

ИСПОЛЬЗОВАНИЕ ВЕЙВЛЕТОВ ДЛЯ АНАЛИЗА СРЕДНЕГОДОВЫХ ТЕМПЕРАТУР

Л. А. Зуйкевич¹⁾, Н. Н. Труш²⁾

¹⁾ *Белорусский государственный университет,
пр. Независимости, 4, 220030, г. Минск, Беларусь*

²⁾ *Белорусский государственный университет,
пр. Независимости, 4, 220030, г. Минск, Беларусь*

В данной работе вейвлеты используются для анализа данных о среднегодовой температуре в г. Минске за период 1888 – 2022 гг. выполняется непрерывное вейвлет-преобразование данных, строится скалограмма вейвлет-коэффициентов. Полученные результаты сравниваются с результатами, полученными с помощью преобразования Фурье. На основе дискретного вейвлет-разложения ряда строится прогноз.

Ключевые слова: вейвлет; вейвлет-анализ; непрерывное вейвлет-преобразование; скалограмма; преобразование Фурье; прогнозирование временных рядов; анализ температур.

WAVELET ANALYSIS OF AVERAGE ANNYAL TEMPERATURES

L. A. Zuikevich¹⁾, N. N. Trough²⁾

¹⁾ *Belarusian State University, Niezaliežnasci Av., 4, 220030, Minsk, Belarus*

²⁾ *Belarusian State University, Niezaliežnasci Av., 4, 220030, Minsk, Belarus*

In this paper, wavelets are used to analyze the average annual temperatures obtained in Minsk from 1888 to 2022. A continuous wavelet transformation of the data is performed, and a scalogram of wavelet coefficients is plotted. The obtained results are compared with the results obtained using the Fourier transform. A data forecast is constructed based on the discrete wavelet decomposition of the series.

Keywords: wavelet; wavelet analysis; continuous wavelet transform; scaleogram; Fourier transform; time series forecast; temperature analysis.

Введение

Вейвлет-анализ – относительно новое направление обработки данных, появившееся в 80-тых годах прошлого века. Вейвлеты, в силу своих свойств, дают возможность обрабатывать особенности данных, которые

при традиционных подходах анализа учесть довольно трудно. Кроме того, результаты, полученные с помощью вейвлет-анализа, зачастую обладают большей информативностью.

Вейвлет-преобразование и вейвлет-анализ широко используются в решении многих практических задач, в частности, для анализа сигналов. Например, вейвлет-преобразование было выбрано для сжатия изображения при разработке стандарта Joint Photographic Experts Group 2000 (JPEG-2000). Также вейвлеты используются для анализа и прогноза среднегодовых температур, например, в работах [1], [2], [3], [4].

Анализ среднегодовых температур показывает, что исходный временной ряд можно разложить на долгосрочный тренд, гармоническую составляющую и аномалии, являющиеся стохастической составляющей. Анализ гармонической составляющей удобно проводить по скалограмме вейвлет-коэффициентов и спектрограмме преобразования Фурье.

Основные определения

Вейвлетом (масштабирующей функцией) называется функция вида:

$$\varphi^{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right),$$

где функция $\psi(\cdot)$ называется материнским вейвлетом, a – сдвиг по времени, b – изменение временного масштаба, а множитель $1/\sqrt{a}$ обеспечивает независимость нормы функции от масштабирующего числа $a > 0, a, b \in R, t \in R$.

Формула непрерывного вейвлет-преобразования:

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

где $f(t)$ – исследуемая функция, $\psi(t)$ – материнский вейвлет, t – время, a – величина масштаба, b – параметр сдвига по временной оси, * – оператор комплексного сопряжения, $a > 0, a, b \in R, t \in R$.

Если a, b принимают только дискретные значения, то получим вейвлет-преобразование в дискретном варианте:

$$W^{m,n}(f) = a_0^{-\frac{m}{2}} \int_{-\infty}^{\infty} f(t) \psi(a_0^{-m}t - nb_0) dt,$$

где $f(t)$ – исследуемая функция, $\psi(t)$ – материнский вейвлет, t – время, a – величина масштаба, b – параметр сдвига по временной оси, $*$ – оператор комплексного сопряжения, $a = a_0^m, b = nb_0a_0^m, m, n \in \mathbb{Z}$ и $a_0 > 1, b_0 > 0$ – фиксированные числа

Скалограмма, или масштабодиграмма (анг. scaleogram, scalogram) – это спектр вейвлет преобразования. Определяется формулой:

$$PW(a,b) = |W(a,b)|^2,$$

где $W(a,b)$ – коэффициенты непрерывного вейвлет-преобразования, a – величина масштаба, b – параметр сдвига по временной оси, $a > 0, a, b \in \mathbb{R}$.

