IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Mathematical Modelling of Engineering Problems
Vol. 11, No. 1, January, 2024, pp. 159-168
Journal homepage: http://iieta.org/journals/mmep

Seasonal Autoregressive Integrated Moving Average Modelling and Forecasting of Monthly Rainfall in Selected African Stations

Timothy Kayode Samson[1*], Francis Olatunbosun Aweda[2]

[1] Statistics Programme, College of Agriculture, Engineering and Science, Bowen University, Iwo 540004, Nigeria
[2] Physics Programme, College of Agriculture, Engineering and Science, Bowen University, Iwo 540004, Nigeria

Corresponding Author Email: kayode.samson@bowen.edu.ng

https://doi.org/10.18280/mmep.110117

Received: 30 September 2023
Revised: 12 November 2023
Accepted: 20 November 2023
Available online: 30 January 2024

Keywords:
rainfall, Seasonal Integrated Moving Average (SARIMA), Africa, modelling, forecasting, RMSE

ABSTRACT

Africa as a continent is blessed with arable land suitable for crop production but this cannot be fully harnessed without proper understanding of the rainfall pattern. Modelling and forecasting rainfall in Africa is even more important now considering the climate change that has brought a new narrative into the rainfall pattern globally, Africa inclusive. This study applied Seasonal Integrated Moving Average (SARIMA) models in modelling and forecasting rainfall across five selected African stations with one station each from the five African regions: West (Abuja, Nigeria), East (Nairobi, Kenya), South (Pretoria, South- Africa), North (Cairo, Egypt) and Central Africa (Yaoundé, Cameroon). Monthly rainfall data for these stations between 1980 and 2022 (42 years) were obtained from the MERRA-2 satellite. However, the data for this study were obtained from the solar radiation data archive website (www.soda-pro.com). The Soda service provides time series of solar radiation data derived from satellites. Furthermore, Modern-Era Retrospective Analysis for Research and Application-2 (MERRA-2) data were extracted from the satellite, which included meteorological and atmospheric data. Since January 1980, the data has been available in hourly, daily, and monthly increments. However, missing data values were checked and removed before implementing the analysis in this study. The determination of the specific SARIMA parameters orders for each city was carried by manual tuning after observing the plots of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). Descriptive analysis revealed that Abuja had the highest variance in the amount of rainfall compared with other major cities in Africa while rainfall in Yaoundé between March and June was higher than that of other stations. Monthly rainfall was stationary in all the stations as shown by the result of Augmented Dickey ($p<.05$) except for Yaoundé which was stationary after the first differencing. Based on the result of out-of-sample forecast performance, different SARIMA models were found to be suitable for rainfall in each of the stations with ARIMA (0,0,1) $(1,0,1)_{12}$ for Abuja (RMSE=70.7044) and Nairobi (RMSE=92.8925), ARIMA (1,0,1) $(1,0,1)_{12}$ for Cairo (RMSE=9.9279), ARIMA (2,0,0) $(1,0,1)_{12}$ for Pretoria (RMSE=42.05462) and ARIMA (1,1,1) $(1,0,1)_{12}$ for Yaoundé (RMSE=79.42084). The findings show that the seasonal terms were statistically significant in all models which justified the use of seasonal ARIMA models in modelling rainfall in these selected stations in Africa. This also underscored the significant role of the season in the rainfall pattern in the selected African stations. Findings also revealed that the previous month's rainfall has a positive influence on the present month's rainfall in some of the stations.

1. INTRODUCTION

The global climate condition in recent years has had a significant impact on rainfall trends. Climate change is known to be driven by natural resource management, which aids in the forecasting and prediction of rainfall patterns [1]. According to the report, climate change has a significant ecological impact on hydrology, natural resource degradation, and human development in the environment [1]. As a result, it is critical to understand how rainfall patterns vary around the world. This is due to the global rainfall pattern, which has a significant impact on climate change [1]. Rainfall plays an important role in the development of any society, providing important information to agriculturalists in planning farming activities [2]. Rainfall reduction has been reported to affect agricultural product decrease as a result of global climate change [3], which includes African countries. According to study [1], information on the pattern of rainfall may have future consequences on the possibilities of prediction and forecasting, which have great implications on the agricultural product globally. According to research, the duration and frequency of rainfall intensity, that is, the amount of water that falls in any country are affected by the period and change in the weather of the environment [1]. The recurrence of rainfall flow volume shall affect the future level of water flow in any country, as a result of an increase in the volume of water downpours as a result of temperature changes [4]. Hyndman and Athanasopoulos [5] and Aweda et al. [6] have emphasized the need for government to put up actionable policies targeted towards mitigating the effects of rainfall on the environment. Researchers such as Hyndman and Athanasopoulos [5] and Aweda et al. [6] have reported that the main reason for forecasting is to be able to predict the future occurrence of any environment so that policymakers can make accurate decisions on rainfall and other climatic occurrences. The rainfall trend can be forecasted using the time series analysis method. As a result, for this study, rainfall prediction and forecasting will be based on the Seasonal Autoregressive Moving Average (SARIMA) as reported by Ramli et al. [7]. According to the report, there are various variations around an average length value, which is a modelling and prediction tool for the use of atmospheric data forecasting tools [8-10]. According to the study made by Heizer et al. [11], time series are various variable values observed at a specific time interval. According to the study made by Suhartono [12], future prediction is based on the forecasting of some parameters, which involved the use of some mathematical tools on historical data. However, using time series analysis [1], it is possible to predict the future trend of data for past events. Autoregressive Integrated Moving Average (ARIMA) model has been reported by Suhartono et al. [12-17] to solve various data issues. It is based on univariate stochastic prediction models [13]. It has been demonstrated that ARIMA time series have an accurate predictive method in comparison to other models [1]. As demonstrated in the forecasting of the periodic intensity of stream flow of the SERIMA model, a series of static analyses were created [14, 15]. Similarly Aweda et al. [6] and Ramli et al. [7] adopted a non-seasonal ARIMA model, ARIMA (p,d,q) while Subbaiah Naidu [16] and Chen et al. [17] employed Seasonal ARIMA (SARIMA) model. However, the SARIMA rainfall prediction model is known to be based on time series analysis which can also be used for the development of various data to follow the seasonal pattern mentioned by Liu et al. [18]. Thereby it has advantages over other models due to its ability to sense and house different datasets [1]. Various researchers such as Ramli et al. [1], Aweda et al. [2], Dindarloo et al. [19], El-Mallah and Elsharkawy [20] applied SARIMA model in their research due to its accuracy and predictive ability.

