# Программа WaveSpectExp

## оценки эволюции вейвлетной спектральной экспоненты для группы временных рядов 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
http://AlexeyLyubushin.narod.ru/Index.htm

Программа предназначена для оценки эволюции значения спектральной экспоненты и постоянной Херста в скользящем временном окне для группы синхронных сигналов.

**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 исходное данные представляют собой таблицы, то временные метки и прочая служебная информация не должны находится в первой колонке таблицы. Небольшие пропуски данных должны быть восполнены какими-то «правдоподобными» значениями. Перед запуском программы 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 запросы:

1) Ввести длину скользящего временного окна в числе отсчетов.
2) Ввести взаимное смещение временных окон в числе отсчетов.
3) Ввести значения $T_{ini}, T_{step}$ - временной метки, соответствующей самому первому отсчету и длине шага по времени, с которым производилась регистрация.
4) Ввести порядок полинома (от 1 до 10), который будет использован в каждом окне для устранения тренда и последующего построения вейвлетного спектра мощности для остатка («шума»).

Отметим, что вводить порядок используемого ортогонального вейвлета нет необходимости, поскольку программа определяет его автоматически в каждом окне из условия минимума энтропии распределения квадратов модулей вейвлет-коэффициентов. Как и в программе **WBRCM** используется словарь из 17 вейвлетов: 10 обычных ортогональных вейвлетов Добеши с порядками от 2 до 20 и 7 т.н. «симлетов» - модификаций вейвлетов Добеши, в которых форма базисных функций является более симметричной, чем для обычных вейвлетов [*Mallat, 1998*]. Порядок вейвлета равен числу коэффициентов в дискретных зеркальных фильтрах, реализующих расщепление уровня детальности на 2 частотных диапазона. Число обнуляемых моментов (или степень гладкости) для материнской базисной функции вейвлета равно половине порядка вейвлета.

Копия экрана диалога программы:

```
D:\Users\Lbshn\Wrk\WaveSpectExp.exe
This is the program for esimating spectral exponent values for
a set of time series using orthogonal wavelet decomposition method
within moving time window. Files with time series to be processed must be
at the same directory as the program and their filenames must be listed
in a column in a file with standard filename "list". For each input file
the correspondent output file is generated which has filename
"WaveSpectExp_<Input_Filename>.dat". Output files have 4 columns:
1-st column: time marks of right-hand end of moving time window,
2-nd column: wavelet-based estimate of spectral exponent;
3-rd column: Hurst exponent values, 4-th column: number
of vanishing moments for the best wavelet within current time window.

Alexey Lyubushin, IPE RAS, Moscow
http://lyubushin.hotbox.ru/Index.htm,  e-mail: lyubushin@yandex.ru
-------------------------------------------------------------
Number of signals=          6
What is the length of time window (number of samples, * => 64) ?
2880
What is the value of window's shift= ?
120
What are the values of initial time mark for series and time step:
Tini, Tstep=?
0 0.5
What is the polynomial order for trend removing ( 1 <= * <= 10 ) ?
8
```

Выводные файлы программы записываются в поддиректорию “WaveSpectExp” рабочей директории. Каждому вводному файлу с именем “InpFileName” (имена которых должны быть перечислены в файле “list”) соответствует выводной файл в поддиректории “WaveSpectExp”, имеющий имя “WaveSpectExp_InpFileName.dat”. Выводные файлы представляют собой числовые таблицы, состоящие из 4-х колонок: 1-я колонка представляет собой временные метки правых концов скользящих временных окон, 2-я колонка – значения спектральной экспоненты $\alpha$; 3-я колонка – значения оценок постоянной Херста по формуле $H=(\alpha-1)/2$; 4-я колонка – число обнуляемых моментов (целое число от 1 до 10, которое может быть интерпретировано также как степень гладкости вейвлета) у вейвлета, найденного автоматически для текущего временного окна из условия минимума энтропии распределения квадратов модулей вейвлет-коэффициентов

