International Journal of Scientific and Innovative Mathematical Research(IJSIMR)
Volume 8, Issue 4, 2020, PP 28-46
ISSN No. (Online) 2347-3142
DOI: https://doi.org/10.20431/2347-3142.0804003
www.arcjournals.org

Monthly Temperature Prediction Based on Seasonal ARIMA Model: A Case Study of Dar es Salaam, Tanzania

Paul Andrew Panga[1*], Shaban Nyimvua[2], Isambi Mbalawata[3]

[1]*National Institute of Transport (NIT)*

[2]*Department of Mathematics University of Dar es salaam*

[3]*African Institute of Mathematical Sciences (AIMS)*

Corresponding Author:** ***Paul Andrew Panga, *National Institute of Transport (NIT), Tanzania*

Abstract: *Temperature is among the main climatic elements which is directly connected with climate change and recently the issues has become an important topic in many parts of the world. This is because the issue leads to variations in water, wind, rain, to mention few, that may results to meteorological disasters like floods and droughts. In this paper Seasonal Autoregressive Integrated Moving Average (SARIMA) model have been setup and carry out prediction of monthly maximum and minimum temperatures in Dar es Salaam region of Tanzania. The Box and Jenkins methodology was used to set up the Seasonal ARIMA models by using the temperature data recorded from the period of January 1985 to December 2015. Temperatures observations were found to have seasonality and non-stationarity and hence differencing and seasonal differencing was used to attain stationarity. Based on the ACF and PACF plots the optimum orders of the Seasonal ARIMA models were determined and evaluated by using the useful information criterion (AIC, AICc and BIC). The analysis reveals that the appropriate models which are useful in describing the temperature observations are SARIMA* $(1, 1, 2)\times(1, 1, 1)_{12}$ *for monthly maximum temperature and SARIMA* $(2, 0, 2)\times(1, 1, 1)_{12}$ *for monthly minimum temperature. After model evaluation and validation, the forecasting was made for the upcoming ten (10) years, from January 2016 to December 2025. In view of the forecasting, there is an increase in maximum and minimum temperature for the upcoming ten years. The increase in temperature suggests that climate change could continue to have negative impacts on different economic sectors including tourism, water resource, to mention few, in Dar es Salaam community and this call for increased adaptive capacity to the community. Moreover, higher temperatures have effects on droughts, changing rainfall patterns and availability of surface water whose consequences range from less food supply to general fewer water supplies in Tanzania particularly Dar es Salaam region .*

Keywords: *Maximum Temperature, Minimum Temperature, Seasonal ARIMA (SARIMA) models, Forecasting, time series, MAPE, RMSE*

1. INTRODUCTION

Change in climate has been observed as a crucial issue over the past decade and temperature was categorized as the leading component for climate change (Roy, 2012). The Intergovernmental Panel on Climate Change (IPCC) report (IPCC, 2012)(Stocker et al, 2013), approved that the overall degree of hotness is increasing while cold days and cold nights are expected to decrease. As the world climate changes, massive effects are seen on the farming based societies as a result of extreme temperature and drought incidences (Stocker et al, 2013). Arguably, the impacts of climate change are still increasing (Shahin et al, 2016). It is evident that countries which are located in the tropical regions are more likely to be affected by change in climate than other regions. Also, developing countries are more affected by climatic change than developed countries because of lack of resources to adapt to changes, relying on rainfed agriculture and natural resource to control their livelihoods (UNFCCC, 2006).

Vital sectors of the Tanzanian economy such as agriculture and fishing, to mention few, depend on climatic conditions. Considering this case, for the survival and growth of the plants, temperature is required, though too much or too little is still a problem (Shahin et al, 2016). On the other hand, increase in temperature may lead to drought, fall in crop production accompanied by food insecurity (Ojija et al, 2017). Thus, early indication may help to solve a number of problems associated with climatic changes.

Over the past decades Tanzania has witnessed increased prevalence of climatic events such as drought and floods, which are linked with grievous ecological and socio-economic intimation like loss of lives and destruction of structural design (Roy, 2012) (Kijazi A and Reason C, 2009). Serious floods that have recently tortured many parts of the country include those of 2006, 2009, 2010, 2011, 2012, 2014, 2016 and 2017 (Chang'a et al, 2017). And it is important to notice that among the top ten catastrophes which cause devastating impacts on the country's economy, epidemics is ranked first followed by floods (Mboera et al. 2012). And it has no doubt that the massive increase in rainfall is directly connected with the increase in green house emission like carbon dioxide which rise the degree of hotness in the atmosphere. Research has declared that the developing country like Tanzania is likely to encounter irregular drought which pose effect to agriculture, water, energy, and livestock. Moreover, drought is associated with backwardness in development and crippled socio-economic activities. Current droughts which harm the country include those of 2003, 2005, 2011, 2014 and 2016 (Chang'a et al, 2017). Hence the need for a suitable prediction method to be applied in forecasting climatic pattern is important.

Variation of climate has been a topic in many parts of the world due to its immediate effects on people's lives (Ghahraman, 2007). Dar es Salaam region located on the coastal areas of Tanzania is characterized by tropical type of climate with higher degree of hotness, high humidity and average annual precipitation over 1000 mm (UARK, 2017). The region is experiencing warmest time during January and February and coolest time during July and August. Thus due to change in climate the trends of climate variables like temperature for the region is dynamic and there is variability in rainfall caused by a number of different time scale from daily to decadal (UARK, 2017). It is evident that the trend and variability of climate will continue at a longer timescale (Roy, 2012). Dar es Salaam has been experiencing destructive rainfall in some rainy seasons which results to loss of lives and destruction of properties (Ngailo et al, 2016). Also studies has declared that the region has been experiencing back to back floods in recently years including 2010, 2011, 2012, and 2013 (Kebede et al, 2012). Therefore, there is a need of understanding the nature and scale of change in climate in Dar es Salaam region associated with finding better forecasting tool for temperature which will be crucial in taking precautions and formulation of policies for mitigation and adaptation measures.

Recently Time series analysis and forecasting was observed to be an important tool when applied in studying the variations and trends of different hydo-meteorological variables such as precipitation, humidity, temperature, streamflow and many other environmental parameters (Nury et al, 2013). Various published papers have analyzed temperatures by using Time series Box and Jenkins Seasonal ARIMA approaches, which gives the usefulness of modelling temperature from different parts of the world. Libya (El-Mallah et al. (2016)), Bangladesh (Nury et al. (2013), Sultana et al. (2015)), Sri lanka (Alibuhtto et al. (2019)), Iran (Machekposhti et al. (2018)), Iraq (Chawsheen et al. (2017)) and Khuzestan (Sarraf et al. (2011)). Most of the observations and time series modelling results of the mentioned studies have declared projected increase in temperature. However, there are limited or no published papers that have attempted to understand, analyze, model and predict temperature by using Box and Jenkins ARIMA approach in Tanzania particularly Dar es salaam. Therefore, this paper would seem to be the first application of the Box and Jenkins ARIMA approach for temperature in Dar es Salaam Tanzania.

Herein in this Paper, we will begin with finding the appropriate time series models for monthly maximum and minimum temperature by using previous available data from 1985 to 2015 of Dar es Salaam region, Tanzania. Second, we will predict the future trends of maximum and minimum temperature values by using the time series model developed. Box and Jenkins methodology will be used in developing the time series model. The approach flows through identification of the model, estimation of the model parameters, diagnostic checking and use the model for forecasting purposes (Box et al, 1976).

Different researchers allude that socio economic development of the developing countries like Tanzania are hindered by the trends and patterns of climatic extremes (Chang'a et al, 2017). Efforts like achieving Millennium Development Goals (MDG), Sustainable Development Goals (SDG) and National Developmental Vision (Visions 2025) which are associated with reducing poverty, hunger and promoting food security are hampered by floods and natural disasters like drought, hence if not managed properly the prolonged impacts will continue in the future. Hence this study is importance for providing information to decision makers, planners, climatologist, meteorologist and others on predicting the future rainfall.

The rest of the paper are organized as follows. Section 2, describes methodology for fitting time series models. Finally, the Study area and type of data used together with the results of the appropriate time series model and their prediction are discussed in Section 3 and then we will give Conclusion.

2. METHODOLOGY

2.1. Stationary and Non stationary Series

A time series is said to be stationary if the joint distribution

$X_{t_1}, X_{t_2}, \dots, X_{t_n}$ and $X_{t_{1+T}}, X_{t_{2+T}}, \dots, X_{t_{n+T}}$ are the same. Note that shifting the position of period T does not intervene the distribution simply because it depends $t_1, t_2, \dots\ t_n$. Time series model which is not stationary is given by:

$$Y_t = \mu_t + \varepsilon_t,$$

where, μ_t is the mean function of time and ε_t is the weakly stationary process.

2.2. Unit Root Test

The test was derived by Dickey and Fuller (1979) to test the presence of non-stationary (unit root) versus stationary process. The model for unit root and stationary process is given by system of equations below:

$$W_t = \phi_1 W_{t-1} + \varepsilon_t$$

$$W_t = \phi_0 + \phi_1 W_{t-1} + \varepsilon_t$$

when $\phi = 1$, then the system is said to have the unit root (non-stationary).

In this Paper we check the existence of unit root by using the following tests.

Augmented Dickey Fuller Test (ADF)

The test statistics is given by:

$H_0 : \phi_1 =$ The series has a unit root

$H_1 : \phi_1 =$ The series has no unit root

If the test statistics of ADF test is less than the critical value then we reject the null hypothesis that the time series data has the unit root.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test

The Hypothesis test is given by:

$H_0 : \phi_1 =$ The series is level or trends stationary

$H_1 : \phi_1 =$ The series is level or trends non stationary,

If the test statistics of the KPSS test is less than the critical values then we accept the null hypothesis that is the time series data has the level or trend stationary.

