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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 475-481, August-2023
https://doi.org/10.47001/IRJIET/2023.708069

Calculating Future Values of Annual Neonatal Mortality Rate for Thailand Using the Box-Jenkins 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 **- This study uses annual time series data on neonatal mortality rate (NMR) for Thailand 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 (2) variable. The optimal model based on AIC is the ARIMA (3,2,4) model. The ARIMA model predictions indicate that neonatal mortality will remain low throughout the out of sample period. Therefore, we encourage policymakers to design local policies that will address various challenges affecting different parts of the country so as to keep neonatal mortality under control.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

The third sustainable development goal (SDG-3) focuses on good health and well-being for all at every stage of life. SDG-3 targets 3.1 and 3.2 aim to reduce maternal mortality ratio (MMR) to less than 70 per 100 000 live births, under five mortality to levels as low as 25 deaths per 1000 live births and neonatal mortality rate (NMR) to at least 12 per 1000 live births (UN, 2020; UNICEF, 2019; WHO; 2019; WHO; 2018). Neonatal mortality is still a challenge for Thailand. Neonatal deaths in this country contributed58 percent of all under 5 deaths in 2019 (countdown 2030, Thailand). The leading causes of neonatal mortality are birth asphyxia, prematurity, sepsis and congenital anomalies. The aim of this paper is to model and forecast neonatal mortality rate for Thailand using the Box-Jenkins ARIMA technique. Linear time series data can be easily analyzed using this traditional statistical method (Nyoni, 2018; Box & Jenkins, 1970). The findings of this study are expected to inform neonatal policies so that appropriate and effective measures are implemented to keep mortality in neonates under control.

II. LITERATURE REVIEW

Li *et al.* (2021) examined the proportion of mothers with history of neonatal deaths using the most recent Demographic and Health Surveys from 56 low- and middle-income countries. Logistic regression models were used to assess the association between maternal history of neonatal death and subsequent neonatal mortality. The adjusted models controlled for socioeconomic, child, and pregnancy-related factors. Country-specific analyses were performed to assess heterogeneity in this association across countries. Study findings suggested that maternal history of neonatal death could be an effective early identifier of high-risk pregnancies in resource-poor countries. In another study by Khader *et al.* (2021)explored the healthcare professionals' perception about the usability of JSANDS.A descriptive qualitative approach, using focus group discussions, was adopted. A total of 5 focus groups including 23 focal points were conducted in five participating hospitals in Jordan. The study findings revealed that JSANDS was perceived positively by the current users. According to them, it provides a formative and comprehensive data on stillbirths and neonatal deaths and their causes. Nath *et al.* (2020) examined the effect of extreme prematurity and early neonatal deaths on infant mortality rates in England. Authors used aggregate data on all live births, stillbirths and linked infant deaths in England in 2006–2016 from the Office for National Statistic. Infant mortality decreased from 4.78 deaths/1000 live births in 2006 to 3.54/1000 in 2014 (annual decrease of 0.15/1000) and increased to 3.67/1000 in 2016 (annual increase of 0.07/1000). This rise was driven by increases in deaths at 0–6 days of life.A descriptive study was carried out by McNamara *et al.* (2018) to reveal intrapartum fetal deaths and unexpected neonatal deaths in Ireland from 2011 to 2014. Anonymised data pertaining to all intrapartum fetal deaths and unexpected neonatal deaths for the study time period was obtained from the national perinatal epidemiology centre. The findings of the study indicated that the corrected intrapartum fetal death rate was 0.16 per 1000 births and the overall unexpected neonatal death rate was 0.17 per 1000 live births.A cross-sectional design was conducted by Soleman (2020) to examine the causes of neonatal death between 2000 and 2017. Data were taken from World Health Organization Maternal Child Epidemiology Estimation (WHO MCEE) database.The collected data were live birth; neonatal mortality rate; and the big five of neonatal mortality etiologies in the eight SEAC. Data were then analyzed descriptively with line chart to describe

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 475-481, August-2023

https://doi.org/10.47001/IRJIET/2023.708069

the trend of NMR. The findings from this study revealed that Indonesia had the highest neonatal mortality rate, yet the trend decreased gradually from 102.700 in 2000 to 60.986 in 2017, followed by Philippines, Vietnam, Myanmar, Cambodia, Thailand, Laos, and Malaysia respectively. On the other hand, the trend of live birth was the lowest in Indonesia and the highest in Philippines.

