ISSN 1069-3513, Izvestiya, Physics of the Solid Earth, 2011, Vol. 47, No. 6, pp. 488–495. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © A.A. Lyubushin, 2011, published in Fizika Zemli, 2011, No. 6, pp. 26–34.

Cluster Analysis of Low-Frequency Microseismic Noise

A. A. Lyubushin

Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Bol'shaya Gruzinskaya ul. 10, Moscow, 123996 Russia

Received April 26, 2010

Abstract—A method is proposed for describing low-frequency microseismic noise from the network of broadband seismic stations in a large seismically active region of the Japan Islands. The median values of daily estimates from each station for seven parameters (three characteristics of the multifractal singularity spectra of the waveforms, their spectral exponents and the smoothness indices, the logarithmic variance and the linear predictability index) are used for the description. These parameters are determined for consecutive daily time intervals from the beginning of January 1997 through the end of February 2010. Since these parameters are taken as the median values of the estimates from each station, they are, actually, the integral statistics of the microseismic field. The present paper is the continuation of two previous works [Lyubushin, 2009; 2010], where the effects of synchronization in the low-frequency microseismic field on a large time scale were analyzed on the data from the *F*-net stations. In the present work, the number of different "behavior modes" of the microseismic field are sought as the number of clusters in the optimal partition of the cloud of 7-dimensional vector parameters, estimated within a moving time window with a width of 2 years. A new characteristic of the geophysical field is introduced, namely, the notion of the cluster exponent, which is the power exponent in the dependence of the value of the compactness function of a cloud of vector parameters on the number of clusters in the optimal partition of this cloud. Previously, a relatively rapid increase was revealed in the level of synchronization of the microseismic field, which started in the middle of 2002 and lasted for approximately one year. The level of synchronization remains high up to now. During the past 4 years (taking into account the width (2 years) of the time window within which the estimates were made), the cluster exponent exhibits a long trend which is similar to the shorter trend before the Hokkaido event ($M = 8.3$) that occurred on September 25, 2003. These facts, together with the pattern of variations in the coefficient of correlation between two multifractal parameters of the field, suggest a hypothesis of the enhancement of the seismic hazard in the region of the Islands of Japan from the second half of 2010.

DOI: 10.1134/S1069351311040057

INTRODUCTION

The low-frequency microseismic oscillations and their correlation with the processes occurring in the hydrosphere and atmosphere of the Earth, which are the major sources of microseismic energy, are a common subject of research in geophysics [Friedrich et al., 1998; Kobayashi and Nishida, 1998; Tanimoto et al., 1998; Tanimoto and Um, 1999; Ekstrom, 2001; Tanimoto, 2001; 2005; Kurrle and Widmer-Schnidrig, 2006; Stehly et al., 2006; Rhie and Romanowicz, 2004; 2006].

It is however evident that the variations in the structure of the microseismic background may also reflect the changes in the properties of the Earth's crust, which is the medium where the microseismic signals propagate. These changes in the parameters of microseisms, especially in relation to a series of strong earthquakes, are the subject of the present paper. The periodic structure of the point process formed by strong bursts and low-frequency asymmetric pulses of microseismic noise before strong earthquakes was studied in [Sobolev et al., 2005; 2008; Sobolev and Lyubushin, 2006]. The authors of [Lyubushin and Sobolev, 2006; Lyubushin, 2007; 2008; 2009; 2010] analyzed the effects of synchronization in the variations of the parameters of the multifractal singularity spectra [Feder, 1988; Kantelhardt et al., 2002] of the waveforms assessed within the moving time windows of a small width. In these works, the parameters were estimated from both the records at the individual stations and as the median values of estimates from spatial clusters of stations [Lyubushin, 2009; 2010], which allowed the authors to use long records of the seismic noise with numerous gaps in signal recording.

The 1-Hz-sampled records of the vertical microseismic components from the *F*-net database over the period from the beginning of 1997 through June 2008 were analyzed in [Lyubushin, 2009]. A statistically significant decrease half a year before the Hokkaido earthquake of September 25, 2003 was revealed in the time-average median value (over the network stations) of the support width of the singularity spectrum $\Delta\alpha$ of the microseismic waveforms. We note that after this event the mean value $\Delta\alpha$ did not relax to the previous level and settled at low values. It was shown that this change was accompanied by the increasing synchronization between the median values of the parameters of the singularity spectra calculated from five groups of stations. The effect of synchronization was estimated in consecutive non-overlapping time windows 2 months wide.

Drawing parallels with the dynamics of the coupled nonlinear oscillators, we can state that the narrowing of the width of the support of the singularity spectrum (which formally means a reduced variety of the patterns of stochastic behavior) is directly related to the synchronization of the elements of a complex system [Pavlov et al., 2003].

In [Lyubushin, 2010], the problem was discussed of identification of the synchronization effects in the parameters of low-frequency microseismic noise using the data from the *F*-net network for the time period from early 1997 to August 2009. In that work, the scope of the parameters to be analyzed was considerably broader. In addition to the earlier considered support width $\Delta\alpha$ of the multifractal singularity spectrum and the generalized Hurst parameter α^*, which is the argument yielding the maximum of the singularity spectrum, in that work, the asymmetry coefficient of the singularity spectrum, the logarithmic variance, the spectral exponent, and the linear predictability index were invoked as well. These parameters are calculated over the realizations within consecutive daily time intervals. When a moving annual time window is used for estimation of the measure of multiple correlation, the diurnal variations of the median values of the statistics of the microseismic background from five spaced clusters of stations exhibit a steady increase in the synchronization shortly before the Hokkaido earthquake on September 25, 2003, which is followed by the ascent into a new level of high synchronization. Based on the analysis of the trends in the linear predictability index, the onset of increasing synchronization was rather accurately estimated to occur in the middle of 2002. The effect, revealed in the variations of different parameters of microseisms, is an independent argument to support the previous inferences concerning the synchronization of the background microseismic field on the Islands of Japan. It is worth noting that the use of the annual time window for calculating the correlation measure is equivalent to averaging the seasonal effects of the cyclones and hurricanes, which are the main sources of generation of low-frequency microseisms; due to this, the estimate of the multiple correlations becomes very stable.

