International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-5, March 2020

Time Series Analysis and Modeling of Monthly Rainfall in Saudi Arabia

Mohamed M. A. Elgazzar, BenBella S. Tawfik

Abstract: ***Managing of water resources is an important future issue. Modeling is fundamental in preparation and organization of water resource system. Forecasting of occasions request identify proper models to be used in this process. Water is the main living source on earth. The most common and fundamental source of water on earth supporting the survival of the majority of life forms is Rainfall. Time arrangement investigation which incorporates modeling and estimating constitutes a instrument of foremost significance with reference to a wide extend of logical purposes in meteorology (e.g. precipitation, stickiness, temperature, sun powered radiation, surges and drafts). The show work applies the Box-Jenkins approach, utilizing SARIMA (Regular Autoregressive Coordinates Moving Normal) demonstrate is utilized to perform brief term estimates of month to month time arrangement such as precipitation. Modeling the past watched precipitation time arrangement values which result in utilized to anticipate long run amounts in agreement to the past. The demonstrate is tried by confirming the past precipitation information. In turn, the research produces a solid future figure. This show is assessed by implies of the AIC-, BIC-, and SBC- demonstrate.***

Keyword: Prediction; Auto Regressive Moving Average Models; SARIMA Models.

I. INTRODUCTION AND MOTIVATION

Rainfall could be well-thought-out as a non-linear natural process raising complexity. This is clear while trying to predict future values. Because of climate spatial variations, the rainfall is considered as a randomized stochastic process. Rainfall is considered as one of the major important issues which include harvesting, water supply limit plan, flood prediction, and many others. A assortment of factual methods are frequently utilized to figure precipitation sums.ARIMA modeling method has been connected on an awesome number of not as it were monetary but too hydrological time arrangement information in arrange to estimate precipitation information.

II. METHODOLOGY DESCRIPTION

2.1. Regional Area of interest

The time series observations for month / two weeks rainfall amounts in Saudi Arabia were used in this work. The data which is utilized for this research was for every two weeks / month (more than 500 observations) starting from 1964. Figure 1 and 2 shows a time series plot of the monthly and annual rainfalls at that station respectively.

Revised Manuscript Received on March 30, 2020.
* Correspondence Author
Mohamed M. A. Elgazzar*, Associate Professor, Delta Technological University, Egypt. E-mail: m_elgazzar1961@yahoo.com,
BenBella S. Tawfik, Associate Professor, Department of Computers and Informatics, Suez Canal University, Ismailia, Egypt. E-mail: benbellat@gmail.com

2.2. Box-Jenkins model building procedure

This model is introduced in 1970 by Box &Jenkins [1]. It is primarily developed for financial time series analysis.
The used model deals with steady state time series, and fitting either autoregressive moving normal (ARMA) or autoregressive coordinates moving normal (ARIMA) or regular autoregressive coordinates moving normal (SARIMA) models with the observed to find the foremost fitting coordinate of a time series perceptions. Moreover, the foremost precise expectation for future estimates.

2.2.1. Identification and selection of the Model

Identification and selection of the model is presented in this phase, the first step, verification the stationary of the variables. Second, positioning seasonality if exists, within the time series. This is introduced by analyzing the plot of autocorrelation and partial autocorrelation functions.

2.2.2. Model parameters estimation

The model is designed with different parameters, seasonal and non seasonal time series. The parameters are selected using iterative algorithm which try all the values and the select the ones with the minimum mean square error between the estimated time series using this model and the original observed values.

2.2.3. Testing and forecasting

This is done by testing whether the estimated model satisfy the stationary univariate process, or not. Specifically, the residuals should not be dependent between each other and must exhibit constancy in terms of mean / variance within the entire length of the time series. ARIMA models have three main types. These types are Autoregressive (AR) and Moving Average (MA) models, Autoregressive Moving Average (ARMA) models, and Seasonal (ARIMA) models.

III. DATA ANALYSIS

The data is being processed via four different phases, namely, data preprocessing, model design, implementation, testing.

3.1 Data Preprocessing:

In order to do a perfect data analysis which includes forecasting with minimum error, some preprocessing operations should be done. First, covering the missed data using linear interpolation or any other techniques is done. In many time series analysis, the log of the data is processed which is easy to be inverted. Five locations in Kingdom of Saudi Arabia are analyzed. In the first location Bahrah, 46 years (1966-2011), the first analysis is done for the sum of rainfall per each month. Another variable is studied which is the storm day within each year. The calendar is arranged as a day number (the month can be estimated) ranges from (1-365) or (1-366) according to February number of days.

Figure 1 illustrates the first time series sum of rainfall per month within the whole interval (552 values). Figure 2 describes the storm day occurrence for each year within the 46 years.
Figure 3 shows the maximum daily rainfall in each year within the interval.

