

Методология классификации временных рядов по наличию тренда в наукометрии

П. Ю. Блинов*

Российский научно-исследовательский институт экономики, политики и права
в научно-технической сфере, 127254, Москва, Россия

*Контактный автор: Блинов Павел Юрьевич, e-mail: p.blinov@riep.ru

Поступила 18 марта 2021 г., доработана 16 сентября 2021 г., принята в печать 28 декабря 2021 г.

Исследователи часто сталкиваются с задачами обработки большого числа данных, и рост интереса к различным количественным метрикам в разных областях вполне объясним. В частности, в наукометрии существует ряд метрик, характеризующих публикационную активность, которая имеет ярко выраженную тенденцию к росту, в особенности в РФ. Для выявления тех или иных причин роста эффективным инструментом является аппарат математической статистики. В данной статье демонстрируется полезность применения статистических критериев случайности или отсутствия тренда на временных рядах для выявления трендов публикационной активности в области наукометрии. Предложена методология совокупной проверки несколькими критериями на наличие или отсутствие тренда, которая является весьма гибким инструментом. Методология ориентирована на повышение точности результата оценки для временных рядов небольшого объема и позволяет при проверке большого количества временных рядов группировать их по типу или наличию тренда.

Ключевые слова: наукометрия, публикационная активность, проверка статистических гипотез, критерии случайности или отсутствия трендов, тренд, уровень значимости.

Цитирование: Блинов П.Ю. Методология классификации временных рядов по наличию тренда в наукометрии. Вычислительные технологии. 2022; 27(1):52–69. DOI:10.25743/ICT.2022.27.1.005.

Введение

Критерии случайности или отсутствия тренда представляют собой статистический инструмент для выявления случайности некоторых колебаний во временном ряду, отсутствия неслучайной закономерности или ярко выраженного тренда.

Временные ряды могут отражать наличие тренда, наличие систематической составляющей, сезонной составляющей, а также случайной составляющей (шума). При этом под трендом понимается основная тенденция изменения временного ряда, направление преимущественного движения исследуемой переменной. Для выявления тенденций в поведении временных рядов необходимо выбрать критерии, которые хорошо проверяют гипотезу об отсутствии тренда в математическом ожидании. Также такие критерии

должны быть устойчивы при нарушении предположения о том, что анализируемая выборка нормально распределена [1].

Применение критериев случайности (в англоязычной литературе — Trend test) широко распространено там, где есть необходимость в изучении временных рядов. С анализом временных рядов сталкиваются при обработке измерений в технических областях, в экономике, сельском хозяйстве, метеорологии [2].

Много научных публикаций в области экологии используют упомянутые критерии. Так, например, в статье [3] критерии случайности используются для изучения изменений климата на примере одного региона Индии. В техническом отчете [4] описывается множество статистических процедур, нацеленных на исследование и мониторинг качества воды в природных водоразделах, на примере реки Миссисипи. Глубокое исследование эффективности китайского проекта по улучшению качества воды, в котором находят применение соответствующие статистические критерии, представлено в публикации [5]. Критерии находят применение в исследовании изменения температур воды и воздуха в Японии за 50 лет [6]. В статье [7] одним из критериев случайности оценивают рост количества пожаров сахарного тростника и других культур в Бразилии, вызванный изменением климата. Применение критериев случайности для изучения колебаний концентрации озона в атмосфере городов и сел представлено в работе [8].

В российской метеорологии также можно найти сведения о применении статистических критериев случайности. К примеру, статья [9], посвященная оценке роста количества событий очень сильных осадков в пределах Европейской части России, опирается на результаты применения критерия Манна–Кендалла. Схожий подход в статье [10] применяется для поиска трендов в популяции земноводных.

В наукометрии критерии случайности используются довольно редко. Большая часть таких работ направлена на оценку перспективных направлений науки, роста или падения интереса к некоторым отраслям. Например, в работе [11] оценивается рост научного интереса к области изменения климата, а также семантический анализ ключевых слов по данной тематике. Полученные, в частности, с помощью критерия Манна–Кендалла результаты исследования демонстрируют спад и рост употребления конкретных терминов. Стоит также отметить статью [12], в которой статистическими методами определены “горячие” и “холодные” темы в информационной науке и наукометрии.

Критерии случайности могут найти значительно больше применений в области наукометрии. Данные методы можно использовать для исследования роста публикационной активности исследователей РФ, оценки результативности научных организаций в динамике. Благодаря проверке гипотезы случайности можно из большого числа временных рядов выделить нетипичные случаи, например отсутствие роста или падение результативности организаций при увеличении количества публикуемых авторов. Рассмотрение отдельных групп организаций, авторов по отдельности позволит заметить более нетривиальные закономерности или построить новые гипотезы, касающиеся исследуемых показателей.

Несмотря на то что анализ ограничивается только проверкой наличия тренда в среднем, выбрать максимально эффективный критерий весьма непросто. Поскольку аппарат математической статистики довольно часто сталкивается с проблемами верификации случайности и неслучайности данных в некоторой выборке (временном ряду), существует множество статистик [13]. Даже разрабатываются новые критерии случайности или отсутствия тренда [14–16].

Значительное число критериев, ориентированных на проверку гипотезы об отсутствии тренда, представлено в [17]. Данное справочное пособие, содержащее большой спектр различных видов статистических критериев, ориентировано на инженеров, и в нем приводятся простые примеры, поясняющие применение этих критериев. Однако данная работа не ставит задачей проведение сравнительного анализа в отдельных группах и не отвечает на вопрос о правильно выбранном критерии.

Задачей анализа критериев проверки гипотез об отсутствии тренда, а также сравнительным анализом мощности (статистический показатель эффективности критерия, вероятность отклонить неверную альтернативную гипотезу) критериев по отношению к ряду альтернатив занимались Б.Ю. Лемешко, А.А. Комиссарова, А.Е. Щеглов [1, 13], И.В. Веретельникова [18]. В диссертационной работе И.В. Веретельниковой [18] приводится солидный список критериев с углубленным исследованием каждого критерия и сравнительным анализом мощностей, которые осуществлялись с применением компьютерного моделирования распределений случайных величин. Зарубежных статей, посвященных сравнительному анализу мощностей критериев, довольно много. Эти работы в большинстве своем ограничены несколькими критериями и ориентированы на сравнение в конкретной тематике. Наиболее цитируются статьи [19, 20].

В разд. 1 приводится краткое описание критериев, ориентированных на проверку отсутствия/наличия тренда во временных рядах. Во втором разделе дается формальное описание методологии совокупной проверки, которая состоит из нескольких этапов и повышает точность оценки. В разд. 3 демонстрируется применение критериев и совокупных проверок на практических данных публикационной активности нескольких научных организаций. Четвертый раздел содержит анализ публикационной активности научных организаций РФ.

1. Данные и методы

1.1. Данные

Для исследования наличия случайности или тренда по научным организациям собраны данные о публикационной активности авторов российских организаций из SCOPUS за период с 2000 по 2019 г. Выборка организаций ограничена, рассмотрены только те, которые имеют хотя бы одну публикацию в SCOPUS за 2017–2019 гг.

