Empirical mode decomposition method in deformation measurement data processing

MD Parushkin¹,², SV Panov¹*, VM Semibalamut³ and YuN Fomin¹,³

¹Institute of Laser Physics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia
²Institute of Mining, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia
³Siberian Division of the Geophysical Service of the Russian Academy of Sciences, Novosibirsk, Russia

E-mail: *lss@laser.nsc.ru

Abstract. The paper describes the method of empirical mode decomposition and its application in the analysis of data obtained by laser measurement of deformation. The issues of the method adaptation to actual data are discussed. The described method allows tracing monotonous trend and slow term of deformation much more efficient as against the conventional techniques. Furthermore, the method is also good in removal of noise and high-frequency components from signals.

1. Introduction

During long-term continuous operation of laser deformograph installed in a gallery of Talaya seismic station in the Baikal Rift Zone, vast experimental information has been accumulated [1, 2]. The observation data are mostly processed using the popular and common methods of geophysical and deformation measurement data analysis. The main problem connected with the application of these classical approaches is their applicability chiefly to linear and stationary systems. On the other hand, the international practice has advanced in the recent decades in the analysis of nonlinear but stationary and deterministic systems, or linear but nonstationary systems (for instance, Wigner–Ville distribution, wavelet analysis, etc.). However, majority of actual physical processes and systems are nonlinear and nonstationary. In this connection, the analysis of recorded data has to be simplified, especially relative to the a priori set basis of analyzed signal.

One of the key conditions of the correct presentation of nonlinear and nonstationary data is forming an adaptable basis governed by the content of the data. Such approach is implemented in the Hilbert–Huang Transform procedure (HHT), which is the decomposition of signal into empirical modes and application of Hilbert Transform to the resultant components of the decomposition. By now there is no rigorous mathematical validation of the method; at the same time, the positive results of the method application to a wide range of practical tasks enables expecting improvement of the situation in the nearest future.

The empirical mode decomposition method was for the first time been proposed for the analysis of surface typhoon waves by Huang in 1995. The method was generalized for the analysis of arbitrary time series in 1998 [3, 4].

Published under licence by IOP Publishing Ltd
The method of decomposition is based on the assumption that all data under analysis are composed of various oscillation processes and at any time a signal can contain a set of overlapping oscillations. Each process, either linear or nonlinear, stationary or nonstationary, is an oscillation that is to a certain degree “symmetrical” relative to a local mean value and, consequently, has extremums and zero intersections. Such oscillation processes can by presented using intrinsic mode functions (IMF) given that:
— the numbers of extremums and zero intersections differ by 1 at the maximum;
— the average value between the mode function envelopes governed by the local maximums and local minimums is zero for all points.

The modes represent oscillation processes; at the same time, unlike harmonics resultant from Fourier transform of a discrete signal, each IMF can have variable amplitude and frequency as functions of time. Any arbitrary signal can be divided into a set of mutually orthogonal intrinsic mode functions which are not set analytically and are exclusively determined by the analyzed sequence.

2. EMD algorithm

The brief description of EMD below seems to be sufficient for understanding the algorithm of EMD. To get a more profound insight into the method and its peculiarities, it is possible to address the works of the method’s author—Huang [3].

In a general case, the algorithm is composed of sequential operations aimed at identifying mode functions, starting with high frequencies, in a signal. Each IMF identification step includes:
1. Detection of all local extremums of the signal and formation of amplitude-coordinate vectors for maximums and minimums;
2. Identification of upper and lower envelopes of the signal using the found local maximums and minimums by means of “natural” cubic spline. Determination of average values function between the maximum and minimum envelopes—$m_i$;
3. Finding IMF function in the first approximation using the formula: $h_1(k) = y(k) - m_1(k)$;
4. Repeating steps 1–3 and replacement of the initial signal by $h_1$ allows the second approximation for IMF function: $h_2(k) = h_1(k) - m_2(k)$;
5. Similarly, the third and next approximations of IMF are found. As the number of iterations is being increased, the functions $h_i$ and $m_i$ tend to their invariable forms. The cessation criterion is chosen as a limit of normalized square difference between two subsequent iterations, or is set as maximum iteration number (as a rule, 7–109, which conforms with the limit of $10^{-4}–10^{-5}$);
6. The value of $h$ obtained during the last iteration is assumed the first, highest frequency mode function $(I)$ $c_1(k) = h$ contained in the initial signal $y$.

When the obtained mode is withdrawn from the signal, low-frequency components are left: $r_1(k) = y(k) - c_1(k)$

The processing of the resultant function using the same procedure produces the second mode function $(I)$ - $c_2$ and so on:
$$r_2(k) = r_1(k) - c_2(k)$$

Finally, the signal is decomposed in the $n$-dimensional mode empirical approximation with the resultant excess $r_{n+1}$:
$$y(k) = \sum_{n} c_n(k) + r_{n+1}(k)$$

The decomposition may be continued until the maximum “straightening” of the excess so that only the trend of the signal is left with no more than 3 extremums out of which it is impossible to derive another IMF.
As a result, the initial signal is decomposed with respect to an adaptive basis obtained from the analyzed data. Although undefined analytically, it satisfies all requirements of a basis. This means that it is complete, convergent, orthogonal and unique. Huang’s statement on the decomposition uniqueness seems disputable as the empirical process of decomposition, due to its adaptivity, is uncontrolled in a general case. Even a monotone local component of a signal, under a certain influence of destabilizing factors (noise, impulse interference, etc.), can split into two or three IMF during decomposition. Naturally, the said local component can be identified in summing-up of these functions but this requires that user has some a priori knowledge on the signal composition.

