OPEN DEFECATION IN MALAWI: 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 Malawi 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 M 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 show that the presented model is stable and its residuals are stationary in levels. The results of the study indicate that the number of people practicing open defecation in Malawi is likely to decline, over the period 2018 – 2022, from approximately 5.1% to almost 2.8% of the total population. Indeed, by 2030, open defecation can be eliminated in Malawi: hence, the country is in the right track with regards to its vision 2030 (on water, sanitation and hygiene). The study suggested a 3-fold policy recommendation to be put into consideration, especially by the government of Malawi.

1.0 INTRODUCTION

A relatively lesser number of people in Malawi, practice open defecation, that is, relieving themselves anywhere on the ground in the open. Just like in any other country, most open defecators in Malawi live in rural areas. 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. 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). Thus, it has become even more instructive for public health researchers and policy makers to model and forecast the number of people practicing open defecation in order to come up with evidence-driven policies to end open defecation. The main goal of this study is to predict the annual number of open defecators in Malawi over the period 2018 – 2021. This study, besides being the first of its kind in the case of Malawi, 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 5% of the total population in Malawi.
  ii. To forecast the number of people practicing open defecation in Malawi for the period 2017 – 2021.
  iii. To examine the trend of open defecation in Malawi for the out-of-sample period.

2.0 LITERATURE REVIEW

In Kenya, Njuguna & Muruka (2017) looked at open defecation trends among the 47 counties in, newly created in 2013. The study relied on four data sets on open defecation, unimproved water supply coverage, poverty levels and population density. Their results basically show that the average open defecation rate across the 47 counties was 23.5% and the median rate was 6.9% and also that poverty was the most significant predictor accounting for 68.4% of the variance in open defecation after controlling for unimproved water supply and population density. Ayalew *et al.* (2018) examined diarrheal morbidity in under-5 children and its related factors in Dangla district in Northwest Ethiopia. A community-based comparative cross-sectional study design with a multistage random sampling technique was conducted. Descriptive and inferential statistics were performed. The study basically showed that child immunization, latrine presence, water shortage in household, and solid waste disposal practices had statistically significant association with diarrhoea occurrence in Ethiopia. In Ghana, Alhassan & Anyarayor (2018) examined 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 done to collect data. The study proved 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 ARIMA method in analyzing open defecation trends in Malawi.

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 M series under consideration.

3.2 The Moving Average (MA) model

Given:

$$M_t = \sum_{i=1}^{q} \alpha_i \mu_{t-i} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [1]$$

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). M is the annual number of people (as a percentage of the total population) who practice open defecation in Malawi at time t, $\alpha_0 \dots \alpha_q$ are estimation parameters, μ_t is the current error term while $\mu_{t-1} \dots \mu_{t-q}$ are previous error terms.

3.3 The Autoregressive (AR) model

Given:

$$M_t = \sum_{i=1}^{p} \beta_i M_{t-i} + \mu_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [2]$$

Where $\beta_1 \dots \beta_p$ are estimation parameters, $M_{t-1} \dots M_{t-p}$ are previous period values of the M 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:

$$M_t = \sum_{i=1}^{p} \beta_i M_{t-i} + \sum_{i=1}^{q} \alpha_i \mu_{t-i} + \mu_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [3]$$

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

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

$$\Delta^d M_t = \sum_{i=1}^{p} \beta_i \Delta^d M_{t-i} + \sum_{i=1}^{q} \alpha_i \mu_{t-i} + \mu_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots [4]$$

