

Research Article

Sunspots Time-Series Prediction Based on Complementary Ensemble Empirical Mode Decomposition and Wavelet Neural Network

Guohui Li and Siliang Wang

School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi 710121, China

Correspondence should be addressed to Guohui Li; lgxcd@163.com

Received 25 January 2017; Accepted 2 March 2017; Published 16 March 2017

Academic Editor: Tomasz Kapitaniak

Copyright © 2017 Guohui Li and Siliang Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The sunspot numbers are the major target which describes the solar activity level. Long-term prediction of sunspot activity is of great importance for aerospace, communication, disaster prevention, and so on. To improve the prediction accuracy of sunspot time series, the prediction model based on complementary ensemble empirical mode decomposition (CEEMD) and wavelet neural network (WNN) is proposed. First, the sunspot time series are decomposed by CEEMD to obtain a set of intrinsic modal functions (IMFs). Then, the IMFs and residuals are reconstructed to obtain the training samples and the prediction samples, and these samples are trained and predicted by WNN. Finally, the reconstructed IMFs and residuals are the final prediction results. Five kinds of prediction models are compared, which are BP neural network prediction model, WNN prediction model, empirical mode decomposition and WNN hybrid prediction model, ensemble empirical mode decomposition and WNN hybrid prediction model, and the proposed method in this paper. The same sunspot time series are predicted with five kinds of prediction models. The experimental results show that the proposed model has better prediction accuracy and smaller error.

1. Introduction

The sunspot numbers are the major targets which describes the solar activity level. The solar activity influences the human health and living environment on earth. Long-term prediction of sunspot activity can provide important reference information for aerospace, communication, disaster prevention, and so on [1]. Therefore, the research of sunspot number time-series prediction method has been an important and hot topic in this field [2]. The data of annual and monthly sunspot activity is considered as typical nonlinear, non-Gaussian, and nonstationary time series, which has obvious characteristics of chaotic sequences [3, 4]. The accurate prediction of sunspot numbers has great research significance [5, 6]. It is difficult to predict using the traditional method and the accuracy is not high.

Currently, BP neural network, wavelet neural network, and empirical mode decomposition (EMD) [7] are widely used in nonlinear, nonstationary time-series prediction [8]. BP neural network has some advantages such as simple

structure, strong learning ability, and nonlinear mapping ability. But the predict model of BP neural network is easy to fall into the local minimum problem, and the convergence rate is slow [9]. Wavelet neural network is combined with the characteristics of artificial neural network and wavelet analysis. It has the advantages of self-learning ability and localization of wavelet transform, so it can effectively solve the local minimum problem [10]. In [11], tight wavelet neural network is used to predict sunspot number. Wavelet neural network is used to predict the charge load in [12]. It is shown that wavelet neural network obviously improves the prediction accuracy compared with BP neural network model. In [13], EMD is used to extract a signal which is more stable, and then the wavelet analysis method is used to get key features of sunspot data time series. However, the mode mixing effect will happen in the process of decomposing signal by the EMD, which leads to unsatisfactory decomposition. Wu and Huang [14] proposes ensemble empirical mode decomposition (EEMD) which is a noise-assisted data analysis method, and the mode mixing can be avoided during the empirical

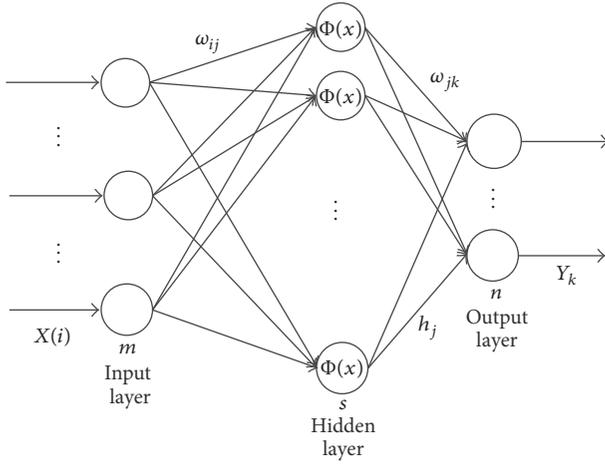


FIGURE 1: The structure of wavelet neural network.

mode decomposition [15]. In [16], the mix model with EEMD and wavelet neural network is proposed. Although the reconstruction error is reduced, the cost of computing is increased [17]. Yeh et al. [18] propose complementary ensemble empirical mode decomposition. The method adds different Gaussian white noise to the remaining components of each intrinsic mode functions to reduce the mode mixing and decrease false component.