2.3. ARIMA Models

A short form ARIMA stands for Auto-Regressive Integrated Moving Average. Here the Lags of differenced series that appear in the forecasting equation are called auto-regressive terms while lags of forecast errors are known as moving average terms. Also the time series which needs to be differenced to be made stationary is said to be an "integrated" version of a stationary series. A classical non-seasonal ARIMA model is written as $ARIMA\ (p, d, q)$ model, where p is the order of autoregressive terms, d is the number of non-seasonal differences and q is the order of lagged forecast errors (moving average) in the prediction equation. A process, X_t is said to be $ARIMA\ (p, d, q)$ if

$$(1 - B)^d\, X_t = \nabla^d\, X_t$$

is $ARMA\ (p, q)$ model. This means that the process is said to be stationary after differencing non stationary process d times. The general form of $ARIMA\ (p, d, q)$ model is given by:

$$\phi(B)\,(1 - B)^d\, X_t = \theta(B)\varepsilon_T$$

If we set $E\left(\nabla^d X_t\right) = \mu$, then the model becomes

$$\phi(B)\,(1-B)^d\, X_t = \alpha + \theta(B)\,\varepsilon_t,$$

Where, $\alpha = \mu\,(1 - \phi_1, - \cdots - \phi_p)$.

2.4 The Box – Jenkins ARIMA Models

The approach flows through identifying, fitting and Checking ARIMA models using the time series data. Forecasts follows directly after making sure all the procedures are true. By using Box-Jenkins, the p^{th} order of Autoregressive model; the general form of $AR\ (p)$ model is given by:

$$X_t = \alpha + \phi_1\, X_{t-1} + \phi_2\, X_{t-2} + \cdots + \phi_p\, X_{t-p} + \varepsilon_t,$$

Where, X_t is the dependent (response) variable at time t, $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$ are the response variables at time lags $t-1, t-2, \ldots\ t-p$ respectively. $\phi_1, \phi_2, \ldots, \phi_p$ are the Coefficients to be estimated and ε_t is the error term at time t.

The q^{th} order of Moving Average Model, $MA\ (q)$ model is given by:

$$X_t = \mu + W_t + \theta_1\, W_{t-1} + \theta_2\, W_{t-2} + \cdots + \theta_q W_{t-q} + \varepsilon_t,$$

Where, X_t is the dependent variable at time t, μ is the constant mean of the process, $\theta_1, \theta_2, \ldots, \theta_q$ are the coefficients to be estimated, ε_t is the error term and $W_{t-1}, W_{t-2}, \ldots, W_{t-q}$ are the errors in previous time periods in which they are normally included in the dependent variable X_t.

Autoregressive Moving Average Model (ARMA)

The general form of $ARMA\ (p,q)$ model is given by:

$$X_t = \alpha + \phi_1\, X_{t-1} + \phi_2\, X_{t-2} + \cdots + \phi_p\, X_{t-p} + \mu + W_t + \theta_1\, W_{t-1} + \theta_2\, W_{t-2} + \cdots + \theta_q W_{t-q} + \varepsilon_t$$

The graph of sample autocorrelation function (ACF) and partial autocorrelation function (PACF) are used to determine the model. The process is summarized in the Table below:

Table1. *Characteristics of the ACF and PACF for ARMA* Models (Shumway et al. 2006)

Model	ACF	PACF
AR(p)	Dies down	Cut off after lag q
MA(q)	Cut off after lag p	Dies down
ARMA(p,q)	Dies down	Dies down

2.5 The Box and Jenkins Seasonal ARIMA Model (SARIMA)

The weakness of ARIMA model is that it does not fit for observations with seasonality. To deal with this, (Box et al, 1976) introduce the general form of ARIMA model which deals with seasonality. The model is known as Seasonal ARIMA (SARIMA) model. Let $X_i\ (i = 1, 2, 3, \ldots, t)$ be a series under consideration. The Seasonal ARIMA (SARIMA) model for the series is given by (Box et al, 1976):

$$\phi(B)\Phi(B^S)[(1-B)^d(1-B^S)^D X_t] - \mu = \theta(B)\Theta(B^S)a_t$$

Or

$$\phi(B)\Phi(B^S)(W_t - \mu) = \theta(B)\Theta(B^S)a_t$$

Where, X_t is the time series observations at time t, t: is the discrete time, S is the seasonal length, μ: is the mean level of the time series process (Usually computed as average of W_t), note when $d + D > 0$ implies $\mu \equiv 0$, at: residual of the series, $NID(0;\delta^2)$, $\Phi(B^S)$: is the seasonal AR operator (polynomial $\Phi(B) = 1 - \Phi_1(B) - \Phi_2(B^2) - \cdots - \Phi_p(B^p)$, $(1-B)^D = \nabla_s^D$: is the seasonal difference operator of order $D\ (D = 0, 1, 2)$, $W_t = \nabla^d \nabla_s^D X_t$: is the stationary series formed after differencing X_t number of terms of W_t series are computed by $n = N - d - SD$, $\Theta(B^S)$: is the seasonal MA operator of order Q (polynomials $\Theta(B^S)$ = 1 - $\Theta_1(B^S)$ - $\Theta_2(B^{2S})$, note that $\Theta_1, \Theta_2, \ldots, \Theta_Q$ are the seasonal MA parameters and when $\Theta(B^S)$ = 0 means the root of the polynomials lies outside the circle.

SARIMA model is represented as $SARIMA(p;\ d;\ q)(P; D; Q)$, where: $(p;\ d;\ q)$ are the non-seasonal operator and $(P; D; Q)$ are the seasonal operator. Note: If the model is non seasonal, then only $(p;\ d;\ q)$ is required and if the model is seasonal then only $(P; D; Q)$ are needed.

Likewise, the sample ACF and PACF plots are used to determine the Seasonal ARIMA model. At the preceding stage, the values of p; q; P; Q are obtained by studying ACF and PACF plots. The characteristics of the plots are shown in the Table 2 below:

Table2. *Characteristics of the ACF and PACF for Pure Seasonal ARMA models (Shumway et al, 2006)*

Plot	AR(P)	MA(Q)	ARMA (P,Q)
ACF	Dies off at lags k's, k=1,2 …	Cur off after lag Q's	Dies off at lag k's
PACF	Cuts off after lag P's	Dies off at lag k's, k=1, 2…	Dies off at lag k's

2.6. The Box and Jenkins Algorithm

Normally, the approach used the previous values to give the predicted values. Box and Jenkins ARIMA time series model has the ability to generate the sequence of historical data and produce mathematical formula which will then be used to generate forecasted values. Studies have declared that the approach is convenient for short and medium predictions (Erhardt, 2001), also some articles have approved Box and Jenkins methodology as a very strong tools for giving solution of the prediction problems due to its ability to provide very tremendous correct prediction of the time series and also it yields a framework to develop the model and do analysis (Montgomery et al, 1967). The aims of using Box and Jenkins Prediction approach are to look for suitable formula that will force the error term to show no change in pattern and must be as small as possible. In this study the approach is used to develop the model and do prediction of monthly maximum and minimum temperature values. The Conceptual framework of Box and Jenkins modelling approach is given in the Table 3 below (Montgomery et al, 1967)(Box et al, 1976)

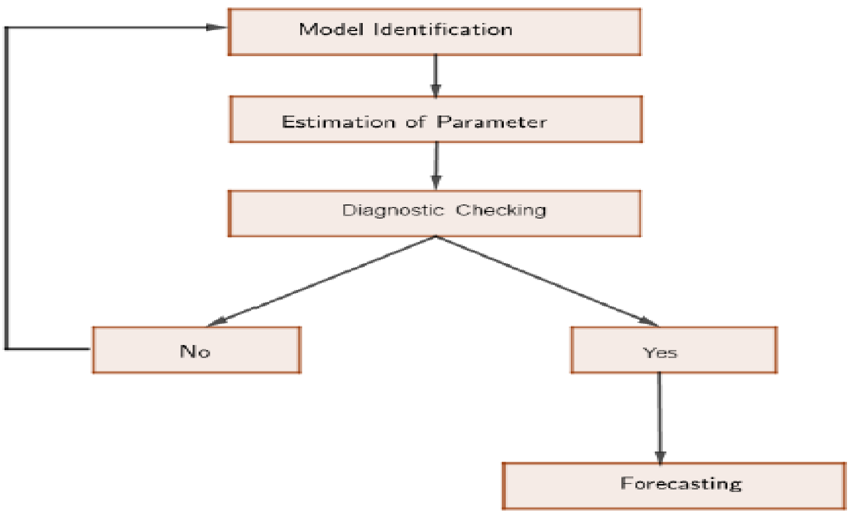

Figure1. *Box and Jenkins Conceptual framework*

In the first phase of Box and Jenkins algorithm is to determine if the time series is stationary and if there is significant seasonality needs to be modelled. The inspection of the time series observation will follows to check the suitable class of ARIMA model by selecting the order of the consecutive and seasonal differencing required making series stationary, as well as specifying the order of the regular and seasonal autoregressive and moving average series of polynomials required to precisely represent the time series model. In this paper the useful elements of time series analysis and forecasting called Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) will be used in determining the time series models. Normally, the ACF measures the amount of linear dependence between observations in a time series that are separated by a lag k while the PACF plot is required to determine how many auto regressive terms are necessary to show one or more of the following characteristics; time lags where high correlations appear, seasonality of the series, trend either in the mean level or in the variance of the series.