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

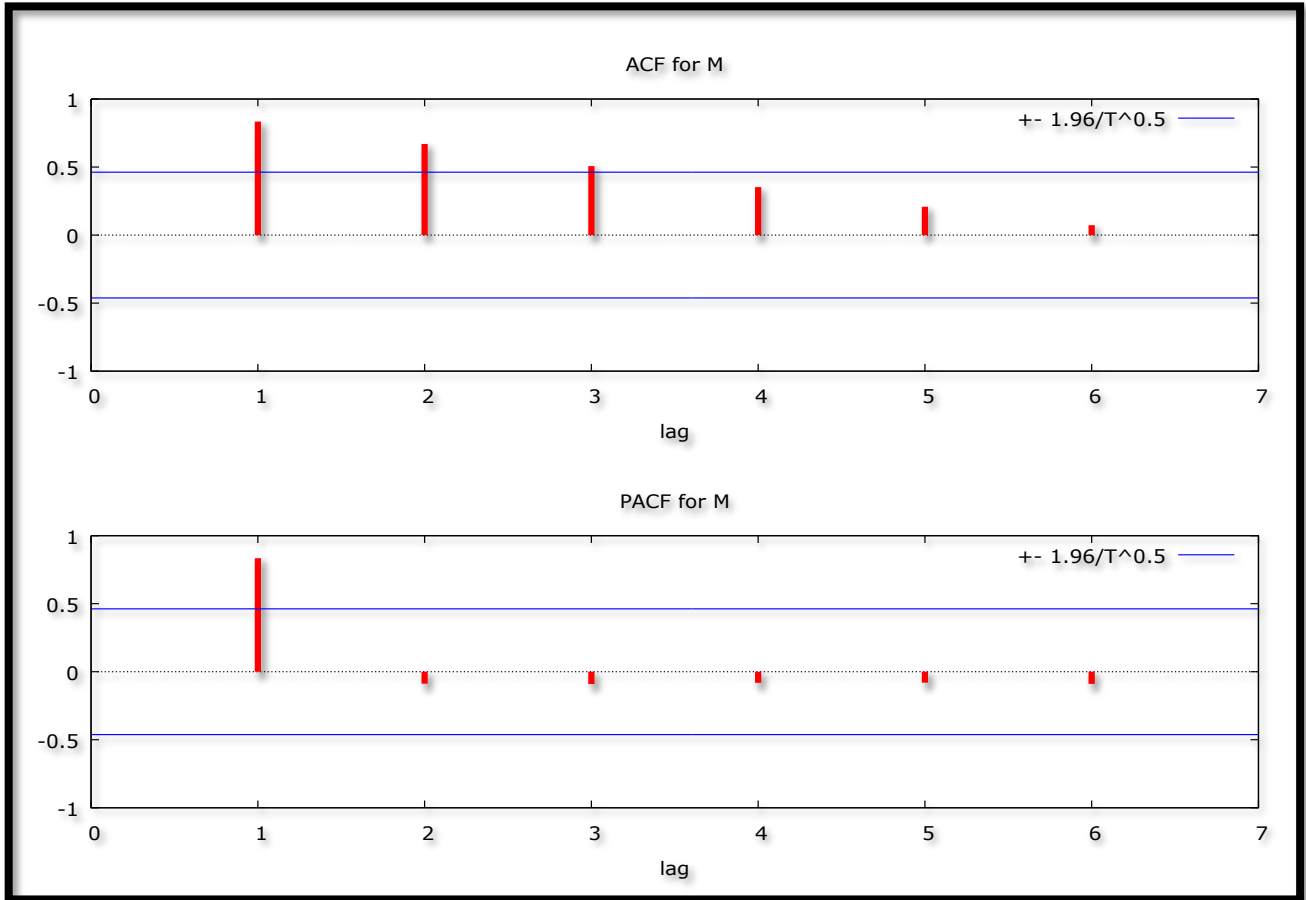### 3.7 Diagnostic Tests & Model Evaluation
3.7.1 Stationarity Tests: Graphical Analysis

