

AN ARMA MODELLING APPROACH ON INFANT MORTALITY RATE IN NIGERIA

ONYEMARIN N.N, OJARIKRE H.I AND JOHN N. IGABARI

onyemarinnelly@gmail.com, ojarikreify@delsu.edu.ng, jn_igabari@delsu.edu.ng

Department of mathematics,

Delta State University, Abraka, Delta State, Nigeria

ABSTRACT

This paper considers infant mortality in Nigeria. A time series plot of the infant mortality rate (IMR) showing trends and seasonality has been considered. A stochastic model that predicts the mortality rate of infant has been constructed. In the work, a non-stationary time series is first converted to a stationary series by obtaining the initial difference of the time series data and then constructing the time series plot. The Augmented Dickey Fuller Test was also determined, a 5 years prediction from 2022-2026 has been forecasted. The work suggest that there is a declining consistent pace over the years and the data used for the analysis is from World Bank and the United Nations Inter-Agency Group for Childhood Mortality Estimation (UN-IGME).

Keywords: ARMA, Bayesian Information Criterion (BIC), Box-Jenkins Approach, Ljung-Box Statistic, Time Series Analysis, infant mortality.

1.0 INTRODUCTION

Infant mortality is the death of a child in their first year of life (W.H.O, 2011), however infant mortality rate is defined as the number of deaths under the age of one year occurring among live births in a given geographical area during a given year per 1,000 live births occurring among the same geographical area's population during the same year (W.H.O, 2011). According to the Organization for Economic Cooperation and

Development (OECD), infant mortality has long been acknowledged as a key indicator of a population's health and health-care system, as well as a barometer for evaluating a county's well-being and health-care facilities (Nyongesa, et al., 2018). The recognition of the importance of infant death and the need to limit its incidence has led to a national and international struggle against infant mortality. Nigeria's neonatal death rate remains unacceptably high despite

major improvements in child health care outcomes during the twentieth century (Adetoro & Amoo, 2014). The current infant mortality rate in Nigeria is estimated to be 70 per 1000 live births, implying that one out of every 15 live births dies before reaching their first birthday (National Bureau of Statistics (NBS), 2017). The fight against infant mortality was highlighted as a public health priority in society, which led to the United Nation Millennium Development Goals (MDG) being implemented. The fourth focal point of MDG, focuses on reducing global mortality by up to two-thirds by 2015. (Sachs & Mc Arthur, 2005), Most European countries made great progress between 1990 and 2015, allowing them to reach the MDG's fourth tenet of reducing childhood mortality. The global community approved the Sustainable Development Goals (SDGs) in 2016, and Goal 3 which is good health and well-being for all ages, which the aim was to improve reproductive and maternal and child health. Countries are intended to reduce infant and mortality to 70 deaths per 100,000 live births by 2030. (W.H.O 2015), As a result, countries all over the world have established programs and policies to address the problem with the resources at hand. Both the government and individuals have been

concerned about the high rate of infant and mother mortality. Without a thorough understanding of the available data's, policies that can sufficiently arrest the situation cannot be implemented. However, the majority of Asian and African countries (including Nigeria) failed to meet the proposed objective for reducing childhood mortality, prompting the adoption of the Sustainable Development Goal. Ude & Ekesiobi (2014) used multiple regression analysis to empirically analyze the effect of per capita health expenditure on child mortality as measured by infant, under-5, and neonatal mortality rates. The findings indicate that Nigeria's infant and neonatal mortality rates are not significantly impacted by per capita health spending. The under-five mortality rate in Nigeria is significantly impacted by per capita education and health investment, according to the results. Similarly, Yaqub and Gul (2013) investigated the impact of governance in Nigeria, focusing on public health expenditure and subsequent health outcomes. They analyzed the data from 1980 to 2008 using Ordinary Least Square (OLS) and Two-Stage Least Square (2-SLS) methodologies, and discovered an inverse link between government health expenditure and infant mortality. David Joseph (2018)

employed the Autoregressive Distributed Lag (ARDL) bounds testing technique with Granger regression to investigate co-integration. The application of a causality technique is to study the nature of the relationship between infant mortality and public health spending in Nigeria from 1980 to 2016. The importance of immunization is also evaluated in the study. Nigerian infant mortality: external health resources and private health spending apart from that, in the presence of bidirectional causality, there is a significant co-integrating (long-run) link between infant mortality and empirical discoveries, government health spending (as well as private health spending), immunizations, and outside health resources. Furthermore, the findings show that government health spending, private health spending, immunization, and external health resources all have a detrimental impact on infant mortality in the long and medium term. Due to the size of the coefficient of private health expenditure, it is revealed to be the largest determinant of the decline of infant death rate in Nigeria. In essence, a thorough revamp of the Nigerian health system is advocated in order to increase efficiency and reduce occurrences of fund mismanagement that have plagued the sector throughout time, as well as the

intensification of immunization programs and activities. In a study conducted by Abraha and Nigatu 2009, changes in Ethiopia's health and health-related variables were examined, and models were created for their prediction. The study covered the years 1995 to 2008, Using ARIMA models in STATA, the causes of the established trends were discovered Only MMR was taken into consideration among the mortality markers in this study, it shown a statistically significant decline during the course of the trial. The Total's trends Rate of fertility, number of doctors per 100,000 people, number of competent birth attendants, and postpartum care were discovered to be significantly correlated with the MMR trend. The authors came to the conclusion that the existing trend suggests the necessity to speed up the indicators to meet MDGs by or before 2015, especially for access to medical care. In a recent study, Usman et al. (2019) looked at the incidence of infant mortality in Nigeria between 1990 and 2017, utilizing ARIMA time series techniques to anticipate. In another work, Kumar & Karthikeyan, (2012) used the ARIMA technique to anticipate sugarcane production in India from 2013 to 2017 using data from the previous 62 years. The understanding that knowledge of trustworthy and precise

projections of infant mortality rate is both necessary and vital for the planning of relevant intervention programs and preventative measures for the reduction of infant mortality in Nigeria serves as the driving force behind this work. Many works have been done using ARIMA, hence, in this current work, ARMA has been adopted.

