Proc. 19th International Conference on Statistical Sciences
Faisalabad, Pakistan – March 17-19, 2022, Vol. 36, pp. 85-92

MODELING AND FORECASTING OF MONTHLY AVERAGE MAXIMUM SURFACE AIR TEMPERATE IN MULTAN USING SEASONAL ARIMA

Muhammad Abdul Hafeez [1], **Omer Bashier**[2], **Ghulam Mustafa**[2],
Khadim Hussian[2], **Muhammad Shazad**[2] and **Abdul Waheed**[2]
[1] Pakistan Meteorological Department, Multan, Pakistan
[2] Crop Reporting Service, Agriculture Department, Vehari, Pakistan

ABSTRACT

The prime objective of this study is to forecast the Monthly Average Maximum Surface Air Temperature in Multan for 5 years from January-2022 to December-2026. This study will help the policy makers for future planning especially in the field of agriculture. For the purpose of analysis of this study R-software is used. For this study the data is taken from Pakistan Meteorological Department for the period of 10 years from January-2012 to December 2021. The data from January-2012 to December-2019 is used as a training data set and data from January-2020 to December-2021 is used as a test data to check the performance of the fitted model. In this study the Box-Jenkins methodology is used and best Seasonal Autoregressive Integrated Moving Average model is developed. On the basis of AIC and BIC, we have found two competing SARIMA models. SARIMA$(1,0,0)(1,1,1)_{12}$ and SARIMA$(1,0,1)(1,1,1)_{12}$ are selected suitable models for Monthly Average Maximum Surface Air Temperature and then on the basis of forecasting accuracy measures i.e. RMSE, MAE and MAPE SARIMA(1,0,0)(1,1,1)12 is selected as a best forecasting model. The forecasted values are fellow the pattern of the past values.

1. INTRODUCTION

Time series analysis and forecasting is one of the major tool used by scientists in meteorology and environmental field to study meteorological phenomenon like Rainfall, temperature and humidity. In this study a most popular time series technique “The Box-Jenkins methodology” is used in order to build an appropriate model. Because temperature is seasonal phenomenon with twelve months period therefore a seasonal autoregressive integrated moving average (SARIMA) model is used.

Significant changes in climate are taking place worldwide, the major cause in climate change is rise in temperature .Temperature is the most significant climatic factors; therefore it is important to understand its nature. Because in the fields of agriculture and business, temperature has a direct impact. So, this study will help the policy makers for better future planning in the field of Agriculture and Business. Excess in temperatures may also influence both human and livestock. As a result, it's essential to know how rapidly the temperature will rise.

Climate change is now one of the most serious environmental threats facing the entire world. It is one of the most serious threats to water resources, livelihoods, and forests. The temperature is a common meteorological variable that indicates how hot or cold it is. It influences not only plant and animal growth and reproduction, but also nearly all other meteorological variables such as evaporation rate, relative humidity, wind speed and direction

Several studies on meteorological phenomenon like rainfall, Minimum temperature, Maximum temperatures , Relative humidity ,Dew point etc. have already be done and in most of the study Time Series Analysis and Forecasting tools have been used in order to built best forecasting models. Nury, Hasan and Alam (2013): The primary goal of this study was to develop ARIMA models for short-term prediction of minimum and maximum temperatures for the years 2010 and 2011 for two temperature stations in the Sylhet division using the Box-Jenkins Methodology. Two meteorological variables Minimum and Maximum temperatures were use in this study. ARIMA (1,1,1)(1,1,1)12, ARIMA (1,1,1)(0,1,1)12 and ARIMA(1,1,1)(1,1,1)12, ARIMA (0,1,1)(1,1,1)12 were develop for Sylhet division and for Moulvibazar respectively. Burney, Barakzai and James (2017): aims of this research was to identify a suitable model for the forecasting of maximum temperature over a period of 12 months for Karachi city by using Box-Jenkins Methodology. One Meteorological variable "Maximum Temperature" was use in this study. On the basis of Likelihood and AIC model SARIMA (0,0,2) (2,1,1)12 finalized to forecast the future maximum monthly temperature of Karachi city of Pakistan. Goswami,

DATA AND METHODOLOGY

In this study two time series variables, Minimum temperature and Maximum temperature are taken. The minimum heating is the lowest temperature recorded within a 24-hour. And the maximum temperature is the highest temperature recorded within a 24-hour. Data is collected for 10 years from January-2012 to December 2021 from the Pakistan meteorological department for the city of Multan.

