Journal of Earth Sciences and Geotechnical Engineering, vol.2, no. 3, 2012, 25-55
ISSN: 1792-9040(print), 1792-9660 (online)
Scienpress Ltd, 2012

ARIMA Models for weekly rainfall in the semi-arid Sinjar District at Iraq

Saleh Zakaria[1], Nadhir Al-Ansari[2], Sven Knutsson[3], Thafer Al-Badrany[4]

Abstract

Time series analysis and forecasting is an important tool which can be used to improve water resources management. Iraq is facing a severe water shortage problem. The use of rainwater harvesting is one of the techniques to overcome this problem. To put this into practice, it is of prime importance to forecast future rainfall events on a weekly basis.

Box-Jenkins methodology has been used in this research to build Autoregressive Integrated Moving Average (ARIMA) models for weekly rainfall data from four rainfall stations in the North West of Iraq: Sinjar, Mosul, Rabeaa and Talafar for the period 1990-2011. Four ARIMA models were developed for the above stations as follow: $(3,0,2)x(2,1,1)_{30}$, $(1,0,1)x(1,1,3)_{30}$, $(1,1,2)x(3,0,1)_{30}$ and $(1,1,1)x(0,0,1)_{30}$ respectively. The performance of the resulting successful ARIMA models were evaluated using the data year (2011).These models were used to forecast the weekly rainfall data for the up-coming years (2012 to 2016). The results supported previous work that had been carried out on the same area recommending the use of water harvesting in agricultural practices.

Keywords: Time series, weekly forecasting, Sinjar, Iraq.

---

[1] Lulea University of Technology, Lulea, Sweden, e-mail: Saleh.Zakaria@ltu.se
[2] Lulea University of Technology, Lulea, Sweden,
e-mail: Nadhir.Al-Ansari@ltu.se
[3] Lulea University of Technology, Lulea, Sweden, e-mail: Sven.Knutsson@ltu.se
[4] Mosul University, Mosul, Iraq, e-mail: thafer@yahoo.com

1 Introduction

Iraq is facing a serious water shortage problem [1]. This implies the need to use new techniques to overcome this problem. Among these methods is the water harvesting technique. The Sinjar area was chosen to evaluate the possibility of using this technique [2; 3; and 4]. To ensure optimal use of the technique, decision makers and farmers require prediction of the future rainfall expected in their locality.

There are several methods dealing with time series forecasting, the most relevant is Box-Jenkins methodology which used in this study. It is discussed in several publications [e.g. 5; 6; 7; 8; and 9].

Chiew et al., [10] conducted a comparison of six rainfall-runoff modeling approaches to simulate daily, monthly and annual flows in eight unregulated catchments in Australia. They concluded that a time-series approach can provide adequate estimates of monthly and annual yields in the water resources of the catchments.

Kuo and Sun, [11] employed an intervention model for a 10 day average of stream flow forecast and synthesis which was investigated to deal with the extraordinary phenomena caused by typhoons and other serious abnormalities of the weather of the Tanshui River basin in Taiwan.

Langu, [12] used time series analysis to detect changes in rainfall and runoff patterns to search for significant changes in the components of a number of rainfall time series.

Al-Ansari et.al [13, 14] and Al-Ansari and Baban [15], examined the rainfall record of all the Jordanian Badia stations, for the period 1967-1998 to determine periodicity and interrelations between stations using power spectral, harmonic analysis, and correlation coefficient techniques. They used an ARIMA model to forecast rainfall trends in individual stations up to 2020. Their results showed that the rainfall intensity has been decreasing with time for most of the stations.

Al-Hajji, [16] used an ARIMA model to forecast time series of monthly rainfall. His results show good sequence of correlation and suitable ARIMA models for the monthly rainfall time series for nine rainfall stations in the north of Iraq.

Al-Dabbagh, [17] used two time series in order to represent and forecast the data using ARIMA methodology. The first was the inflow series to the three reservoirs dams (Mosul, Dokan, and Derbendikhan) in the north of Iraq. The second was a rainfall series for rain falling on the above reservoirs. The results showed that a seasonal ARIMA model is suitable to describe the monthly rainfall and inflow series to the three reservoirs.

Weesakul and Lowanichchai [18] used both of ARMA and ARIMA models to fit the time series of annual rainfall during 1951 to 1990 at 3l rainfall stations in Thailand. They concluded that the ARIMA model is more suitable to predict inter-annual variation of annual rainfall.

Somvanshi et al. [19] made a comparative study of rainfall behaviour as obtained by ARIMA and the artificial neural network (ANN) techniques, using mean annual rainfall data from year 1901 to 2003 in India. Their research established that the ANN method should be favored as an appropriate forecasting tool to model and predict annual rainfall rathar than that using an ARIMA model.

The ARIMA model has been used by various authors to forecast rainfall and other meteorological variables [20, 21, 22, 23 and 24].

Previous work carried out on the Sinjar area [2, 3 and 4] indicated that water harvesting techniques can be successfully used in this area. In order to put these techniques into practice, decision makers and farmers should have an idea about likely future rainfall events in the area.

ARIMA models, included in this research, could be harnessed with management of water resources by estimating future amounts of runoff to be obtained from weekly forecasting rainfall. These weekly runoff data can support supplemental irrigation for the rain-fed farms of wheat crop in the Sinjar district to overcome the problem of water scarcity.

It will also lead to an increase of the efficiency of the irrigation process and consequently the crop yield.

2 Study Area and Rainfall Data

The rainfall stations of Sinjar, Mosul, Rabeaa and Talafar are located within the vicinity of the Sinjar district in the North West of Iraq (Fig.1). The data were provided by Iraqi Meteorological and Seismology Organization for the period 1990-2011, assuming no missing data for all stations except the Talafar station, where data for several years was missing, the inverse distance method [25] was used to estimate the missing data.

Sinjar district is characterized by its semi-arid climate, where rainfall totals are low and there is an uneven distribution. The rainy season extends from November to May [4].

Rainfall on weekly basis of these 4 rainfall stations has been used to develop the ARIMA models. The mean weekly rainfall for the above stations was as follow: 10.2, 11.3, 9.9, and 10.3 (mm) respectively for the period 1990-2011.

The data of these stations, fluctuate greatly with wide variation where its value may ranged between zero to maximum value (more than 100 mm) in addition the maximum value rarely repeated, so it is not easy to find suitable ARIMA models to represent them unless enough trials have been applied. The complexity may be increased due to using weekly period to represent a seasonal period (S).

In general there are about 30 rainy weeks every year for the Sinjar district, which represent the S term in the general form of ARIMA model of $(p,d,q)x(P,D,Q)_S$.

Seasonal differencing was applied on all the time series from the 4 stations after taking a log transformation function, and then applying ARIMA models using Minitab software Release 14.1, the ARIMA model describes the seasonal differencing time series.

Fig. 1: (A) Map of Iraq shows the Sinjar District, (source: lonelyplanet.com).
(B) Meteorological stations: Rabeaa, Sinjar, Talafar and Mosul, (source: Google map).

The time series was divided into three periods:

The first period, from 1990 to 2010, where the data were used to analyze the characteristics of the rainfall and selecting the most appropriate rainfall forecast models. *The second period*, is the final year, 2011, that was used for evaluating the performance of the selected models. *The third period*, the selected model was used to forecast rainfall time series for the up-coming 5 years (2012-2016).

3 ARIMA Model

The multiplicative ARIMA model, as a short term, stands for Autoregressive Integrated Moving Average.

The acronyms AR (p) is known for an Autoregressive model of order (p), and represented by:

$$x_t = \sum_{j=1}^{p} \Phi_j x_{t-1} + \varepsilon_t \qquad (1)$$

Where x = observation at time=t, Φ =jth autoregressive parameter.

ε_t =independent random variable represent the error term at time t,

x_{t-1} = time series at the time (t-1), p= order of autoregressive process.

The acronyms MA (q) is known for a moving average model of the order q and is represented by:

$$x_t = \varepsilon_t - \sum_{j=0}^{q} \Theta_j \varepsilon_{t-1} \qquad (2)$$

Where: Θ =jth moving average parameter, q= order of moving average process.

The combination between AR (p) and MA (q) models is called the ARMA (p, q) model, and is represented by:

$$x_t = \sum_{j=1}^{p} \Phi_j x_{t-1} - \sum_{j=0}^{q} \Theta_j \varepsilon_{t-1} \tag{3}$$

To achieve stationary case in the time series, it may be differenced ARMA model for d times to obtain ARIMA (p,d,q), similarly an ARMA model may be seasonal differenced for D times to obtain seasonal ARIMA $(P,D,Q)_S$ for S seasonal period. So when they are coupled together that will give ARIMA $(p,d,q)x(P,D,Q)_S$ The regular difference is written as

$$(1-B)^d x_t \tag{4}$$

B=backward operator, d=the non-seasonal order of differences.
The seasonal difference of order D with period S is written as

$$(1-B^S)^D x_t \tag{5}$$

In general, the differencing operation may be done several times but in practice only one or two differencing operation are used [8].

