NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 6, ISSUE 11, Nov. -2020

OPEN DEFECATION IN ZAMBIA: A BOX-JENKINS ARIMA APPROACH

DR. SMARTSON. P. NYONI
ZICHIRe Project, University of Zimbabwe, Harare, Zimbabwe

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

ABSTRACT:

Using annual time series data on the number of people who practice open defecation in Zambia from 2000 – 2017, the study predicts the annual number of people who will still be practicing open defecation over the period 2018 – 2021. The study applies the Box-Jenkins ARIMA methodology. The diagnostic ADF tests show that the K series under consideration is an I (1) variable. Based on the AIC, the study presents the ARIMA (1, 1, 0) model as the optimal model. The diagnostic tests further reveal that the presented model is quite stable and its residuals are stationary in levels and also normally distributed. The results of the study indicate that the number of people practicing open defecation in Zambia is likely to decline, but generally, slightly, over the period 2018 – 2022, from approximately 18.9% to almost 17.7% of the total population. The study finally put forward a 3-fold policy recommendation to be put into consideration, especially by the Zambian government.

INTRODUCTION:

Zambia is a landlocked country in Sub-Saharan Africa with one of the lowest rates of access to safe water sources in the world (WSP, 2012). The country has a population of approximately 13 million people; 61% of whom reside in rural areas (USAID, 2006). On average, 4.8 million Zambians do not have access to clean water. Fecal-oral illnesses are common in Zambia, as open defecation is widely practiced, especially in the Eastern province (WaterAid, 2010). Approximately 8700 Zambians, including 6600 children under 5, die each year from diarrhea (WSP, 2012). In fact, fecal contamination of the environment is the root cause of an annual average of 3200 cases of cholera affecting Zambia (ibid). Furthermore, children are particularly exposed to improper Water, Sanitation and Hygiene (WASH) facilities in diarrheal diseases being one of the leading causes of child mortality (Lozano et al., 2013). Open defecation costs the Zambia approximately US$71 million per year and yet eliminating the practice would require less than 420000 latrines to be built and used (WSP, 2012). Thus, it has become even more instructive for researchers to model and forecast the number of people practicing open defecation in order to formulate evidence-driven policies to end open defecation. The main purpose of this study is to predict the annual number of open defecators in Zambia over the period 2018 – 2021. This study, besides being the first of its kind in the case of Zambia, will go a long way in uncovering the possibility of ending open defecation in the country.

1.2 OBJECTIVES OF THE STUDY

i. To investigate the years during which open defection was practiced by people more than 19% of the total population in Zambia.
ii. To forecast the number of people practicing open defecation in Zambia for the period 2018 – 2021.

iii. To examine the trend of open defecation in Zambia for the out-of-sample period.

LITERATURE REVIEW:

In study carried out in Burkina Faso, Klutse et al. (2010) compared the capital expenditure and the operational and maintenance expenditure for sanitation facilities in rural and peri-urban areas and basically found that the pit latrine is not promoted in Burkina Faso, but it is the toilet used by the vast majority of those who have access to one. In a Ghananian study, Alhassan & Anyarayor (2018) looked at 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 carried out to collect data. The research revealed that while effective communication of innovation resulted in widespread awareness, low income levels significantly accounted for households' inability to sustain and utilize latrines. In a Zambian study, Kojo et al. (2019) examined Community Led Total Sanitation (CLTS) and basically found out that the CLTS initiative in Zambia led to improvements in access to improved sanitation facilities, reduced open defecation, and better hand-washing practices. This piece of work adopts the ARIMA method in analyzing open defecation trends in Zambia and is apparently the first of its kind in the country.

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 used to analyze the K series under consideration.

3.2 The Moving Average (MA) model:

Given:

$$K_t = \sum_{i=1}^{q} \alpha_i \mu_{t-i} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad [1]$$

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

$$K_t = \sum_{i=1}^{p} \beta_i K_{t-i} + \mu_t \ldots\ldots\ldots\ldots\ldots\ldots [2]$$

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

$$K_t = \sum_{i=1}^{p} \beta_i K_{t-i} + \sum_{i=1}^{q} \alpha_i \mu_{t-i} + \mu_t \ldots \ldots \ldots \ldots \ldots [3]$$

3.5 The Autoregressive Integrated Moving Average (ARIMA) model:

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

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

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 as K] (as a percentage of total population) in Zambia. Out-of-sample forecasts will cover the period 2018 – 2021. All the data was gathered from the World Bank online database.

