

Л. А. Ковальчук

Последовательный статистический анализ экстремальных аномалий температуры воздуха

(Представлено членом-корреспондентом НАН Украины В. И. Осадчим)

Осуществлен последовательный статистический анализ ежедневных аномалий температуры воздуха с 1900 по 2013 гг. в Киеве. Посредством критерия Вальда установлено, что положительные экстремальные аномалии температуры воздуха в течение прошедшего века были статистически значительно выше $5,00^\circ\text{C}$ и ниже $6,67^\circ\text{C}$, а отрицательные экстремальные аномалии ниже $-5,00^\circ\text{C}$ и выше $-9,50^\circ\text{C}$. Отмечено увеличение частоты положительных экстремальных аномалий и уменьшение частоты отрицательных аномалий температуры воздуха за последние два десятилетия.

Состояние проблемы. Доминирующие суждения о современном состоянии климата и его колебаниях основаны преимущественно на анализе линейных и гармонических трендов температуры воздуха, хотя общеизвестно, что статистическая значимость тренда существенно зависит от крайних значений анализируемого периода и может резко изменяться по мере поступления новых измерений температуры. В то же время были предложены методы последовательного статистического анализа, позволяющие осуществлять объективную проверку альтернативных гипотез с учетом поступающих данных [1–3].

Цель работы — посредством “последовательного критерия отношения вероятностей” Вальда проверить с заданной достоверностью гипотезу об увеличении частоты экстремальных температур воздуха за последние десятилетия.

Материал и методика. В основу исследования положены ежедневные значения температуры воздуха в Киеве с 1900 по 2013 гг., полученные Гидрометеорологической службой Украины по результатам стандартных наблюдений, которые были очищены от детерминированной составляющей — годового хода. Из ежедневных значений температуры воздуха вычитались среднегодовые значения температуры соответствующего дня, т. е. значения медианы.

Суть критерия Вальда сводится к следующему: выбирая одно за другим m значений x_1, x_2, \dots, x_m из совокупности $f(x, \theta)$, на любом этапе отношение вероятностей данной выборки для гипотез H_0 ($\theta = \theta_0$) и H_1 ($\theta = \theta_1$) будет равно [3]:

$$L_m = \frac{\prod_{i=1}^m f(x_i, \theta_1)}{\prod_{i=1}^m f(x_i, \theta_0)}, \quad (1)$$

что действительно для любых распределений. В этой связи, задача была сведена к различению двух гипотез H_0 и H_1 , относящихся к вековому ряду аномалий температуры воздуха ΔT , которые принимали значения ΔT_0 и ΔT_1 . Выборочное пространство аномалий температуры делилось на три взаимно непересекающиеся области: а) область вероятностей

успеха ω_0 , при попадании в которую выборочного значения ΔT мы принимали гипотезу H_0 и отвергали H_1 ; б) область вероятностей успеха ω_1 , при попадании в которую выборочного значения ΔT принимали гипотезу H_1 и отвергали H_0 ; в) оставшаяся часть выборочного пространства ω_c , при попадании в которое значения ΔT продолжали выбор. Вероятности успеха ω_0, ω_1 определены соответственно значениям вероятностей попадания ΔT в диапазон аномалий температур выше нормы (ВН) или в диапазон температур значительно выше нормы (ЗВН). Статистика Вальда вычислялась, согласно выражению [3]:

$$\log \frac{\beta}{1-\alpha} \leq (n-k) \log \frac{1-\omega_1}{1-\omega_0} + k \log \frac{\omega_1}{\omega_0} \leq \log \frac{1-\beta}{\alpha}, \quad (2)$$

где n — число испытаний; k — число “успехов”; α — вероятность ошибки первого рода; β — вероятность ошибки второго рода.

