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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 527-533, August-2023
https://doi.org/10.47001/IRJIET/2023.708076

Utilization of ARIMA Model Forecasts to Design and Implement Appropriate Neonatal Healthcare Interventions in the Republic of Zambia

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

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

***Abstract* - Since the beginning of the era of sustainable development goals the Zambian government has been making frantic efforts to control maternal and under five mortality. Despite a noticeable decline of under-five mortality, neonatal mortality remains a huge public health problem especially in the rural areas. Forecasting future trends of neonatal mortality will inform current neonatal healthcare solutions and allocation of resources to the maternal and child health program in the country to improve neonatal survival. Hence, this study uses annual time series data on neonatal mortality rate (NMR) for Zambia from 1969 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 ARIMA model predictions indicate that neonatal mortality will slightly drop and remain high throughout the out of sample period. Therefore, it is important for the Zambian government to channel adequate resources to maternal and child health programs in the country with special emphasis being given to improving health infrastructure particularly in the rural areas, ensuring availability of adequate and trained medical staff, and medical supplies at all levels of care. The referral system should be strengthened so that medical cases which require higher levels of care are referred appropriately and timeously.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

Neonatal mortality is the death of a newborn within the first 28 days of life (Lukonga & Michelo, 2015) and the probability of dying of a newborn baby within the first 28 days of life is known as neonatal mortality rate (NMR). NMR is expressed as the number of neonatal deaths per 1000 live births. Zambia just like any other developing country is facing the challenge of neonatal mortality due to several factors which include neonatal sepsis, prematurity, birth asphyxia and congenital anomalies (AMANHI, 2018; Lukonga & Michelo, 2015; Liu *et al*. 2015; Ganatra *et al*. 2010). In addition, unsafe home deliveries mainly in the rural areas of Zambia have contributed significantly to the death of newborns (CSO, 2015). The government is making frantic efforts to address this national health problem by implementing appropriate maternal and child health policies and is committed to the 2015 sustainable development goal 3 which aims to reduce maternal mortality ratio (MMR) to less than 70 per 100 000 live births and neonatal mortality to at least 12 per 1000 live births by 2030.The aim of this piece of work is to project future trends of neonatal mortality rate for Zambia using the popular Box-Jenkins ARIMA model (Nyoni, 2018; Box & Jenkins, 1970). The findings of this research are envisioned to guide planning, decision making and allocation of the limited resources to the maternal and child health program with the goal of reducing maternal deaths and improving child survival. Furthermore, this will help in tracking progress towards achieving set targets by 2030.

II. LITERATURE REVIEW

Park *et al*. (2021) applied logistic regression to analyse the effect of clean delivery Kits (CDKs) on perinatal and neonatal mortality in Zambia. The study findings showed that clean delivery kit use both at home and health facility deliveries were associated with lower neonatal mortality especially when certain components were used. A retrospective review was performed by Chaibva *et al.*(2019) in Zimbabwe using 346 patient records of women who delivered at Sakubva hospital and those referred from Mutare district facilities to Mutare Provincial Hospital, between January and June 2014.The study concluded that birth weight, gestational age, delivery complications and delivery method were significant contributors to adverse pregnancy outcomes. Makate & Makate (2016)examined the effect of the quality of prenatal care and its individual components on neonatal, infant and under-five mortality. Their findings indicated that a one-unit increase in the quality of prenatal care lowers the prospect of neonatal,

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 527-533, August-2023

https://doi.org/10.47001/IRJIET/2023.708076

infant and under-five mortality by approximately 42.33, 30.86 and 28.65%, respectively. In a 2015 cross-sectional study by Lukonga & Michello, it was established that low birth weight and overweight were significant predictors of neonatal mortality.

