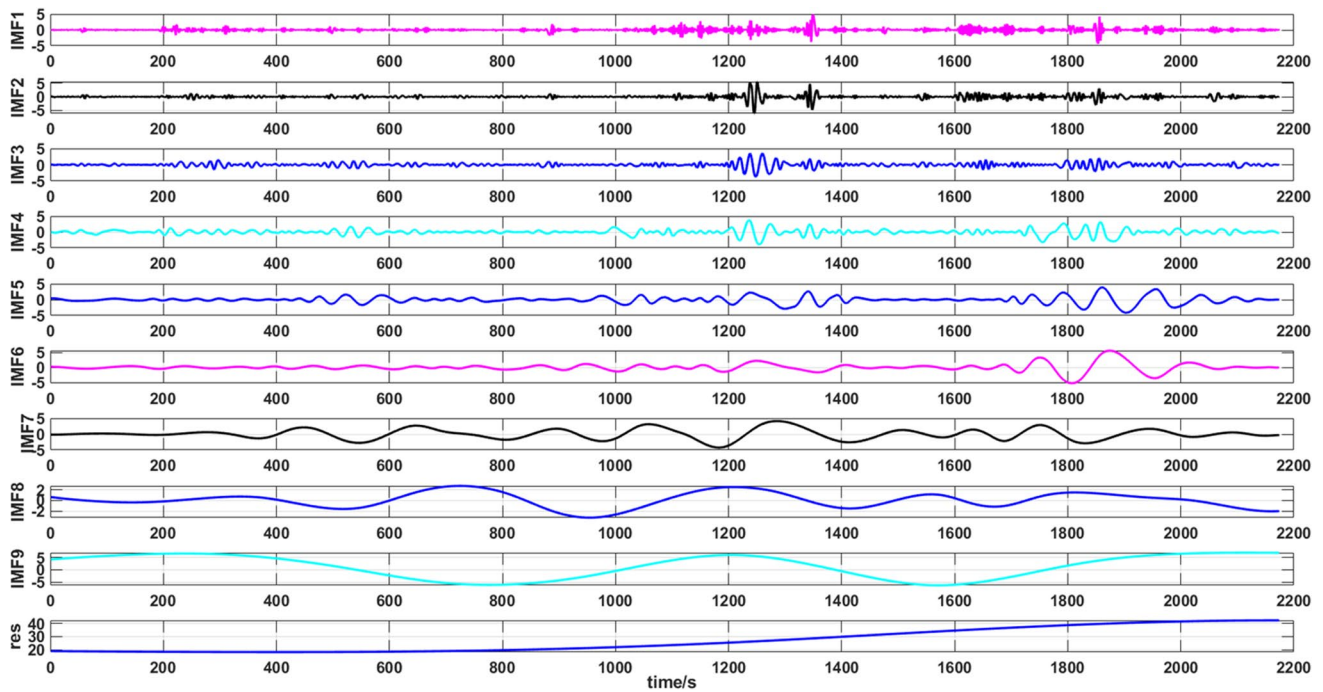**Fig. 14** The primary decomposition of carbon price based on the EEMD model

Fig. 15 The primary decomposition of carbon price based on the EMD model

Appendix 2. The mode statistic of carbon price signal after the primary decomposition

Table 9 The mode statistic of carbon price signal after the CEEM-DAN primary decomposition

IMFs	VarR	Average period	Pearson correlation	FDE
IMF1	0.002	2.812	0.059	1.5865
IMF2	0.003	3.755	0.046	1.9978
IMF3	0.003	3.490	0.041	2.0514
IMF4	0.004	6.194	0.064	1.2452
IMF5	0.004	12.282	0.069	0.6481
IMF6	0.005	24.156	0.099	0.4744
IMF7	0.010	48.311	0.148	0.1886
IMF8	0.015	108.700	0.168	0.1056
IMF9	0.015	310.571	0.309	0.0417
IMF10	0.083	724.667	0.512	0.0177
res	0.707	2174.000	0.901	0.0026

Table 10 The mode statistic of carbon price signal after the EEMD primary decomposition

