NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 5, ISSUE 8, Aug. -2019

PAEDIATRIC NEW HIV INFECTIONS IN TANZANIA: 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 children (ages 0 – 14) newly infected with HIV in Tanzania from 1990 – 2018, the study predicts the annual number of children who will be newly infected with HIV over the period 2019 – 2030. The paper applies the Box-Jenkins ARIMA methodology. The diagnostic ADF tests show that, N, the series under consideration is an I (2) variable. Based on the AIC, the research presents the ARIMA (0, 2, 1) model as the best model. The inverse root of the MA polynomial further reveals that the presented model is quite stable. Moreso, the residual correlogram of the model indicates that the residuals are not serially correlated; while the normality test shows that the residuals of the model are also normally distributed. This is overwhelming proof that the applied model stable, adequate and suitable for forecasting new pediatric HIV infections in Tanzania. The results of the study indicate that the number of new HIV infections in children in the country will continue to decline dramatically, over the out of sample period. In fact, our model shows that Tanzania could win the war against pediatric HIV in at least 6 years' time! Yes, it is possible to end the HIV scourge in children in Tanzania.

1.0 INTRODUCTION:

Human Immunodeficiency Virus (HIV) infection could be defined as the major epidemic of our century, with dramatic human, social, and economic implications. It is a chronic infection that is first characterized by an asymptomatic phase that can stay unchanged for years and subsequently, by the appearance of the first symptoms due to immunosuppression. In the end, it can lead to Acquired Immunodeficiency Syndrome (AIDS). The cause of the symptoms is to be found in the destructive effect of the HIV virus on T-helper lymphocytes, in which the virus completes its replication cycle (Lauritano et al., 2020). Paediatric HIV infection is a growing health challenge worldwide, with an estimated 1500 new infections every day (Prendergast et al., 2007). Globally, there are about 3.3 million children under the age of 15 years living with HIV. Of this number 88% live in Sub-Saharan Africa (Gyamfi et al., 2017). In 1990, new HIV infections in Tanzania were approximately 14000 as shown in figure 1 below. The new infections continued on an upwards trajectory until they reached an all-time high of 27000 new HIV infections for 3 consecutive years, that is, 1999 – 2001. Since then, new HIV infections in Tanzanian children have declined significantly up to as low as 8600 new infections by 2018. This is largely attributed to the introduction of the Prevention of Mother-To-Child Transmission (PMTCT) of HIV programme in 2003 in the country. The significant reduction in new pediatric HIV infections in the country, points to the efficacy of the PMTCT programme and other healthcare interventions in Tanzania.

Children of today are the youth of tomorrow. HIV affects this very precious generation and bear grave consequences to our future, our nation, the continent and the world

at large. It will adversely impact the health statistics, economic growth and above all the morale of nations (Kaushal & Upadhyay, 2013). The vast majority of children with HIV are infected during pregnancy (10-25%), child birth (35-40%) or breast-feeding (35-40%). Hence, HIV hits them when their immune system is still premature, naïve and developing. The timing of infection and the inability to raise a sufficient immune response, result in a persistently high viral load. Clinically, the consequences are multiple and devastating. Immature, developing organs are more susceptible to the demanding effects of the virus itself; common childhood infections are more frequent, severe and do not respond well to treatment and the presence of a chronic infection increases nutritional and metabolic demands leading to malnutrition, poor growth and development. Unlike in adults, the progression to advanced disease is rapid, leading to high early mortality.

Without appropriate intervention, over 50% of HIV infected children will die before they are two years of age. Amongst those infected during pregnancy and delivery, peak mortality rate is around 2 to 3 months of age. Early Infant Diagnosis (EID) and early Antiretroviral Treatment (ART) are crucial to improve the survival of young children infected with HIV. Efficacious interventions to Prevent Mother-To-Child HIV Transmission (PMTCT) are crucial and can reduce transmission rates to less than 2% (MSF, 2015). Actually, the most effective way to address paediatric HIV pandemic is through PMTCT (Prendergast et al., 2007). Unfortunately, the number of paediatric patients affected by HIV still remains high, mainly in developing countries, where the main cause of infection is vertical transmission from the mother (Lauritano et al., 2020; Evans et al., 2020). The main goal of this study is to predict the number of children newly infected with HIV in Tanzania over the period 2019 – 2030. This study will go a long way in investigating the possibility of ending the pediatric HIV scourge in the country.

