

УДК 53.05, 519.6, 004.02

АНАЛИЗ ТЕМПЕРАТУРНЫХ СИГНАЛОВ И АЛГОРИТМ ИХ КЛАСТЕРИЗАЦИИ

© Ю. В. Волков

*Институт мониторинга климатических и экологических систем СО РАН,
634055, г. Томск, просп. Академический, 10/3
E-mail: yvvolkov@mail.ru*

Приведены результаты анализа температурных сигналов. Представлена математическая модель температурного сигнала в форме колебания, модулированного по амплитуде и фазе. Вычисление амплитуды и фазы реализовано на основе теории аналитического сигнала. Амплитуда и фаза температурных сигналов применены в качестве критериев выделения климатических кластеров. Описан алгоритм климатической кластеризации.

Ключевые слова: анализ сигналов, математическая модель, аналитический сигнал, приземная температура, классификация климата, кластеризация.

DOI: 10.15372/AUT20190305

Введение. Изучение изменений в климатической системе Земли в целом или отдельных её регионах является актуальной задачей. Наличие в этой системе глобальных процессов, способных кардинально менять условия жизни биологических объектов, делают задачу наиболее значимой. Сложность исследования и прогнозирования изменений в климатической системе заключается в её многокомпонентности. Количественная оценка вклада каждого климаторегулирующего фактора в общий климатический процесс — задача, не имеющая на данный момент однозначного решения [1].

Энергия Солнца является источником вынуждающего воздействия и модулятором различных изменений параметров геосферы Земли. Она инициирует и синхронизирует природно-климатические процессы [2]. Поля параметров климата трансформируются в приземном слое Земли в зависимости от географического положения местности и условий подстилающей поверхности [3].

Изучение климатической системы основано на выделении статистически достоверных отклонений характеризующих её параметров. Востребованы исследования характеристик физических процессов, происходящих в климатической системе, на основе создания новых эффективных методов анализа данных [4–7]. Важной задачей является кластеризация климата, установление границ между классами и оценка их изменчивости. Работы [8, 9] развивают разные подходы к определению критериев и к численным алгоритмам кластеризации климата. Известны исследования по определению границ климатических кластеров на основе выделения однотипных колебаний среднегодовых температур по знаку и фазе [10]. К недостаткам [8–10] нужно отнести отсутствие чёткой формулировки критериев кластеризации, однозначного вычисления амплитуды и фазы, оценки однотипности колебаний.

Целью данной работы является создание формализованного подхода к анализу климатических данных и кластеризации климата.

В основу разрабатываемого подхода положена математическая модель, используемая для описания и исследования сигналов приземных температур.

Такой подход позволил применить фазы и огибающие в качестве критериев для кластеризации температурных сигналов, создать новый алгоритм кластеризации темпера-

