OPEN DEFECATION IN NAMIBIA: 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

In this study, which is the first of its kind in Namibia, the Box-Jenkins ARIMA model was applied in examining open defecation. The data was collected from the online World Bank data base and covers the period 2000 – 2017. The out-of-sample forecasts cover the period 2018 – 2022. The diagnostics tests employed in this study show that the open defection series under consideration is an I (1) variable. The study finally presents the ARIMA (1, 1, 0) as the optimal model in forecasting the number of people practicing open defecation in Namibia. The model, through the residual ADF tests, and the inverse roots of the AR/MA polynomials; has been shown to be very stable and suitable for forecasting and control of open defecation in Namibia. The results of the study indicate that by 2022 the number of open defecators will be approximately 47% of the total population. Indeed, open defecation is persistent in Namibia, especially in rural areas. The study offers a three-fold policy recommendation for consideration by the government of Namibia.

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

1.0 INTRODUCTION

Despite having achieved the Millennium Development Goal targets in drinking water, Namibia struggling to cope with its sanitation problems. At 34%, the country has the lowest levels of sanitation coverage in southern Africa, a situation that has not improved since 2006. Approximately, half of all Namibians practice open defecation, a rate that is one of the highest in Africa, just behind Somalia and South Sudan. Inequities in access to proper sanitation facilities are also glaring between rural and urban areas. In fact, the majority of people in rural areas have no choice but to defecate in the open, a practice that is highly unsanitary and harmful to health. Open defecation causes cholera, typhoid, hepatitis, polio, diarrhea, worm infestation, reduced physical growth, impaired cognitive function and malnutrition. Young children pay the price for open defecation. When they drink contaminated water; they get sick and quickly malnourished. In Namibia, 17% of children under the age of 5 suffer from diarrhea and repeated episodes of diarrhea contribute to the country's high levels of childhood stunting (UNICEF, 2016). The main of this research is to model and forecast the number of people practicing open defecation in Namibia. This study is important in helping policy makers in ending open defecation in the country. There is no doubt, the elimination of open defecation is essential for human development in the sense that it can decrease visits to health facilities, child deaths, as well as missed school and work days. Hence, the need for this study cannot be undermined.

2.0 LITERATURE REVIEW

In Thailand, Guterres *et al*. (2014) examined determinants that influence household to use and maintain latrines and basically found out that 47.2% of the households continued to use and maintain latrines and 52.8% had stopped by one year after the open defecation free declaration in Haupu village. Level of education is one of the most important factors seen to be influencing household to use and maintain latrines. In Nigeria, Abubakar (2017) investigated access to sanitation facilities and explored the socioeconomic and local factors that influence the type of facility used by households and found out that 44.2% of the households used various kinds of pit latrines, followed by toilets that flash to septic tanks (10.3%). Osumanu *et al*. (2019) examined sociocultural and economic factors determining open defecation in the Wa Municipality in Ghana. The study employed a mixed method approach involving questionnaire administration to 367 households systematically selected from 21 communities, observation, and eight key informant interviews. The mixed logit model was used to determine the factors that significantly influence open defecation. The findings basically revealed that 49.8% of the households had no form of toilet facility at home and were either using communal/public toilets or practicing open defecation. The study also revealed that six factors (education, household size, occupation income, traditional norms, and beliefs and owners of a toilet facility) were positively significant in determining open defecation. In another Nigerian study Nyoni (2019a) modeled and forecasted total population growth dynamics over the period 1960 – 2017 using the ARIMA approach and found out that annual total population in Nigeria is likely to continue rising sharply. The projected rise in total population in Nigeria is a real threat to natural resources in the country. Given the high levels of open defecation in the country, and the projected population explosion, Nigeria is likely to be in a worse scenario from a public health perspective as well as from a natural resource economics perspective.

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 ODN series under consideration.

