

ИССЛЕДОВАНИЕ РЕЖИМА ТЕМПЕРАТУРЫ АТМОСФЕРНОГО ВОЗДУХА С ПОМОЩЬЮ МОДЕЛИ
АВТОРЕГРЕССИИ-СКОЛЬЗЯЩЕГО СРЕДНЕГО

Научная статья

Созаева Л.Т.^{1*}, Шунгаров И.Х.²

¹ ORCID : 0000-0002-9840-3566;

² ORCID : 0009-0009-3358-3699;

¹Высокогорный геофизический институт, Нальчик, Российская Федерация

²Российский университет дружбы народов, Москва, Российская Федерация

* Корреспондирующий автор (ljk_62[at]rambler.ru)

Аннотация

Предложена модель авторегрессии-скользящего среднего для прогнозирования значений метеорологических параметров атмосферы, в частности температуры атмосферного воздуха за осенний период г. Нальчика Юга России. Показано, что предлагаемая модель позволяет прогнозировать значения средних осенних температур воздуха с высокой точностью (6%). Во временном ряду средних осенних температур воздуха прослеживается цикл, характерный 11-летнему циклу солнечной активности. Все критерии качества модели удовлетворяют требованиям, предъявляемым к качеству и адекватности моделей прогноза. Предложенная модель может быть применена к анализу и прогнозированию средних температур воздуха за весенний, летний и зимний сезоны.

Ключевые слова: временные ряды, модель авторегрессии-скользящего среднего, метеопараметры, температура атмосферного воздуха.

A STUDY OF THE ATMOSPHERIC AIR TEMPERATURE REGIME USING THE AUTOREGRESSION-MOVING-AVERAGE MODEL

Research article

Sozaeva L.T.^{1*}, Shungarov I.K.²

¹ ORCID : 0000-0002-9840-3566;

² ORCID : 0009-0009-3358-3699;

¹High-Mountain Geophysical Institute, Nalchik, Russian Federation

²Peoples' Friendship University of Russia, Moscow, Russian Federation

* Corresponding author (ljk_62[at]rambler.ru)

Abstract

An autoregression-moving-average model for predicting the values of meteorological parameters of the atmosphere, in particular, the atmospheric air temperature for the autumn period of Nalchik, Southern Russia, is suggested. It is shown that the proposed model allows predicting the values of average autumn air temperatures with high accuracy (6%). The time series of mean autumn air temperatures shows a cycle characteristic of the 11-year cycle of solar activity. All quality criteria of the model meet the requirements for the quality and adequacy of forecast models. The offered model can be applied to the analysis and forecasting of average air temperatures for spring, summer and winter seasons.

Keywords: time series, autoregression-moving-average model, meteorological parameters, atmospheric air temperature.

Введение

Временная трансформация осредненных значений температуры атмосферного воздуха (годовые, осенние, зимние, весенние и летние) носит сложный характер и является одной из ключевых позиции не только потребностей климатологии, но и нужд народного хозяйства. Такие метеопараметры используются для принятия научно обоснованных решений по адаптации сельскохозяйственных культур к изменяющимся природно-климатическим условиям. Несмотря на широкое применение временных рядов для анализа и прогнозирования различных метеорологических параметров атмосферы [1], [3], [7], еще недостаточно изучена природа временных рядов выше отмеченных осредненных значений температуры воздуха.

Этим определяется актуальность детального анализа и прогнозирования временных рядов температуры атмосферного воздуха.

Целью работы является анализ и прогнозирование значений временных рядов метеорологических параметров. Для достижения которой поставлены следующие задачи:

- анализ временного ряда динамики значений средних температур воздуха за осенний период;
- выбор модели прогнозирования будущих значений временного ряда на ретроспективном участке;
- оценка качества модели с помощью показателей и критериев точности.

Материалом данного исследования являются значения средних температур атмосферного воздуха за осенний период, характерных для г. Нальчика Юга России.

Исследование проводится с помощью модели ARIMA, реализованной в пакете статистических программ SPSS [5].

Методы и методы исследования

Обычно для понимания изменения климата прогнозируется динамика климатических параметров. Что касается температурного режима, то это значения средних сезонных и годовых температур, а также экстремально высоких и низких температур воздуха. Материалом данного исследования являются многолетние значения температур,

осредненные за три осенних месяца (сентябрь, октябрь и ноябрь). Выбор этих параметров оправдан их применением в климатологии, особенно строительной, а также в сельском хозяйстве.

