

# Comparing Weighted Markov Chain and Auto-Regressive Integrated Moving Average in the Prediction of Under-5 Mortality Annual Closing Rates in Nigeria

Phillips Edomwonyi Obasohan

Correspondence: School of Health and Related Research (SchARR), The University of Sheffield, Sheffield, United Kingdom; Department of Liberal Studies, College of Administration and Business Studies, Niger State Polytechnic, Bida, Niger State, Nigeria. E-mail: philiobas@yahoo.com

Received: January 29, 2020 Accepted: March 9, 2020 Online Published: March 25, 2020

doi:10.5539/ijsp.v9n3p13

URL: <https://doi.org/10.5539/ijsp.v9n3p13>

## Abstract

In developing countries, childhood mortality rates are not only affected by socioeconomic, demographic, and health variables, but also vary across regions. Correctly predicting childhood mortality rate trends can provide a clearer understanding for health policy formulation to reduce mortality. This paper describes and compares two prediction methods: Weighted Markov Chain Model (WMC) and Autoregressive Integrated Moving Average (ARIMA) in order to establish which method can better predict the annual child mortality rate in Nigeria. The data for the study were Childhood Mortality Annual Closing Rates (CMACR) data for Nigeria from 1964-2017. The CMACR provides random values changing over time (annually), so we can analyze the mortality closing rate and predict the change range in the next state. Weighted Markov Chain (WMC), a method based on Markov theory, addresses the state and its transition procedures to describe a changing random time series. While the Autoregressive Integrated Moving Average (ARIMA) is a generalization of an Autoregressive Moving Average (ARMA) model. The findings indicate that the ARIMA model predicts CMACR for Nigeria better than WMC. The WMC entered in a loop after two iterations, and we could not use it effectively to predict the future values of CMACR.

**Keywords:** weighted Markov chain, autoregressive integrated moving average, under-five mortality rates, forecasting, prediction

## 1. Introduction

Childhood mortality reflects Social Economic Status (SES) and Quality of Life (QoL). In developing countries, mortality rates are not only affected by social, economic, demographic, and health variables, but also vary across regions and districts (Odimegwu & Mkwanzani, 2016). Globally, nearly 7 million children die each year before they reach their fifth birthday. While India (24%) and Nigeria (11%) together account for more than one-third of all under-five deaths (UNICEF, 2014; You et al., 2011). Nigeria Demographic and Health Survey (NDHS) reports indicated that the under-5 mortality rate decreased from 201 deaths per 1,000 live births in 2003 to 128 deaths per 1,000 live births in 2013 (National Population Commission (NPC) & ICF International, 2004, 2014). However, by 2015, Nigeria could not achieve the millennium development goal (MDG) target of reducing under-5 mortality to 64 deaths per 1,000 live births and infant mortality to 30 deaths per 1,000 live births (Government of Federal Republic of Nigeria, 2010). The estimated infant mortality rate in 2003 was 100 deaths per 1,000 live births (National Population Commission (NPC) & ICF International, 2004), which decreased to 75 deaths per 1,000 live births in 2008 (National Population Commission (NPC) & ICF International, 2009) and further to 69 deaths per 1,000 live births in 2013 (National Population Commission (NPC) & ICF International, 2014). The pattern shows that about 29 per cent of deaths in children aged under five occur during the neonatal period (37 deaths per 1,000 live births), and 24 per cent occur during the post-neonatal period (31 deaths per 1,000 live births) (National Population Commission (NPC) & ICF International, 2004). Just as important as childhood mortality is to the economic status of a country, so are many other childhood morbidities. Accurate prediction of these future patterns is of great value to these countries to plan to alleviate the burden of these health conditions. Many new prediction models are available now in the literature that may provide more accurate prediction of time series data. However, model performance still needs to be compared and contrasted. One such model is the Weighted Markov Chain (WMC). As enumerated above, it is clear that Nigeria, has experienced a gradual decline in Child Mortality Annual Closing Rates (CMACR), but the extent to which this will continue to decline into the future and the model which can make this prediction better is not very clear. The aims of

this paper, therefore, are to describe and compare two statistical models (Weighted Markov Chain (a relatively new model) and Auto-regressive Integrated Moving Average (an established model)) for predicting time series data using World Bank dataset of annual childhood (aged<5) mortality rates in Nigeria between 1964 and 2017 (The World Bank, 2018). Of importance in the dataset is that there are elements of seasonality and trends in it. Our focus for both models is to forecast the possibility of a future state of under-5 mortality in Nigeria based on historical data of childhood mortality. However, both methods have different approaches to transforming the value of mortality to the range of dynamic variables.

