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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 452-459, August-2023

https://doi.org/10.47001/IRJIET/2023.708066

Improving Neonatal Survival Rates in Sri Lanka through Utilization of 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* - This study uses annual time series data on neonatal mortality rate (NMR) for Sri Lanka 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 (1,1,5) model. The study results indicate that neonatal mortality will remain low throughout the forecast period. Therefore, we encourage the government of Sri Lanka to design local policies that will keep neonatal deaths under control by focusing on improving quality, affordability and accessibility of maternal and child health care services at all levels particularly primary healthcare.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

The global sustainable development goals (SDGs) established in 2015 included a specific target for the reduction of neonatal mortality rate to at least 12 per 1000 live births in every country by 2030 (UN, 2020; WHO, 2019; UNICEF, 2019; UNICEF, 2018). High rates of neonatal mortality are being reported in Sub-Saharan Africa and Southern Asia (OCHA, 2018; Lawn *et al*. 2016).Previous studies have shown that the determinants of mortality in neonates can be classified into maternal, newborn and health system related factors. Factors such as prematurity, asphyxia, sepsis and congenital anomalies are the newborn factors that influence deaths among neonates. The majority of neonatal deaths are preventable only if proper management guidelines are crafted and implemented timeously and correctly (Lawn *et al*. 2016). The Box-Jenkins ARIMA methodology is employed in this study to model and project future trends of NMR for Sri Lanka. The statistical and econometric model is suitable for analyzing linear time series data (Nyoni, 2018; Box & Jenkins, 1970). The findings of this piece of work will inform neonatal policies & decision making and allocation of resources to maternal and child health (MNCH) programs so as to effectively control neonatal mortality in the country. Furthermore the results will assist in tracking Sri Lanka's progress towards achieving the SDG-3 target 3.2 by 2030.

II. LITERATURE REVIEW

Schellekens (2021) estimated the contribution of maternal education to infant mortality decline in Indonesia. A longitudinal, individual-level analysis of the determinants of trends in infant mortality in Indonesia was done by utilizing pooled data from all available phases of the Demographic and Health Survey (1980-2015).The study findings revealed that maternal education explains 15% of the infant mortality decline in Indonesia from 1980 to 2015. Soleman *et al*. (2020) conducted a cross-sectional study in Indonesia to describe trends and main causes of children mortality in Indonesia from 2000 to 2017.The data was taken from World Health Organization Maternal Child Epidemiology Estimation from 2000 to 2017. The study found that the trend of three parameters of child mortality declined within 17 years and the main causes of mortality were premature birth in neonates, ARI in post neonates and premature birth in under five children. A prospective, population-based research study was conducted by Dhaded *et al*. (2020) to investigate neonatal deaths in rural Karnataka, India for the period 2014–2018.Study staff collected demographic and health care characteristics on eligible women enrolled with neonatal outcomes obtained at delivery and day 28. Cause of neonatal mortality at day 28 was assigned by algorithm using prospectively defend variables. Study found that infants who were preterm and low-birth weight remained at highest risk for 28-day neonatal mortality in India. In another study, Khan *et al. (*2020) assessed the extent to which maternal histories of newborn danger signs independently or combined with birth weight and/or gestational age (GA) can capture and/or predict post second day (age>48 hours) neonatal death. Prognostic multivariable models showed that maternally recalled danger signs, coupled to either birth weight or GA, can predict and capture post-second day neonatal death with high discrimination and sensitivity. A comparison of Pakistan's under-five mortality, neonatal mortality, and postnatal newborn care rates with those of other countries was performed by Ahmed *et al*. (2017). Neonatal mortality rates and postnatal newborn care rates from the Demographic and Health Surveys (DHSs) of nine low- and middle-

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 452-459, August-2023

https://doi.org/10.47001/IRJIET/2023.708066

income countries (LMIC) from Asia and Africa were analyzed. Pakistan's maternal, newborn, and child health (MNCH) policies and programs, which have been implemented in the country since 1990, were also analyzed. The results highlighted that postnatal newborn care in Pakistan was higher compared with the rest of countries, yet its neonatal mortality remained the worst.

