Special Issue on Basis of Applied Sciences and Its Development in the Contemporary World
Published in Association with
Department of Technology and Organization of Construction, Samarkand State Architectural and Civil Engineering Institute, Uzbekistan
Department of Mechanization of livestock, Samarkand Institute of Veterinary Medicine, Uzbekistan
Novateur Publication India's International Journal of Innovations in Engineering Research and Technology [IJIERT]
ISSN: 2394-3696, Website: www.ijiert.org, 15th June, 2020

OPEN DEFECATION IN ETHIOPIA: 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 Ethiopia 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 applies the Box-Jenkins ARIMA methodology. The diagnostic ADF tests indicate that the 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 indicate 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 Ethiopia is likely to decline over the period 2018 – 2022, from as high as 19% to as low as 5.69% of the total population. In order to sustainably maintain this desirable downwards trend, the study suggested a three-fold policy recommendation to be put into consideration, especially by the Ethiopian government.

INTRODUCTION

Open defecation is a widespread problem in the developing world and this practice facilitates the transmission of diarrheal diseases (Ayalew et al., 2018). Ethiopia has high morbidity and mortality linked with acute diarrhea (Abera et al., 2018) which is directly linked to open defecation. In fact, in Ethiopia, diarrhea kills half a million of under-five children annually next to pneumonia (Tecklemichael et al., 2014). Open defecation has already been studied in Ethiopia (Njuguna & Muruka, 2017; Ayalew et al., 2018; Ashenafi et al., 2018); however, no study has been done to model and forecast the number of people practicing open defecation in the country. It is important to forecast the number of people practicing open defecation in

Ethiopia in order to help public health policy makers in future planning as well as assessment of the effectiveness of current open-defecation-free-society initiatives in the country. The major objective of this study is to model and forecast the number of people practicing open defecation in Ethiopia.

LITERATURE REVIEW

In Kenya, Njuguna & Muruka (2017) examined open defecation trends among the 47 counties in Kenya, newly created in 2013. The study used four data sets on open defecation, unimproved water supply coverage, poverty levels and population density. Their results indicate 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) investigated diarrheal morbidity in under-five children and its associated factors in Dangla district in Northwest Ethiopia. A community-based comparative cross-sectional study design with a multistage random sampling technique was used. Descriptive and inferential statistics were presented. The study 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 another Ethiopian study, Ashenafi et al. (2018) assessed the latrine utilization coverage of the kebeles who have already declared open defecation free. Community-based cross-sectional study design with multistage sampling technique was employed. Bivariate and multivariate logistic regression models were fitted to identify factors associated with latrine utilization. Their results basically indicate that the extent of latrine utilization was high in the community.

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

Novateur Publication India's International Journal of Innovations in Engineering Research and Technology [IJIERT]
ISSN: 2394-3696, Website: www.ijiert.org, 15th June, 2020

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 ODE series under consideration.

3.2 The Moving Average (MA) model

Given:

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

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

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

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

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

It is not unimportant to always remember that the ARMA (p, q) model, just like the AR (p) and the MA (q) models; can only be applied for stationary time series data. In real life, many time series are non – stationary. In fact, in this study, the ODE series has been found to be an I (1) variables (that is, it only became stationary after first differencing). Based on that simple logic, ARMA models are not suitable for modeling and forecasting non – stationary time series data. In such as a case, the model described below is the one that should ideally be used.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model

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

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

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 ODE] (as a percentage of total population) in Ethiopia. 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

