European Journal of Research Development and Sustainability (EJRDS)
Available Online at: https://www.scholarzest.com
Vol. 1 No. 1, September 2020,
ISSN: 2660-5570

OPEN DEFECATION IN BOTSWANA: 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

Article history:

Received August 23st 2020
Accepted: September 11th 2020
Published: September 30th 2020

Abstract:

Using annual time series data on the number of people who practice open defecation in Botswana from 2000 – 2017, the study predicts the annual number of people who will still be practicing open defecation over the period 2018 – 2021. The research applies the Box-Jenkins ARIMA methodology. The diagnostic ADF tests show that the BOD series under consideration is an I (2) variable. Based on the AIC, the study presents the ARIMA (1, 2, 0) model as the optimal model. The diagnostic tests further show that the presented model is stable and its residuals are stationary in levels. The findings of the study indicate that the number of people practicing open defecation in Botswana is likely to decline, over the period 2018 – 2022, from approximately 10.4% to almost 8.1% of the total population. Indeed, it is possible to create an open defecation free society in Botswana. The study suggested a 3-fold policy recommendation to be put into consideration, especially by the government of Botswana.

Keywords: Defecation in Botswana, research, Box-Jenkins ARIMA methodology.

1.0 INTRODUCTION

A significant number of people, especially in rural Botswana, practice open defecation, that is, relieving themselves anywhere on the ground in the open, usually in the bush (Odirile *et al*., 2018). This may be due to a limited awareness of safe hygienic practices coupled with the availability of ample open space around their homesteads and villages (UNICEF, 2015). Open defecation is terrible from a public health perspective (UNICEF, 2018), particularly, in terms of the spread of bacterial, viral and parasitic infections including diarrhoea, polio, cholera, soil-transmitted helminth, trachoma infection, schistosomiasis and hookworm and is also an important cause of child stunting (Megersa *et al*., 2019) and deaths (Thiga & Cholo, 2017). It has, thus, become fundamental for public health researchers and policy makers 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 Botswana over the period 2018 – 2021. This study, besides being the first of its kind in the case of Botswana, 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 10% of the total population in Botswana.
ii. To forecast the number of people practicing open defecation in Botswana for the period 2018 – 2021.
iii. To examine the trend of open defecation in Botswana for the out-of-sample period.

2.0 LITERATURE REVIEW

Guterres *et al*. (2014) analyzed factors that influence household to use and maintain latrines in Thailand using a cross-sectional survey, based on a quantitative data design. Their study generally 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. Sintondji *et al*. (2017) investigated the influence of socio-demographic factors on household hygiene and sanitation behaviour in Benin using interviews and the results of their research basically showed that 68% of households did not cover their containers during the transport of water, 58% of respondents defecated in water and 31% in the open air. Osumanu *et al*. (2019) studied sociocultural and economic factors determining open defecation in the Wa Municipality in Ghana. The research applied 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 estimated to analyze the factors that significantly influence open defecation. The results generally show 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. No study has been done to forecast the number of open defecators in Botswana. This study is the first of its kind in the case of Botswana and is envisioned to enhance the eradication of open defecation in Botswana.

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 used to analyze the BOD series under consideration.

3.2 The Moving Average (MA) model

Given:

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

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

$$\mathrm{BOD_t} = \sum_{i=1}^{p} \beta_i \mathrm{BOD_{t-i}} + \mu_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots [2]$$

