International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 321-328, August-2023

https://doi.org/10.47001/IRJIET/2023.708047

Making Use of ARIMA Model Forecasts to Reconfigure Current Neonatal Health Policies to Address High Neonatal Mortality Rates in Kenya

[1]Dr. Smartson. P. NYONI, [2]Thabani NYONI

[1]ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe

[2]Independent Researcher & Health Economist, Harare, Zimbabwe

***Abstract* - Aggressive maternal and neonatal healthcare policies and interventions are required in scaling up efforts to address neonatal mortality in low-middle income countries. Novel approaches should take into account existing challenges and various needs of different populations. In addition, it is important to respect and recognize local traditions or cultures. Achieving set targets for health-related SDGs will accelerate progress towards achieving substantial reduction of maternal, under-five and neonatal mortality. This study utilizes annual time series data on neonatal mortality rate (NMR) for Kenya from 1960 to 2019 to predict future trends of NMR over the period 2020 to 2030. Unit root tests have shown that the series under consideration is an I (1) variable. The optimal model based on AIC is the ARIMA (2,1,2) model. The findings of this research revealed that neonatal mortality will decline from approximately 21 in 2020 to around 17 deaths per 1000 live births by the end of 2030. Therefore, we implore Kenyan policy makers to design appropriate neonatal policies to address major drivers of mortality in neonates and such strategies should improve the quality of health care services during ANC, delivery and postnatal periods.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

Neonatal mortality refers to the death of a newborn within the first 28 days of life. It remains a global health challenge particularly in developing countries. Several factors have been shown to cause neonatal deaths namely neonatal sepsis, birth asphyxia, prematurity, respiratory distress syndrome and congenital malformations. Preterm birth has been found to be the leading cause of mortality in neonates (Koffi *et al*. 2016; UN, 2015;Brankovic*et al.* 2013; Lawn *et al*. 2006). Furthermore, previous studies have also proved that neonatal mortality increases with decreasing delivery gestational age or birth weight (Ludvigsson *et al*. 2018; Katz *et al*. 2013). The 2014 Kenyan demographic health survey revealed that five years prior to the survey, under-five mortality rate dropped from 74/1000 to 52/1,000 live births and infant mortality fell from 52/1,000 to 39/1,000 live births however, neonatal mortality rates declined from 31/1,000 to only 22/1,000 live births (Kenya DHS 2014). Despite the progressive decline in under five and infant mortality, there has been slow progress in reducing neonatal mortality (Olack *et al.* 2021).The majority of factors which cause neonatal mortality can be prevented by improving the quality of maternal and child healthcare services. This includes quality ANC, safe delivery and essential newborn care. The 3rd sustainable development goaltarget 3.2 aims to reduce neonatal mortality rate (NMR) to at least 12 per 1000 live births in every country. Therefore developing countries are encouraged to channel more resources to Maternal and Child health (MNCH) programs so that they achieve the set SDG-3 targets by 2030. In this study we propose the ARIMA model to project neonatal mortality rate for Kenya since it is useful in modelling linear data. The findings of this paper are envisioned to guide policy, planning and resource allocation to MNCH programs.

II. LITERATURE REVIEW

Olack *et al.* (2021) investigated the causes of neonatal LBW and preterm mortality in Migori County, among participants of the PTBI-K (Preterm Birth Initiative-Kenya) study. Verbal and social autopsy (VASA) interviews were conducted with caregivers of deceased LBW and preterm neonates delivered within selected 17 health facilities in Migori County, Kenya. The probable cause of death was assigned using the WHO International Classification of Diseases (ICD-10). The findings revealed that deaths among LBW and preterm neonates occur early in life due to preventable causes. A similar study was conducted in Kenya by Irimu *et al*.(2021). The researchers collected routine patients' data from structured paper record forms for all admissions to newborn units (NBUs) from 16 purposively selected Kenyan public hospitals that are part of a clinical information network and

