Advances in Aeronautical Science and Engineering
ISSN: 1674-8190

Modeling Daily Temperature Extremes with Autoregressive Time Series Techniques

Simone Keller*[1] & Dr. Maria Rinaldi[2]
***[1]Research Scholar, University of Milan, Milan, Italy**
***[2]Assistant Professor, Department of Applied Sciences, University of Milan, Milan, Italy**

ABSTRACT

Time series analysis and forecasting has become a major tool for analyzing, forecasting and in other applications in various meteorological and hydro logical data like temperature, rainfall, humidity, river flow, stream flow and many more others. In this paper our goal is to forecast the future trend of maximum and minimum temperature of Punjab, using non-seasonal and seasonal Autoregressive moving average and Seasonal autoregressive integrated moving average models. For this we collect and analyze the past values and developed appropriate models which describe the inherent structure and characteristics of the series. We calculated the accuracy of our results on the basis of error measures and residual normality test . The forecast results obtained by us provide a detailed explanation of model selection and forecasting accuracy is presented. Autoregressive time series models are widely used in forecasting and analyzing of various data that are also applicable in different research areas as our forecasting results can also be used by various agencies for planning and development of various meteorological and hydro logical decisions.

***Keywords**: ARMA, SARMA, SARIMA, MAE, RMSE.*

I. INTRODUCTION

Autoregressive model is a type of random process which is used to describe certain time varying process. There are various types of autoregressive time series models used as a statistical tool for analyzing various type of time series data. Yule and Walker introduced Autoregressive Moving Average (ARMA) model then later on, Box and Jenkins proposed the methodology of Autoregressive Integrated Moving Average (ARIMA) model [8]. For the seasonal time series forecasting, proposed a quite successful variation of the ARMA and ARIMA models are called Seasonal ARMA (SARMA) and seasonal ARIMA (SARIMA) models . Brockwell P. J. and Davis R. A [2] and Agarwal R.K et. al. [1] gives a detail knowledge about time series and forecasting using autoregressive models in their book. Chujai P. et. al. [5] have been carried out their work on ARIMA and ARMA model for analysis of household electric consumption.

In recent two decades, the resultant changes in global climate were one of the main issues among weather changes in some countries. Temperature is a common meteorological variable which indicates a fair idea of climate. It does not only affect the growth and reproduction of plants and animals, but also has a influence on other meteorological variables such as relative humidity, rate of evaporation, wind speed, wind direction and precipitation patterns. Many researchers has been applied autoregressive models for analyzing meteorological data on various purposes in past as well as in present. Taylor J.W. and Buizza R. [19] has built model about daily UK temperature and forecasted it. Romilly P. [17] had carried out research related to modelling of global mean temperature and many more [15], [18]. There are also several studies related to temperature analysis and forecast of various region [7], [9], [12], [20]. Some researchers applied SARIMA model for forecasting temperature data of Ghana [14], Nanjing[4], Mt. Kenya region [13]. In recent years, research paper of Papacharalampous G. et. al. [16] and Iqelan B. M.[11] had done their work on monthly temperature data. There are verius types of error measures on accuracies are discussed in the paper of Hyndman R. J. and Koehler A. B. [10]. Chai and Draxler R. R. [3] were discussed about MAE and RMSE error measures which are standard statistical metric to measure model performance .

This study aims to provide forecasting information of temperature data. The non seasonal and seasonal autoregressive moving average and Seasonal Autoregressive Integrated Moving Average models are applied to analyzed the maximum and minimum temprature of Punjab, using R software. Cowpertwait P. S. P. et al. [6] discussed in their book about time series analysis with R. We calculated the accuracy of our models on the basis of minimum error measures and residual normality test. The forecast results obtained by us provide a deatailed explanation of model selection and forecasting accuracy is also presented.

