SOFTWARE FOR DETECTING HIDDEN PERIODICITIES WITHIN POINT PROCESS AND FOR DATA MINING SCALAR TIME SERIES PROPERTIES.

Alexey A. Lyubushin

Institute of the Physics of the Earth, Russian Academy of Sciences,
123995, Moscow, Russia, Bolshaya Gruzinskaya, 10; fax: +7 (499) 766-26-54;
e-mail: lyubushin@yandex.ru
http://alexeylyubushin.narod.ru/

The instructions for the user of two programs for the data analysis are presented below. The instructions are accompanied by examples of the data analysis. The description of used algorithms and methods is submitted in the Appendix containing 7 items: **A1-A7**, on which the references in the instructions for the user will be made.

The programs are working in operational systems Microsoft Windows 98/NT/2000/Me/XP/Win7.

1. Program PPPeriod for detecting hidden periodicities within sequence of events

The program PPPeriod (**P**oint **P**rocesses **Period**icity) is written on Compaq Visual Fortran and is a console application. The dialogue with user is organized as a simple question – answer sequence.

The program is intended for computing increments of maximum logarithmic likelihood function for the periodic intensity model of the point process with respect to pure random (Poissonian) model with constant intensity. Probe values of the period are varying within given boundaries and those periods which provide essential peak of the maximum log-likelihood increments detect periodic components within flow of events. The method is described in details in the appendix **A1**.

Input file for the program could be any ASCII table. In the 1st column there must be the sequence of monotonously non-decreasing numbers which are interpreted as time moments. Any other columns of input file (if any) are ignored.

The program could work optionally in 2 regimes:
1) estimate using all available data sample;
2) estimate performed within moving time window of the given length and with given mutual windows' shift (in dimensional time units).

After the launching the program it is necessary to answer the following questions:

1. Input file name. The file must be at the same directory as the executable module of the program. If file is absent or it has a wrong structure (for instance the 1st column of the input table contains decreasing numeric values) then the program terminates with corresponding message.
2. To introduce $T_{\min}$ и $T_{\max}$ - minimum and maximum values of the considered period's range in the dimensional time units, i.e. the same units as in the input file.
3. To introduce the value N_p - a number of tested period values covering the range $[T_{\min}, T_{\max}]$ with steps uniformly distributed in logarithmic scale i.e. the following values of the periods are considered:

$$T_j = 10^{\varphi_j},\ \varphi_j = \lg(T_{\min}) + (j-1)\cdot\frac{\lg(T_{\max}) - \lg(T_{\min})}{(N_p - 1)},\quad j = 1,...,N_p$$

4. To set regime of estimating – using all sample or within moving time window.
5. If moving time window regime was set then it is necessary to answer additional questions:
   5.1. To introduce the length T_L of moving time window in dimensional time units.
   5.2. To introduce the value ΔT_L of mutual windows' shift in dimensional time units.
   5.3. To introduce the value T_{shift} of time marks shift in the output GRID-file.

If the whole sample (static) estimate regime was chosen then program creates output file named "**PPP_out.dat**" within current directory. It is a text table with 3 numerical columns with N_p rows. The 1st column of the output file consists of the sequence of tested (probe) values of periods T_j, the 2nd column presents the difference $\Delta\ln(L)$ between maximums of log-likelihood functions (formula (1.7) from **A1**), the 3rd column consists of the values of dimensionless parameter $a, 0 < a \le 1$ of intensity harmonic amplitude in the model (1.2).

If the moving time window (dynamic) estimate regime was chosen then program creates output file named "**PPP_out.grd**" within current directory. This is a symbolic GRID-file for further using in the program Surfer for plotting either 2D or 3D time-frequency diagrams of $\Delta\ln(L)$ evolution. Time marks of the output file correspond to the values:

$$T_{shift} + T_L + (k-1)\cdot\Delta T_L,\ k = 1,...,N_W,$$

The index k enumerates time windows, N_W is the general number of windows. Thus, time marks correspond to right-hand ends of moving time windows taken with mutual shift T_{shift}. The parameter T_{shift} could be useful for instance in the situation when events time moments are given as years from 1900 whereas we want to have time marks in years A.D. – then we can set T_{shift} =1900.

EXAMPLES OF APPLICATIONS. I present a number of graphics of the static estimates of $\Delta\ln(L)$ function and 2D time-frequency maps for global seismic process, M_0 is the value of minimum magnitude for analysis. The data were taken from the site http://neic.usgs.gov/neis/epic/ . Statistics $\mathbf{R(\omega,\tau\,|\,T_w)}$ is given by the formula (1.7) within appendix **A1**.

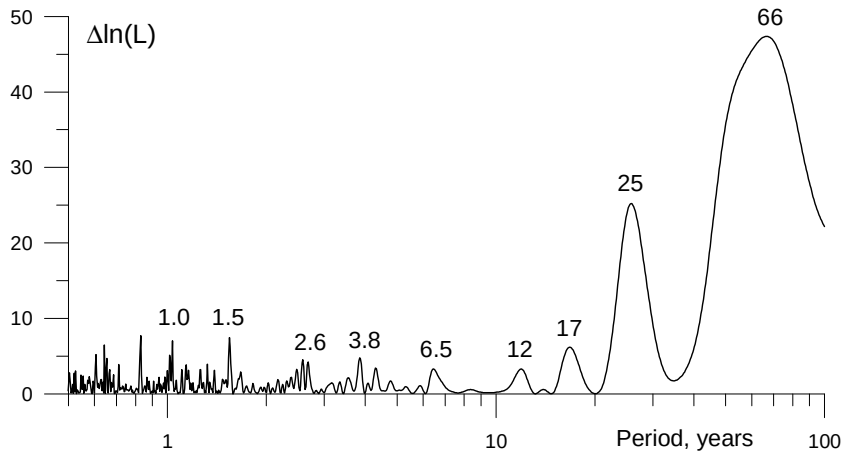

Fig.1. The static estimate for the whole world, 1901-2005, M_0=7.0, depth ≤ 100 km.

Peaks are statistically significant with 90% probability above threshold 2.3.

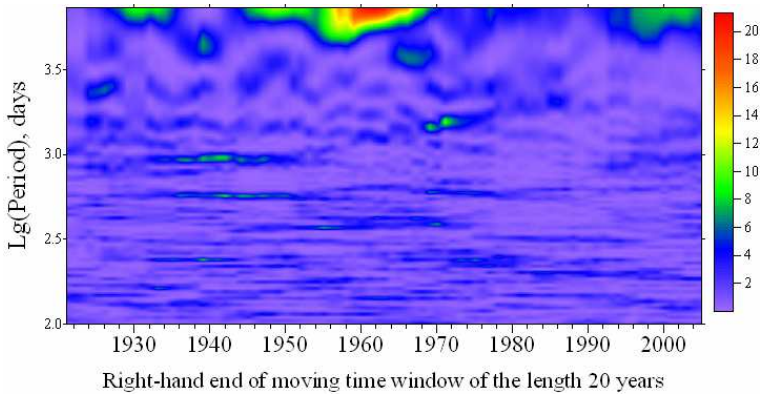

Fig.2. The time-frequency estimate for the whole world, 1901-2005, M_0=7.0, depth ≤ 100 km.

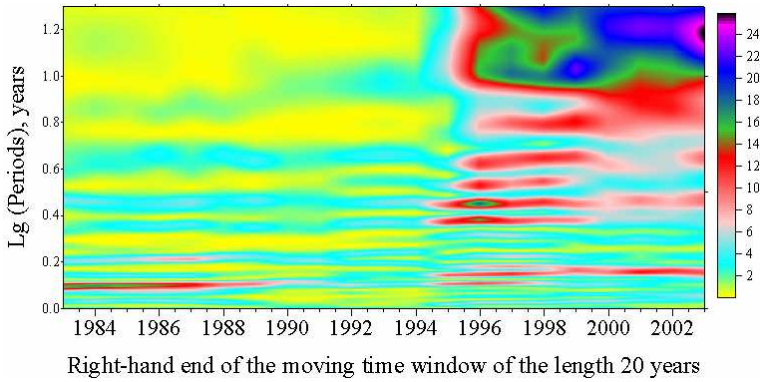

Fig.3. The time-frequency estimate for the region: (Japan + Kuril Islands + South Kamchatka), 1963-2003, M_0=6.0, depth ≤ 100 km.

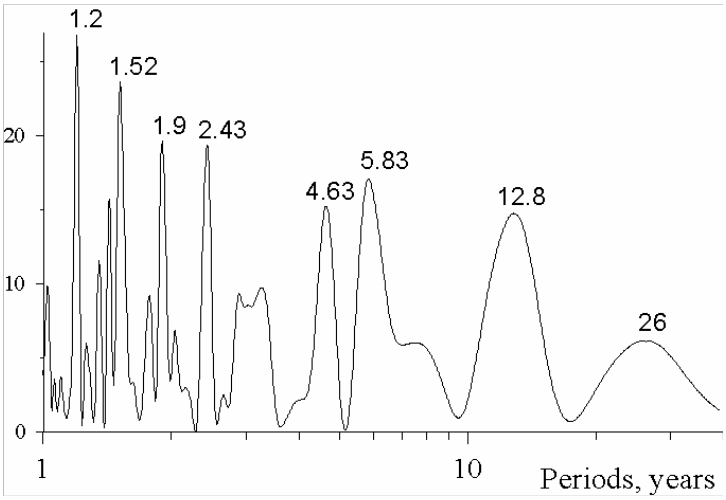

Fig.4. The static estimate for California, 1963-2001, M_0=4.5, depth ≤ 100 km.

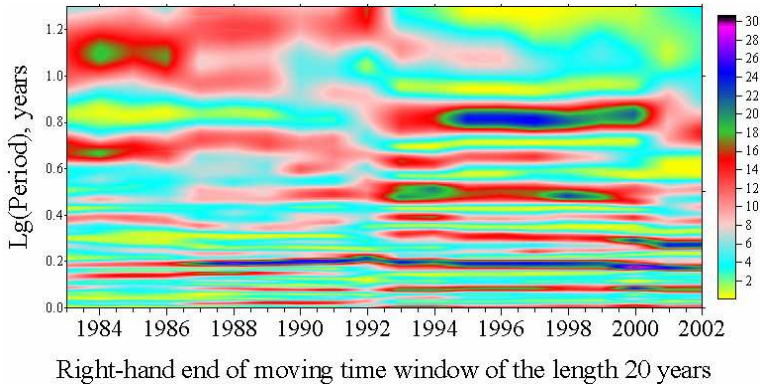

Fig.5. The time-frequency estimate for California, 1963-2001, M_0=4.5, depth ≤ 100 km.

Seismic process at the vicinity of the Sumatra, 26.12.2004, M=9, earthquake for time interval 1963-2004.

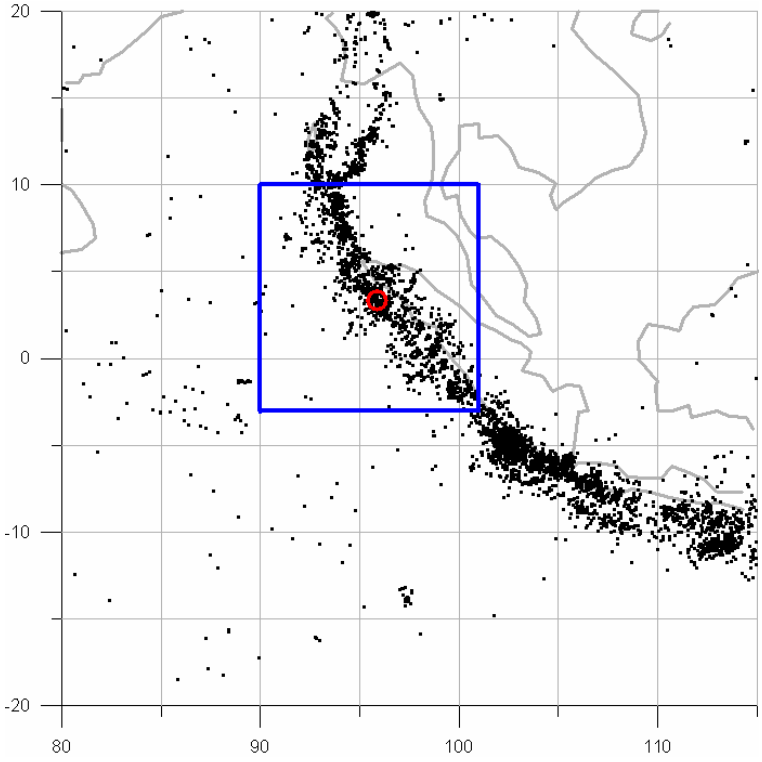

Fig.6. Distribution of earthquakes epicenters with $M \geq 4.5$, depth ≤ 100 km in South-Eastern Asia, 1963-2004. Red cycle – epicenter of the Sumatra, 26.12.2004, M=9, earthquake. Blue rectangular domain is the vicinity of this shock to be investigated. Data were taken from http://neic.usgs.gov/neis/epic/ Aftershocks were not removed.

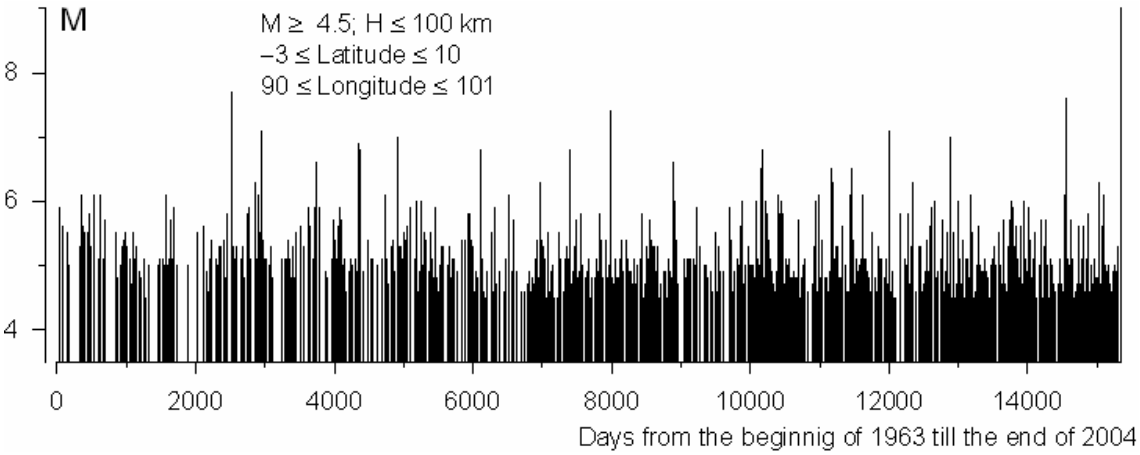

Fig.7. Seismic process within rectangular domain at the Fig.6. The general number of events before the 26.12.2004 shock equals 1387.

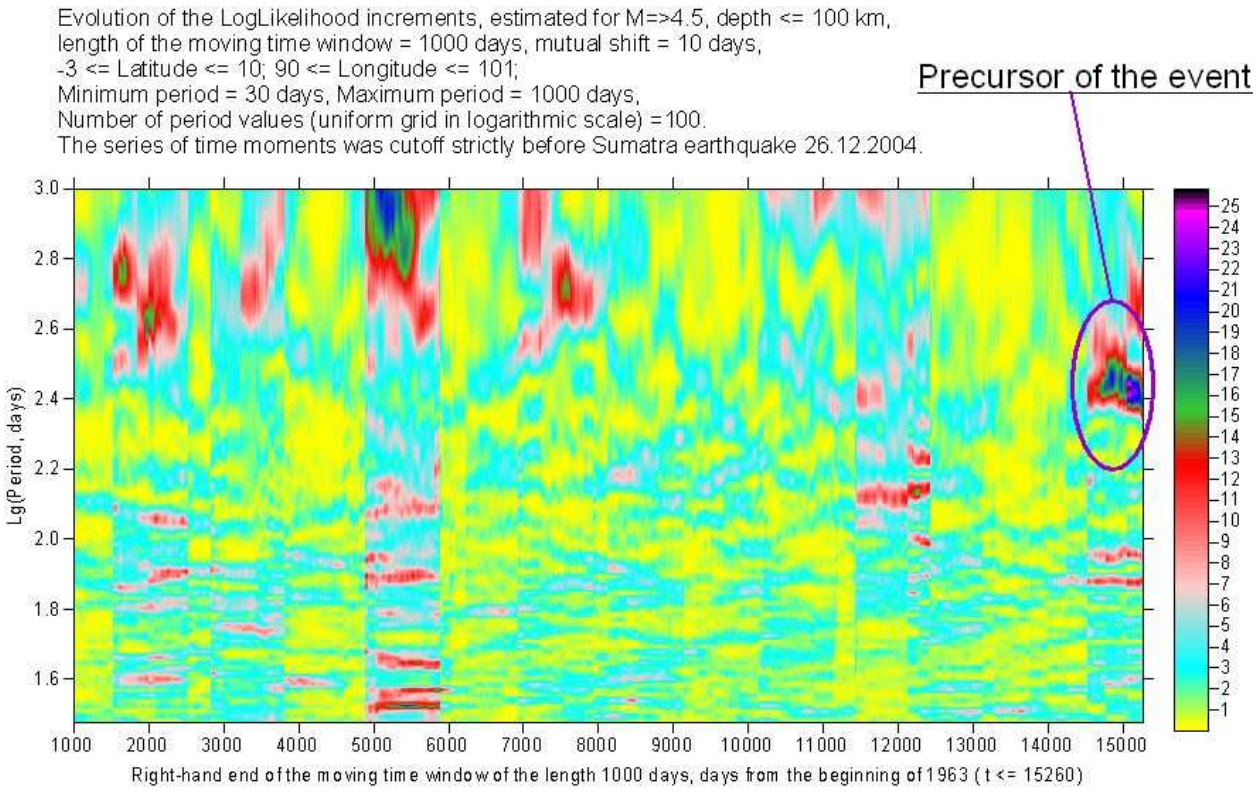

Fig.8. Statistics $R(\omega,\tau\,|\,T_W)$ for the length of moving time window 1000 days. The "strip-like" structure of the pattern is due to the influence of aftershocks series.

Increasing of periodic oscillations of seismic process intensity with the periods 250-320 days turns to be statistically significant precursor of the Sumatra, 26.12.2004, M=9, earthquake. Statistical significance of the result follows from the following experiment: the artificial Poissonian series of time moments of the same average intensity as for the sequence presented at the Fig.2 was generated. The length of this artificial time series was taken 10^5 events. Afterward the method of estimating statistics $R(\omega,\tau\,|\,T_W)$ was applied to this artificial series with the same parameters as at the Fig.8. The maximum value of $R(\omega,\tau\,|\,T_W)$ for this experiment equals to 10.8. Thus, the precursory peak at the Fig.8 is statistically significant. It should be underlined that aftershocks series were not removed before processing. The 2D diagram at the Fig.8 evidently contains characteristic strip-like structure because of the presence of aftershock series after rather strong earthquake. The main reason for leaving aftershock is that periodic component of seismic intensity can be most strong within these aftershocks series. The analogy with Earth's oscillations after strong earthquakes can be presented: although these oscillations take place at each time interval they could be reliably extracted after strong earthquakes only. The same situation seems to be with periodic component of seismic intensity: it became sufficiently strong within aftershock series of some moderate earthquakes. The precursor of Sumatra earthquake is a peak of periodic intensity within aftershock series of previous earthquake 02.11.2002, M=7.6, Latitude=2.82, Longitude=96.08. At this sense this earthquake could be regarded as generalized foreshock of Sumatra earthquake.

2. Interactive program Spectra_Analyzer for data mining properties of scalar time series

This program includes options of spectral and time-frequency analysis, band-pass filtering, wavelet-packet decomposition and nonlinear wavelet threshold filters, estimating trends by Gaussian kernel and local polynomial smoothing, etc.

The user's interface of the program was written on Visual Basic. All computational work is performed by dynamic link library SigSub.dll which includes all numerical procedures and was written in Compaq Visual Fortran. The file SigSub.dll must be within the same directory as module Spectra_Analyzer.exe. The following DLL-files are necessary for working the program: Dforrtd.dll, Msvcrtd.dll and msvbvm60.dll. These DLL-files could be already written into the directories c:\windows\system or c:\windows\system32

The program:

1) estimates power spectra of scalar time series of an arbitrary length (appendix **A5**);

2) performs smoothing and de-trending by Gaussian kernel or local polynomials ($0 \le$ order ≤ 10) within moving time window of the given length or for the whole series (appendix **A6**);

3) makes Fourier band-pass filtering after selecting frequency band at the spectra graphic ('Zoom Spectra' operation);

4) estimates power spectra evolution within moving time window using AR-model (appendix **A5**);

5) makes orthogonal wavelet packet decomposition by splitting detail levels into 1, 2, 4 or 8 sub-levels (appendix **A2**);

6) makes wavelet threshold filtering by selecting 'large' and 'small' values of absolute wavelet coefficients (appendix **A3**);

7) finds long chains of Wavelet Transform Maximum Modulus (WTMM) skeleton of the selected signal fragment using wavelets equal to derivatives of Gaussian of the orders 0 (usual kernel), 1 and 2 ('Mexican hat') (appendix **A6**);

8) estimates time-frequency map of logarithm of squared Morlet continuous wavelet transform (CWT) coefficients (appendix **A4**);

9) estimates time-frequency map of logarithm of squared orthogonal wavelet-packet coefficients ('Heisenberg boxes') (appendix **A2**);

10) estimates best fitted low-frequency harmonic with unknown period;

11) seeks for period and 2 extreme points of 'maximum wave' of the signal for fragment length $\le$ 500 samples;

12) estimates multifractal singularity spectra by detrended fluctuation analysis (appendix **A7**).

Operations 7 – 12 are performed for selected fragment of the signal. All results of linear and nonlinear filtering could be saved on the disk at the same directory as the initial data file with long 'mnemonic' file names.

Auxiliary preprocessing operations as down-sampling, coming to incremental time series, winzorization of outliers and extracting fragments are possible.

Winzorization procedure means calculating average value $\overline{x}$, standard deviation σ, clipping values exceeding or being lower thresholds $\overline{x} \pm 4 \cdot \sigma$ and repeating this sequence of operations until values of $\overline{x}$ and σ will not cease yet to vary.

The input file could be any ASCII table. The 1st column of the input file must be successive values of time series samples, all other columns of the table (if any) are ignored.

If it is necessary to "reach" other columns of the input file table for their processing it could be done by pressing the button "**D**" (de-multiplexing input table). At this case program will find the number of numerical columns from left to right and will output them into the disk as ASCII files with one-column structure and filenames with postfixes "_Col01", "_Col02" and so on. If the input file has one column only this operation has no effect.

