



International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

Forecasting Typhoid Fever Trend Using Arima Models

Adamu Abdulmumin¹

Department of Mathematics and Statistics, Faculty of Science, Yobe State University, Damaturu, Nigeia

Email: adamuabdulmumin34@gmail.com

DOI : <https://doi.org/10.55248/gengpi.07.1026.3114>

ABSTRACT

Typhoid fever remains a significant public health concern globally, with approximately 21 million cases and 200,000 deaths annually. Despite efforts to control the disease, its incidence exhibits complex patterns, influenced by climate variables, sanitation practices, vaccination coverage, population density, and economic factors. The impact of this disease is particularly severe in rural areas, where there's lack of sanitation hygiene. Preventing this disease in rural areas is significant challenges due to the lack of hospitals and inadequate disease prevention. Time series analysis assist health authorities in anticipating and preparing for potential outbreaks, leading to more effective and timely interventions. This study examined the trend of the reported cases of Typhoid fever. According to data on the reported cases of typhoid fever retrieved from dataset expository called Kaggle used for the analysis. The data space for the period of Nine (9) years, started from January, 2015 to December, 2023. Seasonal ARIMA models were developed for the typhoid fever trend, but ARIMA (2, 0, 0) with non-zero mean was selected as the best performed model based on the AIC value. Furthermore, the results of the ARIMA (2, 0, 0) forecast shows an increased in the cases of the typhoid fever due to the series trend. And this study Develop national guidelines for typhoid fever prevention and control.

Keyword: Time series, Trend, Typhoid Fever, Time Series Model, and Forecasting

1. INTRODUCTION

Typhoid fever persists as a global health problem, with an estimated 11 to 21 million cases and over 128,000 deaths annually worldwide. While progress has been made in reducing its burden, challenges such as antibiotic resistance and access to safe water and sanitation persist, underscoring the on-going need for comprehensive prevention and control measures. Typhoid fever has likely been present since ancient times, although it was not until relatively recently that its cause and transmission were understood. Historical records suggest that outbreaks resembling typhoid fever occurred in ancient civilizations such as Egypt, Greece, and Rome. The understanding of typhoid fever improved significantly in the 19th century. In the mid-1800s, British physician William Budd proposed the theory that typhoid fever was transmitted through contaminated water and food. His work laid the foundation for understanding the importance of sanitation in preventing the spread of the disease.

Typhoid refers to a bacterial infection caused by the bacterium *Salmonella enterica* serotype Typhi. The disease is primarily transmitted through the ingestion of contaminated food or water that has been contaminated by the feces of an infected person. Typhoid fever is characterized by symptoms such as high fever, abdominal pain, headache, and gastrointestinal issues. It is a serious illness that can lead to complications if not treated promptly with antibiotics. Typhoid is a significant public health concern, especially in areas with poor sanitation and limited access to clean water. Vaccines are available for the prevention of typhoid fever, and measures such as proper hygiene, sanitation, and safe food handling play essential roles in controlling its spread. Symptoms typically include sustained fever, headache, abdominal pain, and sometimes a rash. Severe cases can lead to complications affecting multiple organs and can be fatal if left untreated. Typhoid fever is prevalent in areas with poor sanitation and hygiene practices. Vaccination and proper sanitation are key preventive measures.

2. RELATED WORK

This section deals with review of previous researches carried out on typhoid fever cases using time series models. Such as Jiaqi Gao et al., (2020) studied time series analysis of cumulative incidences of typhoid and paratyphoid fevers in China using both Grey and SARIMA models. Cumulative incidence of both fevers displayed significant decrease nationally, displaying a drop of 73.9% for typhoid and 86.6% for paratyphoid in 2016 compared to 2004. Cumulative incidence fell in all age subgroups and the 0–4 years-old children were the most susceptible ones in recent years. A cluster of three south western provinces (Yunnan, Guizhou, and Guangxi) were the top high-incidence regions. Grey model GM (1, 1) and seasonal autoregressive integrated moving average (SARIMA) model were employed to extract the long-term trends of the diseases. The predictive models for both fevers forecasted cumulative incidence to continue the slightly downward trend and maintain the cyclical seasonality in near future years. Such data-driven insights are informative and actionable for the prevention and control of typhoid and paratyphoid fevers as serious infectious diseases.

