OPEN DEFECATION IN MOZAMBIQUE: A BOX-JENKINS ARIMA APPROACH

Dr. Smartson. P. NYONI[1], Thabani NYONI[2]

[1] *Medical Doctor, ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe*
[2] *MSc Economics Scholar, Department of Economics, University of Zimbabwe, Harare, Zimbabwe*

ABSTRACT

Based on annual time series data on the number of open defecators in Mozambique from 2000 – 2017, the study predicts the annual number of people who will still be practicing open defecation over the period 2018 – 2022. The study uses the Box-Jenkins ARIMA methodology. The diagnostic ADF tests show that the ODZ series under consideration is an I (1) variable. Based on the AIC, the study presents the ARIMA (1, 1, 0) model as the parsimonious model. The diagnostic tests further uncover that the presented model is very stable and its residuals are stationary in levels. The results of the study indicate that the number of open defecators in Mozambique is likely to decline over the period 2018 – 2022, from 25.14% to 17.61% of the total population. This implies that, open defecation will remain a public health issue in Mozambique, even though it is quite possible for the country to significantly reduce open defecation by 2025. However, it is possible to completely end open defecation in Mozambique, especially if the government remains committed to do so. The study suggested a basic three-fold policy recommendation to be put into consideration.

Keyword: - *Box-Jenkins ARIMA, Forecasting, Open Defecation*

1.0 INTRODUCTION

Despite the fact that Mozambique has made steady progress towards the achievement of Millennium Development Goals (MDGs), it still remains that only half of Mozambicans have access to improved water supply and only 1 in 5 use improved sanitation facilities. In Mozambique, open defecation is still a significant public health issue. Open defecation is dangerous to anyone, from a public health perspective; particularly, in terms of the spread of bacterial, viral and parasitic infections. However, Mozambique is committed to eliminating open defecation by 2025 and achieving universal access to adequate and sustainable sanitation and hygiene services by 2030, including effective services for the management of wastewater and faecal sludge. The main aim of the study is to model and forecast the number of people practicing open defecation in Mozambique.

2.0 LITERATURE REVIEW

Gitau & Flachenberg (2016) presented the lessons learnt from a pilot CLTS intervention in the challenging context of Somalia and generally found out that it is still possible to trigger to a great extent a community "with a desire for change with regards to open defecation and facilitate them to build their own household latrines without subsidization". Their study also indicated that commitment, attitude and mind-set of the implementers, community and the local authority are key to success in any context – fragile or stable. Ashenafi *et al.* (2018) analyzed the latrine utilization coverage of the kebeles who have already declared open defecation free in Ethiopia. Community-based cross-sectional study design with multistage sampling technique was used. Bivariate and multivariate logistic regression models were applied to identify factors associated with latrine utilization. Their results show that the extent of latrine utilization was high in the community. Alhassan & Anyarayor (2018) investigated the adoption of sanitation innovations introduced in Nadowli-Kaleo district in Upper West region of Ghana as part of the efforts to attain Open Defecation Free (ODF) status. Interviews were employed to gather data. The study showed that while effective communication of innovation resulted in widespread awareness, low income levels significantly accounted for households' inability to sustain and utilize latrines. This study will adopt the Box-Jenkins ARIMA method in analyzing open defecation trends in Mozambique and is apparently the first of its kind in the country.

3.0 METHODODOLOGY

3.1 The Box – Jenkins (1970) Methodology

The first step towards model selection is to difference the series in order to achieve stationarity. Once this process is over, the researcher will then examine the correlogram in order to decide on the appropriate orders of the AR and MA components. It is important to highlight the fact that this procedure (of choosing the AR and MA components) is biased towards the use of personal judgement because there are no clear – cut rules on how to decide on the appropriate AR and MA components. Therefore, experience plays a pivotal role in this regard. The next step is the estimation of the tentative model, after which diagnostic testing shall follow. Diagnostic checking is usually done by generating the set of residuals and testing whether they satisfy the characteristics of a white noise process. If not, there would be need for model re – specification and repetition of the same process; this time from the second stage. The process may go on and on until an appropriate model is identified (Nyoni, 2018c). This approach will be employed to analyze the ODZ series under consideration.

3.2 The Moving Average (MA) model

Given:

$$ODZ_t = \alpha_0\mu_t + \alpha_1\mu_{t-1} + \cdots + \alpha_q\mu_{t-q} \ldots\ldots\ldots\ldots [1]$$

where μ_t is a purely random process with mean zero and varience σ^2. Equation [1] is reffered to as a Moving Average (MA) process of order q, usually denoted as MA (q). ODZ is the annual number of people (as a percentage of the total population) who practice open defecation in Mozambique at time t, $\alpha_0 \ldots \alpha_q$ are estimation parameters, μ_t is the current error term while $\mu_{t-1} \ldots \mu_{t-q}$ are previous error terms.