III. METHODOLOGY

The Autoregressive (AR) Model

A process S_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:

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

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

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

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:

$$S_t = \emptyset S_{t-1} + Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [3]$$

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:

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

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

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

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

Equation [4] can also be expressed as follows:

$$S_t - \sum_{j \le 1} \pi_j S_{t-j} = Z_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [6]$$

for some constant π_j such that:

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

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

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 7, Issue 8, pp 452-459, August-2023
https://doi.org/10.47001/IRJIET/2023.708066

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

Thus:

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

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.

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

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:

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

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

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

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, 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 Sri Lanka 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

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 452-459, August-2023

https://doi.org/10.47001/IRJIET/2023.708066

Table 2: Criteria Table

Model Selection Criteria Table
Dependent Variable: DLOG(S)
Date: 01/29/22 Time: 11:33
Sample: 1960 2019
Included observations: 59

Model	LogL	AIC*	BIC	HQ
(1,5)(0,0)	141.950298	-4.540688	-4.258988	-4.430724
(2,3)(0,0)	140.670732	-4.531211	-4.284724	-4.434993
(3,0)(0,0)	138.564927	-4.527625	-4.351562	-4.458897
(2,5)(0,0)	142.420184	-4.522718	-4.205806	-4.399008
(5,1)(0,0)	141.035827	-4.509689	-4.227989	-4.399725
(5,3)(0,0)	142.942576	-4.506528	-4.154403	-4.369073
(3,3)(0,0)	140.912088	-4.505495	-4.223795	-4.395530
(4,0)(0,0)	138.759277	-4.500314	-4.289039	-4.417841
(3,1)(0,0)	138.755356	-4.500182	-4.288907	-4.417708
(4,3)(0,0)	141.697658	-4.498226	-4.181313	-4.374516
(2,4)(0,0)	140.677652	-4.497548	-4.215848	-4.387583
(4,2)(0,0)	140.628383	-4.495877	-4.214177	-4.385913
(4,5)(0,0)	143.482100	-4.490919	-4.103581	-4.339718
(3,5)(0,0)	142.421282	-4.488857	-4.136732	-4.351402
(3,4)(0,0)	141.174278	-4.480484	-4.163572	-4.356774
(5,4)(0,0)	143.084018	-4.477424	-4.090087	-4.326223
(3,2)(0,0)	138.793638	-4.467581	-4.221093	-4.371362
(5,0)(0,0)	138.775810	-4.466977	-4.220489	-4.370758
(4,1)(0,0)	138.762941	-4.466540	-4.220053	-4.370322
(5,5)(0,0)	143.606269	-4.461229	-4.038679	-4.296283
(5,2)(0,0)	140.425609	-4.455105	-4.138193	-4.331396
(2,1)(0,0)	136.353826	-4.452672	-4.276610	-4.383944
(1,3)(0,0)	137.268915	-4.449794	-4.238519	-4.367321
(0,4)(0,0)	137.245000	-4.448983	-4.237708	-4.366510
(4,4)(0,0)	141.232515	-4.448560	-4.096435	-4.311104
(2,2)(0,0)	137.096654	-4.443954	-4.232679	-4.361481
(1,4)(0,0)	137.972389	-4.439742	-4.193255	-4.343523
(0,2)(0,0)	134.961931	-4.439387	-4.298537	-4.384405
(0,5)(0,0)	137.780640	-4.433242	-4.186755	-4.337023
(1,2)(0,0)	135.512669	-4.424158	-4.248096	-4.355431
(0,3)(0,0)	135.305034	-4.417120	-4.241057	-4.348392
(2,0)(0,0)	131.057059	-4.307019	-4.166169	-4.252037
(1,1)(0,0)	129.487791	-4.253823	-4.112973	-4.198841
(1,0)(0,0)	128.404420	-4.250997	-4.145360	-4.209761
(0,1)(0,0)	122.963200	-4.066549	-3.960912	-4.025313
(0,0)(0,0)	110.071770	-3.663450	-3.593025	-3.635959