Paul *et al.*, (2023) worked on buys ballot approach in time series analysis of typhoid fever. The fever is caused by Salmonella typhi bacteria. Therefore, they had adopted the Buys Ballot approach by using time series analysis for the estimation of trend of typhoid fever disease from the year 2011 to 2020 and predict the future occurrence of the disease in Mubi South Local Government Area (LGA) of Adamawa State, Nigeria.

Jillian *et al.*, (2022) studied on rainfall anomalies and typhoid fever in Blantyre, Malawi. A quasi-Poisson generalised linear modelling framework was used to explore the relationship between rainfall and typhoid incidence at biologically plausible lags of 1–4 weeks. They found a protective effect of rainfall anomalies on typhoid fever, at a two-week lag ($P = 0.006$), where a 10 mm lower-than-expected rainfall anomaly was associated with up to a 16% reduction in cases (95% CI 7.6, 26.5). Extreme flooding events may cleanse the environment of *S. Typhi*, while unusually low rainfall may reduce exposure from sewage overflow.

Masoud *et al.*, (2020) investigated on spatial and spatiotemporal patterns of typhoid fever and investigation of their relationship with potential risk factors in Iran, 2012–2017. To explore the spatial and temporal patterns of typhoid fever and examine their relationship with climatic and socio-economic variables; they have employed a spatiotemporal zero-inflated Poisson (ZIP) model in a Bayesian framework.

Qi *et al.*, (2021) worked on impact of temperature and rainfall on typhoid/paratyphoid fever in Taizhou, china: effect estimation and vulnerable group identification. Significant temperature-typhoid/paratyphoid fever and rainfall-typhoid/paratyphoid fever associations were found in both genders and those aged 0–4 years old, 15–60 years old, farmers, and children. Characterized with a lagged, nonlinear, and cumulative effect, high temperature and rainfall could increase the risk of typhoid/paratyphoid fever in regions with a subtropical climate. Public health interventions such as early warning and community health education should be taken to prevent the increased risk of typhoid/paratyphoid fever, especially for the vulnerable groups.

Olumuyiwe *et al.*, (2020) worked on global stability analysis of typhoid fever model. The computed basic reproductive number R_0 was used to establish that the disease-free equilibrium is globally asymptotically stable when its numerical values are less than one while the endemic equilibrium is locally asymptotically stable when its values are greater than one. In addition, the Lyapunov function was applied to investigate the stability property for the (DFE). The model was numerically simulated to validate the results of the analysis.

Kaigama *et al.*, (2023) worked on time series analysis for the treatment of typhoid (enteric) fever in Maiduguri. Using autocorrelation and partial autocorrelation function (ACF and PACF), the data can also be utilized to determine the model's order. The model obtained is subjected to model diagnostics in order to determine its efficiency and the model is used to forecast the typhoid fever. From the forecast graph shows that there may be a decrease in future years due to the pattern of the series the impression, they obtain from the graph is that predicted series seems to be trend upward and then downward.

Okechukwu *et al.*, (2023) proposed on statistical analysis on diagnosed cases of malaria and typhoid fever in Enugu-Nigeria. Typhoid fever, on the other hand, is caused by the bacterium Salmonella Typhi transmitted through contaminated food and water. These illnesses typically present with flulike symptoms such as fever, body aches, headaches, and fatigue. Malaria and typhoid disproportionately affect vulnerable populations such as children, pregnant women, and those living in impoverished areas with limited access to healthcare. Misdiagnosis and mismanagement can lead to severe complications such as cerebral malaria, kidney failure, and intestinal perforation in the case of typhoid. A sample of two hundred and forty nine patients in Enugu was taken and data obtained were analysed using binary logistic regression. The data obtained are presented in tables and bar charts. The odds ratio, relative risk, chi-square statistic for test of independence of Malaria and Typhoid, trend are calculated.