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 321-328, August-2023
https://doi.org/10.47001/IRJIET/2023.708047

this data wasanalyzed together with data from all pediatric admissions ages 0–13 years from 14 of these hospitals. The study concluded that majority of newborns died of preventable causes (>95%). Masaba & Phetoe (2020) described the trends of neonatal mortality within the two sub-Saharan countries. The authors found out thatin 2018, the neonatal mortality rate for Kenya was 19.6 deaths per 1000 live births. The neonatal mortality rate had fallen gradually from 35.4 deaths per 1000 live births in 1975. On the other hand, South Africa had its neonatal mortality rate fall from 27.9 deaths per 1000 live births in 1975 to 10.7 deaths per 1000 live births in 2018. Gayawan *et al*.(2016)examined the residual geographical variations in infant and child mortality and how the different categories of the risk factors account for the spatial inequality in West African countries. Researchers pooled data for 10 of the countries extracted from Demographic and Health Surveys and used the spatial extension of discrete-time survival model to examine how the variables exert influence on infant and child mortality across space. Inference was Bayesian based on the computational efficient MCMC technique. They found different geographical patterns for infant and child mortality. In the case of children under five, demographic factors inherent to the mother and child as well as maternal status variables when explained away a good part of the huge variations observed in the crude rates.

III. METHODOLOGY

The Autoregressive (AR) Model

A process K_t (annual NMR at time t) is an autoregressive process of order p, that is, AR (p) if it is a weighted sum of the past p values plus a random shock (Z_t) such that:

$$K_t = \emptyset_1 K_{t-1} + \emptyset_2 K_{t-2} + \emptyset_3 K_{t-3} + \cdots + \emptyset_p K_{t-p} + Z_t \quad [1]$$

Using the backward shift operator, B, such that $BK_t = K_{t-1}$, the AR (p) model can be expressed as in equation [2] below:

$$Z_t = \emptyset(B) K_t \quad [2]$$

where $\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3 - \cdots - \emptyset_p B^p$

The 1st order AR (p) process, AR (1) may be expressed as shown below:

$$K_t = \emptyset K_{t-1} + Z_t \quad [3]$$

Given $\emptyset = 1$, then equation [3] becomes a random walk model. When $|\emptyset| > 1$, then the series is reffered to as explosive, and thus non-stationary. Generally, most time series are explosive. In the case where $|\emptyset| < 1$, the series is said to be stationary and therefore its ACF (autocorrelation function) decreases exponentially.

The Moving Average (MA) Model

A process is reffered to as a moving average process of order q, MA (q) if it is a weighted sum of the last random shocks, that is:

$$K_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \quad [4]$$

Using the backward shift operator, B, equation [4] can be expressed as follows:

$$K_t = \theta(B) Z_t \quad [5]$$

where $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

Equation [4] can also be expressed as follows:

$$K_t - \sum_{j \leq 1} \pi_j K_{t-j} = Z_t \quad [6]$$

for some constant π_j such that:

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 321-328, August-2023
https://doi.org/10.47001/IRJIET/2023.708047

$$\sum_{j\leq 1} |\pi_j| < \infty$$

This implies that it is possible to invert the function taking the Z_t sequence to the K_t sequence and recover Z_t from present and past values of K_t by a convergent sum.

The Autoregressive Moving Average (ARMA) Model

While the above models are good, a more parsimonious model is the ARMA model. The AR, MA and ARMA models are applied on stationary time series only. The ARMA model is just a mixture of AR (p) and MA (q) terms, hence the name ARMA (p, q). This can be expressed as follows:

$$\emptyset(B)K_t = \theta(B)Z_t \ldots \ldots \ldots [7]$$

Thus:

$$K_t(1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p) = Z_t(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) \ldots \ldots [8]$$

where $\emptyset(B)$ and $\theta(B)$ are polynomials in B of finite order p, q respectively.

