

## Research Article

# A Deep Learning Prediction Model Based on Extreme-Point Symmetric Mode Decomposition and Cluster Analysis

**Guohui Li, Songling Zhang, and Hong Yang**

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

Correspondence should be addressed to Hong Yang; [uestcyhong@163.com](mailto:uestcyhong@163.com)

Received 14 July 2017; Accepted 5 December 2017; Published 27 December 2017

Academic Editor: Simone Bianco

Copyright © 2017 Guohui Li et al. 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.

Aiming at the irregularity of nonlinear signal and its predicting difficulty, a deep learning prediction model based on extreme-point symmetric mode decomposition (ESMD) and clustering analysis is proposed. Firstly, the original data is decomposed by ESMD to obtain the finite number of intrinsic mode functions (IMFs) and residuals. Secondly, the fuzzy  $c$ -means is used to cluster the decomposed components, and then the deep belief network (DBN) is used to predict it. Finally, the reconstructed IMFs and residuals are the final prediction results. Six kinds of prediction models are compared, which are DBN prediction model, EMD-DBN prediction model, EEMD-DBN prediction model, CEEMD-DBN prediction model, ESMD-DBN prediction model, and the proposed model in this paper. The same sunspots time series are predicted with six kinds of prediction models. The experimental results show that the proposed model has better prediction accuracy and smaller error.

## 1. Introduction

At present, there are still many difficulties in predicting nonlinear signal such as sunspots and underwater acoustic signal. Sunspots are the basic parameters of the solar activity level. They are closely related to the geomagnetic disturbance and ionospheric electron concentration. Prediction of sunspots is an important part of spatial forecast which can provide important reference information for communication, navigation, and positioning. Some scholars have conducted extensive research on the theory of forecasting [1, 2]. In the time-frequency signal analysis, the commonly used method is Fourier transform which is mainly mapping the time domain signal to the frequency domain energy spectrum space, but Fourier transform only applies to the stationary signal. Artificial neural network has the characteristics of independent learning compared with the previous regression analysis which is especially suitable for nonlinear signal processing. However, due to the limitation of synchronous instantaneous input, the time cumulative effect of continuous signal cannot be reflected, and the prediction accuracy is low [3]. Wavelet neural network is combined with the characteristics of artificial neural network and wavelet analysis which has been widely applied to the processing of nonlinear

signal. Li and Wang [1] propose the prediction model based on complementary ensemble empirical mode decomposition and wavelet neural network. Although its prediction accuracy is improved to a certain extent, there is room for further improvement. The emergence of empirical mode decomposition [4] (EMD) provides an idea for the processing of nonlinear signal. It does not need to select a basis function, but it is difficult to determine the number of screenings and there are many defects in Hilbert spectral analysis. The extreme-point symmetric mode decomposition [5–7] (ESMD) method is a further improvement of the EMD, whose envelope interpolation from extreme points of the original external changes to internal upper and lower extreme symmetric interpolation. The residual modal component is optimized by the least squares method, which has the characteristics of adaptive global to determine the number of screenings. ESMD uses the direct interpolation (DI) method, which is different from Fourier transform only by the idea of transformation of the integral algorithm. In view of the advantages of ESMD, this paper selects the ESMD method to decompose the nonlinear time series. Then, fuzzy  $c$ -means [8, 9] clustering analysis is used to aggregate the data of the same membership to facilitate the prediction analysis of the model. Finally, the

deep belief network [10–13] (DBN) is trained to achieve the expected output value, and then the predicted output value is reconstructed to obtain the final predicted value.