Figure 1
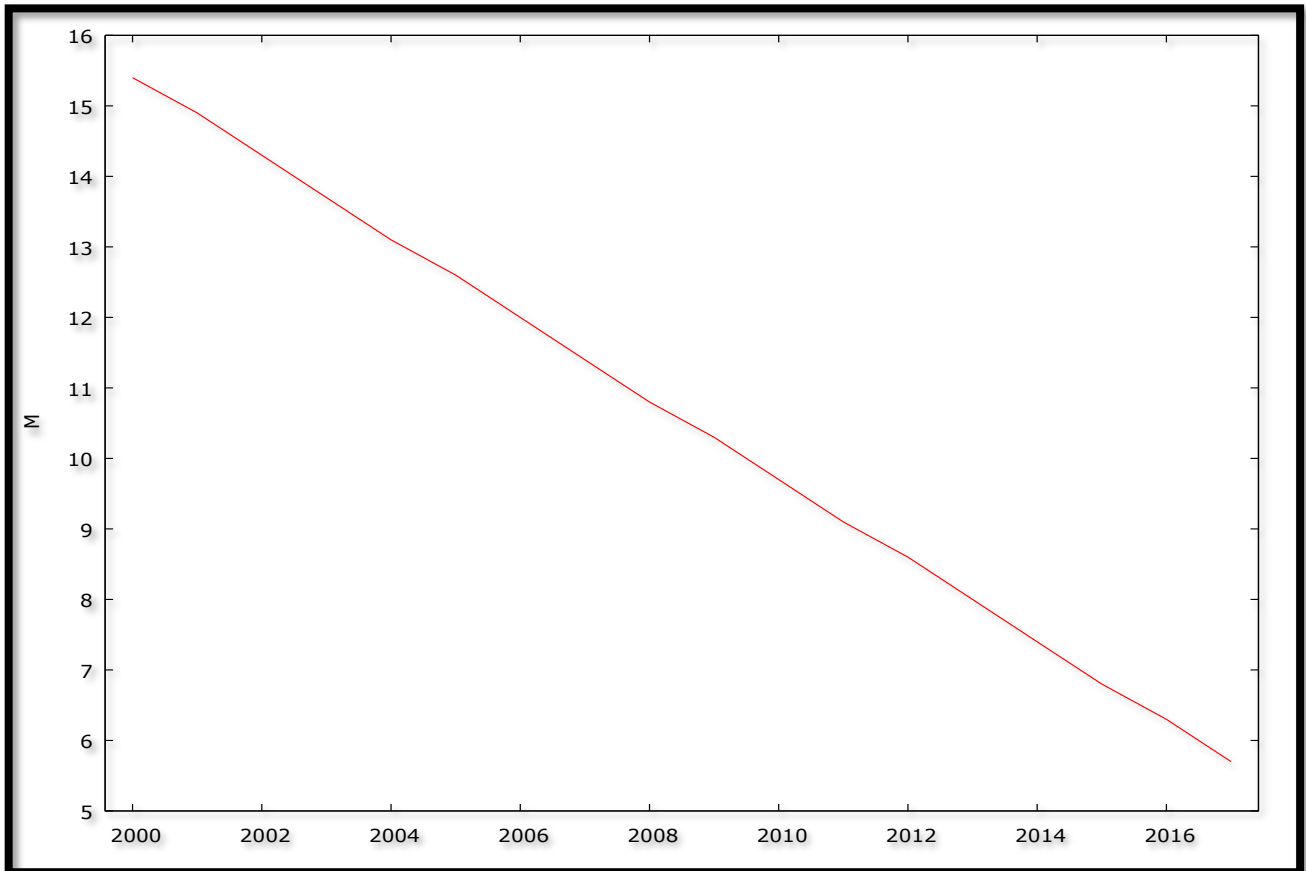
3.7.2 The Correlogram in Levels

Figure 2: Correlogram in Levels

3.7.3 The ADF Test in Levels

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
M	-0.538238	0.8555	-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
M	-5.092470	0.0048	-4.667883	@1%	Stationary
			-3.733200	@5%	Stationary
			-3.310349	@10%	Stationary

Table 1 shows that M is not stationary in levels while table portrays exactly the opposite; in line with figures 1 and 2.

