Carbon price prediction for China’s ETS pilots using variational mode decomposition and optimized extreme learning machine

Shanglei Chai1 · Zixuan Zhang1 · Zhen Zhang2

Accepted: 29 October 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
With the national goal of “carbon peak by 2030 and carbon neutral by 2060 in China”, studies on carbon prices of China’s Emissions Trading System (ETS) pilots have shown growing interest in the related fields. Carbon price fluctuations reflect the scarcity of carbon resources, and accurate prediction can improve carbon asset management capabilities. Therefore, in order to clarify the dynamics of carbon markets and assign carbon emissions allocation rationally, we propose a hybrid feature-driven forecasting model with the framework of decomposition-reconstruction-prediction-ensemble. In this paper, the non-stationary, nonlinear and chaotic characteristics of carbon prices in China’s ETS pilots have been verified, and then the prediction model is built based on the tested features. Firstly, the original carbon price series are decomposed by Variational Mode Decomposition (VMD), and then reconstructed by Sample Entropy (SE). Next, Extreme Learning Machine (ELM) optimized by Particle Swarm Optimization (PSO) is conducted to predict the subsequences. Lastly, the forecasting series of every subseries are summed to obtain the final results. The empirical results based on carbon prices of China’s ETS pilots proved that the proposed model performs more efficiently than the current benchmark models. As carbon prices are expected to increase across all ETS during the post-COVID-19 recovery stage, the new prediction model will be useful for improving the guiding principles of the existing government policies including the likely introductions of Border Carbon Adjustment (BCA) in the EU and the US, and governing the large global public companies to deliver their “net zero” commitments.

Keywords Carbon price forecasting · Emissions trading system (ETS) · Variational mode decomposition (VMD) · Particle swarm optimization (PSO) · Extreme learning machine (ELM)
1 Introduction

Climate change is one of the global problems that is threatening the sustainable social development and safety of human lives, hence drawing profound attention of the communities. The academic community tends to proclaim a high degree of certainty about the causes of climate change (Macdonald et al., 2019), evolution trends (Westerhold et al., 2020) and the risks caused by it (Trisos et al., 2020). However, the main inducing factor of climate change is carbon emissions, which has been used as a significant reference point for the evaluation of future climate change in the prediction models (Tierney et al., 2020). In this connection, it can be perceived that reducing carbon emissions is the key to mitigating climate change and enactments of a variety of environmental regulation mechanisms, such as mandatory, voluntary, and market-oriented systems have been observed to reduce emissions (Chen et al., 2021). As a market-oriented policy instrument in particular, carbon emissions trading has been recognized to be an effective mechanism to achieve low-cost emissions reduction (Bauer et al., 2020; Cui et al., 2014; Daskalakis, 2013; Zhang & Wei, 2010).

In the recent decades, China has been facing serious environmental problems. It is predicted that carbon intensity in the middle region will reach the highest in 2023 (Ma et al., 2020). The introduction of the Emissions Trading System (ETS) to tackle climate change has therefore been put forward on the policy agenda. Since 2013, China’s ETS pilots have been launched successively in Shenzhen, Shanghai, Beijing, Guangdong, Tianjin, Hubei, Chongqing and Fujian provinces. These pilots have covered nearly 3000 key emissions enterprises with a cumulative trading volume of 455 million tons, and a turnover of 10.55 billion yuan (approximately) till December 2020. The booming development of eight pilots thus promote the realization of “carbon peak by 2030 and carbon neutral by 2060 in China” (Liu & Sun, 2021; Qie et al., 2021; Xiao et al., 2021).

However, the carbon price as a signal that transmits information in carbon markets, fluctuates violently due to a large number of potential factors such as macroeconomic events (Duc et al., 2021; Koch et al., 2014), energy prices (Adekoya, 2021; Balcilar et al., 2016), politically-driven policies (Duan et al., 2018), and overall situation of industry development from the macro and long-term perspectives as well as the supply and demand of carbon assets from the micro and short-term perspectives. Excessive price fluctuations would create risks for the market and weaken the effectiveness of ETS (Liu et al., 2020; Lyu et al., 2020; Song et al., 2019). Hence, it is necessary to forecast carbon prices accurately so as to enhance the enthusiasm of market participants. Evidence shows that carbon prices exhibit non-stationarity, nonlinearity, multi-scale and chaos (Fan et al., 2019; Huang et al., 2021; Manaf et al., 2016; Tian & Hao, 2020; Zhu et al., 2017; Zou & Zhang, 2020) and these make predictions difficult. Therefore, capturing features of carbon prices in order to maximise the forecasting precision is of great significance and it has become a challenging task for academic researchers.

There are various carbon price prediction approaches that have emerged over the last few years. These approaches/models can be broadly classified into three categories: traditional econometric methods (Byun & Cho, 2013; Chang et al., 2017), artificial intelligence (AI) algorithms (Fan et al., 2015) and hybrid models. However, individual prediction models of econometrics or AI have shown lack of prediction robustness due to inherent shortcomings (Zhu & Wei, 2013; Zhu et al., 2015). To overcome the disadvantages above, decomposition and ensemble hybrid models have been introduced in carbon price forecasting. These models...
have grasped various characteristics and influencing factors of carbon price in a comprehensive manner and, therefore, been able to generate more accuracy in prediction outcomes (Li et al., 2021; Niu et al., 2019; Sun et al., 2021; Zhu et al., 2021).

The core idea of hybrid models is to integrate the abilities of each algorithm, diversify the risk of single methods, and explore the complex internal information from various angles. Considering the high non-stationarity and nonlinearity features of carbon prices, decomposition by empirical mode decomposition (EMD) has been widely employed for maximising forecasting precisions (Zhu et al., 2018). As the optimization of EMD, variational mode decomposition (VMD) shows strong noise robustness and signals decomposition capabilities. Sun et al. (2016) proposed a hybrid model based on VMD and spiking neural networks (SNN) for ensuring precisions in the carbon price predictions. In an empirical study, Tian and Hao (2020) processed the original carbon price series into subsequences by VMD and illustrated an improved prediction performance of VMD. It is now widely used in the carbon price prediction models, implying its significant influence in carbon price decomposition (Sun & Huang, 2020; Wang et al., 2019; Zhu et al., 2019).