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

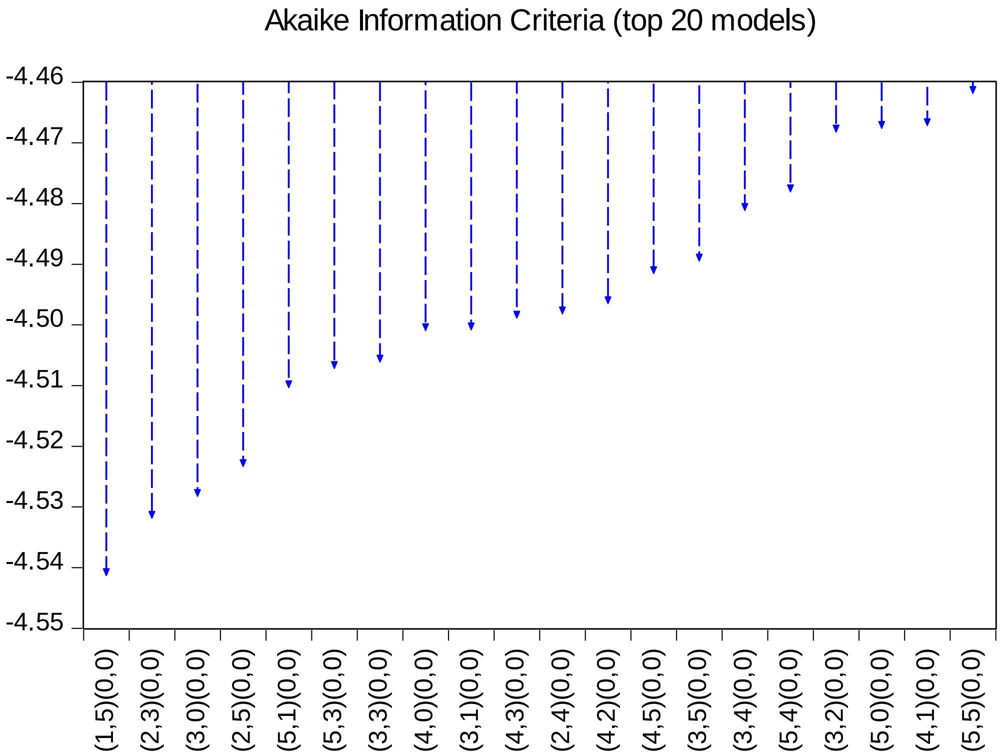
ISSN (online): 2581-3048

Volume 7, Issue 8, pp 452-459, August-2023

https://doi.org/10.47001/IRJIET/2023.708066

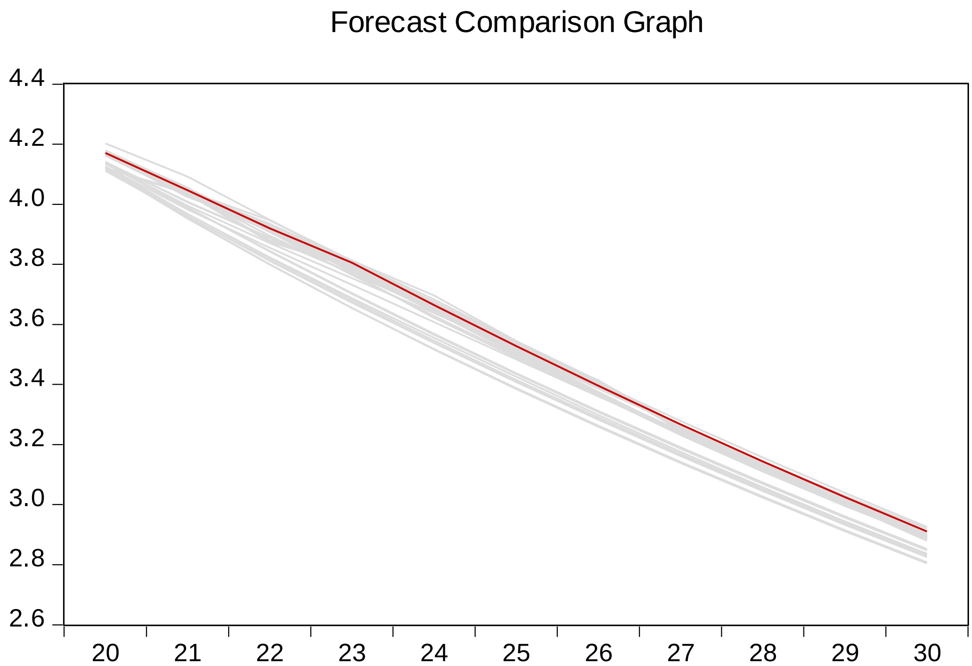
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



ISSN (online): 2581-3048
Volume 7, Issue 8, pp 452-459, August-2023
https://doi.org/10.47001/IRJIET/2023.708066

Table 2 and Figure 1 indicate that the optimal model is the ARIMA (1,1,5) 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 (1,1,5) model.

