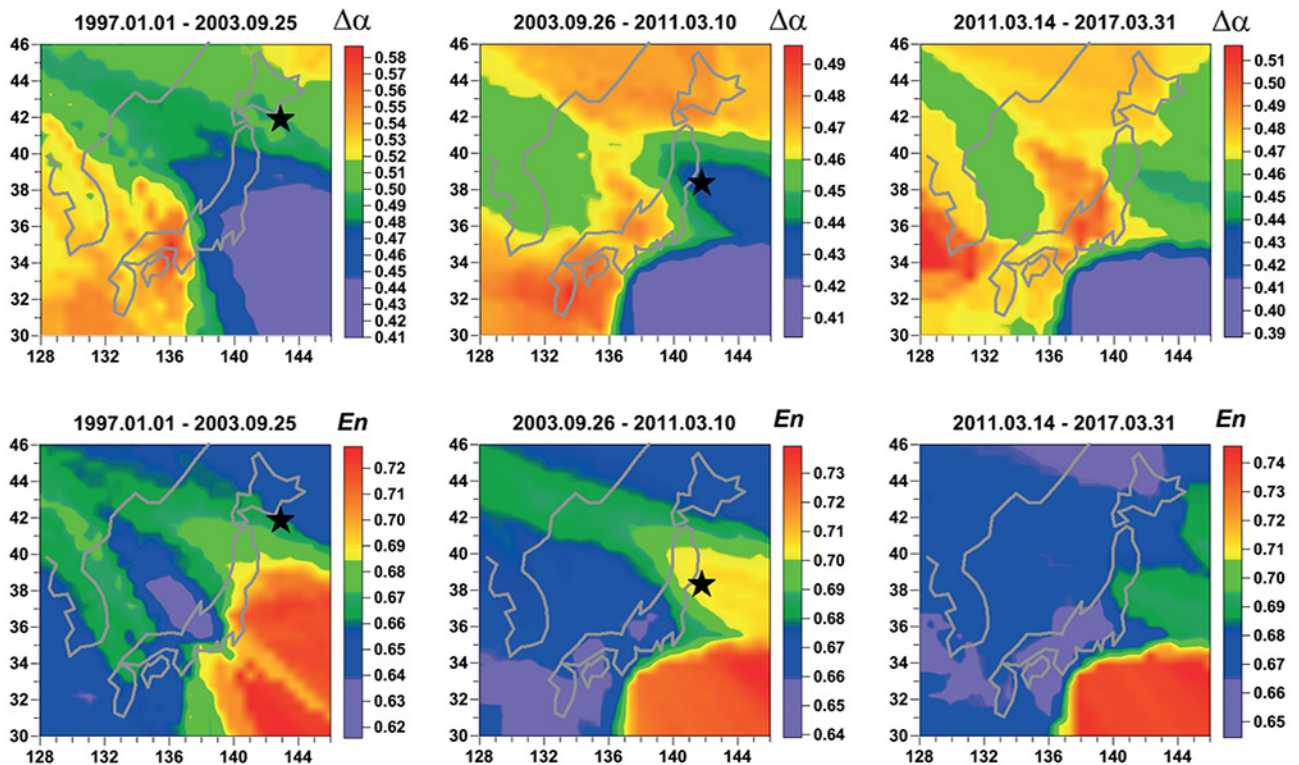# COMPLEXITY OF SEISMIC TIME SERIES

Measurement and Application

Edited by
Tamaz Chelidze, Filippos Vallianatos, Luciano Telesca

Synchronization of Geophysical Fields Fluctuations

Alexey Lyubushin

Institute of Physics of the Earth, Russian Academy of Sciences, Moscow
lyubushin@yandex.ru

Abstract

The methods for investigating synchronization of multiple geophysical monitoring time series are presented, which are based on using wavelet-based and spectral measures of coherence, estimated within moving time window. The methods are applied to the sequences of multifractal and entropy parameters of initial data calculated for adjacent time intervals of certain length. Precursory properties of coherence and multifractal structure of seismic noise at global and regional networks are investigated. Analysis of global seismic noise coherence extracts effect of progressively increasing synchronization after Sumatra mega-earthquake, which could be a precursor of the further rise in the intensity of the strongest seismic events. Analysis of seismic noise at Japan gave a possibility for prediction of Tohoku earthquake on 2011.03.11. According to the analysis of seismic noise after 2011.03.11 the next mega-earthquake could occur at the region of Nankai Trough.

Keywords: seismic noise, synchronization, coherence, multifractals, entropy, earthquake prediction, spots of seismic danger.

Acknowledgement

Introduction

Random fluctuations in geophysical fields carry important information about the processes occurring in the Earth's crust, including preparation of major geological disaster. Synchronization of noise measurements obtained in different points of the monitoring network, is an important indicator of approximation of a complex systems to a drastic change in its properties by virtue of their own dynamics. Therefore, the development and testing of methods of analysis of monitoring data with purposes to allocate time intervals of increased coherence and to determine the characteristic frequencies of the synchronization of a large number of time series of geophysical monitoring systems is a very urgent task in the geosciences. This chapter outlines the methods of analysis of multivariate time series, allowing explore the dynamics of change in the degree of coherence of the noise in the data stream from the monitoring systems.

One of the fundamental problems of geophysical monitoring is 'complexation' of different observations and measurements. This term means joint analysis of observations of different geophysical fields or observations of the same field, but at different measurement points, or both simultaneously. The idea of complexation is based on the hypothesis that using of a large

number of monitored parameters can help extracting a weak common signal, which 'drowns' in the strong noise when each individual measurement is considered separately. The main feature of such a common signal must be its coherence (correlation) in a variety of observations, the use of which allows one to detect the very existence of the common components, despite the fact that the frequency content of the common signal may coincide with the frequency content of the strong local noise.

Identification of precursors of earthquakes or other geological catastrophes is among the most challenging problems of complexing measurements. In such a problem, a weak common signal is related to earthquake preparation processes, such as consolidation of the Earth's crust matter in a future earthquake focus and around it [Rice, 1980]. The search for precursors of catastrophes as the occurrence of synchronous components in a variety of observations is the general idea about increasing the correlation radius of the random fluctuations of the parameters of a complex system, when it approaches a sharp change in its properties as a result of its own dynamics [Gilmore, 1981; Nicolis, Prigogine, 1989].

The idea of complexation requires using methods of analysis of multivariate time series. In this chapter a set of algorithms for the analysis of multivariate time series, obtained by monitoring systems will be presented, which are intended for discovering the hidden relations between processes, including those of different nature and structure. An important part of the developed algorithms is a preliminary analysis of time series with different scales for a purpose to extract the dimensionless characteristics within adjacent time intervals of some short length. These characteristics were chosen in the dimensionless forms which are independent on the physical meaning of initial time series. Such transformation of initial high-frequency time series into low-frequency series is an important tool, which helps to detect hidden coherence effects, that could not be found, when initial data are processed. Analysis of the noise is often neglected, although statistical regularities of the noise structure are an important source of hidden information about upcoming sharp changes in the properties of the objects under consideration. The methods are based on the analysis of canonical coherences, multidimensional spectral matrices and canonical correlation coefficients of the wavelet decomposition of signals in moving time windows. The purpose of these algorithms is the detection of a very weak non-stationary signals of common origin, having both harmonic oscillation behavior or sharply non-stationary, wavelet character, and find time intervals and frequency bands of these common signals.

Wavelet-based robust coherence measure

The robust wavelet-based measure of coherence is a modification of the approach to the analysis of multidimensional time series proposed in [Lyubushin, 2000; Lyubushin and Kopylova, 2004]. In the paper [Lyubushin, 2015] this measure was applied to investigation of global seismic noise coherence. Let $Z(t) = (Z_1(t),...,Z_q(t))^T$ be multidimensional time series of the dimensionality $q \geq 2$, $t = 0,1,...$ is a discrete time index which numerates successive samples. The scale-dependent measure of coherent behavior in a moving time window of a given length of N samples is constructed. Let's consider that time index t within each time window is varying from 0 up to $(N-1)$. The analysis is performed independently for each position of the time window (moved to the right by one sample). Before the wavelet decomposition of the analyzed time series fragments, which is presented in the current time window, the following sequence of operations is applied to each fragment:
  (i)   the general linear trend within the time window is removed;
  (ii)   a sample estimate of the standard deviation is obtained, and each value is divided by this estimate;
  (iii) a tapering operation is performed within each time window;

(iv)   the window fragment is padded by zeros to the full length of $M = \min\{2^m : 2^m \geq N\}$ samples.

Operation (i) removes the strongest low frequency variations in signals, which cannot be statistically representative within the window. The division of each signal within the window by its standard deviation mutually adjusts different time series by reducing the total energy of their variations to the same value. Tapering in the point (iii) is a usual preliminary operation in spectral and wavelet analysis before applying discrete Fourier or wavelet transform and consists in multiplying the samples within current time window by the positive function, which is tending to zero, when samples approach the left and right ends of the window. We use a cosine tapering function which equals $(1-\cos(\pi t / L))/2$ for $0 \leq t \leq L$ and $(1-\cos(\pi(t-(N-1))/L))/2$ for $(N-1)-L \leq t \leq (N-1)$. Here L is the length of time intervals at the beginning and at the end of time window, where tapering operation is performed. We used the value $L = N/8$. Tapering operation (iii) is necessary for reducing negative "wrap-around" effect of finite discrete wavelet transform [Press et al., 1996]. Finally, the last operation (iv) is necessary for the subsequent application of the fast discrete wavelet transform.

Let τ be the position of the right-hand end of a moving time window N samples wide. After performing discrete orthogonal wavelet transform within window after preliminary operations (i)-(iv) we will have a set of tables of wavelet-coefficients:

$$c_j^{(\beta,\tau)}(k), \quad j=1,...,q; \quad k=1,...,M_\beta = 2^{(m-\beta)}; \quad \beta=1,...,m \qquad (6.1)$$

We chose the Haar wavelet from the family of orthogonal wavelets as the most compact and suitable for the analysis of the most abrupt variations in signals. Here β is the number of detail level of wavelet decomposition, m is a power of 2 in the presentation $M = 2^m$ of minimum integer which is not less than N. The general number of wavelet coefficients at the detail level β equals $M_\beta = 2^{(m-\beta)}$. The index j corresponds to different scalar components of the multidimensional time series $Z(t)$, whereas index k successively enumerates the coefficients belonging to the level $\beta=1,...,m$. However, some of these coefficients may correspond to the zero padding of the sample in point (iv) of the preliminary transformations. Therefore, the actual number of coefficients at the level β, reflecting the behavior of the signal inside the window, is equal to $L_\beta(N) = 2^{(m-\beta)}(N/M) = 2^{-\beta}N$. Each coefficient $c_j^{(\beta,\tau)}(k)$ reflects the signal behavior in the frequency band with approximate bounds $[\Omega_{\min}^{(\beta)}, \Omega_{\max}^{(\beta)}] = [1/(2^{(\beta+1)}\Delta s), 1/(2^\beta \Delta s)]$, where Δs is the length of the sampling interval, in the neighborhood of the sample with the number $\tau_k^{(\beta)} = k \cdot 2^\beta$, $k=1,...,M_\beta$, measured from the position of the left end of the time window. The width of this neighborhood (the temporal "zone of responsibility" of the coefficient) is equal to $\Delta s \cdot 2^\beta$.

The number of wavelet coefficients is reducing exponentially with increasing of the number of detail level β. That is why $L_\beta(N)$ is decreasing with the same rate and starting from some detail level number β the number of wavelet coefficients $L_\beta(N)$ could be equal zero. Therefore, it is natural to introduce the parameter of statistical significance $L_{\min}$ as the minimum possible value of the number of wavelet coefficients $L_\beta(N)$ that allows one to perform statistical estimates using wavelet coefficients at the β-th detail level. It is possible to determine the maximum possible detail level $\beta_{\max}$ by the formula $\beta_{\max} = \max\{\beta : L_\beta(N) \geq L_{\min}\}$. Thus, the

length of the window N and the significance threshold L_{min} together set the maximum possible detail level β_{max}, the wavelet coefficients of which can be included into the analysis.

Now we address a scalar time series j_0 from multiple series $Z(t)$ and construct the measure describing the relationship between this series and all other scalar signals within the current time window. Naturally, this relationship depends on the scale of the variations in question and, therefore, should be sought at various detail levels between wavelet expansion coefficients. The problem to be solved for this purpose is

$$\sum_{k=1}^{L_\beta} |c_{j_0}^{(\beta,\tau)}(k) - d_{j_0}^{(\beta,\tau)}(k\,|\,\gamma)| \to \min_{\gamma_j} \qquad (6.2)$$

where $d_{j_0}^{(\beta,\tau)}(k\,|\,\gamma)$ is a linear combination of wavelet coefficients at the detail level β from all other scalar signals except signal with number j_0 with unknown coefficients γ_j:

$$d_{j_0}^{(\beta,\tau)}(k\,|\,\gamma) = \sum_{j=1, j\neq j_0}^{q} \gamma_j \cdot c_{j_0}^{(\beta,\tau)}(k) \qquad (6.3)$$

We should emphasize that the sum in (3) is a linear combination of expansion coefficients of time series except the chosen series j_0. The problem (6.2) is solved by the method of generalized gradient [Clarke, 1975]. Finding the vector γ from the solution of problem (6.2), we obtain certain values of $d_{j_0}^{(\beta,\tau)}(k\,|\,\gamma)$. Now we can find the correlation coefficient between samples of the values of $c_{j_0}^{(\beta,\tau)}(k)$ and $d_{j_0}^{(\beta,\tau)}(k\,|\,\gamma)$ for $k = 1,...,L_\beta$; however, instead of the classic Pearson formula for calculating the sample value of the correlation coefficient, we use its robust modification [Huber, Ronchetti, 2009], according to which the correlation coefficient between samples $x(k)$ and $y(k)$, $k = 1,...,n$, can be calculated by the formula

$$\rho(x, y) = \frac{S(\hat{z}^2) - S(\breve{z}^2)}{S(\hat{z}^2) + S(\breve{z}^2)} \qquad (6.4)$$

where

$$\hat{z}(r) = a \cdot x(r) + b \cdot y(r), \quad \breve{z}(r) = a \cdot x(r) - b \cdot y(r),$$
$$a = 1/S(x), \quad b = 1/S(y), \quad S(x) = med\,|\,x - med(x)| \qquad (6.5)$$

$med(x)$ is the median value of the sample $x(k)$, $S(x)$ is its absolute median deviation.

Substituting $x(k)$ for $c_{j_0}^{(\beta,\tau)}(k)$, $y(k)$ for $d_{j_0}^{(\beta,\tau)}(k\,|\,\gamma)$, and n for $L_\beta(N)$, we obtain the robust value $v_{j_0}(\beta,\tau)$ of the correlation coefficient describing the degree of connection of the process j_0 with all other signals from multiple time series $Z(t)$. If we replace in (6.2) the sum of the moduli of deviations by the sum of their squares, the problem can be reduced to the classic Hotelling problem of canonical correlations [Hotelling, 1936; Rao, 1965]. Therefore, the quantity $v_{j_0}(\beta,\tau)$ is here referred to as the robust canonical correlation of the time series j_0. The need to replace the classic scheme of the calculation of canonical correlations by its robust variant is dictated by the strong instability of the result of the classic calculations with respect to outliers in wavelet coefficients. The presence of such outliers is due to the well-known fact that

the wavelet decomposition is capable of accumulating maximum information about the signal behavior in a relatively small number of wavelet coefficients. We should emphasize that the method is robust in two procedures: the solution of minimization problem (6.2) by the method of least moduli rather than by least squares and the calculation of the correlation coefficient by formula (6.4).

Since, with an increase in the number of the detail level, the number of wavelet coefficients involved in the estimation of $v_k(\beta, \tau)$ exponentially decreases, we reduce statistical fluctuations in estimates by introducing additional averaging over a certain number of coefficients obtained within preceding windows:

$$\overline{V}_k(\beta, \tau) = \sum_{s=1}^{m_\beta} V_k(\beta, \tau - s + 1) / m_\beta, \quad m_\beta = 2^\beta \tag{6.6}$$

The higher the detail level, the deeper the averaging (6.6) over the past time windows; this fact considerably decreases the dependence of the variance of statistical fluctuations in estimation (6.6) on the detail level number and makes this variance nearly the same for different values of β. According to formula (6.6), the effective width of the time window becomes scale-dependent and equal to $(N + 2^\beta - 1)$.

We define the robust wavelet-based measure of coherence by the formula

$$\rho(\tau, \beta) = \prod_{k=1}^{q} |\overline{V}_k(\tau, \beta)| \tag{6.7}$$

The values of measure (6.7) range from 0 to 1. The larger the value of (6.7), the stronger the overall connection between all analyzed processes on scales corresponding to the number β. We should emphasize that the value of (6.7) is the product of q nonnegative values with moduli less than unity. Therefore, the greater the number q of the series analyzed, the lower the absolute values of $\rho(\tau, \beta)$. As a consequence, the absolute values of statistic (6.7) can be compared only for the same number of series q. Most interesting are not the absolute values of measure (6.7) but its relative values for different values of τ. Thus, with a fixed Haar wavelet in use, the method has two free parameters: the time window length N and the significance threshold $L_{\min}$. Further on we will use the threshold $L_{\min} = 16$.

It should be noticed that the value (6.7) could be calculated without wavelet decomposition of the signals within moving time window as well. At this case it is not dependent on the level β and it is not necessary to make additional smoothing (6.6) using values within preceding time windows. Thus, the by-component robust correlations v_k become level independent and the value (6.7) becomes the formula for robust multiple correlation coefficient:

$$\rho(\tau) = \prod_{k=1}^{q} |v_k(\tau)| \tag{6.8}$$

If $q = 2$ then the value (6.8) presents robust estimate of squared correlation coefficient between two time series.

Multiple spectral coherence measure

The multiple spectral coherence is analogous to wavelet-based multiple coherence but it is based on classic Fourier basis functions instead of orthogonal wavelets. It uses frequency-dependent canonical coherences, which are similar to level-dependent wavelet-based canonical correlations.

