Empirical Mode Decomposition and Significance Tests of Temperature Time Series

Zhihua Zhang[1], John Moore[2,3,4]

1. College of Global Change and Earth System Science, Beijing Normal University, China, Email: zhangzh@bnu.edu.cn

2. State Key Laboratory of Earth Surface Processes and Resource Ecology/College of Global Change and Earth System Science, Beijing Normal University, China

3. Arctic Centre, University of Lapland, Finland

4. Department of Earth Sciences, Uppsala University, Sweden

Abstract. Empirical mode decomposition (EMD) has become a powerful tool for adaptive analysis of non-stationary and nonlinear time series. In this paper, we will perform a multi-scale analysis of the Central England Temperature and the proxy temperature from Greenland ice core time series by using EMD. We will make a significance test against the null hypothesis of red noise and determine both the dominant modes of variability and how those modes vary in time.

Keywords: Empirical mode decomposition, Temperature time series, Significance test

1. Introduction

Climatic time series are nonlinear and nonstationary while most traditional time series analysis methods are only valid when applied to stationary time series [1-3, 15]. In order to tackle this issue, Huang [9] introduced Empirical Mode Decomposition (EMD) in 1998. Since then, EMD has become a widely used and powerful tool for adaptive analysis of non-stationary and nonlinear time series. EMD algorithm is based on the direct extraction of the energy associated with various intrinsic time scales, the most important parameters of the system. Since EMD can decompose adaptively a time series into time-frequency space, we are able to determine both the dominant modes of variability in climatic time series and how those modes vary in time.

In this study, we will examine two temperature time series which play a key role in climate change research. The first is the Central England Temperature time series [12]. The other is the Vinther's proxy Greenland winter temperature which is obtained by stable oxygen isotopic composition of ice, specifically the winter value of δ^{18} data from 7 ice cores drilled in southern, western, eastern and central Greenland [8, 11, 16-18]. The isotopic ratio δ^{18} measured in ice cores has long been established as a temperature proxy because of the temperature dependent fractionation of oxygen isotopes that takes place while moisture travels from its evaporation area to the Greenland ice sheet [8, 11, 16-18]. We will perform a multi-scale analysis for the Central England Temperature time series and the proxy Greenland winter temperature time series by using EMD method. After that, we will do a significance test against a suitable hypothesis for climate noise and determine the time variation in the dominant modes of Temperature Time Series.

2. Empirical Mode Decomposition

Huang's Empirical mode decomposition (EMD) method [9-10] is based on the direct extraction of the energy associated with various intrinsic time scales. Since EMD makes full use of the local characteristic time scale of the data, it is suitable to be applied to nonlinear and non-stationary processes.

The EMD method is motivated by the computation of instantaneous frequency defined in terms of the Hilbert transform. In detail, it decomposes a time series into a finite sum of intrinsic mode functions. Huang [9-10] showed that for a function to have physically meaningful instantaneous frequency, it should satisfy:

(a)The number of the extrema and the number of the zero crossings are equal or differ at most by one

(b)At any point, the mean value of the envelopes defined by the local extrema is zero.

Such a function is called an intrinsic mode functions (IMF).

The algorithm for EMD is very simple. Consider a time series $x(t)$. The rule to extract the first IMF from $x(t)$ is as follows:

1)Find the upper envelope of $x(t)$ as the cubic spline interpolant of its local maxima, and the lower envelope, as the cubic spline interpolant of its local minima.

2)Compute the envelope mean $m(t)$ as the average of the upper and lower envelopes.

3)Compute $h(t) = x(t) - m(t)$.

4]If the sifting result $h(t)$ is an IMF, stop and let $c_1(t) = h(t)$. Otherwise, treat $h(t)$ as a new time series and iterate on $h(t)$ through Steps 1-4. The stopping condition is

$$\sum_t \frac{|h_k(t) - h_{k-1}(t)|^2}{|h_k(t)|^2} < SD$$

where $h_k(t)$ is the sifting result in the *k*th iteration, and SD is typically set between 0.2 and 0.3. Finally let $c_1(t) = h_k(t)$. The EMD extracts the next IMF by applying the above procedure to the residue

$$r_1(t) = x(t) - c_1(t);$$

where $c_1(t)$ denotes the first IMF. This process is repeated until the last residue $r_n(t)$ has at most one local extremum. The residue $r_n(t)$ characterizes the nonlinear trend of the time series. For climatic time series, then the residue reflects climate change trend. Finally, we obtain

$$x(t) = \sum c_k(t) + r_n(t)$$

Thus, we achieved a decomposition of the time series into IMFs and a residue which can be either the mean trend or a constant.

