

## Программа SpectCohMes

оценки эволюции спектральной меры когерентности поведения многомерного временного ряда в скользящем временном окне.

А.А. Любушин, доктор физ.-мат. наук  
Институт физики Земли РАН им. О.Ю.Шмидта,  
123995, Москва, Большая Грузинская, 10; факс: +007-499-2556040;  
e-mail: [lyubushin@yandex.ru](mailto:lyubushin@yandex.ru)  
<http://AlexeyLyubushin.narod.ru/Index.htm>

Программа предназначена для выделения скрытых когерентных элементов поведения, которые могут иметь фазовый сдвиг и наблюдаться сразу для нескольких рядов. Метод использует оценку канонических когерентностей в скользящем временном окне, разработанный в [Любушин, 1998, 2007] для поиска предвестников землетрясений по данным низкочастотного геофизического мониторинга. В работах [Любушин и др., 2003, 2004] этот метод был применен к анализу многомерных гидрологических и океанографических (уровнемерных) временных рядов. Метод состоит в оценке частотно-зависимой меры когерентного поведения компонент многомерных временных рядов и его основные конструкции кратко изложены ниже. Если число временных рядов равно 2, то программа фактически строит эволюцию во времени параметрической оценки квадрата модуля спектра когерентности.

### Описание программы.

Обрабатываемые временные ряды должны представлять собой результаты синхронных наблюдений в виде числовых текстовых файлов, имеющих структуру «одна запись – один отсчет» (или «длинная колонка чисел»). Каждому временному ряду должен соответствовать свой файл. Если исходные данные представляют собой таблицы, то временные метки и прочая служебная информация не должны находиться в первой колонке таблицы. Небольшие пропуски данных должны быть восполнены какими-то «правдоподобными» значениями. Перед запуском программы **SpectCohMes** в рабочей директории должен быть создан вспомогательный файл со стандартным именем "list", содержащий имена анализируемых файлов, перечисленных «в столбик» (см. программу **MakeList**).

Программа считывает из файла "list" имена файлов, содержащих анализируемые временные ряды, открывает их, последовательно считывает из них временные отсчеты и закрывает. Если файл "list" отсутствует, то программа останавливается с соответствующим сообщением. Если исходные файлы содержат разное число отсчетов, то обработка будет производиться по выборке длины, равной минимальной длине временных рядов.

Далее пользователю необходимо ответить на следующие вопросы:

- 1) Надо ли переходить внутри каждого скользящего окна к рядам в приращениях («дифференцировать ряды»). Если ряды носят существенно низкочастотный характер, то такой переход необходим. В противном случае, если доминируют низкие частоты, всегда будет наблюдаться высокая когерентность на самых низких для данной длины окна частотах.
- 2) Следует ввести длину скользящего временного окна в числе отсчетов.
- 3) Следует ввести взаимное смещение временных окон в числе отсчетов.
- 4) Следует ввести порядок векторной авторегрессионной модели для оценки спектральной матрицы. Чем больше этот порядок, тем оценка чувствительнее, но,

одновременно, тем она более подвержена статистическим флуктуациям. Значение этого параметра выбирается путем перебора нескольких вариантов и сравнением результатов.

- 5) Следует ввести значения  $T_{ini}, T_{step}$  - временной метки, соответствующей самому первому отсчету и длине шага по времени, с которым производилась регистрация.

Ниже приведена копия экрана такого диалога:

```

C:\APP\SGNL\spectcohmes.exe
of coherent behavior for a multidimensional time series
within moving time window.
Output file - "SpectCohMes.grd".

Alexey Lyubushin, IPE RAS, Moscow
http://lyubushin.hotbox.ru/Index.htm, e-mail: lyubushin@yandex.ru
-----
Number of series to process =          3
Number of samples to be processed <= 1000
From the serie aq1          1000 samples were read.
From the serie aq2          1000 samples were read.
From the serie aq3          1000 samples were read.
Real number of samples to be processed = 1000
Do you want to come to incremental <diff.> series <0/1> ?
1
What is the length of the time window ?
256
What is the length of the intervals` shift= ?
16
What is the AR-order <1 <= *> ?
3
What are the values of initial time mark for series and time step:
Tini, Istep=?
0 1