IMFs	VarR	Average period	Pearson correlation	FDE
IMF1	0.002	2.791	0.057	1.387
IMF2	0.002	5.924	0.076	0.7074
IMF3	0.003	12.213	0.085	0.4953
IMF4	0.003	25.576	0.085	0.3549
IMF5	0.006	51.762	0.121	0.1856
IMF6	0.013	127.882	0.198	0.1057
IMF7	0.012	241.556	0.278	0.0536
IMF8	0.142	1087.000	0.548	0.0168
IMF9	0.013	1087.000	0.496	0.0121
IMF10	0.005	1837.219	0.590	0.0053
res	0.532	2174.000	0.846	0.0005

**Table 11** The mode statistic of carbon price signal after the EMD primary decomposition

IMFs	VarR	Average period	Pearson correlation	FDE
IMF1	0.003	3.106	0.049	1.0814
IMF2	0.004	6.141	0.049	0.7362
IMF3	0.004	12.282	0.047	0.6618
IMF4	0.006	23.630	0.027	0.5736
IMF5	0.010	47.261	0.044	0.3459
IMF6	0.014	86.960	0.158	0.2223
IMF7	0.022	181.167	0.139	0.2251
IMF8	0.017	434.800	0.133	0.0872
IMF9	0.183	724.667	0.525	0.0217
res	0.672	2174.000	0.857	0.0017

Appendix 3. The VMD secondary mode decomposition process of the high-complexity signals that recognized by the primary decomposition

Fig. 16 The VMD secondary mode decomposition after the CEEMDAN primary decomposition (IMF0 means the high complexity mode signal that needs to be decomposed)
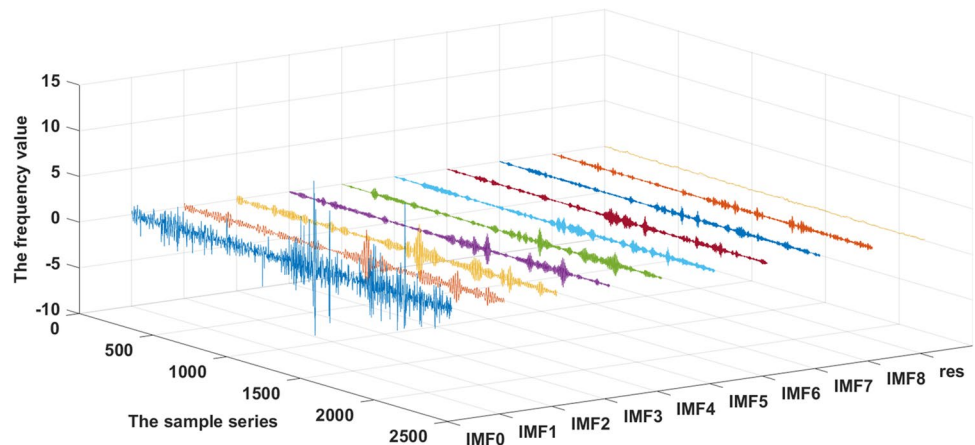
Fig. 17 The VMD secondary mode decomposition after the EEMD primary decomposition (IMF0 means the high complexity mode signal that needs to be decomposed)
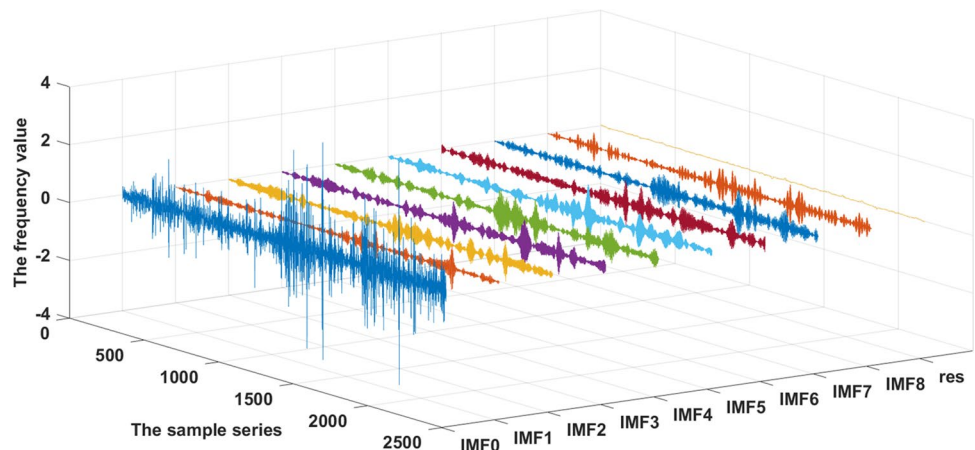
Fig. 18 The VMD secondary mode decomposition after the EMD primary decomposition (IMF0 means the high complexity mode signal that needs to be decomposed)