3.3 The Autoregressive (AR) model

Given:

$$ODZ_t = \beta_1 ODZ_{t-1} + \cdots + \beta_p ODZ_{t-p} + \mu_t \ldots\ldots\ldots\ldots [2]$$

Where $\beta_1 \ldots \beta_p$ are estimation parameters, $ODZ_{t-1} \ldots ODZ_{t-p}$ are previous period values of the ODZ series and μ_t is as previously defined. Equation [2] is an Autoregressive (AR) process of order p, and is usually denoted as AR (p).

3.4 The Autoregressive Moving Average (ARMA) model

An ARMA (p, q) process is just a combination of AR (p) and MA (q) processes. Thus, by combining equations [1] and [2]; an ARMA (p, q) process may be specified as shown below:

$$ODZ_t = \beta_1 ODZ_{t-1} + \cdots + \beta_p ODZ_{t-p} + \mu_t + \alpha_1\mu_{t-1} + \cdots + \alpha_q\mu_{t-q} \ldots\ldots\ldots\ldots [3]$$

While ARMA models just like AR and MA models are meant for stationary series, reality indicates that most time series data is either I (1) or I (2). In fact, in this study, the ODZ series has been found to be an I (1) variables (that is, it only became stationary after first differencing). Therefore, in this paper, the model presented below is the one that will be applied.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

A stochastic process ODZ_t is referred to as an Autoregressive Integrated Moving Average (ARIMA) [p, d, q] process if it is integrated of order "d" [I (d)] and the "d" times differenced process has an ARMA (p, q) representation. If the sequence $\Delta^d ODZ_t$ satisfies an ARMA (p, q) process; then the sequence of ODZ_t also satisfies the ARIMA (p, d, q) process such that:

$$\Delta^d ODZ_t = \sum_{i=1}^{p} \beta_i \Delta^d ODZ_{t-i} + \sum_{i=1}^{q} \alpha_i \mu_{t-i} + \mu_t \ldots \ldots \ldots [4]$$

where Δ is the difference operator, vector $\beta \in \mathbb{R}^p$ and $\alpha \in \mathbb{R}^q$.

