CrossMark

A Novel Decomposition-Ensemble Based Carbon Price Forecasting Model Integrated with Local Polynomial Prediction

Quande Qin[1,2] (ID) · Huangda He[1] · Li Li[1] · Ling-Yun He[3]

Abstract

This study proposes a decomposition-ensemble based carbon price forecasting model, which integrates ensemble empirical mode decomposition (EEMD) with local polynomial prediction (LPP). The EEMD method is used to decompose carbon price time series into several components, including some intrinsic mode functions (IMFs) and one residue. Motivated by the fully local characteristics of a time series decomposed by EEMD, we adopt the traditional LPP and regularized LPP (RLPP) to forecast each component. This led to two forecasting models, called the EEMD-LPP and EEMD-RLPP, respectively. Based on the fine-to-coarse reconstruction principle, an auto regressive integrated moving average (ARIMA) approach is used to forecast the high frequency IMFs, and LPP and RLPP is applied to forecast the low frequency IMFs and the residue. The study also proposes two other forecasting models, called the EEMD-ARIMA-LPP and EEMD-ARIMA-RLPP. The empirical study results showed that the EEMD-LPP and EEMD-ARIMA-LPP outperform the two other models. Furthermore, we examine the robustness and effects of parameter settings in the proposed model. Compared with existing state-of-art approaches, the results demonstrate that EEMD-ARIMA-LPP and EEMD-LPP can achieve higher level and directional predictions and higher robustness. The EEMD-LPP and EEMD-ARIMA-LPP are promising approaches for carbon price forecasting.

Keywords Carbon price forecasting · Ensemble empirical mode decomposition · Decomposition-ensemble framework · Local polynomial prediction

✉ Ling-Yun He
lyhe@amss.ac.cn

[1] Department of Management Science, Shenzhen University, Shenzhen 518060, China

[2] Center for Energy & Environmental Policy Research, Beijing Institute of Technology, Beijing 100081, China

[3] Institute of Resource, Environment and Sustainable Development Research, School of Economics, JiNan University, Guangzhou 510632, China

1 Introduction

Global climate change has become a serious global challenge in recent decades. To address this escalating climate change problem, the European Union (EU) launched the EU Emissions Trading Scheme (ETS) in January 2005 (Ellerman and Buchner 2007). China has been the largest carbon emission emitter since 2009, contributing almost 25% of total worldwide emissions (Qin et al. 2017). The ETS has had a global impact on emission reductions, with the goals set out in the Paris Agreement. In December 19, 2017, China formally launched an initial national carbon trading market. The first phase of the market covers power generation; the new cap is approximately twice the size of the EU's carbon market and ten times the size of California's cap and trade system in the United States.[1]

The emergence of ETS has transformed carbon emission permits into a new asset class, called "carbon emission allowances" (Neuhoff et al. 2006; Hintermann 2010). Like other commodities, carbon emission allowances can be bought and sold in the carbon markets. Corporations, organizations, and investors can trade both the spot and future of emission allowances. The establishment of an appropriate ETS has become an important issue for China to allow it to effectively compete in global carbon finance and achieve sustainable development (Lo 2016; Jotzo and Löschel 2014). Accurate carbon price forecasting has been recognized as an effective way to formulate relevant policies for carbon emissions trading and to reduce the risk of carbon assets in the carbon markets (Koop and Tole 2013; Zhu et al. 2015). As such, carbon price forecasting is critically important for governments and traders in the carbon trading market. Carbon price is affected by many factors, including short-term market fluctuations, the effects of significant trend breaks, and long-term trends (Benz and Trück 2009; Wei et al. 2010; Tang et al. 2017; Chen et al. 2013). The carbon price time series is nonlinear, non-stationary, and is highly volatile. This makes accurately forecasting carbon prices a significant challenge (Feng et al. 2011; Zhu and Wei 2013).

Existing carbon price forecasting studies can be divided into two categories. The first category can be interpreted as multivariate models, which model the relationship between carbon price and factors (García-Martos et al. 2013; Chevallier 2009, 2010, 2011a, b; Fezzi and Bunn 2009). The second category is univariate models, also known as the time series forecasting approach. These models generally include two kinds of methods: the traditional econometric method (Byun and Cho 2013; Benz and Trück 2009) and artificial intelligence (AI) method (Zhu and Wei 2013; Fan et al. 2015; Atsalakis 2016). Because it is difficult to determine the appropriate factors affecting carbon price, time series forecasting approach received widely concerns. In the time series forecasting approach, the decomposition-ensemble framework is becoming increasingly popular. This framework first decomposes the complex time series into several components that have simple structures, which are more easily forecasted. These components are individually forecasted and then we ensemble the results to obtain the final forecasting result (Yu et al. 2008; Zhang et al. 2015; Zhu et al. 2016).

---

[1] Data from "http://www.chinadaily.com.cn/a/201803/27/WS5ab9b38ca3105cdcf65148ca.html".

Decomposition-ensemble forecasting models have been shown to perform better than the single econometric and AI model (Yu et al. 2008; Tang et al. 2011). Over the past decades, some decomposition-ensemble approaches have been proposed to forecast the carbon price. Zhu et al. (2016) proposed a novel decomposition-ensemble paradigm, which incorporates least square support vector machine (LSSVM), auto regressive integrated moving average (ARIMA), and particle swarm optimization. Sun et al. (2016) proposed a combined forecasting model based on variational mode decomposition and spiking neural networks to forecast carbon price.

Lately, empirical mode decomposition (EMD), invented by Huang et al. (1998), decomposes a raw non-linear and non-stationary signal into a set of intrinsic mode functions (IMFs) and one residue, without leaving the time domain. EMD is easy to implement and offers the ability to estimate subtle changes in frequency (Huang et al. 2011). The decomposition scale of EMD depends on the characteristics of the signal itself and does not require a priori basic function. EMD is based on local extreme points: cubic spline interpolation, based on the local extreme point signal envelope to strike down the average envelope (Chen et al. 2006). Therefore, EMD is a fully local and self-adaptive approach, and can decompose the local characteristics of the raw time series into several fluctuations and trend items. EMD-based forecasting models have been widely adopted within the forecasting research community (Yu et al. 2008; Xiong et al. 2013; Chen et al. 2012; Ren et al. 2015). Considering the characteristics discomposed by EMD, combining EMD-based decomposition with a local forecasting method, which can capture the local characteristics of a time series, is a novel attempt to improve forecasting performance.

The local polynomial prediction (LPP), developed by Farmer and Sidorowichl (1988), works effectively by analyzing the local characteristics of a time series in the frame of nonlinear dynamics. LPP displays promising forecasting performance with fast speed in low dimensional and smoothing problems (Regonda et al. 2005; Lu 2002; Beran et al. 2002; Su and Li 2015). Similar to AI model, the forecasting process of LPP is also based on the technique of phase space reconstruction (Packard et al. 1980; Takens 1981). LPP does not require a long time series and it can effectively model different scales of nonlinearity in different period of time series because of using the localization method (Farmer and Sidorowich 1988; Fan and Yao 2008). The decomposition feature of EMD motivated us to investigate the forecasting ability of LPP and RLPP approaches under the EMD framework. This study proposes a novel carbon price forecasting model integrated with the promising and relatively simple LPP approach. The ensemble EMD (EEMD), developed by Wu and Huang (2009) to overcome the mode mixing effect in EMD, is used to decompose the original carbon price time series. The LPP approach is used to individually forecast each component. The obtained model is called EEMD-LPP in this study.

