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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 467-474, August-2023
https://doi.org/10.47001/IRJIET/2023.708068

Identifying Solutions to Address Adverse Neonatal Outcomes in Tanzania Using Forecasts Produced By the ARIMA Model

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

[1]ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe
[2]Independent Researcher & Health Economist, Harare, Zimbabwe

***Abstract* - The emergence of new global health challenges such as outbreaks of infectious diseases, rise in chronic medical conditions, climate change, civil wars and economic recessions have stalled efforts to end all preventable deaths among under five children and older age groups around the world. Addressing maternal and child mortality is a global health priority hence there is need to channel adequate resources to maternal and child health programs in every country. Appropriate health policies can be designed by utilizing time series forecasting approaches such as the ARIMA model for the early detection of abnormal future trends of health events. In addition, substantial progress made in the achievement of health-related SDGs such as infrastructure development, poverty alleviation, sustainable agriculture, economic growth, peace and security has a positive impact on health. This study uses annual time series data on neonatal mortality rate (NMR) for Tanzania from 1968 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 (3,1,0) model. The ARIMA model predictions indicate that neonatal mortality is anticipated to drop from approximately 20 in 2020 down to around 15 deaths per 1000 live births by the end of 2030. Therefore, the Tanzanian government should craft appropriate neonatal policies to effectively deal with the problem of death of newborns. Special attention should be given to the promotion of institutional deliveries, capacitating primary healthcare and retention of healthcare workers.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

The death of a newborn that occurs between the time of birth and 28 days of life is referred to as a neonatal death (Mangu *et al*. 2021). It is a major public health problem in low and middle income countries (Hug *et al.* 2019; UN, 2017; Wang *et al*. 2016). The Sub-Saharan African (SSA) region has the largest burden of neonatal mortality reporting one death for every 38 newborns below the age of 28 days (UNICEF, 2017). Population based surveys done in the past in Tanzania revealed that neonatal mortality rate (NMR)for the country ranges between 26-40 deaths per 1000 live births (Mboera *et al*. 2015). The most important leading causes of mortality in neonates are prematurity, birth asphyxia, and neonatal sepsis (Vogel *et al.* 2014). It has been proven that levels of education among women, limited access to clean water, sanitation and quality antenatal care are associated with adverse neonatal health outcomes (WHO, 2015; Benova *et al*. 2014). The objective of this study is to model and forecast future trends of neonatal mortality rate (NMR) for Tanzania using the popular Box-Jenkins ARIMA methodology. Public health practitioners in this country and SSA are not fully utilizing this very useful statistical technique to inform their decisions and policies in public health programming (Nyoni, 2018; Box & Jenkins, 1970). This study being the first of its kind in the country to analyze neonatal mortality by using the ARIMA model, is envisioned to guide policy formulation, decision making and resource allocation to maternal and child health programs with the aim of achieving the set sustainable development goal 3 target 3.2 by 2030.

II. LITERATURE REVIEW

Several previous studies in Tanzania and the region examined factors associated with neonatal mortality although forecasting studies are rare. A retrospective cohort study was conducted by Mangu *et al*.(2021) to investigatetrends, patterns and causes of neonatal mortality in hospitals in Tanzania during 2006–2015. This retrospective study was conducted in 35 hospitals. Mortality data were obtained from inpatient registers, death registers and International Classification of Diseases-10 report forms. Annual specific hospital-based neonatal mortality rates were calculated and discussed. Two periods of 2006–2010 and 2011–2015 were assessed separately to account for data availability and interventions. It was found that neonatal mortality rate was 3.7/1000

