Faculty of Natural and Applied Sciences Journal of Scientific Innovations
Print ISSN: 2814-0877 e-ISSN: 2814-0923
www.fnasjournals.com
Volume 5; Issue 2; December 2023; Page No. 55-68.

STATISTICAL MODELLING OF TEMPERATURE AND RAINFALL IN IJEBU ODE NIGERIA USING SARIMA

Adewole, A.I.
Department of Mathematics, Tai Solarin University of Education Ijagun Ogun State Nigeria.

***Corresponding author email:** hayorhade2005@gmail.com

Abstract
Rainfall and Temperature remain natural climatic phenomena that are essential in the economic and agricultural productivity of developing nations. The predictions of these climatic phenomena were known to be challenging tasks due to their chaotic and complex natures. This research analyzed and modelled the seasonal autoregressive integrated moving average (SARIMA) of Annual Minimum, Maximum Temperature and Rainfall in the southwestern part of Nigeria using Ijebu Ode City as a case study with annual time series of the variables ranging from 1989- 2018. The study presented the methodology of the ARIMA model that integrates the seasonality of the series. Tentative numbers of SARIMA models were proposed for the variables built on the visualization of autocorrelation and partial autocorrelation functions of the series, a seasonal ARIMA $(4,1,1)(1,1,1)_{12}$, ARIMA $(1,1,1)(1,1,1)_{12}$, and ARIMA $(2,1,1)(0,1,1)_{12}$ for maximum Temperature, minimum Temperature and rainfall respectively were adopted using the information selection criterion. The result of the forecast models shows that there is a tendency for an increasing pattern of annual rainfall and temperature over the forecast period from year 2019 to year 2028. The developed model can be of assistance in planning likely future strategies associated with the weather conditions of Ijebu Ode City and its immediate environment.

Keywords: SARIMA, Box Jerkins, Climatic Data, Model Diagnostics and Forecast

Introduction
Climatic change can be categorized as one of the major environmental challenges confronted by the world today. It is associated with various undesirable effects on water resources, farming increase in temperature levels, rise in the ocean level, flood, drought cycles, forests and biodiversity (Chung & Lee, 2011). Climatic fluctuations are mostly ruled by the deviations that occur with natural phenomena such as rainfall and the weather warmth temperature are important climatic variable that plays a devastating effect on the atmospheric and socioeconomic circumstances of developing countries part of the world through extreme weather events. Rainfall and Temperature remain typical climatic phenomena whose forecasts can be inspiring and demanding due to their nature (Nirmala & Sundaram, 2010). Variability in time and space contribute to difficulties encountered in forecasting rainfall and temperature, moreover, the inability to access all the parameters that influence the rainfall and changes in the temperature of a locality also make climatic predictions a task. Climate forecast is of great significance to the economic development of a region in terms of agriculture, watershed management, human health and management. Time series predicting is the applicability of a specific model to predict upcoming values built on the information of past observed values (Raicharoen et al.,2003). Numerical weather predictions utilize the use of atmospheric models to forecast imminent weather conditions based on existing weather conditions (Barker et al.,2012; Shan et al., 2016; Xu et al., 2016).

Researchers have developed numerous time series models in the literature to advance the efficacy and precision of time series modelling and forecasting. The Autoregressive Integrated Moving Average (ARIMA) model introduced by Box and Jenkins (1970) is among the most broadly used and statistically recognized procedures of time series model forecasting, it is characterized by the advantage of simple applications for distinguished forecasting exactness and accordingly requires only the endogenous variables without necessarily in need for other exogenous variables (Xiaojuan et al., 2007) Autoregressive Integrated Moving Average (ARIMA), efficiently deliberate serial linear correlation amongst observations, whereas Seasonal Autoregressive Integrated Moving Average (SARIMA) models can acceptably define time series that display non-stationary behaviours within and across seasons (Box & Jerkins 1976). Box and Jerkins designed an expansion ARIMA model named Seasonal Autoregressive Integrated Moving Average (SARIMA) models for modelling univariate times series data characterized with seasonal components. (Udo

& Shittu, 2022).

Several studies have been conducted on modelling climate change in Nigeria and across the globe in general. Datong and Goltong (2017) fit a SARIMA model for weather conditions of Heipang Airport- Jos Plateau in Nigeria by developing models that involved monthly records of five years (2011-2016). The records consist of the average weather warmth and rainfall and their wind velocity. Chen et al., (2018) modelled time series prediction of temperatures using SARIMA in Shathir et al., (2019) worked on forecasting monthly maximum temperatures in Kerbala using SARIMA, Chukwudike et al. (2020) studied comparison of SARIMA and seasonal artificial neural network (SANN) that was used for prediction of the weather forecast in Umuahia, their results revealed SARIMA method of modelling and forecasting outperformed its contemporary in terms of minimal forecast errors. Ubaka et al. (2021) study was practically focused on the use of optimal stochastic models for predicting rainfall in selected states of the South-Western part of Nigeria. The work modeled the forecast of the pattern of the rainfall stochastically for the selected part of the region in Nigeria with the aid of SARIMA and ARFIMA (fractional autoregressive integrated moving average) models. Their results also proved the efficacy of seasonal ARIMA models in forecasting rainfall in the investigated cities with output of optimal stochastic models

Numerous researchers such as Murat et al. (2018), Jibril and Sanusi (2019), Adams and Bamaga (2020); Nnoka et al. (2020), Amjad et al. (2023) among others worked on modelling meteorological variables in different places using time series analysis. It is noteworthy to know that ARIMA/SARIMA is not only limited to model meteorological variables but cut across all times series data, researchers have explored the applicability of ARIMA/SARIMA in other fields such as Kahaun and Megersa (2017) (agricultural sectors). Wiri and Essi (2018) (economic sectors), and Adewole et al. (2023) (health sectors) among others.

This research work aims to present an iterative method of analysing, and modelling the seasonal ARIMA model for predicting the average annual minimum temperature, maximum temperature and rainfall in southwest of Nigeria using Ijebu Ode City as a case study. This was chosen with a view of having more insight into the impact of these variables on climate variation in our immediate environment. The study area was chosen due to its peculiarity in extreme weather event occurrences and its location within the transitional zone between the Precambrian basement rocks of southwestern Nigeria and the Nigerian part of the Dahomey basin (Osinowo 2013). Hence this research is essential in providing adequate information needed for meteorologists, agriculturists and climatologist decision-makers in their future planning in Ijebu Ode and its immediate environment. Data used in the study is annual data for 30 years (2019 -2018) including rainfall data, minimum temperature and maximum temperature obtained from the Nigeria Meteorological Department. The iterative procedures of Box Jerkins for seasonal ARIMA

Material and Methods

The study area is Ijebu Ode City, a town in Ogun state, southwestern geographical zone in Nigeria, it is situated 110 km by road Northwest of Lagos and 100 km from the Atlantic Ocean in the eastern part of Ogun state. The total area covers an area extent of approximately 190 km^2 located between longitude $3^0 52^!$ E of $3^0 57^!$ E and latitude $6^0 49^!$ N to $6^0 51^!$ N. The population of the inhabitants is estimated to be about 620,000. The community and its environment are confronted with many challenges due to their geological settings; it has a weather condition of mild and winter also with hot summer. The annual average temperature and rainfall data covering the period of 1989 – 2018 obtained from NIMET (Nigeria Metrological Agency) data management unit was used for the analysis. The analysis adopted the procedure of the Box–Jenkins method of ARIMA (autoregressive integrated moving average) modelling.