2.0 MATERIALS AND METHODS

AREA OF THE STUDY

Nigeria is the focus of this study. Nigeria is divided into six geopolitical zones: North Central, North-East, North-West, South-East, South-South, and South-West. It also has a Federal Capital Territory (FCT).

2.1 METHODOLOGY

DATA SOURCE

The data for this study came from the World Bank and was estimated by the United Nations Inter-Agency Group for Childhood Mortality Estimation (UN-IGME), from the World Development Indicators- World Bank Data – 2021 and the World Bank (web: www.childmortality.org), 1981-2021.

Stationary and Non-stationary Series

A time series is said to be strictly stationary if the joint distribution of $W_{(t_1)}, W_{(t_2)}, \dots, W_{(t_n)}$ is the same as the joint distribution of $W_{(t_1+T)}, W_{(t_2+T)}, \dots, W_{(t_n+T)}$, for all $t_{(1+T)}, \dots, t_{(n+T)}$. Thus, shifting the time position by T periods has no effects on the joint distributions, which depends on the interval between t_1, \dots, t_n . If a time series is not stationary, then it is said to be non-stationary. If a non-stationary series is differenced one or more times it becomes stationary and that series is then said to be homogeneous.

3.0 TIME SERIES MODEL

Autoregressive model AR (p)

Models are based on the idea that the current value of the series Z_t can be explained as a function of \mathcal{P} past values $Z_{(t-1)}, Z_{(t-2)}, \dots, Z_{(t-p)}$, where \mathcal{P} determines the number of steps into the past needed to forecast the current value. A general, \mathcal{P} th - order AR model is given as

$$Z_{(t)} = \beta + \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_p Z_{t-p} + \varepsilon_t \tag{1}$$

where ε_t is the white noise

AR (P) can be form by using the backward shift operator β and re-arranging equation we have

$$Z_{(t)} = 1 - \varphi_1 \mathcal{B} - \varphi_2 (\mathcal{B}^2) \dots \varphi(\mathcal{P}) \mathcal{B}^p$$

Or

$$\varphi(\mathcal{B}) Z_t = \alpha + \varepsilon_t \tag{2}$$

Where $\varphi(\mathcal{B}) = 1 - \varphi_1 \mathcal{B} - \varphi_2 (\mathcal{B}^2) \dots \varphi(\mathcal{P}) \mathcal{B}^p$

The AR (P) time series (Z_t) in equation (10) is casual and stationary if the roots of the associated polynomial

$$a^p - \varphi_1 a^{p-1} - \varphi_2 a^{p-2} - \dots - \varphi(\mathcal{P}) \tag{3}$$

is less than one in absolute value

Moving Average Model MA (q)

The moving average model of order q or a finite q th order MA model is given as

$$Z_t = \alpha + \omega_t - \theta_1 \omega_{t-1} - \dots - \theta_q \omega_t \tag{4}$$

Where ω_t is white noise,

A MA (q) process is always stationary regardless of value of the weights in terms of the backward shift operator, the MA (q) model is

$$Z_t = \alpha + (1 - \theta_1 B - \dots - \theta_q B^q) \omega_t$$

Or

$$\alpha + (1 - \sum_{i=1}^q \theta_i B^i) \omega_t \tag{5}$$

The autocorrelation function (ACF)

The autocorrelation function does a great job of describing the general process needed to build a forecasting model. It measures how

closely a time series of observations are related to one another. The definition of the autocorrelation at any lag K is defined as $COR(Z_{(t)}, Z_{(t-k)})$ and is measured by

$$\mathcal{P}_k = \frac{COV(Z_t, Z_{t-k})}{\sigma_{Z_t} \sigma_{Z_{t-k}}} = \frac{E[(Z_t - \mu_z)(Z_{t-k} - \mu_z)]}{[E(Z_t - \mu_z)^2 (Z_{t-k} - \mu_z)^2]} \tag{6}$$

Where Z_t is the observation at time t ,

Z_{t-k} Is the observation at time $t - k$ and μ is the observed mean.

partial autocorrelation coefficient when the effects of the other time lags remained constant. When we are unsure of the proper order for the autoregressive process to match the time series, we take partial autocorrelation into account. Indicating PACF is denoted by $Q_{k,k}$ and defined by

Partial Autocorrelation Function (PACF)

The strength of the connection between an observation Z_t and Z_{t-k} is measured by a

$$Q_{k,k} = \frac{|\mathcal{P}_k^*|}{|\mathcal{P}_k|} \tag{7}$$

Where \mathcal{P}_k is a $\mathcal{K} \times \mathcal{K}$ auto correlation matrix and \mathcal{P}_k^* is \mathcal{P}_k with the least column replaced by $[\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_k]^T$

RESULTS AND DISCUSSION

The data used for this analysis is the yearly data of infant and maternal mortality rates in Nigeria from 1981 to 2021 for infant

mortality and 2000 to 2021 for maternal mortality. The data is obtained from World Bank and estimated by United Nation Inter-Agency Group for Childhood Mortality

Estimation (UN-IGME). The SPSS software, version 26, was used to analyze the data collected in this study. The forecast was made using time series analysis in particular, the Auto Regressive Moving

Average (ARMA) method, which was developed by Box - Jenkins in 1976. It is possible to obtain the forecast confidence intervals and the distribution of prediction errors.

