Uniport Journal of Engineering and Scientific Research (UJESR)

Vol. 8, Issue 1, 2023, Page 116-122
ISSN: 2616-1192

(www.ujesr.org)

Forecasting Drought with Autoregressive Integrated Moving Average (ARIMA) model and Standardised Precipitation Index (SPI)

Ibrahim, U.A[1*], Mshelia, J.P[1] and Muhammad, H[2]
[1]Department of Civil and Water Resources Engineering, University of Maiduguri, Maiduguri, Borno State
[2]Department of Civil Engineering Technology, The Federal Polytechnic Damaturu, Damaturu, Yobe State
*Corresponding author's email: *uaibrahim@unimaid.edu.ng*

Abstract
Climate change is expected to exacerbate drought conditions in many regions, making it important to understand the potential impacts and develop effective adaptation strategies. Estimating and predicting droughts, however, is not always simple. An objective drought evaluation method known as a Drought Index (DI) was developed, and the DI was later utilized to develop a drought forecasting tool to foretell future drought conditions. In this study, an evaluation of existing DIs was first conducted in terms of their suitability for the assessment of drought conditions in Maiduguri. Based on the findings of the evaluation, Autoregressive Integrated Moving Average Model (ARIMA) was chosen. The ARIMA considered rainfall variable that affects drought. Investigating the scope and severity of drought, predicting future drought conditions, and mitigating drought were the study's primary breakthroughs. Following the research, there was a noticeable decrease in mean precipitation as well as the future occurrence of droughts of varying severity in Maiduguri. To mitigate the occurrence of drought, a comprehensive plan that addresses both the short and long-term causes of drought should be used to lessen their effects. Overall, a multi-pronged strategy that combines several tactics will be most successful in lowering the incidence and effects of drought in Maiduguri.

Keywords: *Drought, Index, Rainfall, Impact, Climate change*

Received: 19th December, 2023 *Accepted: 31st December, 2023*

1. Introduction

Drought is a complex natural phenomenon and has significant impacts on effective water resources management. Among various definitions of drought, a more recognized and logical definition is that drought can be caused by a period of severe scarcity of water resources with respect to normal conditions corresponding to the place and time or a period of abnormal dry conditions that last long enough to create an imbalance in the hydrological condition (Bahrami *et al.*, 2021; Dai, 2011). In general, drought gives an impression of water scarcity due to insufficient precipitation, high evapotranspiration and over-exploitation of water resources or a combination of all the aforementioned (Caloiero and Veltri, 2019; Hao *et al.*, 2018).

Droughts can be categorized into four classes depending on the conditions: meteorological, agricultural, hydrological and socioeconomic droughts. Conceptually, a meteorological drought is expressed as a precipitation deficit over a region for a period of time (Buttafuoco *et al.*, 2014; Mishra *et al.*, 2011). A drought is considered to be specific to a region for the fact that weather conditions of low precipitation, dry winds and high temperature are highly variable and do vary from region to region. When atmospheric moisture is reduced to a level where the soil moisture is affected, the onset of an agricultural drought is imminent (Zargar *et al.*, 2011). During this period, crops and animals are affected as the decline in soil moisture content leads to reduction in crop production, which subsequently affects the balance of the food chain in the ecosystem. In a situation where water supply is insufficient to meet water demand, socio-economic drought is evident. This results in negative consequences for the society, the economy and the environment (Hayes *et al.*, 2012).

There are several methods that have been used in the past as drought assessment tools such as measurement of lack of rainfall, shortage of stream flow, drought indices (DIs) among others. However, traditionally the estimation of future dry conditions (or drought forecasting) has been conducted using DIs as the most common drought assessment tools. This is because the DI is expressed by a number which is believed to be far more functional than raw data during decision making (Khan *et al.*, 2020; Hayes *et al.*, 2012). The DI in general is a function of several hydro-meteorological variables such as rainfall, temperature, streamflow and storage reservoir volume. Thus, in modelling and forecasting drought index, time series are of great importance. Time series methods provide an important approach in drought forecasting. One of the most widely used time series models is the Autoregressive Integrated Moving Average (ARIMA) for its remarkable accuracy in forecasting time-oriented occurrences (Rezaiy and Shabri, 2023; Zhang, 2003.) The ARIMA model has several advantages over other models, in particular its predictability and richer information about changes over time (Rezaiy and Shabri, 2023; Shatanawi *et al.*, 2013; Hu *et al*., 2007).

