Data analysis of sunspot time series with SSA and HHT information adaptive methods

Nikolai T. Safiullin[1], Sergey V. Porshnev[1], and Nathan I. Kleeorin[2]

[1] Ural Federal University, Yekaterinburg, Russia
n.t.safiullin@urfu.ru, sergey_porshnev@mail.ru,
[2] Ben-Gurion University of Negev, Beer-Sheva, Israel
nat@bgu.ac.il

Abstract. In the paper, characteristics of the monthly Wolf numbers time series data, containing information about individual solar cycles, are investigated by decomposition into components with two time series methods, namely, the Singular Spectrum Analysis (SSA) and Huang-Hilbert Transform (HHT). These methods do not require any a priori knowledge about analyzed data, making them information adaptive. As a result, some of the known cycles such as Schwabe-Wolf Cycle, Hale Cycle, Gleisberg Cycle, and Suess Cycle have been identified. These components and their properties are compared with each other, as well as with known characteristics of sunspot cycles.

Keywords: data analysis, information handling, time series, empirical mode decomposition, singular spectrum analysis, sunspot numbers

1 Introduction

Solar activity defines strength of solar flares, coronal mass ejections, amount of energetic accelerated particles, and power of solar electromagnetic field. Moreover, it affects valuable low orbit satellites and other sensitive instruments in space [1], as well as terrestrial climate [2] and people's health to some extent [3]. To measure value of the solar activity, the Wolf sunspot numbers R (relative Zurich sunspot numbers) are often used

$$R = k \cdot (10g + s)\,, \tag{1}$$

where s is the number of isolated sunspots, g is the number of sunspot group, k is the observatory factor taking into account various observation conditions. At present, all the works of estimation and spreading the Wolf numbers from 1749 year are stored by the Royal Observatory of Belgium in Brussels [4]. Represented time series with averaged monthly Wolf sunspot numbers from January 1749 till January 2017 are shown in Fig. 1.

It is important to note that the main difficulty with analysis of these time series is caused by its non-stationary characteristic [3]. This fact significantly reduces number of methods for analysis and forecasting the mentioned time series.

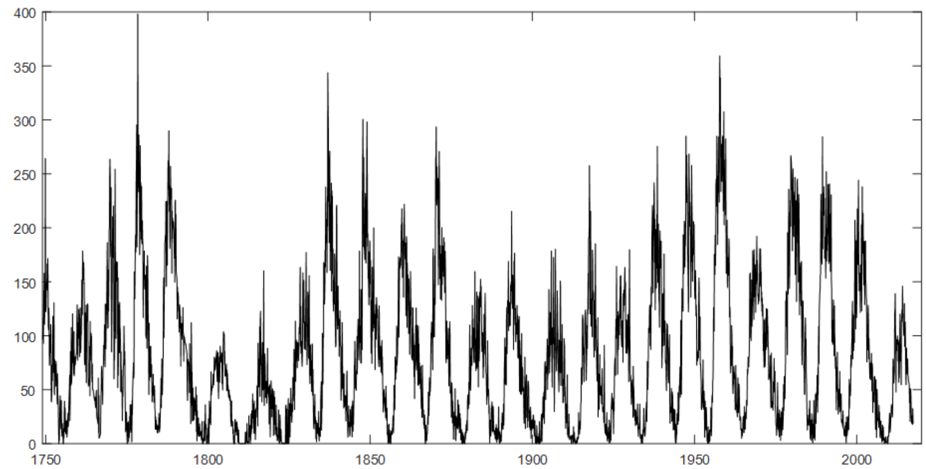

Fig. 1. Averaged monthly Wolf (sunspot) numbers between Jan. 1749 and Jan. 2017

Majority of time series analysis methods (smoothing and regression methods, autoregressive models, etc.) are based on the *a priori* assumption of weak stationarity [6]. So, it would be more appropriate to apply to Wolf numbers the time series analysis and forecasting methods with no *a priori* assumptions (they are called the adaptive methods), such as Singular Spectrum Analysis [7] and novel Huang-Hilbert Transform [8, 9]. These methods decompose the initial data into components

$$u\left(t_j\right) = \sum_{i=1}^{n-1} c_i\left(t_j\right) + r_n\left(t_j\right), \tag{2}$$

where $u\left(t_j\right)$ is the initial time series, $c_i\left(t_j\right)$ is the i-th component, $r_n\left(t_j\right)$ is the residual time series, at the j-th time point t_j. Some of these components contain meaningful (from a physical viewpoint) time-and-frequency characteristics.

