Transactions of the Nigerian Association of Mathematical Physics
***Volume 11, (January - June, 2020 Issue*), pp189 – 194**
© Trans. of NAMP

AN APPLICATION OF TIME SERIES AUTO REGRESSIVE INTEGRATED MOVING AVERAGE FORECASTING MODEL FOR PREDICTING NEONATAL MORTALITY RATE IN NIGERIA

Friday Ewere and Donalben Onome Eke

Department of Statistics, Faculty of Physical Sciences, University of Benin, Benin City, Edo State, Nigeria.

Abstract

The number of preventable child deaths that occurs during the first month of life in Nigeria is high in comparison to the global standard set by the United Nations in the Millennium Development Goals (MDG) and subsequently the Sustainable Development Goals (SDG) for childhood mortality. Following the failure of Nigeria to meet the recommended MDGs target of reducing childhood mortality by two-third in 2015, there exist genuine concerns as to whether the SDGs target of 12 deaths per 1000 live births for neonates can be achieved.

In this paper, the Auto Regressive Integrated Moving Average (ARIMA) Model for time series analysis was used to make forecast of neonatal mortality up to 2030 employing data obtained from the United Nation's Inter Agency Group for Childhood Mortality Estimate (UN-IGME).

The ARIMA (2, 1, 0) model predicted a reduction of up to 19.4% by 2030 at 95% confidence interval. Results from this study also showed that Nigeria will not be able to achieve the SDG goal for child mortality unless there is a drastic reduction of over 200% in neonatal mortality.

***Keywords*:** Time Series, ARIMA, Forecasting, Neonatal Mortality, Nigeria

1. INTRODUCTION

Neonatal mortality is the death of a child within the first 28 days of life [1]. The prevalence of preventable child death within the neonatal period remains a global public health concern as it is the most vulnerable time for a child's survival. In 2018, 46% of all deaths in children less than five years of age occurred during the first month of life. This translates to as high as 2.6 million deaths implying that on the average, 7,000 neonates die daily [2]. Although, there has been a 51% reduction in neonatal mortality rate globally, sub Saharan Africa and South Asia still had alarmingly high rates of mortality in neonates [3].

In Nigeria about half (51%) of all deaths among children under five years of age takes place before the child's first birthday with 30% occurring during the first month of life [4]. Currently, the proportion of child death that occurs in the neonatal period is estimated to be about 39 in every 1000 live birth and is comparatively high [5, 6]. With an estimated population of about 200 million people which is about 3.5% of the world's population, Nigeria accounts for one of the highest rates of neonatal mortality and was among the countries that failed to attain the United Nation's Millennium Development Goal for the reduction of childhood mortality by two third in 2015. Based on the current rates of neonatal mortality which stands at 39 deaths per 1000 live births [4] and owing to the fact that Nigeria could not meet up with MDG 4, genuine fears that Nigeria may fall short of the Sustainable Development Goals 3 for neonatal mortality are not out of place. Nigeria's chances of attaining the Sustainable Development Goal for survival of neonates born within its geographical constraint cannot be met without substantial reduction in the neonatal mortality rate and without a measure of her progress using comprehensive and inclusive assessment.

Hence, the aim of this paper is to provide forecast of neonatal mortality in Nigeria up to the year 2030 with a view to ascertaining if Nigeria will be able to achieve the Sustainable Development Goal for neonatal mortality by applying the ARIMA technique in time series analysis.

The ARIMA model has been identified as a flexible predictive model that has been successfully used in a wide range of applications for forecasting the future successive values in time series analysis [7].

This study will be beneficial to women's health as well as infant and child survival in Nigeria as it will help shape future preventive and intervention programmes and policies meant for the reduction of mortality rates of neonates in the country.

Correspondence Author: Friday E., Email: Friday.ewere@uniben.edu, Tel: +2348066785972

2. METHODOLOGY

2.1 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.

2.2 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)

2.3 Data Analysis

Data analysis was conducted using SPSS software version 23. The stages of model identification, model estimation and diagnostic checking in an ARIMA model in time series analysis as proposed by [8] were employed for forecasting.

