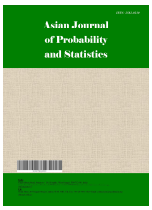

Asian Journal of Probability and Statistics

6(3): 61-73, 2020; Article no.AJPAS.54310
ISSN: 2582-0230

The Effects of Malaria Control Interventions on Out-patient and In-patient Malaria Cases in the Northern Region of Ghana

Emmanual Mohammed Dokurugu[1*], Suleman Nasiru[2] and Benson Abdul Majeed[3]

[1]Department of Mathematics, E. P. College of Education, P.O.Box 16, Bimbila, Ghana.
[2]Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies, Navrongo, Ghana.
[3]Department of Records and Health Information, Tamale Teaching Hospital, Tamale, Ghana.

Authors' contributions

This work was carried out in collaboration with all authors. Author EMD designed the study, performed statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Authors SN and BAM managed the analysis of the study. They also managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2020/v6i330164
Editor(s):
(1) Dr. S. M. Aqil Burney, Department of Computer Science, University of Karachi, Pakistan.
Reviewers:
(1) Sabelo Nick Dlamini, University of Eswatini, Eswatini.
(2) Hussin Jose Hejase, Al Maaref University, Lebanon.
(3) Irshad Ullah, Pakistan.
Complete Peer review History: http://www.sdiarticle4.com/review-history/54310

Original Research Article

Received: 02 December 2019
Accepted: 07 February 2020
Published: 14 February 2020

Abstract

In this study, change detection in Out-patient and In-patient malaria cases in the Northern Region of Ghana was examined using time series intervention analysis. Data on monthly Out-patient and In-patient malaria cases obtained from the Northern Regional Health Directorate were modelled using Seasonal Autoregressive Integrated Moving Average with an Independent variable (SARIMAX) and Autoregressive Integrated Moving Average with an Independent variable (ARIMAX) models. The results revealed that SARIMAX $(1, 1, 1)(1, 1, 1)_{12}$ was the best model for predicting Out-patient malaria cases while SARIMAX $(1, 1, 1)(2, 1, 1)_{12}$ emerged as the best model for predicting the In-patient cases in the region. Diagnostic checks of the two models with the Ljung-Box test and Autoregressive Conditional Heteroscedasticity Lagrange Multiplier (ARCH-LM) test revealed that both models were free from higher-order serial correlation and conditional heteroscedasticity respectively. A chi-square goodness-of-fit test also revealed that there was no significant difference between the predicted values from the models and the observed values for the year 2018. The study further revealed that the coefficients of the

**Corresponding author: E-mail: emmanueldoks@gmail.com;*

intervention variable for the Out-patient and In-patient cases were both negative, which suggest that the intervention policy the government of Ghana implemented brought about a decline in the number of Out-patient and In-patient cases in the region.

Keywords: Out-patient; In-patient; SARIMAX; ARIMAX; ARCH-LM test; intervention.

1 Introduction

Malaria has been an important public health problem in sub-Saharan Africa (SSA) and the main cause of morbidity, mortality and permanent disability in non-immune children below five years of age and pregnant women. Malaria causes considerable pain and weakness among affected people, thereby affecting their performance at work and in many cases causing absenteeism from work or school among children. The consequences from severe malaria such as convulsions or brain dysfunction can hamper long-term development and performance of children in school (Multiple Indicator Cluster Survey) [1]. Due to the debilitating effects of malaria, scale-up of proven malaria prevention, control and treatment continue to be central to successive governments of Ghana, international, and local organizations and other stakeholders in the health sector. The interventions cover the area of improving multiple prevention (through indoor residual spraying, scale-up of the distribution of insecticide treated nets, intermittent preventive treatment for pregnant women, improving sanitation, increasing access to healthcare and improving data reporting).

Myriad of researches have been carried out all over the world on malaria morbidity and mortality. Landoh et al. [2] examined morbidity and mortality due to malaria in Est Mono district, Togo, from 2005 to 2010 using chi-square test of independence and student's t-test. In another study, Nyarango et al. [3] investigated the steep decline of malaria morbidity and mortality trends in Eritrea after the Roll Back Malaria Programme was implemented by employing Autoregressive Integrated Moving Average (ARIMA) models. Furthermore, Karema et al. [4] assessed the trends in malaria cases, hospital admissions and deaths following scale-up of anti-malarial interventions, 2000–2010, in Rwanda. Briet et al. [5] also carried out a research using Generalized Seasonal Autoregressive Integrated Moving Average (GSARIMA) models for count data with application to malaria time series with low case numbers in Sri Lanka. The models were applied to monthly malaria case time series in a district in Sri Lanka, where malaria had decreased drastically in recent years. Additionally, Osadolor et al. [6] employed the SARIMA intervention time series analysis to investigate the effect of malaria control intervention in the Kwazulu-Natal Province in South Africa. The intervention was the re-introduction of dichlorodiphenyltrichloethane (DDT) on confirmed malaria cases. The result showed an abrupt and permanent decline of malaria cases following the implementation of the intervention policy. Again, Anokye et al. [7] carried out a research on forecasting future malaria incidence in the Kumasi Metropolis, Ghana using ARIMA models. Trends of malaria prevalence was analysed and compared by years and months. It was revealed that July had the highest number of cases whereas January recorded the lowest number of cases. The predicted number of cases for the first and second halves of the year 2018 were 61, 371.8 and 77,842.0 respectively. Also, Hassan and Bin [8] used the Box-Jenkins SARIMA model approach to investigate monthly malaria infections in the Kass Zone, South Darfur State, Sudan. An ARIMA forecasting model was obtained from the analyses to predict the monthly malaria infections. Furthermore, Alhassan et al. [9] researched on time series analysis of malaria cases in the Kasena Nankana Municipality, Navrongo, Ghana. They developed an ARIMA model that can adequately forecast future trends of malaria cases in the Municipality. Anwar et al. [10] also used time series ARIMA models to predict future trends in malaria incidence in Afghanistan. Two (2) predictive models were obtained that can accurately forecast malaria incidence in that country. Enhanced vegetation index was also found to have increased the predictive accuracy of the models in the long-term. Additionally, Ankamah et al. [11] used vector autoregression (VAR) to model the impact climatic variability malaria in Ghana. The study revealed that malaria is highly influenced by three (3) main climatic variables that include maximum temperature, rainfall and humidity. Again, Perez and Ceballos [12] conducted a study to develop an appropriate model that could predict the weekly reported malaria incidence in the Philippines using the Box-Jenkins method. Based on the results of their analysis, ARIMA (2, 1, 0) was selected as the model for predicting the weekly malaria incidence in the Philippines.

