NOVATEUR PUBLICATIONS
INTERNATIONAL JOURNAL OF INNOVATIONS IN ENGINEERING RESEARCH AND TECHNOLOGY [IJIERT]
ISSN: 2394-3696
VOLUME 7, ISSUE 6, June-2020

OPEN DEFECATION IN SOMALIA: A BOX-JENKINS ARIMA APPROACH

DR. SMARTSON PUMULANI NYONI
ZICHIRe, 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 Somalia 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 employs the Box-Jenkins ARIMA approach. The diagnostic ADF tests show that the ODX series under consideration is an I (1) variable. Based on the AIC, the study presents the ARIMA (2, 1, 0) model as the best model. The diagnostic tests further reveal that the presented model is indeed stable and its residuals are stationary in levels. The results of the study indicate that the number of people practicing open defecation in Somalia is likely to decline over the period 2018 – 2022, from 26.16% to 18.65% of the total population. In order to effectively sustain this desirable downwards trend, the study suggested a three-fold policy recommendation to be put into consideration, especially by the Somalian government.

INTRODUCTION
Open defecation rates in Somalia are some of the highest in the world and hand-washing practices with soap are very low (UNICEF, 2020). Access to safe water is low in Somalia. Access to sanitation facilities is again very low in Somalia. In fact, 38% of people have access to basic sanitation facilities, 20% rural and 61% urban, with 28% people defecating in the open, 49% in rural and 1% in urban areas (UNICEF, 2019). Open defecation has negative consequences for health (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). Therefore, it has become inevitable for public health researchers to forecast the number of people practicing open defecation in order to formulate effective policies to end open defecation. The main objective of this study is to predict the number of people practicing open defecation in Somalia over the period 2018 – 2022. This study will go a long way in assessing the possibility of ending open defecation in Somalia.

1.2 OBJECTIVES OF THE STUDY
i. To analyze the pattern of open defecation in Somalia over the period 2000 – 2017.
ii. To forecast the number of people practicing open defecation in Somalia for the period 2018 – 2022.
iii. To examine the trend of open defecation in Somalia for the out-of-sample period.

LITERATURE REVIEW
Gitau & Flachenberg (2016) presented the lessons learnt from a pilot CLTS intervention in the challenging context of Somalia and generally found out that it is still possible to trigger to a great extent a community "with a desire for change with regards to open defecation and facilitate them to build their own household latrines without subsidization". Their study also revealed that commitment, attitude and mind-set of the implementers, community and the local authority are key to success in any context – fragile or stable. 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 employed to gather data. The study showed 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) forecasted total population in Somalia using the Box-Jenkins ARIMA technique based on annual time series data on total population in Somalia from 1960 to 2017. The study presented the ARIMA (7, 2, 1) model and concluded that total population in Somalia will

continue to sharply rise in the next three decades, thereby posing a threat to both natural and non-renewable resources; without proper sanitation and hygiene (in the midst of very high open defecation rates), that would be a time bomb in terms of disease transmission, especially cholera and faecal-related diseases. This study will adopt the ARIMA method in analyzing open defecation trends in Somalia 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 ODX series under consideration.

3.2 The Moving Average (MA) model

Given:

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

where L is the lag operator.

or as:

$$\mathrm{ODX_t} = \alpha(L)\mu_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots [2]$$

where:

$\alpha(L)=\theta(L)$ …………………………………………………………………………………….. [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). ODX is the annual number of people (as a percentage of the total population) who practice open defecation in Somalia 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:

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

Or that:

$$\beta(L)\mathrm{ODX_t} = \mu_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots [5]$$

where:

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

or that :

$$\mathrm{ODX_t} = (\beta_1 L + \cdots + \beta_p L^p)\mathrm{ODX_t} + \mu_t \ldots \ldots \ldots \ldots \ldots \ldots \ldots [7]$$

