Geodesy and Geodynamics 2012,3(1):52 -56

http://www.jgg09.com
Doi:10.3724/SP.J.1246.2012.00052

Signal prediction based on empirical mode decomposition and artificial neural networks

Wang Yong[1], Liu Yanping[2] and Yang Jing[3]

[1] *School of Surveying & Land Information Engineering, Henan Polytechnic University, Jiaozuo* 454000, *China*

[2] *School of Civil Engineering, Central South University, Changsha* 410075, *China*

[3] *College of Mining Engineering, Hebei United University, Tangshan* 063009, *China*

Abstract: In view of the usefulness of Empirical Mode Decomposition (EMD), Artificial Neural Networks (ANN), and Most Relevant Matching Extension (MRME) methods in dealing with nonlinear signals, we propose a new way of combining these methods to deal with signal prediction. We found the results of combining EMD with either ANN or MRME to have higher prediction precision for a time series than the result of using EMD alone.

Key words: EMD (Empirical Mode Decomposition); ANN (Artificial Neural Networks); MRME (Most Relevant Matching Extension); IMF (Intrinsic Mode Function); endpoint problem; RBF (Radial Basis Function)

1 Introduction

Empirical mode decomposition (EMD) is a method of transforming an empirical time series into a few Intrinsic Mode Function (IMF) components and a tendency term, which is the final drab and smooth part of the original sequence[1-7]. It is usually applied to deal with some nonlinear or non stationary series. Because of its certain characteristics, such as parallel processing, self adaptivity, self-organization, associative memory, fault tolerance, robustness, it is suitable for application to prediction studies.

In this paper, we show how to use EMD to decompose a simulation signal into several IMF components and a tendency, how to treat the endpoint problem in two ways, how to do signal prediction by using RBF (Radial Basis Function) neural network for each component separately, and how to reconstruct the final prediction results.

Received:2012-02-05; Accepted:2012-02-12

Corresponding author: Tel: +86-13785091437, E-mail: Wangyongiz@126.com

This work was supporteal by the Notional Natural Scince Foundation of Hebei Province (D201000921)

2 EMD and endpoint problem

During the EMD decomposition, the resultant IMF components must meet the following conditions: 1) The number of maximum and minimum points and the number of zero-crossing in different directions must be approximately equal; 2) the mean value of the maximum and minimum at any point must be zero.

The decomposition process is as follows[3]:

1) For a signal $x(t)$, connecting all the maxima with a 3-order spline curve to get the upper envelope, and from the minima to get a lower envelope similarly. Generating a new signal by subtracting the mean of the upper and lower envelope from the original signal.

2) Checking whether the new signal meets the above-mentioned basic requirement of IMF, or whether the residual r is a monotonic function. If not, repeating step 1).

3) At the end, $x(t)$ is decomposed into a time series consisting of a number of IMF components and $r(t)$.

In order to shorten the screening process, a *SD* parameter is defined. When *SD* is less than a certain value, usually chosen to be 0.2 – 0.3, the screening process ends. Since the algorithm uses a cubic spline interpolation, when the number of extreme values is less than two, the process also ends. Cubic spline interpolation introduces error which will spread to the internal data and affect the low-frequency data, causing the so-called endpoint problem. To deal with this problem, we apply the radial basic function neural network and the waveform matching methods.

2.1 Direct EMD method

First we show how to apply EMD directly to a time series $x(t)$ by using toolbox in Matlab programming. Since in the real world, a data sequence usually contains noise, so we add some white Gaussian noise $n(t)$ with *SNR* = 10 dB into our simulation signal.

$$x(t)=0.7\cos(\frac{2\pi}{10}t)+0.8\cos(\frac{2\pi}{30}T)+\cos(\frac{2\pi}{60}t)+0.6\sin(\frac{2\pi}{180}t)+n(t), t=[5,95] \quad (1)$$

where $x(t)$ consists of the following four terms and the noise

$$x_1=0.7\cos\frac{2\pi}{10}t, x_2=0.8\cos(\frac{2\pi}{30}t), x_3=\cos(\frac{2\pi}{60}t),$$

$$x_4=0.6\sin(\frac{2\pi}{180}t), n(t) \quad (2)$$

Figure 1 shows the simulation signal and the IMF components after direct EMD. It may be seen that IMF1 is drawn out as noise, and the remaining components tend to be basically the same as the corresponding original signal components, except at the ends, especially after the third-order.

