Bima Journal of Science and Technology, Vol. 8(2B) July, 2024 ISSN: 2536-6041

DOI: 10.56892/bima.v8i2B.736

Time Series Analysis on Malaria Disease and Contorl in Bauchi State

Usman Rabiu Haruna[2], U.F Abbas[2], M. Abdulhamid[2], Musa Bawa[2] Abdulkadir Abdulkadir[1]

[1]Department of Mathematical Science, Faculty of Science, Gombe State University, Nigeria
[2]Department of Mathematical Science Abubakar Tafawa Balewa University, Bauchi, Nigeria

Corresponding Author: rabiuharun@gmail.com

ABSTRACT

In many developing countries in Africa, the control of non-infectious diseases can be quite challenging due to a combination of factors such as poor housing, inadequate health care facilities, and poor sanitation. Nigeria has been recorded as the second highest country facing such challenges. Despite the measures that were taken by individuals, government and N.G.Os through vector control interventions with the use of insecticide-treated bed nets (ITN) and indoor residual spraying (IRS), but seasonal cases of Malaria couldn't be eradicated completely. Furthermore, 120 samples (monthly) data were captured and analyzed for 10year using R-Statistical Software from the Bauchi State Agency for the control of HIV/AIDS, Tuberculosis, Leprosy and Malaria's (BACATMA) register and different Time Series Model(TSM) were estimated and compared to explain the scenario of malaria cases in Bauchi. ARIMA, SARIMA, and ARFIMA models were fitted separately, model parameters were estimated and the model with the best fit was selected using AIC in each of the separate models, and model stationarity and independence were analyzed using the moving average method. Among these good models, an optimal model was chosen using both Likelihood test and AIC statistically. The ARFIMA(Phi) model performs better than ARIMA(1,0,1) and SARIMA(1,1,1). Unit root tests were conducted using ADF and PP tests. Results confirmed that the data series is stationary since the p-values 0.01 are less than alpha (0.05) for ADF test statistics and test statistics -5.3369 and -4.2434 which is less than critical-value (0.148) at 5% level of significance for PP test respectively and variables in the dataset were tested for normality using the Jargue-Bera test. The result showed that none of the variables is normally distributed as indicated by the p-values for all variables being less than selected level of significance for the study (0.05). The forecasting model suggests that the number of malaria cases in Bauchi state is decreasing probably to extinction. The model has used the moving average method and trend analysis for time series analysis to predict the future rate of malaria cases in the state.

Keywords: Time Series Model, Malaria, ARIMA, SARIMA, ARFIM, Bauchi

INTRODUCTION

Malaria is a vector-borne infectious disease caused by the Plasmodium parasite and transmitted to humans through the bites of infected Anopheles mosquitoes. It is a major public health problem, with an estimated 229 million cases and 409,000 deaths worldwide in 2019, mainly in sub-Saharan Africa (World Health Organization [WHO], 2020). Malaria is a complex disease influenced by a variety of factors, including climate, geography, host genetics, and human behavior, making the control and eradication processes a difficult challenge. Malaria control efforts have been focused on reducing the burden of the disease through prevention and treatment. Prevention measures include vector control, such as insecticide-treated bed nets (ITNs) and indoor residual spraying (IRS), and chemoprophylaxis for high-risk populations. Treatment includes the use of antimalarial drugs, such as artemisinin-based combination therapies (ACTs). Despite significant progress in malaria control and elimination in some parts of the world, challenges remain, including the emergence of drug-resistant strains of the parasite and insecticide-resistant mosquito populations. In addition, climate change and other environmental factors have the potential to influence the transmission dynamics of the disease and further complicate efforts to control and eliminate it (Paaijmans, Blanford, & Bell, 2019).

To address these challenges mentioned above and improve malaria control and suggest possible elimination efforts, ongoing research is focused on improving our understanding of the epidemiology, transmission dynamics, and pathogenesis of the disease. This includes studies on the genetic diversity of the parasite (Ménard et al., 2018), mosquito biology and ecology (Cator et al., 2019), and the social and behavioral factors that influence human exposure and vulnerability to the disease (Das et al., 2019).

In summary, malaria remains a significant global public health challenge, with ongoing efforts focused on reducing the burden of the disease through prevention and treatment. However, the complexity of the disease and the challenges associated with the disease control and elimination require ongoing research and innovation to advance our understanding and develop new tools and strategies to overcome these obstacles.

Malaria is a significant public health concern in Nigeria, with an estimated 100 million cases reported annually, leading to about 300,000 deaths, with a high burden in Bauchi state (Ntaji et al., 2019). This disease is transmitted through the bite of infected female Anopheles mosquitoes and is caused by the Plasmodium parasite. The transmission of malaria in Bauchi state is influenced by various factors, including climate change, socio-economic factors, and others.

Climate change is one of the primary factors influencing malaria transmission in Bauchi state. With the increase in temperature and rainfall, the breeding sites of mosquitoes increase, leading to an increase in the number of malaria cases (Okeke et al., 2019). Socio-economic factors such as poor housing, inadequate health care, and poor sanitation are also contributing to the high incidence of malaria in Bauchi state (Adebayo et al., 2018). Additionally, inadequate vector control interventions such as the use of insecticide-treated bed nets and indoor residual spraying (IRS) have led to the persistence of malaria transmission in the region (Ntaji et al., 2019).

