

ПРИМЕНЕНИЕ СПЕКТРАЛЬНО-СИНГУЛЯРНОГО ПОДХОДА ДЛЯ РАЗРАБОТКИ ДОЛГОСРОЧНОГО ПРОГНОЗА ТЕМПЕРАТУРЫ ВОЗДУХА В ХОЛОДНОЕ ПОЛУГОДИЕ

Л. М. Акимов, Т. Н. Задорожная

*Воронежский государственный университет, Россия
Военный учебно-научный центр ВВС «Военно-воздушная академия
им. профессора Н. Е. Жуковского и Ю. А. Гагарина», Россия*

Поступила в редакцию 30 апреля 2014 г.

Аннотация: в статье рассматриваются вопросы возможности разработки прогноза суммарной за январь температуры воздуха в холодное полугодие на основе метода выборочных компонент. Показана успешность данного подхода.

Ключевые слова: суммарная температура воздуха, январь, сингулярное разложение ряда, компоненты, точность восстановления, прогноз.

Abstract: The article presents the issues of possibility to forecast total for January air temperature in half-year cold period on basis of sample component method. The effectiveness of the approach is also illustrated in the work.

Key words: total air temperature, January, the singular value decomposition of a number, components, accuracy of reconstruction, prediction.

Проблема долгосрочного прогноза погоды, особенно такой ее составляющей как температуры воздуха, остается задачей актуальной. Сложность проблемы в ее многофакторности, которую не всегда можно установить, а тем более учесть.

Большинство современных методов базируется в основном на использовании закономерностей атмосферной циркуляции (аналоги, ритмы, преемственные связи, циклы, климатические вероятности, календарные особенности, использование памяти предикторов, инерция).

Наиболее четко в атмосфере проявляется ритмичная деятельность, которая с успехом используется в прогностических целях. Это связано с многофакторностью атмосферных процессов, которые, обладая различной периодичностью, возвращающихся через определенные промежутки времени в свое первоначальное состояние. Накладываясь друг на друга и взаимодействуя между собой, они создают определенный образ погоды.

В этом случае задача долгосрочного прогноза погоды сводится к выявлению потенциальных цик-

лов и выбору наиболее эффективного подхода их обнаружения.

В последние годы для выявления ритмичности в атмосферных процессах используется фундаментальный статистический аппарат. В его основу положен метод анализа выборочных главных компонент [2, 5, 6]. Основное требование данного метода – наличие длинных рядов наблюдений. В случае анализа коротких рядов предложен метод, в котором прогностические решения осуществляются за счет сдвигов на одно значение. Этот метод описан в работах [3, 4, 11] и реализован в [1, 7, 8, 9, 10]. Данный подход помогает увидеть и выделить в поведении временных рядов разного рода закономерности, часто априори не очевидные. Суть метода заключается в том, что одномерный временной ряд преобразуется в матрицу развертки с помощью сдвиговой процедуры. С этой целью в исследуемом временном ряде размером N в качестве первого шага выбирается отрезок L ($L = x_1, x_2, x_3, \dots, x_L$), названной «длиной окна».

На основе исследуемого ряда строится выборочная траекторная матрица A , столбцами которой являются скользящие отрезки длины ряда N : с первой точки по L , со второй по $L+1$ и т.д.

Затем осуществляется сингулярное разложение данной матрицы. Из полученного набора выбираются главные компоненты, по которым можно восстановить исходный ряд. Реализация данного подхода осуществляется в несколько этапов.

На первом этапе проводится стандартизация элементов матрицы A , с последующим переходом к матрице A^* . Для получения матрицы A^* вычисляется матрица коэффициентов корреляции $S = A^*(A^*)^T$. Для матрицы S определяются собственные значения I_1 и собственные векторы U_i , исходя из соотношения: $SU_i = I_i U_i$, ($i = 1, 2, \dots, k$).