Box and Jenkins, generalized the above model and obtained the multiplicative Autoregressive Integrated Moving Average where the general form is {seasonal ARIMA (p,d,q) x $(P,D,Q)_s$ } which is written as:

$$\left.\begin{aligned}&(1-\Phi_1 B^S - \Phi_2 B^{2S} - \ldots - \Phi_P B^{PS})(1-B^S)^D x_t = \\ &(1-\Theta_1 B^S - \Theta_2 B^{2S} - \ldots - \Theta_Q B^{QS})\alpha_t\end{aligned}\right\} \tag{6}$$

The residuals α_t are in turn is represented by an ARIMA (p,d,q) model

$$\left.\begin{aligned}&(1-\Phi_1 B - \Phi_2 B^2 - \ldots\ldots - \Phi_p B^p)(1-B)^d \alpha_t = \\ &(1-\Theta_1 B - \Theta_2 B^2 - \ldots\ldots - \Theta_q B^q)\varepsilon_t\end{aligned}\right\} \quad (7)$$

The general multiplicative ARIMA (p,d,q) x $(P,D,Q)_S$ model is obtained by solving Eq.7 for α_t and replacing in Eq.6 as:

$$\Phi(B^S)\Phi(B)(1-B^S)^D(1-B)^d x_t = \Theta(B^S)\Theta(B)\varepsilon_t \quad (8)$$

4 Box-Jenkins ARIMA Model, Methodology

In 1976, Box and Jenkins, give a methodology (Fig. 2) in time series analysis to find the best fit of time series to past values in order to make future forecasts.

The methodology consists of four steps:

1) Model identification. 2) Estimation of model parameters. 3) Diagnostic checking for the identified model appropriateness for modeling and 4) Application of the model (i.e. forecasting).

The most important analytical tools used with time series analysis and forecasting are the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF). They measure the statistical relationships between observations in a single data series. Using ACF gives big advantage of measuring the amount of linear dependence between observations in a time series that are separated by a lag k. The PACF plot is used to decide how many auto regressive terms are necessary to expose one or more of the time lags where high correlations appear, seasonality of the series, trend either in the mean level or in the variance of the series [7].

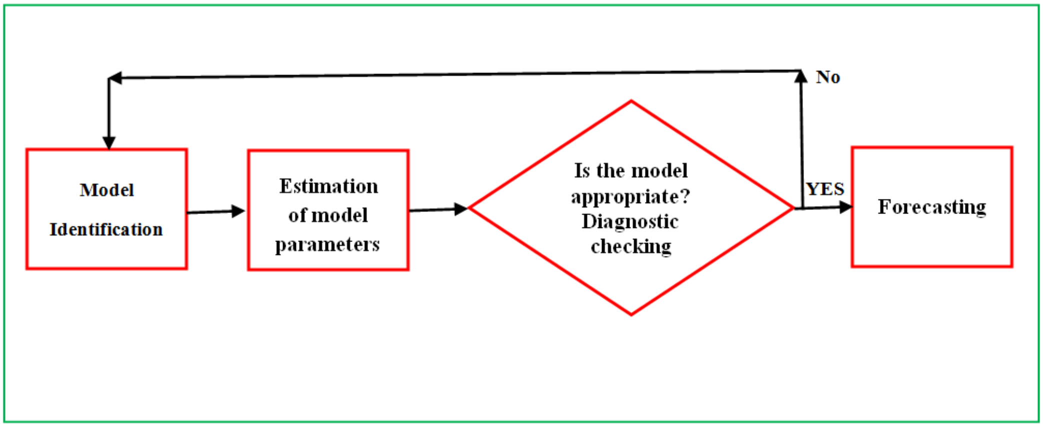

Fig. 2: Outline of Box-Jenkins methodology.

In order to identify the model (step 1), ACF and PACF have to be estimated. They are used not only to help guess the form of the model, but also to obtain approximate estimates of the parameters [26].

As identifying a tentative model is completed. The next step is to estimate the parameters in the model (step 2) using maximum likelihood estimation. Finding the parameters that maximize the probability of observations is main goal of maximum likelihood.

The next, is checking on the adequacy of the model for the series (step 3). The assumption is the residual is a white noise process and that the process is stationary and independent.

Model diagnostic checking is accomplished, in this work, through careful analysis of the residual series, the histogram of the residual, sample correlation and a diagnosis test [27].

Ljung-Box, Q-test, is used to check the assumptions of model residuals and could be written as:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{r_k^2}{n-k} \tag{9}$$

Where:

h= the maximum lag being considered, n=the number of observations in the series and r_k =the autocorrelation at lag k.

The statistic Q has a chi-square (x^2) distribution with degrees of freedom (h-m) where m is the number of parameters in the model which has been fitted to the data, the chi square value has been compared with the tabulated values; in order to evaluate the valid model otherwise the model will be rejected.

For successful models, it should be noted that a model with the less number of variables gives the best forecasting results, i.e. for a time series having more than one successful ARIMA model, in this case it should be consider the model with less variables (number of AR and/or MA), this is achieved by using Akaike's Information Criterion (AIC) [28], in order to select the best ARIMA model among successful models. The smallest value of AIC should be chosen.

Akaike's Information Criterion (AIC) may be written as:

$$AIC = -2 \log_e (L) + 2(p + q + P + Q + C) \quad (10)$$

Where:

L= Maximum likelihood, p= non-seasonal Autoregressive order,

q= non-seasonal Moving average order, P= seasonal Autoregressive order,

Q= seasonal Moving average order, C= constant of the model.

5 Results and Discussion

The data of Sinjar station was chosen as a sample of calculations. The first step in the application of the methodology is to cheek whether the time series (weekly rainfall) is stationary and has seasonality.

The weekly rainfall data (Fig. 3) shows that there is a seasonal cycle of the series and it is not stationary. The plots of ACF and PACF of the original data (Fig. 4) show also that the rainfall data is not stationary, where both ACF and PACF have significant values at different lags.

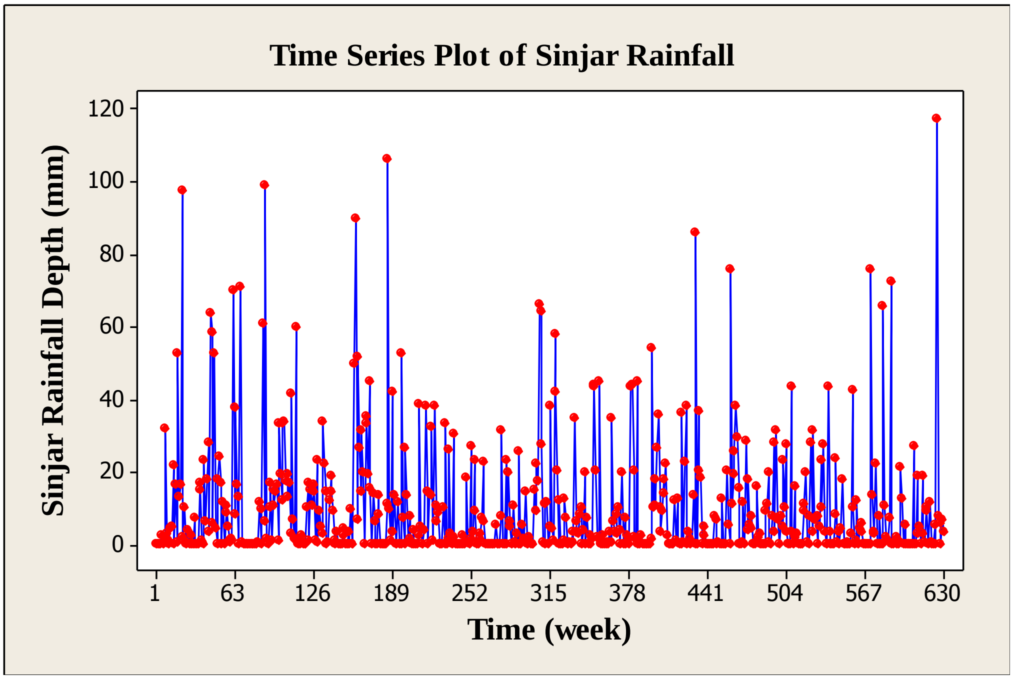

Fig. 3: Weekly rainfall data for Sinjar station for the period (1990-2011).

A stationary time series has a constant mean and has no trend over time. However it could satisfy stationary in variance by having log transformation and satisfy stationary in the mean by having differencing of the original data in order to fit an ARIMA model.

Seasonal trend could be removed by having seasonal differencing (D) through subtracting the current observation from the previous thirtieth observation, as described before that rainfall in Sinjar district, it nearly extend for 30 weeks annually. In general the seasonality in a time series is a regular pattern of changes that repeats over time periods (S).

