Scholars Journal of Physics, Mathematics and Statistics OPEN ACCESS
Abbreviated Key Title: Sch J Phys Math Stat
ISSN 2393-8056 (Print) | ISSN 2393-8064 (Online)
Journal homepage: http://saspjournals.com/sjpms

Modeling Typhoid Mortality with Box-Jenkins Autoregressive Integrated Moving Average Models

Obubu Maxwell[1*], Oyafajo Oyindamola Abubarkri[2], Anyawu Ifeyinwa Fidelia[3], Olayemi Joshua I[4]

[1]Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria
[2]Department of Bioethics, Center for Bioethics and Research, 102, Basorun Rd, Akobo, Ibadan, Nigeria
[3]Department of Operations, National Identity Management Commission, Nigeria
[4]Department of Mathematical Sciences, Federal University Gusau, Nigeria

***Corresponding author:** Obubu Maxwell
| **Received:** 05.03.2019 | **Accepted:** 10.03.2019 | **Published:** 30.03.2019
DOI: 10.21276/sjpms.2019.6.3.2

Abstract **Original Research Article**

In this paper, modeling and forecasting typhoid mortality rate using ARIMA models was examined. Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) was employed to analyze typhoid mortality rate in Delta State. The study intended mainly to forecast the typhoid mortality rate for the coming years. Series of tentative models were developed to forecast the mortality rate, but based on minimum AIC and BIC values and after the estimation of parameters and series of diagnostic test were performed, ARIMA(0,1,0) model was proved to be the best model for forecasting after satisfying the model assumptions. The forecasted results revealed a decreasing pattern of typhoid mortality rate 2019 to 2022.

Keywords: Typhoid mortality, ARIMA models, augmented dickey-fuller test, ACF/PACF plots, Forecasting.

INTRODUCTION

Typhoid fever, also known simply as typhoid, is a bacterial infection due to *Salmonella typhi*, that causes symptoms [1]. Symptoms may vary from mild to severe and usually begin six to thirty days after exposure [2, 3]. Often there is a gradual onset of a high fever over several days [2]. Weakness, abdominal pains, constipation, and headaches also commonly occur [3-4]. Diarrhea is uncommon and vomiting is not usually severe [4]. Some people develop a skin rash with rose-colored spots [3]. In severe cases there may be confusion [4]. Without treatment symptoms may last weeks or months [3]. Other people may carry the bacterium without being affected; however, they are still able to spread the disease to others [5]. Typhoid fever is a type of enteric fever along with paratyphoid fever.[1] The cause is the bacterium *Salmonella typhi*, also known as *Salmonella Enterica* serotype *Typhi*, growing in the intestines and blood [3, 4]. Typhoid is spread by eating or drinking food or water contaminated with the faeces of an infected person [5]. Risk factors include poor sanitation and poor hygiene [1]. Those who travel to the developing world are also at risk [4]. Humans are the only animal infected [5]. Diagnosis is by either culturing the bacterium or detecting the bacterium's DNA in the blood, stool, or bone marrow [1, 3, 6] Culturing the bacterium can be difficult [7]. Bone marrow testing is the most accurate [6]. Symptoms are similar to that of many other infectious diseases [4]. In 2010, there was 27 million cases reported [1]. The disease is most common in India, and children are most commonly affected [1, 5]. Rates of disease decreased in the developed world in the 1940s as a result of improved sanitation and use of antibiotics to treat the disease [5]. About 400 cases are reported and the disease is estimated to occur in about 6,000 people per year in the United States [4, 8]. In 2013, it resulted in about 161,000 deaths - down from 181,000 in 1990 (about 0.3% of the global total) [9]. The risk of death may be as high as 25% without treatment, while with treatment, it is between 1 and 4% [1, 5]. Therefore, the objective of this paper are: To select the best model for Typhoid Mortality (TM) in the light of the available data, and to use the best model selected for forecasting TM until the year 2022.

MATERIALS AND METHODS

In this paper, we have used the Time series data on Typhoid Mortality Cases for past 22 years (1996 -2018). The mortality data were sourced from the records unit of Federal Medical Centre Asaba, Delta State, Nigeria. We have used GRETL (Gnu Regression, Econometrics and Time-series Library) software for plotting the graphs and analysis of the data set.