However, various subgroups produced by over-decomposition are observed to cause problems of poor economic implications and difficult calculations. Reconstruction method has therefore been introduced into carbon prices forecasting field and to avoid errors of subjective selection, sample entropy (SE) has been used to reconstitute subgroups into new components. This algorithm is applied to the reconstruction process of carbon price series after decomposition and it has produced more satisfactory prediction results in the comparative experiments (Hao et al., 2020; Sun & Duan, 2019).

In the prediction process, extreme learning machine (ELM) with fewer training parameters has demonstrated an outstanding generalization ability and a faster execution speed in comparison to the artificial neural network algorithms. Nonetheless, it has displayed some shortcomings in parameter settings (Ali et al., 2021), resulting in numerous applications of ELM, combined with intelligent optimization algorithm in predicting carbon prices (Sun & Duan, 2019; Sun & Zhang, 2018; Zhang et al., 2019; Zhou et al., 2019). In addition, Hao et al. (2020) used the improvements of ELM in the decomposition and ensemble model. The ELM optimization models not only remedied the limitations of dealing with nonlinearity and chaos problems, but also performed robustly in the field of related energy price forecasting tasks.

In light of the above backdrop, this paper presents a neoteric forecasting model with the framework of decomposition-reconstruction-prediction-ensemble, based on the non-stationary, non-linear, multi-scale and chaotic features of carbon prices. The main innovations and contributions are as follows: (1) Carbon price data of eight pilots in China are taken as samples. The characteristics of non-stationarity, non-linearity and chaos are verified in turn by Augmented Dickey–Fuller (ADF) test, Brock–Decher–Scheikman (BDS) test and the maximal Lyapunov exponent (MLYE). (2) The VMD-SE-PSO-ELM decomposition-reconstruction-prediction-ensemble hybrid model was established on the data features of carbon prices series in China’s ETS pilots. The proposed model combines the advantages of each algorithm and makes it possible to improve the prediction performance. (3) Five related models are used for empirical research to explain the effectiveness and robustness of the VMD-SE-PSO-ELM model. Furthermore, four diverse evaluation criteria are used to certify the forecasting performance of models, which contains root mean square error (RMSE), average absolute error (MAE), average absolute percentage error (MAPE) and goodness of fit ($R^2$). (4) As the GHG Market Sentiment Survey 2021 of the International Emissions Trading Association (IETA) and PwC (IETA-PwC, 2021) indicates, global carbon markets have been resilient to the effects of COVID-19 and a further rise in the carbon prices is expected. The new
prediction model can serve useful results in the post-COVID recovery process due to greater accuracy in carbon pricing. (5) Last but not the least, this paper provides the government departments in China a theoretical foundation to support further development and scope of necessary revisions of the pilots, and comprehensively build a national carbon market. Also, in the global context, the use of the prediction model will be supportive to the purpose-built formulations and likely introductions of Border Carbon Adjustment Mechanism (BCAM) in the EU and the US.

The remaining sections of this paper are organised as follows: the related methodologies and framework of the proposed model are presented in Sects. 2 and 3, respectively. Section 4 introduces datasets and data features. Section 5 describes the empirical experiments and Sect. 6 illustrates conclusions.

2 Methodology

Aiming at non-stationary, non-linear and chaotic features of carbon prices in China’s ETS pilots, in this study, we designed a decomposition-reconstruction-prediction-ensemble forecasting model. The proposed model includes variational mode decomposition, sample entropy, particle swarm optimization, and extreme learning machine (VMD-SE-PSO-ELM) algorithms. In this section below, the related technologies of VMD-SE-PSO-ELM model are explained in detail.

2.1 Variational mode decomposition (VMD)

VMD, improved by Dragomiretskiy and Zosso (2014), overcomes shortcomings of EMD and shows strong noise robustness and signals decomposition capabilities, which is adaptive for decomposing non-stationary and non-linear time series. As the evidence of this study proves that the non-stationarity and non-linearity exist in carbon prices in China’s ETS pilots, we use VMD for original series decomposition. VMD divides the time series into multiple intrinsic mode functions (IMF) with various frequencies, to enhance the volatility regularity and improve the prediction performance. IMFs with high frequency show relatively weak influence and short duration, while IMFs with low frequency exhibit greater impact and longer duration. The IMFs with low frequency generally last from two weeks to more than one year, and contributes more to carbon prices.

The core problem of it is to construct a variational mode and solve its problems. The main steps are as follows.

(1) Converting every component \( u_m(t) \) into an analysable signal through the Hilbert transform, and then transforming them into one-sided spectrums.

(2) Estimating center frequency of the modal function of every analysable signal by adjusting the exponential term \( e^{-j\omega_m t} \), and then adjusting the frequency spectrum of every component with the relevant baseband.

(3) Calculating the bandwidth of each component by Gaussian smoothing.

Hence, the construction problem of the variational model is expressed as follows.

\[
\begin{align*}
\min_{\{u_m\}|w_m} & \left\{ \sum_{m=1}^{M} \left\| \frac{\partial}{\partial t} \left( \frac{\delta(t)}{\pi t} + \frac{i}{\pi t} \right) \cdot u_m(t) e^{-j\omega_m t} \right\|^2 \right\} \\
\text{s.t.} \quad & \sum_{m=1}^{M} u_m = f
\end{align*}
\]
where $M$ represents the number of modes, $w_m$ is the center frequency in keeping with the modes, $\partial_t$ is the differential operator, $\delta_t$ represents Dirac distribution, $j$ is an imaginary number, and $f$ is the input signal.

