NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 6, ISSUE 5, May -2020

OPEN DEFECATION IN ZIMBABWE: A BOX-JENKINS ARIMA APPROACH

DR. SMARTSON PUMULANI NYONI,
ZICHIRe, Harare, Zimbabwe

MR. THABANI NYONI
Department of Economics, University of Zimbabwe, Harare, Zimbabwe

ABSTRACT:

Employing annual time series data on the number of people who practice open defecation in Zimbabwe 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 applies the Box-Jenkins ARIMA approach. The diagnostic ADF tests show that the series under consideration is an I (1) variable. Based on the AIC, the study presents the ARIMA (3, 1, 0) model as the optimal model. The diagnostic tests further reveal that the presented parsimonious model is indeed stable and its residuals are not only stationary in levels but also normally distributed. The results of the study indicate that the number of people practicing open defecation in Zimbabwe is likely to decline over the period 2018 – 2022, from 25% to approximately 23.7% of the total population. In order to sustain this desirable downwards trend, the study suggested a three-fold policy recommendation to be put into consideration, especially by the government of Zimbabwe and its partners in water, sanitation and hygiene (WASH) related programmers.

INTRODUCTION:

Despite the international recognition of a right to sanitation, the content and progress of this right lags behind the right to water. In Sub-Saharan Africa, open defecation has actually increased over the last 20 years. Globally, 15% of the population still practices open defecation (Human Rights Watch, 2013). In Zimbabwe, especially in rural areas, many people still defecate in the open, especially in the bush (Morgan, 2010). In most of Zimbabwe's urban areas, including the capital city, Harare; people resort to open defecation because they are unable to flush their toilets as a result of lack of water, or their toilets were clogged and overflowing, rending the toilets unusable. Open defecation and poor sanitation make children and adults sick, which disrupts education and time at work, impacting on a community's development (Human Rights Watch, 2013) and the overall economic productivity of the country at large. Timely modeling and forecasting of the number of open defecators in the country is important in order to limit not only the prevalence of various diseases linked to human waste and lack of sanitation but also to reduce the occurrence of recurrent cholera outbreaks that have characterized Zimbabwe over the years. Closely monitoring open defecation trends in the country is also important for purposes of tracking the effectiveness of Water, Sanitation and Hygiene (WASH) programmers in reducing the number of open defecators in Zimbabwe.

1.2 OBJECTIVES OF THE STUDY:

i. To investigate open defecation trends in Zimbabwe over the period 2000 – 2017.
ii. To forecast the number of people practicing open defecation in the country for the period 2018 – 2022.
iii. To examine the trend of open defecation for the out-of-sample period.

LITERATURE REVIEW:

Whaley & Webster (2011) examined sanitation dynamics in Zimbabwe based on surveys, interviews and focus groups and basically found out that, a household's ability to own a latrine depends heavily on its ability to afford one. Whaley & Webster (2011) suggest that Zimbabweans, especially those in rural areas, practice open defecation largely because they do not afford to build sanitary toilets. Moyo & Moyo (2017) aimed to establish factors contributing to low sanitation and hygiene coverage as well as determining the knowledge, attitudes and cultural practices of the community members on sanitation and hygiene. The study adopted a qualitative research paradigm with the descriptive research design being preferred. Their study indicates that low sanitation and hygiene was due to poverty within the community and that the community under consideration had a negative attitude towards sanitation and hygiene programmes. In line with Whaley & Webster (2011), Moyo & Moyo (2017) suggest that open defecation in rural areas of Zimbabwe could be attributed to poverty and this is quite reasonable given the fact that poverty is rampant in the country. Kugedera & Machikicho (2017) examined the determinants of attaining Open Defecation Free (ODF) in four rural districts where it was implemented. The study found out that expecting subsidies from the project, having and enforcing community constitutions, existence of income savings, having active sanitation action groups and community health clubs were statistically significant factors associated with attainment of ODF status. Nyoni (2019) forecasted total population in Zimbabwe using the Box-Jenkins ARIMA technique based on annual time series data on total population in Zimbabwe from 1960 to 2017. The study presented the ARIMA (2, 2, 2) model and basically found out that total population in Zimbabwe will continue to rise in the next three decades, thereby posing a threat to both natural and non-renewable resources. This will be a worse threat if the open defecation problem is not addressed. No similar study has been done in Zimbabwe. Therefore, this study will be the first of its kind in the country and will go a long way in consolidating existing open defecation policy frameworks.