Kim *et al.*, (2023) worked on associations of water, sanitation, and hygiene with typhoid fever in case-control studies: a systematic review and meta-analysis. Updated the previous review conducted by Brockett *et al.* they included new findings published between June 2018 and October 2022 in Web of Science, Embase, and PubMed. They used the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) tool for risk of bias (ROB) assessment. they classified WASH exposures according to the classification provided by the WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation, and Hygiene (JMP) update in 2015. They conducted the meta-analyses by only including studies that did not have a critical ROB in both Bayesian and frequentist random-effects models.

Egolo *et al.*, (2022) worked on forecasting the trend of typhoid fever in selected teaching hospital in Akwa, Anambara state, the study findings demonstrated health disparities within the community as a result of terrible settings, insufficient understanding of the risk factors for typhoid fever, inadequate hygiene practices, and necessary supplies such as water, based on data obtained from the participants. The quadratic trend model, which had the lowest MAD and MSD values, was found to be the best fit for the data when fitting relevant trend models to typhoid fever cases in Akwa, Anambara state. The quadratic trend model was the best fit for assessing the progression of the typhoid fever outbreak in selected teaching hospital in Akwa, Anambara state. It has been persuasively proved that Typhoid fever cases are expected to rise during the next 12 months, from January to December 2022, according to the findings.

Mathew and Idi (2024) studied on statistical model for occurrence of typhoid fever in Takum, Taraba state. This study successfully applied time series analysis to forecast Typhoid Fever cases at General Hospital Takum from 2012 to 2021. By employing an ARIMA(4,1,1) model, the analysis addressed the non-stationary of the data and ensured the robustness of the model through rigorous diagnostic testing. The model's effectiveness was demonstrated through its significant parameter estimates and favourable diagnostic results, including tests for serial correlation, normality, and heteroscedasticity. The model was evaluated on the basis of AIC, AICC, BIC and it was found that ARIMA (4, 1, 1) is the best. All parameters were estimated by maximum likelihood and all the model coefficients were statistically significant and were within the bounds of stationarity and invertibility. An extensive diagnostic study verified the strength of the model. The Box-Ljung test for serial correlation, Shapiro-Wilk test for normality and ARCH-LM test for heteroscedasticity confirmed the suitability of the ARIMA (4, 1, 1) model. The model was then applied to make predictions for Typhoid Fever

for the next period of 24 months. The predictions, summarized in the tables and figures of the study, show how effective the model is in predicting future trends in Typhoid Fever cases, thus making the model reliable for public health planning and response.

En Muhammad *et al.*, (2020) investigate on burden of typhoid fever Klang valley, Malaysia, in total, 507 typhoid fever cases were reported in Klang Valley; however, only 265 cases were confirmed by culture tests. The crude incidence rates of confirmed cases were between 0.5 to 0.7 but peaked at 1.42 per 100,000 populations in 2015. Most typhoid fever cases were observed among men (55.6%), individuals aged 21 to 30 years (27.6%), Malaysians (86.3%) and individuals of Malay ethnicity (52.1%). The reporting coverage of confirmed cases was 78.9% and non-reporting coverage of unconfirmed typhoid cases was 79.5%. The predictive value positive (PVP) was 89.3, and 7.5% were detected as MDR Salmonella typhi. Statistical significance was found in gender, citizenship and ethnicity regarding MDR Salmonella typhi ($p = 0.004$, $p = 0.008$ and $p = 0.034$, respectively). The local transmission of typhoid is still prevalent in the Klang Valley despite rapid urbanization and development in recent years. These findings are essential for policy makers to plan and implement focused and effective preventative activities to curb typhoid infection in urban areas.