Материалы и методы исследования

В данной работе для анализа используются данные средних годовых температур в г. Минске (метеорологическая станция находится в г. Минске, Беларусь; 53.96 с.ш., 27.70 в.д., высота над уровнем моря - 231 м.) за период 1888 – 2022 гг. (данные за 135 лет).

Анализ ряда температур выполнен с помощью языка программирования python.

С помощью функции plot библиотеки matplotlib построим график исходного ряда (рис. 1):

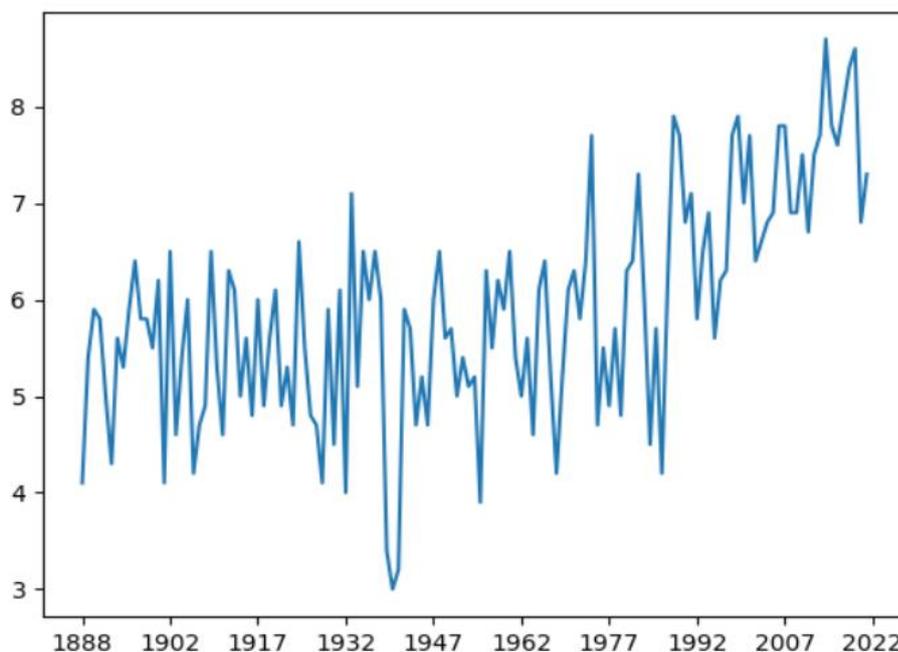


Рис. 1. График исходного ряда

Далее, исходя из результатов, полученных в работах [1], [3], произведем сглаживание ряда скользящей средней с интервалом сглаживания 10 лет. По мнению авторов статьи [5], такой интервал сглаживания уменьшает влияние локальных аномальных значений на циклические особенности поведения значений среднегодовых температур. Сглаживание выполнено с помощью функций `rolling` и `mean` библиотеки `pandas`. Построим графики исходного ряда и ряда 10-летнего скользящего среднего (рис. 2):