III. METHODOLOGY

The Autoregressive (AR) Model

A process Y_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:

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

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

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

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:

$$Y_t = \emptyset Y_{t-1} + Z_t \ldots\ldots\ldots [3]$$

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:

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

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

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

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

Equation [4] can also be expressed as follows:

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

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 Y_t sequence and recover Z_t from present and past values of Y_t by a convergent sum.

The Autoregressive Moving Average (ARMA) Model

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 475-481, August-2023

https://doi.org/10.47001/IRJIET/2023.708069

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:

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

Thus:

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

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.

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

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:

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

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

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

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 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 public 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 Thailand 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.

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 475-481, August-2023

https://doi.org/10.47001/IRJIET/2023.708069

Evaluation of ARIMA Models

Criteria Table

Table 2: Criteria Table

Model Selection Criteria Table
Dependent Variable: D(Y, 2)
Date: 01/29/22 Time: 11:51
Sample: 1960 2019
Included observations: 58

Model	LogL	AIC*	BIC	HQ
(3,4)(0,0)	72.435277	-2.187423	-1.867699	-2.062884
(5,2)(0,0)	72.427910	-2.187169	-1.867445	-2.062630
(1,4)(0,0)	70.371291	-2.185217	-1.936543	-2.088353
(4,5)(0,0)	74.073682	-2.174955	-1.784181	-2.022740
(2,1)(0,0)	67.905344	-2.169150	-1.991525	-2.099961
(5,0)(0,0)	69.892450	-2.168705	-1.920031	-2.071842
(2,0)(0,0)	66.720764	-2.162785	-2.020685	-2.107434
(5,3)(0,0)	72.707591	-2.162331	-1.807082	-2.023954
(3,3)(0,0)	70.695700	-2.161921	-1.877722	-2.051219
(0,5)(0,0)	69.694563	-2.161881	-1.913207	-2.065018
(2,4)(0,0)	70.616377	-2.159185	-1.874986	-2.048484
(5,4)(0,0)	73.523741	-2.155991	-1.765217	-2.003777
(3,5)(0,0)	72.508425	-2.155463	-1.800214	-2.017086
(0,4)(0,0)	68.289739	-2.147922	-1.934773	-2.064896
(4,0)(0,0)	68.234966	-2.146033	-1.932884	-2.063007
(2,2)(0,0)	68.191948	-2.144550	-1.931401	-2.061524
(5,5)(0,0)	74.117784	-2.141993	-1.715694	-1.975941
(4,1)(0,0)	69.036435	-2.139187	-1.890513	-2.042324
(3,1)(0,0)	68.014003	-2.138414	-1.925265	-2.055388
(3,0)(0,0)	66.944249	-2.136009	-1.958384	-2.066820
(5,1)(0,0)	69.910551	-2.134847	-1.850648	-2.024145
(4,2)(0,0)	69.718204	-2.128214	-1.844015	-2.017513
(2,5)(0,0)	70.647349	-2.125771	-1.806047	-2.001232
(1,5)(0,0)	69.500846	-2.120719	-1.836520	-2.010018
(2,3)(0,0)	68.415102	-2.117762	-1.869088	-2.020899
(3,2)(0,0)	68.230841	-2.111408	-1.862734	-2.014545
(1,2)(0,0)	65.559117	-2.088245	-1.910621	-2.019057
(0,2)(0,0)	64.331543	-2.080398	-1.938299	-2.025047
(1,3)(0,0)	65.631789	-2.056269	-1.843119	-1.973243
(0,3)(0,0)	64.523611	-2.052538	-1.874914	-1.983350
(1,1)(0,0)	63.319870	-2.045513	-1.903413	-1.990162
(1,0)(0,0)	60.722754	-1.990440	-1.883865	-1.948927
(0,0)(0,0)	59.674852	-1.988788	-1.917738	-1.961113
(0,1)(0,0)	60.212496	-1.972845	-1.866270	-1.931332
(4,3)(0,0)	65.883023	-1.961484	-1.641760	-1.836945
(4,4)(0,0)	66.630354	-1.952771	-1.597522	-1.814394

