

Article

A Novel Multiscale Ensemble Carbon Price Prediction Model Integrating Empirical Mode Decomposition, Genetic Algorithm and Artificial Neural Network

Bangzhu Zhu

School of Economics and Management, Wuyi University, Jiangmen 529020, Guangdong, China;
E-Mail: wpzbz@126.com; Tel.: +86-0750-3296568; Fax: +86-0750-3299335

Received: 8 December 2011; in revised form: 8 February 2012 / Accepted: 8 February 2012 /

Published: 17 February 2012

Abstract: Due to the movement and complexity of the carbon market, traditional monoscale forecasting approaches often fail to capture its nonstationary and nonlinear properties and accurately describe its moving tendencies. In this study, a multiscale ensemble forecasting model integrating empirical mode decomposition (EMD), genetic algorithm (GA) and artificial neural network (ANN) is proposed to forecast carbon price. Firstly, the proposed model uses EMD to decompose carbon price data into several intrinsic mode functions (IMFs) and one residue. Then, the IMFs and residue are composed into a high frequency component, a low frequency component and a trend component which have similar frequency characteristics, simple components and strong regularity using the fine-to-coarse reconstruction algorithm. Finally, those three components are predicted using an ANN trained by GA, *i.e.*, a GAANN model, and the final forecasting results can be obtained by the sum of these three forecasting results. For verification and testing, two main carbon future prices with different maturity in the European Climate Exchange (ECX) are used to test the effectiveness of the proposed multiscale ensemble forecasting model. Empirical results obtained demonstrate that the proposed multiscale ensemble forecasting model can outperform the single random walk (RW), ARIMA, ANN and GAANN models without EMD preprocessing and the ensemble ARIMA model with EMD preprocessing.

Keywords: carbon price; multiscale prediction; empirical mode decomposition; artificial neural network; genetic algorithm; partial autocorrelation function

1. Introduction

Climate change has been a common challenge in the last few decades. In order to reduce greenhouse gas emissions at the lowest overall cost, the European Union Emissions Trading Scheme (EU ETS) was launched within European Union covering around 12000 installations in 25 countries and six major industrial sectors in 2005. EU ETS is the largest carbon market in the World to date [1], which has proven to be not only an important tool for human beings to address climate changes, but also a major choice for investors to decentralize their investment risks [2]. Therefore, the need for more accurate forecasts of carbon price is driven by the desire to reduce risk and uncertainty.

Recently, although there is much literature on carbon price analysis [1–4], seldom existing literature regarding carbon price forecasting can be found. In fact, carbon price forecasting is a kind of time series forecasting. During the past few decades, various approaches have been developed for time series forecasting, among which the so-called autoregressive integrated moving average (ARIMA) method has been found to be one of the most effective forecasting methods. The popularity of the ARIMA method is due to its statistical properties as well as the well-known Box-Jenkins methodology in the modeling process. However, the ARIMA method is only a class of linear model and thus it can only capture linear patterns of time series. In order to overcome the limitations of the linear models and account for the nonlinear patterns existing in real problems, numerous nonlinear models have been proposed, among which artificial neural networks (ANNs) have shown excellent nonlinear modeling capability. Although a large number of successful applications have shown that ANNs can be a very useful tools within the stationary forecasting domain [5–7], however, carbon price data are highly nonstationary [3,4], which will make its forecasting precision still unsatisfactory.

The Empirical Mode Decomposition (EMD) [8], proposed by Huang *et al.* in 1998, appears to be a new adaptive data analysis approach to improving forecasting precision for nonlinear and nonstationary carbon price data. EMD can capture the physical properties of the observed data accurately and has strong local performance capacity. Therefore, EMD is effective in dealing with nonlinear and nonstationary carbon price data [9]. If the original carbon price data are directly fed into an ANN, carbon price data will not present any outstanding characteristic quantities and it will take a longer time for the ANN to understand and grasp the data's characteristics. Through EMD, carbon price data are decomposed into several independent intrinsic mode functions (IMFs), thus simplifying the interference and coupling across characteristic information of different scales in carbon price data. Meanwhile, each IMF itself describes different local characteristics of the carbon price data, thus an ANN can better understand and grasp the IMF's characteristics so as to improve the efficiency of learning as well as accuracy of forecasting if any IMF is used as an input of ANN [3]. In recent years, some studies have applied an EMD based ANN model for time series forecasting and obtained good results [3,10,11]. However, they often use the traditional back-propagation ANN (BPANN) as predictor, which may lead to overfitting of the data. Moreover, existing literature regarding carbon price analysis has not adopted EMD processes, and this study thus aims to fill this gap.

The contributions of this paper are twofold. Firstly, we establish an EMD-based ANN multiscale ensemble forecasting model to forecast carbon price. Using EMD, carbon price data are decomposed into several IMFs and a residue, which are composed into a high frequency component, a low frequency component and a trend component using the fine-to-coarse reconstruction algorithm. Then

those three components are predicted using an ANN trained by genetic algorithm (GA), *i.e.*, GAANN, respectively and the forecasting values are summarized as the final forecasting results of carbon price. Secondly, we evaluate the forecasting performance of the single random walk (RW), ARIMA, ANN and GAANN models without EMD preprocessing and the ensemble ARIMA model with EMD preprocessing, for forecasting the carbon price in the European Climate Exchange (ECX) market. Empirical results obtained demonstrate that the proposed multiscale ensemble forecasting model can outperform the single RW, ARIMA, ANN and GAANN models without EMD preprocessing and the ensemble ARIMA model with EMD preprocessing.

The remainder of this study is organized as follows. Section 2 describes the EMD, the GAANN and proposed multiscale ensemble forecasting models. Section 3 reports the empirical results. Section 4 provides some conclusions.