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

The ARMA (p, q) process is just a combination of AR (p) and MA (q) processes. Thus, by combining equations [1] and [4]; an ARMA (p, q) process can be specified as shown below:

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

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

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

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

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 application, many time series are non – stationary. In this study, the ODX series has been found to be an I (1) variables (that is, it only became stationary after first differencing). Based on that, ARMA models are not suitable for modeling and forecasting non – stationary time series data. In such a scenario, the model described below is the one that should ideally be used.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

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

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

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, through out this paper, denoted as ODX), as a percentage of total population in Somalia. 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

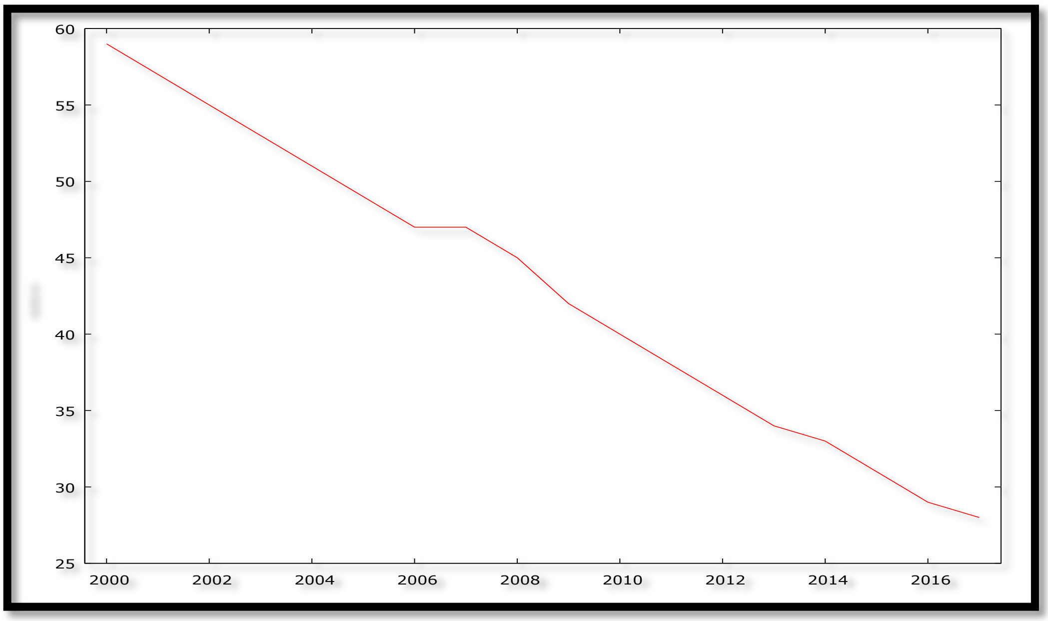

Figure 1

3.7.2 The Correlogram in Levels

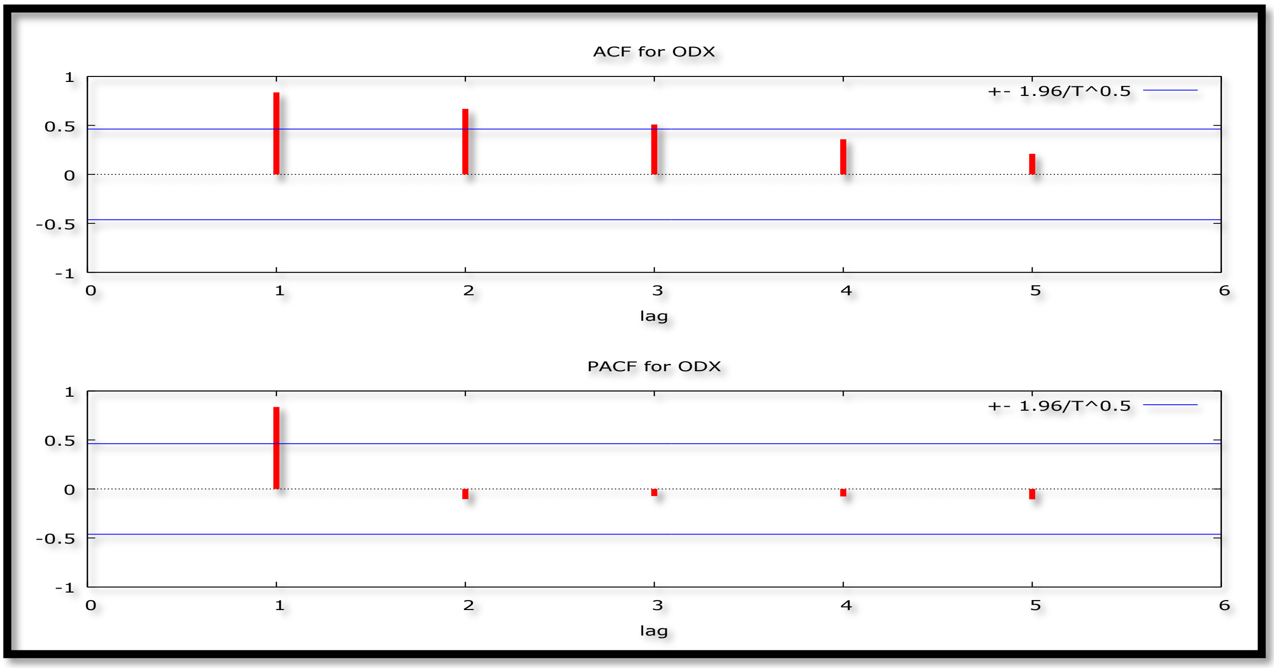

Figure 2: Correlogram in Levels

3.7.3 The ADF Test in Levels

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODX	-0.809929	0.7904	-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
ODX	-2.448490	0.3453	-4.616209	@1%	Non-stationary
			-3.710482	@5%	Non-stationary
			-3.297799	@10%	Non-stationary