3.7 Diagnostic Tests & Model Evaluation:

3.7.1 Stationarity Tests: Graphical Analysis:

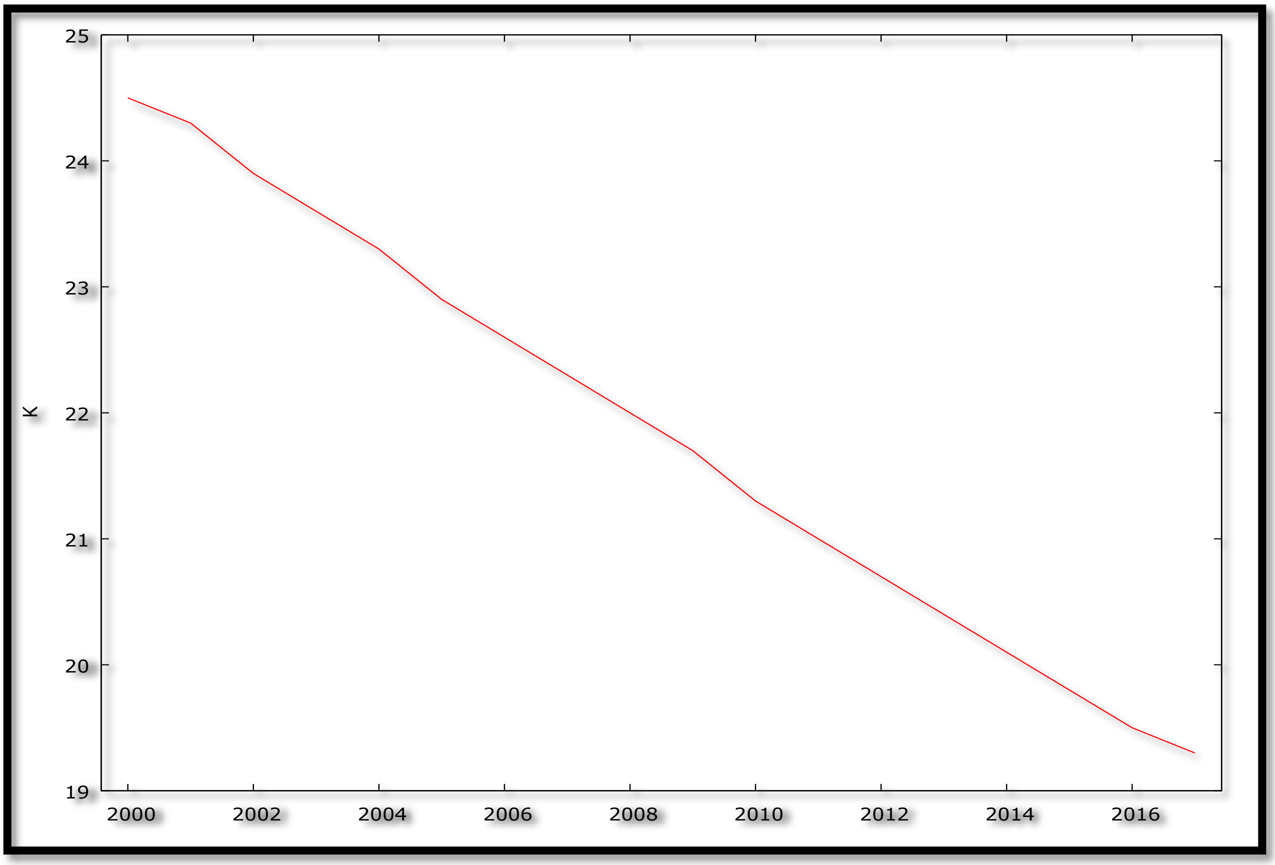

Figure 1

NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 6, ISSUE 11, Nov. -2020

3.7.2 The Correlogram in Levels

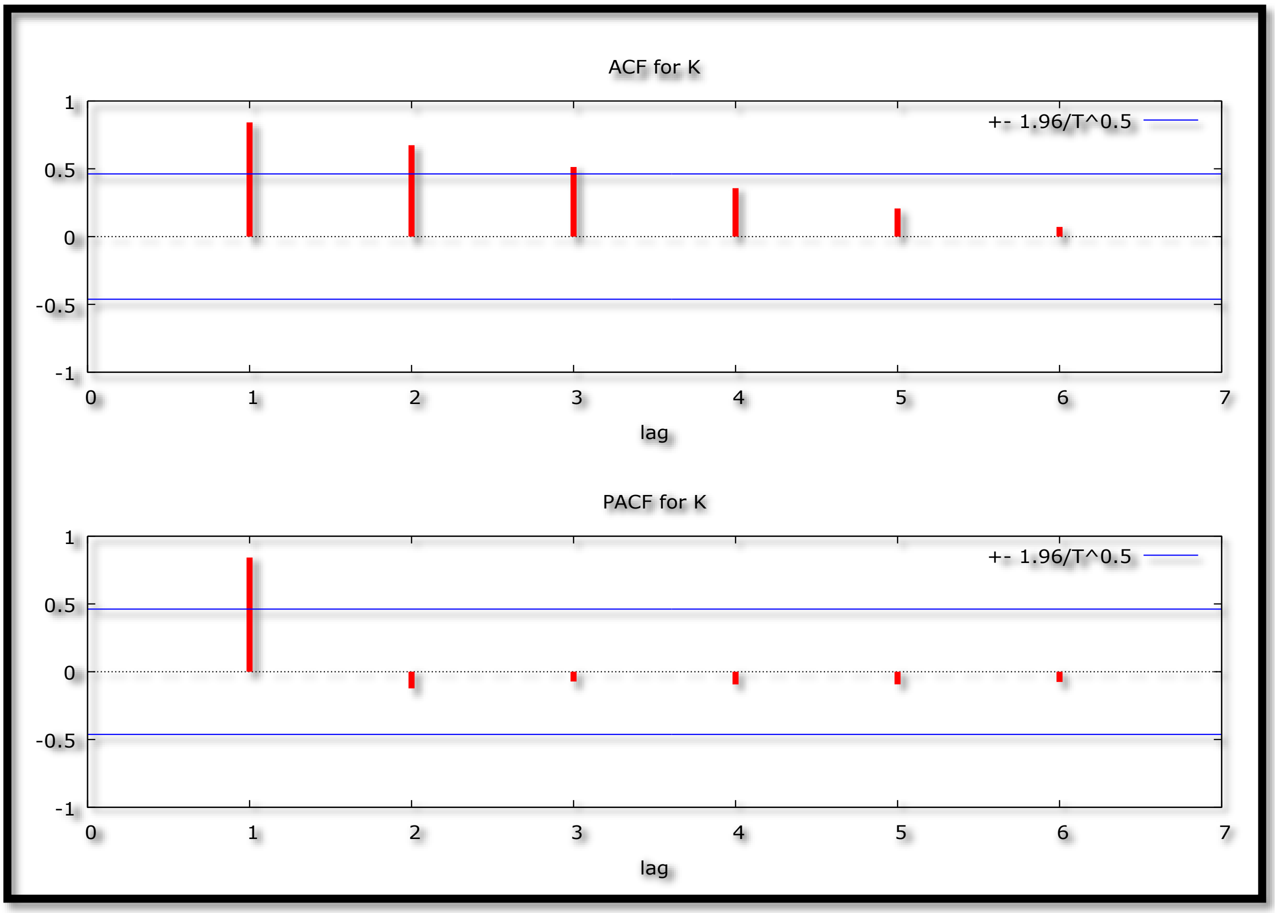

Figure 2: Correlogram in Levels

3.7.3 The ADF Test in Levels

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
K	-0.922593	0.7551	-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
K	-1.630215	0.7368	-4.616209	@1%	Non-stationary
			-3.710482	@5%	Non-stationary
			-3.297799	@10%	Non-stationary