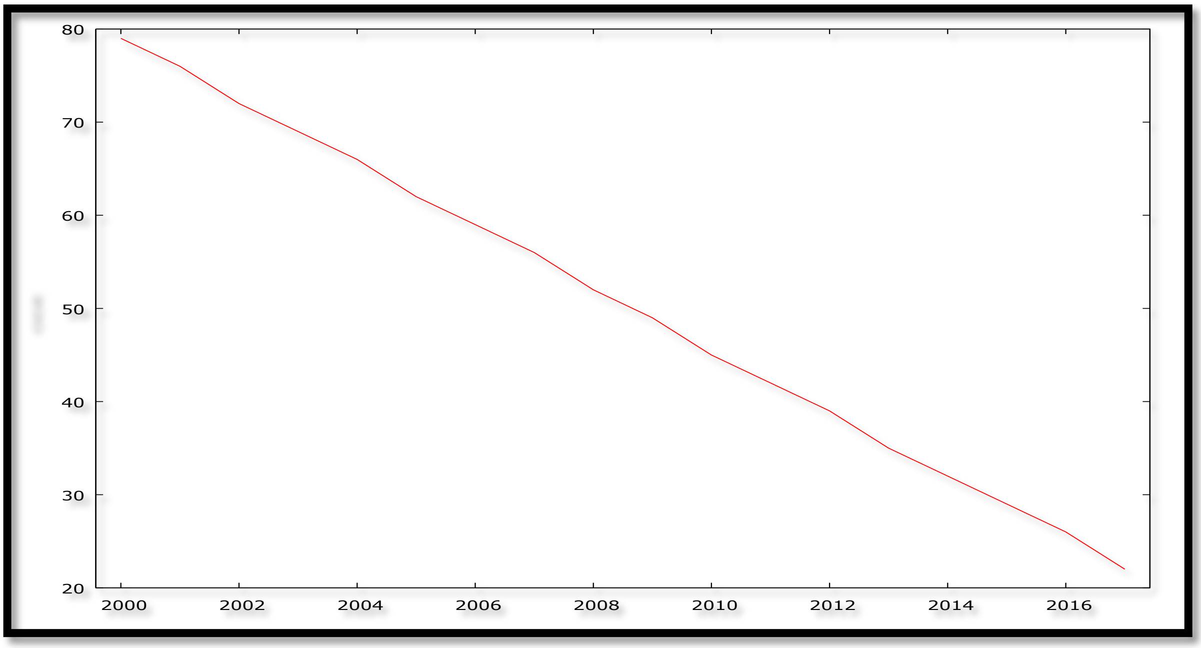

Figure 1

3.7.2 The Correlogram in Levels

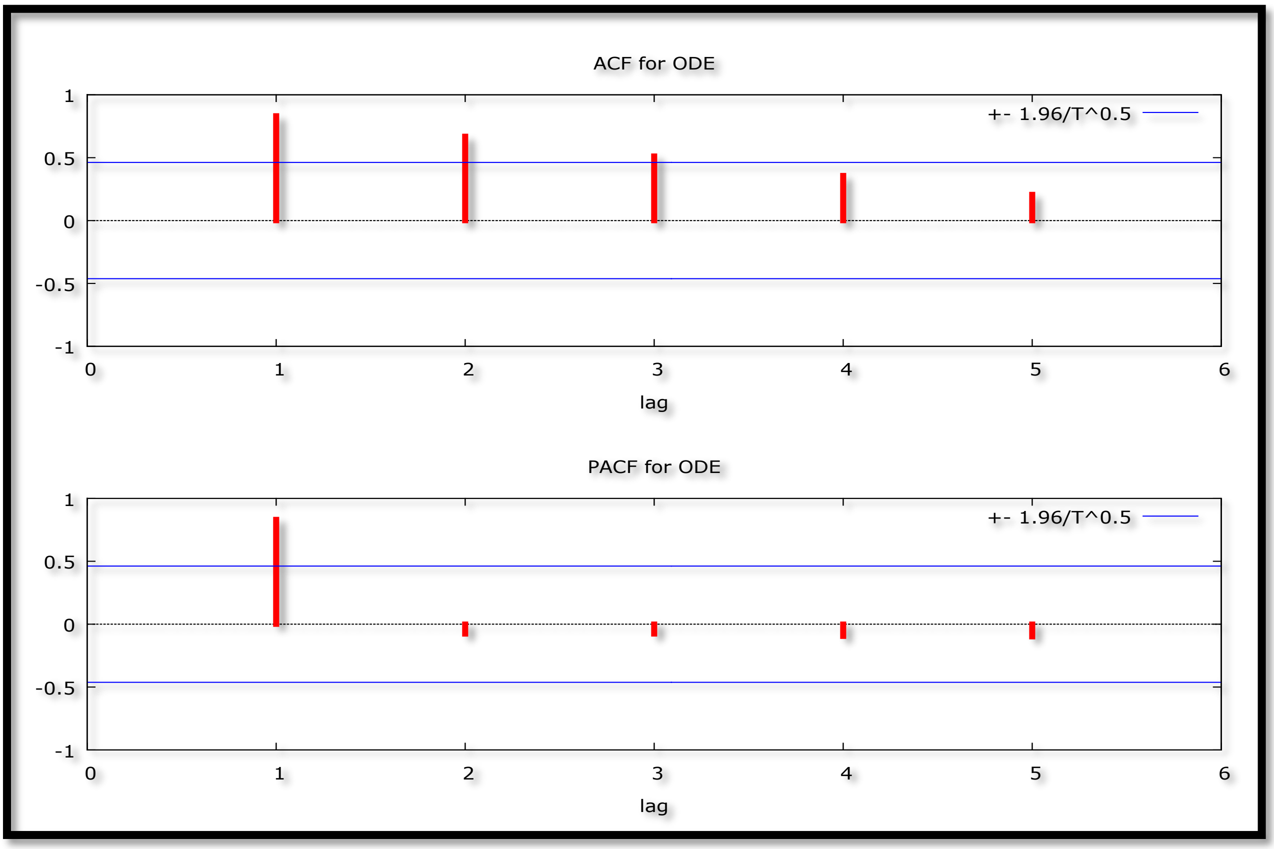

Figure 2: Correlogram in Levels

3.7.3 The ADF Test in Levels

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODE	-0.649436	0.8309	-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
ODE	-4.490293	0.0126	-4.616209	@1%	Non-stationary
			-3.710482	@5%	Stationary
			-3.297799	@10%	Stationary