## 2. The Theoretical Frame

### 2.1 The Weighted Markov Chain (WMC)

WMC method is based on Markov theory and mainly studies the state and its transition procedures, describing a dynamic change process of a random time series (Zhou, 2014). The difference between WMC and the traditional Markov Chain (MC) prediction method is in the weight computed for the initial state and that it is less dependent on historical data (Zhou, 2014). Researchers using WMC believe that predicted values derived from it, the accuracy and practicability are very significant, and as such recommend its study and promotion (Peng et al., 2010; Zhou, 2014). While using this model for prediction in previous studies, some researchers used it solely, and some in combination with other mathematical models (such as fuzzy mathematical models, linear time series models) (Peng et al., 2010; Zhou, 2014). An earlier study has reported that WMC was able to forecast Chinese stock closing prices (Zhou, 2014). The result obtained was closer to the actual value when compared with the result derived using the traditional Markov Chain model (Zhou, 2014). Also, in another study, while predicting the exchange rate ratio between the US Dollar and the Euro, the US Dollar, and the Swiss Franc, the prediction errors using WMC were as low as between 1.4% and 3.0% (Kordnoori et al., 2015). However, Sadeghifar et al. (Shahdoust et al., 2015) used WMC and compared it with two other time series models including Holt Exponential Smoothing (HES) and Seasonal Auto-regressive Integrated Moving Average (ARIMA) in the prediction of hepatitis B (HB) monthly incidence rates in Hamadan Province, western Iran, and the result showed that HES most accurately predicted of HB incidence rates (Shahdoust et al., 2015).

### 2.2 Auto-regressive Integrated Moving Average (ARIMA)

The ARIMA model, also known as the Box-Jenkins model, is widely regarded as the most efficient forecasting technique in social science and is used extensively for time series analysis (Adebiyi et al., 2014). ARIMA is a generalization of an ARMA model whose basic assumption is that the time series is stationary. If that is not the case, Box and Jenkins proposed differencing it to achieve stationarity by incorporating 'I', hence 'Integrated' (Opore, 2015). So, ARIMA is a model built to serve as a basis of standard structures in time-series data, and as a result, provides a simple yet powerful method for making significant time-series predictions (Adebiyi et al., 2014). We can use the ARIMA model in a situation where data exhibit evidence of non-stationarity and where initial differencing steps correspond to eliminate the non-stationarity (Chetty & Narang, 2017a; Katchova, 2013). Researchers have also used ARIMA in the future forecast of disease burden. In China, when used on the incidence of Haemorrhagic Fever with Renal Syndrome (HFRS), it fit the fluctuations in HFRS frequency (Adebiyi et al., 2014). On animal infectious diseases, the predicted incidence of the ARIMA model was consistent with the actual incidence of Newcastle disease (Liu et al., 2011).

The focus of this present study is not to establish that these models (WMC and ARIMA) can predict under-5 mortality rate in Nigeria better than the models currently being used by UNICEF and its allies (The World Bank, 2018), but rather to identify the strengths and weaknesses of this relatively new prediction model (WMC) in comparison with another well-established time series prediction model (ARIMA) on a data set that is presumed to be seasonal and non-stationary. Given the above, the study hypothesized that the Weighted Markov Chain (WMC) model predicts the time series data more accurately than Auto-regressive Integrated Moving Average (ARIMA) while using World Bank Child Mortality Annual Closing Rates (CMACR) in Nigeria.

## 3. Data and Methods

### 3.1 Data

The data used for this study are from UNICEF/World Bank annual report for Nigeria Under-5 mortality (U5M) from 1964 to 2017 (UNICEF, 2018). The Child Mortality Annual Closing Rate (CMACR) of under-5 is a random value changing over time (annually). It is the probability per 1000 live birth that a new-born child will die before the fifth birthday (The World Bank, 2018).

$$\text{Mortality Rates} = \frac{\text{No of death of children aged<5 in the yr}}{\text{r pop estimates of children aged<5 in the yr}} \times 1000 \quad (1)$$

The UNICEF/World Bank sourced mortality data from countries specific vital registration systems and estimates based on sample surveys and censuses (The World Bank, 2018). The computations used all the available parameters that reconcile differences across countries to be able to make comparisons (The World Bank, 2018).

Table 1 represents the data set of annual median (The World Bank, 2018; UNICEF, 2018) estimates for Nigeria under-5 mortality rates from 1964 to 2017. The data witnessed a sharp, steady decline from 1964 to 1982. It began to rise slowly from 1984 (with a mortality rate of 207.7) to 1989 (with a mortality rate of 211.1).