2. Methodology

2.1. EMD

EMD assumes that carbon price data simultaneously have many modes of different oscillations. Each mode, treated as an IMF, can be extracted from the data based on local characteristic scale of the data themselves. IMF meets two conditions [8]: (a) IMF has the same number of extrema and zero-crossings or differs by one at the most; (b) IMF is symmetric with the local zero mean. EMD can extract the IMFs through a sifting process as follows:

- (1) Identify all the maxima and minima of carbon price data $x(t)$;
- (2) Generate their upper and lower envelopes, $e_{\max}(t)$ and $e_{\min}(t)$, with cubic spline interpolation;
- (3) Calculate the point-by-point mean $m(t)$ from the upper and lower envelopes:

$$m(t) = [e_{\max}(t) + e_{\min}(t)] / 2$$

- (4) Extract the mean from carbon price data and define the difference between $x(t)$ and $m(t)$ as $d(t)$:

$$d(t) = x(t) - m(t)$$

- (5) Check the properties of $d(t)$:
 - (a) If it is an IMF, denote $d(t)$ as the i th IMF and replace $x(t)$ with the residue $r(t) = x(t) - d(t)$. The i th IMF is often denoted as $c_i(t)$ and the i is called its index;
 - (b) If it is not an IMF, replace $x(t)$ with $d(t)$;
- (6) Repeat steps (1)–(5) until the residue satisfies some stopping criteria.

Rilling *et al.* [12] proposed one stopping criterion by introducing two thresholds θ_1 and θ_2 , and defining the mode amplitude $\alpha(t)$:

$$\alpha(t) = |e_{\max}(t) + e_{\min}(t)| / 2$$

and the evaluation function $\sigma(t)$:

$$\sigma(t) = |m(t) / \alpha(t)|$$

so that sifting is iterated until $\sigma(t) < \theta_1$ for some prescribed fractions $(1 - \alpha)$ of the total duration, while $\sigma(t) < \theta_2$ for the remaining fractions. One can typically set $\alpha = 0.05$, $\theta_1 = 0.05$ and $\theta_2 = 10\theta_1$.

At the end of this sifting procedure, carbon price data $x(t)$ can be expressed as:

$$x(t) = \sum_{i=1}^m c_i(t) + r_m(t)$$

where m is the number of IMFs and $r_m(t)$ is the final residue. Thus, we can achieve the decomposition of carbon price data into m IMFs and one residue.

2.2. Fine-to-Coarse Reconstruction

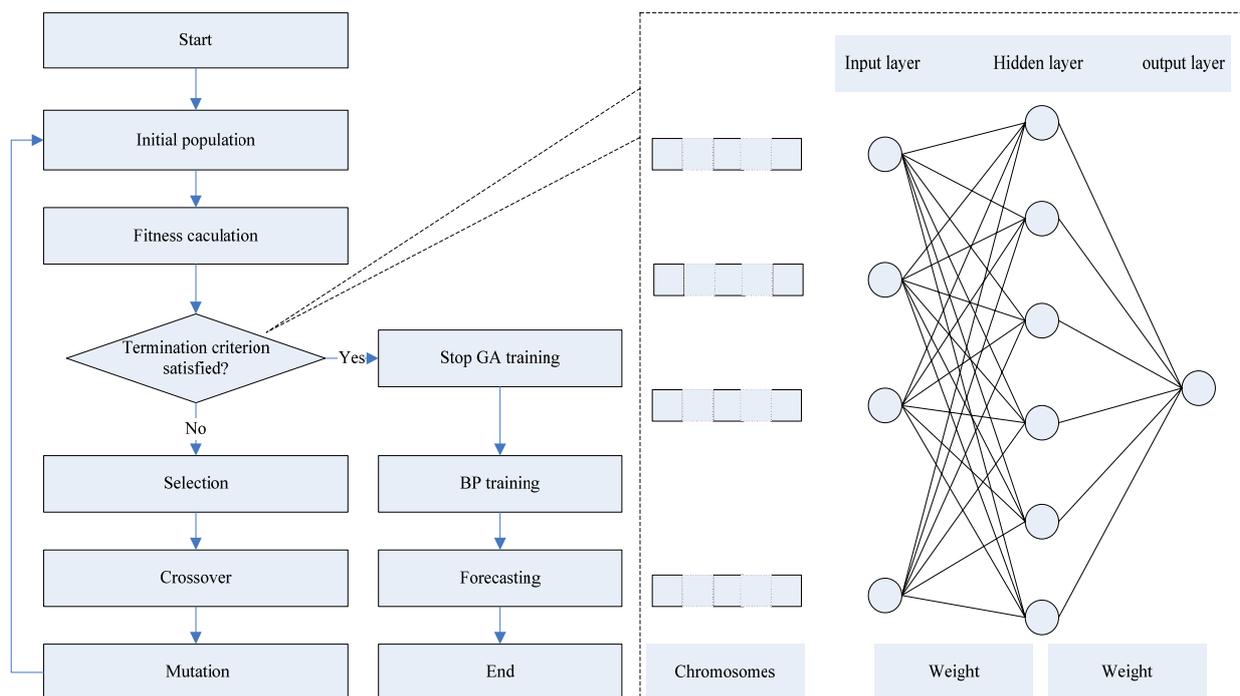
Carbon price data are decomposed into m IMFs and one residue by EMD, which can be decomposed into a high frequency component, a low frequency component and a trend component based on the fine-to-coarse reconstruction algorithm [13]:

- (1) Compute the mean of the sum of c_1 to $c_i (1 \leq i \leq m)$, i.e., $s_i = \sum_{k=1}^i c_k$ for each component (except for the residue);
- (2) Select the significance level α and employ t -test to identify for which i the mean significantly departs from zero for the first time;
- (3) Once i is identified as a significant change point, partial reconstruction with IMFs from this to the end is identified as a low frequency component, and the partial reconstruction with other IMFs is identified as a high frequency component. The residue is identified as a trend component.

2.3. Combining ANN and GA for Regression

In this study, we develop a hybrid ANN and GA model for regression, i.e., GAANN model, as seen in Figure 1.

Figure 1. Framework of combining ANN and GA for regression.