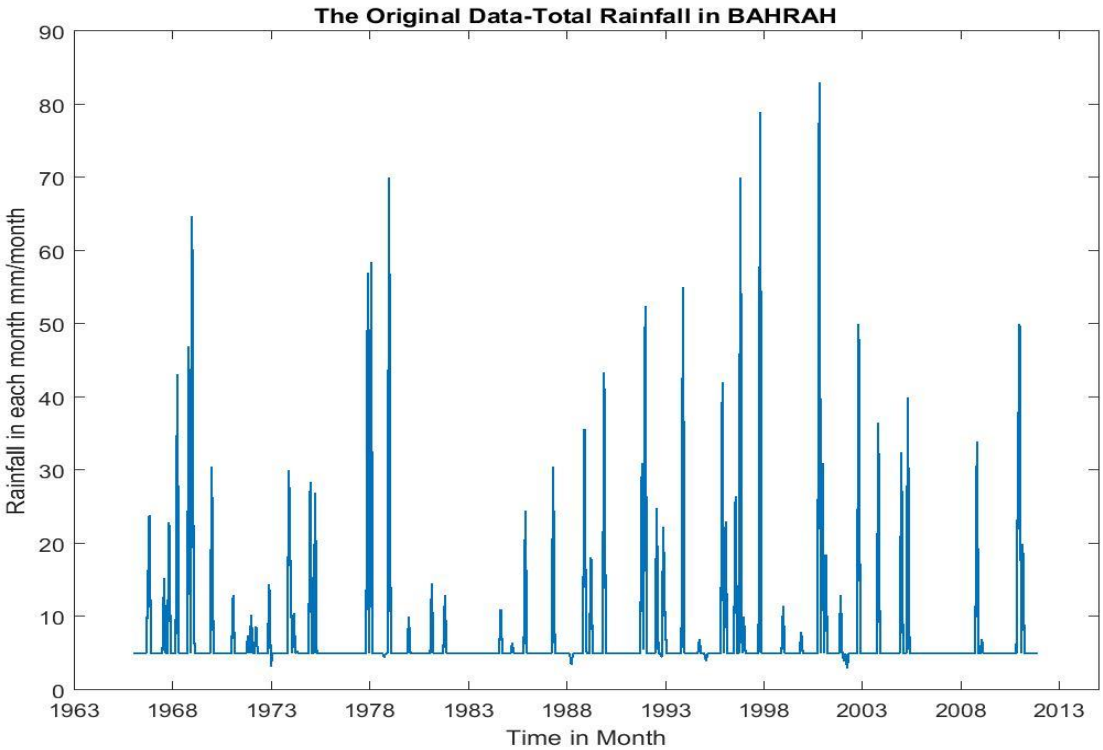

Figure 1: The sum rainfall mm/month within the 552 months in Bahrah Location

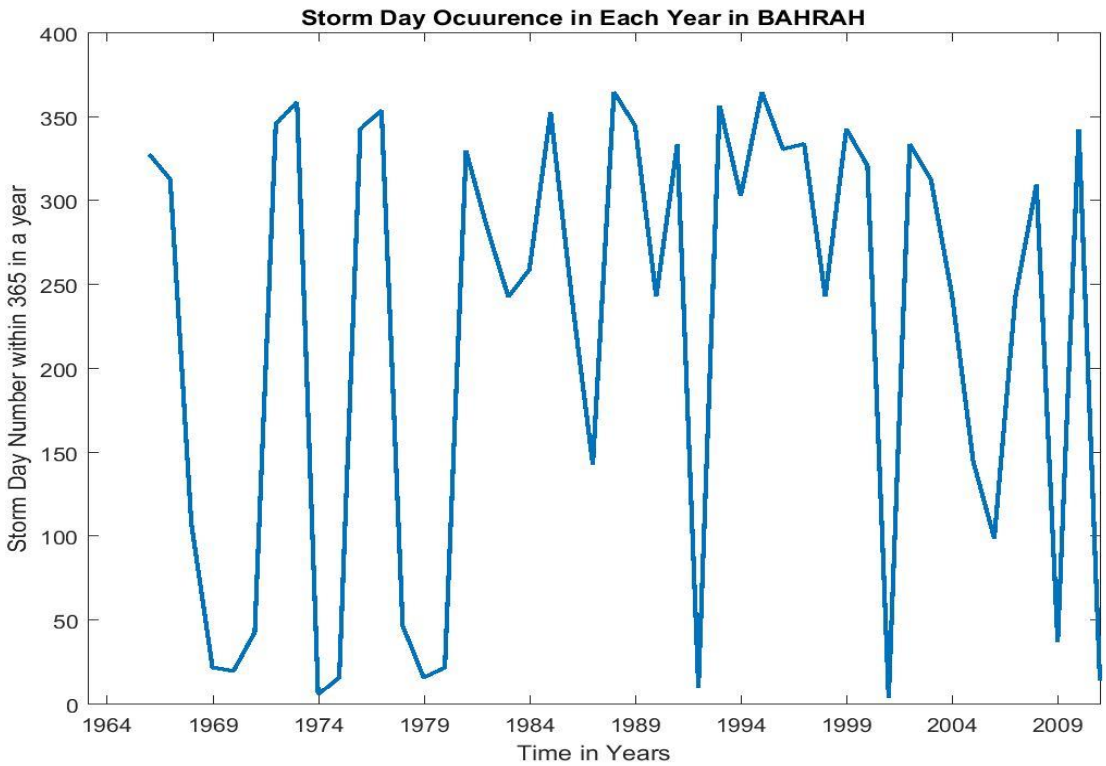

Figure 2: Storm day number within the years in Bahrah

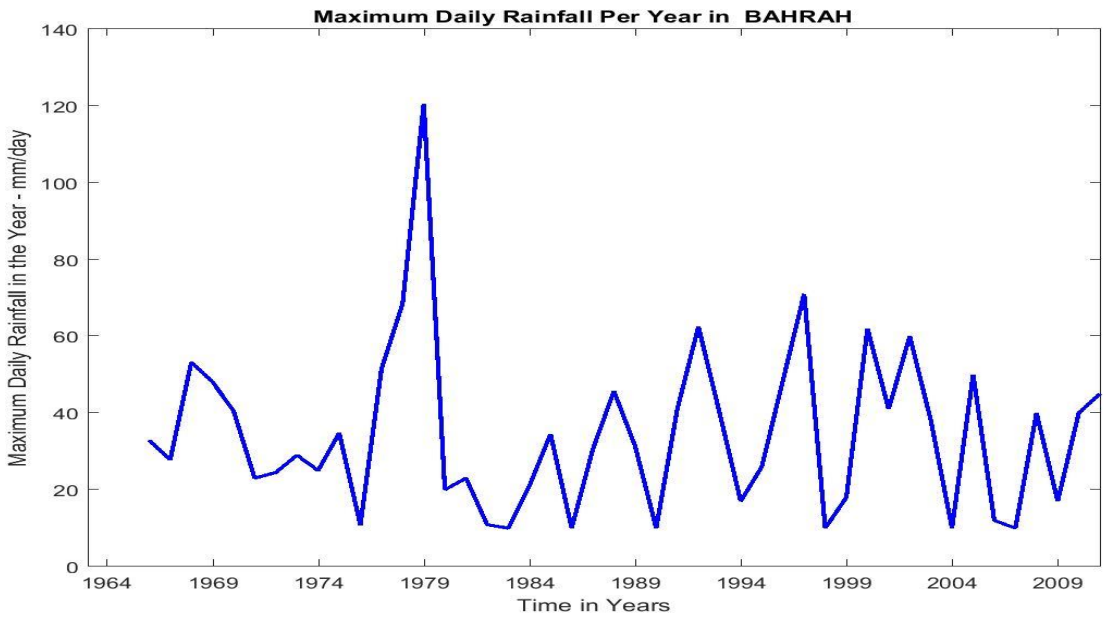

Figure 3: Maximum daily rainfall mm/day in each year

3.2 Model Design:

Much iteration are done in this phase to select the best fit parameters of the model (ARIMA)

ARIMA(0,1,1) Model Seasonally - Integrated with Seasonal MA(12):

Distribution: Name = 'Gaussian'
P1: 13
D1: 1
Q1: 13
Constant: 0
AR1: {}
SAR1: {}
MA1: {NaN} at Lags [1]
SMA1: {NaN} at Lags [12]
Season: 12
Variance: NaN

Second, check the residuals for normality. One presumption of the fitted demonstrates is that the advancements take after a Gaussian dissemination. Induce the residuals, and check them for typicality.