III. METHODOLOGY

The Autoregressive (AR) Model

A process Q_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:

$$Q_t = \emptyset_1 Q_{t-1} + \emptyset_2 Q_{t-2} + \emptyset_3 Q_{t-3} + \cdots + \emptyset_p Q_{t-p} + Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1]$$

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

$$Z_t = \emptyset(B) Q_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [2]$$

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:

$$Q_t = \emptyset Q_{t-1} + Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3]$$

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:

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

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

$$Q_t = \theta(B) Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [5]$$

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

Equation [4] can also be expressed as follows:

$$Q_t - \sum_{j \leq 1} \pi_j Q_{t-j} = Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [6]$$

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 Q_t sequence and recover Z_t from present and past values of Q_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:

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 527-533, August-2023
https://doi.org/10.47001/IRJIET/2023.708076

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

Thus:

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

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.

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

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:

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

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

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

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.

Data Issues

This study is based on annual NMR in Zambia for the period 1969 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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 527-533, August-2023
https://doi.org/10.47001/IRJIET/2023.708076

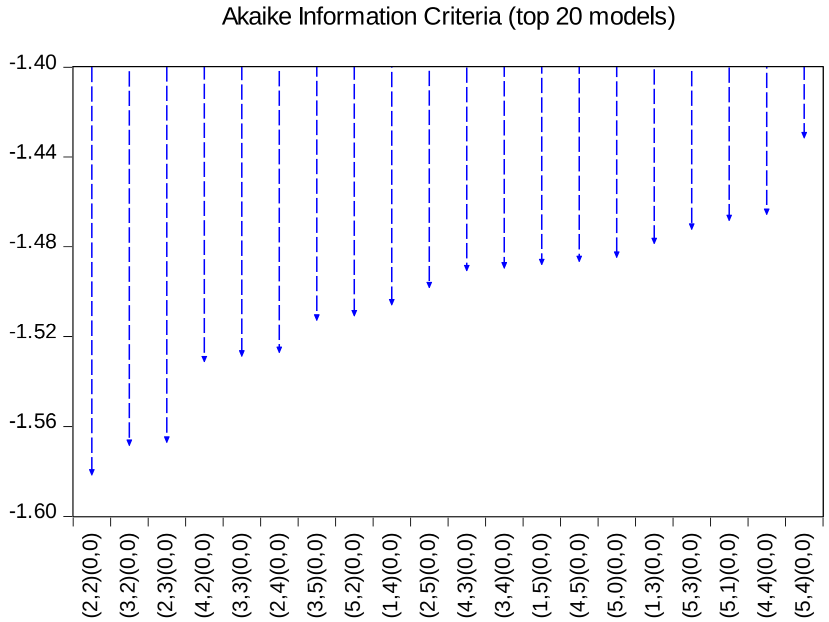
Criteria Table

Table 1: Criteria Table

Model Selection Criteria Table				
Dependent Variable: D(Q)				
Date: 01/29/22 Time: 12:38				
Sample: 1969 2019				
Included observations: 50				
Model	LogL	AIC*	BIC	HQ
(2,2)(0,0)	45.504846	-1.580194	-1.350751	-1.492821
(3,2)(0,0)	46.176715	-1.567069	-1.299385	-1.465133
(2,3)(0,0)	46.141300	-1.565652	-1.297969	-1.463717
(4,2)(0,0)	46.243950	-1.529758	-1.223834	-1.413261
(3,3)(0,0)	46.182638	-1.527306	-1.221382	-1.410808
(2,4)(0,0)	46.142288	-1.525692	-1.219768	-1.409194
(3,5)(0,0)	47.783203	-1.511328	-1.128924	-1.365706
(5,2)(0,0)	46.735158	-1.509406	-1.165242	-1.378347
(1,4)(0,0)	44.613160	-1.504526	-1.236843	-1.402591
(2,5)(0,0)	46.418750	-1.496750	-1.152586	-1.365690
(4,3)(0,0)	46.231658	-1.489266	-1.145102	-1.358207
(3,4)(0,0)	46.202927	-1.488117	-1.143953	-1.357057
(1,5)(0,0)	45.162671	-1.486507	-1.180583	-1.370009
(4,5)(0,0)	48.128909	-1.485156	-1.064511	-1.324972
(5,0)(0,0)	44.083440	-1.483338	-1.215654	-1.381402
(1,3)(0,0)	42.928077	-1.477123	-1.247680	-1.389750
(5,3)(0,0)	46.770190	-1.470808	-1.088403	-1.325186
(5,1)(0,0)	44.673986	-1.466959	-1.161036	-1.350462
(4,4)(0,0)	46.604091	-1.464164	-1.081759	-1.318542
(5,4)(0,0)	46.753331	-1.430133	-1.009488	-1.269949

