

УДК 550.334

## КЛАСТЕРНЫЙ АНАЛИЗ СВОЙСТВ НИЗКОЧАСТОТНОГО МИКРОСЕЙСМИЧЕСКОГО ШУМА

© 2011 г. А. А. Любушин

*Институт физики Земли им. О.Ю. Шмидта РАН, г. Москва*

Поступила в редакцию 26.04.2010 г.

Излагается метод описания состояния низкочастотного микросейсмического фона в крупном сейсмоактивном регионе (Японские острова) на основе данных сети широкополосных сейсмических станций. Для описания используются медианы по ежесуточным оценкам от каждой станции сети для 7 параметров: 3-х характеристик мульти-фрактальных спектров сингулярности волновых форм, их спектральной экспоненты, индекса их гладкости, логарифма дисперсии и индекса линейной предсказуемости. Эти величины определяются для последовательных временных интервалов длиной 1 сутки от начала января 1997 года по конец февраля 2010 года. Поскольку эти параметры являются результатом взятия медианы по их оценкам от каждой работающей станции сети, то они несут характер интегральных статистик поля микросейсм. Статья является продолжением двух предыдущих работ [Любушин, 2009; 2010], основанных на использовании данных сети F-net для анализа эффектов синхронизации поля низкочастотных микросейсм на длительных временных масштабах. Ставится задача выделения числа различных “мод поведения” поля микросейсм как числа кластеров, на которые оптимальным образом разбивается облако 7-мерных векторов параметров при оценке в скользящем временном окне длиной 2 года. Вводится понятие кластерной экспоненты как показателя степени зависимости величины функции компактности облака векторов параметров от числа кластеров при его оптимальном разбиении в качестве новой характеристики геофизического поля. Ранее было выделено сравнительно быстрое повышение уровня синхронизации поля микросейсм, которое началось в середине 2002 года и заняло времени около 1 года. Это высокий уровень синхронизации сохраняется по настоящее время. В последние 4 года (с учетом 2-годовой длины временного окна, в котором производились оценки) наблюдается продолжительный тренд в значении кластерной экспоненты, аналогичный более короткому тренду перед событием на Хоккайдо 25.09.2003 г.,  $M = 8.3$ . Эти факты, в совокупности с особенностями изменения коэффициента корреляции между 2-мя мульти-фрактальными параметрами поля, позволяют выдвинуть гипотезу о повышении сейсмической опасности на Японских островах со второй половины 2010 года.

### ВВЕДЕНИЕ

Низкочастотные микросейсмические колебания и их связи с процессами в гидросфере и атмосфере Земли, как основными источниками энергии микросейсм, являются довольно традиционным объектом исследования в геофизике [Friedrich et al., 1998; Kobayashi, Nishida, 1998; Tanimoto et al., 1998; Tanimoto, Um, 1999; Ekstrom G., 2001; Tanimoto, 2001; 2005; Kurrle, Widmer-Schmidrig, 2006; Stehly et al., 2006; Rhie, Romanowicz, 2004; 2006].

Очевидно, однако, что изменения структуры микросейсмического фона может отражать также и изменения свойств земной коры, как среды распространения микросейсм. Именно эти изменения свойств микросейсм, особенно в связи с последовательностью сильных землетрясений, является объектом исследования в данной статье. В работах [Соболев и др., 2005; 2008; Соболев, Любушин, 2006] исследовались особенности периодической структуры точечного процесса, образуемого сильными выбросами и низкочастотными асимметричными импульсами микросейсмического шума перед сильными землетрясениями. В работах [Любушин, Соболев,

2006; Любушин, 2007; 2008; 2009; 2010] анализировались эффекты синхронизации вариаций параметров мульти-фрактальных спектров сингулярности [Feder, 1988; Kantelhardt et al., 2002] волновых форм, оцененных в скользящих временных окнах “малой” длины, причем параметры оценивались как по записям отдельных станций, так и как медианы оценок параметров от пространственных групп станций [Любушин, 2009; 2010], что давало возможность использовать длительные записи сейсмического фона, содержащие множество разрывов в регистрации.

В работе [Любушин, 2009] был произведен анализ записей вертикальных компонент из базы данных F-net с частотой дискретизации 1 Гц за период с начала 1997 г. по июнь 2008 г. Было обнаружено статистически значимое уменьшение среднего (по времени) значения медианы (по станциям сети) ширины носителя спектра сингулярности  $\Delta\alpha$  волновых форм микросейсм за полгода до землетрясения на Хоккайдо 25.09.2003 г., причем после этого события среднее значение  $\Delta\alpha$  не вернулось на прежний уровень и стабилизировалось на низких значениях. Было показано, что это изменение сопровождалось увеличением эффекта синхронизации между медианами

нами параметров спектров сингулярности, вычисленными по 5 группам станций. Эффект синхронизации оценивался в последовательных непересекающихся временных окнах длиной 2 месяца. Аналогии с динамикой связанных нелинейных осцилляторов позволяют утверждать, что сужение ширины носителя спектра сингулярности (что формально означает уменьшение разнообразия типов стохастического поведения) непосредственно связано с синхронизацией элементов сложной системы [Pavlov et al., 2003].