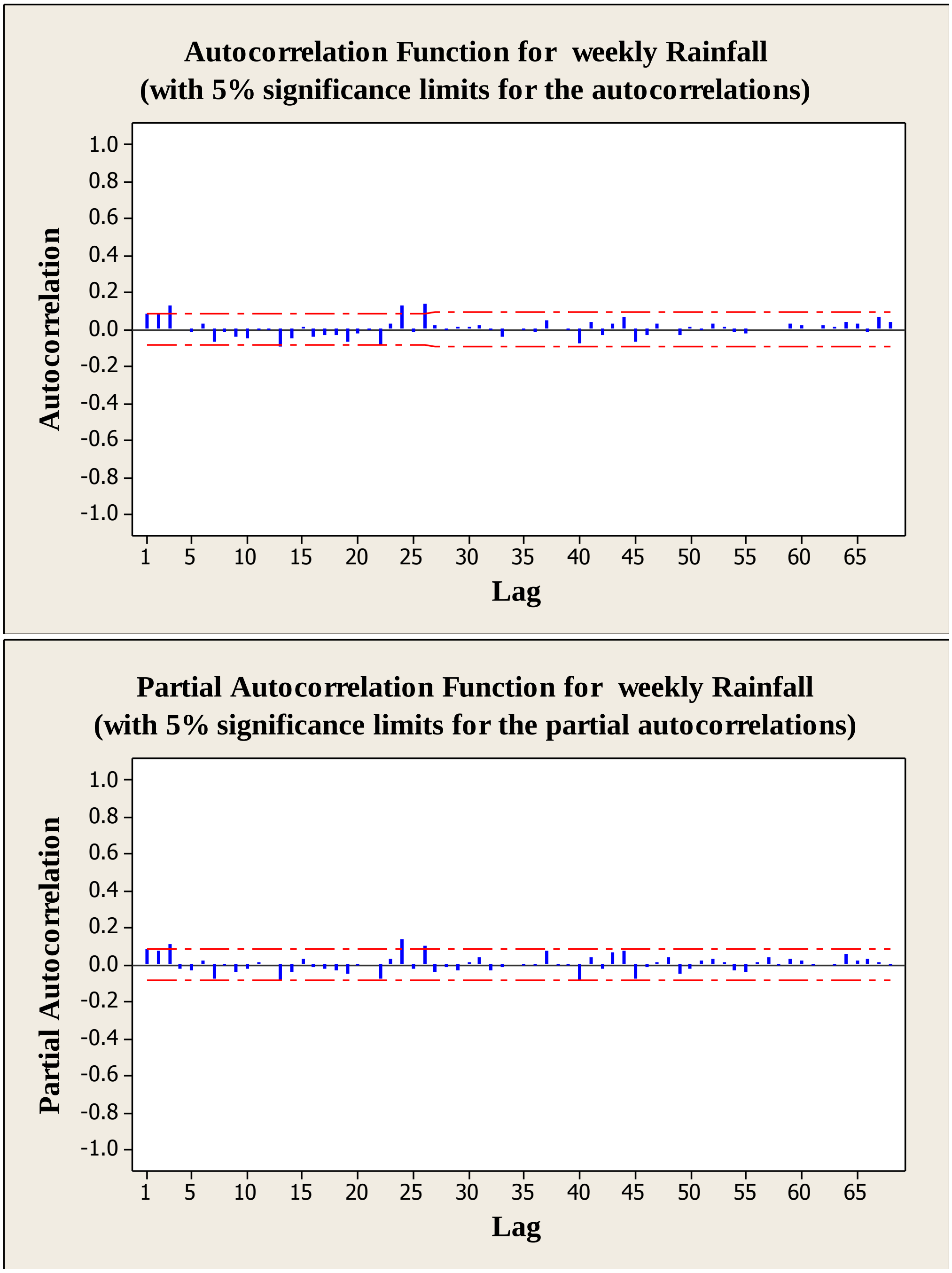

Fig. 4: ACF and PACF for original Sinjar weekly rainfall data.

However, if differenced transformation is applied only once to a series, that means data has been "first differenced" (D=1). This process essentially eliminates the trend for a time series growing at a fairly constant rate. If it is growing at an increasing rate, the same procedure (difference the data) can be applied again, then the data would be "second differenced" (D=2). If a trend is present in the data, then non-seasonal (regular) differencing (d) is required.

The weekly rainfall data, of Sinjar station, required to have a first seasonal difference of the original data in order to have stationary series. Then, the ACF and PACF for the differenced series should be tested to check stationary.

From all of the above, an ARIMA model of $(p, 0, q) \times (P, 1, Q)_{\mathbf{30}}$ could be identified.

In the Box-Jenkins methodology, the estimated model will be depending on the ACF and PACF (Fig. 5). After ARIMA model was identified, the *p, q, P* and Q parameters need to be identified for Sinjar weekly rainfall time series.

The data were tested to check the constraction of the ARIMA model by selecting the required order of the D that making the series stationary, as well as specifying the necessary order of the p,P,q and Q to adequately represent the time series model.

It should be noted that, even if the ARIMA model has been correctly identified and gives good results, this will not mean that it is the only model that can be considered where most documentations of time series dealing with ARIMA models [7, 8 and 26] indicated that other ARIMA models with values of AR and/or MA less than same parameters of the considered ARIMA (for the seasonal or nonseasonal varaibels) might be available. In this case these models should be identified and tested. Then apply AIC to select the best ARIMA model.

After selecting the most appropriate model (step 1), it was found that ARIMA model $(3,0,2)x(2,1,1)_{\mathbf{30}}$ is among several models that passed all statistic tests required in the Box-Jenkins methodology.

The model parameters are estimated (step 2) using Minitab software.

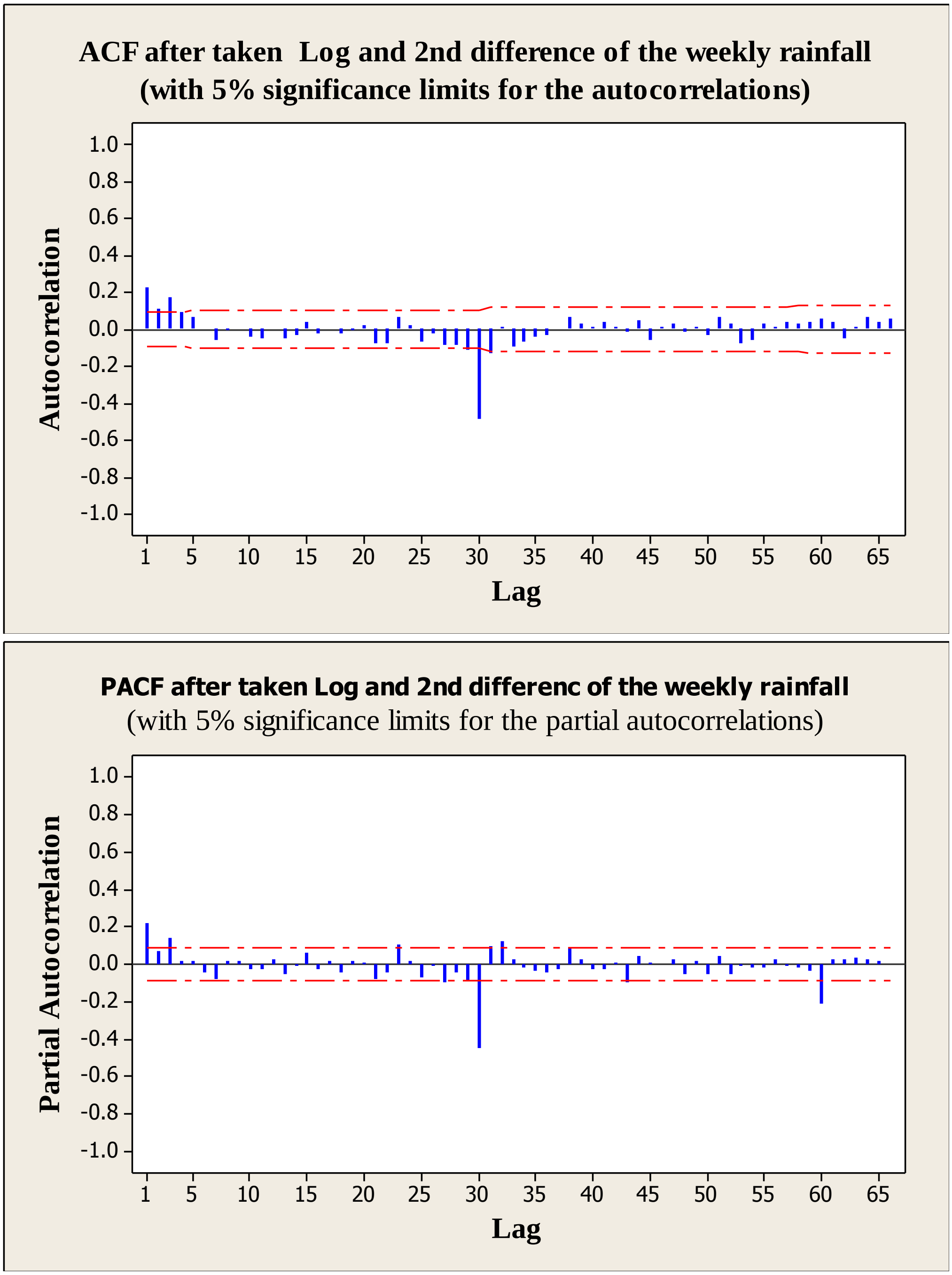

Fig. 5: ACF and PACF after taking difference and Log of Sinjar rainfall data.

Table 1 shows the estimated Parameters for the successful Sinjar ARIMA model $(3,0,2)x(2,1,1)_{30}$.

Table 1: ARIMA Coefficients for Sinjar model $(3, 0, 2)x(2, 1, 1)_{30}$.

