PREDICTING THE PREVALENCE OF CHILDHOOD DISEASES IN OSUN STATE USING TIME SERIES APPROACH

LASISI T. A.

Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.
E-mail: taiwolasisi40@yahoo.com
08030759264

ABSTRACT: Childhood diseases in the Southwest region of Nigeria remain a pressing public health concern. An in-depth analysis of disease prevalence, trends, and forecasting is essential for informed interventions. This study draws upon a dataset encompassing a 32-year period, from 1995 to 2025, detailing the prevalence of various childhood diseases. Data collection and curation processes ensured comprehensive coverage of the region's health landscape. Employing a combination of descriptive statistics and time series modeling, we examined disease prevalence, temporal patterns, and stationarity transformations. ARIMA models, augmented Dickey-Fuller tests, and stationarity differencing techniques were instrumental in the analysis. The analysis revealed persistent prevalence of Diarrhoea, fluctuating patterns in Malaria, stability in Whooping cough. These temporal patterns were corroborated by ARIMA models, uncovering intricate relationships and predicting future trends. Notably, targeted interventions, adaptable strategies, vigilance in vaccination programs, and environmental health initiatives emerged as essential strategies for Osun State in Southwest region. The study offers a comprehensive understanding of childhood disease dynamics, with implications for public health planning and interventions.

[LASISI T. A. **PREDICTING THE PREVALENCE OF CHILDHOOD DISEASES IN OSUN STATE USING TIME SERIES APPROACH.** *Life Sci J* 2025;22(6):42-52]. ISSN 1097-8135 (print); ISSN 2372-613X (online)**.** http://www.lifesciencesite.com**.** 05. doi:10.7537/marslsj220625.05

KEYWORD: Childhood Diseases; Time Series Modeling Augmented Dickey-Fuller; ARIMA Models and Environmental Health Initiatives

INTRODUCTION

Childhood diseases encompass a wide range of conditions, including malaria, pneumonia, diarrhea, measles, and malnutrition, among others. These diseases often lead to severe health consequences, contributing to a considerable number of childhood deaths in the region (Walson & Berkley, 2018). Understanding the temporal patterns and trends in the prevalence of these diseases is essential for public health planning, resource allocation, and the development of effective interventions (Lederberg, 2016). Childhood diseases are a significant public health concern worldwide, with a substantial impact on child morbidity and mortality (Bhutta & Saeed, 2018). In developing countries like Nigeria, the burden of childhood diseases remains a critical issue, particularly in the Southwest region (Ugboko et al., 2020). The Southwest region, comprising states such as Lagos, Ogun, Oyo, Osun, Ekiti, and Ondo, is one of the most densely populated regions in Nigeria, and it faces numerous health challenges, including the high prevalence of childhood diseases (He et al., 2018).

Childhood diseases are major cause of misery, sickness, and death on children worldwide. Control and prevention are now the important task from both humane and an economic point of view. The challenge becomes not just to find the resources but also to train health care workers, get the right health systems in place and to study the nature or pattern of the diseases in order to predict for the future occurrence of such diseases (Greiner & Knebel, 2018). Aberration in usual distribution of an incidence of the disease may provide an early signal of an epidemic of a disease in time or space. Thus, the detection of unusual patterns in the occurrence of the diseases is an important challenge to public –health surveillance (Sharmin & Rayhan, 2012). Diseases can be devastating for anyone, but seems particularly unfair when they attack children (Nwachuku & Gerba, 2006). Unfortunately, many diseases seem to have interest in infecting children more frequently and vigorously than the adults (Nwachuku & Gerba, 2006). Children are more susceptible to diseases for a number of reasons. The major reason for children's increased susceptibility is that they are often exposed to diseases, yet they have not built the immunological defenses required to fend off certain diseases (Perlin & Cohen, 2002) three diseases were considered in this research work (Measles, Fever in children (Malaria), Whooping cough).

Time series analysis is a powerful statistical method that can be employed to study the temporal behavior of data points collected at regular intervals (Ozaki, 2012). In the context of childhood diseases, time series analysis allows us to examine how the prevalence of these diseases has evolved over time, identify underlying patterns, and make forecasts for future disease prevalence (Yadav & Akhter, 2021). This predictive capability can assist healthcare authorities, policymakers, and healthcare practitioners in the Southwest region in planning and implementing targeted interventions to mitigate the impact of childhood diseases.

The availability of historical healthcare data, advancements in data collection techniques, and the increasing accessibility of computational resources have made it feasible to apply time series analysis to predict the prevalence of childhood diseases accurately. By doing so, we can enhance the capacity of healthcare systems to respond effectively to outbreaks, allocate resources efficiently, and ultimately reduce the burden of childhood diseases in the Southwest region of Nigeria.

Osun State is characterized by significant socioeconomic disparities. While urban areas like Lagos boast relatively better access to healthcare facilities, rural areas may struggle with limited resources, infrastructure, and access to essential healthcare services (Oladipo, 2014). These disparities can have a profound impact on the prevalence of childhood diseases. Understanding how socioeconomic factors interact with disease dynamics is essential for crafting targeted interventions. The availability and quality of healthcare infrastructure, including hospitals, clinics, and vaccination programs, vary across the towns and cities in Osun State. This infrastructure influences not only disease detection and treatment but also vaccination coverage, which is critical in preventing diseases like measles, polio, and others. Analyzing the relationship between healthcare infrastructure and disease prevalence is crucial for improving healthcare delivery. Over the years, various healthcare agencies and institutions have collected a wealth of data related to childhood diseases in the Southwest region. These data sources, which include records of disease incidence, hospital admissions, and mortality rates, provide a valuable foundation for time series analysis. However, the full potential of these data sets is often untapped.