To get the optimal solution of the constrained variational model, the penalty factor $\alpha$ and Lagrangian factor $\lambda$ are imported on the basis of Eq. (1), transforming constrained variational problems into unconstrained variational problems. The new equation is obtained as follows.

$$L(u_m, w_m, \lambda) = \alpha \sum_{m=1}^{M} \left\| \partial_t \left[ (\delta(t) + \frac{j}{\pi t}) \cdot u_m(t) \right] \right\|^2 e^{-jw_m t} + \left\| f(t) - \sum_{m=1}^{M} u_m(t) \right\|^2 + \left\langle \lambda(t), f(t) - \sum_{m=1}^{M} u_m(t) \right\rangle$$

(2)

Using the alternating direction method of multipliers (ADMM) to solve Eq. (2), the optimization problem can be concluded as follows.

$$\hat{u}_{m+1}^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i \neq m} \hat{u}_i(w) + \hat{\lambda}(w)}{1 + 2\alpha(w - w_m)^2}$$

(3)

$$u_{m+1}^{n+1} = \int_0^{\infty} w \left| \hat{u}_m(w) \right|^2 dw$$

(4)

2.2 Sample entropy (SE)

SE is designed by Richman and Moorman (2000) as an updated scheme of Approximate Entropy (AE). This method quantifies regularity and complexity of the input signal by calculating probability of generating a new pattern. Compared with AE, SE has the advantages of time series length independence and better consistency and stability. The greater the calculated SE values of IMFs, the higher the complexity, and the greater the probability that the IMFs tend to become a random sequence (Wang et al., 2021). In the feature extraction process, the adoption of SE increases the prediction accuracy and balances the computational efficiency. The calculation formula of SE is given by,

$$SE(m, r) = \lim_{N \to \infty} \left\{ - \ln \left( \frac{B^{m+1}(r)}{B^m(r)} \right) \right\}$$

(5)

where $N$ is the length of input signal, $m$ is the dimension, and $r$ is the similarity tolerance. When the length of input signal is limited, i.e., $N$ is a finite value, the formula changes into:

$$SE(m, r, N) = - \ln \left( \frac{B^{m+1}(r)}{B^m(r)} \right)$$

(6)

This paper sets parameter $m = 2$ and $r = 0.25 \cdot \text{std}$. IMFs with similar complexity are reconstructed into new subsequences for the follow-up prediction.

2.3 Particle swarm optimization (PSO)

PSO algorithm is a global random search algorithm that is simulated by bird predation behavior, and holds the features of fast convergence and integral optimization (Kennedy &
Eberhart, 1995). The particle is viewed as the optimal solution of optimization problem. The primary steps of PSO include:

1. Initializing the speeds and positions of particle swarms randomly where \( s \) represents the number of particles in the swarm, and the speed and position of the \( i \)-th particle in the \( D \)-dimensional space are expressed as \( v_i = (v_{i1}, v_{i2}, \ldots, v_{iD}) \) and \( x_i = (x_{i1}, x_{i2}, \ldots, x_{iD}) \), \( i = 1, 2, \ldots, D \), respectively; and then calculating the fitness value of \( x_i \) according to the preset objective function and judgement of its pros and cons.

2. Updating the personal optimal position \( P_{best} = (P_1, P_2, \ldots, P_D) \) and global optimal position \( G_{best} = (G_1, G_2, \ldots, G_D) \).

3. Using Eq. (7) and Eq. (8) to dynamically calculate the personal and global optimal position of each particle, and updating its speed and position at the same time.

\[
\begin{align*}
v_{ij}(t+1) &= w(t) \cdot v_i(t) + c_1 \cdot r_1 \cdot [P_j(t) - x_{ij}(t)] + c_2 \cdot r_2 \cdot [G_j(t) - x_{ij}(t)] \\
x_{ij}(t+1) &= x_{ij}(t) + v_{ij}(t+1)
\end{align*}
\]

(7) (8)

where \( t \) is the current number of iterations, \( w \) is the inertia weight to improve the search range, \( c_1 \) and \( c_2 \) are acceleration factors, \( r_1 \) and \( r_2 \) are random numbers uniformly among the interval \([0, 1]\). With each iteration update, each particle approaches the optimal position. When the preset maximum number of iterations is attained, PSO algorithm terminates and produces the optimal solution.

2.4 Extreme learning machine (ELM)

ELM, proposed by Huang et al. (2006), has single hidden layer, fewer training parameters, better generalization ability and faster convergence speed. The gradient descent method overcomes the problems of slow training speed and local extreme values. It is suitable for almost all non-linear and complex activation functions. The core problems of ELM are:

1. Setting the connection weight of the input and the hidden layer and the threshold of the hidden layer neurons randomly, where the adjusted parameters are only the number of hidden nodes.

2. Using Moore–Penrose inverse to calculate the matrix of the hidden layer, and transforming the training process into solving the least square problem.

The output of ELM is as follows.

\[
f(x) = \sum_{i=1}^{N} \beta_i G(w_i, b_i, x) = \beta \cdot h(x)
\]

(9)

where \( \beta_i \) is the output weight of the \( i \)-th hidden layer node and the output layer neuron, \( w_i \) is the input weight of the input layer neuron and the \( i \)-th hidden layer node, \( b_i \) is the threshold of the \( i \)-th hidden layer node, \( h(x) = [G(w_1, b_1, x_1), \ldots, G(w_N, b_N, x_N)] \) is the output vector for the hidden layer relative to the input \( x \). The output weight is obtained by solving the least squares solution of Eq. (10):

\[
\min \sum_{i=1}^{N} \| \beta \cdot h(x_i) - y_i \|
\]

(10)
The least squares solution of Eq. (10) is:

$$\beta = H^+ Y$$  \hspace{1cm} (11)$$

where $H^+$ is the Moore–Penrose inverse of the hidden layer output matrix $H$.

However, non-optimal weights and thresholds may reduce the prediction performance and cause poor stability of ELM. As the chaotic feature of carbon markets requires appropriate forecasting tools, PSO which has the ability to adjust the input layer weight and hidden layer deviation, is adopted to solve the defects of ELM, while giving full play to the advantages of ELM and PSO.

3 Framework of the proposed model

With reference to the flowchart of the decomposition-reconstruction-prediction-ensemble model (Fig. 1), the main elements of the VMD-SE-PSO-ELM model are explained as follows:

(1) Decomposition: variational mode decomposition (VMD) resolves original carbon prices into many intrinsic mode functions (IMFs), which effectively reduce the complexity and irregularity.

(2) Reconstruction: the complexity of IMFs is calculated by sample entropy (SE) algorithm and, following this, the IMFs with similar SE values are integrated in a new subsequence. This step is introduced to enhance precision and speed of prediction.

(3) Prediction: extreme learning machine (ELM) optimized by particle swarm optimization (PSO) predicts all subsequences, which enable obtaining satisfactory results with complex features.
This study designs a decomposition and ensemble hybrid prediction model driven by characteristics that influence carbon prices, and proves its accuracy and robustness through real carbon price data from eight pilots in China.