Table 1. Under-5 Mortality Rates for Nigeria per 1000 live births (1964-2017)

Years	Rates	Years	Rates	Years	Rates	Years	Rates
1964	324.8	1978	223.1	1992	210.8	2006	151.5
1965	318	1979	217.8	1993	209.9	2007	145.7
1966	311.2	1980	213.7	1994	208.8	2008	140.1
1967	304.7	1981	210.6	1995	206.9	2009	134.7
1968	297.9	1982	208.8	1996	204.3	2010	129.6
1969	291	1983	208	1997	200.7	2011	124.7
1970	283.7	1984	208.1	1998	196.4	2012	119.9
1971	275.9	1985	208.8	1999	191.5	2013	115.6
1972	267.9	1986	209.9	2000	186.2	2014	111.6
1973	259.7	1987	210.8	2001	180.6	2015	107.5
1974	251.5	1988	211.6	2002	174.9	2016	103.8
1975	243.6	1989	212	2003	169.1	2017	100.2
1976	236.1	1990	211.9	2004	163.2		
1977	229.2	1991	211.4	2005	157.3		

Extracted from the graph: <https://data.unicef.org/country/nga/>(UNICEF, 2018)

The table also revealed that Nigeria has continued to witness a steady decline in the median estimates of under-5 deaths from 1990 to 2017. Between 1990 and 1999, the decline was slower However between 2000 and 2015, the period of millennium development goals was in operations, the under-5 deaths declined sharply. So, from the data, we can analyze the mortality closing rate and predict the change range in the next state. However, 50 annual reports from 1964 to 2013 were purportedly selected and used in the analysis to derive the WMC and the ARIMA models, while we used 2014 to 2017 (Zhou, 2014) for confirmation of prediction accuracy.

### 3.2 Methods

#### 3.2.1 Weighted Markov Chain

The approach adopted here is as proposed in Peng et al. (Peng et al., 2010) modified and used in (Kordnoori et al., 2015; Shahdoust et al., 2015; Zhou, 2014). The steps were as follows:

- i. We computed the number of clusters (i.e. the number of data points that can be combined for similarities) to divide the data into, using k-mean clustering methods. We can also use the mean and mean square error of the data values (Zhou, 2014).
- ii. Divided the data values  $\{X_n\} = X_1, X_2 \dots X_n$  into the chosen number of clusters (say  $n$ ), such that  $f_{ij}$ , are the elements in a shift matrix starting with cluster  $i$  moving to cluster  $j$ .
- iii. In order to make use of WMC, we carry our ‘Markov property’ verification of the known data time series (Zhou, 2014).

The marginal probability estimates, which is the marginal frequency (sum of elements along each column in the shift matrix) divided by the total, are determined by:

$$p_i = \frac{\sum_{j=1}^n f_{ij}}{\sum_i \sum_j^n f_{ij}} \tag{2}$$

The test statistic for large  $n$ , follows a chi-square distribution with  $(n-1)^2$  degree of freedom is given as:

$$\chi^2 = 2 \sum \sum f_{ij} \left| \ln \frac{p_{ij}}{p_i} \right| \tag{3}$$

Where  $p_{ij}$  is the transition matrix from the cluster,  $i$  to cluster  $j$ . The data values  $\{X_n\}$  is said to conform to the Markov property, provided  $\chi^2 = \chi^2_{\alpha, (n-1)^2}$ .

- iv. Compute each order autocorrelation coefficient  $r_k$  of the data values  $\{X_n\}$  and the corresponding weight  $w_k$  of Markov Chain for the various steps using:

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \tag{4}$$

and

$$w_k = \frac{|r_k|}{\sum_{k=1}^m |r_k|} \tag{5}$$

where  $m$  is the maximum order of prediction inquiry;  $r_k$  is the  $k$  order autocorrelation coefficient, such that  $|r_k| \geq 0.3$ ;

$x_i$  is the CMACR for the  $i^{\text{th}}$  year;  $\bar{x}$  is the mean of data values  $\{X_n\}$ ;

$n$  is the length of the reference sample data values;

$w_k$  is the weight of Markov Chain with a step of  $k$  orders.

- v. The Prediction of the CMACR Range.

Here, we combine the initial state  $S_i$  as the corresponding state of CMACR in the one-year past with the row vector of its corresponding transition probability matrix results in state transition probability vector in the year as follows (Kordnoori et al., 2015):

$$p_i^{(k)} = (p_{i1}^{(k)}, p_{i2}^{(k)}, \dots, p_{im}^{(k)}), \quad i \in E \tag{6}$$

We obtain the  $m$ -order weighted state transition probability matrix as follows:

$$\begin{pmatrix} p_{\alpha 1}^{(k)} & p_{\alpha 2}^{(k)} & \dots & p_{\alpha m}^{(k)} \\ \vdots & \ddots & \ddots & \vdots \\ p_{\beta 1}^{(k)} & p_{\beta 2}^{(k)} & \dots & p_{\beta m}^{(k)} \end{pmatrix} \alpha, \dots, \beta \in S, \alpha, \dots, \beta \leq m \tag{7}$$

Then, predict the state of the data values  $\{X_n\}$  in the future using the data values  $\{X_n\}$  for 2012 and 2013 as the initial state ( $S$ ) and the corresponding transition probability matrix to predict the data values  $\{X_n\}$  for 2014 using the formula:

$$p_i = \sum_{k=1}^n w_k p_i^{(k)}, \quad i \in E \tag{8}$$

The forecasted state of data values  $\{X_n\}$  by WMC is  $\max \{P_i, i \in S\}$ . We used R-Statistical Programming Package version 3.5.3 (R Version 3.5.3, 2019) for the analysis of WMC,.