METHODODOLOGY:

3.1 The Box – Jenkins (1970) Methodology:

The first step towards model selection is to difference the series in order to achieve stationary. 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 judgment 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 used to analyze the Y series under consideration. The mathematical intuition behind the Box-Jenkins approach to modeling and forecasting is given below:

3.2 The Moving Average (MA) model:

Given:

$$Y_t = \sum_{i=1}^{q} \alpha_i L^i \mu_t + \mu_t \ldots\ldots\ldots\ldots\ldots\ldots.\ldots\ldots\ldots\ldots\ldots\ldots\ldots..[1]$$

where L is the lag operator.
or as:

$$OD_t = \alpha(L)\mu_t \ldots\ldots\ldots\ldots [2]$$

where:

$$\alpha(L)=\theta(L) \ldots\ldots\ldots\ldots [3]$$

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

$$Y_t = \sum_{i=1}^{p} \beta_i L^i Y_t + \mu_t \ldots\ldots\ldots\ldots [4]$$

Or that:

$$\beta(L)Y_t = \mu_t \ldots\ldots\ldots\ldots [5]$$

where:

$$\beta(L)=\phi(L) \ldots\ldots\ldots\ldots [6]$$

or that :

$$Y_t = (\beta_1 L + \cdots + \beta_p L^p)Y_t + \mu_t \ldots\ldots\ldots\ldots [7]$$

Where $\beta_1 \ldots \beta_p$ are estimation parameters, $Y_{t-1} \ldots Y_{t-p}$ are previous period values of the Y series and μ_t is as previously defined. Equation [4] 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 a combination of AR (p) and MA (q) processes. Thus, by combining equations [1] and [4]; an ARMA (p, q) process may be specified as shown below:

$$\phi(L)Y_t = \theta(L)\mu_t \ldots\ldots\ldots\ldots [8]$$

where $\phi(L)$ and $\theta(L)$ are polynomials of orders p and q respectively, algebraically defined as:

$$\phi(L) = 1 - \beta_1 L \ldots \beta_p L^p \ldots\ldots\ldots\ldots [9]$$

$$\theta(L) = 1 + \alpha_1 L + \cdots + \alpha_q L^q \ldots\ldots\ldots\ldots [10]$$

It is paramount to remember that the ARMA (p, q) model, just like the AR (p) and the MA (q) models; can only be applied for stationary time series data. However, in real life, many time series are non – stationary. In fact, in this study, the Y series has been found to be an I (1) variable (that is, it only became stationary after first differencing). Based on that simple reason, ARMA models are not suitable for modeling and forecasting non – stationary time series data. In such cases, the model described below is the one that should ideally be used.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model:

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

$$\Delta^d Y_t = \sum_{i=1}^{p} \beta_i \Delta^d L^i Y_t + \sum_{i=1}^{q} \alpha_i L^i \mu_t + \mu_t \ldots\ldots\ldots\ldots [11]$$

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 as variable Y] (as a percentage of total population) in Zimbabwe. 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:

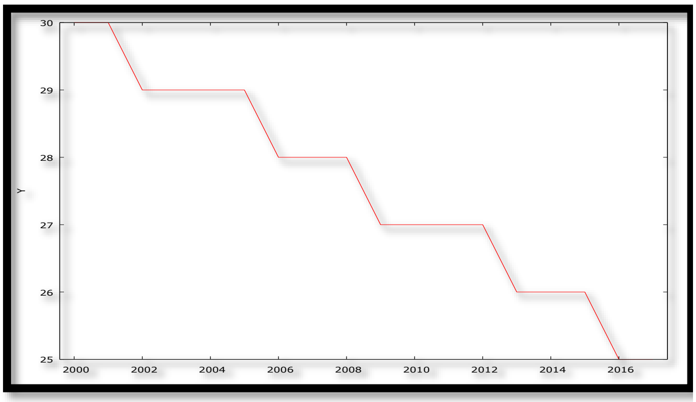

Figure 1

Figure 1 shows the time series plot of Y. Important to note is the fact that Y is downwards trending over the period under study. Hence, it is reasonable to suspect that Y is non-stationary in levels. The Augmented-Dickey-Fuller (ADF) test will be carried out in order to verify the level of stationarity of Y.