Therefore, the prediction model based on complementary ensemble empirical mode decomposition and wavelet neural network is proposed, which is combined with the advantages of CEEMD and wavelet neural network and adopts the idea of a combined model.

2. Basic Principle

2.1. Complementary Ensemble Empirical Mode Decomposition. CEEMD is a noise-assisted analysis method [19]. Two opposite white noise signals are added to the time series signal, and j th mode is obtained by the EMD. $E_j(\cdot)$ represents the j th intrinsic modal function of the EMD. $\omega^i(n)$ represents zero mean Gaussian white noise with unit variance, that is, $N(0, 1)$. ε_k represents the signal-to-noise ratio which is selected for each stage. $x(t)$ is set as target signal, and then the signal added to white noise is $x(t) + \varepsilon_0\omega^i(n)$. The concrete steps of the CEEMD algorithm [20] are as follows:

- (1) Decompose $x(t) + \varepsilon_0\omega^i(n)$ by the EMD for N times, then IMF_1 is obtained:

$$\text{IMF}_1(t) = \frac{1}{N} \sum_{i=1}^N \text{IMF}_1^i(t). \quad (1)$$

- (2) Calculate the first-order residual r_1 :

$$r_1(t) = x(t) - \text{IMF}_1(t). \quad (2)$$

- (3) Decompose $r_1(t) + \varepsilon_1 E_1(\omega^i(t))$, $i = 1, 2, \dots, n$. The first intrinsic modal function is used as the IMF_2 of the CEEMD:

$$\text{IMF}_2(t) = \frac{1}{n} \sum_{i=1}^n E_1(r_1(t) + \varepsilon_1 E_1(\omega^i(t))). \quad (3)$$

- (4) In the same way, the k th residual r_k is obtained:

$$r_k(t) = r_{k-1}(t) - \text{IMF}_k(t). \quad (4)$$

- (5) Decompose $r_k(t) + \varepsilon_k E_k(\omega^k(t))$. The first intrinsic modal function is used as the $\text{IMF}_{(k+1)}$ of the CEEMD:

$$\text{IMF}_{(k+1)}(t) = \frac{1}{n} \sum_{i=1}^n E_1(r_k(t) + \varepsilon_k E_k(\omega^k(t))). \quad (5)$$

- (6) Repeat steps (4) and (5), until the residual can no longer be decomposed, and then the final residuals $R(t)$ is

$$R(t) = x(t) - \sum_{k=1}^K \text{IMF}_k(t). \quad (6)$$

Therefore, the target signal $x(t)$ can be expressed as

$$x(t) = \sum_{k=1}^K \text{IMF}_k(t) + R(t), \quad (7)$$

where K is the sum of $\text{IMF}(t)$. The complete CEEMD decomposition signal may be obtained through the above formula, and the original signal can be precisely reconstructed.

2.2. Wavelet Neural Network. Wavelet neural network is a multilayer feedforward neural network which is based on BP neural network and wavelet theory, and it takes the wavelet basis function as the hidden layer excitation function [10]. So WNN has better generalization ability than neural network [21]. WNN is a good predictive method that can process nonlinear data, and it has characteristic of wavelet analysis and neural network.

The structure of the wavelet neural network is similar to a feedforward neural network [10]. The input node layer has one or more inputs. There is an implicit layer in the middle of the network. The function of basis wavelet is used as the activation function of the hidden layer. The output layer consists of one or more linear combiners. The structure of wavelet neural network is shown in Figure 1.

In Figure 1, m is the number of input layer nodes, x_i is input signal, s is the number of hidden layer nodes, n is the number of output layer nodes, $\phi(t)$ is wavelet basis function, h_j is the output of hidden layer, y_k is output signal, ω_{ij} is the connection weight between input layer and hidden layer, and ω_{jk} is the connection weight between output layer and hidden layer.

The prediction algorithm process of wavelet neural network is as follows:

- (1) Construction of wavelet neural network: initialize network, select the appropriate structure of wavelet

neural network, and randomly initialize the scale factor, the displacement factor, the network connection weight, and the learning rate of wavelet function.