LITERATURE REVIEW

Sub-Saharan Africa remains the world's region with the greatest malaria burden despite massive efforts over the past decades to lower or eliminate malaria (WHO, 2020). Though poor health care systems and low socio-economic status (Degarege et al., 2019; Yadav et al., 2014) are contributing factors, the climate suitability of the region for malaria transmission has a major influence (Caminade et al., 2014). Generally, climate variables such as temperature, rainfall and relative humidity are known to have a significant influence on the development and survival of both the malaria parasites and their vectors. Malaria parasite development is not possible at temperatures below 16°C and temperatures above 40°C have adverse effects on mosquito population turnover (Blanford et al., 2013; Mordecai et al., 2013; Parham & Michael, 2010; Shapiro et al., 2017). Rainfall provides the environment for vector breeding (Ermert et al., 2011; Kar et al., 2014; Tompkins & Ermert, 2013) and relative humidity of at least 60% appears necessary for vector survival (Thompson et al., 2005). Rainfall, therefore, affects the availability, persistence and dimensions of Anopheles vectors and their larval habitats (Afrane et al., 2012; Asare, Tompkins,

Amekudzi, & Ermert, 2016; Boyce et al., 2016; Fournet et al., 2010). Previous work studying the relationship between sporozoite development and the survival of infectious mosquitoes found optimal temperatures for efficient malaria transmission between 25°C and 27°C (Bayoh, 2001; Lunde, Bayoh, & Lindtjorn, 2013; Lunde, Korecha, et al., 2013).

In Sub-Saharan Africa, most countries have annual mean temperatures between 20°C and 28°C (Lunde, Bayoh, & Lindtjorn, 2013). Given Sub-Saharan Africa's warm tropical climate, a plethora of efficient and effective malaria parasite and vectors thrive in this setting (Murray et al., 2012; Sinka et al., 2010). Understanding the relative importance of climate drivers of malaria seasonality is crucial for describing the geographic patterns of the heterogeneous risk and burden of malaria across the sub-region (Gething et al., 2011; Reiner et al., 2015). This could translate to substantial public health gains, taking into account the seasonality in malaria control and prevention interventions, by helping to determine when, where and how to apply vector and parasite control measures.

An overview of the malaria disease and its causes; Malaria is a parasitic disease that is transmitted to humans through the bite of infected female Anopheles mosquitoes. Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, and Plasmodium ovale are the four primary malaria parasites that cause infection in humans, whereas Plasmodium knowlesi is a zoonotic species prevalent in Southeast Asia (Awosolu et al., 2021). P. falciparum is the most pathogenic and, together with P. vivax, causes most death, while P. ovale and P.malariae cause a milder form of malaria that is rarely lethal (Beare et al., 2006; Gething et al., 2012). Even though preventable, malaria has continued to cause significant morbidity and mortality worldwide, especially in sub-Saharan Africa. The breeding of mosquitoes and the spread of malaria are aided by the environmental conditions of tropical and subtropical African countries, such as constant high temperature, humidity and copious stagnant waters due to poor drainage systems (Ogomaka, 2020). An estimated 3.2 billion individuals worldwide are at risk of acquiring malaria each year (Awosolu et al., 2021). Furthermore, in 2017, nearly 219 million cases were reported in 87 countries, resulting in approximately 435,000 deaths (WHO, 2018). Malaria caused the majority of worldwide morbidity and mortality, with nearly 3.1 billion dollars spent on malaria control and elimination projects in Sub- Saharan Africa in 2017 (WHO, 2018). Malaria is spread all year in Nigeria, where more than 194 million people are at risk of contracting malaria disease. As a result, Nigeria had the highest malaria prevalence worldwide in 2007(FMOH, 2009). Malaria significantly contributes to the rise in hospital visits throughout Nigeria's six geo-political zones (Adeyemo et al., 2014). The Nigerian Federal Ministry of Health estimated 110 million clinical malaria cases each year, costing the country around NGN 132 billion in treatment and preventative measures (FMOH, 2009). The intensity of malaria transmission is determined by demographic (age and gender) and environmental factors (presence or absence of bushes and forests that enhance mosquito breeding). Climate elements such as temperature, humidity, and rainfall, which may favour mosquito vectors' rapid growth and development, are also risk factors for malaria transmission. These factors have been well reported in previous studies in Nigeria (Gunn et al., 2015; Morakinyo et al.,

DOI: 10.56892/bima.v8i2B.736

2018) and elsewhere (Dejazmach et al., 2021; Kalinga et al., 2019) regarding the prevalence of malaria infection. Most malaria studies and interventions in Nigeria have focused on pregnant women (Agomo et al., 2009; Fana et al., 2015) and infants (Morakinyo et al., 2018; Olasehinde et al., 2010). It is challenging to create risk-based preventive interventions in our locality.

In Bauchi State, malaria transmission occurs throughout the year, but the peak transmission season is during the rainy season, which lasts from May to October. The predominant malaria parasite species in the state are Plasmodium falciparum and Plasmodium vivax (Duru et al., 2018). The transmission of malaria is influenced by several factors, including climate, topography, and human behavior.

The Climate and Topography; Bauchi State is located in the savannah region of Nigeria, with a tropical climate characterized by two distinct seasons: the rainy season and the dry season. The rainfall pattern in the state is bimodal, with a major peak in July and a minor peak in September. The topography of the state is characterized by plateaus and hills, with altitudes ranging from 200 to 2,000 meters above sea level. The highland areas are cooler and have lower mosquito densities compared to the lowland areas (Gyamfi et al., 2019).

The Human Behavior; Human behavior also plays a significant role in malaria transmission in Bauchi State. Many people in the state do not use insecticide-treated bed nets (ITNs) consistently, and some do not use them at all. Additionally, many people in the state do not seek prompt and effective treatment for malaria when they experience symptoms, which can lead to complications and death (Aworh et al., 2018).

