Izvestiya, Physics of the Solid Earth, Vol. 38, No. 9, 2002, pp. 745–755. Translated from Fizika Zemli, No. 9, 2002, pp. 37–48.

A Robust Wavelet-Aggregated Signal for Geophysical Monitoring Problems

A. A. Lyubushin, Jr.
Institute of Physics of the Earth, Russian Academy of Sciences, Bol'shaya Gruzinskaya ul. 10, Moscow, 123995 Russia

Abstract—A modification of the method of wavelet-aggregated signal intended for the analysis of multidimensional time series of monitoring systems, with account taken of the existence of large outliers in the values of wavelet coefficients, is considered. This property follows from the ability of wavelet expansions to accumulate a maximum amount of information in a relatively small amount of coefficients, as well as from the nature of the data analyzed, in which outliers are not due to errors of recording systems. The non-Gaussian property of the distribution of wavelet coefficients requires modification of conventional schemes used for the multidimensional analysis of canonical and principal components in order to make results of this analysis statistically more reliable. An approach to the solution of this problem is presented, along with its applications to the analysis of geophysical data (time series of monitoring systems and seismic catalogs) and comparison of their results with those obtained by applying the previous nonrobust approach.

INTRODUCTION

A modified method of constructing the aggregated signal for multidimensional time series of monitoring systems is presented. Qualitatively, the aggregated signal can be defined as a scalar signal accumulating in its own variations only those spectral (or wavelet) components that are simultaneously present in each scalar time series of the multidimensional signal analyzed. Moreover, an algorithm of aggregations proposed in this work suppresses components that are present in some of processes but absent in the others (these components could be called local disturbance signals, for example, of the anthropogenic nature). The main purpose of constructing the aggregated signal is to make clearer a common tendency in low-frequency data flow from geophysical networks, indicating an increase in the collective behavior. Previously, I proposed two schemes of aggregation for multidimensional time series, one based on the Fourier transform (Fourier-aggregated signal) [Lyubushin, 1998] and another based on expansions in full finite orthogonal basis functions (the wavelet-aggregated signal) [Lyubushin, 1999a, 2000a]. An advantage of using wavelet expansions in the analysis of geophysical data follows from their main property, compactness, enabling the analysis of strongly nonstationary and non-Gaussian signals [Lyubushin, 2000b, 2001].

One of the main properties of wavelets, making them attractive for using in problems of information compression, is that they accumulate a maximum amount of information in a relatively very small number of wavelet coefficients [Daubechies, 1992; Press *et al.*, 1996, Mallat, 1998]. As a result, sets of wavelet coefficients are characterized by large outliers that cannot be interpreted in terms of measurement uncertainties or failures of recording systems. The presence of these outliers imposes restrictions on methods of data analysis based on transition from the time domain to the domain of wavelet coefficients (similar to the transition to the frequency domain in the classical Fourier analysis), because the set of wavelet coefficients is an essentially non-Gaussian sample. Thus, conclusions based on the use of classical regression procedures, the least-squares method, and classical methods of the multidimensional analysis using ordinary sample estimates of covariance matrices and their eigenvalues and eigenvectors should be accepted with caution, bearing in mind the fact that a part of useful information can be missed exactly because these methods are not fully adequate to the nature of data analyzed.

In statistics, the problem of sensitivity of results to the violation of assumptions on the nature of data is known as the robustness problem of statistical methods; for the first time, this problem was systematically treated in the classical monograph [Huber, 1981], although the sensitivity of least-squares methods to a small number of outliers has long been known and was realized by many specialists in statistics, including geophysicists (e.g., by H. Jeffreys). At the same time, methods improving the stability with respect to outliers were proposed; their main property is the replacement of a quadratic measure used for estimating the fit quality by other measures increasing less rapidly (for example, the modulus of discrepancies). The tradeoff for increasing the stability of statistical processing results is essentially more complicated numerical algorithms and an increase in the computation time.

Wavelet-aggregated signal is constructed in two stages [Lyubushin, 2000a]. The first stage initially involves the calculation of 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 into series in increments and are normalized in order to enable joint processing of diverse physical signals of different scales. The initial wavelet coefficients are converted into 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 to 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 main 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 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 suitable for the 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, the number of which depends on the window width and significance threshold. Below, I use the value $L_{\min} = 10$.

Thus, as is clear from the above, the previously proposed procedure of wavelet aggregation is based on classical methods of the multidimensional analysis in the space of wavelet coefficients and is not robust. In this work, a robust modification of the wavelet-aggregated signal and results of its application to the analysis of real geophysical data are presented.

DESCRIPTION OF THE METHOD

Orthogonal multiresolution analysis of a signal $x(t)$ is defined by the formula [Daubechies, 1992; Mallat, 1998]

$$x(t) = \sum_{\alpha=-\infty}^{+\infty} x^{(\alpha)}(t), \quad x^{(\alpha)}(t) = \sum_{j=-\infty}^{+\infty} c^{(\alpha)}(\tau_j^{(\alpha)}) \times \psi^{(\alpha)}(t-\tau_j^{(\alpha)}), \quad \tau_j^{(\alpha)} = j2^{\alpha}. \tag{1}$$

Here, α is the number of a detail level,

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

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

$$\psi^{(\alpha)}(t) = (\sqrt{2})^{-\alpha}\Psi(2^{-\alpha}t), \quad \psi^{(\alpha)}(t-\tau_j^{(\alpha)}) = (\sqrt{2})^{-\alpha}\Psi(2^{-\alpha}t-j). \tag{3}$$

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

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

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

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

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

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

Now I address the situation when $x(t)$ is a signal of N samples in length with a discrete time $t = t_j = j\Delta t$, $j = 1, \ldots, N$. Let N be an integer of the form 2^m; this is convenient for further applications of the fast wavelet transform. If N is not equal to 2^m, $x(t)$ is complemented by zero values up to a length of 2^m, where m is the minimal integer such that $N \le 2^m$. The formula for the multiresolution analysis in the case of a finite number of samples and a discrete time is

$$x(t) = d + \sum_{\alpha=1}^{m} x^{(\alpha)}(t), \quad x^{(\alpha)}(t) = \sum_{j=1}^{2^{(m-\alpha)}} c^{(\alpha)}(\tau_j^{(\alpha)}) \times \psi^{(\alpha)}(t-\tau_j^{(\alpha)}), \quad \tau_j^{(\alpha)} = j2^{\alpha}\Delta t. \tag{6}$$

The least-scale detail level in the discrete case is the first one, and the total number of detail levels m depends on the length N of the signal. The values of $c^{(\alpha)}(\tau_j^{(\alpha)})$ and d can be calculated by the direct fast wavelet transform [Daubechies, 1992; Mallat, 1998; Press *et al.*, 1996]. These coefficients uniquely define the initial signal $x(t)$, which can be retrieved from given values of $c^{(\alpha)}(\tau_j^{(\alpha)})$ and d by inverse fast wavelet transform.

