

A new method to remove the tree growth trend based on ensemble empirical mode decomposition

Xianliang Zhang¹ · Zhenju Chen¹

Received: 25 June 2015 / Revised: 18 August 2015 / Accepted: 10 September 2015 / Published online: 3 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract

Key message A new method was developed to remove the tree growth trend, which can be used as an alternative to the traditional method.

Abstract The ensemble empirical mode decomposition (EEMD) method can be used to decompose a non-stationary series into its intrinsic variations and its mean trend. For a tree-ring series, the responses of tree growth to external forces, such as climate factors and disturbances, were shown to be the intrinsic variations of trees. The mean trend obtained using EEMD was calculated as the difference between the tree-ring width and the intrinsic variations of trees (i.e., the external forces) and thus could be regarded as the growth trend of the trees. Having determined the tree growth trend via EEMD, the chronology could then be obtained. This method was compared to the traditional methods of using linear and exponential curves [standard method (STD)] and spline smoothing detrending. The results showed that the chronologies calculated based on EEMD, STD, spline and signal-free detrending were almost the same, apart from slight differences at their ends. This end effect might be an artificial cause of the common problem of “divergence”. However, the end effect was alleviated by the ensemble approach of the EEMD method.

Therefore, the new method based on EEMD is a candidate detrending approach that could be used as an alternative to traditional methods in tree-ring chronology development.

Keywords Ensemble empirical mode decomposition · Chronology · Tree ring · Intrinsic mode function · Growth trend

Introduction

Tree radial growth is controlled by its intrinsic growth trend (increments of age), climate factors such as temperature and precipitation, environmental disturbances, and unexplained variability not related to the other signals (Fritts 1976; Cook and Kairiukstis 1990). However, the response to disturbances is transient and it is short for the growth trend (Cook and Kairiukstis 1990), so the trend component may be composed of a large number of disturbances. Thus, the radial growth is a function of the age trend component and disturbances which can be considered as the stochastic perturbation of the age trend (Cook and Kairiukstis 1990). In climate-related research, the common climatic component among trees is the signal of interest, while the growth trend is considered as non-climatic variance or noise. Therefore, the growth trend should be removed from the ring width series for dendroclimatic studies (Fritts 1976).

The methods used to estimate the growth trend fall into two classes: deterministic and stochastic methods. The deterministic methods of growth trend estimation involve fitting a prior defined mathematical model of radial growth to the ring width series. The most commonly used growth trend-fitting approaches are the linear trend model and the negative exponential curve model (Fritts et al. 1969).

Communicated by E. Liang and A. Bräuning.

Electronic supplementary material The online version of this article (doi:10.1007/s00468-015-1295-z) contains supplementary material, which is available to authorized users.

✉ Zhenju Chen
zhenjuchen@hotmail.com

¹ Tree-Ring Laboratory, College of Forestry, Shenyang Agriculture University, 110866 Shenyang, China

Functions such as the power function (Kuusela and Kilkki 1963) and the “Hugershoff” function (Warren 1980) have also been used deterministically to fit the curve of the growth trend. The stochastic methods are data adaptive and are often chosen by a posteriori selection criteria (Cook and Kairiukstis 1990). These methods are low-pass digital filtering type of methods, such as the cubic-smoothing spline (Cook and Peters 1981), exponential smoothing (Barefoot et al. 1974), and differencing (Box and Jenkins, 1970). In general, stochastic methods fit the behavior of the observed tree-ring width, whereas low-frequency signals can be also removed when using stochastic methods. The deterministic methods have the tree growth behave as the curve model, while the stochastic methods reduce the low-frequency noise unassociated with the signals. Each class of model has its advantages and disadvantages.

In the commonly used program ARSTAN (Cook 1985), which was used to develop the chronologies in many tree-ring researches, the linear trend model and negative exponential curve model were used prior to fit the growth trend of every ring series. The spline curve was used to fit the growth trend when both the above methods failed to fit the growth trend. Thus, the linear trend model, negative exponential curve model, and the spline function have become most commonly used among all the deterministic and stochastic methods (Cook 1985; Jacoby and D’Arrigo 1989; Luckman et al. 1997; Zhang et al. 2011). However, the linear trend model and negative exponential curve model failed to fit the growth trend in some cases, while the spline curve model could not retain as much low-frequency signal. Thus, we aim to develop a new method which could be used as an alternative to the above three traditional methods.

