




Predicting the average annual temperature of Tabriz city by Time series models

Nafiseh Rahimi^{1✉}  | Abdollah Faraji²

1. Department of Geography, Faculty of Humanities, Zanjan University, Zanjan, Iran. rahimi.nafiseh2@mail.com

2. Department of Geography, Faculty of Humanities, Zanjan University, Zanjan, Iran. abfaraji@znu.ac.ir

Article Information

Research Paper

Vol: 14
No: 51
P: 99-113
Received: 2022-07-14
Revised: 2023-03-01
Accepted: 2023-04-15
Published: 2023-04-21

Keywords:

- Forecast
- Tabriz
- Annual temperature
- Time series
- Model

Cite this Article:

Rahimi, N., Faraji, A. (2023). Predicting the average annual temperature of Tabriz city by time series models. *Journal of Arid Regions Geographic Studies* 14(51): 99-113.
 doi: 10.22034/jargs.2023.373963.0

Publisher: Hakim Sabzevari University

© The Author(s)



Abstract

Aim: The main purpose of this research was to model the temperature data of Tabriz synoptic station with the help of an integrated auto-correlated moving statistical model in time coverage of 1951-2020.

Material & Method: First, pre-processing was done on the studied data; in the next step, in order to analyze a time series and build a suitable model were exerted three Box-Cox methods, the first-order difference method and the least square method for stationery to decide on the best model based on the autocorrelation and partial autocorrelation diagrams, the minimum Akaic information criterion (AIC) and the Bayesian information criterion (BIC). Next, in order to ensure the appropriateness of the best-selected model were also applied normal probability diagrams, residuZ² 2als against time, autocorrelation and partial autocorrelation, Kolmogorov-Smirnov test and evaluation criteria (MAE), (MSE), (RMSE) and (NRMSE).

Finding: While time series models are a suitable method for modeling climatic parameters, by examining the efficiency of the Box-Cox, first-order differential, and least squares stationery methods for the temperature data of Tabriz, it was determined that the first-order differential method is the best stationery method for the annual temperature data of Tabriz synoptic station due to having the lowest slope of the fitting line and the complete elimination of autocorrelation between temperature values.

Conclusion: The results of this research showed that the ARIMA (0, 1, 1) Con model is the most suitable model for temperature forecasting of the coming years. Therefore, the average annual temperature of Tabriz synoptic station will increase in the next 18 years.

Innovation: In this research, the effectiveness of different stabilization methods was simultaneously evaluated for Tabriz Synoptic station in order to remove the trend, so that by evaluating the performance of different data stabilization methods and different models of the ARIMA family, the appropriate model for temperature forecasting can be determined and a clearer picture of the temperature conditions in the future can be obtained and its results can be exerted in related planning.

Extended Abstract

1. Introduction

The climate is a complex system that mainly various natural and human factors cause fluctuations in all components of the climate system, especially temperature. Temperature, as one of the most important and determining climatic elements, is a significant indicator in the climate classification that stability or variability in a geographical context has different effects in different dimensions, so recognizing and predicting this element is very important for management and planning of the future. Researchers use the concepts of time series in predicting climatic phenomena, especially temperature, due to its random nature; the most important type of time series is the ARIMA model. While recognizing the relationship between past and present values of the time series, we can predict the temperature of Tabriz's synoptic station for the next 18 years and get a clearer picture of temperature behavior in the future.

2. Materials and Methods

In the present study, the average annual temperature of Tabriz station in the period of 1951-2020 was received from the meteorological site of the whole country, and with the help of ARIMA models, temperature forecasting was done in several stages, the characteristics of each stage are as follows.

- **Step 1: Select and process the data**

Due to the real nature of the temperature time series, the first preprocessing was done on the studied data to check the outliers and other unacceptable values in the study time series and analyze.

- **Step 2: Prepare the data**

At this stage, after analyzing and identifying the instability of the Box-Cox methods, the first-order differential and the least squares method were used in data stationary, then in order to evaluate the normality of the statistical distribution of the static data in each method Stagnation, Kolmogorov-Smirnov test was used at 95% significance level.

- **Step 3: Modeling**

At this stage, the static time series were modeled according to the range of auto-regression changes and the moving average, and different models were tested by error and with varying degrees between zero, 1, and 2.

- **Step 4: Select the model**

After identifying the model and estimating its parameters, in the next step, the suitability of the model was determined by BIC and AIC. Then, the model residuals were examined by the normal diagram, ACF and PACF diagram, and Kalmogov-Smirnov test with a 95% confidence level.

Error metric indices

In order to examine the adequacy of the fitted model more accurately and predict the values (real data) and select the best method of predicting were used the ergo metric indices such as the mean absolute error value (MAE), the average Error square (MSE), root mean square root of error (RMSE), and root mean square of normalized error (NRMSE).

- **Step 5: Prediction**

The final criterion for selecting the final model was that the selected model would behave more appropriately for forecasting, that is, behave along the forecast horizon as it did in the recent past.

3. Results and Discussion

One of the critical conditions in the analysis of time series is data stationary. One of the essential factors that cause time series instability is the trend component, so in this study, a linear trend was identified. Time series instability was confirmed by drawing a diagram and regression, and Box-Cox, least square, and first-order differential method removed this inconsistency. Then, for each of the three staging methods, the values of d, p, and q were determined by PACF and ACF diagrams, and different patterns were examined by trial and error. In the next step, the first-order differential method was selected as the best stationing method, and the IM ARIMA (0, 1, 1) Con model was introduced as the final model. Then, we analyze the static residues by normal distribution diagram and Kolmogorov-Smirnov normality test of statistic (equal to 0.7 at 95% level) and ACF and PCF diagrams also indicate the suitability of IM ARIMA (0, 1, 1) Con patterns. The selected model among the studied models has the lowest value of AIC and BIC and the lowest value of absolute mean error (MAE), mean square error (MSE), root mean square error (RMSE), and root mean normalized error (NRMSE). So, the small error criteria for this template indicate that the values

predicted by this template are closer to the actual values of the variable series in question. Therefore, this model has a higher power of prediction. According to this model, the temperature will increase in the next 18 years, following the general characteristics of Tabriz temperature. It should be noted, however, that temperature prediction is intended to determine the values with the highest probability of occurrence and, in any way, does not imply the exact presentation of temperature values.

4. Conclusions

The main purpose of this research is to model the prediction of Tabriz synoptic location data by the ARIMA2 model with coverage of 1951-2020. For this purpose, according to the real nature of the time series, first t temperature preprocessing was done on the studied data, then for modeling and predicting the time series by ARIMA, the data is stationing by Box-Cox methods, one order differential and minimum squares for the data of synoptic station of Tabriz in order to increase the accuracy prediction of the model by selecting the most accurate and appropriate stationing method. The results showed that the first order differential method is the best method and the ARIMA (0, 1, 1) Con model is the most appropriate model considering the changes in auto-regression and moving average, Akaike and Bayesian evaluation criteria and residual analysis for predicting the temperature of the future. Based on the Predictive evaluation of the selected model, the average annual synoptic station in Tabriz for the next 18 years will have an increasing trend. However, due to the sharp change of temporal and spatial acceptance in temperature, prediction of exact value is impossible, but it also represents the probability of incidence and can be used for awareness, management, and planning.

5. Conflict of Interest

The authors declare no conflict of interest.

پیش بینی متوسط دمای سالانه شهر تبریز با استفاده از مدل های سری زمانی

نفیسه رحیمی^۱، عبدالله فرجی^۲

۱- نویسنده مسئول، گروه جغرافیا، دانشکده علوم انسانی، دانشگاه زنجان، زنجان، ایران. rahimi.nafiseh2@mail.com

۲- گروه جغرافیا، دانشکده علوم انسانی، دانشگاه زنجان، زنجان، ایران. abfaraji@znu.ac.ir

اطلاعات مقاله

مقاله پژوهشی

شماره: ۱۴

دوره: ۵۱

صفحه: ۹۹-۱۱۳

تاریخ دریافت: ۱۴۰۱/۰۴/۲۳

تاریخ ویرایش: ۱۴۰۱/۱۲/۱۰

تاریخ پذیرش: ۱۴۰۲/۰۱/۲۶

تاریخ انتشار: ۱۴۰۲/۰۲/۰۱

کلیدواژه‌ها:

- پیش بینی
- تبریز
- دمای سالانه
- سری زمانی
- مدل

نحوه ارجاع به این مقاله:

رحیمی، نفیسه، فرجی، عبدالله. (۱۴۰۲). پیش بینی دمای متوسط سالانه شهر تبریز با استفاده از مدل های سری زمانی. *مطالعات جغرافیایی مناطق خشک*، ۵۱(۱): ۹۹-۱۱۳.

doi: 10.22034/jargs.2023.373963.0

ناشر: دانشگاه حکیم سبزواری

© نویسنده(گان).



چکیده:

هدف: هدف اصلی این پژوهش، مدل سازی داده های دمای ایستگاه سینوپتیک تبریز به کمک مدل آماری میانگین متحرک خودهمبسته یکپارچه با پوشش زمانی ۱۹۵۱-۲۰۲۰ بود.