Поскольку матрица S является симметрической и неотрицательно определенной, то ее собственные значения будут действительными и неотрицательными ($I_i \geq 0$), что позволяет выстроить их в убывающем порядке $I_1 \geq I_2 \geq \dots \geq I_k \geq 0$. Собственные векторы матрицы S будут ортонормированными ($U_i U_j^T = d_{ij}$) в силу линейной независимости и образуют ортогональную матрицу $U = (U_1 U_2 \dots U_k)$ или матрицу вращения. Для $I_i > 0$ строится вектор $V_i = I_i^{-1/2} (A^*)^T U_i$, который будет собственным вектором матрицы $(A^*)^T A^*$, соответствующим собственному числу I_i . В этом случае векторы V_i и U_i являются правым и левым сингулярными векторами матрицы A^* , а числа $\sqrt{I_i}$ – сингулярными числами. При этом будет иметь место равенство:

$$A^* = \sum_{i=1}^k \sqrt{I_i} U_i V_i^T \quad (1)$$

При введении в рассмотрение матрицы $A_i^* = \sqrt{I_i} U_i V_i^T$, разложение представляется в виде $A^* = \sum_{i=1}^k A_i^*$. Это разложение есть сингулярное разложение матрицы A^* . Набор $(\sqrt{I_i}, U_i, V_i)$ является i -ой собственной тройкой матрицы A^* .

На основе разложения (1) производится отбор матриц $A_i^* = \sqrt{I_i} U_i V_i^T$. Тем самым исходное разложение в сгруппированном виде может быть записано как:

$$A^* = A_{j_1}^* + A_{j_2}^* + \dots + A_{j_k}^*, \quad (2)$$

где $J_1 + J_2 + \dots + J_k = K$.

На данном этапе производится отбор из общего множества тех собственных троек, которые войдут в восстановление ряда. С помощью визуального и аналитического анализа компонент, полученных в результате разложения, делается вывод о медленно меняющихся трендовых составляю-

щих в исходном временном ряде, высокочастотных циклических колебаниях и шумовом фоне.

На последнем этапе каждая матрица полученного разложения переводится в новый ряд длины N . Предварительно выполняется денормирование путем умножения всех элементов столбцов матрицы A^* на соответствующие значения стандартных отклонений S_j , затем производится децентрирование путем добавления к элементам каждого столбца соответствующих выборочных средних \bar{a}_i . В результате производится преобразование матрицы N путем диагонального усреднения, которое выполняется по формуле:

$$f_k = \begin{cases} \frac{1}{C} \sum_{j=1}^c a_{j,c-j+1}, & \text{если } 1 \leq c < L^* \\ \frac{1}{L^*} \sum_{j=1}^{L^*} a_{j,c-j+1}, & \text{если } L^*-1 \leq c < k^* \\ \frac{1}{N-c} \sum_{j=c-k^*}^{N-k^*+1} a_{j,c-j+1}, & \text{если } k^* \leq c < N \end{cases} \quad (3)$$

Выражение (3) соответствует усреднению элементов матрицы вдоль «диагоналей» $i + j = c + 1$, т.е.

при $c = 1$, $f_1 = a_{11}$; при $c = 2$, $f_2 = \frac{a_{12} + a_{21}}{2}$ и т.д.

Применив диагональное усреднение к полученным ранее матрицам, производится разложение

исходного ряда в сумму: $a_i = \sum_{j=1}^K f_j$ ($i = 1, 2, \dots, n$).

В данной статье, описанная выше процедура была применена к исходному ряду, составленному из суммарных за месяц максимальных суточных температур воздуха по данным пункта Воронеж за период с 1946 по 2010 годы. Ставилась задача выявить циклические закономерности в многолетнем ходе максимальной температуры воздуха с целью выработки прогностических рекомендаций на длительные сроки по верхнему пределу.

В соответствии с вышеописанной методикой данный временной ряд разделен на отдельные составляющие, вносящие различный вклад в дисперсию исходного ряда. В данном случае отобрано 3 составляющих, описывающих до 75 % дисперсии. С целью их интерпретации и выявления основных особенностей в статистической структуре исследуемого ряда был произведен визуальный анализ самих собственных векторов, соответствующих тренду и возможным гармоникам при условии их хорошей делимости.