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

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

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

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

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

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

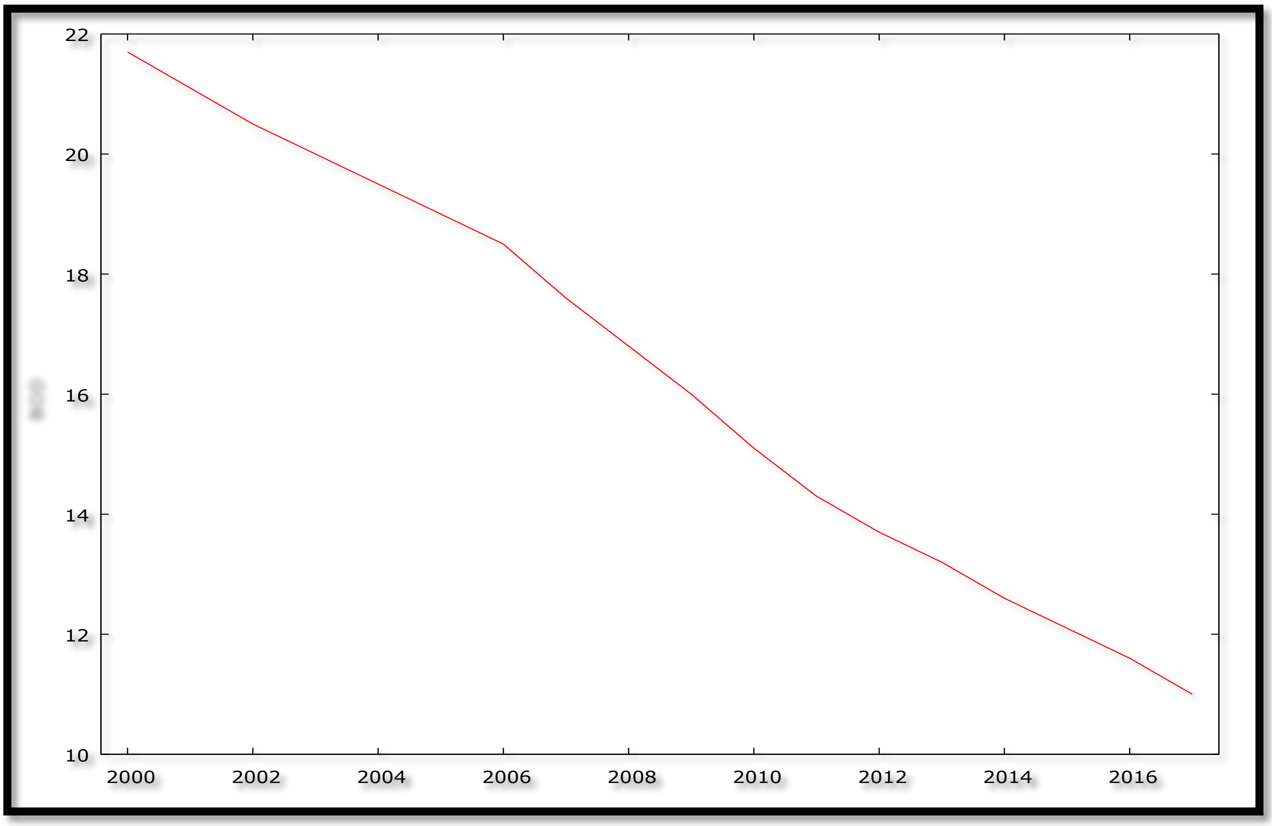
3.6 Data Collection

This study is based on annual observations (that is, from 2000 – 2017) on the number of people practicing Open Defecation [OD, denoted as BOD] (as a percentage of total population) in Botswana. Out-of-sample forecasts will cover the period 2017 – 2021. All the data was gathered from the World Bank online database.

3.7 Diagnostic Tests & Model Evaluation

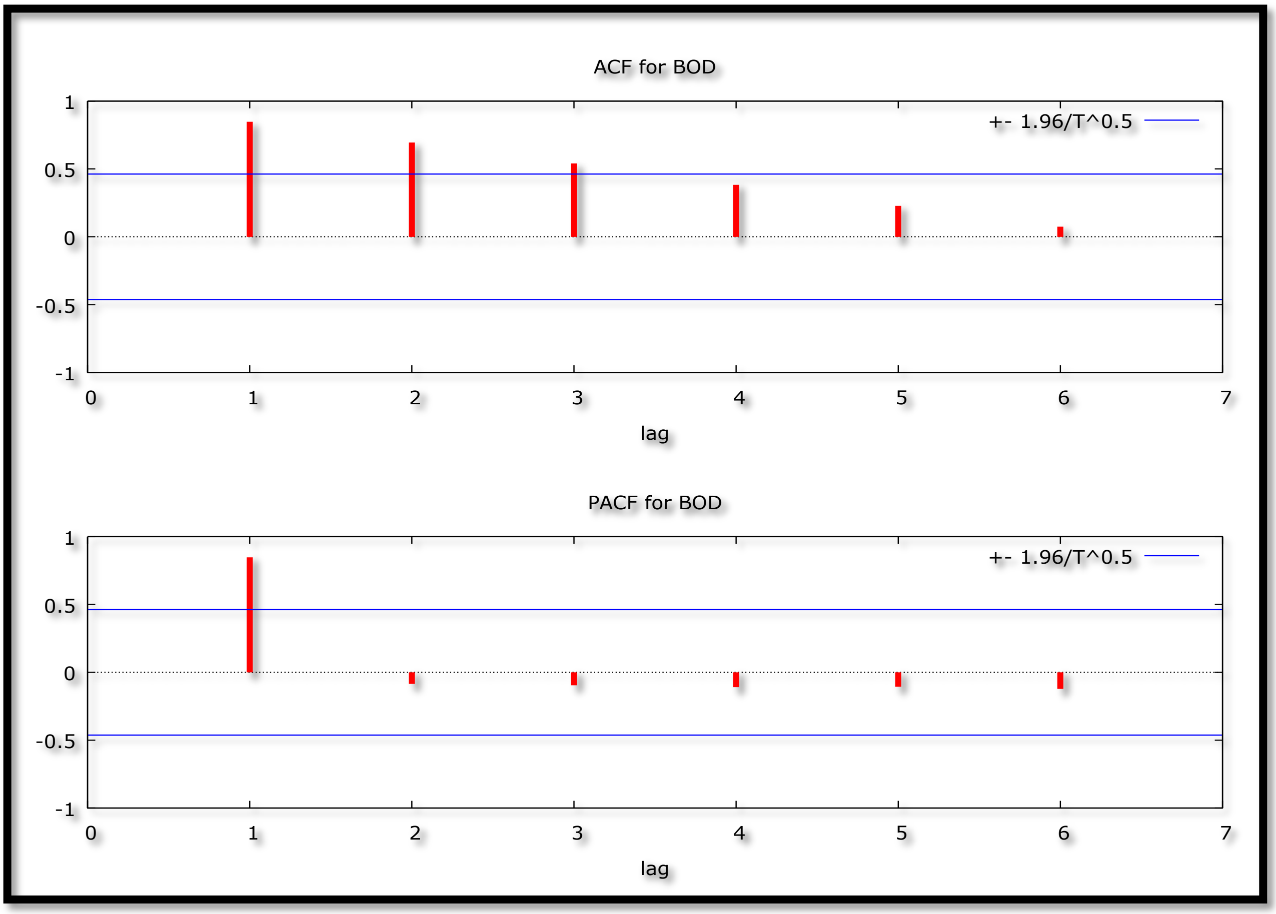
3.7.1 Stationarity Tests: Graphical Analysis