## 2. ESMD Method

ESMD is a new development of the Hilbert-Huang transform, and its algorithm is as follows:

- (1) Find all the extreme points (maximum and minimum) of the data  $Z$  and record them as  $X = \{x_1, x_2, \dots, x_n\} \in R^{pn}$ .
- (2) Connect the adjacent poles with lines, and the midpoint is recorded as  $B_i$  ( $i = 1, 2, \dots, n - 1$ ).
- (3) Supplement the left and right border midpoint  $B_0, B_n$  by certain methods.
- (4) Use the obtained  $n + 1$  midpoints to construct  $p$  bar differential lines and calculate their mean curves  $X = \{x_1, x_2, \dots, x_n\} \in R^{pn}$ .
- (5) Repeat the above steps until the number of screenings reaches the preset maximum value; then the first decomposed empirical mode is recorded as  $M_1$ .
- (6) Repeat the above steps for  $Z - M_1$  to obtain  $M_2, M_3, \dots$  until the final margin  $R$  only has a certain number of poles.
- (7) Let the maximum number of screenings  $D$  process and cycle of the above process in the integer interval  $X = \{x_1, x_2, \dots, x_n\} \in R^{pn}$  to get some components, and then calculate the variance ratio  $\sigma/\sigma_0$  and draw it with the  $D$  change map, where  $\sigma$  is the relative standard deviation of  $Z - R$  and  $\sigma_0$  is the standard deviation of the original data.
- (8) Select the maximum number of screenings  $D_0$  which corresponded to the minimum variance ratio  $\sigma/\sigma_0$  in the interval  $[D_{\min}, D_{\max}]$ , and repeat the first six steps to output the decomposition results.

## 3. Clustering Algorithm

The fuzzy clustering algorithm was originally proposed by Dunn [14] and further introduced by Bezdek [15], which is now being applied to many fields. Its operation steps can be expressed as follows: the sample set  $X = \{x_1, x_2, \dots, x_n\} \in R^{pn}$  is divided into  $c$  class. Membership degree of any element  $x_k$  in the sample on the  $i$  class is recorded as  $u_{ik}$ . The fuzzy membership matrix is used in the matrix after clustering, which is recorded as  $U = \{u_{ik}\} \in R^{cn}$  and satisfies the following conditions:

$$\begin{aligned} u_{ik} &\in [0, 1], \quad \forall i, k \\ 0 &< \sum_k u_{ik} < n, \quad \forall i \\ \sum_i u_{ik} &= 1, \quad \forall k. \end{aligned} \quad (1)$$

The fuzzy  $c$ -means clustering is obtained by minimizing the purpose function  $J_m(U, V)$ . The purpose function is as follows:

$$J_m(U, V) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m d_{ik}^2(x_k, v_i), \quad (2)$$

where  $U = \{u_{ik}\}$  is the membership matrix,  $V = \{v_1, v_2, \dots, v_c\} \in R^{pc}$  represents  $c$  clustering center point sets, and  $m \in [1, \infty)$  is the weighted index. The fuzzy clustering is transformed into hard mean clustering [14] when  $m$  is 1. The ideal range of  $m$  is [1.5, 2.5], usually  $m = 2$ .

The distance from the  $k$ th sample to the  $i$ th class center is

$$d_{ik}^2(x_k, v_i) = \|x_k - v_i\|_A^2 = (x_k - v_i)^T A (x_k - v_i), \quad (3)$$

where  $A$  is the positive definite matrix of  $p \times p$ , special conditions  $A = I$ , and (3) is the Euclidean distance. FCM [16] is achieved by continuously optimizing the objective function. FCM algorithm process is as follows:

- (1) Initialize the cluster center  $V = \{v_1, v_2, \dots, v_c\}$ .
- (2) Calculate membership matrix:

$$u_{ik} = \left[ \sum_{j=1}^c \left[ \frac{d_{ik}(x_k, v_i)}{d_{jk}(x_k, v_j)} \right]^{2/(m-1)} \right]^{-1} \quad k = 1, 2, \dots, n. \quad (4)$$

- (3) Calculate the new cluster center:

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m} \quad i = 1, 2, \dots, c. \quad (5)$$

- (4) Repeat steps (2) and (3) until (2) reaches convergence. When  $d_{ik}(x_k, v_i) = 0$ , a singular value is generated, and membership cannot be calculated by (4). A class of nonsingular values will appear when the membership value is 0. The class of singular value appears, and then the membership is calculated according to (1).

