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

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 519-526, August-2023
https://doi.org/10.47001/IRJIET/2023.708075

Calculating Expected Future Values of Annual Neonatal Mortality Rate for Yemen 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 uses annual time series data on neonatal mortality rate (NMR) for Yemen from 1962 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 (2,2,5) model. The ARIMA model predictions indicate that neonatal mortality is expected to slightly increase and remain high throughout the out of sample period. Therefore, it is crucial for the authorities in Yemen to design and implement appropriate neonatal policies to urgently address causes of neonatal deaths across the country. Strategies should focus on improving the quality, affordability and accessibility of maternal and neonatal healthcare services.**

Keywords: ARIMA, Forecasting, NMR.

I. INTRODUCTION

The persistence of war in Yemen has led to massive displacement of thousands of people and left many people in need of humanitarian aid (OCHA, 2019; Al-Mekhlafi, 2018; Burki, 2015). The war has resulted in the destruction of infrastructure and is aggravating existing problems such as poverty, poor health, shortage of basic human needs like water, food and medical supplies (El Bcheraoui *et al*. 2018; Eshaq *et al.* 2017; Qirbi & Ismail, 2017). The crisis in Yemen has negatively impacted on the quality of maternal and child health programs (Eze*et al.* 2020). Therefore it is not surprising for Yemen to report higher maternal and child mortality rates. A neonatal mortality rate of 26 deaths per 1000 live births was reported over the period 2009-2013 (Yemen MOH, 2015). This paper aims to model and forecast neonatal mortality rate for Yemen using the Box-Jenkins ARIMA approach. The model has been found to be useful in modelling linear data sets (Nyoni, 2018; Box-Jenkins, 1970). This surveillance tool can be utilized in public health programming to inform decisions, policy and resource mobilization. Forecast results are expected to assist health authorities in tracking the country's progress towards achieving sustainable development goal 3 target 3.2 by 2030 and trigger implementation of appropriate neonatal interventions to control neonatal deaths across the country.

II. LITERATURE REVIEW

Jawad *et al.* (2021) applied regression analysis to assess the association between conflict and maternal and child health globally. Data for 181 countries (2000–2019) from the Uppsala Conflict Data Program and World Bank were analyzed using panel regression models. The study findings showed that armed conflict is associated with substantial and persistent excess maternal and child deaths globally. Eze *et al*. (2020) conducted a retrospective study in Yemen to examine morbidities & outcomes of a neonatal intensive care unit in a complex humanitarian conflict setting, Hajjah for the period 2017-2018. A 2-year retrospective study of admissions into the Neonatal Intensive Care Unit (NICU) in Al-Gomhoury Hospital Hajjah, Northwest Yemen was conducted. Data was analyzed with IBM SPSS® version 25.0 statistical software using descriptive/inferential statistics. The study findings revealed that preterm newborns bear disproportionate burden of neonatal morbidity and mortality in this setting which is aggravated by difficulties in accessing early neonatal care. A simulation model was constructed by Jenkins *et al*. (2018) to estimate indirect child mortality attributable to war. Yemen was chosen as the example case because indirect child mortality from war likely outpaces direct casualties in the Yemen conflict. A fixed effects panel regression was used to estimate elasticities between child mortality rate (CMR) (the rate of deaths among children under five years of age, per 1,000 live births) and two effects of war assumed to have the greatest explanatory power toward CMR: economic deterioration (measured by changes GDP per capita) and conflict magnitude (via the Major Episodes of Political Violence dataset). These elasticities were then used in a model to estimate the CMR in Yemen up to the year 2020. The study concluded that Yemen's CMR increased by more than 50 per cent from 54.2 in 2010 to 83.9 in 2017.A cross-sectional study in Ghana was carried out by Annan & Asiedu (2018) who applied logistic regression model to assess the maternal, neonatal, and health system related factors that influence

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 519-526, August-2023
https://doi.org/10.47001/IRJIET/2023.708075

neonatal deaths in the Ashanti Region, Ghana. The authors concluded that there was a high number of neonatal deaths which were mainly caused by birth asphyxia, infections, congenital anomalies and respiratory distress syndrome.

III. METHODOLOGY

The Autoregressive (AR) Model

A process X_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:

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

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

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

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:

$$X_t = \emptyset X_{t-1} + Z_t \ldots \ldots [3]$$

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:

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

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

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

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

Equation [4] can also be expressed as follows:

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

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 X_t sequence and recover Z_t from present and past values of X_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 519-526, August-2023

https://doi.org/10.47001/IRJIET/2023.708075

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

Thus:

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

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.

