

В. А. Андриенко,
профессор, д.ф.-м.н., профессор кафедры математического анализа,
Одесский национальный университет им. И.И.Мечникова
В. М. Андриенко,
доцент, к.е.н., доцент кафедры экономической кибернетики и информационных технологий,
Одесский национальный политехнический университет
А. Ш. Тулякова,
аспирантка, Одесский национальный университет им. И.И.Мечникова

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ФОНДОВЫХ РЫНКОВ

В данной публикации кратко изложены теоретические основы интеллектуального исследования временных рядов и приведена практическая реализация теоретических положений на примере анализа украинского фондового индекса ПФТС.

In this publication theoretical basis of intellectual research of temporal rows is briefly expounded and practical realization of theoretical positions is resulted on the example of analysis of the Ukrainian fund index of PFTS.

Ключевые слова: случайный процесс, временной ряд, долговременная память, фрактальная размерность, хаотическая составляющая, нелинейная динамическая система.

Keywords: casual process, temporal row, of long duration memory, fractal dimension, chaotic constituent.

Введение.

Статистические данные экономических показателей, в том числе и фондового рынка, поступают в форме *временных рядов*. *Временным рядом* называется совокупность наблюдений экономического показателя в разные моменты времени. Анализ взаимосвязей экономических данных, представленных в виде временных рядов, является необходимой составной частью современных исследований в области макроэкономической динамики переходной экономики и эконометрики финансовых рынков. Учет временной структуры для данных о реальных экономических процессах позволяет адекватно отразить их в математических и эконометрических моделях.

1. Постановка задачи.

Принципиальные отличия временного ряда от последовательности наблюдений, образующих случайную выборку, заключаются в следующем: члены временного ряда не являются независимыми и не обязательно являются одинаково распределенными. Это означает, что свойства и правила статистического анализа случайной выборки нельзя распространять на временные ряды. С другой стороны, взаимозависимость членов временного ряда создает свою специфическую базу для построения прогнозных значений анализируемого показателя по эмпирическим значениям.

Обычно временной ряд рассматривают как выборку из последовательности случайных величин X_t , где t принимает целочисленные значения от 1 до T . Совокупность случайных величин $\{X_t, t \in [1, T]\}$ называют *дискретным случайным процессом* или *стохастическим процессом*. Стохастический процесс называется *стационарным в широком смысле* если его свойства не изменяются во времени. В частности, он имеет постоянное математическое ожидание $M(X_t)$ и постоянную дисперсию $D(X_t)$, а ковариация X_{t_1} и X_{t_2} зависит от разности $(t_1 - t_2)$ [1].

В 1938 году Вольд доказал следующий фундаментальный результат [2]. Чисто недетерминированный стационарный в широком смысле случайный процесс X_t может быть представлен в следующем виде: $X_t - M(X_t) = \sum_{\tau=0}^{\infty} \psi_{\tau} \varepsilon_{t-\tau}$, а ε_t - белый шум с конечными математическим ожиданием и дисперсией. То есть всякий стационарный в широком смысле случайный процесс

представляется в виде линейной комбинации белых шумов. При этом должно выполняться условие $\sum_{\tau=0}^{\infty} |\psi_{\tau}| < \infty$. Поскольку реализации белого шума не наблюдаемы, весовые коэффициенты определены с точностью до множителя. Без потери общности можно считать, что $\psi_0 = 1$. Чем больше весовой коэффициент ψ_{τ} , тем больше влияние случайного возмущения в момент $t - \tau$ на текущий момент t . Оказалось, что во многих случаях достаточно рассматривать не общее представление Вольда, а его частные случаи, когда число слагаемых конечно. Такими частными случаями являются *белый шум*, *гауссовский белый шум*, известные в эконометрике процессы $AR(p)$, $MA(q)$, $ARMA(p, q)$ и др. [1, 6].

Экономические показатели, в частности показатели фондовых рынков, не всегда ведут себя стационарным образом. Из макроэкономики известно *сезонное* и *циклическое* поведение экономических показателей, кроме того, они могут иметь *тренд*. Тренд представляет собой общую систематическую линейную или нелинейную компоненту, которая может изменяться монотонно во времени. Аналитически данное положение можно выразить уравнением вида:

$$X(t) = f(t) + S(t) + e(t).$$

где $f(t)$ - тренд (долговременная тенденция) развития;

$S(t)$ - сезонная компонента;
 $e(t)$ - случайная величина (случайная компонента).

Функция $f(t)$ определяет общую тенденцию развития изучаемого явления. Тренд может быть выражен как детерминированной, так и стохастической функциями, либо их комбинацией. Компоненты временного ряда $f(t)$, $S(t)$, $e(t)$ не наблюдаемы. Они являются теоретическими величинами. Выявление этих компонент и является задачей интеллектуального анализа. Детерминированную и периодическую компоненты можно определить с помощью спектрального анализа. Статистическую зависимость значений временного ряда выявляют на основе корреляционного и *R/S* анализа. Названные выше методы описаны разрозненно, в разных изданиях и публикациях. В настоящее время интеллектуальная технология, применяемая для анализа динамики показателей на фондовом рынке, представляет собой сочетание всех названных выше методов. Для этого пользуются специальными компьютерными программами. Популярными являются пакеты *Eviews*, *Fractan*, наиболее простым – надстройка *Microsoft Excel*, пакет *AtteStat*. Однако, не зная теоретических основ методов, невозможно правильно задать исходные параметры для расчетов и сделать правильные выводы по полученным результатам. В данной работе обобщен зарубежный и отечественный опыт по теоретическому обоснованию названных методов в аспекте их комплексного применения.

2. Теоретические основы методов интеллектуального анализа.

2.1. Корреляционный анализ.

Ковариация между значениями X_t и X_{t+k} , отделенными интервалом в k единиц времени, называется автоковариацией с лагом (задержкой) k и определяется формулой

$$R(k) = \text{cov}(X_t, X_{t+k}) = M[(X_t - M(X_t))(X_{t+k} - M(X_{t+k}))], \quad k = 1, 2, \dots, n.$$

Для стационарных процессов автоковариация зависит только от лага k и $R(0) = \sigma^2$, где $\sigma^2 = D(X_t)$. Нормированная автоковариация с лагом k называется *автокорреляцией* и равна

$$\rho(k) = \frac{R(k)}{\sigma^2} = \frac{R(k)}{R(0)}$$

и обладает свойством

$$-1 \leq \rho(k) \leq 1.$$

то есть является коэффициентом корреляции и характеризует степень линейной взаимосвязи между X_t и X_{t+k} . Совокупность значений ковариаций при всевозможных значениях k называется *автоковариационной функцией* случайного процесса, а совокупность автокорреляций называется *автокорреляционной функцией*. Автокорреляционная функция показывает, насколько статистически зависимы значения временного ряда при различных сдвигах времени [1]. Статистические оценки приведенных показателей вычисляются по эмпирическим данным временного ряда и обозначаются соответственно $\bar{R}(k)$, $\bar{\rho}(k)$. Оценка автокорреляционной функции называется *коррелограммой* (Рис.1).