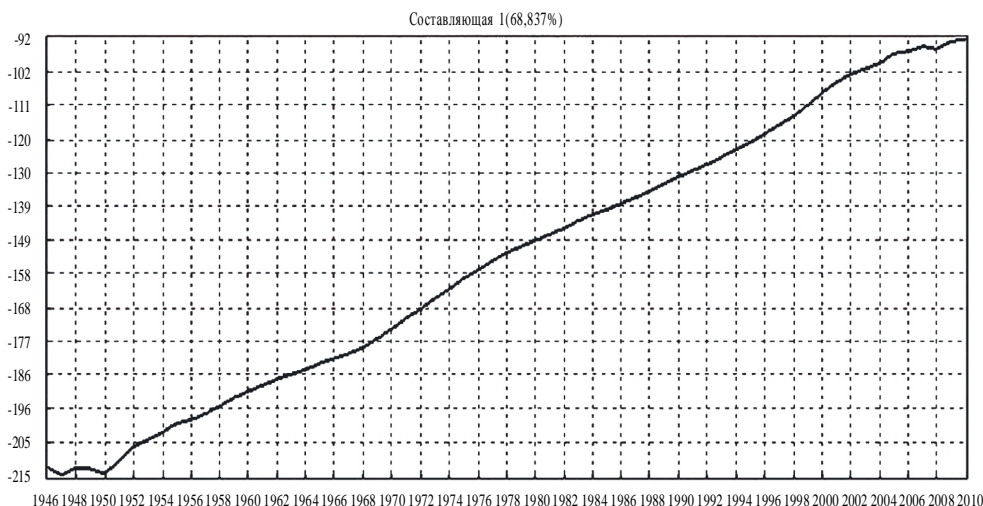


Рис. 1. Одномерная диаграмма первого собственного векторов временного хода суммарной за месяц максимальной температуры воздуха

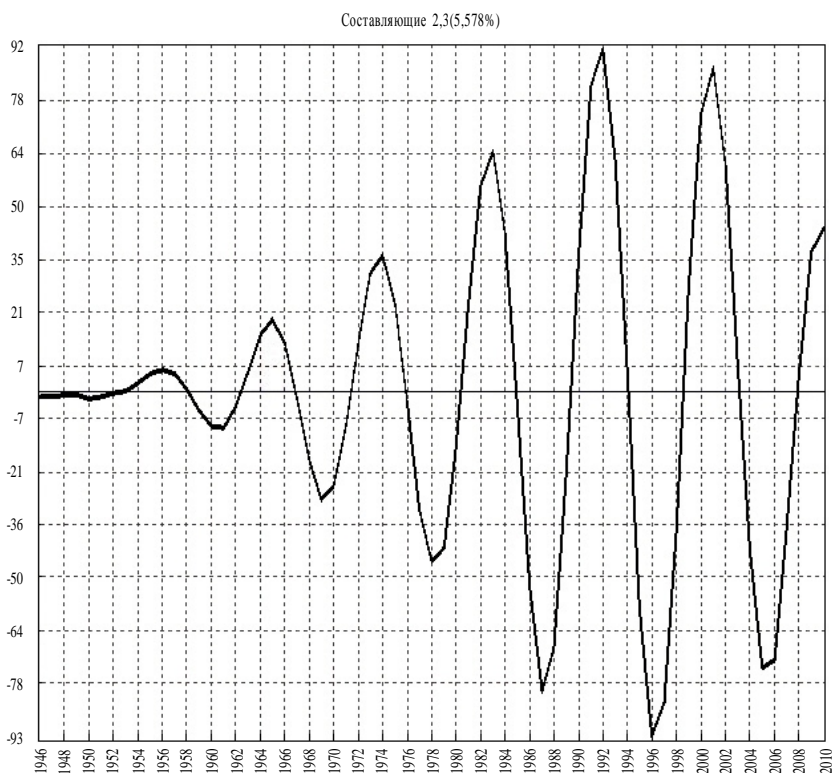


Рис. 2. Одномерная диаграмма главных компонент со вторым и третьим и порядковыми номерами

Согласно установленному понятию, трендом является медленно меняющаяся компонента ряда, не содержащая колебаний. Исходя из такого определения, на одномерных диаграммах собственных векторов для выявления тренда необходимо найти медленно меняющиеся собственные векторы.

