http://wojast.org

https://dx.doi.org/10.4314/wojast.v15i2.35

ISSN: 2141 – 3290
www.wojast.com

A TIME SERIES MODELLING AND FORECASTING OF UNDER-5 MORTALITY RATE IN NIGERIA

DAVID, A. E. AND RAHEEM, M.A.*
Department of Statistics, University of Uyo, Nigeria
***Correspondences:** *rahemarsac@yahoo.com*

ABSTRACT

This paper applies an appropriate univariate time series model to examine the trend in the Under-5 Mortality Rate in Nigeria. Fifty-seven years data, ranging from 1964-2020, were obtained from the United Nations Inter-agency Group for Child Mortality Estimation website for this study. The empirical study revealed that the most adequate model for the under-5 mortality rates is ARIMA(1,1,0). The adequacy of the fitted model was determined and 10 years forecast of Under-5 mortality rates in Nigeria was made. With the forecast series, we found a steady decrease in the rate, which was consistent with the trends in the original series; indicating that the government and stakeholders' efforts over the years in Nigeria are yielding encouraging results that would in the long run, lead to a significant drop in the under-5 mortality rate in the country. However, given the forecast results of this research (of approximately 56 deaths per 1000 live births in 2030), the dreams of achieving SDG goals of 25 per 1000 live births by 2030 in Nigeria may not be realistic unless drastic measures are taken.

KEYWORDS: $ARIMA(p,d,q)$, Out-of-Sample forecast, Under-5 Mortality Rate, ACF, PACF

INTRODUCTION

According to UNICEF (2023), a child born in sub-Saharan Africa was 10 times at risk of dying within the first month than those born in a high-income country. Globally, the risk of dying in the first month of life was about 53 times higher for children born in developing countries with the highest mortality compared to those born in developed countries, with higher life expectancies. Under-5 mortality is a significant barometer for underpinning the general health status of any country. It serves as evidence of the priorities any given country places on the general well-being of its citizens and future generations. You (2012) and Abir, *et al*. (2015), reported that the number of under-5 deaths globally declined from nearly 12 million in 1990 to 6.9 million in 2011; indicating that about 14 000 fewer children dying each day in 2011 than in 1990. In 2019, an estimated 5·30 million Under-5 children were recorded, which represents a significant reduction from the estimated 9·92 million Under-5 mortality that was reported in 2000. The decline was a reflection of many interventions by different organizations and governments between 2000 and 2015; representing a decline in under-5 mortality rate from 75 deaths per 1000 livebirths in 2000, to 38 deaths per 1000 livebirths in 2019 (Perin, *et*, *al*. 2022). The drop in the under-five mortality rate(U5MR) was also attributed to the higher concentration of under-five deaths during the neonatal period. Meanwhile, 47% of the world under-five deaths in 2021, were attributed to neonatal deaths with sub-Saharan Africa having 38% of the neonatal deaths.

The study conducted by Odah (2021), attempted to adapt time series models on U5MR in Iraq from 1985 to 2020, the study found ARIMA (1,0,0) as the most appropriate in predicting U5MR, model fit Statistics and Lung-Box Q. Statistics, the forecast values, which was compared using RMSE and MAPE for evaluation confirm ARIMA model to have the highest forecasting performance. Rasheed (2008), examined the trend and pattern of under-five deaths in Lagos state, Nigeria using secondary data. The study showed that the projection of 55 deaths per 1000 live births for 2015 may be unattainable given the persistent increase in the mortality recorded in the past five years. The study further found that 73.8% of infant mortality and 37% of neonatal deaths were recorded. The study found some of the common causes of Under-5 were bronco, pneumonia, sepsis, anemia, and malaria.

Similar study conducted by Girma, Nesredin and Ketema (2021) to examine the U5MR in Ethiopia using Box-Jenkin's approach found ARIMA (2,2,5) to be the best fitted model. The seven years out of sample forecast from 2019 to 2025 showed a decline in the under-5 mortality trend. The study concludes that with the forecast value of 41.68 deaths per 1000 live births, Ethiopia may be far from achieving the SDG goals of reducing the trend by 2025. In the meantime, despite the remarkable progress, efforts at enhancing child's survival rates remains a matter of global concern, particularly across the sub-Saharan African countries. Nigeria with a population of over 200 million and about 53.5% of the population living below $1.9 a day (WHO/CCU/18.02/ Nigeria), is one of the five countries accounting for half of global burden of infant mortality occupying a second place behind India; thus, posing a serious challenge towards achieving the SDGs goal of reducing U5MR to 25 deaths per 1000 live births by 2030. Given the position of Nigeria with respect to these various mortalities and the importance of U5MR mortality statistics in the assessment of the overall health status of the country, and the studies investigated so far, it becomes necessary to examine the trend in the U5MR historically and identify appropriate time series model that could reproduce the original series and make reliable out-of-sample forecast for at least 10 years ahead.

