Expert Systems With Applications 166 (2021) 114097

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Forecasting wavelet neural hybrid network with financial ensemble empirical mode decomposition and MCID evaluation

Yu Yang, Jun Wang *

School of Science, Beijing Jiaotong University, Beijing 100044, PR China

A R T I C L E I N F O

Keywords:
hybrid neural network prediction model
Energy market
Ensemble empirical mode decomposition
Multiscale complexity invariant distance

A B S T R A C T

By considering the properties of nonlinear data and the impact of historical data, this paper combines ensemble empirical mode decomposition (EEMD) into wavelet neural network with random time effective (WNNRT) to establish a hybrid neural network prediction model to improve the prediction accuracy of energy prices The EEMD is a noise-aided data analyze method, since it can effectively suppress pattern confusion and restore signal essence. Different from traditional models, the random time effective function that considers the timeliness of historical data and the random change of market environment is applied to the wavelet neural network to establish the WNNRT model. Moreover, multiscale complexity invariant distance (MCID) is utilized to evaluate the predicting performance of EEMD-WNNRT model. Further, the proposed model which is tested in predicting the impact on the global energy prices has carried on the empirical research, and it has also proved the corresponding superiority.

1. Introduction

Energy is an indispensable material basis for sustainable economic and social development. With the development of economic globalization and financial liberalization, the energy futures market has been developing rapidly, showing an increasingly important trend in the energy market. The erratic fluctuations of energy market prices have increased the investment risks of investors and brought influence to the development of economic and social (He et al., 2010; Huang et al., 2020; Kang et al., 2015). In addition, energy market prices are closely related to other markets, such as the stock market. For example, some scholars have analyzed the relationship between energy prices and stock prices in some countries (Cunado & de Gracia, 2014; Delgado et al., 2018; Fang & You, 2014; Sun et al., 2019; Wen et al., 2012). Thus, the forecast analysis of energy market prices is particularly important. Due to the complex nature of energy prices, forecasting becomes more difficult. Efforts to improve forecasting accuracy have been the focus of research by many scholars in recent years (Cen & Wang, 2019; Wang & Wang, 2016).

In terms of time series forecasting, artificial neural network has made great progress, which is a mathematical method of simulating human actual neural networks (Xu et al., 2019; Yao, 1999; Zhang et al., 1998). Generally, the neural network construction model is divided into the following two cases: forward modeling and reverse modeling. So far, the number of neural network models has been about forty, such as back-propagation neural networks, convolutional neural networks (Goh, 1995; Kristjanpoller et al., 2014; Lin et al., 2009; Ting et al., 2019). Artificial neural network has been applied to various fields in time series prediction. Adamowski and Chan (2011) put forward a new method based on coupling discrete wavelet transforms and artificial neural networks for groundwater level forecasting applications. Kristjanpoller and Minutolo (2015) applied the artificial neural network to the generalized autoregressive conditional heteroskedasticity method to generate a hybrid model to forecast the gold price volatility. Besides, they extended previous research to the area of oil price volatility and the hybrid model increased the volatility forecasting precision (Kristjanpoller & Minutolo, 2016). Mo et al. (2016) proposed the exponent back propagation neural network to predict the cross-correlations between two financial time series. A novel clustering-enhanced adaptive artificial neural network model was used to forecast 24h-ahead building cooling demand in subtropical areas (Luo, 2020). Due to its strong generalization and learning ability as well as adaptability, in this paper, we use the artificial neural network to improve the forecasting accuracy of energy prices.

Wavelet neural network (WNN) is an artificial neural network established in recent years which is based on wavelet transform (Zhang & Benveniste, 1992; Zhang & Kon, 2017). The wavelet transform uses dilation and translation for multi-scale analysis of signals, furthermore it can efficaciously from time domain and frequency domain of stationary or non-stationary signals to extract information (Ford, 2003;

* Corresponding author.
E-mail addresses: 18121645@bjtu.edu.cn (Y. Yang), wangjun@bjtu.edu.cn (J. Wang).

https://doi.org/10.1016/j.eswa.2020.114097
Received 22 April 2019; Received in revised form 29 August 2020; Accepted 3 October 2020
Available online 6 October 2020

Nomenclature

BRE	Brent Crude Oil
BPNN	Back-Propagation Neural Network
CID	Complexity Invariant Distance
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
MCID	Multiscale Complexity Invariant Distance
NG	Natural Gas
NSW	New South Wales
OXY	Occidental Petroleum Corporation
SNP	China Petroleum Chemical Corporation
WNN	Wavelet Neural Network
WNNRT	Wavelet Neural Network with Random Time Effective
WTI	West Texas Intermediate Crude Oil

Minutolo et al., 2018; Stocchi & Marchesi, 2018). Compared with the forward neural network, WNN has obvious advantages. First of all, the basic element and overall structure of WNN model can avoid blindness compared with the back-propagation neural network and other structures. What is more, WNN has stronger learning ability and higher precision. Generally, for a learning task, WNN structure is simpler, the speed of convergence is quicker, and the precision is greater (Zhang et al., 1995). Wavelet neural network was found to be more effective and robust than multilayer perceptron, which was employed to predict three gasoline properties in 2008 (Balabin et al., 2008). Zhang and Yu (2010) has proposed a prediction model of gold price based on wavelet neural network, which has the higher precision and speed than back-propagation neural network. In recent years, in order to improve the prediction accuracy, more and more researchers combine decomposition method with neural network to build a hybrid model to predict time series (Tan et al., 2018). For instance, an empirical mode decomposition based neural network ensemble learning paradigm was proposed for world crude oil spot price forecasting (Yu et al., 2008). Huang and Wang (2018) combined the discrete wavelet transform and stochastic recurrent wavelet neural network to structure a novel hybrid model to improve the prediction accuracy of energy prices. Wu et al. (2018) proposed a novel model based on ensemble empirical mode decomposition and long short-term memory for forecasting crude oil price.

In fact, there are many factors interacting in the energy futures price series. The traditional prediction model only conducts forecasting based on historical data without considering the behavior of the market, resulting in poor accuracy. Since the early historical data represents the market information at that time, it will also have an impact on the current forecast results. Therefore, both early data and recent data should be considered. But the degree of data impact at different times is also different. In general, closer to the present time of the historical data in the market, the influence is greater. Thus, we use stochastic process theory to express the extent to which historical data influences the market (Niu & Wang, 2013; Yu & Wang, 2012). In the operation of random time-dependent neural network algorithm, each set of historical data will be given different weights due to its appearance time and importance to the late prediction. In addition, be grounded on principle of noise-assisted signal processing, the ensemble empirical mode decomposition method decomposes the original time series into different components with high and low frequency. By considering the properties of nonlinear data and the impact of historical data, this paper puts forward a hybrid forecasting model which is used to predict energy price series called the EEMD-WNNRT forecasting model to improve the prediction accuracy of energy prices. It integrates ensemble empirical mode decomposition (EEMD) into the wavelet neural network with random time effective (WNNRT) model. It provides a more accurate forecasting method for energy prices prediction.

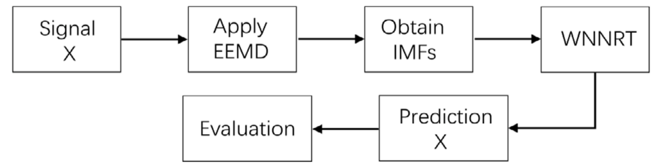


Fig. 1. A simple flowchart of EEMD-WNNRT model.

The rest of the paper is listed as follows. Section 2 presents the algorithm of proposed predictive model. Section 3 introduces the selected data, which are West Texas Intermediate Crude Oil (WTI), Brent Crude Oil (BRE), Occidental Petroleum Corporation (OXY), China Petroleum Chemical Corporation (SNP), Natural Gas (NG), and New South Wales (NSW), and the training and forecasting on these data by the proposed model. In addition, some original models are selected as the comparison models and error assessment methods are used to evaluate the prediction effect of each model. Section 4 uses multiscale complexity invariant distance to evaluate the predicting performance of each model. Finally in Section 5, the conclusions and future works are stated.

2. Methodology

A main prediction system is designed. Firstly, the ensemble empirical mode decomposition (EEMD) method decomposes the original time series into different components with high and low frequency. Then a suitable prediction WNNRT model is constructed for each subseries by adjusting parameters. The predicted results of all the extracted components are combined to generate an aggregated output, and then the final predicted result is obtained. Fig. 1 shows its flow chart.

2.1. Ensemble empirical mode decomposition

Based on the linear and steady-state spectral analysis of Fourier transform, the empirical mode decomposition (EMD) method for signal decomposition is proposed. Moreover, EMD method breaks through the limitation of Fourier transform (Furlaneto et al., 2017; Lei et al., 2013; Oladosu, 2009). Since the EMD decomposes the signal according to the time-scale characteristics of the sequence itself, all signal sequences can be decomposed theoretically. Therefore, for the processing of nonlinear and non-stationary signal sequences, the EMD has a strong advantage and a high signal-to-noise ratio. After the EMD decomposition, the intrinsic mode function (IMF) of the data signal can be obtained, which is composed of local characteristic signals in dissimilar time scales of the initial signal. There are two requirements for each IMF: (i) The amount of zero crossing points is equivalent to or at most one different from the amount of local extreme points; (ii) The mean value of the upper envelope composed of local maxima and the lower envelope composed of local minima is zero.