```

Далее программа начинает работать и создает выводной файл со стандартным именем "SpectCohMes.grd" - это GRID-файл в символьной форме, который затем может быть использован для построения 2-мерных или 3-мерных частотно-временных диаграмм эволюции меры когерентного поведения в пакете Surfer. Структура символьных GRID-файлов следующая (в виде последовательности записей):

DSAA	Стандартная служебная аббревиатура
$N_x \ N_y$	$N_x$ – целое число сеточных узлов по оси X (ось времен) $N_y$ – целое число сеточных узлов по оси Y (ось частот)
$X_{min} \ X_{max}$	$X_{min}$ – минимальное значение X (времени) $X_{max}$ – максимальное значение X (времени)
$Y_{min} \ Y_{max}$	$Y_{min}$ – минимальное значение Y (частоты) $Y_{max}$ – максимальное значение Y (частоты)
$Z_{min} \ Z_{max}$	$Z_{min}$ – минимальное значение Z (меры когерентности) $Z_{max}$ – максимальное значение Z (меры когерентности)
Последовательность значений Z для 1-го (минимального) значения Y (частоты)	
.....	
Последовательность значений Z для максимального значения Y (частоты) с номером $N_y$	

В выводном файле "SpectCohMes.grd" значения X по оси времени соответствуют положению правого конца скользящего временного окна, а по оси частот Y идут в единицах, равных  $1/T_{step}$ . Поскольку выводной файл является обычным текстовым файлом, то его легко редактировать в любом текстовом редакторе типа Notepad, Aditor или WinEdit. Пользуясь

этим, можно изменить те значения  $(\mathbf{X}_{\min}, \mathbf{X}_{\max})$  и (или) значения  $(\mathbf{Y}_{\min}, \mathbf{Y}_{\max})$ , которые генерируются программой автоматически, исходя из введенных значений параметров  $T_{ini}, T_{step}$ . Например, если шаг по времени  $T_{step}$  равен 1 часу, то частоты по оси  $Y$  на диаграмме будут идти в единицах  $\text{час}^{-1}$ , а временные метки по оси  $X$  – в часах, с учетом смещения  $T_{ini}$ . Бывает удобно единицы частот оставить неизменными, а временные метки переделать, например, в сутки или даже годы – в этом случае надо рассчитать новые значения  $(\mathbf{X}_{\min}, \mathbf{X}_{\max})$  и в редакторе их заменить.

### Описание метода.

Квадратичный спектр когерентности двух процессов можно определить, как квадрат коэффициента корреляции этих процессов на частоте  $\omega$ . Канонические когерентности являются обобщением понятия спектра когерентности на ситуацию, когда вместо пары скалярных временных рядов необходимо исследовать связь на различных частотах между двумя векторными временными рядами:  $m$ -мерным рядом  $X(t)$  и  $n$ -мерным рядом  $Y(t)$ . Без ограничения общности будем считать, что  $n \leq m$ . Величина  $\rho_1^2(\omega)$ , называемая квадратом модуля первой (максимальной) канонической когерентности рядов  $X(t)$  и  $Y(t)$  [Brillinger, 1975; Hanan, 1970], которая в данном случае заменяет обычный спектр когерентности, вычисляется как максимальное собственное число эрмитовой матрицы

$$U(\omega) = S_{xx}^{-1/2} S_{xy} S_{yy}^{-1} S_{yx} S_{xx}^{-1/2} \quad (1)$$

Здесь  $\omega$  - частота,  $S_{xx}(\omega)$  - спектральная матрица размером  $m \times m$  временного ряда  $X(t)$ ,  $S_{xy}(\omega)$  - кросс-спектральная прямоугольная матрица размером  $m \times n$ ,  $S_{yx}(\omega) = S_{xy}^H(\omega)$ , “ $H$ ” – знак эрмитова сопряжения. Величина  $\rho_1^2(\omega)$  заменяет квадрат модуля спектра когерентности в случае двух многомерных сигналов.