METHODOLOGY

This study would explore Box-Jenking's time series procedures of model fitting by starting with obtaining the time plot to identify the inherent time series components (such as Trend, seasonality, cyclical and irregular variations), characterizing the series. Next is to investigate series for stationarity by testing for presence (or otherwise) of unit root. Having ensured stationarity, correlograms (autoregressive and partial autoregressive functions plots) are obtained to identify possible models that could better underpin the series behaviour. After fitting of the most appropriate model (using Akaike or Bayesian information criteria), a ten-year out-of-sample forecast shall be made for us to achieve the set objective of this study.

Stationarity in Time series

The first task in time series is to obtain the time plot and examine the series for stationarity before fitting appropriate Box-Jenking's model to the series. Though there are two types of stationarity, namely strictly and weakly stationarity in time series, it is expected that any stationary series should at least be weakly (second-order) stationary. And for a weakly stationary series; it is likely that (i) the mean and the variance of such series should be constant over time, and that (ii) it should be covariance-stationary; i.e., the covariance should rather depend on the relative difference in time (time lags) between series and not on time.

Autoregressive processes

The autoregressive process of order p is denoted AR(p), and defined

$$X_t = \sum_{r=1}^{p} \varphi_r X_{t-r} + \epsilon_t \qquad (2.1)$$

where $\varphi_1, \dots, \varphi_r$ are autoregressive coefficients and $\{\epsilon_t\}$ is a sequence of independent (or uncorrelated) random variables with mean 0 and variance σ^2. That is, ϵ_t is a white noise.

Consider an $AR(1)$ process, defined by;

$$X_t = \varphi_1 X_{t-1} + \epsilon_t . \qquad (2.2)$$

Obtaining the Mean, the Variance and Covariance functions of (2.1)

Consider re-writing (2.1) as a function of the residuals

$$X_t = \epsilon_t + \varphi_1(\epsilon_{t-1} + \varphi_1(\epsilon_{t-2} + \cdots))$$
$$= \epsilon_t + \varphi_1\epsilon_{t-1} + \varphi_1^2\epsilon_{t-2} + \cdots \qquad (2.3)$$

The fact that $\{X_t\}$ is second order stationary follows from the observation that $E(X_t) = 0$ and that the autocovariance function can be calculated as follows:

$$\gamma_0 = E[(\epsilon_t + \varphi_1\epsilon_{t-1} + \varphi_2^2\epsilon_{t-2} + \dots)]^2$$
$$= (1 + \varphi^2 + \varphi^4 + \dots)\,\sigma^2 \qquad (2.4)$$
$$= \frac{\sigma^2}{(1-\varphi_1^2-\varphi_2^2-\cdots)} \qquad (2.5)$$
$$\gamma_k = Cov\,(X_t, X_{t-k}) \equiv Cov\,(X_t, X_{t+k}) = E(X_t . X_{t-k})$$
$$\gamma_k = E\{(\textstyle\sum_{r=0}^{\infty} \varphi_1^r \varphi_{t-r} \sum_{s=0}^{\infty} \varphi_1^s \varphi_{t-s+k})\} = \frac{\sigma^2 \varphi_1^k}{1-\varphi_k^2} \qquad (2.6)$$

For example, the mean for $AR(1)$ is 0, while its variance is

$$\gamma_0 = \frac{\sigma^2}{(1-\varphi_1^2)} \qquad (2.7)$$

Thus, its autocovariance at lag k is;

$$\gamma_k = \varphi_1\gamma_{k-1}, \quad \forall\, k = 1,2,\dots \qquad (2.8)$$

Investigating the stationarity of the process

Consider the AR(p)

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \epsilon_t \qquad (2.9)$$

Multiplying (2.9) by X_{t-k}; take the expectation and divide by γ_0, to produce the Yule-Walker equation

$$\rho_k = \varphi_1\rho_{k-1} + \varphi_2\rho_{k-2} + \cdots + \varphi_p\rho_{k-p}, \forall\, k = 1,2 \qquad (2.10)$$

These are linear recurrence relations, with the general solution of the form

$$\rho_k = C_1 w_1^{|k|} + \cdots + C_p w_p^{|k|} \qquad (2.11)$$

where $\omega_1, \dots, \omega_p$ are the roots of the characteristic equation

$$\omega^p - \varphi_1\omega^{p-1} - \varphi_2\omega^{p-2} - \cdots - \varphi_p = 0 \qquad (2.12)$$

and $C_1, \dots, C_p$ in (2.11) are determined by setting $\rho_0 = 1$ for $k = 1, \dots, p-1$.

It is natural that $\gamma_k \to 0\ as\ k \to \infty$, in which case the inverse roots must lie within the unit circle; that is, $|\omega_i| < 1$. Thus one can conclude that the process is covariance stationary for any chosen value of $\varphi_1, \dots, \varphi_p$.

Moving Average Process

The moving average process of order q is denoted MA(q) and defined by

$$X_t = \sum_{s=0}^{q} \theta_s \epsilon_{t-s} \qquad (2.13)$$

where $\theta_1, \dots, \theta_q$ are fixed constants, $\theta_0 = 1$, and $\{\varepsilon_t\}$ is a sequence of independent (or uncorrelated) random variables with mean 0 and variance σ^2.

