

# Making Use of the Box-Jenkins ARIMA Model to Forecast Future Values of Annual Neonatal Mortality Rate for Brazil

<sup>1</sup>Dr. Smartson. P. NYONI, <sup>2</sup>Thabani NYONI

<sup>1</sup>ZICHIRE Project, University of Zimbabwe, Harare, Zimbabwe

<sup>2</sup>Independent Researcher & Health Economist, Harare, Zimbabwe

**Abstract** - The Box-Jenkins ARIMA methodology was proposed by Box and Jenkins in 1970. This technique has gained tremendous popularity in various fields such as economics, engineering and human medicine as result of its accuracy and reliable forecasts. This study uses annual time series data on neonatal mortality rate (NMR) for Brazil from 1963 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 (2) variable. The optimal model based on AIC is the ARIMA (0,2,5) model. The ARIMA model predictions indicate that NMR will decline throughout the out-of-sample period. Therefore, policy-makers are encouraged to draft and implement neonatal policies that will effectively address factors which significantly contribute to mortality in neonates.

**Keywords:** ARIMA, Forecasting, NMR.

## I. INTRODUCTION

Reducing neonatal mortality is critical in the control of under 5 mortality (UN, 2020). The sustainable development goal 3 (SDG-3) target 3.2 aims to reduce global neonatal mortality rate (NMR) to at least 12 per 1000 live births (UN, 2020; UNICEF, 2019). Brazil's neonatal mortality rate declined from 37 deaths per 1000 live births in 1990 to 18 deaths per 1000 live births in 2017 (Hug *et al.* 2019; DATASUS, 2020). As revealed by the World Bank, the country's infant mortality rate declined from 47 per 1000 live births in 1990 to 12 per 1000 live births in 2017. The occurrence of neonatal deaths indicates the quality of prenatal care, childbirth, and neonatal resuscitation (Gatt *et al.* 2016). In addition, socioeconomic and demographic factors influence mortality in neonates (Mosley *et al.* 1984). The aim of this paper is to model and forecast future trends of neonatal mortality rate (NMR) for Brazil using the Box-Jenkins ARIMA model. The model is an appropriate technique for analyzing linear time series data (Nyoni, 2018; Box & Jenkins, 1970). Forecast results are expected to detect abnormal future trends of NMR and assist in crafting appropriate neonatal healthcare strategies to keep neonatal mortality under control.

## II. LITERATURE REVIEW

Reis *et al.* (2021) evaluated the fetal and infant mortality rates due to congenital anomalies (CA) in Maranhão from 2001 to 2016 in Brazil. Data were obtained from the SINASC, and SIM databases. The study used simple linear regression, Poisson distribution, and ANOVA (Bonferroni's post hoc test) and analyzed the public data (2001–2016) of 1934858 births and determined the fetal, neonatal, perinatal, and post-neonatal mortality rates associated with CA by mesoregions. The results indicated mortality rates due to CA in Maranhão increased over the period 2001–2016 possibly as a result of improved maternal-infant health conditions eliminating other causes of death. Baroni *et al.* (2021) outlined an integrated dataset containing monthly data in a historical series from 1996 to 2017 with information on all births, neonatal deaths, and NMR (total, early and late components) enriched with information related to the municipality. It is a dataset of historical data with information on the number of births, the number of neonatal deaths, the neonatal mortality rate (including early and late), and geographic information for each month (between January 1996 and December 2017) and Brazilian municipality. Schuurmans *et al.* 2021 conducted a retrospective cohort study in Haiti to determine the prevalence of maternal death, stillbirth and low birthweight in women with (pre-) eclampsia and complicated pregnancies or deliveries in Centre de Références des Urgences Obstétricales, an obstetric emergency hospital in Port-au-Prince, Haiti, and identified the main risk factors for these adverse pregnancy outcomes. The study included pregnant women admitted to Centre de Référence des Urgences Obstétricales between 2013 and 2018 using hospital records. Risk factors investigated were age group, type of pregnancy (singleton, multiple), type of delivery and use of antenatal care services. The findings showed that of all admissions, 10 991 (34.9%) were women with (pre-) eclampsia and the main

predictors of adverse pregnancy outcomes were not attending antenatal care, low birth weight and caesarian section in patients with complicated pregnancy. In a 2019 study, Souza *et al* investigated the determinants of neonatal mortality in Foz do Iguassu in Brazil. The authors analyzed all neonatal deaths that occurred in Foz do Iguassu from 2012 to 2016. Birth and mortality data were extracted from two national governmental databases (SINASC and SIM). It was found that high rate of neonatal death in Foz do Iguassu is strongly associated with newborn characteristics and not associated with maternal socio-demographic characteristics. Boulos *et al.* (2017) investigated the etiology of severe bacterial infections in neonates. Researchers conducted a secondary retrospective analysis of a de-identified database from the Neonatal Intensive Care Unit (NICU) at Nos Petit Frères et Soeurs-St. Damien Hospital (NPFS-SDH). Records from 1292 neonates admitted to the NICU at NPFS-SDH in Port-au-Prince Haiti from 2013 to 2015 were reviewed. Sepsis accounted for 708 of 1292 (54.8%) of all admissions to the NICU. The most common organism cultured was *Streptococcus agalactiae*, followed by *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Enterobacter aerogenes*, *Staphylococcus aureus* and *Proteus mirabilis*.