Figure 1



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
BOD	-0.331322	0.8997	-3.920350	@1%	Non-stationary
			-3.065585	@5%	Non-stationary
			-2.673459	@10%	Non-stationary

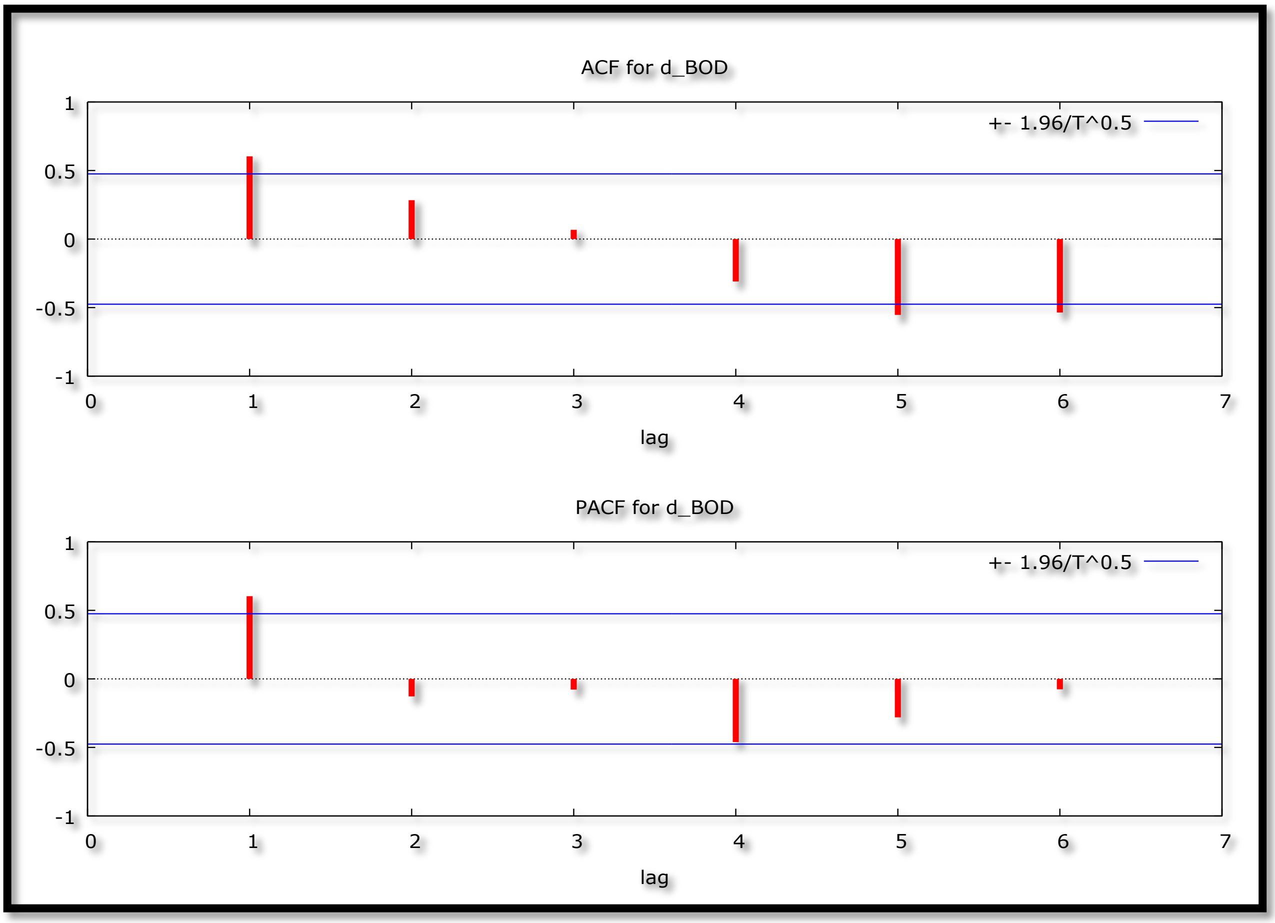
Table 2: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
BOD	-4.089762	0.0312	-4.800080	@1%	Non-stationary
			-3.791172	@5%	Stationary
			-3.342253	@10%	Stationary