However, the estimates of the multiple measures of coherence and correlation within moving time windows, applied in [Lyubushin, 2009; 2010], yielded no other significant anomalies in the parameters of the microseismic field over the observation period of 1997–2010, but the relatively fast attainment of a high synchronization level in the second half of 2002 and early in 2003. The present work addresses a different method of the data analysis, namely, the cluster analysis of the cloud of vector parameters of the microseismic noise within the biannual moving window. The goal of this analysis is to answer the question on how many "behavior modes" of the microseismic field can be revealed and how the number of these modes varies with time. Here, the behavior mode is understood as a cluster (compact group) of the vector parameters in the current long biannual window. In the context of this approach, we can introduce a new parameter, namely, the cluster exponent μ, and to identify the anomalous region characterized by the positive trend of μ, which is similar to the shorter trend before the Hokkaido event of September 25, 2003. In this approach, the network stations are not divided into a given number of spatial groups for the subsequent calculation of the measures of coherence or correlation between the variations of different microseismic parameters from each group of stations in the moving time windows, as it was done in [Lyubushin, 2009; 2010]. Thus, the median values of the parameters of microseismic signals are taken over all network stations, which render the used parameters integral characteristics of the microseismic field in the region covered by the observations.

THE STATISTICS

We used the broadband seismic data from the *F*-net network [http://www.hinet.bosai.go.jp/fnet] for the period from the beginning of 1997 through the end of February 2010 inclusive (more than 13 years). The layout of 83 network stations can be accessed via the cited link and is presented in the works [Lyubushin, 2009; 2010] as well. The input for the analysis was the vertical components sampled at 1 s. These seismic records were decimated to 1 min by calculating the mean values over 60 successive samples.

Below, the seven statistics used in the cluster analysis are briefly described. The detailed description of these is given in [Lyubushin, 2010], therefore, in what follows, many technical details of their calculation are omitted.

Spectral exponent β, smoothness index ξ of a waveform, and logarithmic variance $\log(Var)$. The spectral exponent β defines the type of variations in the logarithmic power spectrum dependent on the logarithmic period; its value is closely related to the fractal characteristics of noise [Feder, 1988; Kantelhardt et al., 2002]. Below, instead of using the classical power spectrum estimates based on the Fourier expansion or the parametric models, we shall use its estimation based on the rate of change of the mean squared modules of the wavelet coefficients W_k [Mallat, 1997] depending on the number of the level of detail $k = 1, \ldots, m$ as described by the formula $\log_2(W_k) \sim k^{\beta}$. This automatically ensures strong averaging of the spectrum, which is necessary for estimation of the spectral exponent. Parameter β was estimated in successive time windows with a length of 1 day (1440 minute samples). This mode of calculation of the spectral exponents was applied in the analysis of the microseismic noise in [Lyubushin, 2008]. In order to eliminate the effect of tidal variations, an 8th order polynomial trend was fitted within each window, and the wavelet power spectrum W_k and the common decimal logarithm of variance $\log(Var)$. were calculated for the residual. Here,

the optimal orthogonal Daubechies' wavelet with the number of vanishing moments varying from 1 to 10 (one vanishing moment corresponds to the Haar wavelet) was used, which provides the minimum entropy of the distribution of the squared wavelet coefficients for the first seven levels of detail in the wavelet expansion (the scales or the periods span from 2 to 256 minutes for a 1 minute sampling step). This technique of calculating the spectral exponent automatically gives another useful characteristic, namely the number ξ of vanishing moments of the optimal wavelet, which can assume integer values from 1 through 10. The larger values of ξ correspond to the more smooth waveform within the current daily interval. The median values over all F-net network stations that were in operation during the day considered are denoted by $\overline{\beta}(s)$, $\overline{\xi}(s)$, and $\overline{\log(Var)}(s)$. In these notations, the horizontal line above the characters designates spatial averaging (taking the median value), and the argument s is the integer-valued index, which numbers the consecutive days from the beginning of measurements (January 1, 1997). Note that after calculation of the median value over all stations, the smoothness index $\overline{\xi}(s)$ is no longer integer.