## 4. Forecasting Model

**4.1. DBN Network Structure.** DBN [17, 18] is organized by a number of restricted Boltzmann machine (RBM) models. The visual layer of the RBM model is similar to the input layer, and the hidden layer is similar to the output layer. Learning between layers and layers of a large numbers of RBM models is used to end the final operation. The specific structure of RBM model is shown in Figure 1. The unit of the visual layer and the unit of the hidden layer can be interconnected with each other. The elements inside the layers are not connected. The units of the hidden layer can obtain a close correlation between the units of the visual layer.

The core of DBN is restricted Boltzmann machine unit, RBM is a typical artificial neural network, and it is a special logarithmic linear Markova random field [19]. The RBM model has three parameters: the offset vector  $V = (V_1, V_2, \dots, V_m)$  represents the offset of each node of the visual

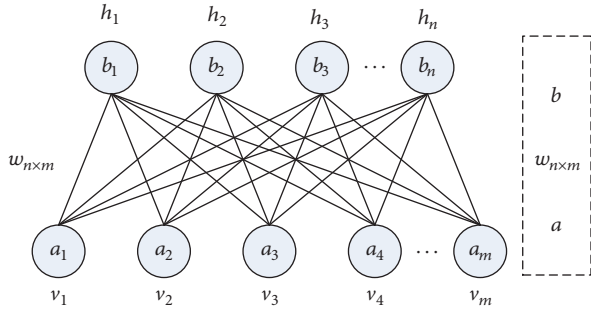


FIGURE 1: Network structure diagram of RBM.

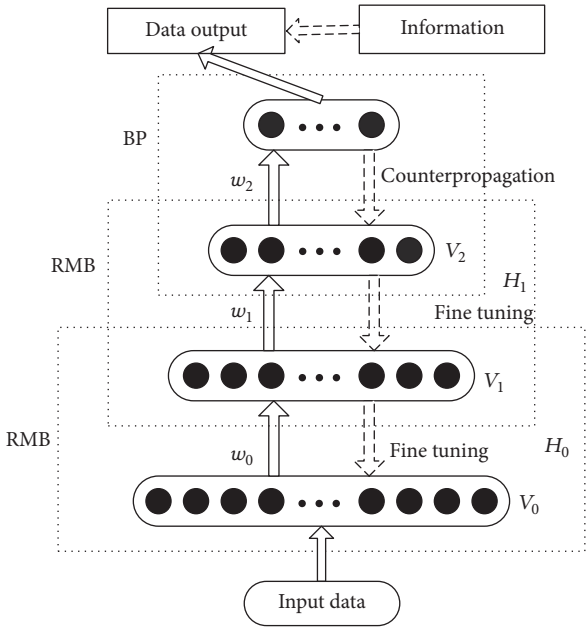


FIGURE 2: Deep belief network architecture.

layer,  $H = (H_1, H_2, \dots, H_n)$  represents the offset of each node of the hidden layer, and  $w$  represents the weight matrix between the nodes of two layer. These three parameters directly determine the model to encode the  $m$  dimension data into  $n$  dimension data; thus the conversion between features is realized.

DBN is composed of a large number of RBM models from the bottom to the top and the top of a layer of BP neural network, which is shown in Figure 2. The bottom is the sample input which is waiting for training.  $V_0$  and  $H_0$  are the nodes of RBM visual layer and hidden layer in the first layer, respectively.  $w_0$  represents the weight between the visual and hidden layers [20].

#### 4.2. DBN Training Process

- (1) The input sample is entered from the bottom level.
- (2) The first RBM model was trained and then passed to the second RBM model for training, followed by

continuous training until the training of the top of the RBM model is also complete.

- (3) After the training is completed, the training data can supervise the operation and adopt the maximum likelihood estimation method to fine-tune the network model.
- (4) Finally, the BP model is used to fine-tune the model parameters of the top layer so as to minimize the value of the loss function.