Объектом данного исследования является температурный режим г. Нальчика, расположенного в предгорной зоне Юга России. Средние показатели температуры воздуха в этом городе колеблются от +25...+28 °С – в июле, до –3...–5 °С – в январе. Среднегодовая температура воздуха составляет 9,6 °С. Самая высокая температура воздуха наблюдается в конце июля, а наиболее низкая – в конце января или в начале февраля. Возможны оттепели зимой до +15 °С и выше. Характерной особенностью климата города Нальчика являются значительные суточные колебания температуры, связанные с горно-долинным ветром.

Многолетние значения средних температур воздуха за осенний период по метеостанции г. Нальчика с 1961 по 2022 гг. были предоставлены Северо-Кавказским управлением по гидрометеорологии и мониторингу окружающей среды.

В ходе исследования временной ряд температуры атмосферного воздуха разбивается на две части: 1961-2016 гг. и 2017-2022 гг. Анализ ряда проводится для первой части, а для второй прогнозируются значения средних температур воздуха и сравниваются с фактическими значениями ряда (ретропрогнозирование).

Для данного исследования применяется одна из наиболее распространенных статистических моделей, получившая название модели Бокса-Дженкинса [4].

Согласно методологии Бокса-Дженкинса модель $ARIMA(p, d, q)$ состоит из трех частей [2], [4]: $AR(p)$ – модель авторегрессии порядка p , $MA(q)$ – скользящей средней порядка q и $I(d)$ – интегрированная часть порядка d (порядок последовательных разностей для преобразования временного ряда в стационарный).

Модель $ARIMA(p, d, q)$ в общем случае имеет вид:

$$Y_t = c + \varepsilon_t + \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} - b_1 \varepsilon_{t-1} - \dots - b_q \varepsilon_{t-q} \quad (1)$$

где c – константа модели;

$\alpha_1, \dots, \alpha_p$ – коэффициенты авторегрессии;

b_1, \dots, b_q – коэффициенты скользящего среднего;

ε_t – случайная ошибка или «белый шум».

Моделирование временного ряда при помощи $ARIMA(p, d, q)$ включает в себя следующие этапы: подготовка данных, выбор и идентификация модели (определение параметров p , q и d), оценка качества модели с помощью показателей и критериев точности и прогнозирование временного ряда на ретроспективном участке [6], [10].

Результаты исследования и обсуждение

Предварительное рассмотрение значений временного ряда показало отсутствие пропущенных значений и выбросов (данные наблюдений аномально далекие от других значений временного ряда). Далее временной ряд исследуется на стационарность. Динамика изменений средних температур воздуха (рис. 1а), свидетельствует о наличии тренда, причем имеет место его увеличение с течением времени, т.е. исходный временной ряд является нестационарным.



Рисунок 1 - Динамика изменения значений средних температур воздуха с нанесенным трендом

Примечание: а) исходный ряд; б) стационарный

Процедура взятия разности первого порядка $d = 1$ преобразовала исследуемый ряд в стационарный $ARIMA(p, 1, q)$ (рис. 1б), в результате появилась постоянная средняя, вокруг которой колеблются значения ряда.

Далее проводилась оценка параметров p и q модели по автокорреляционной и частной автокорреляционной функциям (рис. 2).

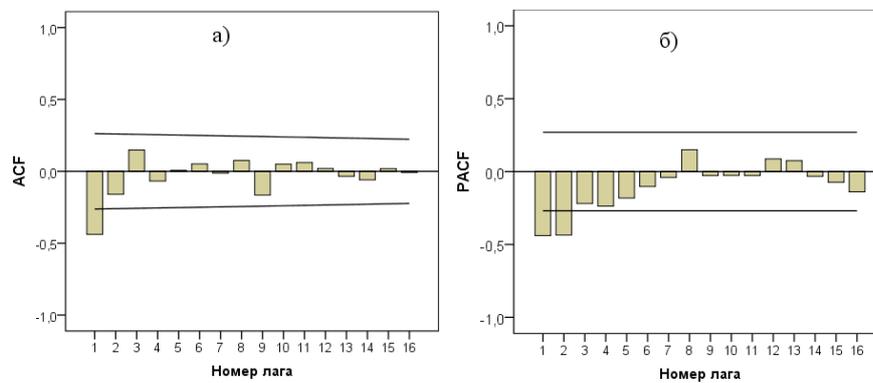


Рисунок 2 - Автокорреляционная ACF (а) и частная автокорреляционная PACF (б) функции