II. DATA SOURCE

The source of temperature data in our study is Indian Metrological Department (IMD), Chandigarh. It provides metrological data for various research and development activities like Monsoon, Climate Change and Agricultural Meteorology. The geography and subtropical latitudinal location of Punjab lead to large variations in temperature. Our study focus on maximum and minimum temprature of Punjab from the period February, 2017 to March, 2018. We analyzed the daily mean of maximum and minimum temperature data of three climate zone of Punjab namely Amritsar, Ludhiana and Patiala.

III. METHODOLOGY

The difference of the series $\{X_t\}$ at lag *s* is a convenient way of eliminating a seasonal component of period '*s*' when we fit an ARMA (p,q) model i.e $\phi(B)X_t = \theta(B)Z_t$. The advantage of the ARMA (p,q) specification in given equation is that it can offer a more parsimonious model than a purely AR(p) process, when the order of p is high in the latter process means that the lag order p+q in the ARMA(p, q) specification will generally be less than the lag order p in the AR(p) specification. To the differenced series $Y_t = (1 - B_s)X_t$, then the model for the original series becomes

$$\phi(B)(1 - B_s)X_t = \theta(B)Z_t$$

The above equation represent the ARIMA (p, d, q) model and given series is stationary if d=0, then it reduce in to ARMA (p, q) process. Since for d ≥ 1, above equations determines the integrated model. When d and D are non-negative integers, then $\{X_t\}$ is a seasonal ARIMA (p, d, f) ×(P, D,Q)$_s$ process with period *s* if the differenced series $Y_t = (1 - B)^d (1 - B_s)^D X_t$ is ARMA. In special case of the general seasonal ARIMA (SARIMA) model defined as following ,

$$\phi(B)\Phi(B_s)Y_t = \theta(B)\Theta(B^s)Z_t,\ (Z_t) \sim WN(0, \sigma^2)$$

Where $\phi(z) = 1 - \phi_{1Z} - ... - \phi_{pZ^p}$, $\Phi(z) = 1 - \Phi_{1Z} - ... - \Phi_{pZ^p}$, $\theta(z) = 1 + \theta_{1Z} + ... + \theta_{pZ^p}$ and $\Theta(z) = 1 + \Theta_{1Z} + ... + \Theta_{QZ^Q}$.

After analysis of our data by using non-seasonal and seasonal autoregressive model we forecast our data. Error measures are used for a clear view of our accuracy in forecasting. For this purpose we proceed through MAE, RMSE and normality test.

$$MAE = \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{n}$$

$$RMSE = \sum_{i=1}^{n} \sqrt{\frac{(\hat{y}_i - y_i)^2}{n}}$$

Where $\hat{y}$ is a vector of n forecast values, y is the vector of the

We also apply the normality test for residual checking.

IV. GRAPHIC ANALYSIS

This section contains analysis of raw data on daily maximum and minimum temperature of Punjab based on Autoregressive time series models. The plot of our daily data is given in fig. 1. Graph can be analysed in to three parts, first part shows the trend of graph values from 0 to 100 which indicates the inceasing trend of temperature with respect to time because of upcoming summer seasons. Second part is the centre area of graph which shows the values from 100-200 observations of maximum tempratue value and minimum temperature values are also high in this area due to summer seasons and after that graph shows fluctuations. In third part trend has decreased because of season variations.