Index of linear predictability ρ is calculated as $\rho = V_0/V_{AR} - 1$. Here, V_0 is the variance of the error $\varepsilon_0(t+1)$ of the one-step-ahead trivial prediction $\hat{x}_0(t+1)$ for the increments $x(t)$ of seismic records, which is equal to the mean value taken over the preceding "short" time window with a length of n samples: $\hat{x}_0(t+1) = \sum_{s=t-n+1}^{t} x(t)/n$. Thus, $\varepsilon_0(t+1) = x(t+1) - \hat{x}_0(t+1)$, while $V_0 = \sum_{t=n+1}^{N} \varepsilon_0^2(t)/(N-n)$, where $N > n$ is the number of readings in the consecutive "long" time windows. The quantity V_{AR} is calculated in a similar way: $V_{AR} = \sum_{t=n+1}^{N} \varepsilon_{AR}^2(t)/(N-n)$, where $\varepsilon_{AR}(t+1) = x(t+1) - \hat{x}_{AR}(t+1)$ is the error of the one-step-ahead linear prediction $\hat{x}_{AR}(t+1)$ in the second order autoregressive model (AR-prediction), whose coefficients are also estimated in the preceding "short" time window containing n samples. The second-order autoregression is selected because this is the minimal order for the AR-model, which enables one to describe the oscillatory motion and provides the maximum spectral density between the Nyquist frequency and zero [Box and Jenkins, 1970; Kashyap and Rao, 1976]. The transition to increments is dictated by the necessity to avoid dominance of low frequencies associated with tides and other trends. The AR-prediction makes use of the correlation property of the nearby values of increments of the recorded signals and, if such a correlation exists, $V_{AR} < V_0$ and $\rho > 0$. In the calculations of the linear predictability index ρ for one-minute data, the estimations were conducted in consecutive long time windows of length $N = 1440$ counts (1 day) and within the short window of length $n = 60$ readings (1 hour).

Parameters α*, Δα, and g of the multifractal singularity spectrum. Let $X(t)$ be a random process. We define the measure $\theta(t, \delta)$ of the signal $X(t)$ behavior in the interval $[t, t+\delta]$ as the span (peak-to-peak amplitude) $\theta(t, \delta) = \max_{t \le \lambda \le t+\delta} X(\lambda) - \min_{t \le \lambda \le t+\delta} X(\lambda)$. The singularity spectrum $F(\alpha)$ can be determined as the fractal dimension of the time moments λ_α, that have the same value of the Holder–Lipschitz local exponent $h(t) = \lim_{\delta \to 0} \frac{\ln(\theta(t, \delta))}{\ln(\delta)}$, that is, $h(\lambda_\alpha) = \alpha$. In the present work, the singularity spectrum in a moving time window was estimated using the version of detrended fluctuation analysis (DFA) [Kantelhardt et al., 2002], which is described in detail in [Lyubushin and Sobolev, 2006; Lyubushin, 2007; 2008; 2009, 2010]. Below, in the analysis of low-frequency microseisms, the estimates of the singularity spectrum in successive nonoverlapping daily time windows were used; the local scale-dependent trends were eliminated by the 8th-order polynomial fitting. Major attention will be attached to the study of variations in two parameters of the singularity spectrum: the generalized Hurst exponent α^* and the width of support of the singularity spectrum $\Delta\alpha$. The parameter α^* characterizes the most typical and most common Holder–Lipschitz exponent, while $\Delta\alpha$ reflects the diversity of the random behavior of a signal and is some kind of a measure of the hidden degrees of freedom of a stochastic system. Besides these two parameters, also the parameter $\gamma = \alpha^* - (\alpha_{min} + \alpha_{max})/2$, will be analyzed, which characterizes the asymmetry of the singularity spectrum.

Similarly to the notations $\overline{\beta}(s)$, $\overline{\xi}(s)$ and $\overline{\log(Var)}(s)$, introduced above, we denote the median values of the corresponding statistics by $\overline{\rho}(s)$, $\overline{\alpha}^*(s)$, $\overline{\Delta\alpha}(s)$, and $\overline{\gamma}(s)$, where the horizontal line above the characters means taking the median value over the data from all network stations (which is equivalent to spatial averaging); the argument s numbers the successive days.

The graphs in Fig. 1 depict the variations in the median values of all seven statistics together with their mean values within a moving window 57 days wide. Such a width of the time window is chosen since it is double the length of the lunar month (28 days is a common modulation period of many geophysical processes), and the additional unity is added to ensure odd value of the width of the moving window of averaging. Figure 2 presents the graphs of robust coefficient of correlation [Lyubushin, 2007; 2010] between the parameters $\overline{\alpha}^*(s)$ and $\overline{\Delta\alpha}(s)$, calculated within a moving annual time window. It is worth noting that all multiple correlation coefficients for the median values of all parameters exhibit a common increase before the event of September 25, 2003, similar to the effects of increas-