Визуальный их анализ показал, что в данном случае первый собственный вектор имеет требуемый вид, который представлен на рисунке 1.

Представленная на рисунке 1 первая главная компонента, описывает 68,837 % дисперсии. Визуальный анализ ряда, изображенного на рисунке 1, свидетельствует о том, что он имеет хорошо выраженный тренд, у которого отмечается положительный знак, свидетельствующий о монотонном увеличении температуры. Кроме того, наблюдаются указания на возможные периодические колебания. В данном случае обнаруженный тренд описывается единственной собственной тройкой.

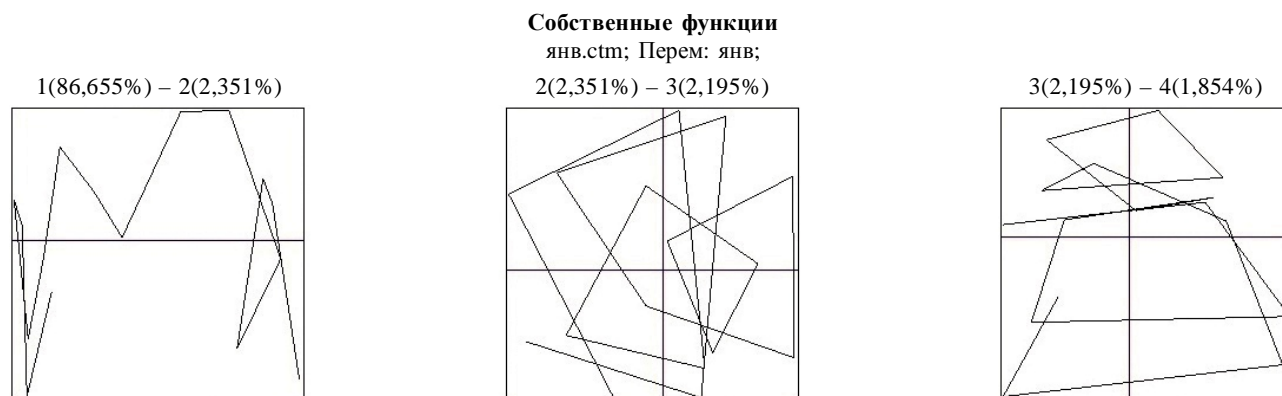


Рис. 3. Двумерные диаграммы, отражающие траектории точек собственных векторов

Наличие в исходных рядах трендовой составляющей позволяет разрабатывать инерционные прогнозы. Такой подход, в случае отсутствия методического прогноза, позволяет довольно эффективно планировать различные мероприятия на длительные периоды.

На следующем этапе проводилось выявление периодических колебаний. С этой целью анализировалась идентификация остальных гармоник, которые должны показать наличие колебаний с возможно меняющейся амплитудой исходного ряда. О наличии колебаний на фоне слабой общей тенденции могут свидетельствовать 2-3 составляющие, представленные на рисунке 2.

Из анализа рисунка 2 следует, что в ходе кривой второго и третьего собственного вектора имеет место регулярное периодическое поведение. Особенно четко это проявилось во второй половине наблюдений. При этом имеет место четко выраженный девятилетний период. Вклад этих составляющих, по сравнению с первой составляющей, значительно меньше (5,58%).

Для уточнения устойчивости регулярного характера выявленных закономерностей рассмотрены двумерные траектории точек с координатами $(z_{ij}, z_{i+1,j})$, для $i = 1, 2, \dots, 6, j = 1, 33$, которые представлены на рисунке 3 для трех первых гармонических компонентов.

Представленные на рисунке 3 изображения – двумерные диаграммы траекторий точек собственных векторов, позволяющих сделать вывод о наличии различных квазирегулярных периодичностей. Траектории собственных функций, представляющие различные формы, образуют двумерные траектории с вершинами, лежащими на кривой, имеющей спиралеобразную форму. Так, в первом случае (для первой и второй собственных функций) точки совершают сложные движения с явным

возвышающим эффектом, который может свидетельствовать о наличии не совсем монотонного положительного тренда.

