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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 311-315, March-2021

https://doi.org/10.47001/IRJIET/2021.503053

Modeling and Forecasting TB Incidence in Tanzania: The Artificial Neural Network Approach

[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 article, the ANN approach was applied to analyze TB incidence in Tanzania. The employed annual data covers the period January 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 Tanzania. The results of the study indicate that TB incidence will sharply rise from 239 cases/100 000 /year in 2019 to 456 cases/100 000/year in 2020 and then decline sharply to around 239 cases per 100 000/year in 2023. In order to contribute meaningfully to the national control strategy of a TB-free Tanzania, the government should, among other things, intensify TB surveillance and control programs.**

Keywords: ANN, Forecasting, TB incidence.

I. INTRODUCTION

Modelling and forecasting in public health is very critical for planning and allocation of resources. Several models have been applied by researchers and these include Box-Jenkins models, machine learning methods, exponential smoothing and hybrid models. Box-Jenkins Seasonal autoregressive integrated moving average (SARIMA) models are widely used in the field of medicine because of their simplicity. The SARIMA model was proposed by Box and Jenkins in the 1970s (George &Jenkins, 1976). The general form of the SARIMA model is as follows:(p,d,q)(P,D,Q)s, where p and q are the autoregressive (AR) and moving average (MA) parts respectively .d is the order of non-seasonal differences. P, D and Q are the corresponding seasonal orders and s represents the steps of the seasonal differences (Nyoni & Nyoni, 2019 (a) &(b)). The SARIMA model assigns the combined effects of multiple risk factors that affect the disease occurrence and prevalence time (Li et al, 2019).This model has become one of the most popular models in time series forecasting. It has been applied in prediction of infectious diseases like malaria, influenza, hepatitis, hand foot mouth disease and other diseases (Wangdi et al, 2010; Ramirez et al, 2011; Briet et al, 2008; Anwar et al, 2016; N'Gatta et al, 2016). The SARIMA model has been previously used to model and forecast the incidence of TB (Permanasari et al, 2011; Li et al, 2013; Wubuli et al, 2017). In recent times artificial neural networks (ANNs) are dominating in the field of time series forecasting (Zhang, 2003). The main attribute of ANNs when applied to time series forecasting problems is their inherent capability of nonlinear modeling without assumption about the statistical distribution followed by the observations. The ideal model is developed based on the given data (Zhang, 2003). The most popular ANN model is the multilayer perceptrons (MLPs) which are characterized by a single hidden layer Feed Forward Network (FNN) (Hamzacebi, 2008; Zhang, 2003).

In this paper the researchers applied the artificial neural network, ANN (9,12,1) model to forecast the annual incidence of TB in Tanzania. The more efficient hyperbolic tangent activation function was chosen. The results of the study are expected to reveal the future trends of TB incidence in the country thereby facilitating planning and enough allocation of resources to the National TB/HIV program.

II. RELATED STUDIES

Nyoni & Nyoni (2019(a)) developed a SARIMA model to forecast TB notifications at Zengeza clinic, Zimbabwe. The study utilized monthly TB notification data covering the period January 2013 to December 2018.The best model, SARIMA (2,0,2)(1,0,1)12 predicted that TB notifications would decline over the out of sample period. In a similar study, Nyoni &Nyoni (2019(b)) modeled and forecasted TB notifications at Silobela District hospital in Zimbabwe. The developed SARIMA (1,0,1) (0,1,1)12 model also suggested that TB notifications would decline in the out of sample period. Patowary & Barman (2017) developed a SARIMA model to forecast TB detection rate in Dibrugarh district Assam, India. The SARIMA (0,0,0) (1,1,0)4 model was found to be the suitable model for the given data set. Kakchapati (2010) applied a negative binomial model to analyse TB incidence in

ISSN (online): 2581-3048
Volume 5, Issue 3, pp 311-315, March-2021
https://doi.org/10.47001/IRJIET/2021.503053

Nepal. The study concluded that TB incidence showed a decreasing trend but the number of cases still very high and gender differences existed in TB incidence in Nepal.

III. METHOD

The Artificial Neural Network (ANN), which we intend to 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 extensively applied artificial intelligence methods that are used for forecasting and prediction purpose. It consists of basically three layers i.e., input layer, hidden layer, and output layer, the present work includes the number of years as input layer and the annual TB incidence in Tanzania as output data for the network. In this paper, our ANN is based on the hyperbolic tangent function.

Data Issues

This study is based on TB incidences (referred to as TB series in this study) in Tanzania. 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.

