Comparative Modeling of Measles Incidences Among Children Below Five Years in Rivers State, Nigeria Using the Holt-Winters, SARIMA, and Hybrid SARIMA-ANN Models

Amala Isabella Smart
Department of Mathematics
Rivers State University,
Port Harcourt
Email: amalaisabella3@gmail.com

Amos Emeka
Department of Mathematics
Rivers State University,
Port Harcourt
Email: amos.emeka@ust.edu.ng

Davis Iyai
Department of Mathematics
Rivers State University,
Port Harcourt
Email: davis.iyai@ust.edu.ng

Anthony I. Wegbom
Department of Public Health Sciences
Rivers State University,Port Harcourt
Email: Anthony.wegbom1@ust.edu

Da-Wariboko, Yvonne Asikiye
Email: yvonne.da-wariboko@ust.edu.ng;
daasiks73@gmail.com
Department of Mathematics, Rivers State University

https://doi.org/10.56201/ijasmt.vol.12.no10.2026pg44.59

Abstract
This research analyzed various time-series forecasting models for measles occurrences in children younger than five in Rivers State, Nigeria, utilizing monthly data sourced from the Rivers State Hospital Management Board and allied partners spanning from January 2015 to December 2024. Python 3.11 software was used for the data analysis. The investigation focused on the temporal patterns of measles cases and assessed the prediction efficacy of additive Holt–Winters, multiplicative Holt–Winters, Seasonal Autoregressive Integrated Moving Average (SARIMA), and hybrid SARIMA–Artificial Neural Network (ANN) models. To explore the data series, methods

such as descriptive statistics, seasonal decomposition, the Jarque–Bera normality examination, the Augmented Dickey–Fuller unit-root evaluation, autocorrelation assessment, and the Ljung–Box test were utilized. The forecasting efficiency was measured via root mean squared error (RMSE), mean absolute scaled error (MASE), and mean absolute percentage error (MAPE). Findings indicated an average of 26.17 monthly cases of measles, with a standard deviation of 21.43 and a peak of 124 cases. The data series displayed positive skewness and did not conform to a normal distribution, while the results from the Augmented Dickey–Fuller test suggested stationarity at level significance ($p < 0.001$). The Ljung–Box analysis uncovered significant serial correlations at lags 12 and 20 ($p < 0.001$). Out of the four forecasting techniques, the multiplicative Holt–Winters method yielded the best results, achieving the lowest RMSE at 16.43 and a reported MASE of 86.68. Conversely, the hybrid SARIMA–ANN approach was the least effective, presenting an RMSE of 22.05 and a reported MASE of 261.66. The study concludes that the multiplicative Holt–Winters method stands out as the most effective model among those tested. It suggests the validation of the reported MASE values, the importance of consistent out-of-sample evaluation, and enhancing regular measles monitoring to aid in timely forecasting, vaccination strategy, and focused public health actions.

Keywords: *Measles, forecasting, Holt–Winters, SARIMA, artificial neural network, Rivers State*

Introduction

Measles continues to pose a significant global health challenge due to its high contagiousness and the fact that it can be prevented with vaccination, particularly impacting children under the age of five. Even with effective vaccines available, the World Health Organization (WHO) and the United States Centers for Disease Control and Prevention (CDC) projected that in 2023, there were 10.3 million cases of measles and approximately 107,500 fatalities worldwide, most of which were among children younger than five years (WHO & CDC, 2024). Ongoing issues such as insufficient vaccination coverage, restricted access to medical care, population migrations, and delays in detecting outbreaks continue to impede efforts to prevent and control measles (Awogbemi, et al,2026). Nigeria is notably affected by ongoing measles spread. WHO statistics indicated that there were 14,999 cases reported in 2024, underscoring the persistent burden of the illness and the necessity for robust surveillance and prevention methods (Awogbemi, et al,2026).

In Rivers State, it is crucial to comprehend the temporal patterns of measles occurrences in children under five years to anticipate heightened transmission periods, manage vaccination distribution, and enhance public health responses. Unpredictable variations in reported cases may complicate strategic planning and increase the social and economic pressures on families and healthcare institutions. Time-series forecasting offers a methodical way to analyze past disease trends and project future incidence rates. Holt-Winters exponential smoothing methods capture the level, trend, and seasonal changes, while Seasonal Autoregressive Integrated Moving Average (SARIMA) models consider serial dependencies and seasonal interactions. A hybrid SARIMA-Artificial Neural Network (ANN) model could further identify nonlinear trends that traditional statistical models might not sufficiently address (Hyndman & Athanasopoulos, 2021; Zhang, 2003). Evaluating these various methods can assist in determining which model yields the most dependable forecasts for planning in public health (Awogbemi, et al,2026).

This study is consistent with Sustainable Development Goal 3 (SDG 3), particularly target 3.2, which aims to eliminate preventable deaths among newborns and children under five by the year 2030. Enhancing the forecasting of measles occurrences can facilitate evidence-based planning for

immunizations, prompt outbreak preparedness, and optimal allocation of limited healthcare resources. Thus, this study assesses Holt-Winters, SARIMA, and integrated SARIMA-ANN models to evaluate their relative effectiveness in forecasting measles occurrences among children under five years in Rivers State, Nigeria.

Methodology

The data source for this study was the Rivers State Hospital Management Board in collaboration with other partners. The Rivers State Hospital Management Board employs a rigorous sampling methodology, making it a reliable source of population-level data on measles. The data spanned from January 2015 to December,2024. The monthly series of measles cases was analyzed to inspect its temporal characteristics, encompassing potential trends, variations, seasonal behaviors, and any irregular incidents. Following this, seasonal decomposition was utilized to break down the observed data into its trend, seasonal, and residual elements. In an additive decomposition approach, the observed data can be represented as: $Y_t = T_t + S_t + E_t$,

where Y_t represents observed measles cases, T_t denotes the trend-cycle component, S_t represents the seasonal component, and E_t is the irregular component. This method aids in recognizing temporal patterns that may influence the selection of forecasting strategies (Hyndman & Athanasopoulos, 2021). Additionally, descriptive statistics such as mean, median, minimum, maximum, standard deviation, skewness, and kurtosis were computed to analyze the distribution and variability of measles cases. The Jarque-Bera test will be implemented to evaluate the null hypothesis regarding the normal distribution of the observations. The test statistic is defined as thus:

$$x^2 = {}^{N}/_{6}\left[S^2 + \left(\frac{k-3}{4}\right)^2\right]$$

where; S-Stands for skewed Statistic, K-Stands for kurtosis and N-Stands the size of the data under investigation.