Short-term forecasting with the ARIMA model is good, but long-term forecasting of rainfall is better [21]. According to research, the SARIMA model and ANN can be used to aid education prediction in the study of rainfall and other parameters [1, 22, 23]. However, for this study, the author employs the ARIMA model to forecast rainfall across selected African stations. According to Afrifa-Yamoah et al. [23] and Sampson et al. [24], the SARIMA model forecasts monthly rainfall for the future and is used to predict what may occur in the future. According to Papalaskaris et al. [25] and Mohamed and Ibrahim [26], time series modelling using the ARIMA model on some meteorological data, hydrological data, and other data shows that some cities are prone to flooding due to the impact of rainfall around the world. As a result, the author intends to use the SARIMA model to forecast rainfall parameters in some selected African stations for this study. These stations were selected based on their climatic region and the activities that are likely to occur in the sub-region. However, the effect of climate change on rainfall has made a significant contribution to the development of Africa Station. As a result, there is a need to study and report on the future occurrence of rainfall in Africa using SARIMA models.

This study is very significant due to the fact that predicting the rainfall can help in decision making regarding agricultural production, weather planning as well as mitigating the risk that could result in excessive or inadequate rainfall. Proper understanding of the future rainfall pattern can inform agricultural decision making regarding what should be planted, when it should be planted and the possibility of having an alternative source of rainfall in order to guarantee food sufficiency. The issue of food insufficiency is a global problem, and Africa is blessed with arable lands suitable for agricultural production. It is believed that the result of this study will provide adequate information that will assist in making best use of rainfall while also providing possible mitigating measures to forestall any negative effect of rainfall on livelihood.

2. RESEARCH METHODOLOGY

2.1 The process of data collection

Monthly rainfall data for selected African stations were obtained from the HelioClim-1 (www.soda-pro.com) archives using the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) technique. This followed the findings made by Aweda and Samson [27] and Gelaro et al. [28]. On May 1st 2023, HelioClim-1 data used for the study years was accessed for all the stations considered, that is, Abuja, Cairo, Nairobi, Pretoria, and Yaoundé. Furthermore, the data used ranges from 1980 to 2022 were collected as a monthly average for January through December of each year in comma-separated value (CSV) format, according to Aweda et al. [2, 29].

Table 1. African station division according to their regions and period of data collection

Station	Division	Country	Longitude	Latitude	Period of Data
Abuja	Hinterland Region	Nigeria	07.399°*W*	09.077°*N*	1980 – 2022
Cairo	Nile Valley Region	Egypt	30.0444°*N*	*31.2357°E*	1980 – 2022
Nairobi	Highland Region	Kenya	1.2921°*S*	36.8219°*E*	1980 – 2022
Pretoria	Coastal Region	South Africa	25.7479°*S*	28.2293°*N*	1980 – 2022
Yaoundé	Hinterland Region	Cameroon	3.8480°*N*	11.5021°*E*	1980 – 2022

2.2 The study area and their locations

The selected African sub-region stations used for this study were divided into various divisions, and the coordinates of each station are listed and shown in Table 1 and Figure 1. These stations are divided into the regions of North, South, East, West, and Central Africa. These stations were chosen to determine the rainfall pattern across African stations based on their region and location. The selected locations for this research were: Abuja (07.399°*W*, 09.077°*N*), Cairo (30.0444°*N*, *31.2357°E*), Nairobi (1.2921°*S*, 36.8219°*E*), Pretoria (25.7479°*S*, 28.2293°*N*), and Yaoundé (3.8480°*N*, 11.5021°*E*). However, the study spanned from 1980 to 2022 across all stations considered.

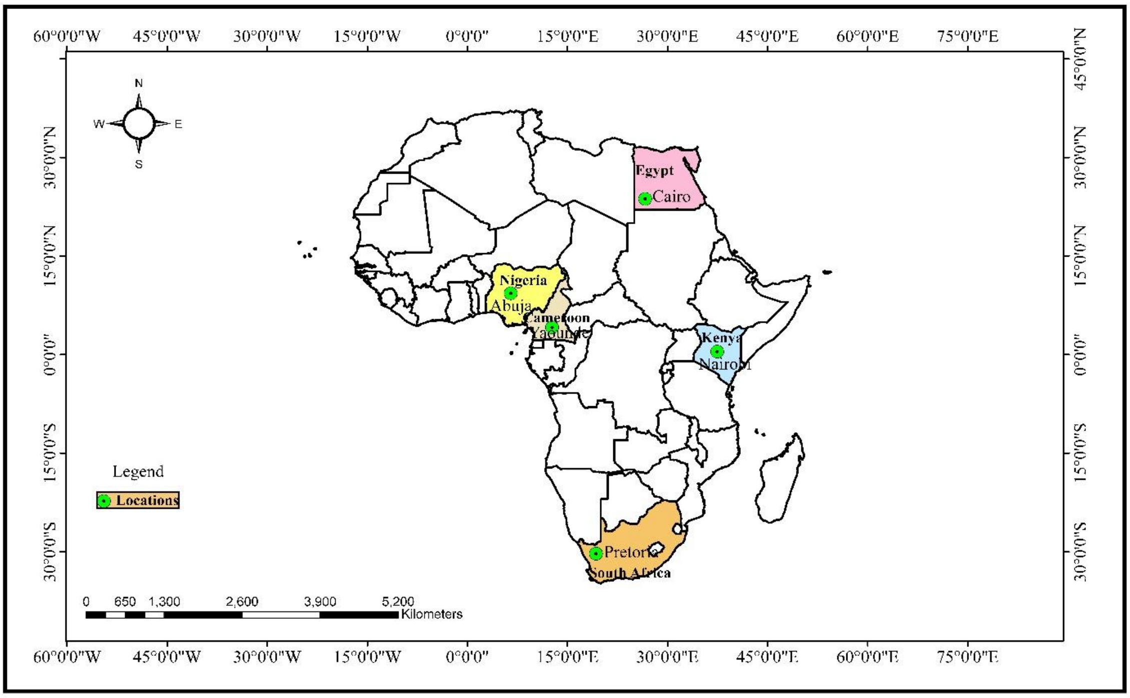


Figure 1. Map of African countries showing the studied locations

2.3 Statistical analysis of data for all stations considered

The data used in this study is monthly rainfall data between 1980 and 2022 (42 years) in five stations in Africa with one station drawn from each of the African regions: West (Abuja, Nigeria), East (Nairobi, Kenya), North (Cairo, Egypt), South (Pretoria, South Africa) and Central Africa (Yaoundé, Cameroun). For this study, the rainfall series of each of the African stations was subjected to a stationarity test by the use of Augmented Dickey-Fuller, where the p-value was less than 0.05 which indicates stationarity. The orders of the tentative SARIMA models were determined based on the plots of the Autocorrelation Function (ACF) as well as the Partial Autocorrelation Function (PACF). After these tentative SARIMA models were identified, their parameters were estimated using the Statistical Package for Social Sciences (SPSS version 20.0) while the goodness of fit of these models was determined using the Ljung-Box test and a p-value greater than 0.05 (p>.05) indicates a good fit. The plots of the ACF and PACF of the residuals were also used as a way of diagnosing the estimated SARIMA models. For forecasting the accuracy of these SARIMA models, the data set was divided into two that is; training which comprised 80% and 20% in testing [30].