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 467-474, August-2023
https://doi.org/10.47001/IRJIET/2023.708068

during 2006–2010 and 10.4/1000 during 2011–2015, both periods indicating a stagnant trend in the years between. The leading causes of early neonatal death were birth asphyxia (22.3%) and respiratory distress (20.8%), while those of late neonatal death were sepsis (29.1%) and respiratory distress (20.0%). Baruwa *et al.*(2021) applied duration models (Kaplan Mier and Cox proportional hazards) to examine the relationship between type of birth attendant and neonatal mortality while controlling for socio-demographic characteristics of mothers in Lesotho. The findings of the study revealed that the risk of neonatal mortality is two times higher among children delivered by non-skilled birth attendants. In another study, Norris *et al.*(2021) examined urban–rural NMR disparities among 21 SSA countries with four or more DHS, at least one of which was before 2000, using the DHS Stat Compiler. For Tanzania DHS 2015–2016, descriptive statistics were carried out disaggregated by urban and rural areas, followed by bivariate and multivariable logistic regression modelling the association between urban/rural residence and neonatal mortality, adjusting for other risk factors. The study concluded that several factors were significantly associated with higher NMR, including multiplicity of pregnancy, being the first child, higher maternal education, and male child sex. In multivariable models, urban residence remained associated with double the odds of neonatal mortality compared with rural. An investigation of trends and determinants of neonatal, post-neonatal, infant, child and under-five mortalities in Tanzania from 2004 to 2016 was carried out by Ogbo *et al.*(2019). The study used combined data from the 2004–2005, 2010 and 2015–2016 Tanzania Demographic and Health Surveys, with a sample of 25,951 singletons live births and 1585 under-five deaths. Age-specific mortality rates were calculated, followed by an assessment of trends and determinants (community, socioeconomic, individual and health service) of neonatal, post neonatal, infant, child and under-five mortalities in Cox regression models.Mothers who gave births through caesarean section, younger mothers (< 20 years), mothers who perceived their babies to be small or very small and those with fourth or higher birth rank and a short preceding birth interval (≤2 years) reported higher risk of neonatal, postneonatal and infant mortalities.

III. METHODOLOGY

The Autoregressive (AR) Model

A process M_t (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:

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

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

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

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:

$$M_t = \emptyset M_{t-1} + Z_t \quad [3]$$

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:

$M_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q} \quad [4]$ Using the backward shift operator, B, equation [4] canbe expressed as follows:

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

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

Equation [4] can also be expressed as follows:

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 467-474, August-2023

https://doi.org/10.47001/IRJIET/2023.708068

$$M_t - \sum_{j \leq 1} \pi_j M_{t-j} = Z_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [6]$$

for some constant π_j such that:

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

This implies that it is possible to invert the function taking the Z_t sequence to the M_t sequence and recover Z_t from present and past values of M_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:

$$\emptyset(B) M_t = \theta(B) Z_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [7]$$

Thus:

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

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.

$$\left.\begin{array}{c} The\ first\ difference\ is\ given\ by: \\ M_t - M_{t-1} = M_t - BM_t \\ The\ second\ difference\ is\ given\ by: \\ M_t(1-B) - M_{t-1}(1-B) = M_t(1-B) - BM_t(1-B) = M_t(1-B)(1-B) = M_t(1-B)^2 \\ The\ third\ difference\ is\ given\ by: \\ M_t(1-B)^2 - M_{t-1}(1-B)^2 = M_t(1-B)^2 - BM_t(1-B)^2 = M_t(1-B)^2(1-B) = M_t(1-B)^3 \\ The\ d^{th}\ difference\ is\ given\ by: \\ M_t(1-B)^d \end{array}\right\} \dots [9]$$

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:

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

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

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

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

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 467-474, August-2023

https://doi.org/10.47001/IRJIET/2023.708068

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 human health. In this paper, hinged on this technique; the researcher will use automatic ARIMA modeling for estimating equation [10].