Since a three-layer feedforward BPANN can map any nonlinear relationship with a desired degree of accuracy, we adopt a three-layer BPANN as predictor of the original series and the restructured components, in which the transfer functions of hidden and output layers are sigmoid and linear, respectively. During the learning process, the error is backward propagated through ANN to adjust the weights of the connections and thresholds, minimizing the sum of the mean squared error (*MSE*) in the output layer [14]:

$$MSE = \sum_{k=1}^n \sum_{j=1}^m [T_j(k) - Y_j(k)]^2$$

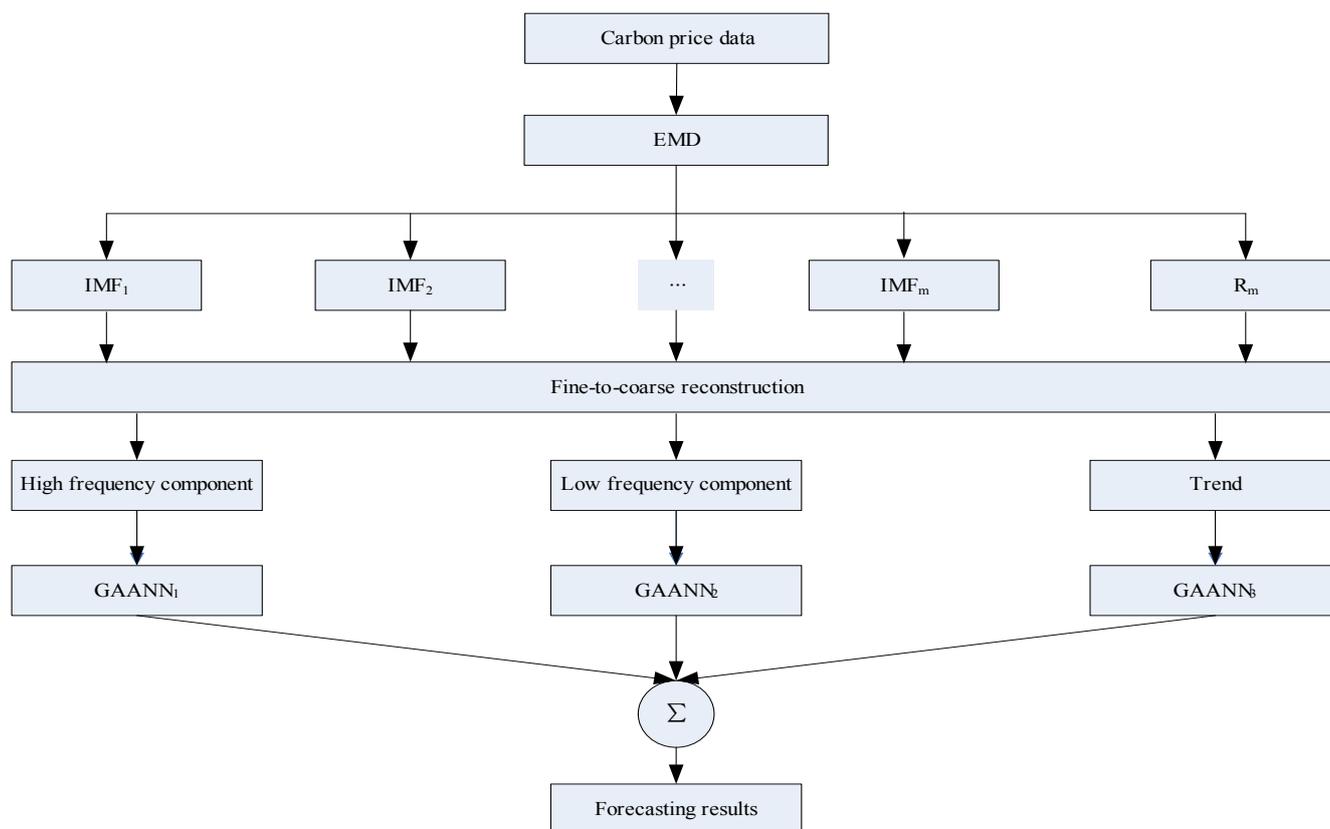
where m is the number of output nodes, n is the number of training samples, $T_j(k)$ is the expected output, and $Y_j(k)$ is the actual output.

However, a potential difficulty in the use of BPANN is the possibility of overfitting the data. To avoid the overfitting, we employ GA to optimize the weights and thresholds of the ANN. The GA can find global optimal solutions by constructing fitter solutions, which processes populations of chromosomes by replacing unsuitable candidates according to the fitness function [15]. In this study, we define the fitness function as $F = 1 / (MSE + 1)$. The objective of the optimization is to maximize the fitness values F which will lead to the minimization of the *MSE*. Thus, the smaller the *MSE*, the closer fitness value to 1 (maximum). Once the fitness values of all chromosomes are evaluated, a population of chromosomes is updated using three genetic operators: selection, crossover and mutation. The selection operator of the GA is implemented by using the roulette-wheel algorithm to determine which population members are chosen as parents that will create offspring for the next generation. Crossover is a mechanism of randomly exchanging information between two chromosomes. We use arithmetical crossover which can ensure the offspring are still within the constraint region. Mutation operation can change the values of randomly chosen gene bits, and this process will continue until some predefined termination criteria are fulfilled. This ensures that we can obtain a good ANN.

2.4. EMD-Based GAANN Multiscale Ensemble Forecasting Model

Figure 2 shows the proposed EMD-based GAANN multiscale ensemble forecasting model, which works as follows:

- Step 1: Use the EMD to decompose the carbon price data into a set of IMFs and one residue.
- Step 2: Apply the fine-to-coarse reconstruction algorithm to reconstruct the IMFs and residue obtained from decomposition into a high frequency component, a low frequency component and a trend component.
- Step 3: Use the GAANN model to forecast the future one-day values of those three reconstructed components.
- Step 4: The forecasting results obtained by the sum of the predicted values in the previous step, can be treated as the final prediction results for the original carbon price.

Figure 2. EMD-based GAANN multiscale ensemble prediction model.

To summarize, the proposed EMD-based GAANN multiscale ensemble forecasting model is actually an “EMD–GAANN– Σ ” ensemble learning approach, which is an application of the “decomposition and ensemble” strategy [10]. In order to verify the effectiveness of the proposed EMD–GAANN– Σ model, two main carbon future prices with maturity in December, 2010 (DEC 10) and December, 2012 (DEC 12) are used for testing purpose in the next section.