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

3.7.4 The Correlogram (at First Differences)

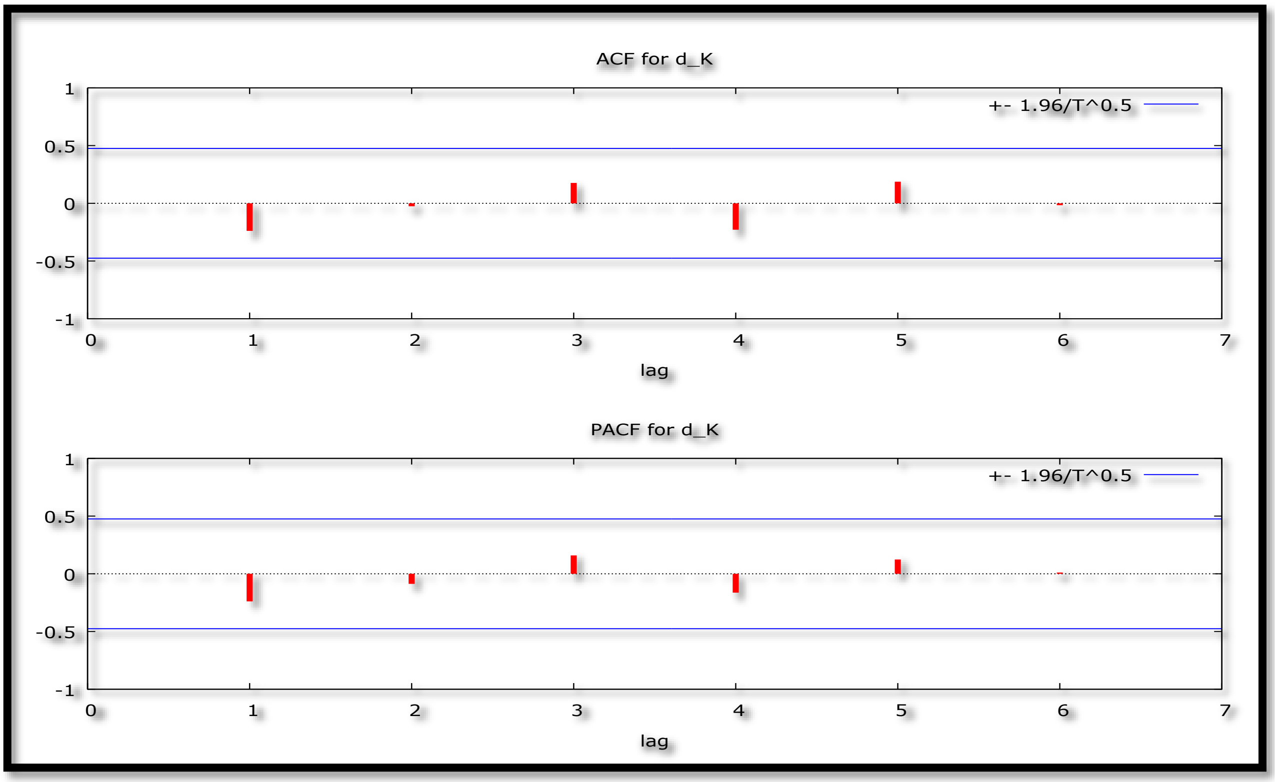

Figure 3: Correlogram (at First Differences)

3.7.5 The ADF Test (at First Differences)

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔK	-5.291503	0.0007	-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
ΔK	-6.298168	0.0006	-4.667883	@1%	Stationary
			-3.733200	@5%	Stationary
			-3.310349	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that K 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	RMSE	MAPE
ARIMA (1, 1, 0)	**-46.91411**	0.15065	0.0027827	0.052009	0.17593
ARIMA (2, 1, 0)	-45.14562	0.14999	0.0028396	0.051689	0.16735
ARIMA (3, 1, 0)	-43.72766	0.1476	0.0026112	0.050939	0.17241
ARIMA (1, 1, 1)	-45.69695	0.15037	0.0028519	0.051856	0.17221
ARIMA (0, 1, 1)	-46.63895	0.15012	0.0024057	0.05183	0.16236
ARIMA (0, 1, 2)	-45.09621	0.15006	0.0029067	0.051774	0.17446
ARIMA (0, 1, 3)	-43.10016	0.1495	0.0027967	0.051655	0.17178

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 finally chosen.

