

АНАЛИЗ ВРЕМЕННОГО РЯДА ЭЛЕКТРОПОТРЕБЛЕНИЯ МЕТОДОМ НОРМИРОВАННОГО РАЗМАХА

Филатова Е.С., Филатов Д.М., Стоцкая А.Д.

ФГАОУ ВО Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И.Ульянова (Ленина), es_filatova@mail.ru

Прогнозы потребления составляют основную информацию для процесса принятия решений о планировании оптимальных режимов работы и развития электроэнергетических систем. В статье рассматривается применение метода нормированного размаха к анализу временного ряда, описывающего получасовые, часовые и среднесуточные колебания электрической нагрузки. Исследуются результаты применения данного метода для зашумленного сигнала и для сигнала прошедшего предварительную фильтрацию и нормирование. Сформулированы рекомендации по применению метода нормированного размаха для анализа временных рядов электропотребления. Доказывается, что фильтрация шума в рассматриваемом временном ряду не оказывает существенного влияния на показатель Хёрста, а увеличение уровней детализации сигнала является безрезультатным и лишь увеличивает время расчета.

Ключевые слова: анализ временных рядов, электропотребление, показатель Хёрста, метод нормированного размаха.

ANALYSIS OF POWER CONSUMPTION TIME SERIES BY NORMALIZED RANGE METHOD

Filatova E.S., Filatov D.M., Stotckaia A.D.

Saint-Petersburg State Electrotechnical University (ETU LETI), es_filatova@mail.ru

Consumption forecasts are the basic information decision making process about the planning of optimal operating conditions and the development of electric power systems. The paper deals with the application of the normalized range method to the analysis of time series describing the half-hour, hour and daily fluctuations of electrical load. The results of applying this method to the noisy signal and the pre-filtering and normalizing signal are discussed. Recommendations on the application of the normalized range method for analysis of power consumption time series are given in detail. It is proved that the filtering of noise in the present time series does not have a material impact on the Hurst index, and the increase of signal detail levels is inconclusive and only leads to computation time increases.

Keywords: time series analysis, power consumption, Hurst index, the method of the normalized amplitude.

Анализ и прогнозирование временных рядов является непрерывно развивающимся направлением теоретических и практических исследований. Тенденцией последних лет стало применение к анализу и прогнозированию временных рядов методов из нелинейной динамики, таких, в частности, как фрактальный анализ.

Применение методов анализа, основанных на фракталах, позволяет находить закономерности во временных рядах, на первый взгляд являющихся совершенно случайными [1]. Наиболее распространённым является фрактальный анализ на основании коэффициента Хёрста, получивший название RS анализ или метод нормированного размаха. В начале прошлого века гидролог Х.Е. Хёрст (Harold Edwin Hurst) при анализе данных о разливах Нила обнаружил что $(R/S)_n = c * n^H$, где R_n – скоррелированный размах, S_n – стандартное отклонение на выборке длиной n , H – коэффициент Хёрста [2].

Нормированного размаха

Коэффициент Хёрста принимает значения от 0 до 1.

Значение $0,50 < H \leq 1,00$ подразумевает персистентный временной ряд. Персистентная система проходит большее расстояние, чем случайная система. Персистентный временной ряд характеризуется эффектами долговременной памяти. Теоретически, то, что происходит сегодня, воздействует на будущее. В терминах хаотической динамики существует чувствительная зависимость от начальных условий. Такая долговременная память имеет место независимо от масштаба времени [3].

Значение $0 \leq H < 0,50$ означает антиперсистентность. Антиперсистентная система проходит меньшее расстояние, чем случайная система. Чтобы система прошла меньшее расстояние, она должна меняться чаще, чем вероятностный процесс.

Анализ Хёрста позволяет достаточно просто выявить периодические циклы и даже не периодические циклы. Для этого необходимо определить, при какой величине n в графике H -траектории, т.е. зависимости $\log R/S = f(\log n)$ возникают разрывы. Или при помощи V -статистики меняет направление тренда [3].

Методика анализа довольно проста и подробно описана в [3]:

1. После подготовки данных временной ряд длины M разделяется на A смежных подпериодов длины n , так чтобы $A \cdot n = M$, для каждого подпериода I_a , $a = 1, 2, 3, \dots, A$ определяется среднее значение e_a .

$$e_a = \frac{1}{n} \sum_{k=1}^n N_{k,a},$$

где $N_{k,a}$ – k -й элемент подпериода I_a .

