OPEN DEFECATION IN MADAGASCAR: 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 research article, the Box-Jenkins ARIMA model was applied in analyzing open defecation in Madagascar. The data was gathered 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 indicate that the open defection series is an I (1) variable. The study finally presents the ARIMA (2, 1, 0) as the best model in forecasting the number of people practicing open defecation in Madagascar. The model, through the residual ADF tests, the Normality Test as well as the inverse roots of the AR/MA polynomials; has been shown to be stable and suitable for forecasting and control of open defecation in Madagascar. The results of the study are surprising in the sense that open defecation has been shown to be increasing in the out-of-sample period. This means that, if nothing is done on the public policy front, by 2022, almost 47% of the total population in Madagascar [that is, at least 9 million people, especially those in rural areas] will be practicing open defecation. This is a serious warning signal of the likelihood of a more dangerous health threat due to open defecation. The study offers the following recommendations: [i] the government of Madagascar should make toilets a status symbol; [ii] the government of Madagascar should create more demand for sanitation; [iii] there is need for the government of Madagascar to encourage a habit of systematic hand-washing, not defecating in the open, as well as keeping toilets fly-proof. This will easily translate into an open-defecation-free Madagascar; [iv] the Community Led Total Sanitation (CLTS) program in Madagascar should continue indefinitely.

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

1.0 INTRODUCTION

Madagascar is one of the poorest countries in the world (UNICEF, 2013) and has some of the worst child development and survival rates in the world; risk factors include open defecation which is a major contributing factor to the high incidence of diarrhea. Unfortunately, 10 400 Malagasy, including 6 900 children under the age of five die each year as a result of diarrheal diseases (Government of Madagascar, 2016). In fact, open defecation, coupled with the lack of safe, clean water and poor hygiene practices like hand washing with soap is estimated to be directly responsible for causing 90% of diarrhea cases and has a direct influence on high levels of chronic malnutrition in Madagascar (UNICEF, 2013 & 2015). Eliminating open defecation is increasingly seen as a key health outcome, with links to reduced stunting, improved educational and positive health outcomes for children (Thomas & Bevan, 2014).

1.2 OBJECTIVES OF THE STUDY

i. To investigate the nature of open defecation in Madagascar.
ii. To forecast the number of people practicing open defecation in Madagascar for the period 2018 – 2022.
iii. To examine the trend of open defecation for the period 2018 – 2022.

2.0 LITERATURE REVIEW

Alhassan & Anyarayor (2018) investigated 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 used to collect data. Their study revealed that while effective communication of innovation resulted in widespread awareness, low income levels significantly accounted for households' inability to sustain and utilize latrines.

Nyoni (2019) predicted total population in India using the Box-Jenkins ARIMA technique based on annual time series data on total population in India from 1960 to 2017. The researcher presented the ARIMA (1, 2, 3) model and concluded that total population in India will continue to sharply rise in the next three decades, hence posing a threat to both natural and non-renewable resources. This is actually a worse threat if the open defecation puzzle is left unpuzzled in India. Adhikari & Ghimire (2020) analyzed various determinants of open defecation in Nepal. Bivariate analysis was done to examine the association between dependent variables (toilet status – having and not having toilets in the household) and independent variables (demographic, socio-economic and geographical characteristics) using the Chi-square test. The multivariate logistic regression model was employed to assess significant predictors for a household not having a toilet after controlling other variables. The results of the study show that Nepal still has a large number of residences without a toilet. No study has been done so far, in Madagascar, to model and forecast the number of people practicing open defecation. This is envisioned to reinforce existing policy frameworks in the fight against open defecation in Madagascar.

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 ODM series under consideration.

