Multidimensional Wavelet Analysis of Seismicity

A. A. Lyubushin

Institute of Physics of the Earth, Russian Academy of Sciences,
Bol'shaya Gruzinskaya ul. 10, Moscow, 123810 Russia,
lubushin@mtu-net.ru

ABSTRACT

A method of joint analysis of seismic regimes is proposed for recognition collective behavior phenomena of seismicity in a group of areas that form a large seismically active region. The method is based on the robust multidimensional wavelet analysis of square root values of earthquake energies released in each of the areas within successive uniform time intervals (thus, increments of so-called cumulative Benioff curves proportional to the values of elastic stresses accumulated and released in an earthquake source). This method is a further development of the method of wavelet-aggregated signals previously proposed by the author to analyze multidimensional time series of geophysical monitoring. It is based on robust multidimensional analysis of canonical components of wavelet coefficients. The method is exemplified by applying it to a number of seismically active regions.

INTRODUCTION

Methodologically, analysis of seismic catalogs is more difficult than processing of such traditional sources of geophysical information as time series (derived from seismic observations or low-frequency geophysical monitoring). This is due to the fact that the analysis of point processes [Cox, Lewis, 1966], including earthquakes sequences, does not allow the direct application of a vast variety of methods, parametrical models, and fast algorithms developed in the theory of signals [Brillinger, 1975; Marple , 1987; Hannan, 1970]. Actually, application of these methods requires a preliminary conversion of seismic catalogs to time series, which are sequences of values with a given constant time step. Formally, this conversion is not difficult and can be realized via calculation of either average values of a certain catalog parameter (for example, energy released during an earthquake) in successive non-overlapping time windows of a constant width or cumulative values of these characteristics with a constant time step (cumulative curves). However, the resulting time series are essentially non-Gaussian and include either outliers or step-like features (in cumulative curves) due to the time non-uniformity of seismic catalogs (gaps and groups of events such as swarms and aftershocks) and concentrating of major seismic energy in rare but strong events (the well-known problem of "heavy tails" of distributions). Although classical methods of the signal analysis, based on the Fourier transformation and calculating of covariances, are formally applicable to the processing of these time series, they are ineffective due to large biases in estimates caused by outliers (or steps).

In this paper, I consider the recognition of collective behavior features in seismic regimes of several areas that form a large seismically active region. I have previously proposed to solve a similar problem of extracting of a synchronizing signal from multidimensional time series of low-frequency geophysical monitoring observations, based on moving-window estimates for the evolution of eigenvalues of spectral matrices and canonical coherences (the synchronization

indicates a peak in the values of these statistics) [Lyubushin, 1993, 1994, 1998]. To solve the same problem, I proposed the concept of aggregated signal, which is a scalar signal carrying the maximum information on the most general variations that are present in all of the analyzed processes and simultaneously suppressing components that are characteristic of only one process and can be interpreted as local noise caused by specific local conditions of measurements, anthropogenic factors, or measurements uncertainties [Lyubushin, 1998]. Lyubushin *et al.* [1996, 1997] and Lyubushin [1998, 1999(a), 1999(b)] gave numerous examples of application of these methods to the recognition of earthquake precursors and to solution of other problems of geophysical monitoring. However, as was already mentioned above, these methods are mainly applicable to Gaussian time series, which reduces their potential in the analysis of seismic catalogs.

Below, to avoid this limitation, the signal is expanded in orthogonal finite functions called wavelets. The compactness of the basis functions involved in the signal expansion makes it possible to analyze not only Gaussian but also essentially non-stationary time series, which allows the application of non-parametric methods of analysis of multidimensional time series to non-Gaussian signals, including series obtained from seismic catalogs. This paper is actually develops the method of wavelet-aggregated signals proposed and developed in [Lyubushin, 1999(a), 2000, 2001].

One of the main properties of wavelets, which makes them attractive for using in problems of information compression is that they accumulates a maximum of an information in very small number (in percentage terms to total) of wavelet-coefficients [Daubechies, 1992; Chui, 1992; Press et al., 1996, Mallat, 1998]. As a consequence of this quality, the sets of wavelet coefficients are characterized by large outliers, which cannot be interpreted as an outcome of errors of measurements or failures of recording systems. The availability of these outliers puts restrictions on using methods of the data analysis based on transition from time area into area of wavelet coefficients (that is similar to transition into frequency area in classic Fourier analysis), as the set of wavelet coefficients represents essentially non-Gaussian sample. Thus, the conclusions based on the use of classic regression procedures, based on a method of least squares as well as classic methods of the multidimensional analysis based on usual sample estimates of covariance matrices and their eigenvalues and eigenvectors, should be accepted with known caution and with comprehension that the part of a useful information can be missed just because of the partial correspondence of used methods to a nature of analyzed data.

In statistics the problem of stability of conclusions to violation of the suppositions about a nature of data is known as a problem of statistical methods robustness and for the first time systematically is stated in the classical monograph [Huber, 1981], although the instability of least squares methods to small number of outliers is known for a long time before and was realized by many statisticians, including geophysicists, for example [Jeffreys, 1932]. At the same time the methods improving stability to availability of outliers were offered which main property consists in refusing of quadratic quality of fitness measure and transition to other measures growing not so fast, for example, to the module of discrepancies. The payment for increasing of stability of results of statistical processing is the essential complicating of computing procedures and magnification of computing time.

Wavelet-aggregating procedure includes two stages [Lyubushin, 2000]. The first stage initially involves the calculation of the wavelet coefficients for each time series under study and at each scale level using the fast discrete wavelet transformation. Before the transformation, the time series are converted to series in increments and are normalized in order to provide for the joint processing of diverse physical signals of different scales. The initial wavelet coefficients are converted to the so-called canonical wavelet coefficients. The latter are obtained from covariance matrices of

wavelet coefficients at each detail level using the method of canonical correlations. This conversion aims at removing individual noise (specific of only an individual series) from the wavelet coefficients and to amplify the common component. This procedure accomplishes the first stage.

