

Стохастические модели временных рядов биоклиматических индексов в Арктической зоне Российской Федерации

М. С. Акентьева¹, Н. А. Каргаполова^{1}*

¹ Институт вычислительной математики и математической геофизики СО РАН,
г. Новосибирск, Российская Федерация
* e-mail: nkargapolova@gmail.com

Аннотация. В статье рассмотрен опирающийся на применении определяющих формул подход к численному стохастическому моделированию временных рядов двух биоклиматических индексов: индекса холодного стресса и эквивалентно-эффективной температуры на метеостанциях, расположенных в Арктической зоне Российской Федерации. На основе рассмотренного подхода построены, численно реализованы и верифицированы имитационные стохастические модели временных рядов рассмотренных биоклиматических индексов. Модели построены в предположении о периодической коррелированности реальных биоклиматических процессов. Использование данного предположения позволяет учесть присущий им суточный ход. Использовано также предположение о негауссовости реальных рядов индекса холодного стресса и эквивалентно-эффективной температуры. Для моделирования нестационарных негауссовских рядов применён метод обратных функций распределения. Результаты верификации разработанных стохастических моделей показали, что модельные траектории обладают многими статистическими характеристиками близкими к соответствующим характеристикам реальных рядов.

Ключевые слова: индекс холодного стресса, эквивалентно-эффективная температура, стохастическое моделирование, негауссовский нестационарный случайный процесс Арктическая зона Российской Федерации

Stochastic models of bioclimatic indices time series in the Arctic zone of the Russian Federation

M. S. Akenteva¹, N. A. Kargapolova^{1}*

¹ Institute of Computational Mathematics and Mathematical Geophysics SB RAS, Novosibirsk,
Russian Federation
* e-mail: nkargapolova@gmail.com

Abstract. In this paper, an approach to the numerical stochastic simulation of time series of two bioclimatic indices (wind chill index and the equivalent-effective temperature) at weather stations located in the Arctic zone of the Russian Federation is considered. The approach is based on the use of so-called “defining formulas”. On the basis of the considered approach, stochastic models of the time series of the considered bioclimatic indices were constructed, numerically implemented and verified. The models are developed on the assumption that the real bioclimatic processes are periodically correlated. The use of this assumption makes it possible to take into account the daily variation of the real processes. We also use the assumption that the real series of the wind chill index and the equivalent-effective temperature are non-Gaussian. To simulate non-stationary non-Gaussian time series, the method of inverse distribution functions is applied. The results of verification of the developed stochastic models showed that many statistical characteristics of simulated trajectories are close to the corresponding characteristics of the real series.

Keywords: wind chill index, equivalent-effective temperature, stochastic simulation, non-Gaussian nonstationary random process, Arctic zone of Russian Federation

Введение

Для описания теплового влияния на человека различных комбинаций метеорологических параметров часто используют биоклиматические индексы (БИ). Их применяют, например, в работе оперативных систем предупреждения о наступлении некомфортных погодных условий [1] и для медико-климатической классификации территорий [2]. Для решения таких задач необходимо, среди прочего, исследовать характеристики различных редких неблагоприятных биоклиматических явлений. Провести такое исследование по реальным данным не всегда представляется возможным в силу ограниченного (и, часто, малого) объема выборки. Однако такие исследования можно провести, используя модельные значения БИ.

Существует 3 основных подхода к моделированию временных рядов, пространственных и пространственно-временных полей БИ: статистический [3], динамический [4] и стохастический [5, 6]. В данной работе используется последний.

Целью данной работы является разработка, численная реализация и верификация имитационных стохастических моделей временных рядов двух БИ, а именно, индекса холодового стресса (ИХС) и эквивалентно-эффективной температуры (ЭЭТ). Предложенные модели должны воспроизводить с высокой точностью основные статистические характеристики реальных рядов биоклиматических индексов на метеостанциях, расположенных на территории Арктической зоны Российской Федерации (АЗРФ).

Отметим, что стохастический подход к моделированию ЭЭТ применяется впервые. Численные стохастические модели временных рядов ИХС ранее разрабатывались при участии авторов, однако не верифицировались по данным из АЗРФ.