Furthermore, the Augmented Dickey-Fuller (ADF) test was utilized to determine if the measles series contained a unit root. The null hypothesis posits that the series contains a unit root, and rejecting it indicates evidence of stationarity based on the chosen test criteria (Dickey & Fuller, 1979). If the initial series is indicated as stationary, it is unnecessary to apply first differencing automatically. The requirement for regular or seasonal differencing can be determined by analyzing time-series plots, the seasonal structure, and model diagnostics. Seasonal differencing at lag 12 is notably significant when monthly data demonstrates a consistent annual trend. Moreover, autocorrelation analysis and model identification were also taken into account. The Ljung-Box test was employed to evaluate if autocorrelations up to predefined lags were collectively equal to zero. The autocorrelation function (ACF) and the partial autocorrelation function (PACF) were also scrutinized to assist in identifying suitable autoregressive and moving average terms. These instruments aid in initial model identification; however, the final model formulation must be established through model estimation and residual diagnostics (Ljung & Box, 1978). To ensure suitable model fitting, both additive and multiplicative Holt-Winters models will be applied. The additive and multiplicative Holt-Winters exponential smoothing models were adopted to effectively capture the level, trend, and yearly seasonality in monthly measles cases. The additive model presumes that seasonal variations remain largely constant in absolute terms, while the multiplicative model suggests that seasonal changes are proportional to the level of the series (Winters, 1960; Hyndman & Athanasopoulos, 2021). The additive model can be represented as:

$$\begin{aligned}\ell_t &= \alpha(Y_t - s_{t-m}) \\ &\quad +(1-\alpha)(\ell_{t-1} + b_{t-1}), \\ b_t &= \beta(\ell_t - \ell_{t-1}) \\ &\quad +(1-\beta)b_{t-1}, \\ s_t &= \gamma(Y_t - \ell_t) + (1-\gamma)s_{t-m}, \\ \hat{Y}_{t+h|t} &= \ell_t + hb_t + s_{t-m+h_m},\end{aligned}$$

where ℓ_t is the estimated level, b_t is the trend, s_t is the seasonal component, $m = 12$ denotes the monthly seasonal period, and α, β, γ are smoothing parameters. For the multiplicative model, the seasonal contribution is incorporated proportionally to the level and trend. Its use requires strictly positive observations and appropriately normalized seasonal indices. The multiplicative model can be represented as

$$\begin{aligned}\ell_t &= \alpha\left(\frac{Y_t}{s_{t-m}}\right) \\ &\quad +(1-\alpha)(\ell_{t-1} + b_{t-1}), \\ b_t &= \beta(\ell_t - \ell_{t-1}) \\ &\quad +(1-\beta)b_{t-1}, \\ s_t &= \gamma\left(\frac{Y_t}{\ell_t}\right) + (1-\gamma)s_{t-m}, \\ \hat{Y}_{t+h|t} &= (\ell_t + hb_t)s_{t-m+h_m},\end{aligned}$$

where ℓ_t is the estimated level, b_t is the trend, s_t is the seasonal component, $m = 12$ denotes the monthly seasonal period, and α, β, γ are smoothing parameters.

The preliminary seasonal vectors were analyzed to comprehend the anticipated seasonal trend; however, the choice of the model was determined by its forecasting effectiveness instead of just the initial seasonal metrics. In a further development, Seasonal autoregressive integrated moving-average (SARIMA) models were utilized to account for both seasonal and non-seasonal interactions in the monthly measles dataset. The overall specification is

$$\text{SARIMA}(p, d, q)\,(P, D, Q)_{12},$$

where p and q refer to the orders for non-seasonal autoregressive and moving-average components, d indicates the order of non-seasonal differencing, while P and Q pertain to the orders of seasonal autoregressive and moving-average terms, respectively, and D represents the order for seasonal differencing. The potential models were assessed using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and measures of forecasting errors. Decreased AIC and BIC values signify a preferable compromise between model fitting and complexity among similar models, and reduced forecasting errors suggest superior predictive capabilities. This method aligns with established Box–Jenkins principles of time-series modeling as outlined by Box et al. in 2015. Ultimately, a hybrid SARIMA-artificial neural network (ANN) model was proposed to capture both linear time-related dependencies and any nonlinear trends that SARIMA may not clarify. This method is aligned with the hybrid forecasting approach described by Zhang in 2003. The model can be formulated as

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t,$$

where $\hat{L}t$ represents the SARIMA prediction and $\hat{N}t$ stands for the ANN prediction concerning the residuals left from the SARIMA model. The ANN configuration included a first hidden layer containing 50 neurons, a second hidden layer with 25 neurons, with ReLU activation functions in both layers, and a linear layer for output. The purpose of the ANN was to identify nonlinear trends within the residual series. Its effectiveness should be evaluated using identical chronological training and testing intervals as the standalone SARIMA model, ensuring that predictions for the future do not contribute to the training of the model. The hybrid model should only be deemed superior if it performs better out of sample than both individual models do. The additive Holt–Winters, multiplicative Holt–Winters, SARIMA, and hybrid SARIMA-ANN models were analyzed by utilizing root mean squared error (RMSE), mean absolute scaled error (MASE), and where applicable, mean absolute percentage error (MAPE). The definitions for these measures are:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}$$

$$MASE = \frac{\frac{1}{n}\sum_{t=1}^{n}|Y_t - \hat{Y}_t|}{\frac{1}{T-m}\sum_{t=m+1}^{T}|Y_t - Y_{t-m}|}$$

RMSE penalizes significant forecasting errors, while MASE evaluates the model's mean absolute error in relation to a seasonal-naïve benchmark for monthly data, with m set at 12. MASE values falling below one indicate superiority over the benchmark on average. MAPE evaluates absolute percentage errors but may be undefined when actual counts are zero and can become erratic when counts are close to zero, as noted by Hyndman and Athanasopoulos in 2021. The findings indicated that the multiplicative Holt–Winters model emerged as the top-performing model based on calculated RMSE values. However, the MASE values reported necessitate scrutiny since they appear exceedingly high for a scaled error metric. All models ought to be assessed using the same test data and with consistent forecast horizons before a final model is chosen. Additionally, the residual autocorrelation should be investigated using the residual ACF and the Ljung–Box test to determine if a significant temporal structure is still unexplained.