At the second stage, the intensity of the common component is additionally increased by calculating the first principal component of the covariance matrices of canonical wavelet coefficients at each detail level. Thus, a scalar sequence of hypothetical wavelet coefficients is obtained at each detail level, which makes it possible to calculate the inverse discrete fast wavelet transform and to obtain the time realization of a scalar signal called the wavelet-aggregated signal of the initial time series. Since sample estimates of the covariance matrices are used, I introduce an algorithm parameter L_{min} (significance threshold) determining the minimum possible number of wavelet coefficients at a detail level corresponding to the time window width that can be used for sample estimation of the covariance matrix. The total number of coefficients decreases twofold as the number of the detail level increases; therefore, the aggregation can only be carried out for several first detail levels whose number depends on the window width and significance threshold. Below I use the value $L_{min} = 10$.

Thus, as it is clear from above-stated, the earlier offered procedure of a wavelet-aggregation is based on classic methods of the multidimensional analysis in space of wavelet coefficients and is not robust. At the present paper the robust modification of the wavelet-aggregated signal [Lyubushin, 2002] is applying to investigation of collective behavior effects of seismic process for two regions: (1) Japan, Kuril's islands and Kamchatka and (2) California.

As a case study, the method is applied to the recognition of collective phenomena in the seismicity of the following large regions:
   - Japan, the Kuril Islands, and southern Kamchatka;
   - California.

DESCRIPTION OF THE METHOD

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

$$x(t) = \sum_{\alpha=-\infty}^{+\infty} x^{(\alpha)}(t), \; x^{(\alpha)}(t) = \sum_{j=-\infty}^{+\infty} c^{(\alpha)}(\tau_j^{(\alpha)}) \cdot \psi^{(\alpha)}(t - \tau_j^{(\alpha)}), \; \tau_j^{(\alpha)} = j \cdot 2^\alpha \qquad (1)$$

Here α is a detail level number,

$$c^{(\alpha)}(\tau_j^{(\alpha)}) = \int_{-\infty}^{+\infty} x(t) \cdot \psi^{(\alpha)}(t - \tau_j^{(\alpha)}) dt \qquad (2)$$

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

$$\psi^{(\alpha)}(t) = (\sqrt{2})^{-\alpha} \cdot \Psi(2^{-\alpha} \cdot t), \;\; \psi^{(\alpha)}(t - \tau_j^{(\alpha)}) = (\sqrt{2})^{-\alpha} \cdot \Psi(2^{-\alpha} \cdot t - j) \qquad (3)$$

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

$$\Psi(t) = -1 \text{ for } t\in(0,\tfrac{1}{2}] \tag{4}$$
$$+1 \text{ for } t\in(\tfrac{1}{2},1] \text{ and zero for other } t,$$

then formula (1) corresponds to wavelet expansion of $x(t)$ by Haar wavelets. The most popular family of orthogonal wavelets are Daubechies wavelet functions $\Psi(t):=D_{2p}(t)$ of the order $2p$, which possess the following properties:

$$D_{2p}(t)=0 \text{ outside interval } [-p+1,p], \tag{5a}$$

$$\int_{-\infty}^{+\infty} t^k \cdot D_{2p}(t)dt=0 \quad \text{for} \quad k=0,1,..,(p-1) \tag{5b}$$

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

Let us consider now situation when $x(t)$ is a signal with discrete time t of the length N samples, $t=t_j = j\cdot\Delta t$, $j=1,..,N$, and suppose that N is an integer number of the form 2^m; this is convenient for further using of fast wavelet transform. If N is not equal to 2^m, then $x(t)$ is appended by zero values up to the length of 2^m, where m is the minimal integer for which $N\leq 2^m$. Formula for multi-resolution analysis in the case of finite number of samples and discrete time is as follows:

$$x(t)=d+ \sum_{\alpha=1}^{m} x^{(\alpha)}(t), \quad x^{(\alpha)}(t) = \sum_{j=1}^{2^{(m-\alpha)}} c^{(\alpha)}(\tau_j^{(\alpha)})\cdot\psi^{(\alpha)}(t-\tau_j^{(\alpha)}), \quad \tau_j^{(\alpha)} = j\cdot 2^{\alpha}\cdot\Delta t \tag{6}$$

The least-scale detail level in the discrete case is the first one and general number of detail levels m depends on the length N of the signal. A set of values $c^{(\alpha)}(\tau_j^{(\alpha)})$ and d could be computed by direct fast wavelet transform [Daubechies, 1992; Mallat, 1998; Press et al, 1996]. These coefficients uniquely define the initial signal $x(t)$, which could be restored by given values of $c^{(\alpha)}(\tau_j^{(\alpha)})$ and d by inverse fast wavelet transform. The wavelet coefficients of discrete transform (6) are equal to convolutions of basis functions $\psi^{(\alpha)}(t-\tau_j^{(\alpha)})$ of the detail level α with some signal of continuous time $\tilde{x}(s)$, which is obtained by an interpolation of a discrete signal $x(t)$ into arbitrary time points s between nodes t_j with the help of so called scaling function [Mallat, 1998]. The less is the sampling interval, the closer is interpolated signal $\tilde{x}(s)$ to an initial signal with continuous time.

Let q be a general number of time series $V^{(k)}(t)$, $t=1,..,N$; $k=1,..,q$; $q>2$; which represent synchronous records of geophysical fields variations (of the same or of the different physical nature) measured in spatially scattered points of monitoring system.

Usually time series of geophysical monitoring accumulate the main power of their variations in low frequencies. That is why we shall come from initial series to series in increments

("differentiated" series) and for preserve the same notation for incremental series as for the initial ones.

Let us introduce now a *moving time-window of adaptation* of the length of **r** samples and scale time series by such a way that they will have a uniform diapason of values (this is necessary for mutual processing of heterogeneous time series, having different scales and physical meaning):

$$U^{(k)}(t):=V^{(k)}(t)/(V^{(k)}_{max}(1,r)-V^{(k)}_{min}(1,r)) \text{ for } 1 \le t \le (r+1) \qquad (7a)$$

$$U^{(k)}(s+r):=V^{(k)}(s+r)/(V^{(k)}_{max}(s,r)-V^{(k)}_{min}(s,r)) \text{ for } s > 1 \qquad (7b)$$

$$V^{(k)}_{min}(s,r) = min_{s \le t \le s+r}V^{(k)}(t), \quad V^{(k)}_{max}(s,r) = max_{s \le t \le s+r}V^{(k)}(t) \qquad (7c)$$

Each time series is scaled in the first time-window to have a unit diapason (formula (7a)). In the following time-windows, shifted towards right-hand side in steps of one sample, scaling is made only for the rightmost sample **(s+r)**, without affecting results of previous scaling operations (formula (7b)). Thus, operations (7) realize adaptation to uniform diapasons of time series in the window of the length of **r** samples, taking into consideration information about variations only from the left-hand direction from the current time points. Such restriction will be hold on in all further computations in order to exclude the influence of post-seismic effects on the "past".