3. Application of EMD method to deformograph data

The practice of introduction of EMD method in processing of deformation measurement data points at the high application potential of the method, including identification of slow deformation changes (plus a monotone trend), detection of nonstationary disturbances in signals, regularization of data and withdrawal of any nature high-frequency noise from signal.

The main difficulty in the application of EMD in processing of deformograph data is the breaks from minutes to days long in the signal. Such breaks worsen the method efficiency, especially when plotting envelopes using cubic splines. When finding first (high-frequency) mode functions, local extremums in the zone of breaks are far from each other as compared with the neighbor extremums; for this reason, under interpolation of envelopes by spline, considerable false bursts can appear in the mode functions, which are then transferred to the excess for the further decomposition. This effect accumulates during later steps, which completely impairs the result. In order to avoid the situation, it was decided to add the signal with the Gaussian noise at a level not higher than 0.01–0.05 of standard scatter of initial signal.

Addition of noise in the initial signal, aside from the problem of breaks in the signal, enables overcoming another rather frequent problem connected with the division of one or a few of closely spaced harmonics into different neighbor mode functions. This situation often takes place in processing of actual data and is uncontrollable in a general case owing to adaptivity of the decomposition basis. The problem can be solved by operating pairs of modes rather than separate modes. On the other hand, there already exists the method of EEMD (Ensemble Empirical Mode Decomposition) [5] which is the multiple variation of the initial signal by means of adding the Gaussian noise, finding of mode functions in each case and their averaging over an ensemble. Depending on the number of variations, the level of noise is varied, and as the ensemble of variations grows, the level of noise can be increased as in the averaging the value of the resultant noise is inversely proportional to the ensemble size root. Thus, when averaging the ensemble of 100 variations, the signal can be added with the Gaussian noise at the level of 25% of standard scatter of the initial signal. Varying the level of the added noise and the size of the ensemble makes it possible to solve the problem connected with the mix of the mode functions. The payment for this is the considerable computer time of the signal processing.

The result of detection of slow deformation component using the described method is depicted in Figure 1. The upper curve in the figure shows the unfiltered deformation signal 85 days long recorded by a measurement arm of the deformograph installed in the Talaya seismic station gallery (Baikal Region). The curve in the middle is the result of the low-frequency Fourier filter operation at a frequency of 0.6 1/day. The slow deformation component is detected sufficiently well only in the center of the curve. At the edges there are so-called “marginal” effects 10–12 days long. The empirical decomposition method allows the total trend and the slow component detection along the whole signal, which is illustrated by the lower curve. The curve is the result of summation of the three last mode functions and the decomposition excess shown in Figure 2.

After the analysis of the modes $\text{IMF}_{11} - \text{IMF}_{13}$ and the excess $R$, it has been concluded that except for the monotone drift (compression of the deformograph arm), there is a gradual oscillation of low frequencies in the signal. The cause of such behavior requires additional analysis, which is beyond the
scope of this paper; on the other hand, detectability of such behavior of deformation deserves attention.

Figure 1. Comparison of the results of slow deformation component detection by the empirical mode decomposition method and digital filtration.

Figure 2. Monotone trend and low-frequency modes in deformation signal by EMD method.

Figure 3. Results of erasing of monotone trend and low-frequency deformation component by EMD method.

The result of removal of monotone trend from the signal is shown in Figure 3. Aside from the straightened plot with the observable tidal oscillations, the figure demonstrates the initial signal. The absence of the “marginal” effect of EMD detection of the monotone part is seen.
Aimed to prove the efficiency of EMD method and the absence of “marginal” effects, Figure 4 shows the deformation signal and the identified slowly changing part and the similarly calculated slow part of signal from the expanded data for 7 days toward each side. Complete coincidence of the signals is evident in the figure.

![Figure 4. Absence of marginal effects in EMD method application.](image)

The key potential of EMD in processing of deformograph data lies in the method applicability to studying nonstationary processes in the range of ultra long period oscillations of earth. Inside this range, in the course of continuous recording of deformation in the Earth’s crust in order to trace precursors of earthquakes, a repeating effect was identified. In the deformation curves plotted for each recording interferometric arms, separately, an unusual disturbance was observed against the background of tidal oscillations 1.5–2 days before an earthquake. This effect is described in more detail in [6–8]. The lithospheric disturbances were clearly recorded as trains of deformation noise in the band of periods of oscillations 0.5–2 hours long; however, due to impermanence of the disturbances, it is very difficult to analyze their structure using classical methods.