The Autoregressive Integrated Moving Average (ARIMA) Model

The AR, MA and ARMA processes are usually not applied empirically because in most cases many time series data are not stationary; hence the need for differencing until stationarity is achieved.

$$\left.\begin{array}{c} \textit{The first difference is given by:} \\ K_t - K_{t-1} = K_t - BK_t \\ \textit{The second difference is given by:} \\ K_t(1-B) - K_{t-1}(1-B) = K_t(1-B) - BK_t(1-B) = K_t(1-B)(1-B) = K_t(1-B)^2 \\ \textit{The third difference is given by:} \\ K_t(1-B)^2 - K_{t-1}(1-B)^2 = K_t(1-B)^2 - BK_t(1-B)^2 = K_t(1-B)^2(1-B) = K_t(1-B)^3 \\ \textit{The } d^{th} \textit{ difference is given by:} \\ K_t(1-B)^d \end{array}\right\} \ldots [9]$$

Given the basic algebraic manipulations above, it can be inferred that when the actual data series is differenced "d" times before fitting an ARMA (p, q) process, then the model for the actual undifferenced series is called an ARIMA (p, d, q) model. Thus equation [7] is now generalized as follows:

$$\emptyset(B)(1 - B)^d K_t = \theta(B)Z_t \ldots \ldots \ldots [10]$$

Therefore, in the case of modeling and forecasting international tourism, equation [10] can be written as follows:

$$\emptyset(B)(1 - B)^d K_t = \theta(B)Z_t \ldots \ldots \ldots [11]$$

The Box – Jenkins Approach

The first step towards model selection is to difference the series in order to achieve stationarity. Once this process is over, the researcher will then examine the correlogram in order to decide on the appropriate orders of the AR and MA components. It is important to highlight the fact that this procedure (of choosing the AR and MA components) is biased towards the use of personal judgement because there are no clear – cut rules on how to decide on the appropriate AR and MA components. Therefore, experience plays a pivotal role in this regard. The next step is the estimation of the tentative model, after which diagnostic testing shall follow. Diagnostic checking is usually done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. If not, there would be need for model re – specification and repetition of the same process; this time from the second stage. The process may go on and on until an appropriate model is identified (Nyoni, 2018). The Box – Jenkins technique was proposed by Box & Jenkins (1970) and is widely used in many forecasting contexts, including health sector. In this paper, hinged on this technique; the researcher will use automatic ARIMA modeling for estimating equation [10].

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 321-328, August-2023
https://doi.org/10.47001/IRJIET/2023.708047

Data Issues

This study is based on annual NMR in Kenya for the period 1960 to 2019. The out-of-sample forecast covers the period 2020 to 2030. All the data employed in this research paper was gathered from the World Bank online database.

Evaluation of ARIMA Models

Criteria Table

Table 2: Criteria Table

Model Selection Criteria Table
Dependent Variable: DLOG(K)
Date: 01/23/22 Time: 09:50
Sample: 1960 2019
Included observations: 59

