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

**ISSN (online): 2581-3048**

**Volume 5, Issue 3, pp 407-411, March-2021**

https://doi.org/10.47001/IRJIET/2021.503070

# Modeling and Forecasting Annual TB Incidence in Mauritania 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 piece of work the ANN approach was applied to analyze TB incidence in Mauritania. 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 Mauritania. The results of the study indicate that TB incidence will continue on a downward trajectory over the period 2019-2023. The government is encouraged to intensify TB surveillance and control programs in order to significantly reduce the incidence of TB.**

***Keywords:*** ANN, Forecasting, TB incidence.

## I. INTRODUCTION

Machine learning methods besides SARIMA models (Yan et al, 2019; Nyoni & Nyoni, 2019a & b) have attracted a lot of attention in the field of Public Health. They are powerful tools to use when dealing with large amounts of data and complex data patterns. Machine learning methods use computational techniques to "learn" information directly from the data without relying on a predetermined equation to model (Weng et al, 2017).The algorithms adaptively improve their performance as the number of samples available for the training process increases. Algorithms of machine learning can be categorized into supervised, unsupervised and semi-supervised learning. Supervised learning trains a model on known inputs and output data to predict future output. Unsupervised learning finds hidden patterns or intrinsic structures in the input data (Weng et al, 2017). Semi supervised learning algorithms use a mixture of both. These algorithms are also put into classes namely regression, classification and clustering algorithms. The table below is a summary of classification of machine learning algorithms:

**Table 1: Machine learning algorithms**

| Classification algorithms | Regression algorithms | Clustering algorithms |
|---|---|---|
| Logistic regression | Linear regression | K-means |
| K-Nearest Neighbour | Nonlinear regression | K-Medoids |
| Support Vector Machine | Support Vector Machine | Fuzzy C means |
| *Neural Network* | Gaussian process | Hierarchical clustering |
| Naïve Bayes | Regression tree | Self-organizing map |
| | | Gaussian mixture model |

In recent decades machine learning methods which have been used in forecasting problems are the artificial neural networks ,the multilayer perceptron (MLPs) ,least square support vector machine (LS-SVM),dynamic least square support vector machine (DLS-SVM) and support vector for regression( Cao L J &Francis E.H,2003 ). The widely used artificial neural networks are the multilayer perceptrons with a single hidden layer feed forward network(FNN).The model is composed of the 3 layers; input, hidden and output layers which are connected by connection weights(Fajnica et al, 2016; Zhang, 2003; Kaushik & Sahi, 2018 ). The support vector machine which was proposed by Vapnik in 1995 is gradually gaining ground in forecasting problems. The method is based on the structural risk minimization principle (Cao &Francis, 2003; Farooq et al, 2007; Raicharoen et al, 2003; Vapnik, 1998). The main idea is to construct an optimal hyperplane through nonlinear mapping of the input space X into a higher dimensional feature space. SVM when applied to classification an optimal hyperplane which separates the two classes of data is found. Whereas in the case of regression a hyperplane is to be constructed that lies close to as many points as much as possible. In this study we apply the ANN (9,12,1) model to predict annual TB incidence in Mauritania. The results of this piece of work are expected to uncover the future trends of annual TB incidence in the country. This is important for health authorities as this will guide them in planning, policy making and resource allocation for health.

 

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

ISSN (online): 2581-3048

Volume 5, Issue 3, pp 407-411, March-2021

https://doi.org/10.47001/IRJIET/2021.503070

## II. LITERATURE REVIEW

Azeez et al (2016) analyzed TB incidence data from January 2010 to December 2015 in an Eastern Cape Healthcare facility, South Africa. The SARIMA and SARIMA-GRNN hybrid model were applied. The study concluded that the hybrid model has a better TB incidence forecasting with a lower AIC. Mao et al (2018) developed a SARIMA model to forecast TB incidence in China. Monthly TB incidence data from January 2004 to December 2015. The study concluded that the SARIMA (1,0,0) (0,1,1)12 was the best model and it was a useful tool for monitoring epidemics. Cao et al (2013) applied the SARIMA and the hybrid model, SARIMA-GRNN to analyse monthly TB incidence cases from June 2005 to December 2011.The study findings revealed that the hybrid model showed better TB incidence forecasting than the SARIMA model. Nyoni & Nyoni (2019a ) constructed a SARIMA model to predict monthly TB notifications at Zengeza 3 clinic in Zimbabwe using monthly data for TB notifications for the period January 2013 to December 2018. The optimal model SARIMA (2,0,2) (1,0,1)12 projected a decline in the monthly TB notifications over the out of sample period. In a similar study Nyoni &Nyoni (2019b) developed a SARIMA model to forecast TB notifications at Silobela District Hospital in Zimbabwe using monthy data covering the period January 2004 to December 2018. The optimal model SARIMA (1,0,1) (0,1,1)12 projected a decline in TB notifications over the out of sample period.