Let q be the total number of time series $V^{(k)}(t)$, $t = 1, \ldots, N$; $k = 1, \ldots, q$; $q > 2$ representing synchronous records of geophysical fields variations (of the same or different physical origins) measured by the spatial network of a monitoring system.

Time series of low-frequency geophysical monitoring usually accumulate the main power of their variations in a low-frequency range. For this reason, the initial series are transformed into series in increments ("differentiated" series), preserving the notation used for the former.

Furthermore, I introduce a moving time window of adaptation of r samples in length and scale time series in order to normalize them to a common range of values (this is necessary for joint analysis of diverse time series differing scale and physical meaning):

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

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

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

Each time series is scaled in the first time window to have a unity range (Eq. (7a)). In the next time windows, each shifted to the right by one sample, scaling is made only for the rightmost sample $(s + r)$, without affecting the results of previous scaling operations (Eq. (7b)). Thus, operations (7) effect the adaptation to a uniform range of time series in an r-sample window, taking into consideration the information about variations only on the left-hand side from the current time point. Such a restriction will be held to below in order to exclude the influence of post-seismic effects on the "past."

Let

$$Y_{kj}^{(\alpha)} = Y_k^{(\alpha)}(\tau_j^{(\alpha)}), \quad a = 1, \ldots, m; \quad k = 1, \ldots, q; \quad j = 1, \ldots, 2^{(m-\alpha)}; \quad \tau_j^{(\alpha)} = j2^{(\alpha)} \tag{8}$$

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

The wavelet-aggregated signal is constructed in two stages. The first stage is to replace wavelet coefficients (after scaling operations (7)) by canonical wavelet coefficients. Let

$$Q^{(\alpha)}(s, r) = \{Y_{kj}^{(\alpha)}: s \le \tau_j^{(\alpha)} \le s + r, k = 1, \ldots, q\}, \quad j = 1, \ldots, L^{(\alpha)}(r) \tag{9}$$

be a set of q-dimensional vectors of wavelet coefficients of the α level in the initial time series, with the time indexes j being such that the corresponding time values lie within the time window of r samples in length with the initial sample s: $s \le \tau_j^{(\alpha)} \le s + r$. Let $L^{(\alpha)}(r)$ denote the number of wavelet coefficients of the level α for which the neighboring time indexes j lie within the same time-window r samples long. Since the number of wavelet coefficients decreases with an increasing detail level index as the geometric progression with a ratio of 2, $L^{(\alpha)}(r)$ decreases by the same law. If $r < N$, this means that, starting from a certain α, all $L^{(\alpha)}(r)$ can be equal to zero; i.e., sets (9) will be empty.

In order to make all subsequent estimates statistically significant, I introduce another parameter of the wavelet-aggregating algorithm, namely the threshold, which is a positive integer number. The meaning of this parameter consists in the fact that all estimates are made only at detail levels α such that

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

The wavelet-aggregating procedure with a the given time window length r cannot be performed at detail levels that do not satisfy condition (10), and appropriate wavelet coefficients for aggregated signal will be set to zero.

The previous, nonrobust, procedure of calculating canonical wavelet coefficients consists in the following. Let $R^{(\alpha)}(s, r)$ denote $q \times q$ matrices representing sample estimates of covariance matrices of nonempty sets of vectors (9):

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

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

Now I divide the components of vectors z into two parts: the scalar z_1 and the $(q-1)$-dimensional column-vector of the other components $\xi = (z_2, \ldots, z_q)^T$. The scalar product of each vector ξ by an unknown vector φ yields the set of scalar values $\zeta_1 = \varphi^T\xi$. The vector φ is found from the condition that the squared value of the correlation coefficient between the sets of scalars z_1 and ζ_1 is maximal. This is a special case of the classical problem of Hotelling ob canonical correlations: the vector φ is defined as an eigenvector corresponding to a maximum eigenvalue (which is exactly the maximum squared correlation coefficient between the sets of z_1 and ζ_1) of the $(q-1) \times (q-1)$ matrix [Hotelling, 1936; Rao, 1965]

$$S_{\xi\xi}^{-1} S_{\xi z_1} S_{z_1 z_1}^{-1} S_{z_1 \xi}, \tag{12a}$$

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

It is evident that the matrices in (12a) and (12b) are sub-matrices of the general covariance $q \times q$ matrix $S_{zz} = \mathrm{cov}(z, z^T)$. Replacing the matrix S_{zz} in (12) by its sample estimate (11), the vector φ and the set of scalar values ζ_1 can be calculated. The values ζ_1 are referred to as canonical wavelet coefficients of the scalar time series $U_1(t)$ of the detail level α in the current adaptation time window [s, s + r].

The meaning of this operation is as follows: if the component $U_1(t)$ corresponding to the detail level α includes noise present solely in this series and absent in variations of other components, the set of canonical wavelet coefficients is free of such noise simply on the strength of their construction. On the other hand, ζ_1 retains all variations of the detail level α that are common for all other components of the initial multiple time series for the sets of wavelet coefficients ξ.