'Mnemonic' abbreviations within output filenames:

Sp_AR(...) – maximum entropy power spectra estimate, based on AR-model;

Sp_Four(...) – power spectra estimate, based on Fourier expansion and averaging peridograms;

BPF – after Band-Pass Filter;

TRN – trend component obtained by Gaussian kernels smoothing or local polynomials within moving time window;

RSD – residual component after removing trend;

INC – increments of the signal;

DNS – after Down-Sampling into the given ratio:
if down-sampling step is integer (≥ 2) then by taking successive mean values in time windows;
if down-sampling step is fractional (> 1) then down-sampling is performed using Fast Fourier Transform;

Var.DNS – sample variance estimates after Down-Sampling into the given integer ratio;

WINZ – after winzorization procedure (iterative clipping of outliers);

PART[Tstart,Tfinish] – operation for selected fragment of the signal, Tstart ≤ Time ≤ Tfinish;

WDC – Wavelet-packet Decomposition Component after double-clicking it on the components map;

WaveSp(...) – wavelet-packet power spectra estimate, i.e. time-averaged values of squared wavelet-packet coefficients;

Lg(Sp)_*.grd – ASCII GRID-file of time-frequency map of power spectra estimates lg-values within moving time windows;

THR-MAIN – result of orthogonal wavelet threshold filtering ('wavelet denoising operation') i.e. removing of given share of minimum modulus wavelet coefficients - 'a signal';

THR-RSD – residual after subtracting the result of above operation from the signal, i.e. 'noise'

Lg(WaveCoeff^2)_*.grd – GRID-file of decimal logarithms of squared Morlet wavelet coefficients;

Sp(WaveCoeff^2)_*.dat – Morlet wavelet coefficients power spectra in the format:
<Period, WaveCoeff.^2>.

Lg(AbsWaveCoeff_HBoxes)_*.grd – GRID-file of decimal logarithms of squared orthogonal wavelet-packet coefficients;

F(Alfa)_*.dat - multifractal singularity spectrum estimated for chosen fragment of the signal by Detrended Fluctuation Analysis.

Spectra estimates files are saved in the sub-directory 'SpectEstimates', grid-files - in the sub-directory 'GRID_files.'

Two spectral estimates are provided:

1) The estimate within overlapping moving time windows of some length LengF samples based on the Fourier expansion, frequency averaging of periodograms within each time window using frequency window of the given radius MM and final averaging these estimates from each time window. Thus, the periodogram averaging has 2-stages: the 1^{st} stage is averaging by frequency values and the 2^{nd} one is averaging by different time windows.

The default value of the frequency averaging radius MM = 1.
The default value of Fourier-based spectral estimate time window LengF = min (2048, Nrec).
The default values of MM and LengF could be changed.

The Fourier-based spectral estimate serves as the background for the 2-nd estimate of maximum entropy and is plotted by green color and is not zoomed.

2) The maximum entropy (Burg's) estimate within overlapping moving time windows of some length LengAR, with some auto-regression (AR) order Nar. The AR spectral estimates from each position of moving time window are averaged by all these windows for obtaining the final estimate.

The default value of AR-based spectral estimate time window LengAR = min (2048, Nrec).
The default value of AR-order Nar = min (LengAR / Fc, 1000), where
Fc = 6.5 - 3.5 * Cos(Pi * (LengAR - 32) / (2048 - 32)):
when LengAR = 32 then Fc = 3, but when LengAR = 2048 then Fc = 10.
The default values of LengAR and Nar could be changed.

The AR-estimate is the main one and is used for seeking typical periods or frequency values. It is plotted by dark blue color and could be zoomed.

The minimum possible value of Nrec, LengF and LengAR equals 32.
The maximum possible value of LengF and LengAR equals Nrec.
(MM)min = 0, (MM)max = LengF / 8; (Nar)min = 1, (Nar)max = LengAR / 4.

The spectral estimates could be visualized for any combination of linear or log scale X and Y axis and for periods or frequency values for X-axis.

The estimates could be saved on the disk in the sub-directory (of data directory) named 'SpectEstimates' in the file with the 'long' names which stores information about signal input filename and estimates parameters

The spectra estimates could be zoomed by using horizontal scroll bars and the signal could be band-pass filtered within chosen frequency band and the result of band-pass filtering could be saved on the disc with the name having 'BPF' abbreviation inside output filename.

The low-frequency component (trend) and residual after trend removing could be estimated and saved at the disc either by using local polynomials of the orders 0–10 and within moving time window of the given radius or by Gaussian kernel smoothing with given averaging radius.

The wavelet-packet decomposition of the given signal with splitting each usual detail level into 2, 4 and 8 sub-levels could be performed and visualized. Orthogonal wavelet order is found automatically from minimum entropy criterion.

The wavelet-packet decomposition component ('WDC' inside filename) could be saved on the disk with a long filename by double-clicking the component graph on the wavelet decomposition components map.

The orthogonal wavelet nonlinear threshold filtering, i.e. removing of the given share of minimum modulus wavelet coefficients ('THR-' inside output filename) is available.

The Wavelet Transform Maximums Modulus (WTMM) skeleton long chains for 3 types of Gaussian kernel derivatives (of the orders 0, 1 and 2) are estimated and visualized for selected (zoomed) fragment of initial time series.

General instruction: move mouse cursor within graphic windows to inquire information about time, frequency or period values, etc. as mouse tool-tips.

The next figure presents a block-scheme of the user’s interface which could be visualized in Help menu:

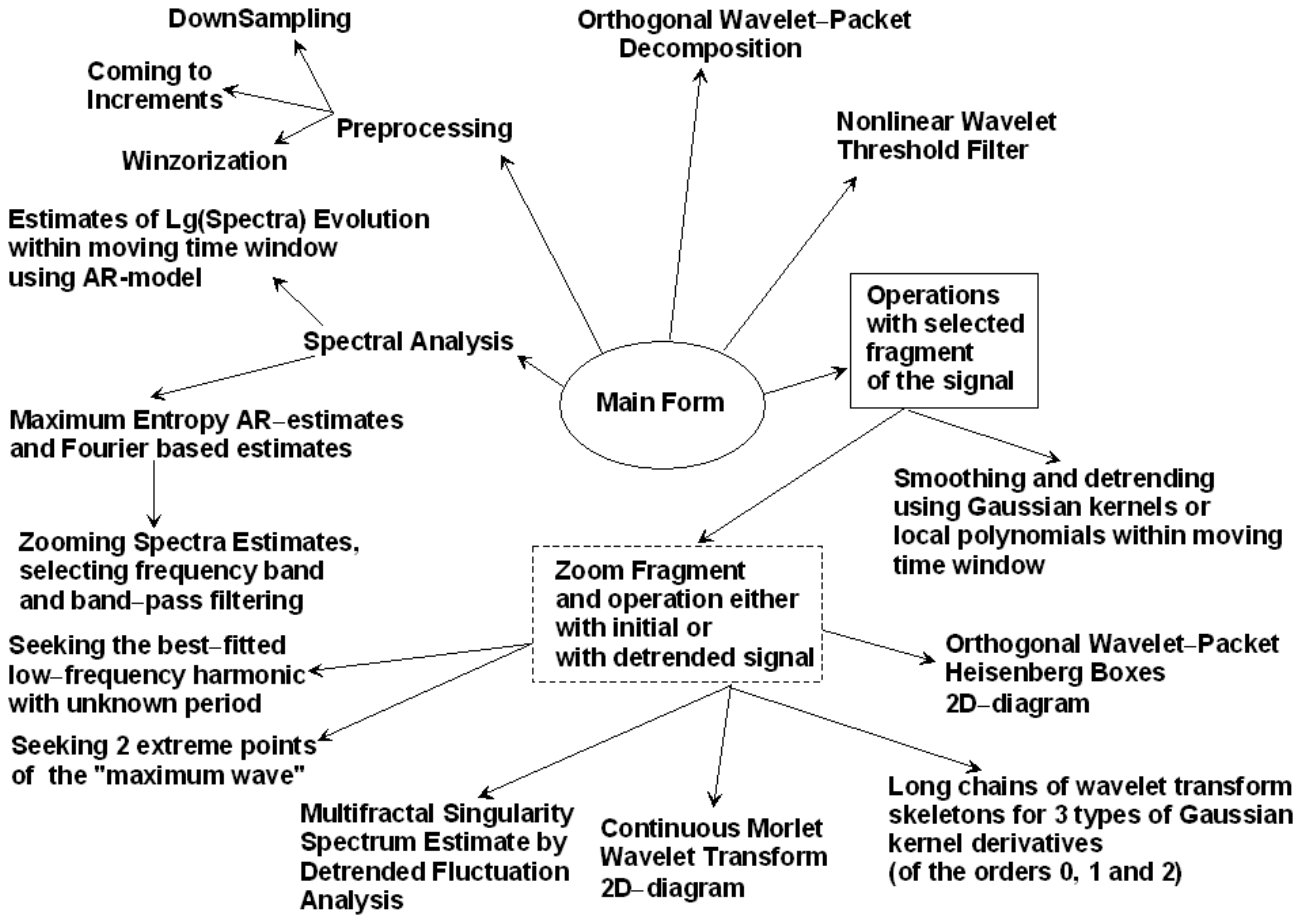

Fig.9. Block-scheme of the user's interface of the program "Spectra_Analyzer".

Examples of screen snapshots illustrating data processing:

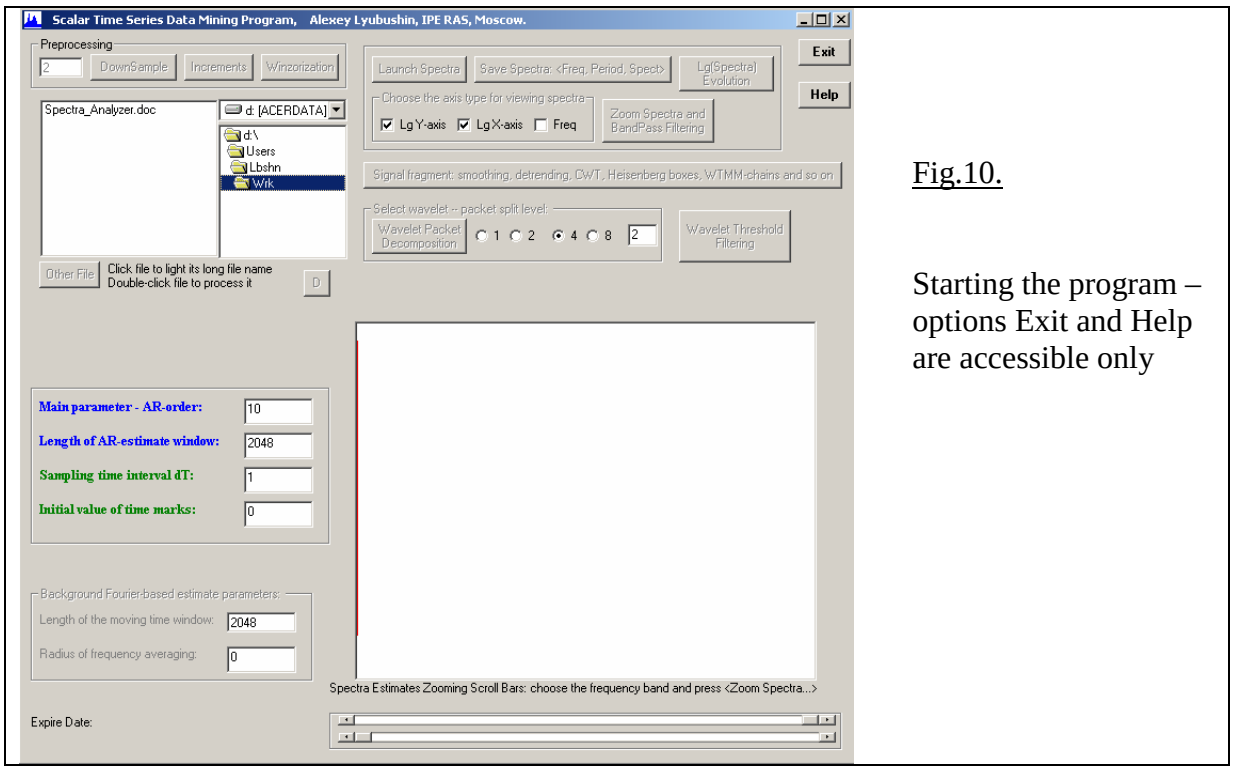

Fig.10.

Starting the program – options Exit and Help are accessible only

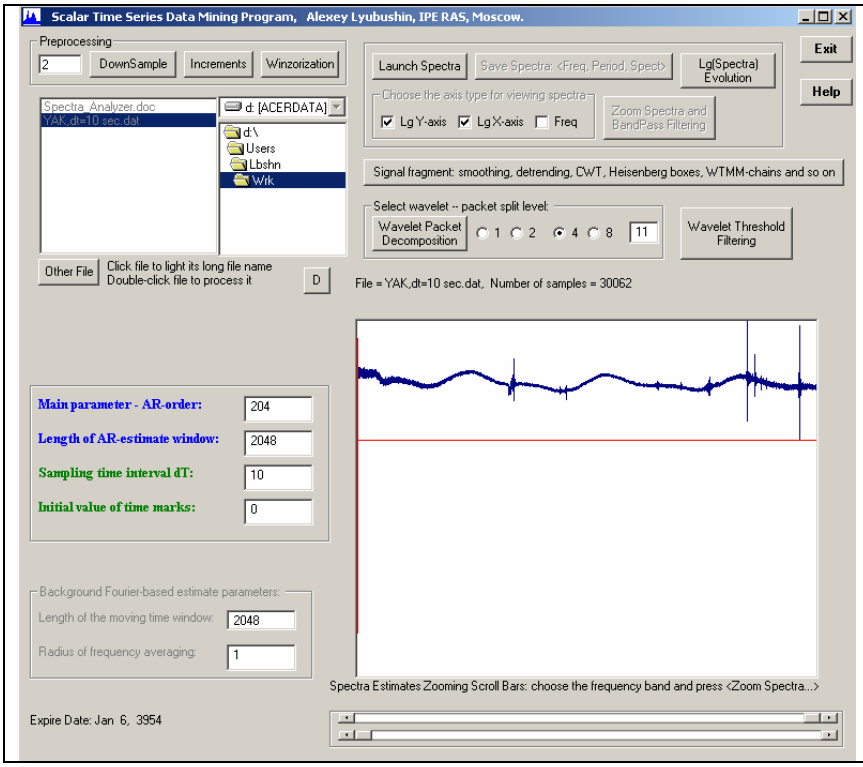

Fig.11.

Here an input file was chosen and sampling time interval 10 sec was set. Initial time mark for the 1st sample is 0. The series contains huge outlier – let us apply winzorization procedure to it.

The result will be the following:

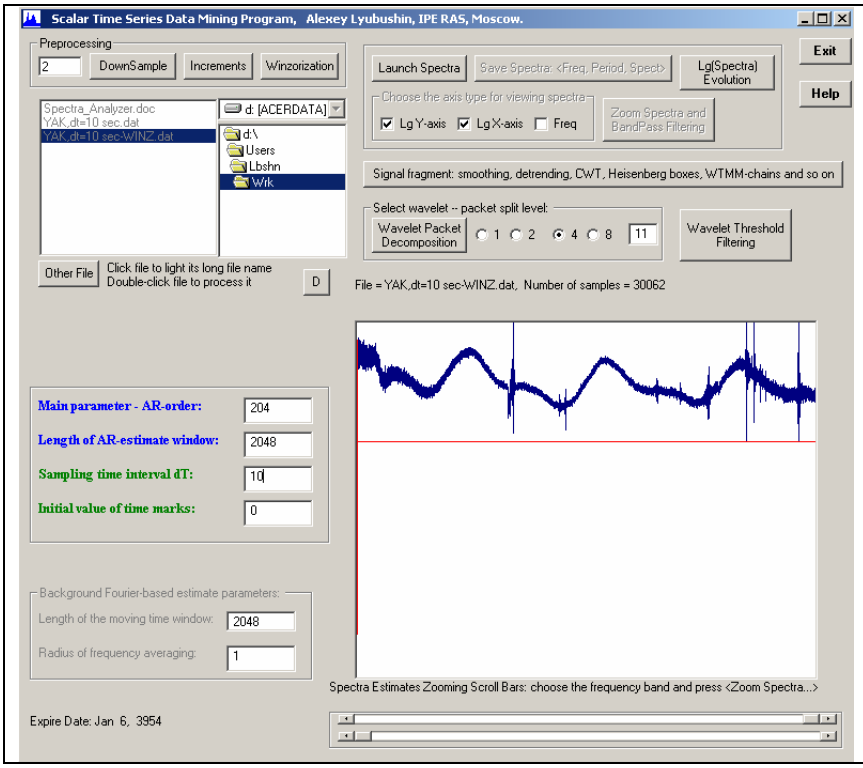

Fig.12.

– the after iterative outliers clipping. Let us estimate its power spectra with default parameters and plot the result in double logarithmic scales in dependence on periods (flag “Freq” is not tagged – otherwise the plot would be in dependence on frequency).

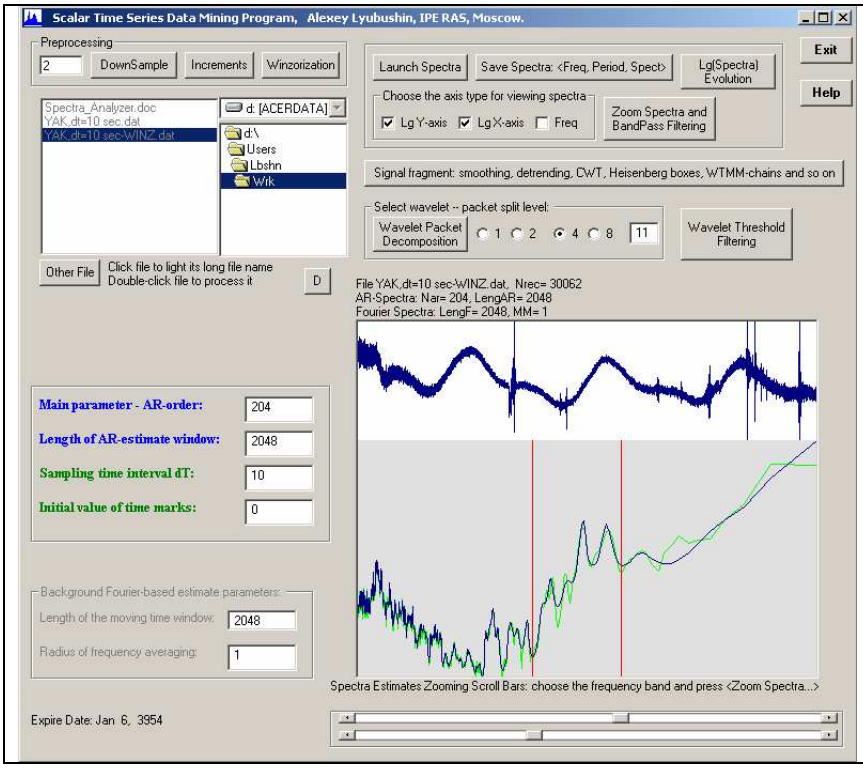

Fig.13.

Two graphics of spectral estimates are presented – AR-estimate of maximum entropy (dark blue color) and background Fourier-based estimate (light-green color). Let us extract some frequency band using horizontal scroll bars and corresponding vertical red boundary line. After selecting band press button Zoom Spectra:

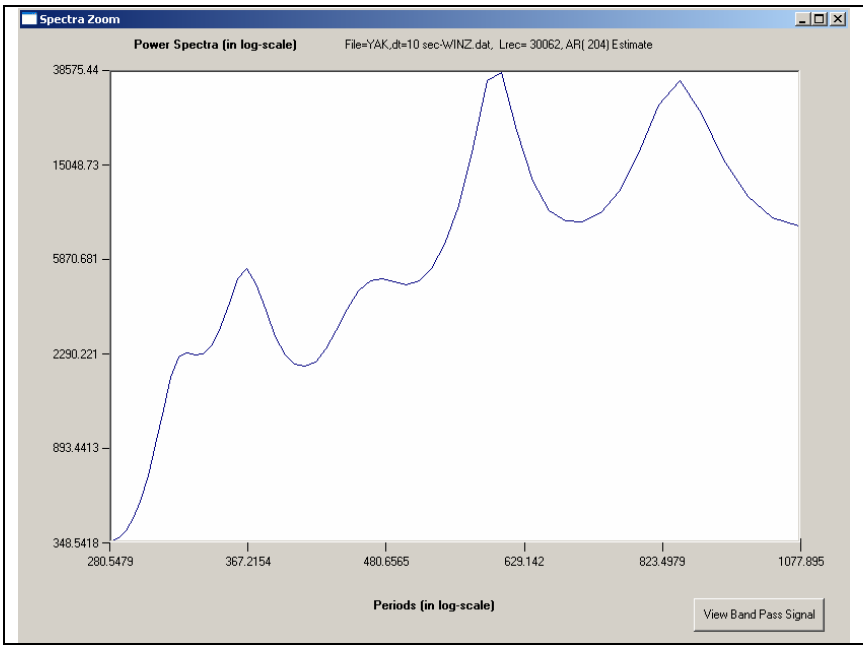

Fig.14.

– Zoom Spectra. By pressing button View Band Pass Signal we can look at the result of Fourier band-pass filtering and save the result on the disk into separate file:

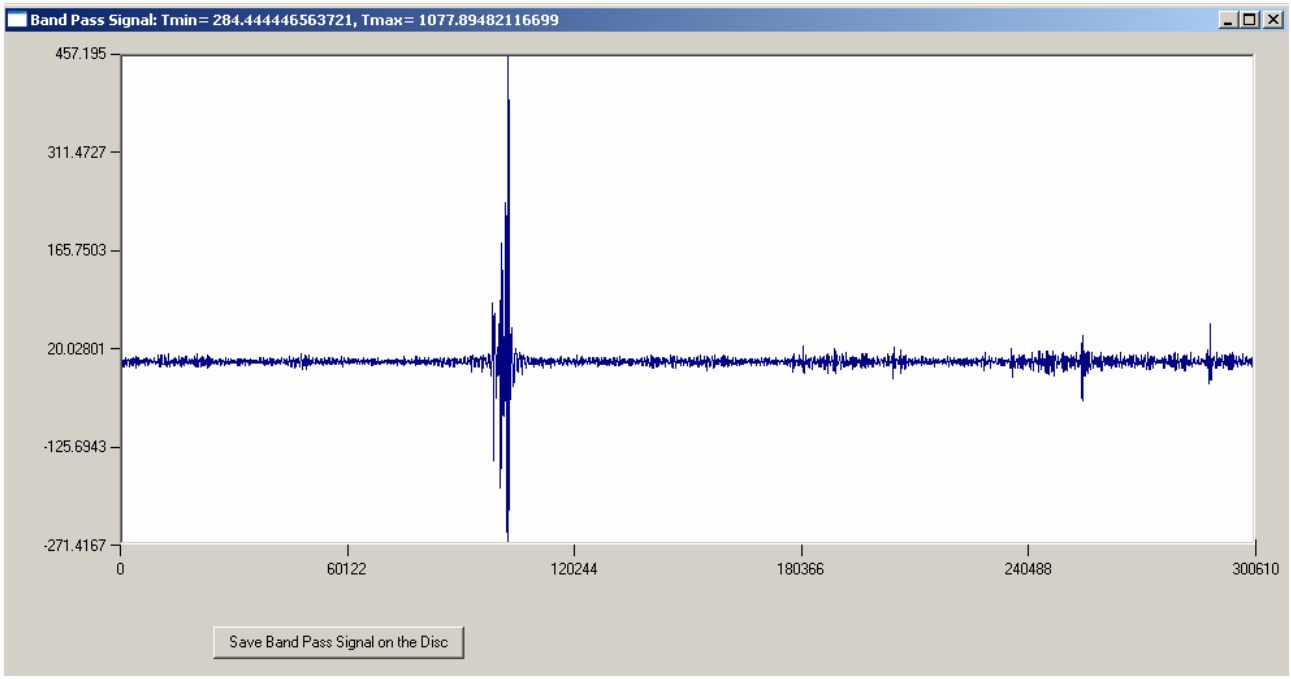

Fig.15. Result of band-pass filtering. It could be saved on the disk at separate file (mnemonic abbreviation BPF within filename).