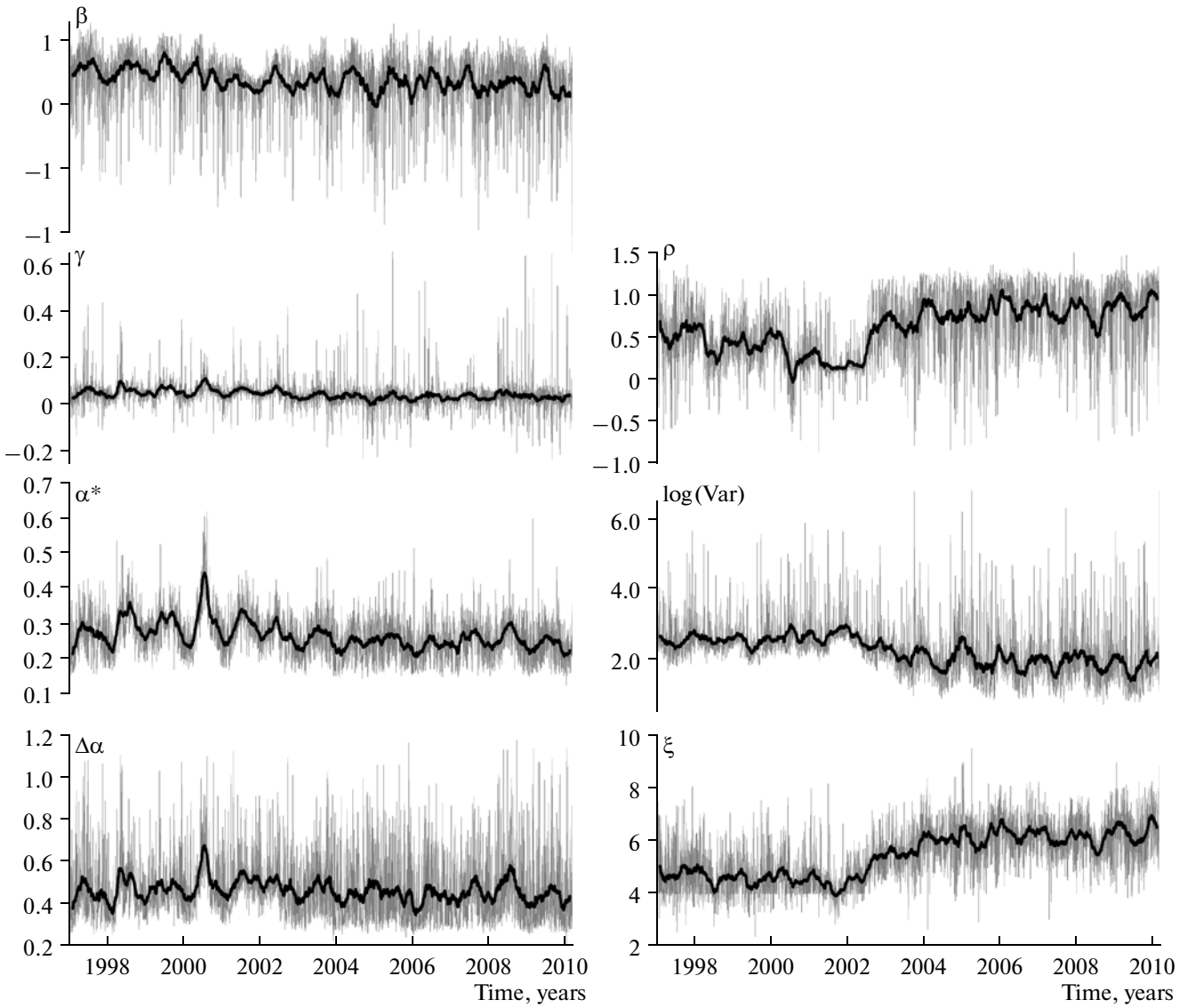

Fig. 1. The dependences of the statistics calculated as median values over all F-net seismic network stations in consecutive time windows 1 day wide: β is the spectral exponent; γ is the asymmetry coefficient of the singularity spectrum; α* is the generalized Hurst exponent; Δα is the width of support of the singularity spectrum; ρ is the linear predictability index; log(*Var*) is the decimal logarithm of variance; ξ is the smoothness index of waveforms. The values of β, log(*Var*) , and ξ are calculated after the removal of the trend by the 8th-order polynomial fitting in each window 1 day long. Solid black lines are the graphs of moving means in a window 57 days long.

ing synchronization between the variations in the median values of the same parameter calculated for the different groups of stations [Lyubushin, 2010]. However, a remarkable feature of Fig. 2 is the fact that it contains two pronounced anomalies in the correlation coefficient, namely the deep minima in 2002 and 2009. Because the first anomaly of 2002 was followed by the large earthquake of September 25, 2009, we may quite naturally suppose that the second deep minimum in the correlation coefficient also might be a precursor of an impending large seismic event in the second half of 2010, and the latter event might be stronger in energy. The latter argument in favor of the expected larger seismic event is supported by the increasing synchronization in the microseismic field starting from the second half of 2002, which has been earlier established in [Lyubushin, 2009; 2010].

Next, for each moving time window with a width of 730 days (2 years), the following sequence of operations over the cloud of 7-dimensional vectors $\boldsymbol{\psi}$ with components $\overline{\beta}$, $\overline{\xi}$, $\overline{\log(Var)}$, $\overline{\rho}$, $\overline{\alpha}^*$, $\overline{\Delta\alpha}$, and $\overline{\gamma}$ was carried out within the window:

(1) Each component of vector $\boldsymbol{\psi}$ has been normalized and subjected to winsorization: first, the sample mean values and the standard deviations were calculated; then, the sample means were subtracted, and the

