**International Journal of Interdisciplinary Research Statistics, Mathematics and Engineering**
**ISSN: 2992-9261| Impact Factor: 6.46**
Volume. 09, Number 4; October-December, 2022;
Published By: Scientific and Academic Development Institute (SADI)
8933 Willis Ave Los Angeles, California
https://sadipub.com/Journals/index.php/ijirsme | editorial@sadipub.com

# OPTIMIZING RESOURCE MANAGEMENT USING SEASONAL ARIMA RAINFALL FORECASTING MODEL FOR ODEDA, NIGERIA

**Olamide Adeboye**
Department of Civil Engineering, Federal University of Agriculture, Abeokuta, Nigeria

**Abstract:** Forecasting rainfall is essential for countries with economies hinged on agriculture, engineering, and tourism, as well as for the supply of water and sanitation. In this study, a seasonal ARIMA model was used to forecast rainfall events in the Odeda region, Southwestern Nigeria. The rainfall records of the OgunOshun River Basin Development Authority weather station from January 1982 to December 2016 were employed to develop the model. A visual inspection of the autocorrelation function (ACF) and partial autocorrelation (PACF) of tentative models was done using the Box-Jenkins approach. The selected preliminary models were subjected to selection criteria of coefficient of determination (R2), Bayesian information criterion (BIC), and Root Mean Square Error (RMSE) and diagnostic checking. The selected best model was Seasonal ARIMA (3,1,3) (2,1,1)12 with a coefficient of determination of 0.68 for predicting rainfall events from January 2017 to December 2018. Hence, the model was utilized to forecast rainfall events till 2022. With the increasingly changing patterns of rainfall occurrence worldwide due to climate change, this study highlights the importance of analyzing and forecasting rainfall.

**Keywords:** ARIMA, Box-Jenkins, Odeda, Rainfall, Time-series.

**Introduction:** Forecasts of rainfall are vital for countries whose economies depend on agriculture, engineering, tourism, and water supply and sanitation. With the changing patterns of rainfall occurrence globally, the importance of analyzing and predicting rainfall cannot be overemphasized. Over the years, numerous statistical models, including ARIMA and SARIMA, have been used to forecast rainfall occurrence successfully, especially with seasonal data. In this study, we employed a seasonal ARIMA model to forecast rainfall events within Odeda, Ogun State, Nigeria, and using data from the Ogun-Oshun River Basin Development Authority weather station from January 1982 to December 2016. A seasonal and ordinary differencing method was adopted to eliminate any non-stationarity in the raw data. The processes applied to select the best model include visual inspection of the autocorrelation and partial autocorrelation plots of each model, coefficient of determination, Bayesian information criterion, and Root Mean Square Error. The selected best model was then subjected to diagnostic checking. The model was found to accurately forecast rainfall events from January 2017 to December 2018, with a coefficient of determination of 0.68 when compared with the actual rainfall events for that period. Thus, this study provides insight into the forecasting of rainfall occurrence in Odeda and highlights how vital it is to analyze and forecast rainfall, especially with the changing patterns of rainfall occurrence worldwide due to climate change.

## 2.0 MATERIALS AND METHOD

### *2.1 Description of Study area*

Odeda is a Local Government Area located in Ogun in southwestern Nigeria. Its coordinates extend from longitude 3.3899E to 3.47950E and latitude 7.1745N to 7.2300N. It is neighbored by Abeokuta North, Abeokuta South, Remo North, and Obafemi Owode Local government Areas (Adeleke *et al*., 2015). The meteorological features of the study area include a distinct wet season which is derived from the Moist Maritime Southerly Monsoon from the Atlantic Ocean and a dry season derived from the Continental North Easterly harmattan winds from the Sahara desert (Makinde *et al*., 2017). The dry season spell extends from November to February while the rainy season occurs predominantly from March to October with a break in precipitation in August. The study area exhibits Temperature records ranging from 25.7oC experienced during the rainy season, and 30.2oC experienced during the dry season. Humidity is lowest during the dry season with readings ranging from (37-54%) and highest during the rainy season with values ranging from (78-to 85 %) (Makinde *et al*., 2017).