Despite the many researches that have been carried out on malaria all over the world, it is observed that no such investigation has been conducted in the Northern Region of Ghana using intervention analysis. This study will therefore employ the intervention analysis models developed by Box and Jenkins [13] to assess a change detection in Out-patient and In-patient malaria cases following the implementation of the malaria control intervention programme in the Northern Region of Ghana.

2 Materials and Methods

This study was carried out in Ghana using data on monthly Out-patient and In-patient cases on malaria obtained from the Northern Regional Health Directorate. The data ranged from January 2004 to December 2018 with the pre-intervention period being January 2004 to December 2007 and the post-intervention period being January 2008 to December 2018. The analysis was done using the data from January 2004 to December 2017 while that of 2018 was used for cross validation. The data was modelled using Seasonal Autoregressive Integrated Moving Average (SARIMA) and ARIMAX models. The modelling was preceded by preliminary tests to determine the presence or absence of seasonality and unit roots in the data.

2.1 Unit root test

Stationarity is a vital aspect of time series analysis. Several approaches have been developed to test for the stationarity or non-stationarity of a time series data which include both graphical and quantitative approaches. In this study, we employed the Augmented Dickey-Fuller (ADF) [14] test to test for unit roots. The presence of a unit root indicates that the time series is not stationary and needs to be differenced.

2.1.1 Augmented Dickey Fuller (ADF) test

The ADF test tests the null hypothesis that a unit root is present in a time series sample. It has the advantage of handling a larger and more complicated set of time series models. The test is based on the regression of the observed variable Y_t on its one-period lagged value Y_{t-1}, sometimes including an intercept and a time trend. The ADF model is given as:

$$\Delta Y_t = \alpha + \beta t + \delta Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \cdots + \gamma_{p-1} \Delta Y_{t-p+1} + \varepsilon_t \quad (1)$$

Where Δ is the difference operator, implying that $\Delta Y_t = Y_t - Y_{t-1}$, $\delta = \phi - 1$, α is a constant, β the coefficient on time trend series, $\gamma_1 \Delta Y_{t-1} + \cdots + \gamma_{p-1} \Delta Y_{t-p+1}$ is the sum of the lagged values of the dependent variable ΔY_t and p is the lag order of the Autoregressive (AR) process. The ADF test is concerned with the value of the parameter δ. If $\delta = 0$, it presupposes that the series contains unit root and hence non-stationary. The test statistic for the ADF test is given by:

$$\mathrm{F}_\tau = \frac{\hat{\delta}}{\mathrm{SE}(\hat{\delta})} \quad (2)$$

Where $\hat{\delta}$ is the least square estimate and $\mathrm{SE}(\hat{\delta})$ is the standard error estimate of $\hat{\delta}$. If the calculated value of the test statistic is greater than the critical value, we reject the null hypothesis of $\delta = 0$.

2.2 Seasonal ARIMA (SARIMA) model

Identification of relevant models and inclusion of suitable seasonal variables is necessary when a time series data exhibit periodic pattern. The SARIMA model therefore has the advantage of capturing both seasonal and non-seasonal components. The general notation for the order of a SARIMA model is ARIMA(p, d, q)(P, D, Q)s and can be expressed using the backshift operator as:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D Y_t = \theta(B)\Theta(B^s)\varepsilon_t \quad (3)$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \quad (4)$$

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps} \quad (5)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \quad (6)$$

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \ldots - \Theta_Q B^{Qs} \quad (7)$$

Where Y_t represents the time series data at period *t*, *B* denotes the backshift operator, ε_t is a sequence of iid variables with mean zero and variance σ^2, *s* is the seasonal order, ϕ_i and Φ_j are the non-seasonal and seasonal AR parameters respectively, θ_i and Θ_j are respectively the non-seasonal and seasonal Moving Average (MA) parameters, *p, d* and *q* denote the non-seasonal AR, Integrated (I) and MA orders respectively and *P, D* and *Q,* respectively represent the seasonal AR, I and MA orders.