2.0 LITERATURE REVIEW:

In a recent study in Asia, Siddiqui et al (2020) examined pediatric HIV in Pakistan based on a matched case-control study was done with 406 cases recruited. Conditional logistic regression was employed to assess the association of a priori defined risk factors with HIV infection. Moreso, global positioning system coordinates of participants' addresses were collected to analyze concordance between the genetic and spatial epidemiology. At the time of preparing this research paper, the final results of the study by Siddiqui et al. (2020) had not been released but their research is expected to provide information on the likely routes of infection and drivers of the HIV outbreak among children in Pakistan. In an African study, Ng'eno et al. (2016) estimated the prevalence of HIV infection among children aged 18 months to 14 years in Kenya. Data was collected and analyzed from the second Kenya AIDS Indicator Survey (KAIS – 2012). Moreso, blood specimens were collected for HIV serology and viral load measurement. The research showed that pediatric HIV care and treatment is still lagging behind in Kenya as than 50% of HIV infected children are still not getting proper care and treatment. More recently in another African study, Badru et al. (2020) investigated the factors associated with comprehensive HIV knowledge, stigma and HIV risk perceptions among youth adolescents aged 10-14 years in Akwa Ibom State, Nigeria. Cross-sectional data from the 2017 Akwa Ibom AIDS Indicator Survey was employed. A multiple logistic regression as used to examine relationships with the research outcomes. Results of the paper indicated that there was low comprehensive knowledge among young adolescents in Nigeria. Studies that forecast

new HIV infections in children in Tanzania are relatively scanty. This is the information hiatus that this paper seeks to address.

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, N, the series under consideration.

3.2 The Applied Box – Jenkins ARIMA Model Specification

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

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

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

3.3 Data Collection

This study is based on annual observations (that is, from 1990 – 2018) on the number of children newly infected with HIV (ages 0 – 14) [denoted as N] in Tanzania. Out-of-sample forecasts will cover the period 2019 – 2030. All the data was collected from the World Bank online database.