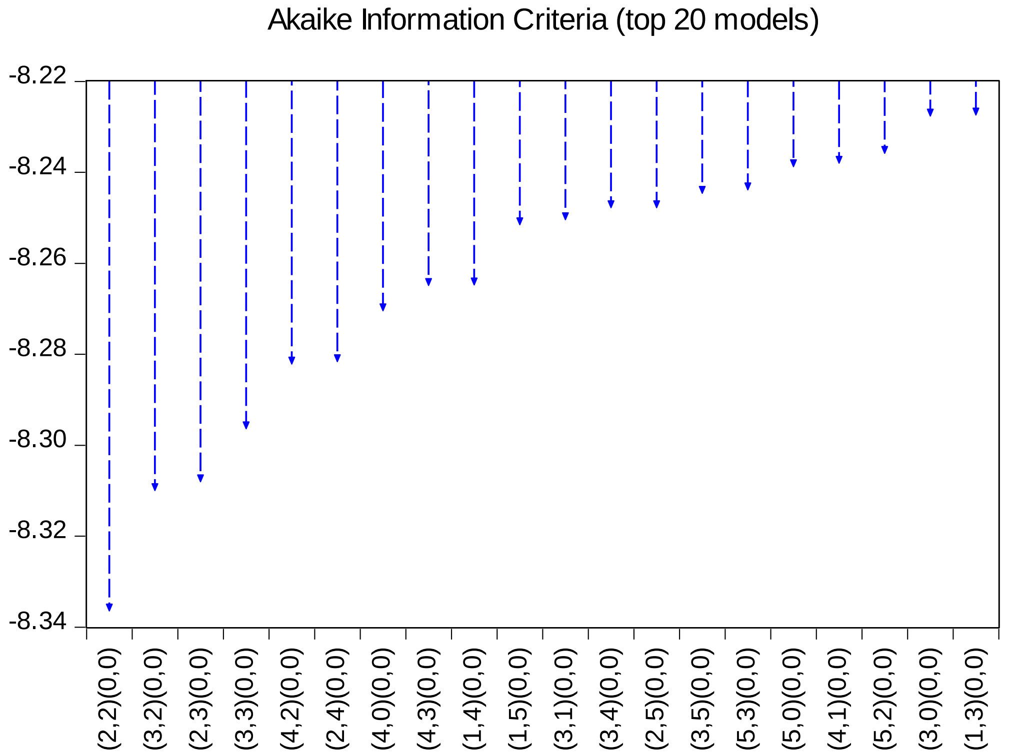
Model	LogL	AIC*	BIC	HQ
(2,2)(0,0)	251.899223	-8.335567	-8.124292	-8.253094
(3,2)(0,0)	252.117816	-8.309078	-8.062591	-8.212860
(2,3)(0,0)	252.061939	-8.307184	-8.060697	-8.210966
(3,3)(0,0)	252.718067	-8.295528	-8.013828	-8.185563
(4,2)(0,0)	252.297831	-8.281282	-7.999582	-8.171318
(2,4)(0,0)	252.283256	-8.280788	-7.999088	-8.170824
(4,0)(0,0)	249.952860	-8.269588	-8.058313	-8.187115
(4,3)(0,0)	252.789166	-8.264040	-7.947127	-8.140330
(1,4)(0,0)	250.785220	-8.263906	-8.017418	-8.167687
(1,5)(0,0)	251.393625	-8.250631	-7.968931	-8.140667
(3,1)(0,0)	249.361049	-8.249527	-8.038252	-8.167054
(3,4)(0,0)	252.283306	-8.246892	-7.929979	-8.123182
(2,5)(0,0)	252.283296	-8.246891	-7.929979	-8.123182
(3,5)(0,0)	253.191341	-8.243774	-7.891649	-8.106319
(5,3)(0,0)	253.167668	-8.242972	-7.890847	-8.105516
(5,0)(0,0)	250.018399	-8.237912	-7.991424	-8.141693
(4,1)(0,0)	249.995860	-8.237148	-7.990660	-8.140929
(5,2)(0,0)	251.931956	-8.234982	-7.918069	-8.111272
(3,0)(0,0)	247.688684	-8.226735	-8.050673	-8.158007
(1,3)(0,0)	248.682580	-8.226528	-8.015253	-8.144055
(4,5)(0,0)	253.316612	-8.214122	-7.826785	-8.062922
(5,4)(0,0)	253.309852	-8.213893	-7.826556	-8.062692
(4,4)(0,0)	252.284092	-8.213020	-7.860895	-8.075565
(5,1)(0,0)	250.023291	-8.204179	-7.922479	-8.094215
(1,2)(0,0)	246.633581	-8.190969	-8.014906	-8.122241
(0,5)(0,0)	248.053512	-8.171305	-7.924818	-8.075087
(5,5)(0,0)	252.997685	-8.169413	-7.746863	-8.004467
(2,1)(0,0)	245.237744	-8.143652	-7.967590	-8.074925
(1,0)(0,0)	242.550222	-8.120347	-8.014709	-8.079110
(2,0)(0,0)	242.604545	-8.088290	-7.947440	-8.033308
(1,1)(0,0)	242.579286	-8.087433	-7.946583	-8.032451
(0,4)(0,0)	240.741612	-7.957343	-7.746068	-7.874870
(0,3)(0,0)	235.264973	-7.805592	-7.629530	-7.736865
(0,2)(0,0)	228.117513	-7.597204	-7.456354	-7.542222
(0,1)(0,0)	214.566996	-7.171763	-7.066125	-7.130526