2.3 Regression with ARIMA errors

ARIMA models with input variables are referred to as regression with ARIMA errors or ARIMAX model. The model combines a regression model with an ARIMA model. The regression component describes the explanatory relationship of the variables whereas the ARIMA component deals with the autocorrelation in the residuals of the regression model. An ARIMAX model is given by:

$$Y_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \frac{\theta(B)\Theta(B^s)}{\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D}\varepsilon_t \quad (8)$$

2.4 Criterion for model selection

It is imperative for model selection criteria to be carried out since there is the possibility of two or more models to compete in the selection of the best model. The Akaike Information Criterion (AIC), the Akaike Information Criterion corrected (AICc) and the Bayesian Information Criterion (BIC) were employed in this study to select the most adequate model [15]. The best model is the one with the smallest AIC, AICc or BIC values, given a set of candidate models. The AIC, AICc, and BIC are generally given by;

$$\text{AIC} = 2k - 2In(L) \quad (9)$$

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{n-k-1} \quad (10)$$

$$\text{BIC} = \log(\sigma_e^2) + \frac{k}{n}\log(n) \quad (11)$$

Where *k* represents the number of parameters in the model, *L* denotes the maximised value of the likelihood function, *n* is the number of observations in the data and σ_e^2 is the error variance.

2.5 Model diagnostics

After a model has been built, it is important to diagnose the model in order to ensure that it truly reflects the real time series observations. When these checks are done the model can be used to make meaningful generalisations or to draw inferences. The Ljung-Box and ARCH-LM tests were employed in this study to diagnose the adequacies of the developed models.

2.5.1 Ljung-Box test

The Ljung-Box test was developed by Ljung and Box [16] as a diagnostic tool to examine autocorrelations of the residuals of a fitted model. It tests the null hypothesis that autocorrelations up to say order *k* equal

zero, which indicates that the values are random and independent over time. The test therefore assumes that the residuals do not contain serial correlation up to order k and is considered as a portmanteau test. The Ljung-Box test was therefore used to determine the presence or absence of serial correlation in the data up to a given order, say k. The test statistic is given by;

$$Q_h = n(n+2)\sum_{k=1}^{h}\frac{r_k^2}{n-k} \tag{12}$$

Where r_k^2 represents the residual autocorrelation at lag k, n is the number of residuals and h is the number of lags being tested.

We reject the null hypothesis if Q_h is greater than the chi-square table value. The model is therefore considered adequate when the p-value associated with Q_h is large.

2.5.2 ARCH-LM test

The ARCH-LM test developed by Eagles [17] is a standard approach used to notice autoregressive conditional heteroscedasticity. It tests the null hypothesis that no ARCH exists up to order k in the residuals. A researcher is therefore likely to be challenged with the issue of conditional heteroscedasticity when fitting models. This problem occurs when the residuals do not have a constant variance. Therefore, the assumption of constant variance must be met in order to obtain an adequate model. This study therefore employed the ARCH-LM test due to its undisputed advantage in detecting autoregressive conditional heteroscedasticity. The test statistic is given as:

$$\text{LM} = nR^2 \tag{13}$$

Where n is the number of observations and R^2 is the coefficient of determination of the auxiliary residual regression. This is given by:

$$e_t^2 = \beta_0 + \beta_1 e_{t-1}^2 + \beta_2 e_{t-2}^2 + \ldots + \beta_q e_{t-q}^2 + v_t \tag{14}$$

Where e_t is the residual. The null hypothesis is rejected when the p-value is greater than the level of significance and hence we conclude that there is no heteroscedasticity in the model residuals.

2.5.3 Jarque-Bera (JB) test

Normality is a common assumption in many statistical analyses. Hence testing the normality of a distribution has become a standard feature in many statistical works (Jula) [18]. This study therefore employed the JB test to confirm the normality of the data. The JB test is very useful when the sample size is large (greater than 2000). The test statistic is given as:

$$JB = n[\frac{S^2}{6} + \frac{(K-3)^2}{24}] \tag{15}$$

Where n= sample size, S= coefficient of skewness and K= coefficient of kurtosis

The JB test has a chi-square distribution with two degrees of freedom. Hence, we reject the null hypothesis of normality if $JB > X^2$ calculated value.

3 Results and Discussion

The ADF test was employed to determine whether or not the two time series data were both stationary. Using the ADF test with only a constant term and a constant with quadratic trend, revealed the presence of

unit roots in the two series since in both cases, the *p*-values were greater than the 0.05 level of significance as illustrated in Table 1.

Table 1. ADF test for Out-patient and In-patient cases

Case	Constant		Constant + Quadratic Trend	
	Test statistic	***P*-value**	**Test statistic**	***P*-value**
Out-patient	-0.8780	0.7958	-2.2981	0.6798
In-patient	-0.8158	0.8142	1.9758	0.8263

The periodic spikes in the ACF plots of both the Out-patient and In-patient series shown in Figs. 1 and 2 gave an evidence of seasonality in both series. Thus, in order to stabilise the variance, the two time series data were transformed logarithmically. The transformed series were then differenced both seasonally and non-seasonally and then tested for stationarity. For both the Out-patient and In-patient series, the ADF test revealed that the transformed seasonal and non-seasonal differenced series were stationary since each had a *p*-value less than the 5% significance level as illustrated in Table 2.