It is worth noting that the parameters of the LPP are typically estimated using ordinary least squares (OLS), which can have large variance under certain conditions. Regularization techniques can modify the OLS solution towards better predictions (Kugiumtzis et al. 1998). In this study, a well-known regularization technique, namely ridge regression, is used to estimate the LPP (RLPP) parameters. A new forecasting model, called the EEMD-RLPP, is developed. The EEMD-RLPP model applies the EEMD method as a decomposition tool and uses RLPP as forecasting tool. Selecting

the optimal model for each component can improve forecasting performance in the decomposition-ensemble framework. Some hybrid models have been developed based on this idea (Zhang et al. 2015; Zhu et al. 2016). The empirical results show that hybrid models demonstrate satisfactory performance. We integrate the ARIMA, which has a strong capability of modeling short-term memory and random process, with LPP and RLPP to develop two new models: EEMD-ARIMA-LPP and EEMD-ARIMA-RLPP. Based on the fine-to-coarse reconstruction principle developed by Zhang et al. (2008), the ARIMA approach is used to forecast the high frequency IMFs, whereas LPP and RLPP are used to forecast the low frequency IMFs and the residue.

The features of this study are represented as follows: (1) Motivated by the fully local characteristics of a time series decomposed by EEMD, we first integrate the LPP and RLPP approach with the EEMD method for carbon price forecasting. (2) Based on the fine-to-coarse reconstruction principle, we combined LPP and RLPP with ARIMA to forecast different components according to the components' characteristics. (3) Experimental results demonstrate the effectiveness of the proposed approach for carbon price forecasting. (4) Under the decomposition-ensemble framework, we analyzed the robustness of the related parameters and how the parameters affect the forecasting result in the proposed model.

The rest of this study is organized as follows: Sect. 2 describes the methods, including EEMD, LPP, and ARIMA. Section 3 presents the proposed EEMD-based forecasting models in detail. Section 4 presents the experimental study, mainly including data sources, evaluation criteria, experimental results, and comparisons with state of the art approaches. Concluding remarks are drawn in Sect. 5.

2 Methods

2.1 EEMD

EEMD was developed to overcome the mode mixing problem of EMD through performing the EMD over an ensemble of the data plus Gaussian white noise (Wu and Huang 2009). EEMD is considered to be a significant improvement over the EMD method. EEMD can decompose a complex signal into finite IMFs and one residue that have simple pattern and stationary fluctuation based on the local characteristic time scale of the signal.

For an original time series $y_0(t)$ $(t = 1, 2, \ldots, n)$, where n is the length of the time series. The EEMD procedures are described as follows:

(1) Initialize the number of ensemble (N) and the standard deviation (σ) of the added white noise.
(2) Add N random white noises $\{n_1(t), n_2(t), \ldots, n_N(t)\}$ to the original data and produce a group of new data $\{y_1(t), y_2(t), \ldots, y_N(t)\}$.
(3) Decompose every new data $y_i(t)(i = 1, 2, \ldots, N)$ using EMD to obtain N groups of IMFs components and N groups of residues. The EMD method is described as follows:

1) Identify all the local minima and local maxima of the a time series $X(t)$.

2) Connect all local extrema by the cubic spline interpolation to individually create the upper envelope $e_{up}(t)$ and the lower envelope $e_{low}(t)$.
3) Compute the point-by-point mean value $m_1(t)$ of the upper envelope and the lower envelope: $m_1(t) = [e_{up}(t) + e_{low}(t)]/2$.
4) Calculate the difference between $x(t)$ and $m_1(t)$: $d_1(t) = X(t) - m_1(t)$.
5) Check the property of $d_1(t)$: (1) If $d_1(t)$ satisfies the conditions of an IMF, then an IMF is derived and denoted as $c_1(t)$, while $x(t)$ is replaced with the residue $r_1(t) = x(t) - d_1(t)$; (2) if $d_1(t)$ is not an IMF, replace $x(t)$ with $d(t)$ and repeat the steps 1–4 until $d(t)$ meets the conditions of an IMF.
6) Repeat steps 1)–5) until the residue $r(t)$ cannot further extract an IMF, then the originally data $X(t)$ can be expressed as the sum of the IMFs and the residue:

$$X(t) = \sum_{i=1}^{v} c_i(t) + r(t)$$

where v is the total number of IMFs, $c_i(t)$ is the ith IMF at time t.

(4) Average N groups of IMFs components and residue respectively to obtain the responding mean value $\{C_1, C_2, \ldots, C_n\}$ and Res, and then the original data can be expressed as:

$$y_o(t) = \sum_{i=1}^{v} C_i + Res$$

2.2 LPP and RLPP

2.2.1 Phase Space Reconstruction

Assume an observed time series of $x_1, x_2, \ldots, x_n$. The first step in the local prediction is phase space reconstruction. Based on the approach introduced by Packard et al. (1980) and Takens (1981), the time series can be embedded in a state space by using delay coordinates. The state vector has the following form:

$$X_t = [x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau}]^{\mathrm{T}}$$

In this expression, m is the embedding dimension, τ is the delay time, and T denotes the transpose. According to phase space reconstruction theory, there is a smooth map $f : R^m \to R^m$ between the current state X_t and the future state X_{t+T_1}. For convenience, the m-dimensional states are usually mapped into one dimensional value, such that:

$$x_{t+T_1} = f(X_t)$$

for all integers t and where the forecasting time T_1 and τ are also assumed to be integers.

2.2.2 LPP

Farmer and Sidorowich (1988) first used local approximation to fit the functional relationship between the state vectors and to compare the forecasting capability between different order approximations. The most general form for a dth degree m-dimensional polynomial is as follows:

$$A_d(x_1, \ldots, x_m) = \sum_{i_1=0,\ldots,i_m=0} a_{i_1\ldots i_d} x_1^{i_1} \ldots x_m^{i_m}$$

where $\sum_{j=1}^{m} i_j \leq d$. The empirical results show that the local quadratic polynomial, i.e. the second order approximation, performs better in most cases (Farmer and Sidorowich 1988). This expression is described as:

$$x_{t+T_1} = \sum_{i=1}^{m} a_{0i}(x_{t-(i-1)\tau}) + \sum_{i=1}^{m}\sum_{j=i}^{m} a_{ij} x_{t-(i-1)\tau} x_{t-(j-1)\tau}$$

where T_1 is the forecasting step, m is the embedding dimension, τ is the delay time, and a_{ij} are coefficients. Unlike global approximation methods, such as the artificial neural network (ANN) and support vector machine, which use global data to model the smooth map. LPP estimates the function coefficient by partitioning the embedding space into the k nearest neighbors $\{X_r\}$ ($r = 1, 2, \ldots, k$). These neighbors are identified by computing the distance between the current state vector X_r and its preceding delay vectors with an imposed metric $\|\cdot\|$ (i.e. the k state vector $X_{r'}$ with an $r' < r$ that minimizes $\|X_r - X_{r'}\|$), and using OLS to linearly fit a LPP model and predict the future value. The key of local approximation is to correctly select the local neighborhood size; this provides sufficient points to make the local parameter fitting stable while adding more points would not lead to significant improvements. To ensure solution stability, the number of neighbors k is consistently larger than embedding dimension m (Farmer and Sidorowich 1987).

2.2.3 RLPP

A typical and well-known regularization technique is ridge regression, proposed by Hoerl and Kennard (1970). This technique balances the bias and the variance of the regression by introducing a regularization term based on least squares, as follows:

$$\sum_{i=1}^{n}\left(y_i - \sum_{j=0}^{m} \omega_j x_{ij}\right)^2 + \lambda \sum_{j=0}^{m} \omega_j^2, \ \lambda > 0$$

In this expression, ω_j is the coefficient and λ is the penalty parameter, which plays a vital role in controlling the bias of the regression, while being able to be specified by users. Here, we combine the LPP with the ridge regression method, i.e. using the ridge regression to fit the LPP to develop a regularized LPP (RLPP), defined as:

$$x_{t+T_1} = \sum_{i=1}^{m} a_{0i}(x_{t-(i-1)\tau}) + \sum_{i=1}^{m}\sum_{j=i}^{m} a_{ij}x_{t-(i-1)\tau}x_{t-(j-1)\tau} + \lambda\left(\sum_{i=1}^{m}\sum_{j=i}^{m} a_{ij}^2 + \sum_{i=1}^{m} a_{0i}^2\right), \; \lambda > 0$$

The penalty parameter λ can be determined by trial and error. The coefficients $a_{ij}(i = 0, \ldots, m, j = 1, \ldots, m)$ can be estimated by nearest neighbor points.