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

3.7.4 The Correlogram (at First Differences)

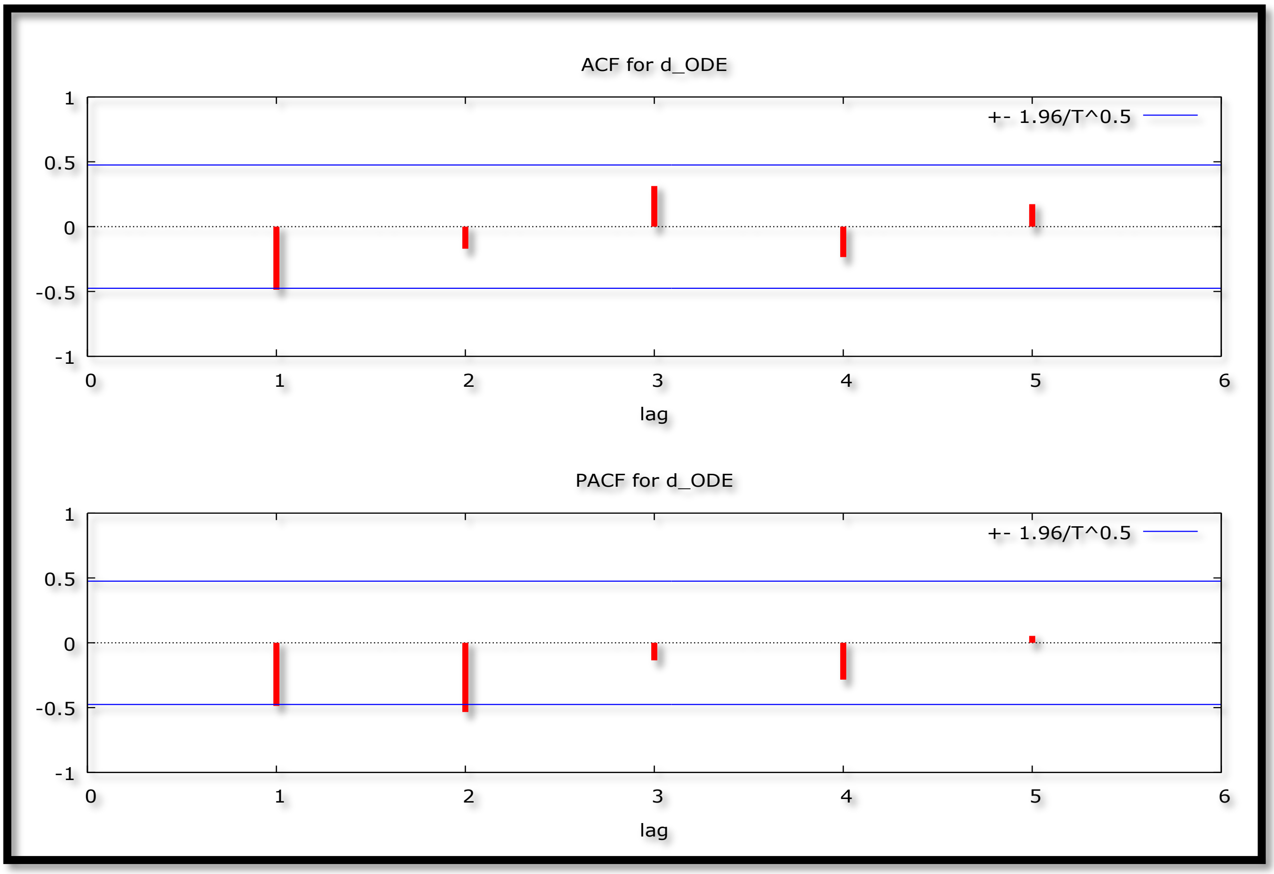

Figure 3: Correlogram (at First Differences)

3.7.5 The ADF Test (at First Differences)

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODE	-6.000000	0.0003	-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
ΔODE	-5.776457	0.0018	-4.728363	@1%	Stationary
			-3.759743	@5%	Stationary
			-3.324976	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that ODE 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)	24.01056	0.12728	0.01081	0.34224	0.40944	0.83956
ARIMA (2, 1, 0)	**19.36099**	0.1124	-0.0014016	0.23731	0.3393	0.59084
ARIMA (3, 1, 0)	21.13502	0.10928	-0.0023309	0.24507	0.33712	0.59339

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 chosen.

Novateur Publication India's International Journal of Innovations in Engineering Research and Technology [IJIERT]
ISSN: 2394-3696, Website: www.ijiert.org, 15th June, 2020

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	-4.010193	0.0084	-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	-4.262742	0.0202	-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 (2, 1, 0) model; are stationary. Hence, the model is stable.

