

DOI: <https://doi.org/10.60797/IRJ.2024.149.129>

A STUDY OF THE ATMOSPHERIC AIR TEMPERATURE REGIME USING THE AUTOREGRESSION-MOVING AVERAGE MODEL (TRANSLATION OF THE ORIGINAL PUBLICATION IN ENGLISH)

Research article

Sozaeva L.T.^{1,*}, Shungarov I.K.²

¹ORCID : 0000-0002-9840-3566;

²ORCID : 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

Translation of the original publication Sozaeva L.T. Issledovanie rezhima temperatury atmosfernogo vozduha s pomoshh'ju modeli avtoregressii-skol'zjashhego srednego [Investigation of the atmospheric air temperature regime using an autoregression model-a moving average] / L.T. Sozaeva, I.H. Shungarov // Mezhdunarodnyj nauchno-issledovatel'skij zhurnal [International Scientific Research Journal]. – 2023. – №8 (134). DOI: 10.23670/IRJ.2023.134.94

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.

ИССЛЕДОВАНИЕ РЕЖИМА ТЕМПЕРАТУРЫ АТМОСФЕРНОГО ВОЗДУХА С ПОМОЩЬЮ МОДЕЛИ АВТОРЕГРЕССИИ-СКОЛЬЗЯЩЕГО СРЕДНЕГО (ПЕРЕВОД ОРИГИНАЛЬНОЙ ПУБЛИКАЦИИ НА АНГЛИЙСКИЙ ЯЗЫК)

Научная статья

Созаева Л.Т.^{1,*}, Шунгаров И.Х.²

¹ORCID : 0000-0002-9840-3566;

²ORCID : 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

Перевод оригинальной статьи Созаева Л.Т. Исследование режима температуры атмосферного воздуха с помощью модели авторегрессии-скользящего среднего / Л.Т. Созаева, И.Х. Шунгаров // Международный научно-исследовательский журнал. – 2023. – №8 (134). DOI: 10.23670/IRJ.2023.134.94

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.

Introduction

The temporal transformation of averaged atmospheric air temperature values (annual, autumn, winter, spring and summer) is complex and is one of the key positions not only for the needs of climatology, but also for the needs of the national economy. Such meteorological parameters are used to make scientifically-based decisions on the adaptation of agricultural crops to changing natural and climatic conditions. Despite the widespread use of time series for the analysis and forecasting of various meteorological parameters of the atmosphere [1], [3], [7], the nature of the time series of the above-mentioned averaged air temperature values is not yet sufficiently studied.

This determines the relevance of a detailed analysis and forecasting of time series of atmospheric air temperature.

The aim of the work is to analyze and forecast the values of time series of meteorological parameters. To achieve this, the following tasks are set:

- analysis of the time series of the dynamics of average air temperatures for the autumn period;
- selection of a forecasting model for future time series values in the retrospective section;

- evaluation of the quality of the model using accuracy indicators and criteria.

The material of this study is the values of average atmospheric air temperatures for the autumn period, characteristic of Nalchik, South Russia.

The study is carried out using the ARIMA model, implemented in the SPSS statistical software package [5].

Research methods and principles

Usually, to understand climate change, the dynamics of climatic parameters are predicted. As for the temperature regime, these are the values of average seasonal and annual temperatures, as well as extremely high and low air temperatures. The material of this study is the long-term values of temperatures, averaged over the three autumn months (September, October and November). The choice of these parameters is justified by their use in climatology, especially in construction, as well as in agriculture.

The object of this study is the temperature regime of Nalchik, located in the foothill zone of the South of Russia. The average air temperature indicators in this city range from $+25...+28$ °C – in July, to $-3...-5$ °C – in January. The average annual air temperature is $9,6$ °C. The highest air temperature is observed in late July, and the lowest – in late January or early February. There can be thaws in winter up to $+15$ °C and higher. A characteristic feature of the climate of the city of Nalchik is the significant daily temperature fluctuations associated with the mountain-valley wind.

