ISSN: 3029-0716

Research Article

Journal of Infectious Diseases & Treatments

Time-Series Forecasting of Lassa fever Outbreaks in Nigeria using Facebook Prophet; Application of Machine Learning in Early Warning, Alert and Response Systems

Franco Apiyanteide*, Adenonmon Monday Osagie, Chaku Shammah Emmanuel, Abubakar Muhammad Auwal and Justus Ajaegbu Ezechinyere

Centre for Cyberspace Studies, Nasarawa State University, Keffi, Nasarawa State, Nigeria

***Corresponding author**
Franco Apiyanteide, Centre for Cyberspace Studies, Nasarawa State University, Keffi, Nasarawa State, Nigeria.

Received: April 13, 2026; **Accepted:** April 20, 2026; **Published:** April 27, 2026

ABSTRACT
Since the first cases of Lassa fever were diagnosed in Northeastern Nigeria in 1969, the disease has been one of the neglected tropical diseases with no definitive cure, no approved vaccines and poor public health measures needed to mitigate the burden of the disease in the country. This study aimed to provide time series forecasting of Lassa fever in Nigeria using Facebook Prophet, to improve the understanding of the seasonal patterns and trends of the disease in Nigeria. Data for the study were obtained from the Nigeria Centre for Disease Control and Prevention (NCDC) weekly situation report website from 2018 to 2025, using a Microsoft Excel template that captured relevant information on the disease (year, epidemiological week, number of cases, and number of deaths) from the agency's situation report. Data collected was processed, split into training and testing datasets that were used for forecasting and performance evaluation of the model. The model was fit and used to make a 5-year forecast using the Facebook Prophet Library in Python version 3.13 to obtain a 5-year forecast for the disease in Nigeria from 2025 to 2030, as well as gain an in-depth understanding of the seasonality and trend of the disease. Our result showed that a total of 7728 Lassa fever confirmed cases and 1288 confirmed deaths, with a Case Fatality Rate (CFR) of 16.7% were recorded in Nigeria from 2018 to 2025. There was an upward trend for the number of cases and deaths from Lassa fever over the study period, with January to March, which corresponds to the dry season in the country, being the peak for the disease. Performance evaluation was conducted for confirmed cases of the disease using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R2) and Mean Poisson Deviance (D) and this showed 9.44, 12.51, 0.57 and 6.3 respectively. Using Facebook Prophet, by 2030, there will be an estimated 9100 (95% Confidence Interval (CI): 1937 – 16268) confirmed cases and 1721 (95% CI: 313 – 3125) confirmed deaths from Lassa fever in Nigeria. There is therefore an urgent need to apply machine learning algorithms in understanding infectious disease trend and seasonality which could guide early warning, alert and response systems which is proactive rather than the traditional reactive response to outbreaks when several morbidity and mortality have already occurred.

Keywords: Time series forecasting, Lassa fever, Facebook Prophet, Early warning, alert and response systems, Machine Learning, Algorithm, Nigeria

Introduction

Since the first cases of Lassa fever were described in 1969 following the death of two missionary nurses in Lassa town, North east Nigeria the disease has continued to cause havoc and spread across Africa with sub-optimal prevention and control measures which has made some countries in the continent to become epicenter of the disease [1-3].

The World Health Organisation described Lassa fever as an acute viral disease caused by a single-stranded RNA virus that is endemic in West Africa, with annual outbreaks amongst countries where the virus is endemic [4]. Ogbu noted that annually, there are estimated 300,000 – 500,000 cases of Lassa fever with 5,000 deaths each year [5]. Estimates from the Africa Centre for Disease Control and Prevention revealed that there are an estimated 10,000 – 30,000 cases of Lassa fever that occur yearly, with 5,000 deaths due to this condition [3]. Data from the Nigeria Centre for Disease Control and Prevention revealed that Nigeria has experienced several outbreaks with a rising trend in the number of confirmed cases and associated mortality reported across some states in the country [6]. In 2012, there was an estimated 273 confirmed cases and 149 confirmed deaths due to Lassa fever, in 2018, 413 confirmed cases with 114 confirmed deaths occurred, in 2023, there was an estimated

Citation: Franco Apiyanteide, Adenonmon Monday Osagie, Chaku Shammah Emmanuel, Abubakar Muhammad Auwal, Justus Ajaegbu Ezechinyere. Time-Series Forecasting of Lassa fever Outbreaks in Nigeria using Facebook Prophet; Application of Machine Learning in Early Warning, Alert and Response Systems. J Infect Dise Treat. 2026. 4(2): 1-10. DOI: doi.org/10.61440/JIDT.2026.v4.70

877 confirmed cases with 152 confirmed deaths and in 2024, the outbreak was severe and fatal with 1059 confirmed cases and 175 deaths across 28 states in the country [6]. The disease is highly endemic in Nigeria, with an estimated 28,780 suspected cases, 4,036 confirmed cases, and 762 deaths with a case fatality rate (CFR) of 18.9% across 34 of the 37 states in Nigeria noted over the years (Al-Mustapha et al., 2024). These reports clearly demonstrate that there is a rising trend in the number of cases and mortality from Lassa fever, and this could be associated with the absence of an effective early warning system that can timely detect and enable effective and efficient preventive measures to mitigate the burden of the disease in Nigeria. In addition to the rising trend in the disease, nosocomial infection due to the disease is well established and the Africa CDC, have reported that outbreaks of the disease in Benin, Liberia and Nigeria in 2012 and 2016 are associated with the spread of the infection to health care workers including community health extension workers, nurses and medical doctors (4 – 10 health care workers) [3].

Lassa fever is an acute viral illness caused by the Lassa virus that belongs to the virus family Arenaviridae [4]. It is a zoonotic disease that is transmitted by multi-mammate rats, and is responsible for deadly epidemics of haemorrhagic fevers in the West African sub-region. Rural communities and in hospital settings have higher vulnerability for the disease with higher outbreaks in this region that is fuelled by socio-cultural practices, poor environmental and personal hygiene and poor practice of infection prevention and control [7]. The clinical presentation is protean, and diagnosis is often delayed, with Ribavirin being the mainstay and effective life-saving treatment for the disease when administered within six days of onset of symptoms [7]. An estimated 80% of people who become infected with the Lassa virus express no or mild symptoms, and 1 in 5 infections result in severe disease that affects several organs like the liver, spleen and kidneys [4]. Diagnosis of the disease can be a challenge as it mimics other endemic diseases like malaria, typhoid fever, yellow fever and other haemorrhagic fevers. A high index of suspicion, confirmatory investigations such as reverse transcriptase polymerase chain reaction (RT-PCR) assay, antibody enzyme-linked immunosorbent assay (ELISA), antigen detection tests, and virus isolation by cell culture are needed to make a confirmatory diagnosis for the disease [8]. To date, there is no licensed vaccine for the disease, and the absence of vaccines, especially for frontline healthcare workers and locations with high endemicity of the virus, is and possibly contributes to the sustained and increasing burden of the disease across Nigeria and other West African countries where the disease is endemic [4,6].

