Available on: https://jnsrr.ep-journals.org/ E-ISSN: 3051-3766 P-ISSN: 3051-3758

Journal of Natural Science Research and Review

Volume: 01 Issue: 03 September-2025

Forecasting Typhoid Fever Cases using Decomposition Method, Holt-Winter's Method, and Sarima Method

Adeniyi Oyewole Ogunmola
Department of Statistics, Faculty of Physical Sciences, Federal University Wukari, Taraba State, Nigeria.
Yunusa Namale Jibo
Department of Statistics, Faculty of Physical Sciences, Federal University Wukari, Taraba State, Nigeria.

Published Date: September 16, 2025

ABSTRACT: This study compares the forecasting accuracy of the three univariate time series models, Decomposition, Winter's and Seasonal autoregressive integrated moving average (SARIMA) method to predict also monthly typhoid fever cases. Secondary data is collected from Federal Medical Centre Jalingo and explored with the three time series models. The study applied time series decomposition method, Holt-Winter's method, and SARIMA method to compare their forecasting accuracy using Root Mean Square Error (RMSE) as the selection criterion. The results for typhoid fever cases revealed that SARIMA (4,0,0) × (1,0,0) best fits and predict the data more than the Decomposition, and the two Holt-Winter's model. The monthly Typhoid fever indices indicates that the month of April, May and July to October has higher typhoid fever cases compared to the annual average. The highest index is in the month of August accounting for nearly 6.6% increase over the annual average.

KEYWORDS: Typhoid Fever, Time Series, Decomposition Method, Holt-Winter Method, SARIMA Method.

INTRODUCTION

Typhoid fever, caused by *Salmonella enterica* serovar Typhi, remains a persistent public health challenge in many low- and middle-income countries, particularly in sub-Saharan Africa and South Asia (Crump et al., 2004; Ugboko et al., 2020). The disease is primarily transmitted through contaminated food and water, and its prevalence is closely linked to inadequate sanitation and limited access to clean drinking water (WHO, 2023). The disease is observed to occur with malaria and treated together in many patients in Nigeria. Typhoid fever like malaria fever in Nigeria, continues to contribute significantly to morbidity, especially among vulnerable populations such as children and young adults (Ogunmola and Jibo, 2025).

Accurate forecasting of typhoid fever incidence is essential for timely public health interventions, resource allocation, and strategic planning. Time series analysis has emerged as a robust approach for modeling infectious disease trends, enabling researchers and policymakers to anticipate outbreaks and mitigate their impact (Hyndman and Athanasopoulos, 2018; Zhang et al., 2013). Among the most widely used techniques are decomposition methods, Holt-Winters exponential smoothing, and Seasonal Autoregressive Integrated Moving Average (SARIMA) models, each offering unique strengths in capturing seasonality, trends, and irregular fluctuations in epidemiological data (Box et al., 2015; Xian et al., 2023).

Recent comparative studies have demonstrated the effectiveness of SARIMA and Holt-Winters models in forecasting various infectious diseases, including typhoid fever, malaria, and foodborne illnesses (Zhang et al., 2013; Ogunmola and Jibo, 2025; Xian et al., 2023). These models have shown strong predictive performance, particularly in capturing seasonal variations and producing low error metrics such as RMSE and MAE. Moreover, decomposition techniques provide valuable insights into the underlying structure of time series data, facilitating better model selection and interpretation (Anwar et al., 2016).

This study applies decomposition, Holt-Winters, and SARIMA models to forecast typhoid fever cases using historical data from a tertiary health facility in Nigeria. By evaluating the predictive accuracy of each method, the research aims to identify the most suitable approach for modeling typhoid fever trends, thereby contributing to improved disease surveillance and public health preparedness.

1. Methodology

Study Design and Data Collection

This study employed a retrospective time series analysis using monthly reported cases of typhoid fever from Federal Medical Centre Jalingo, Taraba State, from January 2008 to December 2023. It compares the predicted values with the data for January 2024 to December 2024 for accuracy. The dataset spans a period sixteen five years, capturing seasonal and trend variations in disease incidence. Data were cleaned to remove missing values and outliers, ensuring consistency and reliability for modeling.

Time Series Preprocessing

The raw data were first visualized to assess stationarity, seasonality, and trend components. Augmented Dickey-Fuller (ADF) tests were conducted to evaluate stationarity (Dickey & Fuller, 1979). Seasonal decomposition of time series by Loess (STL) was applied to separate the series into trend, seasonal, and residual components (Cleveland et al., 1990).