Let

$$Y^{(\alpha)}_{kj}=Y^{(\alpha)}_{k}(\tau^{(\alpha)}_{j}), \alpha = 1,...,m; k = 1,...,q; j = 1,...,2^{(m-\alpha)}; \tau^{(\alpha)}_{j} = j \cdot 2^{\alpha} \qquad (8)$$

be wavelet coefficients on the detail level α for wavelet decomposition of time series $U^{(k)}(t)$.

Let us denote by

$$Q^{(\alpha)}(s,r)=\{Y^{(\alpha)}_{kj}: s \le \tau^{(\alpha)}_{j} \le s+r, k = 1,...,q\}, j = 1,...,L^{(\alpha)}(r) \qquad (9)$$

- a set of **q**-dimensional vectors of wavelet coefficients of the level α of initial time series for those time indexes **j**, for which corresponding time values $\tau^{(\alpha)}_{j}$ lay inside time-window of the length **r** samples with initial sample **s**: $s \le \tau^{(\alpha)}_{j} \le s+r$. Let $L^{(\alpha)}(r)$ be a number of wavelet coefficients of the level α, for which the neighboring time indexes **j** lay within the same time-window of the length **r** samples. As for number of wavelet coefficients decreases with increasing of detail level index α as $2^{-\alpha}$, then $L^{(\alpha)}(r)$ decreases by the same law. It means that if **r<N**, then starting from some value of α all $L^{(\alpha)}(r)$ could be equal to zero and sets (8) could be empty.

In order to make all the next estimates statistically significant we shall introduce another parameter of wavelet-aggregating algorithm, the threshold L_{min}, which is a positive integer number. The sense of this parameter L_{min} consists in the fact that all estimates are performed only on those detail levels α for which

$$L^{(\alpha)}(r) \ge L_{min} \qquad (10)$$

For those detail levels α not satisfying condition (10), the wavelet-aggregating procedure for the given time-window length r could not be performed and appropriate wavelet coefficients for aggregated signal will be set to zero.

The former, non-robust, procedure of calculation of canonical wavelet-coefficients consist in the following. Let us denote by $R^{(\alpha)}(s,r)$ matrices of the size qxq, which represent sample estimates of covariance matrices of non-empty sets of vectors (9):

$$R^{(\alpha)}(s,r) = \frac{1}{L^{(\alpha)}(r)} \sum_{z \in Q^{(\alpha)}(s,r)} z \cdot z^T \qquad (11)$$

where z are q-dimensional column vectors of wavelet coefficients of the detail level α, laying inside adaptation time-window $[s,s+r]$. When computing (11), estimates of mean values of z are not subtracted from z because mean values of wavelet coefficients are equal to zero.

Let us now divide the components of vectors z into two parts: scalar z_1 and $(q-1)$-dimensional column vector of other components $\xi = (z_2,...,z_q)^T$. By scalar multiplying each vector ξ by some unknown still vector φ, we shall obtain a set of scalar values $\varsigma_1 = \varphi^T \cdot \xi$. Now we shall find vector φ from the condition that squared value of correlation coefficient between sets of scalar values z_1 and ς_1 be maximal. This is a special case of classic problem of Hotteling about canonical correlations: vector φ is defined as eigenvector, corresponding to maximal eiagenvalue (which equals to maximal correlation coefficient between sets of z_1 and ς_1) of the following matrix of the size $(q-1)x(q-1)$ [Hotelling, 1936; Rao, 1965]:

$$S_{\xi\xi}^{-1} \cdot S_{\xi z_1} \cdot S_{z_1 z_1}^{-1} \cdot S_{z_1 \xi} \qquad (12a)$$

$$S_{z_1 z_1} = cov(z_1, z_1), \quad S_{z_1 \xi} = S_{\xi z_1}^T = cov(z_1, \xi^T), \quad S_{\xi\xi} = cov(\xi, \xi^T) \qquad (12b)$$

It is evident that matrices in formulae (12a) and (12b) are sub-matrices of general covariance qxq-matrix: $S_{zz} = cov(z,z^T)$. Replacing matrix S_{zz} in (12) by its sample estimate (11), one can really computes vector φ and the set of scalar values ς_1. Let us call values ς_1 as *canonical wavelet coefficients* of scalar time series $U_1(t)$, corresponding to detail level α in the current time-window of adaptation $[s,s+r]$.

The meaning of this operation is as follows: if the component $U_1(t)$ in its variations, corresponding to the detail level α has some noises, which are intrinsic only for $U_1(t)$ and are absent in variations of other components, then in the set of canonical wavelet coefficients ς_1 they are absent simply as the consequence of algorithm. At the same time ς_1 accumulates all variations of $U_1(t)$ on the detail level α, which are common for all other components of the initial multiple time series for the set of wavelet coefficients ξ.

Performing similar operations with all other components of the vector z, we shall obtain a set of q-dimensional vectors of canonical wavelet coefficients $\varsigma = (\varsigma_1,...,\varsigma_q)^T$. For the given length of adaptation window r and for detail levels α that satisfy the condition (10), in the first adaptation

window (*s*=1), we shall remember all $L^{(\alpha)}(r)$ number of *q*-dimensional vectors ς. Further, shifting the adaptation window towards the right-hand side at steps of one sample, we shall repeat independently in each new window all the procedure of computing vector φ, and will compute and remember not all *q*-dimensional "cloud" of $L^{(\alpha)}(r)$ vectors ς for each window, but the only one vector that corresponds to the rightmost end of the window, i.e., to the sample *s+r*. In this way, the values of canonical wavelet coefficients are adapted to the collective behavior on the past time-window, which is of the length *r* of multi-dimensional signal components. Shifting of adaptation window to the right at steps of one sample is going on until *(s+r)=N*. As a result, a set of initial wavelet coefficients (8) for various detail levels that satisfies (10) will be replaced by a similar set of canonical wavelet coefficients.