2. Формируется временной ряд накопленных отклонений от среднего значения $X_{k,n}$ для каждого подпериода I_a

$$X_{k,a} = \sum_{i=1}^k (N_{i,a} - e_a) \quad k=1, 2, 3, \dots, n$$

3. Диапазон R_k определяется как разность максимального и минимального накопленного отклонения

$$R_a = \max(X_{k,a}) - \min(X_{k,a}),$$

где $1 \leq k \leq n$

4. Для каждого подпериода определяется стандартное отклонение S_a

5. Каждый диапазон R_a нормализуется путём деления на S_a и определяется среднее значение R/S для длины n .

$$(R/S)_n = 1/A * \sum_{a=1}^A (R_a/S_a)$$

6. Процесс повторяется при увеличении n до величины $M/2$.

Результаты исследований

Следует отметить, что при малых n (единицы) велика погрешность метода, так что требуется довольно большая величина выборки. Так же не обязательно выполнение условия $n \cdot A = M$, в этом случае - если M не кратно n последние члены ряда отбрасываются, при малых n влияние отброшенных элементов ряда будет невелико, однако при росте n погрешность будет расти. К тому же при таком алгоритме время обработки многократно возрастает.

Таким образом, для точного анализа важен не только объём выборки, но и чтобы объём этой выборки был кратен как можно большему количеству чисел.

В качестве исходных данных возьмём данные об энергопотреблении, данные фиксировались на протяжении одного года каждые тридцать минут.

Рассмотрим временной ряд среднесуточного энергопотребления за год и проанализируем его при помощи RS-анализа.

На рис. 1 представлена V-статистика и траектория показателя Хёрста среднесуточного энергопотребления, в графиках мы наблюдаем разрывы соответствующие семидневным циклам энергопотребления, что ожидаемо т.к. соответствует недельному циклу.

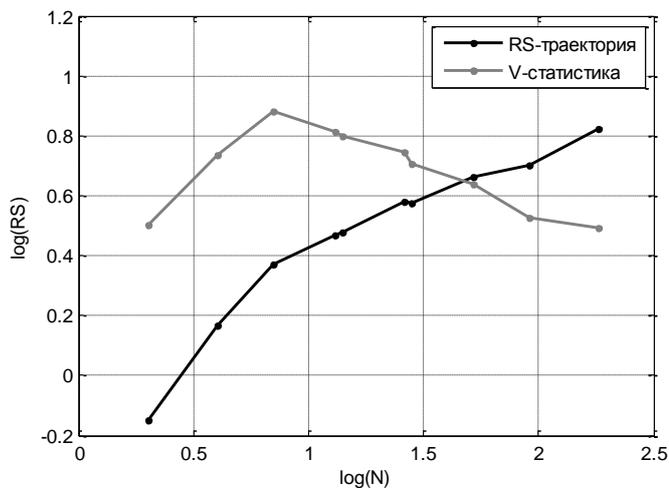


Рис. 1. - V-статистика и траектория показателя Хёрста среднесуточного энергопотребления

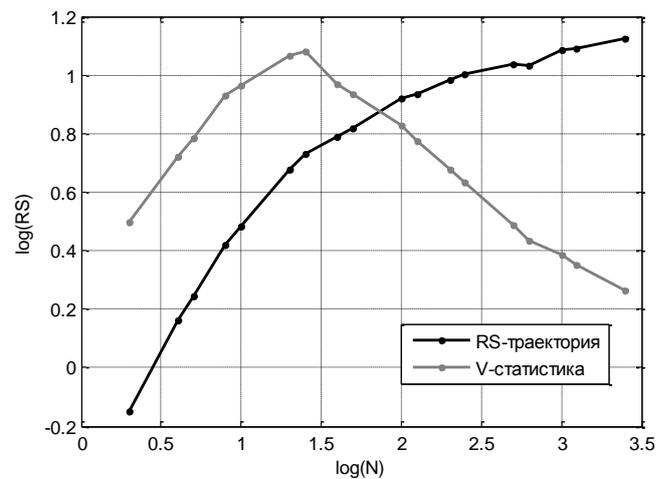


Рис. 2. – график V статистики энергопотребления за год

Коэффициент Хёрста равен $H=0.95$, временной ряд сильно персистентный, а значит должен достаточно легко поддаваться прогнозированию на основании данных о его предыдущих значениях.

