Time Series Analysis of Tuberculosis Cases in Mining Communities in Ghana – A Case Study at Tarkwa Nsuaem Municipality

[1]C. C. Nyarko, [1]P. K. Nyarko and [1]A. Buabeng

[1]University of Mines and Technology, P. O. Box 237, Tarkwa, Ghana

Nyarko, C. C., Nyarko, P. K. and Buabeng, A. (2020), "Time Series Analysis of Tuberculosis Cases in Mining Communities in Ghana – A Case Study at Tarkwa Nsuaem Municipality", *Proceedings of 6th UMaT Biennial International Mining and Mineral Conference*, Tarkwa, Ghana, pp. 153-158.

Abstract

The main purpose of this paper was to examine the incidence rate of Tuberculosis (TB) in Tarkwa Nsuaem Municipality, a mining community in Western Ghana and to identify some of the demographic factors that necessitate the spread and trend of the diseases. Monthly reported case of TB from January 2009 to December 2013 was obtained from the Tarkwa Municipal Government Hospital which was analysed using the Box-Jenkins approach. An Autoregressive Integrated Moving Average, ARIMA (1, 1, 0) was found to be the best among the competing models with relatively minimum information loss (Bayesian Information Criteria (BIC) and Akaike Information Criteria (AIC)). A forecast result shows a constant decrease in the number of TB cases in the municipality for the first six months of 2014. Exposure and environment were strongly identified as risk factors necessitating the transmission of TB in the communities. The paper concluded that, the decrease in TB cases in the municipality could be sustained if TB patients adhere to all the rules governing TB medications.

Keywords: Tuberculosis (TB), Bayesian Information Criteria, Forecast, Medications, and Multi Drug Resistant TB

1 Introduction

Tuberculosis (TB), an ancient disease that has caused more suffering and death still remains a major public health problem worldwide (Mohajan, 2014). Globally, an annual incidence of nine million new cases of TB bacilli infection are reported with more than 1.5 million deaths occur in sub-Saharan Africa (Addo *et al*., 2010). However, in Ghana, an estimate of over 46 000 new cases of TB are being reported annually (Anon., 2014). Moreover, less than a third of the estimated number of cases is officially reported each year to the health facilities in the country despite the availability of a cheap and effective treatment (Addo *et al*., 2010). Averagely 122 new TB cases develop every day in Ghana and out of that, 27 people die, that is an average of one person dying every day in Ghana (Appiah, 2017). Although TB is a curable disease, one of the main challenges in TB control is the emergence and spread of drug resistance. According to the 2014 World Health Organisation (WHO) global tuberculosis control report, out of 9 million TB cases, nearly 480 000 had multidrug-resistant to TB (MDR-TB) (Anon., 2010). This emergence and spread of drug resistance cases (MDR-TB) has surfaced in Ghana, and currently accounting for almost 2 percent of all TB cases. As of 2015, over 640 new form of tuberculosis cases has been reported, which according to health professionals is considered as dangerous (Appiah, 2017).

For instance, in 2014, the incidence of TB infection in Ghana was estimated at 44 000, including 11 000 individuals with HIV coinfection, 1 799 cases were tested for further resistance. In 2015, out of the 14 632 cases of TB diagnosed and placed on treatment, 60 were Multi Resistant TB. This number increased to 77 in 2016, and out of the 77 TB patients, 12 have died, 15 declared cured while 50 are still on treatment (Anon., 2017).

The Tarkwa-Nsuaem Municipality is one of the seventeen districts in the Western Region of Ghana known to be a busy mining community. The municipality contributes about 35% of Tuberculosis cases report to the Region (Anon, 2009). This is as a result of the dominating activities of small-scale miners commonly called ('Galamsey'). It is estimated that over 90% of reported cases have a history of "Galamsey operators" (either they are currently involved or have ever worked in small scale mining). Research has shown that mineral mining is a high risk for tuberculosis morbidity and mortality (Rees and Murray, 2007; Sonnenberg *et al*., 2005: Akcom, 2016).

