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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 316-320, March-2021

https://doi.org/10.47001/IRJIET/2021.503054

Forecasting TB Incidence in Yemen Using the Multilayer Perceptron Neural Network

[1]Dr. Smartson. P. NYONI, [2]Thabani NYONI

[1]ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe
[2]Department of Economics, University of Zimbabwe, Harare, Zimbabwe

***Abstract* - In this research paper, the ANN approach was applied to analyze TB incidence in Yemen. The employed data covers the period 2000-2018 and the out-of-sample period ranges over the period 2019-2023. The residuals and forecast evaluation criteria (Error, MSE and MAE) of the applied model indicate that the model is stable in forecasting TB incidence in Yemen. The results of the study indicate that TB incidence will be around 48 cases/100 000 /year over the period 2019-2023. In order to contribute meaningfully to the national control strategy of a TB-free Yemen, authorities should, among other things, intensify TB surveillance and control programmes in order to reduce incidence to below 30 cases per 100 000/year.**

Keywords: ANN, Forecasting, TB incidence.

I. INTRODUCTION

Modelling and forecasting in the medical field is very important as a tool for early surveillance and control of infectious diseases such as Tuberculosis. There are many studies that have been done in order to understand the evolution of the TB epidemic (Liao et al, 2019; Carvajal, 2018; Mao et al, 2018; Anokye, 2018; Withanage et al, 2018; Siregar et al, 2018; Tohidinik et al, 2018; Xu et al, 2018; Wang et al, 2018; Box et al, 2015; Box G et al, 2010). TB is one of the biggest health problems in the World especially in poor and low income countries. It is an infectious disease caused by mycobacterium tuberculosis and spreads through air droplets by sneezing and coughing of the infected person (WHO, 2018). Around 10.4 million people were infected with TB in 2016 with a high mortality of 2million deaths (Azeez et al, 2016; WHO, 2016). The disease occurs via two mechanisms: transmission and reactivation of latent TB infection. Many studies in the past have revealed seasonal variation in TB incidence. Some studies found out high TB case notification in spring and summer (Willis et al, 2012; Khaliq et al, 2015; Naranbart et al, 2009; Wubuli et al, 2017)

Several studies have been conducted to model and forecast TB incidence. Mohammed S.H et al (2018) modelled the monthly incidence rates of PTB cases in Kubala, Iraq. The study utilized data covering the period January 2010 to December 2016.The SARIMA, SARIMA-ETS, SARIMA-NNAR and SARIMA-ANFIS were applied. The study concluded that there is a seasonal characteristic of pulmonary TB incidence with peaks during spring and winter. All models predicted a marginal decrease in TB trends over the period 2016-2018.The SARIMA-ANFIS the best model. Khaliq et al (2015) conducted a study to find out the temporal and seasonal pattern of TB incidence in Lahore, Pakistan from 2006 to 2013 in newly diagnosed pulmonary TB cases. The SPSS version 21 software was used for correlation to determine the temporal relationship and time series analysis for seasonal variation. The study revealed that there is a significant temporal relation with a seasonal pattern and declining trend in the TB incidence. In another study, Klotz et al (2013) modelled and forecasted the TB incidence in Province of Quebec, China using the SELIR compartmental model. The results from the study showed that the incidence of TB in the Province of Quebec was expected to decrease overall, certain populations will remain at risk. In this paper we applied the artificial neural network method to model and forecast the annual TB incidence in Yemen. The findings from this study will help the Yemen government to assess the impact of TB prevention and control measures. This will then facilitate the implementation of appropriate corrective measures by the National response team.

II. METHOD

The Artificial Neural Network (ANN), which we apply; is a data processing system consisting of a large number of simple and highly interconnected processing elements resembling a biological neural system. It has the capability of learning from an experimental or real data set to describe the nonlinear and interaction effects with great accuracy. ANN-based curve fitting technique is one of the most applied artificial intelligence methods that are used for forecasting and prediction purpose. It

© 2021-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 5, Issue 3, pp 316-320, March-2021
https://doi.org/10.47001/IRJIET/2021.503054

consists of basically three layers, that is., input layer, hidden layer, and output layer, the present work includes the number of years as input layer and the annual TB incidence in Yemen as output data for the network. In this study, our ANN is based on the hyperbolic tangent function.