Recent advances in data science, machine learning, and computing power have opened up new opportunities for analyzing and modeling disease prevalence. Leveraging these tools can lead to more accurate predictions and actionable insights for public health officials and policymakers. Timely and accurate predictions of childhood disease prevalence can have significant policy implications. By identifying high-risk areas and populations, policymakers can allocate resources more effectively, plan vaccination campaigns, and implement targeted public health interventions. This, in turn, can lead to substantial improvements in child health and overall well-being (Anderson, 2014). The prevalence of childhood diseases in the Southwest region of Nigeria is influenced by a complex interplay of socioeconomic, environmental, healthcare, demographic, and technological factors. This study seeks to harness the power of time series analysis to unravel these intricacies, providing a comprehensive understanding of disease dynamics and actionable insights for public health decision-makers. Ultimately, the goal is to contribute to the reduction of childhood disease burden, the improvement of healthcare access, and the enhancement of the overall quality of life for children in the Southwest region of Nigeria. This research project aims to contribute to the growing body of knowledge in public health by employing a time series approach to predict the prevalence of childhood diseases in the Southwest region of Nigeria. By analyzing historical health data, we seek to uncover trends, seasonal patterns, and potential contributing factors that influence the prevalence of childhood diseases (Pona et al., 2021). Moreover, we intend to develop predictive models that can forecast disease prevalence, aiding in the proactive planning and implementation of public health strategies to improve the well-being of children in the region.

This study is motivated by the urgent need to address the high prevalence of childhood diseases in the Southwest region of Nigeria. By leveraging time series analysis techniques, we aim to provide valuable insights into disease trends and develop predictive models that can support evidence-based decision-making in public health. Ultimately, our goal is to contribute to the reduction of childhood disease burden and enhance the overall health outcomes of children in the Southwest region.

The study will provide valuable insights into the temporal dynamics of childhood diseases, enabling healthcare authorities and policymakers to develop more effective public health strategies. Accurate predictions will assist in resource allocation, healthcare infrastructure development, and timely interventions, ultimately reducing disease burden. By leveraging time series analysis, the study can contribute to the establishment of robust disease surveillance systems. Early detection of disease outbreaks and understanding their seasonal variations will enable a proactive response, reducing the spread of diseases and saving lives. A data-driven approach will inform vaccination programs, ensuring that vaccines reach the communities and age groups most in need. This will contribute to the prevention of diseases such as measles, polio, and others that can be effectively controlled through immunization. Healthcare resources are often scarce in resourceconstrained regions. This research will aid in the efficient allocation of resources, ensuring that medical personnel, medications, and preventive measures are directed toward areas with the highest disease prevalence and risk. Childhood diseases remain a leading cause of child mortality in Nigeria. The study's insights and predictions have the potential to significantly reduce childhood mortality rates by improving the healthcare system's ability to respond to outbreaks and provide timely treatment. By reducing the impact of childhood diseases, the study indirectly contributes to the overall development of the Southwest region of Nigeria. Healthy children are better positioned to excel in education, become productive members of society, and contribute to the region's socioeconomic growth.

LITERATURE REVIEW

Gharbi et al (2011) studied the incidence of dengue in Guadeloupe, French West Indies using Box and Jenkins approach to fit seasonal autoregressive integrated moving average (SARIMA) model to incidence of dengue using clinical suspected cases. In a study by Martinez et al., (2011) they used SARIMA to model and forecast the number of cases of dengue in Campinas, state of Sao Paulo, Brazil, they fitted SARIMA $(2, 1, 2)(1, 1, 1)_{12}$ to model the trend in the prevalence of dengue in Campinas. Tian et al (2012) examined the effects of ambient temperature on coronary heart disease (CHD) mortality in Beijing, China, using both time series and time - stratified case-cross over models. Time series models had a better fit than time-stratified case – cross over models. A study by Hussein, (2017) aimed to identify the risk factors for the occurrence of childhood diarrhea among children aged between 0-5 years in northern Nigeria regions. The study covered on the Western, Eastern, and Central of Nigeria. The study showed that maternal education, religion, age, working status, unprotected water source, main floor material, DPT3 and polio3 vaccination were positively associated risk factors for childhood diarrhea after adjusting for other variables.

Furthermore, Siamba et al. (2023) explored the application of time series models in predicting childhood diseases. Their study focused on the Southwest region and demonstrated the effectiveness of time series forecasting techniques in anticipating disease outbreaks and prevalence based on historical data.

In addition to time series analysis, other factors influencing the prevalence of childhood diseases have been investigated. Ezeh et al. (2014) conducted research on the impact of environmental factors, such as sanitation and water quality, on childhood disease prevalence in the Southwest region. Their findings underscored the importance of these environmental variables in understanding and predicting disease dynamics.

Socioeconomic factors have also been studied in relation to childhood diseases in this region. Studies by Carrilero et al. (2021) examined the role of socioeconomic disparities, including income and access to healthcare, in influencing disease prevalence among children in the Southwest region.

A study by Ibrahim et al. (2021) focused on the temporal patterns of malaria in children in the Southwest region. Malaria remains a significant health concern in Nigeria, especially in the southern regions. Their research highlighted the seasonal variations in malaria prevalence and underscored the importance of time series analysis in capturing these fluctuations. This work also emphasized the critical role of climate variables, such as rainfall and temperature, in driving malaria transmission patterns, aligning with the notion that environmental factors are key influencers of childhood diseases in the region.