To prevent and control the spread of the Lassa fever virus in Nigeria, the WHO and NCDC are implementing a joint, multi-sectoral action plan to combat Lassa fever, focusing on strengthening surveillance, enhancing laboratory capacity, improving case management (including infection prevention and control), and boosting public awareness. The strategy, targeting a reduction in the case fatality rate to below 10%, utilises a "One Health" approach, incorporating rodent control and rapid, coordinated response teams, particularly during peak seasons. Recent efforts for the control and prevention of Lassa fever and other infectious diseases with pandemic potential have focused on the implementation of artificial intelligence and machine learning techniques for early warning and proactive response to infectious diseases [9].

Lassa fever is a dreaded viral disease with a pandemic potential that the Global Health Security Agenda is designed to address, among other diseases. Moore, in their Lassa fever research for effective countermeasures, noted that in 2016, the WHO designated Lassa fever as a priority disease for epidemic preparedness as part of a blueprint for Action in the prevention of epidemics [10]. In Nigeria, various efforts have been put in place for the control and prevention of the disease despite a few shortcomings. Among these efforts are the establishment of a case management guideline, which was developed in 2018 and the setup of a technical working group that is mandated to provide Lassa fever Public Health Advisory aimed at improving preparedness, readiness and response to the disease [11]. Other public health measures carried out include the conduct of training for health care workers and community members on signs and symptoms of the disease for early detection, prevention and control of the infection, as well as guidance against nosocomial transmission and spread of the disease during outbreaks. The institution has also established a website where regular situation reports for outbreaks of the disease are provided to the general public, with the overwhelming goal of encouraging the use of data to guide decision-making processes aimed at the prevention and control of the disease [6].

Despite these measures that are designed and being implemented across Nigeria, gaps still exist, especially in the application of machine learning techniques that can help in forecasting disease outbreaks and guide early detection for timely prevention and control of the disease, which is needed to minimise morbidity and mortality from Lassa fever in the country. Prediction models such as the Autoregressive Integrated Moving Average (ARIMA) and machine learning models such as random forest have played crucial roles in time series forecasting of Lassa fever outbreaks and other diseases such as COVID-19 across Nigeria and other regions of the world with great accuracy. These models come with a lot of rigour that requires highly trained and skilled individuals in data science, for which the skill is often lacking. To address this challenge, Meta (formerly known as Facebook) developed the Facebook Prophet, which provides time series forecasts promptly as well as ensures easy decomposition of the time series into seasonality, trend, and holiday components that could affect predictions of events such as disease outbreaks using historical data [12,13]. The Prophet has the advantage of addressing missing data and has been shown to be great in forecasting disease outbreaks from time series data with missing components, as often observed in real-world data collected over a long period of time.

Time series forecasting is very crucial to understanding disease trends and burden across the world, and several studies have been done to forecast various diseases such as COVID-19, and other diseases using various statistical, mathematical and machine learning techniques such as ARIMA and Random Forest to improve understanding of disease outbreaks, trends and seasonality [12-17]. The results from these forecasts are often encouraging and provide an in-depth understanding of disease trends and patterns, which is necessary to guide logistics and public health response to a disease outbreak. Facebook Prophet has the peculiarity of providing timely forecasts and understanding of the trend, seasonality and other factors that could affect disease outbreaks, such as the holiday period [13]. This makes Prophet an important tool for time series forecasting, early warning and alert systems and effective surveillance of diseases across the globe.

Despite the flexibility and effectiveness of Facebook Prophet in time series forecasting, which it was specifically designed for, the integration of this novel machine learning technique for early warning systems for timely detection of outbreaks and understanding of seasonality and trend is limited in Nigeria and other low- and middle-income countries [13]. This study, therefore, sought to make a time series forecast for Lassa fever in Nigeria using Facebook Prophet with the ultimate goal of integrating findings into early warning systems for a robust and proactive surveillance system in Nigeria and other West African countries where the disease is endemic.

Methods

Study Design

Study design refers to the plan or the blueprint for conducting a study in a systematic manner and involves data collection, data cleaning, data analysis and presentation of the results and discussion of findings from the study. Basically, study designs are grouped into observational studies, such as cross-sectional studies and experimental study such as randomised controlled trials [18]. For this study, a retrospective ecological time series modeling study design that involves the collection and analysis of time series weekly data obtained from the weekly epidemiological record of Lassa fever cases in Nigeria over a period of eight years (2018 – 2025) was employed. Time series (weekly) data on the number of confirmed cases and number of confirmed deaths from Lassa fever were collected from the Nigeria Centre for Disease Control and Prevention (NCDC) website that provides the situation report for the disease.

Study Setting

This study was carried out in Nigeria, which is the most populous black nation in the world that accounts for almost half of the population in West Africa, situated between the Sahel to the North and the Gulf of Guinea in the Atlantic Ocean to the South. In 2024, there were an estimated 232,679,478 persons [19]. The country has multiple ethnic groups who are dispersed across the thirty-six states and the Federal Capital Territory in the nation. Nigeria is endowed with a large land mass of about 923,769 square kilometres (356,669sqmi) with several natural resources such as crude oil, gas, gold and others [20]. Despite the abundance of natural resources and a large pool of young people, the country is among the poorest nations on earth, with one of the weakest health systems, and it is ranked 11th out of the 12 countries with the worst healthcare systems globally [21]. This has contributed to the vulnerability of various populations across Nigeria to infectious disease such as Lassa fever which can be controlled by the implementation of proactive surveillance that is based on early warning and alert systems which integrates innovations such as machine learning techniques like Facebook Prophet to forecast and understand disease seasonality and trend promptly to guide logistics, operational and policy design and implementation aimed to protect lives, prevent diseases and improve the health outcomes of populations across the country.