روش و داده: ابتدا پیش پردازش بر روی داده های مورد مطالعه صورت گرفت، در گام بعدی به منظور تجزیه و تحلیل یک سری زمانی و ساختن یک الگوی مناسب از سه روش باکس - کاکس، روش تفاضل گیری مرتبه یک و روش حداقل مربعات برای ایستاسازی استفاده شد تا بر اساس نمودارهای ACF و PACF، کمینه معیار اطلاع آکائیک (AIC) و معیار اطلاع بیزی (BIC) در مورد بهترین الگو تصمیم گیری شود. در ادامه به منظور اطمینان از مناسب بودن بهترین مدل انتخاب شده، از نمودارهای احتمال نرمال، نمودار باقی مانده ها در برابر زمان، ACF و PACF و آزمون کلموگروف - اسمیرنوف و معیارهای ارزیابی (MAE)، (MSE)، (RMSE) و (NRMSE) نیز استفاده گردید.

یافته ها: ضمن این که مدل های سری زمانی روش مناسبی در مدل سازی پارامترهای اقلیمی هستند، با بررسی کارایی روش های ایستاسازی باکس کاکس، تفاضلی مرتبه یک و حداقل مربعات برای داده های دما به صورت موردی برای ایستگاه سینوپتیک تبریز مشخص شد که روش تفاضلی مرتبه یک به دلیل دارا بودن کمترین شیب خط برازش و حذف کامل خودهمبستگی میان مقادیر دما بهترین روش ایستاسازی برای داده های دمای سالانه ایستگاه سینوپتیک تبریز است.

نتیجه گیری: نتایج این پژوهش نشان داد که مدل $ARIMA(0,1,1)_{Con}$ مناسب ترین مدل برای پیش بینی دما برای سال های آتی است. لذا دمای متوسط سالیانه ایستگاه سینوپتیک تبریز در ۱۸ سال آینده روند افزایشی خواهد داشت.

نوآوری، کاربرد نتایج: با توجه به پیشینه مطالعات صورت گرفته و ضرورت آگاهی از وضعیت دما در تحقیقات مختلف روش های منفرد متفاوتی از جمله روش های تفاضل گیری، حداقل مربعات، تبدیل باکس - کاکس که هر یک به روش خاصی عوامل ناپستایی را حذف می کنند برای حذف روند استفاده شده بود. اما در تحقیق پیش رو، کارایی روش های مختلف ایستاسازی به منظور حذف روند، به طور هم زمان به صورت موردی برای ایستگاه همدید تبریز مورد ارزیابی قرار گرفت، تا با ارزیابی عملکرد روش های مختلف ایستاسازی داده ها و مدل های مختلف خانواده ARIMA الگوی مناسب جهت پیش بینی دما تعیین گردد و تصویر واضح تری از شرایط دما در آینده حاصل گردد و نتایج آن بتواند در برنامه ریزی های مرتبط مورد استفاده قرار گیرد.

۱- مقدمه

در میان پارامترهای مختلف اقلیمی، دما یکی از مهم‌ترین عناصر اقلیمی است و بر روی سایر عناصر اقلیمی نقش تعیین‌کننده‌ای دارد، لذا ثبات و یا تغییرپذیری آن در یک بستر جغرافیایی می‌تواند ساختار آب‌وهوایی هر محل را دگرگون سازد و تأثیرات متفاوتی را در ابعاد مختلف به همراه داشته باشد. با توجه به این که فعالیت‌های انسانی اعم از کشاورزی، صنعت و سیستم‌هایی که وابسته به دما هستند، بر مبنای پایداری دما طراحی شده‌اند؛ لذا نوسان دما می‌تواند ثبات و پایداری ذکر شده را به هم بریزد و یا حتی مختل کند و انسان را با مشکلات جدیدتری روبه‌رو سازد. این مشکلات که عمدتاً ریشه در نوسانات اقلیمی دارند یکی از چالش‌های مهم قرن ۲۱ میلادی به‌ویژه در کشورهایی با اقلیم خشک است. کشور ایران در مناطق خشک و نیمه‌خشک و در منطقه اقلیمی و جغرافیایی خاصی از کره زمین واقع شده است و با توجه به رشد جمعیت و نیاز به محصولات کشاورزی افزایش تبخیر و تعرق و افزایش خشکسالی‌های اخیر، مسئله کمبود آب، تکثیر بیش از حد حشرات موزی و مسائلی از این دست که جملگی ریشه در افزایش دما دارند، مسئله افزایش دما و معضلات محیطی ناشی از آن مسائل متعددی را فراروی برنامه‌ریزان این کشور قرار داده است؛ لذا پیش‌بینی و برآورد دما به کمک مدل‌های اقلیمی در صورتی که از دقت کافی برخوردار باشد، می‌تواند تا حدودی با برنامه‌ریزی‌ها و مدیریت صحیح در کاهش این مشکلات و خسارت‌های ناشی از آن مفید واقع شود. مدل‌های اقلیمی که به‌منظور پیش‌بینی پارامترهای اقلیمی به کار برده می‌شوند عموماً در دو گروه کلی قرار دارند، مدل‌های فیزیکی که مبتنی بر مدل‌سازی فرآیندهای فیزیکی مؤثر در دما است و مدل‌های آماری - احتمالی مبتنی بر داده که با استفاده از داده‌های مشاهده شده تاریخی، مقدار دما در آینده را پیش‌بینی می‌کند. این مدل‌ها اصطلاحاً مدل‌های تحلیل سری‌های زمانی نامیده می‌شوند (Abbot & Marohasy, 2014). با توجه به این که مدل‌های سری زمانی، مدل‌هایی تجربی و ابزاری قدرتمند برای شبیه‌سازی و پیش‌بینی رفتار تصادفی عنصر اقلیمی دما است، لذا در صورت فراهم بودن شرایط، پیش‌بینی‌های صحیح و مطلوبی ایجاد خواهد کرد. مهم‌ترین نوع سری زمانی مدل ترکیبی اتو رگرسیون و میانگین متحرک، ARIMA است (Khazayi & Mirzayi, 2014). اساس عملکرد این الگو بر این فرض اولیه استوار است که رابطه مشخص و واضحی بین مقادیر آینده سری زمانی با مقادیر گذشته و فعلی سری زمانی مدل برقرار است. با توجه به اهمیت مدل‌سازی رفتار عناصر اقلیمی مطالعات متعددی در ایران و جهان به کمک روش‌های آماری انجام شده است که می‌توان به مدل‌سازی میانگین دمای سالانه شهر تبریز در بازه زمانی ۲۰۰۵-۱۹۵۱ با الگوسازی آریمای توسط عساکره اشاره کرد، وی بر اساس مدل منتخب آریمای $ARIMA(0,1,2)_{con}$ پیش‌بینی کرد که میانگین دمای سالانه شهر تبریز با ۰.۵ خطا روند افزایشی خواهد داشت (Asakere, 2014). رحیمی و غیور به مطالعه دبی رودخانه کارون با استفاده از روش‌های سری‌های زمانی در بازه زمانی ۱۳۷۸-۱۳۳۵ برای ایستگاه شالو، بارز و ارمند و بازه زمانی ۱۳۴۸-۱۳۷۸ برای ایستگاه مرغک پرداختند و مدل $ARIMA(1,4)$ برای ایستگاه شالو، $ARIMA(3,0,3)$ برای ایستگاه ارمند، $ARIMA(1,0)$ برای ایستگاه بارز و $ARIMA(0,0)$ را برای ایستگاه مرغک مناسب تشخیص دادند (Rahimi & Ghayor, 2009). سلیقه و همکاران با توجه به روند و چرخه‌های سری زمانی بارش سالانه ۳۷ ایستگاه حوضه‌های آبریز حله و مند که دارای آماری بیش از ۳۰ سال از بدو تأسیس تا سال ۲۰۱۱ بودند، به این نتیجه رسیدند که روند بارش در تمامی ایستگاه‌های حوضه مورد مطالعه کاهش یافته است (Saligheh et al., 2014). بارندگی سالانه ایستگاه‌های منتخب آذربایجان غربی توسط سلطانی گردفرامری و همکاران به‌منظور انتخاب بهترین مدل سری زمانی برای پیش‌بینی بارندگی سالانه با پوشش زمانی ۱۳۹۲-۱۳۶۳ بررسی شد و نتایج نشان داد که برای ایستگاه ارومیه مدل $ARIMA(1,0,0)$ و برای ایستگاه ماکو و مهاباد و مدل $ARIMA(0,1,1)$ مدل مناسبی بوده است (Soltani Gardframarzi et al., 2016). یونسی و همکاران به بررسی جریان متوسط ماهانه رودخانه بادآور - نورآباد شهرستان دلفان با استفاده از مدل‌های سری زمانی در بازه زمانی ۱۳۴۷-۱۳۹۵ پرداختند و به این نتیجه رسیدند که مدل برای دوره‌ی ۲۴ و ۳۶ ماهه، با ضریب نش سات کلایف ۰.۶۴ شبیه‌سازی مناسب‌تری در مقایسه با دوره‌های ۱۲ و ۳۶ ماهه داشته است (Younesi et al., 2017). پورکریم برآبادی و همکاران، نیز عملکرد مدل $ARIMA$ و $SARIMA$ را در پیش‌بینی میانگین دمای سالانه ایستگاه یزد و شیراز طی دوره ۲۰۱۷-۱۹۵۳ مورد بررسی قرار دادند، نتایج آن‌ها حاکی از آن بود که برای هر دو ایستگاه برازش مدل $ARIMA(0,1,2)$ مناسب بوده است، ولی مدل‌های $SARIMA(1,1,2)$ برای ایستگاه شیراز و مدل $SARIMA(0,1,2)$ برای ایستگاه یزد خارج از مرز معناداری بوده است (Pourkarim Barabadi et al., 2018). غفوریان و همکاران به مدل‌سازی داده‌های بارش ماهواره‌ای $TMPA(3B43)$ برای اقلیم‌های خشک و نیمه‌خشک ایران (فارس، اصفهان، کرمان، خوزستان) با استفاده از مدل آریمای با پوشش زمانی ۲۰۱۷-۱۹۹۸ پرداختند. نتایج پیش‌بینی بر اساس