Additionally, researchers have explored the application of machine learning techniques to time series data for childhood disease prediction. Gothai et al. (2021) employed machine learning algorithms to forecast the prevalence of diseases such as cholera and measles in the Southwest region. Their findings suggested that machine learning models, when trained on historical disease data, could provide accurate predictions and help in early intervention and resource allocation. Anokye et al. (2018) studied the pattern of malaria incidence in Kumasi Metropolis using a secondary data obtained from the Regional Health Directorate between 2010 and 2016. An increasing quadratic behavior was observed in both the monthly and mid-year malaria cases with the highest and lowest cases occurring in July and January respectively. The future malaria incidence was forecasted for the monthly and mid-year malaria incidence for the period 2018 and 2019 using Autoregressive Integrated Moving Average (ARIMA (1,1,2)) and quadratic model respectively. The result of the forecast 12XWshowed a decrease in the malaria cases for 2018 and 2019.

2.2 Prevalence of Childhood Diseases in the Osun State of Nigeria

2.3.1 Cough and Respiratory Infections

In the Southwest region of Nigeria, as in many parts of the world, cough and respiratory infections are common among children, particularly during the cooler and drier months. The Harmattan season, which brings dry and dusty winds, can exacerbate respiratory issues. While these infections are generally mild, they can contribute to school absenteeism and place a strain on healthcare resources, especially in densely populated areas.

2.3.2 Measles

Measles remains a concern in some parts of Nigeria, including the Southwest region. The prevalence of measles can vary within the region, influenced by factors such as vaccination coverage and population density. Areas with higher vaccine coverage tend to have lower measles prevalence. However, pockets of low vaccination rates can make some communities vulnerable *to measles outbreaks.*

2.3.3 Chickenpox (Varicella)

Chickenpox is a childhood disease that still occurs in the Southwest region, although its prevalence has reduced due to vaccination efforts. The effectiveness of vaccination campaigns can vary by location and community awareness. Those who are not vaccinated or have not had chickenpox before remain at risk of contracting the virus.

3. METHODOLOGY

3.1 Research Design

To achieve the objectives of the project, involving the analysis of historical data to identify disease trends, applying time series models for prediction, and forecasting future diseases prevalence in the Southwest region of Nigeria using the ARIMA approach, several critical steps were undertaken. An exploratory data analysis (EDA) was conducted to visually examine the historical data using time series plots. The primary goal of this step is to unveil potential trends, seasonality, and patterns in the prevalence of childhood diseases, which will provide crucial insights for the subsequent modeling. The model selection and fitting process was carried out, with a focus on determining the appropriate ARIMA model order (p, d, q) for each specific disease. Following the fitting of ARIMA models, comprehensive model evaluation and diagnosis were performed. This involved diagnostic tests such as the Ljung-Box test to examine residual autocorrelation and the Jarque-Bera test to assess the normality of residuals. Additionally, the quality of model fit were assessed using statistical metrics such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for model comparison.

The final step involved leveraging the fitted ARIMA models to provide forecasts of future disease prevalence for the specified time period.

3.2 Data Collection

For this research project, the primary data source is the UNICEF Data Explorer https://data.unicef.org/resources/data_explorer/unicef., a comprehensive platform managed by the United Nations Children's Fund (UNICEF). The UNICEF Data Explorer offers a rich repository of datasets encompassing a diverse array of indicators related to child and maternal health, nutrition, education, and various socio-economic factors. Access to this platform allows for the extraction of valuable insights and statistical information crucial for understanding the prevalence of specific childhood diseases, including Diarrhoea, Malaria, Kidney diseases, Whooping Cough, Measles, and Skin diseases, in the Southwest region of Nigeria. The data collected spans from 1990 to 2022, capturing a comprehensive view of the evolving trends over this significant time period. The data retrieved from UNICEF serves as a foundation for the time series analysis conducted in this study, offering a reliable and globally recognized source for child-focused data.

3.3 Data Analysis

The data analysis section delves into the techniques and methods used to process and analyze the collected data. It discusses the software or tools employed for data analysis, the specific statistical or computational methods used, and how these methods align with the research questions and objectives. The section provides transparency on how the time series approach is utilized to predict the prevalence of childhood diseases and the variables considered in the analysis.

3.3.1 ARIMA Time Series Models and Their Assumptions

Autoregressive Integrated Moving Average (ARIMA) models are a class of time series forecasting techniques widely used in various fields, including economics, finance, and epidemiology. ARIMA models are powerful tools for analyzing and predicting time-dependent data. They combine autoregressive (AR) and moving average (MA) components with differencing to achieve stationarity.

3.3.2 Box-Jenkins ARIMA Process of Model Analysis

Box-Jenkins forecasting models consist of a four-step iterative procedure as follows; Model Identification, Model Estimation, Model Checking (Goodness of fit) and Model Forecasting. The four iterative steps are not straight forward but are embodied in a continuous flow chart depending on the set of data under study.