Table 1: Infant mortality data 1981-2021

S/N	YEAR	INFANT MORTALITY	S/N	YEAR	INFANT MORTALITY
1	1981	124.517	21	2001	106.799
2	1982	123.480	22	2002	103.788
3	1983	123.038	23	2003	100.819
4	1984	123.09	24	2004	97.818
5	1985	123.394	25	2005	94.958
6	1986	123.897	26	2006	92.279
7	1987	124.379	27	2007	89.821
8	1988	124.671	28	2008	87.716
9	1989	124.769	29	2009	85.869
10	1990	124.604	30	2010	84.325
11	1991	124.318	31	2011	83.000
12	1992	123.859	32	2012	81.885
13	1993	123.287	33	2013	80.953
14	1994	122.518	34	2014	80.104
15	1995	121.399	35	2015	79.302

16	1996	119.886	36	2016	78.330
17	1997	117.904	37	2017	77.123
18	1998	115.442	38	2018	75.665
19	1999	112.725	39	2019	74.012
20	2000	109.819	40	2020	72.244
			41	2021	70.00

Table 2: Descriptive statistics of infant mortality

statistics	values	statistics	values
No of observation	41	Range	54.77
Sum	4227.81	Standard deviation	19.89021
Mean	103.1172	Skewness	-0.253
Maximum	124.77	Kurtosis	-1.636
Minimum	70.00	SE mean	3.10633

It could be observed that the average infant mortality in Nigeria over 41 year period was approximately 103 per One Hundred Thousand with a standard error of 3.10633, The extent to which the observed data fall within the Centre of the distribution was -1.636 which implies that the distribution was platykurtic with a wide midrange on either side of the mean and a lower peak as compared to the standard normal distribution. The degree of asymmetry (Skewness) of the distribution of

the infant mortality rate was -0.253 which is negatively skewed to the left.

Figure 1 displays a time series graphic of Nigeria's infant mortality rates from 1981 to 2021. The time series plot in Figure 1 appears to be declining at a consistent pace, which suggests that the series exhibits trends and seasonality, which are two traits common to non-stationary time series. The plots of the autocorrelation and partial autocorrelation functions, also known as the correlogram are shown in

Figure 2 and examined to validate the series' plot in Figure 1.
of stationary as shown by the time series

