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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 371-375, March-2021

https://doi.org/10.47001/IRJIET/2021.503064

Prediction of the Annual TB Incidence in Niger Using Artificial Neural Networks

[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 Niger. The employed annual 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 Niger. The model predictions suggest that TB incidence will continue to decline over the period 2019-2023. Therefore the government is encouraged to intensify TB surveillance and TB control programs among other measures.**

Keywords: ANN, Forecasting, TB incidence.

I. INTRODUCTION

Prediction of infectious diseases is very important in order to understand the future trends. Accurate and reliable forecasts enable proper planning and adequate allocation of resources for health. In recent decades predictive modeling has emerged as one of the pillars in public health and therefore it is imperative in this paper to briefly discuss the trending models which are being used in infectious disease modeling globally. Public health specialists have applied various models which include compartmental epidemiological models, Autoregressive integrated moving average (ARIMA), machine learning methods, exponential smoothing and hybrid models. In this paper, the researchers will focus on machine learning methods and ARIMA models. Machine learning is the field of study in which computers perform specific tasks without being explicitly programmed or in which computers learn through experience .There are many machine learning algorithms namely artificial neural networks ,support vector machine(SVM),decision trees, ensembles and graphical models. These models can be classified into two classes i.e supervised and unsupervised learning algorithms. The support vector machine was proposed by Vapnik and his co-workers at the A. T & T. Bell laboratories in 1995 (Cao & Francis, 2003; Farooq et al, 2007; Raicharoen et al, 2003; Vapnik, 1998). The support vector machine is based on the structural risk minimization principle (SRM). The idea is to find an optimal separating hyperplane through a nonlinear mapping of the input space into a higher dimension feature space (H). The training process is equivalent to solving a linearly constrained quadratic optimization problem (Cao & Francis, 2003).

Another important machine learning algorithm is the artificial neural network. The algorithm is biologically inspired and mimics the function of the brain. The most widely used ANNs is the the multilayer perceptron (MLPs) which use a single hidden layer Feed Forward Network (FNN) (Zhang et al, 2003). The model is made up 3 layers; input, hidden and output layers which are connected by acyclic links called connection weights. The nodes in different layers are called processing elements (Kaushik & Sahi, 2018). The ARIMA model is a popular stochastic series .It assumes that the time series under consideration is linear and follows a known statistical distribution. The subclasses of the model are the Autoregressive (AR), Moving average (MA) and the ARMA model. For the seasonal prediction the SARIMA model was then proposed by Box and Jenkins. The basic SARIMA model is represented as follows :(p,d,q) (P,D,Q)s .p and q are the autoregressive(AR) and moving average(MA) parts, d is the non -seasonal differencing order. P and Q are the seasonal AR and MA parts, D are the seasonal differences .S represents the steps of the seasonal differences (Yan et al, 2019; Nyoni & Nyoni, 2019(a); Nyoni & Nyoni, 2019(b)). The Box-Jenkins methodology is an ARIMA model building procedure and comprised of 3 steps which are model identification, parameter estimation and diagnostic checking. When an optimal model is identified through an iterative process forecasting can be performed so that control measures are implemented guided by the forecast results.

II. LITERATURE REVIEW

Yan et al (2019) developed an ARIMA model to forecast the incidence of TB in China .Monthly TB incidence data covering the period January 2005-December 2017 was analyzed .The study concluded that the SARIMA $(0,1,2)(0,1,0)_{12}$ was the best model and the model predicted well the incidence of TB and thus the model can be used for short term prediction and dynamic analysis

ISSN (online): 2581-3048
Volume 5, Issue 3, pp 371-375, March-2021
https://doi.org/10.47001/IRJIET/2021.503064

of TB in China. Nyoni & Nyoni (2019 (a)) developed a SARIMA model to forecast TB notifications at Zengeza clinic, Zimbabwe. The best model SARIMA (2,0,2) (1,0,1)12 predicted that TB notifications would decline over the out of sample period .In related study Nyoni & Nyoni (2019 (b)) forecasted monthly TB notifications at Silobela District Hospital in Zimbabwe .Monthly TB notifications recorded over the period January 2014 to December 2018 was analaysed.The SARIMA (1,0,1) $(0,1,1)_{12}$ model predicted that TB notifications would generally decline over the out of sample period .Kusimo &Ladipo (2015) Investigated the seasonality in TB notification in Nigeria. A trend analysis of annual TB notifications from 2004 to 2013 was done. The study results suggested that there seasonal variations in TB notifications which has a significant impact in the implementation of TB control strategies.

