

Estimation of Growth model for Population of Ethiopia using Least square method

Mulugeta Tesfa¹, Ashager Adane², Muluken Admasu³

¹Statistics Stream, Department of Pre-Engineering, Engineering College, Wollo University

²Mathematics Stream, Department of Pre-Engineering, Engineering College, Wollo University

³Mathematics Stream, Department of Pre-Engineering, Engineering College, Wollo University

Abstract: The purpose of this study focused on modeling the population of Ethiopia using different models and estimating the models parameters via least square method. The models, that were applied for the population growth, were Malthus growth model, Logistic growth model and General growth model. To identify the models which performed effectively in prediction of the actual population, the measure accuracy has been used, such that the models satisfying the criteria of the measure of accuracy is the best statistical model. The results of the analysis were presented using tables and graphical form which are very good to perform comparison for the effectiveness of the models. In this study, MAPE, RSE, MAD and R2 which are considered to measure the accuracy of the models. Malthus growth model, Logistic growth model and General growth model used the population of Ethiopia from 1980 to 2020 inclusive, the data was obtained from international data base (IDB). R studio 3.6.3 were used to estimate the models parameters using simple codes. The study proposed to project the population of Ethiopia via General growth model which performed best in measure of accuracies that makes it effective and efficient as compare with the other models. The model had the smallest RSE (492,155), MAPE (0.75%) and MAD (379,942) as well as the highest R2(99.97%) relative to the other models.

Keyword: Logistic model, parameter estimation, Malthus model, General model, selection methods

Date of Submission: 11-08-2020

Date of Acceptance: 27-08-2020

I. Introduction

Ethiopia population is estimated to be about 114,310,670 in the year 2020 (World meters, 2020) that rose in rapid population growth, which in many instances far outstrips economic growth and environmental sustainability, is the reality in most developing countries of sub-Saharan Africa (UNDP, 1992). The type of development which involves complex interconnected processes of expanding education, health, employment and other basic needs provides the chief ingredient for lasting demographic transition with markedly reduced fertility and mortality. Sub-Sahara is among the major regions of the developing world that have not yet undergone a general decline in fertility (Jehan I., 2009). Ethiopia is the second most populous country in the sub-Saharan Africa next to Nigeria, with population growth rate of 2.6 percent per year and the total fertility rate approximate to 5 (CSA, 2007, Sintayehu A. 2016).

The most recent census was conducted in 2007 and estimated a population of 73.5 million, corresponding to an average annual growth rate of 2.9 percent. In 2012, an Inter-Censual Population Survey (ICPS) estimated the population at 83.7 million, corresponding to an average annual growth rate of 2.6 percent (Alemayehu B. and Yihunie, 2014). Predicting the future population is the main activity in order to prepare the effective and efficient development planning for the next development. Population estimation supplies the future estimates of Ethiopia population which is basic for planners and policy makers who works on the improvement of quality of life of citizens (Sintayehu A., 2016).

The aim of this paper was to identify the population growth model for Population of Ethiopia and select the best model for projection (prediction) the future population of Ethiopia.

II. Methodology

This study used a data of population of Ethiopia in international Data Base (IDB) from 1980 to 2020 (December, 2019). The study used the real data to estimate and select a model out of three models namely Malthus, Logistic and General growth model (Sintayehu A, 2016). The results obtained from the models were used to predict the future population of Ethiopia. The data analysis was done using a statistical Software (R studio version 3.6.3) in order to estimate the model parameters with simple built in function of R studio. The results conveyed using tables and graphs to explore the result of the analysis which are very crucial to create essential insight about the model parameters and to make a comparison among the models for Ethiopian population. The estimation method of the model parameters of the models was obtained using least square

technique for all models (Mulugeta et al, 2020). The next section explained about the models development, parameter estimation and model selection methods.