Stationary of the time series observations can also be identified by performing Portmanteau test, which is used to check whether observations is significantly different from a zero set. A common portmanteau test is the Box-Pierce test, designed by Box and Pierce (1970). This residual from a forecast model test is based on the Box-Pierce statistic:

$$Q = n \sum_{k=1}^{h} r_k^2$$

where, h is the maximum number of lags to be computed, n is the number of data set and r_k is the autocorrelation at $lag\ k$. If the residuals follows a white noise, the statistic Q has a chi-square (χ^2) distribution with degrees of freedom $(h - m)$ where m is the number of parameters in the model which has been fitted to the time series observations. An alternative portmanteau test is the Ljung-Box due to Ljung and Box (Kijazi A and Reason C, 2009). They argued that the alternative statistic:

$$Q^* = n\,(n+2)\ \sum_{k=1}^{n} {r_k^2}/{n-k},$$

the model has a distribution closer to the chi-square distribution with $(h - m)$ degrees of freedom than does the Q statistic. It is normal to conclude that the data are not white noise if the value of Q or Q^* lies in the extreme 5% of the right-hand tail of the χ^2 distribution.

To check whether the chosen model satisfies the normality, Shapiro-wilks test was used. The test statistics is given by:

$$W = \frac{(\sum_{i=1}^{n}(e_i m_i))^2}{\sum_{i=1}^{n}(m_i - \bar{m})^2},$$

where, m_i is the i^{th} order statistics, $\bar{m} = \frac{(m_1 + m_2 + \cdots + m_n)}{n}$ is the sample mean and constant e_i is given by: $e_1, e_2, \dots, e_i = \frac{k^T.V^{-1}}{(k^T.W^{-2}.k)^{1/2}},$

here, $k = (k_1, k_2, \dots, k_n)^T$ and $k_1, k_2, \dots, k_n$ are the expected value of the order statistics of IID random variables sampled from standard normal distribution and W is the covariance statistics of those order statistics.

The choice of the best model among the class of plausible models are done by using the informations criterion called Akaike's Information Criterion (AIC), proposed by Akaike (1974), Corrected Akaike Information Criterion (AICc) and Bayesian Information Criterion (BIC).

The AIC deals with minimizing the following quantity:

$$AIC = -2\ln(L) + 2\,(p + q + P + Q + C),$$

where, L is maximum likelihood, p and P are non-seasonal and seasonal autoregressive order respectively, q and Q are non-seasonal and seasonal moving average order respectively and C is the constant term of the model.

BIC is written mathematically as:

$$BIC = -2\ln(L) + 2\,(p + q + P + Q + C)\ln(N),$$

here, N is the sample size.

The model which has the minimum Information criterion value is our model of interest. Generally, the model with fewer number of parameters gives accurate forecasting (Chawsheen et al, 2017). After choosing the most suitable model (step 1 above), the next step is to estimate the model parameters (step 2) by using either the least square method or maximum likelihood estimator. In this step, we choose them values of the parameters to make the Sum of the Squared Residuals (SSR) between the observed data set and the estimated values as small as possible. Normally, non-linear estimation method is used to estimate the parameters specified to maximize the likelihood (probability) of the observed series given the parameter values. In this paper

Maximum likelihood estimation (MLE) is generally the preferred technique. In diagnose checking step (step three), we examine the residuals of the fitted models against adequacy. Normally, this is done by correlation analysis through the residual ACF plots. If the residuals are correlated, then the model should be refined as directed in step one above. Otherwise, the autocorrelations are white noise and the model is right choice for the set of time series data observed. After the application of the previous procedure for a given time series, a calibrated model will be developed which has enclosed the basic statistical properties of the time series into its parameters (step four).

3. Numerical Analysis and Discussion of Results

3.1. Overview

In this section, we present, discuss and give the interpretation of the results acquired from the study. In order to study suitable models for forecasting maximum and minimum temperature data of Dar es Salaam, Tanzania were used in the time series analysis. Thirty one (31) years daily maximum and minimum temperature data from 1985 to 2015 obtained from Tanzania Meteorological Agency (TMA) were used. The given daily temperature data was converted into monthly data by using the averaging method. Since temperature data are time dependent data then monthly data was converted into time series data and then smoothing moving average of five data points was applied. In this study, all numerical analyses were done by a statistical programming language R and in modal validation SPSS package was used. The rest of the section are subdivided into study area, preceding analysis, fitting the model, doing prediction and discussing the prediction accuracy.

3.2. Study Area

Dar es Salaam is one among the thirty regions of Tanzania, lying at the latitudes of 6°52' South and longitudes of 39°12' East. It is among highly populated coastal regions with population of 6,368,272 covering the area of 1,393km^2 (WPR, 2019). The region constitutes of five districts which are Kinondoni, Ubungo, Kigamboni, Ilala and Temeke.

Dar es Salaam region is characterized by tropical type of climate with higher degree of hotness, high humidity and average annual precipitation of over 1000 mm (UARK, 2017). The region is characterized by bimodal rainy seasons. The longer rain falls from March to May (MAM) and shorter rains fall from October to December (OND). The map of Tanzania and the extract of Dar es Salaam region from the map are exhibited in Figure

Figure2. *A map of Dar es Salaam region extracted from the map of Tanzania (IRA, GIS Laboratory UDSM)*

3.3.Time Plots for Temperature Data

Time series graphs are used to display the variation of temperature series. The smoothed plots of maximum and minimum temperature in Dar es Salaam region were shown in Figure 3 and Figure 4. From Figure 3, for maximum temperature there was clear validation on systematic change in mean, also the increasing trend is clearly seen. Moreover, in Figure 4 for minimum temperature the positive trend is seen in the plot.

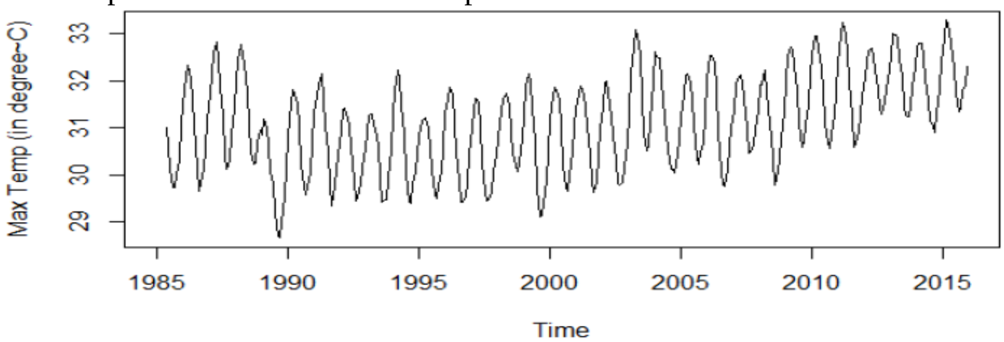

Figure3. *Smoothed Time Series Plot of Maximum Temperature from January 1985 to December 2015*

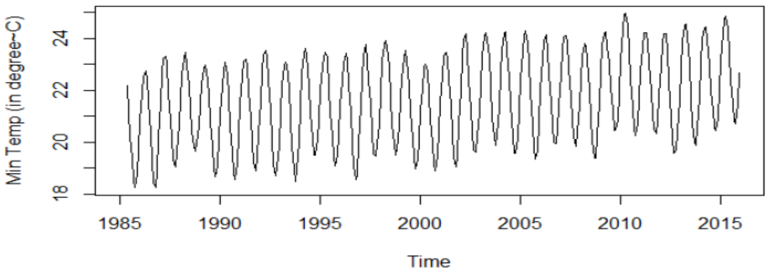

Figure4. *Smoothed Time Series Plot of Minimum Temperature from January 1985 to December 2015*

Both plots show evidence of change in temperature patterns and seasonal characteristics was observed in both plots, hence it is evident that the plots have strong yearly circle.

3.4 Descriptive Analysis

Summary of maximum and minimum temperature records (from January 1985 to December 2015) are illustrated in the Table below.

Table4. *Summary Statistics of Monthly Maximum and Minimum Temperature Series*

	Range	Highest value	Lowest value	Mean	Standard Deviation (SD)	Variance
Maximum Temperature (°C)	5.093	34.09	28.16	31.13	1.325018	1.755673
Minimum Temperature (°C)	9.0	26.02	17.02	21.65	1.230344	4.974434

The temperature summary shown in Table 4, reveals that the highest maximum temperature in Dar es Salaam is 34.09°C and was recorded in February 2003 while the lowest maximum temperature is 28.16°C and was recorded in July 1989. The lowest minimum temperature is 17.02°C and was recorded in July 1986 and the highest minimum temperature is 26.02°C and was recorded in December 2015. It was observed that the minimum temperature was more varying (Standard deviation (SD) = 2.230344) compared to the maximum temperature (Standard deviation (SD) = 1.32018).

The months of November, December, January and February recorded the highest maximum temperature while the lowest maximum temperature were recorded in June, July and August. The lowest minimum temperature was recorded in June, July and August while the highest minimum temperature was recorded in November and December.

With the fluctuations observed, both maximum and minimum temperature were seen to be not stable throughout the year. However, from March both minimum and maximum temperature dropped significantly till June and started rising again from July to November.

3.4 Decomposition of Time Series Data

We need to know whether time series data has trends, seasonal, cyclical and random components. For the case of monthly maximum and minimum temperature (in Figure 5 and Figure 6) the results reveal that both plots for maximum and minimum temperature series have random, seasonal and trend components. Also the positive trends were observed clearly for both maximum and minimum temperature series.