3.4 Diagnostic Tests & Model Evaluation

3.4.1 Stationarity Tests: Graphical Analysis

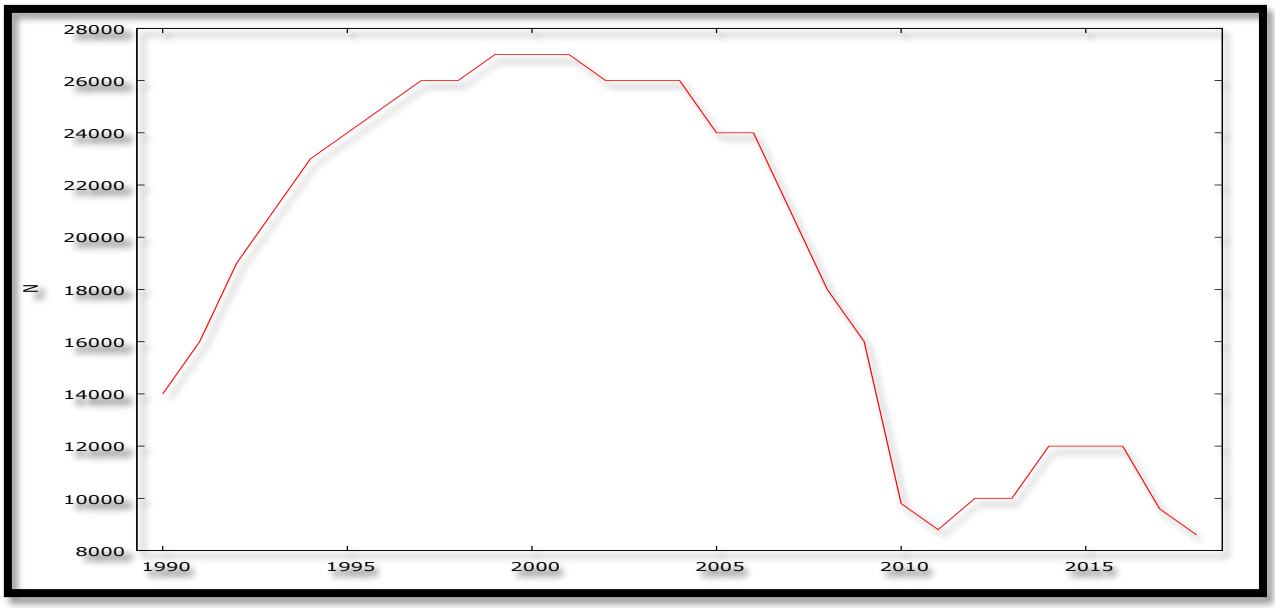

Figure 1

3.4.2 The Correlogram in Levels

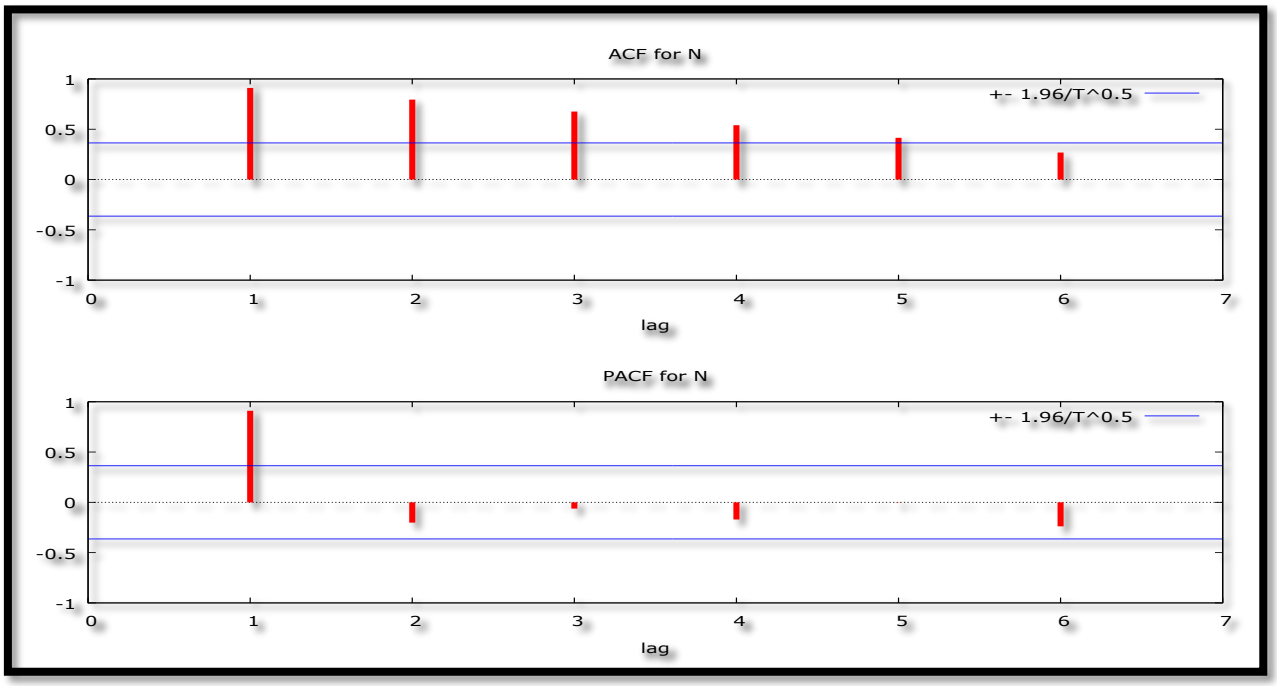

Figure 2: Correlogram in Levels

3.4.3 The ADF Test in Levels

Table 1: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
N	-0.742411	0.8191	-3.699871	@1%	Non-stationary
			-2.976263	@5%	Non-stationary
			-2.627420	@10%	Non-stationary

Table 1 shows that N is not stationary in levels; this has already been hinted in figures 1 and 2.

3.4.4 The Correlogram (at First Differences)

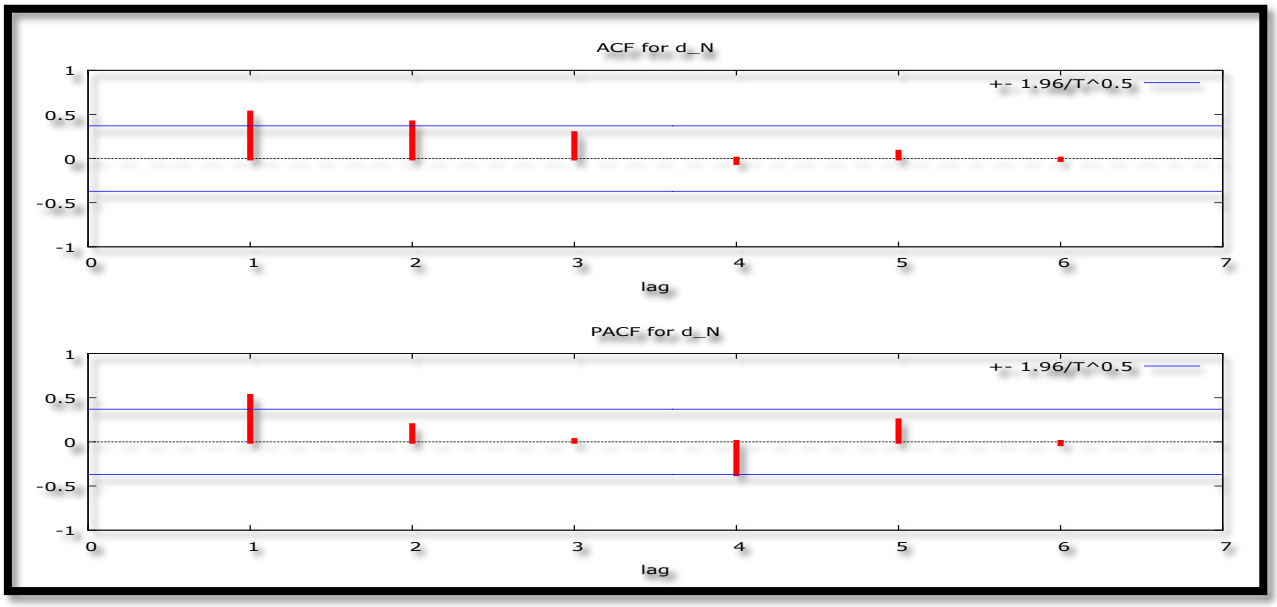

Figure 3: Correlogram (at First Differences)

3.4.5 The ADF Test (at First Differences)

Table 2: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
ΔN	-2.872537	0.0619	-3.699871	@1%	Non-stationary
			-2.976263	@5%	Non-stationary
			-2.627420	@10%	Stationary

Figure 3 shows that N is not yet stationary. However, table 2 indicates that the series under consideration is now stationary at 10% level of significance. It is ell known that a series should desirably be stationary at 1 – 5% level of significance, not 10%. Given, the inconsistency between the figure 3 and table 2, we proceed to test for stationarity after taking first differences.