Введем понятие покомпонентных канонических когерентностей  $\nu_i^2(\omega)$   $q$ -мерного временного ряда  $Z(t)$  как квадратов модулей максимальной канонической когерентности в ситуации, когда в формуле (1) в качестве ряда  $Y(t)$  берется  $i$ -ая скалярная компонента  $q$ -мерного ряда  $Z(t)$ , а в качестве  $X(t)$  –  $(q-1)$ -мерный ряд, состоящий из прочих компонент. Величина  $\nu_i^2(\omega)$ , таким образом, характеризует связанность на частоте  $\omega$  вариаций  $i$ -ой компоненты с вариациями совокупности всех прочих компонент. Введение покомпонентных канонических когерентностей позволяет определить еще одну частотно-зависимую статистику  $\kappa(\omega)$ , характеризующую связанность на частоте  $\omega$  вариаций всех компонент векторного ряда  $Z(t)$ :

$$\kappa(\omega) = \prod_{i=1}^q \nu_i(\omega) \quad (2)$$

Заметим, что, в силу построения, значение величины  $\kappa(\omega)$  принадлежат интервалу  $[0,1]$  и чем ближе соответствующее значение к единице, тем сильнее связь между вариациями компонент многомерного временного ряда  $Z(t)$  на частоте  $\omega$ . Следует подчеркнуть, что сравнение абсолютных значений статистики  $\kappa(\omega)$  возможно лишь для одного и того же числа  $q$  одновременно обрабатываемых временных рядов – поскольку, в силу формулы

(2.3.2), при росте  $q$  величина  $\kappa(\omega)$  убывает, как произведение  $q$  величин, меньших единицы. Если  $q=2$ , то мера (2) становится обычным квадратом модуля спектра когерентности.

Чтобы оценить изменчивость во времени взаимодействия регистрируемых процессов, необходимо производить вычисления в скользящем временном окне заданной длины. Пусть  $\tau$  - временная координата окна длиной  $L$  отсчетов. Вычисляя спектральные матрицы для выборок, попавших во временное окно  $\tau$ , получим двухпараметрическую функцию  $\kappa(\tau, \omega)$ . Всплески величины  $\kappa(\tau, \omega)$  будут определять частотные полосы и временные интервалы увеличения коллективного поведения совместно анализируемых процессов.

Для реализации этого алгоритма необходимо иметь в каждом временном окне оценку спектральной матрицы  $S_{zz}(\tau, \omega)$  размером  $q \times q$ . Ниже предпочтение отдается использованию модели векторной авторегрессии (Marple (Jr.), 1987). Метод заключается в оценке параметров модели:

$$Z(t) + \sum_{k=1}^p A_k \cdot Z(t-k) = e(t) \quad (3)$$

Здесь  $A_k$  - матрицы авторегрессионных параметров размером  $q \times q$ ,  $p$  - порядок авторегрессии,  $e(t)$  -  $q$ -мерный временной ряд остатков идентификации, относительно которого предполагается, что он является последовательностью независимых случайных векторов с нулевым средним и неизвестной ковариационной матрицей  $\Phi = M\{e(t)e^T(t)\}$ , которая считается независимой от времени.

Важно отметить, что модель (3) строится после предварительных операций устранения общего линейного тренда, перехода к приращениям (для увеличения стационарности на коротких временных окнах) и нормировки каждой скалярной компоненты на единичную дисперсию. Эти операции осуществляются независимо в каждом временном окне обработки и для каждой скалярной компоненты многомерного ряда. Их смысл состоит в том, чтобы устранить влияние разномасштабности обрабатываемых рядов. Для оценки матриц  $A_k$  и  $P$  используется рекуррентная процедура Дарбина-Левинсона [Marple, 1987], для которой необходимо предварительно вычислить выборочные оценки ковариационных матриц.

