Determining Population Projections Using Leslie, Lee Carter and ARIMA Models: A Case Study of Rwandan Population Dynamics

Dieudonne HATEGEKIMANA[1], Dr. Joseph K.Mung'atu[2], Dr. Marcel NDENGO[3]

[1,2,3] Jomo Kenyatta University of Agriculture and Technology

Abstract: **In 2012, National Institute of statistics of Rwanda conducted the 4th population and housing census whereby the population projection was done about fertility and mortality .Based on the assessment of past mortality, it is assumed that life expectancy at birth in Rwanda will increase linearly up to 70 years with of course a significant decrease in infant mortality and under five mortality rates by the end of the projection period 2032.Therefore,the forecasts results should be inaccurate as long as the forecasting period becomes too big. Objective of this study was to apply Leslie matrix, Lee carter and ARIMA models in forecasting population dynamics in Rwanda. Leslie matrix model was used to determine the annual growth rate of female population which was r =2.33% .By ARIMA models, Forecasts showed a decreasing trend in IMR and U5MR from 2012 to 2015. The results showed that models used had a higher predictive ability because of their small root mean square errors where RMSE was found to be 0.124456 for IMR and 0.068576 for under 5 mortality rates (‰).**

Keywords: **Leslie matrix, Infant Mortality Rates, Under 5 Mortality Rates, Lee-Carter model, ARIMA models, modeling and Forecasting.**

1. INTRODUCTION

Leslie matrix models are useful tools in population dynamics to predict population growth. They are particularly suitable to describe the evolution of age structured population. These models can even be used to determine the evolution of human age pyramids (Gonze, 2015).

This study aimed at applying Leslie, Lee carter and ARIMA models for projection of the population dynamics in Rwanda where the annual growth rate was determined by using Leslie matrix model and mortality rates were forecasted by ARIMA models.

However, Population projections are inevitable tool for decision makers and planners. The government ministries, particularly health, education, transport, environment, social welfare and housing, constantly seek projections of future demographic parameters for planning purposes and resource allocation.

The Lee and Carter model, also named LC herein after, is a demographic and statistical model that is used to project mortality rates. The equation describing the model is expressed as ln $(m_{x,t})=\alpha_x + \beta_x\ k_t + \varepsilon_{x,t}$, **(1.1)**

In the equation (1.1) x is the age and t is the time, α_x: Average age-specific pattern of mortality,β_x: Pattern of deviations from the age of profile as the k_t varies, k_x : A time-trend index of general mortality level and finally ε_t is the residual term at age x and time t which has zero mean and constant variance that is E(ε_t)=0 and var (ε_t)=σ (Jenny, 2007).

In forecasting by using ARIMA models, The input series for an ARIMA needs to be stationary, that is, it should have a constant mean, variance, and autocorrelation through time. Therefore, the series usually needs to be differenced first until

it is stationary. The number of times the series needs to be differenced to achieve stationary is reflected in the d parameter. In order to determine the necessary level of differencing, one should examine the plot of the data and autocorrelogram, that displays graphically and numerically the autocorrelation function (Haberman, 2015).

2. LITERATURE RIVIEW

2.1 Introduction:

Life expectancy at birth has increased significantly in the least developed countries in recent years. The six-year average gain in life expectancy among the poorest countries, from 56 years in 2000-2005 to 62 years in 2010-2015, is roughly double the increase recorded in the rest of the world. (UN, 2015).

Infant mortality rates in Rwanda were estimated at 92‰ in 1990 and reduced to 39‰ in 2012.The number of infant deaths was 30000 in 1990 and this number has reduced to 17000 in 2012.Under five mortality rate was estimated at 151‰ in 1990 ,182‰in 2000 and reduced to 55‰ in 2012.in addition, under 5 deaths were estimated at 49,000 in 1990 and 24,000 in 2012.(World Bank, United Nations &World Health Organization,2013).

National Institute of Statistics of Rwanda is basing on the assumptions regarding fertility and mortality as components of population dynamics whereby three possible scenarios have been derived to project the Rwandan population from 2012 to 2032. According to the high scenario, TFR would decrease from 4.0 children per woman in 2012 to 3.5 at the end of the period of projections, while the life expectancy at birth would increase from 62.6 years in 2012 for men and 66.2 years for women to 69.1 years for men and 72.3 years for women in 2032.

According to the medium scenario, TFR would decrease from 4.0 children per woman in 2012 to 3.0 at the end of the period of projections, while the life expectancy at birth would increase from 62.6 years in 2012 for men and 66.2 years for women to 69.3 years for men and 73.4 years for women in 2032.Low scenario states that TFR would decrease from 4.0 children per woman in 2012 to 3.5 in 2020, and to 2.5 by the end of the projection period, while the life expectancy at birth would increase from 62.6 years in 2012 for men and 66.2 for women to 70.6 years for men and 74.5 years for women in 2032. (NISR,2012)

2.2 Evolution of Rwandan population   from 1978 to 2012:

The First national census was conducted in 1978, the second in 1991, the third in 2002 and then the fourth in 2012.

Rwanda is the most densely populated country in Africa, with about 10,515,973 Population of which 5,064,868 male and 5,451,105 females. Data shows the sex ratio of 93 male per 100 females .The population density is then estimated at 415 population /km^2 (NISR,2012).

Table 2.1: Population size and population growth

Period	Population size	Annual average growth rate
1978	4,831,527	3.1% from 1978 to 1991
1991	7,157,551	1.2 from 1991 to 2002
2002	8,128,553	2.6 % from 2002 to now
2012	10,515,973	2.6%

Source: Secondary data (NISR, 2012)

Table 2.1 shows a decrease of annual average growth rate of the population from 1978 to 2012
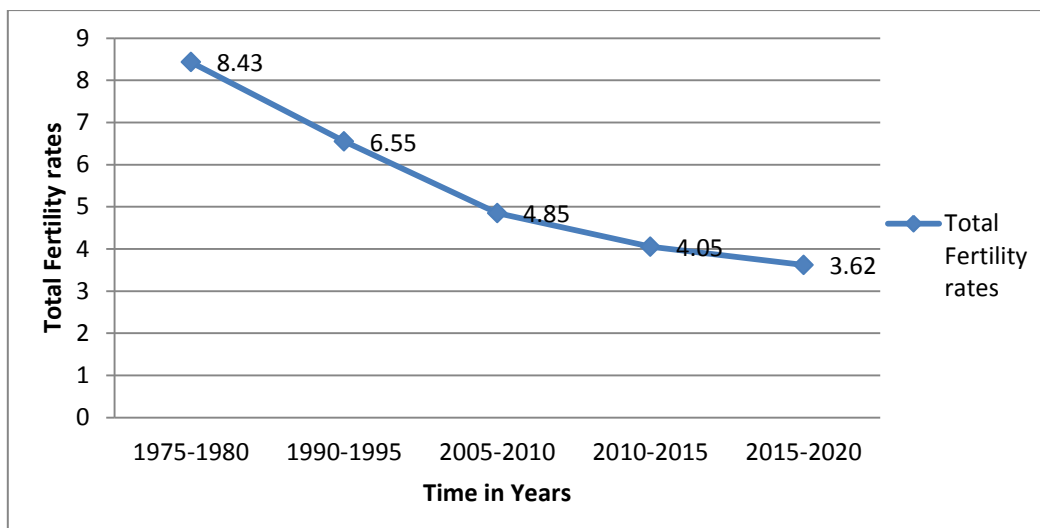
That rate decreased from 3.1 % in 1978 to 1.2% in 1991 and increased to 2.6% in 2012.
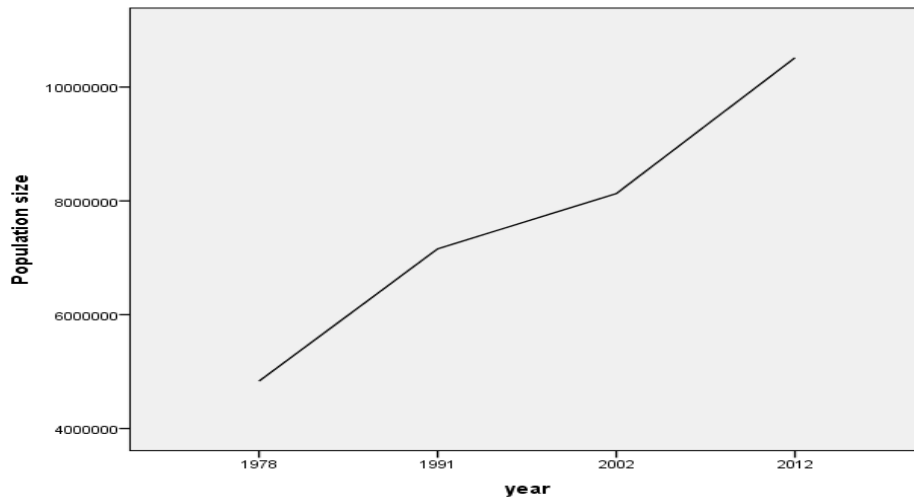
Source: (NISR, 2012)