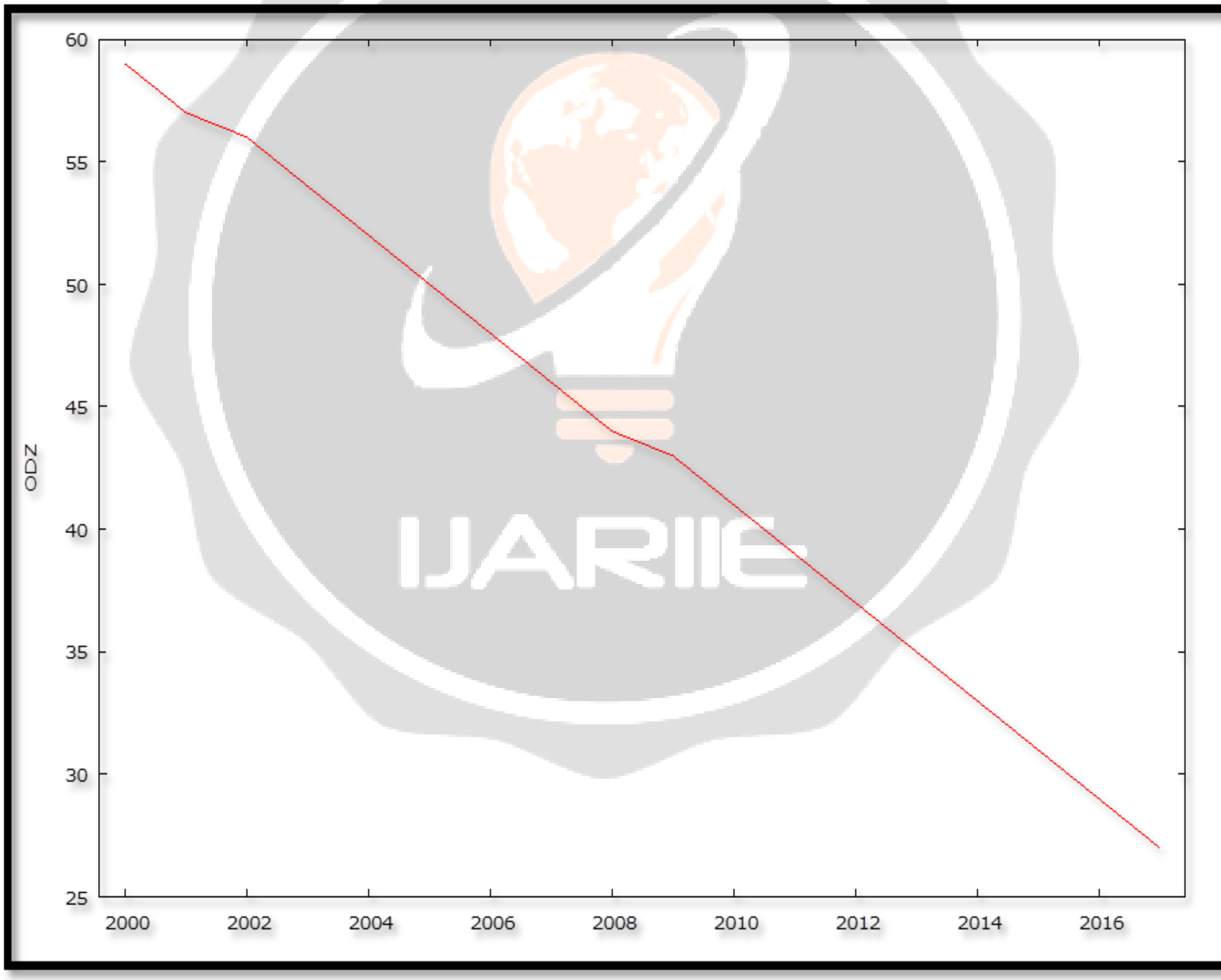
3.6 Data Collection

This study is based on annual observations (that is, from 2000 – 2017) on the number of people practicing Open Defecation [OD, denoted ODZ] (as a percentage of total population) in Mozambique. Out-of-sample forecasts will cover the period 2018 – 2022. All the data was gathered from the World Bank online database.

3.7 Diagnostic Tests & Model Evaluation

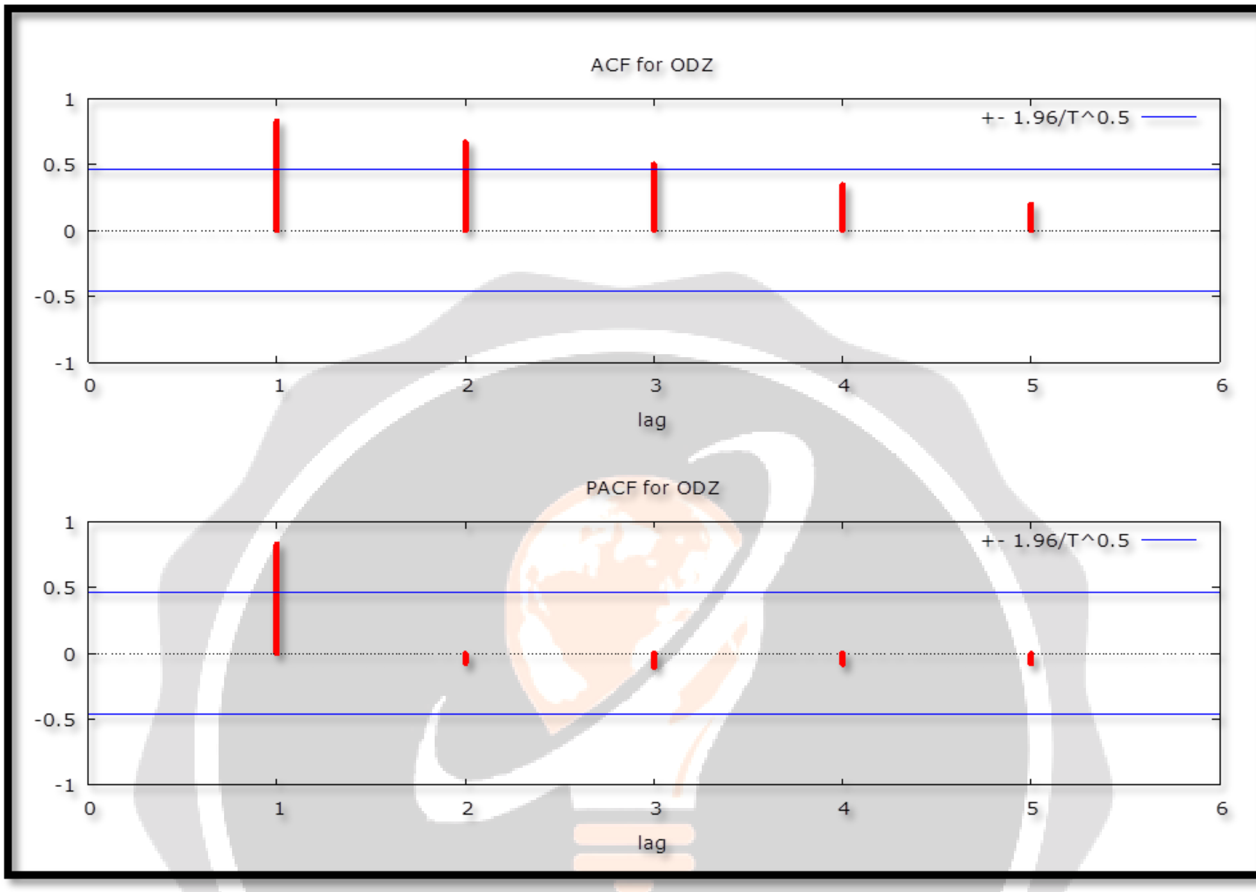
3.7.1 Stationarity Tests: Graphical Analysis

