Canadian Journal of Pure and Applied Sciences
Vol. 14, No. 2, pp. 5049-5059, June 2020
Online ISSN: 1920-3853; Print ISSN: 1715-9997
Available online at www.cjpas.net

TIME SERIES ANALYSIS AND FORECAST OF INFANT MORTALITY RATE IN NIGERIA: AN ARIMA MODELING APPROACH

*Friday Ewere and Donalben Onome Eke
Department of Statistics, Faculty of Physical Sciences, University of Benin
Benin City, Edo State, Nigeria

ABSTRACT

Childhood mortality in general and infant mortality in particular has long been a public health menace in Nigeria. Identified as one of the barometers for the measurement of any population's state of health, health facilities and well being, relevant authorities in government and stakeholders in public health have all moved to reduce and possibly eliminate its occurrence with little success. This is evident in the fact that Nigeria was one of the countries that failed to meet the Millennium Development Goal (MDG) for the reduction of childhood mortality by two-thirds in 2015. Having failed to achieve MDG 4, genuine concerns of her ability to achieve the Sustainable Development Goal (SDG) 3.2 by 2030 has led to an inquest into the country's chances of reducing childhood mortality rate occurring within the first year of life. The present study utilized the Auto-Regressive Integrated Moving Average (ARIMA) model for to make forecast of infant mortality in Nigeria up to the year 2030 using data obtained from the United Nation's Inter Agency Group for Childhood Mortality Estimation (UN-IGME). The ARIMA (1, 1, 1) model selected predicted a reduction of up to 30% by 2030 at 95% confidence interval.

Keywords: Nigeria, ARIMA, forecasting, time series analysis, infant mortality.

INTRODUCTION

Infant mortality is the death of a child within the first year of life (World Health Organization- WHO, 2011) while Infant mortality rate is the number of deaths under one year of age occurring among the live births in a given geographical area during a given year per 1,000 live births occurring among the population of the given geographical area during the same year (Organization for Economic Co-operation and Development- OECD, 2001).

Infant mortality has long been identified as an important tool for the evaluation of a population's health and health care system and a barometer for the measurement of a county's well being and the state of health and health facilities (Madise *et al*., 2001). The realization of the importance of infant mortality and the need to reduce its occurrence has led to the fight against mortality in infants both at the national and international level.

Highlighted as an area of public health concern in the society, the fight against infant mortality led to the implementation of the United Nations Millennium Development Goals. The fourth MDG in particular, addresses the issue of reducing mortality worldwide by up to two thirds in 2015 (Sachs and Mc Arthur, 2005). This global fight against the deaths of children within the first year of their lives which has been on for a few decades saw most European countries make significant progress between 1990 and 2015 sufficient enough to meet the childhood mortality reduction target set in the fourth tenet of the MDG. However, majority of the countries in Asia and Africa (Nigeria inclusive) failed to meet the recommended target for the reduction of childhood mortality and as a result, the Sustainable Development Goal was rolled out.

In Nigeria, infant mortality rate still remains unacceptably high despite modest improvement in child health outcomes during the 20th century (Adetoro and Amoo, 2014). The current infant mortality rate in Nigeria is estimated at 70 per 1000 live births implying that 1 in 15 live births in Nigeria die before their first birthday (National Bureau of Statistics- NBS, 2017). These rates remain high in comparison to the proposed target set by the United Nations as contained in the SDG 3.2 objective.

The motivation for this study is drawn from the realization that knowledge of reliable and accurate forecasts of infant mortality rate is both necessary and important for the planning of suitable intervention programs and preventive measures for the reduction of infant mortality in Nigeria. Results from several studies have shown that the ARIMA model if carefully selected

*Corresponding author e-mail: friday.ewere@uniben.edu

can be fitted to the time series data of single variable to forecast, with better accuracy, the future values in the series.

In a recent study Usman *et al*. (2019) critically examined the incidence of neonatal mortality in Nigeria between 1990 and 2017 using ARIMA techniques of time series to make forecast. Another study Kumar and Anand (2012) employed the ARIMA technique to make forecast of sugarcane production in India between 2013 and 2017 using the past 62 years sugarcane production data. The result of their research which went further to statistically test and validates the forecast errors suggested that the ARIMA procedure is an adequate predictive model capable of making reliable forecast of univariate time series data.

This paper attempts to forecast the future rate of occurrence of infant mortality in Nigeria using the ARIMA model. This study will be beneficial to both women's health and infant and child survival in Nigeria.

MATERIALS AND METHODS

Methodology

Study Area

This paper is focused on Nigeria. Nigeria is made up of 36 states and a Federal Capital Territory (FCT), grouped into six geopolitical zones; North Central, North-East, North-West, South-East, South-South, and South-West.

Data Source

This study is conducted using data obtained from World Bank and estimated by United Nation Inter-Agency Group for Childhood Mortality Estimation (UN-IGME). The data, which is the mortality estimates for infants in Nigeria between 1964 and 2018, is presented in Table 1.

Table 1. Estimate of infant mortality rate in Nigeria (1964 - 2018).

S. No.	Year	Infant Mortality	S/N	Year	Infant Mortality	S/N	Year	Infant Mortality
1	1964	195	**20**	1983	123	**38**	2001	107.8
2	1965	191	**21**	1984	123	**39**	2002	104.8
3	1966	187	**22**	1985	124	**40**	2003	101.6
4	1967	182	**23**	1986	124	**41**	2004	98.6
5	1968	178	**24**	1987	125	**42**	2005	95.6
6	1969	174	**25**	1988	125	**43**	2006	92.8
7	1970	169	**26**	1989	125	**44**	2007	90.2
8	1971	164	**27**	1990	125.1	**45**	2008	87.9
9	1972	159	**28**	1991	124.9	**46**	2009	85.9
10	1973	154	**29**	1992	124.5	**47**	2010	84.1
11	1974	149	**30**	1993	124	**48**	2011	82.7
12	1975	144	**31**	1994	123.4	**49**	2012	81.5
13	1976	140	**32**	1995	122.3	**50**	2013	80.5
14	1977	135	**33**	1996	120.9	**51**	2014	79.6
15	1978	132	**34**	1997	118.9	**52**	2015	78.7
16	1979	129	**35**	1998	116.5	**53**	2016	77.9
17	1980	126	**36**	1999	113.7	**54**	2017	76.9
18	1981	125	**37**	2000	110.9	**55**	2018	75.7
19	**1982**	**124**						
Source: world bank, www.childmortality.org								

Data Analysis

The data used for this study was analyzed using SPSS software, version 22. Time series analysis using the Auto Regressive Integrated Moving Average (ARIMA) technique as proposed by Box and Jenkins (1976) was employed to make the forecast. The distribution of the prediction errors and the prediction confidence intervals are obtained.

