Science Journal of Applied Mathematics and Statistics
2014; 2(1): 31-41
Published online March 20,2014 (http://www.sciencepublishinggroup.com/j/sjams)
doi: 10.11648/j.sjams.20140201.15

Modelling and Forecasting Malaria Mortality Rate using SARIMA Models (A Case Study of Aboh Mbaise General Hospital, Imo State Nigeria)

Ekezie Dan Dan[1], Opara Jude[1], Okenwe Idochi[2]

[1]Department of Statistics, Imo State University, Owerri, Nigeria
[2]Department of Statistics, Rivers State Polytecnic, Bori, Rivers State, Nigeria

Email address:

danekezie@yahoo.com (E. D. Dan), judend88@yahoo.com (O. Jude), nwonda@yahoo.com (O. Idochi)

To cite this article:

Ekezie Dan Dan, Opara Jude, Okenwe Idochi. Modelling and Forecasting Malaria Mortality Rate using SARIMA Models (A Case Study of Aboh Mbaise General Hospital, Imo State Nigeria). *Science Journal of Applied Mathematics and Statistics.* Vol. 2, No. 1, 2014, pp. 31-41. doi: 10.11648/j.sjams.20140201.15

Abstract: This paper examined the modeling and forecasting malaria mortality rate using SARIMA Models. Among the most effective approaches for analysing time series data is the method propounded by Box and Jenkins, the Autoregressive Integrated Moving Average (ARIMA). In this paper, we employed Box-Jenkins methodology to build ARIMA model for malaria mortality rate for the period January 1996 to December 2013 with a total of 216 data points. The model obtained in this paper was used to forecast monthly malaria mortality rate for the upcoming year 2014. The forecasted results will help Government and medical professionals to see how to maintain steady decrease of malaria mortality in other to combat the predicted rise in mortality rate envisaged in some months.

Keywords: ARIMA Model, SARIMA Model, Forecasting, ARMA Model, Box-Jenkins Methods, Malaria Mortality, Akaike Information Criteria, Bayesian Information Criterion

1. Introduction

Malaria is a mosquito borne disease caused by a parasite called plasmodium (Henderson, 1999). This plasmodium has four species which include plasmodium falciparum, plasmodium vivax, and plasmodium ovale and plasmodium malariae. Malaria parasite is transmitted from one person to another through the bite of a female Anopheles Mosquito which require blood to nurture her eggs. When Malaria parasites enter the blood stream of a person, they infect and destroy the red blood cells. The destruction of these essential cells leads to fever and flu-like symptoms such as chills, headache, muscle aches, tiredness, nausea, vomiting and diarrhea. Malaria, when not treated, can lead to coma and hence death.

Globally, Malaria is increasingly becoming a disease of serious concern to everybody. This is because day by day, the impact of Malaria in human existence, the world over, becomes more ravaging and damaging as a result of high morbidity and mortality experienced in different parts of the globe especially the developing countries of which Nigeria is one.

Malaria parasite has been with man since the dawn of time. Hippocrates, a physician born in ancient Greece, today regarded as the “father of medicine” was the first to describe the manifestation of the disease.

The association with stagnant water (breeding grounds for the Anopheles Mosquito) led the Romans to begin drainage program, the first intervention against Malaria. The first recorded treatment of Malaria dates back to 1600, when the bitter bark of cinchona tree in peru was used by the native Indians. Not until 1889 was the protozoa (single celled parasite) cause of Malaria discovered by Alphonse Laveran and only in 1987 was the Anopheles Mosquito demonstrated to be the vector for the disease by Ronald Ross. The discovery of Ronald Ross was followed by a series of important works which not only enlarged the understanding of Malaria but also supplied useful knowledge in the combat against Malaria and prevention of Malaria. Despite initial success, there was a complete failure to eradicate Malaria in many countries (Mills et al; 2008).

According to World Health Organization (WHO), Center for Disease Control and Prevention (CDCP), Roll Back Malaria Partnership (RBM) (2010), 3.3 billion people-half

the world's population- are at risk of Malaria; one million people die each year from Malaria; every 30 seconds a child dies from Malaria. Also, in Africa, 91% of all Malaria death cases occur in Sub- Sahara Africa; 1 in 5 childhood deaths are caused by Malaria; 10, 000 pregnant women and 200, 000 infants die from Malaria every year.

Further more, one in ten infant's deaths and 25% of deaths in children below the age of four years is attributable to Malaria in Africa (Ofovwe and Erejie, 2001 and Ezedinachi et al, 1998). The country records about 1858 deaths per 100, 000 population from Malaria and Malaria is responsible for 60% of patients visits to health facilities and also about 30% and 11% of childhood and adult deaths, respectively (National Malaria Control Programme; N M C P, 2011).

Malaria accounts for an estimated 2 to 3 million deaths annually and is also responsible for untold morbidity in approximately 300 to 500 million people annually. Susceptible groups are children and adults who have host or never acquired immunity (Smith et al, 2002). Malaria is said to kill about one African (whether child or adult) every 15 secs and roughly 300, 000 Nigerian children annually (Salako, 2002).

Malaria is responsible for over 10% of the overall African disease burden. Children under five years of age (22% of the population) and pregnant women (20% of the population) are the most vulnerable to Malaria disease (Guillet et al, 2001). Nigeria is known for a high prevalence of malaria (Federal Ministry of Health, 2001 and Onwujekwe et al, 2000) and it is a leading cause of morbidity and mortality in the country (Federal Ministry of Health, 2001). Available records show that at least 50% of the population of Nigeria suffers from at least one episode of Malaria each year and Malaria accounts for over 45% of all out-patient visits (Federal Ministry of Health, 2001 and Ejezie et al, 1991).

It was reported that malaria prevalence (notified cases) in 2000 was about 2.4 million and responsible for an estimated average annual reduction of 1.3% in economic growth for the countries with the highest burden, Nigeria inclusive (Federal Ministry of Health, 2001 and Onwujekwe et al, 2000). Therefore, it imposes a great burden on the country in terms of pains and trauma suffered by its victims as well as loss in output and cost of treatments (onwujekwe et al, 2004).

2. Literature Review

Many researches have been done in the past regarding incidence and mortality in Malaria. The need to review some of these previous works and other related topics is necessary as it will add flavour to this study.

Durueke (2005) carried out a research on the incidence, management and bionomic of malaria in children under 5years of age in parts of Isiala Mbano L.G.A, Imo State, from November 2004 to August 2005. Using a chi-square test for proportion, the result revealed that the incidence of malaria in the studied area was inversely proportional to the socio-economic levels of the areas under study. Also, the incidence of malaria increased with decrease in socio-economic level and decreased with improvement in standard of living.

