15th International Research/Expert Conference
"Trends in the Development of Machinery and Associated Technology"
TMT 2011, Prague, Czech Republic, 12-18 September 2011

# ON-LINE EMPIRICAL MODE DECOMPOSITION OF ENVIRONMENTAL DATA

**Milan Hofreiter**
**CTU in Prague, Faculty of Mechanical Engineering**
**Technicka 4, 166 07 Prague 6**
**Czech Republic**

**Pavel Trnka**
**CTU in Prague, Faculty of Mechanical Engineering**
**Technicka 4, 166 07 Prague 6**
**Czech Republic**

## ABSTRACT
*This paper is devoted to an analysis of environmental time series using the empirical mode decomposition (EMD) method. This is a new method, which was developed for the off-line analysis of non-stationary and stochastic signals. The purpose of the work was to modify the EMD algorithm for on-line analysis and to find out the states of ecosystems in real time. The practical realization of the algorithm was implemented using the MATLAB software and its simulation toolbox, SIMULINK. The modified algorithm was verified using the environmental data that was measured by meteorological stations deployed in the southern part of the Czech Republic. The results obtained are shown graphically.*
**Keywords:** environmental series, on-line, EMD, non-stationarity

## 1. INTRODUCTION
Environmental time series represent most non-stationary non-linear stochastic processes. The available data analysis methods are for linear but non-stationary processes or for nonlinear but stationary and statistically deterministic processes [2, 3]. However, these results can lead to incorrect conclusions. The Hilbert-Huang transform (HHT) is a new method, which was developed by N.E.Huang et al. [1] for the analysis of non-stationary and non-linear signals. The HHT consists of two parts, the empirical mode decomposition (EMD) followed by the Hilbert spectral analysis. The EMD algorithm decomposes a real signal into components, called intrinsic mode functions (IMF) to which the Hilbert analysis can be applied. The IMFs have time variable amplitudes and frequencies. The process which decomposed the original signal $x(t)$ is called sifting. The EMD algorithm decomposes adaptively a signal $x(t)$ into the intrinsic mode functions and a residue $r(t)$:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r(t), \tag{1}$$

The residue $r(t)$ is a monotonic functions which reflects the average trend of a signal $x(t)$. The intrinsic mode functions $c_i(t)$, $i=1,2,...,n$ are signals with following two properties:
a) In the whole data set the number of extremes and the number of zero crossings must either be equal or they can differ by one at most.
b) The mean value of the envelope defined by the local maximums and the envelope defined by the local minimums is zero everywhere.
The sifting process is described by the following steps:

(1) $n$=1; $r(t)=x(t)$;
(2) Identify all extremes (maximums and minimums) of the signal $x(t)$ and connect these maximums (minimums) with the cubic spline line to construct an upper envelope $E_u(t)$ (a lower envelope $E_l$).
(3) Calculate the mean of the upper and lower envelope

$$m(t)=\frac{E_u(t)+E_l(t)}{2}. \qquad (2)$$

(4) Subtract the mean $m(t)$ from $x(t)$

$$h(t)=x(t)-m(t). \qquad (3)$$

(5) If $h(t)$ satisfies condition for IMFs go to the following step else $x(t)=h(t)$ and repeat steps 2 to 4.
(6) $n=n+1$; $c_n(t)=h(t)$; $r(t)=r(t)-c_n(t)$;
(7) If $r(t)$ is not a monotonic function then $x(t)=r(t)$ and go to step (2).
(8) End of EMD

## 2. ON-LINE ANALYSIS

The algorithm described in Chapter 1 is calculated off-line over the entire measured data range. Since the dataset can be very large, the sifting process can be time-consuming and computationally very demanding [5]. In many cases it would be useful to process the data on-line. The algorithm should then be based on observation up the current time. The on-line computation of the algorithm must also be done in such a way that the processing of the measurements can be completed during one sampling interval. For accelerating of the decomposition, the original EMD algorithm was corrected in the following way:

- The floating time window is created. The window range is $\langle (t-T_w),t \rangle$, where $t$ is the actual time point and $T_w$ is the length of the window. The decomposition process is, at any point in time, evaluated only in the appropriate window range. So the time needed for the sifting process does not increase with the current time.
- To reduce a distortion of the decomposition at the current time $t$, the currently known courses of IMFs are used for the estimates of their future courses.

The window range $T_w$ affects frequencies, which will be detected during the sifting process in the IMFs. Generally, the longer window, the longer frequencies that will appear in IMFs.

