https://doi.org/10.26682/csjuod.2023.26.2.69
Journal of University of Duhok, Vol. 26, No.2 (Pure and Engineering Sciences), Pp 784-794, 2023
4th International Conference on Recent Innovations in Engineering (ICRIE 2023) (Special issue)

GENERATING THE RAINFALL TIME SERIES USING ARIMA MODEL IN KURDISTAN REGION, IRAQ

GAHEEN SARMA and EVAN HAJANI
College of Engineering, University of Duhok, Kurdistan Region, Iraq

(Accepted for Publication: November 27, 2023)

ABSTRACT

Generating time series data are an important tool in operations research, as this data is often the basis of model decision-makers. In this study, the annual maximum rainfall (AMR) data from three rainfall stations (i.e., Duhok, Erbil, and Sulaymaniya) located in the Kurdistan Region of Iraq have been used to build auto-regressive integrating moving average (ARIMA) models. For this reason, the rainfall data series from the years 1991 to 2021 was used. The Box-Cox transformation was used to make the rainfall time series stationary and normal. Several statistical tests were used to evaluate how well the successful ARIMA models performed. Results revealed that the most suitable model for the Duhok station was ARMA (0, 3), and for both Erbil and Sulaymaniya stations, it was the ARMA (0, 4) model. The AMR data for the following five years was predicted using these models (2022 to 2026). The study found that in a semi-arid region like the Kurdistan Region of Iraq, the ARIMA models were a useful tool for generating future rainfall.

KEYWORDS: *Time series; ARIMA; Rainfall; Forecast; Modelling Rainfall data*

1. INTRODUCTION

Estimating rainfall is crucial for climate impact studies and evaluations of hydrological processes [23]. The occurrence of rainfall is a complicated phenomenon that makes forecasting it difficult [5, 29]. Projects for managing water resources, agriculture, and flood control are all based on accurate rainfall forecasts [13]. Accurate rainfall prediction helps decision-makers to develop appropriate mitigation plans [27]. Various models, methods, and techniques are available in the literature for evaluating, simulating and predicting hydrological variables. Each prediction model differs in terms of accuracy, range, duration, and cost [22]. For example, [7] forecasted the monthly rainfall data on three reservoir dams in northern Iraq (i.e., Mosul, Dokan, and Derbendikhan) using Autoregressive Integrated Moving Average (ARIMA) models. Based on the results, it was clear that the ARIMA model could accurately predict the monthly rainfall series in the three reservoirs. In order to predict the amount of rainfall that will fall in India during the months of June to August, [10] developed an ARIMA model using the rainfall data for the years 1871 to 1999. As a result of the findings, it can be concluded that the ARIMA (0, 1, 1) is a suitable model for predicting monthly rainfall data.

In another study, [12] presented Time Series Modeler (TSM) for forecasting rainfall in an Indian coastal region. The main characteristics include a five-year dataset (2009–2013) that includes rainfall, maximum and minimum temperature, dew point, visibility, and wind speed. A reliable model for predicting rainfall is therefore feasible because the performance criteria for this model's evaluation are based on MAD, MSE, MAPE, and RMSE. With a prediction accuracy range of 80%, the results produced by this model are widely accepted. [21] used generalized linear models (GLMs) to predict daily rainfall from 1955 to 2010 on two Greek islands (Milos and Naxos). Depending on the accuracy of the prediction of rainfall data, the results revealed that the GLMs are suitable as inputs in future hydrological applications. [3] used an ARIMA model to predict the changes in precipitation for the projected years. The best-fit ARIMA models were identified using the Box-Jenkins method. The daily precipitation data from five stations were used to predict the amount of precipitation for the Jordanian Wadi Shuaib catchment. The most accurate ARIMA models, which had been tested with 10 years of data (2007–2016), were used to forecast precipitation up to 2026. It has been found that

gaheen.mohammed@uod.ac
evan.haiani@uod.ac

for the Jordanian Wadi Shuaib catchment area, which is already water-stressed, this observed pattern calls for efficient water management strategies. [26] improved forecasts for rainfall data in Makassar, Indonesia, using ARIMA and Kalman Filter techniques. The information was gathered from January 2010 to December 2020. Use should be made of the model with the lowest mean absolute percentage error (MAPE). The results showed that the best ARIMA model is ARIMA (0,1,1) (0,1,1), with a MAPE value of 111.48, while the MAPE value obtained by applying the Kalman Filter algorithm is 47.00, indicating that Kalman Filter has better prediction than ARIMA model.

For hydrologists and designers, Rainfall is considered an important water resource. In the last decade, the Iraqi Kurdistan Region has been facing a severe water problem, and there has been a lack of meteorological information about the characteristics and patterns of the current and upcoming rainfall time series. Therefore, rainfall time series analysis and forecasting are essential for planners, decision-makers, and flood risk management to provide information on a variety of design issues relating to the management of water resource structures. This study aims to use the ARIMA models to predict the AMR data for the five years from 2022 to 2026 in the three selected sites in the Kurdistan Region, Iraq (i.e., Duhok, Erbil, and Sulaymaniya). The current study's quantitative analysis of the future rainfall time-series data can help hydrologists diagnose the factors that influence how they analyze the current and future rainfall situation, which helps select the most appropriate hydrological estimates and hydraulic designs.

2. STUDY AREA AND DATA

In this study, three rainfall stations within three major governorates, namely, Duhok, Erbil, and Sulaymaniya in the Kurdistan Region located north of Iraq are examined (see Figure 1). The Kurdistan Region is located in the Arabian Plate's northeastern corner [1, 16]. Between October and April, there is a very seasonal amount of rainfall during the winter [2]. The annual maximum rainfall (AMR) data series for 32 years, from 1990 to 2021, has been used in the current study (see Figure 1). The maximum amount of rainfall that could be expected in a year was calculated using the rainfall data (i.e., 365 days period). It should be possible to characterize changes in rainfall using the threshold record length of 32 years, which was determined at random. The gaps in the rainfall data of a particular station were filled by the linear regression analysis method with a nearby station that had no data gap using Excel software. It involved contrasting the rainfall data gaps at one station with those at a nearby station that didn't have any [14, 15]. Table 1 displays the statistical analysis of the rainfall data series for the three stations selected for this study.