Canonical coherences are the generalization of usual squared coherence spectrum between two scalar time series for the case when two vector time series are considered: m-dimensional time series $X(t)$ and n-dimensional time series $Y(t)$. Here t is an integer time index. Without loss of generality let us suppose that $m \leq n$. Squared maximum canonical coherence $\mu^2(\omega)$ between multiple time series $X(t)$ and $Y(t)$ is computed as maximum eigenvalue of the following frequency-dependent matrix [Brillinger, 1975; Hanan, 1970]:

$$U(\omega) = S_{xx}^{-1} S_{xy} S_{yy}^{-1} S_{yx} \qquad (6.9)$$

Here ω is the frequency, $S_{xx}(\omega)$ is spectral matrix of the size $m \times m$ of time series $X(t)$, $S_{xy}(\omega)$ is cross-spectrum matrix of the size $m \times n$ between time series $X(t)$ and $Y(t)$, $S_{yx}(\omega) = S_{xy}^{H}(\omega)$, "$H$" is the sign of Hermitian conjunctions (i.e. transposition of the matrix and complex conjugation), $S_{yy}(\omega)$ is spectral matrix of the size $n \times n$ of time series $Y(t)$. The value of $\mu^2(\omega)$ is used instead of usual squared coherence spectrum when 2 scalar time series are regarded, i.e. when $m = n = 1$.

If we take $X(t)$ as scalar i-th component of q-dimensional time series $Z(t)$ and $Y(t)$ as $(q-1)$-dimensional time series composed of all other scalar components of $Z(t)$, then function (6.9) became scalar which could be named by-component canonical coherence $\nu_i^2(\omega)$

The value $\nu_i^2(\omega)$ is the measure of connection of variations of the i-th component of q-dimensional time series $Z(t)$ with variations of all other scalar components of $Z(t)$ at the frequency ω. The inequality $0 \leq |\nu_i(\omega)| \leq 1$ is fulfilled, and the closer the value of $|\nu_i(\omega)|$ to unity, the stronger the linear relation of variations at the frequency ω of the i-th scalar series to analogous variations in all other series. Now we can define the multiple spectral coherence measure by formula:

$$\lambda(\omega) = \prod_{i=1}^{q} |\nu_i(\omega)| \qquad (6.10)$$

The value (6.10) provides a frequency-dependent measure of linear joint synchronization of variations of all scalar components of time series $Z(t)$ at the frequency ω. Because the dimensionality of series $X(t)$ in the formula (6.9) equals 1, $m = 1$, the matrix $U(\omega)$ in fact is a scalar. Thus, its "maximum eigenvalue" is the value of the following quadratic form divided by power spectrum of i-th component:

$$\nu_i^2(\omega) = S_i^H(\omega)(S_{ZZ}^{(i)}(\omega))^{-1} S_i(\omega) / P_i(\omega) \qquad (6.11)$$

Here $S_{ZZ}^{(i)}(\omega)$ is a Hermitian matrix of the size $(q-1) \times (q-1)$, which is obtained from the full spectral matrix $S_{ZZ}(\omega)$ of the size $q \times q$ of multiple time series $Z(t)$ by removing its i-th column and i-th row, $S_i(\omega)$ is a $(q-1)$-dimensional vector consisting of cross-spectrums

between i-th component of $Z(t)$ with all other its scalar components. It is evident that vector $S_i(\omega)$ is composed of elements of spectral matrix $S_{ZZ}(\omega)$ from i-th column except the elements in the i-th row. Finally $P_i(\omega)$ is a power spectrum of i-th component of $Z(t)$, i.e. the i-th element on the main diagonal of the matrix $S_{ZZ}(\omega)$. The matrix $S_{ZZ}^{(i)}(\omega)$ is Hermitian and positively defined – that is why quadratic form $S_i^H(\omega)(S_{ZZ}^{(i)}(\omega))^{-1}S_i(\omega)$ is real and positive.

For calculating the measure (6.10) using values (6.11) it is necessary to estimate spectral matrix $S_{zz}(\omega)$ of the size $q{\times}q$. For this purpose we use vector autoregression model [Marple, Jr., 1987]:

$$Z(t) + \sum_{k=1}^{p} A_k \cdot Z(t-k) = e(t) \tag{6.12}$$

where p is an autoregression order, A_k are matrices of autoregression coefficients of the size $q{\times}q$, $e(t)$ is q-dimensional residual signal with zero mean and covariance matrix $\Phi = M\{e(t)e^T(t)\}$ of the size $q{\times}q$. Matrices A_k and Φ are defined using Durbin-Levinson procedure and the spectral matrix is calculated using formula:

$$S_{ZZ}(\omega) = F^{-1}(\omega) \cdot \Phi \cdot F^{-H}(\omega), \quad F(\omega) = E + \sum_{k=1}^{p} A_k \cdot \exp(-i\omega k) \tag{6.13}$$

where E is a unit matrix of the size $q{\times}q$.

When $q = 2$ the value (6.10) equals to usual squared coherence spectrum:

$$\lambda(\omega) = |S_{12}(\omega)|^2 \big/ \big(S_{11}(\omega) \cdot S_{22}(\omega)\big) \tag{6.14}$$

where $S_{11}(\omega)$ and $S_{22}(\omega)$ are diagonal elements of the matrix (6.13), i.e. parametric estimates of the power spectra of two signals, and $S_{12}(\omega)$ is their mutual cross-spectrum.

Let us consider moving time window of the certain length and let τ be the time coordinate of right-hand end of moving time window. If the function (6.10) will be estimated within each time window independently then we will have time-frequency function:

$$\lambda(\tau,\omega) = \prod_{i=1}^{q} |v_i(\tau,\omega)| \tag{6.15}$$

The value (6.15) presents the evolution of linear synchronization measure for multiple time series $Z(t)$. It is important that, before calculating the spectral matrix, each scalar component of the multidimensional time series was subjected (independently in each time window) to the following preliminary operations. First of all, the general linear trend was eliminated and optionally conversion to the increments was carried out. Then, the obtained data were winsorized [Huber, Ronchetti, 2009]: the sample mean and standard deviation σ were iteratively calculated, the mean was subtracted from the sample, after which the counts were divided by σ and all the values that fell beyond the limits of $\pm 3\sigma$ were replaced by their limiting values. The

iterations were repeated until σ stopped changing. These procedures ensure the robustness of the estimate of the coherence measure to the outliers (extreme values).

Besides the frequency–time dependence $\lambda(\tau,\omega)$, also the pure time-dependent measures of the maximum and mean coherence in the current time window with coordinate τ were used:

$$\lambda_{\max}(\tau) = \max_{\omega_{\min} \leq \omega \leq \omega_{\max}} \lambda(\tau,\omega), \quad \lambda_{\mean}(\tau) = \sum_{\omega_{\min} \leq \omega \leq \omega_{\max}} \lambda(\tau,\omega) / m(\omega_{\min},\omega_{\max}) \tag{6.16}$$

Here $m(\omega_{\min},\omega_{\max})$ is the number of discrete frequency values within frequency band $[\omega_{\min},\omega_{\max}]$. We note that quantities (6.16) are certain analogs of the coefficient of multiple correlation $\rho(\tau)$ from (6.8) calculated in the moving time window. However, since the maximum and mean values in formula (6.16) are taken over the frequencies, these coefficients allow account for the time shifts between the scalar components of multidimensional time series within the current time window.

Multiple spectral coherence measure in the form (6.10)-(6.11) was suggested in [Lyubushin, 1998] for multidimensional time series processing in the problems of geophysical monitoring. In the papers [Lyubushin, 1999; Lyubushin et al., 2003, 2004; Lyubushin, Sobolev, 2006; Lyubushin, 2008(b), 2009, 2010(a,b); Lyubushin, Klyashtorin, 2012; Lyubushin, 2014(a); Lyubushin et al, 2016] this spectral measure was applied to different problems of multidimensional time series analysis in geophysics, meteorology, hydrology and climate sciences.

Statistics of time fragments

For characterizing changing of statistical properties of analyzed geophysical time series estimated in moving time windows we have chosen 3 parameters: entropy of distribution of squared orthogonal wavelet coefficients En and 2 multifractal parameters – singularity spectrum support width $\Delta\alpha$ and generalized Hurst exponent α^{*}.

Minimum normalized entropy En of squared wavelet coefficients. Let $x(t)$ be the finite sample of the signal $t = 1,...,N$ - index, numerating the counts. The normalized entropy is defined by the formula:

$$En = -\sum_{k=1}^{N} p_k \cdot \log(p_k) / \log(N), \quad p_k = c_k^2 / \sum_{j=1}^{N} c_j^2, \quad 0 \leq En \leq 1 \tag{6.17}$$

Here $c_k, k = 1, N$ are the orthogonal wavelet coefficients which were found from minimization of the value (6.17). We try 17 orthogonal wavelets [Mallat 1998]: 10 usual wavelets of Daubechies (number of vanishing moments equals to integer numbers from 1 up to 10) and 7 so called symlets with numbers of vanishing moments varying from 4 up to 10. For geophysical monitoring time series the parameters En was estimated within adjacent "short" time windows of the certain length after removing trend by polynomial of some order. This operation provides one of the possible transformations of high-frequency initial time series to low-frequency series of its properties.

Minimum normalized entropy En was suggested in [Lyubushin, 2012] and was used for investigating seismic noise properties in [Lyubushin, 2013(a,b), 2014(a); Lyubushin et al., 2014]. This entropy measure has some common features with multiscale entropy which was introduced in [Costa et al., 2003, 2005] for analysis of time series. In particular orthogonal wavelet

transform of the signal, which is used in (6.17), is multiscale as well because it provides decomposition into discrete dyadic time-frequency "atoms" with energy which is equal to c_k^2.

Multifractal parameters $\Delta\alpha$ and α^*. Let $x(t)$ be a random signal. Let us define its measure of variability $\mu_x(t,\delta)$ on the time interval $[t,t+\delta]$ as the difference between maximum and minimum values $\mu_x(t,\delta) = \max\limits_{t \le u \le t+\delta} x(u) - \min\limits_{t \le u \le t+\delta} x(u)$ and calculate the mean value of its power degree q: $M(\delta,q) = M[(\mu_x(t,\delta))^q]$. A random signal is scale-invariant [Taqqu, 1988] if $M(\delta,q) \sim \delta^{\rho(q)}$ when $\delta \to 0$, that is, the following limit exists:

$$\rho(q) = \lim_{\delta \to 0}\left(\ln M(\delta,q)/\ln \delta\right) \tag{6.18}$$

If $\rho(q)$ is a linear function $\rho(q) = Hq$, where $H = const, 0 < H < 1$, the process is called monofractal. In the case where $\rho(q)$ is a nonlinear concave function of q, the signal is called multifractal. To estimate the value of $\rho(q)$ using a finite sample $x(t), t = 0,1,...,N-1$ we used the method, which is based on the approach of detrended fluctuation analysis (DFA) [Kantelhardt et al., 2002]. Let us split the entire time series into non-overlapping intervals of length s:

$$I_k^{(s)} = \{t : 1 + (k-1)s \le t \le ks, \quad k = 1,...,[N/s]\} \tag{6.19}$$

and let

$$y_k^{(s)}(t) = x((k-1)s+t), \quad t = 1,...,s \tag{6.20}$$

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 the local trend:

$$\Delta y_k^{(s,m)}(t) = y_k^{(s)}(t) - p_k^{(s,m)}(t), \quad t = 1,...,s \tag{6.21}$$

and calculate the values

$$Z^{(m)}(q,s) = \left(\sum_{k=1}^{[N/s]} (\max_{1 \le t \le s} \Delta y_k^{(s,m)}(t) - \min_{1 \le t \le s} \Delta y_k^{(s,m)}(t))^q \middle/ [N/s]\right)^{1/q} \tag{6.22}$$

that can be regarded as the estimate of $(M(\delta_s,q))^{1/q}$. Let us define the function $h(q)$ as a coefficient of linear regression between $\ln(Z^{(m)}(q,s))$ and $\ln(s)$: $Z^{(m)}(q,s) \sim s^{h(q)}$ fitted for scales range $s_{min} \le s \le s_{max}$. It is evident that $\rho(q) = qh(q)$ and, for a monofractal signal, $h(q) = H = const$. The multifractal singularity spectrum $F(\alpha)$ is equal to the fractal dimensionality of the set of time moments t for which the Hölder-Lipschitz exponent is equal to α i.e. for which $|x(t+\delta) - x(t)| \sim |\delta|^\alpha$, $\delta \to 0$ [Feder, 1988]. The singularity spectrum can be estimated using the standard multifractal formalism, which consists in calculating the Gibbs sum:

$$W(q,s) = \sum_{k=1}^{[N/s]} (\max_{1 \le t \le s} \Delta y_k^{(s,m)}(t) - \min_{1 \le t \le s} \Delta y_k^{(s,m)}(t))^q \tag{6.23}$$

and in estimating the mass exponent $\tau(q)$ from the condition $W(q,s) \sim s^{\tau(q)}$. From (6.22) it follows that $\tau(q) = \rho(q) - 1 = qh(q) - 1$. In the next step, the spectrum $F(\alpha)$ is calculated with the Legendre transform:

$$F(\alpha) = \max \left\{ \min_{q}(\alpha q - \tau(q)), 0 \right\} \qquad (6.24)$$

If the singularity spectrum $F(\alpha)$ is estimated in a moving window, its evolution can give useful information on the variations in the structure of the "chaotic" pulsations of the series. In particular, the position and width of the support of the spectrum $F(\alpha)$, i.e., the values $\alpha_{min}, \alpha_{max}$, $\Delta\alpha = \alpha_{max} - \alpha_{min}$, and α^*, such that $F(\alpha^*) = \max_{\alpha} F(\alpha)$, are characteristics of the noisy signal. The value α^* can be called a generalized Hurst exponent and it gives the most typical value of Lipschitz-Holder exponent. Parameter $\Delta\alpha$, singularity spectrum support width, could be regarded as a measure of variety of stochastic behavior. In the case of a monofractal signal, the quantity $\Delta\alpha$ should vanish and $\alpha^* = H$. Usually $F(\alpha^*) = 1$, but there exist time windows for which $F(\alpha^*) < 1$. Estimates of minimum Hölder-Lipschitz exponent α_{min} are mainly positive. Nevertheless negative values of α_{min} are quite possible as well [Telesca et al., 2005; Currenti et al, 2005; Telesca, Lovallo, 2011; Chandrasekhar et al., 2016] for time fragments which are characterized by high-amplitudes spikes and steps.

In the calculation of $\Delta\alpha$ and α^* we were guided by the following considerations. The exponent q in the formula (6.23) was varying within the interval $q \in [-Q, +Q]$ where Q is a certain sufficiently large number, for example $Q = 10$. For each probe value of α within interval $\alpha \in [A_{min}, A_{max}]$ where $A_{min} = \min_{q \in [-Q,+Q]} d\tau(q)/dq$ and $A_{max} = \max_{q \in [-Q,+Q]} d\tau(q)/dq$ we calculated the value $\tilde{F}(\alpha) = \min_{q \in [-Q,+Q]} (\alpha q - \tau(q))$. If the value of α is close to A_{min} then $\tilde{F}(\alpha) < 0$, and this value is unsuitable as an estimate of the singularity spectrum, which must be non-negative. However, beginning from some certain α, the value of $\tilde{F}(\alpha)$ becomes non-negative, and this condition defines the α_{min} value. At a further α increase, the value $\tilde{F}(\alpha)$ increases, reaches its maximum when $\alpha = \alpha^*$, then begins to decrease, and finally, attains a certain value $\alpha_{max} < A_{max}$, at which $\tilde{F}(\alpha)$ again becomes negative for $\alpha > \alpha_{max}$. Thus, the condition $\tilde{F}(\alpha) \geq 0$ determines the interval of singularity spectrum support $\alpha \in [\alpha_{min}, \alpha_{max}]$, where $F(\alpha) = \tilde{F}(\alpha)$. The derivative $d\tau(q)/dq$ is calculated numerically from the values of $\tau(q)$, $q \in [-Q, +Q]$, and the accuracy of its calculation is of little significance, because this derivative is used for a rough determination of an a priori interval of possible values of exponent q. Minimum value of scale s_{min} within formulae (6.22-6.23) was chosen 20 samples, maximum scale equals $s_{max} = N/5$.

Multifractal analysis is a rather popular tool in geophysical studies [Ramirez-Rojas et al. 2004; Ida et al. 2005; Currenti et al. 2005; Telesca et al. 2005; Lyubushin et al. 2012, 2014; Chandrasekhar et al., 2016]. In the paper [Lyubushin et al., 2012] the multifractal analysis of geo-mechanical monitoring time series was applied for its fragmentation into intervals with different behaviour. In the papers [Lyubushin, 2008(b), 2009, 2010(a,b), 2011(a,b), 2012, 2013(a,b), 2014(a,b)] estimates of multifractal properties $\Delta\alpha$, α^* and α_{min} of low-frequency seismic noise were used for the purposes of earthquake prediction and dynamic estimate of seismic danger.

First principal component

There is a necessity to aggregate the used sequences of properties of time series such as $(En, \Delta\alpha, \alpha^*)$ into time series of scalar characteristics, which carries the most common properties from the set of initial properties. We used here the most popular approach of principal components [Jolliffe, 1986]. Let $P(t) = (P_1(t),...,P_m(t))^T$, $t = 0,1,...$ be a multiple time series of dimensionality m. Let L be a number of samples within time window, which is moving from left to right with minimum mutual shift 1 which we will name a "window of adaptation". Let s be a number of the sample corresponding to right-hand end of moving time window. It means that time window contains samples with time indexes t which obey the condition $s - L + 1 \leq t \leq s$. Let's calculate a correlation matrix $\Phi(s)$ of the size $m \times m$ within each time window after normalization of multiple time series components:

$$\Phi(s) = \left(\varphi_{ab}^{(s)}\right), \quad \varphi_{ab}^{(s)} = \sum_{t=s-L+1}^{s} q_a^{(s)}(t) q_b^{(s)}(t) / L, \quad a,b = 1,...,m \tag{6.25}$$

where

$$q_a^{(s)}(t) = (P_a(t) - \overline{P}_a^{(s)}) / \sigma_a^{(s)}, \quad \overline{P}_a^{(s)} = \sum_{t=s-L+1}^{s} P_a(t) / L,$$

$$\left(\sigma_a^{(s)}\right)^2 = \sum_{t=s-L+1}^{s} (P_a(t) - \overline{P}_a^{(s)})^2 / (L-1), \quad a = 1,...,m \tag{6.26}$$

First principal component $\psi^{(s)}(t)$ is calculated using formula

$$\psi^{(s)}(t) = \sum_{\alpha=1}^{m} \theta_a^{(s)} \cdot q_a^{(s)}(t) \tag{6.27}$$

Here m-dimensional vector $\theta^{(s)} = (\theta_1^{(s)},...,\theta_m^{(s)})^T$ is an eigenvector of the correlation matrix $\Phi(s)$ corresponding to its maximum eigenvalue. Let us define a scalar time series of first principal components $\psi(t)$ within window of adaptation of the length L samples by the rule:

$$\psi(t) = \begin{cases} \psi^{(L-1)}(t) \text{ for } 0 \leq t \leq (L-1) \\ \psi^{(t)}(t) \text{ for } t \geq L \end{cases} \tag{6.28}$$

Thus, within the 1st time window of adaptation time series $\psi(t)$ is composed of values calculated by (6.27) whereas for all further time indexes $\psi(t)$ equals to the value (6.27) corresponding to the most right-hand end of time window, i.e. outside the 1st window of adaptation $\psi(t)$ depends on past values of $P(t)$ only.