In this paper, we present the results of the comparative analysis of the Wolf sunspot time series examination by two methods: Singular Spectrum Analysis (SSA) and Huang-Hilbert Transform (HHT) that are also compared with contemporary astrophysical facts of the solar activity cycles.

2 Basic SSA algorithm

Following [7], we present a brief description of the basic SSA algorithm. Consider a real-valued time series $F = (f_0, ..., f_{N-1})$ of length $N > 2$. Assume that F is a non-zero series. So, it is usually assumed that $f_i = f\left(i\Delta t\right)$ for a certain function of time $f\left(t\right)$ with some time interval Δt. Moreover, the numbers $0, ..., N-1$ can be interpreted not only as discrete time instants but, also, as labels of any other sequentially ordered structure. The Basic SSA consists of two complementary stages: decomposition and reconstruction.

2.1 First stage: decomposition

The 1st step. Embedding

The embedding procedure maps the original time series to a sequence of multidimensional lagged vectors. Let L be an integer (the window length), $1 < L < N$. The embedding procedure forms $K = N - L + 1$ lagged vectors

$$X_i = (f_{i-1}, ..., f_{i+L-2})^T, 1 \leq i \leq K,$$

which have dimension L. If we need to emphasize the dimension of the X_i, then we shall call them the L-lagged vectors. The L-trajectory matrix (or simply trajectory matrix) of the series F is

$$X = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} f_0 & f_1 & f_2 & \cdots & f_{K-1} \\ f_1 & f_2 & f_3 & \cdots & f_K \\ f_2 & f_3 & f_4 & \cdots & f_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & f_{L+1} & \cdots & f_{N-1} \end{pmatrix}. \tag{3}$$

Obviously, $x_{ij} = f_{i+j-2}$ and the matrix X has equal elements on the diagonals $i+j = const$. Thus, the trajectory matrix is a Hankel matrix. Certainly, if N and L are fixed, then there is a one-to-one correspondence between the trajectory matrices and the time series.

The 2nd step. Singular value decomposition

The result of this step is the Singular Value Decomposition (SVD) of the trajectory matrix (3). Let $S = XX^T$. Denote by $\lambda_1, ..., \lambda_L$ the eigen-values of S taken in the decreasing order of their magnitudes ($\lambda_1 \geq ...\lambda_L \geq 0$) and by $U_1, ..., U_L$ the orthonormal system of the eigen-vectors of the matrix S corresponding to these eigen-values.

Let $d = \max\{i : \lambda_i > 0\}$. If we denote $V_i = X^T U_i / \sqrt{\lambda_i}, i = 1, ..., d$, then the SVD of the trajectory matrix X can be written as

$$X = X_1 + ... + X_d, \tag{4}$$

where $X_i = \sqrt{\lambda_i} U_i V_i^T$. The matrices X_i have rank 1; therefore, they are elementary matrices. The collection $\left(\sqrt{\lambda_i}, U_i, V_i\right)$ will be called the i-th eigen-triple of the SVD (4).

2.2 Second stage: reconstruction

The 3rd step. Grouping

Once the expansion (4) has been obtained, the grouping procedure breaks down the set of indices $\{1, ..., d\}$ into m disjoint subsets $I_1, ..., I_m$.

Let $I = \{i_1, ..., i_p\}$. Then the resultant matrix X_I corresponding to the group I is defined as $X_I = X_{i_1} + ... + X_{i_p}$. These matrices are computed for $I = I_1, ..., I_m$, and expansion (4) leads to a new decomposition. The procedure of choosing the sets $I_1, ..., I_m$ is called the eigen-triple grouping.

The 4th step. Diagonal averaging
The last step in the Basic SSA transforms each matrix of the grouped decomposition $X_I = X_{I_1} + ... + X_{I_m}$ into a new series of length N.

