

Проблема больших данных в геофизике: иерархически агрегированные сигналы и обобщенная когерентность шумов

А.А. Любушин, доктор физ.-мат. наук
Институт физики Земли РАН им. О.Ю. Шмидта,
e-mail: lyubushin@yandex.ru
<http://AlexeyLyubushin.narod.ru/Index.htm>

Доклад на Ученом Совете ИФЗ РАН, 10 февраля 2016 г.

В этой книге много интересных примеров того, как самые сложные технологии Big Data — методы анализа огромных объемов данных — применяются для решения важных задач из нашей повседневной жизни.

Сергей Мацюцкий, председатель правления компании IBS

ВИКТОР МАЙЕР-ШЕНБЕРГЕР | КЕННЕТ КУКЬЕР

BIG

БОЛЬШИЕ ДАННЫЕ

DATA

РЕВОЛЮЦИЯ, КОТОРАЯ
ИЗМЕНИТ ТО, КАК МЫ ЖИВЕМ,
РАБОТАЕМ И МЫСЛИМ

Viktor Mayer-Schönberger
Kenneth Cukier

BIG DATA

A Revolution That Will Transform
How We Live, Work, and Think

Полезные цитаты из этой книги:

«Имея дело с большими данными, как правило, приходится довольствоваться общим представлением, а не пониманием явления вплоть до молекулы.»

«Люди привыкли во всем искать причины, даже если установить их не так просто или малополезно. В мире больших данных мы больше не обязаны цепляться за причинность. Вместо этого мы можем находить корреляции между данными, которые открывают перед нами новые знания. Корреляции не могут сказать нам точно, почему происходит то или иное событие, зато предупреждают о том, какого оно рода. И в большинстве случаев этого вполне достаточно.»

«Вопрос не в том **почему**, а в том **что**. В мире больших данных нам не всегда нужно знать причины, которые стоят за теми или иными явлениями. Лучше позволить данным говорить самим за себя.»

«Использование огромных массивов данных вместо их малой доли и выбор количества в ущерб точности открывают путь к новым способам понимания мира. Это подталкивает общество отказаться от освященного веками поиска причинности и в большинстве случаев пользоваться преимуществами корреляций.»

«Поиск причин стал своего рода религией современности. Большие данные в корне меняют это мировоззрение, и мы снова оказываемся в таком историческом тупике, где «Бог умер». То, в чем мы были непоколебимо уверены, в очередной раз меняется»

Корреляции ~ Когерентность ~ Синхронизация

Некоторые источники геофизических «больших данных» в интернете:

Ежесуточные 3-компонентные временные ряды GPS:

http://gf9.ucs.indiana.edu/daily_rdahmmexec/daily/

<http://sideshow.jpl.nasa.gov/post/series.html>

<ftp://sideshow.jpl.nasa.gov/pub/usrs/mbh/point/>

Трех-компонентные временные ряды GPS с шагом по времени 5 минут:

ftp://gneiss.nbmq.unr.edu/rapids_5min/kenv/

Данные широкополосных сейсмических сетей:

<http://www.fnet.bosai.go.jp/top.php>

http://www.iris.edu/mda/_GSN

<http://www.iris.edu/mda/G>

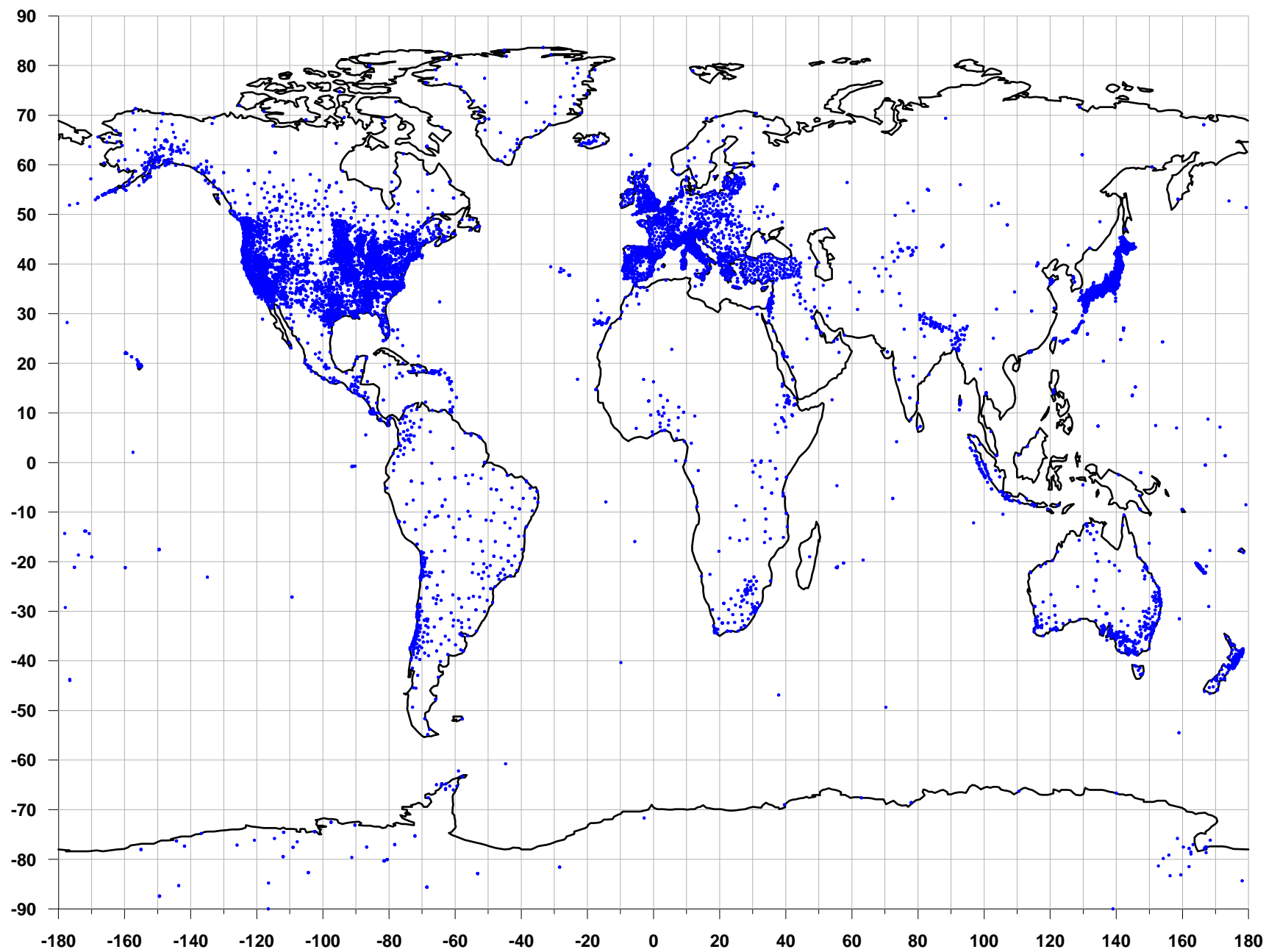
<http://www.iris.edu/mda/GE>

<http://ds.iris.edu/mda/AZ>

<http://ds.iris.edu/mda/BK>

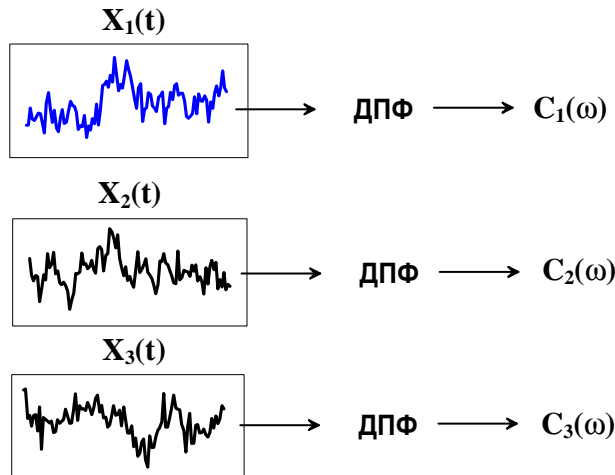
<http://ds.iris.edu/mda/CI>

Положения 10590 стационарных станций GPS по всему миру



Фурье-агрегированный сигнал

1-я стадия агрегации:
вычисление канонических компонент

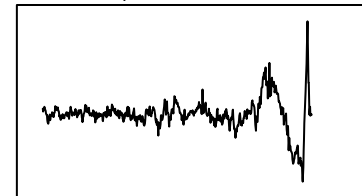
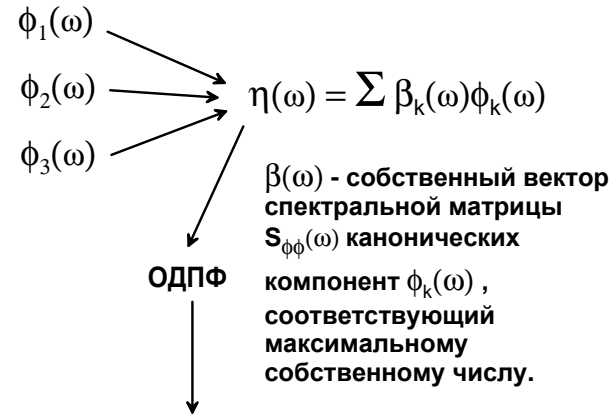


Каноническая компонента $X_1(t)$
в частотном представлении:

$$\phi_1(\omega) = \gamma_2 C_2(\omega) + \gamma_3 C_3(\omega)$$

Коэффициенты γ находятся из условия
максимума квадрата когерентности между
 $\phi_1(\omega)$ и $C_1(\omega)$

2-я стадия агрегации:
главная компонента
от канонических компонент



Агрегированный сигнал

На 1-ой ступени агрегации каждый из анализируемых временных рядов заменяется на так называемые канонические компоненты, в которых подавляются «индивидуальные» вариации, присущие только тому или иному сигналу и отсутствующие в других сигналах.

На 2-й ступени агрегации из набора канонических компонент строится первая главная компонента, которая усиливает общие вариации, присутствующие сразу во всех временных рядах.

Программа вычисления Фурье-агрегированного сигнала и руководство пользователя на русском языке вместе с описанием метода и примером применения свободно доступны на сайте ИФЗ:

http://old.ifz.ru/scient_directions/data_monitoring/06/AggF.pdf

http://old.ifz.ru/scient_directions/data_monitoring/06/AggF.zip

Техника агрегирования большого числа временных рядов с пропусками

Однако непосредственно применять процедуру агрегации нескольких тысяч временных рядов с пропусками невозможно. Для этой цели была разработана процедура иерархической агрегации.

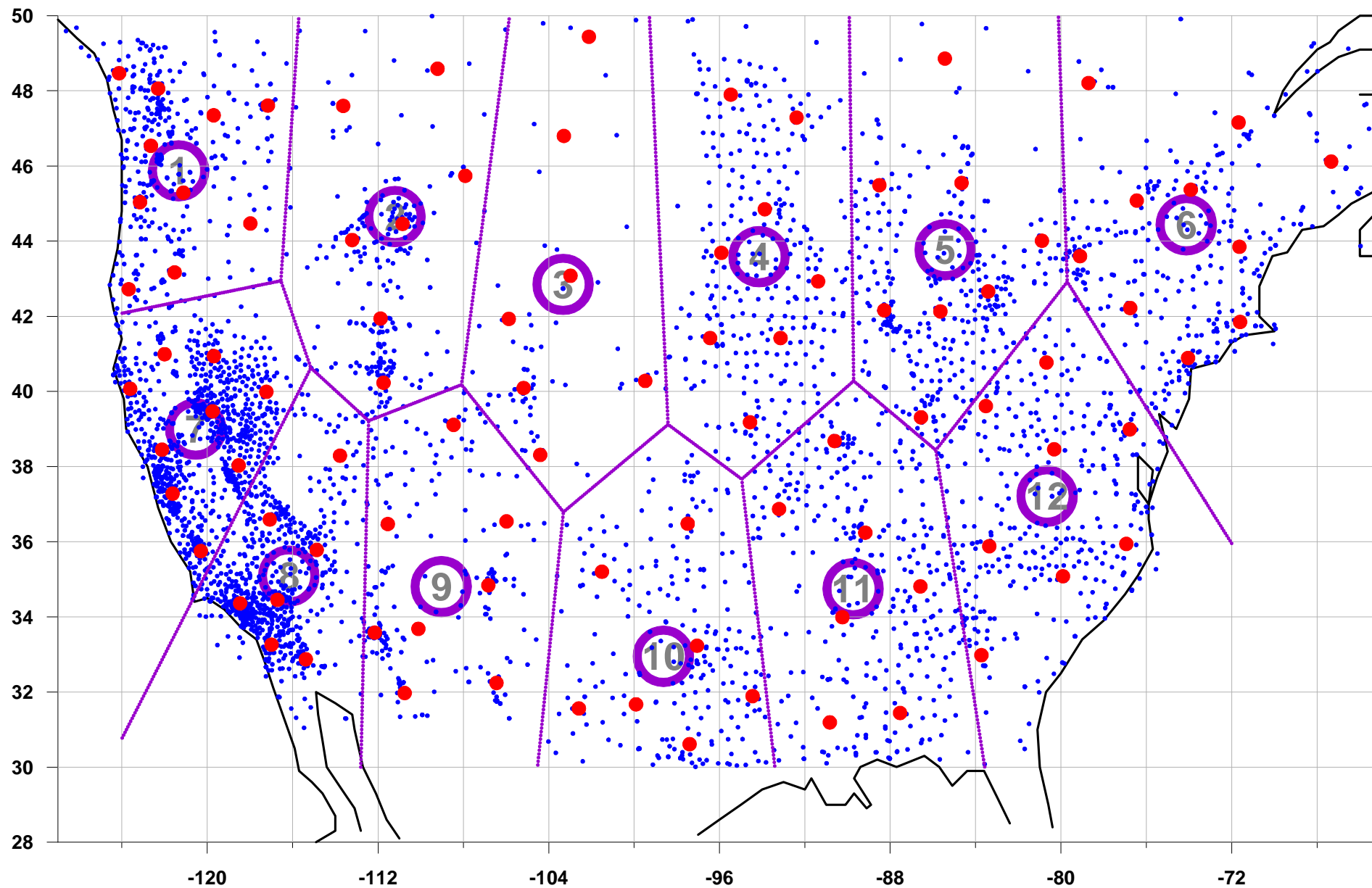
Станция регистрации считается работоспособной в течение текущего временного окна заданной длины (далее используется длина окна 730 суток) если суммарное число пропусков в этом окне не превышает заданного малого процента от длины окна. Далее работоспособной считалась станция, для которой в текущем временном окне число пропусков не превышало 5% от длины окна. Таким образом, для окна длиной 730 суток суммарное число пропусков не должно превышать 37 отсчетов. Пропущенные значения восполняются «правдоподобными» значениями по данным из соседних интервалов времени такой же длины, что и пропуск.

Центром агрегации может быть любая географическая точка. Положения центров агрегации находятся путем разбиения множества положений станций GPS на заданное число кластеров методом *k-средних*. После решения задачи *кластерного анализа* центрами агрегации назначались центры масс кластеров.

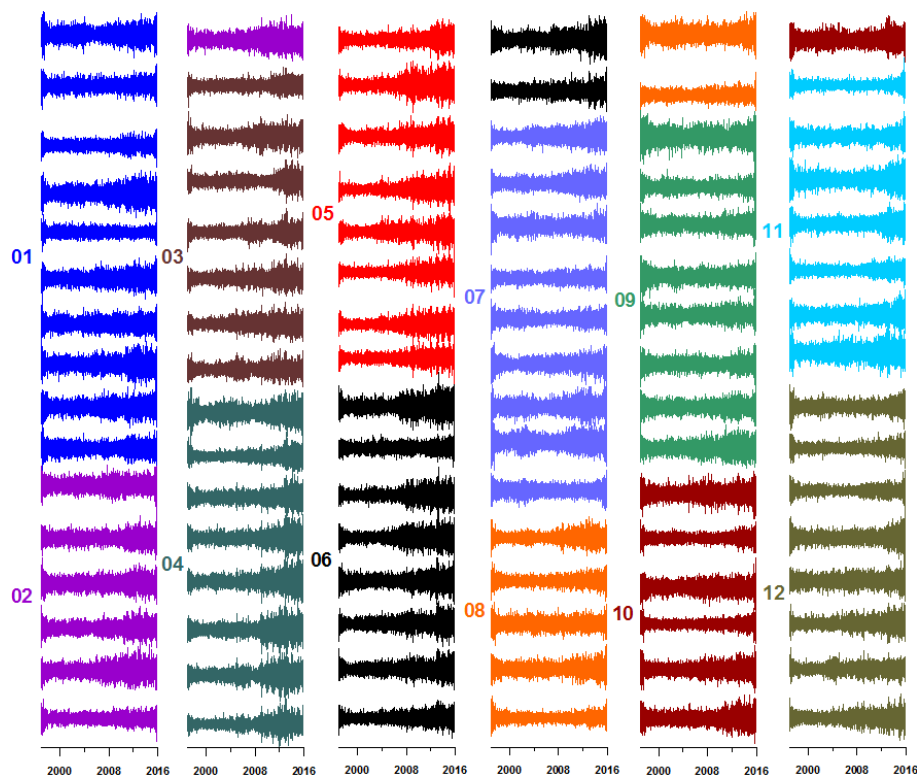
Для каждого временного окна и для каждого центра агрегации определялось заданное число ближайших работоспособных станций. Далее число таких ближайших станций бралось равным 20. Для данных от ближайших работоспособных станций вычислялись Фурье-агрегированные сигналы. Таким образом, для каждого из центров агрегации получается последовательность агрегированных сигналов, соответствующая последовательности временных окон заданной длины (730 суток), взятых с некоторым смещением. Далее использовалось смещение временных окон равное 28 суткам.

При такой технике агрегирования временных рядов каждому временному отсчету, то есть каждым суткам, соответствуют значения агрегированных сигналов от нескольких смежных временных окон, которые включают в себя этот временной отсчет (сутки). Усредненный агрегированный сигнал получается усреднением всех таких значений от смежных временных окон. В результате каждому центру агрегации соответствует свой усредненный агрегированный сигнал.

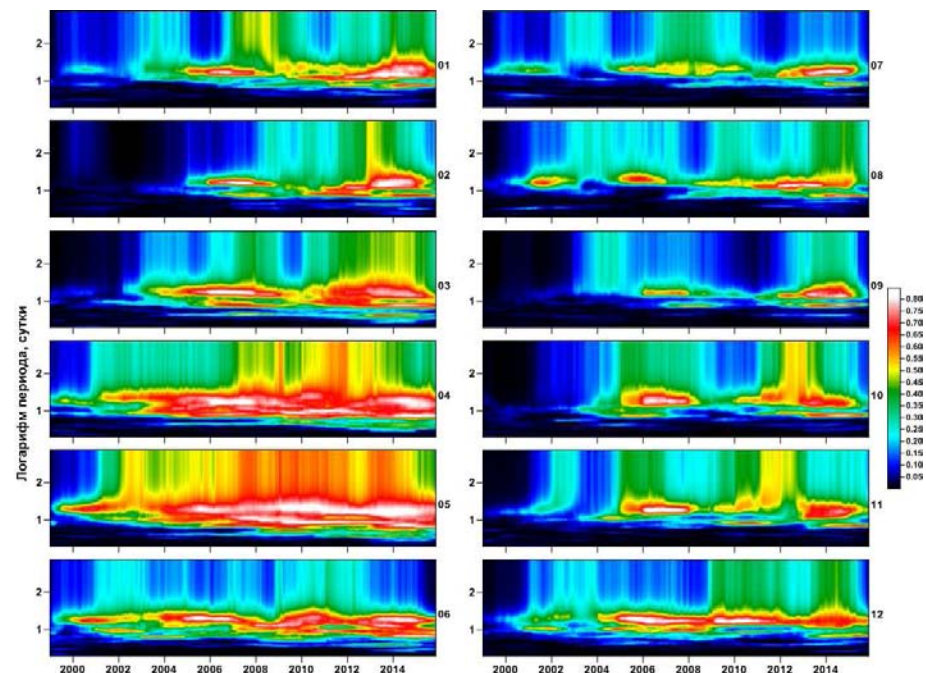
Сеть 4512 стационарных станций GPS на территории США. Красные точки - 96 центров агрегации 1-го порядка. Пурпурные кружки - 12 центров агрегации 2-го порядка. Показано разбиение на 12 областей агрегации 2-го порядка в виде полигонов Вороного.



Графики агрегированных сигналов для 96 центров первичной агрегации ежесуточных временных рядов приращений вертикальных компонент GPS измерений на сети из 4512 станций на территории США. Разными цветами показаны агрегированные сигналы принадлежащие 12 различным областям вторичной агрегации.



Частотно-временные диаграммы эволюции множественной спектральной меры когерентности для агрегированных сигналов 1-го порядка из 12 областей вторичной агрегации приращений вертикальных компонент на сети GPS станций на территории США.