- (2) Classification of samples: divide the original data into training samples and test samples, use the training samples to train network, and use test samples to predict.
- (3) Predicted output: output training samples and corresponding expected outputs, the error ε of network output and expected output can be expressed as

$$\varepsilon = \sum_{k=1}^m g(k) - f(k), \quad (8)$$

where $g(k)$ is expected output and $f(k)$ is predicted output.

- (4) Correction of weight parameters: in order to make the predicted value more close to the expected value, the network parameters can be corrected according to error ε . The adjustment formula is

$$\begin{aligned} \omega_{ij}^{(d+1)} &= \omega_{ij}^{(d)} + \Delta\omega_{ij}^{(d+1)} \\ \omega_{jk}^{(d+1)} &= \omega_{jk}^{(d)} + \Delta\omega_{jk}^{(d+1)} \\ \alpha_j^{(d+1)} &= \alpha_j^{(d)} + \Delta\alpha_j^{(d+1)} \\ b_j^{(d+1)} &= b_j^{(d)} + \Delta b_j^{(d+1)}, \end{aligned} \quad (9)$$

where ω_{ij} and ω_{jk} , respectively, denote the connection weight between input layer and hidden layer and hidden layer and output layer in training stage, $\Delta\omega_{ij}$ and $\Delta\omega_{jk}$, respectively, denote the connection weight correction term between input layer and hidden layer and hidden layer and output layer in training stage. α_j and b_j denote, respectively, the scale factor and the displacement factor of the j th wavelet.

- (5) When the error reaches the preset accuracy, return to step (3).

2.3. Prediction Model Based on Complementary Ensemble Empirical Mode Decomposition and Wavelet Neural Network. For nonlinear and nonstationary time series, the normalization processing is conducted firstly. Secondly, the CEEMD method is used to decompose, and a set of the IMF components and residuals are obtained. Then, all IMF components and residuals are trained and predicted by WNN method. Finally, the final prediction results can be obtained by accumulating prediction results. The block diagram of prediction model based on CEEMD and WNN is as shown in Figure 2.

Prediction model process based on CEEMD and WNN is as follows:

- (1) Load the original data, use CEEMD to decompose original data, and obtain the IMF components and residuals.

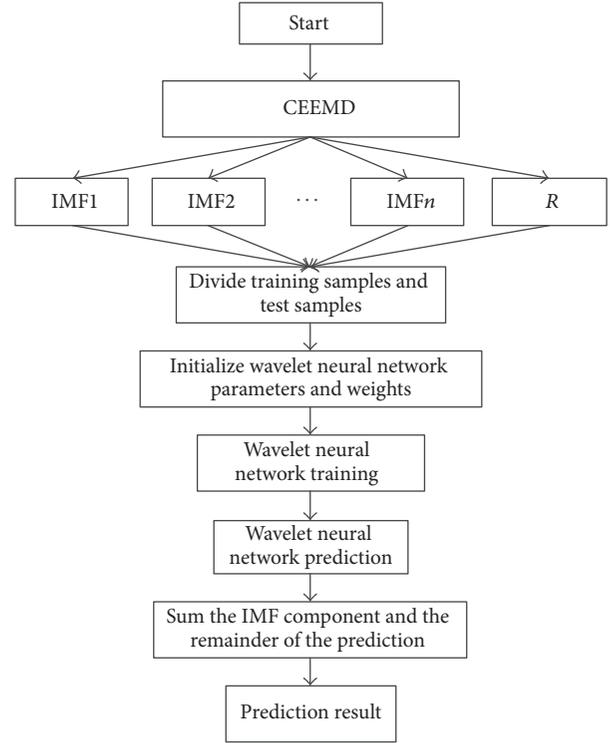


FIGURE 2: The block diagram of prediction model based on CEEMD and WNN.

- (2) Design the structure of wavelet neural network for every IMF component. Initialize the number of input layers, hidden layers, and output layers. Initialize the weight and learning rate, divide the data of each IMF component into training samples and predicting samples, and normalize input/output samples.
- (3) Train wavelet neural network.
- (4) Predict each IMF component and residual by WNN.
- (5) Accumulate the prediction of each IMF component. That is the predicted value of the original time series.

3. Data Simulation and Analysis

In this paper, the sunspot number smoothing monthly mean time series is derived from the sunspot index data center of the Royal Observatory Belgium (<http://sidc.oma.be/silso/datafiles>). The data used for calculation is the sunspot number data from January 1900 to December 2014, and there are 1380 datums. The number of original sunspots is shown in Figure 3.

