ДОКЛАДЫ РОССИЙСКОЙ АКАДЕМИИ НАУК. НАУКИ О ЗЕМЛЕ, 2022, том 502, № 1, с. 38–44

ГЕОГРАФИЯ

УДК 551.583

АНАЛИЗ ИЗМЕНЕНИЯ РЕЖИМА ОСАДКОВ НА ТЕРРИТОРИИ РОССИЙСКОЙ ФЕДЕРАЦИИ ВО ВТОРОЙ ПОЛОВИНЕ 20–НАЧАЛЕ 21 ВЕКА С ПРИМЕНЕНИЕМ БАЙЕСОВСКОЙ ОЦЕНКИ ПАРАМЕТРОВ МАРКОВСКОЙ ЦЕПИ

© 2022 г. И. О. Попов[1, *], Е. Н. Попова[2, **]

Представлено академиком РАН В.М. Котляковым 23.09.2021 г.
Поступило 23.09.2021 г.
После доработки 05.10.2021 г.
Принято к публикации 06.10.2021 г.

Произведена байесовская оценка изменения вероятностей однородных переходов в бинарной марковской цепи первого порядка как модели изменения временнóго режима осадков на территории России в 2000–2019 гг. по сравнению с 1966–1985 гг. Показано достоверное изменение этих вероятностей на значительной части территории России, причем в большинстве случаев наблюдаются увеличение вероятности однородных переходов 0→0 и уменьшение вероятности однородных переходов 1→1, что может означать увеличение продолжительности непрерывных периодов без осадков и уменьшение непрерывных периодов с осадками. Построены карты пространственного распределения этих изменений и статистически доказана его неслучайность.

Ключевые слова: изменение климата, осадки, марковские цепи, байесовская статистика

DOI: 10.31857/S2686739722010054

ВВЕДЕНИЕ

Осадки являются важнейшим компонентом климата, влияющим на природные экосистемы и хозяйственную деятельность человека [1]. В связи с наблюдаемыми в настоящее время климатическими изменениями неоднократно отмечено изменение режима осадков, заключающееся, в том числе, в изменении их объема за определенный период времени, а также в изменении числа дней с осадками и без осадков, изменении характера и интенсивности осадков (ливневые или обложные) и др. [2, 3]. Важнейшей характеристикой режима осадков в том или ином географическом локусе является средняя продолжительность непрерывных периодов с осадками и без осадков. Для анализа и моделирования этого режима часто применяются марковские цепи: достаточно простые, но эффективные модели описания дискретных процессов, в ходе которых исследуемая система в один момент времени может принимать лишь одно из конечного множества возможных состояний, а вероятность перехода в некоторое состояние в следующий момент времени определяется состояниями системы в один или несколько предыдущих моментов и только ими. Число таких моментов определяет порядок марковской цепи [4–6].

Очевидно, что метеорологическая система может характеризоваться одним из двух состояний: отсутствием (состояние 0) или наличием (состояние 1) осадков в некоторый период времени. Обычно в качестве такого периода берутся сутки. В ряде исследований было показано, что наибольшей точностью при описании режима осадков обладают марковские цепи второго порядка. Однако марковские цепи первого порядка также демонстрируют высокую точность, обладая при этом простотой структуры [7]. Марковские цепи первого порядка с двумя возможными состояниями системы 0 или 1 характеризуются четырьмя переходами состояний: 0→0, 0→1, 1→1, 1→0 и соответствующими четырьмя вероятностями перехода p, причем $p_{0\to0} + p_{0\to1} = 1$ и $p_{1\to1} + p_{1\to0} = 1$.

Вероятности однородных переходов (0→0 и 1→1) при описании режима осадков являются важными характеристиками климата, определяющими продолжительность непрерывных периодов с осадками или без них [8]. Их изменения во

[1] *Институт глобального климата и экологии им. Ю.А. Израэля, Москва, Россия*
[2] *Институт географии Российской академии наук, Москва, Россия*
**E-mail: igor_o_popov@mail.ru*
***E-mail: en_popova@igras.ru*

времени составляют важный компонент климатических изменений. При этом недостаточно фиксации только наблюдаемых частотных изменений. Для заключений о достоверных изменениях вероятностей переходов или об отсутствии таковых необходим статистический анализ.