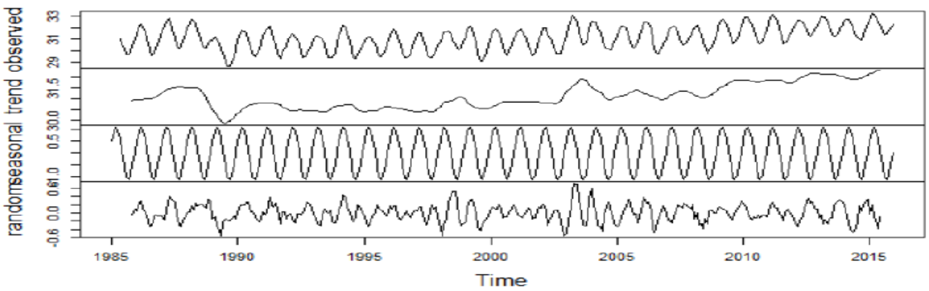

Figure5. *Decomposition of Smoothed Maximum Temperature Series*

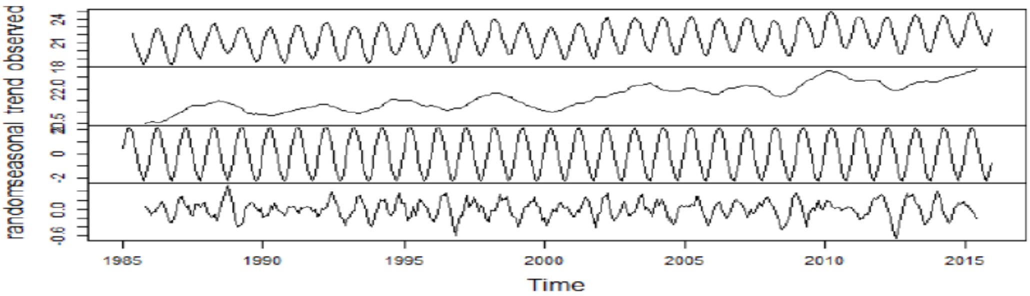

Figure6. *Decomposition of Smoothed Minimum Temperature Series*

Ordinarily, the Box and Jenkins methodology works under assumptions that the time series are stationary and serially correlated. In this study, graphical inspection and unit root tests were the most selected techniques for stationarity test of monthly maximum and minimum temperature observations.

3.5 Stationarity Tests

Graphical Approach

For the case of monthly maximum and minimum temperature, Figure 7 and Figure 8 exhibits the plots of ACFs and PACFs. The plots revealed a very strong seasonal sinusoidal patterns that decay slowly. The observations show that non-seasonal lags declined so fast.

Generally, ACF and PACF plots presented in Figure 7 and Figure 8, show almost all spikes at different lags are not within the confidence limits implying that the series are not stationary. Also the series are observed to have seasonality fluctuations that means seasonal differencing should be applied to convert non stationary time series data into stationary.

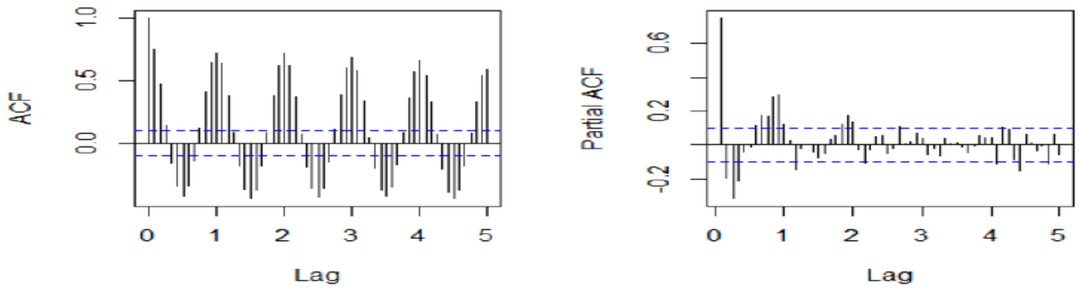

Figure7. *Autocorrelation and Partial Autocorrelation Function of Maximum Temperature*

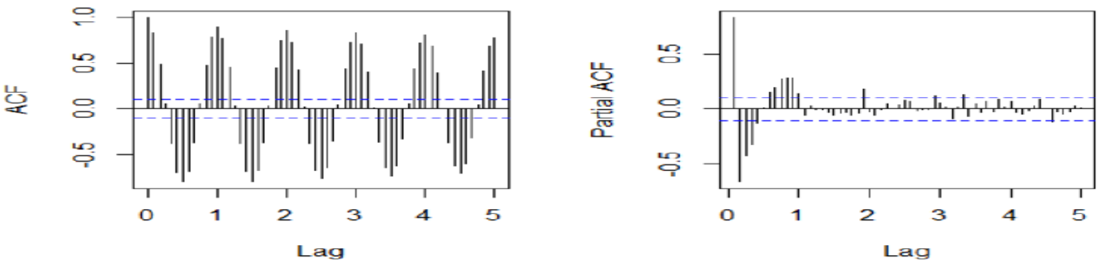

Figure8. *Autocorrelation and Partial Autocorrelation Function of Minimum Temperature*

The visual inspection technique used by plots showed that the monthly temperature time series were non-stationary. This was guaranteed by the fact that in ACF, as the number of lags increase the plots do not decay quickly. So statistical tests were conducted in order to verify the results of visual inspection technique. In this study ADF and KPSS tests were used for stationarity check.

Unit Root Test

Augmented Dickey Fuller Tests

The hypothesis testing for the stationary series was formulated as:

H_0: the time series has a unit root problem (It is non-seasonal and non-stationary).

H_1: the time series is stationary

Table5. *Augmented Dickey-Fuller Test for Monthly Maximum and Minimum Temperature Series*

	Dickey Fuller	Lag Order	p-value
Maximum Temperature	-7.5935	7	0.3901
Minimum Temperature	-10.599	7	0.0631

The ADF tests results of both maximum and minimum temperature series observed in Table 5, the null hypothesis is not rejected and therefore the two series are not stationary. This is guaranteed by the fact that both p-values for both maximum and minimum temperature series were greater than 5% (level of significance).

KPSS Test

The formal test with the KPSS test for stationarity was given by:

H_0: the time series has trend-stationarity against

H_1: the time series is non-stationary.

Table6. *KPSS Test for Monthly Maximum and Minimum Temperature Series*

	KPSS Level	Lag Parameter	p-value
Maximum Temperature (℃)	1.5498	5	0.01
Minimum Temperature (℃)	0.056401	5	0.01

It is important to note that, lack of unit root in ADF test does not necessarily mean that the series has trends stationarity. More statistical test analyses are required. The KPSS test results for temperature in Table 6, for both maximum and minimum temperature series we reject the null hypothesis, that is the series trend stationary since the p-value are 0.01 which is less than 0.05 level of significance. Hence conclude that monthly maximum and minimum temperature series are non-stationary.

Therefore the above three test results agree and suggest that there is non-stationarity in the original monthly maximum and minimum Temperature series. Natural logarithmic transformation and seasonal differencing are the mostly used techniques for eliminating non-stationarity from the time series observations (Box et al, 1976). Thus in this study seasonal differencing was used to remove non-stationarity effects.

3.6 Seasonal Differencing

Figure 9 and Figure 10 exhibit the time series plots for monthly maximum and minimum temperature series after performing seasonal differencing. It was observed that there is no clear trends or repetitive cycles for monthly maximum and minimum series and seasonality was seen in both monthly maximum and minimum temperature time series. Thus conclude that monthly maximum and minimum temperature series have now achieved stationarity.

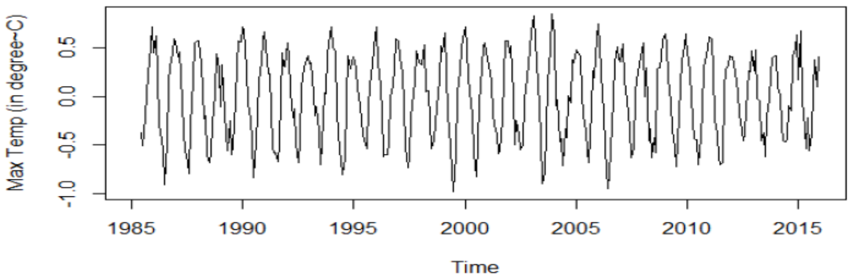

Figure9. *Smoothed Time Series Plot for Seasonal Differenced Maximum Temperature*

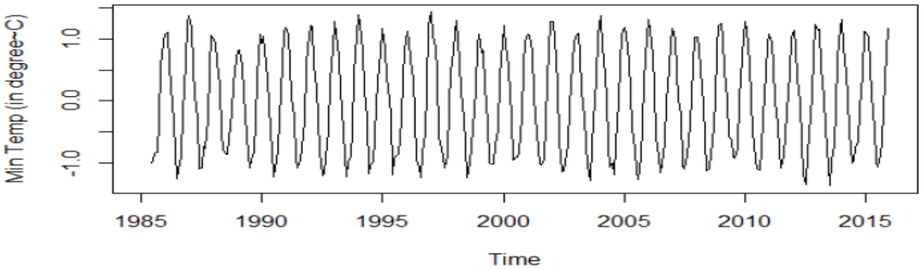

Figure10. *Smoothed Time Series Plot for Seasonal Differenced Minimum Temperature*

So more tests are required to ascertain that the monthly maximum and minimum temperature series have now attained stationarity. Janacek et al, (1993), suggests that the ACF and PACF plots can also be used to examine if the condition of stationarity is attained. The stationarity condition is achieved when the time series sequence converges very fast as the number of lags increases (Takele, 2012).

By using the visual inspection technique from Figure 11 and Figure 12, as the number of lags increases the autocorrelation function for series converges to zero. Also almost all spikes are within the 95% confidence limit for all plots which confirm that the monthly maximum and minimum temperature series are now stationary after performing seasonal differencing.