Реальные данные

Обозначим

$$\begin{aligned}\bar{W}^T &= (W_1, W_2, \dots, W_{8ND})^T = (W_{1,1}, \dots, W_{8,1}, W_{1,2}, \dots, W_{8,ND})^T, \\ \bar{E}^T &= (E_1, E_2, \dots, E_{8ND})^T = (E_{1,1}, \dots, E_{8,1}, E_{1,2}, \dots, E_{8,ND})^T\end{aligned}$$

моделируемые ряды ИХС и ЭЭТ, соответственно. Здесь и всюду далее, $i = 1, 2, \dots, 8$ – срок измерения внутри суток, $j = 1, 2, \dots, ND$ – номер дня на рассматриваемом временном интервале длительностью ND дней.

В качестве исходных данных при построении моделей используются данные восьмисрочных наблюдений на метеорологических станциях, расположенных в АЗРФ [7]. Период наблюдения – с 1992 по 2020 г. По данным наблюдений

за скоростью ветра $V_{i,j}^k$ м / с на высоте 10 м над уровнем земли, приземной температурой $T_{i,j}^k$ °С и относительной влажностью воздуха $H_{i,j}^k$ % вычисляются реальные значения ИХС [8, 9]

$$W_{i,j}^k = \begin{cases} 13.12 + 0.6215T_{i,j}^k - 11.37(\tilde{V}_{i,j}^k)^{0.16} + 0.3965T_{i,j}^k(\tilde{V}_{i,j}^k)^{0.16}, & \text{если } \tilde{V}_{i,j}^k \geq 5, \\ T_{i,j}^k + \frac{-1.59 + 0.1345T_{i,j}^k}{5}\tilde{V}_{i,j}^k, & \text{если } \tilde{V}_{i,j}^k < 5, \end{cases}$$

и ЭЭТ [10]

$$E_{i,j}^k = 37 - \frac{37 - T_{i,j}^k}{0.68 - 0.0014\hat{H}_{i,j}^k + \left(1.76 + 1.4(\hat{V}_{i,j}^k)^{0.75}\right)^{-1}} - 0.29T_{i,j}^k(1 - H_{i,j}^k),$$

где

$$\hat{H}_{i,j}^k = 100H_{i,j}^k, \hat{V}_{i,j}^k = \frac{2}{3}V_{i,j}^k, \tilde{V}_{i,j}^k = 3.6V_{i,j}^k, k = 1992, 1993, \dots, 2020.$$

По определению, единица измерения ИХС и ЭЭТ – градус Цельсия (°С).

Необходимо отметить специфику рассматриваемых данных. В середине зимы, суточный ход временных рядов БИ выражен неявно. Однако, в другие периоды времени суточный ход заметен (в том числе, в летнее время, когда наблюдается полярный день) В качестве примера, в табл. 1 приведены оценки $EW_i, EE_i, i = 1, \dots, 8$ математического ожидания БИ, полученные по реальным данным (здесь и всюду далее, σ – оценка статистической погрешности рассматриваемой характеристики).

Стохастическая модель

Необходимо смоделировать временные ряды \vec{W}^T, \vec{E}^T . Моделирование этих рядов будем осуществлять в рамках метода обратных функций распределения (МОФР) [11]. Для того, чтобы применить этот подход, необходимо определить плотности одномерных распределений (ПОР) моделируемых величин и корреляционные матрицы временных рядов.

Для того чтобы учесть суточный ход реальных временных рядов ИХС и ЭЭТ, используем предположение о периодической коррелированности векторов \vec{W}^T, \vec{E}^T . Это эквивалентно предположению, что векторные последователь-

ности $\vec{W}_1, \vec{W}_2, \dots, \vec{W}_{ND}$ и $\vec{E}_1, \vec{E}_2, \dots, \vec{E}_{ND}$, где $\vec{W}_j^T = (W_{1,j}, \dots, W_{8,j})^T$, $\vec{E}_j^T = (E_{1,j}, \dots, E_{8,j})^T$, стационарны.