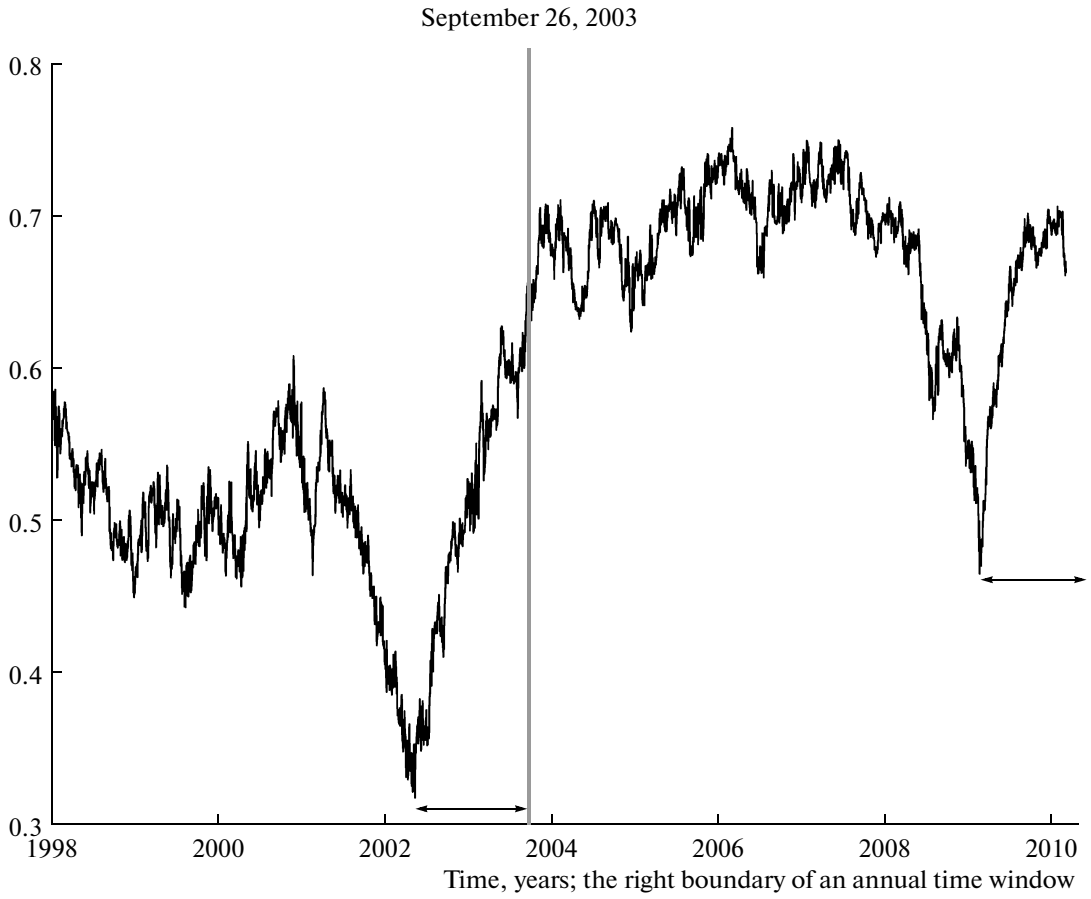

Fig. 2. The graph of the robust coefficient of correlation between the variations in α^* and $\Delta\alpha$ (see Fig. 1) in a moving window with a width of 365 days (one year) as a function of the position of the right boundary of the time window. The time of the Hokkaido earthquake of September 25, 2009 ($M = 8.3$) is shown by the vertical line; the bidirectional arrows indicate equal time intervals from the first minimum in the correlation coefficient until the occurrence time of the September 25,2003 earthquake and from the second minimum until June, 2010.

values falling beyond $\pm 4\sigma$, were cut, after which the obtained values were divided by σ; this process was repeated iteratively until the value of σ stopped to change.

(2) For the obtained cloud of the normalized 7-dimensional vectors, the first four principal components were calculated as the projections on the eigenvectors of the covariance matrix within the current window, which correspond to the four maximal eigenvalues (this ensured the additional noise suppression and conserved 91 to 95% of the common variance).

(3) For the obtained cloud of 4-dimensional principal vector components, partition into a given number q of clusters Γ_k, $k = 1, \ldots, q$ was carried out. The trial number of clusters successively varied from 40 to 2. The partition was carried out by the hierarchical farthest neighbor clustering (which provides compact round clusters) and the following iteration of the K-mean method [Aivazyan et al., 1989; Duda and Hart, 1973].

Let N be the total number of 4-dimensional vectors ζ of the principal components (of the normalized 7-dimentional vectors) in the current time window, and ζ_0 be the common center-of-mass vector for the cloud of the principal components (due to the preliminary normalization and winsorization, $\zeta_0 = 0$). Let $\bar{\zeta}_k$, $k = 1, \ldots, q$ be the center-of-mass vectors of clusters, and n_k be the number of elements in each cluster, $\sum_{k=1}^{q} n_k = N$. Partition of a cloud of N vectors into a given number q of clusters is estimated by the following quantities:

$$\sigma_0^2(q) = \frac{\sum_{k=1}^{q} \sum_{\zeta \in \Gamma_k} |\zeta - \zeta_k|^2}{N - q} \qquad (1)$$

is the measure of intracluster compactness;

$$\sigma_1^2(q) = \frac{\sum_{k=1}^{q} \nu_k |\zeta_k - \zeta_0|^2}{q - 1}, \quad \nu_k = \frac{n_k}{N} \qquad (2)$$

is the weighted measure of divergence between the centers of clusters;

$$PFS(q) = \frac{\sigma_1^2(q)}{\sigma_0^2(q)} \quad (3)$$

is the so-called pseudo-F-statistics [Vogel and Wong, 1978].

The aim of partitioning a cloud into q clusters is to minimize the quantity $\sigma_0^2(q)$. Formally, the value of $\sigma_0^2(q)$ can be defined for $q = 1$ as well: $\sigma_0^2(1) = \sum_{\bar{\xi}} |\zeta - \zeta_0|^2/(N-1)$. The value $\sigma_0^2(q)$ monotonically increases as q decreases, and the dependence of $\log(\sigma_0^2(q))$ on $\log(q)$ is close to linear, that is, $\sigma_0^2(q)$ scales as $q^{-\mu}$. This is illustrated in Fig.3. The quantity μ will be referred to as the cluster exponent. Its value for the given window can be estimated as the slope of a straight line least-squares-fitted to the dependence of $\log(\sigma_0^2(q))$ on $\log(q)$.

