Asian Journal of
Science, Technology, Engineering, and Art
e-ISSN : 3025-4507
p-ISSN : 3025-5287
Index : Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc
https://doi.org/10.58578/AJSTEA.v2i6.4126

Predictive Autoregressive Integrated Moving Average Model; Meningitis Death and Alzheimer Death

Adeniyi Oyewole Ogunmola[1], James Oruarooghene Onowuzou[2], Danjuma Idi[3]
Federal University Wukari, Taraba State, Nigeria
ogunmola@fuwukari.edu.ng; oruarooghenejames@gmail.com

Article Info:

Submitted:	Revised:	Accepted:	Published:
Oct 19, 2024	Nov 2, 2024	Nov 15, 2024	Nov 20, 2024

Abstract

Nigeria, being a developing country and by its limited resources, priority must be placed on deaths due to ill-health-related research. This study aims to investigate the temporal pattern and forecast mortality associated with Alzheimer's disease and meningitis using ARIMA model techniques. Results showed that the best fitted for the Alzheimer series is the ARIMA (1, 2, 0) model and the best for the meningitis series is the ARIMA(1,1,0) model. The forecasted values revealed that there will be an inconsistent slight decrease in meningitis deaths while there will be an increase in Alzheimer's deaths over the years. For meningitis, the predicted deaths for 2020, 2021, and 2022 respectively, were 45401, 45216, and 45287, and their 95% confidence intervals were obtained. For Alzheimer's, predicted deaths for 2020, 2021, and 2022 respectively, are 12472, 12832, and 13191, and their corresponding 95% confidence intervals were also obtained. Diagnostic checks for the predictive models were carried out and assumptions were sastisfied.

Keywords: Alzheimer, Meningitis, ARIMA, Nigeria, Forecast, Predicted

INTRODUCTION

Meningitis, a serious infection affecting the brain and spinal cord's protective membranes (meninges), is caused by both protozoa and bacteria [1]. Viral and bacterial infections are the most common triggers, with some strains being contagious and spreading through coughing, sneezing, or close contact. Symptoms vary with children often experiencing decreased appetite, irritability, and gastrointestinal issues, while adults may suffer from headaches, fever, and stiff neck. Bacterial meningitis symptoms include altered mental status, nausea, chills, and skin discoloration resembling bruises. In Africa, meningitis remains a significant public health concern, with geographical variations in its impact [2]. Neisseria meningitidis (meningococcus), Streptococcus pneumoniae (pneumococcus), and Haemophilus influenzae are the main culprits. These bacteria are usually carried asymptomatically in the nasopharynx and spread by respiratory droplets. [3]

Alzheimer's disease is a neurological disorder that progressively impairs memory, cognitive function, and ultimately the ability to carry out routine tasks. Dementia encompasses a broader loss of cognitive and behavioral functions, significantly impeding daily life. According to [4] Dr. Alois Alzheimer first described Alzheimer's disease in the early 20th century. Alzheimer's disease has a significant negative impact on the quality of life for both the affected person and their caretakers. Although it shortens life expectancy, the World Alzheimer Report from 2015 emphasizes that the difficulties it presents on a daily basis are what have the most impact.

Related Work

According to research conducted by [1]. in their study titled "Four Decades Temporal Trend in Meningitis Cases in Africa (1971-2010): A Tale of Three Countries," the objective was to examine the temporal distribution of meningitis cases within three African nations located in the meningitis belt: the Democratic Republic of Congo (DR Congo), Ghana, and Nigeria. The Statistical Information System of the World Health Organization provided the majority of the data used in the study. Plotting annual meningitis cases against the preceding year allowed for the visualization of temporal patterns, and Poisson regression models were used to assess trends throughout the 40-year span from 1971 to 2010. Nigeria ranked highest in the absolute number of reported cases, followed by DR Congo and Ghana. While DR Congo initially experienced a low growth rate in cases, there was a significant surge from 1996 to 2010. Ghana's case growth rate remained relatively stable, with peaks observed in 1997 and intermittent spikes between 1981-1986 and 1996-2001. Nigeria's temporal variability showed no clear pattern, with peaks recorded in 1978 and 2009. The study revealed an increase in reported cases of meningitis in Nigeria and DR Congo over the four-decade surveillance period, emphasizing the critical need for effective preventive and control measures.

Based on Intervention Model Based on Exponential Smoothing Methods and ARIMA Modelling of the Nigerian Naira Exchange Rates carried out by Elisha J. Inyang 2024, the study aimed at use intervention modeling to simulate exchange rates between the Pakistani rupee and the Nigerian naira during a financial downturn. method built around the ARIMA model and exponential smoothing techniques in order to facilitate comparisons. The daily exchange rates between Pakistani rupees and Nigerian naira from January to December of 2016 comprised the dataset used in this study. Findings showed that the naira was negatively impacted by the 2016 economic slump, with a drop in percentage of 47.51. This suggests that, in comparison to the times before and after the intervention, the rupee rose against the naira to the point where Rs1 PKR = 1.9524 NGN. The economic recession was regarded as a steady, permanent effect with a decay rate of 0.60 and a step function with a delay of one period. However, in comparison, the ARIMA-intervention modelled better than the ETS-intervention model.

In a study conducted by [(6)]. in 2017, focusing on the cerebrospinal meningitis outbreak in Kebbi state, Nigeria, the aim was to characterize the outbreak's pattern in 2015. Data on cases and fatalities were collected throughout the outbreak period across all affected local governments in Kebbi state. Over the 18-week duration of the outbreak, 1,992 suspected cases of cerebrospinal meningitis were reported, with 57.0% being males and 43.0% females, resulting in a case fatality rate of 4.0%. The highest proportion of cases occurred in individuals aged 15 and above (31.0%), with 62.9% of cases having received immunization against Neissseria meningitidis type A. Sixteen out of the state's LGAs were affected, with Aliero LGA reporting approximately half (55.5%) of the cases. The majority (77.3%) of analyzed samples tested positive for Nm type C. The study concluded that an effective surveillance system and mass vaccination using polyvalent vaccines containing serogroup C could prevent future outbreaks.