3.7.2 The Correlogram in Levels:

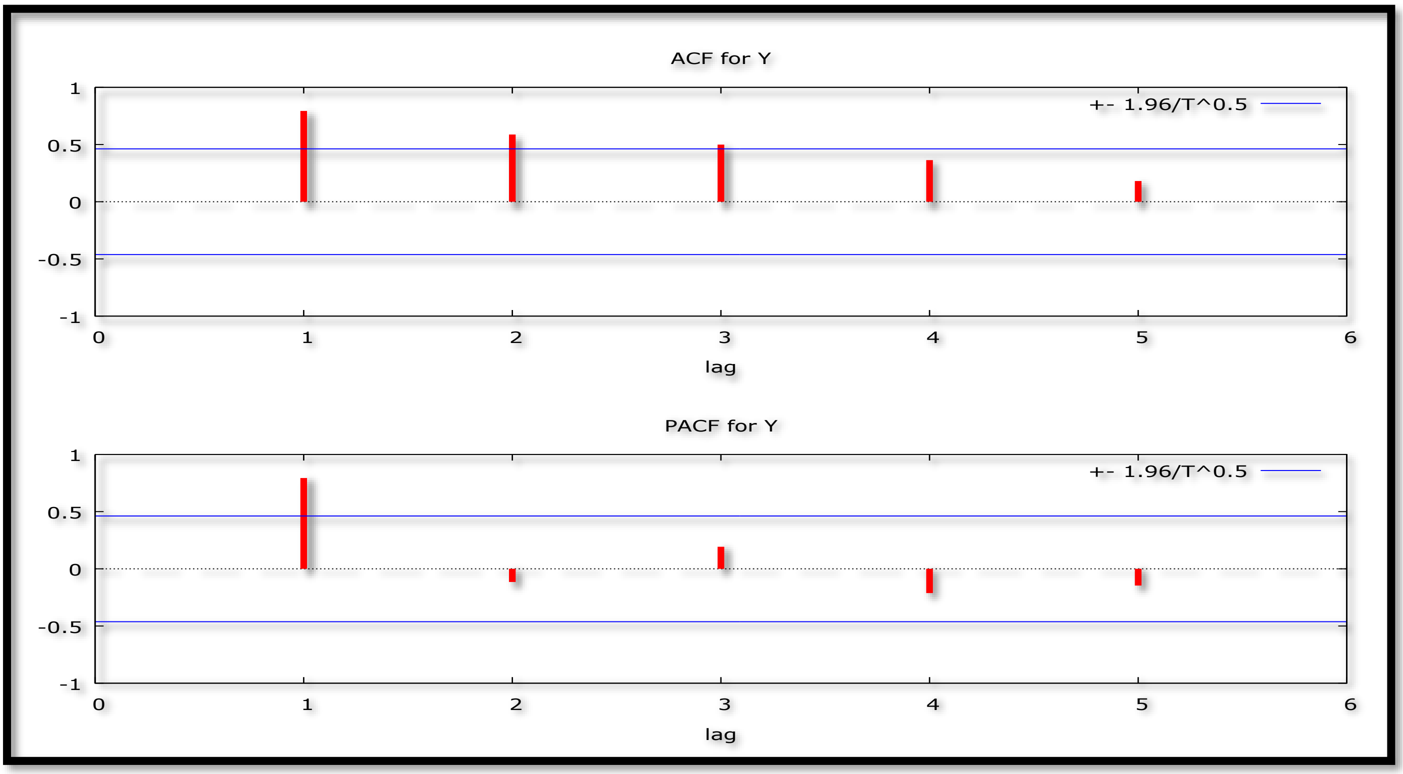

Figure 2: Correlogram in Levels

Figure 2 shows the correlogram of Y and basically points to the fact that Y is not an I (1) variable as shown by the ACF and PACF lags which are long as to be outside of the 5% confidence intervals. Hence, figure 2 is consistent with figure 1.