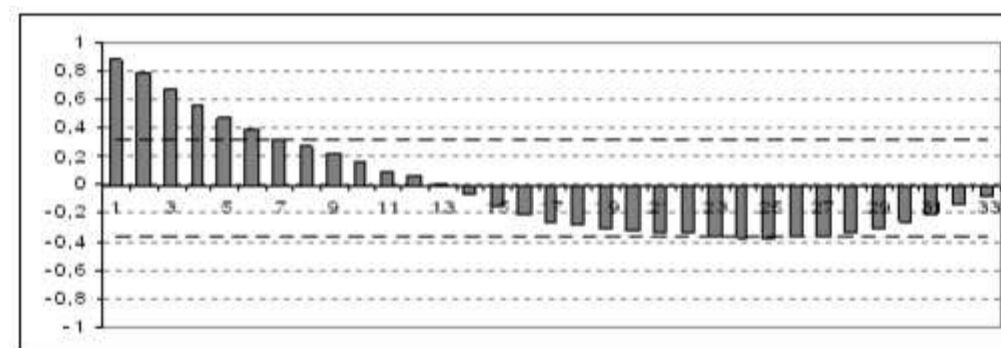


Рис. 1. Коррелограмма

Следует отметить, что при больших значениях k оценка для коэффициента автокорреляции малонадежна, это связано с заменой полного суммирования частичным. Поэтому поведению коррелограммы при больших значениях k не следует придавать большого значения. Для заданного уровня вероятности α (обычно при компьютерных вычислениях берут $\alpha = 0.05$ (5%)) можно вычислить границы доверительного интервала, в котором нахождение значения функции автокорреляции при заданном лаге k с вероятностью $p = 1 - \alpha$ не противоречит предположению об отсутствии корреляции сечений с этим лагом. При графическом изображении функции автокорреляции или ее оценки – коррелограммы, эти интервалы дают две граничные кривые (выше и ниже оси абсцисс). Выход за эти граничные кривые рассматривается как указание на значимость корреляции с соответствующим лагом. Коррелограмма на Рисунке 9 выходит за пределы доверительного интервала на лагах 1- 6. Следовательно для этих лагов имеет место корреляция между X_t и X_{t+k} , $k = \overline{1,6}$. Если случайный процесс или временной ряд близок к белому шуму, то коррелограмма близка к горизонтальной оси, а ее значения близки к нулю. Если ряд стационарный, то автокорреляционная функция, начиная с некоторого момента, затухает или совсем исчезает [1]. Например, для процессов $AR(p)$ и $MA(q)$ она либо имеет вид затухающей синусоиды, либо экспоненциально убывает. Если функция медленно убывает на достаточно продолжительном временном периоде, то можно предположить, что ряд характеризуется долговременной памятью. Долговременная (длинная) память – это свойство, описывающее корреляционную структуру высокого порядка временного ряда, то есть существование значимой автокорреляции в моментах высших порядков. Подтверждение или отклонение этой гипотезы осуществляется на основе *R/S анализа*.

2.2. Спектральный анализ.

Спектральный анализ временных рядов включает:

- спектральное разложение ряда;
- совокупность статистических приемов, позволяющих оценить вероятностные характеристики временного ряда по результатам наблюдений, отображаемых через его спектральное представление.

В настоящее время для спектрального анализа используются *метод Фурье* и *вейвлет-анализ*.

Метод Фурье [3,4]. Любой стационарный процесс можно рассматривать как наложение некоррелированных друг с другом гармонических колебаний различных частот со случайными фазами и амплитудами.

Спектральное разложение случайной функции (случайного процесса) - это разложение в ряд или интеграл по той или иной специальной системе функций такое, что коэффициенты этого разложения представляют собой некоррелированные случайные величины. Говорят, что случайный процесс X_t допускает спектральное разложение, если его можно представить в виде суммы (вообще говоря, бесконечной, то есть в виде бесконечного функционального ряда) элементарных гармонических случайных процессов $C_k \sin(\omega_k t + \varphi_k)$. Такое представление иногда называют полигармонической моделью временного ряда.

Теорема о спектральном разложении. Предположим, что ковариационная функция стационарного случайного процесса X_t может быть представлена в виде $R(\tau) = \sum_{n=0}^{\infty} C_n \cos \omega_n \tau$ (при этом можно доказать, что все коэффициенты C_n будут неотрицательны). Тогда X_t допускает спектральное разложение вида

$$X_t = M(X_t) + \sum_{n=0}^{\infty} A_n \cos \omega_n \tau + B_n \sin \omega_n \tau. \quad (1)$$

где A_n, B_n - некоторые случайные величины, независимые между собой, имеющие нулевое математическое ожидание и равные (при каждом отдельном n) дисперсии, а $M(X_t)$ - математическое ожидание процесса X_t .

Набор $\{\omega_n\}$ неслучайных частот, фигурирующих в этом разложении, называется спектром случайного процесса X_t . В некоторых случаях разложение (1) функции $R(\tau)$ можно найти с помощью ряда Фурье. А именно, если $R(\tau)$ периодична, то она может быть разложена (при некоторых, довольно несущественных на практике, ограничениях на $R(\tau)$) на отрезке $[-T, T]$ в ряд Фурье, причем только по косинусам (так как $R(\tau)$ - четная функция): $R(\tau) = \sum_{n=0}^{\infty} a_n \cos \pi n / l$. Тут спектр имеет вид $\left\{ \omega_n = \frac{\pi n}{l} \right\}$. Коэффициенты a_n вычисляются по обычным формулам

$$\text{Фурье: } a_n = \frac{1}{T} \int_{-T}^T R(\tau) \cos \frac{\pi n}{l} \tau d\tau$$

На самом же деле периодической функция $R(\tau)$ бывает довольно редко (и практически никогда для стационарных процессов). Поэтому предположение о ее периодичности довольно ограничительно и обычно может рассматриваться как довольно грубое приближение к реальности. Но если имеются основания считать $R(\tau)$ периодической, то тут применима спектральная теорема и мы получаем разложение $X(t)$ в сумму элементарных гармоник, частоты которых имеют вид $\frac{\pi n}{l}$. При $l \rightarrow \infty$ спектр становится все более похож на непрерывный. Для непериодических $R(\tau)$ используется не ряд Фурье, а преобразование Фурье (2), которое дает возможность исследования функций во временной ($|t| < \infty$) и частотной ($|\nu| < \infty$) областях.

$$\mathcal{F}(\nu) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi \nu t} dt \quad (2)$$

где $f(t)$ - случайный процесс, $\mathcal{F}(\nu)$ спектр процесса, ν - частота, $i = \sqrt{-1}$.