Using a meta-regression epidemiology model, [(7)]. (2019) conducted a study on the prevalence of dementia in Nigeria. Using United Nations demographic data for Nigeria, this model calculated the total number of dementia patients in the nation between 1995 and 2015. Four of the five studies, all of which were deemed to be of high quality, were carried out in the South-West region. In Nigeria, the combined crude prevalence of dementia was determined to be 4.9% (95% CI: 3.0–6.9). Women had a considerably greater prevalence of dementia (6.7%, 3.6–9.9) than men (3.1%, 1.2–5.0). In Nigeria, significant risk factors for dementia have been discovered, including age 80 years or older, female sex, and BMI ≤18.5. The number of dementia cases among people aged 60 and older rose by almost 400% over a 20-year period, from 63,512 in 1995 to 318,011 in 2015, according to estimates made using the epidemiologic model. These results point to a significant rise in dementia cases and prevalence in Nigeria during the previous 20 years.

Adeleye et al., (2023) conducted research on "Time Series Analysis of Crude Oil Production in Nigeria between the Years 2010 to 2020". The Nigerian National Petroleum Company Limited (NNPC Limited) website provided the data used for this study, which included the total number of barrels of crude oil produced monthly from 2010 to 2020. Python was the computer language used for the analysis for this study. The information gathered indicates that the volume of crude oil produced in a given year varies significantly from the year before. For instance, comparing 2019 and 2020, it was seen that, on average, 61,270,340 barrels of crude oil were produced in 2019 and 53,696,864 barrels in 2020. The data indicates that there was a 12.3% drop in crude oil output between 2019 and 2020. They observed that there was a difference in the production as well as a loss in the production of crude oil between the years 2019 and 2020. the model SARIMA (0, 1, 1) x (1, 1, 1, 11) was found to be the best fit for the data with the lowest AIC. It was suggested that the Nigerian government build new refineries, fix the ones it already has, and create plans for maintaining both the new and old refineries.

Trend Analysis Forecast of Life Expectancy at Birth in Nigeria was conducted by I.M Ogolo 2022, the study aimed at examining patterns in the birth life expectancy of Nigeria. Examples of life expectancy at birth in Nigeria were fitted using four different trend models of time series: linear, quadratic, exponential, and s-curve trend models (year data set from 1950 to 2021). Then, utilizing model accuracy criteria such as Mean Absolute Percentage (MAPE), Mean Absolute Deviation (MAD), and Mean Squared Deviation (MSD), the best trend model was selected. The findings demonstrated that the quadratic and s-curve trend models outperform in each of the three model accuracy approaches investigated. The s-curve trend model, on the other hand, best fit the data and had the lowest MAD and MAPE. Using the s-curve trend model, a six-year estimate of life expectancy at birth in Nigeria was produced. The results indicated that cases of life expectancy at birth in Nigeria will climb (grow) during the next six years, from 2022 to 2027.

1. Prince et al. (2016) conducted research on current worldwide patterns in dementia prevalence, incidence, and survival. They identified studies tracking dementia or Alzheimer's disease prevalence and incidence since 1980 within specific populations, totaling nine studies for prevalence and eight for incidence. Additionally, four studies tracked mortality among individuals with dementia. The findings hinted at a potential decline in dementia incidence in high-income countries, although evidence was moderately consistent. The project aims to explore the temporal patterns and forecast mortality linked to Alzheimer's disease and meningitis using ARIMA model techniques.

2. Theories/Calculations

The global database from which the data for this study was drawn is available on Kaggle, a well-known website for those interested in data science and machine learning. Specifically, the data set focuses on causes of death in Nigeria, offering a nuanced perspective on mortality pattern in Nigeria. The variables to be used are Alzheimer, and meningitis. The data covers 1990 to 2019.

Stationary And Non-Stationary Time Series

The majority of observed time series in the business and economic domains are non-stationary; these conditions include outliers, random walks, drift, deterministic trends, and non-constant variance. A stationary time series, on the other hand, has a zero mean, a constant variance, and a constant autocorrelation structure over time with no periodic fluctuations (seasonality). A stationary stochastic process is any of various types whose joint probability distribution remains constant with respect to time shifts. If a series has a constant auto-covariance structure in addition to a fixed mean and constant variance, it is considered rigorously stationary. The mean and variances of a weak stationary series, also known as second-order stationarity, do not change with time, but the auto-covariance is dependent on the series delays. When a time series data is non-stationary with respect to either the mean or the variance, it can be made stationar by using techniques like logarithmic or square-root transformation to stabilize the variance of the series, or by differentiating the original series once (linear trend).

Augmented Dickey-Fuller (ADF) Unit Root Tests

The dickey-fuller test investigates the presence of unit root in a time series using a regression expression that nets the mean, the lagged and deterministic time trend term.

$$Y_t - \varrho Y_{t-1} + \varepsilon_t \quad (3)$$

Where the variable of interest is Y_t and (t), (ϱ),(ε_t) represent the time index, the parameter coefficient and the residual (error term) respectively. The hypothesis of the ADF test is given

below as:

$$H_o : \varrho = 0$$

$$H_1 : \varrho < 0$$

The null hypothesis ϱ=0 is used when doing the unit root test. In contrast to the alternative (ϱ<0), which indicates that the original series is stationary, the original series is nonstationary. $DF_{t\,=}$ is used to compute the statistic, which is then compared to the pertinent critical value found in the Dickey-Fuller table. Differentiating is used until the ACF displays a discernible pattern with just a few significant autocorrelations if the test is unable to reject the null hypothesis.