3.8 Residual & Stability Tests:

3.8.1 ADF Test (in levels) of the Residuals of the ARIMA () Model:

Table 6: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-3.257895	0.0351	-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	-3.963136	0.0339	-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:

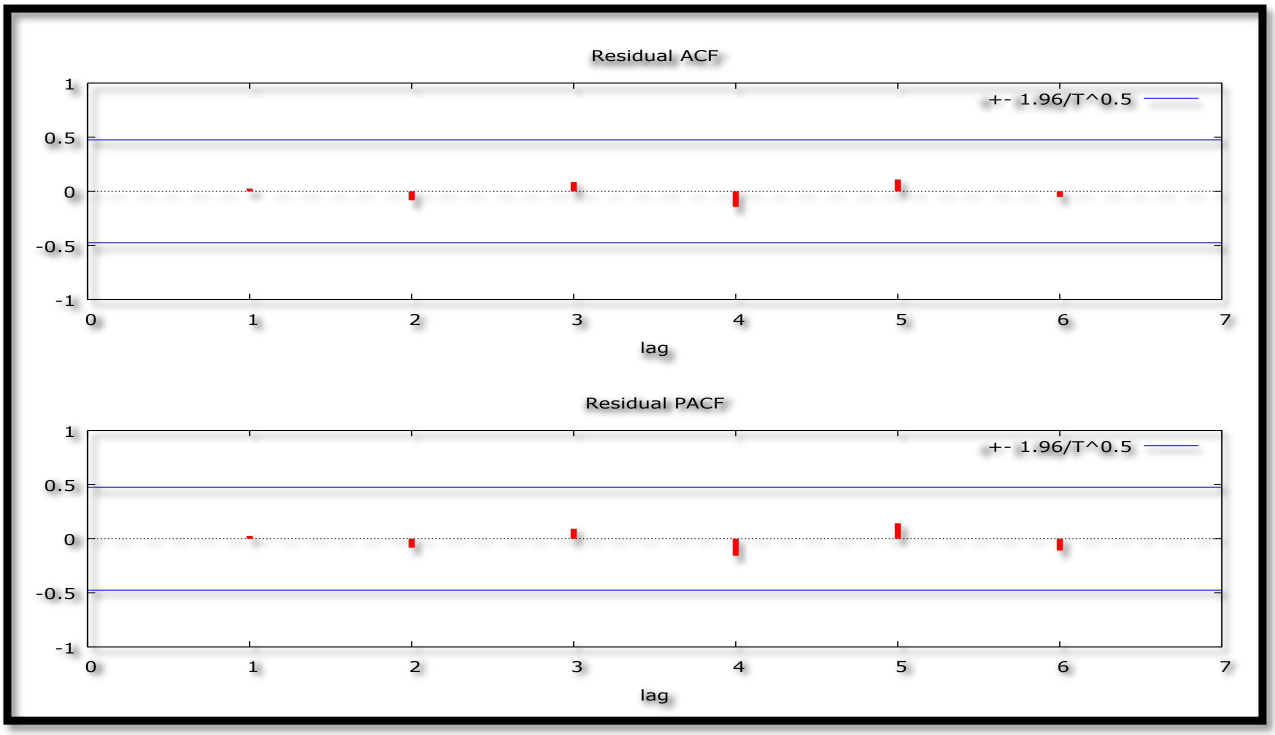

Figure 4: Correlogram of the Residuals

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

NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 6, ISSUE 11, Nov. -2020

