International Journal of Statistical Distributions and Applications
2025, Vol. 11, No. 3, pp. 98–105
https://doi.org/10.11648/j.ijsda.20251103.12

Research Article

A Statistical Analysis of Typhoid Fever in Aliade

Tamber Abraham Jighjigh[1, *], Tamber Dooshima Deborah[2], Ikpom Emmanuel Terhemba[3], Okafor Linus Uchenwe[4]

[1]Department of Statistics, Joseph Sarwua Tarka University, Makurdi, Nigeria
[2]Mathematics and Computer Science Department, Benue State University, Makurdi, Nigeria
[3]Benue State Health Insurance Agency, Makurdi, Nigeria
[4]Mathematics Department, Nigerian Defence Academy, Kaduna, Nigeria

Abstract

Typhoid fever remains a serious public health problem in many underdeveloped regions, with children and young adults (5–19 years) at highest risk. In Benue State, Nigeria, hospital records suggest a rising trend in reported typhoid cases, yet few studies have examined future incidence under current conditions. This research modeled and forecasted daily typhoid fever cases at Saint Vincent Hospital. Aliade (Gwer-East LGA) using an Auto-Regressive Integrated Moving Average (ARIMA) approach, thereby providing policymakers with data-driven projections for resource planning. The research also analyzed 501 daily observations of laboratory-confirmed typhoid cases collected from October 2019 to February 2021. Descriptive statistics (mean = 15.95 cases/day; SD = 8.81) characterized the series, which exhibited non-stationarity (ADF $p = 0.301$) at level. First and second differencing achieved stationarity (ADF $p = 0.01$). Autocorrelation (ACF) and partial autocorrelation (PACF) plots guided tentative model selection, and the final ARIMA (2,1,4) model was chosen based on minimum Akaike Information Criterion (AIC = 837.11). Model adequacy was confirmed via residual diagnostics (Ljung–Box test), and forecast accuracy was benchmarked against mean and naïve methods using mean absolute deviation and mean squared error. The ARIMA (2,1,4) model demonstrated strong fit (log-likelihood = –410.56; $\sigma^2 = 0.2998$) and outperformed benchmark forecasts in both MAD and MSE. A 15-year forecast indicates a slight upward trajectory in daily typhoid cases, suggesting continued burden without enhanced interventions. ARIMA modeling provides reliable short- and long-term forecasts of typhoid fever incidence. These projections can inform targeted prevention strategies, resource allocation, and public health planning to mitigate future outbreaks in the region.

Keywords

Typhoid Fever, Autoregression Integrated Moving Average (ARIMA), Optimal Model, Forecast, Prediction

1. Introduction

Typhoid Fever is a worldwide disease which creates a very serious public health problem. In many underdeveloped countries, the World Health Organization estimated that about one person out of every hundred (100) is suffering from typhoid fever. WHO also identifies typhoid as a serious public health problem whose incidence is highest in children and

*Corresponding author: abraham.tamber@uam.edu.ng (Tamber Abraham Jighjigh)

Received: 7 May 2025; **Accepted:** 21 May 2025; **Published:** 24 December 2025

young adults between 5 to 19 years old [17, 9]. Typhoid fever is a systematic disease caused by the dissemination of Salmonella Typhi or Salmonella Para-typhi mainly characterized by fever. It is called Salmonella Typhi because of its common similarity to typhus. It is clearly defined pathologically as a unique disease of its own, it is a common worldwide illness, transmitted by the ingestion of food or water contaminated with the faces of an infected person, which contains the bacterium Salmonella enterica. The bacteria then perforate through the intestinal wall and are Phagocytosed by macrophages [9, 16, 12, 5-7].