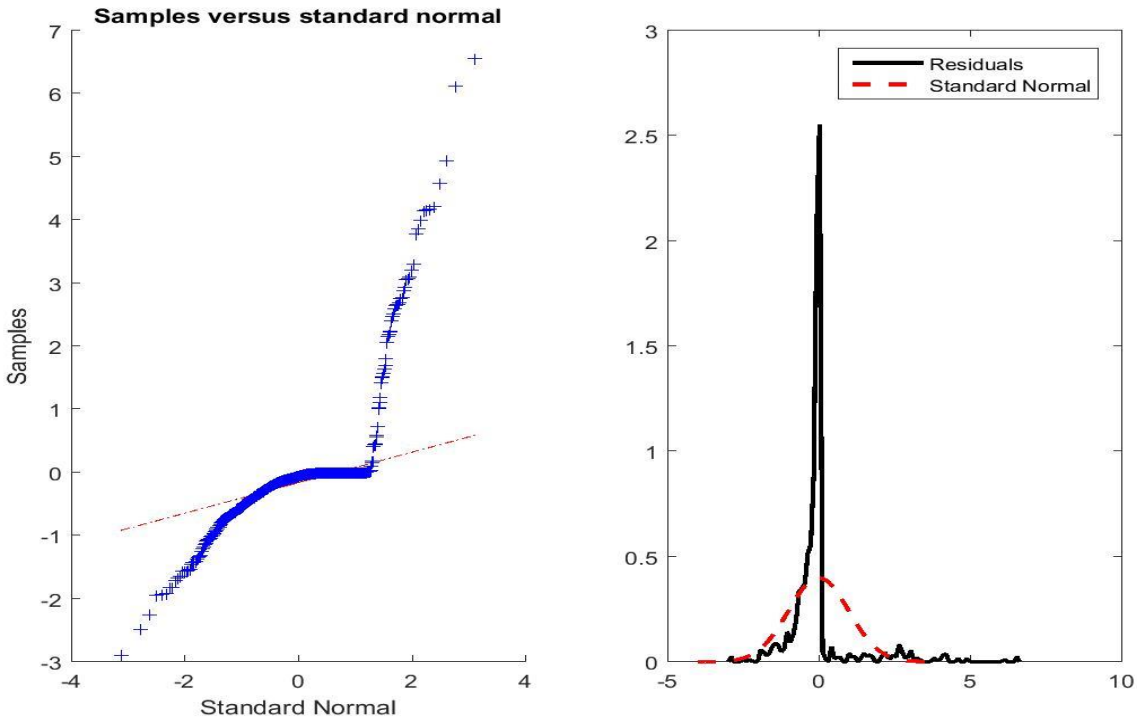

Figure 4: Quantile and Quantile plot (QQ-plot) and Kernel Density

Quantile and Quantile plot (QQ-plot) and Kernel Density estimate illustrates the normality assumption. The next step, the residuals are checked for autocorrelation. The objective is to verify that the residuals are uncorrelated. See at the test autocorrelation work (ACF) and fractional autocorrelation work (PACF) chart for the standardized residuals depicted in figure 5.