2.2.4 Procedures of LPP and RLPP

For a time series $x_1, x_2, \ldots, x_n$, to predict the future value x_{n+T_1}, the LPP and RLPP procedures are as follows:

(1) Determine the embedding dimension m, delay time τ, and the nearest neighbors number k (and the penalty parameter λ when using RLPP).
(2) Calculate the distance of delay vector X_n from other vectors X_j, $1 + (m - 1)\tau \leq j \leq n - T_1$, in the state space. For computational efficiency, we use the maximum norm (i.e. $\|A - B\| = \|a_1 - b_1, a_2 - b_2, \ldots, a_m - b_m\| = \max(|a_1 - b_1|, |a_2 - b_2|, \ldots, |a_m - b_m|)$) to measure the distance between state vectors.
(3) Rank the distance d_j, identify the k nearest neighbors $X_{j1}, X_{j2}, \ldots, X_{jk}$, and fit a model of the form:

$$LPP: \; x_{t+T_1} = \sum_{i=1}^{m} a_{0i}(x_{t-(i-1)\tau}) + \sum_{i=1}^{m}\sum_{j=i}^{m} a_{ij}x_{t-(i-1)\tau}x_{t-(j-1)\tau}$$

$$RLPP: \; x_{t+T_1} = \sum_{i=1}^{m} a_{0i}(x_{t-(i-1)\tau}) + \sum_{i=1}^{m}\sum_{j=i}^{m} a_{ij}x_{t-(i-1)\tau}x_{t-(j-1)\tau}$$
$$+ \lambda\left(\sum_{i=1}^{m}\sum_{j=i}^{m} a_{ij}^2 + \sum_{i=1}^{m} a_{0i}^2\right), \; \lambda$$
$$> 0$$

In these expressions, the parameters $a_{ij}(i = 0, \ldots, m, j = 1, \ldots, m)$ of LPP are computed using the ordinary least squares. The RLPP is fitted using ridge regression.

(4) Use the fitted model to estimate a T_1 step ahead forecast $\hat{x}_{n+T_1}$ for the vector X_n.

2.3 ARIMA

Theoretically, ARIMA is the most general class of models for forecasting a time series. The ARIMA is a generalization of an autoregressive moving average (ARMA) (Box et al. 2015). ARMA is generally applied to stationary series. ARIMA can be applied in non-stationarity time series using an initial differencing, which is used to eliminate the non-stationarity. An ARIMA (p, d, q) typically consists of three parts:

autoregression (order p), difference (order d), and moving average MA (order q). This can be expressed as:

$$\nabla^d x_t = c + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots \varphi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

In this expression, $\nabla = (1 - B)$, B refers to the backward shift operator for $B(x_t) = x_{t-1}$; x_t is the observation data at time t; c is the constant; p and q are the numbers of autoregressive and moving average terms in the ARIMA, respectively; $\varphi_1, \varphi_2, \ldots, \varphi_p$ and $\theta_1, \theta_2, \ldots, \theta_q$ are the autoregressive and the moving average parameters to be estimated, respectively; ε_t is the random error at time t; and $\varepsilon_t \sim N(0, \sigma^2)$.

3 Proposed EEMD-Based Carbon Price Forecasting Model

There are three main steps in a typical decomposition-ensemble model: (1) decomposition of the original complicated series; (2) individual prediction for each component; and (3) aggregating the forecasting result of all components to generate a final forecasting result of the original time series. Based on the general framework of the "Divide-and-Conquer" principle, we developed four different hybrid models based on the EEMD method to forecast the nonlinear, non-stationary carbon price time series.

3.1 Model 1: EEMD-LPP

The EEMD-LPP forecasting model generally consists of the following three steps:

Step 1 The original carbon price time series is first decomposed by EEMD method into *n* IMFs and one residue series, denoted as Res.

Step 2 The LPP is used to model the extracted IMF components and residue, and to individually forecast each component.

Step 3 Simple addition has been shown to be an effective aggregation method (Tang et al. 2012). As such, the forecasting result of all extracted IMF components and the residue are added, to generate a final forecasting result of the original time series.

3.2 Model 2: EEMD-RLPP

The main step of EEMD-RLPP model is like the EEMD-LPP, except for in Step 2, where the RLPP is used as an individual forecasting tool to forecast the extracted IMF components and residue, instead of using LPP.

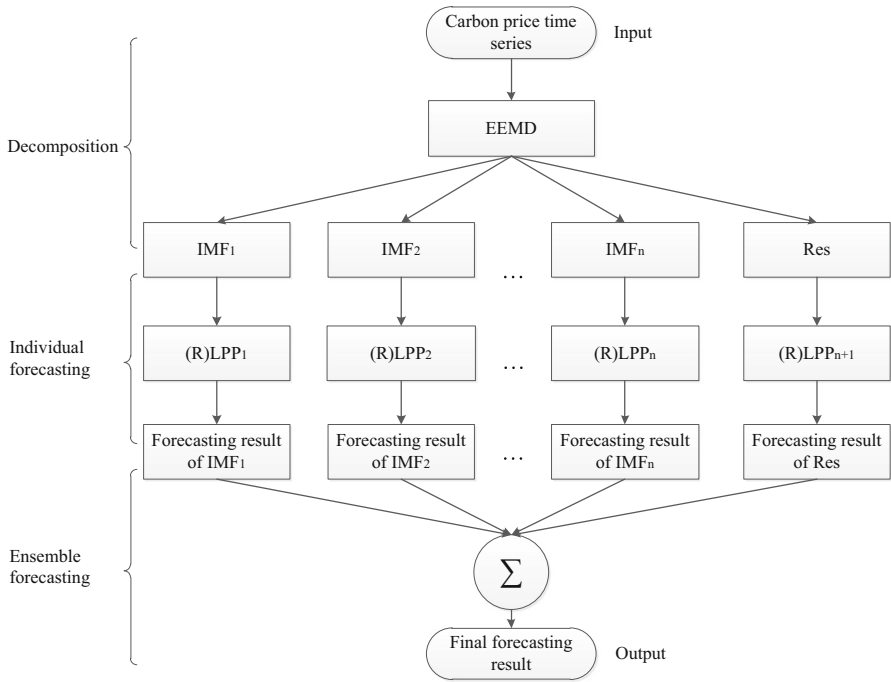Figure 1 illustrates the methodological procedures associated with Models 1 and 2.

Fig. 1 The methodological procedures of the proposed EEDM-(R)LPP model

3.3 Model 3: EEMD-ARIMA-LPP

Unlike EEMD-LPP and EEMD-RLPP, EEMD-ARIMA-LPP incorporates the ARIMA and the LPP to forecast different components, based on their data characteristics. The IMFs can be divided into two parts based on their frequency (Zhang et al. 2008): (1) High frequency parts (HFs), which are characterized by highly fluctuations and low amplitude, reflecting the information of normal market fluctuations which have a short-term impact on the carbon price; (2) Low frequency parts (LFs), characterized by low fluctuations, expressing the effects of significant trend breaks, with a medium-term impact on carbon price. The residue drives the major long-term trends in carbon price. Here, we use the fine-to-coarse reconstruction algorithm proposed by Zhang et al. (2008) to identify the HFs and LFs of the carbon price to determine the forecasting model for each component. The identification process is as follows:

(1) Compute the mean value $\bar{s}_i$ of the cumulative sum series $s_i = \sum_{p=1}^{i} IMF_i$ ($i = 1, 2, \ldots, n$) from IMF_1 to IMF_i, where n is the total number of IMFs.

(2) Apply a t test to determine from which IMF_i the mean value $\bar{s}_i$ most significantly departs from zero. The significance level α is generally selected as 0.05.

(3) If $\bar{s}_i$ start to deviate from zero, then IMF_1 to IMF_i are identified as HFs, and the remaining $IMFs$ are identified as LFs

The ARIMA model has a strong ability to model a short-term memory and random process. This makes it suitable for forecasting HFs. The LPP is used to forecast the LF components and the residue. The procedures of the EEMD-ARIMA-LPP forecasting model can be summarized as follows:

Step 1 Decompose the original carbon price time series using EEMD.