Performing similar operations with all other components of the vector z, we obtain the set of q-dimensional vectors of canonical wavelet coefficients $\zeta = (\zeta_1, \ldots, \zeta_q)^T$. For a given length of adaptation window r and detail levels α that satisfy condition (10), in the first adaptation window (s = 1), we save all $L^{(\alpha)}(r)$ q-dimensional vectors ζ. Further, displacing the adaptation window to the right by one sample, we repeat independently in each new window the entire procedure of determining the vector φ; however, in doing so, we calculate and remember not the entire q-dimensional cloud of the vectors ζ for each window but 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 of the multidimensional signal in the past time window of r samples in length r. The adaptation window is displaced to the right by one sample until the condition $(s + r) = N$ is satisfied.

As a result, a set of initial wavelet coefficients (8) at various detail levels that satisfies (10) will be replaced by a similar set of canonical wavelet coefficients. This accomplishes the first stage of the wavelet-aggregation procedure.

With regard to the robust modification of the calculation of canonical wavelet coefficients, the following should be noted. Consider the problem of regression of $(q-1)$-dimensional random vector $\xi = (z_2, \ldots, z_q)^\mathrm{T}$ on a scalar random variable z_1, i.e., the problem of evaluating the 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 the regression residual. If the vector u is defined in terms of the least-squares method,

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

its estimate is easily obtained as

$$\hat{u} = S_{\xi\xi}^{-1} S_{\xi z_1}. \tag{15}$$

Let

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

denote the estimate of the regression contribution in formula (13). Since

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

it is easy to show that the squared correlation coefficient between (16) and z_1 is equal to $S_{z_1 \xi} S_{\xi\xi}^{-1} S_{\xi z_1} S_{z_1 z_1}^{-1}$, which is the same as the maximum eigenvalue of matrix (12a) [Rao, 1965].

Thus, it is shown that the first canonical wavelet component can be determined as $\hat{\xi}_1 = \hat{u}^T \xi$ from the solution of regression problem (13)–(14). Obviously, a similar statement is also valid for all other canonical wavelet components. This fact opens the possibility of determining robust canonical wavelet components as solutions of regression problem (13) in its robust modification, i.e., to solve, rather than (14), the problem

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

The solution of (18) is evidently much more complicated than (14), cannot be expressed by a simple formula of type (15), and requires an iterative procedure. In the given realization, problem (18) was solved as follows. A sequence of iterations by a gradient method was organized, in which the generalized gradient of nondifferentiable function (18) [Shor, 1979] with respect to the components of the vector u was taken. The step along the generalized antigradient was calculated by solving the problem of one-dimensional minimization with the use of the method of golden section. As an initial approximation for the iterative gradient procedure, the least-squares solution (15) was taken; the covariance matrices in (15) were evaluated by usual formulas of type (11), with the data being preliminarily subjected to the vinzorization procedure [Huber, 1981].

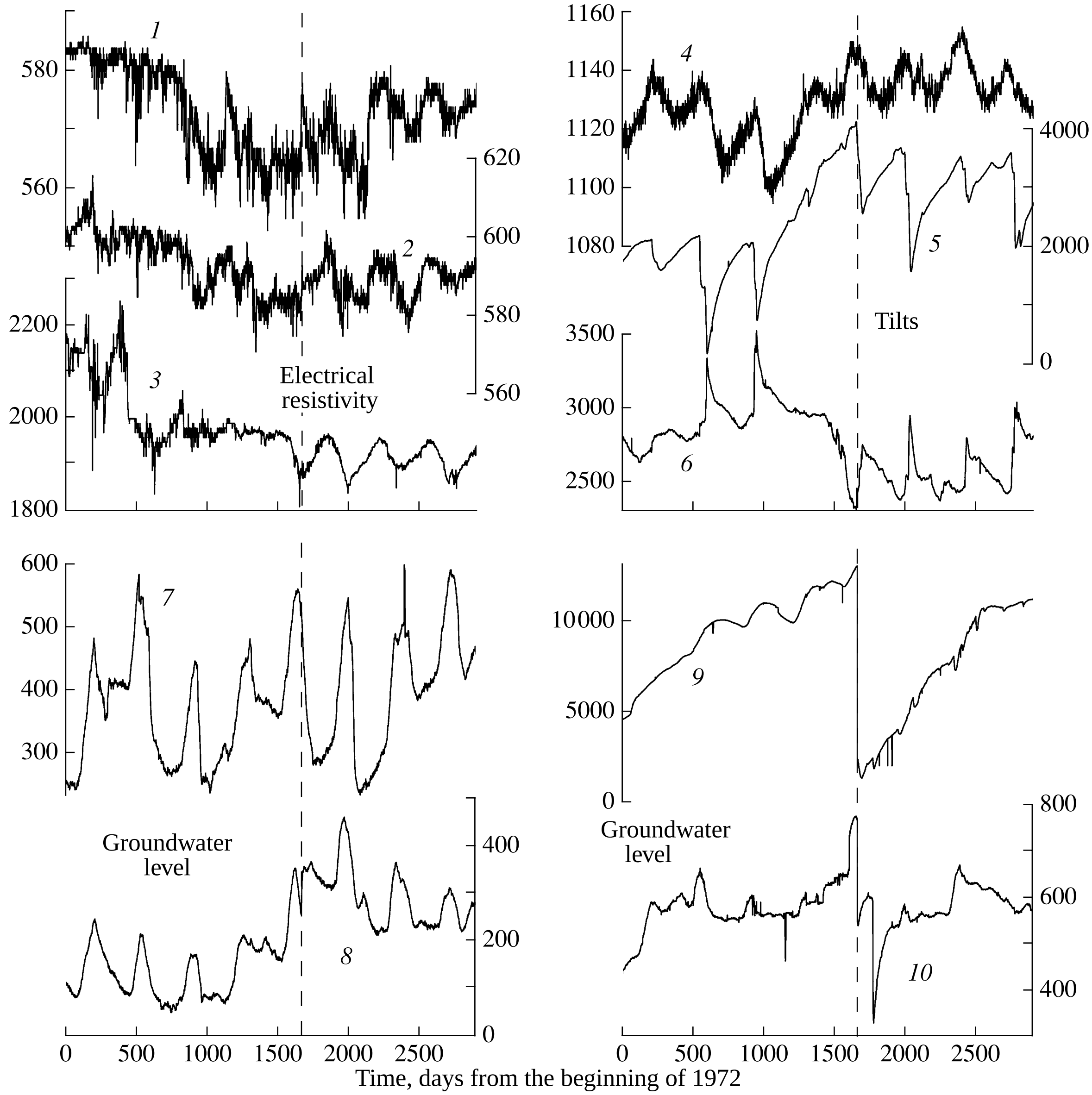

Fig. 1. Plots of ten initial time series (*1–10*) from the geophysical monitoring network in the northeastern China. The observations were conducted from January 1, 1972, to December 31, 1979, with a sampling time interval of 1 day. Vertical dashed lines show the time moment of the Tangshan, $M = 7.8$ earthquake of July 28, 1976 (1671st day from the beginning of 1972).