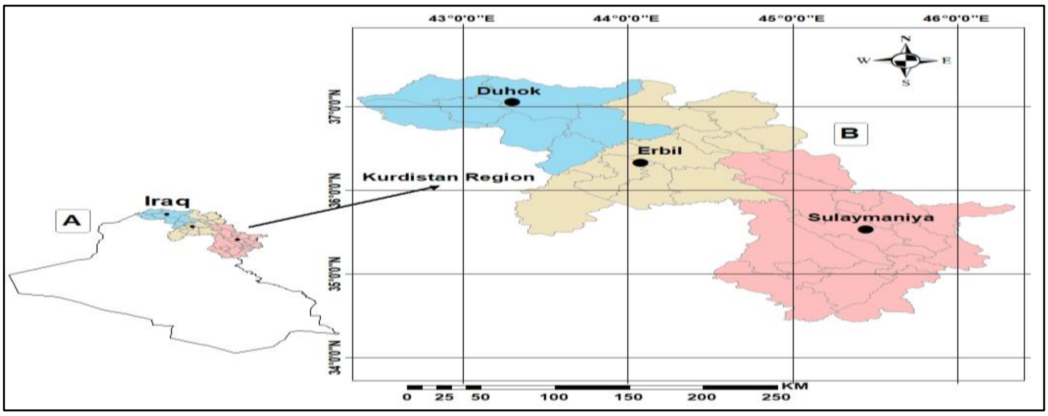


Fig.(1):- A map of the study area's location.

Table(1):- A statistical summary of the data series.

Stations	Station ID	Longitude (E)	Latitude (N)	Elevation	Max Rainfall (mm)
Dohuk	370428	43° 17' 60"	37° 02' 60"	795	150
Erbil	361441	44° 00' 00"	36°10' 00"	439	103.9
Sulaymaniya	354453	45° 27' 00"	35° 31' 48"	886	131.8

3. METHODS

In this study, the procedure for estimating the ARIMA models involves the following steps:

3.1 Stationary and Normality Identification:

By using differencing and transformation of the time series data, the first step in time series modeling with the ARIMA model is to establish stationary and normal time series. The Box-Cox method and the Differencing method must be used to convert the time series into stationary and normal series for the ARIMA model [4]. Three tests for stationary were used: the Phillips-Perron (PP), Augmented-Dickey Fuller (ADF), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests [11, 28, 18]. A unit root test can be used to determine whether differencing of the data series is necessary. Based on the outcomes of the three stationary tests mentioned above, this will be done. In addition, the Shapiro-Wilk, Anderson-Darling, and Jarque-Bera tests for normality were used [24, 8]. Table 2 displays the critical values (CV) for the stationarity and normality tests as a function of sample size (n=32) at the previously mentioned 5% significance level.

Table(2):- The CV for the tests of stationarity and normality [18]

Stationary Tests	ADF	PP	KPSS
CV	-2.960	-2.96	0.463
Normality Tests	SW	AD	JB
CV	0.934	2.492	5.991

3.2 ARIMA model

The autoregressive integrated moving average (ARIMA) time series model [9] combines moving average, integration, and autoregressive functions (MA). The ARIMA model is a class of statistical models used for analyzing and predicting time series data. The ARIMA (p, d, q) model has the following form:

$$\Delta^d y_t = C + \emptyset_1 \Delta^d y_{t-1} + \emptyset_2 \Delta^d y_{t-2} + \cdots + \emptyset_p \Delta^d y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (1)$$

A ARMA (p, q) model is appropriate for time series data that are stationary and normal.

$$y_t = \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (2)$$

Where $\Delta^d y_t$ denotes a differenced series. p denotes the AR parameters, the nonstationary time series' order of differentiation is given by d, and the MA parameter is symbolized by q. The parameters $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ denotes as AR constant coefficients, the θ_j (j = 1, 2, ..., q) are MA constant coefficients, ε_t is a distinct series of white noise.

3.3 Application of the ARIMA model

The steps taken in this study's application of ARIMA modeling and data series forecasting are detailed below:

1. Utilizing the various statistical tests mentioned in subsection 3.1 to analyze the stationarity and normality of the rainfall time series.

2. Calculating the order of the integer (d) in the non-stationary time series data and examining the time series' normality. The original time series is transformed using the Box-Cox transformation [19] method if d is not equal to zero or if a data set appears to be out of the ordinary.

3. The ARMA model for the stationary time series will be used. The ARMA model developed in this step is a special case of the ARIMA model with d = 0 if the time series is stationary. The stationarized time series' autocorrelation function (ACF) and partial autocorrelation (PACF) plots, the order of auto-regression (p) and the order of moving average (q) will be calculated [25].

4. The Akaike Information Criterion (AIC) [6], the Schwarz Criterion (SC) [17], the Standard Error (S.E.) of Regression, and the Probability (p-value) are used to determine which model is the best.

5. Once the best model has been chosen, the time series data will be predicted using this model. The procedure as a flowchart of ARIMA modeling which involves the above steps is given in Fig. 2.