2.2 EMD with ANN method

Reference[4] used a single layer, single neuron, and linear neural network to generate two additional extreme values at ends of the extension of the data.

In this paper, we use a radial basic function neural network, select 15 consecutive data samples as input

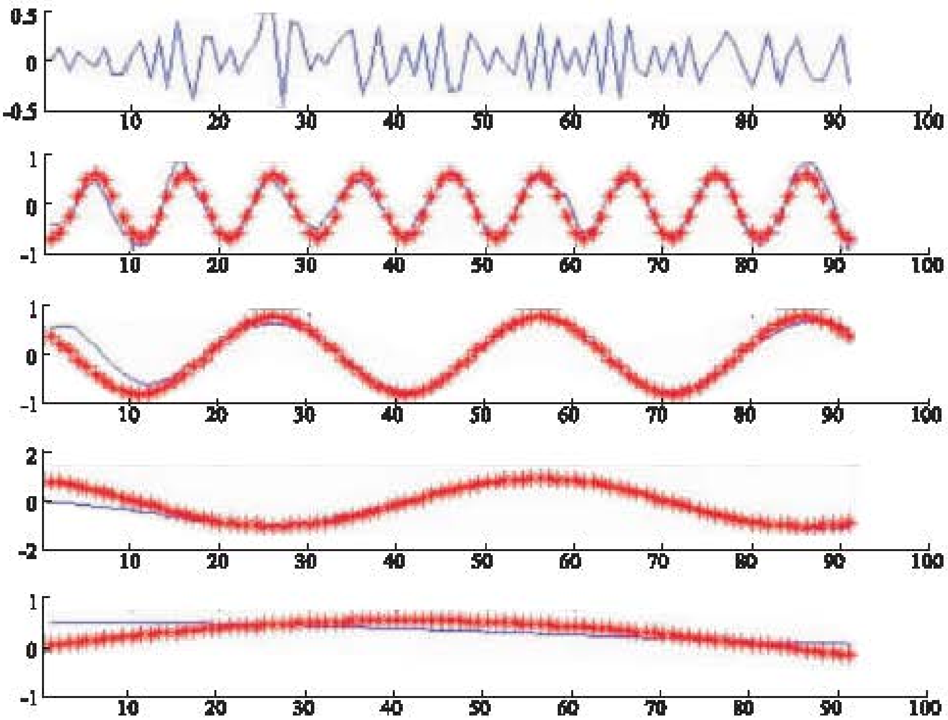

Figure 1 Comparison of simulation signal (dotted lines) with its IMF components (thin lines) after direct EMD

and the next 12 samples as the output, and by reiteration generate 60 groups of training samples. We then select 5 groups as test samples and feed them into the Matlab neural network toolbox. After testing different spread values we choose the the network which has the best performance to generate extreme values at both ends of the data.

As shown in figure 2, after the treatment, the difference between the IMF components and the real signal is significantly reduced, indicating that the neural network endpoint continuation method is effective in solving the problem.

2.3 EMD with MRME method

Based on references [5,6], we propose the most-relevant-matching-extension (MRME) method as follows:

1) Mark all maxima from left to right as $M_1, M_2, \cdots, M_k$ with corresponding time series as $T_{M1}, T_{M2}, \cdots, T_{Mk}$; mark all of the minimum point as $N_1, N_2, \cdots, N_k$ and the corresponding time series as $T_{N1}, T_{N2}, \cdots, T_{Nk}$, mark left point of the signal as x_1 and the corresponding time series as T_1. Let the waveform from x_1 to N_1 be ω_1 and set its length as N; let the length from x_1 to M_1 be M, and $N > M$.

2) Make T_{M1} as the reference point in the waveform and let ω_1 move to the right to overlap T_{M1} and T_{Mi}; calculate ω_i and its correlation coefficient of ω_1.

3) Take the waveform corresponding to the maximum correlation coefficient as the most recent wave. Denote the corresponding maximum value of M_1 at this time as M_p, and calculate the difference between M_1 and M_p; and denote the difference as η. Then the time series corresponding to x_1 should be

$$T_p = T_{Mp} - (T_{M1} - T_1)$$

Add to the real data and extend the actual waveform to the left.

