Digital Innovations & Contemporary Research in SCIENCE & ENGINEERING
Vol. 4 No. 3, Sept. 2016

Article Citation Format
Obanla O. J & Awariefe, C. (2016): Modeling Nigeria's Life Expectancy Using ARIMA Model Journal of Digital Innovations & Contempt Res. In Sc., & Eng Vol. 4, No. 4. Pp 59-66

Article Progress Time Stamps

Article Type: Research Article
Manuscript Received: 19rd June , 2016
Review Type: Blind
Word Count Post Review: 2584
Review/Acceptance Information Sent : 2nd July, 2016
Final Acceptance:: 27th Aug, 2016
DOI Prefix: 10.22624

Modeling Nigeria's Life Expectancy Using ARIMA Model

Obanla O. J[1], Awariefe C[2],
[1] Department of Statistics, Ogun State Institute of Technology, Igbesa
[2] Department of Mathematics & Statistics, Delta State Polytechnic, Ozoro

ABSTRACT

This paper fit a time series model to the life expectancy in Nigeria between 1980 and 2014 and provided five years forecast for the expected life expectancy in Nigeria. The Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model was estimated and the optimal fitted ARIMA model was adopted to obtain the post-sample forecasts. It was discovered that the best fitted model is ARIMA (1, 2, 0). The model was further validated by the adequacy of the AR and MA models. The ARIMA model revealed that the Nigeria life expectancy is moving gradually upward. Finally, the five years forecast was made from 2014 to 2019 with the highest life expectancy being estimated as 55.038 in 2019.

Keywords: Life Expectancy, ARIMA, Stationarity, Differencing, ACF, PACF.

1. INTRODUCTION

Life expectancy at birth signifies the number of years an infant would live if current patterns of mortality at the time of being given birth were to remain the same all through its life. Sede and Ohemeng (2015) postulated that Life expectancy measures the length of life expected to be lived by an individual at birth. Life expectancy is very sacrosanct to the developing economies who are committed to achieve socio-economic improvement by investing extensively in social sectors such as health care, education, sanitation and environmental management, and social safety nets. Touili et al (2015) posited that Life expectancy at birth shows the number of years a newborn infant would live if current patterns of mortality at the time of its birth were constant throughout its life. Life expectancy is the approximate average years a person is expected to live before he/she dies. Jie etal (2001) compared West-African sub-region in terms of life expectancy regional estimates, Nigeria ranked 15th out of 17 countries in the, with Cape Verde having the highest life expectancy at birth that is, 73 years and Sierra Leone having the lowest life expectancy. In Nigeria, as in other less developed economies, disparities in mortality have been connected with a broad array of measures of socio-economic status including per capita GDP, fertility rate, high level of adult illiteracy rate, per capita calorie intake, poor health-care system, poor portable drinking water, urban residents, unemployment rate and the insignificant exchange rate.

Life expectancy is very crucial to the less developed country like Nigeria that is earnestly striving to attain socio-economic progress through investment in social sectors like health, education, sanitation, environmental management and sustainability, and social youth empowerment. Improvement of Life expectancy to at least 70 years in the future should be one of Nigeria's health policy targets, based on this it becomes imperative to develop a model that will best forecast life expectancy which may improve macroeconomic performance and make government improve on health care and well-being of Nigerians.

Therefore, an effective health policy depends largely on the ability to develop a reliable model that could help understand the ongoing socio-economic processes and predict future developments. In this regard, this study is important since it is aimed at forecasting Nigeria Life Expectancy, which is a component of the sustainable development goals (SDG) by 2030 in the Nigeria.

This study therefore seeks to fit an ARIMA model to the yearly data on Nigeria Life Expectancy from 1980 – 2014. The research shall in addition present five years forecast. In many practical applications, the autoregressive integrated moving average (ARIMA) model is the most widely used Box-Jenkins models since it can handle non-stationary data Shafiee and Topal (2010). ARIMA model is employed for the reason that it can generally handle many series in spite of stationarity or non-stationarity, with seasonal or without seasonal components.