Преобразование Фурье обладает способностью фокусировать в точку распределенную по времени информацию о периодичности функции при переходе из временной области в частотную. Достигается это за счет того, что ядро преобразования Фурье, функция $e^{-i2\pi \nu t}$, не локализовано во времени, но имеет предельную локализацию в частотной области. Это обстоятельство делает преобразование Фурье прекрасным инструментом для изучения процессов, свойства которых не меняются со временем, но плохим методом для исследования функций, характеристики которых эволюционируют во времени. Для устранения этого недостатка нужно локализовать преобразование Фурье на промежутках конечной длины. Формализация такого подхода может быть описана, например, с помощью оконного преобразования Габора [7]

$$GT(\nu, b, \alpha) = \int_{-\infty}^{\infty} f(t) e^{\frac{(t-b)^2}{\alpha^2}} e^{-i2\pi \nu t} dt \quad (3)$$

Введенное под знак интеграла выражение временное окно $e^{\frac{(t-b)^2}{\alpha^2}}$ выделяет лишь небольшой отрезок исходного временного ряда с центром в произвольной точке b (Рис.2) и позволяет получить эволюцию спектра во времени.

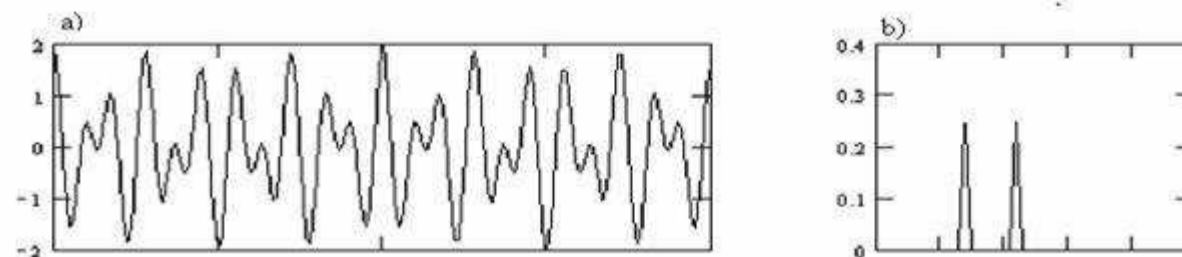


Рис. 2. а) временной ряд , б) выделенное окно

Окно Габора имеет постоянную ширину, которая определяется параметром α . Эффективная ширина окна определяет длину интервала ΔT , который дает главный вклад в значение интеграла в выражении (3). Длина ΔT является мерой временного разрешения, в то время как ширина спектральной линии $\Delta \nu$ определяет меру частотного разрешения. Эти характеристики связаны между собой соотношением $\Delta \nu = \frac{1}{\Delta T}$.

Преобразование Фурье для $R(\tau)$ (если оно существует, то есть если сходится соответствующий несобственный интеграл), называется *спектральной плотностью* стационарного случайного процесса X_t и вычисляется по формуле: $S(\omega) = \frac{2}{\pi} \int_0^{\infty} R(\tau) \cos \omega \tau d\tau$. Спектральная плотность - функция всегда неотрицательная, ее график является важной характеристикой стационарного случайного процесса. Сказанное здесь о случайных процессах относится и к временным рядам. Для них тоже используется спектральное разложение, которое широко применяется при практическом исследовании временных рядов.

На практике используется оценка спектральной плотности, которая называется *периодограммой*. Периодограмма является асимптотически несмещенной, но несостоятельной оценкой и применяется для выявления скрытых периодичностей. Вычисления периодограммы основаны на БПФ (Быстром Преобразовании Фурье) и проводятся только при наличии компьютера с соответствующей программой. При анализе периодограммы нужно обращать особое внимание на ее пики. Большой пик в области некоторой частоты ν_0 указывает на то, что в спектральном разложении автокорреляционной функции присутствует цикл. Чем выше и резче выделен пик, тем большая часть мощности сосредоточена около частоты ν_0 и тем большую роль играет эта частота в описании соответствующего случайного процесса или временного ряда. На Рис. 3 представлена периодограмма ряда, которая указывает на то, что спектр амплитуд содержит два пика, то есть анализируемый ряд имеет два цикла.



Рис. 3. Периодограмма

На практике периодограмма применяется для поиска оптимально подходящего параметра усреднения скользящих средних, с целью фильтрации информации. Напомним, что фильтрацией называют такое преобразование исходной функции, которое приводит к изменению ее спектрального состава. Рынок не постоянен и на нём присутствует много шумов. Шумы формируют ложные впадины и пики. Частично проблема решается различными фильтрами, например подбором скользящих средних. Так, скользящая средняя с параметром усреднения M дней может обрезать (сократить) амплитуду колебаний любой из применяемых гармоник, период которой будет меньше значения M .

Условно можно разделить спектры для дневных графиков поведения активов на рынке на несколько спектральных полос:

- низкочастотные - диапазон частот от 0 до 4 циклов в год. Как правило, это долгосрочные тренды, им соответствует цикличность более 60 дней;
- среднечастотные - диапазон частот от 5 до 40 циклов в год. Обычно это среднесрочные и краткосрочные циклы. Им соответствует период цикличности в 6-60 дней;
- высокочастотные - диапазон частот от 41 до 130 циклов в год. Это рыночный шум, на который приходится большая часть спектра. Соответствующий период цикличности менее шести дней.

Таким образом, можно фильтровать частоты, неинтересные для торговли, заменяя исходный ряд данных на скользящую среднюю с соответствующим параметром усреднения M . Для трейдеров, работающих на долгосрочных частотах, анализ спектра помогает подобрать период скользящей, сохраняющий силу низких частот, и снижает помехи оказываемые средними и высокими частотами, убирая их пики. Трейдеры, работающие на средних частотах могут без проблем фильтровать высокие и низкие частоты, оставляя нужные сильные пики средних частот.

Для детерминированной непериодической составляющей периодограмма бесконечно возрастает на низких частотах. Если случайный процесс или временной ряд близок к белому шуму, то периодограмма близка к горизонтальной оси, а ее значения близки к нулю.