4) Extend the actual waveform to the right similarly. Choose 0.9 or more as correlation coefficients generally. If the requirement is not meet, we may add directly the average value of three neighboring extreme values at the endpoints.

In this way, we can get the IMF components shown in figure 3. From this figures we may see that the difference between IMF components and the true signal is

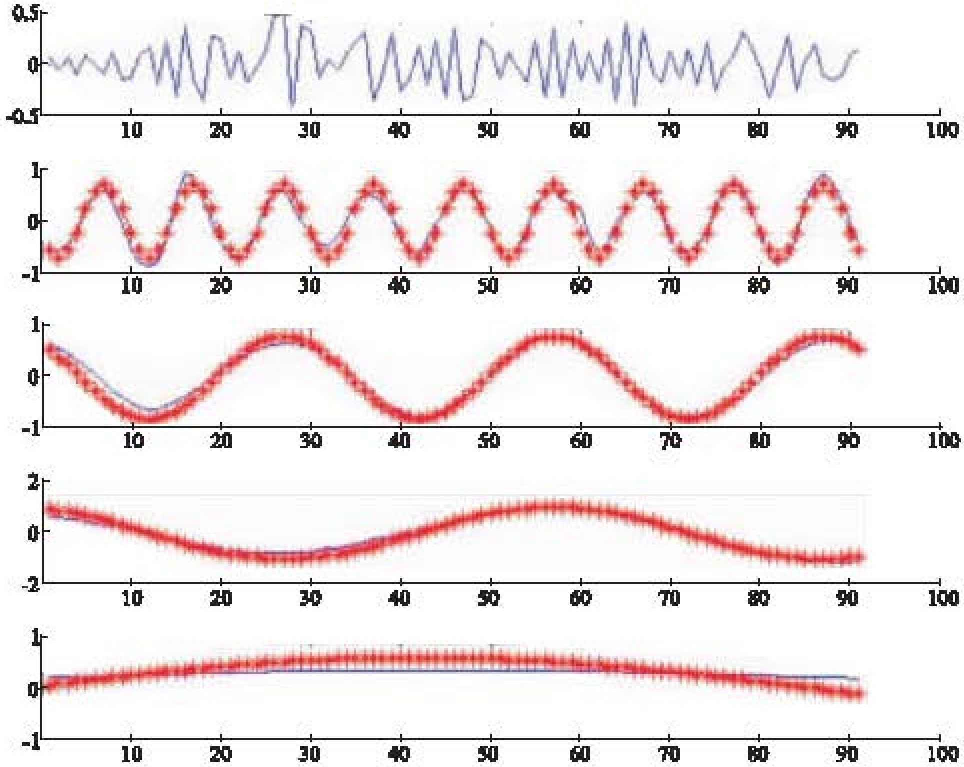

Figure 2 Comparison of simulation signal (dotted lines) with its IMF components (thin lines) after treatment with extension of neural network