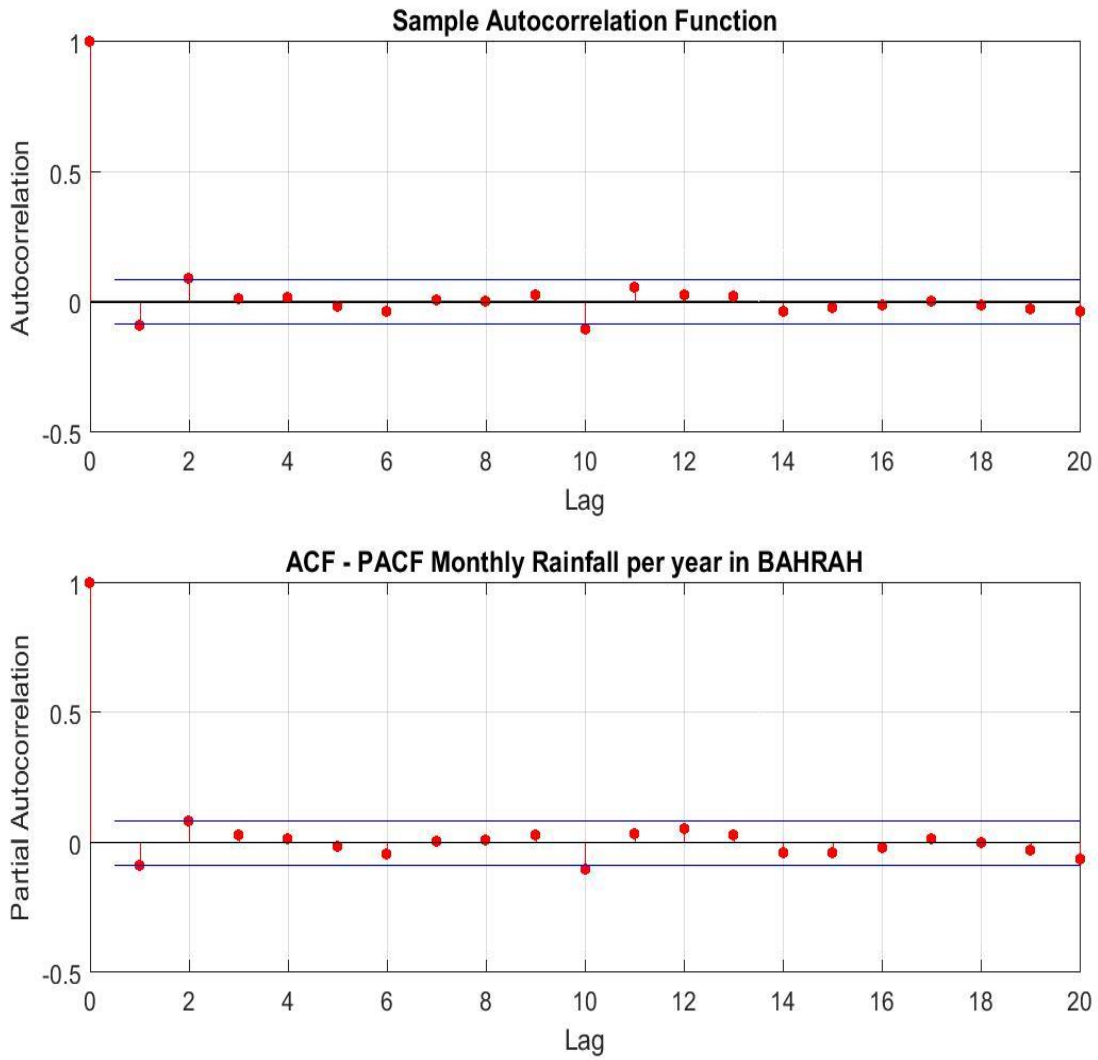

Figure 5: Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots versus the Standardized Residuals

The ACF and PACF plots are illustrate in figure 5. It can be concluded that there is no significant autocorrelation. More formally, conduct a Ljung-Box Q-test at slacks 5, 10, and 15, with degrees of flexibility 3, 8, and 13, individually. The degrees of flexibility account for the two assessed moving normal coefficients. The Ljung-Box Q-test affirms the test ACF and PACF comes about. The invalid theory that all autocorrelations are together rise to zero up to the tried slack which isn't rejected (h = 0) for any of the three slacks.

3.3 Model Testing:

This phase is concerned with checking the performance of predictive. Use a holdout sample to compute the predictive MSE of the model. In order to achieve this goal, a comparison between the estimated time series and the observed one is illustrated in figure 6.

ARIMA(0,1,1) Model Seasonally Integrated with Seasonal MA1(12):

```
---------------------------------------------------------------
Conditional Probability Distribution: Gaussian
                          Standard                    t
 Parameter      Value       Error              Statistic
 -----------   -----------  ------------       -----------
 Constant          0         Fixed              Fixed
   MA1{1}         -1       0.00879247        -113.734
 SMA1{12}     -0.830567    0.0213176         -38.9616
 Variance      0.146902    0.00496624         29.5802
pmse =      0.13287
```

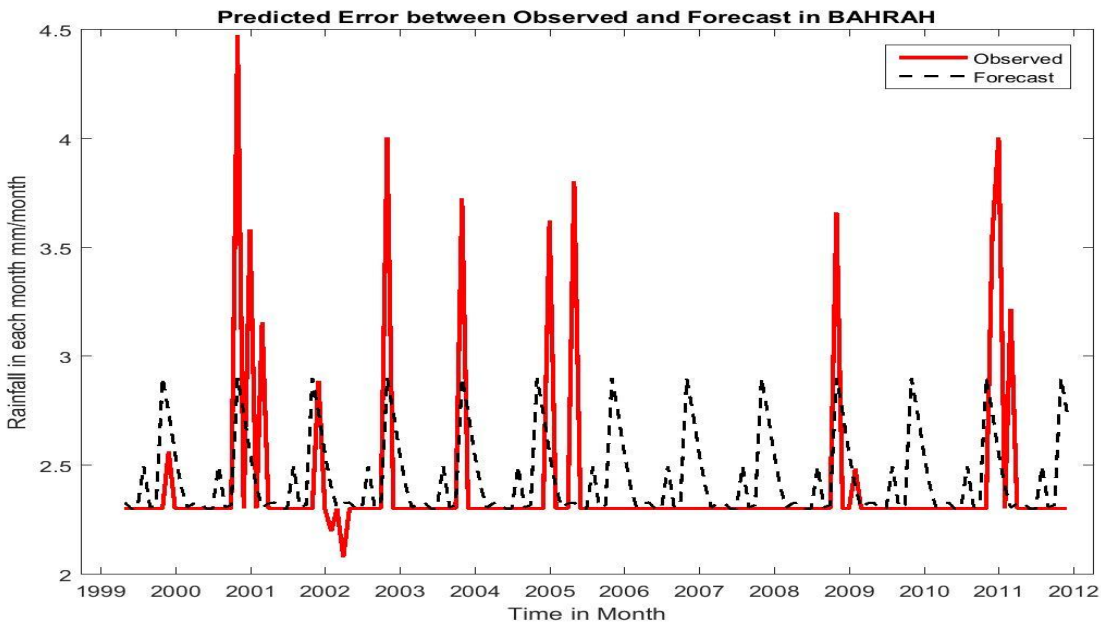

Figure 6: Predicted and observed time series-monthly rainfall

The residuals are checked for the normality. One presumption of the fitted demonstrates is that the advancements take after a typical conveyance. Gather the residuals, and check them for typicality as appeared in figure 7.