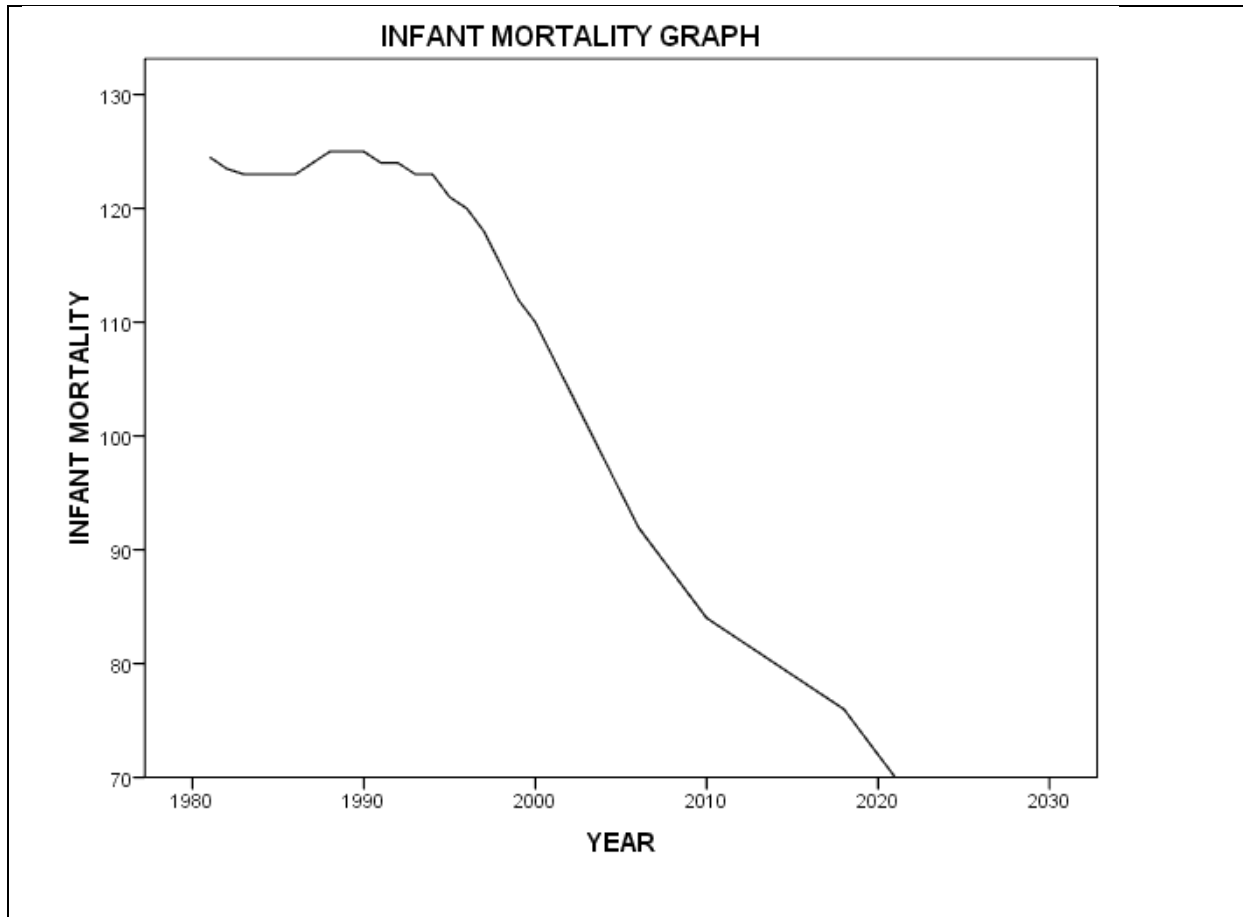


Figure 4.1 Time series plot of infant mortality rate in Nigeria

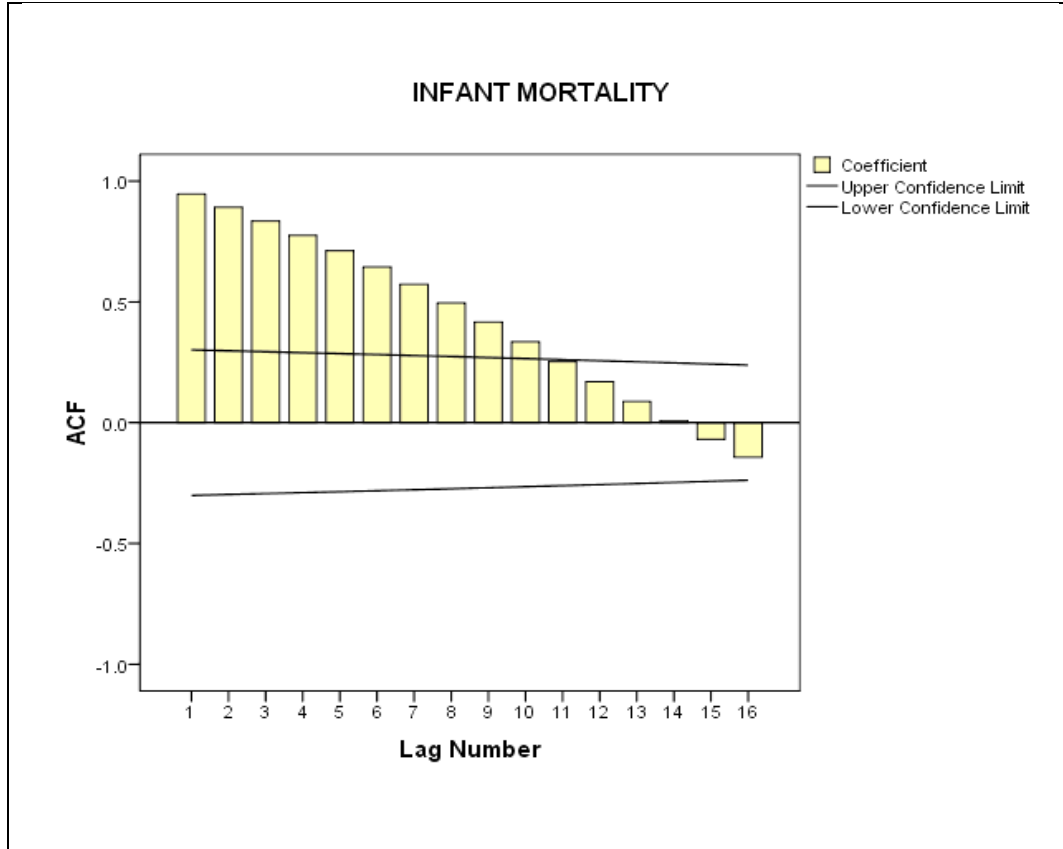


Figure 1a Correlagram of the time series

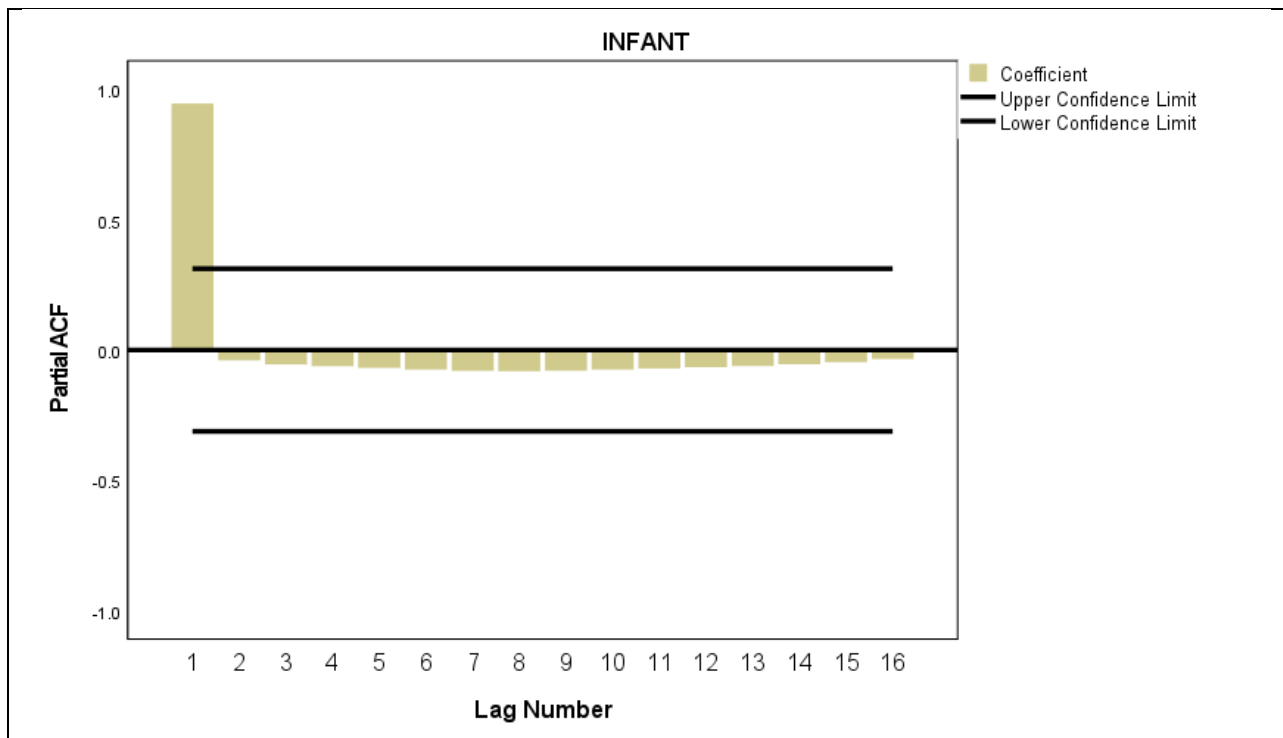


Figure 1b Correlagram of the time series

For the time series plot of the infant mortality rate in Nigeria, Figure 4.2 graphically illustrates the Auto Correlation Function (ACF) 4.2a and the Partial Auto Correlation Function (PACF) 4.2b. The time lags in Figure 4.2 have substantial significant ACFs that gradually shrink in size but do not decay to zero (slow decay). As a result, the ACF exhibits a pattern resembling a non-stationary time series. The partial autocorrelations for time lags 1 through 16 are close to zero in the PACF plot, which is likewise typical of non-stationary series, whereas the partial autocorrelation at time lag 1 is close to one.

By taking the first difference of the time series data and then creating the time plot, the series is afterwards changed into a stationary series.

The differenced series' time plot in Figure 4.3 appears stationary and devoid of trend upon visual inspection, indicating that

the mean and variance now appear to be stable over time.