Data Collection

Data for this study were collected from the Nigeria Centre for Disease Control and Prevention (NCDC) situation report website and span from January 2018 to December 2025. Time series (TS) weekly data for each epidemiological week were collected from the website and recorded on a Microsoft Excel Sheet. These datasets (time series) are data points which are indexed in time order and are equally spaced in time from one data point to the other [22]. They form a sequence of discrete-time datasets with a chronological sequence of observations that demonstrates the changing state of any variable over time [23,24]. A key feature of time-series data is temporal dependency, which explains the fact that observations that are closer together in time are more closely related than those further apart, and this is often considered in time series forecasting [22]. Time-series data also exhibit properties such as trend, periodicity, seasonality, and randomness. Trend refers to long-term changes, periodicity to regular rises or falls with unfixed frequency, seasonality to changes at a known and fixed frequency which could be influenced by seasons, and randomness to irregular changes that satisfy statistical laws [25]. These types of data are prevalent in clinical medicine and public health, where they are employed for forecasting of disease outbreaks needed to guide public health decision-making and timely response [23,24].

Data Quality and Limitation of Surveillance Data

Data quality is an important component of machine learning and predictions made from machine learning techniques are as good as the quality of the data. To ensure data quality, several steps were observed. First, data used for this study are exclusively confirmed cases and confirmed deaths due to Lassa fever that are verified by the NCDC before uploading to their website. A confirmed case of Lassa fever is defined as defined as any suspected case that has been validated through laboratory testing, specifically showing positive IgM antibodies, PCR (polymerase chain reaction), or virus isolation [4]. This was the case definition used by the NCDC for classifying a case as confirmed case or death from Lassa fever. The use of the case definition for the disease reduces misdiagnosis that could introduce into the dataset. Secondly, a two-stage data collection and review process was ensured by engaging a research assistance who was trained in the data collection process and was given a predefined Microsoft Excel template to collect data from the NCDC website. This was followed by the review of the data collected by the principal investigator for this study to ensure double checking and correctness of the dataset collected from the website. Manual collection of data was done as it was difficult to use Python commands and packages such as Beautiful Soup to automatically obtain the data from this website due to the variation in the manner in which the table for the dataset were presented over the years and other design challenges. Missing values observed in the dataset were dropped using the Python code dropna() before splitting the dataset into training and testing dataset used for modeling and performance evaluation of the metrics for forecasting Lassa fever cases in

Nigeria. Also, for a week in 2018 where the dataset on the website was inaccessible, this was tracked through other reports by the NCDC while 2016 to 2017 dataset was not used because it has a lot of missing data which could affect the performance of the model. Hence data was limited to manually retrieved dataset within a time span of 8years (2018 to 2025) from the reported cases of Lassa fever uploaded on the NCDC website. Also, there are possible unreported cases in hard-to-reach locations across Nigeria which could affect the exact picture of the burden of the disease. Howbeit, the available and collected dataset is adequate and good for the purpose of this study.

Statistical Analysis

Data obtained from this study was used to conduct a time series forecast of Lassa fever for a period of 5years (2025 – 2030) using Facebook Prophet. Relevant packages, libraries and modules used include Pandas, Numpy, Matplotlib, Seaborn, and Facebook Prophet in Python version 3.13.

8-years weekly epidemiological dataset from 2018 to 2025 was collected to carry out time series analysis of Lassa fever, and then a 5-year forecast for 2025 – 2030 for Lassa fever in Nigeria using the historical dataset collected from the NCDC Website. Data collection was done by research assistance and thereafter a review of the dataset was done by the principal investigator for this study before time series analysis and forecasting of Lassa fever cases and deaths in Nigeria.

Time series analysis is a sophisticated statistical method that can reveal patterns and give a clear direction of the observed dataset. It allows for the examination of data points collected and recorded at specific time intervals, enabling the identification of trends, patterns, and seasonal variations crucial for making informed predictions and decisions in public health and other sectors. Such analysis is critical in understanding the period/ season in the year where there are more cases and deaths from the disease, which could guide logistics and operational plans and strengthen proactive response to infectious diseases.

Several statistical approaches, such as Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) and machine learning techniques, such as Random Forest and Support Vector Machine, have been used to forecast infectious disease outbreaks. For this study, the Facebook Prophet was used to make a forecast of Lassa fever using the dataset that was collected over the 8-year study period (2018 – 2025). Prophet is an open-source time series forecasting library that was developed by Facebook (now Meta) Core Data Science team in 2017 and made publicly available as an open-source software for Python and R in February 2017. It is an additive model introduced for forecasting time series data [26]. It is capable of detecting non-linear trends in the time series, like yearly, monthly, weekly, and daily trends. It is also capable of detecting seasonality and holiday effects and thereafter combines these features to obtain the forecast value. The method employs a decomposable time series model with three components, namely, trend, seasonality, holidays and an error term, where trend represents growth, seasonality describes periodic effects within the trend, holiday effects capture sudden events that are predictable in time and the error term represents noise which are idiosyncratic changes that are not accounted for by trend, seasonality, or the holiday component [23]. It is a robust library in time series that is capable of addressing missing data, outliers and change points in the trend of a dataset, which are common in the real world. The model functions like a generalised additive model with time as a regressor parameter, which fits the linear and non-linear functions of the time component. Essentially, it achieves suitable estimates of the mixed data without putting much effort [27,28]. These reasons justify the use of Facebook Prophet as the forecasting tool for this study. There are multiple benefits to using FB-Prophet as a public health modelling tool, and these include the fact that it is designed to outperform common time series models through increased flexibility and by accounting for seasonal and holiday variations [23]. It is also easily integrated into common statistical software, including R and Python.

Mathematically, the Facebook Prophet is represented as:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Where:

y(t) is the forecast, g(t) refers to trend which describes the increase or decline of a variable over a long period of time, s(t) refers to seasonality which is the recurring patterns such as weekly, monthly, or yearly fluctuations observed in the dataset, h(t) refers to irregular but predictable events, such as national holidays that could affect the dataset for a particular period in time and ϵt - epsilon refers to the error term which accounts for any unusual changes not accommodated by the model [23].

To conduct a time series forecast using Prophet, the following steps were employed based on the method described by Chugh [23]. They include: import relevant libraries and modules for the forecast, upload the dataset, pre-process the dataset including data cleaning, set the data in a date time format, prepare the data for Prophet, split data into training and test set, fit the prophet model, make your forecast, compare predicted data with actual dataset, conduct model performance evaluation, plot your forecast, plot your seasonal component, review the forecast samples and develop your report.