![Figure 5. Excitation of ultra long period oscillations on the eve of regional earthquakes: A—earthquake on 3 Aug 2006, 11:58 a.m., φ = 55.40, λ = 110.41, K = 12.0; B—earthquake on 3 Aug 2006 16:57 p.m., φ = 55.41, λ = 110.40, K = 12.2.](image)
An illustration of EMD application to a section of deformation plot 5 days long, containing the described above effect is given in Figure 5. At the top of the figure, there is the unfiltered deformation plot with clearly seen deformation excitations against the background of half-day tide, which have comparable amplitudes with the tides. The second curve in the upper part of Figure 5 is the superposition of four modes in which during decompositions the mentioned disturbances were detected; these modes separately are presented at the bottom of the figure. It is possible to characterize the onset and duration of oscillations (grey-colored rectangles in the figure), and the amplitude–frequency dynamics of the event; also, it is possible to identify the most representative range. It is seen that the effect appears at different time in different modes and has the largest amplitude in mode 8, which conforms with the range of periods of 1–2 hours for this decomposition. Inside this mode, there are traceable harmonics with the periods of 62, 73 and 85 min. The effect is present in the higher frequency modes (e.g. 5 and 6) but is less expressed and has much smaller amplitude. The application of EMD method has allowed identification of the wanted signal in the test section and the analysis of its behavior in different frequency ranges (modes), which later on can offer additional information required for the interpretation of the event and for the determination of its prediction potential.

Finally, the result of EMD application for the removal of high frequencies, including different nature noise, from the signal is depicted in Figure 6. To this effect, first three modes \( \text{IMF}_1 \)–\( \text{IMF}_3 \) were withdrawn from the initial signal. The method is efficient and yields to none of the most widespread rivals, is free from the marginal effect and is good in processing of single bursts. In particular, EMD has proved the suggestion [9] that the daily variation of amplitude of the high-frequency component in the deformation signal is not of the noise nature as it is totally absent in the first identified mode.

![Filtration of noise from the signal by EMD.](image)

**Figure 6.** Filtration of noise from the signal by EMD.

### 4. Conclusion

Despite being comparatively new, the method of empirical mode decomposition has already some modifications and is widely applied in various research areas connected with nonlinear and nonstationary data. The authors believe that the method presents a promising tool for operation with geophysical data and deformographs, in particular, to detect local features of signals and their internal structure, to remove noise, to carry out regularization and to identify slow part in a signal, including a monotone trend. It is highly expected that the method can be efficient in studying excitation of ultra long period oscillations before earthquakes as this phenomenon is of short duration and is nonlinear both in terms of amplitude and frequency.

### Acknowledgments

The studies under the Section on Application of EMD Method to Deformograph Data of this paper have been supported by the Russian Science Foundation, Project No. 16-17-10121.
References

[1] Panov SV, Parushkin MD, Semibalmut VM and Fomin YuN 2016 Heterodyne laser deformograph for geophysical research *Current Challenges of Modern Seismology* pp. 293–299 (in Russian)

[2] Orlov VA, Panov SV, Parushkin MD, Semibalmut VM and Fomin YuN 2013 Capabilities of highly sensitive laser measurements in studying seismic and geodynamic events *Pros. Int. Conf. Geodynamics and Stress State of the Earth’s Interior* Novosibirsk: IGD SO RAN pp 25–31 (in Russian)

[3] Huang NE and Shen Z *The Hilbert–Huang Transform and Its Applications* World Scientific Publishing Co. Pte. Ltd. Singapore 596224

[4] Huang NE, Shen Z, Long SR, Wy MC, Shih HH, Zheng Q, Yan N-C, Tung CC and Liu HH 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis *Proc. R. Soc. London Ser. A* 454 pp 903–995

[5] Wu Z and Huang NE 2009 Ensemble empirical mode decomposition: A noise-assisted data analysis method *Advances in Adaptive Data Analysis* Vol 1 No 1 pp 1–41

[6] Orlov VA, Panov SV, Parushkin MD and Fomin YuN 2008 Connection of the Earth’s seismicity and solar activity by the precision deformograph data *Pros. Int. Conf. Geodynamics and Stress State of the Earth’s Interior* Novosibirsk: IGD SO RAN pp 31–40

[7] Orlov VA, Panov SV, Parushkin MD and Fomin YuN 2008 Solar activity, oscillations of the inner core of earth and seismicity of the planet *Proc. IV Int. Symp. Geodynamics of Intracontinental Orogenes and Geoecology* Moscow pp 321–326 (in Russian)

[8] Orlov VA, Panov SV, Parushkin MD, Fomin YuN, Tokmoldin SZh, Nevmershitsky IS, Uzbekov MB and Tereshkin DO 2011 Spaced highly sensitive deformation observations in studying ultra long period oscillations of earth *Pros. Int. Conf. Geodynamics and Stress State of the Earth’s Interior* Novosibirsk: IGD SO RAN pp 145–150 (in Russian)