3.2 The Moving Average (MA) model

Given:

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

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

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

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

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

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 ODM series has been found to be an I (1) variables (that is, it only became stationary after first differencing). Therefore, in this research paper, the model presented below is the one that will be applied.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

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

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

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

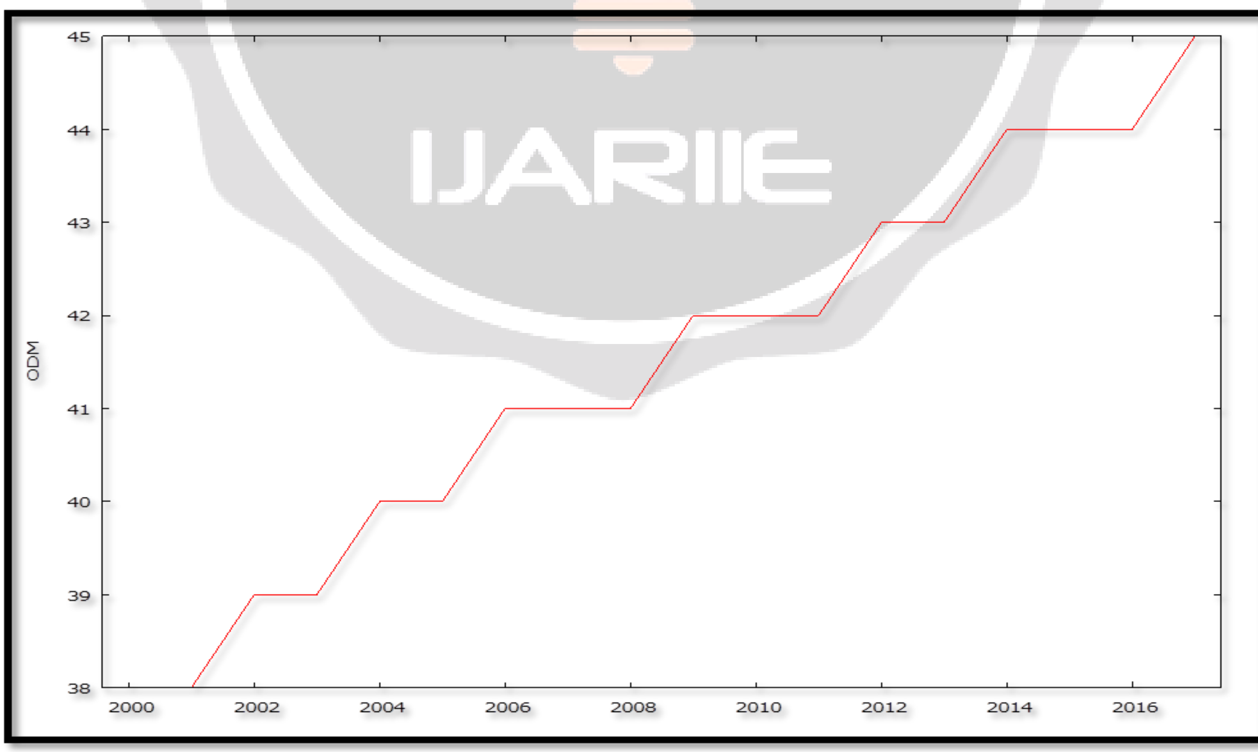
3.6 Data Collection

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

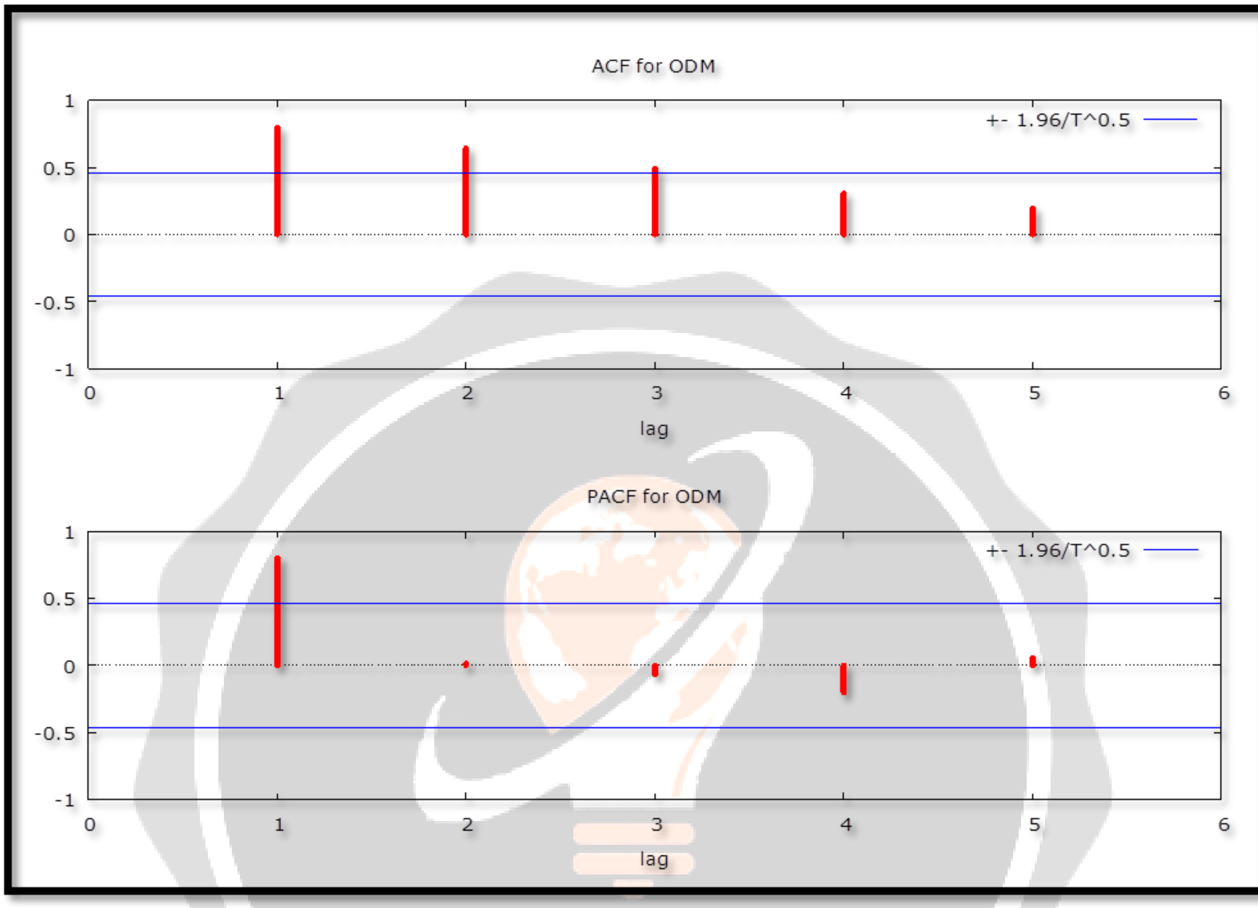
3.7.1 Stationarity Tests: Graphical Analysis

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
ODM	-1.136257	0.6790	-3.959148	@1%	Non-stationary
			-3.081002	@5%	Non-stationary
			-2.681330	@10%	Non-stationary

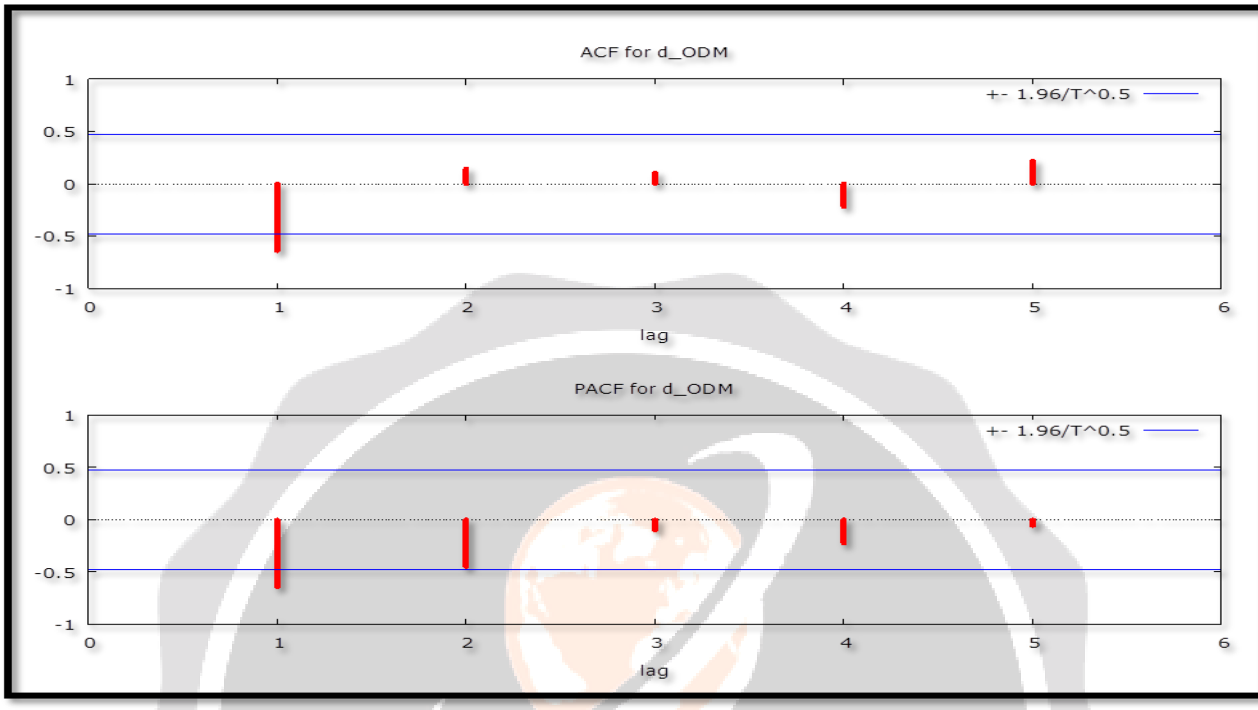
Table 2: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODM	-5.252411	0.0032	-4.616209	@1%	Stationary
			-3.710482	@5%	Stationary
			-3.297799	@10%	Stationary