Вероятности попадания $\Delta T(\omega_0, \omega_1)$ в диапазон ВН или ЗВН определены на основе функции распределения вероятностей (интеграла вероятностей) ежедневных аномалий температуры воздуха с 1900 по 2013 гг. Закон статистического распределения ежедневных аномалий температуры идентифицирован при помощи модели экспоненциального распределения [4], модифицированного для асимметричного случая:

$$p(x) = \frac{\nu}{2\lambda\sigma\Gamma(1/\nu)} \exp\left(-\left|\frac{x-m}{\lambda\sigma}\right|^\nu\right), \quad (3)$$

где $\lambda = \sqrt{\Gamma(1/\nu)/\Gamma(3/\nu)}$; $\Gamma(z)$ — гамма-функция; $\mu = 0$ — центр распределения; $\sigma = (\sigma_{\Delta T < 0}, \sigma_{\Delta T > 0})$ — стандартные отклонения, $\nu = (\nu_{\Delta T < 0}, \nu_{\Delta T > 0})$ — некоторые характеристики распределения, которые однозначно определяют параметр формы экспоненциальных распределений — эксцесс $\varepsilon = (\varepsilon_{\Delta T < 0}, \varepsilon_{\Delta T > 0})$:

$$\varepsilon = \frac{\Gamma(1/\nu)\Gamma(5/\nu)}{(\Gamma(3/\nu))^2}. \quad (4)$$

Выбор значений вероятности ошибки первого рода α и вероятности ошибки второго рода β был предопределен двумя правилами: первое — $\alpha + \beta < 1$; второе — изучаемая проблема связана со значительными материальными ущербами. Согласно рекомендациям, нами приняты: $\alpha = 0,01$, $\beta = 0,05$.

Процедуре различения были подвергнуты три пары гипотез: первая — $H_0: 3,98 \text{ }^\circ\text{C} < \Delta T \leq 5,00 \text{ }^\circ\text{C}$, $H_1: \Delta T > 5,00 \text{ }^\circ\text{C}$; вторая — $H_0: 3,98 \text{ }^\circ\text{C} < \Delta T \leq 6,00 \text{ }^\circ\text{C}$, $H_1: \Delta T > 6,00 \text{ }^\circ\text{C}$; третья — $H_0: 3,98 \text{ }^\circ\text{C} < \Delta T \leq 6,67 \text{ }^\circ\text{C}$, $H_1: \Delta T > 6,67 \text{ }^\circ\text{C}$. После решения выражения (2) с учетом значений $\alpha = 0,01$, $\beta = 0,05$ и ω_0, ω_1 (табл. 1) для различения каждой пары гипотез расчетные формулы статистик Вальда приобрели вид соответственно:

$$63 \leq 15k - n \leq 96; \quad 92 \leq 14k - n \leq 141; \quad 42 \leq 15k - n \leq 65.$$

Аналогичное исследование было выполнено для отрицательных аномалий температур, но вероятности успеха ω_0, ω_1 определены соответственно значениям вероятностей попадания ΔT в диапазон аномалий температур ниже нормы или в диапазон аномалий температур значительно ниже нормы. Процедуре различения были подвергнуты также три пары гипотез: первая — $H_0: -4,00 \text{ }^\circ\text{C} > \Delta T \geq -5,00 \text{ }^\circ\text{C}$, $H_1: \Delta T < -5,00 \text{ }^\circ\text{C}$; вторая — $H_0: -4,75 \text{ }^\circ\text{C} > \Delta T \geq -7,00 \text{ }^\circ\text{C}$, $H_1: \Delta T < -7,00 \text{ }^\circ\text{C}$; третья — $H_0: -4,75 \text{ }^\circ\text{C} > \Delta T \geq -9,50 \text{ }^\circ\text{C}$,

$H_1: \Delta T < -9,50$ °C. Иные по длине интервалы отрицательных аномалий температуры проверяемых пар гипотез, чем положительных аномалий ΔT , обусловлены асимметрией распределения аномалий температуры. Для различения каждой пары гипотез расчетные формулы статистик Вальда имели вид соответственно:

$$26 \leq 10k - n \leq 41; \quad 457 \leq 13k - n \leq 698; \quad 27 \leq 13k - n \leq 41.$$