Results

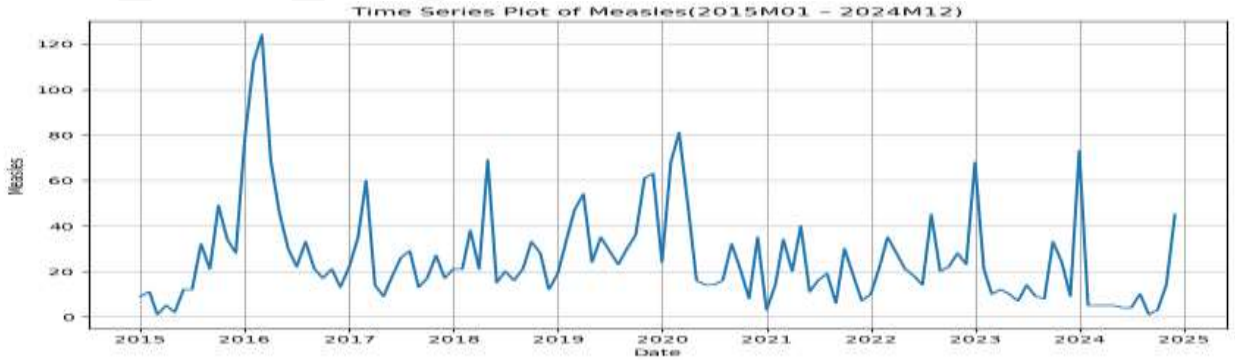


Figure 4.1: Seasonal time series plot for Measles

Figure 4.1 illustrates the seasonal time-series graph of measles cases in children younger than five in Rivers State.

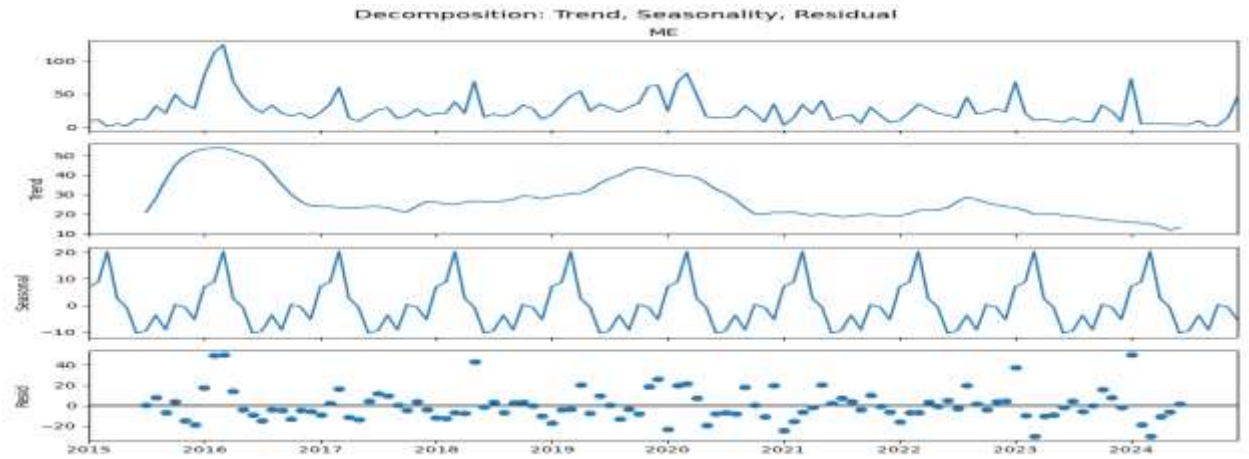


Figure 4.2: Time Series Decomposition (Trend, Seasonality, Residual) for Measles

Figure 4.2 shows the decomposition of measles cases into three components: trend, seasonal, and residual. The trend component is valuable for detecting long-term changes in measles occurrence, while the seasonal component reveals any recurring patterns that exist within the year. The residual component indicates the variation that the estimated trend and seasonal structure cannot explain. Nevertheless, one cannot determine the specific direction and size of the trend, the intensity of the seasonal component, or the behavior of the residuals purely from the figure's caption; these aspects should be assessed in conjunction with the actual plotted components. Should the seasonal component reveal consistent peaks at similar times across different years, this would reinforce the need to incorporate seasonal elements in the forecasting model. Conversely, if the residual component displays systematic patterns or unusually large swings, it would indicate that the decomposition has not completely captured the underlying temporal dynamics of measles incidence.

Table 4.1: Descriptive Statistics of the Raw and Differenced Childhood Diseases (Measles) Under 5 years of Age in Rivers State.

Statistics	ME	DME
Mean	26.16667	26.16667
Median	21.00000	21.00000
Maximum	124.0000	124.0000
Minimum	1.000000	1.000000
Std. Dev.	21.43324	21.43324
Skewness	1.917271	1.917271
Kurtosis	7.683979	7.683979
Jarque-Bera	183.2169	183.2169
Probability	0.000000	0.000000
Sum	3140.000	3140.000
Sum Sq. Dev.	54666.67	54666.67

Table 4.1 indicate that the average number of cases was 26.17, while the median value stood at 21 cases. This suggests that the mean value surpasses the middle value of the data. Such a difference aligns with a distribution that is positively skewed, as evidenced by a skewness coefficient of 1.9173. The highest recorded figure was 124 cases, contrasted with a minimum of just one case, reflecting significant fluctuations in the number of measles cases throughout the observation

period. A standard deviation of 21.43 cases further illustrates the considerable spread around the mean. The kurtosis value of 7.684, which greatly exceeds the normal distribution threshold of three, indicates a distribution with heavy tails, where extremely high case counts occur more frequently than would be presumed under a normal distribution. The Jarque–Bera statistic of 183.2169, accompanied by a p-value of 0.0000 (to be interpreted as $p < 0.001$ rather than strictly zero), offers compelling evidence against the null hypothesis of observations being normally distributed. From a statistical modeling standpoint, these results imply that a forecasting model predicting normally distributed errors without proper diagnostics or transformations may not effectively capture the measles data. Thus, models that can address non-normality, variability changes, and possible seasonal trends should be explored.