Forecasting Models

Decomposition Method

The classical multiplicative decomposition model was used to isolate the underlying patterns in the time series. This method helps identify seasonal cycles and long-term trends, which are crucial for understanding disease dynamics (Hyndman & Athanasopoulos, 2018).

Holt-Winters Exponential Smoothing

The Holt-Winters method was applied in both additive and multiplicative forms, depending on the nature of the seasonality. This model accounts for level, trend, and seasonal components and is suitable for short-term forecasting of seasonal diseases (Chatfield, 2004).

SARIMA Model

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model was fitted using the Box-Jenkins methodology. Model parameters (p, d, q, P, D, Q, s) were selected based on autocorrelation function (ACF) and partial autocorrelation function (PACF) plots, along with Akaike Information Criterion (AIC) for model selection (Box et al., 2015). Diagnostic checks were performed to validate residual normality and independence.

Model Evaluation

Forecast accuracy was assessed using standard error metrics including:

Mean Absolute Error (MAE),

Root Mean Squared Error (RMSE), and

Mean Absolute Percentage Error (MAPE).

These metrics provided a comparative basis for evaluating the performance of each model (Zhang et al., 2013; Xian et al., 2023).

The details of these methods are shown and explained in Ogunmola and Jibo, 2025.

2. RESULTS AND DISCUSSION

The descriptive statistics the Typhoid fever cases from January 2008 to December 2023 is presented in table 3.1

Table 3.1: Descriptive Statistics for Malaria Fever and Typhoid Fever Cases

Mean	Median	Std Dev	Skewness	Kurtosis	Minimum	Maximum	N
29.95	30	4.05	0.24	-0.36	21	30	192

The sample data contain 192 total months. Descriptively the monthly mean case of typhoid fever is 29.95 with a standard deviation of 4.05. The monthly median of typhoid fever case is 30, approximately equal to the monthly mean. The typhoid fever case is positively skewed with a small magnitude of 0.24, and the monthly kurtosis shows to be platykurtic with also a small magnitude of -0.36.

Figure 3.1 shows the monthly plot of typhoid fever cases. This did not show clear circles repeating itself within a year, but it appears to be showing two circles with each spanning between 5 and 7 years, likely cyclical patterns. Generally, there is no clear trend over the time showing.