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

3.7.4 The Correlogram (at First Differences)

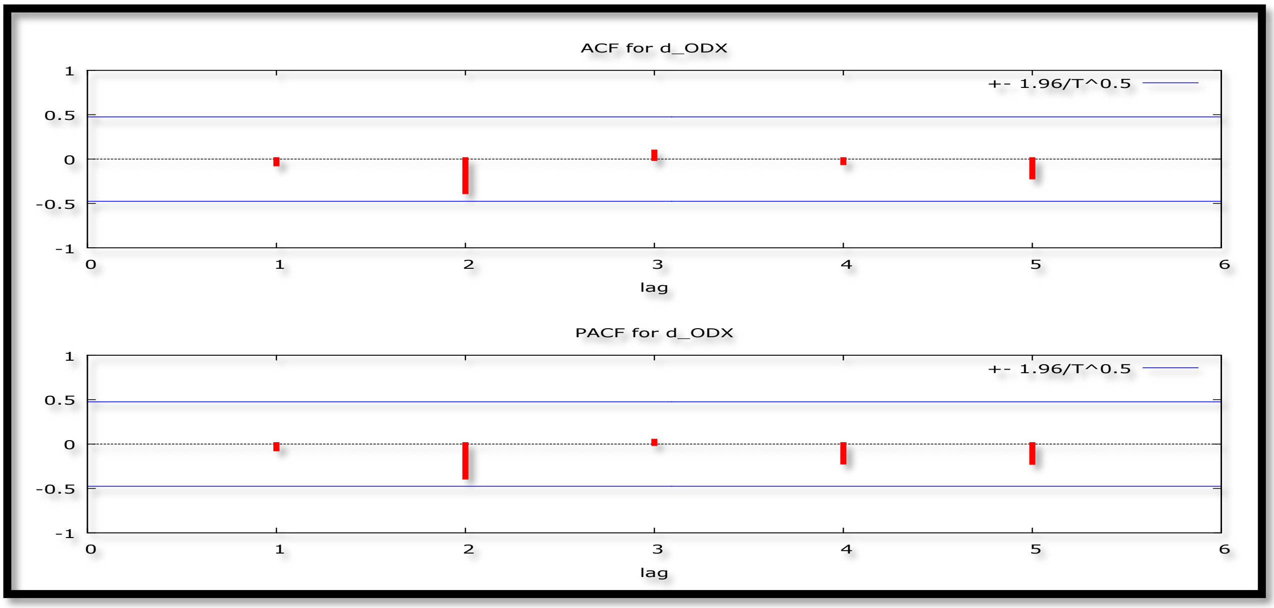

Figure 3: Correlogram (at First Differences)

3.7.5 The ADF Test (at First Differences)

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODX	-3.774019	0.0132	-3.920350	@1%	Non-stationary
			-3.065585	@5%	Stationary
			-2.673459	@10%	Stationary

Table 4: with intercept and trend & intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODX	-3.709658	0.0520	-4.667883	@1%	Non-stationary
			-3.733200	@5%	Non-stationary
			-3.310349	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that ODX 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)	37.75934	0.33313	-0.00064343	0.40589	0.61573	1.0306
ARIMA (2, 1, 0)	**37.11112**	0.30775	-0.0081507	0.40681	0.56488	1.0309
ARIMA (3, 1, 0)	39.08686	0.30628	-0.0071394	0.40547	0.56442	1.0259
ARIMA (4, 1, 0)	40.14607	0.2973	-0.017035	0.4048	0.5455	1.0311

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. Hence, 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	-3.456494	0.0242	-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	-3.452927	0.0796	-4.667883	@1%	Non-stationary
			-3.733200	@5%	Non-stationary
			-3.310349	@10%	Stationary

Tables 6 and 7 indicate that the residuals of the selected optimal model, the ARIMA (2, 1, 0) model; are stationary. Thus, the model is stable.