В работе [Любушин, 2010] исследовалась задача выделения эффектов синхронизации параметров низкочастотного микросейсмического фона по данным сети F-net за период с начала 1997 г. по август 2009 г. При этом список анализируемых параметров был существенно расширен: помимо ранее рассмотренных ширины носителя  $\Delta\alpha$  мульти-фрактального спектра сингулярности и обобщенного показателя Херста  $\alpha^*$  (аргумента, реализующего максимум спектра сингулярности), рассматривались также коэффициент асимметрии спектра сингулярности, логарифм дисперсии, спектральная экспонента и индекс линейной предсказуемости. Эти параметры вычисляются по реализациям внутри последовательных интервалов времени длиной 1 сутки. При использовании скользящего временного окна, равного 1 году для оценки множественной меры корреляции, суточные вариации медианных значений статистик фона от 5 пространственных кластеров станций демонстрируют устойчивый рост степени синхронизации незадолго до землетрясения на Хоккайдо 25.09.2003 г. с последующим выходом на новый уровень высокой синхронизации. Анализ поведения трендов индекса линейной предсказуемости позволил довольно точно определить начало процесса усиления синхронизации — середина 2002 года. Полученный эффект для вариаций различных параметров микросейсм является независимым подтверждением сделанного ранее вывода о синхронизации поля микросейсмического фона на Японских островах. Заметим, что использование годового окна для вычисления меры корреляции эквивалентно усреднению сезонных воздействий циклонов, штормов и ураганов как основных генераторов низкочастотных микросейсм и делает оценку множественных корреляций весьма устойчивой.

Однако, оценки множественных мер когерентности и корреляции в скользящих временных окнах, использованные в работах [Любушин, 2009; 2010], не позволили выделить других значимых аномалий поведения параметров поля микросейсм за время наблюдений 1997–2010 гг., помимо сравнительно быстрого перехода на уровень высокой синхронизации во второй половине 2002–начале 2003 гг. В настоящей работе используется другой метод анализа данных, а именно кластерный анализ облака векторов параметров фона в скользящем окне длиной

2 года. Исходной целью такого анализа является поиск ответа на вопрос, сколько “мод поведения” микросейсмического фона можно выделить и как число этих мод меняется во времени. Здесь под модой поведения имеется в виду кластер (компактная группа) векторов параметров в текущем “длинном” 2-годовом окне. Использование этого подхода позволяет ввести новый параметр — “кластерную экспоненту”  $\mu$  и выделить аномальный участок поведения фона 2007–2010 гг., характеризующийся положительным трендом значений  $\mu$ , аналогичный более короткому тренду перед событием на Хоккайдо 25.09.2003 г. При этом подходе станции сети не делятся на заданное число пространственных групп с целью последующего вычисления мер когерентности или корреляции между вариациями различных параметров фона от каждой группы станций в скользящих временных окнах, как это делалось в работах [Любушин, 2009; 2010]. Таким образом, медианы параметров фона берутся по всем станциям сети, что делает используемые параметры интегральными характеристиками поля микросейсм в области, покрываемой сетью наблюдений.

#### ИСПОЛЬЗУЕМЫЕ СТАТИСТИКИ

Для анализа использовались данные широкополосной сейсмической сети F-net, <http://www.hinet.bosai.go.jp/fnet>, за период наблюдения с начала 1997 года по конец февраля 2010 года включительно (более 13 лет). Схема расположения 83-х станций сети имеется как по вышеуказанной ссылке в сети, так и в работах [Любушин, 2009; 2010]. Анализируемые данные — вертикальные компоненты с шагом по времени 1 с. Эти сейсмические записи приводились к шагу по времени 1 минута путем вычисления средних значений по последовательным 60 отсчетам.

Ниже кратко описаны 7 статистик, используемые для кластерного анализа. Их подробное описание имеется в работе [Любушин, 2010] вследствие чего многие технические детали их вычисления ниже опущены.

**Спектральная экспонента  $\beta$ , индекс гладкости волновой формы  $\xi$  и логарифм дисперсии  $\lg(Var)$ .** Спектральная экспонента  $\beta$  определяет тип изменения логарифма спектра мощности в зависимости от логарифма периода и ее значение тесно связано с фрактальными характеристиками шума [Feder, 1988; Kantelhardt et al., 2002]. Ниже вместо классических оценок спектра мощности на основе Фурье-разложения или параметрических моделей использовалась ее оценка по скорости изменения средних значений квадратов модулей вейвлет-коэффициентов  $W_k$  [Mallat, 1998] в зависимости от номера уровня детальности  $k = 1, \dots, m$  согласно формуле  $\log_2(W_k) \sim k^\beta$ , что автоматически обеспечивает сильное сглаживание спектра, необходимое для оценки спектральной экспоненты. Параметр  $\beta$  оце-

нивался в последовательных временных окнах длиной 1 сутки (1440 минутных отсчетов). Ранее такой способ вычисления спектральных экспонент при анализе микросейсмического шума использовался в [Любушин, 2008]. Чтобы исключить влияние приливных вариаций, в каждом окне устранялся полиномиальный тренд 8-го порядка и для остатка вычислялся вейвлетный спектр мощности  $W_k$  и десятичный логарифм дисперсии  $\lg(Var)$ . При этом выбирался оптимальный ортогональный вейвлет Добеши с числом обнуляемых моментов от 1 до 10 (число обнуляемых моментов 1 соответствует вейвлету Хаара), реализующий минимум энтропии распределения квадратов вейвлет-коэффициентов для первых 7 уровней детальности вейвлет-разложения (масштабы или “периоды” от 2 до 256 минут при шаге по времени 1 минута). Автоматически для данного метода вычисления спектральной экспоненты получается еще одна полезная характеристика  $\xi$  – число обнуляемых моментов оптимального вейвлета, с возможными целочисленными значениями от 1 до 10. Чем больше значение  $\xi$ , тем более гладкой является волновая форма в пределах текущих суток. Значения медиан по всем станциям сети F-net, работавшим в течение рассматриваемого дня будем обозначать  $\bar{\beta}(s)$ ,  $\bar{\xi}(s)$  и  $\overline{\lg(Var)}(s)$ . В этих обозначениях верхняя черта означает пространственное усреднение (взятие медианы), а аргумент  $s$  является целочисленным индексом, нумерующим последовательные сутки от начала наблюдений – 01.01.1997 г. Отметим, что после вычисления медианы по всем станциям показатель гладкости  $\bar{\xi}(s)$  перестает быть целым числом.