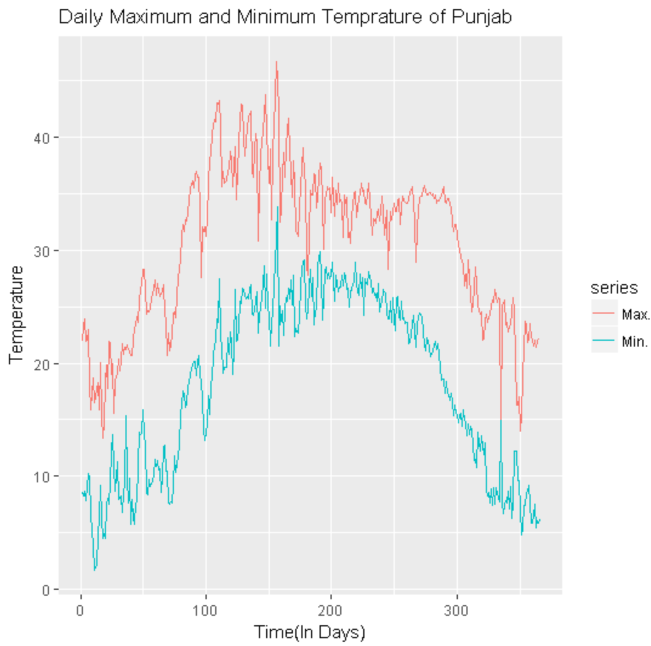

Figure 1: Graph of Daily Temperature Data of Punjab

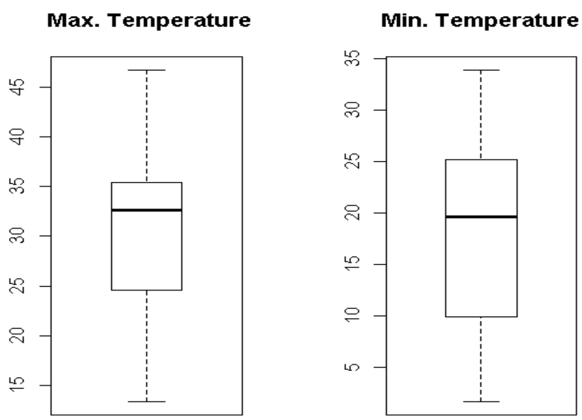

Figure 2: Box plot of temperature data

In fig 2 we examine the outliers in our data, the first plot is related to maximum temperature data and second plot is related to minimum temperature data. After observing the plots we find out that there exist no outliers in our data.

V. ACF AND PACF PLOTS

Autocorelation function (ACF) and Partial Autocorrelation function (PACF) are primary tool for identify the relations that occur between time series at various lags.

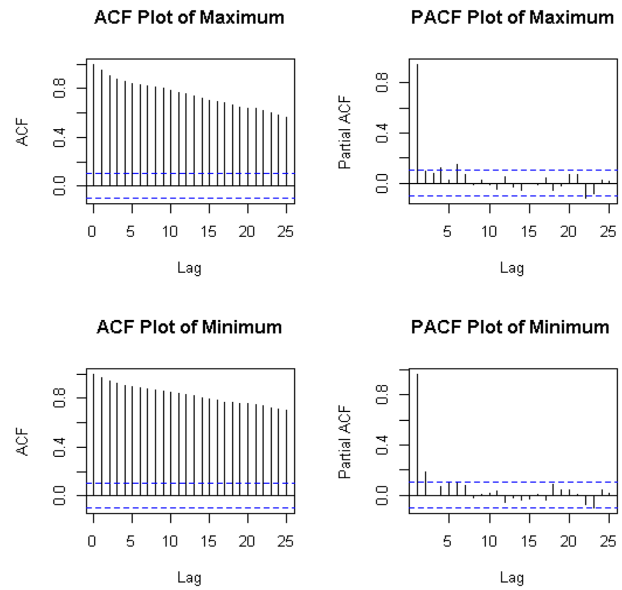

Figure 3: ACF and PACF plot of maximum and minimum temperature

In fig. 3 plot of ACF and PACF show the correlatin between observations of time series. The series has shown the variation with time over a constant mean and variance, series is not stationary and also occur seasonality.

VI. STATIONARITY TEST

A time series is stationary when the mean and variance are constant over time. We use the ADF test and the KPSS test for testing the stationarity. In KPSS test, the null hypothesis is that the data is stationary and also specified a trend stationarity for maximum and minimum temperature data means that the data follow a straight line time trend with stationary errors. The p-value is 0.1, so the null hypothesis is not rejected at the usual 1% level and ADF test also shown that there is no unit root in our time series data.