آریما برای دوره چهار ساله بیانگر یک درصد کاهش میانگین بارش سالانه (به طور تفکیکی در مناطق کاشان، سیرجان، بهبهان و زرقان به ترتیب برابر با ۱.۳، ۱.۱، ۱، ۱.۸- درصد)، و برای بارش فصلی ثابت نسبی در فصل تابستان و کاهش در کاشان، سیرجان، زرقان و افزایش در بهبهان در فصل زمستان بود. ایمانی و همکاران، به بررسی بارش ماهانه ایستگاه سینوپتیک اردبیل با پوشش زمانی ۲۰۱۹-۱۹۹۰ با استفاده از مدل‌های تصادفی سری زمانی پرداختند و به این نتیجه رسیدند که مقدار بارندگی سالانه در ۴ سال ۵ سال آینده نسبت به میانگین بارندگی ۲۰ سال گذشته، بین ۳ تا ۱۷ درصد کاهش خواهد داشت که بیشترین کاهش مربوط به سال ۲۰۲۲ میلادی است (Imani et al., 2021).

میشرا و دیسای، مدل ARIMA و SARIMA را برای پیش‌بینی خشکسالی رودخانه کانسباتی^۱ هند (در ناحیه پورولیا در ایالت بنگال غربی در شرق هند)، در بازه زمانی ۲۰۰۱-۱۹۶۵ بکار بردند و به این نتیجه رسیده‌اند که این مدل‌ها برازش مناسبی را در بررسی مشاهدات داشته است، اما دقت مقادیر پیش‌بینی شده با افزایش میزان پیش‌بینی کاهش یافته، بنابراین مناسب‌ترین برآورد برای پیش‌بینی خشکسالی تا حدود دو ماه بوده است (Mishra & Desai, 2005). تکتاس، یک مطالعه مقایسه‌ای از مدل‌های شبکه‌های آماری و عصبی فازی برای پیش‌بینی آب و هوای گوزتپه^۲ استانبول انجام داد و از داده‌های میانگین دمای روزانه (خشک و مرطوب)، فشار هوا و سرعت باد در بازه زمانی ۲۰۰۸-۲۰۰۰ استفاده کرد و به این نتیجه رسید که مدل SNFIS نسبت به روش ARIMA با توجه به معیارهای RMSE, R², MAE نتایج بهتری را ارائه می‌دهد (Tektaş, 2010). پیش‌بینی بارش ماهانه (مارس، آوریل، می) شش ایستگاه سینوپتیک هند در بازه زمانی ۲۰۰۹-۱۹۴۹ توسط نارایانان صورت گرفت، آن‌ها نیز ضمن بررسی روند و شیب به کمک آزمون من - کندال و آزمون تیل و سن در غرب هندوستان، بر اساس پیش‌بینی به کمک آریما به این نتیجه رسیدند که روند بارش مونسون در دو ایستگاه (شهر اجمیر^۳-بیکانر^۴)، افزایشی خواهد بود (حدود ۱۰ درصد) (Narayanan, 2013). اجادی و همکاران، نیز میانگین دمای ماهانه شهر ظهران^۵ در عربستان سعودی را با استفاده از مدل‌های ARIMA برای سال‌های ۲۰۱۰-۱۹۵۱ مدل‌سازی کردند، و مدل (۰,۱,۱) (۲,۱,۱) ARIMA را به‌عنوان بهترین الگو انتخاب کردند (Ajadi et al., 2017). نیاتوام و آفودزو بارش و دما در حوضه توردیز در غنا را با استفاده از مدل ARIMA برای پوشش زمانی ۱۹۸۴-۲۰۱۴ پیش‌بینی کردند و بهترین مدل آریما برای بارش حوضه کپتو و توردیز^۶ را به ترتیب (۳,۱,۳) و (۳,۰,۳) اعلام کردند (Nyatume, Agodzo, 2018). همچنین البلاسمه، به پیش‌بینی و بررسی روند بارش (۲۰۱۹-۱۹۷۳) در حوضه وادی شعیب^۷ در اردن با آزمون من - کندال، شیب سن و ITA و ARIMA پرداختند، و به این نتیجه رسیدند که بارش شدید روند کاهش و باران با شدت کم روند افزایشی خواهد داشت (Al Balasmeh et al., 2019). آل سیاح و همکاران نیز تغییرات اقلیم را در حوضه آبخیز نهر ابراهیم^۸، در ۳۰۹ km^۲ لبنان، به کمک شاخص‌های خشکسالی طی دوره ۱۹۹۰-۲۰۱۸ با آریما بررسی کردند و به این نتایج رسیدند که میانگین، حداقل و حداکثر دما به ترتیب ۰.۹، ۰.۸، ۰.۷ افزایش و میانگین بارش ۶ میلی‌متر کاهش تا سال ۲۰۳۰ خواهد داشت (Al Sayah et al., 2020).

با توجه به پیشینه مطالعات صورت گرفته در ایران و خارج از ایران و ضرورت آگاهی از وضعیت دما در تحقیقات مختلف روش‌های منفرد متفاوتی از جمله روش‌های تفاضل‌گیری، حداقل مربعات، تبدیل باکس - کاکس که هر یک به روش خاصی عوامل ناپیوستایی را حذف می‌کنند برای حذف روند استفاده شده بود. اما در تحقیق پیشرو، کارایی روش‌های مختلف ایستاسازی به‌منظور حذف روند، به‌طور هم‌زمان به‌صورت موردی برای ایستگاه همدید تبریز مورد ارزیابی قرار گرفت، تا با ارزیابی عملکرد روش‌های مختلف ایستاسازی داده‌ها و مدل‌های مختلف خانواده ARIMA الگوی مناسب جهت پیش‌بینی دما با توجه به داده‌های ۷۰ ساله ایستگاه تبریز (۲۰۲۰-۱۹۵۱) برای ۱۸ سال آتی با روش‌های سری زمانی تعیین گردد تا تصویر واضح‌تری از شرایط دما در آینده حاصل گردد و نتایج آن بتواند در برنامه‌ریزی‌های مرتبط مورد استفاده قرار گیرد.

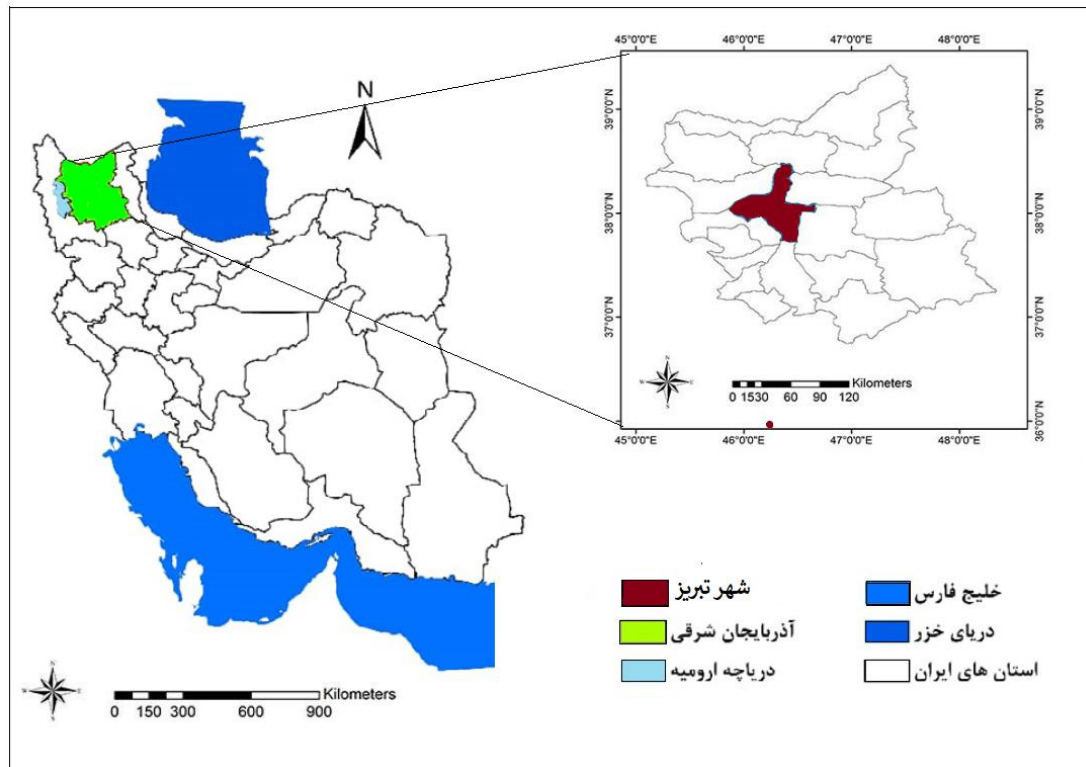
1. Kansabati
2. Göztepe
3. Ajmer
4. Bikaner

5. Dhahran
6. Kpetoe & Tordzinu
7. Wadi Shueib
8. Nahr Ibrahim

۲- مواد و روش

۲-۱- منطقه مورد مطالعه

منطقه مورد بررسی این مطالعه به صورت موردی برای ایستگاه همدید تبریز انجام گردید. شهر تبریز در غرب استان آذربایجان شرقی و در منتهی‌الیه شرق و جنوب شرق جلگه تبریز قرار گرفته است و از سمت شمال به کوه عینالی، از جنوب به دامنه کوه سهند، از سمت غرب به جلگه تبریز و دریاچه ارومیه محدود است. ارتفاع آن از سطح دریا بین ۱۳۰۰ تا ۲۱۰۰ متر و وسعت آن به طور تقریبی ۱۷۸۱ کیلومترمربع است. موقعیت کشوری و استانی ایستگاه مورد مطالعه در شکل ۱ نشان داده شده است.