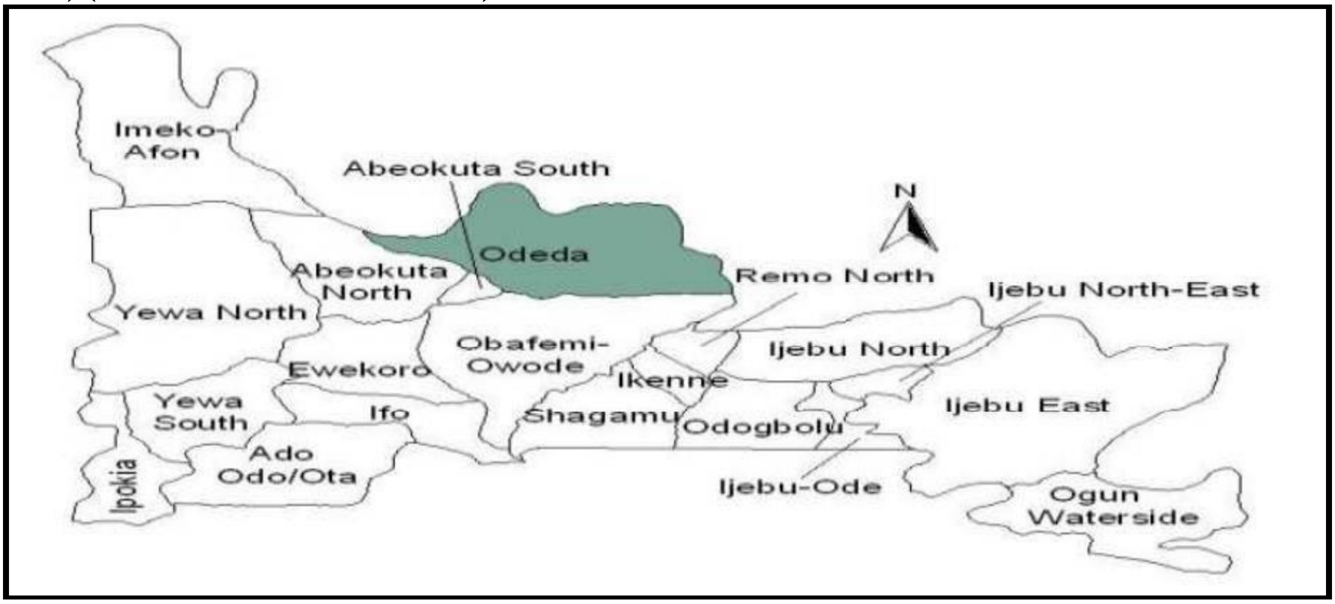


Figure 1: Map of Odeda Local Government Area, Ogun State, Nigeria (Source: Adeleke *et al*., 2015).

### *2.2 ARIMA principles and criteria for model selection*

A typical ARIMA model made is up of three parameters, p (number of autoregressive values), d (the order of differencing) which also indicates the number of times needed to subject the time series to a sort of equilibrium, and q (the number of moving average values). The Box-Jenkins model also known as the (p-d-q) model is the non-seasonal part of the model (Asadi, *et al.*, 2013). The ARIMA model is applied to a nonstationary series of data (Jakasa *et al*., 2011). Analysis of the time series is made in four stages; the first stage involves identification, which is followed by estimation, diagnostic checks, and finally, forecasting (Bari *et al*., 2015). The seasonal ARIMA model can apply both seasonal and non-seasonal factors in a multiplicative model. This can be illustrated by the notation shown below; ARIMA (p, d, q) × $(P, D, Q)_S$ Where, p - non-seasonal Auto-Regressive (AR) order d - Nonseasonal differencing q - Non-seasonal Moving Average (MA) order

P - Seasonal (AR) order

D - Seasonal differencing, and

Q - Seasonal Moving Average (MA) order.

S is the duration of the seasonal pattern (Bari *et al*., 2015; Box *et al*., 1994).

### *2.3 Model Identification and selection*

The rainfall data for the period from January 1982 – to December 2016 was obtained from the Ogun-Oshun River Basin Development Authority (OORBDA) and there should be an elimination of any form of nonstationarity of the time series plot if found to be present in the time plot. This can be done by differencing (or seasonal differencing). The series should also be checked for seasonality which can also be eliminated by the differencing process. A time series that had been differenced to become a stationery series might still exhibit some auto-correlated errors which can be removed by adding AR terms ($p \geq 1$) and MA terms ($q \geq 1$) in the forecasting equation (Dimri *et al*., 2020).