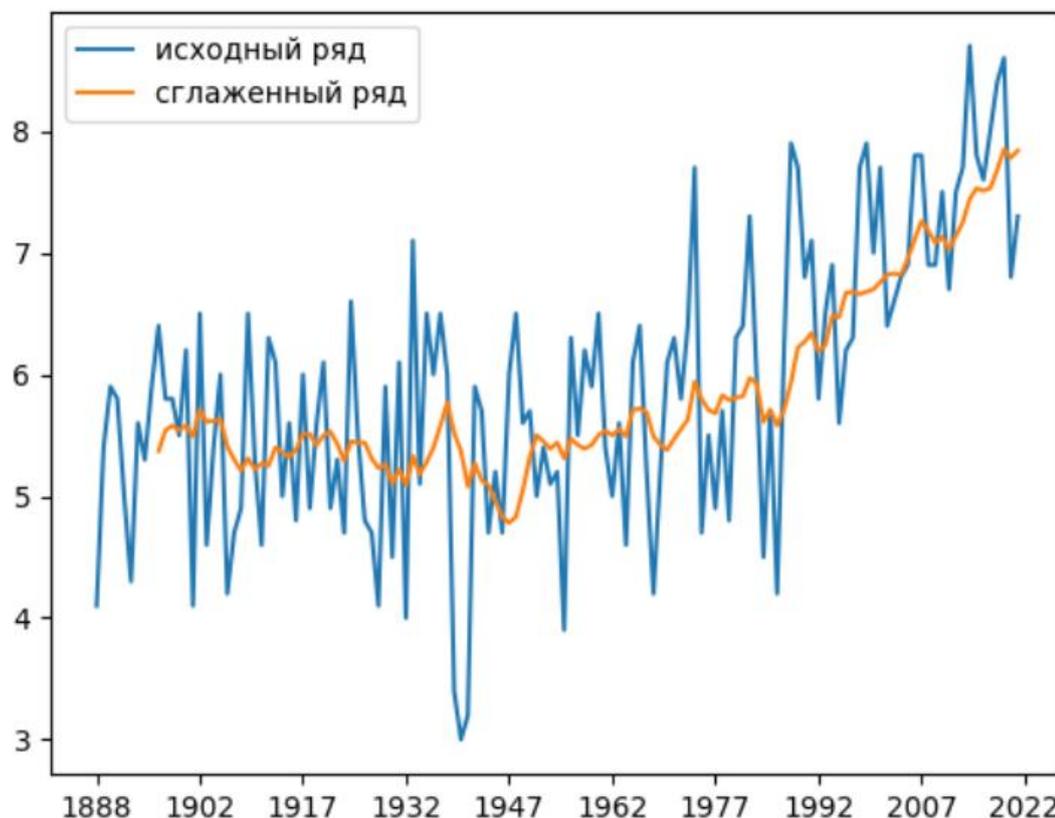


Рис. 2. Исходный ряд температур и ряд десятилетнего скользящего среднего

Далее у сглаженного ряда выделим и удалим квадратичный тренд. Для этого использованы функции `PolynomialFeatures`, `fit_transform`, `LinearRegression` и `predict` библиотеки `sklearn`. Построим графики сглаженного ряда и модели тренда (рис. 3), а также сглаженного ряда с удаленным трендом (рис. 4):