The training data is from 31/01/1980 and 31/05/2014 while the validation data covers the period of 30/06/2014 to 31/12/2022. This is very important in order to see how the model performed on data not seen (out of sample data). The use of the later date data in testing is to see the performance of the data on the recent trend in rainfall so as to capture the changing dynamic of rainfall in these selected countries as a result of climate change. Data analyses were performed using the Statistical Package for Social Sciences (SPSS version 20.0) and Eview 9.0. The estimation of SARIMA model parameters was estimated on SPSS version 20.0 while a test of stationarity was done on Eview 7.0. The performance evaluation of the SARIMA models was carried out based on fitness performance measures (adjusted R^2, NBIC) and forecasting accuracy measures (RMSE) as defined below:

$$R^2_{adj.} = 1 - \left[\frac{(1-R^2)(m-1)}{m-p-1} \right] \quad (1)$$

where, $R^2_{adj.}$ is the adjusted coefficient of determination, *m* is the number of observations and *p* is the number of parameters in the model.

$$BIC(p) = m \ln\left(\frac{\hat{\sigma}_e^2}{m} \right) + p + p \ln(m) \quad (2)$$

In Eq. (2), BIC is the Bayesian Information Criteria.

$$RMSE = \frac{1}{n} \sqrt{\sum_{t=1}^{n} \left(R_t - \hat{R}_t \right)^2} \quad (3)$$

where, R_t and $\hat{R}_t$ are the actual and predicted rainfall.

The Ljung-Box test statistic for assessing the goodness of fit of the SARIMA models is given as:

$$Q^* = m(m+2) \sum_{i=1}^{w} \frac{r_i^2}{m-1} \quad (4)$$

where, m is the number of observations, w is the length of the series and r_i is the autocorrelation coefficients at lag i, $-1 \leq r_i \leq 1$.

The Root Mean Square Error (RMSE) as defined in Eq. (3) measures the mean of the squared difference between the actual and predicted rainfall while adjusted R^2 and NBIC are the statistical measures of the goodness of fit. These metrics were chosen to assess the performance of the SARIMA models.

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is given as:

$$\psi_p\left(B^S\right)\lambda_p\left(B\right)\left(1-B^S\right)^D\left(1-B\right)^d R_t = \theta_Q\left(B^S\right)\theta_q^*\left(B\right)\varepsilon_t \tag{5}$$

where,

$$\psi_p\left(B^S\right) = 1-\psi_1 B^s - \psi_2 B^s - \psi_3 B^s - \ldots - -\psi_p B^{ps} \tag{6}$$

$$\lambda_p\left(B\right) = 1-\lambda_1 B - \lambda_2 B - \lambda_3 B - \ldots - -\lambda_p B^p \tag{7}$$

$$\theta_Q\left(B^S\right) = 1+\theta_1 B^s + \theta_2 B^s + \theta_3 B^s + \ldots + \theta_Q B^{Qs} \tag{8}$$

$$\theta_q\left(B\right) = 1+\theta_1^* B + \theta_2^* B + \theta_3^* B + \ldots + \theta_q^* B^q \tag{9}$$

The confidence interval for the RMSE is given as:

$$\sqrt{\frac{n}{\chi^2_{1-\frac{\alpha}{2},n}}}RMSE < \sigma < \sqrt{\frac{n}{\chi^2_{\frac{\alpha}{2},n}}}RMSE \tag{10}$$

where, n is the number of observation and $\chi^2_{1-\frac{\alpha}{2}}$ and $\chi^2_{\frac{\alpha}{2}}$ are obtained from the Chi-square distribution with α level of significance.

3. RESULTS AND DISCUSSION

Table 2 presents the mean and coefficient of variation for rainfall in some of the cities in Africa and the results show that the peak of rainfall in Abuja is in August (536.11), in Cairo, it is in March (7.95) while in Nairobi (168.04), Pretoria (107.65) and Yaoundé (493.39) the peak of the rainfall is in April, January and October respectively. The dispersion around the mean was highest in December in Abuja (224.56%), March in Cairo (204.66%), February in Nairobi (114.54%), July in Pretoria (177.76%) and in January in Yaoundé (97.29%). The result also shows that Abuja which is the capital of Nigeria recorded the highest dispersion in rainfall compared with other major cities in Africa. This could be because Nigeria, in the west African region, is closer to the ocean than any other African country except South Africa. Furthermore, research has shown that Abuja has a high temperature, which could cause evaporation, condensation, and transpiration, followed by rainfall. However, Abuja is Nigeria's transition center between ocean breeze and Sahara breeze, giving it an advantage in terms of high rainfall. Furthermore, as reported in study [27], the higher the temperature, the more rainfall. The result also established that there was a higher rainfall in Yaoundé between March and June compared to other major cities in Africa while Abuja reported the highest rainfall in July and August than other African cities (Figure 2). In both Yaoundé and Abuja, a very significant downward trend in rainfall was observed towards the end of the year (September to December) as shown in Figure 2.

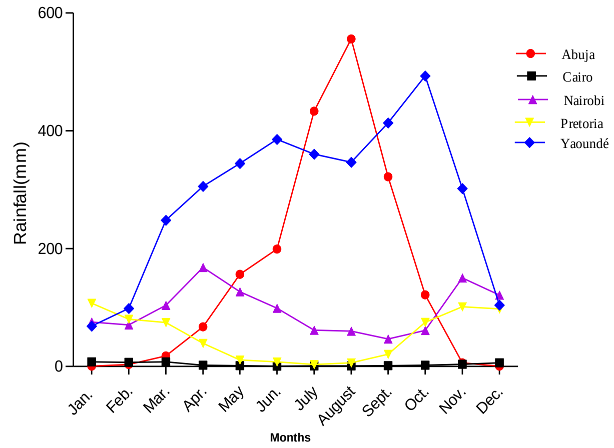


Figure 2. Mean monthly rainfall in the selected countries in Africa

Table 2. Descriptive statistics for the monthly rainfall in some selected countries in Africa

	Abuja		Cairo		Nairobi		Pretoria		Yaoundé	
Month	**Mean**	**COV (%)**	**Mean**	**COV (%)**	**Mean**	**COV (%)**	**Mean**	**COV (%)**	**Mean**	**COV (%)**
Jan.	0.62	193.09	7.92	78.87	75.48	111.85	107.65	62.31	68.61	97.29
Feb.	3.48	147.59	7.08	87.60	70.75	114.54	80.55	83.49	98.50	69.53
Mar.	18.06	77.71	7.95	204.66	103.66	109.37	74.63	88.61	248.53	41.15
Apr.	67.61	44.01	2.31	137.77	168.04	66.47	39.58	87.13	305.75	36.39
May	156.29	41.30	1.26	91.53	126.91	56.35	11.04	129.78	344.57	32.15
Jun.	199.21	39.17	0.70	75.33	99.14	69.96	7.73	147.08	385.81	28.42
July	433.49	31.71	0.85	41.88	61.82	54.64	3.56	177.76	360.66	26.61
August	556.11	27.18	0.98	56.31	60.29	46.84	6.13	135.08	346.90	28.56
Sept.	321.91	42.01	1.30	152.29	46.79	68.12	20.99	162.81	413.60	31.91
Oct.	121.64	63.58	2.04	177.86	61.16	49.47	75.13	85.35	493.39	30.52
Nov.	6.23	99.07	3.98	149.98	150.31	39.11	101.27	47.94	302.35	42.73
Dec.	0.60	224.56	6.43	127.26	121.28	69.50	97.76	66.21	104.13	71.04