3. RESULTS AND DISCUSSIONS

Table 1: Estimates of neonatal mortality rates in Nigeria between 1960 and 2018

S/N	Year	Neonatal Mortality	S/N	Year	Neonatal Mortality	S/N	Year	Neonatal Mortality
1	1960	78.9	21	1980	53.7	41	2000	47.5
2	1961	77.1	22	1981	52.6	42	2001	46.5
3	1962	75.5	23	1982	51.7	43	2002	45.3
4	1963	74.7	24	1983	51	44	2003	44.1
5	1964	73.2	25	1984	50.4	45	2004	42.8
6	1965	71.9	26	1985	50	46	2005	41.8
7	1966	71	27	1986	49.8	47	2006	40.8
8	1967	70.7	28	1987	49.8	48	2007	40
9	1968	69.6	29	1988	49.8	49	2008	39.3
10	1969	68.3	30	1989	49.9	50	2009	38.8
11	1970	67.1	31	1990	50	51	2010	38.3
12	1971	65.8	32	1991	50.3	52	2011	37.9
13	1972	64.5	33	1992	50.6	53	2012	37.6
14	1973	63.1	34	1993	50.7	54	2013	37.3
15	1974	61.7	35	1994	50.8	55	2014	37
16	1975	60.3	36	1995	50.7	56	2015	36.8
17	1976	59	37	1996	50.5	57	2016	36.6
18	1977	57.5	38	1997	50.1	58	2017	36.4
19	1978	56.2	39	1998	49.4	59	2018	36
20	1979	54.9	40	1999	48.5			
Source: world bank, www.childmortality.org								

Table 1 gives an estimate of the mortality rate for neonates per 1000 live births in Nigeria between 1960 and 2018. To check if the time series is stationary, a plot of the series is presented in Figure 1.

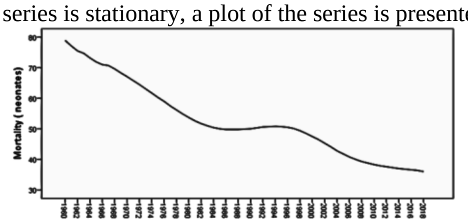

Figure 1: A time series plot of neonatal mortality in Nigeria from 1960 to 2018

A visual inspection of the time series plot of neonatal mortality in Figure 1 shows clearly that the data is not stationary. The plot reveals that there is a decreasing trend in the time series. In order to stabilize the variance and make the time series stationary, the series is transformed by $W_t = \Delta y_t - y_{t-1}$. That is, the first order difference of the time series data was computed. After taking the first difference, the differenced series has n-1 observations. Figure 2 provides the plot of the resulting differenced time series.

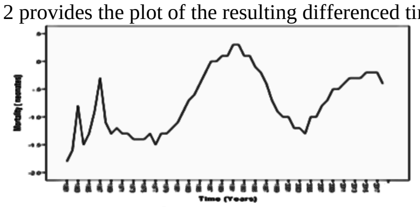

Figure 2 Time series plot of the first difference (d = 1)

A look at the time series plot of the differenced series in Figure 2 shows that the time series appears to be stationary and devoid of trend when compared to the Figure 1. To validate this, the unit root test otherwise known as the Augmented Dickey-Fuller test was applied [9]. The unit root test for stationarity 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. This is because it violates the basic requirement of zero mean and constant variance necessary for time series modeling. To test for stationarity in the time series data, the following hypotheses are adopted:

H_0**:** the series contains a unit root (it is non-stationary)

H_1**:** there is no unit root in the series (it is stationary)

The null hypothesis is rejected and the alternative hypothesis is true if the p-value is greater than the significance level: $\alpha = 0.01$

Table 2: Augmented Dickey-Fuller Test

Dickey-Fuller	Sig.	ACF coeff. of e (d=1)	DF
-2.391	0.090	0.334	16

Since the computed p-value = 0.09 is greater than $\alpha = 0.01$ the null hypothesis is rejected and the alternative hypothesis H_1 is true implying that the series is stationary after taking the first difference (d = 1).

The Auto Correlation Function (ACF) and the Partial Auto Correlation Function (PACF) for the neonatal mortality time series data is examined to confirm stationarity and to determine the order of the AR and MA components of the ARIMA model. The ACF and PACF which can give insight into what model to fit to the data are presented in Figures 3 and 4 respectively.