The CEEMD method is used to decompose the original data, where the mean of a white noise sequence is 0, the standard deviation coefficient of amplitude is 0.01, and the general average times N is 100. The results of CEEMD decomposition are shown in Figure 4.

Eight IMF components and one remainder are obtained by CEEMD. It can be seen that the IMF component is extracted from high frequency to low frequency. IMF1–IMF3

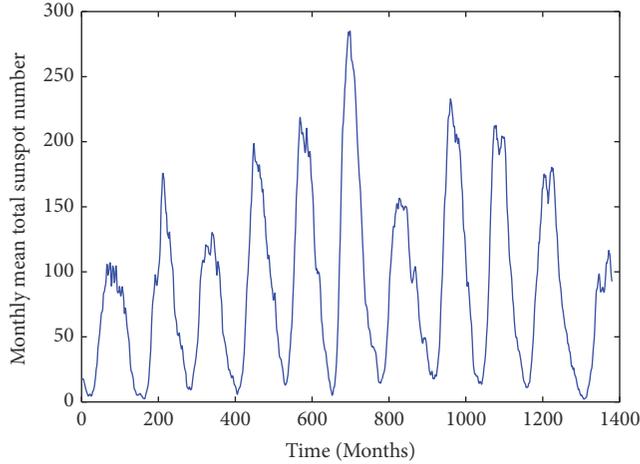


FIGURE 3: Monthly mean total sunspot number in 1900–2014.

frequency is large which is no periodicity; this part is the random component in sunspot sequence, which includes the useful components of the signal and also includes a lot of noise. IMF4–IMF6 show cyclicity. IMF7–IMF8 are a low-frequency component and have less influence on the sunspot number of fluctuations in the curve. Remainder R reflects the changing trend.

The IMF1–IMF8 and residuals are, respectively, predicted by WNN. The first four sunspot values of each component are used to predict the value of the fifth sunspot. 1380 datums can be divided into 1372 sets of data, where the former 996 sets of data are used as the test data and the latter 376 sets of data are used as the forecast data. Then the parameters of wavelet neural network are set up. The best hidden layer nodes and the number of layers are determined by trial-and-error method. The number of nodes in the input layer is 4, the number of nodes in the hidden layer is 9, and the number of nodes in the output layer is 1. That is, the structure of the wavelet neural network is 4-9-1. The learning probabilities are 0.01 and 0.001, respectively, and the number of iterations is 100. The wavelet basis function of the hidden layer is usually the Morlet mother wavelet function. Morlet wavelet function has advantages in time-frequency local characteristics and symmetry, so that the model has good linear approximation ability, and the formula of mother wavelet function is

$$f(t) = \exp\left\{-\frac{t^2}{2} \times \cos(1.75 \times t)\right\}, \quad (10)$$

where t is input variable and $f(t)$ is output variable.

The partial derivative of mother wavelet function is

$$\frac{dy}{dt} = -1.75 \cdot \sin(1.75 \cdot t) \cdot \exp\left\{-\frac{t^2}{2}\right\} - t \cdot \cos(1.75 \cdot t) \cdot \exp\left\{-\frac{t^2}{2}\right\}. \quad (11)$$

The results of the prediction are accumulated, which is the result of the final prediction. The prediction results with

TABLE 1: Error comparison of RMSE and MAE in five models.

Models	RMSE	MAE
BP	50.14155	4.75006
WNN	28.36957	3.44447
EMD-WNN	24.705	3.0122
EEMD-WNN	15.2656	1.7076
CEEMD-WNN	12.64374	1.58413

CEEMD and WNN are shown in Figure 5. For the sake of comparison, the BP neural network prediction model, WNN prediction model, EMD and WNN hybrid prediction model, EEMD and WNN hybrid prediction model, and the proposed method are used to predict the same sunspot time series.

The red line in Figure 5 represents the predicted number of sunspots and the blue line represents the practical number of sunspots. It can be seen that the CEEMD-WNN model proposed in this paper has good fitting to the original data and can predict the number of sunspots well. Figure 6 shows the comparison results of BP neural network prediction model, WNN prediction model, EMD-WNN prediction model, EEMD-WNN prediction model, and CEEMD-WNN prediction model, and the partial enlargement is shown in Figure 7.

In order to verify the prediction result, the RMS error (RMSE) and mean absolute error (MAE) are used to estimate the result of prediction model. RMS error can be used to measure the deviation between the observed value and the true value, which reflects the discrete degree of the data. The mean absolute error is a good reflection of the actual situation of the prediction error.

