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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 489-496, August-2023

https://doi.org/10.47001/IRJIET/2023.708071

Approximation of Anticipated Future Values of Annual Neonatal Mortality Rates for Tunisia Using 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 utilizes annual time series data on neonatal mortality rate (NMR) for Tunisia from 1965 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 (3,2,2) model. The ARIMA model predictions indicate that neonatal mortality is expected to remain below 12 deaths per 1000 live births throughout the forecast period. Therefore, authorities in Tunisia are encouraged to address local issues that contribute to neonatal deaths especially in marginalized regions of the country.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

Time series modelling and forecasting has been utilized by many researchers around the world and has proved to be an essential tool to detect abnormal disease trends early to facilitate resource allocation, planning and drafting of policies (Zhang*et al*. 2014). There are various modelling techniques that have been outlined in literature such as the Box-Jenkins ARIMA model which is regarded by many researchers as one of the popular statistical techniques for modeling linear data (Nyoni, 2018; Yu *et al*. 2014; Box & Jenkins, 1970). The procedure is a 3 step iterative process involving model identification, parameter estimation and diagnostic checking (Box & Jenkins, 1970). The model is as expressed as ARIMA (p, d, q) where p and q represent the non-seasonal autoregressive and moving average parts & d' represents the number of times the time series under consideration needs to be differenced in order to become stationary (Nyoni & Nyoni, 2020; Nyoni & Nyoni , 2019 a & b; Yu *et al.* 2014; Box & Jenkins, 1970). In maternal and child health programs it is important to assess program performance by carrying out surveillance of maternal, under five and neonatal mortality. One of the surveillance tools is the application of time series modeling and forecasting of neonatal mortality rate (NMR) with the aim to have an insight of the likely future trends of NMR to inform policy, decision making and resource allocation. Therefore in this paper we proposed the Box-Jenkins ARIMA approach to analyze NMR for Tunisia and the findings will form the basis of drafting neonatal policies that are effective in controlling neonatal deaths in the country.

II. LITERATURE REVIEW

Tesema and Worku (2021) investigated the individual and community-level determinants of neonatal mortality in the Emerging regions of Ethiopia. A multilevel binary logistic regression was fitted to identify the significant determinants of neonatal mortality. The findings of the study revealed Neonatal mortality in Emerging regions of Ethiopia was unacceptably high. Significant predictors of neonatal mortality were maternal education, women's autonomy in making decisions for health care, sex of a child, type of birth, preceding birth interval, ANC visit, and community media exposure. A cross-sectional study by Tanou *et al*.(2021) assessed the effect of geographical accessibility to health facilities on antenatal care and delivery services utilization in Benin, with an emphasis on geographical zones. The study employed multivariate logistic regression for analysis and the findings indicated that the distance to the closest health center had adverse effects on the likelihood of a woman receiving appropriate maternal healthcare. The estimates showed that one km increase in straight line distance to the closest health center reduces the odds of the woman receiving at least one antenatal care by 0.042, delivering in facility by 0.092, and delivering her baby with assistance of skilled birth attendants by 0.118.An investigation of factors associated with neonatal mortality at the Referral Hospital in Nouakchott, Mauritania was carried out by Weddih *et al*.(2019). Across-sectional study was conducted between January 2013 and December 2013 and included neonatal patients hospitalized at the National Referral Hospital (NRH). Data were collected by reviewing the medical charts and through questionnaires administered to the parents. The findings of the study

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 489-496, August-2023
https://doi.org/10.47001/IRJIET/2023.708071

indicated that significant predictors of neonatal mortality were low birth weight, hypothermia and birth outside NRH. Merabet *et al*. (2018) described neonatal deaths and identified risk factors at the Al Hoceima Provincial Hospital in Morocco. The study concluded that neonatal mortality in the Al Hoceima hospital remains high and is mainly related to the course of pregnancy and childbirth as well as the characteristics of the newborn at birth. Brault *et al*.(2018)examined factors contributing to the reductions of under-five mortality. A case study mixed methods approach drawing on data from quantitative indicators, national documents and qualitative interviews was used to describe factors that enabled Liberia to rebuild their maternal, neonatal and child health (MNCH) programmes and reduce under-five mortality following the country's civil war.The authors found out that three prominent factors contributed to the reduction in under-five mortality: national prioritization of MNCH after the civil war; implementation of integrated packages of services that expanded access to key interventions and promoted inter-sectoral collaborations; and use of outreach campaigns, community health workers and trained traditional midwives to expand access to care and improve referrals.