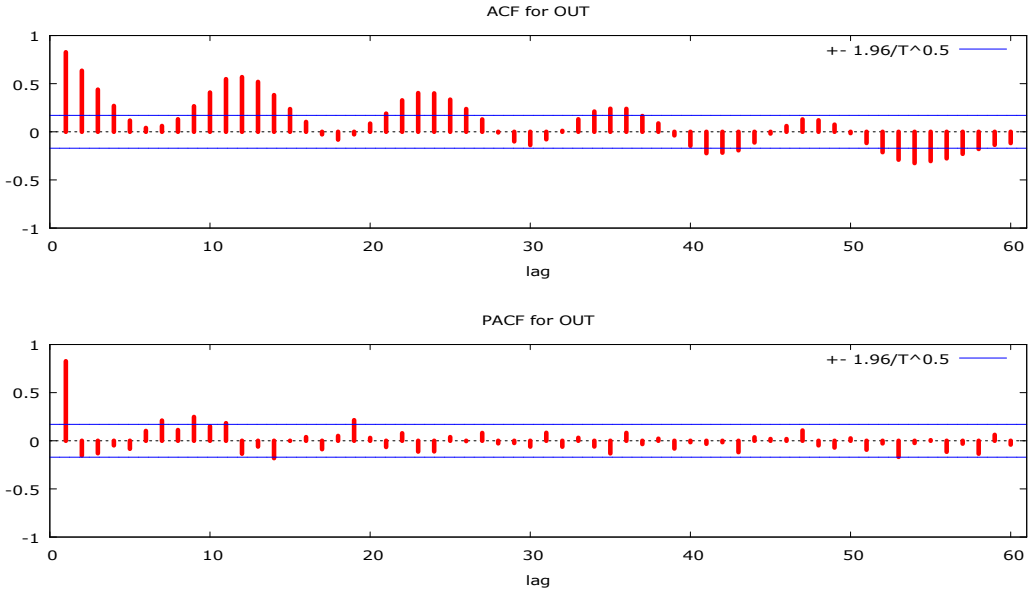

Fig. 1. ACF and PACF plots of Out-patient cases

Table 2. ADF test for seasonal and non-seasonal differenced of Out-patient and In-patient cases

Case	Constant		Constant+ Quadratic Trend	
	Test statistic	***P*-value**	**Test statistic**	***P*-value**
Out-patient	-4.5070	.0001	-4.4587	.0074
In-patient	-4.3826	.0001	-4.2476	.0149

3.1 Estimating the SARIMAX model for the out-patient

The ACF plot in Fig. 3 shows significant spikes at the non-seasonal lag 1 and seasonal lag 12 with significant spikes at other non-seasonal lags. The PACF plot also has significant spikes at the non-seasonal lags 1 and 2 and seasonal lags 12, 24, and 36. Using the lower significant lags of both the ACF and PACF and their respective seasonal lags, tentative SARIMAX models were fitted with an independent variable as

the intervention variable in order to obtain the best model as shown in Table 3. The SARIMAX model is also referred to as regression with ARIMA errors. Among these possible models, SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$ was adjudged the best since it had the least AIC, AICc and BIC values as compared to the other models.