Порядок скользящей средней $MA(q)$ определялся как число значимых лагов автокорреляционной функции ACF (рис. 2а). Наиболее высоким оказался коэффициент автокорреляции первого порядка $MA(1)$. Порядок модели авторегрессии $AR(p)$ выбирался по графику PACF (рис. 2б), на котором выделяется лаг 2, как последний ненулевой элемент, т.е. имеем модели $AR(1)$ и $AR(2)$. Выбор пал на модели-кандидаты $ARIMA(1, 1, 1)$ и $ARIMA(2, 1, 1)$ с константами и без.

Значения ACF демонстрируют медленное затухание и наличие во временном ряду средних температур воздуха периодической составляющей. Величина периода колебания равна величине лага, при котором коэффициент корреляции выделяется. В данном случае это 9 лаг, но с учетом взятия разности первого порядка мы имеем 10-летний цикл. Не подлежит сомнению, что это – результат проявления 11-летнего цикла солнечной активности [8].

Отбор наилучшей из этих моделей проводился по:

- информационному критерию BIC;
- тесту Льюнга–Бокса на отсутствие автокорреляции в остатках;
- тесту Колмогорова–Смирнова на нормальность распределения остатков.

Нормализованный байесовский информационный критерий BIC [9], рассчитываемый в программе SPSS, предназначен для работы с выборками больших размеров. В случае небольших выборок рассчитывается вероятность того, что рассматриваемая модель сведет к минимуму потерю информации:

$$P = \exp((BIC_{min} - BIC_i)/2) \quad (2)$$

где BIC_{min} – самый низкий показатель BIC;

BIC_i – показателей моделей-кандидатов.

Из двух моделей предпочтение отдается модели с меньшим значением p .

Значения вероятностей для всех моделей-кандидатов, оказались равными $P = 0,01$.

Уровни значимости теста Льюнга–Бокса оказались очень высокие ($p > 0,05$) для всех моделей-претендентов (табл. 1), следовательно, можно принять нулевую гипотезу об отсутствии автокорреляции в остатках по всем моделям.

Отклонение остатков от нормального распределения считается существенным, если уровень значимости теста Колмогорова–Смирнова принимает значение $p < 0,05$. В рассматриваемом случае $p > 0,72$, т.е. остатки не отклоняются от нормального распределения.

Это свидетельствует о том, что все модели-кандидаты равнозначны.

Самыми популярными критериями выбора модели являются средняя абсолютная ошибка MAPE, характеризующая точность прогноза и коэффициент детерминации R^2 , показывающий доля дисперсии объясняемую моделью.

Анализ значений MAPE и R^2 отменил модели $ARIMA(1, 1, 1)$ с константой и без (табл. 1).

Таблица 1 - Критерии оценки моделей $ARIMA(p, d, q)$

Модель	Тест Льюнга-Бокса			MAPE		
	Статистика	Количество степеней свободы	p-уровень значимости			
1	2	3	4	5	6	
ARIMA(1,1,1)	с константой	7,77	16	0,96	9,56	0,45
	без	8,51		0,93	9,61	0,44
ARIMA(2,1,1)	с константой	7,56	15	0,94	9,37	0,46
	без	7,89		0,93	9,36	0,45

Таким образом, на этапе идентификации остались две модели-кандидаты $ARIMA(2, 1, 1)$ с константой и без константы.

Дальнейший выбор адекватной модели основывался на прогнозировании значений временного ряда на ретроспективном участке с 2017 по 2022 гг. (рис. 3). На рисунке 3 приведены фактические значения средних осенних температур воздуха и их прогнозные значения.