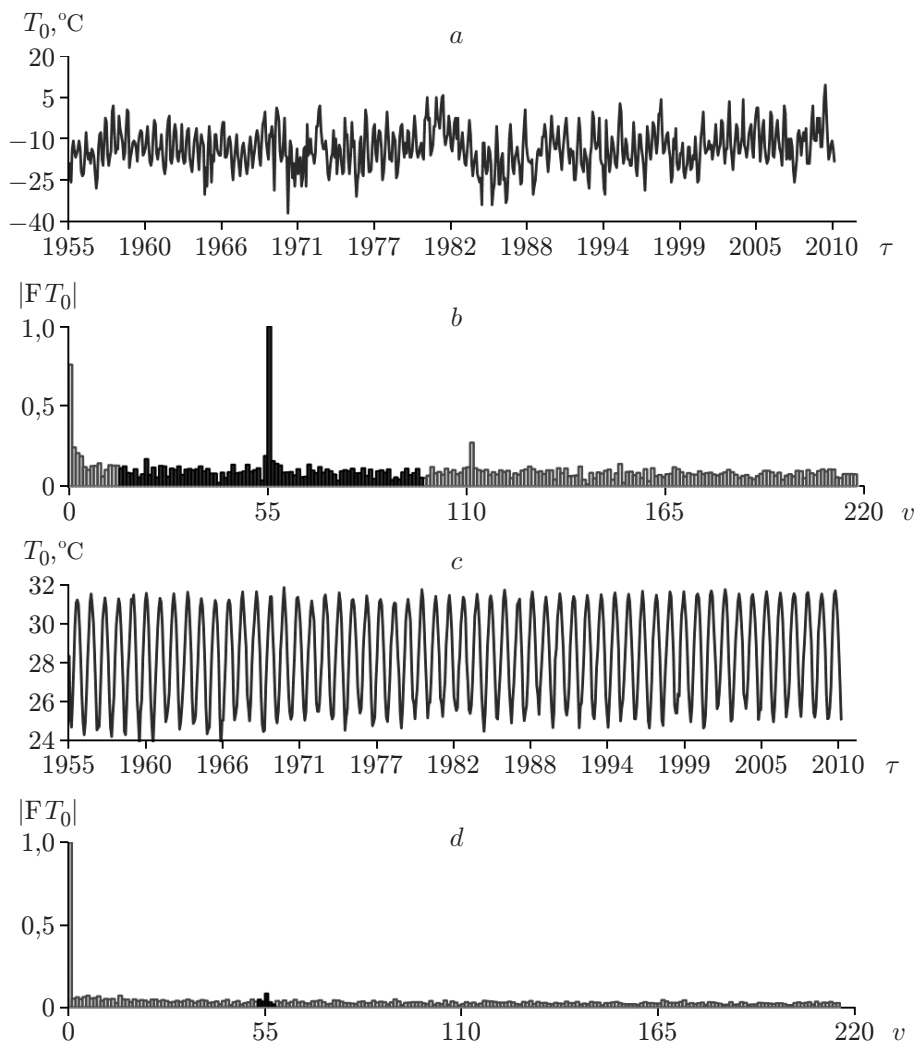


Рис. 1. Температурные сигналы $T_0(\tau)$ и их нормированные модули преобразования Фурье $|FT_0|$ за период 1955–2010 гг., полученные на метеостанциях: *a, b* — «Томск», *c, d* — Ciudad Bolivar. Представлены первые 220 частот из 2048

турных сигналов, использующий синхронность как основное свойство кластеров. Предлагаемое исследование является развитием методов, разработанных в [11, 12].

Предварительный анализ данных инструментальных измерений. В представленной работе использованы среднемесячные значения приземной температуры — температурные сигналы $T_0(\tau)$, полученные на 818 метеостанциях Северного полушария за период с 1955 по 2010 гг., интервал наблюдений составлял 55 лет [13]. На рис. 1 показаны два характерных температурных сигнала $T_0(\tau)$ и их фурье-спектры $|FT_0|$. Первый сигнал (рис. 1, *a*) получен на метеостанции «Томск» (Россия; синоптический индекс 29430; широта $56,5^\circ\text{N}$; долгота $84,9^\circ\text{E}$; высота над уровнем моря 139 м), расположенной в умеренных широтах, второй (рис. 1, *c*) — на метеостанции Ciudad Bolivar (г. Сьюдад-Боливар, Венесуэла; синоптический индекс 80444; широта $8,2^\circ\text{N}$; долгота $63,6^\circ\text{W}$; высота над уровнем моря 48 м), расположенной в тропических широтах.

Численная реализация оператора фурье-преобразования F осуществлялась с применением алгоритма быстрого преобразования Фурье (БПФ), который предназначен для пери-

одических сигналов. Для температурных сигналов характерен период в один год. Сигналы были периодически продолжены за пределами интервала наблюдения. Дополнительно 660 отсчётов температуры интерполировались сплайнами до 2048. Данные действия выполнялись для минимизации погрешностей реализации БПФ.

В частотном спектре температурного сигнала умеренных широт (рис. 1, *b*) наиболее выражен модуль на частоте 55, соответствующей несущей частоте колебания с периодом в один год на всём интервале наблюдения. Наименее выражены модули на кратных частотах, которые могут быть следствием нелинейных искажений некоторого начального колебания, как результат наличия модулирующих процессов. Несущая частота в частотном спектре сигнала тропических широт (рис. 1, *d*) выражена слабо. В частотных спектрах обоих сигналов (см. рис. 1, *b, d*) присутствуют также значительные постоянные составляющие на нулевой частоте. Имеется широкополосный частотный фон, величина которого для сигнала, характеризующего умеренные широты, составляет не более 10 % от уровня несущей частоты (см. рис. 1, *b*). Сравнительный анализ сигналов разных метеостанций показывает, что в умеренном поясе в спектре температурного сигнала преобладает несущая частота (см. рис. 1, *b*), тогда как в спектре сигналов тропического пояса несущая частота отличается не более чем на 8 % от фоновых частот (см. рис. 1, *d*).