Figure 2.1: Trend in population size

The figure 2.1 shows an increase in total population from 4831527 in 1978 up to 10515973 in 2012.

Figure 1.2: Trend in Total Fertility rates

Figure 2.2 shows a decreasing trend in Total Fertility rates where it is 8.4 in 1975 and 3.6 in 2020. This shows that TFR is projected at 3.6 children per woman in 2020(UN ,2015)

2.3 Measurement of childhood mortality:

Infant and child mortality rates are basic indicators of a country's socioeconomic situation and quality of life. Estimates of childhood mortality are based on information collected in the birth history. The rates are estimated directly from the information in the birth history on a child's birth date, survivorship status, and age at death for children who died; and are expressed per 1,000 live births. This information is used to directly estimate mortality rates (NISR, 2015).

Table 2.2 : Infant mortality and under five mortality in Rwanda

Period	Infant mortality rate	Under 5 mortality rate
1990-1995	289	466
2005 -2010	59	90
2010-2015	49	73
2015-2020	43	61

Table 2.2 shows that From 1990-1995: infant mortality was 289 and under 5 mortality 466, 2005-2010:Infant mortality rate was 59 and under 5 mortality 90 ,2010-2015 :IMR was 49 and U5MR 73,2015-2020:IMR is estimated at 43 and U5MR at 61(UN,2015)

Infant mortality is still high in Rwanda. The Infant Mortality Rate (IMR) in 2012 is 48.6 ‰ and is higher among boys (53‰) than girls (44%). However the IMR has decreased a lot and more quickly during the last decade: from 139‰ in 2002 to 48.6‰ in 2012. The decrease is more important among girls (67%) than among boys (63%).

2.4. Overview of Lee Carter method in mortality forecasting:

The basic LC model of age-specific death rates (ASDRs, and denoted m x,t is: $\ln(m_{x,t}) = \alpha_x + \beta_x\, k_t + \varepsilon_{x,t}$, **(2.1)**

Here α_x describes the general age shape of the ASDRs, while k_t is an index of the general level of mortality. The coefficients β_x describe the tendency of mortality at age x to change when the general level of mortality k_t changes. When β_x is large for some x, then the death rate at age x varies a lot when the general level of mortality changes and when β_x is small, then the death rate at that age varies little when the general level of mortality changes .

Over the past ten years, a number of new approaches have been developed for forecasting mortality using stochastic models such as Lee and Carter model. Recently, the Lee-Carter model became more and more popular and was applied for long-run forecasts of age specific mortality rates from many countries and time periods. This model is computationally simple to apply and it has given successful results for various countries, for instance USA, Canada , Japan and Italy. (Jenny, 2007).

2.5 Leslie matrix models:

The basic Leslie matrix model is deterministic. The parameters as survival   and reproductive rates are constant over time.Those parameters may change with many factors. For human population, the survival rates, especially among the elderly, will change with new medical advances or better feeding habits, and the fecundity rates will be affected by changing social attitudes toward marriage and family.If the parameter values for years are randomly generated from a specific probability distribution, the model will be stochastic.The size of the current population will play an important role in determining those parameters and the initial size has a lasting effect on the population's future chances of reproducing and surviving.

2.6. Life expectancy:

Life expectancy is an important parameter in determining the size of a population on account of a given birth rate and the number of people is proportional to it.

The symbol T_x in a life table denotes the total number of years lived beyond agex by the survivorship group with l0 initial members. We have $T_x = \int_0^\infty l_{x+t}\, dt = L_x + L_{x+1} + L_{x+2} +$. .**(2.2)**

Let e_x^0 denote the life expectancy of (x), i.e. the average number of years or the future life time lived by (x). Therefore, life expectancy e_x^0 can be calculated as $e_x^0 = \frac{T_x}{l_x}$ **(2.3)**

Life expectancy is an important indicator for the mortality level of a population. It has been widely used to measure overall mortality changes in a region or to compare mortality differences between cohorts. (Ofosuhene ,2009)

2.7 Survival distributions:

Let X be a nonnegative random variable representing the lifetime of an individual in a cohort or Population.

All distribution functions related to the random variable X, unless stated otherwise, are defined over the interval [0, ∞). Let f(x) denote the probability density function (p.d.f) of X and let the cumulative distribution function (CDF) be $F(x) = P(X \le x) = \int_0^x f(t)dt$ **(2.4)**

The probability of an individual surviving to age x is given by the survival function (s. f)

$$s(x) = P(X > x) = \int_x^\infty f(t)dt$$ **(2.5)**

A very important concept in mortality modeling is the force of mortality (often referred to as the hazard function which is defined as:

$$H(x) = \lim_{\Delta x \to 0} \frac{P(x < X \le x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{S(x)} \qquad (2.6)$$

The force of mortality specifies the instantaneous rate of death at age x, given that the individual survives up to age x.

Any one of the functions f(x), F(x), s(x), or H(x) can be used to specify the distribution of X. It is easy to see that, given an expression for any one of the above

Four functions, the other three can be derived. For example, in terms of the force of Mortality H(x), we have: $s(x) = e^{-\int_0^x H(t)dt}$ and F(x) = 1-S(x) $\qquad (2.7)$

2.8 Effects of private income levels on mortality:

Richer countries not only have richer people but in general ,they have larger and more effective social economic programs. Some indication of the importance of private living standards for international mortality differences may be gained by examining the importance of income distribution as a factor in those differences (Samuel.H,1980).

It is reasonable to expect that mortality also responds nonlinearly to individual income levels, in which case the distribution of income within a nation should influence its aggregate level of mortality. In particular, suppose that the relation between individual income and life expectancy is log linear with respect to the following equation:

$$e_i^0 = \alpha + \beta \ln Y_i \qquad (2.8)$$

In the equation (2.8) e_i^0 is the life expectancy at birth in income group i, Y_i stands for the level of income received by group i and then α and β are constants (Samuel .H,1980).