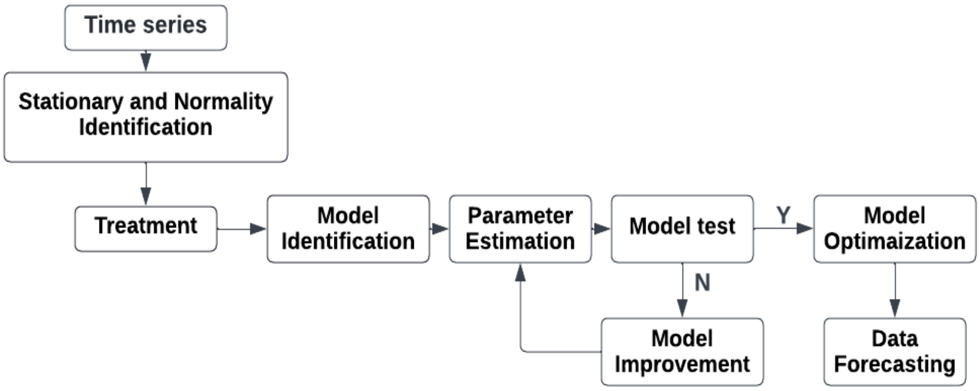


Fig. (2): - The procedure of applying the ARIMA model.

4. RESULTS AND DISCUSSION

4.1 The Stationarity and Normality Results on the Amr Data Series

The outcomes of the stationary and normality tests mentioned using EViews software in section 3.1 on the AMR data series of the three stations that were adopted are shown in Tables 3 and 4. Table 3 shows that none of the three stationary tests used in this study (i.e., the ADF, PP, and KPSS) results at a significance level of 0.05 or less (Table 2). The three stations that were adopted have stationary initial AMR data series, in other words. The data series' non-stationarity was not, however, supported by any statistical test (including the ADF, PP, and KPSS). The normality of the data series still needs to be verified using objective normality tests. Table 4 shows the results of the three normality tests that were used in this study (SW, AD, and JB), and by comparing them to the critical value of the significance level of 0.05 (Table 2), it is decided to reject the null hypothesis and come to the conclusion that the AMR data series for the three adopted stations do not follow a normal distribution. In order to confirm the non-normal AMR data to a normal distribution and enable additional capability analysis and hypothesis testing, the Box-Cox transformation method was used.

Table(3):- Results of stationarity tests of AMR data series

	ADF		PP		KPSS	
	t-statistic	P-value	t-statistic	P-value	t-statistic	P-value
Duhok	-3.987	0.004	-4.617	<0.001	0.197	0.651
Erbil	-6.572	<0.001	-6.768	<0.001	0.083	0.944
Sulaymaniya	-6.672	<0.001	-6.772	<0.001	0.222	0.729

Table (4):- Results of normality tests of AMR data series at a 5% significant level

Stations	SW		AD		JB	
	t-statistic	P-value	t-statistic	P-value	t-statistic	P-value
Duhok	0.859	<0.001	1.330	0.002	20.048	<0.001
Erbil	0.914	0.014	0.740	0.041	10.480	0.005
Sulaymaniya	0.870	<0.001	2.001	<0.001	11.512	0.003

4.2 The Box-Cox Result

Depending on the results of the three normality tests in subsection 4.1, it has been concluded that there is a need for transformation using the Box-Cox method. According to the Box-Cox method, the suitable lambda value (λ) has to be determined for the rainfall data for the Duhok, Erbil, and Sulaymaniya stations. According to the Box-Cox plot in Fig. 3, the best value for λ in a 95% confidence interval has been estimated for the three selected stations. Depending on the Box-Cox method, the Transformed function for each adopted station and depending on the values of λ are shown in Table 5. Once the transformation of the AMR data has been done, the stationarity and normality of the AMR data transformation need to be verified.