Let Y be an $L \times K$ matrix with elements y_{ij}, $1 \le i \le L$, $1 \le j \le K$. We set $L^* = \min(L, K)$, $K^* = \max(L, K)$ and $N = L + K - 1$. Let $y^*_{ij} = y_{ij}$ if $L \le K$ and $y^*_{ij} = y_{ji}$ otherwise. The diagonal averaging transfers the matrix Y to the series $(g_0, ..., g_{N-1})$ by the formula

$$g_k = \begin{cases} \frac{1}{k+1} \sum\limits_{m=1}^{k+1} y^*_{m,k-m+2}, & \text{for } 0 \le k < L^* - 1, \\ \frac{1}{L^*} \sum\limits_{m=1}^{L^*} y^*_{m,k-m+2}, & \text{for } L - 1^* \le k < K^*, \\ \frac{1}{N-k} \sum\limits_{m=k-K^*+2}^{N-K^*+1,} y^*_{m,k-m+2} & \text{for } K^* \le k < N. \end{cases} \tag{5}$$

Expression (5) corresponds to averaging of matrix elements over the diagonals $i + j = k + 2$: the choice $k = 0$ gives $g_0 = y_{11}$, for $k = 1$ we have $g_1 = (y_{12} + y_{21})/2$, and so on. Note that if the matrix Y is the trajectory matrix of some series $(h_0, ..., h_{N-1})$ (in other words, if Y is the Hankel matrix), then $g_i = h_i$ for all i.

The diagonal averaging (5) applied to a resultant matrix X_{I_k} produces the series $\tilde{F}^{(k)} = \left(\tilde{f}^{(k)}_0, ..., \tilde{f}^{(k)}_{N-1}\right)$, and, therefore, the initial time series $(f_0, ..., f_{N-1})$ is decomposed into the sum of m series like in the additive model (2).

3 Huang-Hilbert transform and its modifications

3.1 The first stage: empirical mode decomposition

Using the Huang-Hilbert transform, components $c_i(t)$ in (2) are obtained by the ensemble empirical mode decomposition (EEMD) [9]. This procedure is a modification of the basic empirical mode decomposition (EMD) [8], which diagram is illustrated in Fig. 2.

The EMD main idea is based on the assumption that any data consist of different simple intrinsic modes of oscillation. These modes can be empirically estimated with enveloping curves (upper curve U and lower curve L) constructed by cubic splines passing through the maximum and minimum points of the time series. The arithmetical mean m of those curves U and L is removed from the analysed data. So, the received residual r is subjected to this enveloping procedure again until calculated residual r fits with a certain criteria and then it is noted as the found intrinsic component c_i. By removing the found component c_i from the initial time series $u_0(t)$, we obtain a new time series $u(t)$ to decompose iteratively based on definition (2). This decomposition is repeated until the residual $r_n(t)$ becomes a monotonic function.

The EEMD procedure [9] is a more complex version of EMD; a diagram for EEMD is presented in Fig. 3. In contrast to EMD, overall decomposition (2) is

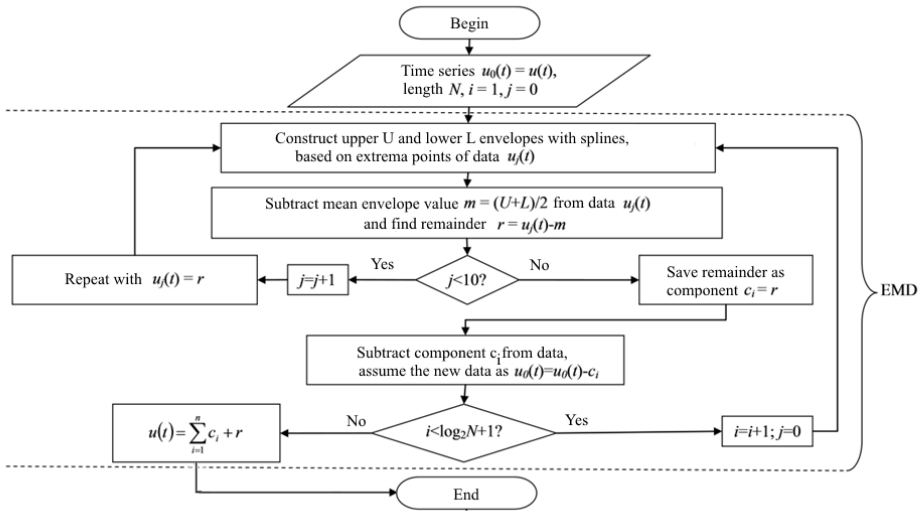

Fig. 2. Empirical Mode Decomposition (EMD)