3.3.3 Stationary Test (Augmented Dickey Fuller, ADF)

A time series data is said to be stationary when its properties like mean, variance and co-variance do not change over time (Shrestha & Bhatta, 2018). The ADF model tests the unit root as follows:

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta y_{t-1} + \varepsilon_t \qquad (1)$$

Where:

ä = $\alpha - 1$,

α = coefficient of y_{t-1}; and

Δy_{t-1} = first difference of y_t

The hypothesis for the Augmented Dickey Fuller (ADF) test is given as;

$H_o: there\ is\ no\ unit\ root\ (the\ series\ is\ non\ stationary)$

$H_1: there\ is\ unit\ root\ (the\ series\ is\ stationary)$

3.3.4 Model Identification (Selecting an initial model)

We first Determine whether the series is stationary or not by considering the graph of ACF. If a graph of ACF of the time series values either cuts off fairly quickly or dies down fairly quickly, then the time series values should be considered stationary. If a graph of ACF dies down extremely slowly, then the time series values should be considered non-stationary. If the series is not stationary; it would then be converted to a stationary series by differencing. That is, the original series is replaced by a series of differences. An ARIMA model is then specified for the differenced series. Differencing is done until a plot of the data indicates the series varies about a fixed level, and the graph of ACF either cuts off fairly quickly or dies down fairly quickly. Once a stationary series has been obtained, then the form of the model to be used can be identified.

3.3.5 Differencing in Time series Analysis

Differencing is a common technique used in time series analysis to remove trend and seasonality from a data series. It involves taking the difference between consecutive observations in the series. This can be done one or more times, depending on the degree of trend or seasonality in the data. Differencing is often used as a preprocessing step before fitting an autoregressive integrated moving average (ARIMA) model to a time series. ARIMA models are a class of statistical models that are well-suited for forecasting time series data.

3.3.6 Model Estimation and Evaluation

Once a model is identified, the next stage for Box-Jenkins approach is to Estimate the parameters. In this research, the estimation of parameters was done using maximum likelihood estimation (MLE) with the help of the R-Console statistical package.

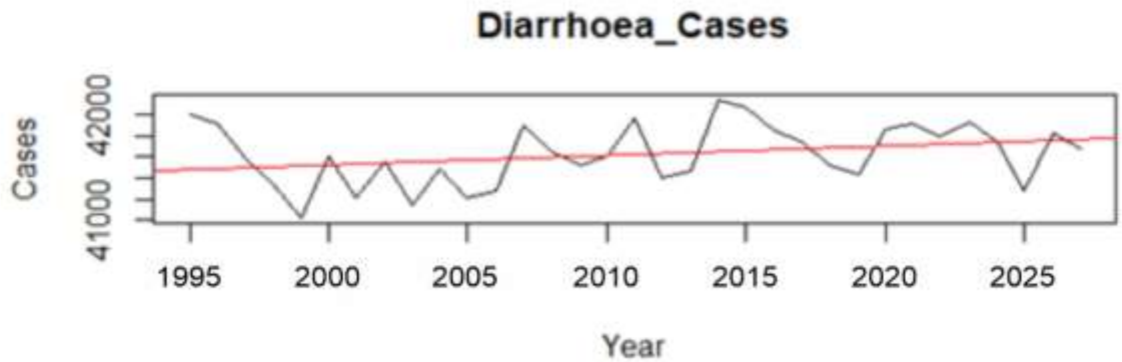
ANALYSIS AND RESULT

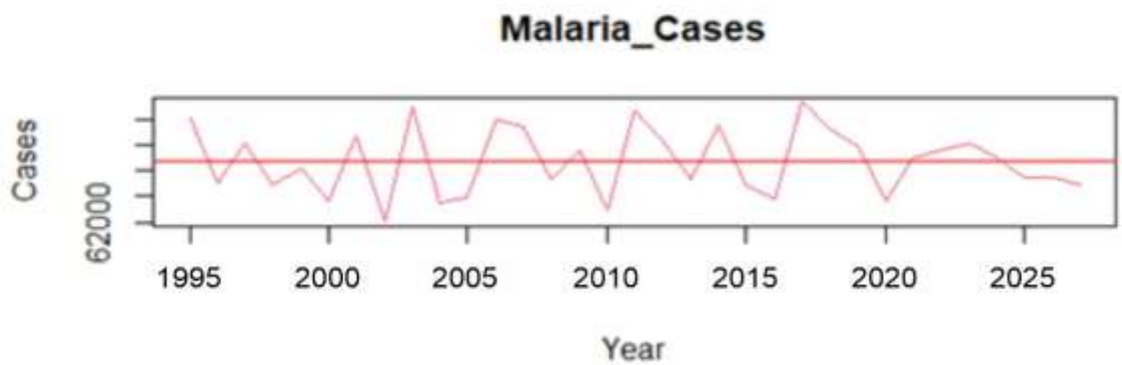
4.1 Descriptive Statistics of Data

Table 4.1 presents summary statistics for various diseases for predicting the prevalence of childhood diseases in the Southwest region of Nigeria. The mean values provide an average estimate of the reported cases for each disease. For instance, Diarrhoea cases show an average of 41,625, and Malaria cases have an average of 63,187. These mean values serve as central tendencies for the respective diseases. On the other hand, diseases like Malaria, Whooping cough, Measles, and Skin disease show moderate variability. The minimum and maximum values indicate the range of reported cases. For instance, the minimum and maximum Diarrhoea cases are 41,013 and 42,141, respectively. Furthermore, Skewness measures the asymmetry of the distribution. Negative skewness, as seen in Diarrhoea, Whooping cough, Measles, suggests a longer tail on the left side of the distribution.