Step 2 Identify the HFs and LFs of the extracted IMF components.

Step 3 Use the ARIMA to forecast each component of HFs and use the LPP to forecast each LF component and the residue.

Step 4 Add the forecasting result of all IMF components and the residue to obtain the final forecasting result of the original time series.

3.4 Model 4: EEMD-ARIMA-RLPP

The main step of EEMD-RLPP model is like the step of the EEMD-ARIMA-LPP model, except for Step 3, where the RLPP is used to forecast each LF component and the residue instead of LPP.

Figure 2 summarizes the procedures associated with proposed Models 3 and 4.

4 Experimental Study

4.1 Data

The European Climate Exchange (ECX) is the largest carbon market under the EU ETS. The ECX includes spot, futures, and options of the EU allowance (EUA) and Certified Emission Reduction (CER), which has the maximum trading volume of EUA. In this study, two daily carbon EUA future prices with maturity dates in December 2014 (Dec 14) and December 2015 (Dec 15) form the ECX. These were selected as the experimental data, and are freely available from the Intercontinental Exchange (ICE) website (http://www.theice.com). For Dec 14, the daily trading data cover the period from September 28, 2010 to December 15, 2014, excluding public holidays, for a total of 1079 observations. For Dec 15, the trading data is from November 29, 2011 to December 14, 2015, excluding public holidays, for a total of 1035 observations.

Figure 3 shows the time series curves of Dec 14 and Dec 15, and demonstrates that carbon prices have highly uncertain, nonlinear, and complicated characteristics. For the convenience of LPP and RLPP modeling and optimal parameter searches, the experimental data are divided into three subsets: the training set, the validation set, and the testing set (70%, 15% and 15% respectively). The training set and the validation set are used to find the best parameters using trial and error, where $\lambda \in (0, 1] m \in [1, 10]$, $\tau \in [1, 4]$ and $k \in [2m + 1, L]$, m, τ, k are integers. In contrast, λ is presented in terms of decimals and L denotes the length of training set after phase space reconstruction. Then, the testing set is used to verify the validity of the newly proposed models using the previously identified parameters. Table 1 reports the divided sample of Dec 14 and Dec 15.
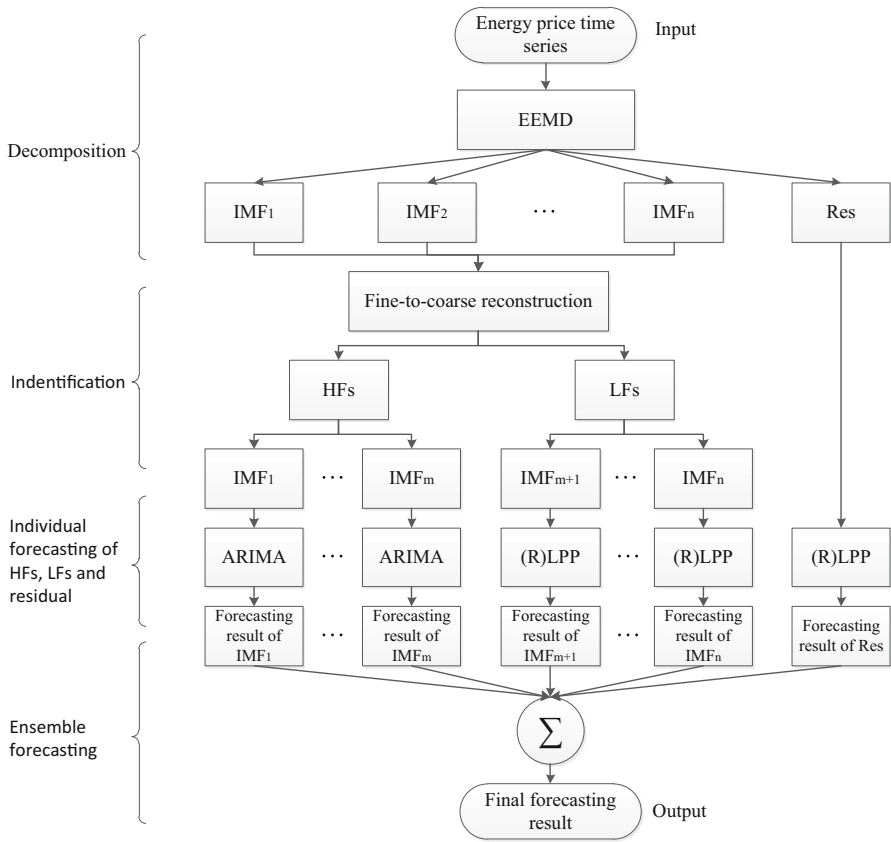
Fig. 2 The methodological procedures of the proposed EEDM-ARIMA-(R)LPP model

4.2 Evaluation Criteria

To measure the forecasting ability of the hybrid models, two main criteria are used to evaluate the level forecasting and directional forecasting. First, the root mean squared error (RMSE) is selected as the evaluation criterion of level forecasting, defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left[\hat{x}(t) - x(t)\right]^2}$$

where $x(t)$ is the actual value, $\hat{x}(t)$ is the predicted value, and n is the number of predictions. Smaller RMSE values are associated with a more accurate model.

Second, the directional prediction statistic (D_{stat}) is used to measure the directional forecasting performance, expressed as:

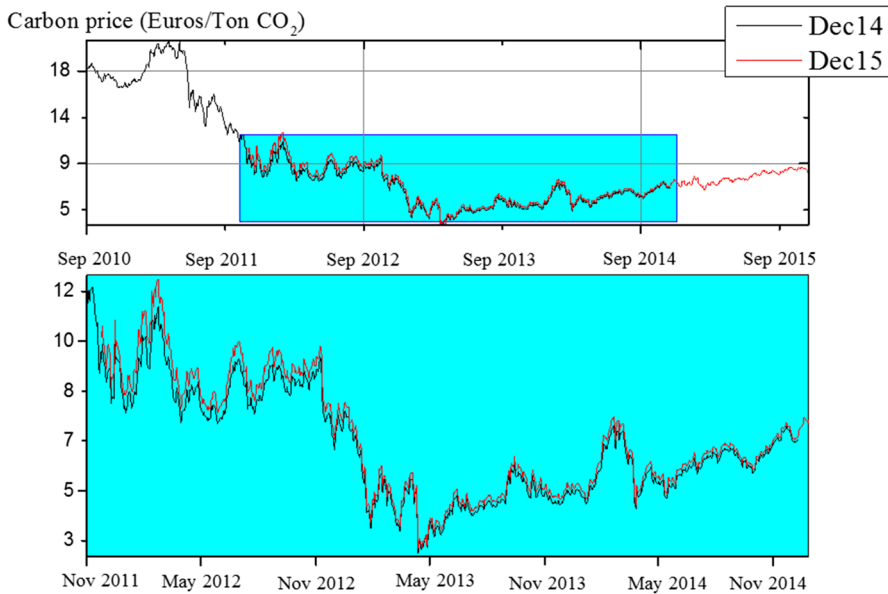$$D_{stat} = \frac{1}{n} \sum_{t=1}^{n} a_t * 100\%$$

Fig. 3 Dec 14 and Dec 15 carbon future price time series

In this expression, $a_t = 1$ if $[x(t + 1) - x(t)] * [\hat{x}(t + 1) - x(t)] \geq 0$; otherwise, $a_t = 0$. The greater the value of D_{stat}, the more accurate the prediction is.