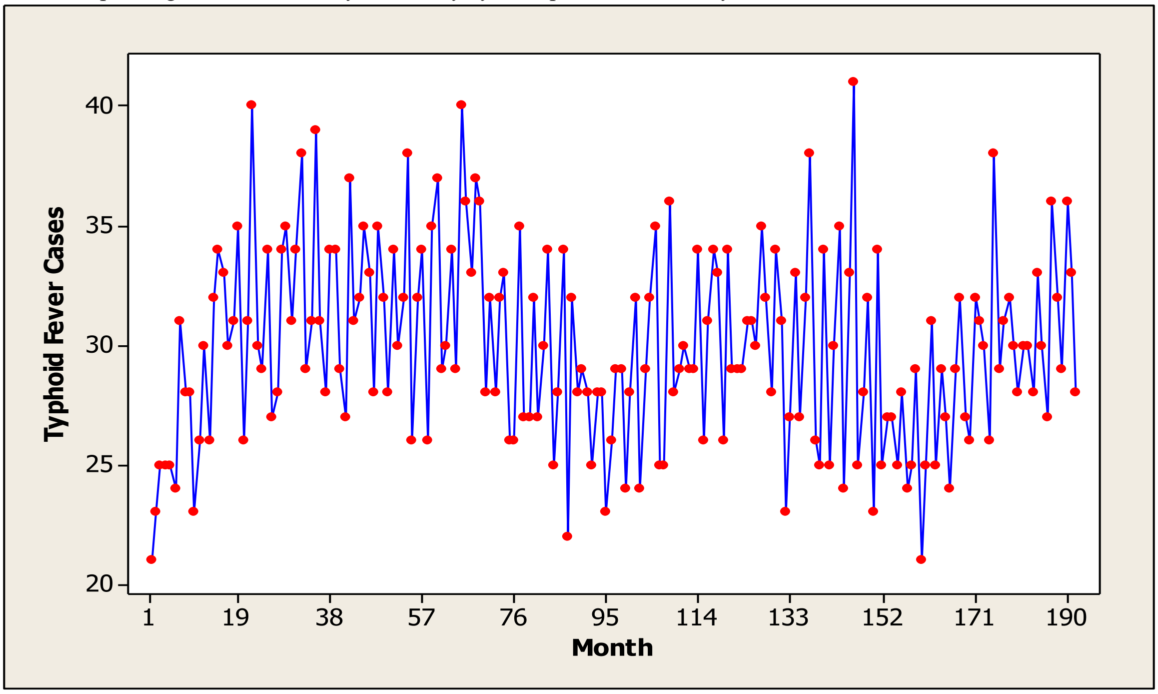

Figure 3.1: Time Series Plot of Typhoid Fever Cases

3.1 Decomposition Method

Fitting decomposition model to the data, the estimated trend equation is $Y_t = 30.620 - 0.006822*t$. The linear trend equation shows that when time increases by a month, the typhoid fever cases will decrease by approximately 0.007, which is much less than a single case of the disease. Table 3.2 shows the estimated monthly indices of typhoid fever cases. Expressing the indices in percentage, the month of April, May and July to October shows higher typhoid fever cases than the annual percentage. The highest cases are in the month of August accounting for nearly 6.6% increase over the annual. While the month of January to March, June, November and December experienced lower typhoid fever cases than the annual. The lowest cases are in the month of November representing 0.7% decrease over the annual.

Table 3.2: Seasonal Indices of Overall Typhoid Fever Cases

Period	Typhoid Fever Index	Percentage (%) index	Remarks
January	0.95661	95.661	Decrease by 4.339
February	0.99501	99.501	Decrease by 0.499
March	0.95014	95.014	Decrease by 4.986
April	1.00572	100.572	Increase by 0.572
May	1.01430	101.430	Increase by 1.430
June	0.98959	98.959	Decrease by 1.041
July	1.01663	101.663	Increase by 1.663
August	1.06572	106.572	Increase by 6.572
September	1.02958	102.958	Increase by 2.958
October	1.01177	101.177	Increase by 1.177
November	0.99637	99.637	Decrease by 0.363
December	0.96855	96.855	Decrease by 3.145

Table 3.3 shows the accuracy measures of decomposition method for forecasting typhoid fever cases. Each of the accuracy measures is relatively small.

Table 3.3: Decomposition Forecasting Model Accuracy

Accuracy Measures	Typhoid Fever
MAPE	10.9613
MAD	3.2141
MSD	15.8741

The monthly forecasted values of typhoid fever cases for the year 2024, using decomposition method are shown in Table 3.4. It can be seen that the forecasted values of typhoid fever cases are lower than the forecasted values of malaria fever cases.

Table 3.4: Decomposition Forecasts

Period	Typhoid Fever
193 (January 2024)	28.0323
194 (February 2024)	29.1508
195 (March 2024)	27.8299
196 (April 2024)	29.4509
197 (May 2024)	29.6952
198 (June 2024)	28.9651
199 (July 2024)	29.7495
200 (August 2024)	31.1787
201 (September 2024)	30.1146
202 (October 2024)	29.5867
203 (November 2024)	29.1295
204 (December 2024)	28.3096

3.2 Holt-Winter's Method

Applying Holt Winter's method for forecasting monthly typhoid fever cases, Table 3.5 shows the best model that fits the data.

Table 3.5: Selection Table for the best Winter's Model that Fits Monthly Typhoid Fever Cases

α	β	γ	MAPE	MAD	MSD
0.1	**0.1**	**0.1**	**10.9276**	**3.2187**	15.7459
0.1	0.2	0.2	11.3459	3.3498	16.5373
0.1	0.1	0.2	11.1930	3.2929	16.5831
0.2	**0.1**	**0.1**	11.0859	3.2765	**15.7265**
0.2	0.2	0.1	11.2124	3.3219	16.0907
0.2	0.1	0.2	11.4317	3.3748	16.6467
0.2	0.2	0.2	11.5735	3.4242	17.0595
0.3	0.2	0.1	11.5835	3.4304	17.3229
0.3	0.2	0.2	11.9436	3.5324	18.2881
0.3	0.3	0.2	12.1118	3.5829	18.9152
0.3	0.3	0.3	12.4903	3.6930	19.9377
0.3	0.1	0.1	11.4461	3.3854	16.7655
0.3	0.1	0.3	12.1329	3.5803	18.6038
0.9	0.1	0.1	15.1833	4.4834	28.8916

Based on MAPE and MAD, the smallest values of the measures of accuracy are obtained when the parameters are at $\alpha = 0.1$, $\beta = 0.1$, and $\gamma = 0.1$. Based on MSD, the smallest value is obtained when the parameters are at $\alpha = 0.2$, $\beta = 0.1$, and $\gamma = 0.1$. Hence, the two models are considered for the forecast of typhoid fever cases. The monthly forecasted values for the year 2024, for the typhoid fever cases are presented in Table 3.6.