Properties of global low-frequency seismic noise

Microseismic oscillations in a wide frequency range are one of the most frequently investigated topics of geophysical studies. This is due to their accessibility, the presence of numerous regional and global seismic networks, and the well-developed practice of seismic observations. Even an approximate review of the literature devoted to analysis of microseisms apparently cannot be made. This is particularly true of the analysis of high frequency (HF) microseisms (from 0.01 to 100 Hz and higher, up to seismoacoustic waves). The widespread occurrence of HF microseismic observations is due to the relative simplicity and mobility of instrumentation

free from rigid requirements on long-term stability of sensors that can by no means be neglected in problems of low frequency (LF) geophysical monitoring. In the paper [McNamara and Buland, 2004] the results were presented of detailed research into microseismic background of natural and industrial origin in the frequency band 0.01–16 Hz, including the construction of estimators for the temporal (diurnal and seasonal) and spatial distribution of power spectrum properties. More recent studies on the composition of short-period microseisms are presented in [Koper and de Foy, 2008; Koper et al., 2010]. With an increase in the period of microseismic background oscillations studied, the role of atmospheric and oceanic waves as main sources of microseisms becomes predominant. The paper [Berger et al., 2004] presented a review of the use of IRIS broadband seismic stations for the study of background microseisms. Microseismic oscillations in the period range 5–40 s were studied in the paper [Stehly et al., 2006], where their oceanic origin was established. Continuously observed microseismic oscillations at periods of 100–500 s were examined in [Friedrich et al., 1998; Kobayashi and Nishida, 1998; Tanimoto, 2001, 2005; Ardhuin et al., 2011]. These oscillations are generated both by weak earthquakes and by processes in the atmosphere and ocean. In the papers [Aster et al., 2008; Grevemeyer et al., 2000; Kedar et al., 2008; Schimmel et al., 2011] variability of field of microseisms due to climate change and ocean processes were studied.

In order for earthquakes to be a source of continuously present microseismic oscillations, at least one earthquake with a magnitude of 6 should occur daily to maintain the observed intensity of such oscillations. The cumulative effect of all weak earthquakes estimated from the Gutenberg–Richter recurrence law yields an energy contribution one to two orders smaller than the observed value. The effect of atmospheric processes (movement of cyclones) and oceanic waves generated by them, as well as the impact of the waves on the shelf and coasts, contributes most to the energy of the low-frequency (LF) microseismic background. The origin of an LF seismic hum with a predominant period of 4 min was studied in [Rhie and Romanowicz, 2004, 2006]. A significant correlation was established between the intensity of these oscillations and the storm wave height in the oceans, and it was shown that the hum intensity is independent of the Earth's seismic activity: the authors presented an example of a seismically quiet time interval (January 31 – February 3, 2000) characterized, however, by anomalously high amplitudes of microseismic background in the vicinity of the 4-min period. As a possible mechanism of excitation of such oscillations, they proposed the perturbation of the gravitational field by high waves resulting in the excitation of LF seismic waves on the sea floor. The main regions of excitation of these oscillations are suggested to be the North Pacific Ocean in winter and the southern Atlantic Ocean in summer. This frequency range of the ambient seismic noise ("seismic hum") was investigated in [Fukao et al., 2010; Nishida et al., 2008, 2009].

In spite of the fact that the main source of energy for LF microseisms is an external one with respect to the Earth's crust, and the latter is merely the propagation medium, the conditions in the Earth's crust affect the statistical characteristics and the specific features in the behavior of LF microseismic vibrations. Consequently, if we study the time variations of the characteristics of seismic noise, this study will hopefully yield important information concerning the changes in the Earth's crust, including those linked with the seismic process and with the preparation of strong earthquakes.

This basically simple idea of the use of low-frequency microseismic oscillations for monitoring the lithosphere, nevertheless, cannot be realized in a simple way. The main difficulty consists in a strong influence of numerous uncorrelated sources of the data. These sources are often diffusely distributed over the Earth's surface. Therefore, it is impossible in this case to investigate the transmitting properties of the lithosphere by controlling input actions and responses. Additionally, the division into "a signal" and "noise," which is typical of the traditional methods used for data analysis, loses its sense, when microseismic oscillations are

processed. Only tidal variations in the amplitude of microseisms, as well as the arrivals and coda from the well-known strong earthquakes, can be related to "signals." These signals have been long and traditionally used in geophysics. All other microseism variations relate to "noise."

In this section of the chapter the main tool for overcoming the influence of uncorrelated random sources is using $(En, \Delta\alpha, \alpha^*)$ statistics calculated within adjacent "short" time fragments. Thus, a seismic noise records are transformed into time series with "big" sampling time step, 1 day for instance. These time series are much more correlated and are more suitable for investigating of synchronization effects.

The seismic records were taken by requests to Incorporated Research Institutions for Seismology (IRIS) data base by the address http://www.iris.edu/forms/webrequest/ from 229 seismic stations of 3 global broadband seismic networks:

Global Seismographic Network: http://www.iris.edu/mda/_GSN
GEOSCOPE: http://www.iris.edu/mda/G
GEOFON: http://www.iris.edu/mda/GE

Vertical components with sampling rate 1 Hz (LHZ-records) were downloaded for 20 years of observation since 01 Jan 1997 up to 31 Dec 2016. The initial LHZ-records were transformed to sampling time step 1 minute by calculating mean values within successive time intervals of the length 60 seconds. A further analysis is based on estimating statistical properties of low-frequency seismic noise waveforms (periods exceeding 2 minutes) within successive daily time intervals of the length 1440 samples with time step 1 minute.
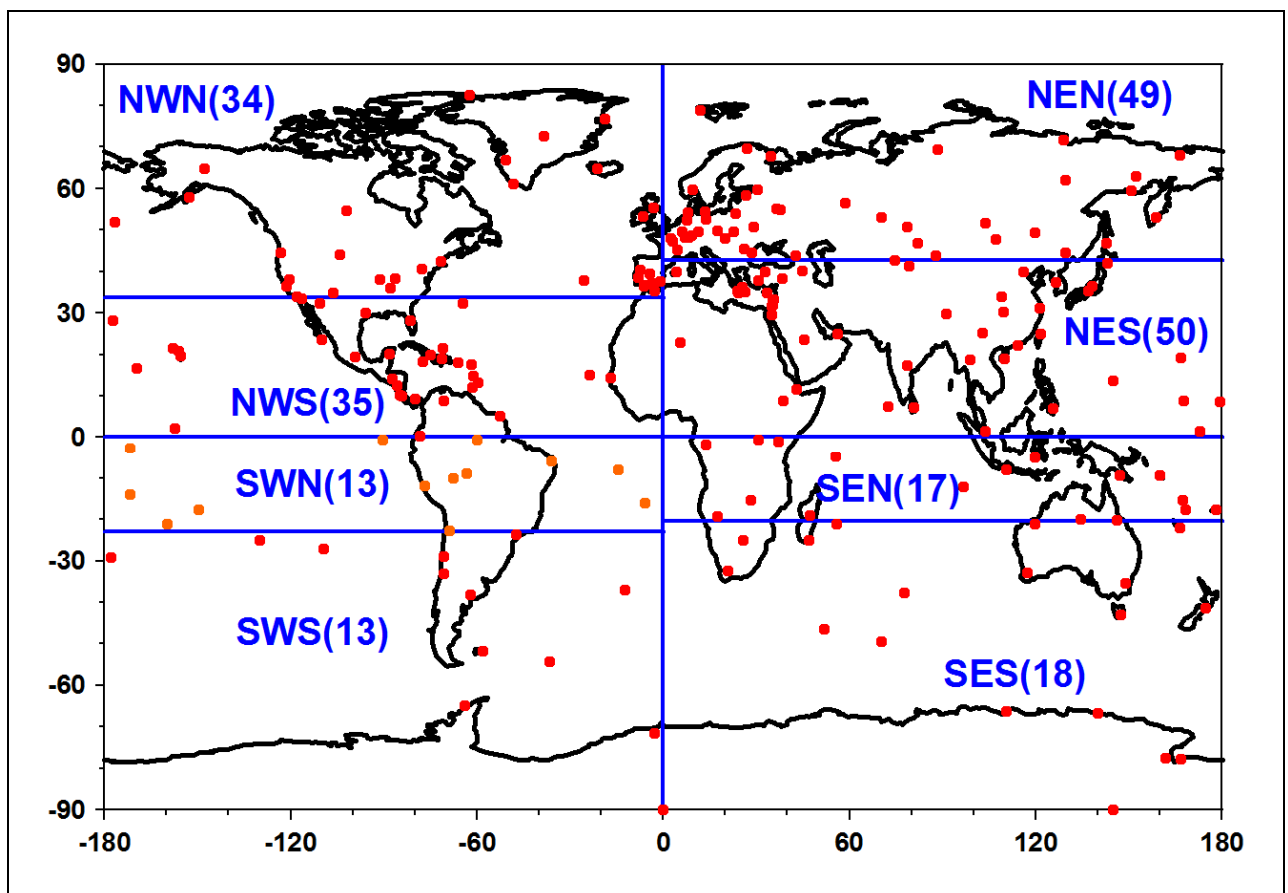

Figure 6-1. Positions of 229 broadband seismic stations and their splitting into 8 groups with number of stations in each group in brackets.

Fig.6-1 presents positions of 229 broadband seismic stations all over the world and their splitting into 8 groups of stations. Each group has 3-letters identification code and the number of stations within each group is given in brackets. The names of the groups have the following abbreviation sense: the first letter is "N" or "S" what means North or South. The second letter is "E" or "W" what means East or West. Thus, initially all station were divided into 4 parts by splitting into North-East, North-West, South-East and South-West quarter-spheres. Finally each of 4 parts was split into North and South parts (the third letter is "N" or "S") by the rule that the number of stations within each part must be approximately equal each other.

The seismic records from each station after coming to 1 minute sampling time step were split into adjacent time fragments of the length 1 day (1440 samples) and for each fragment 3 parameters of low-frequency daily seismic noise waveforms were calculated. Two of them are multifractal parameters: generalized Hurst exponent α^* and singularity spectrum support width $\Delta\alpha$. For removing scale-dependent trends (which are mostly caused by tidal variations) in method of singularity spectrums estimates a local polynomials of the 8-th order were used in the formula (6-21). Other seismic noise parameter is minimum normalized entropy En of squared orthogonal wavelet coefficients. Before computing entropy En in each daily time window a polynomial trends of 8th order were removed from seismic noise waveforms. Thus, time series of α^*, $\Delta\alpha$ and En values with sampling time step 1 day were obtained from each of 229 seismic stations which are presented at the Fig.6-1. The Fig.6-2 illustrates the sequence of data transform operations.
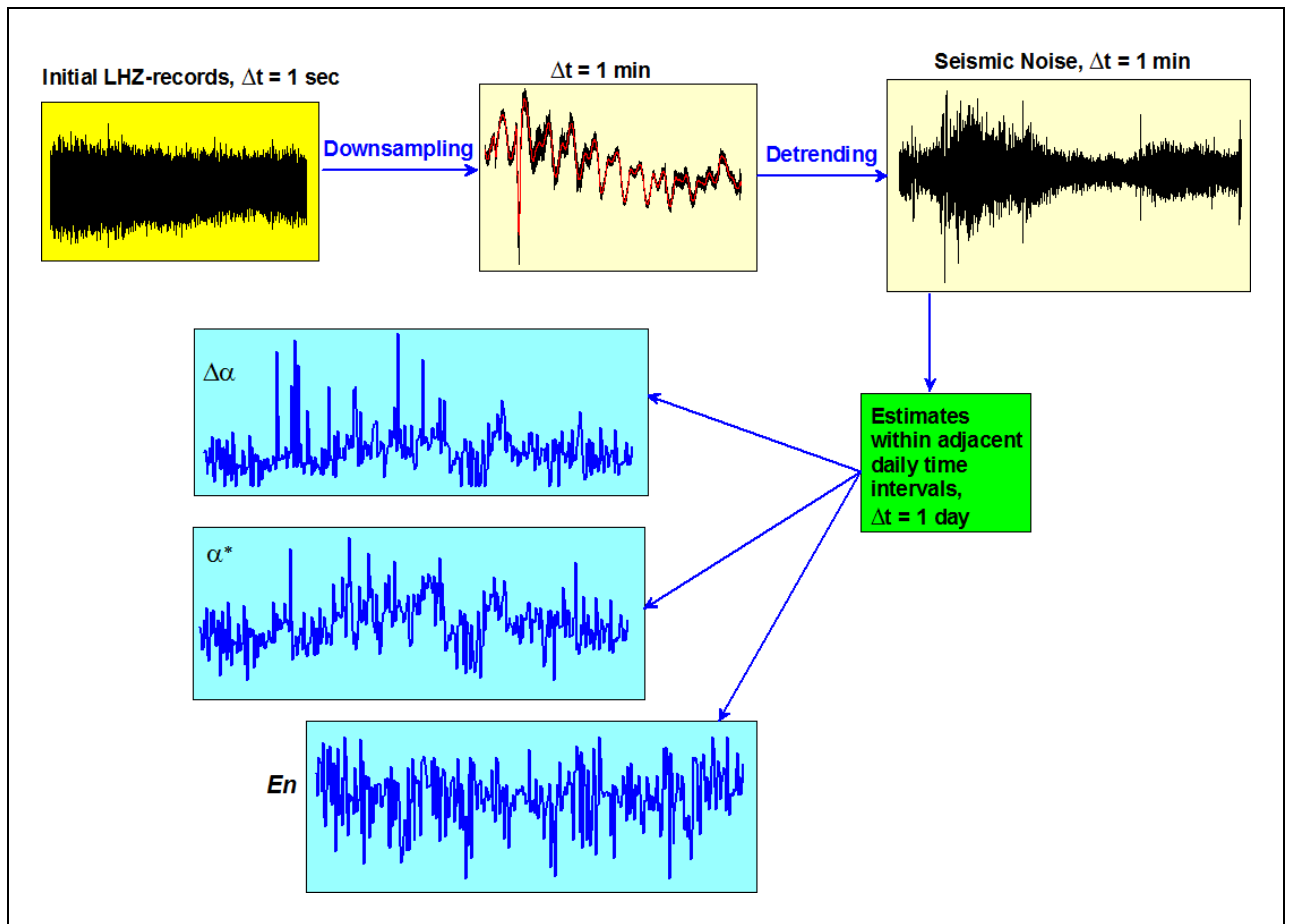
Figure 6-2. Scheme of data transform from initial seismic records with sampling rate 1 Hz to time series of multifractal and entropy properties with sampling time step 1 day.
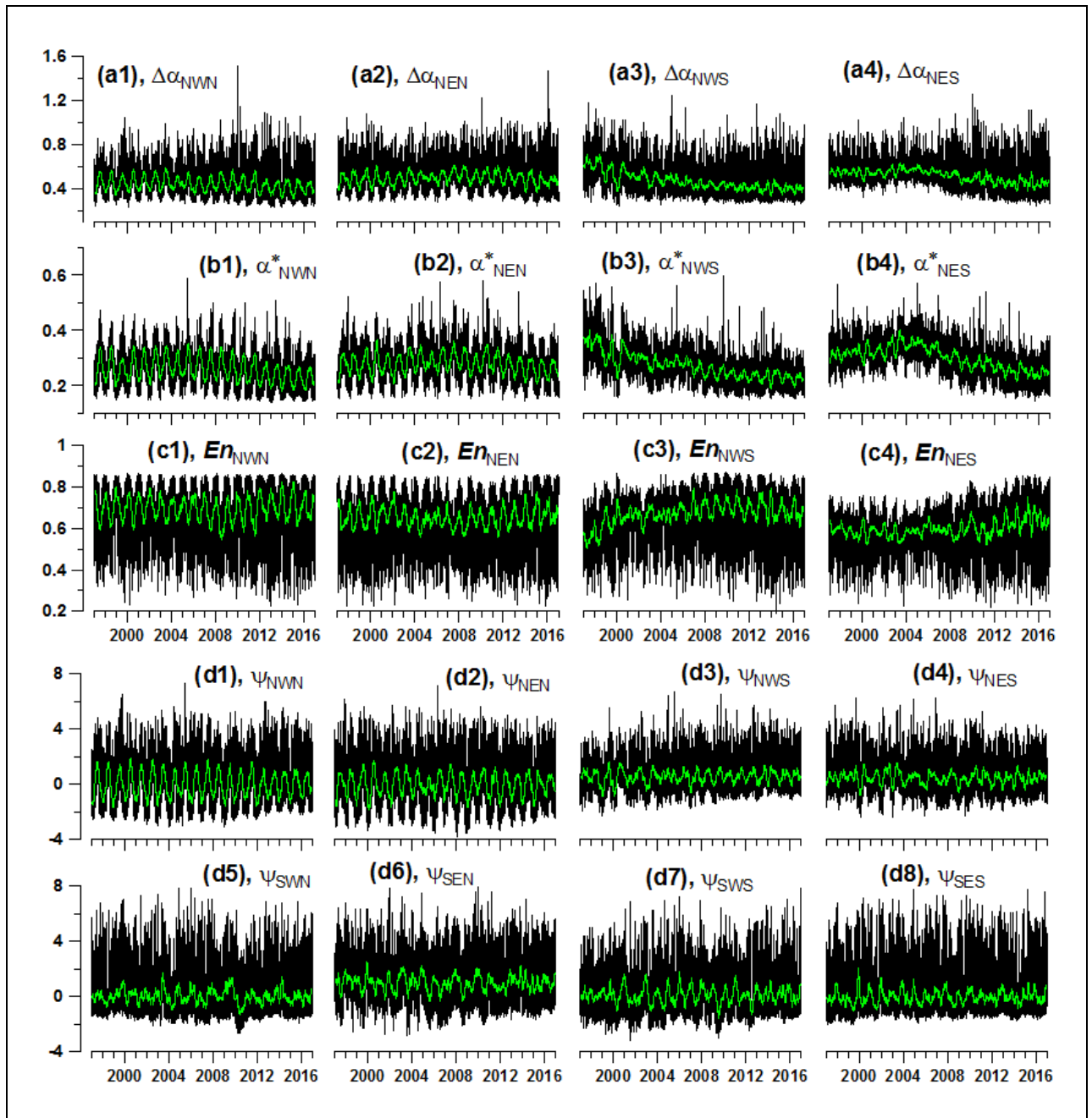
Figure 6-3. ((a1)–(a4), black lines - plots of daily median values of for four regions (NWN, NEN, NWS and NES) in the Northern hemisphere (Fig.6–1); (b1)–(b4), black lines - plots of daily median values of for regions NWN, NEN, NWS and NES; (c1)–(c4), black lines - plots of daily median values of for regions NWN, NEN, NWS and NES; (d1)–(d4), black lines - plots of values of first principal component , calculated for daily median values of for regions NWN, NEN, NWS and NES respectively within the length of adaptation of 365 days; (d5)–(d8) as for (d1)–(d4) but calculated for four regions in the Southern hemisphere SWN ("South-West-North"), SEN ("South-East-North"), SWS ("South-West-South") and SES ("South-East-South") (Fig.6–1). Bold green lines represent the running average within the time window of 57 days for each curve.

