УДК 519.25, 53.023,52-17

В. М. Журавлев, В. М. Морозов, М. С. Петряков, В. В. Самойлов

МЕТОД УСЛОВНОЙ ДЕКОМПОЗИЦИИ ЭМПИРИЧЕСКИХ РАСПРЕДЕЛЕНИЙ И ЕГО ПРИМЕНЕНИЕ К ЗАДАЧЕ АНАЛИЗА РЯДОВ НАБЛЮДЕНИЙ[1]

Аннотация.

Актуальность и цели. Одним из способов анализа рядов наблюдений является анализ эмпирических вероятностных распределений (гистограмм). Задача при таком подходе – выяснение фундаментальных свойств физических процессов, ответственных за изменчивость наблюдаемых параметров физических и других систем. Одним из способов обнаружения всех действующих физических факторов в наблюдаемой системе является метод декомпозиции эмпирических распределений. Декомпозиция позволяет представить гистограмму в виде смеси, каждый из компонент которой может интерпретироваться как вероятностное распределение одного из механизмов со специфическими признаками. Стандартным подходом к декомпозиции является метод моментов в сочетании с заранее заданным набором теоретических распределений, которые выбираются до проведения самой декомпозиции. В этом случае сами признаки разделения распределения на компоненты фактически исключаются из анализа, что часто приводит к трудностям в интерпретации полученных результатов. Поэтому актуальная задача обработки рядов – разработка метода декомпозиции гистограммы с помощью эмпиричесих признаков, которые непосредственно участвуют в обработке данных. Цель данной работы – построение метода декомпозиции рядов наблюдений с помощью формирования эмпирических признаков разделения значений ряда на основе статистических характеристик самого ряда.

Материалы и методы. Для реализации метода декомпозиции важным является требование выработки статистически устойчивых признаков, подлежащих проверке во время работы алгоритмов. Устойчивые признаки на базе самого исходного ряда наблюдений можно построить, используя те или иные статистики. Поскольку каждый признак должен относиться к каждому отдельному элементу ряда, то в данной работе используются два метода. Это метод регрессионных моделей и метод вычисления базовых статистик скользящих рядов.

Результаты. Основным результатом работы является создание математических алгоритмов проведения условной декомпозиции и его применение к задаче декомпозиции эмпирических распределений ряда чисел Вольфа (ежемесячное число пятен на Солнце) и ежечасного ряда атмосферного давления за 2009 г. Найдены компоненты распределений, и на основе скользящих рядов проанализирована изменчивость параметров эмпирических распределений и эволюция априорных вероятностей.

Выводы. Предложенный метод условной декомпозиции дает значительно более эффективный способ разделения гистограмм на компоненты, чем методы декомпозиции, основанные на методе моментов для теоретически заданных распределений смеси. Метод может применяться для большинства систем при условии, что сформулированы основные принципы выявления устойчивых признаков на основе самих рядов наблюдений. Показано, что для этого

[1] Работа выполнена при поддержке Министерства образования и науки РФ (в рамках Государственного задания и проекта № 14.Z50.31.0015), а также при частичной финансовой поддержке РФФИ в рамках проекта 13-01-97067 р_поволжье_а.

можно использовать как методы регрессионных моделей, так и методы вычисления базовых статистик, таких как дисперсия, для скользящих рядов наблюдений.

Ключевые слова: статистические характеристики рядов наблюдений, эмпирические распределения вероятностей, метод декомпозиции распределений, эволюция статистических параметров, скользящие ряды наблюдений.

V. M. Zhuravlev, V. M. Morozov, M. S. Petryakov, V. V. Samoylov

THE METHOD OF CONDITIONAL DECOMPOSITION OF EMPIRICAL DISTRIBUTIONS AND ITS APPLICATION TO THE ANALYSIS OF A SERIES OF OBSERVATIONS

Abstract.

Background. One of the ways to analyze a series of observations is to analyze the empirical probability distributions (histograms). The challenge with this approach is to find the fundamental properties of the physical processes responsible for the variability of the observed parameters of physical and other systems. One of the ways to identify all existing physical factors in the observed system is the decomposition method of empirical distributions. Decomposition enables a histogram to be represented as a mixture, each component of which can be interpreted as the probability distribution of one of the mechanisms; with specific characteristics. The standard approach to the decomposition is the method of moments in conjunction with a predetermined set of theoretical distributions that are chosen to undertake the actual decomposition. In this case the signs of separation are actually excluded from the analysis, which often leads to difficulties in the interpretation of results. Therefore, the topical problem of series processing is the development of a decomposition method using a histogram of empirical features that are directly involved in the data processing. The aim of this work is to construct the decomposition method, a series of observations using the empirical evidence of the formation of a number of separate values based on the statistical characteristics of the series.

Materials and methods. In implementation of the decomposition method there is an important requirement to generate statistically stable signs to be checked at the time of the algorithms. Persistent signs on the basis of the initial series of observations can be constructed using these or other statistics. As each characteristic must relate to each individual series element, in this study the authors used two methods. These are the method of regression models and the method of calculating the basic statistics of moving rows.

Results. The main result is the creation of mathematical algorithms of conventional decomposition and its application to the decomposition of the empirical distributions of a number of Wolf numbers (monthly number of sunspots) and hourly series of atmospheric pressure in 2009. The authors found components of distributions and analyzed the variability of the parameters of empirical distributions and the evolution of the a priori probabilities using moving series.

Conclusions. The proposed method of conditional decomposition gives a much more efficient method of separating components in the histogram than decomposition methods, based on the method of moments for the theoretical given mixture distributions. The method is suitable for the majority of systems, provided the basic principles of sustainable identifying features are formulated on the basis of the series of observations themselves. It is shown that for this purpose one can use both the methods of regression models and the methods of basic statistics calculation, such as variance, for moving series observations.

Key words: statistical characteristics of a series of observations, empirical probability distribution, method of decomposition of distributions, evolution of the statistical parameters, sliding series of observations

Введение

Принцип декомпозиции эмпирических распределений вероятностей (гистограмм) состоит в представлении этого распределения в виде смеси теоретических распределений, каждое из которых отвечает за отдельный физический механизм изменения изучаемого параметра системы. Под смесью понимается формальная сумма отдельных плотностей распределений $\rho_a(x \mid q_a)$, $a = 1,\ldots,K$, физического параметра x с весовыми коэффициентами p_a:

$$\rho(x) = \sum_{a=1}^{K} p_a \rho_a(x \mid q_a). \tag{1}$$

Это соотношение можно рассматривать как формулу полной вероятности, в которой параметры p_a представляют собой априорные вероятности появления признака q_a с номером a в системе, а $\rho_a(x \mid q_a)$ – условные вероятности появления значения x изучаемого параметра при условии, что реализуется признак q_a. Признаки q_a являются несовместными и отражают реализацию в системе определенного физического механизма появления значения x, изучаемого параметра. Задача декомпозиции состоит в том, чтобы с помощью каких-либо вычислений отыскать значения априорных вероятностей p_a и всех параметров q_a теоретических распределений. С математической точки зрения чаще всего задача декомпозиции эмпирических распределений в смесь теоретических распределений осуществляется с помощью метода моментов, хорошо известного в математической статистике [1, 2]. При реализации такого подхода задают явный вид распределений смеси, а параметры этих распределений вычисляют с помощью решения системы нелинейных алгебраических уравнений, число которых и их порядок зависят от выбора теоретических распределений смеси. При этом приходится сталкиваться с серьезной проблемой решения такой системы уравнений и отбора корней, которые реально соответствуют решаемой задаче. Пример реализации такого подхода приведен в работах [3, 4], где проводилась декомпозиция распределения ряда чисел Вольфа солнечной активности с целью выявления механизмов, влияющих на характеристики цикла солнечной активности.