Gerritsen et al (2008) carried out an analysis on malaria incidence in Limpopo Province South Africa from 1998 to 2007, using chi-square test of independence and time series analysis, the result showed that out of 58768 cases of malaria reported including 628 deaths, the mean incidence of malaria was 124.5 per 100, 000 person and the mean mortality rate was 1.1% per season. Also, there was a decreasing trend in the incidence over time, and the mean incidence in males was higher than in females. Finally, the result revealed that incidence in malaria peaked at the age of 35 to 39 years, decreased with age from 40 years and is lowest in 0 – 4years old. The Cohort Fertility Rate (CFR) increased with increasing age.

Ayeni (2011) conducted a research titled "Malaria Morbidity in Akure South West, Nigeria: A temporal observation in climate change scenario, from 2000 to 2008". Applying the method of time series analysis, the result revealed that malaria morbidity was generally low before 2004 and that the reported cases of malaria increased from 43, 533 in 2004 to about 62, 121 case in 2008. From the result also, malaria morbidity index revealed an increase of 0.005 annually between 2000 and 2008.

Yeshiwodim, et al (2009) carried out a research on spartial analysis of malaria incidence at the village level in areas with unstable transmission in Ethiopia from September, 2002 to August, 2006. Applying the method of poisson regressios analysis, the result showed the presence of significant spartio-temporal variation and also showed a decrease in the incidence of malaria with increasing age. The conclusion was that incidence of malaria varies according to gender and age, with males age 5 and above showing a statistically higher incidence.

Korenromp et al (2007) carried out a study titled "Forecasting Malaria Incidence based on monthly case reports and Environmental Factors in karuzi Burudi, from 1997 to 2003". Using time series analysis, the result revealed that the exploration of the incidence of malaria, precipitation, temperature and vegetation for 1997 to 2003 showed no clear trend, and suggests a seasonal dependency in the series with a 6-month period for the incidence and a 12-month period for rainfall, temperature and vegetation.

Nwankwo and Okafor (2009) carried out a research on the effectiveness of insecticide treated bed nets (ITN_s) in malaria prevention among children aged 6months to 5 years in Umungwa Obowo L.G.A, Imo State of Nigeria between June and September 2006. From the 100 children selected and randomly assigned either treated bed nets or traditional bed nets, and using a chi-square test of independence, the result revealed that there was a significant difference in the malaria morbidity situation among the two groups. That is to say, morbidity due to malaria was higher in children that used traditional bed nets than the other group.

Opara (2001) carried out a study titled "The effects of malaria during pregnancy on infant mortality in Abia State Nigeria between 1993 and 1999". Using chi-square test for independence, the result showed that malaria during pregnancy increased neonatal mortality by lowering birth weight.

Adebola and Okereke (2007) conducted a study titled "Increasing Burden of Childhood Severe Malaria in a Nigerian Tertiary Hospital: Implication for control, between January 2000 and December 2005". Using logistic Regression, the result showed that severe Malaria constituted an important cause of hospital admission among Nigerian children especially those aged below 5years. The result also revealed that there was significant increase in the proportion of cases of severe malaria from 2000 to 2005.

Greenwood et al (2009) carried out a research on the evolution of malaria mortality and morbidity after the emergence of chloroquine resistance in rural area of the Gambia, West Africa between 1992 and 2004. Applying the method of univariate logistic regression and time series analysis, the result revealed that mortality attributable to malaria did not continue to increase dramatically, in spite of the growing resistance to chloroquine as first-line treatment, until 2003. The result also showed that malaria morbidity and mortality followed parallel trends and rather fluctuated accordingly to rainfall.

Baird, et al (2002) conducted a research on the seasonal malaria attack rates in infants and young children in northern Ghana from 1996 to 1997. Using fisher's exact test and chi-square test of independence, the result showed that the mean parasitemia count at the time of reinfection in the dry season roughly equaled that in the wet season.

Having reviewed some of these related literatures, we shall now in this paper examine the modeling and forecasting malaria mortality rate using SARIMA models.

3. Materials and Methods

In this paper, the methodology and the theorems propounded by Box and Jenkins called the Autoregressive Integrated Moving Average (ARIMA) was extensively explored. This is an advance forecasting technique that takes into account historical data and decomposes it into an Autoregressive (AR) process, where there is a memory of past values, an Integrated (I) process, which accounts for stabilizing or making the data stationary plus a Moving-Average (MA) process, which accounts for previous error terms making it easier to forecast.

3.1. Autoregressive Moving Average Process (ARMA) or Mixed Process

According to (6), autocorrelation patterns may require more complex models. A more General model is a mixture of the AR(p) and MA(q) models and is called autoregressive moving-average model, ARMA(p, q) model . He explained further that this model forecasts Y as both a linear combination of actual past values and a linear combination of past errors. The general ARMA (p, q) model is given by

$$\begin{aligned} Y_i &= \mu + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \mathrm{K} + \alpha_p Y_{t-p} \\ &\quad + e_i - \theta_1 e_{i-1} - \theta_2 e_{1-2} - \mathrm{K} - \theta_q e_{i-q} \end{aligned} \tag{1}$$

$$Y_i = \sum_{k=1}^{p} \alpha_k Y_{t-k} - \sum_{k=1}^{q} \theta_k e_{t-k} + \mu + e_t \tag{2}$$

Like the AR (p) model, the ARMA (p, q), has autocorrelation that diminish as the distance between residuals increases.

3.2. The Autoregressive Integrated Moving Average Model (ARIMA)

The order of the autoregressive component is p, the order of differencing needed to achieve stationarity is d, and the order of the moving average component is q. In general the ARIMA process (8) is of the form

$$Z_t = \sum_{k=1}^{p} \alpha_k Z_{t-k} - \sum_{k=1}^{q} \theta_k e_{t-k} + \mu + e_t \tag{3}$$

3.3. The Backshift and Difference Operators for ARIMA Representation

To express and understand differenced ARIMA models the concept of the backshift (lag) operator, B, and difference operator, ∇, is used, These has no mathematical meaning other than to facilitate the writing of different type of models that would otherwise be extremely difficult to express. The backshift is defined as $B^m Y_t = Y_{t-m}$ .For example $BY_t = Y_{t-1}$.