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 321-328, August-2023
https://doi.org/10.47001/IRJIET/2023.708047

(0,0)(0,0)	191.806646	-6.434124	-6.363699	-6.406633

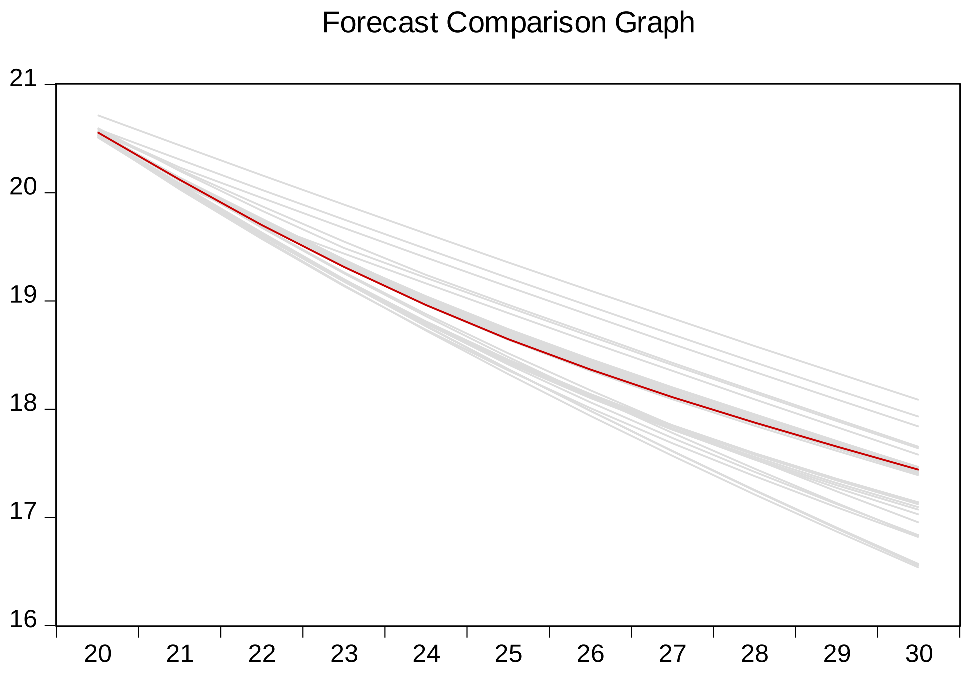
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



ISSN (online): 2581-3048
Volume 7, Issue 8, pp 321-328, August-2023
https://doi.org/10.47001/IRJIET/2023.708047

Table 2 and Figure 1 indicate that the optimal model is the ARIMA (2,1,2) model. Figure 2 is a combined forecast comparison graph showing the out-of-sample forecasts of the top 25 models evaluated based on the AIC criterion. The red line shows the forecast line graph of the optimal model, the ARIMA (2,1,2) model.

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: DLOG(K)
Date: 01/23/22 Time: 09:50
Sample: 1960 2019
Included observations: 59
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (2,2)(0,0)
AIC value: -8.33556688932

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: DLOG(K)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/23/22 Time: 09:50
Sample: 1961 2019
Included observations: 59
Convergence achieved after 11 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.014296	0.002495	-5.730046	0.0000
AR(1)	1.668206	0.118508	14.07678	0.0000
AR(2)	-0.775551	0.120373	-6.442906	0.0000
MA(1)	-0.942143	0.152239	-6.188595	0.0000
MA(2)	0.630631	0.120821	5.219555	0.0000
SIGMASQ	1.08E-05	2.53E-06	4.272857	0.0001