VII. MODELING OF TIME SERIES DATA

In this section we discuss the fitted autoregressive models used for analyzed our temperature data. There are three autoregressive models used for analysis of data such as ARMA, SARMA, SARIMA.The model selection is based on minimum value of information criterion which is discussed and shown in following table for maximum and minimum temperature for different given models.

ARMA model

In ARMA model we are taking bivariate time series of temperature data which contain daily maximum and minimum temperature. We apply ARMA model on our original series for maximum and minimum temperature in Table1 and Table2 as given below which shows the parameter estimation of both the series with their corresponding standard error :

Table 1: Parameter estimation of the ARMA(1,1) model for maximum temperature

	ar1	**ma1**
Coefficients	0.5712	-0.8361
S. E.	0.0788	0.0517

Table 2: Parameter estimation of the ARMA(3,2) model for minimum temperature

	ar1	ar2	ar3	ma1	ma2
Coeffi.	**1.0210**	**-0.2194**	**-0.2172**	**-1.1805**	**0.3842**
S. E.	**0.1985**	**0.2332**	**0.0619**	**0.1945**	**0.2618**

SARMA model

Seasonal ARMA model is used for decomposing our time series then we obtain the parameter estimation for maximum and minimum temperature which is given in Tables3 and 4.

Table 3: Parameter estimation of the SARMA(5,0,1)(2,0,0) model for maximum temperatue

	ar1	ar2	ar3	ar4	ar5	ma1	sar1	sar2
Coeff.	0.46	-0.01	-0.03	0.04	-0.15	-0.73	0.07	-0.02
S. E.	0.13	0.07	0.06	0.06	0.05	0.13	0.06	0.05

Table 4: Parameter estimation of the SARMA(3,0,3)(2,0,2)model for minimum temperature

	ar1	ar2	ar3	ma1	ma2	sar1	sar2
Coeff.	-0.68	0.24	0.69	0.5	-0.17	0.28	0.03
S. E.	0.22	0.12	0.10	0.23	0.16	0.13	0.06
	sma1	sma2					
Coeff.	-0.25	0.83					
S.E.	0.22	0.23					

SARIMA model

Seasonal ARIMA model has apllied on our transformed data and then obtain the parameter estimation of our model for maximum and minimum which is shown for in Table5 and 6 below:

Table 5: Parameter estimation of the SARIMA(5,1,1)(2,0,0) model for maximum temperature

	ar1	ar2	ar3	ar4	ar5	ma1	sar1	sar2
Coeff.	0.46	-0.01	-0.03	0.03	-0.15	-0.73	0.07	-0.02
S. E.	0.13	0.07	0.06	0.06	0.05	0.13	0.06	0.05

Table 6: Parameter estimation of the SARIMA(1,1,0)(1,0,0) model for minimum temperature

	ar1	sar1
Coeff.	-0.5671	0.1047
S. E.	0.0432	0.0531

From Table6 we get the results of parameters estimation of the SARIMA model for minimum temperature time series.

VIII. MODEL COMPARISON

For the model comparision we used and calculated the two error measures which are MAE and RMSE of our applied models are given below in table7:

	Models→	ARMA	SARMA	SARIMA
For Max. Temperature	MAE	0.052	0.051	**0.050**
	RMSE	0.08	0.08	**0.07**
For Min. Temperature	MAE	0.15	**0.14**	0.19
	RMSE	0.09	**0.09**	0.12

From Table 7 it can be analyzed that SARIMA (5,1,1)(2,0,2) is the best model for maximum temperature and SARMA(3,0,3)(2,0,2) is the best model for minimum temperature on the basis of their minimum error measures.

IX. RESIDUAL CHECKING

In time series modeling, the selection of a best model fit to the data is based on residual checking. There are various statistical test to check the residual performance of fitted model. In this paper we are using the ACF and PACF residual plots and normality test for residual checking of selected model.