The latter is an iterative procedure calculating sample estimates of the average s and standard deviation σ at each step, after which all outliers exceeding the level $s \pm 3\sigma$ are rejected. These iterations are repeated until the average and standard deviation estimates are stabilized. The purpose of the vinzorization is to avoid biases in the estimates of covariance matrices caused by the presence of large outliers. It is necessary to emphasize that the vinzorization of wavelet coefficients was used only for obtaining the initial approximation of covariance matrices; in all other respects, the analysis was made with the initial wavelet coefficients (after performing preliminary scaling operations (7)). The gradient procedure stopped when either the total number of iterations (i.e., calculations of the gradient) reached 10000 or the step along the antigradient found by the method of golden section became less than 10^{-8}. After the gradient procedure had been stopped, problem (18) was regarded to have been solved and the canonical wavelet component was calculated by formula (16). A similar procedure was repeated for all components. In all other details, the method of calculating canonical wavelet coefficients did not vary.

The second stage of the aggregation procedure is the calculation of principal components of canonical wave-

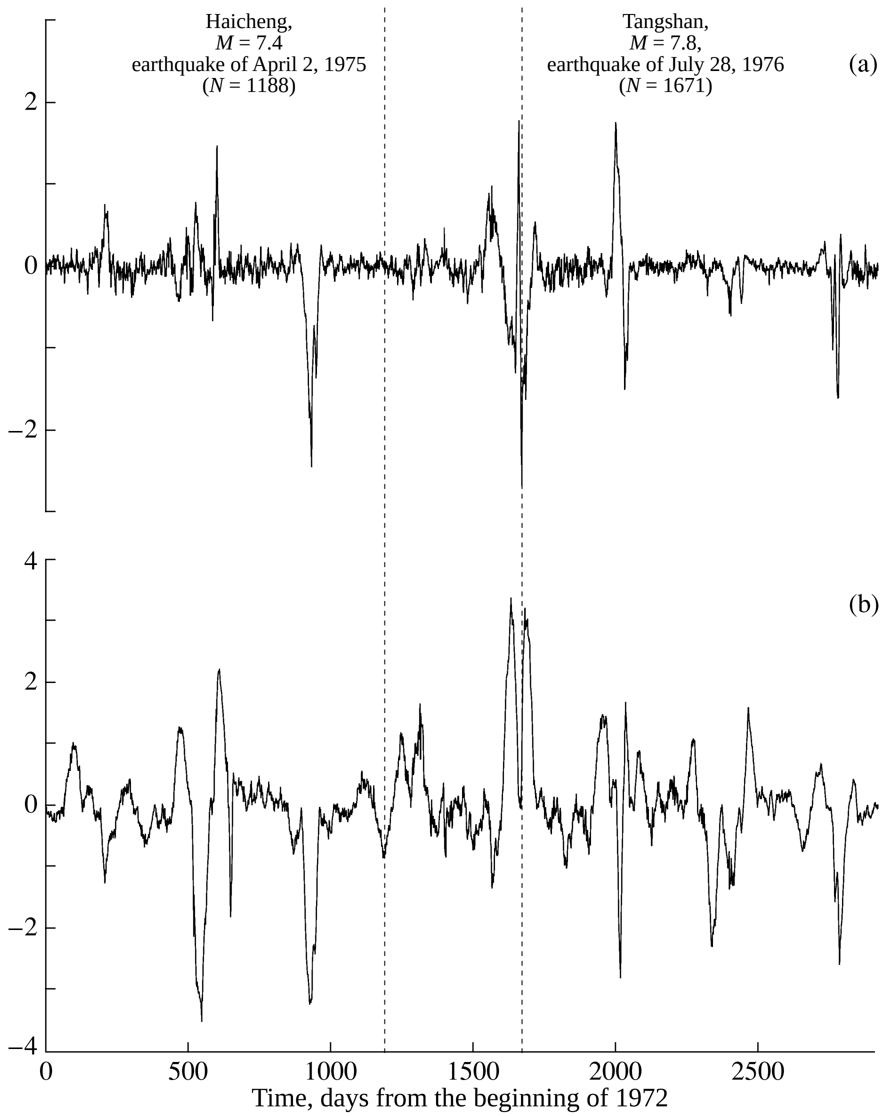

Fig. 2. Plots of the wavelet-aggregated signals from the time series presented in Fig. 1 and estimated for an adaptation window length of 700, $L_{\min} = 10$, and Haar wavelets: (a) nonrobust wavelet-aggregated signal; (b) its robust modification. The vertical dashed lines indicate the time moments of the Haicheng and Tangshan earthquakes.

let coefficients. In the previous, nonrobust realization of the algorithm, this stage consisted in the following.