It is clear from the definition that this is a second-order stationarity and that

$$\gamma_k = \begin{cases} 0 & \forall\, |k| > q \\ \sigma^2 \sum_{s=0}^{q-|k|} \theta_s\theta_{s+k} & \forall\, |k| \le q \end{cases} \qquad (2.14)$$

We remark that two moving average processes can have the same autocorrelation function. For instance,

$$X_t = \epsilon_t + \theta\epsilon_{t-1} \qquad (2.15)$$
$$X_t = \epsilon_t + \left(\frac{1}{\theta}\right)\epsilon_{t-1} \qquad (2.16)$$

both have $\rho_1 = \frac{\theta}{(1+\theta^2)}, \quad \rho_k = 0, |k| > 1 \qquad (2.17)$

However, re-writing the first (in (2.15) as an inverse function of error terms (in terms of X_t) gives

$$\epsilon_t = X_t - \theta\epsilon_{t-1}$$
$$= X_t - \theta = X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \qquad (2.18)$$

The (2.18) is called Invertible process, with $|\theta| < 1$ for the process to be stationary. No two invertible processes have the same autocorrelation function.

A R M A processes

The autoregressive moving average model provides a parsimonious description of a weakly stationary stochastic process compared to the autoregressive and the moving average processes. It is usually referenced as the $ARMA\,(p,q)$ model, consisting of the p-lags of autoregressive (AR) terms and q-lags of moving average (MA) terms, mathematically expressed as:

$$X_t - \sum_{r=1}^{p} \varphi_r X_{t-r} = \sum_{s=0}^{q} \theta_s \epsilon_{t-s} \qquad (2.19)$$

where again $\{\epsilon_t\}$ is white noise.

A R I M A processes

When the original series $\{Y_t\}$ is not stationary, we then explore differencing the series, starting with

https://dx.doi.org/10.4314/wojast.v15i2.35

first-order differencing until the series becomes stationary. For example, the first order differenced series is achieved using:

$$X_t' = \nabla Y_t = Y_t - Y_{t-1} \quad (2.20)$$

The d-order differenced process

$$X_t^d = \nabla^d Y_t = \nabla(\nabla^{d-1} Y_t) \quad (2.21)$$

While differencing a series, we test for the stationarity and once the stationarity if achieved, the differenced series is adopted in fitting an appropriate ARMA model.

The process $\{Y_t\}$ is said to be an autoregressive integrated moving average process, $ARIMA(p, d, q)$, such that $X_t = \nabla^d Y_t$.
In the meantime, having ensured stationarity of the series, by testing for the presence or otherwise of the unit root using Augmented Dickey and Fuller (1979) or Kwiatkowski *et al.*, (1990). test statistic; next is to obtain the correlogram, where both the autocorrelation (ACF) and partial autocorrelation (PACF) functions have to be inspected for possible model(s) that best reproduce the original series. Next is the model estimation and selection, using (corrected) Akaike information criterion proposed by Akaike (1974) or Schwarz Bayesian information criterion elaborated by Bumham and Anderson (2002). The model with the lowest AIC or BIC value is usually favoured. After that, the fitted model is examined to ensure no serial correlation in the residuals, using the Durbin Watson Statistic.

Results and Discussions

Table 1 presents the summary statistics of the Under-5 mortality rates, with a total of 54 observations, representing yearly statistics from 1964 to 2020.

Table 1: Summary Statistics of the Under-5 Mortality Rates

Sample/Series: Under-5 Mortality rate	1964-2020
Number of observations (N)	54
Mean	200.5561
Median	207.3000
Minimum	113.8000
Maximum	325.2000
Standard deviation	57.06000
Skewness	0.383154
Kurtosis	2.429524
Jarque Bera	2.167595
Probability	0.338308

From the table, there is an average of about 201(7.76 SE) under-5 deaths recorded per 1000 live births. Across the 54 years, the least number of recorded under-5 mortality was approximately 114 deaths, while the highest was about 326 deaths per 1000 live births. The coefficient of variation about the mean is approximately 28.5%, which reflects low variability around the mean. Furthermore, the series, with a skewness of 0.38, though positively skewed, is about zero; and with a kurtosis of about 2.17 (less 3), shows the series is mesokurtic. Thus, with such characteristics exhibited by the series, we can conclude the original series is approximately normal. Meanwhile, to ascertain the normality of the series further, Jacque Bera test was conducted at 5% significant level; and with a p-value of 0.338 (> 0.05), it can be concluded that the original series is approximately normally distributed.

Figure 1 displays the time plots for the original (and the differenced) under-5 mortality series from 1964 to 2020. From the plot (on the right panel), it could be observed that there is a negative trend over the years in the series; indicating persistent drop in the rate of mortality from the highest point around 1964 to the lowest point around 2020. Meaning that the impact of interventions by all the relevant stakeholders, including the Nigerian government have paid of overtime, leading to a steady and persistent decrease witnessed in the series. Meanwhile, with the downward trends characterizing the series, one can say that the original series is non-stationary series, which would require appropriate transformation for the series to be stationary. Due to the behaviour exhibited by the series, we obtained the differenced series as displayed in the (left) panel. The plot though does not exhibit a long-term trend and appear to be stationary around the median level (of about -4). The oscillatory movement around the mean level is an indication of mean reversion expected of stationary series.