Чаще всего причиной, по которой проводится декомпозиция распределений, является многомодовый характер эмпирических гистограмм, выходящий за уровень случайных флуктуаций. Пример приведен на рис. 1, где показано эмпирическое распределение температуры воздуха на одной из метеостанций за 2009 г. Как видно, гистограмма имеет характерный двухмодовый вид. Каждая мода (или страта) с точки зрения физики обычно отвечает за некоторый особый механизм динамики системы и выявление свойств этих мод позволяет получить некоторые данные об этих механизмах. Однако основной трудностью при реализации декомпозиции является даже не математическая сложность решения систем уравнений (как правило, нелинейных), а окончательная интерпретация самих механизмов формирования отдельных распределений смеси, что является некорректной задачей в смысле явной неоднозначности возможных интерпретаций.

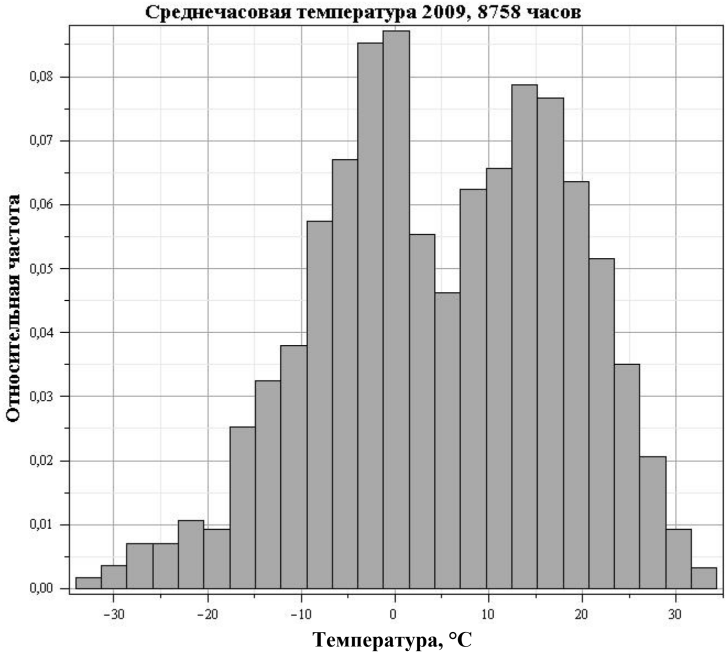


Рис. 1. Пример гистограммы распределений, построенной по ряду температуры, длиной 12 месяцев с шагом 1 ч

В настоящей работе развивается общая методология метода условной декомпозиции распределений. Основная идея метода состоит в представлении исходного эмпирического распределения не в виде смеси теоретических распределений, а в виде именно эмпирических распределений. Распределения смеси формируются в процессе вычислений с помощью проверки соответствия каждого значения исследуемого ряда некоторому эмпирическому признаку. Признаки должны выбираться таким образом, чтобы проверку на соответствие можно было бы осуществлять в ходе вычислений непосредственно по самим рядам наблюдений или их совокупности. В работе метод применяется к задачам анализа ряда чисел Вольфа и метеорологического ряда наблюдений давления.

1. Принцип условной декомпозиции

Основной принцип условной декомпозиции распределений состоит в выработке эмпирических признаков, соответствующих каждому конкретному измерению исследуемого ряда, на основе всей совокупности измерений. Поскольку признаки должны быть относительно статистически устойчивыми, то в качестве значений признаков предлагается использовать какие-либо статистические модели параметров наблюдаемого ряда. Простейшим способом выработки признаков могут служить простые условия, связанные с попаданием значения ряда в некоторые заранее заданные интервалы значений. В этом случае процедура построения декомпозиции исходной гистограммы может быть представлена следующим образом. Пусть x_i, $i=1,\dots,N$, – исходный ряд наблюдений с дискретным временем $t_i = i\Delta$ (Δ – шаг по времени), а X_i – усредненная его модель, построенная с помощью, например, метода наименьших квадратов (МНК) [1, 2], синхронизированная с самим рядом:

$$X_i = A_1 f_1(i\Delta) + A_2 f_2(i\Delta) + \dots + A_L f_L(i\Delta),$$

где $A_1,\dots,A_L$ – коэффициенты регрессии; $f_1(t),\dots,f_L(t)$ – заданные функции модели. Исходя из требований статистической устойчивости признаков декомпозиции, в качестве условий можно выбрать условия попадания в заданные интервалы не самих значений ряда, а именно значений усредненной модели. Такие условия в дальнейшем будут называться **условиями декомпозиции**. Пусть $s_a, a=0,\dots,K,$ – заранее заданная совокупность границ интервалов $[s_{a-1}, s_a]$, $a=1,\dots,K$, $s_K = \max(x_i, i=1,...,N)$, попадание в которые значений усредненного процесса X_i соответствует наступлению события q_a. Это означает, что при выполнении условия

$$s_{a-1} \le X_i < s_a \quad (2)$$

значение x_i исходного ряда относится к эмпирическому распределению $\rho_a(x \mid q_a)$ и учитывается в соответствующей ему гистограмме. В результате исходная гистограмма распадается на K отдельных гистограмм, соответствующих условиям (2). Каждая из гистограмм описывает случайные флуктуации в системе при попадании ее в некоторую область усредненных значений, наблюдаемых ее параметров.

Другим способом определения условий декомпозиции является использование скользящих рядов наблюдений и их статистических параметров. Под скользящим рядом понимается разбиение исходного ряда длиной N на множество отрезков одинаковой длины $L < N$, начало и конец которых смещаются друг относительно друга последовательно на одну и ту же величину T, которую будем называть сдвигом. В результате такой процедуры в обработку попадают отрезки ряда $\{Y_i^{(a)}\}$, $i=1,\dots,L; a=1,\dots,M$ ($M = [(N-L)/T]+1$), сформированные следующим образом:

$$Y_i^{(a)} = X_{i+(a-1)T}, i=1,\dots,L; a=1,\dots,M. \quad (3)$$

Здесь $[x]$ – целая часть вещественного числа x. Отрезки данных перекрываются и содержат основную информацию о процессе на отрезке. При использовании скользящих рядов в качестве текущих признаков можно использовать значения различного рода статистик самих отрезков рядов, например, их средние значения, дисперсии, коэффициенты корреляций и т.д. В дальнейшем в качестве примера будет проведена декомпозиция ряда чисел Вольфа с помощью дисперсии скользящих рядов:

$$S_a = \sum_{i=1}^{L} (Y_i^{(a)} - \overline{Y^{(a)}})^2 / (L-1), \; a=1,\dots,M. \quad (4)$$

