JOURNAL OF STATISTICAL MODELING & ANALYTICS (JOSMA)
(ISSN: 2180-3102)

UNIVERSITI MALAYA

Bayesian Structural Time Series Model and SARIMA Model for Rainfall Forecasting in Nigeria

Ogundeji, Rotimi Kayode[1*] and Okemakinde, Sherif Sunday[2]

[1,2]*Department of Statistics, Faculty of Science, University of Lagos, Akoka, Nigeria*

Corresponding author: rogundeji@unilag.edu.ng

RESEARCH ARTICLE

Received 7 January 2024
Accepted 07 April 2025
Published 29 May 2025

Abstract

Nigeria is recognized as being susceptible to climate change, and global warming if not taken care of, will lead to serious problems on livelihoods in Nigeria, especially in the area of agricultural activities. Rainfall is a major determinant of climate change the world over and climate change is one of the foremost global challenge facing humans at the moment. Using monthly time series rainfall data, Bayesian structural time series (BSTS) methodology was applied to fit models through MCMC algorithm. Also, Seasonal Autoregressive Moving Average (SARIMA) models were fitted to the same dataset using Box-Jenkins approach. The two models are considered based on their respective capacities to capture trend, seasonal and structural components of rainfall data. On the basis of model evaluation criteria (RMSE, MAE, MAPE and MASE), the SARIMA model had values that were clearly significantly smaller than that of the BSTS time series model. This implies that the SARIMA model is more robust in its estimations and forecasting abilities. Similarly, the R squared was larger for the SARIMA model than the BSTS (MCMC) model indicating that the SARIMA model was a better fit for the rainfall data. This study shows that SARIMA model is a more precise and robust in dealing with this type of dataset than BSTS (MCMC) model. It is better because its computational process using differencing, lags and moving averages ensure that the underlying components of the model are properly identified and estimated.

Keywords: Bayesian Methods, Climate Change, MCMC Algorithm, Model Selection Criterion, Time Series.

2.1 Introduction

Rainfall prediction is a difficult endeavor due to the non-linear character of climate processes. In recent years, data-driven (empirical) approaches have surpassed knowledge-driven (physical) approaches in terms of popularity (Ogundari *et al.*, 2021). Rainfall is a climatic parameter whose prediction is challenging and demanding as the world continues to experience climate change. It affects every component of the ecological system including wildlife and vegetation. Therefore, rainfall investigation is vital and cannot be over emphasized. Climate events such as flood have been increasing recently all over the world and this trend has been attributed to climate change and global warming. One indicator of climate change is rainfall (Ogungbenro and Morakinyo, 2014). The study of rainfall is significant to the existence of man and seasonal rainfall patterns are very important for continuous supply of water for domestic and industrial uses. Naturally, rainfall variability is of spatial and temporal forms (Mohammed and Alehile, 2022). Nigeria is recognized as being susceptible to climate change. Climate change and global warming if not taken care of, will lead to serious problems on livelihoods in Nigeria, especially in agricultural activities, because the rainfall seasons will be alternated. Floods which devastate everywhere, change in temperature and decrease in humidity will lead to natural disasters which destroy lives and properties, bringing down the total Gross Domestic Products of the nation.

In different studies, traditional methods like the autoregressive (AR), moving average (MA), autoregressive moving average (ARMA) and the autoregressive integrated moving average models (ARIMA) models have been adopted, as seen in (Bari *et al.*, 2015; Onyeka-Ubaka *et al.*, 2021). These models are based on assumptions that the seasonal component is deterministic and independent of other non-seasonal components with a belief that empirical knowledge of the data be known. A comparative study using machine learning techniques has been used to build models for rainfall prediction (Ayodele and Precious, 2019). To predict long-term pollution trends in Kolkata, India using limited data accessibility, a study of different statistical methods and Bayesian structural time series methods were conducted (Narasimha *et al.*, 2018). Almarashi and Khan, (2020) focused on modeling times series using the Bayesian Structural Time Series technique (BSTS) on a univariate data-set and compared the results with using autoregressive integrated moving average models (ARIMA) models. Cowden *et al.*, (2010) examined stochastic rainfall modeling in West Africa. The study examined two stochastic rainfall models: Markov Models (MM) and Large Scale Weakening (LARSWG). A first order Markov occurrence model with mixed exponential amount was selected as the best option for unconditional Markov models. It was concluded that there was no clear advantage in selecting Markov models over the LARSWG model for Domestic Rainfall in West Africa.

Many researchers (within Nigeria and beyond) have study SARIMA models as very robust for rainfall time series analysis in comparison with other family of ARIMA models (Narasimha, *et al.*, 2018; Xinghua, *et al.*, 2012). Several SARIMA/ adjusted SARIMA models of different dimensions have been established as appropriate optimal forecasting models for rainfall time series (Xinghua *et al.*, 2012; Nwokike *et al.*, 2020). Identified models were evaluated on the basis of Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). For more studies on SARIMA models, see (Amaefula, 2021; Ogunrinde, 2012; Bari *et al.*, 2015; Chonge *et al.*, 2015 and Ogundeji *et al.*, 2021). Similarly, several other researchers have studied rainfall data using SARIMA or BSTS independently such as: (Dabral and Murry, 2017; Lunn *et al.*, 2000; Lee and Heghinian, 1977; Jayawardene, 2005; Aliyu *et al.*, 2021 and Ogundeji *et al.*, 2023). In this study, two models (BSTS and SARIMA) are considered based on their respective capacities to capture trend, seasonal and structural components of rainfall data and thus compared.

South West Nigeria is endowed with the tropical rain forest climate. Abeokuta, Ogun state, Nigeria is characterized by a tropical climate with distinct wet and dry seasons (Obot and Onyeukwu, 2010). Rainfall begins in earnest around March ending and early April each year. Rainfall rest a little in August during the summer and retreats towards the end of October and early November. Thus, the months of April and May are the first two months into the raining season, while the last two months to the end of rainy season in the Sub regions are September and October. Abeokuta is the capital of Ogun State, situated in the South Western region of Nigeria. Abeokuta typically receives about 142.49 millimeters (5.61 inches) of precipitation and has 225.62 rainy days (61.81%of the time) annually.