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: DLOG(S)
Date: 01/29/22 Time: 11:33
Sample: 1960 2019
Included observations: 59
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (1,5)(0,0)
AIC value: -4.54068806654

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: DLOG(S)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/29/22 Time: 11:33
Sample: 1961 2019
Included observations: 59
Convergence achieved after 216 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.038668	0.001782	-21.70288	0.0000
AR(1)	0.597163	0.180627	3.306050	0.0017
MA(1)	0.066606	149.8311	0.000445	0.9996
MA(2)	0.216832	359.3223	0.000603	0.9995
MA(3)	-0.638391	908.9362	-0.000702	0.9994
MA(4)	-0.154865	268.0558	-0.000578	0.9995
MA(5)	-0.490179	1293.560	-0.000379	0.9997
SIGMASQ	0.000442	0.049956	0.008844	0.9930

R-squared	0.685103	Mean dependent var	-0.037459
Adjusted R-squared	0.641882	S.D. dependent var	0.037779
S.E. of regression	0.022608	Akaike info criterion	-4.540688
Sum squared resid	0.026067	Schwarz criterion	-4.258988
Log likelihood	141.9503	Hannan-Quinn criter.	-4.430724
F-statistic	15.85112	Durbin-Watson stat	1.888148
Prob(F-statistic)	0.000000		

Inverted AR Roots .60

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 452-459, August-2023

https://doi.org/10.47001/IRJIET/2023.708066

Inverted MA Roots	1.00	.05+.79i	.05-.79i	-.59+.67i
	-.59-.67i			

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	4.170435155035974
2021	4.046857689506466
2022	3.920207222421853
2023	3.805235928994024
2024	3.664852127041842
2025	3.528115450694396
2026	3.395600071731582
2027	3.267556063424247
2028	3.144049774244792
2029	3.025044743371636
2030	2.91044820830868

Table 5 clearly indicates that neonatal mortality will remain low throughout the forecast period.

V. POLICY IMPLICATION & CONCLUSION

It is apparent that low and middle income countries will continue to fight the problem of neonatal mortality for a couple of years from now due to several challenges that affect them especially persistence of poverty, hunger, armed conflicts and mass exodus of seasoned and experienced healthcare professionals. In addition, dilapidated health infrastructure, poor road networks and natural disasters aggravate the worrisome situation. Of note, there is underutilization of early surveillance tools such as time series forecasting techniques that usually help to inform policies, decisions and allocation of resources. In this piece of work we applied the ARIMA model to project future trends of NMR for Sri Lanka and the findings indicate that neonatal mortality will remain low throughout the forecast period. Therefore, we encourage the government of Sri Lanka to design local policies that will keep neonatal deaths under control by focusing on improving quality, affordability and accessibility of maternal and child health care services at all levels particularly primary healthcare.

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] UN Office for Coordination of Humanitarian Affairs (OCHA) (2018). Global Humanitarian Overview: 2018. Geneva.

[6] Lawn J. E., Blencowe H., Kinney M.V., Bianchi F., and Graham W. J (2016). Evidence to inform the future for maternal and newborn health. Best Pract Res Clin Obstetr Gynaecol. 2016, 36, 169–83.

[7] UNICEF (2019). Child Mortality 2019. New York: United Nations Children's Fund.

[8] UN (2020) sustainable development goals. https://www.un.org/sustainabl development/development-agenda.

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 452-459, August-2023

https://doi.org/10.47001/IRJIET/2023.708066

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Improving Neonatal Survival Rates in Sri Lanka through Utilization of Forecasts Produced By the ARIMA Model" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 452-459, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708066

****