The catastrophic impact and growth of TB epidemic in the municipality have necessitated the need to improve on TB prevention and control measures. In literature, several qualitative researches have been conducted to investigate the prevalence of TB and MDR-TB in Ghana, however, little research has been done to modelled the dynamics of TB taking into account the MDR-TB as well as the associated risk factors. This alarming situation necessitates a critical look at the current dynamics and forecast of the incidence of Tuberculosis (TB) from the multivariate point of view. This is because, in the presence of complexity mathematical models offer valuable tools for synthesising information to understand epidemiological patterns, and for developing quantitative evidence for decision-making in global health (Heesterbeek, 2015).

Hence, this paper seeks to model the current dynamics and forecast of the incidence of Tuberculosis (TB) in the Tarkwa Nsuaem Municipality.

1.1 Risk Factors that increase the Probability of TB Transmission

The probability of the risk for transmission of tuberculosis is increased as a result of various environmental factors such as:

i. exposure to TB in small, enclosed spaces and duration of contact.

ii. environmental factors that affect the concentration of tuberculosis organisms.

iii. inadequate local or general ventilation that results in insufficient dilution or removal of infectious droplet nuclei.

iv. recirculation of air containing infectious droplet nuclei.

Also, people with health conditions such as weakened immune system are at a greater risk for developing TB disease since they cannot fight the TB germs. That is, TB and HIV have a strong fatal correlation; each drives the progress of the other. Worldwide, TB is one of the leading causes of death among people living with HIV. According to WHO, the following are susceptible to TB:

i. people with HIV infection,

ii. people with malnutrition,

iii. people who have diabetes,

iv. people who have chronic liver disease,

v. patients who are at the end stage of kidney disease,

vi. infants and young children under 5 years,

vii. people who became infected with TB bacteria in the last 2 years,

viii. people who inject illegal drugs or drink alcohol,

ix. use of tobacco greatly increases the risk of getting TB and dying from it,

x. elderly people, and

xi. people who are not treated correctly for TB in the past.

2 Materials and Methods Used

This study is carried out using monthly recorded cases of TB data from Tarkwa Municipal Hospital from 2009 to 2013. There were a total number of 250 patients who were diagnosed of TB in Tarkwa-Nsuaem municipality. Out of the 250 patients, 201 were males and 49 females. The data was then modelled using Autoregressive Integrated Moving Average (ARIMA), a stochastic model popularised by Box *et al*. (2015). The ARIMA (p, d, q) model is a combination of Autoregressive (AR) model and the Integrated Moving Average which shows that, the present value has something to do with the past residuals. The ARIMA process can be defined as:

$$\phi(B)(\Delta^d Y_t) = \theta(B)e_t \qquad (1)$$

where

Y_t is the number of TB cases recorded at time t;

Δ^d is the order of differencing;

$\phi(B) = (1-\phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)$ is the Autoregressive (AR) characteristic operator; and

$\theta(B) = (1-\theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)$ is the Moving Average (MA).

The estimation of the model consists of three steps, namely: model identification, estimation of parameters and diagnostic checking

2.1 Model Identification

For any ARIMA (*p, d, q*) process, the identification step involves the use of the techniques to determine the values of *p*, *d* and *q*. The values are determined by using Autocorrelation function (ACF) and Partial Autocorrelation function (PACF). In this study, several tentative models were fitted and an appropriate model was selected based on their information criteria (Akaike Information Criteria (AIC) or Bayesian Information Criteria (BIC)). However, for a non-stationary series the data is differenced to make the series stationary. The number of times the series is differenced determines the order of *d*.

2.2 Estimation of Parameters and Diagnostic Checking

The second step is the estimation of the model parameters for the "best model" that have to be selected. The estimated model is then checked to verify if it adequately represents the series. Diagnostic checks are performed to ascertain whether the residuals of the selected model are randomly and normally distributed. In this paper, the standardised residual, the ACF of residuals and Ljung-Box statistic plots were employed. Lastly, an overall check of the model adequacy was explored using the Box-Pierce and the Ljung-Box test statistics. Collectively, these tests establish the adequacy of the selected model.

3 Results and Discussions

Figure 1 shows the pattern of monthly TB cases reported in Tarkwa Municipal Hospital, in a mining community in Western Ghana. As observed, the series displays a downward trend with considerable variation indicating little evidence of seasonality. However, the change in variations from months to months throughout the period indicates that the TB cases reported is not stationary. This was confirmed by the Autocorrelation Function (ACF) plot in Figure 2. The plot shows a very slow linear decay which is typical of non-stationarity.