Let's decrease the number of samples by averaging and down-sampling into 6 times, i.e. come to sampling time interval 1 minute (DownSample: 6):

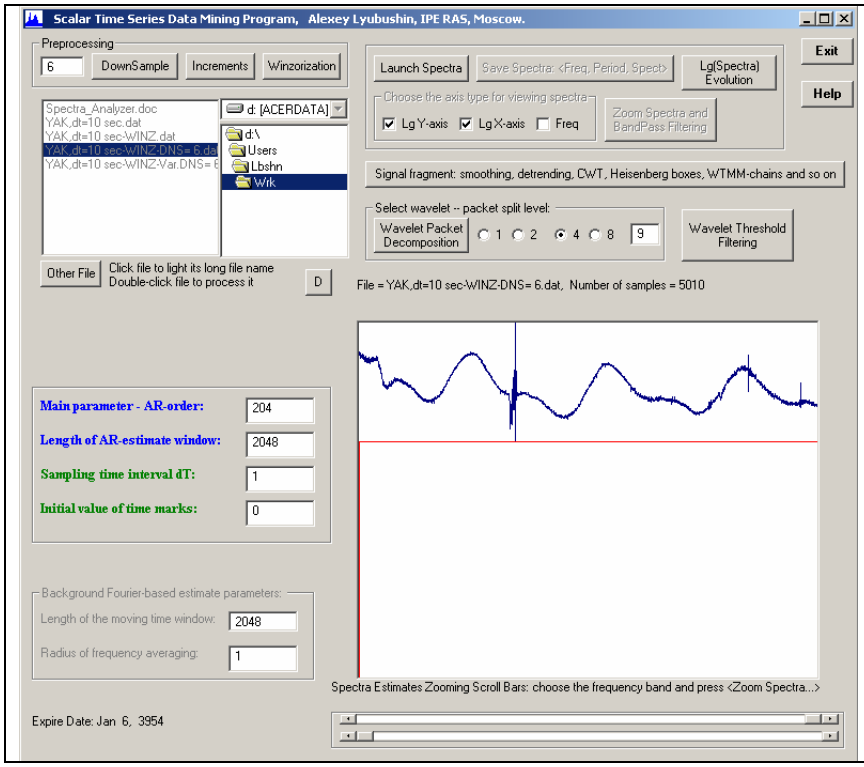

Fig.16.

– time series after averaging and down-sampling into 6 times, new δt = 1 minute.

Let's pass this signal through the operation of Wavelet Threshold Filtering:

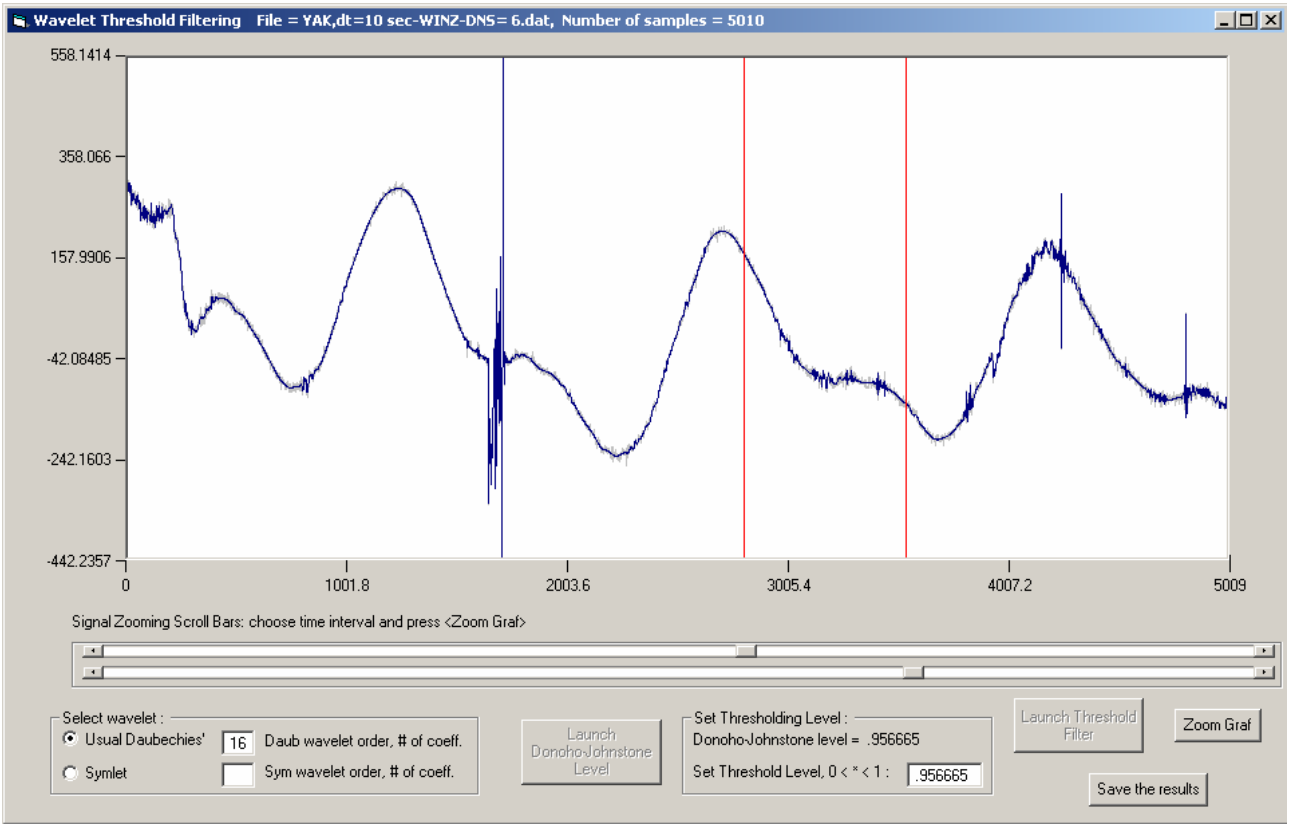

Fig.17. The program defined the optimal wavelet – Daub16 (the default basis, it can be changed) and the default Donoho-Johnstone threshold 0.956665 (it can be changed by the user as well). By applying these parameters for threshold filtering 4.34% of wavelet coefficients only provided the main variations of the signal (deep-blue line). The initial signal is plotted by light-grey color as background. The result could be visualized in details by zooming any fragment using horizontal scroll bars:

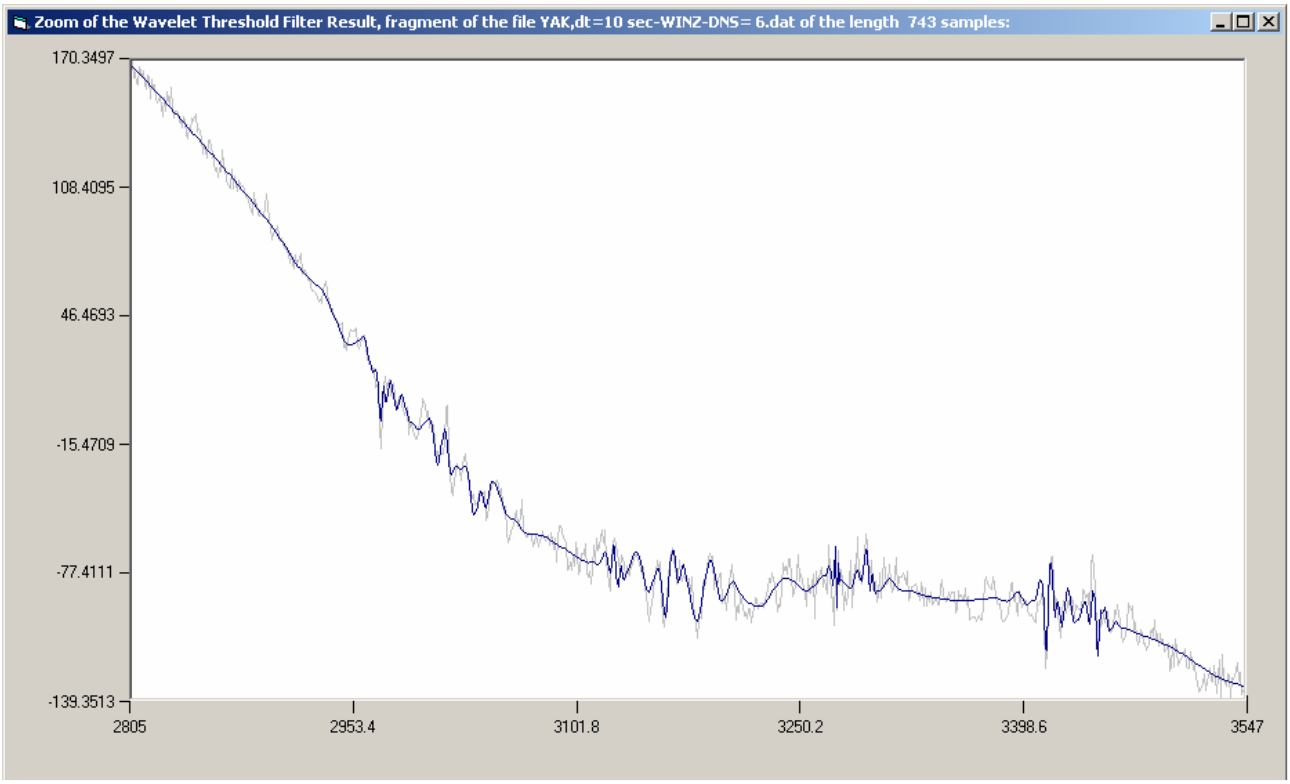

Fig.18. After saving the result of threshold filtering on the disc it is possible to look in details at the difference between initial signal and the most informative variations (a noise components):

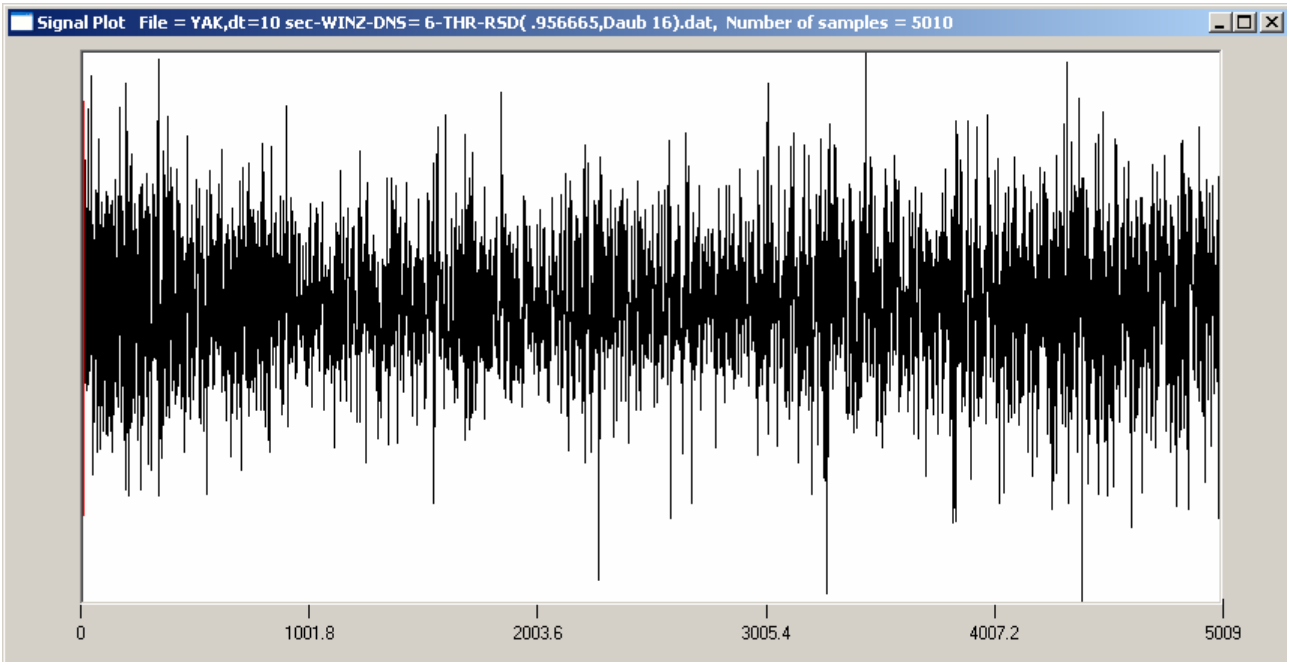

Fig.19. It should be underlined that the header of each graphic form contains a long filename which presents the whole history of the signal processing. Here you can see that this is a result of winzorization, down-sampling into 6 times, threshold filtering using wavelet Daub16 and shrinkage level 0.956665, a residual (THR-RSD).

Let's return to initial signal with 1-minute sampling and remove its trend. For this purpose we must go to sub-menu Signal Fragment:

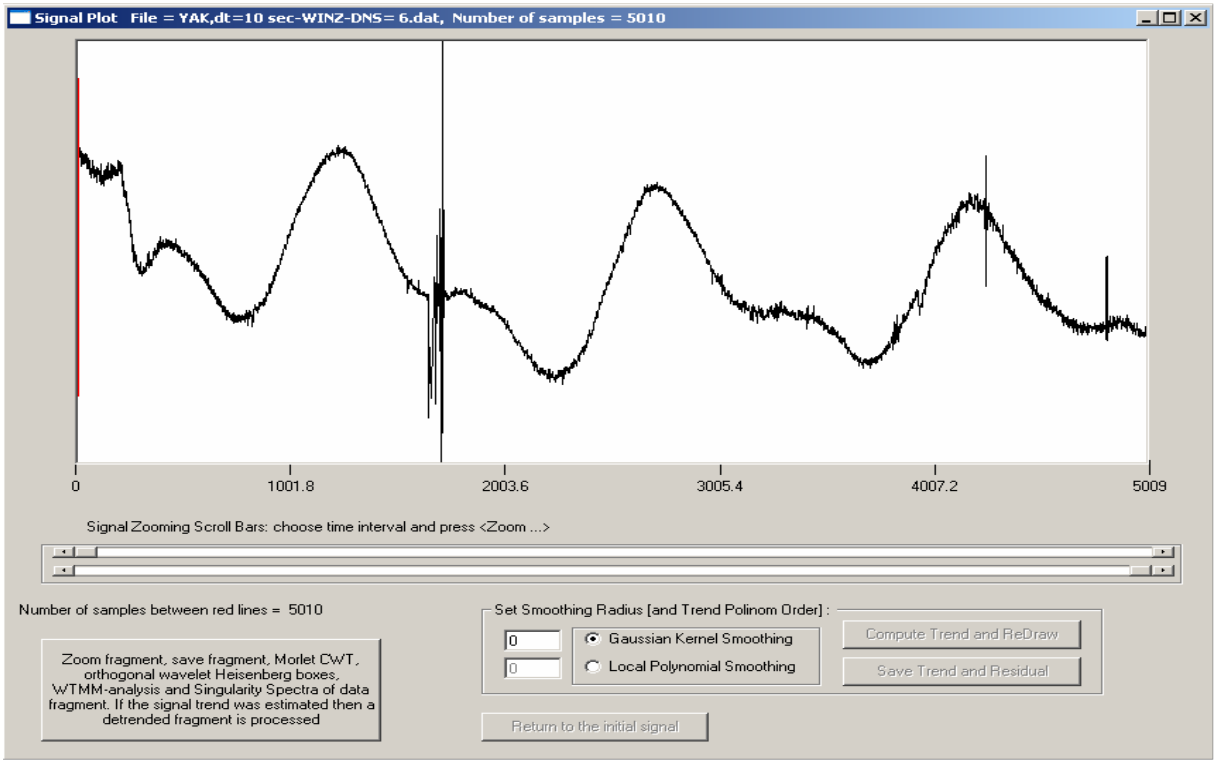

Fig.20.

Let us chose a Gaussian kernel smoothing with radius 100 and look the result:

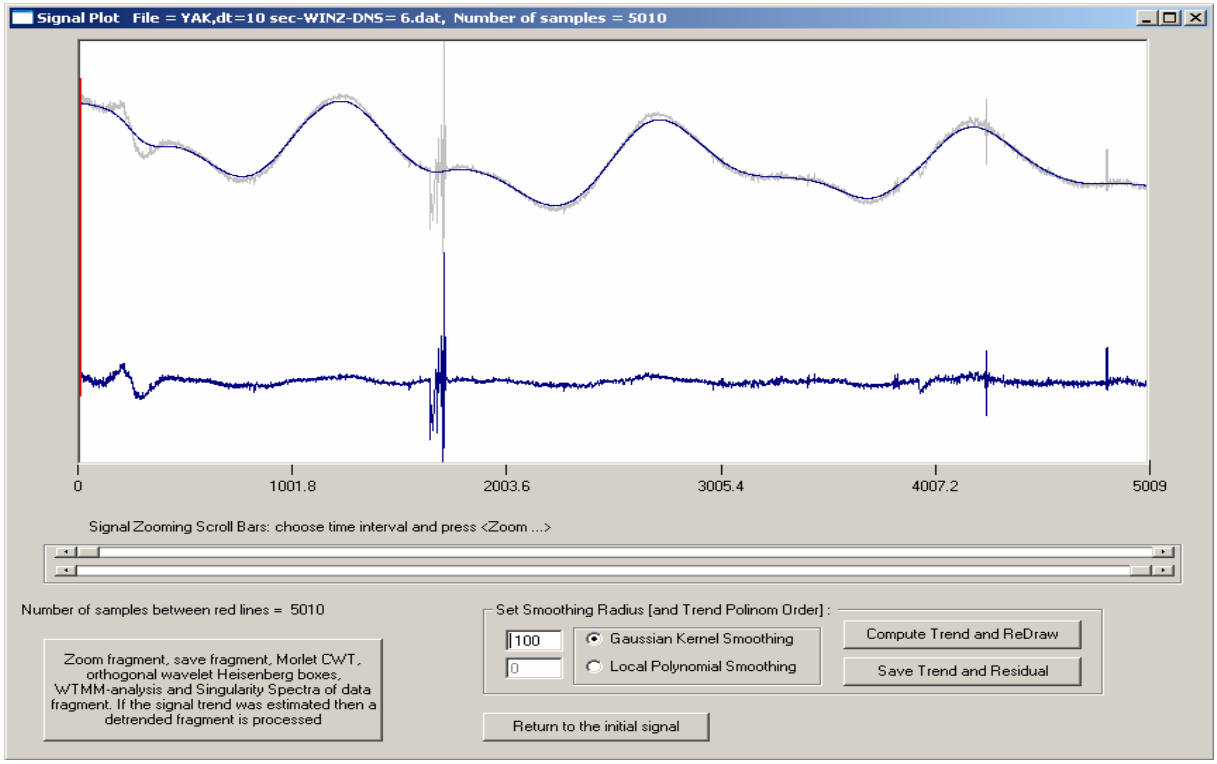

Fig.21. – The result is rather satisfactory except the vicinity of local extreme points of the smoothed curve where it has large deflections from initial signal. This lack could be improved by decreasing the radius of Guassian kernel but the other possibility is using local polynomial trends. At this case parameter 100 will be the radius of moving time window and let us set the polynomial order to 3. At this case the trend estimate will be the next:

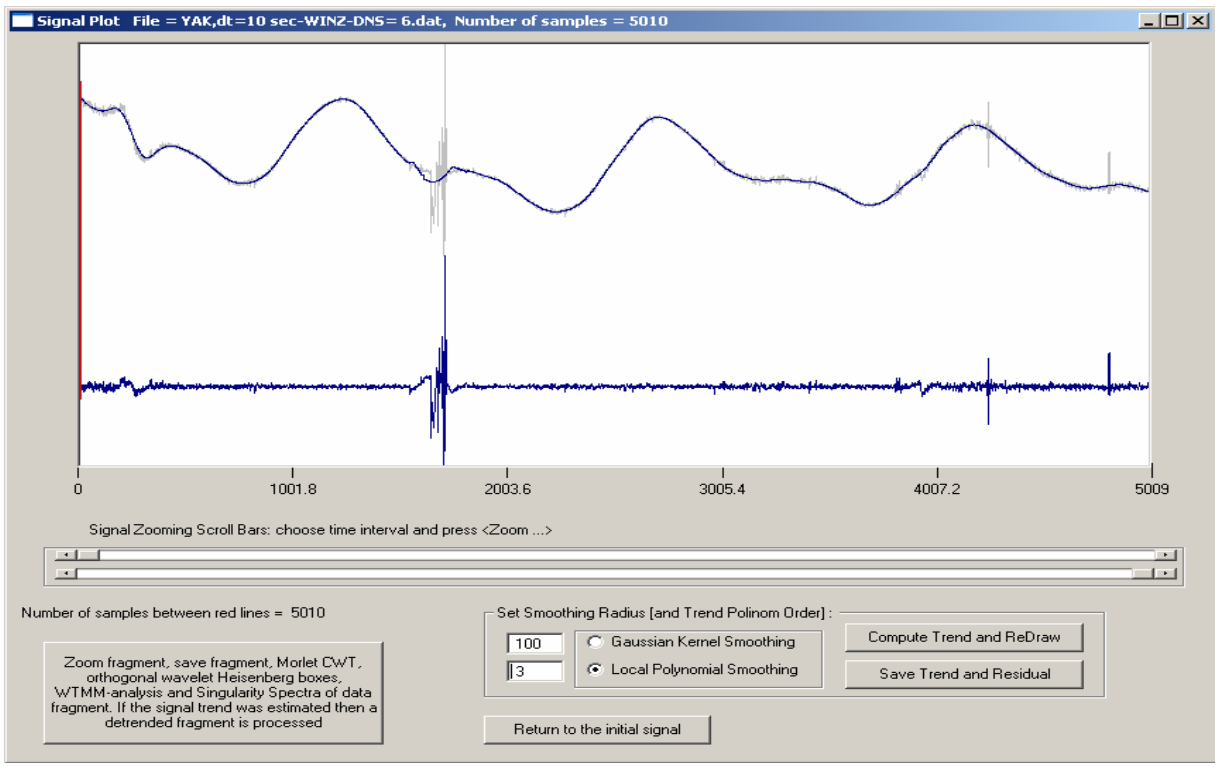

Fig.22.