Fig.6-3 presents graphics of daily median values of multifractal singularity spectrum support width $\Delta\alpha$, generalized Hurst exponent α^* and minimum normalized entropy En of squared orthogonal wavelet coefficients for 4 regions NWN ("North-West-North"), NEN ("North-East-North"), NWS ("North-West-South") and NES ("North-East-South") in Northern hemisphere and first principal components calculated by using formula (28) with length of adaptation 365 days from daily time series of $(En, \Delta\alpha, \alpha^*)$ for all 8 regions both from Northern and Southern hemispheres. From the plots of moving averages in window of the length 57 days it could be noticed that a strong annual periodicity is characterized for all regions especially for NWN and

NEN. We suppose that this periodicity is caused by influence of periodic seasons of strong oceanic storms which are known as the important source of energy for permanent seismic noise [Rhie and Romanowicz, 2004, 2006; Fukao et al., 2010; Nishida et al., 2008, 2009]

First principal components time series from 8 parts of the world compose multiple time series, which is the object of applying wavelet-based and spectral coherence measures with a purpose to detect coherence effects in global seismic noise properties. The length of moving time window 365 days is quite natural for such application.
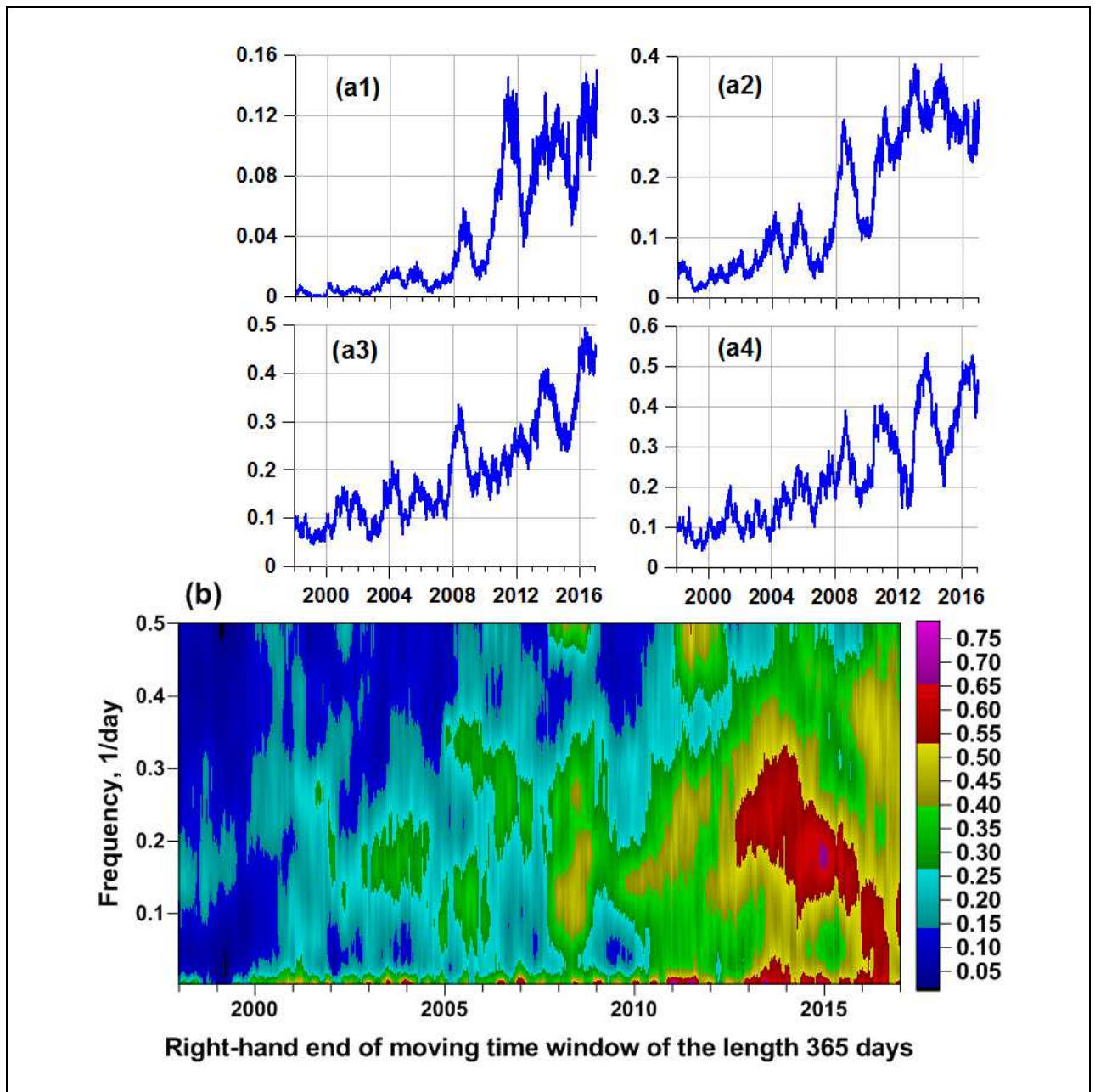
Figure 6-4. (a1)-(a4) – plots of multiple wavelet-based coherence measure for 8-dimensional time series of first principal components (Fig.6-3, (d1)-(d8)) for detail levels 1-4 depending on right-hand end of moving time window of the length 365 days; (b) – time-frequency diagram of multiple spectral measure of coherence for the same 8-dimensional time series estimated within moving time window of the length 365 days.

Fig.6-4 presents results of estimating measures of coherence. Fig.6-4(a1)-(a4) demonstrate evolution of the wavelet-based measure $\rho(\tau, \beta)$ (formula (7)) with using of Haar wavelets for 4 detail levels with range of scales 2-4, 4-8, 8-16 and 16-32 days correspondently. The occurrence

of first 4 detail levels of wavelet decomposition is a consequence of the length 365 samples of time window and the significance threshold $L_{min} = 16$. The graphs Fig.4(a1)-(a4) are plotted versus time position of right-hand end of moving time window. The main peculiarity of these graphs is increasing of coherence, which is observed starting from the window position 2007-2008. This increasing is not gradual and have rather strong fluctuations with time scale 2-3 years at the background of general positive trend. The time-frequency diagram of spectral measure of coherence $\lambda(\tau, \omega)$ (formula (6-15)) at the Fig.6-4(b) confirms this conclusion. The diagram Fig.6-4(b) was obtained by estimating $\lambda(\tau, \omega)$ within moving time window of the same length 365 samples using vector autoregression model (6.12) of the 5th order. Thus, using both wavelets (compact basic functions) and Fourier spectral approach independently extract the same effect of coherence increasing for variations of seismic noise properties from the stations all over the world.

Table 1 presents information about 20 strongest earthquakes, which occurred from beginning of 20th century.

Table 1. Strongest earthquakes, $M \geq 8.4$, from the beginning of 20th century Source: https://earthquake.usgs.gov/earthquakes/search/							
Date	Magnitude	Latitude	Longitude	Date	Magnitude	Latitude	Longitude
1906.01.31	8.8	0.955	-79.369	1963.10.13	8.5	44.872	149.483
1922.11.11	8.5	-28.293	-69.852	1964.03.28	9.2	60.908	-147.339
1923.02.03	8.4	54.486	160.472	1965.02.04	8.7	51.251	178.715
1933.03.02	8.4	39.209	144.59	**2001.06.23**	8.4	-16.265	-73.641
1938.02.01	8.5	-5.045	131.614	**2004.12.26**	9.1	3.295	95.982
1946.04.01	8.6	53.492	-162.832	**2005.03.28**	8.6	2.085	97.108
1950.08.15	8.6	28.363	96.445	**2007.09.12**	8.4	-4.438	101.367
1952.11.04	9	52.623	159.779	**2010.02.27**	8.8	-36.122	-72.898
1957.03.09	8.6	51.499	-175.626	**2011.03.11**	9.1	38.297	142.373
1960.05.22	9.5	-38.143	-73.407	**2012.04.11**	8.6	2.327	93.063

According to information from the Table 1 among 20 strongest earthquakes which occurred from beginning of 20th century 7 events took place starting from middle of 2001 from which 4 events – after September 2007. It turns out that the increasing of coherence of global seismic noise properties coincides with dramatic increasing of strongest earthquakes rate, which is observed starting from Sumatra mega-earthquake at 26 Dec of 2004, especially starting from 2007. Taking into account that we investigate a range of periods from 2 minutes up to 500 minutes this coherence increasing could not be the direct consequence of aftershocks of strongest earthquakes. Our hypothesis is that slow movements of small Earth's crust blocks are synchronized in the regions of preparing huge earthquakes [Lyubushin 2009, 2010(b), 2011(a,b), 2012, 2013(a,b)] and we see that this synchronization is a global phenomenon starting from the beginning of 2000s [Lyubushin, 2014(a), 2015]. The increasing of seismic noise synchronization is observed till now, what could be a precursor of strongest earthquakes occurrence in the near future.

Low-frequency seismic noise at Japan islands

The main attention in this section of the chapter will be focused on the processing seismic noise data from network at Japan islands. This peculiarity is following from the fact that one of the strongest mega-earthquakes with magnitude 9.1 in the latest period of instrumental seismology had happened on 11 March, 2011 in Japan which is the region with extremely dense network of geophysical observation which are at open access via internet. Such combination gives a unique possibility to test different hypothesizes about the ways in which processes of seismic catastrophe preparation influence on the statistical properties of seismic noise in the active region.

For the analysis a vertical broadband seismic oscillations components with 1-second sampling time step (LHZ-records) from the broadband seismic network F-net stations in Japan were downloaded from internet address http://www.fnet.bosai.go.jp starting from the beginning of 1997 up to August 31, 2017. The whole list of F-net seismic stations includes 84 positions. We considered the stations, which are located northward from 30°N and, thereby excluding the data from six solitary stations located on remote small islands. The locations of 78 stations, which were chosen for analysis are shown in Fig.6-5 with epicenters of 2 the strongest earthquakes which occurred during observations: near Hokkaido at 25 of September 2003 with magnitude 8.3 and Tohoku mega-earthquake on March 11, 2011 with magnitude 9.1. These 78 stations were split into 5 clusters which are presented at Fig.6-5 by numbers of clusters with different colors.
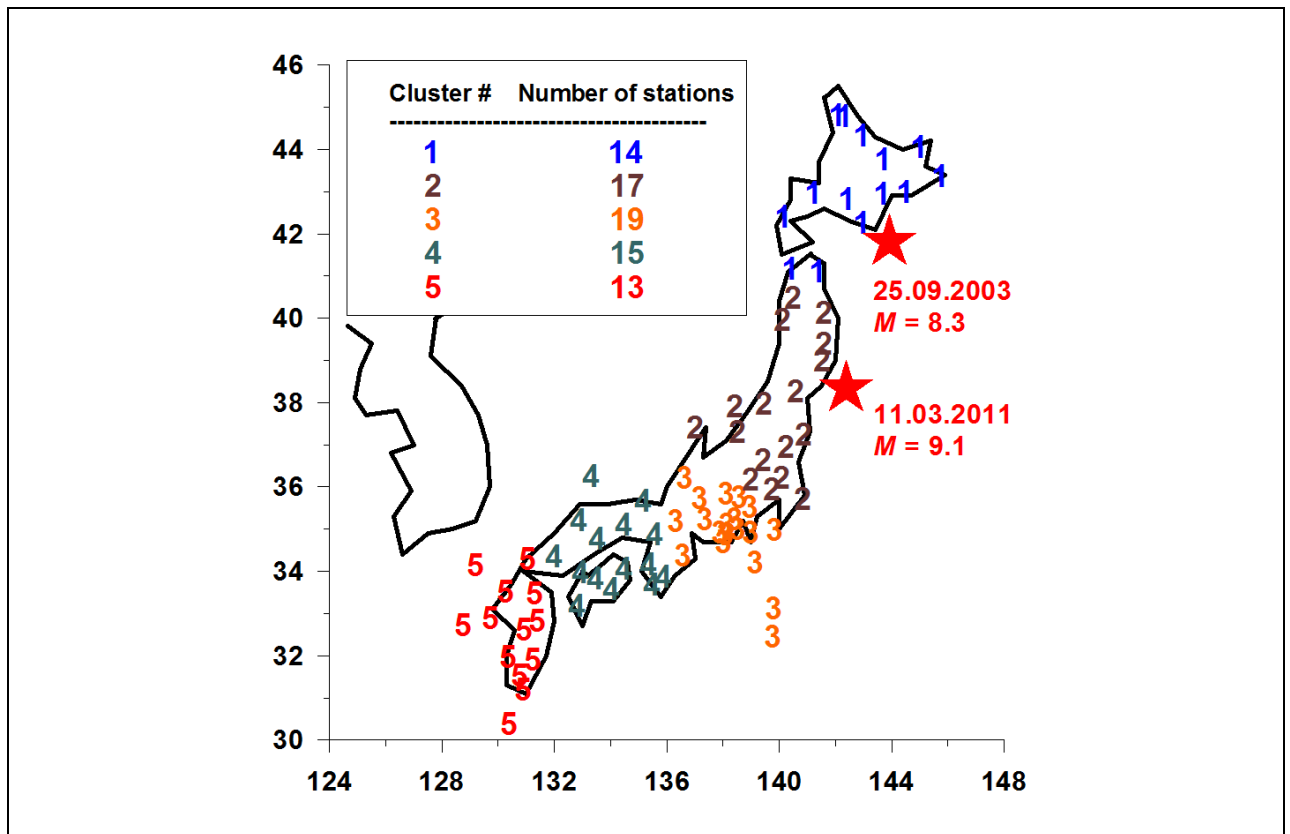

Figure 6-5. Positions of 78 broadband seismic stations of the network F-net in Japan and their splitting into 5 clusters. Positions of stations belonging to the same cluster are indicated by the same number from 1 up to 5. Number of stations within each cluster are given in the frame. Red stars shows epicenters of two strongest earthquakes since the beginning of 1997.

The F-net seismic data were analyzed after transforming them to sampling time step 1 minute by calculating mean values in adjacent time windows of the length 60 sec and further calculating values of $(En, \Delta\alpha, \alpha^*)$ within adjacent time windows of the length 1440 samples with 1-minute

sampling time step, i.e. of the length 1 day (see Fig.6-2).

Let us analyze in details time series of $\Delta\alpha$. For each time window we will calculate median values of $\Delta\alpha$ estimates from all operable stations and thus, scalar time series of median values of $\Delta\alpha^{(m)}$ will be obtained as integral characteristics of seismic noise from the whole network. Our special attention to $\Delta\alpha^{(m)}$ is following from the fact that reducing of singularity spectrum support width is known as the indicator of transforming the complex system to critical state with more simple structure of its ambient noise (less multifractal) preceding some catastrophic changes.
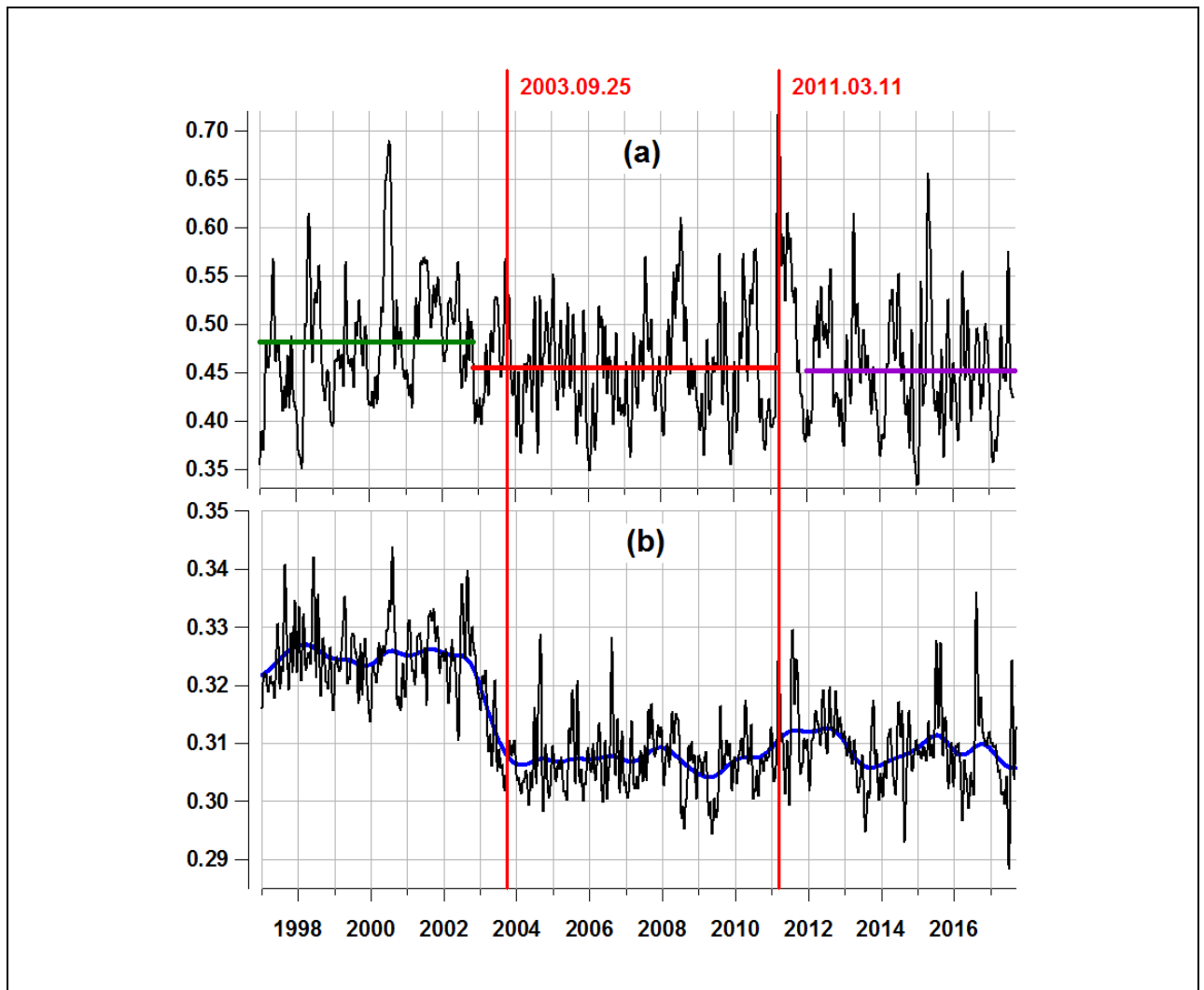
Figure 6-6. Graph (a) – plot of Gaussian kernel smoothing with averaging radius 13 days of median values of singularity spectrum support width $\Delta\alpha^{(m)}$ from network F-net after coming to sampling time step 1 minute within adjacent time windows of the length 1 day. Vertical red lines indicate time moments of earthquakes on September 25, 2003, $M = 8.3$ and on March 11, 2011, $M = 9.1$. Parallel bold lines of green, red and purple colors show mean values for time intervals 1997.01.01-2002.11.08, 2002.11.09-2011.03.10 and 2012.01.01-2017.08.31. Graph (b), black line – plot of Gaussian kernel smoothing with averaging radius 13 days of median of $\Delta\alpha$ from all F-net network for vertical seismic noise waveforms with sampling time step 1 sec within adjacent time windows of the length 30 minutes; graph (a), bold blue line – the same with averaging radius 0.5 years.