#### 4.3. The Training Method of Deep Learning

(1) *Unsupervised Learning from Bottom Up (Pretraining)*. Using unlabeled data to train each parameter hierarchically, this is an unsupervised training method, which is the biggest difference from the traditional neural network, and also can be regarded as the process of feature learning. The first layer is first trained with unlabeled data, and the first layer parameters are obtained. The output of the first layer is used as the input of the second layer, so as to train the second layers and finally obtain the parameters of each layer.

(2) *Top-Down Supervised Study (Tuning)*. After the first step is completed, the network adopted discriminative training using labeled data, and the error is transmitted from top to bottom. The first step is similar to the random initialization of the traditional neural network. The difference is that the first step of deep learning is obtained through the study of unlabeled data, rather than random initialization. So the initial value is closer to the overall optimal, so the effect of deep learning is mainly the pretraining of the first step.

4.4. *ESMD and DBN Prediction Model Based on Clustering*. ESMD and DBN prediction model based on clustering is proposed, whose structure is described as follows:

- (1) The original sequence is decomposed by ESMD; then the finite number of IMFs and residuals is obtained.
- (2) The fuzzy  $c$ -means clustering analysis is performed for each IMF component and residual; then the frequency fluctuation rule is got.
- (3) The DBN model is established for each IMF component and residual, respectively; then the predicted value of each component is obtained.
- (4) Reconstruct IMF predicting value to obtain the final predicting results.

### 5. Data Simulation and Analysis

The monthly mean total sunspot number from 1963 to 2012 was used as the original data. There are a total of 600 data points shown in Figure 3. The original data is decomposed by EMD and ESMD, respectively, shown in Figures 4 and 5.

In Figure 5, modal components IMF1~IMF6 are shown from top to bottom, and the instantaneous frequency of modal components IMF3~IMF6 is basically stable. The modal components can achieve relatively high prediction

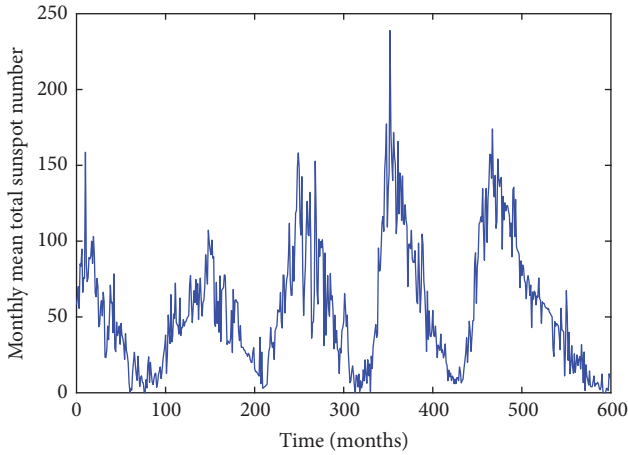


FIGURE 3: Monthly mean total sunspot number in 1963–2012.

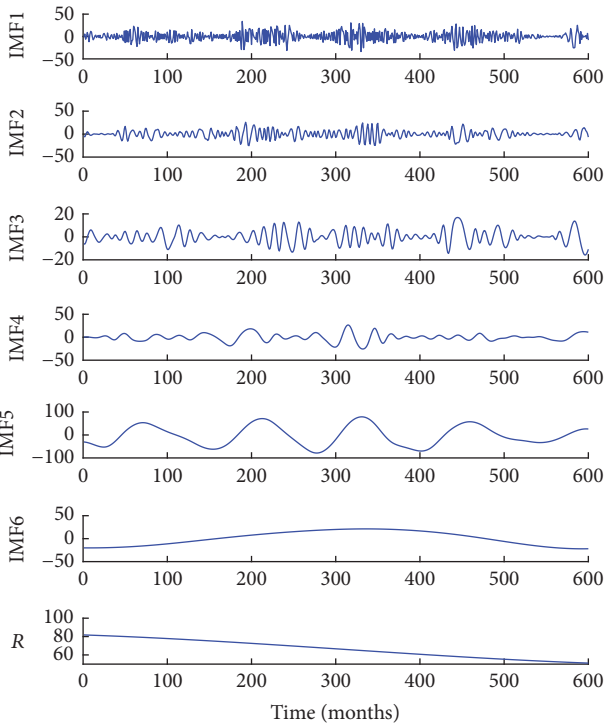


FIGURE 4: The results of EMD decomposition.