Use ACF and PACF residual plot

The selection of best model fit to the data is depends on its performance of residual analysis. Residual must have zero mean, constant variance and also uncorrelated. Fist plot of figure 4 shows the standarized residuals for maximum temperature which has zero mean and constant variance since the residuals are constrained around -0.2 to 0.2. Again ACF and PACF of residuals shows that the autocorrelation of the residuals are all zero i.e. uncorrelated for maximum temperature. Similarly figure 5 shows the standarized residuals of minmum temperature constrained around -5 to 5 with zero mean and constant variance. Also the ACF and PACF residual of the model for minmum temperature shows that the autocorrelation of the residuals are zero i.e. uncorrelated for minimum temperature.

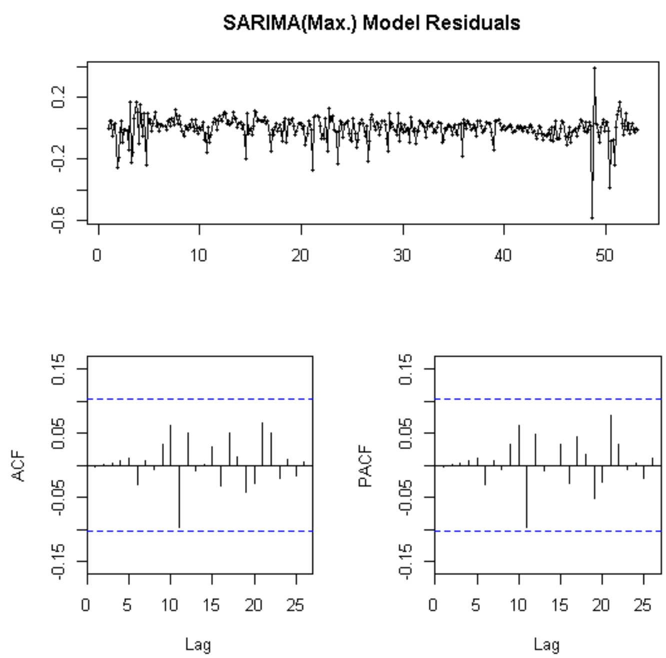

Figure4: Residul plot for Maximum Temperature

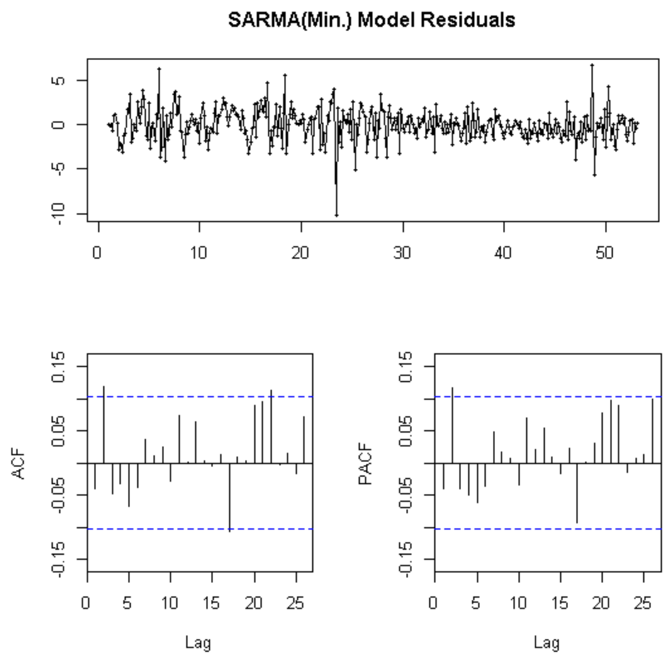

Figure5: Residul plot for Minimum Temperature

Normality Test for Residuals

For the normality test we use the Q-Q plot and histogram plot for the residual checked. From the Q-Q plot it is observed that the most of the points passes through the straight line which indicates the residual in the model are normal and form histogram plots of residual the distribution of the residuals can be clearly observed normal with a bell shape distribution.