شکل ۱. موقعیت جغرافیایی منطقه مورد مطالعه

۲-۲- روش پژوهش

• تشریح مدل‌های خانواده ARIMA

سری‌های زمانی داده‌های ترتیبی اند که به صورت گسسته در بازه‌ی زمانی مساوی جمع‌آوری شده‌اند. داده‌های سری زمانی تا زمان t به صورت $X = (X_1, X_2, \dots, X_{t-1}, X_t)$ می‌باشد، بنابراین پیدا کردن رابطه بین سری زمانی و ساختن یک الگوی مناسب در آنالیز سری زمانی اهمیت زیادی دارد. زیرا یکی از اهداف ساختن یک الگوی مناسب برای یک سری زمانی این است که بتوانیم مقدار آینده سری را نیز با توجه به اطلاعات گذشته آن پیش‌بینی کنیم (Ahmadi Shali et al., 2016). یکی از معروف‌ترین مدل‌ها برای پیش‌بینی سری زمانی، که توسط باکس و جنکینز در سال ۱۹۷۰ مطرح شد، مدل ARIMA با مرتبه p, q, d است و به صورت زیر است:

• فرایند میانگین متحرک $(MA(q))$

مقادیر مربوط به $MA(q)$ مرتبه‌ی میانگین متحرک را معلوم می‌سازد و به وسیله‌ی آن می‌توان وابستگی سری‌های اقلیمی را به عنصر تصادفی همان لحظه و زمان‌های مربوط به قبل از آن را تعریف کرد (مونتگمری و دیگران، ۱۹۷۹:۳۶۱). به عبارت دیگر، چنانچه Y_{t-1} و Y_{t-q} به ترتیب سری زمانی تصادفی نرمال و استاندارد در زمان $t-1$ و $t-q$ (مجموعه متغیرهای تصادفی که برحسب زمان مرتب شده‌اند) باشد، فرایند (Y_t) یک فرایند میانگین متحرک مرتبه q ، خواهد بود، و به صورت رابطه (۱) بیان می‌شود:

$$Y_t = C + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_q Y_{t-q} + e_t \quad \text{رابطه (۱)}$$

که در آن θ_i پارامترهای مدل، C مقدار ثابت مدل، e_t مقدار باقیمانده تصادفی در زمان t است (Papalaskaris et al., 2016).

• فرایند اتو رگرسیو (AR(p))

اساس این مدل بر پایه زنجیره مارکف در زنجیره زمانی بنا شده است و مقادیر مربوط به p در جمله AR در این مدل برحسب مقادیر پیشین خود تعیین می‌شود. در این فرایند (Y_t) متغیر پیش بینی شده در زمان t ام، C مقدار ثابت مدل و φ_i پارامترهای مدل، e_t مقدار باقیمانده تصادفی در زمان t ام است و به صورت زیر تعریف می‌شود: (Papalaskaris et al., 2016):

$$Y_t = C + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t \quad \text{رابطه (۲)}$$

• فرایند اتو رگرسیو میانگین متحرک مرکب (ARMA)

فرایند اتو رگرسیو - میانگین متحرک مرکب حاصل ادغام فرایندهای MA و AR است و به شکل زیر بیان می‌شود:

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + e_t - \theta_1 Y_{t-1} - \theta_2 Y_{t-2} - \dots - \theta_q Y_{t-q} \quad \text{رابطه (۳)}$$

در رابطه بالا Y_t متغیر پیش‌بینی شده در زمان t ام، $\varphi_1, \varphi_2, \dots, \varphi_p$ پارامترهای مدل AR، $\theta_1, \theta_2, \dots, \theta_q$ پارامترهای مدل MA، C مقدار ثابت مدل و e_t مقدار باقیمانده تصادفی در زمان t ام است (Papalaskaris et al., 2016)

• فرایند اتو رگرسیو میانگین متحرک یکپارچه (ARMA)

اگر d را به ARMA(p,q) اضافه کنیم آنگاه الگویی خواهیم داشت که قابلیت بیان بعضی از انواع سری‌های زمانی را دارد، چنین مدلی، "مدل یکپارچه" نامیده می‌شود، و ساختار آن به شکل رابطه زیر است:

$$\varphi_p(B) \nabla^d Y_t = \mu + \theta_q(B) e_t \quad \text{رابطه (۴)}$$

که در آن Y_t متغیر پیش‌بینی شده در زمان t ام، μ میانگین داده‌ها، e_t مقدار باقیمانده تصادفی در زمان t ام، ∇ عملگر تفاضل و B عملگر پسر است، مقادیر $\theta_q(B)$ و $\varphi_p(B)$ نیز به ترتیب با استفاده از روابط ۵ و ۶ تعیین می‌شود (Imani et al., 2021):

$$\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad \text{رابطه (۵)}$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad \text{رابطه (۶)}$$

• مراحل پیش‌بینی دما

مدل‌سازی دمای ایستگاه سینوپتیک تبریز به کمک الگوهای آماری پرکاربرد، یعنی الگوهای خانواده ARIMA در طی چندین مرحله انجام شد که جزئیات هر مرحله در زیر آمده است.

مرحله اول: انتخاب و پیش‌پردازش داده‌ها

در پژوهش حاضر میانگین دمای سالانه ایستگاه تبریز با پوشش زمانی ۱۹۵۱-۲۰۲۰ از سایت هواشناسی کل کشور دریافت گردید، سپس با توجه به ماهیت واقعی سری زمانی دما ابتدا پیش‌پردازش بر روی داده‌های مورد مطالعه صورت گرفت تا مقادیر پرت و سایر مقادیر غیر قابل قبول در سری زمانه مورد مطالعه بررسی شد.

مرحله دوم: آماده‌سازی داده‌ها

متدولوژی روش ARIMA برای پیش‌بینی سری‌های زمانی مبتنی بر این فلسفه است که داده‌ها باید ایستا باشند. لیکن با توجه به ماهیت سری زمانی دما نایستایی عمدتاً در بخش اصلی روند (تغییرات افزایشی و یا کاهشی خطی و یا اغلب غیرخطی (سهمی) در کل داده‌ها) وجود دارد، لذا در این پژوهش برای شناسایی نایستایی از نمودار و الگوسازی در خانواده الگوهای چندجمله‌ای استفاده شد (رابطه ۷)، تا به صورت بصری و علمی با امتحان و بررسی دقیق‌تر سری زمانی دید کلی در مورد وجود یا عدم وجود روند، نوع روند و... حاصل گردد.

$$Z_t = \beta_1 t + \beta_2 t^2 + \beta_k t^k + e_t \quad \text{رابطه (۷)}$$

Z_t متغیر پاسخ، β متغیر مجهول و e_t باقیمانده است (Asakere, 2009).

در گام بعدی، پس از تحلیل و شناسایی نایستایی روش‌های باکس کاکس، تفاضلی مرتبه یک و روش حداقل مربعات به دلیل کارایی بالا برای ایستاسازی داده‌ها به کار گرفته شد، سپس به منظور انتخاب روش مناسب ایستاسازی ضمن ارزیابی نرمال بودن توزیع آماری داده‌های ایستاده در هر یک از روش ایستاسازی با آزمون کلموگروف - اسمیرنوف (با سطح معنی‌داری ۹۵ درصد)، شیب معادله خطی سری ایستا شده و نمودارهای ACF برای هر سه روش مذکور بررسی گردید.

روش تبدیل توانی باکس - کاکس یکی از رایج‌ترین روش‌ها برای برطرف کردن نایستایی است که عمدتاً برای ایستاسازی در واریانس برای داده‌های مثبت و غیرصفر بکار گرفته می‌شود. در این روش λ عمدتاً بین ۲ و ۳- است و مقدار آن به گونه‌ای تعیین می‌شود که میانگین داده‌ها برابر با صفر و داده‌ها فاقد روند شوند و مطابق رابطه (۸) است:

$$y_t = \begin{cases} \frac{x_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(X_t) & \lambda = 0 \end{cases} \quad t = 1, 2, 3, \dots, N \quad \text{رابطه (۸)}$$

که در آن λ پارامتر تبدیل، X_t مقدار مشاهده شده در زمان t و y_t داده‌های تبدیل شده در زمان t است (Ahmadi Shali et al., 2016).