The suggestion that private incomes are very crucial in determining national levels of life expectancy at a moment in time does not imply that changes in private incomes have been the dominant factor in mortality changes. Therefore , before trying to establish the role played by changes in private living standards in Less Developed Countries mortality declines, it is useful to make an assessment of the causes of death responsible for those declines(Richard.A,1980).

3. METHODOLOGY

3.1. Introduction:

The Leslie matrix population model is a discrete and age dependent model. This matrix population model is widely used in demography in order to determine the growth of a population, as well as the age distribution within the population over time, where age class i corresponds to ages i-1≤ x ≤i.

The Leslie model describes the dynamics of an age-structured population and is based on 3 main elements:

1. The age x is a continuous variable starting from 0 and subdivided into discrete age classes, from 0 to w: the age class i thus corresponds to individuals whose age satisfy $i-1 \le x < i$ for $i = 1,2,\dots\dots\dots w$

2.Time is a discrete variable. We will denote by t the time step (also called the projection interval.

3.The time step is exactly equal to the duration of each age class, meaning that from t to t+1 all individual go from class I to class i+1.

The female population can be divided into several categories by age or by size in the simplest Leslie matrix model. If it's grouped by age, the group intervals are supposed to have equal length of time.

We assume that the survival and fecundity rates of each category are constant over time and therefore not dependent on population density.(Gonze Didier ,2015)

The Leslie matrix model is expressed as :

$$A = \begin{bmatrix} F_1 & F_2 & F_3 \dots \dots F_{i-1} \cdots & F_i \\ P_1 & \ddots & & \vdots \\ \dots 0 & \cdots & & 0 \end{bmatrix} \qquad (3.1)$$

In the equation (3.1) A represents the n × n Leslie matrix, or projection matrix. F_i are fertility rates with i=1,…,n . Leslie matrix is thus characterized by a first line with the fecundity of the different classes and a sub-diagonal with contains the

survival probabilities from one class to the next one. A Leslie matrix is non-negative, i.e. if $m_{i,j}$ denote the elements (i, j) of M, then $m_{i,j} \geq 0 \; \forall_{i,j}$ (Gonze Didier,2012)

3.2 Parameter estimation:

Let us start by examining the survival probability, designated by the letter *P*. *P* is the probability that an individual in age class *i* will survive to age class *i + 1*. The small letter *l* gives

the number of individuals in the population at a given time: $P_{i} = \frac{l_i}{l_{i-1}}$ **(3.2)**

3.3. Dominating eigenvalues and the properties of a stable vector:

Since the Leslie model is an nxn matrix, it can be concluded that there are n possible eigenvalues and eigenvectors which satisfy the equation: $AV = \lambda V$ **(3.3)**

In the equation (3.3) λ is any eigenvalue and v is an eigenvector corresponding to λ. Eigenvalues and eigenvectors are usually used to study the change in a population over time in a dynamical system. The aim is to determine the long term dynamics of the population. That is to say to demonstrate whether the population is increasing, decreasing or staying constant.After computing the eigenvalues from a projected matrix using the analytical method, the eigenvalue of interest is the one which is more positive in comparison with the others.

This eigenvalue is called the eigenvalue of greatest magnitude, or the dominating eigenvalue or a latent root.

The parameter $\boldsymbol{\lambda}$ is very important because it defines the rate of population growth; the significance of the dominant eigenvalue is supported by the Peron Frobenius theorem for non-negative and irreducible matrices whose properties are below:

1. There exists one eigenvalue that is greater than or equal to any of the others in magnitude, called the dominant eigenvalue of A,

2. There exists an eigenvector such that its element are non-negative,

3. The parameter $\boldsymbol{\lambda}$ is greater or equals to the smallest row sum of A and less or equals to the largest row sum.

The above properties, especially the last one, does not always satisfy the Markov theory because there is a possibility that the values of F_i of the Leslie matrix may sometimes sum up to a value greater than 1 which can't happen in Markov theory .After obtaining Eigen values from the equation:$|A - \boldsymbol{\lambda}\,I| = 0$ **(3.4)**

In the equation **(3.4)** λ is an eigen value of Leslie matrix and I is the identity matrix.

If$\lambda \leq 1$, the population is decreasing   and $\lambda \geq 1$ the population is increasing.

If $\lambda=1$, the population is said to be stationary.

The annual rate of increase of the population will be given by the logarithm of the dominant eigen value , $r = \ln(\lambda)$.

3.4 ARIMA model in forecasting mortality rates:

Given a time series of data X_t, the ARMA model is a tool for understanding and   predicting future values in series. The model consists of two parts, an autoregressive (AR) part and a moving average (MA) part as explained. The model is usually then referred to as the ARMA (p,q) model where p is the order of the autoregressive part and q is the order of the moving average part. In order that the ARMA model functions properly the roots of the AR (p) should be stationary.

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$ **(3.5)**

The equation (3.5) shows an ARMA (p, q) model where p refers to the model with p autoregressive terms and q moving average terms. This model contains the AR (p) and MA(q) models. ARMA models in general can, after choosing p and q, are fitted by least squares regression to find the values of the parameters which minimize the error term. It is generally considered good practice to find the smallest values of p and q which provide an acceptable fit to the data (Shahzad, 2007)

These models are fitted to time series data either to better understand the data or to predict future points in the series. The model is generally referred to as an ARIMA (p,d,q) model where p, d, and q are integers greater than or equal to zero and refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. Since we already

discussed the AR and MA part in precise, the integrated aspect of the data leads to differencing of the series in order to achieve a subsequent stationary series fit ARMA process and together they are mentioned as ARIMA.

The input series for an ARIMA needs to be stationary, that is, it should have a constant mean, variance, and autocorrelation through time.Therefore, the series usually needs to be differenced first until it is stationary. The number of times the series needs to be differenced to achieve stationarity is reflected in the d parameter.

In order to determine the level of differentiating, we examined the plot of data and autocorrelogram that displays graphically and numerically the autocorrelation function (Haberman & Russolillo,2005).

Data to be used in this study was captured from the 4th Rwandan population and housing census and demographic Health surveys conducted by National Institute of Statistics. That is to say, secondary data were used.

4. RESULTS AND DISCUSSION

4.1 Introduction:

This chapter presents the analysis and the discussion of the results obtained from the study. In this study, age specific fertility rates of females aged 15 to 49 and survival probabilities from life tables were used to construct Leslie matrix model for the purpose of estimating the annual growth rate which in turns may be used in population projection. The analysis was carried out using SPSS and E-views 4.0 statistical software. Here E-views 4.0 was used to plot various curves and analyze descriptive statistics, estimation, and forecasting of infant mortality and under 5 mortality rates taking observations from 2010 to 2011 and making forecasts from 2012 to 2015.

Table 4.1.Age Specific Fertility rates

5-year age-group	Age-Specific Fertility Rates ‰	Contribution (%) to the general fertility
15-19	27	3.4
20-24	150	18.7
25-29	202	25.1
30-34	185	23
35-39	142	17.7
40-44	79	9.8
45-49	19	2.4

The table 4.1: is showing age specific fertility rates and its contribution to the general fertility .it is 0.027‰ for female aged 15-19**,**0.150‰ for female aged 20-24 years ,0.202‰ for female aged 25-29years ,0.185‰ for female aged 30-34 years ,0.142‰ for female aged 35-39,0.079‰ for female aged 40-44 years ,0.19‰ for female for female aged 45 -49 years .

