

УДК 550.334

## **СЕЙСМИЧЕСКАЯ КАТАСТРОФА В ЯПОНИИ 11 МАРТА 2011 г.: ДОЛГОСРОЧНЫЙ ПРОГНОЗ ПО НИЗКОЧАСТОТНЫМ МИКРОСЕЙСМАМ**

© 2011 г. А.А. Любушин

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

Излагается технология анализа многомерных временных рядов микросейсмических колебаний на сети широкополосных сейсмических станциях F-net в Японии за более чем 14 лет непрерывных наблюдений, с начала 1997 г. по февраль 2011 г. Анализ мультифрактальных параметров низкочастотного микросейсмического шума позволил еще в 2008 г. выдвинуть гипотезу о приближении Японских островов к крупной сейсмической катастрофе, признаком которой является статистически значимое уменьшение ширины носителя мультифрактального спектра сингулярности. В дальнейшем, по мере поступления новых данных и включения в совместный анализ других статистик микросейсмического фона (таких как логарифм дисперсии шума и индекс линейной предсказуемости), были получены новые результаты, свидетельствующие о все большей синхронизации параметров микросейсмического фона (начало процесса синхронизации было оценено как середина 2002 г.) и, таким образом, о постоянном повышении сейсмической опасности. Анализ свойств кластеризации параметров фона позволил сделать вывод, что именно с середины 2010 г. Японские острова вступают в критически опасную фазу развития сейсмического процесса. Прогноз катастрофы, сначала с оценкой лишь магнитуды (середина 2008 г.), а затем с оценкой времени (середина 2010 г.) был заблаговременно опубликован в серии статей и тезисов на международных конференциях.

*Ключевые слова:* синхронизация, микросейсмический фон, мультифракталы, предвестники землетрясений.

### **Введение**

Низкочастотные микросейсмические колебания являются важным источником информации о процессах в земной коре, несмотря на то, что основная энергия этих колебаний обусловлена процессами в атмосфере и океане: вариациями атмосферного давления и воздействием океанических волн на берег и шельф. Связь низкочастотных микросейсм с периодами 5–500 с с интенсивностью океанических волн подробно исследована в работах [Friedrich *et al.*, 1998; Kobayashi, Nishida, 1998; Tanimoto *et al.*, 1998;

*Tanimoto, Um*, 1999; *Ekstrom*, 2001; *Tanimoto*, 2001, 2005; *Berger et al.*, 2004; *Rhie, Romanowicz*, 2004, 2006; *Kurrle, Widmer-Schmidrig*, 2006; *Stehly et al.*, 2006]. Фактически земная кора является средой распространения энергии от атмосферных и океанических процессов и, поскольку передаточные свойства коры зависят от ее состояния, то и статистические свойства микросейсм отражают изменения свойств литосферы.

Эта простая в своей основе идея использования низкочастотных микросейсмических колебаний для мониторинга литосферы, тем не менее, не так просто реализуема. Основная сложность заключается в сильном влиянии на данные большого числа некоррелированных источников, которые часто распределены диффузно по поверхности Земли. Поэтому исследовать передаточные свойства литосферы, контролируя входные воздействия и отклик, в данном случае не представляется возможным. Кроме того, типичное для традиционных методов анализа данных разделение на «сигнал» и «шум» при обработке микросейсмических колебаний теряет смысл. Априорно к «сигналу» могут быть отнесены лишь приливные вариации амплитуды микросейсм, а также вступления и кода от известных сильных землетрясений. Эти сигналы уже давно используются в геофизике. Все прочие вариации микросейсм «уходят в шум».

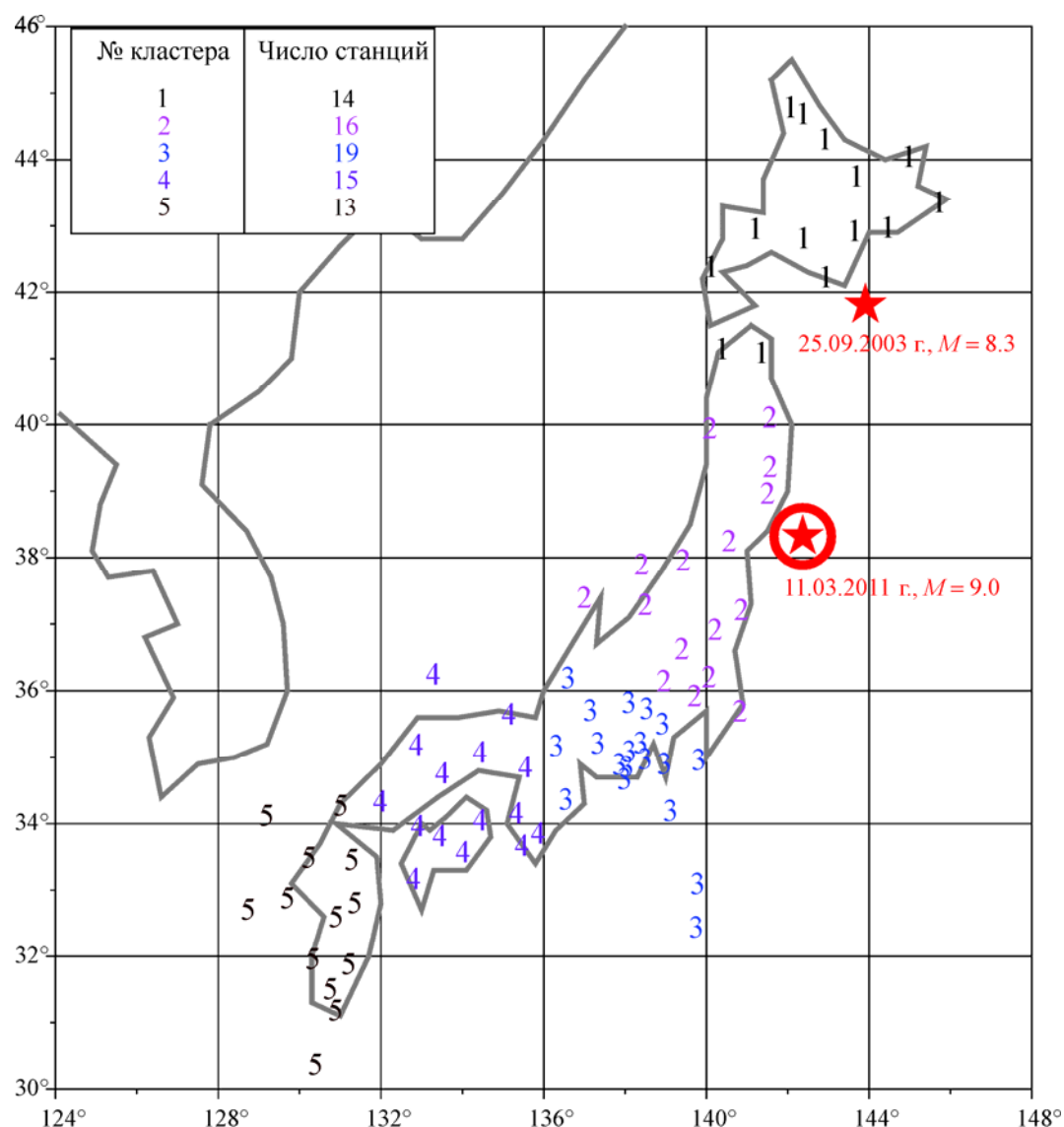
Традиционно используемый в геофизической практике спектральный анализ в исследовании шума является неудачным выбором потому, что шум не содержит ни монохроматических компонент, ни узкополосных сигналов. Поэтому в излагаемом подходе для анализа используется аппарат мультифрактальных спектров сингулярности [*Mandelbrot*, 1982; *Feder*, 1988], позволяющий наиболее детально описать структуру шума. Для анализа геофизических временных рядов оценки спектров сингулярности применялись в работах [*Kantelhardt et al.*, 2002; *Ramírez-Rojas et al.*, 2004; *Currenti et al.*, 2005; *Ida et al.*, 2005; *Telesca et al.*, 2005; Любушин, Соболев, 2006; Любушин, 2007, 2008].

Следует подчеркнуть, что в данном случае имеет место очень редкая ситуация, когда прогноз задолго до случившейся сейсмической катастрофы был опубликован в нескольких статьях [Любушин, 2009, 2010, 2011] (последняя статья была подана в конце апреля 2010 г.), в тезисах международных конференций [Любушин, 2008a,b, 2010a,b,c] и зафиксирован в заявке, поданной 26 апреля 2010 г. в Российский экспертный совет по прогнозу землетрясений и сейсмической опасности.

Далее последовательно излагаются этапы обработки данных и результаты этих работ, которые позволили сформулировать прогноз о подготовке катастрофы и дать оценку времени сильнейшего землетрясения.

### Исходные данные: сеть F-net

Данные широкополосной сейсмической сети F-net свободно доступны в Интернете (<http://www.hinet.bosai.go.jp/fnet>). Общее число станций равно 83. Анализируемые данные – вертикальные компоненты с шагом по времени 1 с (LHZ-записи) – содержат интервалы пропусков, а также неправильных данных (типа постоянных нулевых значений) по причине сбоев в измерительной и регистрирующей аппаратуре. Рассматривались лишь станции, расположенные выше 30° с.ш., что исключает из анализа данные шести уединенных станций, расположенных на удаленных небольших островах. Положение остальных 77 станций представлено на рис. 1 цифрами от 1 до 5, которые являются метками принадлежности станций к пяти пространственным кластерам, вместе с числом станций в каждом кластере. Разбиение станций на пространственные кластеры преследует две цели: пространственное усреднение параметров микросейсм



**Рис. 1.** Положения 77 широкополосных сейсмических станций сети F-net и их разбиение на пять про- странственных кластеров с указанием числа станций в каждом кластере. Звездами обозначены гипоцен- тры землетрясений 25.09.2003 г. ( $M=8.3$ ) и 11.03.2011 г. ( $M=9.0$ )

(посредством взятия медианы по станциям) и обеспечение непрерывности усредненных по кластеру значений в силу достаточно большого числа станций (т.е. всегда работали несколько станций внутри кластеров).

### Параметры спектра сингулярности низкочастотных микросейсм

Ниже кратко изложены технические детали используемых оценок спектра сингулярности [Любушин, Соболев, 2006; Любушин, 2007, 2008]. Важный элемент этой оценки – удаление масштабно-зависимых трендов локальными полиномами, что позволяет избавиться от детерминированных трендов (в нашем случае – приливных и температурных вариаций) и исследовать лишь относительно высокочастотные пульсации ряда, т.е. именно шумовую составляющую.

Пусть  $X(t)$  – случайный процесс. Определим в качестве меры  $\mu_X(t, \delta)$  поведения сигнала  $X(t)$  на интервале  $[t, t + \delta]$  размах:  $\mu_X(t, \delta) = \max_{t \leq s \leq t + \delta} X(s) - \min_{t \leq s \leq t + \delta} X(s)$  и вычислим среднее значение модуля таких мер в степени  $q$ :

$$M(\delta, q) = M\{(\mu_X(t, \delta))^q\}. \quad (1)$$

Случайный процесс называется масштабно-инвариантным, если  $M(\delta, q) \sim |\delta|^{\kappa(q)}$  при  $\delta \rightarrow 0$ , т.е. существует предел

$$\kappa(q) = \lim_{\delta \rightarrow 0} \frac{\ln M(\delta, q)}{\ln |\delta|}. \quad (2)$$

Если зависимость  $\kappa(q)$  линейная:  $\kappa(q) = Hq$ , где  $H = \text{const}$ ,  $0 < H < 1$ , то процесс называется монофрактальным [Taqqi, 1988].