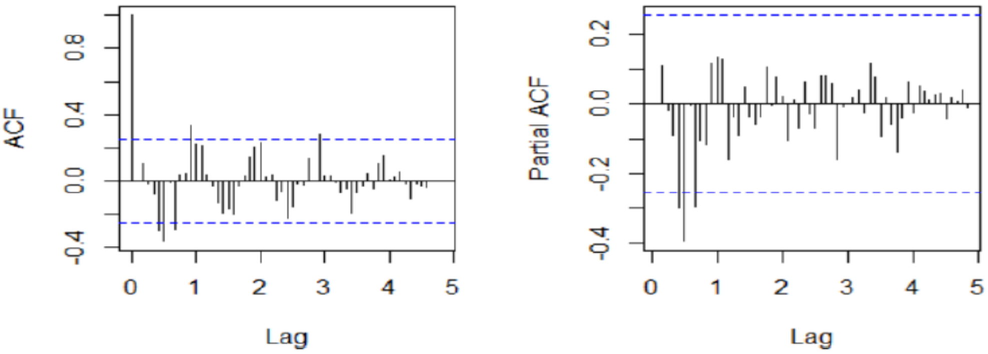

Figure11. *ACF and PACF Plot for Seasonal Differenced Maximum Temperature Series*

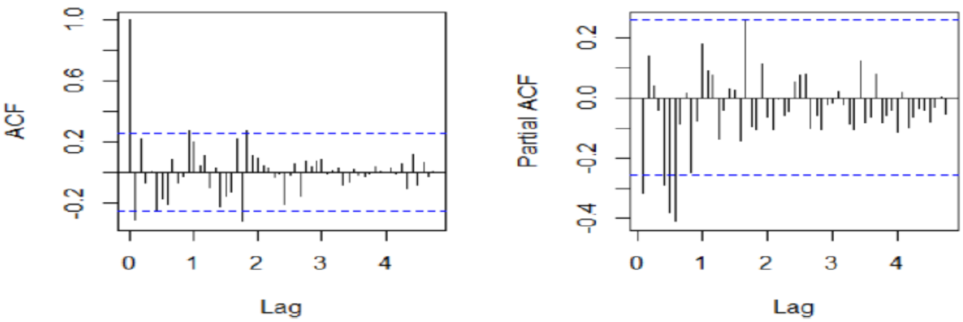

Figure12. *ACF and PACF Plot for Seasonal Differenced Minimum Temperature Series*

Formal statistical tests were performed again to validate the stationarity condition of the monthly maximum and minimum temperature series. The ADF and KPSS tests are used validate the stationarity of monthly maximum and minimum temperature series. According to ADF tests results for monthly maximum and minimum temperature series shown in Table 7, the null hypothesis is strictly rejected, which means there is a unit root at some confidence level. The test results is supported by the fact that the p-values for the tests are less than 5% and the DF values attenuate to negative. Thus conclude that the monthly maximum and minimum temperature series are now stationary.

Table7. *Augmented Dickey-Fuller Test for Differenced Seasonal Maximum and Minimum Temperature Series*

	Dickey Fuller	Lag Order	p-value
Maximum Temperature (°C)	-15.134	7	0.01
Minimum Temperature (°C)	-20.004	7	0.01

From the results of the KPSS test for monthly maximum and minimum temperature series shown in Table 8, the null hypothesis are not rejected because the p-values are greater than 5% (level of significance). This affirms that monthly maximum and minimum temperature series achieved trend-Stationarity.

Table8. *KPSS Test for Seasonal Differenced Maximum and Minimum Temperature Series*

	KPSS Level	Lag Parameter	p-value
Maximum Temperature (°C)	0.010589	5	0.1
Minimum Temperature (°C)	0.01229	5	0.1

3.7 Model Building for Monthly Temperature Data

It is crucial to note that when developing the model, it is important to plot the graphs of observations, do transformation if required, look for dependence coefficients of the model, estimate parameters of the model, do diagnostic checking for appropriateness of the model and select the appropriate model (Takele, 2012). In this sub-section, univariate Seasonal ARIMA (SARIMA) models was used to model monthly maximum and minimum temperature data.

3.7.1 Model Identification

After obtaining the value of differencing, the next step was to select the coefficients of autoregressive and moving average parameters by critically scrutinizing the sample autocorrelation and partial autocorrelation plots.

The value of p, q, P and Q were chosen based on Figure 7, Figure 8, Figure 11 and Figure 12, so as to get the speculative models. Since the study deals with approximate values, there is no fixed way of identifying if ACF or PACF dies down or cuts off (Shumway et al, 2006). Therefore at this stage, the numbers of provisionary values for p, q, P and Q were identified, then followed by the parameter estimation of the models. Since the climate variables (maximum and minimum temperature) data follows seasonality (annual cycle) then the suitable model is

$Seasonal\ ARIMA$ $(\mathrm{p, d, q})(\mathrm{P, D, Q})_{12}$. At first the rainfall and temperature time series data were non stationary hence the seasonal differencing was used to make it stationary. From Figure 11 and Figure 12, the ACF (at low lags) i.e at lag 1 and lag 2 are significantly different from 0 since the spikes passes out of the confidence limits. Hence the order of non-seasonal MA term is 2 and that of seasonal MA occurs at lags that are multiples of 12. Only one spikes (Figure 12) are significant at lag 12. Hence the orders of seasonal MA terms is 1. Similarly for the case of AR, significant spikes in the PACF (at lower lags) indicated possible non seasonal AR terms. The order of non-seasonal AR parts is 2 and that of seasonal AR part is 2. The order of seasonal differencing for monthly data is 12 ($S = 12$). Following the nature of ACF and PACF plots, the number of models has been identified and the most competing ones were suggested. So the following models were suggested:

Suggested Models for Monthly Maximum Temperature data

$SARIMA\ (0, 1, 1)(0, 1, 1)_{12}$,

$SARIMA\ (1, 0, 1)(2, 1, 1)_{12}$,

$SARIMA\ (2, 0, 1)(1, 1, 1)_{12}$

$SARIMA\ (1, 0, 1)(1, 1, 1)_{12}$

$SARIMA\ (2, 0, 2)(1, 1, 1)_{12}$

Suggested Models for Monthly Minimum Temperature data

$SARIMA\ (0, 1, 1)(1, 1, 1)_{12}$

$SARIMA\ (1, 0, 1)(1, 1, 1)_{12}$

$SARIMA\ (0, 1, 1)(1, 0, 1)_{12}$

$SARIMA\ (1, 1, 0)(2, 1, 1)_{12}$

$SARIMA\ (1, 1, 0)(1, 1, 1)_{12}$

$SARIMA\ (1, 1, 2)(1, 1, 1)_{12}$

$SARIMA\ (1, 1, 0)(2, 1, 0)_{12}$

3.8 Parameter Estimation

After identifying the competing models, the next step is to perform efficient estimation of the model parameters. The model parameters for autoregressive and moving average should agree with two conditions; they should be stationarity and invertibility (Nury et al, 2013). In this study it was assumed that the process X_i $(i = 1, 2, \dots, n)$ followed a normal invertible Gaussian $ARMA(p, q)(P, Q)$ process.

The parameter estimation by Box and Jenkins method was accompanied by the maximum likelihood approach computed, which relied on asymptotic condition for any time series observation (Brockwell et al, 2013). So maximum likelihood approach was used for monthly maximum and minimum temperature observations, in order to estimate the parameters of the models. The results for parameter estimates and selection criteria were presented in the Table 9 and Table 10

Table9. *Summary of Parameter Estimates and Selection Criteria for Maximum Temperature Models*

Model	Parameter	Estimate	SE	t-value	p-value	Criteria
SARIMA(0,1,1)(0,1,1)	MA1	-0.7245	0.0496	-14.6014	0.0000	AIC = -0.004796853
	SMA1	-1.0000	0.0442	-22.6148	0.0000	AICc = 0.0007548065
						BIC = -0.9837275
SARIMA(1,0,1)(2,1,1)	AR1	0.9018	0.0467	19.3119	0.0000	AIC = -0.05915081
	MA1	-0.6199	0.0876	-7.0779	0.0000	AICc = -0.05294734
	SAR1	-0.1400	0.0586	-2.3879	0.0175	BIC = -0.9959428
	SMA1	-1.0000	0.0962	-10.3930	0.0000	
	Constant	0.0041	0.0009	4.3681	0.0000	
SARIMA(2,0,1)(1,1,1)	AR1	1.1247	0.1044	10.7760	0.0000	AIC = -0.06808378
	AR2	-0.1603	0.0841	-1.9063	0.0574	AICc = -0.06188031
	MA1	-0.8007	0.0829	-9.6543	0.0000	BIC = -1.004876
	SAR1	-0.1704	0.0559	-3.0486	0.0025	
	SMA1	-0.9999	0.1861	-5.3725	0.0000	
	Constant	0.0042	0.0012	3.5491	0.1923	
SARIMA(1,0,1)(1,1,1)	AR1	0.9109	0.0433	21.0578	0.0000	AIC =-0.04268998
	MA1	-0.6339	0.0851	-7.4448	0.0000	AICc = -0.03669499
	SAR1	-0.1564	0.0619	-2.5272	0.0119	BIC = -0.9900167
	SMA1	-0.9735	0.1624	-9.5052	0.0000	
	Constant	0.0041	0.0010	4.2222	0.0000	
SARIMA(2,0,2)(1,1,1)	AR1	1.6087	0.3212	5.0085	0.0000	AIC = -0.6451125
	AR2	-0.6141	0.3065	-2.0036	0.0459	AICc = -0.05806852
	MA1	-1.2866	0.3372	-3.8151	0.0002	BIC = -0.9907686
	MA2	0.3371	0.2580	1.3064	0.0004	
	SAR1	-0.1612	0.0577	-2.7967	0.0054	
	SMA1	-1.0000	0.0788	-12.6831	0.0000	
	Constant	0.0040	0.0016	2.5209	0.0021	

Table10. *Summary of Parameter Estimates and Selection Criteria for minimum temperature models*