In this research, the theoretical frameworks and fitting of both the BSTS and SARIMA models to monthly time series rainfall data from Abeokuta were compared. Bayesian structural time series (BSTS) methodology was applied to fit models through MCMC algorithm while Seasonal Autoregressive Moving Average (SARIMA) models were fitted to the same dataset using Box-Jenkins approach. On the basis of model evaluation criteria (RMSE, MAE, MAPE and MASE), forecasts were generated from a better and more robust model.

This study is geared towards the choice of an appropriate model for rainfall forecasts in the areas of agricultural activities.

2. Materials and Methods

2.1 Data

The data for this work are from weather archive of Ogun state, Nigeria website www.meteoblue.com/en/weather/archive/yearcomparison/ogun-state_nigeria. The dataset is monthly rainfall recorded in Abeokuta between 2010 and 2022.

2.2 Bayesian Structural Time Series Model (BSTS)

Bayesian Structural Time Series (BSTS) model, which was introduced by Scott and Varian, 2013, for the estimation of weather rainfall using contemporaneous-predictors. The BSTS allows for

i. decomposing the time series data into several latent components that can describe the underlying dynamics of the data, such as trend, seasonality and regression,
ii. variable selection and
iii. Bayesian model averaging.

Generally, the Bayesian structural model can be written as (Qiu, *et al.*, 2018): Let Y_t denote observation t in a real-valued time series,

$$Y_t = \mu_t + x_t\beta + S_t + \epsilon_t, \qquad e_t \sim N(0, \sigma_e^2) \tag{1}$$

where x_t denotes a set of regressors. S_t represents seasonality, β represents regression coefficients and μ_t is the local level term.

The structural time series models are in fact the building blocks of BSTS. In Structural time series model the data comes from some unobserved process known as state space and the data which is observed is generated from the state-space with added noise.

A structural time series model can be described by a pair of equations relating y_t to a vector of latent state variables α_t

$$y_t = Z_t^T \alpha_t + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_e^2) \tag{2}$$

$$\alpha_t = \begin{bmatrix} T_t \\ S_t \end{bmatrix}, \qquad Z_t = [1\ 1] \tag{3}$$

$\boldsymbol{\alpha}_t$ is the state of the system at t.

Equation (1) is called the observation equation, because it links the observed data y_t with unobserved latent state α_t

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t \qquad \eta_t \sim N(0, \sigma_e^2) \tag{4}$$

Equation (2) is called the state/transition equation because it defines how the latent state evolves over time. The model matrices Z_t, T_t and R_t typically contain a mix of known values (often 0 and 1), and unknown parameters. Equations (2) and (3) together are called state-space model.

The first step in the application of the Bayesian framework on the Structural time series model is specifying the prior distribution for each parameter in the model (i.e., for error variances). The error variances are treated as parameters of the models in all the equations and have conditionally independent inverse Gamma full conditional distributions assuming independent inverse Gamma priors. The second step is to obtain posterior distributions. In BSTS, the forecasts are worked out from the posterior predictive distribution but for fitting the model Kalman filtering and MCMC procedures are used. Filtering simple means updating our knowledge of the system whenever a new observation is added (Almarashi and Khan, 2020).

2.2.1 Model Estimation

Let $\phi \equiv (\theta; \alpha)^T$ and ψ denote the set of parameters other than β *and* σ_ε^2 considering an initial $\theta = \theta^{(0)}$generated from the prior distributions, we obtain a stationary distribution $p(\phi|y)$ using MCMC as follows (Broemeling, 1985; Katarina and Gunardi, 2023):

(i) Simulate α from $p(\alpha|y, \theta)$ using the simulation smoother.
(ii) Simulate ε from $p(\psi|y, \alpha, \beta, \sigma_\epsilon^2)$
(iii) Simulate β and σ_ε^2 from $p(\beta, \sigma_\epsilon^2|y, \alpha, \psi)$

By cycling through Steps (i) – (iii) for M times, a sequence of MCMC draws $\phi^{(1)} - \phi^{(M)}$, is obtained. The first m samples (also known as burn-in samples) may not be representative for the target posterior distribution and hence will be discarded. The remaining sequence of MCMC draws is used to estimate the posterior distribution of y_t by means of (1) and the t (Hung, 2017):

$$y_t = \frac{1}{M-m} \sum_{i=m+1}^{M} \widetilde{y_t} \tag{5}$$

The rainfall dataset does not have any regressors, and therefore a simple Bayesian structural model using MCMC algorithms is fitted as follows;

(i) 1000 MCMC draws
(ii) Trend and seasonality
(iii) Prediction created by averaging across the MCMC draws
(iv) Credible interval generated from the distribution of the MCMC draws

2.3 Seasonal Autoregressive Moving Average (SARIMA) Model

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model fitted to the average monthly rainfall data is of the form (Aliyu *et al.*, 2021):

$$SARIMA(p, d, q)(P, D, Q)s \tag{6}$$

where,
p and P = The order of non-seasonal and seasonal Autoregressive process respectively.
q and Q= The order of non-seasonal and seasonal Moving Average process respectively.
d and D = Non-seasonal and seasonal difference order.
s = Seasonal period.

Let $\{X_t\}$ be a time series. Suppose that it is stationary. It is said to follow an autoregressive moving average model of order p and q (denoted by $ARMA$ (p,q)) if it satisfies the following equation

$$X_t - \alpha_1 + X_{t-1} + \alpha_2 X_{t-2} + \cdots + \alpha_p X_{t-p} = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \tag{7}$$

where $\alpha's$ and $\beta's$ are constants such that (5) be both stationary and invertible and $\{\varepsilon_t\}$ is a white noise process. Equation (5) may as well be written as:

$$A(L)X_t = B(L)\varepsilon_t \tag{8}$$

where

$$A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \cdots \alpha_P L^p \text{ } and \text{ } B(L) = 1 + \beta_1 L + \beta_2 L^2 + \cdots + \beta_q L^q \text{A (L)}$$

L is the backshift operator defined by $L^k X_t = X_{t-k}$. If $\{X_t\}$ is not stationary according to Box and Jenkins (1976, 2004), a certain difference of $\{X_t\}$,$\nabla^d(X_t)$, might be, where $\nabla = 1 - L$ and d is a positive integer.