### 3.2.2 Auto-regressive Integrated Moving Average (ARIMA)

ARIMA model has a unique condition that the time series data are either stationary or can be transformed into it (Chetty & Narang, 2017a; Opare, 2015). The prediction equation is a linear regression such that the independent variables are the lags of the dependent variable including/or lags of error terms (Nau, 2014).

ARIMA is simply the ‘Integrated series (I)’ of two lags series: the stationary part known as ‘auto-regressive (AR)’ and the forecast errors part known as ‘moving average (MA)’. When an ARMA ( $p, q$ ) process carries a  $d$ -order differencing, it is then known as ARIMA( $p, d, q$ ) (Zaiontz, 2012). The  $p$  refers to the number of autoregressive terms,  $d$  as the degree of differencing, while  $q$  refers to the number of moving average terms. The mathematical formulation of the ARIMA model is well established in the literature (Nau, 2014). In general, with lag-operator set as  $B$ , ARIMA ( $p, d, q$ ) is given as (Deljac et al., 2011):

$$(1 - \sum_{i=1}^p \phi_i B^i)(1 - B)^d Y_t = (1 - \sum_{i=1}^q \theta_i B^i) \varepsilon_t \tag{9}$$

In order to use ARIMA, the model requires that we transform the original series into stationary data through differencing at different levels.

The steps involved in formulating ARIMA model, as stated in (Opare, 2015) include:

1. **Model Identification:** We carried out a visual inspection on the graphical plot to confirm the stationarity in the series. Stationarity means the existence of constant mean, and variance
2. **Model Estimation:** We used Augmented Dickey-Fuller (ADF) to test the null hypothesis that the data has a unit root and, therefore, non-stationary (Opare, 2015). ADF usually will check the correlation in error terms by adding lags. Two values are most important to make confirmation:  $Z(t)$  and Mackinnon  $p$ -value for  $Z(t)$  (Chetty & Narang, 2017a). For stationarity,  $Z(t)$  should have a sizeable negative number, and the  $p$ -value should be significant, at least on a 5% level (Chetty & Narang, 2017b; Opare, 2015). To achieve this resolution, we add lags and subsequently taking the differencing (Chetty & Narang, 2017b).

For instance, a series that needs no difference ( $d=0$ ) has the forecasting equation as  $y_t = Y_t$ , where  $Y$  is the original series and  $y$  is the difference series at time  $t$ . Also, for  $d=1$  (first difference):  $y_t = Y_t - Y_{t-1}$ , this addresses linear trends, for  $d=2$  (second difference):  $y_t = Y_t - 2Y_{t-1} + Y_{t-2}$  addresses quadratic trends, this is usually the first-order difference of the first-order difference (Zaiontz, 2012). For  $d = 3$  (third difference):  $y_t = Y_t - 3Y_{t-1} + 3Y_{t-2} - Y_{t-3}$  addresses cubic trends, etc., so that,  $d = n$  ( $n^{\text{th}}$  difference) will address the  $n^{\text{th}}$  - ordered trends. The rule of thumb is that over differencing could introduce into the model large variances that could result in greater errors in the prediction value.

The model estimation also involves the determination of the time lags ( $p, q$ ) for auto-regression (AR) and Moving average (MA). We can achieve it by plotting the correlogram for Autocorrelation Function (ACF) and Partial Auto-Correlation (PAC) for third differencing

- 3. **Model Diagnostic checks:** We carried out a Portmanteau test for white noise to know if the residual variable follows a white noise process (Stata, 2017). Furthermore, we plot the 'fitted' values with the actual values on a two-way graph with standard error as scattered points (Chetty et al., 2018),
- 4. **Forecasting:** In general, the forecast equation for the differencing series is:

$$y'_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} \dots - \theta_q e_{t-q} + \epsilon_t \tag{10}$$

In the situation where differencing has been introduced, there was need to reverse the differencing to obtain the original series. However, the software does it automatically (Nau, 2014).