results after they are decomposed and predicted. But IMF1~IMF2 are still quite complex compared to other components, the instantaneous frequency is very large, and nonstationary is strong. So the fuzzy *c*-means clustering analysis is performed, and the results are shown in Figure 6.

The DBN structure contains two hidden layers. The number of neurons is 2 and 12, and the learning rate is 1. The DBN network includes two hidden layers. The number of neurons is 20 and 10, the learning rate is 0.1, the cycle number is set to 100, and the momentum is set to 0. After the training is completed; each layer of the RBM model can obtain initialization parameters which constitute the simplest

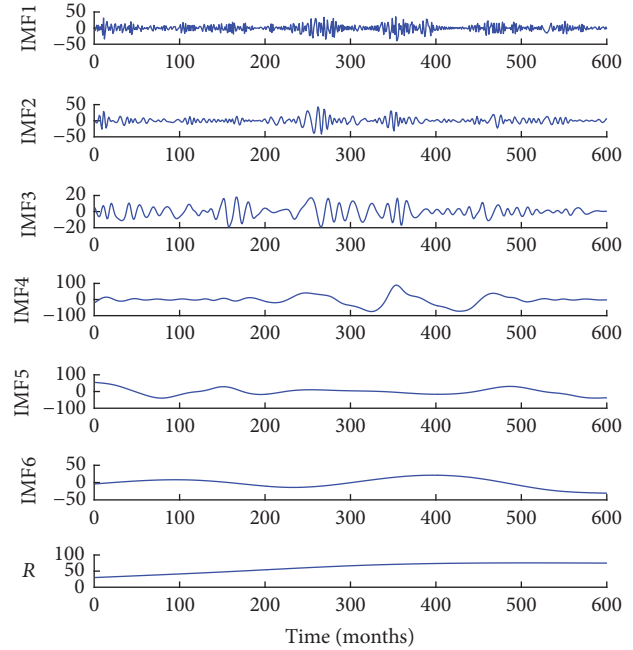


FIGURE 5: The results of ESMD decomposition.

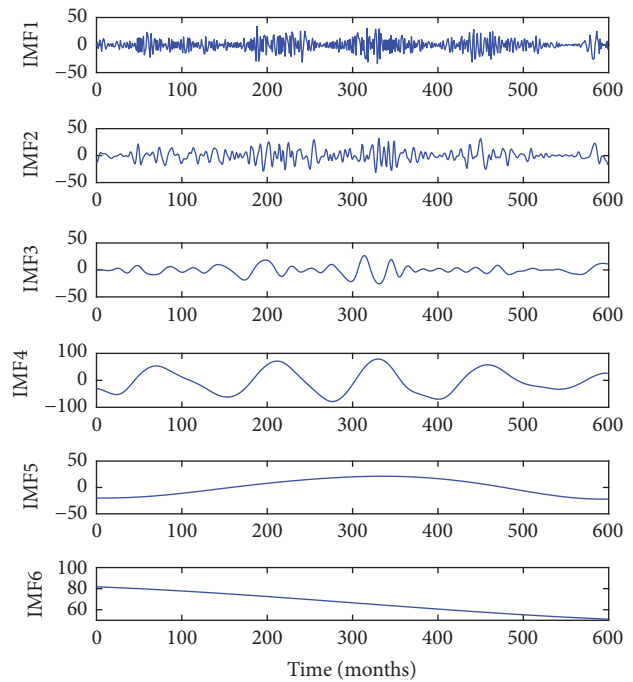


FIGURE 6: Modal components after fuzzy clustering.

model of DBN. The prediction experimental steps are shown in Figure 7.