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

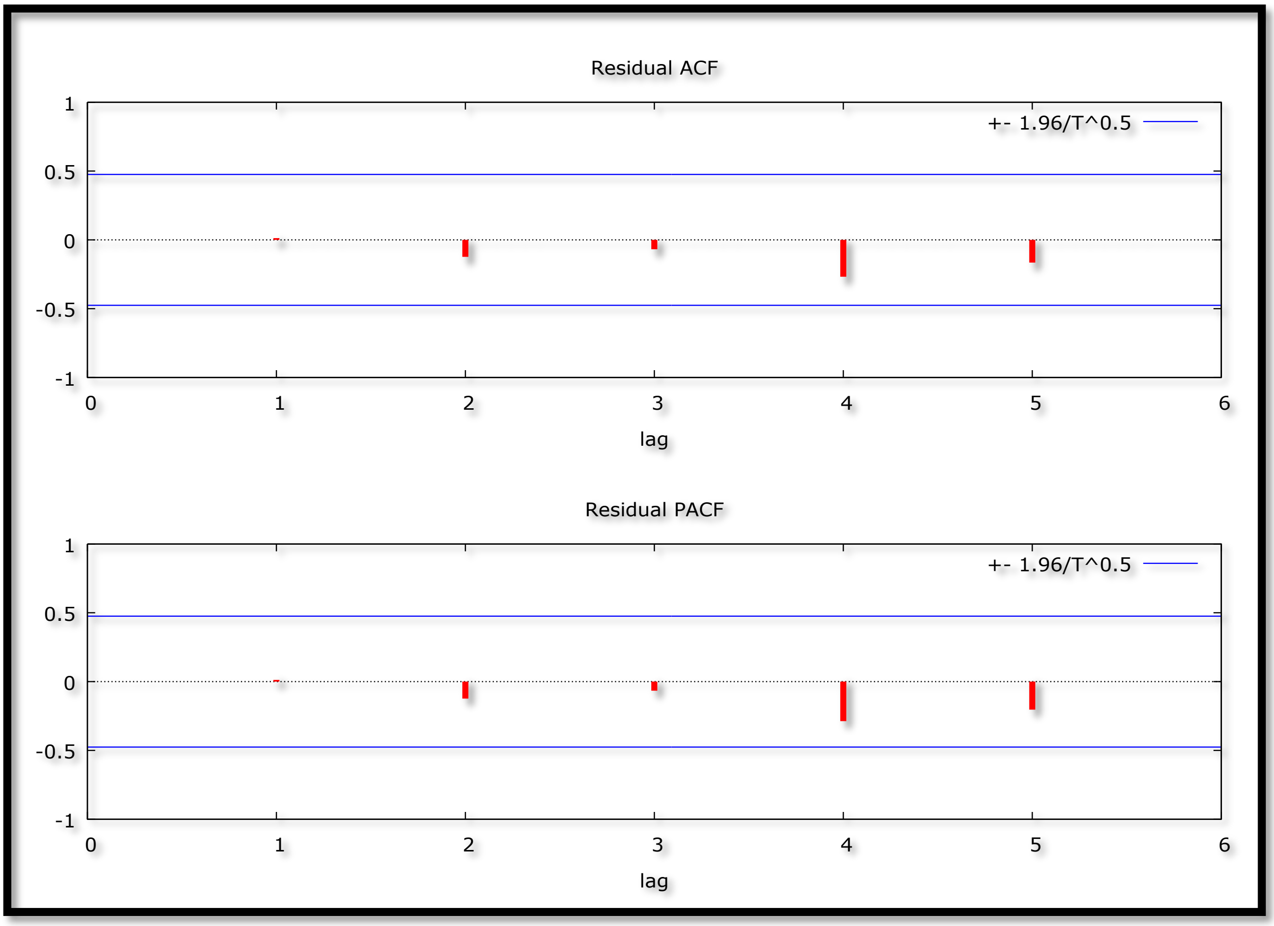

Figure 4: Correlogram of the Residuals

Figure 4 indicates that the chosen ARIMA (2, 1, 0) model is adequate since ACF and PACF lags are quite short and within the bands. This shows that the "no autocorrelation" assumption is not violated in this study.

