

# Application of Autoregressive Integrated Moving Average Model and Weighted Markov Chains on Forecasting Under-Five Mortality Rates in Nigeria

**Abstract:** The aim of this paper is to obtain the best model that will be used to predict Under-Five Mortality Rate (U5MR) between Autoregressive Integrated Moving Average (ARIMA) model and Weighted Markov Chains (WMC). The annual dataset of U5MR in Nigeria for the period 1980-2019 is obtained from the official website of World Bank. The descriptive statistics and the unit root test for the stationarity of data were carried on the data series. ARIMA was modelled to U5MR using the techniques of Box-Jenkins while WMC was modelled using the techniques of k-means cluster analysis, Chi-Square, and Correlation. The best ARIMA model was obtained using Bayesian Information Criterion (BIC) while the best forecast model was obtained using Theil's U Statistics and Mean Absolute Percentage Error (MAPE). U5MR attained stationarity after third differencing under ARIMA model dynamics. ARIMA(0,3,2) is considered the best ARIMA model with BIC of -2.679, and was selected as the best forecast model with Theil's U Statistic of 0.000014 and MAPE of 0.174336%. The fitted model was used to make out-sample forecast for the period 2020-2030, which showed a steady decline. The findings of this paper will help in establishment and implementation health policies.

**Keywords:** U5MR, ARIMA, Weighted Markov Chain (WMC), MAPE, Theil's U Statistic, K-Mean Cluster

## 1. Introduction

Under-Five Mortality Rate is the probability of a child dying between birth and exactly 5 years of age, expressed per 1,000 live births in a given year for a particular geographical area [4,13]. In developing countries, childhood mortality rate is affected by socioeconomic, demographic, health variable [9], but also varies across regions and districts [14]. Nigeria and other countries in Sub Saharan Africa though experienced a decline in U5MR from 1980 to 2019, still maintain relatively and unacceptable high Mortality compared to many countries in Europe and America [2,3,8]. [1] compared the effect of ARIMA, Artificial Neural Networks, and Exponential Smoothing. [5] studied the U5MR of Malaysia by gender and developed a forecasting model for future prediction; the result showed that the U5MR for both genders decreased slowly. [9] analyzed the Under-5 mortality annual closing rate (CMACR) in Nigeria using Weighted Markov Chain and ARIMA model, the findings showed that ARIMA predicts CMACR better than WMC.

This paper attempts to establish an adequate model using the Autoregressive Integrated Moving Average (ARIMA) and Weighted Markov Chain (WMC) to forecast the Under-five Mortality Rate in Nigeria.

## 2. Materials and Methods

### 2.1 Autoregressive Integrated Moving Average (ARIMA) Model

ARIMA is a statistical model which is used to predict future values based on past values. The 'AR' stands for Autoregressive, 'MA' stands for Moving Average, and 'I' stands for Integrated (that is the data values are replaced by difference between the data values and the previous values). ARIMA model is denoted by  $ARIMA(p, d, q)$  and it is written as

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (1)$$

where  $\phi_1, \phi_2, \dots, \phi_p$  are Autoregressive model's parameters;  $\theta_1, \theta_2, \dots, \theta_q$  are Moving Average model's parameters;  $c$  is a constant;  $\varepsilon_t$  is a white noise, and  $y'_t$  is the differenced series which might been differenced more than once

#### 2.1.1 Autoregressive Moving Average (ARMA) Model

When the time series data is stationary and however does not require differencing, then the resultant model is an Autoregressive Moving Average (ARMA) model. ARMA model is denoted by  $ARMA(p, q)$  and it is written as

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

### 2.1.2 Autoregressive (AR) Model

AR model is the regression of the current observations against one or more past observations. That is the current observation  $y_t$  are generated by a weighted averages of past time series data going back  $p$  periods, together with a random disturbance in the current period. The AR of order  $p$  denoted by  $AR(p)$  is defined as

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (3)$$

Where  $\varepsilon_t$  is a white noise;  $\phi_1, \phi_2, \dots, \phi_p$  are the parameters of the AR model;  $y_t$  is the current observation,  $y_{t-1}, y_{t-2}, \dots, y_{t-p}$  are past observations.

### 2.1.3 Moving Average (MA) Model

MA is a linear combination of error terms occurring at various times in the past. MA model of order  $q$  is denoted as  $MA(q)$  and it is written as

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (4)$$

## 2.2 ARIMA Fitting

ARIMA model is fitted to the time series data (historical data) using the Box-Jenkins method. Four (4) steps are employed here, which are: Model Identification using ACF and PACF, Estimation of Parameters, Model Adequacy Check, and Forecasting

### 2.2.1 Model Identification using ACF and PACF Plots

- If the Autocorrelation Function (ACF) drops to zero relatively quickly as the number of lags increases, the time series data is stationary
- If the Autocorrelation Function (ACF) drops very slowly as the number of lags increases, the time series data is not stationary
- If there is presence of a unit root in the time series data, then the time series is not stationary. This study adopts Augmented Dickey-Fuller (ADF) Test for Unit Root Test

If the time series data fails to be stationary, then, differencing the data series at least once will bring the data to stationarity.

Differencing is a process that stabilizes the mean of time series by removing the changes in the series and eliminating or reducing trend and seasonality.