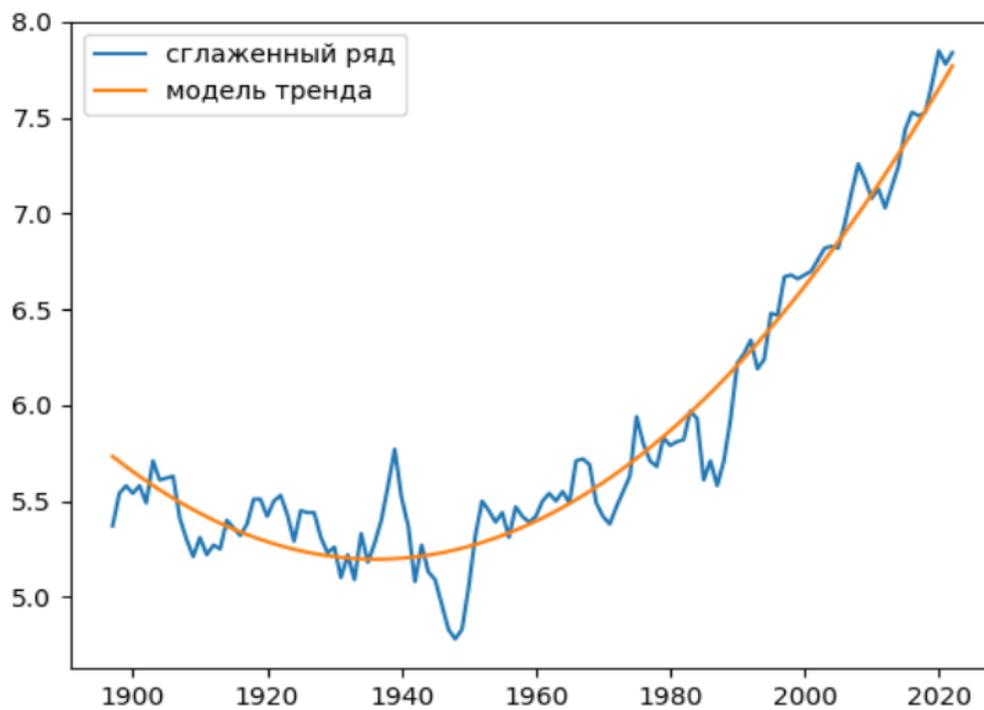


Рис. 3. Ряд десятискользящего среднего и модель квадратичного тренда

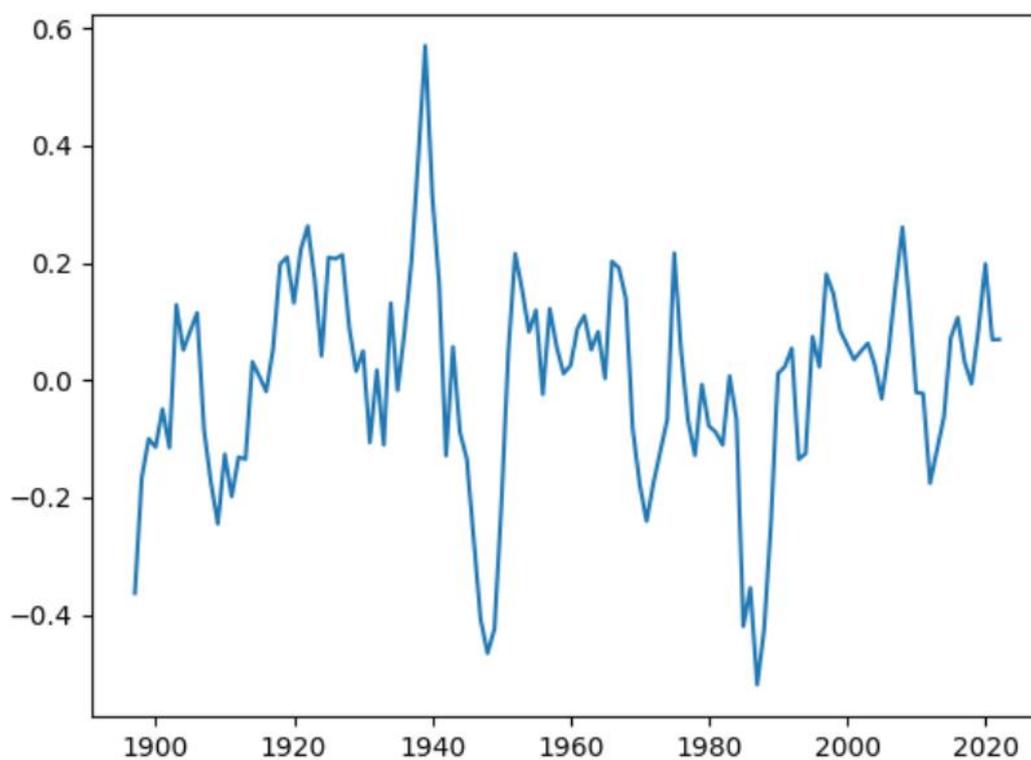


Рис. 4. Ряд десятилетнего сглаживания после удаления тренда

Построим скалограмму сглаженного ряда после удаления тренда. Рассмотрим временные масштабы от 2 до 60 лет. С помощью функции `cwt` библиотеки `ruwt` найдем коэффициенты непрерывного вейвлет-преобразования. Для преобразования использован вейвлет Морле, поскольку в работах [6], [7] при анализе ряда температур было замечено, что вейвлет Морле подходит для выявления масштабов колебаний и временной локализации экстремумов ряда. При построении скалограммы с помощью функции `imshow` библиотеки `matplotlib` была использована интерполяция (возможность функции `imshow`). Получили скалограмму ряда 10-летнего сглаживания после удаления тренда (рис. 5):

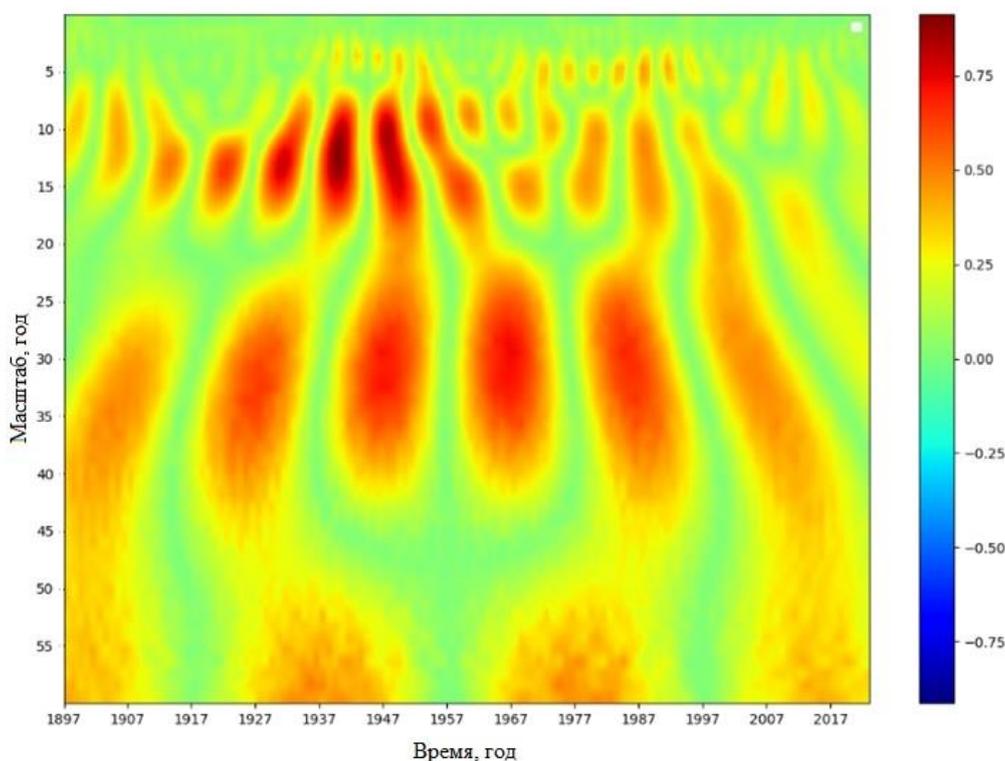


Рис. 5. Скалограмма абсолютных значений вейвлет-коэффициентов ряда десятилетнего скользящего среднего после удаления квадратичного тренда