3.2 The Moving Average (MA) model

Given:

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

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). ODN is the annual number of people (as a percentage of the total population) who practice open defecation in Namibia 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:

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

Where $\beta_1 \ldots \beta_p$ are estimation parameters, $ODN_{t-1} \ldots ODN_{t-p}$ are previous period values of the ODN 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:

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

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 ODN 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 used.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

A stochastic process ODN_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 $\Delta^d ODN_t$ satisfies an ARMA (p, q) process; then the sequence of ODN_t also satisfies the ARIMA (p, d, q) process such that:

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

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

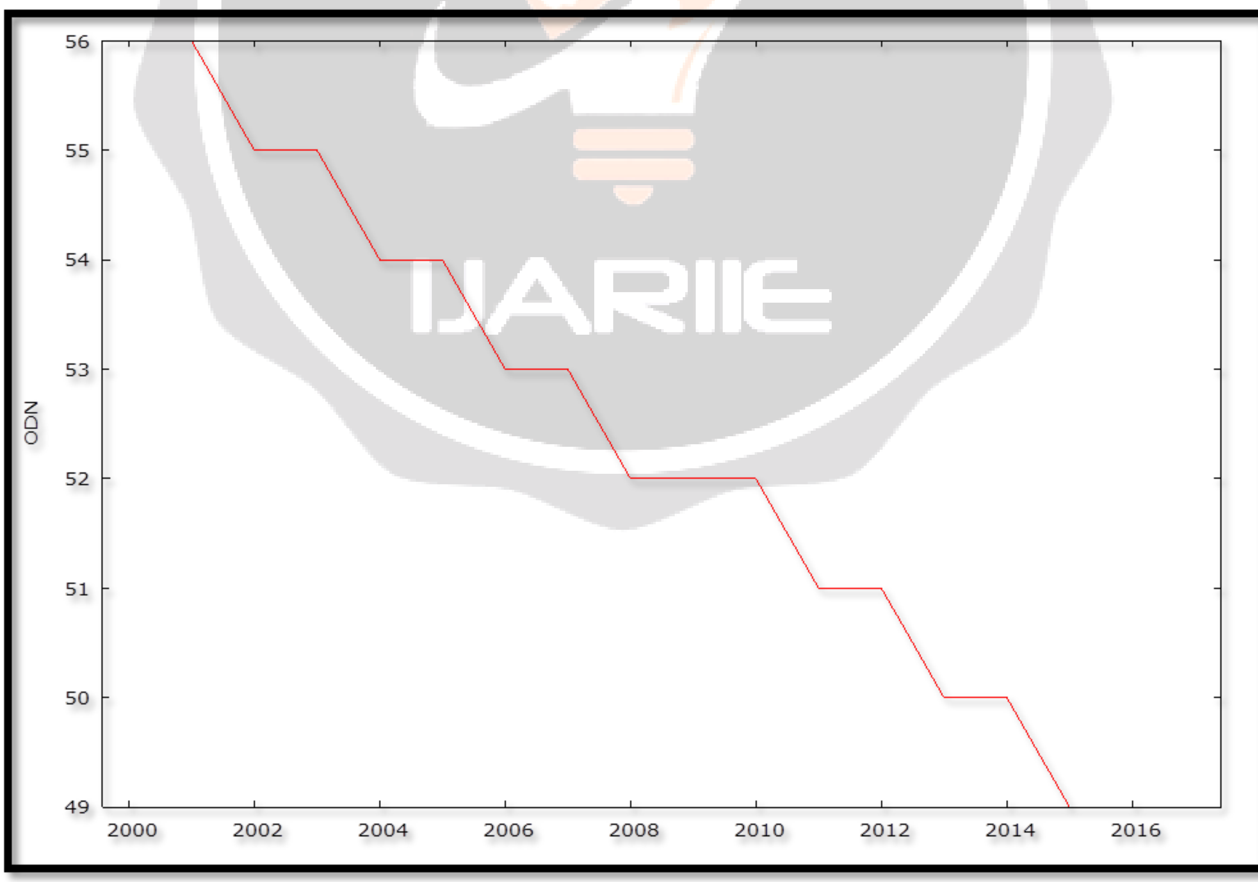
3.6 Data Collection

This study is based on annual observations (that is, from 2000 – 2017) on the number of people practicing Open Defecation [OD, denoted ODN] (as a percentage of total population) in Namibia. 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