**Описание метода.**

Эффективным способом анализа информации является переход от исходных временных рядов к каким-нибудь интегральным безразмерным характеристикам данных, вычисленных внутри скользящего временного окна заданной длины. Важную роль в анализе сигналов играют так называемые фрактальные характеристики, такие как показатель Херста, наклон графика спектра мощности в двойном логарифмическом масштабе, параметры мультифрактального спектра сингулярности, корреляционная размерность [*Hurst, 1951; Feder, 19887*]. Открытие эмпирического закона Херста для среднегодового режима расхода воды в реках [*Hurst, 1951; Feder, 1988*] и последующие его приложения к изучению различных случайных процессов в природе, социологии и финансах явилось началом широкого применения фрактальных методов в анализе временных рядов.

Первоначально постоянная Херста $H$ для временного ряда определялась так называемым RS-методом: как коэффициент линейной регрессии между значениями $\ln(RS(s))$ и $\ln(s)$. Здесь $s$ - длина временного интервала, $RS(s)$ - среднее значение отношения размаха (разности между максимальным и минимальным значениями) накопленной суммы отклонений от выборочного среднего к выборочной оценке стандартного отклонения на всех интервалах времени длиной $s$. При вычислении значения $RS(s)$ берется усреднение по всем интервалам этой длины, укладывающимся в имеющуюся выборку временного ряда. То есть, имеет место формула $RS(s) \sim s^H$. Следует подчеркнуть, что эти операции надо выполнять для приращений исследуемого временного ряда. Эмпирический закон Херста заключается в том, что значение $H$ для большого числа метеорологических, гидрологических и геологических наблюдений лежит в окрестности 0.7.

Близость значений постоянной Херста, оцененных для различных процессов, является доводом в пользу того, что их статистическая структура одинакова и близка к свойствам самоподобных случайных процессов. Случайный процесс $Y(t)$ с непрерывным временем называется самоподобным с индексом $H, H > 0$, если для любых $a > 0$ функция распределения (ф.р.) любых конечных выборок случайной величины $Y(a \cdot t)$ совпадает с ф.р. конечных выборок величины $a^H \cdot Y(t)$ [*Taqqu, 1988*]. То есть, если взять произвольное конечное число моментов времени $t_1,...,t_n$, а потом эти значения изменить в $a$ раз, то ф.р. $n$-мерного случайного вектора с компонентами $Y(a \cdot t_1),...,Y(a \cdot t_n)$ будет совпадать с ф.р. вектора с компонентами $a^H \cdot Y(t_1),..., a^H \cdot Y(t_n)$. Растяжение (при $a > 1$) или сжатие (при $a < 1$) временной оси приводит соответственно к увеличению или уменьшению вероятности появления больших значений $Y$. Для самоподобных процессов того же самого можно добиться простым растяжением или сжатием оси ординат в $a^H$ раз. Параметр $H$ называется масштабирующей экспонентой или параметром Херста.

Если самоподобный процесс $Y(t)$ имеет стационарные приращения (для любого шага по времени $h$ ф.р. величин $\Delta_t Y(h) = Y(t+h) - Y(t)$ зависит только от $h$, но не зависит от $t$), то последовательность значений $z(k) = Y((k+1) \cdot \delta t) - Y(k \cdot \delta t)$ с равномерным шагом по времени $\delta t$ образует стационарный временной ряд с нулевым средним. Пусть $0 < H < 1$ и $Y(t)$ - гауссовский процесс. Тогда $Y(t)$ называется фрактальным броуновским движением и обозначается $B_H(t)$. Если $H = 0.5$, то $B_H(t)$ является обычным броуновским движением или винеровским процессом. Без ограничения общности положим шаг по времени $\delta t = 1$. Если $Y = B_H(t)$, то временной ряд $z(k) = Y(k+1) - Y(k), \quad k = 0, \pm 1, \pm 2,...$ называется фрактальным гауссовским шумом.