Из самого принципа декомпозиции вытекает, во-первых, что исходная полная гистограмма распределения, представленная в виде (1), имеет смысл полной вероятности для полной совокупности признаков q_a. Во-вторых, априорные вероятности могут быть вычислены по формулам:

$$p_a = \frac{N_a}{N},$$

где N_a – число событий, соответствующих появлению признака q_a, а N – общее число событий (длина ряда): $N = \sum_{a=1}^{K} N_a$. Такой способ декомпозиции является точным, в отличие от вычисления параметров q_a смеси, исходя из заданной формы теоретических распределений $\rho_a(x \mid q_a)$. В этом случае, если есть необходимость строить теоретическую декомпозицию, то ее можно проводить независимо для каждого отдельного эмпирического распределения $\rho_a(x \mid q_a)$, построенного в результате условной декомпозиции. При этом необходимо применить стандартные методы проверки гипотез о соответствии отдельных эмпирических распределений заданным теоретическим.

Недостатком такого способа декомпозиции является возможность строить разбиение, основываясь лишь на сравнительно простых признаках достижения определенных граничных значений самими рассматриваемыми переменными. В реальности отдельные компоненты в распределениях могут порождаться разными по сути физическими процессами, различие между которыми нельзя установить, изучая только один вид процесса и, тем более, по пересечениям им граничных значений. Однако предлагаемый подход может дать очень полезную информацию в тех случаях, когда диапазоны изменений являются достаточным признаком явления. Примером таких явлений могут служить атмосферные явления, связанные с прохождением областей повышенного давления – антициклоны, и пониженного – циклоны. В этом случае в более общем случае, когда в качестве ряда наблюдений используются несколько отдельных рядов наблюдения различных параметров среды или системы, в качестве признака декомпозиции могут быть использованы более сложные числовые функции, зависящие от значений нескольких параметров, взятых в различные моменты времени.

2. Декомпозиция ряда чисел Вольфа

В качестве первого примера рассмотрим задачу декомпозиции ряда чисел Вольфа солнечной активности. Это позволит сравнить результаты декомпозиции с уже имеющимися результатами теоретической декомпозиции, которые были получены в работах [3, 4]. Исходный ряд ежемесячных значений чисел Вольфа за период с января 1749 по май 2014 г. приведен на рис. 2.

На рис. 3, кроме самой гистограммы, приведено показательное распределение $\rho_W(w) = \lambda e^{-\lambda w}$ со средним значением $\overline{W} = 1/\lambda$, равным среднему значению ряда чисел Вольфа. Видно, что хотя теоретическое распределение по форме хорошо согласуется с эмпирическим, тем не менее имеются отклонения в области значений $w = 40 - 50$. Это и явилось причиной проведения декомпозиции данного распределения в работах [3, 4]. Основная идея декомпозиции состояла в том, что кроме «равновесного» механизма генерации пятен, который определяет показательное распределение числа пятен на Солнце, существует еще и «взрывной» механизм, который и приводит к появлению отклонений. Гистограмма распределения чисел Вольфа исходного ряда представлена на рис. 3.

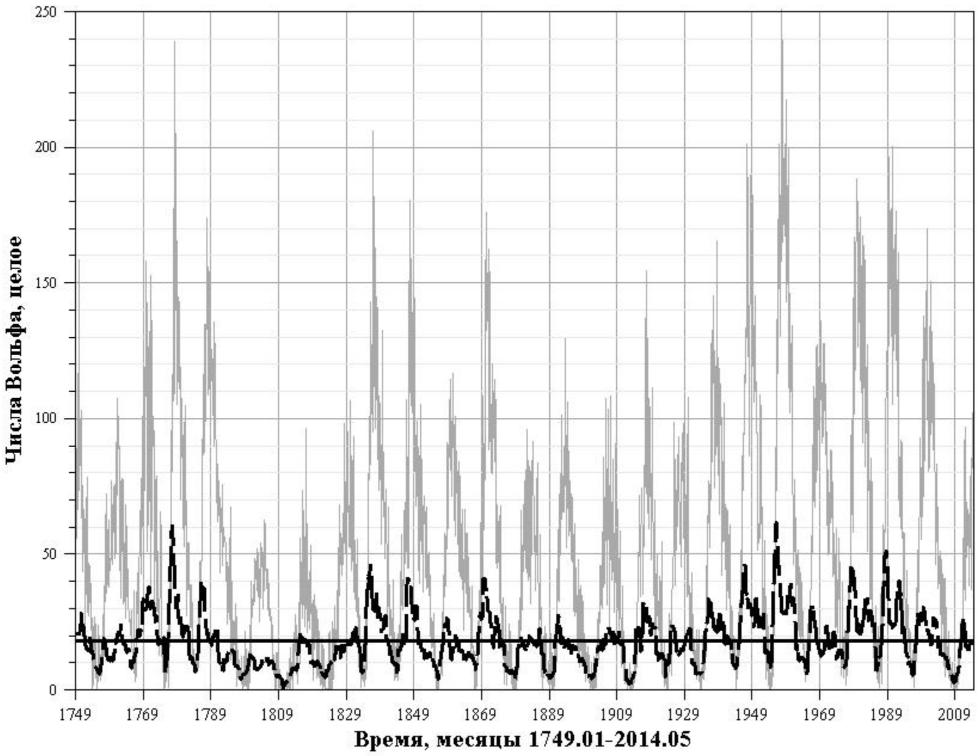


Рис. 2. Ряд чисел Вольфа (серая сплошная кривая) и значения выборочных значений стандартных отклонений скользящих рядов $L = 24, T = 1$ (жирная пунктирная кривая) и среднее значение ряда стандартных отклонений за весь период (сплошная черная горизонтальная линия)

В работах [3, 4] предполагалось выявить характеристики взрывного механизма с помощью декомпозиции распределения путем разделения на показательное и дополнительное распределения типа распределения Максвелла с нулевой вероятностью при $w = 0$. Задача решалась с помощью метода моментов. В данной работе проведем декомпозицию с помощью разделения значений на основании сравнения стандартных отклонений скользящих рядов длиной 24 месяца со средним по всему ряду стандартным отклонением. Дисперсия скользящего ряда может служить индикатором величины энергии возмущений процесса на данном отрезке времени. Превышение некоторого уровня можно идентифицировать как признак появления «взрывного» процесса образования пятен. На рис. 2 вместе с самим рядом чисел Вольфа жирной пунктирной линией приведен ряд стандартных отклонений для каждого скользящего ряда длиной 24 месяца со сдвигом 1 месяц. Именно этот ряд использовался для идентификации признака проявления механизма «взрывного» образования пятен. В качестве граничного значения выбиралось среднее значение стандартного отклонения по всему ряду. При выбранных параметрах среднее значение стандартного отклонения было равно $\overline{W} \simeq 17$. Этот уровень на рис. 2 представлен сплошной черной горизонтальной линией. В результате все значения исходного ряда со значениями $x_i = w_i > \overline{W}$ относились к действию механизма взрывного типа, а со значениями $x_i = w_i \le \overline{W}$ к равновесному состоянию. Соответствующие гистограммы представлены на рис. 4. В качестве теоретических распределений для сравнения с эмпирическими использовались следующие распределения:

$$\rho_1(w) = \frac{1}{s_1} e^{-w/s_1}, \tag{5}$$

$$\rho_2(w) = \frac{4}{3\sqrt{\pi} s_2^{5/2}} w^{3/2} e^{-w/s_2}. \tag{6}$$

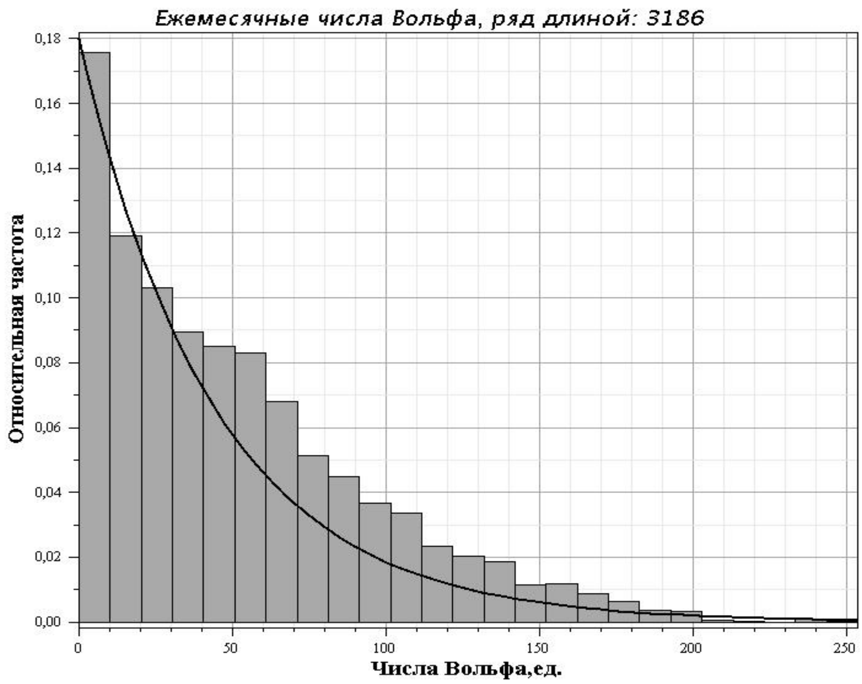
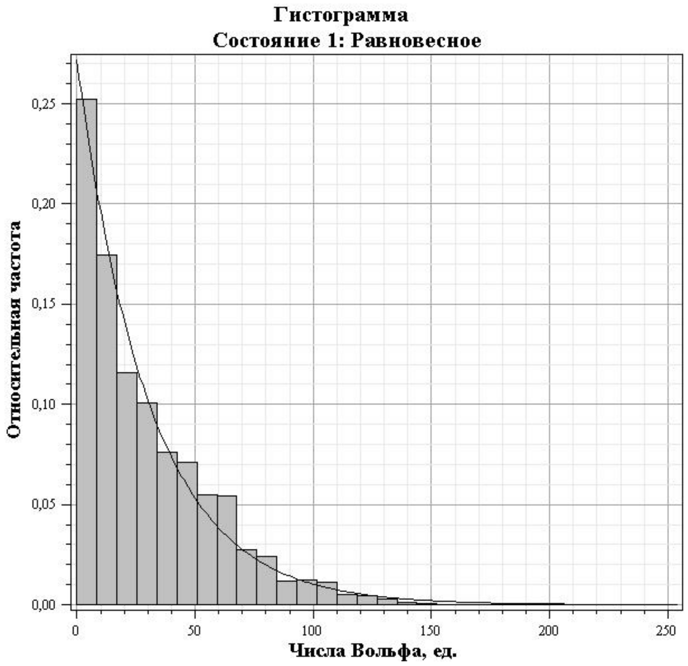


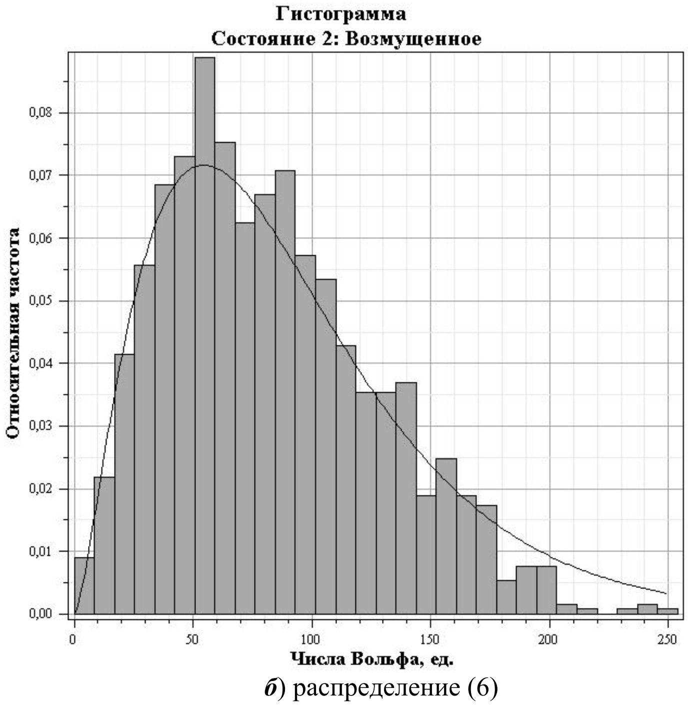
Рис. 3. Гистограмма распределения чисел Вольфа (серая) и теоретическое показательное распределение с $\lambda = 1/\overline{W}$ (сплошная черная)

Выбор распределения $\rho_2(w)$ производился среди большого класса распределений двух видов: 1) $\rho(w) = Z_1 w^k e^{-\alpha w^2}$, 2) $\rho(w) = Z_2 w^k e^{-\alpha w}$, на основе критерия χ^2. Выбор (6) является наилучшим при вычислении параметра s_2

на основе среднего значения эмпирического распределения на рис. 4,*б*: $s_2 = 2\overline{W}_2 / 5$. Выбор распределения (5) производился на основе вычисления среднего по формуле $s_1 = \overline{W}_1$



а) распределение (5)



б) распределение (6)

Рис. 4. Гистограммы компонент распределения чисел Вольфа (серая) и теоретические распределения (сплошная черная)

По аналогии с работами [3, 4] приведем анализ изменения параметров компонент распределений, полученных описанным способом для скользящих рядов длиной 100 лет со сдвигом 1 год. Для каждого отрезка длиной 100 лет производилась декомпозиция распределения на основе признака превышения стандартным отклонением скользящего ряда среднего значения стандартных отклонений. Результаты представлены на рис. 5 и 6.