1. Development of the models

A. Malthus growth Model

Thomas R. Malthus (1798), proposed a mathematical model for the population growth with the assumption that the population rises at a constant rate proportional to the original population size. Malthus model applied for Ethiopia population and performed better for estimation (projection) (Sintayehu A, 2016) as well as the model proved that it was a model to projection of Tanzania population (Andongwisye et al, 2019). Malthus did not take into account the fact that in any given environment, the growth of population may stop due to the density of population or competition of resources (Al-Eideh and AlOmar 2019). The model is expressed in simple differential equation as follows;

$$\frac{dP}{dt} = \beta P \tag{1}$$

Where P the total population size and β is the constant growth rate defined as the difference between the birth rate and death rate for a certain population size. We rearrange and integrate both side of equation (1)

$$\int_{P_0}^{P(t)} \frac{dP}{P} = \int_{t_0=0}^t \beta dt$$

Finally, the exponential model is expressed as

$$P(t) = P_0 e^{\beta t} \tag{2}$$

Where, P_0 represents the population at some specified time, $t = 0$ that means P_0 is constant number. By considering $z(t) = \frac{P(t)}{P_0}$ and rescaling equation (2) expressed as:

$$z(t) = e^{\beta * t} \tag{3}$$

B. Logistic growth model

Logistic growth model was proposed by Verhulst in 1845 (Marsden et al. 2003). The model is the extension of Malthusian model. The main feature of logistic model is the inclusion of the limiting value which is the size of population that an environment can support (Andongwisye J and Allen R, 2019). Ali et al. (2015) studied census data and predicted population of Bangladesh by using logistic model. They used a curve fitting method and tried to compare the prediction between the case when carrying capacity is known and not known. The model suggested that the population growth depends on carrying capacity and the maximum rate of growth. The model was applied for Ethiopia population and less preformed than the exponential model (Sintayehu A, 2016). The logistic model is an extension of exponential model that includes the ideal conditions that some were excluded in the model as follows

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right), \quad r, K > 0 \tag{4}$$

Where K is the maximum sustainable population (carrying capacity), it is assumed to constant value and r is the growth rate which is unknown parameter in this case, r and r/K are vital constants. The logistic model reduces to exponential growth signifying that as P is greater K then the rate of growth becomes negative and population decreases.

From equation (1.4) we can find the solution of the non-linear differential equation as follows;

$$\begin{aligned} \frac{dP}{P \left(\frac{K-P}{K}\right)} &= r dt \\ \int_{P_0}^{P(t)} \frac{K}{P(K-P)} dP &= \int_0^t r dt \\ \int_{P_0}^{P(t)} \left[\frac{1}{P} + \frac{1}{K-P} \right] dP &= \int_0^t r dt \end{aligned}$$

By applying the concept of calculus and some rearrangement, the logistic model for the population is given as (Sintayehu A, 2016, Mussa A. and Jung I, 2019)

$$P(t) = \frac{K}{1 + \left(\frac{K}{P_0} - 1\right) e^{-rt}} \tag{5}$$

$$P_{max} = \lim_{t \rightarrow \infty} P(t) = \lim_{t \rightarrow \infty} \left(\frac{K}{1 + \left(\frac{K}{P_0} - 1\right) e^{-rt}} \right) = K$$

Let P_0 : Population at $t = t_0 = 0$, P_T : Population at $t = T$ and P_{2T} : Population at $t = 2T$, then

$$\frac{1}{K} [1 - e^{-rT}] = \frac{1}{P_T} - \frac{e^{-rT}}{P_0} \quad (6)$$

$$\frac{1}{K} [1 - e^{-2rT}] = \frac{1}{P_{2T}} - \frac{e^{-2rT}}{P_0} \quad (7)$$

By dividing equation (6) to (7) gives

$$\frac{[1 - e^{-rT}]}{[1 - e^{-2rT}]} = \frac{\frac{1}{P_T} - \frac{e^{-rT}}{P_0}}{\frac{1}{P_{2T}} - \frac{e^{-2rT}}{P_0}}$$

So that $e^{-rt} = \frac{P_0(P_{2T}-P_T)}{P_{2T}(P_T-P_0)}$

This implies that, If $0 < \frac{P_0(P_{2T}-P_T)}{P_{2T}(P_T-P_0)} < 1$

$$r = \frac{1}{T} \ln \left(\frac{P_0(P_{2T} - P_T)}{P_{2T}(P_T - P_0)} \right)$$