Osagi reveals that typhoid fever is an infection that is usually spread by contamination of food, milk, or water supply with Salmonella Typhi (S. Typhil) either directly by sewage or indirectly by flies or by faulty personal hygiene [12]. Symptomless carriers harboring the germ in the gall bladder and excreting it in their stools are the main sources of outbreaks of disease in this country. The average incubation period is 10-14 days. A progressive feverish illness marks the onset of disease which develops as the germ invades typhoid tissues, including that of the small intestine to profuse diarrheal (pea soup) stools which may become frankly hemorrhagic, ultimate recovery usually begins at the end of the third week. A rose-colored rash may appear on the upper abdomen and back at the end of the first week [10, 11, 8, 14, 13, 2].

Most researches conducted on the disease were done by medical practitioners, who majorly reported the rate of occurrence and major cause of typhoid fever in the region, little or no research from the region focuses on the future incidence of the disease if the present situation exists. Such information is important for policymakers to be able to plan ahead of time. Thus, this research seeks to investigate the pattern of occurrence of typhoid fever in the region over different periods of the year to forestall or reduce the danger posed by typhoid fever and provide adequate information about the past, present and future of the disease by adopting statistical tools the Auto Regression Integrated Moving Average (ARIMA) model. ARIMA model is a class of models used to forecast stationary time series data. A time series data is said to be stationary if the statistical properties like mean, variance, etc. are constant over time. ARIMA model comprises three parts namely: Auto Regression (A. R), Integrated (I), and Moving Average (M. A) [15, 3]. In this research work, we use the ARIMA model because time series data of typhoid fever obtained from Saint Vincent Hospital Aliade is stationary. ARIMA (Box-Jenkins) Model is a mathematical model designed to forecast data from specified time series estimation and diagnostic checking the identification process involves identifying or determining the appropriate order of the Auto-Regressive Integrated moving Average (ARIMA) (p,d,q) required to capture the pattern of the time series using the Autocorrelation function (ACF) and Partial Autocorrelation function (PACF). After the appropriate ARIMA (pdq) model, the parameter of the model will be estimated after identification and estimation, it is required to determine whether the model is adequate, and this is done through diagnostic checking [1].

Over the last decade, there has been an alarming increase in the rate of recorded cases of typhoid fever in our various hospitals especially in Benue State. This has led to a series of questions in the minds of the citizenry and the professionals in the medical field in general. Ordinary, since typhoid fever is generally associated with a lot of symptoms such as fever headache, cough, etc. These however make it a complicated disease with a very serious economic and social effect on the victims. This study seeks to predict the prevalent rate of typhoid fever by: developing an ARIMA model using the time series data of St Vincent Hospital; predicting the rate of prevalent in the future, using the developed model and fitting the appropriate model to capture the pattern or behavior of typhoid fever and its types to aid the decision makers for decision making in the health sector.

2. Methodology

2.1. Research Method

This research seeks to find the rate of typhoid fever to understand the underlying reasons for such rates, the quantitative research method is used. The method is adequate for the study because it provides useful insight into the data and a possible prediction about typhoid fever and its types in society.

2.2. Area of the Study

This research was conducted in Saint Vincent Hospital at Aliade Gwer-East Local Government Area (LGA), Benue State.

2.3. Method of Data Collection

The data collection here was basically from a secondary source, the medical report of a patient on the rate of typhoid fever and its type from Saint Vincent Hospital. The data was collected with due permission. The data was collected for seventeen (17) months.

2.4. Method of Data Analysis

The data used in the study was presented using tables, graphs and charts. The graph shows the trend of typhoid fever and its types for the period of seventeen (17) months. The R Statistical Computing package was used to analyze the time series data of Typhoid Fever from Saint Vincent Hospital Aliade Gwer-East LGA Benue State.