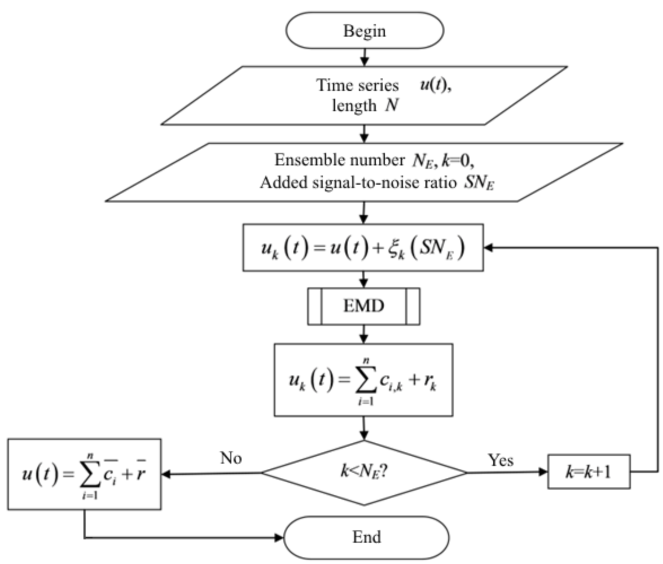

Fig. 3. Ensemble Empirical Mode Decomposition (EEMD)

applied to the initial time series $u_0(t)$ not only once, but several N_E (Ensemble Number) times. Also, in order to achieve the initial discrepancy of time series for the subsequent ensemble averaging procedure, the white Gaussian noise (wGn) with known signal-to-noise ratio SN_E is added to the initial data $u_0(t)$.

3.2 The second stage: time-and-frequency characteristics

The decomposition is only the first step in the HHT method (and to some extent in SSA as well). Time-and-frequency characteristics of received components that were found meaningful from a physical viewpoint are extremely significant in investigation of sunspot time series; for example, for comparison of the results with known aspects of the sunspot cycles (especially, periods) and in further analysis of non-stationary time series and forecasting of the Wolf numbers.

On the second stage of the HHT method, a concept of analytical signal is used to calculate time-and-frequency characteristics of data. This idea was introduced in 1946 by Gabor [5]. Analytical signal allows one to introduce simple and unambiguous definitions for amplitude $a(t)$, phase $\varphi(t)$, and instantaneous frequency $\omega(t)$ of signal $u(t)$.

Based on [5], the analytical signal $w(t)$ is defined as

$$w(t) = u(t) + jv(t) = a(t)\, e^{j\varphi(t)}, \tag{6}$$

where $v(t)$ is the signal conjugated to the initial data $u(t)$ with Hilbert transform

$$v(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{u(\tau)}{t - \tau} d\tau. \tag{7}$$

Definition (6) allows one to estimate naturally the amplitude $a(t)$ for the signal $u(t)$

$$a(t) = |w(t)| = \sqrt{u^2(t) + v^2(t)}, \tag{8}$$

phase $\varphi(t)$

$$\varphi(t) = \arctan\left(\frac{v(t)}{u(t)}\right) = \arccos\left(\frac{u(t)}{a(t)}\right) = \arcsin\left(\frac{v(t)}{a(t)}\right), \tag{9}$$

and instantaneous frequency $\omega(t)$

$$\omega(t) = \dot{\varphi}(t) = \frac{u(t)\,\dot{v}(t) - \dot{u}(t)\,v(t)}{a^2(t)}. \tag{10}$$

The same method (6 – 10) can be used to calculate time-and-frequency characteristics of components received by the SSA decomposition and reconstruction. Note that in the case of the monthly sunspot numbers (1), we are more interested not in the instantaneous frequency $\omega(t)$, but in the function of time for the instantaneous period $T(t) = \omega^{-1}(t)$.

4 Analysis of results for the monthly Wolf sunspot numbers

With decomposition of the initial Wolf sunspot numbers data (Fig. 1) by two methods (SSA and EEMD), we received five significant (from physical point of view) components (Fig. 4).

In the SSA method, we used the dimension window $L = 1592$ in accordance with recommendations from [7]. Grouping was made with eigen-triples 1, 2–5, 6–7, 16–17, 18–19. Reconstructed data justify 84.09% of dispersion from the initial time series. In the HHT method, we used $N_E = 100$, $SN_E = 11$dB in accordance with technique for the Huang-Hilbert transform [9].