The data in Table 1 was examined for non stationarity. Furthermore, the Augmented Dickey Fuller test (Dickey and Fuller, 1979) was conducted.

Forecast for infant mortality rate in Nigeria was made up to the year 2030. The forecast error and the degree of accuracy of the forecast were calculated. To do this, the last five observed values of infant mortality rates (2014 -2018) were not included in the analysis.

RESULTS AND DISCUSSION

Figure 1 shows the time series plot of the mortality rates of infants in Nigeria between 1964 and 2013. A visual inspection of the time series plot in Figure 1 indicates that the series appears to be decreasing at a constant rate implying that the series exhibits trends and seasonality, which are characteristic of non stationary time series. To confirm the non stationarity of the series as indicated by the time series plot in Figure 1, the plots of the Auto Correlation and Partial Autocorrelation Functions also known as the correlogram are presented in Figure 2 and inspected.

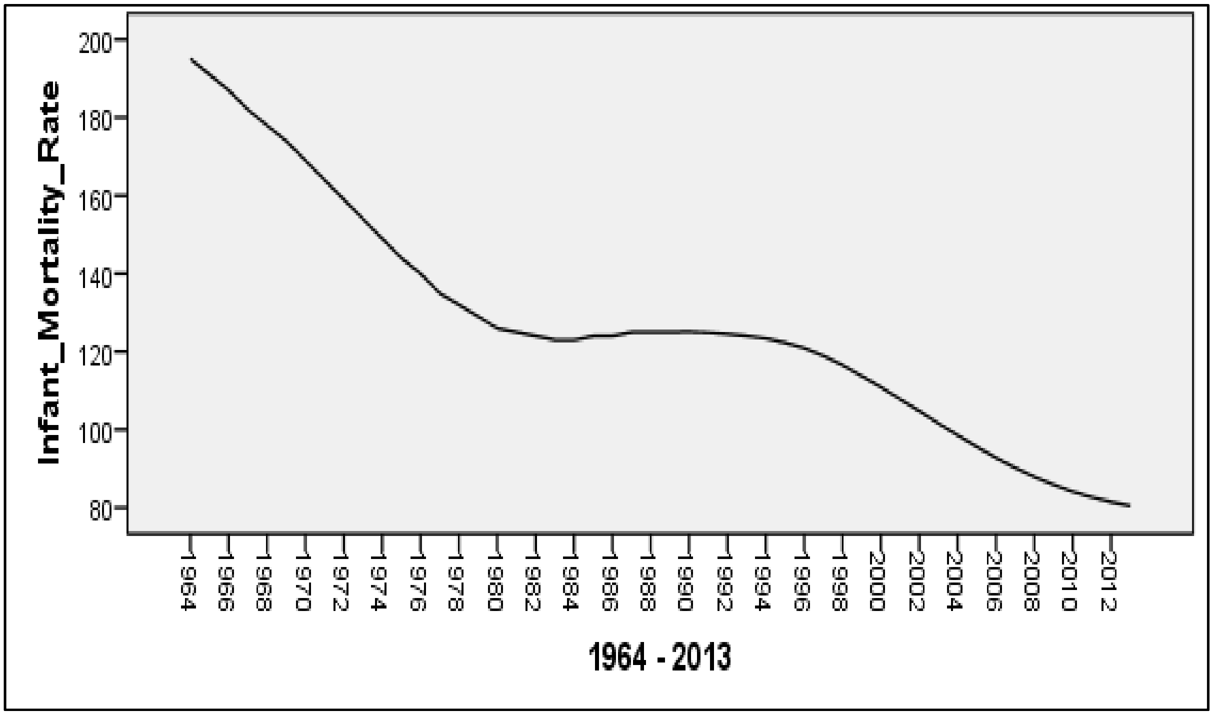
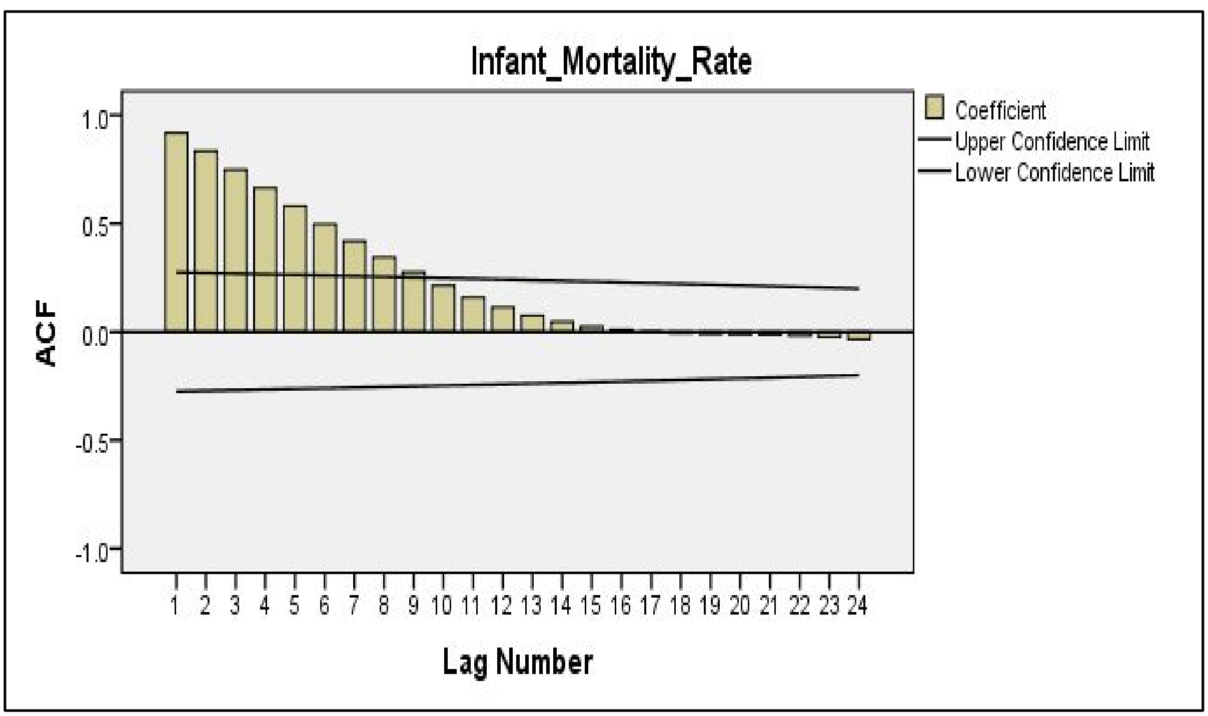


Fig. 1. Time series plot of infant mortality rate in Nigeria.