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

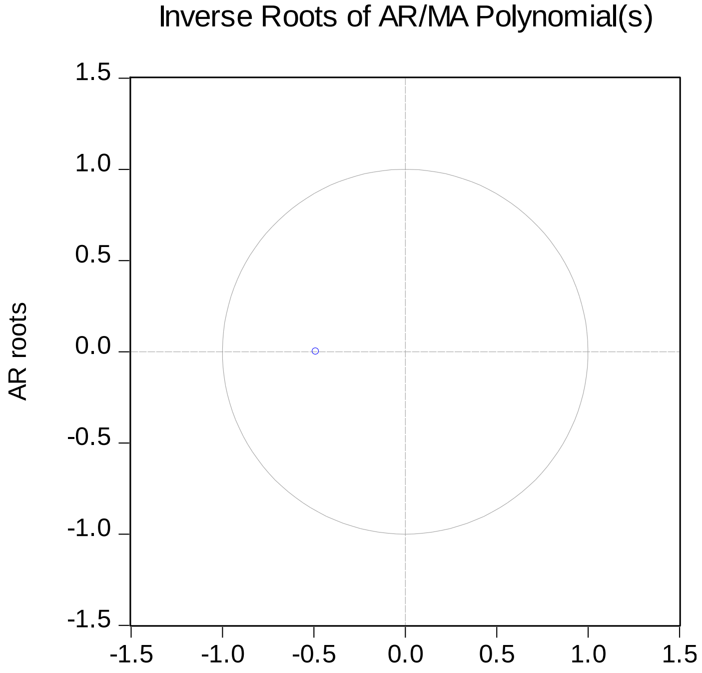

Figure 5: Inverse Roots

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 really stable and suitable for forecasting annual number of people practicing open defecation in Zambia.

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

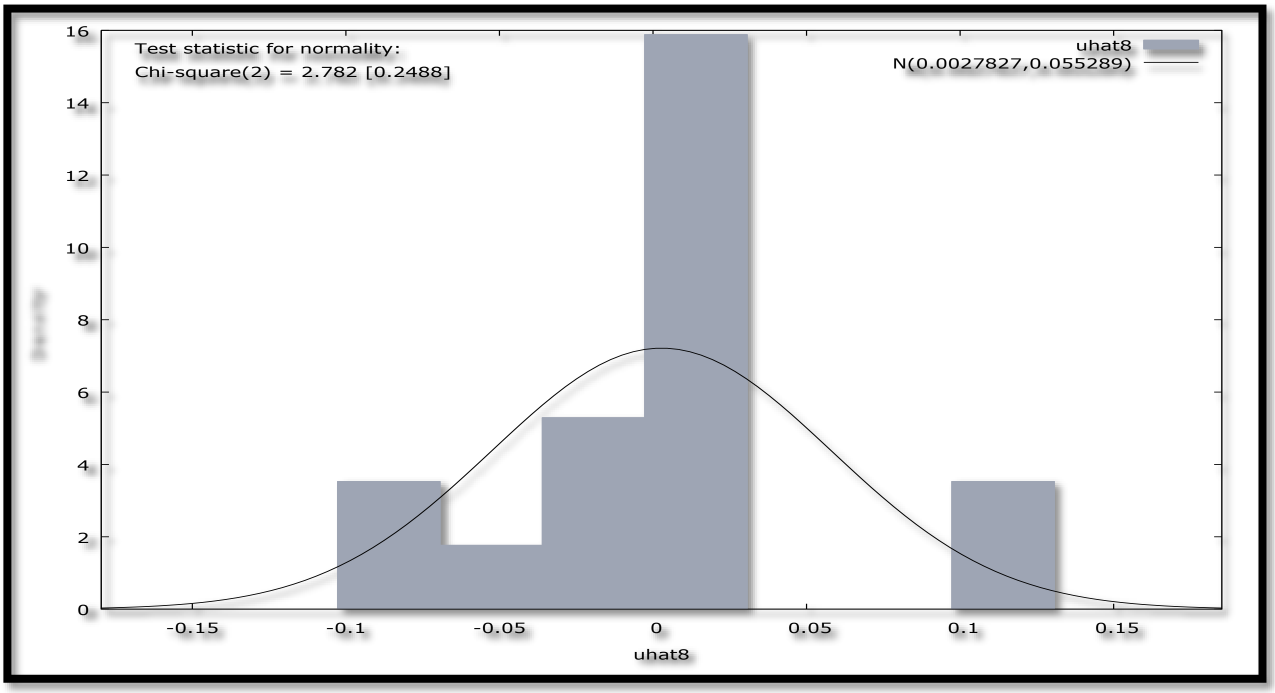

Figure 6: Normality Test

Figure 6 indicates the residuals of the optimal model are normally distributed as shown by the insignificance of the p-value of the chi-square statistic.