IV. FINDINGS OF THE STUDY

DESCRIPTIVE STATISTICS

Table 1: Descriptive statistics

Mean	Median	Minimum	Maximum
423.84	452.00	253.00	510.00
Std. Dev.	C.V.	Skewness	Ex. Kurtosis
89.767	0.21179	-0.59280	-1.0618
5% Perc.	95% Perc.	IQ range	Missing obs.
Undefined	510.00	170.00	0

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

Table 2: ANN model summary

Variable	TB
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	0.038452
MSE	30.141808
MAE	4.514835

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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 311-315, March-2021

https://doi.org/10.47001/IRJIET/2021.503053

Residual Analysis for the ANN model

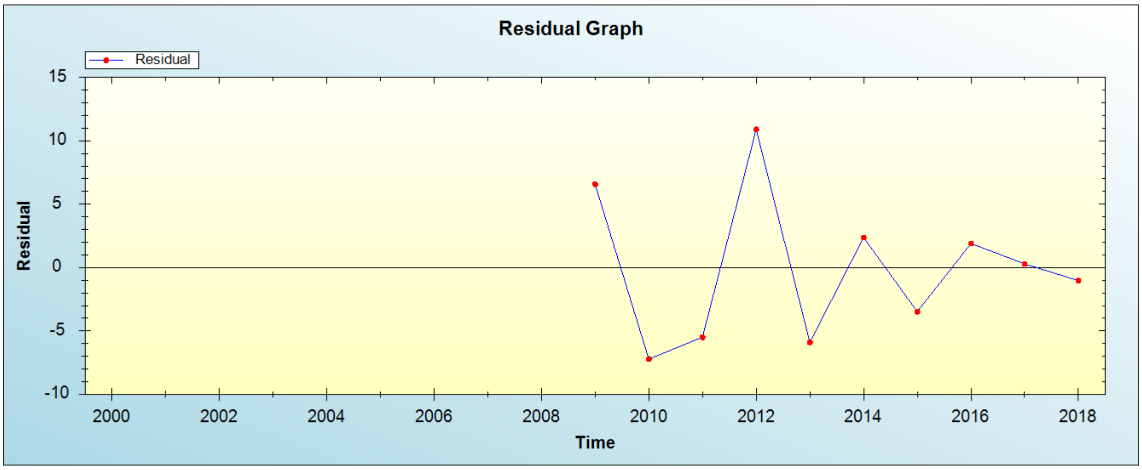

Figure 1: Residual analysis

In-sample Forecast for TB

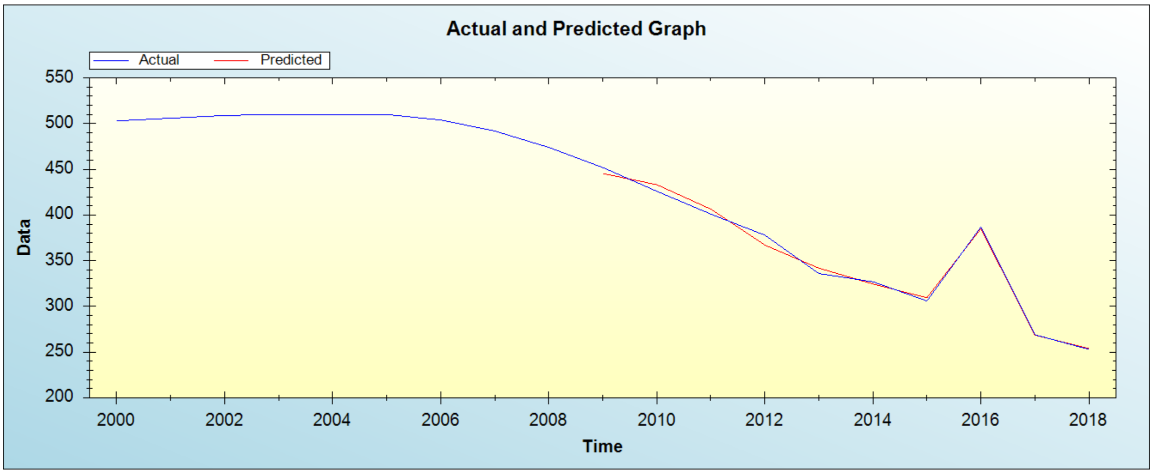

Figure 2: In-sample forecast for the TB series

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

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

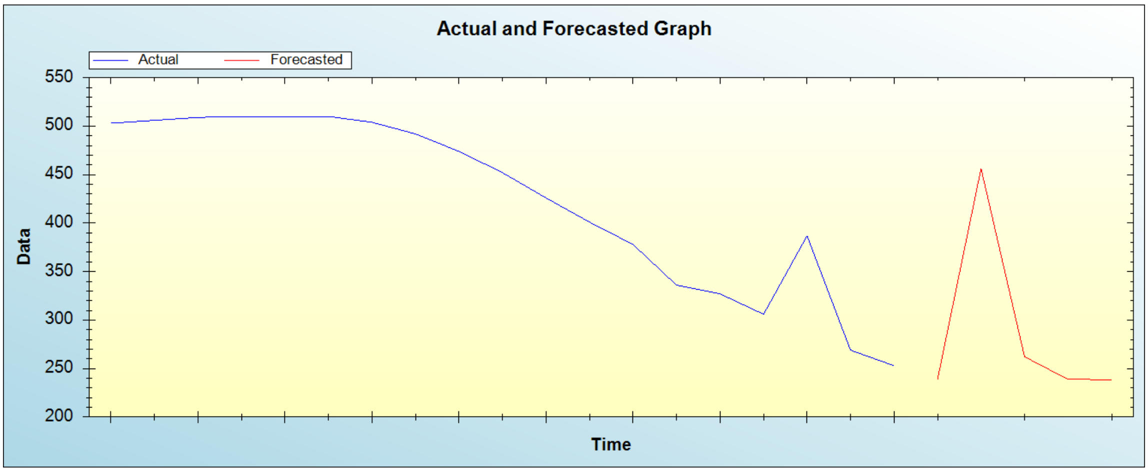

Figure 3: Out-of-sample forecast for TB: actual and forecasted graph

ISSN (online): 2581-3048
Volume 5, Issue 3, pp 311-315, March-2021
https://doi.org/10.47001/IRJIET/2021.503053

Out-of-Sample Forecast for TB: Forecasts only

Table 3: Tabulated out-of-sample forecasts

Year	Forecasts
2019	238.7668
2020	456.2800
2021	262.2107
2022	238.8056
2023	238.7223

Over the study period 2000-2018 the country experienced a gradual decline in TB incidence. The minimum and maximum incidence was 253 and 510 cases per 100 000 population/year. The applied data is negatively skewed and with an excess kurtosis of -1.0618 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 Tanzania. Figure 2 shows that the applied ANN (9,12,1) for Tanzania's TB incidence simulates the observed data very well. The model predicts that TB incidence will sharply rise from 231 cases per 100 000 population/year in 2019 up to 456 cases per 100 000/year in 2020 and then decline sharply to 238 cases per 100 000 /year by 2023.

V. CONCLUSION & RECOMMENDATIONS

Tanzania has made great strides in the fight against TB epidemic. The country recorded a significant drop in TB incidence between 2006 and 2018 as a result of effective TB control measures implemented by the government. However, the model predicts a sharp increase in TB incidence in 2020 and sharp decline in 2022 and 2023. Although the nation is facing limited funding for the HIV/TB program, the authorities are obliged to prioritize TB/HIV programs, intensify TB surveillance and control amongst other measures.