A sample estimate of the covariance matrix of canonical wavelet coefficients at a given detail level (and in a current adaptation window) is calculated by a formula analogous to (11). Let $\chi^{(\alpha)}$ denote a q-dimensional eigenvector corresponding to the maximum eigenvalue of this Hermitean, positively defined matrix. Multiplying it scalarly by each vector ζ yields a set of values of the first principal component of the vectors ζ that possesses properties of preserving the maximum information [Rao, 1965] about correlations between components of the q-dimensional vectors of canonical wavelet coefficients within the current adaptation window:

$$W_j^{(\alpha)} = \sum_{k=1}^{q} \chi_k^{(\alpha)} \zeta_{kj}^{(\alpha)}. \tag{19}$$

Quite similar to the first stage of aggregation, all values of $W_j^{(\alpha)}$ $(1 \le j \le (1 + r))$ obtained in the first adaptation window are saved for each detail level α satisfying (10). In subsequent windows of adaptation, only the values at right-hand end of a window, i.e., at $j = (s + r)$, are saved. Thus, the values of $W_j^{(\alpha)}$ are determined for all α satisfying (10) and for all j: $1 \le j \le N2^{-\alpha}$. We set

$W_j^{(\alpha)} = 0$ for α that do not satisfy condition (10) and for j in the interval $N2^{-\alpha} < j \leq 2^{(m-\alpha)}$.

The resulting values $W_j^{(\alpha)}$ are regarded as wavelet coefficients of a certain scalar signal $W(t)$ that can be found by the inverse fast wavelet transform and is called a wavelet-aggregated signal. Consequently, the wavelet-aggregated signal is a signal whose wavelet coefficients are the first principal components of the canonical wavelet coefficients of the initial time series.

To make the second stage also robust, one should resort to the known fact related to the method of principal components [Rao, 1965]: the vector $\chi^{(\alpha)}$ is the solution of the problem on conditional maximums:

$$\sum_j \left(\sum_{k=1}^{q} \chi_k^{(\alpha)} \zeta_{kj}^{(\alpha)} \right)^2 \longrightarrow \max_{\chi^{(\alpha)}}, \quad \sum_{k=1}^{q} (\chi_k^{(\alpha)})^2 = 1. \quad (20)$$

The formulation of the problem on robust determination of principal components follows from (20) in a natural way:

$$\sum_j \left| \sum_{k=1}^{q} \chi_k^{(\alpha)} \zeta_{kj}^{(\alpha)} \right| \longrightarrow \max_{\chi^{(\alpha)}}, \quad \sum_{k=1}^{q} (\chi_k^{(\alpha)})^2 = 1. \quad (21)$$

The solution of problem (21) was found by an iterative procedure. As an initial approximation of the vector $\chi^{(\alpha)}$, the eigenvector corresponding to the maximum eigenvalue of the covariance matrix of robust canonical wavelet coefficients found from the solution of problem (18) was taken. The sample estimates of covariance matrices were also derived only after application of the vinzorization procedure. Problem (21) was solved by the gradient method (using generalized gradients) with the projection of the gradient step onto the unit sphere. A method for choosing the along-gradient step and the conditions for stopping iterations were same as in the solution of problem (18).

This accomplishes the description of the procedure of constructing a robust wavelet-aggregated signal. Parameters of the latter are the Daubechies wavelet order p, length of the adaptation moving time-window r (measured in terms of the number of samples), and threshold of significance $L_{\min}$ for estimating covariation matrices (10).

EXAMPLES OF GEOPHYSICAL DATA ANALYSIS

1. Geophysical monitoring time series from northeastern China. Figure 1 presents the plots of the initial time series. These data consist of 10 time series representing synchronous observations of the following geophysical parameters:

electrical resistivity of rocks (3 time series, plots *1–3* in Fig. 1);

tilts (3 time series, plots *4–6* in Fig. 1);

groundwater level variations in wells (four time series, plots *7–10* in Fig. 1).

These data were kindly placed at my disposal by Prof. Zhang Zhaocheng, Center for Analysis and Prediction of Earthquakes, State Seismological Bureau, China. A characteristic linear size of the observational network is about 200 km. The observations were conducted during eight years, from January 1, 1972, to December 31, 1979. The sampling time interval was 1 day. Thus, the length of each time series is 2922 samples. The observation period included a catastrophic (M = 7.8) Tien-Shan earthquake of July 28, 1976. This event coincided in time with the 1671st day from the beginning of 1972 and has the most distinct signature in the postseismic response of the groundwater level variations (plot *9* in Fig. 1).

Previously, these data were analyzed using the Fourier aggregated signal in [Lyubushin, 1999b] and non-robust variant of the wavelet-aggregated signal in [Lyubushin, 1999a, 2000a]. The wavelet-aggregated signal was constructed an adaptation window length of 700 samples, $L_{\min}$ =10, and Haar wavelets. It was shown that 100 days before the Tangshan earthquake a distinct precursory anomaly started to develop and became most pronounced 5 days before the event.

Figure 2a plots the older variant of the wavelet-aggregated signal exhibiting a well-resolved anomaly preceding the Tine Shan earthquake (the second vertical dashed line). Figure 2b presents the plot of the robust modification of the aggregated signal that also reliably resolves the precursor of the Tine Shan earthquake. I emphasize once more that the algorithm of time series aggregation is left-oriented. This means that, if only the initial segment of an arbitrary length were processed, not the whole sample, the rejection of the remaining data segment would have no effect on the result of the analysis of the first part: the aggregated signal from the initial segment would be the same as that obtained by processing the entire information.

At first glance, Fig. 2a is more impressive, because the anomalies observed in this plot have higher amplitudes relative to the general background noise of statistical fluctuations. However, the Tangshan earthquake was preceded in this region by the Haicheng (M = 7.4) earthquake of April 2, 1975, i.e., on the 1188th day from the beginning of 1972. The latter appears to be the only earthquake that was successfully predicted (by using other data) [Kasahara, 1981; Sobolev, 1993]. It is marked by another vertical dashed line in Fig. 2. Discouragingly, no statistically significant effect of synchronization pre- or postdates this event in Fig. 2a. However, its both precursory and postseismic effects are noticeable in Fig. 2b. Moreover, all anomalies noticeable in Fig. 2a are also present in Fig. 2b, where they are even more contrasting.