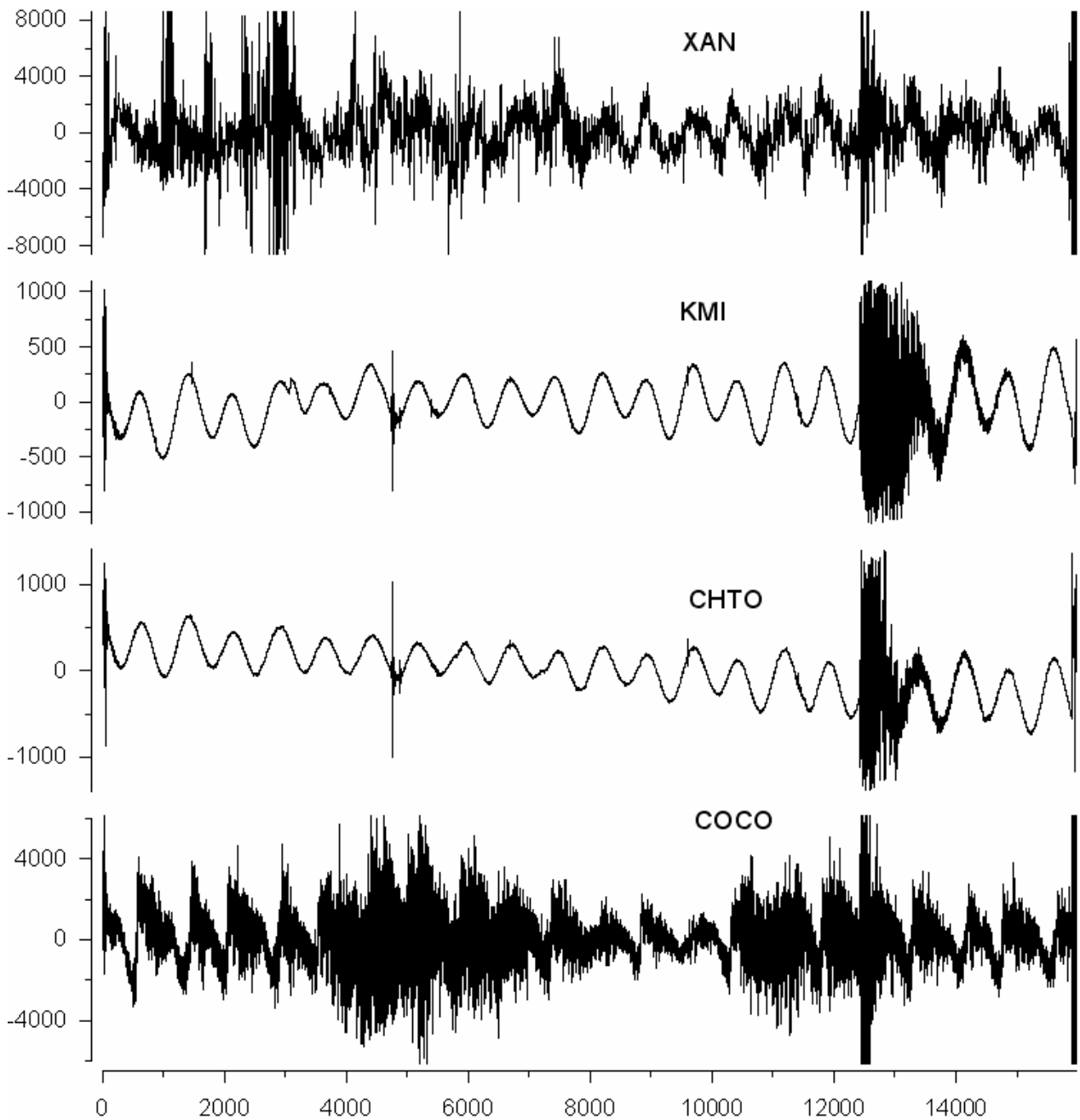
Оценка спектральной матрицы вычисляется по формуле:

$$S_{zz}(\omega) = F^{-1}(\omega) \cdot \Phi \cdot F^{-H}(\omega), \quad F(\omega) = E + \sum_{k=1}^p A_k \cdot \exp(-i\omega k) \quad (4)$$

$E$  - единичная матрица размером  $q \times q$ . Оценка (4) обладает хорошей разрешающей способностью по частоте для коротких выборок. Для выбора порядка авторегрессии  $p$  не существует надежных формализованных процедур. В расчетах выбор  $p$  производился методом проб, как такое минимальное значение, для которого дальнейшее увеличение не приводит к существенному изменению основных элементов поведения зависимости  $\kappa(\tau, \omega)$ .