### III. METHODOLOGY

#### 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 Brazil for the period 1963 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 1: Criteria Table

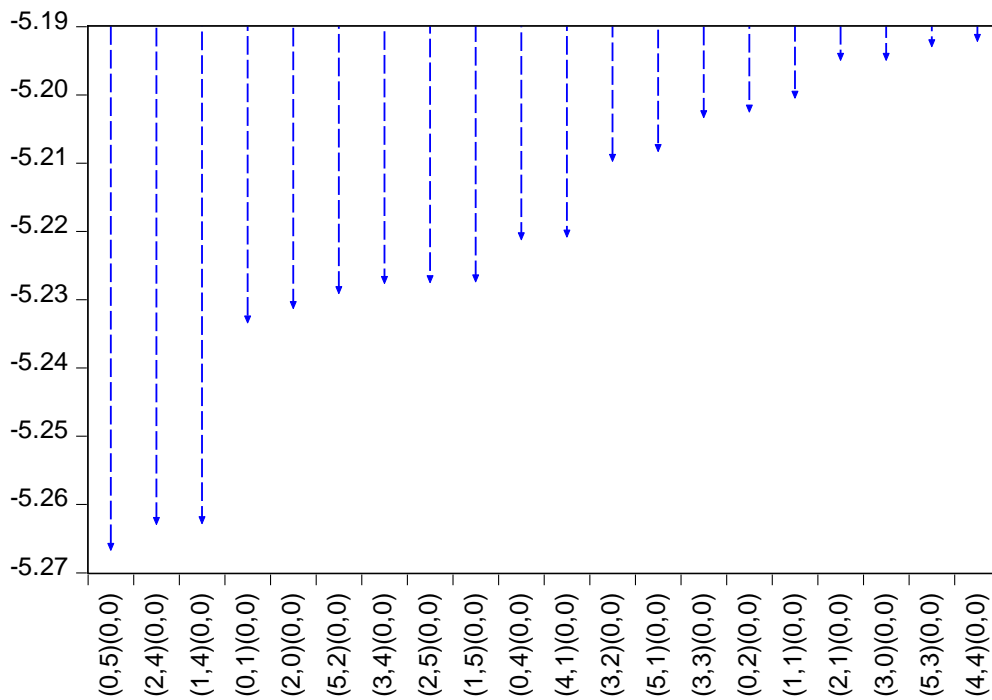
Model Selection Criteria Table				
Dependent Variable: DLOG(B, 2)				
Date: 01/22/22 Time: 12:53				
Sample: 1963 2019				
Included observations: 55				
Model	LogL	AIC*	BIC	HQ
(0,5)(0,0)	151.818205	-5.266117	-5.010638	-5.167321
(2,4)(0,0)	152.714275	-5.262337	-4.970362	-5.149428
(1,4)(0,0)	151.710346	-5.262194	-5.006716	-5.163399
(0,1)(0,0)	146.902577	-5.232821	-5.123330	-5.190480
(2,0)(0,0)	147.844540	-5.230711	-5.084723	-5.174256
(5,2)(0,0)	152.784449	-5.228525	-4.900053	-5.101502
(3,4)(0,0)	152.743113	-5.227022	-4.898550	-5.099999
(2,5)(0,0)	152.740183	-5.226916	-4.898443	-5.099893
(1,5)(0,0)	151.736921	-5.226797	-4.934821	-5.113888
(0,4)(0,0)	149.567790	-5.220647	-5.001665	-5.135965
(4,1)(0,0)	150.556190	-5.220225	-4.964746	-5.121429

(3,2)(0,0)	150.251842	-5.209158	-4.953679	-5.110362
(5,1)(0,0)	151.211618	-5.207695	-4.915719	-5.094786
(3,3)(0,0)	151.075245	-5.202736	-4.910760	-5.089827
(0,2)(0,0)	147.052489	-5.201909	-5.055921	-5.145454
(1,1)(0,0)	146.997637	-5.199914	-5.053926	-5.143459
(2,1)(0,0)	147.844880	-5.194359	-5.011874	-5.123791
(3,0)(0,0)	147.844849	-5.194358	-5.011873	-5.123790
(5,3)(0,0)	152.789453	-5.192344	-4.827374	-5.051207
(4,4)(0,0)	152.768204	-5.191571	-4.826601	-5.050434