Для вычисления функции  $\kappa(q)$  по конечной выборке из временного ряда  $X(t)$ ,  $t = 1, \dots, N$  можно применить метод DFA [Kantelhardt et al., 2002]. Пусть  $s$  – число отсчетов, ассоциированное с варьируемым масштабом  $\delta_s$ :  $\delta_s = s \cdot \Delta t$ . Разобьем выборку на непересекающиеся малые интервалы длиной  $s$  отсчетов:

$$I_k^{(s)} = \{t : 1 + (k-1) \cdot s \leq t \leq k \cdot s, \quad k = 1, \dots, [N/s]\}; \quad (3)$$

$$y_k^{(s)}(t) = X((k-1)s + t), \quad t = 1, \dots, s \quad (4)$$

участок временного ряда  $X(t)$ , соответствующий интервалу  $I_k^{(s)}$ . Пусть  $p_k^{(s,m)}(t)$  – полином порядка  $m$ , подогнанный методом наименьших квадратов к сигналу  $y_k^{(s)}(t)$ . Рассмотрим отклонения от локального тренда:

$$\Delta y_k^{(s,m)}(t) = y_k^{(s)}(t) - p_k^{(s,m)}(t), \quad t = 1, \dots, s \quad (5)$$

и вычислим значение

$$Z^{(m)}(q, s) = \left( \sum_{k=1}^{[N/s]} (\max_{1 \leq t \leq s} \Delta y_k^{(s,m)}(t) - \min_{1 \leq t \leq s} \Delta y_k^{(s,m)}(t))^q \right)^{1/q} / [N/s], \quad (6)$$

которое будем рассматривать как оценку для  $(M(\delta_s, q))^{1/q}$ . Определим функцию  $h(q)$  как коэффициент линейной регрессии между значениями  $\ln(Z^{(m)}(q, s))$  и  $\ln(s)$ :  $Z^{(m)}(q, s) \sim s^{h(q)}$ . Очевидно, что  $\kappa(q) = qh(q)$ , а для монофрактального процесса  $h(q) = H = \text{const}$ .

Следующим шагом в мультифрактальном анализе [Feder, 1988] после определения функции  $\kappa(q)$  является вычисление спектра сингулярности  $F(\alpha)$ , который можно определить как фрактальную размерность моментов времени  $\tau_\alpha$ , имеющих одно и то же

значение локального показателя Гельдера – Липшица:  $\lambda(t) = \lim_{\delta \rightarrow 0} \frac{\ln(\mu(t, \delta))}{\ln(\delta)}$ , т.е.

$\lambda(\tau_\alpha) = \alpha$ . Стандартный подход состоит в вычислении статистической суммы Гиббса:

$$W(q, s) = \sum_{k=1}^{[N/s]} (\max_{1 \leq t \leq s} \Delta y_k^{(s,m)}(t) - \min_{1 \leq t \leq s} \Delta y_k^{(s,m)}(t))^q \quad (7)$$

и определении показателя массы  $\tau(q)$  из условия  $W(q, s) \sim s^{\tau(q)}$ , после чего спектр  $F(\alpha)$  вычисляется согласно формуле

$$F(\alpha) = \max_q \{ \min(\alpha q - \tau(q)), 0 \}. \quad (8)$$

Сравнивая уравнения (6) и (7), нетрудно заметить, что  $\tau(q) = \kappa(q) - 1 = qh(q) - 1$ . Таким образом,  $F(\alpha) = \max_q \{ \min(q(\alpha - h(q)) + 1, 0) \}$ .

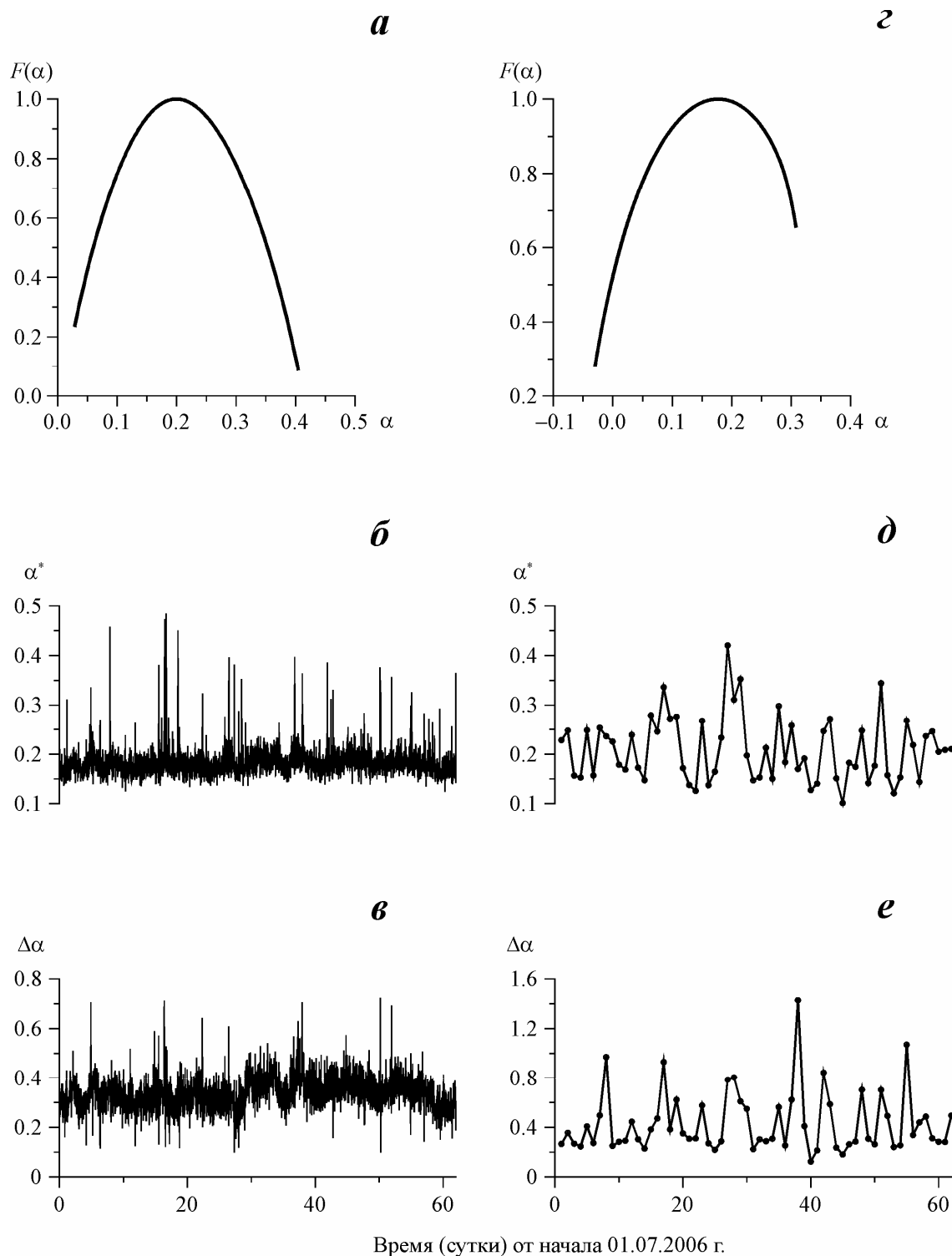
Если оценивать спектр сингулярности  $F(\alpha)$  в скользящем временном окне, то его эволюция дает информацию об изменении структуры шума. В частности, положение и ширина носителя спектра  $F(\alpha)$ , т.е. значения  $\alpha_{\min}$ ,  $\alpha_{\max}$ ,  $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$  и  $\alpha^*$  – то значение, которое доставляет функции  $F(\alpha)$  максимум:  $F(\alpha^*) = \max_{\alpha} F(\alpha)$ , являются характеристиками шума. Величина  $\alpha^*$  называется обобщенным показателем Херста. Для монофрактального сигнала значение  $\Delta\alpha$  должно быть равно нулю, а  $\alpha^* = H$ . Обычно  $F(\alpha^*) = 1$ , но встречаются окна, для которых  $F(\alpha^*) < 1$ . В общем случае величина  $F(\alpha^*)$  равна фрактальной размерности носителя мультифрактальной меры [Feder, 1988].

При вычислении величин  $\Delta\alpha$  и  $\alpha^*$  использовались следующие соображения. Показатель степени  $q$  варьировался внутри интервала:  $q \in [-Q, +Q]$ , где  $Q$  – некоторое достаточно большое число, например  $Q = 10$ . Для каждого значения  $\alpha$  внутри интервала  $\alpha \in [A_{\min}, A_{\max}]$  (где  $A_{\min} = \min_{q \in [-Q, +Q]} \frac{d\tau(q)}{dq}$ ,  $A_{\max} = \max_{q \in [-Q, +Q]} \frac{d\tau(q)}{dq}$ ) вычислялось значение  $\tilde{F}(\alpha) = \min_{q \in [-Q, +Q]} (\alpha q - \tau(q))$ . Если значение  $\alpha$  близко к  $A_{\min}$ , то  $\tilde{F}(\alpha) < 0$  и эта величина не годится в качестве оценки спектра сингулярности, который должен быть неотрицательным. Но начиная с некоторого  $\alpha$ , величина  $\tilde{F}(\alpha)$  становится неотрицательной – это условие выделяет значение  $\alpha_{\min}$ . При дальнейшем увеличении  $\alpha$  значение  $\tilde{F}(\alpha)$  увеличивается, достигает максимума при  $\alpha = \alpha^*$ , а затем начинает уменьшаться и, наконец, достигает некоторого значения  $\alpha_{\max} < A_{\max}$ , такого, что при  $\alpha > \alpha_{\max}$  оно опять становится отрицательным:  $\tilde{F}(\alpha) < 0$ . Таким образом,  $F(\alpha) = \tilde{F}(\alpha)$  при условии, что  $\tilde{F}(\alpha) \geq 0$ , которое определяет интервал носителя спектра сингулярности  $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ . Производная  $\frac{d\tau(q)}{dq}$  находится численно по значениям  $\tau(q)$ ,  $q \in [-Q, +Q]$ , и при ее вычислении точность не имеет особого значения, так как она используется лишь для грубого определения априорного интервала возможных показателей степени  $q$ .

Ниже при анализе низкочастотных микросейсм использовались оценки спектра сингулярности в следующих последовательных непересекающихся временных окнах: для исходных 1-секундных данных – в окне длиной 30 мин (1800 отсчетов) и для 1-минутных данных – в окне длиной 1 сут (1440 отсчетов). Локальные тренды удалялись в первом случае полиномами 4-го порядка, а во втором – полиномами 8-го порядка. Величина  $\alpha^*$  характеризует наиболее типичный и часто встречающийся показатель Гельдера – Липшица, тогда как  $\Delta\alpha$  отражает разнообразие случайного поведения сигнала и, как будет обсуждаться ниже, это является своего рода мерой числа скрытых степеней свободы стохастической системы.

### Вариации ширины носителя спектра сингулярности

На рис. 2 представлены примеры графиков оценок спектра сингулярности  $F(\alpha)$  для одного из окон для 1-секундных (рис. 2,а) и 1-минутных (рис. 2,б) данных для одной из



**Рис. 2.** Графики оценок спектра сингулярности  $F(\alpha)$  и вариаций его параметров ( $\alpha^*$  – обобщенного показателя Херста и  $\Delta\alpha$  – ширины носителя спектра сингулярности) для станции KSK, для двухмесячного фрагмента 01.07–31.08.2006 г.

*a*, *б*, *в* – для 1-секундных данных в последовательных интервалах длиной 30 мин (1800 отсчетов); *z*, *д*, *е* – для 1-минутных данных в последовательных интервалах длиной 1 сут (1440 отсчетов). *a*, *z* приведены для первых интервалов времени длиной 30 мин и 1 сут соответственно для 2-месячного фрагмента 01.07–31.08.2006 г.

станций сети (KSK). Кроме того, ниже каждого графика  $F(\alpha)$  помещены графики вариаций параметров  $\alpha^*$  и  $\Delta\alpha$  для двухмесячного фрагмента 01.07 – 31.08.2006 г. для 1-секундных (рис. 2,б,в) и 1-минутных данных (рис. 2,д,е). Наблюдаемые на рис. 2,б,в резкие выбросы отражают влияние вступлений от различных близких и удаленных землетрясений. Эти выбросы отсутствуют на рис. 2,д,е вследствие двойного усреднения: сначала при переходе от 1 с к 1 мин в шаге по времени, а затем при использовании окна длиной 1 сут вместо 30 мин.