### Примеры применения.

Ниже на 2-х рисунках представлены исходные временные ряды – низкочастотные широкополосные вертикальные сейсмограммы на 4-х сейсмических станциях сети IRIS после перехода к шагу по времени 30 сек перед Суматранским землетрясением 26 декабря 2004 года и их частотно-временная диаграмма эволюции спектральной меры когерентности во временном окне длиной 12 часов. Пример взят из работы [Соболев, Любушин, 2007]



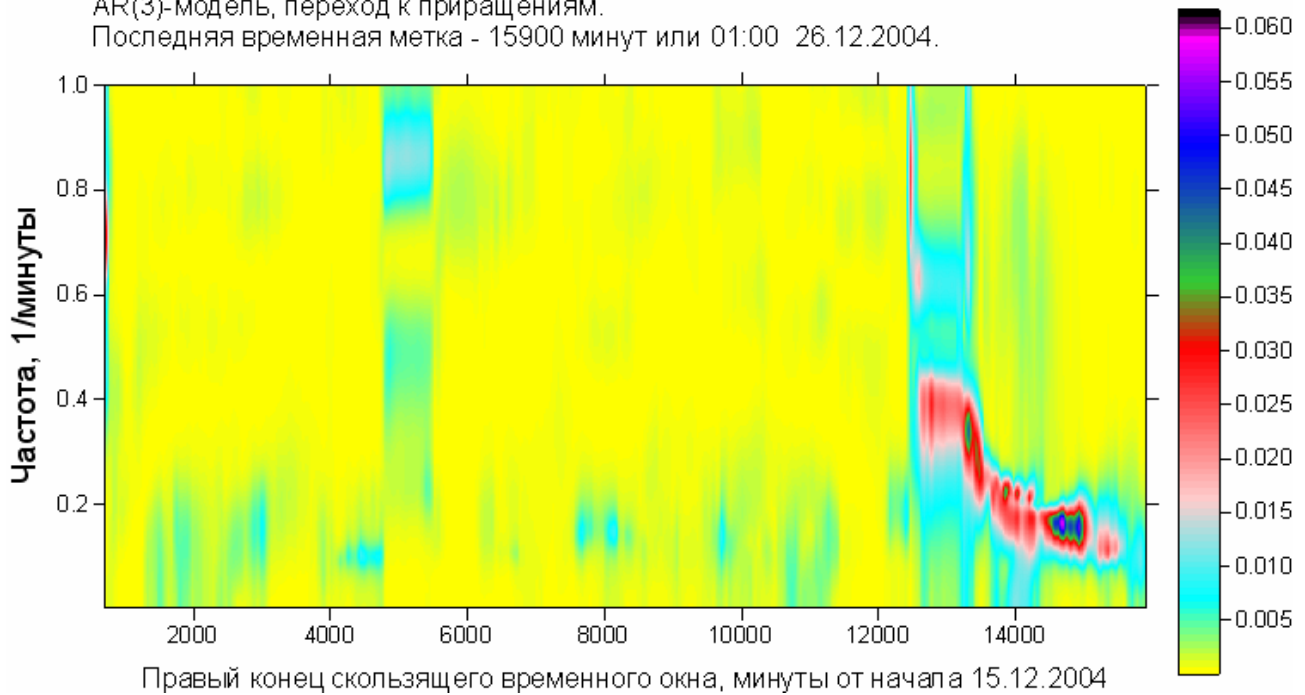
Время, минуты от начала 15.12.2004 (350-й день) по 02:15 26.12.2004 (361-й день),  $\Delta t = 30$  сек

– графики исходных временных рядов.

Эволюция спектральной меры когерентного поведения сейсмических записей после перехода к 30-секундным интервалам. Оценка в скользящем временном окне длиной 1440 отсчетов (12 часов), взятых со взаимным смещением 120 отсчетов (1 час) - строго до вступления от землетрясения на Суматре.