Table 1 shows that BOD is not stationary in levels. However, table 2 indicates the opposite. We therefore, proceed to test for the existence of a unit root after taking first differences.

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
ΔBOD	-1.856382	0.3423	-3.920350	@1%	Non-stationary
			-3.065585	@5%	Non-stationary
			-2.673459	@10%	Non-stationary

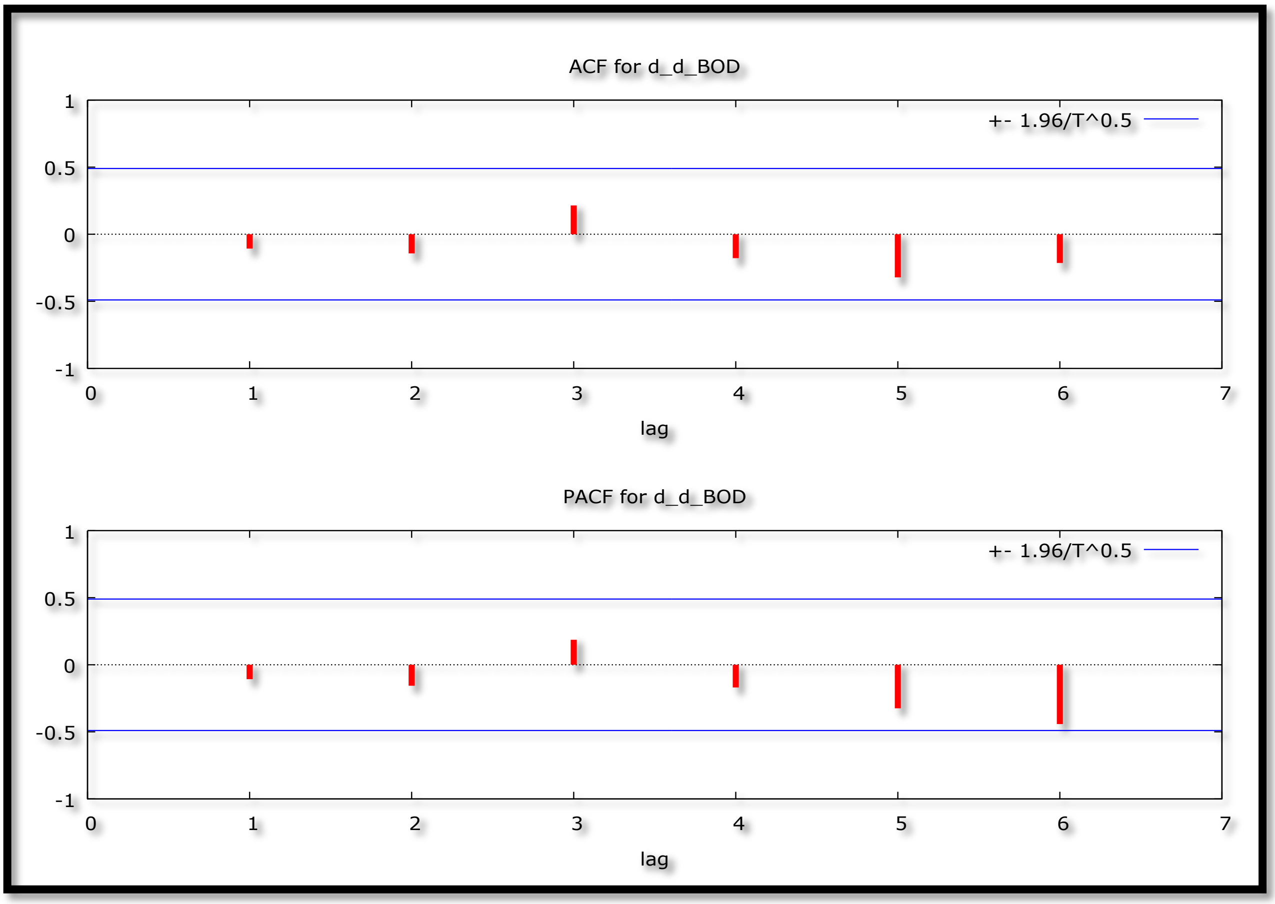
Table 4: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔBOD	-1.786504	0.6632	-4.667883	@1%	Non-stationary
			-3.733200	@5%	Non-stationary
			-3.310349	@10%	Non-stationary