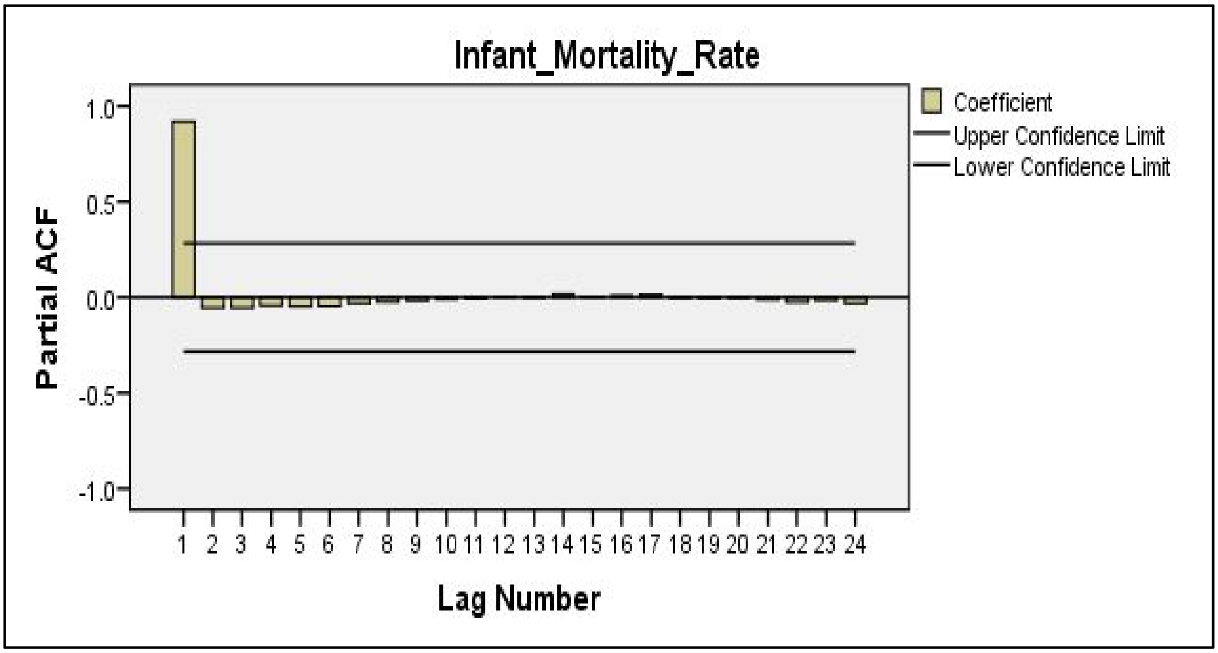


Fig. 2. Correlogram of the time series.

Figure 2 displays graphically the Auto Correlation Function (ACF) and the Partial Auto Correlation Function (PACF) for the time series plot of infant mortality rate in Nigeria. Figure 2 shows large significant ACF for the time lags which gradually decreases in size, but do not decay to zero (slow decay). The ACF thus shows a pattern typical of a non-stationary time series. In the PACF plot, the partial autocorrelation at time lag 1 is close to one and the partial autocorrelations for the time lag 2 through 10 are close to zero which is also typical of non-stationary series.

The series is therefore transformed into a stationary series by taking the first difference of the time series data and then making the time plot. To obtain the first difference of the time series data, we use the transformation $W_t = \Delta y_t - y_{t-1}$.

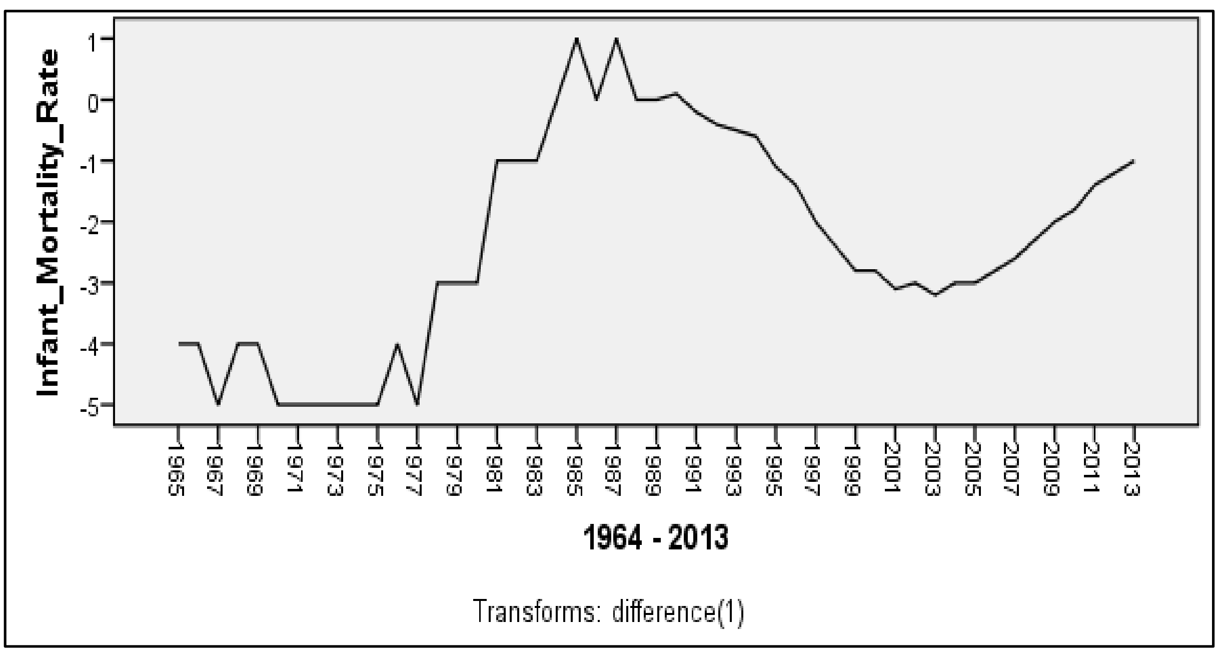


Fig. 3. Time series plot of the differenced series (d=1).

A visual inspection of the time plot of the differenced series in Figure 3 suggests that the time series appears stationary and devoid of trend, an implication that the mean and the variance now seem consistent over time.

We apply the unit root test for stationarity. The unit root test proposes that if $\emptyset_1 = 1$ in the simple autoregressive scheme $Z_t = \mu + \emptyset_1 Z_{t-1} + \varepsilon_t$ where $\emptyset_1$ is the coefficient of the auto-regression process, then the series is said to possess a unit root and is thus not stationary as it violates the basic requirement of zero mean and constant variance necessary for time series modeling (Dickey and Fuller, 1979).