Это означает, что соответствующая пара собственных векторов порождена моделированной гармонической компонентой исходного ряда, каждая из которых имеет свой период и амплитуду. Устойчивость и надежность таких периодов могут быть оценены величиной дисперсии соответствующей составляющей.

Во втором случае, который описывает совместное движение второй и третьей составляющей, наблюдаются четко выраженные ритмические почти круговые движения. Амплитуда их достигает симметричности на некотором расстоянии по оси U . В третьем случае (для третьей и четвертой собственных функций) точки также совершают регулярные движения, возвращаясь в исходное положение через определенное время.

Следует отметить, что в первом и втором случаях движение имеет более устойчивый регулярный характер по сравнению с движением в третьем и четвертом случаях.

С точки зрения инерционного прогноза наибольший интерес представляет первая и вторая точки, соответствующие первым парам выборочных главных компонент скользящего отрезка, так как эти точки связаны не только с большим процентом общей дисперсии, но и имеют заметно большую регулярность движения. Здесь следует отметить, что с ростом номеров проекция тенденции и нерегулярности, как правило, возрастает, что компенсируется снижением соответствующего % общей дисперсии.

Подтверждением эффективности разложения временного ряда на главные компоненты, а именно: выделение тренда, сигнала и шума, является оценка точности восстановления его с помощью

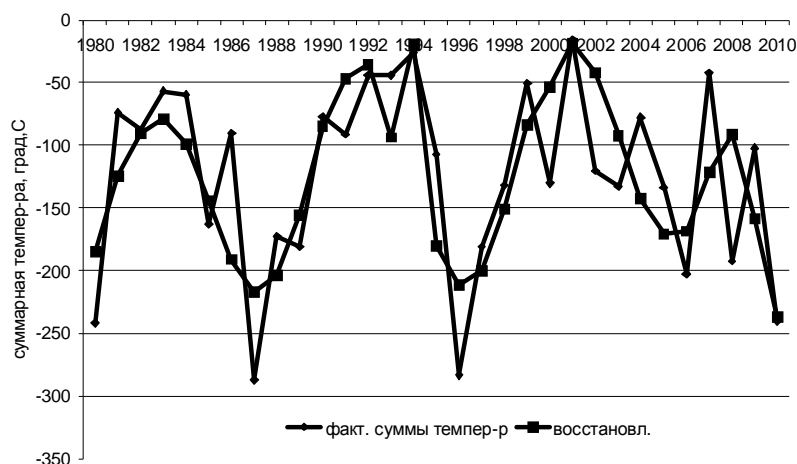


Рис. 4. Фактический и восстановленный ряды (январь, Воронеж, сумма максимальных температур)