Thus, the robust wavelet-aggregated signal in the given special case is more sensitive. However, it is premature as yet to conclude that the robust aggregation

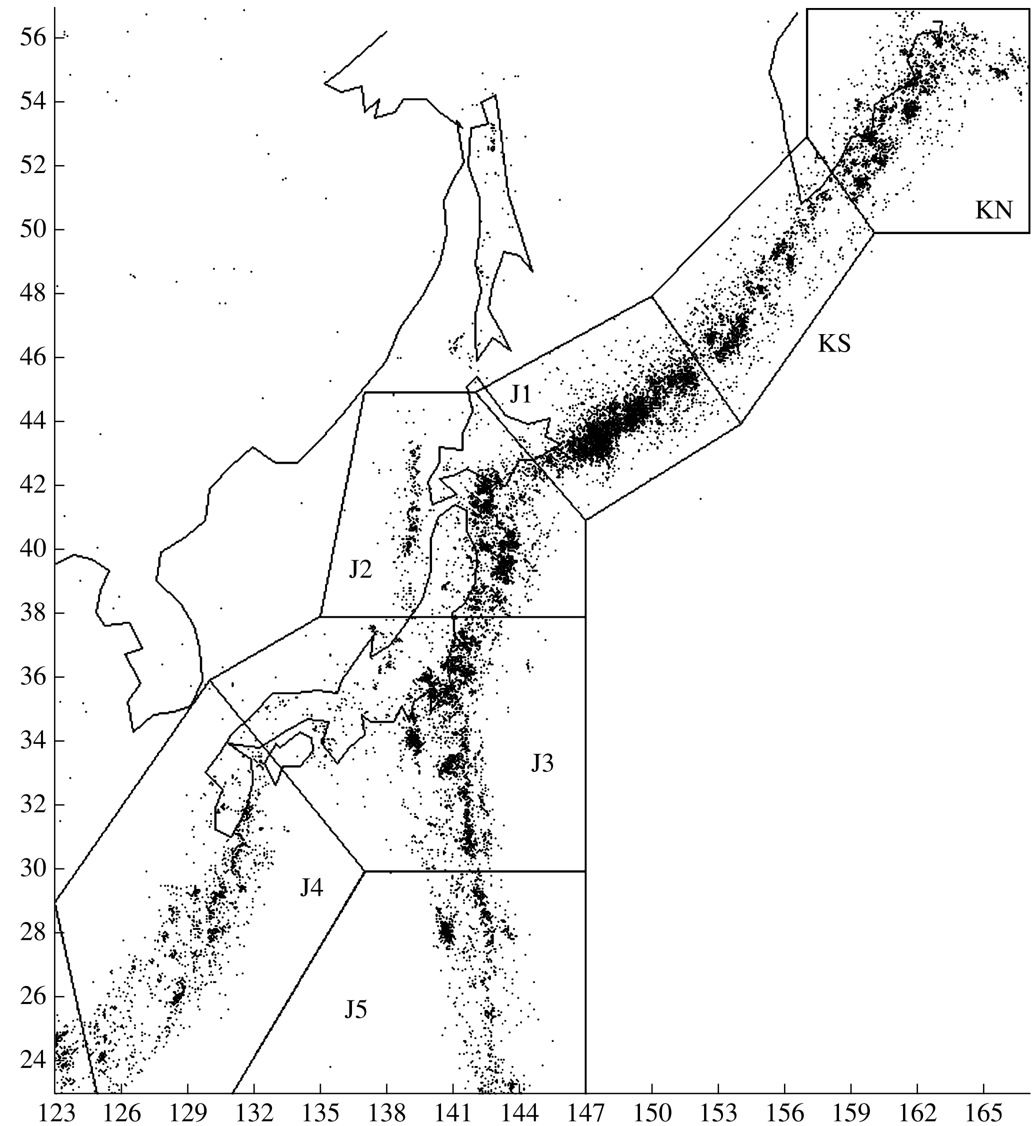

Fig. 3. Hypocenters of $M \geq 4.5$ earthquakes that occurred at depths of ≤100 km from 1963 through 2000 in the region of Japan, Kuril Islands, and Kamchatka, subdivided into seven areas.

should always be used in the analysis of time series of low-frequency geophysical monitoring. The problem does not reduce solely to the fact that the robust variant of the method requires much more computation time. The excessive sensitivity of the method can lead to the extraction of not only strong synchronization effects (preceding strong earthquakes which are of the most interest) but also weaker common variations that can raise the noise background level. Comparison of Figs. 2a and 2b confirms this supposition, although the response of the robust signal to the Haicheng earthquake is undoubtedly a strong argument in favor of the application of the robust scheme of time series aggregation.

2. Analysis of increments in the Benioff curves in the regions of Japan, Kuril Islands, and Kamchatka. The method of joint analysis of seismic regimes in a group of areas composing a large seismically active region was proposed in [Lyubushin, 2000b] for the purpose of detecting collective behavior phenomena in seismicity. The method is based on the multidimensional wavelet-analysis of cumulated values of the square roots of energies released by earthquakes in each of the areas (the so-called cumulative Benioff curves which are proportional to the magnitude of stresses accumulated and released in the process of an earthquake sequence). Thus, the problem of the joint analysis of several point processes was reduced to a simpler problem of processing the multidimensional time series of Benioff curves. However, the distribution of wavelet coefficients obtained for these curves is characterized by even larger outliers compared to usual monitoring time series. This circumstance is connected not only with the property of wavelet expansions to collect main