To determine the performance of the model, various performance evaluation metrics can be used. For this study, model evaluation was done using specific time-dependent metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and Coefficient of Determination also known as R-squared (R2) to determine the performance of the model using 20% of the dataset as the test dataset while 80% of the training dataset was used for the forecasting of Lassa fever outbreaks in Nigeria in a train-test split of the 470 data points uploaded into the Python Notebook. These metrics are discussed below.

Mean Absolute Error (MAE)

$$MAE = \frac{\sum_{i=1}^{n} |yi - xi|}{n}$$

The mean absolute error (MAE) is the mean of the observed errors when comparing observations, in this case, predicted vs. observed cases of Lassa fever, where yi is the actual value for the ith observation, xi is the actual value for the ith observation, and n is the total number of observations. The values of MAE range from 0 to infinity, with 0 being a difference in forecast and actual value without observed errors. No absolute range of values has been established for accurate vs. inaccurate forecasting; thus, this metric was used in a relative and comparative fashion [17].

Mean Poisson Deviance (D)
This is represented mathematically as:

$$D = 2\sum (y_i \log(y_i / y_i) - (y_i / \mathrm{y}_i))$$

Where D, is the Mean Poisson Deviance, y is the actual value and y-hat (ŷ) is the predicted value from the cross-validation.

Root Mean Square Error (RMSE)
The RMSE defines the average difference when comparing the predicted and actual values. This is shown mathematically with the formula below:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(yi - yi)^2}$$

RMSE describes the standard deviation among residuals, where yi is the actual value, ŷi is the predicted value, and n is the number of observations [17].

R-Squared (R2)
The coefficient of determination can be interpreted as the proportion of the variance in the dependent variable that is predictable from the independent variables.

It is mathematically presented as:

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(X_i - Y_i)^2}{\sum_{i=1}^{m}(\overline{Y} - Y_i)^2}$$

Where:
R2 (R-Squared) = Proportion of the variance in the observed data that is explained by the model.
m = Total Number of observations in the sample (sample size).
i = Index that represents each of the observations (from 1 to m).
Xi = Fitted (Predicted) value from the regression model for observation i. Ŷi.
Yi = Actual (observed) value of the dependent variable for observation i.
Y = Average (Mean) of all observed values of Yi

The standard R2 ranges from 0 to 1 in regression, with an R2 of 1 described as a Perfect fit with the ability of the model to explain 100% of the variation, while an R2 of 0 implies that the model did not explain any variation. The worst value = −∞ while the best value = +1 [29].

Ethical Consideration
The data used for this study were secondary data retrieved from the NCDC weekly situation report for Lassa fever which is made available to the public. The dataset were anonymised and based on the Helsinki Declaration, it is impractical to obtain informed consent from individual Lassa fever cases hence informed consent was exempted. The study did not involve any intervention and data were used exclusively for public health and research purposes. All efforts by the NCDC to provide this dataset on their website for surveillance and to the public are well acknowledged.

Results
The results for this study are presented below with appropriate figures (Figures 1 – 5) and tables (Tables 1- 3).

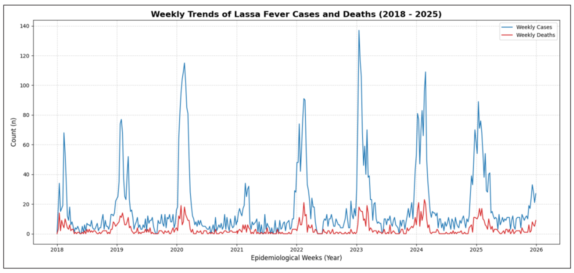


Figure 1: Weekly trend of Lassa fever Confirmed Cases and Confirmed Deaths in Nigeria from 2018 – 2025

Figure 1 show the weekly trend of Lassa fever confirmed cases and confirmed deaths in Nigeria over a period of 8-years from 2018 to 2025. The result showed that 2023, 2020 and 2024 had the highest number of cases and deaths, while 2021, 2018 and 2019 had the least number of confirmed cases and deaths from Lassa fever over the study period.

Table 1: Yearly Distribution of Lassa fever Confirmed Cases and Deaths in Nigeria from 2018 – 2025

S/N	Year	Confirmed Cases	Confirmed Deaths
1	2018	540	135
2	2019	796	155
3	2020	1190	159
4	2021	480	72
5	2022	1015	140
6	2023	1267	209
7	2024	1311	207
8	2025	1129	211
Total		7728	1288

Table 1 shows the yearly distribution of confirmed cases and deaths from Lassa fever in Nigeria from 2018 to 2025. The number of confirmed cases of Lassa fever in Nigeria is 540, 796, 1190, 480, 1015, 1267, 1311 and 1129 for 2018, 2019, 2020, 2021, 2022, 2023, 2024 and 2025, respectively, with a total of 7728. The number of confirmed deaths of Lassa fever in Nigeria is 135, 155, 159, 72, 140, 209, 207, and 211 for 2018, 2019, 2020, 2023, 2024 and 2025, respectively, with a total of 1288. The Case Fatality Rate (CFR) for Lassa fever over the 8-year study period is 16.7%.

Table 2: Model Performance for Lassa fever Confirmed Cases

S/N	Performance Metric	Confirmed Cases	Confirmed Deaths
1	MAE	9.44	20.21
2	RMSE	12.51	24.44
3	R2	0.57	`0.57
4	D	6.3	29.84

Table 2 shows the model performance for Lassa fever confirmed cases and deaths in Nigeria over the study period using the test data set held out for evaluating the model. The performance metrics for the confirmed cases are: 9.44 for MAE, 12.51 for RMSE, 0.57 for R^2 and 6.3 for D while the performance metrics for modeling deaths from Lassa fever are: 20.21 for MAE, 24.44 for RMSE, 0.57 for R^2, and 29.84 for D.