Data Issues

This study is based on annual NMR in Tanzania for the period 1968 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: D(M)
Date: 01/29/22 Time: 11:45
Sample: 1968 2019
Included observations: 51

Model	LogL	AIC*	BIC	HQ
(3,0)(0,0)	47.286660	-1.658300	-1.468906	-1.585927
(2,1)(0,0)	46.932137	-1.644398	-1.455003	-1.572024
(2,2)(0,0)	47.872907	-1.642075	-1.414801	-1.555227
(4,0)(0,0)	47.843225	-1.640911	-1.413637	-1.554063
(3,1)(0,0)	47.712219	-1.635773	-1.408500	-1.548925
(1,3)(0,0)	47.426454	-1.624567	-1.397293	-1.537719
(5,2)(0,0)	50.149734	-1.613715	-1.272805	-1.483443
(1,2)(0,0)	46.030076	-1.609023	-1.419628	-1.536649
(3,2)(0,0)	47.875347	-1.602955	-1.337802	-1.501632
(2,3)(0,0)	47.875241	-1.602951	-1.337798	-1.501628
(5,0)(0,0)	47.846444	-1.601821	-1.336669	-1.500499
(4,1)(0,0)	47.846359	-1.601818	-1.336665	-1.500495
(2,0)(0,0)	44.819352	-1.600759	-1.449243	-1.542860
(1,4)(0,0)	47.450702	-1.586302	-1.321150	-1.484979
(1,0)(0,0)	43.342489	-1.582058	-1.468422	-1.538634
(1,1)(0,0)	44.246301	-1.578286	-1.426771	-1.520388
(4,3)(0,0)	49.226867	-1.577524	-1.236614	-1.447252
(5,4)(0,0)	51.182422	-1.575781	-1.159113	-1.416560
(3,3)(0,0)	48.167085	-1.575180	-1.272148	-1.459383
(3,5)(0,0)	50.117661	-1.573242	-1.194452	-1.428495
(4,2)(0,0)	47.887191	-1.564204	-1.261172	-1.448406
(2,4)(0,0)	47.875279	-1.563736	-1.260705	-1.447939
(5,1)(0,0)	47.867027	-1.563413	-1.260381	-1.447616
(4,5)(0,0)	50.666330	-1.555542	-1.138874	-1.396321
(4,4)(0,0)	49.507262	-1.549304	-1.170515	-1.404558
(1,5)(0,0)	47.497804	-1.548933	-1.245902	-1.433136
(2,5)(0,0)	48.473655	-1.547986	-1.207076	-1.417715
(5,3)(0,0)	49.154852	-1.535484	-1.156695	-1.390738
(3,4)(0,0)	48.054792	-1.531560	-1.190650	-1.401289

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

ISSN (online): 2581-3048

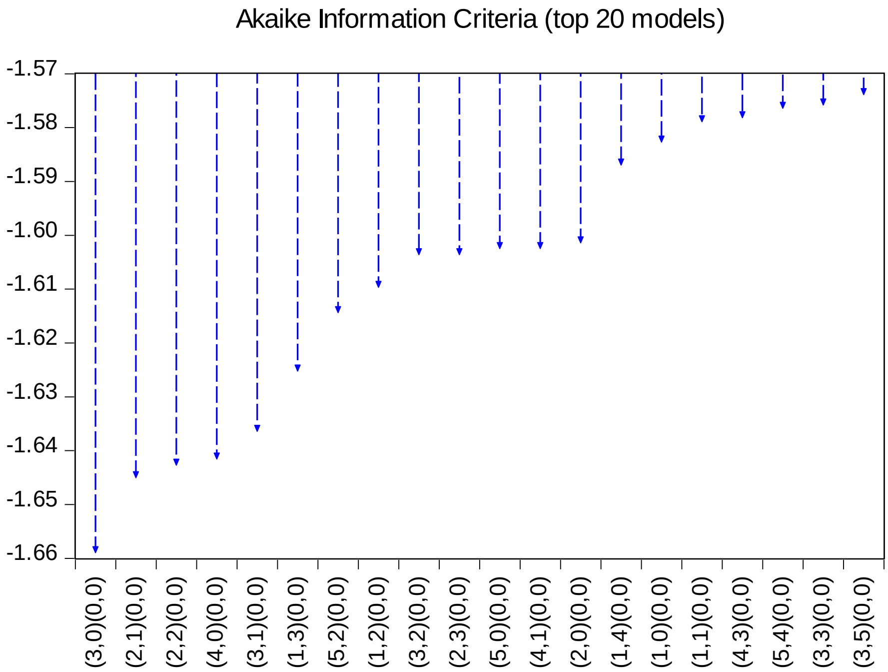
Volume 7, Issue 8, pp 467-474, August-2023