Для ковариационной функции фрактального шума справедлива формула [*Taqqu, 1988*]:

$$\gamma_{zz}(k) = M\{z(i) \cdot z(i+k)\} = \sigma^2 (|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H})/2 \qquad (1)$$

где $M\{\cdot\}$ - знак математического ожидания (усреднения), $\sigma^2 = Mz^2(k)$ - дисперсия фрактального шума, причем, $M\{|Y(t+h) - Y(t)|^2\} = \sigma^2 \cdot |h|^{2H}$. Заметим, что при $k \neq 0$, если $H = 0.5$, то $\gamma_{zz}(k) = 0$ (обычный белый шум), если $0 < H < 0.5$, то $\gamma_{zz}(k) < 0$ (отрицательная корреляция, антиперсистентность), а если $0.5 < H < 1$, то $\gamma_{zz}(k) > 0$ (положительная корреляция, персистентность). При $H \neq 0.5$ справедлива асимптотическая формула:

$$\gamma_{zz}(k) \sim \sigma^2 H(2H-1) \cdot |k|^{2H-2}, \quad |k| \to \infty \qquad (2)$$

Пусть $S_{zz}(\omega)=\frac{1}{2\pi}\sum_{k=-\infty}^{+\infty}\gamma_{zz}(k)e^{-ik\omega}$ - спектральная плотность стационарной случайной последовательности $z(k)$. Тогда из формулы (2) следует, что $S_{zz}(\omega)\sim\omega^{-(2H-1)}$ при $\omega\to 0$, т.е. если $0.5<H<1$, то $S_{zz}(\omega)\underset{\omega\to 0}{\to}\infty$ и временной ряд $z(k)$ носит низкочастотный характер.

Если теперь по фрактальному шуму $z(k)$ вычислить его кумулятивное значение, то есть рассмотреть временной ряд $X(k)$, удовлетворяющий соотношению $X(k+1)=X(k)+z(k)$, то формально спектр мощности такого ряда будет удовлетворять асимптотике $S_{XX}(\omega)\sim\omega^{-(2H+1)}$, $\omega\to 0$, причем дисперсия любой реализации такого процесса длиной $N$ отсчетов будет стремиться к бесконечности с ростом $N$ по закону $\sim N^{2H}$, то есть понятие спектра мощности для ряда $X(k)$ строго говоря неприменимо. Очевидно также, что многие процессы, встречающиеся в природе и микросейсмический фон, в частности, не являются самоподобными с точки зрения чистой математики. Тем не менее, в практическом фрактальном анализе временных рядов делается предположение, что анализируемые сигналы в определенные интервалы времени могут обладать некоторыми свойствами самоподобных процессов и рассматривают приращения рядов как фрактальный шум. То есть, те характеристики временных рядов, которые получаются из их фрактального анализа, имеют смысл и поддаются интерпретации, в том числе и физической. Например, по конечной выборке ряда оценивается спектр мощности, определяется наклон $\alpha$ графика спектра мощности в двойном логарифмическом масштабе, как коэффициент линейной регрессии между значениями $\ln(S_{XX}(\omega))$ и $-\ln(\omega)$ исходя из формулы $S_{XX}(\omega)\sim\omega^{-\alpha}$. Далее получается оценка постоянной Херста $H=(\alpha-1)/2$. Эта оценка не обязана удовлетворять ограничению $0<H<1$ и может быть даже отрицательной. Но те временные окна, когда она удовлетворяет этому ограничению, могут быть интерпретированы как интервалы самоподобного поведения.

Альтернативным способом вычисления спектральной экспоненты $\alpha$ является использование ортогонального вейвлет-разложения фрагментов сигнала в текущем временном окне. Оценку постоянной Херста можно найти по скорости роста средних значений квадратов модулей вейвлет-коэффициентов [*Mallat, 1997*]:

$$W_k=\sum_{j=1}^{N^{(k)}}|c_j^{(k)}|^2/N^{(k)} \tag{3}$$

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

$$\log_2(W_k)\sim k^{(2H+1)} \tag{4}$$

Таким образом, значение коэффициента наклона прямой, подогнанной методом наименьших квадратов к парам значений $(\log_2(W_k),k)$, дает оценку для величины $2H+1$.