3. Empirical Analysis

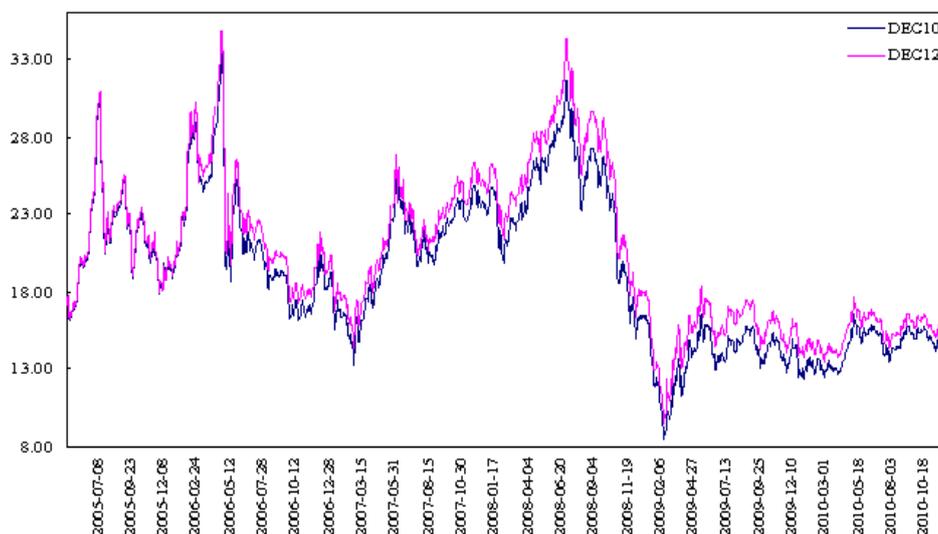
3.1. Data

The ECX, located in London, is the largest carbon trading market affiliated under EU ETS, since its daily trading volume generally accounts for over 80% of the total trading volume. It goes without saying that the state of ECX can reflect the overall state of EU ETS to a great extent.

As is known to all, there are a great number of carbon prices in ECX. In this study, two carbon price series, DEC10 and DEC12 from April 22, 2005 to December 3, 2010, excluding public holidays, with a total of 1438 observations, are chosen as experimental samples. For convenience of GAANN modeling for DEC10 and DEC12, we take daily data from April 22, 2005 to March 18, 2009, excluding public holidays, with a total of 1000 data points are used as the in-sample training sets, and the remainder with a total of 438 data points are used as the out-of-sample testing sets, which are used to check the forecasting ability based on evaluation criteria. The main reason of selecting these two carbon prices is that they are the longest two future contracts covering the entire operating segment of EU ETS. The data of two carbon prices used in this paper are daily data, freely available from the ECX

website [16]. Figure 3 shows the curve of daily carbon prices for DEC10 and DEC12 in unit of Euros/ton, which shows that carbon price movements appear to be nonlinear and nonstationary in that the means are changing over time.

Figure 3. DEC10 and DEC12 from April 22, 2005 to December 3, 2010.



3.2. Evaluation Criteria

For the sake of measuring the forecasting performance, two main criteria, root mean squared error (RMSE) and directional prediction statistic (D_{stat}), are used to evaluate the level prediction and directional forecasting, respectively:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n [\hat{x}(t) - x(t)]^2} \quad \text{and} \quad D_{stat} = \frac{1}{n} \sum_{t=1}^n a_t \times 100\%$$

where $x(t)$ is the actual value, $\hat{x}(t)$ is the predicted value, n is the number of predictions, $a_t = 1$ if $[x(t) - x(t-1)][\hat{x}(t) - \hat{x}(t-1)] \geq 0$, and $a_t = 0$.

For comparing the prediction capacity of the proposed EMD–GAANN– Σ model with other widely used forecasting approaches, we employ the single RW, ARIMA, ANN and GAANN models as benchmark models. Moreover, a variant of the ensemble model, the EMD–ARIMA– Σ model, is also used to predict carbon price for the purpose of comparison.

3.3. Forecasting Results

We conduct the prediction experiments following the previous steps as shown in Section 2.4. Firstly, we decompose each of the two typical carbon price series into a set of IMFs and a residue. Before that, the thresholds and tolerance level of the stop criterion are determined by $[\theta_1, \theta_2, \alpha] = [0.05, 0.5, 0.05]$. We get graphical representations of the decomposition results through EMD, as illustrated in Figures 4 and 5. Obviously, DEC10 is decomposed into eight IMFs and one residue, DEC12 is decomposed into seven IMFs and one residue.

Figure 4. The decomposition of DEC10 derived from EMD.

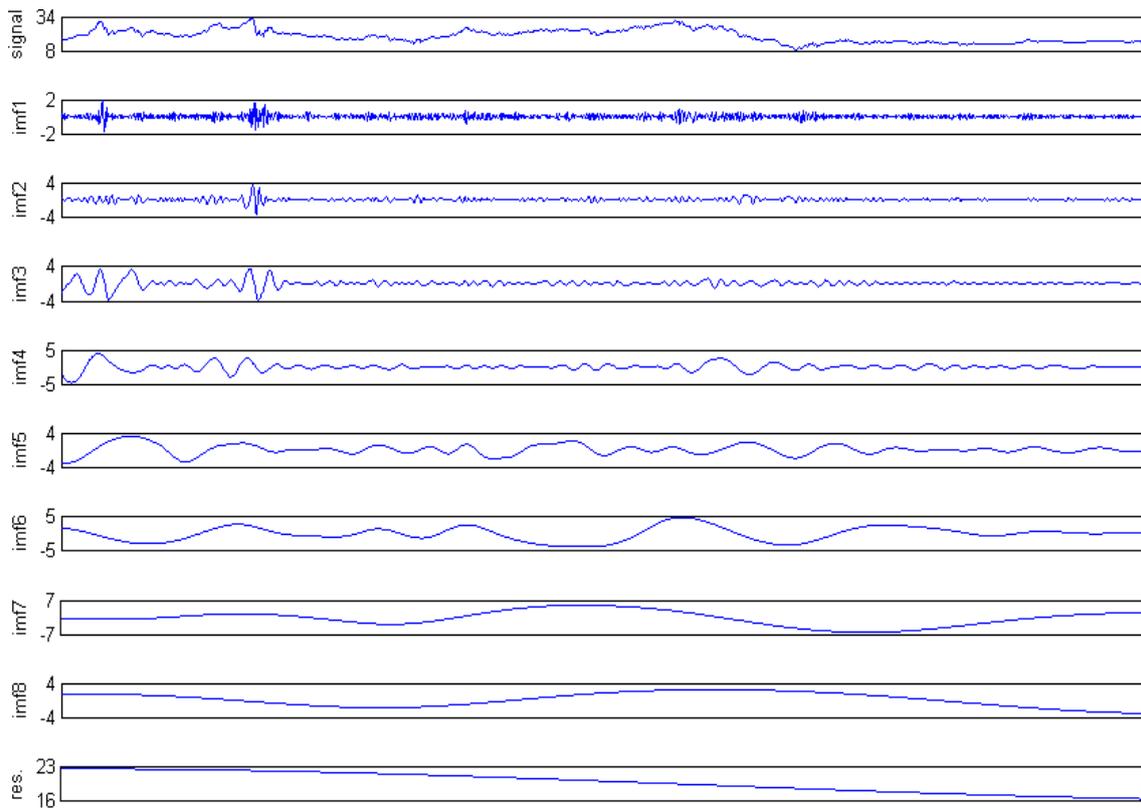


Figure 5. The decomposition of DEC12 derived from EMD.