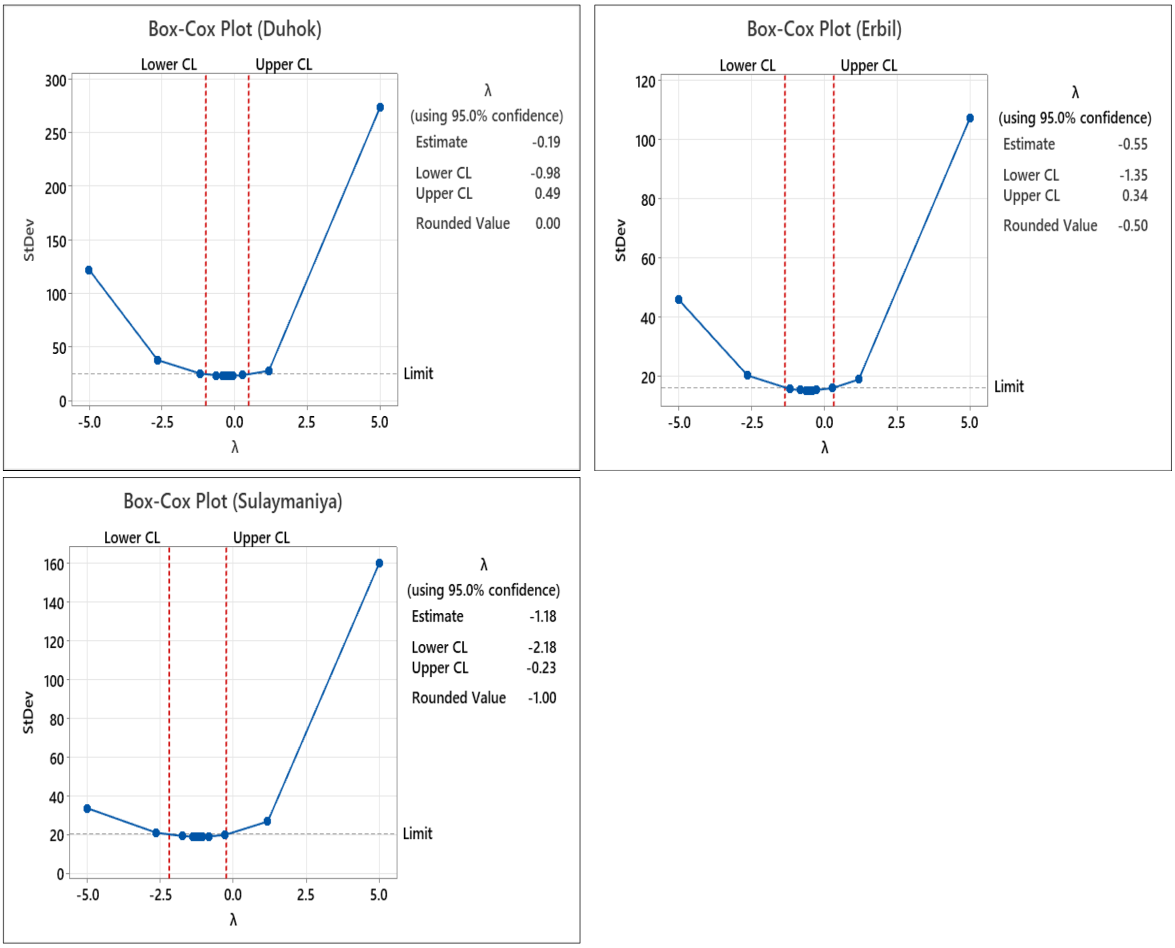


Fig.(3):- Plot Box-Cox of the AMR data for Duhok, Erbil, and Sulaymaniya.

Table(5):-Lambda value (λ) and the Transformed function depend on the Box-Cox method

Station	Lambda value (λ)	Transformed function
Duhok	0	Log (AMR)
Erbil	-0.5	1/Sqrt (AMR)
Sulaymaniya	-1	1/(AMR)

4.3 Stationary and Normality Results on the Transformed Data Series

Tables 6 and 7 provide the findings of the stationary and normality tests applied in section 3.1 of this study to the transformed (AMR) data series of three adopted stations. The transformed (AMR) time-series data for the three adopted stations are stationary and have a normal distribution, as shown in Tables 6 and 7, and by comparing them with the critical value of the significance level of 0.05 (Table 2), respectively. The P-P and Q-Q plots, which display the graphical results of the normality test, were also used in this study's visual inspection of the normality test. In general, the P-P and Q-Q in Fig. 4 show that the transformed (AMR) data for the three adopted stations fits the normal distribution. The transformed (AMR)

gaheen.mohammed@uod.ac
evan.haiani@uod.ac

data will be incorporated into the ARIMA model based on the outcomes of the stationary and normality tests.

Table(6):- Results of stationarity tests of Transformed (AMR) data series

Stations	ADF		PP		KPSS	
	t-statistic	P-value	t-statistic	P-value	t-statistic	P-value
Duhok	-4.666	<0.001	-6.866	<0.001	0.049	0.964
Erbil	-6.238	<0.001	-6.823	<0.001	0.071	0.822
Sulaymaniya	-7.158	<0.001	-7.210	<0.001	0.151	0.512

Table (7):- Results of normality tests of Transformed (AMR) data series

Stations	SW		AD		JB	
	t-statistic	P-value	t-statistic	P-value	t-statistic	P-value
Duhok	0.967	0.411	0.476	0.222	1.021	0.600
Erbil	0.992	0.997	0.111	0.992	0.343	0.842
Sulaymaniya	0.979	0.784	0.196	0.882	0.373	0.830

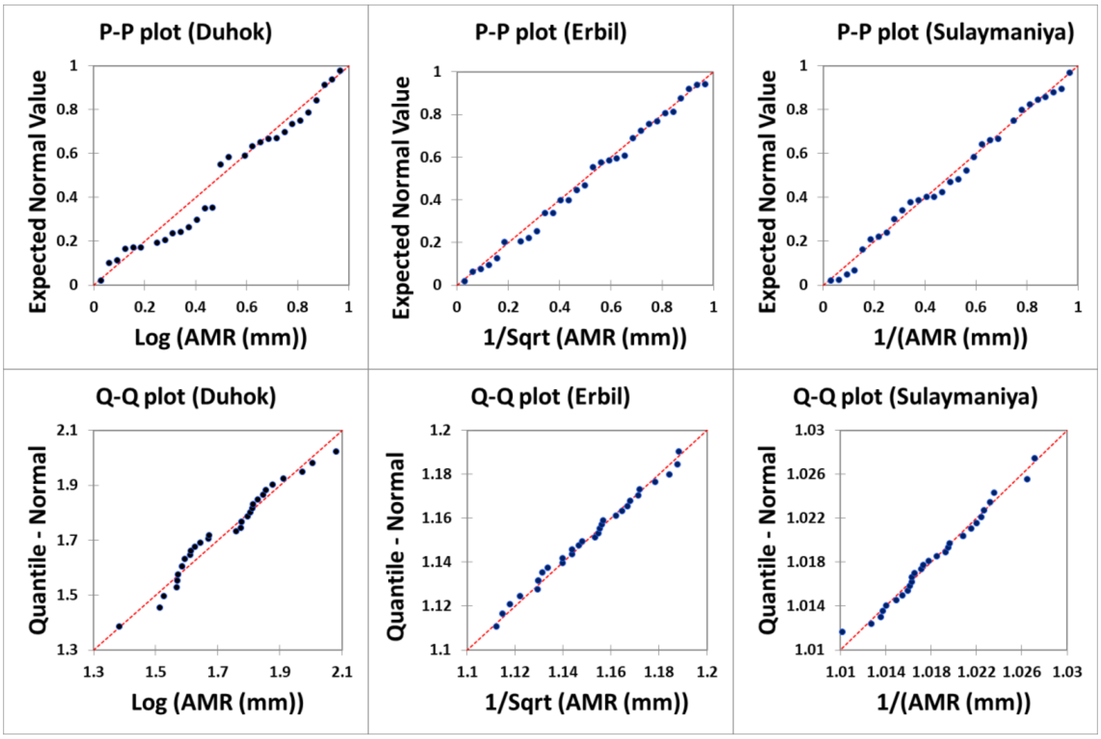


Fig.(4):- Visual inspection of P-P and QQ plots for Transformed (AMR) data for Duhok, Erbil and Sulaymaniya.