The ARIMA modelling constitutes three main parts which include; autoregressive, integrated (I) and moving average. The autoregressive part represents the autocorrelation between current and previous data, In contrast, the moving average operator narrates the autocorrelation framework of the residuals in the model. Reports from HASMIDA [2009] works revealed that the integrated (I) part indicates the number of differences needed to achieve a stationarity series from non-stationarity data... The ARIMA model usually takes the form (*p*, *d*, *q*) where the p signifies the AR fragment of the model which represents no. of lag observations in the model (known as the lag order), d signifies the level of appropriate differencing and q describes the size of moving average section. the MA operator in the ARIMA model is expressed as;

$$\emptyset(U) = 1 - \emptyset_1 U - \emptyset_2 U^2 - \cdots, -\emptyset_q U^q \quad (1)$$

q is the size of the MA operator $\emptyset_p, p = 1, 2, \dots, q$ is the Moving Average parameters and M is the backward shift operator in such a way that

$$UY_t = Y_{t-1} \quad (2)$$

The Autoregressive operator is expressed in the form of;

$$\theta(U) = 1 - \theta_1 U - \theta_2 U^2 - \cdots, -\emptyset_p U^p \quad (3)$$

p denotes the order of Autoregressive operators and θ_i is the non-seasonal AR parameters, $i = 1, 2, \dots, n$
ARIMA model defined for an average data size can be defined in the form of

$$Y = (Y_1, Y_2, \dots, Y_n)$$

written as;

$$\theta(U)(1 - U)^d (Y_t) = \theta(U) e_t \quad (4)$$

d represents the level of modifications; t denotes the definite time and e_t denotes the residual. ARIMA model can be generalized and defined as;

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} +, \dots, \theta_p X_{t-p} + \epsilon_t - \emptyset_1 \epsilon_{t-1} - \emptyset_2 \epsilon_{t-2} -, \dots, \emptyset_q \epsilon_{t-q} \quad (5)$$

$$X_t = Y_t - Y_{t-d} \quad (6)$$

θ_p is the autoregressive parameter; ϵ_t denotes the white noise or residual while $\emptyset_q$ is the moving average parameter and Y_t is the dependent variable; X_t is the dth difference of the dependable variable.

Seasonal Autoregressive Integrated Moving Average (SARIMA*)*
A seasonal ARIMA model is generated by including additional seasonal terms to the ARIMA model, generally, it is expressed as;
SARIMA (p, d, q) (P, D, Q)-

The seasonal component terms of the model are related to the non-seasonal component but with a difference of back shift during the season

$$\theta_p(U)\theta_p(U^s)(1 - U)^d (1 - U^s)^D X_t = \emptyset_q(U)\emptyset_q(U^s)\epsilon_t \quad (7)$$

Where $\theta_p(U) = 1 - \theta_1 U - \theta_2 U^2 -, \dots, \theta_p U^p$ (8)

$$\theta_p(U^s) = 1 - \theta_p U^s, \dots, \theta_p U^{ps} \quad (9)$$

$$\emptyset_q(U) = 1 + \emptyset_1 U + \theta_2 U^2 -, \dots, +\emptyset_q U^q \quad (10)$$

$$\emptyset_q((U^s) = 1 + \emptyset_1 U^s + \theta_2 U^2 -, \dots, +\emptyset_p U^{qs} \quad (11)$$

X_t represent the actual data; θ_p and $\emptyset_q$ are the coefficients of AR component and the MA component Coefficient correspondingly; c is the constant rate; μ is the average value of the series and ε is the white noise. U represents the non-seasonal backshift operators and d is the non-seasonal differencing order. For the seasonal part, θ_p is the seasonal AR component coefficients while $\emptyset_q$ is the seasonal moving Average component coefficients, U^s is the seasonal backshift operators and D represents the seasonal differencing order.

The SARIMA models are generated using the methodology described then the order of differencing is evaluated by unit root stationarity tests and the principal model is selected using the selection criterion.

Box Jenkins Methods of SARIMA Model

Identifying a perfect Seasonal ARIMA model for a specific time series analysis, Box and Jenkins (1970) proposed a methodology that consists of four major steps, namely,

i) identification of the model
ii) parameters of the model estimation
iii) checking the goodness of fit of the model
(iv) Utilization of the final model in forecasting.

The iterative procedures adopted are illustrated in the flow chart below;

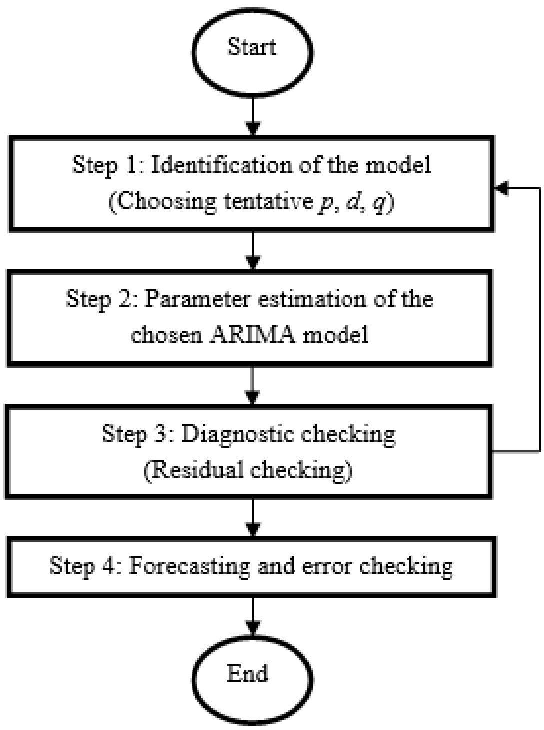


Fig 1: Flow chart of the ARIMA modelling.

Model Identification**:** Identifying the appropriate model comprises specifying the relevant form of AR, MA or ARMA model in order. Non-stationarity tests such as autocorrelation function, partial autocorrelation function, Augmented dickey fuller, Kwiatkowski–Philips–Schmidt–Shin (KPSS) and Phillips–Perron (PP) at 0.05 level of significant level were used to check the stationarity of the data involved. The stationarity was also evaluated by the Mann–Kendall trend test (MK) [Kendall, 1975] as proposed by Box -Jenkins (1976), seasonality identification in the series was done by examining the ACF and PACF series in identifying preliminary values of autoregressive order P, the order of differencing d, the moving average order q and to disclose the correlative patterns among variables.