Total fertility is computed as TFR=$\sum_{x=15}^{49} ASFR(x)$ Hence from the above table,

TFR=5(0.027+0.15+0.202+0.185+0.142+0.079+0.019) = 5*804=4.02 **.**It means that total fertility in Rwanda is 4 children per woman.

4.2 Leslie matrix model:

$$A=\begin{pmatrix} 0.027 & 0.15 & 0.202 & 0.185 & 0.142 & 0.079 & 0.019 \\ 0.991 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.989 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.987 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.985 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.983 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.98 & 0.973 \end{pmatrix}$$

(4.1)

The equation (4.1) is showing Leslie matrix model where the first row shows ASFR for females.

Others non zero values are the survival probabilities and so ,the dominant eigen value for Leslie matrix was found $\lambda=1.0236$ .The results showed that female population is increasing at growth rate is r= ln λ ,That is r =2.33%.

This rate may be used to forecast the number of female population by Malthusian growth model where $P_t = P_i\, e^{rt}$
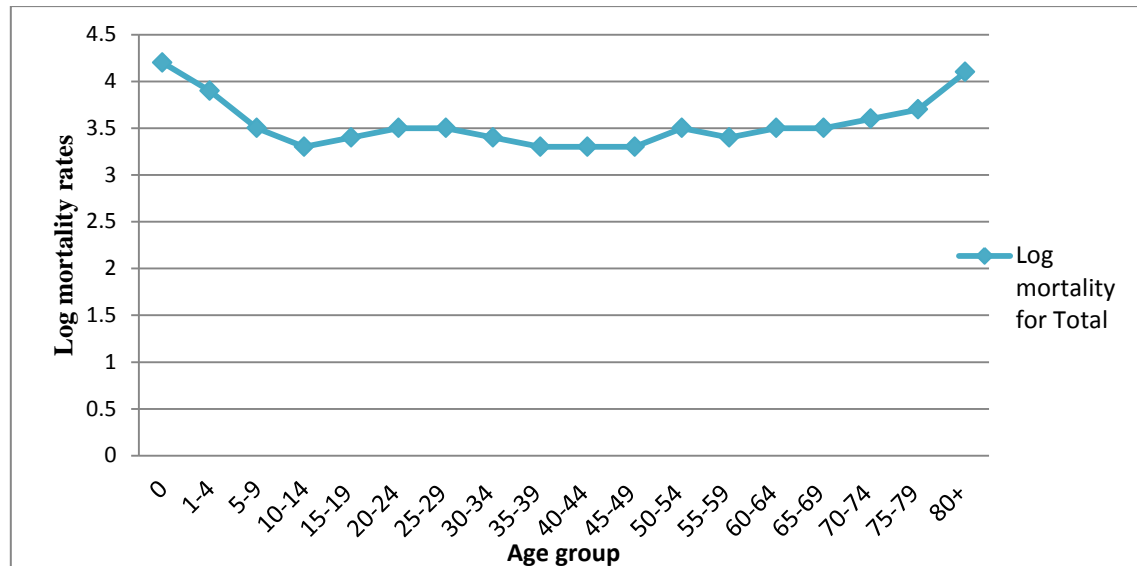
Figure 2.1: Log mortality rates for total population

The figure 4.1.shows logarithmic transformation of expected deaths for both sexes. However, Log transformations show that a number of expected deaths is observed in individual aged under 1 year and above 80 years old. Figure 4.2 was plotted based on 4 th Rwandan Population and housing Census data (NISR,2012).

4.2. General assessment of Lee Carter models:

Recall the basic LC model of age-specific death rates (ASDRs, and denoted m $_{x,t}$ is:

$$\ln(m_{x,t})=\alpha_x + \beta_x\ k_t + \varepsilon_{x,t}, \tag{4.2}$$

Whereby α_x describes the general age shape of the ASDRs, while k_t is an index of the general level of mortality. The coefficients β_x describe the tendency of mortality at age x to change when the general level of mortality k_t changes.

4.3 Methods used in Parameter estimation of Lee Carter model:

To study the efficiency of the Lee-Carter method, the model's parameters may be estimated with the Singular Value Decomposition (SVD) approach and the Maximum Likelihood Estimate (MLE) In using the Lee-Carter approach, the time-series process for the mortality index (parameter k_t) is of critical importance because the entire mortality forecast is determined by an extrapolation of k_t.

4.3.1 Singular value decomposition method (SVD):

In order to obtain a unique solution for the system of equations of the model ln $(m_{x,t})=\alpha_x + \beta_x\ k_t + \varepsilon_{x,t}$, α_x is set equal to the means over time of ln [m(x,t)] and β_x is constrained to sum to unity. The k(t) values sum to zero. Under this normalization, β_x is the proportion of the change in overall mortality on the logarithmic scale attributable to age x. To find a solution, α_x is subtracted from ln [m(x,t)] and the model becomes log-additive.

4.3.2 Maximum likelihood estimation:

Alternative means of fitting Lee carter model is specify a probabilistic mode whose parameters can be estimated by the method of maximum likelihood .Let D_{xt} denote a random variable representing the death count at age x and time t and dxt be the corresponding number of deaths actually observed. Therefore, D_{xt} can be approximated by a poisson distribution with mean λ_{xt},where $\lambda_{xt}=m_{xt}\ E_{xt}$ and E_{xt} denotes the exposure to risk at age x and time t.

Likelihood function for a single age time combination can be written as:L(d, λ)=$\frac{\lambda^d e^{-\lambda}}{d!}$ and the full log likelihood is

$$l=\sum_x \sum_t [d_{x,t}\ \ln(\lambda_{x,t}) - \lambda_{x,t} - \ln(d_{x,t}!)] \tag{4.3}$$

Maximum likelihood estimates are the values of λ_{xt} that maximize the equation (4.3.)It is sufficient to maximize the equation $\sum_x \sum_t [d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t}$. If no restrictions on the form of λ_{xt}, then it is easy to verify that the equation $\sum_x \sum_t [d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t}.]$ attain its maximum values when $\lambda_{xt} = d_{x,t}$. The results obtained in figure 4.1 showed that Lee carter methodology couldn't be easily applied in forecasting the mortality trend index since age group and year to year based data on deaths were not available.