Рассмотрим временной ряд энергопотребления за год, данные взяты с часовым интервалом.

График V статистики (см. рис. 2) показывает, что в графике показателя Хёрста присутствуют разрыв при $N=24$, который соответствует суточному циклу энергопотребления. Показатель Хёрста в данном случае равен $H=0.91$.

На рис. 3 представлена траектория показателя Хёрста, для временного ряда получасовых интервалов электропотребления. На графике V-статистики виден разрыв, соответствующий суточному циклу электропотребления. Также на обоих графиках RS-траектории присутствуют разрывы, соответствующие недельному циклу.

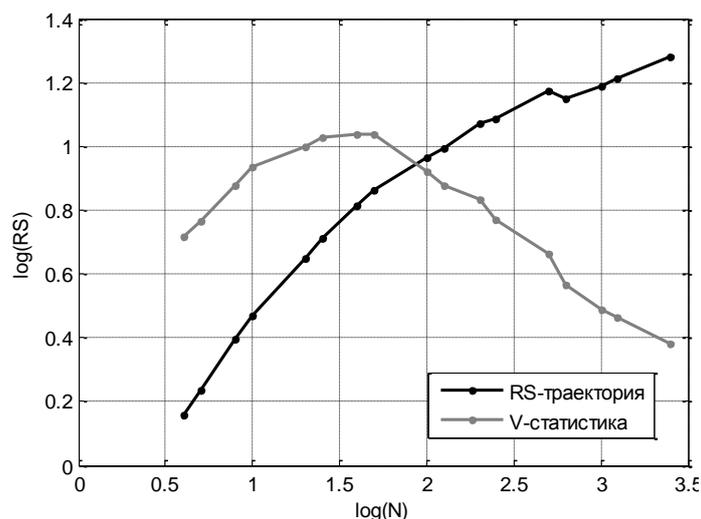


Рис. 3. - траектория показателя Хёрста для временного ряда получасовых интервалов электропотребления

Из графика H-траектории представленного на рисунке 3 видно, что показатель Хёрста $H=0.8197$, а значит временной ряд получасового потребления менее персистентный, чем ряд среднесуточного потребления. Таким образом, методика RS-анализа, предложенная Хёрстом, позволяет находить циклы энергопотребления, и полученные данные совпадают с ожидаемыми циклами рабочего дня, суток и недели.

Однако, RS анализ также может быть использован для оценки сложности (изрезанности) временного ряда. На основании такой оценки возможно принятия ряда мер, таких как фильтрация шума. И наоборот, применение некоторых операций, так как нормирование, может сделать анализируемый временной ряд менее персистентным, а значит, менее прогнозируемым. Покажем это на примере. Но до этого немного адаптируем алгоритм применительно к решаемой задаче.

Использование в расчёте только подвыборки кратных мощности множества приводит к снижению точности анализа и возможности потери одного из циклов. Модифицируем алгоритм анализа таким образом, чтобы рассматривались все подвыборки, объёмом не больше половины мощности множества. Если величина подвыборки не кратна величине выборки, лишние члены выборки не будут учитываться. Так как наиболее важны последние

члены ряда – они оказывают влияние на будущие значения, отбрасываться будут первые члены временного ряда среднесуточного энергопотребления. При анализе временного ряда электропотребления широко применяют различные типы нормирования. Предыдущие результаты анализа были показаны для временных рядов электропотребления, нормированных по методу логарифмических разностей.

Рассмотрим, как нормирование повлияет на результаты анализа временного ряда методом нормированного размаха. На рис. 4 приведены графики RS-траектории для нормированной и ненормированной выборки среднесуточного электропотребления при анализе по всем подвыборкам.