Trend is one of the basic components of a time series. The trend is the general movement (upward or downward) of a series over a long period. There are several methods of estimating the trend of a given series. However, least square estimation (LSE) and moving average (MA) are commonly used in many literatures. The LSE method will be used in this work and the trend line equation is given as:

$$Y_t = \alpha + bX_t \tag{1}$$

Where,
Y_t is the series
α is the intercept
b is the slope
X_t is the time variable.

$$\alpha = \bar{Y}_t - \hat{b}\bar{X}_t \tag{2}$$

$$b = \frac{n\sum_t X_t Y_t - \sum_t X_t \sum_t Y_t}{n\sum_t X_t^2 - (\sum_t X_t)^2} \tag{3}$$

2.5. ARIMA Model (Box and Jekins)

The Box and Jenkins approach involves three stages: identification, estimation and diagnostic checking. The identification process consists of identifying or determining the appropriate order of the Auto-Regressive Integrated Moving Average (ARIMA) (p,d,q) required to capture the pattern of the time series using the autocorrelation function (ACF) and partial autocorrelation function (PACF). After the appropriate ARIMA (p,d,q) model, the model's parameters will be estimated. After identification and estimation, it is required to determine whether the model is adequate, and this is done through diagnostic checking.

The ARIMA model is a class of models used for forecasting stationary time series data. A time series data is said to be stationary if the statistical properties, like mean, variance, etc., are constant over time. ARIMA model comprises three parts, namely: Auto-Regressive (AR), Integrated (I) and Moving Average (MA) can be classified generally as an ARIMA (*p, d, q*) model, where;

p is the order of the autoregressive term;
d is the differencing order required to achieve stationarity;
q is the order of the moving average term.
The ARIMA (*p, d, q*) can be expressed mathematically as

$$X_t = \nabla^d x_t \tag{4}$$

$$X_t = \mu + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \ldots + \varphi_p x_{t-p} - \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_p \varepsilon_{t-p} - \varepsilon_t \tag{5}$$

Where X_t follows ARIMA (*p, d, q*) and follows ARMA (*p, q*) of [4].

If the series is stationary, *d* will be equal to zero and leads to the ARIMA (*p, q*) model. However, if the series is differenced times to achieve stationarity, that leads to the ARIMA (*p, d, q*) model. To determine whether or not the series is stationary, the test for the presence of a unit root A unit root helps to identify some features of a series and when a unit root is present in a series, it is characterized as stationary; otherwise, it is non-stationary.

***Table 1**. The Comparison between the ACF and PACF plots.*

MODEL	ACF	PACF
AR (*p*)	Exponential decay/or damped sinusoid	Cuts off after lag *p*
MA (*q*)	Cuts off after lag *q*	Exponential decay/or damped sinusoid
ARMA (*p,q*)	Exponential decay/or damped sinusoid	Exponential decay/or damped sinusoid

The identified model using the ACF and PACF plots is called the tentative model, which will be used as a starting model to search for an adequate model with the least Akaike Information Criteria (AIC). After the model with the least AIC has been identified, the model will be fitted by estimating the parameters of the model. The significance of the model coefficients will be examined using *the z-test*.

$$Z = \frac{Coefficient - 0}{S.E(Coefficient)} \tag{6}$$

2.6. Forecasting Methods

The future occurrence of typhoid fever in the region will be forecasted using the identified ARIMA models. However, the following forecasting measures will be used as benchmarks.

2.7. Measurement of Forecasting Error

The main purpose of Measurement of Forecasting Error is to examine and evaluate different forecasting error measurements, traditional measurements of forecast errors are: mean absolute deviation (MAD), mean square error (MSE), together with new suggested error and bias measurement: “periods in stock” (PIS) and “number of shortages” (NOS).

2.8. Mean Method

The mean method assumes that all future values are constant by using the mean of the historical data as the forecast of all future values.

$$X_1+X_2+\ldots+X_t \tag{7}$$

Given the historical data, the forecast value using the mean method is expressed as;

$$\hat{X}_{t+h\backslash T} = \bar{X} = \frac{(X1 + X2 + \cdots + Xt)}{T} = \frac{\sum_1^t X}{T} \tag{8}$$

2.9. Naïve Method

The naïve method assumes that all future values are constant by taking the last observation of the historical data as the forecast of all future values.