For robust modification of canonical wavelet coefficients calculation we shall remark the following. Let's consider a problem of a regression of *(q-1)*-dimensional random vector $\xi = (z_2 ,..., z_q)^T$ on a scalar random variable z_1, i.e. the problem of an evaluation of a vector *u* of regression coefficients in the linear formula:

$$z_1 = \sum_{i=1}^{q-1} u_i z_{i+1} + \varepsilon_1 = u^T \xi + \varepsilon_1 \tag{13}$$

where ε_1 is a regression residual. If the vector *u* is defined by the least squares method:

$$\sum_{z \in Q^{(\alpha)}(s,r)} (\sum_{i=1}^{q-1} u_i z_{i+1} - z_1)^2 = \sum_{z \in Q^{(\alpha)}(s,r)} (u^T \xi - z_1)^2 \rightarrow min_u \tag{14}$$

then its estimate is easily obtained:

$$\widehat{u} = S_{\xi\xi}^{-1} \cdot S_{\xi z_1} \tag{15}$$

Let us denote by

$$\widehat{\xi}_1 = \widehat{u}^T \xi \tag{16}$$

– an estimate of regression share in the formula (13). Inasmuch as

$$cov(z_1 , \widehat{\xi}_1) = cov(z_1 , S_{\xi\xi}^{-1} \cdot S_{\xi z_1} \cdot \xi) = S_{z_1\xi} \cdot S_{\xi\xi}^{-1} \cdot S_{\xi z_1} \tag{17}$$

then it is easily to show that the square of correlation coefficient between (16) and z_1 is equal to the value $S_{z_1\xi} \cdot S_{\xi\xi}^{-1} \cdot S_{\xi z_1} \cdot S_{z_1 z_1}^{-1}$. The last term is nothing else as a maximum eigenvalue of the matrix (12a) [Rao, 1965].

Thus, it is shown, that the value of the first canonical wavelet-component can be determined as $\widehat{\xi}_1 = \widehat{u}^T \xi$ from a solution of regression problem (13)-(14). The similar statement, obviously, takes a place and for all other canonical wavelet components. This fact opens a possibility of determination of robust canonical wavelet components as a solutions of regression problem (13) in its robust modification, i.e. instead of (14) to decide a problem:

$$\sum_{z \in Q^{(\alpha)}(s,r)} | \sum_{i=1}^{q-1} u_i z_{i+1} - z_1 | = \sum_{z \in Q^{(\alpha)}(s,r)} | u^T \xi - z_1 | \rightarrow min_u \qquad (18)$$

It is obvious that the solution of a problem (18) is much more complicated, than (14), and can not be expressed by the simple formula of a type (15) and requires an iterative procedure. At the present realization the problem (18) was solved as follows. The sequence of iterations by a gradient method was organized, in which the generalized gradient of non-differentiable function (18) [Shor, 1979] with respect to components of vector **u** was taken. The step along a generalized anti-gradient was calculated by a solution of a problem of one-dimensional minimization by a method of a golden section. As an initial approximation for gradient iterative procedure the solution of a method of least squares under the formula (15) was taken, in which the covariance matrixes were evaluated by the usual formulas of a type (11), but before an evaluation of matrixes the data were subjected to a *windzorization* procedure [Huber, 1981]. The last operation presents an iterative procedure, for which on each step the sample estimates of average **s** and standard deviation σ are calculated. Afterwards a clipping operation was made for all outliers over the level **s±3σ**. These iterations are repeated so long as the evaluations of an average and standard deviation are not stabilized. The sense of *windzorization* is to avoid displacements of estimates of covariance matrices caused by availability of strong outliers. It is necessary to underline, that *windzorization* of wavelet coefficients was used only for obtaining the initial approximation of covariance matrices - all other analysis was made with initial wavelet coefficients (after preliminary scaling operations (7)). The stop condition for gradient procedure consists that the total number of iterations (i.e. calculations of a gradient) has achieved 10000, or that the step along an anti-gradient found in a method of a golden section, became less than 10^{-8}. After the gradient procedure has been stopped the problem (18) consider to be solved and formula (16) was calculated. The similar procedure was repeated for all components.

Performing cyclically calculation of $\widehat{\xi}_k$ for all **k=1,…,q** we can compute robust correlation coefficient $v_k(\tau, \alpha)$ between $\widehat{\xi}_k$ and z_k according to the formula [Huber, 1981]:

$$v_k(\tau, \alpha) = (S(a \cdot \widehat{\xi}_k + b \cdot z_k)^2 - S(a \cdot \widehat{\xi}_k - b \cdot z_k)^2) / (S(a \cdot \widehat{\xi}_k + b \cdot z_k)^2 + S(a \cdot \widehat{\xi}_k - b \cdot z_k)^2) \qquad (19)$$

where $\tau = (r + s)$ - is a right-hand time coordinate of the moving time window of adaptation; $S(X)$ is an absolute median deviation of the random variable X:

$$S(X) = med(|X - med(X)|), \ a = 1/S(\widehat{\xi}_k), \ b = 1/S(z_k) \qquad (20)$$

Argument α in the formula (19) underlines the fact that calculations are performed for wavelet coefficients, which belong to the detail level α.

Having values of $v_k(\tau, \alpha)$ we can estimate a robust wavelet coherence measure:

$$\mu(\tau, \alpha) = \prod_{k=1}^{q} v_k(\tau, \alpha) \qquad (21)$$

that indicates a measure of collective behavior of analyzed time series within current moving time window $s \leq \tau_j^{(\alpha)} \leq s + r$ on a given detail level α.

Estimating dependences $\mu(\tau, \alpha)$ is a final aim of the method. The more is the value of $\mu(\tau, \alpha)$, the more collective are variations of analyzed time series within current time window for the given detail level (which indicates the time scale or "period" of variations.

CASE STUDY

1. Region of Japan, Kuril islands and Kamchatka.

On the Fig.1 the hypocenters distribution of earthquakes with magnitudes $M \geq 4.5$, depths ≤ 100 km for the period 1963-2001 is presented for the region of Japan, Kuril islands and Kamchatka and partition of the region into 7 areas: J1-J5, KN and KS. The global catalogue NEIC [Global hypocenters data base, http://wwwneic.cr.usgs.gov/neis/] was used, for which the used lower threshold magnitude is representative on a considered space of time. The partition was made only visually using reasons for selection of large clusters of seismicity. The mnemonics for names of areas is rather transparent: "J" - Japan, "K" - Kamchatka and Kurils, accordingly northern and southern parts.