**Индекс линейной предсказуемости  $\rho$**  вычисляется по формуле:  $\rho = V_0/V_{AR} - 1$ . Здесь  $V_0$  – дисперсия ошибки  $\varepsilon_0(t+1)$  тривиального прогноза  $\hat{x}_0(t+1)$  на 1 шаг вперед для приращений  $x(t)$  сейсмических записей, который равен среднему по предыдущему “малому” временному окну длиной  $n$  отсчетов:  $\hat{x}_0(t+1) = \sum_{s=t-n+1}^t x(s)/n$ . Таким образом,  $\varepsilon_0(t+1) = x(t+1) - \hat{x}_0(t+1)$ , а  $V_0 = \sum_{t=n+1}^N \varepsilon_0^2(t)/(N-n)$ , где  $N > n$  – число отсчетов в последовательных “больших” временных фрагментах. Величина  $V_{AR}$  вычисляется по аналогичной формуле  $V_{AR} = \sum_{t=n+1}^N \varepsilon_{AR}^2(t)/(N-n)$ , в которой  $\varepsilon_{AR}(t+1) = x(t+1) - \hat{x}_{AR}(t+1)$  есть ошибка линейного прогноза  $\hat{x}_{AR}(t+1)$  на 1 шаг вперед с помощью модели авторегрессии 2-го порядка (AR-прогноз), коэффициенты которой оцениваются также по предыдущему “малому” временному окну длиной  $n$  отсчетов. Выбор 2-го порядка авторегрессии обусловлен тем, что этот порядок – минимальный для AR-модели, при котором описывается колебательное движение и допускается положение максимума спектральной плотности в

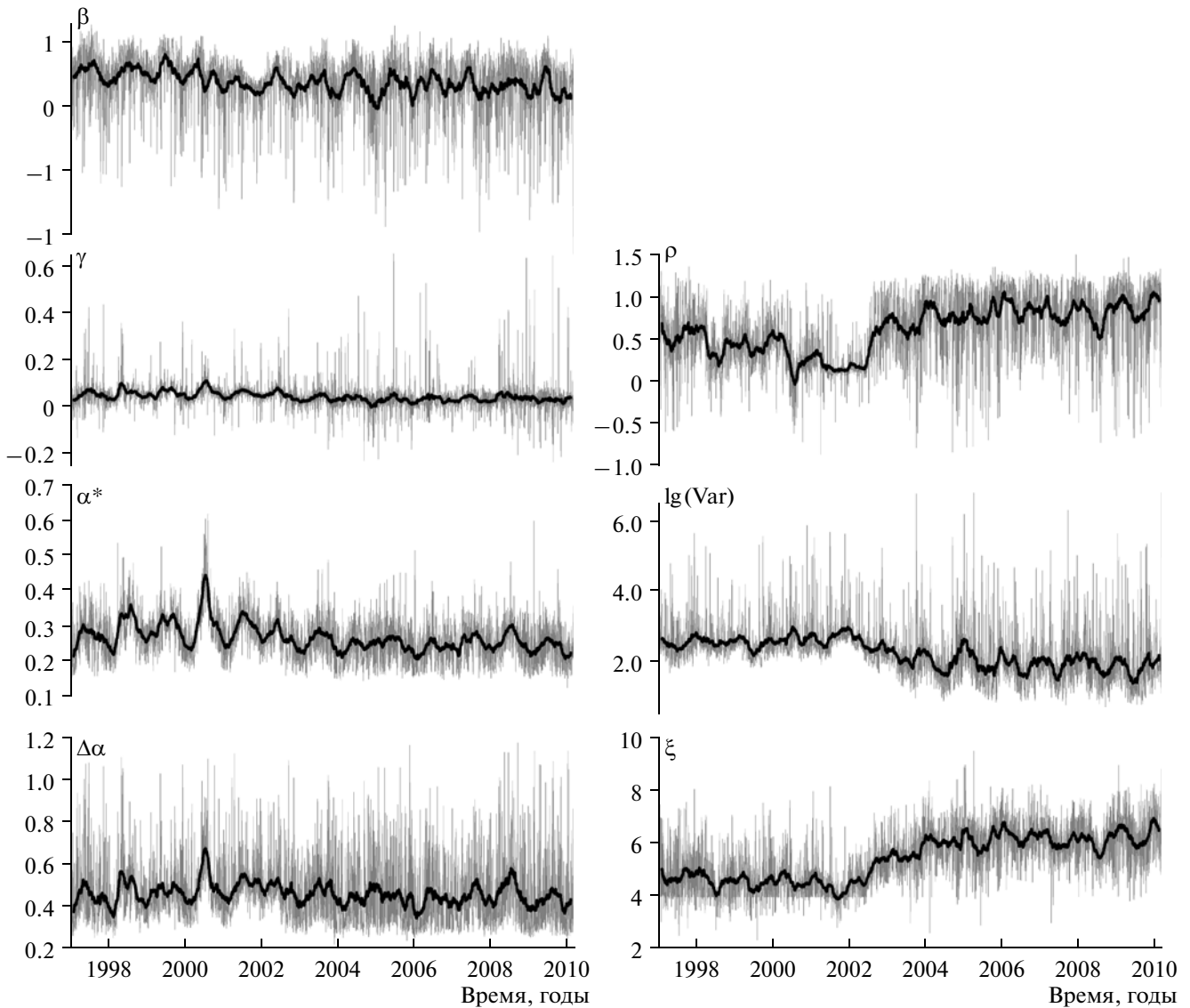
значениях частот между частотой Найквиста и нулевой [Box, Jenkins, 1970; Kashyap, Rao, 1976]. Переход к приращениям связан с необходимостью избавления от доминирования низких частот (приливной компоненты и прочих трендов). AR-прогноз использует свойство коррелированности соседних значений приращений записей и, если она имеет место, то  $V_{AR} < V_0$  и  $\rho > 0$ . При вычислении индекса линейной предсказуемости  $\rho$  для 1-минутных данных оценки производились в последовательных “длинных” временных окнах размера  $N = 1440$  (1 сутки) и при “коротком” окне длиной  $n = 60$  отсчетов (1 час).