3. EMD-based Significance Test

The climate system is undoubtedly a multi-scale system where a multitude of vastly different time and space scales nonlinearly interact with each other. The EMD procedure allows for the analysis of non-stationary time series to extract physically meaningful intrinsic mode functions (IMFs) and nonlinear trends. In order to determine both the dominant modes of variability and how those modes vary in time, the significance of IMFs and trends should be tested against the null hypothesis of some climate noise. However, choice of noise model is crucial to reliable significance testing.

For many geophysical phenomena, a plausible null hypothesis to test against is red noise [7]. A simple model for red noise is the univariate lag-1 autoregressive [AR(1)] process [4] as follows:

$$x_0 = 0, \quad x_{n+1} = \lambda x_n + \varepsilon_n, \qquad n = 1,2,3,\ldots$$

where λ is called AR(1) coefficient and ε_n is independent Gaussian white noise with mean 0 and variance σ^2.

For a given time series, one can use the following formula to estimate AR(1) coefficient and noise variance [1]:

$$\lambda = \frac{\frac{1}{N-1}\sum_{t=1}^{N-1}(x_t - \bar{x})(x_{t+1} - \bar{x})}{\frac{1}{N}\sum_{t=1}^{N}(x_t - x)^2}$$

$$\sigma^2 = \frac{1-\lambda^2}{N}\sum_{t=1}^{N}(x_t - \bar{x})^2$$

Where

$$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t$$

Based on these AR(1) estimators of the climatic time series, we can use the familiar Monta-Carlo method to do the significance test and discover the intrinsic feature in variability in the particular climate time series. That is, we will generate many realizations of the AR(1) model with the same AR(1) coefficient and noise variance as climatic time series. Then, by comparing the variances of each IMF of original climatic time series with that of multiple AR(1) model realizations, we can discover the dominant modes with significant variability and how those modes vary in time.

Many statistical tests assume that the probability density function (pdf) is close to normal. Before one uses the EMD to make the significance test on typical climatic time series, one needs to transform the original time series such that the pdf of the transformed data is normal [13, 14]. The main reason for doing this in our study is because otherwise the normally distributed red noise null hypothesis we use is wrong. So, if the data are not normalized, then the obtained dominant modes with significant variability will be misleading simply because the null-hypothesis can be trivially rejected.

4. Applications

We will examine two important temperature time series with the help of EMD, i.e., we will make a significance test against the null hypothesis of climate noise and determine both the dominant modes of variability and how those modes vary in time.

A.Greenland Winter Temperature Index. We will apply EMD method to the proxy Greenland Winter Temperature [11, 16-18]. We firstly transform the original time series such that the pdf of the transformed data is Normal. A practical way of doing this is by taking the inverse normal cumulative distribution function (cdf). Fig. 1 shows the normalized Winter Greenland Temperature Index. After that, we do the EMD analysis for this normalized temperature time series. Fig. 2 displays all the IMFs and the residue in Empirical Mode Decomposition of the Normalized Greenland Winter Temperature index.