The unit root test is used to check stationarity, The unit root test suggests that the series is said to have a unit root and is therefore not stationary as it violates the fundamental requirement of zero mean and constant variance required for time series modeling if $\alpha_1 = 1$ in the simple autoregressive scheme $q_t = \alpha_0 + \alpha_1 q_{t-1} + \varepsilon_t$ where α_1 is the coefficient of the autoregression process (Dickey and Fuller, 1979). We use the following approach to examine the time series data for stationarity ADF

$H_0: \alpha_1 = 1$ Series has unit root

$H_a: \alpha_1 \neq 1$ Series has no unit root

We reject the null hypothesis that the data has a unit root if the test statistic of the ADF test is smaller than the crucial value.

Table 3: Augmented dickey-fuller tests

Time series tests for variable	
Infant mortality	
	Values
Test	Augmented Dickey-fuller
Alternative hypothesis	stationary
p-value	0.1
note	p-value smaller than printed p-value
Truncation lag	1
Test	KPSS test for Stationarity
Null hypothesis	level
p-value	0.01
Note	p-value smaller than printed p-value
Truncation lag	3

From the table, we conclude that alternative hypothesis H_a , which implies that the series is stationary, are valid and reject the null hypothesis.

The computation and plotting of the autocorrelation and partial autocorrelation function at various time lags provide additional evidence of the series' stationarity

Table 4: ACF and PACF of infant mortality

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic			Partial Autocorrelation	Std. Error
			Value	Df	Sig. ^b		
1	0.961	0.152	39.798	1	0.000	0.961	0.158
2	0.881	0.150	74.107	2	0.000	-0.564	0.158
3	0.763	0.148	100.517	3	0.000	-0.376	0.158
4	0.620	0.146	118.440	4	0.000	-0.106	0.158
5	0.466	0.144	128.852	5	0.000	0.039	0.158
6	0.310	0.142	133.596	6	0.000	-0.005	0.158
7	0.159	0.140	134.885	7	0.000	-0.063	0.158
8	0.014	0.138	134.894	8	0.000	-0.181	0.158
9	-0.122	0.136	135.696	9	0.000	-0.092	0.158
10	-0.244	0.134	139.019	10	0.000	-0.019	0.158
11	-0.347	0.131	145.986	11	0.000	0.051	0.158
12	-0.425	0.129	156.805	12	0.000	0.088	0.158
13	-0.475	0.127	170.844	13	0.000	0.012	0.158
14	-0.495	0.124	186.653	14	0.000	0.025	0.158
15	-0.488	0.122	202.664	15	0.000	-0.60	0.158
16	-0.457	0.120	217.308	16	0.000	0.005	0.158

a. The underlying process assumed is independent (white noise).

b. Based on the asymptotic chi-square approximation.

The ACF and PACF plots of the differenced series of infant mortality in Table 4.3 from

lags 1 to 16 indicate that we have an AR process.

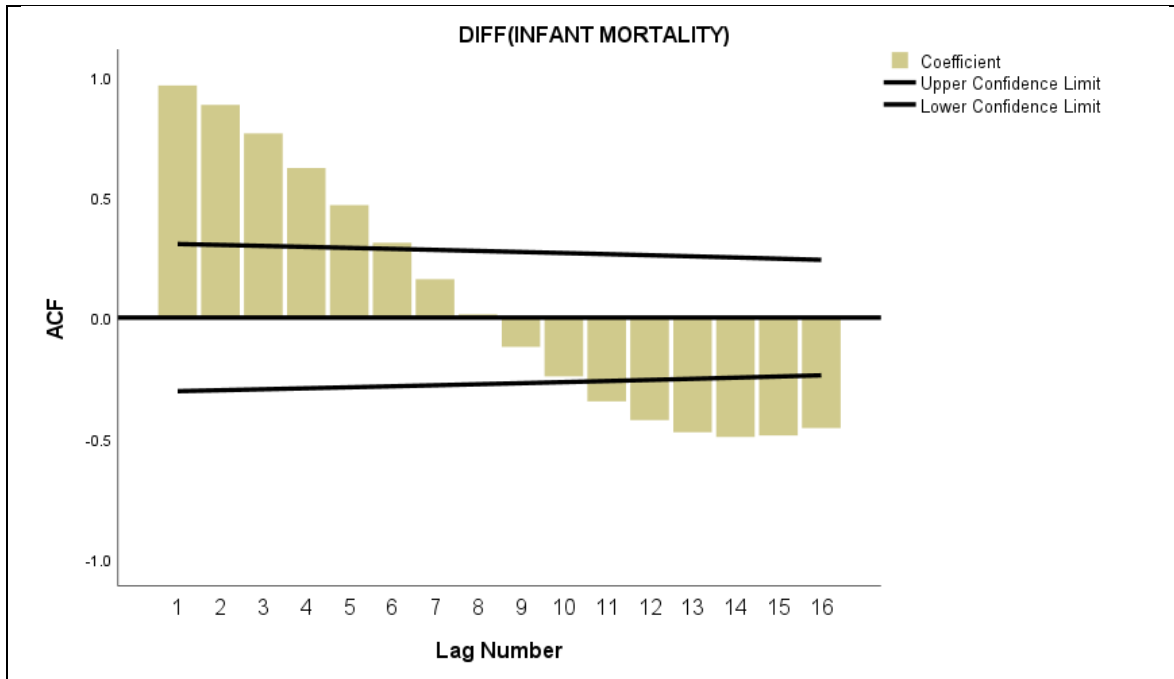


Figure 3a ACF of the first differenced series

The autocorrelation function of the first order differenced time series of infant mortality rate in Nigeria is shown in Figure 4.4a from lags 1 to 16. As the time lag k

rises, Figure 4.4a demonstrates that the ACF steadily decays to zero, demonstrating a geometric decay representative of an AR process.