4.4. ANALYSIS ON INFANT MORTALITY AND UNDER 5 MORTALITY RATES:

4.4.1 Time plot of infant mortality rates from 2000 to 2015:

The figure 4.2 is showing the time plot of Infant Mortality Rates (IMR) taking observations from 2000 to 2015
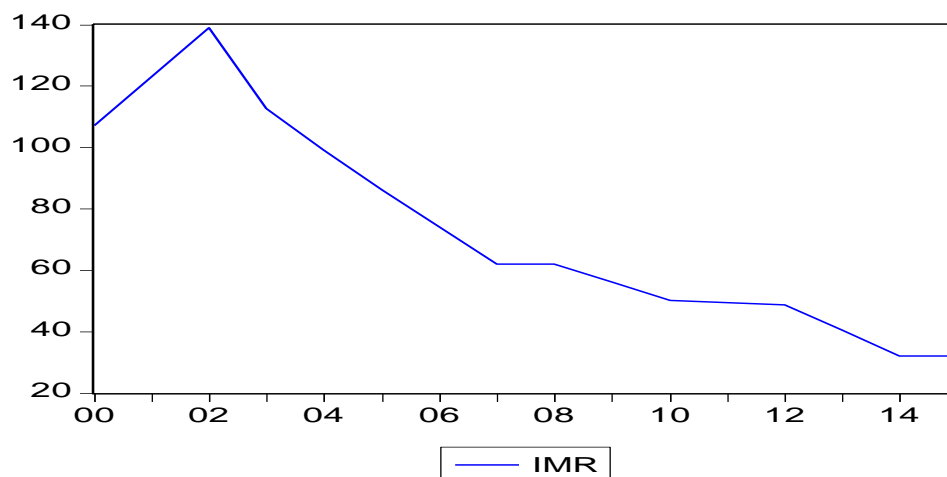

Figure 4.2: Time plot of infant mortality rate 2000-2015

Figure 4.2 is showing a decrease trend in infant mortality rates from 2000 to 2015 where a high IMR was observed in 2002(139 deaths per 1000 live births).
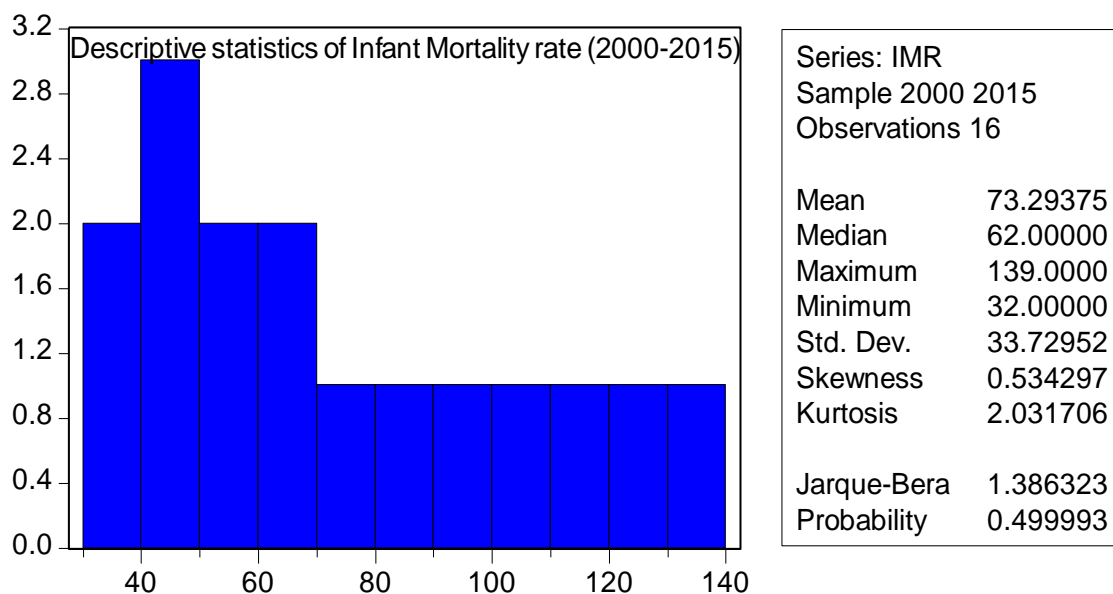

Figure 4.3: Descriptive statistics of infant mortality rates (‰) from 2000 to 2015

Mean rate of infant mortality data is 73 and the dispersion is 33.7 .The highest rates in the series is 139 and the lowest 32. Results also shows that Jarque Bera 1.386323 and the probability is 0.4999.
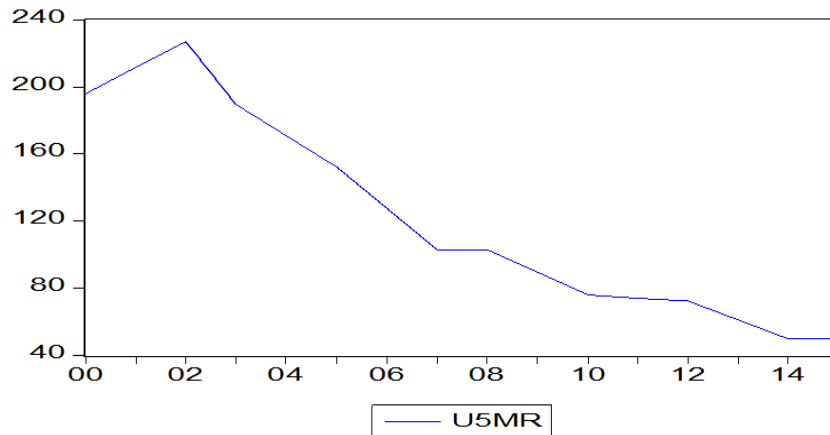
4.4.2 Time plot of under 5 mortality rates:

Figure 4.4: Time plot of Under 5 mortality rates from 2000 to 2015

Figure 4.4 is indicating that under 5 mortality rate was high at 227‰ in 2002 whereas the lowest under 5 mortality rate is observed in 2014 and 2015.
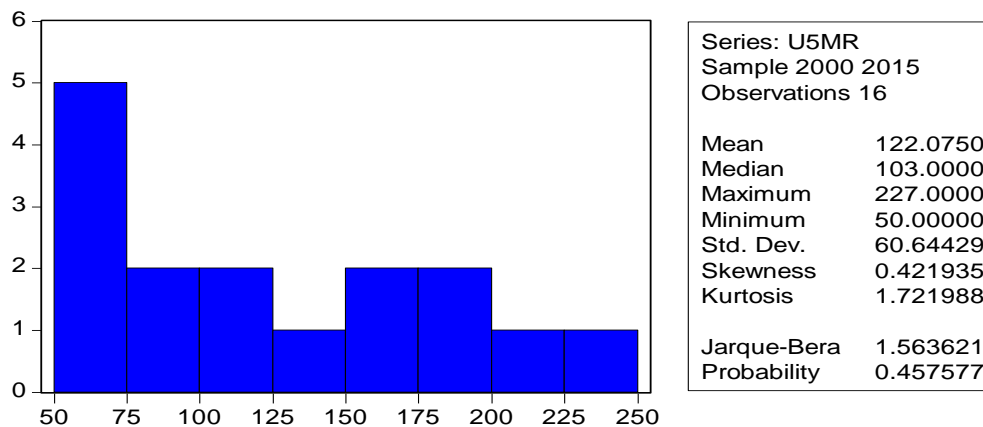

Series: U5MR	
Sample 2000 2015	
Observations 16	
Mean	122.0750
Median	103.0000
Maximum	227.0000
Minimum	50.00000
Std. Dev.	60.64429
Skewness	0.421935
Kurtosis	1.721988
Jarque-Bera	1.563621
Probability	0.457577

Figure 4.5 : Descriptive statistics of under 5 mortality rates (‰) (2000-2015)

Figure 4.5 shows that the mean rate of under 5 mortality is 122‰. Data also shows a highest under 5 mortality rate of 227‰ and the lowest of 50 ‰ and so the dispersion of data around the mean is 60.