Table 4.2: Unit Root Tests of the Raw Series on Childhood Diseases (Pneumonia, Diarrhea and Measles) Under 5 years of Age in **Rivers State.**

Variable	t-Statistic	P-Value	Remarks
ME	-5.9444	0.0000	1(0)
△(ME)	-5.9444	0.0000	1(1)

Note: **△ Differenced**

Table 4.2 displays the outcome of the unit-root examination for the unprocessed measles data, labeled ME. The calculated test statistic is −5.9444, accompanied by a p-value of 0.0000, which presents significant evidence against the unit-root null hypothesis at standard significance levels, contingent on the validity of the test specification and the assumptions about critical values. This finding falls under the category of (I(0)), indicating that the raw data series is stationary in levels according to the adopted test. This result is crucial as stationarity plays a vital role in time-series analysis: it suggests that differencing is not necessarily required just to eliminate a unit root. However, having a stationary mean does not guarantee consistent variance, normally distributed data points, or the lack of seasonal patterns. The notable skewness and high kurtosis found in Table 4.1, along with the autocorrelation findings in Table 4.3, indicate that other aspects of the series still need to be addressed. The differenced series, denoted as △(ME), lacks a noted test statistic or p-value in Table 4.2. Hence, it is impossible to make any determinations regarding its stationarity from that particular row, and results for the differenced series should be included if differencing was part of the analytic process.

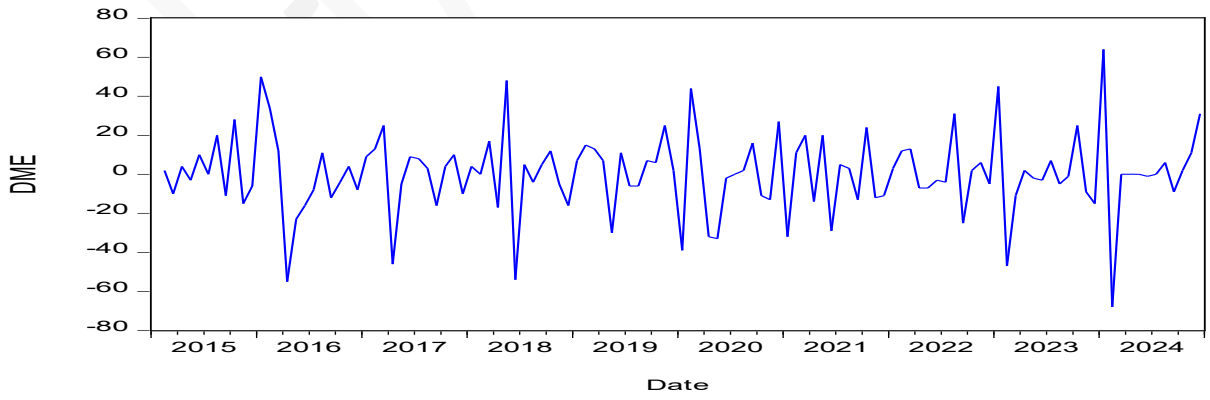


Figure 4.3: Differenced Seasonal time series plot for Measles

Figure 4.3 shows the seasonal time-series visualization of the differenced measles data. Differencing is often employed to analyze the variations between consecutive observations and can aid in eliminating trends if they exist. In this research, though, the unit-root analysis in Table 4.2 reveals that the raw series is already stationary in levels. Thus, it should not be assumed that differencing is essential without clear diagnostic reasoning. An analysis of the differenced plot is required to see if it diminishes ongoing movements or merely introduces unwarranted volatility. If differencing leads to frequent changes in signs and pronounced short-term variation without enhancing the statistical characteristics of the series, it could result in over-differencing and a less efficient model. Additionally, if the initial dataset consists of monthly or quarterly data points, it's important to differentiate seasonal differencing from basic first differencing, as they address distinct time-related structures.

Table 4.3: **Autocorrelation Test Results (Ljung-Box + ACF) for Childhood Diseases (Measles) Under 5 years of Age in Rivers State.**

	Measles		
Lag	Stat	**P-value**	**Remarks**
12	64.848	0.000	Autocorrelated
20	68.358	0.000	Autocorrelated

Table 4.3 provides the results of the Ljung–Box test for the measles data. For lag 12, the calculated test statistic is 64.848, with a p-value of 0.000 reported, while at lag 20, the statistic is 68.358, also with a p-value of 0.000. Both outcomes present substantial evidence against the null hypothesis stating that the autocorrelations up to the specified lags are collectively zero. Therefore, the measles data demonstrates a significant level of serial correlation. This indicates that the observations do not act as independent random variables and that historical case counts may provide valuable insights for predicting future counts. The findings advocate for the exploration of autoregressive, moving-average, seasonal, or other time-series models depending on the frequency of observations and the patterns highlighted by the correlograms. However, the Ljung–Box test does not confirm that the observed dependence is necessarily seasonal; the autocorrelation might be due to non-seasonal persistence, recurring seasonal traits, or a blend of both. The table should clarify whether the examinations were performed on the raw dataset, the differenced dataset, or the model residuals, as these will yield varied interpretations