Table 1. Patterns of Seasonal ACF and PACF

Model	ACF	PACF
AR(p)	Exponentially declining or damped trend size	Spikes to lag p then zero
MA(q)	Spikes to lag q then zero	Exponentially declining or damped wave size
ARMA (p, q)	Exponentially declining or damped trend size following (q-p) lags	Exponentially declining or damped trend size following (p-q) lags
SMA(Q)s	Spikes for lag Q then zero	Spikes to lag P then zero
SARIMA (P, Q)s	Exponentially declining or damped trend size following lag (Q-P)s	Exponentially declining or damped trend size following lag (P-Q)s

Note: AR(p) is the Autoregressive operator, MA(q) is the Moving average operator, ARMA(p,q) is the Autoregressive Moving Average operator, SMA(Q) is the Seasonal Moving Average Operator and SARIMA(P,Q) is the Seasonal Autoregressive Integrated Moving Average operator. Also, ACF is the Autocorrelation function and PACF is the Partial Autocorrelation function.

The model selection was done using the optimum selection criteria by selecting the model with Akaike Information Criteria (Akaike, 1978), Schwarz Information Criterion (SIC) and Hannan-Quinn Information Criteria (HQIC)

Estimation of Parameters**:** After detecting a tentative model, the coefficients of the models were estimated by maximum likelihood estimation methods. The likelihood function for the parameters of known observation data is given as;

$$L\left(\theta, \emptyset\sigma_{\epsilon}^{2} \middle| X\right) = (2\pi\sigma_{\epsilon}^{2})^{-\frac{1}{2}}$$
$$(\exp(\frac{1}{2}\sigma_{\epsilon}^{2}S(\theta,\emptyset)) \quad (12)$$

Diagnostic checking (Goodness of fit)**:** With the diagnostic checking procedure, the estimated model was investigated for statistical significance and acceptability, in other words, if it "best" fits the data. The diagnostic testing of the model used involves the normality test (Jarque & Bera, 1980), autocorrelation test (Ljung & Box, 1978 statistic), and ARCH (squared residuals'). The lung box test for the residual examination is expressed as equation (13) below

$$Q = n(n+2)\sum_{k=1}^{k} n - k^{-1} \hat{\rho}_{k}^{2} \quad (13)$$

If $Q > \chi^2$ or the ρ- value is less than the Significant value, then the null hypothesis of no residual with white noise property is rejected.

Model Forecasting and Performance evaluation**:** Adopting the last phase of Box Jenkins methods of ARIMA Modelling, the precision of the forecast model was evaluated using the model performance validation criteria such as; Akaike Information criteria (AIC), Schwarz Information Criterion (SIC) and Hannan-Quinn Information Criteria (HQIC)

$$AIC = 2K - \log l \quad (14)$$

$$SIC = 2Klogn - logl \quad (15)$$

$$HQIC = -2logl + 2klogn \quad (16)$$

where k symbolizes the total of estimable parameters, l denotes the maximum likelihood and n is the digits of samples.

Moreover, the performances of the model were also examined with the validation statistic. Mean absolute error (*MAE*), *Mean Absolute Percentage Error (MAPE)* and Root mean square error (*RMSE*). MAE is the absolute value of the disparity between the forecasted value and the real value. It estimates the average absolute deviation of predicted values from actual values. MAE is estimated as follows:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|\hat{y}_f - y_t\right| \qquad (17)$$

MAPE is estimated as the average absolute percent error for each time period minus actual values divided by actual values. It calculates the percentage of mean absolute error that occurred in the model formation.it is expressed as follows;

$$MAPE = \frac{100}{n}\sum_{t}^{n}\left|\frac{\hat{y}_f - y_t}{y_t}\right| \qquad (18)$$

RMSE describes the absolute fit of the model to the observed data, it is expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\hat{y}_f - y_t} \qquad (19)$$

Results

Table 2. Descriptive Statistics

	ANNUAL MAX TEMP	**ANNUAL MIN TEMP**	**ANNUAL RAINFALL**
Mean	379.021	278.715	1673.43
Median	378.161	278.601	1683.7
Maximum	389.604	285.36	2125.6
Minimum	371.478	278.622	1173.3
Std. Dev.	3.89293	3.860	223.2367
Skewness	0.65701	-0.0610	-0.08298
Kurtosis	3.47137	2.2390	2.842877
Jarque-Bera	2.43607	0.5432	0.6528
P- value	0.5833	0.9012	0.8410
Observations	30	30	30

The summary of maximum, minimum temperature and rainfall observations from 1989 to 2018 is described in Table 1 above. The highest and the lowest values for the maximum temperature were recorded in the year 1991 and 1998 respectively while the maximum and the minimum values for the minimum temperature were recorded in the year 2016 and 1989 respectively. Also, the maximum and the minimum values for the annual rainfall were documented in the years 2006 and 2015 respectively. The series is normally distributed for annual minimum temperature and the annual rainfall as showed by the high p-value and also their Jarque-Bera test values are low. The various stationarity tests at the level are presented in Table 3. The tests showed that the average annual maximum temperature, minimum temperature and rainfall data series exhibit non-stationary characteristics respectively. Moreover Mann–Kendall (MK) test in Table 4 also confirmed a trend in the data series supporting non-stationarity.

Table 3: Stationarity test results at level

Variables	ADF Test Stat	Prob.	PP Test Stat.	Prob.	KPSS Test Stat.	Prob.
Max. Annual Temp.	-2.195	0.263	-6.336*	0.027	-6.662	0.0005
Min. Annual Temp.	0.119	0358	0.034	0.249	-2.961	0..0001
Annual Rainfall	-0.782*	0.016	-4.620	0.613	-5,225	0.0024

Note * indicates significance at $\alpha = 0.05$ at the level.

Table 4: Seasonal Mann Kendall Trend Analysis

Parameters	Annual Maximum Temp.	Annual Minimum Temp.	Annual Rainfall
Kendall's tau	0.297	0.333*	0.209
Sen's Slope	0.245	0.189	6.869
S	129	145	91
P value	0.226	0.009	0.109

Note * indicate significance at $\alpha = 0.05$

Stationarity and invertibility (Nury et al., 2013) are necessary conditions required to proceed with the use of the ARIMA procedure for the modelling, the non-stationarity in the series was transformed into a stationary series by data differencing. Table 5 shows the results of the stationarity test at first difference.

Table 5: Stationarity test results at first difference

Variables	ADF Test Stat	Prob.	PP Test Stat.	Prob.	KPSS Test Stat.	Prob.
Max. Annual Temp.	-6.292*	0.000	-5.721*	0.0001	0.027*	0.309
Min. Annual Temp.	-3.699*	0.000	-3.117*	0.0000	0.063*	0.511
Annual Rainfall	-2.976*	0.002	-2.552*	0.0000	0.071*	0.192

Note * indicates significance at $\alpha = 0.05$ in the first difference. Moreover, from Table 5 above, KPSS confirmed the stationarity of the annual data series.