4.4 Parameters of the ARIMA model

As mentioned in section 3, the ARIMA model is represented by three parameters (AR, I, and MA), which stand for autoregressive, difference order, and moving average. These three parameters also have statistical significance for the model's accuracy (p, d, and q). Given that the transformed (AMR) data for the three adopted stations are stationary and normal by the analysis of stationary in the preceding subsection, the ARMA (p, q) models are applied to the transformed (AMR) data for the three adopted stations, which results in d=0. For each adopted station in this study, the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the transformed (AMR) data series determine the order of both AR (p) and MA (q) in the ARMA model (Fig. 5). The range of p and q values is appropriately loosened in order to create multiple ARMA (p, q) models, as shown in Tables 8, 9, and 10, to create a more accurate ARMA model for each adopted station based on Fig. 5. Additionally, the five tests (p-value, Adjusted R-squared, AIC value, SC value, and S.E. of regression) results of ARMA (p, q) models are listed in these three tables. These tests are all crucial factors in determining the best ARMA model. Tables 8 to 10 show that the ARMA (0, 3) model is the best optimal model for the Duhok station while the ARMA (0, 4) model is the best optimal model for the Erbil and Sulaymaniya stations based on the results of the five tests (i.e., p-value, Adjusted R-squared, AIC value, SC value, and S.E. of regression) used to determine the optimal ARMA model for each station. For these three optimal models (i.e., ARMA (0, 3) for Duhok station and ARMA (0, 4) for Erbil and Sulaymaniya stations) the Adjusted R-squared shows the higher value, and the other tests (i.e., p-value, AIC value, SC value, and S.E. of regression) show the lowest values among the other models. The model in Tables 8, 9, and 10 that failed the parameter significance test (p-value > 0.05) test was denoted by an asterisk (*). The results of fitting the transformed (AMR) data for the three stations using the best optimal models (ARMA (0, 3), and ARMA (0, 4)) are displayed in Fig. 6.

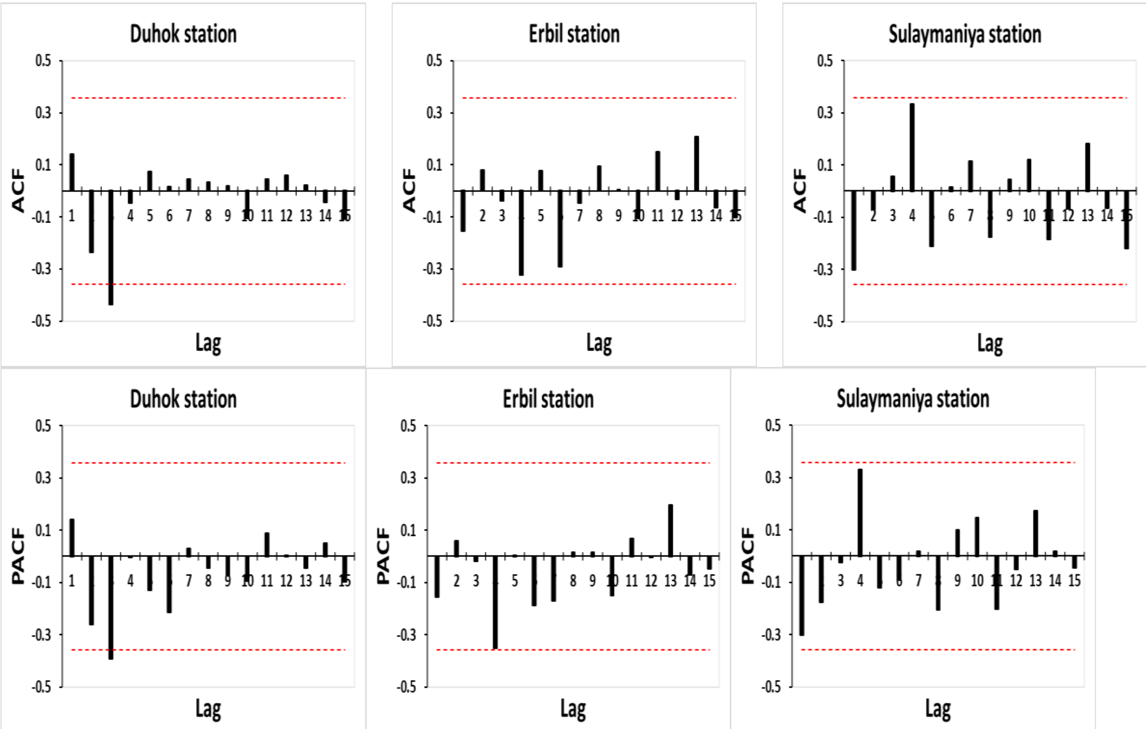


Fig.(5):- The ACF and PACF plot of the Transformed AMR data series for three adopted stations with 5% critical values at $\pm 1.96N^{-0.5}$ (red dotted lines).

Table (8):- Results of the ARMA tests for the Duhok station

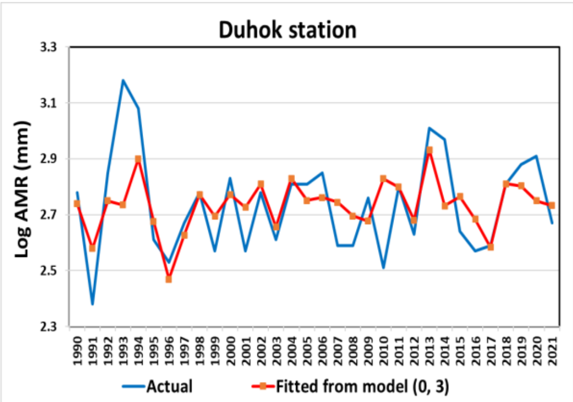
(p,q)	P (AR; MA)	Adjusted R^2	AIC	SC	S. E. of Regression
(0, 2)*	(---; 0.126)	0.010	-0.526	-3.889	0.177
(0, 3)	(---; 0.006)	0.520	-0.663	-0.626	0.154
(2, 0)*	(0.153; ---)	0.002	-0.519	-0.382	0.178
(2, 2)*	(0.915; 0.650)	-0.025	-0.464	-0.281	0.181
(2, 3)	(0.237; 0.005)	0.265	-0.750	-0.566	0.153
(3, 0)	(0.014; ---)	0.187	-0.703	-0.566	0.161
(3, 2)	(0.031; 0.303)	0.189	-0.679	-0.495	0.160
(3, 3)*	(0.848; 0.109)	0.223	-0.703	-0.519	0.157