4 Data

4.1 Data collection

Starting from June 18, 2013, eight provinces and municipalities of Shenzhen, Shanghai, Beijing, Guangdong, Tianjin, Hubei, Chongqing, and Fujian in China have launched the pilot projects of carbon emission trading markets one after another. In this paper, we chose the daily transaction prices of eight pilots for empirical analysis. The carbon price data from the opening date of each pilot until December 31, 2020 were downloaded from the Wind Database (http://www.wind.com.cn/). According to extant literature (e.g., Huang et al., 2021; Sun & Huang, 2020; Sun & Zhang, 2018), the training set and testing set are divided on the basis of an approximate proportion 4:1. Table 1 summarizes the basic situation of each dataset.

4.2 Data characteristics

Figure 2 displays carbon prices of the eight pilots in China, which show complex and irregular trends in general. Table 2 lists statistical descriptions of carbon prices in eight pilots, and displays discrete distribution and volatility of the prices. The Jarque–Bera values of eight pilots carbon prices are far greater than 0, indicating that none of them follows normal distributions, and this also confirmed by the P-values. Except for SHEA, the skewness values of other carbon prices are all greater than 0, showing right skewed distribution, whereas SHEA showing left skewed distribution. The kurtosis values of SZA, SHEA, GDEA, TJEA, HBEA are greater than 3, showing sharp peak characteristics. In addition, the P-values of Augmented Dickey–Fuller (ADF) test in the pilots of BEA, SHEA, TJEA, HBEA, CQEA and FJEA are greater than 0.05, indicating non-stationary features of the series.

(4) Ensemble: the results of each subsequence are linearly added to obtain the final carbon price forecasting results.
To verify whether carbon price series of China’s ETS pilots have non-linear feature, Brock–Decher–Scheikman (BDS) test is employed in this paper. Since the parameter \( \epsilon \) and the embedded dimension \( m \) directly affect the results of BDS test, this paper has set \( \epsilon = 1.5\sigma \), \( m = 2, 3, 4 \ldots 6 \), according to the carbon market actual condition and related literature. As shown in Table 3, the P-values of BDS test are all 0.0000 under different embedding dimensions. It indicates that carbon prices of eight pilots have non-linear characteristics at the 1% significance level.

To examine the chaotic property of carbon prices in eight pilots, firstly, C–C method is applied to reconstruct the phase space of each original carbon price series and hence obtain their \( m \) (the embedded dimension), \( \tau \) (the delay time) and \( \tau_w \) (the delay time windows). Following this, the extensively used ‘Wolf’ method is selected to calculate the maximal Lyapunov exponent (MLYE) of carbon prices in each pilot. As shown in Table 4, the MLYE values of eight pilots are all greater than 0, confirming the existence of chaos in carbon price time series.

![Fig. 2 The time series of carbon prices in China’s ETS pilots](image-url)
**Table 3** The results of Brock–Decher–Scheikman (BDS) test in China’s ETS pilots

| ETS pilots | m (The embedded dimension) |  |  |  |  |  |  |  |
|------------|---------------------------|---|---|---|---|---|---|---|
|            | m = 2 BDS P                | m = 3 BDS P | m = 4 BDS P | m = 5 BDS P | m = 6 BDS P |  |
|            | BDS P             | BDS P | BDS P | BDS P | BDS P |  |
| SZA        | 0.1627 0.0000     | 0.2843 0.0000 | 0.3712 0.0000 | 0.4320 0.0000 | 0.4735 0.0000 |  |
| BEA        | 0.1809 0.0000     | 0.3066 0.0000 | 0.3912 0.0000 | 0.4485 0.0000 | 0.4863 0.0000 |  |
| SHEA       | 0.1619 0.0000     | 0.2866 0.0000 | 0.3808 0.0000 | 0.4512 0.0000 | 0.5035 0.0000 |  |
| GDEA       | 0.1094 0.0000     | 0.2047 0.0000 | 0.2873 0.0000 | 0.3589 0.0000 | 0.4208 0.0000 |  |
| TJEA       | 0.1720 0.0000     | 0.2959 0.0000 | 0.3827 0.0000 | 0.4431 0.0000 | 0.4833 0.0000 |  |
| HBEA       | 0.1961 0.0000     | 0.3325 0.0000 | 0.4258 0.0000 | 0.4890 0.0000 | 0.5314 0.0000 |  |
| CQEA       | 0.1759 0.0000     | 0.2985 0.0000 | 0.3832 0.0000 | 0.4403 0.0000 | 0.4771 0.0000 |  |
| FJEA       | 0.1816 0.0000     | 0.3056 0.0000 | 0.3874 0.0000 | 0.4398 0.0000 | 0.4720 0.0000 |  |
To determine prediction performance of the proposed model with the framework of decomposition and ensemble, carbon price series of eight pilots in China are utilized for simulation prediction and analysis. The average value of the 5 runs prediction results is taken as the last result of the prediction model, so as to reduce the randomness.

### 5.1 Performance evaluation criteria

In order to directly judge the forecasting performance of the proposed hybrid model and benchmark models, this study introduces four widely used error evaluation indicators and tests the prediction precision, including root mean square error (RMSE), average absolute error (MAE), average absolute percentage error (MAPE), and goodness of fit (R²). The smaller the values of the RMSE, MAE and MAPE, the more accurate the prediction (Yang et al., 2020). The closer the R² value is to 1, the better the model fit condition (Sun & Xu, 2021). The definition formulae of the error evaluation indicators are expressed as follows.

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}
\]

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |(y_i - \hat{y}_i)|
\]

\[
MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%
\]

\[
R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y}_i)^2}
\]

where \(y_i\) represents i-th actual value of time series, \(\hat{y}_i\) represents i-th prediction value, \(\bar{y}_i\) represents the average of the actual data.

### 5.2 Carbon price forecasting in the Hubei market

According to the statistical descriptions and characteristics verification, carbon prices in Hubei market (HBEA) have significant non-stationarity, non-linearity, multi-scale and chaotic
features. The Hubei carbon emissions trading market has received extensive attention, due to its leading position in terms of trading volume, turnover and liquidity among eight pilots. Therefore, HBEA is employed as a typical sample for a detailed analysis.