Figure 1



3.7.2 The Correlogram in Levels

Figure 2: Correlogram in Levels



3.7.3 The ADF Test in Levels

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODZ	0.975444	0.9939	-3.886751	@1%	Non-stationary
			-3.052169	@5%	Non-stationary
			-2.666593	@10%	Non-stationary

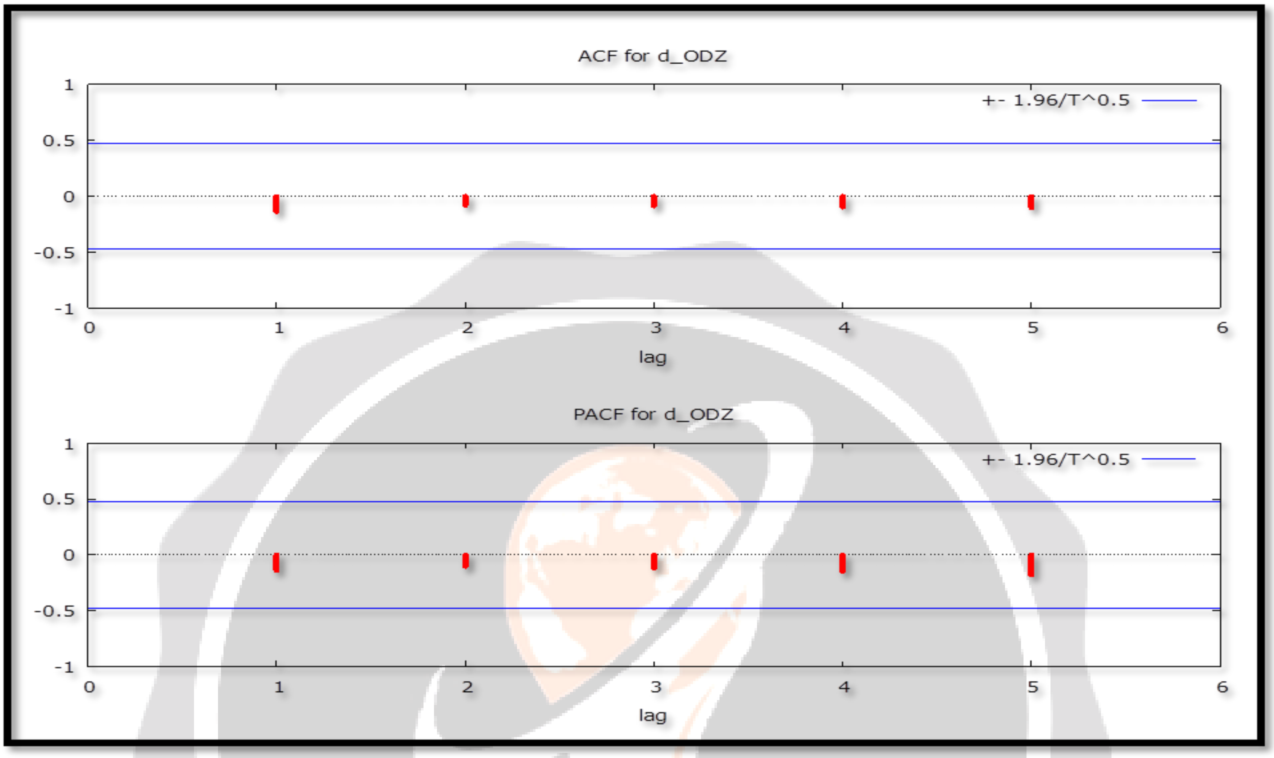
Table 2: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODZ	-2.324568	0.4005	-4.616209	@1%	Non-stationary
			-3.710482	@5%	Non-stationary
			-3.297799	@10%	Non-stationary

Tables 1 and 2 show that ODZ is not stationary in levels as already suggested by figures 1 and 2.