در ادامه روش حداقل مربعات با هدف ایستاسازی داده‌ها با حذف روند خطی سری زمانی دما مورد استفاده قرار گرفت (رابطه ۹)، در این روش Z_t روند داده در زمان t ، α شیب خط روند و β عرض از مبدأ خط روند و t داده‌ی مشاهده شده در زمان t ، است (Imani et al., 2021).

$$Z_t = \alpha t + \beta \quad \text{رابطه (۹)}$$

یکی دیگر از روش‌های رایج برای ایستاسازی سری‌های زمانی که در میانگین نایستا هستند، روش تفاضلی کردن است که در آن X_t مقدار مشاهده شده در زمان t ، $(\nabla^d X_t)$ روش تفاضل گیری مرتبه d ام از داده‌ها است. در ادامه روش‌های محاسبه آن در رابطه (۱۰) آورده شده است (Ahmadi Shali et al., 2016):

$$\begin{aligned} \nabla^d X_t &= X_t - X_{t-1} && \text{تفاضلات اولیه} \\ \nabla^2 X_t &= X_t - 2X_{t-1} + X_{t-2} && \text{تفاضلات ثانویه} \end{aligned} \quad \text{رابطه (۱۰)}$$

مرحله سوم: مدل‌سازی

در این مرحله سری‌های زمانی ایستا شده با توجه به محدوده تغییرات اتورگرسیون و میانگین متحرک، الگوسازی شدند، الگوسازی سری زمانی یک روش تکراری است که با شناخت الگو و برآورد پارامترها شروع می‌شود و مدل‌های مختلفی به صورت آزمون و خطا و با مرتبه‌های متغیر بین صفر، ۱ و ۲ انتخاب می‌شوند. بر پایه مراحل انجام شده چنانچه مدل پیشنهادی نامناسب باشد، باید مدل دیگری را در نظر گرفت، اما در صورتی که بعد از مراحل فوق به دو یا چند مدل مناسب دست یافتیم و تحلیل برای تشخیص این که کدام مدل بهتر است، کافی نبود، در آن صورت از روش‌های دیگری از جمله معیار اطلاعات آکائیک^۱ (AIC) و معیار اطلاع بیزی^۲ (BIC) برای شناسایی و انتخاب مدل مناسب باید استفاده گردد و در نهایت الگوی مناسب، الگویی خواهد بود که اضافه کردن و یا حذف هر پارامتری معنی‌داری الگوها را تحت تأثیر قرار ندهد. روابط ریاضی مربوط به دو معیار ذکر شده در رابطه (۱۱) آمده است:

$$AIC = n\{\text{Ins}_a^2\} + 2m \quad \text{رابطه (۱۱)}$$

$$BIC = n\{\text{Ins}_a^2\} + m\ln(n)$$

که در آن‌ها s_a^2 برآورد حداکثر درستی واریانس جامعه (σ_a^2) ، n طول سری و m تعداد پارامترهای مدل (بدون احتساب (s_a^2)) است (Asakere, 2009) معیار انتخاب بهترین مدل داشتن کمترین مقدار است (Marouphi et al., 2013).

مرحله چهارم: بررسی مناسبیت مدل

پس از تشخیص مدل و برآورد پارامترهای آن، در گام بعدی باید کیفیت مدل با بررسی این که آیا فرض‌های الگو برقرار هستند مورد ارزیابی قرار گیرد؛ لذا در این مرحله رسایی و مناسب بودن مدل با تجزیه و تحلیل باقیمانده‌های مدل بر اساس روش‌های ذیل مورد آزمون قرار گرفت (Ghafurian et al., 2019):

1. Akaike Information Criterion, AIC
2. Bayesian Information Criterion, BIC

۱. نرمال بودن مانده‌ها (باقیمانده‌ها): سنجش نرمال بودن و استقلال باقیمانده‌ها با به‌کارگیری نمودارهای احتمال نرمال و نمودار باقیمانده‌ها در برابر زمان صورت گرفت، همچنین برای بررسی دقیق‌تر کفایت مدل برازش داده شده، نرمال بودن باقیمانده‌های مدل، به کمک آزمون کلمورگروف - اسمیرنوف با سطح اطمینان ۹۵ درصد نیز بررسی شد.
۲. مقادیر باقیمانده مدل‌ها از نظر عدم وجود همبستگی با استفاده از نمودارهای (ACF) و (PACF) نیز بررسی شد Marouphi et al., 2013). چنانچه در نمودارها، حداکثر پنج درصد خودهمبستگی‌های باقیمانده‌ها در تأخیرهای متفاوت بیرون از فاصله اطمینان ۹۵ درصد واقع شوند، فرض تصادفی بودن باقیمانده‌ها در سطح پنج درصد رد نمی‌شود.
۳. شاخص‌های خطا سنجی: در ادامه برای بررسی دقیق‌تر کفایت مدل برازش داده شده برای پیش‌بینی مقادیر (داده‌های واقعی) و انتخاب بهترین روش پیش‌بینی از شاخص‌های ارزیابی عملکرد مدل همانند میانگین قدرمطلق خطا (MAE)، میانگین مجذور خطا (MSE)، جذر ریشه میانگین مربع خطا (RMSE)، و جذر ریشه میانگین خطای نرمال شده (NRMSE) استفاده شد که توابع معیار و روابط ریاضی آن‌ها به صورت زیر تعریف می‌شوند:

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |\hat{Y}_t - y_t| \quad \text{رابطه (۱۲)}$$

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (\hat{Y}_t - y_t)^2 \quad \text{رابطه (۱۳)}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (\hat{Y}_t - y_t)^2}{n}} \quad \text{رابطه (۱۴)}$$

$$\text{NRMSE} = \sqrt{\frac{\sum_{t=1}^n (\hat{Y}_t - y_t)^2}{y_{\max} - y_{\min}}} \quad \text{رابطه (۱۵)}$$

در روابط (۱۲) تا (۱۵) تعداد کل مشاهدات برای دوره پیش‌بینی، \hat{Y}_t و y_t به ترتیب ارزش پیش‌بینی شده و ارزش واقعی در زمان t است. y_{\max} و y_{\min} به ترتیب بیشترین و کمترین مقدار ارزش واقعی را نشان می‌دهد. کوچکی معیارهای خطا برای هر الگو نشانگر این نکته می‌باشد که مقادیر پیش‌بینی شده توسط این الگوها، به مقادیر واقعی نزدیک‌تراند و قدرت بالاتری در پیش‌بینی دارند (Khodaparast Shirazi et al., 2013).

مرحله پنجم: پیش‌بینی

بعد از تحلیل باقیمانده مدل‌های برازش شده، آخرین ملاک برای انتخاب مدل نهایی این بود که مدل منتخب، رفتار مناسب‌تری را برای پیش‌بینی ارائه دهد. در مدل‌های ARIMA خطی تنها بر اساس الگوی گذشته، آینده عناصر اقلیمی پیش‌بینی می‌شود؛ یعنی الگوی به دست آمده از مقادیر گذشته، در طول افق پیش‌بینی، همان رفتاری را خواهد داشت که در گذشته‌ی نزدیک داشته است در برخی مواقع، این فرض صحیح نبوده و فرایند مورد بررسی با زمان تغییر می‌کند. از این جهت، انتظار می‌رود دقت بازآفرینی و پیش‌بینی به سمت انتهای افق پیش‌بینی کاهش یابد با این وصف، مدلی که ثبات بیشتری در بازه اطمینان پیش‌بینی ارائه دهد، از مقبولیت بیشتری برخوردار است (Asakere, 2009).

۳- یافته‌ها

برای مدل‌سازی پیش‌بینی داده‌های دمای ایستگاه سینوپتیک شهر تبریز (۲۰۲۰-۱۹۵۱)، ابتدا پیش‌پردازش بر روی داده‌ها صورت گرفت و ایستایی سری‌های زمانی تحلیل شد. یکی از عوامل مهمی که سبب نایستایی سری زمانی می‌شوند، مؤلفه روند است، لذا در این بررسی برای به‌کارگیری روش مدل‌سازی ARIMA، ابتدا روند و نوع آن (الگوی خطی و سهمی) به کمک ترسیم نمودار و الگوسازی در خانواده چند جمله‌ای بررسی شد. بر اساس نمودار ترسیم شده، الگوی سری‌های زمانی مورد بررسی به صورت غیر افقی به سمت بالا نوسان داشتند و دارای روند افزایشی خطی بودند. با الگوسازی در خانواده چندجمله‌ای (رگرسیون) نیز خطی بودن روند الگوهای سری‌های زمانی آشکار گردید. در این الگوسازی با استفاده از نتایج حاصل از تحلیل رگرسیون مشخص گردید که مقدار $(p < 0; T > 2)$ است، لذا با سطح اطمینان ۰.۵ معنی‌دار الگوی خطی تأیید شد. الگوی خط برازش شده به سری زمانی طبق رابطه (۷) به صورت زیر است:

$$Y_t = 11.1 + 0.0357t + e_t$$

برای سنجش سهمی بودن الگوی خطی متغیر t^2 نیز به الگوی فوق برازش داده شد و با توجه به مقدار $(P > 0.05; T < 2)$ الگوی سهمی معنی‌داری نبود. لذا از بین دو الگوی فوق، الگوی خط برتری داشت از این رو $d = 1$ است.