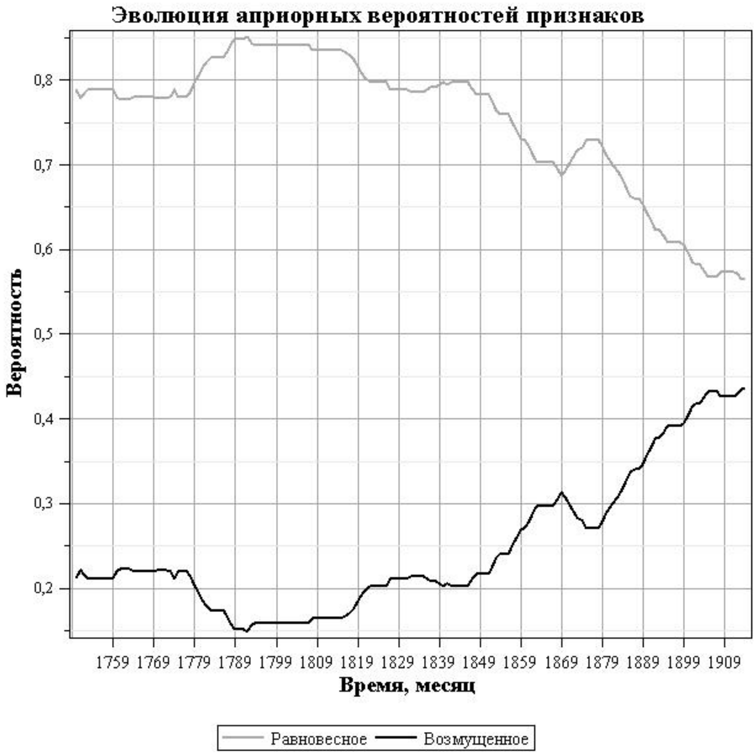


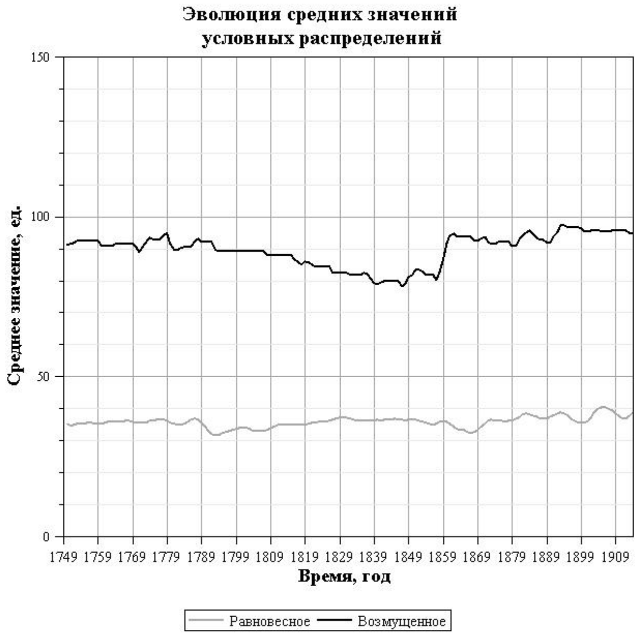
Рис. 5. Изменение априорных вероятностей компонент распределения чисел Вольфа скользящих рядов длиной 100 лет с шагом 1 год

Основные выводы анализа результатов в основном совпадают с тем, что было сделано в [3, 4]. В частности, обнаруживаются скачки свойств априорных вероятностей, средних и дисперсий в те же моменты времени, что и при обработке, выполненной в работах [3, 4]. Это говорит об устойчивости найденных закономерностей. Однако основным достижением данного подхода является явное указание признака, в результате которого получена декомпозиция. Для выбора конкретных типов распределений для их теоретического анализа появляются реальные данные в форме самих гистограмм компонент, которые получены в явном виде. Этот результат дает основания для получения более глубоких выводов относительно характера эволюции солнечной активности.

3. Описание модели ряда атмосферного давления

В качестве примера использования моделей регрессии для декомпозиции эмпирических распределений рассмотрим условную декомпозицию распределения значений давления на одной из метеостанций Ульяновской области за 2009 г.

а)



б)

Рис. 6. Изменение средних значений (***а***) и стандартных отклонений (***б***) компонент распределения чисел Вольфа скользящих рядов длиной 100 лет с шагом 1 год

На рис. 7 приведен график изменения давления за 2009 г. (сплошная серая кривая) вместе с моделью, построенной с помощью МНК (сплошная черная кривая). Кроме этого, на графике проведены линии, соответствующие границам областей циклонической деятельности (нижняя черная штриховая), антициклонической (антициклонической) (верхняя черная штриховая) и среднее значение за весь период (сплошная черная). Условные границы были выбраны в соответствии со следующими соотношениями: $S_1 = \overline{P} - 0{,}26\sigma_p$, $S_2 = \overline{P} + 0{,}26\sigma_p$, где $\overline{P}$ – оценка среднего, σ – оценка стандартного отклонения. Такой выбор позволяет заранее оценить вероятность попадания отдельного измерения в интервал $[S_1, S_2]$ в предположении о нормальности распределения.

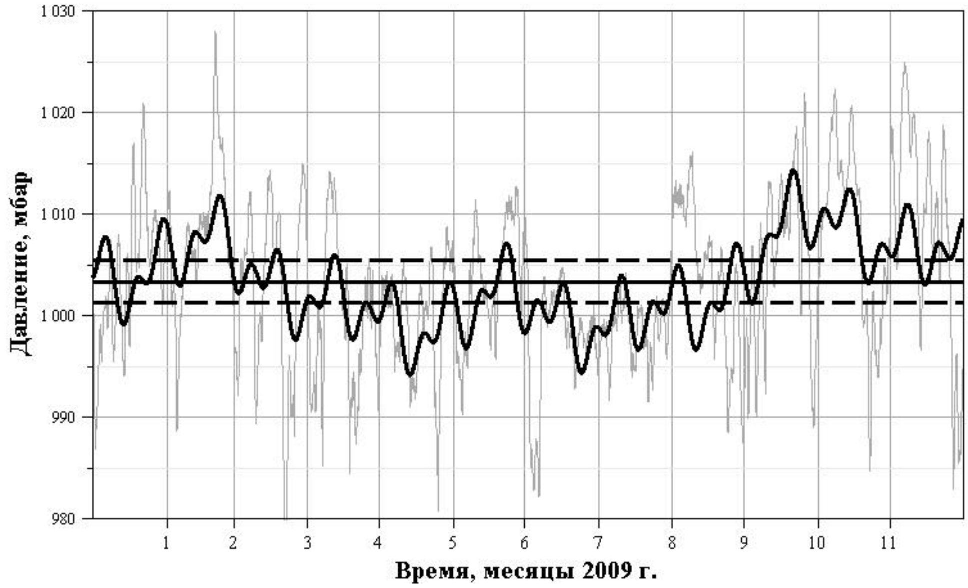


Рис. 7. Изменение давления за 2009 г. (серая сплошная кривая), модель МНК (черная сплошная кривая), границы областей (черные штриховые) и среднего значения за весь период (сплошная черная)

На рис. 8 приведена гистограмма распределения давления за весь 2009 г. и нормальное распределение, имеющее среднее и дисперсию, равные среднему и дисперсии самого ряда (сплошная кривая).