Emperical Studies

A study conducted by Ojurongbe et al. in 2023 analyzed the prevalence of malaria in Osogbo, Nigeria using time series analysis. They identified an appropriate model, ARIMA (4, 1, 4), to describe the trend of malaria and forecast its future occurrence in the region. The study suggests that based on the data from May 2002 to December 2019, malaria prevalence in Osogbo may not disappear but may remain relatively constant in the next few years, approximately 5 years.

In the study conducted by Adekola et al. in the Al-Hayat: Journal of Biology and Applied Biology to assess the prevalence of malaria in peri-urban areas. They performed a time-series analysis of 1,141 samples with suspected febrile illness that were collected over a year from a peri-urban health center in the study area. The findings revealed that 273 individuals were seropositive for malaria, with males (162) having a higher prevalence than females (111). The study used a simple percentile and shows an overall prevalence of 24% ($p<0.05$).

In a similar study by Segun et. al. (2020), conducted a research to investigate the impact of weather factors on malaria occurrences in Abuja. The study analyzed the statistical relationship between weather variables and malaria incidence using the monthly data collected in Abuja between 2000 and 2013. The analysis was based on generalized linear models, and Pearson correlation was undertaken at the bivariate level. The study found that the months with the highest rainfall (June-August) recorded more cases of malaria due to an increase in humidity and rainfall. In contrast, a significant decrease in malaria was observed due to temperature increase in the study area at different lags in the time series analysis. The study recommends the need to prioritize vector control activities and to create public health awareness about the proper usage of intervention measures.

In the study to determine an appropriate statistical technique for forecasting the monthly Malaria cases in Ghana, Twumasi-Ankrah et. al., 2019 obtained the monthly data on malaria spanning from January 2008 to December 2017 from the District Health Information Management System (DHIMS) 2, Ghana Health Service where four(4) unique forecasting techniques were applied to the Malaria cases data. These techniques include the Seasonal Autoregressive Integrated Moving Average (SARIMA), Artificial Neural Network (ANN), Exponential smoothing (ETS) and a Combination technique. The four competing forecasting techniques were compared using their respective forecast accuracy measures in order to choose the appropriate technique for forecasting Malaria cases in Ghana and results shows that SARIMA technique was the appropriate statistical technique which happens to be the “best” model for forecasting the monthly malaria cases in Ghana. The fitted model was SARIMA (2, 1, 0) (2, 0, 0)12 which passed all the required diagnostic tests of AIC, Root-Mean-Square Error (RSME) and the Mean Absolute Percentage Error (MAPE) where applicable.

Twumasi-Ankrah et. al., (2019), collected monthly data on malaria cases in Ghana from January 2008 to December 2017 from the District Health Information Management System (DHIMS) 2, Ghana Health Service. The researchers applied four unique forecasting techniques to the malaria cases data, including the Seasonal Autoregressive Integrated Moving Average (SARIMA), Artificial Neural Network (ANN), Exponential smoothing (ETS), and a Combination technique. The researchers compared the four forecasting techniques using their respective forecast accuracy measures to determine the appropriate technique for forecasting malaria cases in Ghana. The results showed that the SARIMA technique was the best statistical technique for forecasting the monthly malaria cases in Ghana. The fitted model was SARIMA (2, 1, 0) (2, 0, 0), which passed all the required diagnostic tests of AIC, Root-Mean-Square Error (RSME), and the Mean Absolute Percentage Error (MAPE) where applicable.

Conceptual Framework

A conceptual framework outlines the key concepts, variables, and relationships between them. In the context of malaria transmission and control in Bauchi State, Nigeria, a possible theoretical framework could be:

The main concept is malaria transmission and control, which is influenced by several factors, including climate, topography, and human behavior. These factors can be further broken down into sub-concepts and variables, such as:

Climate:

I Temperature

II Rainfall

II Humidity

Topography:

I Altitude

II Vegetation cover

III Water bodies

Human behavior:

I Use of insecticide-treated bed nets (ITNs)

II Indoor residual spraying (IRS)

III Prompt and effective treatment of malaria

These factors can interact with each other in complex ways, leading to variations in malaria

DOI: 10.56892/bima.v8i2B.736

transmission and control across different regions of Bauchi State. For example, in areas with high rainfall and humidity, mosquito populations may be higher, and the risk of malaria transmission may be greater. In areas with low altitude and dense vegetation cover, mosquito breeding sites may be more abundant, increasing the risk of malaria transmission.

To control malaria transmission and reduce the burden of the disease in Bauchi State, effective interventions must be implemented. These interventions can be conceptualized as part of the theoretical framework, with the following variables:

I Distribution of insecticide-treated bed nets (ITNs)

II Indoor residual spraying (IRS)

III Prompt and effective treatment of malaria

The effectiveness of these interventions may depend on several factors, including the level of community engagement and participation, the availability of resources and infrastructure, and the strength of the health system.

Overall, the theoretical framework for malaria transmission and control in Bauchi State highlights the complex interactions between environmental, social, and health-related factors that influence the spread of the disease. By understanding these factors and their relationships, policymakers and public health practitioners can design and implement effective interventions that can reduce the burden of malaria in Bauchi State.