Tables 1 and 2 show that ODM 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
ΔODM	-5.477226	0.0006	-3.959148	@1%	Stationary
			-3.081002	@5%	Stationary
			-2.681330	@10%	Stationary

Table 4: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODM	-5.503654	0.0028	-4.728363	@1%	Stationary
			-3.759743	@5%	Stationary
			-3.324976	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that ODM 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)	20.03844	0.53556	-0.016457	0.27415	0.36522	0.65741
ARIMA (2, 1, 0)	**17.73907**	0.47528	-0.010426	0.21973	0.32266	0.53304
ARIMA (3, 1, 0)	19.73907	0.47528	-0.010427	0.21973	0.32266	0.53304
ARIMA (4, 1, 0)	21.21321	0.47107	-0.003278	0.21405	0.31805	0.52083
ARIMA (5, 1, 0)	23.21321	0.47107	-0.0032799	0.21405	0.31805	0.52082
ARIMA (6, 1, 0)	25.12536	0.47052	0.0054713	0.21877	0.31741	0.53176

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 (2, 1, 0) model is finally 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	-5.239205	0.0008	-3.920350	@1%	Stationary
			-3.065585	@5%	Stationary
			-2.673459	@10%	Stationary

Table 7: without intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
R	-5.404978	0.0028	-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 (2, 1, 0) model; are stationary. Hence, the model is indeed stable.