Наличие тренда искалось по трем разным, но связанным мерам. Первая мера — количество публикаций целым счетом. Под данной мерой понимается количество публикаций в определенном году для организаций без учета количества авторов или организаций, которые аффилированы в публикациях. Вторая мера — количество публикаций организаций фракционным счетом. Фракционный счет отличается от первой меры учетом количества авторов и аффилиаций. Более подробно ознакомиться с оценкой результативности данных индикаторов можно в статье [21]. Третья мера представляет собой некоторый аналог комплексного балла публикационной результативности (КБПР) [22, 23]. Исследования о применимости КБПР в отдельных сферах наук весьма актуальны [24, 25]. Эта мера отличается от второй еще учетом весового коэффициента разных публикаций по типу квартиля журнала (по JCR за актуальный год) или по типу публикации. Изменение квартиля журнала по годам учитывается.

Тип журнала	Без квартиля	Q4	Q3	Q2	Q1
Весовой коэффициент	1	2.5	5	10	20

Первую и вторую меру обозначим аналогично тому, как это сделано в статье [21], — AC (Article Count) и FC (Fraction Count) соответственно. Третью меру обозначим FP (Fraction Point).

Безусловно, данные метрики не могут в полной мере корректно оценивать результаты по авторам и научным организациям, к ним возникает множество вопросов. Тот же КБПР подвергается конструктивной критике [26, 27] и дорабатывается. Но все же исследование этих индикаторов в динамике может быть информативным. Исследование временных рядов может помочь определить тренд, доказать или опровергнуть наличие значимой тенденции, проверить, являются флуктуации случайными или нет, выделить аномальные показатели.

1.2. Используемые статистические критерии

Среди всего многообразия критериев проверки гипотезы отсутствия тренда в математическом ожидании выбраны три: критерий Кокса – Стюарта [28], критерий инверсий [18, 29] и критерий Манна – Кендалла [30, 31]. Первые два хорошо зарекомендовали себя в сравнительном анализе, проведенном в упомянутой ранее работе И.В. Веретельниковой [18], и являются наиболее мощными критериями. Критерий Манна – Кендалла широко применяется в зарубежных публикациях [3–8, 11, 12] и по статистике очень схож с критерием инверсий.

Стоит отметить, что критерий Манна – Кендалла зачастую применяется вместе с оценочной функцией Тейла – Сена [32, 33] для получения информации о скорости изменения величины и построения линейного приближенного тренда. В данной работе оценочная функция Тейла – Сена применяться не будет, однако ее использование вместе с тремя остальными статистиками оправдано и позволит упорядочить временные ряды с трендами по первой производной. Однако стоит учитывать, что данная статистическая процедура не чувствительна к наличию выбросов, поскольку вычисляет медианную скорость во временном ряду.

Общая постановка задачи для всех рассматриваемых критериев одинаковая. Пусть дан набор наблюдений X_1, X_2, \dots, X_n , который представляет собой временной ряд, упорядоченный по времени, n — количество наблюдений. Проверяем гипотезу H_0 — наличие случайности или отсутствие тренда против конкурирующей гипотезы H_1 — наличия тренда в математическом ожидании. Если некоторая статистика S_i , полученная по временному ряду X_1, X_2, \dots, X_n , попадает в определенный доверительный интервал, гипотеза об отсутствии тренда не отклоняется, иначе — отклоняется. Отклонение гипотезы означает, что во временном ряду, вероятно, присутствует некоторый тренд.

Доверительные интервалы для критериев в математической статистике бывают разными, они зависят от размера временного ряда. В нашем случае все три статистики обладают удобным преимуществом — все критерии двухсторонние и их доверительные интервалы в пределе не отличаются от нормального гауссова распределения. У критерия Манна – Кендалла и инверсионного критерия достаточно десяти наблюдений [18, 34] для применения доверительного интервала стандартного нормального закона. У критерия Кокса – Стюарта желательно применение хотя бы 15–20 наблюдений, но в данной работе он будет применяться и для временных рядов с десятью и более наблюдениями (большинство временных рядов будет состоять из двадцати наблюдений). В совокупной проверке эта погрешность будет несколько компенсирована. Доверительные интервалы

Т а б л и ц а 1. Доверительные интервалы для стандартного нормального распределения
Table 1. Confidence intervals for standard normal distribution

Уровень значимости, α	0.001	0.01	0.05	0.1
Доверительный интервал	$[-3.291, 3.291]$	$[-2.567, 2.567]$	$[-1.96, 1.96]$	$[-1.645, 1.645]$

Т а б л и ц а 2. Статистики и оценки математического ожидания и дисперсии критериев
Table 2. Test statistics and estimates of the mathematical expectation and variance of tests

Критерий	Ненормализованная статистика	Математическое ожидание	Дисперсия
Кокса – Стюарта	$S_1 = \sum_{i=1}^{\lfloor n/2 \rfloor} (n - 2i + 1)h_{i,n-i+1}$	$E[S_1] = \frac{n^2}{8}$	$D[S_1] = \frac{n(n^2 - 1)}{24}$
Инверсий	$S_2 = \sum_{i=1}^{n-1} \sum_{j=i}^n h_{i,j}^*$	$E[S_2] = \frac{n(n-1)}{4}$	$D[S_2] = \frac{2n^3 + 3n^2 - 5n}{72}$
Манна – Кендалла	$S_3 = \sum_{i=1}^{n-1} \sum_{j=i}^n \tilde{h}_{i,j}$	$E[S_3] = \text{sign}(S_3)$	$D[S_3] = \frac{n(n-1)(2n+5) - G}{18}$

для стандартного нормального распределения представлены в табл. 1. В данной работе все ограничено доверительным интервалом при уровне значимости α , равном 0.1.

Все статистики рассмотренных критериев имеют вид

$$S_i^* = \frac{S_i - E[S_i]}{\sqrt{D[S_i]}}, \quad (1)$$

где $E[S_i]$ и $D[S_i]$ — оценки математического ожидания и дисперсии, необходимые для вычисления нормализованной статистики, а S_i — ненормализованная статистика. В табл. 2 представлены формулы этих величин для каждого из критериев.

Во всех ненормализованных статистиках применяются h -функции, которые имеют различия для каждой статистики. Для критерия Кокса – Стюарта $h_{i,j} = 1$, если $X_i \leq X_j$ (при $i < j$), иначе 0. Этот критерий рассчитан так, что он сравнивает самое раннее и самое позднее наблюдения, присваивает результату сравнения некоторый вес, затем отбрасывает эти наблюдения и переходит к следующим.

Для критерия инверсий $h_{i,j}^* = 1$, если $X_i < X_j$ (при $i < j$), иначе 0. Отличается только наличием строгого неравенства. Это изменение введено только в данной работе, чтобы различия в критериях помогали классифицировать определенный редкий вид тренда во временных рядах при совокупной проверке. Этот критерий позволяет сравнить все наблюдения со всеми наблюдениями, которые произошли позднее.