III. METHODOLOGY

The Autoregressive (AR) Model

A process M_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:

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

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:

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

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:

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

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:

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

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

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

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

Equation [4] can also be expressed as follows:

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

for some constant π_j such that:

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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 489-496, August-2023
https://doi.org/10.47001/IRJIET/2023.708071

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:

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

Thus:

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

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.

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

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:

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

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

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

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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 489-496, August-2023
https://doi.org/10.47001/IRJIET/2023.708071

Data Issues

This study is based on annual NMR in Tunisia for the period 1965 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, 2)
Date: 01/29/22 Time: 12:04
Sample: 1965 2019
Included observations: 53

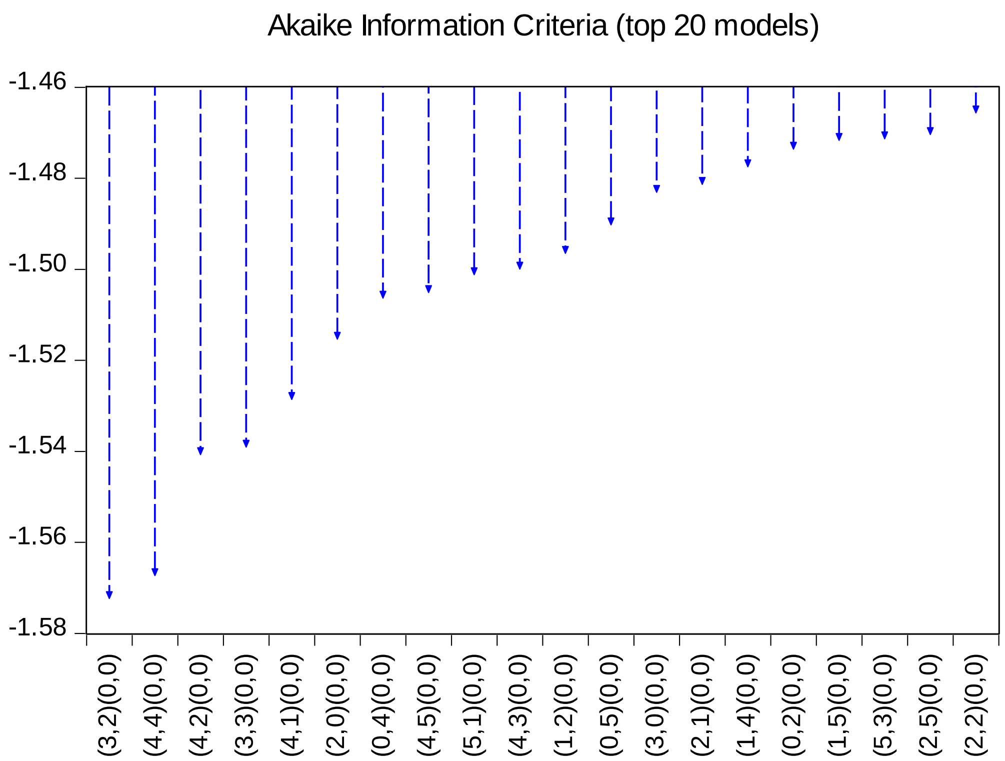
Model	LogL	AIC*	BIC	HQ
(3,2)(0,0)	48.644834	-1.571503	-1.311276	-1.471432
(4,4)(0,0)	51.510830	-1.566446	-1.194693	-1.423488
(4,2)(0,0)	48.805499	-1.539830	-1.242428	-1.425463
(3,3)(0,0)	48.762687	-1.538215	-1.240812	-1.423848
(4,1)(0,0)	47.485349	-1.527749	-1.267522	-1.427678
(2,0)(0,0)	44.133968	-1.514489	-1.365788	-1.457306
(0,4)(0,0)	45.895522	-1.505491	-1.282439	-1.419716
(4,5)(0,0)	50.862970	-1.504263	-1.095335	-1.347009
(5,1)(0,0)	47.758955	-1.500338	-1.202935	-1.385971
(4,3)(0,0)	48.726218	-1.499103	-1.164525	-1.370440
(1,2)(0,0)	44.633791	-1.495615	-1.309738	-1.424136
(0,5)(0,0)	46.468735	-1.489386	-1.229159	-1.389315
(3,0)(0,0)	44.279470	-1.482244	-1.296368	-1.410765
(2,1)(0,0)	44.232586	-1.480475	-1.294598	-1.408996
(1,4)(0,0)	46.130540	-1.476624	-1.216397	-1.376553
(0,2)(0,0)	43.027241	-1.472726	-1.324025	-1.415543
(1,5)(0,0)	46.975619	-1.470778	-1.173376	-1.356411
(5,3)(0,0)	48.967695	-1.470479	-1.098726	-1.327521
(2,5)(0,0)	47.942560	-1.469531	-1.134953	-1.340868
(2,2)(0,0)	44.815733	-1.464745	-1.241693	-1.378970
(0,3)(0,0)	43.780397	-1.463411	-1.277535	-1.391932
(1,3)(0,0)	44.680000	-1.459623	-1.236571	-1.373848
(5,2)(0,0)	47.613183	-1.457101	-1.122523	-1.328439
(3,1)(0,0)	44.498497	-1.452773	-1.229722	-1.366998
(2,4)(0,0)	46.475947	-1.451923	-1.154520	-1.337556
(4,0)(0,0)	44.466938	-1.451583	-1.228531	-1.365808
(2,3)(0,0)	45.022277	-1.434803	-1.174576	-1.334732
(3,5)(0,0)	47.950073	-1.432078	-1.060325	-1.289120
(1,1)(0,0)	41.522080	-1.415928	-1.267226	-1.358744
(5,0)(0,0)	44.473408	-1.414091	-1.153864	-1.314020
(5,4)(0,0)	48.232751	-1.405009	-0.996081	-1.247755
(0,0)(0,0)	38.965538	-1.394926	-1.320575	-1.366334
(1,0)(0,0)	39.626153	-1.382119	-1.270593	-1.339231
(5,5)(0,0)	48.408333	-1.373899	-0.927796	-1.202349
(0,1)(0,0)	39.336146	-1.371175	-1.259649	-1.328288