No.	Type	Coefficients	Probability
1	AR 1	0.9672	0.001
2	AR 2	-0.6016	0.009
3	AR 3	0.1638	0.014
4	SAR 30	-0.7191	0.000
5	SAR 60	-0.4307	0.000
6	MA 1	0.8013	0.004
7	MA 2	-0.5505	0.004
8	SMA 30	0.8676	0.000

Not: AR and MA represent non-seasonal coefficients and SAR and SMA represent seasonal coefficients.

The residuals from the fitted model are examined for the adequacy (step 3). This is done by testing the residual ACF and PACF plots that shows all the autocorrelation and partial autocorrelations of the residuals at different lags are within the 95 % confidence limits.

Table 2 shows the Ljung-Box Q-test of the residuals for the successful Sinjar ARIMA model $(3,0,2)x(2,1,1)_{30}$ .This model appears to fit the Sinjar data.

Only the model with no significant residuals should be considered to indicate that the model is adequate to represent the considered time series. Figure 6 shows the residual ACF and PACF plots.

Figure 7 represents four graphical measures for the adequacy of the model.

Table 2: Ljung-Box Q test of the residuals for Sinjar ARIMA model $(3, 0, 2)x(2, 1, 1)_{30}$.

No.	**Details**	**Values**			
1.	Lag	12	24	36	48
2	Chi-Square	6.1	22.9	34.5	48.5
3	Degree of Freedom	4.0	16.0	28.0	40.0
4	Probability –Value	0.189	0.116	0.185	0.168

The first measure is the normal probability plot of the residuals (top-left of Fig.7) which is good as required for an adequate model and most of the residuals are on the straight line.

The second measure for adequacy of model is the histogram of the residuals (bottom-left of Fig.7) which shows good normality of the residuals.

The third measure is the plot of residuals against fitted values (top-right of fig.7). The variance of the error terms must be constant, and they must have a mean of zero. If this is not the case, the model may not be valid.

To check these assumptions, the plot (Fig. 7 top-right) of the residuals versus fitted values of the Sinjar ARIMA model $(302)x(211)_{30}$**,** showed that the errors have constant variance, with the residuals scattered randomly around zero. If, for example, the residuals decrease or increase with the fitted values in a pattern , then that means the errors may not have constant variance. The points on the plot (same figure) appear to be randomly scattered around zero, indicating that the error terms have a mean of zero which is reasonable. The vertical width of the scatter doesn't appear to increase or decrease across the fitted values, this suggests that the variance in the error terms is constant.

The fourth measure is the plot of residuals against fitted order of the data (bottom-right of fig.7).

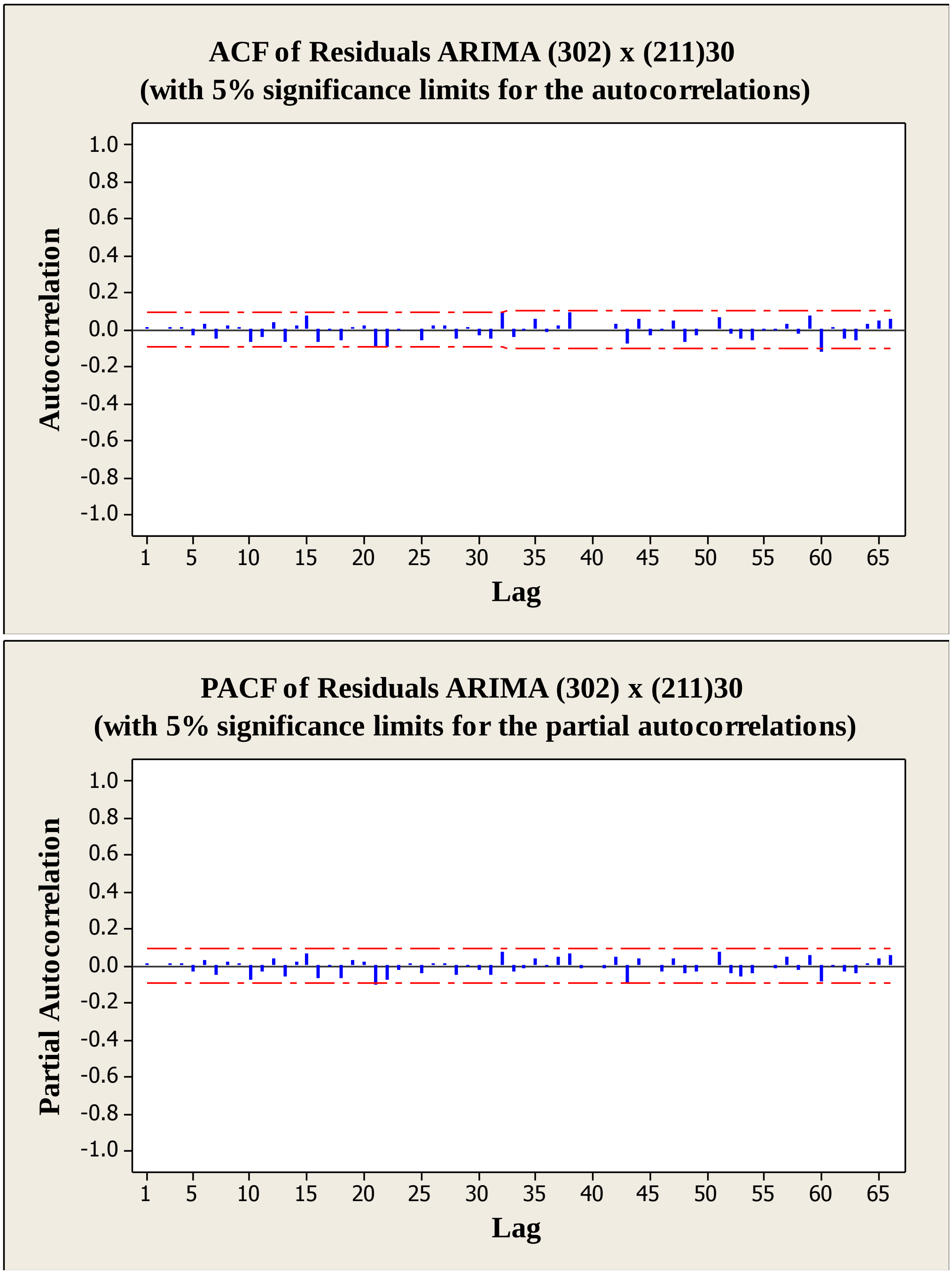

Fig. 6: ACF and PACF of residuals of Sinjar ARIMA model $(302)x(211)_{30}$.

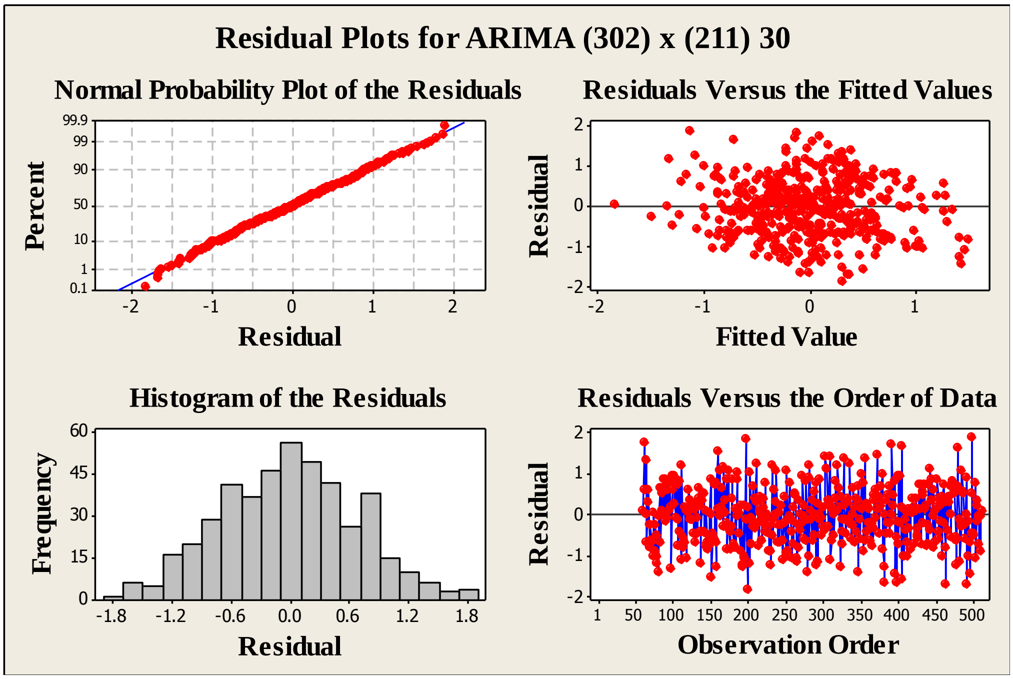

Fig .7: Residual plots of Sinjar ARIMA model $(3,0,2)x(2,1,1)_{30}$.

In this plot the data does not follow any symmetric pattern with the run order value. It shows almost randon behaviour of residuals with the increasing run order which indicates that the model is a good fit.

Almost all of the residuals are within acceptable limits which indicate the adequacy of the recommended model.