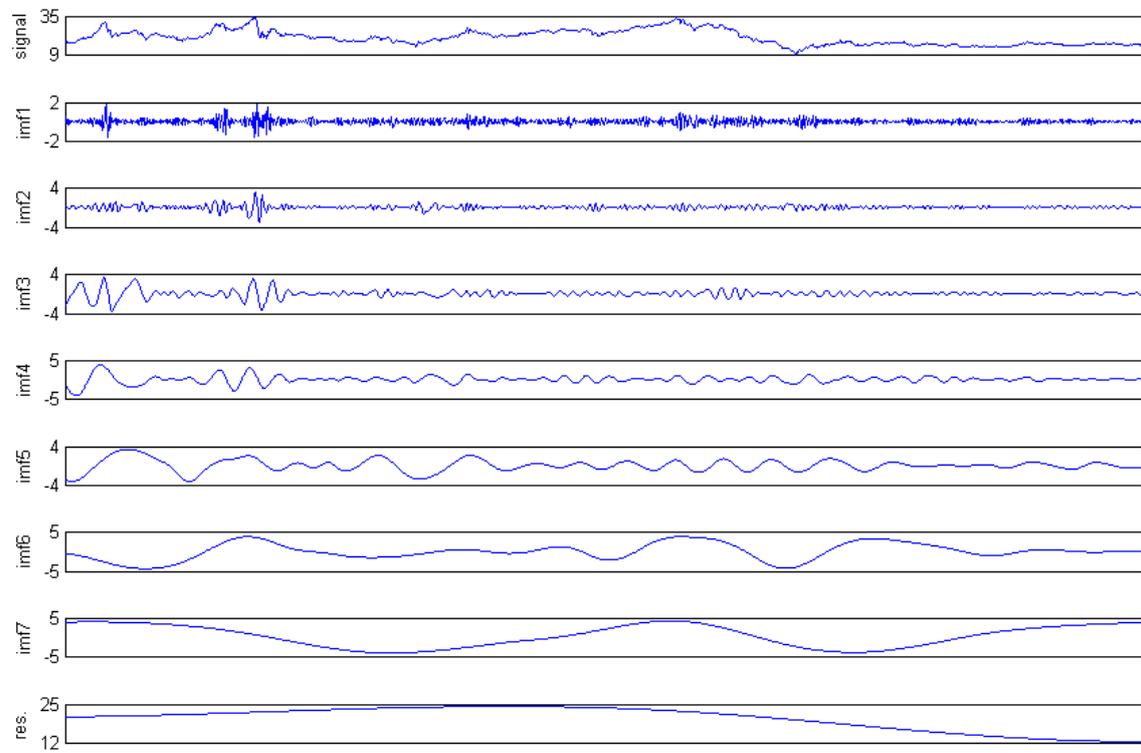


Table 1. Mean of s_i and t value of DEC10.

Item	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
Mean	-3.23×10^{-4}	2.54×10^{-3}	-4.52×10^{-4}	-4.90×10^{-3}	1.20×10^{-2}	-4.13×10^{-2}	-9.71×10^{-2}	-6.55×10^{-2}
t value	-0.044	0.399	-0.058	-0.551	1.273	-3.463	-6.325	-4.527

Table 2. Mean of s_i and t value of DEC12.

Item	s_1	s_2	s_3	s_4	s_5	s_6	s_7
Mean	-1.24×10^{-4}	-1.81×10^{-3}	-5.27×10^{-3}	-1.74×10^{-2}	-6.21×10^{-3}	-1.28×10^{-2}	2.52×10^{-2}
t value	-0.016	-0.276	-0.682	-2.031	-0.657	-1.073	1.698

Figure 6. The reconstruction of DEC10 derived from EMD.