Thus, the Augmented Dickey-Fuller (ADF), as well as the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, were therefore computed to validate this claim. As indicated in Table 1, it was shown that the series under consideration was nonstationary. Thus, the series became stationary after taking the first difference, which indicates an integrated of order one.

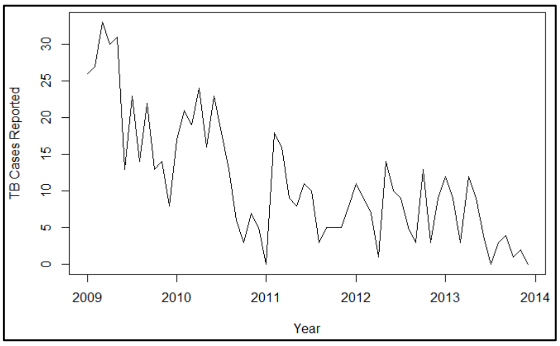

Figure 1 TB Cases Reported

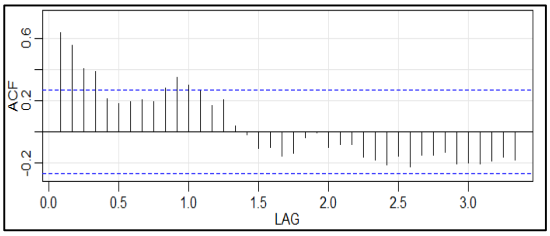

Figure 2 ACF Plot

Table 1 Summary of Stationarity Test

Test	Hypothesis	Differencing (P-Value)	
		Order 0	Order 1
ADF	H_0: Not Stationary H_1: Stationary	-3.2 (0.0965)	-4.4923 (0.0100)
KPSS	H_0: Stationary H_1: Not Stationary	1.6857 (0.0100)	0.0310 (0.1000)

3.1 Model Identification

In selecting the 'best' ARIMA model, nine competing models were fitted and the result shown in Table 2. The 'best' model is the model with the minimum information criteria. Thus, the 'best' model is ARIMA (1,1,0).

Table 2 Competing ARIMA models and Information Criteria

Model (*p*, 1, *q*)	AIC	BIC
(0,1,0)	356.76	358.78
(1,1,0)	**345.69**	**349.93**
(0,1,1)	347.87	351.84
(1,1,1)	347.81	353.77
(2,1,0)	349.30	355.26
(2,1,1)	348.63	356.59
(0,1,2)	347.91	353.88
(1,1,2)	349.17	357.12
(2,1,2)	351.51	361.46

3.2 Estimation of Parameters and Diagnostic Checking

The ARIMA model estimated was given in Table 3. Regarding the model diagnostic checking, the p-values from the Box-Pierce test (>0.05) indicates that the residuals of the selected model is now uncorrelated. Also, the Shapiro-Wilk test (p-value >0.05), indicates that the residuals are normally distributed. This is also confirmed by the diagnostic plots in Figure 3.

Table 3 Summary ARIMA Model

Variable	Parameter Estimate	Standard Error	t value	P value
AR1	-0.4253	0.1217	-3.4958	0.001

AIC=347.87, BIC=351.84, Box-Pierce=2099 (0.6468), Shapiro-Wilk= 0.9783 (0.4201)

As observed in Figure 3(a), the standardised residual plot displays a normally distributed residuals as the points indicate zero trace of trend, no outliers and no changing variance across time. In Figure 3 (b), the ACF residuals plot shows that, only two (2) lags out of the seventeen (17) lags of the series residuals exceeds the significant bounds. This is negligible, since the probability of a spike being significant by chance is minimal. This simply gives an indication of non-significant and autocorrelation, since it is expected that at most two (2) out of seventeen (17) sample autocorrelations to exceed the 95% significance bounds. Furthermore, in Figure 3 (c), the p-value for Ljung-box statistic plot; shows that all the p-values show an indication of an adequate model. Collectively, these tests suggest that the model is very significant.