2.10. Data Presentation

The daily data was collected for seventeen (17) months from October 2019 to February 2021, as depicted in Table 1 below:

Table 2. *St Vincent Hospital Aliade, Typhoid positive cases between October 2019 and February 2021.*

2019			2020												2021	
oct	**Nov**	**Dec**	**Jan**	**Feb**	**March**	**April**	**May**	**Jun**	**July**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**	**Jan**	**Feb**
	3	6	4	2	7	7	4	3	5	2	5	4	4	5	4	6
	7	5	6	8	7	2	2	8	6	5	2	2	8	3	2	8
	3	2	4	6	4	2	5	2	5	2	5	3	4	3	3	6
	7	6	3	14	8	2	4	3	8	8	3	2	7	3	3	6
	6	4	11	6	7	2	3	4	2	3	3	3	4	4	7	6
	9	4	2	9	12	4	3	5	2	2	2	4	6	4	3	3
	10	7	2	2	3	4	4	3	8	4	6	3	3	4	3	5
	6	3	5	11	3	3	2	3	7	3	3	8	2	2	4	3
	4	2	3	12	10	5	6	2	2	3	3	4	5	6	3	3
	2	2	2	2	5	6	4	7	7	8	4	3	4	4	3	3
	2	5	4	2	7	3	3	5	5	7	5	4	7	3	4	3
	8	2	2	7	9	2	3	3	4	4	5	3	9	2	4	3
	4	2	4	9	7	4	2	8	5	4	3	3	5	9	5	3
	2	7	6	14	2	2	2	8	6	4	4	5	5	5	3	3
	2	4	2	10	9	5	2	7	5	2	5	3	2	2	3	3
	7	7	4	8	10	5	4	3	7	6	4	6	6	2	3	3
10	5	5	6	15	5	4	6	2	2	6	2	3	5	2	4	3
8	3	5	7	6	11	6	2	3	4	8	3	3	5	3	2	3
8	5	3	5	15	4	5	5	4	3	5	5	3	6	4	7	3
6	3	5	6	3	3	3	3	6	7	7	3	5	2	3	4	3
4	2	3	3	2	2	2	2	4	5	3	4	6	4	2	6	3
16	5	3	2	2	3	8	3	5	2	5	3	7	4	4	7	3
7	2	2	10	7	4	2	6	4	5	2	6	4	2	7	2	3
4	2	2	5	7	7	7	7	3	7	3	4	6	5	5	3	3
3	4	5	5	7	2	9	3	4	3	4	3	3	7	8	2	3
2	3	7	6	7	15	4	3	4	3	5	2	4	2	2	6	3

2019			2020												2021	
oct	Nov	Dec	Jan	Feb	March	April	May	Jun	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb
7	6	22	8	7	7	4	5	4	8	5	3	4	6	3	7	3
12	3	22	22	7	2	4	5	4	5	5	4	4	3	4	3	3
5	6	22	22	8	5	2	4	6	5	4	4	5	4	7	3	
4	6	22	22	4	4	2	4	3	5	4	4	2	4	4		
7	22	22	7	3	5	5	4	4	4							

Sources: St Vincent Hospital Aliade (2021)

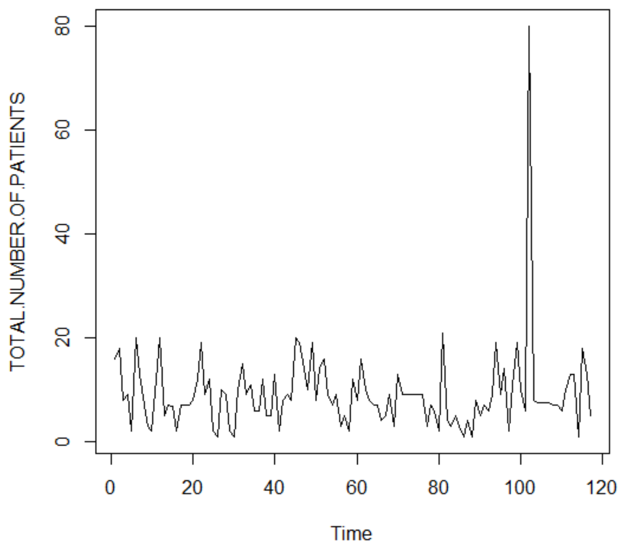

Figure 1. *St Vincent Hospital Aliade Typhoid positive cases between October 2019 and February 2021.*