Figure 3 as well as tables 3 and 4; indicate that BOD is non-stationary in first levels. We therefore, proceed to test for stationary after taking second differences.

3.7.6 The Correlogram (at Second Differences)

Figure 4: The Correlogram (at Second Differences)



3.7.7 The ADF Test (at Second Differences)

Table 5: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
$\Delta^2 BOD$	-3.953755	0.0101	-3.959148	@1%	Non-stationary
			-3.081002	@5%	Stationary
			-2.681330	@10%	Stationary

Table 6: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
$\Delta^2 BOD$	-3.804724	0.0465	-4.728363	@1%	Non-stationary
			-3.759743	@5%	Stationary
			-3.324976	@10%	Stationary

Table 5 and 6, in line with figure 4, shows that the BOD series is an I(2) variable.

3.7.8 Evaluation of ARIMA models (with a constant)

Table 7: Evaluation of ARIMA Models (with a constant)

Model	AIC	U	ME	RMSE	MAPE
ARIMA (1, 2, 0)	**-13.50319**	0.1989	-0.0000039135	0.1315	0.58859
ARIMA (2, 2, 0)	-11.88107	0.19511	0.00091457	0.12982	0.58095
ARIMA (3, 2, 0)	-10.42220	0.1899	-0.00013836	0.12735	0.55246
ARIMA (0, 2, 1)	-13.56216	0.19848	0.00012317	0.13123	0.5818
ARIMA (0, 2, 2)	-12.04586	0.18896	-0.010889	0.12799	0.56041
ARIMA (0, 2, 3)	-12.48098	0.16949	-0.0037741	0.11971	0.47797
ARIMA (0, 2, 4)	-13.02989	0.15926	-0.010249	0.11285	0.41701
ARIMA (0, 2, 5)	-12.10438	0.15101	-0.0079867	0.10836	0.37464

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, 2, 0) model is finally chosen.

3.8 Residual & Stability Tests

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

Table 8: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-3.599170	0.0195	-3.959148	@1%	Non-stationary
			-3.081002	@5%	Stationary
			-2.681330	@10%	Stationary

Table 9: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-3.464024	0.0805	-4.728363	@1%	Non-stationary
			-3.759743	@5%	Stationary
			-3.324976	@10%	Stationary

Tables 8 and 9 indicate that the residuals of the chosen optimal model, the ARIMA (1, 2, 0) model; are stationary. Hence, the model is stable.