Box-Jenkins ARIMA Model

$$\Delta y_t = y_t - y_{t-1}$$

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2}$$

Where, y_t is time series at time $t, y_{t-1}$is the proceeding time series of y_t , Δy_t is the first order difference, $\Delta^2 y_t$ is the second order difference of the current observation, y_t is the current observation and y_{t-2} is the preceding time series to y_{t-1} in the same series. After the appropriate differencing, the expected time series is expected to exhibit features of a stationary time series so that the appropriate ARIMA (p, d, q) process can be used to model the remaining serial correlation in the series. Where p is the number of auto regressive terms, d is the number of non seasonal differences, q is the number of lagged forecast errors in the prediction equation.

For a time series process y_t , *ARIMA (0,0,1) / AR(1)* is the first order auto-regressive process and is given by;

$$y_t = \mu + \emptyset_1 y_{t-1} + e_t$$

and a first order moving average process ARIMA (0,0,1) / MA(1) and is given by;

$$y_t = \mu - \theta_1 \varepsilon_{t-1} + e_t$$

Where $\emptyset$ and θ are coefficients of polynomial with order *p* and *q* respectively.

Alternatively, the model ultimately derived may be a mixture of these processes and of higher orders, in that case, a stationary ARMA (p, q) process is defined by;

$$y_t = \mu + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \ldots\ldots \theta_q \varepsilon_{t-q} + e_t$$

Where y_t is the degree of the differencing, ε_t is independently and normally distributed residual with zero mean and constant variance for $t = 1,2,3, \ldots, n$.

The Augmented Dickey - Fuller Test

The augmented Dickey–Fuller (ADF) test is most widely used test for checking Stationarity of a series. If d equals 0, the model becomes ARMA, which is linear stationary model. ARIMA (i.e. d > 0) is a linear non-stationary model. If the underlying time series is non-stationary, taking the difference of the series with itself predecessor to determine d makes it stationary, and then ARMA is applied onto the differenced series. A stationary process has a constant mean and variance over the time period. There are various methods available to make a time series stationary. Normally differencing techniques are used to transform a time series from a non-stationary to stationary by subtracting each datum in the series from its predecessor.

Model Identification Criteria

At the identification stage different ARIMA are formulated and tested on the data then their respective Akaike Information Criterion, Schwarz-Bayesian Information Criteria (BIC) and Hannan-Quinn Criteria (HQC) were considered and recorded. In each case, the model with the least AIC, BIC and HQC values were selected and subjected to diagnostic check to ensure that they fit well with the data.

$$AIC = (-2logL + 2m) \; where \; m = p + q + P + Q$$

And L is the likelihood function

$$-2logL = n(1 + log\, 2\pi) + n \log \sigma^2$$

Where σ^2 *is the mean square error, this implies that;*

$$AIC = \{ n(1 + log\, 2\pi) + n \log \sigma^2) + 2m \}$$
$$BIC = log\, \sigma^2 + \{ (m \log n) / n \}$$