Models were identified by the difference between the autocorrelation function (ACF) and partial autocorrelation (PACF) functions which can determine the order number of the preliminary model and a subsequent trial and error method was adopted for different models based on the coefficient of determination ($R^2$), Bayesian information criterion (BIC) and Root Mean Square Error (RMSE) before a final model was then selected. Thus, the identification of p and q values based on ACF and PACF for non-seasonal time series and P and Q values based on ACF and PACF for seasonal time series ensures that preliminary models are selected tentatively using appropriate (pdq) and/or (PDQ) values.

### *2.4 Model Diagnostic Checking*

Diagnostic checks are carried out after the models selected are fit the data with the model's residuals which are expected to be uncorrelated with constant variance. The ACF and PACF of residuals were plotted for standard checks against the occurrence of spikes on the error correlogram. It is often assumed that the error is normally distributed and therefore the residuals are expected to be normally distributed. The Q-Q plot and error histogram were adopted to ensure that there was a normal distribution of the residuals (Bari *et al*., 2015). Any inadequacies observed indicate a need to adopt an alternative model, and then repeat the whole iterative cycle of identification, estimation, and application until an acceptable model is obtained (Graham and Mishra, 2017).

### *2.5 Rainfall Forecast*

Having passed through the aforementioned processes, the selected model was used to forecast future rainfall values.

## 3.0 RESULTS AND DISCUSSION 3.1 Preliminary Data Analysis

The data analysis was implemented in the SPSS software version 21 and from the time plot shown below in Figure 2, it can be observed that the dataset is seasonal with a periodicity of twelve months and appears nonstationary due to the nature of the mean and variance of the time series.

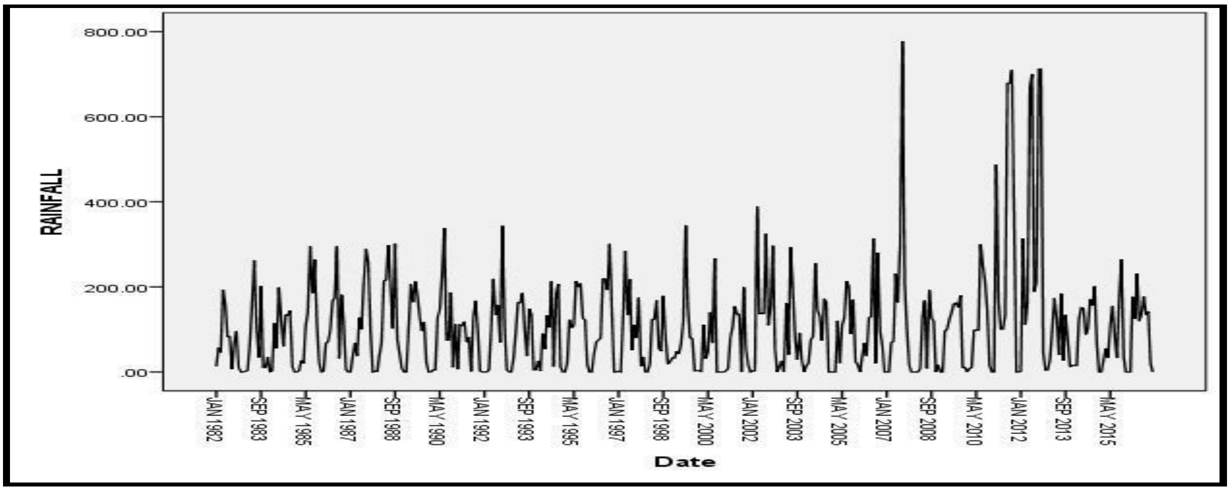


Figure 2: Time series Plot of raw Rainfall data for Odeda.