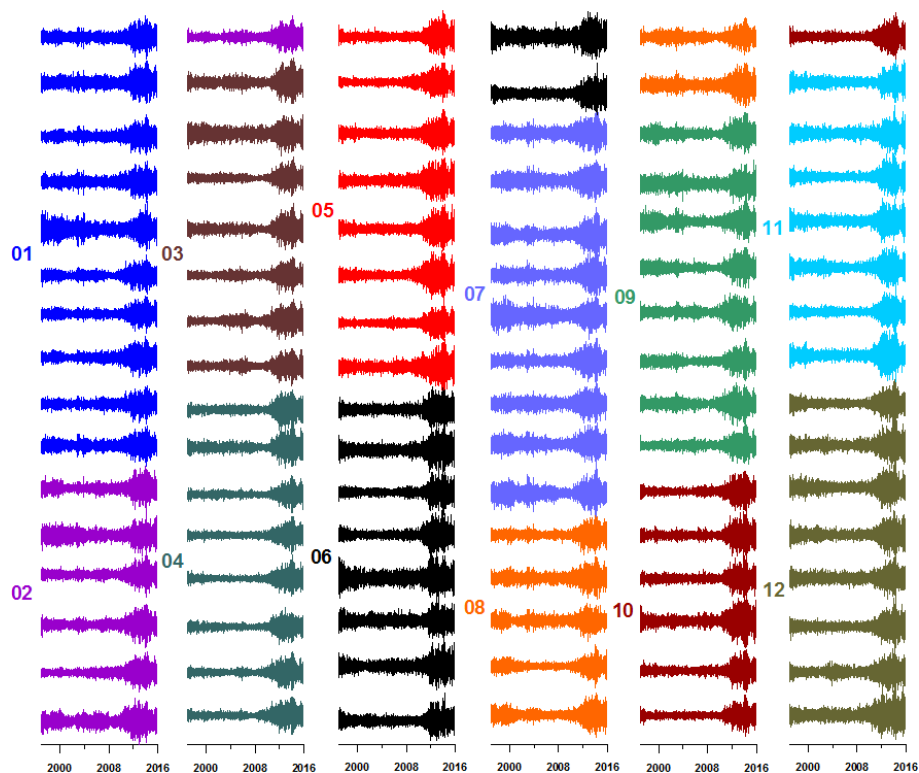


Программа вычисления множественной меры когерентности многомерных сигналов и руководство пользователя на русском языке вместе с описанием метода и примерами применения свободно доступны на сайте ИФЗ:

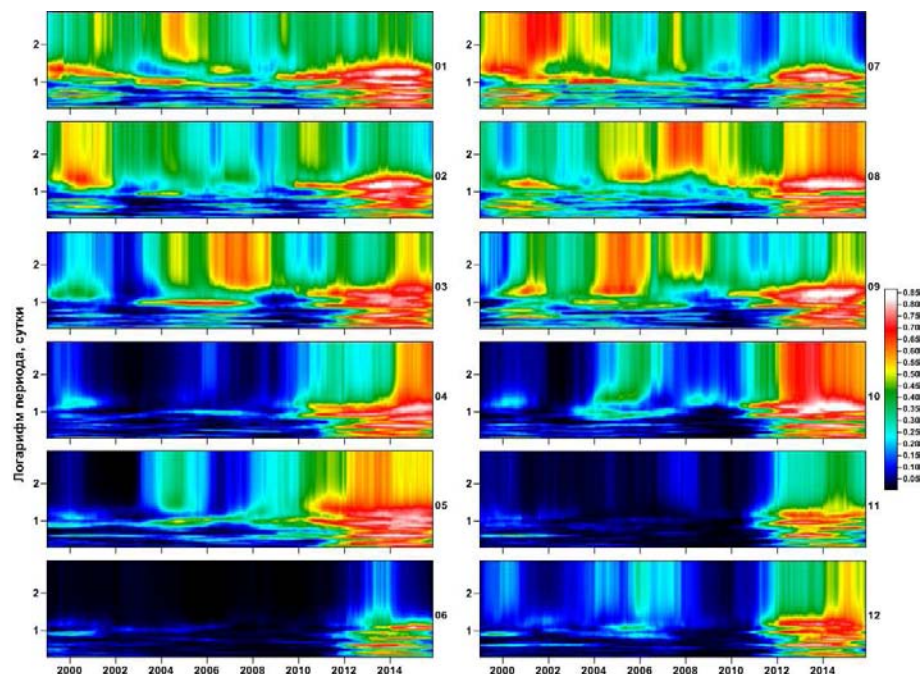
http://old.ifz.ru/scient_directions/data_monitoring/05/SpectCohMes.pdf

http://old.ifz.ru/scient_directions/data_monitoring/05/SpectCohMes.zip

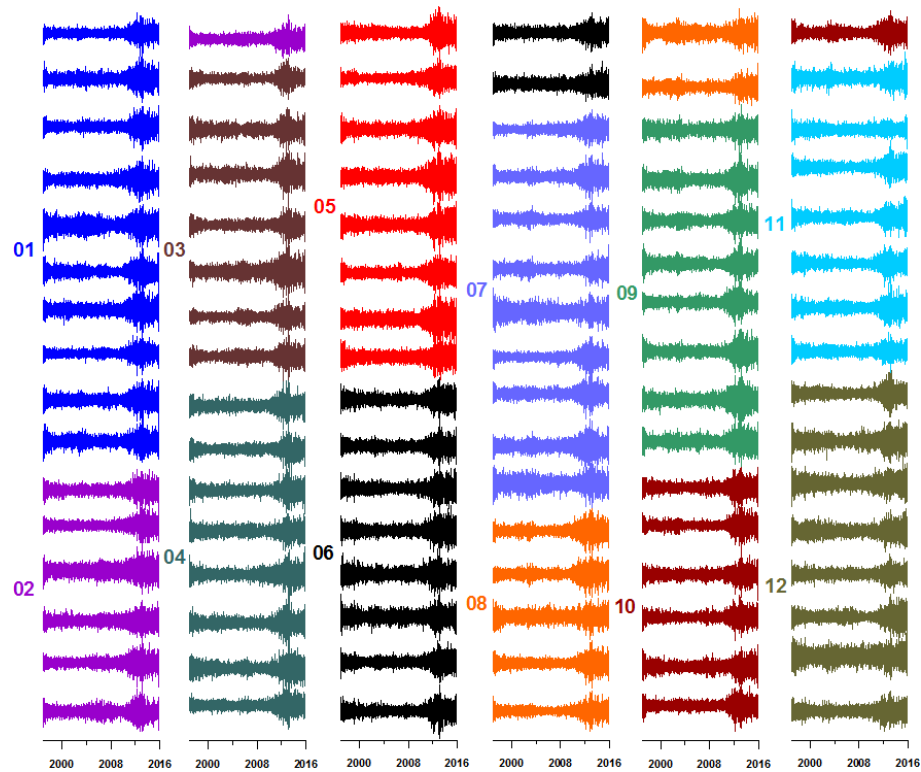
Графики агрегированных сигналов для 96 центров первичной агрегации ежесуточных временных рядов приращений компонент Запад-Восток GPS измерений на сети из 4512 станций на территории США. Разными цветами показаны агрегированные сигналы принадлежащие 12 различным областям вторичной агрегации.



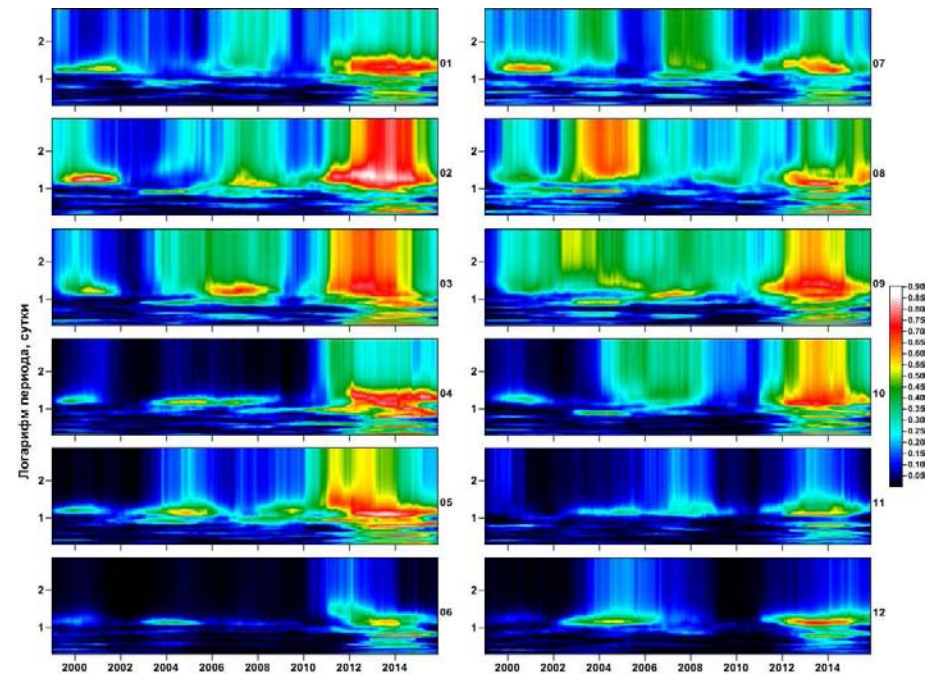
Частотно-временные диаграммы эволюции множественной спектральной меры когерентности для агрегированных сигналов 1-го порядка из 12 областей вторичной агрегации приращений компонент Запад-Восток на сети GPS станций на территории США.



Графики агрегированных сигналов для 96 центров первичной агрегации ежесуточных временных рядов приращений компонент Север-Юг GPS измерений на сети из 4512 станций на территории США. Разными цветами показаны агрегированные сигналы принадлежащие 12 различным областям вторичной агрегации.

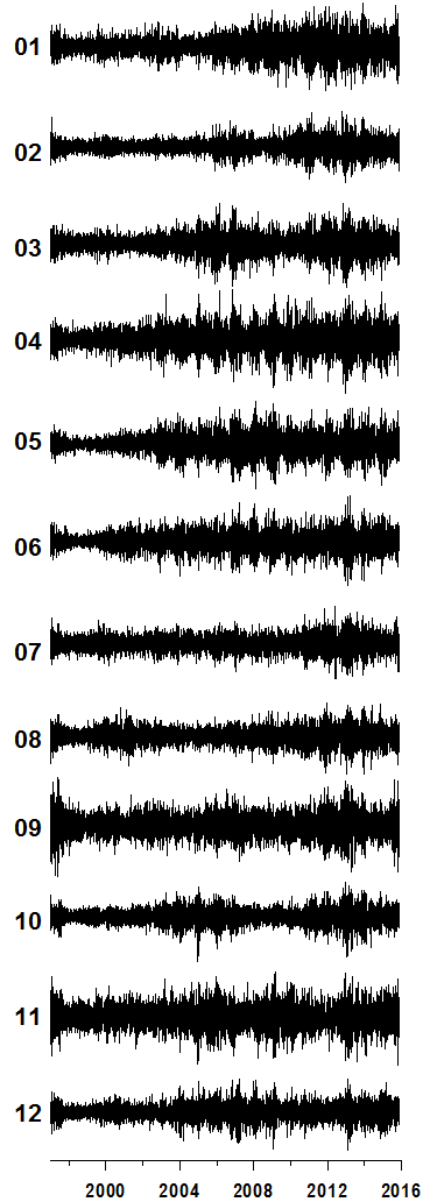


Частотно-временные диаграммы эволюции множественной спектральной меры когерентности для агрегированных сигналов 1-го порядка из 12 областей вторичной агрегации приращений компонент Север-Юг на сети GPS станций на территории США.

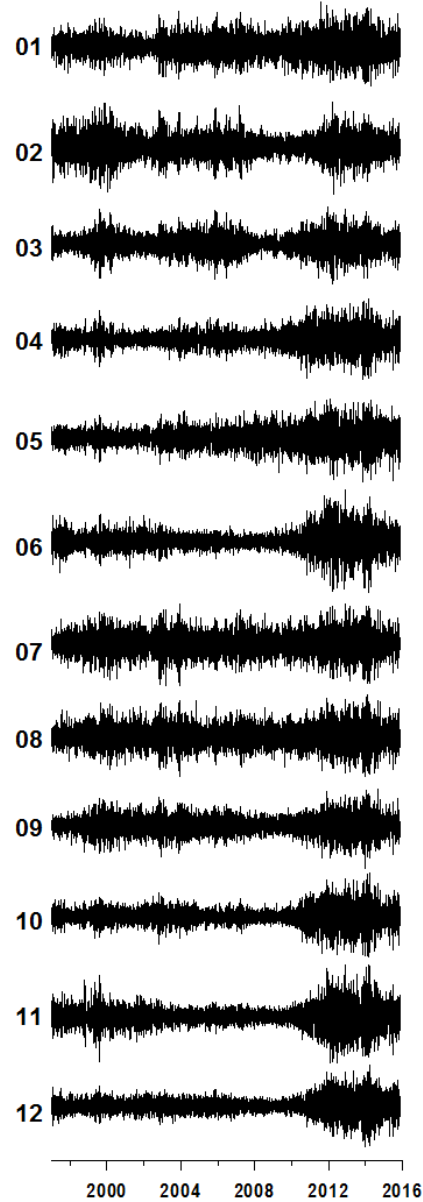


Фурье-агрегированные сигналы 2-го порядка для каждого из 12 центров вторичной агрегации для приращений ежесуточных временных рядов GPS на территории США. Усреднение оценок во временных окнах длиной 730 суток со смещением 28 суток. Агрегация в каждом окне от 20 ближайших работоспособных станций.

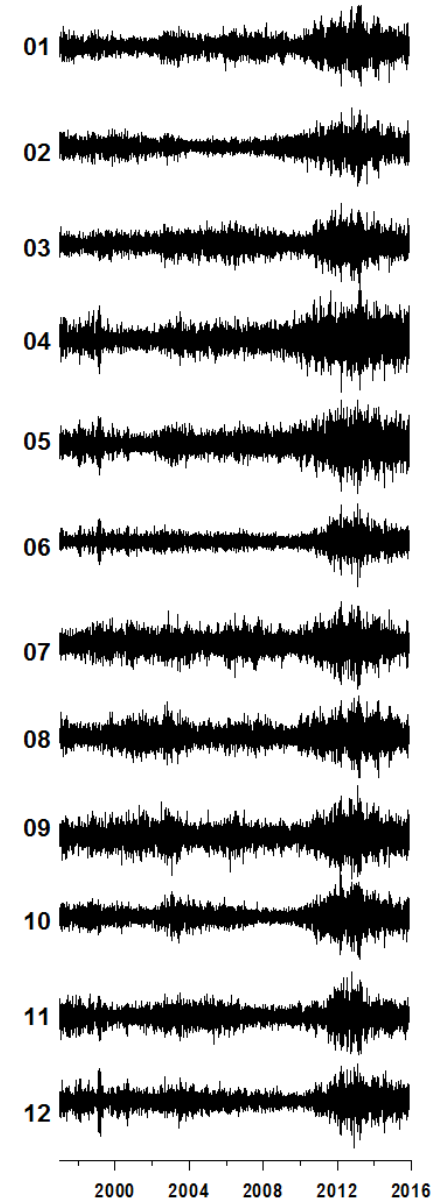
Вертикальная компонента



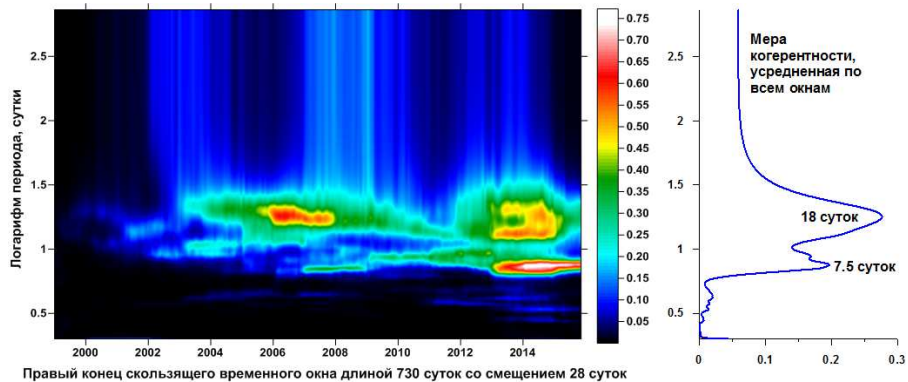
Запад-Восток



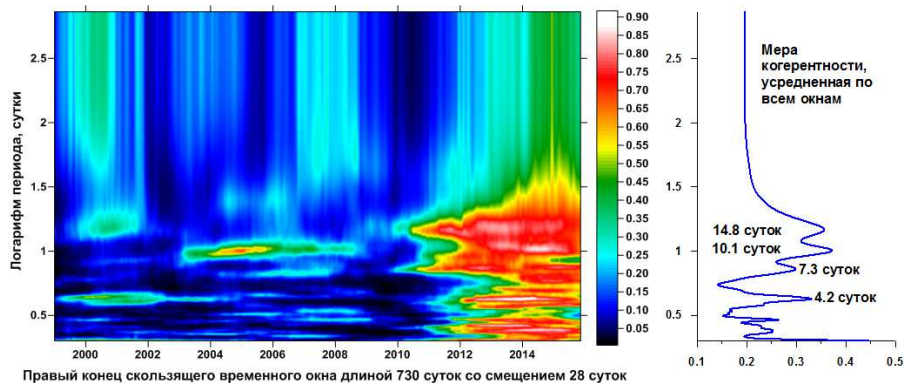
Север-Юг



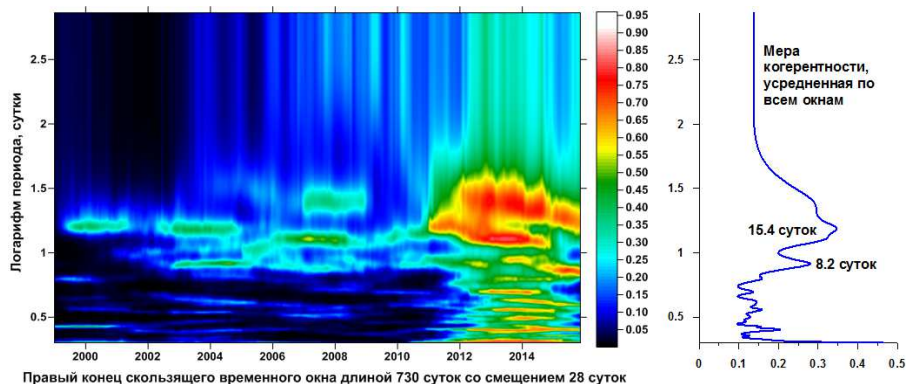
Частотно-временные диаграммы эволюции множественных спектральных мер когерентности 12-мерных временных рядов агрегированных сигналов 2-го порядка на территории США.



Для приращений вертикальной компоненты временных рядов GPS.

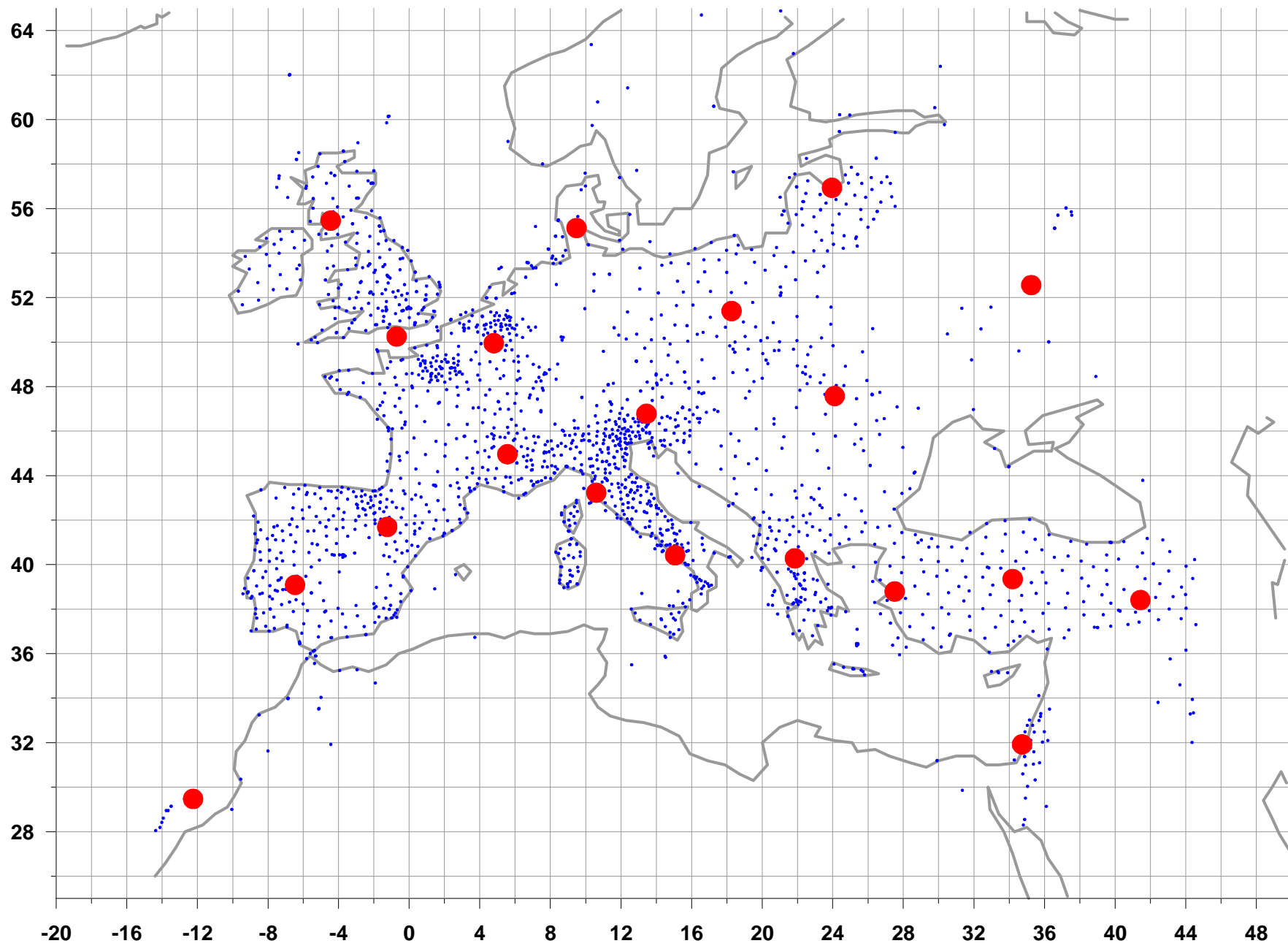


Для приращений компоненты Запад-Восток временных рядов GPS.



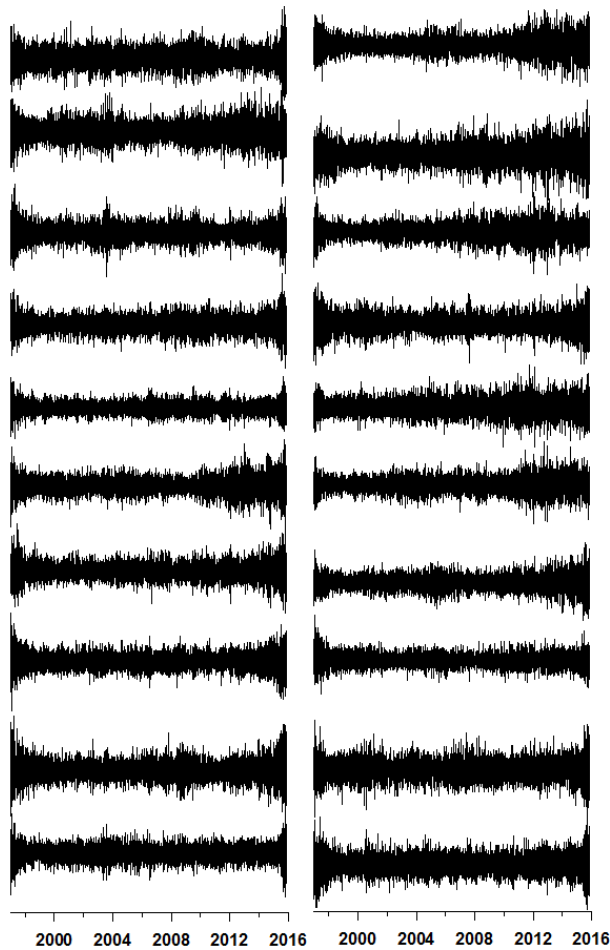
Для приращений компоненты Север-Юг временных рядов GPS.

Сеть 2122 GPS станций в Европе, 20 центров агрегации

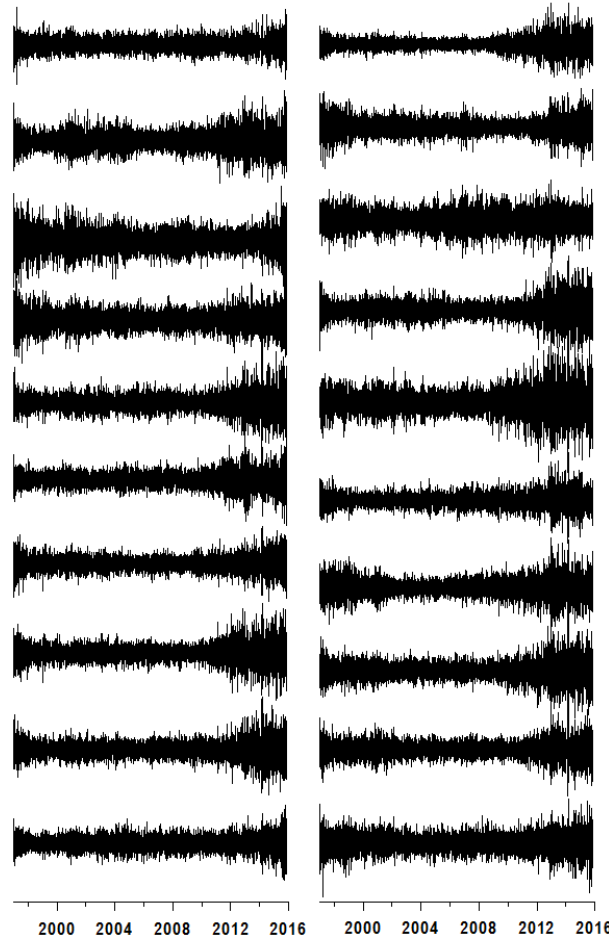


Фурье-агрегированные сигналы 1-го порядка для каждого из 20 центров агрегации для приращений
ежесуточных временных рядов GPS на территории Европы. Усреднение оценок во временных окнах длиной
730 суток со смещением 28 суток. Агрегация в каждом окне от 20 ближайших работоспособных станций.