5.2.1 Decomposition and reconstruction

Due to the complexity of carbon price features, VMD algorithm could process original carbon price series into multiple IMFs with relatively stationarity and different frequencies, resulting in improved forecasting precision. However, there is no adaptive processing procedure or clear standard to confirm the best number of modes so far (Bisoi et al., 2019a, 2019b; Jiang et al., 2019). Existing studies have determined the parameter by selecting the number of IMFs generated from EMD or EEMD algorithm (Bisoi et al., 2019a, 2019b; Zhu et al., 2019). Tian and Hao (2020) proposed that the parameter settings of EEMD may influence decomposition results, whereas EMD demands no specific parameter to be set. Therefore, the number of modes produced by VMD is set to 10 in line with the decomposition of EMD and the results, as shown in Fig. 3.

To improve the computational efficiency and avoid the information loss caused by over-decomposition, IMFs generated by VMD are reconstructed by SE algorithm according to their complexity. This provides a more ideal method to make the prediction results more
scientific and theoretically sound. The larger SE values, the higher the complexity of IMFs. IMFs with approximate SE values will be merged to obtain new subseries. Table 5 shows the SE values of IMFs and the reconstruction results. Hence, four reconstructed subsequences are obtained, as shown in Fig. 4 below.

### 5.2.2 Forecasting results

Aimed at comparing and verifying the advantages of the designed model in prediction, ELM, PSO-ELM, and three decomposition and ensemble hybrid models of VMD-SE-ELM, VMD-PSO-ELM, EMD-SE-PSO-ELM are selected for benchmark models. Table 6 highlights four evaluation criteria indicators of RMSE, MAE, MAPE and $R^2$ of each model. Figure 5 displays the comparative results between actual data and forecasting results of each model in Hubei market.

Some important findings are described as follows:

1. As delineated in Fig. 5, the trends of actual data and prediction model results look consistent. Nonetheless, there is a big difference in the evaluation criteria values of forecasting error. The proposed model (VMD-SE-PSO-ELM) has the best predictive performance on RMSE, MAE, MAPE and $R^2$ of all tested prediction models, and the values are 0.2360, 0.1755, 0.0062, and 0.9929, respectively. Therefore, the designed model could be used as an excellent approach for analysis and prediction in Hubei market.

2. The purpose of the comparison between ELM and PSO-ELM is to illustrate the superiority of PSO-ELM. The four values of PSO-ELM are significantly greater than these of ELM, and the four values of VMD-SE-PSO-ELM are significantly greater than these of VMD-SE-ELM. These results endorse that PSO-ELM has an advantage in forecasting carbon prices.

3. When comparing ELM, PSO-ELM with the models combined with EMD or VMD decomposition algorithm, the decomposition and ensemble hybrid models have shown remarkable enhancements in the forecasting accuracy. For example, the four values of PSO-ELM are 0.3437, 0.2461, 0.0085, and 0.9849, respectively whereas the four values of VMD-PSO-ELM are 0.2696, 0.2122, 0.0075, and 0.9907.
Table 6 Prediction evaluation indicators in Hubei market

| Models            | RMSE  | MAE   | MAPE  | R²    |
|-------------------|-------|-------|-------|-------|
| ELM               | 0.6787| 0.5062| 0.0179| 0.9410|
| PSO-ELM           | 0.3437| 0.2461| 0.0085| 0.9849|
| VMD-SE-ELM        | 0.5123| 0.3607| 0.0127| 0.9664|
| VMD-PSO-ELM       | 0.2696| 0.2122| 0.0075| 0.9907|
| EMD-SE-PSO-ELM    | 0.2428| 0.2133| 0.0075| 0.9925|
| Proposed model    | 0.2360| 0.1755| 0.0062| 0.9929|

(4) When comparing EMD-SE-PSO-ELM with VMD-SE-PSO-ELM, the prediction evaluation criteria values of the latter look definitely better. It depicts that VMD has better decomposition ability than EMD.

(5) When comparing VMD-PSO-ELM without reconstruction algorithm with VMD-SE-PSO-ELM, the latter optimizes 12.46%, 17.30%, 17.33%, 0.22% in RMSE, MAE, MAPE and R², respectively. It is hence evident that SE algorithm reduces the difficulty in calculations and enhances the prediction accuracy of the models.
5.3 Carbon price forecasting in the Shanghai market

The carbon price series in Shanghai market (SHEA) also displayed non-stationary, non-linear, and chaotic characteristics. Moreover, the Shanghai carbon market hosts intense market activities and it is the second pilot in operation in China. Also, the upcoming national carbon emissions trading system will be established in Shanghai. Taking the actual data of Shanghai pilot as a typical case, Figs. 6 and 7 exhibit the decomposition and reconstruction results, respectively.

Table 7 illustrates four evaluation indicators of the prediction results, and Fig. 8 displays the prediction results of each model in Shanghai market. We can get similar prediction results as Hubei market.

5.4 Forecasting results of other pilots

To further verify the stability of the proposed model, Table 8 summarizes carbon price forecasting results of six other pilots. Overall, according to the forecasting results of eight pilots, the following conclusions can be drawn.

(1) The VMD-SE-PSO-ELM, i.e., developed in this study is the optimal model in the comparative empirical tests of eight pilots in China. The proposed model performs outstandingly in the eight different pilots, according to the four evaluation criteria. The results of empirical analysis also demonstrate the robustness and validity of the proposed model for forecasting carbon prices.

(2) The decomposition algorithms are able to improve the forecasting precision of models. This finding is consistent with Huang and He (2020), who insist that the decomposition algorithms effectively enhance the forecasting accuracy due to the poor property of the non-decomposed forecasting model. In the results of eight pilots, the models combined with VMD or EMD algorithms perform better than the single forecasting models. It is possible that the decomposition and ensemble hybrid models could process the carbon prices with high non-stationary, non-linearity and chaos into regular subsequences, which makes the prediction more accurate and easier.
Fig. 6 VMD decomposition results in Shanghai market

Fig. 7 VMD-SE results in Shanghai market
Table 7 Prediction evaluation indicators in Shanghai market

| Models               | RMSE   | MAE    | MAPE   | R²     |
|----------------------|--------|--------|--------|--------|
| ELM                  | 0.6611 | 0.4971 | 0.0122 | 0.9587 |
| PSO-ELM              | 0.4953 | 0.3629 | 0.0090 | 0.9768 |
| VMD-SE-ELM           | 0.4641 | 0.3557 | 0.0087 | 0.9797 |
| VMD-PSO-ELM          | 0.2871 | 0.1874 | 0.0047 | 0.9922 |
| EMD-SE-PSO-ELM       | 0.2878 | 0.2021 | 0.0049 | 0.9926 |
| Proposed model       | 0.1418 | 0.1096 | 0.0027 | 0.9981 |

Fig. 8 Prediction results of numerous models in the Shanghai market

(3) It is important to employ SE reconstruction algorithm to reduce calculation complexity and improve prediction performance of the models. This point has been verified in the empirical experiments of China’s ETS pilots. Consistent with the conclusions of Xu et al. (2019), it can be implied that China’s ETS pilots are featured with longer duration relatively. It may be related to annual cycle design, consists of compliance, allowance allocation, ‘measurement, reporting and verification’ (MRV) process, etc. Therefore, it is necessary to set up a data preprocessing process in carbon price prediction models.