3.7.3 The ADF Test in Levels:

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
Y	0.750000	0.9885	-4.004425	@1%	Non-stationary
			-3.098896	@5%	Non-stationary
			-2.690439	@10%	Non-stationary

Table 2: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
Y	-5.439413	0.0027	-4.667883	@1%	Stationary
			-3.7332	@5%	Stationary

		00		
		-3.310349	@10%	Stationary

Table 1, unlike table 2, shows that Y is not stationary in levels as already suggested by figures 1 and 2.

3.7.4 The Correlogram (at First Differences):

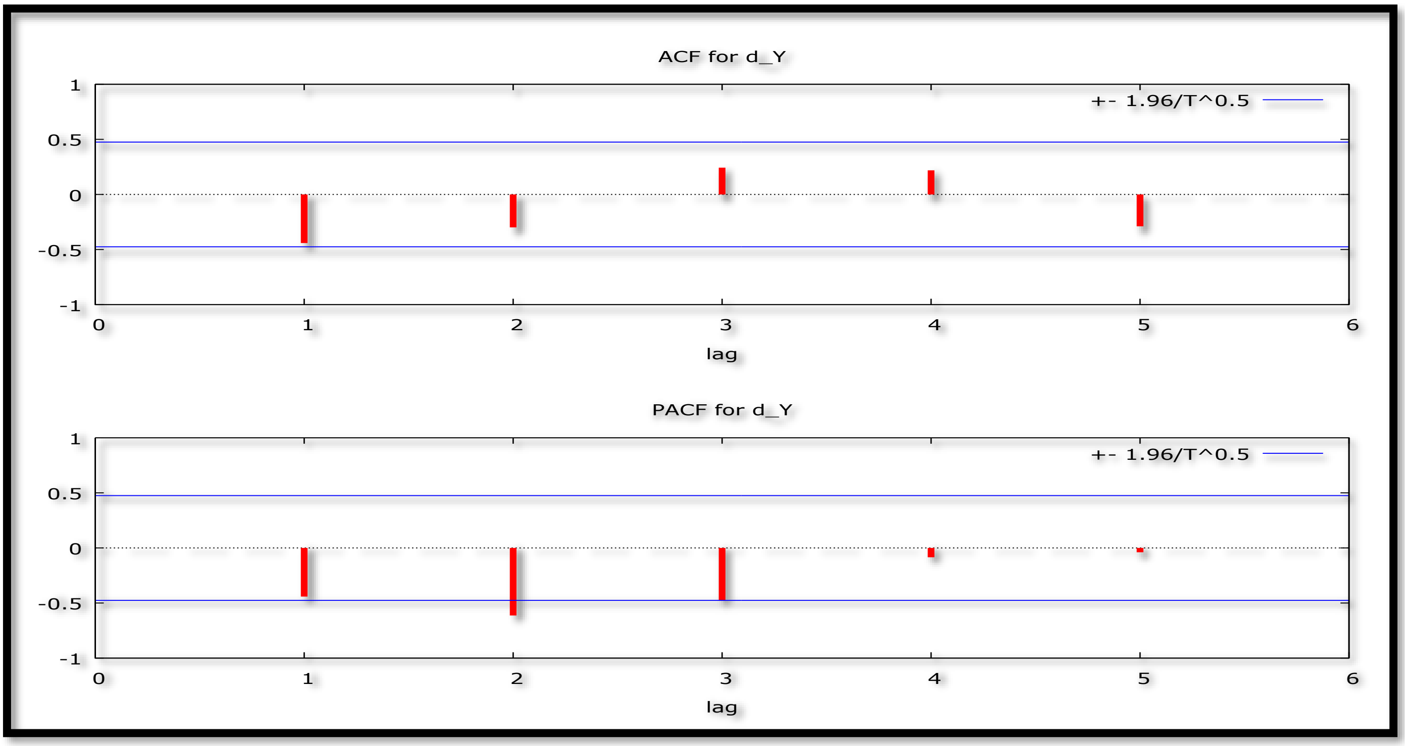

Figure 3: Correlogram (at First Differences)