Таблица 1

Оценки математического ожидания ИСХ и ЭЭТ

Срок i	EW_i		EE_i	
	Индига, 1-10 декабря	Архангельск, 1-10 февраля	Мурманск, 1-10 июля	Архангельск, 16-26 июня
1	-16.38 ± 1.13	-11.78 ± 0.66	4.70 ± 0.51	5.74 ± 0.50
2	-16.70 ± 1.17	-12.11 ± 0.66	5.11 ± 0.53	7.41 ± 0.51
3	-16.69 ± 1.15	-12.05 ± 0.68	6.77 ± 0.60	9.53 ± 0.58
4	-16.82 ± 1.12	-11.08 ± 0.67	8.48 ± 0.60	11.05 ± 0.57
5	-16.37 ± 1.08	-10.14 ± 0.68	9.34 ± 0.60	11.73 ± 0.57
6	-16.64 ± 1.08	-10.76 ± 0.67	9.06 ± 0.60	11.74 ± 0.58
7	-16.78 ± 1.01	-11.12 ± 0.66	8.06 ± 0.59	10.68 ± 0.54
8	-16.92 ± 1.11	-11.42 ± 0.67	6.38 ± 0.55	7.65 ± 0.51

В качестве корреляционных матриц R_W, R_E рядов \vec{W}^T, \vec{E}^T будем использовать выборочные корреляционные матрицы, полученные по реальным данным $W_{i,j}^k, E_{i,j}^k$. Отметим, что в силу использованного предположения о периодической коррелированности, матрицы R_W, R_E являются блочно-тёплицевыми с блоками размерности 8×8 , а объём выборки, по которой оценивается каждый коэффициент корреляции, равен $ND \times NY$, где NY – количество лет, в которые проводились наблюдения (в данной работе, $NY = 29$).

Перейдём к определению ПОР ИХС и ЭЭТ. Поскольку последовательности $\vec{W}_1, \vec{W}_2, \dots, \vec{W}_{ND}$ и $\vec{E}_1, \vec{E}_2, \dots, \vec{E}_{ND}$ стационарны, достаточно определить ПОР $fw_{i,1}(x), fe_{i,1}(x)$, $i = 1, \dots, 8$ первых восьми компонент векторов \vec{W}^T, \vec{E}^T , соответственно. Оставшиеся ПОР определяются равенствами $fw_{i,j}(x) = fw_{i,1}(x)$, $fe_{i,j}(x) = fe_{i,1}(x)$, $i = 1, \dots, 8$, $j = 2, \dots, ND$.

Для построения стохастической модели нецелесообразно использовать выборочные ПОР (гистограммы). Действительно, гистограммы определены лишь на конечных интервалах (т.е. не имеют хвостов). Если их использовать при моделировании, то среди модельных значений будут отсутствовать те значения биоклиматических индексов, которые физически возможны, но не попали в выборку реальных данных в силу ограниченности периода наблюдений. В связи с этим, следует аппроксимировать выборочные ПОР какими-либо аналитически

заданными плотностями, которые с одной стороны, не сильно искажают вид и моменты одномерного распределения, а с другой – обладают хвостами.

Проведённый статистический анализ данных $W_{i,j}^k, E_{i,j}^k$ показал, что в качестве аппроксимирующих плотностей $fw_{i,1}(x), fe_{i,1}(x)$ могут быть использованы смеси нормальных распределений вида

$$\theta_i \frac{1}{b_{i1} \sqrt{2\pi}} \exp\left(-\frac{(x-a_{i1})^2}{2b_{i1}^2}\right) + (1-\theta_i) \frac{1}{b_{i2} \sqrt{2\pi}} \exp\left(-\frac{(x-a_{i2})^2}{2b_{i2}^2}\right),$$

параметры $\theta_i, a_{i1}, b_{i1}^2, a_{i2}, b_{i2}^2$ которых для каждого срока $i = 1, \dots, 8$ и рассматриваемого биоклиматического индекса выбираются с помощью EM-алгоритма [12].

Как уже было сказано выше, моделирование временных рядов \vec{W}^T, \vec{E}^T с заданными ПОР $fw_{i,j}(x), fe_{i,j}(x)$ и корреляционными матрицами R_W, R_E , соответственно, осуществляется методом обратных функций распределения. Заметим, что благодаря использованию предположения о периодической коррелированности рядов, вычисление корреляционной матрицы вспомогательного гауссовского процесса при реализации МОФР не является трудоёмким. Моделирование гауссовского процесса проводилось методом, основанном на спектральном разложении корреляционных матриц [11].

Результаты верификации

Для верификации предложенной модели временных рядов \vec{W}^T, \vec{E}^T было проведено сравнение оценок различных статистических характеристик рассматриваемых рядов, не являющихся входными параметрами моделей, по реальным (РД) и модельным данным (МД). Ниже приведены несколько примеров характеристик, использованных для верификации. Для оценок по МД использовалось 10^6 модельных траекторий.