The predicted results of ESMD-DBN prediction model based on clustering are shown in Figure 8. The same sunspot time series is predicted by DBN directly named DBN prediction. The same sunspot time series is decomposed, respectively, by EMD, EEMD, CEEMD, and ESMD, and the finite number of IMF components and residue are obtained.

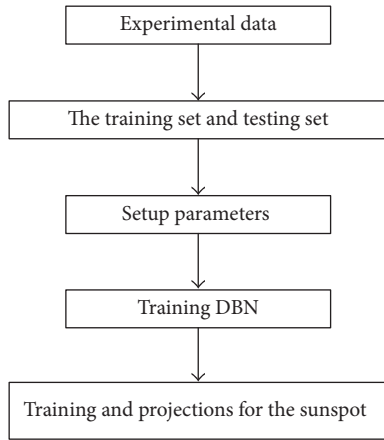


FIGURE 7: Steps of sunspot prediction based on DBN.

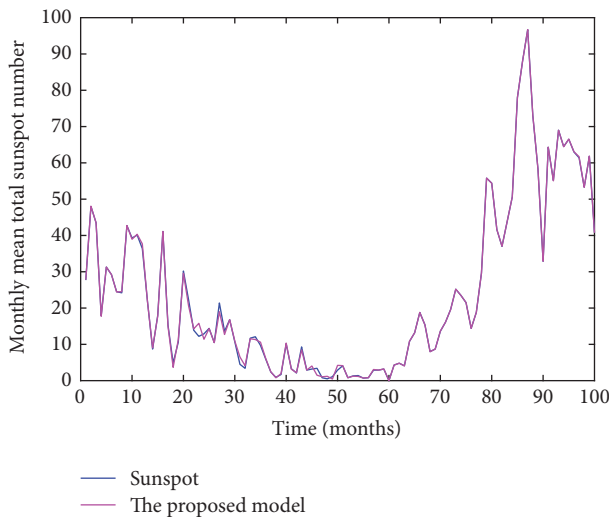


FIGURE 8: The predicted results based on clustering ESMD-DBN.

The DBN model is established for each IMF component and residue, respectively; then the predicted value of each component is obtained. They were named, respectively, by the prediction models of EMD-DBN, EEMD-DBN, CEEMD-DBN, and ESMD-DBN.

The purple line in Figure 8 represents the predicted number of sunspots and the blue line represents the practical number of sunspots. It can be seen that the clustering ESMD-DBN model proposed in this paper has good fitting to the original data and can predict the number of sunspots well. Figure 9 shows the comparison results of the prediction models of DBN, EMD-DBN, EEMD-DBN, CEEMD-DBN, and ESMD-DBN. In order to identify predicted results, local predicted results are shown in Figure 10.

In order to verify the prediction result, the root mean square error (RMSE) and the mean absolute percentage error

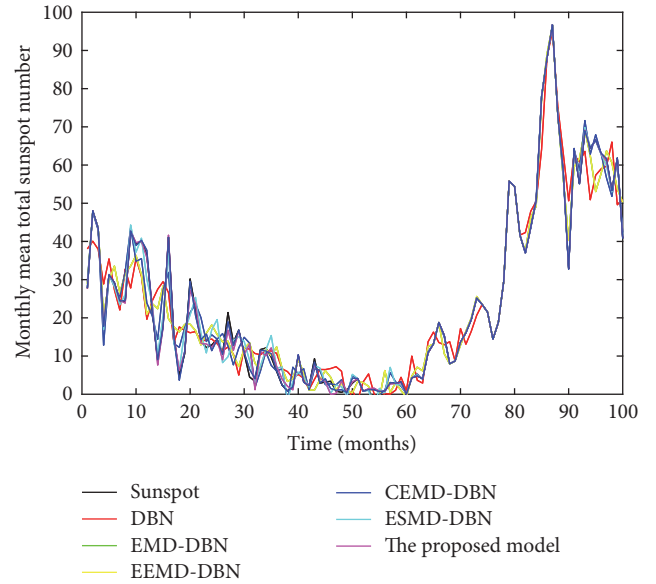


FIGURE 9: Predicted results of sunspot numbers for each model.