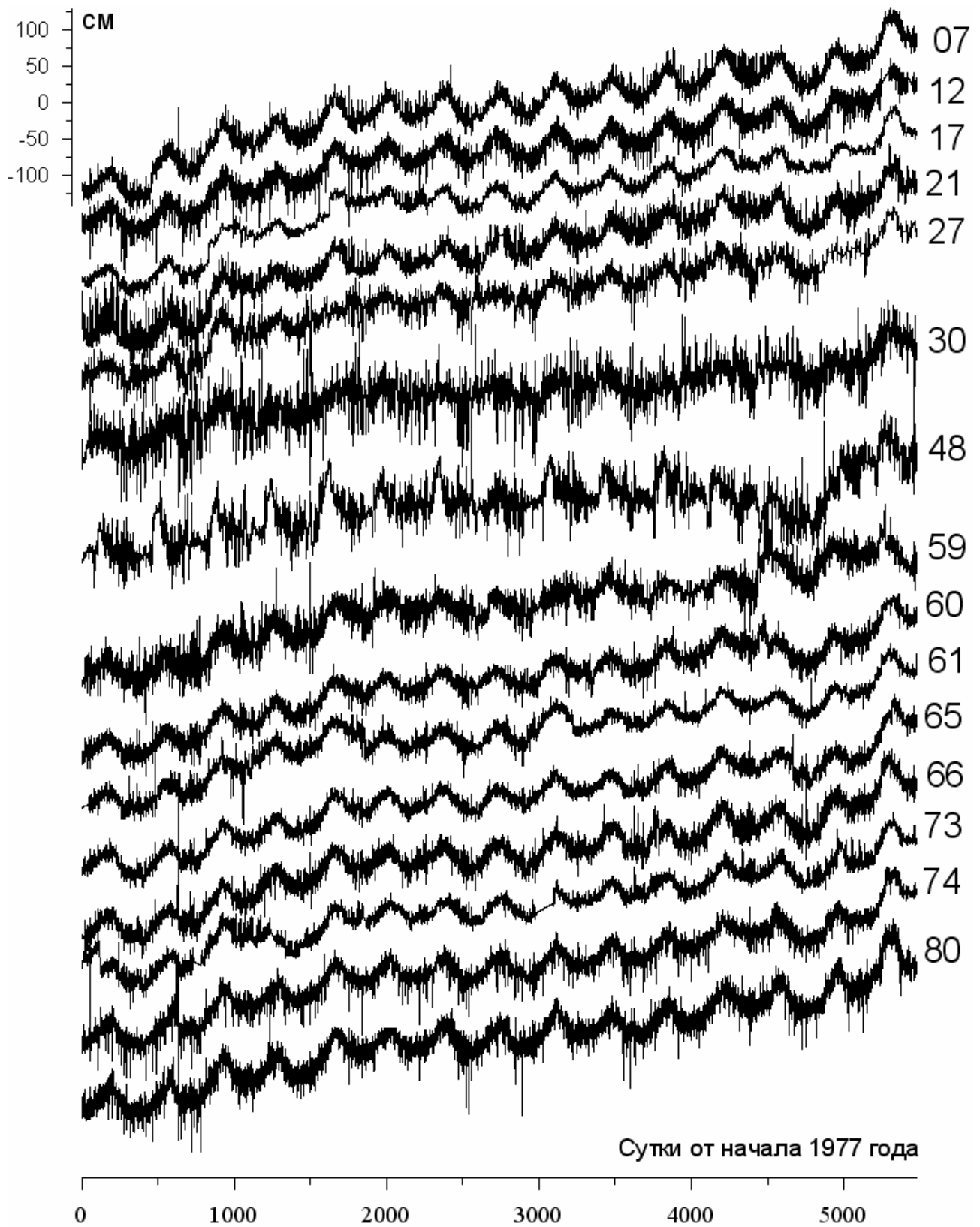
AR(3)-модель, переход к приращениям.

Последняя временная метка - 15900 минут или 01:00 26.12.2004.