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

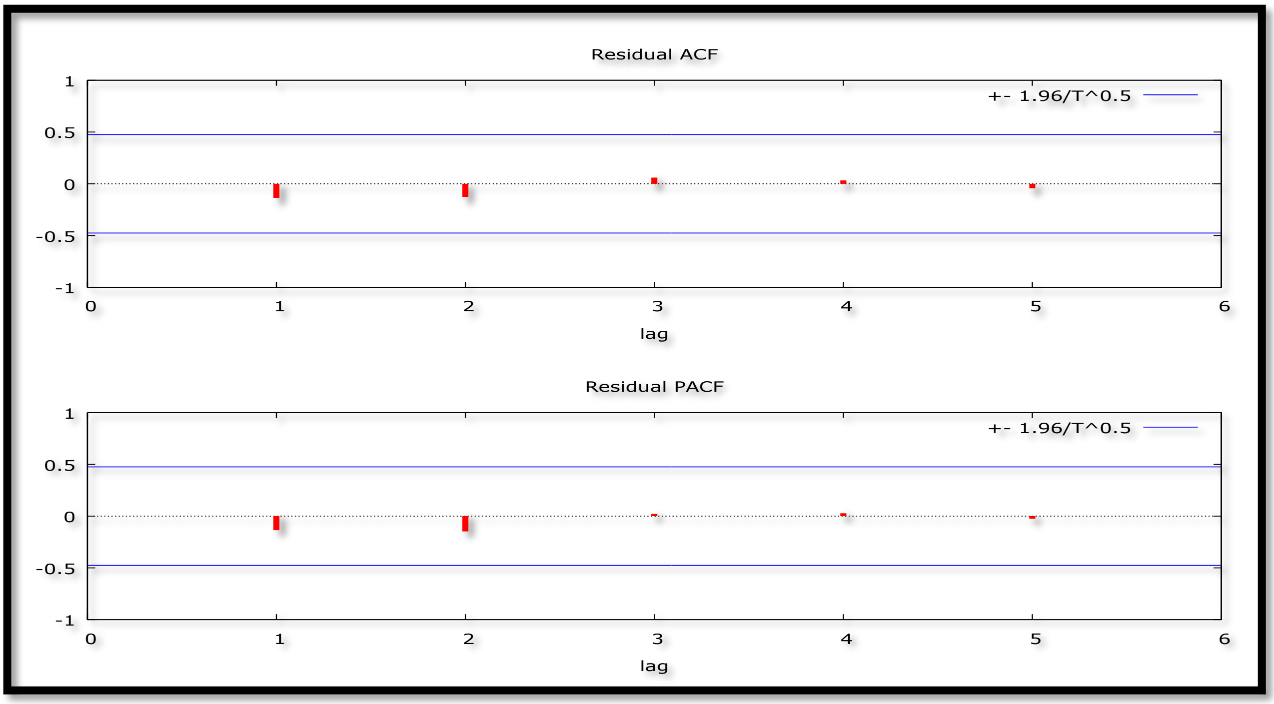

Figure 4: Correlogram of the Residuals

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

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

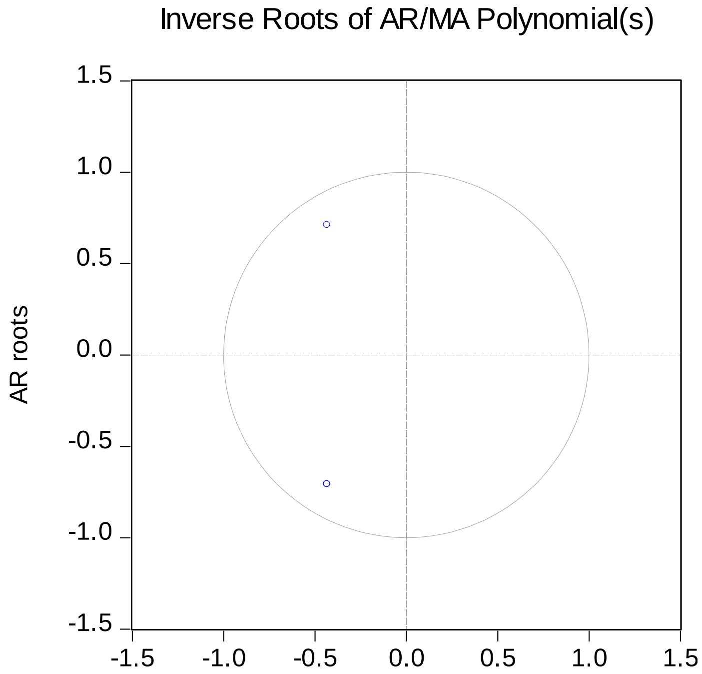

Figure 5: Inverse Roots

Because 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 stable and suitable for forecasting annual number of people practicing open defecation in Ethiopia.