## 3. APPLICATION

The algorithm is programmed in the MATLAB environment and its simulation toolbox SIMULINK. For this purpose the function „eemd( )" formed by [3, 4] was modified and created the s-function called "s-emd( )".
The on-line EMD algorithm was demonstrated on the soil temperature collected at the meteorological station called "Vrt Domanín" near the town of Třeboň in the southern part of the Czech Republic. Figure 1 shows the soil temperature $\vartheta$ during September 2010. The sampling period was 10 minutes; the position of sensor was 10 mm under the surface.
The experimental analysis has been divided into two parts. An off-line EMD analysis was carried out in the first part. The resulting functions are represented by the thin lines in Figure 2 and they serve as reference samples of the IMF functions.

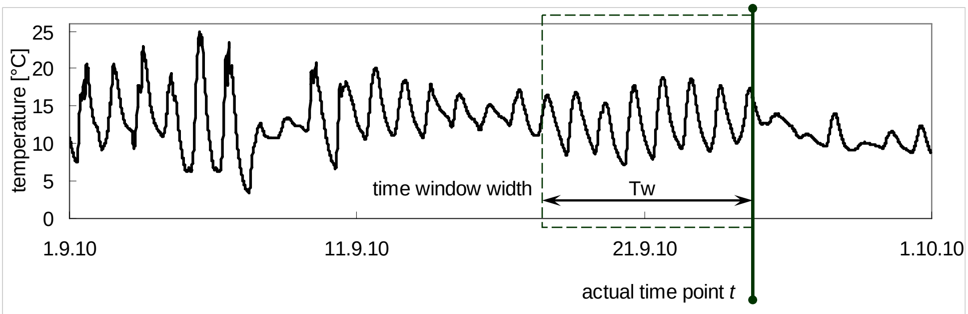


*Figure 1. The measured temperature $\vartheta$ with symbolic representation of moving time window.*