MATERIALS AND METHODS

AR (p) model

This is the kind of model that calculates the regression of past time series and calculates the present or future values in the series. The model is defined as:

$$X_t = \sum_{i=1}^{p} \varphi_1 X_{t-i} + \varepsilon_t$$

Where X_t is the value of the time series, c is the constant (intercept) term, $\theta_1 ... \theta_i$ are the autoregressive coefficients for lag 1 to lag i, x_{i-1}, x_{i-2}, **…..,** y_{i-p}, are the lagged values of the time series, ε_i is the error term or white noise at time 0f i.

The AR (p) model assumes that the current value of the time series is a linear combination of its past p values, adjusted by the autoregressive coefficients, plus some random noise.

MA (q) model

Moving average (MA) models uses past forecast errors rather than values in a regression. The model is defined as

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

where the $\theta_1 ... \theta_p$ are the parameters of the model, μ is the expectation of X_t (often assumed to be zero), and then ε_t, ε_{t-1}…**.** Are again, white noise error terms.

ARMA (p,q) model

Autoregressive moving average (ARMA) is a model of forecasting in which the methods of auto regressive (AR) model and moving average (MA) model are both applied to time series data that is well behaved. In ARMA it is assumed that the time series is stationary and when it fluctuates, it does so uniformly around a particular time. The model is given by

:

DOI: 10.56892/bima.v8i2B.736

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

Where X_t is the value of the time series, m is a constant (intercept) term, $\theta_1 ... \theta_i$ are the moving average for lag 1 to lag i, $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$ are the lagged values of the time series, $\varepsilon_{t-I}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-p}$, are the past forecast errors (residuals), ε_t is the error term or white noise at time t

The ARMA (p, q) model combines the autoregressive and moving average components to capture both the dependencies on past values of the time series and past forecast errors. It's important to note that while the ARMA model can capture many patterns in time series data, it assumes that the underlying process is stationary. If the data is non-stationary, you might consider using the more comprehensive ARIMA (Autoregressive Integrated Moving Average) model, which includes an integration component to handle non-stationary data.

ARIMA (p,d,q) model

Autoregressive integrated moving average (ARIMA) is the generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better un derstand the data or to predict future points in the series (forecasting). The model is giving by:

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right) X_t = \left(1 - \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t$$

Where Y_t is the value of the time series at time t, μ is the mean of the time series, $\Phi_1, \Phi_2, \ldots, \Phi_p$ are the autoregressive coefficient of lags 1 through p, $\Theta_1, \Theta_1, \ldots, \Theta_k$, are the moving average coefficient at lags 1 trough q**,** ε_t is wide noise error term at time t.

SARIMA (p,d,q) model

SARIMA models takes seasonality into account by essentially applying an ARIMA model to lags that are integer multiples of seasonality. Once the seasonality is modeled, an ARIMA model is applied to the leftover to capture non-seasonal structure. The model is given by:

$$\Phi_p(L^s)\varphi_p(L)z_t = \Theta_Q(L^s)\theta_q(L)\varepsilon_t$$

z_t is the seasonality differenced series.

ARFIMA (p,d,q) model

An ARFIMA model (AutoRegessive Fractionally Integrated Moving Average) model is a time series model that extends the traditional ARIMA model to account a long range dependence in the data. It is characterized by three component: AutoRegressive (AR) Component, the Fractional Integration (I) Moving Average (MA) component.

***AutoRegessive (AR) componen*t:** this represent the correlation between the current value of the time series and its past values. It is donated by AR(p) where p is the order of

Bima Journal of Science and Technology, Vol. 8(2B) July, 2024 ISSN: 2536-6041

DOI: 10.56892/bima.v8i2B.736

AutoRegessive component. The formula for AR(p) is:

$$X_t = \sum_{i=1}^{p} \varphi_1 X_{t-i} + \varepsilon_t$$

Where X_t is the value of the time series, c is the constant (intercept) term, $\theta_1 ... \theta_i$ are the autoregressive coefficients for lag 1 to lag i, x_{i-1}, x_{i-2}, ….., y_{i-p}, are the lagged values of the time series, ε_i is the error term or white noise at time 0f i

The AR (p) model assumes that the current value of the time series is a linear combination of its past p values, adjusted by the autoregressive coefficients, plus some random noise.

***Fractional Integration (I) component*:** this account for long range dependence and is donated by I(d), where d is the differencing parameter. The fractional integration component introduces fractional differences in the time series, and the formula is:

$$D^d Y_t = \varepsilon_t$$

Where D^d represent a fractional difference operator, Y_t is the difference time series, and ε_t is the wide noise.

Moving average (MA) component: This represent the correlation between the current values of the time series and the past wide noise or error terms. It is donated by MA(q), where q is the order of the moving average component. The formula for MA(q) is:

:

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

Where the $\theta_1 ... \theta_p$ are the parameters of the model, μ is the expectation of X_t (often assumed to be zero), and then ε_t, ε_{t-1}…. Are again, white noise error terms.

ARFIMA model combined these three component to capture the dynamics of the time series while considering fractional integration. The model is specified as ARFIMA (p,d,q) where p is the AR order, d is the differencing parameter and q is the MA order

Performance Measure

We employed a number of measures to evaluate the variance prediction of a specific model by comparing the model-predicted variance with the monthly Rv estimated as the sum of the squared daily log returns within each month. We used two loss functions: the Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE).