3.7.4 The Correlogram (at First Differences)

Figure 3: Correlogram (at First Differences)
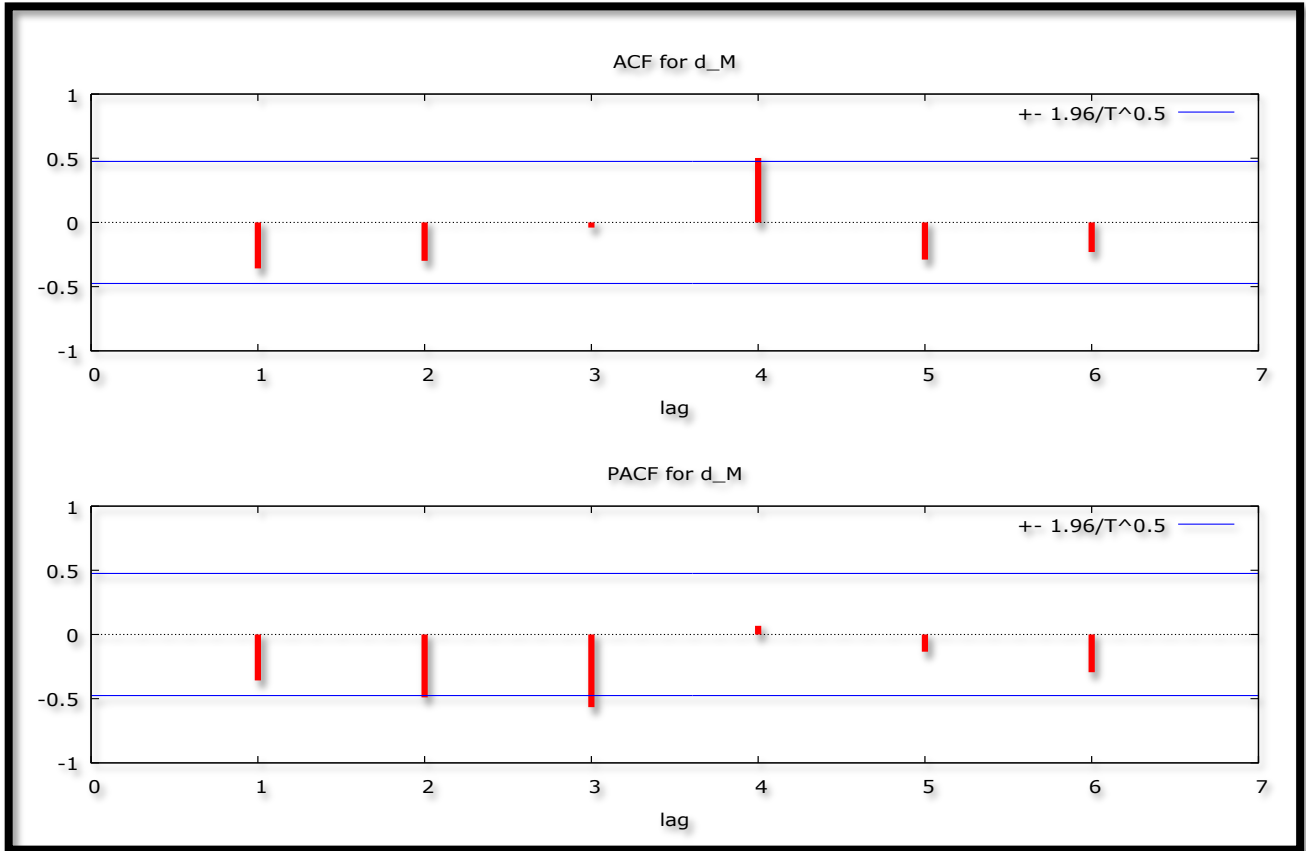
3.7.5 The ADF Test (at First Differences)