Table 3: 5-Year Time series Forecast for Lassa fever Confirmed Cases and Deaths in Nigeria from 2026 – 2030

S/N	Year	Confirmed Cases with 95% Confidence Interval			Confirmed Deaths with 95% Confidence Interval		
		Confirmed Cases	Lower 95%	Upper 95%	Confirmed Deaths	Lower 95%	Upper 95%
1	2026	1563	135	3001	282	4	557
2	2027	1689	253	3125	312	35	591
3	2028	1810	391	3237	342	61	620
4	2029	1985	535	3441	382	96	670
5	2030	2055	637	3476	402	117	685
Total		9100	1937	16268	1721	313	3125

Table 3 shows the yearly forecast for Lassa fever confirmed cases and deaths in Nigeria for 2026 – 2030. There are: 1563 (95% CI: 135 – 3001), 1689 (95% CI: 253 – 3125), 1810 (95% CI: 391 – 3237), 1985 (95% CI: 535 – 3441), and 2055 (95% CI: 637 – 3476) confirmed cases of Lassa fever forecast for 2026, 2027, 2028, 2029, and 2030 respectively, with an overall 9100 (95% CI: 1937 – 16268) confirmed cases over the forecasted period. For confirmed deaths, there were 282 (95% CI: 4 – 557), 312 (95% CI: 35 – 591), 342 (95% CI: 61 – 620), 382 (95% CI: 96 – 670), and 402 (95% CI: 117 – 685) for 2026, 2027, 2028, 2029, and 2030 respectively, with an overall number of 1721 (95% CI: 313 – 3125) confirmed deaths during the forecast period.

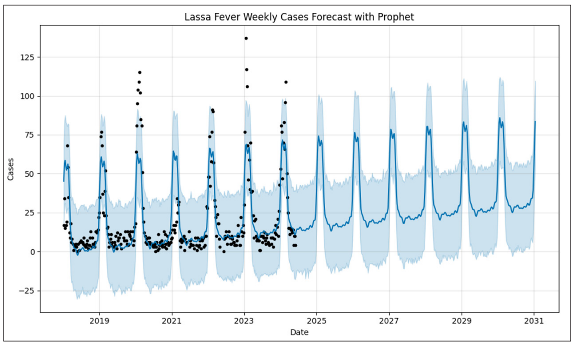
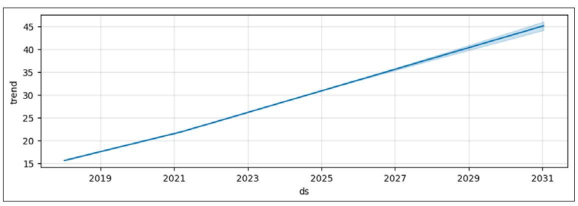


Figure 2: Weekly Forecast of Confirmed Cases of Lassa fever in Nigeria

Figure 2 showed a steady rise in the number of confirmed cases of Lassa fever in Nigeria from 2018 to 2020, then a decline in 2021 and then a rise upto 2024 which was followed by a rise in 2025 with a final steady rise from 2026 to 2030.



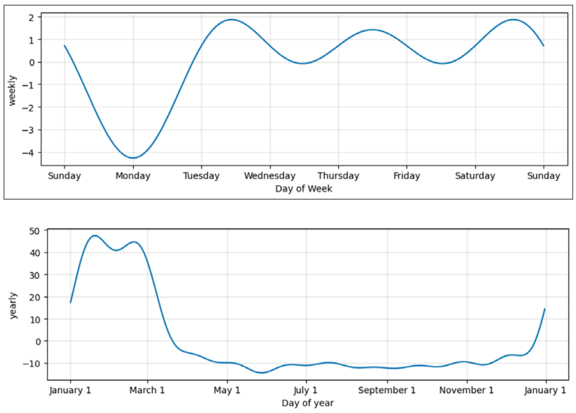


Figure 3: Weekly Forecast of Lassa fever Cases in Nigeria

Figure 3 shows that on the overall, there is a steady rise in the number of confirmed cases of Lassa fever in Nigeria from 2018 to 2030. Most cases occurred in January to March of each year.

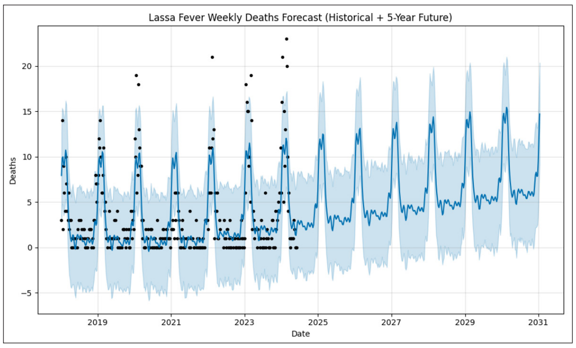


Figure 4: Weekly Forecast of Confirmed Deaths from Lassa fever in Nigeria

Figure 4 showed a steady rise in the number of confirmed deaths from Lassa fever in Nigeria.

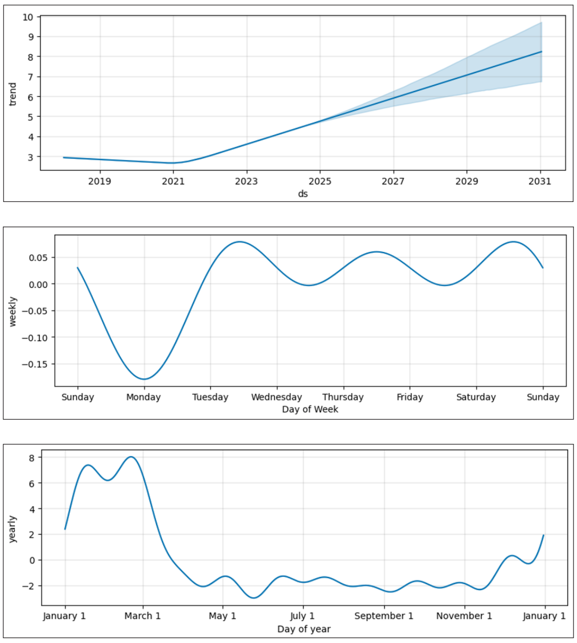


Figure 5: Decomposition of Lassa fever Confirmed Deaths in Nigeria using Facebook Prophet

Figure 5 showed the decomposition of Lassa fever Confirmed Deaths in Nigeria. On the overall, there is increasing trend in the number of confirmed deaths from Lassa fever in Nigeria, with most deaths reported in January–March of each year.

Discussion

Time series Analysis and Case Fatality Rate for Lassa fever in Nigeria

Time series analysis is a specific way of analysing a sequence of data points collected over an interval of time. In time series analysis, the researcher records data points at consistent intervals over a set period of time. Time series analysis shows how variables change over time, and this was applied to this study. For this study, time series weekly data from 2018 to 2025 for Lassa fever outbreaks in Nigeria were collected to identify the trend and make a time series forecast over a period of 5years.

