NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 6, ISSUE 6, June -2020

OPEN DEFECATION IN NEPAL: A BOX-JENKINS ARIMA APPROACH

DR. SMARTSON PUMULANI NYONI,
ZICHIRe, Harare, Zimbabwe

MR. THABANI NYONI
Department of Economics, University of Zimbabwe, Harare, Zimbabwe

ABSTRACT:

Employing annual time series data on the number of people who practice open defecation in Nepal from 2000 – 2017, the study forecasts the annual number of people who will still be practicing open defecation over the period 2018 – 2022. The study applies the Box-Jenkins ARIMA approach. The diagnostic ADF tests show 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 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 Nepal is likely to sharply decline over the period 2018 – 2022, from as high as 19% to as low as 8% of the total population. Therefore, it is very possible to end open defecation in Nepal. In order to significantly sustain this desirable downwards trend, the study suggested a three-fold policy recommendation to be put into consideration, especially by the government of Nepal.

INTRODUCTION:

Globally, 2.5 billion people, including 840 million children, do not use improved sanitation; 1.2 billion, almost a 5th of the world's population, practice open defecation. In rural areas, this is the case for nearly 1 in 3 people. In Nepal, 41% of the population uses an improved sanitation facility while 50% defecate in the open. Human faeces are the main source of diarrhoeal pathogens, which cause many common gastrointestinal infections: one gram of human faeces can contain 10 million viruses and 1 million bacteria (UNICEF, 2009). Open defecation is still ongoing in Nepal despite the rise in efforts for increasing latrine coverage and its use (Bhatt et al., 2019). However, the elimination of open defecation is important because it reduces diarrhoeal morbidity and mortality by approximately 36% (UNICEF, 2009) and also avoids violence against women in developing countries (Caruso et al., 2017), including rape among women and girls (Clair et al., 2018). The main purpose of this study is to model and forecast the number of people practicing open defecation in Nepal.

LITERATURE REVIEW:

Kafle & Pradhan (2018) analyzed the situation of water, sanitation and hygiene and diarrhoeal diseases after open defecation declaration. Their research was a cross-sectional study among randomly sampled 178 households using interviews and observations. The study basically revealed that water, sanitation and hygiene in the study area was lower than the criteria for open defecation free declaration. Bhatt et al. (2019) explored different motivations of people who practice open defecation in Hattimudha village in Nepal. The maximum variation sampling method was used to recruit participants for 20 in-depth interviews and 2 focus group discussions. A content analysis approach was used to analyze data. The study basically found out that open defecation is either a voluntary choice or a compulsion and that this choice is closely linked with personal preferences, cultural and traditional norms with special concerns for privacy for women and girls in different communities. 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 used to determine 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 Nepal, to model and forecast the number of people practicing open defecation. This study is the first of its kind, and is expected to strengthen existing policy frameworks in the fight against open defecation in Nepal.

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 ODP series under consideration.

3.2 The Moving Average (MA) model:

Given:

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

where L is the lag operator.

or as:

$$ODP_t = \alpha(L)\mu_t \ldots \ldots [2]$$

where:

$$\alpha(L)=\theta(L) \ldots \ldots [3]$$

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

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

Or that:

$$\beta(L)ODP_t = \mu_t \ldots \ldots [5]$$

where:

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

or that :

$$ODP_t = (\beta_1 L + \cdots + \beta_p L^p)ODP_t + \mu_t \ldots \ldots [7]$$

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

An 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 may be specified as shown below:

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

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

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

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

ARMA (p, q) models, just like the AR (p) and the MA (q) models; can only be applied for stationary time series data. But, in real life, many time series are non – stationary. In fact, in this study, the ODP series has been found to be an I (1) variables (that is, it only became stationary after first differencing). Owing to that, ARMA models are not suitable for modeling and forecasting non – stationary time series data. In such instances, the model described below is the one that should ideally be used.