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔM	-6.993787	0.0001	-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
ΔM	-6.759356	0.0005	-4.800080	@1%	Stationary
			-3.791172	@5%	Stationary
			-3.342253	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that M 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)	-53.42307	0.069858	0.0016851	0.042498	0.35781
ARIMA (2, 1, 0)	-57.29043	0.057074	0.0031851	0.036363	0.24789
ARIMA (3, 1, 0)	**-66.84058**	0.038707	0.0040004	0.027516	0.12855
ARIMA (4, 1, 0)	-64.34364	0.038707	0.0040004	0.027516	0.12855
ARIMA (5, 1, 0)	-62.84058	0.038707	0.0040001	0.027516	0.12855

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 (0, 1, 1) model is finally 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	-3.647938	0.0168	-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	-3.524510	0.0708	-4.667883	@1%	Non-stationary
			-3.733200	@5%	Non-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 stable.

3.8.2 Correlogram of the Residuals of the ARIMA (3, 1, 0) Model
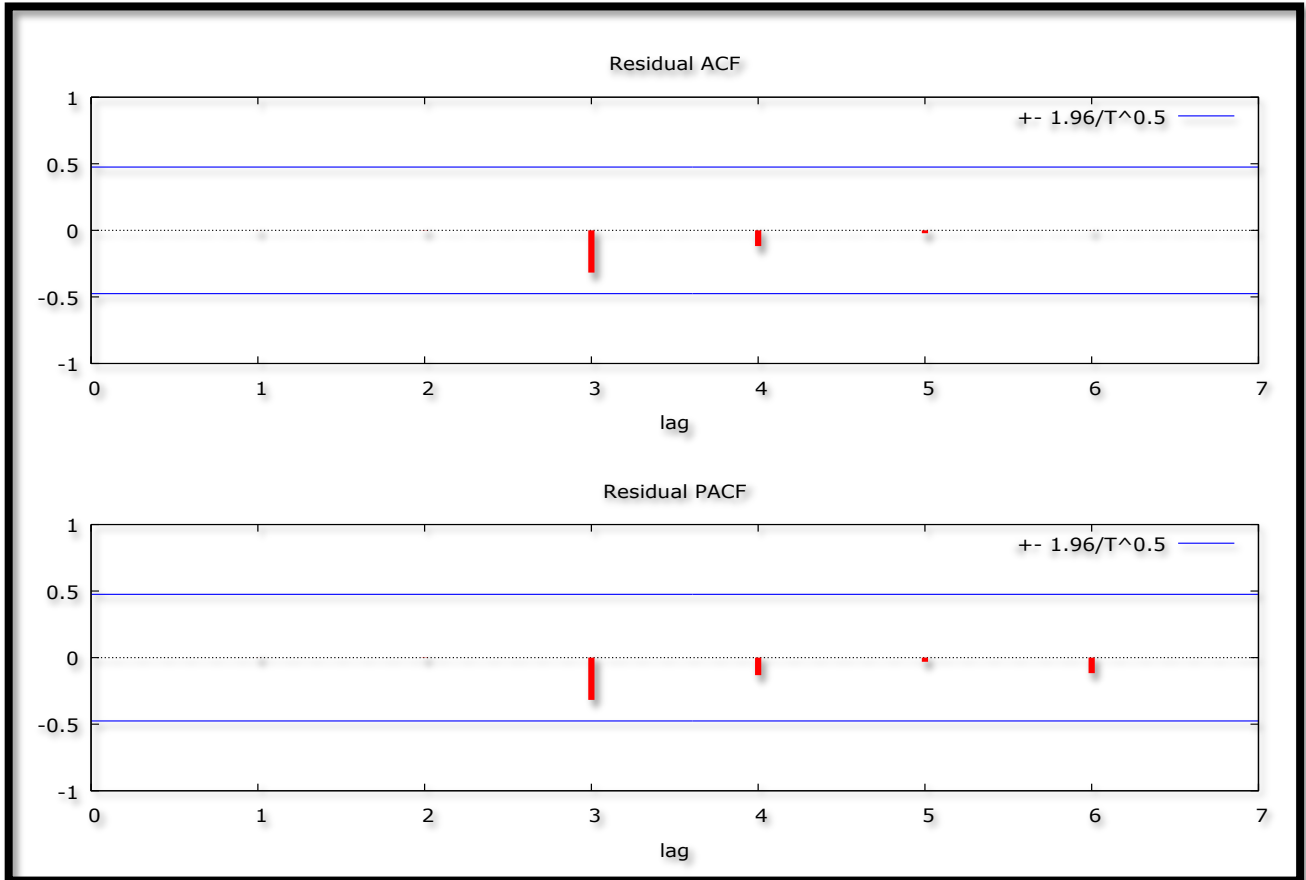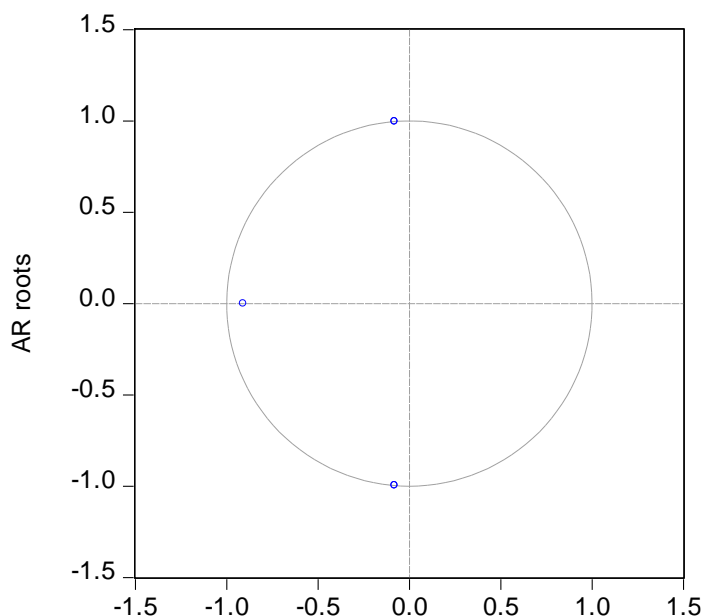
Figure 4: Correlogram of the Residuals

