



Carbon trading and COVID-19: a hybrid machine learning approach for international carbon price forecasting

Xingmin Zhang¹ · Zhiyong Li^{1,2}  · Yiming Zhao¹ · Lan Wang¹

Accepted: 29 March 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Accurate carbon price forecasting can better allocate carbon emissions and thus ensure a balance between economic development and potential climate impacts. In this paper, we propose a new two-stage framework based on processes of decomposition and re-estimation to forecast prices across international carbon markets. We focus on the Emissions Trading System (ETS) in the EU, as well as the five main pilot schemes in China, spanning the period from May 2014 to January 2022. In this way, the raw carbon prices are first separated into multiple sub-factors and then reconstructed into factors of ‘trend’ and ‘period’ with the use of Singular Spectrum Analysis (SSA). Once the subsequences have been thus decomposed, we further apply six machine learning and deep learning methods, allowing the data to be assembled and thus facilitating the prediction of the final carbon price values. We find that from amongst these machine learning models, the Support vector regression (SSA-SVR) and Least squares support vector regression (SSA-LSSVR) stand out in terms of performance for the prediction of carbon prices in both the European ETS and equivalent models in China. Another interesting finding to come out of our experiments is that the sophisticated algorithms are far from being the best performing models in the prediction of carbon prices. Even after accounting for the impacts of the COVID-19 pandemic and other macro-economic variables, as well as the prices of other energy sources, our framework still works effectively.

Keywords Carbon price prediction · Carbon emission · COVID-19 · Singular spectrum analysis · Ensemble learning method

1 Introduction

Since the signing of the 2015 Paris Agreement, international attention has shifted to focus on the potentially devastating impact of global climate warming, with the agreement being

✉ Zhiyong Li
liz@swufe.edu.cn

¹ School of Finance and Fintech Innovation Center, Southwestern University of Finance and Economics, Chengdu, China

² Collaborative Innovation Center of Financial Security, Southwestern University of Finance and Economics, Chengdu, China

focused on mitigating the consequences of climate change by reducing greenhouse gas emissions. Greenhouse gases are widely regarded as the major culprit of global warming, which poses a great threat not just to the survival of human society, but also its continued economic development. The establishment of a sound carbon allowance trading system is commonly viewed as an effective and low-cost instrument which would reduce carbon emissions and control the excessive rising of global temperatures (Bauer et al., 2020). To regulate these carbon emissions, companies which generate more greenhouse gases as a result of their industrial activities would need to buy emission quotas in a public trading system, while those with a higher carbon emission allowance who do not exceed their quota would also be able to sell their emission allowance to other companies. Carbon would therefore become a fundamental trading asset with its prices fluctuating from time to time, as in the current global stock market. An effective pricing model should therefore be the basis of such a trading system. An accurate estimation of carbon prices would thus provide an invaluable insight into the carbon markets, which would be helpful for companies to manage their ‘climate costs’. With this in mind, this study proposes a new ensemble framework for international carbon markets, based on data from the European Union (EU) and China, which are currently the principal global carbon trading markets.

As a crucial indicator of market supply and demand, the price of carbon is affected by a lot of drivers including national or international allowances, restrictions in place, as well as local/global energy prices and other macroeconomic conditions (Byun & Cho, 2013; Nguyen et al., 2021; Zhao et al., 2018). Fluctuating carbon prices may affect the running costs of a company and therefore its operational strategy, since extreme volatility in carbon prices can be risky and create uncertainty in production activities. Hence, effective carbon price forecasting will not only help manage the market risk but also contribute to broader political objectives to tackle climate change.

The currently extensive carbon forecasting methods can be grouped into three streams: traditional *econometric models*, such as GARCH (Liu & Huang, 2021), ARMA/ARMA (Conejo et al., 2005; Sun & Zhang, 2018; Zhou et al., 2006), *machine learning models* such as neural networks (Chiu et al., 2020; Mayer et al., 2015; Yu et al., 2014; Zhu et al., 2017), and *ensemble models* (Lu et al., 2020; Safari & Davallou, 2018; Zhu et al., 2018b). Overall, the integrated model, with its two-stage framework of decomposition-prediction, has come to dominate the field of carbon price forecasting. Given that carbon prices are non-linear and non-stationary, taken alongside the overfitting and local optimization problems of machine learning, many studies apply a combination of econometric models to predict the linear part of the decomposition, and machine learning models to predict the non-linear part (Zhang & Wu, 2022; Zhang et al., 2018; Zhu & Wei, 2013; Zhu et al., 2018a, 2018b). While so far there has been an abundance of studies to explore carbon price forecasting, it is currently difficult to guarantee that any single algorithm proposed in the literature will consistently outperform others all the time. The optimal prediction model may change as the data sample varies. A two-step technique is widely used in the existing literature, specifically, the process of decomposition followed by a single learning method, in an effort to more accurately predict the price outcome of a given set of data. However, the data distribution characteristics of the different sub-series are distinct and so the prediction results may not always be optimal when using the same approach. This paper therefore proposes a new two-step framework for predicting international carbon prices which can serve as a reference for better-informed investment decisions and climate governance in both China and the EU.

More precisely, we first decompose the carbon prices into factors based on trend and period, using the Singular Spectrum Analysis (SSA) approach. The second stage is then to find the best-performing model from among the numerous available prediction models for

the factors of trend and period. We select four generally best-performing learning models and two common deep learning models, Recurrent Neural Network (RNN) and Long short-term memory (LSTM). We then ensemble the results for each sub-factor using the best performing prediction model to obtain the final prediction values for the carbon prices. The advantages of this type of framework are that we do not assume in advance that any learning model will perform better in terms of prediction, but instead try all algorithms which have performed historically well in previous experiments. This openness allows our results to be better adapted to the characteristics of the data distribution of the carbon price series. Such data-driven algorithms are more appropriate and applicable to a wider variety of scenarios. The final prediction results are equivalent to incorporating the optimal prediction values of each decomposition sequence, and therefore clearly outperform the prediction results of one single algorithm. Moreover, the framework in this paper is easily interpretable and extendable, allowing for the inclusion of more feature variables, to further improve the prediction efficiency of the algorithm.

Our study is closely related to the literature that applies the structure of first decomposing the raw carbon price series and then inputting the results into machine learning models, for instance Li et al. (2021), Chai et al. (2021), Hao et al. (2020) and Lu et al. (2020). Chai et al. (2021) explore China's carbon price prediction using Variational Mode Decomposition (VMD) and others, and found it to be limited by boundary effects and sudden signals (Li et al., 2021; Wang et al., 2017). Zhang and Wu (2022) propose a hybrid forecasting method for carbon price in the EU ETS, which incorporates the empirical mode decomposition (EMD) and the integration of ARMA and LSSVMs. Their work shows that this kind of model integration does improve the prediction efficiency of algorithms. Regarding the decomposition step, the popular methods include Empirical Mode Decomposition (EMD) or VMD (Huang, et al., 1998; Zhang et al., 2018; Zhu et al., 2018a), as well as their extensions, such as Multivariate Empirical Mode Decomposition (MEMD) (He et al., 2017), Ensemble EMD (EEMD) (Wu et al., 2019), and Complementary Ensemble Empirical Mode Decomposition (CEEMD) (Li et al., 2021). Zhu (2012) claims that the widely used EMD methods are inherently weak at distinguishing noise terms. In addition, these decomposition methods severely restrict data distribution. The SSA approach gives the best performance in denoising, with the help of its distribution features for subsequences (Hassani, 2007; Zhang et al., 2019). Compared to EMD, SSA can be applied to a wider range of data generating processes, regardless of whether they are linear or nonlinear, stable or unstable, and it also does not require primary functions to be pre-set. More importantly, the SSA applies singular values as a criterion, and its ability to extract the main factors in a given sequence is more effective and intuitive than the alternatives (Vautard et al., 1992).

The second contribution of our paper is that we explore carbon pricing from two different carbon trading markets, namely the European ETS and five main carbon trading pilots in China, including the key markets of Beijing, Shanghai, Guangdong, Shenzhen and Hubei. The EU Emissions Trading System (EU ETS), which was established as early as 2005, is today the largest carbon emissions trading market in the world (Aspinall et al., 2021; Dutta, 2018; Veith et al., 2009). It is well developed, and many private companies, acting as investors in Europe, hedge carbon in this market. China's carbon trading markets are relatively young, with pilot schemes having only been established in a few of the country's larger cities since 2013, with a national carbon trading market being set up in 2021. The eight pilot schemes are based in Beijing, Shanghai, Guangdong, Shenzhen, Hubei, Chongqing, Fujian and Tianjin. The cumulative trading volume of carbon emissions is about 371 million tons, as of the end of 2021. In fact, China produces the highest amount of carbon emissions, accounting for about 30% of the global market in 2018, according to the World Bank. The Chinese carbon

market is therefore of great importance for the country's climate change governance and can also act as a reference for the rest of the world (Leippold et al., 2022; Liu & Sun, 2021; Wang et al., 2019).

Though several studies (Chai et al., 2021; Hao et al., 2020; Li et al., 2021) have discussed carbon prices in either the EU ETS or China, little work comparing international carbon pricing predictions between the EU and China has been done thus far. The environmental challenges, regulatory concerns and policy frameworks in Europe and China are very different, and it is therefore necessary to compare the pricing of these two representative carbon trading systems separately, in order to more thoroughly address global climate change. Hao et al. (2020) show that there are large discrepancies between EU carbon price forecast errors and China's, with their use of a hybrid model containing both feature selection and a multi-objective optimization algorithm.

Our final contribution is in our discussion of the impacts of COVID-19 on global carbon prices, as well as the impact of other key macro-economic variables and the prices of other energy sources.

While a substantial part of the literature is devoted to the social and economic impacts of the unprecedented COVID-19 pandemic, mostly focused on the influence on financial markets, e.g., Akhtaruzzaman et al. (2021a, 2021b, 2022), Boubaker et al. (2022), Cheng and Yen (2020), Hong et al. (2021), Mazur et al. (2020), Pandey and Kumari (2021), Zhang and Mao (2022), Batten et al. (2022), Hassan et al., (2021, 2022) and Choudhury et al. (2022). Akhtaruzzaman et al. (2021a) suggest that there was increased financial contagion during COVID-19. Cheng and Yen (2020) discuss the significant predictive effects of new confirmed cases in the US on S&P 500's return. Hassan et al., (2021, 2022) explore the relative safety of investing in gold during the pandemic. Furthermore, some studies have started to focus on the impacts of unexpected external shocks such as the Russia-Ukraine conflict on the financial markets, for instance, Bounou and Yatić (2022). However, these current studies only discuss the impacts of these types of sudden shock on financial markets like the gold future markets or stock markets, with no analysis given for carbon prices, though the impact on the energy market is evident. Nevertheless, it is already clear that the distribution of carbon emissions changed significantly during the pandemic, when a large number of manufacturers were forced to shut down production. To test if this is indeed the case, our sample spans the period from May 2014 to January 2022, thus allowing for the analysis of the impacts of a lasting event such as the current global public health crisis. We compare the predictive performance of our model with and without the pandemic variable (in this case, the number of new confirmed cases per day) and find that its predictive power is slightly different but not significantly. We can thus conclude that our model is able to capture the various influences of the event.