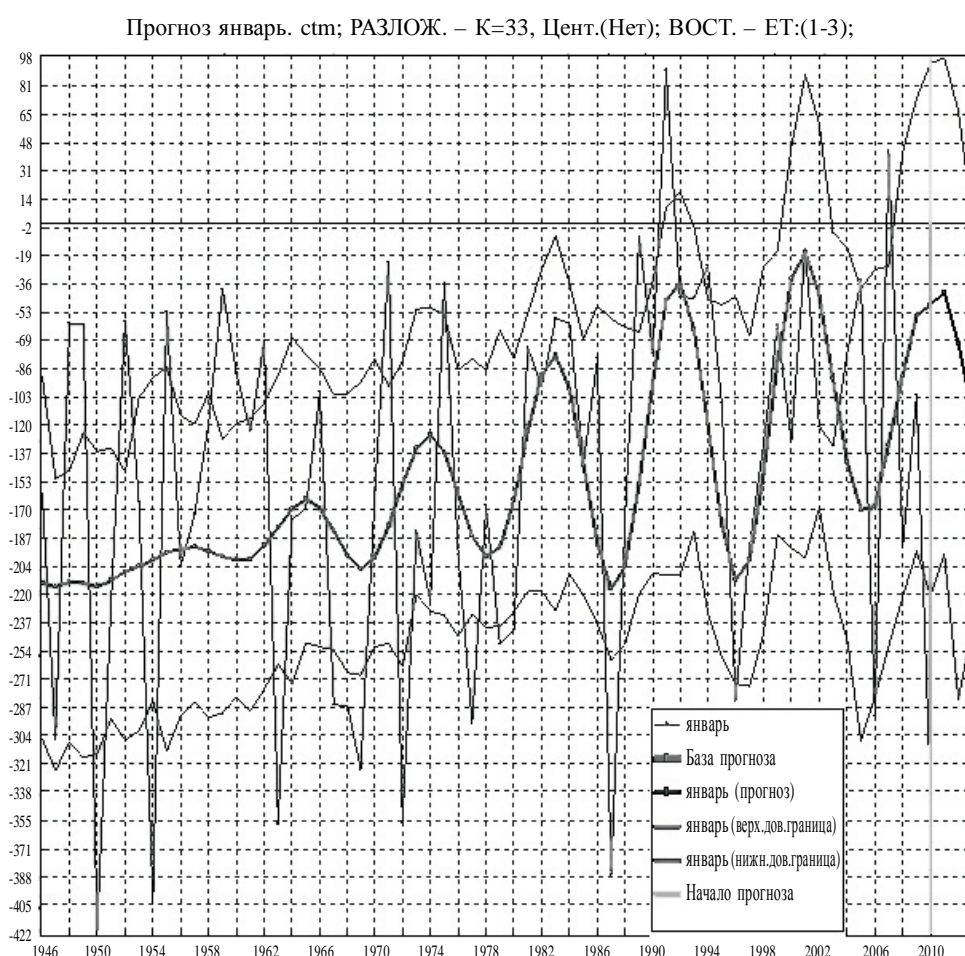


Рис. 5. Прогноз суммарной за месяц максимальной температуры воздуха в пункте Воронеж на январь по трем первым главным компонентам

нескольких первых составляющих. В данном случае восстановление, произведенное по первым трем выборочным главным компонентам, представлено на рисунке 4, на котором для сравнения приведен ряд исходных данных.

Нетрудно видеть, что кривые имеют вид близкий к синусоидальному, наблюдаемый по сглажен-

ной кривой восстановленных значений, где нестабильная часть колебаний исключена. Характер кривой восстановленных значений отражает регулярность вращения точек, указывая на ее синусоидальность. Ошибка восстановления составляет $56,6^{\circ}\text{C}$.

Следует заметить, что среднемесячные значения суммарной температуры воздуха в январе со-

ставляют $-123,8^{\circ}\text{C}$. При этом экстремальные максимальные значения находились на уровне $-50-70^{\circ}\text{C}$, а экстремально низкие температуры близки к -300°C . Важно отметить, что каждый из них был успешно восстановлен.

Таким образом, с помощью выбранных составляющих достаточно хорошо выявлены отдельные особенности ряда. Это позволяет предположить, что по данным составляющим можно попытаться разработать прогноз с некоторой заблаговременностью. С помощью данного подхода с учетом трех первых главных компонент был разработан прогноз суммарной за январь месяц температуры воздуха, представленный на рисунке 5 жирной линией.

С этой целью исходная выборка была разделена на две части – первая половина периода являлась известной, а вторая, начиная с 2010 года, прогнозировалась с помощью первых 3-х главных компонент.

Как следует из рисунка 5, с помощью данного подхода удалось предусмотреть как основные колебания, так и структуру ряда в целом. Согласно графическим показаниям, температура воздуха в январе после 2010 года начала несколько понижаться.

Ошибка прогноза за данный отрезок времени, оцениваемая по значениям среднеквадратического отклонения s , составляет $62,7^{\circ}\text{C}$. Учитывая тот факт, что естественные колебания суммарной за январь максимальной температуры воздуха в Воронеже составляют $-74,6^{\circ}\text{C}$, можно полагать, что использование для долгосрочного прогноза предложенной методики будет эффективнее, чем традиционный климатический прогноз на 16%. Итак, использование предложенного подхода может быть полезным при планировании температурно-зависимых мероприятий на период, превышающий годовую заблаговременность.