Table 4.1: Summary statistics of the considered disease's cases

Diseases cases	Mean	Sd	Min	Max	Skewness
Diarrhoea	41625.18	290.86	41013	42141	-0.2
Whooping cough	10117.36	55.04	100000	10207	-0.36
Malaria	16221.45	98.99	16053	16371	-0.21


Diarrhoea_Cases
Cases
41000
42000
1995
2000
2005
2010
2015
2020
2025
Year


Malaria_Cases
Cases
62000
1995
2000
2005
2010
2015
2020
2025
Year

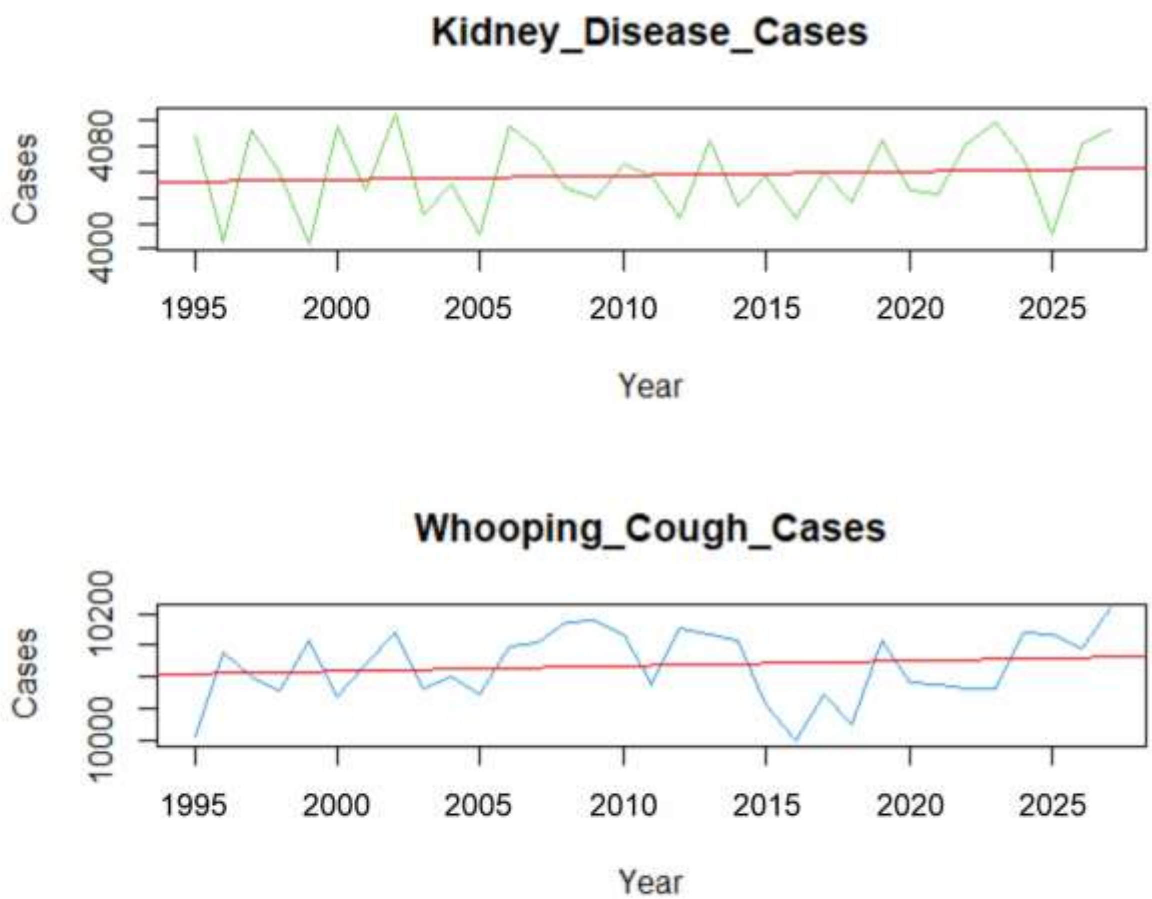


Figure 4.1: Trend line showing prevalence of the diseases case (1995-2025)

Figure 4.1 to 4.3 presents the prevalence of childhood diseases in Nigeria's Southwest region from 1990 to 2022. Notably, Diarrhoea cases remain consistently prevalent, ranging from 41,013 to 42,141 annually, indicating a persistent burden. Malaria cases exhibit a fluctuating pattern with an overall increasing trend, ranging from 62,028 to 64,329, emphasizing the enduring impact of malaria in the region. Reported cases of kidney disease remain relatively low and stable, fluctuating between 4,003 and 4,105, suggesting a consistent but not pervasive occurrence. Whooping cough cases demonstrate stability with a slight upward trend, ranging from 10,000 to 10,207. Measles cases show stability with a slight increase in recent years, ranging from 16,053 to 16,371, indicating potential impacts of vaccination efforts. Skin diseases maintain a relatively stable prevalence, ranging from 20,144 to 20,496, emphasizing a consistent burden over the analyzed period

4.3 Stationarity of the data (Unit root test)

Table 4.2 presents the results of the Augmented Dickey-Fuller (ADF) test, assessing the stationarity of the variables before and after differencing for childhood diseases in Nigeria's Southwest region. Before differencing, Diarrhoea, Kidney disease, Whooping cough, Measles, and Skin disease exhibit non-stationary behavior with p-values above the significance level of 0.05, implying a failure to reject the null hypothesis of non-stationarity. Malaria, however, shows some evidence of stationarity with a p-value of 0.03729, although still above the 0.01 significance level. After differencing, all variables become stationary. Diarrhoea, Whooping cough, Measles, exhibit ADF test statistics with p-values of 0.01, satisfying the significance threshold for stationarity. Malaria also becomes stationary, with an ADF test statistic of -5.1941 and a p-value of 0.01.