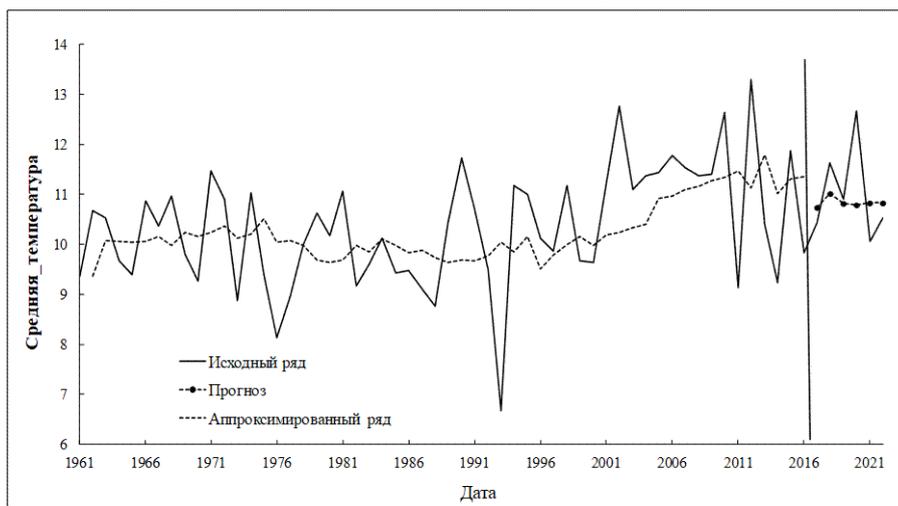


Рисунок 3 - Динамика значений средних осенних температур воздуха

Примечание: исходный ряд (сплошная кривая), аппроксимированный ряд (штриховая кривая) и прогнозные значения с 2017 по 2022 гг. (штрихпунктирная)

Были рассчитаны абсолютные ошибки модели прогноза на ретроспективном участке – MAPE (табл. 2) по фактическим и прогностическим значениям исследуемого ряда.

Таблица 2 - Результаты прогнозирования средних температур атмосферного воздуха на ретроспективном участке 2017-2022 гг. по модели $ARIMA(2,1,1)$

Дата	Фактические значения	Прогностические значения по модели		Абсолютные ошибки модели	
		с константой	без	с константой	без
1	2	3	4	5	6
2017	10,43	10,91	10,74	4,60	2,97
2018	11,63	11,18	11,02	3,90	5,27
2019	10,90	10,98	10,82	0,73	0,73
2020	12,67	10,98	10,79	13,32	14,82
2021	10,07	11,03	10,83	9,57	7,58
2022	10,53	11,05	10,83	4,91	2,82
Средняя абсолютная ошибка				6,17	5,70

Модель $ARIMA(2, 1, 1)$ без константы оказалась наилучшей, так как средняя абсолютная ошибка прогноза по этой модели оказалась минимальной и равной 5,70; сама процедура взятия первой разности временного ряда также отменяет константу модели.

Из литературы известно, что значение $MAPE < 10\%$ свидетельствует о высокой точности моделей прогнозов.

Заключение

Анализ динамики временного ряда значений средних температур воздуха за осенний период показал наличие периодичности, вызванной 11-летним периодом солнечной активности.

Средняя относительная ошибка прогноза на ретроспективном ряде, равная $\approx 6\%$, свидетельствует о высокой точности предлагаемой модели $ARIMA(2, 1, 1)$.

Анализ эффективности выбранной модели показал, что критерии качества прогнозирования значений средних осенних температур атмосферного воздуха по предлагаемой модели $ARIMA(2, 1, 1)$ удовлетворяет всем критериям и требованиям качества и адекватности, предъявляемым к моделям прогнозирования.