Table 3.6: Holt Winter's Forecasts

Period	Typhoid Fever $\alpha = 0.1, \beta = 0.1, \gamma = 0.1$	Typhoid Fever $\alpha = 0.2, \beta = 0.1, \gamma = 0.1$
193 (January 2024)	30.9888	30.7465
194 (February 2024)	32.9052	32.5509
195 (March 2024)	31.3261	30.9045
196 (April 2024)	32.2365	31.7275
197 (May 2024)	34.1761	33.6211
198 (June 2024)	32.0923	31.5922
199 (July 2024)	34.5787	33.8230
200 (August 2024)	34.2240	33.5108
201 (September 2024)	33.5134	32.7605
202 (October 2024)	33.8378	33.0235
203 (November 2024)	33.8212	32.9560
204 (December 2024)	33.2901	32.3661

3.3 SARIMA Method

Different SARIMA models were used to model the monthly typhoid fever cases and the best model that fits the monthly typhoid fever cases is selected using the lowest value of the Akaike Information Criterion (**AIC**). The results are shown in table 3.7.

Table 3.7: AIC measures for all possible SARIMA models for monthly typhoid fever cases

SARIMA (p, d, q)×(P, D, Q)

Model	AIC
SARIMA (4,0,1) × (1,0,1)	1021.23
SARIMA (2,0,0) × (1,0,1)	1090.45
SARIMA (4,0,0) × (1,0,1)	1020.8
SARIMA (3,0,1) × (2,0,2)	1023.86
SARIMA (4,0,0) × (2,0,0)	1020.65
SARIMA (4,0,0) × (2,0,1)	1022.95
SARIMA (3,0,0) × (1,0,0)	1022.14
SARIMA (5,0,0) × (1,0,0)	1020.32
SARIMA (4,0,1) × (1,0,0)	1019.17
***SARIMA (4,0,0) × (1,0,0)**	**1018.17**

SARIMA (4,0,0) × (1,0,0) yields the minimum **AIC of 1018.17** among the computed models. Hence SARIMA (4,0,0) × (1,0,0) is the best fitted model for the monthly typhoid fever cases. This model has five key parameters to estimate which are four non-seasonal autoregressive orders and one seasonal autoregressive order, Table 3.8, gives the estimates of the parameters and the tests of their significances. The significant estimated parameter in the model is Autoregressive at lag 3 with a coefficient of 0.58, which means for every 100 additional cases reported previously, we expect 58 additional typhoid fever cases now due to a recurring three months cycle inherent in the disease's epidemiology.

Table 3.8: Final Estimates of Parameters

Coefficient	Estimate	Std-Error	T-Value	P-Value
AR1	0.02	0.07	0.29	0.77
AR2	0.00	0.06	0.04	0.97
AR3	0.58	0.06	9.61	0.00
AR4	0.08	0.07	1.05	0.29
SAR1	-0.05	0.08	-0.72	0.47

Testing for the residuals after fitting the model, Table 3.9 shows the Ljung Box Chi-Square statistic with its p-value at lags 12, 24, 36, and 48. The null hypothesis is that the residuals are independently distributed, that is, no autocorrelation for the tested lags. In the table, the p-values of the lags are all greater than 0.05, indicating no evidence of autocorrelation at those lags. There is no pattern left in the residual that the fitted model has not captured.

Table 3.9: Modified Box-pierce (Ljung Box) Chi-square statistic

Lag	12	24	36	48
Chi- square	8.0278	15.46	25.33	37.004
DF	12	24	36	48
p-value	0.783	0.9066	0.9079	0.8754

The monthly forecasted typhoid fever cases of the model for the year 2024, and their 95% confidence intervals are presented in Table 3.10