Для критерия Манна – Кендалла:

$$\tilde{h}_{i,j} = \begin{cases} 1, & \text{если } X_i < X_j, \\ 0, & \text{если } X_i = X_j, \\ -1, & \text{если } X_i > X_j. \end{cases}$$

У данного критерия более сложные формулы дисперсии и математического ожидания. В математическом ожидании под $\text{sign}(S_3)$ понимается формула, которая принимает значения 1 или -1 , если значение S_3 положительно или отрицательно. Величина G является поправкой на число повторов наблюдений во временном ряду и вычисляется по формуле

$$G = \sum_{k=1}^g t_k(t_k - 1)(2t_k + 5),$$

где t_k — число повторов k -го наблюдения, g — число уникальных наблюдений с повторами во временном ряду. Данный критерий довольно эффективный, поправка на число повторов сильно выделяет его среди других. К сожалению, стоит отметить, что применение данного критерия на выборке, состоящей из одного уникального значения, приводит к тому, что $D[S_3] = 0$ и в формуле (1) происходит ошибка.

2. Совокупная проверка несколькими критериями

Несмотря на то что при проверке гипотезы отсутствия тренда все рассмотренные критерии являются высокомоощными, ошибки при проверке не исключены. Зачастую на практике при проверке статистических гипотез наборы данных проверяют несколькими критериями для повышения уверенности в правдивости результата проверки и минимизации вероятности ошибки. Четких правил и рекомендаций для применения совокупности критериев практически не существует, поэтому в данной работе использована совокупность трех критериев, которая будет состоять из трех последовательных этапов.

Первым этапом будет основная совокупная проверка с наименьшим риском ошибки в классификации. Если все критерии не отклоняют гипотезу об отсутствии тренда (значения всех статистик попадают в доверительный интервал при уровне значимости 0.1), то временной ряд классифицируется как бестрендовый. Если все критерии отклоняют гипотезу об отсутствии тренда (значения всех статистик не попадают в доверительный интервал), то во временном ряду, вероятно, есть некоторый тренд. При всех остальных исходах совокупной проверки временные ряды отмечаются как “неклассифицированные” и переходят на следующий этап.

Изначально критерии не классифицируют временные ряды по типу, однако их все-таки можно разделить на возрастающие и убывающие по значению статистик. Для приведенных в первой главе статистик, если все статистики критериев принимают значения больше значений доверительного интервала, то временной ряд классифицируется как возрастающий, если полученные значения меньше значений интервала, — как убывающий. Важное замечание — существуют и применяются критерии, у которых значение статистик характеризует временной ряд с точностью до наоборот. Ситуации, когда все значения статистик не попадают в доверительный интервал, но имеют разный знак, потенциально возможны. Однако изначально на модельных (моделирование проведено аналогично работе [18], общую методику моделирования можно найти в [35]) и на практических данных для текущих трех критериев не установлено ни одного временного ряда такого типа. Такие ситуации обнаружены только при совокупной проверке с критерием Фостера–Стюарта [18, 36]. Этот критерий, несмотря на то что довольно распространен в эконометрике, является малоэффективным [1, 13, 18]. Включение данного критерия в совокупную проверку приводило к значительному числу конфликтов между критериями, а это свидетельствует о том, что выбор критерия очень важен.

Поскольку с выбранной совокупностью критериев ни разу не возникала подобная ситуация, данной вероятностью решено пренебречь. Одна из рассмотренных статистик намеренно изменена, чтобы такие ситуации (значения статистик с разными знаками при непопадании в доверительный интервал) возникали и классифицировали еще один тип

тренда во временном ряду — постоянный. Первоначально критерий Кокса–Стюарта и критерий инверсий имели одинаковую h -функцию. После изменения в критерии инверсий $h_{i,j}^* = 1$, если $X_i < X_j$ вместо $X_i \leq X_j$. Таким образом, один из критериев совпадение наблюдений воспринимает как рост, другой — как падение оцениваемой меры. Вероятность возникновения такой ситуации также мала и проявляется в сильно дискретных данных. Тем не менее довольно полезно отделить данную ситуацию, чтобы временные ряды, состоящие преимущественно из одинаковых наблюдений, не классифицировались как растущие. Для потенциально же “пострадавших” временных рядов, которые не смогут пройти совокупную проверку после введения изменения в статистику критерия инверсий, классификация должна произойти на последующих этапах совокупной проверки.

Второй этап совокупной проверки ориентирован на классификацию временных рядов, которые близки к тому, чтобы быть классифицированными на первом этапе. На данном этапе проверку проходят временные ряды, не классифицированные на первом этапе и состоящие из не менее десяти наблюдений (классификация меньших временных рядов подобным способом рискованна). Если две статистики попали в доверительный интервал и среднее значение p -value $> \alpha$ ($\alpha = 0.1$ и соответствует доверительному интервалу в табл. 1), то тренда во временном ряду нет. Если две статистики не попали в доверительный интервал и среднее значение p -value $< \alpha$, то, вероятно, тренд есть, тип тренда определяется по знаку статистик. Все остальные временные ряды остаются “неклассифицированными” и переходят на последний третий этап совокупной проверки.

Третий этап совокупной проверки ориентирован на классификацию временных рядов с дискретными данными, число повторов в которых недостаточно велико, чтобы быть обнаруженными на первых двух этапах. Этот этап не отличается от первого, однако все временные ряды проходят предобработку. Для того чтобы избавиться от повторяющихся значений, используется метод экспоненциального сглаживания [37] с малым показателем коэффициента сглаживания. Каждый новый элемент сглаженного временного ряда вычисляется по формуле

$$X_i^* = \begin{cases} X_i & \text{для } i = 1, \\ pX_{i-1}^* + (1-p)X_i & \text{для } i > 1, \end{cases}$$

где $0 < p < 1$ — коэффициент сглаживания, который в нашем случае равен 0.01.

После третьего этапа все временные ряды делятся на пять типов: возрастающие, убывающие, ряды с постоянным трендом, бестрендовые и неклассифицированные. При необходимости можно объединить группы бестрендовых и временных рядов с постоянным трендом. Тем не менее во временных рядах с постоянным трендом могут быть обнаружены некоторые стратегии. К примеру, если автор публикует ровно одну статью в год на протяжении десяти лет и более — это маловероятно случайный исход.

Все неклассифицированные временные ряды после трех этапов совокупной проверки являются очень рискованными для однозначной классификации и отмечаются как “неопределенность”. Конечно, их можно классифицировать каким-то одним критерием, но это чревато частыми ошибками. При малом числе таких временных рядов их можно классифицировать “вручную” без автоматизации.

Как было упомянуто ранее, можно использовать оценочную функцию Тейла–Сена в качестве четвертой статистики для временных рядов с возрастающим и убывающим трендами. Данная формула позволит оценить примерную скорость роста/падения показателя за весь временной диапазон, чтобы проранжировать временные ряды и выделить

самые быстрорастущие. Также для проверки можно применить функцию Тейла – Сена на “неопределенных” временных рядах. Если среди таких рядов будет обнаружена относительно высокая скорость изменения показателя — временной ряд, возможно, был ошибочно забракован. Оценочную функцию Тейла – Сена не принято применять на временных рядах без тренда, поскольку она может найти ненулевую скорость из-за колебаний шума, особенно в случае, если временной ряд с малым числом наблюдений.