Criteria Graph

Figure 1: Criteria Graph



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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 527-533, August-2023

https://doi.org/10.47001/IRJIET/2023.708076

Forecast Comparison Graph

Figure 2: Forecast Comparison Graph

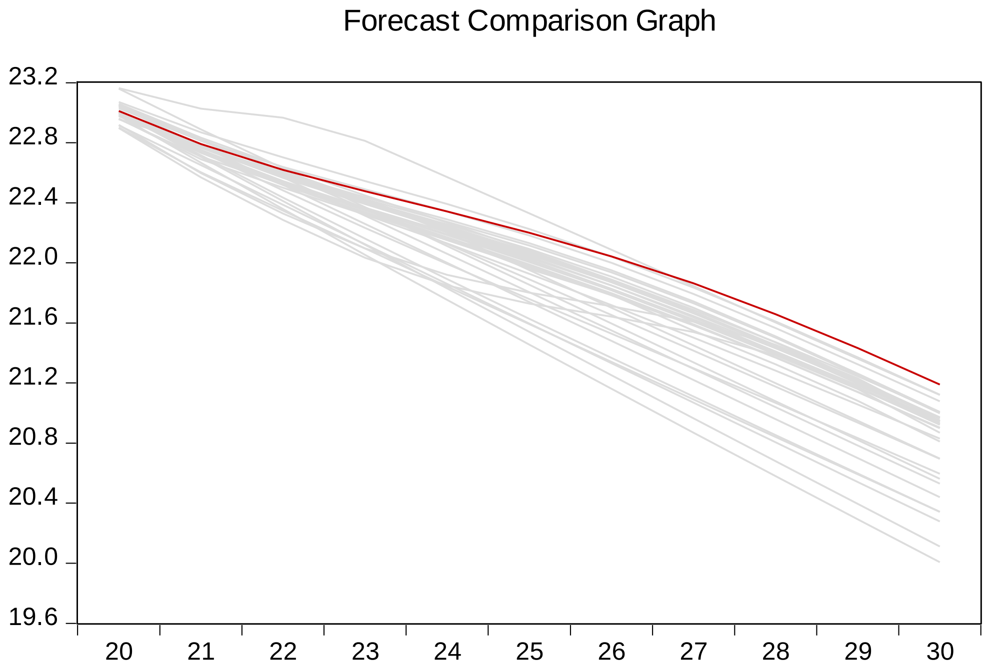


Table 1 and Figure 1 indicate that the optimal model is the ARIMA (2,1,2) model. Figure 2 is a combined forecast comparison graph showing the out-of-sample forecasts of the top 25 models evaluated based on the AIC criterion. The red line shows the forecast line graph of the optimal model, the ARIMA (2,1,2)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

IV. RESULTS

Summary of the Selected ARIMA () Model

Main Results of the Selected ARIMA () Model

Table 2: Main Results of the Optimal Model

Dependent Variable: D(Q)				
Method: ARMA Maximum Likelihood (BFGS)				
Date: 01/29/22 Time: 12:38				
Sample: 1970 2019				
Included observations: 50				
Convergence achieved after 10 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.231318	0.125165	-1.848111	0.0713
AR(1)	1.609368	0.100084	16.08018	0.0000
AR(2)	-0.745555	0.105297	-7.080507	0.0000
MA(1)	-0.572114	0.133862	-4.273920	0.0001
MA(2)	0.703921	0.142395	4.943426	0.0000
SIGMASQ	0.008621	0.002132	4.044343	0.0002
R-squared	0.943132	Mean dependent var		-