To test for stationarity in the time series data, we adopt the following hypothesis:
H_0**:** the series contains a unit root
H_1**:** there is no unit root in the series

We shall reject the null hypothesis and conclude that the alternative hypothesis is true if the p-value is greater than the significance level: $\alpha = 0.01$.

Table 2. Augmented Dickey-Fuller Tests.

Dickey-Fuller	Sig.	ACF coeff. of e (d=1)	DF
-4.182	0.038	0.140	16

Since the computed p-value = 0.038 in Table 2 is greater than $\alpha = 0.01$, we reject the null hypothesis and conclude that the alternative hypothesis H_1 is true implying that the series is stationary after taking the first difference (d = 1).

Further proof of the stationarity of the series is obtained by the computation and plots of the autocorrelation and partial autocorrelation function at multiple time lags.

Table 3. Autocorrelation and Partial Autocorrelation of the differenced series.

Lag	Autocorrelation	Std. Error[a]	Partial Autocorrelation	Std. Error
1	-.253	.140	-.253	.144
2	.157	.138	.100	.144
3	.412	.137	.510	.144
4	-.017	.135	.280	.144
5	.026	.134	-.105	.144
6	.167	.132	-.211	.144
7	-.008	.131	-.124	.144
8	-.187	.129	-.292	.144
9	.113	.127	-.115	.144
10	-.139	.126	-.029	.144
11	-.169	.124	.020	.144
12	-.138	.122	-.216	.144
13	.001	.121	.015	.144
14	-.302	.119	-.058	.144
15	-.121	.117	-.117	.144
16	-.042	.115	-.123	.144
17	-.152	.114	.130	.144
18	-.143	.112	.035	.144
19	.021	.110	-.017	.144
20	-.175	.108	-.286	.144
21	.019	.106	-.033	.144
22	.002	.104	-.046	.144
23	-.050	.102	.026	.144
24	.061	.100	.039	.144

The ACF and PACF plots of the differenced series of infant mortality in Table 3 from lags 1 to 24 indicate that we have an AR process.

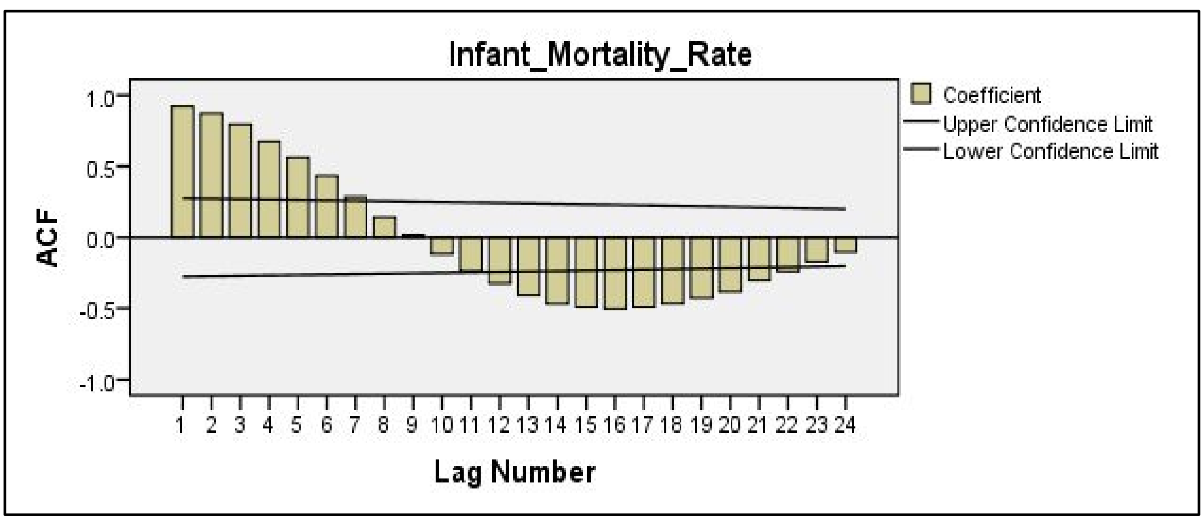


Fig. 4a. ACF of the first differenced series.

Figure 4a displays the autocorrelation function from lags 1 to 24 of the first order differenced time series of infant mortality rate in Nigeria. Figure 4a shows that the ACF dies down to zero slowly as the number of time lags k increases indicating a geometric decay symbolic of an AR process.

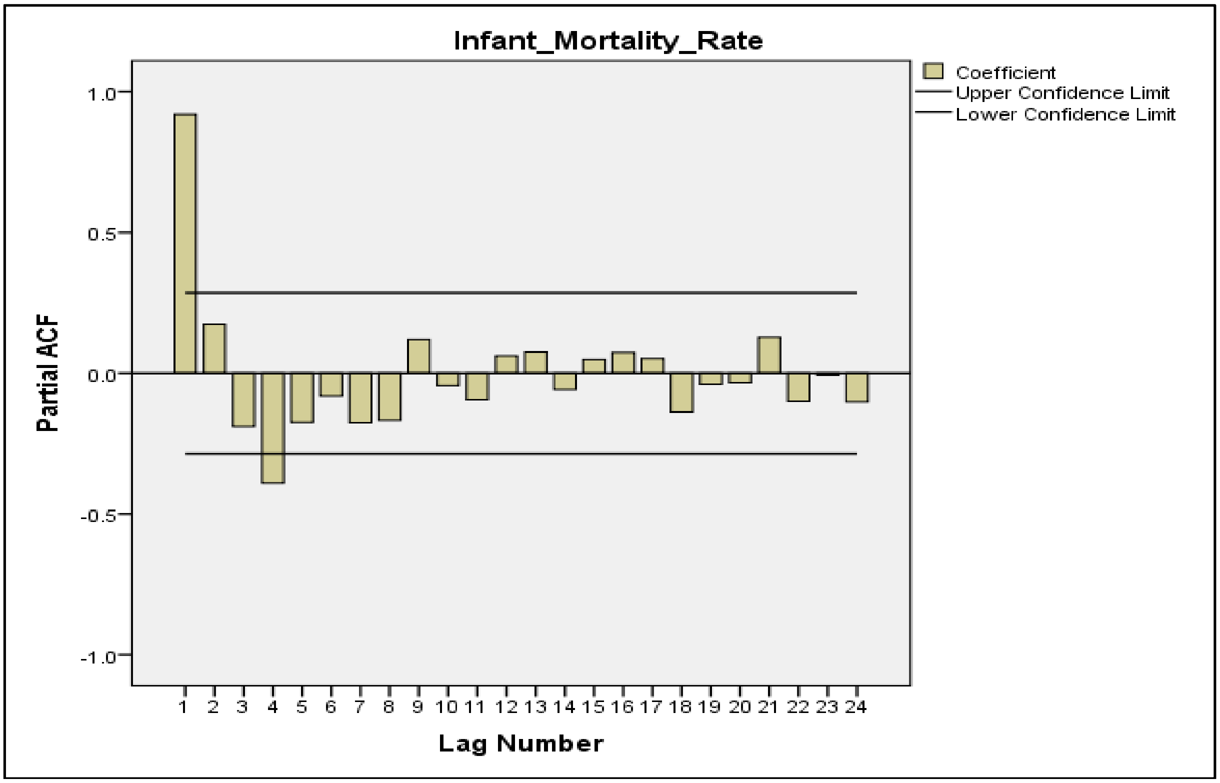


Fig. 4b. PACF of the first differenced series.