Many existing methods could be used to determine the trend in time series, such as auto-regressive integrated moving average, moving mean method, regression analysis, and Fourier-based filtering method. However, the trend was extrinsically determined by all of these methods (Wu et al. 2007). As a method which can intrinsically fit the data, ensemble empirical mode decomposition (EEMD) is an adaptive, data-driven, and highly efficient algorithm used to decompose non-stationary signals into their intrinsic modes of oscillation and mean trends (Huang et al. 1998, 2003; Lei et al. 2009). It is widely used as an effective method in geophysical studies (e.g., Wu et al. 2007, 2011; Huang and Wu 2008; Zhang and Yan 2014). Tree-ring series are non-stationary signals, because a trend was contained in them. EEMD can therefore be used to decompose a tree-ring series into its intrinsic variations and mean trend. The mean trend can be regarded as the growth trend of the tree-ring series, which, having been identified, can then be removed. Accordingly, the main aims of the present reported study were to (1) identify the growth trend

based on EEMD and (2) compare the results with those obtained using traditional approaches.

Data and method

Tree-ring data

Tree-ring raw data with long records were randomly selected from the international tree-ring data bank (ITRDB) (<http://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets/tree-ring>). This resulted in the tree-ring raw data provided by D’Arrigo et al. (2006) (ID: cana312) being selected to illustrate the performance of our proposed method. Another long tree-ring raw dataset (ID: cana163), provided by Zhang (2000), was also used to develop the chronologies obtained by different detrending methods. The two sample sites of the trees were located in Canada (Fig. 1).

Methods to develop the chronology

The tree-ring chronologies were developed based on the EEMD method. Three steps were involved to develop the chronology. First, the growth trend was identified by the EEMD method in this study. Second, the growth trend was removed from the ring width series. The previous two steps were called as standardization, and tree-ring indices with a defined mean of 1.0 could be obtained by standardization. Third, the chronology could be estimated by averaging the tree-ring indices using the bi-weight robust mean method.

Growth trend identified by EEMD

EEMD is an improved version of empirical mode decomposition (EMD), which was developed by Huang et al. (1998). EMD is an adaptive, data-driven, and highly efficient algorithm used to decompose nonlinear and non-stationary time series into their intrinsic modes of oscillation. A time series $x(t)$ can be decomposed into a finite and often small number of intrinsic mode functions (IMFs) based on EMD,

$$x(t) = \sum_{j=1}^n \text{IMF}_j + r_n,$$

where r_n is the residue of the time series $x(t)$ which reflects the mean trend of $x(t)$. An IMF is defined as any function having an equal number of extrema and zero-crossings (or differing at most by one) and also having symmetric envelopes defined by the local minima and maxima, respectively.

The EMD procedure is implemented through a sifting process: (1) All of the local extremes are identified in the