https://doi.org/10.47001/IRJIET/2023.708068

(5,5)(0,0)	50.244555	-1.499786	-1.045239	-1.326091
(0,4)(0,0)	43.585924	-1.473958	-1.246684	-1.387110
(0,5)(0,0)	39.832651	-1.287555	-1.022402	-1.186232
(0,3)(0,0)	34.910671	-1.172967	-0.983573	-1.100594
(0,2)(0,0)	30.516846	-1.039876	-0.888361	-0.981978
(0,1)(0,0)	15.206816	-0.478699	-0.365062	-0.435275
(0,0)(0,0)	-9.516668	0.451634	0.527392	0.480583

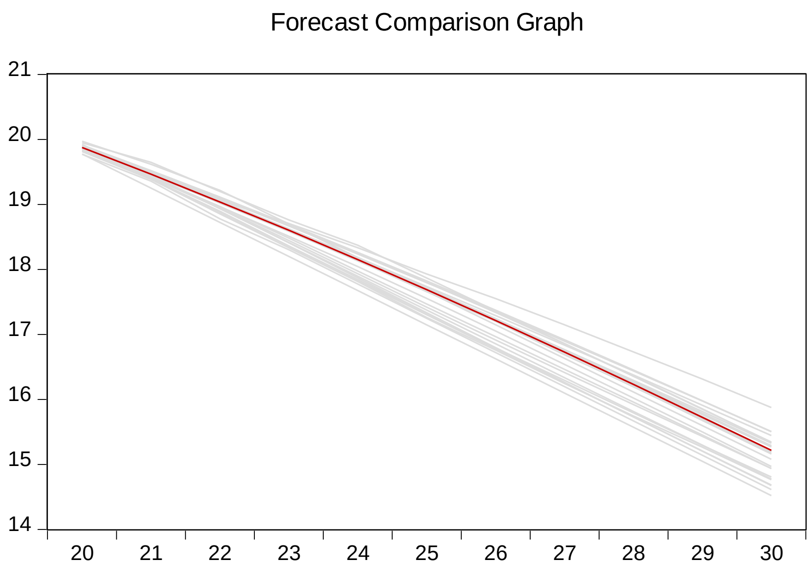
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 467-474, August-2023

https://doi.org/10.47001/IRJIET/2023.708068

Table 2 and Figure 1 indicate that the optimal model is the ARIMA (3,1,0) 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 (3,1,0) model.

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: D(M)
Date: 01/29/22 Time: 11:45
Sample: 1968 2019
Included observations: 51
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (3,0)(0,0)
AIC value: -1.65830038007

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: D(M)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/29/22 Time: 11:45
Sample: 1969 2019
Included observations: 51
Convergence achieved after 7 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.519411	0.132033	-3.933933	0.0003
AR(1)	1.065603	0.181859	5.859496	0.0000
AR(2)	0.129776	0.271693	0.477656	0.6352
AR(3)	-0.303336	0.138728	-2.186557	0.0339
SIGMASQ	0.008743	0.001795	4.871900	0.0000

R-squared	0.897183	Mean dependent var	-0.525490
Adjusted R-squared	0.888242	S.D. dependent var	0.294512
S.E. of regression	0.098456	Akaike info criterion	-1.658300
Sum squared resid	0.445905	Schwarz criterion	-1.468906
Log likelihood	47.28666	Hannan-Quinn criter.	-1.585927
F-statistic	100.3488	Durbin-Watson stat	2.029917
Prob(F-statistic)	0.000000		