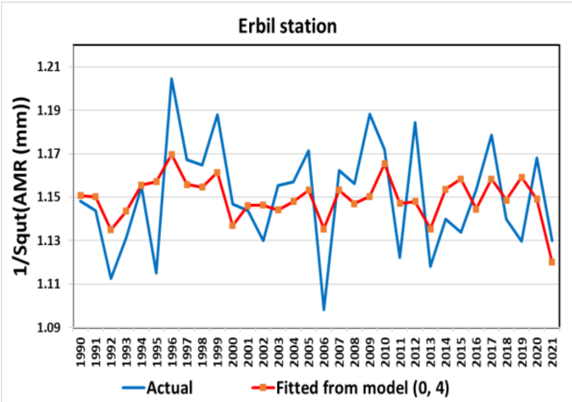
Table (9):- Results of the ARMA tests for the Erbil station

(p,q)	P (AR; MA)	Adjusted R^2	AIC	SC	S. E. of Regression
(0,4)	(---; 0.022)	0.011	-0.972	-0.756	0.147
(4, 0)	(0.021; ---)	0.008	-0.8301	-0.763	0.132
(4, 4)*	(0.576; 0.808)	0.022	-0.841	-0.657	0.149
(6, 0)*	(0.251; ---)	0.011	-0.898	-0.761	0.147
(6, 4)*	(0.166; 0.143)	0.082	-0.926	-0.742	0.141

Table(10):-Results of the ARMA tests for the Sulaymaniya station

(p,q)	P (AR; MA)	Adjusted R^2	AIC	SC	S. E. of Regression
(0, 1)*	(---; 0.058)	0.065	-1.745	-1.608	0.138
(0, 4)	(---; 0.012)	0.155	-1.808	-1.671	0.005
(1, 0)	(0.047; ---)	0.032	-1.019	-0.882	0.139
(1, 1)*	(0.856; 0.450)	0.033	-0.965	-0.782	0.141
(1, 4)	(0.051; 0.013)	0.211	-1.064	-0.881	0.105
(4, 0)*	(0.118; ---)	0.053	-0.725	-0.757	0.110
(4, 1)*	(0.265; 0.069)	0.006	-1.023	-0.830	0.086
(4, 4)*	(0.689; 0.305)	0.127	-1.010	-0.827	0.135





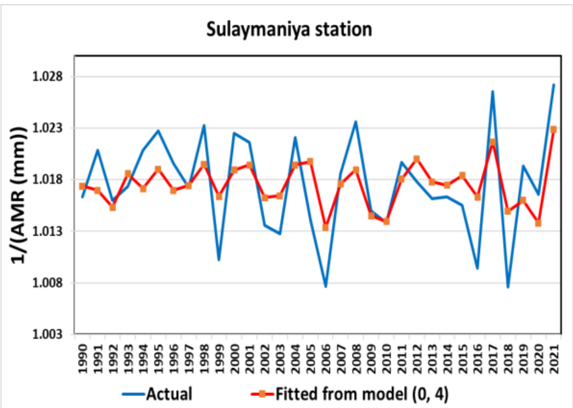


Fig.(6):- Time series data was generated using the ARMA model for the Duhok, Erbil, and Sulaymaniya stations; the actual data is given in blue color and the red line color corresponds to the fitted values.

A white noise test (i.e., ACF and PACF plots with the Q-statistic of the Ljung-Box test) is used on the residual (i.e., In time series analysis, residuals are the difference between the fitted values and the actual values) after fitting the best optimal ARMA models for each station.

White noise means that a variable does not have autocorrelation. The ACF and PACF function plots of the residual series are shown in Fig. 7. The estimated AMR data series from the three optimal models essentially have white noise, as can be seen in Fig. 7, where neither plot (ACF nor PACF) for the three stations shows any significant correlation between lags. It is common to use a Ljung-Box test to check that the residuals from a time series model resemble white noise. Also, after applying for the Ljung-Box test [20], the Q-statistics for Duhok station were 2.756, for Erbil station were 6.558, and for Sulaymaniya station were 8.798. The residual is seen to be white noise, demonstrating the validity of all three optimal ARMA models (i.e., they are normal, have a zero mean, or are serially autocorrelated). In addition, three statistical tests (i.e., RE, RMSE, and MAE) are evaluated between the actual AMR data and the fitted AMR data by the ARMA model, as shown in Table 11. It can be seen in Table 11 that the lower values of MAE, MSE, and RMSE imply higher accuracy of an ARMA model for each adopted station. Finally, the three selected ARMA models for the three adopted stations are used to predict the AMR data values for future periods from 2022 to 2026 (five years), as shown in Table 12.