Declarations

Ethics approval This study follows all ethical practices during writing.

Consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

References

Bengio Y, Courville A, Vincent P (2013) Representation learning: a review and new perspectives. IEEE T Pattern Anal 35(8):1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Byun SJ, Cho H (2013) Forecasting carbon futures volatility using GARCH models with energy volatilities. Energy Econ 40:207–221. https://doi.org/10.1016/j.eneco.2013.06.017

Chevallier J (2009) Carbon futures and macroeconomic risk factors: a view from the EU ETS. Energy Econ 31:614–625. https://doi.org/10.1016/j.eneco.2009.02.008

Colominas MA, Schlotthauer G, Torres ME (2014) Improved complete ensemble EMD: a suitable tool for biomedical signal processing. Biomed Signal Process 14:19–29. https://doi.org/10.1016/j.bspc.2014.06.009

Dragomiretskiy K, Zosso D (2013) Variational mode decomposition. IEEE T Signal Process 62(3):531–544. https://doi.org/10.1109/TSP.2013.2288675

Guo W, Liu Q, Luo Z, Tse Y (2022) Forecasts for international financial series with VMD algorithms. J Asian Econ 80:101458. https://doi.org/10.1016/j.asieco.2022.101458

Hao Y, Tian C, Wu C (2020) Modelling of carbon price in two real carbon trading markets. J Clean Prod 244:118556. https://doi.org/10.1016/j.jclepro.2019.118556

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Huang Y, Dai X, Wang Q, Zhou D (2021) A hybrid model for carbon price forecasting using GARCH and long short-term memory network. Appl Energ 285:116485. https://doi.org/10.1016/j.apenergy.2021.116485

Hubert M, Vandervieren E (2008) An adjusted boxplot for skewed distributions. Comput Stat Data An 52(12):5186–5520. https://doi.org/10.1016/j.csda.2007.11.008

Jianwei E, Ye J, He L, Jin H (2021) A denoising carbon price forecasting method based on the integration of kernel independent component analysis and least squares support vector regression. Neurocomputing 434:67–79. https://doi.org/10.1016/j.neucom.2020.12.086

Junior PO, Tiwari AK, Padhan H, Alagidede I (2020) Analysis of EEMD-based quantile-in-quantile approach on spot-futures prices of energy and precious metals in India. Resour Policy 68:101731. https://doi.org/10.1016/j.resourpol.2020.101731

Kong F, Song J, Yang Z (2022) A novel short-term carbon emission prediction model based on secondary decomposition method and long short-term memory network. Environ Sci Pollut Res 29(43):64983–64998. https://doi.org/10.1007/s11356-022-20393-w

Li G, Yin S, Yang H (2022) A novel crude oil prices forecasting model based on secondary decomposition. Energy 257:124684. https://doi.org/10.1016/j.energy.2022.124684

Li H, Jin F, Sun S, Li Y (2021) A new secondary decomposition ensemble learning approach for carbon price forecasting. Knowl-Based Syst 214:106686. https://doi.org/10.1016/j.knosys.2020.106686

Li J, Liu D (2023) Carbon price forecasting based on secondary decomposition and feature screening. Energy 278:127783. https://doi.org/10.1016/j.energy.2023.127783

Liu YL, Zhang JJ, Fang Y (2023) The driving factors of China's carbon prices: evidence from using ICEEMDAN-HC method and quantile regression. Financ Res Lett 54:103756. https://doi.org/10.1016/j.frl.2023.103756