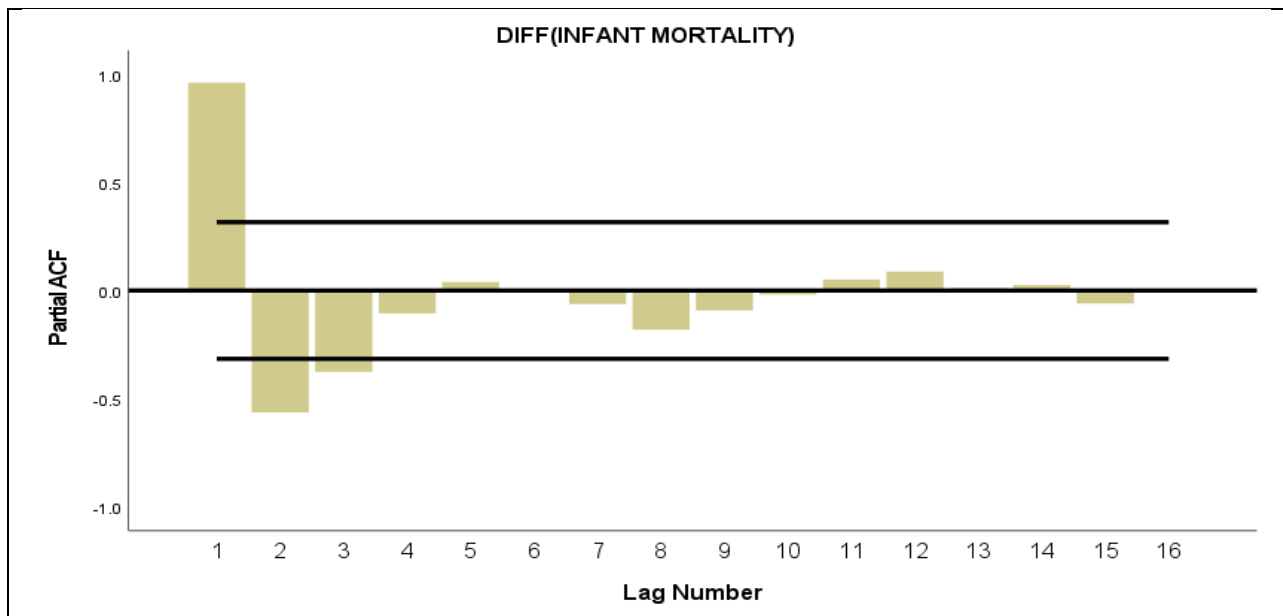


Figure 3b PACF of the first differenced series

The differenced series' autocorrelation and partial autocorrelation functions showed no need for further differencing because they typically tail off quickly. Since they don't repeat at lags that are multiples of the number of periods every season, they also showed no evidence of seasonality.

Once more, the ACF plot of the series in Figure 4.3 was used to check for

the existence of white noise. The series is said to be characterized by white noise if all of the sample ACFs are near to zero. The series of yearly IMR was consequently determined to be characterized by white noise. The ARMA (0, 1), ARMA (1, 0), and ARMA models are recommended by the ACF and PACF as suitable for estimating the infant mortality rate in Nigeria.

Table 5: ARMA model parameters

	Estimate	S.E	T-ratio	Sig.
Constant	96.405	111.958	0.861	0.395
AR lag1	0.998	0.016	60.932	<0.001
MA lag1	-0.973	0.534	-1.823	0.76

The SPSS Modeler was used to fit the model to the data. The coefficient of both the AR and the MA were not

significantly different from zero with values of 0.998 and -0.973 respectively. This model enables us to write the model equation as

$$Z_{(t)} = 0.998Z_{t-1} - 0.973\omega_{t-1} + \varepsilon_t$$

That is, the AR coefficient φ was estimated to be 0.998 with standard error of 0.016 and a t-ratio of 60.932 while the MA

coefficient θ was estimated to be -0.973, with standard error of 0.534 and a t-ratio of -1.823.

Table 6: Model statistics

Model Fit Statistics				Ljung-Box Q (18)			
R-squared	RMSE	MAPE	BIC	Statistics	D.F	SIG	Number of outliers
0.948	4.656	1.341	3.348	0.732	16	1.000	0

In this model $Q = 0.732$, the 10% and 5% points of chi-square with 16 degree of freedom are 32.00 and 26.30 respectively. Therefore, the model is extremely acceptable and significantly appropriate because Q is not excessively huge and the evidence does not refute the hypothesis of White Noise behavior in the residuals. By

analyzing the Normalized Bayesian Information Criterion, the model's sufficiency and considerable appropriateness were established (BIC). The ARMA (1, 1) model, among a class of statistically significant ARMA (p, q) models fitted to the series, has the lowest BIC value of 3.348