3.7.4 The Correlogram (at First Differences)

Figure 3: Correlogram (at First Differences)



3.7.5 The ADF Test (at First Differences)

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODZ	-4.320494	0.0046	-3.920350	@1%	Stationary
			-3.065585	@5%	Stationary
			-2.673459	@10%	Stationary

Table 4: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODZ	-4.722443	0.0091	-4.667883	@1%	Stationary
			-3.733200	@5%	Stationary
			-3.310349	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that ODZ is an I (1) variable.

3.7.6 Evaluation of ARIMA models (with a constant)

Table 5: Evaluation of ARIMA Models (with a constant)

Model	AIC	U	ME	MAE	RMSE	MAPE
ARIMA (1, 1, 0)	15.40626	0.14526	-0.0095178	0.20438	0.31894	0.47635
ARIMA (2, 1, 0)	17.15686	0.14544	0.005884	0.19636	0.3178	0.46264
ARIMA (3, 1, 0)	18.80749	0.14553	0.012394	0.18767	0.31578	0.44722

A model with a lower AIC value is better than the one with a higher AIC value (Nyoni, 2018b) Similarly, the U statistic can be used to find a better model in the sense that it must lie between 0 and 1, of which the closer it is to 0, the better the forecast method (Nyoni, 2018a). In this research paper, only the AIC is used to select the optimal model. Thus, the ARIMA (1, 1, 0) model is chosen.

3.8 Residual & Stability Tests

3.8.1 ADF Test (*in levels*) of the Residuals of the ARIMA (1, 1, 0) Model

Table 6: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-3.788739	0.0129	-3.920350	@1%	Non-stationary
			-3.065585	@5%	Stationary
			-2.673459	@10%	Stationary

Table 7: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-4.222023	0.0217	-4.667883	@1%	Non-stationary
			-3.733200	@5%	Stationary
			-3.310349	@10%	Stationary

Tables 6 and 7 indicate that the residuals of the chosen optimal model, the ARIMA (1, 1, 0) model; are stationary. Thus, the model is stable.