Table 3.10: Forecasts and 95 percent Limit

Year	Month	Forecast	Lower	Upper	Actual
2024	Jan	33.38179	26.80868	39.95490	29
2024	Feb	32.22065	25.64611	38.79519	39
2024	Mar	29.10904	22.53447	35.68361	32
2024	Apr	31.56574	23.96094	39.17054	32
2024	May	31.52551	23.89156	39.15947	34
2024	Jun	29.72913	22.09504	37.36322	30
2024	Jul	30.53758	22.58555	38.48961	28
2024	Aug	30.86888	22.88351	38.85425	30
2024	Sep	29.88259	21.89661	37.86856	29
2024	Oct	30.08204	21.89661	38.17339	25
2024	Nov	30.42033	22.30504	38.53561	28
2024	Dec	30.03723	21.92091	38.15356	23

3.4 Comparing the Forecasts of the Three Methods for Forecasting Monthly Typhoid Fever Cases

Root mean square error is used to determine the best method that forecasts the monthly typhoid fever cases. Table 3.11 shows the actual monthly typhoid fever cases and the year 2024 forecasted values from the Decomposition, Winter's, and SARIMA methods.

Table 3.11: The Actual Monthly Typhoid Fever Cases and the Forecasts from the Decomposition, Winter's and SARIMA methods.

Month	**2024**	**Forecast (Decomposition)**	**Forecast Winter's (α = 0.1, β = 0.1, γ = 0.1)**	**Forecast Winter's (α = 0.2, β = 0.1, γ = 0.1)**	**Forecast (SARIMA)**
January	29	32.0555	30.9888	30.7465	33.38179
February	39	29.4038	32.9052	32.5509	32.22065
March	32	31.2374	31.3261	30.9045	29.10904
April	32	30.0043	32.2365	31.7275	31.56574
May	34	30.1885	34.1761	33.6211	31.52551
June	30	33.1998	32.0923	31.5922	29.72913
July	28	32.5549	34.5787	33.8230	30.53758

August	30	28.9426	34.2240	33.5108	30.86888
September	29	33.0550	33.5134	32.7605	29.88259
October	25	32.5720	33.8378	33.0235	30.08204
November	28	31.6116	33.8212	32.9560	30.42033
December	23	33.1556	33.2901	32.3661	30.03723

Table 3.12 shows the computed RMSE values of Decomposition, Winter's, and SARIMA methods for monthly malaria fever cases.

Table 3.12: RMSE of Typhoid Fever Cases for Decomposition, Winter's, and SARIMA methods

Month	**Decomposition** $(y_t - \hat{y}_t)^2$	**Winter's (α = 0.1, β = 0.1, γ = 0.1)** $(y_t - \hat{y}_t)^2$	**Winter's (α = 0.2, β = 0.1, γ = 0.1)** $(y_t - \hat{y}_t)^2$	**SARIMA** $(y_t - \hat{y}_t)^2$
January	9.3361	3.9553	3.0503	19.2001
February	92.0871	37.1466	41.5909	45.9596
March	0.5816	0.4541	1.2001	8.3576
April	3.9828	0.0559	0.0743	0.1886
May	14.5275	0.031	0.1436	6.1231
June	10.2387	4.3777	2.5351	0.0734
July	20.7471	43.2793	33.9073	6.4393
August	1.1181	17.8422	12.3257	0.755
September	16.443	20.3708	14.1414	0.779
October	57.3352	78.1067	64.3766	25.8271
November	13.0437	33.8864	24.5619	5.858
December	103.1362	105.8862	87.7238	49.5226
Total	**342.5771**	**345.3922**	**285.631**	**169.0834**
RMSE	**5.3430**	**5.3649**	**4.8788**	**3.7537**

The RMSE for Decomposition method is 5.3430. The RMSE for Winter's (α = 0.1, β = 0.1, γ = 0.1) method is 5.3649. The RMSE for Winter's (α = 0.2, β = 0.1, γ = 0.1) method is 4.8788. The RMSE for SARIMA method is 3.7537. The least RMSE value is from SARIMA method. SARIMA (4,0,0) × (1,0,0) is the best method in predicting monthly typhoid fever cases.

3. SUMMARY

Decomposition method was first fitted to the monthly cases. The trend equation for Typhoid fever cases shows about 0.7 percent decrease for an additional month in time. The monthly Typhoid fever indices indicates that the month of April, May and July to October has higher typhoid fever cases compared to the annual average. The highest index is in the month of August accounting for nearly 6.6% increase over the annual average. While the month of January to March, June, November and December experienced lower typhoid fever index to the annual average. The lowest cases are in the month of November representing 0.7% decrease over the annual average.