4.4.3. Unit root test of mortality rates by ADF test:

Table 4.2 Unit root test for mortality rates for original data

ADF Test Statistic	-2.368987	1% Critical Value*	-4.7315
		5% Critical Value	-3.7611
		10% Critical Value	-3.3228
*MacKinnon critical values for rejection of hypothesis of a unit root. Augmented Dickey-Fuller Test Equation Dependent Variable: D(IMR) Method: Least Squares Sample(adjusted): 2001 2015 Included observations: 15 after adjusting endpoints			
Variable	Coefficient	Std. Error	t-Statistic	Prob.
IMR(-1)	-0.554738	0.234167	-2.368987	0.0355
C	69.74789	31.27310	2.230284	0.0456
@TREND(2000)	-4.070236	1.727944	-2.355538	0.0363
R-squared	0.323958	Mean dependent var	-5.000000

Table 4.2 is showing that infant mortality rates (IMR) data are not stationary at levels. In other words there is a unit root since ADF statistic -2.4 is not less than critical value of -3.7611 at 5% significance level. The first difference was necessary to make data stationary.

Table 4.3 First differencing for infant mortality rates ((‰)

ADF Test Statistic	-5.381061	1%	Critical Value*	-4.0681
		5%	Critical Value	-3.1222
		10%	Critical Value	-2.7042

*MacKinnon critical values for rejection of hypothesis of a unit root.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(IMR,2)
Method: Least Squares
Sample(adjusted): 2003 2015
Included observations: 13 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(IMR(-1))	-1.366235	0.253897	-5.381061	0.0003
D(IMR(-1),2)	0.238590	0.185386	1.286995	0.2271
C	-10.34843	2.601644	-3.977651	0.0026
R-squared	0.764671	Mean dependent var		-1.230769
Adjusted R-squared	0.717605	S.D. dependent var		13.81891
S.E. of regression	7.343490	Akaike info criterion		7.024680
Sum squared resid	539.2685	Schwarz criterion		7.155053
Log likelihood	-42.66042	F-statistic		16.24682
Durbin-Watson stat	0.762538	Prob(F-statistic)		0.000722

We rejected a null hypothesis that there is a unit root in IMR because ADF Test statistic (-5.381) is less that critical value of -3.1222 at $\alpha = 5\%$. Hence first differencing is making data stationary. ADF test also showed that original data of Under 5 mortality rates were not stationary .Hence fist difference .

Results are then presented in the table 4.4 which showed that data became stationary by first difference.

Table 4.4: First differencing of Under 5 mortality rates data

ADF Test Statistic	-4.303207	1%	Critical Value*	-4.0681
		5%	Critical Value	-3.1222
		10%	Critical Value	-2.7042

Method: Least Squares
Sample(adjusted): 2003 2015
Included observations: 13 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(U5MR(-1))	-1.310873	0.304627	-4.303207	0.0016
D(U5MR(-1),2)	0.209699	0.219887	0.953666	0.3627
C	-17.04993	4.792534	-3.557602	0.0052
R-squared	0.686459	Mean dependent var		-1.192308

Table 4.4 shows that the null hypothesis of existence of unit root is rejected at $\alpha = 5\%$. The reason is that ADF test statistic (-4.3) is less than critical value of -3.12 at the said significance level α. In order words we reject the hypothesis of a unit root because |ADF|> Critical value in absolute value Therefore under 5 mortality rate data becomes stationary after first difference.

4.4.4 Autocorrelation Function and Partial Autocorrelation Function for mortality rates:

Table 4.5. ACF AND PACF of mortality rates

Lags	ACF	PACF	Q-stat	Probability
1	0.856	0.856	14.077	0.000
2	0.635	-0.367	22.377	0.000
3	0.396	-0.144	25.850	0.000
4	0.214	0.092	26.946	0.000
5	0.050	-0.190	27.011	0.000
6	-0.078	-0.033	27.185	0.000
7	-0.177	-0.040	28.182	0.000
8	-0.254	-0.127	30.498	0.000

Table 4.5 is showing that ACF is dying off geometrically with increasing lag k. Hence low order autoregressive process. As PACF is very significantly positive of 0.856 at lag 1 and close to 0 thereafter. This is the fact that the patterns of autocorrelation can be captured by auto regression of order 1 AR (1). The probability value for all lags are 0 .Therefore ARIMA model to be used in mortality forecasting is ARIMA (p,d,q) where p=1, d=1 and q=0 .

Table 4.6.ARIMA models for Infant mortality rates from 2000 to 2011

Dependent Variable: X_t Method: Least Squares Sample(adjusted): 2001 2011 Included observations: 11 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.117490	0.500516	-0.234738	0.8197
Xt(-1)	1.010629	0.112767	8.962094	0.0000
R-squared	0.899238	Mean dependent var		4.355701
Adjusted R-squared	0.888042	S.D. dependent var		0.369861
S.E. of regression	0.123756	Akaike info criterion		-1.178042

Table 4.6 shows   Model: $X_t = -0.117490 + 1.010629 X_{t-1} + \varepsilon_t$,   **(4.4)**

The model may be used to forecast for next period it mean from 2012 to 2015 since the coefficient of determination R^2 is 89.9%.. The equation in 4.1 shows that the model is statistically significant since 89.9% is high and this can also be explained by residual plots in figure 4.6.
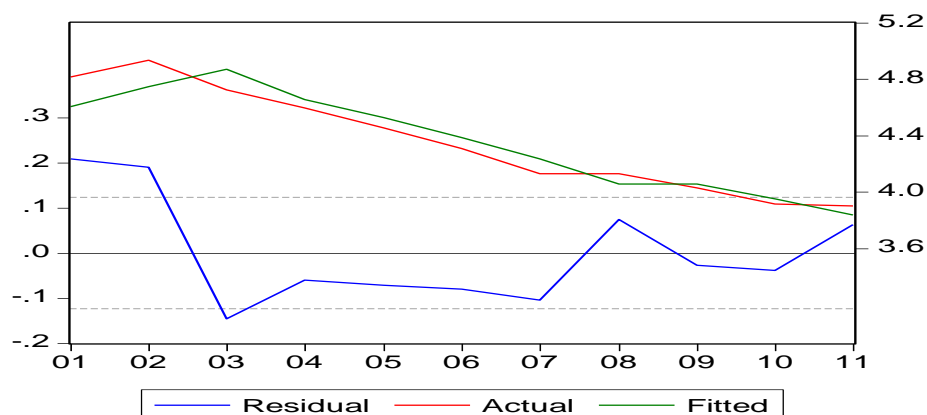

Figure 4.6: Residual plot for infant mortality rates from 2000 to 2011

Figure 4.6 shows that Residual, Actual and fitted plots of Infant mortality rates are showing a decreasing trend and so the model $X_t = -0.117490 + 1.010629 X_t + \varepsilon_t$ may be used to forecast next 4 years basing on sample from 2000 to 2011.However, ex post forecasting of data is done where data is already available from 2000 to 2015 and the sample is taken to be able to forecast.

Table 4.7: ARIMA model of under 5 mortality rates from 2000 to 2011

Dependent Variable: Y_t				
Method: Least Squares				
Sample(adjusted): 2001 2011				
Included observations: 11 after adjusting endpoints				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.281125	0.444377	-0.632629	0.5427
Yt(-1)	1.038933	0.089602	11.59492	0.0000
R-squared	0.937257	Mean dependent var		4.857793
Adjusted R-squared	0.930285	S.D. dependent var		0.405206
S.E. of regression	0.106989	Akaike info criterion		-1.469221
Sum squared resid	0.103019	Schwarz criterion		-1.396877
Log likelihood	10.08072	F-statistic		134.4422
Durbin-Watson stat	1.605815	Prob (F-statistic)		0.000001

Table 4.7 shows ARIMA model for under 5 mortality rates of data from 2010 to 2011 is $Y_t = -0.281125 + 1.038933Y_{t-1} + \varepsilon_t$.   **(4.5)**

In equation (4.5) the coefficient of determination R^2=93.7% meaning that under 5 mortality rates are explained by lag values at 93.7%. Therefore, this model may be used to forecast U5MR from 2012 to 2015.