Существуют также периодический и смешанный типы тренда, которые данными критериями классифицировать не рекомендуется. Поэтому если в проверяемых данных присутствует периодичность или тренд в дисперсии, рекомендуется использовать совокупность других критериев, эффективных для иных типов трендов (например, критерий Хсу [18, 38]).

3. Применение критериев на практических данных

В качестве примера работоспособности совокупности критериев рассмотрим временные ряды шести научных организаций из общего числа рассмотренных. В табл. 3 даны показатели публикационной активности этих организаций. Первые три примера представляют временные ряды каждого типа, классифицированные на первом этапе (рис. 1), последние три примера демонстрируют более редкие ситуации, связанные с последующими этапами проверки (рис. 2).

В табл. 4 приведены результаты применения критериев для каждого временного ряда, а также результат классификации совокупностью критериев. У первой организации все значения больше доверительного интервала (см. табл. 1, $\alpha = 0.1$), следовательно, временной ряд возрастающий. Все значения статистик по второй организации отрицательные и также не попали в доверительный интервал, следовательно, временной ряд классифицируется как убывающий. Для третьей организации все значения статистик попали в доверительный интервал, что характеризует временной ряд как ряд “без тренда”.

Т а б л и ц а 3. Количество публикаций за год по нескольким организациям
Table 3. Number of publications per year by several organizations

Организация	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
1	83	87	82	84	94	82	79	64	87	67
2	207	216	173	160	156	157	132	99	120	94
3	176	190	186	205	189	164	165	154	181	146
4	49	42	54	38	51	56	51	54	33	35
5	–	–	–	–	–	1	1	0	2	0
6	104	82	118	101	95	97	74	92	84	85
Организация	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
1	84	85	106	128	210	290	256	243	350	352
2	96	93	88	82	74	65	70	76	79	45
3	157	156	151	149	174	170	193	182	198	191
4	31	41	20	31	23	28	22	20	17	96
5	0	0	2	0	2	0	8	9	11	18
6	77	87	99	92	107	133	98	128	105	113

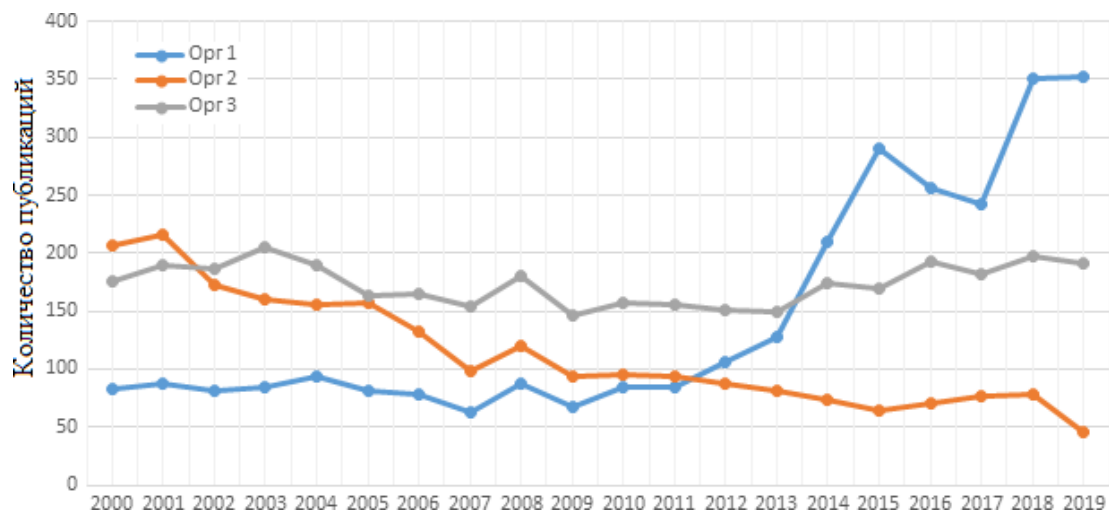


Рис. 1. Временные ряды количества публикаций организаций 1–3
 Fig. 1. Time series for the number of publications for organizations 1–3

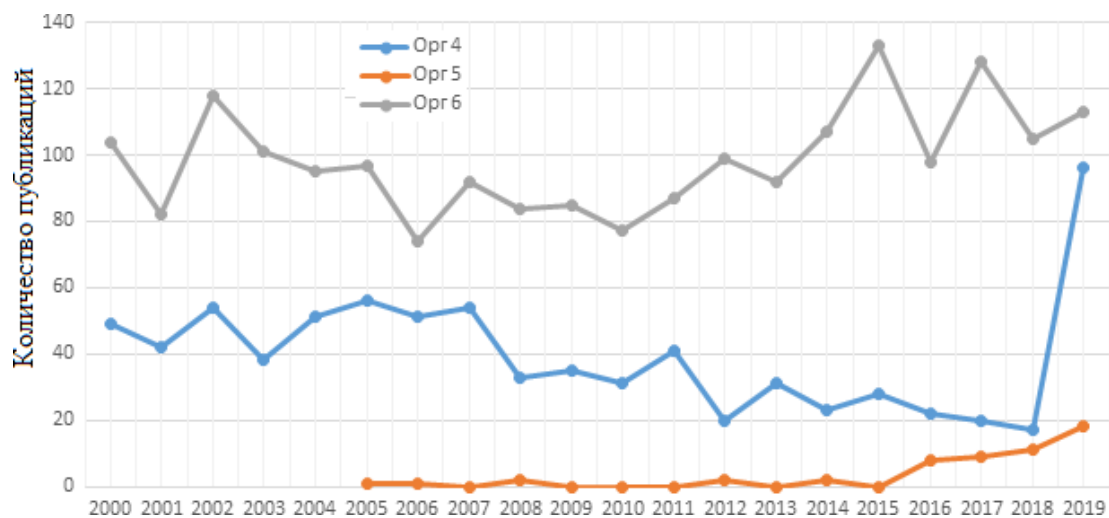


Рис. 2. Временные ряды количества публикаций организаций 4–6
 Fig. 2 Time series for the number of publications for organizations 4–6

Т а б л и ц а 4. Результаты проверки временных рядов организаций
 Table 4. Results of hypothesis testing of the organizations time series

Организация	Кокс – Стюарт	Инверсия	Манн – Кендалл	Совокупная проверка	
				Этап	Классификация
1	2.58	3.50	3.57	1	Возрастающий ряд
2	-2.69	-5.39	-5.35	1	Убывающий ряд
3	-0.22	-0.06	-0.03	1	Бестрендовый ряд
4	-1.54	-3.11	-2.96	2	Убывающий ряд
5	до сглаж.	2.36	2.43	3	Возрастающий ряд
	после	2.36	2.43		
6	1.97	1.49	1.49	–	Неопределенность

Временной ряд четвертой организации не классифицируется на первом этапе проверки. Причиной является статистика критерия Кокса – Стюарта, которая попала в доверительный интервал в отличие от двух других критериев. Произошло это из-за того, что во временном ряду присутствует выброс в правом краевом значении (рис. 2). Тем не менее на втором этапе две другие статистики перевешивают прогноз критерия и временной ряд классифицируется как убывающий. Временной ряд пятой организации также не классифицируется на первом этапе из-за статистики критерия инверсий, которая вследствие большого количества повторов не считает рост значимым. На втором этапе две другие статистики не перевешивают прогноз критерия, однако на третьем этапе значение статистики меняется от 1.34 до 2.42. После сглаживания эффект дискретных данных устраняется и временной ряд классифицируется как возрастающий. Последний временной ряд не получается однозначно классифицировать ни на одном этапе. В данном ряду присутствует как падение (с 2002 по 2010 г.), так и рост (с 2011 по 2019 г.) — у двух критериев эти два промежутка взаимно устраняются, первый критерий же улавливает небольшое различие в левых и правых краевых значениях. В неопределенности попадают такие временные ряды и временные ряды с недостаточной для классификации размерностью (организации, в которых сотрудники начали публиковать статьи с 2013 г. и позднее).