Auto Regressive Integrated Moving Average (ARIMA)

The acronym ARIMA stands for auto regressive integrated moving average. It is basically generalization of ARMA class of Models and used in the situation where the observed time series is non-stationary. To make the observed series stationary we apply ARIMA class of models where term "I" called integrated term and represents the order of difference that is used to make the series stationary. Mathematically form of ARIMA = ARIMA (p, d, q) Where, P = is the order of a.r terms d = is the order of differencing to make the series stationary q = is the order of m.a terms.

Use of SARIMA Model

Because Maximum temperature is a seasonal phenomenon with twelve months period that's why The (SARIMA) will be accustomed for this study. SARIMA is created by increasing seasonal variables to ARIMA models. Mathematical form of SARIMA (p, d, q) (P, D, Q)m. (p, d, q) and (P, D, Q)m, are the non-seasonal and seasonal parts the model respectively represent the number of season. Because the data utilized in this study is monthly data with a period of 12 month, we will set the value of m to 12.

Box Jenkins Methodology

Box Jenkins methodology is model building procedure mostly used in ARMA class of models. Prime objective of this approach is "To built and select the best forecast model among list of candidate's models". Box-Jenkins Methodology is a 4-step procedure. (1) Identification (2) Estimation (3) Diagnostic Checking (4) Forecasting Box-Jenkins Methodology is applicable only for stationary time series. Therefore very first step to test the Stationarity of the observed time series.

RESULTS AND DISCUSSION

Detection of Stationarity

To determine whether the underlying historic data of the maximum surface air temperature is stationary or not graphical and statistical testing approach is used as discussed below.

Graphical Analysis

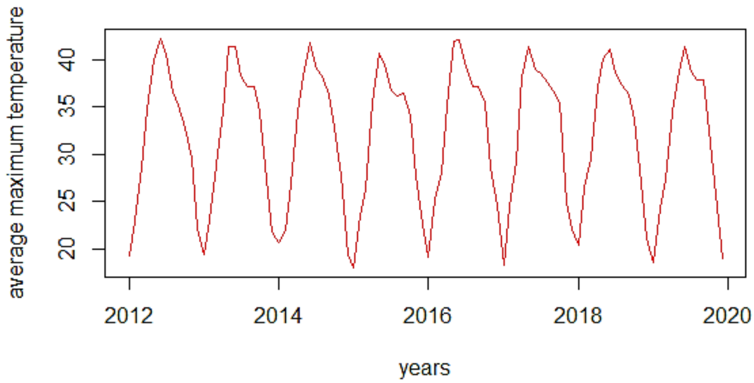


Figure 1: **Time Series Plot of Observed Mean Monthly Maximum Temperature Series**

Monthly average Maximum surface air temperature from January-2012 to December-2019 is plotted in Figure 1. The graphical analysis of monthly average Maximum surface air temperature represents that there is no trend in the given data. However a seasonal component can be observed.

Testing Approach to Assess Stationarity

Table 1
ADF and KPSS tests results of the Observed Monthly Average Maximum Surface Air Temperature

Test Type	Test Statistic	P-value
ADF	-7.9395	0.01
KPSS	0.017545	0.1

Because both tests and graphical analysis shows that the underlying time series data is stationary. Hence it is concluded that observed monthly average maximum surface air temperature is stationary.

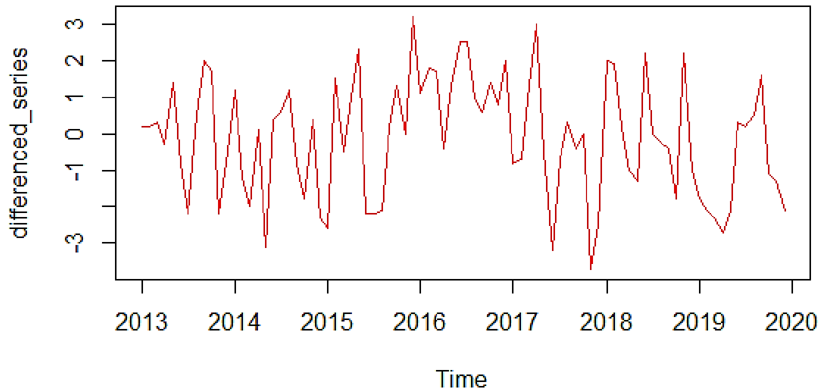


Figure 2: Time Series Plot after First Seasonal Difference of Monthly Average Maximum Surface Air Temperature Series

Seasonal Difference

After ensure the Stationarity and taking the seasonal difference of the original time series now we will apply the formal four steps of Box-Jenkins Methodology.

Identification

In the step of Identification first of all a model is specified among ARMA class of models and then tentative orders of AR and MA terms are identified on the basis of PACF and ACF respectively and finally on the basis of AIC and BIC best model is selected.