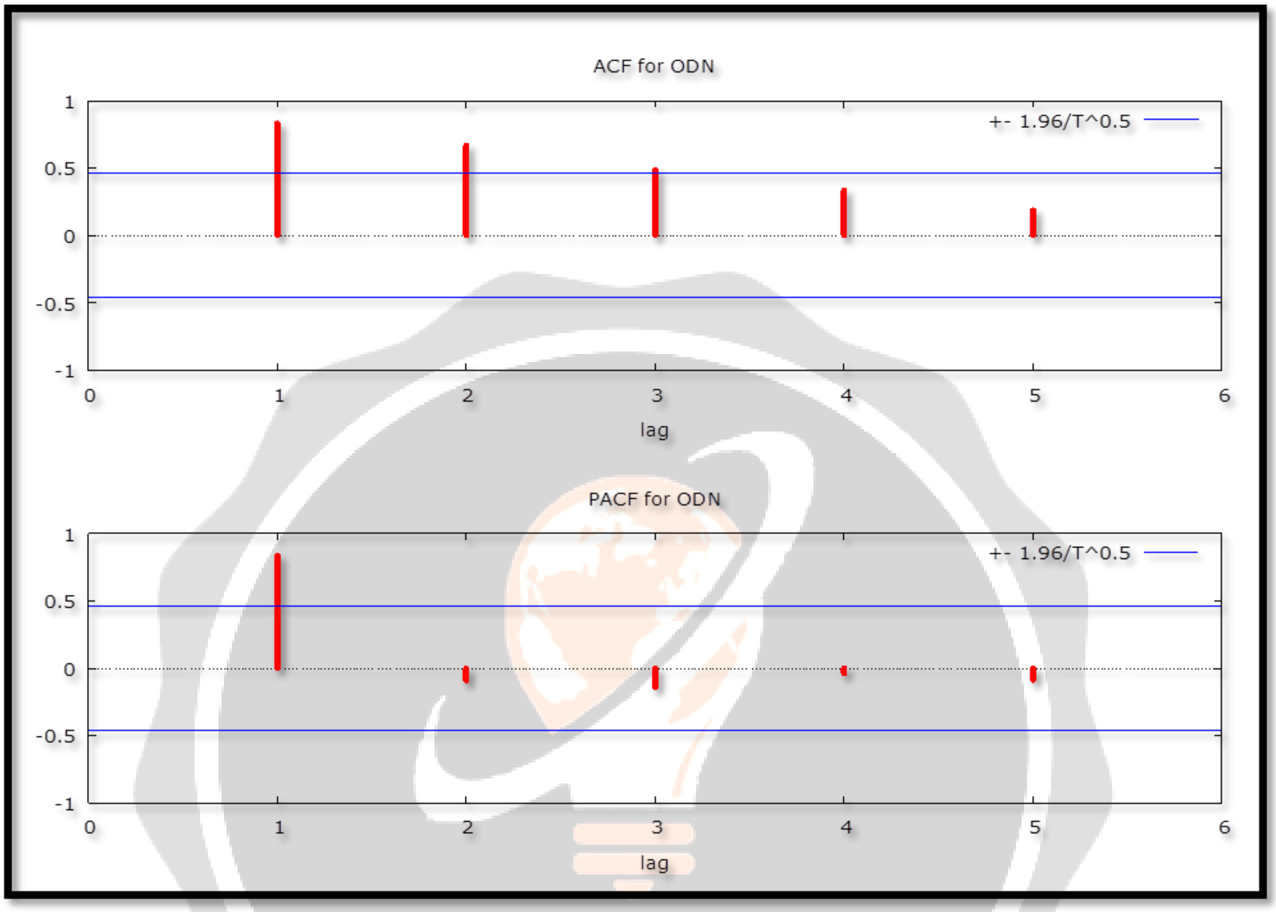
3.7.1 Stationarity Tests: Graphical Analysis