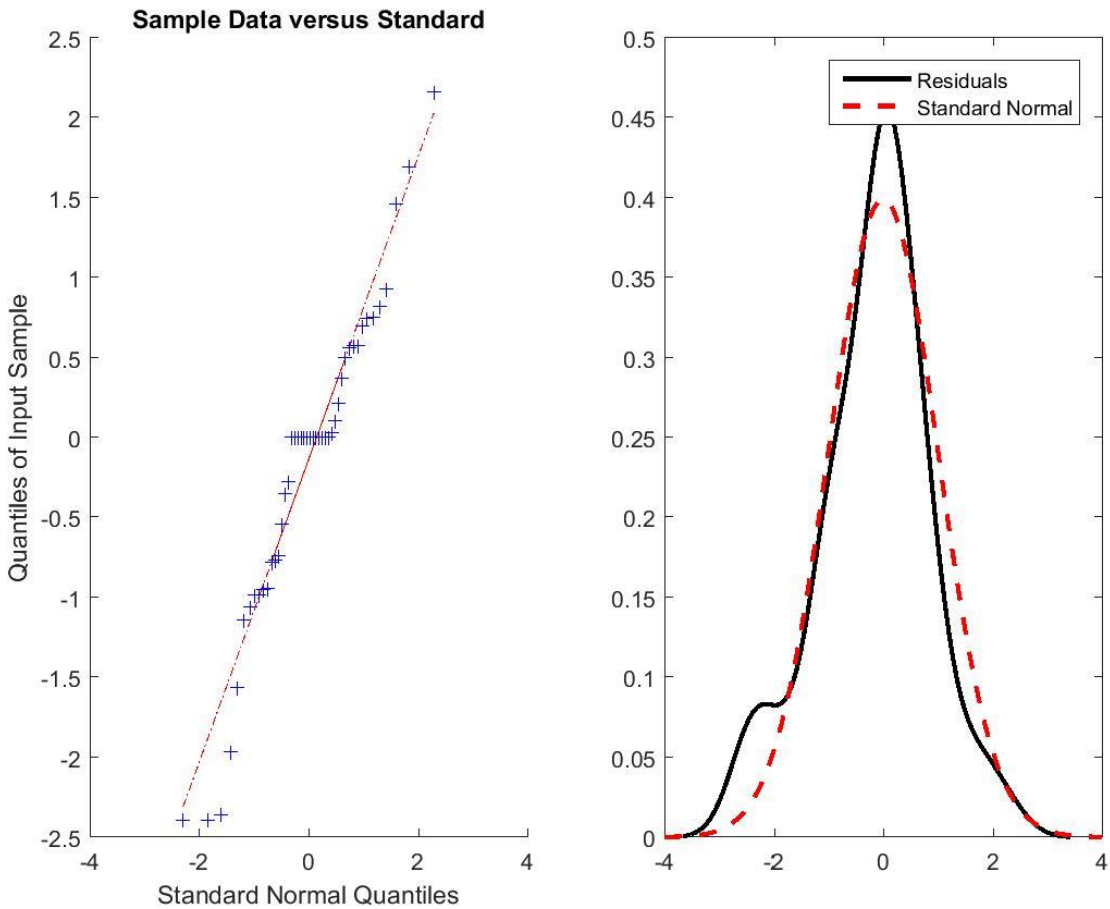

Figure 7: Check the normality of the residuals.

Quantile and Quantile plot (QQ-plot) and Kernel Density estimate confirm the normality assumption. Also, from checking the residuals for autocorrelation, it is clear that the residuals are uncorrelated. Also, the sample autocorrelation function (ACF) is shown.

At that point, halfway autocorrelation work (PACF) plots for the standardized residuals. Figure 8 appears a Test Autocorrelation Work (ACF) and Halfway Autocorrelation Work (PACF).