R-squared	0.876852	Mean dependent var	-0.013582
Adjusted R-squared	0.865234	S.D. dependent var	0.009454
S.E. of regression	0.003470	Akaike info criterion	-8.335567
Sum squared resid	0.000638	Schwarz criterion	-8.124292
Log likelihood	251.8992	Hannan-Quinn criter.	-8.253094
F-statistic	75.47534	Durbin-Watson stat	1.910909
Prob(F-statistic)	0.000000		

Inverted AR Roots	.83-.28i	.83+.28i
Inverted MA Roots	.47+.64i	.47-.64i

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 321-328, August-2023

https://doi.org/10.47001/IRJIET/2023.708047

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	20.5579126984814
2021	20.11938312231597
2022	19.70120699694543
2023	19.31398223279138
2024	18.96261370221408
2025	18.64730673489107
2026	18.36477449220001
2027	18.10946855544977
2028	17.87471127466095
2029	17.65365658321651
2030	17.44003775890002

Table 2 clearly indicates that neonatal mortality will decline from approximately 21 in 2020 to around 17 deaths per 1000 live births by the end of 2030.

V. POLICY IMPLICATION & CONCLUSION

Sub-Saharan African countries continue to bear a huge burden of neonatal deaths due to preventable causes. The launching of sustainable development goals in 2015 was meant to establish a roadmap to effectively address challenges in resource limited settings to improve the quality of life and living standards of different human populations across the globe including ensuring the availability of high quality maternal and child healthcare services. Setting of SDG targets was meant to stimulate a sense of responsibility and commitment to improving the quality of life for women and children who are usually victims of circumstances in different parts of the world. Several factors have been found to influence neonatal mortality, however these can be addressed timeously to avoid adverse maternal and neonatal outcomes. In this paper we proposed the Box-Jenkins ARIMA approach to forecast NMR for Kenya and the findings indicated that neonatal mortality will decline from approximately 21 in 2020 to around 17 deaths per 1000 live births by the end of 2030. Therefore, we implore Kenyan policy makers to design appropriate and country specific neonatal policies to address major drivers of mortality in neonates and such strategies should improve the quality of health care services during ANC, delivery and postnatal periods.

REFERENCES

[1] Kof AK., Maina A., Yaroh AG., Habi O., Bensaïd K., and Kalter HD (2016). Social determinants of child mortality in Niger: Results from the 2012 National Verbal and Social Autopsy Study. J Glob Health, 6,1:010603.

[2] UNICEF (2015). Levels and Trends in Child Mortality. Estimates Developed by the UN Inter-610 agency Group for Child Mortality Estimation. UN IGME Report. New York.

[3] Ludvigsson JF., Lu D., Hammarström L., Cnattingius S., and Fang F (2018). Small for gestational age and risk of childhood mortality: A Swedish population study. PLoS Med, 15,12: e1002717.

[4] Katz J., Lee AC., Kozuki N., Lawn JE., Cousens S., and Blencowe H (2013). Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. Lancet, 38,9890:417–25.

[5] Lawn J.E., Wilczynska-Ketende K., and Cousens S.N (2006). Estimating the causes of 4 million neonatal deaths in the year 2000. Int J Epidemiol. 35(3):706–18

[6] Brankovic S., Hadziomerovic A.M., Rama A., and Segalo M (2013). Incidence of morbidity and mortality in premature infants at the Department of Neonatal Intensive Care of Pediatric Clinic, Clinical Center of Sarajevo University. Med Arch. 67(4):286–8.

[7] Kenya National Bureau of Statistics and, I.C.F.I., Kenya Demographic and Health Survey, Key Indicators. 2015, ICF International Rockville.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 321-328, August-2023
https://doi.org/10.47001/IRJIET/2023.708047

[8] Olack et al (2021). Causes of preterm and low birth weight neonatal mortality in a rural community in Kenya: Evidence from verbal and social autopsy.

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Making Use of ARIMA Model Forecasts to Reconfigure Current Neonatal Health Policies to Address High Neonatal Mortality Rates in Kenya" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 321-328, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708047

****