**Параметры  $\alpha^*$ ,  $\Delta\alpha$  и  $\gamma$  мульти-фрактального спектра сингулярности.** Пусть  $X(t)$  – случайный процесс. Определим в качестве меры  $\theta(t, \delta)$  поведения сигнала  $X(t)$  на интервале  $[t, t + \delta]$  размах:  $\theta(t, \delta) = \max_{t \leq \lambda \leq t + \delta} X(\lambda) - \min_{t \leq \lambda \leq t + \delta} X(\lambda)$ . Спектр сингулярности  $F(\alpha)$  можно определить как фрактальную размерность моментов времени  $\lambda_\omega$  имеющих одно и то же значение локального показателя Гельдера–Липшица:  $h(t) = \lim_{\delta \rightarrow 0} \frac{\ln(\theta(t, \delta))}{\ln(\delta)}$ , то есть  $h(\lambda_\omega) = \alpha$ .

Для оценки спектра сингулярности в скользящем временном окне в статье использовалась разновидность метода DFA [Kantelhardt et al., 2002], которая подробно описана в работах [Любушин, Соболев, 2006; Любушин, 2007; 2008; 2009; 2010]. Ниже при анализе низкочастотных микросейсм использовались оценки спектра сингулярности в последовательных непересекающихся суточных временных окнах, локальные масштабно-зависимые тренды удалялись полиномами 8-го порядка. Основное внимание будет уделено исследованию изменений 2-х параметров спектра сингулярности – обобщенного показателя Херста  $\alpha^*$  и ширины носителя спектра сингулярности  $\Delta\alpha$ . Величина  $\alpha^*$  характеризует наиболее типичный и часто встречающийся показатель Гельдера–Липшица, тогда как  $\Delta\alpha$  отражает разнообразие случайного поведения сигнала и является своего рода мерой числа скрытых степеней свободы стохастической системы. Помимо этих 2-х параметров будет анализироваться также величина  $\gamma = \alpha^* - (\alpha_{\min} + \alpha_{\max})/2$ , которая характеризует степень асимметрии спектра сингулярности.

Аналогично обозначениям  $\bar{\beta}(s)$ ,  $\bar{\xi}(s)$  и  $\overline{\lg(Var)}(s)$ , введенным выше, будем обозначать через  $\bar{\rho}(s)$ ,  $\bar{\alpha}^*(s)$ ,  $\bar{\Delta\alpha}(s)$  и  $\bar{\gamma}(s)$  значения медиан соответствующих статистик, где верхняя черта символизирует операцию взятия медианы по данным от всех станций сети (что эквивалентно пространственному усреднению), а аргумент  $s$  нумерует последовательные сутки.

На рис. 1 представлены графики изменения медиан всех 7 рассматриваемых статистик вместе с их средними значениями в скользящем окне длиной

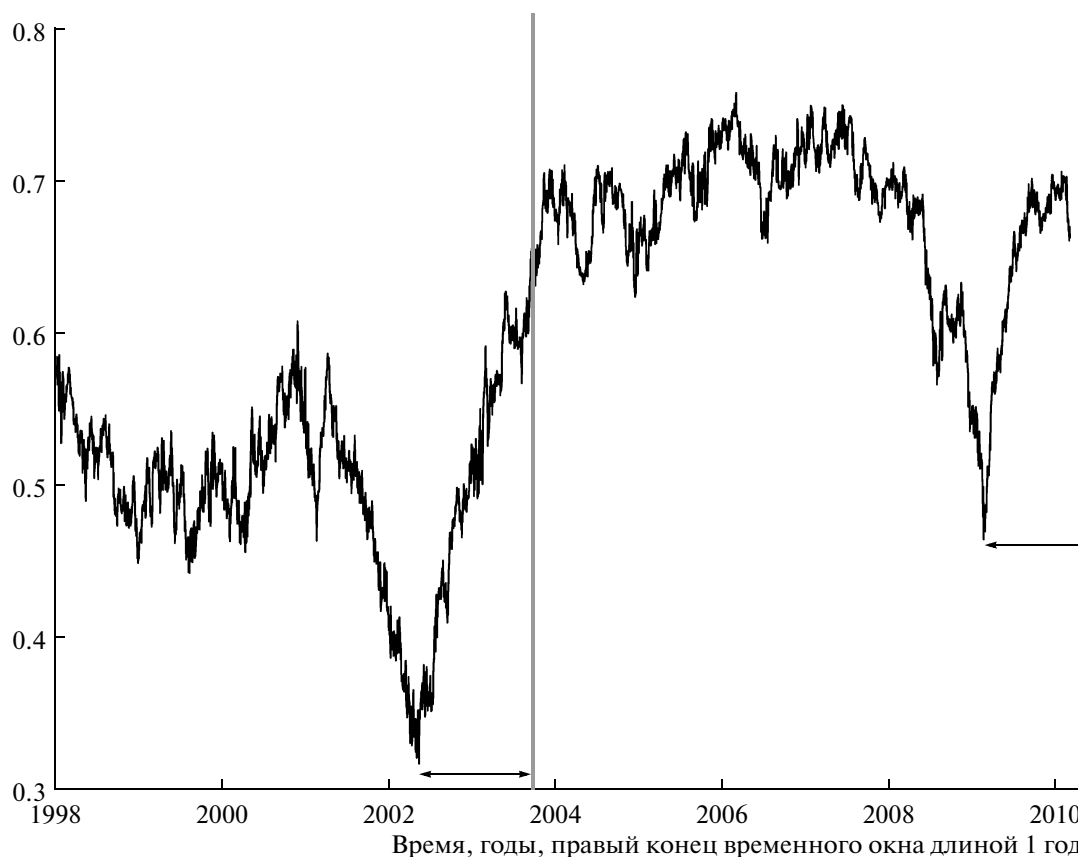


**Рис. 1.** Зависимости статистик, вычисляемых как медианы по всем сейсмическим станциям сети F-net в последовательных временных окнах длиной 1 сутки:  $\beta$  – спектральная экспонента;  $\gamma$  – коэффициент асимметрии спектра сингулярности;  $\alpha^*$  – обобщенный показатель Херста;  $\Delta\alpha$  – ширина носителя спектра сингулярности;  $\rho$  – индекс линейной предсказуемости;  $\lg(\text{Var})$  – десятичный логарифм дисперсии;  $\xi$  – индекс гладкости волновых форм. Величины  $\beta$ ,  $\lg(\text{Var})$  и  $\xi$  вычислялись после удаления тренда полиномом 8-го порядка в каждом окне длиной 1 сутки. Черными толстыми линиями даны графики скользящих средних в окне длиной 57 суток.