The paper is prepared as follows; section one consists of the introduction, section two presents the methodology. The data description and analysis are presented in section three, section four dealt with the fitting the model and forecasting while the last section focused on the result obtained and conclusion.

2. METHODOLOGY

The Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) methodology was adopted in this research. This method is a set of procedures for identifying and estimating time series models contained by the category of (ARIMA) models. We refer also of AR models, MA models and ARMA models which are unique cases respectively. The development and designing of ARIMA models as forecasting tools of variables is known as Box-Jenkins Methodology (1976).

The identification, fitting, checking is undertaken systematically using integrated autoregressive, moving average (ARIMA) time series models. The ARIMA methodology consists of four iterative steps of model identification, parameter estimation, diagnostic checking and forecasting. If the time series is non stationary, it is then transformed in the identification step to make it stationary.

Stationary process is a requirement in developing an ARIMA model. When time series under investigation exhibits trends and non-seasonal behaviour, transformation and differencing are applied to the observed series in order to stabilize variance and to remove the trend before an ARIMA model is applied, Ruiz-Aguilari et al (2015).

2.1 Source of Data for The Study

Secondary source of data was adopted for this study and the data was obtained from the website of the World Health Organization (WHO). The data used for this study consists of yearly data of Nigeria life expectancy collected over the period from 1980 to 2014. The data was analyzed with Autoregressive Integrated Moving Average made popular by Box-Jenkins; all analysis performed with R programme.

2.2 Method of Estimation: ARIMA Methodology

According to Box-Jenkins (1976), ARIMA model has three parts:

I. The autoregressive part (AR) is a linear regression that relates past values of data series to future values.
II. The integrated part indicates how many times the data series has to be differenced to get a stationary series
III. The moving average (MA) part that relates past forecast errors to future values of data series. The processes types are AR (p), MA (q), ARMA (p, q), ARIMA (p, d, q).

The three basic Box-Jenkins models for Yt are:

1. Autoregressive model of order p (AR (p)):

$$Yt= \delta+ 1Yt-1+ 2Yt-2+ \ldots\ldots\ldots\ldots+ pYt-p+ t \qquad (1)$$

The autoregressive part is a linear regression that relates past values of data series to future values (i.e., Yt depends on its p previous values)

2. Moving average model of order q (MA (q)):

$$Yt= \delta+ t-\theta 1t-1-2t-2-\ldots\ldots\ldots\ldots-pt-p \qquad (2)$$

The moving average part that relates past forecast errors to future values of data series (i.e., Yt depends on its q previous random error terms)

3. Autoregressive moving average model of orders p and q (ARMA (p; q)):

$$Yt=\delta+ 1Yt-1+ 2Yt-2+ \ldots\ldots+ pYt-p+t-\theta 1t-1-2t-2-\ldots\ldots-pt-p \qquad (3)$$

(i.e., depends on its p previous values and q previous random error terms)

Obviously, AR and MA models are special cases of ARMA.
t is typically assumed to be "white noise"; i.e., it is identically and independently distributed with a common mean 0 and common variance 2 across all observations.
We write t˜ i.i.d (0,2).The white noise assumption rules out possibilities of serial autocorrelation and heteroscedasticity in the disturbances.

2.3 General ARIMA Representation

If Yt is integrated of order d, we write Yt˜I(d). Now, suppose that Yt˜I(d) and the stationary series after a dth order differencing, Wt , is represented by an ARMA (p; q) model.

Then we say that Yt is an ARIMA (p; d; q) process, that is, (1-B)dYt = Wt

$$= \delta+ 1Wt-1+ 2Wt-2+ \ldots\ldots+ pWt-p+ t-1t-1-2t-2-\ldots-pt-p \qquad (4)$$

According to Adams etal (2014) the Box –Jenkins model building techniques consist of the following four steps:

Step 1:

Preliminary Transformation:

If the series violate stationarity assumption, then it may be necessary to make a transformation in order to generate a series that is well-suited with the assumption of stationarity, if after appropriate transformation has been done and the sample autocorrelation function (ACF) appears to be non-stationary, differencing may be carried out.