Although some other ARIMA models have been applied on the Sinjar data such as $(0,0,1)x(3,1,1)_{\mathbf{30}}$ ' $(1,0,0)x(3,1,1)_{\mathbf{30}}$ and $(0,0,3)x(1,1,1)_{\mathbf{30}}$ that pass the probability and Ljung-Box Q-tests but they still contain little residuals which might not be significant and occur at late or delayed lag (60). These were not considered as the successful models. On the basis of the above, the selected ARIMA $(3,0,2)x(2,1,1)_{\mathbf{30}}$ model is adequate to represent the Sinjar data and could be used to forecast the future rainfall data. After finding a valid model, weekly rainfall depth for the Sinjar station is forecast (step 4).

The performance of the Sinjar ARIMA model $(3,0,2)x(2,1,1)_{30}$ is evaluated by forecasting the data for the year 2011. Both the forecasted and real weekly rainfall depth of the Sinjar station for the year 2011 were fitted on the same plot to indicate the model adequacy, performance and comparison purposes (Fig. 8).

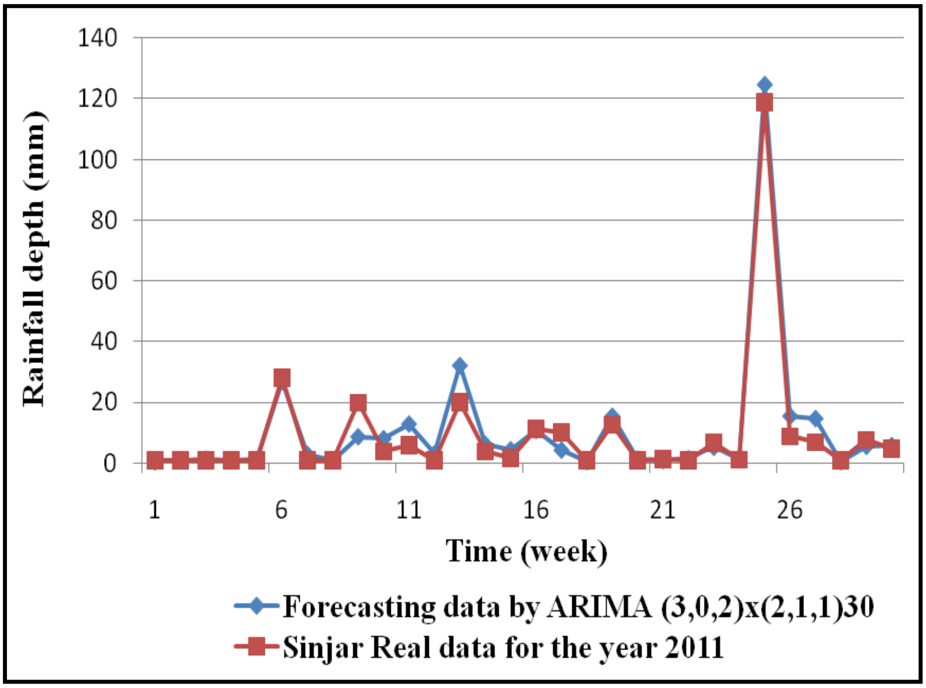

Fig. 8: Real and forecasting Sinjar rainfall depth for the year 2011.

The similarity and matching between the forecasted and real rainfall depth were good. The above comparison increases confidence with the ARIMA $(3,0,2)x(2,1,1)_{30}$ to represent the rainfall data at Sinjar station and can be use for forecasting the future rainfall data. Fig. 9 shows the forecasting rainfall depth for the years 2012-2016 using ARIMA $(3,0,2)x(2,1,1)_{30}$.

The same procedures of the Box-Jenkins methodology were followed for the other three stations (Mosul, Rabeaa and Talafar) to forecast future rainfall.

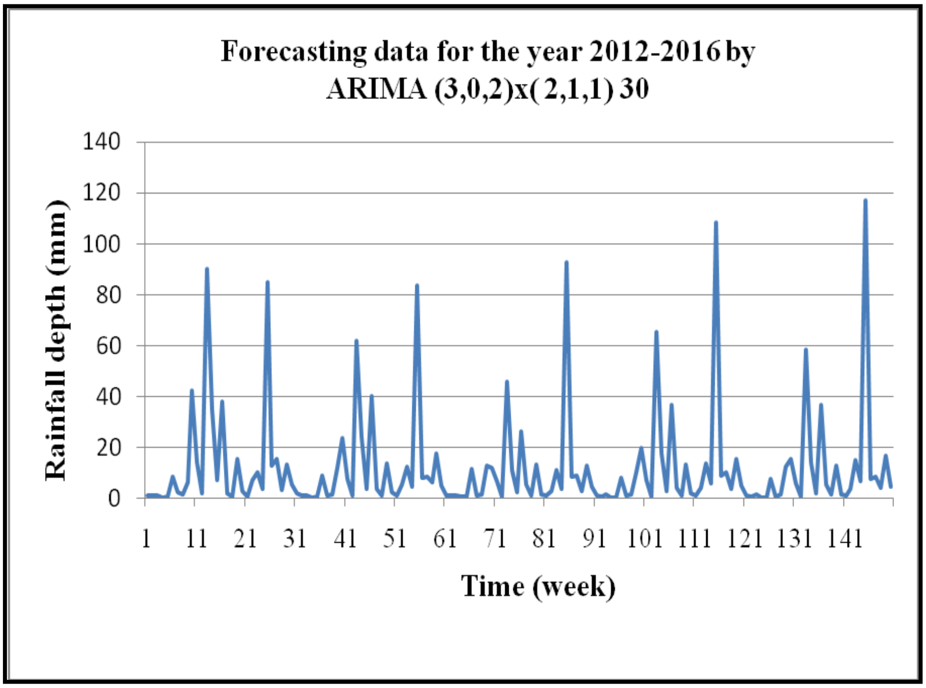

Fig. 9: Sinjar forecast rainfall depth for the years 2012-2016 using ARIMA $(3,0,2)x(2,1,1)_{30}$.

For all three stations, it was found that some ARIMA models passed all the statistical tests required in the Box-Jenkins methodology without significant residuals for ACF and PACF plots. These models are the following.

For Mosul data the ARIMA model $(1,0,1)x(1,1,3)_{30}$, for Rabeaa data the ARIMA model $(1,1,1)x(3,0,1)_{30}$ and for Talafar data the ARIMA model $(1,1,1)x(0,0,1)_{30}$. Table 3 shows the ARIMA coefficients for the above models.

Table 4 shows the results of the Ljung-Box Q-test of the residuals for the ARIMA models according to the stations.

The selected ARIMA models for the three stations (Mosul, Rabeaa and Talafar) have the smallest values of AIC (Table 5) among other successful ARIMA models. Thus, these models can represent rainfall data and can be used with confidence to forecast future rainfall data.

Table 3: ARIMA Coefficients for the selected models according to the stations.

No.	Type	Coefficients	Probability
Mosul station ARIMA model (101) (113)$_{30}$			
1	AR 1	0.6085	0.000
2	SAR 30	-0.9061	0.000
3	MA 1	0.4513	0.012
4	SMA 30	-0.6892	0.000
5	SMA 60	0.6137	0.000
6	SMA 90	-0.4534	0.000
Rabeaa station ARIMA model (112) (301)$_{30}$			
1	AR 1	-0.5882	0.000
2	SAR 30	-0.4542	0.000
3	SAR 60	-0.4280	0.000
4	SAR 90	-0.3972	0.000
5	MA 1	0.3057	0.020
6	MA 2	0.6353	0.000
7	SMA 30	0.5569	0.000
Talafar station ARIMA model (111) (001)$_{30}$			
1	AR 1	0.1305	0.004
2	MA 1	0.9540	0.000
3	SMA 30	0.9007	0.000

Table 4: Ljung-Box Q test of the residuals for the ARIMA models according to the stations

No.	**Details**	**Location and Model**			
		Mosul station ARIMA model (1,0,1) x (1,1,3)30			
1	Lag	12	24	36	48
2	Chi-Square	9.9	26.8	36.2	44.4
3	DF	6	18	30	42
4	P-Value	0.129	0.083	0.201	0.370
		Rabeaa station ARIMA model (1,1,2) x $(3,0,1)_{30}$			
1	Lag	12	24	36	48
2	Chi-Square	10.6	27.4	36.9	54.4
3	DF	5	17	29	41
4	P-Value	0.061	0.052	0.148	0.078
		Talafar station ARIMA model (1,1,1) x $(0,0,1)_{30}$			
1	Lag	12	24	36	48
2	Chi-Square	9.6	26.1	34.3	50.4
3	DF	9	21	33	45
4	P-Value	0.386	0.203	0.406	0.268

Table 5: Akaike's Information Criterion (AIC) values for the successful ARIMA models according to the stations.

No.	**Station**	**Model**	**AIC Value**
1	**Sinjar**	**(3,0,2)x $(2,1,1)_{30}$**	17.44
2	**Mosul**	**(1,0,1)x$(1,1,3)_{30}$**	13.03
3	**Rabeaa**	**(1,1,2)x$(3,0,1)_{30}$**	16.31
4	**Talafar**	**(1,1,1)x$(0,0,1)_{30}$**	8.3

Figures 10,13 and 16 show the four graphical measures for the adequacy of the selected ARIMA models for the three stations (Mosul, Rabeaa and Talafar) respectivly. The normal probability plots of the residuals show that most of the residuals are on the straight line having a good histograms shape. Follow-up the plots of residuals vs. fitted values and order of the data respectively, show that the errors have constant variance, the points on the plot appear to be randomly scattered around zero.