Вейвлет анализ [5]. При использовании преобразования Габора (3) возникает проблема выбора ширины окна во временной области. Слишком широкое окно может обеспечить разумное представление низкочастотных компонент ряда, но его ширина будет избыточной для гармоник с высокой частотой. Наоборот, достаточно узкое окно даст возможность изучить вариации во времени высокочастотных компонент, но оно не будет адекватным для низкочастотных гармоник. Если сделать оконную функцию зависящей от частоты так, чтобы для низких частот окно становилось шире, а для высоких уже, то оконное преобразование Фурье переходит в новый класс преобразований, который получил название вейвлет-преобразования. Термин «вейвлет» происходит от английского слова wavelet, буквальный перевод которого означает «маленькая волна» или «всплеск».

Интегральным вейвлет – преобразованием функции $f(t) \in L^2(R)$ называется выражение

$$W(a,b) = \frac{1}{|a|^2} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt ,$$

где $a, b \in R, a \neq 0$, символом * обозначена процедура комплексного сопряжения, функция $\psi(t)$ называется *вейвлетом (анализирующим, базисным или материнским вейвлетом)*. Параметр a определяет размер вейвлета и называется масштабом (*scale*). Его аналогом в Фурье-анализе является период (частота) гармонического колебания. Понятие *масштаба* – более широкое, чем понятие периода. Это связано с тем, что в Фурье-преобразовании функциональный вид ядра преобразования зафиксирован раз и навсегда, в то время как вейвлет-преобразование одной и той же функции можно получить с помощью различных базисных вейвлетов, то есть в разных системах масштабов. Параметр b задает временную локализацию вейвлета и называется *сдвигом (shift)*. Этот параметр не имеет аналога в Фурье-преобразовании.

Двухпараметрическая функция $W(a,b)$ дает информацию об изменении относительного вклада компонент разного масштаба во времени и называется спектром коэффициентов вейвлет-преобразования. Таким образом для каждого значения масштабного коэффициента $a > 0$ величины $W(a,b)$, определенные с помощью вейвлет-преобразования, представляют собой результат фильтрации исходной функции в диапазоне частот, центр которого определяется значением масштабного коэффициента, а размер - свойствами принятого анализирующего вейвлета.

Помимо интегральных вейвлет - преобразований в отдельных приложениях (численное моделирование, сжатие изображений и т.д.) нашли широкое применение *дискретные вейвлет-преобразования*, основанные на использовании целочисленных сдвигов и задании масштабов степенями двойки. Опыт показал, что для анализа временных рядов непрерывные преобразования более удобны, чем дискретные.

Располагая вейвлет-спектром, можно рассчитать полную энергию ряда

$$E_f = \int_{-\infty}^{\infty} f^2(t) dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |W|^2(a,b) \frac{da db}{a^2} .$$

Величина

$$E_w(a, b) = |W|^2(a, b) \quad (4)$$

называется плотностью спектра энергии, она определяет спектральную характеристику не только для заданного масштаба, но и для параметра сдвига b . По этой причине ее называют *локальным спектром энергии*. В противоположность этому величину

$$E_w(a) = \int_{-\infty}^{\infty} |W|^2(a, b) db = \int_{-\infty}^{\infty} E_w(a, b) db \quad (5)$$

называют *глобальным спектром энергии*.

Оценку локального спектра энергии (4) называют *скалограммой (scalogram)*, подчеркивая тем самым ее способность описывать распределение энергии по масштабам. Значения скалограммы лежат на поверхности, для изображения которой используют два способа:

- 1) представление поверхности в трехмерном пространстве координат (a, b, S) ;
- 2) плоское изображение поверхности в виде топографической карты.

Масштаб третьего измерения указывается либо оцифровкой изолиний, либо раскраской областей между изолиниями. Такие графики можно построить с помощью графических пакетов *Mathcad, Matlab, Surfer* и др. (Рис.4а, 4б).

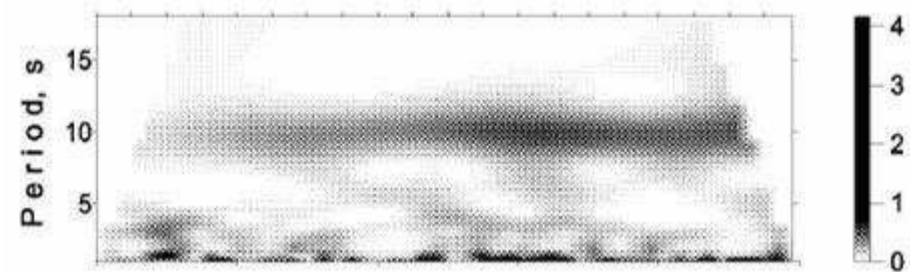


Рис. 4а . Скалограмма временного ряда в виде топографической карты

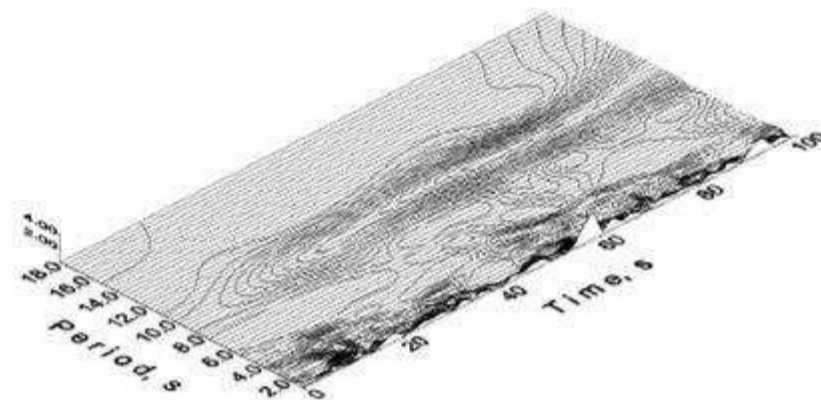


Рис. 4б. Скалограмма в виде поверхности

Бывает так, что широкие контуры линий, близких по частоте гармоник мешают проследить за эволюцией их частот по времени. Чтобы отсечь влияние контуров, можно выделить те точки скалограммы, в которых она имеет максимумы по переменным a и b . Такую функцию называют *скелетом*. В случае синусоидального ряда, точки скелетона располагаются вдоль линий, идущих параллельно оси времени. Если в данных имеются гармонические или квазигармонические компоненты, то топографическая карта скелетона будет состоять из линий, ориентированных вдоль оси периодов b . В случае шумовой компоненты, линии скелетона вытягиваются в перпендикулярном направлении, то есть параллельно оси времени a . Если в данных присутствуют и гармонические компоненты и шум, то карта скелетона позволяет увидеть их раздельно (Рис. 5,6).