A lot of researches have used ARIMA model in drought prediction (Rezaiy and Shabri, 2023; Rahman *et al.*, 2017; Bazrafshan *et al.*, 2015; Shatanawi *et al.*, 2013; Han *et al.*, 2010). Thus, the aim of this research is to use the ARIMA model to lay and predict future drought conditions (2023 – 2050) for Maiduguri, Nigeria. In addition, it will provide a subjective method to deal with climate-related parameters in the study area.

2. Materials and methods

2.1 Study area

Maiduguri (Latitude 11.8311º N, Longitude 13.1510º E) with an area of 15.1 km^2 is located in the Northern part of Nigeria (Fig. 1). The climate of Maiduguri is characterized by a long dry season with high evaporation rate from October to May and a short wet season for the remaining part of the year (Waziri, 2012). However, the native of the city identified four seasons. These are rainy season (June to September), harvest season (September to November), harmattan or cool season (December to February) and hot season (March to May). Generally, the mean monthly temperature is always above 20°C but the daily extremes vary in a wide range reaching up to 47°C in April (Waziri, 2012).

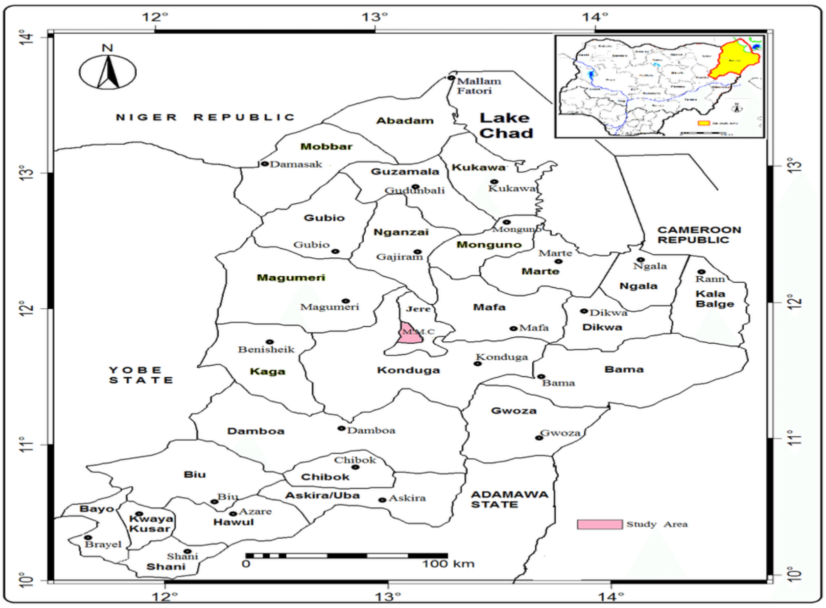

Fig. 1: Location of the study area

2.2 Data collection

Monthly historical precipitation data of Maiduguri meteorological station was obtained from the Nigerian Meteorological Agency (NiMet) for the period of 41 years (1980 – 2020). The mean monthly precipitation of the 40 years data series was 49.71mm.

2.3 Standardized precipitation index

Standardized Precipitation Index (SPI), developed by McKee *et al.* (1993) for the purpose of drought monitoring was used in this study. It was recommended by World Meteorological Organization (WMO) in 2009 as the main meteorological drought index to be used by countries in monitoring and following drought conditions (Fluixá-sanmartín *et al.*, 2018). SPI was chosen because it's simple and uses only precipitation data, making it very easy to be used than other drought index methods. SPI is obtained by dividing the difference between precipitation and mean at a selected period to standard deviation using Equation (1) to fit a distribution.

$$SPI = \frac{x_i - \bar{x}_i}{\sigma} \quad (1)$$

where x_i is the observed precipitation at a station, $\bar{x}_i$is the mean of the observed precipitation, and σ is the standard deviation.