Almost all of the residuals are within acceptable limits which indicate the adequacy of the recommended models.

The performance of ARIMA models for the above three stations are evaluated by forecasting the data for the year 2011 to indicate the models adequacy, performance and comparison purposes (figures 11, 14 and 17).

Figures 12, 15 and 18 show forecast rainfall depth for the years 2012-2016 for three stations data Mosul, Rabeaa and Talafar, respectively.

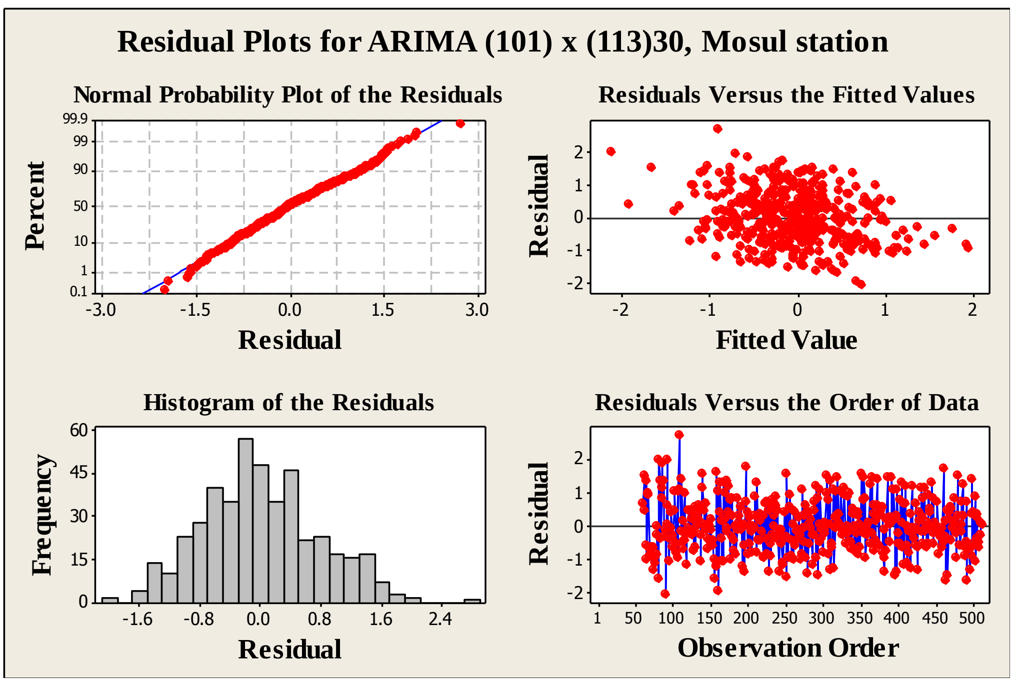

Fig. 10: Residual plots of Mosul ARIMA model (1,0,1)x(1,1,3)30.

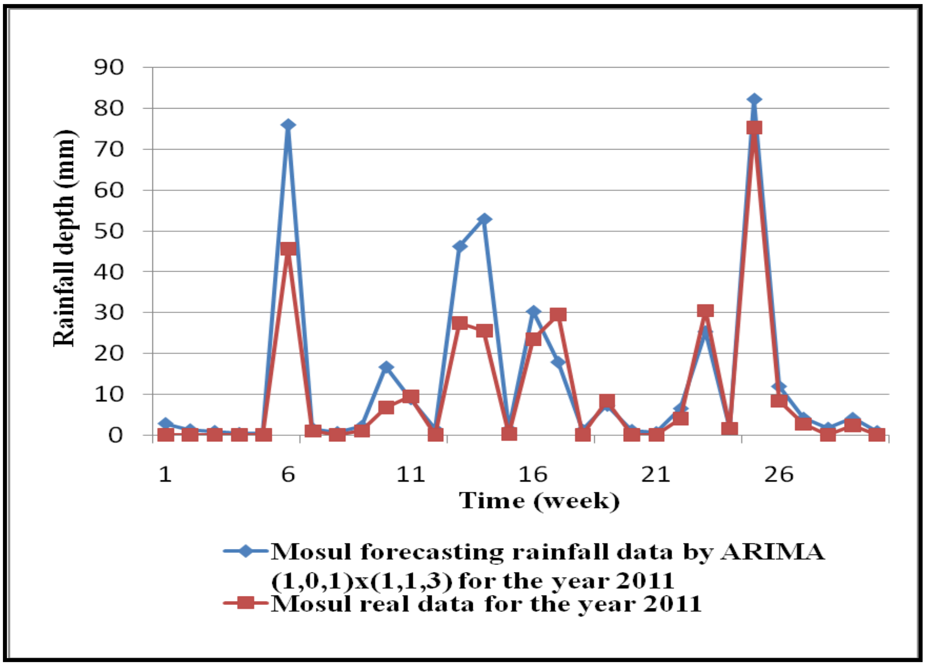

Fig. 11: Real and forecasting Mosul rainfall depth for the year 2011.

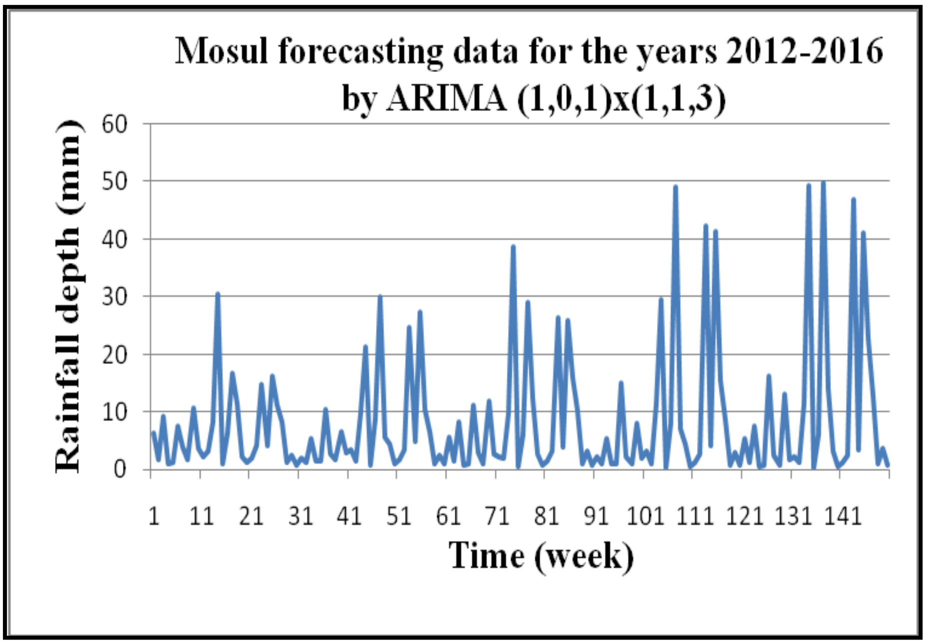

Fig. 12: Mosul forecast rainfall depth for the years 2012-2016 using RIMA(1,0,1)x(1,1,3)$_{30}$.

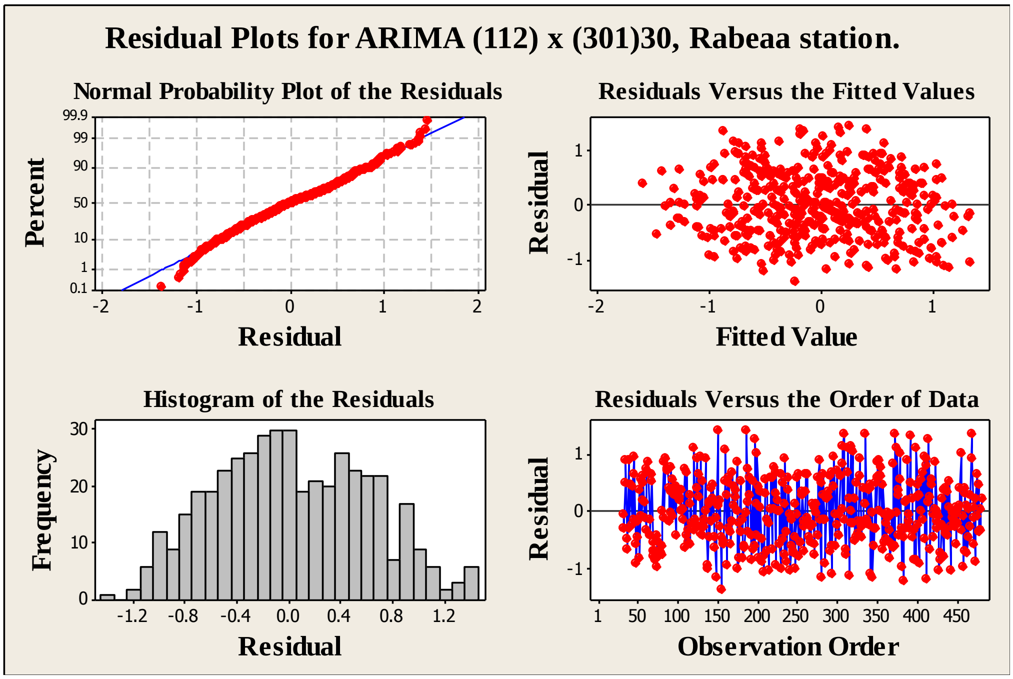

Fig. 13: Residual plots of Rabeaa ARIMA model $(1,1,2)x(3,0,1)_{30}$.