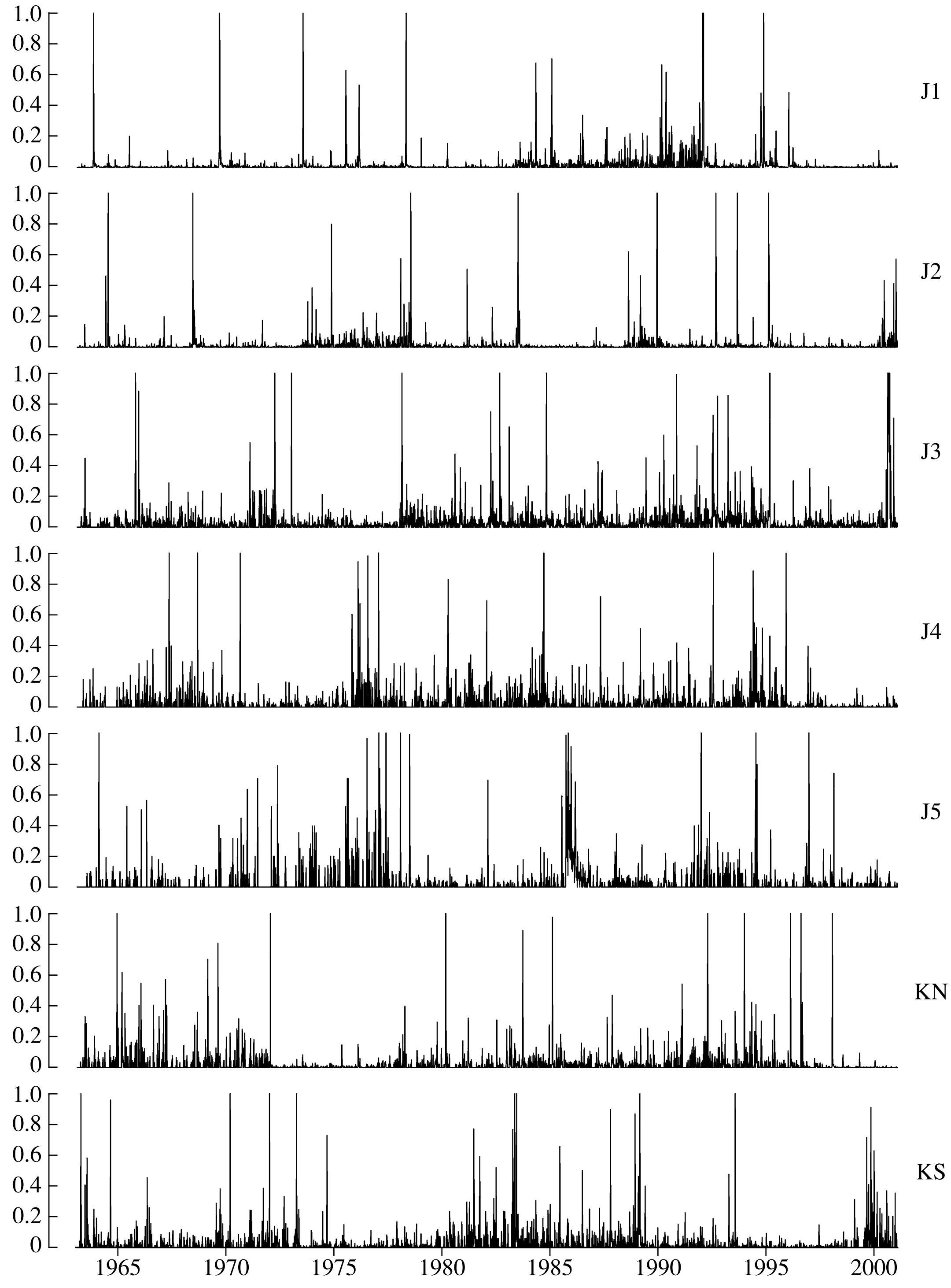

Fig. 4. Time series of increments of Benioff curves in areas J1–J5, KN, and KS (Fig. 3) with a time step of 5 days after application scaling operations (7) in the adaptation window of 365 samples in length (1825 days, or five years).

information by a small amount of the coefficients but also with the essentially non-Gaussian behavior of initial time series that include large jumps associated with strong earthquakes. Therefore, the application of the robust method of aggregating to these data appears to be always preferable to the nonrobust method.

Benioff curves were analyzed in [Lyubushin, 2000b] without the use of increments, which had the purpose of attenuating the influence of catastrophic outliers arising in the analysis of elastic stresses released over small uniform time intervals. The nonrobust analysis of time series including such outliers