Then if $\{X_t\}$ is replaced by$\nabla^d(X_t)$, in eqn (5), the model becomes an autoregressive Integrated moving average of orderp, d, q, denoted by an ARIMA $(p, d, q,)$ in $\{X_t\}$. If the series is seasonal of period s follow a seasonal autoregressive integrated moving average model of order (p, d, q) $x(P, D, Q)$ (denoted by a $SARIMA(p, d, q) \times (P, D, Q, S)s)$.

If

$$\text{A(L)}\Phi(L^s)\Delta^d \text{ } \nabla_s^D X_t = \text{B(L}\Theta(L^s)\varepsilon_t \tag{9}$$

$\theta_q L_s^Q$ and ∇_s^D is the seasonal difference operator such that $\nabla_s = 1 - L^s$. Here p is the seasonal autoregressive order, Q is the seasonal moving average order, the $\phi's$ and the $\theta's$ are the seasonal autoregressive and the seasonal moving average parameters (Ayodele and Precious, 2019; Ameafula, 2021).

2.3.1 Box-Jenkins Approach for the SARIMA Model Estimation

The Box-Jenkins technique involves finding the best fit of a time series data. It can be used on seasonal, non-seasonal, stationary and non-stationary series. The approach starts with model identification which include construction of a time plot of the data and inspection of the graph for any anomalies (Cryer and Chan, 2008). If the variance grows with time, it will be necessary to stabilize the variance. The next step is to identify preliminary values of autoregressive order P, the order of differencing d, the moving average order q and their corresponding seasonal parameters P, D and Q.

Here, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) are the most important elements (Shumway and Stoffer, 2008). The ACF measures the amount of linear dependence between observations in a time series that are separated by a lag q.

The PACF helps to determine how many autoregressive terms p is necessary. The parameter d is the order of difference frequency from non-stationary time series to stationary series. Furthermore, a time series plot and ACF of data will typically suggest whether any differencing is needed. If differencing is called for, the time plot will show some kind of linear trend. when preliminary values of D and d have been fixed. Furthermore, parameters can be chosen using Akaike's Information Criterion(AIC) to determine the values of the parameters (Shumway and Stoffer, 2008).

The Box-Jenkins' method is concluded with parametric test and diagnostic checks which include the analysis of the residuals as well as model comparisons. If the model fits well, the standardized residuals should behave as an independent and identically distributed sequence with mean zero and variance one Cryer and Chan, 2008). A standardized residuals plot or a Q-Q plot can help in identifying the normality (Shumway and Stoffer, 2008). Once a model has been identified and all the parameters have been estimated, future values of a time series with these models are generated.

2.3.2 Model Selection Criterion

The following criteria were considered in this study to compare the most appropriate fitted models either using the BSTS or the SARIMA models (Ogundeji *et al.*, 2022);

(i) Root Mean Sum of Square Error (RMSE)

(ii) Coefficient of determination, $\mathbf{R^2}$

(iii) Mean Sum of Square Error (MSE)
(iv) Mean Absolute Error (MAE)
(v) Mean Absolute Percentage Error (MAPE)
(vi) Akaike Information Criteria (AIC)

3. Results and Discussion

3.1 Descriptive Analysis

The dataset is monthly rainfall recorded in Abeokuta between January 2010 and December 2022 (i.e. 156 months). Table 1 is the summary statistics for the average monthly rainfall (millimeters). The highest average monthly rainfall recorded in Abeokuta between 2010 and 2022 is 569.42mm and the lowest was 0.60mm and mean was 136.98mm. The time series plot of the average monthly rainfall in Figure 1 shows trend and seasonal components. That is, recurring patterns/cycles that appear to occur annually.

Table 1: Descriptive Statistics for Average Monthly Rainfall

Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
0.60	45.02	121.06	136.98	212.28	569.42

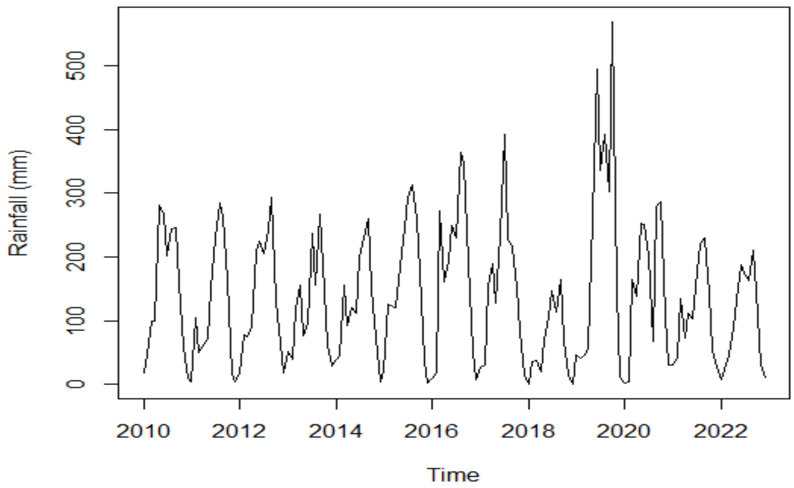

Figure 1: Time Series Plot of Average Monthly Rainfall (Millimeters)

3.2 BSTS Model Using Markov Chain Monte Carlo (MCMC)

Table 2 below shows the summary statistics of the posterior distribution of the BSTS (MCMC) model on average monthly rainfall with one thousand number draws, two number chains and one thousand number of tunes.

Table 2: BSTS (MCMC) Model Summary Statistics Result.

