THE MODIFIED EMPIRICAL MODE DECOMPOSITION METHOD FOR ANALYSING THE CYCLICAL BEHAVIOR OF TIME SERIES

Vladimir Sebesta, Roman Marsalek and Jitka Pomenkova,
Department of Radio Electronics
Brno University of Technology
BRNO, Czech Republic
Email: marsaler@feec.vutbr.cz

KEYWORDS

empirical mode decomposition; instantaneous frequency; cyclical behavior

ABSTRACT

This paper is devoted to the analysis of time series using the Empirical Mode Decomposition (EMD) method. This method decomposes the analyzed time series into a small set of narrow-band components (modes) that fully represent the original time series. The modified EMD method that eliminates excessive changes of individual mode periods is proposed and evaluated on one example application of industrial production data. In contrast to other decomposition methods, like the singular value decomposition, the empirical mode decomposition can describe the time-variation of the period of individual components.

INTRODUCTION

The Empirical Mode Decomposition (EMD) method followed by the Hilbert spectrum estimation represents an important tool for time-frequency analysis of nonstationary and nonlinear random processes (Huang N.E. et al., 1998). Its advantage is the ability to analyze the local behavior of estimated parameters and thus to quantify the time evolution of the process. EMD has been first proposed for earthquake/wind data analysis (Huang N.E. et al., 1998), the application to time series in economy can be found in (Huang N.E. et al., 2003). Some improvements such as using the beating phenomena of waves (Chen, Y. and M.Q. Feng, 2003) have been researched in the past.

The results of EMD are obtained in the form of so-called Intrinsic Mode Functions (IMF) and the Hilbert spectrum. The Modified EMD (MEMD) method described in this paper is based on the decomposition of the time series to several modes described by IMF. The first step is decomposition using classic EMD followed by the determination of mode boundaries, local periods and local significances. Whenever the local period of the IMF significantly differs from the average mode period, the corresponding segment of the IMF is considered as the different mode. This is an improvement over the existing EMD described in (Huang N.E. et al., 1998, 2003), in which some modes can contain parts with strongly different values of period.

Several other methods can be used for the time-frequency analysis as the Short Time Fourier Transform or Singular Value Decomposition (SVD) (Carvalho, M. et al., 2012) . The first mentioned method requires the time series to be de-trended (e.g. using high-pass filters). Due to non-ideal filter approximation, the de-trending operation influence the analysis results. The SVD does not require de-trending prior to its application similarly to EMD, but it has other disadvantages that will be discussed later.

EMD ALGORITHMS

Original EMD

The EMD method is based on the decomposition of analyzed time series $s(n)$ into IMFs obtained through the sifting process (Huang N.E. et al., 1998). Sifting makes use of two signal envelopes - the one defined by the local minima and the second by the local maxima of the time series. These extremes are connected with cubic splines. From the signal processing theory, the IMFs correspond to narrow-band signals with amplitude and phase modulation and randomly varying parameters (Haykin, S. and B. Van Veen, 2003). The algorithm can be described as:

Sifting:

$$h_{j,k}(n) = h_{j,(k-1)}(n) - m_{j,k}(n), \tag{1}$$

IMF definition:

$$c_j(n) = h_{j,\hat{k}}(n), \tag{2}$$

Remainder calculation:

$$r_j(n) = r_{j-1}(n) - c_j(n), \tag{3}$$

Proceedings 27th European Conference on Modelling and Simulation ©ECMS Webjørn Rekdalsbakken, Robin T. Bye, Houxiang Zhang (Editors)
ISBN: 978-0-9564944-6-7 / ISBN: 978-0-9564944-7-4 (CD)

where n is discrete-time step , j is the index of the j-th IMF $c_j(n) = h_{j,\hat{k}}(n)$, $h_{j,\hat{k}}(n)$ is sifted time series after k iterations, $m_{j,k}(n)$ is the mean between the two signal envelopes, $r_j(n)$ is the $j-$th remainder and $r_0(n) = s(n)$. The sum of all IMFs and the residue is the original time series $s(n)$.

Modified EMD

The application of the EMD method can result in a situation where the instantaneous period of the IMF locally significantly exceeds its mean value and corresponds to the period of another mode. This is evident from the results in (Huang N.E. et al., 2003) and figure 2 below. This effect will be suppressed by the MEMD method described below.