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
ODN	-1.445362	0.5341	-3.920350	@1%	Non-stationary
			-3.065585	@5%	Non-stationary
			-2.673459	@10%	Non-stationary

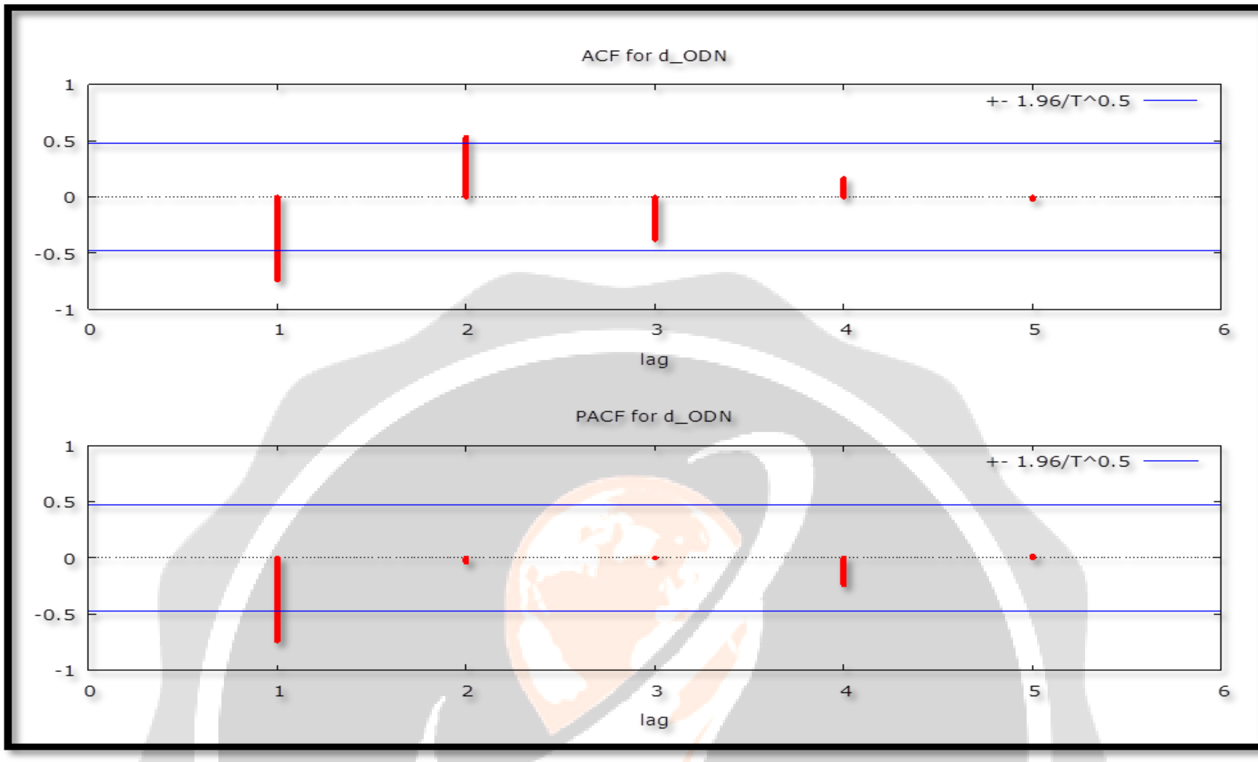
Table 2: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODN	-5.80947	0.0012	-4.616209	@1%	Stationary
			-3.710482	@5%	Stationary
			-3.297799	@10%	Stationary

Tables 1 and 2 show that ODN 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
ΔODN	-10.58301	0.0000	-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
ΔODN	-10.8739	0.0000	-4.667883	@1%	Stationary
			-3.733200	@5%	Stationary
			-3.310349	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that ODN 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)	**15.31592**	0.48901	0.01962	0.21202	0.32037	0.40981
ARIMA (2, 1, 0)	16.54375	0.47534	0.018376	0.20484	0.31337	0.39547
ARIMA (3, 1, 0)	18.54374	0.47534	0.018374	0.20483	0.31337	0.39547
ARIMA (4, 1, 0)	19.02959	0.44815	0.011672	0.19321	0.29888	0.3718

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. Therefore, the ARIMA (1, 1, 0) model is eventually chosen.