The PACF in Figure 4b shows that there are only two significant time lags. Only time lags 1 and 4 exceeds the significant threshold which is the error bound indicating an AR (2) model. All other time lags fall within the significant threshold.

To determine the best ARIMA model to fit the infant mortality rate series, the Aikaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are calculated and compared for a set of tentative ARIMA models chosen based on the principle of parsimony and recognizable patterns of the ACF and PACF of the time series data.

Table 4. Tentative ARIMA models for infant mortality.

Models	**AIC**	**BIC**
ARIMA (1, 1, 0)	23.24	34.51
ARIMA (1, 1, 1)	23.45	25.10
ARIMA (2, 1, 1)	28.73	33.81
ARIMA (1, 1, 2)	36.14	43.01
ARIMA (2, 1, 0)	28.45	24.91

A comparison of the AIC and the BIC values in Table 4 for the possible models shows that ARIMA (1, 1, 0) has the least AIC value. However, the ARIMA (1, 1, 1) produced least values for both the AIC and the BIC and is therefore selected as the best fit model for the data.

Table 5. Estimated Summary for the ARIMA (1, 1, 1) model.

ARIMA Model Parameters

	Estimate	SE	t
Constant	-2.406	1.204	-1.997
AR Lag 1	.950	.046	20.443
Difference	1		
MA Lag 1	.151	.158	.958

The ARIMA (1, 1, 1) model for infant mortality in Nigeria is thus given as

$Z_t = \mu + Z_{t-1} + \emptyset_1(Z_{t-1} - Z_{t-2}) - \theta_1 \varepsilon_{t-1}$ and

$$Z_t = -2.406 + Z_{t-1} + 0.950\,(Z_{t-1} - Z_{t-2}) - 0.151\varepsilon_{t-1}$$

where Z_t is the stationary series observed, μ is the mean of the time series (constant), $\emptyset_i$ and θ_1 are the parameter estimates and ε_t is the white noise with zero mean and constant variance.

The residual ACF and PACF plots given in Figure 5 are used to check the adequacy of the model as a good fit of the time series data before it is used for forecast. A model is adequate if the residuals left over after fitting the model are simply white noise. In this case, the residuals should be uncorrelated with constant variance.

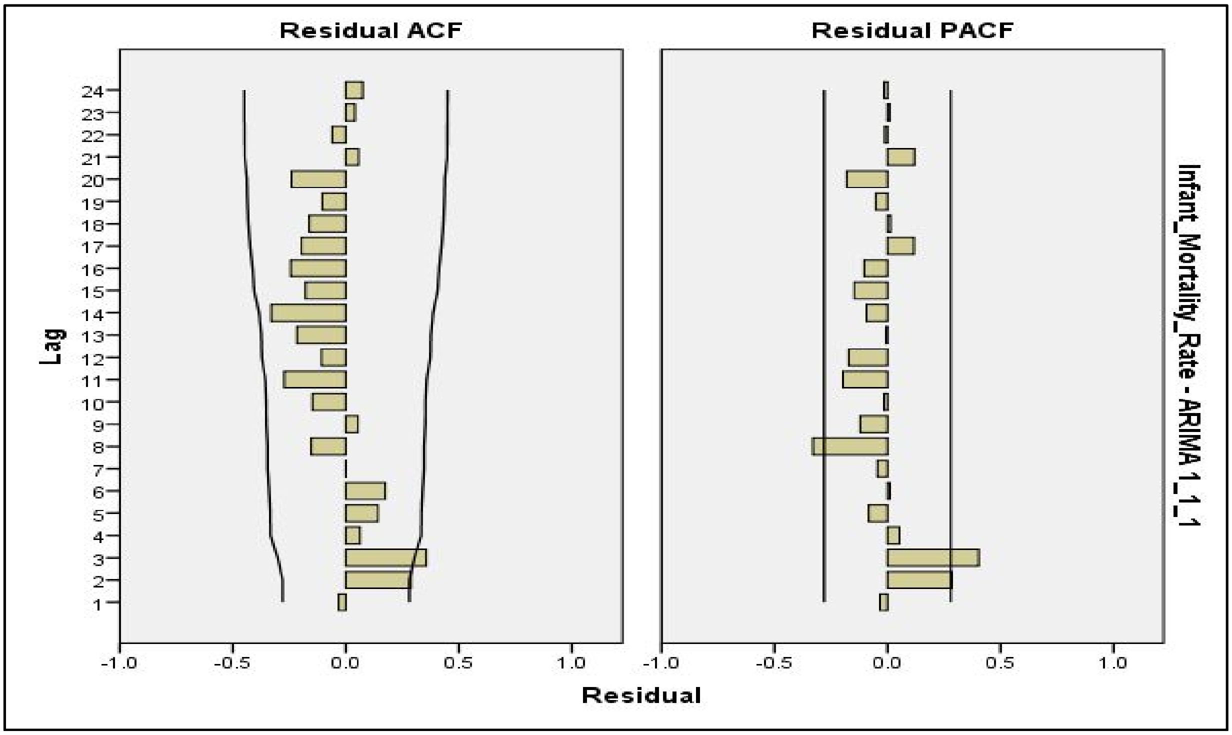


Fig. 5. Residual ACF and PACF plots.

A diagnostic of the residuals by the ACF and PACF shows that their values are all within the 5% zero bound indicating that there is no correlation amongst the residuals.

The K period ahead forecast based on the ARIMA (1, 1, 1) model is given by

$$\hat{Z}_{t+k} = \mu + Z_{t+k-1} + \emptyset_1(Z_{t+k-1} - Z_{t+k-2}) - \theta_1\varepsilon_{t+k-1}$$

where

$\mu = -2.406, \emptyset_1 = 0.950$ and $\theta_1 = 0.151$

so that

$$\hat{Z}_{t+k} = -2.406 + Z_{t+k-1} + 0.950\ (Z_{t+k-1} - Z_{t+k-2}) - 0.151\varepsilon_{t+k-1}$$

After fitting the ARIMA (1, 1, 1) model to forecast for the future values of mortality in infants up to year 2030, the following forecast estimates in Table 6 are obtained.