By direct substitution into equation (1.6) yields

$$K = \frac{P_T(P_0P_T - 2P_0P_{2T} + P_T P_{2T})}{(P_T)^2 - P_0P_{2T}} \quad (8)$$

The value of K in equation (8) is constant as per the values in the equation are given. This implies that equation (5) can be formulated by some rearrangement

$$\frac{P(t) * (1 - \frac{K}{P_0})}{P(t) - K} = e^{r * t} \quad (9)$$

Let $v(t) = \frac{P(t) * (1 - \frac{K}{P_0})}{(P(t) - K)}$ is, then equation (9) expressed as

$$v(t) = e^{r * t} \quad (10)$$

C. General growth model

This is a modification of Malthus growth model by considering P_0 as the unknown parameter. The model was applied for Tanzania population and Ethiopia population as a candidate for the population projection (Sintayehu A, 2016, Mussa A. and Jung I, 2019) and indicated that the model performed less as compared to the Malthus and logistic models. On the other hand, a review (Mulugeta and et al, 2020) indicates that the general growth model was the best for the prediction of united republic of Tanzania. Amare and Mulugeta (2017) studied about Ethiopian population and suggested that exponential trend model was selected for projecting Population of Ethiopia. The general expression of the model is given as follows (Sintayehu A., 2016 and Mussa A. and Jung I, 2019)

$$p = \alpha * e^{\beta * t} \quad (11)$$

Where α and β are the parameters to be estimated for the prediction of the population.

2. Parameter Estimation of the Models

A French mathematician Adrien-Marie Legendre (1805) developed method of least squares is an algorithm in regression analysis that has most important application in data fitting. This algorithm involves minimization of sum of squared residuals for the sake of maximizing the model parameters. In various areas of experimental sciences, maximization theory is associated with accuracy and precision of the predicted output is attained by error reduction. For many observed data point, the method of least square is reasonably most systematic procedure to fit the unique curve. Suppose we have set of observation, $((t_0, P_0), (t_1, P_1), \dots, (t_{n-1}, P_{n-1}))$.

A. Malthus growth Model

Using equation (3), we can estimate the parameter using method of least square technique. We need to transform the Malthus growth model the exponential model ($z(t) = e^{\beta * t}$) into linear model by taking natural logarithm (ln) both and the linear model is given as

$$\ln(z(t)) = \beta * t$$

Suppose that $y(t) = \ln(z(t))$ and thus the linear model is

$$y_i = \beta * t_i \quad (12)$$

The sum square error of the equation (2.2) which is formulated as

$$E(\beta) = \sum_{i=1}^n (y_i - \beta * t_i)^2$$

To estimate the parameter β , we need to minimize the sum square error by driving $E(\beta)$ with respect to β and equal with 0.

$$\frac{dE(\beta)}{d\beta} = 0$$

So that the estimated value for β can be calculated by using the following equation

$$\hat{\beta} = \frac{\sum_{i=1}^n y_i * t_i}{\sum_{i=1}^n t_i^2} \quad (13)$$

B. Logistic growth Model

Using equation (10), we estimate the model parameter using method of least square. We need to transform the Logistic growth model the exponential model ($v(t) = e^{r*t}$) into linear model by taking natural logarithm (ln) both and the linear model is given as

$$\ln(v(t)) = r * t$$

Suppose that $y(t)=\ln(v(t))$ and thus the linear model is

$$y_i = r * t_i \quad (14)$$

The sum square error of the equation (14) which is formulated as

$$E(r) = \sum_{i=1}^n (y_i - r * t_i)^2$$

To estimate the parameter r , we need to minimize the sum square error by driving $E(r)$ with respect to r and equal with 0.