**Пример применения.**

Приведенный ниже пример взят из работы [*Любушин, 2008*], другие примеры использования изменчивости оценок постоянной Херста а задачах геофизики см. в [*Любушин, 2007*].

На рис.1 изображены графики вертикальных широкополосных сейсмограмм на 6 станциях сети IRIS после перехода, путем усреднения и прореживания, к шагу по времени 30 секунд. Все станции, за исключением Obn (Обнинск, Калужская область) находятся в Дальневосточном регионе. По оси времени отложены минуты от начала 15 сентября 2003 года по конец 25 сентября 2003 года, когда недалеко от острова Хоккайдо (Япония) произошло сильнейшее землетрясение с магнитудой 8.3. Вступления от этого сейсмического события заметны на графиках на рис.1 в конце приведенного временного интервала.

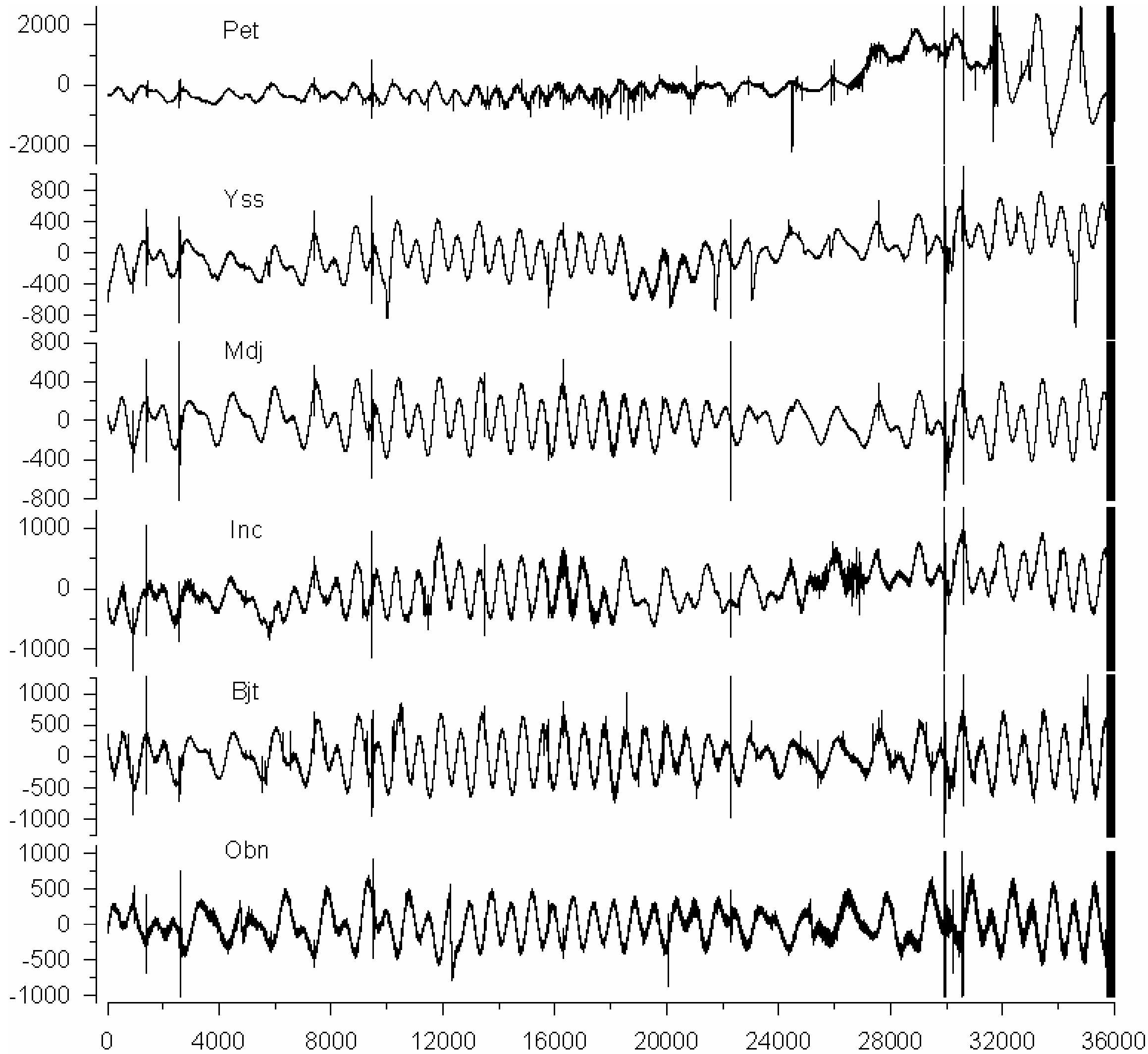


Рис.1.

На Рис.2. (а) представлены графики вариаций оценок постоянной Херста по формуле (4) для указанных станций в скользящем временном окне длиной 1 сутки (2880 30-секундных отсчетов) со смещением 1 час после удаления в каждом окне полиномиального тренда 8-го порядка.