$BY_t = Y_{t-1}$, and $B^{12} Y_t = Y_{t-12}$. The difference operator takes the form $\nabla^d = (1-B)^d$, when d differences are taken to achieve stationarity in the time series data. Using these notations,

1 The general AR(p) model $Y_t = \sum_{k=1}^{p} \alpha_k Y_{t-k} + \mu + e_t$ is expressed as $Y_t - \alpha_1 Y_{t-1} - \alpha_2 Y_{t-2} - \ldots - \alpha_p Y_{t-p} = \alpha(B) Y_t = e_1 + \mu$, where $\alpha(B)$ is the autoregressive operator of order p, defined by

$$\alpha(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \ldots - \alpha_p B^p \tag{4}$$

2 The general MA (q) model $Y_t = \sum_{k=1}^{q} \theta_k e_{t-k} + \mu + e_t$ is expressed as $Y_t = e_t - \theta_t e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} = \theta(B) e_t + \mu$ where (B) is the moving average operator of order q, defined by

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \tag{5}$$

3 The general ARMA (p, q) model,

$$Y_t = \mu + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-p} + \text{K} + \alpha_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \text{K} - \theta_q e_{t-q},$$

is expressed as

$$\begin{aligned} Y_t &= \alpha_1 Y_{t-1} - \alpha_2 Y_{t-p} - \text{K} - \alpha_p Y_{t-p} \\ &= e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \text{K} - \theta_q e_{t-q} + \mu \\ &(1 - \alpha_1 B - \alpha_2 B^2 - \text{K} - \alpha_p B^p) Y_t \\ &= (1 - \theta_1 B - \theta_2 B^2 - \text{K} - \theta_q B^q) e_1 + \mu \\ &\alpha(B) Y_t = \theta(B) e_t + \mu \end{aligned} \quad (6)$$

4 Stationary series Z_t obtained after d differencing of Y_t T is given by

$$Z_t = \nabla^d Y_t = (1 - B)^d Y_t \quad (7)$$

5 A general ARIMA (p, d, q) model is expressed

$$\begin{aligned} &(1 - B)^d (1 - \alpha_1 B - \alpha_2 B^2 - \text{K} - \alpha_p B^p) Y_t \\ &= (1 - \theta_1 B - \theta_2 B^2 - \text{K} - \theta_q B^q) e_t \\ &(1 - B)^d \alpha(B) Y_t = \theta(B) e_t \end{aligned} \quad (8)$$

Table 1 gives the summary of the general non seasonal time series models and their statistical properties. The table summarizes discussions on general AR, MA, and mixed ARMA (8) models.

Table 1. *General Time Series Models.*

Model	Stationarity Condition	Invertibility Condition	Acf Coefficients	Pacf Coefficients
AR(p)	Yes	No	Die down	Cuts off after lag p
MA(q)	No	Yes	Cuts off after lag q	Die down
ARM(p, q)	Yes	Yes	Die down	Die down

3.4. Seasonal Autoregressive Models

A purely seasonal time series is the one that has only seasonal AR or MA parameters. Seasonal autoregressive models are built with parameter called seasonal autoregressive (SAR) parameters. The SAR parameters represent the autoregressive relationships that exist between time series data separated by multiples of the number of periods per season. A general AR model with P SAR parameters is given by $Y_t = \sum_{i=1}^{p} \alpha_{is} Y_{t-is}$ where Y_{t-s} is of order s, Y_{t-2s} is of order 2s and Y_{t-ps}, is of order ps. A model with one SAR parameter is written as

$$Y_t = \alpha_s Y_{t-s} + e_t \quad (9)$$

Seasonal moving Average (SMA) models are built with seasonal moving average (SMA) parameters, and the general SMA model with Q parameters is given by:

$$Y_t = \sum_{i=1}^{Q} \theta_{is} e_{t-is} + e_t \quad (10)$$

The general mixed SAR and SMA model is given by

$$Y_t = \sum_{i=1}^{p} \alpha_{is} Y_{t-is} + \sum_{i=1}^{Q} \theta_{is} e_{t-is} + e_t \quad (11)$$

The order the seasonal ARMA process is given in terms of both Ps and Qs

Table 2 gives the summary of the stationarity and invertibility conditions of some specific seasonal time series models and the behaviour of their theoretical ACF and PACF.

Table 2. *Specific Pure Seasonal Time Series Models.*

Arma Model	Stationarity Condition	Invertibility Condition	Acf Coefficients	Pacf Coefficients
$(1,D,0)^s$	$-1 < \alpha_s < 1$	None	Die down	Cuts off after one seasonal lag
$(1,D,0)^s$	$\alpha, + \alpha_{2s} < 1$	None	Die down	Cuts off after one seasonal lag
$(0,D,1)^s$	None	$-1 < \theta_s < 1$ $\theta_s + \theta_{2s} < 1$	Cuts off after one seasonal lag	Die down
$(0,D,2)^s$	None	$\theta_{2s} - \theta_s < 1$ $\theta_{2s} < 1$	Cuts off after two seasonal lag	Die down
$(1,D,1)^{s`}$	$-1 < \alpha_s < 1$	$-1 < \theta_s < 1$	Die down	Die down

4. Data on Malaria Mortality

Looking at Table in the Appendix, it shows the data of malaria mortality from January 1996 to December 2013, totaling two hundred and sixteen (216) monthly observations. The data were obtained from the Records Department, Aboh Mbaise General Hospital, Imo State Nigeria. Figures 1 and 2 show the plot of monthly malaria mortality and the trend analysis plot respectively. Figures 3 and 4 also describe the features of the data that is the autocorrelation plot and the partial autocorrelation plot respectively.

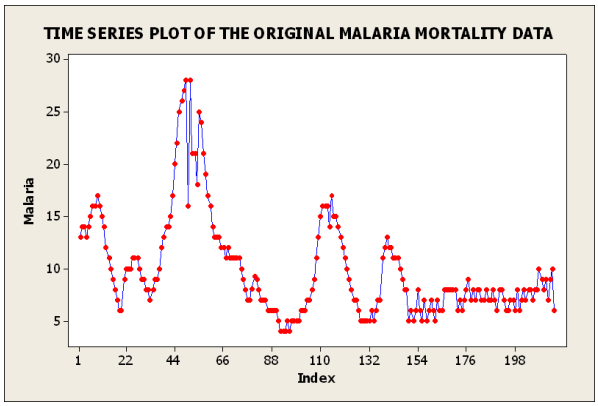

Figure 1. *Time Series Plot of Malaria Mortality.*

Looking at the time series plot of the original data that is in Fig. 1, it suggests that the series is non-stationary. Moreover, the trend analysis as shown in Fig. 2 reveals a decreasing trend. Hence, the ACF plot as shown in Fig. 3 slightly dies down in a sinewave fashion and the PACF plot as shown in Fig. 4 tails of at lag 2. Therefore an AR(2) model is suspected. The result of estimates of parameters, the ACF and the PACF of the residuals obtained by MINITAB version 15.0 Statistical software package are shown in Tables 3(a) and 3(b), Figs. 5 and 6 respectively.