Выполним преобразование Фурье сглаженного ряда без тренда. Рассчитаем коэффициенты преобразования с помощью функции `fft` библиотеки `scipy`, которая использует алгоритм быстрого преобразования Фурье. Возведем полученные значения в квадрат, чтобы построить спектрограмму преобразования Фурье данного ряда и выведем ее график (рис. 6):

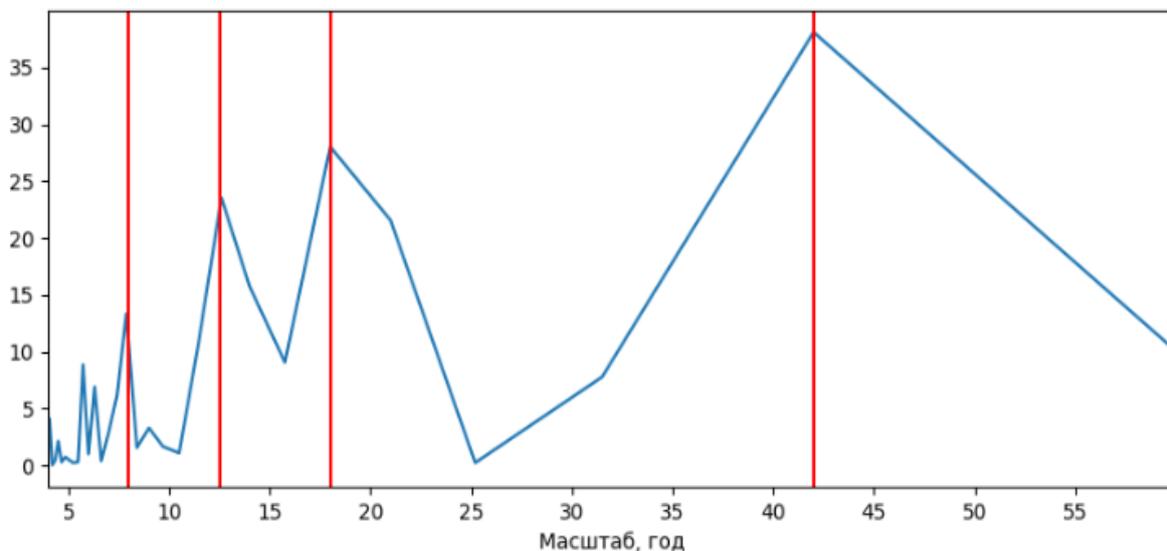


Рис. 6. Спектр преобразования Фурье ряда 10-летнего скользящего среднего после удаления квадратичного тренда

Далее построим прогноз рассматриваемого ряда. Существует несколько основных способов построения прогноза значений временного ряда с помощью вейвлетов [8]. Мы будем использовать метод, основанный на вейвлетном разложении.

После выполнения вейвлет-преобразования временного ряда $S_t, t = 1, T$ мы получаем матрицу W вейвлет-коэффициентов $W^{a,b}(S_t)$, в которой каждому масштабу $a \in A = \{a_1, a_2, \dots, a_n\}, n \in N$ соответствует вектор $W(a) = \{W^{a,1}, \dots, W^{a,T}\}$. Каждый такой вектор $W(a)$ может быть рассмотрен в качестве временного ряда, для прогноза значений которого применимы классические методы прогнозирования (такие как использование моделей ARMA, ARIMA). Тогда идея данного метода прогнозирования состоит в том, что для каждого вектора $W(a)$ мы строим прогноз с помощью классических методов, а после восстанавливаем временной ряд по прогнозу коэффициентов с помощью обратного вейвлет-преобразования.

Так как анализируется среднегодовая температура, период, на который делается прогноз, было решено выбрать небольшим, а именно 6 лет. Для того, чтобы сравнить прогноз с реальными данными, будем использовать для его построения часть ряда, и полученные прогнозные значения сверять с оставшейся частью ряда. Для начала выполним дискретное вейвлет-преобразование части ряда, получим два массива коэффициентов, которые будем рассматривать как временные ряды и строить по ним прогноз. Массивы коэффициентов были проверены на стационарность с помощью теста Дики-Фуллера и KPSS-теста, а также на нормальность с

помощью тестов Шапиро-Уилка и Колмогорова-Смирнова. Массивы коэффициентов оказались ненормальными и стационарными, что позволяет строить прогноз их значений с помощью модели авторегрессии скользящего среднего (ARMA). Для подбора параметров модели используется функция `auto_arima` библиотеки `pmdarima`. Результаты подбора оптимальных параметров модели ARMA для масштабирующих (сA2) и детализирующих (сD2) коэффициентов при разложении с помощью симлета `sym2` приведены на рис. 7. При подборе параметров модели ARIMA(p, d, q) видим, что параметр d равен нулю, что соответствует модели ARMA (рис. 7).