Data Issues

This study is based on TB incidences (referred to as Y series in this study) in all age groups at GPH. The annual data covers the period 2000-2018 while the out-of-sample forecast covers the period 2019-2023. All the data employed in this research paper was gathered from the World Bank online database.

III. FINDINGS OF THE STUDY

DESCRIPTIVE STATISTICS

Table 1: Descriptive statistics

Mean	Median	Minimum	Maximum
67.474	55.000	48.000	116.00
Std. Dev.	C.V.	Skewness	Ex. kurtosis
23.682	0.35098	0.85814	-0.73158
5% Perc.	95% Perc.	IQ range	Missing obs.
Undefined	116.00	39.000	0

TB incidence, on average has been around 67 cases per 100 000 people, annually; in the country.

ANN MODEL SUMMARY FOR TB INCIDENCE (new cases per 100 000 population/year) IN YEMEN

Table 2: ANN model summary

Variable	Y
Observations	10 (After Adjusting Endpoints)
Neural Network Architecture:	
Input Layer Neurons	9
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Function	Hyperbolic Tangent Function
Back Propagation Learning:	
Learning Rate	0.005
Momentum	0.05
Criteria:	
Error	0.008227
MSE	0.096590
MAE	0.189937

Residual Analysis for the ANN model

© 2021-2017 IRJIET All Rights Reserved

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 5, Issue 3, pp 316-320, March-2021
https://doi.org/10.47001/IRJIET/2021.503054