It should be noticed that using of multifractal singularity spectrum support width has a rather long history in investigation of nonlinear systems behavior. Particularly the "loss of multifractality" i.e. decreasing of singularity spectrum support width, is a well-known effect before the abrupt change of different system properties. Mainly this effect was investigated in

biological and medicine systems [Ivanov et al., 1999; Humeaua et al., 2008; Dutta et al., 2013], but in [Pavlov, Anishchenko, 2007] it was shown that it has a rather universal character and is observed in physical systems as well. The analogy between effect of singularity spectrum support narrowing of seismic noise waveforms and the loss of multifractality in the behavior of other nonlinear systems gave an impulse to the author to make a hypothesis about approaching Japanese island to seismic catastrophe [Lyubushin, 2008(a)].

Before analyzing peculiarities of $\Delta\alpha^{(m)}$ let us compute Gaussian kernel smoothing $\xi(t)$ [Hardle, 1990] of the signal $\Delta\alpha^{(m)}(t)$ by the formula

$$\xi(t\,|\,h) = \int \Delta\alpha^{(m)}(t + h\,s) \cdot \psi(s)\,ds, \quad \psi(s) = \exp(-s^2/2)/\sqrt{2\pi} \tag{6.29}$$

where $h > 0$ is a smoothing parameter which could be called averaging radius of Gaussian smoothing. The smoothed value of $\Delta\alpha^{(m)}$ with radius of averaging 13 days is presented at the Fig.6-6(a) by black line.

Two red vertical lines indicate time moments of strongest earthquakes and they split the history of $\xi(t\,|\,h)$ into sequence of three fragments. The 1st time fragment belongs to time interval before Hokkaido earthquake on September 25, 2003. It could be noticed that mean value of $\xi(t)$ for this fragment is bigger than mean value for other time interval before Tohoku earthquake on March 11, 2011. Let us find time point t_C of maximum change of mean values of $\xi(t)$ at the vicinity of time moment of Hokkaido earthquake by using Fisher criterion from analysis of variance (ANOVA) [Rao, 1965]. Let us calculate general mean value of $\xi(t)$ by all analyzed time interval $\bar{\xi}_0 = \sum_{t=1}^{N} \xi(t)/N$ where N is the general number of samples and mean values from left and right sides of probe time moment t_C: $\bar{\xi}_1 = \sum_{t=1}^{t_C} \xi(t)/t_C$ and $\bar{\xi}_2 = \sum_{t=t_C+1}^{N} \xi(t)/(N - t_C)$. Change point t_C is found from the condition

$$F(t_C) = S_1^2(t_C)\big/S_2^2(t_C) \rightarrow \max_{t_C} \tag{6.30}$$

where

$$S_1^2(t_C) = t_C \cdot (\bar{\xi}_1 - \bar{\xi}_0)^2 + (N - t_C) \cdot (\bar{\xi}_2 - \bar{\xi}_0)^2,$$
$$S_2^2(t_C) = \left(\sum_{t=1}^{t_C} (\xi(t) - \bar{\xi}_1)^2 + \sum_{t=t_C+1}^{N} (\xi(t) - \bar{\xi}_2)^2 \right)/(N - 2) \tag{6.31}$$

The using of criterion (6.30) detects day 2002.11.08 as the change point t_C with mean values 0.482 and 0.454 for $\bar{\xi}_1$ and $\bar{\xi}_2$. These mean values are shown at the Fig.6-6 by bold green and red parallel lines. This drop of $\Delta\alpha^{(m)}$ detects some transient geodynamic process to unstable state, which began with preparation of Hokkaido earthquake on September 25, 2003. After Tohoku earthquake, the smoothed value of $\Delta\alpha^{(m)}$ shows rapid increasing and this is a reaction to processes of stress relaxation in the Earth crust immediately after mega-earthquake. However, starting from the beginning of 2012 the mean smoothed value of $\Delta\alpha^{(m)}$ returns approximately to its previous level 0.452. It means that Tohoku mega-earthquake on March 11, 2011 did not return the region to its "normal" state as it was before Hokkaido earthquake and the situation of high seismic danger is continuing till now.

The difference between mean values of $\Delta\alpha$ before and after Hokkaido earthquake on 2003.09.25 becomes more explicit if they will be estimated for initial vertical seismic records with 1 sec sampling time step in the adjacent time windows of the length 30 minutes (1800 samples). For this case multifractal singularity spectra were calculated using DFA by removing local trends by polynomials of the 4th order. At the Fig.6-6(b) black line presents result of Gaussian smoothing of these $\Delta\alpha$ values with averaging radius 13 days whereas bold blue line – with using averaging radius 0.5 years. Fig.6-6(b) shows the same drop of $\Delta\alpha$ before the earthquake on 2003.09.25 as the Fig.6-6(a).
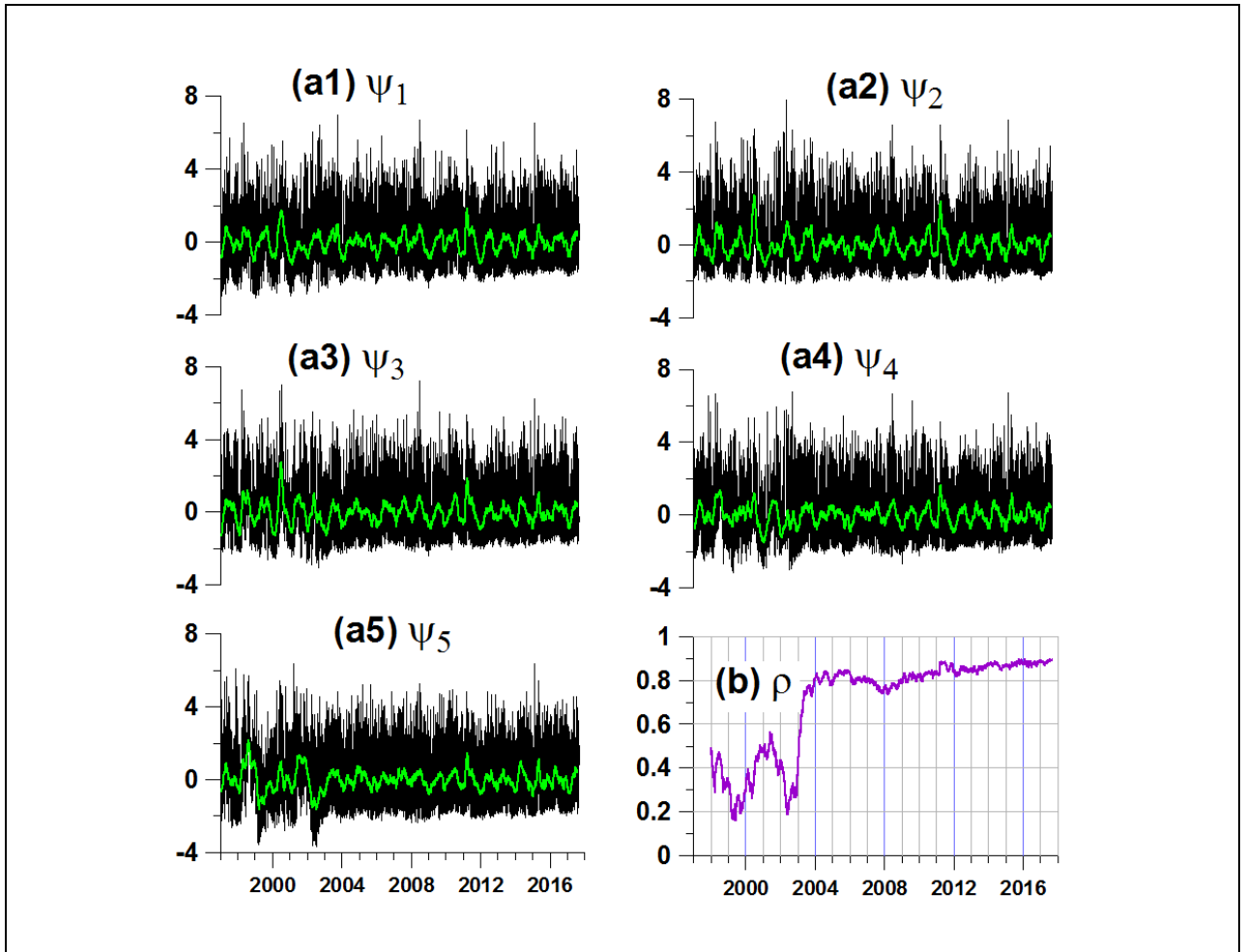
Figure 6-7. (a1)-(a5) – plots of values of first principal components ψ_k, $k = 1,..,5$, calculated for daily median values of multifractal singularity spectrum support width $\Delta\alpha$, generalized Hurst exponent α^* and minimum normalized entropy En of squared orthogonal wavelet coefficients for 5 clusters of stations in Japan (Fig.6-5) after coming to sampling time step 1 minute (look Fig.6-2) correspondently within length of adaptation 365 days; green lines present running average within time window of the length 57 days for each curve. Plot (b) presents values of multiple robust correlation coefficient ρ which was calculated within moving time window of the length 365 days. The values of ρ are plotted in dependence on time moments corresponding to right-hand end of moving time window.

Let us consider 5 groups of stations which are presented at the Fig.6-5. We are interested in time-dependent evolution of coherence effects between median values of daily seismic noise properties calculated from these 5 parts of the network. For this purpose let us compute 1st principal components $\psi_k(t)$, $k = 1,...,5$, from median values of parameters $(En, \Delta\alpha, \alpha^*)$ estimated daily from all operational stations within each cluster of stations at the Fig.6-5. These 1st principal components were computed within moving window of adaptation of the length 365

days using formula (6.28). Plots at Fig.6-7(a1)-(a5) present graphs of these 1st principal components. The simplest measure of coherence is the squared robust multiple correlation coefficient ρ (6.8) which is shown at the Fig.6-7(b), estimated in the moving time window of the same length 365 days. We see that the value of ρ demonstrates rapid increasing for time windows laying in the interval 2002-2004, i.e. at the vicinity of time moment of Hokkaido earthquake on September 25, 2003. This peculiarity of seismic noise correlation independently confirms the conclusion, which was made from the analysis of the Figure 6: 2002 was a year of rapid changing in the structure of seismic noise at Japan islands which is an indicator of preparing future seismic catastrophe on March 11, 2011. At this sense, the Hokkaido earthquake 2003.09.25 could be interpreted as a foreshock for Tohoku earthquake 2011.03.11 nevertheless of its strength. If we continue comparing Figures 6 and 7 we can notice one more conformity: the Tohoku mega-earthquake has a short-term response both in the median value of Δα and in the value of multiple correlation ρ. It means that Tohoku mega-earthquake has not changed the correlation structure of seismic noise, and moreover, this correlation becomes more strong because the value of ρ has a slight positive trend after 2012.
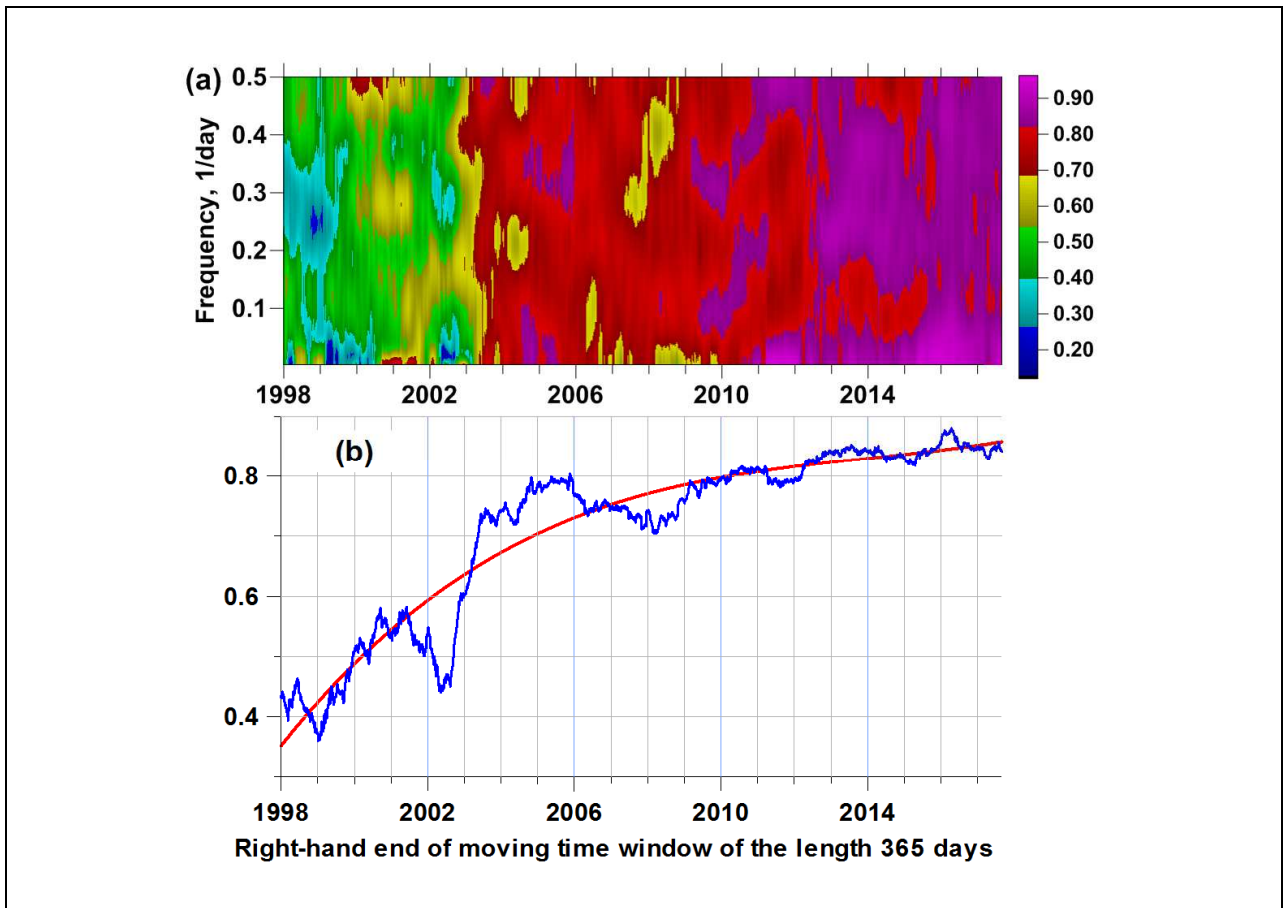

Figure 6-8. (a) – time-frequency diagram of multiple spectral measure of coherence for the 5-dimensional time series of first principal components $\psi_k, k = 1,..,5$ presented at Fig.6-7 ((a1)-a(5)) estimated within moving time window of the length 365 days; (b) – plot of mean values $\lambda_{mean}(\tau)$ of multiple spectral coherence calculated by averaging all frequency-dependent values within each time window; bold red line presents best-fitted polynomial of the 3rd order.

It is interesting to obtain independent confirmation about conclusion on growth of seismic noise correlation using other approach – estimate of coherence with the help of multiple spectral measure (6.10). For this purpose, let us consider 5-dimensional time series of 1st principal components $\psi_k(t), k = 1,...,5$ and calculate multiple spectral coherence measure (6.10) within moving time window of the length 365 days using vector autoregression model (6.12) of 5th

order. Time-frequency 2D diagram at Figure 8(a) presents result of such estimate, whereas Fig.6-8(b) shows graph of time-dependent averaging $\lambda_{mean}(\tau)$ of the coherence measure (6.10) by using all frequency values within each time window (formula (6.16)).

Results of coherence estimates presented at Fig.6-8 confirms the conclusion, which was made from the analysis of behavior of multiple correlation ρ at Fig.6-7(b): seismic noise coherence effects became more strong after transient process at 2002-2004 with slightly positive trend which is continuing till current time despite the occurrence of Tohoku mega-earthquake on 2011.03.11. Using of time-frequency diagram at Fig.6-8(a) gives a possibility to visualize frequency decomposition of coherence growth and we can see that this increasing practically is independent on the frequency band.