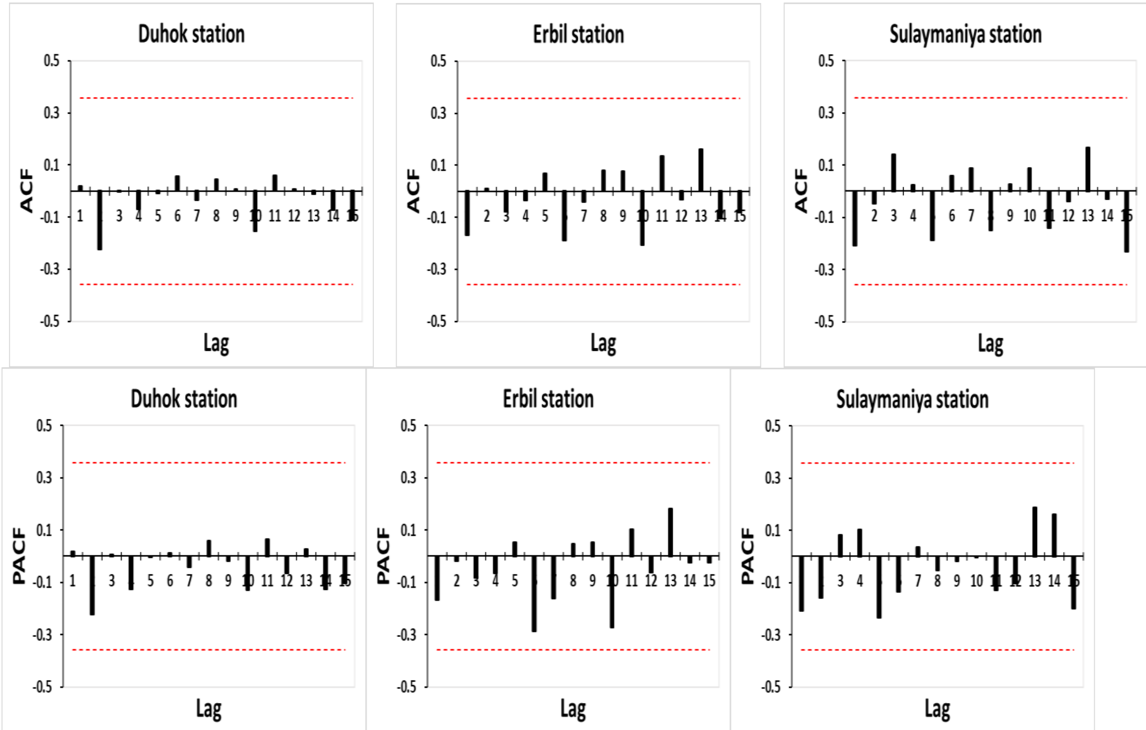


Fig.(7):- The ACF and PACF plot of the residual series for three adopted stations with 5% critical values at $\pm 1.96N^{-0.5}$ (red dotted lines).

Table (11):- Results Average RE (%), RMSE, MAE between actual AMR data and fitted AMR data by ARMA model.

Stations; model	RE (%)	RMSE	MAE
Duhok; ARMA (0,3)	2.747	0.154	0.113
Erbil; ARMA (0,4)	1.494	0.024	0.019
Sulaymaniya; ARMA (0,4)	1.306	0.004	0.003

Table (12):-Time series data forecasted using the ARMA model from 2022 to 2026.

Stations; model	2022	2023	2024	2025	2026
Duhok; ARMA (0,3)	48.419	41.639	59.148	54.057	54.031
Erbil; ARMA (0,4)	41.947	40.133	46.877	41.659	43.842
Sulaymaniya; ARMA (0,4)	80.013	51.779	52.569	50.369	57.052

5. CONCLUSION

The analysis of the rainfall time series in this study has been done using the Box-Jenkins (ARIMA) model methodology. The models were created and tested using the annual maximum rainfall data series for the three rainfall stations in the Kurdistan Region of Iraq (Duhok, Erbil, and Sulaymaniya) from 1990 to 2011. The parameters of the ARIMA models were estimated using plots of the autocorrelation function and partial autocorrelation function. For stationary, the tests Phillips-Perron, Kwiatkowski-Phillips-Schmidt-Shin, and Augmented-Dickey Fuller were all applied. Additionally, the normality tests Shapiro-Wilk, Anderson-Darling, and Jarque-Bera were applied. The Box-Cox transformation was used to make the rainfall time series stationary and normal. Different statistical tests were applied to evaluate the performance of the successful ARIMA models. The findings showed that the ARMA (0, 3) model was best suited for the Duhok station, and the ARMA (0, 4) model was best suited for Erbil and Sulaymaniya. These models were used to forecast the AMR data for the subsequent five years (2022 to 2026). Depending on the results of average statistical test errors (i.e., RE (%), RMSE, and MAE) between actual AMR data and fitted AMR data by the ARMA model, it has been decided that the ARIMA model is suitable for projecting future annual rainfall values in the Kurdistan Region.

6. ACKNOWLEDGEMENTS

The authors are grateful to the Directorate of Meteorology and Seismology in the governorates of Erbil and Duhok for providing the daily rainfall data used in this study. The staff at the College of Humanities and Geographical Department at Duhok University also deserves praise for contributing some essential maps of the Kurdistan Region.

Author Contributions

Dr. Evan Hajani designed this research. Both Authors contributed to the analysis and writing this research.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research is not supported by any funding agency.