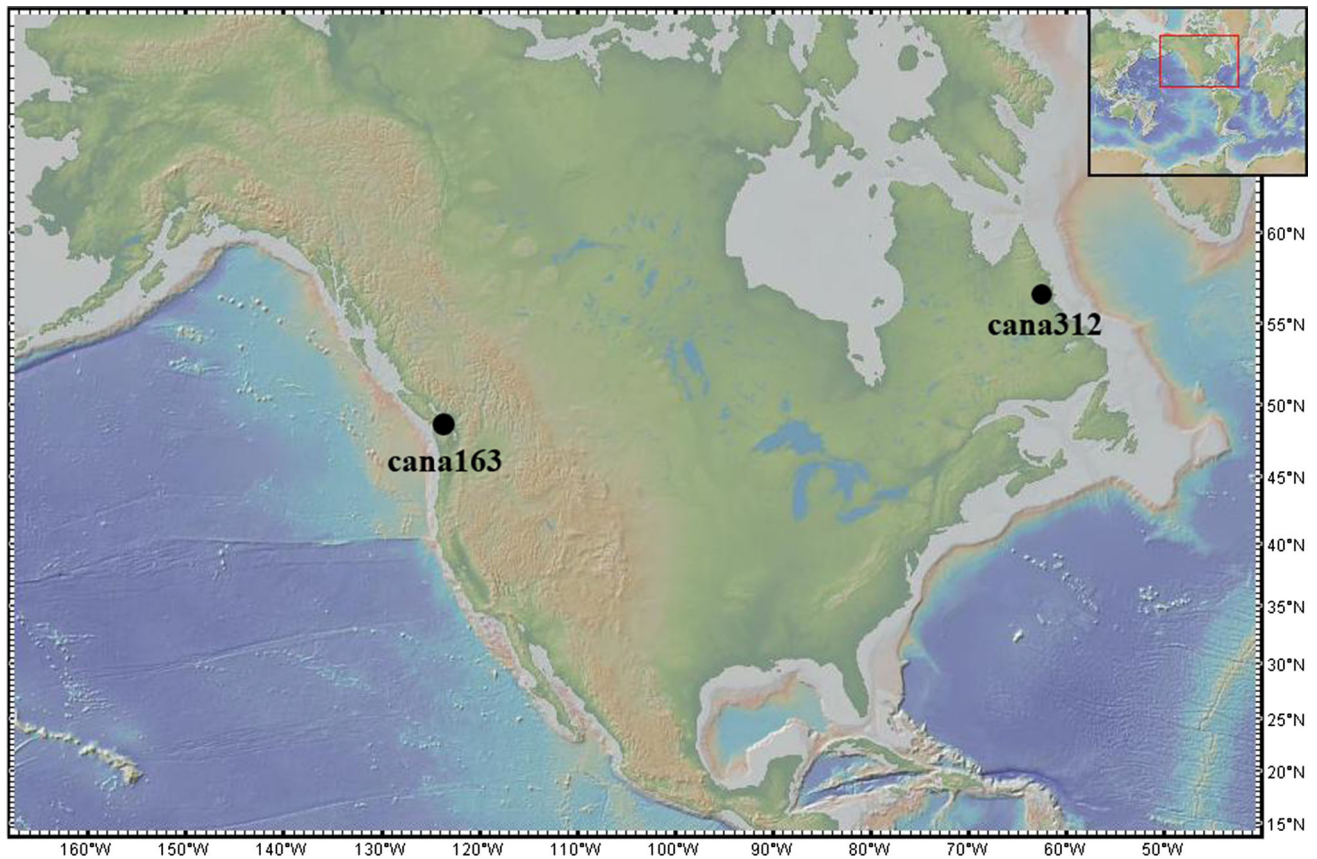


Fig. 1 Sample sites of the two sets of raw tree-ring data (cana163 and cana 312)

time series $x(t)$ and the upper (lower) envelope is obtained by connecting all of the local maxima and minima with a cubic spline. (2) The difference between the data $x(t)$ and the local mean of the upper and lower envelopes is calculated as the first component $h1$. (3) Treating $h1$ as the original data, steps (1) and (2) are then repeated until the upper and lower envelopes are symmetric with respect to zero mean under certain criteria. Then, the final $h1$ is designated as IMF_j and the sifting process continues. The sifting process is completed when the residue r_n becomes a monotonic function. Finally, the data $x(t)$ is decomposed into several IMFs and the residue r_n . More details of the EMD method can be found in Huang et al. (1998, 1999).

Certain weaknesses, however, such as the frequent appearance of mode mixing, could not be solved by EMD. The EEMD method, introduced by Wu and Huang (2009), can be used to alleviate some of the common problems of EMD, such as mode mixing and the increasing robustness of EMD. EEMD involves averaging numerous EMD runs with the addition of some Gaussian noise. The noise is averaged out by averaging the different decompositions, and an estimate of the true decomposition is calculated along with a confidence estimate.

The ensemble number and the standard deviation of the added noise should be assigned in the EEMD procedure. The sensitivity of the decomposition of the data to the amplitude of the noise is often small, within a certain window of noise amplitude (Wu and Huang 2009). The ensemble number and the standard deviation of added noise were tested. Noise with a standard deviation of 0.2 and the ensemble size of 1000 could obtain the optimal results, and then they were used in the present study.

Every tree-ring series was adaptively decomposed to a certain number of components by the EEMD method. These components contained several IMFs and a mean trend (the last component was the mean trend for every tree-ring series). Thus, the growth trend was identified for every tree-ring series.

Standardization

The procedure used to estimate and remove the growth trend was standardization. Standardization transforms the ring width series into tree-ring indices that have a defined mean of 1.0 and a relatively constant variance. After the growth trend was identified for every tree based on EEMD,

tree-ring indices were obtained by dividing each measured ring width by its growth trend. They were calculated as

$$I_t = R_t/G_t,$$

where I_t represents the tree-ring indices, R_t the measured ring width every year, and G_t the expected growth ring width every year, as estimated by the growth trend.

Estimation of the mean chronology

Estimation of the mean chronology was achieved by the bi-weight robust mean method to ensemble the tree indices. In the tree-ring indices, the presence of suspected outliers or extreme values affects the results of the arithmetic mean chronology. The arithmetic mean chronology is no longer a minimum variance estimate of the population mean and may be biased. Thus, the bi-weight mean was used to reduce the variance and bias caused by the outliers. The bi-weight mean for year t is calculated by iteration as

$$\bar{I}_t = \sum_{j=1}^m \omega_j I_{t,j},$$

where

$$\omega_t = \left[1 - \left[\frac{I_t - \bar{I}_t}{cS_t} \right]^2 \right]^2,$$

when

$$\left[\frac{I_t - \bar{I}_t}{cS_t} \right]^2 < 1.$$

Otherwise, $\omega_t = 0$ and $S_t = \text{median}\{|I_t - \bar{I}_t|\}$. Here, c is a constant, often taken as 6 or 9 (Mosteller and Tukey 1977). The iteration stops when an estimate of \bar{I}_t does not change by more than 10^{-3} . Cook (1985) suggested that c should equal 9 in developing tree-ring indices, and so that was the value used in present study.