	Posterior Mean	Sd	hdi_2.5%	hdi_97.5%	mcse_mean	mcse_sd	ess_bulk	ess_tail
μ	136.982	8.549	120.829	152.256	0.194	0.137	1956.0	1186.0
σ	107.083	6.042	95.821	118.292	0.148	0.105	1655.0	1249.0

Table 2 gives the BSTS (MCMS) model statistics in terms of the overall mean **μ** and standard deviation **σ** which include the posterior mean, standard deviation and effective sample sizes (ess). The hdi_2.5% and 97.5% percentiles of the posterior sample for each parameter give a 95% posterior credible interval (the range of values within which the parameter falls with probability 0.95). The Monte Carlo Standard Error (MCSE) is an indication of how much error is in the estimate due to the fact that MCMC is used. As the number of iterations increases the MCSE$\rightarrow$0.

Also obtained are model evaluation criterion statistics on identified BSTS (MCMC) model on the average monthly rainfall (2010 – 2022) as shown in Table 3.

Table 3: Model Evaluation Criterion Statistics for Estimated BSTS (MCMC) Model

RMSE	R^2	MSE	MAE	MASE	MAPE
155.064	-0.079	24044.794	125.746	3.960	986.182

Table 3 gives the results of the estimated BSTS (MCMC) model based on the adopted criteria for the purpose of comparison with the SARIMA model.

3.3 Seasonal Autoregressive Moving Average (SARIMA) Model

3.3.1 SARIMA Model of Average Monthly Rainfall from January 2010 to December 2022

Figure 2 shows the time plots of the rainfall data and inspection of the graphs shows the decomposition of the series data through the breakdown of trend, seasonality, and residuals.

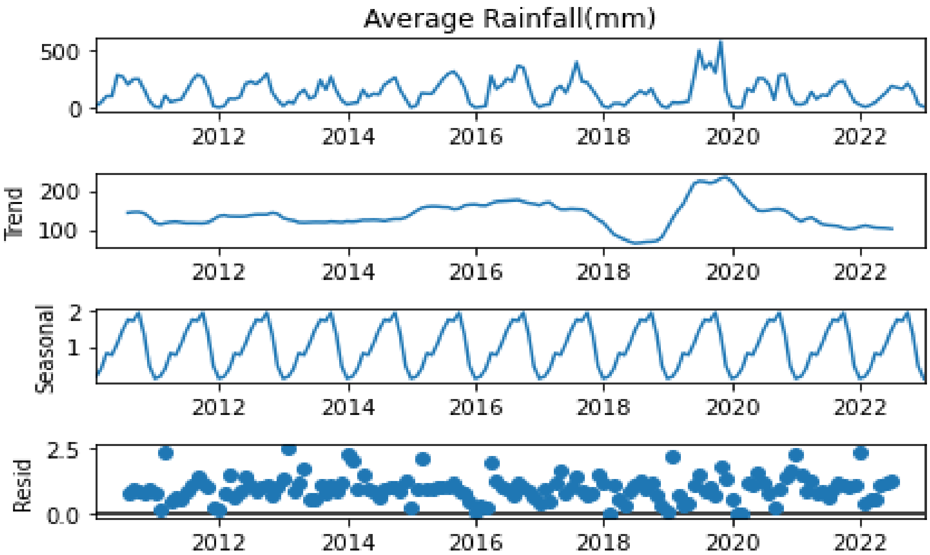

Figure 2: Time Series Decomposition Plots of Average Monthly Rainfall.

Using results of Dickey-Fuller Test to check for stationary of the rainfall data, Table 4 shows that the series is not stationary given that the test Statistics is higher than the critical value of 5% level of significance.

Table 4: Results of Dickey-Fuller Test for Stationary of Rainfall Series

Test Statistics	p-Value	Lags Used	Observations	Critical Value (5%)
-2.7553	0.0649	13.0000	106	-2.8892

In Figure 3, the plots show the ACF, PACF and the distribution of the dataset, The ACF drops to zero relatively slowly and show that the series is non-stationary.

The series has an interesting behavior, there is a sequential significant negative autocorrelation starting at lag 6 and repeating each 12 months, it is because of the difference in the seasons.

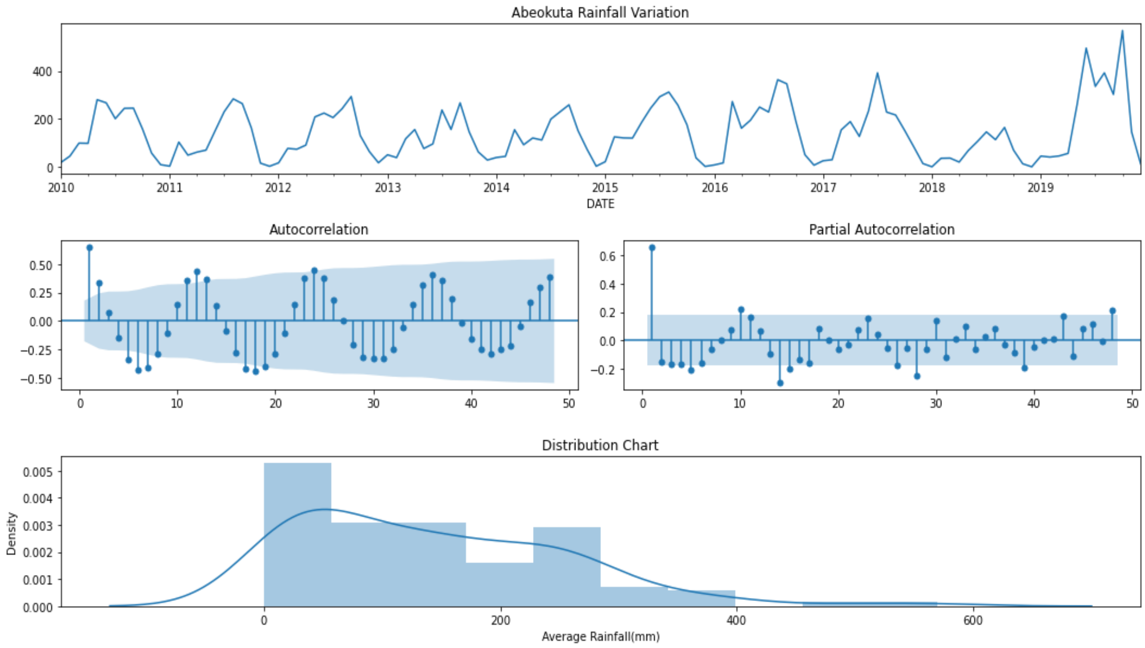

Figure 3: Time Series Plot of ACF, PACF, and Distribution of the Average Monthly Rainfall

Also, from lag 12 and sequentially from every 12 lags there is a significant positive autocorrelation. The PACF shows a negative spike in the first lag and a drop to zero PACF in the following lags.