Модель метода наименьших квадратов, приведенная на рис. 7, строилась как совокупность их шести основных гармоник, периоды которых находились с помощью предварительного спектрального анализа ряда наблюдений. Спектральная плотность оценивалась с помощью метода максимальной энтропии в сочетании с методом векторизации исходного ряда [5]. Предварительно проводилось сглаживание ряда косинусным фильтром Тьюки с окном $P = 12$ и прореживанием с шагом $D = 12$. Спектральная плотность в диапазоне периодов от 2 суток и до 40 суток приведена на рис. 9. На рис. 9 видна характерная гармоника с периодом $T = 33$ сут. и кратные ей гармоники. По всей видимости, эти гармоники связаны с лунным приливом и его обертонами. Более низкочастотные гармоники, в частности годовой период и его обертоны, на этом графике не видны, но могут быть обнаружены с помощью сглаживания и прореживания ряда с большим шагом.

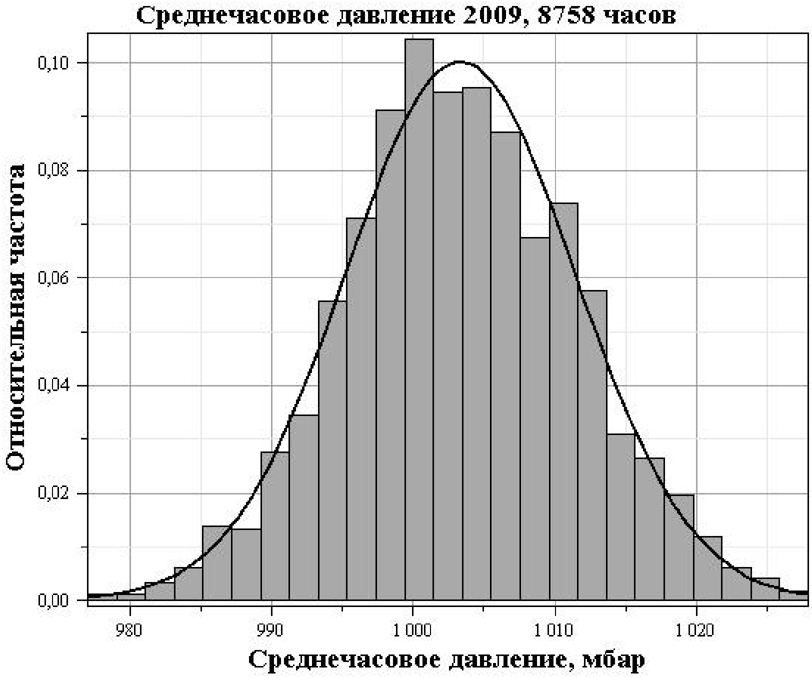


Рис. 8. Гистограмма распределения давления за весь 2009 г.
Сплошная кривая – нормальное распределение со средним и дисперсией, равными их оценкам по всему ряду

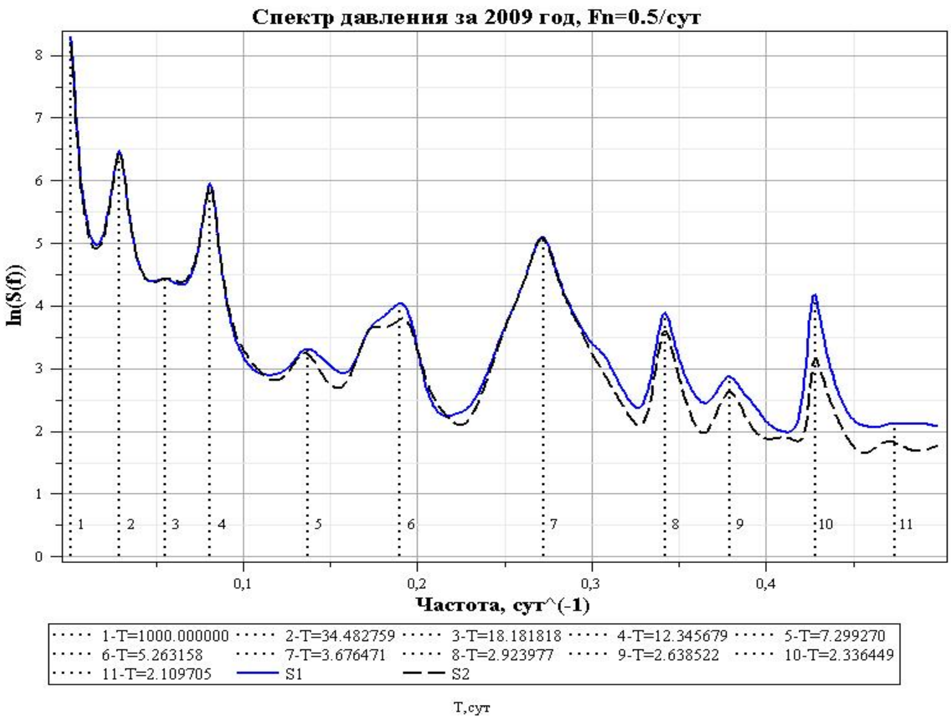


Рис. 9. График спектральной плотности ряда давления за весь 2009 г.
Частота Найквиста 0,5 сут$^{-1}$. Периоды пиков приведены в сутках

В соответствии со спектром модель МНК строилась в виде

$$P(t)=\sum_{k=0}^{13} A_k F_k(t),$$

где $F_0=1$, $F_{2k-1}(t)=\cos(\Omega_k t)$, $F_{2k}(t)=\sin(\Omega_k t)$, $k=1,\ldots,6$. Частоты выбраны следующим образом: $\Omega_1=\pi/3800$, $\Omega_2=\pi/2500$, $\Omega_3=\pi/1488$, $\Omega_4=\pi/744$, $\Omega_5=\pi/288$, $\Omega_6=\pi/144$.

4. Условная декомпозиция распределений вероятностей атмосферного давления

Пример декомпозиции рассмотрим для отрезка ряда длиной в 4 месяца, гистограмма которого представлена на рис. 10, из которого видно, что гистограмма имеет сложный характер. Декомпозиционные границы выбраны аналогично рис. 7: $S_1=\overline{P}-0,26\sigma_p$, $S_2=\overline{P}+0,26\sigma_p$. Результаты декомпозиции представлены на рис. 11.