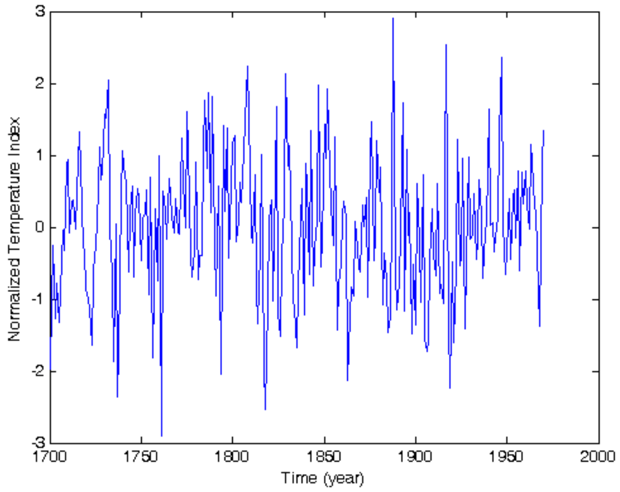

Fig. 1. Normalized Greenland Winter Temperature index

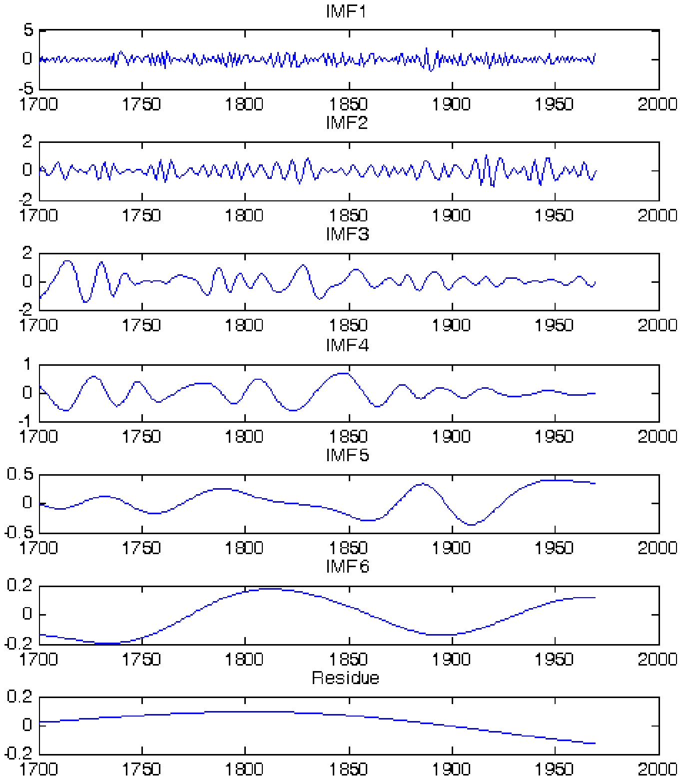

Fig. 2. Empirical Mode Decomposition of Normalized Greenland Winter Temperature index

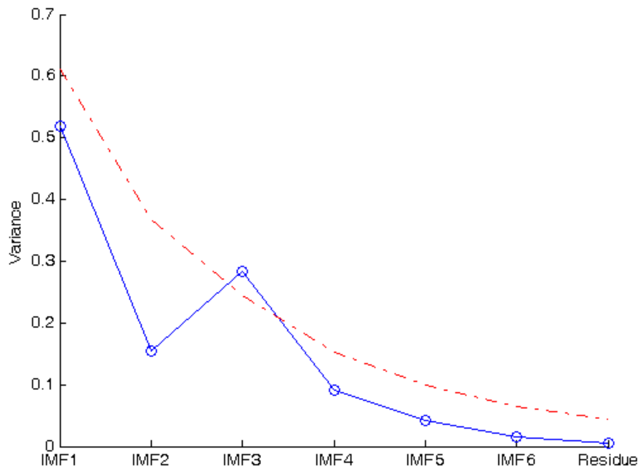

Fig. 3. Statistical significance test for the normalized Greenland winter temperature index.

The variances of each IMF and the residue in the EMD of Greenland winter temperature are marked by circle, while the dashed line denotes 95% percentiles of each IMF variance distribution of AR(1) process. The significant mode for Greenland Winter Temperature index is IMF 3.

The statistical significance of the IMFs will be tested against a climate noise null hypothesis by using Monte-Carlo Method. First of all, we will estimate its AR(1) coefficient and noise variance. After that, we will generate 1000 realizations of the AR(1) model with the same AR(1) coefficient and noise variance as Greenland winter temperature index. In Fig. 3, the variances of each IMF and the residue of Greenland winter temperature are marked by circle, while the dashed line denotes 95% percentiles of each IMF variance distribution of AR(1) process. The statistical significance test reveals that the 95% significant mode is IMF 3 for Greenland Winter Temperature index while the rest IMFs and the residue cannot be distinguished from a climate noise process. In the other words, IMF 3 reflects an intrinsic feature of Greenland Winter Temperature index.