Mean Square Error

The mean square error is a quadratic loss function and gives a larger weight to large prediction errors. The estimator measures the average of the squares of the errors, that is, the average squared difference between the estimated values and actual values, which can be defined as

$$MSE = \frac{1}{T}\sum_{t=1}^{T} (\delta_{t+1}^2 - E(\delta_{t+1}^2))^2 \qquad (3.29)$$

T-number of data point

δ_{t+1}^2 − Predicted value

$E(\delta_{t+1}^2)$ − Observed values

Mean Absolute Percentage Error

The Mean Absolute Percentage Error is a measure of how accurate a forecast system is; it measures this accuracy as a percentage and can be calculated as the average absolute percent error for each time period minus actual values divided by actual values.

$$M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{3.30}$$

n- is the number of fitted points

A_t -is the observed value

F_t -is the Predicted value

Diagnostics Test

Diagnostic tests for normality, long-memory test and unit root test will be carried out for the estimated model.

Jarque-Bera Test

Jarque and Beran (1987) provide a method for investigating the normality of the time series models' residuals. The null hypothesis can be defined as the time series not normally distributed and the test statistics being given as;

$$\text{JB}= n\ [(\sqrt{b_1})^2/6 + (b_2 - 3)^2/24 \tag{3.31}$$

where:

n- is the sample size,

$\sqrt{b_1}$ - is the sample skewness coefficient,

b_2 - is the kurtosis coefficient

Unit Root Test

Testing for unit roots has recently become a standard procedure in time series studies and used to determine if trending data should be first differenced or regressed on deterministic functions of time to render the data stationary.

The unit root test is important because the absence of a unit root indicates that the series has some variances that are not being determined by time and that the effects of shocks dissolve over time. Besides that, the existence of non-stationary variables will cause spurious regression, which has a high t-statistic and is significant, but the results will not have any economic meaning. We are going to adopt Augmented Dickey Fuller test (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for testing the presence of unit root in the data set.

The Augmented Dickey Fuller test

The Dickey-Fuller (DF) test is used for testing the presence of a unit root. The approach involves testing the null hypothesis that a series does contain a unit root (i.e., it is non-stationary) against the alternative of stationarity. The ADF test is comparable with the simple Dickey-Fuller test, but it involves adding an unknown number of lagged first differences of the dependent variable to capture autocorrelated omitted variables that would otherwise, by default, enter the error term. In this way, one can validly apply unit root tests when the underlying data-generating process is quite general.

$$\Delta y_i = \beta y_{i-1} + \varepsilon_t \quad (3.32)$$

where $\beta \leq 0$, ε_t is the error term, $\beta -$ coefficients and $\Delta y_i = (y_i - y_{i-1})$

Kwiatkowski-Phillips-Schmidt-Shin Test

Kwiatkowski et al. (1992) proposed that trend stationarity should be considered the null hypothesis and the unit root should be the alternative. Rejection of the null of trend stationarity could then be viewed as convincing evidence in favour of a unit root. It was soon realised that the KPSS test of Kwiatkowski et al. (1992) has a much broader utility. For example, Lee and Schmidt (1996) and Giraitis et al. (2003) used it to detect long memory, with short memory as the null hypothesis; de Jong et al. (1997) developed a robust version of the KPSS test. The KPSS test is based on linear regression. It breaks up a series into three parts: a deterministic trend (βt) a random walk (r_t), and a stationary error (ε_t), with the regression equations :

$$X_t = r_t + \beta_t + \varepsilon_t \quad (3.33)$$

If the data is stationary, it will have a fixed element for an intercept, or the series will be stationary around a fixed level.

Akaike Information Criterion

The idea of the Akaike information criterion (AIC) is based on the minimization of the Kullback-Leibler distance but to penalize the high-dimensional models by finding some $\tilde{K}*$ such that the expectation of $\tilde{K}*$ is asymptotically identical to the expectation of $\hat{K}$, i.e., $E(\tilde{K}* - \hat{K}) = 0$, while $\dim(\theta)$ is considered as penalty, and also, AIC can be write as:

$$\text{AIC} = -2\sum_{i-1}^{n}(\ln f \frac{y_t}{x_t};\theta) + 2p \quad (3.35)$$

Bayesian Information Criterion (Schwarz Criterion)

Another criterion for model selections is the Bayesian Information Criterion (BIC), also known as Schwarz Criterion. Schwarz (1987) developed the criterion from Bayesian likelihood maximization. Schwarz also proved that the BIC is valid since it does not depend on the prior distribution. The BIC is defined as

$$\text{BIC} = l(\theta) - \frac{p}{2}\ln(n) \quad (3.36)$$

where $l(\theta)$ is the log-likelihood, $p = \dim(\hat{\theta})$ and n is the number of observations.

RESULTS AND DISCUSSION

Time Series Graph of Malaria Cases

The time series plot which display observation on the Y-axis against equally time interval on the X-axis used to evaluate patterns and

Bima Journal of Science and Technology, Vol. 8(2B) July, 2024 ISSN: 2536-6041

DOI: 10.56892/bima.v8i2B.736

behaviors of the data over the time is displayed below:

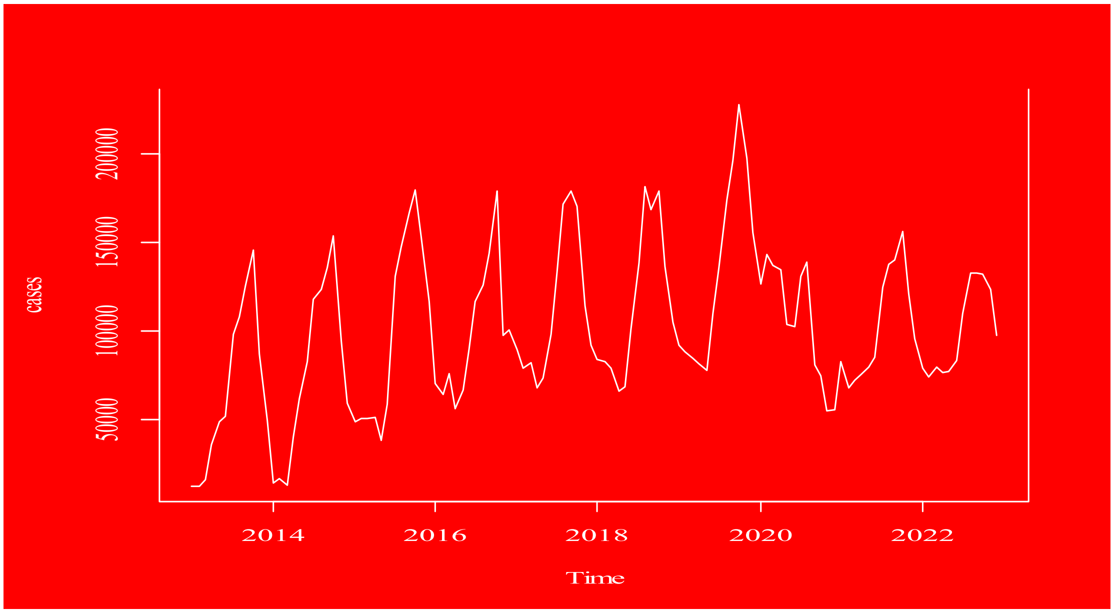


Figure 1: Time series plot of cases of malaria

Test of Stationary

Stationarity Tests

In line with the methodology of the study, unit root tests were conducted using ADF and PP tests. This is necessary in order to determine the nature of the series and to avoid getting spurious results. The table below summarizes the results of the tests.

Table 1: Stationarity Test on the Original Series (cases of malaria)

Unit root test	Test statistics	p-value	Critical -value
ADF test	-5.3369	0.01	
PP test	-4.2434	-	0.148

At level the result confirmed that the data series is stationary since the p-values (0.01) is less than alpha (0.05) for ADF test statistics and test statistics -5.3369 and -4.2434 which is less than critical-value (0.148) at 5% level of significance for PP test respectively

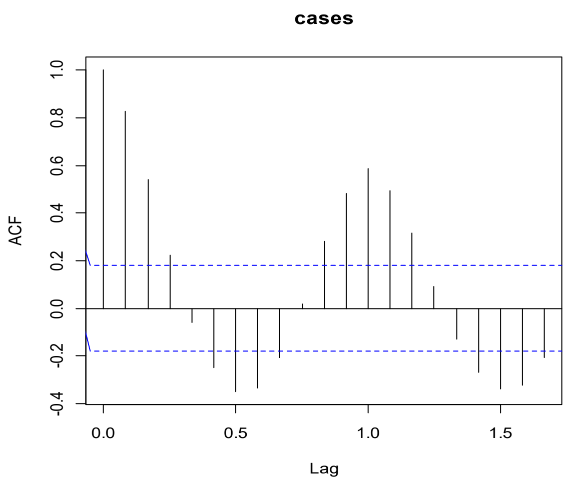
Table 2: Jargue Bera Test of Normality for the variables

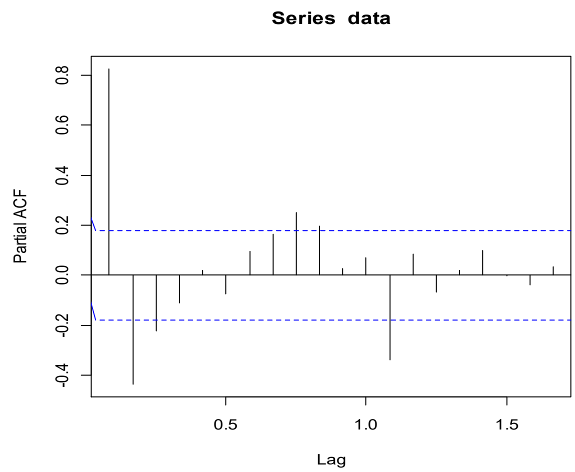
Variable	Statistics	p-value
Malaria cases	1.7151	0.4242

First of all, variables in the dataset were tested for normality using the Jargue Bera test. The result showed that none of the variables is normally distributed as indicated by the p-values for all variables being less than selected level of significance for the study (0.05). The Jargue Bera statistics for normality are presented in .Hence, we would reject the null hypothesis that variables are normally distributed.

DOI: 10.56892/bima.v8i2B.736

Test Present of Autocorrelations functions and Partial Autocorrelations functions





The ACF plot dies down slowly. The autocorrelation function provides a measure of temporal correlation between data points with different time lags. For a purely random event, all autocorrelation coefficients (r) are zero, apart from r(0), which is equal to 1.

The ACF of a stationary ARMA process falls exponentially with a rising time lag. However, these Malaria cases time series variables are strongly correlated. A time series with this attribute is referred to as a series with long-range dependence and is generated by stochastic processes called processes with a long memory. It is usually characterized by slowly decaying ACF values with an increasing time lag, which is hyperbolic rather than exponential decay.

The ACF plot dies down slowly. The autocorrelation function provides a measure of temporal correlation between data points with different time lags. For a purely random event, all autocorrelation coefficients (r) are zero, apart from r(0), which is equal to 1.

The ACF of a stationary ARMA process falls exponentially with a rising time lag. However, malaria case time series variables are strongly correlated. A time series with this attribute is referred to as a series with long-range dependence and is generated by stochastic processes called processes with a long memory. It is usually characterized by slowly decaying ACF values with an increasing time lag, which is hyperbolic rather than exponential decay.