I have taken sum of square root values of earthquakes energies released in each of the areas within successive uniform time intervals of the length 5 days as time series to be analyzed. These values are proportional to released values of elastic stresses [Kasahara, 1981; Sobolev, 1992]. On the Fig.2 these values for corresponding areas after scaling operations (7) in the window of adaptation of the length 365 samples (i.e. 1825 days, that are equal to 5 usual years) are presented. It is visible, that the analyzed time series have essentially non-Gaussian and non-stationary character. Thus, the problem is to detect time intervals when these time series have hidden collective component. It is obvious also, that most suitable wavelet for the analysis of such data is the Haar's one.

Fig.3 contains graphs of dependences $\mu(\tau, \alpha)$ for the first 5 detail levels. More big-scale detail levels could not be processed as they do not satisfy the condition (10) for a given length of time window. It is quite natural that the amplitude of coherence measure variations increases with increasing of detail level number (i.e. with increasing of period). The 1st detail level indicates a negligible correlation. Nevertheless, it is interesting a burst of small correlation, corresponding to time interval 1983-1990 (note that we must subtract 5 years from the time mark in order to obtain the time interval). Other detail levels contain more or less prolonged time interval of correlation "bursts". Of course, they should be investigated from the point of view of their correspondence to some remarkable time points of seismic history of this region. But this is going beyond methodological scope of the article and could be a theme of independent investigation. From another point of view Fig.3 could give a scheme of hierarchical segmentation of the seismic history of the region into time intervals between peaks of correlation on different detail levels.

Another interesting peculiarity of the Fig.3 is a visible migration of the correlation peaks from big-scale detail levels to more and more "high-frequency" detail levels. This process begins in 1975-1980 and is not finished yet (it does not reach the 1st detail level till the end of 2001). Such behavior could be classified as precursor of oncoming seismic catastrophe in connection with hypothesis of log-periodic behavior before critical phenomena [Johansen et all, 1996; Sornette et all, 1996].

2. <u>California.</u>

Fig.4 presents spatial distribution of hypocenters of earthquakes in California with magnitude $M{\geq}3.0$ for 1963-2001. At first the region was divided into 8 areas Q1-Q8 and analyzed with lower magnitude threshold 3.0. But the results for such choice turn to be very noisy and not interesting. Afterwards the regions were joined into 3 more big ones: Q13, Q24 and Q5678 and the magnitude lower threshold was taken the same as for the 1st case – 4.5. This value of lower magnitude threshold corresponds to the seismic source size which is equal to the lithosphere thickness in this region and this choice seems to be more natural.

Fig.5 contains graphs of dependences $\mu(\tau,\alpha)$ for the first 5 detail levels. The first conclusion from Fig.5 is that seismic regime in California is much more correlated than Japan + Kurils + Kamchatka and this is quite natural because California is much less than those huge seismically active region, which was studied in the 1st case. At the same time the Fig.5 gives a rather noise pattern of correlation and it difficult to divide the seismic history into non-overlapping segments taking peaks of correlation as time marks. The following interesting features could be extracted from Fig.5: the burst of correlation on the 1st detail level occurs approximately at the same time interval as in the 1st case; there is essential increasing of correlation on the 4th and 5th detail level at the end, thus, corresponding to the interval 1995-2001.

CONCLUSION.

The robust scheme of wavelet aggregation of geophysical time series was suggested. The method is based on multidimensional analysis of wavelet coefficients. The analysis of geophysical data (increments of Beniof's curves for a set of seismically active regions) with a purpose to detect hidden signals of collective behavior was performed. The result gives a hierarchical pattern of arising of correlation peaks on different time intervals and for different time scales which could be used for extracting time segments of seismic history of the region and for searching strong earthquakes precursors. The method seems to be more effective for analyzing collective behavior effects in large seismically active regions than for medium and small regions.

REFERENCES

BRILLINGER D.R. (1975): *Time series. Data analysis and theory*. Holt, Rinehart and Winston, Inc., N.Y., Chicago, San Francisco.

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

COX D.R., LEWIS P.A.W. (1966) *The statistical analysis of series of events*. London, Methuen.

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

Global hypocenters data base, http://wwwneic.cr.usgs.gov/neis/

HANNAN E.J. (1970): *Multiple time series*. John Wiley and Sons, Inc., N.Y., London, Sydney, Toronto.

HOTELLING H. (1936). Relations between two sets of variates - Biometrika. Vol. 28. pp. 321-377.

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

JEFFREYS H. (1932) An alternative to the rejection of observations - Proc. Roy. Soc. Lond., v. 187, pp. 78-87.

JOHANSEN A., D.SORNETTE, H.WAKITA, U.TSUNOGAI, W.I.NEWMAN, H.SALEUR (1996): Discrete scaling in earthquake precursory phenomena: evidence in the Kobe earthquake, Japan, *Journ. Phys. I. France*, **6**, 1391-1402.

KASAHARA K. (1981): *Earthquake mechanics*. Cambridge University Press.

LYUBUSHIN A.A. (1993) Multidimensional Analysis of Time Series for a Geophysical Monitoring Systems. - Fizika Zemli. No. 2, pp. 103-108. *English translation*: Izvestiya, Physics of the Solid Earth, Vol. 29, 1994, pp. 297-301.

LYUBUSHIN A.A. (1994) Classification of States of Low-Frequency Systems of Geophysical Monitoring. - Fizika Zemli. 1994. No. 7/8, pp. 135-141. *English translation*: Izvestiya, Physics of the Solid Earth, Vol. 30, 1995, pp. 706-712.

LYUBUSHIN A.A, G.N.KOPYLOVA and YU.M.KCHATKEVICH. (1996) The Use of Multivariate Analysis for Processing Groundwater Precursors of Kamchatkan Earthquakes at the Petropavlovsk Site. - Vulcanologiya i Seismologiya. No. 1, pp. 79-97. *English translation*: Volcanology and Seismology, Vol. 18, 1996, pp. 87-106.

LYUBUSHIN A.A., V.A.MALUGIN and O.S.KAZANTCEVA. (1997) Monitoring of Tidal Variations of the Underground Water Level in a Group of Water-bearing Horizons - Fizika Zemli. No. 4, pp. 52-64. *English translation*: Izvestiya, Physics of the Solid Earth, Vol. 33, 1997, pp. 302-313.

LYUBUSHIN A.A. (1998) An Aggregated Signal of Low-Frequency Geophysical Monitoring Systems. - Fizika Zemli. No. 3, pp. 69-74. *English translation*: Izvestiya, Physics of the Solid Earth, Vol. 34, 1998, pp. 238-243.