Our result (Figure 1, Table 1) showed the time series analysis of Lassa fever confirmed cases and confirmed deaths in Nigeria over a period of 8-years from 2018 to 2025, which is a large dataset that is reliable and shows consistency in the data points. Our study revealed that 2023 had the highest number of cases and deaths from confirmed cases of Lassa fever, while 2021 had the lowest number of confirmed cases and deaths from Lassa fever. This could be due to the increased susceptibility of vulnerable populations to the virus, which could be from an interplay of environmental factors and the primary rodent infectivity and spread across the various states in Nigeria which are more active in 2023 than 2021 [30].

Overall, a total of 7728 confirmed cases and 1288 confirmed deaths from Lassa fever were recorded over a period of 8years from 2018 to 2025. This translates to a Case Fatality Rate (CFR) of 16.7%. Our findings are similar to studies by Yaro, who noted a steady increase in the incidence of Lassa fever in Nigeria in their study on infection pattern, case fatality rate and spread of the virus across Nigeria, although they reported a higher annual case fatality rate of 18.5% [31]. This decline in CFR is similar to studies by Akpede who observed a decline in the CFR for Lassa fever from an annual CFR rate of 94% in 2001 to 15% in 2018 [32]. This could be due to improve awareness on infection prevention and control and the improvement in care for patients with the disease over the years.

Model Performance

The result from our study (Table 3) showed the performance of the model. While the model performed moderate based on some performance indicators such as R-squared, its performance on other performance metrics such as Poisson Mean Deviance that measures the deviation of a prediction from the mean was not satisfactory. The model demonstrated moderate predictive performance by explaining 57% of the cases and deaths from Lassa fever from the weekly epidemiological data and this categorization corresponds to James who graded models performance using R-squared to be <0.25 as weak, 0.25 – 0.50% as moderate, 0.50 – 0.75 to be moderate to good while 0.75 was considered strong [33]. The mean absolute error (9.44) and root mean square error (12.51) for confirmed cases indicate an acceptable prediction accuracy while the Poisson Mean Deviance of 6 did not show a good fit as a score of zero is considered perfect [34]. On the overall, based on the first three predictors, the Prophet is a moderate model for predicting Lassa fever cases and this is in-line with the WHO assertion that such models are considered useful for epidemiological surveillance and trend monitoring. Despite an R-squared value of 57.7 for number of deaths, on the overall Prophet did not provided great forecast for Lassa fever deaths and this could be attributable to the high missing values for some weeks.

When compared with study on Lassa fever that made use of the Box-Jenkins methods such as SARIMA, the Facebook Prophet performed better based on R-squared where our result was able to explain 57% of the forecast unlike previous results by Tahmo that explained 48.1% of the forecast [34,35]. It also performed better that studies that used ARIMA and ARIMAX in the prediction of COVID-19 across performance metrics such as R-squared, RMSE and MAE [36].

Even though our model outperformed Box-Jenkins method such as ARIMA, and SARIMA, models like the Gaussian-Matern 5/2

Gaussian Process Regression (GPR) did relatively better than the Prophet as noted by Samson who evaluated multiple machine learning models on confirmed Lassa fever cases in Nigeria and reported MAE of approximately 7.72, RMSE of approximately 11.62, and R^2 of approximately 0.85 for confirmed case of Lassa fever in Nigeria [9]. These metrics are better than the results from our study and also outperformed simpler regression models and many other ML models tested in their study. This observation where GPR outperformed ML models could be explained by the fact that the data points from Samson from (2018 – 2020) are less than a hundred (<100), and under such circumstances, ML models struggle to perform well [9,37]. The implication of this finding is that models such as GPR that perform better are recommended for a smaller amount of dataset while ML techniques such as Facebook Prophet can be used for modeling epidemiological data for surveillance purpose when the dataset is much with limited missing data.

Furthermore, when performance metrics for performance are compared between confirmed cases and deaths from Lassa fever are compared, the metrics in confirmed cases outperformed the metrics in confirmed deaths from Lassa fever. This implies that even though the Prophet model captures broad trends and seasonality very well, the week-to-week precision for low-count weeks is limited mostly in the forecasting of deaths from Lassa fever. This finding is similar to result by Samson who noted that the ML models performed better in predicting cases of Lassa fever than deaths [9]. Despite these nuances, the need to use a forecasting model that is relatively easier to use (Facebook Prophet) cannot be overemphasised, as this can improve the use of time series data to forecast and plan out break responses needed to limit the spread and impact of Lassa fever in Nigeria.

Forecast of Confirmed Cases and Deaths using Historical Dataset

One of the powers of machine learning is the capacity to use historical data to predict future outcomes of a disease which can guide outbreak response. Our result (Table 3) showed the yearly forecast for Lassa fever confirmed cases and deaths in Nigeria from 2026 to 2030. This include 1563 (95% CI: 135 – 3001), 1689 (95% CI: 253 – 3125), 1810 (95% CI: 391 – 3237), 1985 (95% CI: 535 – 3441), and 2055 (95% CI: 637 – 3476) confirmed cases of Lassa fever forecast for 2026, 2027, 2028, 2029, and 2030 respectively, with an overall 9100 (95% CI: 1937 – 16268) confirmed cases over the forecasted period. This implies that by the end of 2030 starting from January 2026, there could be upto 10000 cases of Lassa fever in Nigeria and this could translate to higher number of deaths from Lassa fever. For confirmed deaths, there were 282 (95% CI: 4 – 557), 312 (95% CI: 35 – 591), 342 (95% CI: 61 – 620), 382 (95% CI: 96 – 670), and 402 (95% CI: 117 – 685) for 2026, 2027, 2028, 2029, and 2030 respectively, with an overall number of 1721 (95% CI: 313 – 3125) confirmed deaths during the forecast period. This means by the end of 2030, there could be almost 2000 deaths from Lassa fever which translates to approximately 150 cases per year and this is a very big burden for the disease and this call for immediate proactive early warning interventions needed to prevent these deaths completely or limit the death associated with Lassa fever in Nigeria.

The forecast pattern showed a steady rise in the number of confirmed cases and deaths from Lassa fever from 2026 upto 2030 (Table 3 and Figure 3 and 5). This finding is similar to findings from another predictive study by Musa, which utilised the Box-Jenkins ARIMA (0,1,1)(0,1,1)_{12} model to analyse historical Lassa fever outbreaks in Nigeria and project future incidence [38]. That study also reported distinct seasonal and temporal patterns, indicating fluctuating incidence rates with periodic surges. This finding implies that there will be a steady increase in Lassa fever confirmed cases and mortality, and this could present a challenge if timely proactive measures are not taken to curb the spread of the disease across endemic and vulnerable states across Nigeria.