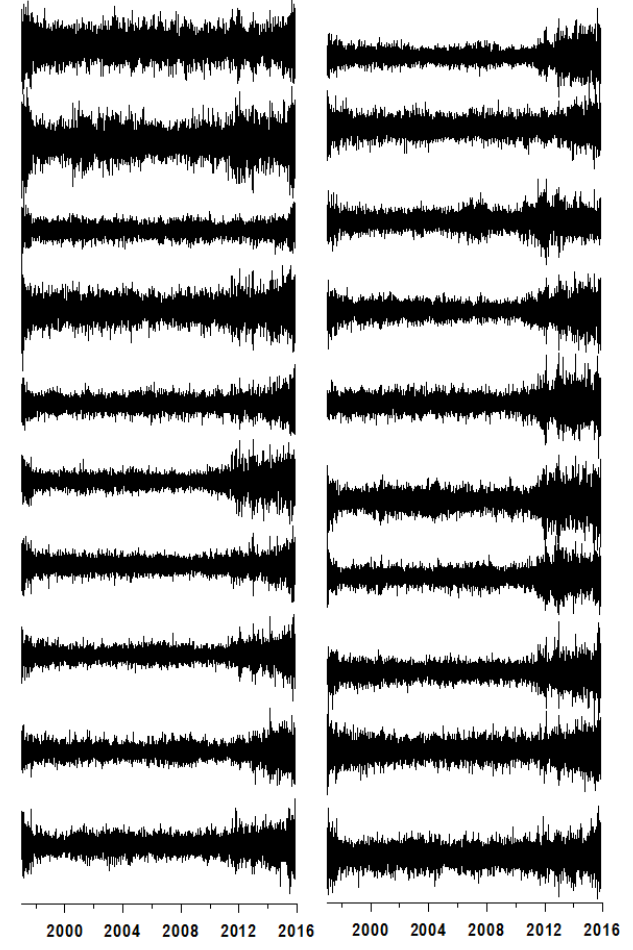
Вертикальная компонента



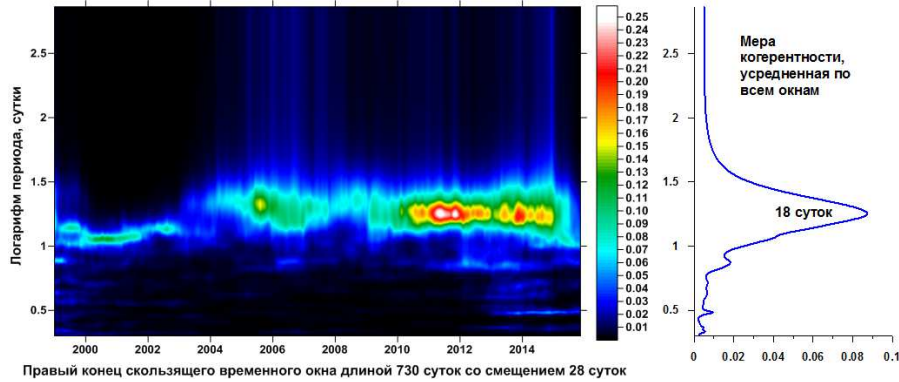
Запад-Восток



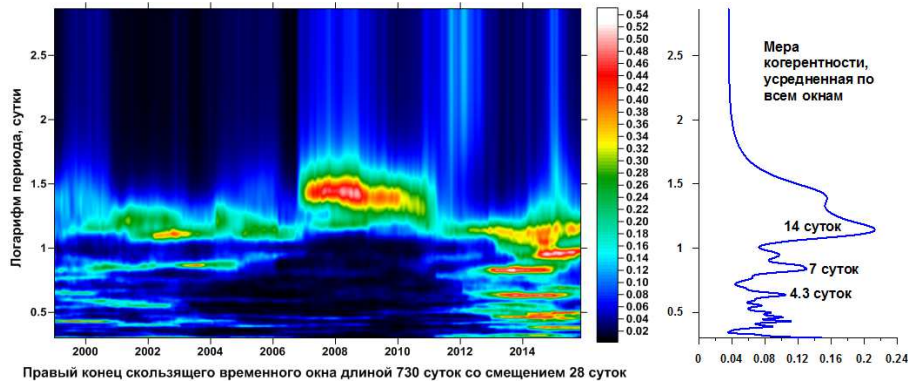
Север-Юг



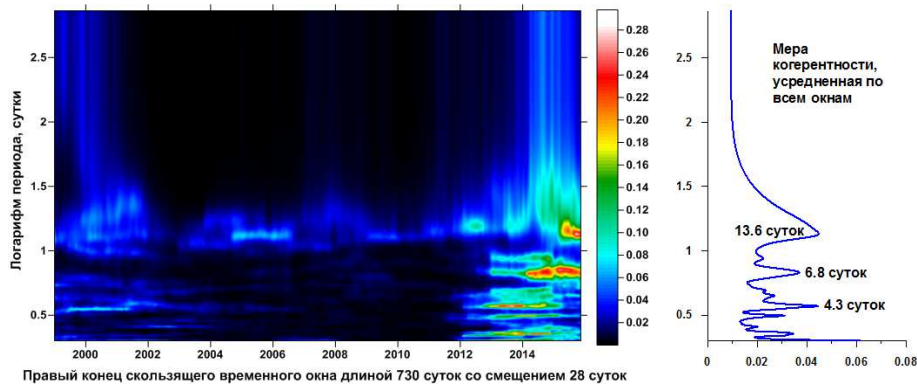
Частотно-временные диаграммы эволюции множественных спектральных мер когерентности 20-мерных временных рядов агрегированных сигналов 1-го порядка на территории Европы.



**Для приращений
вертикальной
компоненты временных
рядов GPS.**

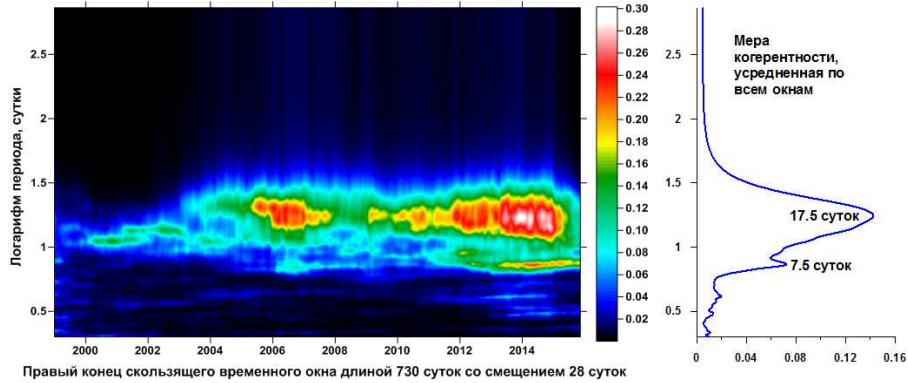


**Для приращений
компоненты Запад-
Восток временных
рядов GPS.**

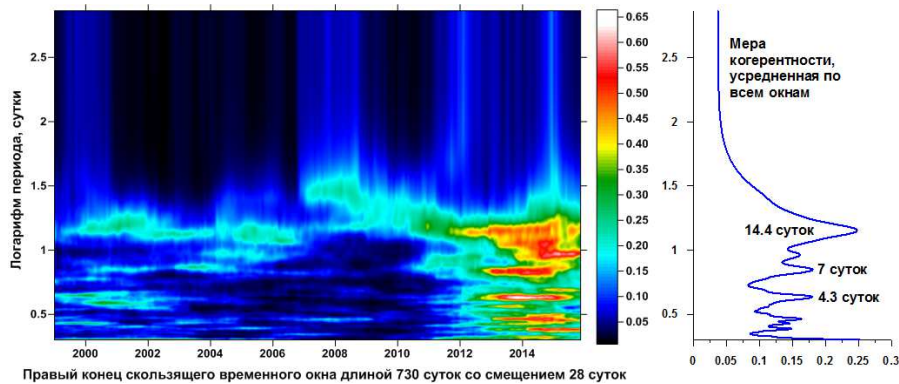
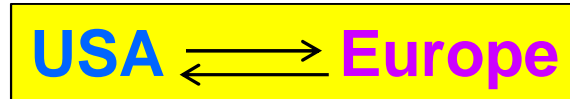


**Для приращений
компоненты Север-Юг
временных рядов GPS.**

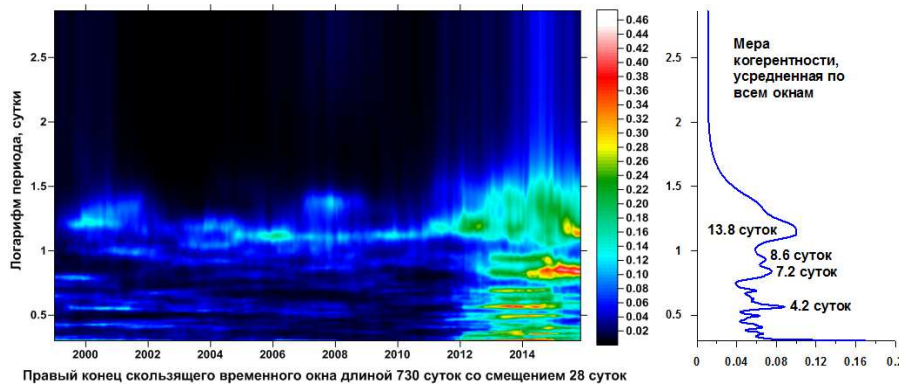
Частотно-временные диаграммы эволюции множественной спектральной меры когерентности 32-мерных временных рядов: 12 агрегированных сигналов 2-го порядка на территории США и 20 агрегированных сигналов 1-го порядка на территории Европы.



Для приращений вертикальной компоненты временных рядов GPS.

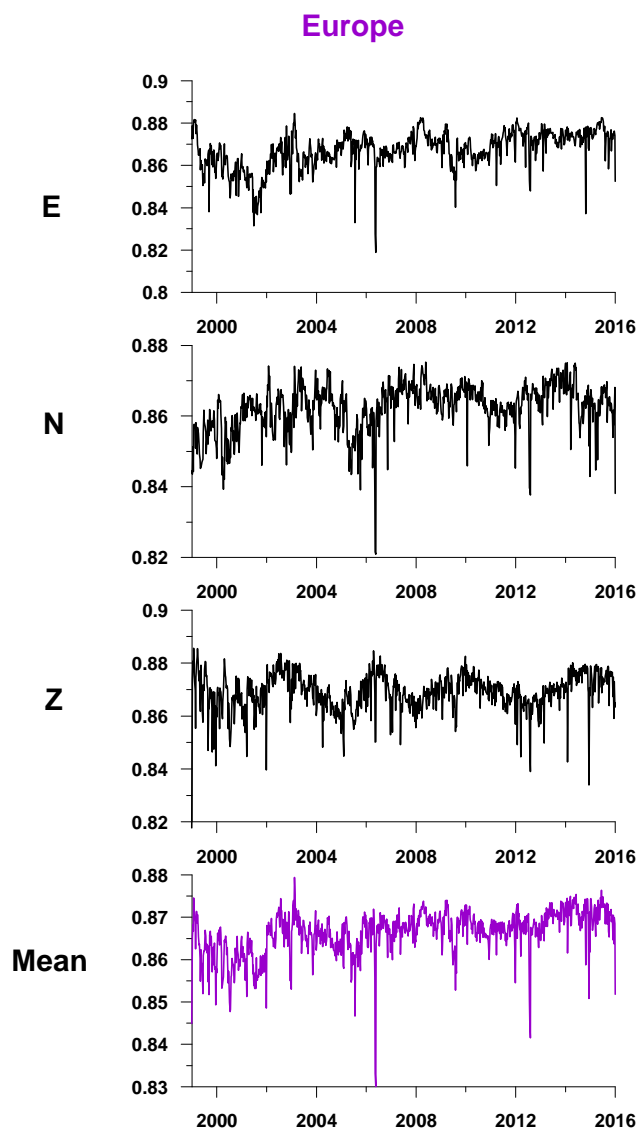
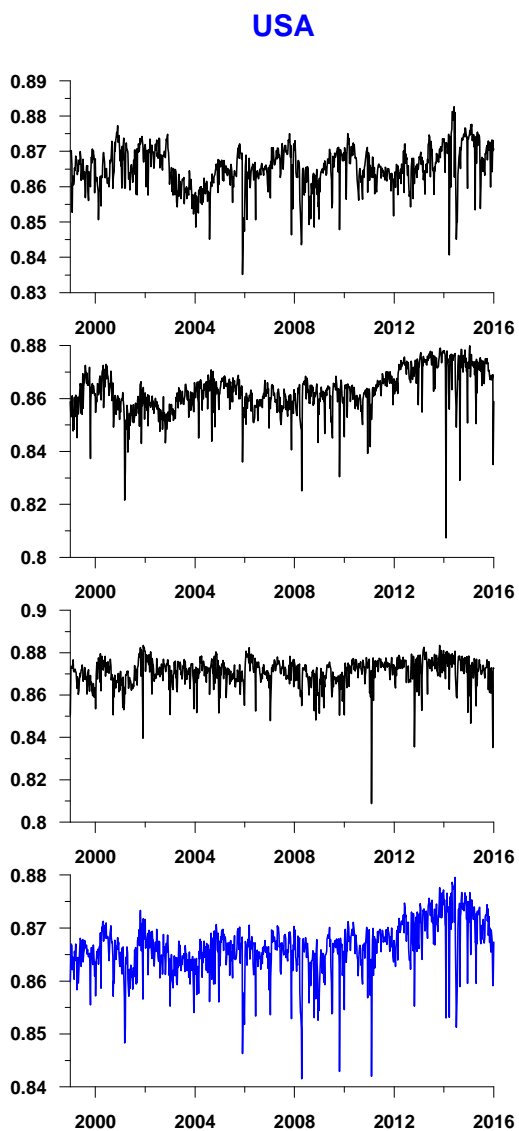


Для приращений компоненты Запад-Восток временных рядов GPS.



Для приращений компоненты Север-Юг временных рядов GPS.

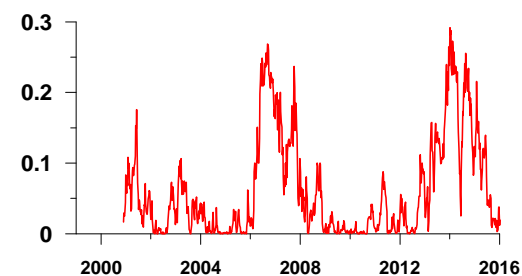
Средние значения нормализованной энтропии скачков в ежесуточных временных рядах GPS, вычисленные по всем станциям на территории США и Европы во временных окнах длины 730 суток со смещением 7 суток.



Любушин А.А., Яковлев П.В. Энтропийная мера скачкообразной составляющей временных рядов GPS – Физика Земли, 2016, №1, с.98–107.
<http://link.springer.com/article/10.1134/S106935131506004X>

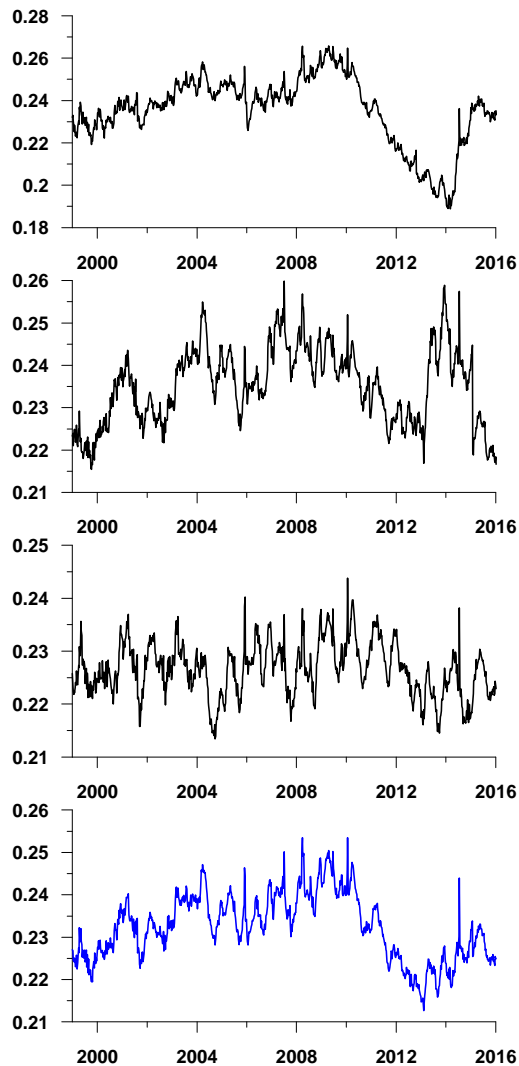
Корреляция (1) между усредненными значениями параметров сейсмических шумов в США и в Европе

Квадрат коэффициента корреляции между средними значениями энтропии скачков (усреднение по станциям и по 3-м компонентам) в скользящем временном окне длиной 700 суток

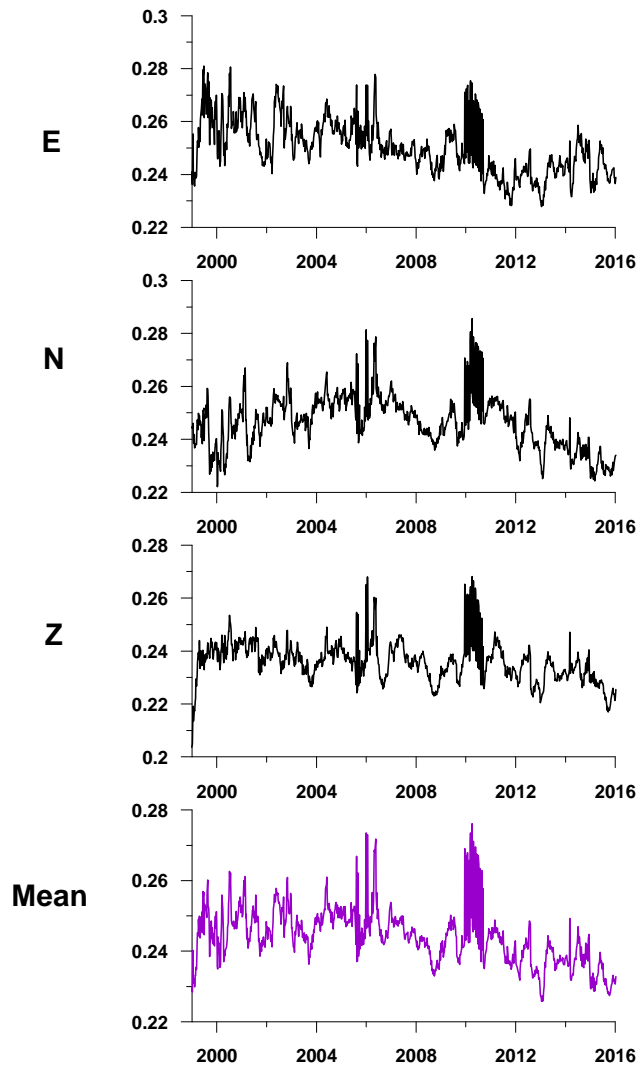


Средние значения обобщенного показателя Херста ежесуточных временных рядах GPS, вычисленные по всем станциям на территории США и Европы во временных окнах длины 730 суток со смещением 7 суток.

USA

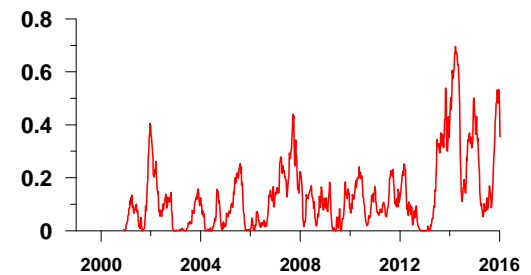


Europe



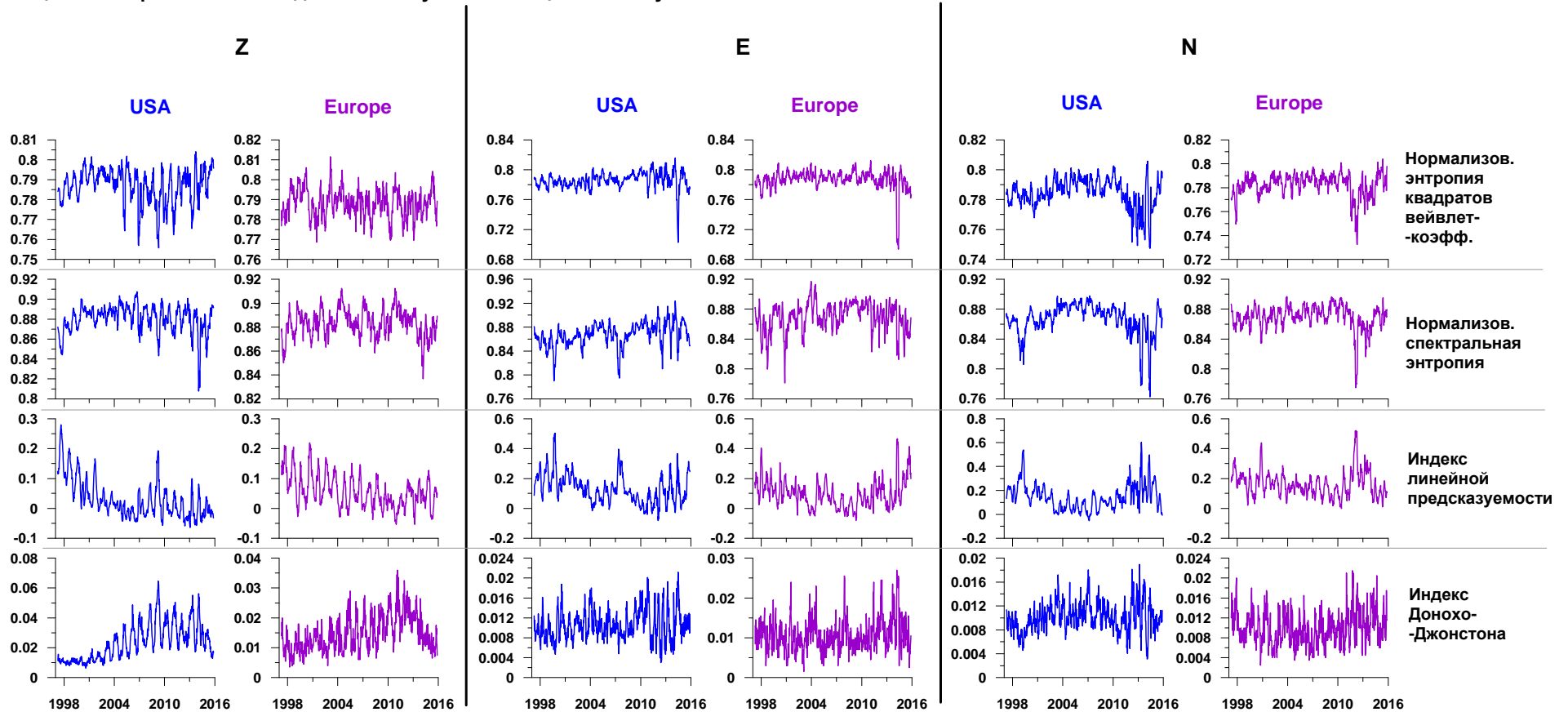
Корреляция (2) между усредненными значениями параметров сейсмических шумов в США и в Европе

Квадрат коэффициента корреляции между средними значениями обобщенного показателя Херста (усреднение по станциям и по 3-м компонентам) в скользящем временном окне длиной 700 суток

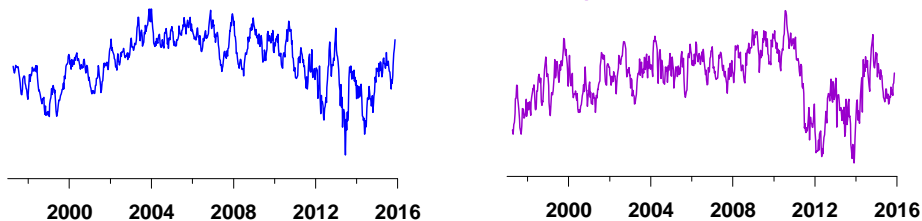


Корреляция (3) между усредненными значениями параметров сейсмических шумов в США и в Европе

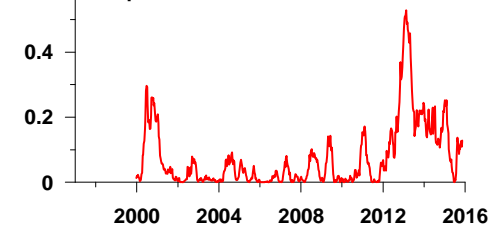
Средние значение различных статистик Фурье-агрегированных сигналов 1-й степени агрегации ежесуточных временных рядов GPS для 96 центров агрегации для сети из 4512 станций GPS на территории США и для 20 центров агрегации для сети из 2122 станций GPS на территории Европы. Оценки во временном окне длиной 100 суток со смещением 10 суток.