Цель данной работы – статистическое исследование достоверности изменений вероятностей однородных переходов в марковской цепи как модели временнóго режима осадков суточного разрешения на территории России к современному периоду 2000–2019 гг. по сравнению с периодом 1966–1985 гг., соответствующим началу глобального изменения климата, и исследование пространственного распределения этих изменений. Для решения этой задачи были применены методы байесовского анализа, использование которых в данном контексте не было обнаружено в проанализированной литературе.

МАТЕРИАЛЫ И МЕТОДЫ

Байесовская статистика в последние десятилетия бурно развивается, предоставляя высокоэффективные и гибкие методы решения задач статистического вывода [9]. Основой данного подхода является применение теоремы Байеса для заключения о характере распределения вероятности значений анализируемой величины (апостериорное распределение) на основе предположений о его распределении без учета наблюдаемых данных (априорное распределение) и распределении наблюдаемой переменной, зависящей от анализируемого параметра (функция правдоподобия) [10].

Предположим наличие марковского процесса первого порядка с двумя состояниями системы 0 и 1 и соответствующей матрицей вероятностей переходов. Число n наблюдаемых однородных переходов $x \to x$ среди N переходов $x \to \{0.1\}$ при вероятности перехода $p_{x\to x}$ можно считать распределенным по закону биномиального распределения:

$$n \sim \mathrm{Binom}(n, N, p_{x\to x})$$

Данное распределение будем использовать в качестве функции правдоподобия. В качестве априорного распределения величины $p_{x\to x}$ можно взять бета-распределение:

$$p_{x\to x,\mathrm{prior}} \sim \mathrm{Beta}(\alpha_{\mathrm{prior}}, \beta_{\mathrm{prior}})$$

Известно, что биномиальное распределение в качестве функции правдоподобия и бета-распределение в качестве априорного распределения являются сопряженными распределениями, порождающими апостериорное бета-распределение анализируемой величины с параметрами, строго выводимыми из значений n, N, α_{prior} и β_{prior} [11, 12]:

$$p_{x\to x,\mathrm{post}} \sim \mathrm{Beta}(\alpha_{\mathrm{post}} = \alpha_{\mathrm{prior}} + n,$$
$$\beta_{\mathrm{post}} = \beta_{\mathrm{prior}} + N - n)$$

Обычная практика в байесовском анализе – использовать в качестве априорного равномерное распределение, являющееся частным случаем бета-распределения с параметрами $\alpha_{\mathrm{prior}} = \beta_{\mathrm{prior}} = 1$. Тогда параметры апостериорного бета-распределения приобретают значения:

$$\alpha_{\mathrm{post}} = n + 1 \quad (1)$$

$$\beta_{\mathrm{post}} = N - n + 1 \quad (2)$$

Формулы (1) и (2) в рамках представленной статистической модели позволяют напрямую вычислять параметры апостериорного распределения, не прибегая к современным вычислительным методам [13].

В качестве источника метеоданных использовалась база данных метеонаблюдений суточного разрешения TTTR, созданная во Всероссийском научно-исследовательском институте гидрометеорологической информации – Мировом центре данных (ВНИИГМИ-МЦД, http://www.meteo.ru). Были отобраны 451 станция, расположенные на территории России, данные по осадкам с которых имели менее 10% пропусков за оба исследуемые периода.

Каждые сутки были охарактеризованы значениями 0 или 1, соответственно, отсутствию или наличию в них осадков. Было рассчитано число четырех видов переходов из одного состояния в такое же или другое для обоих периодов. Для предварительной оценки происшедших изменений были рассчитаны частоты переходов $0\to0$ относительно событий $0\to X$ и $1\to1$ относительно события $1\to X$, $X \in \{0, 1\}$, представляющие собой эмпирические оценки исследуемых вероятностей $p_{0\to0}$ и $p_{1\to1}$.