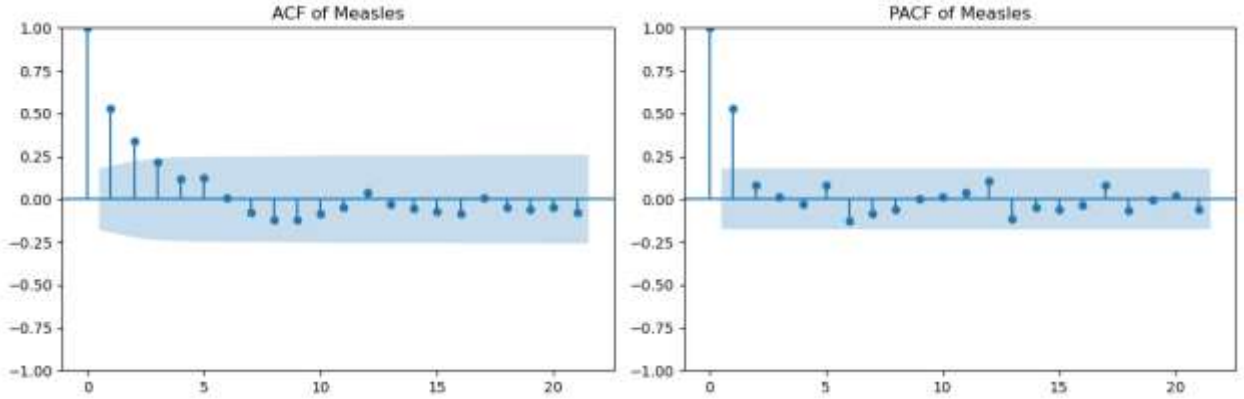


Figure 4.4: ACF and PACF plots on Measles

Figure 4.4 displays the graphs for the autocorrelation function (ACF) and partial autocorrelation function (PACF) concerning measles cases. The ACF evaluates how the current series relates to its past values, while the PACF assesses the correlation at a specific lag by considering the impacts of earlier lags. These diagrams are instrumental in recognizing possible autoregressive and moving-average models, as well as in determining if significant seasonal lags exist. The notable Ljung–Box statistics found in Table 4.3 highlight the importance of exploring the temporal relationships rather than assuming the independence of measles data points. When choosing a model, it is essential to analyze the ACF and PACF alongside the stationarity test, the frequency of observations, and the seasonal decomposition. Subsequently, potential models should be evaluated against appropriate information criteria, like AIC or BIC, and examined through residual diagnostics. An adequately fitted model should produce residuals that do not display significant autocorrelation, which can be verified using the residual ACF and a Ljung–Box test on the residuals from the fitted model. In cases where the outcome consists of non-negative counts and the distribution is markedly skewed, consideration may also be given to count-based time-series models, provided their assumptions are satisfied. The data set was fitted to Holt-Winters Model for modeling actual, level, trend, seasonal and forecast of selected childhood disease (Measles) in Rivers State. The Additive Holt-Winters Model for measle is given as:

$$\begin{aligned} L_t &= 0.4775\,(Y_t - S_{t-12}) + (1 - 0.4775)(L_{t-1} + T_{t-1}) \\ T_t &= 0.0\,(L_t - L_{t-1}) + (1 - 0.0)\,T_{t-1} = T_{t-1} \\ S_t &= 0.0\,(Y_t - L_t) + (1 - 0.0)\,S_{t-12} = S_{t-12} \\ \hat{Y}_{t+h} &= L_t + hT_t + S_{t-12+h_m} \end{aligned}$$

Initial conditions: $L_0 = 1.5568, T_0 = 0.2429,$ S_0= [5.89, 7.24, 16.20, 0.15, −3.60, −11.64, −11.19, −5.33, −11.47, −2.91, −2.95, −4.29]

Multiplicative Holt-winters model for measle is given as:

$$\begin{aligned} L_t &= 0.4181\left(\frac{Y_t}{S_{t-12}}\right) + (1 - 0.4181)(L_{t-1} + T_{t-1}) \\ T_t &= 0.0\,(L_t - L_{t-1}) + (1 - 0.0)\,T_{t-1} = T_{t-1} \\ S_t &= 0.0\left(\frac{Y_t}{L_t}\right) + (1 - 0.0)\,S_{t-12} = S_{t-12} \\ \hat{Y}_{t+h} &= (L_t + hT_t) \cdot S_{t-12+h_m} \end{aligned}$$

Initial Conditions: $L_0 = 25.12\,, T_0 = 2.485,$ S_0=[0.0947, 0.0983, 0.1317, 0.0828, 0.0645, 0.0437, 0.0459, 0.0596, 0.0501, 0.0669, 0.0725, 0.0726

Table 4.4: Comparison of Additive and Multiplicative Holt–Winters Models for Measles

Model	Evaluation period	RMSE	RMSE Rank	MASE	MASE Rank
Additive Holt–Winters	Jan. 2016–Dec. 2024	17.54	2nd	99.56	2rd
Multiplicative Holt–Winters	Jan. 2016–Dec. 2024	16.43	1st	86.68	1st

The additive Holt–Winters technique for measles employs initial seasonal components ([5.89, 7.24, 16.20, 0.15, -3.60, -11.64, -11.19, -5.33, -11.47, -2.91, -2.95, -4.29]) to illustrate the anticipated monthly seasonal influences. This form of modeling indicates seasonal effects within the same metrics as the actual measles cases, where positive figures signify months in which the seasonal aspect raises the projected cases above the basic level and trend, while negative figures denote months where it decreases the projected cases. The maximum positive seasonal influence is 16.20, and the most negative influence is −11.64, revealing significant fluctuations in the initial

seasonal trend. The figures imply that measles occurrences might show variation throughout the months, yet the exact months linked to these impacts need to be recognized from the sequence of the seasonal vector and validated against the actual data. These seasonal components can aid in predictions if the pattern remains stable over time. Conversely, if the seasonal smoothing parameter is set to zero, the model preserves these original seasonal effects without adjustments as new data comes in, potentially hindering its capacity to reflect evolving trends in measles transmission.