К построенному прогнозу ряда (Рис. 8) также добавим тренд и сравним со значениями исходного ряда (Рис. 9). Для построения прогноза были использованы разные вейвлеты, такие как вейвлеты Добеши Db1, Db2, Db3, Db4, Db5, Db6, койфлет `coif3`, вейвлет Хаара, симлеты `sym2`, `sym3`. Лучших результатов удалось добиться при использовании симлета `sym2` (рис. 8, рис. 9) и вейвлета Хаара `harr` (рис. 10).

```

Choose optimal arima parameters for cA2:
Performing stepwise search to minimize aic
ARIMA(2,0,2)(0,0,0)[0]           : AIC=inf, Time=0.13 sec
ARIMA(0,0,0)(0,0,0)[0]           : AIC=3.598, Time=0.01 sec
ARIMA(1,0,0)(0,0,0)[0]           : AIC=-17.432, Time=0.01 sec
ARIMA(0,0,1)(0,0,0)[0]           : AIC=-20.825, Time=0.01 sec
ARIMA(1,0,1)(0,0,0)[0]           : AIC=-21.648, Time=0.02 sec
ARIMA(2,0,1)(0,0,0)[0]           : AIC=-20.018, Time=0.03 sec
ARIMA(1,0,2)(0,0,0)[0]           : AIC=inf, Time=0.11 sec
ARIMA(0,0,2)(0,0,0)[0]           : AIC=-22.435, Time=0.04 sec
ARIMA(0,0,3)(0,0,0)[0]           : AIC=inf, Time=0.09 sec
ARIMA(1,0,3)(0,0,0)[0]           : AIC=inf, Time=0.14 sec
ARIMA(0,0,2)(0,0,0)[0] intercept : AIC=-20.505, Time=0.04 sec

Best model: ARIMA(0,0,2)(0,0,0)[0]
Total fit time: 0.633 seconds
Choose optimal arima parameters for cD2:
Performing stepwise search to minimize aic
ARIMA(2,0,2)(0,0,0)[0]           : AIC=-148.955, Time=0.12 sec
ARIMA(0,0,0)(0,0,0)[0]           : AIC=-143.456, Time=0.01 sec
ARIMA(1,0,0)(0,0,0)[0]           : AIC=-142.995, Time=0.01 sec
ARIMA(0,0,1)(0,0,0)[0]           : AIC=-142.736, Time=0.09 sec
ARIMA(1,0,2)(0,0,0)[0]           : AIC=inf, Time=0.10 sec
ARIMA(2,0,1)(0,0,0)[0]           : AIC=inf, Time=0.10 sec
ARIMA(3,0,2)(0,0,0)[0]           : AIC=inf, Time=0.13 sec
...

Best model: ARIMA(3,0,3)(0,0,0)[0]

```

Рис. 7. Результаты выполнения теста Дики-Фуллера



Рис. 8. Прогноз ряда 10-летнего сглаживания без тренда с помощью симплета sum2 и наблюдаемые значения