The long-term values of the average air temperatures for the autumn period at the meteorological station of the city of Nalchik from 1961 to 2022 were provided by the North Caucasus Administration for Hydrometeorology and Environmental Monitoring.

In the course of the study, the time series of atmospheric air temperature is divided into two parts: 1961-2016 and 2017-2022. The analysis of the series is carried out for the first part, and for the second part, the values of the average air temperatures are predicted and compared with the actual values of the series (retroforecasting).

For this study, one of the most common statistical models, called the Box-Jenkins model, is used [4].

According to the Box-Jenkins methodology, the model $ARIMA(p,d,q)$ consists of three parts [2], [4]: $AR(p)$ – the autoregressive model of order p , $MA(q)$ – the moving average of order q and $I(d)$ – the integrated part of order d (the order of the successive differences for transforming the time series into stationary).

The model $ARIMA(p,d,q)$ in the general case has the form:

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

where c - model constant;

$\alpha_1, \dots, \alpha_p$ - autoregression coefficients;

b_1, \dots, b_q - moving average coefficients;

ε_t - random error or "white noise".

Modeling a time series using $ARIMA(p,d,q)$ includes the following stages: data preparation, model selection and identification (determining the parameters p , d and q), model quality assessment using accuracy indicators and criteria, and time series forecasting in the retrospective section [6], [10].

Main results

A preliminary examination of the values of the time series showed the absence of missing values and outliers (observation data abnormally far from other values of the time series). Further, the time series is investigated for stationarity. The dynamics of changes in average air temperatures (Fig. 1a) indicates the presence of a trend, and its increase over time, i.e. the original time series is non-stationary.

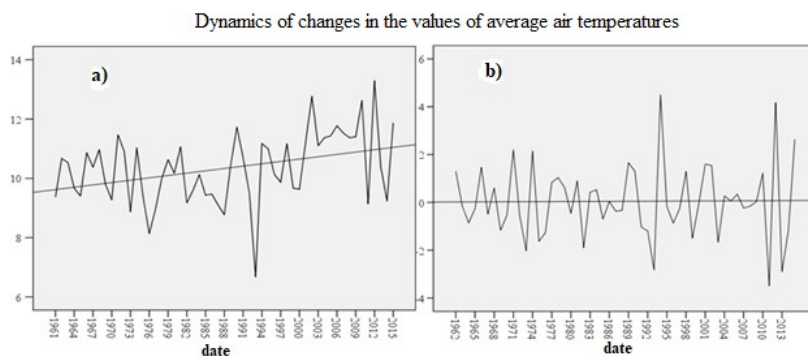


Figure 1 - Dynamics of changes in the values of average air temperatures with a trend applied:

a - original series; *b* - stationary

DOI: <https://doi.org/10.60797/IRJ.2024.149.129.1>

The procedure of taking the first-order difference $d=1$ transformed the studied series into a stationary $ARIMA(p,1,q)$ (Fig. 1b), resulting in a constant mean around which the series values fluctuate.

Further, the estimation of the p and q parameters of the model was carried out according to the autocorrelation and partial autocorrelation functions (Fig. 2).

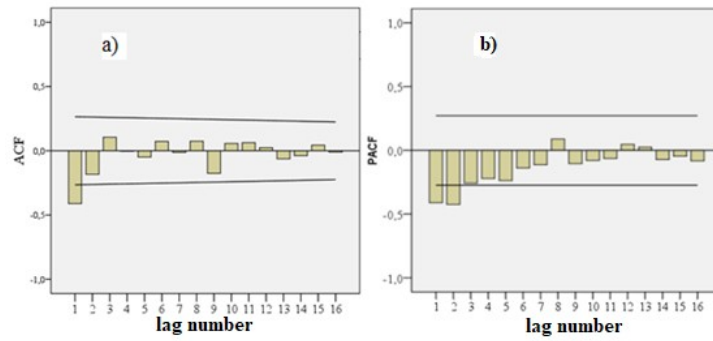


Figure 2 - Autocorrelation $ACF(a)$ and partial autocorrelation $PACF(b)$ functions
 DOI: <https://doi.org/10.60797/IRJ.2024.149.129.2>