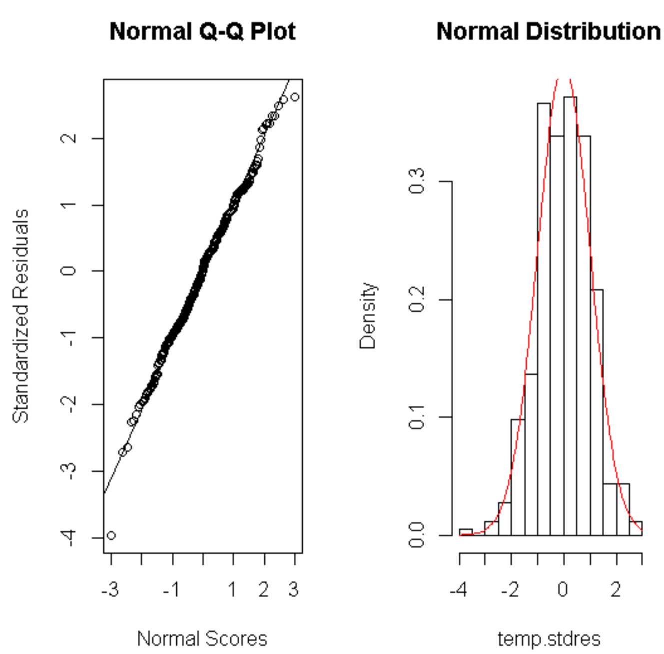

Figure 6. Graph of normality test for the residuals

From the above discussion therefore our selected models are good for the forecasting.

X. FORECASTING PERFORMANCE

In this section, we complete our ultimate goal of seasonal modeling for forecasting the temperature data. In figure 7 shown the graph of predicted values of maximum and minimum temperature for year 2019 to 2021. It also observed that the seasonal models are usually applicable for short term forecasting as there are very less variations observe for maximum and minimum temperature of forecasting plot. The diagnostic checking confirms the accuracy of the models. The chosen models predict the estimated values of maximum and minimum temperature data.

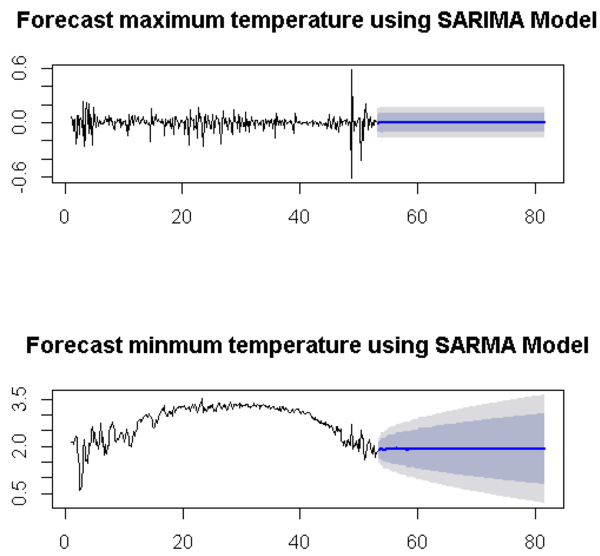

Figure 7: Forecasted plot of maximum and minimum temperature

XI. CONCULSION AND FUTURE SCOPE

In this study, the stationarity of the series is verified through the plot the sample of ACF and PACF and tested by using KPSS and ADF tests. Since the seasonal components are observed from the decomposition of the series, so the seasonality is removed. Among three models used in this paper , the best selected model for maximum temperature is obtained as SARIMA(5,1,1)(2,0,2) and for the minimum temperature and SARMA(3,0,3)(2,0,2) is the best model fitted model on the basis of minimum error measures. For the model diagnostics were also done through the performance of model residuals and Q-Q plot and histogram proves it to follow normality . Lastly it can be concluded that identified model was found to be very adequate and best for predicting future values of maximum and minimum temperature. Our results can be help for decision makers to establish better strategies against upcoming weather changes.