Step 2:

Identification: The problem at the identification stage is to find the most satisfactory ARIMA (p,q) model to represent the stationary series in step 1. next is the determination of the integer parameters (p,q) that governs the underlying process of the stationary series by examining the autocorrelations function (ACF) and partial autocorrelations (PACF) of the stationary series. This step sometimes poses some difficulties and involves a lot of subjectivity, it is useful to entertain more than one structure for further analysis. Salau (1998) stated that this decision can be justified on the ground that the objective of the identification phase is not to rigidly select a single correct model but to narrow down the choice of possible models that will then be subjected to further examination.

Step 3:

Estimation of the model: This deal with estimation of the tentative ARIMA model identified in step 2. Parameters of the can be estimated by the conditional least squares and maximum likelihood.

Step 4: Diagnostic checking: Having chosen a particular ARIMA model, and having estimated its parameters, the adequacy of the model is checked. if the model is adequate; we accept the model, else we go to step 1 again and start over.

3. DATA ANALYSIS

In this section the ARIMA modeling strategy discussed in section 2.3 is applied to analyze the data on Life Expectancy. In this framework, model building commences with the examination of the plot of the series, the second and third difference plot, and sample plot of the autocorrelation (ACF), partial autocorrelation (PACF), model description and forecast value using the fitted model.

3.1 Results and Discussion

Assessing the life expectancy data using a graphical descriptive measure below, one could have a clue that the data is non stationary.

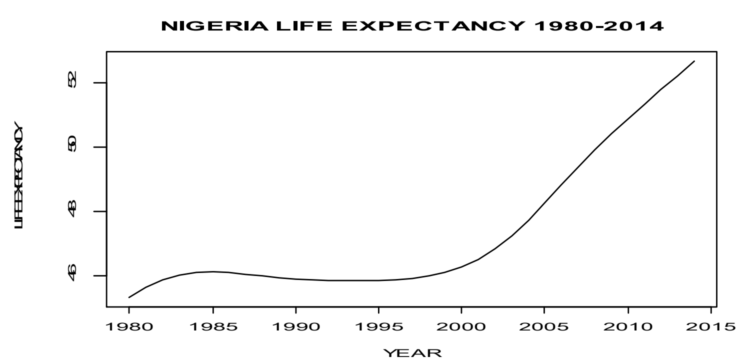

Figure 1: Time series trend of Nigeria life expectancy 1980-2014.

Augmented Dickey-Fuller (ADF) Test with Dickey-Fuller = 0.087408, Lag order = 3 and p-value 0.99 which is greater than 0.05 level of significant leads to infer that the series is non-stationary as such support the hint obtained from figure 1 above. In order to make the data stationary, differencing was introduced on the life expectancy data (1980-2014). First differencing shows that Dickey-Fuller = -1.9694, Lag order = 3, with p-value = 0.5849> 0.05 confirms the data remain non stationary. Meanwhile, on second differencing the series become stationary with Dickey-Fuller = -3.8857, Lag order = 3 and p-value = 0.02666 which is less than 0.05, as such making us conclude that there is a sufficient evidence to say the series is now stationary, leading us to obtain the optimal parameters (p,q) of ARIMA with d=2.

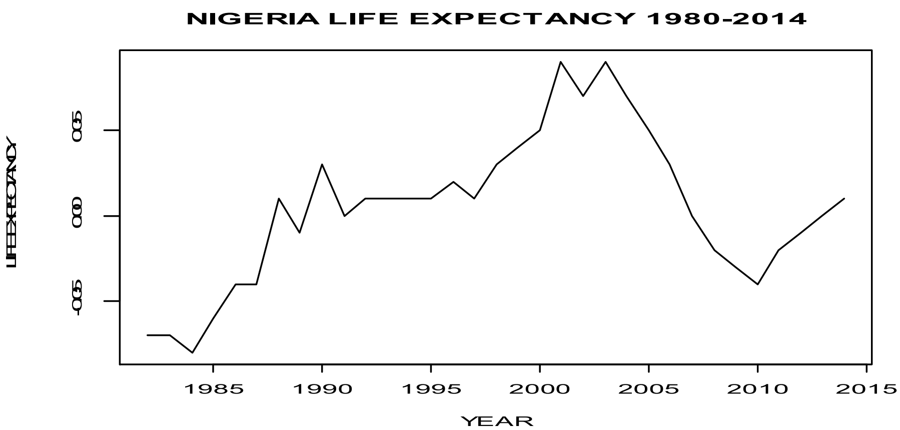

Figure 2: Figure 1: Time series plot of Nigeria life expectancy 1980-2014 on second differencin