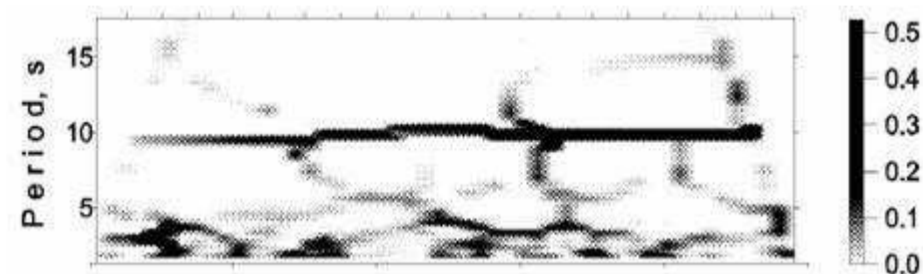


Рис. 5. Скелетон скалограммы в виде топографической карты

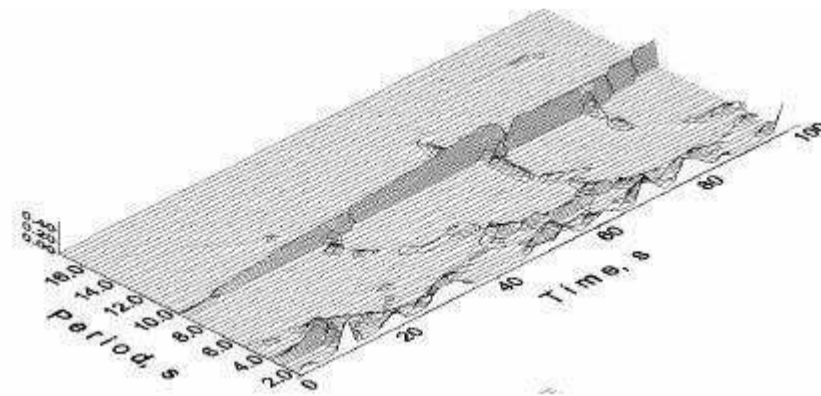


Рис. 6. Скелетон скалограммы в виде поверхности.

Рис. 6 показывает линии скелетона в направлениях осей времени и периодов. Можно сделать вывод, что в состав исследуемой последовательности входят как квазигармонические компоненты, эволюционирующие во времени - линии, идущие по направлению оси времени, так и чисто шумовые компоненты, линии, идущие перпендикулярно оси времени.

Выбор анализирующих функций диктуется тем, какую информацию нужно извлечь из временного ряда. Для более полного выявления особенностей анализируемого ряда, используют различные вейвлеты.

В спектральном анализе временных рядов выделение значимой компоненты из шума производится с заданной вероятностью на основе статистических критериев. Обычно в качестве стохастической компоненты временного ряда принимается *белый шум*. Критерий выделения шума позволяет построить массив значений скалограммы, которые с заданной вероятностью $p = 1 - q$ принадлежат значимой, а не шумовой составляющей. Иногда выделяют шум не из самой скалограммы, а из ее скелетона. На рисунках 7 и 8 представлены линии скелетона в виде топографической карты и поверхности соответственно.

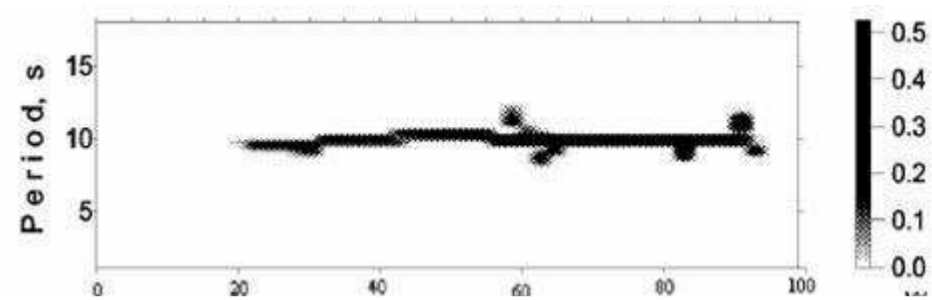


Рис. 7. Значимая компонента скелетона, выделенная из шума с вероятностью 0,99 в виде топографической карты

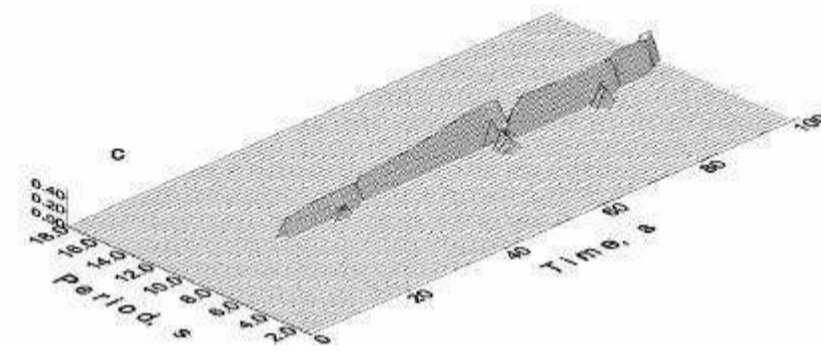


Рис. 8. Значимая компонента скелетона, выделенная из шума с вероятностью 0,99 в виде поверхности

Оценку глобального спектра энергии (5) называют *нормированной скалограммой* или *скейлограммой (scalegram)*. Скейлограмма является прямым аналогом сглаженной периодограммы в Фурье-анализе. График скейлограммы показан на Рис.9.

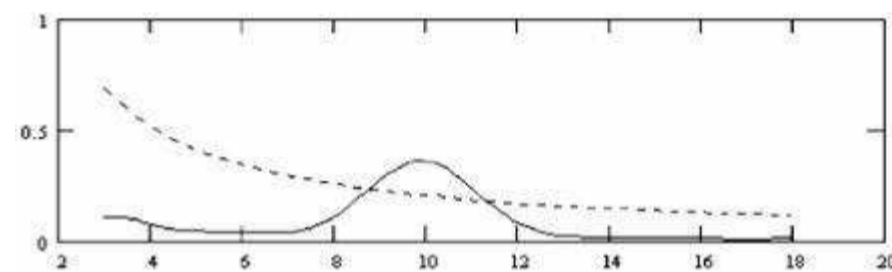


Рис.9. Скейлограмма

Штриховой линией обозначен порог обнаружения шума. Значимой компоненте соответствует часть графика, расположенная выше штриховой линии, а все остальное - шумы.