Details of de-trending operation could be visualized by extracting any fragment by scroll bars:

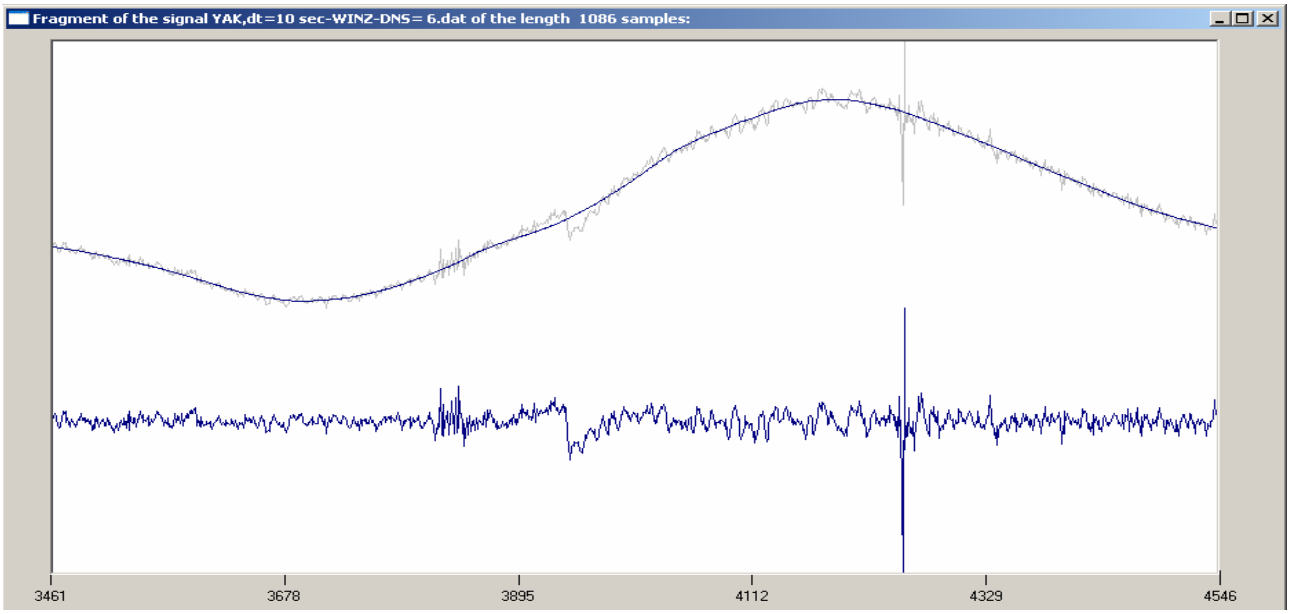

Fig.23.

Let us save the results of trend and residual estimates and work with residual signal. Let's go to sub-menu Operations with selected fragment → Zoom Fragment and operation either with initial or with detrended signal:

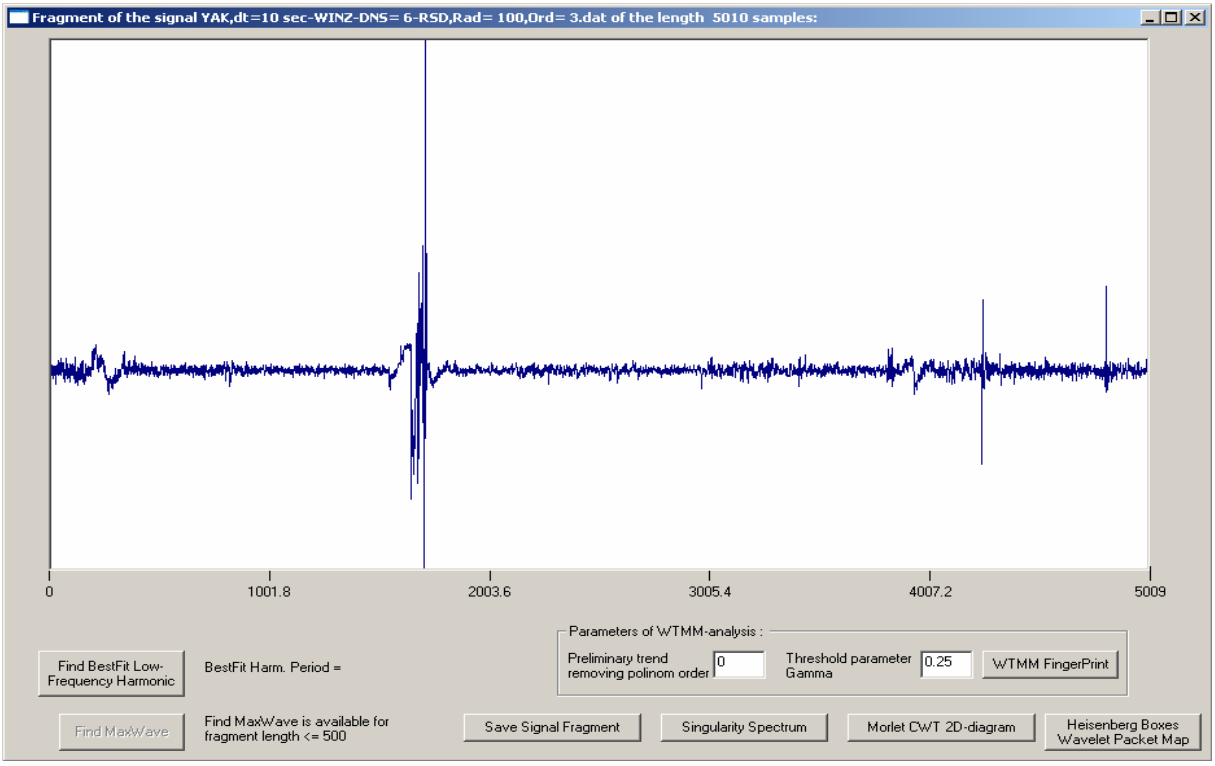

Fig.24.

Let's press the button Morlet CWR 2D diagram:

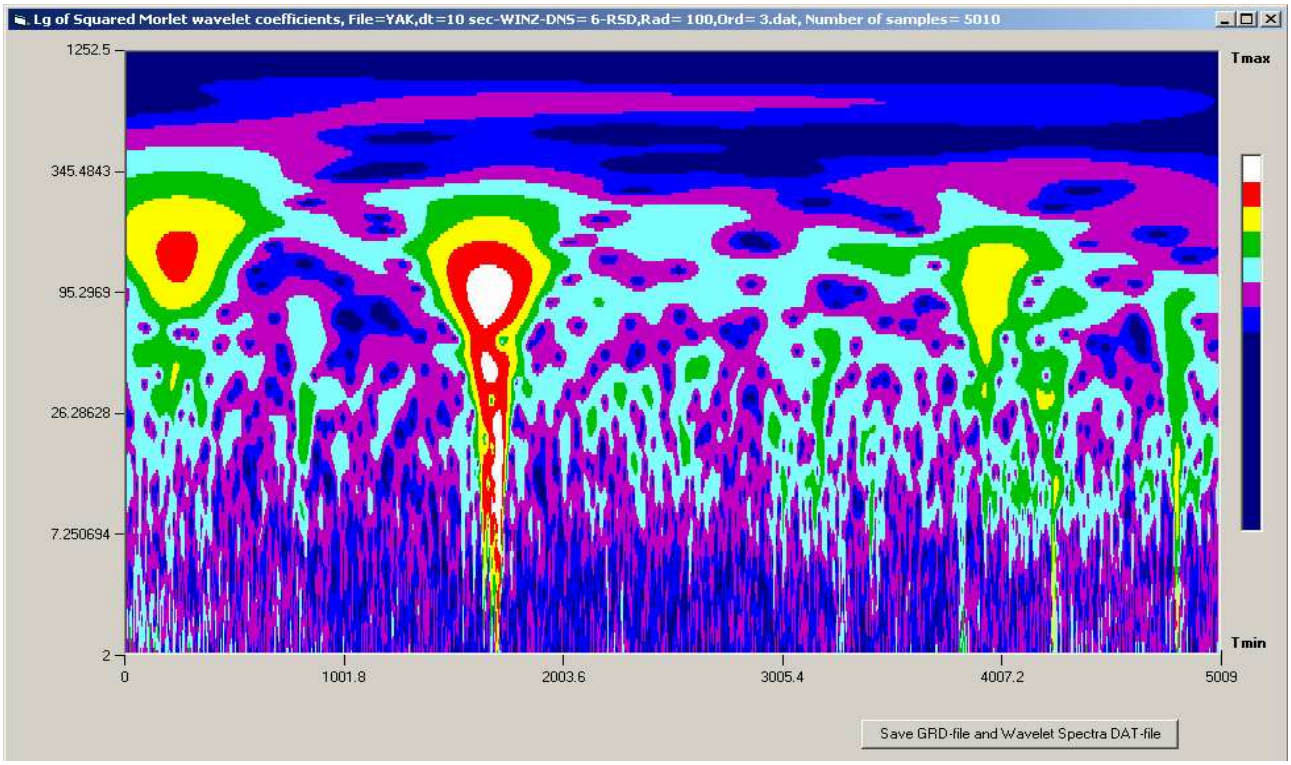

Fig.25. – this 2D diagram presents distribution of logarithm of energy of variations in dependence on time and periods. This image could be saved as GRID-file for further plotting in Surfer.

Other form of presentation – Heisenberg's boxes:

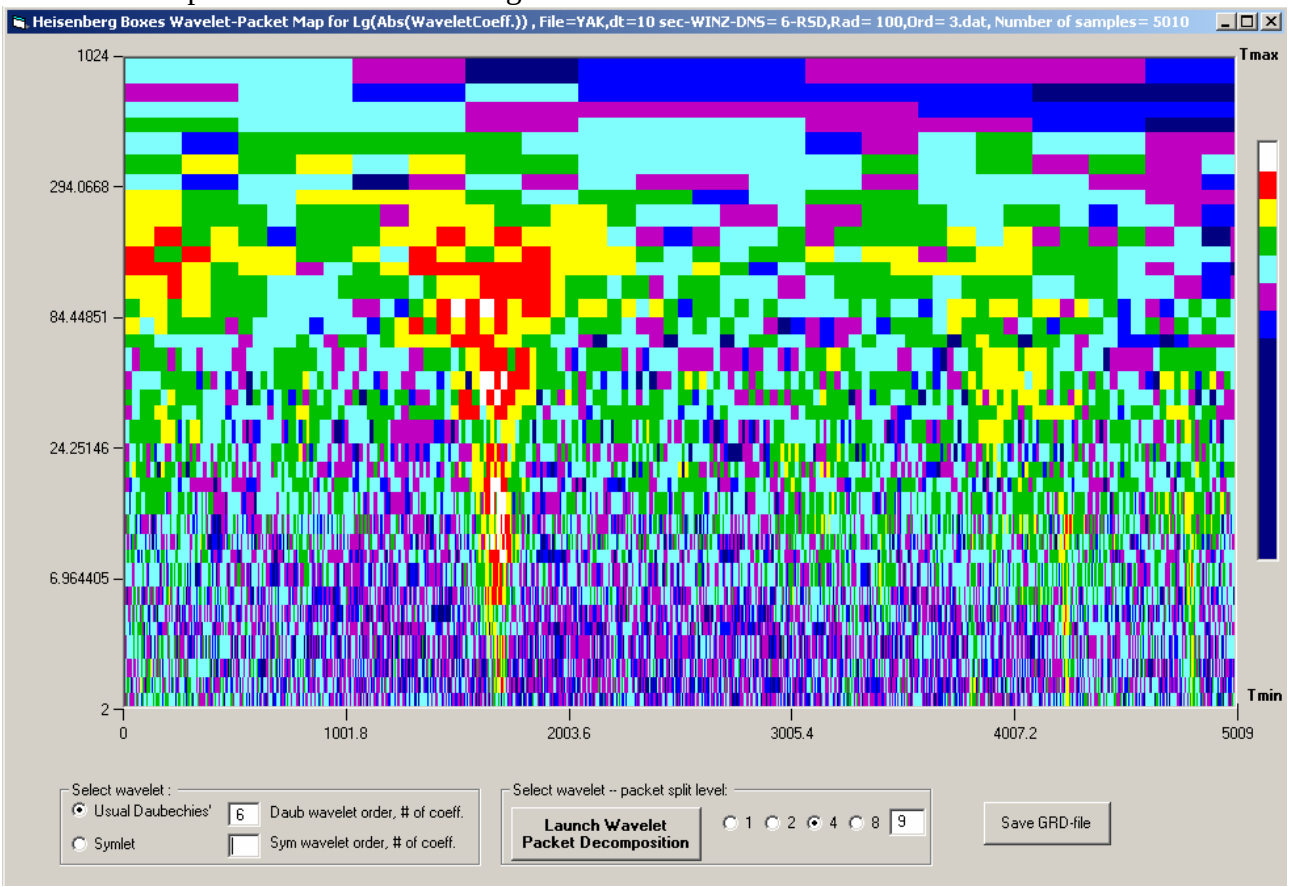

Fig.26. – here the first 9 detail levels are split into 4 time each level, the whole number of sub-levels = 36.

If the inverse wavelet-transform be taken from these wavelet coefficients then a sequence of frequency-ordered signals will be presented (sub-menu Wavelet-Packet Decomposition):

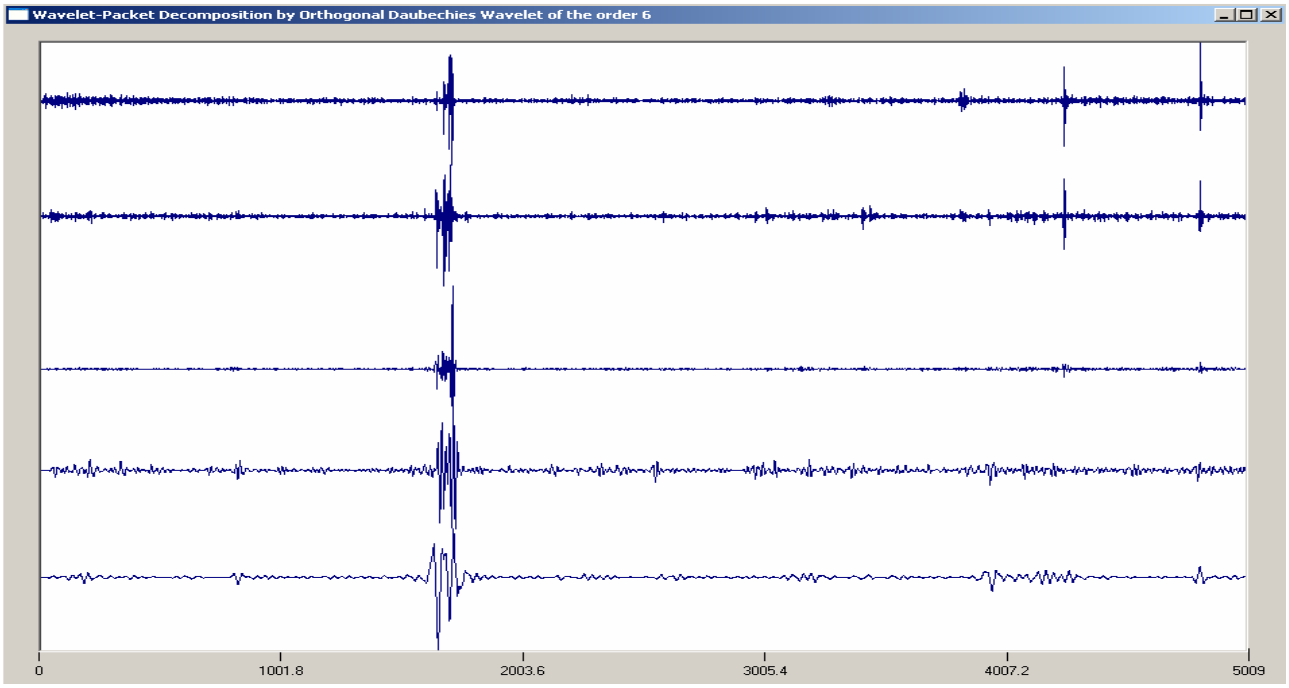

Fig.27. – here it is a usual wavelet decomposition by the first 5 detail levels. If each level will be split into 8 sub-levels that we will obtain:

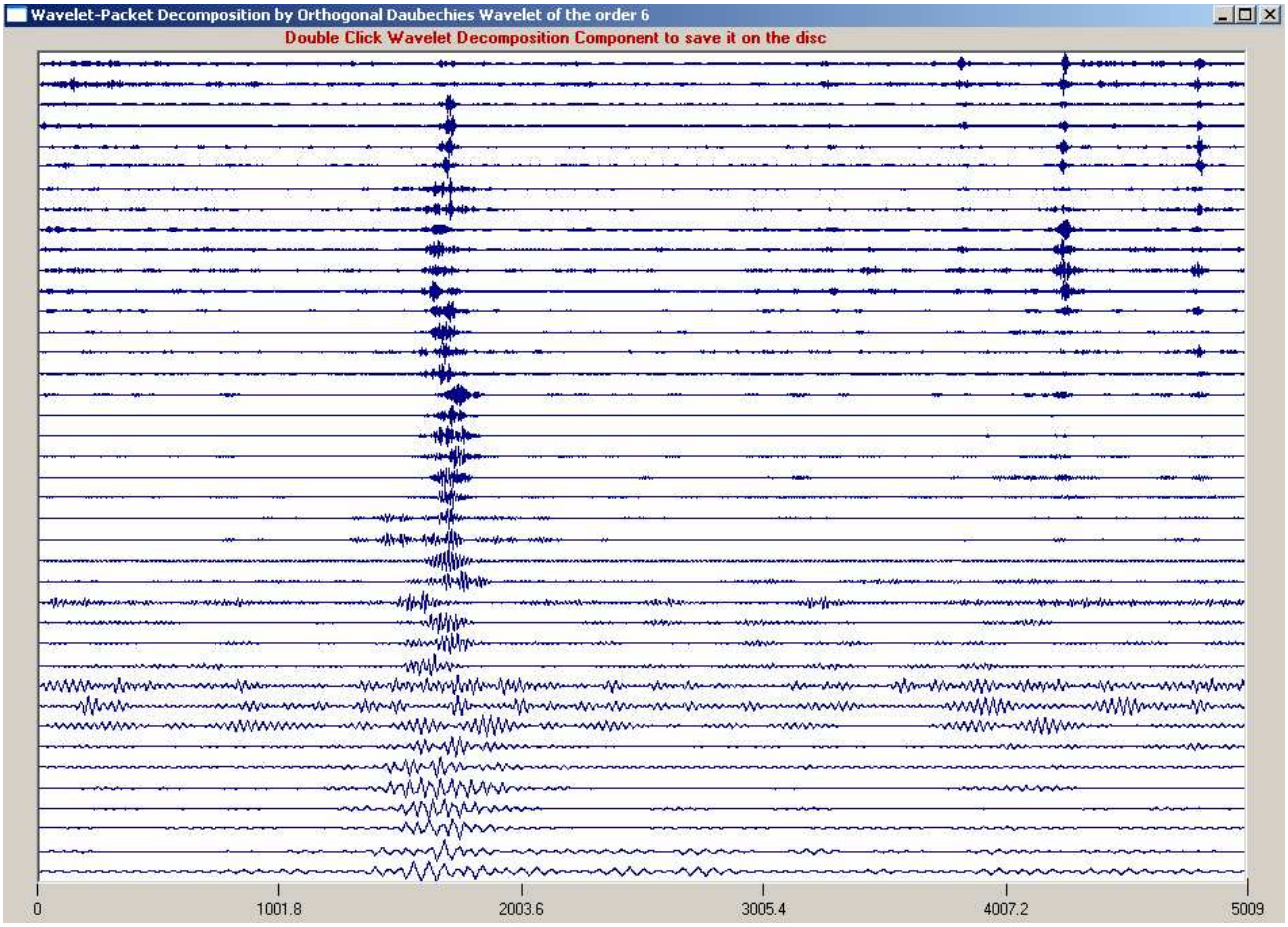

Fig.28.

It is necessary to notice that the most suitable orthogonal wavelet is found automatically. The program outputs wavelet-packet decomposition power spectra and the form of the best wavelet basis function:

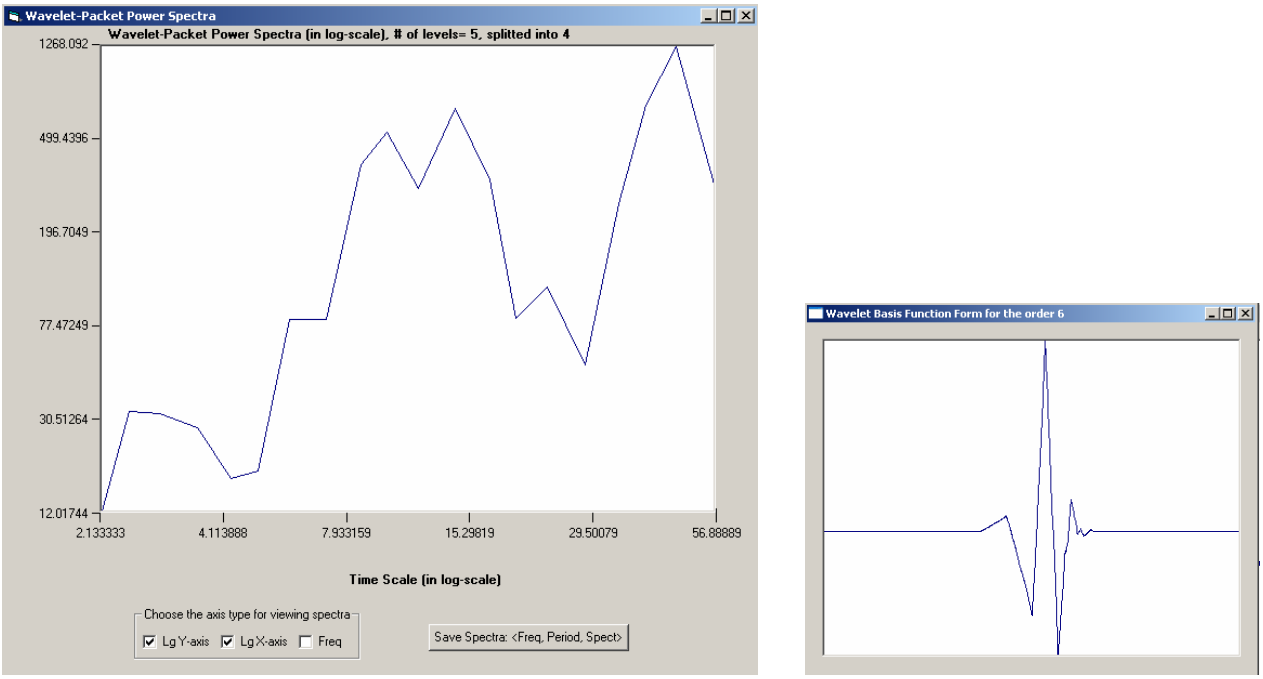

Fig.29.