First Order Differenced series denoted as  $y'_t$  is the change between consecutive observations in the original series. It is written as

$$y'_t = y_t - y_{t-1} \quad (5)$$

If the first differenced series fails to be stationary, there is need to carry out second differencing

Second Order Differenced series denoted as  $y''_t$  is written as

$$y''_t = y'_t - y'_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2} \quad (6)$$

The general  $d$ th order difference using the Backshift Operator  $B$ , is written as

$$y_t^d = (1 - B)^d y_t \quad (7)$$

## 2.2.2 Estimation of Parameters using Ordinary Least Squares Method

The Yule Walker's Simultaneous Equations given in equation (8) helps in obtaining the parameters of Autoregressive model

$$\left. \begin{aligned} \hat{\rho}_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \phi_4\rho_3 + \dots + \phi_n\rho_{n-1} \\ \hat{\rho}_2 &= \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 + \phi_4\rho_2 + \dots + \phi_n\rho_{n-2} \\ \hat{\rho}_3 &= \phi_1\rho_2 + \phi_2\rho_1 + \phi_3 + \phi_4\rho_1 + \dots + \phi_n\rho_{n-3} \\ &\vdots \\ \hat{\rho}_k &= \phi_1\rho_{n-1} + \phi_2\rho_{n-2} + \phi_3\rho_{n-3} + \dots + \phi_n \end{aligned} \right\} \quad (8)$$

In the matrix notation, these equations can be written as  $\rho = \phi\rho^*$

where  $\rho$  is a  $k \times k$  matrix;  $\phi$  is a  $k \times 1$  matrix, and  $\rho^*$   $k \times 1$  matrix

$$\rho = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \rho_2 & \dots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \rho_1 & \dots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \rho_{n-4} & \dots & 1 \end{pmatrix}, \quad \phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_k \end{pmatrix} \quad \text{and} \quad \rho^* = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{pmatrix} \quad (9)$$

The parameters of Autoregressive (AR) Model is obtained using  $\phi_k = \frac{\rho_k^*}{\rho_k}$ , where

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)} = \frac{\gamma_k}{\gamma_0}, \quad \text{var}(y_t) = \gamma_0 = \sigma^2 = \frac{\sum(y_t - \bar{y}_t)^2}{N} \quad \text{and} \quad \text{cov}(y_t, y_{t-k}) = \gamma_k = \frac{\sum(y_t - \bar{y}_t)(y_{t-k} - \bar{y}_t)}{N} \quad (10)$$

and  $\rho_k^*$  is obtained by replacing the last column of matrix  $\rho$  by  $\begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{pmatrix}$ , the first two parameters  $\hat{\phi}_1$  and  $\hat{\phi}_2$  of AR are expressed as

$$\hat{\phi}_1 = \frac{\rho_1^*}{\rho_1} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \quad \text{and} \quad \hat{\phi}_2 = \frac{\rho_2^*}{\rho_2} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \quad (11)$$

The parameters of Moving Average (MA) Model is obtained using  $\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)} = \frac{\gamma_k}{\gamma_0}$ , where

$$\text{var}(y_t) = \gamma_0 = \sigma^2(1 + \theta_1^2 + \dots + \theta_q^2) \quad (12)$$

$$\text{cov}(y_t, y_{t-k}) = \gamma_k = \sigma^2(-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q) \quad (13)$$

For MA(1),  $\rho_k = 0$ , for  $k > 1$  and under MA(2),  $\rho_k = 0$ , for  $k > 2$

## 2.2.3 Model Adequacy Check

The Akaike Information Criterion (AIC) and/or Bayesian Information Criterion (BIC) is/are used to check for model adequacy. The AIC is written as

$$AIC = n \log(\hat{\sigma}^2) + 2k \quad (14)$$

$k$  is the number of model parameters;  $\hat{\sigma}^2$  is the residual sum of squares, and  $n$  is the sample size  
Bayesian Information Criterion (BIC) is written as

$$BIC = n \log(\hat{\sigma}^2) + k \log(n) \quad (15)$$

The ARIMA or ARMA model with the lowest AIC and/or BIC are/is considered the best model among others.

## 2.2.4 Forecast Equation

The estimated ARIMA model used in forecasting time series data is expressed as

$$F_t = c + \hat{\phi}_1 y_{t-1} + \dots + \hat{\phi}_p y_{t-p} + \hat{\theta}_1 \varepsilon_{t-1} + \dots + \hat{\theta}_q \varepsilon_{t-q} \quad (16)$$

## 2.3 Weighted Markov Chain (WMC)

### 2.3.1 Markov Chain

Markov chain is a stochastic process  $X_t, t = 0, 1, 2, \dots$  having the property that given the present state of the system, the past and the future are conditionally independent.

Whenever the process is in state  $i$ , there is a fixed probability  $P_{ij}$  that it will be in state  $j$  next, then the property of Markov chains is define as

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P_{ij} \quad (17)$$

for all states  $i_0, i_1, \dots, i_{n-1}, i, j$  and all  $n \geq 0$

where  $j$  is the future state,  $i$  is the present or current state,  $i_{n-1}, i_{n-2}, \dots, i_1, i_0$  are the past states

### 2.3.2 Weighted Markov Chain

The method of Weighted Markov Chain adopted in this study for forecasting the U5MR is the one expressed by [6,7,10-12], which is categorized into seven (7) steps:

- Set up a classification standard for the historical data, Under-five Mortality Rate (U5MR) using the K-means Cluster Analysis
- Determine the  $m$  states, that is, the states of the historical data (U5MR) according to the classification standard
- Obtain the Frequency Matrix (or Transition Matrix) according to step 2

$$F = \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1m} \\ f_{21} & f_{22} & \dots & f_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m1} & f_{m2} & \dots & f_{mm} \end{pmatrix} \quad (18)$$