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 489-496, August-2023
https://doi.org/10.47001/IRJIET/2023.708071

(3,4)(0,0)	38.433383	-1.110694	-0.776116	-0.982031

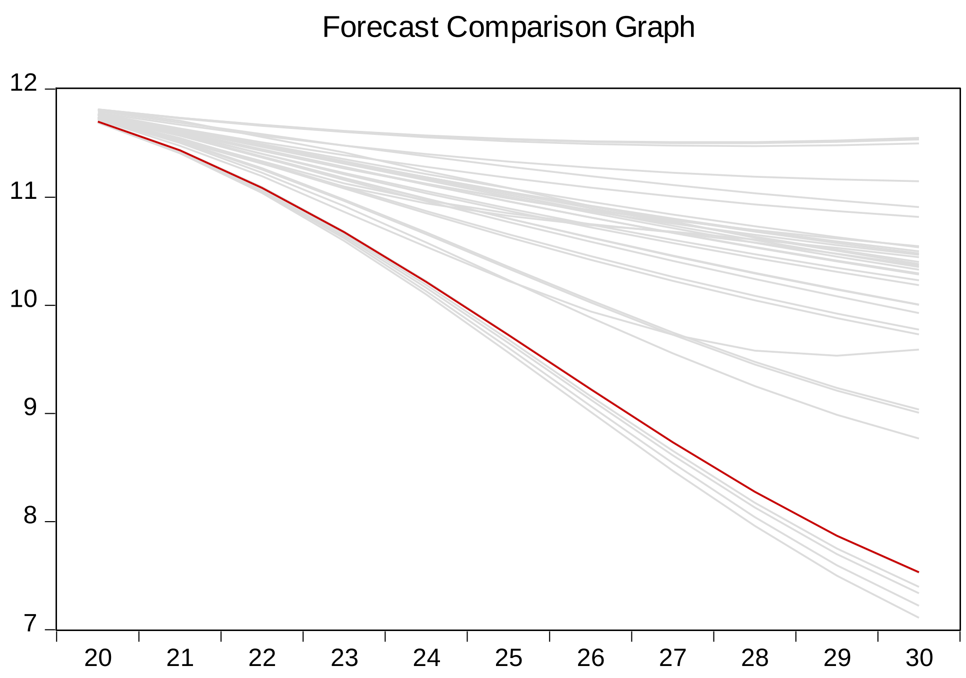
Criteria Graph

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 489-496, August-2023
https://doi.org/10.47001/IRJIET/2023.708071