But this method has problems such as modal aliasing. In view of the shortcomings of EMD, the ensemble empirical mode decomposition (EEMD) method is put forward in recent years (Lei et al., 2011; Wu & Huang, 2009; Zhang et al., 2010). Modal aliasing occurs when the signal distribution is not uniform, and the problem of modal aliasing is solved by adding white noise in EEMD method to automatically separate a signal to its adapted reference scale. The uniform distribution of white noise spectrum can complement some missing dimensions, which makes the signal decomposed better. Although white noise is added, the final calculation result can be directly used as the final result due

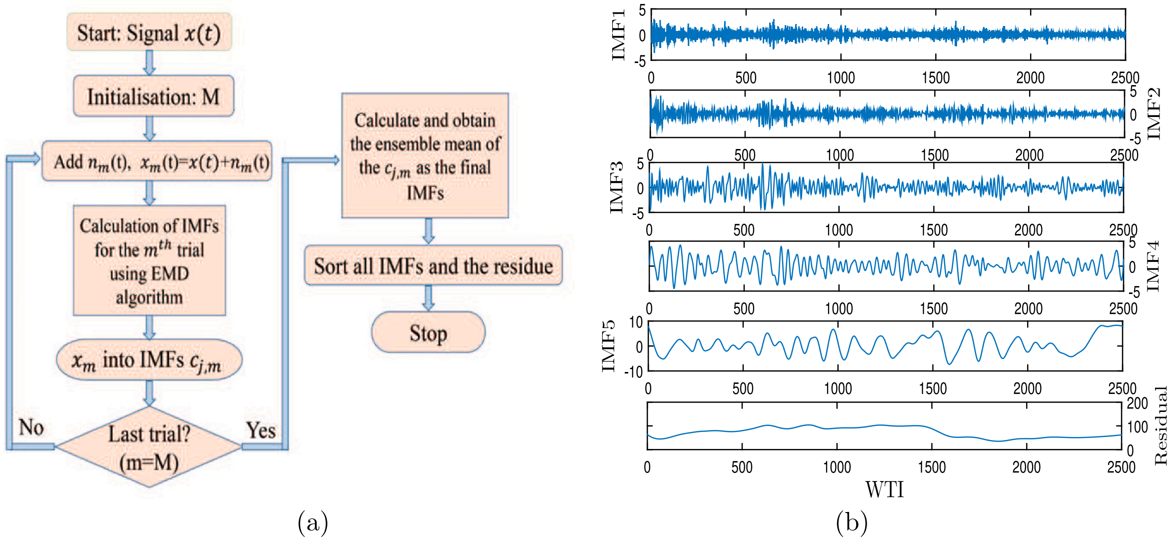


Fig. 2. (a) The flowchart of EEMD algorithm. (b) The corresponding IMFs and the residual of WTI index by the EEMD.

to the nature of zero mean noise. The brief algorithm steps are under following.

Step 1: In the mth test, a sequence with white noise n_m is added to the original signal x to obtain a new signal x_m

$$x_m(t) = x(t) + n_m(t). \tag{1}$$

Step 2: Use the EMD method to decompose the new signal x_m into IMFs $c_{j,m}$

$$x_m(t) \rightarrow \sum_{j=1}^{n} c_{j,m}(t) + r(t) \tag{2}$$

where r is called the surplus of this signal x_m, which are derived after p times of $c_{j,m}$.

Step 3: Repeat Steps 1 and 2 for a preset M of tests, and each test uses varying new white noise sequences which have a consistent amplitude.

Step 4: The integrated IMF of obtained IMFs is used as the final result, m as the final IMFs, as following

$$IMF_j(t) = \frac{1}{M}\sum_{m=1}^{M} c_{j,m}, \quad m = 1,2,\ldots,M, \; j = 1,2,\ldots,p. \tag{3}$$

The core idea of EEMD is to use the statistical characteristics of the white noise, which is uniform frequency distribution. When the white noise signal is added, it will make the signal have continuity at different scales, change the characteristics of the signal extreme point, and promote anti-aliasing decomposition. The flow chart of EEMD algorithm is presented by Fig. 2(a), and the components of IMFs and residual of WTI index by the EEMD is presented by Fig. 2(b). Through the sifting process, it is found that the vibration frequency from IMF1 to IMF6 gradually decreases. The first pattern obtained is IMF1 with the highest frequency, which is also the main structure and feature of WTI. When the data is recalculated through the sifting process, the previous frequency data is higher.

2.2. Wavelet neural network with random time effective

When investing, investors mainly rely on the observation of actual energy security and energy futures markets. Therefore, analyzing historical data is a very effective method. For investors, the most recent information is more important than the distant information. So this time effect is considered in the prediction model of this paper. A random time effective function is established and utilized into wavelet neural network (WNN) to construct a wavelet neural network with random time effective (WNNRT). The structure of the neural network has a parallel distribution structure. When a neuron is stimulated strongly enough, it activates the next neuron to release the transmitter, then activates the next neuron, and so on, ultimately transmitting the signal where it is needed. Artificial neural network consists of a neuron model, and each neuron has an output, which can be connected with other neurons. There are a variety of output connection methods. The wavelet basis function is used to replace traditional sigmoid function as an activation function, and the neural network with no change in structure is called the wavelet neural network. After selecting a specific wavelet, any signal can be mapped to a set of basis functions. These base functions are stretched and translated by the mother wave, and the traversal analysis can be performed along the time axis of the signal at different frequencies while ensuring that the initial information is not lost. Therefore, the WNN overcomes the shortcomings of being prone to partial sub-optimal advantages, and the training speed is also greatly improved. Based on the wavelet analysis theory, mother wavelets can be selected in many forms. The mother wavelet satisfies the below conditions

$$\int H(x)\,dx = 0, \quad H_{a,b}(x) = H\left(\frac{x-b}{a}\right) \tag{4}$$

where a is the scale scaling factor and b is the time shift factor. In this paper, the selected morlet wavelet function has symmetric structure and simple expression. The expression of Morlet wavelet function which is selected is given by the following formula

$$H(x) = \cos(1.75x)\exp\left\{\frac{-x^2}{2}\right\}. \tag{5}$$

A three-layer wavelet neural network with an input layer, a hidden layer and an output layer is constructed as shown in Fig. 3. The inputs are represented as x_{it} ($i = 1,2,\ldots,m$) at time t, and the output is represented as y_{t+1} at time $t+1$. There is also a layer of hidden nerve cells between the inputs and the output as processing units, and the outputs of the layer of hidden nerve cells are represented as z_{jt} ($j = 1,2,\ldots,n$) at time t. Moreover, the weight of this process which from an input layer to a hidden layer, is represented as w_{ij}. So the inputs of the layer of hidden nerve cells are represented as

$$net_{jt} = \sum_{i=1}^{m} w_{ij}x_{it}. \tag{6}$$

The outputs of the layer of hidden nerve cells are denoted by

$$z_{jt} = H_{a,b}(net_{jt}) = H\left(\frac{net_{jt} - b_j}{a_j}\right) \tag{7}$$

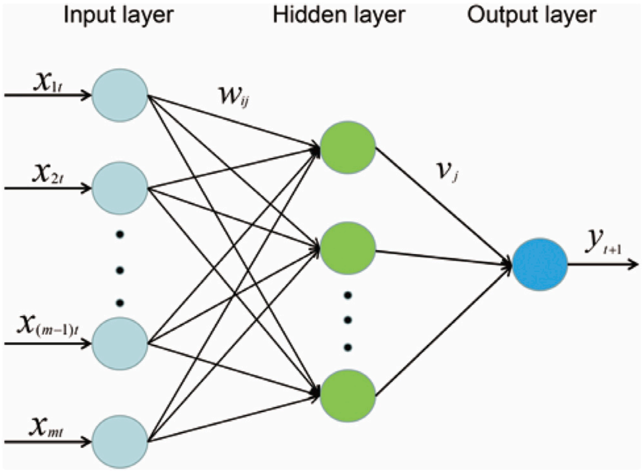


Fig. 3. A three-layer neural network.

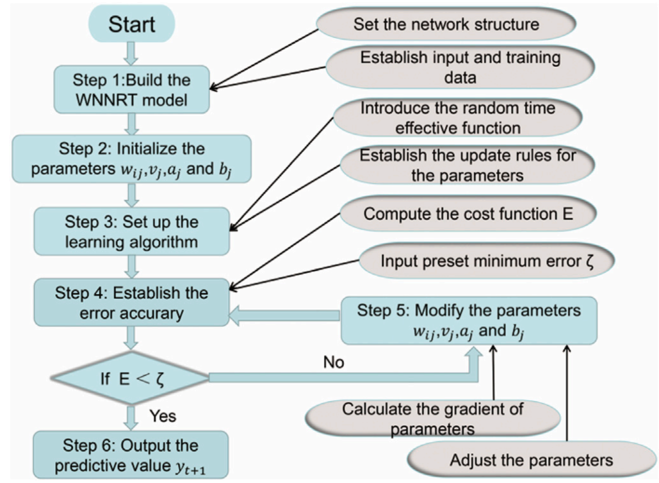


Fig. 4. Training algorithm of WNNRT model.

where a_j is the scale scaling factor and b_j is the time shift factor of the nerve cell j in hidden layer separately. $H(x)$ is Morlet wavelet function and the activation function in hidden layer is represented as $H_{a,b}(x)$ which is the matching wavelet basis function. Moreover, the weight of this process which from an input layer to a hidden layer, is represented as v_j. So the output of this network is expressed as

$$y_{t+1} = \sum_{j=1}^{n} v_j z_{jt} = \sum_{j=1}^{n} v_j H_{a,b}\left(net_{jt}\right). \tag{8}$$