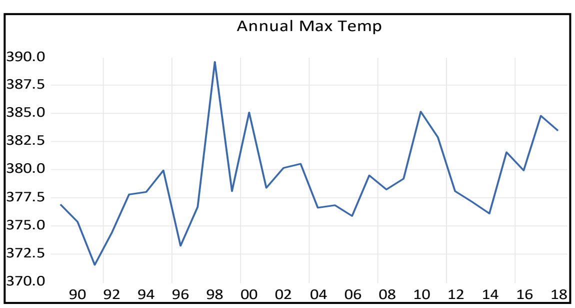


***Fig 2*. Seasonality of the Annual Max. Temp.**

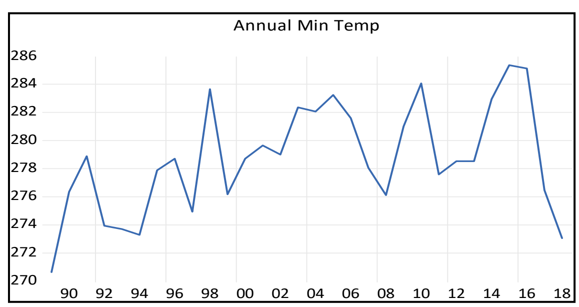


***Fig 3*. Seasonality of the Annual Min. Temp**

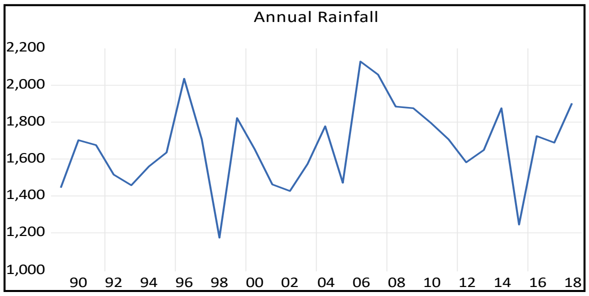


Fig 4. Seasonality of the Annual Rainfall.

Figures 2, 3, and 4 display the seasonality of the average annual maximum, minimum and rainfall data. The illustrations below allow the usage of seasonal differencing which will now give motive for the choice of SARIMA model based on the ACF and PACF of the seasonally differenced series.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.682	-0.682	14.457	0.000
		2	0.212	-0.472	15.913	0.000
		3	0.040	-0.165	15.967	0.001
		4	-0.242	-0.434	18.015	0.001
		5	0.351	-0.159	22.521	0.000
		6	-0.263	-0.080	25.166	0.000
		7	0.071	-0.175	25.369	0.001
		8	0.057	-0.165	25.504	0.001
		9	-0.112	-0.140	26.059	0.002
		10	0.178	0.022	27.539	0.002
		11	-0.218	-0.129	29.880	0.002
		12	0.133	-0.165	30.812	0.002
		13	0.061	0.118	31.017	0.003
		14	-0.196	0.017	33.315	0.003
		15	0.208	-0.002	36.115	0.002
		16	-0.233	-0.186	39.916	0.001
		17	0.241	0.004	44.343	0.000
		18	-0.125	-0.062	45.663	0.000
		19	0.004	0.006	45.665	0.001
		20	0.043	0.000	45.858	0.001
		21	-0.062	0.081	46.320	0.001
		22	0.042	-0.108	46.564	0.002
		23	0.005	-0.083	46.568	0.003
		24	-0.021	0.064	46.662	0.004
		25	0.030	0.059	46.920	0.005
		26	-0.030	0.034	47.303	0.006
		27	0.008	0.061	47.353	0.009

***Fig 5*: Average Annual Max. Temp.**

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.551	-0.551	9.4398	0.002
		2	0.097	-0.297	9.7418	0.008
		3	-0.011	-0.167	9.7462	0.021
		4	-0.182	-0.393	10.911	0.028
		5	0.269	-0.115	13.561	0.019
		6	-0.150	-0.101	14.420	0.025
		7	0.125	0.052	15.045	0.035
		8	-0.145	-0.095	15.930	0.043
		9	0.072	0.013	16.162	0.064
		10	-0.035	-0.063	16.220	0.093
		11	-0.014	-0.089	16.230	0.133
		12	0.142	0.058	17.291	0.139
		13	-0.238	-0.141	20.452	0.084
		14	0.189	-0.069	22.585	0.067
		15	-0.042	0.061	22.699	0.091
		16	0.021	0.145	22.729	0.121
		17	-0.116	-0.171	23.753	0.126
		18	0.088	-0.012	24.410	0.142
		19	-0.043	-0.080	24.580	0.175
		20	0.072	0.036	25.117	0.197
		21	-0.002	-0.030	25.118	0.242
		22	-0.100	-0.015	26.512	0.230
		23	0.015	-0.183	26.552	0.275
		24	-0.019	-0.173	26.626	0.322
		25	0.090	-0.088	28.909	0.268
		26	-0.017	-0.097	29.028	0.310
		27	-0.015	-0.109	29.222	0.350

***Fig 6:* Average Annual Min. Temp.**

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.589	-0.589	10.781	0.001
		2	-0.005	-0.538	10.782	0.005
		3	0.159	-0.348	11.630	0.009
		4	-0.066	-0.248	11.780	0.019
		5	0.079	0.056	12.006	0.035
		6	-0.232	-0.228	14.061	0.029
		7	0.291	-0.018	17.439	0.015
		8	-0.047	0.250	17.532	0.025
		9	-0.328	-0.203	22.287	0.008
		10	0.430	0.011	30.931	0.001
		11	-0.237	-0.078	33.718	0.000
		12	0.052	-0.035	33.862	0.001
		13	-0.019	-0.019	33.883	0.001
		14	0.016	-0.018	33.899	0.002
		15	0.093	0.029	34.458	0.003
		16	-0.259	-0.077	39.161	0.001
		17	0.261	-0.047	44.381	0.000
		18	-0.057	-0.049	44.655	0.000
		19	-0.092	0.116	45.443	0.001
		20	0.107	0.112	46.646	0.001
		21	-0.072	0.134	47.274	0.001
		22	0.016	-0.086	47.309	0.001
		23	0.013	-0.004	47.337	0.002
		24	0.011	0.063	47.365	0.003
		25	-0.031	-0.145	47.637	0.004
		26	0.016	0.069	47.741	0.006
		27	-0.010	-0.044	47.821	0.008

***Fig 7*: Average Annual** Rainfall

Seasonal ARIMA Model

Following the visualization of the correlogram in Fig 5,6 and 7 different tentative Seasonal ARIMA models for variables were fitted to the series and the model with the lowest value of AIC, SIC and HQIC was selected as the utmost model among the competitors. The competitive models and their respective values for the selection criteria are tabulated in Table 6. The best model is in bold print and asterisk mark for easier identification.