In general, the transformation is done using this substitution (Opare, 2015):

$$y_t = (1 - B)^d Y_t \tag{11}$$

For a 3<sup>rd</sup> differencing, for instance, the reverse equation will be:

$$Y'_t = y'_t + 3Y_{t-1} - 3Y_{t-2} + Y_{t-3} \tag{12}$$

The chosen ARIMA ( $p, d, q$ ) model was used to forecast values for the period 2014 to 2017. We used Stata 14 SE for academic user (Stata Corporation, 2014) in the computations

### 3.3 Measure of Comparison

The Mean Absolute Percentage Error (MAPE), is widely used statistical measure of forecast accuracy (Stephanie, 2017). The fact that MAPE is computed in percentage makes it easy to interpret (ForecastPRO, 2011). In this study, we apply MAPE as a measure of accuracy by converting the errors between the actuals and the estimates into percentages (Opare, 2015). The lower the MAPE, the more accurate the forecast.

$$MAPE = \left( \frac{1}{n} \sum_{i=1}^n \frac{|Actual - Forecast|}{|Actual|} \right) 100 \tag{13}$$

## 4. Results

### 4.1 WMC Model in Forecasting (CMACR) in Nigeria

Changes in Nigeria CMACR was divided into two states blocks according to k-means cluster analysis as 'rate below' and 'rate above' with the corresponding state space as  $E = \{1, 2\}$ . Table 2 showed the results.

Table 2. State Division Cluster of CMACR

Cluster	interval	Designation
1	$X < 200.6$	Rate Below
2	$x > 200.5$	Rate Above

To verify the Markov property to justify the use of WMC, the chi-square of 68.98 was higher than the table value of 3.84 at 0.05 level of significance, so the annual closing CMACR conformed to the 'Markov property'; therefore WMC prediction theory can be applied (Zhou, 2014).

Table 3. Each Order Autocorrelation Coefficient and the Weight of each Step

<b>k</b>	<b>1</b>	<b>2</b>
$r_k$	0.981	0.952
$w_k$	0.508	0.492

Table 3 displayed the results of the computation of each order autocorrelation coefficient  $r_k$  of the sequence parameter values and the corresponding weights. Table 4 displayed the result of the prediction of CMACR for 2014 using 2012 and 2013 as the initial years

Table 4. The Prediction about CAMCR for 2014

Initial year	state	Retention time	State of weight	1	2
2013	2	1	0.508	1.000	0.000
2012	2	2	0.492	0.059	0.941
$p_i$				0.537	0.463

The result in table 4 showed that  $\max \{p_i, i \in E\} = 0.537$ , indicating that CMACR annual closing rate in 2014 was in cluster 1 (i.e., ‘Rate below’), which met the block interval of  $X < 200.6$  with a probability of 53.7%. From the historical data, the actual closing rate was 111.6 and consistent with the prediction.

Also, we perform similar iteration such that the state in 2014 served as one of the initial states for predicting the annual closing rate in 2015.

Table 5. The Prediction about CMACR for 2015

Initial year	State	Retention time	State of weight	1	2
2014	2	1	0.508	1.000	0.000
2013	2	1	0.492	1.000	0.000
$p_i$				1.000	0.000

The result in table 5 showed that  $\max \{p_i, i \in E\} = 1.000$ , indicating that CMACR closing in 2014 was also in cluster 1 (i.e., ‘Rate below’), which met the block interval of  $X < 200.6$  with a probability of 100%. From the historical data, the actual closing rate was 107.5 and consistent with the prediction.

#### 4.2 ARIMA Model in Forecasting (CMACR) in Nigeria

##### 4.2.1 Model Identification

Fig 1 indicated that there is trending downward during 1964-2013 with minor fluctuations.

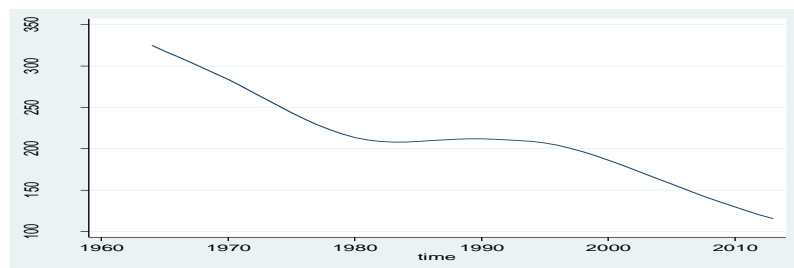


Figure 1. Diagram for CMACR Trend 1964 - 2013

So, CMACR cannot have constant mean and variance; hence, the primary assumption of stationarity cannot be confirmed (Adebiyi et al., 2014; Katchova, 2013; Liu et al., 2011; Opere, 2015).