По формулам 1 и 2 для каждой станции были рассчитаны параметры апостериорных бета-распределений значений вероятностей двух однородных переходов для обоих периодов 1966–1985 и 2000–2019 гг. С целью оценки достоверности изменения вероятностей однородных переходов во втором периоде по сравнению с первым была осуществлена оценка вероятности превосходства выборки из апостериорного распределения значений оцениваемых вероятностей для второго периода над распределением для первого периода [10]. Для этого были сгенерированы выборки случайных величин {*p} из всех четырех бета-распределений объемом 5×10^5 экземпляров каждая. Полученные случайные величины объединялись в пары отдельно для каждого вида перехода. Рассчитывалась частота случаев $*p_{(\text{период 2})} > *p_{(\text{период 1})}$. Если эта частота была более 0.975, делалось заключение о достоверном увеличении значения вероятности однородного перехода $p_{x\to x}$ ($x \in \{0, 1\}$)

Таблица 1. Статистические характеристики наблюдаемых частот переходов суточных режимов осадков 0→0 и 1→1 в периоды 1966–1985 и 2000–2019 гг. и их постанционных изменений

Тип перехода	Период	Среднее значение	Стандартное отклонение
0→0	1966–1985	0.702	0.069
	2000–2019	0.709	0.069
	изменение	0.007	0.019
1→1	1966–1985	0.598	0.084
	2000–2019	0.59	0.088
	изменение	–0.009	0.026

во втором периоде по сравнению с первым. Если частота была менее 0.025, делался вывод о достоверном уменьшении значения вероятности однородного перехода. В остальных случаях делался вывод о недостоверности изменений.

В результате этих исследований каждая станция была охарактеризована двумя признаками, отражающими достоверность и направление изменения величин $p_{0\to0}$ и $p_{1\to1}$. Эти признаки имели одно из значений: “изменение недостоверно”, “достоверное уменьшение”, “достоверное увеличение”. Очевидно, что всего может быть девять комбинаций этих признаков для обоих видов вероятности.

Для визуализации пространственного распределения анализируемых изменений для каждой из вероятностей $p_{0\to0}$ и $p_{1\to1}$ были построены две географические карты, на которых кружками отмечены локализации станций, при этом цвет кружка отображает достоверность и направление изменений.

С помощью диаграммы (полигонов) Вороного [14] была осуществлена интерполяция полученных для станций меток на всю территорию Российской Федерации. Каждый полигон на картах закрашивался цветом, соответствующим метке его станции.

Для оценки совместного пространственного распределения характера изменения обоих видов вероятностей была построена карта с полигонами Вороного, цвет которых отображает характер изменений вероятностей $p_{0\to0}$, а вид штриховки – вероятностей $p_{1\to1}$.

Для оценки неслучайности пространственного распределения меток над множеством локализаций станций в географическом пространстве была проведена триангуляция Делоне [15]. В результате каждая станция была связана с ближайшими станциями двумя или более отрезками (“связями”). Каждая связь может быть охарактеризована как однородная, если соединенные ею станции имеют одинаковые метки, или как разнородная в противоположном случае. Очевидно, что при случайном пространственном распределении меток доля однородных связей меньше по сравнению со случаем пространственной кластеризации станций с одинаковыми метками. Для проверки гипотезы о неслучайном распределении меток был проведен пермутационный тест, в каждом цикле которого метки случайным образом перемешивались между станциями и высчитывалась доля однородных связей. Всего было проведено 5×10^4 циклов. В результате было создано эмпирическое распределение доли однородных меток и определен его 95%-й процентиль. В случае, если наблюдаемое значение доли однородных связей больше данного значения 95%-го процентиля, гипотеза о случайном характере наблюдаемого распределения меток отвергается. Тест проводился как отдельно для переходов 0→0 и 1→1, так и для их совместного пространственного распределения.

Подготовка данных и расчеты проводились с помощью программ, написанных на языке Python с применением библиотек numpy, Pandas, scipy. Построение полигонов Вороного, триангуляция Делоне и отрисовка карт осуществлялись в ГИС QGIS.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В табл. 1 приведены статистические характеристики (среднее арифметическое и стандартное отклонение) наблюдаемых частот переходов режима осадков типа 0→0 и 1→1 в исследуемые периоды и их изменений. Из этих данных видно, что во второй период среднее значение наблюдаемых частот переходов 0→0 несколько увеличилось по сравнению с первым периодом, а частота переходов 1→1 несколько уменьшилась.