Results

Tree growth trend identified based on EEMD

One raw tree-ring series was randomly selected to illustrate how the growth trend was identified using this method. The raw data were adaptively decomposed by EEMD into nine components (eight IMFs and the mean trend r_n) for the example tree-ring series (Fig. 2). The different IMFs represented the intrinsic variations of the data at different timescales. A general separation of the data was decomposed into locally non-overlapping timescale components.

Thus, signals of the same timescale would never occur at the same location in two different IMF components. The IMFs represented the intrinsic variations of the data at timescales shorter than the age of the tree. The component r_n of $x(t)$ reflected the mean trend of the tree-ring data and represented the variation of the data at a timescale longer than the age of the tree. In fact, the growth trend of a living tree should be longer than the length of the tree sample. The component r_n of $x(t)$ can be regarded as the tree growth trend.

Comparison of the growth trends identified by EEMD and other methods

Four selected tree cores/radii were used to identify the growth trends using the EEMD method and the methods that conserve the low-frequency signals (hereafter referred to as ‘conservation methods’), such as the linear trend and negative exponential curve method (Fig. 3). The different growth trends of the same trees were identified using different methods, and differences in low-frequency variation of the growth existed at both ends of the tree samples. The most significant result using the different methods was that the tree indices obtained with the different detrending methods showed similar variations, despite the differences at the ends of the tree-ring indices. Moreover, the conservation methods failed to fit the trend for the tree core in Fig. 3d. This tree core was only fitted by the EEMD method. All the tree cores, which were failed to be fitted by conservation methods, could be fitted by the EEMD method.

Comparison of EEMD-based chronologies and other chronologies

Four versions of tree-ring width chronology, based on the samples of cana325, were obtained with the EEMD method, conservation methods (STD, linear trend, negative exponential curve), SPLINE, and signal-free (Ssf) method (Fig. 4). The variations of the four chronologies were very similar ($r = 0.90\text{--}0.97$, $p < 0.01$), especially the high-frequency growth variation in the same year (Fig. 4). Similar to the results of detrending single sample trees, the differences in the chronologies also existed at their beginnings and ends. In terms of the end effect of the chronologies, the values of the EEMD chronology mostly stayed among the values of the STD chronology, the SPLINE chronology, and the signal-free chronology. Similar results were found when using the EEMD-based method and traditional detrending approaches for the trees at site cana163, albeit with slight differences in the end effects among them.