##### 4.2.2 Augmented Dickey-Fuller (ADF) Test

The results are shown in Table 5 below:

Table 6. The Differencing of CAMCR for 1964 - 2013

S/N	Lags/Differencing	Z(t)	P-values	Remarks
1	Lag=0	-2.03	0.27	Not Significant
2	Lag=10	0.63	0.99	Not Significant
3	Differencing = 1	-0.68	0.97	Not Significant
4	Differencing = 2	-1.44	0.84	Not Significant
5	Differencing = 3	-7.90	<0.0001	Significant
6	Differencing = 4	-19.22	<0.0001	Significant

We observe that from table 6, the third differencing Z(t) value (-7.90) is a sizeable negative & corresponding p-value is also significant (0.0001), so the null hypothesis, which states that the series has a unit root was rejected. Therefore, we considered the third differencing for further studying.

4.2.3 Model Estimation

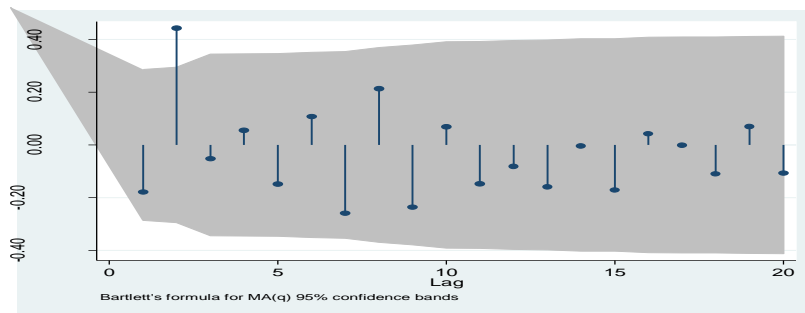


Figure 2. Diagram for ACF

From the diagram in Fig 2, only the 2<sup>nd</sup> lag is outside the shaded portion (acceptable region), so, we set AR at 2. This AR (2) means that the AR model needs only two autoregressive terms (i.e., AR of order 2).

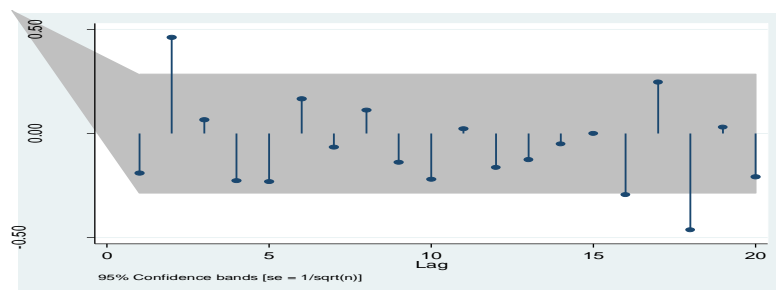


Figure 3. Diagram for PACF

Similarly, in figure 3, for the 3<sup>rd</sup> differencing correlogram of PACF, the 2<sup>nd</sup> lag is coming out of the shaded region, so we set MA at 2. So far we are able to identify only one plausible model (2, 3, 2); therefore, there may be no need to make a choice using Akaike’s Information Criterion (AIC) and Bayesian information criterion (BIC)

4.2.4 Model Adequacy

The portmanteau Q statistic is 26.81 with a p-value of 0.26, so we cannot reject the null hypothesis indicating that the residual variance is indeed a white noise process

Furthermore, after plotting the ‘fitted’ values with the actual values on a two-way graph with standard error as scattered points (Chetty et al., 2018), it can be seen clearly that the predicted values of CAMCR do not differ from the actual values.

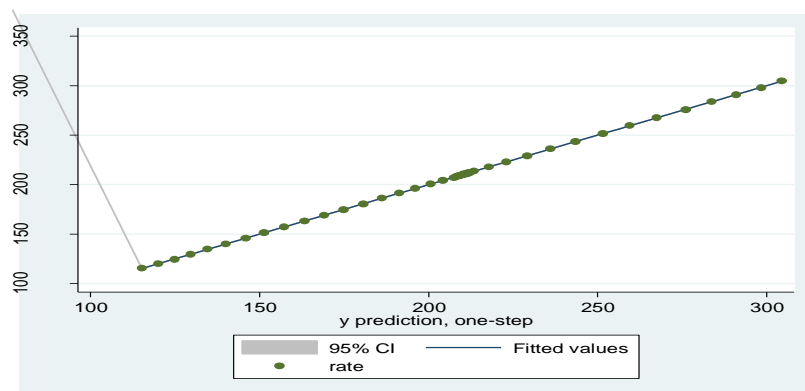


Figure 4. Two-way graph of Actual and Fitted Values of CAMCR

4.2.5 Forecasting

The chosen ARIMA model is used to forecast values for the period 2014 to 2017, as shown on the table below:

Table 6. Comparing the Actual and Forecast Values for CAMCR per 1000 live births per year

Period	2014	2015	2016	2017	.....	2030
CAMCR Actual	111.6	107.5	103.8	100.2	.....	
CAMCR Forecast	111.58	108.18	105.36	103.23	.....	139.43

The Mean Absolute Percentage Error (MAPE) for the forecast (2014 to 2017) is 1.3% signifying that the ARIMA model was off by a yearly average of 1.3%. This error is relatively low and therefore near a perfect fit. However, using the ARIMA model to forecast mortality rate for 2030 resulted in a rate of 139/1000 live births showing a rise in CAMCR for Nigeria.