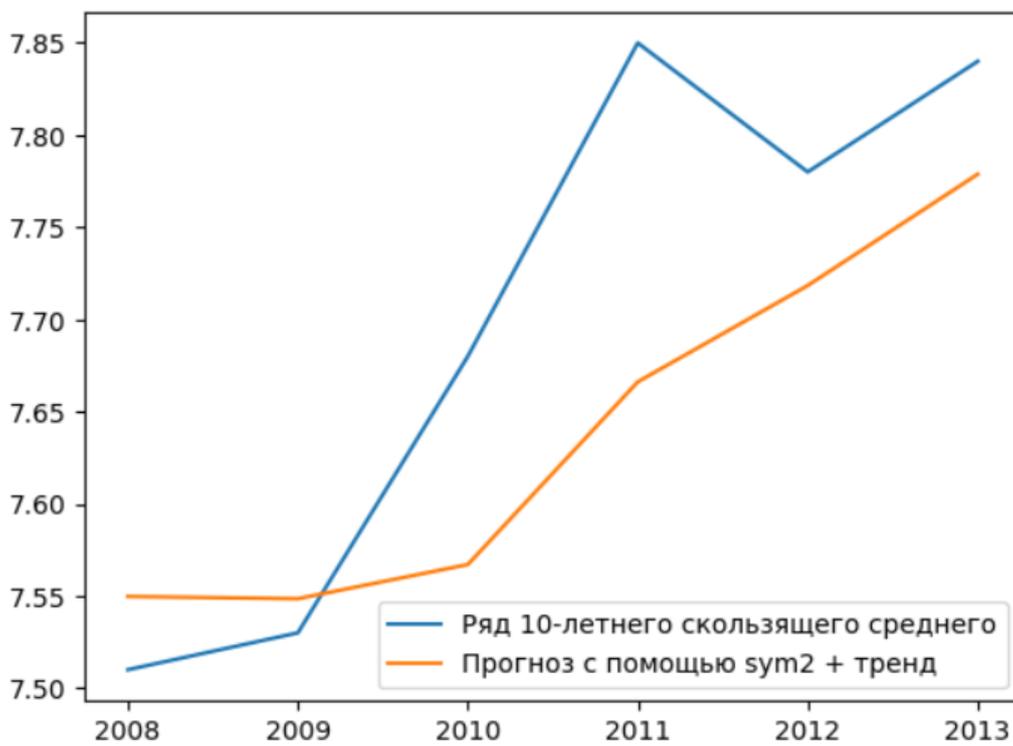


Рис. 9. Прогноз ряда 10-летнего сглаживания без тренда с помощью симплета sum2 с добавлением тренда и наблюдаемые значения ряда десятилетнего сглаживания

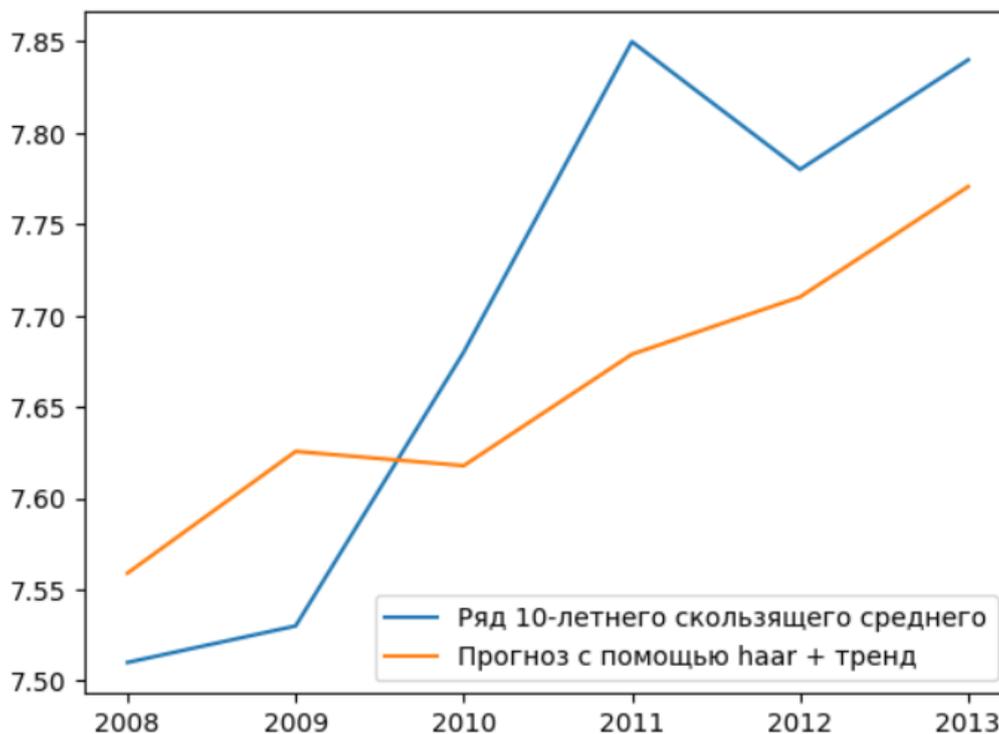


Рис. 10. Прогноз ряда 10-летнего сглаживания без тренда с помощью симплета $sym2$ с добавлением тренда и наблюдаемые значения ряда десятилетнего сглаживания

Анализ результатов

Анализ скалограммы показывает, что для распределения коэффициентов вейвлет-преобразования характерно наличие квазипериодических структур различных временных масштабов.

На графике Фурье-преобразования (рис. 6) выделены некоторые периодичности масштабов, а именно 8, 12.5, 18 лет и 42 года. Эти значения соответствуют гармоникам, на которые может быть разложен наш ряд. Такие же периодичности масштабов мы можем увидеть на скалограмме анализируемого ряда (рис. 5), однако распределение коэффициентов вейвлет-преобразования дает нам также информацию о временных периодах проявления колебаний различных масштабов. Также на скалограмме видна гармоника с периодом в 15 лет, которой соответствуют наибольшие значения вейвлет-коэффициентов, им соответствует временной отрезок с 1927 по 1952 г.