Let us extract the main impulse of this signal as a fragment:

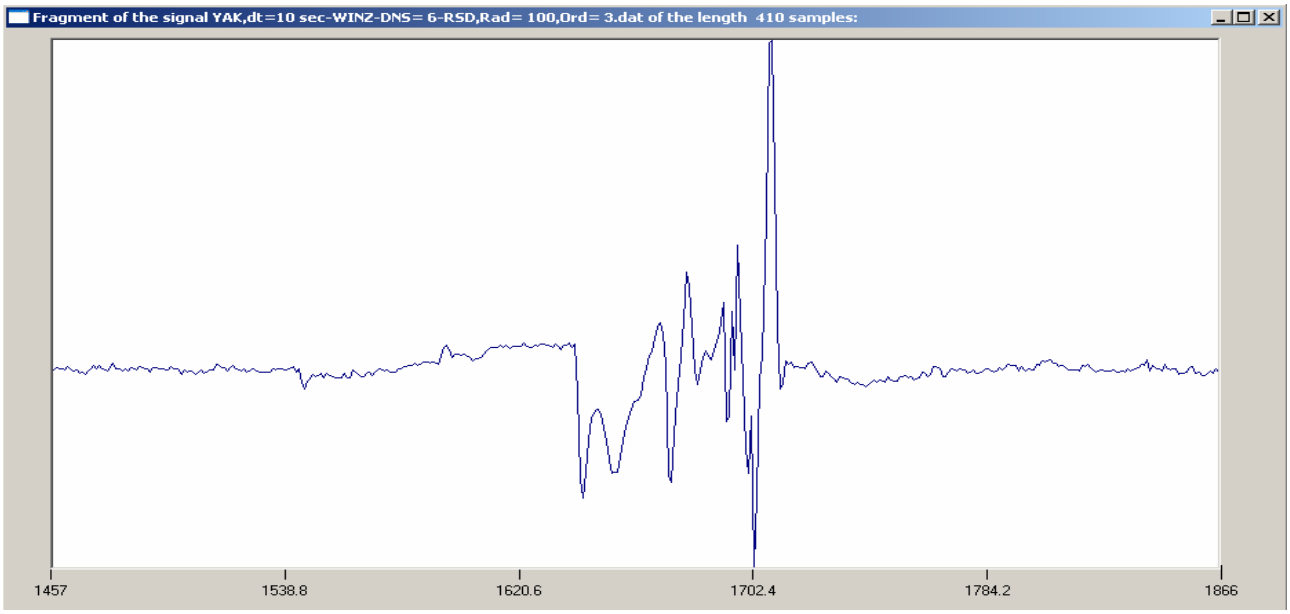

Fig.30.

- and look its Morlet diagram in details:

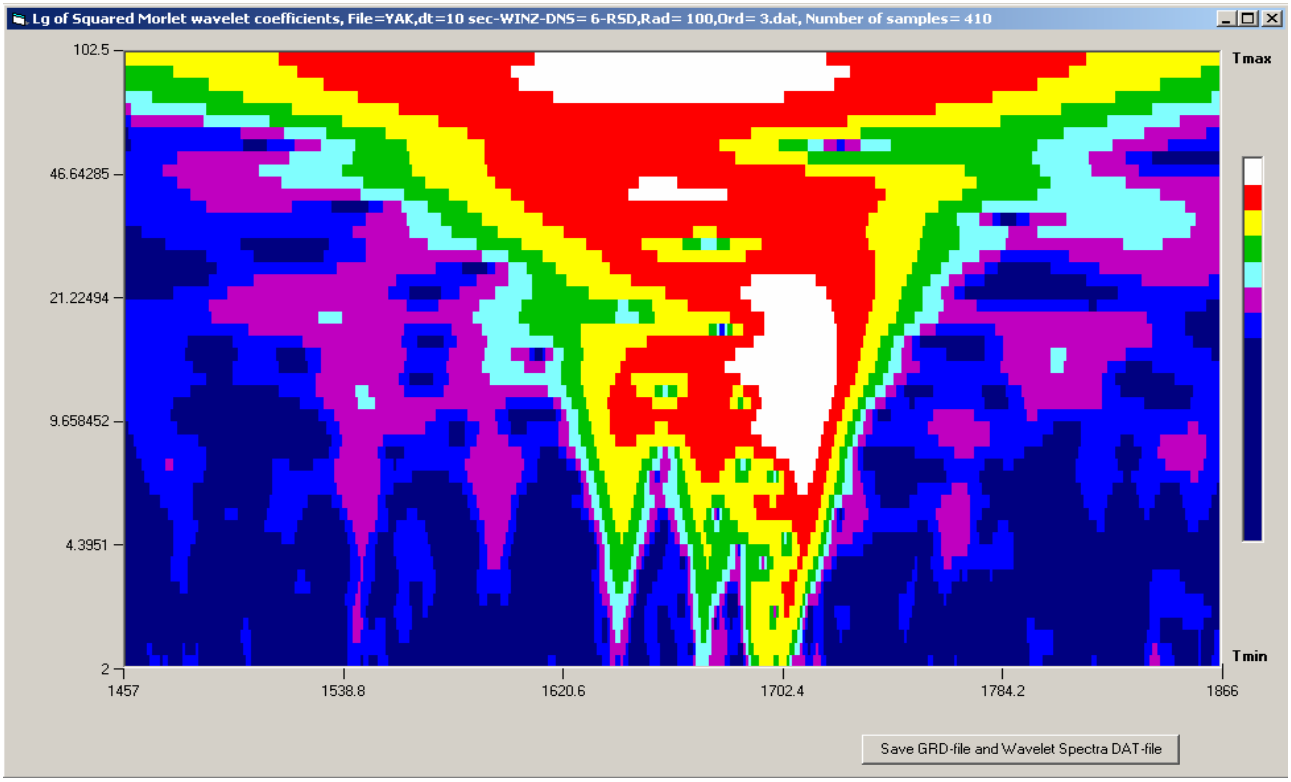

Fig.31.

The structure of long WTMM-skeletons for wavelets as derivative of Gaussian with orders 0 (simple smoothing), 1 and 2 (Mexican hat):

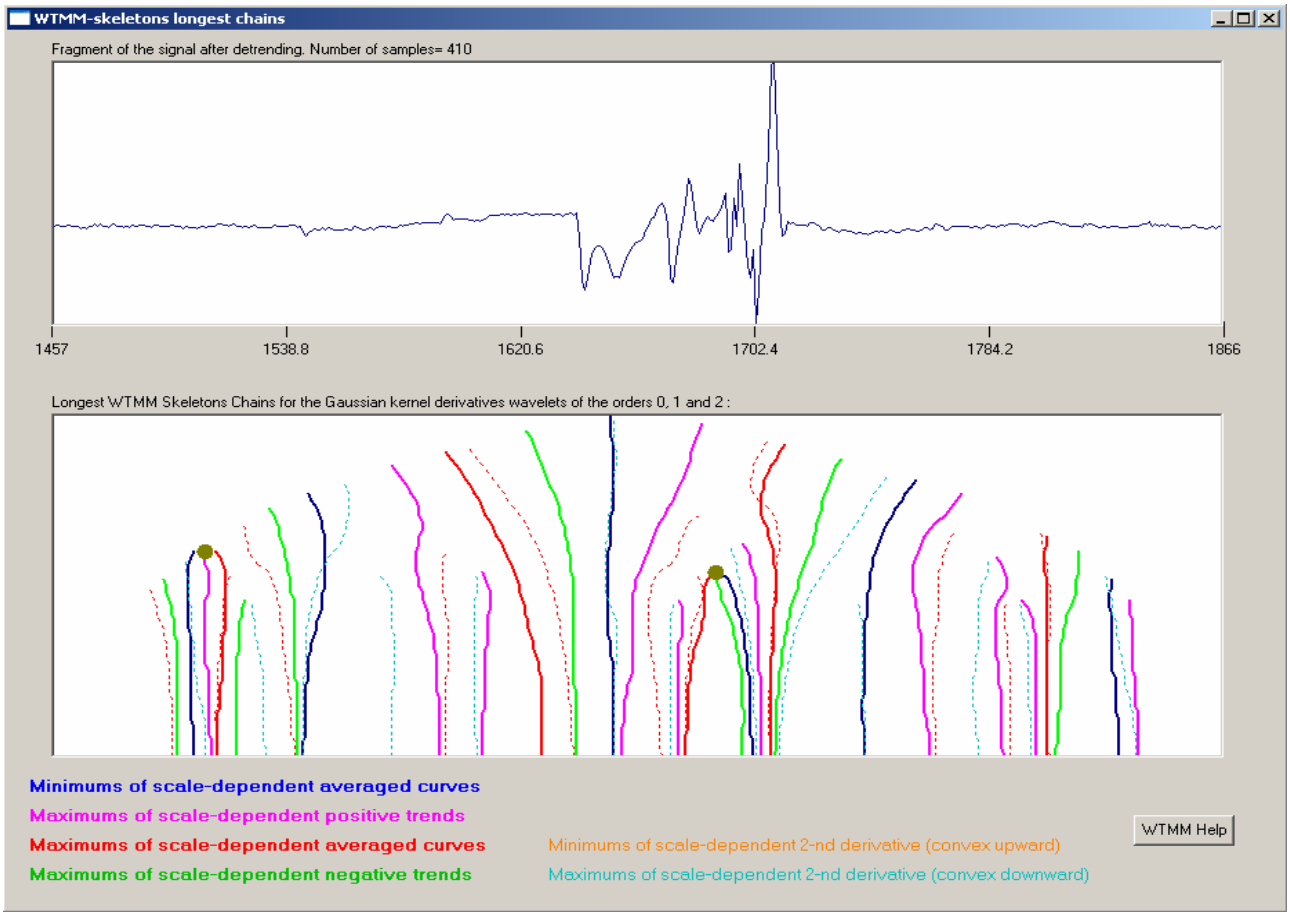

Fig.32.

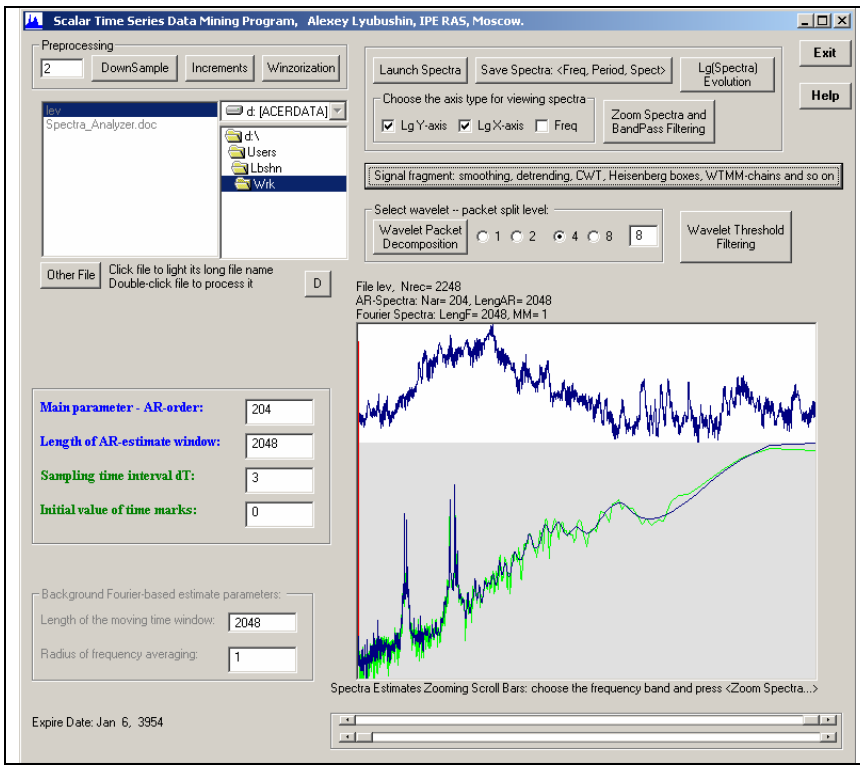

Fig.33.

For illustrating evolution of power spectra it would be better take a signal containing explicit harmonic components, for example variations of underground water well level with 3-hours sampling time interval

←

Its spectral time-frequency map with default parameters values:

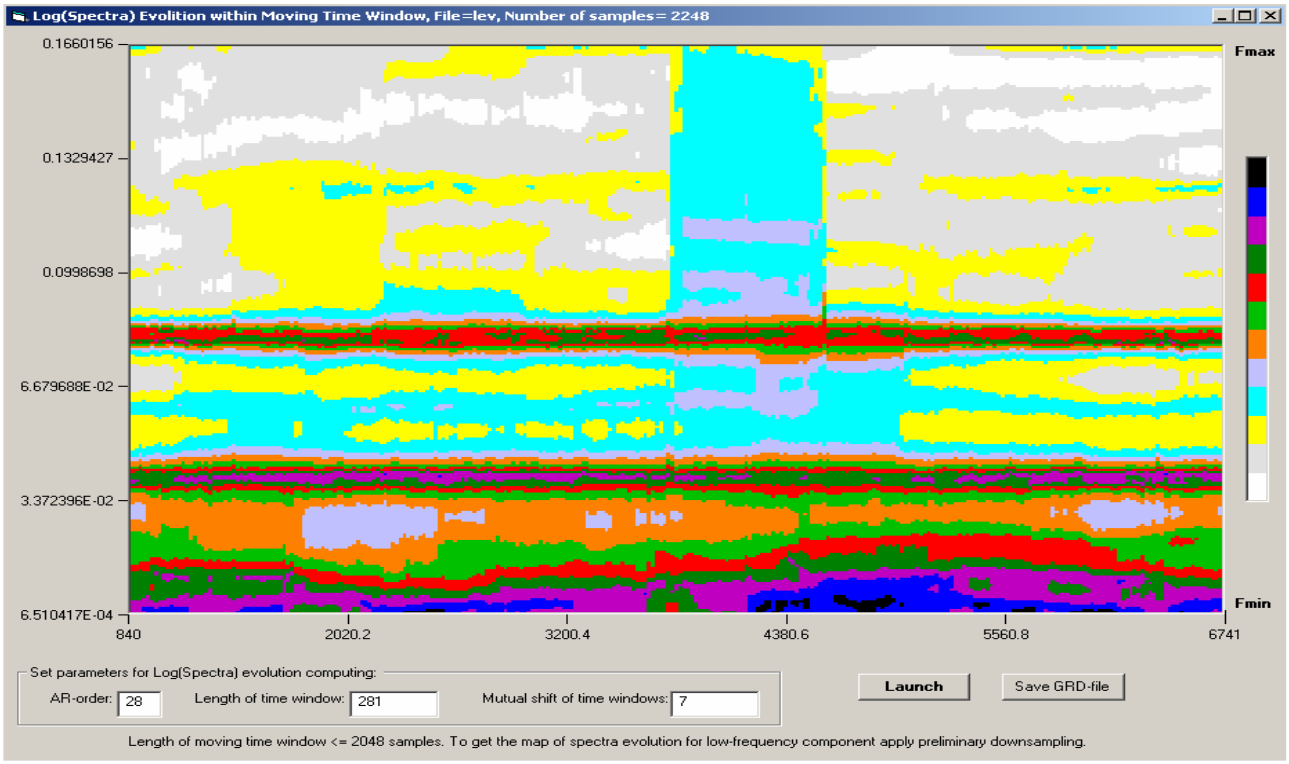

Fig.34.

The next spectral time-frequency map is presented for a very short moving time window:

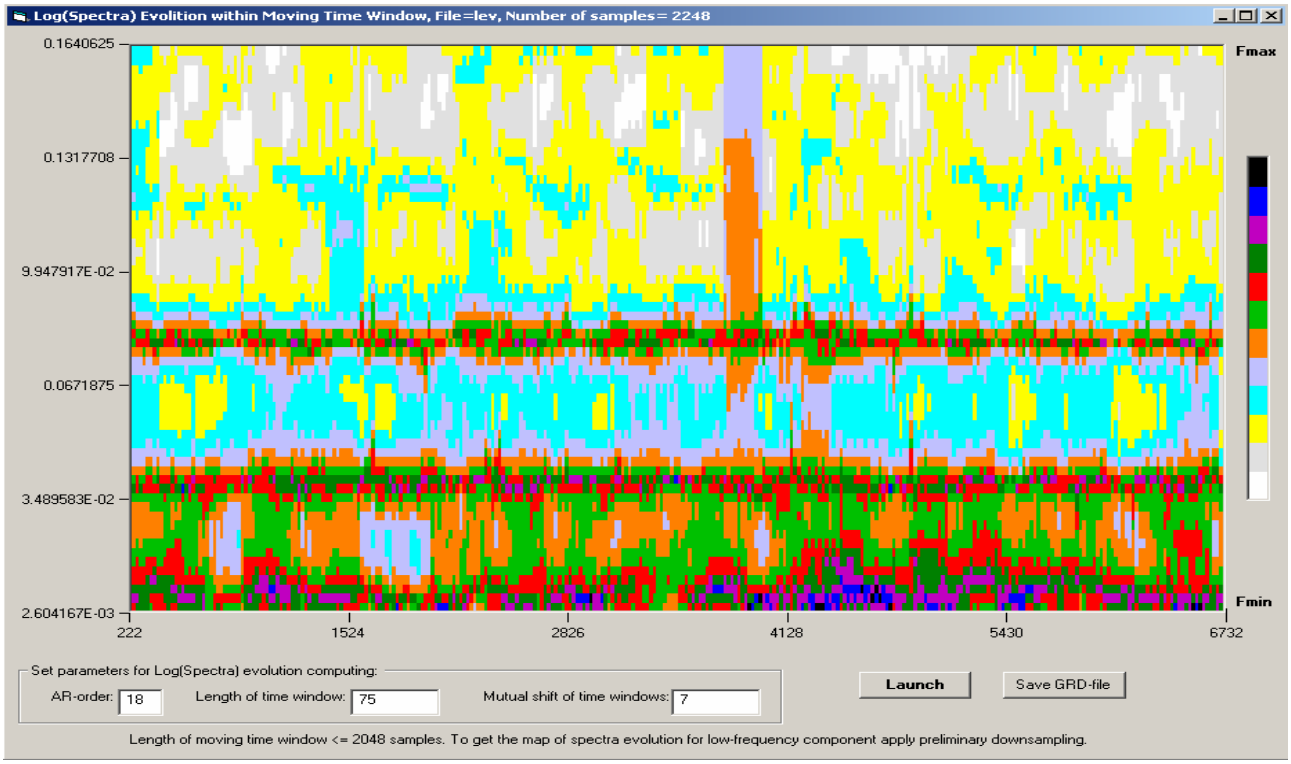

Fig.35.

For comparison – its Morlet wavelet diagram:

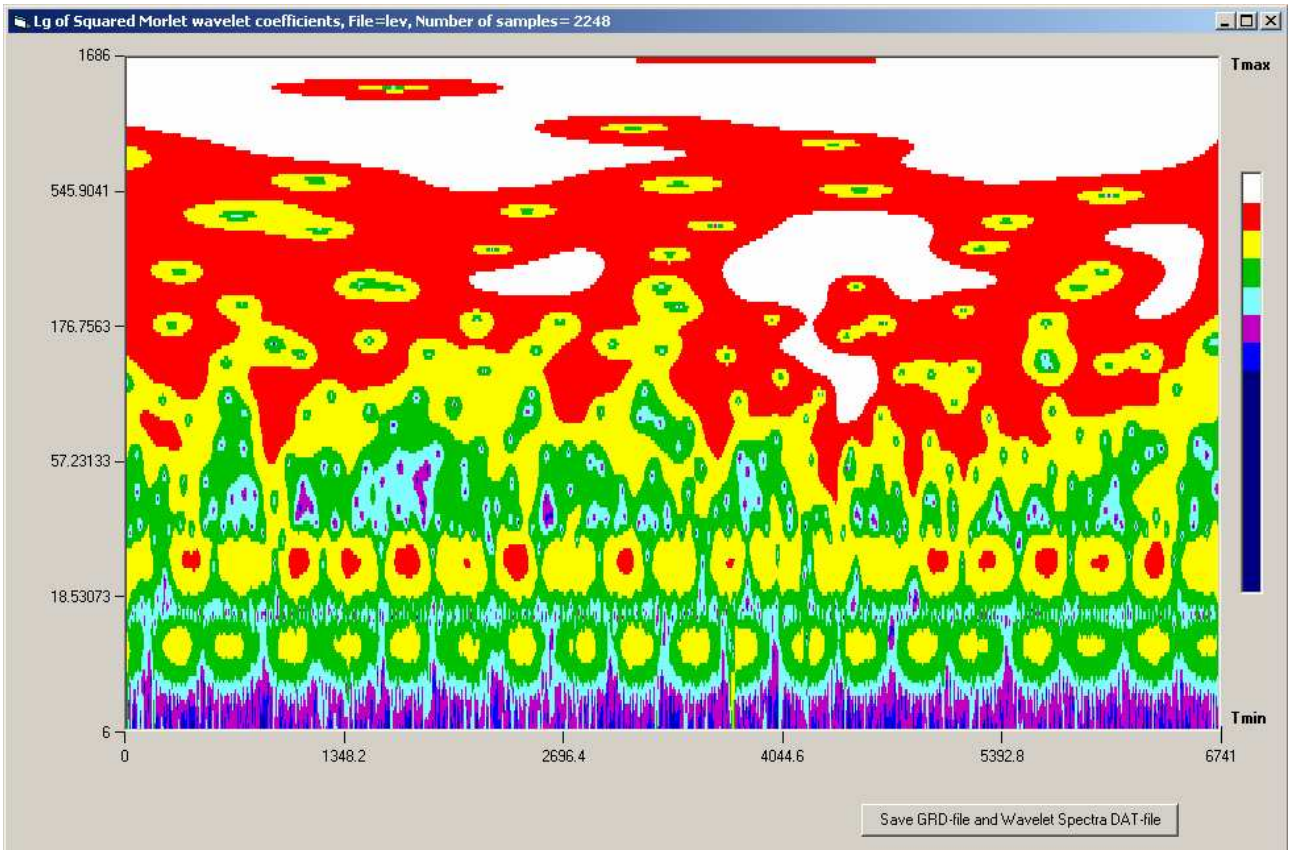

Fig.36.