LYUBUSHIN A.A. (1999(a)) Wavelet-Aggregated Signal in Earthquake Prediction //Earthquake Research in China (English Edition). Vol.13. No.1, pp.33-43.

LYUBUSHIN A.A. (1999(b)) Analysis of multidimensional geophysical monitoring time series for earthquake prediction //Annali di Geofisica, Vol.42, No.5, pp.927-937.

LYUBUSHIN A.A. (2000) Wavelet-Aggregated Signal and Synchronous Peaked Fluctuations in Problems of Geophysical Monitoring and Earthquake Prediction. - Fizika Zemli. No. 3, pp. 20-30. *English translation*: Izvestiya, Physics of the Solid Earth, Vol. 36, 2000, pp. 204-213.

LYUBUSHIN A.A. (2001) Multidimensional Wavelet Analysis of Geophysical Monitoring Time Series - Fizika Zemli. No.6, pp. 41-51. *English translation*: Izvestiya, Physics of the Solid Earth, vol. 37, 2001, pp. 474-483.

LYUBUSHIN A.A. (2002) Robust wavelet-aggregated signal for geophysical monitoring problems - Fizika Zemli. No.9. *English translation*: Izvestiya, Physics of the Solid Earth, vol. 38, 2002.

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

MARPLE S.L.(JR.) (1987): *Digital spectral analysis with applications*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

PRESS W.H., B.P.FLANNERY, S.A.TEUKOLSKY, and W.T.VETTERING. (1996) *Numerical Recipes, 2-nd edition, Chapter 13, Wavelet Transforms*, Cambridge Univ. Press, Cambridge.

RAO C.R. (1965) *Linear statistical inference and its applications*. John Wiley & Sons, Inc. N.Y., London, Sydney.

SHOR N.Z. (1979) *Methods of minimization non-differentiable functions and their applications (in Russian)*, Kiev, Naukova Dumka.

SOBOLEV G.A. (1993) *Fundamentals of Earthquake Prediction (in Russian)*, Moscow, Nauka.

SORNETTE D., A.JOHANSEN, J.-PH.BOUCHAUD (1996): Stock market crashes, precursors and replicas, *Journ. Phys. I. France*, **6**, 167-175.