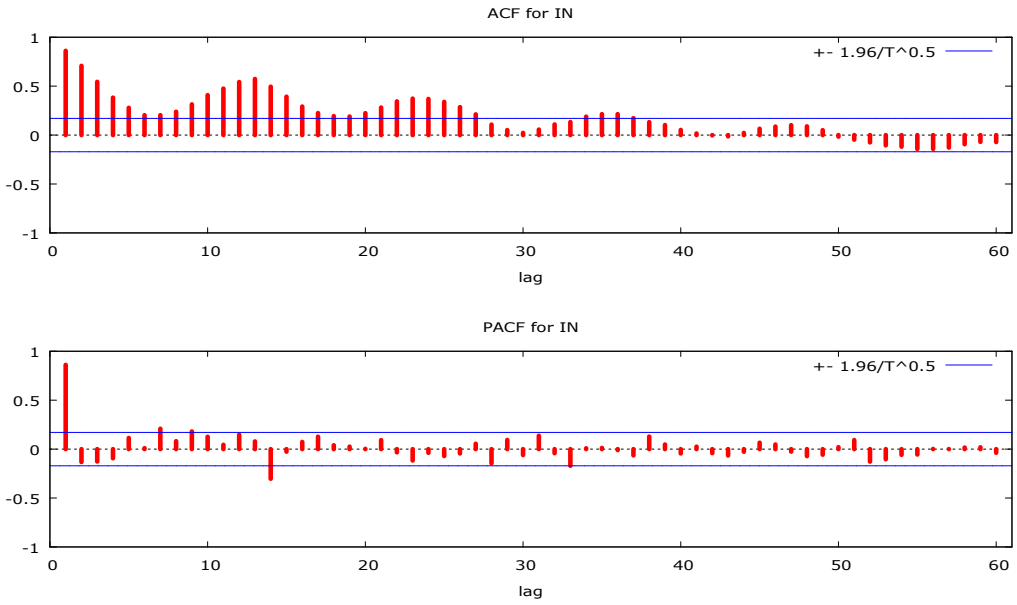

Fig. 2. ACF and PACF plots of In-patient cases

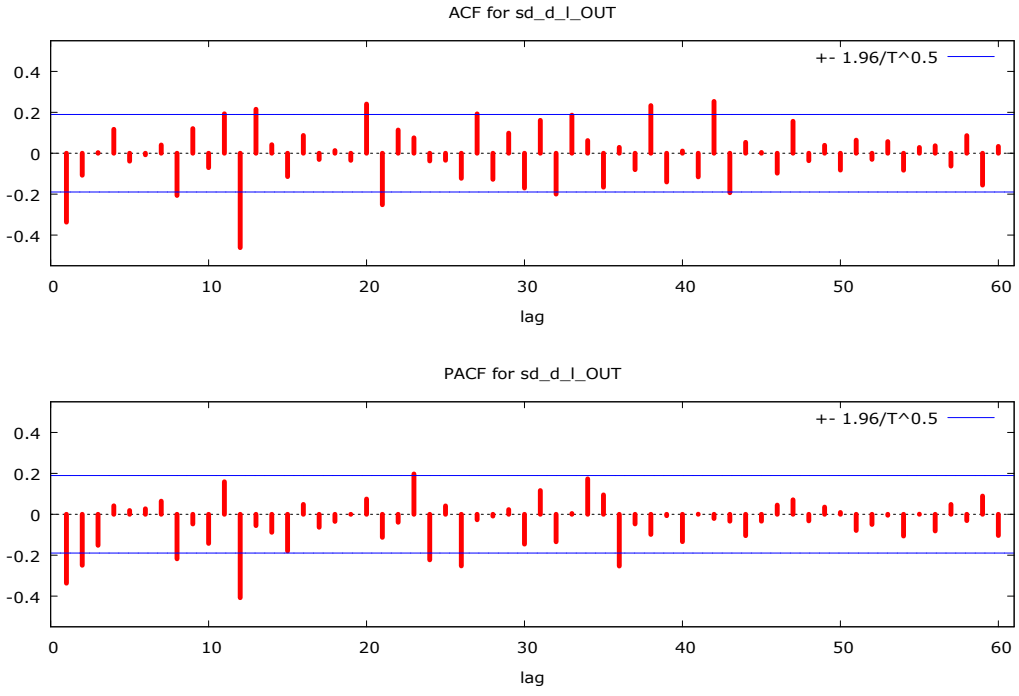

Fig. 3. ACF and PACF plot of differenced series for Out-patient cases

Table 3. Tentative SARIMAX models for Out-patient cases

Model	AIC	AICc	BIC
SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$	-53.2994*	-52.9515*	-39.9352*
SARIMAX (2, 1, 1)(1, 1, 1)$_{12}$	-51.3490	-51.0012	-35.3120
SARIMAX (1,1 ,1)(2, 1, 1)$_{12}$	-51.3324	-50.9846	-35.2954
SARIMAX (2, 1, 1)(2, 1,1)$_{12}$	-49.4036	-49.0558	-30.6983

**: Means best based on the selection criteria*

The selected model, SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$ can be expressed in terms of the backshift operator based on the parameters of the model shown in Table 3 as:

$$\ln \text{OUT} \quad = -0.1467D + \frac{(1 + 0.8666B)(1 + 0.817B^{12})}{(1 - 0.4997B)(1 - 0.0368B^{12})(1 - B)(1 - B^{12})}\varepsilon_t \tag{16}$$

Where D= Dummy.

It is observed from Table 4 that the *p*-values of the parameters of the selected model for the seasonal and non-seasonal AR and MA components were highly significant at the 5% level of significance. The model thus appears to be the best model among the proposed models.

Table 4. Estimated parameters for SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$

Variable	Coefficient	Standard error	*z*-statistic	*P*-value
Dummy	-0.1467	0.1307	1.1220	.0262
ϕ	0.4997	0.1350	0.2692	.0007
Φ	0.0368	0.1367	3.7020	.0002
θ	0.8666	0.0860	-10.0765	.0000
Θ	0.8170	0.1723	-4.5563	.0000

The selection of the best model among competing models to fit a data in time series analysis depends heavily on the performance of the residuals of the model. In ARIMA modelling, it is assumed that the residuals of a good model follow a white noise process. This means that the residuals must have zero mean, constant variance and uncorrelated. It was observed from the diagnostic plot in Fig. 4 that the standardised residuals of the model have zero mean and constant variance. Further, the ACF plot of the model residuals revealed that not less than 95% of the residual autocorrelations lie within the significance bounds indicating that they are uncorrelated. In addition, the Ljung-Box statistic clearly shows that the *p*-values of the test statistic exceed the 5% level of significance for all lag orders which suggests that there is no significant departure from white noise for the residuals.

An ARCH-LM test was carried out to test for the assumption of constant variance and zero mean in order to further confirm the information depicted in Fig. 4. The ARCH-LM test results showed that there was no ARCH effect in the residuals of the selected model. Also, the JB test for normality was carried out to test for the normality of the model residuals and it was clearly shown that the model residuals were normally distributed. It can therefore be concluded that the selected model, SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$ was the best model since it satisfied all the diagnostic conditions.

Table 5. ARCH-LM test of residuals of SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$

Lag	Chi-squared	Df	*P*-value
12	42.667	12	.5765
24	50.925	24	.8046
36	53.039	36	.8989

JB Test: Chi-square=18.574, p-value=.4783

A chi-square goodness-of-fit test was carried out to cross validate the model. This was done to determine whether there is a significant difference between the monthly expected numbers of Out-patient cases and that of the observed values for the year 2018. The test produced a chi-square calculated value of 14.436 and a critical value of 19.675 at 11 df and 5% significance level. We therefore accept the null hypothesis of no significance difference between the observed and the expected values since the calculated value is less than the critical value and conclude that there is no significant difference between the predicted values from the model and the observed values for the year 2018.