Criteria Graph

Figure 1: Criteria Graph

Akaike Information Criteria (top 20 models)



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph

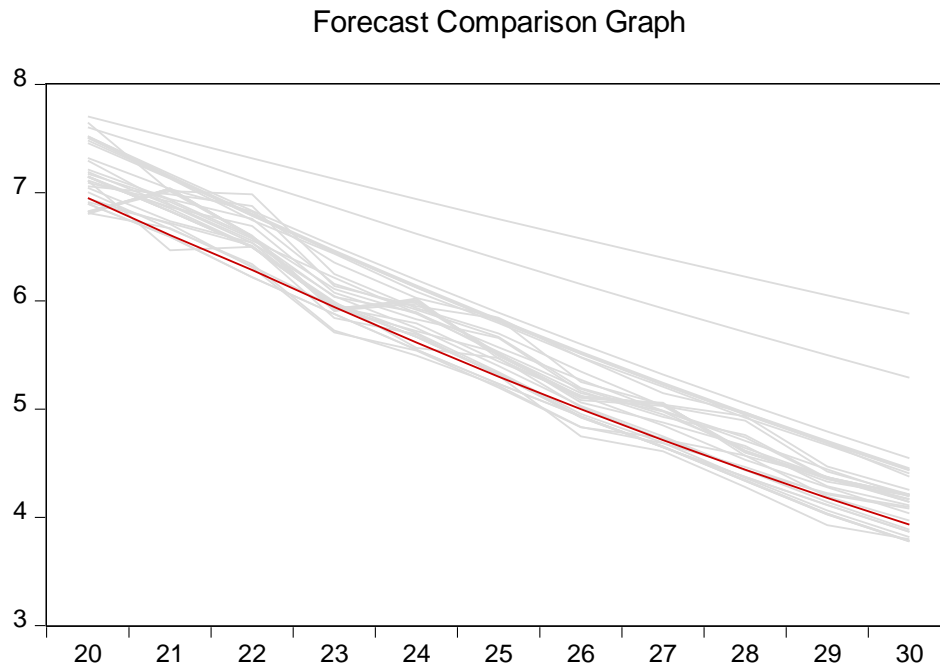


Table 1 and Figure 1 indicate that the optimal model is the ARIMA (0, 2, 5) model. Figure 2 is a combined forecast comparison graph showing the out-of-sample forecasts of the top 25 models evaluated based on the AIC criterion. The red line shows the forecast line graph of the optimal model, the ARIMA (0, 2, 5) model.

IV. RESULTS

ARIMA (0,2,5) Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

Year	Forecasts
2020	6.949724308959731
2021	6.607096623687637
2022	6.28469599454626
2023	5.942593973771727
2024	5.613887932589742
2025	5.299596891225719
2026	4.999347807197365
2027	4.71275957500717
2028	4.439444534472884
2029	4.179009915724878
2030	3.931059218522218

In line with previous studies such as Reis *et al.* (2021), Table 5 indicates that NMR is likely to continue declining over the decade from 2020 to 2030.

## V. POLICY IMPLICATION & CONCLUSION

The Brazilian government is committed to sustainable development goals as evidenced by the remarkable progress that has been reported over the past decades with the country reporting a decline in NMR over the period 1990 to 2017 from 37 neonatal deaths per 1000 live births to 18 deaths per 100 live births. Several previous studies done established that prematurity, sepsis and congenital anomalies are major causes of deaths in neonates. In this study we propose the Box-Jenkins ARIMA model to project future trends of neonatal mortality rate and the findings indicate that neonatal mortality will continue to drop throughout the forecast period. Therefore, policy-makers in this country are encouraged to draft and implement neonatal policies that will effectively address drivers of deaths among neonates.

## 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] UNICEF (2019). *Child Mortality 2019*. New York: United Nations Children's Fund.
- [4] UN (2020) sustainable development goals. <https://www.un.org/sustainabledevelopment/development-agenda>.
- [5] Hug L., Alexander M., You D., and Alkema L (2019). For Child Mortality Estimation UI-aG. 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, 6, 710–20.
- [6] DATASUS (2020). *Information systems on live births and mortality*. Brazilian Ministry of Health, 2020
- [7] Gatt M., England K., Grech V., and Calleja N (2015). Contribution of congenital anomalies to neonatal mortality rates in Malta. *Paediatr Perinat Epidemiol*, 2015, 29, 401–6
- [8] Mosley W. H., and Chen L. C (1984). *An Analytical Framework for the Study of Child Survival in Developing Countries*, 1984, 81
- [9] World Bank (2020). *Infant mortality rate*.

### Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Making Use of the Box-Jenkins ARIMA Model to Forecast Future Values of Annual Neonatal Mortality Rate for Brazil" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 225-229, August 2023. Article DOI <https://doi.org/10.47001/IRJIET/2023.708030>

\*\*\*\*\*