Furthermore, we discuss whether our results change when taking more macro-economic variables into consideration. In the existing literature on carbon price forecasting, little attention has been paid to the impact of macro variables, or other energy prices such as oil prices, on carbon price forecasting. In fact, all these factors may potentially influence the supply and demand of carbon emissions. Thus, in our extension analysis, we include macro-economic variables such as GDP growth, oil price and coal price. It is found that the predictive effect of our model is somewhat improved with the incorporation of these macro variables. At this point it is worth mentioning the 26th Conference of the Parties (COP 26), an important meeting which took place in 2021. The Glasgow Climate Pact, negotiated during COP26, set out a number of new agreements on climate change, in an attempt to maintain the targets set by the Paris Agreement and further encourage the global transition to lower carbon emissions. These new agreements stress the urgency of taking immediate action, committing to cutting global

CO2 emissions by 45% by 2030, and appeal to all countries to deliver more rigorous plans to reduce emissions. Undoubtedly, this pact places force on the cutting of carbon emissions. Carbon allowances might be referred to as a scarce resource and be allocated differently in distinct industries. However, limited research data available at the time of writing the current paper prevent an in-depth analysis of the impact of this pact on the carbon allowance price market, but it is a matter worth discussing in the future.

Given our new hybrid machine learning framework, we predict the international carbon prices in China and EU spanning from May 2014 to January 2022, with a total of 2011 daily observations. Overall, we find that the classic machine learning methods of SVR and LSSVR outperform other models in forecasting the subsequences of carbon prices, while more complicated deep learning methods turn out to not be the best forecasting algorithm. After considering the impacts of the COVID-19 pandemic (using the number of new cases per day as a proxy) and other macro-state variables such as Term Spread, Oil price and Coal price, our method remains valid.

The remainder of this paper is organized as follows. Section 2 shows the methodology used, followed by a description of the data in Sect. 3. Sections 4 and 5 present the main empirical analysis, and an extended analysis. The last section provides with some implications for future work.

2 Methodology

This section introduces the structure of our new hybrid prediction approach, including a two-stage module, consisting of data preprocessing and forecasting.

2.1 Data preprocessing

SSA, first proposed by Broomhead and King (1986), has been a method of decomposition and denoising of time series data. SSA is a non-parametric method for both stationary and non-stationary time series (Hassani & Zhigljavsky, 2009), decomposing and reconstructing the time series into trend, periodic and noise terms. More precisely, it includes two stages with two small steps for each stage (Hassani, 2007). For more details about SSA decomposition, please see the work of Golyandina and Korobeynikov (2014).

2.1.1 Stage one: decomposition

Step 1: Window embedding

Consider a time series $\mathbb{X} = (x_1, x_2, \dots, x_T)$. Let the integer $L(1 < L < T)$ be the moving window length. Transform \mathbb{X} into $T - L + 1$ vectors:

$$X_i = (x_i, x_{i+1}, \dots, x_{i+L-1})^T, \quad i = 1, 2, \dots, T - L + 1 \tag{1}$$

Then the trajectory matrix \mathcal{X} of \mathbb{X} can be defined as:

$$X = [X_1, X_2, \dots, X_{T-L+1}] = \begin{bmatrix} x_1 & x_2 & \cdots & x_{T-L+1} \\ x_2 & x_3 & \cdots & x_{T-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_T \end{bmatrix}_{L \times (T-L+1)} \tag{2}$$

Step 2: Singular Value Decomposition (SVD)

The eigen triples (s_i, U_i, V_i) of matrix $\mathcal{X}\mathcal{X}^T$ are calculated by SVD. Denote U_i and V_i as the i th left and right eigenvectors of $\mathcal{X}\mathcal{X}^T$. Let $d = \text{rank}(\mathcal{X})$. Then the trajectory matrix can be decomposed as:

$$X = \sum_{i=1}^d \mathcal{X}_i = s_i U_i V_i^T \quad (3)$$

where s_i is the i th singular value of \mathcal{X} .

2.1.2 Stage two: reconstruction

Step 1: Grouping

Group the indices from 1 to d into M disjoint subsets L_1, L_2, \dots, L_M . Accordingly, divide the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_L$ into M groups, that is $\lambda_{L_1}, \lambda_{L_2}, \dots, \lambda_{L_M}$. We thus have:

$$X = \sum_{i=1}^M \mathcal{X}_{L_i} \quad (4)$$

Step 2: Diagonal averaging

In this step, we transform each \mathcal{X}_{L_i} into a new sequence. Define z as the element in each matrix \mathcal{X}_{L_i} ($1 \leq i \leq M$) and z_{ik}^* as the value after diagonal averaging in each \mathcal{X}_{L_i} , that is:

$$z_{ik}^* = \begin{cases} \frac{1}{n+1} \sum_{m=1}^{n+1} z_{m, n-m+2} & 0 \leq n < L-1 \\ \frac{1}{L} \sum_{m=1}^L z_{m, n-m+2} & L-1 \leq n < T-L \\ \frac{1}{T-n} \sum_{m=n-T+L+1}^{T-n} z_{m, n-m+2} & T-L \leq n \leq T \end{cases} \quad (5)$$

By this method, each matrix corresponds to a subsequence. This means that we perform diagonal averaging in matrix \mathcal{X}_{L_i} ($1 \leq i \leq M$) to get $s_{L_i} = (z_{i_1}^*, z_{i_2}^*, \dots, z_{i_T}^*)$ ($1 \leq i \leq M$). Finally, the original time series $\mathbb{X} = (x_1, x_2, \dots, x_T)$ is decomposed into a sum of M subsequences:

$$X = \sum_{i=1}^M s_{L_i} \quad (6)$$

Denote the sum of variance contribution of all singular values as $C(L) = \sum_{i=1}^L s_i$ and the threshold to distinguish the main part and the noise term as μ . Let k be an integer between 1 and L . If $\frac{c(k-1)}{C(L)} < \mu$ and $\frac{C(k)}{C(L)} \geq \mu$, the sequences corresponding to the first k largest singular values are considered the 'main part' and the remaining part are the 'noise term'. The largest singular value is usually understood to reflect the trend of the time series. In this paper, the sequence corresponding to the largest singular value is considered the trend term, and the sum of the remaining part of the first k sequences is considered the periodic term.

2.2 Forecasting module

This part introduces six machine learning models used in this study, namely SVR, LSSVR, RF, XGB, RNN, and LSTM.

(1) Support Vector Regression

Support Vector Machine (SVM) is a good classifier when dealing with binary outcomes, and it can be extended to regression problems to better fit the continuous response variables. Its earliest study was made by Vapnik and Lerner (1963) and its corresponding introduction was completed by Vapnik (1999). In this paper, we use Support Vector Regression (SVR) for a continuous outcome (i.e., the carbon price) instead of the binary form of SVM.

SVR works for both linear and nonlinear relations between response variables and feature variables. For the non-linear data distribution, SVR maps the input space into a high-dimensional feature space. The SVR model is expressed as

$$f(x) = w^T \varphi(x) + b \tag{7}$$

where w and b are the weight vector and interception respectively, and φ is a mapping function. The model collapses to the linear relations when $\varphi(x)$ is x . However, using SVR is a time-consuming process when handling big samples. The alternative LSSVM was proposed by Suykens and Vandewalle (1999) to overcome these disadvantages, which replaces the former's quadratic programming with linear conditions. Similarly, LSSVM's regression framework is defined as least squares support vector regression (LSSVR).

Consider a training data set $\{(x_i, y_i)\}, i = 1, 2, \dots, N$. LSSVR can be expressed as the following problem:

$$\min J(w, \varepsilon) = \frac{1}{2} \left(w^T w + c \sum_{i=1}^N \varepsilon_i^2 \right) \tag{8}$$

subject to

$$y_i = w^T \varphi(x_i) + b + \varepsilon_i, i = 1, 2, \dots, N \tag{9}$$

where c is the penalty parameter, ε_i is the approximation error. Denote λ_i as the Lagrange multiplier, then the Lagrange function can be easily obtained:

$$L(w, \varepsilon, \zeta, b) = J(w, \varepsilon) - \sum_{i=1}^N \zeta_i (w^T \varphi(x_i) + b + \varepsilon_i - y_i) \tag{10}$$

Using the KKT conditions and eliminating w and ε_i , we obtain following expressions:

$$\begin{bmatrix} b \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 & I_1^T \\ I_1 & K + c^{-1}I \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{11}$$

where $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$, I_1 is a unit vector consists of ones and I is an identity matrix.

Finally, the LSSVR model is derived as follows:

$$f(x) = \sum_{i=1}^N \zeta_i K(x, x_i) + b, \tag{12}$$

where $K(x, x_i)$ is the kernel function.

(2) Random forest

Random forest (RF), an ensemble learning method based on a decision tree, was first properly introduced by Breiman (2001). For regression relations, its output is the average of all the decision trees' outputs. Consider a RF model composed of a number of N decision trees. Define $T_j(x)$, $j = 1, 2, \dots, N$ as the output of the j th decision tree, so the RF model can be expressed in the following form:

$$f(x) = \frac{1}{N} \sum_{j=1}^N T_j(x). \quad (13)$$

(3) XGBoost

XGBoost (XGB), a learning method based on the gradient tree boosting, was first proposed by Chen and Guestrin (2016). Boosting is an ensemble algorithm that integrates the multiple weak learning models into a strong learning model, as

$$\hat{y}_i = \sum_{m=1}^M f_m(x_i) \quad (14)$$

where $f_m(\cdot)$ is a weak learning model and M represents the number of weak learning models. It will be a tree-based boosting process if all weak learning models are decision trees.

The objective of XGB is to minimize the regularized objective function L as

$$L = \sum_{i=1}^N l(\hat{y}_i, y_i) + \sum_{m=1}^M \Omega(f_m), \quad (15)$$

where $\Omega(f_m) = \gamma T_m + \frac{1}{2} \kappa \sum_{j=1}^T \omega_{m,j}^2$ represents the penalty term for the complexity of the m th tree; γ and κ represents hyperparameters. T_m indicates the number of nodes in the m th tree. $\omega_{m,j}$ is the value of the j th node in the m th tree.

It is difficult to minimize the regularized objective function by gradient descend method to obtain continuous results, so an additive method is adopted to change the regularized objective function into

$$L = \sum_{i=1}^N l(\hat{y}_i^{(t-1)} + f_t(x_i), y_i) + \Omega(f_t), \quad (16)$$

where $\hat{y}_i^{(t-1)}$ represents the output at the $(t-1)$ th iteration. f_t is the tree at the t th iteration.