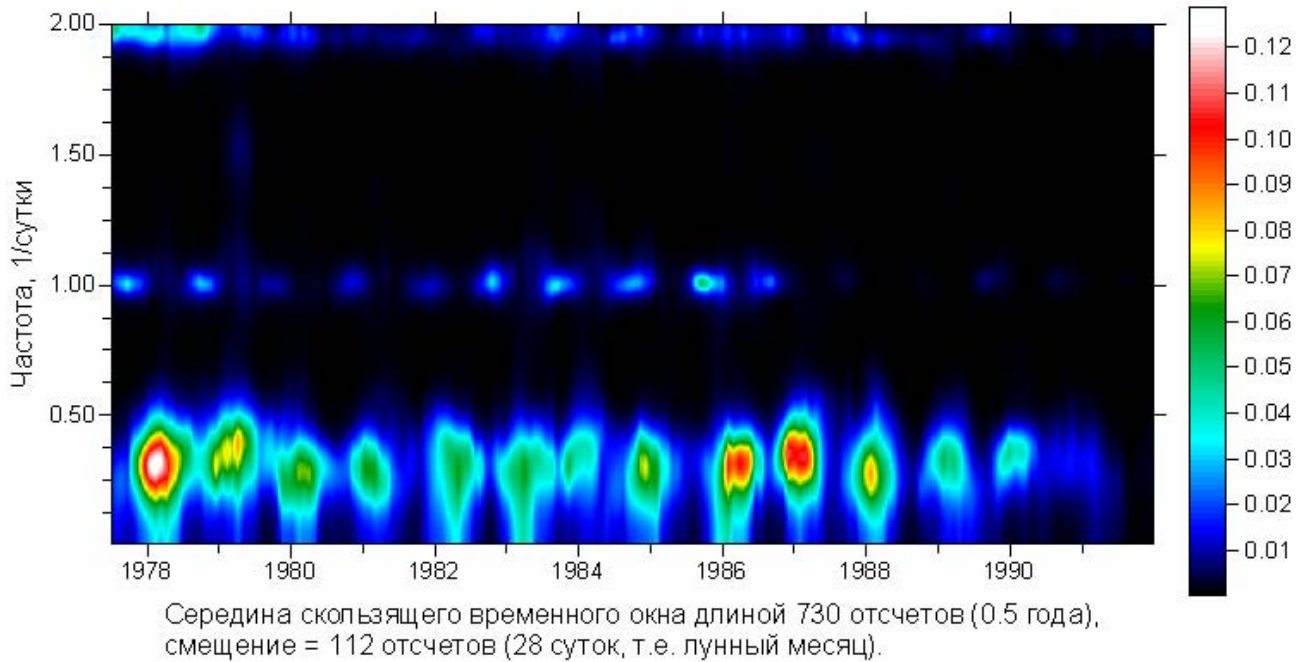
– частотно-временная диаграмма статистики  $\kappa(\tau, \omega)$ , демонстрирующая эволюцию частоты, на которой наблюдается синхронизация, от высоких к низким, по мере приближения к моменту землетрясения на Суматре.

Следующие 2 рисунка демонстрируют применение программы к данным наблюдений за уровнем Каспийского моря на 15 береговых станциях [Любушин и др., 2004].