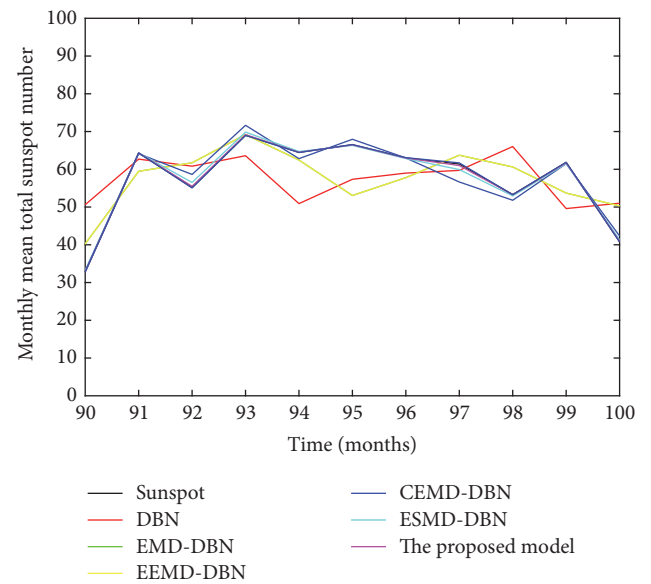


FIGURE 10: Local predicted results of sunspot numbers for each model.

(MAPE) are used to measure the prediction performance. The formulas are as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [\hat{x}(i) - x(i)]^2} \tag{6}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x(i) - \hat{x}(i)}{x(i)} \right| \times 100\%$$

where  $n$  is the number of sample datasets,  $\hat{x}(i)$  is the  $i$ th value of the predicted data, and  $x(i)$  is the  $i$ th value of the actual data. Performance comparison of the six models is shown in Table 1.

TABLE 1: Performance comparison of the six models.

Models	RMSE	MAPE
DBN	8.4398	26.52%
EMD-DBN	5.8374	12.36%
EEMD-DBN	3.0837	7.58%
CEEMD-DBN	1.3405	3.67%
ESMD-DBN	1.0543	2.45%
The proposed model	0.8760	1.36%

As shown in Table 1, the RMSE and MAPE of the proposed model are smaller than the other five models. Therefore, the proposed model can predict sunspot number and the trend of sunspot time series better, and it is an effective prediction model.

## 6. Conclusions

In this paper, a deep learning prediction model based on extreme-point symmetric mode decomposition and clustering analysis is proposed to predict the sunspot monthly mean time series. Comparing with the other models such as DBN, EMD-DBN, EEMD-DBN, CEEMD-DBN, and ESMD-DBN, the RMSE and MAPE of the proposed model are the smallest. The experimental results show that the proposed model can improve the prediction precision and reduce the error compared with other models in predicting the same sunspot time series. It can also be applied to other fields after conducting some modification and has high application value.

## Conflicts of Interest

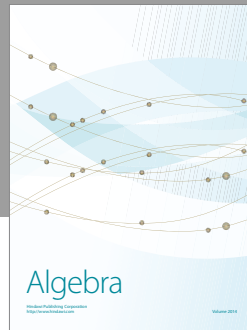
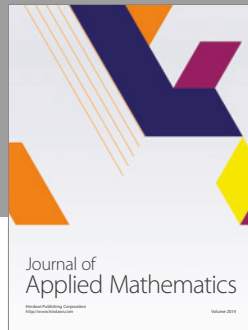
The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51709228).