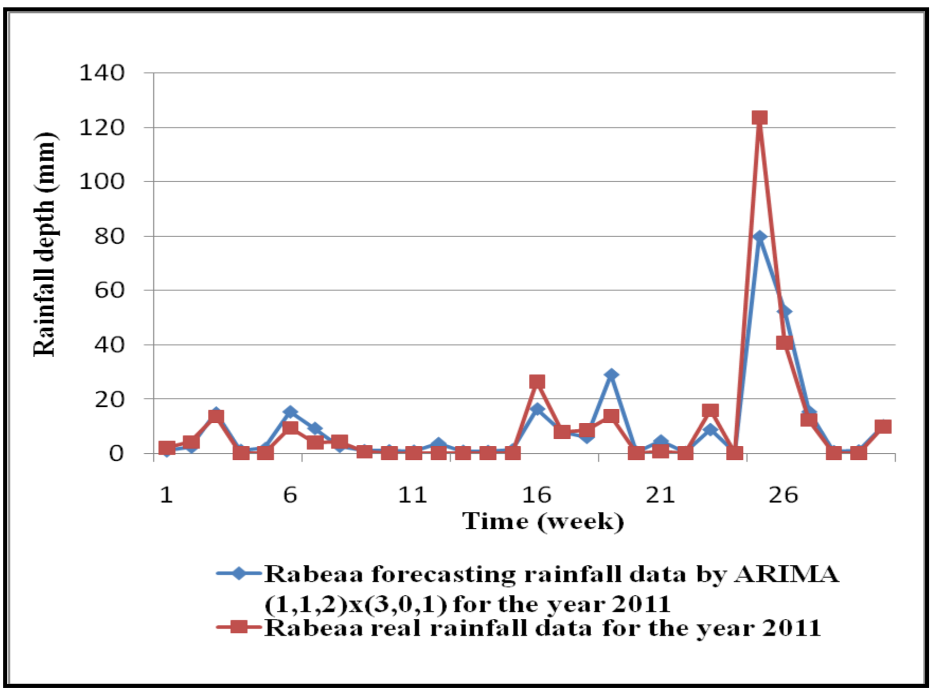

Fig. 14: Real and forecasting Rabeaa rainfall depth for the year 2011.

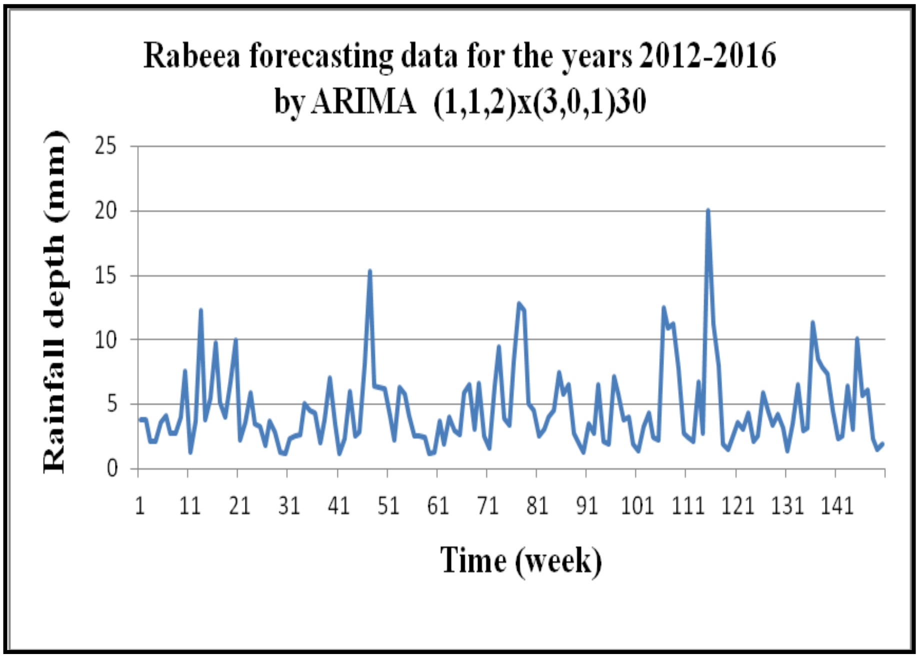

Fig. 15: Rabeaa forecast rainfall depth for the years 2012-2016 using RIMA$(1,1,2)x(3,0,1)_{30}$.

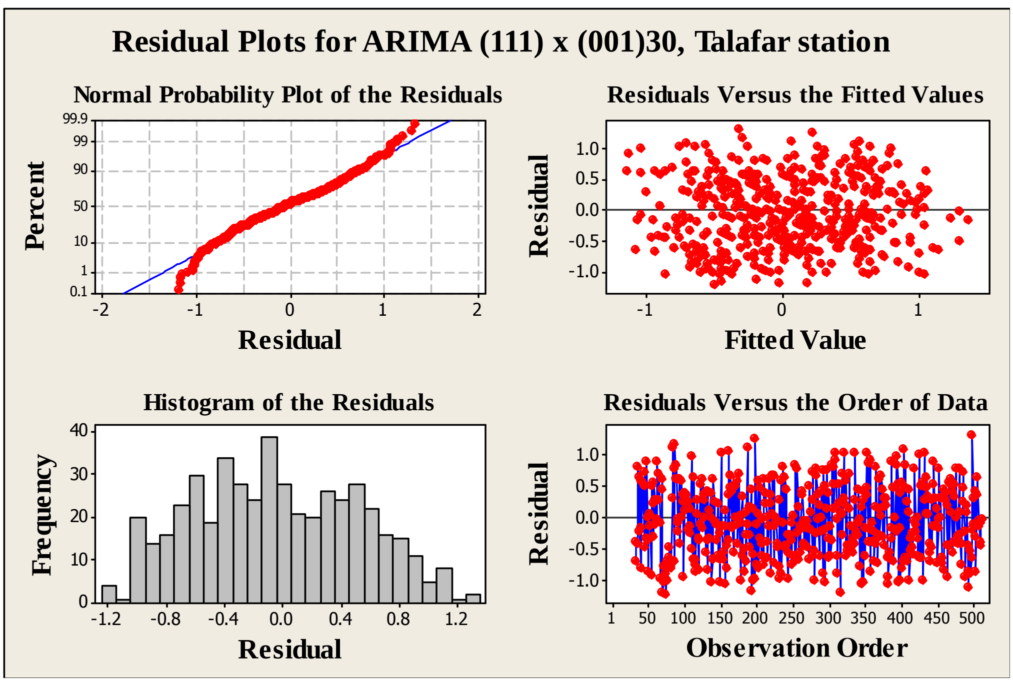

Fig. 16: The residual plots of Talafar ARIMA model $(1,1,1)x(0,0,1)_{30}$.

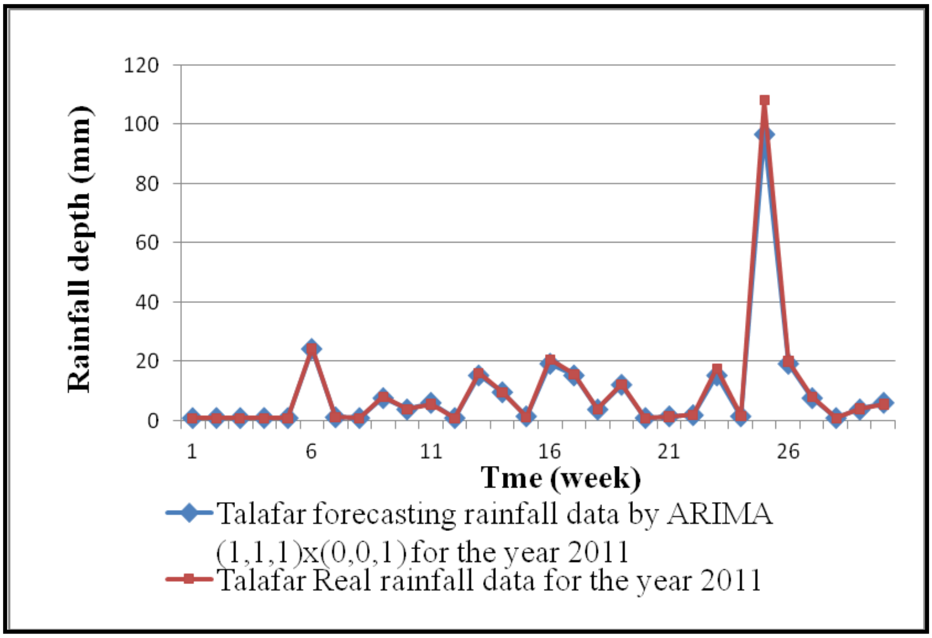

Fig. 17: Real and forecasting Talafar rainfall depth for the year 2011.