Note: SD- standard deviation, COV- coefficient of variation

The result of the test of stationarity of the series was carried out using the Augmented Dickey-Fuller (ADF) test and the result obtained is presented in Table 3. The result reveals that rainfall in all selected African cities ($p<0.05$) was stationary at levels except for Yaoundé where stationarity was achieved after the first differencing (test statistic=-24.77805, p-value=0.0000, $p<0.05$). This implies that the order of differencing for Abuja, Cairo, Nairobi and Pretoria will be 0 while that of Yaoundé will be 1.

Different tentative models were identified based on the plots of the autocorrelation function (ACF) and partial autocorrelation (PACF) and the performances of these tentative models are presented in Table 3. The goodness of fits of each of these tentative models was examined using the Ljung box statistic and the result as presented in Table 4 reveals p-values greater than 0.05 ($p>0.05$) for each of the tentative models indicating that the error follows a white noise, implying that these models are all of the good fit ($p>0.05$).

Table 3. Augmented Dickey-Fuller (ADF) test for the rainfall series in the selected countries in Africa

Cities	At Level		After First Differencing	
	Test Statistics	p-Value	Test Statistics	p-Value
Abuja	-4.202271	0.0007**	-	-
Cairo	-19.13706	0.0000**	-	-
Nairobi	-17.33174	0.0000**	-	-
Pretoria	-4.874466	0.0000**	-	-
Yaoundé	-2.573407	0.0992	-24.77805	0.0000**

**Significant at 1% ($p<0.01$)

Table 4. Performance evaluation for tentative models for forecasting rainfall in selected cities in Africa

Cities	Time Series Models	R^2_{adj}	NBIC	RMSE	95% C.I	Ljung Box Test	
						Test Statistic	p-Value
Abuja	ARIMA (1,0,1) $(1,0,1)_{12}$	0.769	9.218	70.7999	63.8672-80.4853	17.795	0.2160
	ARIMA (0,0,1) $(1,0,1)_{12}$	0.769	9.203	**70.7044**	63.7810- 80.3768	18.546	0.2350
	ARIMA (1,0,0) $(1,0,1)_{12}$	**0.769**	**9.201**	70.7344	63.8081-80.4109	18.824	0.2480
	ARIMA (0,0,0) $(1,0,1)_{12}$	0.768	20.437	70.7099	63.7860-80.3830	30.437	0.201
Cairo Natural log	ARIMA (1,0,1) $(1,0,1)_{12}$	0.142	3.490	**9.9279**	8.9558-11.2860	12.994	0.527
	ARIMA (0,0,1) $(1,0,1)_{12}$	0.144	3.472	10.0345	9.0519-11.4072	14.984	0.453
	ARIMA (1,0,0) $(1,0,1)_{12}$	0.144	3.472	10.03874	9.0557-11.4120	14.484	0.489
	ARIMA (0,0,0) $(1,0,1)_{12}$	**0.146**	**3.453**	10.03029	9.0481-11.4024	19.032	0.267
Nairobi	ARIMA (1,0,1) $(1,0,1)_{12}$	0.187	8.518	93.0889	83.9736-105.8235	12.640	0.555
	ARIMA (1,0,0) $(1,0,1)_{12}$	**0.189**	**8.501**	93.2667	84.1340-106.0256	12.872	0.612
	ARIMA (0,0,1) $(1,0,1)_{12}$	0.189	8.502	**92.8925**	83.7965-105.6002	12.616	0.632
Pretoria	ARIMA (1,0,1) $(1,0,1)_{12}$	0.361	7.895	42.06133	37.9427- 47.8153	17.549	0.228
	ARIMA (1,0,0) $(1,0,1)_{12}$	**0.367**	**7.886**	42.06061	37.9420-47.8145	22.224	0.102
	ARIMA (2,0,0) $(1,0,1)_{12}$	0.362	7.893	**42.05462**	37.9366-47.8077	16.883	0.265
Yaoundé	ARIMA (1,1,1) $(1,0,1)_{12}$	**0.641**	**9.349**	**79.42084**	71.6440-90.2856	18.352	0.191
	ARIMA (2,1,1) $(1,0,1)_{12}$	0.640	9.367	79.42085	71.6440-90.2856	18.076	0.155
	ARIMA (1,1,2) $(1,0,1)_{12}$	0.640	9.368	80.62303	72.7284-91.6523	18.076	0.155

Note: Bolded values are the least normalized Bayesian Information Criteria and highest adjusted R^2. NBIC- Normalized Bayesian Information Criteria.

Table 5. A summary estimate of the best model for each of the selected cities in Africa

Cities	Best Models	Constant	AR 1 (SE)	AR 2 (SE)	MA (1) (SE)	SAR (1) (SE)	SMA (1) (SE)
Abuja	ARIMA (0,0,1) $(1,0,1)_{12}$	161.029** (53.885)	-	-	-.020 (.050)	.999** (.001)	.905** (.031)
Cairo (Natural log)	ARIMA (1,0,1) $(1,0,1)_{12}$	.509** (.253)	.534* (.0415)	-	0.435 (.435)	1.000** (0.002)	0.975** (0.045)
Nairobi	ARIMA (0,0,1) $(1,0,1)_{12}$	88.140** (12.127)	-	-	-.161 (.049)	.995** (.010)	. .952** (.052)
Pretoria	ARIMA (2,0,0) $(1,0,1)_{12}$	52.279** (12.700)	.160** (.0440)	.071 (.044)	-	.999** (.003)	.976** (.053)
Yaoundé	ARIMA (1,1,1) $(1,0,1)_{12}$	-.355** (.131)	.235** (.049)	-	1.000** (0.036)	.995** (.003)	.882** (0.033)

*Significant at 5% ($p<.05$), **significant at 1% ($p<0.01$).