## III. METHOD

The Artificial Neural Network (ANN), which we intend to employ in this paper; 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 Mauritania 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 Mauritania. 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 2: Descriptive statistics**

| Mean | Median | Minimum | Maximum |
|---|---|---|---|
| 153.58 | 143.00 | 93.000 | 250.00 |
| Std. Dev. | C.V. | Skewness | Ex. kurtosis |
| 48.811 | 0.31782 | 0.55198 | -0.90131 |
| 5% Perc. | 95% Perc. | IQ range | Missing obs. |
| Undefined | 250.00 | 82.000 | 0 |

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

**Table 3: ANN model summary**

| Variable | F |
|---|---|
| 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 |

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

**ISSN (online): 2581-3048**

**Volume 5, Issue 3, pp 407-411, March-2021**

https://doi.org/10.47001/IRJIET/2021.503070

| | |
|---|---|
| Back Propagation Learning: | |
| Learning Rate | 0.005 |
| Momentum | 0.05 |
| Criteria: | |
| Error | 0.009891 |
| MSE | 0.744282 |
| MAE | 0.686418 |

*Residual Analysis for the ANN model*

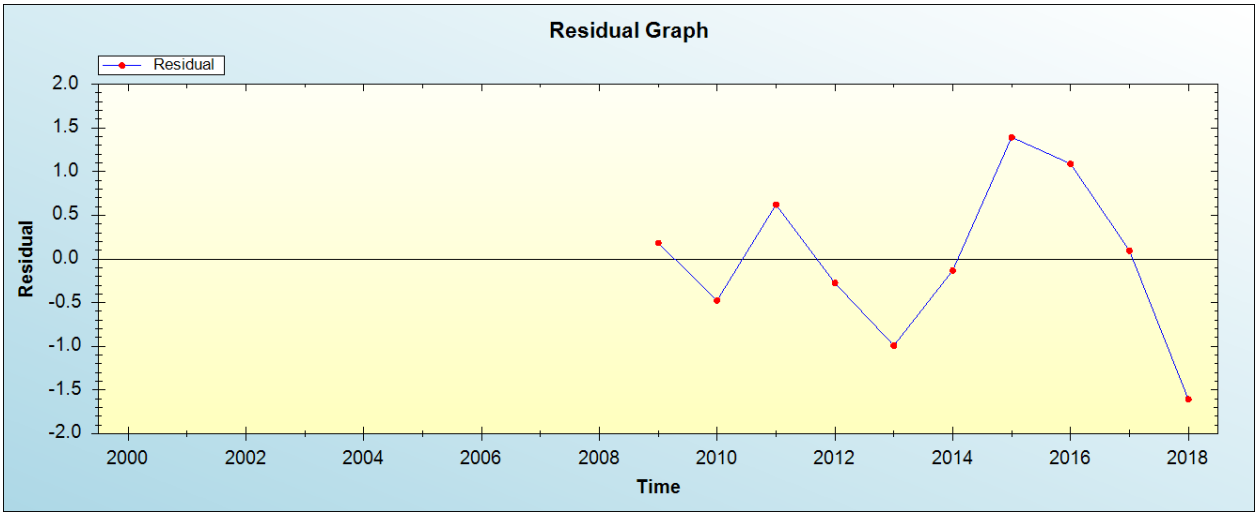


**Figure 1: Residual analysis**

*In-sample Forecast for F*

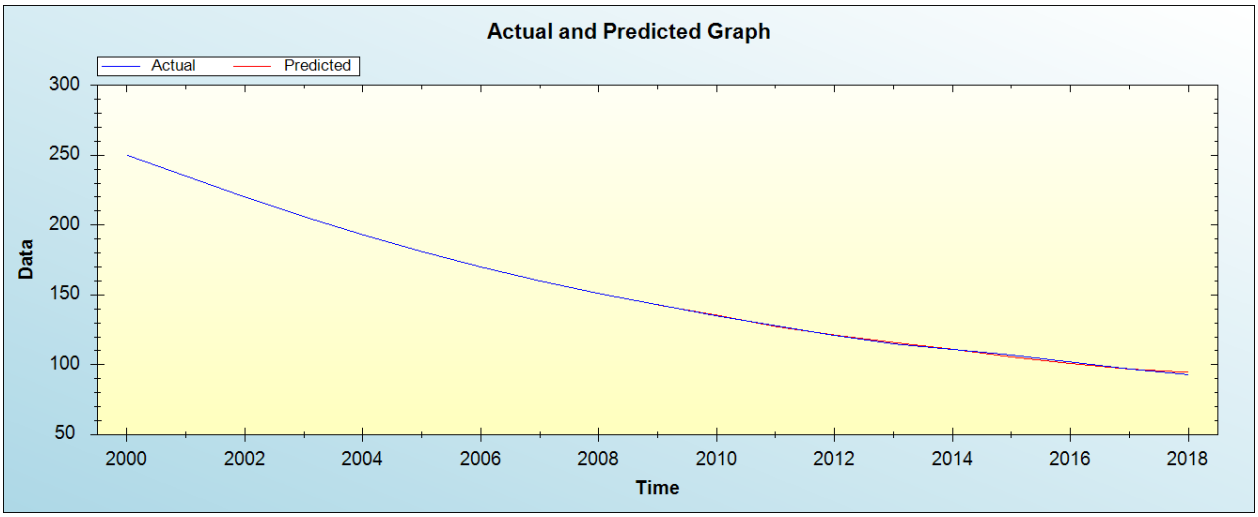


**Figure 2: In-sample forecast for the F series**

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

*Out-of-Sample Forecast for F: Actual and Forecasted Graph*

 