3.8.3 Stability Test of the ARIMA (2, 1, 0) Model

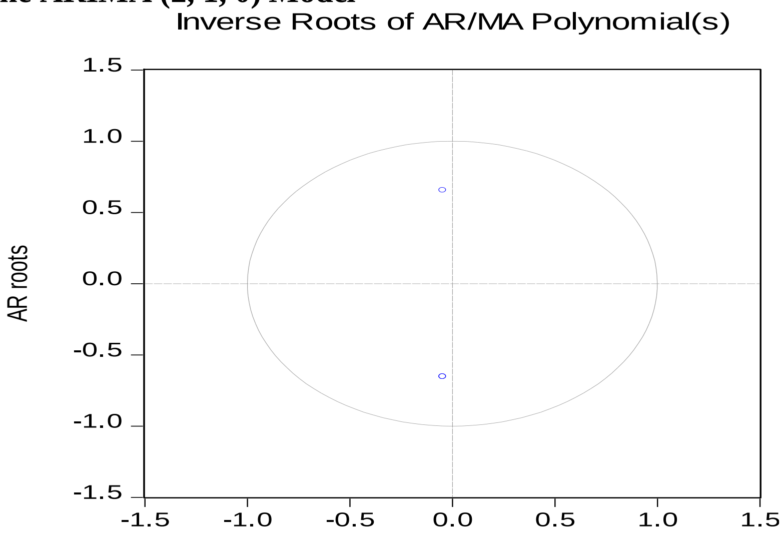

Figure 5: Inverse Roots

Since all the AR roots lie inside the unit circle, it implies that the estimated ARIMA process is (covariance) stationary; hence indicating that, indeed, the ARIMA (2, 1, 0) model is stable and suitable for forecasting annual number of people practicing open defecation in Somalia.