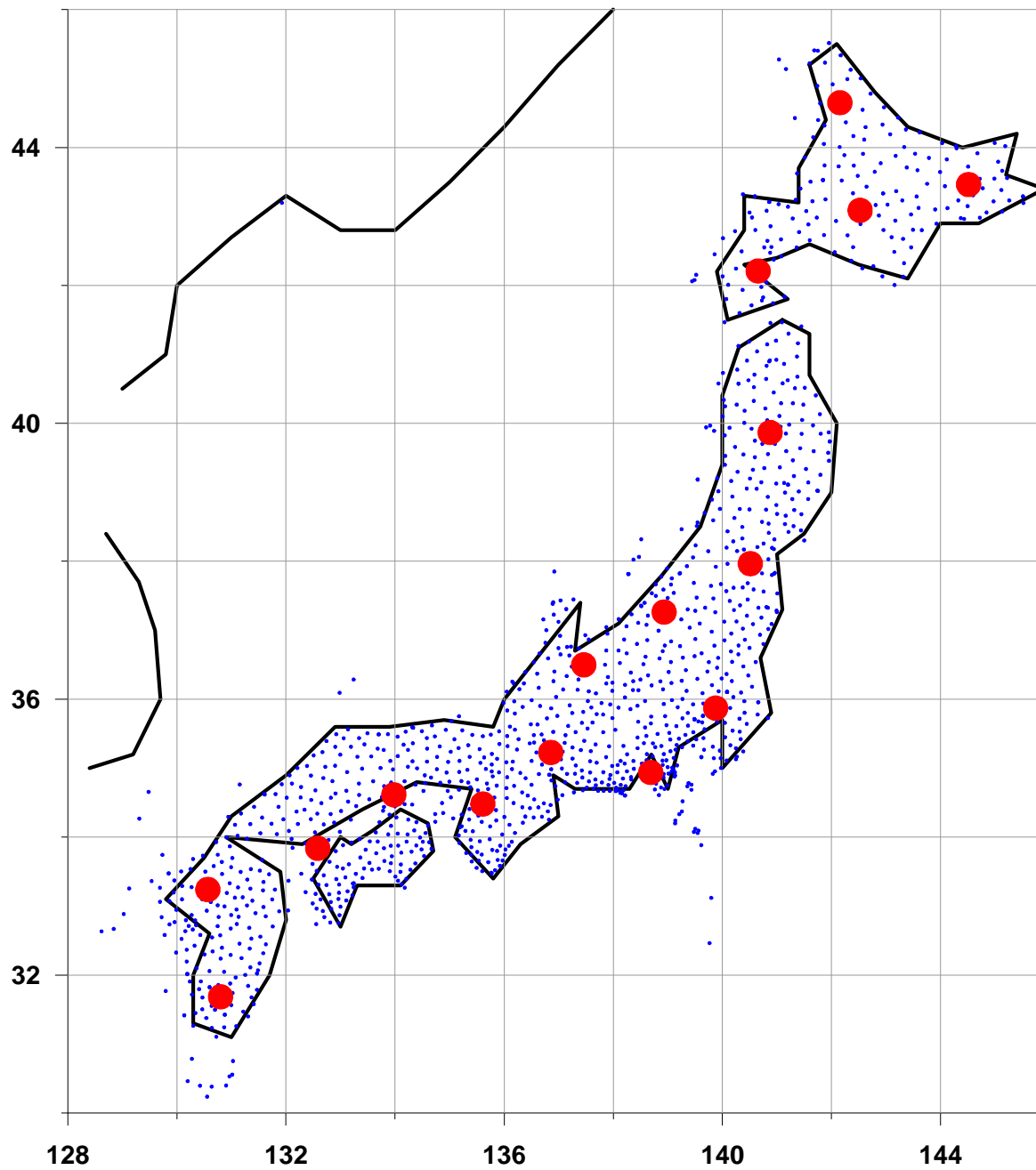
Агрегированные сигналы средних свойств иерархических агрегированных сигналов GPS в США и в Европе



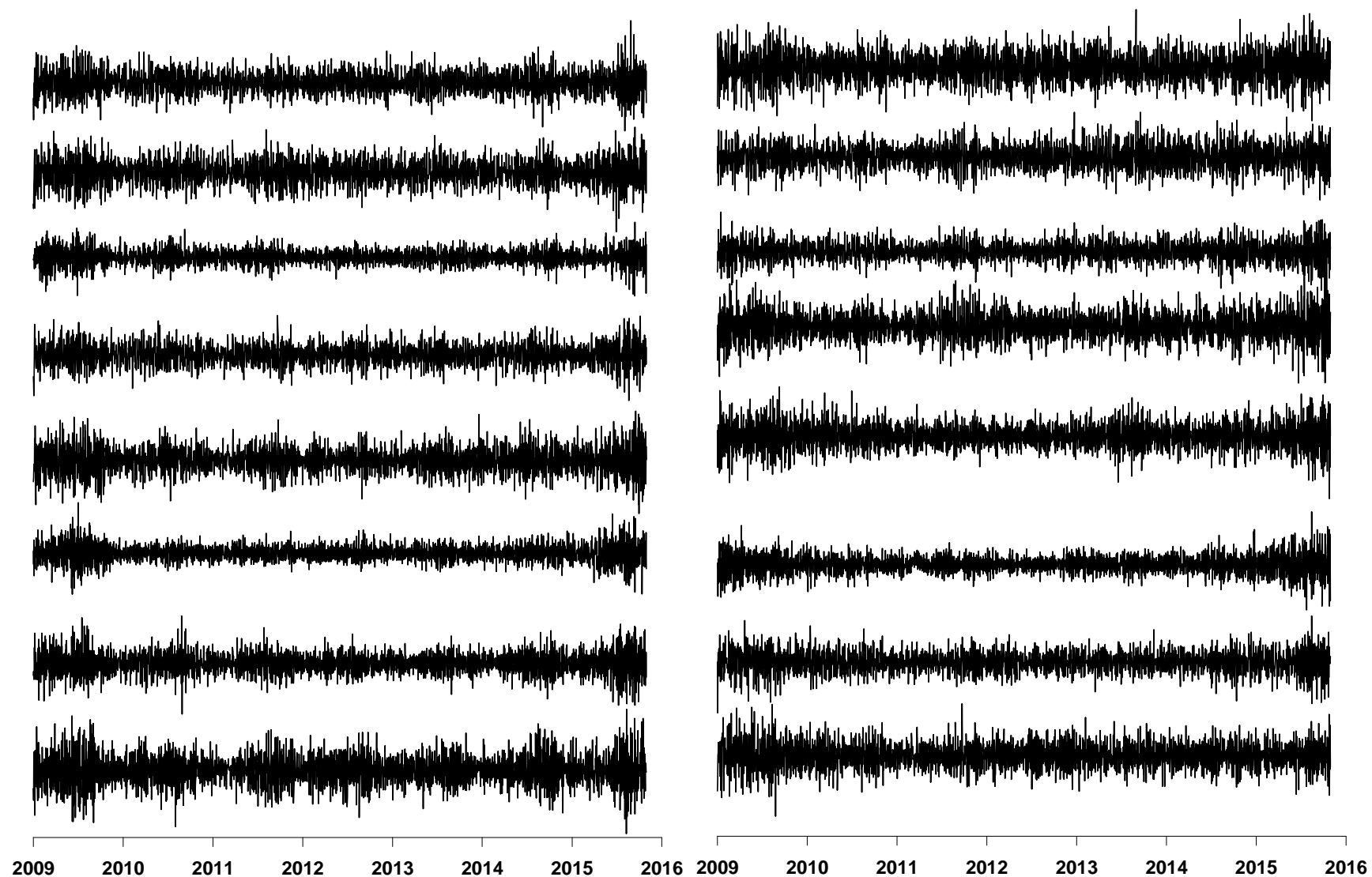
Квадрат коэффициента корреляции между агрегированными сигналами средних свойств в окне длиной 1000 суток



Сеть 1270 GPS станций в Японии, 16 центров агрегации

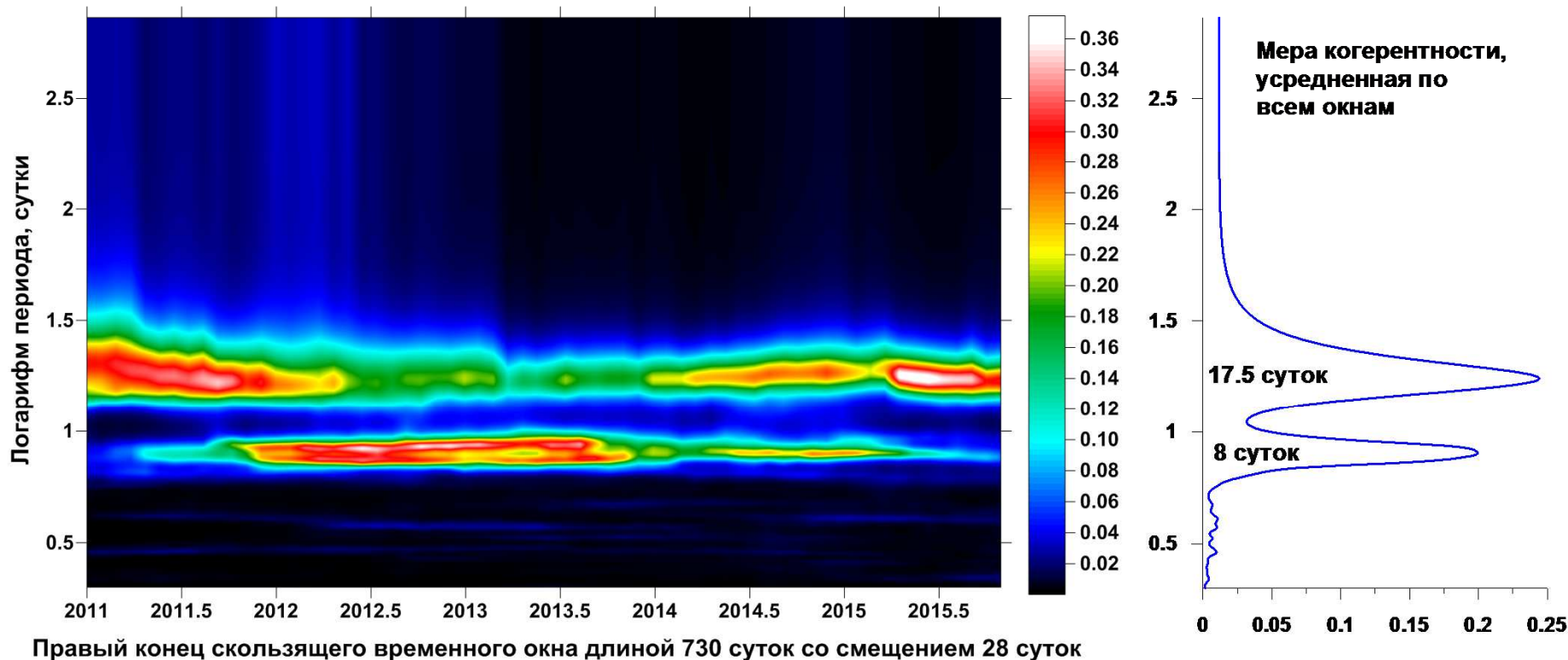


Фурье-агрегированные сигналы для каждого из 16 центров агрегации приращений вертикальных компонент ежесуточных временных рядов GPS на Японских островах. Усреднение оценок во временных окнах длиной 730 суток со смещением 28 суток. Агрегация в каждом окне от 20 ближайших работоспособных станций.

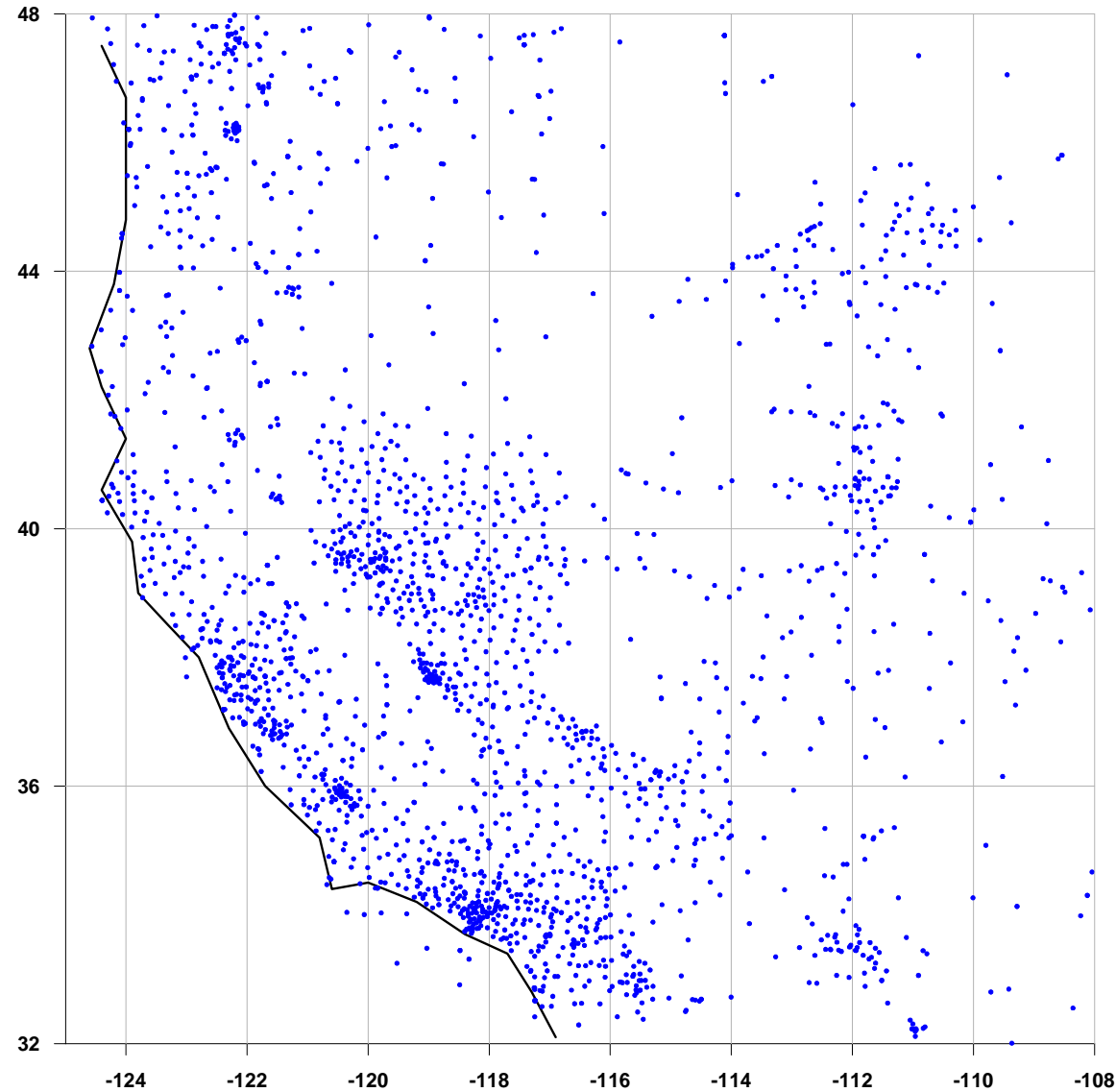


На частотно-временной диаграмме на этом слайде мы опять видим максимум когерентности на периодах 17-18 суток, аналогичный всплескам на этих периодах на диаграммах обобщенной когерентности шумов GPS в Северной Америке и Европе. Таким образом, в Японии мы наблюдаем проявления того же механизма синхронизации шумов, что и на территории США и в Европе.

Частотно-временная диаграмма эволюции множественной спектральной меры когерентности между агрегированными сигналами приращений ежесуточных временных рядов GPS для вертикальных смещений для 16 центров агрегации сети GPS станций в Японии



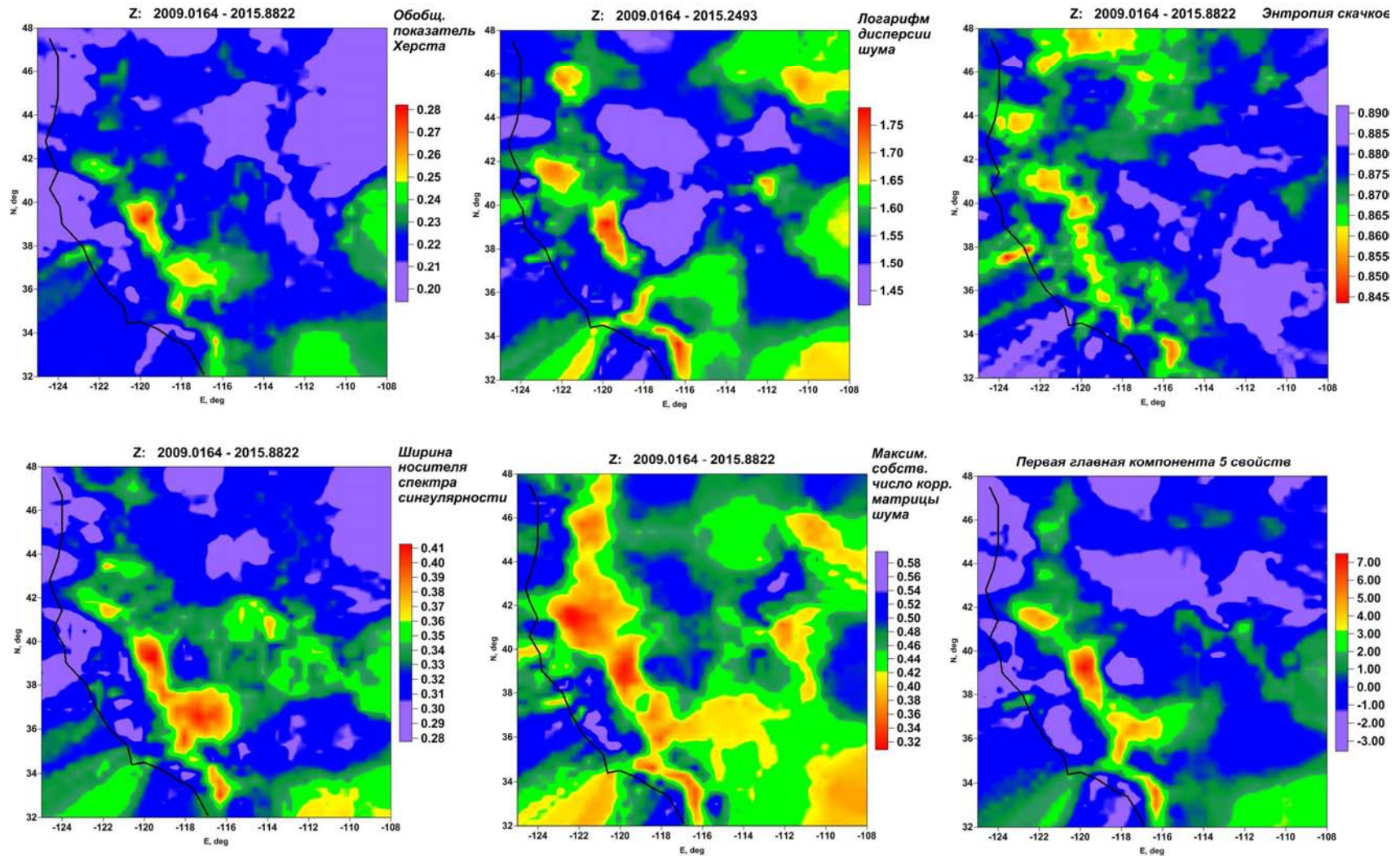
Запад США, 2176 станций GPS – естественный полигон для разработки методов поиска зон интенсивных «медленных движений» с помощью анализа свойств временных рядов GPS.



Усредненные карты свойств шума вертикальной компоненты ежесуточных временных рядов GPS для данных сети 2176 станций GPS на западе США. Усреднение от оценок в скользящих временных окнах длиной 730 суток со смещением 7 суток. Для каждого узла карты вычислялись медианы от 10 ближайших станций работоспособных в текущем окне.