Наименование организаций, чьи временные ряды количества публикаций в SCOPUS были рассмотрены:

- 1) Башкирский государственный университет;
- 2) Научно-исследовательский физико-химический институт им. Л.Я. Карпова;
- 3) Институт теоретической физики им. Л.Д. Ландау РАН;
- 4) Национальный медицинский исследовательский радиологический центр;
- 5) Институт конструкторско-технологической информатики РАН;
- 6) Пушинский научный центр РАН.

Примечание. В представленных примерах критерий Манна – Кендалла фактически не ошибается, однако при рассмотрении большого числа временных рядов возникали ситуации, когда критерий ошибался. Поэтому применение данного критерия в совокупности с двумя другими оправдано.

4. Обсуждение результатов

Из полученных данных по 787 научным организациям рассмотрены 783 (четыре организации исключены из рассмотрения, поскольку не опубликовали ни одной статьи с 2017 по 2019 г. и, вероятно, ликвидированы). Временные ряды начинаются с того года, в котором зафиксирована первая публикация организации в периоде с 2000 по 2019 г. Нули показателей до первой публикации могут сильно исказить классификацию, поскольку такие организации заведомо будут растущими. В табл. 5 представлены результаты классификации временных рядов по трем мерам публикационной активности.

Большая часть временных рядов классифицируется на первом этапе, который среди всех этапов является наименее рискованным. Последующие этапы, однако, сокращают число “неопределенностей” (неклассифицированные временные ряды) раза в 2–3. Единственным постоянным временным рядом стала организация, которая после своей единственной работы в 2000 г. не публиковалась в SCOPUS в 2001–2014 и 2016 гг. На рис. 3 представлена диаграмма итогового результата для трех рассматриваемых мер (постоянный ряд добавлен в категорию “нет тренда”).

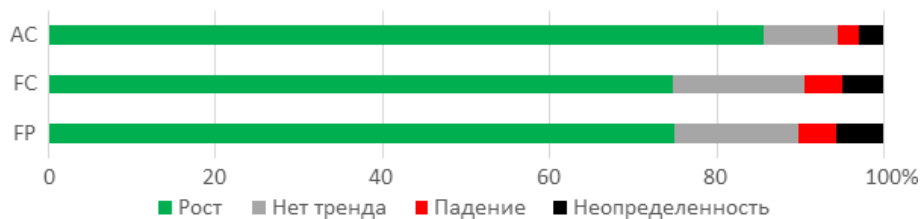


Рис. 3. Классификация научных организаций по мерам публикационной активности
 Fig. 3. Classification of scientific organizations by measures of publication activity

Т а б л и ц а 5. Классификация научных организаций (всего 783) по мерам публикационной активности

Table 5. Classification of scientific organizations (783 total) by measures of publication activity

Этап	Публикации целым счетом (АС)			Публикации фракционным счетом (FC)			КБПР (FP)		
	№ 1	№ 2	№ 3	№ 1	№ 2	№ 3	№ 1	№ 2	№ 3
Возрастающий	645	665	671	557	580	585	558	582	586
Убывающий	16	19	19	33	35	36	31	35	35
Нет тренда	52	67	68	97	121	122	100	112	116
Постоянный ряд	1	1	1	1	1	1	1	1	1
Неопределенность	69	31	24	95	46	39	93	53	45

Т а б л и ц а 6. Подмножества организаций по классификации трех метрик
 Table 6. Subsets of organizations divided by metric classification

АС	FC	FP	Кол-во	Комментарий
Рост	Рост	Рост	541	Рост всех показателей
То же	Рост	Нет тренда	28	Падение "качества" статей
«	Нет тренда	Рост	31	Рост "качества" статей
«	То же	Нет тренда	21	Коллаборации
«	«	Падение	2	Коллаборации/Падение "качества"
«	Падение	Рост	1	Коллаборации/Рост "качества"
«	Падение	Падение	2	Мегаколлаборации
Нет тренда	Нет тренда	Рост	5	Рост "качества" статей
То же	То же	Нет тренда	39	Стабильные организации
«	«	Падение	5	Падение "качества" статей
«	Падение	Нет тренда	7	Дефицит ресурсов
«	То же	Падение	6	То же
Падение	«	Нет тренда	1	«
Падение	«	Падение	17	«
Остальные			77	Есть какая-то неопределенность

Может показаться, что результаты классификации по публикациям фракционным счетом и КБПР совершенно не различаются, однако различия присутствуют. У 80 организаций классификации по FC и FP различаются. В табл. 6 представлены разные подмножества организаций по классификации трех мер одновременно.

Основная масса организаций находится в подмножестве полного роста. Тем менее разные классификации по метрикам могут выявить наиболее яркие примеры в стратегиях развития организаций. Например, организации, которые не растут в числе публикаций, но растут по КБПР, скорее всего ставят в приоритет рост качества публикаций. И наоборот, рост числа публикаций при отсутствии роста КБПР, вероятно, характери-

зуются преобладанием количества над качеством (но необязательно, могут быть выявлены жертвы методики КБПР или организации, которые достигли верха “совершенства” с точки зрения квартиля журнала).

В группу “Рост – Нет тренда – Падение” попали два института, которые очень часто участвуют в мегаколлорациях. Это Петербургский институт ядерной физики в составе Курчатовского института и Институт ядерной физики им. Г.И. Будкера СО РАН. Средний балл на статью у них несколько упал, но основные потери баллов связаны не с квартилями журналов, в которых они публикуются, а с уменьшением фракционной доли.

Уникальная ситуация в группе “Рост – Падение – Рост” — Национальный научный центр морской биологии ДВО РАН, где резко вырос КБПР, несмотря на небольшой рост числа публикаций и сильное увеличение доли коллабораций. Причина такого явления заключается в росте среднего балла на статью в пять раз. Если в нулевых годах большая часть статей была в бесквартильных журналах (в лучшем случае Q4), то к 2019 г. статьи организации стали публиковаться в журналах первых трех квартилей.

В группу “Рост – Падение – Падение” попали два института во ведении Курчатовского института — Институт физики высоких энергий и Институт теоретической и экспериментальной физики им. А.И. Алиханова. Число публикаций немалое, средний балл и качество высокое, но фракционная доля падает. В среднем в каждой публикации институтов только 12 % авторов относятся к данным организациям. Эти институты — яркие примеры того, что мегаколлорации с точки зрения КБПР являются плохой практикой.