Спектральный анализ позволяет выявить наличие детерминированных компонент во временном ряде. Если эти компоненты удалить из ряда, то можно получить стационарный ряд. Бокс и Дженкинс [8] предложили компоненту $f(t)$ исключать дифференцированием, то есть вычислением последовательных разностей. Если ряд после вычисления d последовательных разностей приводится к стационарному, то такой ряд называют *ARIMA(p,d,q)* (*авторегрессионный проинтегрированный процесс скользящего среднего – Autoregressive Integrated Moving Average*) порядка (p,d,q) , которые моделируют различные ситуации, встречающиеся при анализе стационарных и нестационарных рядов. При этом p параметр *AR* - части, d – степень интеграции, q – параметр *MA* - части. Периодическую составляющую удаляют вычитанием ее из ряда. Однако, следуя методологии Бокса и Дженкинса, можно сделать ошибочные выводы [9]. Дело в том, что в экономике долгое время игнорировалась возможность дробного значения параметра d . В работах зарубежных ученых, в первую очередь, *C.W.Granger, J.R.Hosking, P.M.Robinson, R. Beran* [10], был предложен новый класс моделей *ARFIMA(p,d,q)* (*авторегрессионный дробно - интегрированный процесс скользящего среднего - Autoregressive Integrated Fractional Moving Average*), допускающий возможность нецелого параметра d . Характеристики таких временных рядов обладают важными свойствами, например, X_t является стационарным и обратимым при $d \in (-1/2, 1/2)$. Если исследуемый ряд имеет тренд и долговременную память, то можно предположить, что он представляет собой авторегрессионный дробно - интегрированный процесс скользящего среднего.

2.3. R/S анализ.

В рамках созданного Э. Херстом *R/S* анализа [11] используется уравнение

$$\log(R/S) = \log c + H \log(n), \quad (6)$$

где $R = \left[\max_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \bar{X}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^k (X_j - \bar{X}_n) \right]$ - размах частичных сумм отклонений значений ряда X_j от его среднего значения, $S = \sqrt{\frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)^2}$ - стандартное отклонение ряда, $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ - среднее значение ряда, H - показатель Херста.

Показатель Херста – это тангенс угла наклона прямой (6) к оси абсцисс в координатах $\log(R/S), \log(n)$. Оценку показателя Херста находят из эмпирического уравнения линейной регрессии. Для этого вычисляют размах R/S для подпоследовательностей различной длины исходного ряда. Затем по методу наименьших квадратов находят оценку угла наклона прямой к оси абсцисс $\log(n)$. Алгоритм предложен Э. Петерсом и подробно описан в [11]. Показатель Херста связан с величиной d модели *ARFIMA(p,d,q)* соотношением $H = d + 0,5$. В результате могут получиться следующие оценки:

1) $0,5 < H \leq 1,00$ - такое значение указывает на то, что ряд обладает персистентным свойством. Ряд характеризуется эффектами долговременной памяти и имеет склонность следовать трендам. Если ряд возрастает (убывает) в предыдущий период, то вероятно, что он будет сохранять эту тенденцию какое-то время в будущем. Трендоустойчивость поведения увеличивается при приближении H к единице. Персистентный временной ряд является самым распространенным типом, встречающимся в природе, а также в экономике, и на фондовых рынках. Такой ряд называют *дробным броуновским движением* или *обобщенным броуновским движением* или *фракталом*.

2) $0 \leq H \leq 0,5$ - означает антиперсистентность, то есть ряд меняет направление чаще, чем случайный ряд. Наблюдения зависимы, каждое из них несет память о предыдущих событиях. Рынок неустойчив. Чем ближе значение H к нулю, тем неустойчивей динамика цен (за подъемом следует спад и наоборот).

3) $H = 0,5$ - соответствует случайному ряду. Наблюдения случайны и независимы. Такой процесс полностью лишен памяти. Рынок стагнирует, нет никаких движений или они циклически с очень большой частотой колебаний.

4) $H > 1$ - очень редкое явление. Возникают независимые скачки амплитуды, распределенные по Леви [12].

Херст сформулировал свой *эмпирический закон* [13]. Он предложил формулу для оценки величины H по значению R/S :

$$H = \log(R/S) / \log(n/2), \quad n - \text{количество наблюдений.}$$

В этой формуле предполагается, что в соотношении (6) константа $c = \frac{1}{2}$. Федер показал [13], что этот эмпирический закон имеет тенденцию преувеличивать H , когда $H > 0,7$, и, наоборот, преуменьшать, если $H < 0,4$. Однако для коротких рядов, где регрессия невозможна, этот эмпирический закон может быть использован как разумное приближение.

При подсчете показателя Херста необходимо проверять гипотезу о его значимости. В работе [15] предлагается рассматривать вопрос о поведении среднего значения $E_0(R/S)$, где усреднение E_0 отвечает справедливости гипотезы H_0 о случайном ряде. Данные по значениям $E_0(R/S)$ можно найти в [11]. К сожалению, R/S анализ чувствителен к краткосрочным зависимостям. Э. Ло обнаружил этот факт и модифицировал статистику таким образом, чтобы ее статистические свойства оставались инвариантными для обширного класса процессов с короткой памятью, но изменялись для процессов с длинной памятью. Дело в том, что дисперсия частичных сумм не равна сумме дисперсий отдельных членов, она должна включать и автоковариации. Следовательно, оценка должна отражать не только суммы квадратов отклонений, но и взвешенные автоковариации до некоторого лага q . Поэтому Э. Ло предложил заменить величину S в знаменателе R/S более сложной суммой [16]:

$$\sigma_n^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)^2 + \frac{2}{n} \sum_{j=1}^q \varpi_j(q) \left(\sum_{i=j+1}^n (X_i - \bar{X}_n)(X_{i-j} - \bar{X}_n) \right), \quad \varpi_j(q) = 1 - \frac{j}{q+1}, q < n.$$

Веса $\varpi_j(q)$ обеспечивают положительность σ_n^2 . Однако слишком большие (по сравнению с n) значения q приводят к тому, что свойства оценок в конечных выборках будут существенно отличаться от их асимптотического поведения. Но брать слишком маленькие значения q тоже нельзя, так как автоковариации за лагом q могут оказаться важными, и их следовало бы включить во

взвешенную сумму. Д. Эндрюс ввел для выбора q правило, основанное на свойствах исходных данных [17]. В качестве q берут целую часть числа $k_n = \left(\frac{3n}{2}\right)^{1/3} \left(\frac{2\rho}{1-\rho^2}\right)^{2/3}$, где ρ

есть оценка коэффициента автокорреляции первого порядка X_1 .