3.4.6 The Correlogram (at Second Differences)

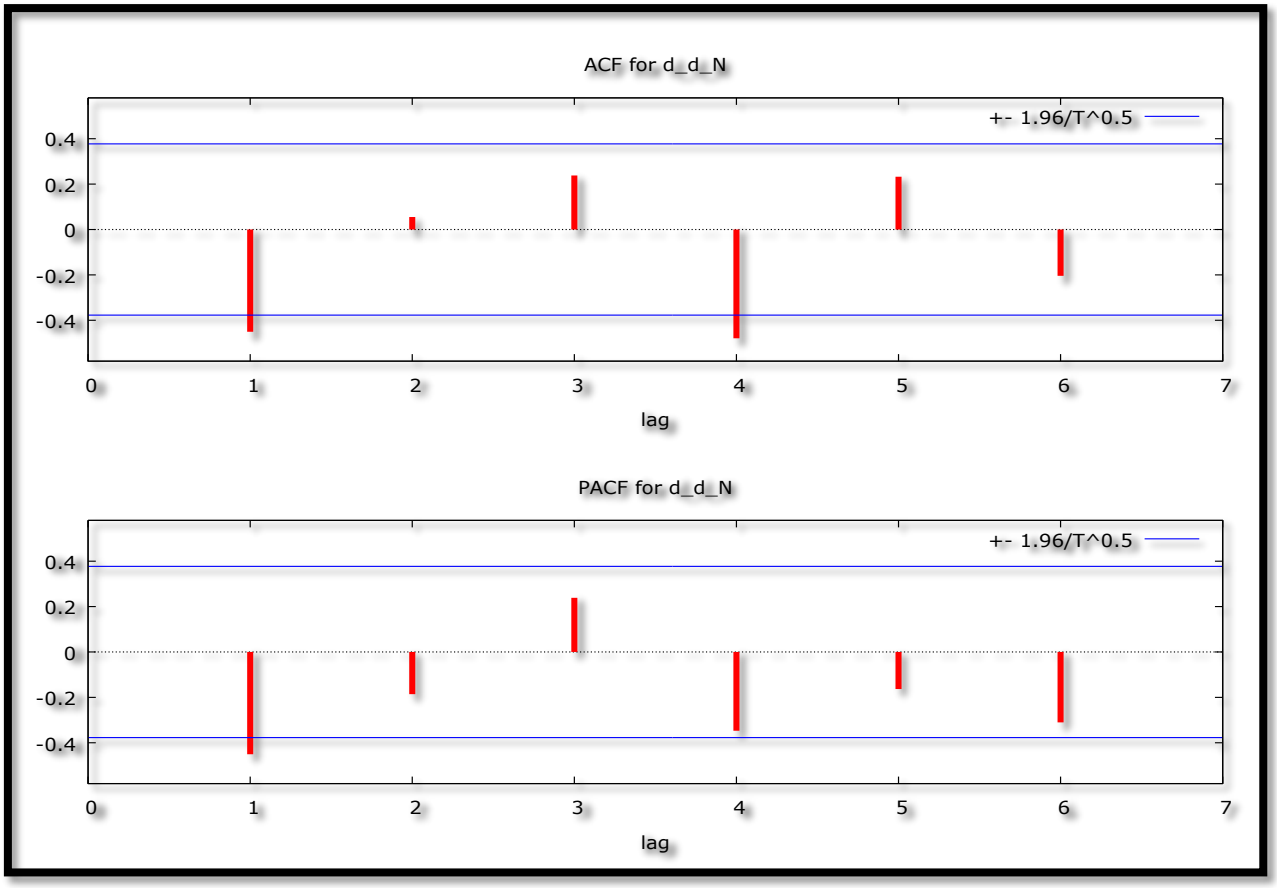

Figure 4: Correlogram (at Second Differences)

3.4.7 The ADF Test (at Second Differences)

Table 3: with intercept

Variable	ADF Statistic	Probability	Critical Values		Conclusion
Δ^2N	-8.026605	0.0000	-3.711457	@1%	Stationary
			-2.981038	@5%	Stationary
			-2.629906	@10%	Stationary

Figure 4 and table 3 indicate that N is an I (2) variable.