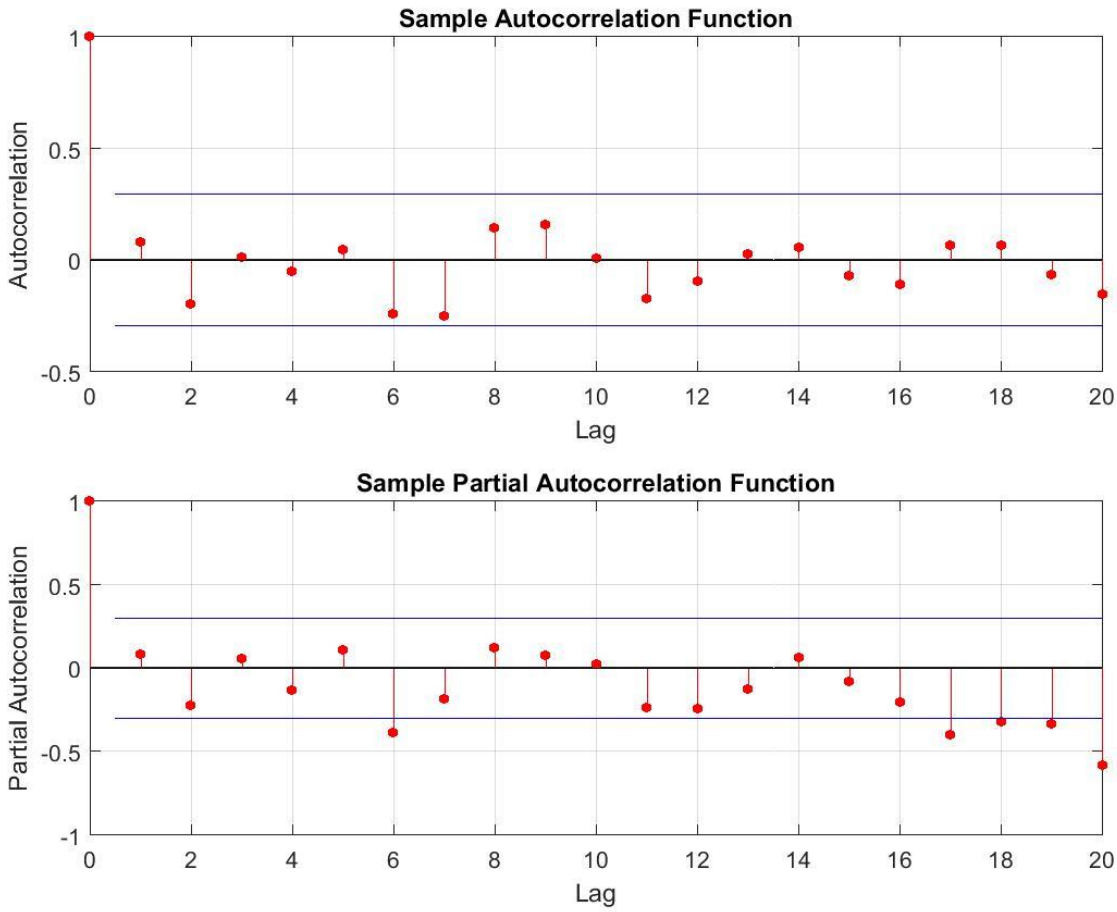

Figure 8: Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for the Standard Residuals.

These steps are repeated for the maximum rainfall per month in the whole time interval. The final result is presented in figure 9, observed time series compared with the predicted. Finally the storm day prediction is illustrated in figure 10.