Рассмотрим совокупность оценок параметра  $\Delta\alpha$  для 1-секундных данных (2,в). Для каждого 30-минутного окна, в котором эти оценки были получены, существует определенное число станций, которые обеспечивают эти оценки своими данными. Это число меняется от одного двухмесячного фрагмента к другому и, кроме того, внутри каждого фрагмента. Вычислим для каждого 30-минутного окна медиану значений  $\Delta\alpha$  по всем станциям, которые имеют данные, пригодные для анализа. Медиана является робастной (устойчивой по отношению к выбросам) альтернативой обычного среднего значения.

Последовательность медиан  $\Delta\alpha$  по всем 77 станциям (см. рис. 1) формирует один непрерывный временной ряд общей длительностью 14 лет с шагом по времени 30 мин, который является своего рода интегральной статистической характеристикой поля микросейсм. Точно такие же ряды можно получить и отдельно от каждой из пяти групп станций на рис. 1. Рассмотрим поведение этого ряда при различном сглаживании. В качестве метода сглаживания выбраны гауссовские тренды, которые обладают определенными оптимальными свойствами [Hardle, 1989]. Назовем гауссовским трендом  $\bar{X}(t|H)$  сигнала  $X(t)$  с параметром (радиусом) сглаживания  $H > 0$  следующее значение:

$$\bar{X}(t|H) = \int_{-\infty}^{+\infty} X(t+H\xi) \cdot \psi(\xi) d\xi \bigg/ \int_{-\infty}^{+\infty} \psi(\xi) d\xi, \quad \psi(\xi) = \exp(-\xi^2). \quad (9)$$

Для временных рядов с дискретным временем вычисление величины (9) может быть эффективно реализовано с помощью быстрого преобразования Фурье. Этот способ усреднения применялся в исследовании геофизических временных рядов в работе [Любушин, 2007]. Согласно формуле (9), среднее значение, грубо говоря, относится к интервалу с центром в точке  $t$  радиуса  $H$ . Для сглаживания медиан  $\Delta\alpha$  использовались два значения радиуса  $H = 13$  сут и 0.5 года. Результаты сглаживания для медиан по всем станциям и по пяти кластерам представлены на рис. 3.

Важная особенность поведения сглаженных при  $H = 0.5$  года значения  $\Delta\alpha$  заключается в существенном падении среднего значения, которое началось в начале 2003 г., за полгода до события на о. Хоккайдо, причем достигнутый средний уровень после этого землетрясения уже не восстанавливается до прежнего значения вплоть до катастрофы 11.03.2010 г. Отметим, что эффект падения среднего значения  $\Delta\alpha$  не зависит от выбора группы кластеров. Этот эффект впервые был опубликован для всех станций сети в работе [Lyubushin, 2008a], но его интерпретация была сформулирована несколько позже.

Величина  $\Delta\alpha$  отражает степень разнообразия случайного поведения сигнала, поэтому, косвенно, уменьшение  $\Delta\alpha$  является признаком подавления определенных степеней свободы среды, уменьшения их числа. В то же время можно найти более прямые аналогии с уменьшением числа степеней свободы, отражающиеся в уменьшении  $\Delta\alpha$ . В работах [Pavlov et al., 2003; Ziganshin, Pavlov, 2005] численно исследовались спектры сингулярности для последовательности времен возврата в сечения Пуанкаре для систем двух связанных осцилляторов Ресслера и Лоренца.

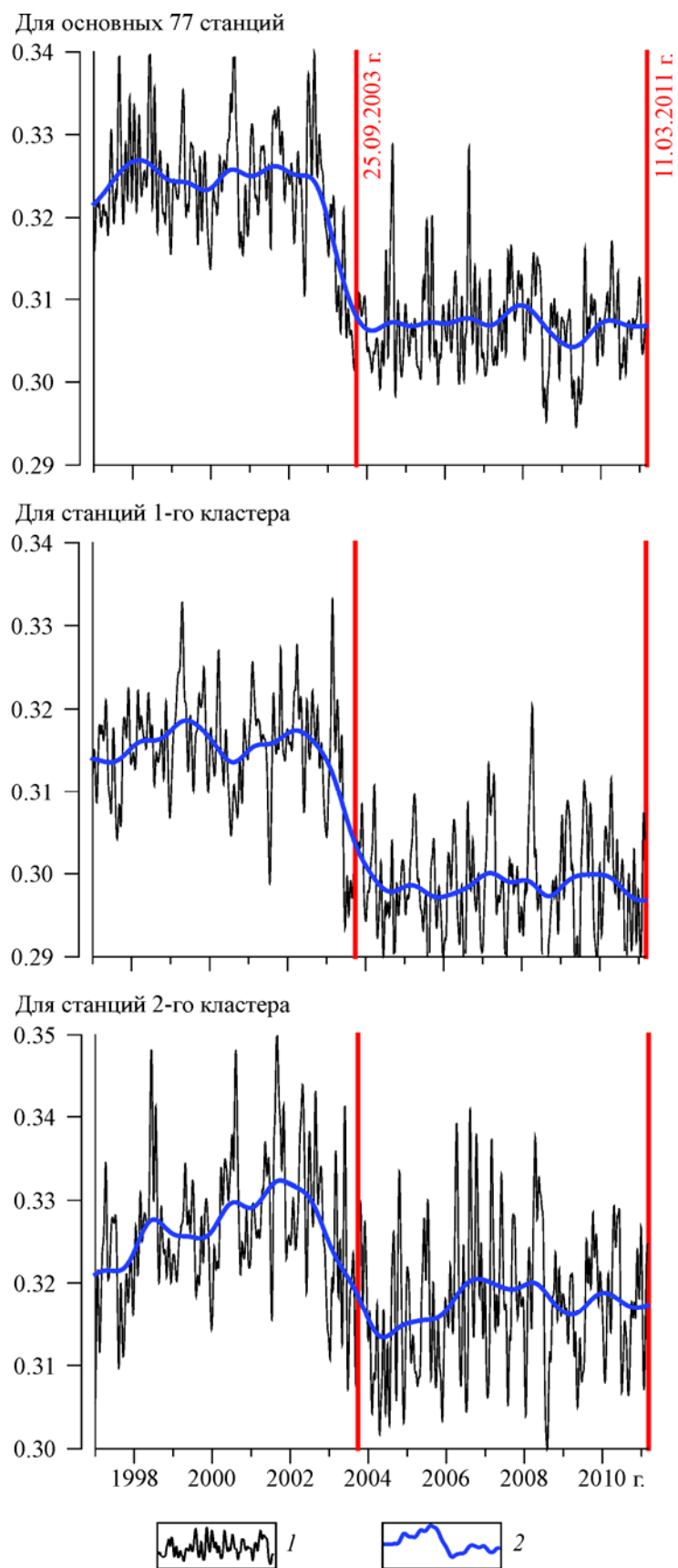
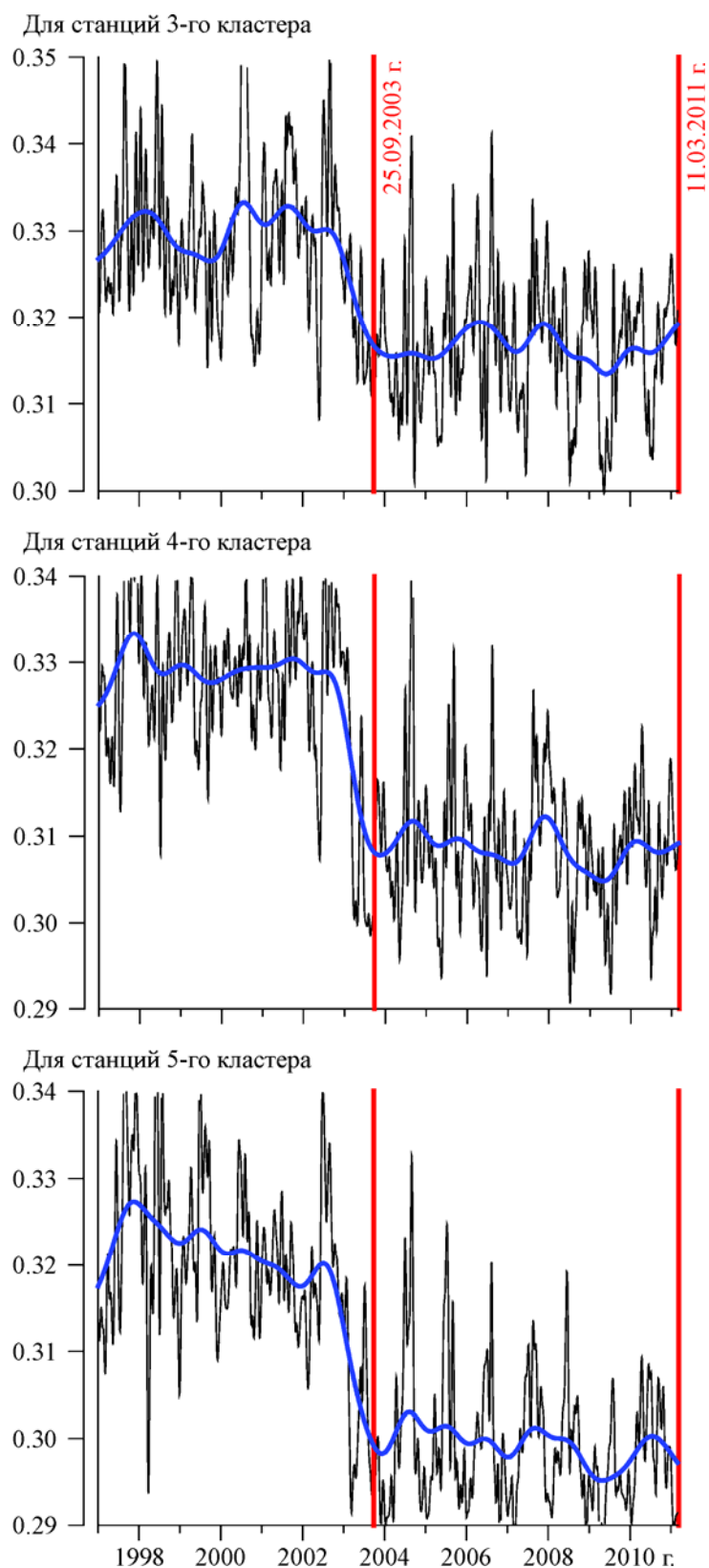


Рис. 3





**Рис. 3.** Сглаженные кривые значений медиан оценок ширины носителя мультифрактального спектра сингулярности  $\Delta\alpha$  в последовательных временных окнах длиной 30 мин для исходных вертикальных сейсмических записей с частотой дискретизации 1 Гц. 1, 2 – гауссовское ядерное сглаживание в окне радиуса 13 сут (1) и 0.5 года (2). На рис. 3–6, 8, 10 вертикальные линии соответствуют моментам землетрясений 25.09.2003 г. ( $M=8.3$ ) и 11.03.2011 г. ( $M=9.0$ )

При наличии достаточно сильной связи эти осцилляторы синхронизируются, что приводит к существенному уменьшению ширины носителя спектра сингулярности  $\Delta\alpha$ . Следовательно, совокупность результатов, представленных на графиках (рис. 3), показывает, что поле микросейсмических колебаний в Японии после события 2003 г. синхронизировалось, и это состояние продолжалось до 11.03.2011 г.

Исходя из известного положения теории катастроф, что синхронизация является одним из свидетельств приближающейся катастрофы [Gilmore, 1981], в работах [Lyubushin, 2008b, 2010; Любушин, 2009] был сформулирован вывод, что событие на о. Хоккайдо, несмотря на свою силу ( $M=8.3$ ), может быть всего лишь форшоком еще более сильного землетрясения, готовящегося в районе Японских островов. Качественная оценка магнитуды будущего толчка как  $M=8.5-9.0$  базируется лишь на здравом смысле: нижняя граница ( $M=8.5$ ) основана на предположении, что магнитуда главного толчка должна быть больше магнитуды форшока, а верхняя ( $M=9.0$ ) выбрана из сообщений, что «больше не бывает».