3.5 The Autoregressive Integrated Moving Average (ARIMA) model:

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

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

Where Δ is the difference operator, vector $\beta \in R^p$ and $\alpha \in R^q$.

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 ODP] (as a percentage of total population) in Nepal. 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:

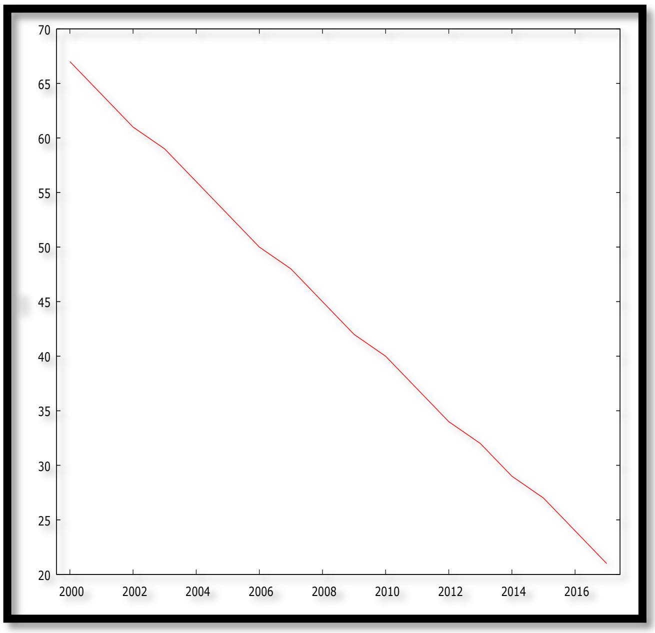

Figure 1

3.7.2 The Correlogram in Levels:

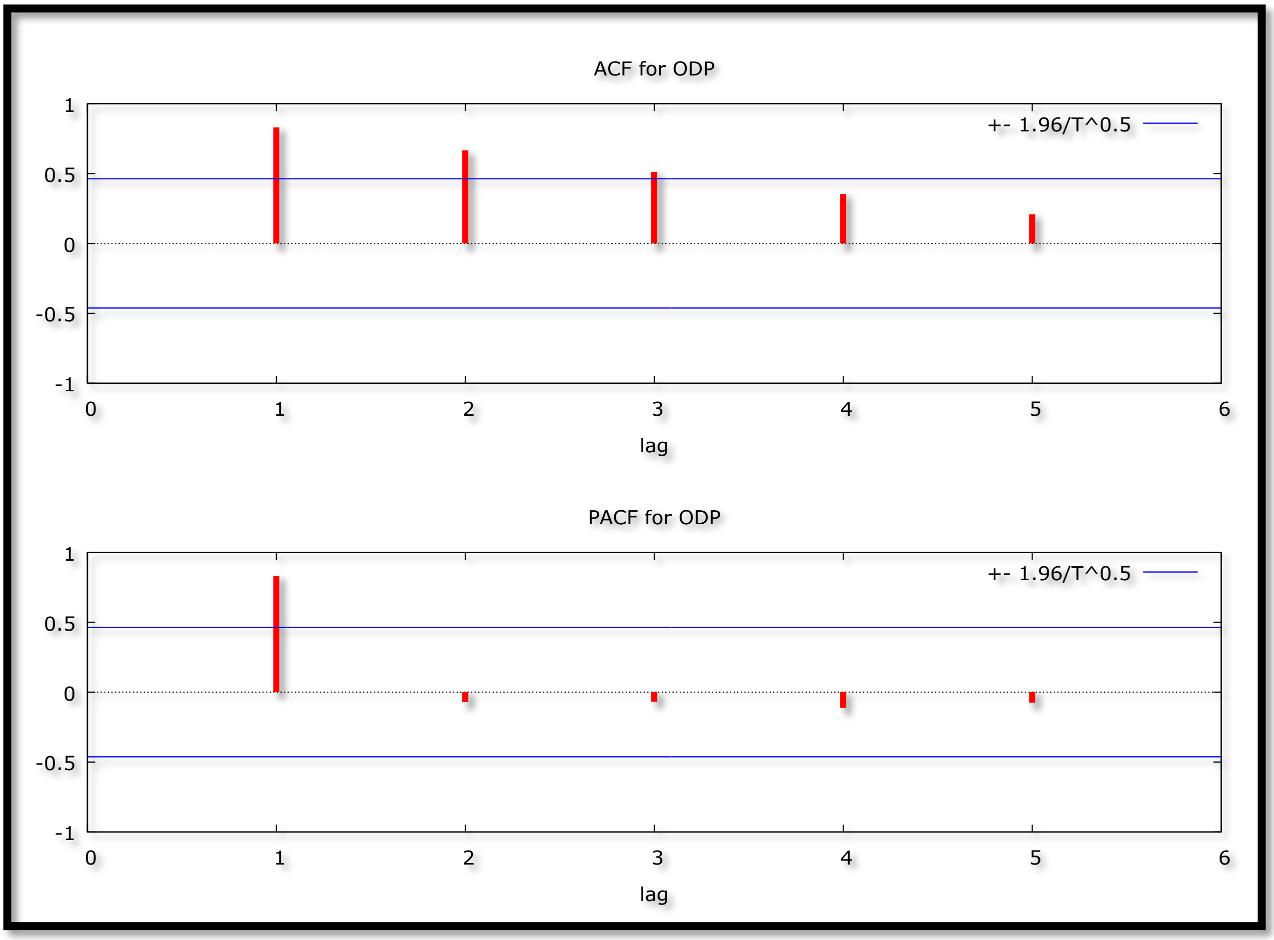

Figure 2: Correlogram in Levels

3.7.3 The ADF Test in Levels:

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ODP	-0.929200	0.7492	-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
ODP	-4.176204	0.02020	-4.616209	@1%	Non-stationary
			-3.710482	@5%	Stationary
			-3.297799	@10%	Stationary