In MEMD, each $j-$th IMF $c_j(n)$ from eq. 2 is analyzed prior to its subtraction from the previous remainder $r_{j-1}(n)$ in eq. 3 as follows:

- $c_j(n)$ is partitioned into L_j segments $c_{j,l}(n)$, where $l = 1 \dots L_j$ with approximately constant instantaneous period per segment
- the instantaneous frequency (equivalently its reciprocal quantity - instantaneous period $T_{j,l}(n)$) and amplitude $a_{j,l}(n)$ of each segment is computed with the use of the Hilbert transform (Haykin, S. and B. Van Veen, 2003), (Huang N.E. et al., 1998). Due to the use of EMD, the instantaneous frequency and amplitude are estimated very locally, practically from three successive samples of data.
- the weighted means $\bar{T}_{j,l}$ of instantaneous period in each segment $c_{j,l}(n)$ are computed as

  $$\bar{T}_{j,l} = \frac{\sum_{n=1}^{N_{j,l}} a_{j,l}(n) T_{j,l}(n)}{\sum_{n=1}^{N_{j,l}} a_{j,l}(n)}, \qquad (4)$$

  where $N_{j,l}$ is the length of the l-th segment. The instantaneous amplitude $a_{j,l}(n)$, corresponding to the period significance is used for weighting.
- if the period $\bar{T}_{j,l_i}$ of any l_i-th segment significantly differs from the other $\bar{T}_{j,l}, \forall l$, the segment $c_{j,l_i}(n)$ is excluded from $c_j(n)$ and its corresponding part remains in the remainder $r_j(n)$.

The principle and usefulness of the proposed modification will be further illustrated on a real time series. Another modification of the original EMD was aimed at improving the performance on a short time series with a small number of maxima and minima. In order not to shorten the analyzed time series in each sifting, both the first local maxima and minima are copied to the beginning of the signal as two new fictive local extremes. A similar arrangement holds also for the end of the sifted time series.

EXPERIMENT AND RESULTS

Analyzed data

The MEMD method has been evaluated and compared with EMD on the industrial production index data of EU15 countries (transformed by natural logarithm) gathered from Eurostat. The period of the analyzed data is 1991/M1-2011/M3 resulting in a total number of 243 samples (months). Note that in all figures, the time axis description corresponds to the month index (discrete time n) with 1 equal to January 2001. For simple notation, n is sometimes omitted in the following text.

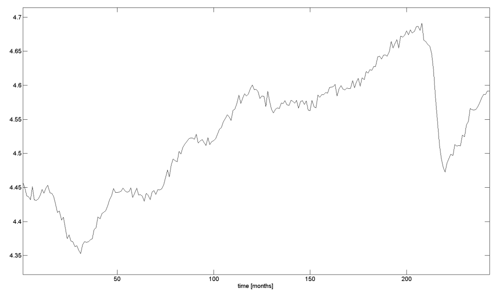

Figure 1: Analysed data

Original EMD results

As the first step, EMD is performed according to (Huang N.E. et al., 1998). The results: IMFs $c_1 - c_5$ are shown in Fig. 2. Apparently the frequency (period) of modes changes. In some cases, e.g. the mode c_1 between month 211 and 221 (marked by an arrow), the instantaneous frequency is much lower and corresponds rather to the instantaneous frequency of mode c_2. This is the motivation for the MEMD method.

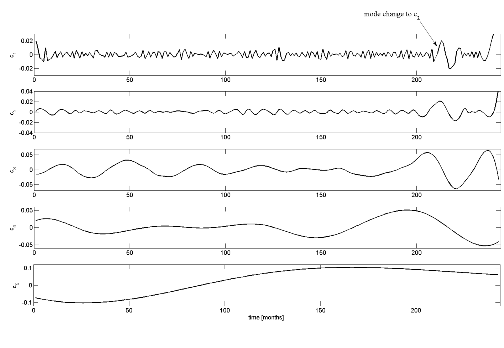

Figure 2: EMD results

Results of modified algorithm

The MEMD analysis will be explained in the example of the first IMF c_1. This IMF has been partitioned into $L_1 = 4$ segments according to the instantaneous frequency.