3.4.8 Evaluation of ARIMA models (with a constant)

Table 4: Evaluation of ARIMA Models (without a constant)

Model	AIC	ME	RMSE	MAPE
ARIMA (1, 2, 1)	482.7269	-200.4	1646.5	9.2234
ARIMA (1, 2, 0)	481.0678	-182.86	1657.3	9.3087
ARIMA (2, 2, 0)	482.4060	-207.78	1636.1	9.2833
ARIMA (3, 2, 0)	482.5092	-167.06	1574.8	8.9282
ARIMA (4, 2, 0)	481.3451	-172.41	1477.2	8.0763
ARIMA (5, 2, 0)	482.8650	-190.35	1462.4	8.0624
ARIMA (0, 2, 2)	482.2226	-137.96	1625.7	9.1564
ARIMA (0, 2, 1)	**481.0151**	-220.13	1654.8	9.2409

A model with a lower AIC value is better than the one with a higher AIC value (Nyoni, 2018b). In this research paper, only the AIC is used to select the optimal model. Therefore, the ARIMA (0, 2, 1) model is chosen at last.

3.5 Residual & Stability Tests

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

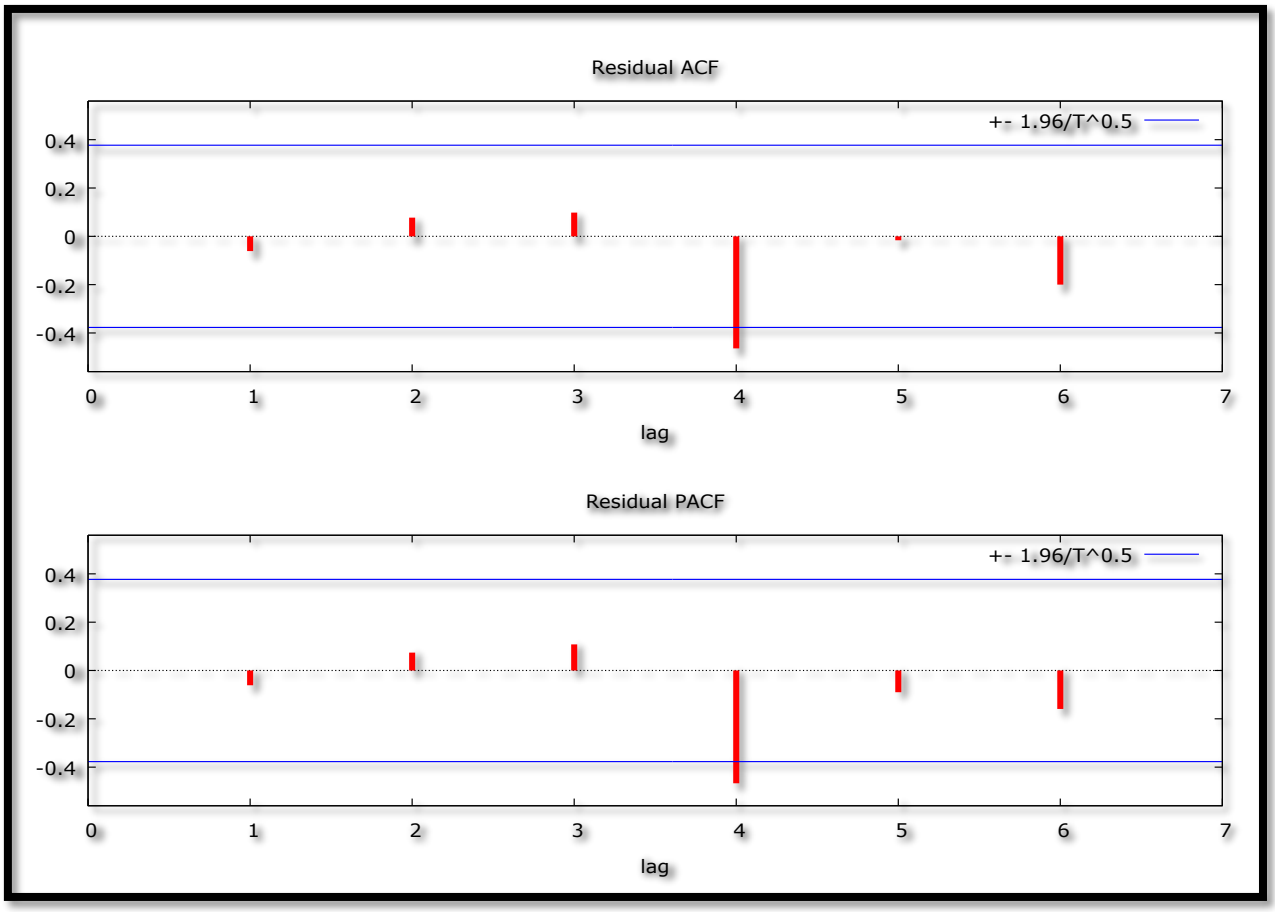

Figure 5: Correlogram of the Residuals

Figure 5 indicates that the estimated best model is adequate since most ACF and PACF lags are quite short and within the bands.