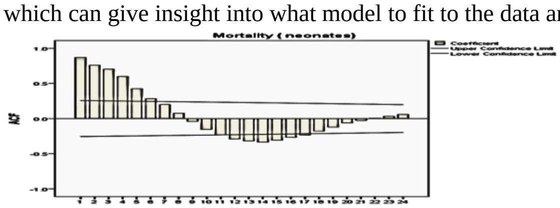

Figure 3 ACF for Neonatal mortality

Figure 3 above displays the ACF from lags 1 to 24 of the first order differenced time series of neonatal mortality rate in Nigeria. Figure 3 shows that the ACF reduces to zero slowly as the number of time lags k increases indicating a geometric decay symbolic of an AR process.

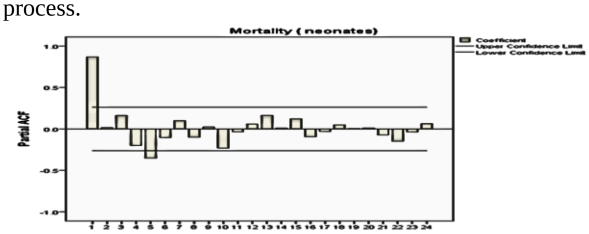

Figure 4 PACF for Neonatal mortality

The PACF in Figure 4 indicates that there are only two significant time lags. Only time lags 1 and 5 exceeds the significant threshold which is the error bound indicating an AR (2) model. All other time lags fall within the significant threshold.

Having identified the theoretical properties of the ACF and PACF as an AR (2) process given that the ACF of the first difference of the series dies down to zero slowly and the dying down pattern produced by the AR component dominate the cutting off pattern produced by the MA component, the inference is that the ARIMA model is an ARIMA (2, 1, 0).Below are other tentative models identified based on the principle of parsimony and recognizable pattern of the ACF and PACF. The best approximation to the data by the member of ARIMA family is obtained by comparing the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) values.

Table 3: Possible Fitted ARIMA Model for Neonatal Mortality

Models	AIC	BIC
ARIMA (1, 1, 0)	23.34	24.77
ARIMA (1, 1, 1)	23.96	23.91
ARIMA (2, 1, 1)	26.73	23.10
ARIMA (2, 1, 0)	23.45	23.91

A comparison of the AIC and the BIC values in Table 3 for the possible models shows that ARIMA (1, 1, 0) has the least AIC value. However, the ARIMA (2, 1, 0) produced least values for both the AIC and the BIC and is therefore selected as the best fit model for the data.

Table 4: Estimated Summary for the ARIMA (2, 1, 0) model

ARIMA Model Parameters

			Estimate	SE	t	Sig.
Neonatal Mortality	Constant		-.839	.334	-2.514	.015
	AR	Lag 1	.891	.135	6.581	.000
		Lag 2	.026	.137	.194	.847
	Difference		1			

The summary of the estimates in Table 4 reveals that AR (1) = 0.891 and AR (2) = 0.026 when the series is differenced once.
The ARIMA (2, 1, 0) model for neonatal mortality in Nigeria is thus given as

$$Z_t = \mu + Z_{t-1} + \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \varepsilon_t$$

That is,

$$Z_t = -0.839 + Z_{t-1} + 0.891\, Z_{t-1} + 0.026\, Z_{t-2} + \varepsilon_t \qquad (1)$$

where Z_t is the stationary series observed, μ is the mean of the time series (constant), $\emptyset_i$ are the parameter estimates and ε_t is the white noise which is assumed to be identically and independently distributed with zero mean and constant variance across all observations. The white noise assumption rules out possibilities of serial autocorrelation and heteroscedasticity in the disturbances.
The residual ACF and PACF plots and the Ljung-Box test [10] are used to check the adequacy of the model as a good fit of the time series data before it is employed for forecasting.