Because of the presence of seasonality, the time series was made stationary by applying both seasonal and ordinary differencing with an order of 1 in the time series plot with the plot shown in Figure 3 below;

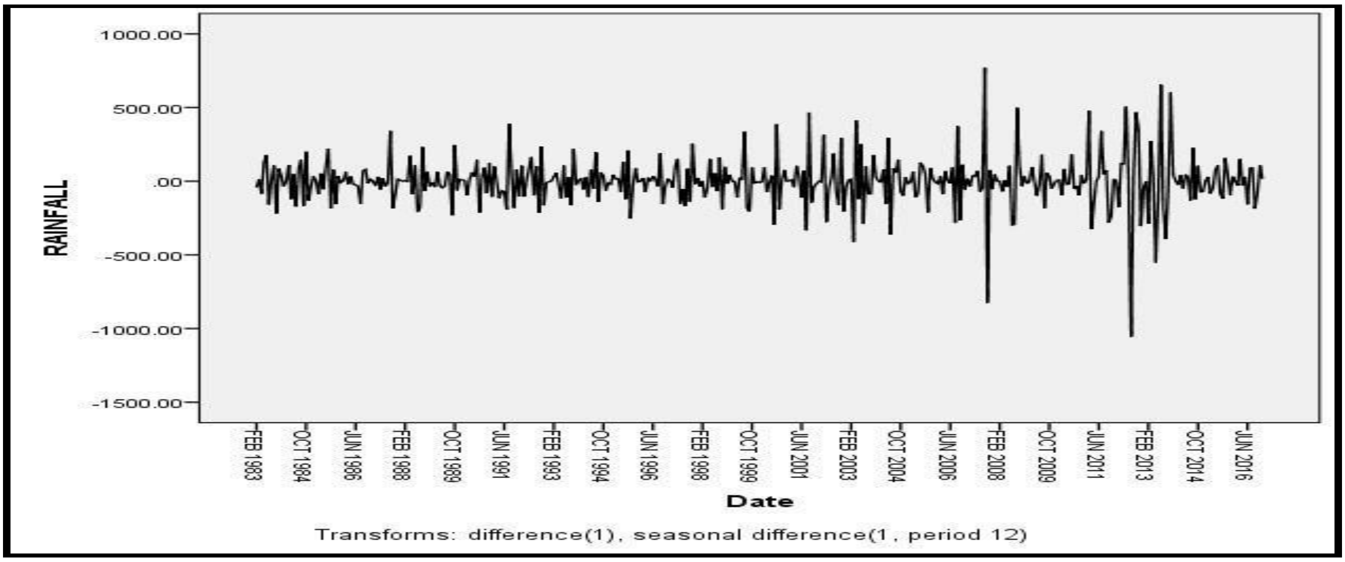


Figure 3: Time series Plot of Rainfall data for Odeda after seasonal differencing of order 1, and ordinary differencing, also of order 1.

***3.2 Model Selection***

The parameters p, q, P and Q were estimated by visual inspection of the autocorrelation (ACF) and partial autocorrelation (PACF) plots of the differenced time series shown in Figures 4(a) and 4(b) of the stationary process in order estimate the order of the model.

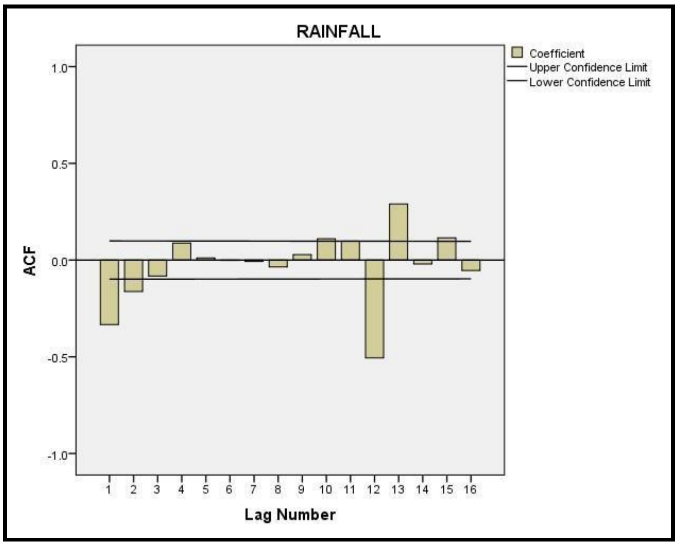


Figure 4(a): ACF plot of rainfall for Ogun-Oshun weather station at Odeda after seasonal differencing of order 1, and ordinary differencing, also of order 1.