### Корреляции других статистик микросейсмического фона

При употреблении термина «статистика» далее имеется в виду его значение как «функция от наблюдений» [Cox, Hinkley, 1974]. Ниже рассматриваются 6 таких статистик, каждая из которых определена на последовательных (непересекающихся) временных интервалах заданной длины. Для микросейсмических данных, приведенных к шагу по времени 1 мин, длина этих интервалов составляет 1 сут (1440 отсчетов). Поскольку данные регистрации микросейсмических колебаний содержат пропуски и сбойные участки различной длины, то для получения непрерывных временных рядов вариаций исследуемых статистик с шагом 1 сут вычислялась их медиана по станциям, обеспечивающим нормальную регистрацию в течение текущих суток.

Таким образом, получалось множество из 30 временных рядов (6 параметров, умноженные на 5 кластеров станций) с равномерным шагом по времени 1 сут. Каждый отчет каждого из этих 30 временных рядов представляет собой результат последовательности из трех операций усреднения: 1) непосредственного временного усреднения – перехода от исходных 1-секундных отсчетов к 1-минутным путем взятия средних значений по последовательным 60 точкам; 2) косвенного временного усреднения, реализуемого как оценка той или иной статистики по последовательным 1440 1-минутным значениям микросейсм от каждой станции (если такие данные имеются); 3) пространственное усреднение полученных суточных оценок статистик путем взятия медианы по значениям от работоспособных станций внутри кластера.

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; ее значение тесно связано с факторными 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

$$W_k = \sum_{j=1}^{N^{(k)}} |c_j^{(k)}|^2 / N^{(k)}. \quad (10)$$

Здесь  $c_j^{(k)}$  – коэффициенты ортогонального дискретного вейвлет-разложения выборки временного ряда ( $k = 1, \dots, m$  – номер уровня детальности разложения);  $N^{(k)}$  – число вейвлет-коэффициентов на уровне детальности  $k$ ,  $N^{(k)} \leq 2^{(m-k)}$ . Тогда, аналогично соотношению для скорости роста спектра мощности,  $W_k \sim (s_k)^\beta$ , где  $s_k$  – характерный временной масштаб уровня детальности  $k$ . Поскольку  $s_k = 2^k \div 2^{(k+1)}$ , то отсюда следует, что

$$\log_2(W_k) \sim k^\beta. \quad (11)$$

Таким образом, значение коэффициента наклона прямой, подогнанной методом наименьших квадратов к парам значений  $(\log_2(W_k), k)$ , дает оценку для величины  $\beta$ . Параметр  $\beta$  оценивался в последовательных временных окнах длиной 1440 отсчетов (1 сут). Ранее такой способ вычисления спектральных экспонент при анализе микро-сейсмического шума использовался в работе [Любушин, 2008]. Для того чтобы исключить влияние приливных вариаций, в каждом окне устранялся полиномиальный тренд 8-го порядка и для остатка вычислялся вейвлетный спектр мощности (11) и десятичный логарифм дисперсии  $\lg(Var)$ . При этом выбирался оптимальный ортогональный вейвлет Добеши с числом обнуляемых моментов от двух до 10, реализующий минимум энтропии распределения квадратов вейвлет-коэффициентов для первых 7 уровней детальности вейвлет-разложения (масштабы, или «периоды», от 2 до 256 мин при шаге по времени 1 мин). 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.

Выше описаны параметры, которые вычисляются для суточных временных интервалов на каждой станции, которая обеспечивает нормальную регистрацию для текущих суток. Далее применяется операция усреднения по всем станциям сети и по станциям внутри кластеров. Значения медиан будем обозначать  $\bar{\beta}_r(s)$ ,  $\overline{\lg(Var)}_r(s)$  и  $\bar{\xi}_r(s)$ . В этих обозначениях верхняя черта означает пространственное усреднение (взятие медианы внутри кластеров), нижний индекс  $r = 1, \dots, 5$  показывает, какому из пяти кластеров принадлежит то или иное значение медианы, а аргумент  $s$  является целочисленным индексом, нумерующим последовательные сутки либо от общего начала наблюдений (начало 1997 г.), либо от начала временного окна, в отличие от временного индекса  $t$ , нумерующего последовательные 1-минутные отсчеты сейсмических записей внутри текущих суток. Заметим, что после вычисления медианы по всем станциям показатель гладкости  $\bar{\xi}_r(s)$  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 ана-

логичной формуле:  $V_{AR} = \sum_{t=n+1}^N \varepsilon_{AR}^2(t) / (N-n)$ . Здесь  $\varepsilon_{AR}(t+1) = x(t+1) - \hat{x}_{AR}(t+1)$  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. То есть, рассматривается модель:

$$x(t) + a_1 \cdot x(t-1) + a_2 \cdot x(t-2) = e(t) + d, \quad (12)$$

в которой вектор  $c = (a_1, a_2, d)^T$  является вектором неизвестных параметров, определяемых по скользящему «малому» окну длиной  $n$  отсчетов из условия минимума суммы квадратов невязок  $e(t)$ . Введем вектор  $Y(t) = (-x(t-1), -x(t-2), 1)^T$ . Тогда авторегрессионную модель можно коротко записать в виде  $x(t) = c^T \cdot Y(t) + e(t)$ .

Для вычисления прогноза на шаг вперед  $\hat{x}_{AR}(t+1)$  найдем вектор  $c$  из условия минимума суммы квадратов невязок  $e(t)$  по предыдущим  $n$  отсчетам:  $\sum_{\lambda=t-n+3}^t e^2(\lambda) = \sum_{\lambda=t-n+3}^t (x(\lambda) - c^T Y(\lambda))^2 \rightarrow \min_c$ , откуда нетрудно получить формулы для оценки метода наименьших квадратов вектора параметров и прогноза на 1 шаг вперед:

$$\hat{c}(t) = A^{-1}(t) \cdot R(t), \quad A(t) = \sum_{\lambda=t-n+3}^t Y(\lambda) \cdot Y^T(\lambda), \quad R(t) = \sum_{\lambda=t-n+3}^t x(\lambda) \cdot Y(\lambda), \quad (13)$$

$$\hat{x}_{AR}(t+1) = x(t+1) - \hat{x}^T(t) \cdot Y(t).$$

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  $\rho > 0$ .

Ниже всюду при вычислении индекса линейной предсказуемости  $\rho$  для 1-минутных данных оценки производились в последовательных «длинных» временных окна длиной  $N = 1440$  отсчетов (1 сут) и при «коротком» окне длиной  $n = 60$  отсчетов (1 ч).

Помимо двух параметров мультифрактального спектра сингулярности  $\Delta\alpha$  и  $\alpha^*$ , введенных выше, будет анализироваться также величина  $\gamma = \alpha^* - (\alpha_{\min} + \alpha_{\max})/2$ , которая характеризует степень асимметрии спектра сингулярности.

Аналогично обозначениям  $\bar{\beta}_r(s)$ ,  $\overline{\lg(Var)}_r(s)$  и  $\bar{\xi}_r(s)$ , введенным выше, будем обозначать через  $\bar{\rho}_r(s)$ ,  $\bar{\alpha}_r^*(s)$ ,  $\bar{\Delta\alpha}_r(s)$  и  $\bar{\gamma}_r(s)$  значения медиан соответствующих статистик, где верхняя черта символизирует операцию взятия медианы, аргумент  $s$  нумерует последовательные сутки, а индекс  $r = 1, \dots, 5$  нумерует кластеры станций.

*Робастная множественная мера корреляции  $\kappa$* . Кратко рассмотрим процедуру вычисления меры, описывающей множественную (совокупную) корреляцию компонент

многомерного временного ряда. Она основана на использовании канонических корреляций [Hotelling, 1936; Rao, 1965], но отличается от классического подхода использованием робастных (устойчивых к выбросам) оценок. Подробно она изложена в работе [Любушин, 2007]. Пусть  $u_r(s)$ ,  $r = 1, \dots, Q$  –  $Q$ -мерный временной ряд,  $s = 1, \dots, L$  – дискретное время.

В нашем случае  $Q = 5$  (число кластеров станций);  $u_r(s)$  – медианные суточные значения  $\bar{\beta}_r(s)$ ,  $\overline{\lg(Var)}_r(s)$ ,  $\bar{\rho}_r(s)$ ,  $\bar{\alpha}_r^*(s)$ ,  $\bar{\Delta\alpha}_r(s)$ ,  $\bar{\gamma}_r(s)$ ;  $s$  – индекс, нумерующий последовательные сутки;  $L$  – общее число последовательных суток, анализируемых одновременно, которое далее будет равно либо 91 (четверть года), либо 365 (1 год).

Выделим компоненту с номером  $p$  и рассмотрим линейную регрессионную модель влияния всех прочих компонент на выбранную  $u_p$ :

$$u_p(s) = w_p(s) + \varepsilon_p(s), \quad w_p(s) = \sum_{r=1, r \neq p}^Q \gamma_r^{(p)} \cdot u_r(s). \quad (14)$$

Регрессионные коэффициенты  $\gamma_r^{(p)}$  найдем из условия минимума суммы модулей  $\sum_{s=1}^L |\varepsilon_p(s)|$ , после чего вычислим оценку коэффициента корреляции  $\mu_p$  между выделенной компонентой  $u_p(s)$  и получившимся регрессионным вкладом  $w_p(s)$  согласно формуле робастной оценки коэффициента корреляции [Huber, 1981]:

$$\mu_p = (S(\hat{z}_p^2) - S(\tilde{z}_p^2)) / (S(\hat{z}_p^2) + S(\tilde{z}_p^2)), \quad (15)$$

где  $\hat{z}_p(t) = a_p u_p(s) + b_p w_p(s)$ ,  $\tilde{z}_p(s) = a_p u_p(s) - b_p w_p(s)$ ,  $a_p = 1/S(u_p)$ ,  $b_p = 1/S(w_p)$ ,  $S(u_p) = \text{med} |u_p - \text{med}(u_p)|$ . Здесь  $\text{med}(u_p)$  означает медиану выборки  $u_p(s)$ ,  $s = 1, \dots, L$ , а  $S(u_p)$ , таким образом, является абсолютным медианным отклонением выборки  $u_p(s)$ . Необходимость использования робастных оценок, т.е. минимизации суммы модулей регрессионных остатков  $\varepsilon_p(s)$ , а не их квадратов (что вычислительно много проще), а также использование формулы (15) для оценки коэффициента корреляции связаны с необходимостью обеспечения устойчивости полученных оценок  $\mu_p$  по отношению к большим выбросам, обусловленным наличием вступлений от близких малых и средних, а также удаленных сильных землетрясений.

Величину  $\mu_p$  назовем робастной канонической корреляцией [Hotelling, 1936; Rao, 1965]  $p$ -й компоненты по отношению ко всем прочим. Проведем эти вычисления последовательно для всех  $p = 1, \dots, Q$ . Затем определим величину

$$\kappa = \prod_{p=1}^Q |\mu_p|, \quad (16)$$

которую назовем робастной множественной мерой корреляции многомерного временного ряда.