3.8 Residual & Stability Tests

3.8.1 ADF Test (*in levels*) of the Residuals of the ARIMA (2, 1, 0) Model

Table 6: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-3.599793	0.0185	-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	-3.631986	0.0593	-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. Hence, the model is stable.

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

Figure 4: Correlogram of the Residuals

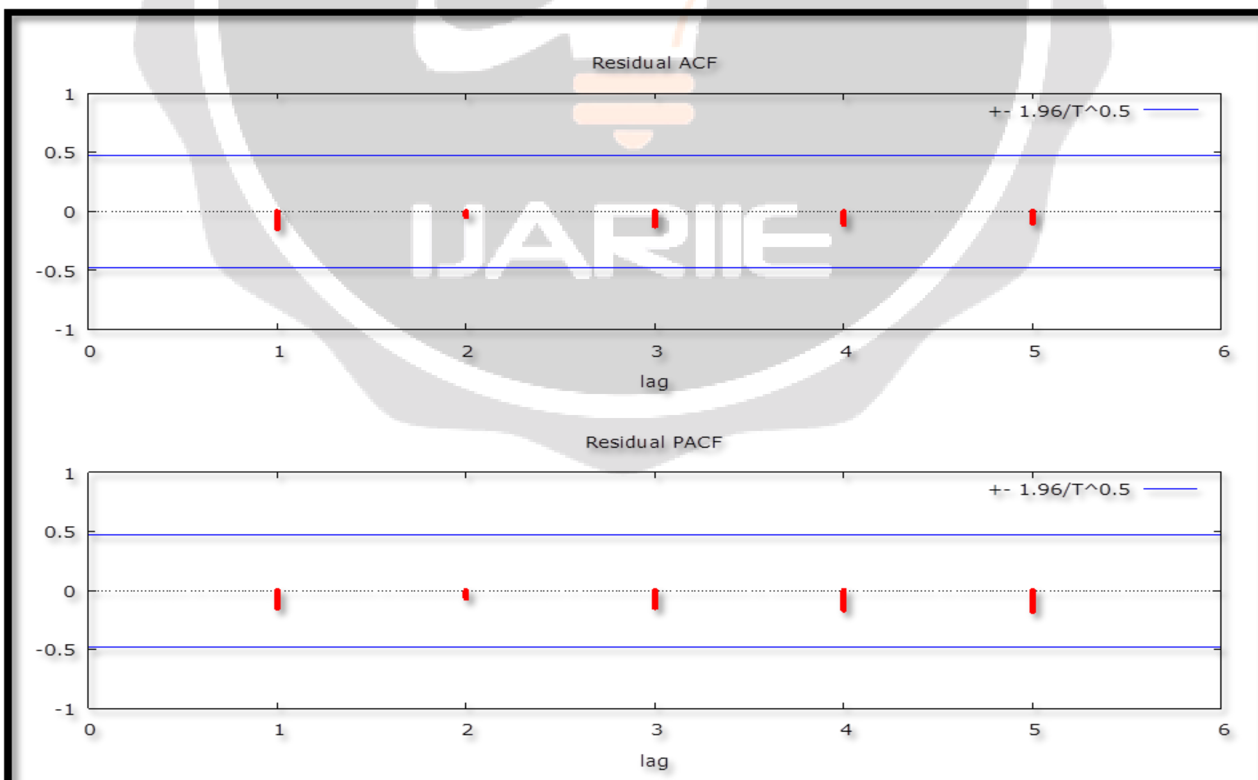
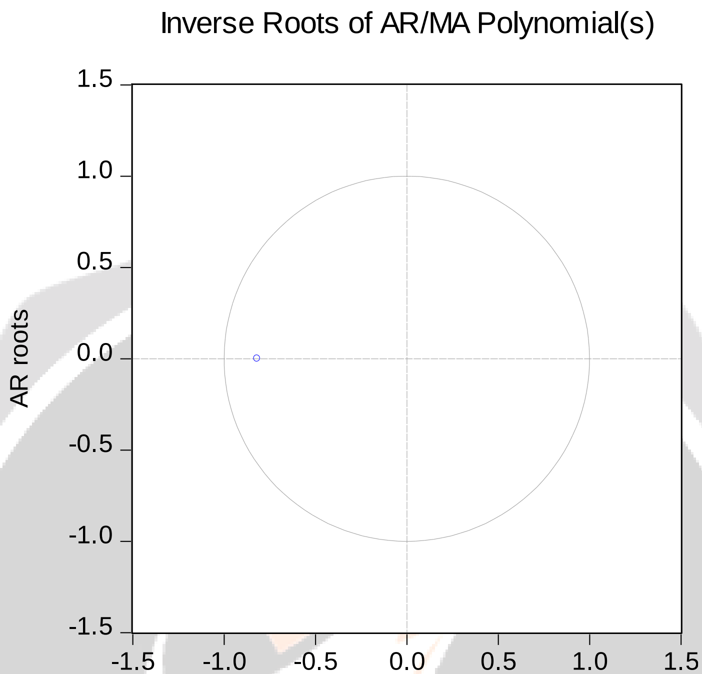