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

3.7.4 The Correlogram (at First Differences):

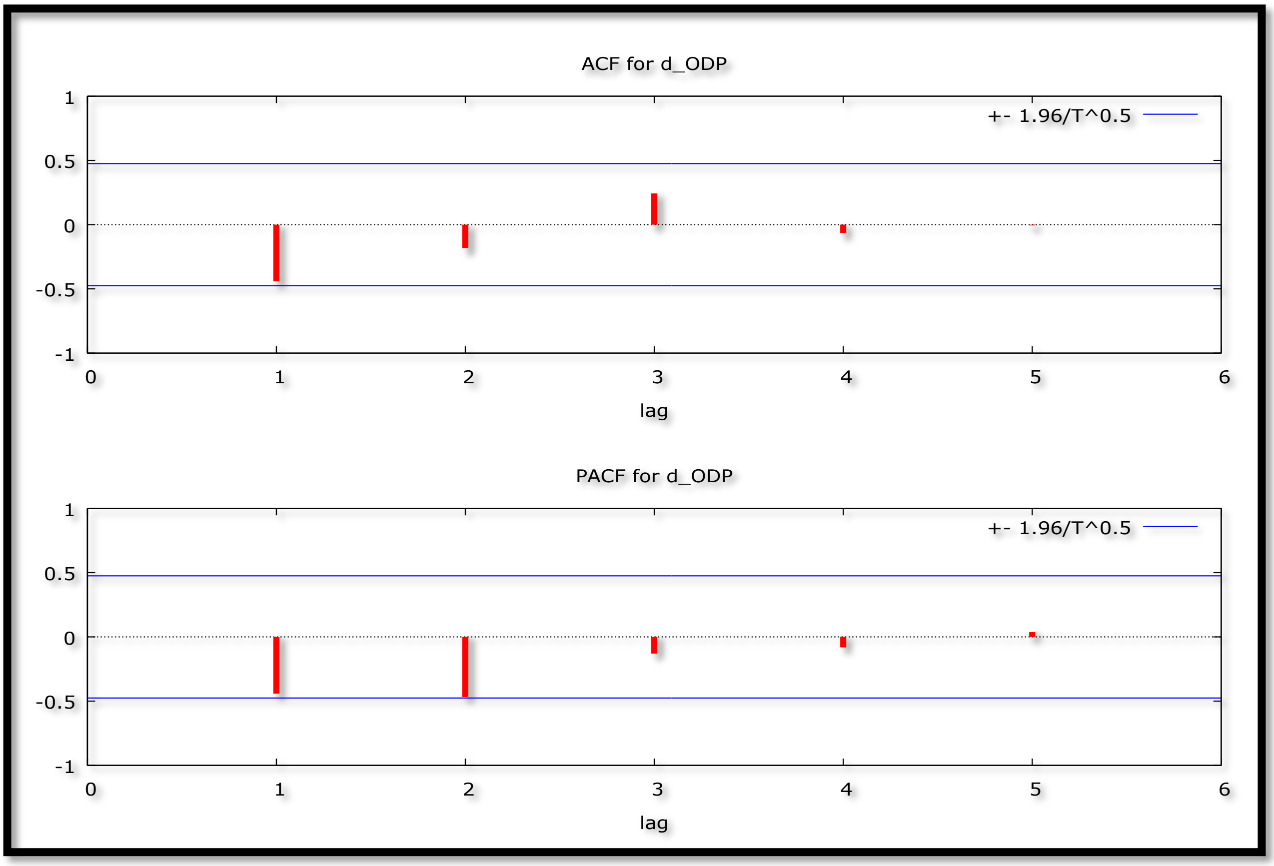

Figure 3: Correlogram (at First Differences)

3.7.5 The ADF Test (at First Differences):

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔODP	-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
ΔODP	-4.144856	0.0310	-4.886426	@1%	Non-stationary
			-3.828975	@5%	Stationary
			-3.362984	@10%	Stationary

Figure 3 as well as tables 3 and 4, indicate that ODP 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)	23.8276	0.14807	-0.0078946	0.33809	0.4075	0.86424
ARIMA (2, 1, 0)	**19.31151**	0.12582	-0.032076	0.23681	0.33842	0.59765
ARIMA (3, 1, 0)	20.94559	0.12654	-0.032123	0.24783	0.33492	0.64271
ARIMA (4, 1, 0)	22.80317	0.12374	-0.032454	0.2474	0.33353	0.62766
ARIMA (5, 1, 0)	23.96678	0.11859	-0.029257	0.24363	0.32494	0.62283

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, finally.

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	-3.882065	0.0108	-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	-4.435409	0.0150	-4.667883	@1%	Non-stationary
			-3.733200	@5%	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. Hence, the model is quite stable.