Зона интенсивных медленных движений земной коры выделяется независимо различными методами как:

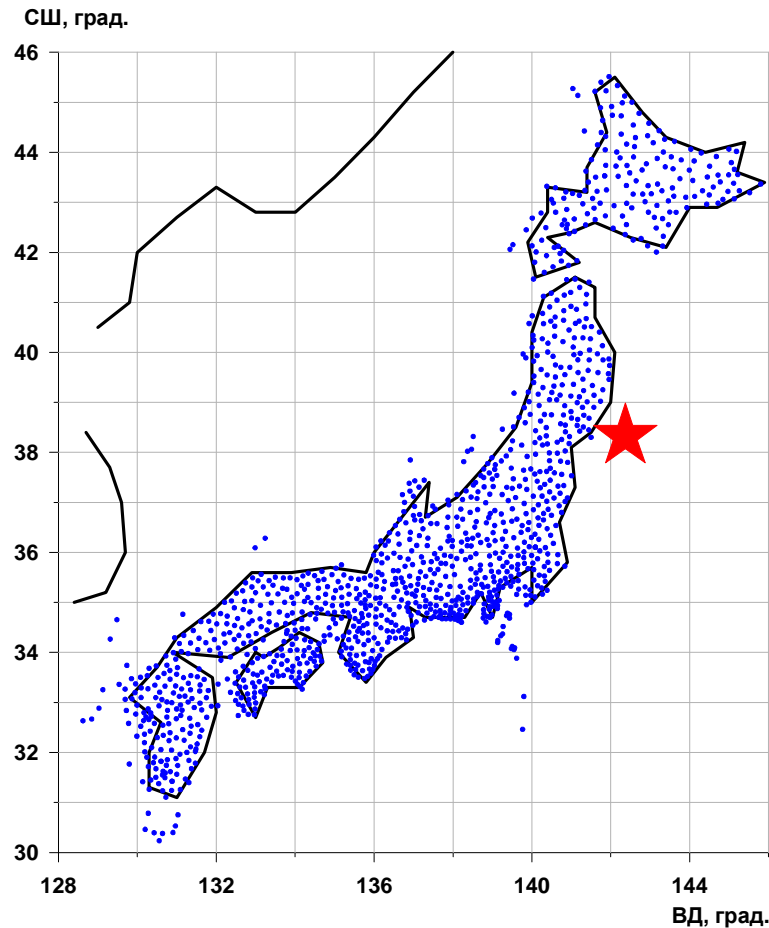
- 1) зона повышенных значений обобщенного показателя Херста;
- 2) зона повышенных значений ширины носителя спектра сингулярности;
- 3) зона пониженных значений максимального нормализованного собственного числа корр. матрицы шума от заданного числа ближайших станций;
- 4) зона повышенных значений логарифма дисперсии шума;
- 5) зона пониженных значений нормализованной энтропии скачкообразной компоненты временных рядов GPS, оцененной с помощью псевдопроизводной.



Анализ данных GPS в Японии за период наблюдений с 30 января 2011 г. по 26 марта 2011 г.

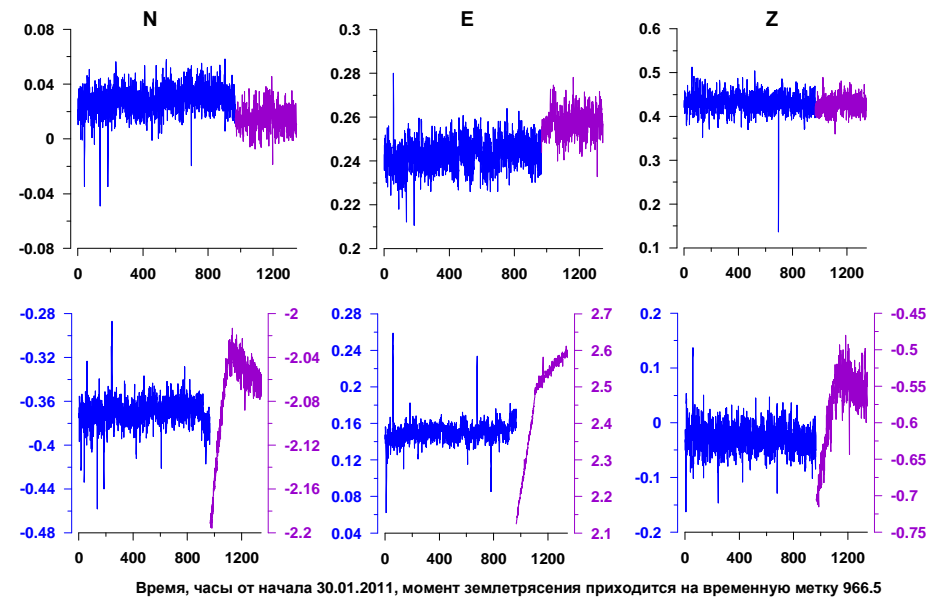
Шаг по времени – 30 минут, до землетрясения 11 марта 2011 г. – 1932 отсчета (чуть более 40 суток), после – 756 отсчетов (почти 16 суток).

Информация свободно доступна по адресу: <http://quakesim.org/tools/timeseries>



Примеры графиков временных рядов GPS:

На станции, удаленной от эпицентра



На станции, близкой к эпицентру. Пришлось использовать разные оси ординат для построения графиков до и после события

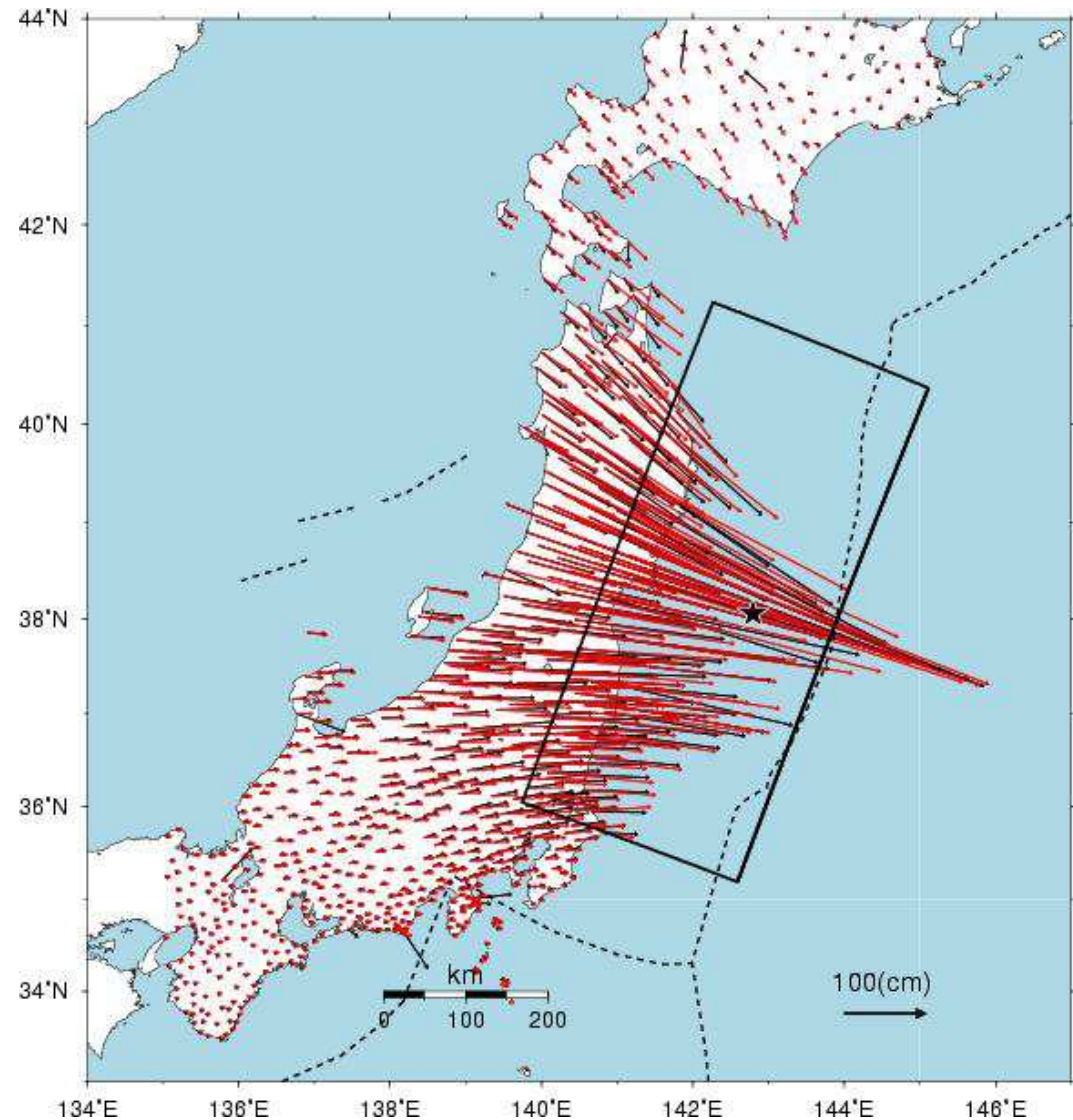
Положения сети GPS станций на Японских островах.
Красная звезда – положение эпицентра землетрясения Тохоку.

Мега-землетрясение Тохоку, 11 марта 2011 года, $M = 9.0$.

После землетрясения Тохоку возник вопрос: «Есть ли толк от GPS для прогноза?»

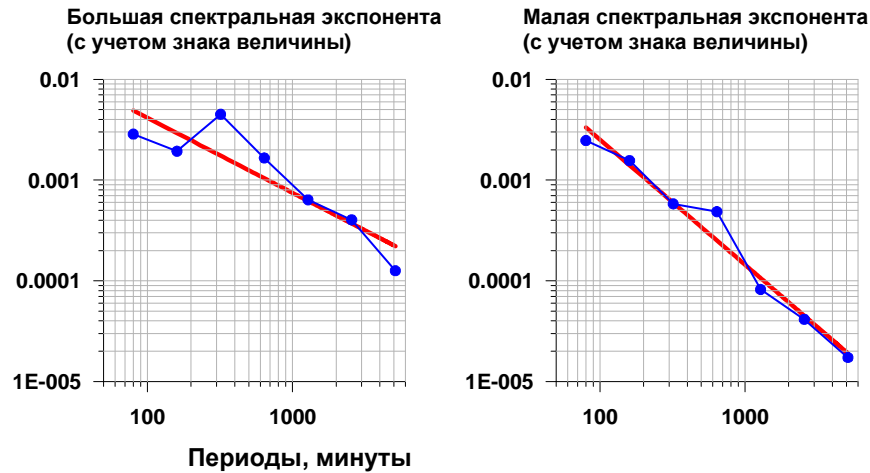
Сеть GPS станций в Японии является уникальной по своей густоте и длительности наблюдений, но создается впечатление, что для прогноза она бесполезна.

Это впечатление не соответствует истине – дело в том, как данные GPS анализируются.



Поле горизонтальных смещений после землетрясения по данным сети GPS

Вейвлет-спектры мощности GPS сигналов после перехода к приращениям



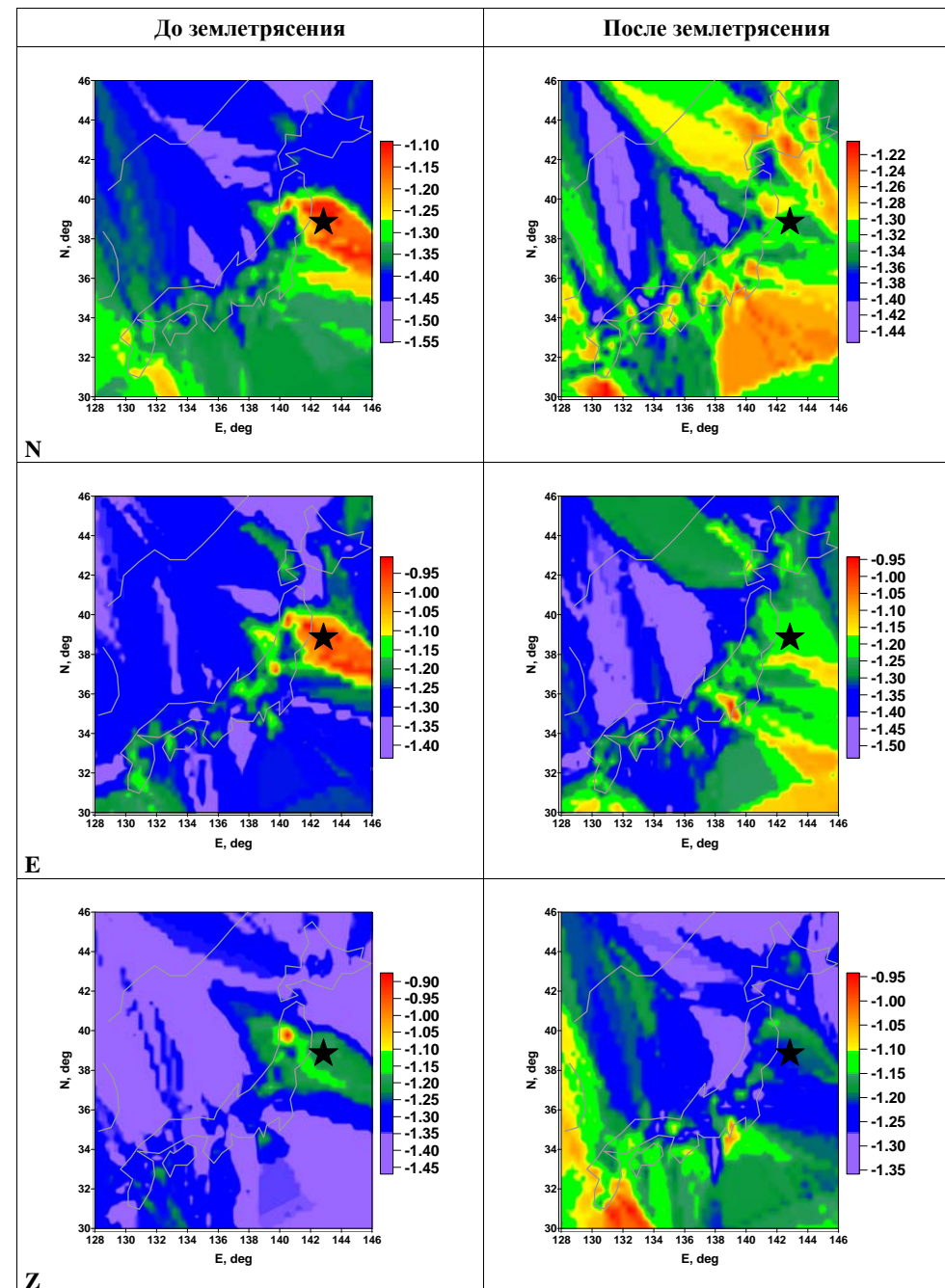
Спектральная экспонента – угол наклона линейного закона, аппроксимирующего зависимость логарифма среднего значения квадратов вейвлет-коэффициентов на уровнях детальности ортогонального вейвлет-разложения («энергии уровня») от логарифма периода, соответствующего центральной частоте уровня.

Карты спектральной экспоненты →

В каждом узле регулярной сетки размером 50×50 узлов бралось среднее от значений спектральной экспоненты в 10 ближайших станциях.

После удаления линейного тренда и перехода к приращениям.

Область будущего мега-землетрясения выделяется как пятно повышенных значений спектральной экспоненты

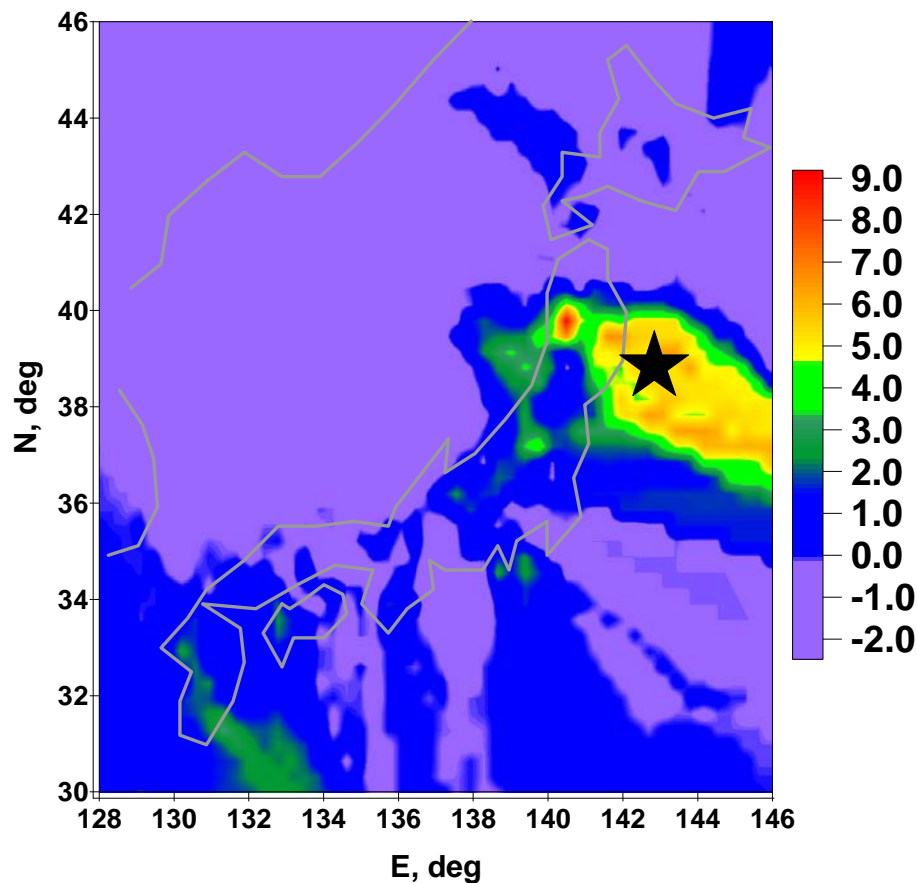


Первая главная компонента значений спектральной экспоненты для флуктуаций компонент GPS-сигналов на регулярной сетке карты после центрирования и нормировки.

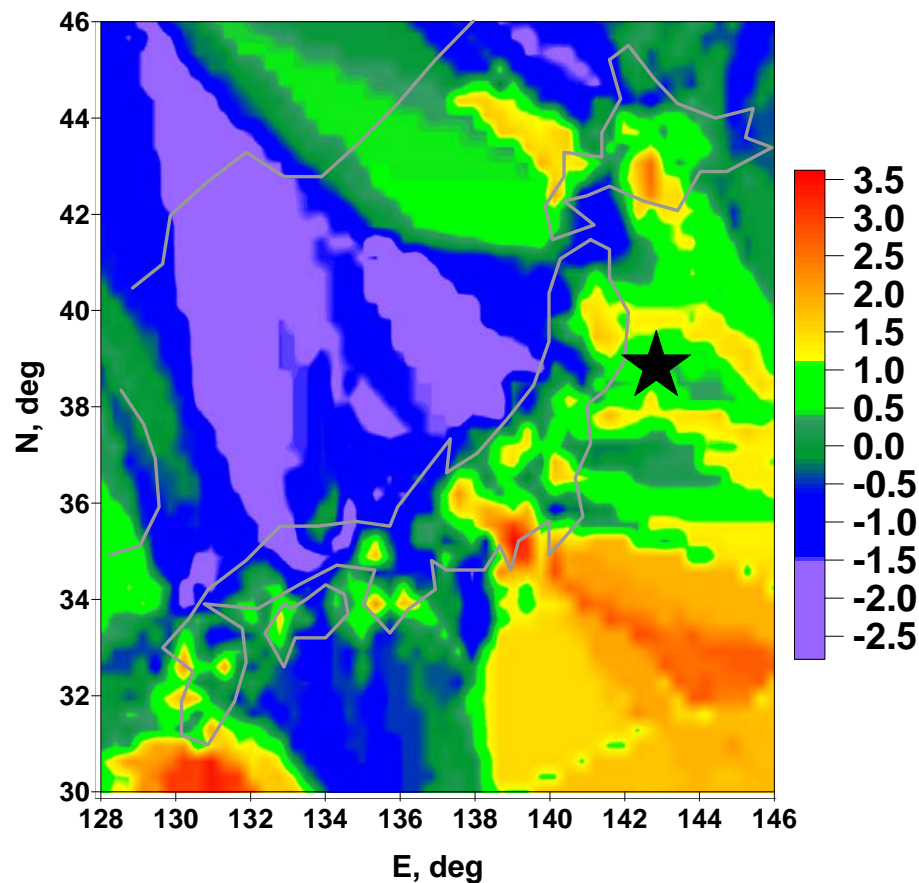
Оказалось, что для прогноза важен не традиционный анализ: куда и с какой скоростью плиты ползут, а то, как они «дрожат».

Alexey Lyubushin and Pavel Yakovlev (2014) Properties of GPS noise at Japan islands before and after Tohoku mega-earthquake. SpringerPlus 2014, 3:364 doi:10.1186/2193-1801-3-364, <http://www.springerplus.com/content/3/1/364>, <http://www.springerplus.com/content/pdf/2193-1801-3-364.pdf> <http://link.springer.com/article/10.1186%2F2193-1801-3-364>

До землетрясения

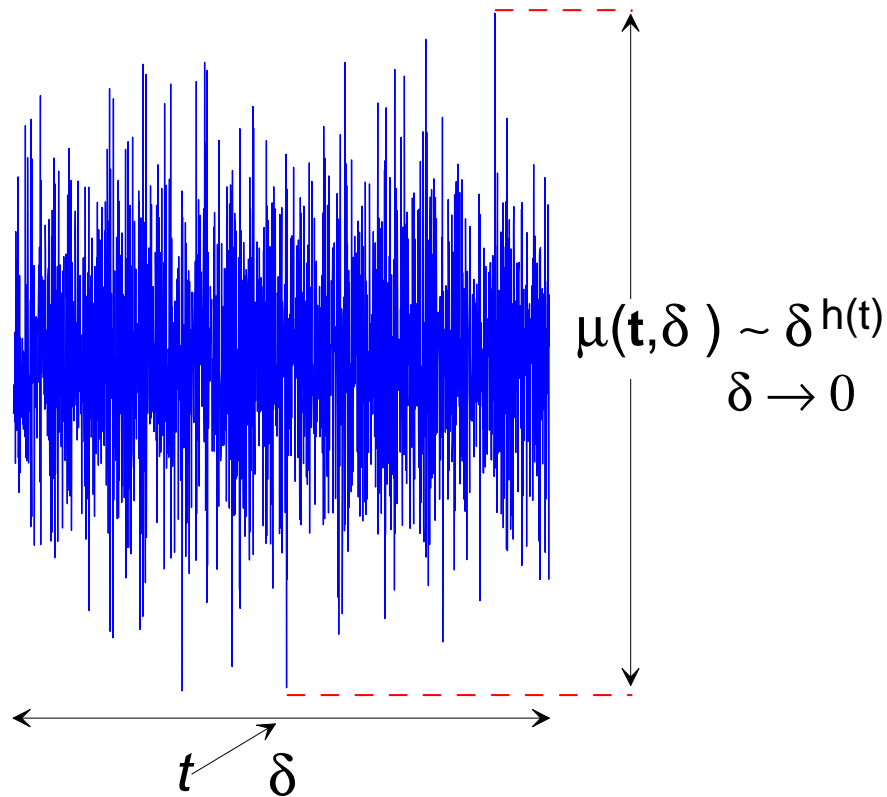


После землетрясения

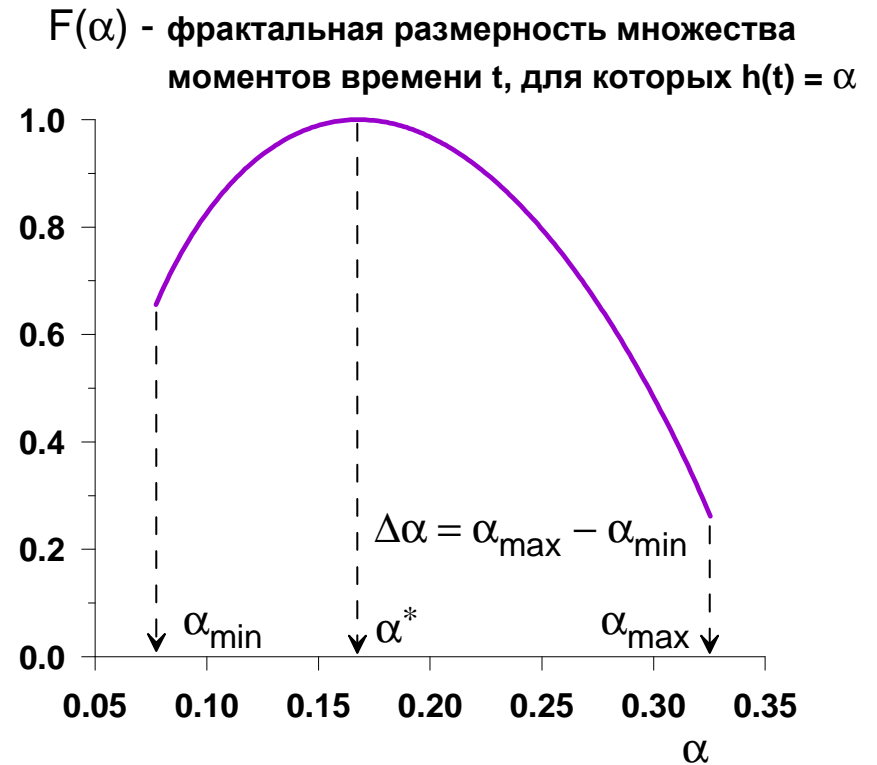


Мульти-фрактальный спектр сингулярности

Мера изменчивости случайного сигнала на интервале времени $[t - \delta/2, t + \delta/2]$



Мультифрактальный спектр сингулярности $F(\alpha)$ и его параметры: $\Delta\alpha$ - ширина носителя и α^* - обобщенный показатель Херста.