(4) Recurrent Neural Network

Recurrent Neural Network (RNN) was first proposed by Elman (1991). RNNs have been widely used for Natural Language Processing (NLP) and time series forecasting over the past few years. Unlike standard neural networks, RNNs allow for the passing the output of hidden layer cells as the input into the next cells. This recurrent structure enables RNNs to construct a memory, which makes RNNs a powerful model to deal with sequential data.

Denote $x = (x_1, x_2, \dots, x_T)$ as a sequence of input, then the hidden state h at time t can be expressed as

$$h_t = \begin{cases} 0, & t = 0 \\ f(w_{xh}x_t + w_{hh}h_{t-1}), & t \geq 1 \end{cases}, \quad (17)$$

where f represents an activation function, w_{xh} represents the weight matrix between the input layer and hidden layer, and w_{hh} represents the weight matrix between the hidden layer at any given time and another hidden layer at an adjacent time.

Finally, the output of RNN at time t can be expressed as

$$y = w_{yh}h_t, \quad (18)$$

where w_{yh} represents the weight matrix between output y and hidden layer h .

(5) Long short-term memory

Long short-term memory (LSTM), proposed by Hochreiter and Schmidhuber (1997), is an improved version of RNN which overcomes the long-term dependencies problem. In an LSTM cell structure, the forget gate f_t , the input gate i_t , the output gate o_t and the cell state candidate \tilde{c}_t are given by (Li et al., 2023):

$$f_t = \sigma(w_{xf}x_t + w_{hf}h_{t-1} + b_f) \quad (19)$$

$$i_t = \sigma(w_{xi}x_t + w_{hi}h_{t-1} + b_i) \quad (20)$$

$$o_t = \sigma(w_{xo}x_t + w_{ho}h_{t-1} + b_o) \quad (21)$$

$$\tilde{c}_t = \tanh(w_{xc}x_t + w_{hc}h_{t-1} + b_c) \quad (22)$$

where w represents the weight matrix, σ is a sigmoid function and b represents the bias term.

Finally, the cell state and the hidden state at the next time step are given by:

$$c_t = c_{t-1} \cdot f_t + i_t \cdot \tilde{c}_t \quad (23)$$

$$h_t = \tanh(c_t) \cdot o_t \quad (24)$$

2.3 Framework of the proposed approach

Our framework, shown in Fig. 1, can be explained as a three-step process:

- (1) *Original series decomposition* we apply the Singular Spectrum Analysis (SSA) to decompose the original carbon prices into several subsequences.
- (2) *Subsequences reconstruction* the subsequence with the largest singular value is considered the 'trend term'. The remaining subsequences are divided into 'period term' and 'noise term' by a threshold of cumulative variance contribution rates.
- (3) *Forecasting* we compare the results of six aforementioned machine learning models by using the trend term and the periodic term after reconstruction, and the best ones are then selected to be the forecasting results of our subsequences. Then we use the simple sum of the predicted subsequences (trend and periodic terms) as the final predictor of carbon prices.

3 Data and predictive features

This section describes the data for analysis and the definitions of predictive features.

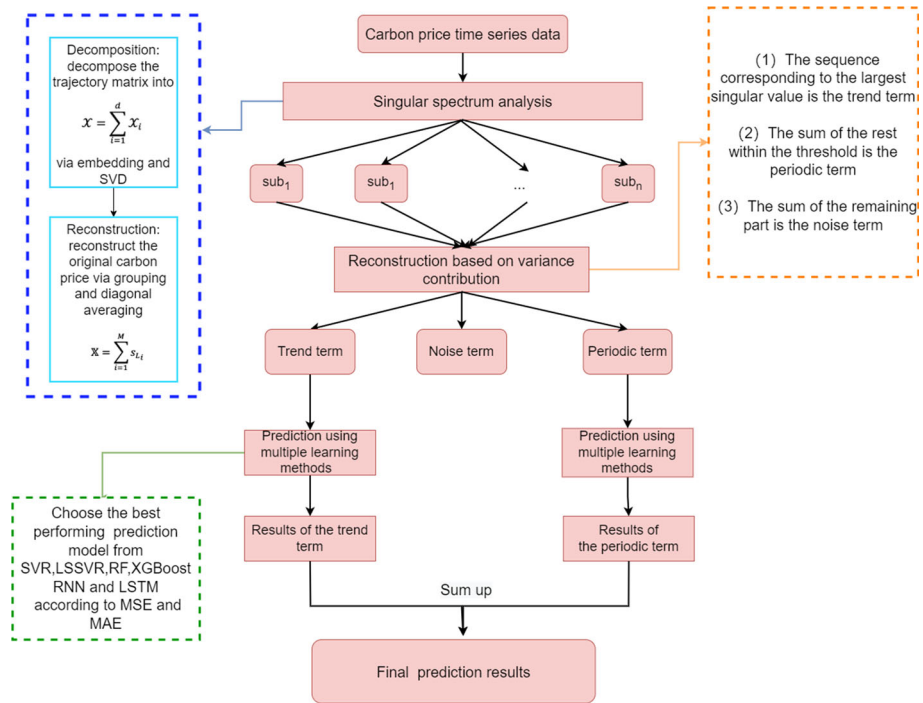


Fig. 1 The framework of our new hybrid machine learning approach. The blue dashed box top left shows the decomposition process of the original carbon valence sequence. The orange dashed box top right shows the decomposed sequences. The green dashed box bottom left gives the machine learning models used to predict the decomposed sequences in the second step. The middle part gives the complete flow of the algorithm. (Color figure online)

3.1 Data description and preparation

We explore the price forecasting in the largest international carbon trading markets, *i.e.* those of the EU and China. More precisely, we use the average transaction price of carbon emission allowances in China. There are eight pilot trading markets which have been set up in China since 2013, namely Beijing (BEA), Shanghai (SHEA), Guangdong (GDEA), Shenzhen (SZA), Hubei (HBEA), Chongqing (CQEA), Tianjin (TJEA), and Fujian (FJEA). We filter out any pilots with small carbon trading turnovers and with more than 50% missing values, retaining five Chinese pilots. These are the BEA, SHEA, GDEA, SZA and HBEA carbon pilots, with the total trading volumes of carbon futures accounting for over 90% of the country's overall volume. In this paper, we also explore the carbon market in EU, *i.e.* the European Union Allowance (EUA) futures prices, taken from ICE (the Intercontinental Exchange). For all carbon transaction prices, we select the daily data for the period from May 2014 to January 2022, giving a total of 2011 observations. Of these, 1810 observations are inputted into the training model. The remaining 201 observations are used to valid the performance on a out-of-time sample. Since we are interested in the prediction of time series of carbon prices, ones which include extreme changes, we do not winsorize our variables. Table 8 in the Appendix shows all variable definitions and data sources in our empirical analysis.

Table 1 describes the price series of six carbon markets in the EU and China. The average carbon price in the EUA ETS is about 18.60 Euro/t, with a standard deviation of 16.51. A high standard deviation usually represents an elevated risk. In China, the BEA holds the highest carbon prices, with the largest standard deviation being 17.22 across all pilots. The cumulative transaction volume of BEA in 2022 is also the largest, with about 3706.6 million Chinese Yuan, followed by HBEA with 1866.8 million Chinese Yuan. The other four pilots in China share a similar mean and standard deviation in their carbon prices. Beyond the mean and standard deviation, the information on price distribution is important. Except for SHEA which is left skewed, all other carbon prices are right skewed. The Kurtosis of EUA and GDEA are 2.40 and 2.51, close to 3, showing sharp peaks. Obviously, all of these price distributions are typically non-normal.

Figure 2 depicts the time-varying prices of six carbon markets in EU and China. The EUA displays a trend of consistently increasing from 2014 to 2022, implying the rising cost of

Table 1 Summary statistics of carbon prices in all carbon markets

	EUA	BEA	SHEA	GDEA	SZA	HBEA
Mean	18.60	60.22	32.35	23.05	27.05	24.58
Std	16.51	17.22	10.66	12.18	13.83	7.32
Minimum	3.93	24.00	4.21	1.27	1.00	9.38
Q1	6.14	50.00	27.32	14.16	15.19	17.45
Median	10.62	53.50	35.99	18.90	27.39	24.95
Q3	25.36	74.60	40.07	28.07	35.64	28.80
Maximum	88.97	107.26	49.93	77.00	77.89	54.64
Skewness	1.61	0.45	-1.16	1.51	0.59	0.40
Kurtosis	2.40	-0.80	0.30	2.51	0.69	0.01

The carbon transaction price of EUA is illustrated with the unit of Euro/t, and that of the various Chinese markets with the unit of Chinese Yuan/t. EUA indicates the European ETS. BEA, SHEA, GDEA, SZA and HBEA are respectively Beijing, Shanghai, Guangdong, Shenzhen and Hubei carbon trading systems in China. Q1 and Q3 are respectively the first and third quartiles of the carbon prices

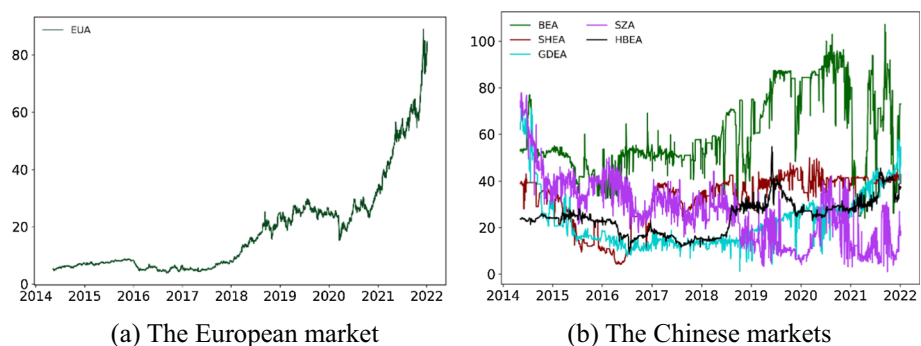


Fig. 2 The time-varying price series for six carbon markets in China and Europe. **a** The time-varying carbon prices in the European carbon trading system. **b** Time-varying carbon prices for the BEA, SHEA, GDEA, SZA and HBEA in China. The sample spans from May 8, 2014 to January 4, 2022

carbon emissions, which is evidently a good incentive for companies to develop low carbon alternatives. However, data from the five pilots in China show different characteristics of time variation. Typically, the BEA exhibits high prices with high volatility. In particular, it sharply declines at the end of 2020, which could well be in response to the stresses of COVID-19. The GDEA, on the other hand, first demonstrates a decreasing trend followed by an increasing trend during the sample period, reaching its lowest point during 2016. One possible reason for this is that the GDEA carbon market was not yet mature at the beginning of our sample, but with the signing of the Paris Agreement, along with the introduction of a range of other pro-climate policies, the carbon trading market is now becoming more active with more demands, meaning that prices are again showing an upward trend. In addition, we can observe from these graphics that all of these carbon price series display non-linear and non-stationary features.