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

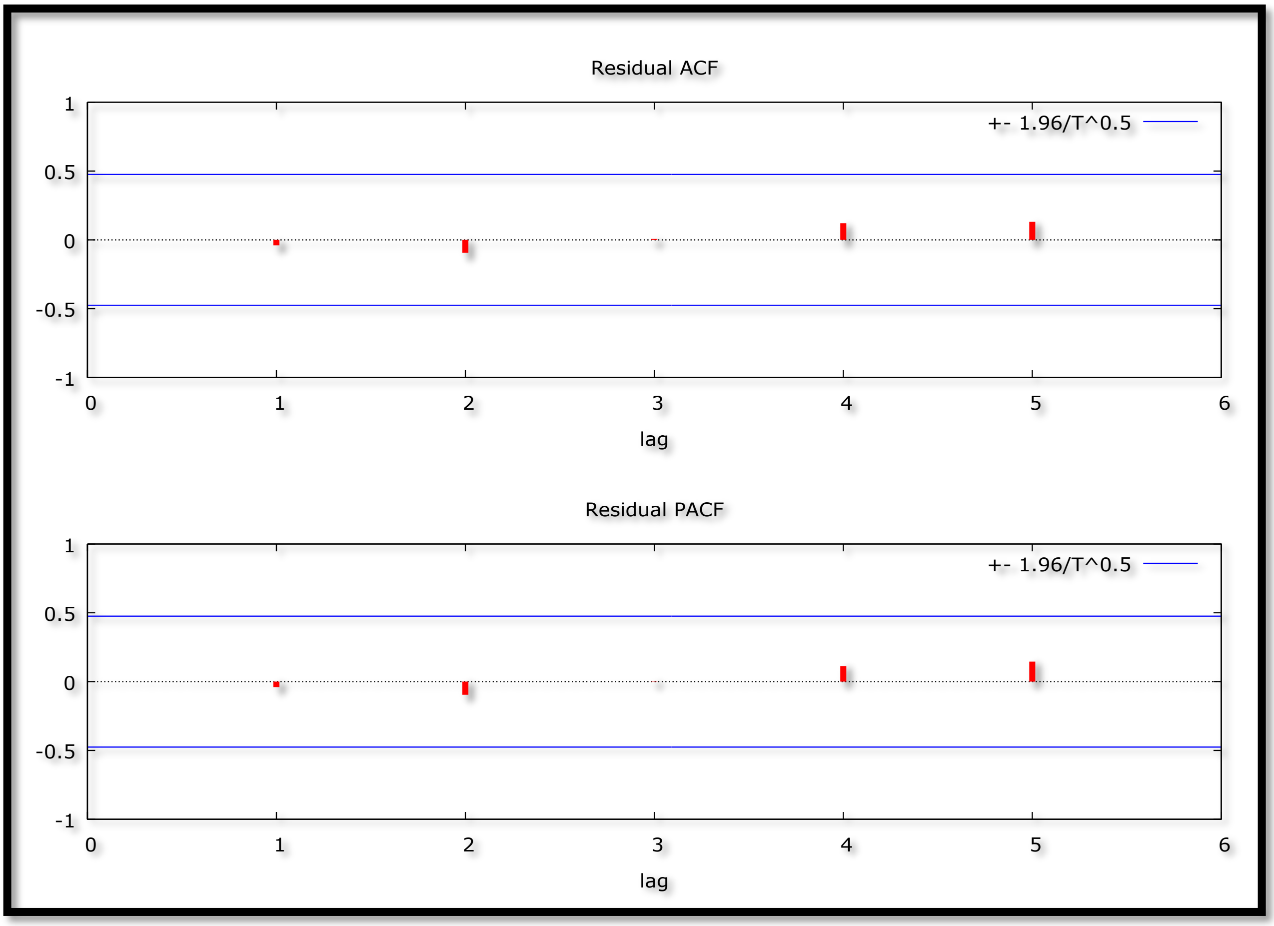

Figure 4: Correlogram of the Residuals

Figure 4 indicates that the estimated model is adequate 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:

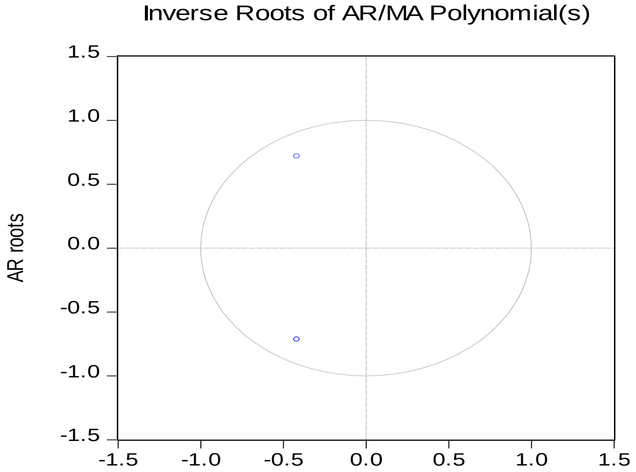

Figure 5: Inverse Roots

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

FINDINGS:

4.1 Descriptive Statistics:

Table 8: Descriptive Statistics

Description	Statistic
Mean	43.833
Median	43.5
Minimum	21
Maximum	67
Standard deviation	14.3
Skewness	0.32624
Excess kurtosis	-1.1971

As shown in table 8 above, the mean is positive, that is, 43.833. This shows that, over the study period, the annual average number of people practicing open defecation in Nepal is approximately 44% of the total population. This is a warning alarm for policy makers in Nepal with regards to the need to promote an open defecation free society. The minimum number of people practicing open defecation in Nepal over the study period is approximately 21% of the total population, while the maximum is 67% of the total population. In fact, the number of people practicing open defecation in Nepal has continued to decline over the years from 67% in 2000 to 21% of the total population in 2017. This is a desirable change and therefore, there is need to intensify policies and strategies that discourage the practice of open defecation in Nepal.

4.2 Results Presentation

Table 9: Main Results

ARIMA (2, 1, 0) Model:

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

$$\Delta ODP_t = -2.67411 - 0.771259\Delta ODP_{t-1} - 0.59481\Delta ODP_{t-2} \ldots\ldots\ldots\ldots\ldots\ldots [12]$$

Variable	Coefficient	Standard Error	z	p-value
constant	-2.67411	0.0349662	-76.48	0.0000***
β_1	-0.771259	0.205361	-3.756	0.0002***
β_2	-0.59481	0.2064	-2.882	0.0040***

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

Forecast Graph:

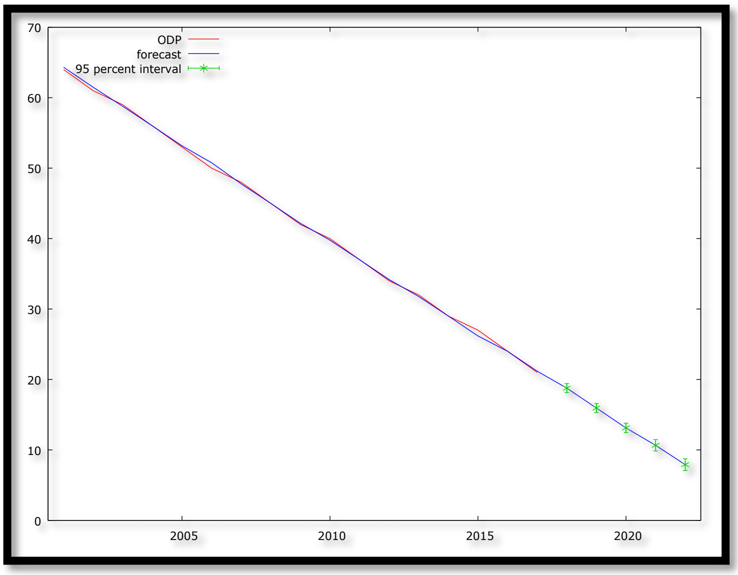

Figure 6: Forecast Graph – In & Out-of-Sample Forecasts

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

Predicted ODP – Out-of-Sample Forecasts Only:

Table 10: Predicted ODP

Year	Predicted ODP	Standard Error	Lower Limit	Upper Limit
2018	18.77	0.326	18.13	19.41
2019	15.95	0.335	15.29	16.6
2020	13.12	0.343	12.45	13.8
2021	10.65	0.41	9.85	11.46
2022	7.91	0.424	7.08	8.74