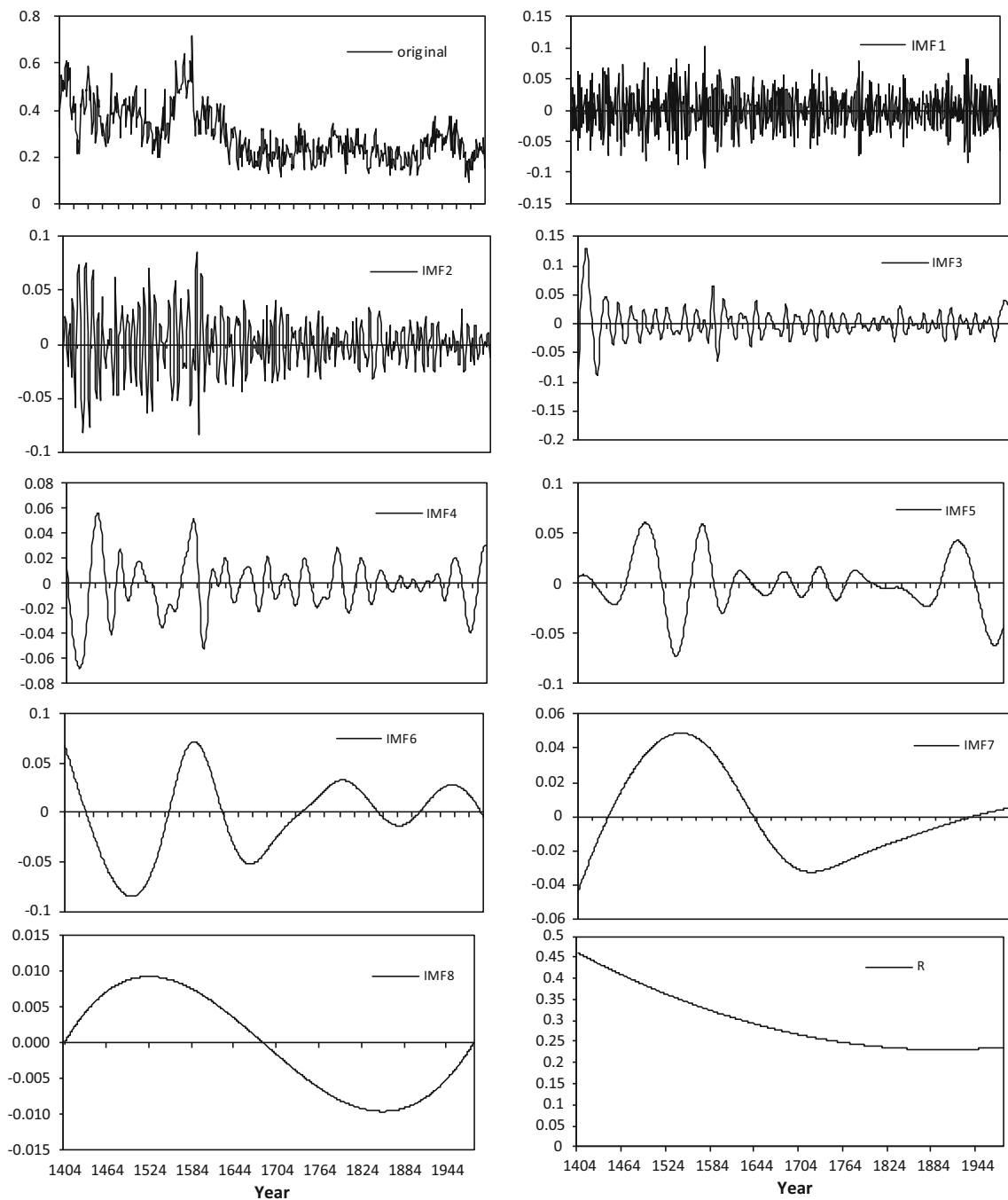


Fig. 2 Different components decomposed by EEMD for an example tree-ring dataset. The IMFs are the different intrinsic variations at different timescales, while *R* is the mean trend of the data

Discussion

In this study, a new dendrochronological detrending method based on EEMD was introduced to remove the tree growth trend. Tree-ring indices can be decomposed into different IMFs and a mean trend by using the EEMD method. The IMFs decomposed by EEMD represented the variations of the original signals at different timescales and

were not mixed over different timescales. Tree-ring samples were usually taken from living trees rather than trees that have died due to old age. Moreover, the timescale of the mean trend was longer than the age of the sample segment. The mean trend contains the long-term trends in growth due to the aging and development of the tree from low-frequency climate signals longer than the age of trees. However, it is impossible to remove the low-frequency