3.8.2 Correlogram of the Residuals of the ARIMA (1, 1, 0) Model

Figure 4: Correlogram of the Residuals

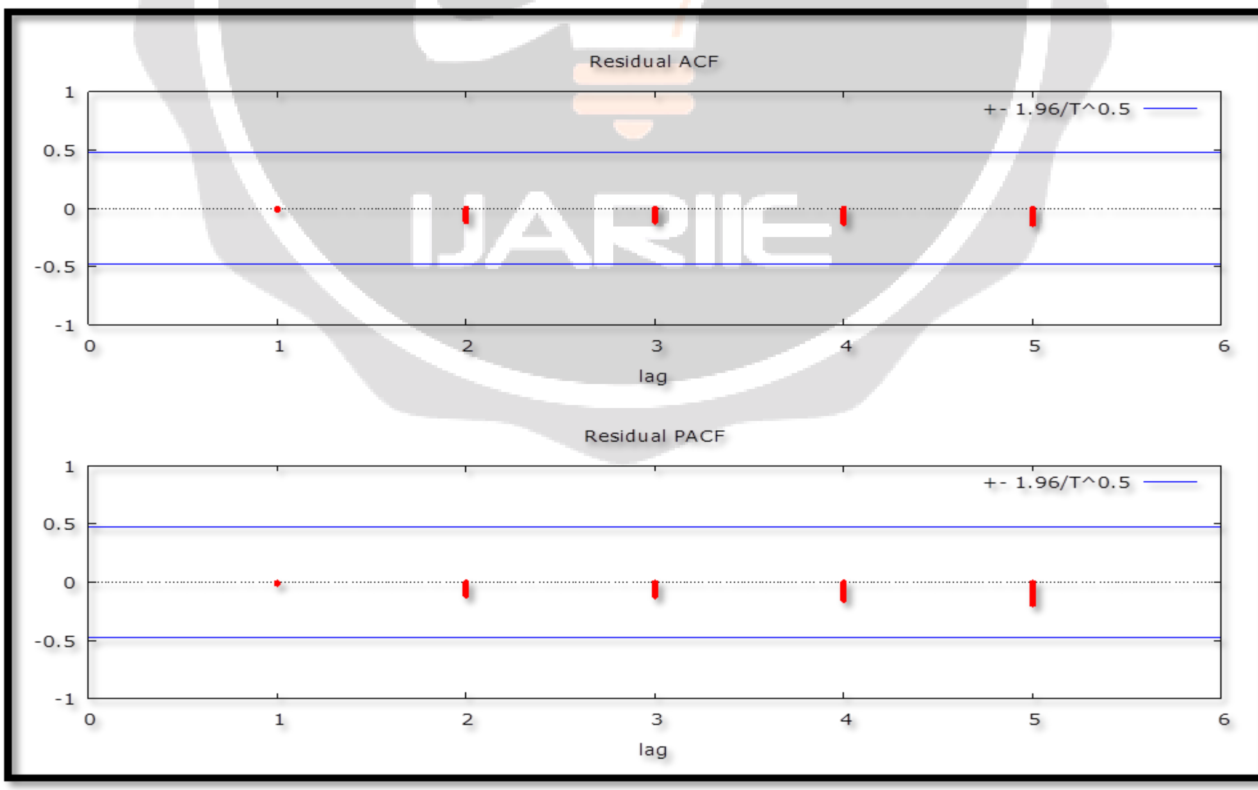
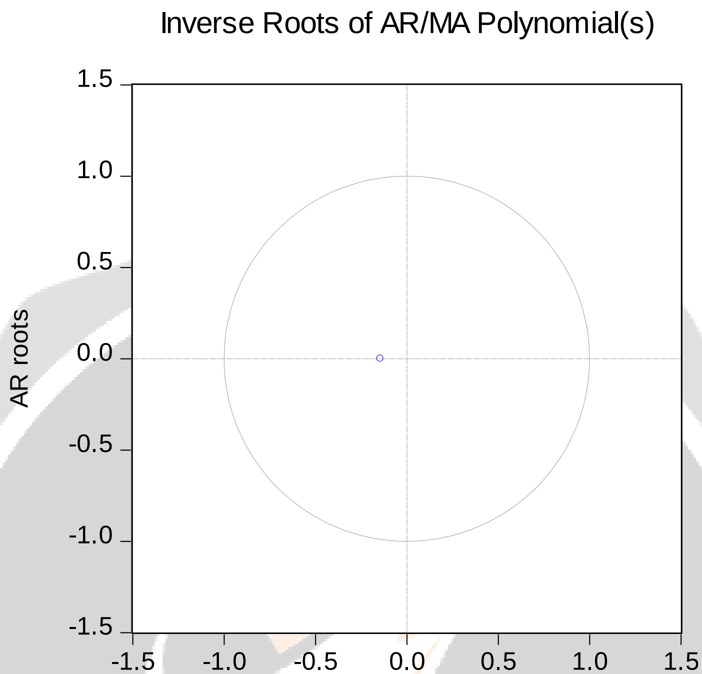


Figure 4 indicates that the estimated model is adequate since ACF and PACF lags are short and within the bands. This shows us that the "no autocorrelation" assumption is not violated in this study.

3.8.3 Stability Test of the ARIMA (1, 1, 0) Model

Figure 5: Inverse Roots



Since all the AR roots lie inside the unit circle, it implies that the estimated ARIMA process is (covariance) stationary; thus confirming that the ARIMA (1, 1, 0) model is indeed stable and suitable for forecasting annual number of people practicing open defecation in Mozambique.

4.0 FINDINGS

4.1 Descriptive Statistics

Table 8: Descriptive Statistics

Description	Statistic
Mean	43.389
Median	43.5
Minimum	27
Maximum	59
Standard deviation	10.059

As shown in table 8 above, the mean is positive, that is, 43.389. This means that, over the study period, the annual average number of people practicing open defecation in Mozambique is approximately 43% of the total population. This is a warning alarm for policy makers in Mozambique with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation in Mozambique over the study period is approximately 27% of the total population, while the maximum is 59% of the total population. In fact, the number of people practicing open defecation in Mozambique has continued to decline over the years from 59% in 2000 to 27% of the total population in 2017.

4.2 Results Presentation[1]

Table 9: Main Results

ARIMA (1, 1, 0) Model:

Guided by equation [4], the chosen optimal model, the ARIMA (1, 1, 0) can be expressed as follows:

$$\Delta ODZ_t = -1.88068 - 0.135598\Delta ODZ_{t-1} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots [5]$$

Variable	Coefficient	Standard Error	z	p-value
constant	-1.88068	0.0685764	-27.42	0.0000***
$\emptyset_1$	-0.135598	0.241076	-0.5625	0.5738

Table 9 shows the main results of the optimal ARIMA (1, 1, 0) model.