Finally, the example of determining low-frequency cyclic trend with unknown period for the Global temperature anomaly:

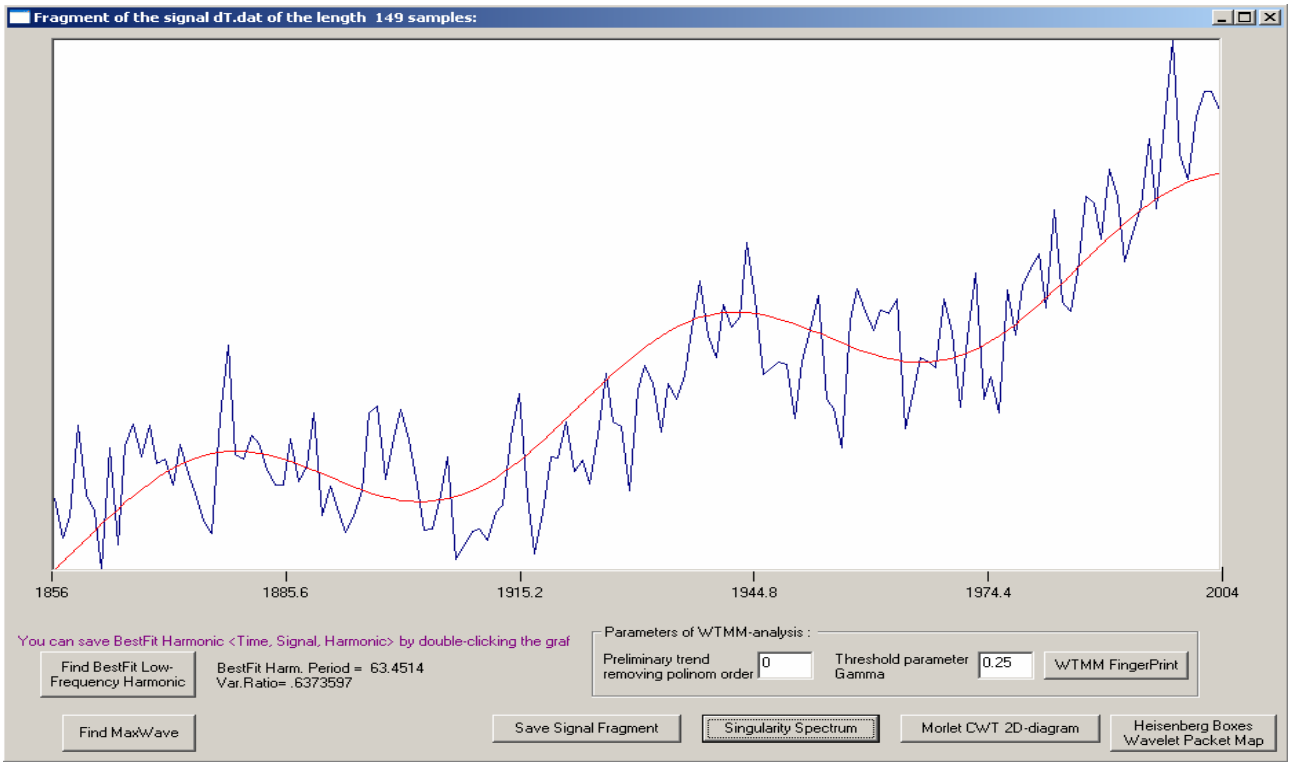

Fig.37.

and automatic detecting of 2 extreme points corresponding to maximum amplitude variation within seismic record:

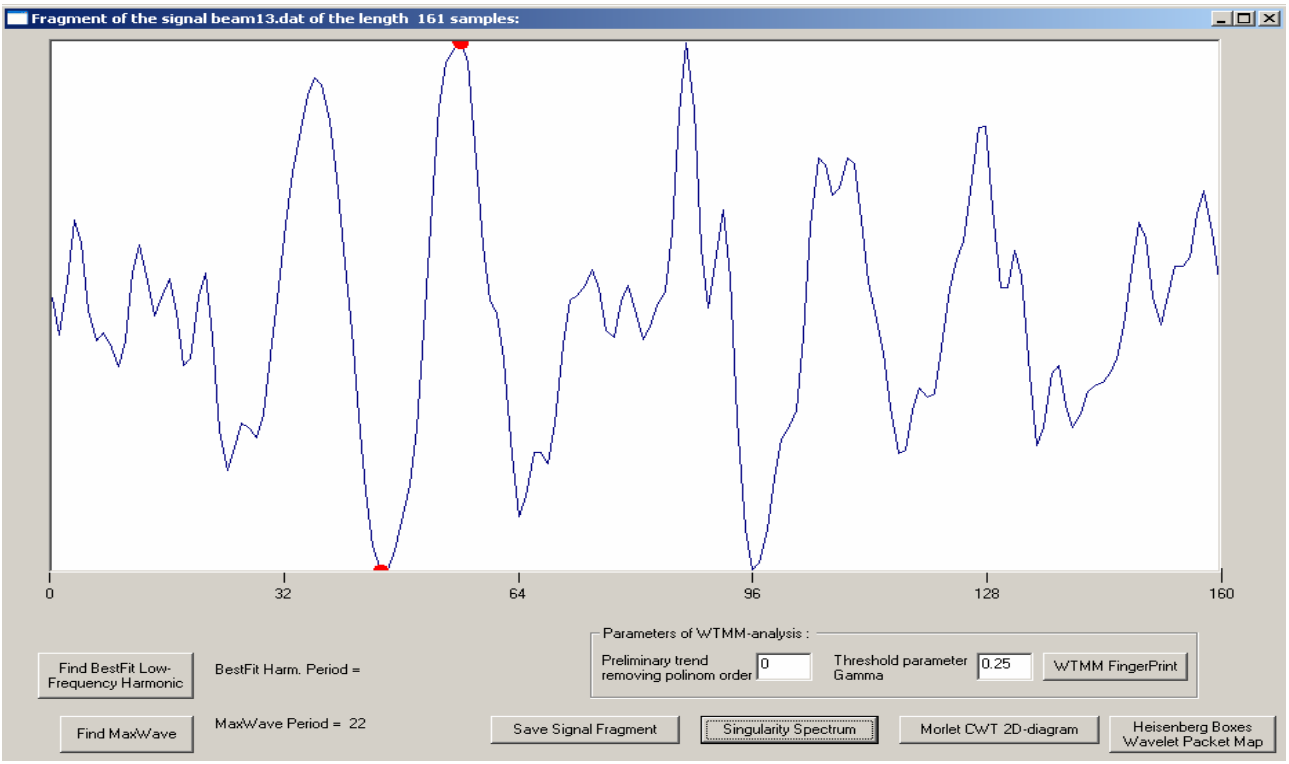

Fig.38.

Estimates of parameters of multifractal singularity spectrum is a powerful tool for understanding the structure of the noise component of geophysical monitoring time series. These estimates could be a solid foundation for earthquake prediction on the basis of continuous monitoring of geophysical fields variations [*Lyubushin, Kopylova, 2004; Lyubushin, Sobolev, 2006 Lyubushin, 2007, 2009, 2010, 2011*].

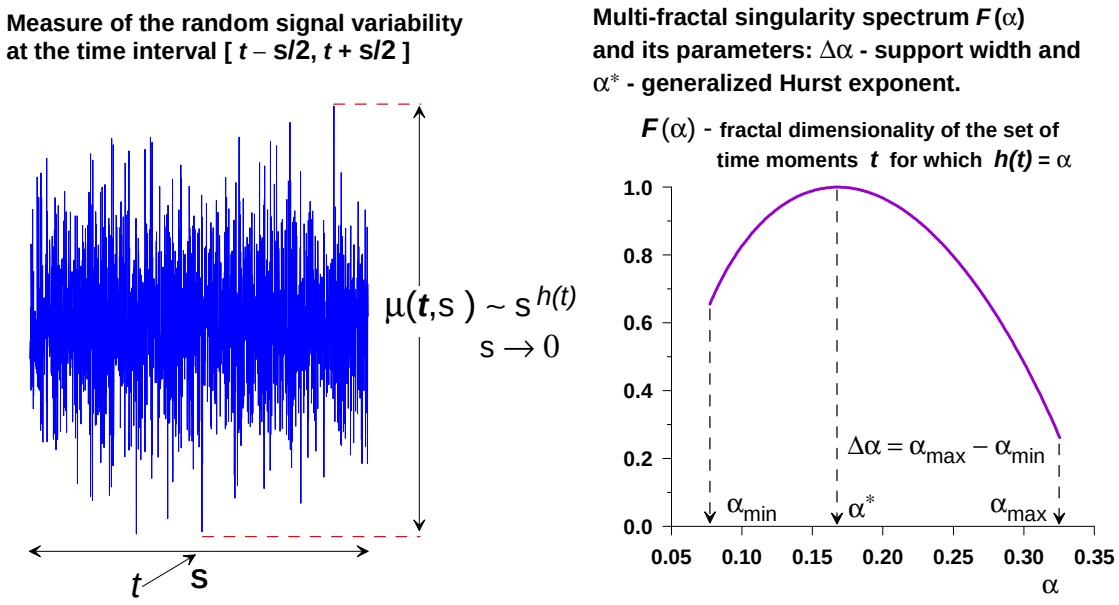

Fig.39. Illustrating of the notion of multifractal singularity spectrum.

For illustrating multifractal singularity spectra estimate let's take a time series of electro-telluric potentials measured on Kamchatka after downsampling it to provide sampling time interval 1 day:

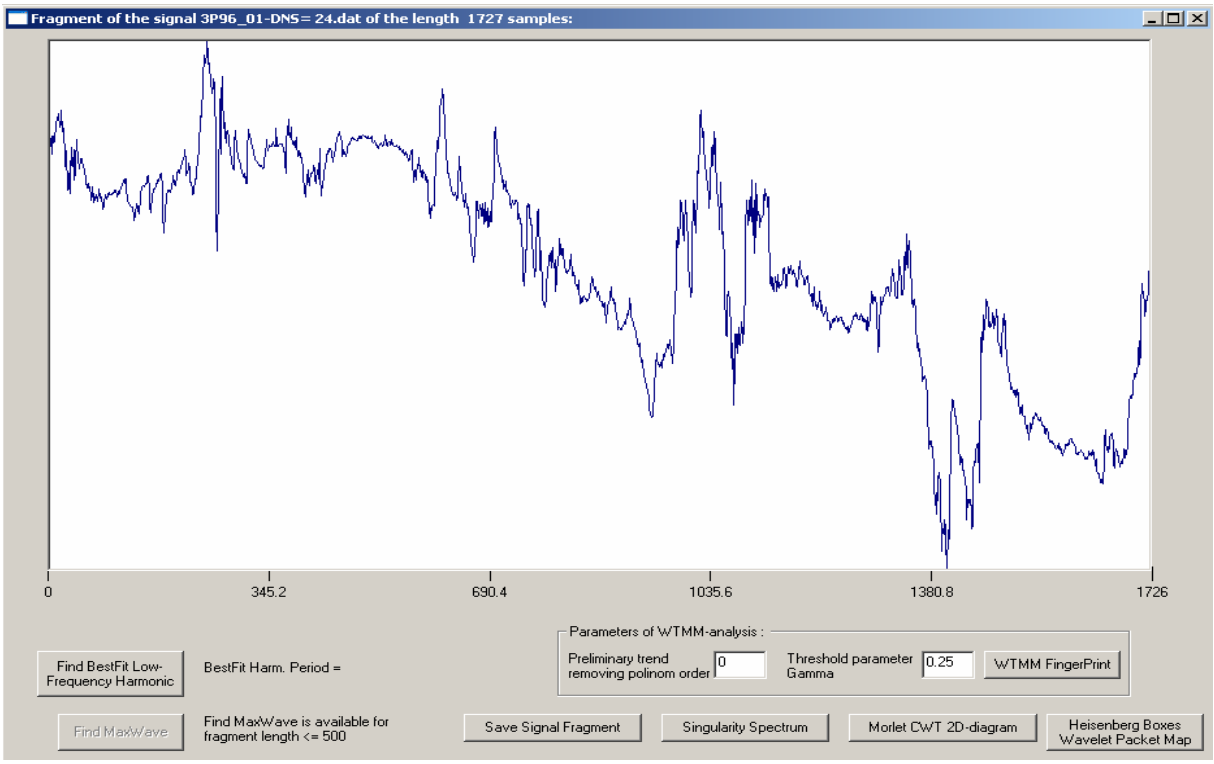

Fig.40. This is the whole sample of the signal. Let's estimate it singularity spectra:

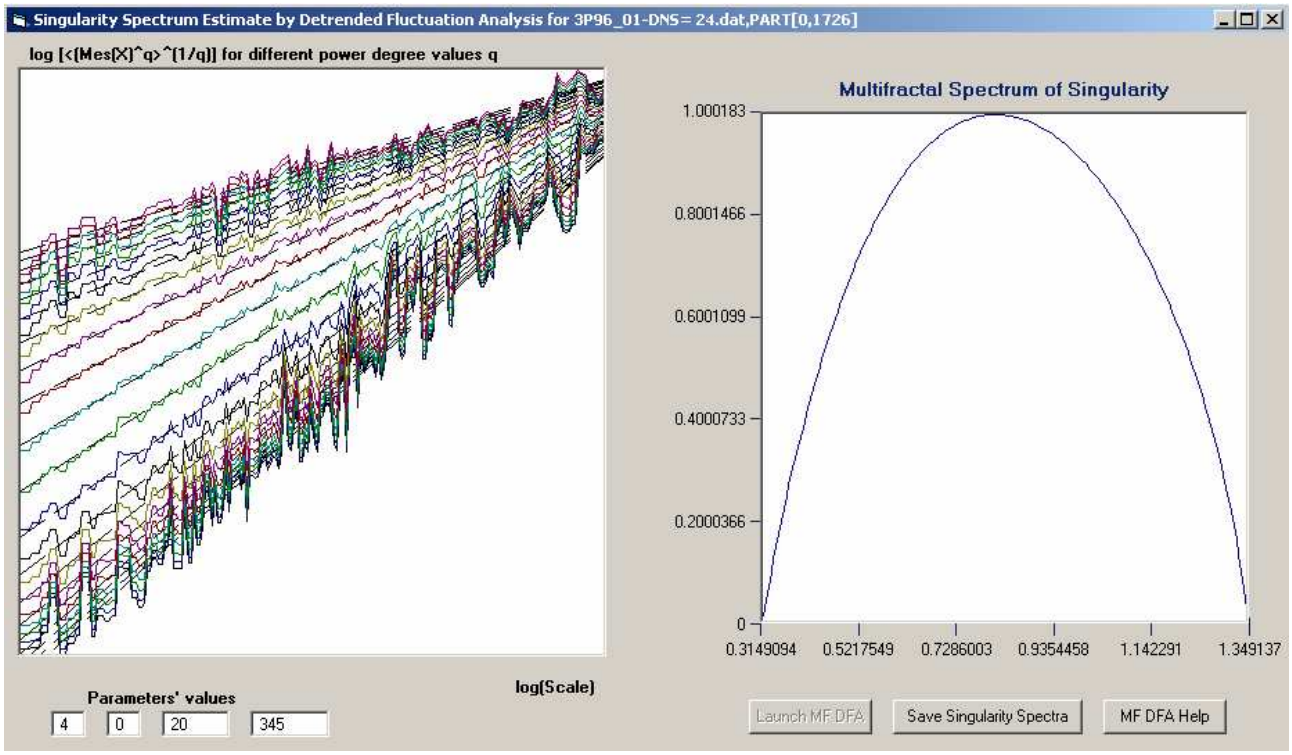

Fig.41.

Left graphic on the Fig.41 presents dependences of $\ln(Z^{(m)}(q,s))$ on $\ln(s)$ (formulae (7.6(a)) or (7.6(b)) – logarithm of mean value of the scale-dependent signal variability measure in the power degree q and subsequent power degree $1/q$, after excluding scale-dependent local trends by polynomials of the order m on logarithm of scale value) for different values of power degree q which varies from -10 up to +10. The slopes coefficients of best fitted linear regression (dashed lines) for these dependences give values of the generalized Hurst exponent $h(q)$. If the signal is monofractal then the linear regression lines to the curves on the left panel on the Fig.41 are approximately parallel each other – it means that $h(q) = H = const$.

Four text boxes under the left panel serve for introducing parameters' values for the procedure of singularity spectra estimate. Which text box serves for introducing which parameter could be recognized through mouse tool-tip after moving the mouse cursor into the corresponding text box.

The most left text box (the 1st one from left to right) is for introducing the order m of orthogonal polynomials for local scale-depended de-trending operation. The possible values: $0 \le m \le 10$, a default value: $m = 4$.

The 2nd text box is for introducing a type of the signal variability measure for further computing $Z^{(m)}(q,s)$ values according to formulae (7.6(a)) or (7.6(b)). Possible type indexes as "0" or "1". If "0" then the measure equals to the difference between maximum and minimum values on the time interval of the length s after excluding local trends (formula (7.6(a)). If "1" then the measure equals to standard deviation (formula (7.6(b)). The default value is "0".

The next 3rd and 4th text boxes are for introducing minimum and maximum values $s_{\min}$ and $s_{\max}$ of possible scale values for defining function $h(q)$ as coefficient of linear regression between $\ln(Z^{(m)}(q,s))$ and $\ln(s)$: $Z^{(m)}(q,s) \sim s^{h(q)}$, $s_{\min} \le s \le s_{\max}$. The default value for $s_{\min}$ is 20. The

default value for $s_{\max}$ equals $L/5$, where L is a number of samples within selected fragment of the signal. The possible values of $s_{\max}$ obey restrictions: $40 \le s_{\max} \le L/5$, whereas values of $s_{\min}$ are under the restrictions: $20 \le s_{\min} \le s_{\max}/2$.

The right graphic window on the Fig.41 presents graphic of the singularity spectra estimate which could be saved on the disc with the filename 'F(Alfa)_*.dat' within subdirectory 'SpectEstimates' of the data directory.

APPENDIX. DESCRIPTION OF THE METHODS.

A1. Analysis of Hidden Periodicities within Sequences of Events.

The method was proposed at the paper [*Lyubushin et al., 1998*] and is intended for detecting periodic components within flow of events.

Let

$$t_i, i = 1, ..., N \tag{1.1}$$

be a sequence of the events' time moments which was observed within time interval $(0, T]$. Let us consider the following model of seismic intensity which has a periodic component:

$$\lambda(t) = \mu \cdot (1 + a \cdot \cos(\omega t + \varphi)) \tag{1.2}$$

where the frequency ω, amplitude $a, 0 \le a \le 1$, phase angle φ, $\varphi \in [0, 2\pi]$ and multipliers $\mu \ge 0$ (which describe a Poissonian part of seismic process intensity) are parameters of the model to be identified. Thus, the Poissonian part of intensity is modulated by harmonic oscillation.

Let us fix some value of the frequency ω. Logarithmic likelihood function [*Cox, Lewis*, 1966] for the set of observations is equal to:

$$\begin{aligned} &\ln L(\mu, a, \varphi \mid \omega) = \sum_{t_i} l\,n(\lambda(t_i)) - \int_0^T \lambda(s)ds = \\ &= N \cdot \ln(\mu) + \sum_{t_i} \ln(1 + a\cos(\omega t_i + \varphi)) - \mu \cdot T - \frac{\mu \cdot a}{\omega}[\sin(\omega T + \varphi) - \sin(\varphi)] \end{aligned} \tag{1.3}$$

Taking maximum value of (1.3) with respect to μ it is easily to find that

$$\hat{\mu} = \hat{\mu}(a, \varphi \mid \omega) = \frac{N}{T + a \cdot (\sin(\omega T + \varphi) - \sin(\varphi))/\omega} \tag{1.4}$$

Substituting (4) into formula (3) we will have:

$$\ln(L(\hat{\mu}, a, \varphi \mid \omega)) = \sum_{t_i} \ln(1 + a\cos(\omega t_i + \varphi)) + N \cdot \ln(\hat{\mu}(a, \varphi \mid \omega)) - N \tag{1.5}$$

It should be noted that $\hat{\mu}(a = 0, \varphi \mid \omega) \equiv \hat{\mu}_0 = N/T$ is the estimate of the uniform Poissonian (pure random) part of intensity.

Thus, the increment of log-likelihood function due to introduction of the harmonic oscillation with given frequency value ω into the model of intensity with respect to zero hypothesis that seismic process is uniform pure random (Poissonian) equals:

$$\Delta \ln L(a,\varphi \mid \omega) = \sum_{t_i \in} \ln(1 + a\cos(\omega t_i + \varphi)) + N \cdot \ln(\hat{\mu}(a,\varphi \mid \omega)/\hat{\mu}_0) \qquad (1.6)$$

Let

$$R(\omega) = \max_{a,\varphi} \Delta \ln L(a,\varphi \mid \omega),\ 0 \le a \le 1,\ \varphi \in [0, 2\pi], \qquad (1.7)$$

The function (1.7) could be regarded as the generalization of the spectra for the sequence of events. The graphic of this function indicates which probe values of the frequency provide the maximum gain in log-likelihood function increment with respect to a pure random model. Thus, the points of maximum of the function (1.7) detect periodic components of the seismic process.

The next generalization of this approach is estimating the function (1.7) not over the whole time interval of observation $(0,T]$ but within moving time window of the certain length T_W. Let τ be a time coordinate of the right-hand end of the moving time window. Then we have the function of 2 arguments: $R(\omega,\tau \mid T_W)$ which could be visualized as 2D map within the plane of (ω,τ)-values. The time-frequency diagrams allow describe the dynamics of periodic component within seismic process.

This time-frequency diagram allows investigate dynamics of occurring and development of periodic components within considered flow of events [*Lyubushin et al. 1998; Sobolev, Lyubushin, 2010, Lyubushin, 2007*].

An important question for application of this method is a statistical significance of the extracted periodicities within intensity of the point process from peak values of the functions $R(\omega)$ and $R(\omega,\tau \mid T_W)$. To solve this problem an asymptotic Wilks theorem could be applied [*Wilks, 1962*]. Let us consider two hypothesis about the distribution of a vector $X^{(N)}$ consisting of N independent observations:

1) hypothesis H_0: $X^{(N)}$ is distributed according to the density $p_0(X^{(N)} \mid \theta_0)$;
2) hypothesis H_1: $X^{(N)}$ is distributed according to the density $p_1(X^{(N)} \mid \theta_1)$.

where θ_0 and θ_1 are vectors of unknown parameters with dimensionality m_0 and m_1 where $m_1 > m_0$ and vector of parameters θ_0 is included entirely into vector θ_1. Let us consider the difference between maximum values of logarithmic likelihood functions:

$$\Delta \ln L(X^{(N)}) = \ln\left(\max_{\theta_1} p_1(X^{(N)} \mid \theta_1)\right) - \ln\left(\max_{\theta_0} p_0(X^{(N)} \mid \theta_0)\right) \qquad (1.8)$$

Just because probability density $p_1(X^{(N)} \mid \theta_1)$ has more parameters than $p_0(X^{(N)} \mid \theta_0)$ it is evident that $\Delta \ln L(X^{(N)}) \ge 0$. According to Wilks theorem if hypothesis H_0 is true then the value (1.8) has an asymptotic distribution:

$$\Delta \ln L(X^{(N)}) \sim \frac{\chi^2_m}{2}, \quad m = m_1 - m_0, \quad N \to \infty \tag{1.9}$$

In our case $m = 2$ i.e. $2 \cdot \Delta \ln L \sim \chi^2_2$. Probability density of χ^2_2 equals $e^{-x/2}/2$, thus, $\Delta \ln L$ has an asymptotic exponential density e^{-x} and its asymptotic probability distribution function equals

$$\Pr\{R(\omega) < x\} = 1 - e^{-x}, \quad N \to \infty \tag{1.10}$$

From formula (1.10) it follows that 90% probability threshold equals 2.3. Thus, if the peak values of (1.7) exceed the threshold 2.3 it means that we can regard these peaks as evidence for existing periodic component with probability not less that 0.9.

A2. Wavelet-analysis of time series. .

Wavelet analysis of geophysical time series became more and more popular at last years as efficient tool of data processing [*Lyubushin, 2000, 2001, 2002; Lyubushin et al., 2004; Lyubushin and Kopylova, 2004; Lyubushin, 2007*].

Orthogonal multi-resolution analysis of the signal $x(t)$ is defined by the formula [*Daubechies, 1992; Mallat, 1998*] :

$$x(s) = \sum_{\alpha=-\infty}^{+\infty} x^{(\alpha)}(s), \quad x^{(\alpha)}(s) = \sum_{j=-\infty}^{+\infty} b^{(\alpha)}(\tau_j^{(\alpha)}) \cdot \psi^{(\alpha)}(s - \tau_j^{(\alpha)}) \ , \quad \tau_j^{(\alpha)} = j \cdot 2^{\alpha} \tag{2.1}$$

Here α is a detail level number,

$$b_j^{(\alpha)} = b^{(\alpha)}(\tau_j^{(\alpha)}) = \int_{-\infty}^{+\infty} x(s) \cdot \psi^{(\alpha)}(s - \tau_j^{(\alpha)}) ds \tag{2.2}$$

are wavelet coefficients on the α-th detail level, corresponding to the time moment $\tau_j^{(\alpha)}$, $\psi^{(\alpha)}(t)$ are basis functions of the α-th level, which are obtained by dilation and translation of the *mother wavelet function* $\Psi(t)$

$$\psi^{(\alpha)}(s) = (\sqrt{2})^{-\alpha} \cdot \Psi(2^{-\alpha} \cdot s), \quad \psi^{(\alpha)}(s - \tau_j^{(\alpha)}) = (\sqrt{2})^{-\alpha} \cdot \Psi(2^{-\alpha} \cdot s - j) \tag{2.3}$$

The function $\Psi(t)$ is constructed in such a way that it is finite supported, has a unit norm in $L_2(-\infty, +\infty)$ and an infinite set of functions $\{\psi^{(\alpha)}(t - \tau_j^{(\alpha)})\}$, which are the copies of the main function translated into time moments $\tau_j^{(\alpha)}$ and dilated into 2^{α} times, is an orthonormal basis in $L_2(-\infty, +\infty)$. For instance if:

$$\Psi(t) = \begin{cases} -1, & t \in (0, \frac{1}{2}] \\ +1, & t \in (\frac{1}{2}, 1] \\ 0 \ \text{for all other } t \end{cases} \tag{2.4}$$

.

formula (2.1) provides the expansion of $x(t)$ in Haar wavelets.