Computation of the SPI involves fitting a gamma probability density function to a given frequency distribution of cumulative precipitation at a station (Thom, 1958). The gamma distribution is defined in terms of its frequency presented in Equation (2).

$$G(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{a-1} e^{\frac{-x}{\beta}} \quad (2)$$

where α > 0, α is a shape factor, β > 0, β is a scale factor, *x* >0, *x* is the precipitation amount, and $\Gamma(\alpha)$ defines the gamma function defined by Equation (3).

$$\Gamma(\alpha) = \int_0^{\infty} y^{a-1} e^{-y} \quad (3)$$

From Thom (1958), the maximum likelihood solutions were used to optimally estimate α and β as presented in Equations (4) to (5).

$$\alpha = \frac{1}{4A}\left(1 + \sqrt{1 + \frac{4A}{3}}\right) \quad (4)$$

$$\beta = \frac{\bar{x}}{\alpha} \quad (5)$$

and *A* can be calculated as:

$$A = \ln \bar{x} - \frac{\sum \ln(x)}{n} \quad (6)$$

where n is the number of precipitation observations. The cumulative probability function is then calculated as:

$$G(x) = \int_0^x g(x)dx = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^x x^{\alpha-1}\, e^{-x/\beta} dx \quad (7)$$

If the precipitation series have zero values which is $x = 0$, the gamma function is undefined, and the cumulative probability distribution is described as:

$$H(x) = q + (1 - q)G(x) \quad (8)$$

where q is the probability for zero value. However, if m is assumed as zero values in a given precipitation series, then q can be defined as:

$$q = \frac{m}{n} \quad (9)$$

The cumulative probability value $H(x)$ is then transformed to standard normal random value, z, which represents the SPI value with a zero mean and a standard deviation. SPI is categorized based on their range values presented in Table 1.

Table 1: Categories of Drought Based on SPI Range Values

SPI	Classification
>2.00	Extremely Wet
1.50 to 1.99	Very Wet
1.00 to 1.49	Moderately Wet
0.00 to 0.99	Near Normal
0.00 to -0.99	Mild Drought
-1.00 to -1.49	Moderate Drought
-1.50 to -1.99	Severe Drought
< 2.00	Extremely dry

(McKee *et al.*, 1993)

2.4 ARIMA model

The ARIMA model is one of the most commonly used methods applied in hydrological and metrological forecasting. It was initially introduced by Box and Jenkins (1976), which is a suitable solution for addressing non-stationarity in historical time series records. The model consists of three parts: an Autoregressive (AR) part, a moving average (MA) part and the differencing part. The model is usually referred to as ARIMA **(**p,d,q**)** model, where P the order of the autoregressive part, d is the order of the differencing part and q is the order of the moving average part. If $d = 0$, the model become ARIMA, which is a linear stationary model. Also, when $d > 0$, it becomes a linear non-stationary model (Etuk, 2012).

The main stages in setting up a forecasting ARIMA model includes model identification, model parameter estimation and diagnostic checking for the identified model appropriateness for modelling and forecasting. Model identification involves examining the data to check the most appropriate class of ARIMA process. This is done through selecting the order of the consecutive and seasonal differencing required in making the series stationary, as well as specifying the order of the regular and seasonal autoregressive and moving average polynomials necessary to adequately represent the time series model (Rezaiy and Shabri, 2023; Al Sayah *et al.*, 2021). The general form of the ARIMA model describing the current value X_t X_t of a time series by its own past is presented as Equation (10):

$$(1-\emptyset_1\mathrm{B})(1-\alpha_1\mathrm{B})^{12}(1-\mathrm{B}^{12})\mathrm{X}_\mathrm{t} = (1-\theta_1\mathrm{B})(1-\gamma_1\mathrm{B}^{12})\mathrm{e}_\mathrm{t} \quad (10)$$

where $1-\emptyset_1\mathrm{B}$ is the non-seasonal autoregressive of order 1, $1-\alpha_1\mathrm{B}$ is the seasonal autoregressive of order 1, X_t is the current value of the time series examined, B is the backward shift operator ($\mathrm{BX}_\mathrm{t} = \mathrm{X}_{\mathrm{t}-_1}$ and $\mathrm{B}^{12}\mathrm{X}_\mathrm{t} = \mathrm{X}_{\mathrm{t}-12}$), B is the first order non-seasonal difference, B^{12} is the seasonal difference of order 1, $\theta_1\mathrm{B}$ is the non-seasonal moving average of order 1 and $1-\gamma_1\mathrm{B}^{12}$ is the seasonal moving average of order 1. This model can be multiplied out and used for forecasting after the model parameters were estimated.

2.5 Parameter estimation

After choosing the most appropriate model (Step 1, Section 2.4), the model parameters were estimated (Step 2) by using the least square method. In this step, values of the parameters were chosen to make the sum of the squared residuals (SSR) between the real data and the estimated values as small as possible. In general, non-linear estimation method was used to estimate the above identified parameters to maximize the likelihood (probability) of the observed series given the parameters values.