3.8.2 Correlogram of the Residuals of the ARIMA (1, 2, 0) Model

Figure 5: Correlogram of the Residuals

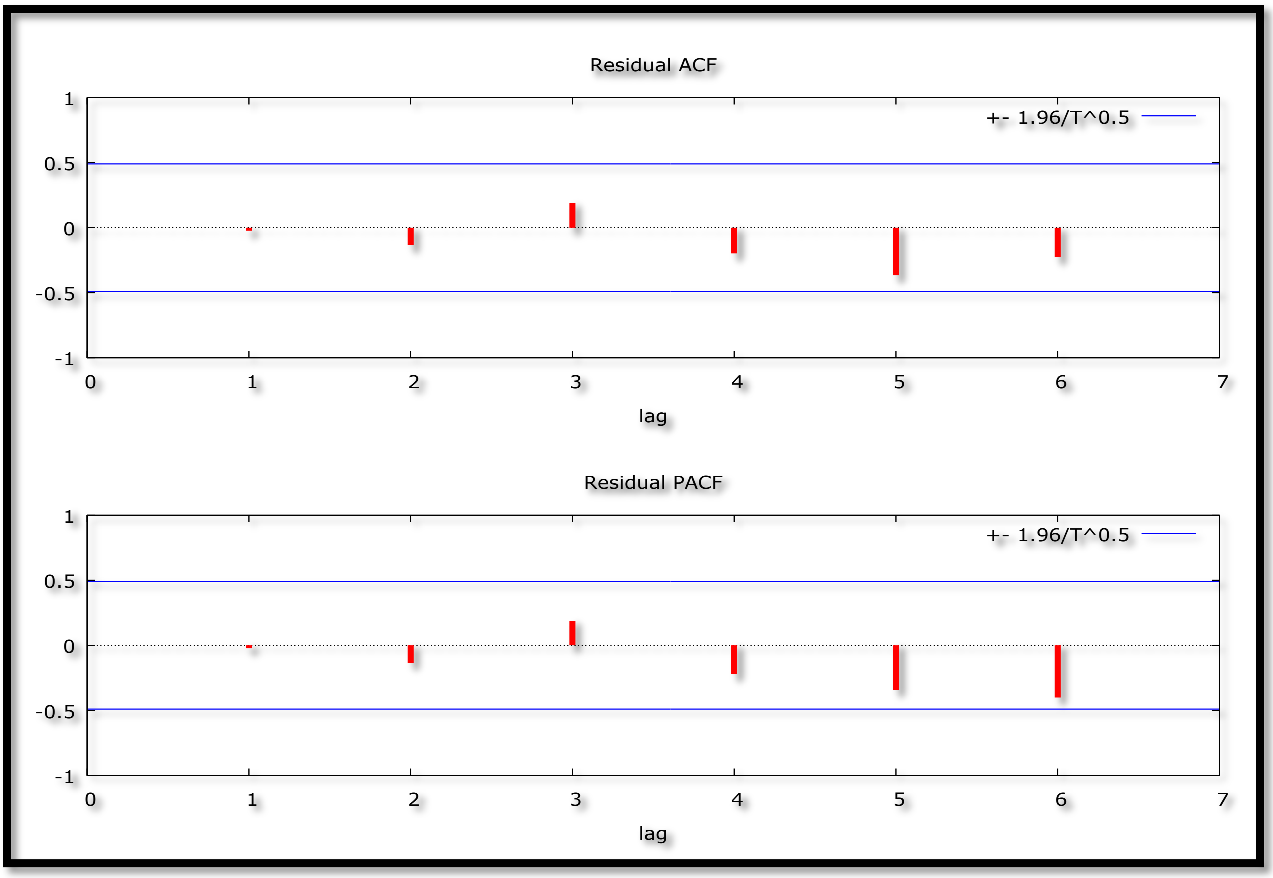
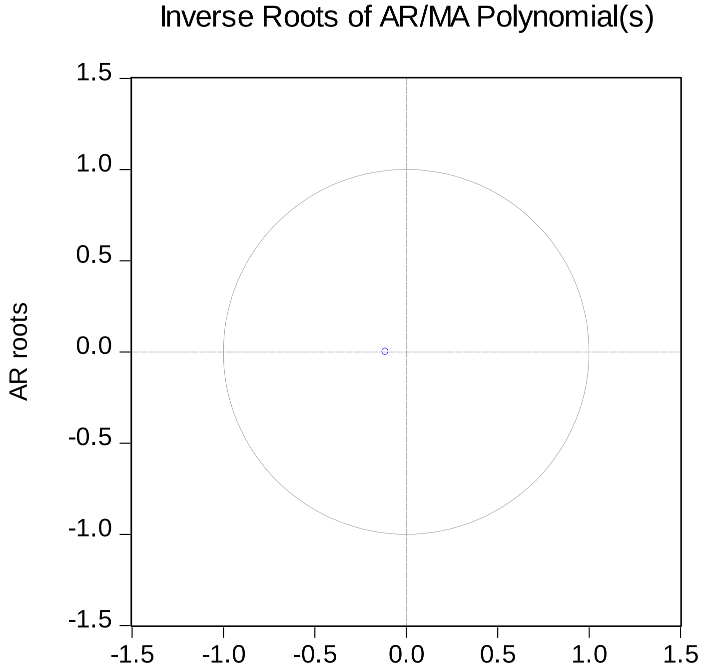


Figure 5 reveals that the estimated model is adequate since ACF and PACF lags are quite short and within the bands. This apparently means that the "no autocorrelation" assumption is not violated in this study.

3.8.3 Stability Test of the ARIMA (1, 2, 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 confirming that the ARIMA (1, 2, 0) model is really stable and suitable for forecasting annual number of people practicing open defecation in Botswana.

4.0 FINDINGS

4.1 Descriptive Statistics

Table 10: Descriptive Statistics

Description	Statistic
Mean	16.35
Median	16.4
Minimum	11
Maximum	21.7

As shown in table 8 above, the mean is positive, that is, 16.35. This means that, over the study period, the annual average number of people practicing open defecation in Botswana is approximately 16% of the total population. The minimum number of people practicing open defecation in Botswana over the study period is approximately 11% of the total population, while the maximum is 21.7% of the total population. In fact, the number of people practicing open defecation in Botswana has declined over the years from 21.7% in 2000 to 11% of the total population in 2017.