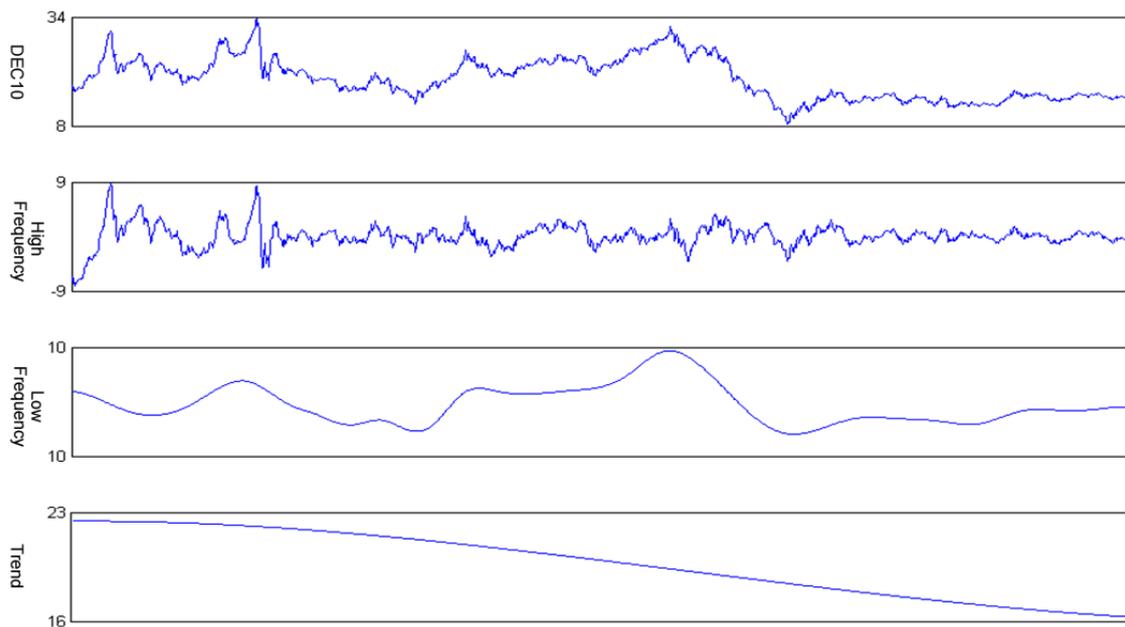
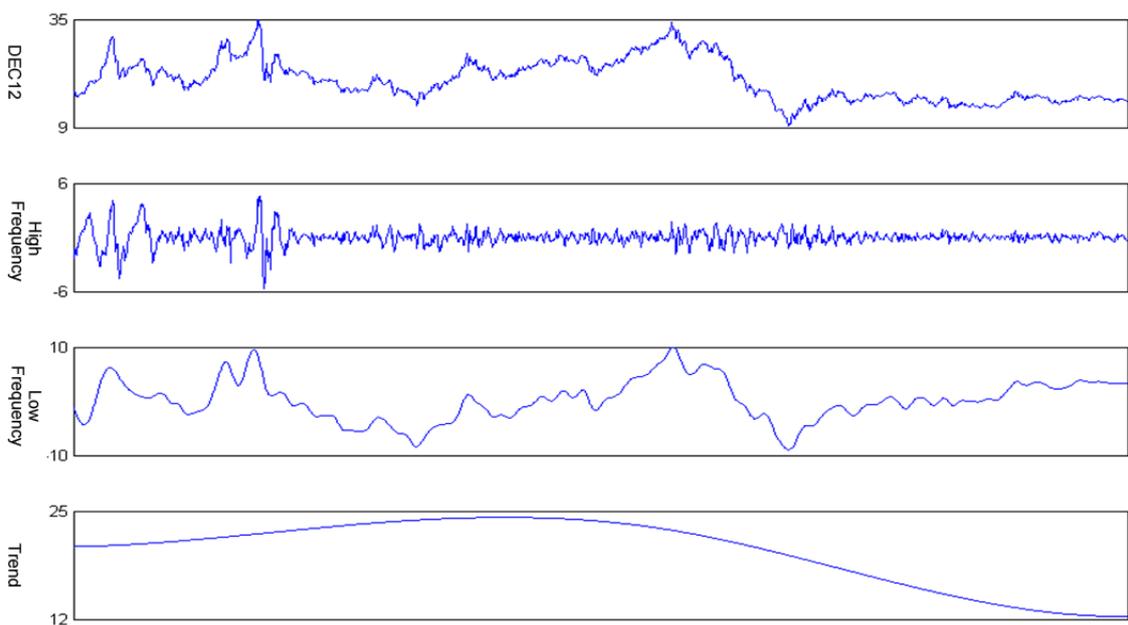


Figure 7. The reconstruction of DEC12 derived from EMD.



Then, the fine-to-coarse reconstruction algorithm is used to reconstruct the IMFs and residue into a high frequency component, a low frequency component and a trend component. The t values of DEC10 and DEC12 corresponding to means of s_i based on fine-to-coarse reconstruction algorithm are shown in Tables 1 and 2. We can find that the means of the fine-to-coarse reconstruction depart significantly from zero at $i = 6$ (DEC10) and $i = 4$ (DEC12) for the first time. Thus, for DEC10, the partial reconstruction with IMF1, IMF2, IMF3, IMF4 and IMF5 represents the high frequency component, the partial reconstruction with IMF6, IMF7 and IMF8 represents the low frequency component and the residue is separately treated as the trend component. For DEC12, the partial reconstruction with IMF1, IMF2 and IMF3 represents the high frequency component, the partial reconstruction with IMF4, IMF5, IMF6 and IMF7 represents the low frequency component and the residue is also separately treated as the trend component. As seen from Figures 6 and 7, each component is more stationary and regular, which can help improve the prediction performance by employing the “decomposition and ensemble” strategy [10].

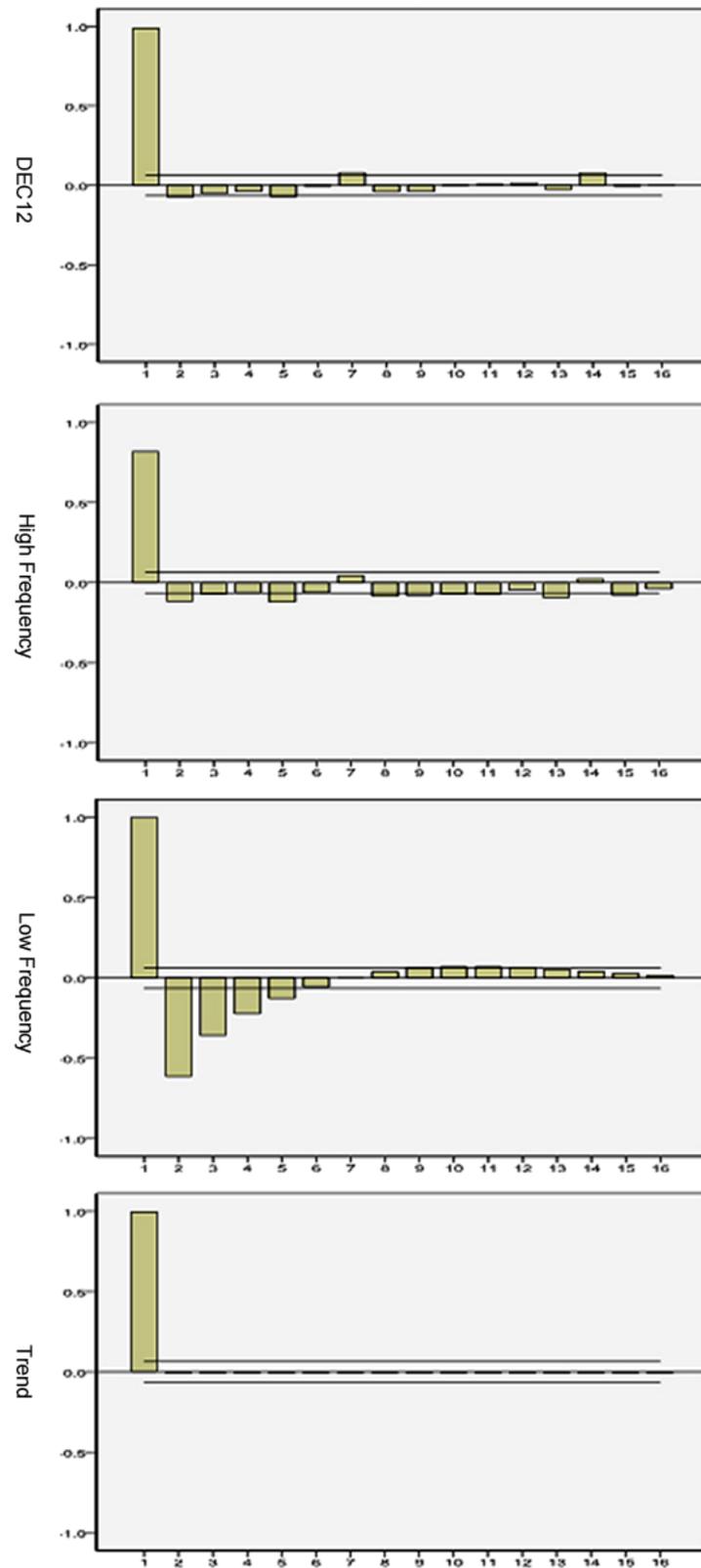
We first take DEC10 for single-step-ahead forecasting. RW modeling is implemented via the Excel 2003 software produced by Microsoft Corporation. ARIMA modeling is implemented via the Eviews statistical software package produced by Quantitative Micro Software Corporation. The model with the lowest Akaike Information Criteria (AIC) and Schwarz criterion (SC) is the best model. Once the optimal ARIMA model has been identified, it can be used to predict the high frequency component, low frequency component, trend component and the original series. Meanwhile we aggregate the forecasting results of those three components to produce an ensemble forecasting result, which is the EMD-ARIMA- \sum modeling process.