Математическая модель. Сигнал, описывающий изменение температуры, можно построить в виде квазипериодического колебания, обусловленного глобальным ежегодным циклом вращения Земли, модулированного по амплитуде и фазе в результате влияния процессов различной природы. Формальная математическая модель такого сигнала представляется в виде

$$T_0(\tau) = a_0(1 + ma_m(\tau)) \cos(2\pi\nu_c\tau + n\varphi_m(\tau)), \quad (1)$$

где τ — время; a_0 — амплитуда несущего колебания; m, n — индексы амплитудной и фазовой модуляций; $a_m(\tau)$ — сигнал, модулирующий амплитуду; $\varphi_m(\tau)$ — сигнал, модулирующий фазу; ν_c — несущая частота.

Несущее колебание имеет период, равный одному календарному году. Функции, описывающие модулирующие $a_m(\tau)$, $\varphi_m(\tau)$ воздействия, в реальных температурных сигналах имеют сложную форму. Наибольшая часть информации для реальных сигналов содержится в симметричном частотном интервале шириной $\Delta\nu$, в котором нет иных выраженных спектральных составляющих, кроме несущей частоты ν_c , относительно которой интервал центрирован (см. рис. 1, *b, d*; спектральный интервал $\Delta\nu$ выделен тёмным цветом). Если справедливо неравенство $\Delta\nu/\nu_c \ll 1$, то спектр на этом интервале соответствует узкополосному гармоническому колебанию, модулированному по амплитуде и фазе. Данное условие выполняется только для температурного сигнала метеостанции умеренных широт (см. рис. 1, *b*). Модулированный по амплитуде и фазе температурный сигнал $T(\tau)$ может быть выделен из исходного температурного сигнала $T_0(\tau)$ с помощью спектрального окна $\Omega(\nu)$ с эффективной шириной $\Delta\nu$. Данное преобразование можно представить следующим образом:

$$T(\tau) = F^{-1}\{\Omega[F T_0(\tau)]\} = A(\tau) \cos \Phi(\tau),$$

где $\Phi(\tau)$, $A(\tau)$ — модулированная полная фаза и амплитуда.

Аналитический сигнал. Законы изменения амплитуды и фазы в выражении (1) требуют определения, которое должно быть конструктивным, т. е. должно поддерживаться вычислительным алгоритмом. Непротиворечивое определение амплитуды и фазы даётся на основе аналитического сигнала, введённого в [14]. Аналитический сигнал $W(\tau)$ являет-

ся комплексной функцией, которая представляет колебание температуры $T(\tau)$ во времени как естественное обобщение гармонических колебаний:

$$W(\tau) = T(\tau) + iV(\tau), \quad V(\tau) = \frac{1}{\pi} \text{v.p.} \int_{-\infty}^{\infty} \frac{T(s)}{\tau - s} ds, \quad (2)$$

где несобственный интеграл определён в смысле главного значения по Коши (v.p.) в тех случаях, когда $s \rightarrow \pm\infty$ и при $\tau = s$. Мнимая составляющая аналитического сигнала $V(\tau)$ является Гильберт-трансформантой [15] его действительной части $T(\tau)$. Введение комплексного представления $W(\tau)$ позволяет определять амплитуду и фазу известным способом:

$$A(\tau) = \sqrt{T^2(\tau) + V^2(\tau)}, \quad \Phi(\tau) = \text{arctg} \frac{V(\tau)}{T(\tau)}. \quad (3)$$

Функция $A(\tau)$ для сигнала, представленного в виде (1), описывает закон изменения его амплитуды. Данная функция имеет смысл огибающей амплитуды.

Величина $\Phi(\tau)$ есть полная фаза, которая определена как главное значение функции arctg и может быть сшита при выполнении условия непрерывности. Для узкополосного сигнала непрерывная фаза $\Phi(\tau)$ всегда реализуется и изменяется монотонно.