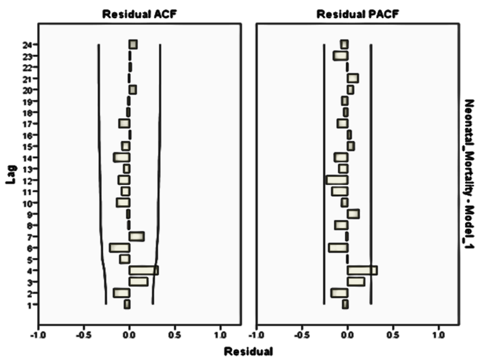

Figure 5 Residual ACF and PACF plots

A look at the plot of the residual ACF in Figure 5 shows that all the values of the autocorrelation coefficients between lags 1 and 24 are within the 5% zero bound implying that all the ACF are within the significant limit.
Similarly, all the PACFs coefficients of the residuals of fitted ARIMA (2, 1, 0) model for lags 1 to 24 are within the significant limits. This means that the ACF and PACF indicates that there is no non-zero autocorrelations in the residuals.
The Ljung-Box test results are given in Table 5 while Figure 6 represents the plot of the standard residual for the fitted model given in eqn (1).

Table 5: Ljung-Box Test Statistics

Model	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
	Stationary R-squared	Statistics	DF	Sig.	
Neonatal Mortality	.768	23.979	16	.090	0

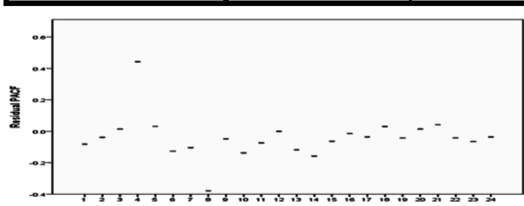

Fig 6: Standard Residual for the fitted ARIMA model
The Coefficient of the Ljung Box statistics and a p-value of 0.09 in Table 5 is an indication that all of the autocorrelation functions in lags 1 to 24 are zero. The conclusion here is that there is no (or almost nil) evidence for non-zero autocorrelations in the residuals at lags 1 to 24 in the fitted model given in eqn (1).

The ARIMA (2, 1, 0) model is now fitted to make forecast for the future values of the time series data. The K period ahead forecast based on the ARIMA (2, 1, 0) model is given by

$$\hat{Z}_{t+k} = \mu + Z_{t+k-1} + \emptyset_1 Z_{t+k-1} + \emptyset_2 Z_{t+k-2} + \varepsilon_t$$

where $\mu = -0.839$, $\emptyset_1 = 0.891$ and $\emptyset_2 = 0.026$

Hence,

$$\hat{Z}_{t+k} = -0.839 + Z_{t+k-1} + 0.891\, Z_{t+k-1} + 0.026\, Z_{t+k-2} + \varepsilon_t \qquad (2)$$

Fitting the ARIMA (2, 1, 0) model to forecast for the future values of mortality in neonates up to year 2030, the following forecast estimates as shown in Table 6 are obtained.

Table 6: Neonatal Mortality Forecast

Forecast

	ARIMA (2,1,0) model		
	Neonatal Mortality		
Years	Forecast	UCL	LCL
2019	36	36	35
2020	35	36	34
2021	35	36	33
2022	34	36	32
2023	34	37	30
2024	33	37	29
2025	32	37	28
2026	32	37	26
2027	31	38	25
2028	30	38	23
2029	30	38	22
2030	29	38	20

Table 6 shows the neonatal mortality forecast for 2019 through to 2030 with 95% upper and lower confidence intervals. From Table 6, it can be seen that neonatal mortality will be reduced to 29 deaths per 1,000 live births by 2030 which will amount to a total reduction of about 19.4% and an average yearly reduction of 1.6%.

The plot of the observed series from 1960 to 2018 and the forecast for 2019 up to year 2030 of neonatal mortality by fitting the ARIMA (2, 1, 0) to the time series data is presented in Figure 7.

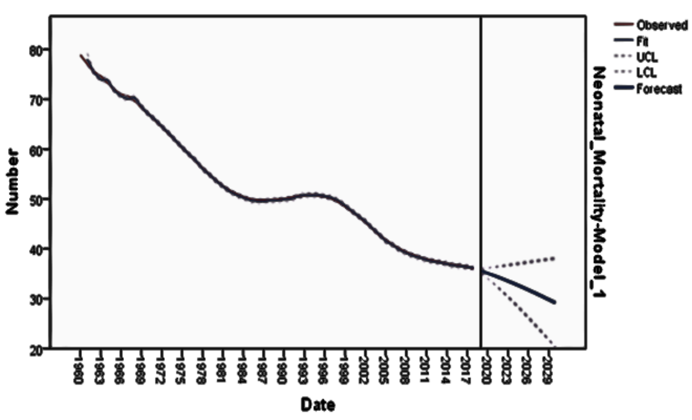

Figure 7: Plot of the observed series and the in-sample forecast value.