FINDINGS

4.1 Descriptive Statistics

Table 8: Descriptive Statistics

Description	Statistic
Mean	43
Median	43.5
Minimum	28
Maximum	59
Standard deviation	9.8876
Skewness	0.022542
Excess kurtosis	-1.2402

As shown in table 8 above, the mean is positive, that is, 43. This implies that, over the study period, the annual average number of people practicing open defecation in Somalia is approximately 43% of the total population. This is a warning alarm for Somalian policy makers with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation in Somalia over the study period is approximately 28% of the total population, while the maximum is 59% of the total population. In fact, the number of people practicing open defecation in Somalia, as already shown in figure 1, has continued to decline over the years from 59% in 2000 to 28% of the total population in 2017. This is a desirable public health outcome and there is need to intensify policies and strategies that discourage the practice of open defecation in Somalia.

4.2 Results Presentation

Table 9: Main Results

ARIMA (2, 1, 0) Model:

Guided by equation [11], the selected optimal model, the ARIMA (2, 1, 0) can be expressed as follows:

$$\Delta ODX_t = -1.83066 - 0.0831696\Delta ODX_{t-1} - 0.37897\Delta ODX_{t-2} \ldots\ldots\ldots\ldots\ldots\ldots [12]$$

Variable	Coefficient	Standard Error	z	p-value
constant	-1.83066	0.0969001	-18.89	0.0000***
β_1	-0.0831696	0.235780	-0.3527	0.7243
β_2	-0.378970	0.235948	-1.606	0.1082

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

Forecast Graph

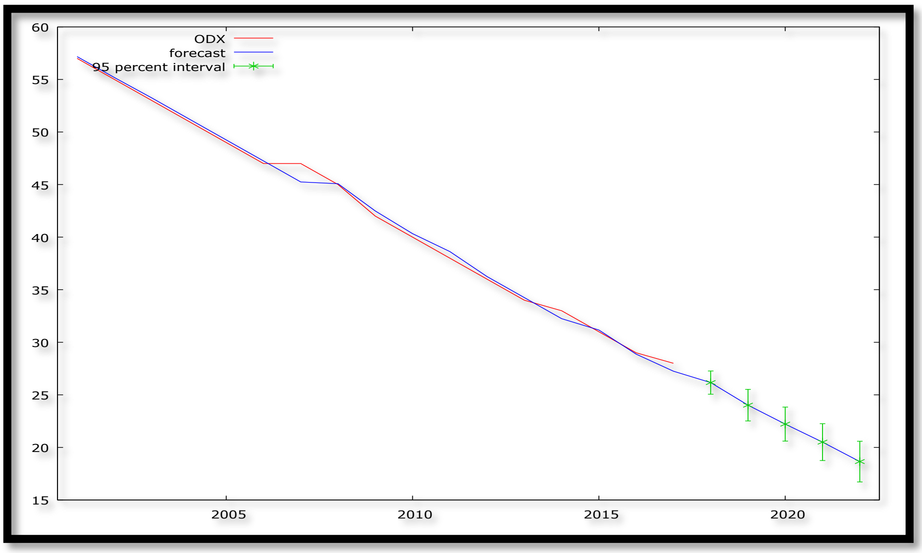

Figure 6: Forecast Graph – In & Out-of-Sample Forecasts

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

Predicted ODX – Out-of-Sample Forecasts Only

Table 10: Predicted ODX

Year	Predicted ODX	Standard Error	Lower Limit	Upper Limit
2018	26.16	0.564	25.06	27.27
2019	24.02	0.766	22.52	25.52
2020	22.22	0.825	20.60	23.83
2021	20.50	0.894	18.75	22.25
2022	18.65	0.987	16.72	20.59