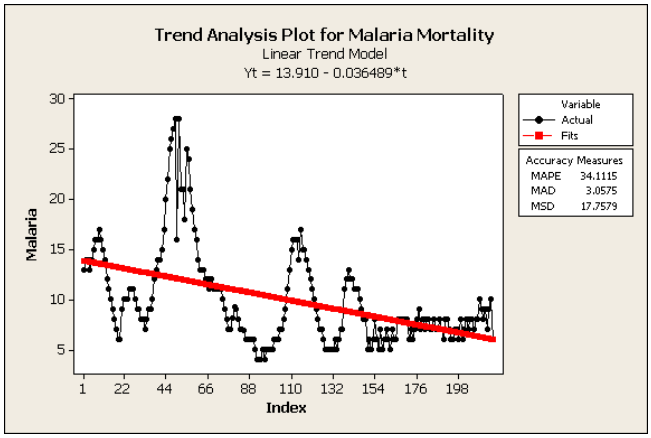

Figure 2. *Trend analysis plot of malaria mortality.*

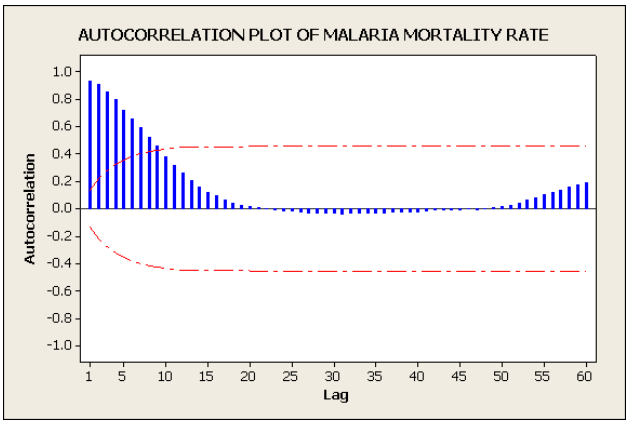

Figure 3. *Autocorrelation plot of malaria mortality.*

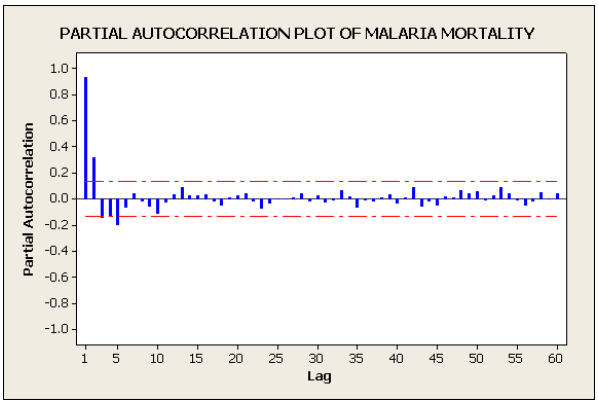

Figure 4. *Partial autocorrelation plot of malaria mortality.*

Table 3(a). *Estimates of parameters for ar (2) model.*

Final Estimates of Parameters				
Type	Coef	SE Coef	T	P
AR 1	0.6355	0.0658	9.66	0.000
AR 2	0.3187	0.0658	4.85	0.000
Constant	0.4571	0.1150	3.98	0.000
Mean	9.984	2.512		

Number of observations: 216
Residuals: SS = 608.278 (back forecasts excluded)
MS = 2.856 DF = 213

Table 3(b). *Modified Box-pierce (Ljung – Box) Chi-Square Statistic.*

Modified Box-pierce (Ljung – Box) Chi-Square Statistic				
Lag	12	24	36	48
Chi - square	19.6	31.1	33.6	39.2
DF	9	21	33	45
P – Value	0.021	0.071	0.438	0.715

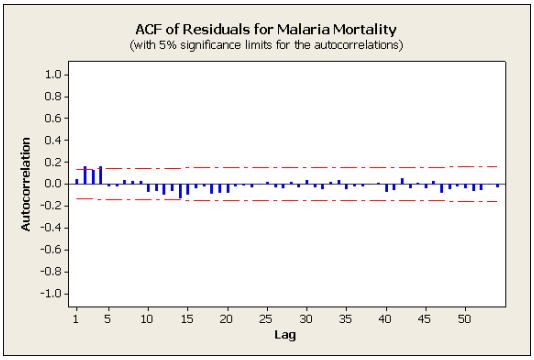

Figure 5. *ACF Plot of residuals of malaria mortality.*

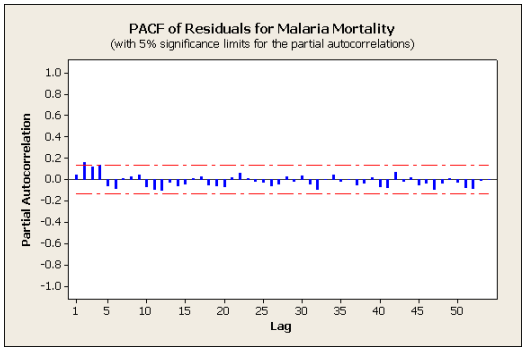

Figure 6. *PACF of residuals of malaria mortality.*

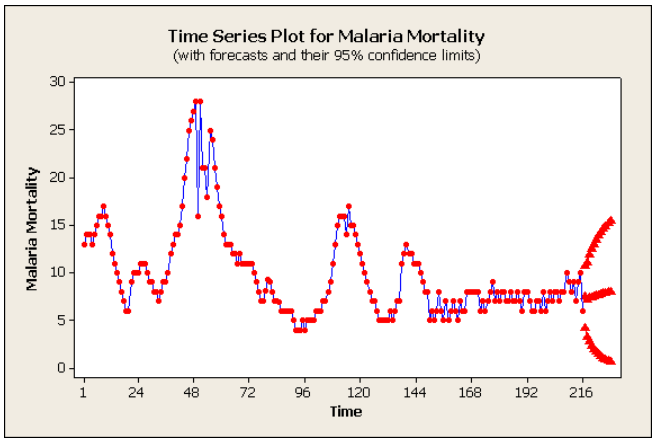

Figure 7. *Time Series plot for Forcast using ar (2).*

Looking at Figs. 5 and 6, it shows some insignificant number of spikes within the limit – 2 suggesting that the residuals are random. Fig. 4, it shows that the P-values for the Ljung-Box statistics are not significant. The forecast as shown in Fig. 7 does not seem to be consistent with the forecast of malaria mortality figures. We then try differencing the data to bring about stationarity in mean.

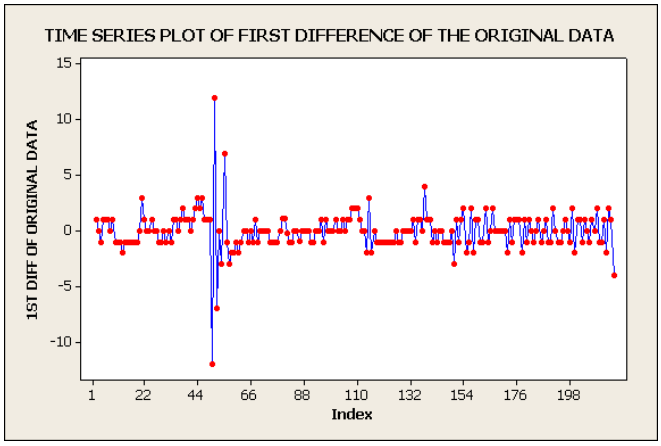

Figure 8. *Time series plot of 1st diff. of the original data.*

Figure 8 shows the time series plot of the first difference of malaria mortality original data. There is stationarity in mean and the existence of seasonality is evident.