Differencing:

If the series is found to be non-stationary, differencing is used to transform it into a stationary series.

The first difference of a series y_t is:

$$\Delta y_t = y_t - y_{t-1}$$

If the first difference is not stationary, higher order differencing can be used, such as the second difference:

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$$

Autocorrelation and Partial Autocorrelation Functions (ACF and PACF)

Once the series is stationary, ACF and PACF plots are used to identify the properties of the series, which help in determining the appropriate order of the ARIMA model.

ACF Plot

The ACF measures the correlation between the time series and its lags. It helps in identifying the degree of dependency between observations.

For a time series y_t, the autocorrelation at lag k is:

$$\rho_k = \frac{cov(y_t, y_{t+k})}{\sqrt{var(y_t)var(y_{t+k})}}$$

PACF Plot

The PACF measures the partial correlation between the series and its lag, removing the effects of intervening lags. It helps in identifying the order of an autoregressive model.

For a time, series y_t, the partial autocorrelation at lag k is the correlation between y_t and y_{t+k} after removing the linear influence of all lags from 1 to k-1.

Box-Jenkins ARIMA Process

A novel model that incorporates an integrated term along with the AR and MA processes was created by Box and Jenkins in 1976. The foundation of the ARIMA method was laid by Yule-walker's earlier research (1927). Just ARIMA (p, d, q) was used to express the model; (p) is the AR process, (d) is the integrated term that handles the non-stationary process, and (q) is the MA process. Three iterative steps may be roughly identified in the Box-Jenkins process of fitting and analyzing univariate time series: model formulation, parameter estimation, and model diagnostic testing.

ARIMA Forecasting Model

A view into the future is one of the main objectives of model building. In this study the forecast estimation function for optimal (h) periods ahead is expressed as:

$$\hat{y}_{T+h/T} = -\theta_1 \hat{\varepsilon}_{T+h-1/T} \quad (4)$$

The error is the difference between the actual series values and the fitted values from the chosen model in the study. The forecast error $\hat{\varepsilon_t}$ (h) at lead time is given by:

$$\hat{\varepsilon_t} = y_{T+h} - \hat{y}_{T+h/T} \quad (5)$$

were y_{T+h} represents the actual series value at T + h.

Model Selection Criteria

To select the best-fitting ARIMA model, model selection criteria are used.

Akaike's Information criterion:

In statistical modeling and model selection, the Akaike's Information Criterion (AIC) is a metric used to compare the goodness of fit of various models. The Japanese statistician Hirotugu Akaike created it.

Both the model's complexity and goodness of fit are taken into consideration by AIC. The fundamental idea is to strike a balance between the number of parameters in the model and how well it fits the data. This formula is used to calculate the AIC:

$$AIC = 2K - 2\ln(L) \quad (1)$$

Where;

The number of variables in the model, including the intercept, is denoted by the letter K.

A measure of how well the model describes the observed data is the log-likelihood.

The AIC penalizes models with more parameters, so it tends to favor simpler models that explain the data well. When comparing models, the model with the lower AIC is generally preferred, as it indicates a better balance between goodness of fit and model complexity. AIC is widely used in various fields, including statistics, econometrics, and machine learning, as a tool for model selection.

If sample size is small, we use the second order AICc:

$$AICc = AIC + \frac{2K(K+1)}{n-K-1`} \quad (2)$$

where;

n is sample size

K is number of parameters.

Bayesian Information Criterion (BIC)

BIC also evaluates model fit with a stronger penalty for the number of parameters:

$$BIC = \ln(n)\,K - 2\ln(L)$$

Residual Diagnostic Checking

After fitting the ARIMA model, it is essential to check the residuals to ensure they resemble white noise, indicating a good model fit.

Box-Ljung Test

The Box-Ljung test checks the null hypothesis that the residuals are independently distributed:

$$Q = n(n+2)\sum_{k=1}^{m} \frac{\widehat{\rho^2}_k}{n-k}$$

where:

n is the sample size,

m is the number of lags,

ρ^2 is the autocorrelation of the residuals at lag k.

Shapiro-Wilk Test

The Shapiro-Wilk test assesses the normality of the residuals:

$$W = \frac{\sum_{i=n}^{n} a_i X_{(i)}}{\sum_{i=1}^{n} (X_i - \bar{x})^2}$$

where:

$X_{(i)}$ are the ordered sample values,

a_i are constants derived from the sample size.

Model Forecasting

Using the fitted ARIMA model, forecasts can be made for future time points. The forecast includes point forecasts and prediction intervals, which account for the uncertainty in the forecasts.

The general formula for the forecast is:

$$\hat{y}_{t+h} = \mu + \emptyset_1 \hat{y}_{t+h-1} + \emptyset_2 \hat{y}_{t+h-2} + \cdots + \emptyset_\rho \hat{y}_{t+h-\rho} + \theta_1 \epsilon_{t+h-1} + \theta_2 \epsilon_{t+h-2} + \cdots + \theta_q \epsilon_{t+h+q}$$

Where:

μ is the mean of series

Φ_i are the AR coefficients,

Θ_i are the MA coefficients,

ϵ_i are the forecast errors.

RESULTS AND DISCUSSION

Descriptive Statistics

The table.1 below displays the descriptive statistics of the meningitis and Alzheimer disease

Table 1 Descriptive statistics for meningitis and Alzheimer series.

	Sample size	Mean	Median	Standard deviation	skewness	kurtosis	Variance
Meningitis	30	50679	51856	5099	-0.30	-0.59	26004553
Alzheimer disease	30	8057	7814	2297	0.26	-1.38	5276454

The table shows the descriptive statistics of deaths due to meningitis and Alzheimer disease. It shows that death occurs more from meningitis than Alzheimer disease because it has a higher mean occurrence. Meningitis disease has higher standard deviation which shows that the points spread out and are highly far away from the mean. While Alzheimer has a lower standard deviation which implies that the points are clustered around the mean.