Обсуждение результатов исследований. Сравнивая вероятности успехов ω_0, ω_1 каждой из проверенных пар гипотез (см. табл. 1), становится очевидным, что по мере уменьшения разницы между ω_0, ω_1 возрастает количество необходимых испытаний для различения гипотез H_0 и H_1 . Если для принятия гипотезы $H_0: 3,98$ °C $< \Delta T \leq 5,00$ °C было достаточно, чтобы ΔT не превысило 5,00 °C в первых последовательных 63 испытаниях, а для отклонения H_0 и принятия $H_1: \Delta T > 5,00$ °C требовалось, чтобы статистика Вальда была больше 96; то для принятия гипотезы $H_0: -4,75$ °C $> \Delta T \geq -7,00$ °C уже было необходимо 457 испытаний, а отклонения H_0 и принятия $H_1: \Delta T < -7,00$ °C требовалось 698 испытаний.

Процедура последовательного различения гипотез $H_0: 3,98$ °C $< \Delta T \leq 5,00$ °C и $H_1: \Delta T > 5,00$ °C, начатая с ноября 1900 г. обнаружила, что аномалии температуры статистически значительно превышали 5,00 °C до 31 октября 1916 г., в связи с чем для периода 1.11.1900–31.10.1916 гг. была принята гипотеза $H_1: \Delta T > 5,00$ °C (табл. 2). По мере включения в расчет последующих значений аномалий температуры до 31.10.2013 г. было обнаружено, что гипотеза $H_1: \Delta T > 5,00$ °C также должна быть принята. Рост значений статистики Вальда до 23 240 за период 1900–2013 гг. свидетельствует только о том, что с достоверностью 95% мы можем утверждать о верности гипотезы $H_1: \Delta T > 5,00$ °C (кривая 1 на рис. 1, а).

Последовательное различение гипотез $H_0: 3,98$ °C $< \Delta T \leq 6,00$ °C, $H_1: \Delta T > 6,00$ °C приводит к выводу о верности гипотезы $H_0: 3,98$ °C $< \Delta T \leq 6,00$ °C для периода 1900–2013 гг. (см. кривую 2 на рис. 1, а). Различение гипотез $H_0: 3,98$ °C $< \Delta T \leq 6,67$ °C, $H_1: \Delta T > 6,67$ °C, когда значения статистик Вальда с 1900 г. по 2013 г. были меньше критического значения, позволили принять гипотезу $H_0: 3,98$ °C $< \Delta T \leq 6,67$ °C с достоверностью 99% (см. кривую 3 на рис. 1, а).

Таблица 1

Гипотеза	Вероятности успеха: ω_0/ω_1	Число испытаний		Статистика Вальда
		H_0	H_1	
$H_0: 3,98$ °C $< \Delta T \leq 5,00$ °C	0,049	<63		$15k - n$
$H_1: \Delta T > 5,00$ °C	0,093		>96	
$H_0: 3,98$ °C $< \Delta T \leq 6,00$ °C	0,086	<92		$14k - n$
$H_1: \Delta T > 6,00$ °C	0,056		>141	
$H_0: 3,98$ °C $< \Delta T \leq 6,67$ °C	0,104	<42		$15k - n$
$H_1: \Delta T > 6,67$ °C	0,039		>65	
$H_0: -4,00$ °C $> \Delta T \geq -5,00$ °C	0,050	<26		$10k - n$
$H_1: \Delta T < -5,00$ °C	0,150		>41	
$H_0: -4,75$ °C $> \Delta T \geq -7,00$ °C	0,083	<457		$13k - n$
$H_1: \Delta T < -7,00$ °C	0,077		>698	
$H_0: -4,75$ °C $> \Delta T \geq -9,50$ °C	0,130	<27		$13k - n$
$H_1: \Delta T < -9,50$ °C	0,030		>41	