Очевидно, что  $0 \leq \kappa \leq 1$ , и чем ближе значение (16) к единице, тем сильнее общая связь вариаций компонент многомерного временного ряда  $u_p(s)$  друг с другом. Вычислив величины (16) не для всей выборки, а в скользящем временном окне заданной дли-

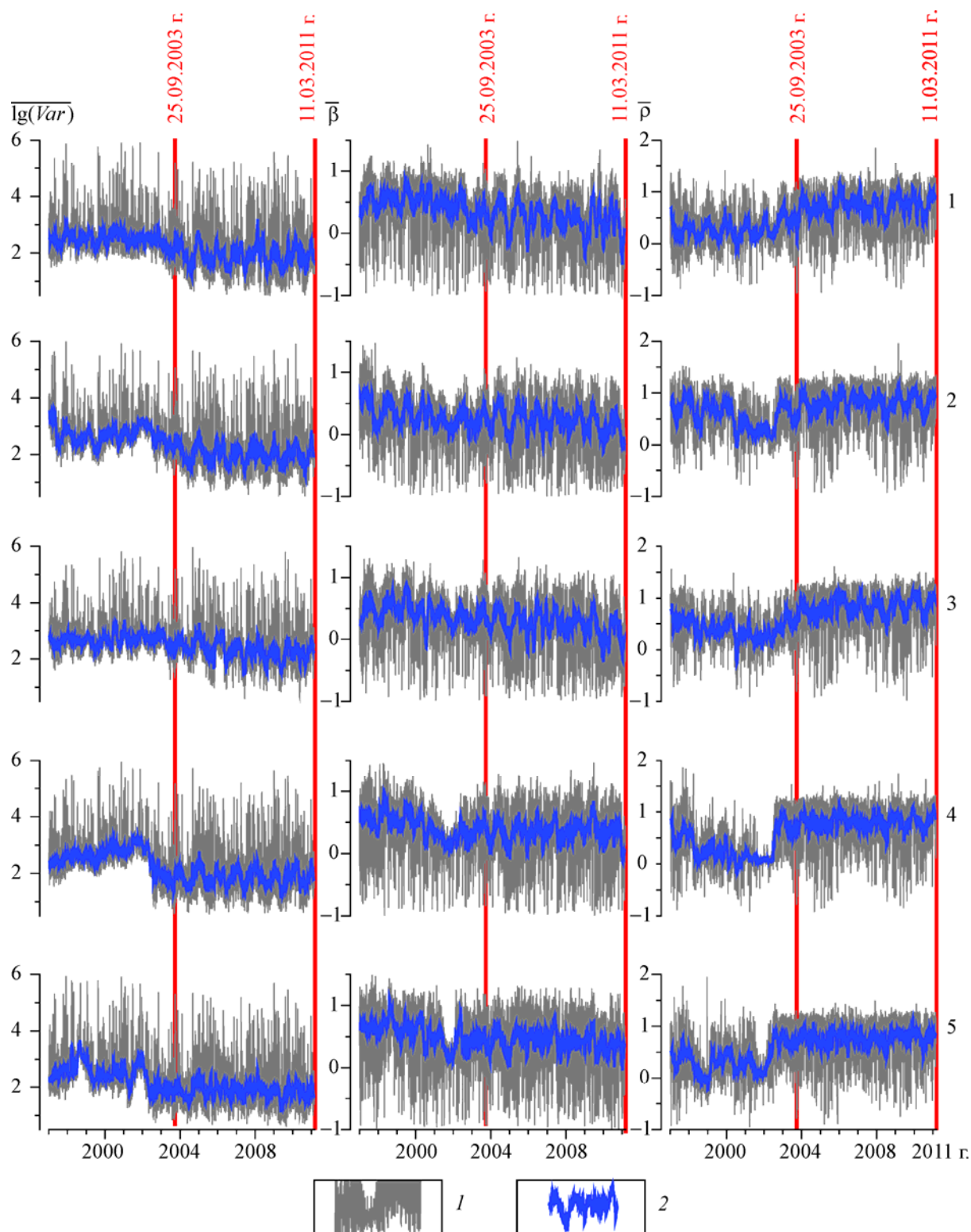
ны  $L$  отсчетов с временной меткой правого конца скользящего окна  $\tau$ , получим эволюцию множественной корреляции (16) в виде зависимости  $\kappa(\tau | L)$ . По числу используемых статистик суточных временных фрагментов получим 6 таких зависимостей, при записи которых для простоты второй аргумент  $L$  будем опускать:  $\kappa_\beta(\tau)$ ,  $\kappa_{\lg(Var)}(\tau)$ ,  $\kappa_\rho(\tau)$ ,  $\kappa_{\alpha^*}(\tau)$ ,  $\kappa_{\Delta\alpha}(\tau)$  и  $\kappa_\gamma(\tau)$ . Заметим также, что величины (15) и (16) вычислялись не для исходных, а для приращений медианных значений с целью обеспечения большей стационарности анализируемых выборок.

Графики медианных величин используемых шести статистик представлены на рис. 4 и 5. Они характеризуются довольно случайным и нерегулярным характером поведения. Поэтому на рис. 4, 5 линиями 2 изображены графики временного усреднения исходных медианных значений в скользящем окне радиусом 14 сут. Эти средние значения часто демонстрируют сильно выраженные сезонные (годовые изменения).

Следует отметить, что, хотя годовые вариации величин, представленных на рис. 4, довольно синхронны, общий низкочастотный тренд уменьшения медианных значений  $\lg(Var)$  сопровождается такими же общим трендом увеличения индекса линейной предсказуемости  $\bar{\rho}$ . Особенно ярко эта тенденция проявляется для 4-го и 5-го кластеров в интервале конец 2002 г. – начало 2003 г. При этом величины  $\bar{\rho}_4$  и  $\bar{\rho}_5$  претерпевают значительный и довольно резкий скачок, приходящийся на июль 2002 г. По всей видимости, этот скачок можно считать началом роста синхронизации. По смыслу своего построения рост индекса линейной предсказуемости означает увеличение временной коррелированности микросейсмических колебаний. Для трех северных кластеров медиана  $\bar{\rho}$  растет довольно плавно, тогда как для двух южных она растет скачкообразно. Медианы спектральных экспонент  $\bar{\beta}$  часто имеют максимумы, приходящиеся на летние месяцы, и минимумы – на зимние, т.е. летом микросейсмсы носят более низкочастотный характер, чем зимой.

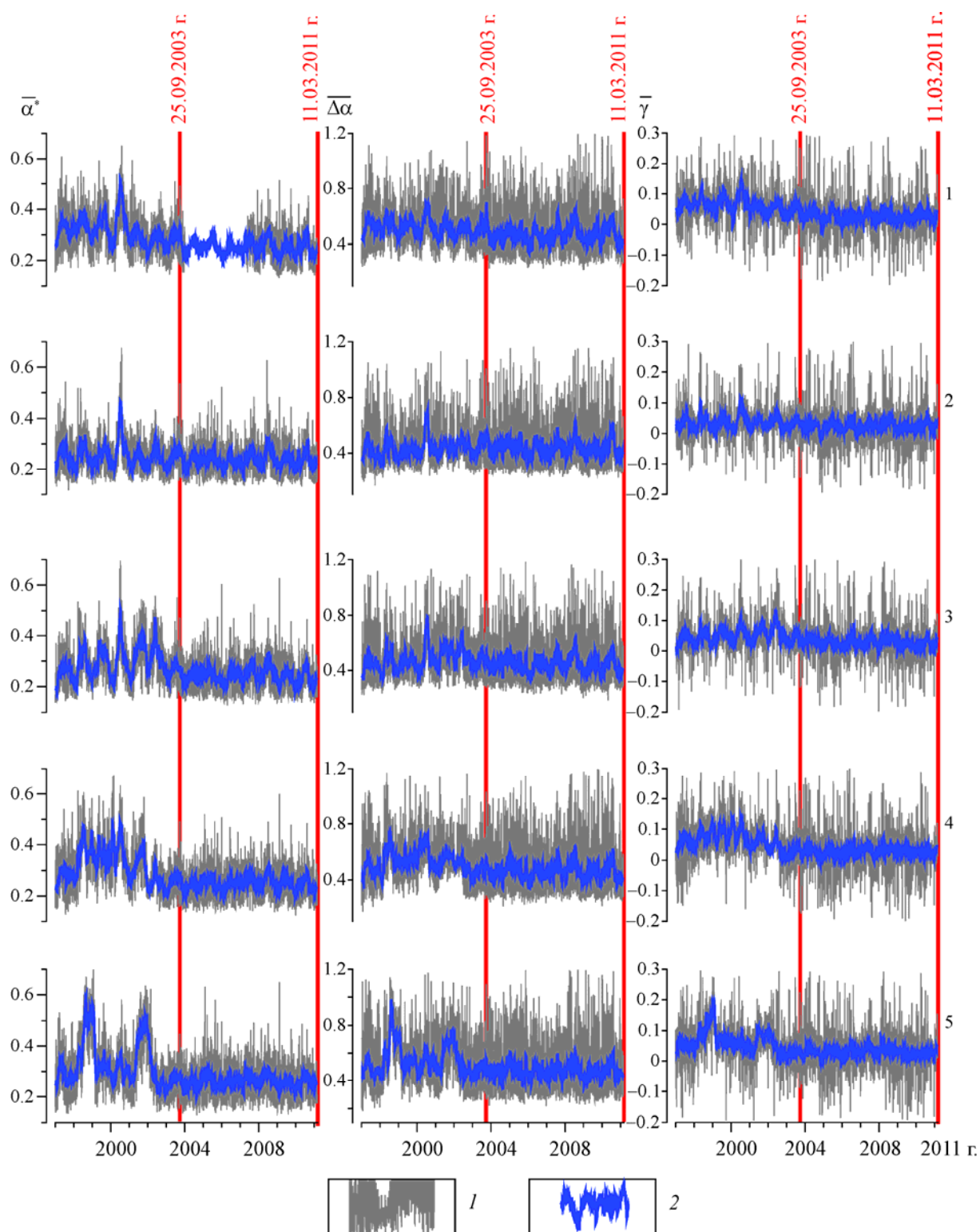
На рис. 6 представлены графики изменений робастной множественной меры корреляции  $\kappa$  для всех статистик. Мера (16) вычислена для двух временных окон – 91 день и 365 дней. Заметим, что использование годового окна для вычисления меры корреляции эквивалентно усреднению сезонных воздействий циклонов, штормов и ураганов как основных генераторов низкочастотных микросейсм, что делает оценку весьма гладкой и устойчивой – это заметно при сравнении линий 1 и 2 на рис. 6. Главная же особенность рассматриваемых графиков – рост степени множественного коэффициента корреляции перед событием 25.09.2003 г. и его стабилизация на новом более высоком уровне вплоть до 11.03.2011 г.

Проведенный анализ показал [Любушин, 2010] что использование больших длин временного окна (1 год) делает выделение эффектов синхронизации очень устойчивым и статистически значимым. Таким образом, независимо (за счет использования не только параметров спектров сингулярности) подтвержден основной вывод работы [Любушин, 2009] о синхронизации параметров поля низкочастотных микросейсм на Японских островах в связи с землетрясением на о. Хоккайдо 25.09.2003 г., причем использование новой статистики – индекса линейной предсказуемости позволило довольно точно указать время начала систематического роста синхронизации – июль 2002 г.



**Рис. 4.** Вариации медиан статистик микросейсмического шума от разных групп станций

1 – графики поведения медиан  $\overline{\lg(Var)}$  десятичных логарифмов дисперсии, спектральных экспонент  $\bar{\beta}$  и индексов линейной предсказуемости  $\bar{\rho}$ , вычисленных для станций внутри пяти кластеров, положения которых представлены на рис.1 (номера кластеров проставлены справа) для 1-минутных сейсмических записей; 2 – их усреднения в скользящем временном окне радиусом 14 сут. Параметры вычислялись внутри последовательных временных окон длиной 1440 отсчетов (1 сут)



**Рис. 5.** Вариации медиан параметров  $\bar{\alpha}^*$ ,  $\overline{\Delta\alpha}$ ,  $\bar{\gamma}$  микросейсмического шума от разных групп станций

1 – графики поведения медиан параметров  $\bar{\alpha}^*$ ,  $\overline{\Delta\alpha}$ ,  $\bar{\gamma}$  мультифрактальных спектров сингулярности, вычисленных для станций внутри пяти кластеров, положения которых представлены на рис. 1 (номера кластеров проставлены справа от графиков) для 1-минутных сейсмических записей; 2 – их усреднения в скользящем временном окне радиусом 14 сут. Спектры сингулярности вычислялись внутри последовательных временных окон длиной 1440 отсчетов (1 сут)