The performances of these models were examined using the adjusted R^2, normalized BIC and Root Mean Square Error (RMSE). In terms of fitness performance, ARIMA (1,0,0) $(1,0,1)_{12}$ ($R^2_{adj.}$=0.769, NBIC=9.201) outperformed other models for Abuja, in Cairo, it was ARIMA (0,0,0) $(1,0,1)_{12}$ ($R^2_{adj.}$=0.146, NBIC=3.453), in Nairobi, it was ARIMA (1,0,0) $(1,0,1)_{12}$ ($R^2_{adj.}$=0.189, NBIC=8.501), in Pretoria, it was ARIMA (1,0,0) $(1,0,1)_{12}$ ($R^2_{adj.}$=0.367, NBIC=7.886) while in Yaoundé, ARIMA (1,1,1) $(1,0,1)_{12}$ ($R^2_{adj.}$=0.641, NBIC=9.349) outperformed other seasonal ARIMA models

(Table 4). In terms of out-of-sample forecasting performance, ARIMA (1,0,0) $(1,0,1)_{12}$ (RMSE=70.7044), ARIMA (1,0,1) $(1,0,1)_{12}$ (RMSE=9.9279), ARIMA (0,0,1) $(1,0,1)_{12}$ (RMSE=92.8925), ARIMA (2,0,0) $(1,0,1)_{12}$ (RMSE=42.05462) and ARIMA (1,1,1) $(1,0,1)_{12}$ (RMSE=79.42084) were the best forecasting seasonal model for rainfall in Abuja, Cairo, Nairobi, Pretoria and Yaoundé respectively. All the fitted SARIMA models were found to be of good fit (Ljung box p-value >.05) and this is also substantiated by the ACF and PACF plots of the residuals as shown in Figures 3-7. The plots of the ACF and PACF of the residuals as presented in Figures 3-7 indicate that the models were of good fit as they were no significant spikes capable of undermining the goodness of fit of these models.

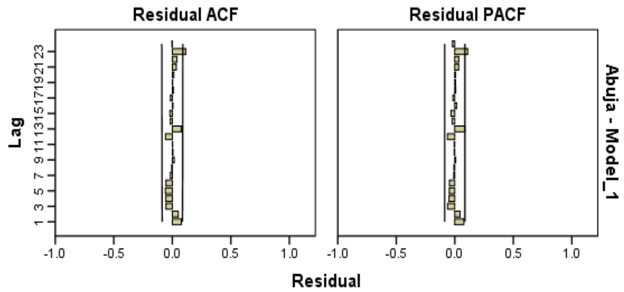


Figure 3. Plot of the ACF and PACF of the residuals for ARIMA$(0,0,1)(1,0,1)_{12}$- Abuja

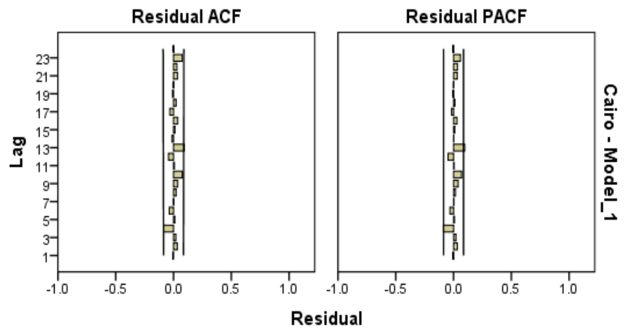


Figure 4. Plot of the ACF and PACF of the residuals for ARIMA$(1,0,1)(1,0,1)_{12}$- Cairo

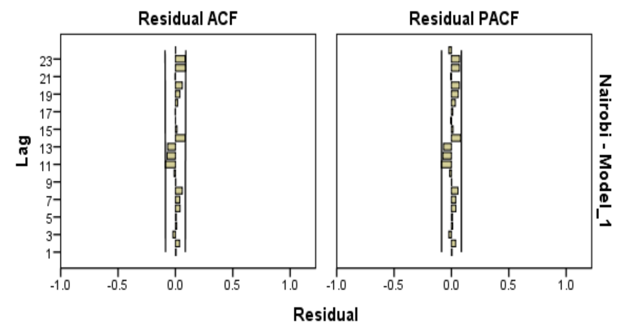


Figure 5. Plot of ACF and PACF of the residuals for ARIMA$(0,0,1)(1,0,1)_{12}$- Nairobi

The parameter estimates of the best forecasting model are as presented in Table 5. The result shows that the seasonal terms were statistically significant in all models which justified the use of seasonal ARIMA models in modelling rainfall in these selected cities in Africa. This also underscored the significant role of the season in the rainfall pattern in the selected African cities. In Cairo, Pretoria and Yaoundé, the AR (1) terms were significant but was more positively significant in Cairo than other locations. This implies that the previous month's rainfall has a positive influence on the present month's rainfall in these locations. For Abuja and Nairobi, there were no AR (1) terms in the optimal model (Table 5).

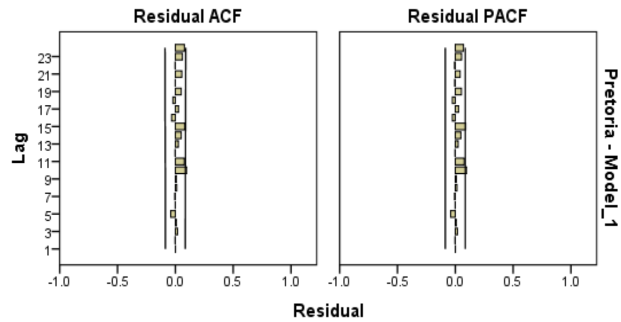


Figure 6. Plot of ACF and PACF of the residuals for ARIMA$(2,0,0)(1,0,1)_{12}$- Pretoria

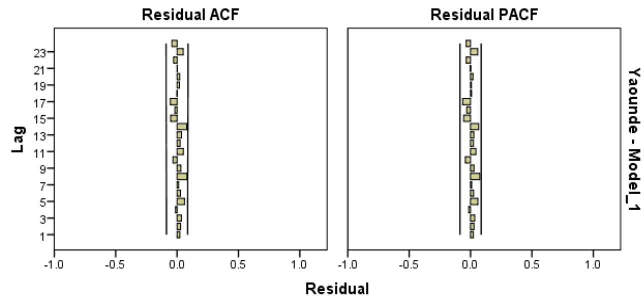


Figure 7. Plot of ACF and PACF of the residuals for ARIMA$(2,0,0)(1,0,1)_{12}$- Yaounde

The SARIMA models for forecasting rainfall in each of the selected African stations are presented below:

For Abuja:

$$\begin{aligned}&\left(1-0.999B^{12}\right)R_t = 161.029\\&+\left(1+0.905B^{12}\right)\left(1-0.020B\right)\varepsilon_t\end{aligned} \tag{11}$$

For Cairo:

$$\begin{aligned}&\left(1-B^{12}\right)\left(1-0.534B\right)R_t^*\\&=0.5090+\left(1+0.975B^{12}\right)\left(1+0.435B\right)\varepsilon_t\end{aligned} \tag{12}$$

where, $R_t^* = ln(R_t)$.