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 5, Issue 3, pp 407-411, March-2021
https://doi.org/10.47001/IRJIET/2021.503070

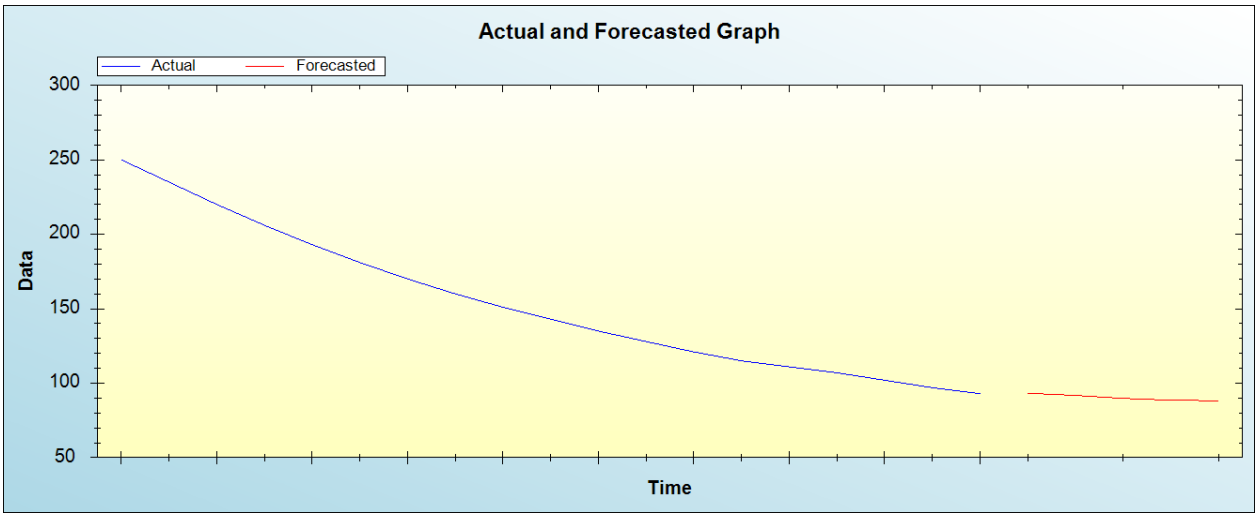


**Figure 3: Out-of-sample forecast for F: actual and forecasted graph**

*Out-of-Sample Forecast for F: Forecasts only*

**Table 4: Tabulated out-of-sample forecasts**

| Year | Forecasts |
|---|---|
| 2019 | 93.2521 |
| 2020 | 91.8092 |
| 2021 | 89.8060 |
| 2022 | 88.5459 |
| 2023 | 88.2729 |

Over the study period 2000-2018 the incidence of TB has been gradually decreasing in Mauritania. This feature has been characterized by a minimum and maximum of 93 and 250 cases per 100 000 population per year respectively with an average incidence of 154.0 cases per 100 000 population/year. The applied data is positively skewed with an excess kurtosis of -0. 90131.The model evaluation criteria (Error, MSE, MAE) and the residuals show that the ANN (9,12,1) model is stable and suitable for forecasting TB incidence in Mauritania. In-sample forecasts indicate that the model simulates the observed data very well. The model predicts that the TB incidence will continue to decline from around 93 in 2019 to 88 cases per 100 000 population in 2023.

## V. CONCLUSION & RECOMMENDATIONS

Mauritania's response to the TB epidemic has been very good as indicated by Figure 3. The incidence of TB has been on a downward trend over the period 2000-2018.This shows government's commitment in the fight against TB. The model predicts a continued decline in incidence over the period 2019-2023.Therefore we encourage the government to intensify TB surveillance and control programs in order to further bring down TB incidence.

## REFERENCES

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

[2] 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.

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

[4] L.J. Cao & 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; 1506-1518.

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

[6] Nyoni & Nyoni (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`

 

ISSN (online): 2581-3048
Volume 5, Issue 3, pp 407-411, March-2021
https://doi.org/10.47001/IRJIET/2021.503070

[7] Raicharoen T., Lursinsap C & Sanguanbhoki P (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] Weng SF., Reps J., Kai J., Garibaldi JM &Qureshi N (2017).Can machine learning improve cardiovascular risk prediction using routine clinical data? Plos One 12 (4).

[10] 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

[11] Zhang G P, "Time series forecasting using a hybrid ARIMA and neural network model", Neurocomputing 50: 159–175.

**Citation of this Article:**

Dr. Smartson. P. NYONI, Thabani NYONI, "Modeling and Forecasting Annual TB Incidence in Mauritania Using Artificial Neural Networks" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 5, Issue 3, pp 407-411, March 2021. Article DOI https://doi.org/10.47001/IRJIET/2021.503070

*******

 