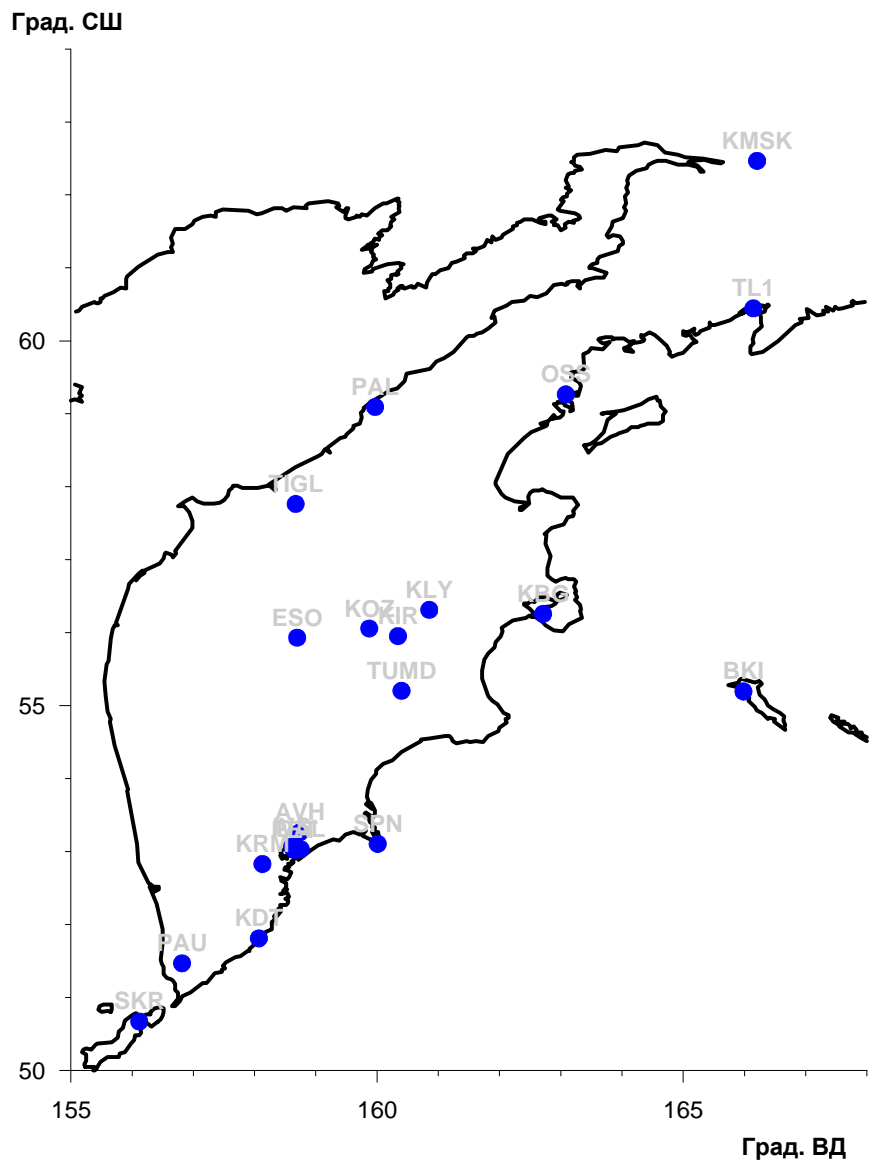
Минимальная нормализованная энтропия :

$$En = - \sum_{k=1}^N p_k \cdot \log(p_k) / \log(N) \rightarrow \min$$

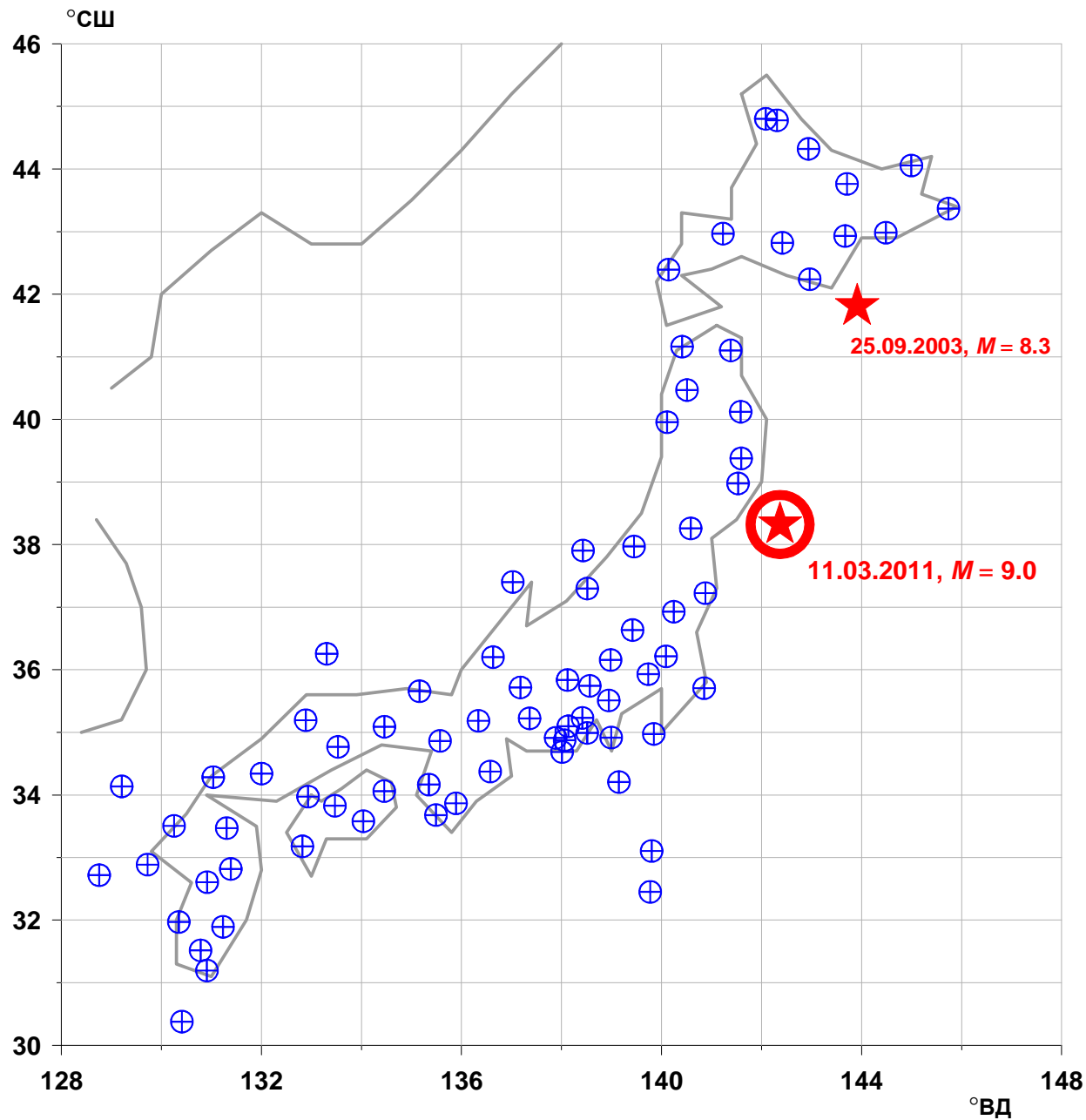
$$0 \leq En \leq 1, \quad p_k = c_k^2 / \sum_{j=1}^N c_j^2,$$

c_j - ортогональные вейвлет-коэффициенты.

Минимум берется по вейвлетам Добеши с числом обнуляемых моментов от 1 до 10.

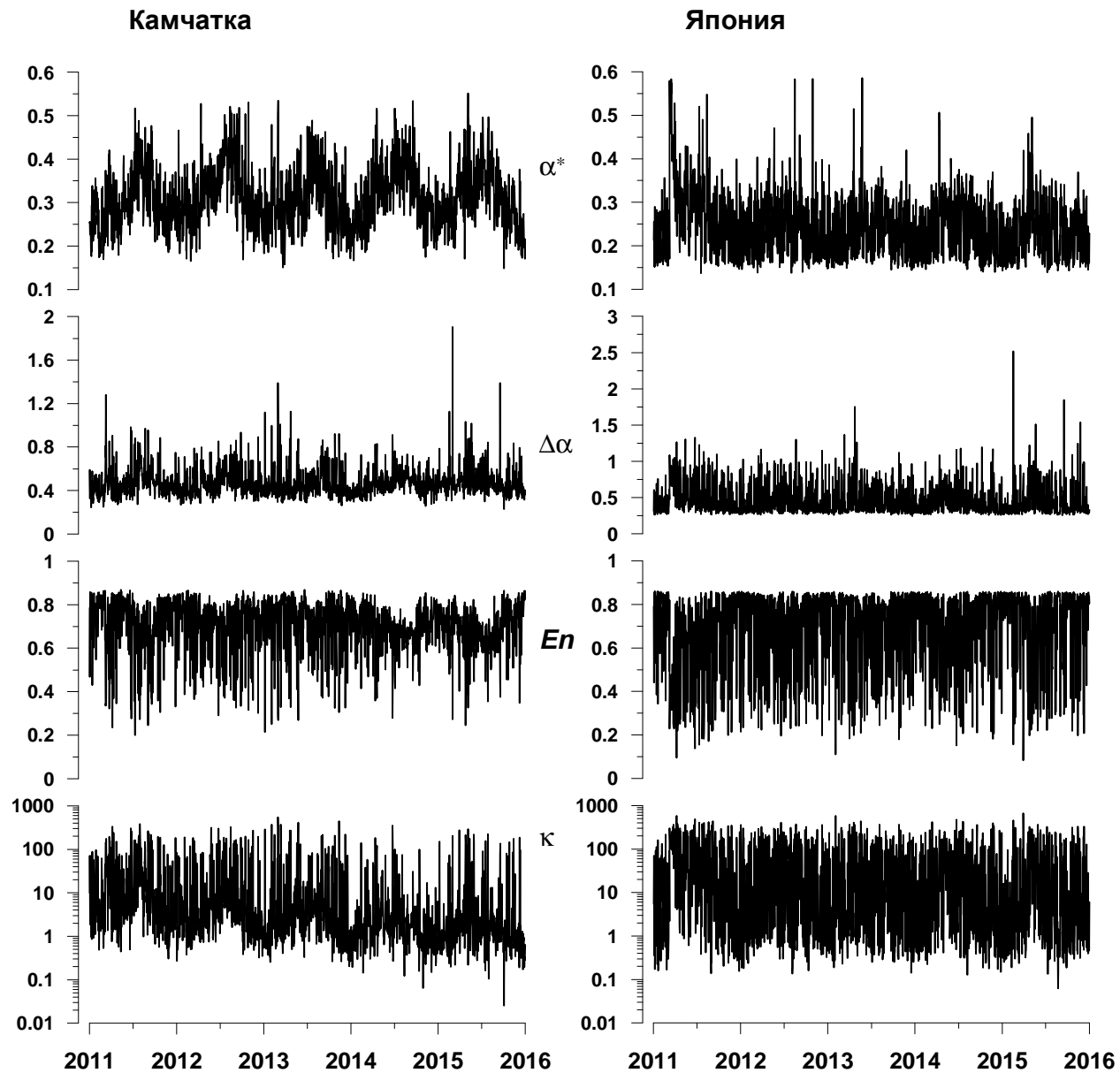


Положения 21 сейсмической станции широкополосной сети КФ ГС РАН на Камчатке



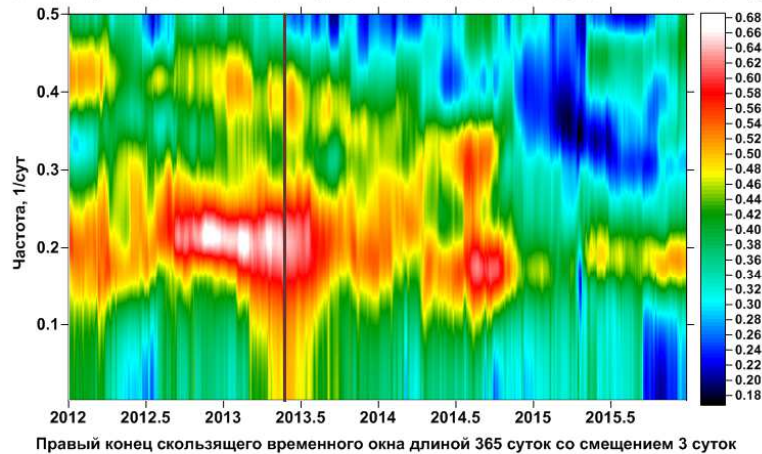
**Положения 78
широкополосных
сейсмических станций
сети F-net на Японских
островах.**

Временные ряды ежесуточных значений медиан характеристик сейсмического шума на Камчатке и в Японии за 2011-2015 гг. :, мульти-фрактальные параметры, минимальная нормализованная энтропия вейвлет-коэффициентов и коэффициент эксцесса.

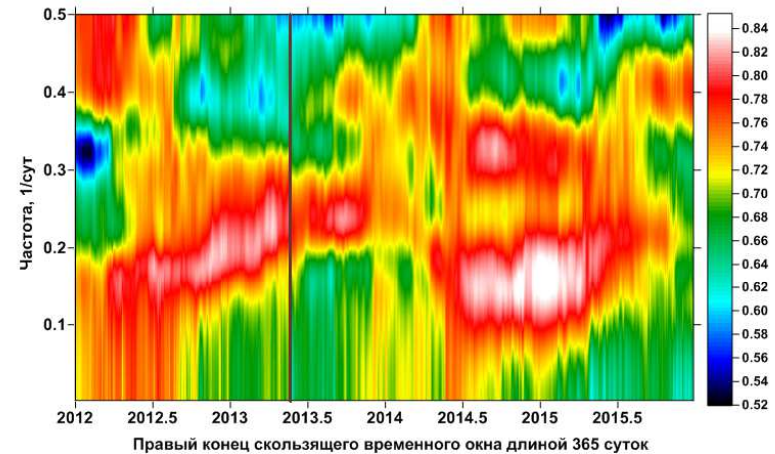


Были ли предвестники Охотоморского мантийного землетрясения 24 марта 2013 года в данных сейсмического шума на Камчатке и в Японии? – Да.

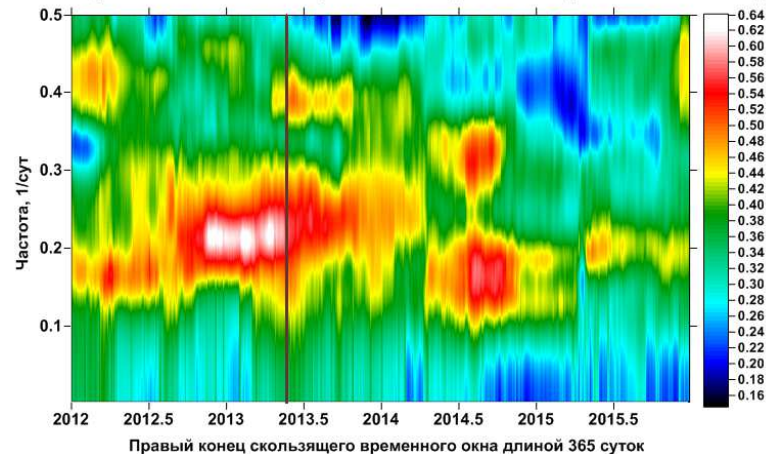
Множественный спектр когерентности между медианами 4 ежесуточных значений свойств сейсмического шума от сети широкополосных сейсмических станций на Камчатке. Список свойств: обобщенный показатель Херста, ширина носителя спектра сингулярности, минимальная нормализованная энтропия квадратов ортогональных вейвлет-коэффициентов и коэффициент эксцесса. Оценка в скользящем временном окне длиной 365 суток со смещением 3 суток, AR(5)-модель, переход к приращениям. Вертикальная коричневая линия - момент времени Охотоморского землетрясения 24 мая 2013 года



Множественный спектр когерентности между медианами 4 ежесуточных значений свойств сейсмического шума от сети F-net широкополосных сейсмических станций в Японии. Список свойств: обобщенный показатель Херста, ширина носителя спектра сингулярности, минимальная нормализованная энтропия квадратов ортогональных вейвлет-коэффициентов и коэффициент эксцесса. Оценка в скользящем временном окне длиной 365 суток, AR(5)-модель, переход к приращениям. Вертикальная коричневая линия - момент времени Охотоморского землетрясения 24 мая 2013 года



Множественный спектр когерентности между медианами 8 ежесуточных значений свойств сейсмического шума от сети широкополосных сейсмических станций на Камчатке и от сети широкополосных станций F-net на Японских островах: по 4 свойства от каждой сети. Список свойств: обобщенный показатель Херста, ширина носителя спектра сингулярности, минимальная нормализованная энтропия квадратов ортогональных вейвлет-коэффициентов и коэффициент эксцесса. Оценка в скользящем временном окне длиной 365 суток, AR(5)-модель, переход к приращениям. Вертикальная коричневая линия - момент времени Охотоморского землетрясения 24 мая 2013 года



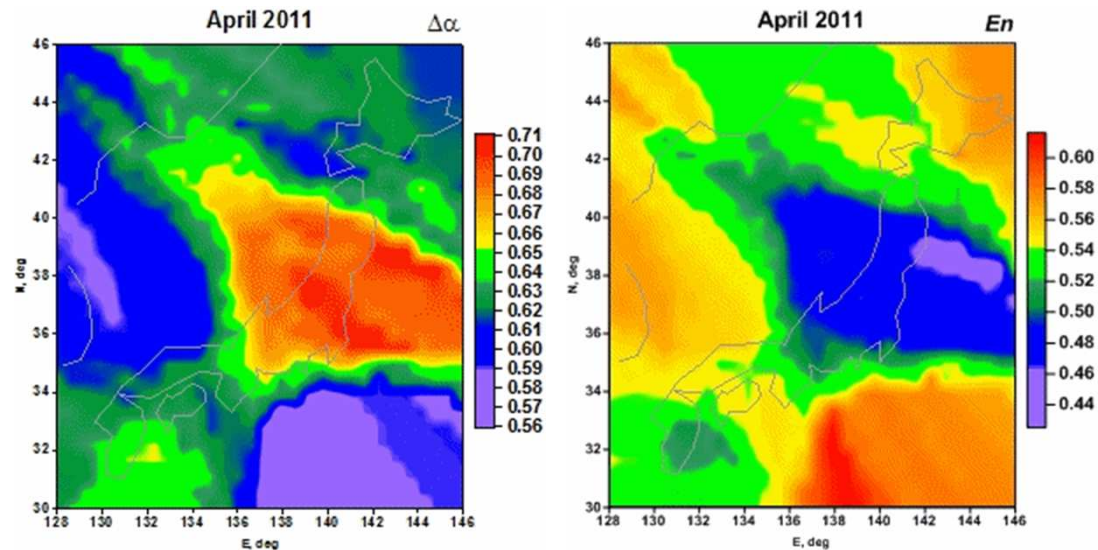
Всплеск множественной когерентности свойств сейсмического шума перед Охотоморским землетрясением с середины 2012 г. на периодах около 5 суток

**Традиционный вопрос для Японии: где и когда
будет следующее мега-землетрясение?**

Естественные флуктуации оценок сейсмической опасности на Японских островах после 11.03.2011 по свойствам сейсмического шума в окне длиной 30 суток

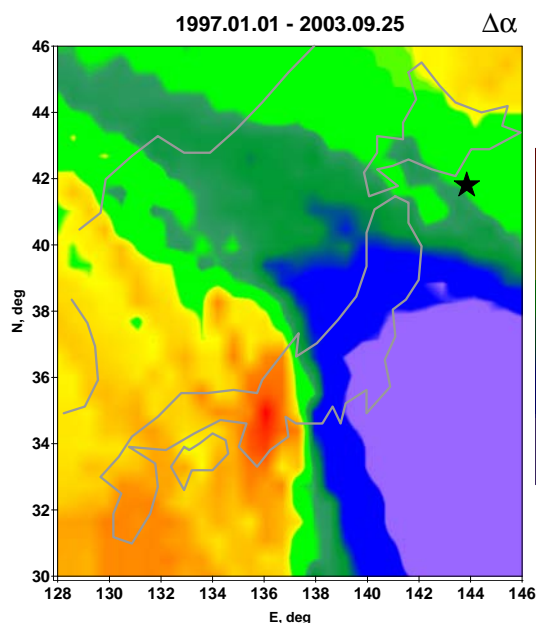


Последовательности ежемесячных карт ширины носителя мульти-фрактального спектра сингулярности и минимальной нормализованной энтропии квадратов вейвлет-коэффициентов волновых форм сейсмического шума на Японских островах после мега-землетрясения Тохоку 11 марта 2011 года. Области пониженных значений ширины носителя и повышенных значений энтропии выделяют регионы, опасные с точки зрения вероятности возникновения следующего мега-землетрясения в Японии

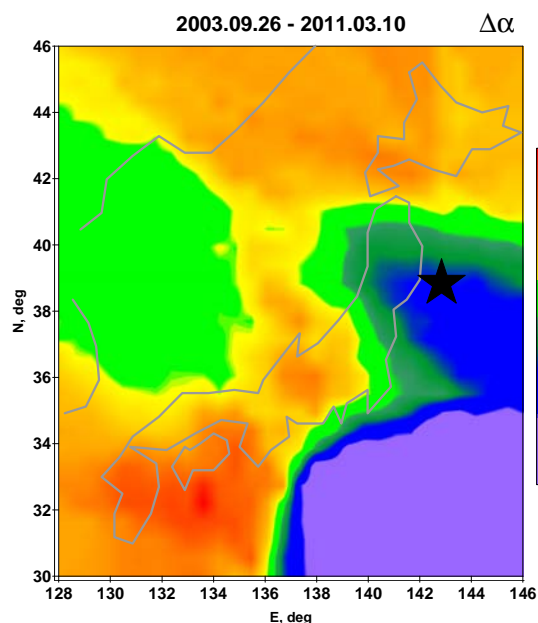


← Gif-анимация

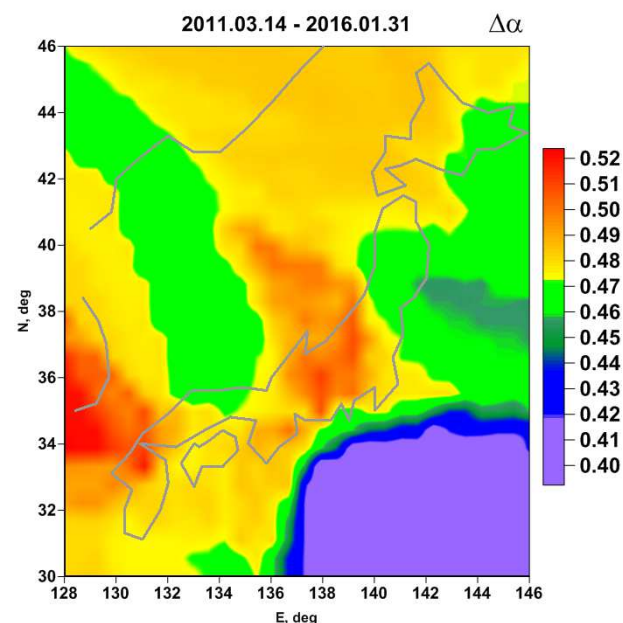
**Карты пространственного распределения ширины носителя спектра сингулярности $\Delta\alpha$.
Области относительно низких значений $\Delta\alpha$ («синие» и «фиолетовые» области) выделяют области повышенной опасности.**



От начала 1997 (начало работы системы F-net) по 25.09.2003: область будущей сейсмической катастрофы характеризуется относительно низкими значениями $\Delta\alpha$ и она еще не разделена на Северную и Южную части.