III. METHOD

The Artificial Neural Network (ANN), which we intend to empirically employ; 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 Niger 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 F series in this study) in Niger. 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
125.74	118.00	87.000	191.00
Std. Dev.	C.V.	Skewness	Ex. kurtosis
31.717	0.25225	0.61716	-0.78422
5% Perc.	95% Perc.	IQ range	Missing obs.
Undefined	191.00	53.000	0

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

Table 2: ANN model summary

Variable	P
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	0.012034
MSE	0.483432
MAE	0.494056

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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 371-375, March-2021

https://doi.org/10.47001/IRJIET/2021.503064

Residual Analysis for the ANN model

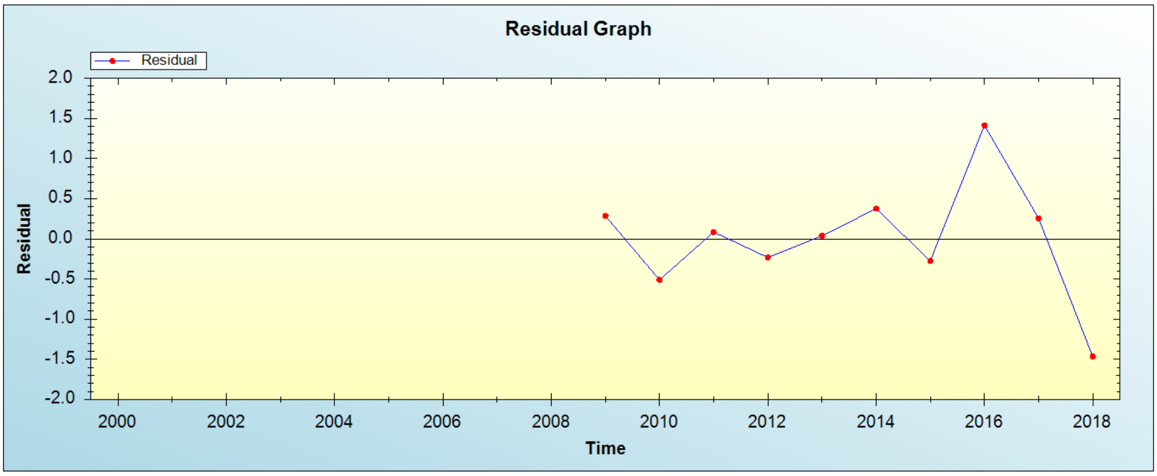

Figure 1: Residual analysis

In-sample Forecast for P

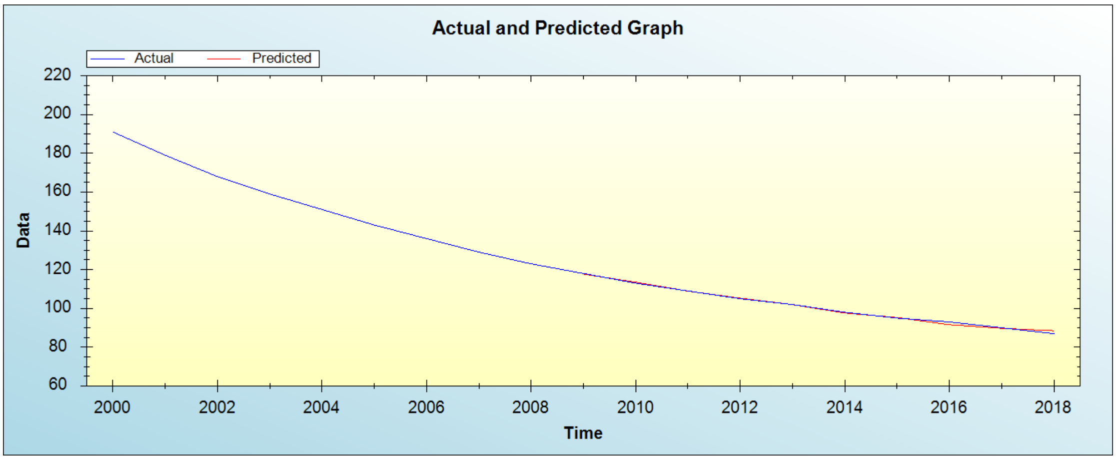

Figure 2: In-sample forecast for the P series

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

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

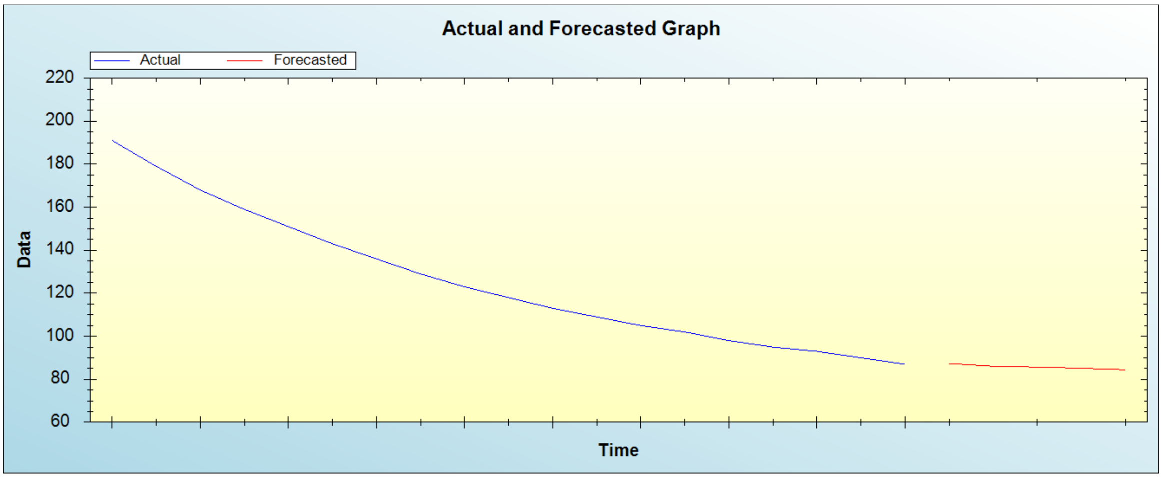

Figure 3: Out-of-sample forecast for P: actual and forecasted graph

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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 371-375, March-2021

https://doi.org/10.47001/IRJIET/2021.503064

Out-of-Sample Forecast for P: Forecasts only

Table 3: Tabulated out-of-sample forecasts

Year	Forecasts
2019	87.3319
2020	86.0745
2021	85.7123
2022	85.1509
2023	84.4712

Over the study period 2000- 2018 the incidence of TB has been gradually declining as a result of effective TB preventive and control measures implemented by the Niger government and its supporting partners. The minimum and maximum TB incidence was 87 and 191 cases per 100 000 population/year with an average of 126 cases per 100 000 population/year. The residual graph and model evaluation statistics indicate that the applied ANN (9,12,1) model is stable. Figure 2 apparently shows that the applied model simulates the observed data very well. The model predicts that over the period 2019-2023 TB incidence will continue to decline from 87 cases per 100 000 population/year in 2019 to 84 cases per 100 000 population/year in 2023.

V. CONCLUSION & RECOMMENDATIONS