- Obtain the One-step Transition Probability Matrix  $P = (p_{ij})$  and the Marginal Matrix  $Q = (q_i)$  using the matrix of step 3, where

$$p_{ij} = \frac{f_{ij}}{\sum_{j=1}^m f_{ij}} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix} \quad (19)$$

$$q_{ij} = \frac{f_{ij}}{\sum_{i=1}^m \sum_{j=1}^m f_{ij}} = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1m} \\ q_{21} & q_{22} & \cdots & q_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ q_{m1} & q_{m2} & \cdots & q_{mm} \end{pmatrix} \quad (20)$$

Hence the Marginal Matrix

$$q_i = \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_m \end{pmatrix} \quad (21)$$

where  $q_1 = q_{11} + q_{12} + \cdots + q_{1m}$ ;  $q_2 = q_{21} + q_{22} + \cdots + q_{2m}$  up to  $q_m = q_{m1} + q_{m2} + \cdots + q_{mm}$

- e. Test whether the Transition Probability Matrix obtained in step 4 has a Markov Property, using Chi-square test, defined as

$$\chi^2 = 2 \sum_{i=1}^m \sum_{j=1}^m f_{ij} \left| \ln \frac{p_{ij}}{q_i} \right| \quad (22)$$

with degree of freedom  $(m-1)^2$ . The stochastic process (Transition Probability Matrix) has a Markov Property if  $\chi^2$  is greater than  $\chi_{\alpha, (m-1)^2}^2$

- f. Obtain the weight of the various steps Markov Chain,  $w_k$  transition probabilities matrices, which is defined as

$$w_k = \frac{|r_k|}{\sum_{k=1}^K r_k} \quad (23)$$

where  $r_k$  is the autocorrelation coefficient of the historical data with  $k \in \{1, 2, \dots, K\}$ , and is computed as

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (24)$$

where  $y_t$  is the time series at time  $t$ ,  $\bar{y}$  is the average of time series  $y_t$ , and  $n$  is the number of time series  $y_t$

- g. Take the weighted average of various predicting probabilities of the same state as predicting probabilities of the U5MR, defined as

$$\hat{p}_{ij} = \sum_{k=1}^K w_k p_{ij}^{(k)} \quad (25)$$

for every  $j \in \{1, 2, \dots, m\}$ ,  $\hat{p}_{ij}$  is the probability for a time series  $y_t$  to be in the state  $j$  in the future.

The forecast result is in the form of a state, state  $j$  obtained by  $arg \max\{\hat{p}_{ij}, j = 1, 2, \dots, m\}$

## 2.4 Forecast Adequacy Check

The measures of forecast accuracy adopted in this study is Theil's U Statistic and Mean Absolute Percentage Error (MAPE).

### 2.4.1 Theil's U Forecast Accuracy

The Theil's U shows how the forecast conforms to the values of the future periods. It is written as

$$U = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n y_t^2 + \frac{1}{n} \sum_{t=1}^n \hat{y}_t^2}} \quad (26)$$

where  $Y_t$  is the actual value of a point for a given time period  $t$ ,  $\hat{Y}_t$  is the forecast value,  $n$  is the number of the data points.

If  $U$  falls within the range  $0 \leq U < 1$ , the proposed model is a good fit

If  $U = 0$ , the proposed model is a perfect fit

If  $U \geq 1$ , the proposed model is not a good fit

### 2.4.2 Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) is used to measure the error of both methods (ARIMA and WMC). The model with the smallest MAPE is considered the appropriate model. It is defined as

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (27)$$

## 3. Results/Findings

Figure 1 shows the timeplot for Under-five Mortality Rate in Nigeria for the period 1980-2019. U5MR in Nigeria shows a partial decrease over the years. In Table 1, the average Under-Five Mortality Rate (U5MR) is 174.315 deaths per 1000 live births, and standard deviation of 34.6777 deaths per 1000 live births.