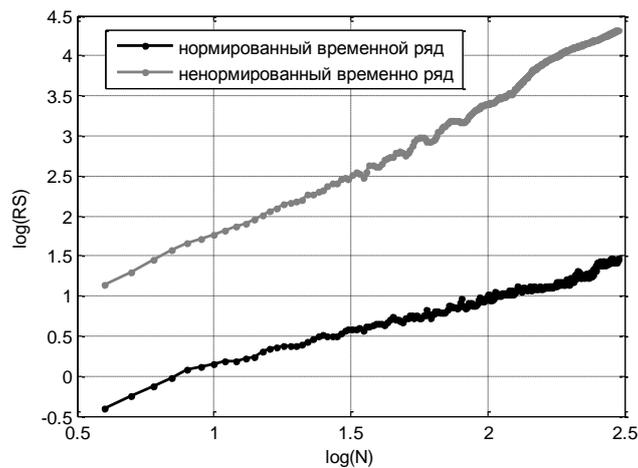


Рис. 4. - графики RS-траектории для нормированной и ненормированной выборки среднесуточного электропотребления

Показатель Хёрста не изменился, однако анализ V-статистики показал, что лучше для анализа RS-траектории после нормирования временного ряда.

Рассмотрим нормированный ряд получасового энергопотребления. На рис. 5 приведены графики RS-траектории при использовании в анализе всех выборок для нормированного и ненормированного временных рядов. На рис. 6 приведены графики V-статистики для аналогичного случая.

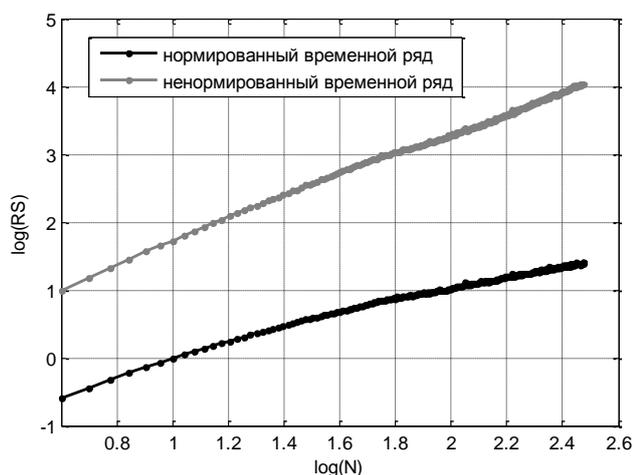


Рис. 5. - графики RS-траектории при использовании в анализе всех выборок для нормированного и ненормированного временных рядов

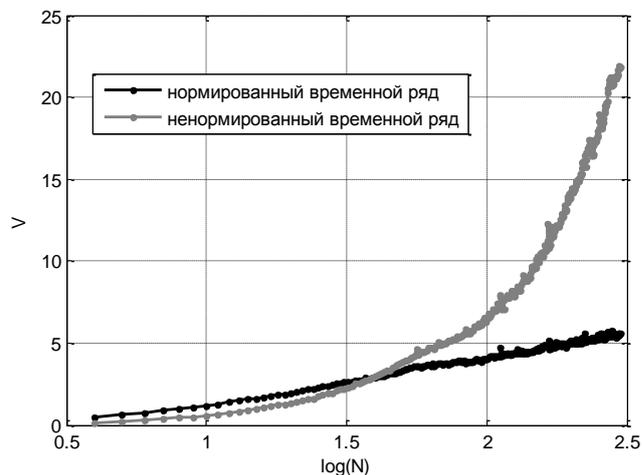


Рис. 6. графики V-статистики при использовании в анализе всех выборок для нормированного и ненормированного временных рядов

В области больших выборок (больше 300) возникают колебания графиков как V статистики, так и RS-траектории, что делает невозможным анализ наличия циклов длиной больше 300. Поэтому, можно сократить время обработки данных, ограничив максимальный объём подвыборки.

На представленных графиках присутствуют разрывы соответствующие суточному и недельному циклам, что совпадает с ожидаемыми циклами ряда энергопотребления.

Таким образом, для временных рядов среднесуточного и получасового ряда энергопотребления можно применять RS-анализ по всем подвыборкам, что позволяем добиться более точного определения длительности цикла, а также позволяет использовать достаточно большую выборку любой длины, не подбирая специально число элементов (даже если длина выборки является простым числом, анализ будет проведен). Предварительное нормирование выборки не является обязательным, но делает график H – траектории более удобным для визуального анализа. V-статистика может быть использована только для случая нормированной выборки.

Еще один традиционный способ предварительной подготовки временного ряда к дальнейшему анализу – фильтрация шума. Применим для фильтрации вейвлет-анализ [4] и проанализируем полученный временной ряд методом нормированного размаха.