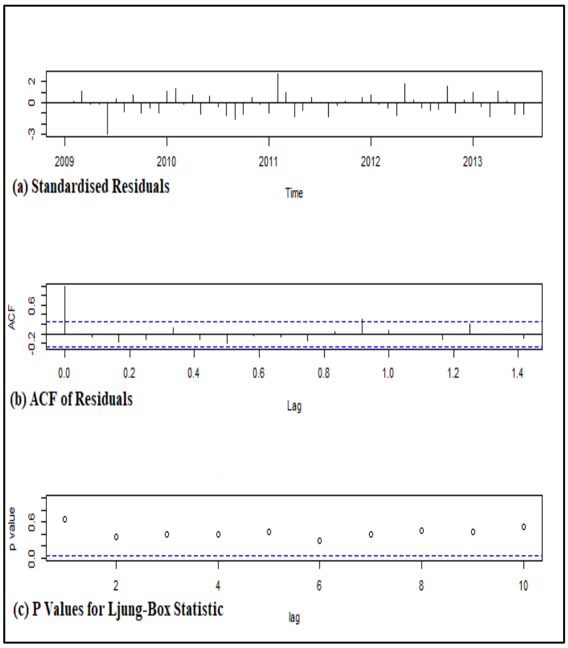

Figure 3 Diagnostic of Regression Model with ARIMA (1,1,0) Errors

Thus, the ARIMA (1, 1, 0) model is written as:

$$Y_t = Y_{t-1} - 0.4253(Y_{t-1} - Y_{t-2}) \qquad (2)$$

Equation (2) shows that the fitted model is a linear combination of previous incidences of TB cases reported. To further ascertain the predictive performance of the model, Figure 4 shows the plot of observed as well as the predicted production values for five months. Table 4 shows the observed and predicted values of the model.

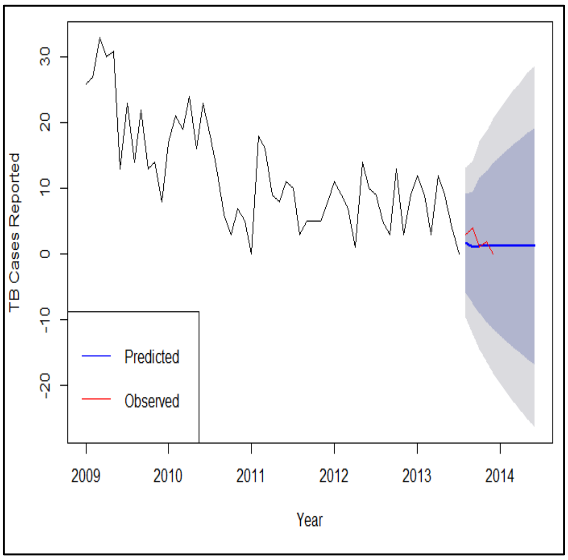

Figure 4 Plot of Observed and Predicted Values

Table 4 Observed and Predicted Values with Respective Errors

Year	Month	Observed	Predicted
2013	Aug	3	1.7014
	Sep	4	0.9777
	Oct	1	1.2855
	Nov	2	1.1546
	Dec	0	1.2103
2014	Jan		1.1866
	Feb		1.1967
	Mar		1.1924
	Apr		1.9421
	May		1.9343
	Jun		1.9376

4 Conclusions

This paper determined the trend and incidence rate of Tuberculosis (TB) in Tarkwa Nsuaem Municipality. Environment and exposure of individuals was found to play a major rule in the transmission of TB in the municipality. An Autoregressive Integrated Moving Average model, ARIMA (1, 1, 0) was found to be the best model among the competing models with relatively minimum information loss (BIC and AIC). All the tests conducted indicate that the model obtained was very significant. The projection for the next six months of 2014 showed a static trend in the reported cases of TB (i.e. revolving around 1 or 2). This could be maintained or reduced in subsequent years if TB patients adhere to all the rules governing its medications and constant supply of drugs by Stakeholders be made available at TB health centres. However, one of the limitations of this study is acquisition of data. Further investigations would perform a comparative study on TB subject to availability of data. It is recommended that knowledge about TB especially in mining communities should be intensified.

References

Ackom, T. K. (2016), *Tuberculosis Treatment Outcomes in the Tarkwa Nsuaem Municipality*, Master's Thesis, Kwame Nkrumah University of Science and Technology, 81 Pp.