FINDINGS:

4.1 Descriptive Statistics:

Table 8: Descriptive Statistics

Description	Statistic
Mean	21.84
Median	21.85
Minimum	19.3
Maximum	24.5

As shown in table 8 above, the mean is positive, that is, 21.84. This means that, over the study period, the annual average number of people practicing open defecation in Zambia is approximately 22% of the total population. The minimum number of people practicing open defecation in Zambia over the study period is approximately 19% of the total population, while the maximum is 25% of the total population. In fact, the number of people practicing open defecation in Zambia has slightly declined over the years from 25% in 2000 to 19% of the total population in 2017.

4.2 Results Presentation

Table 9: Main Results

ARIMA (1, 1, 0) Model:				
Guided by equation [4], the chosen optimal model, the ARIMA (1, 1, 0) model can be expressed as follows: $\Delta K_t = -0.309772 + 0.430947\Delta K_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [5]$				
Variable	Coefficient	Standard Error	z	p-value
constant	-0.309772	0.00874385	-35.43	0.0000***
β_1	-0.430947	0.261946	-1.645	0.0999*

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

Forecast Graph

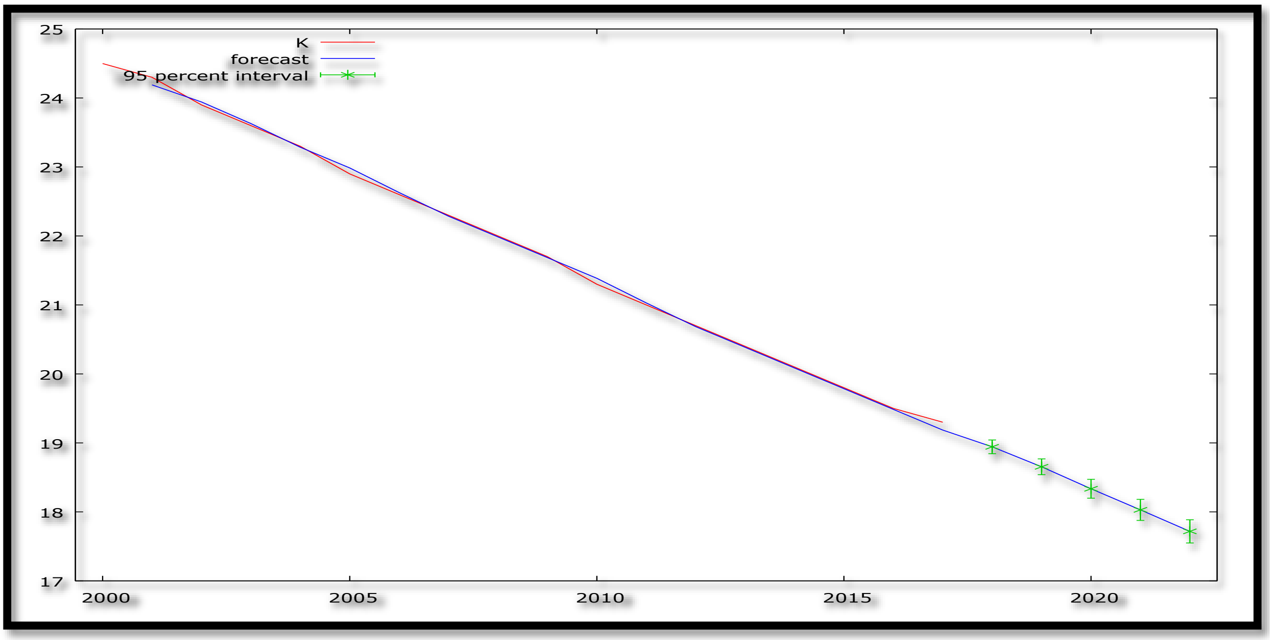

Figure 7: Forecast Graph – In & Out-of-Sample Forecasts

Figure 7 shows the in-and-out-of-sample forecasts of the K series. The out-of-sample forecasts cover the period 2018 – 2022.