Группа “Нет тренда – Нет тренда – Рост” содержит только пять институтов РАН. Самой крупной организацией здесь является Институт физиологии им. И.П. Павлова РАН, самый сильный рост среднего балла зафиксирован у Института теплофизики УрО РАН.

Группа “Нет тренда – Нет тренда – Падение” разнообразна. В нее попали Институт минералогии, геохимии и кристаллохимии редких элементов, Физико-энергетический институт им. А.И. Лейпунского, РАЕН, “Гиредмет” и Институт физики твердого тела РАН.

Группа “Нет тренда – Падение – Нет тренда” также довольно разнообразна. Эти организации скорее всего испытывают недостаток ресурсов, но при этом среднее “качество” статей не падает или даже растет. Среди таких организаций Институт теоретической и экспериментальной биофизики РАН, Институт высокомолекулярных соединений РАН, Санкт-Петербургский государственный технологический институт.

Единственной организацией в группе “Падение – Падение – Нет тренда” является АО Радиевый институт им. В.Г. Хлопина.

Подмножество полного падения наблюдается в большинстве своем у предприятий, ведущих опосредованную научную деятельность (АО, ООО, организации, подконтрольные Госкорпорации “РосКосмос”). Вполне вероятно, что у этих организаций стоят другие цели, кроме публикаций статей. Однако помимо этих организаций три имеют отношение к РАН: Бурятский научный центр СО РАН, Амурский научный центр ДВО РАН, Институт физических проблем им. П.Л. Капицы РАН. Также в эту группу попали НИИ органических полупродуктов и красителей, НМИЦ радиологии, ГНИИ органической химии и технологии, Всероссийский теплотехнический институт, РНЦ “Прикладная химия”.

Стоит отметить, что в большинстве своем в группы роста “качества” гораздо чаще попадают институты РАН, а в группу падения “качества” — университеты.

Заключение и выводы

В работе предложена совокупность статистических критериев, ориентированная на проверку отсутствия или наличия трендов во временных рядах. Полное описание рассмотренных критериев предоставлено, их преимущества и недостатки освещены. На практических данных демонстрируется применение совокупной проверки, ориентированной на повышение точности оценки при проверке на малых выборках. Рассматриваемый подход является весьма гибким инструментом, который может быть настроен под нужды исследователей и инженеров.

На примере публикационной активности 783 научных организаций РФ показано, как можно использовать критерии случайности и их совокупности для автоматизированного определения наличия/отсутствия трендов. Рост показателей по стране наблюдается у большей части рассмотренных организаций в период 2000–2019 гг., однако показатели ряда организаций не растут, а более того переживают спад. Пересечение по разным мерам демонстрирует существование разных паттернов развития публикационной активности организаций.

Классифицированы только российские научные организации по трем мерам публикационной активности с целью продемонстрировать практическую ценность применения статистических критериев случайности в наукометрии. Таким же образом можно классифицировать не только научные организации, но и журналы, авторов. Также можно классифицировать те же организации по другим показателям: уровень коллабораций, абсолютное и относительное (доля от общего числа публикаций) число статей в сборниках трудов конференций, количество активных авторов и т. п.

Применение совокупности критериев случайности или отсутствия тренда возможно везде, где имеет место анализ временных рядов. Использование этого раздела математической статистики способно помочь исследователю ответить на ряд вопросов, связанных с динамическими изменениями некоторых показателей, проводить сравнительные анализы, сужать область исследований и многое другое.

Благодарности. Автор выражает благодарность к. т. н. А.Е. Гуськову и Д.В. Косякову за ценные советы при подготовке статьи.

Список литературы

- [1] Лемешко Б.Ю., Комиссарова А.С., Щеглов А.Е. Вопросы применения некоторых критериев проверки случайности и отсутствия тренда. Метрология. 2010; (12):3–25.
- [2] Андерсон Т. Статистический анализ временных рядов. М.: Мир; 1976: 756.
- [3] Jaiswal R.K., Lohani A.K., Tiwari H.L. Statistical analysis for change detection and trend assessment in climatological parameters. Environmental Processes. 2015; 2(4):729–749. DOI:10.1007/s40710-015-0105-3.
- [4] Meals D.W., Spooner J., Dressing S.A., Harcum J.B. Statistical analysis for monotonic trends, Tech Notes 6. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA; 2011; 23. Available at: https://www.epa.gov/sites/default/files/2016-05/documents/tech_notes_6_dec2013_trend.pdf (accessed December 28, 2021).
- [5] Pan Y., Yuan Y., Sun T., Wang Y., Xie Y., Fan Z. Are the water quality improvement measures of China's south-to-north water diversion project effective? A case study of Xuzhou

- section in the East Route. *International Journal of Environmental Research and Public Health*. 2020; 17(17). DOI:10.3390/ijerph17176388.
- [6] **Kamimura K., Miyamoto H.** Trend analysis of long-term change in annual mean equilibrium water temperature in Japan by Mann–Kendall test. *Journal of Japan Society of Civil Engineers*. 2020; 8(1):154–160.
- [7] **De Andrade C.F., Delgado R.C., Barbosa M.L.F., Teodoro P.E., Silva Jr.C.A., Wanderley H.S., Capristo-Silva G.F.** Fire regime in Southern Brazil driven by atmospheric variation and vegetation cover. *Agricultural and Forest Meteorology*. 2020; 295.
- [8] **Sicard P., Paoletti E., Agathokleous E., Araminiene V., Proietti C., Coulibaly F., De Marco A.** Ozone weekend effect in cities: Deep insights for urban air pollution control. *Environmental Research*. 2020; 191. DOI:10.1016/j.envres.2020.110193.
- [9] **Шабанов П.А., Матвеева Т.А., Маркина М.Ю.** Межгодовые изменения событий очень сильных осадков на Европейской части России. *Фундаментальная и прикладная климатология*. 2017; (4):106–123. DOI:10.21513/2410-8758-2017-4-106-123.
- [10] **Dixon P.M., Pechmann J.H.K.** A statistical test to show negligible trend. *Ecological Society of America*. 2005; 86(7):1751–1756.
- [11] **Deng J., Zhang Y., Qin B., Yao X., Deng Y.** Trends of publications related to climate change and lake research from 1991 to 2015. *Journal of Limnology*. 2017; 76(3):439–450.
- [12] **Marrone M.** Application of entity linking to identify research fronts and trends. *Scientometrics*. 2020; 122(1):357–379.
- [13] **Лемешко Б.Ю., Комиссарова А.С., Щеглов А.Е.** Свойства и мощность некоторых критериев случайности и отсутствия тренда. *Научный вестник НГТУ*. 2012; 46(1):53–66.
- [14] **Ghods M., Amiri S., Hassani H., Ghods Z.** An enhanced version of Cochran–Armitage trend test for genome-wide association studies. *Meta Gene*. 2016; (9):225–229.
- [15] **Hu Z., Liu S., Zhong G., Lin H., Zhou Z.** Modified Mann–Kendall trend test for hydrological time series under the scaling hypothesis and its application. *Hydrological Sciences Journal*. 2020; 65(14):2419–2438. DOI:10.1080/02626667.2020.1810253.
- [16] **Tanil H.A.** Generalized version of Foster and Stuart’s d-statistic. *Communications Faculty of Sciences University of Ankara. Series A1 Mathematics and Statistics*. 2019; 68(1):98–110. DOI:10.31801/cfsuasmas.4436483.
- [17] **Кобзарь А.И.** *Прикладная математическая статистика. Для инженеров и научных работников*. М.: ФИЗМАТЛИТ; 2006: 816.
- [18] **Веретельникова И.В.** Исследование и применение критериев проверки гипотез об отсутствии тренда и критериев однородности. Дис. ... канд. тех. наук. Новосибирск: НГТУ; 2019: 203.
- [19] **Onoz B., Bayazit M.** The power of statistical tests for trend detection. *Turkish Journal of Engineering and Environmental Sciences*. 2003; 27(4):247–251.
- [20] **Yue S., Pilon P.** A comparison of the power of the t test, Mann–Kendall and bootstrap tests for trend detection. *Hydrological Sciences Journal*. 2016; (49):21–37.
- [21] **Гуськов А.Е., Косяков Д.В.** Национальный фракционный счет и оценка научной результативности организаций. *Научные и технические библиотеки*. 2020; (9):15–42.
- [22] Методика подсчета комплексного балла публикационной активности на портале Минобрнауки. Адрес доступа: https://minobrnauki.gov.ru/upload/Methodika_novaya.pdf (дата обращения 28.12.2021).