С вычислительной точки зрения преобразование Гильберта эквивалентно умножению в частотной области ν на функцию $i \text{sgn} \nu$, поэтому аналитический сигнал, соответствующий данной действительной функции $T(\tau)$, получается путём обнуления одной половины фурье-спектра этой функции. Такой односторонний спектр называют причинным или каузальным. Эта операция использует свойства преобразования Гильберта и справедлива не только для узкополосных сигналов, но и в более общих случаях, когда $T(\tau)$ по своей природе является двухполосным сигналом, т. е. несущая частота ν_c высока настолько, что спектры функций $e^{2\pi i \nu_c \tau}$ и $A(\tau) e^{i\Phi(\tau)}$ не пересекаются. Для реальных сигналов это условие выполняется приближённо. В нашем случае сигнал $T(\tau)$ формируется искусственно, для него всегда можно подобрать спектральное окно $\Omega(\nu)$ таким, что фаза $\Phi(\tau)$ будет монотонной [10]. Однако для сохранения полезной информации размер спектрального окна нужно выбирать максимально широким при выполнении условия монотонности фазы.

Следующей задачей является установление закономерностей изменений в климатической системе Земли по параметрам амплитуд и фаз температурных сигналов.

Анализ расчётных данных. Рассчитанные в соответствии с формулами (3) для двух характерных сигналов $T(\tau)$ огибающие (рис. 2, *a*, *b*) имеют разный характер поведения. Огибающая, характеризующая сигнал умеренных широт (см. рис. 2, *a*), на выбранном интервале времени изменяется больше, чем огибающая сигнала экваториальной зоны (см. рис. 2, *b*). Подобные отличия имеют также и преобразованные фазовые функции $\Phi'(\tau)$, полученные путём вычитания из полных фаз прямых, характеризующих фазу несущего колебания (рис. 2, *c*). Данный результат — следствие применения малого размера спектрального окна при фильтрации частотного спектра экваториального сигнала в целях соблюдения условия монотонности фазы. Огибающая и фаза сигнала экваториальных широт содержат недостаточное для дальнейшего анализа количество информации. Пригодными для дальнейшего анализа являются огибающая и фаза сигнала умеренных широт. В частотных спектрах фазы (рис. 2, *d*) и огибающей (рис. 2, *e*) этого сигнала наиболее выражены частоты в интервале от 0 до 40, что соответствует колебаниям с периодами более одного года. В полученных частотных комбинациях содержится информация, которую можно использовать для оценки сходства сигналов разных метеостанций и для построения алгоритма климатической кластеризации.