3. Result

Table 3. *Descriptive statistics.*

	Descriptive statistics
Nobs	501
Nas	0
Minimum	1
Maximum	31
1. Quartile	8
3. Quartile	24
Mean	15.948104
Median	16
Sum	7990
SE Mean	0.393473
LCL Mean	NaN
UCL Mean	NaN
Variance	77.565301
Stdev	8.807117
Skewness	-0.031766
Kurtosis	-1.201302

3.1. Testing for Stationarity

Figure 1 clearly indicates that the series is stationary. The verification to confirm that the data is stationarity is shown in Table 4, Table 5 and Table 6 below: Augmented Dickey-Fuller (ADF) Test.

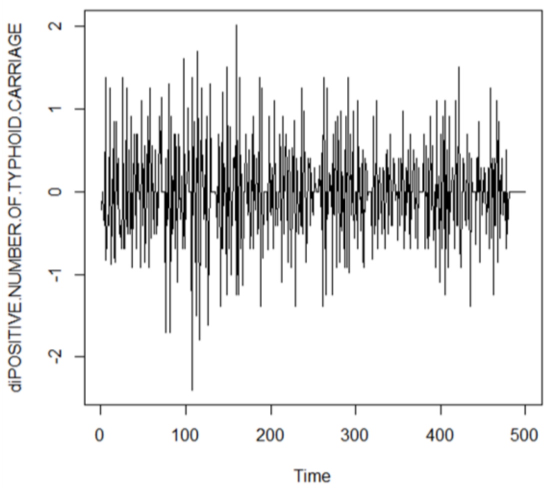

Figure 2. *The first difference of the positive number of typhoid fevers.*

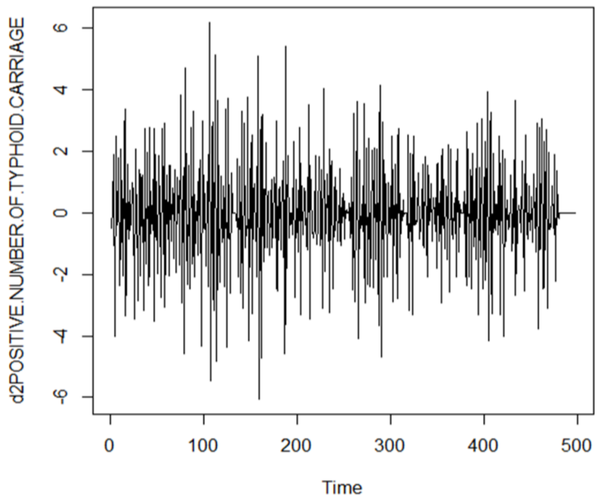

Figure 3. *The second difference in the positive number of typhoid fevers.*