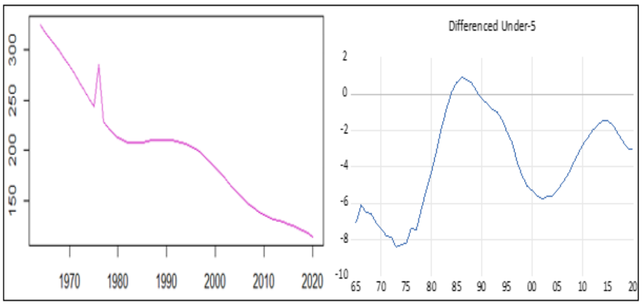


Figure 1: Time Plot of Under-5 Mortality Rate-1964-2020 (Original(left) and Differenced (Right) Series)

Figure 2 displays the normal plots for both original and the differenced series for the under-5 mortality rates in Nigeria from 1964 to 2020.

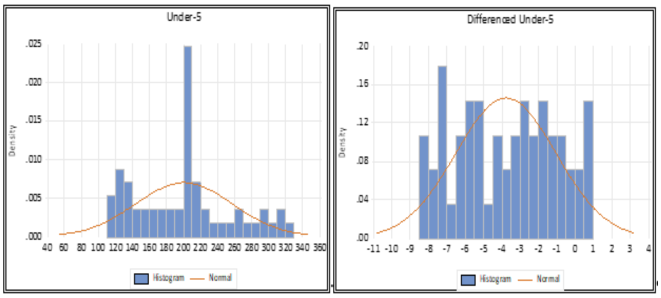
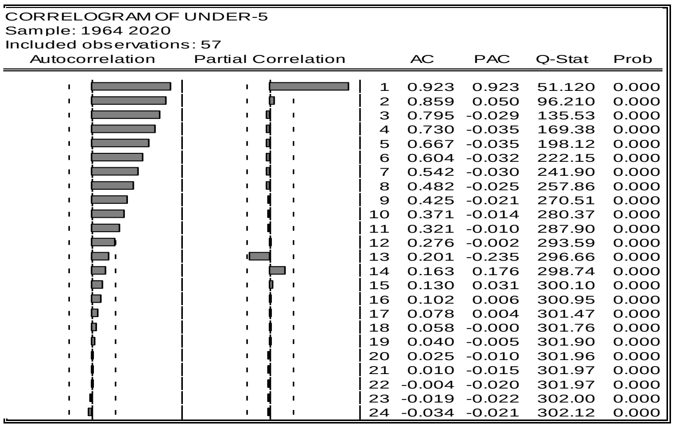


Figure 2: Normality Plots for both original (right panel) and the Differenced (left panel) series for the Under-5 Mortality Rate (1964-2020)

From the plot (on the right panel) shows that though the original series seem symmetric about the mean, the kurtosis appears to be higher than that of the Normal distribution. However, the differenced series (on the left panel) appears more symmetric with approximate moderate kurtosis. Q-Q plot presented in the supplementary result 1 displays a clearer picture on the Normality of the series.

Figure 3 is the correlogram, where the ACF, PACF and values of test statistic for determination of the degree of serial correlation in the original series are displayed. While the ACF dies off slowly at long lags, the PACF cuts off after lag one. Besides, the significant p-values at long lags, suggest the presence of a trend, which is a sign of non-stationarity of the original series, and that the series is likely to be characterized by the AR process of order one.



	AC	PAC	Q-Stat	Prob
1	0.923	0.923	51.120	0.000
2	0.859	0.050	96.210	0.000
3	0.795	-0.029	135.53	0.000
4	0.730	-0.035	169.38	0.000
5	0.667	-0.035	198.12	0.000
6	0.604	-0.032	222.15	0.000
7	0.542	-0.030	241.90	0.000
8	0.482	-0.025	257.86	0.000
9	0.425	-0.021	270.51	0.000
10	0.371	-0.014	280.37	0.000
11	0.321	-0.010	287.90	0.000
12	0.276	-0.002	293.59	0.000
13	0.201	-0.235	296.66	0.000
14	0.163	0.176	298.74	0.000
15	0.130	0.031	300.10	0.000
16	0.102	0.006	300.95	0.000
17	0.078	0.004	301.47	0.000
18	0.058	-0.000	301.76	0.000
19	0.040	-0.005	301.90	0.000
20	0.025	-0.010	301.96	0.000
21	0.010	-0.015	301.97	0.000
22	-0.004	-0.020	301.97	0.000
23	-0.019	-0.022	302.00	0.000
24	-0.034	-0.021	302.12	0.000

Figure 3: Correlogram of the Original series of Under-5 Mortality

Stationarity test of the series, the Augmented Dickey-Fuller test was conducted to check for the presence of unit root in the differenced U5MR series, the result is presented in Figure 4.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(UNDER_5)
Method: Least Squares
Date: 10/04/23 Time: 16:43
Sample (adjusted): 1968 2020
Included observations: 53 after adjustments