4.2 Results Presentation[1]

Table 11: Main Results

ARIMA (1, 2, 0) Model:

Guided by equation [4], the chosen optimal model, the ARIMA (1, 2, 0) model can be expressed as follows:

$$\Delta^2 BOD_t = 0.000598855 - 0.104585\Delta^2 BOD_{t-1} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots [5]$$

Variable	Coefficient	Standard Error	z	p-value
constant	0.000598855	0.0299674	0.01998	0.9841
β_1	-0.104585	0.245375	-0.4262	0.6699

Table 11 shows the main results of the ARIMA (1, 2, 0) model.

[1] *The *, ** and *** imply statistical significance at 10%, 5% and 1% levels of significance; respectively.*

Forecast Graph

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

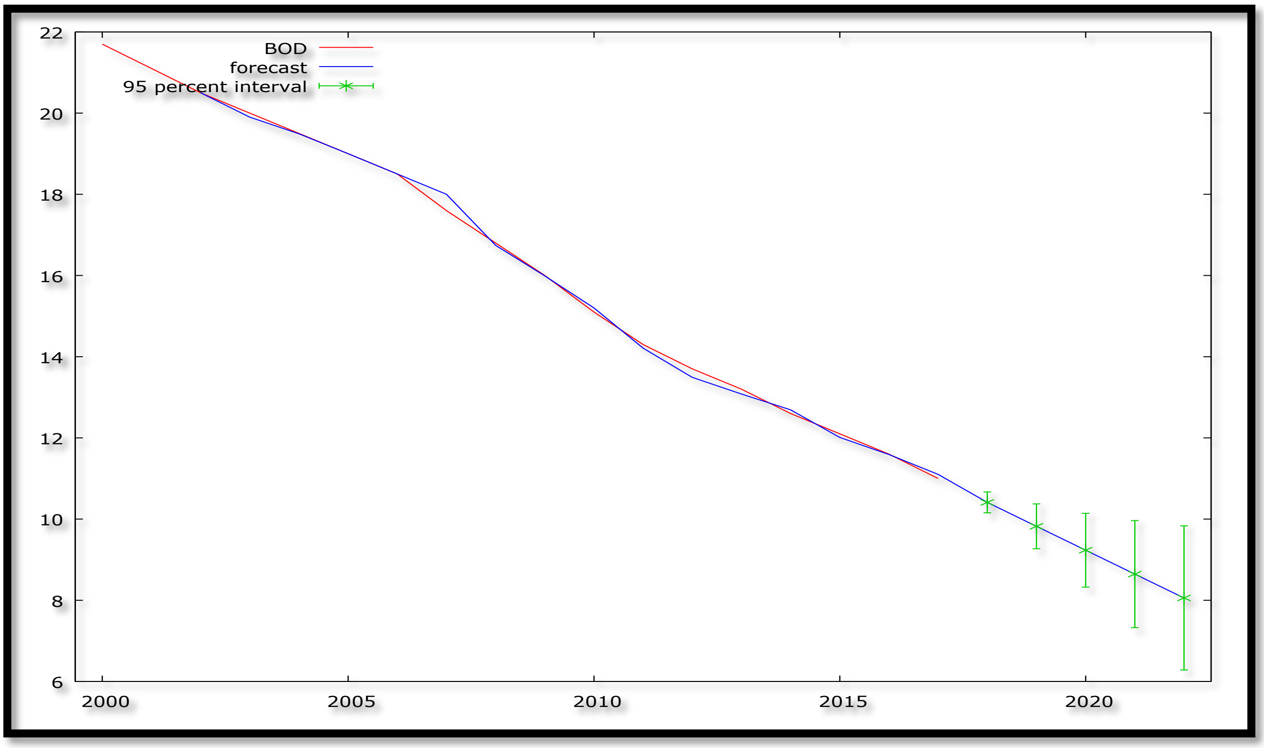


Figure 7 shows the in-and-out-of-sample forecasts of the BOD series. The out-of-sample forecasts cover the period 2018 – 2022.