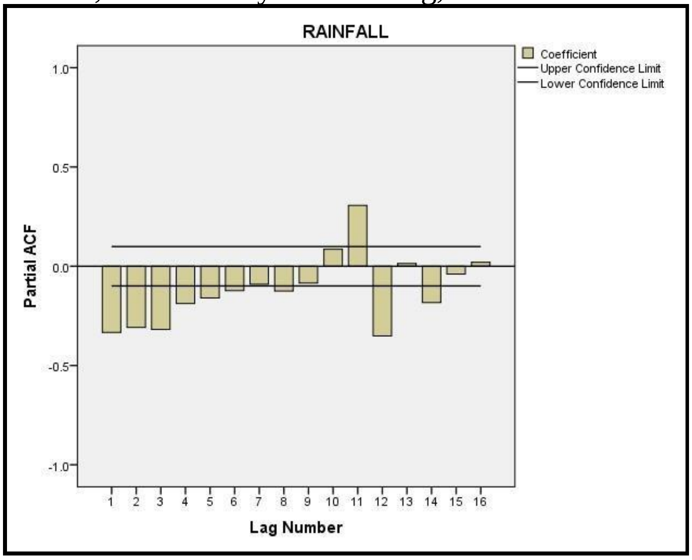


Figure 4(b): PACF plot of rainfall for Ogun -Oshun weather station at Odeda after seasonal differencing of order 1, and ordinary differencing, also of order 1.

From the visual inspection of the ACF and PACF plots, several models were tentatively selected and the model (in the following order) with the best coefficient of determination ($R^2$), lower RMSE, and lower BIC as well as ability to generate white noise residuals, were applied to pick the best model. These are listed below in Table 1 while the best models that were selected are given below in Table 2;

Table1: Tentative Seasonal ARIMA Models selected

| S/No | MODEL | $R^2$ | RMSE | BIC |
|---|---|---|---|---|
| 1 | $(2,1,1)\,(3,1,1)_{12}$ | 0.38 | 97.58 | 9.29 |
| 2. | $(3,1,3)\,(2,1,1)_{12}$ | 0.40 | 97.29 | 9.28 |
| 3. | $(1,1,1)\,(1,1,1)_{12}$ | 0.38 | 97.68 | 9.25 |
| 4. | $(3,1,3)\,(2,1,2)_{12}$ | 0.40 | 96.43 | 9.30 |
| 5. | $(3,1,3)\,(2,1,3)_{12}$ | 0.40 | 96.76 | 9.32 |

Note; only the best five models were listed after applying rules of parsimony. Table 2: The best seasonal ARIMA models for Forecasting Rainfall in Odeda

| S/No | MODEL | $R^2$ | RMSE | BIC |
|---|---|---|---|---|
| 1. | $(3,1,3)\ (2,1,1)_{12}$ | 0.40 | 97.29 | 9.28 |
| 2. | $(3,1,3)\ (2,1,2)_{12}$ | 0.40 | 96.43 | 9.30 |
| 3. | $(3,1,3)\ (2,1,3)_{12}$ | 0.40 | 96.76 | 9.32 |

These three Seasonal ARIMA models were also selected because the ACF and PACF plots of the residuals of all the models were investigated for any significant spikes after applying the rule of parsimony with the results indicating that it was only the three models listed in Table 2 that had no significant correlation observed which, in turn, is an exhibition of white noise properties. However, model **$(3,1,3)\ (2,1,1)_{12}$** was preferred based on the criteria for selection. The other models did not pass these last criteria.

### *3.3 Diagnostic Check*

An inspection of the autocorrelation and partial autocorrelation charts of the residuals indicates that there are no spikes that extend beyond the continuous lines as shown in Figure 5 below which implies the approximate 95% confidence limits. This also indicates that the residualss are white noise and therefore, adequate.