3.7.5 The ADF Test (at First Differences):

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔY	-4.830459	0.0023	-4.004425	@1%	Stationary
			-3.098896	@5%	Stationary
			-2.690439	@10%	Stationary

Table 4: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔY	-4.828758	0.0096	-4.800080	@1%	Stationary
			-3.791172	@5%	Stationary
			-3.342253	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that Y 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)	23.8276	0.74223	0.0078946	0.33809	0.4075	1.2343
ARIMA (2, 1, 0)	17.07814	0.57425	-0.010058	0.20643	0.32274	0.74462
ARIMA (3, 1, 0)	**15.6748**	0.5179	-0.014051	0.17041	0.29111	0.61651
ARIMA (4, 1, 0)	17.06748	0.5179	-0.014051	0.17041	0.29111	0.61651
ARIMA (5, 1, 0)	19.06748	0.5179	-0.014045	0.1704	0.29111	0.61649

A model with a lower AIC value is better than the one with a higher AIC value (Nyoni,

NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 6, ISSUE 5, May -2020

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 (3, 1, 0) model is eventually chosen.

3.8 Residual & Stability Tests:

3.8.1 ADF Test (in levels) of the Residuals of the ARIMA (3, 1, 0) Model:

Table 6: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-6.190560	0.0003	-4.057910	@1%	Stationary
			-3.119910	@5%	Stationary
			-2.701103	@10%	Stationary

Table 7: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-5.439413	0.0027	-4.667883	@1%	Stationary
			-3.733200	@5%	Stationary
			-3.310349	@10%	Stationary

Tables 6 and 7 indicate that the residuals of the chosen optimal model, the ARIMA (3, 1, 0) model; are stationary. Hence, the model is very stable.