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

Figure 4: Normality of the Residuals

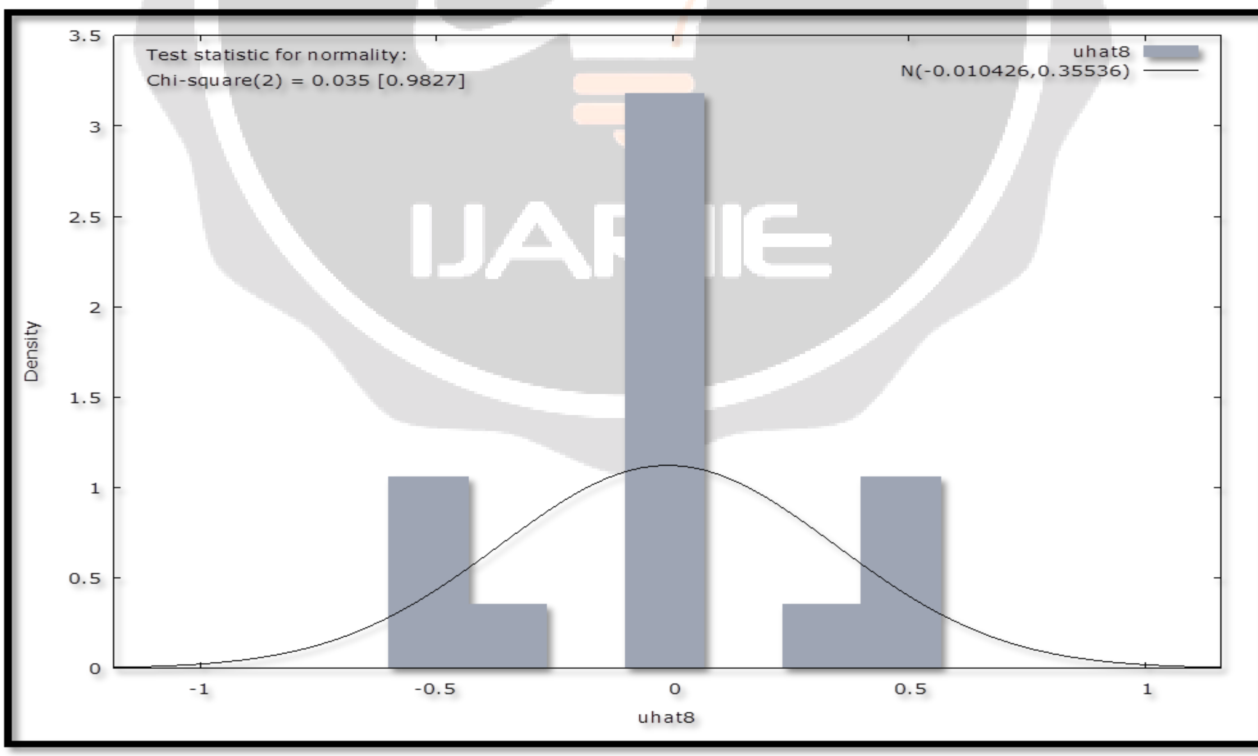


Figure 4 shows that the residuals of the ARIMA (2, 1, 0) model are normally distributed as shown by the p-value of the Chi-square statistic, which is statistically insignificant. This implies that the model is stable as confirmed by figure 6 below.

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

Figure 5: Correlogram of the Residuals

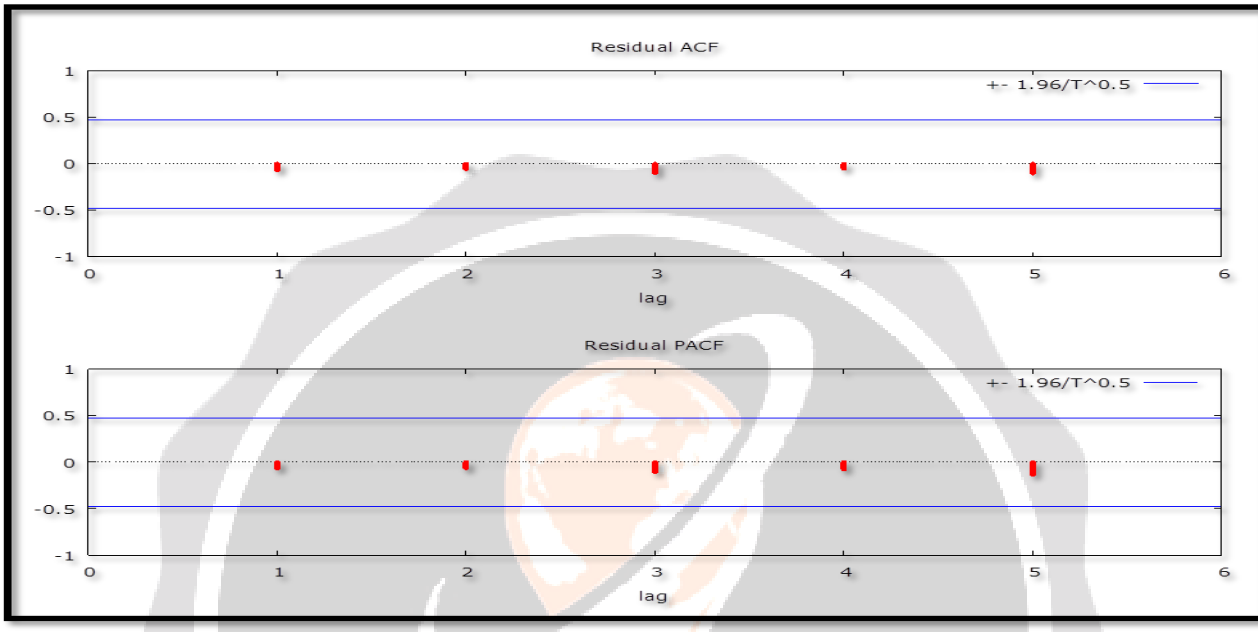
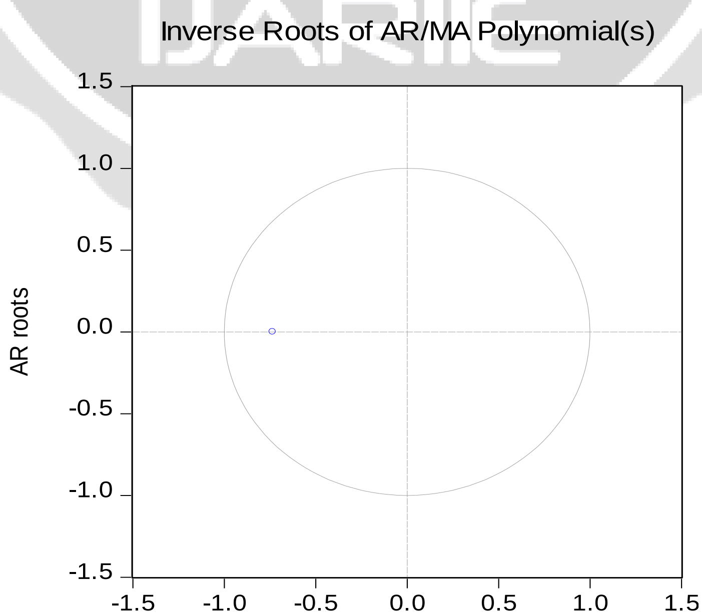