This behavior between the ACF and PACF plots suggests an AR (1) model and also a first seasonal difference ($Y_t - Y_t - 12$). There is need to plot the stationarity function again with the first seasonal difference to see if we will need some SAR (P) or SMA (Q) parameter:

3.3.2 Stationarity and Differencing

After differencing the series once, the results of Dickey-Fuller Test are shown in Table 5 that the Test Statistic is lower than the Critical Value of 5%. Hence the series seems to be stationary.

Table 5: Results of Dickey-Fuller Test for Stationary of Rainfall Series

Test Statistics	p-Value	Lags Used	Observations	Critical Value (5%)
-3.4919	0.0082	12.0000	106	-2.8892

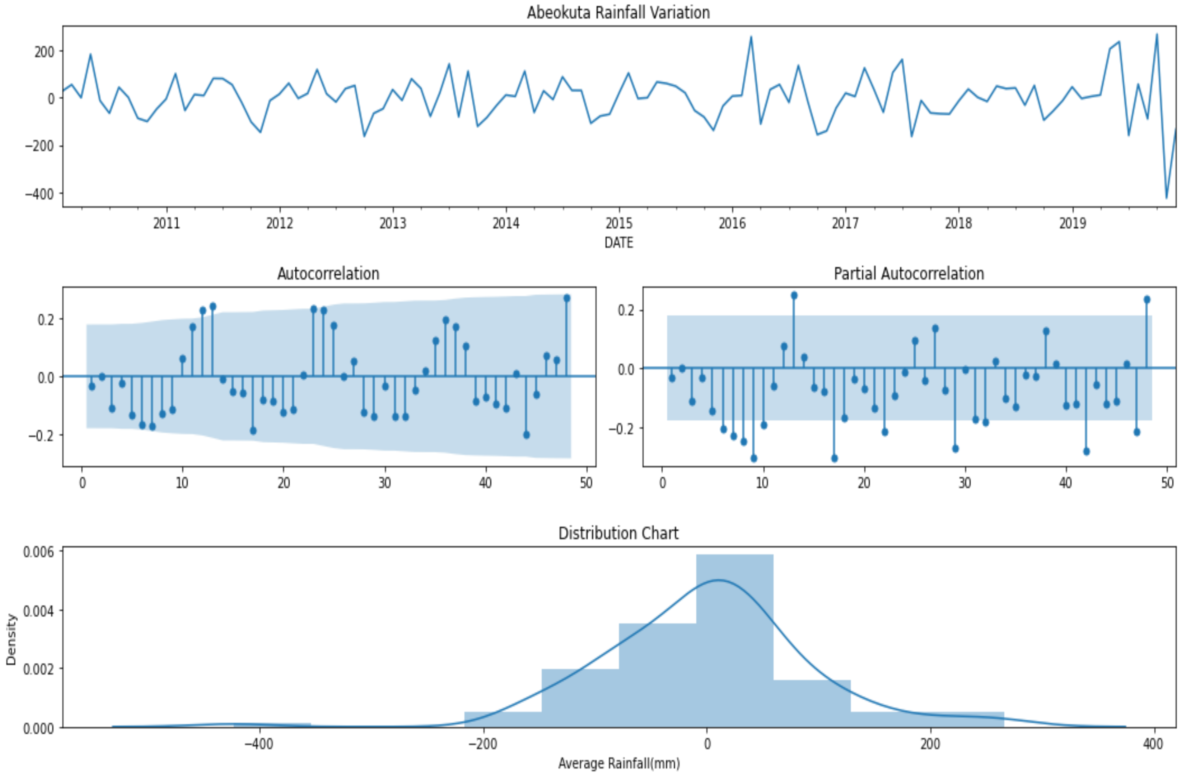

Figure 4: Differencing Plot, ACF and PACF Plots and Distribution of the Average Monthly Rainfall

In Figure 4, the seasonal differencing plots shows the ACF, PACF and the distribution of the dataset, the series is stationary.

Also, after seasonal differencing the series once, the results of Dickey-Fuller Test as shown in Table 6 is that the test statistic is lower than the Critical Value of 5%. Hence the series seems to be stationary.

Table 6: Results of Dickey-Fuller Test for Stationary of Rainfall Series

Test Statistics	p-Value	Lags Used	Observations	Critical Value (5%)
-4.4960	0.0002	11.0000	96	-2.8892