Далее рассматривается статистика $V_n(q) = \frac{R}{\sigma_n \sqrt{n}}$. Асимптотическое распределение стандартизованной статистики V_n сходится к распределению случайной величины V - размаху броуновского моста на единичном интервале [17]. Классификация ряда производится на основе проверки статистической гипотезы H_0 - исследуемый ряд является случайным. В таблице 1 приведены интервалы, соответствующие области принятия гипотезы H_0 . Гипотеза принимается на соответствующем уровне значимости, если статистика V_n попадает в какой-либо из интервалов таблицы. В противном случае, можно считать, что с вероятностью $1 - \alpha$ имеет место долговременная зависимость.

В работе [18] отмечено, что описанный тест склонен принимать гипотезу H_0 , если ее на самом деле нет. Тем не менее, в научных исследованиях, как правило, обращаются именно к этому тесту.

Таблица 1.
Область принятия гипотезы H_0 для статистики V_n

Уровень значимости α	0,005	0,025	0,05	0,1	0,2	0,3	0,4
Интервал принятия гипотезы	0,721 - 2,098	0,809 - 1,862	0,861 - 1,747	0,927 - 1,620	1,018 - 1,473	1,090 - 1,374	1,157 - 1,294

R/S анализ позволяет определить среднюю длину цикла, необходимую для оценки инерции развития, то есть интервал времени, по истечении которого теряется память о начальных условиях. Если график функции $\log R/S$ имеет отчетливо выраженный линейный участок, то это указывает на существование цикла на этом участке. Временной диапазон, соответствующий линейному участку, является длиной цикла. Длину цикла можно определить и по графику зависимости статистики V_n . Рост V_n - статистики вместе с возрастанием n означает, что данному ряду соответствует показатель $H > 0,5$, и имеет место цикл. Длина как периодического, так и непериодического цикла равна длине интервала, на котором сохраняется тенденция роста статистики. [11]. Если же $V_n \rightarrow const$ (стабилизируется) при росте n , то это соответствует белому шуму.

3. Анализ фондового рынка Украины.

Основными индикаторами фондового рынка являются индексы, рассчитываемые на основании котировок определенной группы ценных бумаг. Индекс фондового рынка является своего рода измерительным инструментом, позволяющим инвестору выносить суждение о состоянии рынка в целом.

Проведем интеллектуальный анализ индекса ПФТС, так как именно этот индекс признан международной финансовой корпорацией (МФК) как единственный индекс, используемый этой организацией при мониторинге внутреннего состояния украинского фондового рынка.

На рисунках 10-13 приведены основные характеристики временного ряда индекса ПФТС. Данные временного ряда - это ежедневные значения индекса на момент закрытия торгов за период 2004-2010г.г. (сайт www.pfts.ua (<http://www.pfts.ua/>))

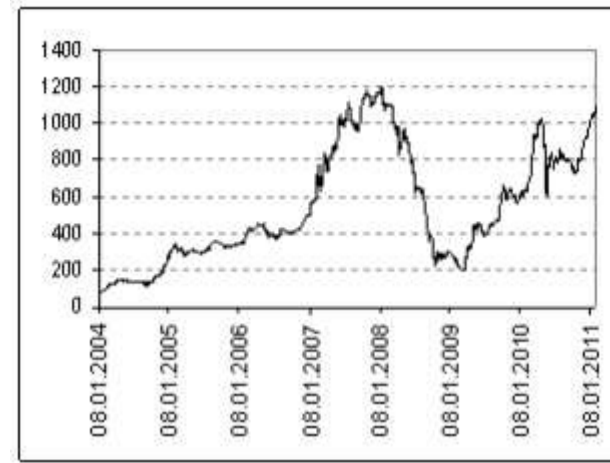


Рис. 10. Временной ряд индекса ПФТС

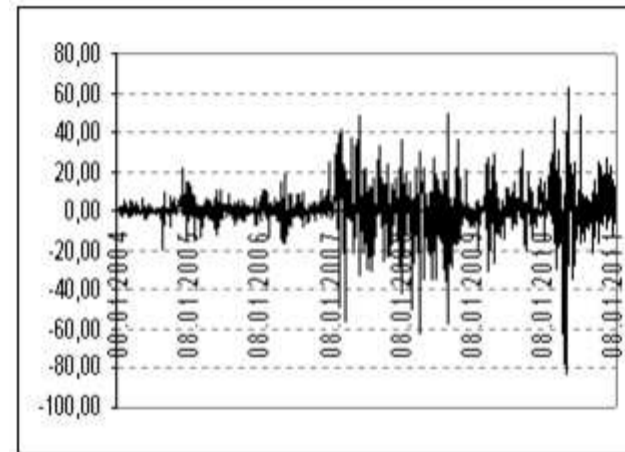


Рис. 11. Скорость изменения индекса ПФТС

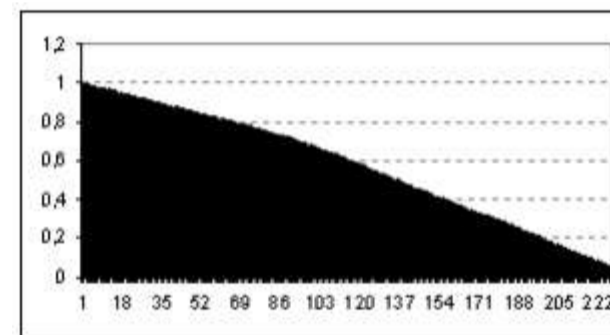


Рис. 12. Коррелограмма

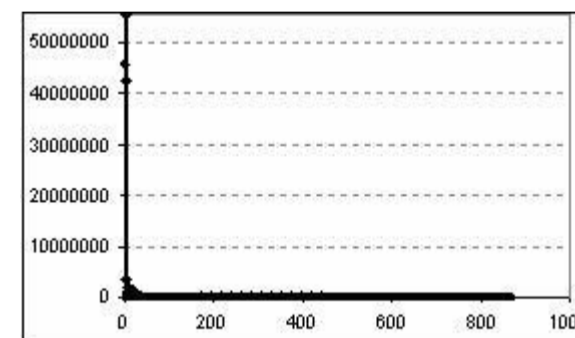


Рис. 13. Периодограмма

По графику временного ряда видно, что он носит скачкообразный характер, изменения могут отличаться друг от друга на порядок. Скорость изменения индекса имеет хаотичный характер, причем в докризисный период скорость оставалась практически постоянной, а для кризисного и послекризисного периодов характерны резкие и более хаотичные колебания. Ряд характеризуется наличием возрастающего тренда, долгой положительной автокорреляционной связью, сосредоточением мощности процесса вблизи нулевой частоты. Указанные характеристики позволяют предположить, что временной ряд обладает эффектом долговременной зависимости (памяти). Показатель Херста $H = 0,9652 \pm 0,1078$ и тест для V_n подтверждают этот вывод ($V_n = 0,3016$ и не попадает ни в один интервал таблицы 1). Это свидетельствует о том, что ряд с большой вероятностью является персистентным и трендоустойчивым (Рис.14).