Table 3: Results for estimated model parameters: ARIMA (p, d, q) Malaria cases

Models	Coefficient	Estimated parameter	AIC
ARIMA (0,1,1)	MA1	0.3095	2739.93
ARIMA (1, 0, 1)	AR1	0.3802	2739.86
	MA1	0.8410	
ARIMA (2,0,0)	AR1	0.3607	2740.64
	AR2	0.7645	
ARIMA(1,0,1)	AR1	0.5432	
	MA1	0.8762	2738.9

From on Table 3 above, ARIMA (1, 0, 1) happens to be the best among the four possible models because it has the lowest AIC of 2738.9 .

SARIMA Model

Table 4: Results for estimated model parameters: SARIMA (p, d, q) Malaria cases

Models	Coefficients	Estimates	Log-likehood	AIC
SARIMA	AR1	0.1186	-1332.98	2673.96
	SAR1	0.9631		
	SMA1	-0.6867		

From the table 4, SARIMA(1,1,1) model have been estimated using malaria data series with order (1,1,1) and has the AIC value of 2673.96 less the ARIMA model due to the seasonality in the data.

ARFIMA Model

Table 5: Results for estimated model parameters: ARFIMA (p, d, q) Malaria cases

Models	Coefficients	Estimates	Log-likehood	AIC
ARFIMA	Phi	0.64132	-1194.45	2408.05

From table 5 above, ARFIMA Model was estimated with AIC value of 2408.05 using malaria data series.

Table 6: Estimated model parameters: ARIMA, SARIMA and ARFIMA (p, d, q) Malaria

Models	Coefficients	Estimates	Log-likehood	AIC
ARIMA	AR1	0.5432	**-1543.54**	2738.9
	MA1	0.8762		
SARIMA	AR1	0.1186	-1332.98	2673.96
	SAR1	0.9631		
	SMA1	-0.6867		
ARFIMA	**Phi**	**0.64132**	**-1194.45**	**2408.05**

From table 6 above, a comparison analysis was carried out between the good models in each series of the models built, where ARIMA(1,0,1), SARIMA(1,1,1) and ARFIMA(Phi) were compared and ARIMA(1,0,1) happens to have the highest log-likelihood ratio value of (-1194.45) and has the least AIC value which qualifies it to be the optimal model since the models are not nested and were estimated using the time series data in the work.

Table 7: Comparison Results of Trend Analysis for the Malaria data

Trend	MAPE	MAD	MSD
Linear	57	35292	178940126
Quadratic	**50**	**34644**	**1577686172**

DOI: 10.56892/bima.v8i2B.736

Growth curve	48	35533	2011920176

Based on the table 7: the measure of forecasting between the three accuracy with shows that Quadratic trend is the best among three trend techniques because it has least value of (MAD and MSD) respectively.

Forecast using the quadratic model

Trend Analysis for cases

Data cases
Length 120
NMissing 0

Fitted Trend Equation

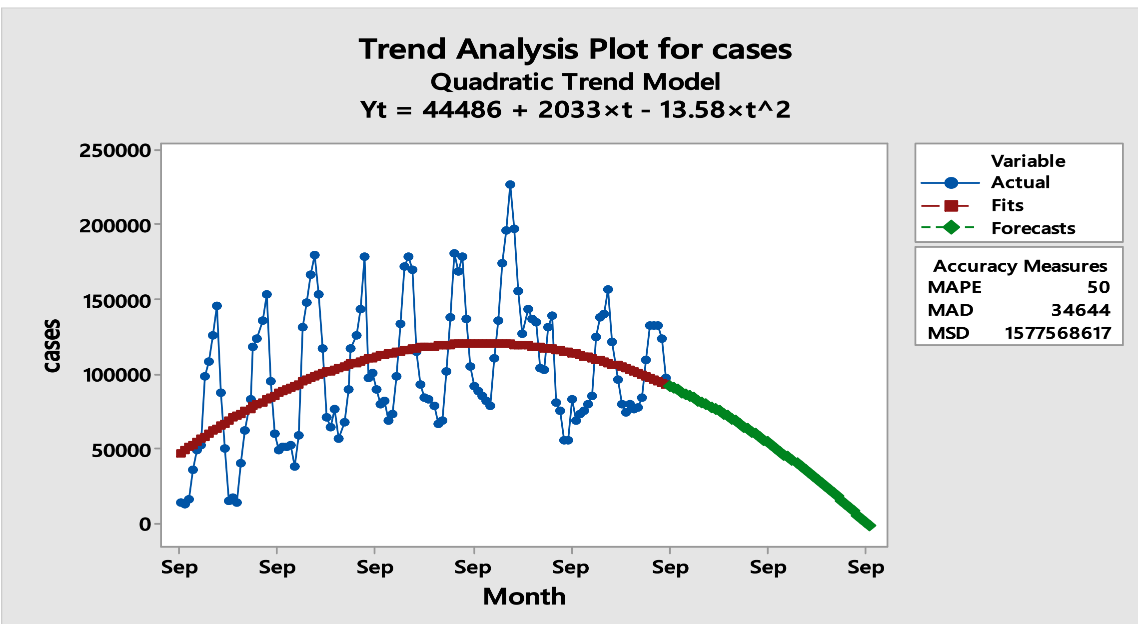
$$Yt = 44486 + 2033\times t - 13.58\times t^2$$

Accuracy Measures

MAPE 50
MAD 34644
MSD 1577568617

Forecasts

Trend Analysis Plot for cases



Discussion

ARIMA(1,0,1) happens to be the best fit for the data among the series of ARIMA(p,d,q) base on (2740.64) AIC criterion, SARIMA(1,1,0) is the best model fit for the data in the series of SARIMA(p,d,q) base on (2673.96) AIC criterion, and ARFIMA (Phi) with 2408.05 is the best model fit for the data as well. Likelihood comparison was carried out and AIC comparison was considered as well and the optimal model was chosen base on accuracy test compared between MAPE, Mad and MSD. A forecast for the next four season was carried out and malaria spread is likely going to extinction.