Mao S, Zeng XJ (2023) SimVGNets: similarity-based visibility graph networks for carbon price forecasting. Expert Syst Appl:120647. https://doi.org/10.1016/j.eswa.2023.120647

Nazifi F, Milunovich G (2010) Measuring the impact of carbon allowance trading on energy prices. Energy Environ 21(5):367–383. https://doi.org/10.1260/0958-305X.21.5.367

Nguyen THT, Phan QB (2022) Hourly day ahead wind speed forecasting based on a hybrid model of EEMD, CNN-Bi-LSTM embedded with GA optimization. Energy Rep 8:53–60. https://doi.org/10.1016/j.egyr.2022.05.110

Pan D, Zhang C, Zhu D, Hu S (2023) Carbon price forecasting based on news text mining considering investor attention. Environ Sci Pollut Res 30(11):28704–28717. https://doi.org/10.1007/s11356-022-24186-z

Qin Q, He H, Li L, He LY (2020) A novel decomposition-ensemble based carbon price forecasting model integrated with local polynomial prediction. Comput Econ 55:1249–1273. https://doi.org/10.1007/s10614-018-9862-1

Rostaghi M, Khatibi MM, Ashory M, Azami H (2021) Fuzzy dispersion entropy: a nonlinear measure for signal analysis. IEEE T Fuzzy Syst 30(9):3785–3796. https://doi.org/10.1109/TFUZZ.2021.3128957

Schneider L, Hoz L, Theuer S (2019) Environmental integrity of international carbon market mechanisms under the Paris Agreement. Clim Policy 19(3):386–400. https://doi.org/10.1080/14693062.2018.1521332

Shen F, Chao J, Zhao J (2015) Forecasting exchange rate using deep belief networks and conjugate gradient method. Neurocomputing 167:243–253. https://doi.org/10.1016/j.neucom.2015.04.071

Sun J, Zhao P, Sun S (2022) A new secondary decomposition-reconstruction-ensemble approach for crude oil price forecasting. Resour Policy 77:102762. https://doi.org/10.1016/j.resourpol.2022.102762

Sun W, Huang C (2020) A carbon price prediction model based on secondary decomposition algorithm and optimized back propagation neural network. J Clean Prod 243:118671. https://doi.org/10.1016/j.jclepro.2019.118671

Sun W, Zhang C (2018) Analysis and forecasting of the carbon price using multi—resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. Appl Energ 231:1354–1371. https://doi.org/10.1016/j.apenergy.2018.09.118

Tang BJ, Gong PQ, Shen C (2017) Factors of carbon price volatility in a comparative analysis of the EUA and sCER. Ann Oper Res 255:157–168. https://doi.org/10.1007/s10479-015-1864-y

Wang D, Tan D, Liu L (2018) Particle swarm optimization algorithm: an overview. Soft Comput 22:387–408. https://doi.org/10.1007/s00500-016-2474-6

Wang J, Cheng Q, Sun X (2022) Carbon price forecasting using multiscale nonlinear integration model coupled optimal feature reconstruction with biphasic deep learning. Environ Sci Pollut Res 29(57):85988–86004. https://doi.org/10.1007/s11356-021-16089-2

Wang J, Sun X, Cheng Q, Cui Q (2021) An innovative random forest-based nonlinear ensemble paradigm of improved feature extraction and deep learning for carbon price forecasting. Sci Total Environ 762:143099. https://doi.org/10.1016/j.scitotenv.2020.143099

Wu Q, Liu Z (2020) Forecasting the carbon price sequence in the Hubei emissions exchange using a hybrid model based on ensemble empirical mode decomposition. Energy Sci Eng 8(8):2708–2721. https://doi.org/10.1002/ese3.703

Wu Y, Zhang C, Yun P et al (2021) Time-frequency analysis of the interaction mechanism between European carbon and crude oil markets. Energy Environ 32(7):1331–1357. https://doi.org/10.1177/0958305X211002457

Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Data Sci Adapt 1(1):1–41. https://doi.org/10.1142/S1793536909000047

Yang H, Yang X, Li G (2023a) Forecasting carbon price in China using a novel hybrid model based on secondary decomposition, multi-complexity and error correction. J Clean Prod 401:136701. https://doi.org/10.1016/j.jclepro.2023.136701

Yang R, Liu H, Li Y (2023b) An ensemble self-learning framework combined with dynamic model selection and divide-conquer strategies for carbon emissions trading price forecasting. Chaos Soliton Fract 173:113692. https://doi.org/10.1016/j.chaos.2023.113692

Yang S, Chen D, Li S, Wang W (2020) Carbon price forecasting based on modified ensemble empirical mode decomposition and long short-term memory optimized by improved whale optimization algorithm. Sci Total Environ 716:137117. https://doi.org/10.1016/j.scitotenv.2020.137117

Yue W, Zhong W, Xiaoyi W, Xinyu K (2023) Multi-step-ahead and interval carbon price forecasting using transformer-based hybrid model. Environ Sci Pollut Res 30:95692–95719. https://doi.org/10.1007/s11356-023-29196-z

Yun P, Huang X, Wu Y, Yang X (2023) Forecasting carbon dioxide emission price using a novel mode decomposition machine learning hybrid model of CEEMDAN-LSTM. Energy Sci Eng 11(1):79–96. https://doi.org/10.1002/ese3.1304

Zhang C, Yun P, Wagan ZA (2019) Study on the wandering weekday effect of the international carbon market based on trend moderation effect. Financ Res Lett 28:319–327. https://doi.org/10.1016/j.frl.2018.05.014

Zhang J, Li D, Hao Y, Tan Z (2018) A hybrid model using signal processing technology, econometric models and neural network for carbon spot price forecasting. J Clean Prod 204:958–964. https://doi.org/10.1016/j.jclepro.2018.09.071

Zhang K, Yang X, Wang T, Thé J, Tan Z, Yu H (2023) Multi-step carbon price forecasting using a hybrid model based on multivariate decomposition strategy and deep learning algorithms. J Clean Prod 405:136959. https://doi.org/10.1016/j.jclepro.2023.136959

Zhang W, Li J, Li G, Guo S (2020) Emission reduction effect and carbon market efficiency of carbon emissions trading policy in China. Energy 196:117117. https://doi.org/10.1016/j.energy.2020.117117

Zhang W, Wu Z (2022) Optimal hybrid framework for carbon price forecasting using time series analysis and least squares support vector machine. J Forecasting 41(3):615–632. https://doi.org/10.1002/for.2831

Zhou F, Huang Z, Zhang C (2022) Carbon price forecasting based on CEEMDAN and LSTM. Appl Energ 311:118601. https://doi.org/10.1016/j.apenergy.2022.118601

Zhou J, Wang Q (2021) Forecasting carbon price with secondary decomposition algorithm and optimized extreme learning machine. Sustainability 13:8413. https://doi.org/10.3390/su13158413

Zhou K, Li Y (2019) Carbon finance and carbon market in China: progress and challenges. J Clean Prod 214:536–549. https://doi.org/10.1016/j.jclepro.2018.12.298

Zhu B, Wang P, Chevallier J, Wei YM (2015) Carbon price analysis using empirical mode decomposition. Comput Econ 45:195–206. https://doi.org/10.1007/s10614-013-9417-4

Zhu B, Ye S, Wang P, He K, Zhang T, Wei YM (2018) A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting. Energy Econ 70:143–157. https://doi.org/10.1016/j.eneco.2017.12.030

Environmental Science and Pollution Research (2024) 31:16530–16553 16553

Zhu J, Wu P, Chen H, Liu J, Zhou L (2019) Carbon price forecasting with variational mode decomposition and optimal combined model. Physica A 519:140–158. https://doi.org/10.1016/j.physa.2018.12.017

Zhu T, Wang W, Yu M (2023) A novel hybrid scheme for remaining useful life prognostic based on secondary decomposition, BiGRU and error correction. Energy 276:127565. https://doi.org/10.1016/j.energy.2023.127565