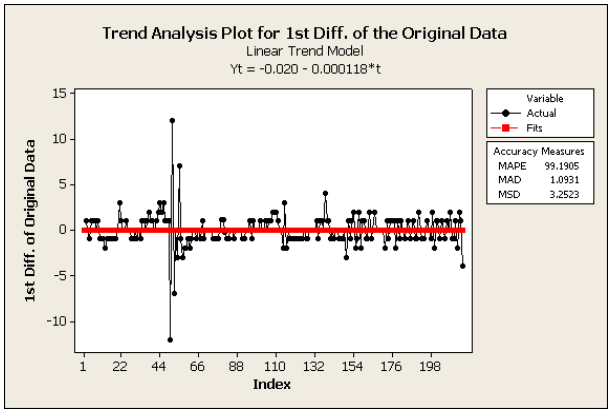

Figure 9. *Trend analysis for 1st diff. of the original data.*

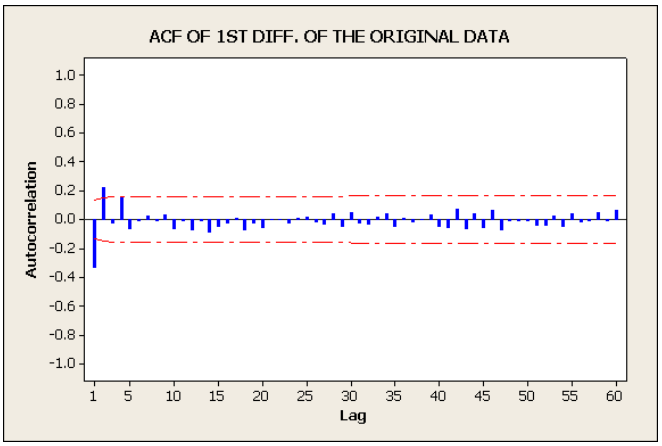

Figure 10. *ACF of 1st diff. of the original data.*

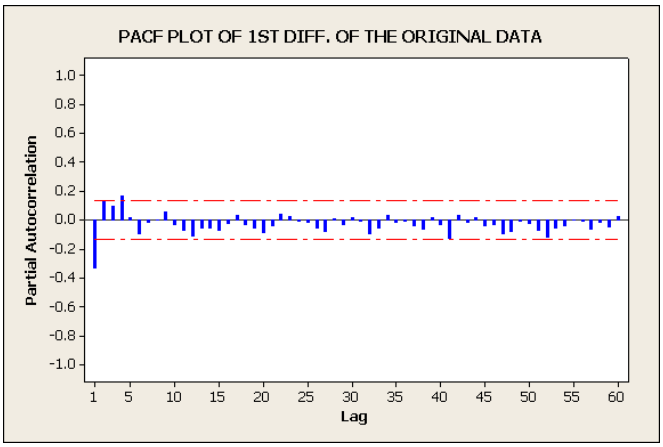

Figure 11. *PACF of 1st diff. of the original data.*

Figs. 10 and 11 show the autocorrelation function and the partial autocorrelation function of the first difference of malaria mortality original data respectively. The ACF dies in a sine wave form and the PACF also shows significant number of spikes dieing down in a sine wave fashion.

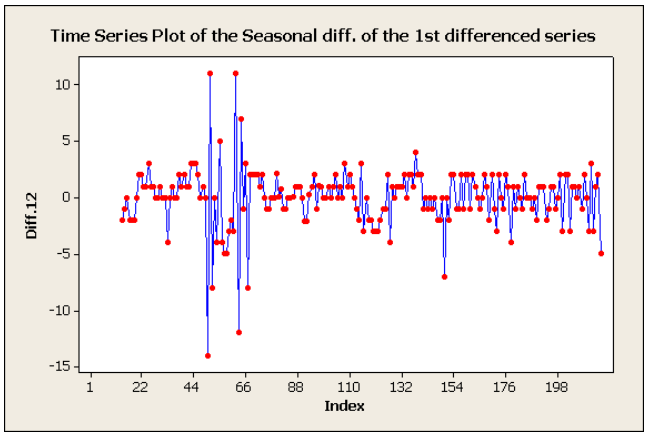

Figure 12. *Time series plot of the seasonal diff. of the 1st diff. data.*

Figure 12 shows the time series plot of the seasonal difference of the first differenced of malaria mortality data which shows stability in mean at both the seasonal and the non-seasonal levels.

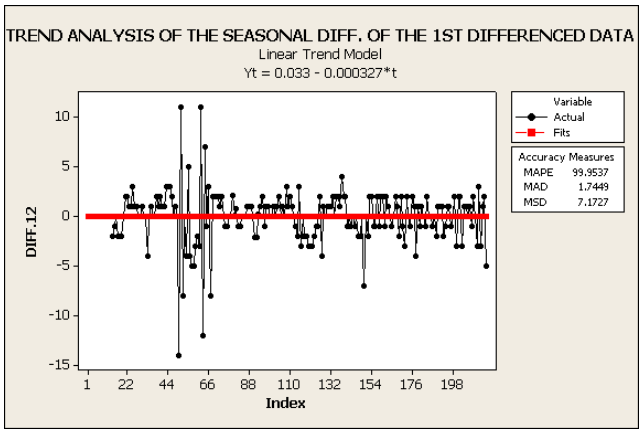

Figure 13. *Trend analysis of the seasonal diff. of the 1st diff. data.*

Figure 13 shows the trend analysis of the seasonal difference of the first differenced of malaria mortality's original data. The trend revealed neither increasing nor decreasing which is an indicative of stationarity in mean.

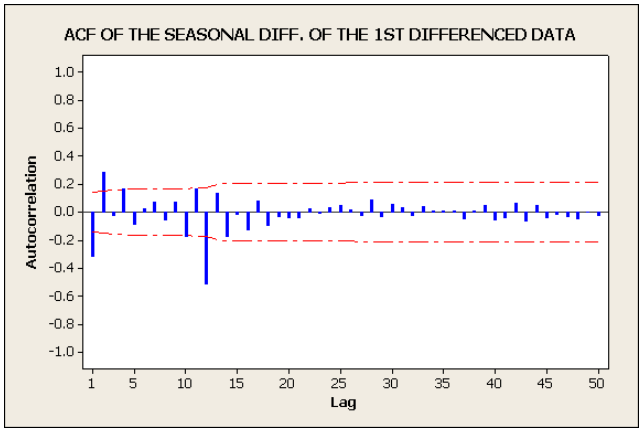

Figure 14. *ACF of the seasonal diff. of the 1st differenced data.*

Figure 14 shows the autocorrelation function of the seasonal difference of the first differenced of malaria mortality's original data moving in a sin wave fashion.