Это свидетельствует о том, что такой сложный по своей природе временной ряд, характеризующий температурный режим атмосферного воздуха может прогнозировать по модели Бокса-Дженкинса.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Алдошкина Е.С. Опыт использования аппарата нейронных сетей для анализа и прогноза временного ряда температуры воздуха / Е.С. Алдошкина, А.Д. Кузнецов, А.А. Пугачев [и др.] // Ученые записки Российского государственного гидрометеорологического университета. — 2009. — № 11. — С. 91-100
2. Алексеева В.А. Анализ временных рядов / В.А. Алексеева. — Ульяновск: УлГТУ, 2020. — 147 с.
3. Бисчоков Р.М. Пути решения проблем прогнозирования климата в КБР / Р.М. Бисчоков // Вестник Курганской ГСХА. — 2019. — № 3(31). — С. 54-58
4. Бокс Дж. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г. Дженкинс; Пер. с англ. А.Л. Левшина. — М.: Мир, 1974. — 406 с.
5. Бююль А. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей / А. Бююль, П. Цефель. — СПб.: ДиаСофтЮп, 2005. — 608 с.
6. Дуброва Т.А. Статистические методы прогнозирования: Учебное пособие для вузов / Т.А. Дуброва. — М.: ЮНИТИ-ДАНА, 2003. — 206 с.
7. Ефременко Д.С. Об одном алгоритме выявления локальных трендов при анализе метеорологических временных рядов / Д.С. Ефременко, А.Д. Кузнецов, О.С. Сероухова // Ученые записки Российского государственного гидрометеорологического университета. — 2016. — № 45. — С.132-141
8. Ривин Ю.Р. Циклы земли и Солнца / Ю.Р. Ривин. — М.: Наука, 1989. — 162 с.
9. Schwarz G. Estimating the Dimension of a Model / G. Schwarz // Annals of Statistics. — 1978. — Vol. 6. — P. 461-464.
10. ARIMA Model Coefficients // Real Statistics Resources. — URL: <https://real-statistics.com/time-series-analysis/arima-processes/arima-model-coefficients> (accessed: 25.05.23).

Список литературы на английском языке / References in English

1. Aldoshkina E.S. Opyt ispol'zovaniya apparata nejronnyh setej dlja analiza i prognoza vremennogo rjada temperatury vozduha [Experience in Using Neural Network Apparatus to Analyse and Forecast Air Temperature Time Series] / E.S. Aldoshkina, A.D. Kuznecov, A.A. Pugachev [et al.] // Uchenye zapiski Rossijskogo gosudarstvennogo gidrometeorologicheskogo universiteta [Scientific Notes of the Russian State Hydrometeorological University]. — 2009. — № 11. — P. 91-100 [in Russian]
2. Alekseeva V.A. Analiz vremennyh rjadov [Time Series Analysis] / V.A. Alekseeva. — Ulyanovsk: UISTU, 2020. — 147 p. [in Russian]
3. Bischokov R.M. Puti reshenija problem prognozirovaniya klimata v KBR [Ways to Address Climate Prediction Problems in the CBD] / R.M. Bischokov // Vestnik Kurganskoj GSHA [Bulletin of Kurgan State Agricultural Academy]. — 2019. — № 3(31). — P. 54-58 [in Russian]
4. Box J. Analiz vremennyh rjadov. Prognoz i upravlenie / J. Box, G. Jenkins; Transl. from Eng. by A.L. Levshin. — M.: Mir, 1974. — 406 p. [in Russian]
5. Bjujul' A. SPSS: iskusstvo obrabotki informacii. Analiz statisticheskikh dannyh i vosstanovlenie skrytyh zakonornostej [SPSS: The Art of Information Processing. Analysing Statistical Data and Recovering Hidden Patterns] / A. Bjujul', P. Cefel'. — SPb.: DiaSoftJup, 2005. — 608 p. [in Russian]
6. Dubrova T.A. Statisticheskie metody prognozirovaniya: Uchebnoe posobie dlja vuzov [Statistical Methods of Forecasting: textbook for universities] / T.A. Dubrova. — M.: UNITI-DANA, 2003. — 206 p. [in Russian]
7. Efremenko D.S. Ob odnom algoritme vyjavlenija lokal'nyh trendov pri analize meteorologicheskikh vremennyh rjadov [On One Algorithm for Detecting Local Trends in Meteorological Time Series Analysis] / D.S. Efremenko, A.D. Kuznecov, O.S. Serouhova // Uchenye zapiski Rossijskogo gosudarstvennogo gidrometeorologicheskogo universiteta [Scientific Notes of the Russian State Hydrometeorological University]. — 2016. — № 45. — P.132-141 [in Russian]
8. Rivin Ju.R. Cikly zemli i Solnca [The Earth and Sun Cycles] / Ju.R. Rivin. — M.: Nauka, 1989. — 162 p. [in Russian]
9. Schwarz G. Estimating the Dimension of a Model / G. Schwarz // Annals of Statistics. — 1978. — Vol. 6. — P. 461-464.
10. ARIMA Model Coefficients // Real Statistics Resources. — URL: <https://real-statistics.com/time-series-analysis/arima-processes/arima-model-coefficients> (accessed: 25.05.23).