The most popular family of orthogonal wavelets are Daubechies wavelet functions $\Psi(t) = D_{2p}(t)$ of the order $2p$, which possess the following properties:

$$D_{2p}(t) = 0 \text{ outside interval } [-p+1, p], \tag{2.5a}$$

$$\int_{-\infty}^{+\infty} t^k \cdot D_{2p}(t)\, dt = 0 \quad \text{for } k = 0,1,...,(p-1) \tag{2.5b}$$

Note that Haar wavelet is a Daubechies wavelet of 2nd ($p = 1$).

With an increase in the number p of vanishing moments in formula (2.5b), the function $D_{2p}(s)$ becomes smoother, although the number of its continuous derivatives is not proportional to the parameter p. For example, the Daubechies function of the 4th order $D_4(s)$ sets to zero the zero-th and first moments and is continuously differentiable at all points except a countable set of points of the form $k \cdot 2^{-l}$, where k, l are integer numbers. At such points, $D_4(s)$ has a left-hand derivative and does not have a right-hand derivative. Note that Haar wavelet (2.4) is a second-order Daubechies wavelet ($p = 1$). We used a dictionary of 17 wavelets: 10 usual orthogonal Daubechies wavelets of orders from 2 to 20 (the use of higher orders entails numerical instability) and 7 so-called "symlets," which are Daubechies wavelets whose basis functions are more symmetric compared to ordinary wavelets [*Chui, 1992; Daubechies, 1992; Mallat, 1998*]. Symlets possess the same properties of compactness, orthogonality, completeness, and smoothness as wavelets (2.5); however, for orders of 2 to 6, they coincide with the ordinary orthogonal Daubechies basis, while, for orders of 8 to 20, some distinctions appear in the form of the basis function. As a consequence, the total number of variants of orthogonal compact basis functions used here is 17.

Now, we address the situation when $z(t)$ is a signal discrete in time t having a length of N measured values, $t = 1,..., N$. We assume that N is an integer of the form 2^m because this is convenient for the subsequent application of the fast wavelet transform. If N is not equal to 2^m, the signal $z(t)$ is complemented by zeros to make it 2^m long, where m is the minimum integer such that $N \le 2^m$.

The formula of the multiple-resolution analysis in the case of a finite sample and discrete time is

$$z(t) = a_1^{(m)} + \sum_{\beta=1}^{m} z^{(\beta)}(t), \quad z^{(\beta)}(t) = \sum_{j=1}^{2^{(m-\beta)}} c^{(\beta)}(\tau_j^{(\beta)}) \cdot \psi^{(\beta)}(t - \tau_j^{(\beta)}), \quad \tau_j^{(\beta)} = j \cdot 2^{\beta} \tag{2.6}$$

where $z^{(\beta)}(t)$ is the component of the signal belonging to the detail level of the number β and $a_1^{(m)}$ is a constant proportional to the mean of the sample [*Chui, 1992; Daubechies, 1992; Mallat, 1998; Press et al., 1996*]. The coefficients $c_j^{(\beta)} = c^{(\beta)}(\tau_j^{(\beta)})$ in (2.6) can be represented, similarly to formula (2.2), as the convolution of the basis function $\psi^{(\beta)}(s)$ of the continuous argument s with a certain signal $\tilde{z}(s)$:

$$c_j^{(\beta)} = \int_{-\infty}^{+\infty} \tilde{z}(s) \cdot \psi^{(\beta)}(s - \tau_j^{(\beta)})\, ds \tag{2.7}$$

The signal $\tilde{z}(s)$ is obtained from the signal $z(t)$ with the discrete time t through the interpolation formula:

$$\tilde{z}(s) = \sum_t z(t) \cdot \Phi(s-t) \tag{2.8}$$

where the function $\Phi(s)$ is called the scaling function of the wavelet expansion. For example, in the case of Haar wavelet (2.4), $\Phi(s) = 1$ if $s \in [0,1)$ and $\Phi(s) = 0$ for all other s; consequently, the interpolated signal will be a piecewise-continuous function. In the general case of orthogonal Daubechies wavelets, the scaling function is orthogonal to the principal basis function $\Psi(s)$, $\int_{-\infty}^{+\infty} \Psi(s)\Phi(s)\,ds = 0$, and has the same properties of smoothness and a compact support of the same length as $\Psi(s)$ (but not coinciding with it completely): $\Phi(s) = 0$ outside the interval $[0, 2p-1)$.

If the discrete signal $z(t)$ is obtained from a signal continuous in time $x(s)$ that is measured at a time step Δs, then, if $\Delta s \to 0$, the interpolated signal $\tilde{z}(s)$ always tends in the mean-square metric L_2 to the initial signal $x(s)$. If the scaling function $\Phi(s)$ corresponding to the Daubechies basis of the order $2p$ is used in formula (2.5), the interpolated signal $\tilde{z}(s)$ will have $(p-1)$ first derivatives continuous almost everywhere, with the possible exception of a countable set of points, regardless of the smoothness of the initial signal $x(s)$. However, if $\Delta s \to 0$, these derivatives will tend in the integral metric L_2 to the derivatives of the initial signal only if $x(s)$ is differentiable almost everywhere also $(p-1)$ times. Thus, the choice of the wavelet for the signal analysis must correspond to its smoothness.

Regardless of the possible origin of the signal $z(t)$ obtained by the discretization of the continuous signal $x(s)$ with a certain time step Δs, the coefficients of discrete expansion (2.6) are a result of the application of a successive linear filtering of the discrete signal. At the first step, the discrete signal is split into two parts: the wavelet coefficients of the first detail level $c_j^{(1)}$ (or the "detail signal" of the first level) and the so-called approximating (smoothed) signal $a_j^{(1)}$, using the formula

$$c_j^{(1)} = \sum_t g(t-2j) \cdot a_t^{(0)}, \quad a_j^{(1)} = \sum_t h(t-2j) \cdot a_t^{(0)}, \quad a_t^{(0)} \equiv z(t), \quad j = 1, \ldots, N/2 \tag{2.9}$$

The coefficients of the linear filter $g(k)$ in (2.9) possess the property of detecting high frequencies, and the coefficients $h(k)$ possess the property of smoothing. Note that formulas (2.9) involve not only the linear filtration but also a twofold down-sampling procedure; therefore, the detail and approximating signals contain half as many samples as the initial signal. Because of a finite length of the sample, technical difficulties are associated with the application of formulas (2.9) to the beginning and the end of the sample. These difficulties can be overcome by various methods such as, for example, the consideration of the sample $z(t)$ on a ring rather than in an interval. In this case, edge distortions of the wavelet filtering results similar to those produced by the cyclic effect of the discrete Fourier transform [*Press et al., 1996*] can arise. Step (2.9) is repeated $(m-1)$ times (recall that $N = 2^m$):

$$c_j^{(\beta+1)} = \sum_t g(t-2j) \cdot a_t^{(\beta)}, \quad a_j^{(\beta+1)} = \sum_t h(t-2j) \cdot a_t^{(\beta)}, \quad j = 1, \ldots, N \cdot 2^{-(\beta+1)} \tag{2.10}$$

According to formula (2.10), at each new detail level of the wavelet expansion, the approximating signal $a_j^{(\beta)}$ of the preceding detail level is split into its high-frequency component $c_j^{(\beta+1)}$ and the progressively more smoothed signal $a_j^{(\beta+1)}$. The number of components in the detail signal (i.e., the number of wavelet coefficients) and in the smoothed (approximating) signal decreases by two times as the number of the detail level increases by unity. The coefficient $a_1^{(m)}$ in (2.6) is the approximating "signal" corresponding to the deepest smoothing at the final detail level m. The coefficients $g(k)$ and $h(k)$ of the linear filters (called the conjugated mirror filters) are interrelated as $g(k) = (-1)^{1-k} h(1-k)$ in accordance with the scaling equations of the orthogonal multi-resolution analysis

$$\frac{1}{\sqrt{2}}\Phi(t/2) = \sum_{k=-\infty}^{+\infty} h(k)\cdot\Phi(t-k), \quad \frac{1}{\sqrt{2}}\Psi(t/2) = \sum_{k=-\infty}^{+\infty} g(k)\cdot\Phi(t-k) \tag{2.11}$$

for the scaling and mother basis functions. For the Daubechies finite basis functions $D_{2p}(s)$, the number of non-vanishing coefficients in the linear mirror filters $g(k)$ and $h(k)$ is equal to the order of the function $2p$. For example, for the Haar wavelet, $h(k) = 1/\sqrt{2}$ for $k = 0,1$ and $h(k) = 0$ for all other k. For the Daubechies functions of orders 4 and 6, the coefficients of mirror filters are determined analytically from linear equations following from condition (2.5b), vanishing a given number of first moments; however, for higher orders, these linear equations are solved numerically. We should note that, as the wavelet order increases, the conditioning of the linear equations deteriorates and the round off error increases. Therefore, calculations are usually restricted to wavelets of an order not exceeding 20.

Algorithmically, the wavelet transform is a linear orthogonal transform of the N-dimensional vector of a sample $z(t)$ into the vector of coefficients $C_z^{(N)} = (a_1^{(m)}, c_1^{(m)}, ..., (c_j^{(\beta)}, j = 1, ..., n_\beta), ...)$, whose length is also equal to N. The vector consists of the constant $a_1^{(m)}$ as a first component and the successively arranged coefficients of all detail levels, beginning from the m-th, $(m-1)$-th, and so on, up to the coefficients of the first detail level, which occupy the entire second half of the vector $C_z^{(N)}$ [*Press et al., 1996*]. The inverse transform of the vector $C_z^{(N)}$ yields the initial sample $z(t)$. Note that the inverse wavelet transformation is also effected in the form of consecutive steps reconstructing the approximating signal of the level β from the wavelet coefficients $c_k^{(\beta+1)}$ and the approximating signal $a_k^{(\beta+1)}$ of the level $(\beta+1)$ with the use of mirror-conjugated filters by the formula:

$$a_j^{(\beta)} = \sum_k h(j-2k)\cdot a_k^{(\beta+1)} + \sum_k g(j-2k)\cdot c_k^{(\beta+1)}, \quad j = 1, ..., N\cdot 2^{-\beta} \tag{2.12}$$

The inverse transform starts with the coefficients $a_1^{(m)}$ and $c_1^{(m)}$ of the final (lowest-frequency) detail level of the number m and ends at the first level. In this case, the number of components in the approximating signal doubles each time, when the number of the detail level decreases until the sequence of reconstructions stops at the first detail level and the number of components becomes equal to N. The forward and inverse wavelet transforms (2.11) and (2.12) can be numerically implemented in the form of fast algorithms requiring $O(N)$ operations and a shorter computation time compared to the fast Fourier transform [*Chui, 1992; Daubechies, 1992; Mallat, 1998; Press et al., 1996*]. The component $z^{(\beta)}(t)$ is a result of the inverse transform of the vector of coefficients of

the $C_z^{(N)}$ type provided that all coefficients, except those corresponding to the level β, are set at zero.

At a sufficiently large value of N, the component $z^{(\beta)}(t)$ is localized within the frequency band:

$$[\Omega_{\min}^{(\beta)}, \Omega_{\max}^{(\beta)}] = [1/(2^{(\beta+1)}\Delta s), 1/(2^{\beta}\Delta s)] \tag{2.13}$$

where Δs is the length of the sampling interval. The value of the coefficient $c_j^{(\beta)}$ reflects the behavior of the signal $z(t)$ in the vicinity of the point $\tau_j^{(\beta)}$ in the interval of a length equal to $p \cdot 2^{\beta}$ values. Consequently, the smoother the wavelet, the wider this interval. However, the main variations (bursts) in the finite basis function $\psi^{(\beta)}(s)$ are always concentrated in the interval 2^{β} long, regardless of the smoothness parameter p. Therefore, we associate each coefficient $c_j^{(\beta)}$ with "a temporal zone of responsibility" of length $\Delta T^{(\beta)} = \Delta s \cdot 2^{\beta}$. The product of the width $\Delta\Omega^{(\beta)} = 1/(2^{\beta+1}\Delta s)$ of frequency range (2.13) by the length of the time interval $\Delta T^{(\beta)}$ gives the area of the so-called "Heisenberg box" on the frequency–time plane; this area is equal to 1/2, regardless of the detail level under consideration.

The smallest scale detail level in formula (2.6) is the first level, and the total number of detail levels m depends on the sample length. The values $c_j^{(\beta)}$ and $a_1^{(m)}$ are calculated by using the fast wavelet transform [*Daubechies, 1992; Mallat, 1998; Press et al., 1996*]. These values uniquely determine the initial sample $z(t)$, which can be reconstructed from given values of $c_j^{(\beta)}$ and $a_1^{(m)}$ by the inverse fast wavelet transform. The detail level can be associated with the number of the frequency (the frequency discrete) in the classical discrete Fourier transform. The wavelet expansion differs from the Fourier analysis by a substantially rarer (uniform on the logarithmic scale) set of "wavelet frequencies." This is the price for such an important property as the compactness of basis functions, which is absent in the Fourier expansion and allows one to locate short-lived anomalies (bursts) much more accurately. Moreover, the compactness of basis functions makes the wavelet analysis applicable to non-stationary and non-Gaussian time series, the Fourier analysis of which, although formally possible, is inefficient.

Although the ordinary wavelet expansion possesses such a useful property as a high-accuracy localization in time of non-stationary signals, the reverse side of this property is a poor resolution in frequency, in accordance with the Heisenberg principle. The wavelet-packet expansion partially eliminates this drawback at the expense of a certain decrease in the time resolution. The realization of the packet splitting is based on the hierarchical scheme of successive wavelet transforms of the initial coefficients $c_j^{(\beta)}$. The orthogonal wavelet-packet expansion of the signal can be, by analogy with formula (2.6), written as the sum:

$$z(t) = a_1^{(m)} + \sum_{\beta=m_q+1}^{m} z^{(\beta)}(t) + \sum_{\beta=1}^{m_q}\sum_{\gamma=1}^{q} z^{(\beta,\gamma)}(t) \tag{2.14}$$

The quantity q can be equal to 2, 4, 8, …, i.e., $q = 2^r$, $r = 1, 2, 3, \ldots$, and it controls the number of sublevels into which an ordinary detail level is split. For a given value of q, the maximum number $m_q < m$ of the detail level β that can be split is determined from the condition that this

level must contain at least q wavelet coefficients. The components $z^{(\beta,\gamma)}(t)$ are frequency-ordered and split frequency band (2.13) of detail level β into q equal parts. Thus, the signal $z^{(\beta,\gamma)}(t)$ is localized in frequency within this band:

$$[\Omega_{\min}^{(\beta,\gamma)}, \Omega_{\max}^{(\beta,\gamma)}], \quad \Omega_{\min}^{(\beta,\gamma)} = \Omega_{\min}^{(\beta)} + (\gamma - 1) \cdot \Delta\Omega^{(\beta)}, \quad \gamma = 1, ..., q; \\ \Omega_{\max}^{(\beta,\gamma)} = \Omega_{\min}^{(\beta,\gamma)} + \Delta\Omega^{(\beta)}, \Delta\Omega^{(\beta)} = (\Omega_{\max}^{(\beta)} - \Omega_{\min}^{(\beta)}) / q \tag{2.15}$$

If n_β ordinary wavelet coefficients correspond to the detail level with the number β, n_β / q wavelet-packet coefficients $c_j^{(\beta,\gamma)}$, $j = 1, ..., n_\beta / q$ correspond to each sublevel γ of the packet expansion. The Heisenberg boxes for these coefficients have a time length q times larger compared to the initial coefficients $c_j^{(\beta)}$, but their frequency side is q times smaller (consequently, the area of the Heisenberg boxes remains unchanged and equal to 1/2).

To obtain the component $z^{(\beta,\gamma)}(t)$, it is necessary to perform successive inverse wavelet transforms starting with the coefficients $c_j^{(\beta,\gamma)}$. The final inverse transform is applied to a certain set of coefficients $w_j^{(\beta,\gamma)}$, $j = 1, ..., n_\beta$ that occupy in an analogue of the vector $C_z^{(N)}$ the same positions as the ordinary coefficients $c_j^{(\beta)}$ of the expansion of the signal $z(t)$. However, the inverse transformation with these coefficients yields the component $z^{(\beta,\gamma)}(t)$ rather than $z^{(\beta)}(t)$. The coefficients $w_j^{(\beta,\gamma)}$ are called modified wavelet-packet coefficients of the signal $z(t)$. Thus, the modified coefficients $w_j^{(\beta,\gamma)}$ are the ordinary wavelet coefficients at the detail level β for the case when only the component $z^{(\beta,\gamma)}(t)$ of the initial signal rather than the signal itself is fed to the input of the forward transformation. The Heisenberg boxes for the coefficients $w_j^{(\beta,\gamma)}$ have the same time length as for $c_j^{(\beta)}$; however, their frequency length is equal to that of the coefficients $c_j^{(\beta,\gamma)}$. Thus, the transition from the ordinary wavelet-packet coefficients $c_j^{(\beta,\gamma)}$ to the modified coefficients $w_j^{(\beta,\gamma)}$ is a procedure narrowing the uncertainty time interval (decreasing the time length of the Heisenberg box). Note that such a transition preserves the energy of the expansion coefficients: $\sum_{j=1}^{n_\beta} (w_j^{(\beta,\gamma)})^2 = \sum_{j=1}^{n_\beta / q} (c_j^{(\beta,\gamma)})^2$.

If a two-dimensional map consisting of the frequency-ordered Heisenberg boxes of the modified coefficients $w_j^{(\beta,\gamma)}$ is constructed in the time–frequency plane and each box is painted in accordance with a palette proportional to the absolute values of the wavelet-packet coefficients (or to their logarithms), we will obtain a diagram visualizing the temporal behavior of the main time scales (or periods) of a non-stationary signal. In its outward appearance, this mosaic of wavelet-package Heisenberg boxes may be inferior in aesthetic perception, for example, to the traditional spectral–temporal diagrams. However, it provides a more accurate and adequate idea of the frequency–time dynamics of an essentially non-stationary signal consisting of a set of short-lived bursts of various scales that can differ significantly in shape from a harmonic oscillation.

A3. Optimal choice of wavelet basis and threshold filtering.

Let $c_j^{(k)}$ be the wavelet coefficients of the analyzed signal $x(t)$, ($t = 1, ..., N$ is the discrete time) expanded in a system of orthogonal finite basis functions. The superscript k is the number of the

detail level of the wavelet expansion, and the subscript j indicates the center of the time vicinity. The greatest possible value m of the detail level number depends on the volume of the sample analyzed. Here, we used a dictionary of 17 wavelets: 10 Daubechies ordinary orthogonal wavelets ranging in order from 2 to 20 (the use of higher orders entails numerical instability) and 7 so-called "symlets"; the latter are modifications of the Daubechies wavelets in which the form of basis functions is more symmetric than in ordinary wavelets [*Chui, 1992; Daubechies, 1992; Mallat, 1998*]. Symlets possess the same properties of compactness, orthogonality, completeness, and smoothness as wavelets do; however, for orders of 2 to 6, they coincide with the ordinary orthogonal Daubechies basis, while orders of 8 to 20 reveal distinctions in the form of a basis function. For these reasons, we used 17 variants of orthogonal compact basis functions.

In choosing the optimal wavelet basis, the criterion of the entropy minimum in the distribution of the squared values of the wavelet coefficients:

$$E(x) = -\sum_{k=1}^{m}\sum_{j=1}^{2^{(m-k)}} p_j^{(k)} \cdot \ln(p_j^{(k)}) \to \min, \quad p_j^{(k)} = |c_j^{(k)}|^2 \Big/ \sum_{l,i} |c_i^{(l)}|^2 \tag{3.1}$$

is commonly used. Method (3.1) selects a basis for the signal $x(t)$ such that the distribution of the signal wavelet coefficients differs most from a uniform distribution. In this case, maximum information concentrates in the minimum number of wavelet coefficients. Usually, the application of criterion (3.1) yields quite satisfactory results. However, a more sophisticated method used below for the choice of the optimal basis has the form of an iterative procedure repeatedly using criterion (3.1). This was dictated by the desire to reveal the finest distinctions in signal structures by applying one or another basis. This method was proposed in [*Berger, Coifman, Goldberg, 1994*] for solving the problem of removing characteristic noises (hissing, cracks, and clicks) from old vocal recordings of opera classics and was called the method of successive coherent basis thresholding. The method can be briefly described as the sequence of the following operations:

1. Initialization: the initial signal $x_0(t)$ is moved into the working buffer $x(t)$.
2. The wavelet order is determined from criterion (3.1) for the signal $x(t)$: $E(x) \to \min$
3. The wavelet coefficients $c_j^{(k)}$ of the signal $x(t)$ are sorted in descending order of their magnitudes for testing the basis determined at stage (2), and the coefficients rearranged in this way are designated as d_j, $j = 0,1,...,(N-1)$. Thus, the coefficient d_0 has the maximum absolute value.
4. The minimum integer $M = 0,1,...,(N-1)$ is determined from the inequality

$$\frac{|d_M|^2}{\sum_{j=M+1}^{N-1} |d_j|^2} \le \frac{2 \cdot \ln(N-M)}{(N-M)} \tag{3.2}$$

5. If condition (3.2) is immediately fulfilled at $M = 0$ then the optimal order is found and the algorithm stops operating.
6. If condition (3.2) is not fulfilled for any $M = 0,1,...,(N-1)$, then the optimal wavelet order is set to be equal to the value found at stage (2) from the condition of the entropy minimum immediately after the initialization and the algorithm stops operating.
7. All coefficients $c_j^{(k)}$ for which $|c_j^{(k)}| \ge |d_M|$ are set at zero, the inverse wavelet transformation is applied to the remaining coefficients, the resulting residual signal is moved into the working buffer $x(t)$, and stage (2) is executed.

The meaning of this procedure reduces to the following. The signal is considered to consist of a desired signal, whose variations are reflected in the values of wavelet coefficients that are fairly large in magnitude, and noise, accounted for by all other coefficients. The problem is to choose the threshold of coefficient modulus above which the coefficients account for the desired signal and below which they account for the noise. Inequality (3.2) is precisely intended for determining such a threshold. This condition is taken from the formula for the probability of asymptotic maximum deviations of the Gaussian white noise values $B(t)$:

$$\lim_{N\to\infty} \Pr\{ \max_{0\le t\le(N-1)} |B(t)|^2 \Big/ \sum_{j=0}^{N-1} |B(j)|^2 \le \frac{2\ln N}{N} \} = 1 \tag{3.3}$$

The following formula, immediately resulting from (3.3), is also used below:

$$\lim_{N\to\infty} \Pr ob \{ \max_{0\le t\le(N-1)} |B(t)| \le \sigma\sqrt{2\cdot\ln N} \} = 1 \tag{3.4}$$

where σ is a standard deviation white noise $B(t)$.