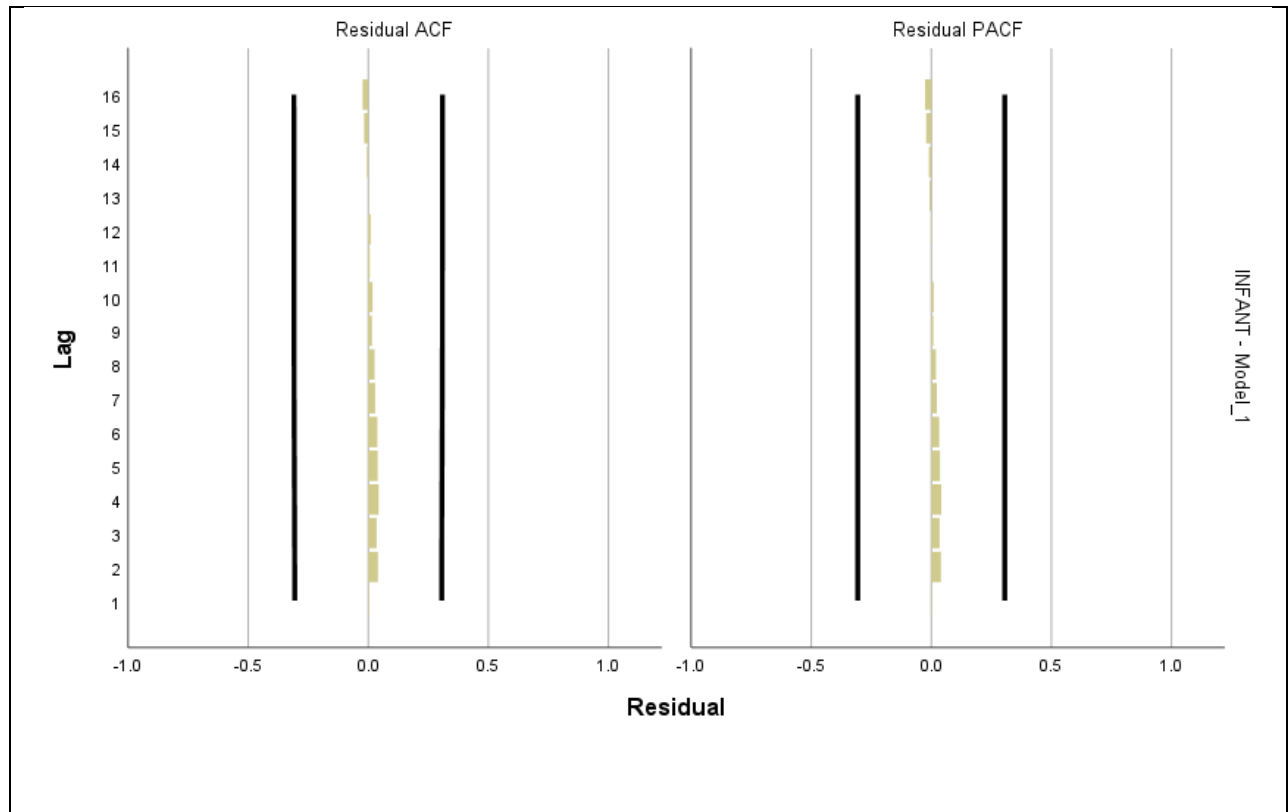


Figure 4 Residual plots of ACF and PACF plots

Before using the model for forecasting, the residual ACF and PACF plots shown in Figure 4 are used to assess the model's suitability as a good fit for the time series data. If all that remains after fitting the model are white noise residuals, the model is appropriate. The residuals in this situation

need to have constant variance and be uncorrelated. The ACF and PACF's diagnostic of the residuals reveal that all of their values fall within the 5% zero bound, showing that there is no correlation between the residuals.

FORECAST OF INFANT MORTALITY

Table 7: Forecast

Year	Model			
	Infant Mortality Rate-ARMA 1,1			
	Forecast	Upper confidence limit	Lower confidence limit	
2022	69	81	59	
2023	68	81	56	
2024	67	82	52	
2025	65	83	48	
2026	64	83	44	
2027	62	84	40	

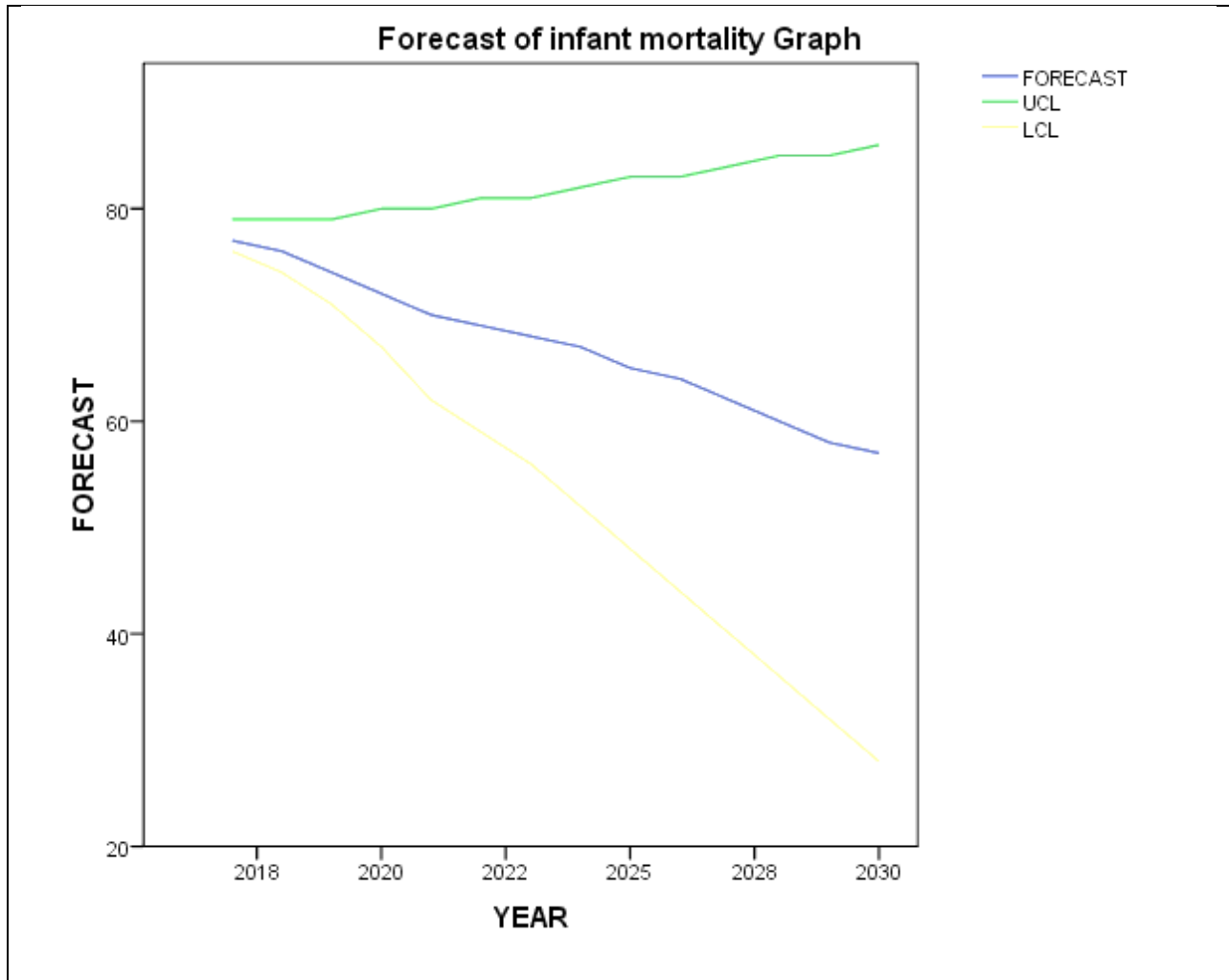


Figure 5 Plot of the observed series

Table 8 Distribution of the forecast error of infant mortality rate in Nigeria

Year	Observed Value	Forecast	Forecast Error
2017	77.12	76.8	0.41
2018	75.67	75.5	0.22
2019	74.01	73.8	0.28
2020	72.24	72.2	0.06
2021	70.00	69.5	0.71