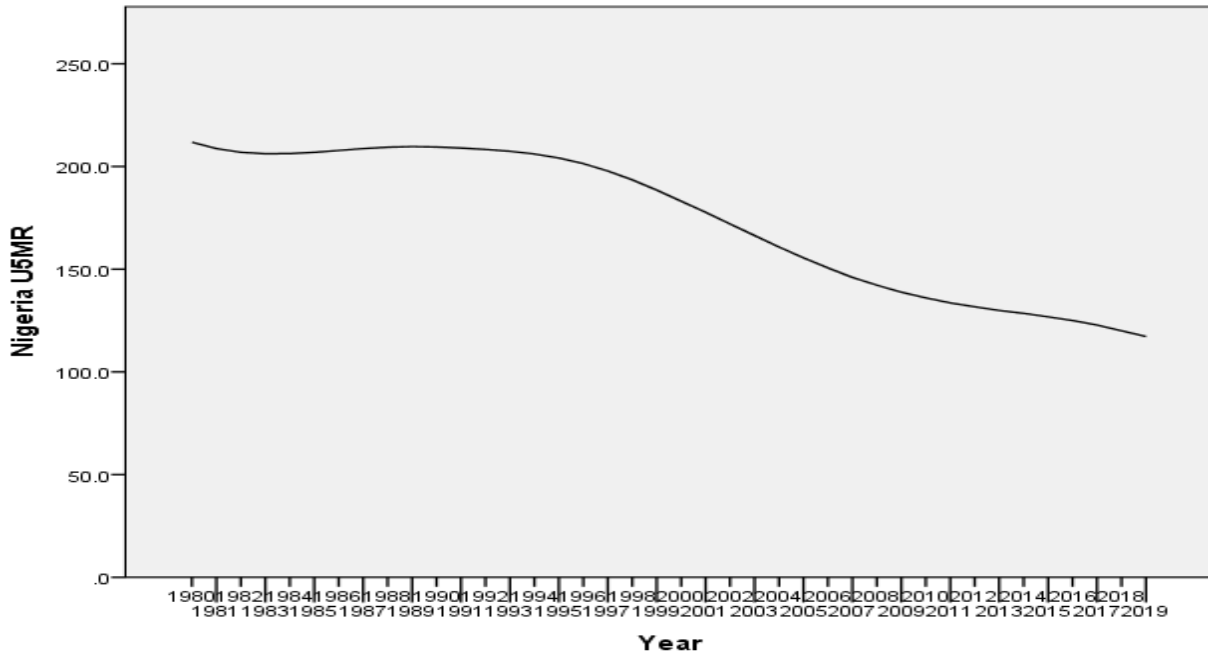


Fig. 1. Timeplot of U5MR in Nigeria for the Period 1980-2019

**Table 1. Descriptive Statistics for U5MR**

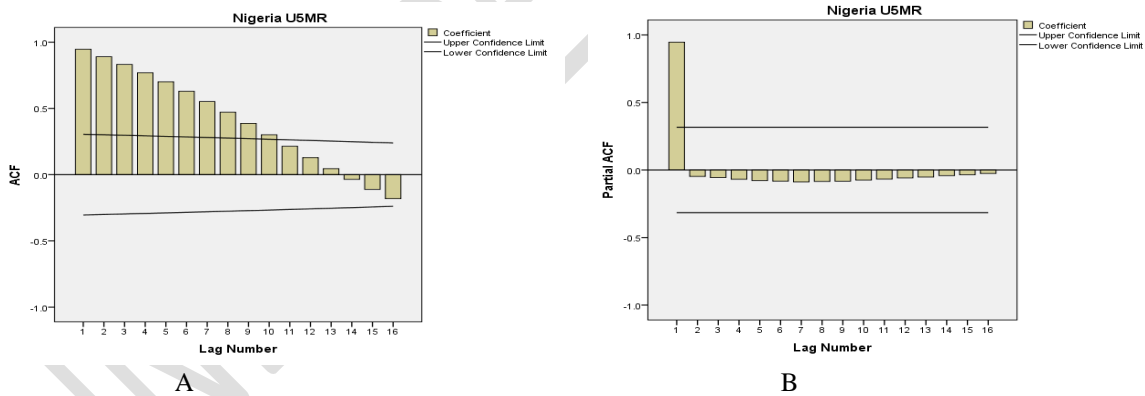
Descriptive Statistics						
	N	Minimum	Maximum	Mean	Std. Deviation	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
U5MR	40	117.2	211.8	174.315	5.4830	34.6777

**Table 2. Unit Root Test in Under-Five Mortality Rate (U5MR)**

Null Hypothesis: U5MR has a unit root			
Exogenous: Constant			
Lag Length: 3 (Automatic - based on SIC, maxlag=9)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		0.111199	0.9623
Test critical values:	1% level	-3.626784	
	5% level	-2.945842	
	10% level	-2.611531	

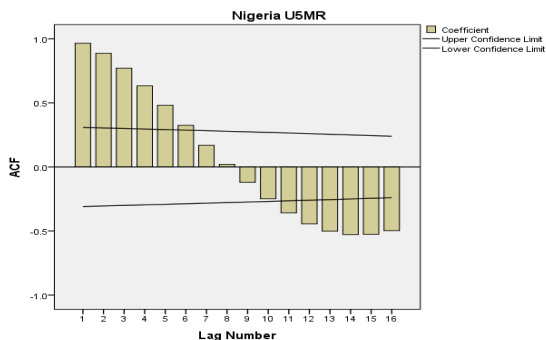
\*MacKinnon (1996) one-sided p-values.

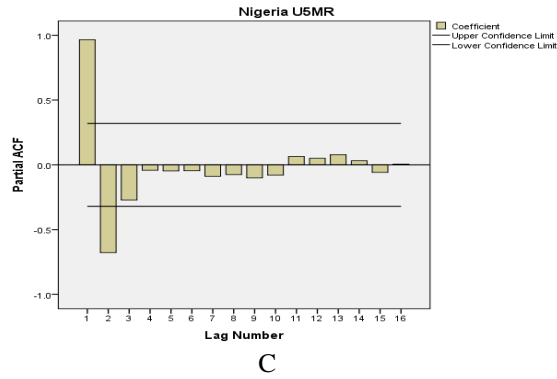
In Table 2, the Augmented Dickey-Fuller (ADF) test statistic is 0.111199 with p-value of 0.9623 greater than 0.05, implying that there is presence of unit root in the Under-Five Mortality Rate (U5MR), which indicates that there is need to difference at least once to get to stationarity.



**Fig. 2. (A) ACF Plot for Nigeria’s U5MR (B) PACF Plot for Nigeria’s U5MR**

Figure 2A shows a slow fall of the lags as the lag number increases, thereby indicating a non-stationarity of Nigeria’s U5MR. However, a first differenced ACF and PACF of Nigeria’s U5MR is obtained as shown in Figure 3.