Time Series Plot

Figure 1 and 2 below represent the time plots for the Alzheimer series and meningitis series respectively. The meningitis curve shows that it has higher occurrence compared Alzheimer disease there 2 major peaks at the years 1996 and 2009. It shows to have been on a steady increase from 1990 up until 1996.

The Alzheimer curve shows a minimal occurrence though there has been a steady increase up till the year 2019.

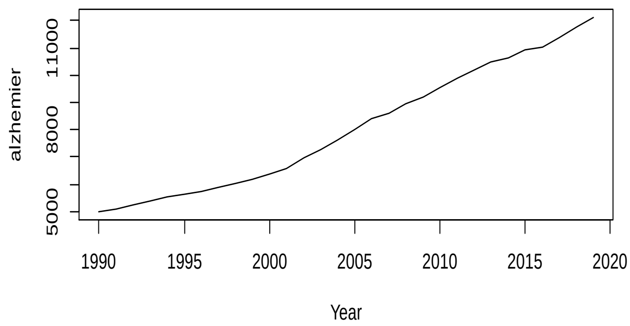

Figure 1. Time series plot for Alzheimer disease and other dementias

Figure 2. Time series plot for the meningitis series.

Stationary Test Analysis

Preliminary tests for stationary property (that is the series remains at a constant level over time) before attempting to fit a suitable time series model were conducted. However, if a trend exists, the series is not stationary. This is examined graphically as well as using the Augmented Dickey-Fuller (ADF) tests. The variables have to be checked for unit roots and the order of integration of each series must be determined. The stationarity of the series was tested using time series plots followed by the Augmented Dickey Fuller (ADF) test the null (H_0) and alternative hypothesis (H1) are:

H_0: The series is not stationary Vs H_1: the series is stationary

From Table 2 below, the p-value for the ADF test is all greater than the significance level of 0.05. Therefore, we do not reject the null hypothesis and conclude that there is an indication of non- stationary per test hypotheses provided above which shows that the data set for

meningitis and Alzheimer disease are non-stationary.

Table 2. Stationarity test for meningitis and Alzheimer

Series fuller			**Augmented dickey**
		ADF test	
Variable	**Test statistic**	**P-value**	**Remarks**
Meningitis	-0.84093	0.9456	Non stationary
Alzheimer	-2.7518	0.2839	Non stationary

ACF and PACF plots

Figure 3 below shows the ACF and PACF plot for Alzheimer series. A few of the spikes are fall outside the 95% confidence interval. The plot shows that the series is not stationary

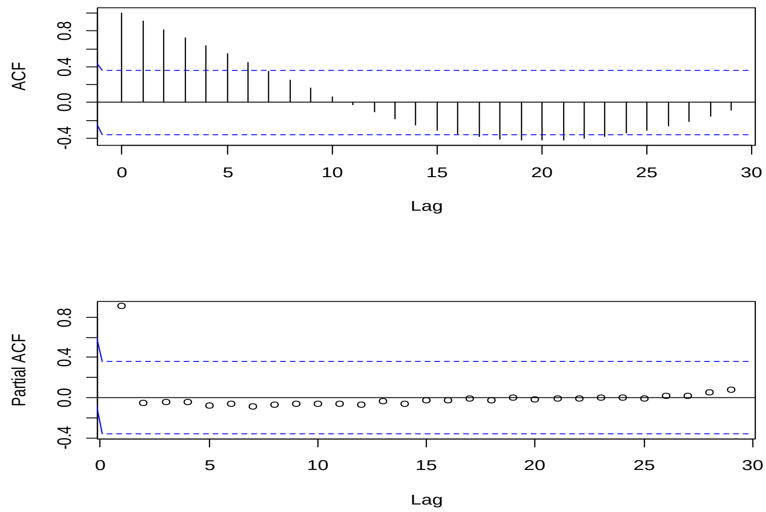

Figure 3 ACF and PACF plots for the Alzheimer series.

Figure 4. Below shows the ACF and PACF plots for meningitis series. The plot also shows that the series is non-stationary. A few of the spikes is outside the 95% confidence interval.

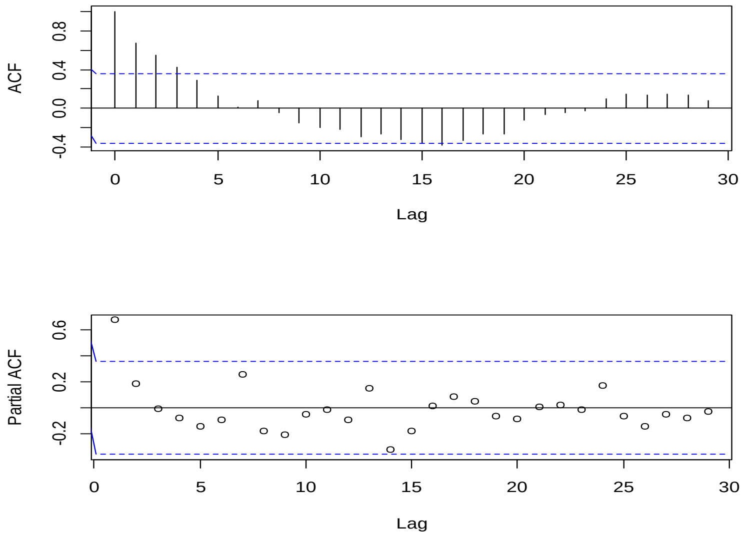

Figure 4. ACF and PACF plot for the meningitis series

The both series are to be differenced to attain stationarity. The Alzheimer series is to be differenced twice while the meningitis series is to be differenced once to attain stationarity. The figure 5 below shows the time plot of the first order differenced series for the Alzheimer series

Figure 5. time plot of the first order differenced Alzheimer series.