The RMS error (RMSE) is

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{\text{pred}} - Y_{\text{real}})^2}. \quad (12)$$

The mean absolute error (MAE) is

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |Y_{\text{pred}} - Y_{\text{real}}|, \quad (13)$$

where Y_{pred} is forecast data and Y_{real} is original data. In order to avoid the difference of each prediction result and the error of randomness, this paper takes the average value of 10 times prediction as the error value of final prediction. Compared with the five models, the RMS error and the mean absolute error of each model are shown in Table 1.

As shown in Table 1, the RMSE and MAE of the BP neural network prediction model are 50.14155 and 4.75006, respectively, and the RMSE and MAE of the WNN prediction model are 28.36957 and 3.44447, respectively, which shows that the WNN has better accuracy than the BP neural network. The RMSE and MAE of EMD-WNN model and EEMD-WNN model are 24.705, 3.0122 and 15.2656, 1.7076 respectively. The RMSE and MAE of EMD-WNN model are smaller than those of WNN. The RMSE and MAE of

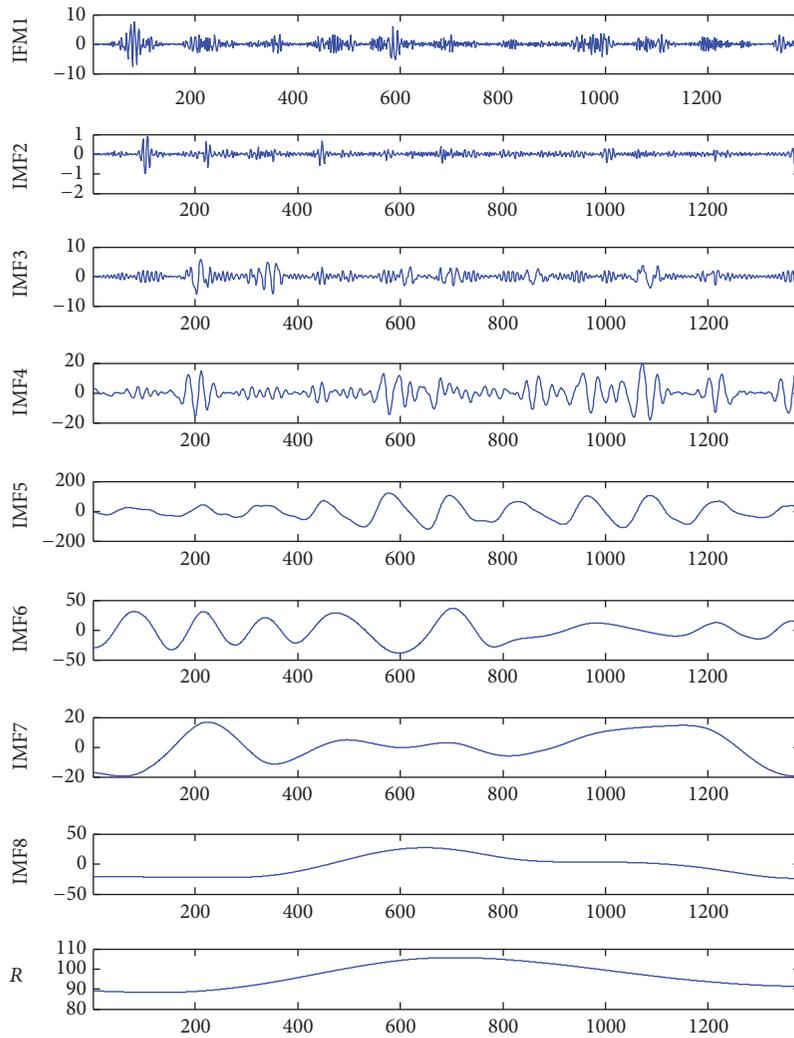
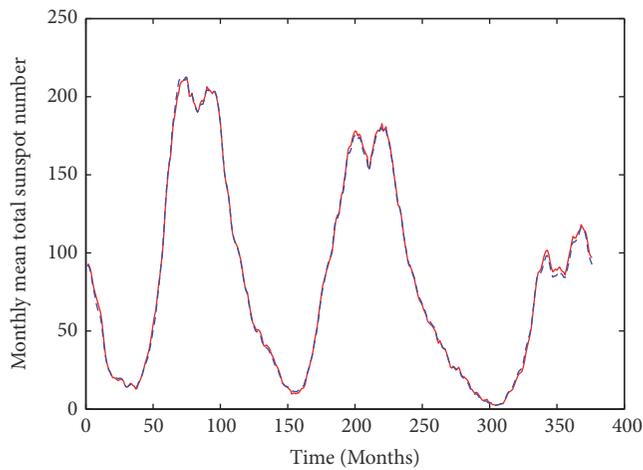
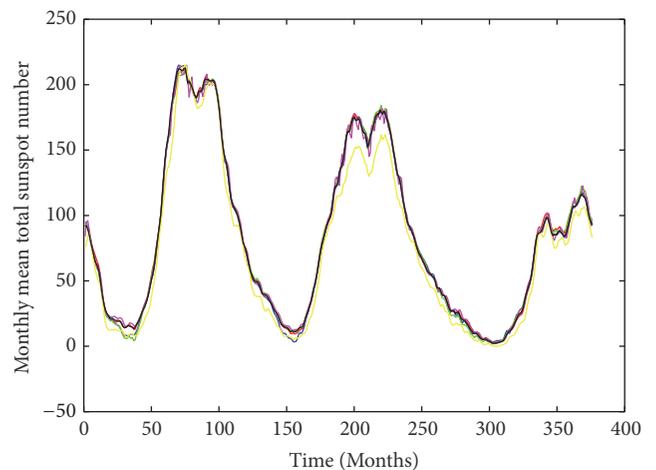