Analysis of c_1 - 1^{st} segment

The time plot of segment $c_{1,1}$, instantaneous period $T_{1,1}$ and instantaneous amplitude $a_{1,1}$ further denoted as significance are shown in Fig. 3. The mean period $\bar{T}_{1,1}$ in this segment is 0.292 years (1,168 quarters). The instantaneous values of period exceeding the value 1 are considered to be incorrectly estimated and have been excluded from the mean calculation. The reason for their incorrect estimation is that the sampling frequency is too close to the Nyquist theorem (Haykin, S. and B. Van Veen, 2003) limit. As the presence of the aliasing effect is not foreclosed, the real period of $c_{1,1}$ could possibly be equal to $1/\left(0.292 + 12 \cdot N\right) = 0.0648$, 0.0365, 0.0254, 0.0194 … years, where N is integer and 12 corresponds to monthly sampling frequency. It is worth noting that the period of 0.0194 corresponds to the length of one week. The interpretation and possible application of this phenomena is out of the scope of this paper.

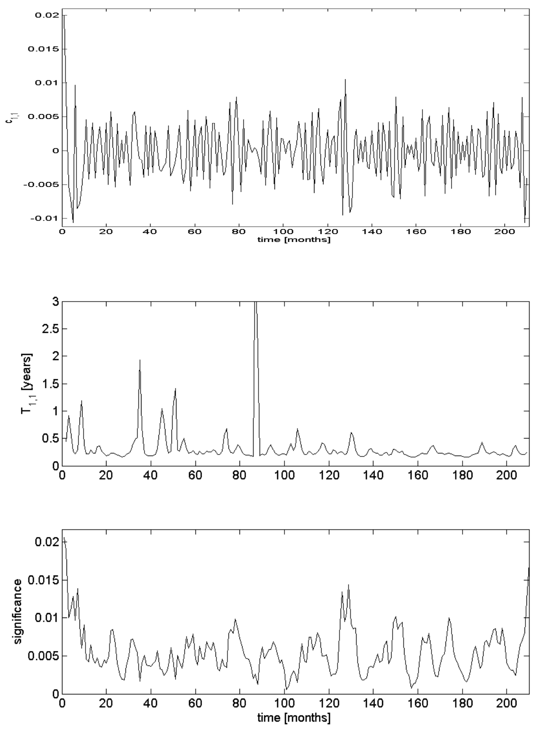

Figure 3: First segment $c_{1,1}$ of the first IMF: time domain plot (top), instantaneous period (middle) and significance (bottom)

Analysis of c_1 - 2^{nd} segment

The time plot of segment $c_{1,2}$, instantaneous period $T_{1,2}$ and amplitude $a_{1,2}$ (significance) are shown in Fig. 4. The mean period $\bar{T}_{1,2}$ in this segment is 0.783 years (3,132 quarters). This period significantly differs from the period in the previous segment $T_{1,1}$. We thus claim that the second segment of the first IMF does not correspond to mode c_1 but rather to a mode with lower instantaneous frequency. The segment $c_{1,2}$ will thus be excluded from c_1 and replaced by zeros. This is also evident from Fig. 7.

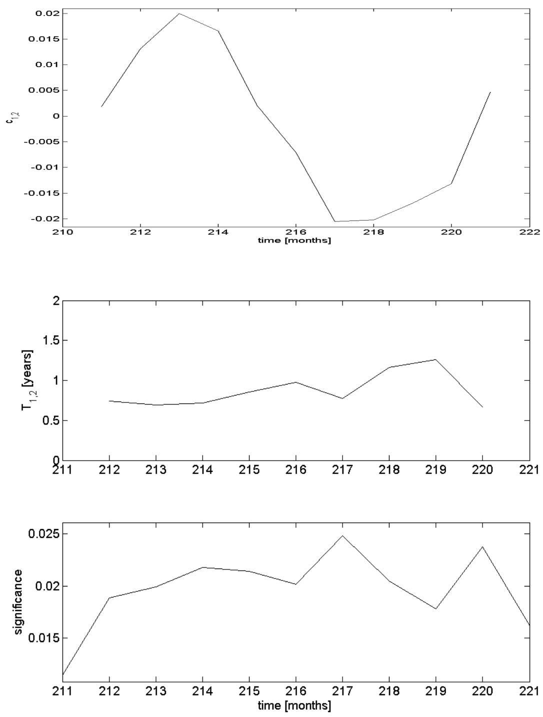

Figure 4: Second segment $c_{1,2}$ of the first IMF: time domain plot (top), instantaneous period (middle) and significance (bottom)

Analysis of c_1 - 3^{rd} segment

In the third segment of c_1 with the analysis results in Fig. 5, the weighted period $\bar{T}_{1,3}$ is similar to the first segment: 0.219 years (0.876 quarters). The mode c_1 is thus present in this segment. The Hilbert transform estimation accuracy suffers from both the short data length and sampling frequency close to the Shannon limit.