Figure 5 indicates that the estimated model is adequate since ACF and PACF lags are indeed short and within the bands. This points to the fact that the "no autocorrelation" assumption is not violated in this study.

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

4.0 FINDINGS

4.1 Descriptive Statistics

Table 8: Descriptive Statistics

Description	Statistic
Mean	41.444
Median	41.5
Minimum	38
Maximum	45

As shown in table 8 above, the mean is positive, that is, 41.444. This means that, over the study period, the annual average number of people practicing open defecation in Madagascar is approximately 41% of the total population. This is a warning alarm for policy makers in Madagascar with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation over the study period is approximately 38% of the total population, while the maximum is 45% of the total population. In fact, the number of people practicing open defecation in Madagascar has continued to increase over the years from 38% in 2000 to 45% of the total population in 2017.

4.2 Results Presentation[1]

Table 9: Main Results

ARIMA (2, 1, 0) Model:

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

$$\Delta ODM_t = 0.402686 - 1.00195\Delta OD_{t-1} - 0.485167\Delta OD_{t-2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.\ldots\ldots\ldots.\ldots.. [5]$$

Variable	Coefficient	Standard Error	z	p-value
constant	0.402686	0.0317176	12.7	0.0000***
$\emptyset_1$	-1.00195	0.219187	-4.571	0.0000***
$\emptyset_2$	-0.485167	0.220309	-2.202	0.0277**

Table 9 shows the main results of the optimal ARIMA (2, 1, 0) model.