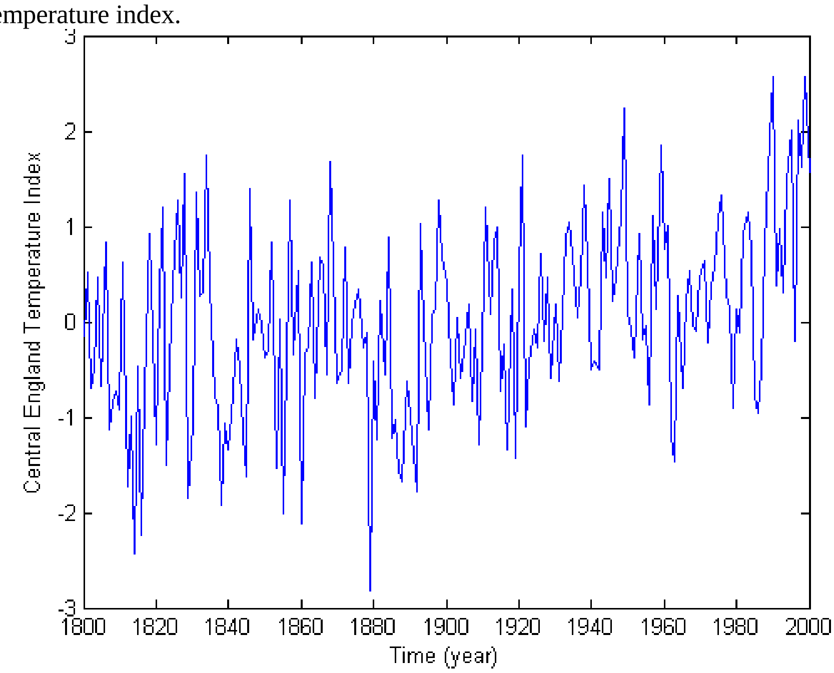

Fig. 4. Normalized Central England Temperature Index

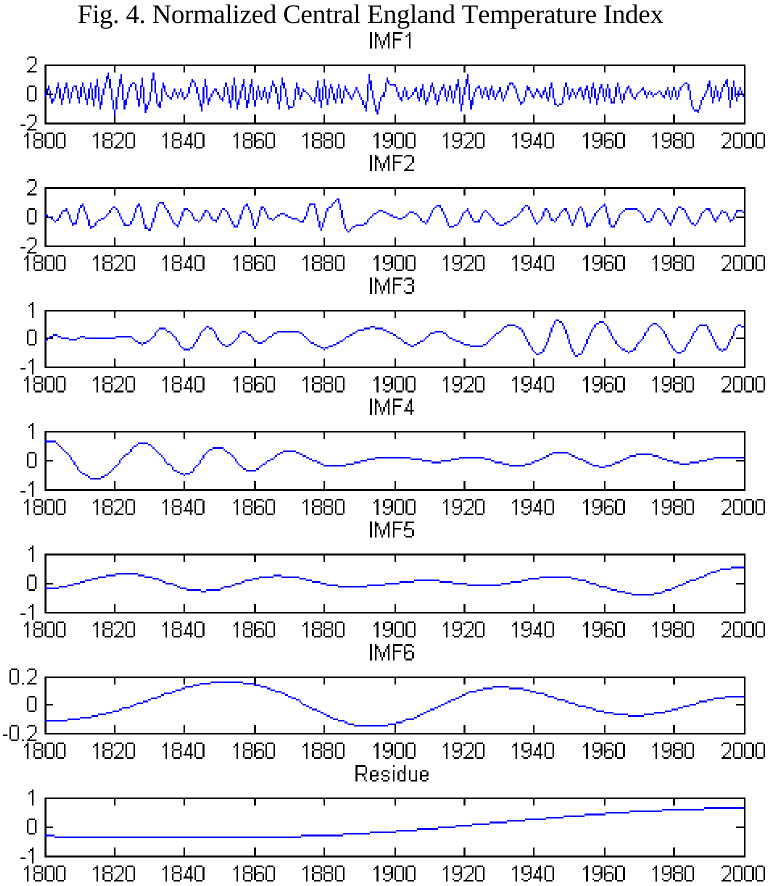

Fig. 5. Empirical Mode Decomposition of the Normalized Central England Temperature index

B. Central England Temperature index.

We examine 1800-2000 Central England Temperature [12]. We will transform the original time series such that the pdf of the transformed data is Normal, as done for Greenland winter temperature (Fig. 4). After that, we decompose the Normalized Central England Temperature Index to six IMFs and one Residue. Fig. 5 displays all the IMFs and the residue in EMD analysis. Finally, we use Monte-Carlo method to do the significance test. Significance testing shows that the only significant mode in EMD analysis is the residue (Fig. 6). Since the residue reflects the temperature change trend, this shows that the climate in Central England has become significantly warmer since 1800. Fig. 5 shows that most of that warming has been through the 20th century.

5. Acknowledgment

This research was partially supported by 973 project No.2010CB950504, 863 project No. 2010AA012305, NSFC No.41076125, Fundamental Research Funds for the Central Universities (Key Program), State Key Lab of Earth Surface Processes and Resource Ecology, and Polar Climate and Environment Key Laboratory.