Cite this article as:
Adewole, A.I. (2023). Statistical modelling of temperature and rainfall in Ijebu Ode Nigeria using SARIMA. *FNAS Journal of Scientific Innovations, 5*(2), 55-68.

Table 6: Seasonal ARIMA Model

Model	Specification(p,d,q)(P,D,Q)$_s$	AIC	SIC	H QIC
ANNUAL MAXIMUM TEMPERATURE				
Model 1	ARIMA (1,1,1) $(1,1,1)_{12}$	6.015	6.205	6.073
Model 2	ARIMA $(2,1,1)(1,1,0)_{12}$	6.199	6.318	6.248
Model 3	ARIMA $(4,1,1)(1,1,1)_{12}$	**5.729***	**5.918***	**5.789***
ANNUAL MINIMUM TEMPERATURE				
Model 1	ARIMA $(1,1,1)(1,1,1)_{12}$	**5.635***	**5.824***	**5.694***
Model 2	ARIMA $(4,1,1)(1,1,1)_{12}$	5.708	5.897	5.767
ANNUAL RAINFALL				
Model 1	ARIMA $(1,1,1)(1,1,1)_{12}$	13.974	14.162	14.683
Model 2	ARIMA $(2,1,1)(0,1,1)_{12}$	**13.902***	**14.121***	**13.041***
Model 3	ARIMA $(10,1,1)(1,1,1)_{12}$	13.932	14.179	13.991

Table 7: Parameter Estimates of the Seasonal ARIMA fitted model.

Parameter	Coefficient	Standard Error	Prob.
ANNUAL MAXIMUM TEMPERATURE			
θ_1	0.0356	0.0026	0.0000
θ_2	0.2217	0.0291	0.0000
θ_3	-0.5470	0.1042	0.0005
θ_4	0.0831	0.0416	0.0000
$\emptyset_1$	-0.1578	0.0327	0.0000
Θ_1	-1.4521	0.0220	0.0010
Φ_1	0.0781	0.0004	
ANNUAL MINIMUM TEMPERATURE			
θ_1	0.1103	0.3631	0.0032
$\emptyset_1$	-0.0571	0.0491	0.0000
Θ_1	0.0205	0.0190	0.0001
Φ_1	0.3167	0.0229	0.0074
ANNUAL RAINFALL			
θ_1	0.0935	0.0448	0.0000
θ_2	-0.1894	0.0619	0.0000
$\emptyset_1$	0.0561	0.0853	0.0029
Θ_1	-2.6730	0.0346	0.0016
Φ_1	0.0602	0.2505	0.0047

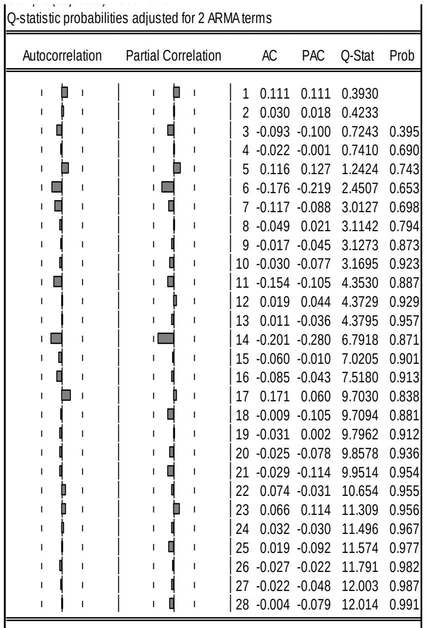
where $\theta_1, \theta_2, \theta_3, \theta_4$, are the autoregressive parameters of non-seasonal components, $\emptyset_1$ is the moving average parameters of non-seasonal components, Θ_1 is the autoregressive parameters of seasonal components and Φ_1 is the moving average parameters of the seasonal component. The results are shown in Table 6. The parameter estimation adopted from Box and Jenkins procedures engaged the maximum likelihood method of estimations, which relied on asymptotic conditions for any time series observation in line with Brockwell et al. (2013). The fitted ARIMA model for the annual maximum temperature, annual minimum temperature, and annual rainfall is specified above in Table 7.

Diagnostic Checking

Having selected the suitable seasonal ARIMA models for the variables, several statistical checks and diverse diagnostic plots were used for the exploration of the appropriateness of the data to the chosen model. Figure 13 and Figure 14, show the correlogram graph for residuals of the annual maximum, minimum temperature and rainfall data models respectively

Cite this article as:
Adewole, A.I. (2023). Statistical modelling of temperature and rainfall in Ijebu Ode Nigeria using SARIMA. *FNAS Journal of Scientific Innovations, 5*(2), 55-68.

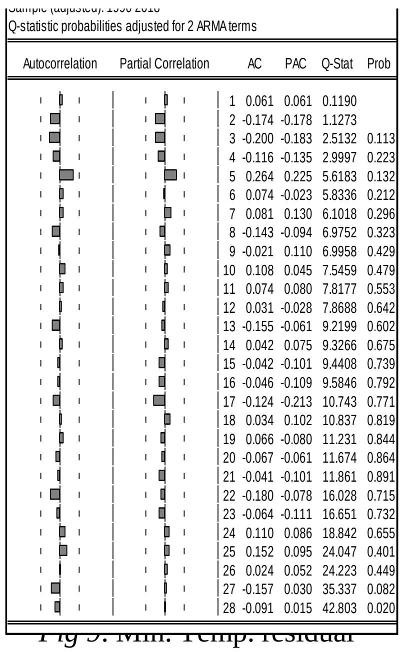
Correlogram of the residual



Q-statistic probabilities adjusted for 2 ARMA terms

	AC	PAC	Q-Stat	Prob
1	0.111	0.111	0.3930	
2	0.030	0.018	0.4233	
3	-0.093	-0.100	0.7243	0.395
4	-0.022	-0.001	0.7410	0.690
5	0.116	0.127	1.2424	0.743
6	-0.176	-0.219	2.4507	0.653
7	-0.117	-0.088	3.0127	0.698
8	-0.049	0.021	3.1142	0.794
9	-0.017	-0.045	3.1273	0.873
10	-0.030	-0.077	3.1695	0.923
11	-0.154	-0.105	4.3530	0.887
12	0.019	0.044	4.3729	0.929
13	0.011	-0.036	4.3795	0.957
14	-0.201	-0.280	6.7918	0.871
15	-0.060	-0.010	7.0205	0.901
16	-0.085	-0.043	7.5180	0.913
17	0.171	0.060	9.7030	0.838
18	-0.009	-0.105	9.7094	0.881
19	-0.031	0.002	9.7962	0.912
20	-0.025	-0.078	9.8578	0.936
21	-0.029	-0.114	9.9514	0.954
22	0.074	-0.031	10.654	0.955
23	0.066	0.114	11.309	0.956
24	0.032	-0.030	11.496	0.967
25	0.019	-0.092	11.574	0.977
26	-0.027	-0.022	11.791	0.982
27	-0.022	-0.048	12.003	0.987
28	-0.004	-0.079	12.014	0.991