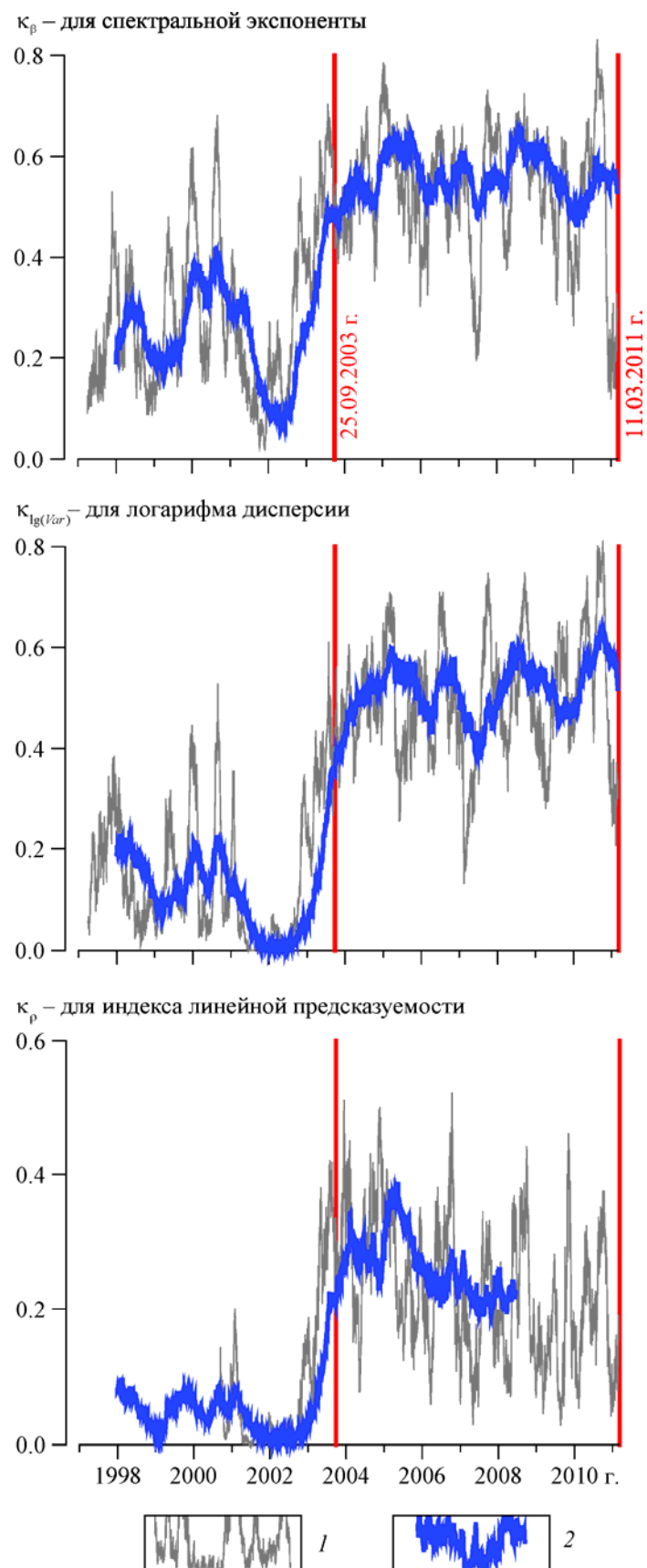
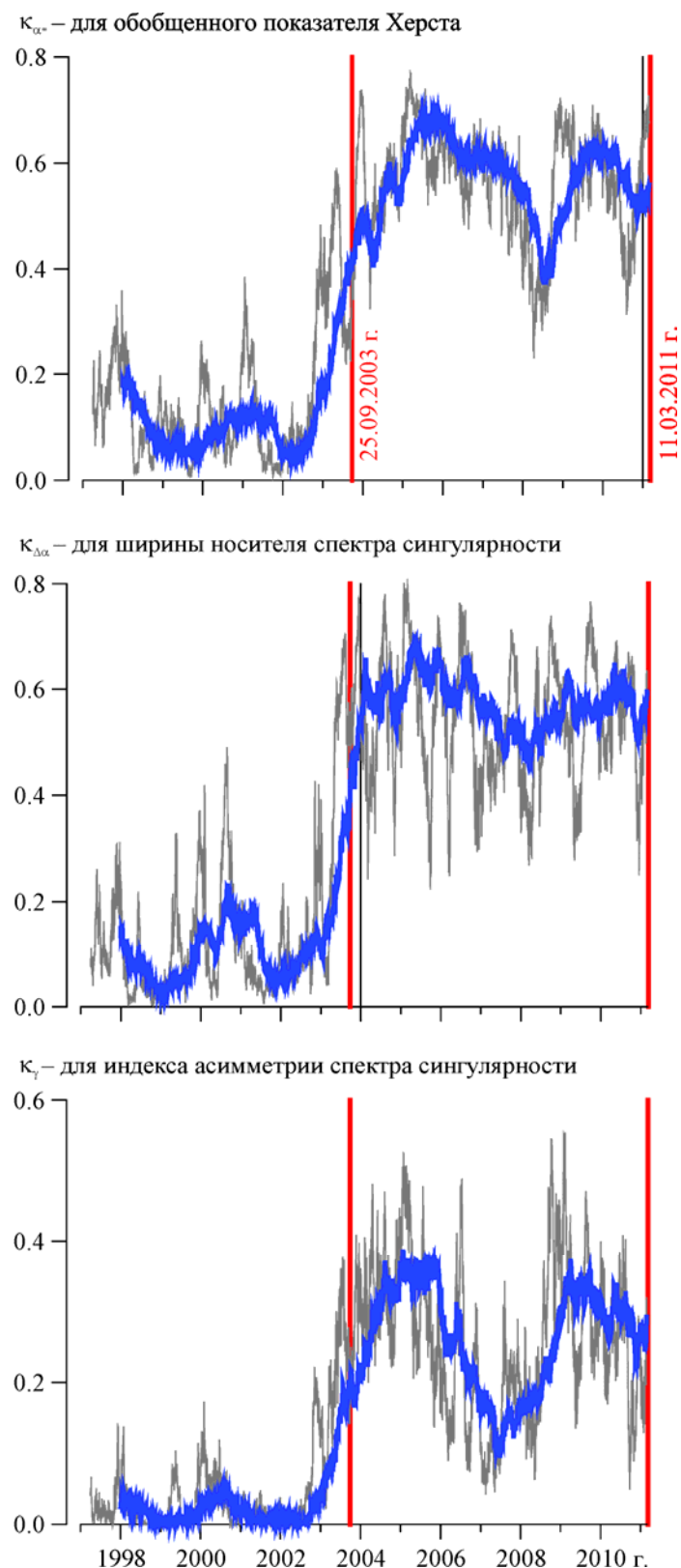


Рис. 6



**Рис. 6.** Графики робастной множественной меры корреляции  $\kappa$ , оцененной для приращений медиан значений статистик  $\bar{\beta}$ ,  $\overline{\lg(Var)}$ ,  $\bar{\rho}$ ,  $\bar{\alpha}^*$ ,  $\bar{\Delta\alpha}$  и  $\bar{\gamma}$  (нижний индекс у величин  $\kappa$ ), вычисленных для сейсмических станций внутри пяти пространственных кластеров (см. рис.1) для 1-минутных данных в последовательных сутках. 1, 2 – оценки  $\kappa$ : 1 – в окне длиной 0.25 года (91 отсчет), 2 – в окне длиной 1 год (365 отсчетов)

## Кластерный анализ параметров низкочастотного микросейсмического фона

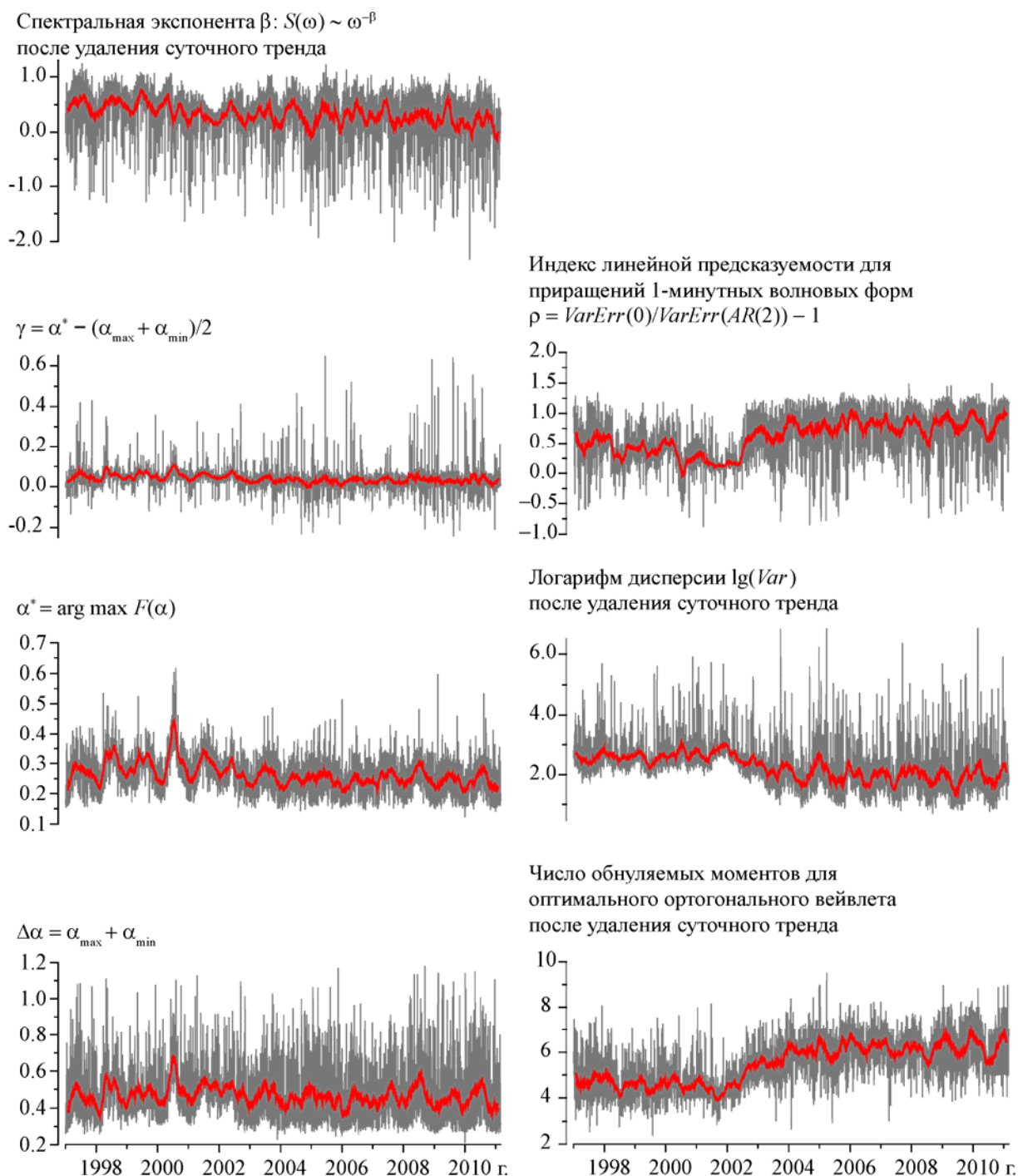
Оценки множественных мер корреляции в скользящих временных окнах, использованные в работах [Любушин, 2009, 2010] и изложенные выше, не позволили выделить других значимых аномалий поведения параметров поля микросейсм за время наблюдений 1997–2010 гг., помимо сравнительно быстрого перехода на уровень высокой синхронизации во второй половине 2002 г. – начале 2003 г. Далее 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, изложенного в работе [Любушин, 2011], позволило ввести новый параметр – «кластерную экспоненту»  $\mu$  и выделить аномальный участок поведения фона 2007–2010 гг., характеризующийся положительным трендом значений  $\mu$ , который аналогичен более короткому тренду перед событием на о. 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.

На рис. 7 представлены графики изменения медиан всех семи рассматриваемых статистик:  $\bar{\beta}$ ,  $\bar{\xi}$ ,  $\overline{\lg(Var)}$ ,  $\bar{\rho}$ ,  $\bar{\alpha}^*$ ,  $\overline{\Delta\alpha}$ ,  $\bar{\gamma}$  вместе с их средними значениями в скользящем окне длиной 57 сут. Такое значение выбрано как кратное двум длинам лунного месяца (28 сут – период модуляции многих геофизических процессов), а дополнительная единица возникла из необходимости иметь нечетное значение длины окна скользящего усреднения.