(4) The PSO algorithm is conducive to optimize the ELM. According to the empirical results, the models optimized by PSO outperform the models without PSO optimization. For examples, PSO-ELM and VMD-SE-PSO-ELM could get more satisfactory forecasting results than ELM and VMD-SE-ELM. The results of this paper prove that PSO is also applicable in the field of carbon price forecasting in China’s ETS pilots.

6 Conclusions

This study established a carbon price prediction model with the framework of decomposition-reconstruction-prediction-ensemble, emerging from the idea of data feature-driven modeling. The key findings and contributions of this research include two aspects as follows. On the
| ETS pilots | Models       | RMSE  | MAE   | MAPE  | R²   |
|-----------|--------------|-------|-------|-------|------|
| SZA       | ELM          | 2.2720| 1.3600| 0.1042| 0.9510|
|           | PSO-ELM      | 1.7531| 1.2985| 0.1260| 0.9708|
|           | VMD-SE-ELM   | 0.7279| 0.4534| 0.0363| 0.9950|
|           | VMD-PSO-ELM  | 1.2208| 1.1917| 0.1032| 0.9858|
|           | EMD-SE-PSO-ELM | 0.5487| 0.3115| 0.0286| 0.9971|
|           | Proposed Model | 0.3138| 0.2407| 0.0198| 0.9991|
| BEA       | ELM          | 1.5774| 1.0886| 0.0134| 0.9728|
|           | PSO-ELM      | 1.1051| 0.8953| 0.0111| 0.9867|
|           | VMD-SE-ELM   | 0.5167| 0.3557| 0.0044| 0.9971|
|           | VMD-PSO-ELM  | 0.4202| 0.3725| 0.0046| 0.9981|
|           | EMD-SE-PSO-ELM | 0.4287| 0.3073| 0.0038| 0.9980|
|           | Proposed Model | 0.4284| 0.2588| 0.0033| 0.9982|
| GDEA      | ELM          | 0.3142| 0.2417| 0.0086| 0.8802|
|           | PSO-ELM      | 0.1540| 0.1141| 0.0041| 0.9712|
|           | VMD-SE-ELM   | 0.1465| 0.1288| 0.0046| 0.9739|
|           | VMD-PSO-ELM  | 0.0366| 0.0271| 0.0010| 0.9984|
|           | EMD-SE-PSO-ELM | 0.0276| 0.0183| 0.0006| 0.9991|
|           | Proposed Model | 0.0257| 0.0180| 0.0006| 0.9992|
| TJEA      | ELM          | 0.3192| 0.2276| 0.0094| 0.9826|
|           | PSO-ELM      | 0.2247| 0.1771| 0.0074| 0.9914|
|           | VMD-SE-ELM   | 0.1904| 0.1411| 0.0057| 0.9938|
|           | VMD-PSO-ELM  | 0.1165| 0.0969| 0.0040| 0.9977|
|           | EMD-SE-PSO-ELM | 0.1300| 0.1014| 0.0042| 0.9971|
|           | Proposed Model | 0.1098| 0.0891| 0.0037| 0.9979|
| CQEA      | ELM          | 0.8790| 0.6707| 0.0248| 0.9907|
|           | PSO-ELM      | 0.7957| 0.6373| 0.0249| 0.9924|
|           | VMD-SE-ELM   | 0.2867| 0.2284| 0.0097| 0.9990|
|           | VMD-PSO-ELM  | 0.2700| 0.2382| 0.0107| 0.9991|
|           | EMD-SE-PSO-ELM | 0.2666| 0.1817| 0.0068| 0.9991|
|           | Proposed Model | 0.1973| 0.1467| 0.0060| 0.9995|
| FJEA      | ELM          | 0.7422| 0.6051| 0.0337| 0.9741|
|           | PSO-ELM      | 0.3735| 0.3160| 0.0208| 0.9934|
|           | VMD-SE-ELM   | 0.2654| 0.2151| 0.0147| 0.9967|
|           | VMD-PSO-ELM  | 0.0869| 0.0696| 0.0048| 0.9995|
|           | EMD-SE-PSO-ELM | 0.1964| 0.1483| 0.0100| 0.9982|
|           | Proposed Model | 0.0863| 0.0696| 0.0048| 0.9996|
one hand, the complex carbon price characteristics of eight pilots in China could be effectively verified and analyzed. On the other hand, the proposed data characteristic-driven model (i.e., VMD-SE-PSO-ELM) presents satisfactory prediction performance. Accordingly, four primary steps are included: First, VMD approach is applied to decompose original carbon price series, which is aimed at reducing the complex characteristics; Second, SE algorithm is applied to reconstruct IMFs on the basis of their SE values; Third, ELM optimized by PSO algorithm is utilized to forecasting the subsequences, which could outperform in forecasting chaos series; Finally, the outputs are acquired by summing the predicting values of subgroups. In order to testify the effective performance of the designed model, four evaluation criteria indicators involving RMSE, MAE, MAPE, $R^2$ are applied. Furthermore, carbon price prediction of eight pilots in China are illustrated to prove the robustness and validity of the VMD-SE-PSO-ELM model.