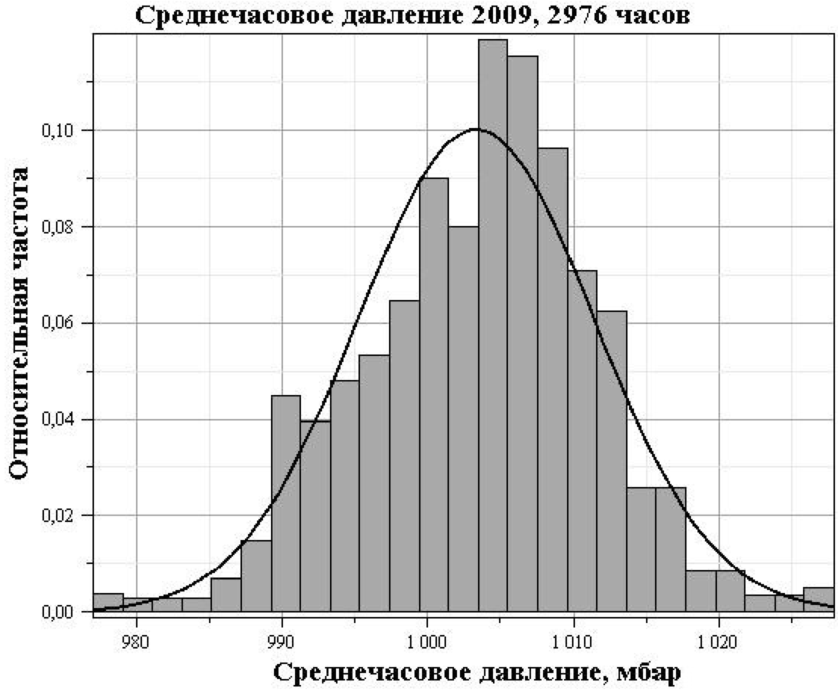


Рис. 10. Распределение давления за первые 4 месяца 2009 г. Сплошной линией нанесено нормальное распределение, имеющее среднее и дисперсию, равные оценкам этих величин по исходному ряду

Представленные рисунки демонстрируют характер распределений в каждом из диапазонов усредненного давления, которые на рисунках условно обозначены так: «циклон» – $p<S_1$, «норма» – $S_1\le p\le S_2$, «антициклон» $p>S_2$. Соответствующие значения априорных вероятностей равны $p_1=0,18$, $p_2=0,56$, $p_3=0,26$.

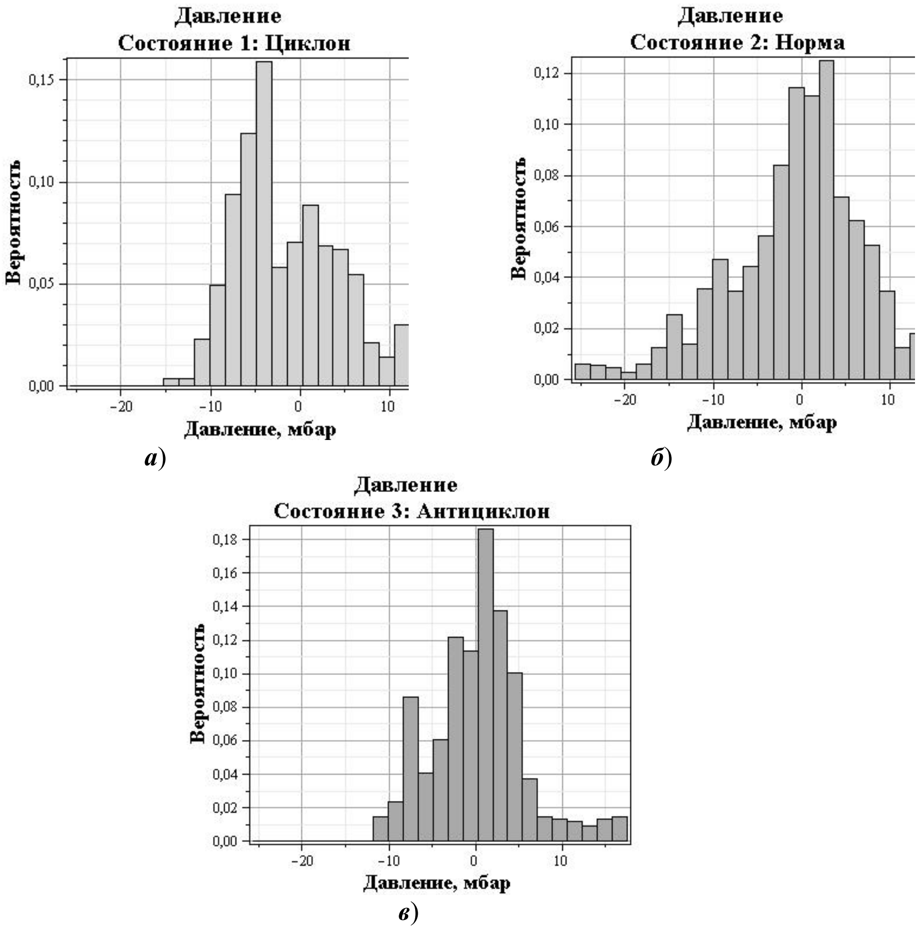


Рис. 11. Нулевое значение соответствует среднему давлению за период циклонической деятельности

5. Условная декомпозиция скользящих рядов давления

Как и для ряда чисел Вольфа, метод условной декомпозиции может быть использован для выяснения динамики изменчивости параметров компонент распределения, полученных в результате декомпозиции. В данной работе для иллюстрации возможностей такого подхода проводилась декомпозиция последовательности скользящих рядов давления за 2009 г. длительностью 4 месяца, взятых со сдвигом 5 сут. Декомпозиция проводилась при одних и тех же границах для всего периода наблюдения (12 месяцев 2009 г.). После декомпозиции вычислялись дисперсии компонент и средние, которые вычислялись как среднее модели и среднее соответствующей компоненты распределения. На рис. 12–14 представлены результаты такой декомпозиции в виде графиков изменения априорных вероятностей, средних и стандартных отклонений компонент.

Анализируя полученные графики на рис. 12–14 и поведение найденных параметров компонент смеси распределений, можно обнаружить интересную информацию. Из графиков априорных вероятностей можно сделать вывод, что циклоническая и антициклоническая деятельности сменяют друг друга в течение года, находясь в противофазе друг к другу.

Рис. 12. Изменение априорных вероятностей за 2009 г.
(длина ряда уменьшена на длину скользящего ряда)