На рис. 1 представлены гистограммы наблюдаемых частот однородных переходов режима осадков и графики ядерных оценок плотности их распределений для двух периодов. Согласно этому

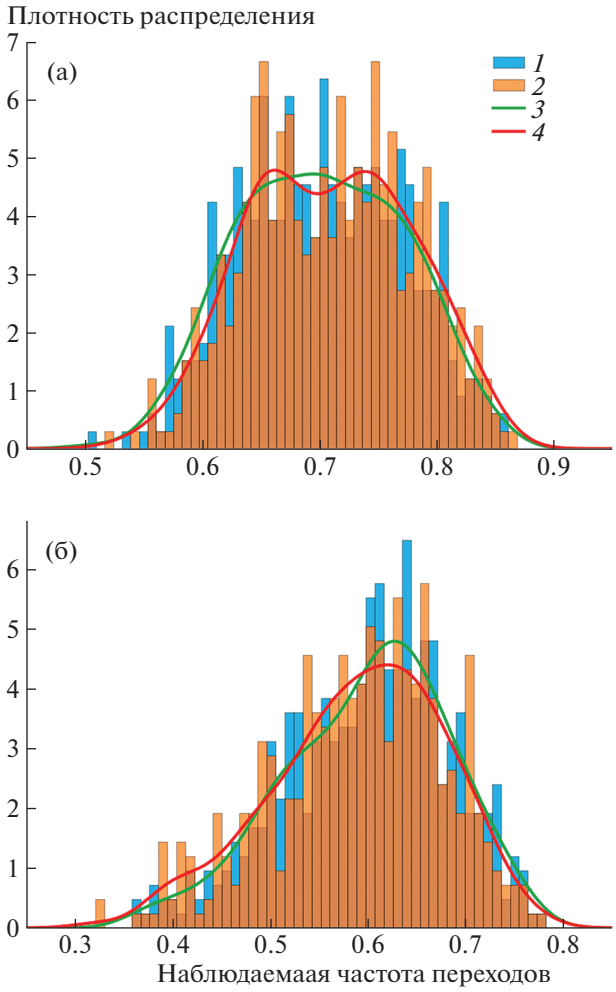

Рис. 1. Плотность распределения значений наблюдаемых частот переходов 0→0 (а) и 1→1 (б). Обозначения: 1 и 2 – гистограммы значений наблюдаемых частот переходов за периоды 1966–1985 гг. (1) и 2000–2019 гг. (2), 3 и 4 – графики ядерной оценки плотности распределения наблюдаемых частот переходов за периоды 1966–1985 гг. (3) и 2000–2019 гг. (4).

рисунку, график плотности распределения частот переходов 0→0 ко второму исследуемому периоду несколько сместился в сторону бóльших значений (рис. 1а), а для плотности распределения частот переходов 1→1 заметно небольшое смещение в сторону уменьшения (рис. 1б), что совпадает с данными, представленными в табл. 1.

В табл. 2 приведено количество станций, характеризующихся различными оценками достоверности и типа изменений вероятностей переходов $p_{0\to0}$ $p_{1\to1}$. Как можно заметить, для большинства станций (58% от общего числа станций) сделан вывод об отсутствии статистически достоверного изменения исследуемых вероятностей переходов. Вероятность $p_{0\to0}$ достоверно уменьшилась на 23 станциях (5%) и достоверно повысилась на 94 станциях (21%). При этом вероятность $p_{1\to1}$ достоверно уменьшилась на 111 станциях (25%) и увеличилась на 42 станциях (9%). Станций, на которых происходит одновременное увеличение вероятности $p_{0\to0}$ и уменьшение вероятности $p_{1\to1}$, отмечено 68 (15%), что более чем в шесть раз превышает число станций с противоположной характеристикой (совместное понижение $p_{0\to0}$ и увеличение $p_{1\to1}$) – 11 (2.5%). При этом не обнаружено станций с одновременным повышением или одновременным понижением обеих вероятностей.

Из полученных результатов можно сделать вывод о наличии преимущественной тенденции к увеличению вероятности $p_{0\to0}$ и уменьшению вероятности $p_{1\to1}$, что может свидетельствовать об общей тенденции к удлинению периодов без осадков и укорочению периодов непрерывных осадков. При этом в отдельных регионах могут наблюдаться иные, в том числе противоположные, тенденции. В целом эти выводы, сделанные в результате проведения байесовской оценки, согласуются с данными описательного статистического анализа наблюдаемых частот однородных переходов, приведенных в табл. 1, и оценки плотности распределения частот переходов на рис. 1.