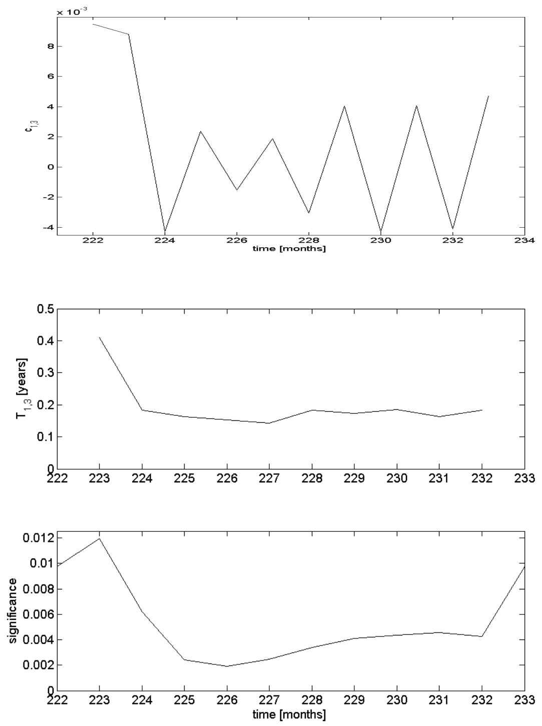

Figure 5: Analysis of third segment $c_{1,3}$ of the first IMF: time domain plot (top), instantaneous period (middle) and significance (bottom)

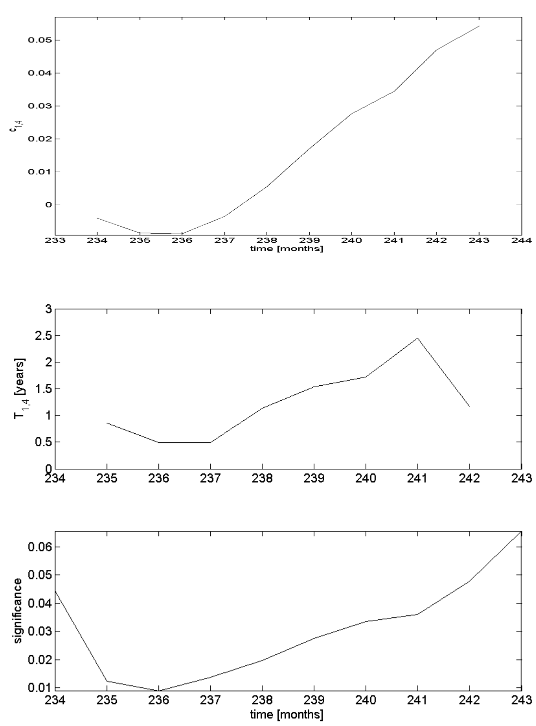

Figure 6: Fourth segment $c_{1,4}$ of the first IMF: time domain plot (top), instantaneous period (middle) and significance (bottom)

Analysis of c_1 - 4^{th} segment

The period in the fourth segment is similar to the period in the second segment - thus it does not correspond to mode c_1 and is excluded. The fourth segment of c_1 is thus replaced by zeros. The $\bar{T}_{1,4}$ is equal to 0.7386 years (2.95 quarters).

$\bar{T}_{\text{column,row}}$	1	2	3	4
1	1.168	*3.132*	0.876	*2.956*
2	2.812	*13.124*	3.864	-
3	13.324	-	-	-
4	*112*	30.88	-	-
5	104	-	-	-

Table 1: Estimated weighted mean periods (in quarters) in segments of modes $c_1 - c_5$

Analysis of $c_2 \ldots c_5$

Because of space limitations, the results of the segment-by-segment analysis of modes in $c_2 - c_5$ are not presented in detail. The identified weighted mean periods for all segments of IMFs c_1 - c_5 are summarized in Table 1. The segments with the values in italic have been excluded from the corresponding IMFs. It can be concluded, that only the modes c_3 and c_4 lie in the business cycle frequency range (6 to 32 quarters). Mode c_3 is present during the whole observed time range, while c_4 does not exist before month 120.

Time plots of all IMFs of MEMD are shown in Fig. 7.

COMPARISON WITH SVD

The application of SVD for band-pass filtering of US business cycles has been demonstrated in (Carvalho, M. et al., 2012). The number of SVD components into which the time series is decomposed depends on the observation window length. Having a small number of components (e.g. 6) results in the superposition of several inseparable narrowband signals in each of them - see Fig. 8 left. On the contrary, a high number of decomposed components (and thus large window length) makes the interpretation of the results unfeasible - it is not possible to distinguish between two adjacent components as the result of their very close period (Fig. 8 right).