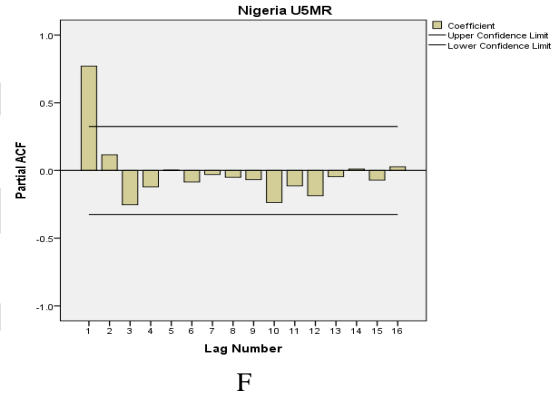
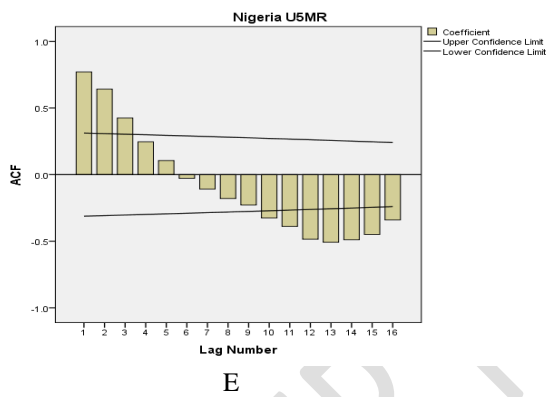




D

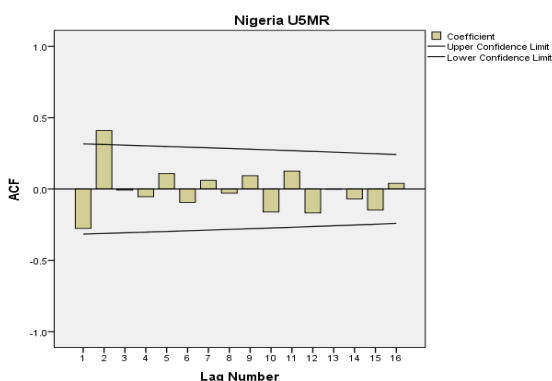
**Fig. 3. (C) First Differenced ACF Plot for Nigeria's U5MR (D) First Differenced PACF Plot for Nigeria's U5MR**

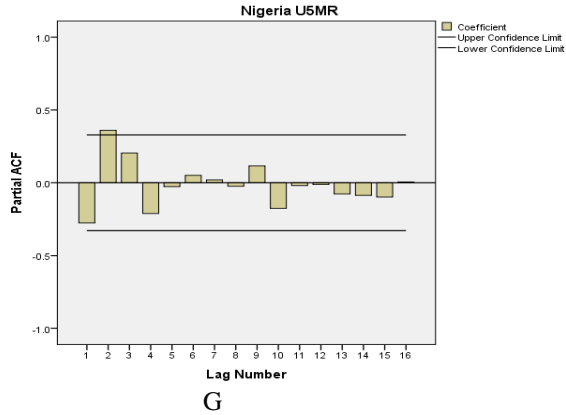
There is still a very slow fall of the lags as the lag number increases in Figure 3C which indicates that the first differenced Nigeria's U5MR is non-stationary, and will however require to be differenced the second time. Second differenced ACF and PACF of Nigeria's U5MR is obtained as shown in Figure 4



**Fig. 4. (E) Second Differenced ACF Plot for Nigeria's U5MR (F) Second Differenced PACF Plot for Nigeria's U5MR**

There is still a very slow fall of the lags as the lag number increases in Figure 4E which indicates that the second differenced Nigeria's U5MR is non-stationary, and will however requires to be differenced the third time. Third differenced ACF and PACF of Nigeria's U5MR is obtained as shown in Figure 5.





H

**Fig. 5. (G) Third Differenced ACF Plot for Nigeria's U5MR (H) Third Differenced PACF Plot for Nigeria's U5MR**

Figure 5G shows a quick fall at lag 1, thereby implying that the third differenced Nigeria's U5MR is now stationary. Lag 2 is significant as it cuts through the upper bound. And in Figure 5H, lag 2 is the only significant lag, which implies that the required ARIMA model is of order 2.

Again, ACF Plot in Figure 5G shows a sharp drop at lag 1 while there is a slowly decreasing trend in PACF Plot in Figure 5H after lag 2 which is the order of the model, indicating a Moving Average (MA) model.

In Table 3, ARIMA(0,3,2) has the smallest Bayesian Information Criterion (BIC) of -2.679 which indicates that the required ARIMA model that will be used to forecast Nigeria's U5MR is the ARIMA(0,3,2). The ARIMA (0,3,2) model that will be used to forecast the Nigeria's U5MR is written as

$$y'_t = -0.581\varepsilon_{t-1} + \varepsilon_t \quad (28)$$

**Table 3. ARIMA Model Adequacy Check**

ARIMA(p, d, q)	BIC
ARIMA(0,3,2)	-2.679
ARIMA(2,3,0)	-2.372
ARIMA(2,3,2)	-2.260
ARIMA(1,3,2)	-2.375
ARIMA(2,3,1)	-2.281
ARIMA(1,3,0)	-2.339
ARIMA(0,3,1)	-2.295

Table 4 shows the classification of Under-Five Mortality Rate (U5MR) in Nigeria into six (6) blocks (class intervals) with their respective states.

