Comparison of abilities of two trend definition techniques for experimental data time series processing

I O Ivanov¹, V V Borog¹, A V Kryanev¹, V G Getmanov² and R V Sidorov²

¹National Research Nuclear University MEPhI, Kashirskoe highway 31, 115409, Moscow, Russia
²Geophysical Center of the Russian Academy of Sciences, Molodezhnaya St. 3, 119296, Moscow, Russia
E-mail: varbright@gmail.com

Abstract. The abilities of two processing techniques for time series of physical data are studied. One of this techniques is based on the implementing of singular spectrum analysis, while the other is based on local approximation approach. The operation of these techniques is analyzed on a model time series. Both methods are compared in terms of accuracy of trend definition for different values of the noise level.

1. Introduction
The problem of trend separation is very important for experimental data time series analysis. There are many physical processes time series for which consist of some slow trend and high-frequency component superimposed on the first one. Since usually only one of them is of interest the problem of trend separation arises. In this paper the two methods are considered: the method of trend separation based on singular spectrum analysis(SSA)[1, 2] and the second one based on local sliding approximation piecewise linear models with weighted average(SLAM)[3, 4]. These methods were tested on the specially generated model time series. To test the methods on different time scales we constructed time series from three consecutive fragments of periodical functions: decaying, amplitude-modulated and long-period ones. After this the model signal was mixed with white noise. Since a noise can significantly affect result of trend definition, we need to check the dependence of quality of trend separation on its level. To achieve this we varied the value of standard deviation of the noise ($\sigma$) and run describing technics on the obtained time series. In the Fig.1 the model time series with noise level $\sigma = 0.01$ is shown.

2. Singular spectrum analysis
Consider experimental data $y_1, \ldots, y_n$. Construct the trajectory matrix of this time series:

$$Y = \begin{pmatrix} y_1 & y_2 & \cdots & y_M \\ y_2 & y_3 & \cdots & y_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_{L+M-1} \end{pmatrix}$$
Figure 1. Model time series with noise level $\sigma = 0.01$.

$M$ stands for the sliding window size, it is the parameter of this method and should be selected based on physical properties of the problem. Perform the singular value decomposition of trajectory matrix. Eigenvalues of matrix $Y^T Y$ are called singular values of matrix $Y$. Sort singular values in decreasing order of magnitude and denote $\psi_j = (\psi_{j1}, \ldots, \psi_{jM})^T, j = 1, \ldots, M$ is the corresponding orthonormal system of eigenvectors. Construct orthogonal matrix $\Omega = (\psi_1, \ldots, \psi_M)$ and define the matrix of principal components: $V = Y \Omega = (V_1, \ldots, V_M)$.

So,

$$Y = V \Omega^T = \sum_{j=1}^{M} (V_j \psi_j^T)$$

is the decomposition of the original time series into principal components $V_j, j = 1, \ldots, M$.

The magnitude of singular values determines the level of "information contribution" of corresponding principal component to the source time series. Components with larger indices corresponds to more frequent fluctuations of the signal. Therefore, for trend separation we need to exclude the last few principal components, corresponding to high-frequency part of time series (noise). After this we get the next expression for time series decomposition.

$$Y' = \sum_{j=1}^{m} (V_j \psi_j^T)$$

where $m$ is the number of the remaining components.

The resulting time series is reconstructed using one-to-one correspondence between trajectory matrix and time series. Applying diagonal averaging to result matrix $Y'$ we get reconstructed series:

$$y'_1 = y'_{11}; y'_2 = (y'_{12} + y'_{21})/2; y'_3 = (y'_{13} + y'_{22} + y'_{31})/3$$

and so on.

Selecting of value for sliding window size is very important here, because it determines number of principal components i.e. "information resolution" of the method. If it is too small, the result trend will be mixed with noise. And vice versa, if the window size is too big, the real trend will
be divided into several components. Because of these reasons, choosing of window size value is very difficult in case of a real problems due to lack of information about the real trend.