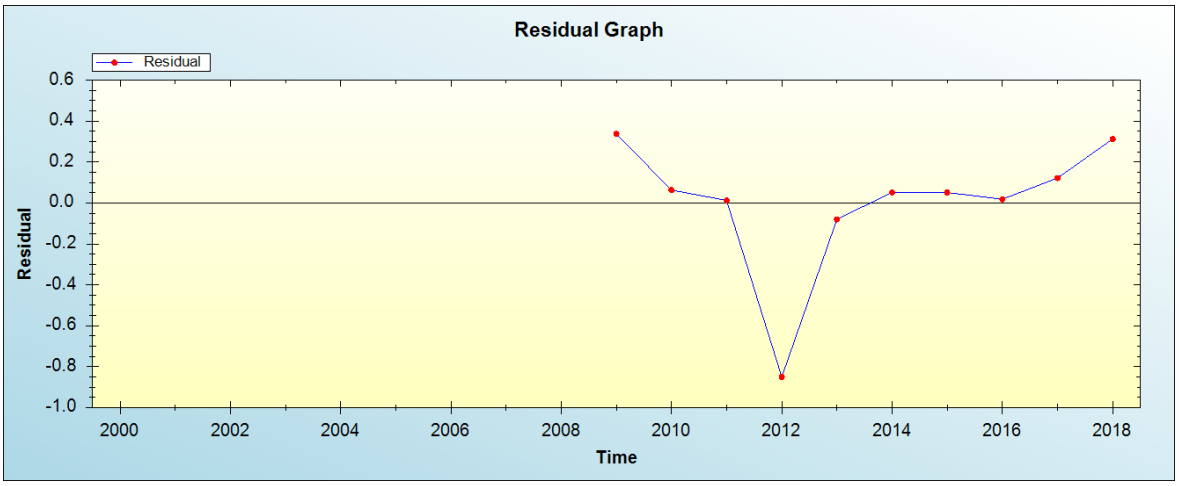

Figure 1: Residual analysis

In-sample Forecast for Y

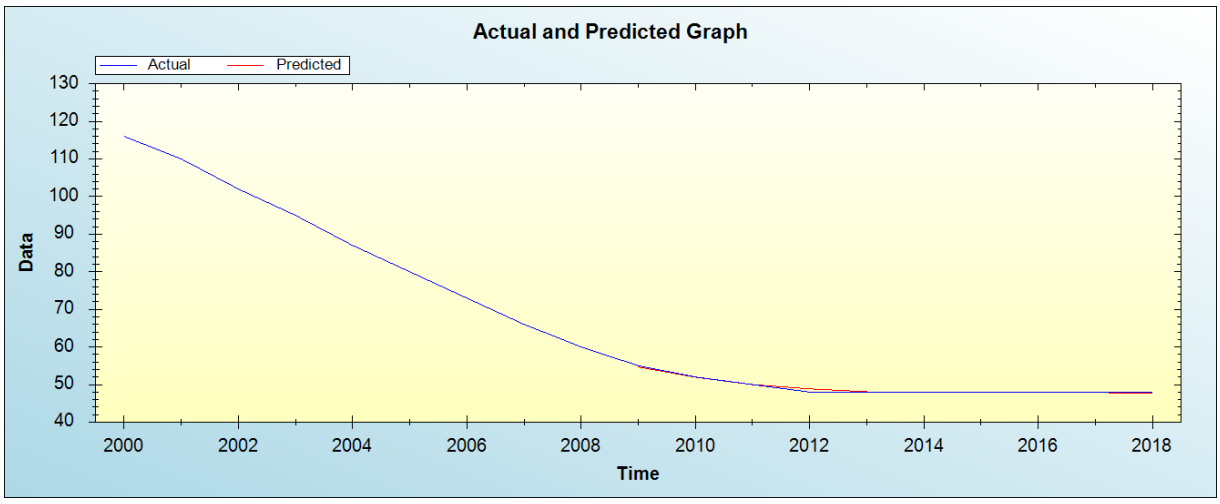

Figure 2: In-sample forecast for the Y series

Figure 2 shows the in-sample forecast for Y series.

Out-of-Sample Forecast for Y: Actual and Forecasted Graph

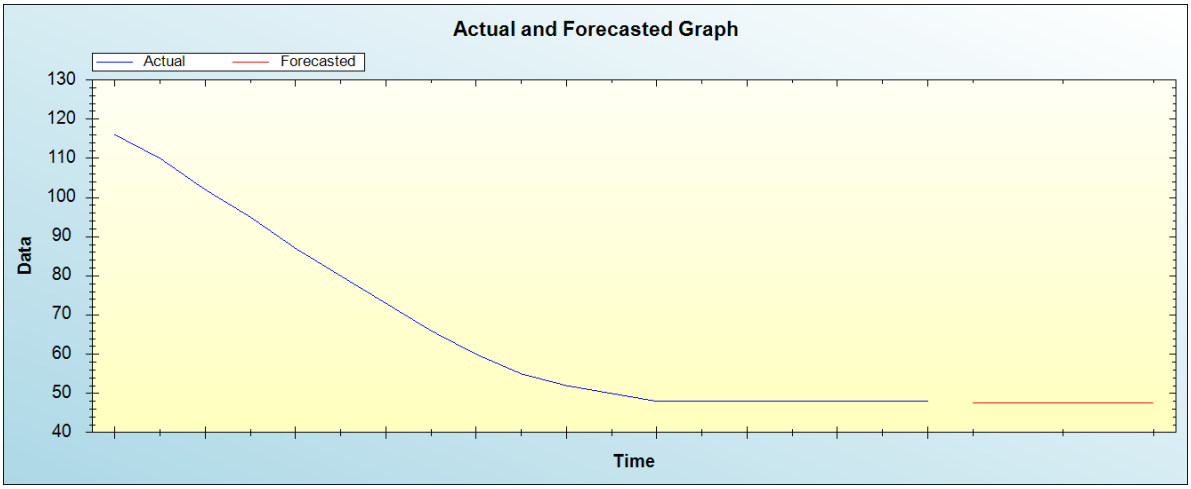

Figure 3: Out-of-sample forecast for Y: actual and forecasted graph

© 2021-2017 IRJIET All Rights Reserved

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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 316-320, March-2021

https://doi.org/10.47001/IRJIET/2021.503054

Out-of-Sample Forecast for Y: Forecasts only

Table 3: Tabulated out-of-sample forecasts

Year	Forecasts
2019	47.6296
2020	47.6652
2021	47.5406
2022	47.5634
2023	47.5999

Table 1 shows that over the study period 2000-2018 the minimum and maximum TB incidence was 67 and 116 cases per 100 000/year respectively. The applied data is positively skewed with an excess kurtosis value of -0.73158 meaning that it is not normally distributed. The residual graph and model evaluation criteria indicate that the model is stable and suitable for forecasting TB incidence in Yemen. Figure 2 shows that the applied model simulates the observed values very well. The model predicts that the TB incidence will be constant with a value around 48 cases per 100 000/year throughout the period 2019-2023.