3.2 Predictive features

In the main empirical analysis, we select the lagged decomposed series offered by SSA as the input features $X_{i,t} = \{P_{i,t-1}, P_{i,t-2}, \dots, P_{i,t-K}\}$ where $P_{i,t-K}$ indicates the subsequence (trend term and periodic term) with a lag of K periods. In a time-series model, past prices provide valuable information for prediction, so that we apply the lagged series as the input feature variables in its learning model. In general, long histories in the past show a fading effect in predicting, so we select the lag period as 4 ($K = 4$) which means that the model is able to include information for the past week. When we apply longer lagged periods, the prediction performance is not significantly increased.

To investigate whether our results depend on other features, in the extension analysis, we include several additional macroeconomic variables which may influence carbon prices. Table 8 in the Appendix reports how we define these macro-state variables, such as term spread, GDP growth, international oil price (WTI), coal index, and CSI 300 return (Leipold et al., 2022; Ye & Xue, 2021). Moreover, we explore the impacts of COVID-19 on carbon trading markets. A substantial number of industries were shut down as a result of the pandemic, creating lower carbon emissions. We collect the data of daily new confirmed COVID cases from January 2020 to January 2022 to gain further insight into the impact of the pandemic. Table 2 shows the summary statistics for these additional feature variables.

4 Empirical analysis

First we decompose the price series into both the trend series and periodic series through the SSA approach. By applying the decomposed series, we are then able to examine the prediction performance of the six advanced machine learning methods discussed in Sect. 2.

4.1 Decomposition and reconstruction

Based on the descriptive analysis, we find that the patterns of carbon prices are complex and non-linear. SSA is thus an effective approach for time series analyses such as forecasting or parameter estimation, since it is able to process multiple sub-factors with different information and frequencies. We separate the raw carbon price series into a trend component and periodic component using the so-called Sequential SSA (Golyandina & Korobeynikov, 2014). The

Table 2 Summary statistics of predictive features in the extension analysis

	Term spread	GDP growth	Oil price	Coal price	CSI300Re	COVID_CN	COVID_EU
Mean	0.60	6.50	56.26	156.06	0.00	186.84	143,590.69
Std	0.28	3.34	16.03	25.55	0.01	802.10	132,703.89
Minimum	-0.07	-6.90	-37.63	124.70	-0.09	1.00	-17,071.00
Q1	0.46	6.30	46.25	137.10	-0.01	25.00	31,756.75
Median	0.58	6.90	53.55	156.20	0.00	49.00	132,825.50
Q3	0.70	7.00	64.30	161.87	0.01	106.25	214,545.75
Maximum	1.94	18.30	107.26	235.80	0.06	15,152.00	1,024,371.00
Skewness	1.45	-0.70	0.72	1.61	-0.96	13.56	1.85
Kurtosis	4.68	10.45	1.80	3.03	6.30	235.68	7.22

All non-daily data are transferred to daily data through the constant interpolation method. Term spread, GDP growth, Oil price, Coal price and CSI300Re are the macro-economic variables used in our extension analysis. COVID_CN and COVID_EU are the proxies of the COVID-19 pandemic. All variable definitions are shown in Table 8

decomposition process of the trend factor is smoothing. To obtain the periodic factor, we choose a window length of 22 trading days, which is almost equivalent to a 1-month sample.

We first report the decomposition results of the EUA. Figure 3 shows six leading eigenvectors and their corresponding reconstructed series. It indicates that the first component in the eigenvalue decomposition represents the trend factor. Thus, the reconstructed trend series, based on the first eigenvector, is shown in Fig. 4. At this stage, we decompose the raw carbon prices with a moving window length of 22 trading days (almost 1 month) through the sequential SSA approach. We find that the trend term captures the main pattern of the raw carbon prices. For the EUA, this trend continues to increase during our sample period, implying a growing emphasis on regulating carbon emissions across the continent, as well as

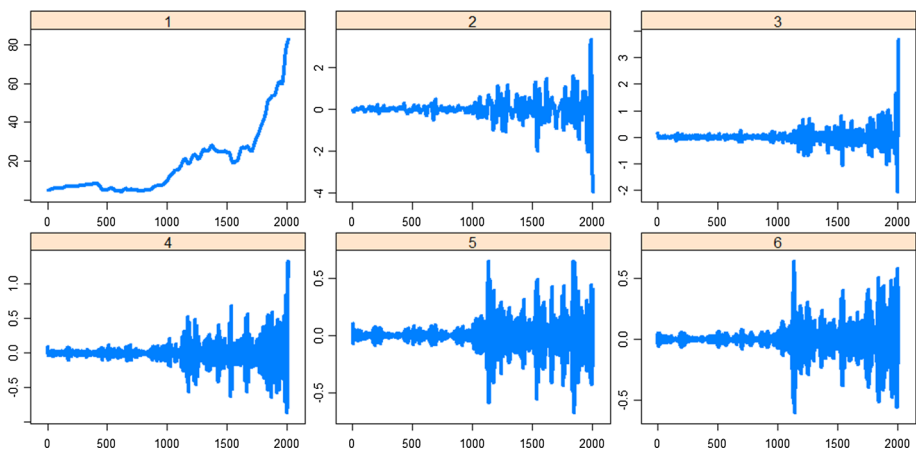


Fig. 3 The reconstructed series of six leading eigenvectors in the first stage of SSA for EUA ($L = 22$). The decomposition window length is 22

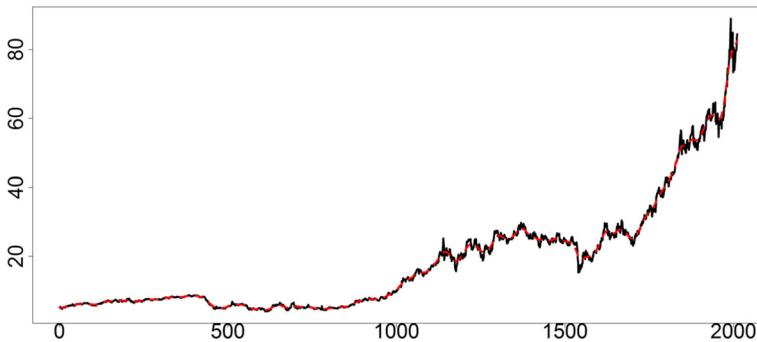


Fig. 4 The trend series of the EUA obtained through the SSA approach. The solid black line is the raw carbon price, and the red dotted line corresponds to the trend factor. The decomposition window length is 22. (Color figure online)

the increased demand for such regulations from industry and other consumers. Based on the residual series in the last stage, we then go on to extract the periodic component. Golyandina and Korobeynikov (2014) claim that the window length should be as long as possible to include more information, suggesting that the window should be half the length of the full sample period, and from that the first moving window length can be divided ($L = 22$ in this case). Thus, we reset the second window length as 990 trading days. Figure 5 depicts the periodic component of the EUA's carbon price. It can be observed that the periodic fluctuations are not significant in the early years but that they have increased significantly in the last few years (i.e. at the end of the sample).

For brevity, we only discuss here the distributional characteristics of the Chinese carbon allowance trading price with an example of BEA. Figure 6 suggests that the first component can be used as an indicator of the overall trend of the BEA carbon prices, so we do the same to obtain the trend term, as shown in Fig. 7. We can see that the pattern of BEA is obviously different to that of the EUA. It can be observed, for instance, that the price of carbon did not rise consistently over the sample period. Within the first 1500 observations, we see a general upward trend, implying a greater demand for carbon allowances. In contrast, in the sample of between 1700 and 1900 observations (taken in the period between November 2020 and June

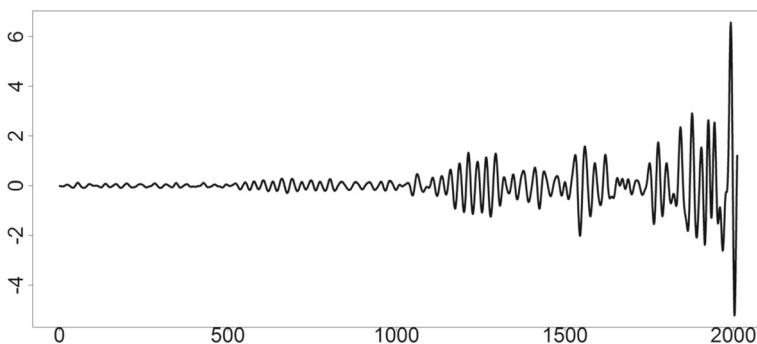


Fig. 5 The periodic component of the EUA through the SSA approach. The decomposition window length is 990

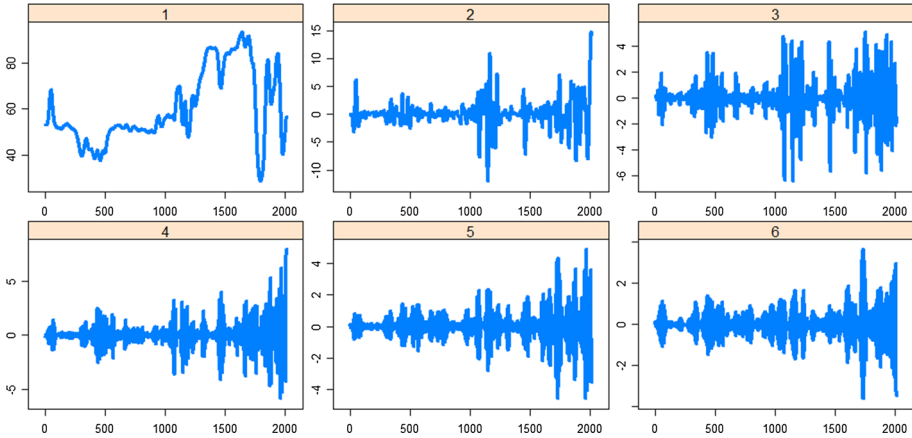


Fig. 6 The reconstructed series of the six leading eigenvectors in the first stage of SSA for BEA ($L = 22$). The decomposition window length is 22

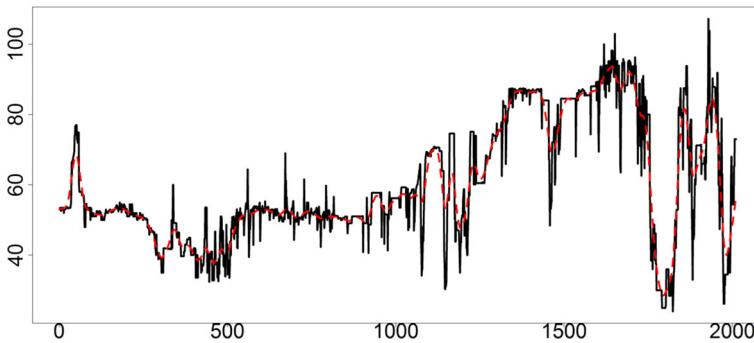


Fig. 7 The trend series of the BEA obtained through the SSA approach. The solid black line is the raw carbon price, and the red dotted line corresponds to the trend factor. The decomposition window length is 22. (Color figure online)

2021), we see a clear trend of precipitous falls and sharp rises. This can be seen as a response to the intensifying climate policy aimed at achieving carbon peak and carbon neutral targets, as proposed by the Chinese government. Figure 8 is the periodic term of BEA which shows a higher volatility than the EUA.