RESULTS AND DISCUSSION

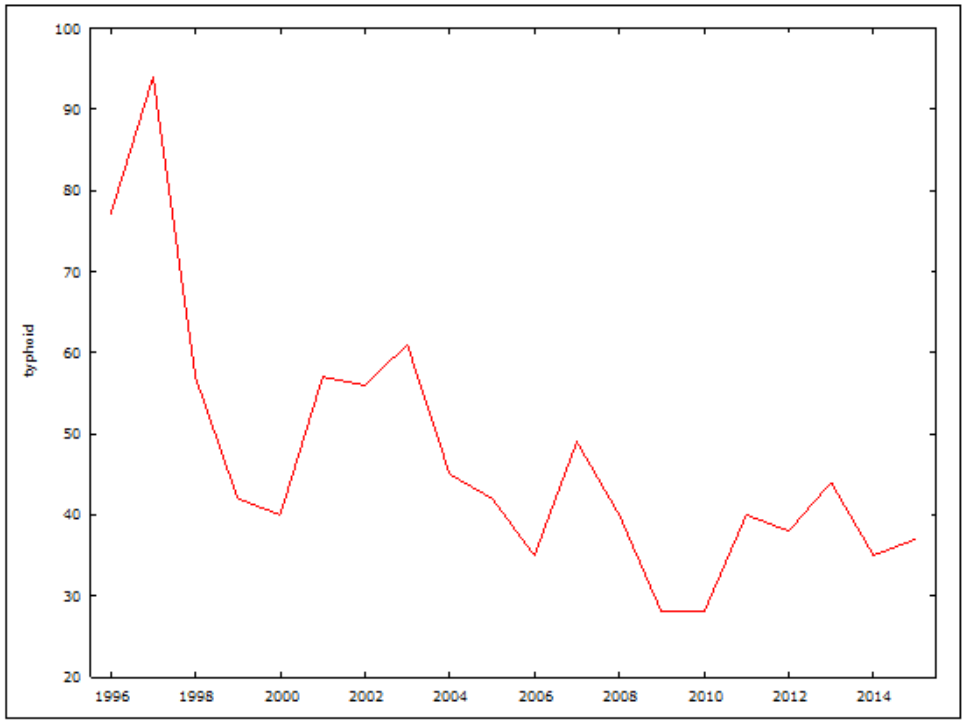

Fig-1: Time Series plot for Typhoid Mortality data series

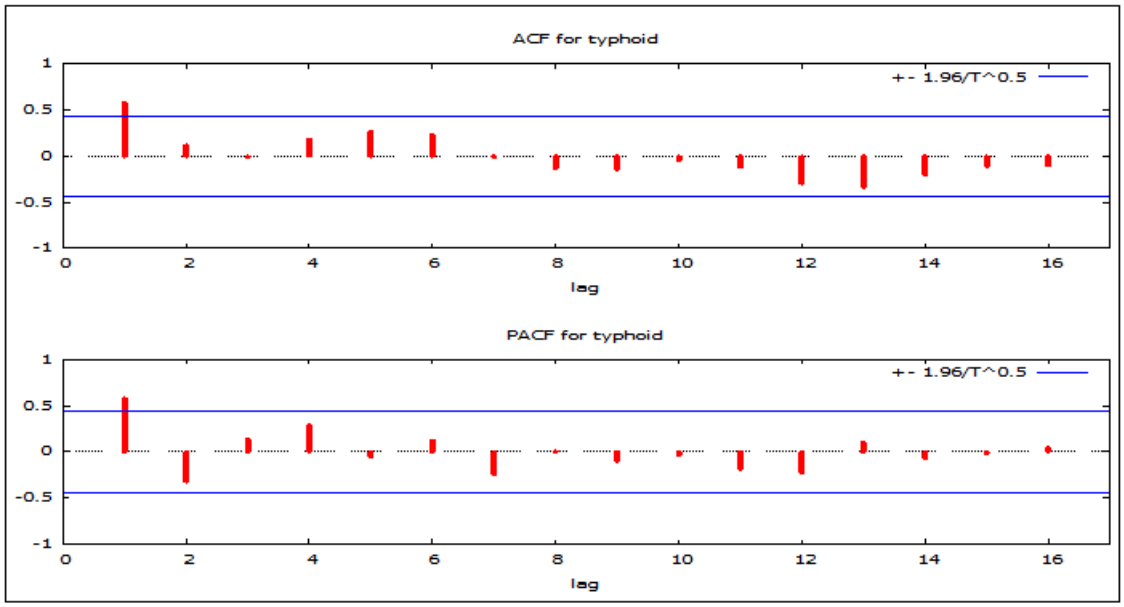

Fig-2: Correlogram for Typhoid mortality data series

Table-1: Augmented Dickey-Fuller test for Stationarity of Typhoid Mortality data series