Model	Parameter	Estimate	SE	t-value	p-value	Criteria
SARIMA(0,1,1)(1,1,1)	MA1	-0.7573	0.0454	-16.6798	0.0000	AIC = -0.2460744
	SAR1	-0.0377	0.0597	-0.6304	0.5289	AICc = -0.240405
	SMA1	0.8896	0.0377	-23.5821	0.0000	BIC = -1.21447
SARIMA(1,0,1)(1,1,1)	AR1	0.7025	0.1167	6.0208	0.0000	AIC = -0.3080283
	MA1	-0.3835	0.1560	-2.4577	0.0145	AICc = -0.3020333
	SAR1	-0.0010	0.0611	-0.0160	0.9873	BIC = -1.255355
	SMA1	-0.9669	0.0442	-20.9808	0.0000	
	Constant	0.0051	0.0006	8.3756	0.0000	
SARIMA(0,1,1)(1,0,1)	MA1	-0.7473	0.0510	-14.6536	0.0000	AIC = -0.2406622
	SAR1	0.9997	0.0002	4042.2735	0.0000	AICc = -0.2348451
	SMA1	-0.9008	0.0337	-26.7128	0.0000	BIC = -1.198524
	Constant	0.0101	0.1549	0.0651	0.9451	
SARIMA(1,1,0)(2,1,1)	AR1	-0.3992	0.0487	-8.2053	0.0000	AIC = 0.1706845
	SAR1	-0.6625	0.0499	-13.2836	0.0000	AICc = 0.1763539
	SAR2	-0.3269	0.0501	-6.5200	0.0000	BIC = -0.7977115
SARIMA(1,1,0)(1,1,1)	AR1	-0.3794	0.0492	-7.7172	0.0000	AIC = -0.1110037
	SAR1	-0.0341	0.0588	-0.5801	0.5622	AICc = -0.1053344
	SMA1	-0.9487	0.0527	-18.0085	0.0000	BIC = -1.0794
SARIMA(1,1,2)(1,1,1)	AR1	0.3697	0.3532	1.0467	0.0052	AIC = -0.2780763
	MA1	-1.0022	0.3642	-2.7515	0.0062	AICc = -0.2720813
	MA2	0.0999	0.2805	0.3562	0.7219	
	SAR1	-0.0076	0.0606	-0.1258	0.0009	BIC = -1.225403
	SMA1	-0.9212	0.0430	-21.4168	0.0000	

Here, Table 9 and Table 10 show the estimates of parameters for the suggested models of monthly maximum and minimum temperature respectively. Each model was classified into its estimated value, standard error (SE), t-value, p-value, AIC, AICc and BIC. Box et al, (1976), alludes that parameter to be estimated should vary significantly from zero and all the significant parameters must be incorporated in the model. The results revealed that autoregressive, moving average together with both seasonal and non-seasonal autoregressive and moving average were significant since their p-values were lower than 0.05, and therefore must be preserved in the models.

3.9 Model Selection

The information criteria (AIC, AICc and BIC) were performed and compared to get the suitable model for monthly temperature data. The model with lower AIC, AICc and BIC was the best suited model. The analysis from Table 9 and Table 10 shows that $SARIMA\ (2,0,2)(1,1,1)_{12}$ and $SARIMA\ (1,1,2)(1,1,1)_{12}$ models has the smaller AIC, AICc and BIC, hence the models best fit the monthly maximum and minimum temperature data respectively. Moreover the parameters for the selected models were all significant since their p-values were less than 5% level of significance, hence they must be incorporated in the models.

3.10 Diagnostic Testing

In this sub-section we evaluates how the selected model effectively agrees with the given monthly maximum and minimum temperature data. Normally, if the model agrees with the data, then residual for the suited model has the behavior of randomness (Chatfield, 2003). Yurekli et al, (2005), asserts that in time series modelling using Seasonal ARIMA models, the selection of the suited model has the correlation with residual analysis computed. Thus, a number of statistical tests and different diagnostic plots can be used to explore how well the selected model fits the data.

The residual analysis began by plotting the residual plots for the monthly maximum and minimum temperature data, followed by examining if the selected model agrees with the data. Shumway et al, (2006), propounded that for the suited model, the residuals must be independent and identically distributed with the condition of having zero mean, constant variance and they are not serially correlated.

3:10:1 Diagnostic Checking for Monthly Maximum and Minimum Temperature Models

Figure 13 and Figure 14, show the statistical plots, standardized plot for residual and ACF plot for residuals and the Q-statistics plots for the residuals of monthly maximum and minimum temperature data models respectively. A normal QQ plot was useful in inspecting clear deviation of residuals from normality. As it was observed in the second panel of Figure 13 and Figure 14, almost all points ow in a straight line with just very few observations close to the line. This indicates that the residuals in the models are normal. Also the Shapiro-Wilk test of normality has a test statistics of W = 0.98769 and W = 0.9945 leading to a p-values of 0.3070 and 0.2031, for monthly maximum and minimum temperature models respectively. Hence the results confirm that normality is not rejected at any of the usual significance levels since the p-values are all greater than 5% (level of significance).

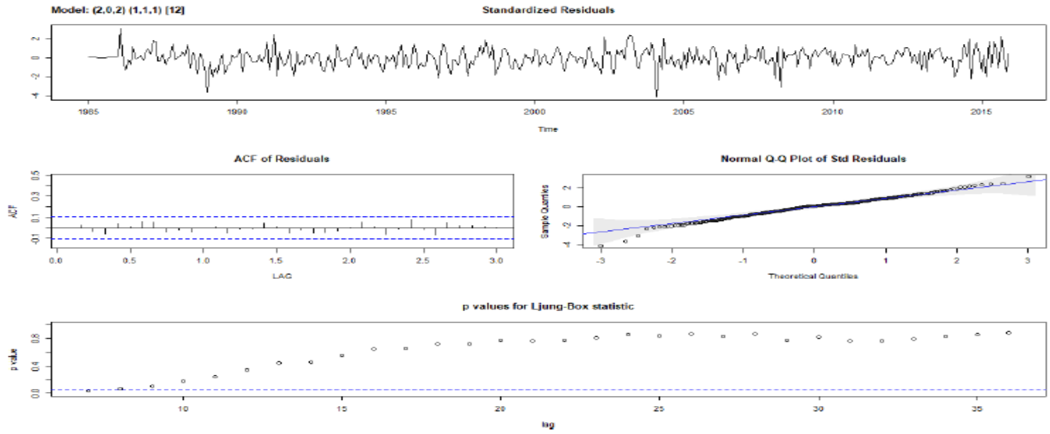

Figure13. *Residuals of Maximum Temperature Model*

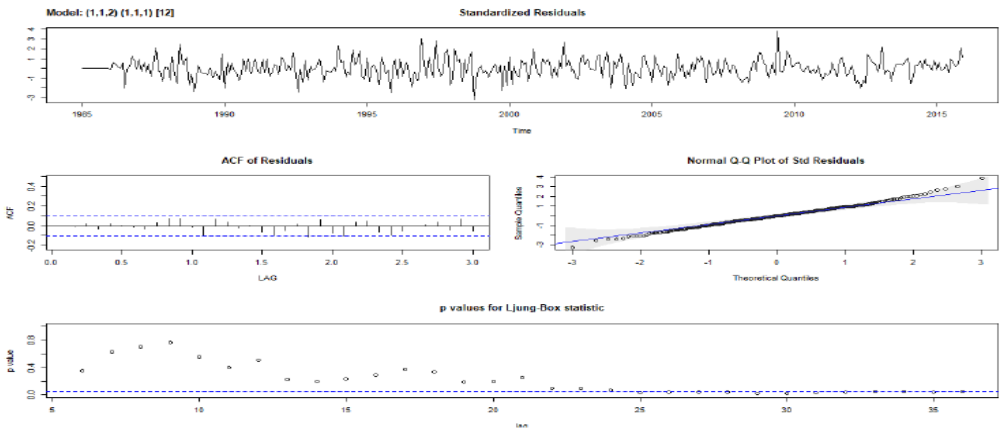

Figure14. *Residuals of Minimum Temperature Model*

Also, the first panel of Figure 13 and Figure 14, shows standardized residuals plots, which exhibit that the models residuals have zero mean and constant variance since the residuals were condensed around -2 to 2. Also plots for autocorrelation function (ACF) for residuals exhibits that, all the spikes for the two models (maximum and minimum temperature) are within the 95% confidence limit. This is an indication that the assumption of zero mean and constant variance for the model residuals were attained hence there was no correlation between residual values. Lastly, in order to test whether the residuals are white noise or not, the Ljung-Box test was computed. The test gives chi-squares of 10.261 and 19.301 under 20 degree of freedom and the p-values of 0.9632 and 0.5023 for monthly maximum and minimum temperature models respectively, which are shown in the third panel of Figure 13 and Figure 14. The Ljung-Box results proves to us that the null hypothesis cannot be rejected since the p-value was greater than the significance level and conclude that the residuals were free from serial autocorrelation. Moreover from the test it is observed that the value of the Ljung-Box test was over 5% for all lag coefficients. This confirmed that there was no clear significant diversion from the Gaussian white noise for the residuals, that is the null hypothesis declared that autocorrelation function up to lag 20 was concurrently equal to zero, meaning that it is correct.

Thus, $SARIMA\ (2,0,2)(1,1,1)_{12}$ and $SARIMA\ (1,1,2)(1,1,1)_{12}$ agree with all model assumptions. Hence they are appropriate models for monthly maximum and minimum temperature data in Dar es Salaam region of Tanzania.

3.11. Model Validation

In order to check the accuracy and the forecasting capability of the picked models $SARIMA\ (2,0,2)(1,1,1)_{12}$ and $SARIMA\ (1,1,2)(1,1,1)_{12}$ for monthly maximum and minimum temperature respectively, the actual observations and the fitted ones were plotted and presented in Figure 15 and Figure 16. The temperature data (maximum and minimum) from January 1985 to December 31, 2015 were designed as the test sets and were used to assess the ability of the models to fit the original data. The red and blue lines are the fitted and actual values respectively. The plots exhibited that, the fitted values (red lines) are very close to the original data (blue lines). This indicates that the selected models for maximum and minimum temperature series were the better ones for the set of data.