The carbon prices of other four pilots in China are also similarly decomposed. Figures 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 and 21 show both the trend and periodic factors for each.

4.2 Carbon prediction results

We employ Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) as two common indicators to evaluate the predictive performance (Chai et al., 2021). A smaller RMSE

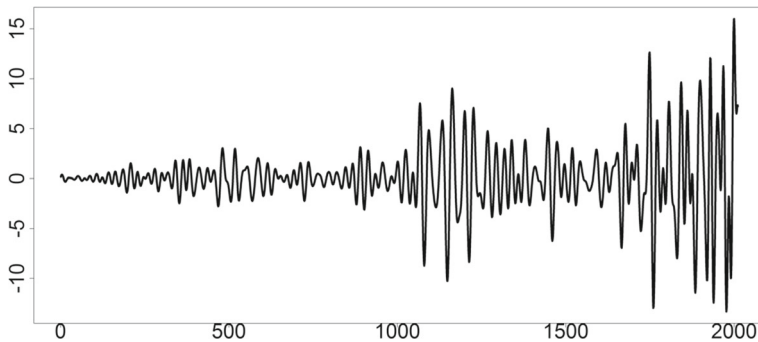


Fig. 8 The periodic component of the BEA obtained through the SSA approach. The decomposition window length is 990

and MAE indicate a better predictive performance. The definitions are given as follows:

$$\text{RMSE} = \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2 / N} \quad (25)$$

$$\text{MAE} = \sum_{i=1}^N |y_i - \hat{y}_i| / N \quad (26)$$

where y_i is the true values of both the trend factor and periodic factor for i th price series. \hat{y}_i corresponds to the estimated values and N is the series length.

To predict carbon prices more accurately, we train our data using six typical machine learning models, including four classic machine learning models and two deep learning methods. More precisely, the four classic machine learning models used are SVR, LSSVR, RF and XGB. The other two models are deep learning models, that is, LSTM and RNN. In practice, a more complicated model does not indicate a better predictive performance, which in reality depends far more on the data distribution. In this paper, we first train several models to discover the best performing prediction model to find both the trend factor and periodic factor. We then ensemble the best models to predict the decomposed series and so the final carbon prices.

Tables 3 and 4 report the out-of-performance RMSE and MAE of all methods for all decomposed series. Some interesting findings are summarized here.

- (1) The predictive performance of these models presents significant heterogeneity. This heterogeneity is determined by the different data distribution profiles. Tables 3 and 4 reveal that for the trend component, the classical support vector regression model (SSA-SVR) performs best in all carbon trading markets. With respect to the periodic component, the classical support vector regression model (SSA-SVR) performs best in the EUA, BEA, SHEA and HBEA carbon trading markets, while the least squares support vector regression model (SSA-LSSVR) performs best in the GDEA and SZA carbon trading markets.
- (2) For the process of data generation in this paper, the forecasting performance of the deep learning algorithm is not significantly improved, and in fact performs the worst out of all the algorithms used.

Table 3 The RMSE for all carbon trading markets

	SSA-LSSVR	SSA-SVR	SSA-RF	SSA-XGB	SSA-RNN	SSA-LSTM
<i>Trend</i>						
EUA	0.04514	0.00879	19.36467	0.53560	0.65031	18.64211
BEA	0.08332	0.06251	0.90825	2.34177	0.23014	15.26186
SHEA	0.00674	0.00391	0.06267	0.09836	0.03971	0.17703
GDEA	0.03376	0.01850	1.30128	0.60485	0.09119	0.37026
SZA	0.04812	0.03088	0.32648	0.75556	0.06405	0.11777
HBEA	0.02695	0.00762	1.02627	0.43804	0.06787	0.58121
<i>Periodic</i>						
EUA	0.03584	0.02979	1.05089	0.17752	0.33777	0.67873
BEA	0.10866	0.04060	1.20526	0.82135	0.23772	0.50631
SHEA	0.00581	0.00440	0.07606	0.07643	0.00799	0.01102
GDEA	0.03524	0.03595	0.25043	0.22103	0.04415	0.04154
SZA	0.10258	0.10336	0.99323	0.48324	0.34457	0.40613
HBEA	0.00803	0.00495	0.25797	0.12590	0.09141	0.05635

Since all machine learning models are applied to the decomposed series from the SSA approach, we denote all models using the prefix SSA. The results of XGB, RNN and LSTM are the average values over 30 time replications. The bold number indicates the minimum RMSE among all methods for each carbon trading market

Table 4 The MAE for all carbon trading markets

	SSA-LSSVR	SSA-SVR	SSA-RF	SSA-XGB	SSA-RNN	SSA-LSTM
<i>Trend</i>						
EUA	0.03379	0.00651	16.55468	0.38306	0.45033	15.75852
BEA	0.06392	0.03449	0.66014	1.97046	0.13829	13.47671
SHEA	0.00540	0.00223	0.05085	0.08008	0.03918	0.16768
GDEA	0.02603	0.01137	1.24263	0.54394	0.08589	0.33595
SZA	0.03626	0.02121	0.25488	0.58543	0.05004	0.09381
HBEA	0.02136	0.00552	0.67749	0.31490	0.06227	0.37341
<i>Periodic</i>						
EUA	0.02515	0.01094	0.53540	0.13352	0.12267	0.26955
BEA	0.08232	0.02640	0.89143	0.64026	0.13329	0.22429
SHEA	0.00425	0.00218	0.06203	0.06279	0.00638	0.00886
GDEA	0.02119	0.02044	0.19697	0.18978	0.03090	0.02363
SZA	0.07880	0.07951	0.70860	0.38625	0.16976	0.18817
HBEA	0.00592	0.00291	0.14866	0.09380	0.02845	0.03464

Since all the machine learning models are applied to the decomposed series from the SSA approach, we denote all models using the prefix SSA. The results of XGB, RNN and LSTM are the average values over 30 time replications. The bold number indicates the minimum MAE among all methods for each carbon trading market

- (3) The optimal prediction models for different carbon price components are distinct, which makes it impossible to consistently and definitively state which algorithm is better. This problem also supports the validity of our framework. In general, the use of integrated learning improves predictive performance.
- (4) The tree model algorithm performs the most poorly among all algorithms chosen. For instance, the RF has a high RMSE of 19.36467 for the trend component of EUA, which is close to being totally unpredictable. Consequently, we conclude that the tree model is not a good choice for time series.

Figure 9 shows the predicted figures for all carbon prices. Since we decompose the raw carbon prices into two factors, the trend factor and periodic factor, the predicted carbon price series is equal to the sum of their predicted series. Our hybrid model can effectively predict the carbon prices in both the EU and China. If we take the example of BEA, we can see that the pattern of its sharp increase from April 2021 to May 2021 is effectively predicted by the model. Table 5 summarizes the overall predictive performance for the price of carbon.

5 Extension analysis

In this section, we further explore the impacts of the macro-economy. In the main analysis, we only include the lag terms of the decomposed series of carbon prices, ignoring the potential influences of external factors. One may argue that during the COVID-19 crisis, a considerable number of companies shut down either temporarily or permanently, with an appreciable drop in CO₂ emissions, resulting in the shift in carbon prices. Thus, the predictive performance of our framework could be adversely impacted due to the pandemic.

To address this concern, we introduce a new feature variable into our machine learning model, ‘new cases per day’ to be a proxy for the pandemic. COVID_CN and COVID_EU are the values allocated for the COVID impact in China and EU respectively. The sample spans the period from January 2020 to January 2022, with a total of 524 observations. Table 6 compares the predictive performance of our model with and without taking the impacts of COVID-19 into consideration. For the sake of simplicity, we only report the RMSE here. We find that the predictive performance of the decomposed carbon price series is not significantly improved after introducing the epidemic variable. This implies that our framework is robust.

Another concern for the integrity of our framework is the influence of macroeconomic variables. The carbon trading market depends on the carbon emissions of firms, which is dictated by their climate policy framework, so important macroeconomic indicators such as GDP growth are potential factors to take into consideration. Here we also add these macroeconomic variables, as shown in Table 8 in the Appendix. Table 7 demonstrates that our predictive performance remains with the introduction of these macro variables, whereas a significant decline in RMSE cannot be observed. While it is clearly important to include these macro variables in prediction models, their impact should not be overestimated.

6 Conclusion

In this paper, we propose a hybrid machine learning-based method to predict international carbon price patterns. This hybrid framework, with its two stages of combining SSA and machine learning methods shows power in effectively forecasting carbon prices in both the

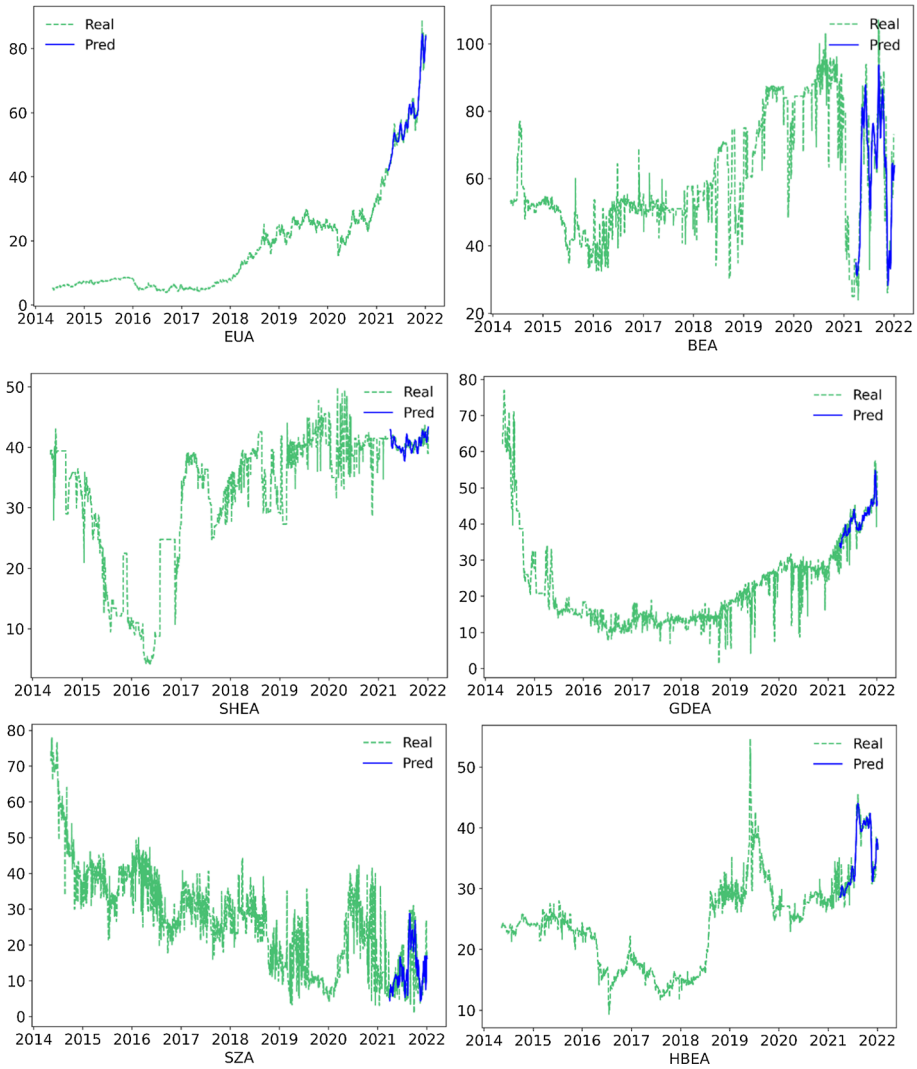


Fig. 9 The predicted carbon prices and real prices for all markets (EUA, BEA, SHEA, GDEA, SZA and HBEA). The green dotted line represents the real carbon price, and the solid blue line is the predicted carbon price. (Color figure online)