One of the used seismic noise parameters is normalized entropy of squared wavelet coefficients En. This quantity is a kind of antipode to $\Delta\alpha$. Fig.6-9 presents examples of 4 daily noise waveforms with different values of $\Delta\alpha$ and En: left-hand panels of graphics, Fig.6-9(a, b), present noise waveforms with high value of $\Delta\alpha$ and low values of En whereas right-hand panels, Fig.6-9(c, d), correspond to 2 noise waveforms with low values of $\Delta\alpha$ and high values of En. The difference in waveforms peculiarities between Fig.6-9(a, b) and Fig.6-9(c, d), is rather evident: high values of $\Delta\alpha$ and low values of En occur because of existence of irregular high-amplitude spikes, which are intermitted with intervals with stationary behavior. This is a typical multi-fractal: different types of stochastic behavior are observed. Low values of $\Delta\alpha$ correspond to much more stationary behavior: the noise structure is more simple and less multi-fractal.
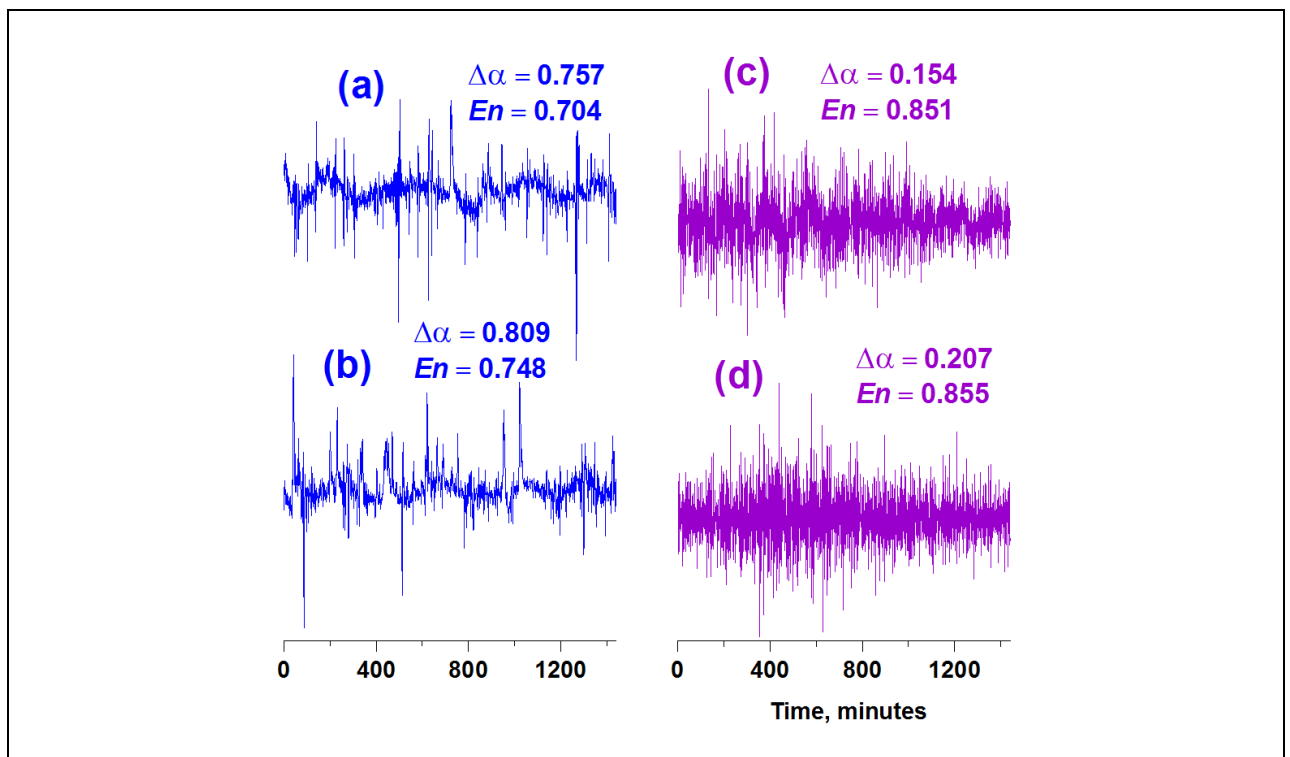

Figure 6-9. Two types of daily low-frequency seismic noise waveforms after removing tidal trends by polynomial of 8th order: (a, b) – with relatively large values of singularity spectrum support width $\Delta\alpha$ and high values of normalized entropy En and (c, d) – with relatively low values of $\Delta\alpha$ and En.

Our hypothesis consists in correlation between low values of $\Delta\alpha$ and growth of seismic danger. Thus, the increasing of entropy En could be connected with increasing of seismic danger as well. A possible physical interpretation of ability of low values of $\Delta\alpha$ and high values of En

extract seismically dangerous regions was given in [Lyubushin, 2012, 2013(a, b)]. It is the consequence of consolidation of small blocks of the Earth's crust into the large one before the strong earthquake. Consolidation implies that seismic noise does not include spikes, which are connected with mutual movements of small blocks. The absence of irregular spikes in the noise follows the decreasing of $\Delta\alpha$ and increasing of entropy En.

Having the values of $\Delta\alpha$ and En from all seismic stations, it is possible to create maps of spatial distribution of these seismic noise statistics. For this purpose let us consider the regular grid of the size 30×30 nodes covering the rectangular domain with latitudes between 30°N and 46°N and longitudes between 128°E and 148°E (see Fig.6-5). For each node of this grid the corresponding daily values of $\Delta\alpha$ and En are found, which are calculated as median for the values of five nearest to the node operable seismic stations. This simple procedure provides the sequence of daily maps of all parameters. The averaged maps are created by averaging daily maps for all days between 2 given dates. Taking into account that almost all stations of the F-net are placed at large Japanese islands these map in the ocean regions have the less significance than at islands of course. But we had to work with those data which we have at our disposal. The method of nearest neighbors, which is used in this chapter provides a rather natural extrapolation of the used values into domains, which have no points of observations.
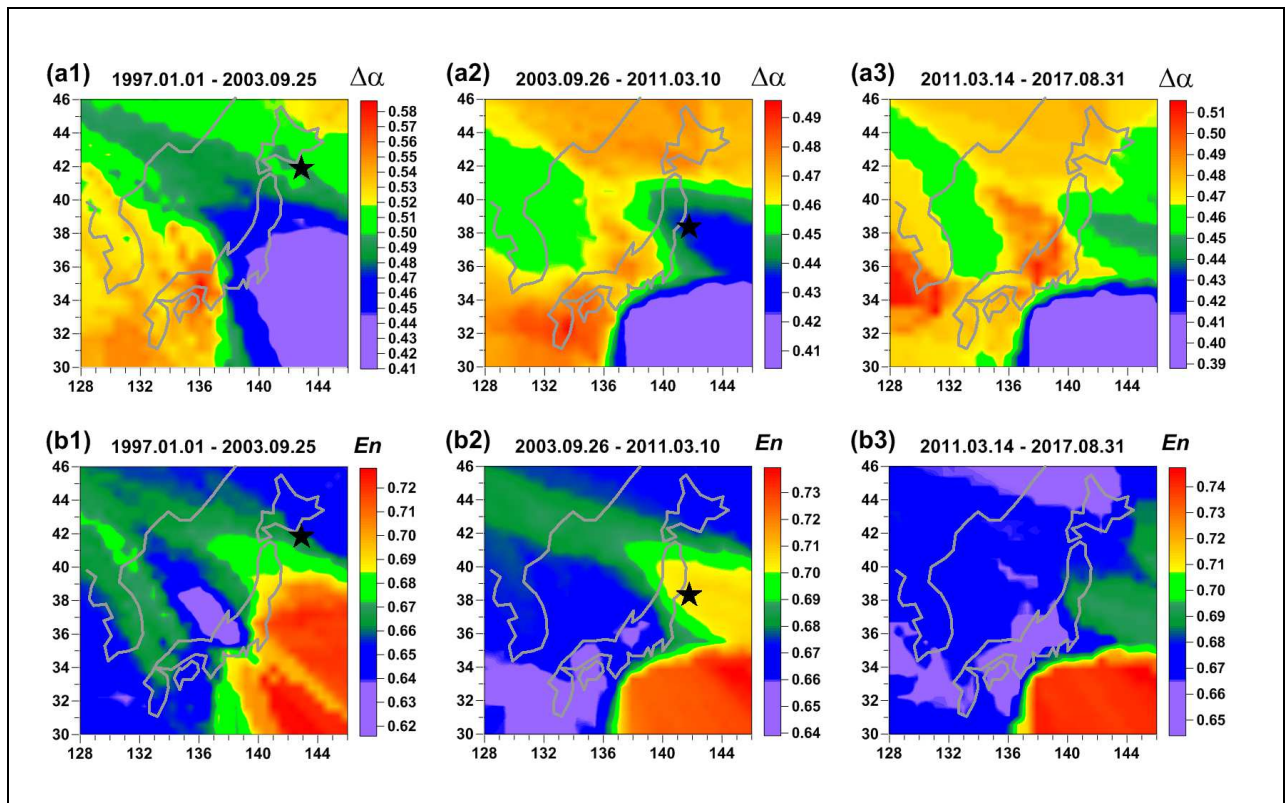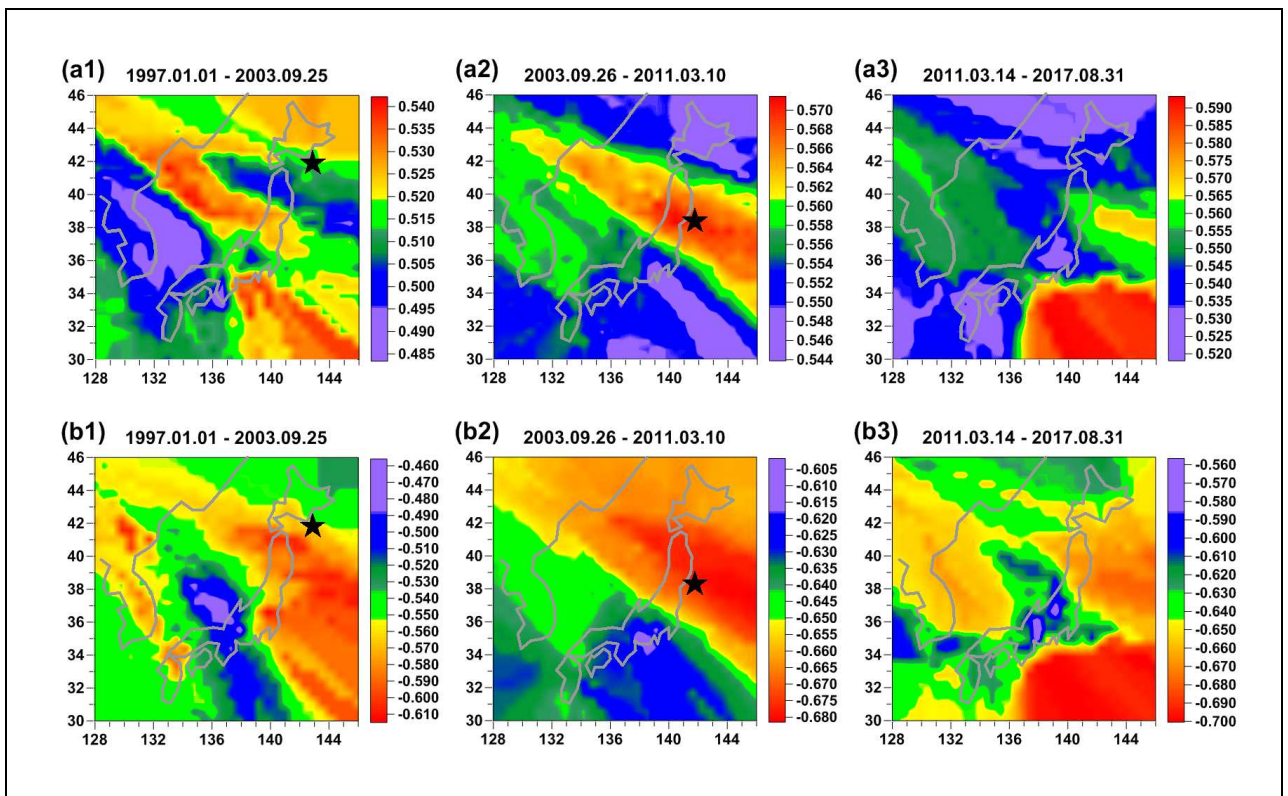

Figure 6-10. Averaged maps of multifractal singularity spectrums support width $\Delta\alpha$ ((a1)-(a3)) and minimum normalized entropy En of squared orthogonal wavelet coefficients ((b1)-(b3)) for 3 time intervals: 1997.01.01-2003.09.25 ((a1) and (b1)), 2003.09.26-2011.03.10 ((a2) and (b2)) and 2011.03.14-2017.08.31 ((a3) and (b3)). Stars indicate epicenters of earthquake on September 25, 2003, $M = 8.3$ ((a1) and (b1)) and on March 11, 2011, $M = 9.1$ ((a2) and (b2)).

Fig.6-10 presents averaged maps of $\Delta\alpha$ and En for 3 adjacent time fragments: from the beginning of 1997 up to 25 of September 2003, the day of earthquake with magnitude 8.3 near Hokkaido; from 26 of September 2003 up to 10 of March 2011, the day before Tohoku mega-earthquake 11 of March with magnitude 9.1 and from 14 of March 2011 up to 31 of August 2017. Three days, 11, 12 and 13 of March 2011 are excluded from the analysis because during

these days a lot of seismic stations of F-net were not working properly after seismic shock of 11 of March 2011.

The Fig.6-10(a2) shows that plotting averaged maps of spatial distribution of singularity spectrum support width $\Delta\alpha$ could extract the place of future catastrophe as the regions with relatively low values of $\Delta\alpha$. Fig.6-10(a1) presents a map where the area of relatively low $\Delta\alpha$ could be noticed which includes the place of future mega-earthquake and it is not split into North and South parts. At the Fig.6-10(a2) we can see that after the event on September 25, 2003 this area was split into North and South parts and the North part turned to be the area of aftershocks of Great Japan earthquake on March 11, 2011, whereas the South part remains to be the region of relatively low values of singularity spectrum support width $\Delta\alpha$ before and after 2011.03.11 (Fig.6-10(a3)). According to interpretation of regions with low values of $\Delta\alpha$ as the seismically dangerous, we could propose a hypothesis that during Tohoku earthquake only a part of accumulated seismic energy was dropped and that the above mentioned South region (the north part of Philippine plate, Nankai Through) could be the area of future mega-earthquake. Such hypothesis explains why the coherence of seismic noise remains high after 2011.03.11 (Fig.6-7(b) and Fig.6-8).

Figure 6-11. (a1)-(a3) – averaged maps of correlation coefficient between increments of generalized Hurst exponent α^* and multifractal singularity spectrums support width $\Delta\alpha$ for 3 time intervals: 1997.01.01-2003.09.25 (a1), 2003.09.26-2011.03.10 (a2) and 2011.03.14-2017.08.31 (a3). Stars indicates epicenters of earthquake on September 25, 2003, $M=8.3$ (a1) and on March 11, 2011, $M=9.1$ (a2). Plots ((b1)-(b3)) present the similar maps for correlation coefficient between α^* and minimum entropy of wavelet coefficients, En.

The Fig.6-10(b1), Fig.6-10(b2) and Fig.6-10(b3) confirm the conclusion, which was made from the analysis of graphs at Fig.6-9 that minimum normalized entropy En is an "antipode" to the parameter $\Delta\alpha$ and almost everything, what was written above about properties of $\Delta\alpha$ could be repeated for En with changing "minimum" to "maximum": relatively maximum values of normalized entropy extract seismically danger domain before the earthquake.

Let us call the regions extracted by low values of $\Delta\alpha$ and high values of En as "spots of seismic danger" – SSD. Mean values of $\Delta\alpha$ and En are strongly anti-correlated – that is why statistics $\Delta\alpha$ and En extract approximately the same SSD. Nevertheless, their mutual considering is expedient because these parameters are based on different approaches. The maps of $\Delta\alpha$ and En could be plotted as the sequence in the moving time window – such estimates provide the possibility to visualize the origin and evolution of SSD.

The problem of predicting strongest earthquakes in Japan at the region of Nankai Trough is a traditional large problem for seismologists in Japan [Rikitake, 1999; Mogi, 2004]. In [Rikitake, 1999] the probability of earthquake with magnitude more than 8.5 at Tokai-Nankai zone, the region where Philippine Sea plate is approaching Central Japan, was estimated as 0.35-0.45 "for a ten-year period following the year 2000". In [Simons et al., 2011] the seismic danger for Japan was estimated immediately after Tohoku earthquake based on the analysis of GPS data and the conclusion was that "estimates … suggest the need to consider the potential for a future large earthquake just south of this event." In the paper [Kagan, Jackson, 2013] the problem why the Tohoku earthquake was a surprise for scientific community is discussed. One of the conclusions in [Kagan, Jackson, 2013] is "A magnitude 9 earthquake off Tohoku should not have been a surprise". This conclusion was made by retrospective analysis of seismic catalogs. Nevertheless, the Tohoku event was a great surprise for all traditional methods of earthquake prediction. Other conclusion in [Zoller et al, 2014] is that even magnitude 10 is quite possible for Japan Trench.

Maps of correlations between pairs of parameters $(\alpha^*, \Delta\alpha)$ and (α^*, En) possess interesting prognostic properties. Similar to the maps presented at the Fig.6-10 we can plot averaged maps of correlation between increments of $(\alpha^*, \Delta\alpha)$. For this purpose, let's estimate evolution of correlation coefficient between $(\alpha^*, \Delta\alpha)$ for each station within moving time window of the length 365 days. For each position of 1-year moving time window, we can plot a map by calculating median of correlation coefficients for 5 operable seismic stations which are nearest to each node of regular grid. The averaged maps are created by averaging maps corresponding to all 1-year time fragments, which lay entirely between and including 2 given dates. In a similar way maps of correlation coefficient could be plotted for pair (α^*, En). Such maps are presented at Fig.6-11. It is interesting to notice that at Fig.6-11(a2,b2) the region of future Tohoku earthquake is extracted by relatively high absolute values of correlations. Note that correlations at Fig.6-11(b1)-(b3) are negative. For the period after Tohoku earthquake the region of SSD according to Fig.6-10(a3,b3) coincides with region of maximum absolute values of correlations – Fig.6-11(a3,b3).

Let us add one more independent multifractal parameter and consider 3 median values of $(\Delta\alpha, \alpha^*, \alpha_{\min})$ where $\alpha_{\min}$ is minimum Hölder-Lipschitz exponent. Fig.6-12(a1-a3) present plots of these values. Considering clustering properties of the clouds of the daily sequence of these 3D vectors within moving time window gives a possibility to estimate natural fluctuations of seismic danger and even discover its periodical structure.