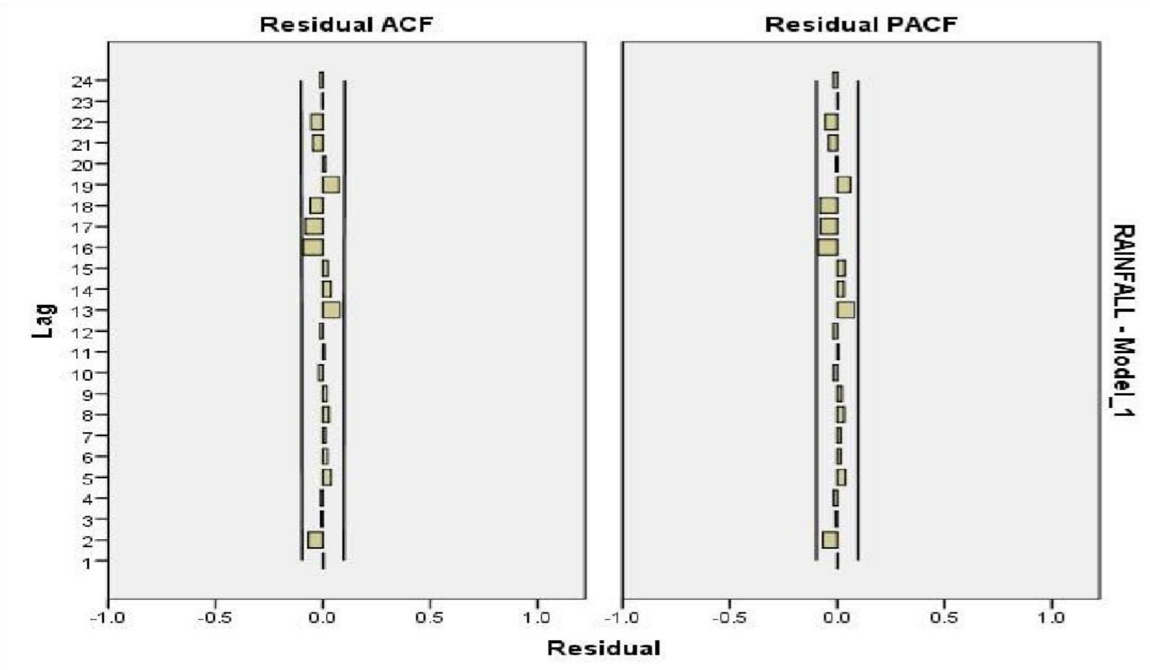


Figure 5: ACF and PACF charts of Seasonal ARIMA model $(3,1,3)\ (2,1,1)_{12}$ residuals.

Further diagnostic check was done on the normality of noise residuals and from figure 6 shown below, the histogram shows a bell-shaped distribution which indicates the normality of the distribution.

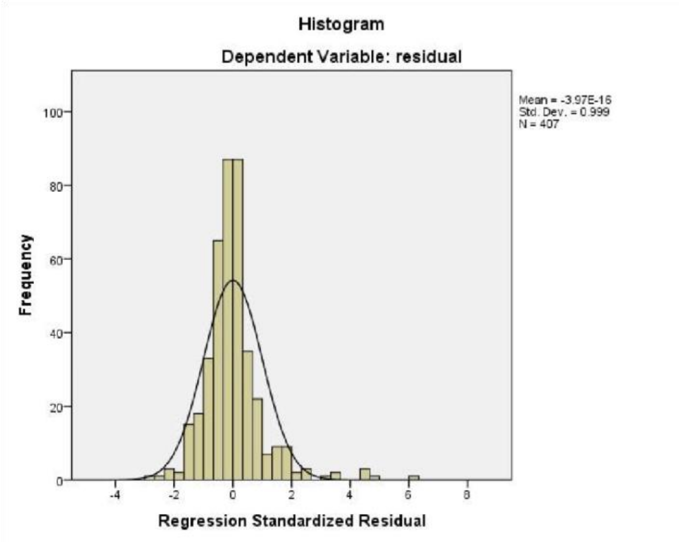


Figure 6: Histogram of Noise residuals distribution.

The normal Q-Q plot of the noise residuals shown below in Figure 7 indicates that the points mostly lie on and around the straight line and that these residuals appear to be normally distributed thus further validating the selected model for forecasting purposes.

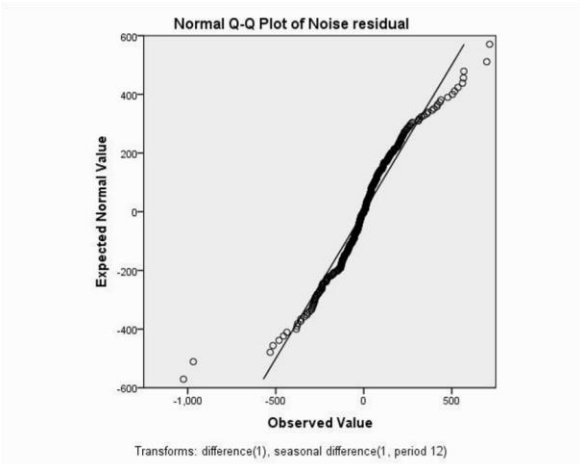


Figure 7: Normality diagnostic plot of the residuals.