D	t-statistic	p-value	A-value
0	−1.394	0.1466	0.05
1	−4.801	7.33e-05	0.05

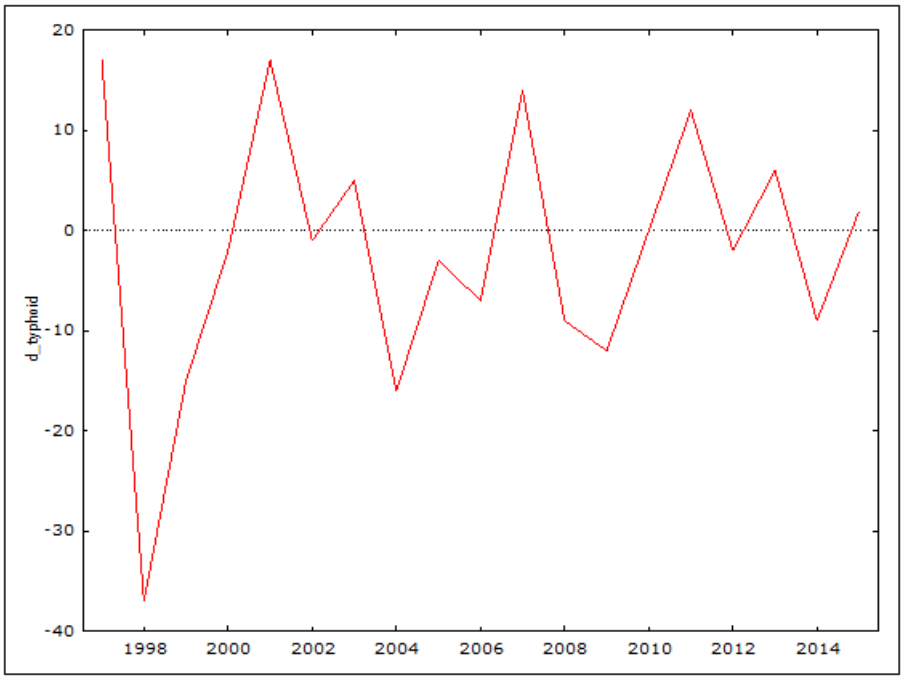

Fig-3: Time Series plot for differenced Typhoid Mortality data series

Table-2: Identification of Best ARIMA model for Typhoid Mortality

ARIMA Model	*AIC*	*BIC*	*HQC*
000	169.1457	170.1414	169.3401
001	164.9584	167.9456	165.5415
010	**152.8524**	**153.7968**	**153.01220**
011	155.8697	158.7030	156.3492
110	156.3915	159.2249	156.8711
111	164.3939	168.3768	165.1714
210	157.7446	161.5223	158.3839
211	153.6928	158.4150	154.4920

Akaike information criteria (AIC), Bayesian information criteria (BIC), and Hannan-Quinn criteria (HQC).