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

Forecast Graph

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

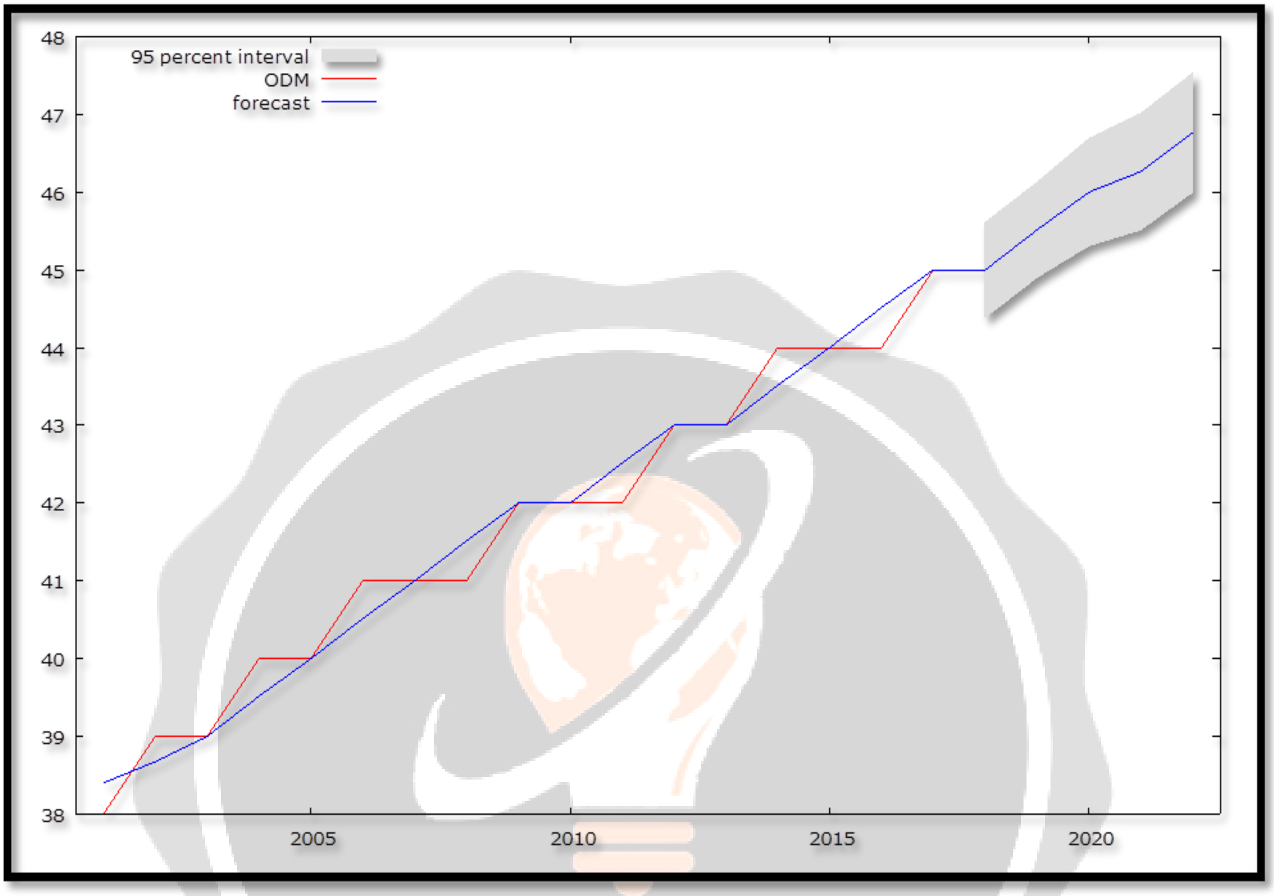


Figure 7 shows the in-and-out-of-sample forecasts of the ODM series. The out-of-sample forecasts cover the period 2018 – 2022.

Predicted ODM – Out-of-Sample Forecasts Only

Table 10: Predicted ODM

Year	Predicted ODM	Standard Error	Lower Limit	Upper Limit
2018	45	0.312	44.39	45.61
2019	45.52	0.312	44.91	46.13
2020	46	0.351	45.31	46.69
2021	46.27	0.382	45.52	47.01
2022	46.77	0.39	46	47.53