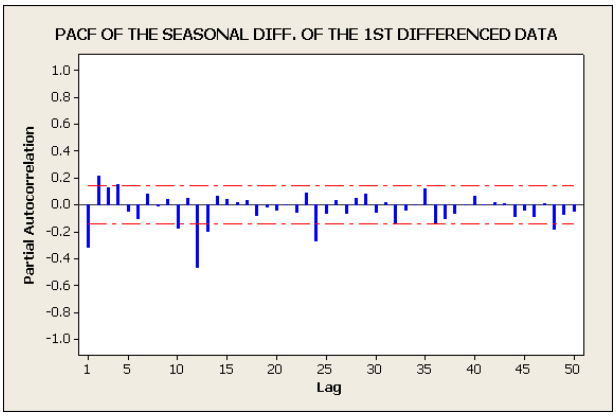

Figure 15. *PACF of the seasonal diff. of the 1st differenced data.*

The time series plot of the 1st differenced data and the trend analysis as shown in Figs. 8 and 9 show stationarity in mean and variance. There were significant spikes in the time series plot at lags 2, etc. This indicated that seasonality is evident in the monthly malaria mortality rates with a period of 12. This calls for seasonal differencing of the 1st non-seasonal differenced data, as shown in Fig. 12. Figs. 10 and 11 are the plots of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the 1st differenced data. The ACF dies down after lag 1 and the PACF tails off after lag 1, suggesting that p=1 and q=1 would be needed to describe these data as coming from a non-seasonal autoregressive and a moving average process respectively. So, the time series model that gives rise to these observations was an ARIMA (1, 1, 1) model, since the data was differenced once (i.e. d=1) to attain stationarity.

Figs. 12 and 13 show the time series plot of the seasonal difference of the 1st differenced series and the trend analysis plot respectively. The trend analysis shows stationarity at the seasonal level. Figs. 14 and 15 show the ACF and the PACF of the seasonal difference of the 1st differenced series respectively. However, a critical look at the seasonal lags show that both ACF and the PACF spikes at seasonal lag 12 dies down to zero for other seasonal lags. Suggesting that p = 1 and q = 1 would be needed to describe these data as coming from a seasonal autoregressive and moving average process. Hence, the time series model that gives rise to these observations is an ARIMA (1, 0, 1). Thus ARIMA (1, 1, 1)(1, 0, 1)12 could be the suggested model for the series at both the non-seasonal and the seasonal levels. With these suggestions encountered in this research work, the appropriate model is thus selected in the next section.

4.1. Identification of the ARIMA Model

Two goodness-of-fit statistics that are most commonly used for the model selection are; Akaike Information Criterion (AIC) and Schwarz Bayesian Information Criterion (BIC). The AIC and BIC are determined based on a likelihood function. The AIC and BIC are calculated using the formulas below: $AIC = In(SSE) + \frac{2k}{n}$ and $BIC = In(SSE) + \frac{k}{n} In(n)$ where n is the total number of observations, SSE is the sum of the squared errors, and $k = (p + q + P + Q + d + s)$. In this paper, n = 216 data points. Four tentative ARIMA models are tested for the data series and the corresponding AIC and BIC values for the models are presented in Table 4.

Table 4. *AIC and BIC values for four Tentative SARIMA Models.*

ARIMA MODEL (p, d, q)	AIC	BIC
$(1\ 1\ 1)\ (1\ 0\ 1)^{12}$	6.45644	6.53457
$(1\ 1\ 1)\ (0\ 0\ 1)^{12}$	6.44718	6.52885
$(1\ 1\ 1)\ (1\ 0\ 0)^{12}$	6.44724	6.53974
$(0\ 1\ 1)\ (1\ 0\ 1)^{12}$	6.48955	6.55206

ARIMA $(1\ 1\ 1)\ (0\ 0\ 1)^{12}$ is the most suitable model since it has the lowest AIC and BIC. We then proceed to the next stage of the Box-Jenkins approach which is the estimation of parameters of the tentative model.

4.2. Parameter Estimation of SARIMA (1, 1, 1) (0, 0, 1)^{12} model

Immediately a suitable SARIMA (P, d, q)(P, D,Q)^{12} structure is identified, the next step is the parameter estimation or fitting stage. The parameters are estimated by the maximum likelihood method. The results of parameter estimations are reported in Table 5.

Table 5(a). *Estimates of parameters of the tentative SARIMA (1, 1, 1) (0, 0, 1)^{12} model.*

Final Estimates of Parameters				
Type	Coef	SE Coef	T	P
AR 1	-0.5497	0.1588	-3.46	0.001
MA 1	-0.2287	0.1855	-1.23	0.219
SMA 12	0.7446	0.0704	1.32	0.189

Differencing: 1 regular difference
Number of observations: Original series 216, after differencing 215
Residuals: SS = 607.979 (back forecasts excluded)
MS = 2.868 DF = 212

Table 5(b). *Modified Box-Pierce (Ljung-Box) Chi-Square Statistic.*

Modified Box-Pierce (Ljung-Box) Chi – Square Statistic				
Lag	12	24	36	48
Chi-Square	11.9	22.8	26.6	32.4
DF	9	21	33	45
P-Value	0.221	0.357	0.779	0.920

We proceed in our analysis to check if the parameters contained in the models are significant. This ensures that there are no extra parameters present in the model and the parameters used in the model have significant contribution, which can provide the best forecast. The estimates of autoregressive, moving average and the seasonal moving average parameters are labeled "AR..1", "MA..1" and "SMA..12", which are -0.5497, -0.2287, and 0.0927, respectively. Based on 95% confidence level, we conclude that all the coefficients of the ARIMA (1, 1, 1) (0, 0, 1)^{12} model are significantly different from zero as shown in Table 3(a). Furthermore, the p-vales for the Ljung-Box statistic clearly all exceed 5% for all lag orders, implying that there is no significant departure from white noise for the residuals. We then proceed to the next step after parameter estimation which is the Diagnostic Checking or model validation.

4.3. Diagnostic Checking and Model Validation

The model verification is concerned with checking the residuals of the model to determine if the model contains any systematic pattern which can be removed to improve on the selected ARIMA model. It is obvious that the selected model may appear to be the best among a number of models considered; it becomes necessary to do diagnostic checking to verify that the model is adequate. Verification of an ARIMA model is tested (i) by verifying the ACF of the residuals, (ii) by verifying the normal probability plot of the residuals.

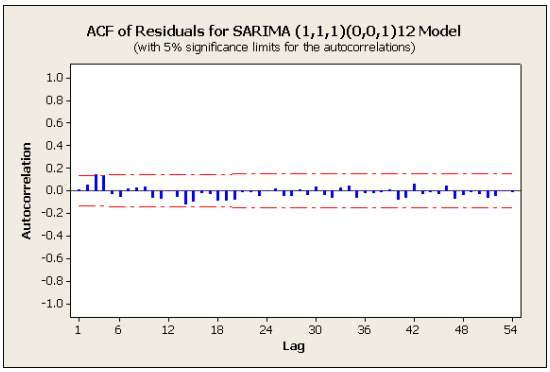

Figure 16. *ACF of Residuals for SARIMA (1, 1, 1) (0, 0, 1)^{12} Model.*

Looking at Figure 16, the autocorrelation checks of the residuals indicate that the model is good because they are white noise process. That is the residuals have zero mean, constant variance and also uncorrelated. Also, the p-values for the Ljung-Box statistic from Table 3 as shown clearly exceed 5% for all lag orders, indicating that there is no significant departure from white noise for the residuals. Since the model diagnostic tests show that all the parameter estimates are significant and the residual series are random, it can then be concluded that (1, 1, 1) (0, 0, 1)^{12} model is adequate for the inflation series. Therefore, (1, 1, 1) (0, 0, 1)^{12} is used to forecast the inflation series of Nigeria.