$$Y_t = 11.4 + 0.0140t + 0.000305t^2 + e_t$$

با شناسایی روند در داده‌های دمای ایستگاه مذکور و تأیید نایستایی آن در مرحله بعد به ایستاسازی داده‌ها پرداختیم، بدین منظور ابتدا با آزمون باکس - کاکس داده‌ها ایستا شدند. در این آزمون مقدار λ برابر با $\lambda = 1.64$ تشخیص داده شد و داده‌های به توان λ رسیدند ($X_t = X_t^{1.64}$)، بدین ترتیب، عملیات تثبیت واریانس به کمک تبدیل باکس کاکس صورت گرفت. در قدم بعدی با تعیین شیب خط روند ($\alpha: 0.035$) و عرض از مبدأ آن ($\beta = 11.1$) ایستاسازی داده‌ها با روش حداقل مربعات انجام شد و در مرحله آخر روش تفاضلی با مرتبه اول به کار گرفته شد. معادله خطی (که در آن x شماره زمان و y مقدار سری ایستا شده است) و سطح معنی‌داری آزمون نرمال بودن (کلموگروف-اسمیرنوف) بعد از ایستاسازی برای هر سه روش مذکور بررسی شد (جدول ۱). سپس برای هر یک از سه روش ایستاسازی، مقادیر p و q به کمک نمودارهای PACF و ACF تعیین شد. در ایستایی‌سازی به روش باکس کاکس رتبه q طبق نمودار ACF پنج مورد و رتبه p در نمودار PACF دو مورد خارج از محدوده ۹۵ درصد اطمینان تعیین شد. در ایستایی‌سازی با روش حداقل مربعات و تفاضلی مرتبه یک نیز به ترتیب روش‌ها رتبه q طبق نمودار ACF پنج مورد و یک مورد و رتبه p در نمودار PACF دو و شش مورد خارج از محدوده ۹۵ درصد اطمینان تعیین شد. برای برازش جامع‌تر مدل‌ها، با تغییر مقدار d به $(d = 0)$ و $(d = 2)$ الگوهای مختلفی با آزمون و خطا بررسی شد. در قدم بعدی به منظور انتخاب بهترین روش ایستاسازی و مدل بهینه از میان انواع مدل‌های اجرا شده با توجه به معیار آکائیک و معیار اطلاع بیزی، هر سه روش ایستاسازی به کمک جدول ۱ و نمودار ACF بررسی شد. طبق جدول (۱) شیب معادله خطی سری ایستای حاصل از روش تفاضلی مرتبه یک در مقایسه با دو روش ایستاسازی دیگر بسیار کم و نزدیک به صفر یعنی حدود -0.0006 بود که نشان دهنده حذف مناسب روند از سری زمانی بود. آنالیز نمودارهای ACF هر سه روش ایستاسازی نیز نشان داد که تغییرات سالانه در سری مشاهداتی در روش تفاضلی مرتبه یک در مقایسه با دو روش دیگر تا حد قابل ملاحظه‌ای حذف شده است. لذا روش تفاضلی مرتبه یک بدلیل دارا بودن کمترین شیب خط برازش و حذف کامل خودهمبستگی میان مقادیر دما به عنوان بهترین روش ایستاسازی برای داده‌های دمای سالانه ایستگاه سینوپتیک تبریز انتخاب شد، در جدول ۲ به اختصار منتخبی از جزئیات الگوهای ایستا شده با روش تفاضلی مرتبه یک ارائه شده است.

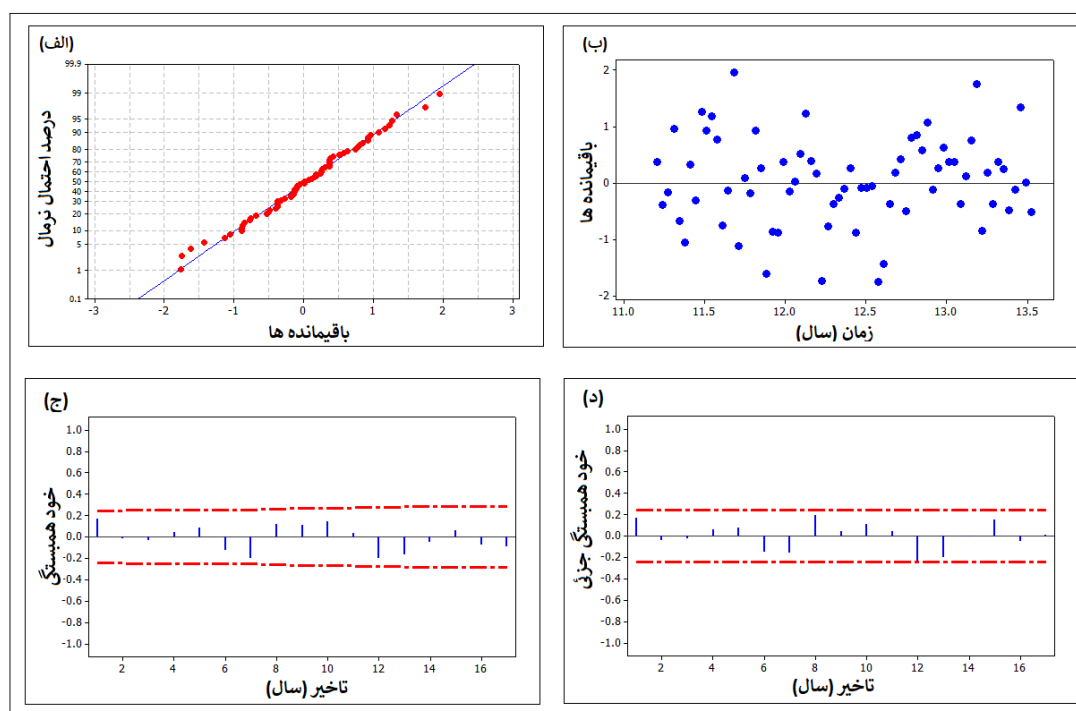
جدول ۱. معادله‌ی خط و سطح معنی‌داری سری‌های زمانی حاصل از روش‌های ایستاسازی مورد استفاده

روش ایستاسازی	معادله خط	سطح معنی‌داری آزمون نرمالیتبه بعد از ایستاسازی
تفاضل مرتبه یک	$y = -0.0006x + 0.0478$	۰.۰۹۲
باکس کاکس	$y = 0.1788x + 31.008$	۰.۰۶۶
حداقل مربعات	$y = 0.0344x - 0.3877$	۰.۰۷۷

جدول ۲. برازش مدل‌های تصادفی برای داده‌های ایستا شده میانگین سالانه دما تبریز (۲۰۲۰-۱۹۵۱)

مدل‌های انتخابی با تفاضل‌گیری مرتبه اول				مدل‌های انتخابی با تفاضل‌گیری مرتبه اول					
	مدل	پارامتر	T		مدل	پارامتر	T		
D1	ARIMA(0,1,1)c	MA1	۱.۰	۳۰	D0		-۰.۳	-۳.۷۳	
		Con	۰.۰۳	۵.۷			MA1	-۰.۷	-۶.۸۹
M1 M2	ARIMA(3,1,0)	AR1	-۰.۵۸	-۴.۹	M1	ARIMA(0,0,5)c	MA2	-۰.۴	-۳.۶۰
		AR2	-۰.۴۳	-۳.۳			MA3	-۰.۵	-۵.۴۲
		AR3	-۰.۲۹	-۲.۴			MA4	-۰.۷	-۷.۵۵
							MA5	۱۲.۶	۴۲.۸۹
		Con							
D2	ARIMA(0,2,2)	MA1	۱.۱	۳۹.۵	M2	ARIMA(1,0,0)c	AR	۰.۵۵	۵.۵
M3		MA	-۰.۲	-۲.۲				con	۵.۴۸
M4	ARIMA(4,2,1)	MA1	-۰.۹	-۳۳					
		AR1	-۲	-۱۷					
		AR2	-۱.۸	-۷.۷					
		AR3	-۱.۲	-۵.۱					
		AR4							

در پایان مدل نهایی با استفاده از معیار آکائیک و معیار اطلاع بیزی الگوی $ARIMA(0,1,1)_{con}$ تشخیص داده شد. در مرحله بعد به تجزیه و تحلیل باقیمانده‌های الگوهای ایستا شده با روش تفاضلی پرداختیم. تحلیل باقیمانده‌های ایستا شده با نمودار توزیع نرمال نشان داد که در الگوهای M1 نقاط در امتداد یک خط مستقیم حول یک محور با شیبی به سمت بالا در نوسان بودند. لذا باقیمانده‌ها از توزیع نرمال تبعیت می‌کنند (۲-الف)، نمودار باقیمانده‌ها در برابر زمان نیز فاقد ساختار خاصی هستند و روند خاصی را نشان نمی‌دهند (شکل‌های ۲-ب)، لذا می‌توان فرض ثابت بودن واریانس باقیمانده‌ها را پذیرفت. همچنین بررسی آماره آزمون نرمال بودن (کلموگروف-اسمیرنوف) در الگوی M1 برابر با ۰.۷ در سطح ۹۵ درصد معنادار بودند که این نشان‌دهنده برازش مناسب و کفایت این الگوها بود. بررسی همبستگی بین باقیمانده‌ها با نمودار ACF و PCF نیز نشان داد که در الگوهای M1 خودهمبستگی در مجذور باقیمانده‌ها وجود ندارد، بنابراین باقیمانده‌ها با سطح اطمینان ۹۵٪، تصافی، مستقل و نرمال بودند (شکل ۲-ج-د). ضمن این که الگوی M1 در بین ۶ مدل بررسی کمترین مقدار AIC و BIC را داشت و در نهایت برای ارزیابی عملکرد مدل منتخب، از دیگر شاخص‌های ارزیابی نظیر (MAE)، (MSE)، (RMSE)، (NRMSE) استفاده شد و مقایسه نتایج طبق جدول ۳ نشان داد که الگوی M1 دارای کمترین مقدار خطاست، لذا کوچکی معیارهای خطا برای این الگو نشانگر این نکته می‌باشد که مقادیر پیش‌بینی شده توسط این الگو به مقادیر واقعی (سری متغیر مورد نظر) نزدیک‌تر می‌باشند. بنابراین این الگو قدرت بالاتری در پیش‌بینی دارد.