In contrary to the expectation that the ETS prices would remain suppressed by COVID-19 over the next few years, carbon prices showed resilience and global ETS displayed an upward tendency. In light of this backdrop, the empirical results of this paper would provide a novel approach to optimize the accuracy of carbon price forecasting, which will not only benefit the emission control companies and investment institutions, but also the relevant government departments in the post-pandemic recovery process towards reaching a stable state of carbon markets. For the emission control companies and investment institutions, establishing a sound risk management mechanism and timely analyzing the fluctuation characteristics of carbon prices will be possible towards avoiding risks that may potentially arise in the future. The use of the new model will result in securing sustainability of over two-thirds of global GDP that are generated by 20% of the world’s largest public companies with “net zero” commitments. For the relevant government departments, developing some relevant policies to control the trading volume and optimizing the quota distribution will be possible to stabilize carbon prices and improve market liquidity.

The findings of this paper have managerial implications in the national economies and on the global institutional levels in the wake of the global pandemic as well as the changing political and economic scenarios in the recent times. Henceforth, the availability of a more definite carbon pricing prediction model will enable formulations of effective policies and their applications in various schemes such as: (a) Biden government’s likely launch of carbon pricing systems to secure the Natural Climate Solutions (NCS) including green aviation, clean electricity and voluntary markets; (b) Chinese government’s plan to kick off trading in nationwide ETS including additional sectors such as petrochemicals, chemical, paper and steel, and associated challenges in establishing monitoring, reporting and verification (MRV); (c) Government initiatives in Japan and Taiwan to implement carbon pricing mechanism, followed by similar approaches in Colombia, Indonesia, Mexico, Thailand and Vietnam between 2022 and 2026; (d) the EU’s likely introduction of Border Carbon Adjustment Mechanism (BCAM), as part of their trajectory towards a net zero region by 2050; (e) working out a feasible pricing of the carbon credits under Article 6 of the Paris Agreement.

This paper takes equal consideration of the fact that the proposed data-characteristic-driven model could also be applied with other time series that show non-stationary, nonlinear and chaotic features. Likewise, more influence factors could be considered in the hybrid prediction models, such as energy prices, relevant policy, and stakeholder behaviors. These factors create scope of further research in cognate areas.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 71704098, 71971039, 72003110), the Natural Science Foundation of Shandong Province (No. ZR2016GQ03).
Authors are indebted to Taimur Sharif, Business School, Newman University, Birmingham B32 3NT, United Kingdom for his contributions in revising, editing and proofreading the paper.

References

Adekoya, O. B. (2021). Predicting carbon allowance prices with energy prices: A new approach. *Journal of Cleaner Production, 282*, 124519.

Ali, M., Prasad, R., Xiang, Y., et al. (2021). Advanced extreme learning machines vs. deep learning models for peak wave energy period forecasting: A case study in Queensland, Australia. *Renewable Energy, 177*, 1031–1044.

Balciar, M., Demirer, R., Hammoudeh, S., et al. (2016). Risk spillovers across the energy and carbon markets and hedging strategies for carbon risk. *Energy Economics, 54*, 159–172.

Bauer, N., Bertram, C., Schultes, A., et al. (2020). Quantification of an efficiency–sovereignty trade-off in climate policy. *Nature, 588*, 261–266.

Bisoi, R., Dash, P. K., & Mishra, S. P. (2019a). Modes decomposition method in fusion with robust random vector functional link network for crude oil price forecasting. *Applied Soft Computing, 80*, 475–493.

Bisoi, R., Dash, P. K., & Parida, A. K. (2019b). Hybrid variational mode decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. *Applied Soft Computing, 74*, 652–678.

Byun, S. J., & Cho, H. (2013). Forecasting carbon futures volatility using GARCH models with energy volatilities. *Energy Economics, 40*, 207–221.

Chang, K., Pei, P., Zhang, C., et al. (2017). Exploring the price dynamics of CO₂ emissions trading scheme pilots. *Energy Economics, 67*, 213–223.

Chen, Z. F., Zhang, X., & Chen, F. L. (2021). Do carbon emission trading schemes stimulate green innovation in enterprises? Evidence from China. *Technological Forecasting and Social Change, 168*, 120744.

Cui, L., Fan, Y., Zhu, L., et al. (2014). How will the emissions trading scheme save cost for achieving China’s 2020 carbon intensity reduction target? *Applied Energy, 2014*(136), 1043–1052.

Daskalakis, G. (2013). On the efficiency of the European carbon market: New evidence from Phase II. *Energy Policy, 54*, 369–375.

Dragomiretskiy, K., & Zosso, D. (2014). Variational mode decomposition. *IEEE Transactions on Signal Processing, 62*, 531–544.

Duan, H., Mo, J., Fan, Y., et al. (2018). Achieving China’s energy and climate policy targets in 2030 under multiple uncertainties. *Energy Economics, 70*, 45–60.

Duc, Khuong, N., Toan, L. D. H., & Nasir, M. A. (2021). Carbon emissions determinants and forecasting: Evidence from G6 countries. *Journal of Environmental Management, 285*, 111988.

Fan, X., Li, S., & Tian, L. (2015). Chaotic characteristic identification for carbon price and a multi-layer perceptron network prediction model. *Expert Systems with Applications, 42*(8), 3945–3952.

Fan, X., Lv, X., Yin, J., et al. (2019). Multifractality and market efficiency of carbon emission trading market: Analysis using the multifractal detrended fluctuation technique. *Applied Energy, 251*, 113333.

Hao, Y., Tian, C., & Wu, C. (2020). Modelling of carbon price in two real carbon trading markets. *Journal of Cleaner Production, 244*, 118556.

Huang, G., Zhu, Q., & Siew, C. (2006). Extreme learning machine: Theory and applications. *Neurocomputing, 70*(1–3), 489–501.

Huang, Y., Dai, X., Wang, Q., et al. (2021). A hybrid model for carbon price forecasting using GARCH and long short-term memory network. *Applied Energy, 285*, 116485.

Huang, Y., & He, Z. (2020). Carbon price forecasting with optimization prediction method based on unstructured combination. *Science of the Total Environment, 725*, 138350.

IETA-PwC. (2021). GHG market sentiment survey 2021, 16th edition. International Emissions Trading Association (IETA), in association with PwC UK.

Jiang, Y., Huang, G., Yang, Q., et al. (2019). A novel probabilistic wind speed prediction approach using real time refined variational model decomposition and conditional kernel density estimation. *Energy Conversion and Management, 185*, 758–773.

Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of the IEEE International Conference on Neural Networks, 4*, 1942–1948.