- [23] **Королёва В.В., Иванов О.В., Ведягин А.А., Лядов А.С., Леонидов А.В., Колобов А.В.** Публикационная активность как показатель эффективности научных организаций на примере организаций химического профиля. Вестник Российской академии наук. 2020; 90(10):948–958.
- [24] **Евдокимов В.И., Глухов В.А.** Комплексный балл публикационной результативности ведущих организаций МЧС России (2005–2019 гг.). Медико-биологические и социально-психологические проблемы безопасности в чрезвычайных ситуациях. 2020; (2):109–119. DOI:10.25016/2541-7487-2020-0-2-109-119.
- [25] **Максимов С.В.** “Новая” методика оценки эффективности научных организаций и перспективы развития конкуренции в российской науке. Информационное право. 2020; (3):4–8. Адрес доступа: <https://infolaw.ru/soderzhanie-nomerov/172-informatsionnoe-pravo-3-2020.html>.
- [26] **Горбунов-Посадов М.М.** Научная публикация — 2020. Научный сервис в сети Интернет. 2020; (22):184–191.
- [27] **Ракин В.И.** Проблемы управления междисциплинарным научным центром. Управление наукой: теория и практика. 2021; (3):55–67. DOI:10.19181/smtp.2021.3.1.3.
- [28] **Cox D.R., Stuart A.** Quick sign tests for trend in location and dispersion. *Biometrika*. 1955; (42):80–95.
- [29] **Химмельблау Д.** Анализ процессов статистическими методами. М.: Мир; 1973: 957.
- [30] **Mann H.B.** Non-parametric tests against trend. *Econometrica*. 1945; (13):245–259.
- [31] **Кендэл М.** Ранговые корреляции. М.: Статистика; 1975: 216.
- [32] **Theil H.** A rank-invariant method of linear and polynomial regression analysis. *Proceedings of the Koninklijke Nederlandse Akademie Wetenschappen, Series A Mathematical Sciences*. 1950; (53):1397–1412.
- [33] **Sen P.K.** Estimates of the regression coefficient based on Kendall’s tau. *Journal of the American Statistical Association*. 1968; (63):1379–1389.
- [34] **Ali R.O., Abubake S.R.** Trend analysis using Mann–Kendall, sen’s slope estimator test and innovative trend analysis method in Yangtze river basin, China: Review. *International Journal of Engineering and Technology*. 2019; 8(2):110–119. DOI:10.14419/ijet.v8i2.29591.
- [35] **Лемешко Б.Ю., Постовалов С.Н.** Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. Новосибирск: Изд-во НГТУ; 2004: 119.
- [36] **Foster F.G., Stuart A.** Distribution-free tests in time series dated on the breaking of records. *Journal of the Royal Statistical Society. Series A*. 1954; 16(1):1–22.
- [37] **Грешилов А.А., Стакун В.А., Стакун А.А.** Математические методы построения прогнозов. М.: Радио и связь; 1997: 112.
- [38] **Hsu D.A.** Test for variance shift at an unknown time point. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 1977; 26(3):279–284.
-

INFORMATION TECHNOLOGIES

DOI:10.25743/ICT.2022.27.1.005

The methodology for classifying time series by trend type in scientometrics

BLINOV PAVEL YU.*

¹Russian Research Institute of Economics, Politics and Law in Science and Technology (RIEPL), 127254, Moscow, Russia

*Corresponding author: Blinov Pavel Yu., e-mail: p.blinov@riep.ru

*Received March 18, 2021, revised September 16, 2021, accepted December 28, 2021.***Abstract**

Nowdays reasearchers often encounter the problems dealing with processing large amounts of data and so there is a growing interest in various quantitative metrics in different areas. In particular, a number of metrics characterizes publication activity in scientometrics. Many countries are currently experiencing an increase in publication activity, moreover, some countries have a tendency of abnormal growth for some reasons. The methods of mathematical statistics are quite effective for identification of certain growth patterns. In this paper, usefulness of statistical tests of randomness or trend tests for time series data in scientometrics is demonstrated. Methodology of aggregate testing by several trend tests is proposed. This methodology, being a very flexible tool, is focused on improving the accuracy of estimation for small size time series as well as allows a large number of time series to be checked and grouped by type of trend.

Keywords: scientometrics, publication activity, statistical hypothesis testing, randomness test, power of test, trends, significance level.

Citation: Blinov P.Yu. The methodology for classifying time series by trend type in scientometrics. Computational Technologies. 2022; 27(1):52–69. DOI:10.25743/ICT.2022.27.1.005. (In Russ.)

Acknowledgements. The author is grateful to PhD A.E. Guskov and D.V. Kosyakov for their valuable advices in preparation of the article.

References

1. **Lemeshko B.Yu., Komissarova A.S., Tsheglov A.E.** Application of tests for trend detection and checking for randomness. *Metrologiya*. 2010; (12):3–25. (In Russ.)
2. **Anderson T.W.** The statistical analysis of time series. N.Y.: John Wiley; 1971: 704.
3. **Jaiswal R.K., Lohani A.K., Tiwari H.L.** Statistical analysis for change detection and trend assessment in climatological parameters. *Environmental Processes*. 2015; 2(4):729–749. DOI:10.1007/s40710-015-0105-3.
4. **Meals D.W., Spooner J., Dressing S.A., Harcum J.B.** Statistical analysis for monotonic trends, Tech Notes 6. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, VA; 2011; 23. Available at: https://www.epa.gov/sites/default/files/2016-05/documents/tech_notes_6_dec2013_trend.pdf (accessed December 28, 2021).
5. **Pan Y., Yuan Y., Sun T., Wang Y., Xie Y., Fan Z.** Are the water quality improvement measures of China's south-to-north water diversion project effective? A case study of Xuzhou section in the East Route. *International Journal of Environmental Research and Public Health*. 2020; 17(17). DOI:10.3390/ijerph17176388.
6. **Kamimura K., Miyamoto H.** Trend analysis of long-term change in annual mean equilibrium water temperature in Japan by Mann–Kendall test. *Journal of Japan Society of Civil Engineers*. 2020; 8(1):154–160.