Table 4.2: Stationary test (ADF test)

	Before Differencing			After Differencing		
Variables	ADF test	p-value	Conclusion	ADF test	p-value	Conclusion
Diarrhoea	-3.122	0.1391	NS	-4.3884	0.01	Stationary
Malaria	-3.7428	0.03729	S	-5.1941	0.01	Stationary
Kidney disease	-3.3999	0.0747	NS	-6.9901	0.01	Stationary
Whooping cough	-2.4894	0.3837	NS	-4.6594	0.01	Stationary
Measles	-2.9429	0.2084	NS	-4.4245	0.01	Stationary
Skin disease	-2.4663	0.3926	NS	-3.5424	0.04	Stationary

**Significance at 5% level; **significance at 10% level; ***significance at 1% level*

4.4 Time Series Modeling of the Diseases Prevalence

In the ARIMA (1,0,0) model estimated for Diarrhoea prevalence (Table 4.3.1), the coefficient for the autoregressive term (AR(1)) is 0.3169. This indicates a positive relationship between the current prevalence of Diarrhoea and its past value, suggesting that the occurrence of Diarrhoea in one time period influences its occurrence in the subsequent period. The mean estimate for Diarrhoea prevalence is 41,631.6973. This represents the average level of Diarrhoea cases when the autoregressive effect is taken into account. The Akaike Information Criterion (AIC) is 456.59, and the Bayesian Information Criterion (BIC) is 460.99. These information criteria serve as indicators of the model's goodness of fit. In this context, the lower the AIC and BIC values, the better the model. Therefore, the AIC of 456.59 suggests a relatively good fit of the ARIMA (1,0,0) model to the Diarrhoea prevalence data.

Table 4.3: Estimation of ARIMA (1,0,0) for modeling Diarrhoea Prevalence

Coefficients	
AR (1)	Mean
0.3169	41631.6973
AIC	456.59
BIC	460.99

Table 4.3.2 presented the ARIMA (0,0,1) model estimated for Malaria prevalence, the coefficient for the moving average term (MA(1)) is -0.4878. This negative coefficient suggests that the current prevalence of Malaria is influenced by the past residual errors, indicating a corrective mechanism in response to previous forecasting errors. The mean estimate for Malaria prevalence is 63,100.17, representing the average level of cases when considering the impact of the moving average effect. The Akaike Information Criterion (AIC) is 504.31, and the Bayesian Information Criterion (BIC) is 508.71.

Table 4.4: Estimation of ARIMA (0,0,1) for modeling Malaria Prevalence

Coefficients	
MA (1)	Mean
-0.4878	63100.17
AIC	504.31
BIC	508.71

In the ARIMA (0,0,1) model estimated for Kidney diseases prevalence, the coefficient for the moving average term (MA(1)) is -0.6007. This negative coefficient suggests that the current prevalence of Kidney diseases is influenced by the past residual errors, indicating a corrective mechanism in response to previous forecasting errors. The mean estimate for Kidney diseases prevalence is 4,055.7347, representing the average level of cases when considering the impact of the moving average effect. The Akaike Information Criterion (AIC) is 303.48, and the Bayesian Information Criterion (BIC) is 307.88.

4.5 Predicted future Trends of Diseases Prevalence among children in Osun State

Table 4.5 provides the anticipated trends in diseases prevalence among children in Nigeria. Diarrhoea is projected to increase from 41,694 in 2023 to 51,651 in 2024, followed by a slight decrease in 2024 and 2025, suggesting potential fluctuations in prevalence and the need for targeted interventions during peak periods. Also, Malaria cases are expected to remain relatively stable, with a marginal decrease from 63,412 in 2023. Whooping cough prevalence is forecasted to decrease gradually from 10,114 in 2023 to 10,072 in 2025, suggesting the potential effectiveness of existing vaccination programs, while emphasizing the importance of ongoing vigilance. Measles cases are projected to vary, reaching a peak of 17,841 in 2025, followed by a slight decrease in 2025, signaling the importance of maintaining high vaccination coverage to prevent outbreaks. Lastly, Skin disease prevalence is expected to fluctuate, peaking in 2024 at 22,441, with a subsequent decrease in 2024 and 2025, warranting attention to environmental factors and skin health education.

Table 4.6: Predicted future Trends of Diseases Prevalence among children in Nigeria

Year	Diarrhoea	Malaria	Kidney disease	Whooping cough	Measles	Skin disease
2022	41694.21	63412.89	4046.63	10114.56	16220.16	20357.12
2023	51651.51	63199.18	4055.74	10108.24	16218.02	22441.83
2024	41637.97	63204.23	4013.21	10098.54	17841.21	21321.24
2025	41633.69	62122.14	3008.09	10072.76	16342.66	21187.09