Diagnostic check on the best model for Typhoid Mortality / Model Verification

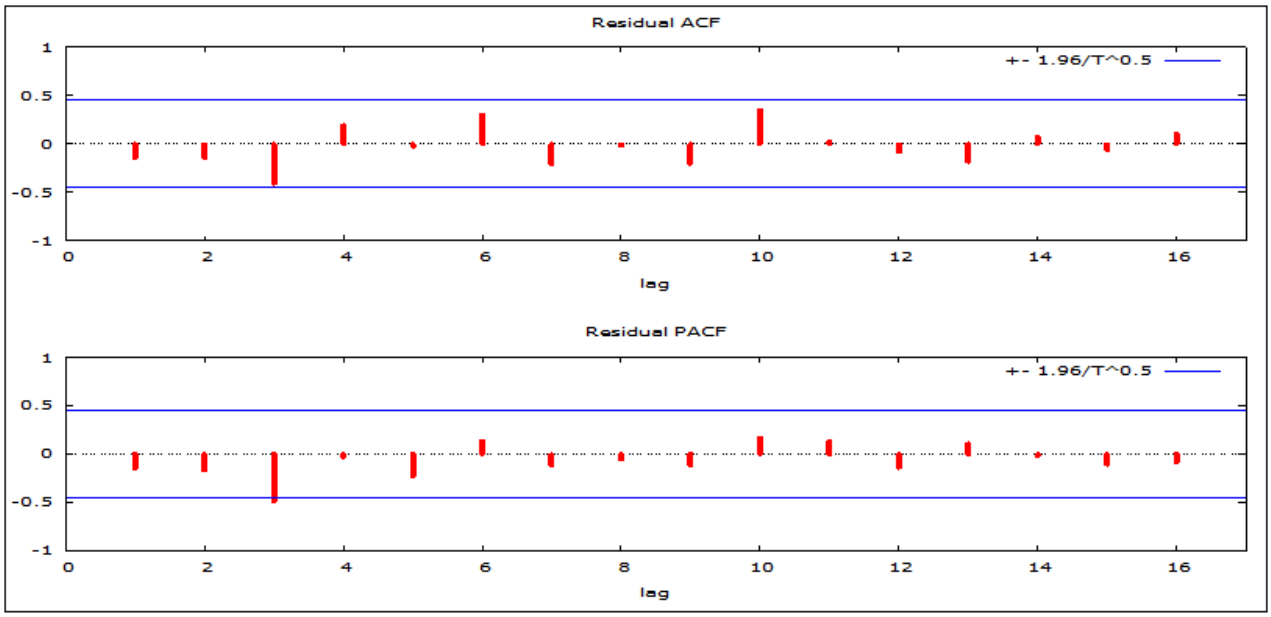

Fig-4: Correlogram of residuals for Typhoid Mortality

Table-3: Residual autocorrelation function

LAG	ACF	PACF	Q-stat.	[p-value]
1	-0.1485	-0.1485	0.4889	[0.484]
2	-0.1477	-0.1736	1.0013	[0.606]
3	-0.4141 *	-0.4909 **	5.2774	[0.153]
4	0.1952	-0.0327	6.2906	[0.178]
5	-0.0267	-0.2258	6.3109	[0.277]
6	0.3064	0.1373	9.1920	[0.163]
7	-0.2153	-0.1173	10.7338	[0.151]
8	-0.0160	-0.0574	10.7431	[0.217]
9	-0.2043	-0.1246	12.4076	[0.191]
10	0.3509	0.1740	17.8662	[0.057]
11	0.0297	0.1399	17.9103	[0.084]
12	-0.0858	-0.1422	18.3303	[0.106]
13	-0.1857	0.1089	20.6232	[0.081]
14	0.0766	-0.0250	21.0909	[0.099]
15	-0.0613	-0.1103	21.4662	[0.123]
16	0.1097	-0.0882	23.0672	[0.112]

***, **,* indicate significance at the 1%, 5%, 10% levels using standard error 1/T^0.5

Table-4: Forecasting for Typhoid Mortality Rate using ARIMA (0,1,0). For 95% confidence intervals, z (0.025) = 1.96

Obs	Typhoid	prediction	Std. error	95% interval
2019	undefined	28.5789	26.3388	(-23.0441, 80.2020)
2020	undefined	26.4737	29.4476	(-31.2426, 84.1900)
2021	undefined	24.3684	32.2583	(-38.8566, 87.5935)
2022	undefined	22.2632	34.8429	(-46.0277, 90.5540)

Fig-5: Correlogram of residuals for Typhoid Mortality

The time series plot and correllogram of the Typhoid Mortality data series on Fig-1 and Fig-2 respectively shows a strong evidence of a non stationary series, the non Stationarity was also confirmed with the help of the augmented dickey-fuller (ADF) test on Table-1, which tests the null hypothesis that Typhoid Mortality data series follows a unit root process which was accepted at *5%* alpha level. Thus, implying that the Typhoid Mortality data series is non stationary. By using first order differencing transformation, we obtained a t-statistic lesser than what was obtained