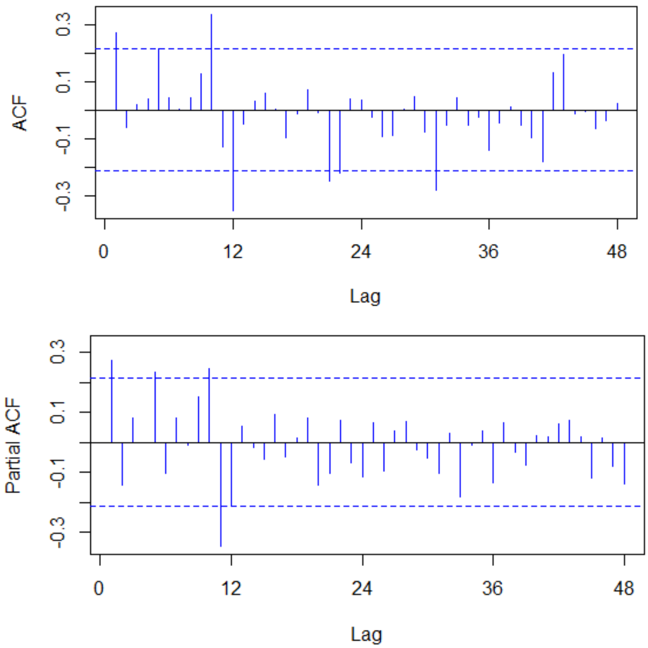


Figure 3: ACF and PACF plot after First Seasonal difference of Monthly Average Maximum Surface Air Temperature Series

From ACF and PACF it is concluded that that SARIMA $(3,0,3)(1,1,1)_{12}$ can be used as a base model.

Selection of Best Model

On the basis of base model there are different SARIMA models are considered with different order as shown in table given blow.

Table 2
Selection of Best Model

Model	AIC	BIC
$(3,0,3)(1,1,1)_{12}$	302.23	324.11
$(1,0,0)(1,1,1)_{12}$	**300.7**	**310.42**
$(1,0,1)(1,1,1)_{12}$	**302.56**	**314.71**
$(2,0,2)(1,1,3)_{12}$	305.83	325.27
$(2,0,0)(0,1,1)_{12}$	302.52	314.67

Above table SARIMA(**(1,0,0)(1,1,1)** and SARIMA **(1,0,1)(1,1,1)$_{12}$** have the lowest AIC and BIC values among list of other candidates models. Hence these two models are considered as competing models. In the final step the forecasting performance of these two competing models will be compared on the basis of some forecasting accuracy measures.

Estimation of SARIMA (1,0,0)(1,1,1)$_{12}$

The parameters of SARIMA(1,0,0)(1,1,1)$_{12}$ are estimated using R-Software and the estimated parameters of SARIMA(1,0,0)(1,1,1)$_{12}$ are given in the following table.

Table 3
Parameters Estimation of SARIMA (1, 0, 0) (1, 1, 1)$_1$ Model

Parameters	AR1	SAR1	SMA1
Estimate	0.2587	0.2291	-0.9984
S.E	0.1077	0.1354	1.2033

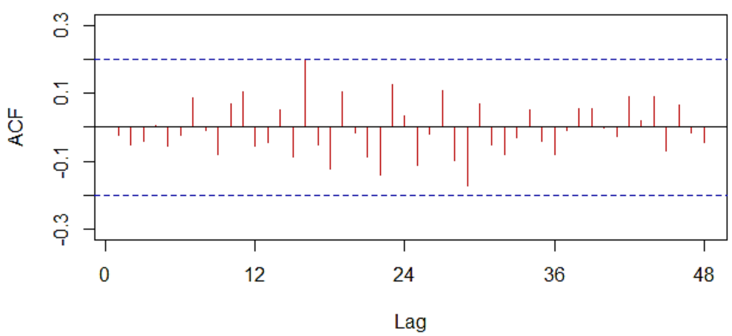
Diagnostic Checking of SARIMA (1,0,0)(1,1,1)$_{12}$
Normality Test for Residuals

Table 4
Shapiro-Wilk and Jarque-Bera tests Results of the Residuals

Test Type	Test Statistic	P-value
Shapiro-Wilk	0.97822	0.1
Jarque-Bera	2.5008	0.3

Because both the tests proved the normality. Hence it can be concluded that residuals of the model **(SARIMA (1, 0, 1)(1, 1, 1)$_{12}$)** follow the assumption of Normality.

Independence/White Noise Testing
ACF Plot of Residuals



In the ACF plot it is observed that all the spikes within confidence band which indicate that residuals are uncorrelated.

Evaluation of Forecasting Accuracy

In this section the accuracy of model SARIMA $(1,0,0)(1,1,1)_{12}$ and SARIMA $(1,0,1)(1,1,1)_{12}$ is compared with the help of graph and on the basis of different accuracy measures such as RMSE, MAE and MAPE.