For Nairobi:

$$\begin{aligned}&\left(1-0.995B^{12}\right)R_t = 88.140\\&+\left(1+0.952B^{12}\right)\left(1-0.161B\right)\varepsilon_t\end{aligned} \tag{13}$$

For Pretoria:

$$\begin{aligned}&\left(1-0.999B^{12}\right)\left(1-0.160B-0.071B^2\right)R_t\\&=52.279+\left(1+0.976B^{12}\right)\varepsilon_t\end{aligned} \tag{14}$$

For Yaounde:

$$\left(1-0.995B^{12}\right)\left(1-B^{12}\right)\left(1-0.235B\right)R_t = \\ -0.355+\left(1+0.882B^{12}\right)\left(1+B\right)\varepsilon_t \tag{15}$$

The forecast of the rainfall in these selected cities is presented in Table 6 and Figure 8. Based on the one-year prediction of rainfall for 2023, the peak of the rainfall is predicted to happen in July in Abuja (531.54), January in Cairo (6.74), April in Nairobi (149.85), January in Pretoria (106.53) and October in Yaoundé (377.66). A downward trend in rainfall was also predicted towards the end of the year in both Abuja and Yaoundé (Figure 8).

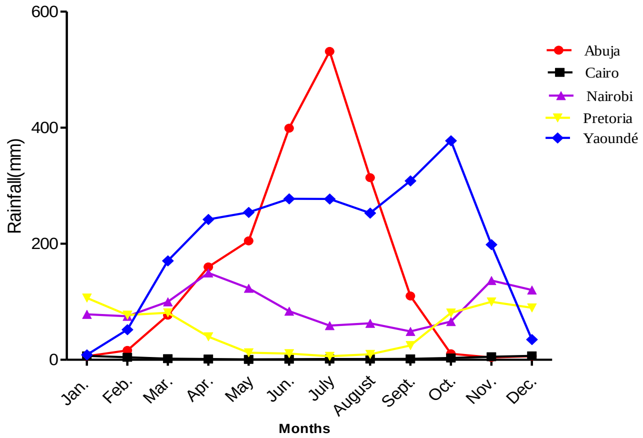
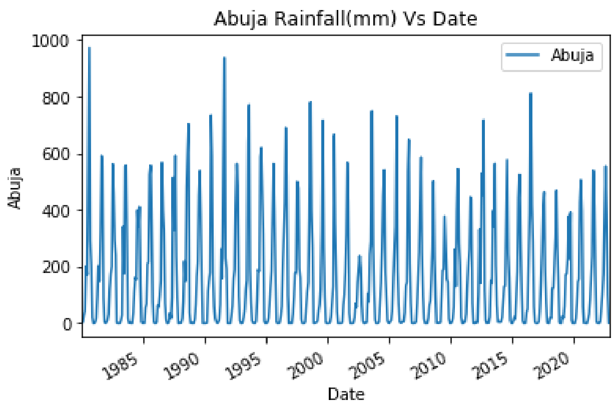


Figure 8. Forecast of the rainfall in some selected major cities in Africa (January 2023- December 2023)

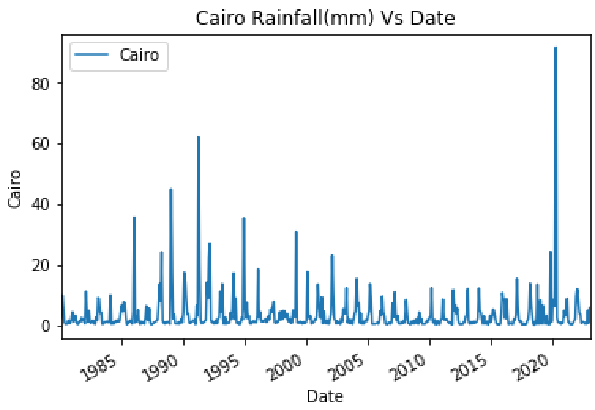
Table 6. Summary of the forecast of rainfall from January 2023 to December 2023 in the selected cities in Africa

Month	Abuja	Cairo	Nairobi	Pretoria	Yaoundé
Jan.	6.08	6.74	78.52	106.53	8.75
Feb.	16.30	4.38	75.03	77.31	52.04
Mar.	76.81	2.05	100.00	81.04	170.77
Apr.	160.01	1.41	149.85	39.78	241.91
May	204.90	0.88	123.40	12.26	254.28
Jun.	399.17	1.18	83.82	10.82	277.55
July	531.54	1.35	59.03	6.08	277.38
August	313.97	1.38	63.07	9.48	252.92
Sept.	109.96	1.57	49.10	24.79	308.52
Oct.	10.43	3.08	66.09	80.93	377.66
Nov.	4.14	5.41	136.79	99.98	198.63
Dec.	6.08	6.74	120.37	89.67	35.13

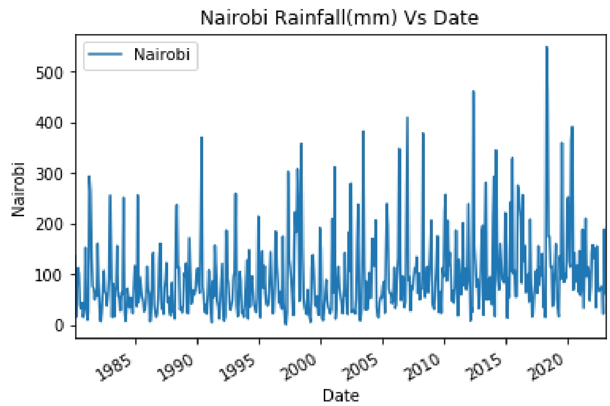
Figure 9 depicts the seasonal variation of rainfall across the African stations studied. It was discovered that all of the stations follow a similar pattern. The results show that Abuja (Figure 9(a)) has deep and evenly distributed rainfall, whereas Cairo (Figure 9(b)) has less rainfall. It was discovered that Abuja (Figure 9(a)) had the highest rainfall in 1980 with 950 mm, and the lowest rainfall was recorded in 2004 with 250 mm. The results also revealed that Cairo (Figure 9(b)) is in Northern Africa with less rainfall than Abuja (Figure 9(a)), with the highest rainfall in 2021 being 75 mm, which is far less than what was experienced in Abuja (Figure 9(a)) compared to the minimum value observed. However, Pretoria (Figure 9(d)) in South Africa receives more rainfall due to its proximity to the ocean, which causes stations in the south to receive more rainfall than stations in the north. However, when compared to Abuja (Figure 9(a)), Pretoria (Figure 9(d)) has less rainfall, with the highest recorded in 2009 at 340 mm and the lowest recorded in 2004 at 45 mm. This refutes the notion that the station in the south should receive more rainfall than the station in the north. The results revealed that Nairobi (Figure 9(c)) has a year-round rainfall pattern, indicating that the rainfall in the station is beneficial to the farmer's planting season. The highest value of rainfall observed in the station is 550 mm in 2020, and the lowest value observed in the station is 100 mm in 1980. Rainfall in Yaoundé (Figure 9(e)) decreases in a sinusoidal pattern throughout the study year. The highest amount of rainfall recorded in Yaoundé (Figure 9(e)) was 870 mm in 1985. However, the seasonal variation of rainfall results shows that Abuja (Figure 9(a)) has the highest rainfall of any African station. As a result of the high rainfall, Abuja (Figure 9(a)) will most likely experience flooding in the coming years.