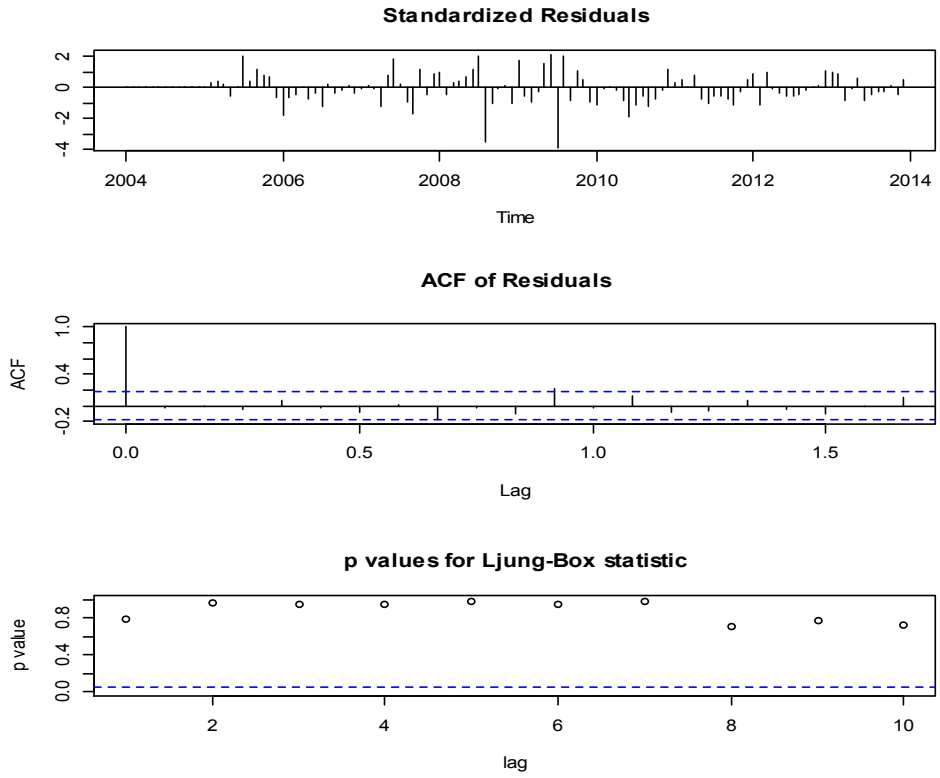

Fig. 4. Diagnostic plot of SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$

3.2 Estimating the SARIMAX model for the In-patient

The ACF plot in Fig. 5 shows significant spikes at the non-seasonal lag 1 and seasonal lag 12, with significant spikes at other non-seasonal lags. Also, the PACF plot has significant spikes at the non-seasonal lags 1 and 2 and seasonal lags 12 and 24.

Table 6. Tentative SARIMAX models for In-patient cases

Model	AIC	AICc	BIC
SARIMAX (1, 1, 1)(1, 1, 1)$_{12}$	38.7354	39.0833	52.0996
SARIMAX (2, 1, 1)(1, 1, 1)$_{12}$	39.3900	39.7378	55.4270
SARIMAX (1,1 ,1)(2, 1, 1)$_{12}$	37.9810*	38.3288*	52.0181*
SARIMAX (2, 1, 1)(2, 1,1)$_{12}$	38.9262	39.2741	53.6361

**: Means best based on the selection criteria*

Using the lower significant lags of both the ACF and PACF and their respective seasonal lags, a number of possible SARIMAX models (regression with ARIMA errors) were identified for the In-patient malaria cases. Comparing the AIC, AICc and BIC values of the various candidate models shown in Table 6, SARIMAX (1, 1, 1)(2, 1, 1)$_{12}$ emerged as the best model.