На рис. 2а и 2б приведены карты пространственного распределения достоверности и характера изменения вероятностей переходов $p_{0\to0}$ и $p_{1\to1}$. На рис. 2в представлено совместное распределение видов изменения вероятностей обоих типов перехода.

Таблица 2. Число станций метеонаблюдений, характеризующихся различными байесовскими оценками достоверности и направления изменения вероятностей переходов $p_{0\to0}$ и $p_{1\to1}$

Изменение $p_{1\to1}$	Изменение $p_{0\to0}$			
	изменение недостоверно	достоверное уменьшение	достоверное увеличение	всего
изменение недостоверно	260	12	26	298
достоверное уменьшение	43	0	68	111
достоверное увеличение	31	11	0	42
всего	334	23	94	451

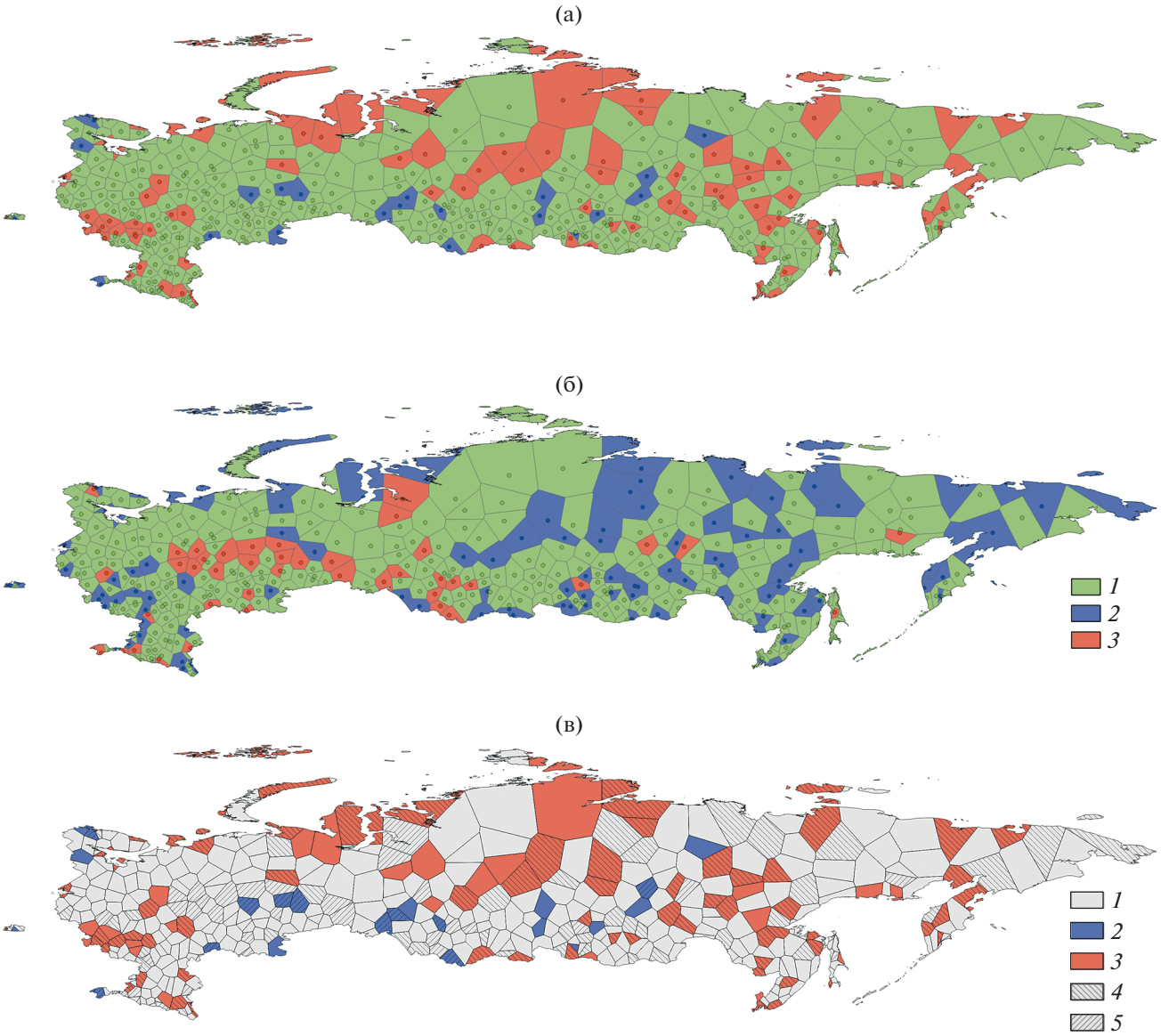

Рис. 2. Пространственное распределение видов достоверности и направления переходов вероятностей $p_{0\to0}$ (а), $p_{1\to1}$ (б) и их совместного распространения (в) на территории России. Обозначения а) и б): *1* – отсутствие достоверного изменения, *2* – достоверное уменьшение, *3* – достоверное увеличение. Обозначения в): *1* – изменение $p_{0\to0}$ недостоверно, *2* – достоверное уменьшение $p_{0\to0}$, *3* – достоверное увеличение $p_{0\to0}$, *4* – достоверное уменьшение $p_{1\to1}$, *5* – достоверное увеличение $p_{1\to1}$, отсутствие достоверного изменения $p_{1\to1}$ показано отсутствием штриховки.

Анализ этих карт позволяет предположить неслучайный характер распределения полигонов с достоверными изменениями вероятностей переходов. Полигоны одного типа образуют агломерации разного размера, что свидетельствует о распространении соответствующих климатических изменений на достаточно протяженные территории.

Неслучайный характер распределения изменений по территории России доказывают и результаты пермутационного теста, приведенные в табл. 3. Из них видно, что во всех трех случаях наблюдаемая доля однородных связей была значительно больше 95%-го процентиля экспериментального распределения, что позволяет отвергнуть гипотезу о случайном пространственном распределении типов изменений между станциями.