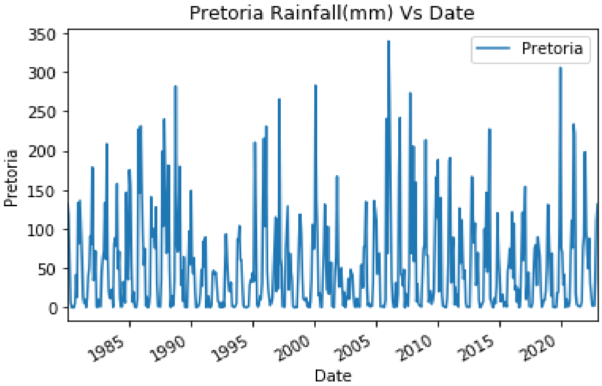
(a)



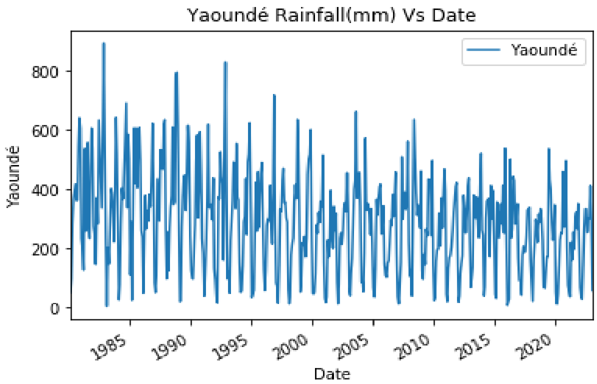
(b)



(c)



(d)



(e)

Figure 9. Seasonal variation of rainfall across the selected African station

This finding has shown that rainfall in most of the African stations considered were stationary at the level which is consistent with that of the study [2] on modelling and forecasting of selected meteorological parameters for the environmental awareness in Sub- Sahel Africa of which rainfall is one of the meteorological parameters considered. The finding also revealed that in most of the fitted models, the previous month's rainfall was found to have a significant impact on the present month's rainfall which is also corroborated [2]. Furthermore, the finding showed a significant effect of the seasonal terms showing evidence of seasonal factors in rainfall in the studied station as reported in other studies on rainfall forecasting [31, 32].

This study based on the forecast rainfall in the different African cities indicate show uniqueness in the peak of rainfall in the different cities (Figure 8). In Cairo, the peak of the rainfall was predicted to be in January and December while in Nairobi, it was predicted to be in April. For Pretoria, Yaoundé and Abuja, the peak of rainfall was predicted to occur in January, October and July respectively. This work contributes to existing literature by providing insights on the trend and future values of rainfall in five major cities that cut across the different regions of Africa. This is very different from other studies as it provides a more informative nature of rainfall that take care of the different regions of Africa: West Africa, North Africa, East Africa, South Africa and Central Africa. This work has provided an all-inclusive modelling of rainfall that reflect the peculiarities of the different sub-regions in Africa. This study has also provided insights on when rainfall is expected to be at the peak in each of these cities. This information, if properly harnessed, could help stimulate better agricultural production in Africa as agriculture is the backbone of Africa economy. Also, these predictions could help address some of the challenges associated with excessive rainfall such as flooding which is also one of the problems in Africa. This could help in formulating policies that would help avert the negative consequences of flooding, thereby reinforce the need for a more proactive measures that can help curb flooding in Africa.

4. CONCLUSIONS

Rainfall data from the MERRA-2 satellite archive was analyzed statistically and environmentally from 1980 to 2022. The collected data were analyzed to determine the rainfall trend and forecast future occurrences, as well as to provide information to farmers and governments in each of the African stations considered. The study's findings show that there is a variation in the pattern of rainfall in the stations studied, indicating that rainfall is said to be conventional. The performances of the models were examined using the adjusted R^2, normalized BIC and Root Mean Square Error (RMSE). Based on the result of out of sample forecast performance, different SARIMA models were found to be suitable for rainfall in each of the station with ARIMA (0,0,1) $(1,0,1)_{12}$ for Abuja (RMSE=70.7044) and Nairobi (RMSE=92.8925), ARIMA (1,0,1) $(1,0,1)_{12}$ for Cairo (RMSE=9.9279), ARIMA (2,0,0) $(1,0,1)_{12}$ for Pretoria (RMSE=42.05462) and ARIMA (1,1,1) $(1,0,1)_{12}$ for Yaoundé (RMSE=79.42084). As a result, the forecast model used indicates that maximum rainfall is observed in Abuja when compared to other stations.

5. RECOMMENDATION

The authors wish to recommend to the water management cooperation under the leadership of the federal republic of Nigeria as well as other stations considered to make available water storage tank so that agriculturists could take the period of low rainfall in their country for their plantation through the medium of irrigation so that there will be all round agricultural products in each of the countries. This is also applicable to other African countries, who should take note and do what is necessary for their countries. It is therefore recommended that more research centers should be built for each country in Africa in order to reduce the continent's food shortage.

ACKNOWLEDGMENT

Based on the data used in this study, the authors express their gratitude to the MERRA-2 satellite for providing data. We are also grateful to Bowen University's management for providing us with the opportunity to conduct the research.

REFERENCES

[1] Ramli, I., Rusdiana, S., Achmad, A., Yolanda, M.E. (2023). Forecasting of rainfall using seasonal autoregreressive integrated moving average (SARIMA) aceh, Indonesia. Mathematical Modelling of Engineering Problems, 10(2): 501-508. https://doi.org/10.18280/mmep.100216

[2] Aweda, F.O., Akinpelu, J.A., Samson, T.K., Sanni, M.,

Olatinwo, B.S. (2022). Modeling and forecasting selected meteorological parameters for the environmental awareness in Sub-Sahel West Africa Stations. Journal of the Nigerian Society of Physical Sciences, 4(3): 820-820. https://doi.org/10.46481/jnsps.2022.820

[3] Angles, S., Chinnadurai, M., Sundar, A. (2011). Awareness on impact of climate change on dryland agriculture and coping mechanisms of dryland farmers. Indian Journal of Agricultural Economics, 66(3): 365-372. https://doi.org/10.22004/ag.econ.204760

[4] Smith, J.A. (1993). Precipitation in Handbook of Hydrology. New York: Mcgraw-Hill, Inc., pp. 3-47.

[5] Hyndman, R.J., Athanasopoulos, G. (2018). Forecasting: Principles and Practice. OTexts.