4.5. Testing adequacy of ARIMA models for mortality forecasting:

ARIMA models to be used in forecasting, there are criteria to be followed such as normality and no serial correlation in errors. However, Jarque Bera test and probabilities were used to show the normality in residuals.

Then after,We have examined various measures of forecasting Errors, namely the mean absolute error (MAE); the root mean squared error (RMSE); and Thieles U to confirm the adequacy of models that were used.
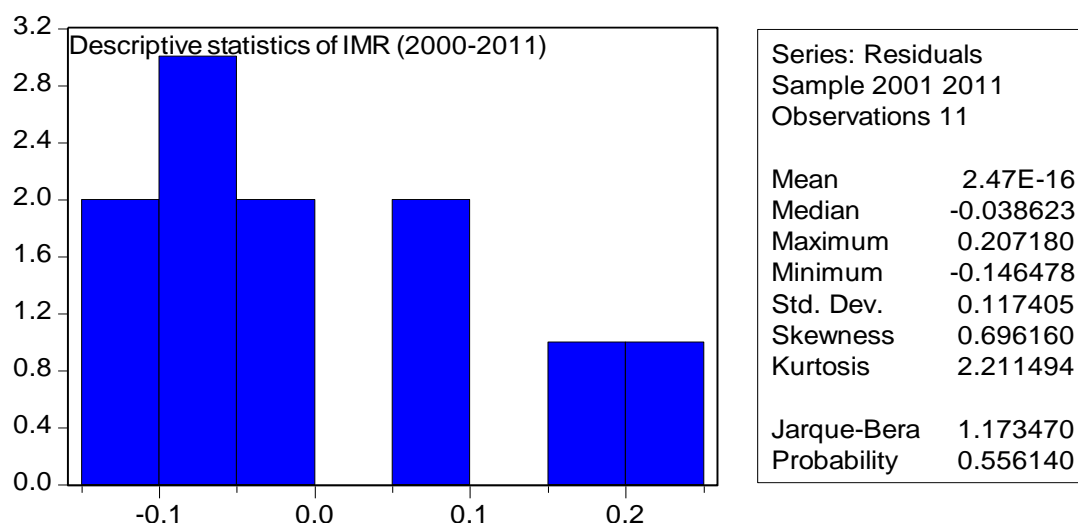

Figure 4.7: Normality test for ARIMA model of Infant Mortality rates

The result of the Normality test shows that Jarque Bera value is 1.173470 with a probability of 0.556, this probability value, however is more than 0.05 meaning that we cannot reject the null hypothesis; instead we reject the alternative hypothesis and fail to reject the null hypothesis that the residual is normally distributed. Therefore residuals for infant mortality rates from 2000 to 2011 are normally distributed and may be used in forecasting for next 4 years.
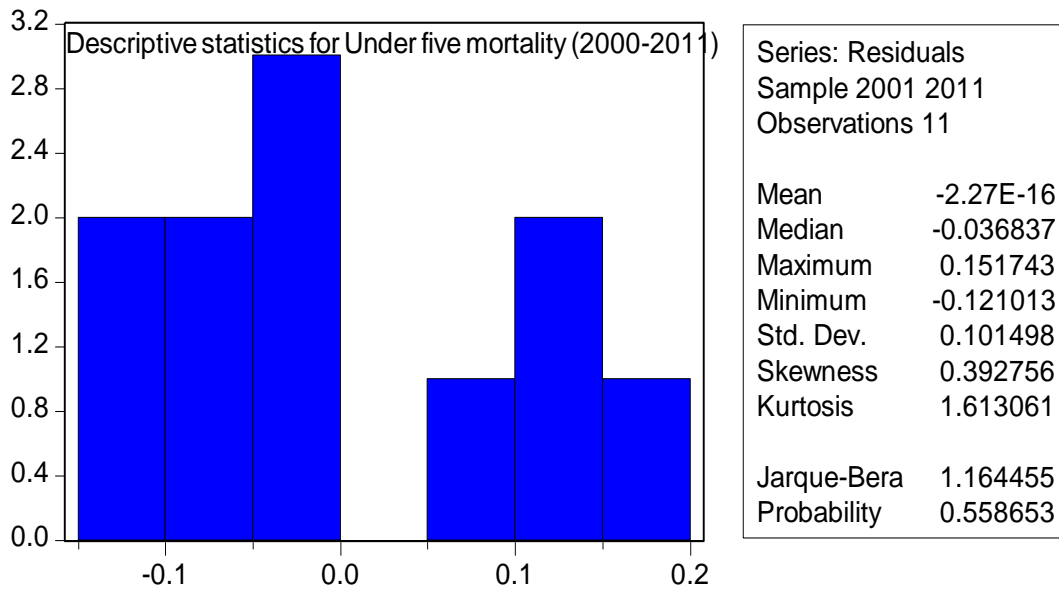
Figure 4.8: Normality test for ARIMA model of Under 5 mortality rates