Moreover, the ANN, GAANN and EMD-GAANN- \sum models are established with the neural network toolbox (Version 5.0) of the Matlab software package produced by the Mathworks Laboratory Corporation. Inspired by the identification of parameter p in ARIMA(p, d, q) model, we use the statistical tools, the partial autocorrelation function (PACF) and the resulting partial autocorrelation graph which is simply the plots of PACF against the lag length, to determine the input variables of (GA)ANN for orienting on the matter [11]. At the same time, we adopt the Kolmogorov theorem $s = 2m + 1$ [17] to determine the number of hidden layer nodes, where m represents the number of input nodes and s represents the number of hidden layer nodes. Furthermore, GA is operated with real code, initial population size of 100, genetic algebra of 1,000, uniform crossover rate of 0.9, uniform mutation probability of 0.1 and other default parameters of GAOT toolbox [18].

We can thus get the partial autocorrelogram of the high frequency component, low frequency component, trend component and the original series, which are shown in Figure 8. According to the input selection method through observing Figure 8, with the output variable x_t , the input variables of these four series for (GA)ANN modeling are as follows:

- DEC10: (x_{t-1}, x_{t-2}) ;
- High frequency component: (x_{t-1}, x_{t-2}) ;
- Low frequency component: $(x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4})$;
- Trend component: (x_{t-1}) .

Figure 9. The PACFs of the original series and the reconstructed components of DEC12.



Having modeled those four series by (GA)ANN and rescaling them, we obtain the predictors of those three components and the original series. Moreover, the EMD-GAANN- Σ model is similar to EMD-ARIMA- Σ , which only uses GAANN to forecast the high frequency component, low frequency

component and trend component and aggregate their forecasting results. Furthermore, we have run the (GA)ANN model ten times and averaged the results to stabilize its outputs, and the final forecasting results obtained from those six models for DEC10 are presented in Tables 3 and 4.

Table 3. RMSE comparisons for different forecasting models.

Models	DEC10		DEC12	
	RMSE	Rank	RMSE	Rank
RW	0.2962	4	0.3176	5
ARIMA	0.3002	6	0.3197	6
ANN	0.2986	5	0.3078	4
GAANN	0.2952	3	0.2986	3
EMD-ARIMA- Σ	0.2886	2	0.2912	2
EMD-GAANN- Σ	0.2817	1	0.2856	1

Table 4. D_{stat} comparisons for different forecasting models.

Models	DEC10		DEC12	
	D_{stat}	Rank	D_{stat}	Rank
RW	48.40	5	49.77	5
ARIMA	47.72	6	48.86	6
ANN	64.38	4	60.96	4
GAANN	67.58	3	64.16	3
EMD-ARIMA- Σ	69.18	2	67.58	2
EMD-GAANN- Σ	70.09	1	69.63	1

For DEC12, in the same way we can obtain the partial autocorrelogram of the high frequency component, low frequency component, trend component and the original series which are shown in Figure 9, so the input variables of these four series for (GA)ANN modeling are as follows:

- DEC12: (x_{t-1}, x_{t-2}) ;
- High frequency component: (x_{t-1}, x_{t-2}) ;
- Low frequency component: $(x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, x_{t-5})$;
- Trend component: (x_{t-1}) .

Through the same process mentioned above, we can obtain the forecasting results of RW, ARIMA, ANN, GAANN, EMD-ARIMA- Σ and EMD-GAANN- Σ models, which are also presented in Tables 3 and 4.

In terms of RMSE, the proposed EMD-GAANN- Σ model performs the best, followed by EMD-ARIMA- Σ model, GAANN model, ANN model, RW model and ARIMA model. Both ANN and GAANN are better than ARIMA, mostly because the former two models are nonlinear and the latter is a class of linear model, which is not suitable to forecast the nonstationary and nonlinear carbon price; GAANN is better than ANN, mainly because the global optimization capacity of GA can improve the ANN's forecasting ability. EMD-based multiscale ensemble forecasting models are better than each of single models, possibly because the EMD decomposition can promote the prediction performance.

With regard to D_{stat} , the proposed EMD-GAANN- Σ model also performs better than other models. EMD-based multiscale ensemble forecasting models are also better than the single forecasting models, possibly because the advantages of ensemble strategy have a great effect on the overall forecasting ability. Both ANNs and GAANN models are better than ARIMA, mostly because highly nonlinear carbon price data have such complex intrinsic characteristics, and the latter is a class of linear model which cannot capture such characteristics well; GAANN is better than ANNs, which demonstrates that GA's global optimization capacity can improve ANN's learning and forecasting ability.