57 суток. Такое значение выбрано как кратное двум длинам лунного месяца (28 суток – период модуляции многих геофизических процессов), а дополнительная единица возникла из необходимости иметь нечетное значение длины окна скользящего усреднения. На рис. 2 представлены графики изменения робастного коэффициента корреляции [Любушин, 2007; 2010] между параметрами  $\bar{\alpha}^*(s)$  и  $\Delta\alpha(s)$ , вычисленного в скользящем временном окне длиной 1 год. Следует отметить, что все множественные коэффициенты корреляций для медиан всех параметров демонстрируют схожее увеличение перед собы-

тием 25.09.2003 г., аналогично эффектам увеличения синхронизации между изменениями медиан одного и того же параметра, но вычисленных от разных групп станций в работе [Любушин, 2010]. Однако рис. 2 примечателен тем, что он содержит 2 яркие аномалии поведения коэффициента корреляции – острые минимумы в 2002 и 2009 гг. Поскольку после первой аномалии 2002 года последовало крупное землетрясение 25.09.2003 г., то логично предположить, что второй острый минимум коэффициента корреляции может быть также предвестником будущего сильного события во 2-й половине 2010 года, причем это событие может быть большим по энер-

25.09.2003 г.



**Рис. 2.** График робастного коэффициента корреляции между вариациями величин  $\alpha^*$  и  $\Delta\alpha$  (см. рис. 1) в скользящем временном окне 365 суток (1 год) в зависимости от положения правого конца временного окна. Вертикальной линией обозначен момент времени землетрясения около о. Хоккайдо 25.09.2009 г.,  $M = 8.3$ , двенаправленными горизонтальными стрелками выделены временные интервалы одинаковой длины от момента первого минимума коэффициента корреляции до момента землетрясения 25.03.2009 г. и от момента второго минимума до июля 2010 года.

гии. В пользу последнего довода об ожидании более сильного события говорит факт увеличения синхронности поведения поля микросейсм, начиная со второй половины 2002 года, установленный ранее в работах [Любушин, 2009; 2010].

Далее бралось скользящее временное окно длиной 730 суток (2 года) и для каждого окна производилась следующая последовательность операций с облаком 7-мерных векторов  $\psi$  с компонентами  $\bar{\beta}$ ,  $\bar{\xi}$ ,  $\overline{\lg(\text{Var})}$ ,  $\bar{\rho}$ ,  $\bar{\alpha}^*$ ,  $\bar{\Delta\alpha}$  и  $\bar{\gamma}$  внутри каждого окна:

1). Каждая компонента вектора  $\psi$  подвергалась операции нормализации и винзоризации: вычислялись выборочные средние значения и стандартные отклонения  $\sigma$ , выборочные значения вычитались, значения, выходящие за  $\pm 4\sigma$ , обрезались, после чего производилось деление на  $\sigma$  — и так итерационно, до тех пор, пока значение  $\sigma$  не перестанет меняться.

2). Для полученного облака нормализованных 7-мерных векторов вычислялись первые 4 главные компоненты как проекции на собственные вектора ковариационной матрицы внутри текущего окна,

соответствующие 4-м максимальным собственным числам (это обеспечивало дополнительное подавление шума и сохраняло от 91 до 95% общей дисперсии).

3). Для полученного облака 4-мерных векторов главных компонент производилось разбиение на заданное число  $q$  кластеров  $\Gamma_k$ ,  $k = 1, \dots, q$ . Пробное число кластеров варьировалось последовательно от 40 до 2-х. Разбиение производилось последовательно иерархической кластеризацией с использованием метрики “дальний сосед” (что обеспечивает получение компактных и “округлых” кластеров) и последующей итерации метода К-средних [Айвазян и др., 1989; Duda, Hart, 1973].

Пусть  $N$  — общее число 4-мерных векторов  $\zeta$  главных компонент (от нормализованных 7-мерных векторов) в текущем временном окне, а  $\zeta_0$  — вектор общего центра масс облака главных компонент (вследствие предварительных операций нормализации и винзоризации  $\zeta_0 = 0$ ). Пусть  $\bar{\zeta}_k$ ,  $k = 1, \dots, q$  — вектора центров масс кластеров, а  $n_k$  — число эле-

ментов в каждом кластере,  $\sum_{k=1}^q n_k = N$ . Разбиение облака из  $N$  векторов на заданное число  $q$  кластеров оценивается следующими величинами:

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

– мера внутрикластерной компактности;

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

– взвешенная мера расхождения между центрами кластеров;

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

– так называемая псевдо- $F$ -статистика [Vogel, Wong, 1978].

Разбиение облака на  $q$  кластеров производится с целью минимизации величины  $\sigma_0^2(q)$ . Формально величина  $\sigma_0^2(q)$  может быть определена и для  $q = 1$ :  $\sigma_0^2(1) = \sum_{\zeta} |\zeta - \zeta_0|^2 / (N - 1)$ . Величина  $\sigma_0^2(q)$  монотонно возрастает с убыванием  $q$  и зависимость  $\lg(\sigma_0^2(q))$  от  $\lg(q)$  близка к линейной, то есть  $\sigma_0^2(q) \sim q^{-\mu}$ . Этот факт иллюстрируется на рис. 3. Назовем величину  $\mu$  кластерной экспонентой. Ее можно оценить для заданного окна как коэффициент наклона прямой линии, подогнанной методом наименьших квадратов к зависимости  $\lg(\sigma_0^2(q))$  от  $\lg(q)$ .