Таблиця 2

Период	Количество		Статистика Вальда			Гипотеза	
	испытаний n (кол-во дней от 1.11.1900 г.)	успехов k	число испытаний для принятия H_0	$15k - n$	число испытаний для принятия H_1	$H_0: 3,98 \leq \Delta T \leq \leq 5,00 \text{ }^\circ\text{C}$	$H_1: \Delta T > 5,00 \text{ }^\circ\text{C}$
1.11.1900							
31.10.1901	365	32	<63	115	>96		Принимается
31.10.1902	730	62	<63	200	>96		То же
31.10.1903	1095	117	<63	660	>96		"
31.10.1904	1460	139	<63	625	>96		"
31.10.1905	1825	168	<63	695	>96		"
31.10.1906	2190	201	<63	825	>96		"
31.10.1907	2555	224	<63	805	>96		"
31.10.1908	2920	239	<63	665	>96		"
31.10.1909	3285	271	<63	780	>96		"
31.10.1910	3650	313	<63	1045	>96		"
31.10.1911	4015	338	<63	1055	>96		"
31.10.1912	4380	352	<63	900	>96		"
31.10.1913	4745	385	<63	1030	>96		"
31.10.1914	5110	420	<63	1190	>96		"
31.10.1915	5475	446	<63	1215	>96		"
31.10.1916	5840	471	<63	1225	>96		"

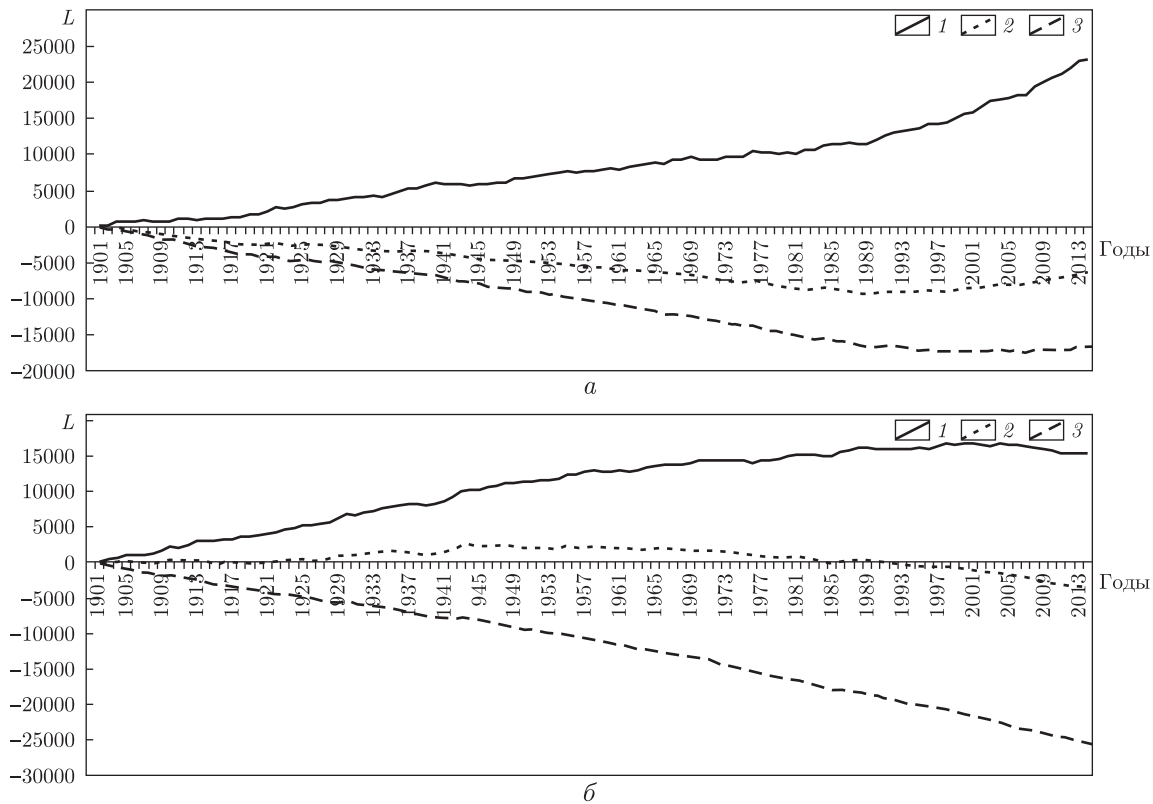


Рис. 1. Статистика Вальда (L) для гипотез: a — $H_1: \Delta T > 5,00$ °C (1); $H_0: 3,98$ °C < $\Delta T \leq 6,00$ °C (2); $H_0: 3,98$ °C < $\Delta T \leq 6,67$ °C (3); b — $H_1: \Delta T < -5,00$ °C (1); $H_0: -4,75$ °C > $\Delta T \geq -7,00$ °C (2); $H_0: -4,75$ °C > $\Delta T \geq -9,50$ °C (3)