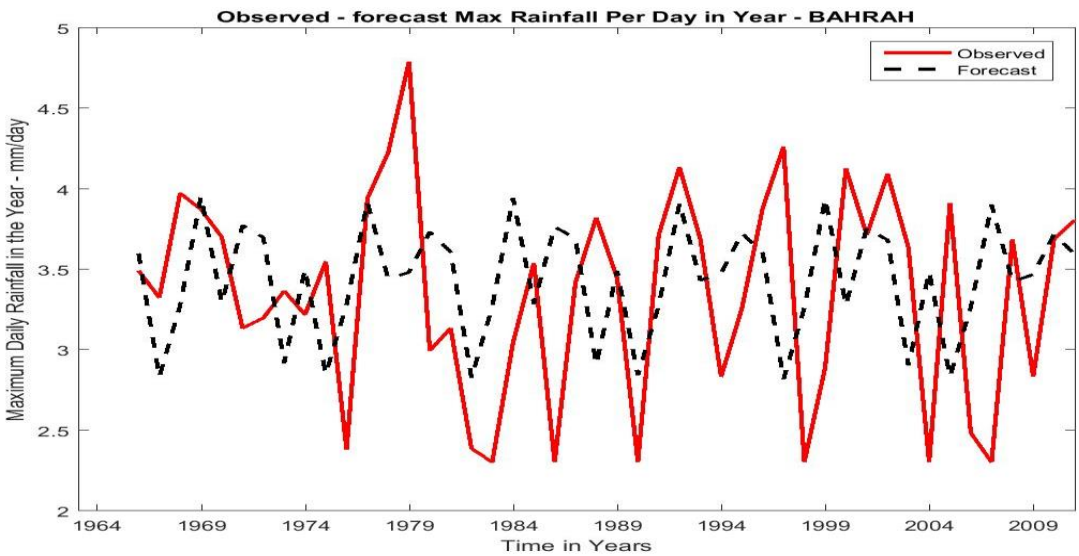

Figure 9: Maximum Daily Rainfall prediction in the year

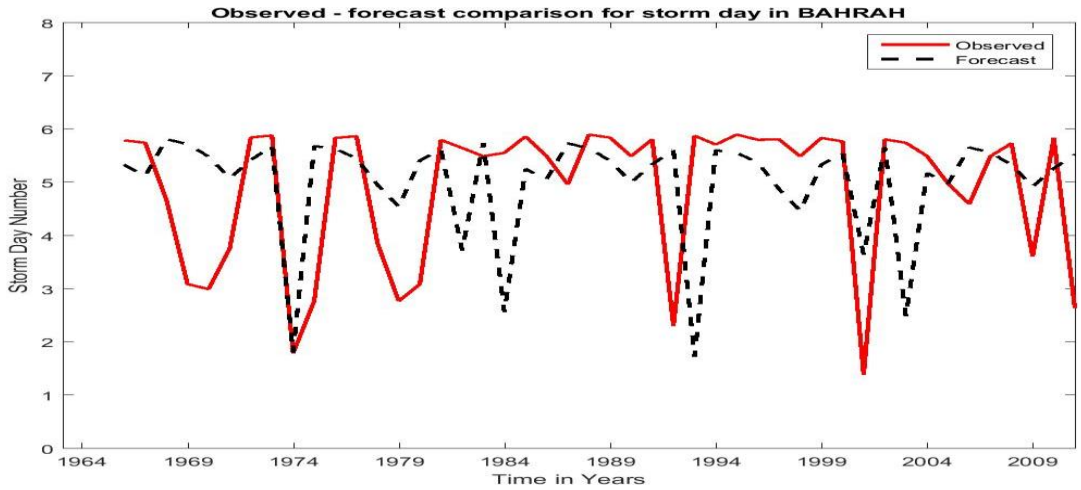

Figure 10: Storm day Prediction

IV. THE RESULTS ANALYSIS

This research introduces a general model for time rainfall analysis.

The first step is data preparation. The data are originally formulated in relational excel sheet. The results of this step are described in figures 1, 2, and 3. Figure 1 illustrates the sum rainfall mm/month within the 552 months in Bahrah Location, figure 2 shows the storm day number within the years in Bahrah, and figure 3 shows maximum daily rainfall mm/day in each year. The second step is parameters selection. This step gives the system generality because the parameters selection is done automatic (using programming module). The selected parameters are tested with Peak Mean Square Error (pmse) with the observations. The work is looking for the minimum pmse. The last step is to illustrate the success of the proposed system. Figure 6 shows the predicted and observed time series-monthly rainfall. Figure 9 illustrates maximum daily rainfall prediction and the observed in the year. Finally, figure 10 describes the storm day Prediction and observed. This analysis proves the generality of our prediction system.

V. CONCLUSIONS

This work introduces a generic model for time rainfall analysis. The results of this research prove that AIRMA modeling is a capable tool for simulation and modeling of Rainfall in arid and semi-arid regions. The proposed models are able to preserve the seasonal statistics of the observed data. The presented work emphasizes the importance of data preprocessing. Also, it is recommended to spend some effort for picking the best parameters. Testing the models by different methods is presenting in this work. Good preprocessing plus finding the best parameters model result in accurate forecasting.

REFERENCES

1. George E. P. Box, Gwilym M. Jenkins, Book 1970Time seris analyses, forcasting and control.
2. Fokianos, K., and Tjostheim, D. (2011), "Log-linear Poisson autoregression," Journal of Mul-tivariate Analysis, 102, 563–578.
3. Adi Nugroho,Simanjuntak , ARMA Model for Prediction of Rainfall in Regency of Semarang - International Journal of Computer Science Issues, Vol. 11, Issue 3, No 1, May 2014 ISSN (Print): 1694-0814 .
4. Morena D, Pueg P, Reos J, Vilella A, Trilla A (2011) A statistcal model for hosptal admissions caused by seasonal deseases.
5. Peng Chen, Aichen Neu, a, Duanyng Leu, Time Series Forecasting of Temperatures using SARIMA: Materials Science and Engineering (2018).
6. A.A. Hanafi-Bojd, M.R. Yaghoobi-Ershadi, Modeling the distribution of cutaneous leishmaniasis vectors : a potential transmission in disease prone areas, 52 (4) (2015), pp. 557-565
7. M Sidiq, Forecasting Rainfall with Time Series Model, INCITEST, IOP Publishing,IOP Conf. Series: Materials Science and Engineering (2018).
8. Duong Tran Anh , Thanh Duc Dang , Improved Rainfall Prediction Using Combined Pre-Processing Methods and Feed-Forward Neural Networks, J 2019.

AUTHORS PROFILE

Mohamed M. A. Elgazzar, Graduated from MTC in 1984 and He got his Master in Computer Networks from Faculty of Engineering, Cairo University, Egypt, in 1991 and the Ph.D. degree in Computer Networks at Faculty of Engineering, from Cairo University, Egypt, in 1995. He has experience of 33 years which includes both academic and research. His research interests Computer Networks, Network Security He published several research papers.

BenBella S. Tawfik, Born in September 1964, Cairo, Egypt, Graduated from Military Technical College in 1986. He got his Master in Computer Engineering in 1991 from the same school of graduation. He earned his Ph.D. from Colorado State University USA in 1998. In 2006, he visited USA-Colorado State University and earned Post Doctor Certificate in Computer Engineering. Besides my research work, he worked as a part time professor in many schools of computer engineering/computers and informatics in Egypt since 1998. Now, he is working as an assistance professor – Suez Canal University - since October 2012. Two years ago, he becomes an associate professor working as a dean of the mentioned department.