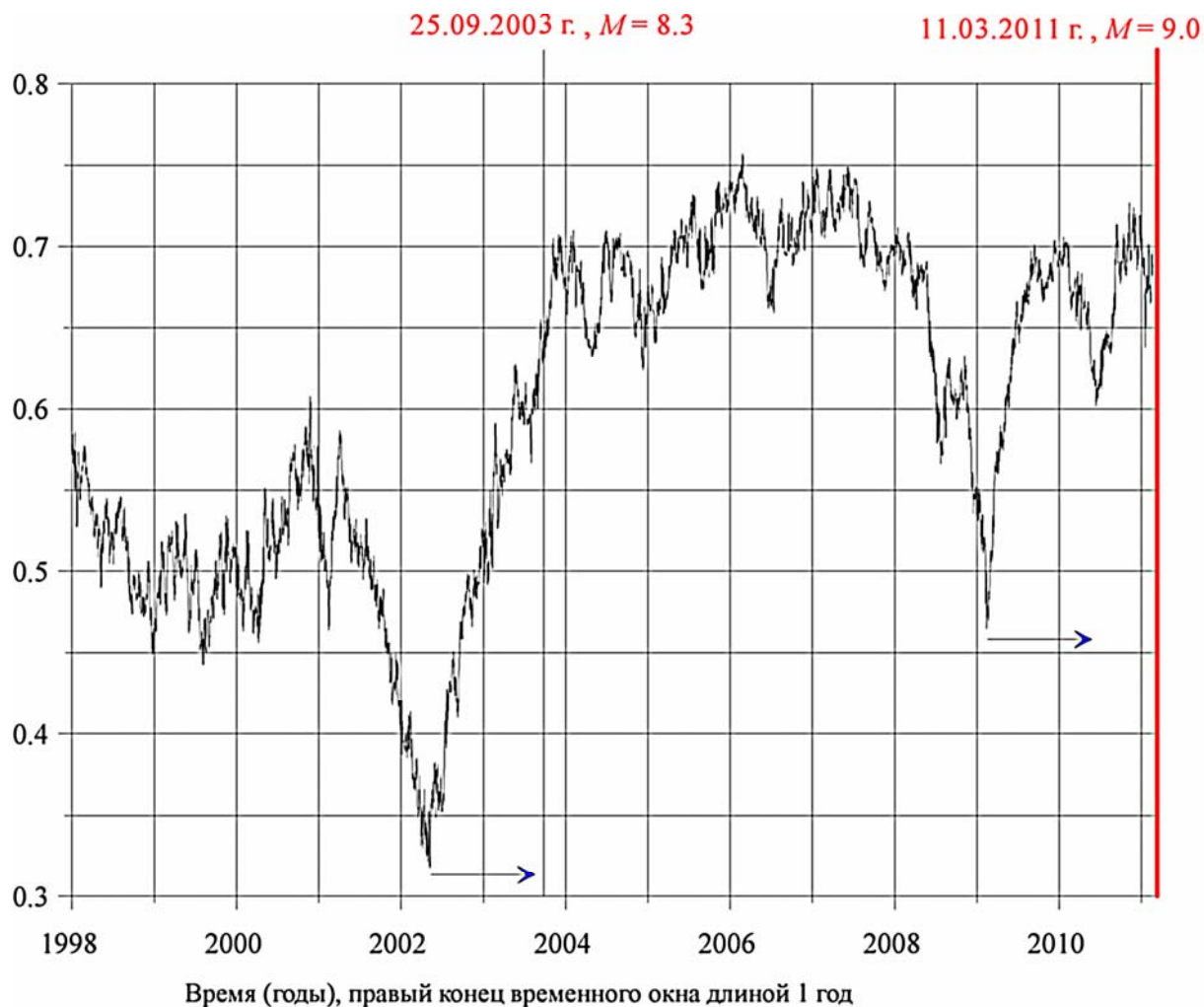
На рис. 8 представлены графики изменения коэффициента корреляции (16) между двумя параметрами  $\bar{\alpha}^*(s)$  и  $\overline{\Delta\alpha}(s)$ , вычисленного в скользящем временном окне длиной 1 год. Рисунок 8 примечателен тем, что он содержит две яркие аномалии поведения коэффициента корреляции – острые минимумы в 2002 и 2009 гг. Поскольку после первой аномалии 2002 г. последовало крупное землетрясение 25.09.2003 г., то логично предположить, что второй острый минимум коэффициента корреляции может быть также предвестником будущего сильного события во второй половине 2010 г., причем это событие может быть большим по энергии. Этот график позволил сделать вывод [Lyubushin, 2010b,c; Любушин, 2011], что начиная с середины 2010 г. следует ждать сильного события на Японских островах с  $M=8.5-9.0$ .

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

1) каждая компонента вектора  $\vec{\psi}$  подвергалась операции нормализации и виндизации: 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;



**Рис. 7.** Зависимости статистик, вычисляемых как медианы по всем сейсмическим станциям сети F-net в последовательных временных окнах длиной 1 сут:  $\beta$  – 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 сут. Утолщенные линии – графики скользящих средних в окне длиной 57 сут



**Рис. 8.** График робастного коэффициента корреляции между вариациями величин  $\bar{\alpha}^*$  и  $\overline{\Delta\alpha}$  (см. рис. 6) в скользящем временном окне 365 сут (1 год) в зависимости от положения правого конца временного окна. Стрелками выделены временные интервалы одинаковой длины от момента первого глубоко «заостренного» минимума коэффициента корреляции до момента землетрясения 25.03.2009 г. и от момента второго минимума, имеющего тот же характер, до июля 2010 г., который оценивается как начало опасного временного интервала ожидания катастрофы

2) для полученного облака нормализованных 7-мерных векторов вычислялись первые четыре главные компоненты как проекции на собственные векторы ковариационной матрицы внутри текущего окна, соответствующие четырем 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. Разбиение производилось последовательно иерархической кластеризацией с использованием метрики «дальний сосед» (что обеспечивает получение компактных и «округлых» кластеров) и последующей итерации метода К-средних [Duda, Hart, 1973].

Пусть  $N$  – общее число 4-мерных векторов  $\vec{\zeta}$  главных компонент (от нормализованных 7-мерных векторов) в текущем временном окне, а  $\vec{\zeta}_0$  – вектор общего центра масс облака главных компонент (вследствие предварительных операций нормализации и винзоризации  $\vec{\zeta}_0 = 0$ );  $\vec{\zeta}_k, k = 1, \dots, q$  – векторы центров масс кластеров, а  $n_k$  – число элементов в каждом кластере,  $\sum_{k=1}^q n_k = N$ . Разбиение облака из  $N$  векторов на заданное число  $q$  кластеров оценивается следующими величинами:

$$\sigma_0^2(q) = \frac{\sum_{k=1}^q \sum_{\vec{\zeta} \in \Gamma_k} |\vec{\zeta} - \vec{\zeta}_k|^2}{N - q} \quad (17)$$

мера внутрикластерной компактности;

$$\sigma_1^2(q) = \frac{\sum_{k=1}^q v_k \cdot |\vec{\zeta}_k - \vec{\zeta}_0|^2}{q - 1}, \quad v_k = \frac{n_k}{N} \quad (18)$$

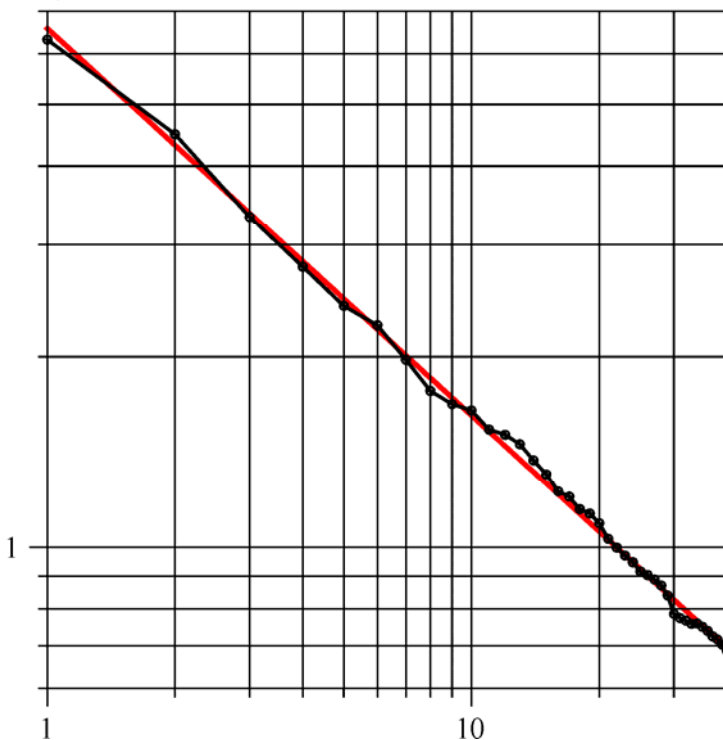
взвешенная мера расхождения между центрами кластеров;

$$PFS(q) = \frac{\sigma_1^2(q)}{\sigma_0^2(q)} \quad (19)$$

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  $q = 1$ :  $\sigma_0^2(1) = \sum_{\vec{\zeta}} |\vec{\zeta} - \vec{\zeta}_0|^2 / (N - 1)$ . Величина  $\sigma_0^2(q)$  монотонно возрастает с убыванием  $q$  и зависимость  $\lg(\sigma_0^2(q))$  от  $\lg(q)$  близка к линейной, т.е.  $\sigma_0^2(q) \sim q^{-\mu}$ .

$$\sigma_0^2(q) \sim q^{-\mu}$$



Этот факт иллюстрируется на рис. 9. 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  $\lg(q)$ .

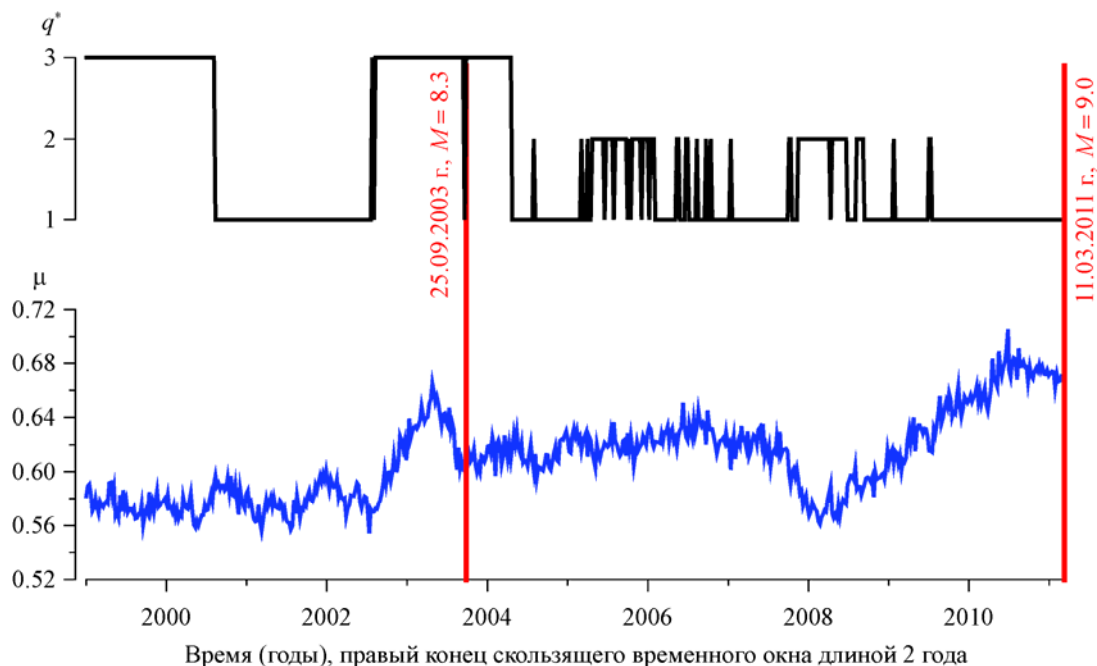
Величина (19) характеризует качество разбиения на заданное число кластеров: чем больше  $PFS(q)$ , тем разбиение лучше.