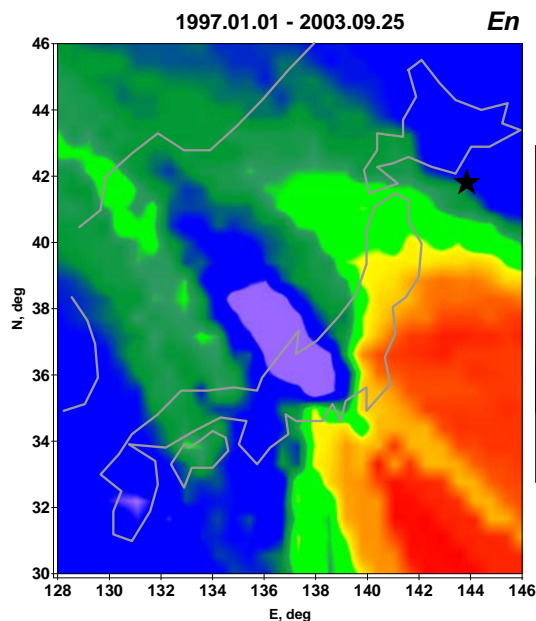


От начала 26.09.2003 по 10.03.2011: область будущей катастрофы 11.03.2011 характеризуется относительно низкими значениями $\Delta\alpha$ и ранее существовавшая единая область низких $\Delta\alpha$ распалась на Северную и Южную части.

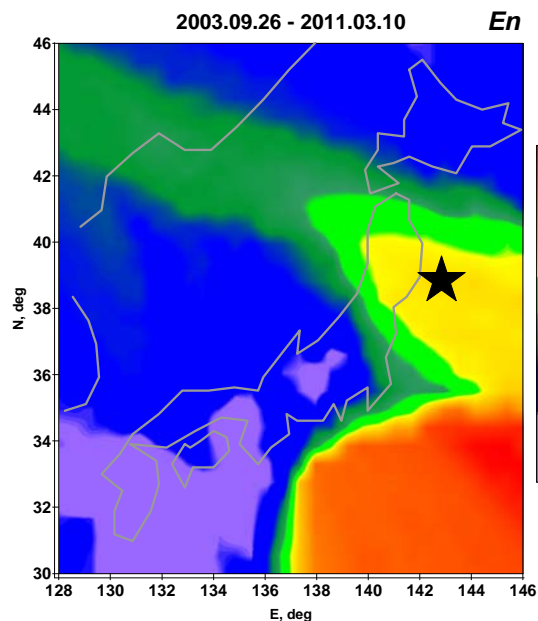


От 14 марта 2011 г. по 31 января 2016 г. включит.: Северная часть низких значений $\Delta\alpha$ реализовалась как область Великого Японского Землетрясения 11.03.2011, $M=9.0$, тогда как Южная часть по прежнему характеризуется относительно низкими значениями $\Delta\alpha$ и представляет опасность как регион будущего землетрясения с магнитудой свыше 8.5.

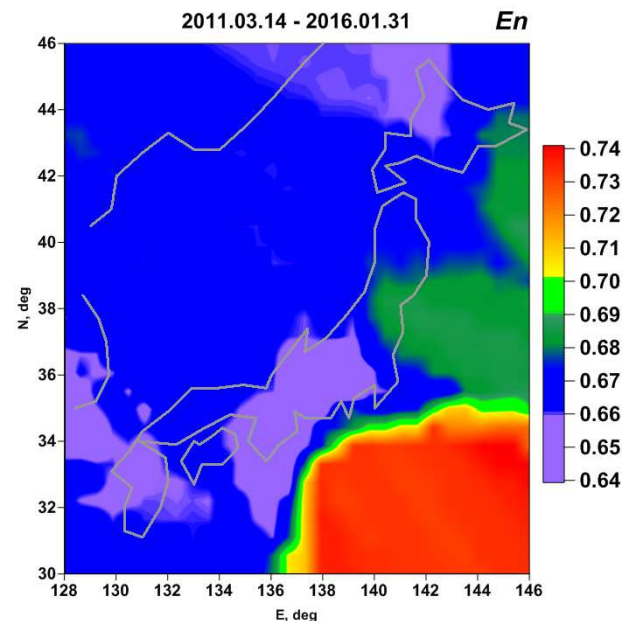
Карты нормализованной энтропии распределения квадратов вейвлет-коэффициентов волновых форм сейсмического шума. Для каждой станции и для каждого дня определялся «наилучший» орт. вейвлет из условия минимума энтропии распределения квадратов его коэффициентов.



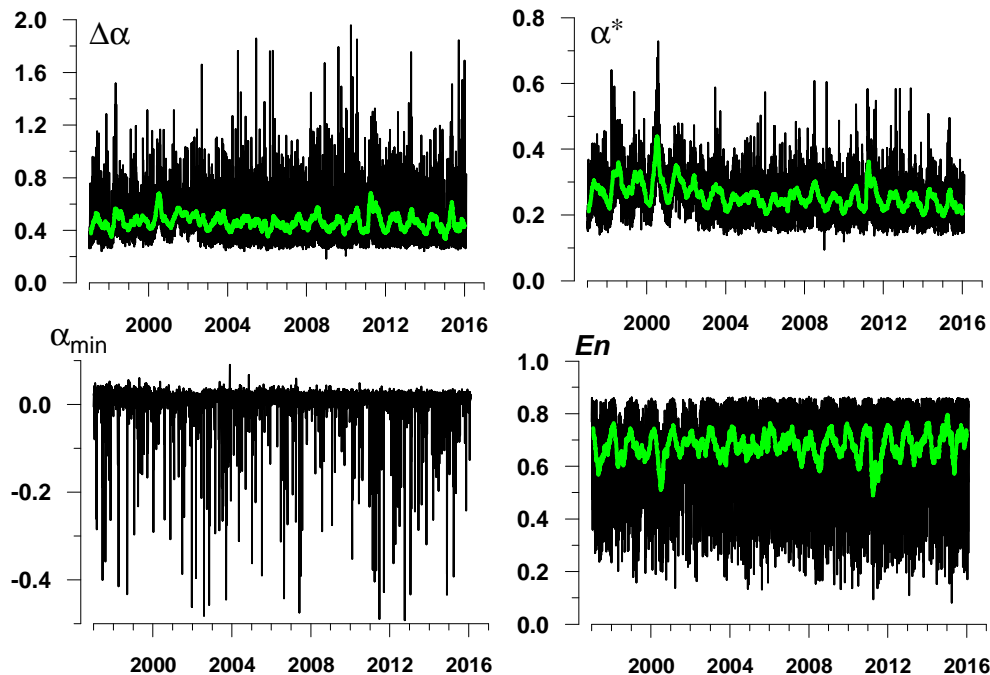
От начала 1997 (начало работы системы F-net) по 25.09.2003: область будущей сейсмической катастрофы характеризуется относительно высокими значениями энтропии и она еще не разделена на Северную и Южную части.



От 26.09.2003 по 10.03.2011: область будущей катастрофы 11.03.2011 характеризуется относительно высокими значениями энтропии и ранее существовавшая единая область больших значений энтропии распалась на две части.



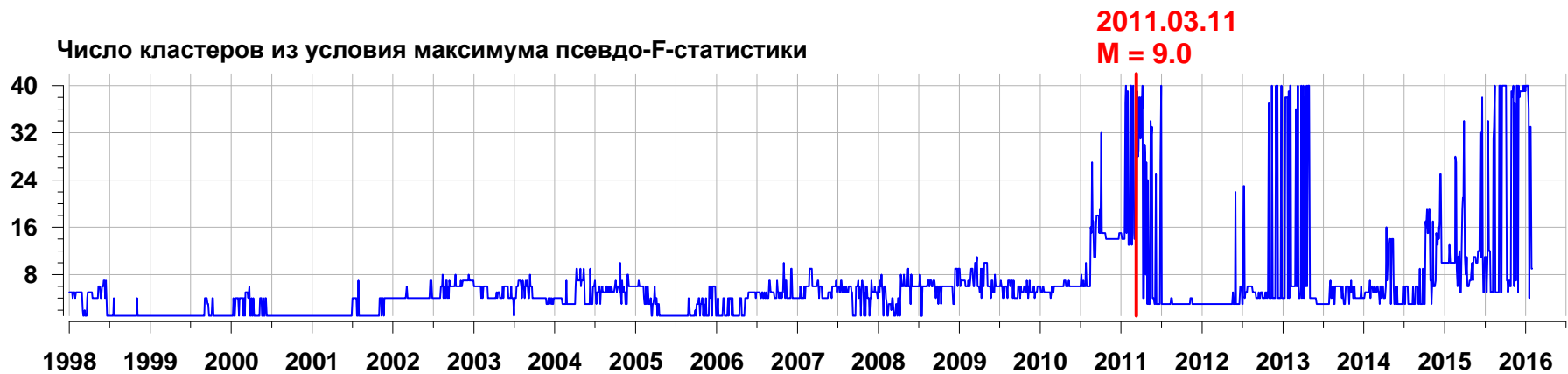
От 14 марта 2011 г. по 31 января 2016 г. включит.: Северная часть высоких значений энтропии реализовалась как область Великого Японского Землетрясения 11.03.2011, $M=9.0$, тогда как Южная часть по прежнему характеризуется относительно высокими значениями энтропии и представляет опасность как регион будущего землетрясения с магнитудой свыше 8.5.



1997 – 31 января 2016

Медианы ежесуточных оценок 4-х параметров сейсмического шума сети F-net:
 $\Delta\alpha$, α^* , α_{\min} - параметры мультифрактального спектра сингулярности;
 En - минимальная нормализованная энтропия квадратов вейвлет-коэффициентов.
 Зеленые линии - скользящие средние в окне длиной 57 суток.

Кластеризация первых 3-х главных компонент значений медиан 4-х параметров сейсмического шума в скользящем временном окне длиной 365 дней со смещением 3 дня.
 Нормализация и винзоризация с пределами $\pm 4\sigma$ в каждом окне.



Правый конец скользящего временного окна длиной 365 дней со смещением 3 дня

Поиск предвестника момента времени следующего мега-землетрясения в Японии.

Для ознакомления с деталями см.:

Lyubushin, A. (2013) How soon would the next mega-earthquake occur in Japan? Natural Science, Vol.5, No.8A1, 1-7.

doi: 10.4236/ns.2013.58A1001. <http://www.scirp.org/journal/PaperInformation.aspx?PaperID=35770>

Псевдо-F-статистика:

$$PFS(q) = \frac{\frac{1}{(q-1)} \sum_{k=1}^q v_k \cdot (z_k - z_0)^2}{\frac{1}{(N-q)} \sum_{k=1}^q \sum_{\zeta \in \Gamma_k} (\zeta - z_k)^2},$$

q - число кластеров, z_k - центр тяжести кластера Γ_k , z_0 - общий центр тяжести
 $v_k = n_k / N$, n_k - число векторов в кластере Γ_k , N - число векторов в облаке

Кластеризация первых 3-х главных компонент значений медиан 4-х параметров сейсмического шума в скользящем временном окне длиной 365 дней со смещением 3 дня.

Нормализация и винзоризация с пределами $\pm 4\sigma$ в каждом окне.

В настоящий момент опасность повторного мега-землетрясения велика

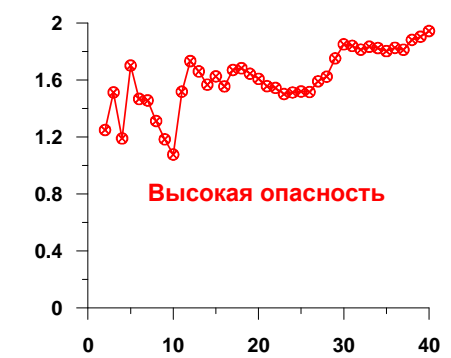
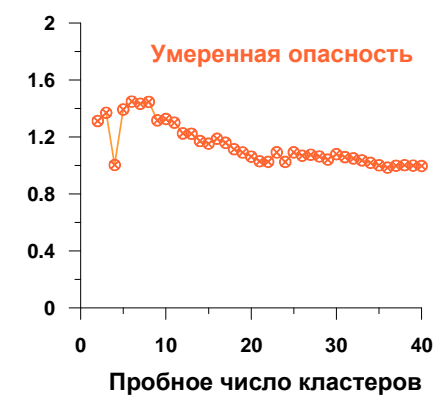
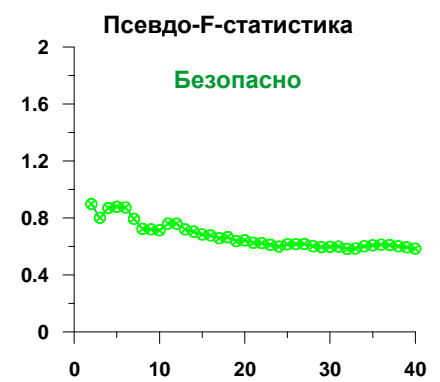
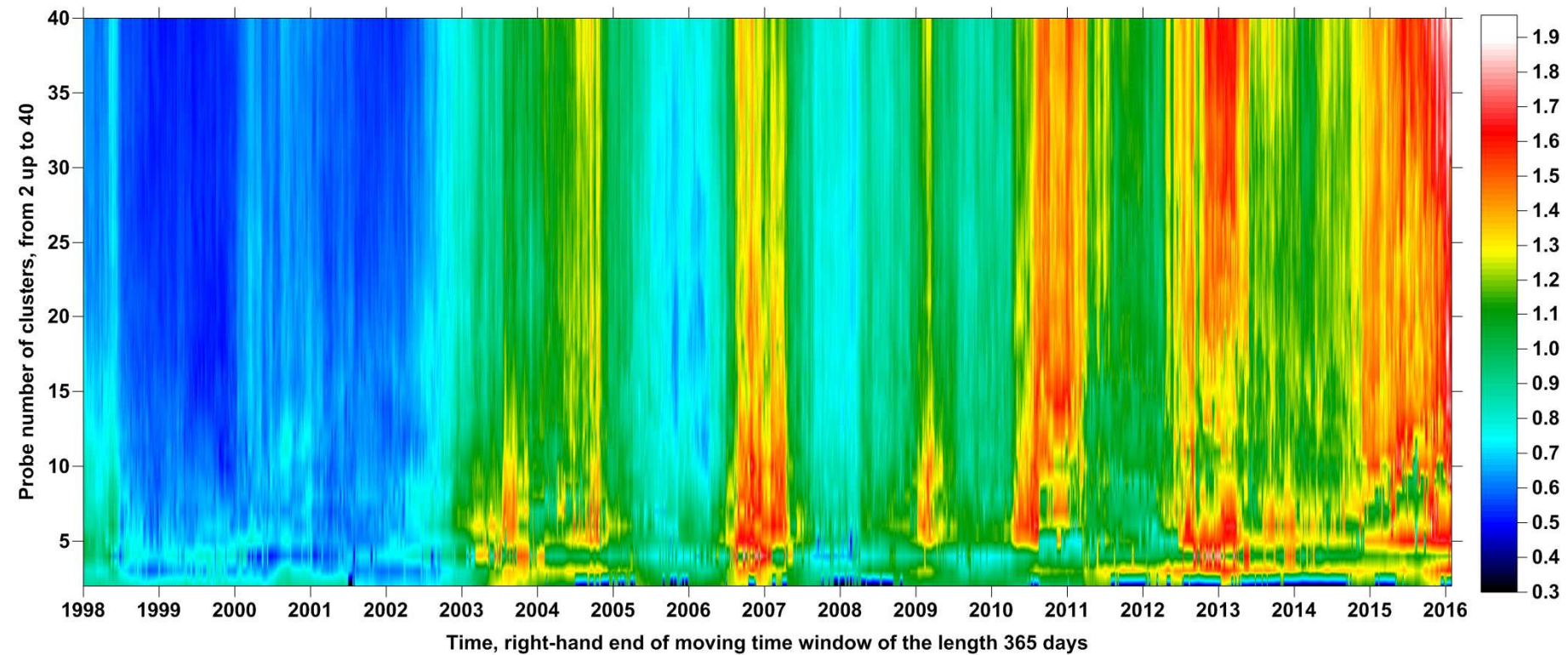
Число кластеров из условия максимума псевдо-F-статистики



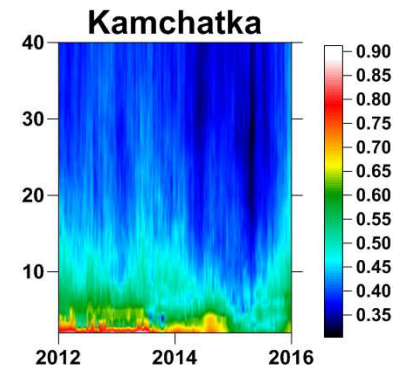
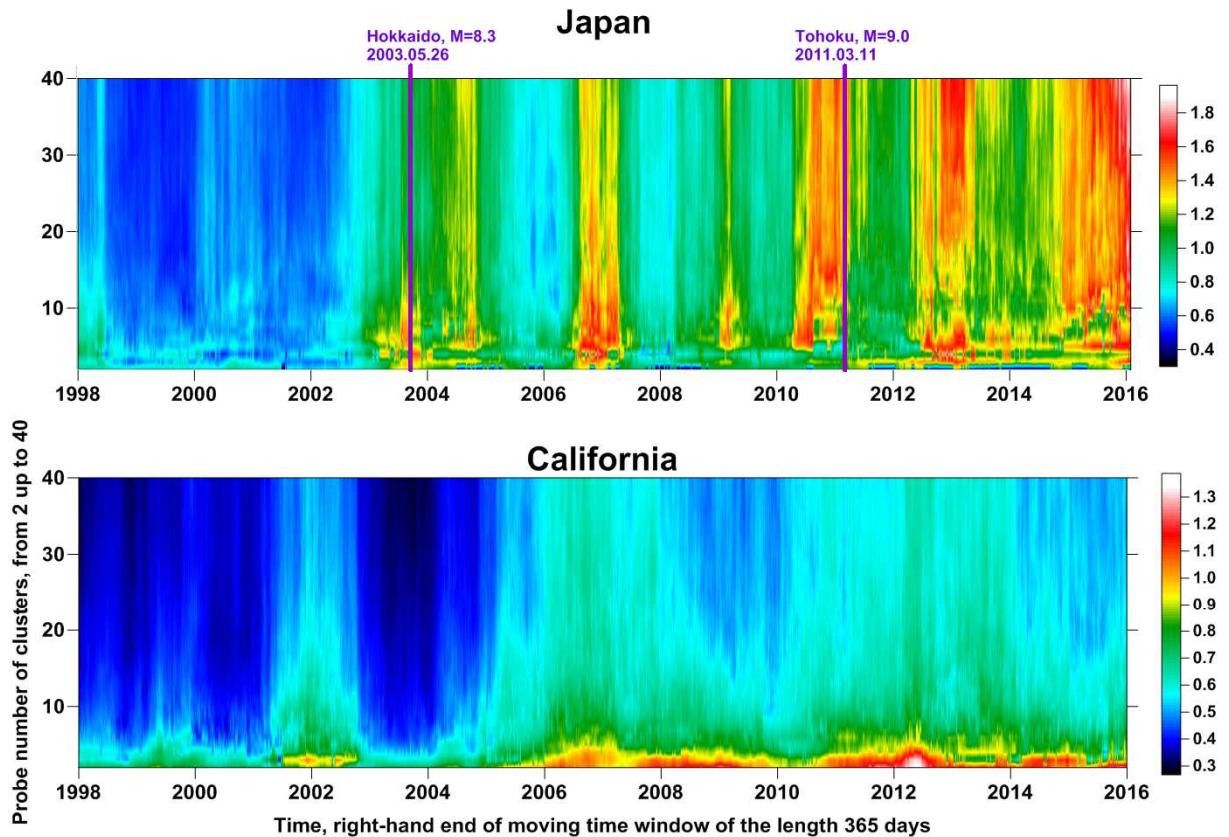
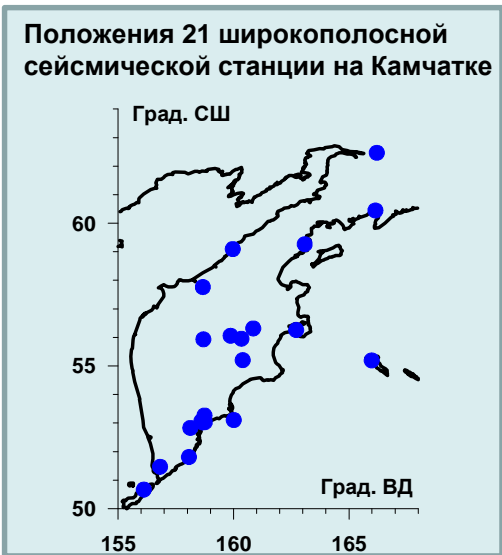
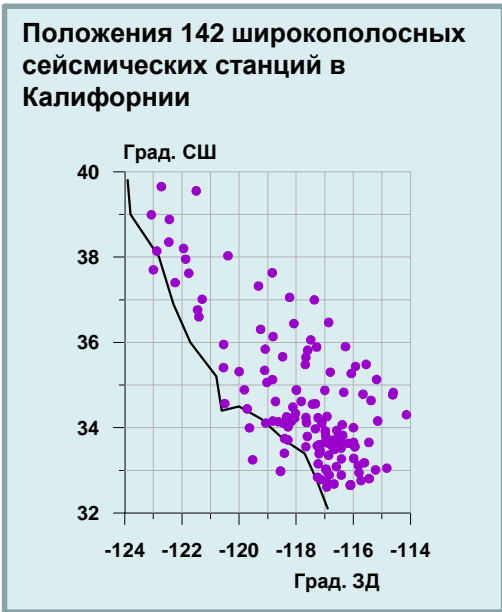
Правый конец скользящего временного окна длиной 365 дней со смещением 3 дня

Карта псевдо-F-статистики как оценка текущей сейсмической опасности

Values of pseudo-F-statistics for clouds of first 3 principal components for median values of 4 parameters of daily seismic noise waveforms: minimum normalized entropy of squared orthogonal wavelet coefficients, minimum Holder-Lipschitz exponent, singularity spectrum support width and generalized Hurst exponent within moving time window of the length 365 days with mutual shift 3 days.