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

3.8.3 Stability Test of the ARIMA () Model

Figure 5: Inverse Roots

Inverse Roots of AR/MA Polynomial(s)

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 (3, 1, 0) model is really stable and suitable for forecasting annual number of people practicing open defecation in Malawi.

4.0 FINDINGS
4.1 Descriptive Statistics

Table 8: Descriptive Statistics

Description	Statistic
Mean	10.561
Median	10.55
Minimum	5.7
Maximum	15.4

As shown in table 8 above, the mean is positive, that is, 10.561. This means that, over the study period, the annual average number of people practicing open defecation in Malawi is approximately 11% of the total population. The minimum number of people practicing open defecation in Malawi over the study period is approximately 5.7% of the total population, while the maximum is 15.4% of the total population. In fact, the number of people practicing open defecation in Eritrea has declined over the years from 15.4% in 2000 to 5.7% of the total population in 2017.

4.2 Results Presentation[1]

Table 9: Main Results

ARIMA (3, 1, 0) Model:				
Guided by equation [4], the chosen optimal model, the ARIMA (3, 1, 0) model can be expressed as follows: $\Delta M_t = -0.573240 - 0.999387\Delta M_{t-1} - 0.998994\Delta M_{t-2} - 0.735813M_{t-3} \dots \dots \dots \dots \dots \dots [5]$				
Variable	Coefficient	Standard Error	z	p-value
constant	-0.573240	0.00161148	-355.7	0.0000***
β_1	-0.999387	0.163216	-6.123	0.0000***
β_2	-0.998994	0.162371	-6.153	0.0000***
β_3	-0.735813	0.175084	-4.203	0.0000***