[6] Aweda, F.O., Adebayo, S., Adeniji, A.A., Samson, T.K., Akinpelu, J.A. (2023). Investigation of wind speed to generate energy using machine learning algorithms approach over selected Nigerian Stations. Journal of Renewable Energy and Environment, 10(3): 81-98. https://doi.org/10.30501/jree.2022.354698.1422

[7] Ramli, I., Rusdiana, S., Basri, H., Munawar, A. (2019). Predicted rainfall and discharge using vector autoregressive models in water resources management in the high hill takengon. In IOP Conference Series: Earth and Environmental Science, 273(1): 012009. https://doi.org/10.1088/1755-1315/273/1/012009

[8] Brockwell, P.J., Davis, R.A. (Eds.). (2002). Introduction to Time Series and Forecasting. New York, NY: Springer New York. https://doi.org/10.1007/0-387-21657-X_8

[9] Montgomery, D.C., Jennings, C.L., Kulahci, M. (2015). Introduction to Time Series Analysis and Forecasting. John Wiley & Sons.

[10] Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M. (2015). Time Series Analysis: Forecasting and Control. John Wiley & Sons.

[11] Heizer, J., Render, B., Munson, C. (2020). Operations management: Sustainability and supply chain management. Pearson. Binus University.

[12] Suhartono, S. (2011). Time series forecasting by using seasonal autoregressive integrated moving average: Subset, multiplicative or additive model. Journal of Mathematics and Statistics, 7: 20-27. http://doi.org/10.3844/jmssp.2011.20.27

[13] Wang, S., Feng, J., Liu, G. (2013). Application of seasonal time series model in the precipitation forecast. Mathematical and Computer modelling, 58(3-4): 677-683. https://doi.org/10.1016/j.mcm.2011.10.034

[14] Salas, J.D., Delleur, J.R., Yevjevich, V., Lane, W.L. (1980). Applied Modeling of Hydrologic Time Series, Water Resor. Pub., Littleton, CO, USA.

[15] Salas, J.D., Tabios III, G.Q., Bartolini, P. (1985). Approaches to multivariate modeling of water resources time series 1. Journal of the American Water Resources Association, 21(4): 683-708. https://doi.org/10.1111/j.1752-1688.1985.tb05383.x

[16] Subbaiah Naidu, K.C.H.V. (2016). SARIMA modeling and forecasting of seasonal rainfall patterns in India. International Journal of Mathematics Trends and Technology, 38(1): 15-22. http://doi.org/10.14445/22315373/IJMTT-V38P504

[17] Chen, P., Niu, A., Liu, D., Jiang, W., Ma, B. (2018). Time series forecasting of temperatures using SARIMA: An example from Nanjing. In IOP Conference Series: Materials Science and Engineering, 394: 052024. https://doi.org/10.1088/1757-899X/394/5/052024

[18] Liu, H., Li, C.X., Shao, Y.Q., Zhang, X., Zhai, Z., Wang, X., Qi, X.Y., Wang, J.H., Hao, Y.H., Wu, Q.H., Jiao, M.L. (2020). Forecast of the trend in incidence of acute hemorrhagic conjunctivitis in China from 2011-2019 using the Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ETS) models. Journal of Infection and Public Health, 13(2): 287-294. https://doi.org/10.1016/j.jiph.2019.12.008

[19] Dindarloo, S., Hower, J.C., Bagherieh, A., Trimble, A.S. (2016). Fundamental evaluation of petrographic effects on coal grindability by seasonal autoregressive integrated moving average (SARIMA). International Journal of Mineral Processing, 154: 94-99. https://doi.org/10.1016/j.minpro.2016.07.005

[20] El-Mallah, E., Elsharkawy, S. (2016). Time-series modeling and short term prediction of annual temperature trend on Coast Libya using the box-Jenkins ARIMA Model. Advances in Research, 6(5): 1-11.

[21] Yusof, F., Kane, I. L. (2012). Modelling monthly rainfall time series using ETS state space and SARIMA models. International Journal of Current Research, 4(9): 195-200.

[22] Jeong, K., Koo, C., Hong, T. (2014). An estimation model for determining the annual energy cost budget in educational facilities using SARIMA (seasonal autoregressive integrated moving average) and ANN (artificial neural network). Energy, 71: 71-79. https://doi.org/10.1016/j.energy.2014.04.027

[23] Afrifa-Yamoah, E., Saeed, B.I., Karim, A. (2016). Sarima modelling and forecasting of monthly rainfall in the Brong Ahafo Region of Ghana. World Environment, 6(1): 1-9. https://doi.org/10.5923/j.env.20160601.01

[24] Sampson, W., Suleman, N., Gifty, A. (2013). Proposed seasonal autoregressive integrated moving average model for forecasting rainfall pattern in the Navrongo municipality of Ghana. Journal of Environment and Earth Science, 3(12): 80-85.

[25] Papalaskaris, T., Panagiotidis, T., Pantrakis, A. (2016). Stochastic monthly rainfall time series analysis, modeling and forecasting in Kavala City, Greece, North-Eastern Mediterranean Basin. Procedia Engineering, 162: 254-263. https://doi.org/10.1016/j.proeng.2016.11.054

[26] Mohamed, T.M., Ibrahim, A.A. (2016). Time series analysis of Nyala rainfall using ARIMA method. SUST Journal of Engineering and Computer Science (JECS), 17(1): 5-11.

[27] Aweda, F.O., Samson, T.K. (2022). Relationship between air temperature and rainfall variability of selected stations in Sub-Sahara Africa. Iranian (Iranica) Journal of Energy & Environment, 13: 248-257. http://doi.org/10.5829/ijee.2022.13.03.05

[28] Gelaro, R., McCarty, W., Suárez, M.J., et al. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). Journal of Climate, 30(14): 5419-5454. https://doi.org/10.1175/JCLI-D-16-0758.1

[29] Aweda, F.O., Adeniji, A.A., Akinpelu, J.A., Abiodun, A.J. (2021). Analysis of rainfall trends and variabilities for three decades in Sub-Sahara Africa. Ruhuna Journal of Science, 12(1): 100. http://doi.org/10.4038/rjs.v12i1.100

[30] Eni, D. (2015). Seasonal ARIMA modeling and forecasting of rainfall in Warri Town, Nigeria. Journal of

Geoscience and Environment Protection, 3(6): 91. http://doi.org/10.4236/gep.2015.36015
[31] Farajzadeh, J., Fard, A.F., Lotfi, S. (2014). Modeling of monthly rainfall and runoff of Urmia Lake basin using "feed-forward neural network" and "time series analysis" model. Water Resources and Industry, 7: 38-48. https://doi.org/10.1016/j.wri.2014.10.003
[32] Joseph, V.R. (2022). Optimal ratio for data splitting. Statistical Analysis and Data Mining: The ASA Data Science Journal, 15: 531-538. https://doi.org/10.1002/sam.11583