The figure below shows the ACF and PACF plots for the first order differencing for the Alzheimer series.

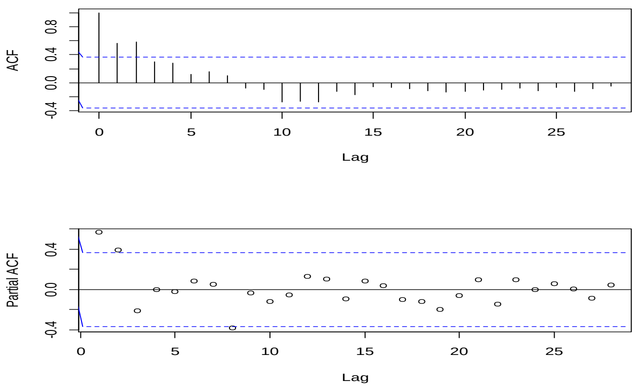

Figure 6. ACF and PACF plot for the first order differenced Alzheimer series.

The ACF and PACF plots above still shows non stationarity. Furthermore, the augmented dickey fuller test is conducted on the differenced series of the Alzheimer series producing the result below.

Table 3. ADF result for the first order differenced Alzheimer series

Series fuller		Augmented dickey	
		ADF test	
Variable	**Test statistic**	**P-value**	**Remarks**
Alzheimer	-2.2979	0.4584	Non stationary

The p value for the ADF test carried out is greater than the level of significance 0.05 which still indicates non stationarity. A second order differencing was carried out along. And it produced the following results.

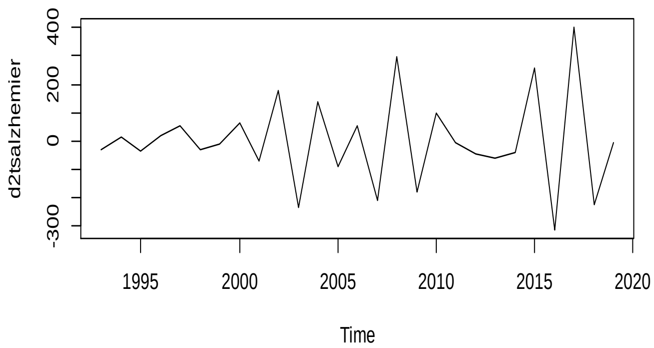

Figure 7. Time plot for the second order differenced Alzheimer series

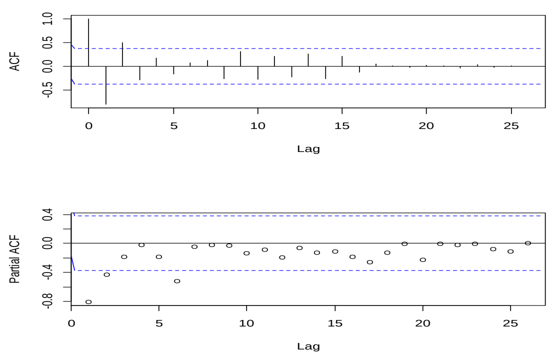

Figure 8. ACF and PACF plot for the second order differenced Alzheimer series.

The figure 8 above represents the PACF and ACF plots of the second order differenced Alzheimer series. The plot still indicates non stationary, the ADF test was conducted on the second order differenced Alzheimer series alongside the first order differenced meningitis series and the following results was generated.

Table 4 ADF test for the first order differenced meningitis series and second order Alzheimer series

Series fuller		**Augmented dickey**	
		ADF test	
Variable	**Test statistic**	**P-value**	**Remarks**
Meningitis	-3.4953	0.01305	Stationary
Alzheimer	-4.2081	0.01572	Stationary

The table shows that the second order differenced series for the Alzheimer series is stationary as the P value 0.01572 is less than the level of significance 0.05. the figures below show the time series plot and PACF and ACF plot for the first order differenced meningitis series respectively

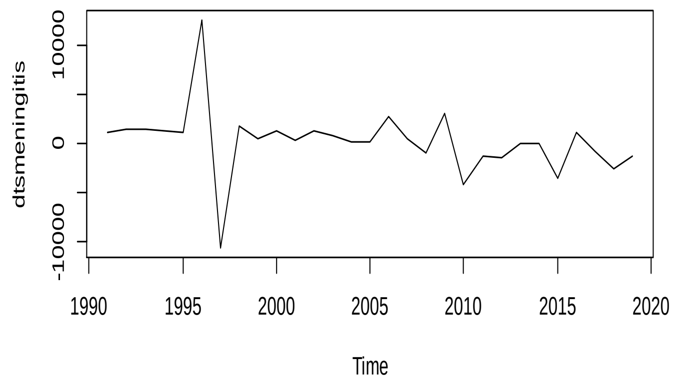

FIGURE 9. Time Series Plot for First Order Differenced Meningitis Series.

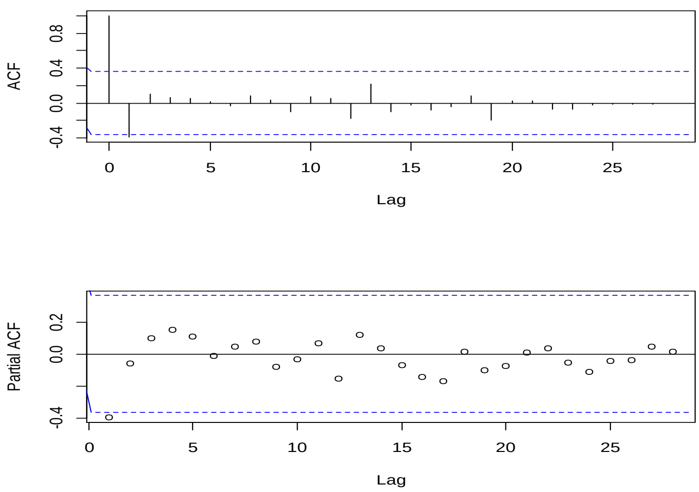

Figure 10. ACF And PACF Plot for The First Order Differenced Meningitis Series.