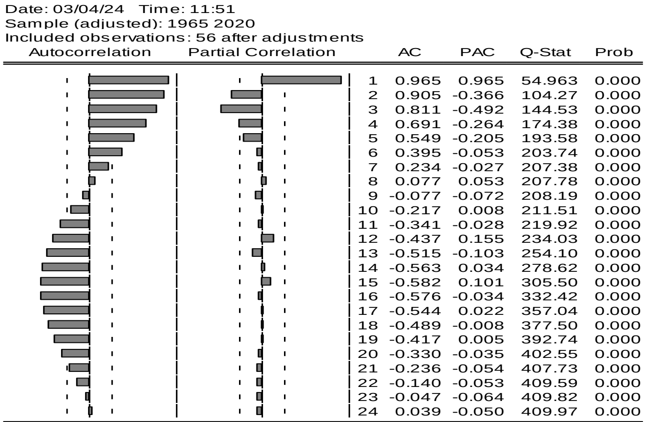
Variable	Coefficient	Std. Error	t-Statistic	Prob.
UNDER_5(-1)	-0.022872	0.004707	-4.859197	0.0000
D(UNDER_5(-1))	1.038238	0.119566	8.683406	0.0000
D(UNDER_5(-2))	0.276794	0.174629	1.585044	0.1197
D(UNDER_5(-3))	-0.384536	0.087187	-4.410488	0.0001
C	6.337210	1.379683	4.593236	0.0000
@TREND("1964")	-0.069865	0.015279	-4.572542	0.0000

R-squared	0.994340	Mean dependent var	-3.616981
Adjusted R-squared	0.993738	S.D. dependent var	2.728886
S.E. of regression	0.215949	Akaike info criterion	-0.121279
Sum squared resid	2.191796	Schwarz criterion	0.101773
Log likelihood	9.213883	Hannan-Quinn criter.	-0.035504
F-statistic	1651.342	Durbin-Watson stat	1.894595
Prob(F-statistic)	0.000000		

Figure 4: Unit root Test for Stationarity Check of the Series

Null Hypothesis: Under-5 series has a unit root. ***Alternative Hypothesis***: Under-5 series does not have a unit root. The results of the test conducted on the differenced series showed that, with the p-value of 0.0000 which is smaller than 0.05, there is no presence of unit root in the series, also, the ADF value at absolute term was higher than that of the t-statistics at all significance level (1%, 5% and 10%) hence the series is stationary and it is integrated of order 1.

Having tested for the stationarity of the differenced series, the correlogram of the series was obtained as presented in **Figure 5**. The plot showed the ACF that dies off to zero as exponential function or damped sinusoidal exponential function while the accompanied PACF present the parameters of the model as depicting AR process, of orders 1, 2 and 3. We then proceeded to fitting the identified ARIMA candidate models that best capture the fluctuations in the series



	AC	PAC	Q-Stat	Prob
1	0.965	0.965	54.963	0.000
2	0.905	-0.366	104.27	0.000
3	0.811	-0.492	144.53	0.000
4	0.691	-0.264	174.38	0.000
5	0.549	-0.205	193.58	0.000
6	0.395	-0.053	203.74	0.000
7	0.234	-0.027	207.38	0.000
8	0.077	0.053	207.78	0.000
9	-0.077	-0.072	208.19	0.000
10	-0.217	0.008	211.51	0.000
11	-0.341	-0.028	219.92	0.000
12	-0.437	0.155	234.03	0.000
13	-0.515	-0.103	254.10	0.000
14	-0.563	0.034	278.62	0.000
15	-0.582	0.101	305.50	0.000
16	-0.576	-0.034	332.42	0.000
17	-0.544	0.022	357.04	0.000
18	-0.489	-0.008	377.50	0.000
19	-0.417	0.005	392.74	0.000
20	-0.330	-0.035	402.55	0.000
21	-0.236	-0.054	407.73	0.000
22	-0.140	-0.053	409.59	0.000
23	-0.047	-0.064	409.82	0.000
24	0.039	-0.050	409.97	0.000

Figure 5: Correlogram of the differenced Logged Series

Model Selection for Under-Five Mortality

To achieve the study's objective, which was to estimate and fit the stationary time series model that will best fit Nigeria's under-5 mortality rate, Information criteria, such as Schwarz criterion (SC), Akaike (AIC), Hannan-Quinn (HQC), and Adjusted-R^2 were examined to determine the best-fitted model for the under-5 mortality rate. Table 4 compares these tentative models using the aforementioned goodness of fit of the model parameters statistics. From the table, $ARIMA$ (1,1,0) was found to be the best model for estimating the series, because it has the highest value of $Adj\ R^2$, smallest; SC, AIC, and HQC values, compared to the other two candidate models.