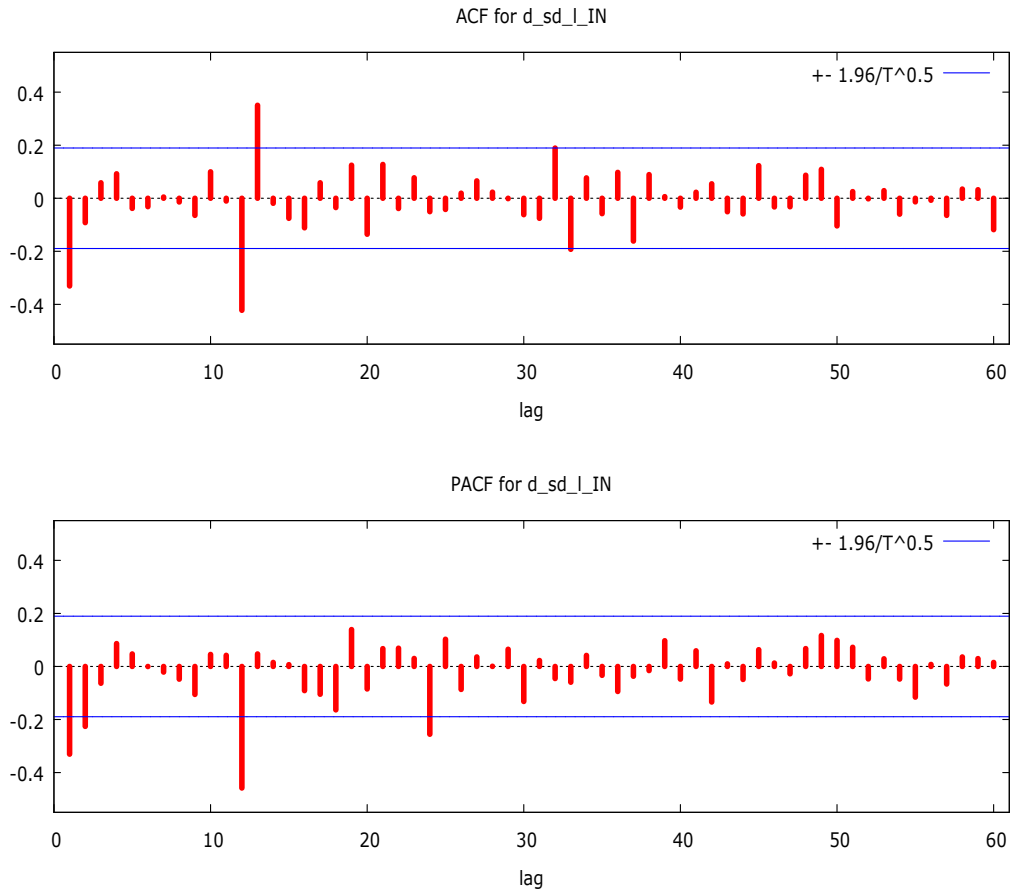

Fig. 5. ACF and PACF plot of differenced series for In-patient cases

Using the backshift operator, the parameters of SARIMAX (1, 1, 1)(2, 1, 1)$_{12}$ shown in Table 7 can be expressed as:

$$\ln \mathrm{Inp} = -0.0262D + \frac{(1 + 0.9334B)(1 + 0.2970B^{12})}{(1 - 0.6824B)(1 - 0.5098B^{12})(1 - B)(1 - B^{12})}\varepsilon_t \qquad (17)$$

Where D= Dummy, Inp= In-patient.

From Table 7, it was observed that the *p*-values of the parameters of the selected model for the non-seasonal and seasonal AR and MA components were highly significant at the 5% level of significance. Thus, the model is regarded as the best model.

We observed from Fig. 6 that the standardised residuals of the model have zero mean and constant variance. Further, the ACF plot of the model residuals revealed that not less than 95% of the residual autocorrelations lie within the significance bounds indicating that they are uncorrelated. In addition, the Ljung-Box statistic

clearly shows that the *p*-values of the test statistic exceed the 5% level of significance for all lag orders which implies that there is no significant departure from white noise for the residuals.

Table 7. Estimated parameters for SARIMAX (1, 1, 1)(2, 1, 1)$_{12}$

Variable	Coefficient	Standard error	*z*-statistic	*P*-value
Dummy	-0.0262	0.0217	0.1193	.0091
ϕ	0.6824	0.0959	7.1160	.0001
Φ	0.5098	0.2226	-2.290	.0022
θ	0.9334	0.0596	-15.6627	.0000
Θ	0.2970	0.2351	-1.2631	.0021

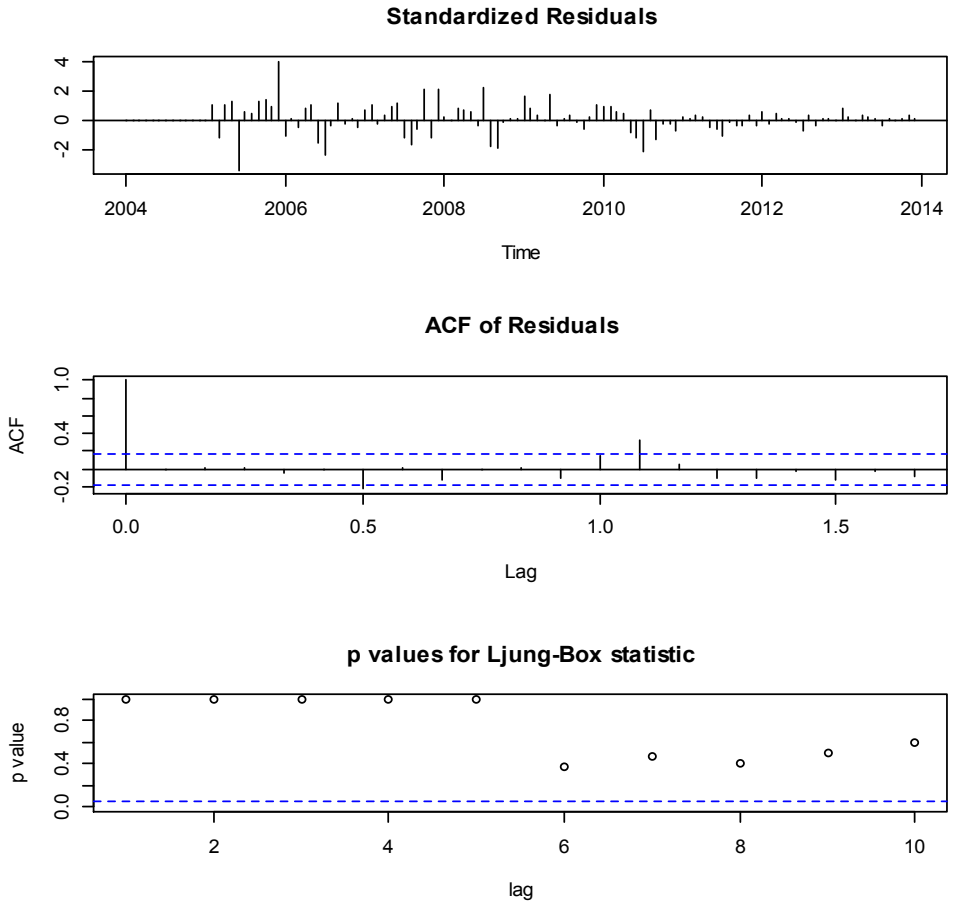

Fig. 6. Diagnostic plot of SARIMAX (1, 1, 1)(2, 1, 1)$_{12}$

On the assumption of constant variance and zero mean, an ARCH-LM test was carried out in order to further affirm the information depicted in Table 8. The results shown in Table 8 revealed that there was no ARCH effect in the residuals of the selected model. Also, the JB test for normality was conducted to test for the normality of the residuals of the model and it was affirmed that the residuals of the model were normally distributed. It can therefore be concluded that the selected model, SARIMAX (1, 1, 1)(2, 1, 1)$_{12}$ was the best model since it satisfied all the diagnostic conditions.