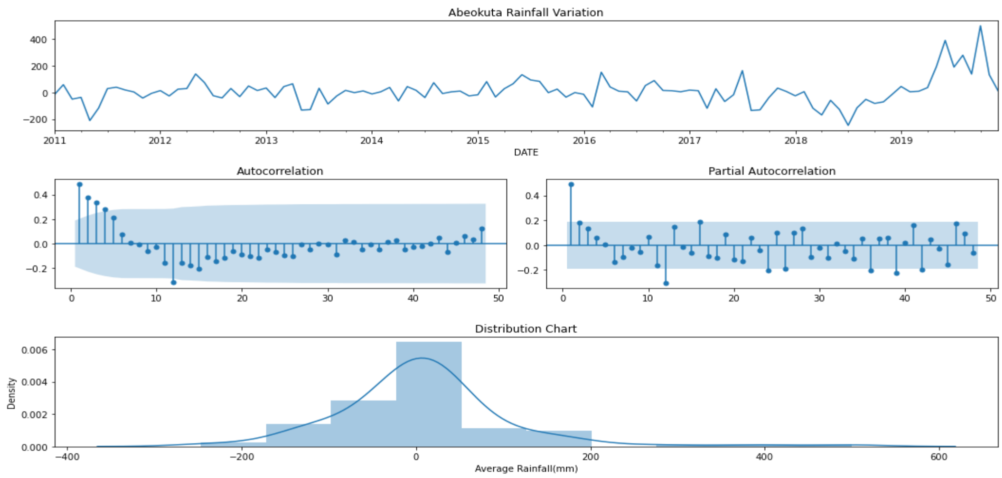

Figure 5: Seasonal Differencing Plot, ACF and PACF Plots and Distribution of the Average Monthly Rainfall

In Figure 5, the seasonal differencing plots shows the ACF, PACF and the distribution of the dataset, the series is stationary. As the plots above showed, the first ACF lags have a gradual decay, while the PACF drops under the confidence interval after the first lag, this is an AR signature with a parameter of 1, so this is an AR (1) model.

3.3.3 SARIMA - Baseline Model

As we used a first seasonal difference, the ACF and PACF showed a significant drop in the 12th lag, it means an SMA signature with a parameter of 1 lag, resuming this is an SAR (1) with a first difference. Initially I'm going to work with the following (p, d, q) orders: (1, 1, 0), and with the following seasonal (P, D, Q, S) orders (0, 1, 1, 12) and as the series has a clear uptrend.
SARIMA (1, 1, 0), (0, 1, 1, 12).

Table 7: Estimated SARIMA Model (1,1,0) x (0,1,1)12

	coef	std err	Z	P>|z|	0.025	0.975
ar.L1	-0.3775	0.049	-7.634	0.000	-0.474	-0.281
Ma.S.L12	-0.7803	0.160	-4.889	0.000	-1.093	-0.467
sigma2	6175.9513	521.018	11.854	0.000	5154.774	7197.129

Based on Table 7, the model chosen was the SARIMA (1,1,0) x(0,1,1)12 model, given by

$$\nabla_{12}X_t = -0.3775X_{t-1} - 0.7803 \in_{t-12} = 6175 \qquad (8)$$

Table 7 presents parameter estimates for the SARIMA (1,1,0)x(0,1,1)12 model, it indicates that both seasonal and non-seasonal coefficients are all significant except the seasonal autoregressive and non-seasonal second components of Moving average because its probability value (0.43654) is greater than common choices of 1%, 5% and 10% level of significance. This could be as a result of the fact that the second spike of ACF in Figure 4 is much away from the interval.

3.4 Model Evaluation

Model evaluation criterion statistics on identified SARIMA (1, 1, 0) x (0, 1, 1)12 model on the average monthly rainfall (2010 – 2022) as shown in Table 8 below.

Table 8: Model Evaluation Criterion Statistics for Estimated SERIMA Model

RMSE	R^2	MSE	MAE	MASE	MAPE
09.488	-0.846	9699.953	92.922	3.061	528.661

Based on the model selection criteria adopted for this study and comparing the results obtained in Table 3 and Table 8, the SARIMA model has values that are clearly significantly smaller than that of the BSTS model apart from R squared. Similarly, the R squared is larger for the SARIMA model than the BSTS (MCMC) model indicating that the SARIMA model is a better fit for the rainfall data.

3.4.1 Diagnostic Check for SARIMA Model

Presented in Figure 6 is (i) the plot of the Current and Predicted values through the time, (ii) Residuals vs. Predicted values in a scatter plot (iii) Q-Q Plot showing the distribution of errors and its distribution (iv) Autocorrelation plot of the Residuals, respectively.

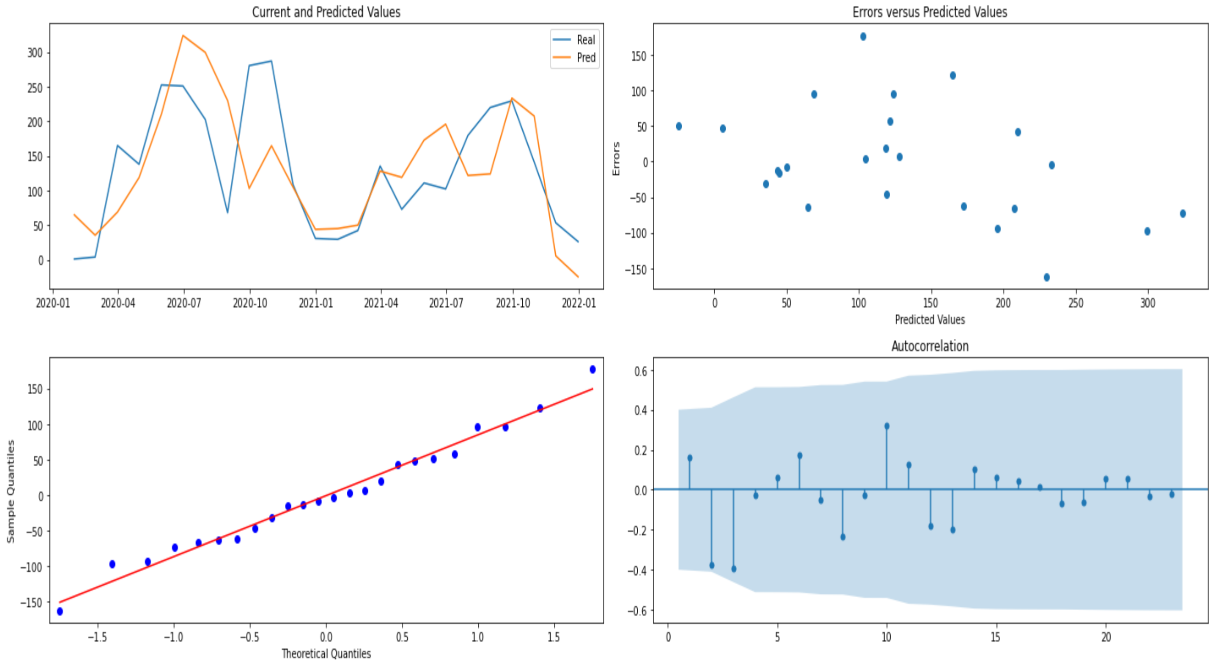

Figure 6: Diagnostic Plots for Adequacy of the Estimated SARIMA Model

The following can be deduced from Figure 6;

(i) Predictions fit very well on the current values.