Forecast Graph

Figure 6: Forecast Graph – In & Out-of-Sample Forecasts

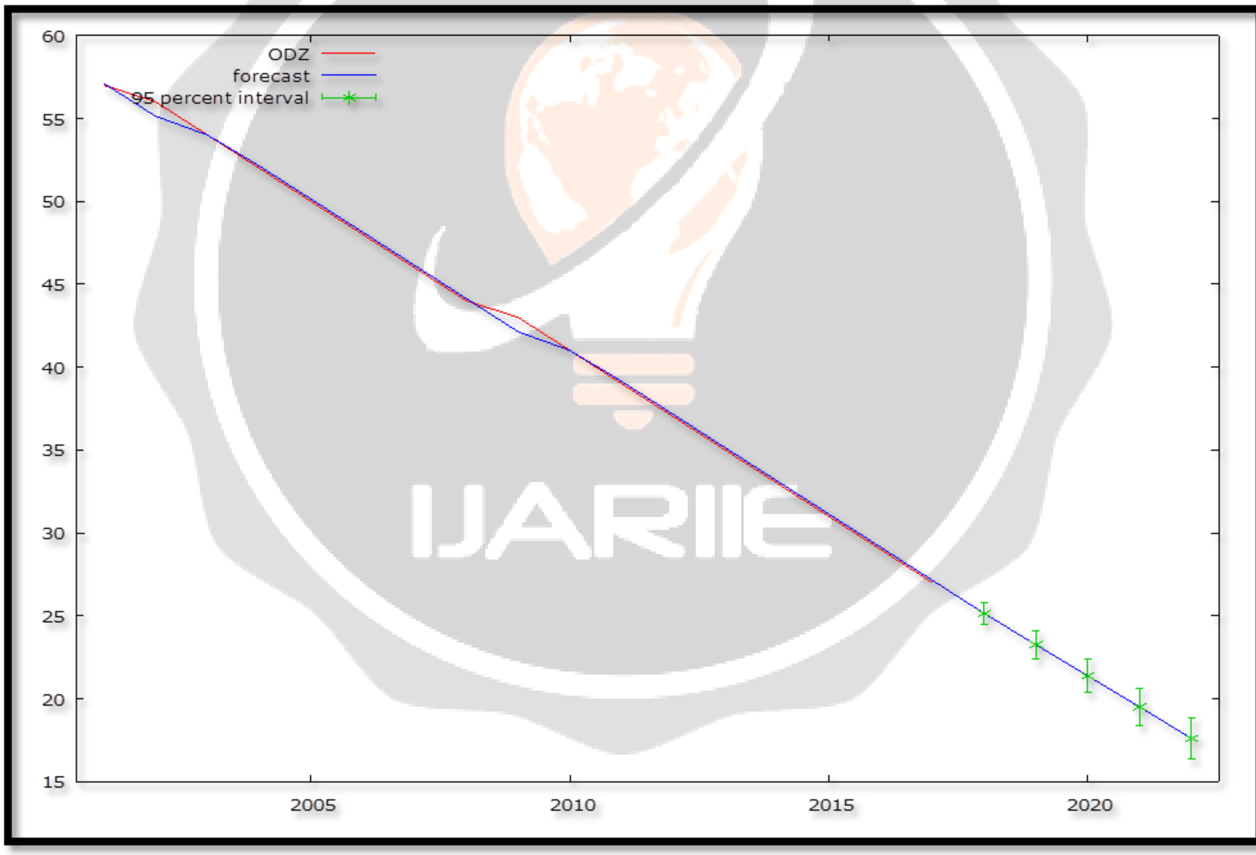


Figure 6 shows the in-and-out-of-sample forecasts of the ODZ series. The out-of-sample forecasts cover the period 2018 – 2022.

[1] *The *, ** and *** imply statistical significance at 10%, 5% and 1% levels of significance; respectively;*
$\emptyset_i = \beta_i$

Predicted ODZ – Out-of-Sample Forecasts Only

Table 10: Predicted ODZ

Year	Predicted ODZ	Standard Error	Lower Limit	Upper Limit
2018	25.14	0.319	24.51	25.76
2019	23.25	0.422	22.43	24.08
2020	21.37	0.507	20.38	22.37
2021	19.49	0.579	18.36	20.63
2022	17.61	0.644	16.35	18.87