From Figure 4 it is seen that

1. it is actually possible to decompose non-stationary Wolf sunspot numbers' time series into physically meaningful components by the SSA and HHT methods without any *a priori* information;
2. corresponding components found by two different adaptive methods are visually close to each other;
3. with increasing order number of components, their frequency domain is decreasing;
4. the fifth component c_5 represents trend for the monthly sunspot numbers;
5. all found components are amplitude-modulated and frequency-modulated signals; so, we need to calculate their time-and-frequency characteristics for their detailed analysis;

Thus, for the received components, we calculated time-and-frequency characteristics, such as instantaneous period $T(t) = \omega^{-1}(t)$ by equations $(6-10)$. Obtained results of instantaneous period were subjected to a smoothing at five points to reduce noise effects; final results are summarized in Table 1. The well-known cycles are also shown in this table. These corresponding cycles were taken from contemporary astrophysical facts of solar activity cycles.

Table 1. The expectation value (E) and variance of instantaneous periods for components received by the SSA and HHT methods.

Method	SSA		HHT		
Cycle	E	Variance	E	Variance	Designation
1	11.10	2.69	11.32	3.13	Schwabe-Wolf Cycle
2	15.89	25.11	17.07	22.79	Hale Cycle
3	39.56	316.05	40.84	153.50	Half-Century Cycle
4	95.22	2182.72	94.09	306.67	Gleisberg Cycle
5	807.52	7014.15	181.58	2217.33	Suess Cycle