Величина (3) характеризует качество разбиения на заданное число кластеров: чем больше  $PFS(q)$ , тем разбиение лучше. Для хорошего разбиения величине  $\sigma_0^2(q)$  внутрикластерной компактности желательнее быть малым, а расхождению  $\sigma_1^2(q)$  между кластерами – большим. Оптимальное число кластеров  $q^*$  находится из условия максимума  $PFS(q)$ . В то же время  $PFS(q)$  (а точнее  $\sigma_1^2(q)$ ) не может быть вычислена для  $q = 1$ . Поэтому для различения случаев  $q = 1$  и  $q = 2$  нужно применять другие соображения. Известно, что оптимальное число кластеров может быть также определено по точке излома монотонной зависимости  $\sigma_0^2(q)$  для  $q = q^*$ : для  $q < q^*$  функция  $\sigma_0^2(q)$  растет быстрее с убыванием  $q$ , чем для  $q > q^*$ . Этот критерий определения  $q = q^*$  более подвержен влиянию шума и работает хуже, чем использование метода  $q^* = \arg \max_{2 \leq q} PFS(q)$ , но для различения случаев  $q = 1$  и  $q = 2$  он является единственной возмож-

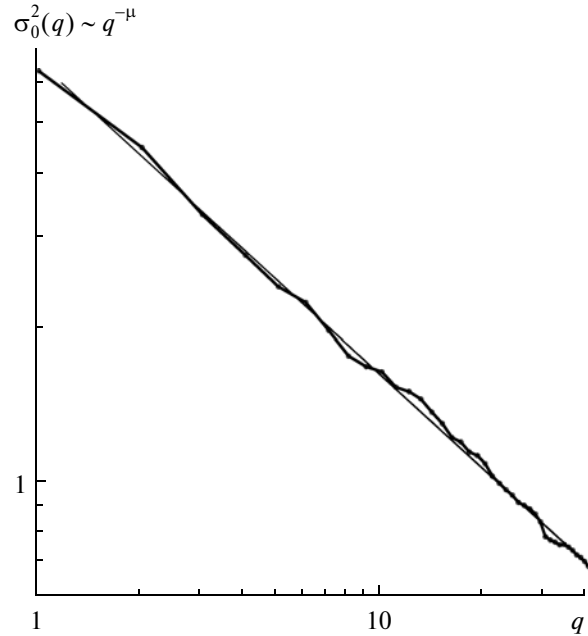


Рис. 3. Зависимость функционала внутри кластерной компактности  $\sigma_0^2(q)$  от пробного числа кластеров  $q$ , тонкой прямой линией представлен график наилучшего линейного приближения зависимости  $\lg(\sigma_0^2(q))$  от  $\lg(q)$ .

ностью. Обозначим через  $\delta \lg(\sigma_0^2(q))$  отклонение величины  $\lg(\sigma_0^2(q))$  от прямой наилучшего приближения зависимости от  $\lg(q)$ . Тогда будем считать, что точка  $q = 2$  является точкой излома зависимости  $\sigma_0^2(q)$ , если  $\delta \lg(\sigma_0^2(1))$  превышает все значения  $\delta \lg(\sigma_0^2(q))$  для  $q \geq 2$ . Таким образом, оптимальное число  $q^*$  кластеров определим согласно правилу:

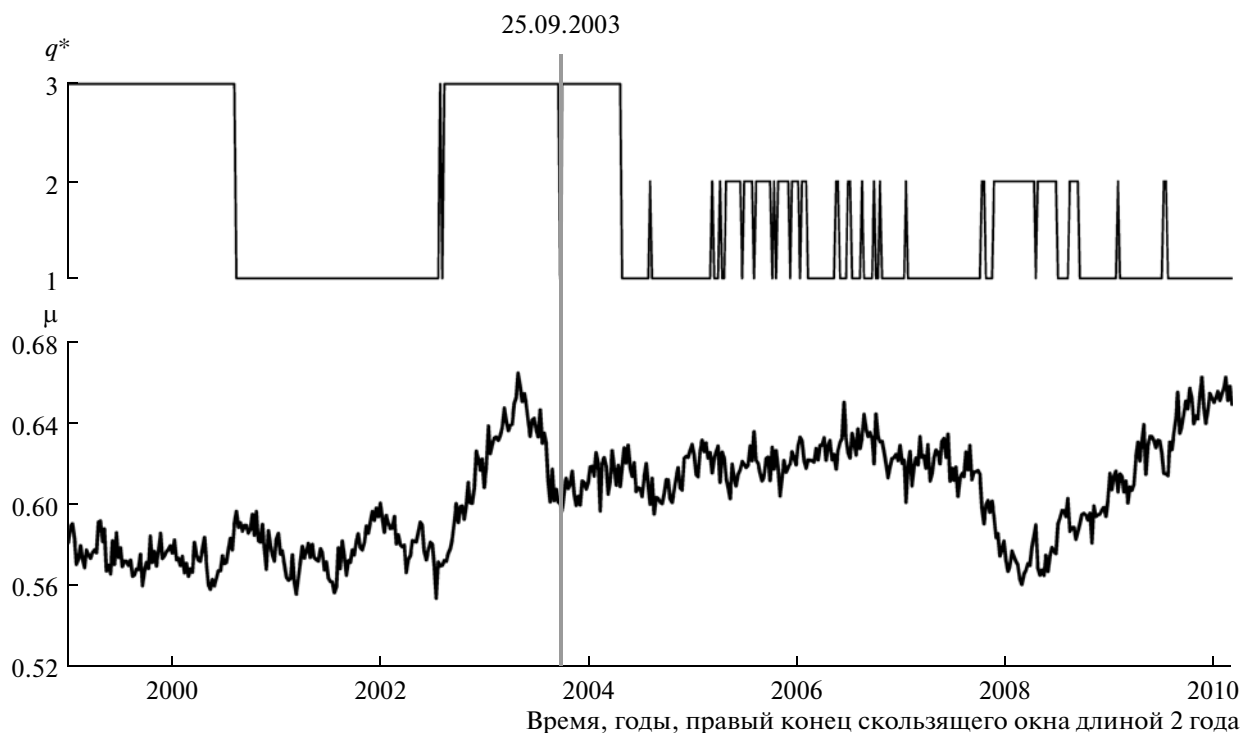
$$\begin{aligned} \text{Пусть } q_0 &= \arg \max_{2 \leq q \leq 40} PFS(q), \\ \text{если } q_0 > 2, & \text{ то } q^* = q_0, \\ & \text{иначе:} \end{aligned} \quad (4)$$