Fig.1. Hypocenters of earthquakes with magnitudes $M{\geq}4.5$, depths $\leq$ 100 km for period 1963-2001 in the region of Japan, Kuril islands and Kamchatka and partition of region into 7 areas.
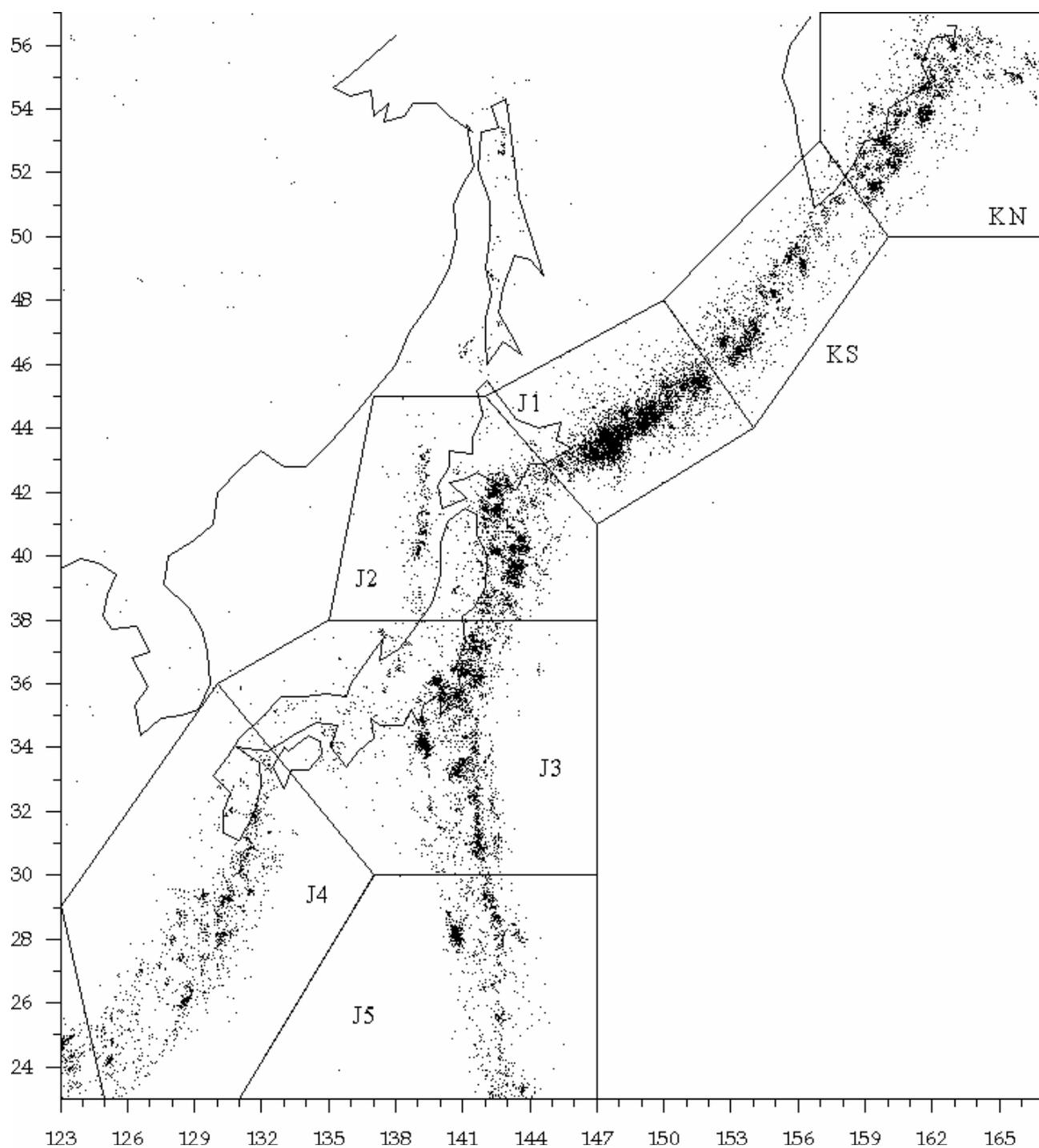
Fig. 2. Time series of increments of Benioff's curves in areas J1-J5 and KN and KS (Fig.3) with a time step 5 day after scaling operations (7) in the window of adaptation of length 365 samples, i.e. 1825 days, that is equal to 5 usual years.
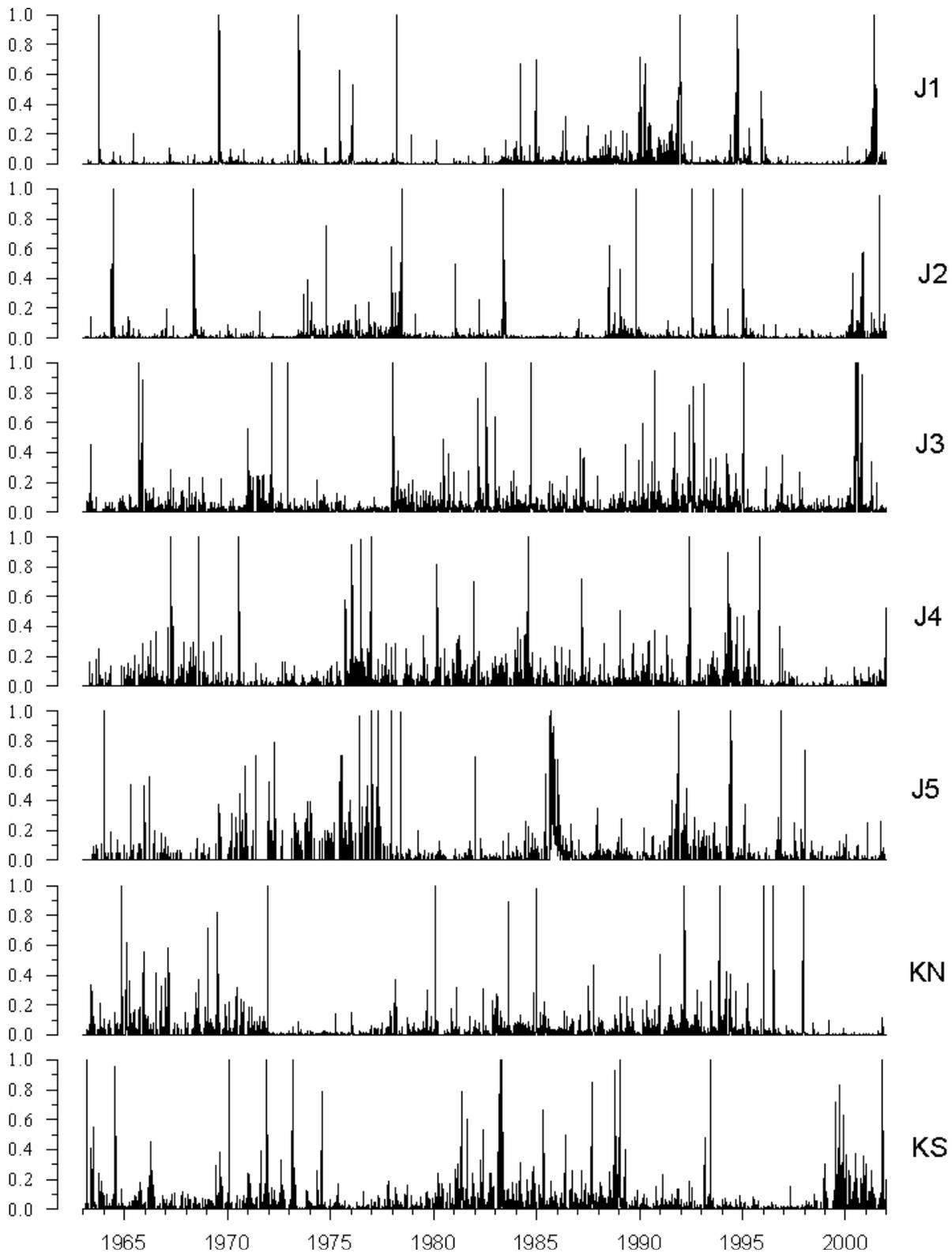
Fig.3. Evolution of the product $\mu(\tau,\alpha)$ of the robust canonical correlation coefficients, estimated in the moving time window. Time marks correspond to right-hand end of the time windows of the length = 5 years. Increments