2.6 Diagnostic checking

After the parameter estimation was completed, a diagnostic check (evaluation matrices) was performed (Step 3). In diagnose checking, the residuals from the fitted model were examined against adequacy. The acceptability or appropriateness of model guarantees that the time series is in the time with model assumptions are well founded. This is usually done by correlation analysis through the residual autocorrelations function (ACF) plots. In this study, three diagnostic statistics were used; Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), and Mean Standard Deviation (MSD) to make sure whether these residuals correspond with error terms or not.

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{x_i-\hat{x}_i}{x_i}\right| \quad (11)$$

$$MAD = \frac{\sum|x_i-\hat{x}_i|}{N} \quad (12)$$

$$MSD = \frac{\sum|x_i-\hat{x}_i|^2}{N} \quad (13)$$

where x_i is the predicted value, $\hat{x}_i$ is actual value and N is the number of total data points. The model with the minimum mean standard deviation was considered the best for modelling and understanding the pattern of rainfall in the study area.

3. Results and discussion

3.1 Drought characteristics examining suitable trend model for forecast

Annual precipitation data (1980 – 2020) were examined for temporal fluctuation and presented in Fig. 2. More so, results of the computed SPI based on period of occurrence is presented in Table 2. From Table 2, SPI of the precipitation data series indicated 0 years extremely wet, 3 years very wet, 2 years moderately wet, 18 years near normal, 12 years mild drought, 3 years moderate drought, 3 years severe drought and 0 years extremely drought conditions. Thus, indicating a fewer drought event occurring in the study area. It was also observed that, 7.3% of the precipitation period is Very Wet, 5% is Moderately Wet, 44% is Near Normal, 29.1% is Mild Drought, 7.3% is Moderate Drought and 7.3% is Severe Drought. The results indicated that Maiduguri experiences mild to severe drought events.

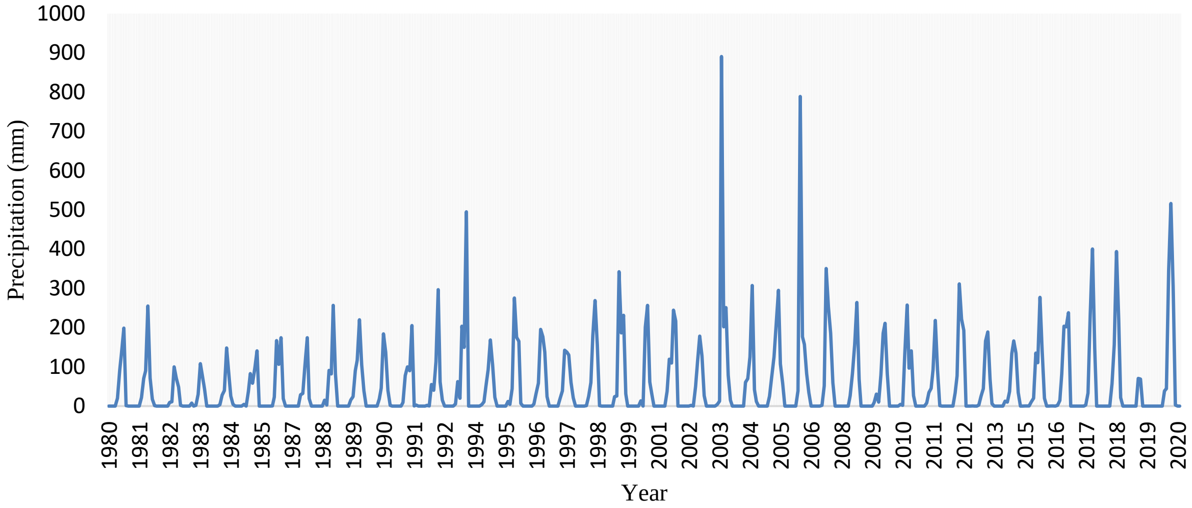

Fig. 2: Annual precipitation (mm) from 1980 to 2020 for Maiduguri Station

Table 2: Classification of Maiduguri's SPI based on period of occurrence

SPI	Classification	Period of occurrence
2 >	Extremely Wet	0
1.5 – 1.99	Very Wet	2003, 2006 and 2007
1 – 1.49	Moderately Wet	2004 and 2018
0 – 0.99	Near Normal	1989, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2005, 2008, 2010, 2011, 2012, 2017 & 2020
0 – (–0.99)	Mild Drought	1980, 1981, 1982, 1985, 1986, 1987, 1988, 1990, 1991, 1992, 2013 and 2019
–1 – (–1.49)	Moderate Drought	1983, 1984 and 2009
–1.5 – (–1.99)	Severe Drought	2014, 2015 and 2016
< 2.00	Extremely Dry	0