REFERENCES

1. *Agrawal R. K. & Adhikari R., "An Introductory study on time series modeling and forecasting",Lambert Academic Publishing, doi:10.13140/2.1.2771.8084, 2013*
2. *Brockwell P. J. and Davis R. A. , "Introduction to time series and forecasting", Springer-Verlag New Yark, 2002*
3. *Chai T. and Draxler, "Root mean square error or Mean absolute error? Arguments against avoiding RMSE in the literature", Geoscientfic Model Development, Vol No. 7, pp. 1247-1250, 2014*
4. *Chen P., Niu A., Liu D., Jiang W. and Ma B., "Time series forecasting of temperatures using SARIMA: An example from Nanjing", IOP Conf. Series: Materials science and engineering, Vol No. 394, 2018*
5. *Chujai P., Kerdprasop N. and Kerdprasop K., "Time series analysis of household electric consumption with ARIMA and ARMA moidels",Procedding of the international MultiConference of engineers and computer scientists, Vol I, pp. 13-15, 2013*

6. *Cowpertwait P. S. P. and Metcalfe A. V., "Introductory Time Series with R",Springer, 2009*
7. *Dabral P. P. , Saring T. and Jhajharia D, "Time series models of monthely rainfall and temperature to detect climate for Jorhat(Assam), India", Global NEST Journal, Vol No. 18, Issue- 3, pp. 494-507,2016*
8. *G. E. P. Box and G. M. Jenkins, "Time series analysis: forecasting and control", Revised Edition , Holden – Day, san Francisco, CA, 1976*
9. *Gillian L. H., Rao S. S. and Rao T. S., "Statistical analysis and time series models for minimum/ maximum temperatures in the Antarctic Peninsula", Proceeding of the Royal Society, doi: 10.1098?rspa.2006.1766*
10. *Hyndman R. J. and Koehler A. B., "Another look at measures of forecast accuracy",International Journal of Forecasting, Vol No. 22, pp. 679-688, 2006*
11. *Iqelan B. M., "Time series modeling of monthly temperarute data of Jerusalem/Palestine", Mathematika, Vol No. 31, Issue- 2, pp. 159-176, 2015*
12. *Jain S. K. , Kumar V. and Saharia, "Analysis of rainfall and temperature trends in norhtest India", International Journal of Climatology , Vol No. 33, pp. 968-978, 2013*
13. *Kibunja H.W., Kihoro J. M., Orwa G. O. and Yodah W. O., "Forecasting precipitation using SARIMA Model: A case study of Mt. Kenya Region", Mathematical Theory and Modeling, Vol No.-4, Issue- 11, pp. 50-58, 2014*
14. *Michael Asamoah- Boaheng, "Using SARIMA to forecast monthely mean surface air temperature in the Ashanti region of Ghana", International Journal of Statistics and Applications, Vol No. 4, Issue-6, pp. 292-299, 2014*
15. *Murat M., Malinowska I., Gos M. and Krzyszczak J.,"Forecasting daily meteorological time series using ARIMA and regression models", International Agrophysics,Vol No. 32, pp.253-264, 2018*
16. *Papacharalampous G., Tyrails H. and Koutsoyiannis, "Predictabilty of monthly temperature and precipitation using automatic time series forecasting methods", acta Geophysics, 2018*
17. *Romilly Peter, "Time series modelling of global mean temperature for managerial decision-making", Journal of Environment Management, Vol No. 76, pp. 61-70, 2005*
18. *Sarraf A. , Vahdat S. F. and Behbahaninia A., "Relative humidity and mean monthly temperature forecasts in Ahwaz Station with ARIMA Model in time series analysis", International Conference on Environment and Industrial Innovation IPCBEE, Vol No. 12,IACSIT Press, Singapore,2011*
19. *Taylor J. W. and Buizzo R., "A comparision of temperature density forecasts from GRACH and Atmosphric Models" , ECMWF Technical Memoranda, 2003*
20. *Ustaoglu B. , Cigizoglu H. K. and Karaca M.,"Forecast of daily mean, maximum and minmum temperature time series by three artificial neural network methods", Royal Meteorological Society , Meteorological Publicaions, Vol No. 15, pp. 431-445, July 2008.*

.