3.8.2 Correlogram of the Residuals of the ARIMA (3, 1, 0) Model:

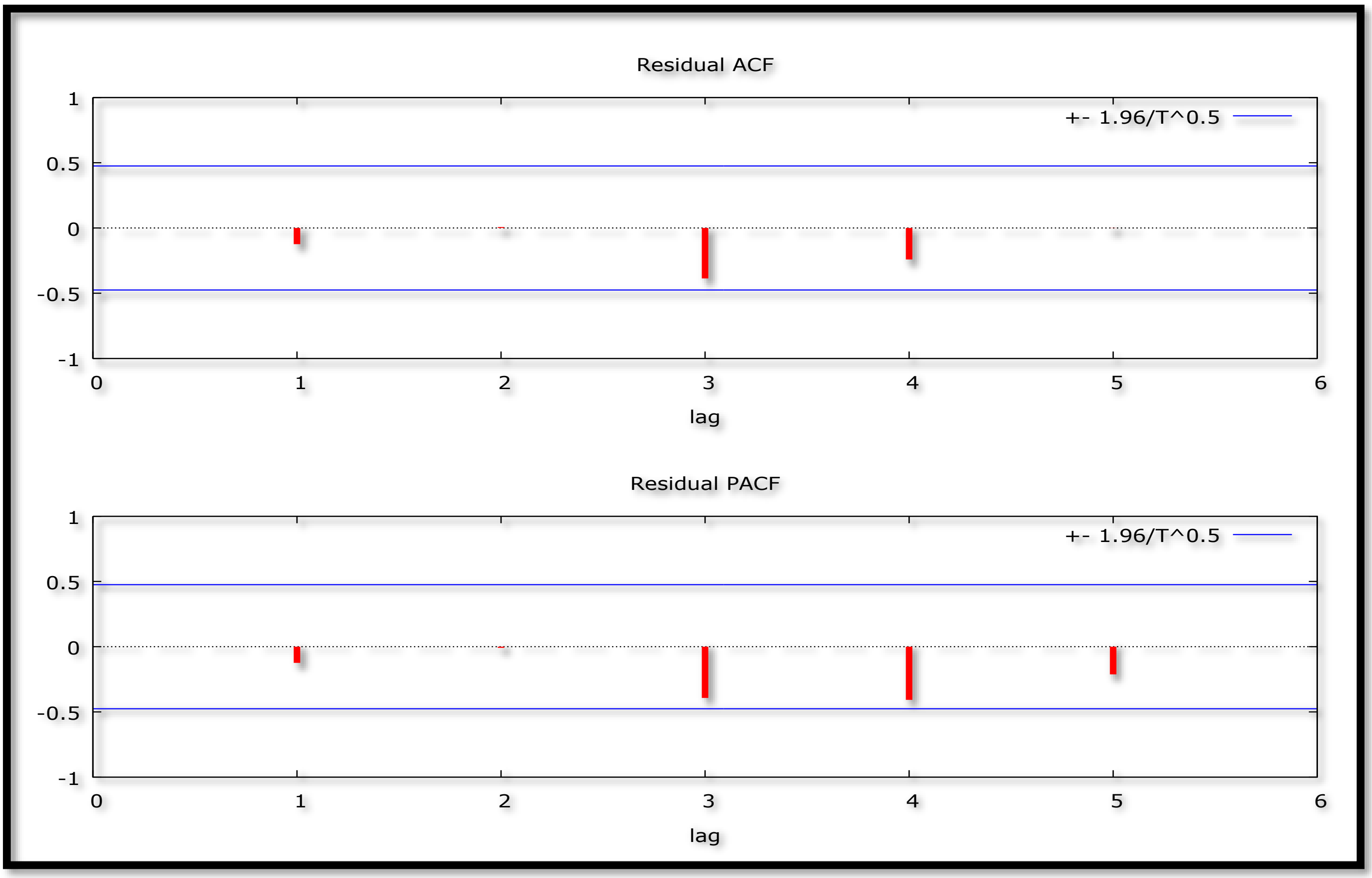

Figure 4: Correlogram of the Residuals

Figure 4 indicates that the estimated ARIMA (3, 1, 0) model is adequate since ACF and PACF lags are quite short and within the bands and this means that the "no autocorrelation" assumption is not violated in this study.

3.8.3 Normality Test of the Residuals of the ARIMA (3, 1, 0) Model:

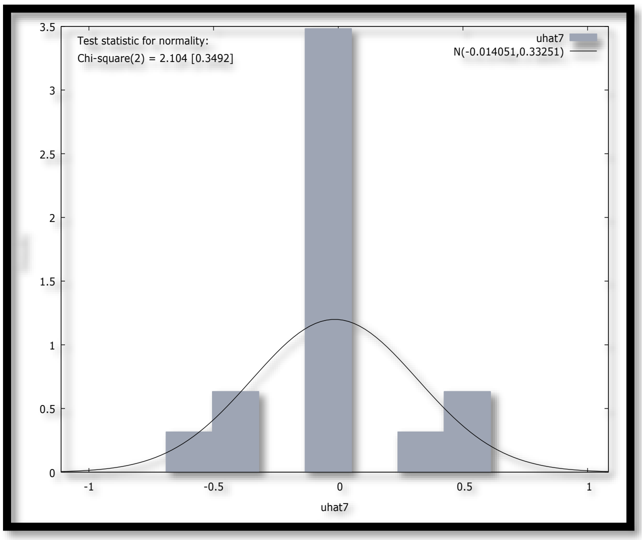

Figure 5: Normality Test

Figure 5 shows that the residuals of the applied model are normally distributed as confirmed by the insignificant p-value of the Chi-square statistic.

3.8.4 Stability Test of the ARIMA (3, 1, 0) Model:

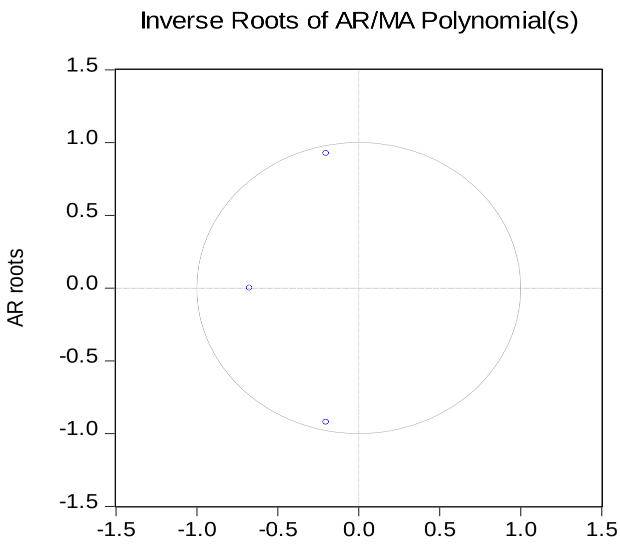

Figure 6: Inverse Roots

Since all the AR roots lie inside the unit circle, it implies that the estimated ARIMA process is (covariance) stationary; thus indicating that the ARIMA (3, 1, 0) model is stable and suitable for forecasting annual number of people practicing open defecation in Zimbabwe.