Table 6. Forecast of Infant Mortality Rate in Nigeria.

Forecast

Year	Model		
	Infant_Mortality_Rate-ARIMA 1_1_1		
	Forecast	UCL	LCL
2014	79	81	78
2015	78	81	76
2016	77	81	73
2017	76	82	70
2018	74	82	66
2019	73	83	63
2020	71	83	59
2021	70	84	56
2022	68	85	52
2023	67	85	48
2024	65	86	44
2025	63	87	40
2026	62	87	36
2027	60	88	32
2028	58	89	28
2029	56	89	24
2030	55	90	19

Table 6 shows the forecast of infant mortality rate in Nigeria for the years 2014 through 2030 with 95% upper and lower confidence intervals. Result from the forecast shows that there will be about 30% reduction in the occurrence of infant mortality by 2030 which is an Annual Reduction Rate (ARR) of 1.9%.

Figure 6 shows the plot of the observed series from 1964 to 2013 (Table 1) and the forecast for 2014 up to year 2030 of infant mortality in Nigeria by fitting the ARIMA (1, 1, 1) model to the time series data.

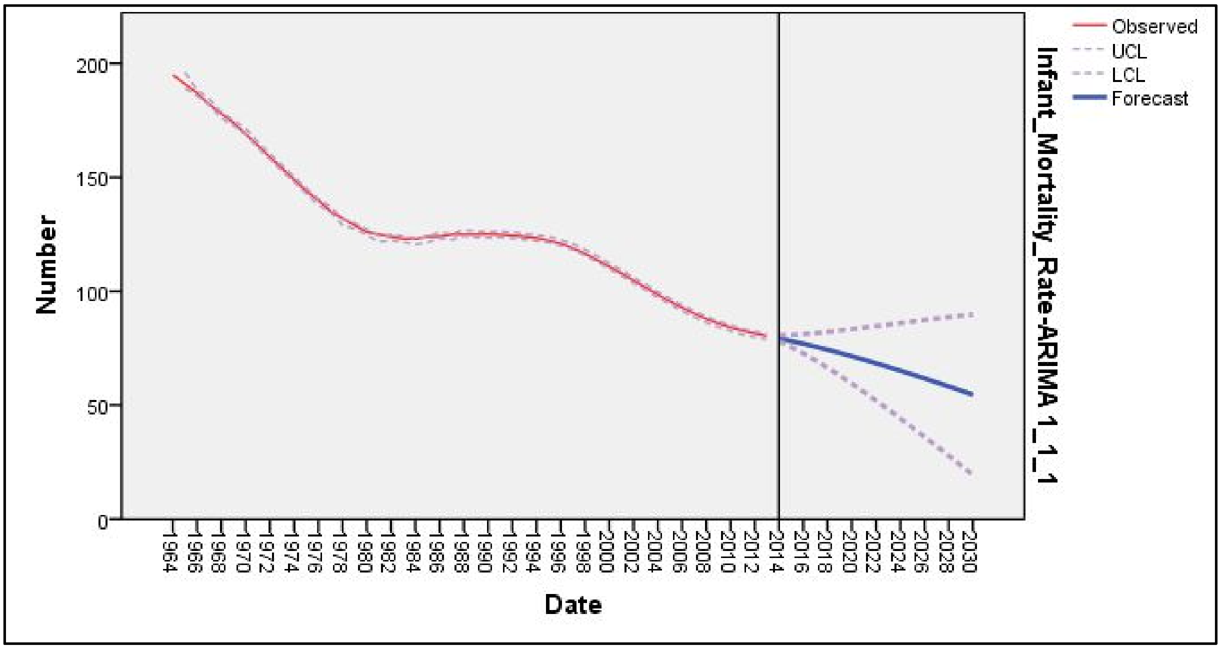


Fig. 6. Plot of the observed series and the in-sample forecast value.

The portion of the plot from the year 2014 to 2030 shows the in - sample estimated and forecast values of infant mortality rate in Nigeria by the ARIMA (1, 1, 1) model produced from the data with the 95% confidence interval for each of the forecast value, depicted by the two opening lines. The forecast estimate shows that by 2030, Nigeria will have a reduced mortality rate of 55 deaths per 1,000 live births for infant mortality. Whilst this would be an improvement compared to previous mortality rates, it would still fall significantly short of the global target for childhood mortality.

To meet the recommended set target for the reduction of child mortality rate in Nigeria as proposed by the United Nations in the Sustainable Development Goal by 2030, there would have to be a significant increase in the ARR of infant mortality in the country.

Distribution of the Forecast Error

We compare the first five forecast values (2014 – 2018) of infant mortality rate in Nigeria obtained from fitting the time series data using the ARIMA (1, 1, 1) model with the observed values in Table 1 and obtained the distribution of the forecast error given in Table 7.

Table 7. Distribution of the forecast error of infant mortality rate in Nigeria.

Year	Observed Value	Forecast	Forecast Error
2014	79.6	79	0.75
2015	78.7	78	0.88
2016	77.9	77	1.16
2017	76.9	76	1.17
2018	75.7	74	2.25

Table 7 shows the forecast error of infant mortality rate in Nigeria for the five years between 2014 and 2018. The Mean Absolute Percentage Error (MAPE) of the forecast for infant mortality rate in Nigeria is 1.241 indicating about 99% forecast accuracy. To further investigate if the mean is normally distributed and to check for non-zero autocorrelation in the forecast residuals, we obtain the ACF and PACF of forecast error (residual) and the QQ plot of forecast error.