Next, with the view of minimizing the global error E, the argument of the model is updated using the steepest descent method (Liao & Wang, 2010). In this paper, we presume the output error of the WNNRT model is represented as $\varepsilon_{t_k} = d_{t_k} - y_{t_k}$ and the error of the sample k is shown as

$$E\left(t_k\right) = \frac{1}{2}\varphi(t_k)\varepsilon_{t_k}{}^2 = \frac{1}{2}\varphi(t_k)(d_{t_k} - y_{t_k})^2 \tag{9}$$

and one of the time points k ($k = 1, \ldots, N$) is represented by t_k, d_{t_k} represents the reality data, y_{t_k} represents the output data, and $\varphi(t_k)$ represents the random time effective function. As we know, the most recent information is more important than the distant information for the energy market. Therefore, the data of each time is given a different weight by the random time effective function. $\varphi(t_k)$ is represented as the following expression

$$\varphi(t_k) = \frac{1}{\beta}\exp\left\{-\int_{t_k}^{t_0}\mu(t)dt - \int_{t_k}^{t_0}\sigma(t)dB(t)\right\} \tag{10}$$

where $\beta(>0)$ is the time intensity modulus. In this training process, t_0 is the newest data, and t_k is an arbitrary time point. It can be seen that as k increases, the farther t_k is from the current time, the value of random time effective function $\varphi(t_k)$ gradually decreases. Explain that the farther away from the current time, the smaller the effect on the training process. $\mu(t)$ is the drift function, $\sigma(t)$ is the volatility function, and $B(t)$ is the standard Brownian motion. $\mu(t)$ is used to guide the overall trend of the model. And $\sigma(t)$ is used to simulate the occurrence of a series of unpredictable events throughout the event. The phenomenon that suspended particles never stop moving irregularly is called Brownian motion (Haw, 2005; Kac, 1947), which is a stochastic process with continuous time parameters and continuous state space. We presuppose that the most recent data is more important than the distant data in the energy prices market. It can raises the influence of the latest data by the random time strength function $\varphi(t_k)$. Especially in this article, we presume that

$$\mu(t) = \exp\{-\alpha t\}, \quad \sigma(t) = \frac{1}{N-1}\left(\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2\right)^{\frac{1}{2}}. \tag{11}$$

According to the historical prices of energy, α is the parameter set in advance, and $\bar{x}$ is the average value of prices. Therefore, the cost function of all the training data is expressed as

$$E = \frac{1}{N}\sum_{k=1}^{N} E(t_k) = \frac{1}{2N}\sum_{k=1}^{N}\frac{1}{\beta}\exp\{\varphi(t_k)\}\left(d_{t_k} - y_{t_k}\right)^2. \tag{12}$$

The training goal of the neural network with random time effective is to continuously correct the weights so that the global error between the prediction result and the reality data is minimized. In this paper, the weights of w_{ij}, v_j, the scale scaling factor a_j and the time shift factor b_j are continuously updated to optimize the WNNRT model. Specifically, the steepest descent method is used for minimizing the value of E by repeating learning until the global error measures up the preset value of ζ. The gradient expression is $\Delta E = \partial E/\partial W$. The gradient of the weight w_{ij} which is connected at the weight nodes of the input layer is represented as

$$\Delta w_{ij} = -\eta\frac{\partial E(t_k)}{\partial w_{ij}} = \frac{1}{a_j}\eta\varepsilon_{t_k}v_j\varphi(t_k)H'_{a,b}(net_{jt_k})x_{it_k} \tag{13}$$

and the gradient of the weight v_j which is connected at the weight nodes of the hidden layer is denoted by

$$\Delta v_j = -\eta\frac{\partial E(t_k)}{\partial v_j} = \eta\varepsilon_{t_k}\varphi(t_k)H'_{a,b}(net_{jt_k}) \tag{14}$$

and the learning rate is represented as η. $H'_{a,b}(net_{jt_k})$ is the derivative of the activation function for this model. The gradient of a_j is expressed as

$$\Delta a_j = -\eta\frac{\partial E(t_k)}{\partial a_j} = -\frac{1}{a_j{}^2}\eta\varepsilon_{t_k}v_j\varphi(t_k)H'_{a,b}(net_{jt_k})\left(net_{jt_k} - b_j\right) \tag{15}$$

and the gradient of b_j is represented as

$$\Delta b_j = -\eta\frac{\partial E(t_k)}{\partial b_j} = -\frac{1}{a_j}\eta\varepsilon_{t_k}v_j\varphi(t_k)H'_{a,b}(net_{jt_k}). \tag{16}$$

So the update rule for these weights are expressed as

$$w_{ij}^{n+1} = w_{ij}^{n} + \Delta w_{ij}^{n} = w_{ij}^{n} + \eta\varepsilon_{t_k}v_j\varphi(t_k)H'_{a,b}(net_{jt_k})x_{it_k} \tag{17}$$

In addition, v_{ij}^{n+1}, a_j^{n+1}, and b_j^{n+1} are calculated in the same way as w_{ij}^{n+1}.

Combining the above knowledge, Fig. 4 shows the algorithm of this neural network.

2.3. The hybrid EEMD-WNNRT forecasting model

From the above theory of EEMD and WNNRT model, a new hybrid modeling EEMD-WNNRT is conducted to predict energy market prices.

Table 1
Distinct subseries with distinct parameters.

Index	WTI		BRE		OXY		SNP		NG		NSW	
Parameter	η	Hidden	η	Hidden	η	Hidden	η	Hidden	η	Hidden	η	Hidden
IMF1	0.01	10	0.05	8	0.02	8	0.02	8	0.01	10	0.05	8
IMF2	0.01	10	0.02	10	0.05	10	0.05	10	0.05	10	0.05	10
IMF3	0.02	8	0.01	10	0.07	9	0.01	8	0.01	9	0.03	8
IMF4	0.05	9	0.05	8	0.05	8	0.05	8	0.02	10	0.05	8
IMF5	0.02	10	0.02	8	0.07	8	0.03	9	0.05	10	0.06	9
Residual	0.005	8	0.001	8	0.002	10	0.001	10	0.02	10	0.07	10

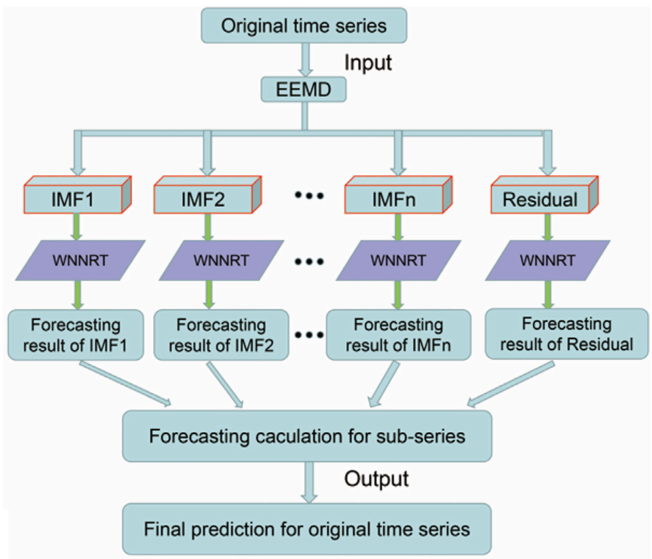


Fig. 5. A detailed flowchart of EEMD-WNNRT model.

Through this work, a new simple hybrid forecasting model with high prediction accuracy can be found for energy market. Then, a brief introduction to the methods used in EEMD-WNNRT is given, and Fig. 5 shows its detailed flowchart.

(i) At the first, input the original energy prices time series $X(t)$, and decompose them into IMF components $C_i(t)$ $(i = 1, 2, \ldots, M)$ and one residual component $R_M(t)$ by the EEMD technique.

(ii) In the next part, establish the most appropriate WNNRT model for each component of IMFs and residual and achieve predictions.

(iii) In the end, the predicted results of all the extracted components are amalgamation to be produced a total output, and then the final predicted result is obtained. The aggregation calculation formula of the subseries is represented as

$$\tilde{X}(t) = \sum_{i=1}^{M} \tilde{C}_i(t) \times \lambda_i + \tilde{R}_M(t) \times \lambda_{M+1}, \quad t = 1, \ldots, N \tag{18}$$

where λ is the overall modulus for every subseries, $C_i(t)$ is the prediction of series of the IMF components, $R_M(t)$ is the prediction of the residual component, and $X(t)$ is the prediction outcome of the energy prices series. And it is assumed $\lambda_i = 1$ $(i = 1, 2, \ldots, M)$.

3. Forecasting and experiment analysis

3.1. Selecting and preprocessing of the data

In this section, we pick out the energy prices series from West Texas Intermediate (WTI), Brent crude oil (BRE), Occidental Petroleum Corp. (OXY), China Petroleum Chemical Corp. (SNP), Natural Gas (NG), and New South Wales (NSW), for training to achieve the performance evaluation of the EEMD-WNNRT prediction model. Currently, there is no widely accepted method to use rigorous training data to build the best predictive model. By comparing the predicted results of samples, we choose the best architecture of the network. Usually, non-trading hours are considered frozen, so we only use the time of the trading period. The empirical research selects 2500 data for each of four oil prices, which are WTI, BRE, OXY and SNP, during 10-year period from 15/10/2008 to 15/06/2018. For natural gas, the research selects 2500 data during 10-year period from 21/12/2010 to 30/07/2020. And the data sets have two parts, a set of training and a set of test, for verification. The training set has 2000 data and the remaining 500 is used as test set data. Specifically, NSW is the price of the electricity of New South Wales, Australia. The time resolution of NSW is 30 min, and the total number of observation values of this dataset is 1478, covering 31 days from 01/07/2020 to 31/07/2020. Moreover, the data from 25 days are used to predict and the data from 6 days to test, which means the number of training data points is 1200 (from 01/07/2020 to 25/07/2020) and the number of testing data points is 278 (from 26/07/2020 to 31/07/2020). Generally, data are standardized before analysis. At present, there are many methods for data standardization, which can be divided into a linear method, a broken line method, and a curved method. Here we normalize the data, that is, the data is uniformly mapped to the $[0, 1]$ interval, which can improve the convergency rate and accuracy of a model (Liu, 2010; Wu & Lo, 2010). The conversion function is as follows