Decomposition of the Time Series

Time series forecasting is a critical tool in data science and disease surveillance, which can be used to improve a country's early warning, alert and response systems by using historical data to forecast and understand seasonal variation of a disease needed to guide public health planning and response [30,39]. It involves the process of using a dependent variable's past values (historical data) to predict its future values (independent variable) [40].

Time series decomposition involves the modelling of the seasonality and trend of a disease using historical data [41]. The results from our study (Figures 3 and 5) clearly showed the decomposition of the time series for Lassa fever in Nigeria into trend and seasonal patterns. There was an increasing trend of Lassa fever outbreaks in Nigeria in terms of confirmed cases and confirmed deaths of the disease, starting from the historical data set in 2018 upto the forecasted dataset that ended in 2030. This could be as a result of the increasing population and increased vulnerability of rural communities to Lassa fever, where most of the disease occurs. Furthermore, the absence of a vaccine needed to prevent this infection, especially among healthcare workers and individuals in endemic locations, could have also contributed to the increasing trend and transmission of the disease across Nigeria. The implication is that more cases and deaths from Lassa fever will be recorded over the years if timely and proactive measures such as health education and awareness creation against the disease, approval of safe and effective vaccines and improvement of sanitary conditions of mostly rural individuals who are exposed to the Lassa fever rodent (Mastomys natalensis) that transmits the infection are not controlled.

Our results also showed that there has been a rise in the number of cases and deaths from Lassa fever, with seasonal variations associated with the incidence of the disease. Specifically, the months of January, February and March has the highest number of cases and deaths from Lassa fever (Figure 3 and 5), and this corresponds to the harmattan seasons, characterized by dry atmospheric conditions with low humidity and sometimes presence of heat waves in some locations that has been associated with increased infectious disease burden in West Africa [42]. A finding that is in line with studies by McKendrick who modeled the seasonality of Lassa fever incidences and vector dynamics in Nigeria and noted that the vector spreads more often during the dry season (January – March) and human contacts with the primary vector (Mastomys natalensis) that is infected with the

Lassa fever virus are higher during this period of the year and calls for a comprehensive approach of controlling the disease incidence by ensuring effective vector control [43].

Implications for Early Warning, Alert, Response and System

The use of a data-driven approach in addressing Lassa fever outbreaks in Nigeria is very important as forecasts clearly give an understanding of the pattern of the disease in terms of long-term variations (trend) and short-term variations (seasons) of infectious diseases, and this can guide operational planning, logistic supply and timely response for an outbreak in Nigeria.

The result from the forecast showed that the models performed moderate in predicting Lassa fever outbreaks in Nigeria, and this makes Facebook Prophet a dependable algorithm for infectious disease prediction that has a trend and a seasonal pattern.

The findings from the dataset and the five-year forecast of confirmed Lassa fever cases and deaths in Nigeria carry important public health, policy, and research implications. The projected burden highlights the persistent endemicity of Lassa fever and underscores the need for proactive, data-driven interventions rather than reactive outbreak responses and this should also include preventing human contacts with the rodent that causes the infection when it contaminants food stuff and other items humans get in contact with [10,44].

The forecasted 9100 confirmed cases and 1721 confirmed deaths between 2025 and 2030 indicate that Lassa fever will remain a significant public health threat in Nigeria over the medium term. The projected steady rise in cases and deaths up to 2030 suggests increasing pressure on healthcare facilities, particularly in endemic states. These findings imply that health authorities must strengthen early preparedness, including surge capacity planning for treatment centres, isolation wards, and diagnostic laboratories, especially during peak transmission seasons.

Following the ease of application of the Facebook Prophet and its ability to provide an understanding of the trend and seasonal variation of the disease, the Prophet can be a tool for increased understanding of disease trend and seasonality for Lassa fever in Nigeria to support early warning, alert and response for effective outbreak response and surveillance system in Nigeria.

Conclusion

Time series forecasting is a critical tool in data science and disease surveillance, which can be used to improve a country's early warning, alert and response systems based on the use of forecasting to understand trend and seasonal patterns of a disease that will guide appropriate proactive interventions, which can save lives.

The result from this study showed that there has been a rise in the number of cases and deaths from Lassa fever, with seasonal variations associated with the occurrence of an outbreak. Specifically, the months of January, February and March have the highest number of cases and deaths from Lassa fever, and this corresponds to the harmathan seasons that are characterised by dry season with low humidity and sometimes presence of heat waves in some locations. This season has been associated with increased infectious disease burden in West Africa and, therefore, calls for proactive measures needed to address it.

The result from the forecast showed that the models performed moderately well in predicting Lassa fever outbreaks (cases) in Nigeria, although not very effective in predicting deaths from the disease. Despite the flaws which could be due to missing data, Facebook Prophet remains a dependable algorithm for infectious disease prediction that has a trend and a seasonal pattern.

The use of a machine learning algorithms for addressing infectious disease (Lassa fever) outbreaks in Nigeria is therefore very important as forecasts clearly give an understanding of the pattern of the disease in terms of long-term variations (trend) and short-term variations (seasons) of infectious diseases. This can guide operational planning, logistic supply and timely response for an outbreak in Nigeria.

References

1. Richmond JK, Baglole DJ. Lassa fever: Epidemiology, clinical features, and social consequences. BMJ. 2003. 327: 1271-1275.
2. Monath TP. Lassa fever: Review of epidemiology and epizootiology. Bulletin of the World Health Organisation. 1975. 52: 577-592.
3. Africa Centre for Disease Prevention and Control. Lassa fever. 2025.
4. World Health Organisation (WHO). Lassa fever. 2024.
5. Ogbu O, Ajuluchukwu E, Uneke CJ. Lassa fever in the West African sub-region: An overview. Journal of Vector Borne Diseases. 2007. 44: 1-11.
6. Nigeria Centre for Disease Control and Prevention (NCDC). 2025.
7. Ogoina D. Lassa fever: A Clinical and Epidemiological Review. Niger Delta Journal of Medicine and Medical Research. 2013. 1: 1-10.
8. World Health Organisation. Lassa fever. 2026.
9. Samson TK, Aromolaran O, Akingbade T. Lassa fever cases and mortality in Nigeria: Quantile Regression vs. Machine Learning Models. Journal of Public Health in Africa. 2024. 14: 2712.
10. Moore KA, Ostrowsky JT, Mehr AJ, Johnson RA, Ulrich AK, et al. Lassa fever research priorities: towards effective medical countermeasures by the end of the decade. The Lancet Infectious Diseases. 2024. 24: 696-706.
11. Nigeria Centre for Disease Control (NCDC). National guideline for Lassa fever case management. 2018.
12. Zhao S, Gao D, Zhuang Z, Chong MKC, Cai Y, et al. Prediction of the global Omicron pandemic using ARIMA, Prophet, and machine learning models. Scientific Reports. 2022.
13. Chugh V. Facebook Prophet: A Modern Approach to Time Series Forecasting. 2025.
14. Pandey G, Chaudhary P, Gupta R, Pal S. SEIR and Regression Model-based COVID-19 outbreak predictions in India. 2020.
15. Adekunle TA, Ogundoyin IK, Akanbi CO. Machine Learning Model for Predicting the Temporal Lassa Fever Confirmed Cases in Nigeria. Jambura Journal of Biomathematics (JJBM). 2025. 6: 166-172.