Obulezi *et al.*, (2023) investigate on statistical analysis on diagnosed cases of malaria and typhoid fever in Enugu Nigeria, This work considers the occurrence of common fevers in school environment. All the independent variables in the model are significant except the moderate infection. For the dummy variable logistic regression: mild and moderate infections as compared to severe infection increases the incidence of typhoid fever though mild infection has more effect; also mild infection as compared to moderate infection increases the incidence of typhoid fever but severe infection does not; and moderate infection as compared to mild infection increases the incidence of typhoid fever but severe infection does not. Also male patients are more likely to be infected by typhoid fever than the female and those 18 years and older are also more likely to be infected. There is a significant association between Typhoid fever and gender, and also for Typhoid fever and age. Also, there is no significant association between Malaria and gender, and Malaria and age. In addition, there is a significant association between malaria and Typhoid fever. This means that there is likelihood that Malaria patients have typhoid fever. More so, male patients are less likely to be infected with malaria fever (both for mild, moderate and severe infection) than the females. Furthermore, the risk of being infected by Typhoid fever is less likely in males than in females; and the risk of having Typhoid fever is more likely in patients aged 18 or more than in patients less than 18 years. Finally, the time series analysis showed that the incidence of malaria and typhoid fever in relation to month has a linear relationship.

Birkhold *et al.*, (2023) investigated on characterization of typhoid intestinal perforation in Africa, in this large case series embedded within a broader typhoid surveillance effort, we describe the demographic, clinical, and surgical characteristics of TIP in 6 culturally and geographically distinct countries in sub-Saharan Africa. Most cases occurred in persons aged a finding consistent with TIP being a late manifestation of typhoid illness. Despite lack of culture confirmation, more than one-third of cases showed classic operative findings for TIP, and the seasonality of TIP cases correlated with the timing of typhoid detections overall. These findings are important and should be considered when determining the burden of typhoid fever and TIP in remote areas that do not have the infrastructure needed to perform bacterial culture diagnostics and pathology of surgical specimens.

Oyengu *et al.*, (2023) investigated on statistical analysis rate of typhoid fever The investigation into the statistical analysis of the rate of typhoid fever in Ijebu-Igbo, Ogun state, has yielded valuable insights into the patterns, trends, and future projections of this infectious disease. A detailed analysis process including linearity tests, trend modelling, data transformation and forecasting has been used to gain a better understanding of the disease dynamics. Descriptive statistics and measures of linearity (such as Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), and Mean Standard Deviation (MSD)) for linear, quadratic and exponential models showed that there was no evidence of linear relationships among the data. This required the research of data transformation methods to improve the accuracy of the modelling. Log transformation of the data was justified by examining the log transformed mean and standard deviation. This allowed for data to be better aligned with the statistical assumptions and allowed for a meaningful relationship to be observed between these transformed variables. The study was able to use the Buys-Ballot procedure to successfully determine underlying trends in the transformed data and found a trend equation for Mt: $Mt = 3.401 - 0.002t$. This trend equation proved to be a good basis for estimating the trend of typhoid fever occurrence from year to year. The trend equation was used to predict the future cases of typhoid fever for 2021 and 2022. The estimates calculated were helpful in understanding the expected range of disease occurrences in different quarters, and helped inform future planning and allocation of resources.

3. MATERIALS AND METHOD

In this section time series monthly data were used on typhoid fever cases, reported from (2015-2023). The data used for this study is a secondary type which downloaded from an online website called “kaggle.com”. Time series model called “ARIMA Model” were developed to forecast typhoid fever cases. The general form of time series models can be written as:

A time series is said to follow a moving average process (MA) if the following equation is satisfied

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

The Equation (1) represents a moving average of order q, denoted by MA (q)

A time series is said to follow an autoregressive process (AR) if the model in Equation (2) is satisfied.

$$Y_t = \mu + y_t + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} \quad (2)$$

The Equation (2) represents an autoregressive model of order p, denoted by AR (p)

ARMA process is a combination of both the autoregressive process AR (p) and the moving average process MA (q) to form a single model denoted by ARMA(p,q). The model enables us to obtain the model with few numbers of parameters useful when time series interests are an aggregate of a number of other series following simple stochastic. A time series x_t follows an ARMA (p,q) process if it is satisfied with this model:

$$y_t = \delta + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (3)$$

Where y_t and ε_t are the white noise process in each AR and MA respectively for an ARMA (1,1), the model is given as

$$X_t = \varphi_0 + \varphi_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (4)$$

$$X_t - \varphi_1 y_{t-1} = \varphi_0 + \varepsilon_t - \theta_1 \varepsilon_{t-1} \quad (5)$$