FINDINGS:

4.1 Descriptive Statistics:

Table 8: Descriptive Statistics

Description	Statistic
Mean	27.556
Median	27.5
Minimum	25
Maximum	30
Standard deviation	1.5801
Skewness	-0.058719
Excess kurtosis	-1.0797

As shown in table 8 above, the mean is positive, that is, approximately 28. This means that, over the study period, the annual average number of people practicing open defecation in Zimbabwe is approximately 28% of the total population. This is a warning signal for Zimbabwe policy makers with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation in Zimbabwe over the study period is approximately 25% of the total population, while the maximum is 30% of the total population. In fact, the number of people practicing open defecation in Zimbabwe has continued to fall over the years from 30% in 2000 to 25% of the total population in 2017. The skewness statistic is -0.058719 and the most important characteristic is that it is negative, meaning that the Y series is negatively skewed and non-symmetric. Excess kurtosis is -1.0797; showing that the Y series is not normally distributed.

4.2 Results Presentation:

Table 9: Main Results

ARIMA (3, 1, 0) Model:

Guided by equation [11], the chosen optimal model, the ARIMA (3, 1, 0) model can be expressed as follows:

$$\Delta Y_t = -0.287151 - 1.00194\Delta Y_{t-1} - 1.00359\Delta Y_{t-2} - 0.484684\Delta Y_{t-3} \ldots\ldots\ldots\ldots\ldots [12]$$

Variable	Coefficient	Standard Error	z	p-value
constant	-0.287151	0.0197447	-14.54	0.0000***
β_1	-1.00194	0.216710	-4.623	0.0000****
β_2	-1.00359	0.225263	-4.455	0.0000***
β_3	-0.484684	0.227954	-2.126	0.0335**

Table 9 shows the main results of the ARIMA (3, 1, 0) model.

Forecast Graph:

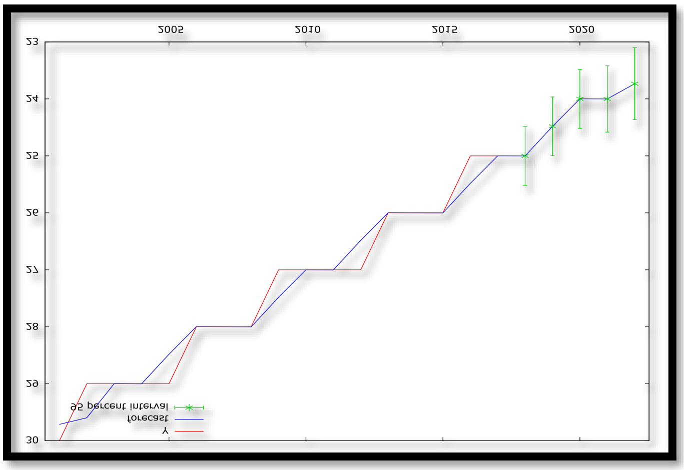

Figure 7: Forecast Graph – In & Out-of-Sample Forecasts

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

Predicted Y – Out-of-Sample Forecasts Only:

Table 10: Predicted Y

Year	Predicted Y	Standard Error	Lower Limit	Upper Limit
2018	25	0.263	24.49	25.52
2019	24.48	0.263	23.97	25
2020	24	0.263	23.48	24.51
2021	24	0.297	23.42	24.58
2022	23.73	0.323	23.1	24.37