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

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:

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

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

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

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 medicine. 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 Yemen for the period 1962 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

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

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 519-526, August-2023

https://doi.org/10.47001/IRJIET/2023.708075

Criteria Table

Table 2: Criteria Table

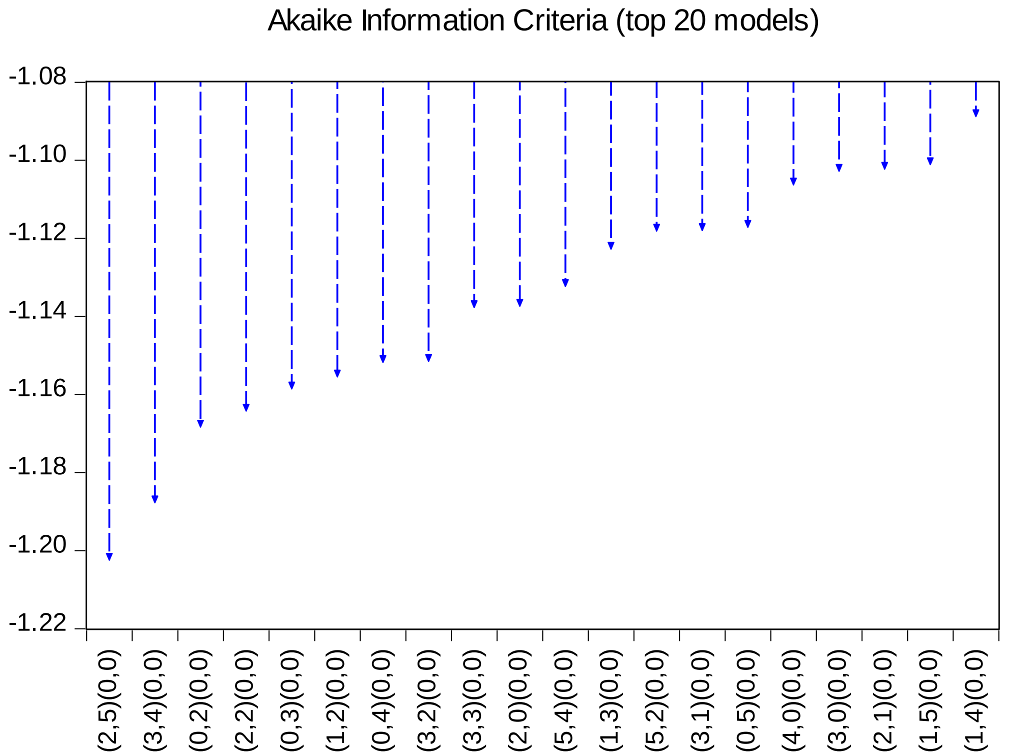
Model Selection Criteria Table
Dependent Variable: D(X, 2)
Date: 01/29/22 Time: 12:32
Sample: 1962 2019
Included observations: 56

Model	LogL	AIC*	BIC	HQ
(2,5)(0,0)	42.641083	-1.201467	-0.875964	-1.075270
(3,4)(0,0)	42.228970	-1.186749	-0.861246	-1.060552
(0,2)(0,0)	36.685708	-1.167347	-1.022679	-1.111259
(2,2)(0,0)	38.571765	-1.163277	-0.946275	-1.079146
(0,3)(0,0)	37.412187	-1.157578	-0.976743	-1.087469
(1,2)(0,0)	37.326955	-1.154534	-0.973699	-1.084425
(0,4)(0,0)	38.223245	-1.150830	-0.933828	-1.066699
(3,2)(0,0)	39.214598	-1.150521	-0.897352	-1.052368
(3,3)(0,0)	39.829682	-1.136774	-0.847438	-1.024599
(2,0)(0,0)	35.819603	-1.136414	-0.991746	-1.080327
(5,4)(0,0)	42.677135	-1.131326	-0.733489	-0.977086
(1,3)(0,0)	37.410324	-1.121797	-0.904795	-1.037666
(5,2)(0,0)	40.280321	-1.117154	-0.791651	-0.990958
(3,1)(0,0)	37.275455	-1.116981	-0.899979	-1.032849
(0,5)(0,0)	38.252316	-1.116154	-0.862985	-1.018001
(4,0)(0,0)	36.948371	-1.105299	-0.888297	-1.021168
(3,0)(0,0)	35.851167	-1.101827	-0.920992	-1.031718
(2,1)(0,0)	35.835405	-1.101264	-0.920429	-1.031155
(1,5)(0,0)	38.804349	-1.100155	-0.810819	-0.987980
(1,4)(0,0)	37.458167	-1.087792	-0.834623	-0.989639
(2,3)(0,0)	37.429458	-1.086766	-0.833597	-0.988613
(5,0)(0,0)	37.235619	-1.079844	-0.826675	-0.981690
(4,1)(0,0)	37.080737	-1.074312	-0.821143	-0.976159
(4,2)(0,0)	38.078033	-1.074215	-0.784880	-0.962041
(5,3)(0,0)	39.913486	-1.068339	-0.706669	-0.928120
(4,3)(0,0)	38.742531	-1.062233	-0.736730	-0.936036
(2,4)(0,0)	37.626422	-1.058087	-0.768751	-0.945912
(1,1)(0,0)	33.429299	-1.051046	-0.906378	-0.994959
(4,4)(0,0)	39.390517	-1.049661	-0.687991	-0.909443
(5,1)(0,0)	37.265429	-1.045194	-0.755858	-0.933019
(3,5)(0,0)	38.806903	-1.028818	-0.667148	-0.888599
(4,5)(0,0)	39.410870	-1.014674	-0.616837	-0.860433
(5,5)(0,0)	40.070613	-1.002522	-0.568518	-0.834259
(1,0)(0,0)	30.930557	-0.997520	-0.889019	-0.955454
(0,1)(0,0)	29.217484	-0.936339	-0.827838	-0.894273
(0,0)(0,0)	27.197959	-0.899927	-0.827593	-0.871883