To further compare the predictive accuracy of different models from a statistical perspective, the Diebold–Mariano (DM) statistic is introduced to test the statistical significance of the forecasting difference (Diebold and Mariano 2002). Assuming e_t^1 and e_t^2 are two sets of forecast errors, generated by two different forecasting models respectively, the DM test is based on the loss differential $d_t = e_t^1 - e_t^2$. In this study, the loss function is set to mean the absolute percentage error (MAPE), defined as:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{\hat{x}(t) - x(t)}{x(t)} \right|$$

4.3 Decomposing Carbon Price Using EEMD and Identifying HFs and LFs

Two important EEMD parameters must be established before it can be used: the number of the ensemble N and the amplitude of white noise σ. Referring to previous studies, this study set N as 100 and σ as 0.1 times the standard deviation of each series (Zhang et al. 2008). Figure 4 shows the decomposition result of Dec 14 and Dec 15.

Figure 4 shows the extracted IMFs for Dec 14 and Dec 15 experienced changing frequency and amplitudes. These IMFs were arranged from high frequency to low frequency. The fine-to-coarse reconstruction algorithm was applied to identify HFs and LFs, with a significance level set at 0.05. Table 2 presents the results, which show that the $\overline{S_i}$ of Dec 14 started to significantly deviate from 0 at the point $i = 6$. Therefore

Table 1 Samples of carbon future prices

Carbon future price	Type	Size	Date
Dec 14	Sample set	1079	September 28, 2010–December 15, 2014
	Training set	757	September 28, 2010–September 12,2013
	Validation set	161	September 12,2013–May 2,2014
	Test set	161	May 5,2014–December 15, 2014
Dec 15	Sample set	1035	November 29, 2011–December 14, 2015
	Training set	725	November 29, 2011–September 29,2014
	Validation set	155	September 30,2014–May 11,2015
	Test set	155	May 12,2015–December 14, 2015

IMF1 to IMF5 belonged to HFs, whereas IMF6 to IMF9 were associated with LFs. For Dec 15, the $\overline{S_i}$ started to significantly deviate from 0 at the point $i = 4$, illustrating that IMF1 to IMF3 were HFs, and the IMF4–IMF 9 were LFs.

4.4 Experimental Results and Analysis

To study the forecasting abilities of the proposed models with other widely used forecasting models, the EEMD-LSSVM, EEMD-ARIMA, and EEMD-ANN served as benchmark models. For LSSVM, the Gaussian RBF kernel function was chosen; the values of the parameters sigma squared and gamma were set using the grid-search method (Fan et al. 2016). For ANN, the hidden nodes were set to 7, and the number of input neurons was determined using autocorrelation and partial correlation analyses (Yu et al. 2016). For ARIMA, the parameters were selected based on the Akaike's information criterion (AIC) (Wang et al. 2015).

All components used one-step-ahead forecasting; results were aggregated as the final forecasted value of the original series. The ARIMA method was realized using EViews 8.0 statistical software; LSSVM and ANN were implemented through programming using the Matlab R2012b platform; and the LPP and RLPP were implemented through programming using the Python 2.7 platform. Table 3 illustrates the optimal parameters of the proposed four models for the two carbon future prices fore-
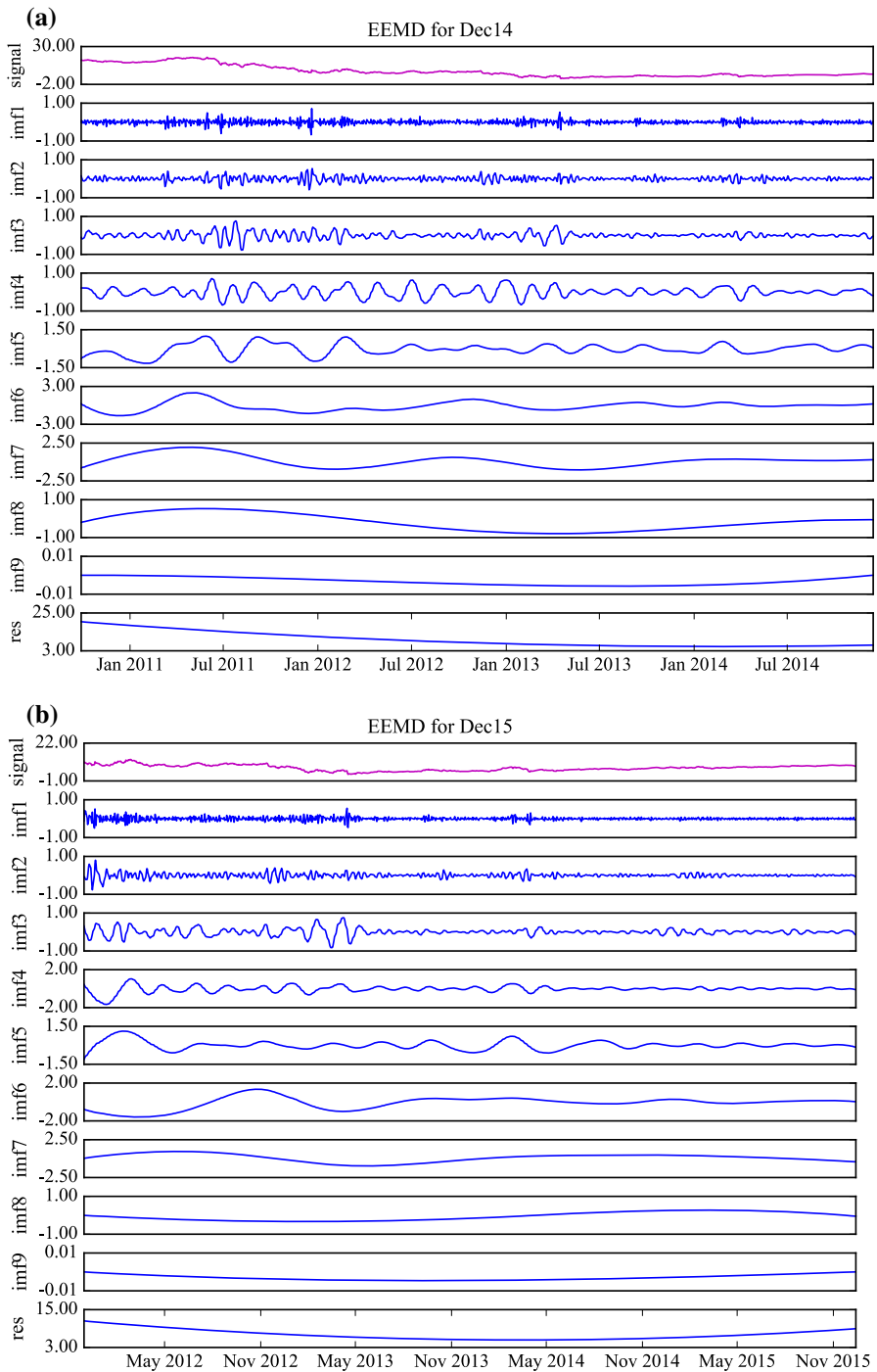
Fig. 4 Decomposition result of carbon price using EEMD **a** Dec 14. **b** Dec 15

Table 2 Identification results of HFs and LF

Carbon price	Parameters	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9
Dec 14	$\bar{s}_i$	− 0.001	− 0.001	0.001	0.005	− 0.031	**− 0.155**	− 0.016	− 0.208	− 0.211
	t	− 0.376	− 0.213	0.131	0.384	− 1.528	**− 4.601**	− 0.334	− 3.985	− 4.044
	p value	0.707	0.832	0.896	0.701	0.127	**0.000**	0.739	0.000	0.000
Dec 15	$\bar{s}_i$	− 0.001	− 0.001	0.001	**− 0.030**	− 0.012	− 0.010	− 0.009	− 0.069	− 0.072
	t	− 0.205	− 0.145	0.164	**− 2.024**	− 0.702	− 4.512	− 0.291	− 2.085	− 2.176
	p value	0.838	0.885	0.870	**0.043**	0.483	0.000	0.771	0.037	0.030

The significance level is 0.05

casting. Tables 4 and 5 compare model performances in terms of RMSE and D_{stat}; Table 6 reports the results of the DM test.