IV. CONCLUSION & RECOMMENDATIONS

The government of Yemen has done great work in its fight against the TB epidemic. The country recorded a significant decrease in the incidence of TB over the period 2000-2018 following implementation of robust measures in order to control the spread of TB among the communities. The model predicts that over the period 2019-2023 the incidence will be around 48 cases per 100 000/year. We encourage the authorities to intensify TB surveillance and control programs in order to reduce incidence to levels below 30 cases/100 000 population/year.

REFERENCES

[1] Anokye R., Acheampong E., & Owusu (2018). Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. Cogent Soc Sci. ;4(1): 1-9.

[2] Azeez A., Obaromi D., Odeyemi A., Ndege J & Muntabayi R (2016). Seasonality and trend forecasting of tuberculosis prevalence data in Eastern Cape, South Africa, using a hybrid model. Int J Environ Res Public Health;13. pii: E757.

[3] Box GEP, Jenkins GM & Reinsel (2015). Time series analysis: forecasting and control, 5th edition. J Oper Res Soc; 22(2): 199–201.

[4] Box., George E.P., Jenkins., Gwilym M, Reinsel & Gregory C (2010). Time series analysis. Forecasting and control. 3rd ed. journal of time. 2010; 31(4):303 - 309.

[5] Carvajal Thaddeus M., Viacrusis Katherine M & Hernandez Lara Fides T (2018). Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines. BMC Infect Dis. 2018; 18: 183-189.

[6] Khaliq A., Batool SA., &Chaudhry MN (2015). Seasonality and trend analysis of tuberculosis in Lahore, Pakistan from 2006 to 2013. J Epidemiol Glob Health 5: 397-403.

[7] Liao Z., Zhang X &Zhang Y (2019). Seasonality and Trend Forecasting of Tuberculosis Incidence in Chongqing, China. Inter discip Sci; 11(1): 77–85.

[8] Mao Q., Zhang K &Yan W (2018). Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. J Infect Public Health;11(5): 707–12.

[9] Naranbat N., Nymadawa P., Schopfer K & Rieder HL (2009). Seasonality of tuberculosis in an Eastern Asian country with an extreme continental climate. Eur Respir J, 34:921-925.

[10] Siregar FA., Makmur T & Saprin S (2018). Forecasting dengue hemorrhagic fever cases using ARIMA model: a case study in Asahan district. In: IOP Conference Series Materials Science and Engineering; p. 300.

[11] Tohidinik HR., Mohebali M., Mansournia MA (2018). Forecasting zoonotic cutaneous leishmaniasis using meteorological factors in eastern Fars province, Iran: a SARIMA analysis. Tropical Med Int Health, 23(8): 860–9.

[12] Wang H., Tian CW., Wang WM (2018). Time-series analysis of tuberculosis from 2005 to 2017 in China. Epidemiol Infect. 2018;146(8):935–9.

[13] WHO (2016). Global Tuberculosis Report 2016.Geneva, Switzerland.

[14] Willis MD., Winston CA., Heilig CM., Cain KP., Walter ND & Mac Kenzie WR (2012). Seasonality of tuberculosis in the United States, 1993-2008. Clin Infect Dis; 54:1553-1560.

© 2021-2017 IRJIET All Rights Reserved

ISSN (online): 2581-3048
Volume 5, Issue 3, pp 316-320, March-2021
https://doi.org/10.47001/IRJIET/2021.503054

[15] Withanage GP., Viswakula SD., Yi SGN (2018). A forecasting model for dengue incidence in the district of Gampaha, Sri Lanka.

[16] Wubuli A., Li Y., Xue F., Yao X., Upur H &Wushouer Q (2017). Seasonality of active tuberculosis notification from 2005 to 2014 in Xinjiang, China. PLoS One, 12: e0180226.

[17] Xu Q., Li R., Liu Y (2017). Forecasting the incidence of mumps in Zibo City based on a SARIMA model. Int J Environ Res Public Health;14(8):925 937.

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Forecasting TB Incidence in Yemen Using the Multilayer Perceptron Neural Network" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 5, Issue 3, pp 316-320, March 2021. Article DOI https://doi.org/10.47001/IRJIET/2021.503054

****

© 2021-2017 IRJIET All Rights Reserved