Table 2: Model Specification for under-5 Series

MODELS	SELECTION CRITERIA				
	No of Significant Coefficient	Adj-R^2	SC	AIC	HQC
ARIMA (1, 1, 0)	2	0.9557	1.95143	1.8429	1.8850
ARIMA (2, 1, 0)	3	0.9828	1.09638	0.9517	1.0078
ARIMA (3, 1, 0)	3	0.900685	3.904435	2.7959	3.8380

Out of the three choice models, ARIMA (2, 1, 0) with the least Information criteria and the highest Adjusted R-squared is found to be the best model and the results of the estimated model parameters are presented in Table 3

D(UNDER_5) = -4.366030 + 1.752193Under_5(-1) - 0.786291Under_5(-2)

Table 3: Estimated Model Parameters

Dependent Variable: D(UNDER_5)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 07/28/24 Time: 18:56
Sample: 1965 2020
Included observations: 56
Convergence achieved after 21 iterations
Coefficient covariance computed using outer product of gradients

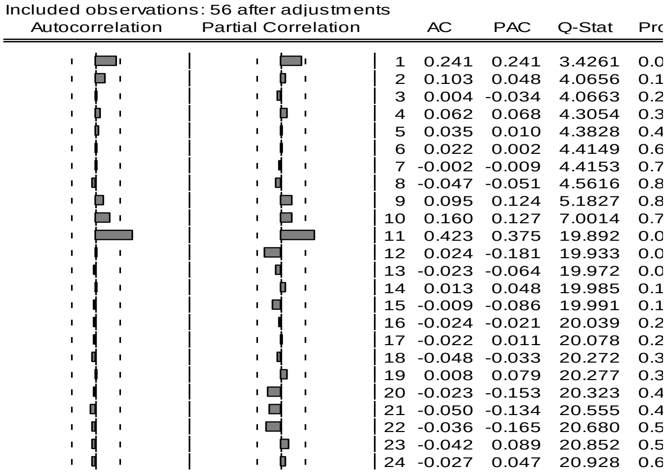
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.366030	1.674498	-2.607367	0.0119
AR(1)	1.752193	0.109717	15.97008	0.0000
AR(2)	-0.786291	0.102036	-7.706033	0.0000
SIGMASQ	0.119804	0.017726	6.758741	0.0000

R-squared	0.983734	Mean dependent var	-3.775000
Adjusted R-squared	0.982795	S.D. dependent var	2.738430
S.E. of regression	0.359193	Akaike info criterion	0.951711
Sum squared resid	6.709015	Schwarz criterion	1.096379
Log likelihood	-22.64791	Hannan-Quinn criter.	1.007799
F-statistic	1048.255	Durbin-Watson stat	2.813624
Prob(F-statistic)	0.000000		

Inverted AR Roots	.88+.14i	.88-.14i

RESIDUAL DIAGNOSTIC CHECK

To examine the adequacy of the selected model, given the correlogram of the squared residuals presented in Figure 6, it is evident no serial correlation left in the residuals of the fitted model, since none of the p-values for the Q-statistic is less than 0.05. Thus, the fact that the residuals are not autocorrelated/serially correlated, is an indication of the model adequacy. Besides, in Figure 7, where the stability plot for the model parameters it is apparent none of the inverse roots of the fitted model lies outside the unit circle. Thus, further reinforcing the adequacy of the fitted model. Additionally, the fact that the fitted and actual series are highly inter-twined, indicates that the fitted model strongly reproduced the original series, with the residual series showing no discerning pattern (see **Figure 10**). Supplementary result II where the original series and the predicted series are approximately the same, this further establishes the fact that the fitted model presents a good reproduction of the original series.


Date: 07/28/24 Time: 18:54
Sample (adjusted): 1965 2020
Included observations: 56 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Pro
		1	0.241	0.241	3.4261	0.0
		2	0.103	0.048	4.0656	0.1
		3	0.004	-0.034	4.0663	0.2
		4	0.062	0.068	4.3054	0.3
		5	0.035	0.010	4.3828	0.4
		6	0.022	0.002	4.4149	0.6
		7	-0.002	-0.009	4.4153	0.7
		8	-0.047	-0.051	4.5616	0.8
		9	0.095	0.124	5.1827	0.8
		10	0.160	0.127	7.0014	0.7
		11	0.423	0.375	19.892	0.0
		12	0.024	-0.181	19.933	0.0
		13	-0.023	-0.064	19.972	0.0
		14	0.013	0.048	19.985	0.1
		15	-0.009	-0.086	19.991	0.1
		16	-0.024	-0.021	20.039	0.2
		17	-0.022	0.011	20.078	0.2
		18	-0.048	-0.033	20.272	0.3
		19	0.008	0.079	20.277	0.3
		20	-0.023	-0.153	20.323	0.4
		21	-0.050	-0.134	20.555	0.4
		22	-0.036	-0.165	20.680	0.5
		23	-0.042	0.089	20.852	0.5
		24	-0.027	0.047	20.928	0.6