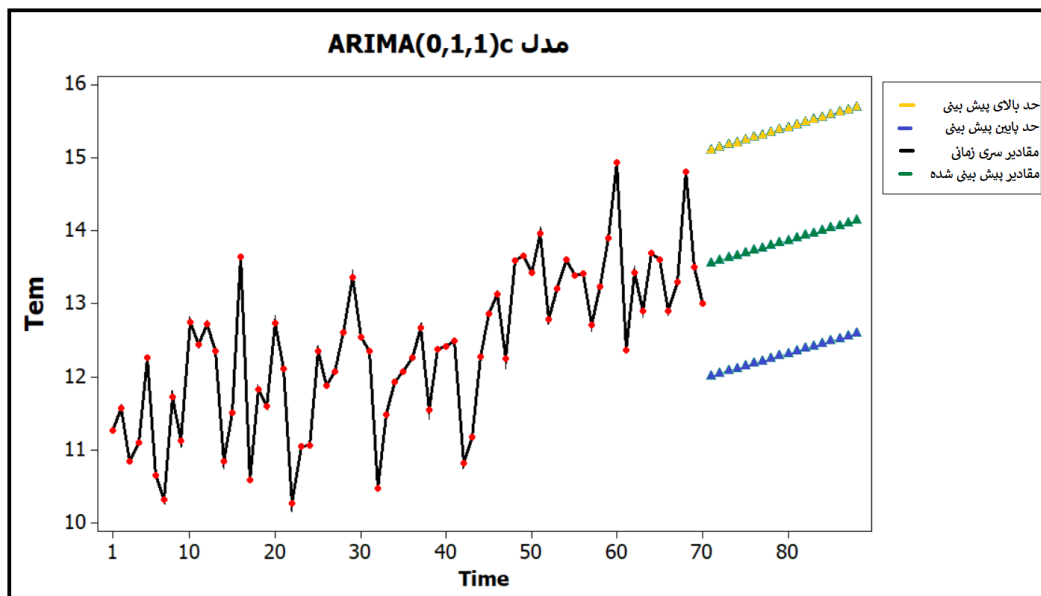


شکل ۲. الف) منحنی احتمال نرمال باقیمانده‌ها: ب) باقیمانده‌ها در برابر زمان ج) ACF باقیمانده‌ها د) PCF باقیمانده‌ها مدل $ARIMA(0,1,1)_c$

جدول ۳. مقایسه‌ی توابع خطا برای میانگین دمای سالانه ایستگاه تبریز.

الگوها	MAE	MSE	RMSE	NRMSE	AIC	BIC
ARIMA(0,1,1)con	۰.۰۹	۰.۰۶	۰.۷	۰.۱	-۲۸۸	-۲۴.۳
ARIMA(3,1,0)	۰.۱	۰.۷	۰.۸۴	۰.۱۵	-۱۳.۹	-۷.۲۷
ARIMA(0,2,2)	۰.۱۱	۰.۹	۰.۹۵	۰.۱۹	-۰.۶۶	۳.۷۷
ARIMA(4,2,1)	۰.۱۲	۱.۱۶	۱.۳۳	۰.۲۵	۲۵.۶۳	۳۶.۷
ARIMA(0,0,5)con	۰.۰۸	۰.۵۳	۰.۷	۰.۱	-۲۴.۹	-۱۱.۴
ARIMA(1,0,0)con	۰.۱	۰.۷۷	۰.۸	۰.۱۶	-۱۲.۲	-۷.۷۴

طبق شکل ۳ ترسیم مدل پیش‌بینی برای الگوی M1 نشان‌دهنده این است که دما در ۱۸ سال آتی روند افزایشی خواهد داشت که این نوع رفتار مطابق با ویژگی‌های عمومی دمای تبریز می‌باشد. هرچند ذکر این نکته لازم است که منظور از پیش‌بینی دما تعیین مقادیر با بیشترین احتمال وقوع است و به هیچ وجه به معنی ارائه دقیق مقادیر دما نیست.



شکل ۳. پیش‌بینی ۱۸ ساله با فاصله اطمینان ۹۵ درصد برای الگوی $ARIMA(0,1,1)$

۴- بحث و نتیجه‌گیری

هدف اصلی این پژوهش، مدل‌سازی داده‌های دمای ایستگاه سینوپتیک تبریز با استفاده از مدل $ARIMA$ با پوشش زمانی ۲۰۲۰-۱۹۵۱ بود. بدین منظور با توجه به ماهیت واقعی سری زمانی دما ابتدا پیش‌پردازش بر روی داده‌های مورد مطالعه صورت گرفت تا مقادیر پرت و سایر مقادیر غیر قابل قبول در سری زمانی مورد مطالعه شناسایی شود. سپس با شناسایی نالیستایی در سری زمانی مورد مطالعه به منظور مدل‌سازی و پیش‌بینی سری زمانی با استفاده از $ARIMA$ ، اقدام به ایستاسازی داده‌ها شد. ایستاسازی داده‌ها با روش‌های مختلفی صورت می‌گیرد و هر کدام از این روش‌ها با تکنیک‌های خاصی بخش‌های قطعی داده‌ها را جداسازی می‌کند و دقت پیش‌بینی‌ها را متأثر می‌سازد. عمدتاً در مطالعات قبلی، از یک نوع روش ایستاسازی داده‌ها استفاده می‌شد. اگرچه ممکن است استفاده از روش‌های ایستاسازی منفرد دقت قابل قبولی در پیش‌بینی ارائه می‌دهند، اما بررسی سایر روش‌های ایستاسازی و تعیین دقیق‌ترین روش نیز خالی از لطف نیست. از این‌رو، در این مطالعه کارایی روش‌های ایستاسازی باکس کاکس، تفاضلی مرتبه یک و حداقل مربعات برای داده‌های دما به صورت موردی برای ایستگاه سینوپتیک تبریز مورد ارزیابی قرار گرفت و روش تفاضلی مرتبه یک به دلیل دارا بودن کمترین شیب خط برازش و حذف کامل خودهمبستگی میان مقادیر دما به عنوان بهترین روش ایستاسازی برای داده‌های دمای سالانه ایستگاه سینوپتیک تبریز انتخاب شد. نتایج بررسی‌ها با توجه محدود تغییرات اتورگرسیون و میانگین متحرک، معیارهای ارزیابی آکائیک و بی‌زی و تحلیل باقیمانده‌ها، نشان داد که مناسب‌ترین مدل برای پیش‌بینی دما برای آینده، مدل $ARIMA(0,1,1)_{Con}$ است. بر اساس پیش‌بینی حاصل از الگوی منتخب $ARIMA(0,1,1)_{Con}$ در این مطالعه، دمای متوسط سالیانه ایستگاه سینوپتیک تبریز در ۱۸ سال آینده روند افزایشی خواهد داشت. نتایج این مطالعه با یافته‌های بسیاری از محققین از جمله یار احمدی و همکاران در بررسی دمای سالانه ایستگاه سینوپتیک تهران با استفاده از مدل سری‌های زمانی و عساکره در بررسی میانگین سالانه دمای شهر تبریز با الگوسازی آرما همسو است و هر دو این محققین ضمن پیش‌بینی افزایش دما برای سال‌های آتی به این نتیجه رسیدند که مدل‌های سری زمانی روش مناسبی در مدل‌سازی پارامترهای اقلیمی است (Yar Ahmadi et al., 2014; Asakere, 2009). به طور کلی، نتایج پیش‌بینی با توجه به تغییرپذیری شدید زمانی و مکانی پارامتر دما، پیش‌بینی دما به

هیچ رو به‌منظور ارائه مقادیر دقیق برای دمای آینده نیست؛ بلکه بیانگر احتمال وقوع است و می‌تواند به‌منظور آگاهی، مدیریت و برنامه‌ریزی، مورد استفاده قرار گیرد.