Figure 4 indicates that the estimated model is adequate since ACF and PACF lags are reasonably short and within the bands and this means that the "no autocorrelation" assumption is not violated in this study.

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 proving beyond any reasonable doubt that the ARIMA (1, 1, 0) model is stable and suitable for forecasting annual number of people practicing open defecation in Namibia.

4.0 FINDINGS

4.1 Descriptive Statistics

Table 8: Descriptive Statistics

Description	Statistic
Mean	52.278
Median	52
Minimum	49
Maximum	56

As shown in table 8 above, the mean is positive, that is, 52.278. This means that, over the study period, the annual average number of people practicing open defecation in Namibia is approximately 52% of the total population. This is a warning alarm for policy makers in Namibia with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation in Namibia over the study period is approximately 49% of the total population, while the maximum is 56% of the total population. The number of people practicing open defecation in Namibia has slightly declined over the years from 56% in 2000 to 49% of the total population in 2017.

4.2 Results Presentation[1]

Table 9: Main Results

ARIMA (1, 1, 0) Model:

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

$$\Delta ODN_t = -0.433952 - 0.768611\Delta OD_{t-1} \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots [5]$$

Variable	Coefficient	Standard Error	z	p-value
constant	-0.433952	0.0433879	-10	0.0000***
$\emptyset_1$	-0.768611	0.15622	-4.92	0.0000***

Table 9 shows the main results of the parsimonious 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 ODN series. The out-of-sample forecasts cover the period 2018 – 2022.

Predicted ODN – Out-of-Sample Forecasts Only

Table 10: Predicted ODN

Year	Predicted ODN	Standard Error	Lower Limit	Upper Limit
2018	48.23	0.31	47.62	48.84
2019	48.05	0.318	47.43	48.68
2020	47.42	0.408	46.62	48.22
2021	47.14	0.423	46.31	47.97
2022	46.59	0.478	45.65	47.53