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 7, Issue 8, pp 519-526, August-2023
https://doi.org/10.47001/IRJIET/2023.708075

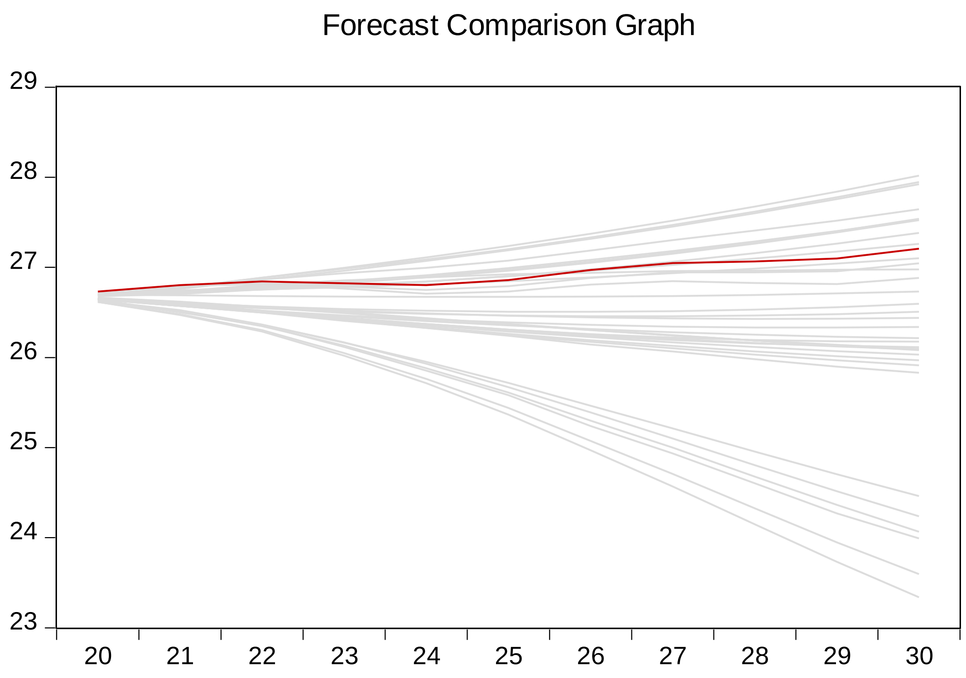
Criteria Graph

Figure 1: Criteria Graph



Forecast Comparison Graph

Figure 2: Forecast Comparison Graph



ISSN (online): 2581-3048
Volume 7, Issue 8, pp 519-526, August-2023
https://doi.org/10.47001/IRJIET/2023.708075

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

IV. RESULTS

Summary of the Selected ARIMA () Model

Table 3: Summary of the Optimal Model

Automatic ARIMA Forecasting
Selected dependent variable: D(X, 2)
Date: 01/29/22 Time: 12:32
Sample: 1962 2019
Included observations: 56
Forecast length: 11

Number of estimated ARMA models: 36
Number of non-converged estimations: 0
Selected ARMA model: (2,5)(0,0)
AIC value: -1.20146726577