Predicted BOD – Out-of-Sample Forecasts Only

Table 12: Predicted

Year	Predicted BOD	Standard Error	Lower Limit	Upper Limit
2018	10.4	0.13	10.2	10.7
2019	9.8	0.28	9.3	10.4
2020	9.2	0.46	8.3	10.1
2021	8.6	0.67	7.3	10
2022	8.1	0.91	6.3	9.8

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

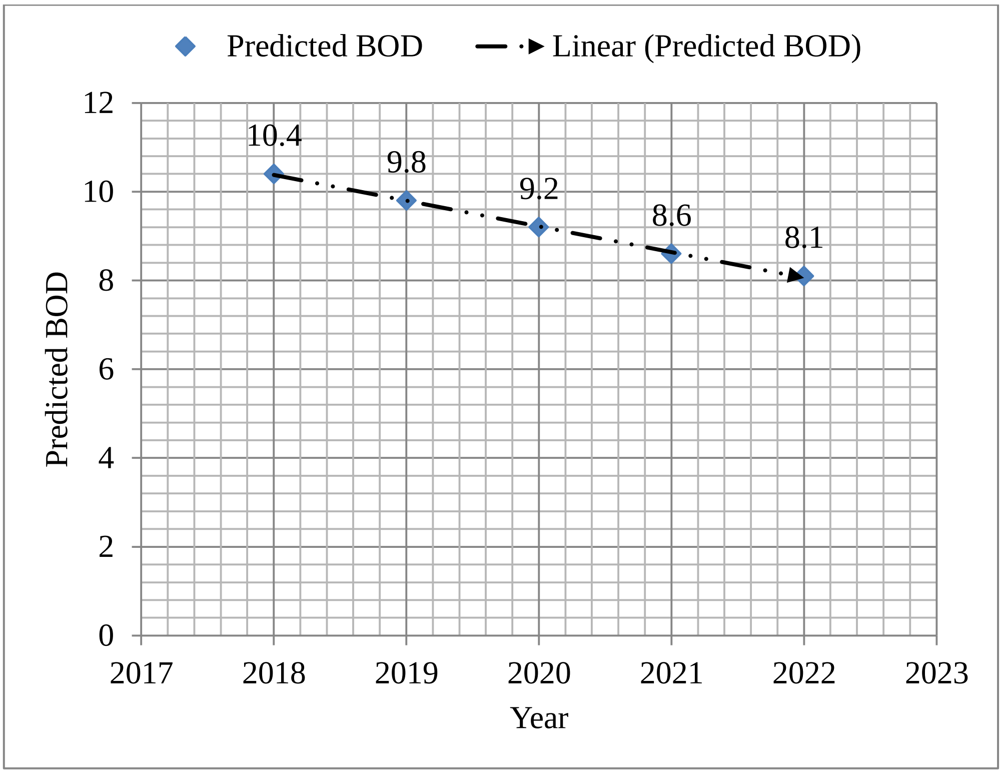


Table 12 and figure 8 show the out-of-sample forecasts only. The number of people practicing open defecation in Botswana is projected to fall from approximately 10.4% in 2018 to 8.1% of the total population by the year 2022. It is possible to do away with the practice of defecating in the open in Botswana. The government of Botswana can benefit from the policy directions derived from this study.

4.3 Policy Implications

i. The government of Botswana should continue making toilets a status symbol.
ii. The government of Botswana should also continue creating demand for sanitation through teaching the public on the importance of investing in toilets.
iii. There is need for the government of Botswana to also continue encouraging a habit of not defecating in the open.

5.0 CONCLUSION

The study shows that the ARIMA (1, 2, 0) model is not only stable but also the most suitable model to forecast the annual number of people practicing open defecation in Botswana over the period 2018 – 2022. The model predicts a commendable decrease in the annual number of people practicing open defecation in Botswana. Hence, open defecation can possibly be eliminated in the country. These results are important for the government of Botswana, especially for 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. 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.
3. Megersa, S., Benti, T., & Sahiledengle, B. (2019). Prevalence of Diarrhoea and Its Associated Factors Among Under-Five Children in Open Defecation Free and Non-Open Defecation Free Households in Goba District Southeast Ethiopia: A Comparative Cross-Sectional Study, *Clinics in Mother and Child Health*, 16 (324): 1 – 9.
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. (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.
6. Nyoni, T. (2018c). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, MPRA Paper No. 87737, *University Library of Munich*, Munich, Germany.
7. Odirile, P. T., *et al*. (2019). Faecal Sledge Management in Botswana: A Review of Current Practices and Policies Using the Case of Gaborone Low Income Areas, *Journal of Environmental Protection*, 9: 122 – 139.
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 Science and Public Health*, pp: 1 – 11.
9. Sintondji, L. O., *et al*. (2017). Socio-Demographic Characteristics of Households and Determinants of Access to Water, Hygiene and Sanitation in So-Ava, Benin, *Journal of Environmental Science and Public Health*, 1 (4): 253 – 267.
10. Thiga, L. M., & Cholo, W. (2017). An Assessment of Open Defecation Among Residents of Thika East Sub-County, Kiambu County, Kenya, *International Journal of Medicine Research*, 2 (3): 9 – 20.
11. UNICEF (2015). National ODF Sustainability Assessment, *UNICEF*, New York.
12. UNICEF (2018). Game Plan to End Open Defecation, *UNICEF*, New York.