Table 5 The overall out-of-performance for all carbon prices in the EU and China

	EUA	BEA	SHEA	GDEA	SZA	HBEA
RMSE	1.34257	6.59269	1.00419	2.29022	4.93781	1.21783
MAE	0.96651	5.18567	0.77259	1.64402	3.64036	0.95507

In the process of carbon price prediction using machine learning methods, we split the whole sample into a training set and a test set. The test set consisting of 201 daily prices is used to examine the validity of our framework

Table 6 The RMSE of the trend component and periodic component for all carbon prices

	EUA	BEA	SHEA	GDEA	SZA	HBEA
<i>Trend</i>						
With COVID	0.01097	0.07899	0.00520	0.02262	0.03766	0.00843
Without COVID	0.01391	0.07901	0.00556	0.02280	0.03769	0.00848
<i>Periodic</i>						
With COVID	0.04085	0.04932	0.00763	0.04586	0.11739	0.00883
Without COVID	0.03751	0.04932	0.00602	0.04587	0.11756	0.00660

All trend series are predicted by the SVR method. The periodic components of EUA, BEA, SHEA, and HBEA are predicted by the SVR model, and those of GDEA and SZA are predicted by the LSSVR model. The test sample accounts for about 20% of the total, with 104 observations

Table 7 The RMSE of the trend component and periodic component for all carbon prices

	EUA	BEA	SHEA	GDEA	SZA	HBEA
<i>Trend</i>						
With macro	0.00884	0.06227	0.00631	0.01806	0.03115	0.00762
Without macro	0.00874	0.06251	0.00389	0.01852	0.03087	0.00761
<i>Periodic</i>						
With macro	0.02995	0.04062	0.00611	0.03528	0.10259	0.00497
Without macro	0.02961	0.04067	0.00438	0.03524	0.10258	0.00495

All trend series are predicted by the SVR method. The periodic components of EUA, BEA, SHEA, and HBEA are predicted by the SVR model, and those of GDEA and SZA are predicted by the LSSVR model. The test sample accounts for about 20% of the total, with 104 observations

EU and China, particularly in terms of out-of-sample prediction. It remains valid in the exceptional circumstances of the COVID-19 pandemic, as well as during different macroeconomic conditions. We find that for predicting carbon price, complex models such as deep learning do not always perform consistently, while support vector regression stands out at the second stage.

There are four steps in total in our two-stage framework, which consists of the separation process and the prediction process. For the separation process, we firstly use the SSA approach to separate the raw carbon prices, eliminating useless noises and then combine the factors into the trend factors and periodic factors as major drivers of price. For the prediction process, various machine learning models are applied to the trend and periodic terms, enabling us to identify the optimal predictive model for each term separately. Finally, the outputs of the corresponding optimal learning models for each factor are integrated into a price series. The framework can accurately predict the time-varying trend of carbon prices.

This paper provides implications for investors, regulators and policy makers. First, investors can apply the pricing model to hedging in the carbon market. Better understanding of the new type of assets can give them an edge in the futures market. Second, since the emerging carbon market calls for increased attention from many countries, regulators may wish to control the underlying risk and its contagion with other financial markets. Some of the trend and periodic factors seen here are shared by other financial assets as well. Third,

global cooperation requires that nations and governments become key players in the management of carbon emissions. Policies related to carbon emission control can be better evaluated and guided by price. Policy makers need reliable information to formulate climate change policies and greenhouse targets. In 2020 China set a goal of achieving a carbon peak in 2030 and full carbon neutrality by 2060. To help achieve this, there were seven operational carbon exchanges open in 2021 in China. An efficient carbon allowance trading system is an essential part of such an ambitious plan.

Finally, we understand that this research is constrained by the limited number of observations made in a time window which stopped at the beginning of 2022, and also that it only covers a selection of carbon markets due to the current availability of data. The proposed prediction framework can obviously be extended to other international carbon markets, including those in the US, Singapore, New Zealand and South Korea. Furthermore, The UN Climate Change Conference (COP26) held at the end of 2021 set new targets on the cutting of carbon emissions, which form the basis of the Glasgow Climate Pact. This shows the global efforts being made to try to mitigate the effects of climate change. International conflicts, including the ongoing Russia-Ukraine crisis, also cause tensions within the energy market and obviously reshape both the energy structure and carbon emissions not just in Europe but all over the world. However, the impact of these events is yet to significantly affect carbon price or volatility. Future studies can focus on these events and test the robustness of our framework in a new context.

Acknowledgements We acknowledge the support of the National Natural Science Foundation of China (No. 72101209) and the Fundamental Research Funds for the Central Universities, China (JBK2103013). This work was also supported by Financial Innovation Centre (Project No. 2022A0011) at Southwestern University of Finance and Economics, China. We thank the editors and the anonymous reviewers for their comments to improve this paper.

Author contributions XZ: writing, methodology, conceptualization, revising, funding acquisition. ZL: conceptualization, revising, resources, supervision, funding acquisition. YZ: formal analysis, software, investigation, revising. LW: data curation, formal analysis, writing.

Availability of data and materials This article uses publicly disclosed data obtained from DataStream Database.

Code availability The programming code can be provided upon request.

Declarations

Conflict of interest None.

Appendix

Variable definitions and data sources

In this section, we show all variables in our sample, including definition, observations, data source and frequency. These variables are the carbon prices in the EU and China's carbon allowance trading systems, macro-economic variables in our extension analysis, as well as pandemic variables (Table 8).

Table 8 Variables definitions and data sources of empirical analysis

	Definition	Observations	Frequency	Data source
<i>Panel A: Carbon prices</i>				
EUA	The average price of European Union Allowance (EUA) futures	2011	Daily	Datastream
BEA	The average price of carbon emission allowances in Beijing, China	2011	Daily	Datastream
SHEA	The average price of carbon emission allowances in Shanghai, China	2011	Daily	Datastream
GDEA	The average price of carbon emission allowances in Guangdong, China	2011	Daily	Datastream
SZA	The average price of carbon emission allowances in Shenzhen, China	2011	Daily	Datastream
HBEA	The average price of carbon emission allowances in Hubei, China	2011	Daily	Datastream
<i>Panel B: Macro-economic variables</i>				
Term spread	The first difference between 10-year government bond yield and 1-year government bond yield	2011	Daily	Datastream
GDP growth	The seasonality GDP growth rate	31	Quarterly	Datastream
Oil price	WTI crude oil future price	2011	Daily	Datastream
Coal price	China coal price index	395	Weekly	Datastream
CSI300Re	The stock logarithm returns of CSI 300 index	2011	Daily	Datastream
<i>Panel C: Pandemic variables</i>				
COVID_CN	The new cases per day in China from Jan 2020 to Jan 2022	524	Daily	Datastream
COVID_EU	The new cases per day in EU from Jan 2020 to Jan 2022	524	Daily	Datastream

Missing values are filled in by the constant interpolation method

Graphic analysis

(1) Decomposition

In this section, we apply the SSA approach to decompose the raw carbon prices. Figures 10, 11, 12 and 13 show the reconstructed series of six leading eigenvectors in the first stage for the four pilots, except for BEA in China. We find that the first component mainly drives the pattern of carbon price.

(2) Reconstruction

In this section, we reconstruct the decomposed subseries in the first stage and obtain the trend factor and periodic factor. Figures 14, 15, 16 and 17 show the trend factors of the four pilots in China, and Figs. 18, 19, 20 and 21 depict the corresponding periodic factors.

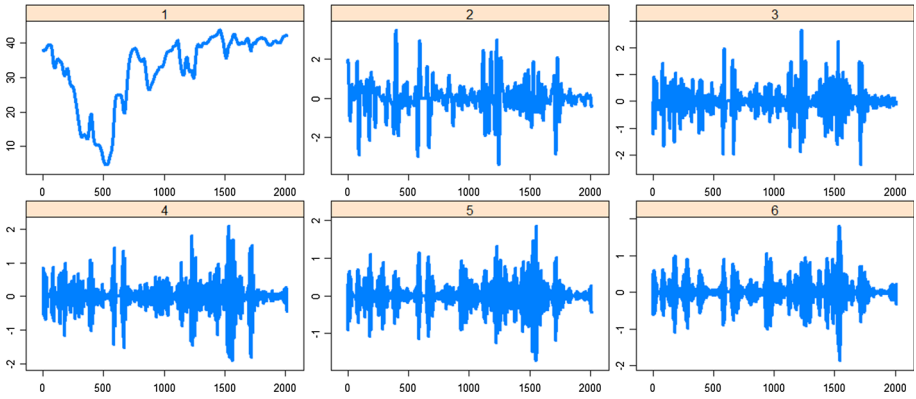


Fig. 10 The reconstructed series of six leading eigenvectors in the first stage for SHEA ($L = 22$). The decomposition window length is 22

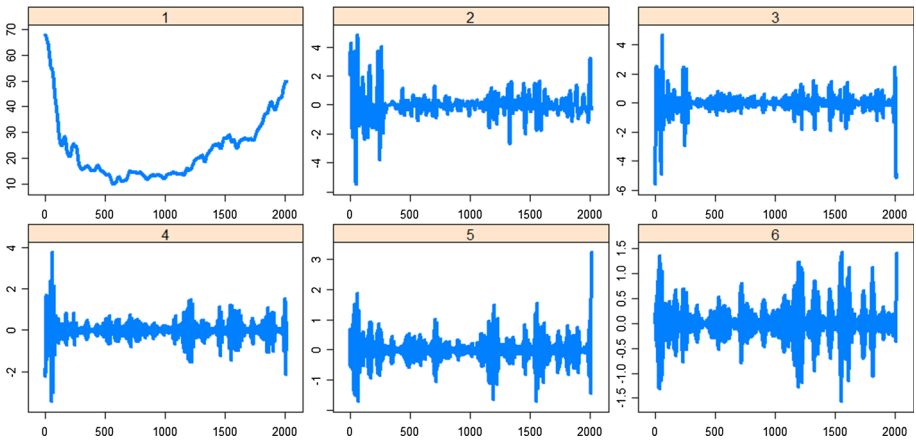


Fig. 11 The reconstructed series of the six leading eigenvectors in the first stage for GDEA ($L = 22$). The decomposition window length is 22