Main Results of the Selected ARIMA () Model

Table 4: Main Results of the Optimal Model

Dependent Variable: D(X,2)
Method: ARMA Maximum Likelihood (BFGS)
Date: 01/29/22 Time: 12:32
Sample: 1964 2019
Included observations: 56
Convergence achieved after 98 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.008764	0.042700	0.205245	0.8383
AR(1)	0.551575	0.115189	4.788421	0.0000
AR(2)	-0.967222	0.089274	-10.83437	0.0000
MA(1)	-0.388431	36.17440	-0.010738	0.9915
MA(2)	1.720911	81.54925	0.021103	0.9833
MA(3)	0.053460	34.76188	0.001538	0.9988
MA(4)	0.663181	66.03332	0.010043	0.9920
MA(5)	0.417773	32.68041	0.012784	0.9899
SIGMASQ	0.010559	1.112464	0.009491	0.9925

R-squared	0.523639	Mean dependent var	0.012500
Adjusted R-squared	0.442556	S.D. dependent var	0.150227
S.E. of regression	0.112163	Akaike info criterion	-1.201467
Sum squared resid	0.591283	Schwarz criterion	-0.875964
Log likelihood	42.64108	Hannan-Quinn criter.	-1.075270
F-statistic	6.458087	Durbin-Watson stat	1.970266
Prob(F-statistic)	0.000012		

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 519-526, August-2023
https://doi.org/10.47001/IRJIET/2023.708075

Inverted AR Roots	.28+.94i	.28-.94i		
Inverted MA Roots	.36+.93i	.36-.93i	.04+1.00i	.04-1.00i
	-.42			
	Estimated MA process is noninvertible			

ARIMA () Model Forecast

Tabulated Out of Sample Forecasts

Table 5: Tabulated Out of Sample Forecasts

2020	26.73097229351558
2021	26.80136152414103
2022	26.84490701067958
2023	26.82219569588329
2024	26.80193197144978
2025	26.85950995123996
2026	26.97006275324609
2027	27.04695158546192
2028	27.06644044526294
2029	27.09923608532172
2030	27.20729656868086

Table 5 clearly indicates that neonatal mortality is expected to slightly increase and remain high throughout the out of sample period.

V. POLICY IMPLICATION & CONCLUSION

The conflict in Yemen has impacted negatively on the country's progress towards achieving set sustainable development goal targets by the end of 2030. The continued destruction of health infrastructure and exodus of health professionals is crippling health service delivery. In this study we proposed the ARIMA model to forecast NMR for Yemen and the findings suggest that neonatal mortality will slightly increase and remain high throughout the out of sample period. Therefore, it is crucial for the authorities in Yemen to design and implement country specific neonatal policies to urgently address causes of neonatal deaths across the country. Strategies should focus on improving the quality, affordability and accessibility of maternal and neonatal healthcare services.

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] Burki T (2015). Yemen health situation "moving from a crisis to a disaster". Lancet, 385, 9978, 1609.

[4] Office for the Coordination of Humanitarian Affairs (OCHA) (2019). Yemen: 2019 Humanitarian Needs Overview, 2019

[5] Al-Mekhlafi H. M (2018). Perspective piece: Yemen in a time of cholera: current situation and challenges. Am J Trop Med Hyg, 98, 6, 1558–62

[6] El Bcheraoui C., Jumaan A. O., Collison ML., Daoud F., and Mokdad A. H (2018). Health in Yemen: losing ground in war time. Glob Health, 14, 1, 42.

[7] Eshaq A. M., Fothan A. M., Jensen E. C., Khan T. A., and AlAmodi A. A (2017). Malnutrition in Yemen: an invisible crisis. Lancet, 389, 10064, 31–2.

[8] Qirbi N., and Ismail S. A (2017). Health system functionality in a low-income country in the midst of conflict: the case of Yemen. Health Policy Plan, 32, 6, 911– 22.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 519-526, August-2023
https://doi.org/10.47001/IRJIET/2023.708075

[9] Yemen MOH (2015). Republic of Yemen Ministry of Public Health & Population and Central Statistical Organization. Yemen National Health & Demographic Survey 2013.

[10] Eze P., Al-Maktari F., Alshehari A.H and LawaniL. O (2020). Morbidities & outcomes of a neonatal intensive care unit in a complex humanitarian conflict setting, Hajjah Yemen: 2017-2018, Conflict and Health, 2020, 14, 53.

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Calculating Expected Future Values of Annual Neonatal Mortality Rate for Yemen Using the ARIMA Model" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 519-526, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708075

****