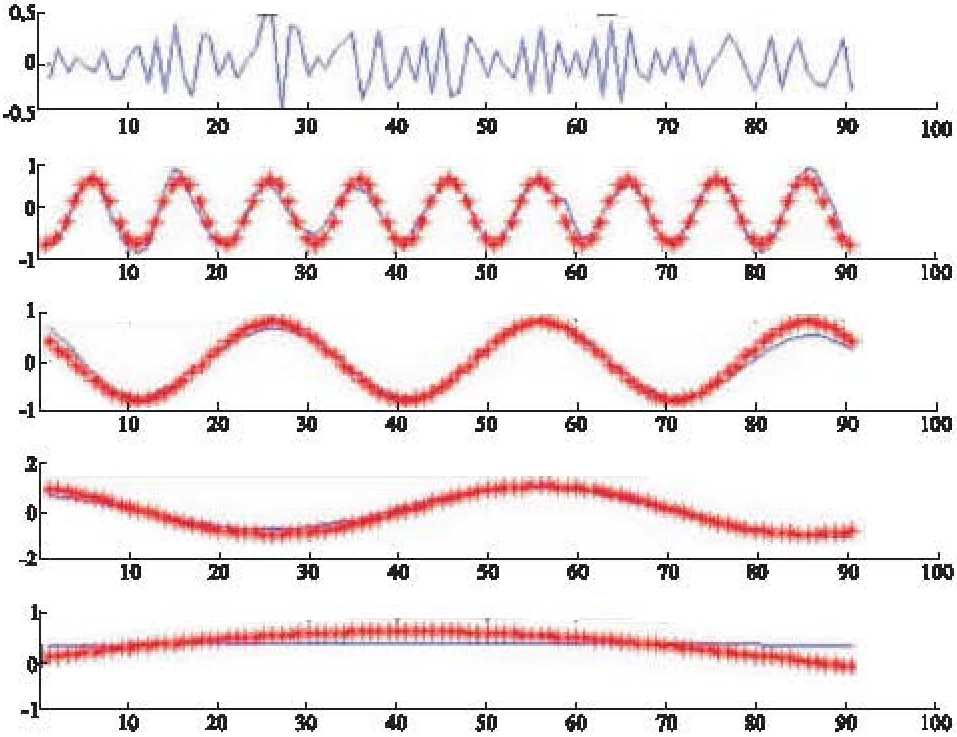

Figure 3 Comparison of simulation signal (dotted lines) with IMF components (thin lines) after treatment with the MRME method.

Table 1 Comparison of the correlations obtained with different methods

	IMF2	IMF3	IMF4	trend
Direct EMD	0.719	0.961	0.925	0.43
Neural networks continuation method EMD	0.955	0.984	0.983	0.97
The most relevant match extension EMD	0.953	0.976	0.985	0.962

smaller than in figure 1, and also the trend lines fit better. This shows the effectiveness of this method in resolving the endpoint problem.

3 Prediction results

After the extension by the different methods and reconstruction of the predicted values of each component, different results are generated. As shown in table 2, the best result is obtained by the ANN method with an mean difference of only 0.45. Prediction from the EMD decomposition result is significantly better than the prediction of the direct results, although some individual errors (such as point 9) are large. The maximum error in the results of the two methods does not exceed 2.4; whereas the maximum error of the direct prediction reaches 3.17.

Table 2 Comparison of errors obtained by different method

	Error of direct EMD	Error of neural networks continuation method EMD	Error of the most relevant match extension EMD
1	2.1	0.164	0.053
2	1.256	0.386	0.006
3	2.002	0.219	0.208
4	3.389	0.136	0.133
5	2.878	0.273	0.159
6	1.656	0.271	0.515
7	1.611	0.389	0.991
8	1.886	0.455	0.845
9	1.445	1.06	1.883
10	2.25	1.18	2.333
11	1.682	0.709	1.771
12	3.171	0.153	1.137
RMS	2.073	0.45	0.836

4 Conclusion

The neural network extension and the most relevant match extension methods are both good solutions to the endpoint-effect problem. EMD decomposition can supply input variables with higher quality to the RBF neural network. The new prediction method presented here can achieve higher precision.

References

[1] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear non-stationary time series analysis. Proceedings of the Poyal Society A, 1998, 454(1971): 903 - 995.

[2] Zhen Zuguang and Liu Hongli. The empirical mode decomposition and wavelet analysis and application. Beijing: China Meteorological Press, 2010, 1 - 83. (in Chinese)

[3] Zhu Jinlong and Qiu Xiaohui. Application with orthogonal polynomial algorithm in the empirical mode decomposition endpoint problem. Computer Engineering and Application, 2006, 23(2): 72 - 74. (in Chinese)

[4] Deng Yongjun, Wang Wei, Qian Chengchun, et al. Treatment the boundary problem in the empirical mode decomposition and the Hilbert transform. Science Bulletin, 2001, 46(3): 257 - 263. (in Chinese)

[5] Shao Chenxi, Wang Jian, Fan Jinfeng, et al. An adaptive extension method of the empirical mode decomposition endpoint. Electronics Journal, 2007, 35(10): 1944 - 1948. (in Chinese)

[6] Wang Ting, Yang Shenyuan and Li Bingbing. A new method to improve the empirical mode decomposition endpoing effect. Harbin University of Science Journal, 2009, 14(5): 23 - 26. (in Chinese)

[7] Dong Changhong. Matlab neural network and application. Beijing: National Defence Industry Press, 2005: 121 - 142. (in Chinese)