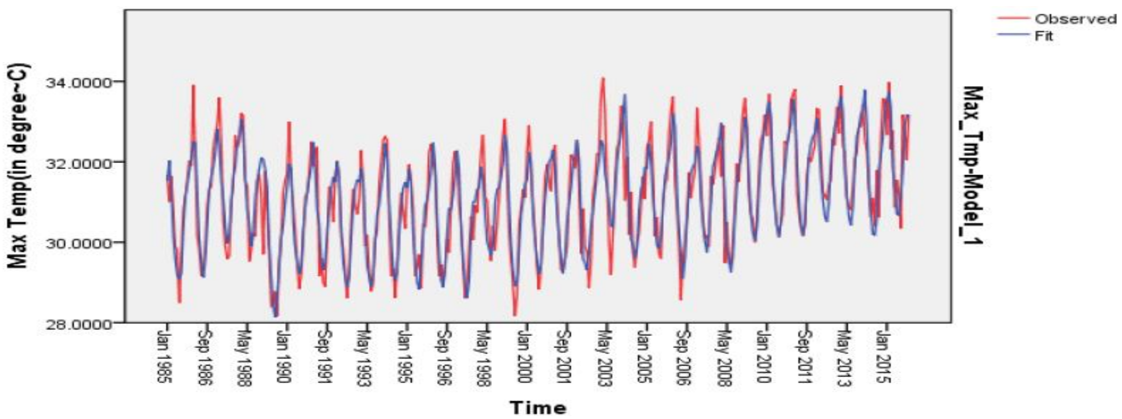

Figure15. *Observed and Fitted Values of Maximum Temperature Series*

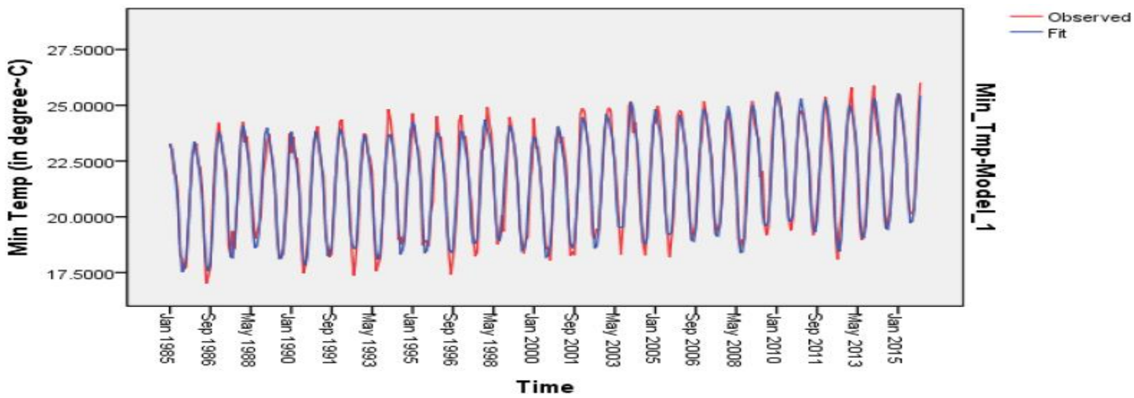

Figure16. *Observed and Fitted Values of Minimum Temperature Series*

3.12 Forecasting

After performing the parameter estimation and conclude that all parameters were significant and then conduct diagnostic analysis and confirm that residuals followed a normal property, the next step was to do forecasting. In this case the term forecast refers to the process of predicting the future monthly maximum and minimum temperature values of the studied time series. It should be noted that forecasting is important in decision making and planning process for all socio-economic sectors. In any time series analysis, getting the suitable model does not mean that it is a better model for prediction. Machekposhti et al, (2018), asserts that the superiority of the model depends on the measure of errors. So in this study, prediction performance were judged by a number of methods, in which the measures of errors such as MAE, MASE and RMSE were used. The performance measures obtained for the monthly rainfall, maximum and minimum temperature models respectively are shown in the Table 11.

Table11. *Forecasting Accuracy Statistic for Temperature Models*

Measure of Error	Maximum Temperature Model	Minimum Temperature Model
RMSE	0.5725927	0.5115882
MAE	0.442904	0.396314
MASE	0.587302	0.3729361

Normally, the best model must show low forecasting inaccuracy (Czerwinski et al, 2007). The performance measure of errors reported in Table 11, revealed that the prediction accuracy is high. This is because the MAE, RMSE and MASE for monthly temperature (maximum and minimum) models are all around zero which means lowest errors for the models. Thus, it is a good indication that $SARIMA$ $(2,0,2)(1,1,1)_{12}$ and $SARIMA$ $(1,1,2)(1,1,1)_{12}$ are appropriate models for forecasting monthly maximum and minimum temperature values respectively.

For the case of predicted maximum temperature specified by the blue line in Figure 17, it was observed that the trends of maximum will increase for the coming ten years. Also, it was noticed that the months of September, October, November, December, January and February in each year will have significant increasing trends. However from March to June the maximum temperature will drop significantly. Despite the variation of temperature, the results show that, February will remain to be the hottest month followed by January and December.

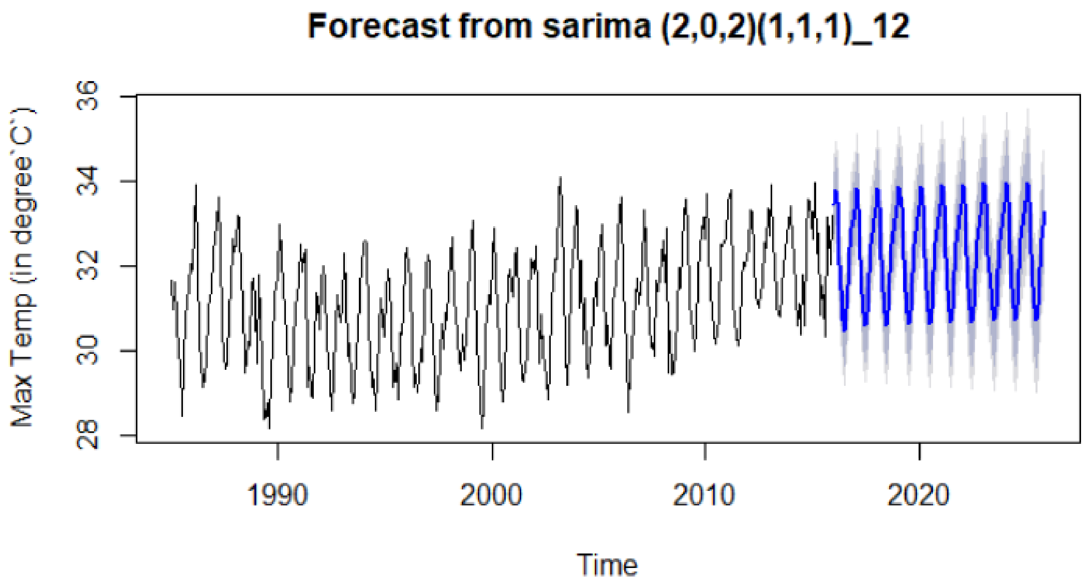

Figure17. *The Forecasted Maximum Temperature using* $SARIMA$ $(2,0,2)(1,1,1)_{12}$ *Model from January 2016 to December 2025*

In predicted minimum temperature designated by the blue line in Figure 18, it was observed that the trends of minimum temperature will increase for the upcoming ten years. Also the minimum temperature was noticed to have increasing trends for all months except from march to June. The findings pointed out January as a month with the highest minimum temperature followed by February and December.

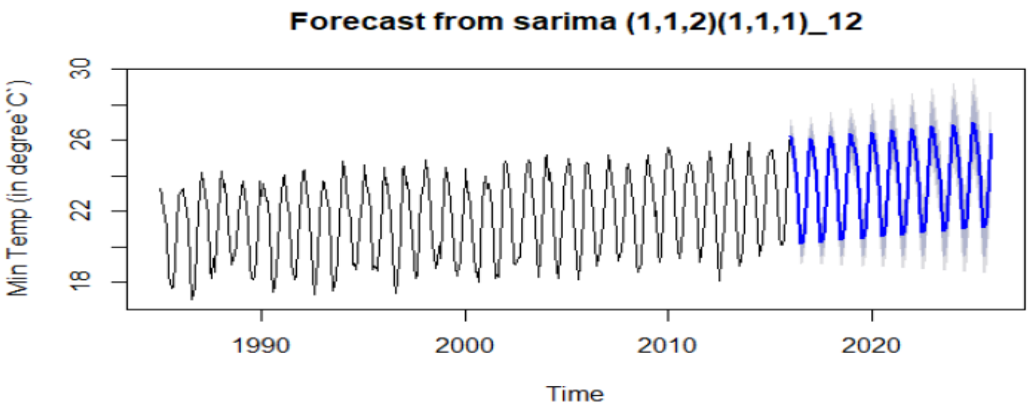

Figure18. *The Forecasted Minimum Temperature using* $SARIMA\ (1,1,2)(1,1,1)_{12}$ *Model from January 2016 to December 2025*

Generally, the maximum temperature is projected to increase by 0.20℃ per decades. Also the minimum temperature is predicted to increase by 0.98℃ per decades. The forecast plots (Figure 17 and Figure 18) indicates that the minimum temperature has a sharp increasing trends than maximum temperature. This result implies that the trends of minimum temperature for Dar es Salaam region will increase rapidly than the maximum temperature. Hence indicates that the region is expected to have warmer night in the future. This results is in line with few studies conducted in this area like (Chang'a et al, 2017), thus proves the fact that global warming is in fact a reality. Also the result gives the implication that there will be changing climate in the whole country. Moreover for an agro-based economy like Tanzania, the increase in forecasted temperature trends will threaten the significant achievements the country has made over the last decades in increasing incomes and reducing poverty. Hence, in view of these changes, it is necessary to uniformly and systematically assemble, examine and analyze the relevant climatic parameters like temperature for assessing the impacts of climate change.