4.4. Point forecast with SARIMA (1, 1, 1) (0, 0, 1)^{12} Model

The ARIMA (1, 1, 1)(0, 0, 1) is selected to forecast the malaria mortality variable, where autoregressive term p = 1(non-seasonal), P = 0(seasonal) [that is, (1 - αB)(1 – 0)]; differencing term d = 1(non-seasonal difference), Q = 0(seasonal difference) [that is (1 - B)(1 – 0)] and moving average term q = 1(non-seasonal), Q = 1(seasonal) [that is $(1 - \theta_1 B)(1 - \theta_{12} B^{12})$. For the dataset in this paper, the fitted model is given by

$$(1-B)(1-\alpha B)y_t = (1-\theta_1 B)(1-\theta_{12}B^{12})e_t \qquad (12)$$

$$\begin{aligned} y_t - \alpha B y_t - B y_t + \alpha B^2 y_t &= e_t - \theta_{12}B^{12}e_t - \theta_1 B e_t + \theta_1\theta_{12}B^{12}e_t \\ y_t &= e_t - \theta_{12}B^{12}e_t - \theta_1 B e_t + \theta_1\theta_{12}B^{13}e_t + \alpha B y_t + B y_t - \alpha B^2 y_t \end{aligned} \qquad (13)$$

Transforming the back operator, equation (13) becomes;

$$\begin{aligned} y_t &= e_t - \theta_{12}e_{t-12} - \theta_1 e_{t-1} + \theta_1\theta_{12}e_{t-13} \\ &\quad + (1+\alpha)y_{t-1} - \alpha y_{t-2} \end{aligned} \qquad (14)$$

4.5. Forecast Results by SARIMA (1, 1, 1)(0, 0,1) 12 model

In order to forecast one period ahead that is, y_{t+1}, the subscript of the equation (14) is increased by one unit throughout as given by

$$\begin{aligned} y_{t+1} &= (1+\alpha)y_t - \alpha y_{t-1} + e_{t+1} \\ &\quad - \theta_{12}e_{t-11} - \theta_1 e_t + \theta_1\theta_{12}e_{t-12} \end{aligned} \qquad (15)$$

The term e_{t+1} is not known because the expected value of future random errors has been taken as zero. There are 216 data points from January 1996 to December 2013 used to build the ARIMA model. From the table 3, using α = -0.5497, θ_1 = -0.2287, θ_{12} = 0.0927, we have $\theta_1\theta_{12}$ = -0.02120. Thus, equation (15) is given as

$$y_{t+1} = 0.4503y_t + 0.5497y_{t-1} - 0.0927e_{t-11} - 0.0212e_{t-12} + 0.2287e_t + e_{t+1}$$

In order to forecast inflation for the period 217 (that is, January 2014), equation (15) is given by

$$\hat{y}_{217} = 0.4503y_{216} + 0.5497y_{215} - 0.0927\hat{e}_{205} - 0.0212\hat{e}_{204} + 0.2287\hat{e}_{216} + \hat{e}_{217}$$

$$\hat{e}_{217} = 0\text{'}\ \hat{e}_{204} = y_{204} - \hat{y}_{204} = 8 - 6.4659 = 1.5341$$

$$\hat{e}_{216} = y_{216} - \hat{y}_{216} = 6 - 6.0280 = -0.028$$
$$\hat{e}_{205} = y_{205} - \hat{y}_{205} = 8 - 6.4294 = 1.5706$$

The forecast quantity for period 217 can now be calculated as follows:

$$\hat{y}_{217} = 0.4503(6) + 0.5497(10) - 0.0927(1.5706) - 0.0212(1.5341) + 0.2287(-0.028) + 0 = 8.01\%$$

Once our model has been obtained and its parameters have been estimated, we can use it to make our prediction. Table 7 summarizes 12 months ahead malaria mortality forecast from January 2014 to December 2014 with 95% confidence interval.

5. Conclusion

In this paper, modeling and forecasting malaria mortality rate using SARIMA models was examined. Box-Jenkins Seasonal Autoregressive Integrated Moving Average (SARIMA) was employed to analyze monthly malaria mortality rate in Imo State from January 1996 to December 2013. The study intended mainly to forecast the monthly malaria mortality rate for the coming period of January, 2014 to December 2014.

Series of tentative models were developed to forecast monthly malaria mortality rate, but based on minimum AIC and BIC values and after the estimation of parameters and series of diagnostic test were performed, ARIMA$(1,1,1)(0,0,1)_{12}$ model was proved to be the best model for forecasting after satisfying the model assumptions.

The forecasted results revealed a decreasing pattern of malaria mortality rate in the last quarter of 2014, except the month of December where it increased.

Table 7. *12- Month Forecasted Malaria Mortality for January 2014 to December 2014.*

Month	Period	Forecast (%)	Lower	Upper
January	217	7.2873	3.9675	10.6072
February	218	6.6853	2.6725	10.6980
March	219	6.9800	2.0640	11.8960
April	220	6.7590	1.2357	12.2822
May	221	6.6895	0.5439	12.8351
June	222	6.7275	0.0554	13.3995
July	223	6.8382	-0.3412	14.0177
August	224	6.7424	-0.9007	14.3855
September	225	6.9307	-1.1548	15.0162
October	226	6.7317	-1.7705	15.2338
November	227	6.6491	-2.2516	15.5499
December	228	7.0097	-2.2717	16.2912

Table 8. *Basic Statistic of Monthly Malaria Mortality Data in Percentages.*

No. of observation	Mean	St. Dev.	Variance	Min.	Max.
120	6.8359	0.1859	0.0346	6.6491	7.2873

Appendix

Month	Malaria	Month	Malaria	Month	Malaria	Month	Malaria	Month	Malaria	Month	Malaria
1	13	41	14	82	8	123	9	164	6	205	8
2	14	42	15	83	7	124	8	165	6	206	7
3	14	43	17	84	7	125	7	166	8	207	8
4	13	44	20	85	7	126	7	167	8	208	8
5	14	45	22	86	6	127	6	168	8	209	10
6	15	46	25	87	6	128	5	169	8	210	9
7	16	47	26	88	6	129	5	170	8	211	8
8	16	48	27	89	6	130	5	171	8	212	9
9	17	49	28	90	6	131	5	172	6	213	7
10	16	50	16	91	5	132	5	173	7	214	9
11	15	51	28	92	4	133	6	174	6	215	10
12	14	52	21	93	4	134	5	175	7	216	6
13	12	53	21	94	4	135	6	176	8		
14	11	54	18	95	5	136	7	177	9		
15	10	55	25	96	4	137	7	178	7		
16	9	56	24	97	5	138	11	179	8		
17	8	57	21	98	5	139	12	180	7		
18	7	58	19	99	5	140	13	181	8		