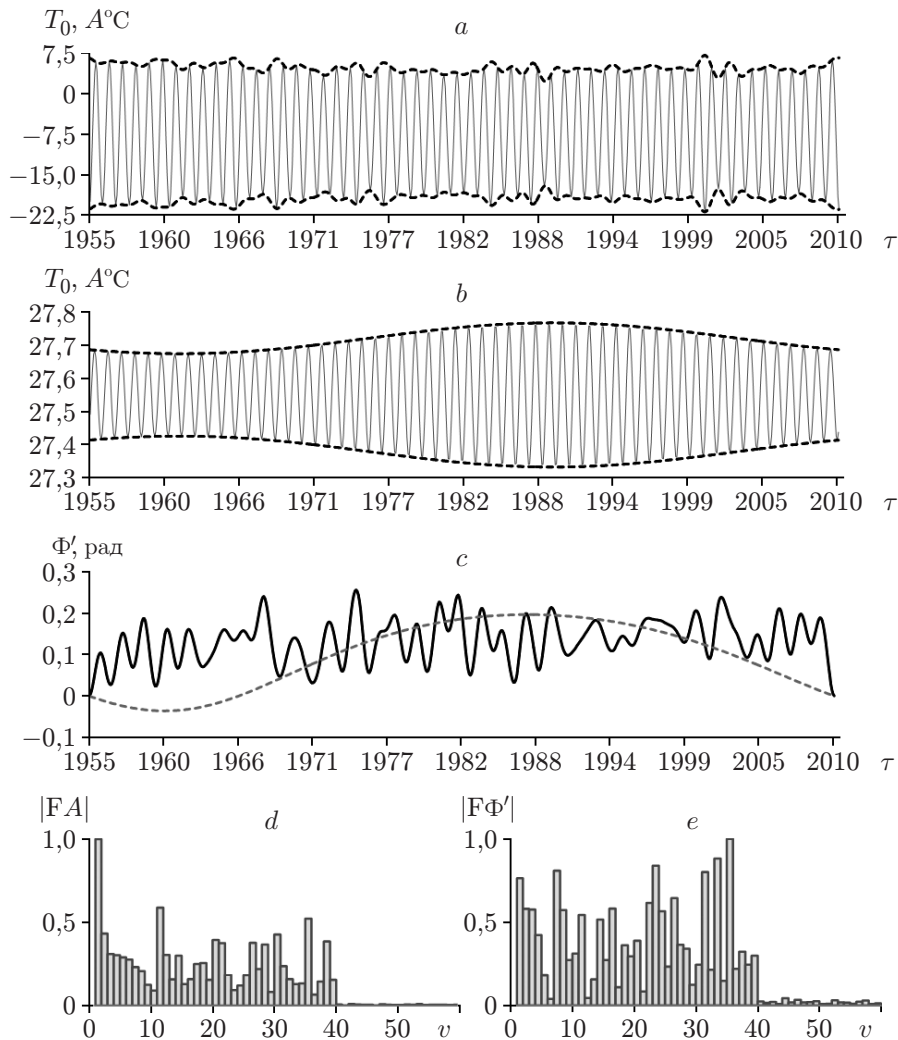


Рис. 2. Фильтрованные температурные сигналы $T(\tau)$ и их характеристики: a — огибающая сигнала, полученного на метеостанции «Томск» (штриховые кривые); b — огибающая сигнала, полученного на метеостанции Ciudad Bolivar (штриховые кривые); c — преобразованные фазы сигналов: метеостанции «Томск» (сплошная кривая), метеостанции Ciudad Bolivar (штриховая кривая); d , e — нормированные модули преобразования Фурье для огибающей и фазы сигнала, полученного на метеостанции «Томск»

Алгоритм климатической кластеризации. Вычисленные характеристики, преобразованные фазы или огибающие создают множество M^i , номера элементов которого соответствуют метеостанциям с номерами $j = 1, 2, \dots, N$ или $k = 1, 2, \dots, N$, $N = 818$; i — порядковый номер итерации. Процедура кластеризации состоит в вычислении корреляционной матрицы $\|R_{j,k}\|$ элементов множества и последующего разбиения этого множества на кластеры. Каждый кластер формируется из элементов, соответствующих условию $R_{j,k} = \langle M_j^i, M_k^i \rangle \geq r$, где r — требуемый уровень взаимной корреляции. Для каждого элемента множества M^i , попавшего в один из кластеров, определяется текущая характеристика, которая вычисляется как среднее значение всех элементов, вошедших в кластер. Формируется новое множество M^{i+1} , элементами которого являются рассчитанные текущие характеристики. Устанавливается уровень отличия между массивами путём расчёта

абсолютной разницы между элементами массивов M^i и M^{i+1} . В случае превышения максимального значения абсолютной разницы между элементами заданного уровня точности $\varepsilon = \pm 0,001$ повторяется итерационный процесс, начинающийся с расчёта корреляционной матрицы $\|R_{j,k}\|$ для элементов массива M^{i+1} . Итерационный процесс останавливается при достижении заданного уровня точности ε .

Сходимость итерационного процесса следует из двух условий: конечности выборки (число метеостанций равно N) и справедливости неравенства $|\langle M_j^i \rangle| \leq \max |M_j^i|$. Число групп не может быть больше числа элементов. В итерационном процессе значения текущих элементов из любого подмножества выборки стремятся к предельному числу кластеров. Количество кластеров зависит от коррелированности участвующих в итерационном процессе исходных и текущих элементов, а также от заданного уровня корреляции r .

Полученные текущие функции характеризуют климатические закономерности, вокруг которых формируются климатические кластеры. В кластеры включаются климатические станции, для которых уровень корреляций исследуемых характеристик температурных сигналов с полученной текущей функцией удовлетворяет заданному критерию r . Кластеры с одинаковым составом станций учитываются 1 раз.

Результаты натурального эксперимента. Применение алгоритма климатической кластеризации к множеству, состоящему из 818 температурных сигналов при заданном уровне корреляции $r = 0,8$ и $\varepsilon = \pm 0,001$, привело к его сокращению до 17 кластеров по фазам (максимальное число итераций $i = 11$) и до 16 кластеров по огибающим (максимальное число итераций $i = 10$). Конечное число итераций при неизменном значении $\varepsilon = \pm 0,001$ является подтверждением наличия сходимости итерационного алгоритма. В распределении элементов выделенных кластеров по географическим координатам существует чёткая географическая локализация. Каждый отдельный кластер характеризует географический район, где изменения температуры происходят синхронно [16]. Климатические кластеры соответствуют известным климатическим классификациям [17], что является подтверждением правильности разработанного формализованного подхода.