The left hand side of the equation is the AR component of the model and the right hand side gives the MA component and φ_0 is a constant term and $\varphi_1 \neq \theta_1$

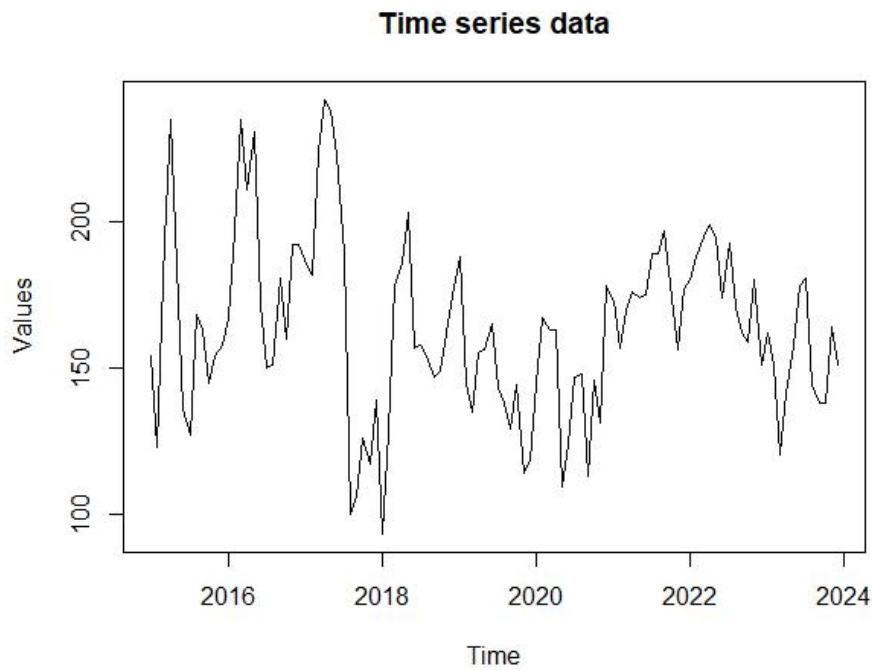


Figure 1 The plot of typhoid fever cases

The original data plot exhibits a random and stationary pattern, characterized by fluctuation around a constant mean with no trend or seasonal variation.