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

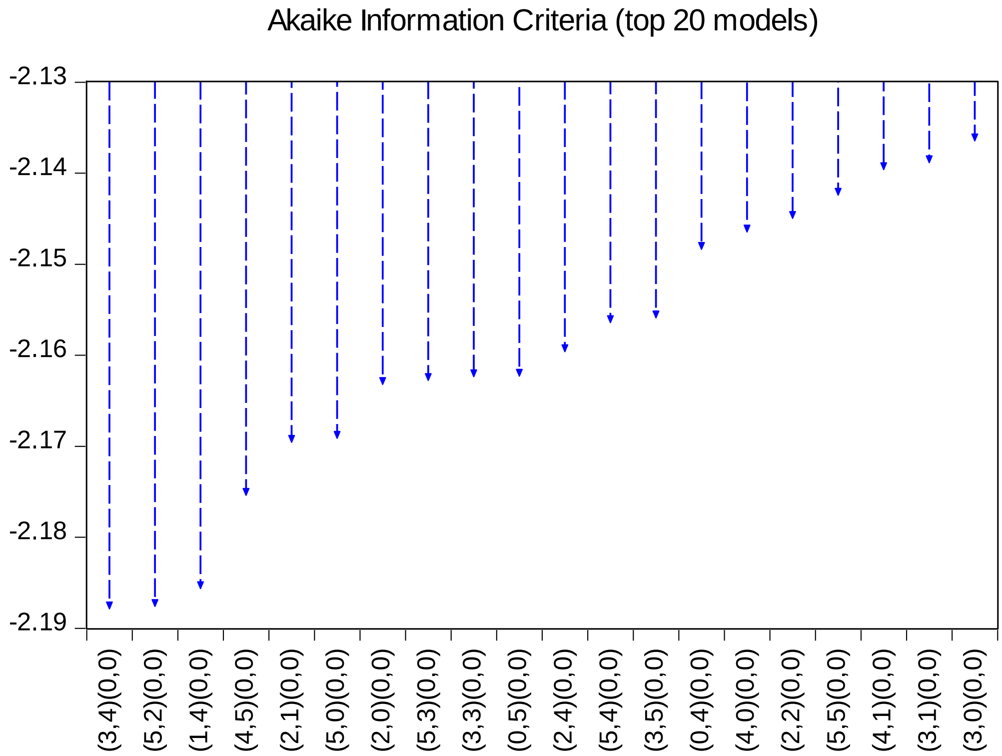
ISSN (online): 2581-3048

Volume 7, Issue 8, pp 475-481, August-2023

https://doi.org/10.47001/IRJIET/2023.708069

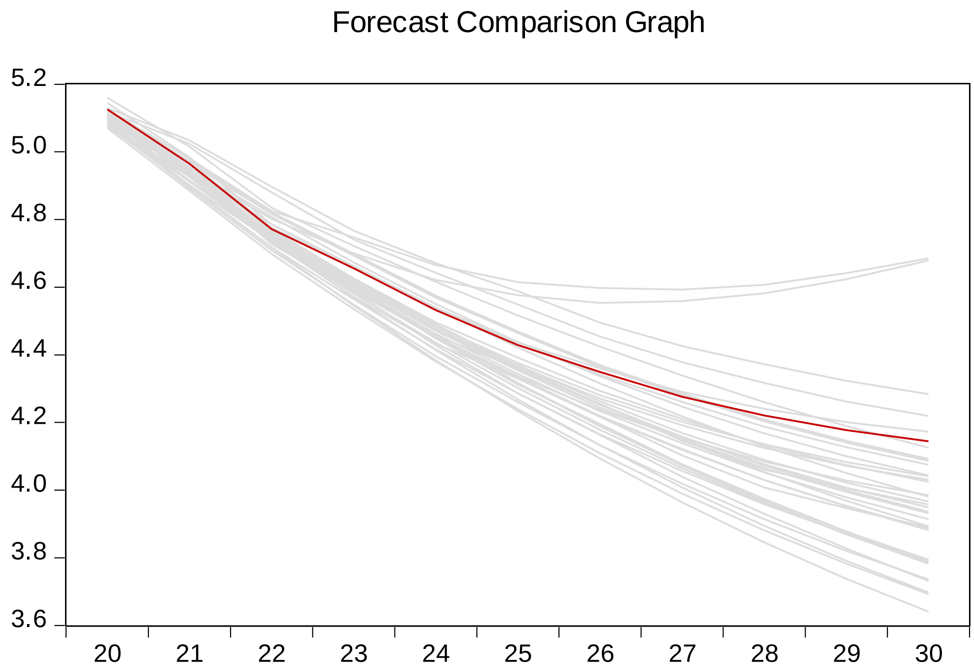
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