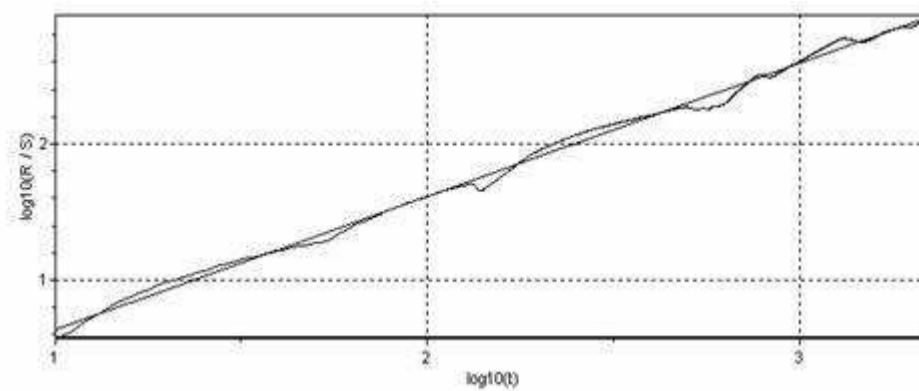


Рис. 14. Зависимость нормированного размаха в двойном логарифмическом масштабе и ее линейная аппроксимация

Циклам соответствуют точки на графике, в которых совпадают линейная аппроксимация и зависимость нормированного размаха. Из рисунка (14) видно, что графики совпадают практически на всем диапазоне. Имеющиеся отклонения можно считать незначительными, они находятся в пределах ошибки. Следовательно, циклы на рынке отсутствуют. Это может быть по двум причинам: либо недостаточно данных, либо цикла не существует. В данном случае, по имеющимся данным, трудно сделать однозначный вывод.

В долгосрочной перспективе можно ожидать, что рынок продолжит восходящую тенденцию. Однако, следуя тренду, рынок многократно менял направление движения, то есть в краткосрочной перспективе поведение рынка является неопределенным. Противоречие между долговременной устойчивостью и кратковременной нестабильностью свидетельствует о том, что имеет место *детерминированный хаос* [19,20], что позволяет классифицировать ряд, как нелинейную динамическую систему [19,20,21]. Система описывается дифференциальными уравнениями, ее невозможно моделировать эконометрическими процессами.

Статистический и спектральный анализ выполнен с помощью программы *AtteStat*, фрактальные характеристики вычислены программой *Fractan*. Эти программы распространяются бесплатно и помещены на многих Internet-сайтах, например, <http://soft.softodrom.ru/ap/Fractan-p44195> (<http://soft.softodrom.ru/ap/Fractan-p44195>), <http://www.twirpx.com/file/166961/>

Выводы.

Проведенное исследование показало, что природа временных рядов динамики фондового рынка достаточно сложная. На рынке присутствует хаотическая составляющая. Для изучения такой динамической системы привлекают теорию хаоса. Эта теория позволяет вычислить характеристики, по которым определяется степень сложности и предсказуемости системы, а также ее эволюция в фазовом пространстве [20,21]. Однако, математическая модель системы, остается неизвестной. В таких случаях, в качестве инструментов моделирования и прогнозирования целесообразно использовать интеллектуальные технологии нечеткой логики, нейронных сетей и генетических алгоритмов.

Литература.

1. Канторович Г.Г. Анализ временных рядов// Экономический журнал ВШЭ -№1. - 2002.- С. 85-116.
2. Wold H. A Study in the Analysis of Stationary Time Series. /H. Wold// Сб.научн.трудов. Stockholm: Almqvist and Wiksel, 1938.
3. Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения.- Вып. 1,2 – М.: Мир, 1971.- 608с.
4. Гренджер К., Хатанака М. Спектральный анализ временных рядов в экономике.- М.: Статистика, 1972.
5. Витязев В.В. Вейвлет-анализ временных рядов/ Учебное пособие.- Изд- во С.-Петербургского университета, 2001. -58с.
6. Гесць В.М., Клебанова Т.С., Черняк О.І., Иванов В.В., Дубровіна Н.А., Ставицький А.В. Моделі і методи соціально-економічного прогнозування/ Підручник. -Х.: ВД «ІНЖЕК», 2005. - 396с.
7. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в Matlab.-М.:ДМК Пресс, 2005.-304с.
8. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление/ Дж.Бокс, Дженкинс Г. - Вып. 1, 2 - М.:Мир, 1974.-197с.
9. Cambell, J.Y.& N.G.Mankiw. Are output fluctuations transitory? Quarterly Journal of economics 102.-1987.- С.875-880.
10. Granger C.W.J. Some Properties of Time Series Data and Their Use in Econometric Model Specification /C.W.J. Granger C.W.J.// Journal of Econometrics, 1981.-Vol.16.- №1.- P. 121-130.
11. Петерс Э. Фрактальный анализ финансовых рисков. Интернет- трейдинг, 2004.- 304 с.
12. Феллер В. Введение в теорию вероятностей.-Т.2.-М.:Мир.-1976.-751с.
13. Hurst H.E. Long term Storage Capacity of Reservoirs//Transactions of the American Society of Civil of engineers.-1951/- N 116/-P/770-799.
14. Федер Е. Фракталы. Пер. с англ.-М.: Мир,1991.- 254с.
15. Anis A.A., Lloyd E.H. The expected value of the adjusted rescaled Hurst range of independent normal summands // Biometrika. 1976. V. 63 №1. P. 111-116.
16. Lo A.W. Long Term Memory in Stock Market Prices//Econometrica.-1991.-№59.-P.1279-1313.
17. Andrews D. Non-Strong Mixing Autoregressive Process//Journal of Probability.-1984.-№21.-P.- 930-934.
18. Teverlosky V., Taqqu M. S., Willinger W. A critical look at Lo modified R/S statistic // Journal of Statistical Planning and Inference. 1999. N 80. P. 211–227.
19. Мандельброт Б. Фракталы, случай и финансы. – М.: Регулярная и хаотическая динамика, 2004 г. - 256 с.
20. Петерс Э. Хаос и порядок на рынках капитала.-М: Мир, 2000.- 238с.
21. Шустер Г. Детерминированный хаос. М.:Мир,1988.-240с.

Стаття надійшла до редакції 29.03.2012 р.