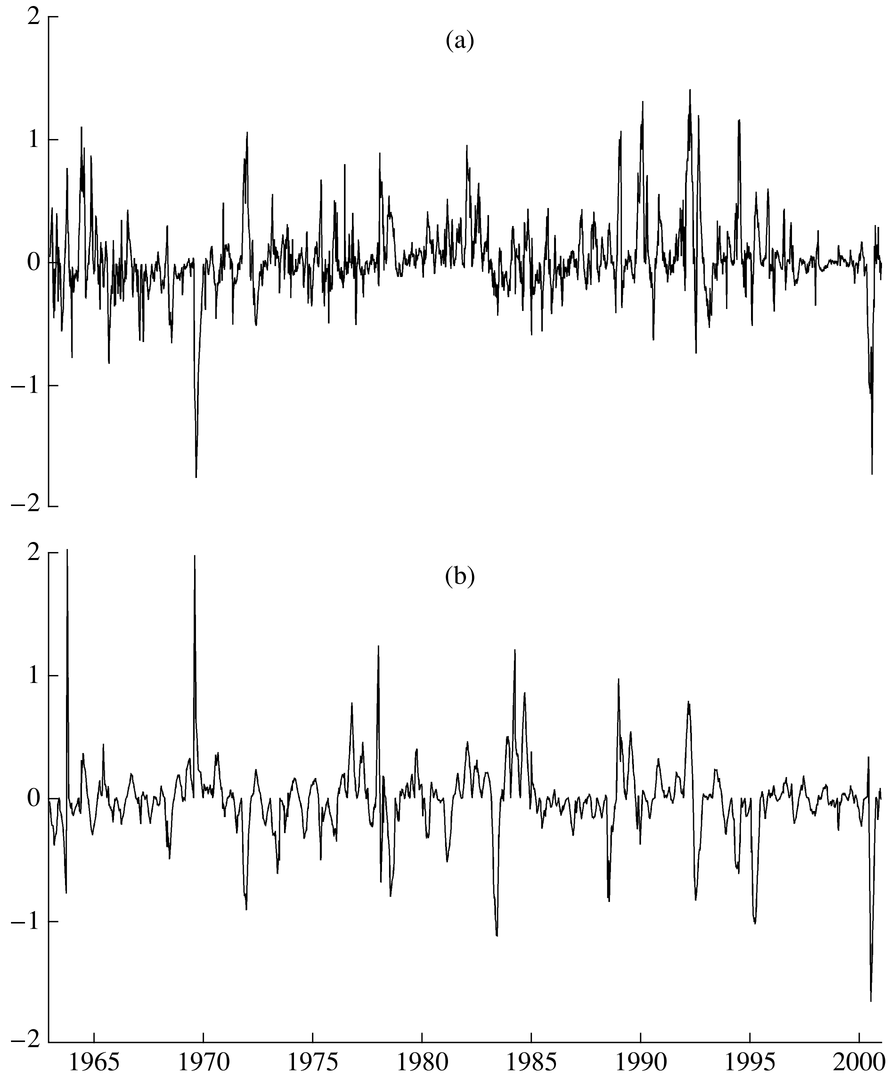

Fig. 5. Nonrobust (a) and robust (b) wavelet-aggregated signals of seven time series plotted in Fig. 4 and estimated with an adaptation window length of 365 (5 years), $L_{\min}$ =10, and Haar wavelets.

seems to be *a priori* problematic, albeit formally possible. This is not the case with the use of the robust aggregation method. For this reason, increments of the curves under study are used here in order to eliminate the dominating influence of low-frequency components.

Figure 3 presents the hypocenter distribution of $M \geq 4.5$ earthquakes that occurred from 1963 through 2000 at depths $\leq$100 km in the region of Japan, Kuril Islands, and Kamchatka (the region is divided into seven areas: J1-J2, KN, and KS). The NEIC global catalog [*Global* …] was used; the lower threshold magnitude used here is representative for this catalog in the time interval studied. The subdivision of the region was made visually in order to delineate large clusters of seismicity. The notation in use is as follows: J, Japan; KN and KS, northern and southern Kamchatka–Kuril areas, respectively.

Figure 4 plots the increments of Benioff curves in these areas with a time step of 5 days after the application of scaling operations (7) in the adaptation window of a length of 365 samples (i.e., 1825 days, or 5 years). The figure shows that the analyzed time series are of the essentially non-Gaussian and nonstationary type. Thus, the problem is to detect time intervals within which these series have a collective component. It is also evident that the most suitable wavelet for the analysis of such data is the Haar one.

Figure 5 plots the nonrobust (Fig. 5a) and robust (Fig. 5b) aggregated signals of the seven time series shown in Fig. 4. Both signals are seen to have much in common (in particular, a substantial increase in synchronization from the middle to the end of 2000), but the robust signal has a much smaller high-frequency noise component, enabling the detection of such features that are nearly unrecognizable against the background of noise fluctuations in Fig. 2a. An average duration of synchronization pulses in Fig. 2b is about half year. Thus, the robust wavelet-aggregated signal is undoubtedly beneficial to the analysis of seismic events.

CONCLUSION

A robust scheme of wavelet aggregation of geophysical monitoring time series is proposed. The method is based on multidimensional analysis of wavelet coefficients (canonical and principal components). The analysis of geophysical data (low-frequency geophysical monitoring time series and increments of Benioff curves for a group of seismically active regions) in order to detect hidden signals of collective behavior was performed, and the comparison with the nonrobust method of wavelet aggregation was conducted. The robust method is shown to have a number of advantages from the standpoint of sensitivity and stability of results.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 99-05-65175, and by the International Association for the Promotion of Cooperation with Scientists from the Independent States of the Former Soviet Union (INTAS), grant no. 99-0099.

REFERENCES

Daubechies, I., *Ten Lectures on Wavelets*, no. 61 in CBMS-NSF Series in Applied Mathematics, Philadelphia: SIAM, 1992.

Global Hypocenters Data Base, http://www-neic.cr.usgs.gov/neis/.

Hotelling, H., Relations between Two Sets of Variates, *Biometrika*, 1936, vol. 28, pp. 321–377.

Huber, P.J., *Robust Statistics*, New York: Wiley, 1981.

Kasahara, K., *Earthquake Mechanics*, Cambridge: Univ. Press, 1981.

Lyubushin, A.A., Jr., An Aggregated Signal of Low-Frequency Geophysical Monitoring Systems, *Izvestiya, Phys. Solid Earth*, 1998, vol. 34, pp. 238–243.

Lyubushin, A.A., Jr., Wavelet-Aggregated Signal in Earthquake Prediction, *Earthquake Res. China* (English Edition), 1999a, no. 1, pp. 33–43.

Lyubushin, A.A., Jr., Analysis of Multidimensional Geophysical Monitoring Time Series for Earthquake Prediction, *Annali Geofis.*, 1999b, vol. 42, no. 5, pp. 927–937.

Lyubushin, A.A., Jr., Wavelet-Aggregated Signal and Synchronous Peaked Fluctuations in Problems of Geophysical Monitoring and Earthquake Prediction, *Izvestiya, Phys. Solid Earth*, 2000a, vol. 36, pp. 204–213.

Lyubushin, A.A., Jr., Multidimensional Wavelet Analysis of Seismicity, *Izvestiya, Phys. Solid Earth*, 2000b, vol. 36, pp.686–696.

Lyubushin, A.A., Jr., Multidimensional Wavelet Analysis of Geophysical Monitoring Time Series, *Izvestiya, Phys. Solid Earth*, 2001, vol. 37, pp. 474–483.

Mallat, S., *A Wavelet Tour of Signal Processing*, San Diego: Academic, 1998.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vettering, W.T., *Numerical Recipes* (Ch. 13, Wavelet Transforms), Cambridge: Univ. Press, 1996.

Rao, C.R., *Linear Statistical Inference and Its Applications*, New York: Wiley, 1965.

Shor, N.Z., *Metody minimizatsii nedifferentsiruemykh funktsii i ikh prilozheniya* (Methods for Minimizing Nondifferentiable Functions and Their Applications), Kiev: Nauk. Dumka, 1979.

Sobolev, G.A., *Osnovy prognoza zemletryasenii* (Fundamentals of Earthquake Prediction), Moscow: Nauka, 1993.