$$S(t)' = \frac{S(t) - \min S(t)}{\max S(t) - \min S(t)}. \tag{19}$$

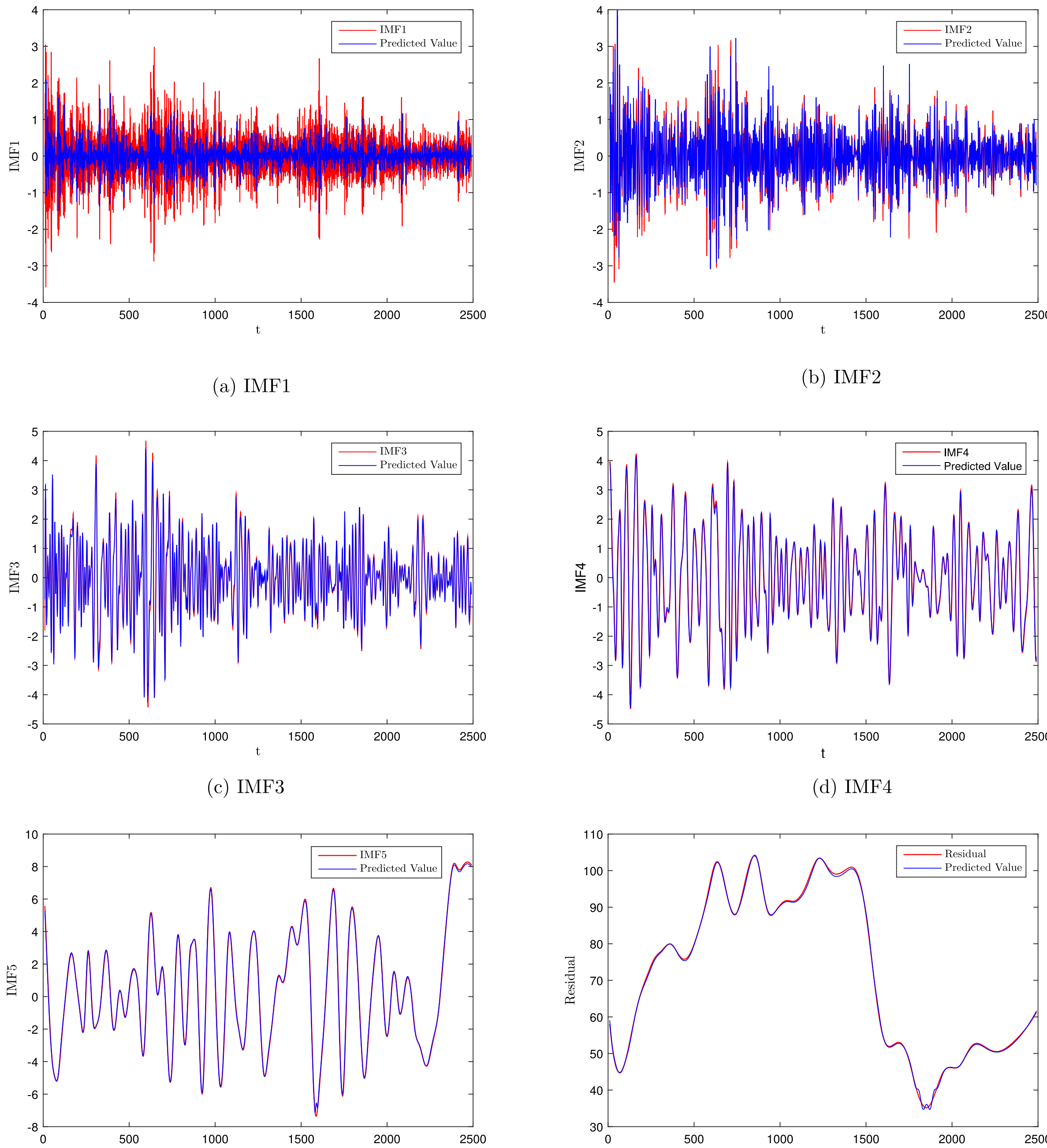
The maximum and minimum values are procured. Then we restore the output variable to the following expression to get the true value after the prediction, that is

$$S(t) = S(t)'(\max S(t) - \min S(t)) + \min S(t). \tag{20}$$

Thereafter, a non-stationary data is transmitted to the network.

3.2. Training and forecasting by EEMD-WNNRT model

In the hybrid EEMD-WNNRT model, firstly we use a classical algorithm of EEMD method to process the original time series. And each index can obtain five IMF components and a residue component. Then every subseries is used for building a suitable WNNRT model for prediction. When establishing the new hybrid predictive EEMD-WNNRT model, according to the rule of selecting the number of hidden layers, that is, if the amount of neural nodes in the input layer is N, the amount of neural nodes in the hidden layer is set to be close to $2N + 1$, and the amount of neural nodes in the output layer is 1 (Han, 2002). In this paper, we select the number of neural nodes in the input as 4 and the number of neural nodes in the output to be 1. And we set the value of the maximum iterations number K to 500 and the value of the minimum error threshold ζ to 10^{-5}. The best prediction model is established for different subseries by adjusting the learning rates and the number of hidden layer nodes. The learning rates ranges from 0.001 to 0.1, and the amount of hidden layer nodes ranges from 6 to 12. The choice of model is on the basis of trial and error, is also on the basis of accumulated experience and expertise of the research scholars. In the repeated experimental results, generally, the parameters of the best prediction results are set as the optimal parameters of the neural network model. We use four different measures (MAE, MAPE, MAPE100,



(a) IMF1 (b) IMF2

(c) IMF3 (d) IMF4

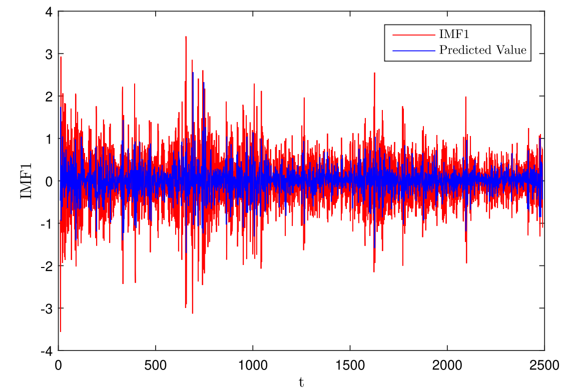
(e) IMF5 (f) Residual

Fig. 6. The prediction data and the real data of IMF1, IMF2, IMF3, IMF4, IMF5 and the residual for WTI.

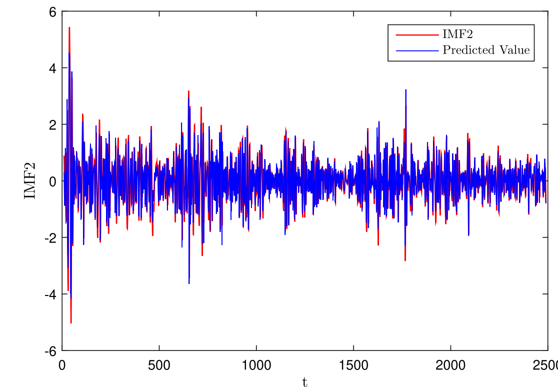
RMSE) to determine the architecture structure which validation set has the best performance at the end. In Table 1, the best parameters chosen are given for all subseries of each index, where the learning rate is represented by η, and the amount of neural nodes in the hidden layer is represented by 'Hidden'.

The data sets of energy prices have been decomposed into a series of distinct subseries by EEMD method, and the corresponding WNNRT prediction model has been established for each subseries. Figs. 6 and 7 respectively show the comparison between the real data and the predicted data of components of IMFs and residual for each subseries of WTI and BRE. To save space, the prediction comparison diagrams for the subseries of OXY, SNP, NG, and NSW are no longer listed here. It can be clearly seen from Figs. 6 and 7 that IMF1 and IMF2 have higher fluctuation frequencies compared with other IMFs', and the predicted results of IMF3, IMF4, IMF5 and the residual are very close to the real data. At one time, it shows that the WNNRT prediction model can get ideal prediction output, and the corresponding prediction value is very near to the real data intuitively. The prediction outcomes of every subseries of each index are weighted and combined to obtain the final prediction result of the index.

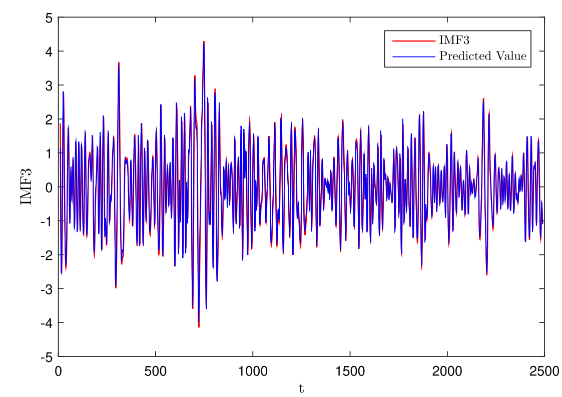
Fig. 8(a)(c)(e)(g)(i)(k) respectively show the prediction outcomes of WTI, BRE, OXY, SNP, NG, and NSW obtained by the prediction model EEMD-WNNRT. It can be seen from these figures that the fluctuations of the predicted data and the real data are exactly the same. In addition, the prediction results of the established EEMD-WNNRT model are compared with the actual data by linear regression analysis. The actual data is represented on the x-axis and the prediction data is represented on the y-axis. So the corresponding linear equation is $y = ax + b$. The