16. Bakare EA, Mogbojuri OA, Oniyelu DO, Abidemi A, Daniel DO, et al. Time series modelling and forecasting of mpox incidence and mortality in Nigeria. BMC Infectious Diseases. 2025. 25: 794.
17. Fuller B, Ssekitoleko R, Kyozira C, Colston JM, Makumbi I, et al. Comparison of ARIMA and FB-Prophet time series models for the prediction of national and regional malaria incidence in Uganda. Malaria Journal. 2025. 24: 289.
18. Ranganathan P, Aggarwal R. Study designs: Part 1–An overview and classification. Perspectives in clinical research. 2018. 9: 184-186.
19. World Bank Group. Population, total – Nigeria. 2025.
20. Federal Republic of Nigeria, Ministry of Foreign Affairs. Natural Resources. 2025.
21. Obokoh A. Nigeria ranks 11th of 12 countries with the worst healthcare globally. 2025.
22. Crompton J. Data Management from the DCS to the Historian. In Machine learning and Data Science in the oil and gas industry. 2021. 83-110.
23. Wang WK, Chen I, Hershkovich L, Yang J, Shetty A, et al. A systematic review of time series classification techniques used in biomedical applications. Sensors. 2022. 22: 8016.
24. Ort J, Hakvoort K, Neuloh G, Clusmann H, Delev D, et al. Foundations of time series analysis. In Machine Learning in Clinical Neuroscience: Foundations and Applications. Cham: Springer International Publishing. 2021. 215-220.
25. Chen Z, Ma M, Li T, Wang H, Li C. Long sequence time-series forecasting with deep learning: A Survey .Information Fusion. 2023. 97: 101819.
26. Taylor SJ, Letham B. Forecasting at scale. The American Statistician. 2018. 72: 37-45.
27. Battineni G, Chintalapudi N, Amenta F. Forecasting of COVID-19 epidemic size in four high-impact nations (USA, Brazil, India and Russia) by the Fb-Prophet machine learning model. Applied Computing and Informatics. 2025. 21: 2-11.
28. Chafiq T, Ouadoud M, Elboukhari K. Covid-19 forecasting in morocco using fbprophet facebook's framework in python. International Journal of Advanced Trends in Computer Science and Engineering. 2020. 9.
29. Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. Peer J Computer Science. 2021. 7: 623.
30. Montgomery DC, Jennings CL, Kulahci M. Introduction to time series analysis and forecasting. John Wiley & Sons. 2015.
31. Yaro CA, Kogi E, Opara KN, Batiha GES, Baty RS, et al. Infection pattern, case fatality rate and spread of Lassa virus in Nigeria. BMC Infectious Diseases. 2021. 21: 149.
32. Akpede GO, Asogun DA, Okogbenin SA, Dawodu SO, Momoh MO, et al. Caseload and case fatality of Lassa fever in Nigeria, 2001–2018: a specialist centre's experience and its implications. Frontiers in public health. 2019. 7: 170.
33. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning: With applications in R (2nd ed.). Springer. 2021.
34. Fox J. Applied regression analysis and generalized linear models. Sage publications. 2015.
35. Tahmo NB, Wirsiy FS, Brett-Major DM. Modeling the Lassa fever outbreak synchronously occurring with cholera and COVID-19 outbreaks in Nigeria 2021: A threat to Global Health Security. PLOS Global Public Health. 2023. 3: 0001814.
36. Somyanonthanakul W, Chalom K, Chalongrat W, Srisaenpang S, Prasert K. Forecasting COVID-19 cases using ARIMA and ARIMAX models in a field hospital: A case study in Thailand. BMC Medical Research Methodology. 2022. 22: 305.
37. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. 2019. 14: 0224365.
38. Musa A, Osuolale K, Lawal D, Salako A, Aponinuola F, et al. Modelling Seasonal Variation and Lassa Fever Outbreak in Nigeria: A Predictive Approach. International Journal of Data Science and Analysis. 2024. 10: 100-108.
39. Tambo E, Adetunde OT, Olalubi OA. Re-emerging Lassa fever outbreaks in Nigeria: Reinforcing" One Health" community surveillance and emergency response practice. Infectious diseases of poverty. 2018. 7: 74-80.
40. Arslan S. A hybrid forecasting model using LSTM and Prophet for energy consumption with decomposition of time series data. Peer J Computer Science. 2022. 8: 1001.
41. Agarwala S, Sufyan MA, Zedni MAA, Agarwala S, Kar R, et al. Improving Monkeypox Outbreak Prediction Through Time-Series Forecasting with Machine Learning Models. In 2024 27th International Conference on Computer and Information Technology (ICCIT). 2024. 2068-2073.
42. Minka NS, Ayo JO. Influence of cold–dry (harmattan) season on colonic temperature and the development of pulmonary hypertension in broiler chickens, and the modulating effect of ascorbic acid. Open Access Animal Physiology. 2014. 1-11.
43. McKendrick JQ, Tennant WS, Tildesley MJ. Modelling seasonality of Lassa fever incidences and vector dynamics in Nigeria. PLoS Neglected Tropical Diseases. 2023. 17: 0011543.
44. Ojo MM, Goufo EFD. Modeling, analyzing and simulating the dynamics of Lassa fever in Nigeria. Journal of the Egyptian Mathematical Society. 2022. 30: 1.

Copyright: © 2026 Franco Apiyanteide, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.