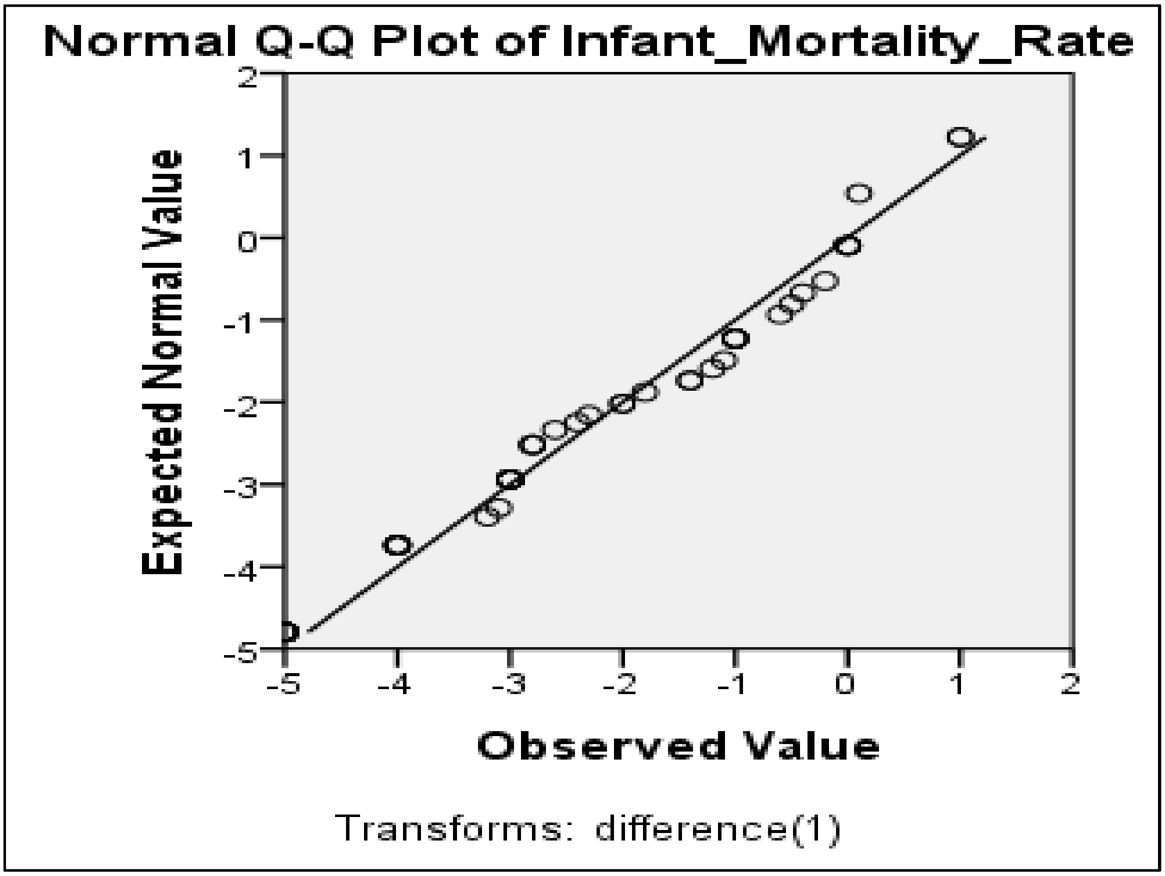
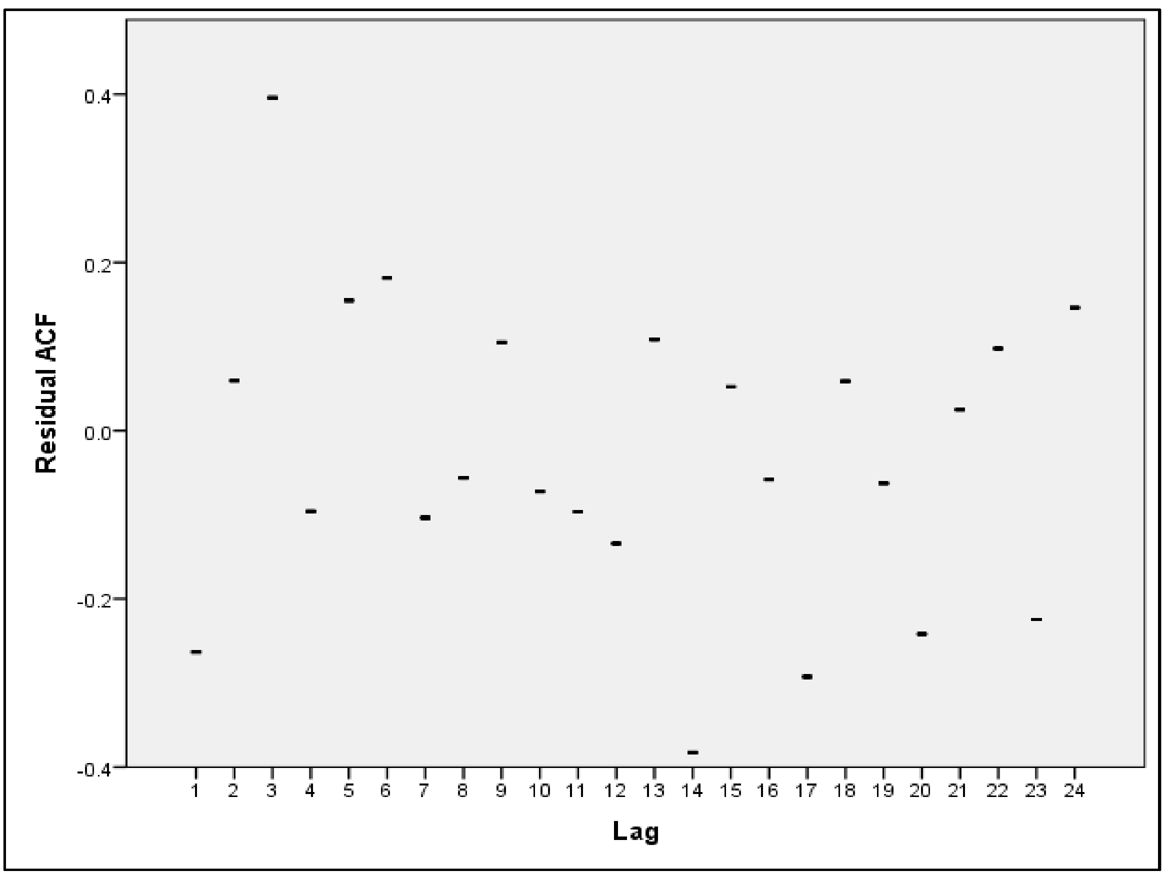


Fig. 7. Normal QQ Plot of residuals (Forecast error).