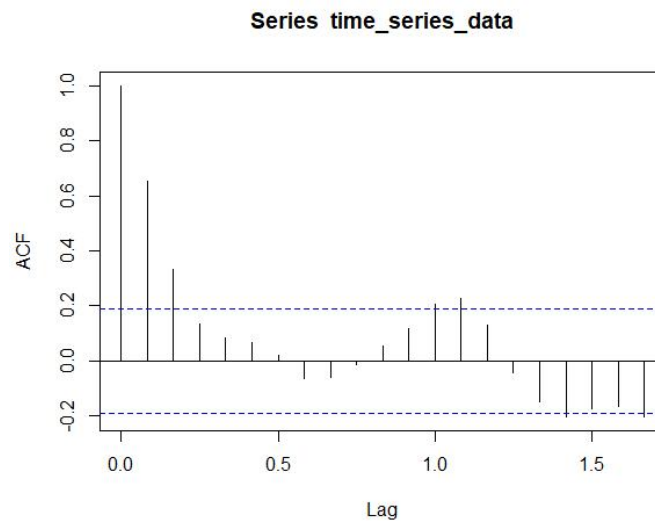


Figure 2 Auto correlation function (ACF) of the typhoid fever cases

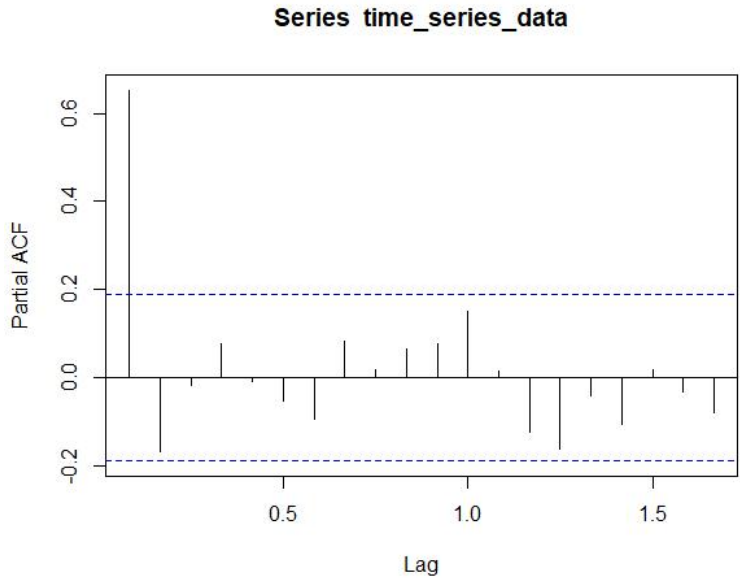


Figure 3 Partial auto correlations (PACF) of the typhoid fever cases

ARIMA MODELS	: AICc
ARIMA(2,0,2) with non-zero mean :	993.874
ARIMA(0,0,0) with non-zero mean :	1048.247
ARIMA(1,0,0) with non-zero mean :	990.7221
ARIMA(0,0,1) with non-zero mean :	1002.52
ARIMA(0,0,0) with zero mean :	1413.469
ARIMA(2,0,0) with non-zero mean :	989.8965
ARIMA(3,0,0) with non-zero mean :	992.0777
ARIMA(2,0,1) with non-zero mean :	992.0871
ARIMA(1,0,1) with non-zero mean :	990.2719
ARIMA(3,0,1) with non-zero mean :	994.3189
ARIMA(2,0,0) with zero mean :	1014.43
Best model: ARIMA (2,0,0) with non-zero mean	