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

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: D(M, 2)
Date: 01/29/22 Time: 12:04
Sample: 1965 2019
Included observations: 53
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (3,2)(0,0)
AIC value: -1.57150315928

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: D(M,2)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/29/22 Time: 12:04
Sample: 1967 2019
Included observations: 53
Convergence achieved after 72 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.017070	0.008413	2.028881	0.0483
AR(1)	1.499037	0.156814	9.559349	0.0000
AR(2)	-0.217817	0.289101	-0.753431	0.4550
AR(3)	-0.373975	0.140944	-2.653357	0.0109
MA(1)	-1.729954	36.56342	-0.047314	0.9625
MA(2)	0.729954	29.36472	0.024858	0.9803
SIGMASQ	0.008428	0.106671	0.079011	0.9374

R-squared	0.373684	Mean dependent var	0.011321
Adjusted R-squared	0.291991	S.D. dependent var	0.117113
S.E. of regression	0.098543	Akaike info criterion	-1.571503
Sum squared resid	0.446693	Schwarz criterion	-1.311276
Log likelihood	48.64483	Hannan-Quinn criter.	-1.471432
F-statistic	4.574226	Durbin-Watson stat	1.895668
Prob(F-statistic)	0.001020		

Inverted AR Roots	.94+.25i	.94-.25i	-.39

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 489-496, August-2023

https://doi.org/10.47001/IRJIET/2023.708071

Inverted MA Roots	1.00	.73

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	11.69948042298777
2021	11.43128206811521
2022	11.08510896273172
2023	10.67596560100672
2024	10.2163052679633
2025	9.725378096364134
2026	9.223716893901016
2027	8.73325086318353
2028	8.275181212836899
2029	7.86883401412276
2030	7.530360790984254

Table 5 clearly indicates that neonatal mortality is expected to remain below 12 deaths per 1000 live births throughout the forecast period.

V. POLICY IMPLICATION & CONCLUSION

Statistical techniques such as the ARIMA model have gained popularity in time series forecasting problems in various fields including Health sciences. The time series under consideration should be linear in nature for one to apply the ARIMA model. In this study we propose to apply this technique to model and forecast future trends of neonatal mortality rate for Tunisia. The results of the study indicate that neonatal mortality is expected to remain below 12 deaths per 1000 live births throughout the forecast period. Therefore, authorities in Tunisia are encouraged to address local issues that contribute to neonatal deaths especially in marginalized regions of the country.

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] Nyoni & Nyoni T (2019). Forecasting TB notifications at Silobela District Hospital, Zimbabwe, IJARIIE 5(6)2395-4396.

[4] Nyoni & Nyoni T (2019). Forecasting TB notifications at Zengeza clinic, Zimbabwe. Online at https://mpra.ub.uni-muenchen.de/97331/ MPRA Paper No. 97331, posted 02 Dec 2019 10:13 UTC

[5] Nyoni S P & Nyoni T (2020).Modelling and forecasting infant deaths in Zimbabwe using ARIMA models. NOVATEUR PUBLICATIONS JournalNX- A Multidisciplinary Peer Reviewed Journal 6(7): 2581 – 4230

[6] Xingyu Zhang., Tao Zhang., Alistair A., Young., & Xiaosong Li (2014). Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data PLOS ONE | www.plosone.org | Volume 9 | Issue 2 | e88075

[7] Yu L J., Zhou L L., Tan L., Jiang H B., Wang Y., Wei S., & Nie S F (2014). Application of a new hybrid model with seasonal auto-regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (ANRNN) in forecasting incidence cases of HFMD in Shenzen, China. PLoS ONE 2014, 9, e98241.

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 489-496, August-2023

https://doi.org/10.47001/IRJIET/2023.708071

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Approximation of Anticipated Future Values of Annual Neonatal Mortality Rates for Tunisia Using the ARIMA Model" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 489-496, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708071

****