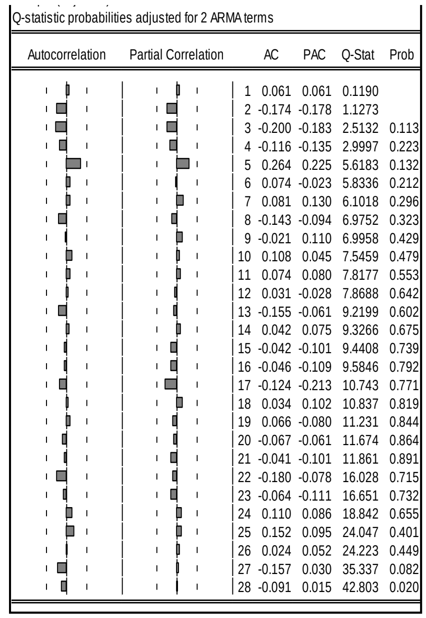
Fig 8 Max. Temp. residual



Q-statistic probabilities adjusted for 2 ARMA terms

	AC	PAC	Q-Stat	Prob
1	0.061	0.061	0.1190	
2	-0.174	-0.178	1.1273	
3	-0.200	-0.183	2.5132	0.113
4	-0.116	-0.135	2.9997	0.223
5	0.264	0.225	5.6183	0.132
6	0.074	-0.023	5.8336	0.212
7	0.081	0.130	6.1018	0.296
8	-0.143	-0.094	6.9752	0.323
9	-0.021	0.110	6.9958	0.429
10	0.108	0.045	7.5459	0.479
11	0.074	0.080	7.8177	0.553
12	0.031	-0.028	7.8688	0.642
13	-0.155	-0.061	9.2199	0.602
14	0.042	0.075	9.3266	0.675
15	-0.042	-0.101	9.4408	0.739
16	-0.046	-0.109	9.5846	0.792
17	-0.124	-0.213	10.743	0.771
18	0.034	0.102	10.837	0.819
19	0.066	-0.080	11.231	0.844
20	-0.067	-0.061	11.674	0.864
21	-0.041	-0.101	11.861	0.891
22	-0.180	-0.078	16.028	0.715
23	-0.064	-0.111	16.651	0.732
24	0.110	0.086	18.842	0.655
25	0.152	0.095	24.047	0.401
26	0.024	0.052	24.223	0.449
27	-0.157	0.030	35.337	0.082
28	-0.091	0.015	42.803	0.020

Fig 9: Min. Temp. residual



Q-statistic probabilities adjusted for 2 ARMA terms

	AC	PAC	Q-Stat	Prob
1	0.061	0.061	0.1190	
2	-0.174	-0.178	1.1273	
3	-0.200	-0.183	2.5132	0.113
4	-0.116	-0.135	2.9997	0.223
5	0.264	0.225	5.6183	0.132
6	0.074	-0.023	5.8336	0.212
7	0.081	0.130	6.1018	0.296
8	-0.143	-0.094	6.9752	0.323
9	-0.021	0.110	6.9958	0.429
10	0.108	0.045	7.5459	0.479
11	0.074	0.080	7.8177	0.553
12	0.031	-0.028	7.8688	0.642
13	-0.155	-0.061	9.2199	0.602
14	0.042	0.075	9.3266	0.675
15	-0.042	-0.101	9.4408	0.739
16	-0.046	-0.109	9.5846	0.792
17	-0.124	-0.213	10.743	0.771
18	0.034	0.102	10.837	0.819
19	0.066	-0.080	11.231	0.844
20	-0.067	-0.061	11.674	0.864
21	-0.041	-0.101	11.861	0.891
22	-0.180	-0.078	16.028	0.715
23	-0.064	-0.111	16.651	0.732
24	0.110	0.086	18.842	0.655
25	0.152	0.095	24.047	0.401
26	0.024	0.052	24.223	0.449
27	-0.157	0.030	35.337	0.082
28	-0.091	0.015	42.803	0.020

Fig 10: Rainfall. Residual

Times Series Plot of the Residual

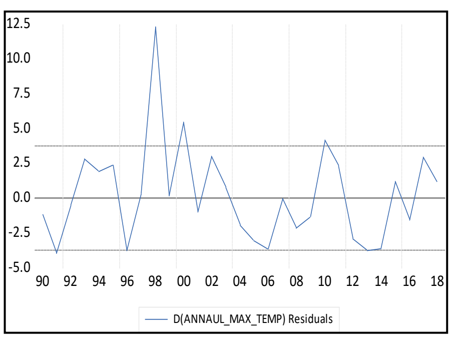


***Fig 11*: Res. Plot of Max. Temp.**

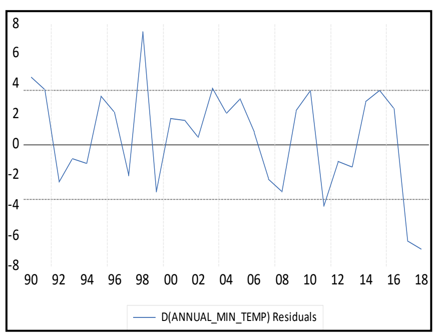


***Fig 12:* Res. Plot of Min. Temp**

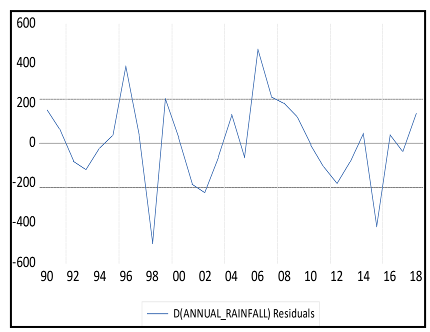


***Fig 13*: Res. Plot of Rainfall**

Following the deductions from the correlogram plot of residuals in Figures 8, 9 and 10, the spikes lie within the bounds, which signifies that none is significant and thus not autocorrelated. For the residual plot, all the plots are supporting that the variance is constant and the average error is zero. Hence, the residual data follows a normal distribution. Thus, it can be deduced that the selected model for each series is appropriate, and the forecasted value is illustrated in the middle red line of the time series plot.

Table 8: Autocorrelation and Heteroskedasticity test of the selected seasonal SARIMA models.