В качестве первого примера, в табл. 2 и 3 приведены оценки вероятностей

$$P_W(lev) = P(\exists l \in \{1, 2, \dots, ND\} : \tilde{W}_l < lev), \quad \tilde{W}_l = \frac{1}{8} \sum_{j=1}^8 W_{8(l-1)+j},$$

$$P_E(lev) = P(\exists l \in \{1, 2, \dots, ND\} : \tilde{E}_l > lev), \quad \tilde{E}_l = \frac{1}{8} \sum_{j=1}^8 E_{8(l-1)+j}$$

того, что среднесуточное значение ИХС было ниже и среднесуточное значение ЭЭТ было выше заданного уровня lev^oC хотя бы единожды на рассматриваемом

временном интервале, соответственно. Численные эксперименты показали, что оценки вероятностей $P_W(lev)$ и $P_E(lev)$ по МД отличаются от оценок по РД не более, чем на 3σ в 97% и 94% случаях, соответственно, т.е. рассмотренные характеристики хорошо воспроизводятся моделью.

Таблица 2

Оценки вероятностей $P_W(lev)$. Янискоски, 1-10 декабря

$lev^{\circ}C$	РД, $P_W(lev) \pm \sigma$	МД
-15	0.79 ± 0.07	0.83
-16	0.75 ± 0.08	0.74
-17	0.68 ± 0.09	0.68
-18	0.57 ± 0.09	0.55
-19	0.50 ± 0.09	0.48

Таблица 3

Оценки вероятностей $P_E(lev)$. Кандалакша, 1-10 июля

$lev^{\circ}C$	РД, $P_E(lev) \pm \sigma$	МД
7	1.00 ± 0.00	1.00
9	1.00 ± 0.01	0.99
11	0.93 ± 0.04	0.96
13	0.79 ± 0.07	0.77
15	0.48 ± 0.09	0.50
17	0.21 ± 0.08	0.19
19	0.00 ± 0.05	0.06
21	0.00 ± 0.05	0.05

Второй пример характеристики, использованный при верификации моделей – вероятность

$$PS_W(lev, h) = P(\exists i \in \{1, 2, \dots, 8ND + h - 1\} : W_i < lev, W_{i+1} < lev, \dots, W_{i+h-1} < lev),$$

$$PS_E(lev, h) = P(\exists i \in \{1, 2, \dots, 8ND + h - 1\} : E_i > lev, E_{i+1} > lev, \dots, E_{i+h-1} > lev),$$

того, что ИХС ниже заданного уровня $lev^{\circ}C$ (или ЭЭТ выше $lev^{\circ}C$) h измерений подряд. Эта характеристика напрямую связана с многомерными распределе-

ниями рядов \vec{W}^T, \vec{E}^T . В качестве иллюстрации, в табл. 4 приведены оценки $PS_W(lev, h)$ для некоторых уровней lev и длительностей h на двух метеостанциях. Оценки $PS_W(lev, h)$, полученные по МД для различных метеостанций, уровней lev и длительностей h , отличаются от советующих оценок по РД не более, чем на σ и 3σ в 78% и 99% проведённых численных экспериментов, соответственно. При этом для ЭЭТ, оценки $PS_E(lev, h)$ попали в аналогичные интервалы в 47% и 100% экспериментов.

Таблица 4

Оценки вероятностей $PS_W(lev, h)$. 1-10 декабря

h	$lev^{\circ}C$	Канин Нос		Полярное	
		РД, $PS_W(lev, h) \pm \sigma$	МД	РД, $PS_W(lev, h) \pm \sigma$	МД
8	-3	0.94 ± 0.04	0.91	0.83 ± 0.04	0.84
8	-6	0.68 ± 0.05	0.71	0.58 ± 0.05	0.56
8	-9	0.48 ± 0.06	0.49	0.38 ± 0.05	0.39
16	-3	0.89 ± 0.05	0.91	0.75 ± 0.05	0.76
16	-6	0.55 ± 0.07	0.53	0.46 ± 0.06	0.44
16	-9	0.37 ± 0.06	0.34	0.29 ± 0.05	0.33
24	-3	0.83 ± 0.06	0.85	0.69 ± 0.06	0.70
24	-6	0.45 ± 0.08	0.52	0.37 ± 0.06	0.35
24	-9	0.28 ± 0.06	0.24	0.19 ± 0.05	0.16
32	-3	0.78 ± 0.08	0.81	0.64 ± 0.07	0.66
32	-6	0.36 ± 0.08	0.43	0.30 ± 0.06	0.29
32	-9	0.19 ± 0.08	0.18	0.15 ± 0.06	0.13
40	-3	0.72 ± 0.09	0.74	0.58 ± 0.08	0.60
40	-6	0.30 ± 0.09	0.35	0.22 ± 0.06	0.19
40	-9	0.15 ± 0.08	0.21	0.11 ± 0.05	0.13