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 527-533, August-2023
https://doi.org/10.47001/IRJIET/2023.708076

			0.260000
Adjusted R-squared	0.936670	S.D. dependent var	0.393311
S.E. of regression	0.098978	Akaike info criterion	-1.580194
Sum squared resid	0.431056	Schwarz criterion	-1.350751
Log likelihood	45.50485	Hannan-Quinn criter.	-1.492821
F-statistic	145.9457	Durbin-Watson stat	1.778582
Prob(F-statistic)	0.000000		
Inverted AR Roots	.80-.31i	.80+.31i	
Inverted MA Roots	.29+.79i	.29-.79i	

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 3: Tabulated Out of Sample Forecasts

Year	Forecasts
2020	23.01199933947852
2021	22.79149034265636
2022	22.61982806376481
2023	22.47645934950567
2024	22.34220745054816
2025	22.20153343410317
2026	22.04372676877443
2027	21.8631354049343
2028	21.65864841131908
2029	21.43269180887499
2030	21.18999822067668

Table 3 clearly indicates that neonatal mortality will slightly decline and remain high throughout the out of sample period.

V. POLICY IMPLICATION & CONCLUSION

Neonatal mortality accounts for approximately 40 percent of under-five mortality globally and this is associated with socio-economic, biological and health system-related factors. Low-middle income countries like Zambia continue to bear a huge burden of deaths of newborns during the neonatal period. It is important for public health practitioners to utilize surveillance tools that will assist in the early identification of abnormal trends of NMR to inform neonatal policies and allocation of adequate resources to maternal and child health programs. This study employed the popular Box-Jenkins ARIMA methodology to predict the likely future trends of NMR for Zambia. This paper established that neonatal mortality will slightly decline and remain high throughout the out of sample period. Therefore, it is important for the Zambian government to channel adequate resources to maternal and child health programs in the country with special emphasis being given to improving health infrastructure particularly in the rural areas, ensuring availability of adequate and trained medical staff, and medical supplies at all levels of care. The referral system should be strengthened so that medical cases which require higher levels of care are referred appropriately and timeously.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 527-533, August-2023
https://doi.org/10.47001/IRJIET/2023.708076

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] Lukonga E &, C Michelo (2015). Factors associated with neonatal mortality in the general population: evidence from the 2007 Zambia Demographic and Health Survey (ZDHS); a cross sectional study, PanAfrican Medical Journal, 1-8.

[4] Liu L., Oza S., Hogan D., Perin J., Rudan I., and Lawn JE (2015). Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet, 385,9966,430–40.

[5] Ganatra HA, Stoll BJ, Zaidi AK. International perspective on early-onset neonatal sepsis. Clin Perinatol, 37,2, 501–23.

[6] Alliance for Maternal and Newborn Health Improvement (AMANHI) mortality study group. Populationbased rates, timing, and causes of maternal deaths, stillbirths, and neonatal deaths in south Asia and sub-Saharan Africa: a multi-country prospective cohort study. Lancet Glob Health. 2018; 6, 12, e1297– 308.

[7] Central Statistical Office (CSO) [Zambia], Ministry of Health (MOH) [Zambia], and ICF International. 2014. Zambia demographic and health survey 2013- 14. Rockville, Maryland, USA: Central Statistical Office, Ministry of Health, and ICF International; 2015: 126.

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, “Utilization of ARIMA Model Forecasts to Design and Implement Appropriate Neonatal Healthcare Interventions in the Republic of Zambia” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 527-533, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708076

****