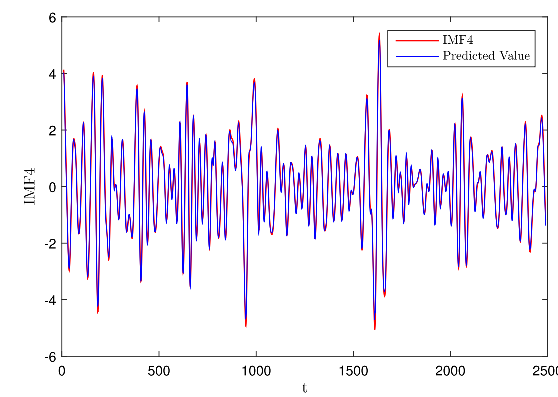
(a) IMF1



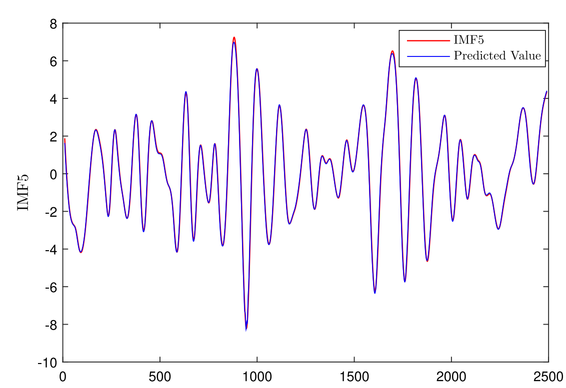
(b) IMF2



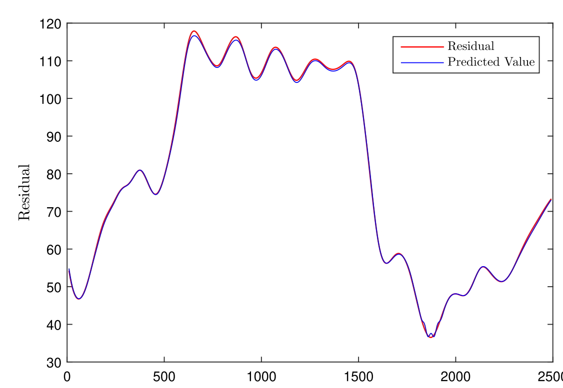
(c) IMF3



(d) IMF4



(e) IMF5



(f) Residual

Fig. 7. The prediction data and the real data of IMF1, IMF2, IMF3, IMF4, IMF5 and the residual for BRE.

linear regression curves of WTI, BRE, OXY, SNP, NG and NSW are respectively shown in Fig. 8(b)(d)(f)(h)(j)(l). The specific values of a, b are also given in Table 2. Besides the data of NSW, it can be concluded by these figures that the curve of the predicted data is near to the curve of the real data. It also can be observed that the slope value a of the linear equation is very close to 1, that is, the deviation between the predicted value and the actual value is not large. It can be explained that through the repeated experiments, the EEMD-WNNRT model can well predict the prices time series of crude oil and natural gas, and the results are also in the ideal range. Due to the nonstationary of the price of electricity and the existence of peak prices of electricity, forecasting the price of electricity with perfect accuracy is more difficult than other indexes. But it can be seen that the prediction of the volatility of the price of electricity is in line with the real data. Therefore, The EEMD-WNNRT model can play a certain role in the prediction of the price of electricity. This can draw a conclusion that the EEMD-WNNRT model has a powerful generalization ability to train the energy time series well. It exhibits the rationality of EEMD-WNNRT model.

Table 2
Linear regression parameters of market indexes.

Index	WTI	BRE	OXY	SNP	NG	NSW
a	0.9926	0.9875	0.9774	0.9604	0.9699	0.8265
b	0.3839	0.6851	6.2507	2.9070	0.0897	10.3395

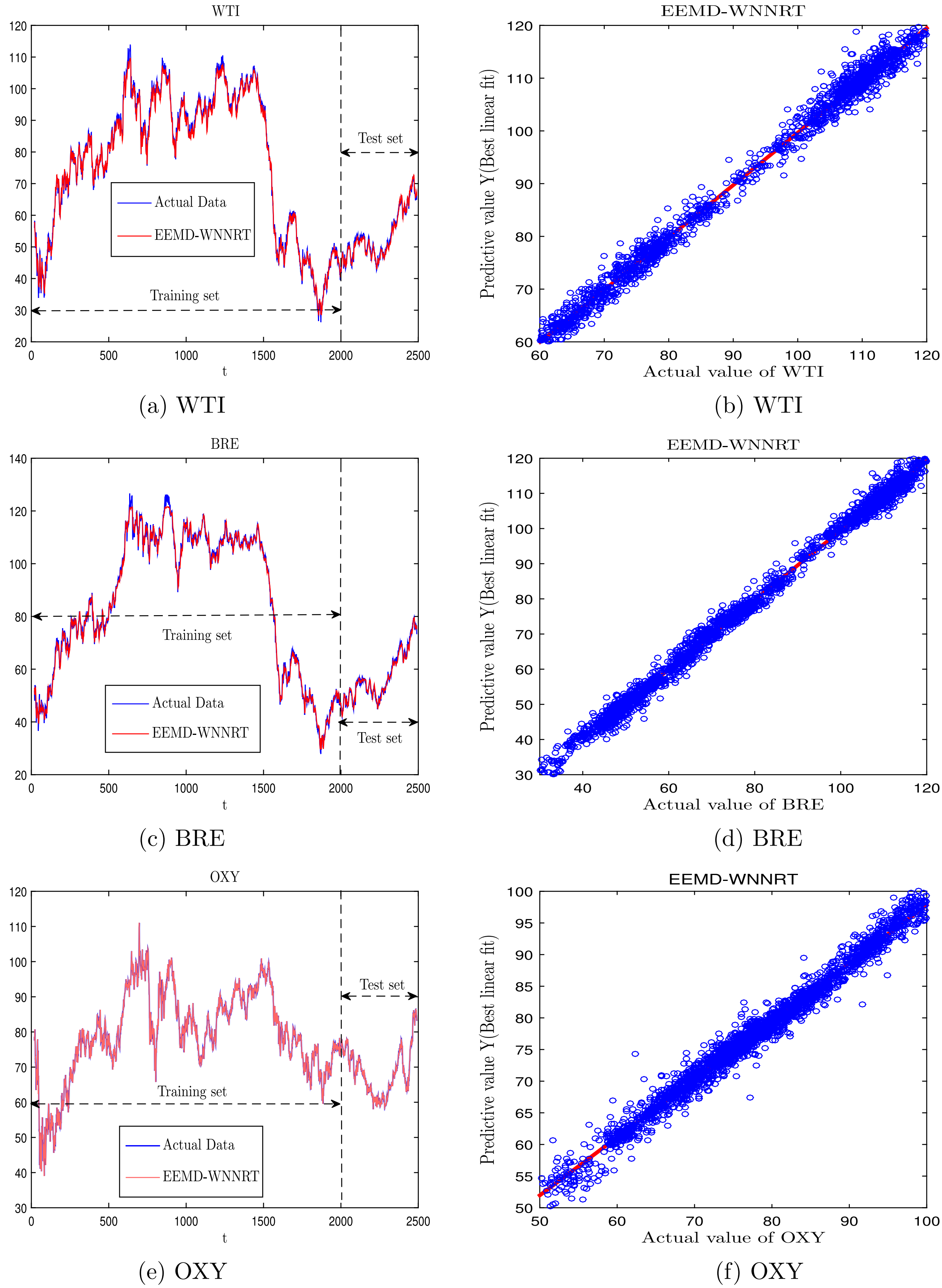


Fig. 8. Comparisons of prediction data and real data for the prediction models.

Within some certain errors, under the prediction of EEMD-WNNRT model, we calculate the accurately prediction days and probability of each index. The results are shown in Table 3. It can be seen from Table 3 that the prediction effects of crude oil and natural gas are