## 5. Discussion

This study used WMC and ARIMA to predict the CMACR for Nigeria Under-five mortality rates. As for WMC, we classified CMACR into 2 clusters after having examined the performances of other clusters (3,4,5,6,7). The Markov Chain property was checked and found to be adequate to apply WMC. However, because of the nature of the data (non-stationary data), WMC could not produce a reasonable conclusion because it entered into a loop after the second iterations (after predicting for two periods, this causes the predicted values to remain in cluster one. Also, because of the open-ended interval created by using two clusters in WMC, the exact value for the forecasts could not be determined; therefore, accurate prediction becomes difficult. This finding agrees with Sadeghifar et al. (Shahdoust et al., 2015). So, WMC is weak in its ability to model effectively a data set including trends and seasonality. ARIMA, on the other hand, proved to be a better and more straight forward approach. The original data were non-stationary. We used third-order differencing to bring it into stationarity. The chosen model, ARIMA (2, 3, 2) gave a near-perfect fit with MAPE of 1.28%. So, because of the findings above, the hypothesis that WMC predicts CAMCR better than ARIMA was rejected. Therefore, we conclude that ARIMA can predict more accurately the CAMCR in Nigeria than WMC. Interestingly, the predicted values with ARIMA shows that the CMACR will continue to decrease up to 2022, after which it will increase such that the rate as of 2030 will be 139.4. However, the limitation from this finding is that the ARIMA model is weak in its ability to predict for a medium-term and long-term range (Deljac et al., 2011). The reason being that ARIMA convergences toward the mean of the auto-regressive part of the series (Huijskens, 2016), so we cannot rely on the efficacy of the 2030 forecast.

## 6. Recommendations for Future Study and Policy Implementation

More studies on WMC are needed to resolve the pitfall identified in this study. However, the implication here for policy-makers or agencies in charge of monitoring the under-five mortality rate in Nigeria is that, if the current approaches are continued, the country will be unlikely to achieve the Sustainable Development Goals (SDGs) for 2030.

## Acknowledgement

The author acknowledged the valuable contribution of Amini for help with the R-codes to analyze the WMC, Prof Walters, Dr Foley and Dr Lawrence for valuable suggestions to improve the work. We also acknowledge UNICEF for the data used in this study.

## Conflict of Interest

The authors declared no conflict of interest

## Funding

The author declared that no funding was received in respect of this project

## References

- Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, 2014. <https://doi.org/10.1155/2014/614342>
- Chetty, P., & Narang, D. (2017a, December 2). Problem of non-stationarity in time series analysis in STATA. *Knowledge Tank*. Retrieved from <https://www.projectguru.in/publications/problem-non-stationarity-time-series-stata/>