7. **De Andrade C.F., Delgado R.C., Barbosa M.L.F., Teodoro P.E., Silva Jr. C.A., Wanderley H.S., Capristo-Silva G.F.** Fire regime in Southern Brazil driven by atmospheric variation and vegetation cover. *Agricultural and Forest Meteorology*. 2020; 295.
8. **Sicard P., Paoletti E., Agathokleous E., Araminiene V., Proietti C., Coulibaly F., De Marco A.** Ozone weekend effect in cities: Deep insights for urban air pollution control. *Environmental Research*. 2020; 191. DOI:10.1016/j.envres.2020.110193.
9. **Shabanov P.A., Matveeva T.A., Markina M.Yu.** Inter-annual variations of heavy precipitation events over European Russia. *Fundamental and Applied Climatology*. 2017; (4):106–123. DOI:10.21513/2410-8758-2017-4-106-123. (In Russ.)
10. **Dixon P.M., Pechmann J.H.K.** A statistical test to show negligible trend. *Ecological Society of America*. 2005; 86(7):1751–1756.
11. **Deng J., Zhang Y., Qin B., Yao X., Deng Y.** Trends of publications related to climate change and lake research from 1991 to 2015. *Journal of Limnology*. 2017; 76(3):439–450.
12. **Marrone M.** Application of entity linking to identify research fronts and trends. *Scientometrics*. 2020; 122(1):357–379.
13. **Lemeshko B.Yu., Komissarova A.S., Tsheglov A.E.** Properties and power of tests for trend detection and checking for randomness. *Science Bulletin of NSTU*. 2012; 46(1):53–66. (In Russ.)
14. **Ghods M., Amiri S., Hassani H., Ghods Z.** An enhanced version of Cochran–Armitage trend test for genome-wide association studies. *Meta Gene*. 2016; (9):225–229.
15. **Hu Z., Liu S., Zhong G., Lin H., Zhou Z.** Modified Mann–Kendall trend test for hydrological time series under the scaling hypothesis and its application. *Hydrological Sciences Journal*. 2020; 65(14):2419–2438. DOI:10.1080/02626667.2020.1810253.
16. **Tanil H.A.** Generalized version of Foster and Stuart’s d-statistic. *Communications Faculty of Sciences University of Ankara. Series A1 Mathematics and Statistics*. 2019; 68(1):98–110. DOI:10.31801/cfsuasmas.4436483.
17. **Kobzar A.I.** *Prikladnaya matematicheskaya statistika. Dlya inzhenerov i nauchnykh rabotnikov [Applied mathematical statistics. For engineers and scientists]*. Moscow: Fizmatlit; 2006: 816. (In Russ.)
18. **Veretel’nikova I.V.** *Issledovanie i primeneniye kriteriev proverki gipotez ob otsutstvii trenda i kriteriev odnorodnost [The research and application of randomness tests and homogeneity tests]*. Novosibirsk: NGTU; 2019: 203. (In Russ.)
19. **Onoz B., Bayazit M.** The power of statistical tests for trend detection. *Turkish Journal of Engineering and Environmental Sciences*. 2003; 27(4):247–251.
20. **Yue S., Pilon P.** A comparison of the power of the t test, Mann–Kendall and bootstrap tests for trend detection. *Hydrological Sciences Journal*. 2016; (49):21–37.
21. **Guskov A.E., Kosyakov D.V.** National fractional calculations and evaluating organization’s science efficiency. *Scientific and Technical Libraries*. 2020; (9):15–42. (In Russ.)
22. **Methodology for calculating the complex score of publication activity on the portal of the Ministry of Education and Science of the Russian Federation.** Available at: https://minobrnauki.gov.ru/upload/Methodika_novaya.pdf (accessed December 28, 2021). (In Russ.)
23. **Koroleva V.V., Ivanov O.V., Vedyagin A.A., Lyadov A.S., Leonidov A.V., Kolobov A.V.** Publication activity as a research efficiency indicator by the example of chemical organizations. *Herald of the Russian Academy of Sciences*. 2020; 90(5):607–617. DOI:10.1134/S1019331620050147.
24. **Evdokimov V.I., Glukhov V.F.** Integrated score of publication performance of leading organizations of EMERCOM of Russia (2005–2019). *Medico-Biological and Socio-Psychological Problems of Safety in Emergency Situations*. 2020; (2):109–119. DOI:10.25016/2541-7487-2020-0-2-109-119. (In Russ.)
25. **Maksimov S.V.** A “new” method of assessment of efficiency of scientific organizations and competition development prospects in the Russian science. *Informatsionnoe Pravo*. 2020; (3):4–8. Available at: <https://infolaw.su/soderzhanie-nomerov/172-informatsionnoe-pravo-3-2020.html>. (In Russ.)
26. **Gorbunov-Possadov M.M.** Scientific publication — 2020. *Scientific Service in Internet*. 2020; (22):184–191. (In Russ.)
27. **Rakin V.I.** Management problems of an interdisciplinary research center. *Management of Science: Theory and Practice*. 2021; (3):55–67. DOI:10.19181/smt.p.2021.3.1.3. (In Russ.)
28. **Cox D.R., Stuart A.** Quick sign tests for trend in location and dispersion. *Biometrika*. 1955; (42):80–95.
29. **Himmelblau D.M.** *Process analysis by statistical methods*. N.Y.: John Wiley; 1970: 463.

30. **Mann H.B.** Non-parametric tests against trend. *Econometrica*. 1945; (13):245–259.
31. **Kendal M.G.** Rank correlation methods. London: Charles Griffin; 1975: 196.
32. **Theil H.** A rank-invariant method of linear and polynomial regression analysis. *Proceedings of the Koninklijke Nederlandse Akademie Wetenschappen, Series A Mathematical Sciences*. 1950; (53):1397–1412.
33. **Sen P.K.** Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*. 1968; (63):1379–1389.
34. **Ali R.O., Abubake S.R.** Trend analysis using Mann–Kendall, sen's slope estimator test and innovative trend analysis method in Yangtze river basin, China: Review. *International Journal of Engineering and Technology*. 2019; 8(2):110–119. DOI:10.14419/ijet.v8i2.29591.
35. **Lemeshko B.Yu., Postovalov S.N.** Компьютерные технологии анализа данных и исследования статистических закономерностей: Учебное пособие [Computer technologies for data analysis and research of statistical laws]. Novosibirsk: Izdatel'stvo NGTU; 2004: 196. (In Russ.)
36. **Foster F.G., Stuart A.** Distribution-free tests in time series dated on the breaking of records. *Journal of the Royal Statistical Society. Series A*. 1954; 16(1):1–22.
37. **Greshilov A.A., Stakun V.A., Stakun A.A.** Математические методы построения прогнозов [Mathematical methods for forecasting]. Moscow: Radio i Svyaz'; 1997: 112. (In Russ.)
38. **Hsu D.A.** Test for variance shift at an unknown time point. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 1977; 26(3):279–284.