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

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

---

[1] *The *, ** and *** imply statistical significance at 10%, 5% and 1% levels of significance; respectively.*
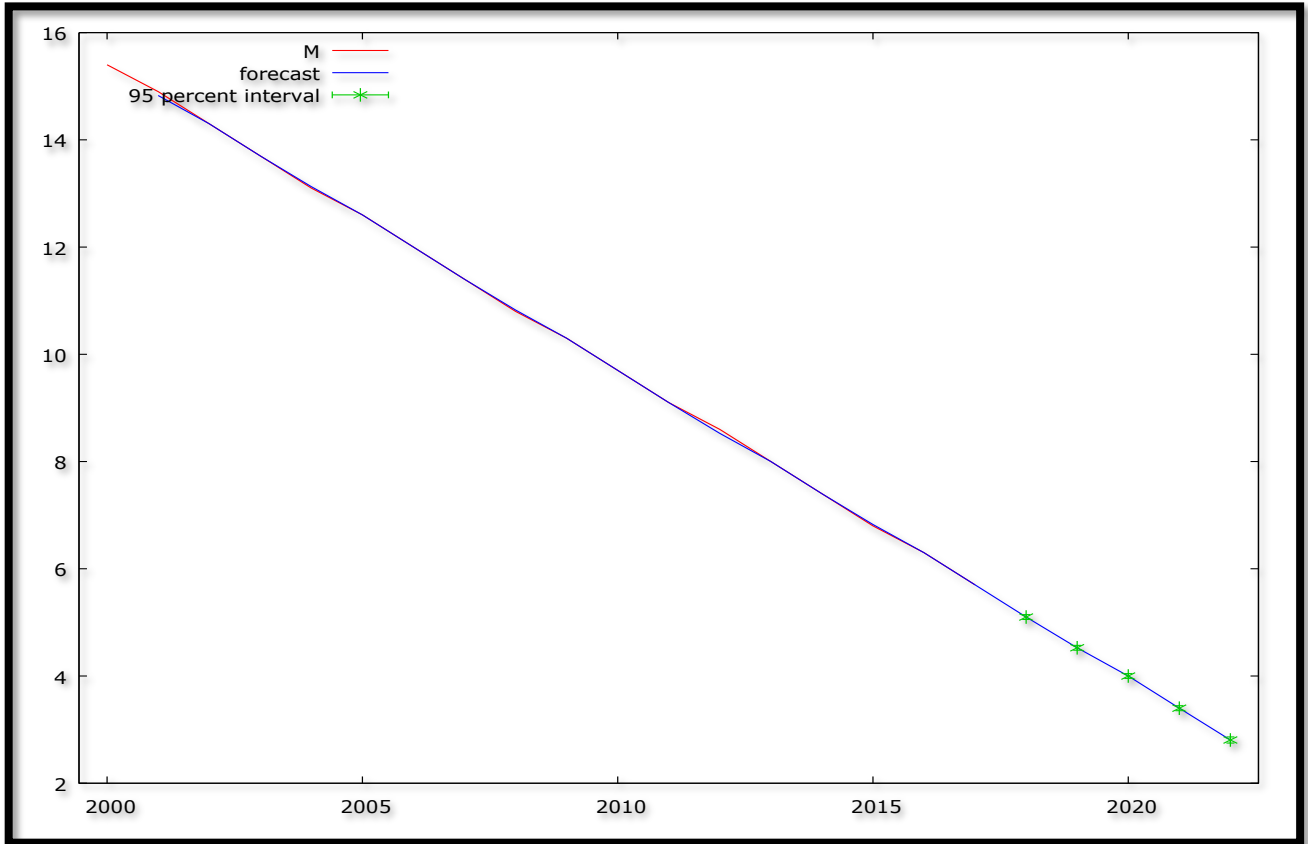
Figure 6 shows the in-and-out-of-sample forecasts of the M series. The out-of-sample forecasts cover the period 2018 – 2022.