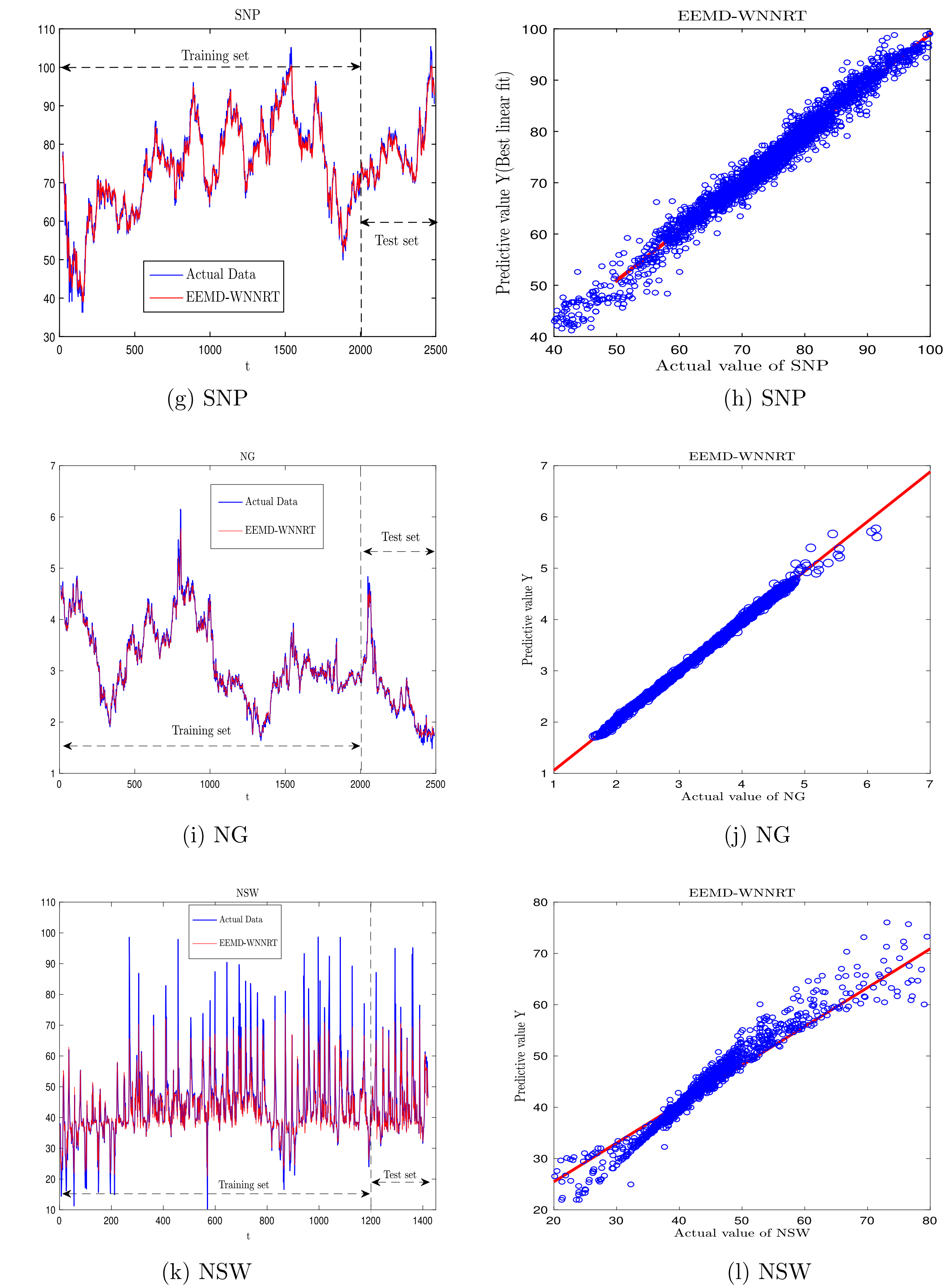


Fig. 8. (*continued*).

similar, and the prediction effect of the price of electricity can be further improved. For energy futures price series, an accuracy of over 90% can be achieved with an error of less than 0.1. Therefore, the EEMD-WNNRT model has a better effect on the prediction of energy futures price series.

Table 3
Prediction accuracy of each index under the proposed model.

Index	WTI				
P. E.	0.041	0.046	0.052	0.063	0.09
A. P. D.	2002	2133	2251	2372	2491
A. P. P.	80.08%	85.32%	90.04%	94.88%	99.64%
Index	BRE				
P. E.	0.049	0.055	0.06	0.075	0.13
A. P. D.	2005	2137	2226	2372	2500
A. P. P.	80.20%	85.48%	89.04%	94.88%	100%
Index	OXY				
P. E.	0.0225	0.025	0.028	0.035	0.06
A. P. D.	2004	2122	2237	2382	2500
A. P. P.	80.16%	84.88%	89.48%	95.28%	100%
Index	SNP				
P. E.	0.074	0.084	0.097	0.118	0.18
A. P. D.	1994	2121	2247	2383	2495
A. P. P.	79.76%	84.84%	89.88%	95.32%	99.80%
Index	NG				
P. E.	0.047	0.055	0.064	0.084	0.3
A. P. D.	2015	2137	2245	2376	2491
A. P. P.	80.60%	85.48%	89.80%	95.04%	99.64%
Index	NSW				
P. E.	1.8	2.3	3.2	6	15
A. P. D.	1134	1213	1276	1346	1402
A. P. P.	79.86%	85.42%	89.86%	94.79%	98.73%

Note: P. E. represents the prediction errors; A. P. D. represents the accurately prediction days;
A. P. P. represents the accurately prediction probability.

3.3. Comparisons of forecasting results

In this section, six indexes WTI, BRE, OXY, SNP, NG, and NSW taken above are trained and predicted under different prediction models, and the prediction effects are compared. The selected model are the EEMD-WNNRT model, the WNNRT model, the WNN model and the BPNN model. For different index sequences, the amount of neurons in the hidden layer is different under different neural network models, and the choice of learning rate is also different. The optimal parameters of each model for each index are obtained through a series of repeated experiments. The detailed parameters selected are given in Table 4. We selected the following error assessment criteria to evaluate the prediction effects of four prediction models: mean absolute error (MAE), root mean square error (RMSE), and mean absolute percent error (MAPE). These are measures of the deviation between predicted and actual values. The smaller these values, the better the prediction performance. When the results in these standards are inconsistent, the relatively stable MAPE can be seen as a benchmark. The corresponding definitions are represented as following

$$\mathrm{MAE}=\frac{1}{N}\sum_{t=1}^{N}|d_t-y_t|,\quad \mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{t=1}^{N}(d_t-y_t)^2} \tag{21}$$

$$\mathrm{MAPE}=100\times\frac{1}{N}\sum_{t=1}^{N}\left|\frac{d_t-y_t}{d_t}\right| \tag{22}$$

where d_t is the expression of the actual value, and y_t is the expression of predict value at time t. The total number of the data is represented by N.

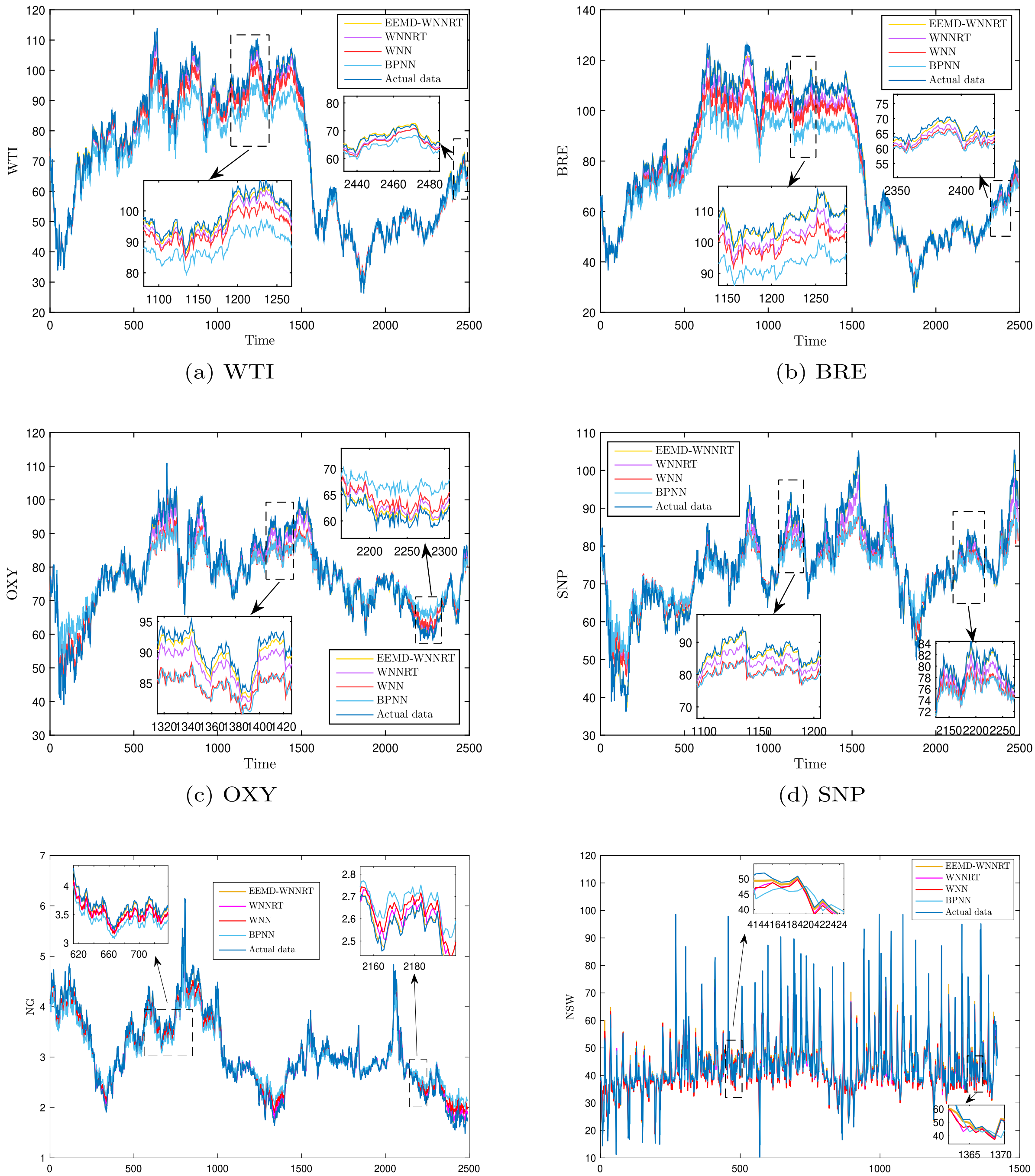
Table 5
Statistical indicators of predicting performance for distinct prediction models.