Заключение

Как показала верификация, предложенная имитационная модель временных рядов ИХС и ЭЭТ воспроизводит с высокой точностью многие статистические свойства реальных рядов биоклиматических индексов. В качестве основных направлений дальнейших исследований в данной области укажем следующие:

1. разработка условных моделей временных рядов биоклиматических индексов и построение вероятностного прогноза на их основе,
2. разработка стохастических моделей безусловных и условных пространственных и пространственно-временных полей ИХС и ЭЭТ на сети метеостанций и произвольных пространственно-временных сетках.

Построение таких моделей позволит провести оценки вероятностей возникновения неблагоприятных (с биоклиматической точки зрения) явлений и исследовать их характеристики при различных погодных сценариях.

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 21-71-00007, <https://rscf.ru/project/21-71-00007/>.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Lowe D., Ebi K.L., and Forsberg B. Heatwave Early Warning Systems and Adaptation Advice to Reduce Human Health Consequences of Heatwaves. – *Int. J. Environ. Res. Public Health.*, 2011, vol. 8., pp. 4623–4648.
2. Гомбоева Н. Г. Климато-рекреационные ресурсы Забайкалья и их влияние на здоровье населения: автореф. дис. ... канд. биол. наук: 14.00.17. / Гомбоева Нина Гындуновна. – М., 1997. – 16 с.
3. Переведенцев Ю. П., Шумихина А. В. Динамика биоклиматических показателей комфортности природной среды в Удмуртской Республике. – *Ученые записки Казанского университета. Серия Естественные науки*, 2016, т. 158, № 4, с. 531–547.
4. Gadzhev G., and Ganev K. Computer Simulations of Air Quality and Bio-Climatic Indices for the City of Sofia. – *Atmosphere*, 2021, vol. 12, No 8, 1078.
5. Kargapolova N. A., Khlebnikova E. I., and Ogorodnikov V. A. Numerical study of properties of air heat content indicators based on the stochastic model of the meteorological processes. – *Russ. J. Num. Anal. Math. Modelling.*, 2019, vol. 34, No 2, pp. 95–104.
6. Sirangelo B., Caloiero T., Coscarelli R., et al. Combining stochastic models of air temperature and vapour pressure for the analysis of the bioclimatic comfort through the Humidex. – *Sci. Rep.*, 2020, vol. 10, 11395.
7. Булыгина О. Н., Веселов В. М., Разуваев В. Н., Александрова Т. М. Описание массива срочных данных об основных метеорологических параметрах на станциях России. / Свидетельство о государственной регистрации базы данных № 2014620549. <http://meteo.ru/data/163-basic-parameters#описание-массива-данных>
8. Mekis É., Vincent L. A., Shephard M. W., and Zhang X. Observed Trends in Severe Weather Conditions Based on Humidex, Wind Chill, and Heavy Rainfall Events in Canada for 1953–2012. – *Atmos. Ocean.*, 2015, vol. 53, pp. 383–397.
9. Osczevski R., and Bluestein M. The New Wind Chill Equivalent Temperature Chart. – *Bulletin of the American Meteorological Society*, 2005, vol. 86, pp. 1453–1458.
10. Кобышева Н. В., Стадник В. В., Ключева М. В., Пигольцина Г. Б., Акентьева Е. М., Галюк Л. П., Разова Е. Н., Семенов Ю. А. Руководство по специализированному климатологическому обслуживанию экономики. – СПб.: Астерион, 2008. – 336 с.
11. Ogorodnikov V. A., and Prigarin S. M. Numerical Modelling of Random Processes and Fields: Algorithms and Applications. – Utrecht: VSP, 1996. – 240 p.
12. Королев В. Ю. EM-алгоритм, его модификации и их применение к задаче разделения смесей вероятностных распределений. Теоретический обзор. – М.: ИПИ РАН, 2007.

© М. С. Акентьева, Н. А. Каргаполова, 2022