In general, according to the experimental results of carbon price forecasting for DEC10 and DEC12 presented in this study, we can draw the following conclusions: (1) The experimental results show that EMD-GAANN- Σ model is superior to RW, ARIMA, ANN and GAANN models, as well as EMD-ARIMA- Σ model, for the test cases of the two main carbon future prices, in terms of accuracy level of prediction, as measured by RMSE, and directional prediction statistics (D_{stat}); (2) The prediction performance of the EMD-GAANN- Σ model and EMD-ARIMA- Σ model are much better than that of ARIMA model. Likewise, EMD-GAANN- Σ and EMD-ARIMA- Σ models perform better than ANN and GAANN models. This indicates that the decomposition-and-ensemble strategy can effectively improve the prediction performance, and the results emphasize that EMD decomposition is meaningful to prediction performance improvement in carbon price forecasting; (3) EMD-GAANN- Σ model is better than RW model in terms of NMSE and D_{stat} , which provides evidence against the efficient market hypothesis and suggests that EMD-GAANN- Σ can forecast carbon prices in the future.

4. Conclusions

This study has proposed an EMD-based GAANN multiscale ensemble forecasting model to predict carbon prices. The main contribution of this study is to present a novel method as well as a simple approach for a stable prediction of nonstationary and nonlinear carbon price data. The proposed method preprocesses the carbon price data and decomposes them into more stationary and regular components (a high frequency component, a low frequency component and a trend component) using the EMD and fine-to-coarse reconstruction algorithms. Furthermore, the corresponding GAANN model for each divided component is easier to build. After the components are forecasted in the built GAANN models, the forecasting values are then summarized the final carbon price forecasting results. The experiments have evaluated two main carbon future prices from the ECX market. This study compared the proposed method with the single RW, ARIMA, ANN, GAANN models and the ensemble EMD-ARIMA- Σ model, using RMSE and D_{stat} as the criteria.

Empirical results shows that the proposed EMD-based GAANN multiscale ensemble forecasting model can produce the lowest RMSE and the highest D_{stat} in the carbon price datasets and exceed the single RW, ARIMA, ANN and GAANN models, as well as the ensemble EMD-ARIMA- Σ model. According to the experiments, EMD which can fully capture the local fluctuations of data can be used as a preprocessor to decompose the complicated raw data into a finite set of IMFs and one residue, which have simpler frequency components and high correlations. By this preprocessing, we can not only advance the simplification of GAANN modeling, but also obtain much more precise than the RW, ARIMA, ANN and GAANN models based on RMSE and D_{stat} . Therefore, the proposed method is very

suitable for prediction with nonlinear, nonstationary and strong complexity data, and is a very promising methodology for carbon price forecasting.

Acknowledgments

I wish to thank the two anonymous referees for insightful comments that led to a greatly improved version of the article. I also wish to thank Yu Lean for meaningful help, as well as the National Science Foundation of China (NSFC) (71020107026, 70733005) and the Ministry of Education of China (11YJC630304) for funding support.

References

1. Montagnoli, A.; de Vries, F.P. Carbon trading thickness and market efficiency. *Energy Econ.* **2010**, doi:10.1016/j.eneco.2010.04.001.
2. Zhang, Y.J.; Wei, Y.M. An Overview of Current Research on EU ETS: Evidence from its Operating Mechanism and Economic Effect. *Appl. Energy* **2010**, *6*, 1804–1814.
3. Seifert, J.; Marliese, U.H.; Michael, W. Dynamic behavior of CO₂ spot prices. *J. Environ. Econ. Manag.* **2008**, *56*, 180–194.
4. Feng, Z.H.; Zou, L.L.; Wei, Y.M. Carbon price volatility: Evidence from EU ETS. *Appl. Energy* **2011**, *88*, 590–598.
5. Wang, W.B.; Fei, P.S.; Yi, X.M. Prediction of China stock based on EMD and neural network. *Syst. Eng. Theory Pract.* **2010**, *6*, 1027–1033.
6. Yu, L.; Wang, S.Y.; Lai, K.K. A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. *Comput. Oper. Res.* **2005**, *32*, 2523–2541.
7. Han, M.; Xi, J.H.; Xu, S.G.; Yin, F.L. Prediction of chaotic time series based on the recurrent network. *IEEE Trans. Sign. Proc.* **2004**, *2*, 3409–3416.
8. Huang, N.E.; Shen, Z.; Long, S.R. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Process R. Soc. Lond.* **1998**, *A454*, 903–995.
9. Huang, N.E.; Wu, M.L.; Qu, W.D.; Long, S.R.; Shen, S.S.P. Applications of Hilbert Huang transform to non-stationary financial time series analysis. *Appl. Stoch. Models Bus. Ind.* **2003**, *19*, 245–268.
10. Yu, L.; Wang, S.Y.; Lai, K.K.; Wen, F.H. A multiscale neural network learning paradigm for financial crisis forecasting. *Neurocomputing* **2010**, *73*, 716–72.
11. Guo, Z.; Zhao, W.; Lu H.; Wang, J.Z. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renew. Energy* **2012**, *37*, 241–249.
12. Riling, G.; Flandrin, P.; Goncalves, P. On Empirical Mode Decomposition and its algorithms. In *Proceedings of the IEEE EURASIP Workshop on Nonlinear Signal and Image Processing*, Grado, Italy, June 2003.
13. Zhang, X.; Lai, K.K.; Wang, S.Y. A new approach for crude oil price analysis based on Empirical Mode Decomposition. *Energy Econ.* **2008**, *30*, 905–918.
14. Sedki, A.; Ouazar, D.; El Mazoudi, E. Evolving neural network using real coded genetic algorithm for daily rainfall–runoff forecasting. *Expert Syst. Appl.* **2009**, *36*, 4523–4527.

15. Irani, R.; Nasimi, R. Evolving neural network using real coded genetic algorithm for permeability estimation of the reservoir. *Expert Syst. Appl.* **2011**, *38*, 9862–9866.
16. ECX EUA Futures Contract: Historic Data 2005-2010. Available online: <http://www.ecx.eu> (accessed on 5 December 2010).
17. Gao, D.Q. On structure of supervised linear basis function feedforward three-layered neural network. *Chin. J. Comput.* **1998**, *1*, 81–86.
18. GAOTV5. Available online: <http://www.ise.ncsu.edu/mirage/GAToolBox/gaot/gaotv5.zip> (accessed on 16 May 2007).

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).