Errors	MAE	MAPE	MAPE100	RMSE
Index	WTI			
BPNN	1.6089	2.8916	2.8648	1.9019
WNN	1.2456	2.3148	2.6346	1.5859
WNNRT	1.0232	1.9082	2.5505	1.3037
EEMD-WNNRT	**0.9686**	**1.8174**	**2.3650**	**1.1716**
Index	BRE			
BPNN	2.3637	3.7633	2.0161	3.0278
WNN	2.1100	3.4049	2.0134	2.6115
WNNRT	1.5728	2.5989	1.9891	1.9019
EEMD-WNNRT	**1.0866**	**1.9123**	**1.7128**	**1.3688**
Index	OXY			
BPNN	3.2068	4.9372	1.9222	3.7672
WNN	1.9439	2.9259	1.6017	2.2353
WNNRT	1.7871	2.6741	1.5076	2.0911
EEMD-WNNRT	**0.8174**	**1.2010**	**1.1230**	**1.0637**
Index	SNP			
BPNN	3.3962	4.0024	1.3789	4.6971
WNN	3.1831	3.7204	1.2190	4.6245
WNNRT	1.8078	2.1493	1.1427	2.4925
EEMD-WNNRT	**0.9825**	**1.2100**	**1.1060**	**1.4032**
Index	NG			
BPNN	0.2147	10.2105	5.0199	0.2659
WNN	0.1226	6.5833	1.5731	0.3490
WNNRT	0.0741	3.6957	1.6285	0.1519
EEMD-WNNRT	**0.0382**	**1.9229**	**0.8892**	**0.0743**
Index	NSW			
BPNN	5.1826	9.9075	10.1949	8.7789
WNN	2.3725	4.3614	4.2464	4.8344
WNNRT	1.9033	3.3353	3.3525	4.7498
EEMD-WNNRT	**1.6453**	**2.8222**	**2.7720**	**4.1778**

After many experiments, the optimal prediction model is constructed for each data to obtain the prediction results. The prediction effects of six indexes WTI, BRE, OXY, SNP, NG, and NSW in the four prediction models are shown in Fig. 9. We can observe that the EEMD-WNNRT mode is superior on prediction effect and its forecast data is closest to the real data. It also proves that the prediction effect of WNN model is better than the common BPNN model. The results of several error assessments are also given in Table 5, where MAPE (100) represents the MAPE of the last 100 days. From the data in the table, it can be seen that the evaluation criteria of the EEMD-WNNRT model are smaller than the evaluation criteria of other models. And the error assessment criteria of the BPNN model are almost greater than the other three models, which also indicates that the prediction of WNN model is better. In the comparison of the four models, we can find that for the complex NSW index, the prediction error of the EEMD-WNNRT model is the smallest and the prediction effect is the best. So it can be observed from Fig. 9 and Table 5 that the prediction effect of EEMD-WNNRT model is the best among the four models. For these six indicators, the prediction effect of the WNN model is better than the BPNN model. By considering the properties of nonlinear data and the impact of historical data, more factors are considered to make the prediction effect more consistent with the actual situation so as to improve the prediction accuracy.

Table 4
Distinct parameters for distinct NN models.

Index	WTI		BRE		OXY		SNP		NG		NSW	
Parameter	η	Hidden	η	Hidden	η	Hidden	η	Hidden	η	Hidden	η	Hidden
BPNN	0.01	8	0.04	8	0.09	8	0.01	8	0.01	8	0.01	8
WNN	0.05	8	0.02	10	0.001	8	0.08	8	0.01	8	0.05	10
WNNRT	0.001	10	0.02	8	0.001	10	0.05	10	0.05	10	0.01	10



(a) WTI (b) BRE

(c) OXY (d) SNP

(e) NG (f) NSW

Fig. 9. Prediction values with distinct models for WTI, BRE, OXY, SNP, NG and NSW.

4. Evaluation of multiscale CID analysis

Next, using another error analysis method to more verify the accuracy of the EEMD-WNNRT model. We use the multiscale complexity invariant distance (MCID) method of analyzing multiple time scales to compare prediction results. Firstly, we introduce the Euclidean distance, which is the basis of this method. The Euclidean distance is a measure which can solve many problems well (Batista et al., 2014, 2011). Suppose there are two time series, P and Q, of length n

$$P = p_1, p_2, \ldots, p_i, \ldots, p_n, \qquad Q = q_1, q_2, \ldots, q_i, \ldots, q_n. \tag{23}$$

The Euclidean distance is represented as

$$ED(P, Q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}. \tag{24}$$

But most of the data will be distorted in some cases, losing its authenticity. Eliminate distortion or use a more robust measurement method before using Euclidean distance. The original definition of complexity invariance is to use the information of complexity difference between two sets of series as a correction factor of the existing distance measure. The complexity correction factor is represented by CF, which further separates time series with different complexity. The complexity

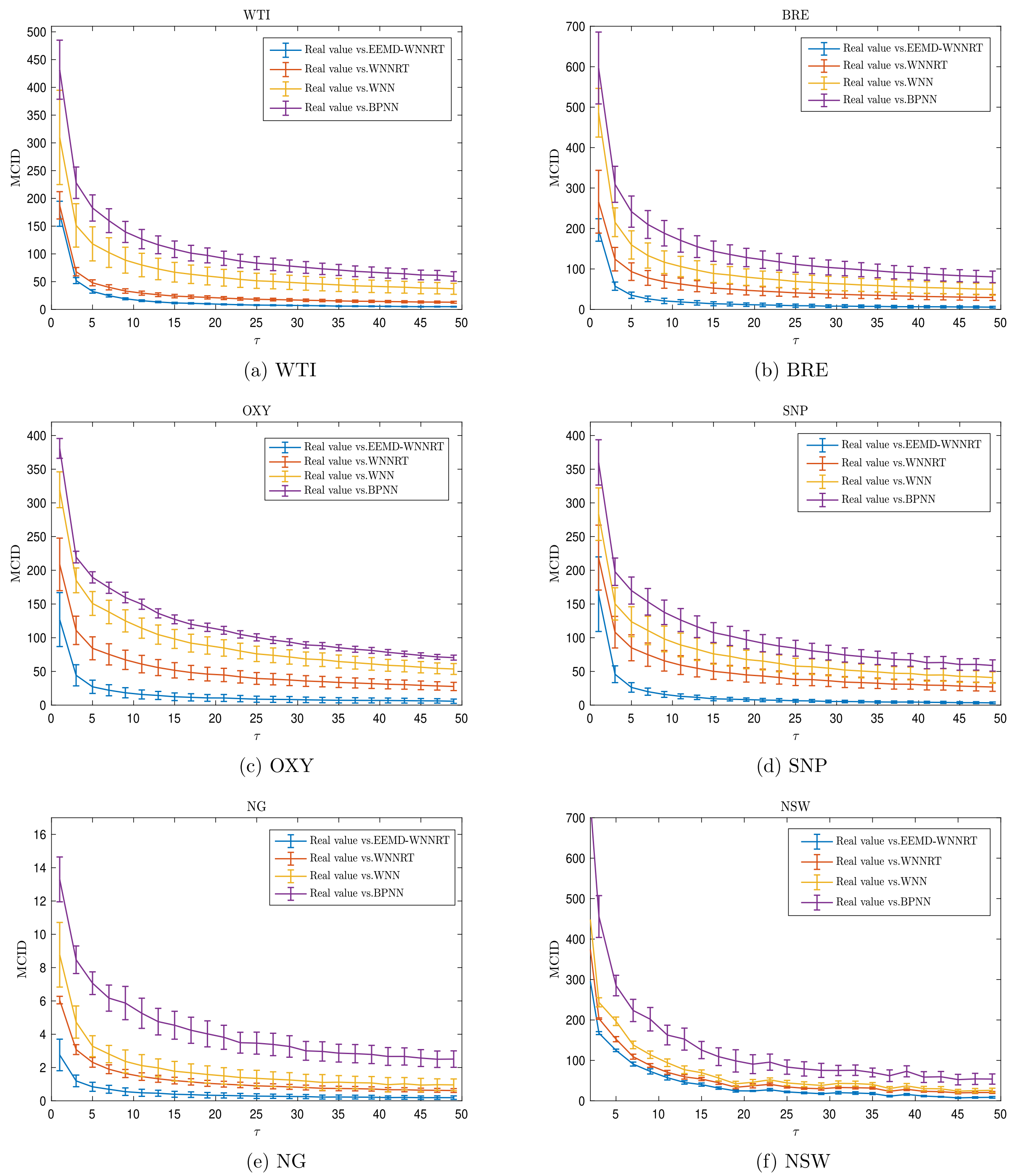


Fig. 10. MCID plots of real market indexes and prediction outcomes of BPNN, WNN, WNNRT and EEMD-WNNRT models.

estimate for time T is represented by $CE(T)$. So we can introduce a correction factor into two series P and Q, which makes $ED(P,Q)$ complexity invariant. The expressions are represented below

$$CE(T) = \sqrt{\sum_{i=1}^{n-1}(t_{i+1} - t_i)^2}, \quad CF(P,Q) = \frac{\max\{CE(P), CE(Q)\}}{\min\{CE(P), CE(Q)\}} \tag{25}$$

$$CID(P,Q) = ED(P,Q) \times CF(P,Q). \tag{26}$$

In particular, when the time series complexity is the same, the CID is the Euclidean distance. The smaller the value of CID, the smaller the distance between the two sets of time series, that is, the predicted data is nearer to the real data.

Considering that a reality complexity is generally multi-scale, this paper compares the prediction performance of several models by considering the multiscale CID (MCID) method of multiple time scales. The analysis of MCID mainly consists of the under two procedures: (i) Constructing the consecutive coarse-grained time series $y^{(\tau)}$, corresponding to the scaling factor τ, for the one-dimensional discrete time series $\{x_1, x_2, \ldots, x_i, \ldots, x_N\}$. The expression is represented as following

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \le j \le \frac{N}{\tau}. \tag{27}$$

The time series $y^{(1)}$ is the original time series, when the scaling factor is 1. Divide the original time series by the scaling factor τ to get the length of every coarse-grained time series. (ii) The CID for every coarse-grained time series is calculated and then plotted as a function of the scaling factor. The error bars of the MCID values between the predicted data of six indexes in BPNN, WNN, WNNRT and EEMD-WNNRT models and the real energy data are given in Fig. 10. It can be seen that the values of MCID in the EEMD-WNNRT model are the smallest, that is, the predicted data of this model is the closest to the real data. It can be seen from the above results that for six indexes WTI, BRE, OXY, SNP, NG, and NSW, the EEMD-WNNRT model is superior to the other three models and has higher prediction accuracy. Once again this proves the superiority of the proposed model.

5. Conclusion

As the huge fluctuations in energy prices will always have an impact on the global economy, the forecast of energy prices is particularly important. In this work, a hybrid prediction model EEMD-WNNRT is constructed by adding the ensemble empirical mode decomposition to the wavelet neural network model of random time effective function. The EEMD method is used to decompose the energy price sequences to obtain different IMF components and a residual component. Then a suitable prediction model is constructed for each subseries by adjusting parameters. The data has been simplified and the random time effective function is made, so that the prediction effect is fast and accurate. Moreover, by means of linear regression analysis, error evaluation criteria, and the MCID method, we demonstrate the predictive superiority of the proposed model in comparison with other models selected for WTI, BRE, OXY, SNP, NG, and NSW indexes.