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

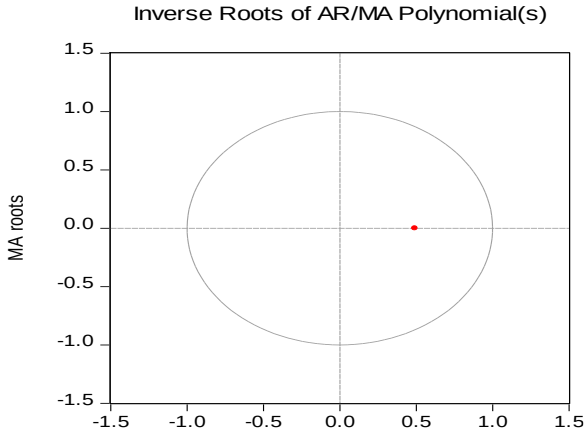

Figure 6: Inverse Roots

Since the MA root lies inside the unit circle, it implies that the estimated ARIMA process is (covariance) stationary; thus confirming that the ARIMA (0, 2, 1) model is stable and suitable for forecasting annual number of new HIV infections in children in Tanzania.

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

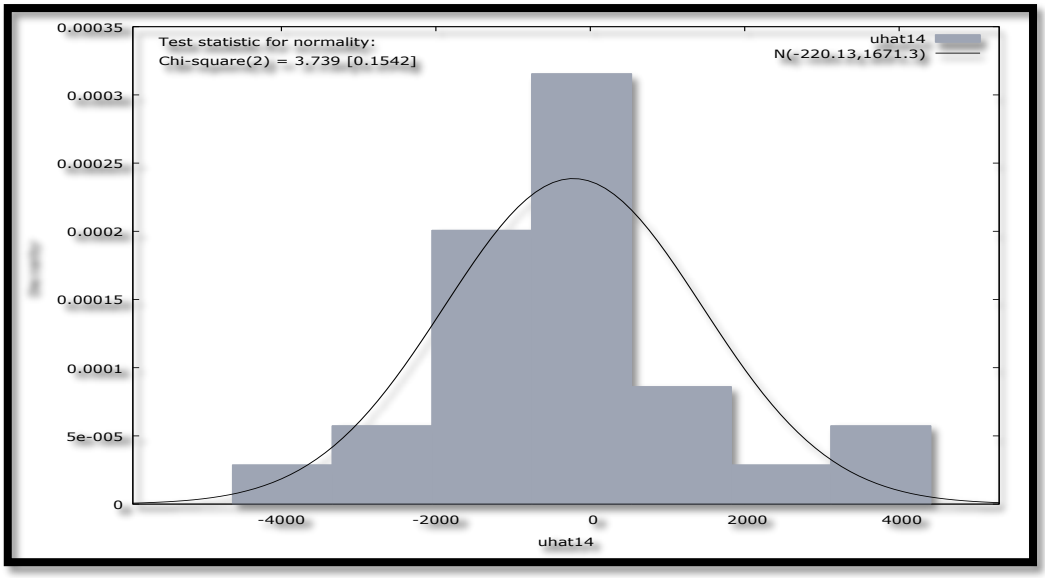

Figure 7: Normality Test

Because the probability value of the chi-square statistic is insignificant, we reject the null hypothesis and conclude that the residuals of the ARIMA (0, 2, 1) model are normally distributed.

4.0 FINDINGS OF THE STUDY:

4.1 Descriptive Statistics

Table 5: Descriptive Statistics

Description	Statistic
Mean	18924
Median	21000
Minimum	8600
Maximum	27000

Over the study period, the annual average number of new HIV infections in children in Tanzania is 18924. The minimum number of new HIV infections in children in the country is 8600 and has recently been recorded for 2018 while the maximum is 27000 infections and was observed for 3 consecutive years, that is; 1999 – 2001.

4.2 Results Presentation

Table 6: Main Results

ARIMA (0, 2, 1) Model:				
The chosen optimal model, the ARIMA (0, 2, 1) model can be expressed as follows: $\Delta^2 N_t = -0.469262\mu_{t-1} \ldots\ldots [2]$				
Variable	Coefficient	Standard Error	z	p-value
α_1	-0.469262	0.173492	-2.705	0.0068***

Table 6 shows the main results of the ARIMA (0, 2, 1) model.