The multiplicative Holt–Winters model starts with the initial seasonal factors ([0.0947, 0.0983, 0.1317, 0.0828, 0.0645, 0.0437, 0.0459, 0.0596, 0.0501, 0.0669, 0.0725, 0.0726]). The highest factor stands at 0.1317, whereas the lowest is 0.0437, illustrating the variability in the relative seasonal weights attributed to different months. Typically, in a standard multiplicative Holt–Winters model, seasonal factors tend to be centered around an average value of 1; thus, a vector with values entirely under 1 necessitates evaluation. These numbers might signify normalized ratios rather than traditional multiplicative seasonal indices. If they are meant to represent standard seasonal factors, it is essential to verify the seasonal normalization and model application prior to interpreting any results. Additionally, the initial estimates of level and trend are not included in the provided information, so their effects on the fitted values and forecasts cannot be evaluated.

From a statistical viewpoint, both models aim to depict the level, trend, and repeated seasonal patterns of measles cases in children under five in Rivers State. The additive model is usually suitable when seasonal changes remain relatively stable in absolute terms, while the multiplicative model fits better when seasonal changes occur in relation to the level of the series. The decision on which model to use should depend on the data's observable pattern and their comparative predictive accuracy, measured by metrics such as MAE and RMSE across a consistent validation timeframe. Considering the provided seasonal values, components of the additive model can be interpreted in case counts, but clarification of the multiplicative seasonal factors is essential before making a credible comparison. It is important to note that neither model can be deemed superior without verified fitted values, prediction errors, and diagnostic checks on residuals.

Table 4.4 compares the predictive performance of the additive and multiplicative Holt–Winters models for measles cases over the evaluation period January 2016 to December 2024. The multiplicative Holt–Winters model performed better, recording the lowest RMSE of 16.43 and MASE of 86.68, compared with 17.54 and 99.56, respectively, for the additive Holt–Winters model. Based on these results, the multiplicative model ranked first for both evaluation metrics, while the additive model ranked second. The multiplicative model therefore provides more accurate forecasts of measles cases and is preferred for subsequent comparative forecasting analysis. Similarly, the seasonal model was estimated and model comparison for measles in Table 4.5:

Table 4.5: Model Comparison for Measles

Disease	Model	AIC	BIC	RMSE	MAPE	Best Fitted
Measles	SARIMA(1,1,0)(0,0,0)[12]	1043.55	1049.09	19.66	88.78	
	SARIMA (0,1,1)(0,0,0)[12]	1030.01	1035.54	19.20	88.64	
	SARIMA (1,1,1)(0,0,0)[12]	1020.69	1028.98	18.31	104.41	
	SARIMA (1,1,1)(1,1,0)[12]	834.66	844.83	19.90	92.24	
	SARIMA (0,1,1)(1,1,1)[12]	827.12	837.25	20.17	89.84	SARIMA (0,1,1) (1,1,1)[12]

Table 4.5 displays the performance comparison of five SARIMA models fitted to the measles time series data from Rivers State. The evaluation of these models was carried out using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), root mean squared error (RMSE), and mean absolute percentage error (MAPE). These metrics offer different perspectives on how suitable each model is: lower values of AIC and BIC point to a more advantageous trade-off between the model's accuracy and its complexity, while reduced RMSE and MAPE values suggest lesser forecasting errors, considering the constraints associated with each metric. The SARIMA (1,1,0) (0,0,0) [12] model had an AIC value of 1043.55, a BIC of 1049.09, an RMSE of 19.66, and a MAPE of 88.78%. The SARIMA (0,1,1) (0,0,0) [12] model showed enhancements in these figures, achieving an AIC of 1030.01, a BIC of 1035.54, an RMSE of 19.20, and a MAPE of 88.64%. Meanwhile, the SARIMA (1,1,1) (0,0,0) [12] model further decreased the AIC and BIC to 1020.69 and 1028.98, respectively, and attained the best RMSE among the three non-seasonal configurations, measuring 18.31. Nevertheless, its MAPE increased to 104.41%, signifying that the relative percentage errors were larger despite the better RMSE. This discrepancy highlights the importance of not relying solely on one performance metric for model selection. The addition of seasonal differencing, as well as seasonal autoregressive or averaging factors, led to a significant drop in the information criteria. The SARIMA (1,1,1) (1,1,0) [12] model achieved an AIC of 834.66 and a BIC of 844.83, along with an RMSE of 19.90 and a MAPE of 92.24%. The SARIMA (0,1,1) (1,1,1)[12] model yielded the lowest AIC and BIC values of 827.12 and 837.25. These outcomes suggest that, out of the five models examined, the latter model offers the most advantageous balance of fit and complexity in line with the information criteria. However, its RMSE of 20.17 is the highest when compared to the other models, and its MAPE of 89.84% exceeds the MAPE values of the initial two models. As a result, it is considered the best model based on the AIC/BIC criteria, yet it does not rank as the top model when evaluated through all measures of forecasting accuracy. The final decision should be guided by the forecasting goals of the study and performance during a standard validation timeframe. The MAPE figures, which range from 88.64% to 104.41%, indicate significant relative forecast errors. However, MAPE can be unreliable when the observed counts of measles are zero or very close to zero, a situation that is particularly pertinent for data concerning childhood diseases. Therefore, these figures should be interpreted with caution, and further evaluation using MAE, RMSE, and an appropriate scale-independent metric like mean absolute scaled error (MASE) would enhance the comparison. The validation process must also ensure that each potential model is evaluated using the same dataset and forecasting timeframe. Also, Hybrid SARIMA–ANN model specification for modeling measles in Rivers State was considered and the results is shown in **Table 4.6.**

Table 4.6: Hybrid SARIMA–ANN Model Specification for modeling Measles in Rivers State.