(a)Graphical Performance of SARIMA $(1,0,0)(1,1,1)_{12}$

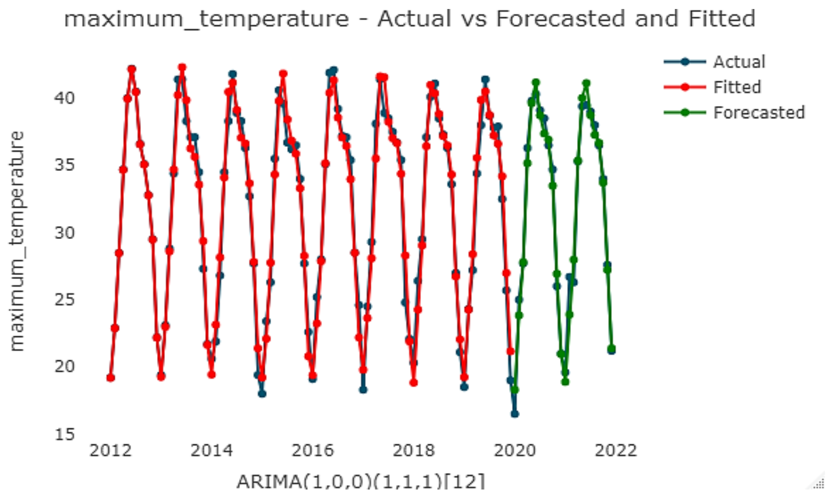


Table 5
In Sample and Out of Sample Forecast Performance

Model	In Sample Forecast Performance			Out of Sample Forecast Performance		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
$\mathbf{(1,0,0)(1,1,1)_{12}}$	1.1464	0.8612	2.9445	1.0174	0.8421	2.7709
$(1,0,1)(1,1,1)_{12}$	1.1475	0.8663	2.9613	1.0300	0.8531	2.8147

Because all the forecasting accuracy measures such as RMSE, MAE and MAPE of model SARIMA$\mathbf{(1,0,0)(1,1,1)_{12}}$ are less than model SARIMA $(1,0,1)(1,1,1)_{12}$ for both cases in sample and for out of sample. Hence it is concluded that on the basis of our study the best forecasting model for Monthly Average Maximum Surface Air Temperature for Multan is SARIMA$\mathbf{(1,0,0)(1,1,1)_{12}}$.

REFERENCES

1. Gerretsadikan, A. and Sharma, M.K. (2011). Modeling and forecasting of rainfall data of Mekele for Tigray region (Ethiopia). *Statistics and Applications*, 9(1-2), 31-53.
2. Nury, A.H., Hasan, K. and Alam, M.J.B. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University-Science*, 29(1), 47-61.
3. Abdul-Aziz, A.R., Anokye, M., Kwame, A., Munyakazi, L. and Nsowah-Nuamah, N.N.N. (2013). Modeling and forecasting rainfall pattern in Ghana as a seasonal ARIMA process: The case of Ashanti region. *International Journal of Humanistic and Social Science*, 3(3), 224-233.
4. John, C., George, U. and Chukwuemeka, O.S. (2014). Time series analysis and forecasting of monthly maximum temperatures in south eastern Nigeria. International Journal of Innovative Research and Development, 3(1), 165-171.
5. Burney, S.A., Barakzai, M.A.K. and James, S.E. (2020). Forecasting Monthly maximum temperature of Karachi city using time series analysis. *Pakistan Journal of Engineering, Technology & Science*, 7(2).
6. Goswami, K., Hazarika, J. and Patowary, A.N. (2017). Monthly temperature prediction based on ARIMA model: A case study in Dibrugarh station of Assam, India. *International Journal of Advanced Research in Computer Science*, 8(8), 292-298.
7. Babazadeh, H. and Shamsnia, S.A. (2014). Modeling climate variables using time series analysis in arid and semi arid regions. *African Journal of Agricultural Research*, 9(26), 2018-2027.
8. Syeda, J.A. (2012). Trend and variability analysis, and forecasting of maximum temperature in Bangladesh. *Journal of Environmental Science and Natural Resources*, 5(2), 119-128.
9. Kabbilawsh, P., Sathish Kumar, D. and Chithra, N.R. (2020). Trend analysis and SARIMA forecasting of mean maximum and mean minimum monthly temperature for the state of Kerala, India. *Acta Geophysica*, 68(4), 1161-1174.
10. Mahsin, M.D., Akhter, Y. and Begum, M. (2011). Modeling rainfall in Dhaka division of Bangladesh using time series analysis. *Journal of Mathematical Modelling and Application*, 1(5), 67-73.