FIGURE 4: The results of CEEMD decomposition.



— Predicted monthly mean total sunspot number
 - - - Practical monthly mean total sunspot number

FIGURE 5: CEEMD-WNN prediction results.



— CEEMD-WNN — WNN
 — EEMD-WNN — BP
 — EMD-WNN — SUNSPOT

FIGURE 6: Predicted results of sunspot numbers for each model.

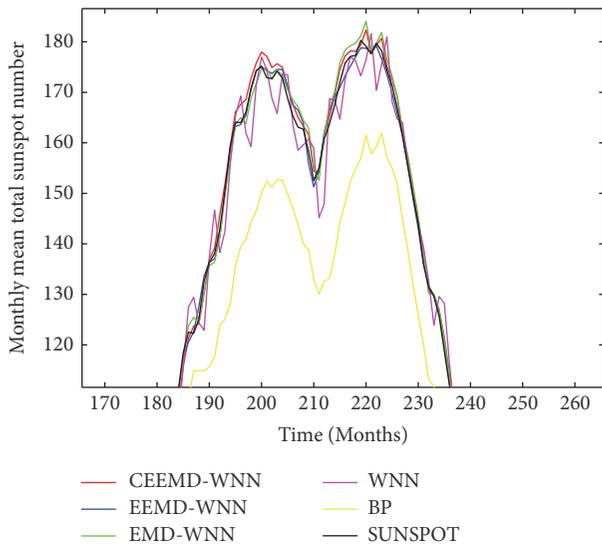


FIGURE 7: Local predicted result graphs of sunspot numbers for each model.

CEEMD-WNN prediction model proposed in this paper are 12.64374 and 1.58413, respectively, which further improves the prediction precision and reduces the prediction error. Therefore, the model of complementary ensemble empirical mode decomposition combined with wavelet neural network proposed in this paper can predict sunspot number and the trend of sunspot time series better, and it is a good forecasting model.