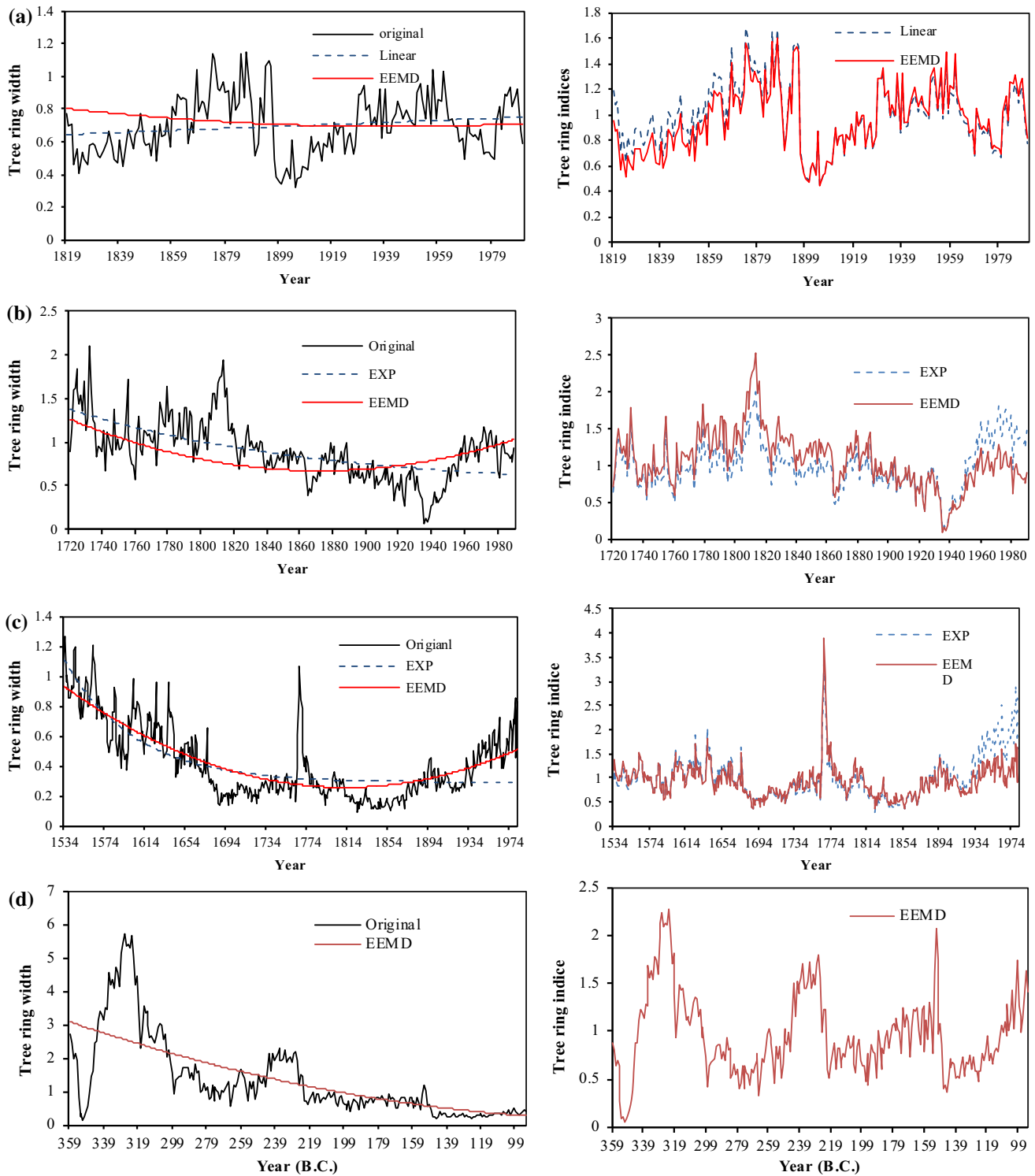


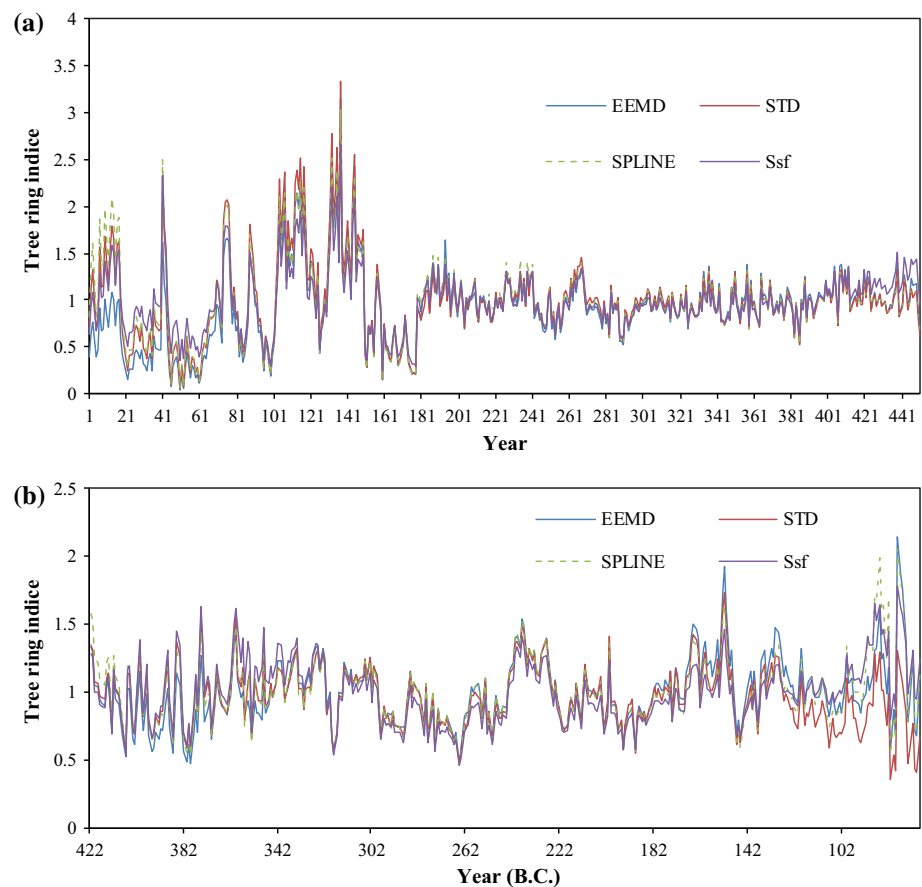
Fig. 3 Growth trends and detrended series for four raw tree-ring datasets. The *left panels* show the growth trends identified by EEMD and traditional methods [linear or exponential curve (EXP)] for four

different raw tree-ring datasets. The *right panels* are the detrended series for each tree-ring dataset. The traditional methods failed to fit the bottom tree-ring series (**d**), so it was only fitted by EEMD

climate signals which were longer than the segment of the ring series with all the detrending methods (e.g., linear trend model, spline curve), which aim to remove the

growth trend for single series. For instance, it is impossible to identify the 1000 years cycle climate signal for the 50 years length of tree core by just fitting a curve for it.