3. Method of local sliding approximation models

The second method of trend removal is based on using so-called local sliding approximation models with weighted averaging, and special correlation analysis for provision of consistency. Suppose that observations of original time series \( y(T_i), i = 0, 1, \ldots, N_f - 1 \) are given. Construct the local sliding intervals:

\[
N_{1j} = N_d(j - 1); \quad N_{2j} = N_{1j} + N - 1
\]

where \( N_d \) is a sliding step, \( N \) is the size of local interval and \( j = 1, \ldots, m; N_d(m - 1) + N = N_f. \)

Let the local sliding models will be represented by the functions \( y_{Mj}(c_j, T_i) \); outside the local interval the models equal to zero. In a particular case local models can be piecewise linear - \( y_{Mj}(c_j, T_i) = c_{1j} + c_{2j} T_i \). Let us take weighted sum of local models as a model of trend.

\[
y_{M}(c, T_i) = \sum_{i=1}^{m} \alpha_j y_{Mj}(c_j, T_i)
\]

\( \alpha_j \) - weights, \( c^T = (c_1^T, \ldots, c_m^T) \) - parameters of model. The optimal model of trend \( y_{M}(c^0, T_i) \) is a result of minimization of functional \( S(c, y) \)

\[
S(c, y) = \sum_{i=0}^{N_f-1} (y(T_i) - y_{M}(c, T_i))^2, \quad c^0 = \arg \min_{c} S(c, y)
\]

The model \( y_{M}(c^0, T_i) \) and difference \( \Delta y(T_i) = y(T_i) - y_{M}(c^0, T_i) \) depend on parameters \( N_d, N \). This ones are being adjusted so that time series \( \Delta y(T_i), (i = 0, 1, 2, \ldots, N_f - 1) \) would be the most statistically similar to white noise. For this purpose we define local sliding autocorrelation functions and calculate ratios of their values for big and small time lags. Then we adjust parameters \( N_d, N \) to minimize the sum of calculated ratios.

4. Trend separation

Let us consider the work of described methods on model series with \( \sigma = 0.01 \). Since we are dealing with nonstationary series, which greatly changes its behavior on different intervals, we selected value for sliding window size as 32. The plot of singular values for model series is shown in Fig.2. As it is evident from the graph, the first singular value is a several tens of times larger than the others, so we can conclude that the first principal component corresponds to the trend. Result of trend definition by SSA method is shown in Fig.3.

Let us consider the work of SLAM method with local interval size equal to 64 and sliding step equal to 4. The result of trend definition using method SLAM is shown in Fig.4.

As you can see, the result is almost the same as in the case of SSA method. Since we use artificial series, and know the real trend, the quality of considered methods can be compared using simple formula based on mean squared error [5]:

\[
\delta = \sum_{i=1}^{N} (y_i^0 - \tilde{y}_i)^2,
\]

where \( y_i^0 \) - source signal without noise and \( \tilde{y}_i \) - resulting trend separated from the series. These calculations were performed for several values of noise level (i.e. standard deviation).

The plot shown in Fig.5 illustrates relationship between trend definition error and noise level for both techniques. Numerical values of the error are shown in Table 1.
Figure 2. Singular values of model series with $\sigma = 0.01$.

Figure 3. Result of trend definition using SSA method.

| $\sigma$  | 0.001     | 0.005     | 0.01      | 0.015     | 0.02      | 0.03      | 0.04      | 0.05      | 0.06      |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| $\delta_{SSA}$ | 0.2233    | 0.2322    | 0.2729    | 0.2666    | 0.3827    | 0.4859    | 0.5552    | 0.7021    | 0.8654    |
| $\delta_{SLAM}$ | 0.0565    | 0.0894    | 0.1489    | 0.1920    | 0.3168    | 0.4527    | 0.5239    | 0.6819    | 0.8768    |

5. Conclusions
In this paper two methods for trend definition are considered. These methods were tested on the specially generated model series. To compare two methods the mean squared error was
Figure 4. Result of trend definition using SLAM method.

Figure 5. Trend definition error depending on noise level

calculated. As we can see from the plot in Fig.5, the method of local sliding intervals is about two times more accurate than SSA for the low noise levels. And for the high ones both methods shows almost equal accuracy.

References
[1] Golyandina N and Osipov E 2007 Journal of Statistical Planning and Inference 137 2642
[2] Kryanev A, Lukin G and Udumyan D 2010 Metric analysis and data processing (Moscow: FIZMATLIT)
[3] Getmanov V G 2013 Optoelectronics, Instrumentation and Data Processing 49 21
[4] Getmanov V G, Sidorov R V and Dabagyan R A 2015 Measurement Techniques 58 1029
[5] Weisberg S 2005 Applied Linear Regression (Hoboken, New Jersey: Wiley-Interscience)