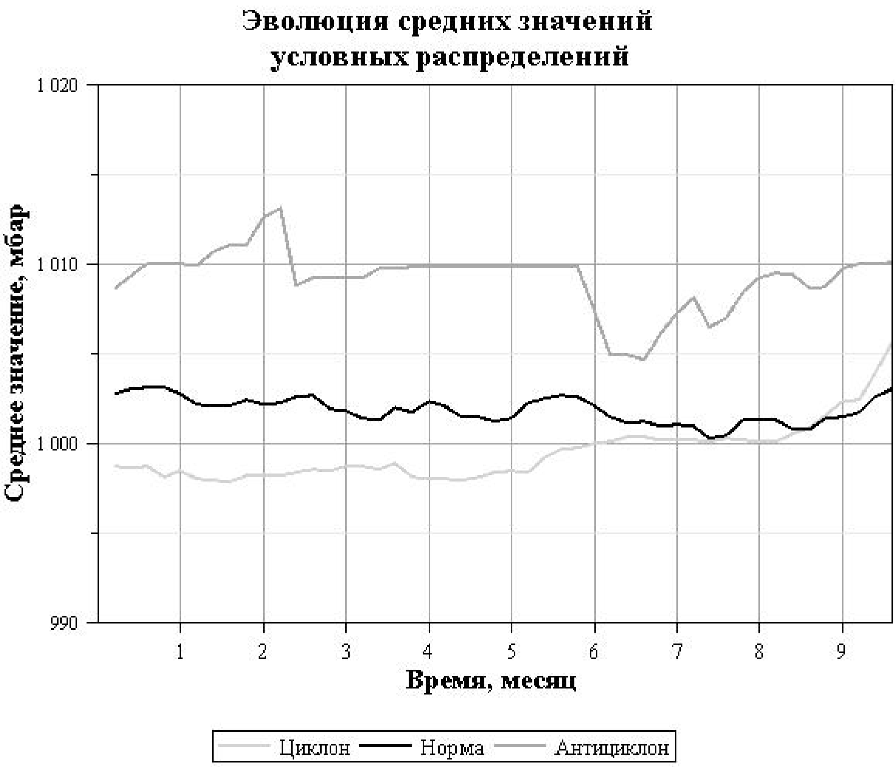


Рис. 13. Изменение средних значений компонент за 2009 г.
(длина ряда уменьшена на длину скользящего ряда)

Рис. 14. Изменение стандартных отклонений компонент за 2009 год
(длина ряда уменьшена на длину скользящего ряда)

Летом превалирует циклоническая деятельность, а зимой – антициклоническая. Из графика стандартных отклонений также можно сделать вывод, что в летний период погода стабилизируется, поскольку величины стандартных отклонений оказываются меньше, чем в остальные периоды времени.

Заключение

В работе предложен и апробирован метод условной декомпозиции на примере рядов ежемесячных чисел Вольфа за период с 1749 по 2014 г., а также ряда ежечасовых значений атмосферного давления за 2009 г. Сам метод реализуется с помощью простых вычислительных процедур, что позволяет его эффективно применять к решению различных задач обработки данных. В работе были продемонстрированы два подхода к формированию признаков условной декомпозиции на основе ряда стандартных отклонений скользящих рядов чисел Вольфа и модели регрессии для ряда давления. Оба подхода демонстрируют возможность получать дополнительную информацию о характере процесса на основе метода декомпозиции. Это позволяет получить новые физические параметры состояния систем в форме параметров компонент смеси эмпирических распределений. Эта дополнительная информация может быть использована для более глубокого выяснения механизмов физических процессов, порождающих те или иные явления в исследуемых системах. Хотя в данной работе использовался лишь простейший вариант условий декомпозиции, данный метод может применяться и для более общих условий, накладываемых на изучаемые процессы.

Авторы выражают искреннюю благодарность метеорологическому центру аэропорта Восточный г. Ульяновска за предоставленную информацию.

Список литературы

1. **Гмурман, В. Е.** Теория вероятностей и математическая статистика / В. Е. Гмурман. – М. : Высшая школа, 2003. – 480 с.
2. **Вентцель, Е. С.** Теория вероятностей / Е. С. Вентцель. – М. : Наука, Физматгиз, 1969. – 576 с.
3. **Журавлев, В. М.** Анализ долговременной эволюции активности Солнца на основе ряда чисел Вольфа (I. Методика) / В. М. Журавлев, С. В. Летуновский // Известия высших учебных заведений. Поволжский регион. Физико-математические науки. – 2010. – № 4 (16). – С. 120–130.
4. **Журавлев, В. М.** Анализ долговременной эволюции активности Солнца на основе ряда чисел Вольфа (II. Результаты) / В. М. Журавлев, С. В. Летуновский // Известия высших учебных заведений. Поволжский регион. Физико-математические науки. – 2011. – № 3 (19). – С. 164–174.
5. Свидетельство государсвенной регистрации программ для ЭВМ № 2012619378. Многомерный метод максимальной энтропии в одномерном спектральном анализе / Журавлев В. М., Валентюк Р. А. Деп. в ВИНИТИ. 09.09.1987. № 6602-В87. – 17 октября 2012 г.

References

1. Gmurman V. E. *Teoriya veroyatnostey i matematicheskaya statistika* [Probability theory and mathematical statistics]. Moscow: Vysshaya shkola, 2003, 480 p.
2. Venttsel' E. S. *Teoriya veroyatnostey* [Probability theory]. Moscow: Nauka, Fizmatgiz, 1969, 576 p.
3. Zhuravlev V. M., Letunovskiy S. V. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* [University proceedings. Volga region. Physical and mathematical sciences]. 2010, no. 4 (16), pp. 120–130.
4. Zhuravlev V. M., Letunovskiy S. V. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Fiziko-matematicheskie nauki* [University proceedings. Volga region. Physical and mathematical sciences]. 2011, no. 3 (19), pp. 164–174.
5. *Svidetel'stvo gosudarsvennoy registratsii programm dlya EVM № 2012619378. Mnogomernyy metod maksimal'noy entropii v odnomernom spektral'nom analize* [Certificate of state registration of the PC programs № 2012619378. Multidimensional method of maximal entropy in one-dimensional spectral analysis]. Zhuravlev V. M., Valentyuk R. A. Dep. in VINITI. 9 September 1987. No. 6602-B87. 17 October 2012.

Журавлев Виктор Михайлович
доктор физико-математических наук, профессор, кафедра теоретической физики, Ульяновский государственный университет (Россия, г. Ульяновск, ул. Льва Толстого, 42)

Zhuravlev Viktor Mikhaylovich
Doctor of physical and mathematical sciences, sub-department of theoretical physics, Ulyanovsk State University (42 Lva Tolstogo street, Ulyanovsk, Russia)

E-mail: zhvictorm@gmail.ru

Морозов Виталий Михайлович
студент, Ульяновский государственный университет (Россия, г. Ульяновск, ул. Льва Толстого, 42)

E-mail: aieler@rambler.ru

Morozov Vitaliy Mikhaylovich
Student, Ulyanovsk State University (42 Lva Tolstogo street, Ulyanovsk, Russia)

Петряков Михаил Сергеевич
аспирант, Ульяновский государственный университет (Россия, г. Ульяновск, ул. Льва Толстого, 42)

E-mail: petraykovms@gmail.com

Petryakov Mikhail Sergeevich
Postgraduate student, Ulyanovsk State University (42 Lva Tolstogo street, Ulyanovsk, Russia)

Самойлов Вадим Владимирович
доцент, кафедра теоретической физики, Ульяновский государственный университет (Россия, г. Ульяновск, ул. Льва Толстого, 42)

E-mail: samoilov_vadim@mail.ru

Samoylov Vadim Vladimirovich
Associate professor, sub-department of theoretical physics, Ulyanovsk State University (42 Lva Tolstogo street, Ulyanovsk, Russia)

УДК 519.25, 53.023,52-17

Журавлев, В. М.

Метод условной декомпозиции эмпирических распределений и его применение к задаче анализа рядов наблюдений / В. М. Журавлев, В. М. Морозов, М. С. Петряков, В. В. Самойлов // Известия высших учебных заведений. Поволжский регион. Физико-математические науки. – 2014. – № 3 (31). – С. 179–197.