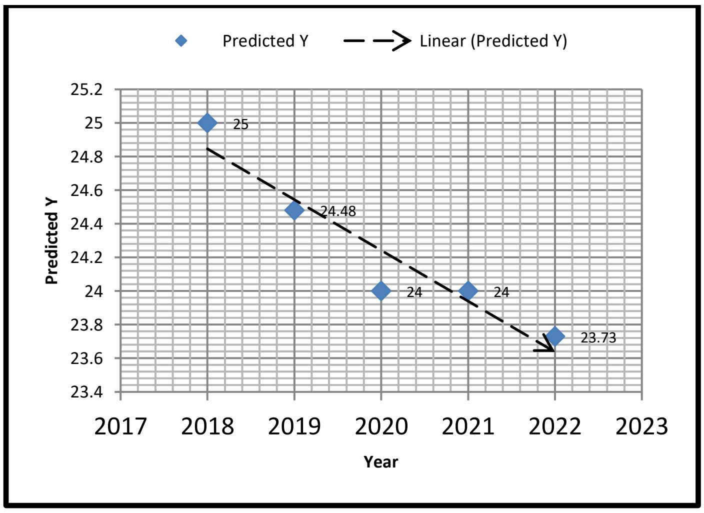

Figure 8: Graphical Analysis of Out-of-Sample Forecasts

Table 10 and figure 8 show the out-of-sample forecasts only. The number of people practicing open defecation in Zimbabwe is projected to fall from approximately 25% in 2018 to around 23.73% of the total population by the year 2022.

4.3 Policy Implications:

i. The government of Zimbabwe should create more demand for sanitation through teaching the public (that is, public awareness campaigns) on the importance of investing in toilets, particularly in light of disease transmission and other risks associated with open defecation.
ii. There is need for the government of Zimbabwe to encourage a habit of not defecating in the open, as well as keeping toilets fly-proof.
iii. The government of Zimbabwe should channel adequate financial resources towards funding open-defecation-related projects and initiatives around the country.
iv. The government of Zimbabwe, through local authorities should repair and maintain sewage infrastructure in urban areas, for example, replacement of burst and old pipes as well as provision of water in adequate amounts.
v. Religious organizations or churches should be provided church stands by local authorities; and in those church stands, availability of water, sanitation and hygiene facilities should be mandatory.

CONCLUSION:

The study shows that the ARIMA (3, 1, 0) model is not only stable but also the most suitable model to forecast the annual number of people practicing open defecation in Zimbabwe over the period 2018 – 2022. The model predicts a decrease in the annual number of people practicing open defecation in

Zimbabwe. The predicted trend can become a reality, especially if the suggested policy directions are put into consideration. The findings of this endeavor are essential for the government of Zimbabwe, especially when it comes to long-term planning with regards to materializing the much needed open defecation free society.

REFERENCES:

1) Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control, Holden Day, San Francisco.
2) Human Rights Watch (2013). Troubled Water: Burst Pipes, Contaminated Wells and Open Defecation in Zimbabwe's Capital, Human Rights Watch, New York.
3) Kugedera, Z., & Machikicho, J. T. (2017). Achieving and Sustaining Open Defecation Free (ODF) Villages: A Study of Four Rural Districts in Zimbabwe, 40th WEDC International Conference, and Lough borough.
4) Morgan, P. (2010). Zimbabwe's Rural Sanitation Programme: An Overview of the Main Events, Government of Zimbabwe, Harare.
5) Moyo, W., & Moyo, M (2017). Factors Contributing to Low Sanitation and Hygiene Coverage: A Case of Blair Ventilated Improved Pit Latrines in Matetsi Ward, Hwange District, International Journal of Academic Research in Business and Social Sciences, 7 (3): 180 – 191.
6) Nyoni, T (2018b). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, Dimorian Review, 5 (6): 16 – 40.
7) Nyoni, T (2019). The Population Question in Zimbabwe: Reliable Projections From The Box-Jenkins ARIMA Approach, MPRA Paper No. 96791, University Library of Munich, Germany.
8) 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.
9) Nyoni, T. (2018c). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, MPRA Paper No. 87737, University Library of Munich, Munich, Germany.
10) Whaley, L., & Webster, J. (2011). The Effectiveness and Sustainability of Two Demand-Driven Sanitation and Hygiene Approaches in Zimbabwe, Journal of Water, Sanitation and Hygiene for Development, 1 (1): 20 – 36.