**Table 4. Classification of Under-Five Mortality Rate (U5MR)**

State	Block of the Nigeria's U5MRs
1	$y \leq 129.4$
2	$129.4 < y \leq 155.9$
3	$155.9 < y \leq 180.3$
4	$180.3 < y \leq 197.6$
5	$197.6 < y \leq 207.9$
6	$y > 207.9$

**Table 5. U5MR and State of Transition**

Year	U5MR	State	State Transition
1980	211.8	6	-
1981	208.7	6	66
1982	206.9	5	65
⋮	⋮	⋮	⋮
2000	183.1	4	44
2001	177.7	3	43
2002	172.0	3	33
⋮	⋮	⋮	⋮
2018	120.0	1	11
2019	117.2	1	11

In Table 5, taking year 1980, it is placed in state 6 but did not transit from another state due to no state preceding it. In the case of 1982, the U5MR transited from state 6 to state 5 in which it is being classified. And again, for 2019, the U5MR is transited from state 1 (U5MR of 2018) which comes before it.

The system in equation (22) is the Transition Matrix of the state transitions in Table 5; the system in equations (23), and (24) are One-Step Transition Probability Matrix, One-Step Marginal Probability Matrix and Marginal Matrix respectively.

$$F = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 1 & 8 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 8 & 1 \\ 0 & 0 & 0 & 0 & 2 & 6 \end{pmatrix} \end{matrix} \begin{matrix} = 5 \\ = 9 \\ = 4 \\ = 3 \\ = 10 \\ = 8 \end{matrix} \quad (29)$$

$$P = (p_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.11 & 0.89 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.75 & 0 & 0 & 0 \\ 0 & 0 & 0.33 & 0.67 & 0 & 0 \\ 0 & 0 & 0 & 0.10 & 0.80 & 0.10 \\ 0 & 0 & 0 & 0 & 0.25 & 0.75 \end{pmatrix} \quad (30)$$

$$q_{ij} = \begin{pmatrix} 0.128 & 0 & 0 & 0 & 0 & 0 \\ 0.026 & 0.205 & 0 & 0 & 0 & 0 \\ 0 & 0.026 & 0.077 & 0 & 0 & 0 \\ 0 & 0 & 0.026 & 0.051 & 0 & 0 \\ 0 & 0 & 0 & 0.026 & 0.205 & 0.026 \\ 0 & 0 & 0 & 0 & 0.051 & 0.153 \end{pmatrix}; \quad Q = (q_i) = \begin{pmatrix} 0.128 \\ 0.231 \\ 0.103 \\ 0.077 \\ 0.257 \\ 0.204 \end{pmatrix} \quad (31)$$

Table 6 shows the computation of the Chi-square for the verification of Markov using the information in the systems in equation (30) and (31).

**Table 6. Chi-square Test for Markov Property**

State, $i$	$f_{ij}$	$p_{ij}$	$q_i$	$\ln \frac{p_{ij}}{q_i}$	$f_{ij} \left( \ln \frac{p_{ij}}{q_i} \right)$
1	5	1.00	0.128	2.0557	10.2786
2	1	0.11	0.231	-0.7419	-0.7419
2	8	0.89	0.231	1.3488	10.7904
3	1	0.25	0.103	0.8867	0.8867
3	3	0.75	0.103	1.9853	5.9560
4	1	0.33	0.077	1.4553	1.4553
4	2	0.67	0.077	2.1635	4.3269
5	1	0.10	0.257	-0.9439	-0.9439
5	8	0.80	0.257	1.1355	9.0843
5	1	0.10	0.257	-0.9439	-0.9439
6	2	0.25	0.204	0.2033	0.4067
6	6	0.75	0.204	1.3020	7.8117
Total					48.3669

In Table 6, the Chi-square computed is 48.3669 and the Chi-square tabulated is 37.652 ( $\chi_{\alpha, (m-1)^2}^2 = \chi_{0.05, 25}^2 = 37.652$ ), this implies that the stochastic process obtained has a Markov Property. Table 7 shows the weights  $w_k$  and the autocorrelation coefficients  $r_k$  for previous five-time series and the second, third, fourth, and fifth step transition probability matrices are shown in the system in Equation (25).

**Table 7. Estimated Autocorrelation Coefficients and Weights of Markov Chain**

	$k$					Total
	1	2	3	4	5	
$r_k$	0.94610	0.89009	0.83132	0.76859	0.70121	4.13731
$w_k$	0.22868	0.21514	0.20093	0.18577	0.16948	

$$P^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.44 & 0.56 & 0 & 0 & 0 & 0 \\ 0.03 & 0.41 & 0.56 & 0 & 0 & 0 \\ 0 & 0.08 & 0.47 & 0.45 & 0 & 0 \\ 0 & 0 & 0.03 & 0.15 & 0.67 & 0.15 \\ 0 & 0 & 0 & 0.03 & 0.39 & 0.58 \end{pmatrix} \quad P^{(3)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.58 & 0.42 & 0 & 0 & 0 & 0 \\ 0.07 & 0.51 & 0.42 & 0 & 0 & 0 \\ 0 & 0.20 & 0.50 & 0.30 & 0 & 0 \\ 0 & 0 & 0.07 & 0.17 & 0.57 & 0.19 \\ 0 & 0 & 0 & 0.06 & 0.46 & 0.48 \end{pmatrix} \quad (32)$$