The results of the Normality test shows that Jarque Bera value is 1.164 with a probability of 0.558, this probability value is more than 0.05 meaning that we cannot reject the null hypothesis that the residual is normally distributed. Therefore residuals for under 5 mortality rates from 2000 to 2011 are normally distributed and may be used in forecasting for next 4 years that is from 2012 to 2015
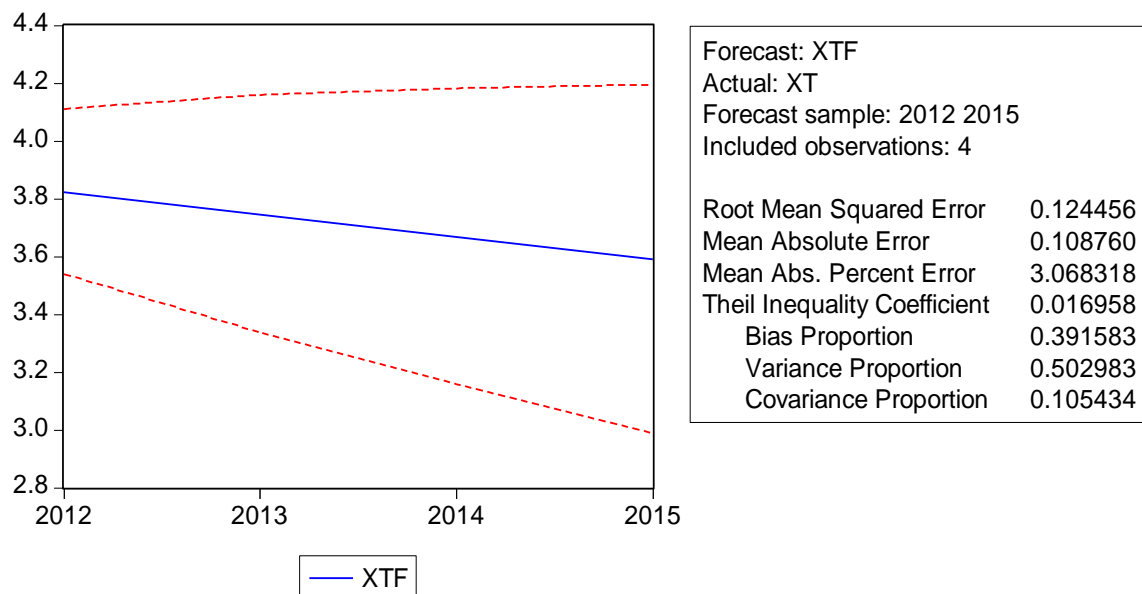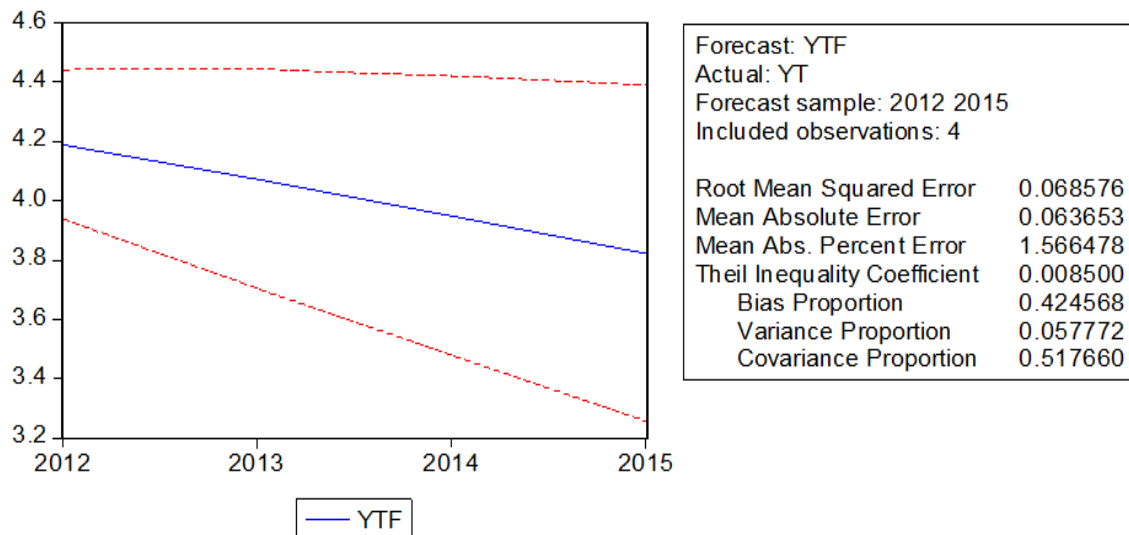
4.6. Ex post forecasting for mortality rates:

Figure 4.9: Forecasting of Infant mortality rates from 2012 to 2015

In ex post forecasting, it is clear that the model for infant mortality rate is good for forecasting because the root mean square error is very small (0.124456) and there is no serial correlation between error terms(probability value for observed R squared is 0.567>5%).

Results indicate various measures of forecasting Errors, namely the mean absolute error (MAE); the root mean squared error (RMSE); and Thieles U, the smaller the error the better the forecasting ability of that model accordingly. The Theil inequality coefficient always lies between zero and one, where zero indicates a perfect fit.

Forecast: YTF
Actual: YT
Forecast sample: 2012 2015
Included observations: 4

Root Mean Squared Error	0.068576
Mean Absolute Error	0.063653
Mean Abs. Percent Error	1.566478
Theil Inequality Coefficient	0.008500
Bias Proportion	0.424568
Variance Proportion	0.057772
Covariance Proportion	0.517660

Figure 4.10 Forecasting: Under 5 mortality rates forecasts from 2012 to 2015

Figure 4.10 shows forecasts from 2012 to 2015 , the mean absolute error (MAE) which is 0.0636; the root mean squared error (RMSE) which is 0.068 show a higher predictive ability of the model Y_t = -0.281125+1.038933 Y_{t-1} +ε_t since they are very small. Therefore there is a decreasing trend of under 5 mortality rates by taking 4 observations from 2012 to 2015.Results indicate various measures of forecasting Errors, namely the mean absolute error (MAE); the root mean squared error (RMSE); and Thieles U, the smaller the error the better the forecasting ability of that model accordingly. The Theil inequality coefficient always lies between zero and one, where zero indicates a perfect fit. Therefore, the figure 4.10 shows a decreasing trend of under 5 mortality rates forecasted from 2012 to 2015.

5. CONCLUSION

The female population growth rater=2.3% obtained from Leslie matrix is a good measure of population projection and therefore the population is increasing since λ>1.

Mortality data was stationary at first diffence .In this model ARIMA model with p=1, d=1 and q=0 was used in forecasting infant mortality and under 5 mortality rates..

Autocorrelation Function for mortality rates was dying off geometrically with increasing lag value k. This meant a low order autoregressive process. As PACF was very significantly positive of 0.856 at lag 1 and close to 0 thereafter. This is the fact that the patterns of autocorrelation can be captured by auto regression of order 1 AR (1). The probability value for all lags are 0 .Therefore ARIMA model to be used in mortality forecasting is ARIMA (p,d,q) where p=1, d=1 and q=0 .

ARIMA models that were used in forecasting Infant Mortality rates and unde 5 mortality rates were $X_t = -0.117490 + 1.010629 + \varepsilon_t$ and $Y_t = -0.281125 + 1.038933 Y_{t-1} + \varepsilon_t$ respectively. Both ARIMA models have a higher predictive ability because of small root mean square errors where RMSE is 0.124456 for Infant mortality rates and 0.068576for under 5 mortality rates.

Jarque Bera test and Probability values for residuals for both Infant mortality rate and under 5 mortality rates were > 0.05 and so this was the evidence of not rejecting the Null hypothesis of normality in residuals. Therefore, ARIMA models used were helpful in forecasting.

Forecasts showed a decreasing trend from 2012 to 2015 both for Infant Mortality and under 5 mortality rates.

REFERENCES

[1] Gonze ,D.(2012) *Discrete age structured models:The Leslie matrices*,1.

[2] Gonze ,D.(2015). *Discrete age structured models:Leslie matrices,3*.

[3] Haberman, S.,& Russolillo.M (2015). Lee Carter Mortality Forecasting: Application to the Italian Population.*Cass business school, 162*, 20-40.

[4] Haberman, S.,& Russolillo.M (2005). Lee Carter Mortality Forecasting: Application to the Italian Population. *Cass business school, 167*,6.

[5] Jenny, Z.(2007).Fitting and forecasting mortality for Sweden :Applying Lee carter model *Stockholm University, 4.*

[6] NISR.(2012).*4th population and housing census*. Kigali - Rwanda.

[7] NISR.(2015). 4 th Demographic and health Survey .Kigali -Rwanda.

[8] Shahzad , N.(2007).Forecasting of inflation rate.

[9] UN.(2015).World population prospects.*New York.*

[10] Ofosuhene , O. (2009) .*Demographic modeling of Human Population Growth,* Lappeenranta.

[11] Richard,A.(1980).Population and economic changes in Developing countries ,*University of Chicago Press.*

[12] Samuel,H.(1980).Causes and consequences of mortality declines in Less developed countries during twentieth century ,*University of Chicago Press*,p.289-360.

[13] World Bank, United Nations &World Health Organization. (2013).*Estimates development by   the UN Inter agency Group for Child Mortality Estimation*,23.