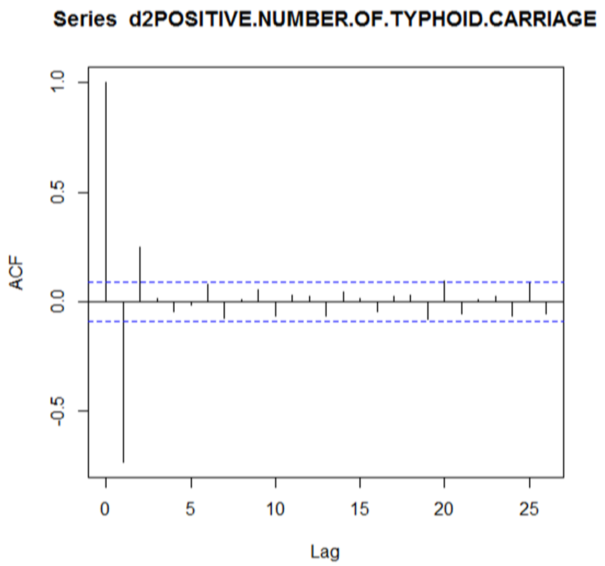

Figure 4. *The ACF of the positive number of typhoid fevers.*

Table 4. *The normal ADF test of the data.*

Dickey-Fuller	Lag order	p-value
= -6.6841	=7	= 0.301

Table 5. *The first differences ADF test of the data.*

Dickey-Fuller	Lag order	p-value
= -13.328	=7	=0.01

Table 6. *The second differences ADF test of the data.*

Dickey-Fuller	Lag order	p-value
= -19.118	=7	=0.01

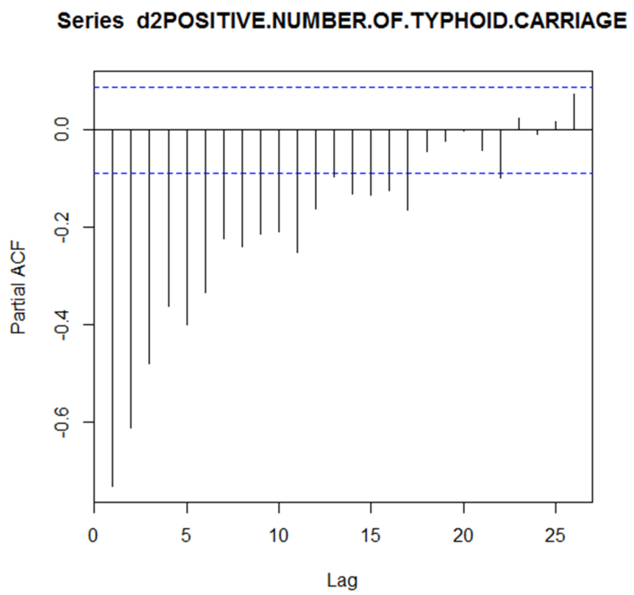

Figure 5. *The Partial ACF positive number of typhoid fevers.*

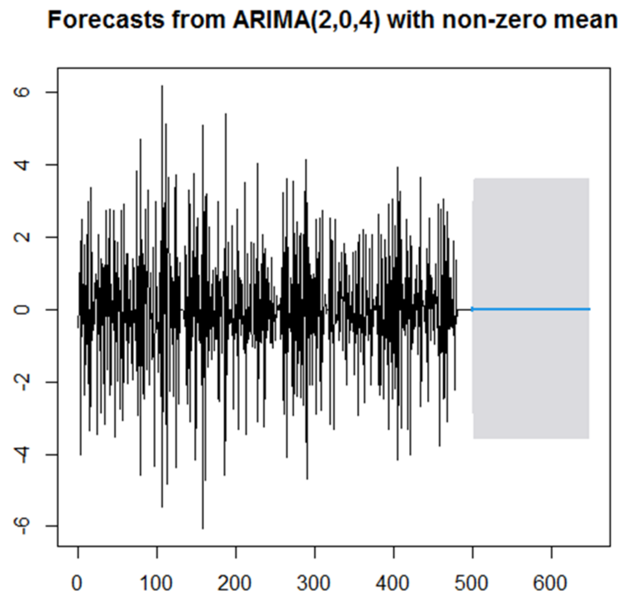

Figure 6. *The Forecast of the Positive Typhoid Cases.*

Table 3 and Figure 3 show that the p-value of the Augmented Dickey-Fuller Test of second differences confirmed that the data is stationarity.