Figure 6: Correlogram for the Squared Residuals

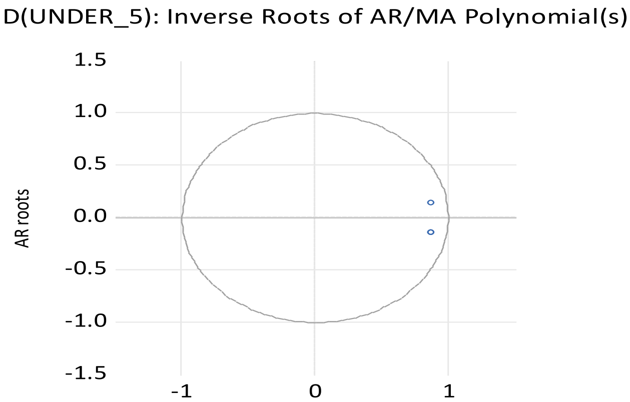


Figure 7: Stability plot for the Fitted Model Parameters

MODEL PREDICTION

Having fitted the identified model, ARIMA (2,1,0), we proceed on to make a ten-year out-of-sample forecast of Under-5 mortality rates under study from 2021 to 2031. The predicted values were compared with the actual values to validate the reproducibility of the fitted model. To further assess the performance ability of the forecast, the Root Mean Square Error was evaluated and found to be the least compared to any other possible model. The lower the RMSE the better the model performance. **Figure 8** presents the forecast evaluation. From the plot, the blue line represents the forecast, while the dotted red lines are the confident limits. Since the forecast succeeded in falling in between the upper and lower limits, we are confident that the forecast is good and reliable.

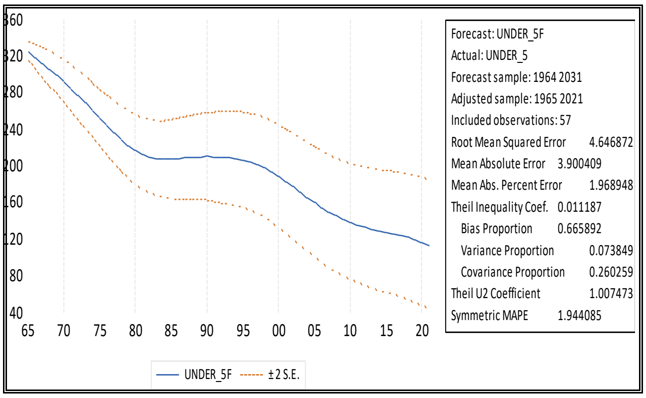


Figure 8: Forecast Performance Evaluation

CONCLUSION

Subject to the objectives of this study, the analysis conducted so far has succeeded in answering the accompanying research question intended for this research, with the result obtained, we have succeeded in establishing that ARIMA (2,1,0) is the best model for estimating Nigeria's Under-5 mortality rate within the study domain. Consequently, the model was applied in predicting the Under-5 mortality rate for the year 2021 to 2031 with a 95% confidence level. The forecast evaluation between the actual series and the predicted values for the predicted years shows that the model is appropriate. The predicted values (supplementary result III) from the ARIMA model revealed that Under-5 mortality rates in Nigeria will gradually decrease in the long run but the SDGs target of reducing U5MR to 25 deaths per 1000 live births by 2030 (Adeyinka and Muhajarine, 2020), may not be realistic except government policies are further strengthened given the importance of U5MR mortality statistics as a vital metrics for assessing the overall health status of a country. The conclusion in this regards stem from the fact that by 2031, according to the forecast, the Nigeria Under-5 mortality rate of about 56 deaths per 1000 live births would have been recorded; the rate that is about twice the expected SDG's target.

REFERENCE

Abir, T., Agho, K. E., Page, A. N., Milton, A. H., and Dibley, M. J. (2015). Risk factors for under-5 mortality: evidence from Bangladesh Demographic and Health Survey, 2004–2011. *BMJ Open*, *5*(8): e006722.

Adeyinka, D., Muhajarine, N., (2020). 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*, 20(292): *https://doi.org/10.1186/s12874-020-01159-9.*

Akaike, H. (1974), Information theory and an extension of the maximum likelihood principle, *2nd International Symposium on Information Theory, B.N. Petrov and F. Csaki (eds.)*, Akademiai Kiado, Budapest, pp. 267–281.

Burnham, K.P. and Anderson, D.R. (2002) Model Selection and Inference: A Practical Information-Theoretic Approach. 2nd Edition, Springer-Verlag, New York. http://dx.doi.org/10.1007/b97636

Dickey, D.A. and Fuller, W.A. (1979), Distribution of the estimators for autoregressive time series with a unit root, *Journal of American Statatistical Association*, 74: 427–431

Chu, T. W. and Shir, M. A. (2004). Evaluation of SWOT Model's hydrology component in the Piedmont physiographic region of Maryland. *Transactions of the ASAE* 47(4): 1057-11075

Girma, W., Nesredin, B., and Ketema, W., (2021). Time Series Analysis for Under-5 Mortality in Ethiopia. *Solid State Technology:* 64(2): 5121- 5133

Odah M. H. (2021). Time Series to Forecasting Under-five Mortality Rate in Iraq. *International Journal of Agriculture and Statistical Sciences*, 17(1): 1873-1877.

Kwiatkowski, D., Philip, P. C., Schmidt, P., and Shin, Y., (1992). Testing the null hypothesis of Stationary against the alternative of a unit root. *Journal of Econometrics*, 54: 159-178.

Perin, J., Mulick, A., Yeung, D., Villavicencio, F., Lopez, G., Strong, K. L., ... and Liu, L. (2022). Global, regional, and national causes of under-5 mortality in 2000–19: an updated systematic analysis with implications for the Sustainable Development Goals. *The Lancet Child* and *Adolescent Health*, *6*(2): 106-115.