7. REFERENCES

S.H. Ahmed. Designation and Study of Anticlines-Kurdistan Region-NE Iraq. 2nd International Science Conference IOP Conf. Series: Journal of Physics: Conf. Series 1294, 082001 IOP Publishing. DOI: 10.1088/1742-6596/1294/8/082001. 2019

O. Agha and S. Bagcaci and N. Şarlak. Homogeneity Analysis of Precipitation Series in North Iraq. IOSR Journal of Applied Geology and Geophysics, 5,57-63. DOI:10.9790/0990-0503025763. 2017

O. Al Balasmeh, R. Babbar and T. Karmaker. Trend analysis and ARIMA modeling for forecasting precipitation pattern in Wadi Shueib catchment area in Jordan. Arabian Journal of Geosciences, 12, 27. DOI: https://doi.org/10.1007/s12517-018-4205-z. 2019

T. Alghamdi, K. Elgazzar, M. Bayoumi, T. Sharaf and S.Shah. Forecasting Traffic Congestion Using ARIMA Modeling. 2019 15th International Wireless Communications and Mobile Computing Conference (IWCMC), 1227-1232. DOI: 10.1109/IWCMC.2019.8766698. 2019

D.T. Anh, T.D. Dang, and S. P. Van. Improved Rainfall Prediction using Combined Pre-Processing Methods and Feed-Forward Neural Networks. Multidisciplinary Scientific Journal, 2, 65-83. DOI: https://doi.org/10.3390/j2010006. 2019

H. Akaike. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19, 716-723. DOI: 10.1109/TAC.1974.1100705. 1974

M.A. Al-Dabbagh. Modeling the Rainfall-Discharge Relation and Discharge Forecasting for some Rivers in Iraq. Unpublished MSc. Thesis, Mosul University. 2005

T.W. Anderson and D.A. Darling. A Test of Goodness of Fit. Journal of American

Statistical Association, 49, 765-769. DOI: https://doi.org/10.2307/2281537. 1954

G.E.P. Box and G.M. Jenkins. Time Series Analysis: Forecasting and Control. Revised Edition, Holden Day, San Francisco. 1976

S. Chattopadhyay and G. Chattopadhyay. Univariate Modelling of Summer-Monsoon Rainfall Time Series: Comparison between ARIMA and ARNN. Comptes Rendus eoscience, 342(2), 100-107. 2010, DOI: 10.1016/j.crte.2009.10.016.

D.A. Dickey and W.A. Fuller. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association 74 427-431. DOI: https://doi.org/10.2307/2286348. 1979

Geetha and G.M. Nasira. Time-series modelling and forecasting: modelling of rainfall prediction using ARIMA model. HomeInternational Journal of Society Systems Science, 8, 4, 361-372, DOI: https://doi.org/10.1504/IJSSS.2016.081411. 2017

J. Gu, S. Liu, Z. Zhou, S.R. Chalov and Q. Zhuang. A Stacking Ensemble Learning Model for Monthly Rainfall Prediction in the Taihu Basin, China. Water, 14 (492), 2-20. DOI: https://doi.org/10.3390/w14030492. 2022

E. Hajani. Climate Change and its Influence on Design Rainfall At-site in New South Wales State, Australia. Journal of Water and Climate Change 11 (S1), 251-269. 2020

E. Hajani and Z. Klari. Trends Analysis in Rainfall Data Series in Duhok city, Kurdistan Region, Iraq. Modeling Earth Systems and Environment, 1-14. DOI: https://doi.org/10.1007/s40808-022-01354-1. 2022

E. Hajani, K. Shajee, F. Kaled and H. Abdulhaq. Characteristics of Changes in Rainfall Data in the Kurdistan Region Iraq. Arabian Journal of Geosciences, 1-21. DOI: https://doi.org/10.1007/s12517-022-09809-6. 2022

A.B. Koehler and E.S Murphree. A Comparison of the Akaike and Schwarz Criteria for Selecting Model Order. Journal of the Royal Statistical Society. Series C (Applied Statistics), 37, 2, 18795. DOI: https://doi.org/10.2307/2347338. 1988

D. Kwiatkowski, P.C.B. Phillips, P. Schmidt and Y. Shin. Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. Journal of Econometrics, 54 (1-3), 159-178. DOI: https://doi.org/10.1016/0304-4076(92)90104-Y. 1992

P. Li. Box-Cox Transformations: An Overview. University of Connecticut, Storrs. 2005

G.M. Ljung and G.E. Box. On a Measure of a Lack of Fit in Time Series Models. Biometrika. 65 (2), 297-303. DOI: https://doi.org/10.1093/biomet/65.2.297. 1978

K. Mammas and D.F. Lekkas. Rainfall Generation using Markov Chain Models; Case Study: Central Aegean Sea. Water, 10, 856. DOI: https://doi.org/10.3390/w10070856. 2018

S. Makridakis, S.C. Wheelwright, and R.J. Hyndman. Forecasting: Methods and Applications, 3rd ed., John Wiley, Hoboken, N. J.

Meher J. (2014). Rainfall and Runoff Estimation Using Hydrological Models and ANN Techniques. Doctor Thesis, National Institute of Technology Rourkela-769008, India. 1998

S.S. Shapiro and M.B. Wilk. An Analysis of Variance Test for Normality (complete samples). Biometrika, 52 (3-4), 591-611. DOI: https://doi.org/10.2307/2333709. 1965

R.H. Shumway and D.S. Stoffer. Time Series Analysis and Its Applications with R Examples. Springer Science, Business Media, Berlin. 2005

S. Sukarna, E. Ananda and M. Wahyuni. Rainfall Forecasting Model Using ARIMA and Kalman Filter in Makassar, Indonesia. Journal of Physics Conference Series 2123, 1, 012044. DOI: 10.1088/1742-6596/2123/1/012044. 2021

M. Peña, A. Va´zquez-Patiño, D. Zhiña, M. Montenegro and A. Alex. Improved Rainfall Prediction through Nonlinear Autoregressive Network with Exogenous Variables: A Case Study in Andes High Mountain Region. Hindawi Advances in Meteorology, 1828319, 17. DOI: https://doi.org/10.1155/2020/1828319. 2020

P.C.B. Phillips and P. Perron. Testing for a Unit Root in Time Series Regression. Biometrika, 75 (2), 335-346. DOI: https://doi.org/10.2307/2336182. 1988

W.M. Ridwan, M. Sapitang, A. Aziz, K.F. Kushiar, A.N. Ahmed and A. El-Shafie. Rainfall Forecasting Model using Machine Learning Methods: Case study Terengganu, Malaysia. Ain Shams Engineering Journal, 12 (2), 1651-1663. 2021, DOI: 10.1016/j.asej.2020.09.011.