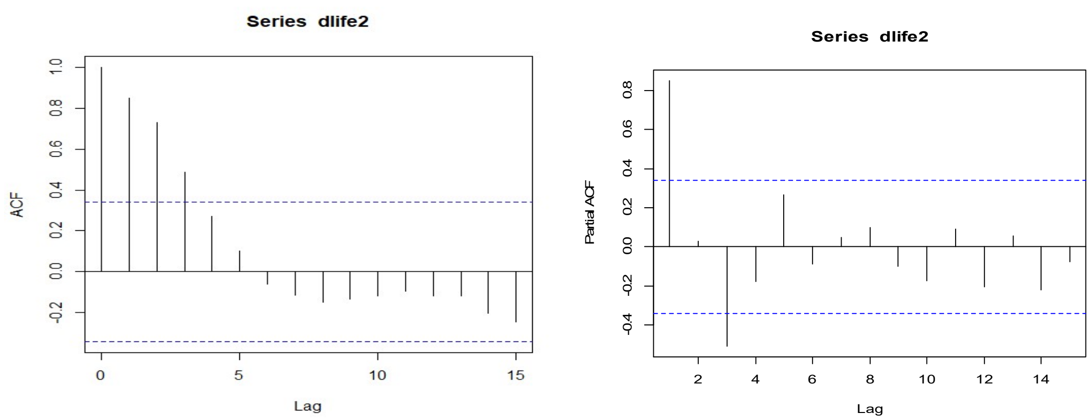

Figure 3: ACF and PACF of the Second differencing

Table 1: Table of ARIMA (1, 2, 0)

Coefficients: ar1	Standard Error	sigma^2 estimated	Log Likelihood	AIC	AICc	BIC
0.8996	0.0710	0.0005246	82	-160.01	-159.61	-157.02

Following the plotted time series and the second difference, the original plot and first differencing depict an apparent non stationarity. However, the second difference looks much stationary as shown in Figure 3 above and supported by Table 1. The ACF cuts off some lags at the beginning. Further test confirms the stationarity of the differenced data. With the help of auto.arima function in R of the forecast package, p and q parameters were obtained and ARIMA (1,2,0) model emerge with least AIC and BIC values.

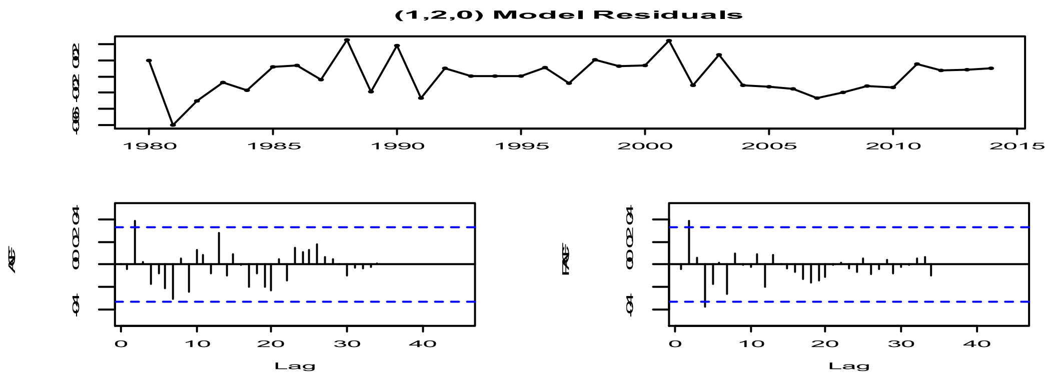

Figure 4: ARIMA (1,2,0)