Predicted K– Out-of-Sample Forecasts Only

Table 10: Predicted

Year	Predicted K	Standard Error	Lower Limit	Upper Limit
2018	18.9	0.05	18.8	19
2019	18.7	0.06	18.5	18.8
2020	18.3	0.07	18.2	18.5
2021	18	0.08	17.9	18.2
2022	17.7	0.09	17.5	17.9

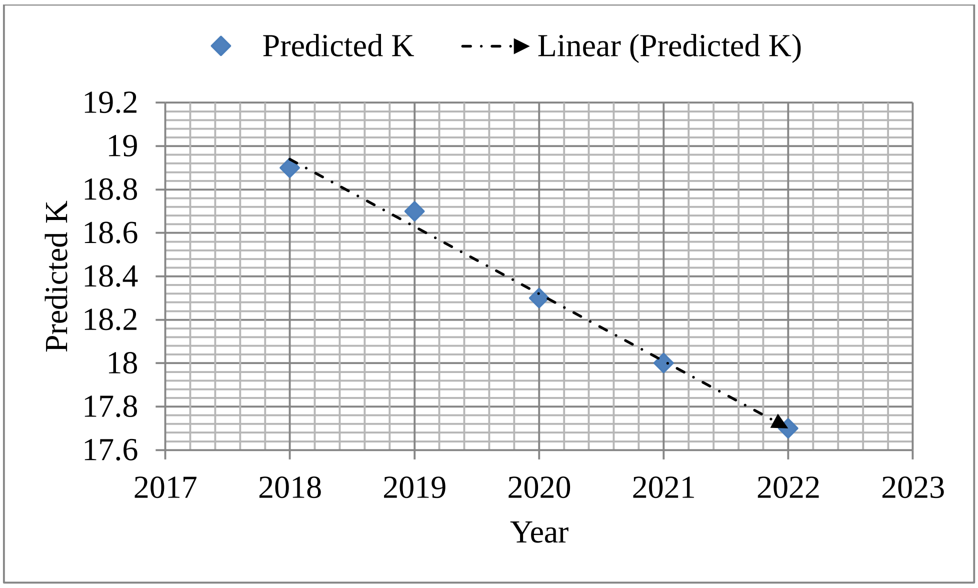

Figure 8: Graphical Analysis of Out-of-Sample Forecasts

Table 10 and figure 8 show the out-of-sample predictions only. The number of people practicing open defecation in Zambia is projected to slightly fall from approximately 18.9% in 2018 to 17.7% of the total population by the year 2022. Open defecation is relatively under control in Zambia. However, it is possible to significantly reduce the number of open defecators in Zambia, especially if the current government considers and intensifies the implementation of the policy directions suggested below.

4.3 Policy Implications:

i. The government of Zambia should not stop making toilets a status symbol throughout the country, especially in rural areas.

ii. The government of Zambia to continue creating more demand for sanitation through teaching the public on the importance of investing in toilets.

iii. There is need for the government of Zambia to continue stimulating a habit of systematic hand-washing, and not defecating in the open.

CONCLUSION:

The study shows 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 Zambia over the period 2018 – 2022. The model predicts a slight decrease in the annual number of people practicing open defecation in Zambia.

These findings are of great use for the government of Zambia, especially for 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) Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control, Holden Day, San Francisco.
3) Klutse, A., Bouraima, Z., & Amegnran, C. (2010). Sanitation Costs Analysis in Burkina Faso, IRC Symposium.
4) Kojo, Y. et al. (2019). Improving Sanitation and Hygiene Through Community Led Total Sanitation: The Zambian Experience, American Journal of Tropical Medicine and Hygiene, 00 (0): 1 – 18.
5) Lozano, R., et al. (2013). Global and Regional Mortality From 235 Causes of Death for 20 Age Groups in 1990 and 2010: A Systematic Analysis for the Global Burden of Disease Study 2010, The Lancet, 380 (9859): 2095 – 2128.
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. (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.
8) Nyoni, T. (2018c). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, MPRA Paper No. 87737, University Library of Munich, Munich, Germany.
9) USAID (2006). Zambia Water and Sanitation Profile, USAID Country Office, Zambia.
10) WaterAid (2010). WaterAid Zambia Country Strategy 2011 – 2015, WaterAid Country Office, Zambia.
11) WSP (2012). Economic Impacts of Poor Sanitation, WSP Country Office, Zambia.