Figure 7: Graphical Analysis of Out-of-Sample Forecasts

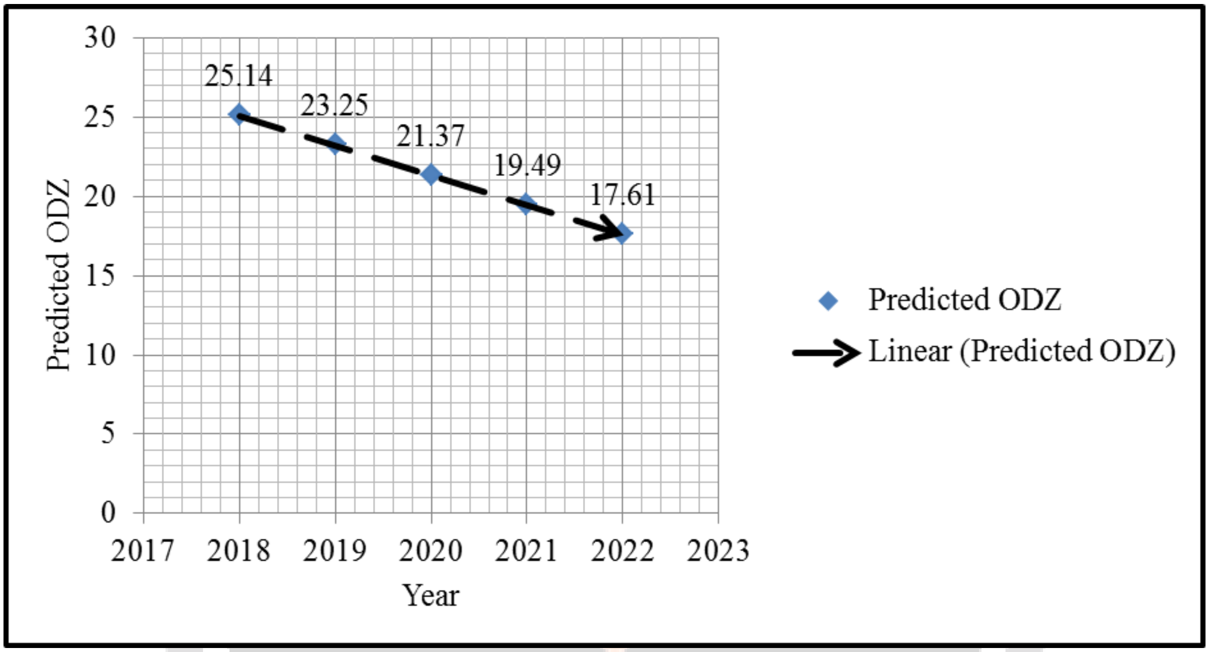


Table 10 and figure 7 show the out-of-sample forecasts only. The number of people practicing open defecation in Mozambique is projected to fall from approximately 25.14% in 2018 to as low as 17.61% of the total population by the year 2022. If the current trend is consistently maintained, Mozambique could achieve her ambitious goal of eliminating open defecation by 2025.

4.3 Policy Implications

i. The government of Mozambique should continue to make toilets a status symbol. In this regard, the government of Mozambique should provide funding for the construction of toilets in rural areas, especially for the poor rural communities.
ii. The government of Mozambique should continue to create more demand for sanitation through teaching the public on the importance of investing in toilets.
iii. The Mozambican government should continue to encourage a habit of systematic hand-washing, not defecating in the open, as well as keeping toilets fly-proof.

5.0 CONCLUSION

The study reveals that the ARIMA (1, 1, 0) model is not only stable but also the most suitable model to forecast the annual number of people practicing open defecation in Mozambique over the period 2018 – 2022. The model predicts a sharp decrease in the annual number of people practicing open defecation in Mozambique. Such a trend must be maintained and in this regard, a three-fold policy implication has been suggested. These findings are essential for the government of Mozambique, especially when it comes to long-term planning with regards to materializing the much needed open defecation free society.

REFERENCES

[1] Alhassan, A., & Anyarayor, B. K. (2018). Determinants of Adoption of Open Defecation Free (ODF) Innovations: A Case Study of Nadowli-Kaleo District, Ghana, *Journal of Development and Communication Studies*, 5 (2): 54 – 69.

[2] Ashenafi, T., *et al*. (2018). Latrine Utilization and Associated Factors Among Kebeles Declared Open Defecation Free in Wondo Genet District, South Ethiopia, 2015, *Journal of Health and Environmental Sciences*, 5 (5): 43 – 51.

[3] Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control, *Holden Day*, San Francisco.

[4] Gitau, M., & Flachenberg, F. (2016). Ensuring Availability and Sustainable Management of Water and Sanitation for All, *39th WEDC International Conference*, Kumasi.

[5] Nyoni, T (2018b). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, *Dimorian Review*, 5 (6): 16 – 40.

[6] Nyoni, T. (2018a). Modeling and Forecasting Naira/USD Exchange Rate in Nigeria: A Box-Jenkins ARIMA Approach, MPRA Paper No. 88622, *University Library of Munich*, Munich, Germany.

[7] Nyoni, T. (2018c). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, MPRA Paper No. 87737, *University Library of Munich*, Munich, Germany.