Predicted M – Out-of-Sample Forecasts Only

Table 10: Predicted

Year	Predicted M	Standard Error	Lower Limit	Upper Limit
2018	5.1	0.02	5.1	5.1
2019	4.5	0.02	4.5	4.6
2020	4.0	0.02	4.0	4.0
2021	3.4	0.02	3.4	3.4
2022	2.8	0.03	2.7	2.9

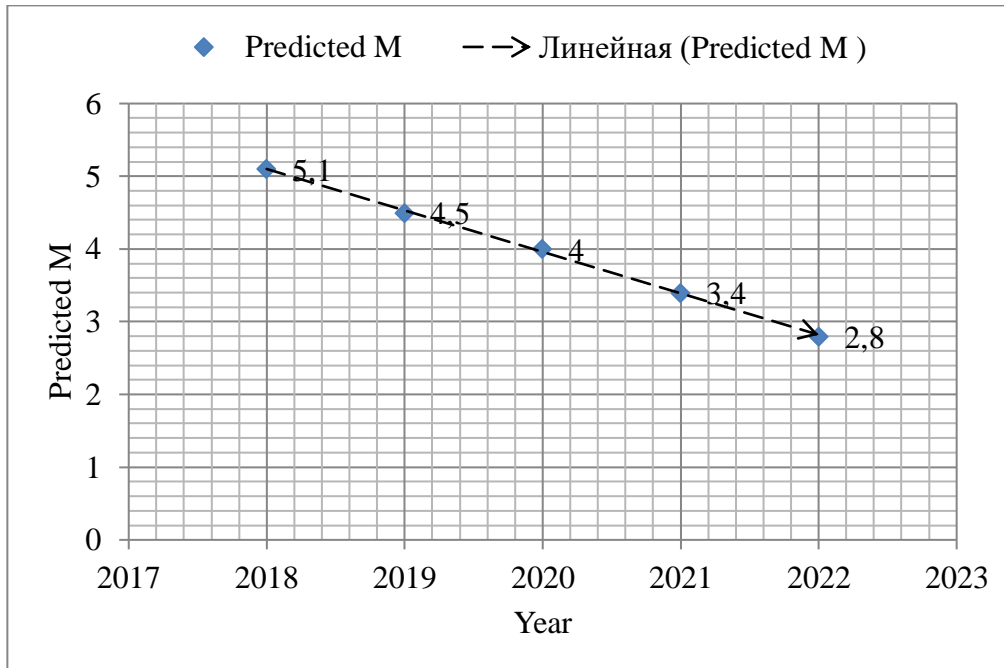Figure 7: Graphical Analysis of Out-of-Sample Forecasts

Table 10 and figure 7 show the out-of-sample forecasts only. The number of people practicing open defecation in Malawi is projected to fall from approximately 5.1% in 2017 to 2.8% of the total population by the year 2022. Indeed, it is possible to significantly reduce the number of open defecators in Malawi, especially if the current government adopts the policy directions suggested below. By 2030, open defecation could be completely eliminated in Malawi.

4.3 Policy Implications

i. The government of Malawi ought to continue making toilets a status symbol, especially through supporting programmes such as the Community Led Total Sanitation (CLTS) initiative.

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

5.0 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 Malawi over the period 2018 – 2022. The model predicts a significant decrease in the annual number of people practicing open defecation in Malawi. Hence, the war against open defecation is almost won in Malawi. These findings are important for the government of Malawi, 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] Ayalew, A. M., *et al*. (2018). Assessment of Diarrhea and Its Associated Factors in Under-five Children Among Open Defecation and Open Defecation Free Rural Settings of Dangla District, Northwest Ethiopia, *Journal of Environmental and Public Health*, pp: 1 – 8.

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

[4] 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.

[5] Njuguna, J., & Muruka, J. (2017). Open Defecation in Newly Created Kenyan Counties: A Situational Analysis, *Journal of Health Care for the Poor and Underserved*, 28 (1): 71 – 78.

[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] 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.

[10] UNICEF (2018). Game Plan to End Open Defecation, *UNICEF*, New York.