Fig. 4 The chronologies developed based on the EEMD, linear or exponential curve (STD), and spline curve (SPLINE) methods for cana312 (a) and cana163 (b)



The climate signals which were longer than the age would always be accompanied by the aging trend when a curve was fitted for it. However, the low-frequency climate signals could be identified using regional curve standardization (RCS; Briffa et al. 1992) and the “signal-free” method (Melvin and Briffa 2008). In the “signal-free” method, the growth trend should be firstly identified for every series with conventional methods (e.g., linear method, negative exponential method, and spline method); then the regional curve could be obtained. Thus, conventional detrending method is still an important way to remove tree growth trend for every ring series; even longer frequency signals than the age of cores were not well identified with these methods. Our method could be used as an alternative to the conventional method to remove the growth trend of every tree ring. In this situation, the mean trend identified by EEMD could be considered as the growth trend of the tree for every raw series.

Extrinsic functions were used to fit the data in the commonly used methods such as the linear method, negative exponential curve, and the cubic-smoothing spline. However, the externally determined trends did not always correspond to embedded growth mechanisms in the tree-

ring data, as the tree radial growth does not follow a fixed function. Thus, it is impossible to choose an ideal function to obtain the best extrinsically growth trend for trees. Compared to almost all the previous tree-ring detrending methods which require predetermined basis functions, EEMD method is an adaptive method which renders intrinsic tree growth trend. The trend is adaptively fitted by EEMD according to tree growth and reflects the natural growth trend. This is the advantage of EEMD in determining the tree growth trend than other methods.

In many cases, the growth trend does not associate with the fitting curve, such as when the “end effect” occurs (D’Arrigo et al. 2004). The detrending-related “end effect” causes fitting problems at the ends of the tree-ring data (Jacoby and D’Arrigo 1995; D’Arrigo et al. 2004, 2008; Melvin 2004). Like many other commonly used detrending methods, the “end effect” is also a large and currently unavoidable problem that needs to be addressed and clarified when using EEMD, although a major part of the tree-ring indices were consistent when different methods were used in the present study. These are considered to be artificially made in some cases and are not considered to be removed in existing methods such as the linear method and

spline method. However, an ensemble approach was proposed to reduce the end effect once EEMD had been used to decompose the data. It is alleviated as much as possible when the new approach proposed by Huang and Wu (2008) is used, although the end effect cannot be eliminated by EEMD. The works of Huang and Wu (2008) and Wu and Huang (2009) document the principles and advantages of this EEMD approach in detail for the alleviation of the “end effect”. The most important point is that this approach applied with EEMD can also reduce the sensitivity to strangely behaving data at the ends of the targeted time series, and this is the key for obtaining an accurate trend of the data.

The differences in the ends of tree-ring indices for single sample radii or trees obtained with different methods can cause differences in the ends of their chronologies, similar to the “population effect” of single samples. However, the variations in EEMD chronology were found to be similar to those in STD, SPLINE, and signal-free chronologies; in other words, the chronologies obtained with different methods have insignificant differences between them. Moreover, the differences between the EEMD chronology and the other three chronologies were smaller than the differences between the STD chronology and SPLINE chronology (Fig. 4). Therefore, the new detrending method, which was developed based on EEMD, could be used as an alternative to traditional methods for tree-ring chronology development.

Similar to the conventional detrending methods, EEMD-based detrending method was not a perfect method to separate the climate signals, which had longer timescale than the length of tree cores, from the tree growth trend. However, the new method could be used as an alternative to the linear method and negative exponential method as it could preserve more low-frequency signals than the spline curve method. Overall, the EEMD-based detrending method outlined in the present study could be used as an efficient method to remove tree growth trend for single series when developing chronologies. A MATLAB program named as “TREC.m” was attached as supplemental material to develop the EEMD-based chronology.

Author contribution statement ZJ Chen and XL Zhang conceived the idea, and XL Zhang conducted the analysis and wrote the manuscript.