Заключение. Результаты анализа показывают, что спектр температурного сигнала, полученного на метеостанции, расположенной вблизи экватора, не имеет достаточно выраженной спектральной компоненты на несущей частоте. Введённое условие монотонности и неразрывности фазы для данного сигнала может быть выполнено при минимальном размере фильтрующего окна.

В работе предложено использовать в качестве критериев климатической кластеризации огибающие и фазы температурных сигналов. Показана применимость данных критериев к построению климатических кластеров на основе разработанного нового численного алгоритма.

Установлено, что в частотных спектрах огибающей и фазы преобладают спектральные составляющие, частоты которых меньше частоты, соответствующей годовому колебанию температурного сигнала. Период закономерностей, определяющих формирование климатических кластеров, составляет более одного года.

Результаты работы можно использовать для оценивания регионального климата в разных временных и пространственных масштабах. На основе разработанного алгоритма может быть построена система мониторинга климатических зон.

СПИСОК ЛИТЕРАТУРЫ

1. **Дымников В. П.** Устойчивость и предсказуемость крупномасштабных атмосферных процессов. М.: ИВМ РАН, 2007. 283 с.
2. **Тартаковский В. А.** Синхронный анализ рядов чисел Вольфа и температуры с метеостанций Северного полушария Земли // Оптика атмосферы и океана. 2015. **28**, № 2. С. 182–188.

3. **Хромов С. П., Петросян М. А.** Метеорология и климатология. М.: Наука, 2006. 582 с.
4. **Peel M. C., Finlayson B. L., McMahon T. A.** Updated world map of the Köppen-Geiger climate classification // Hydrol. Earth Syst. Sci. 2007. **11**, Is. 2. P. 1633–1644.
5. **Franziska H., Körper J., Spangehl T., Cubasch U.** Shifts of climate zones in multi-model climate change experiments using the Köppen climate classification // Meteorologische Zeitschrift. 2012. **21**, N 2. S. 111–123.
6. **Zhang X., Yan X.** Spatiotemporal change in geographical distribution of global climate types in the context of climate warming // Climate Dynamics. 2014. **43**, N 3–4. P. 595–605.
7. **Морозов Ю. В., Спектор А. А.** Классификация объектов на основе анализа спектральных характеристик огибающих сейсмических сигналов // Автометрия. 2017. **53**, № 6. P. 49–56.
8. **Unal Y., Kindap T., Karaca M.** Redefining the climate zones of Turkey using clusteranalysis // Intern. Journ. Climatol. 2003. **23**, N 9. P. 1045–1055.
9. **DeGaetano A. T.** Spatial grouping of United States climate stations using a hybrid clustering approach // Intern. Journ. Climatol. 2001. **21**, N 7. P. 791–807.
10. **Салугашвили Р. С.** Колебание климата на территории первого естественного синоптического района и климатическое районирование // Уч. зап. Казан. ун-та. Сер. Естественные науки. 2012. **154**, Кн. 1. С. 216–227.
11. **Tartakovsky V. A.** Causality and demodulation of optical monotone-phase signal // Atmos. Oceanic Opt. 2002. **15**, N 1. P. 78–86.
12. **Тартаковский В. А., Волков Ю. В., Исаев Ю. Н. и др.** Математическая модель радиального сечения годовых колец деревьев // Автометрия. 2003. **39**, № 5. С. 118–127.
13. **Met Office.** Weather and climate. URL: <http://www.metoffice.gov.uk> (дата обращения: 6.03.2019).
14. **Gabor D.** Theory of communication // Journ. IEEE. 1946. **3**, N 3. P. 429–441.
15. **Вакман Д. Е., Вайнштейн Л. А.** Амплитуда, фаза, частота — основные понятия теории колебаний // УФН. 1977. **123**, № 5. С. 657–682.
16. **Cheredko N. N., Tartakovsky V. A., Krutikov V. A., Volkov Yu. V.** Climate classification in the Northern hemisphere using phases of temperature signal // Atmos. Oceanic Opt. 2017. **30**, N 1. P. 63–69.
17. **Rohli R. V., Joyner T. A., Reynolds St. J. et al.** Globally extended Köppen–Geiger climate classification and temporal shifts in terrestrial climatic types // Phys. Geography. 2015. **36**, Is. 2. P. 142–157.

Поступила в редакцию 06.03.2019

После доработки 29.03.2019

Принята к публикации 16.04.2019