$$\begin{aligned} \text{если } \delta \lg(\sigma_0^2(1)) \leq \max_{2 \leq q \leq 40} \delta \lg(\sigma_0^2(q)), & \text{ то } q^* = 1, \\ \text{иначе } & q^* = 2. \end{aligned}$$

На рис. 4 представлены результаты кластеризации 4-х главных компонент 7 ежедневных медианных характеристик микросейсмического фона сети F-net для 13 лет наблюдений (1997–2009 гг.) при оценке в скользящем временном окне длиной 2 года ( $N = 730$ ) со смещением 7 суток.

### ОБСУЖДЕНИЕ

Следует отметить изменение характера переключения между числом оптимальных кластеров после



**Рис. 4.** Результаты кластерного анализа 4-х первых главных компонент в скользящем временном окне длиной 2 года со смещением 7 суток в зависимости от положения правого конца временного окна:  $q^*$  — оптимальное число кластеров;  $\mu$  — значение кластерной экспоненты. Вертикальной линией обозначен момент времени землетрясения около о. Хоккайдо 25.09.2009,  $M = 8.3$ .

2004 года на рис. 4 — режим переключений стал более хаотичным, причем вариант 3-х кластеров, который был преобладающим ранее, полностью исчез. Этот факт может быть интерпретирован как “замораживание” определенной внутренней степени свободы поля микросейсм после события 25.09.2003 г. перед возможным будущим сильным толчком. Кластерная экспонента  $\mu$  имеет выраженный линейный тренд с последующим достижением максимума и возвратным движением на новый средний уровень значений перед событием 25.09.2003 г. Эта особенность поведения значительно отличается от предыдущего фона статистических флуктуаций вокруг среднего значения. Начальная стадия этой аномалии — линейный тренд, повторяется с положения правого конца 2-годового окна в начале 2008 года, причем положительный тренд  $\mu$  имеет более длительный характер и продолжается по настоящее время.

Результаты анализа получены после применения следующей последовательности операций: а) оценка 7-мерного вектора параметров  $\psi$  поля микросейсмических шумов; б) перехода к 4-мерному вектору  $\zeta$  их первых главных компонент; в) предварительных операций нормализации и винзоризации компонент вектора  $\zeta$  независимо в каждом 2-годовом временном окне; г) оценка оптимального числа кла-

стеров и кластерной экспоненты. Смысл этой последовательности заключается в получении некоторых параметров поля, сглаженных как по пространству (в результате взятия медианы по станциям сети), так и по времени (сначала в результате оценки в окнах длиной 1 сутки, а затем путем анализа свойств кластеризации облака векторов параметров в 2-годовом окне). Таким образом, общая цель является информационной, и это является трудностью в попытках ответить на вопрос, а каков же физический смысл полученных результатов.

Впрочем, смысл отдельных компонент вектора  $\psi$  довольно прозрачен (далее речь пойдет о средних значениях, что символизируется верхней чертой):  $\overline{\lg(\text{Var})}$  — логарифм энергии шума;  $\overline{\beta}$  — спектральная экспонента (наклон графика спектра мощности в двойном логарифмическом масштабе);  $\overline{\alpha^*}$ ,  $\overline{\Delta\alpha}$  и  $\overline{\gamma}$  — параметры мульти-фрактального спектра сингулярности, который уже давно используются при анализе физических систем, демонстрирующих случайное поведение, вытекающее из их внутренней нелинейной динамики (например, при исследовании турбулентности). Что же касается параметров  $\overline{\xi}$  (среднее значение индекса гладкости волновых форм шума) и  $\overline{\rho}$  (среднее значение индекса линей-

ной предсказуемости шума), введенных в работе [Любушин, 2010], то они описывают геометрические и динамические свойства случайных микросейсмических колебаний и могут быть интерпретированы как показатели их “обратной сложности”: чем меньше значение  $\bar{\xi}$  и  $\bar{\rho}$ , тем сложнее устроен шум. Заметим, что параметр  $\overline{\Delta\alpha}$  также является мерой сложности, но уже “прямой”: чем больше  $\overline{\Delta\alpha}$ , тем сигнал сложнее (напомним, что для моно-фрактального сигнала  $\overline{\Delta\alpha} = 0$ ).

Таким образом, каждому графику на рис. 1 может быть дана та или иная физическая интерпретация. Отметим, что быстрое изменение среднего по времени значения  $\bar{\rho}$  на рис. 1 дало возможность оценить начало синхронизации как середину 2002 года в работе [Любушин, 2010]. Дальнейшие операции нормализации, перехода к главным компонентам и кластерного анализа “смешивают” исходные сущности компонент вектора  $\psi$ . В результате, например, кластерная экспонента  $\mu$  имеет довольно абстрактный характер и ее трудно привязать к какому-нибудь “первичному” физическому закону или системе уравнений. Но такая ситуация является общей при попытке реального описания сложной многокомпонентной системы типа земной коры. Вопрос о физической интерпретации поведения  $\mu$  на рис. 4 или коэффициента корреляции между  $\bar{\alpha}^*$  и  $\overline{\Delta\alpha}$  на рис. 2 является открытым.