Совместно: карты псевдо-F-статистики для оценки текущей сейсмической опасности по свойствам сейсмического шума в Японии, Калифорнии и на Камчатке.



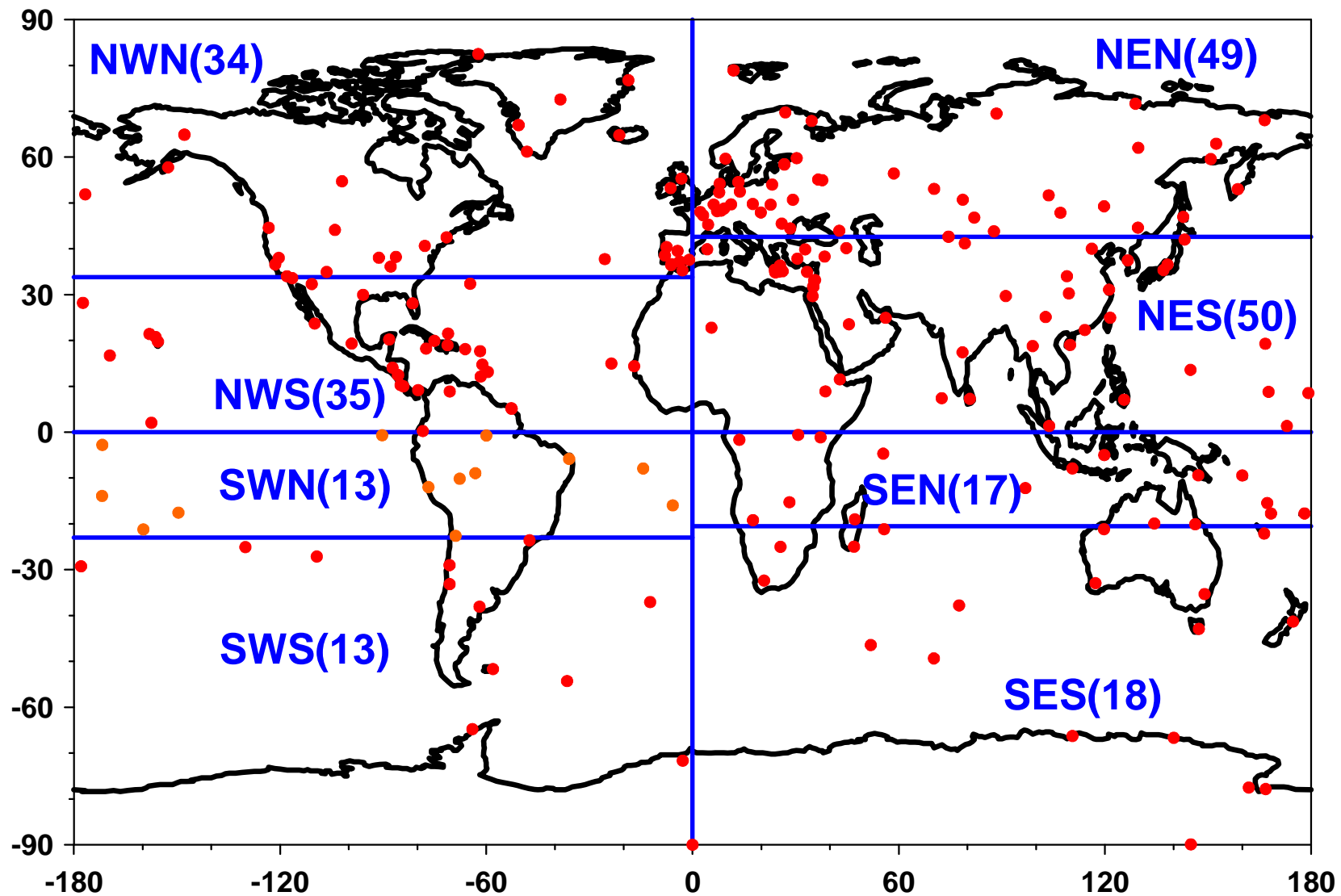
Когерентность свойств глобального поля сейсмических шумов, 1997 – 2015

Любушин А.А. (2014) Анализ когерентности глобального сейсмического шума, 1997-2012. Физика Земли. 2014, №3, С. 18-27.

<http://link.springer.com/article/10.1134%2FS1069351314030069>

Lyubushin A.A. (2015) Wavelet-based coherence measures of global seismic noise properties - Journal of Seismology, April 2015, Volume 19, Issue 2, pp 329-340.

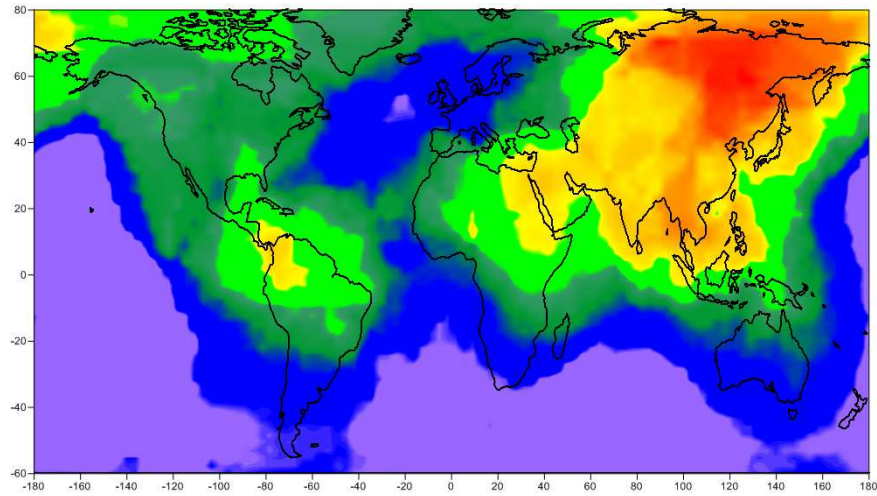
<http://link.springer.com/article/10.1007/s10950-014-9468-6>



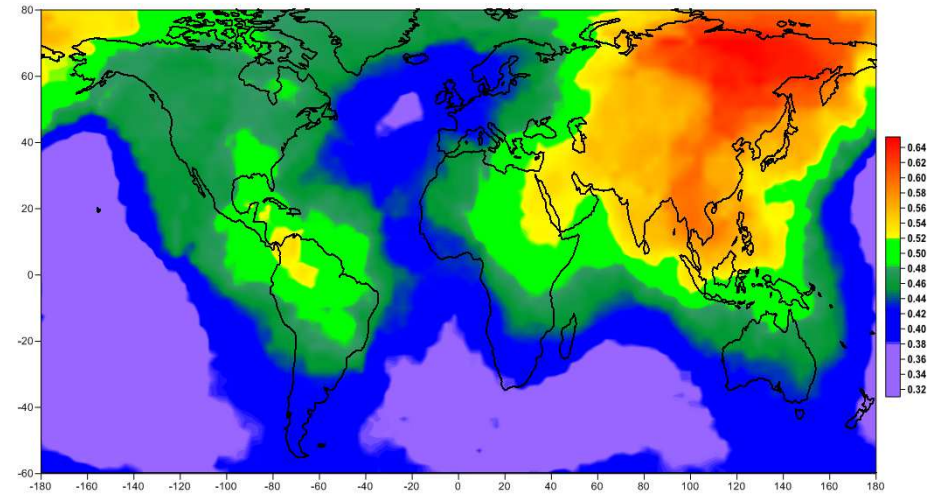
Положения 229 широкополосных станций объединения 3-х глобальных сейсмических сетей и их разбиение на 8 групп. В скобках после имени группы указано число станций в группе.

Усредненные карты свойств глобального сейсмического шума за 19 лет наблюдений, 1997 – 2015

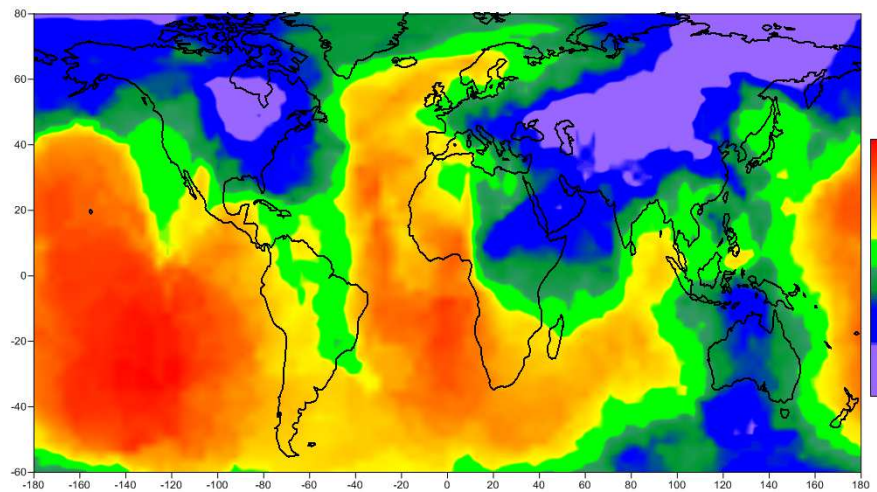
Обобщенный показатель Херста α^*



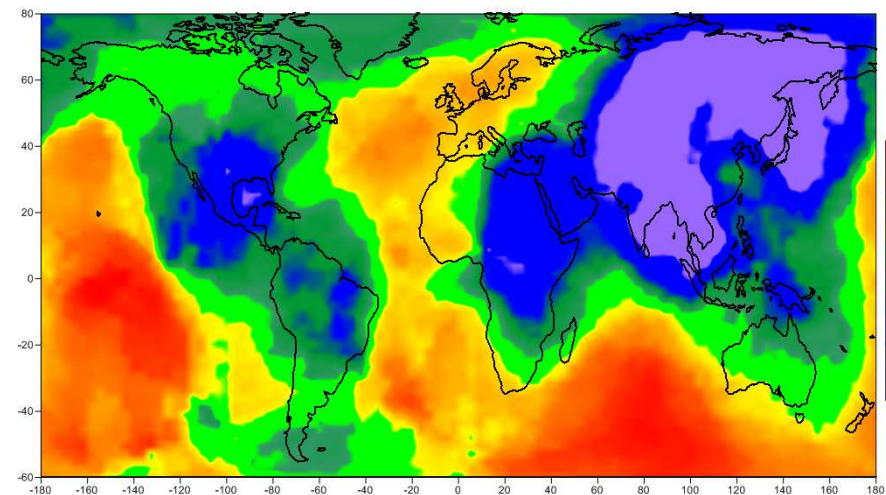
Ширина носителя спектра сингулярности $\Delta\alpha$

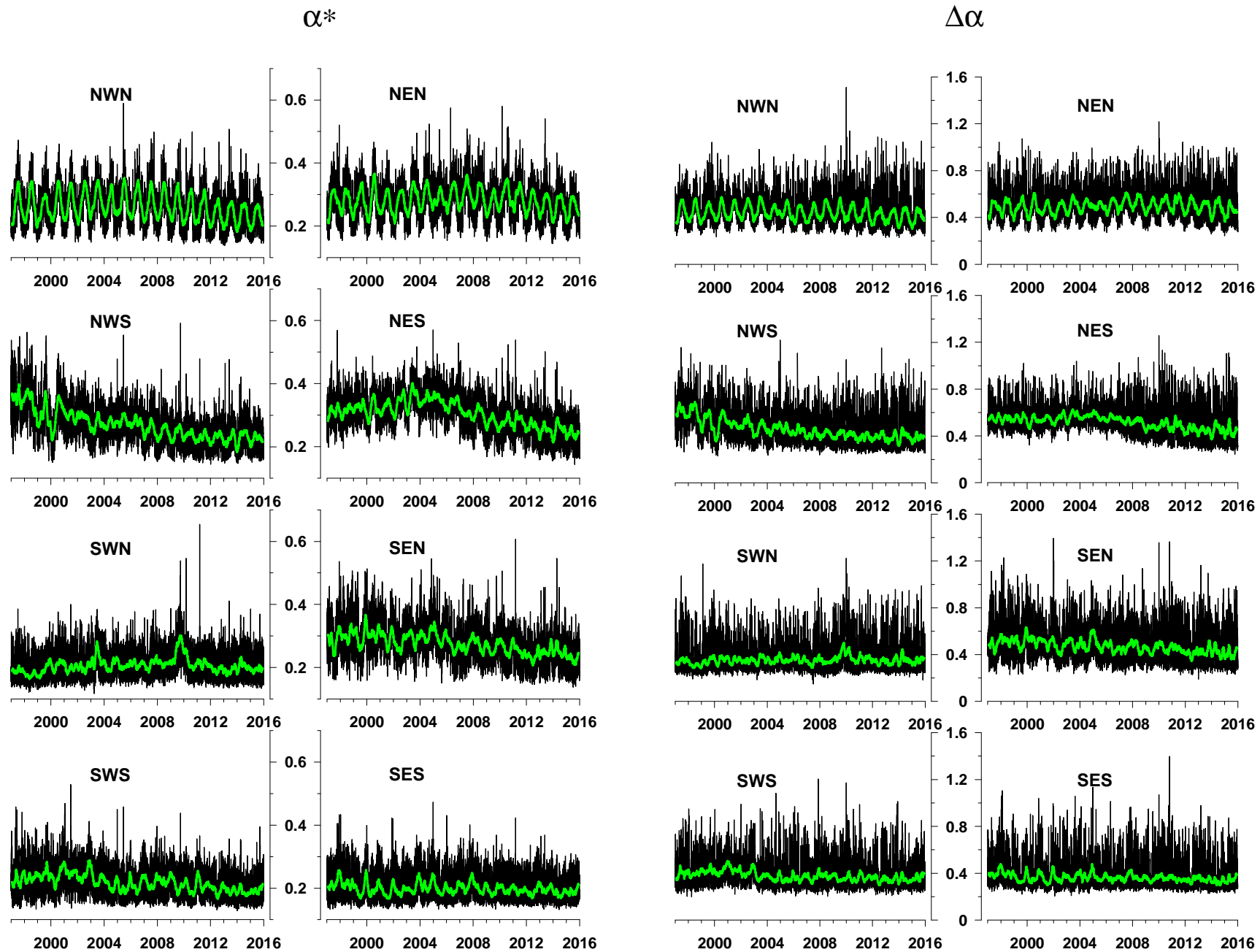


Минимальный показатель Гельдера-Липшица α_{\min}



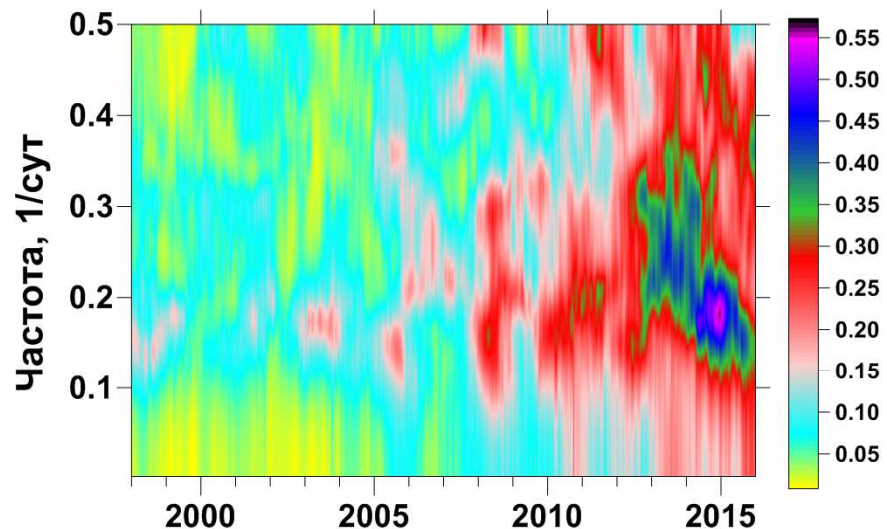
Минимальная нормализованная энтропия квадратов вейвлет-коэффициентов



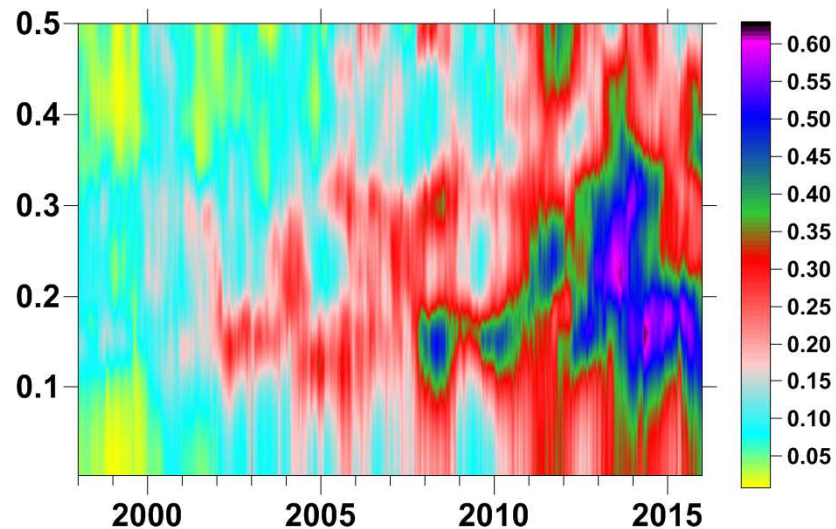


Графики ежесуточных медиан мульти-фрактальных свойств сейсмического фона, вычисленных для 8 групп станций. Толстые зеленые линии – средние значения в скользящем временном окне длиной 57 суток.

Множественная мера когерентности для обобщенного показателя Херста



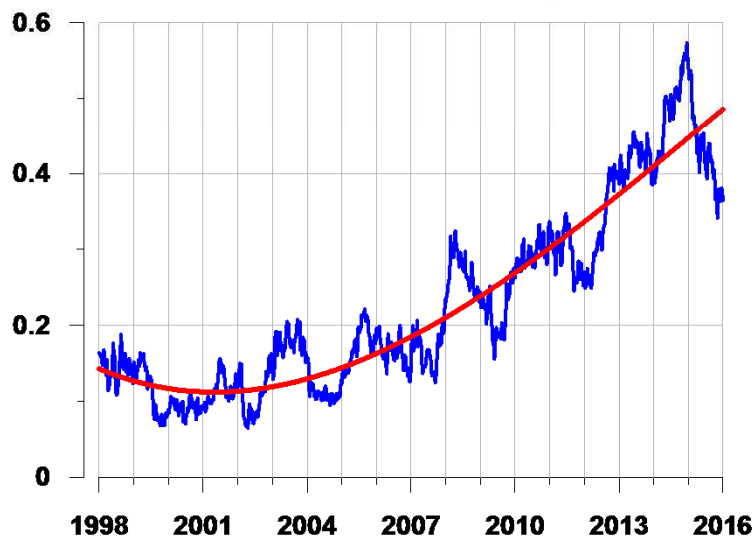
Множественная мера когерентности для ширины носителя спектра сингулярности



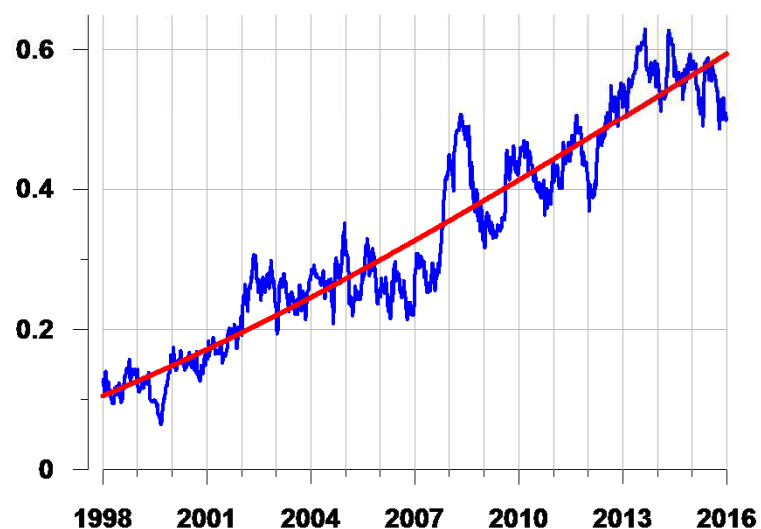
Правый конец скользящего временного окна длиной 365 суток

Красные линии - полиномиальные тренды 3-го порядка

Обобщенный показатель Херста



Ширина носителя спектра сингулярности



Правый конец скользящего временного окна длиной 365 суток со смещением 7 суток

Сильнейшие ($M \geq 8.5$) землетрясения с начала 20-го века

Источник: http://earthquake.usgs.gov/earthquakes/world/10_largest_world.php

Дата	Магнитуда	Широта, град.	Долгота, град.	Дата	Магнитуда	Широта, град.	Долгота, град.
1906.01.31	8.8	1	-81.5	1964.03.28	9.2	61.02	-147.65
1922.11.11	8.5	-28.55	-70.5	1965.02.04	8.7	51.21	178.5
1923.02.03	8.5	54	161	2004.12.26	9.1	3.3	95.78
1938.02.01	8.5	-5.05	131.62	2005.03.28	8.6	2.08	97.01
1950.08.15	8.6	28.5	96.5	2007.09.12	8.5	-4.438	101.367
1952.11.04	9.0	52.76	160.06	2010.02.27	8.8	-35.846	-72.719
1957.03.09	8.6	51.56	-175.39	2011.03.11	9.0	38.322	142.369
1960.05.22	9.5	-38.29	-73.05	2012.04.11	8.6	2.311	93.063
1963.10.13	8.5	44.9	149.6				

Выводы по анализу глобального сейсмического шума.

Можно выдвинуть гипотезу, что нарастание среднего уровня синхронизации параметров глобального поля сейсмических шумов за всю историю сейсмического мониторинга с начала 1997 г. предвещало всплеск сейсмической активности с конца 2004 г. по текущее время.

Следует также отметить, что рост среднего уровня синхронизации наблюдается за весь анализируемый период наблюдений, причем к концу 2014 г. средняя мера синхронизации достигла максимума.

Последнее может быть предвестником дальнейшего роста интенсивности сильнейших сейсмических событий, подобно тому, как это произошло в 1950-1965 гг.