Forecast Graph

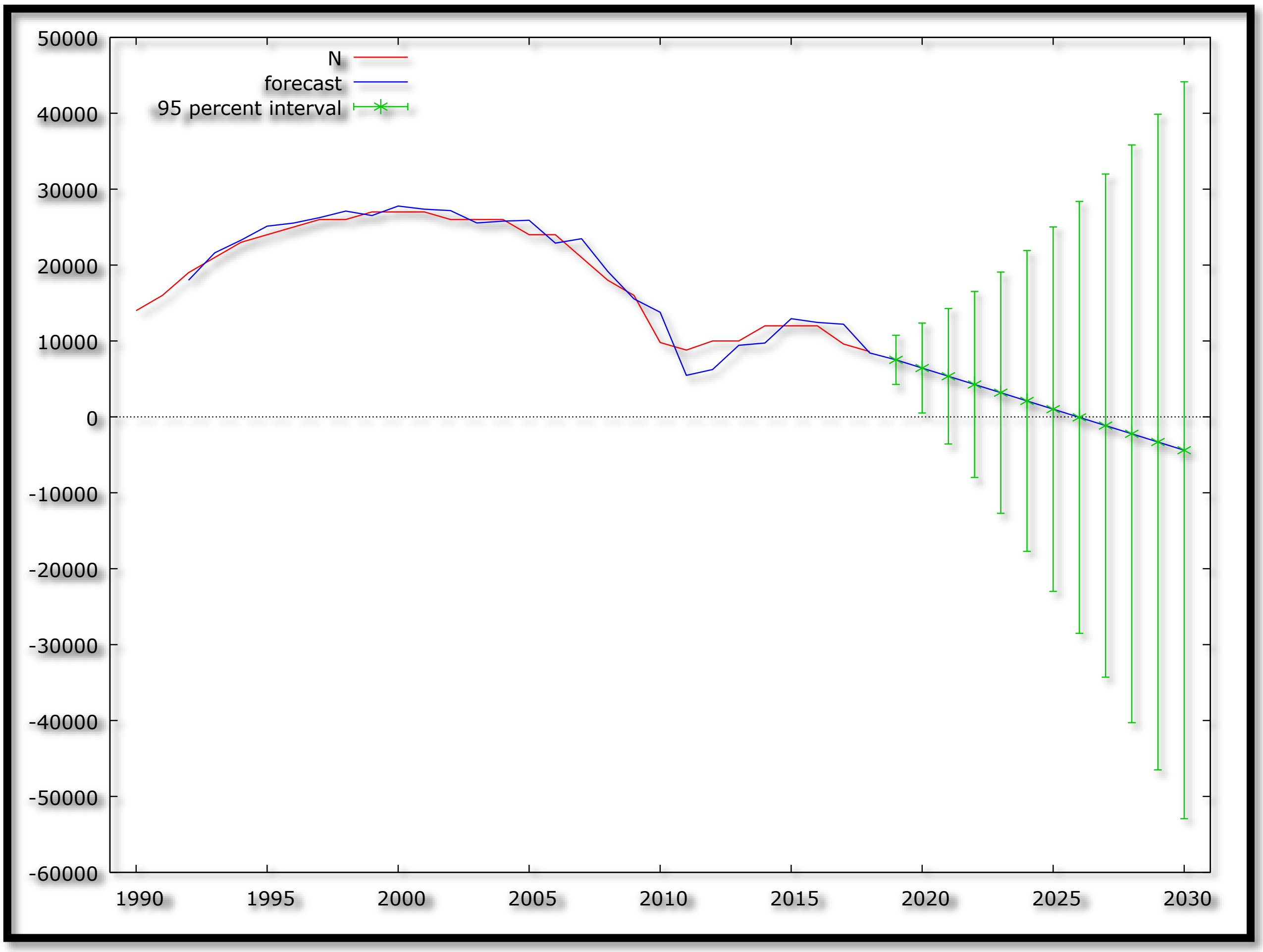

Figure 8: Forecast Graph – In & Out-of-Sample Forecasts

Figure 8 shows the in-and-out-of-sample forecasts of the N series. The out-of-sample forecasts cover the period 2019 – 2030.

Predicted N– Out-of-Sample Forecasts Only

Table 7: Predicted N (2019 – 2025)

Year	Predicted N	Standard Error	95% Confidence Interval
2019	7516.96	1652.54	(4278.05, 10755.9)
2020	6433.93	3021.55	(511.793, 12356.1)
2021	5350.89	4553.59	(-3573.98, 14275.8)
2022	4267.86	6251.84	(-7985.53, 16521.2)
2023	3184.82	8106.75	(-12704.1, 19073.8)
2024	2101.79	10108.2	(-17709.9, 21913.4)
2025	1018.75	12247.1	(-22985.1, 25022.6)