The portion of the plot from the year 2019 to 2030 shows the in - sample forecast values by the ARIMA (2,1,0) model produced from the data with the 95% Confidence Interval for each of the forecast value, depicted by the two opening lines.

With the forecast for neonatal mortality as high as 29 deaths per 1000 live births by 2030 as opposed to the proposed target of 12 per 1,000 live births set by SDG, Nigeria is clearly not in line to meeting the target unless there is some drastic reduction in the rate of neonatal deaths. To meet the set target for neonatal mortality by 2030, the country would have to experience a reduction of over 200%.

4. CONCLUSION

In this study, we have applied the ARIMA (2, 1, 0) model to make forecast for neonatal mortality rate in Nigeria up to 2030 using data obtained from the United Nation's Inter Agency Group for Childhood Mortality Estimate (UN-IGME). The time series plot of the forecast from the ARIMA (2, 1, 0) model showed a downward movement suggesting that there would be a continuous decrease in neonatal mortality rate. The ARIMA (2, 1, 0) model predicted a reduction of up to 19.4% by 2030 at 95% confidence interval. Despite this reduction, Nigeria is set to fall short of the Sustainable Development Goal for childhood mortality by 2030. To be able to significantly combat childhood mortality at the neonatal level in Nigeria and attain the SDG target of 12 deaths per 1,000 live births each year, it is recommended that:

1. the government improve existing health facilities and possibly create new ones to help increase the ratio of skilled birth attendants to expectant mothers which would in turn help to reduce the risk of death at the neonatal level.
2. concerted efforts be made by the government to increase access to health facility in rural areas and also keep such hospitals, clinic and maternity homes fully equipped and the skilled staffs well motivated.
3. expectant mothers visit the hospital during pregnancy for quality antenatal care.
4. the government religiously implement the Integrated Maternal, Newborn and Child Health (IMNCH) strategies.

The provision of Neonatal Intensive Care Unit (NICU) services and increase in training on Early Neonatal Care (ENC) as well as Neonatal resuscitation in all hospitals as a measure of reducing neonatal mortality in Nigeria had also been recommended [11].

AKNOWLEDGEMENT

The authors would like to thank the World Bank Group for granting access to the United Nation Inter-Agency Group for Childhood Mortality Estimation (UN-IGME) dataset that was used for the analysis.

REFERENCES

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

[2] World Health Statistics, 2018: Monitoring health for the SDGs sustainable development goals. Geneva: World Health Organization; 2018.

[3] UN-IGME (2019). Levels and Trends in Child Mortality. Report 2019. Estimates developed by the UN inter-agency Group for Child Mortality Estimation. United Nations Children's Fund. New York 2019.

[4] National Population Commission (NPC) [Nigeria] and ICF.2019. 2018 Nigerian DHS Key Findings. Abuja, Nigeria and Rockville, Maryland, USA: NPC and ICF.

[5] 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.

[6] Lancet, 2018: National, regional and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. Volume 7, Issue 6, PE710 – E720, June 01, 2019. DOI: https://doi.org/10.1016/S2214-109X(19)30163-9

[7] 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, Vol 9 (1), 81-94.

[8] Box, G.E.P. and Jenkins, G.M. (1976), Time Series Analysis, Forecasting and Control, Holden-Day, Oakland, CA.

[9] Dickey, D. A., and Fuller, W. A (1979). "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." Journal of the American Statistical Association 77 (366): 427–31.

[10] Ljung, G.M. and Box, G.E.P. (1978), On a measure of lack of fit in time series models, Biometrika, Vol. 67, pp. 297-303

[11] Usman, A; Sulaiman, M. A, Abubakar, I (2019), Trends of Neonatal Mortality in Nigeria from 1990 to 2017 using Time Series Analysis. J. Appl. Sci. Environ. Manage. Vol. 23 (5) 865-869 May 2019 **https://www.ajol.info/index.php/jasem**.