### ВЫВОДЫ

Предложен метод анализа состояния поля микросейсмического шума на основе исследования кластеризации его параметров, усредненных по станциям сети, в скользящем временном окне. Применительно к анализу поля микросейсмических колебаний на Японских островах этот метод выделил как аномалию перед землетрясением на Хоккайдо 25.09.2003 г., так и современную аномалию, начавшуюся в 2007–2008 гг. Делается предположение, что последняя особенность поведения может быть индикатором существенного увеличения сейсмической опасности в Японии со второй половины 2010 года.

Работа выполнена при поддержке гранта РФФИ 09-05-00134.

### СПИСОК ЛИТЕРАТУРЫ

Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Классификация и снижение размерности. М.: Финансы и статистика. 1989. 607 с.  
 Любушин А.А., Соболев Г.А. Мультифрактальные меры синхронизации микросейсмических колебаний в минутном диапазоне периодов // Физика Земли. 2006. № 9. С. 18–28.

Любушин А.А. Анализ данных систем геофизического и экологического мониторинга. М.: Наука. 2007. 228 с.  
 Любушин А.А. Микросейсмический шум в минутном диапазоне периодов: свойства и возможные прогностические признаки // Физика Земли. 2008. № 4. С. 17–34.  
 Любушин А.А. Тренды и ритмы синхронизации мультифрактальных параметров поля низкочастотных микросейсм // Физика Земли. 2009. № 5. С. 15–28.  
 Любушин А.А. Статистики временных фрагментов низкочастотных микросейсм: их тренды и синхронизация // Физика Земли. 2010. № 6. С. 86–96.  
 Соболев Г.А., Любушин А.А., Закржевская Н.А. Синхронизация микросейсмических колебаний в минутном диапазоне периодов // Физика Земли. 2005. № 8. С. 3–27.  
 Соболев Г.А., Любушин А.А. Микросейсмические импульсы как предвестники землетрясений // Физика Земли. 2006. № 9. С. 5–17.  
 Соболев Г.А., Любушин А.А., Закржевская Н.А. Асимметричные импульсы, периодичности и синхронизация низкочастотных микросейсм // Вулканология и сейсмология. 2008. № 2. С. 135–152.  
 Box G.E.P., Jenkins G.M. Time series analysis. Forecasting and control. Holden-Day. San Francisco, Cambridge, London, Amsterdam. 1970 (Русский перевод: Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. М.: Мир, в 2-х выпусках. 1974. 406 с. и 197 с.).  
 Duda R.O., Hart P.E. (1973) Pattern classification and scene analysis, John Wiley & Sons, N.Y., London, Sydney, Toronto (Русский перевод: Дуда Р., Харт П. (1976) Распознавание образов и анализ сцен. М.: Мир. 511 с.).  
 Ekstrom G. Time domain analysis of Earth's long-period background seismic radiation // J. Geophys. Res. 2001. V. 106. № B11. P. 26483–26493.  
 Feder J. Fractals. Plenum Press, New York, London. 1988 (Русский перевод: Федер Е. Фракталы. М.: Мир. 1991. 254 с.).  
 Friedrich A., Krüger F., Klinge K. Ocean-generated microseismic noise located with the Gräfenberg array // Journal of Seismology. V. 2. Number 1 / March 1998. P. 47–64.  
 Kantelhardt J.W., Zschiegner S.A., Koncsienly-Bunde E., Havlin S., Bunde A., Stanley H.E. // Multifractal detrended fluctuation analysis of nonstationary time series, Physica A, 316, 87. V. 114. 2002.  
 Kashyap R.L., Rao A.R. Dynamic stochastic models from empirical data. Academic Press. N.Y., San Francisco, London. 1976 (Русский перевод: Кашьяп Р.Л., Рао А.Р. Построение динамических стохастических моделей по экспериментальным данным. М.: Наука. 1983. 384 с.).  
 Kobayashi N., Nishida K. Continuous excitation of planetary free oscillations by atmospheric disturbances // Nature. 1998. V. 395. P. 357–360.  
 Kurrle D., Widmer-Schmid R. Spatiotemporal features of the Earth's background oscillations observed in central Europe // Geophys. Res. Lett. 2006. V. 33. P. L24304.  
 Mallat S. A wavelet tour of signal processing. Academic Press. San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto. 1998. 577 p. (Русский перевод: Малла С. Вэйвлеты в обработке сигналов. М.: Мир. 2005. 671 с.).



- Pavlov A.N., Sosnovtseva O.V., Mosekilde E.* Scaling features of multimode motions in coupled chaotic oscillators. *Chaos, Solitons and Fractals*. 2003. V. 16. P. 801–810.
- Rhie J., Romanowicz B.* Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling // *Nature*. 2004. V. 431. 552//554.
- Stehly L., Campillo M., Shapiro N.M.* A study of the seismic noise from its long-range correlation properties // *J. Geophys. Res.* 2006. V. 111. B10306.
- Tanimoto T., Um J., Nishida K., Kobayashi N.* Earth's continuous oscillations observed on seismically quiet days // *Geophys. Res. Lett.* 1998. V. 25. P. 1553–1556.
- Tanimoto T., Um J.* Cause of continuous oscillations of the Earth // *J. Geophys. Res.* 1999. 104:28,723//39.
- Tanimoto T.* Continuous Free Oscillations: Atmosphere-Solid Earth Coupling // *Annu. Rev. Earth Planet. Sci.* 2001. 29:563//84.
- Tanimoto T.* The oceanic excitation hypothesis for the continuous oscillations of the Earth // *Geophys. J. Int.* 2005. V. 160. 276//288.
- Vogel M.A., Wong A.K.C.* PFS Clustering Method. *IEEE Trans. Pattern Analysis. Mach. Intell. PAMI-1.* 1978. P. 237–245.