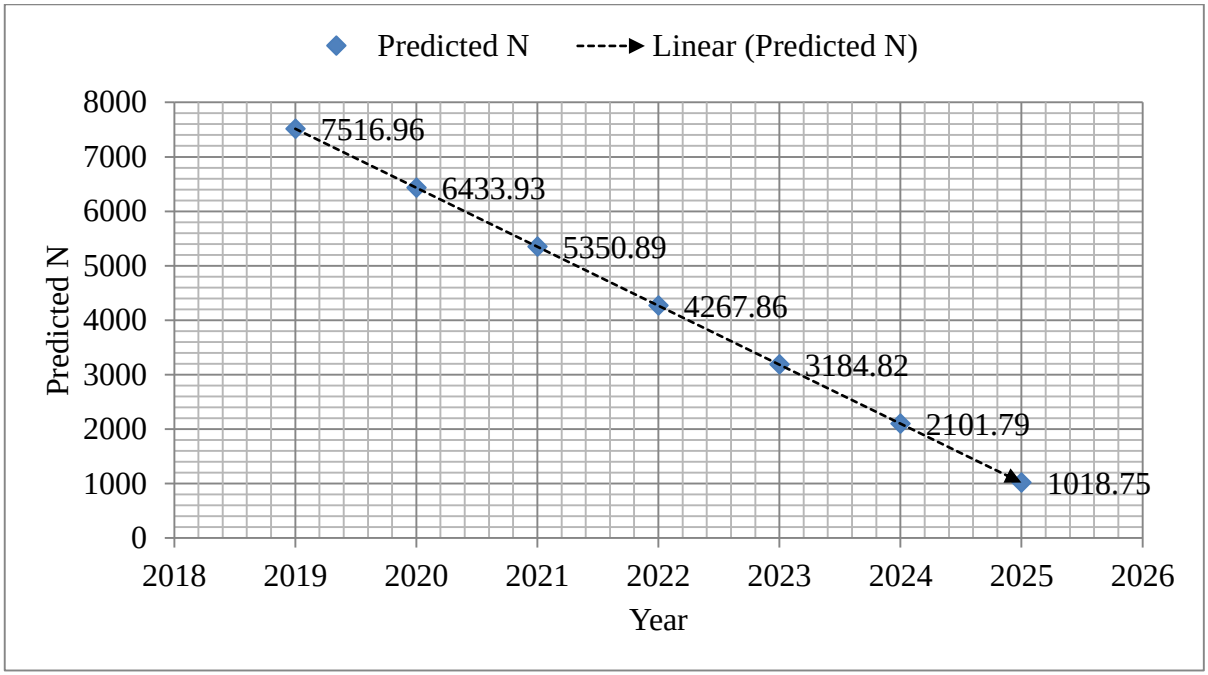

Figure 9: Graphical Analysis of Out-of-Sample Forecasts (2019 – 2025)

Table 7 and figure 9 show the out-of-sample forecasts only. The number of new pediatric HIV infections in Tanzania is projected to fall dramatically from an estimated 7517 new infections in 2019 to an amazing zero new infections by around 2026! Tanzania is indeed in the right direction and is likely to win the war against the paediatric HIV scourge in the near future.

5.0 CONCLUSION:

The research shows that the ARIMA (0, 2, 1) model is not only stable but also the most suitable model to forecast the annual number of new HIV infections in children in Tanzania over the period 2019 – 2030. The model predicts a dramatic decrease in the annual number of new HIV infections in children in Tanzania. The study recommends that the government of Tanzania should continue to intensify HIV prevention and control programmes in the country, especially the PMTCT programme.

REFERENCES:

1) Badru, T., et al. (2020). HIV Comprehensive Knowledge and Prevalence Among Young Adolescents in Nigeria: Evidence From Akwa Ibom AIDS Indicator Survey, 2017, BMC Public Health, 20 (45): 1 – 10.

2) Evans, C., et al. (2020). Mortality, Human Immunodeficiency Virus (HIV) Transmission, and Growth in Children Exposed to HIV in Rural Zimbabwe, Clinical Infectious Diseases, XX (XX): 1 – 9.
3) Gyamfi, E., et al. (2017). Prevalence of, and Barriers to the Disclosure of HIV Status to Infected Children and Adolescents in a District of Ghana, BMC International Health and Human Rights, 17 (8): 1 – 8.
4) Kaushal, A., & Upadhyay, Y. (2013). Pediatric HIV Infection, World Journal of Dentistry, 4 (1): 77 – 79.
5) Lauritano, D., et al. (2020). Oral Manifestations in HIV-Positive Children: A Systematic Review, Pathogens, 1 – 15.
6) MSF (2015). Paediatric HIV Handbook, MSF, Brussels.
7) Ng'eno, B., et al. (2016). Burden of HIV Infection Among Children Aged 18 Months to 14 Years in Kenya: Results From a Nationally Representative Population-based Cross-sectional Survey, Journal of Acquired Immunodeficiency Syndrome, 66 (1): 1 – 13.
8) Nyoni, T (2018b). Modeling and Forecasting Inflation in Kenya: Recent Insights from ARIMA and GARCH analysis, Dimorian Review, 5 (6): 16 – 40.
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) Prendergast, A., et al. (2007). International Perspectives, Progress and Future Challenges of Paediatric HIV Infection, Lancet, 370: 68 – 80.
11) Siddiqui, A. R., et al. (2020). Investigation of an Extensive Outbreak of HIV Infection Among Children in Sindh, Pakistan: Protocol for a Matched Case-Control Study, BMJ Open, 10: 1 – 8.