Finally, the forecasted plots from Figure 17 and Figure 18 were observed to have minimal spread of confidence intervals from 2016 to 2020. However as time goes for example, from 2021 to 2025 the spread of confidence intervals seems to be higher implying that uncertainty of prediction becomes larger. Hence we again realized that the Box and Jenkins Seasonal ARIMA approach is the good method for short period of time forecasting of meteorological variables such as rainfall and temperature. This result is in line with few studies like that of (Erhardt, 2015).

4. Conclusion

The monthly temperature records of Dar es Salaam station in Tanzania has been studied using the Box and Jenkins methodology. Dar es Salaam monthly maximum and minimum temperature have shown to follows $SARIMA\ (2,0,2)(1,1,1)_{12}$ and $SARIMA\ (1,1,2)(1,1,1)_{12}$ respectively. The estimation and diagnostics analysis reveals that the models adequately fit the original data. Ljung-Box statistic indicated that the model is the better one for modelling of maximum and minimum temperature data. The residual analysis confirm that there is no violations of assumption connecting to model adequacy. The adequacy of the model also shown to be suitable and a forecast from 2016 to 2025 was made. The results show that there will be an increase in maximum and minimum temperature for the upcoming ten years. The increase in temperature suggests that climate change will continue to bring the negatively impact on different economic sectors and livelihood options in Dar es salaam community and this call for increased adaptive capacity for the community. With this magnitude of future climate change as forecasted in this study the less concerned of marginalized social groups would continue to remain to be attacked by the impacts of climate change unless deliberate efforts are put in place to help the community to adapt to climate change effects. Higher temperatures have effects on droughts, changing rainfall patterns and availability of surface water whose consequences range from less food supply to general fewer water supplies in Dar es Salaam region of Tanzania.

Acknowledgement

The authors are grateful to Tanzania Meteorological Agency (TMA) for supplying data. We would like also to thank African Institute of Mathematical Science (AIMS) – Tanzania for financial support.

REFERENCES

[1] Akaike, H. (1974). A new look at the statistical model identification. IEEE transactions on automatic control

[2] Box, G.E. & Jenkins, G.M. (1976). Time series analysis, control, and forecasting. San Francisco, CA: Holden Day, 3226(3228): 10.

[3] Brockwell, P. J., & Davis, R. A. (2013). Time series: theory and methods. Springer Science and Business Media,

[4] Chang'a, L.B., Kijazi, A.L., Luhunga, P.M., Ng'ongolo, H.K. & Mtongori, H.I. (2017). Spatial and Temporal Analysis of Rainfall and Temperature Extreme Indices in Tanzania.

[5] Chatfield, v. (2003). The analysis of time series: An introduction, CRC press, 6th edition

[6] Chawsheen, T. A., and Broom, M., (2017). Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq Journal of Statistical Theory and Practice, 11(4): 604-633

[7] Czerwinski, I. A., Gutierrez, J. C., Estrada. & Hernando, J. A. (2007). Casal Short-term forecasting of halibut CPUE, Linear and non-linear univariate approaches. Fisheries Research

[8] Erhardt, Erik, (2015). Box-Jenkins Methodology vs Rec.Sport. Unicycling 1999-2001. Available at : http : www:statacumen:com=pub=proj=WPI=ErhardtEriktsaproj.pdf (accessed 10- August)

[9] Ghahraman, Bijan, (2007). Time trend in the mean annual temperature of Iran. Turkish journal of agriculture and forestry 30: 439-448.

[10] IPCC. (2012). Climate Change, Managing the risks of extreme events and disasters to advance climatic change adaptation. The Physical Science Basic, Summary for policymakers. Contribution of working Group 1 to the Fourth Assessment Report, Paris

[11] Janacek, G., & Swift, L. (1993). Time Series Forecasting, Simulation, Applications. Ellis Harwood Limited, Chichester, West Sussex.

[12] Kebede, A. S. & Nicholls, R. J. (2012). Exposure and vulnerability to climate extremes: Population and asset exposure to coastal flooding in Dar es Salaam, Tanzania. Regional Environmental Change 12(1):81(94).

[13] Kijazi, A.L. & Reason, C.J.C. (2009). Analysis of the 1998 to 2005 drought over the northeastern highlands of Tanzania. Climate Research, 38(3): 209-223.

[14] Ljung, G. M. & Box, G. E. P. (1978). On a measure of lag of _t in time series model, Biometrika, pp 65:2, 297-303

[15] Mboera, L. E., Mayala, B. K., Kweka, E. J. & Mazigo, H. D. (2012). Impact of climate change on human health and health systems in Tanzania: a review. In Tanzania journal of health research, 13(5)

[16] McQuarrie, A.D. & Tsai, C. L. (1998). Regression and time series model selection. World Scientific

[17] Machekposhti, K. H., Sedghi, H., Telvari A & Babazadeh, H. (2018). Modeling Climate Variables of Rivers Basin using Time Series Analysis (Case Study: Karkheh River Basin at Iran) Civil Engineering Journal, 4(1): 78-92.

[18] Mishra, A. K. & Desai, V. R. (2005). Drought Forecasting Using Stochastic Models. Stochastic Environmental Research and Risk Assessment, pp 19(5): 326-339

[19] Mishra, U. & Jain, V. K. (2010). ARIMA forecasting of ambient air pollutants, Stochastic Environ. Res. Risk Assessment: 24:751-760. DOI: 10.1007/s00477-009-0361-8

[20] Montgomery, D. C. & Johnson, L. A. (1967). Forecasting and Time Series Analysis. McGraw-Hill Book Company, http://www.abebooks.com/Forecasting-Time Series-Analysis-Montgomery- Douglas/1323032148/bd, McGraw-Hill

[21] Ngailo, T. J., Reuder, J., Rutalebwa, E., Nyimvua, S. & Mesquita D. S. M. (2016). Modelling of Extreme maximum Rainfall using Extreme Value Theory for Tanzania. International Journal of Scientific and Innovative Mathematical Research (IJSIMR), PP 34-45 ISSN 2347-307X (Print)

[22] Nury, A.H., KochMand. & Alam, M.J.B. (2013). Time Series Analysis and Forecasting of Temperatures in the Sylhet Division of Bangladesh. In 4th International Conference on Environmental Aspects of Bangladesh (ICEAB), August: 24-26).

[23] Ojija, F., Abihudi, S., Mwendwa, B., Leweri, C.M. & Chisanga, K. (2017). The Impact of Climate Change on Agriculture and Health Sectors in Tanzania: A review. International Journal of Environment, Agriculture and Biotechnology, 2(4).

[24] Orindi, V.A. & Murray, L.A. (2005). Adapting to climate change in East Africa: a strategic approach (No. 117). International Institute for Environment and Development.

[25] Pankratz, A. (1983). Forecasting With Univariate Box-Jenkins Models Concepts and Cases. ISBN 0-471-09023- 9, pp: 414, John Wiley and Sons, Inc. New York, USA

[26] Roy, (2012). Time series analysis of Dibrugarh air temperature

[27] Shahin, M.A., Huq, M.M. & Ali, S. (2016). Time Series analysis, MOD Climate variable rainfall District in BANGLADESH.

[28] Shumway, R. H., & Stoffer, D. S. (2006). Time Series Analysis and Its Applications with R Examples 2Ed, Springer Science +Business Media, LLC, NY, USA.

[29] Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y, Bex, V., & Midgley, P. M. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate ChangeSummary for Policymakers. Cambridge University Press, Cambridge, United Kingdom; New York, New York, USA, 24 p

[30] Takele, R. (2012). Statistical Analysis of Rainfall pattern in Dire Dawa, Eastern Ethiopia (Doctoral dissertation, M. Sc. Thesis submitted to the Department of Statistics, Addis Ababa University)

[31] United Republic Of Tanzania (URT, 2007). Tanzania National Adaptation Programme of Action 2007, Vice President's Office-Dar es Salaam

[32] UNFCCC (2006, July 25). Climate change: Impacts, vulnerability and adaptation in developing coun- tries. Retrieved December 2, 2018, from https://unfccc.int/resource/docs/publications/impacts.pdf.

[33] Urban Africa Risk Knowledge (UARK, 2017). Dar es Salaam Climatic Profile: Full Technical Version

[34] World Population Review (WPR, 2019): http://worldpopulationreview.com/world-cities/dares-Salaam-population

[35] Yurekli, K., Kurunc, A. & Ozturk, F. (2005). Testing the residuals of an ARIMA model on the Cekerek Stream Watershed in Turkey. Turkish Journal of Engineering and Environmental Sciences

[36] Yusof, F. & Kane, I. L. (2012). Modelling monthly rainfall time series using ETS state space and SARIMA models. International Journal of Current Research

[37] Zakaria, S., Al-Ansari, N., Knutsson, S. & Al-Badrany, T. (2012). ARIMA Models for weekly rainfall in the semi-arid Sinjar District at Iraq

Citation: *Paul Andrew Panga, et.al., Monthly Temperature Prediction Based on Seasonal ARIMA Model: A Case Study of Dar es Salaam, Tanzania, International Journal of Scientific and Innovative Mathematical Research (IJSIMR), vol. 8, no. 4, pp. 28-46, 2020. Available : DOI: https://doi.org/10.20431/2347-3142.0804003*