Times series	Autocorrelation Test		Heteroskedasticity Test	
	Lung Box *Q*	Durbin Watson	Breusch Pagan	White

Annual Max. Temp.	0.2903	0.3304	0.6613	0.4275
Annual Mini. Temp	0.5720	0.2046	0.4563	0.2753
Annual Rainfall	0.3185	0.4718	0.3378	0.5204
α	0.05	0.05	0.05	0.05

Table 9: Normality Test of the selected seasonal SARIMA models

Times series	**Shapiro Wiki Test p-value**	**Jarque Bera Test p-value**
Annual Max. Temp.	0.2672	0.1136
Annual Mini. Temp	0.1956	0.3891
Annual Rainfall	0.2284	0.2903
α	0.05	0.05

Table 8 presents the autocorrelation and Heteroskedacity check results and the P values for each selected SARIMA model for the variables. The Ljung-Box and the Durbin Watson P value for all the variables surpassed 0.05(the significant level). These results revealed that there is no autocorrelation among the residual of the model's forecast errors. Also, the results of heteroscedasticity tests of residuals for the variables presented in Table 8 show that the residuals are homoscedastic. Moreover, the normality check in Table 9 shows that the residuals generated from the selected SARMA models are normally distributed, the SARIMA model accuracy is further investigated with the forecast accuracy measures of the models

Table 10: Forecast Accuracy Measures

	Annual Max Temp. SARIMA model (ARIMA (4,1,1) $(1,1,1)_{12}$)	Annual Min. Temp. SARIMA model (ARIMA (1,1,1) $(1,1,1)_{12}$)	Annual Rainfall SARIMA model (ARIMA (2,1,1) $(0,1,1)_{12}$)
MAE	0.2075	0.5152	0.1490
MAPE	0.3771	0.1079	0.5137
RMSE	10.478	8.5527	17.618
R^2	0.7827	0.9176	0.6051

The low values of an unbiased statistic MAPE revealed the adequacy of the selected models to predict accurately. Moreover, the error measures showed indication that the SARIMA models will generate better forecasts.

Forecasting using the selected SARIMA model for the variables.

However, with the aid of the derived model for the variables, the following forecasts were made for the years 2019 to 2028 as shown in Tables 11, 12 and 13 with their lower and upper limits respectively for average annual max temperature, average annual minimal temperature and average annual rainfall. The forecasted values are within the lower and the upper limits. The graphical illustrations are presented in Fig 14, 15 and 16 respectively, the graphical analysis of the forecast revealed that the models seem to yield almost identical forecasts. The values for all the models are within 95% confidence limits which shows that the models have good predictive abilities

Table 11: Annual Max. Temperature Forecast

YEAR	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
FORECAST VALUES	381.429	381.949	381.193	381.877	382.241	382.336	382.680	382.743	382.868	383.047
Lower Limits	347.194	362.180	369.00	370.198	357.031	370.259	366.972	359.145	378.206	344.250
Upper limits	396.541	394.306	415.22	395.284	388.248	394.015	389.452	402.942	396.342	399.422

Table 12: Annual Min. Temperature Forecast

YEAR	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
FORECAST VALUES	281.231	281.76	282.05	282.28	282.47	282.67	282.85	283.04	283.23	283.41
Lower Limits	278.671	280.18	278.22	265.03	279.04	269.52	253.09	277.77	280.69	272.30
Upper Limits	285.093	295.26	284.37	288.00	283.96	293.62	290.11	284.05	291.28	284.27

Table 13: Annual Rainfall Forecast

YEAR	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
FORECAST VALUES	1768.45	1772.6	1778.8	1785.2	1791.5	1797.8	1804.1	1810.4	1816.7	1822.9
Lower Limits	1753.06	1718.2	1770.5	1733.9	1729.4	1746.7	1779.3	1801.9	1795.6	1811.7
Upper limits	1788.42	1790.7	1816.0	1820.4	1862.9	1810.2	1882.0	1873.6	1847.3	1838.2

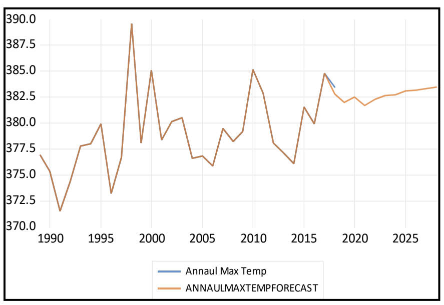


***Fig 14*. forecast of Max. Temp.**

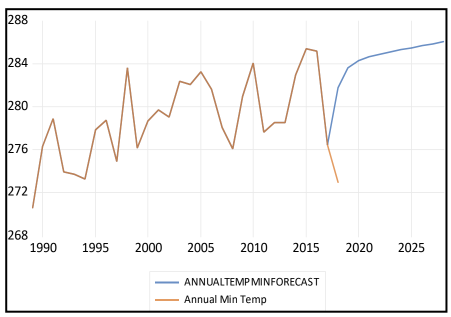


***Fig 15*: forecast of Min. Temp.**

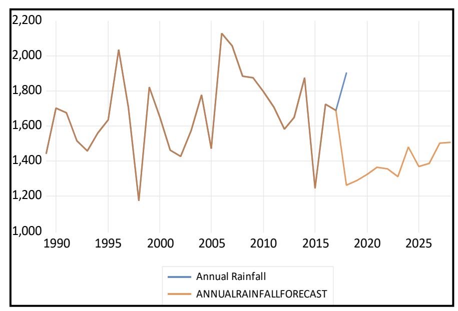


***Fig 16:* forecast of Rainfall**

Discussion

The highest and the lowest values for the maximum temperature were obtained in the year 1991 and 1998 respectively while year 2016 and 1989 recorded the maximum and the minimum values for the minimum temperature respectively. Also, the maximum and the minimum values for the annual rainfall were recorded in the years 2006 and 2015 respectively. A seasonal ARIMA $(4,1,1)(1,1,1)_{12}$, ARIMA $(1,1,1)(1,1,1)_{12}$, and ARIMA $(2,1,1)(0,1,1)_{12}$ for maximum Temperature, minimum Temperature and rainfall respectively were adopted from the Box Jerkins methods. The low values of an unbiased statistic MAPE justifies the adequacy of the selected models to predict accurately. Moreover, the error measures showed indication that the SARIMA models will generate better forecasts, likewise, the residual data follows a normal distribution which signifies that the selected model for each series is appropriate. The result of the forecast models reveals that there is a tendency for an increasing pattern of annual rainfall and temperature over the forecast period from year 2019 to year 2028

Conclusion

This research analyzed and modelled the seasonal autoregressive integrated moving average (SARIMA) of Annual average Minimum, Maximum Temperature and Rainfall in the southwestern part of Nigeria using Ijebu Ode City as a case study using annual average time series of the variables ranging from 1989- 2018. This study presented the methodology of ARIMA model that integrates the seasonality of the series. Tentative numbers of SARIMA models were proposed for the variables based on the visualization of ACF and PACF of the series, a seasonal ARIMA

$(4,1,1)(1,1,1)_{12}$, ARIMA $(1,1,1)(1,1,1)_{12}$, and ARIMA $(2,1,1)(0,1,1)_{12}$ for maximum Temperature, minimum Temperature and rainfall respectively was adopted using the information selection criterion. The estimation and diagnostics examination shows that the models effectively fitted the original data, the performance of each model was evaluated using the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), the error measures results showed indication that the selected SARIMA models will generate a more enhanced predictions. The results reveal that there will be a series of rises in maximum and minimum temperatures for the next 10 years and this demands preparatory adaptive measures for the community. The increase in temperature proposes that climate change may continually affect the livelihood and the economic sectors of Ijebu Ode and its environment if adequate preparations are not in order. Also, the seasonal rise in rainfall patterns reveals the possibility of flooding problems in the coming years, this developed model can aid in planning probable future approaches associated with the weather situations of Ijebu Ode City and its immediate environment.