Therefore, the meaning of condition (3.2) consists in the division of wavelet coefficients into noisy and useful. Coefficients responsible for noise have rather small absolute values (the lower limit of the sum in (3.2) is $M+1$) lying within the asymptotic limits for the white noise (formula (3.3)). However, such an extraction of noise from the signal depends on the basis used (thus, noise in one basis does not necessarily satisfy criterion (3.2) in another). Therefore, upon choosing the basis from the entropy minimum condition (stage 2), noise is determined in terms of this (stage 7); an optimal basis is then again determined, this time for the residual (depleted in information) signal (stage 2); and so on, until the residual signal becomes the noise, even with respect to its own optimal basis (stage 5). The last optimal basis determined in this way is considered as optimal because it is capable of recovering at least something from the most depleted residual signal. This approach makes sense because information can be retrieved from the initial signal $x_0(t)$ with the use of any basis, whereas only the best basis is effective in the case of a depleted residual signal. Note that, as regards the seismic records treated below, the basis found from such an iterative procedure coincides in 90% of cases with the basis determined immediately at the initialization stage; however, this cannot be stated a priori for any given signal.

After the optimal wavelet basis is determined for a given signal, we define the set of wavelet coefficients that are smallest in modulus and can be rejected in the inverse wavelet transformation because they account for noise. For this purpose, we assume that noise concentrates mostly in variations at the first, highest-frequency detail level with the exception of a small number of points at which high-frequency features of the desired signal behavior concentrate and to which, consequently, large values of the first-level wavelet coefficients correspond. Due to the wavelet transform orthogonality, the variance of wavelet coefficients is equal to the variance of the initial signal. Therefore, we estimate the standard deviation of the noise σ as the standard deviation of wavelet coefficients at the first detail level. This estimation must be robust, i.e., insensitive to outliers in desired values of wavelet coefficients at the first level. For example, we can use the robust median estimate of the standard deviation for a normal random value [*Huber, 1981*]

$$\sigma = med\{|c_j^{(1)}|,\ j = 1,...,N/2\}/0.6745 \tag{3.5}$$

Now, the estimate σ being found from (3.5), we can use (3.4) for estimating the threshold of wavelet coefficient moduli below which they can be set at zero because they are carriers of noise

variations. This threshold is equal to $\sigma\sqrt{2\cdot\ln N}$. As a result, one can readily determine the Donoho-Johnstone level α for the signal shrinkage, namely, the ratio of the number of coefficients for which the condition $|c_j^{(k)}|\le\sigma\sqrt{2\cdot\ln N}$ is fulfilled to their total number N .

This dimensionless value $\alpha, 0<\alpha<1$, is proposed as a default parameter for nonlinear wavelet filtering: the share α of all wavelet-coefficients which have absolute values less than the mentioned above threshold are nullified and the inverse wavelet transform is performed.

A4. Continuous Morlet wavelet transform diagrams.

Let $x(t)$ is the times series being analyzed. We are interested in its time-and-frequency structure. A continuous wavelet analysis is the most sensitive tool for the purpose [*Chui, 1992; Daubechies, 1992; Mallat, 1998*]. Let $\psi(t)$ is a certain function satisfying the admissibility condition: $\int_{-\infty}^{+\infty}\psi(t)\,dt=0$ and the normality condition: $\int_{-\infty}^{+\infty}|\psi(t)|^2\,dt=1$ A value which depends on two parameters ($(t,a), a>0$ is called a continuous wavelet transform:

$$Wx(t,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}x(s)\cdot\psi\left(\frac{s-t}{a}\right)ds=\sqrt{a}\int_{-\infty}^{+\infty}x(t+av)\cdot\psi(v)\,dv \qquad (4.1)$$

Here t – is the time point, $a>0$ – is the scale parameter, which we will further give a more usual term as “period”. The value (4.1) reflects behavior of the signal under study about a point $\boldsymbol{t}$ with a typical period of variations a. It is reasonable that the value (4.1) depends greatly on the choice of the function $\psi(t)$. Further we will use the so-called Morlet wavelet or a complex-valued modulated Gaussian:

$$\psi(t)=\frac{1}{\pi^{1/4}}\exp(-t^2/2-i\pi t) \qquad (4.2)$$

This wavelet is adapted best of all for distinguishing short-lived harmonic peaks (trains) and has certain optimality properties in the search of a compromise between the frequency and time resolution (yielding the so-called Heisenberg limit). Our primary aim is to construct a two-dimensional map of the squared module of value (4.1) $\ln(|Wx(t,a)|^2)$, which gives a pictorial presentation of the dynamics of the onset, evolution and disappearance of typical periods of the harmonic peaks of the signal under study.

Besides that a useful statistics is a continuous wavelet transform power spectra:

$$P(a)=<|Wx(t,a)|^2>_t \qquad (4.3)$$

where $<...>_t$ means a temporal averaging. The value (4.3) presents a distribution of mean energy of variations in dependence on their scales.

A5. Spectral and time-frequency spectral analysis.

Let $x(t)$ be time series to be analyzed. For estimating its power spectra an auto-regression (AR) model was applied [*Marple(Jr), 1987*]:

$$x(t)+\sum_{j=1}^{q}\alpha_j x(t-j)=\eta(t) \tag{5.1}$$

where α_j are auto-regression parameters, $\eta(t)$ is a residual signal which proposed to be a white noise with zero mean value and unknown variance σ^2. After identifying parameters $\alpha_j,\ j=1,...,q;\sigma^2$ of the model (5.1) spectral estimate is computing using formula:

$$S_{xx}(\omega)=\frac{\sigma^2}{2\pi\left|1+\sum_{j=1}^{q}\alpha_j\exp(-i\omega j)\right|^2} \tag{5.2}$$

where ω is a cyclic frequency: $\omega=2\pi/T$, T – period, measured in the units of sampling time interval δt, i – imaginary unit. AR-methods of spectral estimation differ from each other by different ways of estimating parameters of the model (5.1). We use a Burg's maximum entropy method as the most reliable and providing the best frequency resolution for short samples [*Marple(Jr), 1987*].

The more is the value of q, the more sensitive is the power spectra estimate. At the same time increasing of AR-order q follows the increasing statistical fluctuations of the estimate. Thus, the choice of q is the result of compromise between sensitivity and stability. We use a following method for choosing AR-order. Let N be a general number of records in the sample. A default spectral estimate is computed within moving time windows of the length L numbers. These time windows are overlapping and their mutual shift is defined from maximizing a number of records involving into spectral estimate. The final estimate is calculated as the average value of estimates from each time window. The default length of time window is taken from formula $L=\mathbf{min}(2048,N)$. The default value of AR-order q is chosen according to empirical rule: $q=\mathbf{int}(\mathbf{min}(L/\varphi,1000))$, where $\varphi=6.5-3.5\cdot\mathbf{cos}(\pi\cdot(L-32)/(2048-32))$. Thus, if $L=32$ (minimum possible general number of records within time series sample for analysis), then $\varphi=3$ (it means that for short length samples the AR-order equals 1/3 of the general length), whereas if $L=2048$ then $\varphi=10$. These default values of parameters L and q could be changed but the restrictions $L\le N, 1\le q\le L/3$ are fulfilled automatically anyway.

For time-frequency spectral analysis the estimate (5.2) is computed within moving time window of the length L, which are taken with mutual shift ΔL and for given AR-order q. Let τ be a time moment of the right-hand end of the moving time window. Then the value (5.2) could be presented at the form:

$$S_{xx}=S_{xx}(\omega,\tau\,|\,L,q),\quad \frac{2\pi}{(L-1)\cdot\delta t}\le\omega<\frac{\pi}{\delta t} \tag{5.3}$$

The values $L,\Delta L,q$ are the parameters of the time-frequency spectral estimate and the value (5.3) could be visualized as 2D image diagram on the plane of ω,τ or as 3D relief. The program proposes default values of the parameters $L,\Delta L,q$ but they could be changed by the user.

A6. Long Chains of Wavelet Transform Modulus Maximums.

Let $x(t)$ be a signal for analysis. Note that at least we will subtract the mean value of the signal $x(t)$ from its values before the analysis but usually we will subtract its linear trend or even fit the trend polynomial of the higher order. The order of the trend polynomial with is removed from the values of the signal within each time window is one of the parameter of the method. The necessity for trend removing is coming from usual considerations about preliminary suppressing dominated low-frequency components which are statistically not significant for analysis.

A scale-dependent kernel smoothed signal [*Hardle, 1989*]:

$$c_0(t,a) = \frac{\int_{-\infty}^{+\infty} x(t+av)\cdot\psi_0(v)\,dv}{\int_{-\infty}^{+\infty}\psi_0(v)\,dv} \tag{6.1}$$

where $a>0$ is a scale value and $\psi_0(t)$ is some fast decay function. Further on we shall use Gaussian: $\psi_0(t)=\exp(-t^2)$. Let us take the wavelet function:

$$\psi_n(t) = (-1)^n\cdot\frac{d^n\psi_0(t)}{dt^n} \equiv (-1)^n\cdot\psi_0^{(n)}(t) \tag{6.2}$$

Using formula for by-part integrating and fast decay properties of the function $\psi_0(t)$ we can obtain a formula for Taylor's coefficients (the n-th derivative of the smoothed signal, divided by $n!$) for the smoothed signal:

$$c_n(t,a) \equiv \frac{1}{n!}\frac{d^n c_0(t,a)}{dt^n} = \frac{\int_{-\infty}^{+\infty} x(t+av)\psi_n(v)\,dv}{a^n\int_{-\infty}^{+\infty} v^n\psi_n(v)\,dv} \tag{6.3}$$

The formula (6.1) is a particular case of the formula (6.3) for $n=0$.

The wavelet transform modulus maximum point (WTMM-point) (t,a) for $n\ge 1$ is defined as the point of local maximum of the value $|c_n(t,a)|$ with respect to time t for given scale a [*Mallat, 1998*]. For $n=0$ WTMM-points are defined as points of local extremes (maximums or minimums) of the smoothed signal $c_0(t,a)$. WTMM-points could by jointed into chains. The set of all chains forms a WTMM-skeleton [*Mallat, 1998*] of the signal. If $\psi_0(t)$ is a Gaussian, then WTMM-skeleton chain could not be aborted when the scale a is decreased [*Mallat, 1998*]. The WTMM-points for the 1st derivative $c_1(t,a)$ indicate time points of the maximum trend (positive or negative) of the smoothed signal $c_0(t,a)$ for the given scale value a.

Let the signal $x(t)$ be given for $t\in[0,T]$. When the time moment t is close to the beginning or to the end of the interval $[0,T]$, then smoothing transform (6.3) is exposed to the absence of information about behavior of the signal $x(t)$ for $t<0$ or for $t>T$. Usually this difficulty is

overcome by regarding the signal $x(t)$ as given not on time interval but on the circle, i.e. by extension the signal outside time interval $[0,T]$ by the rule: if $t<0$ then $x(t)=x(T+t)$ and if $t>T$ then $x(t)=x(t-T)$. This circular extension of the signal provides the ability to compute smoothing transform (6.3) for all time moments and is useful from the points of view of applying fast Fourier transform for fast computing the values (6.3). Nevertheless the values (6.3) are garbled at the ends of time interval $[0,T]$ and it would be better to introduce some "dead intervals" at the beginning and at the end of $[0,T]$ such that for time moments within these dead intervals WTMM-points are excluded from the analysis and from chaining procedure. For Gaussian $\psi_n(t)\approx 0$ for $|t|\geq 3$. Thus, we can introduce the following rule for dead end interval:

$$\begin{array}{l}\text{if} \quad 0\leq t\leq 3a \quad \text{or} \quad T-3a\leq t\leq T \\ \text{then} \quad c_n(t,a) \quad \text{are excluded from analysis}\end{array} \tag{6.4}$$

From the rule (6.4) it follows the value of maximum possible scale value $a_{\max}$ which is suitable for the analysis: $a_{\max}=T/6$ – for this value the only admitted time point is $t=T/2$. The right-hand ends of the dead time intervals adjacent to $t=0$ and the left-hand end of dead time intervals adjacent to $t=T$ form a cupola-like area of permissible points (t,a) on the 2D time-scales diagrams on the plane of $(t,\ln(a))$-values. From the property of continuality of smoothed curves $c_0(t,a)$ it follows that long chains of their local minimums or maximums could have their ends either at the bifurcation point or on the upper boundary of possible values of the scale which is followed from the condition (6.4).

The long chains of local extreme (minimum and maximum) values of $c_0(t,a)$ and maximum absolute values of its 1st derivative $c_1(t,a)$ present the most interest for characterizing the main features of the signal behavior for various scales because they give some kind of "fingerprint" of the signal. Characterizing chains of 1st derivative modulus maximums $|c_1(t,a)|$ we must distinguish chains with negative from positive signs of $|c_1(t,a)|$ as chains of maximum scale-dependent negative (decreasing) or positive (increasing) trends. For time series with sampling time interval Δt the minimum scale is equal to the period of Nyquist: $a_{\min}=2\cdot\Delta t$. Further on we must define the criteria of long chain: it must attain or exceed certain threshold level $\gamma\cdot a_{\max}$ where parameter of the method γ must satisfy the condition: $\frac{a_{\min}}{a_{\max}}\equiv\gamma_{\min}\leq\gamma<1$. The closer is the value of γ to 1, the less is the number of "long chains". For $\gamma=\gamma_{\min}$ all chains are considered to be "long".

If the value of the parameter γ is gradually increasing then a hierarchical set of scale-dependent change points of time moments of the beginning of "long chains" occur. The study of such hierarchical change points is an alternative to usual approach based on spectral characteristics of the signals.

A7. Multifractal Singularity Spectra Estimate by Detrended Fluctuation Analysis.

Let $x(t)$ be a random signal. Let's define its measure of variability $\mu_X(t,\delta)$ on the time interval $[t,t+\delta]$ as absolute value of increment $\mu_x(t,\delta)=|x(t+\delta)-x(t)|$ and calculate its mean value in the power degree q: $M(\delta,q)=E\{(\mu_x(t,\delta))^q\}$. Random signal is scale-invariant if $M(\delta,q)\sim|\delta|^{\rho(q)}$ when $\delta\to 0$ that is the following limit exists:

$$\rho(q) = \lim_{\delta \to 0} \left(\ln M(\delta, q) / \ln |\delta| \right) \tag{7.1}$$

In this definition the measure $\mu_x(t, \delta)$ could be taken as the difference between maximum and minimum values as well:

$$\mu_x(t, \delta) = \max_{t \le u \le t+\delta} x(u) - \min_{t \le u \le t+\delta} x(u) \tag{7.2}$$

If $\rho(q)$ is a linear function: $\rho(q) = Hq$, where $H = const, 0 < H < 1$, then the process is called monofractal. For estimating the value of $\rho(q)$ using a finite sample $x(t), t = 1, ..., N$ a *detrended fluctuation analysis* could be used (*Kantelhardt et al., 2002*). Let s be a number of samples corresponding to the scale δ_s: $\delta_s = s\Delta t$. Let's split the whole sample into non-overlapping time intervals of the length s values:

$$I_k^{(s)} = \{t : 1 + (k-1)s \le t \le ks, \quad k = 1, ..., [N/s]\} \tag{7.3}$$

and let

$$y_k^{(s)}(t) = x((k-1)s + t), \quad t = 1, ..., s \tag{7.4}$$

be a part of the signal $x(t)$, corresponding to interval $I_k^{(s)}$. Let $p_k^{(s,m)}(t)$ be a polynomial of the order m, best fitted to the signal $y_k^{(s)}(t)$. Let us consider the deflections from local trend:

$$\Delta y_k^{(s,m)}(t) = y_k^{(s)}(t) - p_k^{(s,m)}(t), \quad t = 1, ..., s \tag{7.5}$$

and calculate the values

$$Z^{(m)}(q, s) = \left(\sum_{k=1}^{[N/s]} (\max_{1 \le t \le s} \Delta y_k^{(s,m)}(t) - \min_{1 \le t \le s} \Delta y_k^{(s,m)}(t))^q \Big/ [N/s] \right)^{1/q} \tag{7.6(a)}$$

that could be regarded as estimate of $(M(\delta_s, q))^{1/q}$. Other variant consists in using standard deviation as the measure $M(\delta, q)$:

$$Z^{(m)}(q, s) = \left(\sum_{k=1}^{[N/s]} \left(\frac{1}{s} \sum_{t=1}^{s} (\Delta y_k^{(s,m)}(t))^2 \right)^{q/2} \Big/ [N/s] \right)^{1/q} \tag{7.6(b)}$$

Let's define function $h(q)$ as coefficient of linear regression between $\ln(Z^{(m)}(q, s))$ and $\ln(s)$: $Z^{(m)}(q, s) \sim s^{h(q)}$ fitted for scales $s_{\min} \le s \le s_{\max}$. It is evident that $\rho(q) = qh(q)$ and for monofractal signal $h(q) = H = const$. Multifractal singularity spectrum $F(\alpha)$ equals to fractal dimensionality of the set of time moments t for which Hölder-Lipschitz exponent equals α i.e. for which $|x(t+\delta) - x(t)| \sim |\delta|^{\alpha}, \quad \delta \to 0$. Singularity spectrum could be estimated by standard multifractal formalism (*Feder, 1988*) which consists in calculating a Gibbs` sum:

$$W(q, s) = \sum_{k=1}^{[N/s]} (\max_{1 \le t \le s} \Delta y_k^{(s,m)}(t) - \min_{1 \le t \le s} \Delta y_k^{(s,m)}(t))^q \tag{7.7}$$

and estimating mass exponent $\tau(q)$ from the condition $W(q,s) \sim s^{\tau(q)}$. After that the spectrum $F(\alpha)$ is calculated by Legendre transform:

$$F(\alpha) = \max \{ \min_q (\alpha q - \tau(q)), 0 \} \quad (7.8)$$

From (7.6) and (7.7) it follows that $\tau(q) = \rho(q) - 1 = qh(q) - 1$.

Thus, $F(\alpha) = \max \{ \min_q (q(\alpha - h(q)) + 1, 0 \}$.

For mono-fractal signal $F(H) = 1, \quad F(\alpha) = 0 \; \forall \alpha \neq H$ but finite sample estimate of singularity spectrum does not obey these rigorous theoretical conditions of course.

References

Berger J., R. Coifman, and M., Goldberg (1994) Removing noise from music using local trigonometric bases and wavelet packets. J. Audio Eng. Soci., Vol. 42, No. 10, pp. 808-818.

Chui, C.K. (1992) An Introduction to Wavelets, Academic Press, San Diego, CA

Cox D.R., Lewis P.A.W. (1966) The statistical analysis of series of events. London, Methuen.

Daubechies I. (1992) Ten Lectures on Wavelets. No.61 in CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia.

Donoho D., I. Johnstone (1994) Ideal spatial adaptation via wavelet shrinkage. Biometrika, vol. 81, pp. 425-455.

Feder J. (1988) Fractals. Plenum Press, New York, London

Hardle W. (1989) Applied nonparametric regression. Cambridge University Press, Cambridge, New York, New Rochell, Melbourne, Sydney.

Huber P.J. (1981) Robust statistics. John Wiley and Sons. New York, Chichester, Brisbane, Toronto.

Kantelhardt J. W., Zschiegner S. A., Konscienly-Bunde E., Havlin S., Bunde A., and Stanley H. E. (2002) Multifractal detrended fluctuation analysis of nonstationary time series, Physica A, 316, 87–114, 2002.

Lyubushin A.A., V.F.Pisarenko, V.V.Ruzich and V.Yu.Buddo (1998) A New Method for Identifying Seismicity Periodicities – Volcanology and Seismology, vol.20, 1998, pp. 73-89.

Lyubushin A.A. (2000) Wavelet-Aggregated Signal and Synchronous Peaked Fluctuations in Problems of Geophysical Monitoring and Earthquake Prediction. – Izvestiya, Physics of the Solid Earth, vol.36, 2000, pp. 204-213.

Lyubushin A.A. (2001) Multidimensional Wavelet Analysis of Geophysical Monitoring Time Series – Izvestiya, Physics of the Solid Earth, vol. 37, 2001, pp. 474-483.

Lyubushin A.A. (2002) Robust wavelet-aggregated signal for geophysical monitoring problems – Izvestiya, Physics of the Solid Earth, vol.38, 2002, pp. 745-755.

Lyubushin A.A., Z.Kalab and N.Castova (2004) Application of Wavelet Analysis to the Automatic Classification of Three-Component Seismic Records – Izvestiya, Physics of the Solid Earth, vol.40, No.7, 2004, pp. 587-593.

Lyubushin A.A. and Kopylova G.N. (2004) Multidimensional Wavelet Analysis of Time Series of Electrotelluric Observations in Kamchatka – Izvestiya, Physics of the Solid Earth, vol.40, No.2, 2004, pp. 163-175.

Lyubushin A.A. and G.A. Sobolev (2006) Multifractal Measures of Synchronization of Microseismic Oscillations in a Minute Range of Periods – Izvestiya, Physics of the Solid Earth, vol.42, No.9, 2006, pp.734-744.

Lyubushin A.A. (2007) Geophysical and ecological monitoring systems data analysis. Moscow, ‘Nauka’, 228 p. (in Russian).

Lyubushin A.A. (2009) Synchronization Trends and Rhythms of Multifractal Parameters of the Field of Low-Frequency Microseisms – Izvestiya, Physics of the Solid Earth, 2009, Vol. 45, No. 5, pp. 381–394.

Lyubushin A. (2010) Multifractal Parameters of Low-Frequency Microseisms // V. de Rubeis et al. (eds.), Synchronization and Triggering: from Fracture to Earthquake Processes, GeoPlanet: Earth and Planetary Sciences 1, DOI 10.1007/978-3-642-12300-9_15, Springer-Verlag Berlin Heidelberg 2010, 388p., Chapter 15, pp.253-272.

Lyubushin A.A. (2011) Seismic Catastrophe in Japan on March 11, 2011: Long-Term Prediction on the Basis of Low-Frequency Microseisms – Izvestiya, Atmospheric and Oceanic Physics, 2011, Vol. 46, No. 8, pp. 904–921.

Mallat S. (1998) A wavelet tour of signal processing. Academic Press. San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto. 577 p.

Marple S.L.(Jr.) (1987) Digital spectral analysis with applications. Prentice-Hall, Inc., Englewood Cliffs, New Jersey

Press W.H., Flannery B.P., Teukolsky S.A. and Vettering W.T. (1996) Numerical Recipes, 2-nd edition, Chapter 13, Wavelet Transforms, Cambridge Univ. Press, Cambridge.

Sobolev G. and A. Lyubushin (2010), Periodical Oscillations of Microseisms before the Sumatra Earthquake of December 26, 2004 // V. de Rubeis et al. (eds.), Synchronization and Triggering: from Fracture to Earthquake Processes, GeoPlanet: Earth and Planetary Sciences 1, DOI 10.1007/978-3-642-12300-9_13, Springer-Verlag Berlin Heidelberg 2010, 388p., Chapter 13, pp.223-241.

Wilks S.S. (1962) Mathematical statistics. John Wiley & Sons, Inc. N.Y., London, Sydney. 644 p.