ISSN (online): 2581-3048
Volume 7, Issue 8, pp 475-481, August-2023
https://doi.org/10.47001/IRJIET/2023.708069

Table 2 and Figure 1 indicate that the optimal model is the ARIMA (3,2,4) 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 (3,2,4) model.

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: D(Y, 2)
Date: 01/29/22 Time: 11:51
Sample: 1960 2019
Included observations: 58
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (3,4)(0,0)
AIC value: -2.18742334256

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: D(Y,2)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/29/22 Time: 11:51
Sample: 1962 2019
Included observations: 58
Convergence achieved after 344 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007069	0.022553	0.313434	0.7553
AR(1)	-0.042381	0.283341	-0.149578	0.8817
AR(2)	0.373520	0.183389	2.036769	0.0471
AR(3)	0.328152	0.259493	1.264591	0.2120
MA(1)	-0.085003	195.2657	-0.000435	0.9997
MA(2)	-0.035909	92.22305	-0.000389	0.9997
MA(3)	-0.555367	1136.101	-0.000489	0.9996
MA(4)	0.622496	1858.972	0.000335	0.9997
SIGMASQ	0.004347	0.888667	0.004891	0.9961

R-squared	0.418829	Mean dependent var	0.010345
Adjusted R-squared	0.323944	S.D. dependent var	0.087238
S.E. of regression	0.071729	Akaike info criterion	-2.187423
Sum squared resid	0.252108	Schwarz criterion	-1.867699
Log likelihood	72.43528	Hannan-Quinn criter.	-2.062884
F-statistic	4.414065	Durbin-Watson stat	1.848180
Prob(F-statistic)	0.000452		

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 475-481, August-2023

https://doi.org/10.47001/IRJIET/2023.708069

Inverted AR Roots	.85	-.45+.43i	-.45-.43i	
Inverted MA Roots	.67+.42i	.67-.42i	-.62-.78i	-.62+.78i

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	5.125762268010975
2021	4.965005220584099
2022	4.771921378488356
2023	4.656108333700374
2024	4.531777864346958
2025	4.428470975137425
2026	4.348856671446589
2027	4.275704430400286
2028	4.220435383992622
2029	4.177005368888886
2030	4.144282345364054

Table 5 clearly indicates that neonatal mortality will remain low throughout the out of sample period.

V. POLICY IMPLICATION & CONCLUSION

The death of a newborn indicates gaps in the quality of healthcare services offered during the antenatal, delivery and postnatal periods. Most newborn deaths occur during the first week of life as a result of birth asphyxia, prematurity, sepsis and congenital anomalies and it is vital to highlight that most causes of neonatal mortality can be prevented if proper management protocols are adhered to and preventive measures are timeously put in place and implemented. This study focuses on predicting future trends of neonatal mortality rate for Thailand and the findings indicate that neonatal mortality will remain low throughout the out of sample period. Therefore, we encourage policymakers to design local policies that will address various challenges affecting different parts of the country so as to keep neonatal mortality under control.

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] World Health Organization (WHO) (2019). SDG 3: Ensure healthy lives and promote wellbeing for all at all ages.
[4] UNICEF (2018). Every Child alive. New York: UNICEF
[5] UNICEF (2019). Child Mortality 2019. New York: United Nations Children's Fund.
[6] UN (2020) sustainable development goals. https://www.un.org/sustainabl development/development-agenda
[7] Countdown 2030, Thailand.

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Calculating Future Values of Annual Neonatal Mortality Rate for Thailand Using the Box-Jenkins ARIMA Model" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 475-481, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708069

****