СПИСОК ЛИТЕРАТУРЫ

1. Акимов Л. М. Применение метода главных компонент к прогнозу облачного покрова установленной градации / Л. М. Акимов, Т. Н. Задорожная // Вестник Воронежского государственного уни-

Акимов Леонид Мусамудинович

кандидат географических наук, доцент, заведующий кафедрой природопользования факультета географии, геоэкологии и туризма Воронежского государственного университета, г. Воронеж, т. (473)266-56-54, 8-951-850-49-82, E-mail: akl63@bk.ru, deanery@geogr.vsu.ru

Задорожная Тамара Николаевна

кандидат географических наук, доцент, старший научный сотрудник 3 научно-исследовательского отдела Военного учебно-научного центра Военно-воздушных сил «Военно-воздушная академия», г. Воронеж

верситета. Сер. География. Геоэкология. – 2013. – № 1. – С. 74-78.

2. Бриллинджер Д. Временные ряды: Обработка данных и теория / Д. Бриллинджер. – Москва : Мир, 1980. – 536 с.

3. Голяндина Н. Э. Метод Гусеница-SSA : анализ временных рядов : учебное пособие / Н. Э. Голяндина. – Санкт-Петербург, 2004. – 76 с.

4. Данилов Д. Л. Главные компоненты временных рядов : Метод «Гусеница» / Д. Л. Данилов, А. А. Жиглявский. – Санкт-Петербург : Санкт-Петербургский университет, 2007. – 308 с.

5. Мещерская А. В. Разложение годового хода ледовитости северных морей по естественным ортогональным функциям времени / А. В. Мещерская, Л. Р. Дмитриева-Араго // Метеорология и гидрология. – 1968. – № 10. – С. 56-64.

6. Пичугин Ю. А. Естественные составляющие годового хода приземной температуры / Ю. А. Пичугин // Метеорология и гидрология. – 1994. – № 12. – С. 34-42.

7. Прогноз внутримесячного хода аномалии температуры для Европейской территории России и Западной Сибири / О. В. Батырева [и др.] // Метеорология и гидрология. – 1995. – № 12. – С. 20-31.

8. Решетов В. Д. Прогнозирование колебаний температуры и давления воздуха в предстоящем месяце методом статистического учета истории колебаний / В. Д. Решетов, Н. Д. Тихомиров // Применение статистических методов в метеорологии : труды 11 Всесоюзного симпозиума по применению статистических методов в метеорологии, Ленинград, 11-13 нояб. 1975 г. – Ленинград : Гидрометеиздат, 1977. – С. 84-90.

9. Шерстюков Б. Г. Метод кратной цикличности анализа временных рядов и сверхдолгосрочных прогнозов / Б. Г. Шерстюков, А. А. Исаев // Метеорология и гидрология, 1999. – № 8. – С. 46-54.

10. Шипко Ю. В. Анализ временных рядов средней температуры воздуха северного полушария в терминах метода «Гусеница» / Ю. В. Шипко, Т. Н. Задорожная, С. Л. Бузмаков // Сборник статей Всероссийской научно-практической конференции, 26 ноября 2009. – Воронеж : ВАИУ, 2010. – С. 96-101.

11. Vautard R. Singular-spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. / R. Vautard, M. Ghil // Physica D. – 1989. – Vol. 35. – P. 395.

Akimov Leonid Musamudinovitch

Candidate of Geographical Sciences, Associate Professor, Head of the chair of nature management, Department of geography, geoecology and tourism, Voronezh State University, Voronezh, tel. 8-951-850-49-82, (473) 266-56-54, E-mail: akl63@bk.ru, deanery@geogr.vsu.ru

Zadorozhnaya Tamara Nikolayevna

Candidate of Geographical Sciences, Associate Professor, senior researcher of 3 Research Division of Military Educational and Scientific Centre of the Air Force, «Air Force Academy», Voronezh