FINDINGS

4.1 Descriptive Statistics

Table 8: Descriptive Statistics

Description	Statistic
Mean	50.556
Median	50.5
Minimum	22
Maximum	79

As shown in table 8 above, the mean is positive, that is, 50.556. This means that, over the study period, the annual average number of people practicing open defecation in Ethiopia is approximately 51% of the total population. This is a warning alarm for Ethiopian policy makers with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation in Ethiopia over the study period is approximately 22% of the total population, while the maximum is 79% of the total population. However, the number of people practicing open defecation in Ethiopia has continued to sharply decline over the years from 79% in 2000 to 22% of the total population in 2017.

4.2 Results Presentation

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:

$$\Delta ODE_t = -3.33696 - 0.812188\Delta ODE_{t-1} - 0.598202\Delta ODE_{t-2} \ldots\ldots\ldots\ldots\ldots\ldots.\ldots..[5]$$

Variable	Coefficient	Standard Error	z	p-value
constant	-3.33696	0.0343389	-97.18	0.0000***
β_1	-0.812188	0.205490	-3.952	0.0001***
β_2	-0.598202	0.206432	-2.898	0.0038***

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

Forecast Graph

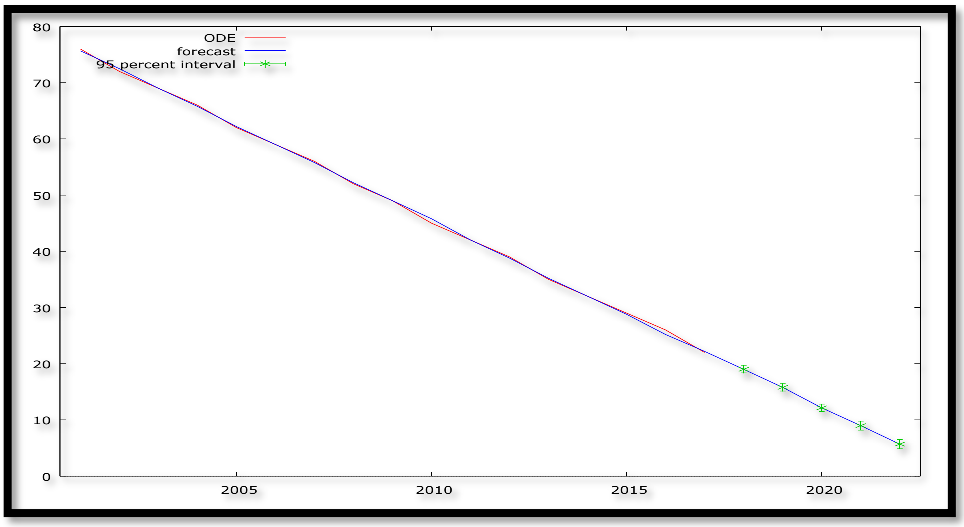

Figure 6: Forecast Graph – In & Out-of-Sample Forecasts

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

Predicted ODE – Out-of-Sample Forecasts Only

Table 10: Predicted ODE

Year	Predicted ODE	Standard Error	Lower Limit	Upper Limit
2018	19	0.326	18.36	19.64
2019	15.79	0.332	15.14	16.44
2020	12.15	0.342	11.48	12.82
2021	8.98	0.409	8.18	9.78
2022	5.69	0.42	4.86	6.51