Component	Parameter/Layer	Specification / Value
ARIMA Model	Order (p, d, q)	(0, 1, 1)
	Seasonal Order (P, D, Q, s)	(1, 1, 1) [12]
	AIC	827.12
	BIC	837.25
	MA(1)	-0.697728
	Seasonal AR(12)	-0.265599
	Seasonal MA(12)	-0.259263
	Sigma²	388.147271
ANN Model	Layer 1	Dense (50 units, ReLU)
	Layer 2	Dense (25 units, ReLU)
	Layer 3	Dense (1 unit, Linear)

Table 4.6 outlines the details of the chosen hybrid SARIMA–ANN model utilized for forecasting measles cases. The SARIMA part features a non-seasonal order of (0,1,1) alongside a seasonal order of (1,1,1) _ {12}. The non-seasonal differencing indicated by d=1 signifies first differencing, while the seasonal differencing represented by D=1 indicates differencing at lag 12 to account for a potential yearly trend in monthly data. The non-seasonal MA (1) coefficient is −0.697728, the seasonal AR (12) coefficient is −0.265599, and the seasonal MA (12) coefficient is −0.259263. These values reflect the relationship structure shown by the fitted SARIMA model; however, statistical significance cannot be determined without standard errors, confidence intervals, or p-values for the coefficients. The reported innovation variance of σ^2=388.147271 translates to an innovation standard deviation of roughly 19.70 cases. This indicates the magnitude of the unaccounted variability remaining in the SARIMA aspect, presuming the variance is represented in the original case count format. It should not be seen as a standalone prediction confidence interval. The provided AIC value of 827.12 and BIC value of 837.25 align with the chosen SARIMA setting displayed in Table 4.4. Nevertheless, these metrics are relevant to the SARIMA model and should not be automatically regarded as the information criteria for the full hybrid SARIMA–ANN model.

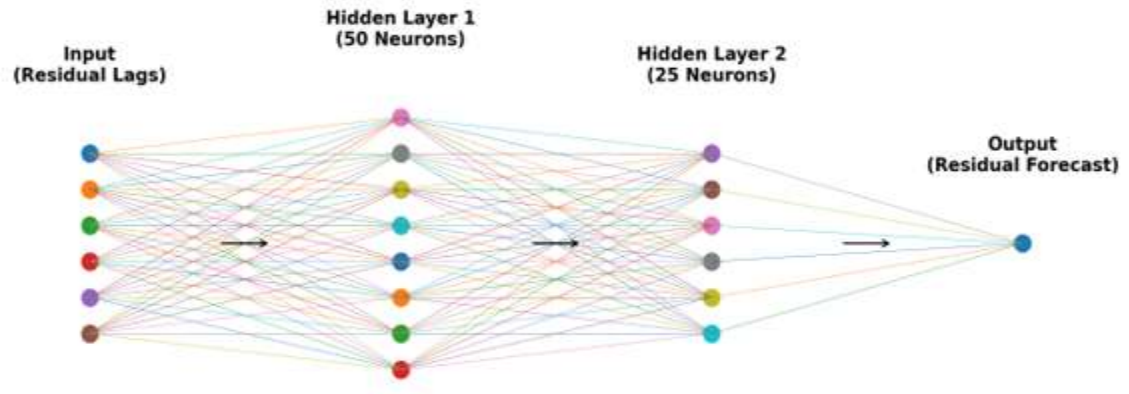


Figure 4.5: Virtual Structural Architecture of Hybrid SARIMA–ANN Model for Measles

Figure 4.5 depicts the theoretical framework of the hybrid SARIMA–ANN model. The SARIMA aspect models linear time-based dependencies and seasonal patterns, while the artificial neural

network aims to capture the nonlinear trends remaining in the SARIMA residuals. The ANN features an initial dense layer comprising 50 units with ReLU activation, a subsequent dense layer with 25 units and ReLU activation, and a final output unit with linear activation. The ReLU activation functions enable the hidden layers to express nonlinear correlations, while the linear output layer produces a continuous numerical forecast. When the ANN is trained on SARIMA residuals, the predicted residual part can be merged with the SARIMA forecast to create the hybrid forecast.

The justification for the hybrid model lies in the observation that time series regarding childhood diseases may exhibit both linear temporal relationships and nonlinear variations that a single SARIMA model may not sufficiently address. However, the model's design by itself does not confirm that the hybrid model enhances forecasting accuracy. Its effectiveness must be validated by juxtaposing its predictions against those of the independent SARIMA model using the same held-out data. Residual diagnostics should verify if the SARIMA residuals possess a significant structure for the ANN to learn from, and the ANN training should occur without utilizing future data that would not have been accessible at the time of the forecast. Additionally, details regarding the number of units within the network, the training methodology, the input features, and any regularization or early-stopping techniques should be documented to ensure reproducibility and mitigate overfitting.

Table 4.7: RMSE and MASE Comparison of Additive, Multiplicative Holt–Winters, SARIMA (0,1,1)(1,1,1)[12] and Hybrid SARIMA–ANN Models for Measles

Model	Evaluation period	RMSE	RMSE Rank	MASE	MASE Rank
Additive Holt–Winters	Jan. 2016–Dec. 2024	17.54	2nd	99.56	3rd
Multiplicative Holt–Winters	Jan. 2016–Dec. 2024	16.43	1st	86.68	1st
SARIMA (0,1,1) (1,1,1)[12]	As reported	20.17	3rd	89.84	2nd
Hybrid SARIMA–ANN	As reported	22.05	4th	261.66	4th

Table 4.7 showcases the relative forecasting effectiveness of the additive Holt–Winters, multiplicative Holt–Winters, SARIMA (0,1,1)(1,1,1)[12], and integrated SARIMA–ANN models for monthly instances of measles. According to the RMSE, the multiplicative Holt–Winters method stands out as the most effective, achieving the lowest RMSE of 16.43, followed by the additive Holt–Winters at 17.54, the SARIMA model at 20.17, and the hybrid SARIMA–ANN model at 22.05. These findings suggest that the multiplicative Holt–Winters approach generates the smallest root mean squared forecast error out of all four methods. In terms of MASE, the multiplicative Holt–Winters model again ranks highest with a score of 86.68, followed by the SARIMA method at 89.84, the additive Holt–Winters at 99.56, and the hybrid SARIMA–ANN at 261.66. Consequently, the multiplicative Holt–Winters model displays the superior performance according to both assessment metrics. The hybrid SARIMA–ANN model is positioned last in terms of both RMSE and MASE, reflecting the lowest effectiveness among the models examined. Based on the RMSE and MASE figures provided, the multiplicative Holt–Winters model is identified as the most favorable. Nonetheless, it is important to validate the listed MASE values prior to their application in a thesis. MASE is typically a scaled, dimensionless metric, and scores ranging from 86.68 to 261.66 are exceptionally high. Verify that MASE was computed by taking the model's mean absolute forecast error and dividing it by the seasonal-naïve mean absolute error from the training series. Additionally, confirm that all models were assessed over the same testing time frame using identical observations before drawing a conclusive comparison.