Holt-Winter's Multiplicative method was fitted to the monthly typhoid fever cases. Fourteen suitable models were derived for typhoid fever cases, and the best model is selected using the least values of the measures of accuracy; MAPE, MAD, and MSD. Based on MAPE and MAD smallest values, the best fitted Holt-Winter's model to the typhoid fever cases is when the weight of level is at 0.1, the weight of trend is at 0.1, the weight of seasonal is at 0.1. Based on MSD smallest value, the best fitted Holt-Winter's model is when the weight of level is at 0.2, the weight of trend is at 0.1, the weight of seasonal is at 0.1.

SARIMA method was fitted to the monthly typhoid fever cases. Ten closely suitable models were derived for monthly typhoid fever cases, and the best model is selected using the least value of AIC. The best fitted model is SARIMA (4,0,0) × (1,0,0), with four non-seasonal parts and one seasonal part. Among the four non-seasonal part, autoregressive of order 3 term was particularly strong and has a significant coefficient of 0.58, meaning for every 100 additional cases reported previously, we expect 58 additional typhoid fever cases now due to a recurring three months cycle inherent in the disease's epidemiology. The one seasonal autoregressive means the current typhoid fever cases are influenced by the cases from the same month in the previous year.

Diagnostic check using the Ljung Box Chi-square statistic determines whether there is autocorrelation in the fitted SARIMA model of monthly typhoid fever cases. The result indicates that there is no autocorrelation for typhoid fever cases.

Comparing the forecast from Decomposition, Winters (α = 0.1, β = 0.1, γ = 0.1), Winter's (α = 0.2, β = 0.1, γ = 0.1), and SARIMA methods for monthly typhoid fever cases, Root Mean Square Error (RMSE) revealed that SARIMA (4,0,0) × (1,0,0) is the best in predicting monthly typhoid fever cases.

4. CONCLUSION

The analysis demonstrates that typhoid fever cases follow a clear seasonal pattern, with elevated incidence during mid-year months and reduced occurrence at the beginning and end of the year. Among the forecasting methods applied, SARIMA(4,0,0) (1,0,0) proved superior, offering the highest predictive accuracy and reliable diagnostic performance. This suggests that SARIMA modeling provides a robust framework for disease forecasting, which can enhance public health preparedness and guide resource allocation. Effective forecasting of typhoid fever trends can enable health authorities to anticipate peak transmission periods, implement targeted preventive measures, and strengthen response strategies. Future research could extend this work by incorporating environmental and socio-demographic factors into the models to improve predictive performance and broaden applicability to other infectious diseases.

REFERENCES

1) Anwar, M., Khan, M. A., & Rizvi, S. A. (2016). Time series analysis of infectious diseases: A review. *International Journal of Epidemiology Research*, 3(2), 101–110.
2) Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control* (5th ed.). Wiley.
3) Chatfield, C. (2004). *The Analysis of Time Series: An Introduction* (6th ed.). CRC Press.
4) Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1), 3–73.
5) Crump, J. A., Luby, S. P., & Mintz, E. D. (2004). The global burden of typhoid fever. *Bulletin of the World Health Organization*, 82(5), 346–353.
6) Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts.
7) Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427–431.
8) Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). Texts.
9) Ogunmola, A. O., & Jibo, Y. N. (2025). Comparing Univariate Time Series Forecast Methods for Malaria Fever Cases. *Journal of Multidisciplinary Sciences*, 3(2), 1018–1030. https://doi.org/10.58578/mikailalsys.v3i2.6366
10) Ugboko, H. U., Deekae, S., & Nwachukwu, E. (2020). Typhoid fever in Nigeria: A review of its epidemiology and management. *African Journal of Clinical and Experimental Microbiology*, 21(1), 1–10.
11) WHO. (2023). *Typhoid fever fact sheet*. World Health Organization. Retrieved from https://www.who.int/news-room/fact-sheets/detail/typhoid
12) Xian, X., Wang, L., Wu, X., Tang, X., Zhai, X., Yu, R., Qu, L., & Ye, M. (2023). Comparison of SARIMA model, Holt-Winters model and ETS model in predicting the incidence of foodborne disease. *BMC Infectious Diseases*, 23, Article 803. https://doi.org/10.1186/s12879-023-08799-4
13) Zhang, X., Liu, Y., Yang, M., Zhang, T., Young, A. A., & Li, X. (2013). Comparative study of four time series methods in forecasting typhoid fever incidence in China. *PLOS ONE*, 8(5), e63116. https://doi.org/10.1371/journal.pone.0063116

Corresponding Author: Adeniyi Oyewole Ogunmola