(ii) The Error vs. Predicted values has a linear distribution.

(iii) The Q-Q Plot shows a normal pattern with some little outliers and,

(iv) The autocorrelation plot shows a positive spike over the confidence interval just at the first lag, as such there is no need for more changes.

Based on the identified SARIMA model the Figure 7 is the extrapolated prediction in the test set for the last 12 months. The SARIMA parameters were well fitted, the predicted values following the real values and also the seasonal pattern. The plot shows the comparison between the actual values and predicted values for SARIMA (1,1,0) x (0,1,1)12.

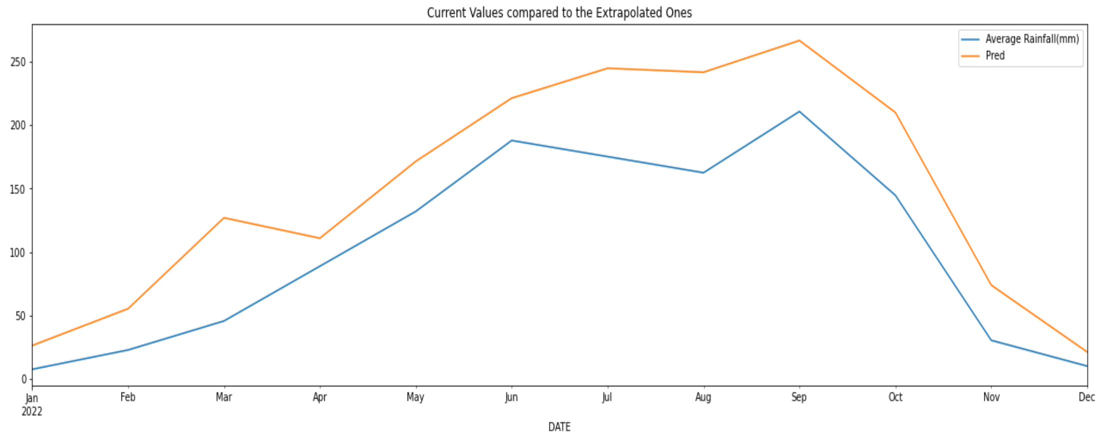

Figure 7: Forecast plot of rainfall using identified SARIMA model

3.5 Comparison BSTS Model and SARIMA Model

Table 9 gives the summary of model selection criteria for selecting a better or more appropriate model between BSTS Model and SARIMA Model.

Table 9: Model Comparison Between BSTS Model and SARIMA Model

MODELS	RMSE	MAE	MAPE	MSE	MASE	R^2
BSTS (MCMC)	155.064	125.746	986.182	24044.794	3.960	-0.079
SARIMA	98.488	92.922	528.661	9699.953	3.061	-0.846

The RMSE, MAE, MAPE and MASE values of the SARIMA time series model are found to be significantly smaller than that of BSTS (MCMC) model implying that the SARIMA time series model is more robust in its estimations and forecasting abilities while the R-squared is larger for the SARIMA time series model than the BSTS (MCMC) model implying that the SARIMA time series model fitted the data better than the BSTS (MCMC) model.

The result of this study is in line with the conclusion in the work of Gianacsa *et al.*, (2023) in which it cautions that using BSTS, as the choice of explanatory time series may lead to incorrect attribution of outcomes to the study effect. BSTS can incorrectly attribute an existing secular trend to the intervention if the secular trend is not represented in the explanatory time series.

4. Conclusion

The Box-Jenkins methodology identified SARIMA (1,1,0) x (0,1,1)12 as the most appropriate model for the rainfall data series after differencing at level of significance of 0.05. The SARIMA time series model out-perform the Bayesian Structural Time Series model implemented by Monte Carlo algorithm in the fitting of rainfall data in Abeokuta, Ogun State, Nigeria.

This study shows that SARIMA model is a more precise and robust in dealing with this type of dataset than BSTS (MCMC) model. It is better because its computational process using differencing, lags and moving averages ensure that the underlying components of the model are properly identified and estimated. However, BSTS (MCMC) computes uncertainty in a way that measure the posterior uncertainty of the model.

5. References

Aliyu, A. S., Auwal, A. M. and Adenomon, M. O (2021). Application of SARIMA Models in Modelling and Forecasting Monthly Rainfall in Nigeria. *Asian Journal of Probability and Statistics*, 13(3): 30-43.

Almarashi, A. M. and Khan, K. (2020). Bayesian Structural Time Series. *Nanoscience and Nanotechnology Letters*,12(1):54-61

Amaefula, C. G. (2021). A SARIMA and Adjusted SARIMA Models in a Seasonal Nonstationary Time Series: Evidence of Enugu Monthly Rainfall. *European Journal of Mathematics and Statistics*, **2**(1): 13 – 18.

Ayodele, A. P. and Precious, E. E. (2019). Seasonal Rainfall Prediction in Lagos, Nigeria Using Artificial Neural Network. *Asian Journal of Research in Computer Science*, 3(4): 1 – 10

Bari, S. H., Hussain, R. and Sourov, M. M. (2015). Forecasting Monthly Precipitation in Sylhet City Using ARIMA Model. *Civil and Environment Research*, **7**(1): 78-82

Broemeling, L.D. (1985). Bayesian Analysis of Linear Models. Marcel Dekker, New York. 23: 66-69.