Table 7. *The model identification.*

Coefficients						
ar1	**ar2**	**ma1**	**ma2**	**ma3**	**ma4**	**Mean**
0.2787	-0.1629	-2.8079	2.8162	-1.1806	0.1727	0
s.e.0.1357	0.0744	0.0445	0.2555	0.3112	0.1153	0
sigma^2 estimated as 0.2998: log likelihood=-410.56						
AIC=837.11 AICc =837.41 BIC =870.8						

3.2. Model Fitting

The optimal ARIMA model for predicting typhoid cases is ARIMA (2,0,4) with a non-zero mean as depicted in equation (9).

$$Y_t=0.2787Y_{t-1} -0.1627Y_{t-2}-2.8079\mathcal{E}_t \tag{9}$$

4. Discussion

The results of this study indicated a stationary trend in the number of days over the period under study. From 17th October 2019 to February 2021 the total number of patients who tested positive for typhoid was 2387 and the highest month of occurrences December, 2019 which recorded 218 positive cases and October 2019 recorded the lowest number of occurrences of 103 positive cases. December, 2019 also recorded the highest in day 22 it recorded five times.

A stationary series must have a constant mean, constant variance and constant autocorrelation structure. The time series plot of the daily reported cases of the positive number of typhoid carriages from October 2019 to February 2021 as shown in Figure 1 above; shows that the time series is stationary. The fact that there are irregular fluctuations in the plot was not a clear indication that the series was non-stationary. The verification to confirm that the data is stationarity is shown in Table 1 above Figure 2 takes the first differences and Figure 3 takes the second differences it is observed that due to the Augmented Dickey-Fuller Test of second differences and the p-value and also the ACF (Figure 4) and PACF (Figure 5) show and confirmed respectively that, the data is stationarity. Also, since the p-value = 0.01 is less than 0.05 for the typhoid cases period one, the null hypothesis is rejected at = 0.05 and we conclude that the typhoid cases are stationary.

The ARIMA model in equation (9) has been proven to be adequate based on the Ljung Box test and Ljung Box plots. The models were used for forecasting the future incidence of typhoid fever as shown in Figure 6 and the forecast indicates a slight increase in the future occurrence of typhoid fever in the region. And we forecast for (15) fifteen years.

5. Summary and Conclusions

Historical data of 503 days was collected and the Box and Jenkins (ARIMA) model was successfully employed to develop a forecasting time-dependent series model. The ARIMA (2,0,4) model developed was used to predict the typhoid fever future prevalence. The forecast indicates a slight increase in the future occurrence of typhoid fever in the region. A forecast over (15) fifteen years was made. The generated forecasts were adequately accurate based on the accuracy of forecasts obtained from the various models built in this research work.

It is concluded that the Aliade District has conditions that accommodate the vector that causes typhoid. The prevalence of typhoid is among the population. Typhoid significantly affects households since it is possible to experience multiple and repeated attacks in a year. In this case, the district must be given priority attention in annual budgets to enable it to combat the disease. In particular, there is a need for strong collaboration among major stakeholders including the Government, District Assemblies, Non-Governmental Organizations and the community to devise holistic, effective, and cost-saving methods for the prevention, control and treatment of typhoid. Although the use of good water for example, good water is identified as the major method of protection due to its availability and affordability for many households, the efficacy of some of these numerous brands on the market may be questionable.

6. Recommendations

It is evident from the historical data used in this research and the results obtained that the incidence of typhoid fever is influenced or affected by seasonal factors as the incidence of typhoid fever tends to be high at some periods of the year and low at some other period. A proper enlightenment program is recommended across genders to enlighten the people of the region about the seasonal pattern and educate them on preventive measures, especially during the period when high incidence is expected. Further research is needed to determine the factors responsible for the high and low incidence at different periods of the year in the region.

Abbreviations

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criteria
ARIMA Model	Auto Regression Integrated Moving Average Model
LGA	Gwer-East Local Government Area
LSE	Least Square Estimation
MAD	Mean Absolute Deviation
MSE	Mean Square Error
NOS	Number of Shortages
PACF	Partial Autocorrelation Function
PIS	Periods in Stock
LCL	Lower Class Limit
UCL	Upper Class Limit

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Adeboye, N. O., & Ezekiel, I. D. (2018). On Time Domain Analysis of Typhoid Morbidity in Nigeria. American Journal of Applied Mathematics and Statistics, 6(4), 170–175.