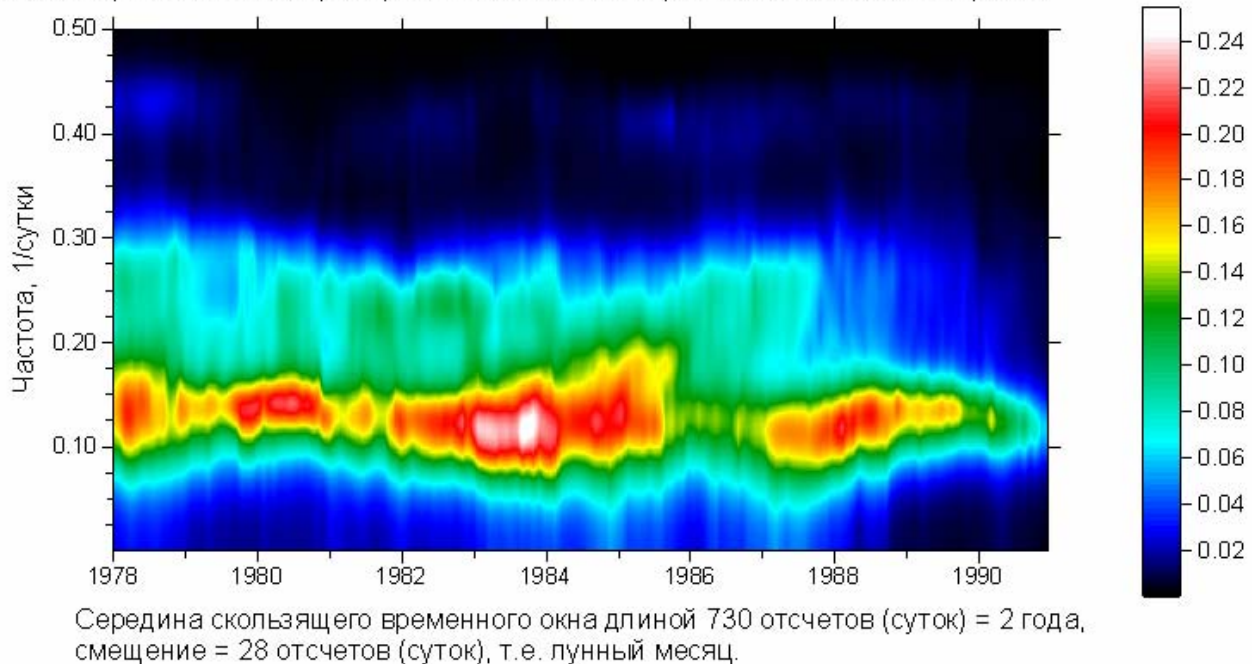


Графики исходных данных. Номера графиков суть идентификационные номера станций наблюдения, расположенных по периметру акватории.

Произведение покомпонентных канонических когерентностей 15-мерного временного ряда синхронных измерений вариаций уровня Каспийского моря на различных береговых станциях. Интервал опроса = 6 часов. Оценка в скользящем временном окне, векторная AR-модель 5-го порядка.



Произведение покомпонентных канонических когерентностей 15-мерного временного ряда синхронных измерений вариаций уровня Каспийского моря на различных береговых станциях. Интервал опроса приведен к 1 суткам, путем фильтрации и прореживания в 4 раза исходных данных. Оценка в скользящем временном окне, векторная AR-модель 5-го порядка.



Основной результат этого анализа заключается в годовой модуляции пиков синхронизации на периоде 4-5 суток, причем по мере приближения к концу 1991 года степень синхронизации уменьшается, что может быть признаком последующей остановки процесса повышения уровня Каспийского моря.



## ЛИТЕРАТУРА.

- Любушин А.А.* (1998) Анализ канонических когерентностей в задачах геофизического мониторинга. – Физика Земли. N1. С.59-66.
- Любушин А.А., В.Ф.Писаренко, М.В.Болгов, Т.А.Рукавишникова* (2003) Исследование общих эффектов вариаций стока рек - Метеорология и гидрология, 2003, N7. С.76-88.
- Любушин А.А., В.Ф.Писаренко, М.В.Болгов, М.В.Родкин, Т.А.Рукавишникова* (2004) Синхронные вариации уровня Каспийского моря по береговым наблюдениям, 1977-1991 гг. – Известия РАН. Физика атмосферы и океана, 2004, N6, том 40, с.821-831.
- Любушин А.А.* (2007) «Анализ данных систем геофизического и экологического мониторинга». М.: Наука, 2007, 228с.
- Соболев Г.А., Любушин А.А.* (2007) Микросейсмические аномалии перед землетрясением 26 декабря 2004 года на Суматре. – Физика Земли, 2007, N5, с.3-16.
- Brillinger D.R.* (1975) Time series. Data analysis and theory. Holt, Rinehart and Winston, Inc., N.Y., Chicago, San Francisco (Русский перевод: *Бриллинджер Д.* (1980) Временные ряды. Обработка данных и теория. М., Мир. 536с.)
- Hannan E.J.* (1970) Multiple time series. John Wiley and Sons, Inc., N.Y., London, Sydney, Toronto (Русский перевод: *Хеннан Э.* (1974) Многомерные временные ряды. М., Мир, 575с.)
- Marple S.L.(Jr.)* (1987) Digital spectral analysis with applications. Prentice-Hall, Inc., Englewood Cliffs, New Jersey (Русский перевод: *Марпл С.Л.* (1990) Цифровой спектральный анализ и его приложения. М., Мир. 584с.).