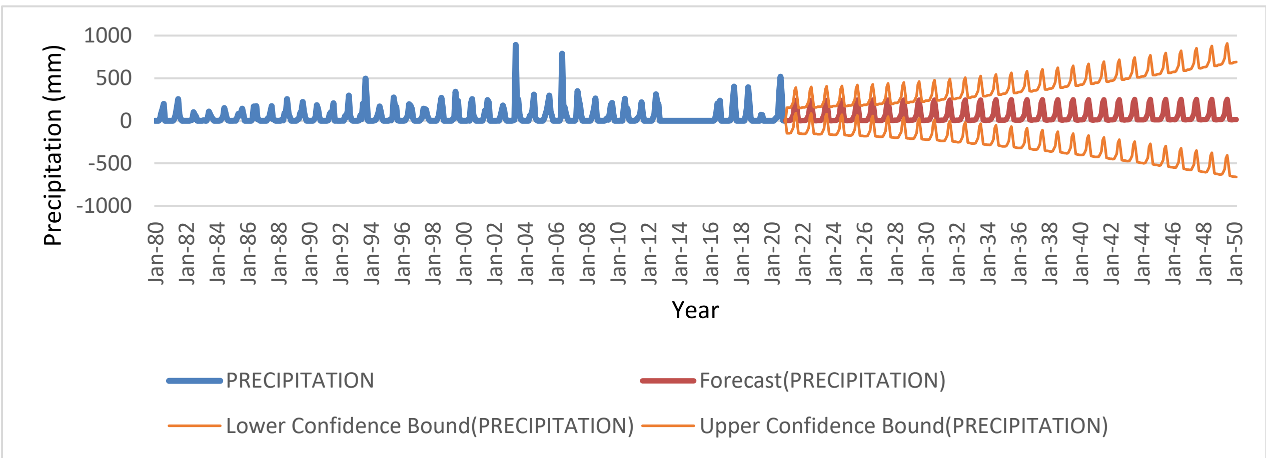

Fig. 3: Forecast result from 2023-2050 at 95% confidence bound

3.2 Parameter estimation and ARIMA model forecasting

The Arima (1, 1, 0) model was used to predict the rainfall from the period 2023 to 2050. It was found that the system changes as time changes, but later becomes static forecast values made by the ARIMA (1, 1, 0) Model as presented in Figure 3.

3.3 Diagnostic checking of residuals

The data were analysed by trend equation to test for the presence of unit root in the time series data, presentation of the trend line equation, i.e., linear trend equation, quadratic trend equation and their related graphs. For the analysis of the forecasting performance or ability of the models to forecast rainfall, the results of the three error measures: MAPE, MAD and MSD are presented in Table 3. Results of the linear and quadratic trends (Table 3) showed a steady straight-line. Indicating that the residuals for the timescales are uncorrelated and exhibited the properties of a white noise process i.e. they are normally distributed, randomly scattered and possess constant variances.

Table 3: Summary of trend analysis accuracy measures

Trend	MAPE	MAD	MSD
Linear	1663.87	50.66	4614.07
Quadratic	1671.52	50.85	4613.38

4. Conclusion

This paper used time series stochastic model (ARIMA) to predict future possible droughts in Maiduguri, Nigeria. The degree and incidence of drought in Maiduguri was predicted to have varying severity. This was evidenced by a noticeable decrease in mean precipitation. The stochastic model developed to predict drought has achieved good simulation results in predicting medium and short-term drought. More so, the linear stochastic model can be used to quickly predict watershed droughts with scarce data, analyze the severity of future droughts, and also be used by decision makers and resource planners to predict drought severity in advance. Therefore, it is important to employ a comprehensive strategy that addresses both the short and long-term causes of drought.

References

Al Sayah, M.J., Abdallah, C.A., Khouri, M. and Darwich, T.A. (2021) Framework for climate change assessment in Mediterranean data-sparse watersheds using remote sensing and ARIMA modeling. Theoretical and Applied Climatology, 143: 639–658.