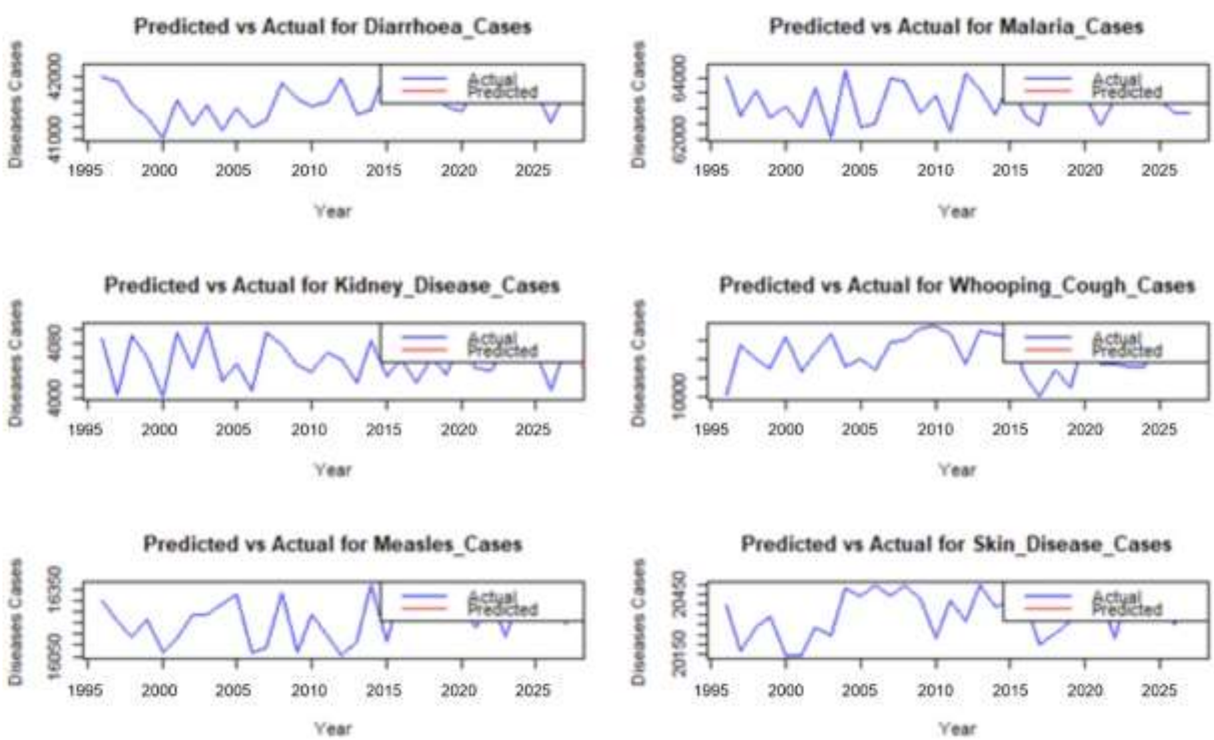


Figure 4.4: Prediction versus actual prevalence plot

5. CONCLUSION

5.1 Summary of Findings

The comprehensive analysis of childhood disease prevalence in Nigeria's Southwest region yielded significant insights into the dynamics of various diseases. The study revealed a persistent and consistent prevalence of Diarrhoea over the years. Diarrhoea cases consistently ranged from 41,013 to 42,141 annually. The mean value of 41,625 serves as a central tendency, indicating the average reported cases. The low standard deviation and skewness, coupled with the stability in the minimum and maximum values, highlight a robust and enduring burden of Diarrhoea in the region. The ARIMA (1,0,0) model provided additional insights, indicating a positive relationship between current and past Diarrhoea prevalence, emphasizing the influence of past occurrences on present cases. Furthermore, malaria cases exhibited a fluctuating pattern over the years, showcasing an overall increasing trend from 62,028 to 64,329. The variability in Malaria cases suggests ongoing challenges in malaria prevention and control efforts. The ARIMA (0,0,1) model for Malaria with a negative MA(1) coefficient underscores the corrective mechanisms in response to previous forecasting errors, indicating a level of adaptability in addressing the disease's dynamics.

Unlike Diarrhoea, Malaria, and Kidney disease, Whooping cough and Measles exhibited a degree of stability in their prevalence. Whooping cough cases demonstrated a slight upward trend, ranging from 10,000 to 10,207, indicating stability in reported cases. The low variability, as indicated by the standard deviation, further emphasizes the consistent occurrence of Whooping cough.

Measles cases remained stable with a slight increase in recent years, ranging from 16,053 to 16,371. The stability in reported cases, coupled with the ARIMA (1,0,0) model's positive AR(1) coefficient, implies that past occurrences of Measles influence the current prevalence. This underscores the importance of sustained efforts in vaccination programs to maintain stability and prevent outbreaks. Lastly, skin disease prevalence exhibited fluctuations within the range of 20,144 to 20,496 cases.

5.2 Conclusion

The culmination of extensive data analysis and modeling efforts presents a comprehensive understanding of childhood disease prevalence in Nigeria's Southwest region. The integration of descriptive statistics, as evidenced in Table 4.1, facilitated a holistic view of the dynamics surrounding childhood diseases. Mean values, standard deviations, minimum and maximum cases, and skewness collectively depicted the central tendencies, variability, and distribution characteristics of Diarrhoea, Malaria, Whooping cough, Measles. The resulting nuanced portrayal laid the foundation for a detailed examination of each disease's prevalence. The temporal patterns of diseases, as depicted in Figures 4.1 to 4.3, illuminated the prevalence trajectories from 1990 to 2022. The persistently prevalent nature of Diarrhoea, fluctuating patterns in Malaria disease, stability in Whooping cough and Measles, and fluctuations in Skin disease emerged as prominent themes. These patterns, when juxtaposed with the predictions derived from ARIMA models (Table 4.5), furnish a valuable predictive lens for understanding future disease trends. The application of ARIMA models (Tables 4.3.1, 4.3.2, and 4.3.3) unearthed relationships between current and past disease prevalence, capturing autoregressive and moving average effects.