Так, гармоники с периодом в 8 лет проявляются только на временном интервале 1987 – 1997 г., а гармоники с периодом 18 лет – на временном интервале 1907 – 1982 г. Гармоники с периодом 42 года проявляются практически на всем рассматриваемом временном отрезке 1897 – 2022 г., а на его концах, на временных интервалах 1897 – 1907 и

2007 – 2022 переходят в гармониками большего масштаба. В этом и заключается один из плюсов вейвлет-преобразования: оно позволяет получить информацию об изменении относительного вклада компонент разного масштаба во времени. С помощью вейвлет-анализа можно получить частотно-временное представление данных, в то время как преобразование Фурье дает либо представление в частотной области, либо – во временной. Также вейвлет-преобразование может обнаружить локальные особенности данных, такие как скачки или выбросы.

Далее были приведены теоретические сведения о способах прогнозирования временных рядов с помощью вейвлетов. Один из способов, основанный на том, чтобы выполнить дискретное преобразование ряда, построить прогноз для каждого вектора вейвлет-коэффициентов с помощью классических методов прогнозирования (например, с помощью моделей ARMA, ARIMA и т.д.) был также реализован с помощью языка программирования Python. Для построения прогноза было использовано дискретное вейвлет-преобразование с помощью различных вейвлетов. Наилучших результатов удалось добиться с использованием симлета $sym2$ (рис. 8, рис. 9), а также вейвлета Хаара (рис. 10), прогноз с помощью вейвлетов Добеши $db2$ и $db1$ полностью идентичен прогнозам с помощью симлета $sym2$ и вейвлета Хаара соответственно. Максимальное отклонение прогноза, построенного с помощью симлета $sym2$, от наблюдаемых значений ряда составляет 2.34%, минимальное отклонение – 0.25%, среднее отклонение – 1.03%. Для прогноза, построенного с помощью вейвлета Хаара, максимальное отклонение прогноза от наблюдаемых значений составляет 2.18%, минимальное отклонение – 0.65%, среднее отклонение – 1.12%.

Заключение

В данной работе для выделения и анализа гармонической составляющей ряда среднегодовых температур в г. Минске было использовано вейвлет-преобразование и преобразование Фурье. Результаты, полученные с помощью вейвлет-преобразования, оказались более информативными, поскольку позволяют не только выделить основные периодичности масштабов, но и получить дополнительную информацию о временных интервалах проявления компонент различного масштаба и частоты. При анализе рядов температур дополнительная информация, которую дает вейвлет-разложение, может быть использована при прогнозе будущих среднегодовых значений температур. Для построения прогноза ряда температур было использовано дискретное вейвлет-преобразование с

помощью различных вейвлетов. Лучших результатов удалось добиться с использованием симплета $\text{sym}2$ и вейвлета Хаара. Так как данные о температуре являются среднегодовыми, прогноз строился на небольшой период (6 лет). Отклонение построенного прогноза от наблюдаемых данных находится в диапазоне от 0,25 % до 2,34 %.

Библиографические ссылки

1. *Ипполитов И.И., Кабанов М.В., Логинов С.В.* Применение вейвлет-преобразования для анализа межгодовых колебаний приземной температуры воздуха в Томске и солнечной активности // *Оптика атмосферы и океана*. 2001. №4(14). С. 280–285.

2. Прогнозирование климатических характеристик с помощью метода вейвлет-преобразования / С.Ю. Золотов [и др.] // *Оптика атмосферы и океана*. 2005. №18(4). С. 349–351.

3. *Золотов С.Ю., Ипполитов И.И., Логинов С.В.* Прогнозные оценки изменения температуры приземного воздуха с использованием метода вейвлет-преобразования // *Оптика атмосферы и океана*. 2009. №22(5). С. 471–475.

4. Time scales and trends in the central England temperature data (1659–1990): A wavelet analysis / S. Baliunas [et al] // *Geophysical Research Letters*. 1997. Vol. 11, № 24. P.1283-1438.

5. *Козлов А.Н., Пармузин С.Ю., Пустовой Г.П.* Вероятностное прогнозирование среднегодовой температуры грунтов и их устойчивости к оттаиванию и промерзанию // *Кристофера Земли*. 1999. №3(4). С. 26–31.

6. *Ипполитов И.И., Кабанов М.В., Логинов С.В.* Вейвлет-преобразование при анализе природно-климатических изменений // *Оптика атмосферы и океана*. 2002. №15(1). С. 21–28.

7. *Гайнутдинова Р.Д., Крымская Д.Н.* Особенности частотного спектра долгопериодных изменений приземной температуры и параметров солнечной активности в регионе Кыргызской Республики // *Труды Гидрометцентра России*. 2017. №366. С. 112–120.

8. *Schluter S., Deuschle C.* Using wavelets for time series forecasting: Does it pay off? // *FAU Discussion Papers in Economics*. 2010. №4. P. 1–28.