I. Linear regression analysis is used to verify the prediction effect and accuracy of the proposed EEMD-WNNRT model. It can be seen from Table 2 that the slopes of the curves of energy futures price series, such as WTI, BRE, OXY, SNP and NG indexes, are close to 1, which shows the rationality of EEMD-WNNRT model. In Table 3, an accuracy of over 90% can be achieved with an error of less than 0.1 for energy futures price series. Therefore, the EEMD-WNNRT model has a better effect on the prediction of energy futures price series.

II. In Table 5, Figs. 9 and 10, through error evaluation criteria and the MCID method which is an effective method to reveal the performance of prediction, it can be seen that the values predicted by EEMD-WNNRT model, which is compared with WNNRT, WNN and BPNN models, are the smallest. These values can exhibit the superiority of the proposed model. Therefore, the EEMD-WNNRT model has the best prediction effect for six selected indexes among the four models.

III. By introducing data decomposition method and random time effective function to construct a new prediction model, the prediction accuracy is improved. Among the three wavelet neural networks, the prediction outcomes of EEMD-WNNRT model are nearer to the real price of energy market. Thus, the research of neural network models in the time series prediction of energy markets is enriched. At the same time, accurate prediction of volatility can help energy market participants adjust their operations in advance. It can reduce investor risks and promote the healthy and stable development of energy market. However, for the data of higher complexity and nonstationary, such as the price of electricity, which have low periods and peak periods, the prediction effect is not very accurate compared with energy futures prices. In the future, we will continue to improve the data decomposition processing and parameters selection of the model to apply it more widely and accurately to time series forecasting.

CRediT authorship contribution statement

Yu Yang: Conceptualization, Methodology, Investigation, Software, Visualization, Writing - original draft, Writing review & editing. **Jun Wang:** Conceptualization, Methodology, Investigation, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Authors were supported by National Natural Science Foundation of China Grant No. 71271026.

References

Adamowski, J., & Chan, H. F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology Amsterdam*, *407*, 28–40.

Balabin, R. M., Safieva, R. Z., & Lomakina, E. I. (2008). Wavelet neural network (wnn) approach for calibration model building based on gasoline near infrared (nir) spectra. *Chemometrics and Intelligent Laboratory Systems*, *93*, 58–62.

Batista, G. E., Keogh, E. J., Tataw, O. M., & De Souza, V. M. (2014). CID: an efficient complexity-invariant distance for time series. *Data Mining and Knowledge Discovery*, *28*, 634–669.

Batista, G. E., Wang, X., & Keogh, E. J. (2011). A complexity-invariant distance measure for time series. In *Proceedings of the 2011 SIAM international conference on data mining.* (pp. 699–710).

Cen, Z., & Wang, J. (2019). Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer. *Energy*, *169*, 160–171.

Cunado, J., & de Gracia, F. P. (2014). Oil price shocks and stock market returns: Evidence for some european countries. *Energy Economics*, *42*, 365–377.

Delgado, N. A. B., Delgado, E. B., & Saucedo, E. (2018). The relationship between oil prices, the stock market and the exchange rate: evidence from Mexico. *The North American Journal of Economics and Finance*, *45*, 266–275.

Fang, C. R., & You, S. Y. (2014). The impact of oil price shocks on the large emerging countries' stock prices: Evidence from China, India and Russia. *International Review of Economics & Finance*, *29*, 330–338.

Ford, M. S. (2003). The illustrated wavelet transform handbook: Introductory theory and applications in science. *Health Physics*, *84*, 667–668.

Furlaneto, D. C., Oliveira, L. S., Menotti, D., & Cavalcanti, G. D. (2017). Bias effect on predicting market trends with EMD. *Expert Systems with Applications*, *82*, 19–26.

Goh, A. T. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, *9*, 143–151.

Han, L. Q. (2002). *Theory, design and application of artificial neural network*. Beijing: Chemical industry publishing company.

Haw, M. (2005). Einstein's random walk. *Physics World*, *18*, 19.

He, Y., Wang, S., & Lai, K. K. (2010). Global economic activity and crude oil prices: A cointegration analysis. *Energy Economics*, *32*, 868–876.

Huang, J. S., Li, W. P., Guo, L. J., Hu, X., & Jim, W. H. (2020). Renewable energy and household economy in rural China. *Renewable Energy*, *155*, 669–676.

Huang, L., & Wang, J. (2018). Forecasting energy fluctuation model by wavelet decomposition and stochastic recurrent wavelet neural network. *Neurocomputing*, *309*, 70–82.

Kac, M. (1947). Random walk and the theory of Brownian motion. *American Mathematical Monthly*, *54*, 369–391.

Kang, W., Ratti, R. A., & Yoon, K. H. (2015). The impact of oil price shocks on the stock market return and volatility relationship. *Journal of International Financial Markets, Institutions and Money*, *34*, 41–54.

Kristjanpoller, W., Fadic, A., & Minutolo, M. C. (2014). Volatility forecast using hybrid Neural Network models. *Expert Systems with Applications*, *41*, 2437–2442.

Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the Artificial Neural NetworkGARCH model. *Expert Systems with Applications*, *42*, 7245–7251.

Kristjanpoller, W., & Minutolo, M. C. (2016). Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Systems With Applications*, *65*, 233–241.

Lei, Y., He, Z., & Zi, Y. (2011). EEMD Method and WNN for fault diagnosis of locomotive roller bearings. *Expert Systems with Applications*, *38*, 7334–7341.

Lei, Y., Lin, J., He, Z., & Zuo, M. J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, *35*, 108–126.

Liao, Z., & Wang, J. (2010). Forecasting model of global stock index by stochastic time effective neural network. *Expert Systems with Applications*, *37*, 834–841.

Lin, C. L., Wang, J. F., Chen, C. Y., Chen, C. W., & Yen, C. W. (2009). Improving the generalization performance of RBF neural networks using a linear regression technique. *Expert Systems with Applications*, *36*, 12049–12053.

Liu, X. T. (2010). Study on data normalization in BP neural network. *Mechanical Engineering and Automation*, *3*, 122–123.

Luo, X. J. (2020). A novel clustering-enhanced adaptive artificial neural network model for predicting day-ahead building cooling demand. *Journal of Building Engineering*, *32*, Article 101504.

Minutolo, M. C., Sierra, A., & Kristjanpoller, W. (2018). Short and long-term dynamic dependencies of main latin American stock indexes and commodity prices : A wavelet approach. *Journal of Accounting and Finance*, *18*(5), 162–178.

Mo, H., Wang, J., & Niu, H. (2016). Exponent back propagation neural network forecasting for financial cross-correlation relationship. *Expert Systems with Applications*, *53*, 106–116.

Niu, H., & Wang, J. (2013). Volatility clustering and long memory of financial time series and financial price model. *Digital Signal Processing*, *23*, 489–498.

Oladosu, G. (2009). Identifying the oil price–macroeconomy relationship: An empirical mode decomposition analysis of US data. *Energy Policy*, *37*, 5417–5426.

Stocchi, M., & Marchesi, M. (2018). Fast wavelet transform assisted predictors of streaming time series. *Digital Signal Processing*, *77*, 5–12.

Sun, C., Ding, D., Fang, X., Zhang, H., & Li, J. (2019). How do fossil energy prices affect the stock prices of new energy companies? evidence from divisia energy price index in china's market. *Energy*, *169*, 637–645.

Tan, Q. F., Lei, X. H., Wang, X., Wang, H., Wen, X., & Ji, Y. (2018). An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach. *Journal of Hydrology*, *567*, 767–780.

Ting, F. F., Tan, Y. J., & Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. *Expert Systems with Applications*, *120*, 103–115.

Wang, J., & Wang, J. (2016). Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations. *Energy*, *102*, 365–374.

Wen, X., Wei, Y., & Huang, D. (2012). Measuring contagion between energy market and stock market during financial crisis: a copula approach. *Energy Economics*, *34*(5), 1435–1446.

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

Wu, G. D., & Lo, S. L. (2010). Effects of data normalization and inherent-factor on decision of optimal coagulant dosage in water treatment by artificial neural network. *Expert Systems with Applications*, *37*, 4974–4983.

Wu, Y. X., Wu, Q. B., & Zhu, J. Q. (2018). Improved EEMD-based crude oil price forecasting using LSTM networks. *Physica A*, *516*, 114–124.

Xu, Q., Zhuo, X., Jiang, C., & Liu, Y. (2019). An artificial neural network for mixed frequency data. *Expert Systems with Applications*, *118*, 127–139.

Yao, X. (1999). Evolving artificial neural networks. *Proceedings of the IEEE*, *87*, 1423–1447.

Yu, Y., & Wang, J. (2012). Lattice-oriented percolation system applied to volatility behavior of stock market. *Journal of Applied Statistics*, *39*, 785–797.

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, Q., & Benveniste, A. (1992). Wavelet networks. *IEEE Transactions on Neural Networks*, *3*, 889–898.

Zhang, Z., & Kon, M. A. (2017). Wavelet sampling and generalization in neural networks. *Neurocomputing*, *267*, 36–54.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*, 35–62.

Zhang, J., Walter, G. G., Miao, Y., & Lee, W. N. W. (1995). Wavelet neural networks for function learning. *IEEE Transactions on Signal Processing*, *43*, 1485–1497.

Zhang, J., Yan, R., Gao, R. X., & Feng, Z. (2010). Performance enhancement of ensemble empirical mode decomposition. *Mechanical Systems and Signal Processing*, *24*, 2104–2123.

Zhang, K., & Yu, Y. (2010). Application of wavelet neural network in prediction of gold price. *Computer Engineering and Applications*, 27.