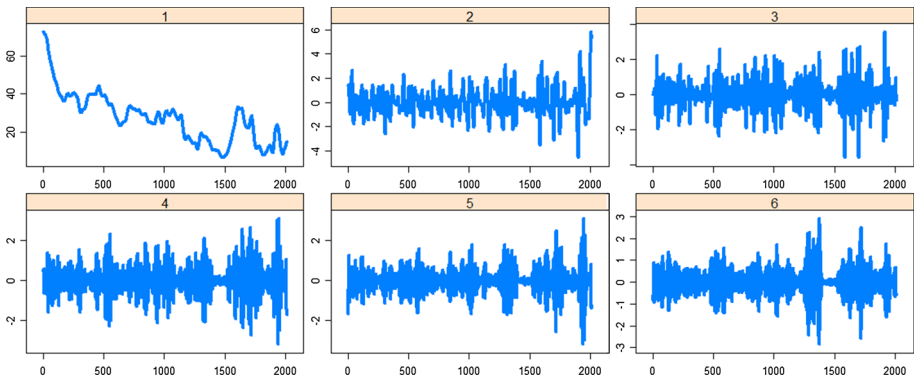


Fig. 12 The reconstructed series of the six leading eigenvectors in the first stage for SZA ($L = 22$). The decomposition window length is 22

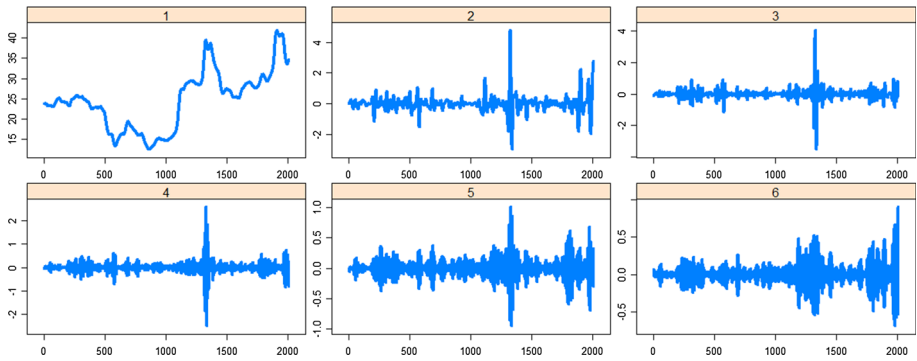


Fig. 13 The reconstructed series of the six leading eigenvectors in the first stage for HBEA ($L = 22$). The decomposition window length is 22

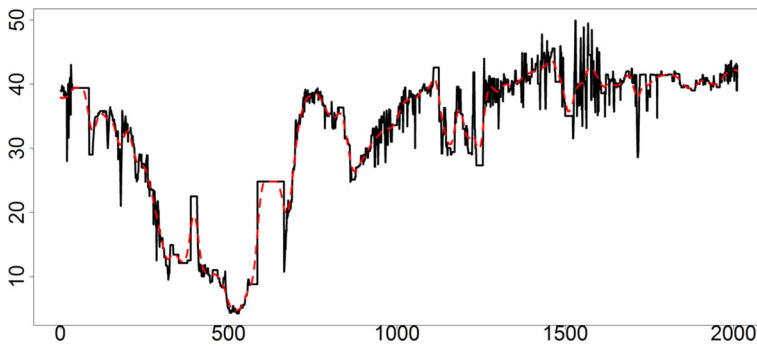


Fig. 14 The trend series of the SHEA through the SSA approach. The solid black line is the raw carbon price, and the red dotted line corresponds to the trend factor. The decomposition window length is 22. (Color figure online)

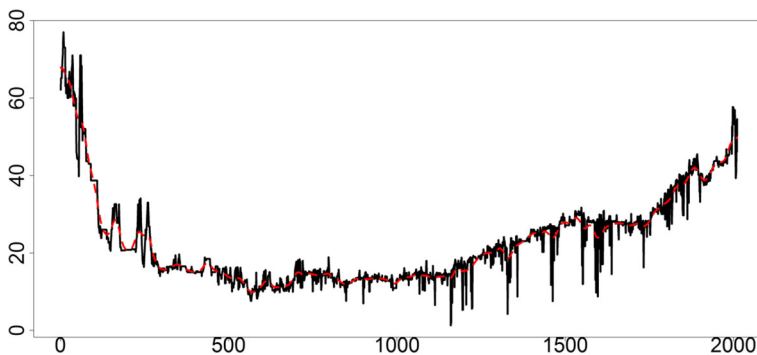


Fig. 15 The trend series of the GDEA through the SSA approach. The solid black line is the raw carbon price, and the red dotted line corresponds to the trend factor. The decomposition window length is 22. (Color figure online)

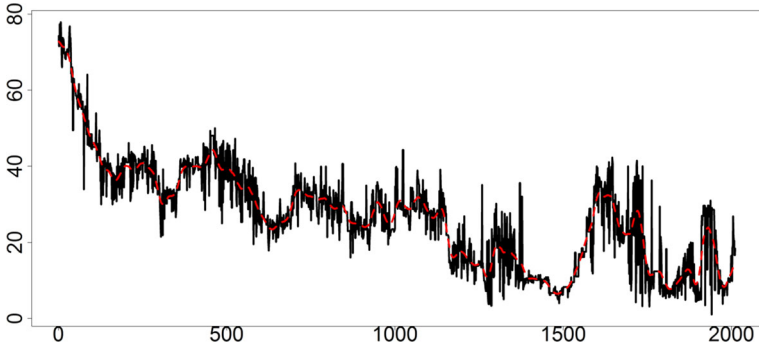


Fig. 16 The trend series of the SZA through the SSA approach. The solid black line is the raw carbon price, and the red dotted line corresponds to the trend factor. The decomposition window length is 22. (Color figure online)

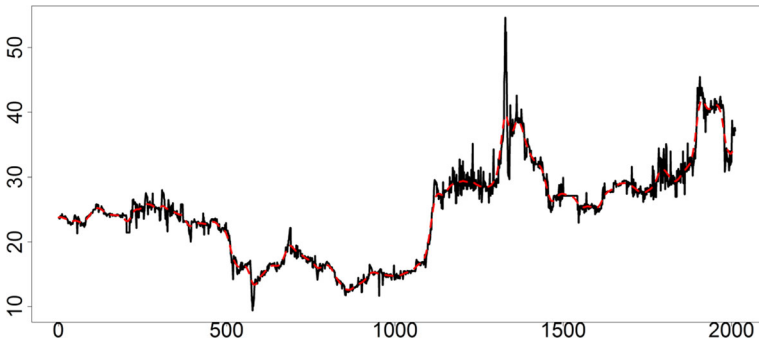


Fig. 17 The trend series of the HBEA through the SSA approach. The solid black line is the raw carbon price, and the red dotted line corresponds to the trend factor. The decomposition window length is 22. (Color figure online)

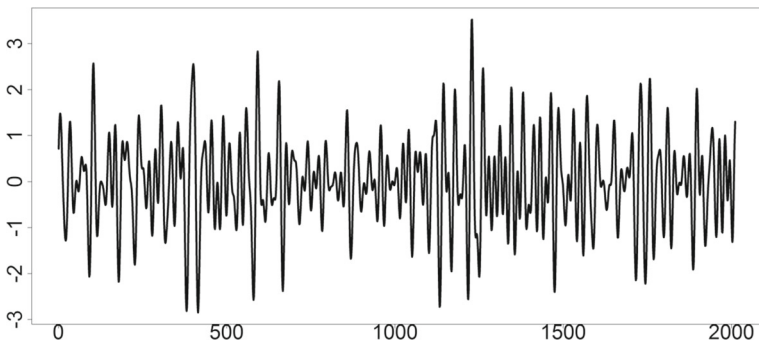


Fig. 18 The periodic component of the SHEA through the SSA approach. The decomposition window length is 990

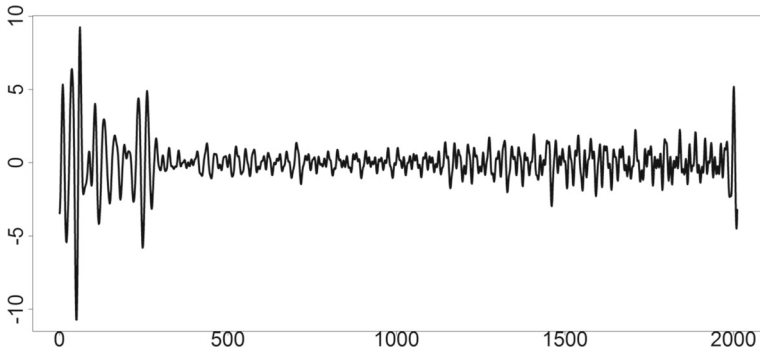


Fig. 19 The periodic component of the GDEA through the SSA approach. The decomposition window length is 990

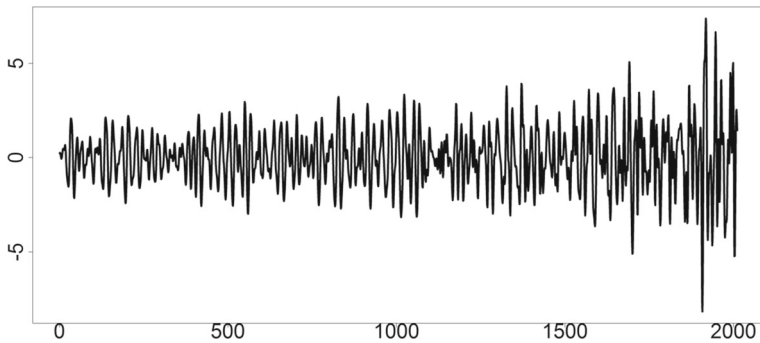


Fig. 20 The periodic component of the SZA through the SSA approach. The decomposition window length is 990