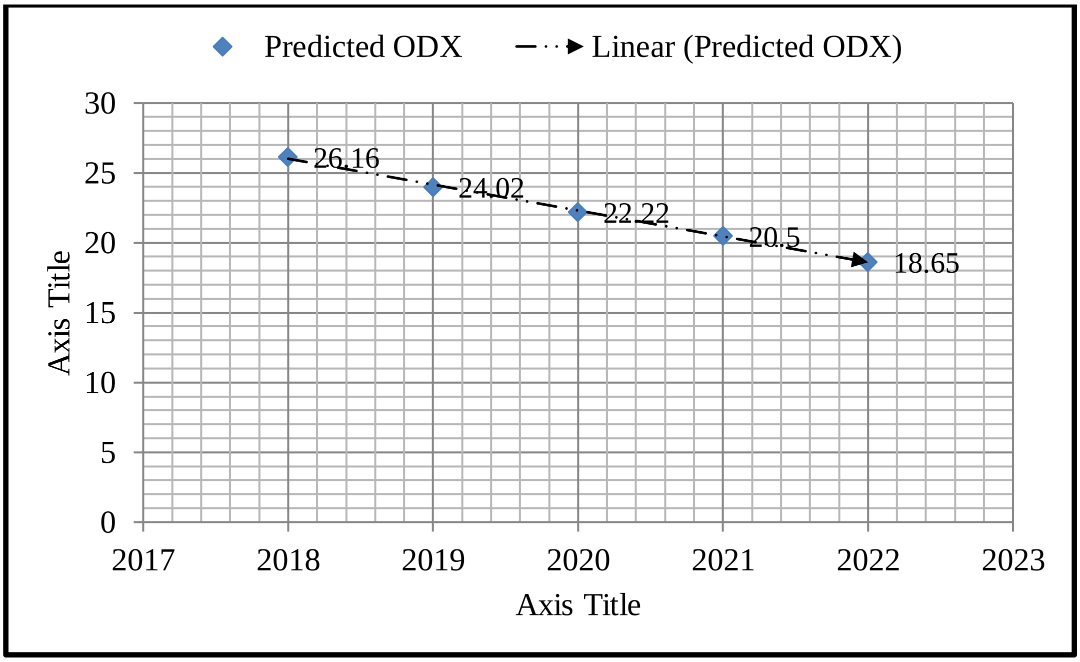

Figure 7: Graphical Analysis of Out-of-Sample Forecasts

Table 10 and figure 7 show the out-of-sample forecasts only. The number of people practicing open defecation in Somalia is projected to fall from approximately 26% in 2018 to as low as 19% of the total population by the year 2022. Clearly, it is possible for Somalians to abandon open defecation.

4.3 Policy Implications

i. The government of Somalia should build toilets for its citizens since most households in Somalia do not afford to build themselves a decent toilet. In this regard the country's WASH programmes are commendable and are expected to continue until open defecation is eliminated.

ii. The government of Somalia should continue to make toilets a status symbol so that people stop the habit of defecating in the open.

iii. The government of Somalia should create more demand for sanitation, especially in light of disease transmission and other risks associated with open defecation.

CONCLUSION

The study shows that the chosen 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 Somalia over the period 2018 – 2022. The model predicts a sharp decrease in the annual number of people practicing open defecation in Somalia. These results are essential for the government of Somalia, especially when it comes to 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) Gitau, M., & Flachenberg, F. (2016). Ensuring Availability and Sustainable Management of Water and Sanitation for All, 39th WEDC International Conference, Kumasi.
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) 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 (2019). Somalia Population Dynamics Versus the Malthusian Population Trap: What Does the ARIMA Approach Tell Us? MPRA Paper No. 92457, University Library of Munich, Germany.
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.
11) UNICEF (2019). WASH Profile, UNICEF, New York.
12) UNICEF (2020). WASH Strategy Note, UNICEF, New York.