[2] Akawu, C., Ikusemoran, M., & Ballah, A. (2018). Analysis of Spatial Patterns of Typhoid Prevalence in Borno State, Nigeria. Academic Research International, 9(2), 24–37.

[3] Babajide, S., & Perry, B. (2015). Meeting the Millennium Development Goal for typhoid by 2015: A Time Series Analysis of typhoid Cases in Ogun State, Nigeria. African Journal of Malaria and Tropical Diseases, 3(9), 245 – 259.

[4] Box, G. E., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control.

[5] Fana, S. A., Bunza, M. D., Anka, S. A., & Imam, A. U. (2015). Prevalence and Risk Factors Associated with Malaria Infection among Pregnant Women in a Semi-Urban Community of North-Western Nigeria. Infectious Diseases of Poverty, 4(24), 1-5.

[6] Houmsou, R. S., Amuta, E. U., Sar, T. T., & Adagba, A. H. (2011). Malaria Infection among Patients Attending a Nigerian Semi-Urban Based Hospital and Performance of HRP-2 PF Rapid Diagnostic Test (RDT) in Screening Clinical Cases of Plasmodium Falciparum Malaria. Translational Biomedicine, 2(1: 5), 1 – 5.

[7] Ibor, U. W., Okoronkwo, E. M., & Rotimi, E. M. (2016). Temporal Analysis of Malaria Prevalence in Cross River State, Nigeria. E3 Journal of Medical Research, 5(1), 1 – 7. 56.

[8] Ishaq, A. A., & Murtala, U. M. (2017). Spatio-Temporal Trends of Typhoid Fever Among Youths Attending Muhammad AbdulahiWase Specialist Hospital in Kano Metropolis, Nigeria. Bayero Journal of Applied Sciences, 10(2), 115 – 121.

[9] Nigeria Malaria Fact Sheet (2011). Nigeria Malaria Fact Sheet. United States Embassy in Nigeria. http://nigeria.usembassy.gov

[10] Obimakinde, E. T., & Simon-Oke, I. A. (2017). The Prevalence of Malaria Infection Among Patients Attending the Health Centre of the Federal University of Technology, Akure, Nigeria. International Journal of Tropical Disease & Health, 27(4), 1 – 7.

[11] Onah, I. E., Adesina, F. P., Uweh, P. O., & Anumba, J. U. (2017). Challenges of Typhoid Elimination in Nigeria; A Review. International Journal of Infections Disease and Therapy, 2(4), 79 – 85.

[12] Osagi, H. D., Anouche, M. R., & Oluchukwu, H. E. (2015). A Study of the Prevalence of Typhoid Fever Co-Infection in Abakaliki, Nigeria. African Journal of Typhoid and Tropical Diseases, 3(9), 162 – 166.

[13] Owoeye, D. O., Akinyemi, J. O., & Yusuf, O. B. (2018). Decomposition of Changes in Typhoid Prevalence Amongst Under-Five Children in Nigeria. Typhoid World Journal, 9(3), 1 – 6.

[14] Ozofor, N. M., &Onos, C. N. (2017). Statistical Analysis of Reported Cases of Typhoid in Parklane Specialist Hospital in Enugu Urban for the Period of 2009 – 2015. Researchjournali's Journal of Mathematics, 4(3), 1 – 10.

[15] Sekubia, A. C., &Mensah, E. K. (2019). Analysis of Typhoid Fever Surveillance Data, Cape Coast Metropolis, 2016. Acta Scientific Medical Sciences, 3(2), 2–8.

[16] Ukaegbu, C. O., Nnachi, A. U., Mawak, J. D., &Igwe, C. C. (2014). Incidence of Concerrent Typhoid Fever Infections in Febrile Patients in Jos, Plateau State Nigeria. International Journal of Scientific & Technology Research, 3(4), 157 – 161.

[17] World Health Organization. (2013). Global action plan for the prevention and control of noncommunicable diseases 2013–2020. Geneva: World Health Organization.