Box, G. E. P. and Jenkins, G. M. (1976). Models for Forecasting Seasonal and Non-Seasonal Time Series in Spectral Analysis of Time Series. New York Wiley.

Box, G. E. P. and Jenkins, G.M. (2004). Time Series Analysis Forecasting and Control. New Jersey: Prentice Hall International Inc.

Chonge, M., Nyongesa, K., Omukaba, M., Makokha, L. and Tireito, F. (2015). Time Series Model of Rainfall Pattern of Uasin Gishu County. *Journal of Mathematics*, 11(5): 77-84.

Cowden, R. J., Watnika, D. W. and Milhucic, R. J. (2010). Stochastic Rainfall Modelling in West Africa: Parsimonious Approaches for Domestic Harvesting Assessment. *Journal of Hydrology*, 36(1): 64-77.

Cryer, J. D. and Chan, K. S. (2008) Time Series Analysis: With Applications in R. Springer. Science, Business Media, London.

Dabral, P. P. and Murry, M. Z. (2017). Modeling and Forecasting of Rainfall Time Series Using SARIMA. *Environmental Processes*, 4(2): 399 – 419.

Gianacas, C., Liu, B., Kirk, M., Tanna, G. L. D., Belcher, J., Blogg, S. and Muscatello, D. J. (2023).
Bayesian Structural Time Series, An Alternative to Interrupted Time Series in the Right Circumstances, *Journal of Clinical Epidemiology*, 163: 102-110.

Hung, C. (2017) Modelling weekly temperatures in Eelde using Bayesian Structural Time Series. *Johann Bernoulli Institute for Mathematics and Computer Science, the Netherlands*, 23: 9-14.

Jayawardene, H. K. (2005). Trends of Rainfall in Sri Lanka over the Last Century. Sri Lanka. *Journal of Physics*, 6**:** 7-17

Katarina, B., Gunardi (2023). Optimization of Bayesian Structural Time Series (BSTS) Applications in Forecasting Stock Prices Through State Components Selection. Proceedings of the 8th International Conference on the Applications of Science and Mathematics. EduTA 2022. Springer Proceedings in Physics, vol 294. Springer, Singapore. https://doi.org/10.1007/978-981-99-2850-7_20

Lee, A. F. S. and Heghinian, S. M. (1977). A Shift of the Mean Level in a Sequence of Independent Normal Random Variables. A Bayesian Approach. *Asia Stat Association*, 19(4): 503-506.

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000) WinBUGS – A Bayesian Modelling Framework: Concepts, Structure and Extensibility. *Statistics and Computing*, 10:325–337.

Narasimha, M. K. V., Saravana, R. and Vijaya K. K. (2018). Modeling and Forecasting Rainfall Patterns of South-West Monsoons in North–East India as a SARIMA Process. *Meteorology and Atmospheric Physics*, 130: 99–106.

Nwokike, C. C., Offorha, B. C., Obubu, M., Ugoala, C. B. and Ukomah, H. I. (2020). Comparing SANN and SARIMA for Forecasting Frequency of Monthly Rainfall in Umuahia. *Scientific African (Elsevier)*, 10**:** 1 – 15.

Mohammed, S. J. and Alehile, K. S. (2022). Impact of Climate Change on Nigerian Agricultural Sector Crop Production. *Journal of Economics and Allied Research*, **7**(1): 105 -115.

Obot, N. and Onyeukwu, N. O. (2010). Trend of Rainfall in Abeokuta, Ogun State, Nigeria: A 2-
Year Experience (2006 – 2007). *Journal of Environmental Issues and Agriculture in Developing Countries*, 2(1): 70 – 81.

Ogundari, K., Ademuwagun, A. A. and Appah, O. (2021). A Note on Rainfall Variability and Trends in Nigeria: Implications for Agricultural Production, *Studies in microeconomics*, 11(3): 1 – 11.

Ogundeji, R. K, Onyeka-Ubaka, J. N. and Akanji, R. A. (2021). Bayesian GARCH Models for
Nigeria Covid-19 Data. *Annals of Mathematics and Computer Science*, 4: 14-27.

Ogundeji, R. K, Onyeka-Ubaka, J. N. and Yinusa, E. (2022)**.** Comparative Study of Bayesian and
Ordinary Least Squares Approaches. *Unilag Journal of Mathematics and Application*, 2(1): 60 – 75.

Ogundeji, R. K., Onyeka-Ubaka, J. N. and E. O. Idowu, E. O. (2023). Modeling and Forecasting
Exchange Rates Volatility Using Selected GARCH Models, *Benin Journal of Statistics*. 6: 84– 102.

Ogungbenro, S. B. and Morakinyo, T. E. (2014). Rainfall distribution and change detection across climatic zones in Nigeria, *Weather and Climate Extremes*, 5(6): 1 – 6.

Ogunrinade, E. D. (2012). Time Series Analysis of Rainfall South-West Part of Nigeria. *Journal of Pure and Applied Sciences*, 22(2): 70 – 75.

Onyeka-Ubaka, J. N. Halid, M, A, and Ogundeji, R. K. (2021). Optimal Stochastic Forecast Models of Rainfall in South-West Region of Nigeria. *International Journal of Mathematical Analysis and Optimization: Theory and Applications*, **7**(2):1 – 20.

Qiu, J.; S.; Rao Jammalamadaka, S. R. and Ning, N. (2018). Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research*, 19, 1-33.

Scott, S. L. and Varian, H. (2013). Predicting the Present with Bayesian Structural Time Series, Pearson Press, Boston, USA.

Shumway, R. H. and Stoffer, D. S. (2008): Time series analysis and its applications with R examples, 2nd Edition. *AStA Advances in Statistical Analysis* 92: 233–234.

Xinghua, C., Meng, G., Yan, W. and Xiyong, H. (2012). Seasonal Autoregressive Integrated Moving, Average Model for Precipitation Time Series. *Journal of Mathematics and Statistics* 8(4): 500-505.