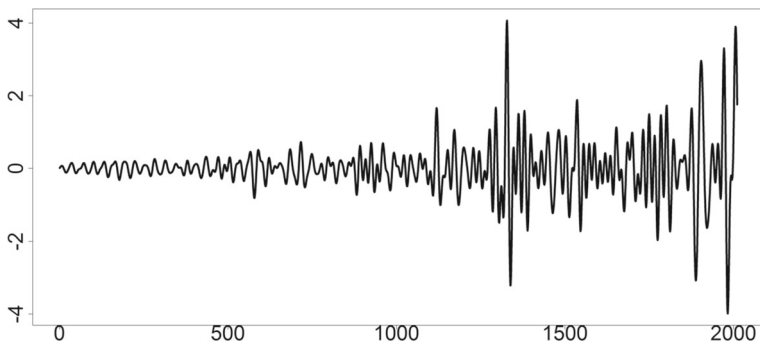


Fig. 21 The periodic component of the HBEA through the SSA approach. The decomposition window length is 990

References

- Akhtaruzzaman, M., Boubaker, S., Chiah, M., & Zhong, A. (2021a). COVID-19 and oil price risk exposure. *Finance Research Letters*, *42*, 101882.
- Akhtaruzzaman, M., Boubaker, S., & Sensoy, A. (2021b). Financial contagion during COVID-19 crisis. *Finance Research Letters*, *38*, 101604.
- Akhtaruzzaman, M., Boubaker, S., & Umar, Z. (2022). COVID-19 media coverage and ESG leader indices. *Finance Research Letters*, *45*, 102170.
- Aspinall, T., Gepp, A., Harris, G., Kelly, S., Southam, C., & Vanstone, B. (2021). Estimation of a term structure model of carbon prices through state space methods: The European Union emissions trading scheme. *Accounting & Finance*, *61*(2), 3797–3819.
- Batten, J. A., Choudhury, T., Kinateder, H., et al. (2022). Volatility impacts on the European banking sector: GFC and COVID-19. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-022-04523-8>
- Bauer, N., Bertram, C., Schultes, A., Klein, D., Luderer, G., Kriegler, E., Popp, A., & Edenhofer, O. (2020). Quantification of an efficiency–sovereignty trade-off in climate policy. *Nature*, *588*(7837), 261–266.
- Boubaker, S., Goodell, J. W., Pandey, D. K., & Kumari, V. (2022). Heterogeneous impacts of wars on global equity markets: Evidence from the invasion of Ukraine. *Finance Research Letters*, *48*, 102934.
- Boungou, W., & Yatié, A. (2022). The impact of the Ukraine-Russia war on world stock market returns. *Economics Letters*, *215*, 110516.
- Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.
- Broomhead, D. S., & King, G. P. (1986). Extracting qualitative dynamics from experimental data. *Physica d: Nonlinear Phenomena*, *20*(2–3), 217–236.
- Byun, S. J., & Cho, H. (2013). Forecasting carbon futures volatility using GARCH models with energy volatilities. *Energy Economics*, *40*, 207–221.
- Chai, S., Zhang, Z., & Zhang, Z. (2021). Carbon price prediction for China's ETS pilots using variational mode decomposition and optimized extreme learning machine. *Annals of Operations Research*, 1–22.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM sigkdd international conference on knowledge discovery and data mining* (pp. 785–794).
- Cheng, H. -P., & Yen, K. -C. (2020). Does COVID-19 affect the financial market? Available at SSRN 3578263.
- Chiu, Y. J., Hu, Y. C., Jiang, P., Xie, J., & Ken, Y. W. (2020). A multivariate grey prediction model using neural networks with application to carbon dioxide emissions forecasting. *Mathematical Problems in Engineering*, *2020*(2), 1–10.
- Choudhury, T., Kinateder, H., & Neupane, B. (2022). Gold, bonds, and epidemics: A safe haven study. *Finance Research Letters*, *48*, 102978.
- Conejo, A. J., Plazas, M. A., Espinola, R., & Molina, A. B. (2005). Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Transactions on Power Systems*, *20*(2), 1035–1042.
- Dutta, A. (2018). Modeling and forecasting the volatility of carbon emission market: The role of outliers, time-varying jumps and oil price risk. *Journal of Cleaner Production*, *172*, 277.
- Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, *7*(2), 195–225.
- Golyandina, N., & Korobeynikov, A. (2014). Basic singular spectrum analysis and forecasting with R. *Computational Statistics & Data Analysis*, *71*, 934–954.
- Hao, Y., Tian, C., & Wu, C. (2020). Modelling of carbon price in two real carbon trading markets. *Journal of Cleaner Production*, *244*, 118556.
- Hassan, M. K., Djajadikerta, H. G., Choudhury, T., & Kamran, M. (2021). Safe havens in Islamic financial markets: COVID-19 versus GFC. *Global Finance Journal*, *54*, 100643.
- Hassan, M. K., Kamran, M., Djajadikerta, H. G., & Choudhury, T. (2022). Search for safe havens and resilience to global financial volatility: Response of GCC equity indexes to GFC and COVID-19. *Pacific-Basin Finance Journal*, *73*, 101768.
- Hassani, H. (2007). Singular spectrum analysis: Methodology and comparison. *MPRA Paper*, *5*(2), 239–257.
- Hassani, H., & Zhigljavsky, A. (2009). Singular spectrum analysis: Methodology and application to economics data. *Journal of Systems Science and Complexity*, *22*(3), 372–394.
- He, K., Chen, Y., & Tso, G. K. (2017). Price forecasting in the precious metal market: A multivariate EMD denoising approach. *Resources Policy*, *54*, 9–24.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.
- Hong, H., Bian, Z., & Lee, C.-C. (2021). COVID-19 and instability of stock market performance: Evidence from the U.S. *Financial Innovation*, *7*(1), 12.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings Mathematical Physical & Engineering Sciences*, *454*(1971), 903–995.

- Leippold, M., Wang, Q., & Zhou, W. (2022). Machine learning in the Chinese stock market. *Journal of Financial Economics*, 145(2), 64–82.
- Li, H., Jin, F., Sun, S., & Li, Y. (2021). A new secondary decomposition ensemble learning approach for carbon price forecasting. *Knowledge-Based Systems*, 214, 106686.
- Li, S., Tian, Z., & Li, Y. (2023). Residual long short-term memory network with multi-source and multi-frequency information fusion: An application to China's stock market. *Information Sciences*, 622, 133–147.
- Liu, Z., & Huang, S. (2021). Carbon option price forecasting based on modified fractional Brownian motion optimized by GARCH model in carbon emission trading. *The North American Journal of Economics and Finance*, 55, 101307.
- 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.
- Lu, H., Ma, X., Huang, K., & Azimi, M. (2020). Carbon trading volume and price forecasting in China using multiple machine learning models. *Journal of Cleaner Production*, 249, 119386.
- Mayer, K., Schmid, T., & Weber, F. (2015). Modeling electricity spot prices: Combining mean reversion, spikes, and stochastic volatility. *The European Journal of Finance*, 21(4), 292–315.
- Mazur, M., Dang, M., & Vega, M. (2020). COVID-19 and the March 2020 stock market crash. Evidence from S&P1500. *Finance Research Letters*, 38, 160190.
- Nguyen, D. K., Huynh, T. L. D., & Nasir, M. A. (2021). Carbon emissions determinants and forecasting: Evidence from G6 countries. *Journal of Environmental Management*, 285, 111988.
- Pandey, D. K., & Kumari, V. (2021). Event study on the reaction of the developed and emerging stock markets to the 2019-nCoV outbreak. *International Review of Economics & Finance*, 71, 467–483.
- Safari, A., & Davallou, M. (2018). Oil price forecasting using a hybrid model. *Energy*, 148, 49–58.
- 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. *Applied Energy*, 231, 1354–1371.
- Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, 9(3), 293–300.
- Vapnik, V. (1999). *The nature of statistical learning theory*. Berlin: Springer.
- Vapnik, V. N., & Lerner, A. Y. (1963). Recognition of patterns with help of generalized portraits. *Avtomat. i Telemekh*, 24(6), 774–780.
- Vautard, R., Yiou, P., & Ghil, M. (1992). Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica d: Nonlinear Phenomena*, 58(1–4), 95–126.
- Veith, S., Werner, J. R., & Zimmermann, J. (2009). Capital market response to emission rights returns: Evidence from the European power sector. *Energy Economics*, 31(4), 605–613.
- Wang, D., Luo, H., Grunder, O., Lin, Y., & Guo, H. (2017). Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. *Applied Energy*, 190, 390–407.
- Wang, H., Chen, Z., Wu, X., & Nie, X. (2019). Can a carbon trading system promote the transformation of a low-carbon economy under the framework of the porter hypothesis? Empirical analysis based on the PSM-DID method. *Energy Policy*, 129, 930–938.
- Wu, Y. X., Wu, Q. B., & Zhu, J. Q. (2019). Improved EEMD-based crude oil price forecasting using LSTM networks. *Physica a: Statistical Mechanics and Its Applications*, 516, 114–124.
- Ye, J., & Xue, M. (2021). Influences of sentiment from news articles on EU carbon prices. *Energy Economics*, 101, 105393.
- Yu, L., Zhao, Y., & Tang, L. (2014). A compressed sensing based AI learning paradigm for crude oil price forecasting. *Energy Economics*, 46, 236–245.
- 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. *Journal of Cleaner Production*, 204, 958–964.
- Zhang, W., & Wu, Z. (2022). Optimal hybrid framework for carbon price forecasting using time series analysis and least squares support vector machine. *Journal of Forecasting*, 41(3), 615–632.
- Zhang, X., Wang, J., & Gao, Y. (2019). A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. *Energy Economics*, 81, 899–913.
- Zhang, Y., & Mao, J. (2022). COVID-19's impact on the spillover effect across the Chinese and U.S. stock markets. *Finance Research Letters*, 47(Part B), 102684.
- Zhao, X., Han, M., Ding, L., & Kang, W. (2018). Usefulness of economic and energy data at different frequencies for carbon price forecasting in the EU ETS. *Applied Energy*, 216, 132–141.

- Zhou, M., Yan, Z., Ni, Y. X., Li, G., & Nie, Y. (2006). Electricity price forecasting with confidence-interval estimation through an extended ARIMA approach. *IEE Proceedings-Generation, Transmission and Distribution*, *153*(2), 187–195.
- Zhu, B. (2012). A novel multiscale ensemble carbon price prediction model integrating empirical mode decomposition, genetic algorithm and artificial neural network. *Energies*, *5*(2), 355–370.
- Zhu, B., Han, D., Wang, P., Wu, Z., Zhang, T., & Wei, Y. M. (2017). Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. *Applied Energy*, *191*, 521–530.
- Zhu, B., Ma, S., Xie, R., Chevallier, J., & Wei, Y. M. (2018a). Hilbert spectra and empirical mode decomposition: A multiscale event analysis method to detect the impact of economic crises on the European carbon market. *Computational Economics*, *52*(1), 105–121.
- Zhu, B., & Wei, Y. (2013). Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. *Omega*, *41*(3), 517–524.
- Zhu, B., Ye, S., Wang, P., He, K., Zhang, T., & Wei, Y. M. (2018b). A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting. *Energy Economics*, *70*, 143–157.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.