Let's consider moving time window of the length L=365 days and let $\vec{\xi}^{(t)} = (\Delta\alpha, \alpha^*, \alpha_{\min})^{(t)}$ be 3D vector within current time window, $t = 1,..., L$, t is time index, numerating vectors. Our purpose is investigating clustering properties of clouds of 3D vectors $\vec{\xi}^{(t)}$ with each 1-year time window. In particular, we are interested, what is the "best" number of clusters.
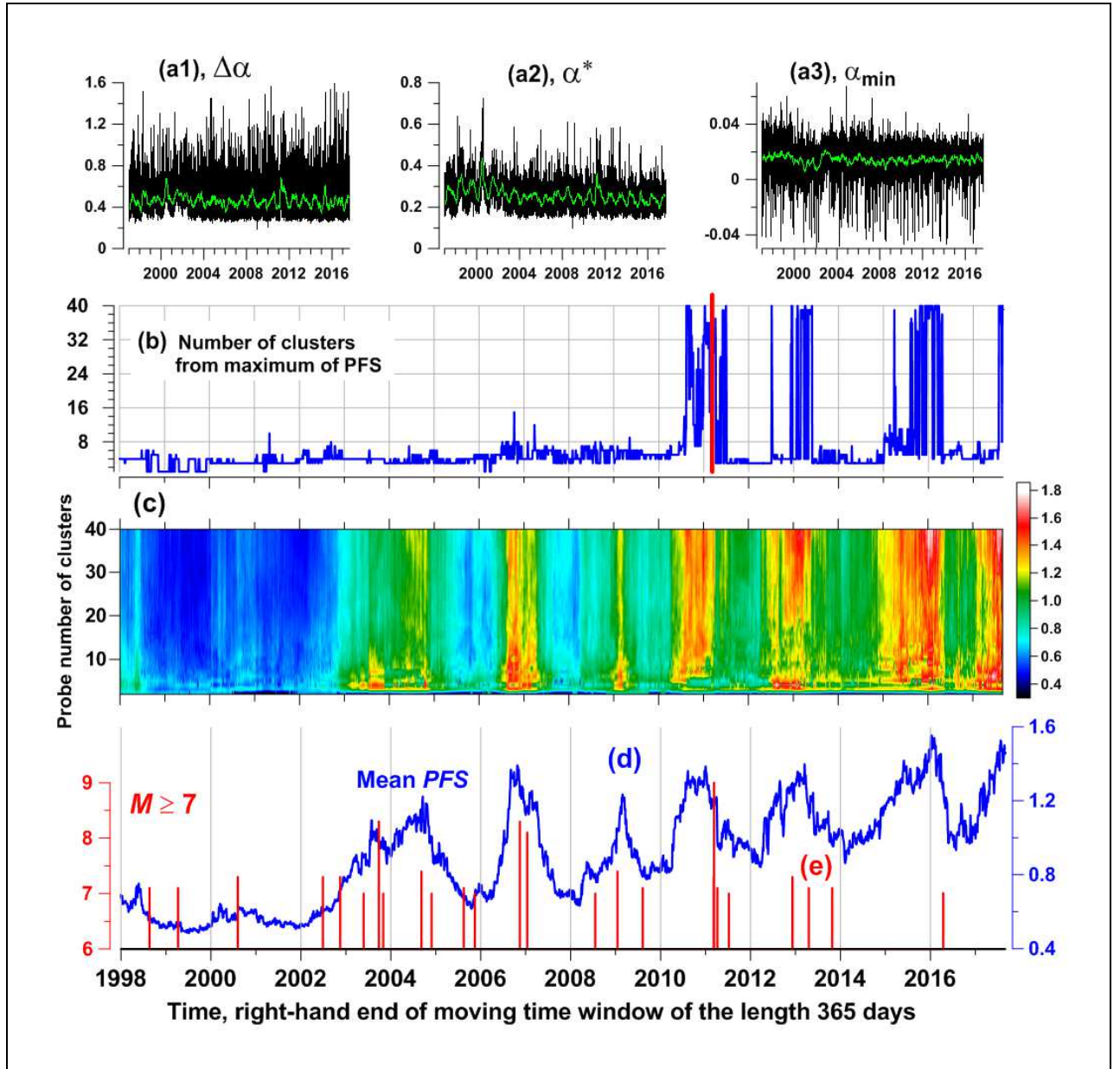
Figure 6-12. (a1)-(a3) – plots of daily median values of multifractal singularity spectrum support width $\Delta\alpha$, generalized Hurst exponent α^* and minimum Hölder-Lipschitz exponent α_{min} from all 78 stations of broadband seismic network F-net in Japan; green lines present running average within time windows of the length 57 days; (b) – plot of the best numbers of clusters for the sequence of clouds consisting of 365 daily 3D vectors $(\Delta\alpha, \alpha^*, \alpha_{min})$ from moving time window of the length 365 days with mutual shift 3 days. The best number of clusters is defined from the maximum of pseudo-F-statistics. Vertical red line indicates time moment of Tohoku mega-earthquake on March 11, 2011, $M = 9.1$. Two-dimensional diagram (c) presents dependence of pseudo-F-statistics on the probe number of clusters, which is varying from 2 up to 40 within each time window. Plot (d) presents mean value of pseudo-F-statistics averaged by all probe numbers of clusters in dependence on right-hand end of moving time window of the length 365 days. Plot (e) presents the sequence of time moments of strong earthquakes $M \geq 7$ in the rectangular domain with coordinates 28°N≤ Latitude≤ 48°N; 128°E ≤ Longitude≤ 156°E, which is a rather broad vicinity of Japan islands.

Before making cluster procedure a preliminary operation of normalizing and iterative clipping of outliers [Huber and Ronchetti, 2009] was performed within each time window for each scalar component $\vec{\xi}_k^{(t)}$ of vectors $\vec{\xi}^{(t)}$. Here $k = 1, 2, 3$ is the index numerating scalar components of the vector $\vec{\xi}^{(t)}$. Let $\bar{\xi}_k = \frac{1}{L}\sum_{t=1}^{L}\xi_k^{(t)}$, $\sigma_k^2 = \frac{1}{(L-1)}\sum_{t=1}^{L}(\xi_k^{(t)} - \bar{\xi})^2$ be sample estimates of mean values

and variance of scalar components of the 3D vector $\vec{\xi}^{(t)}$. Let's perform iterations which consist in coming to values $\zeta_k^{(t)} = (\xi_k^{(t)} - \overline{\xi}_k)/\sigma_k$ and clipping values $\zeta_k^{(t)}$ exceeding thresholds $\pm 3\sigma_k$. These iterations are stopped, when the values $\overline{\xi}_k$ and σ_k became stable and equal to the following values: $\overline{\overline{\xi}}_k = 0$, $\sigma_k = 1$. After this preliminary operation at each current time window we have a cloud consisting of L 3D vectors $\vec{\zeta}^{(t)}$.

Let's split some cloud into given probe number q of clusters using standard k-means cluster procedure [Duda et al, 2000]. Let $\Gamma_r, r = 1, ..., q$ be clusters, $\vec{z}_r = \sum_{\vec{\zeta} \in \Gamma_r} \vec{\zeta}/n_r$ – vector of the center of cluster Γ_r, n_r be a number of vectors $\vec{\zeta}^{(t)}$ within cluster Γ_r, $\sum_{r=1}^{q} n_r = L$. Vector $\vec{\zeta}^{(t)} \in \Gamma_r$ if the distance $|\vec{\zeta}^{(t)} - \vec{z}_r|$ is minimum among all positions of clusters' centers. K-means procedure minimizes sum

$$S(\vec{z}_1, ..., \vec{z}_q) = \sum_{r=1}^{q} \sum_{\vec{\zeta} \in \Gamma_r} |\vec{\zeta} - \vec{z}_r|^2 \rightarrow \min_{\vec{z}_1, ..., \vec{z}_q} \tag{6.32}$$

with respect to positions of clusters' centers $\vec{z}_r$. Let $J(q) = \min_{\vec{z}_1, ..., \vec{z}_q} S(\vec{z}_1, ..., \vec{z}_q)$. We try probe number of clusters within range $2 \leq q \leq 40$. The problem of selecting the best number of clusters q^* was solved from maximum of pseudo-F-statistics [Vogel, Wong, 1979], which is similar to F-criterion (6.30) from analysis of variance:

$$PFS(q) = \sigma_1^2(q) / \sigma_0^2(q) \rightarrow \max_{2 \leq q \leq 40} \tag{6.33}$$

where

$$\sigma_0^2(q) = J(q)/(L-q), \quad \sigma_1^2(q) = \sum_{r=1}^{q} \nu_r \cdot |\vec{z}_r - \vec{z}_0|^2$$

$$\nu_r = n_r / L, \quad \vec{z}_0 = \sum_{t=1}^{L} \vec{\zeta}^{(t)} / L \tag{6.34}$$

The $PFS \rightarrow \max$ rule is not working if we try to distinguish cases $q^* = 1$ and $q^* = 2$ because the value $\sigma_1^2(q)$ is not defined for $q = 1$. These cases could be distinguished by existing of break point of the monotonous function $J(q)$ at the argument $q = 2$ [Lyubushin, 2011(a)]. The value $\sigma_0^2(q)$ monotonically increases as q decreases, and usually the dependence of $\log(\sigma_0^2(q))$ on $\log(q)$ is close to linear, that is, scales as $q^{-\mu}$. As is known, the optimal number of clusters can also be determined from the break point of the monotonic dependence $\sigma_0^2(q)$ for $q = q^*$: as q decreases, the function $\sigma_0^2(q)$ increases faster at $q < q^*$ than at $q > q^*$. This rule for determining number of clusters is known as "elbow method" [Ketchen, Jr; Shook, 1996]. This criterion of identification of $q = q^*$ is more susceptible to noise and exhibits a poorer performance compared to the technique $q^* = \arg \max PFS(q)$ but this is the only possibility to discern the case $q^* = 1$ from the case $q^* = 2$. Let $\delta(q)$ denote the deviation of $\log(\sigma_0^2(q))$ from the best fit straight line approximating the dependence on $\log(q)$, i.e. be defined from formula $\delta(q) = \log(\sigma_0^2(q)) - (a \log(q) + b)$, where coefficients (a,b) are found by least squares: $\sum_{q=1}^{40} \delta^2(q) \rightarrow \min_{a,b}$. Then, we assume that the point $q = 2$ is the break point of the dependency

$\sigma_0^2(q)$ if $\delta(1)$ exceeds all values of $\delta(q)$ for $q \geq 2$. Let $q_0 = \arg\max\limits_{2 \leq q \leq 40} PFS(q)$. Thus, we define the optimal number q^* of clusters according to the following rule:

If $q_0 > 2$ then $q^* = q_0$. Else, if $\delta(1)/\max\limits_{2 \leq q \leq 40} \delta(q) \leq 1$ then $q^* = 1$, else $q^* = 2$.

Graphic at Fig.6-12(b) presents evolution of the estimates of the best number of clusters q^* in dependence on the right-hand end of moving time window of the length 1 year.

The values (6.33) computed within moving time window for all probe number q of clusters are dependent on position of time window. Thus, pseudo-F-statistics could be presented as 2D map as dependence on q and the right-hand end of moving time window. This 2D map is shown at the Fig.6-12(c). From Fig.6-12(c) it is following that the q^* before Tohoku mega-earthquake on March 11, 2011, has strongly chaotic regime with jumps from minimum up to maximum values in the time interval 1 year before the event and this time interval was characterized by high $PFS(q)$ values. Let's consider the sequence of mean values $\bar{P} = \sum_{q=2}^{40} PFS(q)/39$ of pseudo-F-statistics in each time window in dependence on time moments of right-hand end of window. This dependence is presented at the plot Fig.6-12(d), which demonstrates positive trend and strong periodicity with almost 2 years period, which was established after Hokkaido earthquake on 2003.09.25.

Our hypothesis consists in assumption that maps of pseudo-F-statistics, which are built for multi-fractal properties of seismic noise, similar to those presented at Fig.6-12(c) and graphs of its mean values, similar to Fig.6-12(d), could be useful for visualization of natural fluctuations of seismic danger in some rather big region. The basis for this hypothesis could be comparison of Fig.6-12(d) with sequence of strong earthquakes $M \geq 7$ within rectangular domain with coordinates 28°N $\leq$ Latitude $\leq 48^{\circ}$N; 128°E $\leq$ Longitude $\leq 156^{\circ}$E, which is presented at the Fig.6-12(e). In particular starting from 2006 all strong events belong to time intervals with large values of mean $PFS(q)$ and the last time interval of large values of q^* and $PFS(q)$ at Fig.6-12(b,c) precedes the Kumamoto earthquake $M = 7$ on 2016.04.15 with hypocenter Latitude = 32.78°N and Longitude = 130.72°E at South Japan. Mean values of pseudo-F-statistics in Japan beginning from 2004 have fluctuations with approximate period 2 years.

Results for surrogate time series

Let us check the stability of conclusions about multiple correlation which is presented at the Fig.6-7 by calculating the same measure for surrogate time series obtained by simple shuffling of samples. For this purpose in order to preserve visible low-frequency structure of the graphs let us compute trends for each principal component by applying Gaussian kernel smoothing (6.29) with averaging radius h 90 days. These Gaussian trends were subtracted from principal components and residuals were shuffled. After this the surrogate time series were constructed as the sum of trends and shuffled residuals. The result of computing multiple correlation similar to the case presented at the Fig.6-7 is presented at the next Fig.6-13.
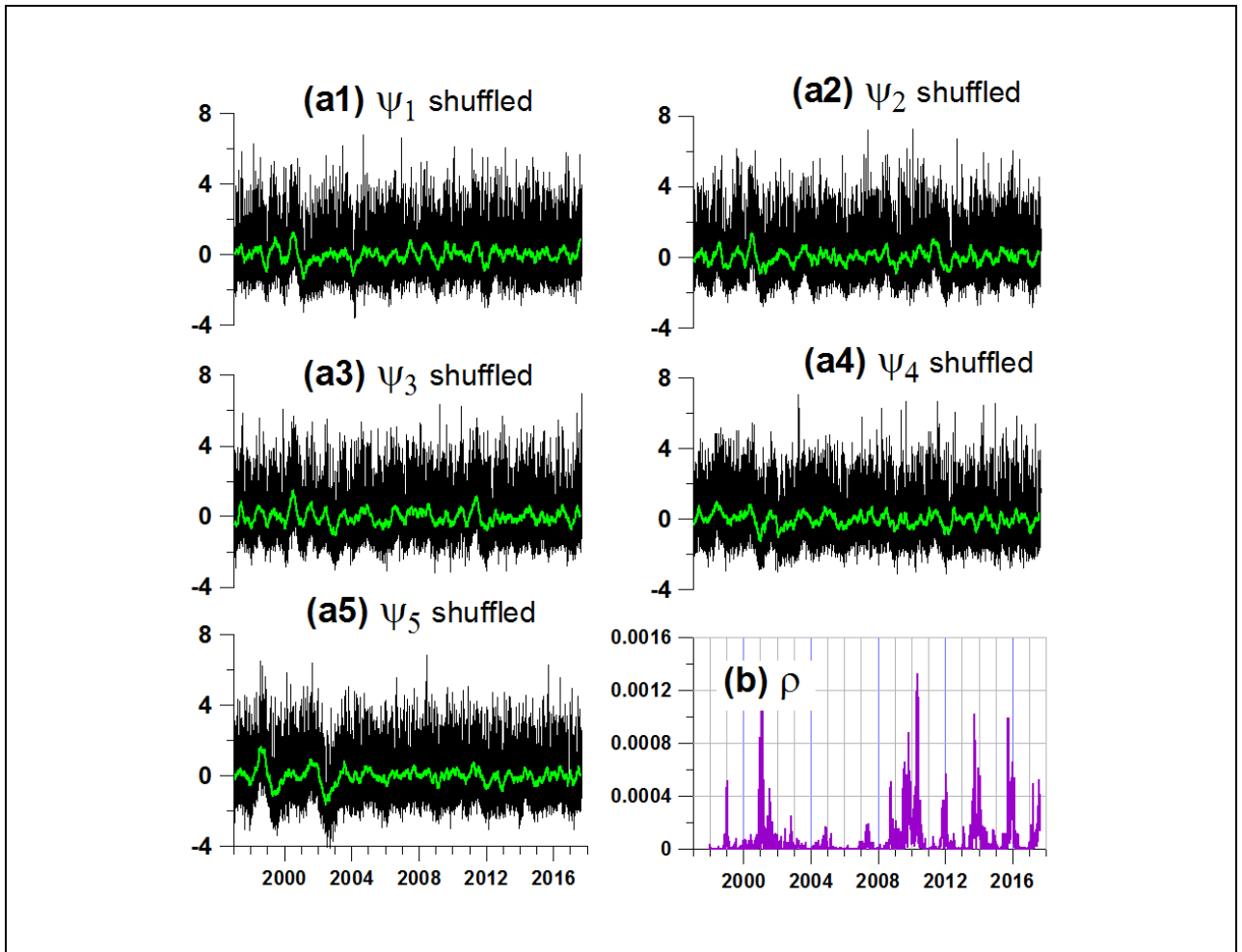
Figure 6-13. (a1)–(a5) black lines - plots of values of shuffled first principal components ψ_k, $k = 1,..,5$ (low-frequency trends were preserved), calculated for daily median values of multifractal singularity spectrum support width $\Delta\alpha$, generalized Hurst exponent α^* and minimum normalized entropy of squared orthogonal wavelet coefficients En for five clusters of stations in Japan (Fig.6–5) for seismic noise waveforms after coming to sampling time step 1 min from initial waveforms with sampling rate 1 Hz (see Fig.6–2), respectively, within a length of adaptation of 365 days; green lines represent running average within the time window of 57 days for each curve. Plot (b) presents values of squared multiple robust correlation coefficient ρ estimated within the moving time window of 365 days for shuffled time series depending on the right-hand end of the moving time window.

Comparing Figs 6–7 and 6–13 shows that shuffling of HF components of time series (HF components with periods less than 180 days) destroys high correlation which is presented at the Fig.6–7.

Conclusion

In studies of such a complex multi-component system as the Earth's crust, it is highly challenging to identify a set of the main deterministic reasons, which would define all the features of the global seismic regime, particularly those, which control long-term changes in the intensity of potential seismic events. Solving this problem may be facilitated by the phenomenological approach based on the use of coherent noise generated by the system in the course of its evolution. For the Earth's crust, the ambient noise is a product of its 'life'. Coherence (or synchronization) of the behavior of characteristics of a complex system, described by data of different nature and structure, is an important feature for assessments of its approach

to rapid changes in the condition, which are often referred to as a 'catastrophe'. Searching for precursors of catastrophes, which may be manifested by the occurrence of synchronous components in a variety of observations, is the general idea for increasing the correlation radius of random fluctuations of parameters of a complex system as it approaches a sharp change in its properties, resulting from its own dynamics [Gilmore, 1981; Nicolis, Prigogine, 1989]. This property of coherence of ambient noise of the Earth is investigated in this chapter.