Parameter (3) characterizes the quality of the partition into a given number of clusters: larger values of $PFS(q)$, mean better partition. To provide a good partition, it is desirable that the value of $\sigma_0^2(q)$ of the intracluster compactness (the intracluster distance) is small while the intercluster distance $\sigma_1^2(q)$ is large. The optimal number of clusters q^* is found from the maximum condition for $PFS(q)$. At the same time, the value of $PFS(q)$ (or, to put it more exactly, $\sigma_1^2(q)$) cannot be calculated for $q = 1$. Therefore, one should invoke other considerations in order to distinguish between $q = 1$ and $q = 2$. 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 $q < q^*$ increases faster at $\sigma_0^2(q)$ than at $q > q^*$. This criterion of identification of $q = q^*$ is more susceptible to noise and exhibits a poorer performance compared to the technique $q^* = \arg\max_{2 \le q} PFS(q)$, but this is the only possibility to discern the case $q = 1$ from the case $q = 2$. Let $\delta\log(\sigma_0^2(q))$ denote the deviation of $\log(\sigma_0^2(q))$ from the best-fit straight line approximating the dependence on $\log(q)$. Then, we assume that the point $q = 2$ is the break point of the dependency $\sigma_0^2(q)$, if $\delta\log(\sigma_0^2(1))$ exceeds all values of $\delta\log(\sigma_0^2(q))$ for $q \ge 2$. Thus, we define the optimal number q^* of clusters according to the following rule:

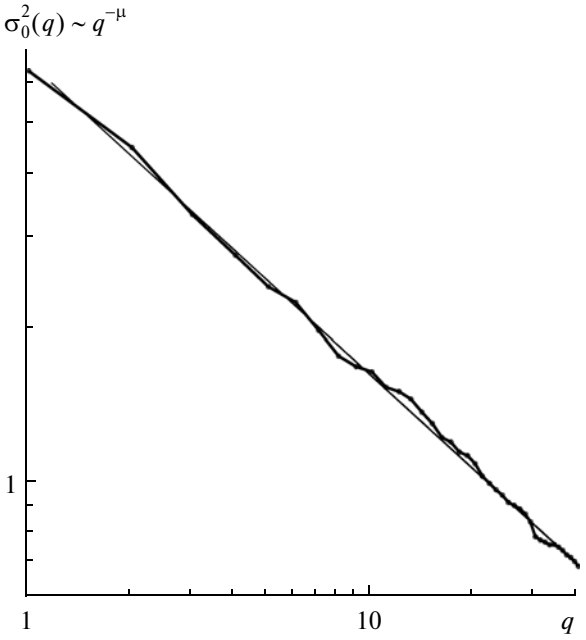

Fig. 3. The dependence of the intra-cluster compactness functional $\sigma_0^2(q)$ on the trial number of clusters q. The graph of the best-fit linear approximation to the dependence of $\log(\sigma_0^2(q))$ on $\log(q)$ is shown by a thin straight line.

$$\begin{gathered}\text{Let } q_0 = \arg\max_{2 \le q \le 40} PFS(q),\\ \text{if } q_0 > 2, \text{ then } q^* = q_0,\\ \text{else:}\\ \text{if } \delta\log(\sigma_0^2(1)) \le \max_{2 \le q \le 40} \delta\log(\sigma_0^2(q)), \text{ then } q^* = 1,\\ \text{else } q^* = 2.\end{gathered} \quad (4)$$

The results of clustering of four principal components of seven daily median characteristics of the microseismic noise for 13 years of the F-net network observations (1997–2009) with estimation in a biannual moving time window ($N = 730$) with a shift 7 days long are presented in Fig. 4.

DISCUSSION

A remarkable feature in Fig. 4 is the changed pattern of switching between the numbers of optimal clusters after 2004: the mode of switching became more chaotic, and the three-cluster structure, which prevailed before, completely disappeared. This fact can be regarded as a certain internal degree of freedom of the microseismic field having been frozen out after the event of September 25, 2003 before the possible impending strong earthquake. The cluster exponent μ has a well-expressed linear trend followed by attaining the maximum μ and subsequent reversal to a new average level before the seismic event of September 25, 2003. This feature considerably differs from the previous back-