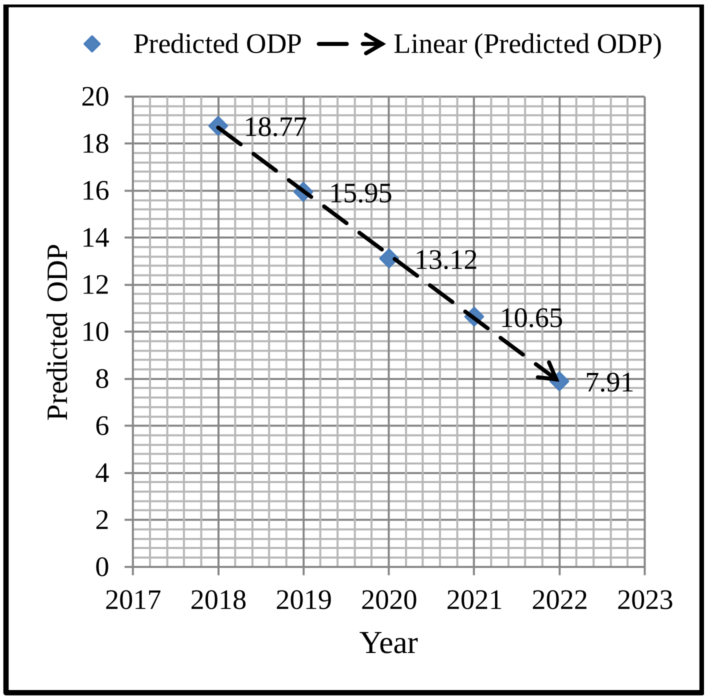

Figure 7: Graphical Analysis of Out-of-Sample Forecasts

Table 10 and figure 7 show the out-of-sample forecasts only. The number of people practicing open defecation in Nepal is projected to fall from approximately 18.77% in 2018 to as low as 7.91% of the total population by the year 2022. Indeed, it is possible to demystify Nepal's open defecation mystery. The results of this study reprove Bhatt et al. (2019) who generally implicate that open defecation in Nepal is persistant.

4.3 Policy Implications:

i. The government of Nepal should continue to make toilets a status symbol so that people consider toilets to be "rooms of happiness". Hence, the Open Defecation Free Movement must be intensified in Nepal.

ii. The government of Nepal should create more demand for sanitation through teaching the public on the importance of investing in toilets, particularly in light of disease transmission and other risks associated with open defecation.

iii. The government of Nepal should continue to encourage a habit of systematic hand-washing, not defecating in the open, as well as keeping toilets fly-proof.

CONCLUSION:

This research shows 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 Nepal over the period 2018 – 2022. The model predicts a sharp decrease in the annual number of people practicing open defecation in Nepal. Such a trend must be maintained. These findings are critical for the government of Nepal, especially when it comes to long-term planning with regards to materializing the much needed open defecation free society.

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) Bhatt, N., et al. (2019). What Motivates Open Defecation? A Qualitative Study From A Rural Setting in Nepal, PLoS ONE, 14 (7): 1 – 15.
3) Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis: Forecasting and Control, Holden Day, San Francisco.
4) Caruso, B. A., et al. (2017). Understanding and Defining Sanitation Insecurity: Women's Gendered Experiences of

Urination, Defecation and Menstruation in Rural Odisha, India, BMJ Global Health, 2 (4): 1 – 9.

5) Clair, R. P., et al. (2018). A Dialectical and Dialogical Approach to Health Policies and Programs: The Case of Open Defecation in India, Healthy Community, 11 (1): 1 – 11.
6) Kafle, S., & Pradhan, B. (2018). Situation of Water, Sanitation and Hygiene and Diarrhoeal Diseases After Open Defecation Free Declaration, Journal of Nepal Health Research Council, 16 (39): 160 – 164.
7) Nyoni, T (2018b). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, Dimorian Review, 5 (6): 16 – 40.
8) 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.
9) Nyoni, T. (2018c). Box – Jenkins ARIMA Approach to Predicting net FDI inflows in Zimbabwe, MPRA Paper No. 87737, University Library of Munich, Munich, Germany.
10) UNICEF (2009). Community Approaches to Total Sanitation, UNICEF, New York.