Acknowledgments This work was funded by the National Natural Science Foundation of China (41271066, 31570632 and 41571094) and the Tianzhu-Shan Scholars Program of Shenyang Agricultural University.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Barefoot AC, Woodhous LB, Hafley WL, Wilson EH (1974) Developing a dendrochronology for Winchester, England. *J Inst Wood Sci* 6:34–40
- Box GE, Jenkins GM (1970) *Time series analysis: forecasting and control*. Holden Day, San Francisco
- Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber FH, Karlen W, Zetterberg P, Eronen M (1992) Fennoscandian summers from AD 500: temperature changes on short and long timescales. *Clim Dynam* 7:111–119
- Cook ER (1985) A time series analysis approach to tree ring standardization. Ph.D. Thesis, University of Arizona, Tucson
- Cook ER, Kairiukstis LA (1990) *Methods of dendrochronology: applications in the environmental sciences*. Kluwer Academic, Dordrecht
- Cook ER, Peters K (1981) The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bull* 41:45–53
- D’Arrigo RD, Kaufmann RK, Davi N, Jacoby GC, Laskowski C, Myneni RB, Cherubini P (2004) Thresholds for warming-induced growth decline at elevational tree line in the Yukon Territory, Canada. *Global Biogeochem Cy* 18. GB3021. doi:10.1029/2004GB002249.
- D’Arrigo RD, Wilson R, Jacoby G (2006) On the long-term context for late twentieth century warming. *J Geophys Res* 111 doi:10.1029/2005JD006352
- D’Arrigo RD, Wilson R, Liepert B, Cherubini P (2008) On the ‘divergence problem’ in northern forests: a review of the tree-ring evidence and possible causes. *Global Planet Change* 60:289–305
- Fritts HC (1976) *Tree rings and climate*. Academic press, London
- Fritts HC, Mosimann JE, Bortorff CP (1969) A revised computer program for standardizing tree-ring series. *Tree-Ring Bull* 29:15–20
- Huang NE, Wu Z (2008) A review on Hilbert-Huang transform: method and its applications to geophysical studies. *Rev Geophys* 46:RG2006. doi:10.1029/2007RG000228
- Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond Series A Math Phys Eng Sci* 454:903–995
- Huang NE, Shen Z, Long SR (1999) A new view of nonlinear water waves: the Hilbert Spectrum 1. *Annu Rev Fluid Mech* 31:417–457
- Huang NE, Wu ML, Qu W, Long SR, Shen SS (2003) Applications of Hilbert-Huang transform to non-stationary financial time series analysis. *Appl Stoch Model Bus* 19:245–268
- Jacoby GC, D’Arrigo R (1989) Reconstructed Northern Hemisphere annual temperature since 1671 based on high-latitude tree-ring data from North America. *Clim Change* 14:39–59
- Jacoby GC, D’Arrigo RD (1995) Tree ring width and density evidence of climatic and potential forest change in Alaska. *Global Biogeochem Cy* 9:227–234
- Kuusela K, Kilkki P (1963) Multiple regression of increment percentage on other characteristics in scotch-pine stands. *Acta Forestalia Fennica* 75(4):1–40
- Lei Y, He Z, Zi Y (2009) Application of the EEMD method to rotor fault diagnosis of rotating machinery. *Mech Syst Signal Pr* 23:1327–1338
- Luckman BH, Briffa KR, Jones PD, Schweingruber FH (1997) Tree-ring based reconstruction of summer temperatures at the Columbia Icefield, Alberta, Canada, AD 1073–1983. *Holocene* 7:375–389

- Melvin TM (2004) Historical growth rates and changing climatic sensitivity of boreal conifers. Unpublished PhD thesis, University of East Anglia Norwich
- Melvin TM, Briffa KR (2008) A “signal-free” approach to dendro-climatic standardisation. *Dendrochronologia* 26:71–86
- Mosteller F, Tukey JW (1977) Data analysis and regression: a second course in statistics. Addison-Wesley Series in Behavioral Science: quantitative Methods
- Warren WG (1980) On removing the growth trend from dendrochronological data. *Tree Ring Bull* 40:35–44
- Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adap Data Anal* 1:1–41
- Wu Z, Huang NE, Long SR, Peng C (2007) On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proc Natl Acad Sci* 104:14889–14894
- Wu Z, Huang NE, Wallace JM, Smoliak BV, Chen X (2011) On the time-varying trend in global-mean surface temperature. *Clim Dynam* 37:759–773
- Zhang Q (2000) Modern and late Holocene climate-tree-ring growth relationships and growth patterns in Douglas-fir, coastal British Columbia, Canada. Ph.D. dissertation, University of Victoria, Victoria, BC, Canada
- Zhang X, Yan X (2014) A novel method to improve temperature simulations of general circulation models based on ensemble empirical mode decomposition and its application to multi-model ensembles. *Tellus A* 66:24846. doi:[10.3402/tellusa.v66.24846](https://doi.org/10.3402/tellusa.v66.24846)
- Zhang X, He X, Li J, Davi N, Chen Z, Cui M, Chen W, Li N (2011) Temperature reconstruction (1750–2008) from Dahurian larch tree-rings in an area subject to permafrost in Inner Mongolia, Northeast China. *Clim Res* 47:151–159