For the five years between 2017 and 2021, Nigeria's infant mortality rate projection error is shown in Table 4.11; the prediction for the infant mortality rate in Nigeria has a Mean Absolute Percentage Error (MAPE) of 0.336, which indicates 99% forecast

accuracy. We derive the ACF and PACF of forecast error (residual) and the QQ plot of forecast error in order to further examine if the mean is normally distributed and to check for non-zero autocorrelation in the forecast residuals.

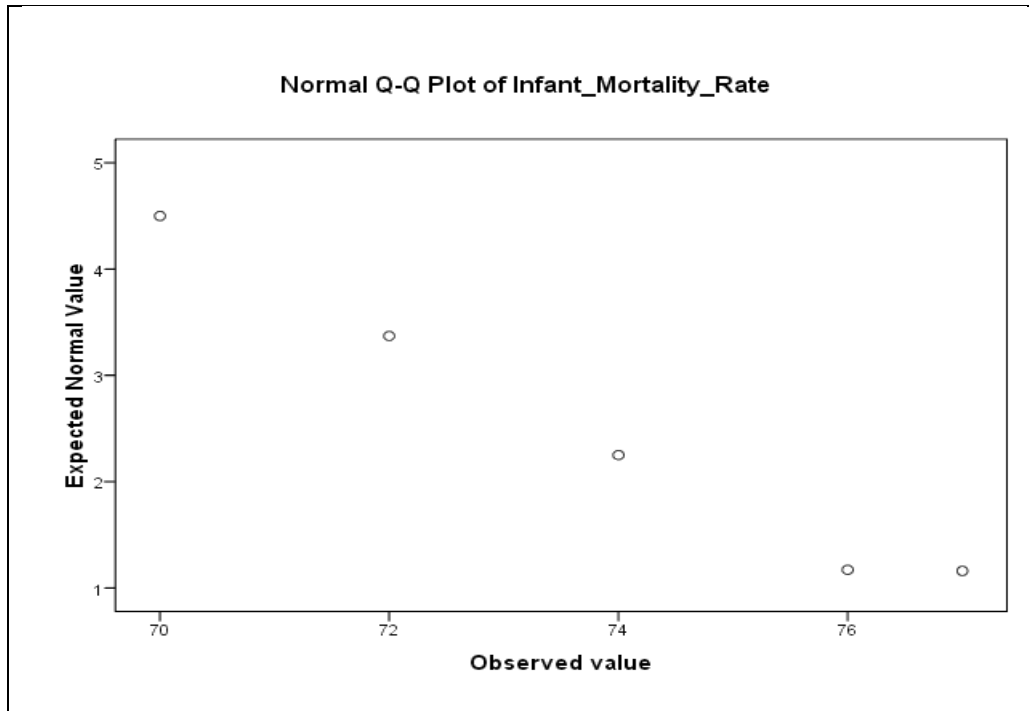


Table 5 shows the forecast of infant mortality rate in Nigeria for the years 2017 through 2030 with 95% upper and lower confidence intervals

CONCLUSION

Using the SPSS 26 Expert modeler, the sample ACF and PACF of the original series of the infant mortality were computed and their graphs were shown. These were taken into consideration when choosing the right

model. Due to the sample ACF'S failure to die out quickly even at high delays, the series displayed non-stationary behavior. Once differencing was applied to the series, stationarity was attained. The differenced series' plot revealed that the data are uniformly distributed around the mean. The Ljung-Box test statistic and the Normalized Bayesian Information Criterion were then used to perform statistical diagnostic testing on the model (BIC). The model is

statistically significant, suitable, and adequate, according to analysis

References

- Abraha, M. W., & Nigatu, T. H. (2009). Modeling trends of health and health related indicators in Ethiopia (1995-2008): a time-series study. *Health Research Policy and Systems*, 7(1), 1-17
- Adetoro, G. W., & Amoo, E. O. (2014). A statistical analysis of child mortality: Evidence from Nigeria. *Journal of Demography and Social Statistics*, 1(1), 110-120.
- David, J. (2018). Infant mortality and public health expenditure in Nigeria: Empirical explanation of the nexus. *Timisoara Journal of Economics and Business (TJE&B)*, 11(2), 149-164.
- Kumar, M. A., & Karthikeyan, S. (2012). Investigating the efficiency of Blowfish and Rejindael (AES) algorithms. *International Journal of Computer Network and Information Security*, 4(2), 22.
- Nyongesa, C., Xu, X., Hall, J. J., Macharia, W. M., Yego, F., & Hall, B. (2018). Factors influencing choice of skilled birth attendance at ANC: evidence from the Kenya demographic health survey. *BMC pregnancy and childbirth*, 18(1), 1-6.
- Sachs, J. D., & McArthur, J. W. (2005). The millennium project: a plan for meeting the millennium development goals. *The Lancet*, 365(9456), 347-353.
- Ude, D. K., & Ekesiobi, C. S. (2014). Effect of per capita health spending on child mortality in Nigeria. *International Journal of Innovative Research and Development*, 3(9), 1-4.

Usman, A., Sulaiman, M. A., & Abubakar, I. (2019). Trend of neonatal mortality in Nigeria from 1990 to 2017 using time series analysis. *Journal of Applied Sciences and Environmental Management*, 23(5), 865-869

Yaquub, A., & Gul, S. (2013). Reasons for failure of exclusive breastfeeding in children less than six months of age. *Journal of Ayub Medical College Abbottabad*, 25(1-2), 165-167.