In terms of level accuracy measured by *RMSE*, EEMD-LPP achieved the best level accuracy on Dec 14; EEMD-LPP-ARIMA performed better on Dec 15. In all, EEMD-LPP and EEMD-ARIMA-LPP outperformed the other five models. Under the decomposition-ensemble framework, LPP generated better forecasts than any other single model when EEMD was used as the decomposition method. Applying ARIMA to forecast HFs may not necessarily improve the forecasting result; this is because LPP may also perform well on HFs. EEMD-RLPP and EEMD-ARIMA-RLPP showed poor level accuracy on Dec 14 and Dec 15. This may be mainly because the components decomposed by EEMD have a simple pattern, stationary fluctuation, and less noisy data. A regularization technique introduces more bias in the forecasting process, increasing the forecasting error.

Regarding the direction forecasting accuracy measured by D_{stat}, Table 5 shows that the EEMD-ARIMA-LPP model performed best on Dec 14, followed by EEMD-LPP. On Dec 15, these two models performed the best. Of the other models, EEMD-ARIMA achieved a higher value of D_{stat}. Compared with level accuracy, EEMD-RLPP and EEMD-ARIMA-RLPP performed satisfactory on direction accuracy. This implies that level accuracy is not necessarily endowed with high directional prediction accuracy. In all, EEMD-LPP and EEMD-ARIMA-LPP were the most accurate and also achieved the highest directional accuracy. This indicates the two models have similar powerful forecasting capability.

Based on the DM test result shown in Table 6, at a 5% significance level, the EEMD-LPP-ARIMA and EEMD-LPP model performed better than other comparative models on both Dec 14 and Dec 15, except EEMD-LPP-ARIMA outperforms EEMD-LSSVM on Dec 14 at a 10% significance level. Additionally, there is no statistically significant difference between EEMD-LPP and EEMD-ARIMA-LPP. EEMD-RLPP and EEMD-ARIMA-RLPP were inferior to EEMD-LPP and EEMD-ARIMA-LPP, and were also inferior to EEMD-ARIMA and EEMD-LSSVM on Dec 15 at a confidence level of 99%. This illustrates a less than satisfying performance by the proposed EEMD-RLPP and EEMD-ARIMA-RLPP model. The introduction of regularization techniques did not improve the forecasting ability of LPP under the EEMD method.

The forecasting results show that among the proposed four models, EEMD-LPP and EEMD-ARIMA-LPP generated the most accurate forecasts. These two models achieved similar forecasting performance, indicating that LPP performs well when forecasting HFs and LFs. EEMD-RLPP and EEMD-RLPP-ARIMA did not perform effectively, and were inferior to the benchmark EEMD-LSSVM, EEMD-ARIMA model.

The components decomposed using the EEMD were endowed with a simpler pattern and more stable fluctuations. After the phase space reconstruction, LPP more precisely approximated the hidden smooth map between the state vectors compared to the complex AI techniques. In this way, EEMD-LPP model could achieve better forecasting result. Compared to the traditional ARIMA method, LPP modeled different scales of nonlinearity in different period because of localization, leading to better performance when forecasting LFs and residue. Introducing regularization techniques to estimate LPP parameters introduces more bias, lowering the forecasting accuracy.

Table 3 Optimal parameters of proposed four models for the two carbon future prices

Carbon price	Parameters	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	Res
Dec 14	m (LPP)	5	8	8	5	6	7	4	6	8	4
	τ (LPP)	1	1	1	1	1	1	1	3	2	2
	k (LPP)	566	750	566	381	483	598	195	361	705	204
	m (RLPP)	7	5	4	8	4	4	4	8	4	7
	τ (RLPP)	1	1	1	1	1	2	3	1	1	1
	k (RLPP)	383	11	9	17	9	9	9	740	381	15
	λ	0.43	0.12	0.09	0.12	0.13	0.10	0.12	0.48	0.61	0.41
	p	3	4	4	4	4					
	d	0	0	0	0	0					
	q	2	4	4	2	3					
Dec 15	m (LPP)	5	6	8	6	6	8	8	4	7	8
	τ (LPP)	1	1	1	1	1	1	3	1	3	1
	k (LPP)	366	544	719	241	335	468	507	343	411	743
	m (RLPP)	5	5	6	5	5	5	8	4	8	4
	τ (RLPP)	1	3	2	1	1	1	1	1	2	1
	k (RLPP)	188	11	13	11	11	11	17	9	543	9
	λ	0.14	0.12	0.11	0.14	0.17	0.12	0.21	0.89	0.32	0.09
	p	2	4	4							
	d	0	0	0							
	q	2	4	4							

Table 4 RMSE comparisons for different models

Model	Dec 14		Dec 15	
	RMSE	Rank	RMSE	Rank
EEMD-LPP	0.047	1	0.027	2
EEMD-RLPP	0.061	5	0.042	5
EEMD-ARIMA-LPP	0.055	2	0.025	1
EEMD-ARIMA-RLPP	0.061	6	0.047	6
EEMD-LSSVM	0.060	4	0.032	3
EEMD-ARIMA	0.058	3	0.033	4
EEMD-ANN	0.079	7	0.063	7

Table 5 D_{stat} comparisons for different models

Model	Dec 14		Dec 15	
	D_{stat}	Rank	D_{stat}	Rank
EEMD-LPP	0.876	2	0.884	1
EEMD-RLPP	0.807	6	0.851	4
EEMD-ARIMA-LPP	0.888	1	0.884	1
EEMD-ARIMA-RLPP	0.857	4	0.800	6
EEMD-LSSVM	0.832	5	0.838	5
EEMD-ARIMA	0.875	3	0.877	3
EEMD-ANN	0.708	7	0.652	7

Thus, RLPP may not effectively forecast the components decomposed by EEMD. Figure 5 shows the difference between the actual values and each one-step ahead predicted values. The predicted values were generally close to their actual values, further demonstrating the forecasting ability of EEMD-LPP and EEMD-ARIMA-LPP model.

4.5 Robustness of Parameters

The experimental results above show there are different parameter settings for different components in the LPP method. The optimal parameter settings of each component in LPP are obtained using trial and error, creating challenges for the real-world applications of the proposed model. As such, we focused on finding a robust and generalized parameter combination for all components and investigated how the parameters of the LPP method affect the forecasting result.

Here, all components are assigned the same parameters i.e. same embedding dimension m, delay time τ, and nearest neighbors k to create individual forecasts. To examine the generalizability of the experimental parameter combination, an orthogonal design (OD) is used to study the effect of multi-factors and multi-level problems (Fang and Wang 1994). Here, we applied the $L_9(3^4)$ orthogonal array for the experiment design;

Table 6 DM test results for different models

| Carbon price	Target model	Benchmark model					
		EEMD-ANN	EEMD-ARIMA	EEMD-LSSVM	EEMD-RLPP	EEMD-RLPP-ARIMA	EEMD-LPP
Dec 14	EEMD-ARIMA	−4.741 (0.000)					
	EEMD-LSSVM	−4.812 (0.000)	0.710 (0.478)				
	EEMD-RLPP	−4.450 (0.000)	1.157 (0.247)	0.184 (0.854)			
	EEMD-RLPP-ARIMA	−4.013 (0.004)	3.744 (0.000)	0.397 (0.691)	0.280 (0.780)		
	EEMD-LPP	−6.998 (0.000)	−3.119 (0.002)	−3.394 (0.000)	−3.935 (0.000)	−4.211 (0.000)	
	EEMD-LPP-ARIMA	−5.366 (0.000)	−2.076 (0.038)	−1.809 (0.071)	−2.086 (0.037)	−3.209 (0.001)	1.768 (0.077)
Dec 15	EEMD-ARIMA	−8.680 (0.000)					
	EEMD-LSSVM	−8.384 (0.000)	1.605 (0.109)				
	EEMD-RLPP	−4.118 (0.000)	4.466 (0.000)	4.088 (0.000)			
	EEMD-RLPP-ARIMA	−2.901 (0.004)	6.148 (0.000)	5.659 (0.000)	2.166 (0.030)		
	EEMD-LPP	−9.419 (0.000)	−2.045 (0.041)	−2.684 (0.007)	−5.942 (0.000)	−6.595 (0.000)	
	EEMD-LPP-ARIMA	−9.873 (0.000)	−3.836 (0.000)	−4.749 (0.000)	−7.558 (0.000)	−8.384 (0.000)	−2.390 (0.017)