۵- فهرست منابع

- احمدی شالی، جعفر و وصفی، مهدی (۱۳۹۶). پیش‌بینی نقدینگی بر اساس برآورد نقطه‌ای و بازه‌ای روش آریمای و مقایسه آن با روش هموارسازی نمایی دوگانه، اقتصاد مالی (اقتصاد مالی و توسعه)، ۱۱ (۴۰)، ۱۷۵-۱۵۹.
- ایمانی، رسول؛ قضاوی، رضا و اسماعیلی اوری، اباذر (۱۴۰۰). بررسی روند، تجزیه و تحلیل، مدل‌سازی و پیش‌بینی بارش ماهانه با استفاده از مدل‌های تصادفی سری زمانی (مطالعه‌ی موردی: ایستگاه سینوپتیک اردبیل)، مطالعات جغرافیایی مناطق خشک، شماره ۱۱ (۴۴)، ۹۸-۸۴.
- پورکریم برآبادی، رویا و حیدری منفرد، زهرا (۱۳۹۸). ارزیابی مدل‌های سری‌های زمانی جهت پیش‌بینی متوسط دما در نیمه جنوبی ایران، دگرگونی‌ها و مخاطرات آب‌وهوایی، ۱ (۲)، ۱۸۹-۱۶۴.
- خداپرست شیرازی، جلیل و صادقی، زهرا (۱۳۹۳). مدل‌سازی و پیش‌بینی صادرات آبزیان دریایی در ایران با استفاده از ARIMA و شبکه‌های عصبی مصنوعی، فصلنامه اقتصاد مقداری، ۹ (۴)، ۱۳۲-۱۱۱.
- خزایی، مجید و میرزایی، محمدرضا (۱۳۹۳). پیش‌بینی متغیرهای اقلیمی به کمک تحلیل سری‌های زمانی حوضه آبخیز زهره، نشریه تحقیقات کاربردی علوم جغرافیایی، ۱۴ (۳۴)، ۲۵۰-۲۳۳.
- رحیمی، داریوش و غیور، حسنعلی (۱۳۸۹). تحلیل دبی رودخانه کارون با تبدیل باکس - کاکس و سری‌های زمانی، تحقیقات جغرافیایی، ۲۵ (۴)، ۱۵۱-۱۳۵.
- سلطانی گردفرامرزی، سمیه؛ صابری، عارف و قیصوری، مرتضی (۱۳۹۶). تعیین بهترین مدل سری زمانی در پیش‌بینی بارندگی سالانه ایستگاه‌های منتخب استان آذربایجان غربی، نشریه تحقیقات کاربردی علوم جغرافیایی، ۱۷ (۴۴)، ۱۰۵-۸۷.
- سلیقه، محمد؛ عساکره، حسین؛ ناصرزاده، محمد حسین و بلیانی، یدالله (۱۳۹۴). تحلیل روند و چرخه‌های سری زمانی بارش سالانه حوضه‌های آبریز حله و مند، تحقیقات کاربردی علوم جغرافیایی، ۱۵ (۳۷)، ۲۷۲-۲۴۵.
- عساکره، حسین (۱۳۸۸). الگوسازی ARIMA برای میانگین سالانه دمای شهر تبریز، تحقیقات جغرافیایی، ۹۳، ۲۴-۳.
- غفوربان، هادی؛ ثنایی نژاد، سیدحسین و جباری نوقایی، مهدی (۱۳۹۹). ارزیابی مدل‌های سری زمانی در پیش‌بینی بارش فصلی مبتنی بر داده‌های سنجش از دور (مطالعه موردی: اقلیم‌های خشک و نیمه‌خشک)، نشریه پژوهش‌های اقلیم‌شناسی، ۱۱ (۴۲)، ۹۴-۷۷.
- معروفی، صفر؛ ختار، بهناز؛ صادقی‌فر، مجید؛ پارسا، نصیرالدین و ایلدورمی، علیرضا (۱۳۹۳). پیش‌بینی خشکسالی با استفاده از سری زمانی SARIMA و شاخص SPI در ناحیه مرکزی استان همدان، پژوهش آب در کشاورزی (علوم خاک و آب)، ۲۸ (۱)، ۲۲۵-۲۱۳.
- یونسی، حجت‌الله؛ ترابی پوده، حسن؛ ارشیا، آزاده و میرزاپور، حافظ (۱۳۹۷). شبیه‌سازی جریان متوسط ماهانه رودخانه بادآور-نورآباد با استفاده از مدل‌های سری زمانی، ششمین کنفرانس علمی پژوهشی مدیریت منابع آب و خاک، کرمان.

References

- Abbot J and Marohasy J, 2014. Input selection and optimization for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. Atmospheric Research 138: 166-178. <https://doi.org/10.1016/j.atmosres.2013.11.002>.
- Ahmadi Shali, Jafar and Wasfi, Mehdi (2016). Liquidity forecasting based on the point and interval criteria of Arima method and its comparison with the dual representation smoothing method, Financial Economy (Financial Economics and Development), 11 (40), 159-175. <https://civilica.com/doc/1570242> [in persian]
- Al Balasmeh, O., Babbar, R., & Karmaker, T. (2019). Trend analysis and ARIMA modeling for forecasting precipitation pattern in Wadi Shueib catchment area in Jordan. Arab J Geosci, 12(27), 1-19. DOI:10.1007/s12517-018-4205-z.

- Al Sayah, M.J., Abdallah, C., Khouri, M., Nedjai, R., & Darwich, T. (2021). A framework for climate change assessment in Mediterranean data-sparse watersheds using remote sensing and ARIMA modeling. *Theor Appl Climatol*, 143, 639–658. DOI:10.1007/s00704-020-03442-7
- Borland, P., & Montana, A. (1996). Forecasting of storm rain fall by combined use of rider, rain gages and linear models. *Atmospheric research*, 42(1), 199-216. Doi: 10.1016/0169-8095(95)00063-1.
- Brockwell P. J., Davis, R. A. (2009). *Time series: theory and methods*, Springer, 44-79. <https://db-nb.info/949251941/04>.
- Ghafurian, Hadi; Sanajnejad, Sayyed Hossein and Jabari Nougabi, Mehdi (2019). Evaluation of time series models in forecasting seasonal precipitation based on remote sensing data (case study: arid and semi-arid climates), *Journal of Climatology Research*, 11 (42), 77-94. https://clima.irimo.ir/article_125162.
[in persian]
- Imani, Rasoul; Qazawi, Reza and Esmaili Avari, Abazar (1400). The process of investigation, analysis, modeling and prediction of monthly rainfall using stochastic time series models (case study: Ardabil synoptic station), *Geographical Studies of Dry Areas*, No. 11(44), 84-98. <https://civilica.com/doc/1405291>. [in persian]
- jadi, N., Ajadi, J., Damisa, S., Asiribo, O., & Dawodu, A. (2017). Modeling Monthly Average Temperature of Dhahran City of Saudi-Arabia Using Arima Models. *International Journal of Data Science and Analysis*, 3(5), 40-45. DOI:10.11648/j.ijdsa.20170305.12.
- Khodaparast Shirazi, Jalil and Sadeghi, Zahra (2013). Modeling and natural forecasting of marine aquatics in Iran using ARIMA and artificial neural networks, *Quarterly of Quantitative Economics*, 9(4), 111-132. magiran.com/p1629509. [in persian]
- Marouphi, Safar; Khatar, Behnaz; Sadeghifar, Majid; Parsafar, Nasr-od-din and Ildurmi, Alireza (2013). Drought prediction using SARIMA time series and SPI index in the central region of Hamedan province. *Water Research in Agriculture (Soil and Water Sciences)*, 28(1), 225-213. SID. <https://sid.ir/paper/487191/fa>. [in persian]
- Mishra, A. K., & Desai, V. R. (2005). Drought forecasting using stochastic models. *Int. J. Climat*, 19(16), 585 – 599. DOI:10.1007/s00477-005-0238-4.
- Narayanan, P., Basistha, A., Sarkar, S., & Sachdeva, K. (2013). Trend analysis and ARIMA modelling of pre-monsoon rainfall data for western India. *Comptes Rendus Geoscience*, 345(1), 22-27. <https://doi.org/10.1016/j.crte.2012.12.001>.
- Nyatuame, M., Agodzo, S. K. (2018). Stochastic ARIMA model for annual rainfall and maximum temperature forecasting over Tordzie watershed in Ghana. *J. Water Land Develop*, 37(1), 127–140. DOI:10.2478/jwld-2018-0032.
- Papalaskaris, T., Panagiotidis, T., & Pantrakis, A. (2016). Stochastic monthly rainfall time series analysis modeling and forecasting in Kavala city, Greece, north-eastern mediterranean basin. *Procedia Engineering*, 162, 254-263. DOI:10.1016/j.proeng.2016.11.054.
- Pourkarim Barabadi, Roya and Heydari Monfared, Zahra (2018). Evaluation of time series models for forecasting the average temperature in the southern part of Iran, *Climate Change and Hazards*, 1(2), 164-189. https://cccd.znu.ac.ir/article_44668. [in persian]
- Rahimi, Dariush and Ghayor, Hassan Ali (2009). Discharge analysis of Karun River with Box-Cox transformation and time series, *Geographical Research*, 25(4), 151-135. <https://sid.ir/paper/29726/fa> [in persian]
- saligheh, Mohammad; Asakere, Hossein; Naserzadeh, Mohammad Hossein and Belyani, Yadullah (2014). Analyzing trends and cycles of annual time series of Helleh and Mand watersheds, *Applied Research of Geographical Sciences*, 15(37), 272-245. <http://jgs.khu.ac.ir/article-1-2428-fa>. [in persian]
- Soltani Gardframarzi, Somaiyeh. Saberi, Aref and Qaisuri, Morteza (2016). The best time series model in forecasting the annual rainfall of selected stations of West Azarbaijan Province, *Applied Research Journal of Geographical Sciences*, 17(44), 105-87. <https://sid.ir/paper/364559/fa>. [in persian]
- Tektaş, M. (2010). Weather Forecasting Using ANFIS and ARIMA MODELS: Case study for İstanbul. *Environmental Research Engineering and Management*, 1(51), 5-10. DOI:10.5755/j01.erem.51.1.58.
- Younesi, Hojatul-lah; Torabi Poteh, Hassan; Arshiya, Azadeh and Mirzapour, Hafez (2017). Simulation of average monthly flow of Badavar-Noorabad River using time series models, 6th Scientific Research Conference on Soil Resources Management, Kerman. <https://civilica.com/doc/869538>. [in persian]