Bahrami, M., Bazrkar, S. and Zarei, A.R. (2019) Modeling, prediction and trend assessment of drought in Iran using Standard Precipitation Index. Journal of Water and Climate Change, 10(1): 181-196.

Bazrafshan, O., Salajegheh, A., Bazrafshan, J., Mahdavi, M. and Maraj, A.F. (2015) Hydrological drought forecasting using ARIMA models (case study: Karkheh Basin). Ecopersia, 3: 1099–1117.

Box, G.E.P. and Jenkins, G.M. (1970) Time Series Analysis: Forecasting and Control, 1st ed.; Holden-Day: San Francisco, CA, USA, 1970.

Buttafuoco, G. and Caloiero, T. (2014) Drought events at different timescales in southern Italy (Calabria). Journal of Maps, 10: 529–537.

Caloiero, T. and Veltri, S. (2019) Drought assessment in the Sardinia region (Italy) during 1922–2011 using the standardized precipitation index. Pure Appl. Geophys., 176: 925–935.

Dai, A. (2011) Drought under global warming: A review. Wiley Interdisciplinary Reviews: Climate Change, 2(1): 45–65.

Etuk, H.E. (2012) A seasonal ARIMA model for Nigerian Gross Domestic Product. Developing Country Studies, 2(3): 1 - 11.

Fluixá-sanmartín, J., Pan, D., Fischer, L., Orlowsky, B., García-hernández, J. and Jordan, F. (2018) Searching for the optimal drought index and timescale combination to detect drought: a case study from the lower Jinsha River basin, China. Hydrology and Earth System Sciences, 22: 889–910.

Han, P, Wang, P.X., Zhang, S.Y. and Zhu, D.H. (2010) Drought forecasting based on the 540 remote sensing data using ARIMA models. Math Comput. Model, 51: 1398–1403.

Hao, Z., Singh, V.P. and Xia, Y. (2018) Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. Reviews of Geophysics, 56: 108-141.

Hayes, M.J., Svoboda, M.D., Wardlow, B.D., Anderson, M.C. and Kogan, F. (2012) Drought monitoring:Historical and current perspectives. In Brian, D.W., Martha, C.A. and James, P.V. (Eds), Remote Sensing of Drought: Innovative Monitoring Approaches (pp. 1 - 19). Florida, United States of America: CRC Press.

Hu, W., Tong, S., Mengersen, K. and Connell, D. (2007) Weather variability and the incidence of cryptosporidiosis: Comparison of time series Poisson regression and SARIMA models. Ann. Epidemiol., 17: 679–688.

Khan, M.M.H., Muhammad, N.S. and El-Shafie, A

(2020) Wavelet based hybrid ANN ARIMA models for meteorological drought forecasting. Journal of Hydrology, 590: 1 - 8.

McKee, T.B., Doesken, N.J. and Kliest, J. (1993) The relationship of drought frequency and duration to time scales. In C. Anaheim (Ed.), 8th Conference on Applied Climatology, 17–22 January, 1993 (pp. 179– 184). Boston, MA: American Meteorological Society.

Mishra, A.K. and Singh, V.P. (2011) Drought modeling - A review. Journal of Hydrology, 403(1–2): 157–175.

Rahman, M.A., Yunsheng, L. and Sultana, N. (2017) Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, spearman's rho tests and ARIMA model. Meteorology and Atmospheric Physics, 129 (4): 409 - 424.

Rezaiy, R. and Shabri, A. (2023) Drought forecasting using W-ARIMA model with standardized precipitation index. Journal of Water and Climate Change, 14(9): 1 - 23.

Shatanawi, K., Bahbeh, M. and Shatanawi, M. (2013) Characterizing, Monitoring and Forecasting of Drought in Jordan River Basin. Journal of Water Resources and Protection, 5(12): 1192 - 1202.

Thom, H.C.S. (1958) A Note on the gamma distribution. Monthly Weather Review. 86. 117-122. Geneva, Switzerland.

Waziri, M. (2012) Spatial Pattern of Maiduguri City: Researchers' Guide. Adamu Joji Publishers. Kano City, p.37.

Zargar, A., Sadiq, R., Naser, B. and Khan, F.I. (2011) A review of drought indices. Environmental Reviews, 19(1): 333–349.

Zhang, G.P. (2003) Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50: 159–175.