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

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

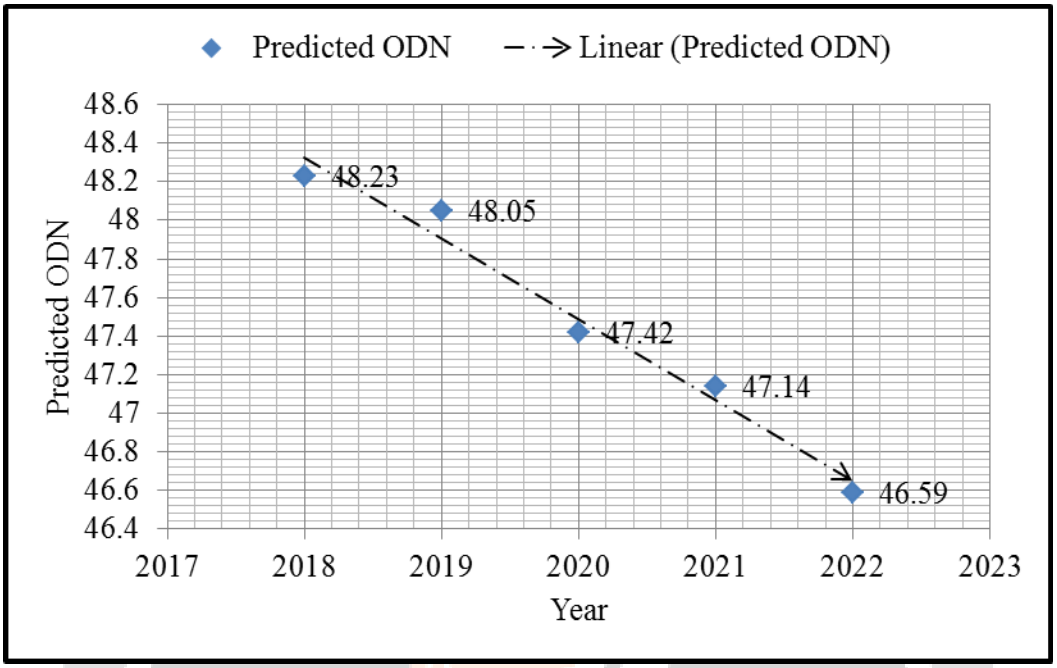


Table 10 and figure 7 show the out-of-sample forecasts only. The number of people practicing open defecation in Namibia is projected to fall from approximately 48% in 2018 to around 47% of the total population by the year 2022. This simply means that the country's main objective of "providing adequate sanitation to all Namibians" enshrined in the country's "Vision 2030" may not be achieved, especially if the below policy implications are not taken seriously.

4.3 Policy Implications

i. The government of Namibia should continue to make toilets a status symbol so that people stop thinking about toilets as "dark, dirty and smelly places" but rather consider toilets to be "rooms of happiness". In this regard, there is need to deliver sustainable hygiene behavior change in Namibia.
ii. The government of Namibia should create more demand for sanitation through teaching the public on the importance of investing in toilets.
iii. The government of Namibia should encourage a habit of systematic hand-washing, not defecating in the open, as well as keeping toilets fly-proof. In the same line of thought, the Namibian government should also construct and maintain water and sewage infrastructure in urban areas.

5.0 CONCLUSION

The study indicates 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 Namibia over the period 2018 – 2022. The model predicts a sharp but slight decrease in the annual number of people practicing open defecation in Namibia. These findings are essential for the government of Namibia, especially when it comes to long-term planning with regards to materializing the much needed open defecation free society.

REFERENCES

[1] Abubakar, I. R. (2017). Access to Sanitation Facilities Among Nigerian Households: Determinants and Sustainability Implications, *Sustainability*, 9 (547): 1 – 17.

[2] Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control, *Holden Day*, San Francisco.
[3] Guterres, I. C. L., Yamarat, K., & Xenos, P. (2014). Factors Influencing Households to Use Latrines After the Open Defecation Free Declaration in Ermera District, Timor-Leste, *Journal of Health Research*, 28 (3): 191 – 198.
[4] Nyoni, T (2018b). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, *Dimorian Review*, 5 (6): 16 – 40.
[5] Nyoni, T (2019a). Population Explosion in Nigeria By 2050: Overwhelming Empirical Evidence From The Box-Jenkins ARIMA Approach, *Dimorian Review*, 6 (2): 15 – 27.
[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.
[8] Osumanu, I. K., Kosoe, E. A., & Ategeeng, F. (2019). Determinants of Open Defecation in the Wa Municipality of Ghana: Empirical Findings Highlighting Sociocultural and Economic Dynamics Among Households, *Journal of Environmental and Public Health*, pp: 1 – 11.
[9] UNICEF (2016). A Namibia Fit for Children – Sanitation, *UNICEF*, New York.