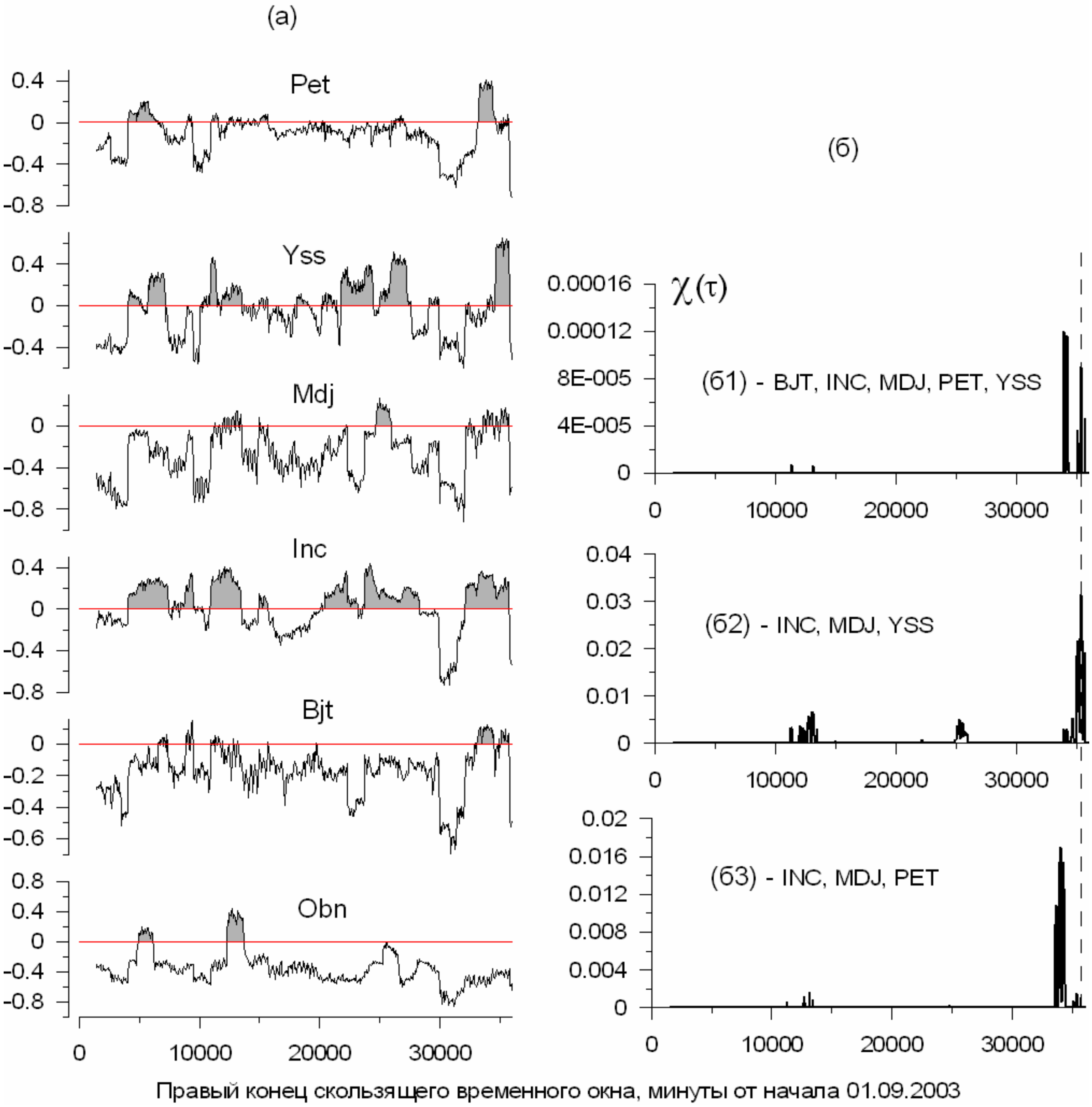


Рис.2.

Обозначим через $H_k(\tau)$ значение оценки параметра Херста для $k$-го временного ряда в зависимости от $\tau$ - временной координаты правого конца скользящего окна. Значения $H_k(\tau)$ для остатка после устранения тренда принимали как отрицательные, так и положительные значения. Выделим те временные окна, для которых эти оценки положительны. Интерес к положительным значениям оценок связан с тем фактом, что для самоподобных процессов значения постоянной Херста должны лежать между 0 и 1. Поэтому неравенство $H_k(\tau) > 0$ косвенно является признаком фрактального самоподобного поведения низкочастотного сейсмического шума. В силу этого представляет интерес выделение таких временных окон, когда сразу для всех совместно анализируемых процессов оценки постоянной Херста стали положительными – это является признаком некоторой низкочастотной синхронизации. Выделить такие окна можно с использованием меры:

$$\chi(\tau) = \prod_k \max(0, H_k(\tau)) \qquad (5)$$