### *3.4 Model validation and forecast of rainfall events*

The seasonal ARIMA model (3,1,3) $(2,1,1)_{12}$ was used to validate a forecast rainfall event between January 2017 and December 2018 with the predicted values along with the lower class limits (LCL) and upper-class limits (UCL) displayed on the table and was found to be adequate with a coefficient of determination of 0.68 when compared predicted values were compared with the actual rainfall events for that period. The plot is shown in figure 8. It was further used to forecast rainfall events up to the year 2022 as shown below in figure 9. The forecasted values show a similar pattern to the original data series. The seasonal rainfall pattern within the study area appears to have a significant impact on the time series analysis as a result of its non-linear nature, thus resulting in the lower values of the coefficient of determination, $R^2$.

Table 3: Comparison between actual and validated values of rainfall events used for validation

| Date | Actual Rainfall | Predicted Rainfall | LCL Predicted | UCL Predicted |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| **Jan-17** | 0 | -7.85 | -191.92 | 176.22 |
| **Feb-17** | 13.2 | -24.64 | -221.28 | 172 |
| **Mar-17** | 63 | 47.22 | -152.41 | 246.85 |
| **Apr-17** | 44 | 94.97 | -104.84 | 294.79 |
| **May-17** | 198.3 | 142.75 | -59.64 | 345.14 |
| **Jun-17** | 134.1 | 156.31 | -47.3 | 359.91 |
| **Jul-17** | 157.1 | 136.63 | -67.23 | 340.49 |
| **Aug-17** | 84 | 85.46 | -120.87 | 291.78 |
| **Sep-17** | 121.4 | 146.35 | -60.86 | 353.56 |
| **Oct-17** | 58.7 | 121.74 | -85.9 | 329.39 |
| **Nov-17** | 16.6 | -15.18 | -225.12 | 194.75 |
| **Dec-17** | 1.3 | -15.19 | -225.83 | 195.44 |