**Рис. 9.** Зависимость функционала внутрикластерной компактности  $\sigma_0^2(q)$  от пробного числа кластеров  $q$ . Прямая линия – график наилучшего линейного приближения зависимости  $\lg(\sigma_0^2(q))$  от  $\lg(q)$

Для хорошего разбиения величине  $\sigma_0^2(q)$  внутрикластерной компактности желательно 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 возможностью. 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  $q \geq 2$ . На основе изложенного, оптимальное число  $q^*$  кластеров определим следующим образом:

$$\begin{aligned} &\text{Пусть } q_0 = \arg \max_{2 \leq q \leq 40} PFS(q), \\ &\text{если } q_0 > 2, \text{ то } q^* = q_0, \\ &\text{иначе:} \\ &\text{если } \delta \lg(\sigma_0^2(1)) \leq \max_{2 \leq q \leq 40} \delta \lg(\sigma_0^2(q)), \text{ то } q^* = 1; \\ &\text{иначе } q^* = 2 \end{aligned} \quad (20)$$

На рис. 10 представлены результаты кластеризации четырех главных компонент 7 ежедневных медианных характеристик микросейсмического фона сети F-net для 14 лет



**Рис. 10.** Результаты кластерного анализа четырех первых главных компонент в скользящем временном окне длиной 2 года со смещением 7 сут в зависимости от положения правого конца временного окна:  $q^*$  – оптимальное число кластеров,  $\mu$  – значение кластерной экспоненты

наблюдений (1997–2010 гг.) при оценке в скользящем временном окне длиной два года ( $N = 730$ ) со смещением 7 сут.

Следует отметить изменение характера переключения между числом оптимальных кластеров после 2004 г. на рис. 10 – режим переключений стал более хаотичным, причем вариант трех 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 будущим сильным толчком 11.03.2011 г.

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) снова повторяется с положения правого конца двухгодового окна в начале 2008 г., причем положительный тренд  $\mu$  имеет существенно более длительный характер. Продолжительный возрастающий линейный тренд кластерной экспоненты, аналогичный более короткому линейному тренду перед событием 25.09.2003 г., точно так же перевалил через локальный минимум и пошел вниз за полгода до мощного землетрясения 11.03.2011 г.

### Выводы

Анализ низкочастотного микросейсмического шума позволил выделить долговременные предвестники сейсмической катастрофы в Японии 11.03.2011 г. Следует отметить, что традиционные методы прогноза землетрясений, основанные на анализе особенностей потока сейсмических событий малой и средней силы перед предыдущими сильными землетрясениями, пропустили подряд две катастрофы магнитудой  $M=9$ : на Суматре 26.12.2004 г. и в Японии 11.03.2011 г. Очевидно, что это связано с отсутствием достаточной статистики для обучения алгоритмов прогноза сильнейших землетрясений в силу малого их количества. Это обстоятельство подчеркивает важность создания и поддержания достаточно густых сетей непрерывного геофизического мониторинга наподобие сети F-net в Японии и обеспечение свободного доступа к их данным через Интернет.

### Благодарности

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 09-05-00134).

### Литература

- 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.
- Любушин А.А. Кластерный анализ свойств низкочастотного микросейсмического шума // Физика Земли. 2011. № 6 (в печати).
- Любушин А.А., Соболев Г.А. Мультифрактальные меры синхронизации микросейсмических колебаний в минутном диапазоне периодов // Физика Земли. 2006. № 9. С. 18–28.
- Berger J., Davis P., Ekstrom G. Ambient Earth Noise: A survey of the Global Seismographic Network // J. Geophys. Res. 2004. V. 109. B11307.
- Box G.E.P., Jenkins G.M. Time series analysis. Forecasting and control. San Francisco; Cambridge; London; Amsterdam: Holden-Day, 1970. (Рус. пер.: Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. В 2-х вып. М.: Мир, 1974. Вып. 1, 406 с.; Вып. 2, 197 с.).
- Cox D.R., Hinkley D.V. Theoretical statistics. London: Chapman and Hall, 1974. (Рус. пер.: Кокс Д., Хинкли Д. Теоретическая статистика. М.: Мир, 1978. 560 с.).
- Currenti G., del Negro C., Lapenna V., Telesca L. Multifractality in local geomagnetic field at Etna volcano, Sicily (southern Italy) // Natural Hazards and Earth System Sciences. 2005. V. 5. P. 555–559.
- Duda R.O., Hart P.E. Pattern classification and scene analysis. New York, London, Sydney, Toronto: John Wiley and Sons, 1973. (Рус. пер.: Дуда Р., Харт П. Распознавание образов и анализ сцен. М.: Мир, 1976. 511 с.).
- Ekstrom G. Time domain analysis of Earth's long-period background seismic radiation // J. Geophys. Res. 2001. V. 106, N B11. P. 26483–26493.
- Feder J. Fractals. New York; London: Plenum Press, 1988. (Рус. пер.: Федер Е. Фракталы. М.: Мир, 199. 1254 с.).
- Friedrich A., Krüger F., Klinge K. Ocean-generated microseismic noise located with the Gräfenberg array // J. Seismol. 1998. V. 2, N 1. P. 47–64.
- Gilmore R. Catastrophe theory for scientists and engineers. New York: John Wiley and Sons, Inc., 1981. (Рус. пер.: Гилмор Р. Прикладная теория катастроф. В 2-х кн. М.: Мир, 1984. Кн. 1, 350 с.; Кн. 2, 285с.).
- Hardle W. Applied nonparametric regression. Cambridge, New York, New Rochell, Melbourne, Sydney: Cambridge University Press, 1989. (Рус. пер.: Хардле В. Прикладная непараметрическая регрессия. М.: Мир, 1993. 349 с.).
- Hotelling H. Relations between two sets of variates // Biometrika. 1936. V. 28. P. 321–377.
- Huber P.J. Robust statistics. New York; Chichester; Brisbane; Toronto: John Wiley and Sons, 1981. (Рус. пер.: Хьюбер П. Робастность в статистике. М.: Мир, 1984. 303 с.).
- Ida Y., Hayakawa M., Adalev A., Gotoh K. Multifractal analysis for the ULF geomagnetic data during the 1993 Guam earthquake // Nonlinear Processes in Geophysics. 2005. V. 12. P. 157–162.
- 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. 2002. V. 316. P. 87–114.
- Kashyap R.L., Rao A.R. Dynamic stochastic models from empirical data. New York; San Francisco; London: Acad. Press, 1976. (Рус. пер.: 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. L24304.

- Lyubushin A.A.* Mean multifractal properties of low-frequency microseismic noise // Proceedings of 31st General Assembly of the European Seismological Commission ESC-2008. Hersonissos, Crete, Greece, 7–12 Sept. 2008. 2008a. P. 255–270.
- Lyubushin A.A.* Multifractal properties of low-frequency microseismic noise in Japan, 1997–2008 // Book of abstracts of 7th General Assembly of the Asian Seismological Commission and Japan Seismological Society. Fall meeting. Tsukuba, Japan, 24–27 Nov., 2008. 2008b. P. 92.
- Lyubushin A.A.* Synchronization of multifractal parameters of regional and global low-frequency microseisms // European Geosciences Union General Assembly 2010, Vienna, 02–07 of May, 2010. Geophys. Res. Abstr. V. 12. EGU2010-696, 2010a.
- Lyubushin A.A.* Synchronization phenomena of low-frequency microseisms // European Seismological Commission, 32nd General Assembly, Sept. 06–10, 2010, Montpellier, France: Book of abstr. Ses. ES6. 2010b. P. 124.
- Lyubushin A.* Multifractal parameters of low-frequency microseisms // Synchronization and Triggering: from Fracture to Earthquake Processes, GeoPlanet: Earth and Planetary Sciences, DOI 10.1007/978-3-642-12300-9\_15. Berlin; Heidelberg: Springer-Verlag, 2010c. 388 p. Chapter 15. P. 253–272.
- Mallat S.* A wavelet tour of signal processing. San Diego; London; Boston; New York; Sydney; Tokyo; Toronto: Academic Press, 1998. 577 p. (Рус. пер.: Малла С. Вэйвлеты в обработке сигналов. М.: Мир, 2005. 671 с.).
- Mandelbrot B.B.* The fractal geometry of nature. New York: Freeman and Co., 1982. (Рус. пер.: Мандельброт Б. Фрактальная геометрия природы. М.: Ин-т компьютер. исслед., 2002. 656 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.
- Ramírez-Rojas A., Muñoz-Diosdado A., Pavía-Miller C.G., Angulo-Brown F.* Spectral and multifractal study of electroseismic time series associated to the  $M_w=6.5$  earthquake of 24 Oct. 1993 in Mexico // Natural Hazards and Earth System Sciences. 2004. V. 4. P. 703–709.
- Rao C.R.* Linear statistical inference and its applications. New York; London; Sydney: John Wiley and Sons, 1965. (Рус. пер.: РАО С.Р. Линейные статистические методы и их применение. М.: Наука, 1968. 548 с.).
- Rhie J., Romanowicz B.* Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling // Nature. 2004. V. 431. P. 552–554.
- Rhie J., Romanowicz B.* A study of the relation between ocean storms and the Earth's hum -  $G^3$ : geochemistry, geophysics, geosystems // An Electronic Journal of Earth Sciences. 2006. V. 7, N 10. <http://www.agu.org/journals/gc/>
- 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.* Continuous Free Oscillations: Atmosphere-Solid Earth Coupling // Ann. Rev. Earth Planet. Sci. 2001. V. 29. P. 563–584.
- Tanimoto T.* The oceanic excitation hypothesis for the continuous oscillations of the Earth // Geophys. J. Int. 2005. V. 160. P. 276–288.
- Tanimoto T., Um J.* Cause of continuous oscillations of the Earth // J. Geophys. Res. 1999. V. 104, N 28. P. 723–739.
- 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.
- Taqqu M.S.* Self-similar processes // Encyclopedia of Statistical Sciences. New York: John Wiley and Sons, 1988. V. 8. P. 352–357.
- Telesca L., Colangelo G., Lapenna V.* Multifractal variability in geoelectrical signals and correlations with seismicity: a study case in southern Italy // Natural Hazards and Earth System Sciences. 2005. V. 5. P. 673–677.

Vogel M.A., Wong A.K.C. PFS Clustering Method // IEEE (Institute of Electrical and Electronics Engineers) Transactions on Pattern Analysis and Machine Intelligence. 1978. V. PAMI1-5. P. 237–245.

Ziganshin A.R., Pavlov A.N. Scaling properties of multimode dynamics in coupled chaotic oscillators // Proceedings of 2-nd Int. Conference «Physics and Control», Aug, 24–26, 2005, St. Petersburg. St. Petersburg, 2005. P. 180–183.

#### *Сведения об авторе*

**ЛЮБУШИН Алексей Александрович** – доктор физико-математических наук, главный научный сотрудник, Институт физики Земли им. О.Ю. Шмидта РАН. 123995, ГСП-5, г. Москва, Д-242, ул. Большая Грузинская, д. 10. Тел.: 8 (499) 254-23-50. E-mail: lyubushin@yandex.ru

## **JAPAN SEISMIC CATASTROPHE OF 11 MARCH, 2011: LONG-TERM PREDICTION BY MICROSEISMIC NOISE PROPERTIES**

**A.A. Lyubushin**

Schmidt Institute of Physics of the Earth, Russian Academy of Sciences, Moscow, Russia

**Abstract.** The software technique for multidimensional time series processing applied to records of microseisms at the broad-band seismic network F-net in Japan (14 years time period, from beginning of 1997 up to the February of 2011) is presented. The analysis of low-frequency microseismic noise multi-fractal properties allowed a hypothesis about approaching Japan Islands a future seismic catastrophe to be formulated at the middle of 2008. The base for such a hypothesis was statistically significant decreasing of multi-fractal singularity spectrum support width mean value. Further on as far as the new information from monitoring system became available and due to including into the analysis of new parameters of microseismic noise (such as logarithm of variance and index of linear predictability) new results confirming increasing of microseisms synchronization were obtained. The process of increasing of synchronization (its beginning was estimated as the middle of 2002) was a flag of an impending large earthquake. The cluster analysis of microseismic noise parameters allowed us to make a decision that starting from July of 2010 Japan islands come to the state of impending strong earthquake. This prediction of Great Japan Earthquake, initially with estimate of lower magnitude as 8.3 only (at the middle of 2008) and further on with estimate of the time beginning of waiting earthquake (from the middle of 2010) was published in advance of the event in a number of scientific articles and abstracts on international conferences.

*Keywords:* synchronization, microseismic noise, multifractals, earthquake precursors.