A chi-square goodness-of-fit test was carried out to cross validate the selected model to determine whether there is a significant difference between the monthly expected numbers of In-patient cases and that of the

observed for the year 2018. The results revealed a chi-square calculated value of 11.794 and a critical value of 19.675 at 11 degrees of freedom and 5% significance level. Hence, we accept the null hypothesis of no significant difference between the observed and the expected values since the calculated value is less than the critical value and conclude that there is no significant difference between the predicted values from the model and the observed values for the year 2018.

Table 8. ARCH-LM test of residuals of ARIMAX (1, 1, 1)(2, 1, 1)$_{12}$

Lag	Chi-squared	df	*P*-value
12	11.6015	12	.9333
24	28.7067	24	.7342
36	23.3458	36	.9487

JB Test: Chi-squared=14.4532, p-value=.5322

4 Conclusion

In this study, a change detection in Out-patient and In-patient malaria cases in the Northern Region of Ghana was assessed following the implementation of the malaria control intervention programme. The interventions covered the area of improving multiple prevention (through indoor residual spraying, scale-up of the distribution of insecticide treated nets, intermittent preventive treatment for pregnant women, improving sanitation, increasing access to healthcare and improving data reporting). It was observed from the results that the coefficients of the intervention variable for the Out-patient and In-patient cases were both significant (*p=0.0262* and *p value=0.0091* respectively) and negative (*-0.1467 and -0.0262* respectively). This is an indication that on the whole, cases decreased by so much between 2008 and 2018 (post-intervention period). The decline in the number of cases presupposes that the intervention policy was fairly well implemented and its plans were also executed fairly well. However, the government of Ghana and other partners in the health sector need to explore other reliable and feasible intervention strategies to support the already existing ones with a long-term goal of eliminating the malaria incidence. There is also the need for continuous monitoring of the forecasting performance of these models, and review of interventions in order to make the use of these models more realistic. Additionally, this study needs to be replicated in other high endemic regions in the country in order to be more conclusive on the effectiveness of the intervention policy.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Multiple Indicator Cluster Survey Final Report. Ghana Statistical Service; 2012.
[Accessed on 23 February 2018]
Available:http://www.statsghana.gov.gh.docfiles/publications/MICS4_MAIN_REPORT.pdf

[2] Landoh ED, Tchamdja P, Bayaki S, Khin ST, Gitta SN, Wasswa P, de Jager C. Morbidity and mortality due to malaria in Est Mono district, Togo, from 2005 to 2010: A times series analysis. Malaria Journal. 2012;11:389.

[3] Nyarango PM, Gebremeskel T, Mebrahtu G, Mufunda J, Abdulmumini U, Ogbamariam A, et al. A steep decline of malaria morbidity and mortality trends in Eritrea between 2000 and 2004: The effect of combination of control methods. Malaria Journal. 2006;5:33.

[4] Karema C, Aregawi CW, Rukundo A, Kabayiza A, Mulindahabi M, Ibrahima SF, et al. Trends in malaria cases, hospital admissions and deaths following scale-up of anti-malarial interventions, 2000–2010. Rwanda. Malaria Journal. 2012;11:236.

[5] Briet OJT, Amerasingbe PHA, Vounatsou P. Generalized seasonal autoregressive integrated moving average models for count data with application to malaria time series with low case numbers. PLos ONE. 2013;8(6):e65761.

[6] Osadolor E, Gebreslasie M, Magubane L. Modelling malaria control intervention effect in Kwazulu-Natal, South Africa using intervention time series analysis. Journal of Infection and Public Health. 2017;10:334-38.

[7] Anokye R, Acheampong E, Owusu I, Obeng EI. Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. Cogent Social Sciences. 2018;4:1461544.

[8] Hassan HE, Bin Y. Time series analysis and forecasting model for monthly malaria infection by Box-Jenkins techniques in the Kass Zone, South Darfur State, Sudan. Journal of Scientific and Engineering Research. 2018;5(9):35-42.

[9] Alhassan EA, Adjei MI, Aidoo E. Time series analysis of malaria cases in Kasena Nankana Municipality. ResearchGate. 2017;312574174.

[10] Anwar MY, Lewnard JA, Parikh S, Pitzer V. Time series analysis of malaria in Afghanistan: Using ARIMA models to predict future trends in incidence. Malaria Journal. 2016;15:566.

[11] Ankamah S, Nokoe KS, Iddrisu WA. Modelling trends of climatic variability and malaria in Ghana using vector autoregression. Hindawi; 2018.

[12] Perez EG, Ceballos RF. Malaria incidence in the Philippines: Prediction using autoregressive moving average models. International Journal of Engineering and Future Technology. 2019;16(4):1-10.

[13] Box GEP, Jenkins GM. Time series analysis: Forecasting and control. Holden-Day, San-Francisco; 1976.

[14] Dickey DA, Fuller WA. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica. 1981;49:1057-1072.

[15] Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974;19(6):716-723.

[16] Eagle RF. Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. Econometrica. 1982;50:987-1008.

[17] Ljung GM, Box GEP. On a measure of lack of fit time series models. Biometrika. 1978;65:297-303.

[18] Jula D. Introduction in econometrics. Ed. Professional Consulting, Bucuresti; 2003.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sdiarticle4.com/review-history/54310