at $d = 0$, and a p-value lesser than *5%* alpha level. Thus, we select the condition that $d = 1$ and transform the data using first order differencing to make it stationary as seen on Fig-3. Eight tentative models were entertained, and the model with the minimum AIC, BIC and HQC, **ARIMA (0,1,0)** was chosen as the best model. To verify that the chosen ARIMA (0,1,0) is an appropriate model for typhoid mortality, a diagnostic check is done using residual ACF/PACF plot at different lags and testing the significance of the correlations up to 16 lags by Q statistic and respective p-values. Fig-4 above clearly shows evidence of random walk as the values are within the bounds and undulate about zero. Hence we uphold the first order differencing. Clearly, non of the correlations is significantly different from zero at a reasonable level. The ACF and PACF of the residuals also indicate good fit of the model and the respective p-values on Table-3 are greater than the α-value which is a desirable result. This proves that the selected ARIMA (0,1,0) is appropriate for modeling Typhoid Mortality in Delta State, Nigeria. Finally, the seven year forecast on Table-4 suggests that typhoid mortality in Delta State would be on the decrease for subsequent years.

CONCLUSION AND RECOMMENDATION

ARIMA (0,1,0) has been successfully used to forecast Typhoid in Delta State, Nigeria. Typhoid Mortality was found to be on a decrease in the forecasted period. However, in order to zero mortality due to typhoid from our society, government and health experts still need to put hands together to sanitize the system in terms of drugs manufacturing, bodies like NAFDAC (National Agency for Food and Drug Administration Control) needs to thoroughly monitor the drug market and ensure that drugs and food meets the necessary standards before they meet the people. There is need to sensitize the people on the use of traditional medicines and herbs.

REFERENCES

1. Wain, J., Hendriksen, R. S., Mikoleit, M. L., Keddy, K. H., & Ochiai, R. L. (2015). Typhoid fever. *Lancet* 385 (9973): 1136-45.
2. Anna, E. N. (2014). 3 Infectious Diseases Related to Travel. CDC health information for international travel 2014: the yellow book.
3. Typhoid Fever. cdc. gov. May 14, 2013. Retrieved 28 March 2015.
4. Typhoid vaccines. (2008). WHO position paper. Wkly Epidemiol Rec. (6): 49-59.
5. Crump, J. A., & Mintz, E. D. (2010). Global trends in typhoid and paratyphoid fever. *Clinical Infectious Diseases*, *50*(2), 241-246.
6. Magill, A. J., (2013). Hunter's tropical medicine and emerging infectious diseases (9th edition). London: Saunders/Elsevier. 568-572.
7. Cunha, B. A. (2004). Osler on typhoid fever: differentiating typhoid from typhus and malaria. *Infectious disease clinics of North America*, *18*(1), 111-125.
8. Anwar, E., Goldberg, E., Fraser, A., Acosta, C. J., Paul, M., & Leibovici, L. (2014). Vaccines for preventing typhoid fever. *Cochrane Database of Systematic Reviews*, (1).
9. Maxwell, O., Friday, A. I., & Chukwudike, N. C. (2019). A theoretical analysis of the odd generalized exponentiated inverse Lomax distribution. *Biom Biostat Int J*, *8*(1), 17-22.
10. Osuji, G. A., Obubu, M., & Nwosu, C. A (2016) Stock Investment Decision in Nigeria; A PC Approach, *World Journal of Multidisciplinary and Contemporary Research*, 2(1):1-11.
11. Osuji, G. A., Obubu, M., & Nwosu, C. A. (2016). Preconception sex selection using proper ovulation timing. *World Journal of Probability and Statistics Research*, *2*(1), 1-12.
12. Osuji, G. A., Okoro, C. N., Obubu, M., & Obiora-Ilouno, H. O. (2016). Effect of Akaike information criterion on model selection in analyzing auto-crash variables. *International Journal of Sciences: Basic and Applied Research*, *26*(1), 98-109.
13. Obubu, M., Valentine, O., Frederick, O., & Oluebube, N. L. (2017). Infant mortality; a continuing social problem in Northern Nigeria: Cox Regression Approach. *American Journal of Innovative Research and Applied Sciences*, *5*(5), 1-5.
14. Maxwell, O., Happiness, O. I., Alice, U. C., & Chinedu, I. U. (2018). An Empirical Assessment of the Impact of Nigerian all Share Index, Market Capitalization, and Number of Equities on Gross Domestic Product. *Open Journal of Statistics*, *8*(3), 584-602.