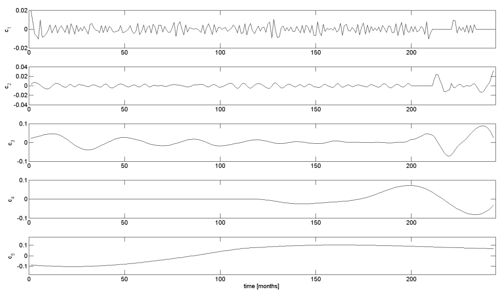

Figure 7: MEMD results

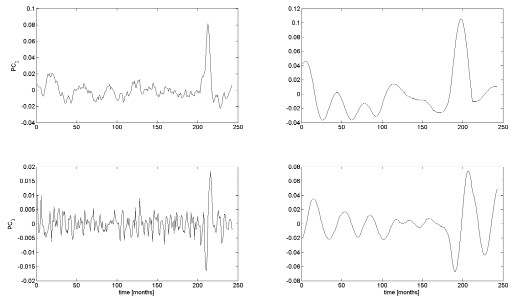

Figure 8: SVD as reference method: 2nd (PC_2) and 3rd (PC_3) components for observation window lengths 6 (left) and 32 (right)

CONCLUSIONS

In this paper, the Modified Empirical Mode Decomposition method has been proposed and evaluated. The proposed modification eliminates the excessive local changes of an instantaneous period of decomposed components. Although the proposed method represents only one possibility among various decomposition techniques, it is advantageous in a small number of resulting components and easier interpretation in comparison with the above mentioned alternatives. Due to using the instantaneous frequency estimation by applying the Hilbert transform, the temporal variation of the components can be easily observed.

From among the wide range of prospective applications ranging from the geophysical or biomedical signal analysis, electrical engineering to economy, the evaluation of described method on the data representing the industrial production index of EU15 is presented. The MEMD method identified two important cycles corresponding to the business cycle frequency range (6-32 quarters) in the analyzed data. The first is the component with a period of 13.3 quarters present in the whole data set, the second is the component with a period of 30.9 quarters present in the period December 1998-March 2011.

Acknowledgements

This research was performed in the laboratories of the SIX research center, reg. no. CZ.1.05/2.1.00/03.0072. It was partially supported by the Czech Science Foundation grant No. P402/11/0570 and by the projects CZ.1.07/2.3.00/20.0007 WICOMT and internal project FEKT-S-11-12 (MOBYS).

References

Huang N.E. et al. 1998. ”The empirical mode decomposition and Hilbert spectrum for non-linear and non-stationary time series analysis”, *Proc. Roy. Soc London A*,Vol. 454, 1998, 903–995.

Huang N.E. et al. 2003. ”Applications of Hilbert-Huang transform to non-stationary financial time series analysis”, *Applied stochastic models in business and industry*,2003; 19, 245-268.

Chen, Y. and M.Q. Feng. 2003. ”A technique to improve the empirical mode decomposition in the Hilbert-Huang transform”, *Earthquake engineering and engineering vibration*, Vol. 2, No. 1, June 2003, 75–85.

Carvalho, M. et al. 2013. ”Tracking the US business cycle with a singular spectrum analysis”, *Economics Letters*, 114 (2012), 32 - 35.

Haykin, S. and B. Van Veen. 2003. *Signals and Systems*, John Wiley, 2003

AUTHOR BIOGRAPHIES

VLADIMIR SEBESTA is the emeritus professor at the Department of Radio Electronics, head of the Signal processing laboratory of Institute of Radio electronics, Brno University of Technology until 2006. His research interests include digital signal processing and analysis, communication theory and wireless communications. His email adress is `sebesta@feec.vutbr.cz`.

ROMAN MARSALEK received his Master degree in control and measurements in 1999 from the Brno University of Technology and the Ph.D. in electronics and signal processing from Universite de Marne la Vallee, France in 2003. His email adress is `marsaler@feec.vutbr.cz`.

JITKA POMENKOVA received the Ph.D. degree in applied mathematics at Ostrava University in 2005, the habilitation degree in Econometric and operational research at Mendelu in Brno in 2010. From 2011 she is the Senior researcher at Department of Radio electronics, Brno University of Technology. Her email adress is `pomenkaj@feec.vutbr.cz`.