From the figure 10 above the spikes in the ACF plot all fall within the 95% confidence interval, therefore indicating stationarity in the series. Though one of the points in the PACF plot falls outside the 95% confidence interval all others fall within the confidence interval thereby supporting the ACF plot and the ADF test in confirming the stationarity of the series.

Identifying the Appropriate ARIMA Model for the Series.

The table below shows the tentative models for the Alzheimer series and meningitis series respectively

Table 5. ARIMA Models fit for Alzheimer series

Model	AIC	AICc	BIC	LIKELIHOOD
ARIMA (2,2,2)	326.65	329.38	333.32	-158.33
ARIMA (0,2,0)	332.54	332.69	333.87	-165.27
ARIMA (1,2,0)* **	**323.6**	**324.08**	**326.27**	**-159.8**
ARIMA (0,2,1)	326.81	327.29	329.48	-161.41
ARIMA (2,2,0)	325.37	326.37	329.37	-159.68
ARIMA (1,2,1)	325.41	326.41	329.41	-159.71
ARIMA (2,2,1)	327.34	329.08	332.67	-159.67

The ARIMA(1, 2, 0) appears to have the least AIC (323.54) thereby making it the best fit for the Alzheimer series out of the 7 ARIMA models of the second order differenced Alzheimer series.

Likewise for the meningitis series, the ARIMA (1, 1, 0) has the minimum AIC (556.37) making it the best fit for the series. Out of the 7 ARIMA models for the first order differenced meningitis series. To estimate the AR coefficient in ARIMA (1,1,0) model, the result from Table 6 below show that the estimated coefficient values for meningitis and Alzheimer are significant, since their p-values of 0.0224 and 0.0002 respectively are less than 0.05 level of significant. The constants for meningitis and Alzheimer are 50679 and

8057 respectively. The estimated coefficients strictly conform to the bounds of stationarity and invertibility since their values lies between -1 and 1. Hence, the chosen model based on the principle of parsimony is expressed as:

Table 7. Model Fit Statistics

For meningitis series				
Coefficient	**Estimate**	**Std.Error**	**t-value**	**p-value** $P(>\|t\|)$
AR(1)	0.3828	0.1677	2.282648	0.0224
For Alzheimer series				
Coefficient	**Estimate**	**Std.Error**	**t-value**	**p-value** $P(>\|t\|)$
AR(1)	0.5582	0.1499	3.723816	0.0002

From equation 1 above, it implies that when lag of the residual increases by one year, the predicted meningitis series increases by 0.3828.

From equation 2 above, it implies that when lag of the residual increases by one year, the predicted Alzheimer series increases by 0.5582.

Diagnostic Check

Checking the adequacy of the fitted model in other to deem the model fit for forecasting the future values of meningitis series and Alzheimer series. The diagnostic checking is performed using the time plots, ACF and PACF plots of the model residuals coupled with other objective diagnostic tests such as the Box-Ljung serial correlation test, Shapiro-Wilk Normality. The model residuals series plots and correlogram plots are shown in Figure 11 and 12 for the meningitis series and the Alzheimer series respectively below;

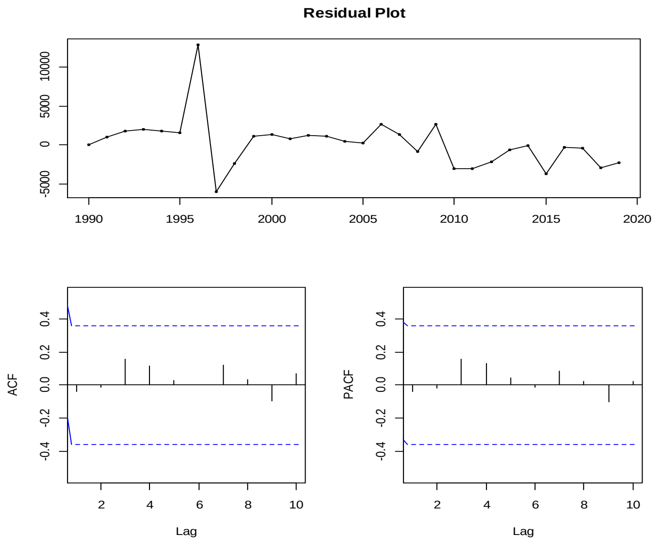

Figure 11. Residual plots for the meningitis series

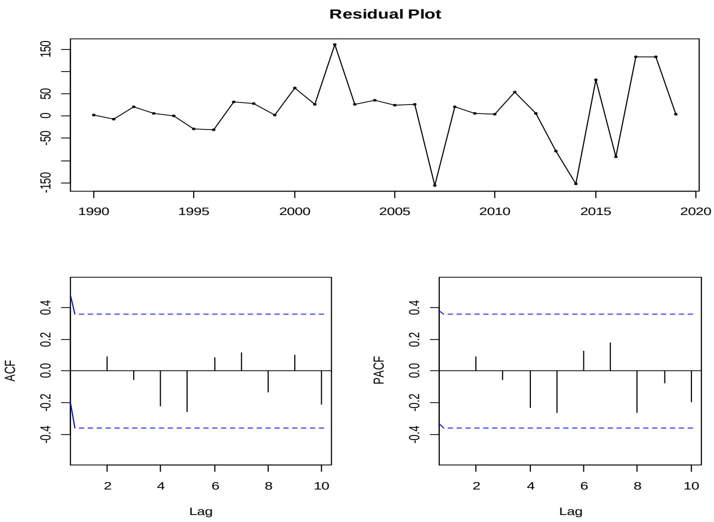

Figure 12 Residual Plot for the Alzheimer series.