The order of the moving average $MA(q)$ was determined as the number of significant lags of the autocorrelation function ACF (Fig. 2a). The highest turned out to be the autocorrelation coefficient of the first order $MA(1)$. The order of the autoregressive model $AR(p)$ was chosen from the $PACF$ plot (Fig. 2b), where the lag 2 is distinguished as the last non-zero element, i.e. we have the models $AR(1)$ and $AR(2)$. The choice fell on the candidate models $ARIMA(1,1,1)$ and $ARIMA(2,1,1)$ with and without constants.

The ACF values show a slow decay and the presence of a periodic component in the time series of average air temperatures. The period of the oscillation is equal to the lag at which the correlation coefficient is distinguished. In this case, this is lag 9, but taking into account the first-order difference, we have a 10-year cycle. There is no doubt that this is the result of the manifestation of the 11-year solar activity cycle [8].

The selection of the best of these models was carried out according to:

- BIC information criterion;
- $Ljung-Box$ test for the absence of autocorrelation in the residuals;
- $Kolmogorov-Smirnov$ test for normality of residual distribution.

The normalized Bayesian information criterion BIC [9], calculated in the SPSS program, is designed to work with large sample sizes. In the case of small samples, the probability that the model under consideration will minimize the loss of information is calculated:

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

where BIC_{min} – the lowest BIC indicator;

BIC_i – indicators of candidate models.

Of the two models, the one with the lower value p is preferred p .

The probability values for all candidate models turned out to be equal to $p=0,01$.

The significance levels of the $Ljung-Box$ test were very high ($p>0,05$) for all candidate models (Table 1), therefore, the null hypothesis about the absence of autocorrelation in the residuals can be accepted for all models.

The deviation of the residuals from the normal distribution is considered significant if the significance level of the $Kolmogorov-Smirnov$ test is $p<0,05$. In the case under consideration, $p>0,72$, i.e. the residuals do not deviate from the normal distribution.

This indicates that all candidate models are equivalent.

The most popular model selection criteria are the mean absolute percentage error $MAPE$, which characterizes the accuracy of the forecast, and the coefficient of determination R^2 , which shows the proportion of the variance explained by the model.

The analysis of the $MAPE$ and R^2 values ruled out the models $ARIMA(1,1,1)$ with and without a constant (Table 1).

Table 1 - Model evaluation criteria $ARIMA(p, d, q)$

DOI: <https://doi.org/10.60797/IRJ.2024.149.129.3>

Model		The Lewing-Box Test			MAPE	R^2
		Statistics	Number of degrees of freedom	p - significance level		
1		2	3	4	5	6
ARIMA (1,1,1)	with a constant	7.77	16	0.96	9.56	0.45
	without a constant	8.51		0.93	9.61	0.44
ARIMA (2,1,1)	with a constant	7.56	15	0.94	9.37	0.46
	without a constant	7.89		0.93	9.36	0.45

Thus, at the identification stage, two candidate models $ARIMA(2,1,1)$ with and without a constant remained.

Further selection of an adequate model was based on forecasting the time series values in the retrospective section from 2017 to 2022 (Fig. 3). Figure 3 shows the actual values of the average autumn air temperatures and their forecast values.