Addo, K. K., Yeboah-Manu, D., Dan-Dzide, M., Owusu-Darko, K., Caulley, P., Mensah, G. I., Minamikawa, M., Lienhardt, C., Bonsu, F. A. and Ofori-Adjei, D. D. (2010), "Diagnosis of Tuberculosis in Ghana - The Role of Laboratory Training", *Ghana Medical Journal*, Vol. 44, No.1, pp. 31-36.

Appiah, P. (2017), Ghana Records Increasing Multi Resistant Tuberculosis Cases, *https://www.myjoyonline.com/news/2017/March-24th/ghana-records-increasing-multi-resistant-tuberculosis-cases.php*. Accessed: June 13, 2018.

Anon. (2009), *Tarkwa-Nsuaem Municipal Health Directorate 2009 Annual Report*, Municipal Health Directorate, 70 pp.

Anon. (2010), World Health Organization, Stop TB Partnership-Tuberculosis Global Facts, *http://www.who.int/tb/publications/2010/factsheet_tb_2010.pdf*. Accessed: June 12, 2018.

Anon. (2014)," Tuberculosis: New Faces of an Old Disease", *http://www.lekari-bez hranic.cz/cz/downloads/diseases/TB*. Accessed: June 12, 2018.

Anon., (2017)," Tuberculosis Report in Ghana", *http://www.Ghanahealthservice.gov.gh*. Accessed: June 19, 2018.

Box, G. E., Jenkins, G. M., Reinsel, G. C. and Ljung, G. M. (2015), *Time Series Analysis: Forecasting and Control*, 5th edition, John Wiley & Sons, 712 pp.

Mohajan, H. (2014), "Tuberculosis is a Fatal Disease among Some Developing Countries of the World", American Journal of Infectious Diseases and Microbiology, Vol. 3, No. 1, pp. 18-31.

Heesterbeek, H., Anderson, R. M., Andreasen, V., Bansal, S., De Angelis, D., C. Dye, K. (2015), "Modelling Infectious Disease Dynamics in the complex Landscape of Global Health", *Science*, Vol. 347, No. 6227, 4339 pp.

Rees, D. and Murray, J. S. (2007), "Silicosis and Tuberculosis", *International Journal of Tuberculosis and Lung Disease*, Vol. 11, No. 5, pp. 474-484.

Sonnenberg, P., Glynn, J. R., Fielding, K., Murray, J., Godfrey-Faussett, P. and Shearer, S. (2005), How soon after infection with HIV does the risk of tuberculosis start to increase? A retrospective cohort study in South African gold miners", *J. Infect. Dis.*, Vol. 191, No. 2, pp. 150-158.

Authors

Mrs Christiana Cynthia Nyarko is a Senior Lecturer at the Mathematical Sciences Department of the University of Mines and Technology (UMaT) Tarkwa, Ghana. She holds a PhD in Mathematics, MPhil in Statistics from University of Mines and Technology and University of Ghana, respectively. She was awarded a BSc. Ed degree in Mathematics at Obafemi Awolowo University, Adeyemi Campus Ondo, Nigeria. She is a member of the Ghana Statistical Association, Mathematical Association of Ghana and IBS Ghana Region. Her research Interest include, Statistical Modelling of Parametric and Non-Parametric Data. Epidemiological Modelling and Mathematical Modeling.

Peter Kwesi Nyarko holds a PhD in Mathematics from the University of Mines and Technology (UMaT), MSc in Mathematics from the Kwame Nkrumah University of Science and Technology, and a BSc in Mathematics from the University of Cape Coast, all in Ghana. He is a member of Ghana Mathematics Society as well as the Mathematical Association of Ghana. He is a Lecturer at the Department of Mathematical Sciences in UMaT. His research areas include dynamical systems and mathematical applications to forest dynamics.

Albert Buabeng holds an MPhil in Mathematics (Statistics) and BSc in Mathematics from the University of Mines and Technology (UMaT), Tarkwa, Ghana. He is a member of the Ghana Mathematics Society. He is an Assistant Lecturer in the Mathematical Sciences Department at UMaT. He is currently pursuing his PhD in Mathematics. His research interests are in Machine Learning, Artificial Neural Networks (ANN), Dimensionality Reduction, Time Series Analysis and Multivariate Prediction for Quality Control.