Для разложения по вейвлету используется избыточное дискретное вейвлет-преобразование (ИДВП), для реализации которого используется алгоритм «Избыточный Хаар» (алгоритм «a trous») [1]. В результате исходный временной ряд можно представить как:

$$x(t) = c_p(t) + \sum_{i=1}^p w_i(t), \quad (4)$$

где $c_p(t)$ – аппроксимация на уровне p , $w_i(t)$ – коэффициенты детализации на уровнях $i = 1 \div p$ соответственно. Фильтрация осуществляется путем отбрасывания последних коэффициентов детализации при восстановлении сигнала.

Величина p называется глубиной вейвлет-разложения (далее глубиной разложения) и обозначает количество уровней декомпозиции, которые используются для приближения исходного сигнала. Этот показатель определяет количество получаемых в результате разложения временных рядов и влияет на вид этих временных последовательностей. В таблице 1 приведены значения коэффициента Хёрста для различных уровней декомпозиции временного ряда получасового электропотребления при различной глубине разложения.

Таблица 1

Значения коэффициента Хёрста для различных уровней декомпозиции временного ряда

	Исходный сигнал	Аппроксимация	Детализация 1	Детализация 2	Детализация 3	Детализация 4	Детализация 5	Детализация 6
H	0,8197	0,812	0,15	0,18	0,2	0,2	0,3	0,3
		0,815	0,18	0,15	0,2	0,22	0,3	
		0,810	0,18	0,15	0,2	0,22		
		0,8104	0,15	0,18	0,2			
		0,813	0,15	0,18				

Как видно из таблицы, фильтрация шума в рассматриваемом временном ряду не оказывает существенного влияния на показатель Хёрста, и даже немного понижает его. Отсюда можно сделать вывод, что данные либо не содержат шума, либо он пренебрежимо мал. Тем не менее, проведенные исследования показали, что для зашумленных сигналов фильтрация позволяет повысить коэффициент Хёрста, а значит предсказуемость временного ряда.

Что касается глубины вейвлет-разложения, то можно сказать, что увеличение уровней детализации сигнала является безрезультатным и лишь увеличивает время расчета. В [5] даны рекомендации по величине p для временных рядов в различными показателями Хёрста.

Заключение

Применение метода нормированного размаха для анализа временного ряда электропотребления выявило ожидаемые суточные и недельные циклы колебания электрической нагрузки. Предварительное нормирование выборки не является обязательным, но делает график H – траектории более удобным для визуального анализа. V -статистика может быть использована только для случая нормированной выборки. Фильтрация шума не дает увеличения показателя Хёрста в случае, если амплитуда данного шума составляет менее 10% от амплитуды сигнала.

Список литературы

1. Солнцев Л.А., Иудин Д.И., Снегирева М.С. Фрактальный анализ векового хода средней температуры воздуха в г. Нижнем Новгороде [Текст] / Л.А. Солнцев, Д.И. Иудин, М.С. Снегирева // Вестник Нижегородского университета им. Н.И. Лобачевского. Нижний Новгород. – 2007.- № 4.- С. 88–91.
2. Hurst H.E., Black R.P., Simaika Y.M. Long-term storage: An experimental study. - L.: Constable, 1965
3. Пэтерс, Э.Э. Фрактальный анализ финансовых рынков [Текст] / Э.Э. Петерс. - М.: Интернет трейдинг. - 2004.- С.304.
4. Фильтрации сигналов и изображений: фурье и вейвлет алгоритмы (с примерами в Mathcad) : монография / Ю. Е. Воскобойников, А.В. Гочаков, А. Б. Колкер ; Новосиб. гос. архитектур.-строит. ун-т (Сибстрин). – Новосибирск : НГАСУ (Сибстрин), 2010. – 188 с.
5. Филатова, Е.С. Определение оптимальной глубины вейвлет-разложения временного ряда электропотребления с использованием метода нормированного размаха [Текст] / Е.С. Филатова, Д.М. Филатов // Тез. докл. междунар. конф. по мягким вычислениям и измерениям SCM-2014. Saint-Petersburg, 25-27 June. – 2014. – Vol. 1. – P. 306–309.

Рецензенты:

Поляхов Н.Д., д.т.н., профессор, профессор ФГАОУ ВО Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), г.Санкт-Петербург;

Путов В.В., д.т.н., профессор, профессор ФГАОУ ВО Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина), г. Санкт-Петербург.