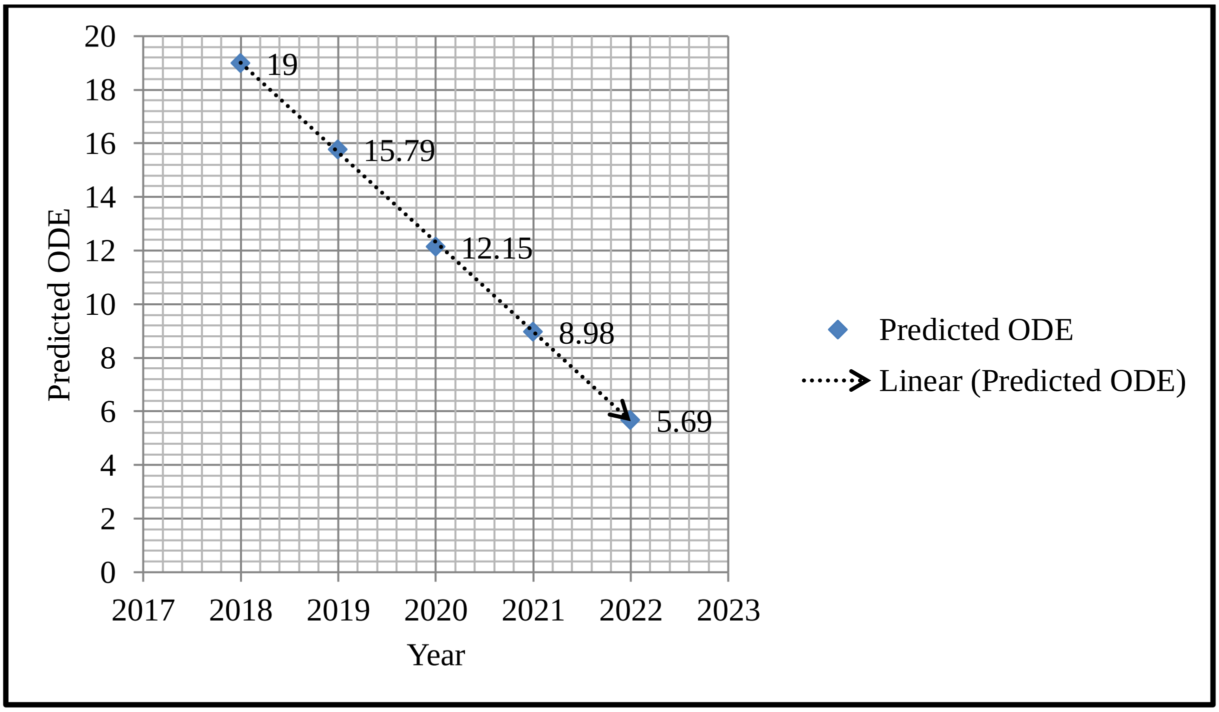

Figure 7: Graphical Analysis of Out-of-Sample Forecasts

Table 10 and figure 7 show the out-of-sample forecasts only. The number of people practicing open defecation in Ethiopia is projected to fall from approximately 19% in 2018 to as low as 5.69% of the total population by the year 2022. This could imply that the current policy framework on open defecation should be continuously implemented in Ethiopia since it is yielding acceptable health outcomes.

4.3 Policy Implications

i. The government of Ethiopia should continue to make toilets a status symbol, especially through the Community-led Total Sanitation and Hygiene (CLTSH) programs.

ii. The government of Ethiopia should continue create more demand for sanitation through teaching the public on the importance of investing in toilets.

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

CONCLUSION

The study indicates that 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 Ethiopia over the period 2018 – 2022. The model predicts a sharp decrease in the annual number of people practicing open defecation in Ethiopia. It is imperative for such a trend to be maintained and in this regard, a three-fold policy implication has been suggested. These findings are essential for the government of Ethiopia, especially when it comes to long-term planning with regards to materializing the much needed open defecation free society.

REFERENCES

[1] Abera, B., et al. (2018). Water Safety, Sanitation and Hygiene Related Knowledge, Attitudes and Practices among Household Residents in Peri-Urban Areas in Northwest Ethiopia, Ethiopian Journal of Health Development, 32 (3): 1 – 7.

[2] Ashenafi, T., et al. (2018). Latrine Utilization and Associated Factors Among Kebeles Declared Open Defecation Free in Wondo Genet District, South Ethiopia, 2015, Journal of Health and Environmental Sciences, 5 (5): 43 – 51.

[3] 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.

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

[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] Tecklemichael, G., et al. (2014). Risk Factors of Diarrheal Diseases in Under-five Children among Health Extend Model and Non-model Families in Sheko District Rural Community, Southwest Ethiopia: Comparative Cross-sectional Study, BMC Public Health, 14 (395): 1 – 9.