Table 1 Model parameter estimation

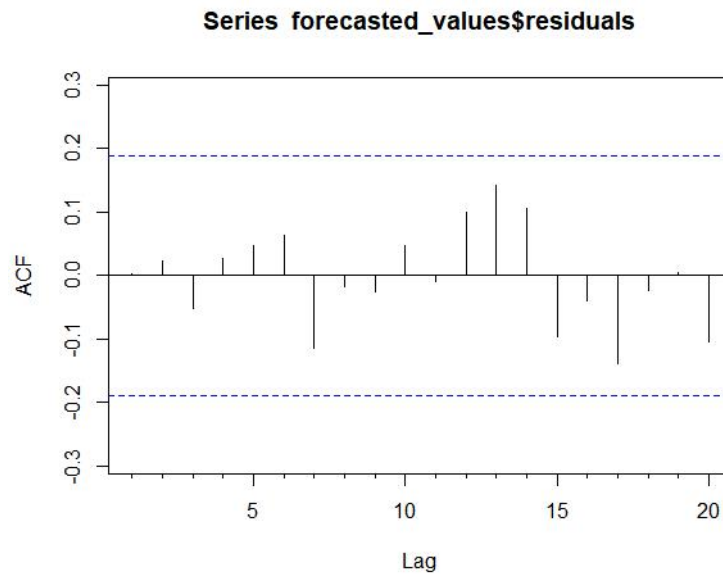


Figure 4: Autocorrelation Function(ACF) for residuals

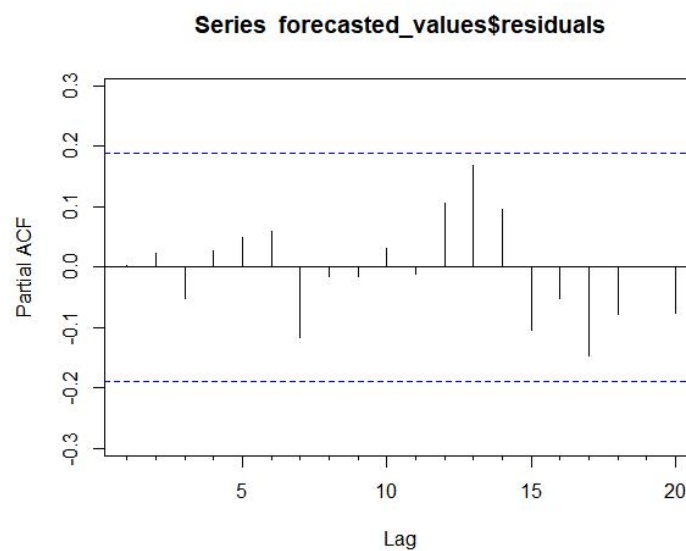


Figure 5: Partial Autocorrelation function(PACF) for residual

From Figure 4 and 5 that is the plot of (ACF) and (PACF) of the residual we noticed that there is no point in the plots outside the significant zones, therefore we conclude that the residuals are random. This suggested the model has adequately captured the underlying structure and no additional lag is needed. The residuals are like white noise, confirming the model adequacy.

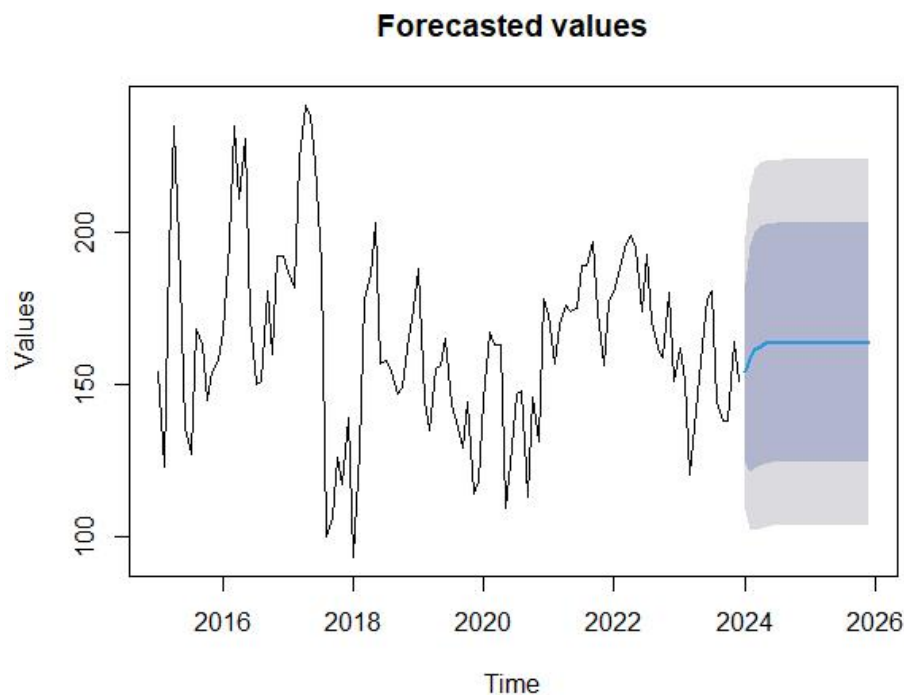


Figure 6 Plot of the forecasted values

RESULTS:

Augmented Dickey-Fuller Test was used on the typhoid fever cases to test the stationarity of the series, From the result obtained p-value is less than 5% level of significance i.e $0.0473 < 0.05$, therefore we reject the H_0 and conclude that the series is stationary. However Autocorrelation function (ACF) and partial autocorrelation function (PACF) were used in building ARIMA models, many models are tested and ARIMA (2, 0, 0) with non-zero mean was selected as the best performed model based on the AIC value. ARIMA (2, 0, 0) with non-zero mean was used to forecast typhoid fever cases, From Figure 6 the graphs suggested that cases of Typhoid fever will increase slightly from January to June and the trend remains relatively stable from July to December 2026.