Rasheed, O. K (2008). Pattern of Under-five Deaths in Lagos State, Nigeria. *Sudanese Journal of Public Health* 3(4):78-89.

UNICEF (2023) *The State of World's Children*. https://www.unicef.org/reports/state-worlds-children-2023 https://data.unicef.org/topic/child-survival/under-five-mortality/

WHO (2018). *Country Co-operation Strategy at a glance.* Report No: WHO/CCU/18.02/Nigeria

You, D., New, J. and Wardlaw T. (2012) Levels and trends in child mortality. Report 2012. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation. 2010

Supplementary Result I

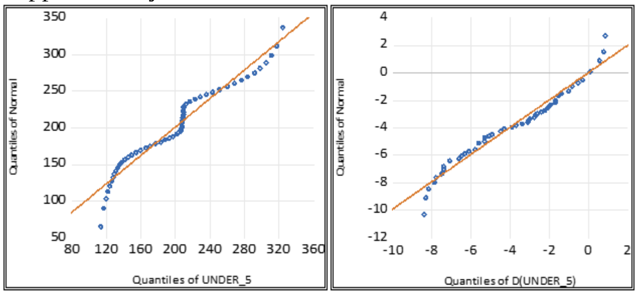


Figure 9: Q-Q plots for both Original and the Differenced series for the Under-5 Mortality rates from 1964-2020

Figure 10: Actual, Fitted and the Residual series for the Under-5 Mortality

Inverse Roots of AR/MA Polynomial(s)
Specification: D(UNDER_5) C AR(1) AR(2)
Date: 07/28/24 Time: 19:53
Sample: 1964 2020
Included observations: 56

AR Root(s)	Modulus	Cycle
0.876096 ± 0.136917i	0.886731	40.52969

No root lies outside the unit circle.
ARMA model is stationary.

Figure 11: Stability Test Results

Supplementary Result II

YEAR	U5MR	Predicted	LCL	UCL
1964.00	325.20			
1965.00	318.10	320.63	315.47	325.79
1966.00	312.00	311.06	309.92	312.21
1967.00	305.50	305.94	304.79	307.08
1968.00	298.90	299.05	297.90	300.19
1969.00	291.80	292.35	291.21	293.50
1970.00	284.40	284.76	283.62	285.91
1971.00	276.60	277.07	275.93	278.22
1972.00	268.70	268.88	267.74	270.03
1973.00	260.30	260.88	259.74	262.03
1974.00	252.00	252.00	250.85	253.14

1975.00	243.80	243.79	242.65	244.94
1976.00	236.40	235.69	234.55	236.84
1977.00	228.90	229.07	227.93	230.22
1978.00	222.60	221.47	220.33	222.62
1979.00	217.30	216.34	215.20	217.49
1980.00	213.00	212.02	210.87	213.16
1981.00	209.90	208.69	207.55	209.84
1982.00	208.00	206.76	205.62	207.91
1983.00	207.20	206.03	204.89	207.18
1984.00	207.30	206.31	205.16	207.45
1985.00	207.90	207.28	206.14	208.43
1986.00	208.80	208.37	207.23	209.52
1987.00	209.60	209.56	208.42	210.71
1988.00	210.20	210.27	209.12	211.41
1989.00	210.30	210.67	209.53	211.82
1990.00	210.00	210.28	209.14	211.43
1991.00	209.50	209.59	208.45	210.74
1992.00	208.70	208.90	207.75	210.04
1993.00	207.70	207.81	206.66	208.95
1994.00	206.30	206.61	205.47	207.76
1995.00	204.20	204.82	203.68	205.97
1996.00	201.50	202.04	200.89	203.18
1997.00	197.80	198.75	197.61	199.90
1998.00	193.30	194.08	192.93	195.22
1999.00	188.20	188.80	187.65	189.94
2000.00	182.90	183.11	181.97	184.26
2001.00	177.30	177.62	176.47	178.76
2002.00	171.50	171.73	170.58	172.87
2003.00	165.90	165.73	164.59	166.88
2004.00	160.30	160.33	159.18	161.47
2005.00	155.00	154.73	153.58	155.87
2006.00	150.10	149.72	148.57	150.86
2007.00	145.60	145.21	144.06	146.35
2008.00	141.70	141.10	139.95	142.24
2009.00	138.30	137.78	136.64	138.93
2010.00	135.50	134.87	133.73	136.02
2011.00	133.10	132.66	131.51	133.80
2012.00	131.10	130.65	129.50	131.79
2013.00	129.40	129.04	127.89	130.18
2014.00	127.90	127.63	126.48	128.77
2015.00	126.40	126.32	125.18	127.47
2016.00	124.70	124.82	123.68	125.97
2017.00	122.50	122.93	121.78	124.07
2018.00	119.90	120.24	119.10	121.39
2019.00	116.90	117.25	116.11	118.40
2020.00	113.80	113.86	112.72	115.01

Supplementary Result III

YEAR	$ARIMA$(1,1,0) FORECAST UNDER-5F
2021	105.68768
2022	100.07308
2023	94.48472
2024	88.92195
2025	83.38413
2026	77.87065
2027	72.38091
2028	66.91433
2029	61.47033
2030	56.04836
2031	56.04836