References

Adams, S. O., & Bamanga, M. A. (2020). Modelling and forecasting seasonal behavior of rainfall in Abuja, Nigeria; A SARIMA approach. *American Journal of Mathematics and Statistics, 10*(1), 10-19.

Adewole, A. I., Amurawaye, F. F., & Oladipupo, J. O. (2023). Time series analysis of malaria fever prevalence in Ogun State. *TASUED Journal of Pure and Applied Sciences*, *2*(1), 201-211.

Akaike, H. (1978). On the likelihood of a time series model. *Journal of the Royal Statistical Society, 24*,(3/4), *217–235,*

Amjad, M., Khan, A., Fatima, K., Ajaz, O., Ali, S., & Main, K. (2023). Analysis of temperature variability, trends and prediction in the Karachi region of Pakistan using ARIMA models. *Atmosphere*, *14*(1), *1-14*. https://doi.org/10.3390/atmos14010088

Barker, D., Huang, X. Y., Liu, Z. Q., Auligné, T., Zhang, X., & Rugg, S., et al. (2012). The weather researches and forecasting model's community variational/ensemble data assimilation system: *Bulletin of the American Meteorological Society, 93* (6), *831 - 843.*

Box, G.E.P., & Jenkins, G.M. (1976). Time series analysis, forecasting and control, *Holden-Day*

Box, G.E.P., & Jenkins, G.M. (1970). Time Series Analysis; *Holden-Day: San Francisco, CA, USA.*

Brockwell, P. J., & Davis, R. A. (2013). Time series: theory and methods. *Springer Science and Business Media*

Chen, P., Niu, A., Liu, D., Jiang, W., & Ma, B. (2018, August). Time series forecasting of temperatures using SARIMA: An example from Nanjing. In *IOP Conference Series: Materials Science and Engineering* (Vol. 394, p. 052024). IOP Publishing.

Chung, E. S., Park, K., & Lee, K. S. (2011). The relative impacts of climate change and urbanization on the hydrological response of a Korean urban watershed. *Hydrological Processes*, *25*(4), 544-560.

Datong, G.M., & Goltong, E.N. (2017). An Application of ARIMA Models in Weather Forecasting: A Case Study of Heipang Airport – Jos Plateau, Nigeria. *International Journal of Innovative Scientific & Engineering Technologies Researchm, 5(2), 1-13*

Jarque, M. C., & Bera, K. A. (1980). Efficient tests for normality, homoskedasticity, and serial independence of regression residuals. *Economic Letters, 6*(3), 255-259.

Jibril, Y. K., & Sanusi, A. J. (2019). A seasonal ARIMA model of Zaria monthly relative humidity, North Western Nigeria. *International Journal of African Sustainable Development*, *11*(2), 77-88.

Kendall, M. (1975). Multivariate analysis. *Charles Griffin & Company.218.*

Ljung, G. M. & Box, G. E. P. (1978). On a measure of lag of t in time series model, *Biometrika, pp 65:2, 297-303.*

Murat, M., Malinowska, I., Gos, M., & Krzyszczak, J. (2018). Forecasting daily meteorological time series using ARIMA and regression models. *International agrophysics*, *32*(2).

Nirmala, M., & Sundaram, S. M. (2010). A seasonal ARIMA model for forecasting monthly rainfall in Tamilnadu. *National Journal on Advances in Building Sciences and Mechanics*, *1*(2), 43-47.

Nnoka, L.C, Lekara-Bayo, I.B., & Etuk, E.H (2020). SARIMA modelling of monthly rainfall in Rivers State of Nigeria. *International Journal of Science and Advanced Innovative Research,* 5(3), 1-11.

Nury, A. H., Koch, M., & Alam, M. J. B. (2013, August). Time series analysis and forecasting of temperatures in the sylhet division of bangladesh. In *4th International Conference on Environmental Aspects of Bangladesh (ICEAB), August* (pp. 24-26).

Nwokike, C. C., Offorha, B. C., Obubu, M., Uguola, C. B., & Ukomah, H. I. (2020). Comparing SANN and SARIMKA for forecasting frequency of monthly rainfall in Umuahia. *Scientific African, 10*.

Onyeke, U., Khalid, J. N., & Ogundeji, R. K. (2021). Optimal stochastic forecast models of rainfall in South-West region of Nigeria. *International Journal of Mathematical Sciences and Optimization: Theory and Applications*, *7*(2), 1-20.

Osinowo, O. O (2013). Application of landsat enhanced thematic mapper plus (etm+) and radar imageries for geological boundary delineation and lineament mapping in a transition zone around Ijebu-Ode, Southwestern Nigeria. *Africa Geoscience Review*, *20*(3& 4), 83-96.

Raicharoen, T., Lursinsap, C., & Sanguanbhokai, P. (2003, May). Application of critical support vector machine to time series prediction. In *Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS'03.* (Vol. 5, pp. V-V). IEEE.

Shan, F., Min, J., & Xu, D. (2016). Assimilation of radar radial velocity data with the hybrid system for the prediction of hurricane . *Atmospheric Research, 169, 127-138.*

Shathir, A., Ali Mohammed Saleh, L., & Amal Al-Din Majeed, S. (2019). Forecasting monthly maximum temperatures in kerbala using seasonal ARIMA models. *Journal of university of Babylon for engineering sciences, 27(2), 223-232.* *https://doi.org/10.29196/jubes.v27i2.2341*

Solar Radiation Data . *Global Scientific Journal, 10 (8), 1461-1475.*

Tadesse, K. B., & Dinka, M. O. (2017). Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa. *Journal of Water and Land Development*, *35*(X–XII), 229–236.

Udo, M. E., & Shittu, O. I. (2022). A comparative study of SARIMA and SARFIMA models: An application to solar radiation data. *Global Scientific Journal, 10*(8). Retrieved from www.globalscientificjournal.com

Wiri, L., & Essi, I. D. (2018). Seasonal Autoregressive Integrated Moving Average (SARIMA) Modelling and Forecasting of Inflation Rates in Nigerian (2003).

Xiaojuan, Z., Yunzhong, J., Changhai, Qin et al., (2007). Application of time-series analysis method on groundwater level prediction, J. South-to-North Water Transfer Water Sci. *Technol. (4) 40–42 (in Chinese).*

Xu, D., Min, J., Shen, F., Ban, J., & Chen, P. (2016). Assimilation of radiance data from the fy-3b satellite with the hybrid-3dvar system for the forecasting of binary typhoons. *Journal of Advances in Modelling Earth Systems, 8 (2), 1-12.*