To determine the adequacy of the fitted model, the ACF, PACF and the residual plots is used. Residuals plot for the meningitis series revealed that the residual is accurate, since the residual plot satisfies stationary condition. The ACF and PACF from figure 5. also reveals that the model is accurate, since the spikes fall within the 95% confidence bound. Likewise, the figure 12

ARIMA Model Residual Diagnostic Test.

Diagnostics for meningitis series

Table below revealed that Ljung-Box test, evaluating autocorrelation in ARIMA (1,1, 0) model residuals, yielded a chi-square statistic value of 6.0897 with a corresponding p-value of 0.9782. Since the p-value is greater than 0.05 level of significance, the null hypothesis of no presence of autocorrelation in the residuals is not rejected. This implies that the assumption of no auto correlation is not violated.

Table 8.Box-Ljung test

Box-Ljung test			
Test Type	Chi-Square	Df	P – Value
Ljung-Box	6.0897	15	0.9782

Table 9. Shapiro wilks normality test for the meningitis residual

Shapiro-Wilk Test Statistics		
Shapiro-Wilk	**Wilks's value**	**p-value**
	0.81394	0.00019

Table 9. Above present the Shapiro-Wilk test for the meningitis series residuals, the result has a p-value of 0.00019 which is less than the 5% level of significance. Hence, we reject the null hypothesis of the sample being drawn from the normal distribution. This implies that the residual is not normally distributed and therefore needs to be transformed.

Table 10. Shapiro-Wilk Test Statistics after Log-Transformation

Shapiro-Wilk Test Statistics After Log-Transformation		
	Wilks's value	**P –value**
Shapiro-Wilk	0.95358	0.3466

Diagnostics for Alzheimer Series

Table 11. Box-Ljung test

Box-Ljung test			
Test Type	Chi-Square	Df	P – Value
Ljung-Box	Inf	15	2.2e-16

Table 12. **Shapiro-Wilk Test Statistics**

Shapiro-Wilk Test Statistics		
	Wilks's value	**p-value**
Shapiro-Wilk	0.91043	0.01526

Above present the Shapiro-Wilk test for the Alzheimer residuals, the result has a p-value of 0.01526 which is less than the 5% level of significance. Hence, we reject the null hypothesis of the sample being drawn from the normal distribution. This implies that the residual is not normally distributed and therefore needs to be transformed.

Shapiro wilks test for Alzheimer series after transformation

Table 13. Shapiro-Wilk Test Statistics after Log-Transformation

Shapiro-Wilk Test Statistics After Log-Transformation		
	Wilks's value	**p-value**
Shapiro-Wilk	0.95358	0.3466

Above shows the Shapiro-wilk test after log transformation. The result revealed that a p-value of 0.3466 which is greater than 5% level of significance. Hence, we do not reject the null hypothesis of the sample being drawn from a normal distribution. This implies that the residual is normally distributed.

Model Forecast

A view into the future is one of the main objectives of model building. Since the models passed all the relevant diagnostic tests, it is used to make prediction for the future deaths from meningitis and Alzheimer disease. The forecast values for the next 3 years.

Table 14. ARIMA forecast for the meningitis and Alzheimer series

Forecast for the meningitis series			
Year	Forecast	95% lower CI	95% upper CI
2020	45401.36	39023.67	51779.05
2021	45216.38	37710.89	52721.86
2022	45286.59	36336.43	54236.75
Forecast for the Alzheimer series			
2020	12472.44	12328.00	12616.89
2021	12832.20	12578.75	13085.64
2022	13191.77	12785.85	13597.70

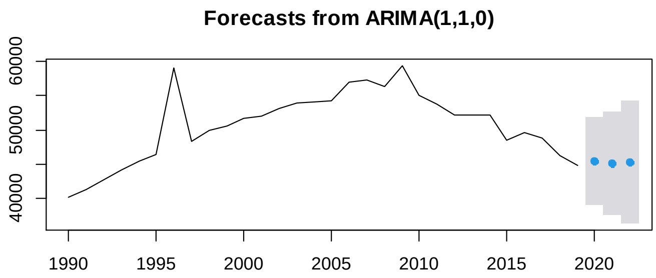

Figure 13. Time plot for meningitis series including 3 years forecast

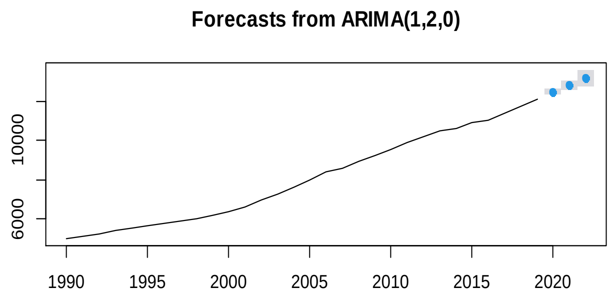

Figure 14. Time plot for the Alzheimer series including the 3 years forecast

The yearly data of meningitis and Alzheimer disease from 1990-2019 was used for this research project. Box-Jenkins (ARIMA) modeling procedure was applied to the data series and the findings from the analysis are obtained.

The time series plot for meningitis and Alzheimer disease. The meningitis curve shows that it has higher occurrence compared Alzheimer disease there two major peaks at the years 1996 and 2009. The Alzheimer curve appears to have been on a steady increase from 1990 up until the year 2019.