(a)

(b)
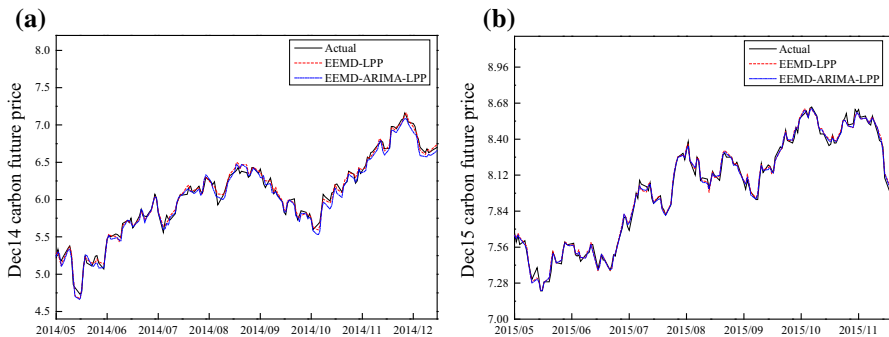
Fig. 5 Difference between the actual values and each one-step ahead predicted values **a** Dec 14. **b** Dec 15

Table 7 Parameter combinations of $L_9(3^4)$

Combinations	Factors		
	m	τ	k
C1	2	1	$0.4n$
C2	2	2	$0.5n$
C3	2	3	$0.6n$
C4	4	1	$0.5n$
C5	4	2	$0.6n$
C6	4	3	$0.4n$
C7	6	1	$0.6n$
C8	6	2	$0.4n$
C9	6	3	$0.5n$

the levels of the three parameters were as follows: $m \in \{2, 4, 6\}$, $\tau \in \{1, 2, 3\}$, $k \in \{0.4n, 0.5n, 0.6n\}$, where n denotes the length of the training data. Details about the orthogonal array are provided in Fang and Wang (1994) and Qin et al. (2015). Table 7 shows the $L_9(3^4)$ orthogonal array; Table 8 shows the experimental results of these parameter combinations for the two series. Table 9 illustrates the range analysis of the orthogonal experiment in terms of the average rank for both series.

Tables 7, 8, and 9 show that the combination ($m=4$, $\tau=1$ and $k=0.4n$) achieves the best average result of the two series. Furthermore, we compared the experimental results obtained by the combination with the parameters obtained using trial and error. Table 10 shows the results; column "T&E" in this table denotes the parameter combination selected by trial-and-error. The differences between the two combinations of the three parameters are extremely small. The optimal combination ($m=4$, $\tau=1$ and $k=0.4n$) is slightly better than the combination obtained using trial and error method in some cases. There may be an over-fitting phenomenon when using trial and error to obtain the best parameters. Thus, the proposed EEMD-LPP mode is shown to be robust for the three parameters.

Table 8 Experiment results in terms of RMSE for the two series

Combinations	Dec 14	Rank	Dec 15	Rank	Average rank
C1	0.047	1	0.028	3	2.0
C2	0.059	6	0.042	6	6.0
C3	0.072	8	0.046	8	8.0
C4	0.052	2	0.026	1	1.5
C5	0.055	4	0.041	5	4.5
C6	0.071	7	0.046	8	7.5
C7	0.055	3	0.027	2	2.5
C8	0.057	5	0.040	4	4.5
C9	0.072	8	0.045	7	7.5

Table 9 Range analysis of the orthogonal experiment in terms of the average rank

Combinations	Factors			Average ranking
	m	τ	k	
C1	2	1	$0.4n$	2.0
C2	2	2	$0.5n$	6.0
C3	2	3	$0.6n$	8.0
C4	4	1	$0.5n$	1.5
C5	4	2	$0.6n$	4.5
C6	4	3	$0.4n$	7.5
C7	6	1	$0.6n$	2.5
C8	6	2	$0.4n$	4.5
C9	6	3	$0.5n$	7.5
K1	16	6	14	$\sum = 44$
K2	14	15	15	
K3	15	23	15	
Level 1	5.3	2.0	4.7	
Level 2	4.5	5.0	5.0	
Level 3	4.8	7.7	5.0	
Optimal level	4	1	$0.4n$	
Order of effect	$\tau\ m\ k$			

4.6 Comparisons of the State of Art Approaches

In order to further investigate and demonstrate the superiority of the proposed EEMD-LPP and EEMD-ARIMA-LPP model, we compare the forecasting accuracy of the proposed models with the state of the art carbon price forecasting models. Based on data availability, we compared the proposed models and a newly adaptive multi-scale ensemble learning paradigm proposed by Zhu et al. (2016). For a fair comparison, we

Table 10 Differences in RMSE and D_{stat} between the two combinations of the three parameters

Time series	RMSE			D_{stat}		
	T&E	$m=4,\ \tau=1,$ $k=0.4n$	Difference	T&E	$m=4,\ \tau=1,$ $k=0.4n$	Difference
Dec 14	0.047	0.052	-0.005	0.876	0.888	-0.008
Dec 15	0.027	0.026	0.001	0.884	0.897	-0.013

Table 11 Comparison results

Model	RMSE	D_{stat}
EEMD-LPP	0.034	0.925
EEMD-ARIMA-LPP	0.032	0.931
EEMD-HLT-Σ	0.047	0.919
EEMD-LSSVM	0.052	0.906
EEMD-ARIMA	0.052	0.894
LSSVM	0.105	0.669
ARIMA	0.111	0.650

used the same historical observations for training and testing to obtain an objective comparison result and the LPP parameters are selected as the experiential parameter mentioned above. The comparison results are illustrated in Table 11.

Table 11 shows that EEMD-LPP and EEMD-ARIMA-LPP are better than the EEMD-HLT-Σ model with respect to both level and directional forecasting. This further demonstrates the superiority of the proposed EEMD-LPP and EEMD-ARIMA-LPP and the effectiveness of the experiential parameter combination.

5 Concluding Remarks

To improve the forecasting performance of carbon price time series, which have highly non-linear and non-stationary characteristics, we proposed four hybrid forecasting models based on the principle of "Divide-and-Conquer." The proposed models use the popular and effective EEMD method as decomposition tool. To investigate the forecasting capability of these newly proposed models, three widely used forecasting models were used as benchmarks: EEMD-SVR, EEMD-ARIMA, and EEMD-ANN. The empirical study shows that EEMD-LPP and EEMD-ARIMA-LPP model statistically outperform other models with higher forecasting accuracy with respect to both level and directional forecasting. The EEMD-LPP and EEMD-ARIMA-LPP are promising to forecast carbon prices; in contrast, EEMD-RLPP and EEMD-RLPP-ARIMA do not perform effectively. This may mainly be because the components decomposed using the EEMD method have stationary fluctuations and less noisy data. Thus, introducing regularization techniques to estimate LPP parameters may intro-

duce more bias and lower the forecasting accuracy. Additionally, several experiments were conducted to analyze the robustness of the parameters in LPP. We further compared the proposed EEMD-LPP and EEMD-ARIMA-LPP model with state-of-the-art approaches using the same historical observations. The experimental results further demonstrated the effectiveness of the proposed EEMD-LPP and EEMD-ARIMA-LPP models. In our future work, we are interested in extending the proposed approach to a multi-step-ahead model.

References

Atsalakis, G. S. (2016). Using computational intelligence to forecast carbon prices. *Applied Soft Computing, 43,* 107–116.

Benz, E., & Trück, S. (2009). Modeling the price dynamics of CO$_2$ emission allowances. *Energy Economics, 31*(1), 4–15.

Beran, J., Feng, Y., Ghosh, S., & Sibbertsen, P. (2002). On robust local polynomial estimation with long-memory errors. *International Journal of Forecasting, 18*(2), 227–241.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. Hoboken: Wiley.