Month	Malaria	Month	Malaria	Month	Malaria	Month	Malaria	Month	Malaria	Month	Malaria
19	6	59	17	100	5	141	12	182	8		
20	6	60	16	101	6	142	12	183	7		
21	9	61	14	102	6	143	11	184	7		
22	10	62	13	103	6	144	11	185	8		
23	10	63	13	104	7	145	11	186	7		
24	10	64	13	105	7	146	10	187	7		
25	11	65	12	106	8	147	9	188	8		
26	11	66	12	107	9	148	8	189	7		
27	11	67	11	108	11	149	8	190	6		
28	10	68	12	109	13	150	5	191	8		
29	9	69	11	110	15	151	6	192	8		
30	9	70	11	111	16	152	5	193	7		
31	8	71	11	112	16	153	6	194	6		
32	8	72	11	113	16	154	8	195	6		
33	7	73	11	114	14	155	6	196	7		
34	8	74	10	115	17	156	5	197	7		
35	9	75	9	116	15	157	7	198	6		
36	9	76	8	117	15	158	5	199	8		
37	10	77	7	118	14	159	6	200	6		
38	12	78	7	119	13	160	7	201	7		
39	13	79	8.12	120	12	161	6	202	8		
40	14	80	9.24	121	11	162	5	203	7		

References

[1] Adebola P.A. and Okereke R.W. (2007). Increasing Burden of childwood severe Malaria in a Nigeria Tertiary Hospital from 2000 to 2005. An unpublished research work.

[2] Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control 19 (6): 716-723.

[3] Ayeni, A. O., (2011). Malaria Morbidity in Akure, Southwest Nigeria: A Temporal Observation in a Climate Change Scenario. Trends Applied Sci. Res., 6: 488-494

[4] Ayodele J.; Oluyemi S.; Amos P.; and Tuoyo O (2007). Quantifying the Economic Burden of Malaria using the Willingness to Pay Approach. An Article submitted to the Department of Economics, University of Illorin, Illorin, Nigeria

[5] Baird, J.K.; Owusu Agyei S; Utz G.C.; Koram, K; Barcus M.J.; Jones, T.R.; Fryauff, D.F.; Binka, F.N.; Hoffman, S.L.; Nkruma, F.N. (2002). Seasonal Malaria attack rates in infants and young children in Northern Ghana. Naval Medical Research Center, Silver Spring, Maryland, USA.

[6] Box, G. E. P and Jenkins, G.M., (1976). "Time series analysis: „Forecasting and control," Holden-Day, San Francisco.

[7] Durueke A.P. (2005). A Research on the incidence, management and bionomc of Malaria in children under 5 years of age in parts of Isiala Mbano L.G.A., Imo State, Nigeria. Unpublished research work.

[8] Fleiss, J . L. (1973). Statistical Methods for Rates and Proportions. John Wiley and Sons, New York.

[9] Fritzer, F., Gabriel, M. and Johann, S. (2002). "Forecasting Austrian HICP and its Components using VAR and ARIMA Models," Working Papers 73, Oesterreichische National bank (Austrian Central Bank).

[10] Gerristsen, A.; Kruger P; Van der Leo, M. and Grobusch, M. (2008). 'Malaria incidence in limpopo provincem, South Africa, 1998 – 2007', Malaria Journal 7(1). URL:http://www.malariajournal. com/content/7/1/162.

[11] Gomez V., and Maravall A., (1998.) "Automatic Modelling Methods for Univariate Series," Banco de EspaÃ±a Working Papers 9808, Banco de España.

[12] Greenwood, B.M.; Bradley, A.K.; Greenwood, A.M.; Byass, P.; Jammeh, K.; Marsh, K.; Tulloch, S.; Old eld, F.S. and Hayes, R. (2009). 'Mortality and morbidity from malaria among children in a rural area of the gambia, West Africa,' American Journal of Tropical Medicine and Hygiene 81 (3), 478 – 486.

[13] Hamidreza M. and Leila S. (2012). Using SARFIMA model to study and predict the Iran's oil supply. International Journal of Energy Economics and Policy. Vol.2, No.1, 2012, pp.41-49.

[14] Hardeo, S. and Mohammed I.A. (2000). The Analysis of Variance. Sheridan Books, Inc; Birkhauser Boston.

[15] Jeffrey J., (1990). "Business forecasting Methods". Atlantic Publishers.

[16] Jack H.R, Bond M.T., and Webb J.R, (1989). "The Inflation- Hedging Effectiveness of Real Estate". Journal of Real Estate Research, Vol.4. Pp. 45-56.

[17] Koram, K.A. and Molyneux, M.E. (2007). 'When is "Malaria" Malaria? The Dierent Burdens of Malaria Infection, Malaria Disease, and Malaria-Like Illnesses'. American Journal of Tropical Medicine and Hygiene 77 (6 Suppl).

[18] Korenromp, E.; Kiniboro, B. and WHO (2007). Forecasting Malaria incidence estimates at Burndi country level for the year 1997 to 2003 – draft report. http://www.who.int/malaria/publications/atoz/incidence estimations2.pdf.

[19] Leila S. and Masoud Y. (2012). An Empirical Study of the Usefulness of SARFIMA models in Energy Science. International Journal of Energy Science. IJES Vol.2 No.2 2012.

[20] Mills, A.; Lubell, Y. and Hanson, K. (2008). “Malaria eradication: the economic, financial and institutional challenge”, Malaria Journal 7 (Suppl 1), S11. URL:http://www.malariajournal.com/content /7/S1/S11

[21] Nomeh F.O. (2008): Statistical Analysis on malaria mortality in Enugu State Nigeria. An unpublished B.Sc. Project submitted to the Department of Statistics, Nnamdi Azikiwe University, Awka, Anambra State.

[22] Opara K.R. (2001). Effect of Malaria during pregnancy on mortality in Abia State Nigeria between 1993 and 1999. An unpublished B.Sc. Project submitted to the department of Statistics, Abia State University, Uturu, Abia State.

[23] Rudiger Dornbusch and Stanley Fischer, (1993). “Moderate Inflation”, The Bank Economic Review, Vol.7, Issue 1, Pp.1-44.

[24] Trevor, W. (2010). Applied Business Statistics, Methods and Excel-Based Application. EBLS/Clarendo Press Oxford London.

[25] Yeshiwondim A.K; Gopal, S.; Hailemarian A.T.; Dengela, D.O.; Patel, H.P.: Spatial analysis of Malaria incidence at the village level in areas with unstable transmission in Ethiopia. Int. J. Health Geogr 2009, 8:5. Pubmed Abstract.