$$P^{(4)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.68 & 0.32 & 0 & 0 & 0 & 0 \\ 0.13 & 0.55 & 0.32 & 0 & 0 & 0 \\ 0.03 & 0.30 & 0.47 & 0.20 & 0 & 0 \\ 0 & 0.03 & 0.11 & 0.17 & 0.50 & 0.19 \\ 0 & 0 & 0.02 & 0.08 & 0.49 & 0.41 \end{pmatrix} \quad P^{(5)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.76 & 0.24 & 0 & 0 & 0 & 0 \\ 0.19 & 0.57 & 0.24 & 0 & 0 & 0 \\ 0.06 & 0.28 & 0.42 & 0.24 & 0 & 0 \\ 0 & 0.05 & 0.12 & 0.16 & 0.46 & 0.20 \\ 0 & 0 & 0.05 & 0.10 & 0.50 & 0.35 \end{pmatrix}$$

**Table 8. Forecast for 2015 Under-Five Mortality Rate (U5MR)**

Initial Year	State	Step, $k$	Weight, $w_k$	1	2	3	4	5	6	Probability Source
2014	1	1	0.22868	1	0	0	0	0	0	$p^{(1)}$
2013	2	2	0.21514	0.44	0.56	0	0	0	0	$p^{(2)}$
2012	2	3	0.20093	0.58	0.42	0	0	0	0	$p^{(3)}$
2011	2	4	0.18577	0.68	0.32	0	0	0	0	$p^{(4)}$
2010	2	5	0.16948	0.76	0.24	0	0	0	0	$p^{(5)}$
$\hat{p}_{ij}$ (Weighted Average)				0.695	0.305	0	0	0	0	

Table 8 shows  $\max\{\hat{p}_{ij}\} = 0.695$ , indicating that the U5MR in 2014 is in state 1 with the highest probability of 0.695, and it satisfies the interval  $y \leq 129.4$ , where the actual U5MR in 2015 which is 126.8 falls within the interval. Taking the average of the interval, the forecast of U5MR is 64.7. Similarly, forecasting the U5MR for 2016 using 2011-2015 as initial states as given in **Table 9**, U5MR is in state 1 with probability 0.815, and it falls within the interval  $y \leq 129.4$ . The actual U5MR in 2016 which is 125 shows that the prediction is also true. The same procedure is used to forecast for the remaining years. **Table 10** shows the In-sample for the period 2000-2019 and Out-of-Sample forecast of U5MR for the period 2020-2030 using ARIMA(0,3,2) and WMC and performance measures (Theil's U statistic and MAPE)

**Table 9. Forecast for 2016 Under-Five Mortality Rate (U5MR)**

Initial Year	State	Step, $k$	Weight, $w_k$	1	2	3	4	5	6	Probability Source
2015	1	1	0.22868	1	0	0	0	0	0	$p^{(1)}$
2014	1	2	0.21514	1	0	0	0	0	0	$p^{(2)}$
2013	2	3	0.20093	0.58	0.42	0	0	0	0	$p^{(3)}$
2012	2	4	0.18577	0.68	0.32	0	0	0	0	$p^{(4)}$
2011	2	5	0.16948	0.76	0.24	0	0	0	0	$p^{(5)}$
$\hat{p}_{ij}$ (Weighted Average)				0.815	0.185	0	0	0	0	

**Table 10. In-sample and Out-of-Sample Forecast of U5MR for 2000-2030 and Performance Measures**

Year	2000	2001	.....	2029	2030	MAPE	Theil's U Statistics
ARIMA(0,3,2)	182.9	177.3	.....	97.4	96.6	0.174 3 36%	0.000014
WMC	182.0	168.1	.....	64.7	64.7	17.5101 30%	0.001702

ARIMA(0,3,2) in Table 10 has a Theil's U Statistic of 0.000014 and MAPE of 0.174336% which are all lower than that of the WMC having a Theil's U Statistic of 0.001702 and MAPE of 17.510130%. This implying that ARIMA(0,3,2) is selected as the best model to forecast U5MR in Nigeria. However, by 2030, the U5MR will drop to 96.6 deaths per 1000 live births, which shows a drop of 20.6%. Figure 6 shows the timeplot for the historical data for the period 1980-2019 and out-of-sample forecast of U5MR for the period 2020-2030.