REFERENCES

[1] Anwar MY, Lewnard JA, Parikh S & Pitzer VE (2016). Time series analysis of malaria in Afghanistan: using ARIMA models to predict future trends in incidence. Malar J.;15(1):566. doi:10.1186/s12936-016-1602-1

[2] Briet OJ., Amerasinghe PH., &Vounatsou P (2013). Generalized seasonal autoregressive integrated moving average models for count data with application to malaria time series with low case numbers. PLoS One;8(6): e65761. doi: 10.1371/journal.pone.0065761

[3] Briet OJ., Vounatsou P., Gunawardena DM., Galappaththy GN &Amerasinghe PH (2008). Models for short term malaria prediction in Sri Lanka. Malar J;7(1):76. doi:10.1186/1475-2875-7-129

[4] G.E.P. Box, G. Jenkins, "Time Series Analysis, Forecasting and Control", Holden-Day, San Francisco, CA, 1970.

[5] Li XX., Wang LX., & Zhang H (2013). Seasonal variations in notification of active tuberculosis cases in China, 2005–2012. PLoS One;8 (7): e68102. doi: 10.1371/journal.pone.0068102

[6] N 'Gattia AK., Coulibaly D & Nzussouo NT (2016). Effects of climatological parameters in modeling and forecasting seasonal influenza transmission in Abidjan, Cote d 'Ivoire. BMC Public Health;16(1):972. doi:10.1186/s12889-016-3503-1

[7] Permanasari AE., Rambli DR., & Dominic PD. Performance of univariate forecasting on seasonal diseases: the case of tuberculosis. Adv Exp Med Biol;696(696):171–179. doi:10.1007/978-1-44197046-6_17

[8] Ramirez AP., Buitrago JI., Gonzalez JP., Morales AH & Carrasquilla G (2014). Frequency and tendency of malaria in Colombia, 1990 to 2011: a descriptive study. Malar J;13(1):202. doi:10.1186/14752875-13-202

[9] Wangdi K., Singhasivanon P., Silawan T., Lawpoolsri S., White NJ &Kaewkungwal J (2010). Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan. Malar J. 2010; 9(1): 251. doi:10.1186/1475-2875-9-251

[10] Wubuli A., Li Y & Xue F (2017). Seasonality of active tuberculosis notification from 2005 to 2014 in Xinjiang, China. PLoS One;12(7): e0180226. doi: 10.1371/journal.pone.0180226

[11] Zhang GP, "Time series forecasting using a hybrid ARIMA and neural network model", Neurocomputing 50 (2003), pages: 159–175.

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 311-315, March-2021

https://doi.org/10.47001/IRJIET/2021.503053

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Modeling and Forecasting TB Incidence in Tanzania: The Artificial Neural Network Approach" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 5, Issue 3, pp 311-315, March 2021. Article DOI https://doi.org/10.47001/IRJIET/2021.503053

****