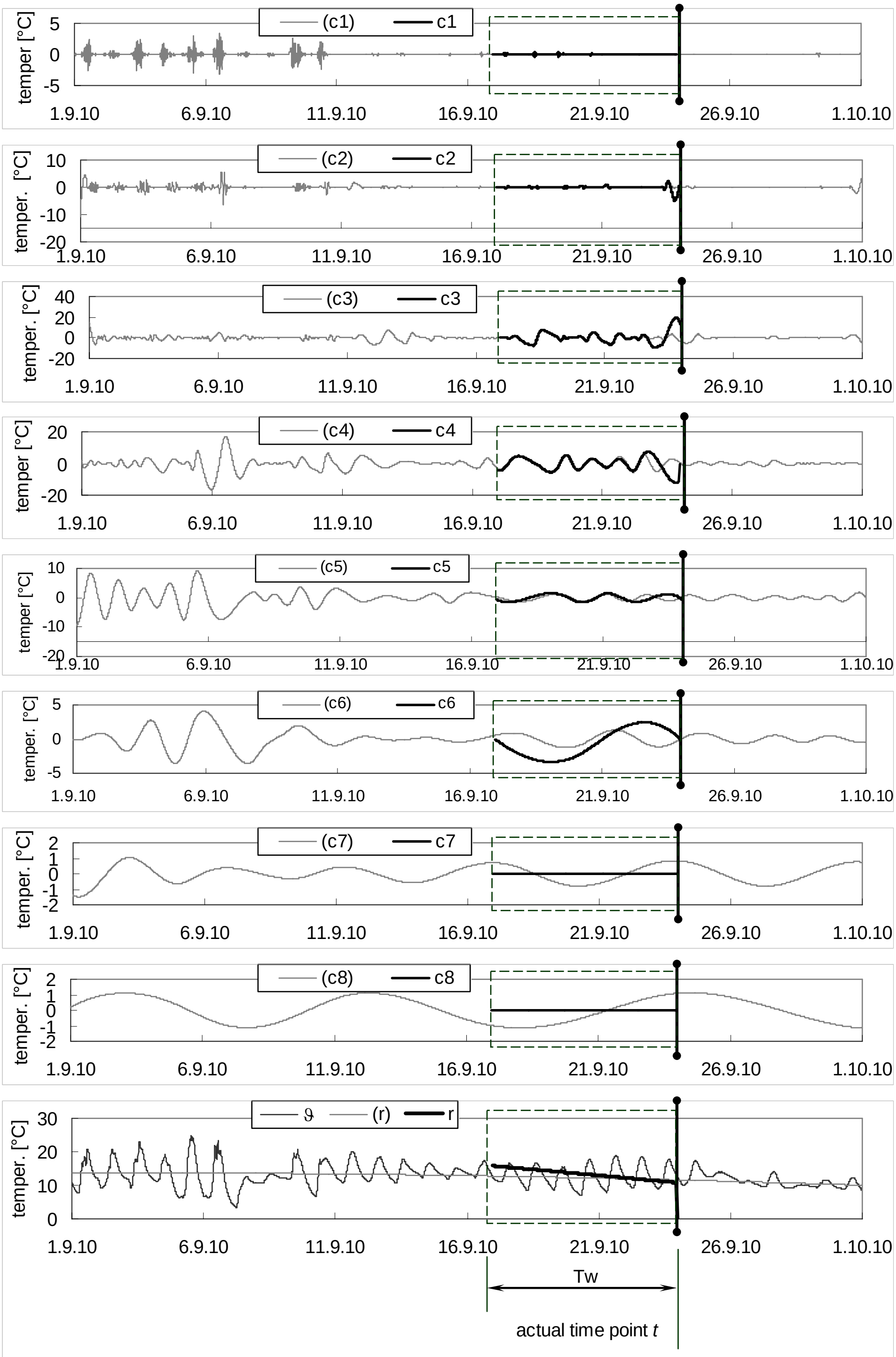


*Figure 2. Comparison of on-line IMF obtained by on-line EMD analysis for one time point t (bold line) and offline IMF (thin line). The number of on-line and off-line IMF differs, so the comparison mainly illustrates how the sifting process depends on the window size.*

The same source data (Figure 1) was processed using the on-line EMD algorithm in the second part of the experiment. The selected time window was seven days which meant a total of 1008 samples were used.

The rectangle shown in Figure 1 represents the time window during the simulation. One set of IMFs that was generated inside the time window is shown in Figure 2. It is obvious that the residue is more curved in comparison to off-line decomposition and also some IMFs have a higher variance. This phenomenon is a necessary and expected consequence of a shorter time range of the analyzed data.

## 4. CONCLUSION

The preliminary results show that the EMD has good performance for analyzing environmental data sets without removing trend or make additional assumptions. The EMD algorithm can be treated as a time-frequency filtering method. Real-time data processing provides results qualitatively similar to off-line analysis. Comparing of both of these methods shows that on-line analysis with a sliding time window is much faster and significantly reduces the computational complexity. Off-line analysis on the contrary provides a slightly more detailed decomposition because it captures even very low frequencies. Data obtained from the experiments provided a number of suggestions for further work. The next study will be focused on on-line identification, prediction and fault detection using this technique as a solution of environmental problems.

## Acknowledgements

***The work is supported by the Ministry of Education of Czech Republic under Project TOKENELEK, no. 2B06023.***

## 6. REFERENCES

[1] Huang N. E., Shen Z., Long S., Wu M. C., Shih H. H., Zheng Q., Yen N.C., Tung C. C., Liu H. H.: The Empirical Mode Decomposition and Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis, Proc. R. Soc. London A 454,pp. 903-995, 1998., [cit. 2010-10-06], Online: <http://keck.ucsf.edu/~schenk/Huang_etal98.pdf> .,

[2] Huang N. E., Shen Z., Long R. S.: A New View of Nonlinear Water Waves—The Hilbert Spectrum. In: Ann. Rev. Fluid Mech. 31, p. 417-457, 1999.,

[3] Zhaohua W., Huang N. E.: Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. In: Advances in Adaptive Data Analysis, Vol. 1, No. 1, p. 1–41, World Scientific Publishing Company,2009.,

[4] Zhaohua W. (2010). HHT MATLAB Program. [cit. 2010-10-06], Online: <http://rcada.ncu.edu.tw/research1_clip_program.htm>.,

[5] Hofreiter, M. (2010). The Application of Hilbert-Huang Transform to Non-Stationary Environmental Data Sets. In: TMT 2010. Zenica: Faculty of Mechanical Engineering in Zenica, 2010, p. 309-312. ISSN 1840-4944.,

[6] Ward, A. D., Trimble S. W.: Environmental Hydrology, CRC Press LLC, Boca Raton. 2004.,

[7] Rao A. R., Hsu E-C.: Hilber-Huang Analysis of Hydrological and Environmental Time Series, Water Science and Technology Library, Vol. 60, Springer Science+Busines Media, 2008.