CONCLUSION

The literature review suggests that ARIMA, SARIMA, and ARFIMA models each have their strengths in capturing specific aspects of

DOI: 10.56892/bima.v8i2B.736

financial time series data. However, each model may not be sufficient in capturing the full dynamics of financial time series data, it is recommended that the ARFIMA model be considered as an alternative approach to modelling financial time series data. Future research can also explore the use of other hybrid models that combine different models to capture the full dynamics of financial time series data. Additionally, further research can investigate the effectiveness of the ARFIMA and SARIMA models in different financial markets and with different financial time series data.

Recommendation

It is recommended that the ARFIMA model be applied in practice by financial practitioners too, to test its effectiveness in real-world scenario.

REFERENCES

Adebayo, A. M., Akinyemi, O. O., Cadmus, E. O., & Ajenifujah-Solebo, S. (2018). Knowledge and use of insecticide treated nets among caregivers of under-five children in rural communities in Southwest Nigeria. PLoS ONE, 13(7), e0201463. doi:10.1371/journal.pone.0201463

Adekola, H. A., Egberongbe, H. O., Olanrewaju, M., Onajobi, I. B., Samson, O. J., & Kareem, W. A. (2023). Time-Series Analysis of Malaria Cases Among Suspected Febrile Patients Attending a Peri-Rural Health Centre Between February 2020-January 2021. Al-Hayat: Journal of Biology and Applied Biology, 6(1), 23-30.

Africa fighting malaria (WHO, 2009) Safety of DDT. pub med (online) accessed 18.8.05Available from URL:http://www.nebi.nlm.nih.gar

Ameh, S., Owoaje, E., & Onyeneho, N. (2014). Malaria parasite prevalence and socioeconomic status in Nigeria: A geospatial analysis. Journal of Vector Borne Diseases, 51(3), 171-177.

B. K. (2019). Malaria and poverty: a socio-economic perspe Ménard, D., Dondorp, A., & Fidock, D. (2013). Antimalarial drug resistance: a threat to malaria elimination. Cold Spring Harbor perspectives in medicine, 3(5), a012438.

Paaijmans, K. P., Blanford, J. I., & Bell, A. S. (2019).ctive. Journal of Public Health, 27(6), 687-693.

Blowing hot and cold: Temperature, immunity and vector-borne disease transmission. Acta tropica, 201, 118-130. World Health Organization (WHO). (2020). Malaria. Retrieved from https://www.who.int/news-room/fact-sheets/detail/malaria

Cator, L. J., Lynch, P. A., Read, A. F., & Thomas, M. B. (2019). Do malaria parasites manipulate mosquitoes? Trends in parasitology, 35(8), 590-603. Das, M., Das, D., & Padhi

Dempsey, P. A, Dempsey, A.D, (2000), (fifth edition), Using Nursing Research Process, Critical Evaluation and Utilization, Lippincott Williams and Williams: Miami, Florida.

Gebremariam Hailemikael, (2007). Malaria in children: implications for the productivity of female caregivers in Nigeria, selected paper for the 2002 annual conference of the Nigerian Economic Society.

Malaria Consortium (2016). Our work in Africa. Where we work/Africa. (Access 10 August 2016) http://www.malariaconsortium.org/where-we-work/Africa-mainpage.htm

Malaria in Bauchi (BACATMA, 2010) Malaria Mortchoti pdf. Available from URL:http://www://bacatma.com

Ojurongbe, T. A., Bello, T. A., Adeboye, N. O., Afolabi, H. A., Aduroja, O. O., &

Bashiru, K. A. (2023). Time Series Analysis on the Prevalence of Malaria in Osogbo, Nigeria. Adeleke University Journal of Engineering and Technology, 6(1), 85-95.

Okeke, T. A., & Okafor, H. U. (2021). Spatial distribution and determinants of malaria infection in Nigeria: A systematic review and meta-analysis. Malaria Journal, 20, 239. doi: 10.1186/s12936-021-03804-1

Roll Back Malaria (RBM, 2007) Malaria in Pregnancy, Modern Medicine, p.53- 62,: Nigeria

Segun, O. E., Shohaimi, S., Nallapan, M., Lamidi-Sarumoh, A. A., & Salari, N. (2020). Statistical modelling of the effects of weather factors on malaria occurrence in Abuja, Nigeria. International journal of environmental research and public health, 17(10), 3474.

Twumasi-Ankrah, S., Pels, W. A., Nyantakyi, K. A., & Addo, D. K. (2019). Comparison of Statistical Techniques for Forecasting Malaria Cases in Ghana.

Twumasi-Ankrah, S., Pels, W. A., Nyantakyi, K., & Addo, D. K. (2008). Comparison of Statistical Techniques for Forecasting Malaria Cases in Ghana. J Biostat Biometric App 4 (1): 102.

WHO (2008). World health statistics 2008.Geneva.http://www.who.int/gho/publications/world_health_statistics/EN_WHS08_Full.pdf (Access 11 October 2014).

World Health Organization (WHO). (2021). World malaria report 2021. Retrieved from https://www.who.int/publications/i/item/9789240020472