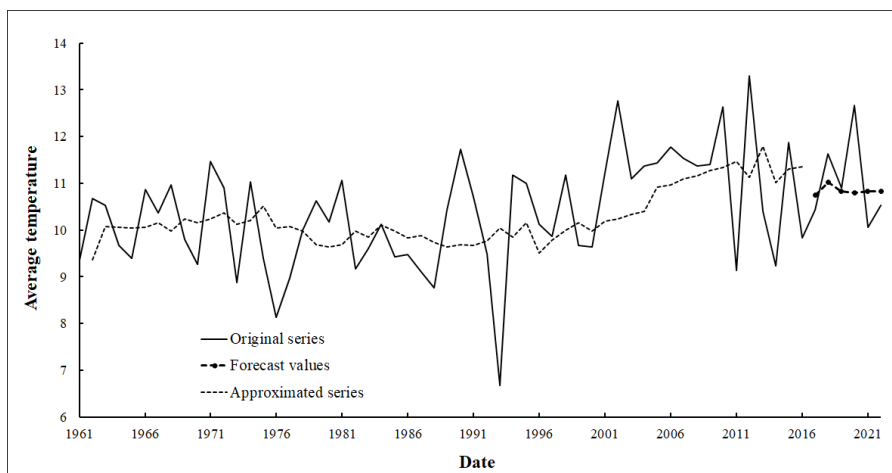


Figure 3 - Dynamics of the values of average autumn air temperatures
DOI: <https://doi.org/10.60797/IRJ.2024.149.129.4>

Note: original series (solid curve), approximated series (dashed curve) and forecast values from 2017 to 2022 (dash-dotted)

The absolute errors of the forecast model in the retrospective section – MAPE (Table 2) were calculated according to the actual and forecast values of the studied series.

Table 2 - Results of forecasting average atmospheric air temperatures in the retrospective section 2017-2022 according to the $ARIMA(2,1,1)$ model

DOI: <https://doi.org/10.60797/IRJ.2024.149.129.5>

Date	Actual values	Forecast values		Absolute forecast error	
		with a constant	without a constant	with a constant	without a constant
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
Mean absolute forecast error				6.17	5.70

The model $ARIMA(2,1,1)$ without a constant turned out to be the best, since the mean absolute forecast error for this model was minimal and equal to 5,70; the procedure of taking the first difference of the time series also cancels the model constant. From the literature, it is known that the $MAPE < 10\%$ indicates a high accuracy of forecast models.

Conclusion

Analysis of the dynamics of the time series of the values of average air temperatures for the autumn period showed the presence of periodicity caused by the 11-year solar activity cycle.

The average relative forecast error in the retrospective series, which is about 6%, indicates the high accuracy of the proposed model $ARIMA(2,1,1)$.

Analysis of the effectiveness of the selected model showed that the quality criteria for forecasting the values of average autumn atmospheric air temperatures using the proposed model $ARIMA(2,1,1)$ meet all the quality and adequacy criteria required for forecast models.

This indicates that such a complex in nature time series, characterizing the temperature regime of atmospheric air, can be predicted using the Box-Jenkins model.

Конфликт интересов

Не указан.

Рецензия

Моисеева Н.О., Санкт-Петербургский Государственный университет гражданской авиации имени Главного маршала авиации А.А. Новикова, Санкт-Петербург, Российская Федерация
DOI: <https://doi.org/10.60797/IRJ.2024.149.129.6>

Conflict of Interest

None declared.

Review

Moiseeva N.O., Saint-Petersburg State university of civil aviation named after Air Chief Marshal A.A. Novikov, Saint-Petersburg, Russian Federation
DOI: <https://doi.org/10.60797/IRJ.2024.149.129.6>

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: 25.05.2023).

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. — Ul'janovsk: UIGTU, 2020. — 147 p. [in Russian]
3. Bischokov R.M. Puti reshenija problem prognozirovanija 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. Boks J. Analiz vremennyh rjadov. Prognoz i upravlenie [Time series analysis. Forecasting and management] / J. Boks, G. Jenkins; transl. 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. Statistical data analysis and recovery of hidden patterns] / A. Bjujul', P. Cefel'. — SPb.: DiaSoftJup, 2005. — 608 p. [in Russian]
6. Dubrova T.A. Statisticheskie metody prognozirovanija: Uchebnoe posobie dlja vuzov [Statistical forecasting methods: A textbook for universities] / T.A. Dubrova. — M.: JuNITI-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 [Cycles of the Earth and the Sun] / 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: 25.05.2023).