Inverted AR Roots	.78-.15i	.78+.15i	-.49

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 467-474, August-2023

https://doi.org/10.47001/IRJIET/2023.708068

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	19.87413123773907
2021	19.46400778599326
2022	19.03697188608529
2023	18.60180430648855
2024	18.15100064364807
2025	17.68760997417448
2026	17.21124399361787
2027	16.72416097273847
2028	16.22779204582837
2029	15.72407308955253
2030	15.21456769281545

Table 2 clearly indicates that neonatal mortality is anticipated to drop from approximately 20 in 2020 down to around 15 deaths per 1000 live births by the end of 2030.

V. POLICY IMPLICATION & CONCLUSION

In Tanzania, prematurity, birth asphyxia, congenital anomalies and neonatal sepsis have been found to be the leading causes of death among neonates. Great efforts are being taken to address maternal and child health challenges in this country, especially that of neonates, to achieve national and international goals. In Sub-Saharan Africa, home deliveries contribute significantly to the death of neonates as this predisposes newborns to infections. Lay midwives or traditional birth attendants fail to identify complications and late referrals compromise the survival of the newborn and mother. In this study we apply the Box-Jenkins ARIMA approach to model and forecast NMR for Tanzania and the model projections indicate that neonatal mortality is anticipated to drop from approximately 20 in 2020 down to around 15 deaths per 1000 live births by the end of 2030. The Tanzanian government should craft local policies specific for this country to effectively deal with the problem of death of newborns. Special attention should be given to the promotion of institutional deliveries, capacitating primary healthcare and retention of healthcare workers.

REFERENCES

[1] Box, D. E., and Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control, Holden Day, London.

[2] Nyoni, T. (2018). Box-Jenkins ARIMA Approach to Predicting net FDI Inflows in Zimbabwe, *University Library of Munich*, MPRA Paper No. 87737.

[3] Wang H., Bhutta ZA., Coates MM (2016). Global, regional, national, and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality, 1980-2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet, 388, 1725–74.

[4] United Nations Inter-Agency Group for Child Mortality Estimation (UN IGME) (2017). Levels & Trends in Child Mortality: Report 2017, Estimates Developed by the UN Inter-Agency Group for Child Mortality Estimation. New York, NY, USA: United Nations Children's Fund.

[5] Hug L., Alexander M., and You D (2019). National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. Lancet Glob Health, 7, e710–20.

[6] United Nations Children's Fund (2017). Levels& Trends in Child Mortality Report 2017. Geneva, Switzerland: UNICEF.

[7] Mboera LEG, Ipuge Y, Kumalija CJ (2015). Mid-term review of national health plans: an example from the United Republic of Tanzania. Bull World Health Organ, 93, 271–8.

[8] Mangu C D., Rumishab S F., Lyimob E P., Mremib IR., Massawed I S., Bwanae V M., Chiduod M G., and MboeraLEG (2021). Trends, patterns and cause-specific neonatal mortality in Tanzania: a hospital-based retrospective survey, International Health 2021; 13: 334–343.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 467-474, August-2023
https://doi.org/10.47001/IRJIET/2023.708068

[9] Vogel JP., Lee ACC., and Souza JP (2014). Maternal morbidity and preterm birth in 22 low- and middle-income countries: a secondary analysis of who global survey dataset. BMC Pregnancy Childbirth 2014, 14,56.

[10] Benova L., Cumming O., and Gordon BA (2014). Where there is no toilet: water and sanitation environments of domestic and facility births in Tanzania. PLoS One 2014,9,e106738.

[11] World Health Organization (2015). Water, sanitation and hygiene in health care facilities status in low- and middle-income countries and way forward. Geneva: World Health Organization.

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Identifying Solutions to Address Adverse Neonatal Outcomes in Tanzania Using Forecasts Produced By the ARIMA Model" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 467-474, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708068

****