CONCLUSION

This study will help in understanding the historical trends and seasonality of typhoid fever cases, and will help health authorities to be better prepared to respond to any future outbreaks, resulting in more effective and timely interventions; in addition, this study will help in better disease monitoring through the identification of critical periods of increased risk, which would facilitate the more focused surveillance during peak seasons or resurgence periods. The forecasted plot ARIMA(2,0,0) graphs suggested that cases of Typhoid fever will increase slightly from January to June, and the trend remains relatively stable from July to December 2026; this will inform national guidelines for typhoid fever prevention and control.

References

- Birkhold, M., Datta, S., Pak, G. D., Im, J., Ogundoyin, O. O., Olulana, D. I., ... & Marks, F. (2023, May). Characterization of typhoid intestinal perforation in Africa: results from the severe typhoid fever surveillance in Africa Program. In *Open Forum Infectious Diseases* (Vol. 10, No. Supplement_1, pp. S67-S73). US: Oxford University Press.
- Dalatu, P. I., & Ibrahim, A. (2023). Buys Ballot Approach in Time Series Analysis of Typhoid Fever. *Sigma&Mu: Journal of Mathematics, Statistics and Data Science*, 1(1), 9-15.
- Gao, J., Li, J., & Wang, M. (2020). Time series analysis of cumulative incidences of typhoid and paratyphoid fevers in China using both Grey and SARIMA models. *PloS one*, 15(10), e0241217.
- Gao, Q., Liu, Z., Xiang, J., Zhang, Y., Tong, M. X., Wang, S., ... & Bi, P. (2022). Impact of temperature and rainfall on typhoid/paratyphoid fever in Taizhou, China: effect estimation and vulnerable group identification. *The American Journal of Tropical Medicine and Hygiene*, 106(2), 532.
- Gauld, J. S., Bilima, S., Diggle, P. J., Feasey, N. A., & Read, J. M. (2022). Rainfall anomalies and typhoid fever in Blantyre, Malawi. *Epidemiology & Infection*, 150, e122.

- Kaigama, A., & Madaki, U. Y. (2023). Time Series Analysis for The Treatment of Typhoid (Enteric) Fever in Maiduguri: Using Arima Model. *Square: Journal of Mathematics and Mathematics Education*, 5(1), 11-22.
- Kim, C., Goucher, G. R., Tadesse, B. T., Lee, W., Abbas, K., & Kim, J. H. (2023). Associations of water, sanitation, and hygiene with typhoid fever in case-control studies: a systematic review and meta-analysis. *BMC infectious diseases*, 23(1), 562.
- Masinaei, M., Eshrati, B., & Yaseri, M. (2020). Spatial and spatiotemporal patterns of typhoid fever and investigation of their relationship with potential risk factors in Iran, 2012–2017. *International journal of hygiene and environmental health*, 224, 113432.
- Mathew, S., & Idi, D. (2024). Statistical Model for the Occurrence of Typhoid Fever in Takum, Taraba State, Nigeria: An Arima Model Approach. *ALSYS*, 4(6), 722-741.
- Muhammad, E. N., Abdul Mutalip, M. H., Hasim, M. H., Paiwai, F., Pan, S., Mahmud, M. A. F., ... & Aris, T. (2020). The burden of typhoid fever in Klang Valley, Malaysia, 2011–2015. *BMC infectious diseases*, 20, 1-10.
- Obulezi, O. J., Chidimma, N. N., Igbokwe, C. P., & Anabike, I. C. (2023). Statistical analysis on diagnosed cases of malaria and typhoid fever in Enugu-Nigeria. *GSIJ*, 11(6).
- OGOLO, I. M., & NKPORDEE, L. (2022). Forecasting the Trend of Typhoid Fever in Selected Teaching Hospital in Awka, Anambra State. *African Journal of Advances in Science and Technology Research*, 4(1), 25-38.
- Oyenuga, A. Y., & Olanrewaju, G. O. (2023). Statistical Analysis of the Rate of Typhoid Fever.
- Peter, O. J., Adebisi, A. F., Ajisope, M. O., Ajibade, F. O., Abioye, A. I., & Oguntolu, F. A. (2020). Global stability analysis of typhoid fever model. *Advances in Systems Science and Applications*, 20(2), 20-31.