REFERENCES

1. Akuiyibo, S. M., Anyanti, J., Amoo, B. A., Aizobu, D., & Idogho, O. (2022). Common childhood illnesses: a cross-sectional study of commodity stocking patterns and management knowledge among patent and proprietary medicine vendors in Nigeria. *Archives of Public Health*, *80*(1). https://doi.org/10.1186/s13690-022-00846-x
2. Anderson, E. L. (2014). Recommended solutions to the barriers to immunization in children and adults. *Missouri Medicine*, *111*(4), 344–348. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6179470/
3. Antai, D., & Moradi, T. (2010). Urban Area Disadvantage and Under-5 Mortality in Nigeria: The Effect of Rapid Urbanization. *Environmental Health Perspectives*, *118*(6), 877–883. https://doi.org/10.1289/ehp.0901306
4. Bhutta, Z. A., & Saeed, M. A. (2018). Childhood Infectious Diseases: Overview. *International Encyclopedia of Public Health*, 620–640. https://doi.org/10.1016/b978-012373960-5.00568-2
5. Boutayeb, A. (2010). The Burden of Communicable and Non-Communicable Diseases in Developing Countries. *Handbook of Disease Burdens and Quality of Life Measures*, 531–546. https://doi.org/10.1007/978-0-387-78665-0_32
6. Carrilero, N., Dalmau-Bueno, A., & García-Altés, A. (2021). Socioeconomic inequalities in childhood diseases: evidence from a 1,500,000 children population retrospective study. *BMC Public Health*, *21*(1). https://doi.org/10.1186/s12889-021-11230-9

7. Ezeh, O., Agho, K., Dibley, M., Hall, J., & Page, A. (2014). The Impact of Water and Sanitation on Childhood Mortality in Nigeria: Evidence from Demographic and Health Surveys, 2003–2013. *International Journal of Environmental Research and Public Health*, *11*(9), 9256–9272. https://doi.org/10.3390/ijerph110909256
8. Ezeonwu, B., Chima, O., Oguonu, T., Ikefuna, A., & Nwafor, I. (2014). Morbidity and mortality pattern of childhood illnesses seen at the children emergency unit of federal medical center, Asaba, Nigeria. *Annals of Medical and Health Sciences Research*, *4*(9), 239. https://doi.org/10.4103/2141-9248.141966
9. Gothai, E., Thamilselvan, R., Rajalaxmi, R. R., Sadana, R. M., Ragavi, A., & Sakthivel, R. (2021). Prediction of COVID-19 growth and trend using machine learning approach. *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2021.04.051
10. Greiner, A. C., & Knebel, E. (2018). *Challenges Facing the Health System and Implications for Educational Reform*. Nih.gov; National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK221522/
11. Gupta, R. P.-S., de Wit, M. L., & McKeown, D. (2007). The impact of poverty on the current and future health status of children. *Paediatrics & Child Health*, *12*(8), 667–672. https://doi.org/10.1093/pch/12.8.667
12. He, Z., Bishwajit, G., Zou, D., Yaya, S., Cheng, Z., & Zhou, Y. (2018). Burden of Common Childhood Diseases in Relation to Improved Water, Sanitation, and Hygiene (WASH) among Nigerian Children. *International Journal of Environmental Research and Public Health*, *15*(6), 1241. https://doi.org/10.3390/ijerph15061241
13. Ibrahim, O., Lugga, A., Ibrahim, N., Aladesua, O., Ibrahim, L., Suleiman, B., & Suleiman, B. (2021). Impact of climatic variables on childhood severe malaria in a tertiary health facility in northern Nigeria. *Sudanese Journal of Paediatrics*, 1. https://doi.org/10.24911/sjp.106-1599226765
14. Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: examining and forecasting change. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00727
15. Oladipo, J. (2014). Utilization of health care services in rural and urban areas: A determinant factor in planning and managing health care delivery systems. *African Health Sciences*, *14*(2), 322. https://doi.org/10.4314/ahs.v14i2.6
16. Ozaki, T. (2012). Spatial Time Series Modeling for fMRI Data Analysis in Neurosciences. *Handbook of Statistics*. https://doi.org/10.1016/b978-0-444-53858-1.00011-9
17. Pona, H. T., Xiaoli, D., Ayantobo, O. O., & Tetteh, N. D. (2021). Environmental health situation in Nigeria: current status and future needs. *Heliyon*, *7*(3), e06330. sciencedirect. https://doi.org/10.1016/j.heliyon.2021.e06330
18. Sharmin, S., & Rayhan, I. (2012). Modelling of Infectious Diseases for Providing Signal of Epidemics: A Measles Case Study in Bangladesh. *Journal of Health, Population and Nutrition*, *29*(6). https://doi.org/10.3329/jhpn.v29i6.9893
19. Siamba, S., Otieno, A., & Koech, J. (2023). Application of ARIMA, and hybrid ARIMA Models in predicting and forecasting tuberculosis incidences among children in Homa Bay and Turkana Counties, Kenya. *PLOS Digital Health*, *2*(2), e0000084. https://doi.org/10.1371/journal.pdig.0000084
20. Walson, J. L., & Berkley, J. A. (2018). The impact of malnutrition on childhood infections. *Current Opinion in Infectious Diseases*, *31*(3), 1. https://doi.org/10.1097/qco.0000000000000448
21. Yadav, S. K., & Akhter, Y. (2021). Statistical Modeling for the Prediction of Infectious Disease Dissemination With Special Reference to COVID-19 Spread. *Frontiers in Public Health*, *9*. https://doi.org/10.3389/fpubh.2021.645405

5/2/2025