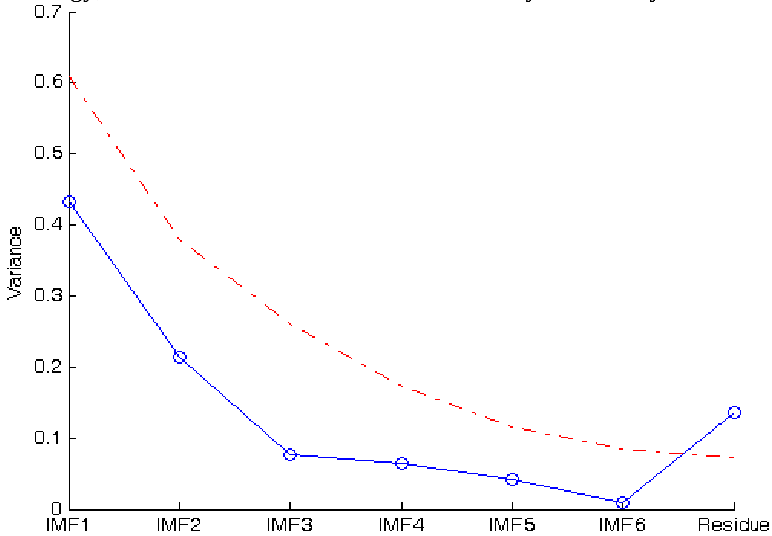

Fig. 6. Statistical significance test for the normalized central England temperature index. The variances of each IMF and the residue in the EMD of the normalized Central England Temperature are marked by circle, while the dashed line denotes 95% percentiles of each IMF variance distribution of AR(1) process. The significant mode for Central England Temperature index is Residue.

References

[1] P.J. Brockwell and R.A. Davis, Time series: theory and methods, Springer-Verlag, New York, 1991

[2] C. Chatfield, The Analysis of Time Series: An Introduction.4th Ed. Chapman and Hall, 1989

[3] M. Ghil, et al., Advanced spectral methods for climatic time series, Rev. Geophys., vol. 40, 2002, pp. 1003-1043.

[4] D.L. Gilman, F. J. Fuglister, and J. M. Mitchell Jr., On the power spectrum of ``red noise", J. Atmos. Sci., vol 20, 1963, pp.182-184.

[5] G.M. Jenkins, and D. G. Watts, Spectral analysis and its applications, Holden-Day, 1968

[6] C. Franzke, Multi-scale analysis of teleconnection indices: climate noise andnonlinear trend analysis, Nonlin. Processes Geophys., vol.16, 2009, pp.65-76.

[7] M.E. Mann, and J. M. Lees, Robust Estimation of Background Noise and Signal Detection in Climatic Time Series. Climate Change, vol. 33, 1996, pp. 409-445.

[8] B.M. Vinther, S. J. Johnsen, K. K. Andersen, H. B. Clausen, and A. W. Hansen, NAO signal recorded in the stable isotopes of Greenland ice cores, Geophysical Research Letters, vol. 30, 2003, pp.1387-1390.

[9] N. E.Huang, et al, The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non- Stationary Time Series Analysis, Proc. R. Soc. Lond. A, vol. 454, 1998, pp. 903-995.

[10] Z. Wu and N. E. Huang, A Study of the Characteristics of White Noise using the Empirical Mode Decomposition Method, Proc. R. Soc. Lond. A, vol. 460, 2004, pp. 1597-1611.

[11] Johnsen, S. J., et al., Stable Isotope Records from Greenland Deep Ice Cores: The Climate Signal and the Role of Diffusion, NATO ASI Series, vol. 156, 1999, pp. 89-107.

[12] G. Plaut, M. Ghil, and R. Vautard, Interannual and interdecadal variability in 335 years of central England temperatures, Science, vol. 268, 1995, pp. 710-713.

[13] S. Jevrejeva, J.C. Moore, and A. Grinsted, Influence of the Arctic Oscillation and El Nieno-Southern Oscillation (ENSO) on ice conditions in the Baltic Sea: The wavelet approach, J. Geophys. Res., vol 108, 2004, pp. 4677-4687.

[14] A. Grinsted, J. C. Moore, and S. Jevrejeva, Application of the cross wavelet transform and wavelet coherence to

geophysical time series, *Nonlin.* Processes *Geophys.*, vol.11, 2003, pp.561-566.
[15] J.H. Zar, Biostatistical analysis, Prentice hall, New Jersey, 1999.
[16] S.J. Johnsen, et al., Stable Isotope Records from Greenland Deep Ice Cores: The Climate Signal and the Role of Diffusion, NATO ASI Series,vol. 156, 1999, pp. 89-107.
[17] C. Appenzeller, et al., The North Atlantic Oscillation and its imprint on precipitation and ice accumulation in Greenland, Geophys. Res. Lett., vol 25, 1998, pp. 1939-1942.
[18] J.W.C. White, et al., The climate signal in the stable isotopes of snow from Summit, Greenland: Results of comparisons with modern climate observations, J. Geophys. Res., vol. 102, 1997, pp.425-26 .