В результате последовательного различения гипотез $H_0: -4,00$ °C > $\Delta T \geq -5,00$ °C и $H_1: \Delta T < -5,00$ °C была принята $H_1: \Delta T < -5,00$ °C, т.е. в течение века отрицательные аномалии температуры воздуха статистически значительно были ниже $-5,00$ °C (см. кривую 1 на рис. 1, б). После различения $H_0: -4,75$ °C > $\Delta T \geq -7,00$ °C и $H_1: \Delta T < -7,00$ °C для периода 1928–1981 гг. была принята гипотеза $H_1: \Delta T < -7,00$ °C, как свидетельство статистически значимого преобладания аномалий температуры воздуха менее $-7,00$ °C (см. кривую 2 на рис. 1, б). Из двух гипотез $H_0: -4,75$ °C > $\Delta T \geq -9,50$ °C, $H_1: \Delta T < -9,50$ °C была принята $H_0: -4,75$ °C > $\Delta T \geq -9,50$ °C (см. кривую 3 на рис. 1, б), что позволяет утверждать с достоверностью 99% об отсутствии статистически значимых различий в частотах аномалий температуры воздуха за вековой период, когда они оказывались ниже ($-9,50$ °C).

На основании изложенного выше можно сделать вывод, что положительные экстремальные аномалии температуры воздуха в течение прошедшего века были выше $5,00$ °C и ниже $6,67$ °C, а отрицательные экстремальные аномалии ниже ($-5,00$ °C) и выше ($-9,50$ °C). За последние два десятилетия отмечено увеличение частоты положительных экстремальных аномалий и уменьшение частоты отрицательных аномалий температуры воздуха.

1. Липцер Р. Ш., Ширяев А. Н. Статистика случайных процессов. – Москва: Наука, 1974. – 696 с.
2. Ширяев А. Н. Статистический последовательный анализ. – Москва: Наука, 1976. – 271 с.

3. Кендалл М., Стюарт А. Статистические выводы и связи. – Москва: Наука, 1973. – 899 с.
4. Булашев С. В. Статистика для трейдеров. – Москва: Компания Спутник+, 2003. – 245 с.

Український гідрометеорологічний інститут
ГСЧС України і НАН України, Київ

Поступило в редакцію 10.07.2014

Л. А. Ковальчук

Послідовний статистичний аналіз екстремальних аномалій температури повітря

Виконано послідовний статистичний аналіз щоденних аномалій температури повітря з 1900 по 2013 рр. в Києві. Завдяки критерію Вальда встановлено, що позитивні екстремальні аномалії температури повітря впродовж минулого століття були статистично значно вищі $5,00\text{ }^{\circ}\text{C}$ та нижчі $6,67\text{ }^{\circ}\text{C}$, а негативні екстремальні аномалії – нижчі $-5,00\text{ }^{\circ}\text{C}$ та вищі $-9,50\text{ }^{\circ}\text{C}$. Відзначено збільшення частоти позитивних екстремальних аномалій та зменшення частоти негативних аномалій температури повітря за останні два десятиліття.

L. A. Kovalchuk

Successive statistical analysis of extreme air temperature anomalies

The statistical analysis of daily air temperature anomalies in Kiev from 1900 to 2013 is consistently implemented. By the Wald test, it is revealed that the positive extreme temperature anomalies during the past century were statistically significantly higher than $5.00\text{ }^{\circ}\text{C}$ or below $6.67\text{ }^{\circ}\text{C}$ and negative below the extreme abnormalities $-5.00\text{ }^{\circ}\text{C}$ and above $-9.50\text{ }^{\circ}\text{C}$. The increase in the frequency of extreme positive anomalies and the decrease in the frequency of negative air temperature anomalies over the past two decades are found.