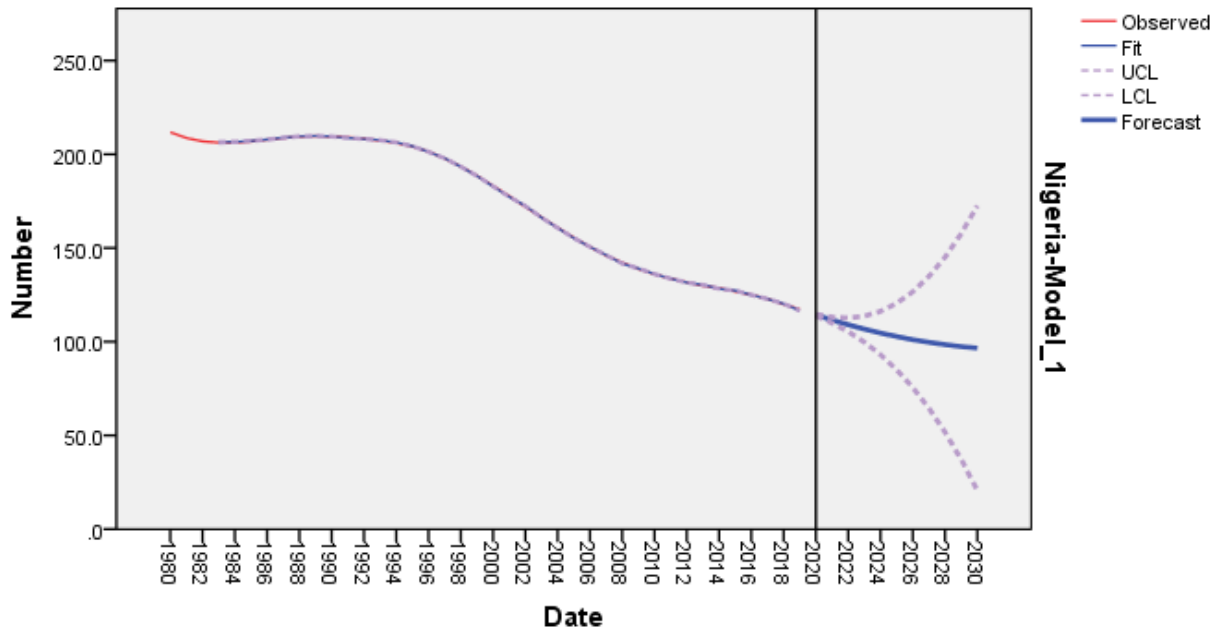


Fig. 6. Timeplot of the Actual U5MR and Out-Sample Forecast of U5MR using ARIMA(0,3,2)

#### 4. Conclusion

The purpose of this paper is to identify the best model that will be used to forecast U5MR in Nigeria. ARIMA(0,3,2) predicts U5MR better than WMC, and based on the modeling and forecasting, the U5MR is showing an intrinsic decrease from year to year and by 2030, the U5MR will drop to 96.6 deaths per 1000 live births, which shows a drop of 20.6%. However, the declined in the U5MR is due to the government involvement through massive sensitization and provision and establishment of health care facilities across the nation. The findings of this study can help promote health policies in order to address and to reduce U5MR in the future, as well as to establish a basis for implementing optimal strategies that can be used to overcome U5MR in order to meet up with the target of SDGs.

#### References

- [1] Adeyinka, D. A. and Muhajarine, N. Time series prediction of under-five mortality rates for Nigeria: Comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models, *BMC Medical Research Methodology*, 2020, 20:292
- [2] Adekanmbi, V. I., Kayode, G. O. and Uthman, O. A. Individual and contextual factors associated with childhood stunting in Nigeria: A multilevel analysis, *Maternal Child Nutr*, 2013, 9(2), 244-259
- [3] Aremu, O., Lawoko, S. and Dalal, K. Neighbourhood Socioeconomic disadvantage, individual wealth status and patterns of delivery care utilization in Nigeria: A multilevel discrete choice analysis, *International Journal of Women's Health*, 2011, 3, 167-174
- [4] Eke, D. O. and Ewere, F. Modeling and forecasting under-five mortality rate in Nigeria using auto-regressive integrated moving average approach, *Earthline Journal of Mathematical Sciences*, 2020, 4(2), 2581-8147
- [5] Husin, W. Z. W., Ramli, R. Z., Muzaffar, A. N., Nasir, N. F. A. and Rahmat, S. N. E. Trend analysis and forecasting models for under-five mortality rate in Malaysia, *Palarch's Journal of Archaeology of Egypt/Egyptology*, 2020, 17(10), 875-889

- [6] Kafi, R. A., Safitri, Y. R., Widyaningsih, Y. and Handari, B. D. Comparison of weighted Markov chain and Fuzzy time series Markov chain in forecasting stock closing price of company X, *AIP Conference Proceedings* 2168, 020033(2019), <https://doi.org/10.1063/1.5132460>
- [7] Kordnoori, S., Mostafaei, H. and Kordnoori, S. Applied SCGM(1,1)c Model and weighted Markov chain for Exchange Rate Ratios, *Hyperion Economic Journal*, 2015, 3(4), 12-22
- [8] Mesike, G. and Mojekwu, N. Environmental determinants of child mortality in Nigeria, *Journal of Sustainable Development*, 2012, 5(1), <https://doi.org/10.5539/jsd.v5n1p65>
- [9] Obasohan, P. E. Comparing weighted Markov chain and autoregressive integrated moving average in the prediction of under-five mortality annual closing rates in Nigeria, *International Journal of Statistics and Probability*, 2020, 9(3), 13-22
- [10] Peng, Z., Changjun Bao, Yang Zhao, Honggang Yi, Letian Xia, Hao Yu, Hongbing Shen and Feng Chen, Weighted Markov chains for forecasting and analysis in incidence of infectious diseases in Jiangsu Province, China, *Journal of Biomedical Research*, 2010, 24(3), 207-214
- [11] Shahdoust, M., Sadeghifar, M., Poorolajal, J., Javanrooh, N. and Amini, P. Predicting hepatitis B monthly incidence rates using weighted Markov chains and time series methods, *Journal of Research in Health Sciences*, 2015, 15(1), 28-31
- [12] Zhou, Qing-xin, Application of weighted Markov chain in stock price forecasting of China sport industry, *International Journal of u- and e-Service, Science and Technology*, 2015, 8(2), 219-226
- [13] UNICEF. (2021). Child Survival. Retrieved from [http://www.data.unicef.org/child\\_survival](http://www.data.unicef.org/child_survival)
- [14] Odimegwu, C. and 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.