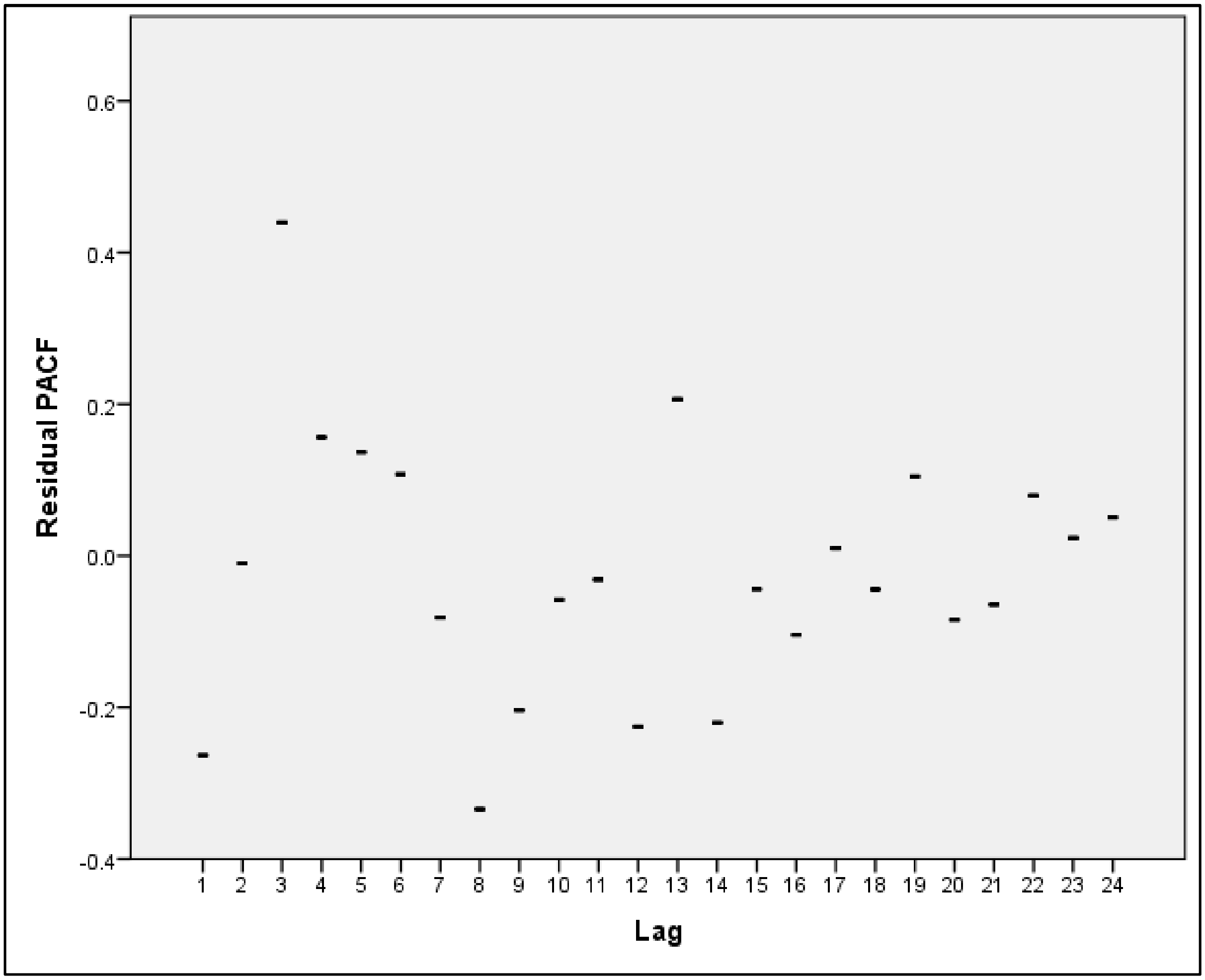


Fig. 8. Estimated ACF and PACF of Residuals (Forecast Errors) – ARIMA (1, 1, 1).

The straight line exhibited by the QQ plot in Figure 7 indicates that the forecast error (residuals) is normally distributed while the residual ACF and PACF of the fitted ARIMA (1, 1, 1) in Figure 8 shows that there is no non-zero autocorrelations between successive forecast errors since all the residuals are within the significant threshold for each of the time lags.

CONCLUSION

This study applied the ARIMA (1, 1, 1) model to make forecast of infant mortality rate in Nigeria up to the year 2030. The time plot of the forecast showed a downward movement suggesting that there would be a continuous decrease in infant mortality rate with a reduction of up to 30% by 2030 at 95% confidence interval. Despite this reduction in infant mortality by 2030, Nigeria may still fall short of the Sustainable Development Goal for the reduction of childhood death if preventive and intervention measures already put in place are not vigorously pursued and implemented.

ACKNOWLEDGEMENT

The authors would like to thank the World Bank Group and the United Nation Inter-Agency Group for Childhood Mortality Estimation (UN-IGME) for access to the data set used in this study.

REFERENCES

Adetoro, GW. and Amoo, EO. 2014. A Statistical Analysis of Child Mortality: Evidence from Nigeria. Journal of Demography and Social Statistics, Department of Demography and Social Statistics, Obafemi Awolowo University, Ile-Ife. Nigeria. 1:110-120.

Box, GEP. and Jenkins, GM. 1976. Time Series Analysis, Forecasting and Control, Holden-Day, Oakland, CA, USA.

Dickey, DA. and Fuller, WA. 1979. Distribution of the Estimators for Autoregressive Time Series with a Unit

Root. Journal of the American Statistical Association. (366):427-431.

Kumar, M. and Anand, M. 2012. An application of time series ARIMA Forecasting Model for Predicting Sugarcane Production in India. Studies in Business and Economics, Lucian Blaga University of Sibiu. 9 (1):81-94.

Madise, N., Magadi, M. and Diamond, I. 2001. Factors Associated with Unfavourable Birth Outcomes in Kenya. Journal of Biosocial Science. 33(2):199-225. https://doi.org/10.1017/S0021932001001997.

MDG Report. 2015. Nigeria Millennium Development Goals. www.mdgs.gov.ng.

National Bureau of Statistics- NBS and United Nations Children's Fund- UNICEF. 2017. Multiple Indicator Cluster Survey (MICS) 2016-17, Survey Findings Report. Abuja, Nigeria. National Bureau of Statistics and United Nations Children's Fund.

Organization for Economic Co-operation and Development- OECD. 2001. Handbook of Vital Statistics Systems and Methods. (vol. 1). :Legal, Organizational and Technical Aspects, United Nations Studies in Methods, Glossary, Series F, No. 35, United Nations, New York 1991 https://stats.oecd.org/glossary/detail.asp?ID=1347.

Sachs, JD. and McArthur, JW. 2005. The Millennium Project: a plan for meeting the Millennium Development Goals. Lancet. London, England: 365(9456):347 -353. https://doi.org/ 10.1016/S0140-6736(05)17791-5

Usman, A., Sulaiman, MA. and Abubakar, I. 2019. Trends of Neonatal Mortality in Nigeria from 1990 to 2017 using Time Series Analysis. J. Appl. Sci. Environ. Manage. 23(5):865-869.

World Health Organization- WHO. 2011. Child Mortality: Millennium Development Goal (MDG) 4. The Partner for Maternal and New Born Birth, World Health Organization. September, 2011. http://www.who.int/pmnch/media/press_materials/fs/fs_mdg4_childmortality/en/.

Received: May 9, 2020; Revised and Accepted: June 14, 2020