Analysis of coherence between properties of global seismic noise, measured at the network of 229 broadband stations all over the world since the beginning of 1997 till current time, extracts effect of progressively increasing synchronization after Sumatra mega-earthquake on 26 Dec 2004, which could be a precursor of the further rise in the intensity of the strongest seismic events [Lyubushin, 2014(a), 2015]. A particular case of this effect is a strong increasing of coherence between behaviour of different parameters of low-frequency seismic noise in Japan and California before Tohoku mega-earthquake on March 11, 2011 in Japan, detected by analysis of seismic noise waveforms from regional broadband seismic networks [Lyubushin, 2016(a,b)].

Plotting the maps of different properties of low-frequency seismic noise (multifractal singularity spectrum support width and minimum normalized entropy of squared orthogonal wavelet coefficients) within moving time window could present a new method of dynamic seismic hazard estimate. It gives a possibility to inspect the origin and evolution of the "spots of seismic danger". Analysis of seismic noise at Japan islands from broad-band seismic network F-net gave a possibility for prediction of Great Japan earthquake of 11 of March 2011. The prediction was published in a number of scientific papers and abstracts at international conferences in advance of the seismic catastrophe [Lyubushin, 2008(a), 2009, 2010(a,b), 2011(a,b,c)]. According to the analysis of seismic noise after 11 of March of 2011 the next mega-earthquake with magnitude 8.5-9.0 could occur at the region of Nankai Trough [Lyubushin, 2012, 2013(a,b), 2014(b)]. For estimating the time interval of occurrence of this seismic event the periodic structure of seismic danger natural fluctuations with period near 2 years (Fig.6-12(d)) could be used.

References

Ardhuin F., Stutzmann E., Schimmel M. and Mangeney A., Ocean wave sources of seismic noise, *J. Geophys. Res.* **116**, 2011, C09004.

Aster R., McNamara D. and Bromirski P., Multidecadal climate induced variability in microseisms, *Seismol. Res. Lett.* **79**, 2008, 194–202.

Berger J., Davis P. and Ekstrom G., Ambient earth noise: a survey of the global seismographic network, *J. Geophys. Res*. 2004 **109**, 2004, B11307.

Brillinger D.R., Time Series. Data Analysis and Theory, 1975, Holt, Rinehart and Winston, Inc, New York, Chicago, San Francisco, 540.

Chandrasekhar E., Sanjana S.P., Gopi K.S. and Nayana S., Multifractal detrended fluctuation analysis of ionospheric total electron content data during solar minimum and maximum, J. Atmos. Sol.—Terr. Phys. **149** (2016), 2016, 31–39. https://doi.org/10.1016/j.jastp.2016.09.007

Clarke E., Generalized gradients and applications, *Trans. Am. Math. Soc*. 1975 **205** (2), 1975, 247–262.

Costa M., Peng C.-K., Goldberger A.L. and Hausdorf J.M., Multiscale entropy analysis of human gait dynamics, *Physica A* **330** (2003), 2003, 53–60.

Costa M., Goldberger A.L. and Peng C.-K., Multiscale entropy analysis of biological signals, *Phys. Rev. E* **71** (2005), 2005, 021906.

Currenti G., del Negro C., Lapenna V. and Telesca L., Multifractality in local geomagnetic field at Etna volcano, Sicily (southern Italy), *Nat. Hazard. Earth Syst. Sci.* **5** (555–559), 2005.

Duda R.O., Hart P.E. and Stork D.G., Pattern Classification, 2000, Wiley-Interscience Publication, New York, Chichester, Brisbane, Singapore, Toronto. 680

Dutta S., Ghosh D. and Chatterjee S., Multifractal detrended fluctuation analysis of human gait diseases, *Front. Physiol.,* 2013 **v.4**, 2013. https://doi.org/10.3389/fphys.2013.00274

Feder J., Fractals, 1988, Plenum Press, New York, London, 284.

Friedrich A., Krüger F. and Klinge K., Ocean-generated microseismic noise located with the Gräfenberg array, *J. Seismol.* **2** (1), 1998, 47–64.

Fukao Y.K., Nishida K. and Kobayashi N., Seafloor topography, ocean infragravity waves, and background Love and Rayleigh waves, *J. Geophys. Res*. **115**, 2010, B04302.

Gilmore R., Catastrophe Theory for Scientists and Engineers, 1981, John Wiley and Sons, Inc, New York, NY, 666.

Grevemeyer I., Herber R. and Essen H.-H., Microseismological evidence for a changing wave climate in the northeast Atlantic Ocean, *Nature* **408**, 2000, 349–352.

Hannan E.J., Multiple Time Series, 1970, John Wiley and Sons, Inc, New York, London, Sydney, Toronto.

Hardle W., Applied Nonparametric Regression. (Biometric Society Monographs No. 19.), 1990, Cambridge University Press, Cambridge, 333.

Hotelling H., Relations between two sets of variates, Biometrika. **28**, 1936, 321–377.

Huber P.J. and Ronchetti E.M., Robust Statistics, second ed, 2009, John Wiley & Sons, Inc, 354. https://doi.org/10.1002/9780470434697.ch1

Humeaua A., Chapeau-Blondeau F., Rousseau D., Rousseau P., Trzepizur W. and Abraham P., Multifractality, sample entropy, and wavelet analyses for age-related changes in the peripheral cardiovascular system: preliminary results, *Med. Phys., Am. Assoc. Phys. Med*. **35** (.2), 2008, 717–727, February 2008.

Ida Y., Hayakawa M., Adalev A. and Gotoh K., Multifractal analysis for the ULF geomagnetic data during the 1993 Guam earthquake, *Nonlinear Processes Geophys*. **vol.12**, 2005, 157–162.

Ivanov P. Ch, Amaral L.A.N., Goldberger A.L., Havlin S., Rosenblum M.B., Struzik Z., et al., Multifractality in healthy heartbeat dynamics, *Nature* **399**, 1999, 461–465.

Jolliffe I.T., Principal Component Analysis, 1986, Springer-Verlag, 487. https://doi.org/10.1007/b98835

Kagan Y.Y. and Jackson D.D., Tohoku earthquake: a surprise?, *Bull. Seismol. Soc. Am.* **103** (.2B), 2013, 1181–1194. https://doi.org/10.1785/0120120110

Kantelhardt J.W., Zschiegner S.A., Konscienly-Bunde E., Havlin S., Bunde A. and Stanley H.E., Multifractal detrended fluctuation analysis of nonstationary time series, *Physica A* **316** (1–4), 2002, 87–114. https://doi.org/10.1016/S0378-4371(02)01383-3

Kedar S., Longuet-Higgins M., Webb F., Graham N., Clayton R. and Jones C., The origin of deep ocean microseisms in the North Atlantic Ocean, *Proc. R. Soc.* A **464**, 2008, 777–793.

Ketchen D.J., Jr and Shook C.L., The application of cluster analysis in strategic management research: an analysis and critique, *Strategic Manage.* J. **17** (6), 1996, 441–458.

Kobayashi N. and Nishida K., Continuous excitation of planetary free oscillations by atmospheric disturbances, *Nature.* **395**, 1998, 357–360.

Koper K.D. and de Foy B., Seasonal anisotropy in short-period seismic noise recorded in South Asia, *Bull. Seismol. Soc. Am.* **98**, 2008, 3033–3045.

Koper K.D., Seats K. and Benz H., On the composition of Earth's short-period seismic noise field, *Bull. Seismol. Soc. Am.* April 2010 **100** (2), 2010, 606–617.

Lyubushin A., Prognostic properties of low-frequency seismic noise, *Nat. Sci.* **4** (8A), 2012, 659–666. https://doi.org/10.4236/ns.2012.428087

Lyubushin A., How soon would the next mega-earthquake occur in Japan?, *Nat. Sci.* **5** (8A1), 2013b, 1–7. https://doi.org/10.4236/ns.2013.58A1001

Lyubushin A.A., Analysis of canonical coherences in the problems of geophysical monitoring, *Izv. Phys. Solid Earth* **34** (1), 1998, 52–58.

Lyubushin A.A., Analysis of multidimensional geophysical monitoring time series for earthquake prediction, *Ann. Geofis.* **42** (5), 1999, 927–937. https://doi.org/10.4401/ag-3757

Lyubushin A.A., Wavelet-aggregated signal and synchronous peaked fluctuations in problems of geophysical monitoring and earthquake prediction, *Izv. Phys. Solid Earth* **36** (2000), 2000, 204–213.

Lyubushin, A.A., 2008a. Multifractal properties of low-frequency microseismic noise in Japan, 1997–2008. In: Book of Abstracts of Seventh General Assembly of the Asian Seismological Commission and Japan Seismological Society, 2008 Fall Meeting, Tsukuba, Japan, 24–27 November 2008, 92.

Lyubushin A.A., Microseismic noise in the low frequency range (periods of 1–300 min): properties and possible prognostic features, *Izv. Phys. Solid Earth* **44** (4), 2008b, 275–290. https://doi.org/10.1134/s11486-008-4002-6.

Lyubushin A.A., Synchronization trends and rhythms of multifractal parameters of the field of low-frequency microseisms., *Izv. Phys. Solid Earth* **45** (5), 2009, 381–394. https://doi.org/10.1134/S1069351309050024

Lyubushin A.A., The statistics of the time segments of low-frequency microseisms: trends and synchronization, *Izv. Phys. Solid Earth* **46** (6), 2010a, 544–554. https://doi.org/10.1134/S1069351310060091.

Lyubushin A.A., 2010b. Synchronization of multifractal parameters of regional and global low-frequency microseisms. In: European Geosciences Union General Assembly 2010, Vienna, 02–07 May 2010. Geophys. Res. Abstr. V. 12. EGU2010-696.

Lyubushin A.A., Cluster analysis of low-frequency microseismic noise, *Izv. Phys. Solid Earth* **47** (6), 2011a, 488–495. https://doi.org/10.1134/S1069351311040057

Lyubushin A.A., Seismic catastrophe in Japan on March 11, 2011: long-term prediction on the basis of low-frequency microseisms, *Izv. Atmos. Oceanic Phys*. **46** (8), 2011b, 904–921. https://doi.org/10.1134/S0001433811080056

Lyubushin A.A., Mapping the properties of low-frequency microseisms for seismic hazard assessment, *Izv. Phys. Solid Earth* **49** (1), 2013a, 9–18. https://doi.org/10.1134%2FS1069351313010084

Lyubushin A.A., Analysis of coherence in global seismic noise for 1997–2012, *Izv. Phys. Solid Earth* **50** (3), 2014a, 325–333. https://doi.org/10.1134/S1069351314030069

Lyubushin A.A., Dynamic estimate of seismic danger based on multifractal properties of low-frequency seismic noise, *Nat. Hazard.* **70** (1), 2014b, 471–483. https://doi.org/10.1007/s11069-013-0823-7

Lyubushin A.A., Wavelet-based coherence measures of global seismic noise properties, *J. Seismol*. **19** (2), 2015, 329–340. https://doi.org/10.1007/s10950-014-9468-6.

Lyubushin A.A., Long-range coherence between seismic noise properties in Japan and California before and after Tohoku mega-earthquake, *Acta Geodaetica Geophys* 2016a. https://doi.org/10.1007/s40328-016-0181-5.

Lyubushin A.A., Coherence between the fields of low-frequency seismic noise in Japan and California, *Izv. Phys. Solid Earth* **52** (6), 2016b, 810–820. https://doi.org/10.1134/S1069351316050086

Lyubushin A.A. and Klyashtorin L.B., Short term global *dT* prediction using (60–70)-years periodicity, *Energy Environ*. **23** (.1), 2012, 75–85. https://doi.org/10.1260/0958-305X.23.1.75

Lyubushin A.A. and Kopylova G.N., Multidimensional wavelet analysis of time series of electrotelluric observations in Kamchatka, *Izv. Phys. Solid Earth* **40** (.2), 2004, 163–175, 2004.

Lyubushin A.A. and Sobolev G.A., Multifractal measures of synchronization of microseismic oscillations in a minute range of periods, *Izv. Phys. Solid Earth* **42** (.9), 2006, 734–744.

Lyubushin A.A., Pisarenko V.F., Bolgov M.V. and Rukavishnikova T.A., Study of general effects of rivers runoff variation, *Russ. Meteorol. Hydrol*. (7), 2003, 59–68.

Lyubushin A.A., Pisarenko V.F., Bolgov M.V., Rodkin M.V. and Rukavishnikova T.A., Synchronous variations in the Caspian sea level from coastal observations in 1977–1991, *Atmos. Oceanic Phys*. **40** (6), 2004, 737–746, 2004.

Lyubushin A.A., Multifractal parameters of low-frequency microseisms, In: de Rubeis V., et al., (Eds.), *Synchronization and Triggering: from Fracture to Earthquake Processes, GeoPlanet*, 2010b, Earth and Planetary Sciences, Springer, Verlag Berlin Heidelberg, 253–272, 2010, 388 p., Chapter 15. https://doi.org/10.1007/978-3-642-12300-9_15

Lyubushin A.A., Kaláb Z. and Lednická M., Geomechanical time series and its singularity spectrum analysis, *Acta Geodaetica Geophys Hungarica* **47** (1), 2012, 69–77, March. https://doi.org/10.1556/AGeod.47.2012.1.6

Lyubushin A.A., Kaláb Z. and Lednická M., Statistical properties of seismic noise measured in underground spaces during seismic swarm, *Acta Geodaetica Geophys* **49** (2), 2014, 209–224. https://doi.org/10.1007/s40328-014-0051-y

Lyubushin A.A., Bobrovskiy V.S. and Shopin S.A., Experience of complexation of global geophysical observations, *Geodyn. Tectonophys*. **7** (1), 2016, 1–21. https://doi.org/10.5800/GT-2016-7-1-0194

Mallat S., A Wavelet Tour of Signal Processing, 1998, Academic Press, San Diego, London, Boston, NY, Sydney, Tokyo, Toronto, 577 p.

Marple S.L., (Jr.), Digital Spectral Analysis with Applications, 1987, Prentice-Hall, Inc., Englewood Cliffs, NJ.

McNamara D.E. and Buland R.P., Ambient noise levels in the continental United States, *Bull. Seismol. Soc. Am*. **2004** (94), 2004, 1517–1527.

Mogi K., Two grave issues concerning the expected Tokai Earthquake, *Earth Planets Space* **56**, 2004, li-lxvi.

Nicolis G. and Prigogine I., Exploring Complexity, An Introduction, 1989, W.H. Freedman and Co., New York, NY, 328.

Nishida K., Kawakatsu H., Fukao Y. and Obara K., Background Love and Rayleigh waves simultaneously generated at the Pacific Ocean floors, *Geophys. Res. Lett.* **35**, 2008, L16307.

Nishida K., Montagner J. and Kawakatsu H., Global surface wave tomography using seismic hum, Science **326** (5949), 2009, 112.

Pavlov A.N. and Anishchenko V.S., Multifractal analysis of complex signals. *Physics – Uspekhi Fizicheskikh Nauk, Russ. Acad. Sci.* **50** (.8), 2007, 819–834. https://doi.org/10.1070/PU2007v050n08ABEH006116

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vettering, W.T., 1996. Numerical Recipes. second ed., Chapter 13, Wavelet Transforms, Cambridge Univ. Press, Cambridge.

Ramirez-Rojas A., Munoz-Diosdado A., Pavia-Miller C.G. and Angulo-Brown F., Spectral and multifractal study of electroseismic time series associated to the Mw=6.5 earthquake of 24 October 1993 in Mexico, *Nat. Hazard. Earth Syst. Sci.* **4**, 2004, 703–709.

Rao C.R., Linear Statistical Inference and Its Applications, 1965, John Wiley and Sons, New York, London, Sydney, 1965.

Rhie J. and Romanowicz B., Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling, *Nature* **2004** (431), 2004, 552–554.

Rhie J. and Romanowicz B., A study of the relation between ocean storms and the Earth's hum, *Geochem. Geophys. Geosyst.* **7** (10), 2006.

Rice J., The mechanics of earthquake rupture, In: Dziewonski A.M. and Boschi E., (Eds.), *Physics of the Earth's Interior. Proceedings of the International School of Physics "Enrico Fermi", Course 78, 1979*, 1980, Italian Physical Society, Amsterdam, North-Holland, 555–649, 1980.

Rikitake, T., 1999. Probability of a great earthquake to recur in the Tokai district, Japan: reevaluation based on newly-developed paleoseismology, plate tectonics, tsunami study, micro-seismicity and geodetic measurements. *Earth Planets Space* 51, 147–157.

Schimmel M., Stutzmann E., Ardhuin F. and Gallart J., Polarized Earth's ambient microseismic noise, *Geochem. Geophys. Geosyst*. **12**, 2011, Q07014.

Simons M., Minson S.E., Sladen A., Ortega F., Jiang J., Owen S.E., et al., The 2011 Magnitude 9.0 Tohoku-Oki earthquake: mosaicking the megathrust from seconds to centuries, *Science* **332** (6032), 2011, 911. https://doi.org/10.1126/science.332.6032.911

Stehly L., Campillo M. and Shapiro N.M., A study of the seismic noise from its long-range correlation properties, *J. Geophys. Res*. **111**, 2006, B10306.

Tanimoto T., Continuous free oscillations: atmosphere-solid earth coupling, *Annu. Rev. Earth Planet. Sci*. **29**, 2001, 563–584.

Tanimoto T., The oceanic excitation hypothesis for the continuous oscillations of the Earth, *Geophys. J. Int.* **160**, 2005, 276–288.

Taqqu M.S., Self-similar processes, *Encyclopedia of Statistical Sciences* **vol. 8**, 1988, Wiley, New York, NY, 352–357.

Telesca L. and Lovallo M., Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane, *J. Stat. Mech.: Theory Exp.* 2011. https://doi.org/10.1088/1742-5468/2011/07/P07001

Telesca L., Colangelo G. and Lapenna V., Multifractal variability in geoelectrical signals and correlations with seismicity: a study case in southern Italy, *Nat. Hazard. Earth Syst. Sci*. **5**, 2005, 673–677.

Vogel M.A. and Wong A.K.C., PFS clustering method, *IEEE Trans. Pattern Anal. Mach. Intell* **1** (3), 1979, 237–245. https://doi.org/10.1109/TPAMI.1979.4766919

Zoller G., Holschneider M., Hainzl S. and Zhuang J., The largest expected earthquake magnitudes in Japan: the statistical perspective, *Bull. Seismol. Soc. Am.* **104** (2), 2014, 769–779. https://doi.org/10.1785/0120130103

Further Reading

Schreiber T. and Schmitz A., Surrogate time series, Physica D 142 (2000), 2000, 346–382.