The above ARIMA (1,2,0) model has an AIC of -160.01 and BIC of -157.02. Checking if the residuals are from a white noise series by performing

Ljung-Box test

data: Residuals from ARIMA(1,2,0)

Q^* = 13.903, df = 6, p-value = 0.03074

Model df: 1. Total lags used: 7

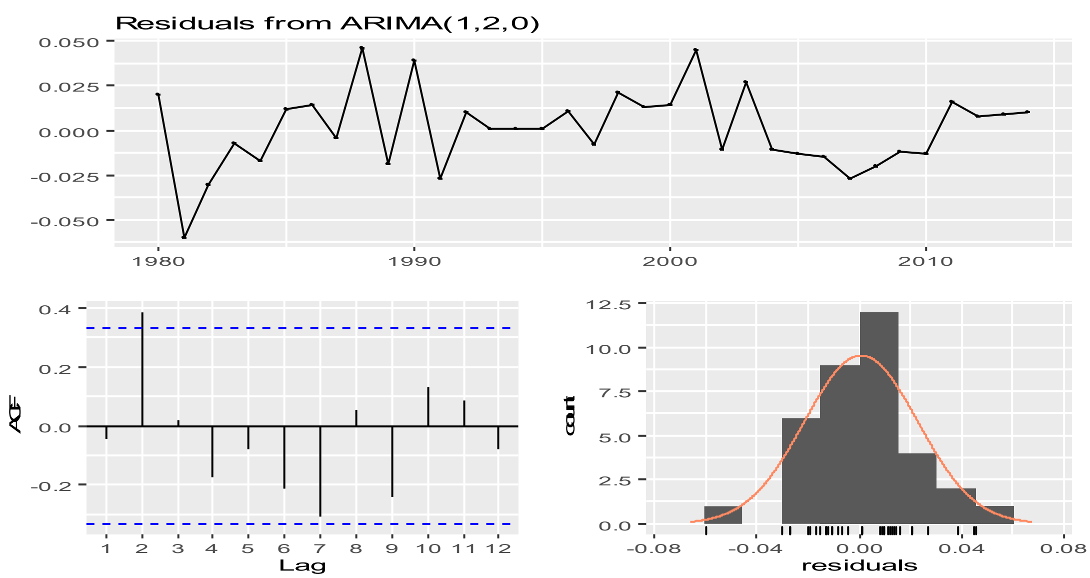

Figure 5: Diagnostics check of Residual from ARIMA (1,2,0)

It is clear to confirm that the residuals are not distinguishable from a white noise series as the results are not significant as displayed in Figure 5 above. Since the ARIMA (1,2,0) is indeed a good model as selected by using the auto.arima test in R, Checking the normality of the residuals of the model. The Figure 5 above shows that the histograms of the residuals does not reveal a serious deviation from normality. Next formal test (Anderson Darling test) was run to support the normality behaviour. The p-value 0.03074 indicates that the residuals are significantly different from normality.