- Chetty, P., & Narang, D. (2017b, December 20). Solution for non-stationarity in time series analysis in STATA. *Knowledge Tank*. Retrieved from <https://www.projectguru.in/publications/solution-non-stationarity-time-series-analysis-stata/>
- Chetty, P., Narang, D., & Choudhury, S. (2018, April 29). How to predict and forecast using ARIMA in STATA? *Knowledge Tank*. Retrieved from <https://www.projectguru.in/publications/prediction-forecasting-using-arima-stata/>
- Deljac, Ž., Kunštić, M., & Spahija, B. (2011). A comparison of traditional forecasting methods for short-term and long-term prediction of faults in the broadband networks. *2011 Proceedings of the 34th International Convention MIPRO*, 517-522.
- ForecastPRO. (2011). *Forecasting 101: A Guide to Forecast Error Measurement Statistics and How to Use Them*. Retrieved from <https://www.forecastpro.com/Trends/forecasting101August2011.html>
- Government of Federal Republic of Nigeria. (2010). MDG Report 2010: Nigeria Millennium Development Goals. *Abuja, Nigeria: Government of Federal Republic of Nigeria*. Retrieved from [https://www.undp.org/content/dam/undp/library/MDG/english/MDG%20Country%20Reports/Nigeria/nigeria\\_2010.pdf](https://www.undp.org/content/dam/undp/library/MDG/english/MDG%20Country%20Reports/Nigeria/nigeria_2010.pdf)
- Huijskens, T. (2016, August 3). *Long-term forecasting with machine learning models*. Thomas Huijskens. Retrieved from <http://thuijskens.github.io/2016/08/03/time-series-forecasting/>
- Katchova, A. (2013). *Time Series ARIMA Models Example—[PDF Document]*. Vdocuments.Mx. Retrieved from <https://vdocuments.mx/time-series-arima-models-example.html>
- Kordnoori, S., Mostafaei, H., & Kordnoori, S. (2015). Applied SCGM(1,1)c Model and Weighted Markov Chain for Exchange Rate Ratios. *Hyperion Economic Journal*, 3(4), 12-22.
- Liu, Q., Liu, X., Jiang, B., & Yang, W. (2011). Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infectious Diseases*, 11(1), 218. <https://doi.org/10.1186/1471-2334-11-218>
- National Population Commission (NPC), & ICF International. (2004). Nigeria Demographic and Health Survey, 2003. *Federal Republic of Nigeria and MeasureDHS*.
- National Population Commission (NPC), & ICF International. (2009). Nigeria Demographic and Health Survey, 2008. *Federal Republic of Nigeria and MeasureDHS*.
- National Population Commission (NPC), & ICF International. (2014, June). Nigeria Demographic and Health Survey, 2013. *Federal Republic of Nigeria and MeasureDHS*.
- Nau, R. (2014). *Lecture notes on forecasting: Introduction to ARIMA models*. Fuqua School of Business, Duke University. Retrieved from <http://people.duke.edu/~rnau/forecasting.htm>
- Nau, R. (2014). *Mathematical structure of ARIMA models*. Fuqua School of Business Duke University.
- Odimegwu, C., & Mkwanzani, S. (2016). Factors Associated with Teen Pregnancy in sub-Saharan Africa: A Multi-Country Cross-Sectional Study. *African Journal of Reproductive Health*, 20(3), 94-107. <https://doi.org/10.29063/ajrh2016/v20i3.14>
- Opare, P. E. (2015). *Time Series Models for the Decrease In Under-Five Mortality Rate in Ghana Case Study 1961—2012* [Department of Mathematics, Kwame Nkrumah University of Science and Technology, Ghana]. Retrieved from <http://ir.knust.edu.gh/bitstream/123456789/6807/1/Prince%20Edward%20Opare.pdf>
- Peng, Z., Bao, C., Zhao, Y., Yi, H., Xia, L., Yu, H., Shen, H., & Chen, F. (2010). Weighted Markov chains for forecasting and analysis in incidence of infectious diseases in Jiangsu Province, China. *Journal of Biomedical Research*, 24(3), 207-214. [https://doi.org/10.1016/S1674-8301\(10\)60030-9](https://doi.org/10.1016/S1674-8301(10)60030-9)
- R Version 3.5.3. (2019). *The R Foundation for Statistical Computing Platform: (Version 3.5.3)* [Computer software].
- Shahdoust, M., Sadeghifar, M., Poorolajal, J., Javanrooh, N., & Amini, P. (2015). Predicting hepatitis B monthly incidence rates using weighted Markov chains and time series methods. *Journal of Research in Health Sciences*, 15(1), 28-31.
- Stata. (2017). *White noise test interpretation*. Cross Validated. Retrieved from <https://stats.stackexchange.com/questions/124346/white-noise-test-interpretation>
- Stata Corporation, 14SE. (2014). *Stata Statistical Software, College Station, TX (14SE)* [Computer software].

- Stephanie. (2017, September 8). *Mean absolute percentage error (MAPE)*. Statistics How To. Retrieved from <https://www.statisticshowto.datasciencecentral.com/mean-absolute-percentage-error-mape/>
- The World Bank. (2018). *Mortality rate, under-5 (per 1,000 live births)—Nigeria | Data*. Retrieved from <https://data.worldbank.org/indicator/SH.DYN.MORT?locations=NG>
- UNICEF (Ed.). (2014). *Every child counts: Revealing disparities, advancing children's rights*. UNICEF.
- UNICEF. (2018). *Key demographic indicators for Nigeria: Under-Five Mortality Rate: UNICEF DATA*. Retrieved from <https://data.unicef.org/country/nga/>
- You, D., New, J. R., & Wardlaw, T. (2011). *Levels & Trends in Child Mortality Report 2012. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation*. Retrieved from [http://www.who.int/maternal\\_child\\_adolescent/documents/levels\\_trends\\_child\\_mortality\\_2012.pdf](http://www.who.int/maternal_child_adolescent/documents/levels_trends_child_mortality_2012.pdf)
- Zaiontz, C. (2012, 2019). *ARIMA Differencing | Real Statistics Using Excel*. Retrieved from <http://www.real-statistics.com/time-series-analysis/arima-processes/arima-differencing/>
- Zhou, Q. (2014). *Application of Weighted Markov Chain in Stock Price Forecasting*. <https://doi.org/10.14257/astl.2014.53.50>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).