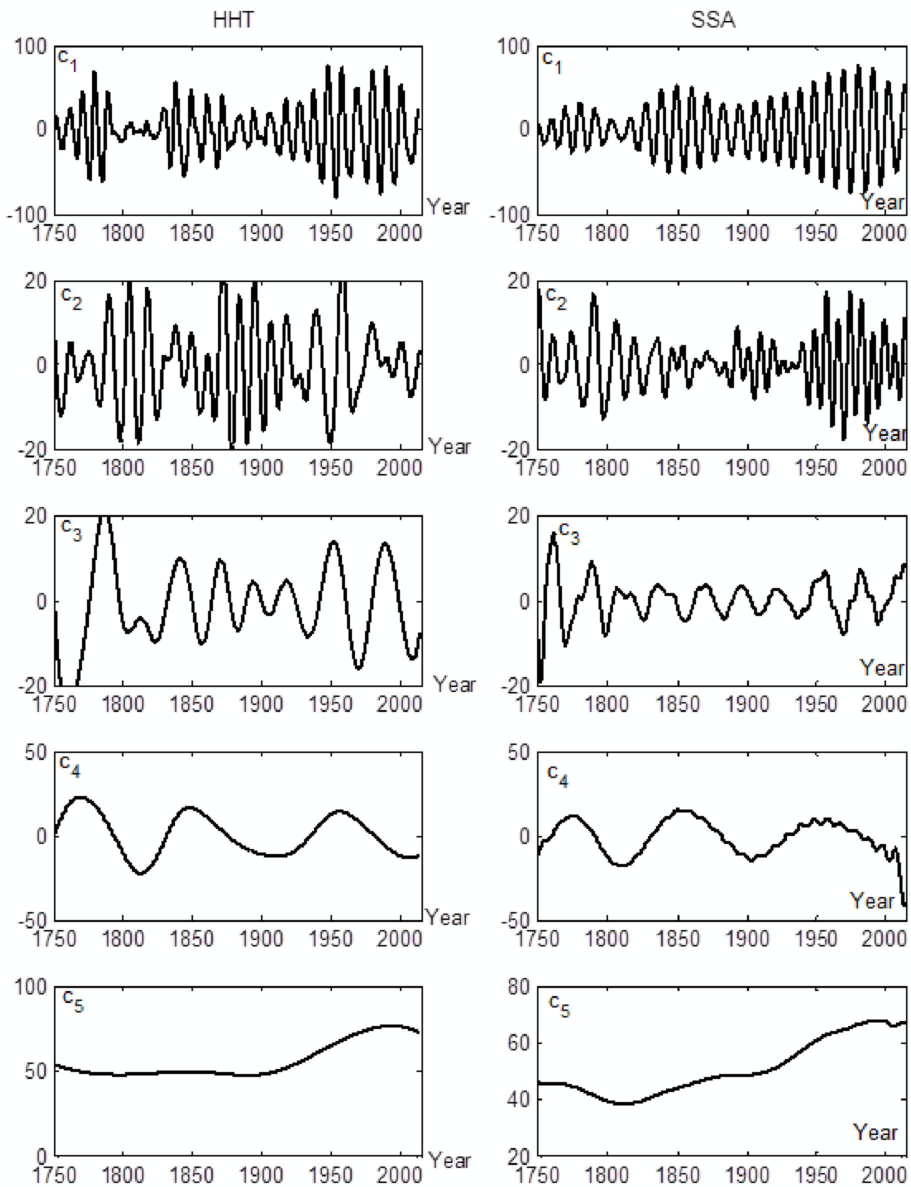

Fig. 4. Components received by decomposition with the HHT (left) and SSA (right) methods

From Fig. 4 and Table 1 some facts can be noted.

1. Instantaneous periods are more distinguished between two methods, because equation (10) uses differentiation procedure.
2. Nevertheless, the mean value and variance (to some extent) of periods for the first four components are quite close to each other. Last component c_5 is a trend; thus, its time-and-frequency characteristics are unreliable due to insufficient time interval of time series analysis.
3. The first component actually represents the 11-year Schwabe-Wolf Cycle and shows a tendency to float between 7 and 15 years. This fact is well-known in the solar activity astrophysics [1–3].
4. The second component corresponds to the 22-year Hale Cycle. This is the most inaccurate component received by those two methods. This fact can be explained by multiplicity of this cycle with 11-year cycle, which leads to inadequate decomposition by adaptive methods that are evidently weak in the case of multiple frequencies.
5. The third component represents the known Half-Century Cycle, which is noted not to be valuable [3], but often found by other methods of time series analysis [6].
6. The fourth component corresponds to the century Gleisberg Cycle and shows a tendency to float between 80 and 120 years (in the HHT method).
7. The latter component represents a trend that can be compared with the known Suess Cycle (period is more than 200 years). Facts about this cycle are not estimated from the solar activity, but from total irradiance research of the terrestrial radio-isotopes [3], hence, it is difficult to compare the results.
8. Distortion in instantaneous period of components c_1, c_3, c_5 next to 1790–1820 years corresponds to the known Dalton Minimum [3].

5 Conclusion

In the paper, comparative analysis was performed for components received by decomposition of Wolf sunspot numbers with the SSA and HHT information adaptive methods. It is possible to make the following conclusions.

1. Both methods (SSA and HHT) can be successfully used in non-stationary time series analysis to decompose the initial data into components with meaningful (from physical viewpoint) properties.
2. Instantaneous periods received by the Hilbert transform after decomposition can be compared with the well-known facts about time-and-frequency characteristics of the solar activity.
3. Distinctions between components calculated by the SSA and HHT methods are connected with different approaches in their basis: the SSA has a more theoretical approach with orthogonal dimensions in its core, while the HHT is clearly the empirical method based on extrema points.

We expect that this information can be helpful in further time series data analysis techniques and, especially, in the forecast of solar activity by the contemporary informational technologies.

References

1. Pulkkinen, T.: Space Weather. Living Rev. Solar Phys. 4(1), (2007)
2. Haigh, J.: The sun and the earths climate. Living Rev. Solar Phys. 4(2), (2007)
3. Hathaway, D.: The solar cycle. Living Rev. Solar Phys. 7(1), (2010)
4. SILSO World Data Center: The international sunspot number. http://sidc.be/silso/datafiles
5. Gabor, D.: Theory of communication. Jour. Inst. Elect. Eng. 93, 429–457 (1946)
6. Box, G. & Jenkins, G.: Time Series Analysis: Forecasting and Control. Prentice Hall, 3rd edition (1994)
7. Golyandina, N.: Analysis of Time Series Structure: SSA and Related Techniques. Boca Raton: Chapman & Hall/CRC (2001)
8. Huang, N. & Shen, Z., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. SOC. London, Ser. A. 454, 903–995 (1998)
9. Huang, N.: Ensemble Empirical Mode Decomposition: a noise assisted data analysis method. Adv. in Adaptive Data Analysis. 1(1), 1–41 (2008)