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

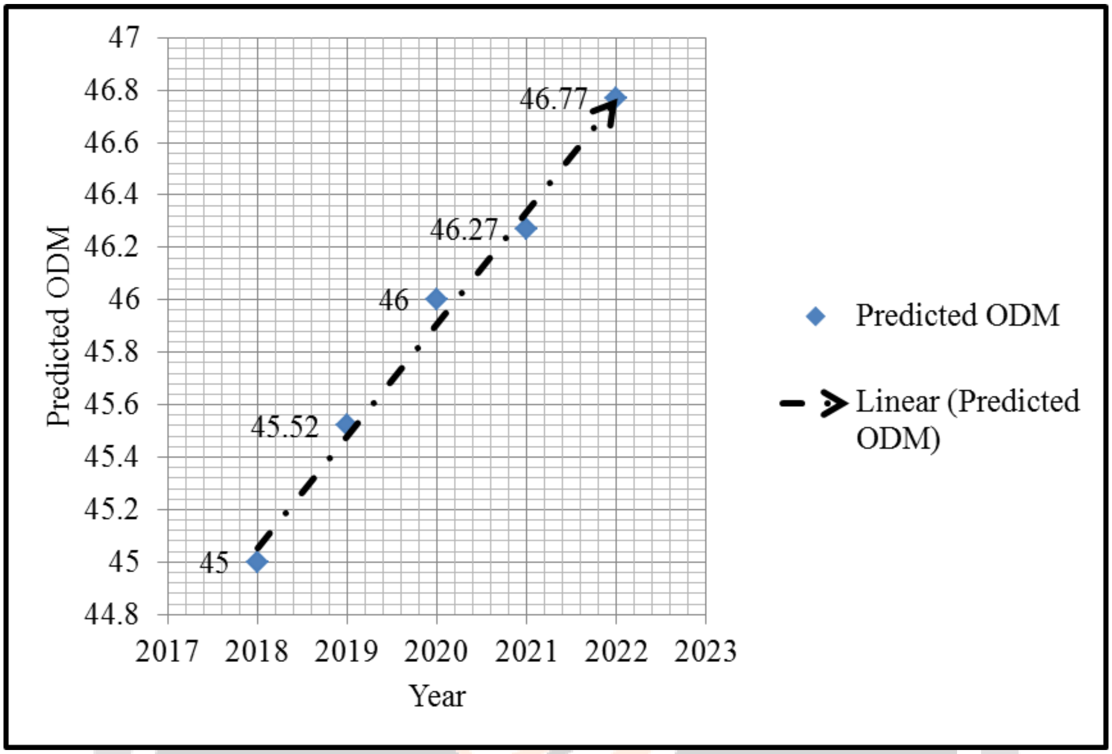


Table 10 and figure 8 show the out-of-sample forecasts only. The number of people practicing open defecation in Madagascar is projected to rise sharply from approximately 45% to nearly 47% of the total population by the year 2022. These results are quite surprising given the current sanitation and hygiene initiatives being implemented in Madagascar, for example the Community Led Total Sanitation (CLTS) program. Further studies should examine determinants of open defecation in Madagascar; that could help uncover the reasons as to why the current initiatives are not producing the intended health outcomes in terms of significantly reducing the number of open defecators.

4.3 Policy Implications

i. The government of Madagascar should make toilets a status symbol.
ii. The government of Madagascar should create more demand for sanitation.
iii. There is need for the government of Madagascar to encourage a habit of systematic hand-washing, not defecating in the open, as well as keeping toilets fly-proof. This will easily translate into an open-defecation-free Madagascar.
iv. The Community Led Total Sanitation (CLTS) program in Madagascar should continue indefinitely.

5.0 CONCLUSION

The study reveals that the optimal model, the ARIMA (2, 1, 0) model is not only stable but also the most suitable model to forecast the annual number of people practicing open defecation in Madagascar over the period 2018 – 2022. The model predicts a sharp increase in the annual number of people practicing open defecation in Madagascar. Such a trend can be reversed if the above policy directions are taken seriously.

REFERENCES

[1] Adhikari, R., & Ghimire, S. (2020). Open Defecation Free: Where Do We Need To Focus? *Health Prospect: Journal of Public Health*, 19 (1): 1 – 6.

[2] 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.

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

[4] Government of Madagascar (2016). Investing in Water, Sanitation and Hygiene in Madagascar – The Business Case, *Government of Madagascar*.

[5] Nyoni, T (2018b). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, *Dimorian Review*, 5 (6): 16 – 40.

[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] Nyoni, T. (2019). Predicting Total Population in India: A Box-Jenkins ARIMA Approach, MPRA Paper No. 92436, *University Library of Munich*, Munich, Germany.

[9] Nyoni, T. (2019). Predicting Total Population in India: A Box-Jenkins ARIMA Approach, MPRA Paper No. 92436, *University Library of Munich*, Munich, Germany.

[10] Thomas, A., & Bevan, J. (2014). Developing and Monitoring Protocol for the Elimination of Open Defecation in Sub-Saharan Africa, *UNICEF*, New York.

[11] UNICEF (2013). A Silent WASH Crisis, *UNICEF*, New York.

[12] UNICEF (2015). Using Social Norms Theory to Strengthen CLTS in Southern Madagascar, *UNICEF*, East and Southern Africa Regional Office.