of Benioff's curves for 7 regions in Japan + Kurils + Kamchatka, $M \geq 4.5$, depth ≤ 100 km, 1963-2001, Haar's wavelets, length of time window = 365 samples (5 years).
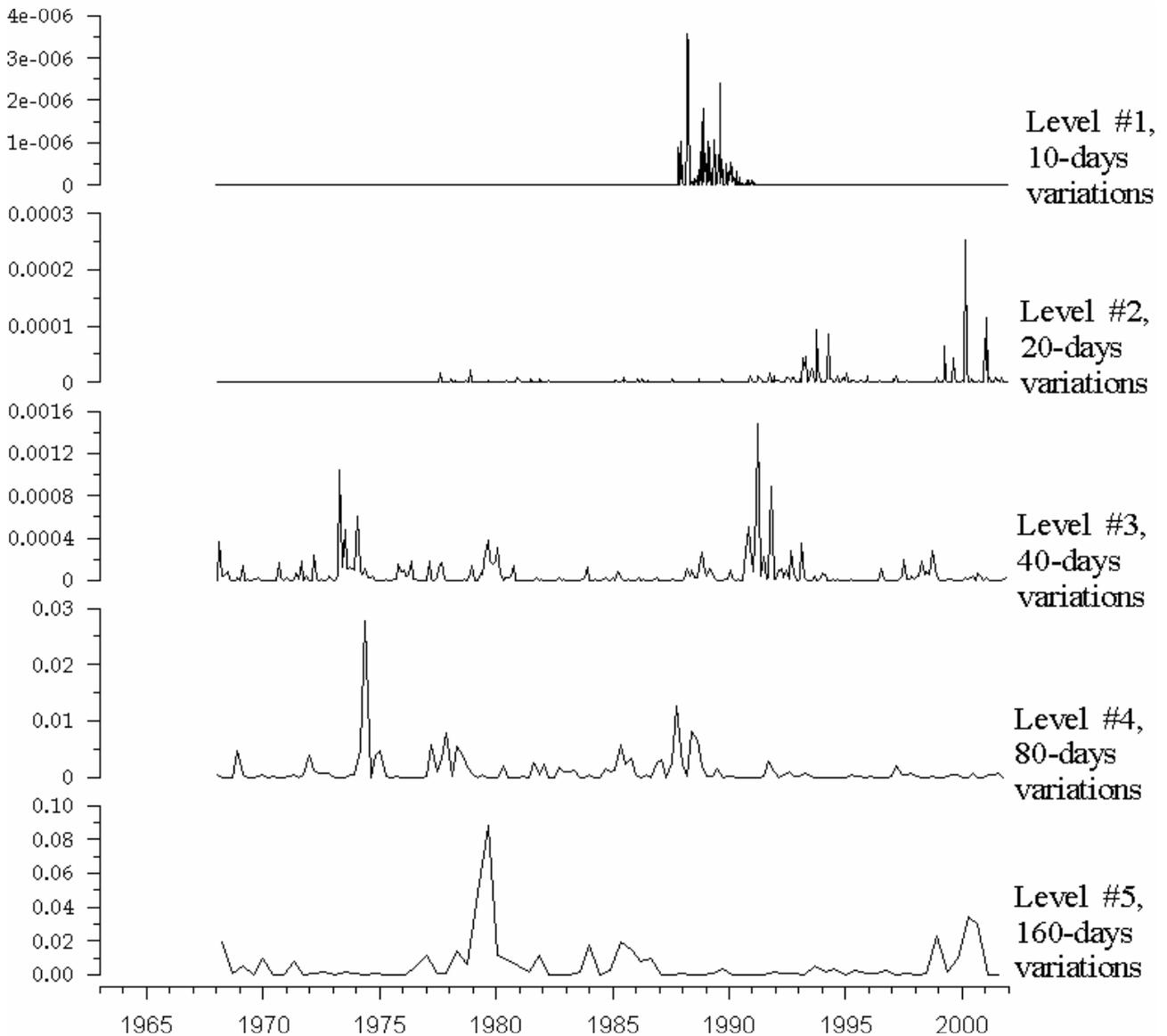
Fig.4. Hypocenters of earthquakes with magnitudes *M*≥3.0, for period 1963-2001 in the California and division of the region.
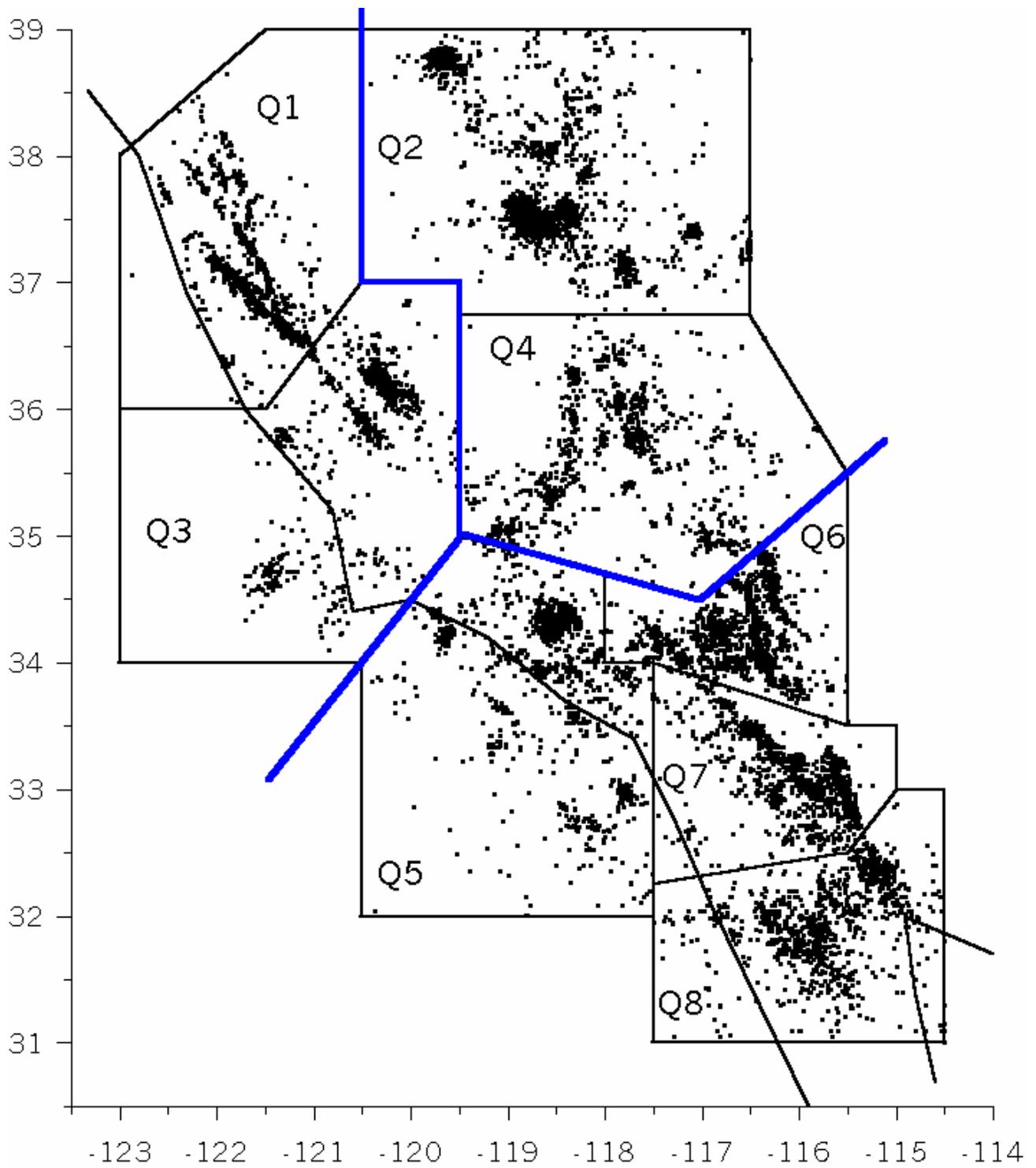
Fig.5. Evolution of the product $\mu(\tau,\alpha)$ of the robust canonical correlation coefficients, estimated in the moving time window. Time marks correspond to right-hand end of the time windows of the length = 5 years. Increments of Benioff's curves for 3 regions in California: Q13, Q24, Q5678, $M \geq 4.5$, 1963-2001, Haar's wavelets, length of time window = 365 samples.