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

Chen, Q., Huang, N., Riemenschneider, S., & Xu, Y. (2006). A B-spline approach for empirical mode decompositions. *Advances in Computational Mathematics, 24*(1), 171–195.

Chen, C. F., Lai, M. C., & Yeh, C. C. (2012). Forecasting tourism demand based on empirical mode decomposition and neural network. *Knowledge-Based Systems, 26,* 281–287.

Chen, X., Wang, Z., & Wu, D. D. (2013). Modeling the price mechanism of carbon emission exchange in the European Union emission trading system. *Human and Ecological Risk Assessment: An International Journal, 19*(5), 1309–1323.

Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics, 31*(4), 614–625.

Chevallier, J. (2010). Volatility forecasting of carbon prices using factor models. *Economics Bulletin, 30*(2), 1642–1660.

Chevallier, J. (2011a). A model of carbon price interactions with macroeconomic and energy dynamics. *Energy Economics, 33*(6), 1295–1312.

Chevallier, J. (2011b). Evaluating the carbon-macroeconomy relationship: Evidence from threshold vector error-correction and Markov-switching VAR models. *Economic Modelling, 28*(6), 2634–2656.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics, 20*(1), 134–144.

Ellerman, A. D., & Buchner, B. K. (2007). The European Union emissions trading scheme: Origins, allocation, and early results. *Review of Environmental Economics and Policy, 1*(1), 66–87.

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

Fan, L., Pan, S., Li, Z., & Li, H. (2016). An ICA-based support vector regression scheme for forecasting crude oil prices. *Technological Forecasting and Social Change, 112,* 245–253.

Fan, J., & Yao, Q. (2008). *Nonlinear time series: Nonparametric and parametric methods*. Berlin: Springer.

Fang, K.-T., & Wang, Y. (1994). *Number-theoretic methods in statistics*. New York: Chapman & Hall.

Farmer, J. D., & Sidorowich, J. J. (1987). Predicting chaotic time series. *Physical Review Letters, 59*(8), 845.

Farmer, J. D., & Sidorowichl, J. J. (1988). Exploiting chaos to predict the future and reduce noise. In *Evolution, learning and cognition* (pp. 277–330). Singapore: World Scientific Publishing Co Pte Ltd.

Feng, Z. H., Zou, L. L., & Wei, Y. M. (2011). Carbon price volatility: Evidence from EU ETS. *Applied Energy, 88*(3), 590–598.

Fezzi, C., & Bunn, D. W. (2009). Structural interactions of European carbon trading and energy prices. *The Journal of Energy Markets, 2*(4), 53–69.

García-Martos, C., Rodríguez, J., & Sánchez, M. J. (2013). Modelling and forecasting fossil fuels, CO_2 and electricity prices and their volatilities. *Applied Energy, 101,* 363–375.

Hintermann, B. (2010). Allowance price drivers in the first phase of the EU ETS. *Journal of Environmental Economics and Management, 59*(1), 43–56.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67.

Huang, N. E., Chen, X., Lo, M. T., & Wu, Z. (2011). On Hilbert spectral representation: A true time-frequency representation for nonlinear and nonstationary data. *Advances in Adaptive Data Analysis, 3*(01n02), 63–93.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the royal society of London A: Mathematical, physical and engineering sciences* (Vol. 454, No. 1971, pp. 903–995). The Royal Society.

Jotzo, F., & Löschel, A. (2014). Emissions trading in China: Emerging experiences and international lessons. *Energy Policy, 75,* 3–8.

Koop, G., & Tole, L. (2013). Forecasting the European carbon market. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(3), 723–741.

Kugiumtzis, D., Lingjærde, O. C., & Christophersen, N. (1998). Regularized local linear prediction of chaotic time series. *Physica D: Nonlinear Phenomena, 112*(3–4), 344–360.

Lo, A. Y. (2016). Challenges to the development of carbon markets in China. *Climate Policy, 16*(1), 109–124.

Lu, Z. Q. (2002). Local polynomial prediction and volatility estimation in financial time series. In A. S. Soofi & L. Cao (Eds.), *Modelling and forecasting financial data. Studies in computational finance* (Vol. 2, pp. 115–135). Boston, MA: Springer.

Neuhoff, K., Martinez, K. K., & Sato, M. (2006). Allocation, incentives and distortions: The impact of EU ETS emissions allowance allocations to the electricity sector. *Climate Policy, 6*(1), 73–91.

Packard, N. H., Crutchfield, J. P., Farmer, J. D., & Shaw, R. S. (1980). Geometry from a time series. *Physical Review Letters, 45*(9), 712.

Qin, Q., Cheng, S., Zhang, Q., Wei, Y., & Shi, Y. (2015). Multiple strategies based orthogonal design particle swarm optimizer for numerical optimization. *Computers & Operations Research, 60,* 91–110.

Qin, Q., Liu, Y., Li, X., & Li, H. (2017). A multi-criteria decision analysis model for carbon emission quota allocation in China's east coastal areas: Efficiency and equity. *Journal of Cleaner Production, 168,* 410–419.

Regonda, S., Rajagopalan, B., Lall, U., Clark, M., & Moon, Y. I. (2005). Local polynomial method for ensemble forecast of time series. *Nonlinear Processes in Geophysics, 12*(3), 397–406.

Ren, Y., Suganthan, P. N., & Srikanth, N. (2015). A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods. *IEEE Transactions on Sustainable Energy, 6*(1), 236–244.

Su, L., & Li, C. (2015). Local prediction of chaotic time series based on polynomial coefficient autoregressive model. *Mathematical Problems in Engineering*, *2015*, 901807.

Sun, G., Chen, T., Wei, Z., Sun, Y., Zang, H., & Chen, S. (2016). A carbon price forecasting model based on variational mode decomposition and spiking neural networks. *Energies, 9*(1), 54.

Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L. S. Young (Eds.), *Dynamical systems and turbulence, Warwick 1980. Lecture notes in mathematics* (Vol. 898, pp. 366–381). Berlin: Springer.

Tang, B. J., Gong, P. Q., & Shen, C. (2017). Factors of carbon price volatility in a comparative analysis of the EUA and sCER. *Annals of Operations Research, 255*(1–2), 157–168.

Tang, L., Wang, S., & Yu, L. (2011). EEMD-LSSVR-based decomposition-and-ensemble methodology with application to nuclear energy consumption forecasting. In *Computational sciences and optimization (CSO), 2011 fourth international joint conference on* (pp. 589–593). IEEE.

Tang, L., Yu, L., Wang, S., Li, J., & Wang, S. (2012). A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. *Applied Energy, 93,* 432–443.

Wang, W. C., Chau, K. W., Xu, D. M., & Chen, X. Y. (2015). Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resources Management, 29*(8), 2655–2675.

Wei, Y. M., Wang, K., Feng, Z. H., & Cong, R. (2010). *Carbon finance and carbon market: Models and empirical analysis*. Beijing: Science Press.

Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis, 1*(01), 1–41.

Xiong, T., Bao, Y., & Hu, Z. (2013). Beyond one-step-ahead forecasting: Evaluation of alternative multi-step-ahead forecasting models for crude oil prices. *Energy Economics, 40,* 405–415.

Yu, L., Dai, W., & Tang, L. (2016). A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Engineering Applications of Artificial Intelligence, 47,* 110–121.

Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics, 30*(5), 2623–2635.

Zhang, X., Lai, K. K., & Wang, S. Y. (2008). A new approach for crude oil price analysis based on empirical mode decomposition. *Energy Economics, 30*(3), 905–918.

Zhang, J. L., Zhang, Y. J., & Zhang, L. (2015). A novel hybrid method for crude oil price forecasting. *Energy Economics, 49,* 649–659.

Zhu, B., Shi, X., Chevallier, J., Wang, P., & Wei, Y. M. (2016). An adaptive multiscale ensemble learning paradigm for nonstationary and nonlinear energy price time series forecasting. *Journal of Forecasting, 35*(7), 633–651.

Zhu, B., Wang, P., Chevallier, J., & Wei, Y. (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, 41*(3), 517–524.