Koch, N., Fuss, S., Grosjean, G., et al. (2014). Causes of the EU ETS price drop: Recession, CDM, renewable policies or a bit of everything?—New evidence. *Energy Policy, 73*, 676–685.

Li, H., Jin, F., Sun, S., et al. (2021). A new secondary decomposition ensemble learning approach for carbon price forecasting. *Knowledge-Based Systems, 214*, 106686.
Liu, J., Huang, Y., & Chang, C. (2020). Leverage analysis of carbon market price fluctuation in China. *Journal of Cleaner Production*, 245, 118557.

Liu, Z., & Sun, H. (2021). Assessing the impact of emissions trading scheme on low-carbon technological innovation: Evidence from China. *Environmental Impact Assessment Review*, 89, 106589.

Lyu, J., Cao, M., Wu, K., et al. (2020). Price volatility in the carbon market in China. *Journal of Cleaner Production*, 255, 120171.

Ma, X., Jiang, P., & Jiang, Q. C. (2020). Research and application of association rule algorithm and an optimized grey model in carbon emissions forecasting. *Technological Forecasting and Social Change*, 158, 120159.

Macdonald, F. A., Swanson-Hysell, N. L., Park, Y., et al. (2019). Arc-continent collisions in the tropics set earth’s climate state. *Science*, 364(6436), 181.

Manaf, N. A., Qadir, A., & Abbas, A. (2016). Temporal multiscalar decision support framework for flexible operation of carbon capture plants targeting low-carbon management of power plant emissions. *Applied Energy*, 169, 912–926.

Niu, W., Feng, Z., Zeng, M., et al. (2019). Forecasting reservoir monthly runoff via ensemble empirical mode decomposition and extreme learning machine optimized by an improved gravitational search algorithm. *Applied Soft Computing*, 82, 105589.

Qi, S., Cheng, S., & Cui, J. (2021). Environmental and economic effects of China’s carbon market pilots: Empirical evidence based on a DID model. *Journal of Cleaner Production*, 279, 123720.

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6), H2039–H2049.

Song, Y., Liu, T., Ye, B., et al. (2019). Improving the liquidity of China’s carbon market: Insight from the effect of carbon price transmission under the policy release. *Journal of Cleaner Production*, 239, 118049.

Sun, G., Chen, T., Wei, Z., et al. (2016). A Carbon price forecasting model based on variational mode decomposition and spiking neural networks. *Energies*, 9(1).

Sun, S., Jin, F., Li, H., et al. (2021). A new hybrid optimization ensemble learning approach for carbon price forecasting. *Applied Mathematical Modelling*, 97, 182–205.

Sun, W., & Duan, M. (2019). Analysis and forecasting of the carbon price in China’s regional carbon markets based on fast ensemble empirical mode decomposition, phase space reconstruction, and an improved extreme learning machine. *Energies*, 12, 2772.

Sun, W., & Huang, C. (2020). A carbon price prediction model based on secondary decomposition algorithm and optimized back propagation neural network. *Journal of Cleaner Production*, 243, 118671.

Sun, W., & Xu, C. (2021). Carbon price prediction based on modified wavelet least square support vector machine. *Science of the Total Environment*, 754, 142052.

Tian, C., & Hao, Y. (2020). Point and interval forecasting for carbon price based on an improved analysis-forecast system. *Applied Mathematical Modelling*, 79, 126–144.

Tierney, J. E., Poulsen, C. J., Montañez, I. P., et al. (2020). Past climates inform our future. *Science*, 370(6517), 639–738.

Trisos, C. H., Merow, C., & Pigot, A. L. (2020). The projected timing of abrupt ecological disruption from climate change. *Nature*, 580, 496–501.

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

Wang, S., E, J. W., & Li, S. G. (2019). A novel hybrid carbon price forecasting model based on radial basis function neural network. *Acta Physica Polonica A*, 135(3), 368–374.

Westerhold, T., Marwan, N., Drury, A. J., et al. (2020). An astronomically dated record of Earth’s climate and its predictability over the last 66 million years. *Science*, 369(6509), 1383–1387.

Xiao, J., Li, G., Zhu, B., et al. (2021). Evaluating the impact of carbon emissions trading scheme on Chinese firms’ total factor productivity. *Journal of Cleaner Production*, 306, 127104.

Xu, J., Tan, X., He, G., et al. (2019). Disentangling the drivers of carbon prices in China’s ETS pilots—an EEMD approach. *Technological Forecasting and Social Change*, 139, 1–9.

Yang, S., Chen, D., Li, S., et al. (2020). Carbon price forecasting based on modified ensemble empirical mode decomposition and long short-term memory optimized by improved whale optimization algorithm. *Science of the Total Environment*, 716, 137117.
Zhang, X., Zhang, C., & Wei, Z. (2019). Carbon Price Forecasting based on multi-resolution singular value decomposition and extreme learning machine optimized by the moth-flame optimization algorithm considering energy and economic factors. *Energies, 12*, 428322.

Zhang, Y., & Wei, Y. (2010). An overview of current research on EU ETS: Evidence from its operating mechanism and economic effect. *Applied Energy, 87*(6), 1804–1814.

Zhou, J., Huo, X., Xu, X., et al. (2019). Forecasting the carbon price using extreme-point symmetric mode decomposition and extreme learning machine optimized by the grey wolf optimizer algorithm. *Energies, 12*, 9505.

Zhu, B., Han, D., Wang, P., et al. (2017). Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. *Applied Energy, 191*, 521–530.

Zhu, B., Wang, P., Chevallier, J., et al. (2015). Carbon price analysis using empirical mode decomposition. *Computational Economics, 45*(2), 195–206.

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

Zhu, B., Ye, S., Wang, P., et al. (2018). A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting. *Energy Economics, 70*, 143–157.

Zhu, B., Ye, S., Wang, P., et al. (2021). Forecasting carbon price using a multi-objective least squares support vector machine with mixture kernels. *Journal of Forecasting*.

Zhu, J., Wu, P., Chen, H., et al. (2019). Carbon price forecasting with variational mode decomposition and optimal combined model. *Physica A-Statistical Mechanics and Its Applications, 519*, 140–158.

Zou, S., & Zhang, T. (2020). Cross-correlation analysis between energy and carbon markets in China based on multifractal theory. *International Journal of Low-Carbon Technologies, 15*(3), 389–397.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.