Достоверное увеличение $p_{0\to0}$ и одновременное понижение $p_{1\to1}$ наблюдаются в центральных и юго-западных районах Европейской части России. Противоположная тенденция выявлена в Крыму и на некоторой части Кольского полуострова. Значительная часть Азиатской территории России также характеризуется областями с достоверными повышениями $p_{0\to0}$ и понижениями

Таблица 3. Результаты пермутационного теста неслучайности пространственного распределения типов изменения вероятностей переходов

Вид перехода	Наблюдаемая доля однородных связей	95% процентиль экспериментального распределения доли однородных связей
0→0	0.579	0.547
1→1	0.593	0.537
совместное распределение	0.41	0.364

$p_{1\to1}$. При этом области с достоверными понижениями вероятности $p_{0\to0}$ и повышениями $p_{1\to1}$ образуют прерывистый пояс, тянущийся от центральной части Уральских гор через Южную Сибирь до центральной части Восточной Сибири. Эти наблюдения могут свидетельствовать о происшедшем в этих регионах увеличении продолжительности непрерывных периодов осадков и сокращении длины непрерывных последовательностей дней без осадков.

Представляется очевидным, что закономерности этого пространственного распределения, так же, как и его связь с изменениями других климатических параметров, требуют дальнейшего изучения.

ИСТОЧНИКИ ФИНАНСИРОВАНИЯ

Анализ результатов выполнен за счет средств темы НИР по Плану Фундаментальных научных исследований государственных академий наук № 0148-2019-0007, АААА-А19-119021990093-8 “Оценка физико-географических, гидрологических и биотических изменений окружающей среды и их последствий для создания основ устойчивого природопользования”. Статистическое моделирование и визуализация результатов осуществлены за счет средств плана НИТР на 2020–2024 гг. Росгидромета, тема 3.2 “Мониторинг глобального климата и климата Российской Федерации и ее регионов, включая Арктику. Развитие и модернизация технологий мониторинга”.