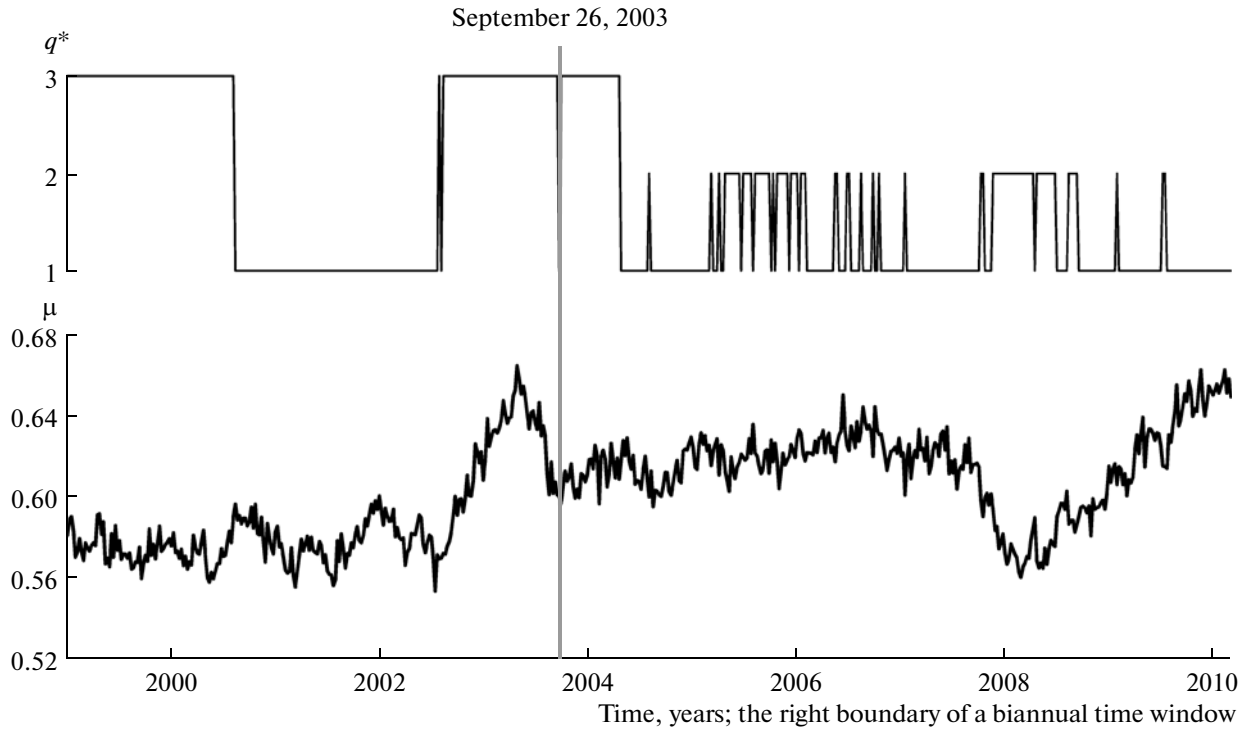

Fig. 4. The results of cluster analysis of four principal components in a biannual moving time window with a shift 7 days long as a function of the position of the right (later) boundary of the time window: q^* is the optimal number of clusters; μ is the value of cluster exponent. The time of the Hokkaido earthquake of September 25, 2009 ($M = 8.3$) is shown by a vertical line.

ground of the stochastic fluctuations about the mean value. The initial stage of this anomaly, which is the linear trend, repeats starting from the moment when the right (later) boundary of the biannual window occurs at the beginning of 2008; here, the positive trend in μ has a longer duration and has continued up to the present day.

The results are obtained after application of the following sequence of procedures: (a) estimation of 7-dimensional vector parameters ψ of the microseismic noise; (b) transition to a 4-dimensional vector ζ of their first principal components; (c) preliminary normalization and winsorization of the vector components of ζ independently in each biannual time window; and (d) estimation of the optimal number of clusters and the cluster exponent. The idea behind application of this sequence of procedures is to obtain some parameters of the field, which are averaged over space (by taking the median values for the network stations) and time (primarily by estimation within windows 1 day long and then by analyzing the clustering properties of the cloud of vector parameters within a biannual window). Thus, the analysis is mainly targeted on informative purposes, and this creates a difficulty when one attempts to gain an insight into the physical implications of the obtained results.

The sense of some individual vector components ψ is rather clear, though (in the further discussion, the average values are meant, which is indicated by a horizontal line above the corresponding character): $\overline{\log(Var)}$ is the log energy of noise; $\overline{\beta}$ is the spectral exponent (the slope of the log-log graph of a power spectrum); $\overline{\alpha}^*$, $\overline{\Delta\alpha}$ and $\overline{\gamma}$ are parameters of multifractal singularity spectrum, which have been long used in the analysis of physical systems with random behavior due to their inherent nonlinear dynamics (for example, in the studies of turbulence). The parameters $\overline{\xi}$ (mean value of the smoothness index of the noise waveforms) and $\overline{\rho}$ (mean value of the linear predictability index of noise) introduced in [Lyubushin, 2010] describe the geometrical and dynamical properties of the random microseismic oscillations and can be regarded as measures of their inverse complexity: the smaller the values of $\overline{\xi}$ and $\overline{\rho}$, the more complex the noise structure. We note that the parameter $\overline{\Delta\alpha}$ also is a measure of complexity, but a direct one: the larger $\overline{\Delta\alpha}$, correspond to a more complex signal (recall that for a monofractal signal $\overline{\Delta\alpha} = 0$).

Thus, each graph in Fig. 1 can be given one or another physical interpretation. We note that the rapid change in the time-average $\overline{\rho}$ in Fig. 1 allowed us to determine the onset of synchronization at middle 2002 [Lyubushin, 2010]. The subsequent procedures of normalization, transition to the principal components and cluster analysis perform some kind of mixing of the

original properties of the vector components of ψ. This results, for example, in the fact that the cluster exponent μ is rather abstract; it is barely possible to associate this parameter with any primary physical law or any equation set. However, this commonly happens when one seeks an adequate description to a complex multicomponent system such as the Earth's crust. The physical interpretation of the behavior of μ in Fig. 4 or the coefficient of correlation between $\bar{\alpha}^*$ and $\overline{\Delta\alpha}$ in Fig. 2 remains an open question.

CONCLUSIONS

A method is proposed for the analysis of microseismic noise by studying the clustering of its parameters, averaged over the network stations in a moving time window. As applied to the analysis of the field of microseismic oscillations on the Japan Islands, this technique revealed an anomaly before the Hokkaido earthquake of September 25, 2003, as well as a recent anomaly that started in 2007–2008. Supposedly, the latter feature is a sign of a noticeable enhancement of seismic hazard in Japan starting from the second half of 2010.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (grant no. 09-05-00134).

REFERENCES

Aivazyan, S.A., Bukhshtaber, V.M., Enyukov, I.S., and Meshalkin, L.D., *Prikladnaya statistika. Klassifikatsiya i snizhenie razmernosti* (Applied Statistics: Classification and Dimension Reduction), Moscow: Finansy i Statistika, 1989.

Box, G.E.R. and Jenkins, G.M., *Time Series Analysis: Forecasting and Control,* San Francisco, CA: Holden-Day, 1970.

Duda, R.O. and Hart, R.E., *Pattern Classification and Scene Analysis,* New York: Wiley, 1973.