## References

- [1] G. Li and S. Wang, "Sunspots time-series prediction based on complementary ensemble empirical mode decomposition and wavelet neural network," *Mathematical Problems in Engineering*, vol. 2017, Article ID 3513980, 7 pages, 2017.
- [2] E. M. Roshchina and A. P. Sarychev, "Approximation of periodicity in sunspot formation of and prediction of the 25th cycle," *Geomagnetism and Aeronomy*, vol. 55, no. 7, pp. 892–895, 2015.
- [3] S. D. Taabu, F. M. D'ujanga, and T. Ssenyonga, "Prediction of ionospheric scintillation using neural network over East African region during ascending phase of sunspot cycle 24," *Advances in Space Research*, vol. 57, no. 7, pp. 1570–1584, 2016.
- [4] 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," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, 1998.
- [5] W.-A. Yang, W. Zhou, W. Liao, and Y. Guo, "Identification and quantification of concurrent control chart patterns using extreme-point symmetric mode decomposition and extreme learning machines," *Neurocomputing*, vol. 147, no. 1, pp. 260–270, 2015.
- [6] J. R. Lei, Z. H. Liu, L. Bai, Z. S. Chen, J. H. Xu, and L. L. Wang, "The regional features of precipitation variation trends over Sichuan in China by the ESMD method," *Mausam*, vol. 67, no. 11, pp. 849–860, 2016.
- [7] X. Tian, Y. Li, H. Zhou, X. Li, L. Chen, and X. Zhang, "Electrocardiogram signal denoising using extreme-point symmetric mode decomposition and nonlocal means," *Sensors*, vol. 16, no. 10, article no. 1584, 2016.
- [8] M. Fatehi and H. H. Asadi, "Application of semi-supervised fuzzy c-means method in clustering multivariate geochemical data, a case study from the Dalli Cu-Au porphyry deposit in central Iran," *Ore Geology Reviews*, vol. 81, pp. 245–255, 2017.
- [9] P. R. Meena and S. K. R. Shantha, "Spatial fuzzy c means and expectation maximization algorithms with bias correction for segmentation of MR brain images," *Journal of Medical Systems*, vol. 41, article 15, 2017.
- [10] H. B. Huang, X. R. Huang, R. X. Li, T. C. Lim, and W. P. Ding, "Sound quality prediction of vehicle interior noise using deep belief networks," *Applied Acoustics*, vol. 113, pp. 149–161, 2016.
- [11] Z. Zhao, L. Jiao, J. Zhao, J. Gu, and J. Zhao, "Discriminant deep belief network for high-resolution SAR image classification," *Pattern Recognition*, vol. 61, pp. 686–701, 2017.
- [12] Y.-J. Hu and Z.-H. Ling, "DBN-based spectral feature representation for statistical parametric speech synthesis," *IEEE Signal Processing Letters*, vol. 23, no. 3, pp. 321–325, 2016.
- [13] A.-R. Mohamed, G. E. Dahl, and G. Hinton, "Acoustic modeling using deep belief networks," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 1, pp. 14–22, 2012.
- [14] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.
- [15] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, USA, 1981.
- [16] S. Niazmardi, S. Homayouni, and A. Safari, "An improved FCM algorithm based on the svdd for unsupervised hyperspectral data classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 831–839, 2013.
- [17] F. Liu, L. Jiao, B. Hou, and S. Yang, "POL-SAR Image Classification Based on Wishart DBN and Local Spatial Information," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3292–3308, 2016.
- [18] A. Courville, G. Desjardins, J. Bergstra, and Y. Bengio, "The spike-and-slab RBM and extensions to discrete and sparse data distributions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 9, pp. 1874–1887, 2014.
- [19] Y. Yuan, J. Lin, and Q. Wang, "Hyperspectral image classification via multitask joint sparse representation and stepwise MRF optimization," *IEEE Transactions on Cybernetics*, vol. 46, no. 10, 2015.
- [20] K. Orphanou, A. Stassopoulou, and E. Keravnou, "DBN-extended: A dynamic Bayesian network model extended with temporal abstractions for coronary heart disease prognosis," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 3, pp. 944–952, 2016.



# Hindawi

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