Очевидно, что величина (5) равна нулю, если хотя бы для одного сигнала оценка $H_k(\tau)$ неположительна. На Рис.2(б1-б3) представлены графики изменения величины (5) для различных указанных комбинаций станций, вертикальная штриховая линия выделяет момент землетрясения на Хоккайдо. Получился эффект синхронного превышения оценками $H_k(\tau)$ нулевого уровня незадолго да землетрясения на Хоккайдо.

ЛИТЕРАТУРА.

*Любушин А.А.* (2007) «Анализ данных систем геофизического и экологического мониторинга». М.: Наука, 2007, 228с.

*Любушин А.А.* (2008) Микросейсмический шум в минутном диапазоне периодов: свойства и возможные прогностические признаки. – Физика Земли, 2008, № 4, с.17-34.

*Feder J.* (1988) Fractals. Plenum Press, New York, London (Русский перевод: *Федер Е.* (1991) Фракталы. М., Мир. 254с.)

*Hurst H.E.* (1951) Long-term storage capacity of reservoirs. – Trans. Amer. Soc. Civ. Eng. Vol.116, pp.770-808

*Mallat S.* (1998) A wavelet tour of signal processing. Academic Press. San Diego, London, Boston, N.Y., Sydney, Tokyo, Toronto. 577 p. (Русский перевод: Малла С. Вэйвлеты в обработке сигналов. - М.: Мир, 2005, 671с.).

*Taqqu M.S.* (1988) Self-similar processes. Encyclopedia of Statistical Sciences, vol.8, pp.352-357, Wiley, New York, 1988.