Ekstrom, G., Time Domain Analysis of Earth'S Long-Period Background Seismic Radiation, *J. Geophys. Res.,* 2001, vol. 106, no. B11, pp. 26 483–26 493.

Feder, J., *Fractals*, New York: Plenum, 1988.

Friedrich, A., Kruger, F., and Klinge, K., Ocean-Generated Microseismic Noise Located with the Gräfenberg Array, *J. Seismol.,* 1998, vol. 2, no. 1, pp. 47–64.

Kantelhardt, J.W., Zschiegner, S.A., Konscienly-Bunde, E., et al., Multifractal Detrended Fluctuation Analysis of Nonstationary Time Series, *Phys. A (Amsterdam, Neth.),* 2002, vol. 316, pp. 87–114.

Kashyap, R.L. and Rao, A.R., *Dynamic Stochastic Models from Empirical Data,* New York: Academic Press, 1976.

Kobayashi, N. and Nishida, K., Continuous Excitation of Planetary Free Oscillations by Atmospheric Disturbances, *Nature,* 1998, vol. 395, pp. 357–360.

Kurrle, D. and Widmer-Schnidrig, R., Spatiotemporal Features of the Earth's Background Oscillations Observed in Central Europe, *Geophys. Rev. Lett.,* 2006, vol. 33, p. L24304.

Lyubushin, A.A. and Sobolev, G.A., Multifractal Measures of Synchronization of Microseismic Oscillations in a Minute Range of Periods, *Fiz. Zemli,* 2006, no. 9, pp. 18–28 [*Izv. Phys. Earth* (Engl. Transl.), 2006, vol. 42, no. 9, pp. 734–744].

Lyubushin, A.A., *Analiz dannykh sistem geofizicheskogo i ekologicheskogo monitoringa* (Geophysical and Ecological Monitoring Data Analysis), Moscow: Nauka, 2007.

Lyubushin, A.A., Microseismic Noise in the Low Frequency Range (Periods of 1–300 min): Properties and Possible Prognostic Features, *Fiz. Zemli,* 2008, no. 4, pp. 17–34 [*Izv. Phys. Earth* (Engl. Transl.), 2008, vol. 44, no. 4, pp. 275–290].

Lyubushin, A.A., Synchronization Trends and Rhythms of Multifractal Parameters of the Field of Low-Frequency Microseisms, *Fiz. Zemli,* 2009, no. 5, pp. 15–28 [*Izv. Phys. Earth* (Engl. Transl.), 2009, vol. 45, no. 5, pp. 381–394].

Lyubushin, A.A., The Statistics of the Time Segments of Low-Frequency Microseisms: Trends and Synchronization, *Fiz. Zemli,* 2010, no. 6, pp. 86–96 [*Izv. Phys. Earth* (Engl. Transl.), 2010, vol. 46, no. 6, pp. 544–554].

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

Pavlov, A.N., Sosnovtseva, O.V., and Mosekilde, E., Scaling Features of Multimode Motions in Coupled Chaotic Oscillations, *Chaos, Solitons Fractals,* 2003, vol. 16, pp. 801–810.

Rhie, J. and Romanowicz, B., Excitation of Earth's Continuous Free Oscillations by Atmosphere-Ocean-Seafloor Coupling, *Nature,* 2004, vol. 431, pp. 552–556.

Sobolev, G.A. and Lyubushin, A.A., Microseismic Impulses as Earthquake Precursors, *Fiz. Zemli,* 2006, no. 9, pp. 5–17 [*Izv. Phys. Earth* (Engl. Transl.), 2006, vol. 42, no. 9, pp. 721–733].

Sobolev, G.A., Lyubushin, A.A., and Zakrzhevskaya, N.A., Asymmetrical Pulses, the Periodicity and Synchronization of Low Frequency Microseisms, *Vulkanol. Seismol.,* 2008, no. 2, pp. 135–152 [*J. Volcanol. Seismol.* (Engl. Transl.), 2008, vol. 2, no. 2, pp. 118–134].

Sobolev, G.A., Lyubushin, A.A., and Zakrzhevskaya, N.A., Synchronization of Microseismic Variations within a Minute Range of Periods, *Fiz. Zemli,* 2005, no. 8, pp. 3–27 [*Izv. Phys. Earth* (Engl. Transl.), 2005, vol. 41, no. 8, pp. 599–621].

Stehly, L., Campillo, M., and Shapiro, N.M., A Study of the Seismic Noise from Its Long-Range Correlation Properties, *J. Geophys. Res.,* 2006, vol. 111, p. B10306.

Tanimoto, T. and Um, J., Cause of Continuous Oscillations of the Earth, *J. Geophys. Res.,* 1999, vol. 104, no. B12, p. 28 723–28 739.

Tanimoto, T., Continuous Free Oscillations: Atmosphere–Solid Earth Coupling, *Annu. Rev. Earth Planet. Sci,* 2001, vol. 29, pp. 563–584.

Tanimoto, T., The Oceanic Excitation Hypothesis for the Continuous Oscillations of the Earth, *Geophys. J. Int.,* 2005, vol. 160, p. 276–288.

Tanimoto, T., Um, J., Nishida, K., and Kobayashi, N., Earth'S Continuous Oscillations Observed on Seismically Quiet Days, *Geophys. Rev. Lett.,* 1998, vol. 25, pp. 1553–1556.

Vogel, M.A. and Wong, A.K.C., PFS Clustering Method, *IEEE Trans. Pattern Analysis. Mach. Intell.,* 1979, vol. PAMI-1, no. 3, pp. 237–245.