СПИСОК ЛИТЕРАТУРЫ

1. *Strangeways I.* Precipitation. Theory, Measurement and Distribution. Cambridge University Press. 2007. https://doi.org/10.1017/CBO9780511535772
2. *Popova E.N., Popov I.O., Semenov S.M.* Assessment of Variations in the Annual Sum of Active Temperatures and Total Precipitation during the Vegetation Period in Russia and Neighboring Countries // Russian Meteorology and Hydrology. 2018. V. 43. № 6. P. 412–417. https://doi.org/10.3103/S1068373918060092
3. *Chernokulsky A., Kozlov F., Zolina O., Bulygina O., Mokhov I., Semenov V.* Observed Changes in Convective and Stratiform Precipitation in Northern Eurasia over the Last Five Decades // Environmental Research Letters. 2019. V. 14. № 4. 45001. https://doi.org/10.1088/1748-9326/aafb82
4. *Турчин В.Н., Турчин Е.В.* Марковские цепи. Днепропетровск: LuzinoffPress. 2016.
5. *Lennartssona J., Baxevania A., Chenb D.* Modelling Precipitation in Sweden Using Multiple Step Markov Chains and a Composite Model // Journal of Hydrology. 2008. V. 363. Iss. 1–4. P. 42–59. https://doi.org/10.1016/j.jhydrol.2008.10.003
6. *Wilks D.S.* Statistical Methods in the Atmospheric Sciences. 3rd Edition. International Geophysics Series (V. 100). Elsevier, Academic Press. 2011. https://doi.org/10.1016/B978-0-12-385022-5.00015-4
7. *Schoof J.T., Pryor S.C.* On the Proper Order of Markov Chain Model for Daily Precipitation Occurrence in the Contiguous United States // Journal of Applied Meteorology and Climatology. 2008. V. 47. Iss. 9. P. 2477–2486. https://doi.org/10.1175/2008JAMC1840.1
8. *Yeh H.-F., Hsu H.-L.* Using the Markov Chain to Analyze Precipitation and Groundwater Drought Characteristics and Linkage with Atmospheric Circulation // Sustainability. 2019. V. 11. Nr 6. Article Number 1817. https://doi.org/10.3390/su11061817
9. *Sprenger J., Hartmann S.* Bayesian Philosophy of Science. Oxford University Press; 2019. https://doi.org/10.1093/oso/9780199672110.001.0001
10. *Bradley C.P.* Bayes and Empirical Bayes Methods for Data Analysis. Boca Raton: Chapman and Hall/CRC. 2000.
11. *Watanabe S.* Mathematical Theory of Bayesian Statistics. Boca Raton: Chapman and Hall/CRC. 2018. https://doi.org/10.1201/9781315373010
12. *Stone J.V.* Bayes’ Rule: A Tutorial Introduction to Bayesian Analysis. Sebtel Press. 2013. https://doi.org/10.13140/2.1.1371.6801
13. *Martin O.* Bayesian Analysis with Python. Packt Publishing. 2016.
14. *Preparata F., Shamos M.* Computational Geometry: An Introduction. Springer. 1985.
15. *Скворцов А.В.* Триангуляция Делоне и ее применение. Томск: Издательство Томского университета. 2002.

ANALYSIS OF CHANGES IN THE PRECIPITATION REGIME ON THE TERRITORY OF RUSSIAN FEDERATION IN THE SECOND HALF OF THE 20th AND IN THE EARLY 21st CENTURY WITH THE USAGE OF THE BAYESIAN ESTIMATION OF THE MARKOV CHAIN PARAMETERS

I. O. Popov[a, #] and E. N. Popova[b, ##]

[a] *Yu.A. Izrael Institute of Global Climate and Ecology, Moscow, Russian Federation*

[b] *Institute of Geography, Russian Academy of Sciences, Moscow, Russian Federation*

[#] *E-mail: igor_o_popov@mail.ru*

[##] *E-mail: en_popova@igras.ru*

Presented by Academician of the RAS V.M. Kotlyakov September 23, 2021

A Bayesian assessment of the change in the probabilities of homogeneous transitions in a binary first-order Markov chain as a model of changes in the temporal regime of precipitation in Russia in 2000–2019 compared with 1966–1985 has been carried out. A significant change in these probabilities is shown over a significant part of the territory of Russia, and in most cases, there is an increase in the probability of homogeneous transitions 0→0 and a decrease in the probability of homogeneous transitions 1→1, which may mean an increase in the duration of continuous periods without precipitation and a decrease in continuous periods with precipitation. Maps of the spatial distribution of these changes have been constructed and its non-randomness has been statistically proven.

Keywords: climate change, precipitation, Markov chains, Bayesian statistics