Niger has recorded an impressive gradual decline in the incidence of TB over the period 2000-2018 indicating the governments commitment in the fight against TB. The model predicts that this desirable downward trend is expected to continue over the period 2019-2023.Health authorities should not be complacent in the fight against the TB epidemic. The researchers strongly encourage the authorities to strengthen TB/HIV program collaboration, continuous training of healthcare workers on emerging trends on HIV and TB and continuous health education among communities in order to significantly reduce TB incidence.

REFERENCES

[1] Cao L J & Francis E.H(2003). Tay "Support Vector Machine with Adaptive Parameters in Financial Time Series Forecasting", IEEE Transaction on Neural Networks, Vol. 14, No. 6, November pages: 1506-1518.

[2] Farooq T, Guergachi A & S. Krishnan (2007). "Chaotic time series prediction using knowledge-based Green's Kernel and least-squares support vector machines", Systems, Man and Cybernetics, 2007. ISIC. 7-10 Oct. 2007, pages: 373-378.

[3] Fojnica A., Osmanoviae & Badnjeviae A (2016). Dynamic model of tuberculosis-multiple strain prediction based on artificial neural network. In proceedings of the 2016 5th Mediterranean conference on Embedded computing pp290-293.

[4] Kaushik AC &Sahi. S (2018). Artificial neural network-based model for orphan GPCRs.Neural.Comput.Appl. 29,985-992.

[5] Nyoni S. P & Nyoni T (2019a). 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.

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

[7] Raicharoen T., Lursinsap C &Sanguanbhoki (2003) "Application of critical support vector machine to time series prediction", Circuits and Systems. ISCAS '03. Proceedings of the 2003 International Symposium on Volume 5, 25-28 May, 2003, pages: V-741-V-744.

[8] Vapnik V (1998) "Statistical Learning Theory", New York: Wiley.

[9] Yan C Q., Wang R B., Liu C H &Jiang Y (2019). Application of ARIMA model in predicting the incidence of tuberculosis in China from 2018-2019.Zhonghua 40(6):633-637

[10] Zhang GP (2003) "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 371-375, March-2021
https://doi.org/10.47001/IRJIET/2021.503064

Citation of this Article:

Dr. Smartson. P. NYONI, Thabani NYONI, "Prediction of the Annual TB Incidence in Niger Using Artificial Neural Networks" Published in *International Research Journal of Innovations in Engineering and Technology* - *IRJIET*, Volume 5, Issue 3, pp 371-375, March 2021. Article DOI https://doi.org/10.47001/IRJIET/2021.503064

****