Discussions of the findings

The research investigates the temporal trends, spread, stationarity, and time dependence of measles incidents in children under five years old in Rivers State. The descriptive analysis reveals notable variation and a positive skew in measles cases, while the time-series breakdown offers a method to analyze the underlying trends and seasonal patterns. This aligns with findings from Ojasanya et al. (2025), who noted recurring peaks in measles cases from January to March in Nigeria, and Nwadiuto et al. (2025), who documented repeated outbreaks in Rivers State during the years 2020 to 2022. Additionally, the Augmented Dickey-Fuller test results ($p < .001$) show stationarity under the conditions of the test, while the significant results from the Ljung-Box test indicate serial dependence. These observations support the use of time-series modeling, although the necessity for differencing should be evaluated through model diagnostics. Ojasanya et al. (2025) similarly utilized time-series decomposition and SARIMA for modeling measles cases in Nigeria.

In other findings, the comparative analysis of additive and multiplicative Holt-Winters models for forecasting revealed that the multiplicative version had a superior performance, recording an RMSE of 16.43 and MASE of 86.68, contrasted with 17.54 and 99.56 for the additive model. This aligns with the broader conclusion of Adnan and Berke (2026), which indicated that Holt-Winters forecasting methods can be competitive for measles monitoring, although their research utilized a different dataset and evaluation criteria. Furthermore, to determine the most suitable SARIMA model for forecasting measles cases, the SARIMA ((0,1,1)(1,1,1)_{12}) had the lowest AIC (827.12) and BIC (837.25), signifying the optimal fit-complexity balance among the five models assessed. Nonetheless, its RMSE was higher than some other competing models, suggesting that out-of-sample forecast accuracy should also be a criterion for selection. Similarly, Ojasanya et al. (2025) chose a SARIMA approach for predicting measles incidence in Nigeria, employing information criteria and validation of forecasts. To assess if a hybrid SARIMA-ANN model enhances forecasting precision, it was found that this model yielded an RMSE of 22.05 and MASE of 261.66, indicating a lesser performance compared to other models evaluated. Consequently, the findings do not endorse its effectiveness for this dataset. Parreño (2025), who analyzed SARIMA, Holt-Winters, and neural network methods for measles prediction in the Philippines, offers a relevant benchmark for comparison; however, the performance is contingent upon the dataset used. Therefore, the findings indicate that the multiplicative Holt-Winters model is the optimal forecasting choice based on the observed RMSE and MASE values. Prior to finalizing the thesis, it is advised to verify the unusually high MASE figures and confirm that all models were assessed using the same test observations and forecasting periods.

Conclusion

The study determined that the number of measles cases each month among children under five in Rivers State showed considerable fluctuation, a lack of normal distribution, and a notable level of serial correlation, indicating that time-series forecasting methods are appropriate. Of the different models assessed, the multiplicative Holt-Winters method yielded the most effective forecasting results, indicated by the lowest RMSE and MASE values, whereas the hybrid SARIMA-ANN model was found to be the least effective. These results imply that simpler seasonal models could be more successful than their more complicated counterparts for this particular dataset; nevertheless, the reported MASE values, validation approaches, and model residuals must be checked before any definitive conclusions are made.

Recommendations

1. Health officials in Rivers State ought to utilize verified measles predictions to enhance vaccine distribution, improve planning for surveillance, and implement targeted immunization efforts during times of heightened incidence.
2. Subsequent analyses need to confirm the accuracy of MASE computations, evaluate all models using the same test periods, and perform residual diagnostics along with rolling-origin validation to guarantee dependable forecasting results.
3. Future research should factor in vaccination rates, weather conditions, demographic details, and information about outbreaks into both statistical and machine-learning models to enhance predictive precision and bolster evidence-based strategies for measles prevention.

References

Awogbemi, C. A., Deebom, Z. D, Amos, E., Augustine, O. E., Victor, R., Adekola, A. S., Kolawole, I. A., Ifeoma, U. K., James, J. O., Kehinde, O. C., Vincent, A. E., Okeke, C. C., Da-Wariboko, Y. A., & Smart, A. I. (2026). Forecasting the occurrence of diarrhea disease using Facebook Prophet, statistics and machine learning models. JP Journal of Biostatistics, 26(2), 251–282.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). John Wiley & Sons.

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(366a), 427–431.

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297–303.

Nwadiuto, I., Mberekpe, O. P., Briggs, C. T., Green, P. A., Onyekwere, N. V., Owhonda, G., & Wekere, F. (2025). Spatiotemporal patterns of measles outbreaks in Rivers State, Nigeria (2020–2022). World Journal of Public Health, 10(3), 398–406.

Ojasanya, R., Awosile, B. B., Adeyemo, P., Jarra, E., & Berke, O. (2025). Time series analysis of measles incidence in Nigeria using surveillance data from 2011 to 2022. McGill Journal of Global Health, 14(1), 23–27.

Parreño, S. J. E. (2025). Forecasting measles incidence in the Philippines: A comparative analysis of SARIMA, Holt-Winters, ESN, and NNAR models. Multidisciplinary Science Journal, 7(7), Article e2025356.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management Science, 6(3), 324–342. https://doi.org/10.1287/mnsc.6.3.324

World Health Organization & Centers for Disease Control and Prevention. (2024, November 14). Measles cases surge worldwide, infecting 10.3 million people in 2023. https://www.who.int/news/item/14-11-2024-measles-cases-surge-worldwide--infecting-10.3-million-people-in-2023

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175. https://doi.org/10.1016/S0925-2312(01)00702-0