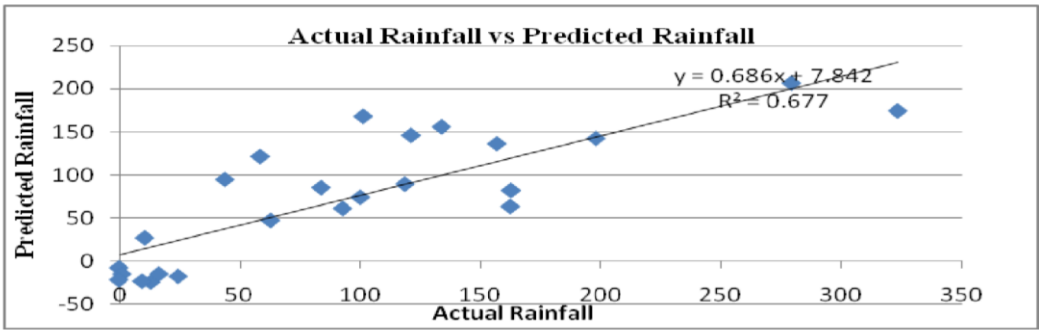


Figure 8: Validation of the seasonal ARIMA model (3,1,3) (2,1,1)$_{12}$

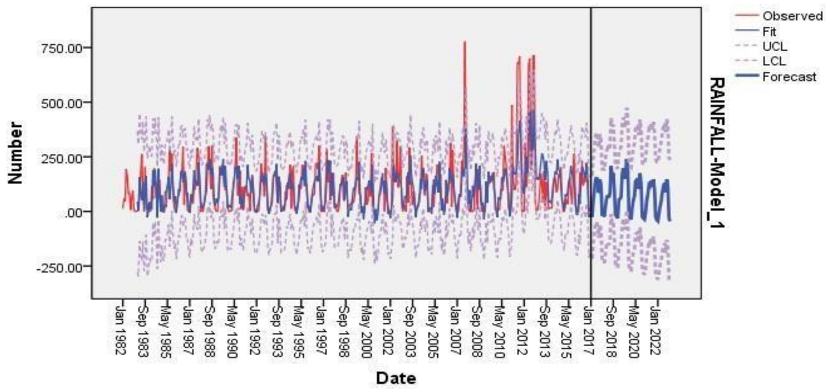


Fig. 9 Forecast of Seasonal ARIMA (3,1,3) (2,1,1)$_{12}$ model

## 4.0 CONCLUSION

The performance of a Seasonal ARIMA was used to develop a monthly weather forecast for Odeda located in Ogun State, Southwestern Nigeria. Seasonality had a significant effect on the rainfall pattern. Two Seasonal ARIMA models were selected from about 10 models based on criteria that included the maximum value of Coefficient of determination, $R^2$, the minimum value of root mean square error (RMSE), Bayesian information criterion (BIC), and a diagnostic check on their noise residuals. The Seasonal ARIMA model (3,1,3) (2,1,1)$_{12}$ was adopted as the most reliable rainfall forecast over the study area. The accuracy of the predictions made was hampered by the marked rainy and dry seasons which generated data values that appeared like outliers and thus resulted in a lower coefficient of determination, $R^2$. It can be concluded that the study showed that a Seasonal ARIMA model can be adopted as an effective tool to forecast rainfall within the study area.

**ACKNOWLEDGMENT**

The authors wish to express their appreciation to the management of the Ogun-Oshun River Basin Development Authority (OORBDA) for releasing the twenty-five-year rainfall data spanning from January 1982 to December 2016 that was used for this study.

**REFERENCES**

Adeleke, O. O., Makinde, V., Eruola, A. O., Dada, O. F., Ojo, A. O. and Aluko, T. J. (2015). Estimation of Groundwater Recharges in Odeda Local Government Area, Ogun State, Nigeria using Empirical Formulae, *Challenges*, 6: 271-281

Asadi, A., Vahdat, S .F. and Sarraf, A. (2013). The forecasting of Potential Evapotranspiration using time series analysis in humid and semi humid regions. *American Journal of Engineering Research*, 2 (12): 296302

Bari, S. H., Rahman, M. T., Hussain, M. M., and Sourav, R. (2015). Forecasting Monthly Precipitation in Sylhet City Using ARIMA Model. *Civil and Environmental Research*, 7 (1): 69-77

Box, G. E .P. and Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control. Holden-Day Publication

Box, G .E. P., Jenkins, G. M. and Reinsel, G. C. (1994). Time Series Analysis, Forecasting and Control, Englewood Cliffs, N. J: Prentice Hall

Dimri, T., Ahmad, S., and Sharif, M. (2020). Time series analysis of climate variables using seasonal ARIMA approach, *J. Earth Syst. Sci*. 129:149

Graham, A. and Mishra, E. P., (2017). "Time series analysis model to forecast rainfall for Allahabad region", *Journal of Pharmacognosy and Phytochemistry*, 6 (5): 1418-1421

Htike, K. K., & Khalifa, O. O. (2010). Rainfall forecasting models using focused time delay neural networks. *In Computer and Communication Engineering (ICCCE)*, 2010 International Conference on (pp. 1-6). IEEE.

Jakasa, T., Androsec, I. & Sprcic, P. (2011). Electricity price forecasting –ARIMA model approach, Proceedings of the International Conference on the European Energy market (EEM) 25-27 May 2011, Zagreb, Croatia page 222-225.

Makinde, V, Adeleke1, O. O., Eruola, A. O., Okeyode, I. C., Akinboro, F. G., Alatise, O. O. and Dada, O. F. (2017). "Impact of Climate Change on Groundwater Resources in Odeda Local Government Area, Ogun State, Southwest Nigeria", *The African Review of Physics* 12: Special Issue on Applied Physics in Africa 0001.

Nayak, D. R., Mahapatra, A. and Mishra, P. (2013). A survey on rainfall prediction using artificial neural network. *International Journal of Computer Applications*, 72(16).

Wang, S., Feng, J., and Liua, G (2013). Application of seasonal time series model in the precipitation forecast, *Mathematical and Computer Modelling*, 58: 677-683.