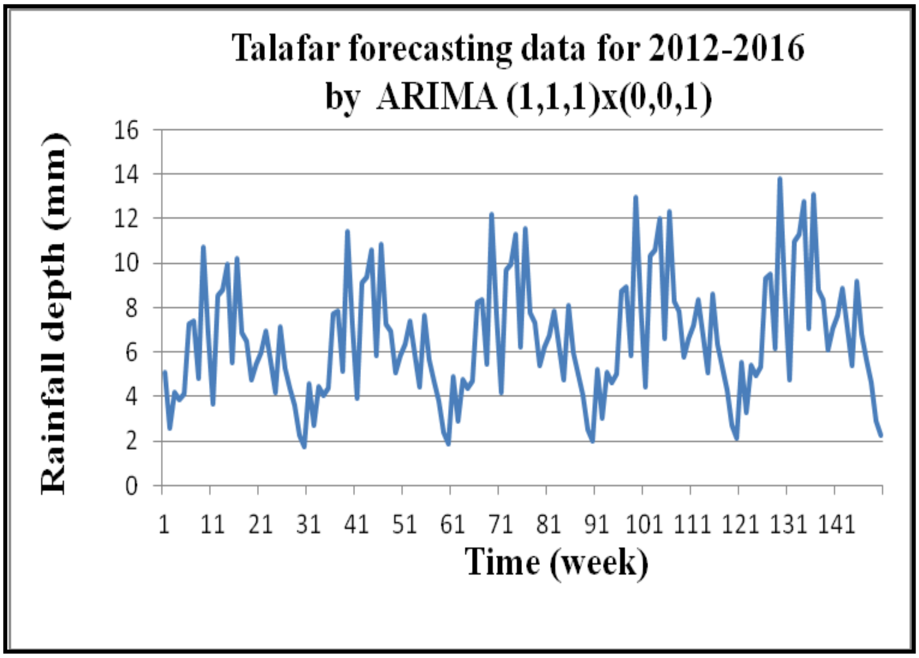

Fig. 18: Talafar forecast rainfall depth for the years 2012-2016 using RIMA$(1,1,1)x(0,0,1)_{30}$.

6 Conclusions

The weekly rainfall record in the Sinjar semi-arid region has been studied using the Box-Jenkins (ARIMA) model methodology. A weekly rainfall record spanning the period of 1990-2011 for four stations (Sinjar, Mosul, Rabeaa and Talafar) at Sinjar district of North Western Iraq has been used to develop and test the models.

The performance of the resulting successful ARIMA models was evaluated by using the data for the year 2011 through graphical comparison between the forecast and actually recorded data. The forecasted rainfall data showed very good agreement with the actual recorded data. This gave an increasing confidence of the selected ARIMA models.

The study reveals that the Box-Jenkins (ARIMA) model methodology could be used as an appropriate tool to forecast the weekly rainfall in semi-arid region like North West of Iraq for the up-coming 5 years (2012-2016).

The results achieved for rainfall forecasting will help to estimate hydraulic events such as runoff, then water harvesting techniques can be used in planning the agricultural activities in that region. Predicted excess rain can be stored in reservoirs and used in a later stage.

The modeling techniques demonstrated in this contribution can help farmers in the area to enlarge the areas of land to be cultivated using supplemental irrigation.

7 Acknowledgements

The authors would like to thank Dr. Erik Vanhatalo, Luleå University of Technology, Sweden, for his support and valuable discussion and advice, and the Iraqi Meteorological and Seismology Organization, Baghdad, Iraq, and especially Dr. Dauod Shaker for providing the meteorological data. Thanks to Professor John McManus for reading the manuscripts and offering fruitful suggestions.

References

[1] Al-Ansari, N.A. and Knutsson, S., Toward Prudent management of Water Resources in Iraq, J. Advanced Science and Engineering Research, Vol.1, (2011), 53-67.

[2] Al-Ansari, N.A., Ezz-Aldeen, M., Knutsson, S. and Zakaria, S., Water harvesting and reservoir optimization in selected areas of south Sinjar Mountain, Iraq, Accepted for publication in ASCE J. Hydrologic Engineering, (2012).

[3] Zakaria, S., Al-Ansari, N., Knutsson, S., Ezz-Aldeen, M., Rain Water Harvesting And Supplemental Irrigation At Northern Sinjar Mountain, Iraq. Journal of Purity, Utility Reaction and Environment **Vol.1** No.3, May, (2012-a), 121-14.

[4] Zakaria, S., Al-Ansari, N., Knutsson, S., Ezz-Aldeen, M., Rain Water Harvesting at Eastern Sinjar Mountain, Iraq. Journal of Geoscience Research, **Vol. 3**, Issue 2, (2012-b), 100-108.

[5] Chatfild, C., The Analysis of Time Series-an introduction, 5th Edn. Chapman and Hall, UK, (1996).

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

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

[8] Salas, J.D., Delleur, J.W., Yevjevich, V., and Lane, W.L., Applied Modeling of Hydrologic Time series. ISBN13:978-0-918334-374, Water Resources Publication, Michigan, USA. (1980).

[9] Vandaele, W. Applied time series and Box-Jenkins Models. Academic Press Inc., Orlando, Florida, ISBN: 10: 0127126503, (1983), 417.

[10] Chiew, F.H.S., Stewardson, M.J., and McMahon, T.A., Comparison of six rainfall-runoff modeling approaches. J. Hydrol., **147**, (1993), 1-36.

[11] Kuo, J.T. and Sun, Y.H., An intervention model for average 10 day stream flow forecast and synthesis. J. Hydrol., **151**, (1993), 35-56.

[12] Langu, E.M., Detection of changes in rainfall and runoff patterns. J. Hydrol., **147**(1993), 153-167, http://cat.inist.fr/?aModele=afficheN&cpsidt=4788801. Lincoln.http://researcharchive.lincoln.ac.nz/dspace/bitstream/10182/4199/3/ edwards_mapplsc.

[13] Al-Ansari, N.A., Salameh, E. and Al-Omari, I., Analysis of Rainfall in the Badia Region, Jordan, Al al-Bayt University Research paper **No.1**, (1999), 66.

[14] Al-Ansari, N.A. Shamali, B. and Shatnawi, A., Statistical analysis at three major meteorological stations in Jordan, Al Manara Journal for scientific studies and research 12, (2006), 93-120.

[15] Al-Ansari, N. and Baban S. M. J., Rainfall Trends in the Badia Region of Jordan. Surveying and Land Information Science; Dec; **65**, 4; ProQuest Science Journals (2005), 233.

[16] Al-Hajji, A.H. , Study of some techniques for processing rainfall time series in the north part of Iraq. Unpublished MSc. Thesis, Mosul University, (2004).

[17] Al-Dabbagh, M.A., Modeling the rainfall-discharge relation and discharge forecasting for some rivers in Iraq. Unpublished MSc. Thesis, Mosul University, (2005).

[18] Weesakul, U. and Lowanichchai, S. Rainfall Forecast for Agricultural Water Allocation Planning in Thailand. Thammasat Int. J. Sc. Tech. **Vol. 10**, No. 3, (2005).

[19] Somvanshi, V.K., Pandey, O.P., P.K.Agrawal, N.V.Kalanker1, M.Ravi Prakash and Ramesh Chand, Modeling and prediction of rainfall using artificial neural network and ARIMA techniques. J. Ind. Geophys. Union

(April) **Vol.10**, No.2, (2006), 141-151.

[20] Edwards, S., Regionally Dissected Temperature and Rainfall Models for the South Island of New Zealand. Master thesis of Applied Science in Environmental Management, pdf Lincoln University Digital Dissertation, (2011).

[21] Mahsin, Md. Akhter,Y. Begum ,M., 2012. Modeling Rainfall in Dhaka Division of Bangladesh Using Time Series Analysis. Journal of Mathematical Modelling and Application, **Vol. 1**, No.5, (2012), 67-73.

[22] Momani, M. Naill, P.E., Time Series Analysis Model for Rainfall Data in Jordan: Case Study for Using Time Series Analysis. American Journal of Environmental Sciences **5** (5), (2009), 599-604.

[23] Shamsnia, S.A., Shahidi, N., Liaghat, A., Sarraf, A. Vahdat, S.F., Modeling of Weather Parameters Using Stochastic Methods (ARIMA model) (Case Study: Abadeh Region, Iran), International Conference on Environment and Industrial Innovation, IPCBEE **vol.12**, IACSIT Press, Singapore, (2011).

[24] Valipour, M., Critical Areas of Iran for Agriculture Water Management According to the Annual Rainfall, European Journal of Scientific Research **Vol.84** No.4 (2012), 600-608

[25] Lam N S., Spatial Interpolation Methods review. The American Cartographer., **10** (1983), 129-149.

[26] Box, G.E.P. and Jenkins, G.M., Reinsel, G.C., Time Series Analysis Forecasting and Control. Prentice-Hall Inc., New Jersey, USA, ISBN 0-13-060774-6, (1976).

[27] Ljung, G.M. and Box, G.E.P., On a measure of lag of fit in time series model, Biometrika (1978), **65**,2, 297-303.

[28] Akaike, H., A new look at the statistical model identification. Transactions on Automatic Control, **Vol.19**, No.16, (1974), 723.