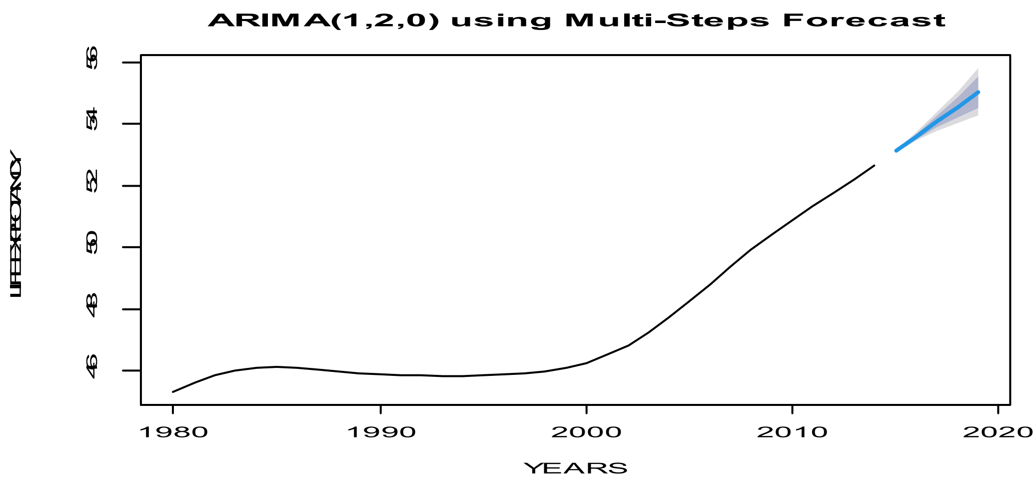

Figure 6: ARIMA (1,2,0) using Multi-Steps Forecast

Table 2: table for Five Year forecast

YEAR	2015	2016	2017	2018	2019
POINT FORECAST	53.12900	53.59609	54.07046	54.55138	55.03820

Prediction is made using Multi-Step Forecast to do the job of forecasting the next five (5) years i.e., 2015-2019 for Nigeria life expectancy data. Table 2 above shows the point forecast output. The forecast from ARIMA (1,2,0) model revealed that the Nigeria life expectancy Nigeria life expectancy is moving up gradually.

4. FINDINGS/SUMMARY

In this research, the Box-Jenkins modelling procedure was discussed to determine an ARIMA model that will be suitable for forecasting. We considered data of the Nigeria life expectancy collected over the period 1980-2014. It was discovered that ARIMA (1, 2, 0) is the most suitable model for the series. The ARIMA regression model was adequate to fit the time series data. The ARIMA model revealed that the Nigeria life expectancy Nigeria life expectancy is moving gradually. Finally, the five years forecast was made from 2015 to 2019 with the highest life expectancy being estimated as 55.038 in

2019.

5. CONCLUSION/RECOMMENDATION

This study has been able to forecast the Nigerian life expectancy. The forecast life expectancy for 2015 – 2019 shows that it would take about five years (2019) to achieve a life expectancy peak of 55.038 of 2019. This does not preclude the possibility of the effect of sound economic policies in rescuing the life expectancy from a trough to a peak. This is especially if government urgently responds to recessions by adopting expansionary macroeconomic policies, such as increasing money supply, increasing government spending and decreasing taxation. Having had a critical study of the life expectancy in the Nigeria, we now profess some recommendations to the government.

Time series analysis and forecasting is an effective statistical tool that has extensive applications in variety of areas including business, economics, government, social science, environmental science, medicine, politics, finance, etc. We therefore recommend the following:
Relevant policy instruments be put in place to enhance life expectancy through the creation of favourable socioeconomic environment. This can be achieved by effective manipulation of the relevant policy instruments such as redistribution of income, employment drive, and diversification of the economy away from oil dependent. These are necessary and highly important in actualizing the 70-year Life expectancy objective of Nigeria

REFERENCES

1. Box, G. E. P.; Jenkins, G. M. (1976) Time Series Analysis: Forecasting and Control. San Francisco. Holden Day.

2. Jie Z, Zhang J and Lee, R (2001) Mortality decline and long run economic growth. J Public Econ, 80:485–507
3. Ruiz-Aguilar, J.J, Ignacio Turias, I., Moscoso-López, J.A., Jiménez-Come, J.A., & Cerbáne, M. (2015). Forecasting of short-term flow freight congestion: A study case of Algeciras Bay Port (Spain). Universidad Nacional de Colombia. DYNA 83 (195), pp. 163-172. Online DOI: http://dx.doi.org/10.15446/dyna.v83n195.47027

4. Salau, M.O. (1998). ARIMA Modelling of Nigeria's Crude Oil Exports. AMSE, Modelling, Measurement & Control, Vol. 18, No. 1, 1 –20.

5. Shafiee, S. & Topal, E. (2010) An overview of global gold market and gold price forecasting. Resources Policy, 35, 178-189.
6. Sede, P., & Ohemeng, W. (2015). Socio-economic determinants of life expectancy in Nigeria (1980– 2011). Health Economics Review, 5(2).
7. Touili, W., Tounmi, H., and Issaoui, F (2015) Effects of CO2 emissions on economic growth, urbanization and welfare. Retrieved from https://mpra.ub.munchen.de/65683