Test for stationarity using the ADF, ACF and PACF plots was carried out to determine the stationarity of both series. The ADF test Suggested that the both series were non stationary. Likewise, the ACF and PACF plots also suggested non stationarity. To achieve stationarity, the two series were differenced. The meningitis series was differenced once to attain stationarity while the Alzheimer series was differenced twice to attain stationarity.

The ADF test was repeated on the differenced series and it was concluded that the condition of stationarity had been met.

Seven ARIMA models were shown to most be suitable for the series out of which the best was selected i.e., ARIMA (1, 1, 0) for the meningitis series and ARIMA (1, 2, 0) for the Alzheimer series on the basis of minimum value AIC selection criteria.

The parameters for the models were estimated and it was observed that when lag of the residual increases by one year, the predicted meningitis series increases by 0.3828 and when lag of the residual increases by one year, the predicted Alzheimer series increases by 0.5582.

The diagnostic checking was performed using the time plots, ACF and PACF plots of the model residuals coupled with other objective diagnostic tests such as the Box-Ljung serial correlation test, Shapiro-Wilk Normality. The assumption of no auto correlation was not violated. The Shapiro wilks test showed that the residuals were not normal which led to carrying out LOG transformation for both series. After all assumptions have been met, the model was used to make predictions for the next 3 years and predictions at 95% confidence interval for the meningitis series appears to be very much higher than that of the Alzheimer series.

CONCLUSION

ARIMA (1, 2, 0) model is the most suitable model for the Alzheimer series and the ARIMA (1, 1, 0) model is the most suitable model for the meningitis series. The fitted models are free from violation of all assumptions. The forecasted values revealed that there will be inconsistent slight decrease in meningitis death while there will be an increase in

Alzheimer deaths over the years. The predicted value for meningitis is 45401.36, 45216.38, and 45286.59.

And the predicted value for Alzheimer is 12472.44, 12832.20, and 13191.77 for the years 2020, 2021, and 2022 respectively all at 95% confidence interval.

Based on the results obtained from this research, the following recommendations were made. To control deaths due to meningitis and Alzheimer's disease in the population, fitting appropriate ARIMA model to see future occurrence will be helpful. From the forecasted values for the Alzheimer's disease, intervention needs to be in place to stop the increase. The forecasted value for meningitis shows that there is a need for interventions to be in place to accelerate the decrease in meningitis deaths.

Data Availability

The Data set used in this research work was extracted from world Data of causes of death in Nigeria from 1990-2019 accessible on www.kaggle.com.

REFERENCES

1. Wu, Y., Fan, H., Su, B., Guo, C., & Feng, L. (2023). *Long-term patterns of meningitis mortality: A continual downward trend and a vulnerable infant population* — China, 1987–2021.
2. Four decades temporal trend in meningitis cases in Africa (1971-2010): a tale of three countries Olalekan A. Uthman1,2,3&, Gbenga.A.Kayode1, Bolaji E. Egbewale3
3. Tagbo, B. N., Bancroft, R. E., Fajolu, I., Abdulkadir, M. B., Bashir, M. F., Okunola, O. P., Isiaka, A. H., Lawal, N. M., Edelu, B. O., Onyejiaka, N., Ihuoma, C. J., Ndu, F., Ozumba, U. C., Udeinya, F., Ogunsola, F., Saka, A. O., Fadeyi, A., Aderibigbe, S. A., Abdulraheem, J., Yusuf, A. G., … Antonio, M. (2019). Pediatric Bacterial Meningitis Surveillance in Nigeria From 2010 to 2016, Prior to and During the Phased Introduction of the 10-Valent Pneumococcal Conjugate Vaccine. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *69*(Suppl 2), S81–S88. https://doi.org/10.1093/cid/ciz474
4. Salawu, F. K., Umar, J. T., & Olokoba, A. B. (2011). Alzheimer's disease: A review of recent developments. Department of Medicine, Neurology Unit, Federal Medical Centre, Yola, Department of Medicine, University of Ilorin Teaching Hospital, Ilorin, Nigeria.
5. Elisha J. Inyang *Intervention Model Based on Exponential Smoothing Methods and ARIMA Modelling of the Nigerian Naira Exchange Rates International Journal of Scientific Research in Mathematical and Statistical Sciences Vol.11, Issue.1, pp.24-33*,
6. Gana, G. J., Badung, S., Bunza, A. U., Gidado, S., & Nguku, P. (2017). Outbreak of cerebrospinal meningitis in Kebbi State, Nigeria. Epidemiogy Unit, Department of

Public Health, Ministry of Health, Birnin Kebbi, Kebbi State. *Annals of Ibadan Postgraduate Medicine*, 15(1), June 2017.

7. Adeloye D, Auta A Ezejimofor M, Oyedokun A, Harhay MO, Rudan I, Chan KY. Prevalence of dementia in Nigeria: a systematic review of the evidence. J Glob Health Rep 2019; 3: e2019014.
8. N.F. Adeleye, H.O. Ilo, T.O. Gbadamosi, *"Time Series Analysis of Crude Oil Production in Nigeria between the Years 2010 to 2020", International Journal of Scientific Research in Mathematical and Statistical Sciences,* Vol.10, Issue.2, pp.38-45, 2023
9. I.M. Ogolo, "*Trend Analysis Forecast of Life Expectancy at Birth in Nigeria*", International Journal of Scientific Research in Mathematical and Statistical Sciences, Vol.9, Issue.4, pp.33-40, 2022
10. Prince, M., Ali, G. C., Guerchet, M., Prina, A. M., Albanese, E., & Wu, Y. T. (2016). Recent global trends in the prevalence and incidence of dementia, and survival with dementia. *Alzheimer's Research & Therapy,* 8(1). https://doi.org/10.1186/s13195-016-0188-8
11. Fletcher, S., & Rees, G. (2015). World Alzheimer Report 2015: *The Global Impact of Dementia.*