4. Conclusions

In this paper, a hybrid prediction model of complementary ensemble empirical mode decomposition and wavelet neural network is proposed to predict the sunspot monthly mean time series. Compared with BP neural network, WNN, EMD-WNN prediction model, and EEMD-WNN prediction model, the complementary ensemble empirical mode decomposition can eliminate Gaussian white noise in the reconstructed signal and reduce the mode aliasing effect effectively. The neural network converges quickly and avoids the local minimum in the iterative process. The experimental results show that the proposed method can improve the prediction precision and reduce the error compared with other models in predicting the same sunspot time series. The model can be applied to other fields after conducting some modification and the model is worth further research and practical application. The two aspects, (1) the Gaussian white noise shall be added in the process of CEEMD decomposition and (2) the weight shall be added in process of the creation of wavelet neural network, will lead to different results of each experiment. Therefore, the next work can focus on improving the decomposition method and optimizing the weight of wavelet neural network.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] J. Tang and X. Zhang, "Prediction of smoothed monthly mean sunspot number based on chaos theory," *Wuli Xuebao/Acta Physica Sinica*, vol. 61, no. 16, Article ID 169601, 2012.
- [2] Z. D. Tian, S. J. Li, Y. H. Wang, X. D. Wang, and Y. Sha, "A hybrid prediction model of smoothed monthly mean sunspot number," *Scientia Sinica Physica, Mechanica & Astronomica*, vol. 46, Article ID 119601, 2016.
- [3] C. Jiang and F. Song, "Sunspot forecasting by using chaotic time-series analysis and NARX network," *Journal of Computers*, vol. 6, no. 7, pp. 1424–1429, 2011.
- [4] R. Arlt and N. Weiss, "Solar activity in the past and the chaotic behaviour of the dynamo," *Space Science Reviews*, vol. 186, no. 1–4, pp. 525–533, 2014.
- [5] R. P. Kane, "An estimate for the size of sunspot cycle 24," *Solar Physics*, vol. 282, no. 1, pp. 87–90, 2013.
- [6] A. Gkana and L. Zachilas, "Sunspot numbers: data analysis, predictions and economic impacts," *Journal of Engineering Science and Technology Review*, vol. 8, no. 1, pp. 79–85, 2015.
- [7] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [8] Y. X. Zhang, Q. T. Xiao, J. X. Xu, and X. L. Sang, "Wavelet neural network prediction model based on empirical mode decomposition," *Computer Applications and Software*, vol. 33, pp. 284–287, 2016.
- [9] S. Li, L. J. Liu, and Y. P. Liu, "Prediction for chaotic time series of optimized BP neural network based on modified PSO," *Computer Engineering and Applications*, vol. 49, pp. 245–248, 2013.
- [10] C. Guan, P. B. Luh, L. D. Michel, Y. Wang, and P. B. Friedland, "Very short-term load forecasting: wavelet neural networks with data pre-filtering," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 30–41, 2013.
- [11] H. Luo, H. J. Wang, and B. Long, "Research about time serial prediction based on tight wavelet neural net," *Application Research of Computer*, vol. 25, pp. 2366–2368, 2008.
- [12] S. Gupta, V. Singh, A. P. Mittal, and A. Rani, "Weekly load prediction using wavelet neural network approach," in *Proceedings of the 2nd International Conference on Computational Intelligence & Communication Technology (CICT '16)*, pp. 174–179, IEEE, Ghaziabad, India, February 2016.
- [13] A. Hu, J. Sun, and L. Xiang, "Mode mixing in empirical mode decomposition," *Zhendong Ceshi Yu Zhenduan/Journal of Vibration, Measurement and Diagnosis*, vol. 31, no. 4, pp. 429–434, 2011.
- [14] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.
- [15] J.-L. Qu, X.-F. Wang, F. Gao, Y.-P. Zhou, and X.-Y. Zhang, "Noise assisted signal decomposition method based on complex empirical mode decomposition," *Wuli Xuebao/Acta Physica Sinica*, vol. 63, no. 11, Article ID 110201, 2014.

- [16] Y. Lei, Z. He, and Y. Zi, "EEMD method and WNN for fault diagnosis of locomotive roller bearings," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7334–7341, 2011.
- [17] Y. Zhao, Y. X. Yue, J. L. Huang, J. Wang, C. X. Liu, and B. Q. Liu, "CEEMD and wavelet transform jointed de-noising method," *Progress in Geophysics*, vol. 30, pp. 2870–2877, 2015.
- [18] J.-R. Yeh, J.-S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method," *Advances in Adaptive Data Analysis*, vol. 2, no. 2, 2010.
- [19] J.-D. Zheng, J.-S. Cheng, and Y. Yang, "Modified EEMD algorithm and its applications," *Journal of Vibration and Shock*, vol. 32, no. 21, pp. 21–46, 2013.
- [20] E. Wang, J. Zhang, X. Ma, and H. Ma, "A new threshold denoising algorithm for partial discharge based on CEEMD-EEMD," *Power System Protection and Control*, vol. 44, no. 15, pp. 93–98, 2016.
- [21] V. Sharma, D. Yang, W. Walsh, and T. Reindl, "Short term solar irradiance forecasting using a mixed wavelet neural network," *Renewable Energy*, vol. 90, pp. 481–492, 2016.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