$$\frac{dE(r)}{dr} = 0$$

So that the estimated value for r can be calculated by using the following equation

$$\hat{r} = \frac{\sum_{i=1}^n y_i * t_i}{\sum_{i=1}^n t_i^2} \quad (15)$$

C. General growth model

By considering equation (11), We transform the general growth model ($P = \alpha * e^{\beta*t}$) into linear model by taking natural logarithm (ln) both in order to apply least square method for parameter estimation and the linear model can be written as

$$\ln P = \ln(\alpha * e^{\beta*t}) = \ln(\alpha) + \beta t$$

Assume that $y = \ln P$, and $\gamma = \ln(\alpha)$, Generally we have

$$y_i = \beta * t_i + \gamma \quad (16)$$

We define the error associated in the set of data with the equation (16) by

$$E(\beta, \gamma) = \sum_{i=1}^n (y_i - (\beta t_i + \gamma))^2$$

For $(n - 1)$ times the variance of the data set $\{(y_0 - (at_0 + \gamma), \dots, y_{n-1} - (at_{n-1} + \gamma))\}$

To get the optimal values of β and γ we have to minimize the residual ($E(\beta, \gamma)$) by taking the partial derivative of the error with respect to β and γ , equal with 0, that can be done by using the concept of Calculus,

$$\frac{\partial E}{\partial \beta} = \frac{\partial E}{\partial \gamma} = 0$$

Differentiating $E(a, \gamma)$ with respect to α and γ and solve for α and γ respectively.

$$\frac{\partial E}{\partial \beta} = 2 \sum_{i=1}^n (y_i - (at_i + \gamma)) (-t_i) = 0$$

$$\frac{\partial E}{\partial \gamma} = 2 \sum_{i=1}^n (y_i - (at_i + \gamma)) = 0$$

After some rearrangement, the estimated value of β and γ are calculated as

$$\hat{\beta} = \frac{\sum_{i=1}^n y_i * t_i - \frac{(\sum_{i=1}^n t_i) * (\sum_{i=1}^n y_i)}{n}}{\sum_{i=1}^n t_i^2 - \frac{(\sum_{i=1}^n t_i)^2}{n}} \quad (17)$$

$$\hat{\gamma} = \frac{\sum_{i=1}^n y_i}{n} - \hat{\beta} * \frac{\sum_{i=1}^n t_i}{n}$$

3. Model Selection criteria

The study applied different models for the same dataset which required various ways to compare the statistical models in order to select best statistical model to fit the dataset. Statistical techniques that are used for the comparison of the models to fit the population of Ethiopia.

A. Mean Absolute Percentage Error (MAPE)

Mean absolute percentage error is one of a measure of accuracy of the models prediction. It is formulated as (Sintayehu A. 2016 and Amare and Mulugeta, 2017)

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|X_i - \hat{X}_i|}{X_i} * 100 \quad (18)$$

Where X_i is the actual observation of the population of Ethiopia at i^{th} time. \hat{X}_i is the fitted value of the population at i^{th} time. n is number of observation for the population. The model having the least MAPE preforms better to predict the population due to having the smallest percentage error.

B. Residual Standard error (RSE)

Residual standard deviation (error) is a way of measuring the variation between the actual and the predicted value of the dataset. This measure can be calculated using the following

$$RSE = \sqrt{\frac{\sum_{i=1}^n (X_i - \hat{X}_i)^2}{n - 1}} \quad (19)$$

Where X_i is the actual observation of Ethiopia population at i^{th} time. \hat{X}_i is the fitted value of the population at i^{th} time. n is number of observation for the population. A model is said to have the best performance if SEE of the model is the smallest that indicates the model is best to predict the population (Mulugeta et al, 2020 and Amare and Mulugeta, 2017).

C. Mean Absolute Residual Deviation (MAD)

Mean absolute Residual deviation, is a measure of accuracy of the models prediction, is the average of the absolute difference between the actual and the fitted value of the population. This can be mathematically expressed as (Mulugeta et al, 2020)

$$MAD = \sum_{i=1}^n \frac{|X_i - \hat{X}_i|}{n} \quad (20)$$

Where X_i is the actual observation of the population of Ethiopia at i^{th} time. \hat{X}_i is the fitted value of the population at i^{th} time. n is number of observation for the population. The model having the least MAPE preforms better to predict the population due to having the smallest percentage error (Amare and Mulugeta, 2017).

D. Coefficient of Determination (R^2)

This measures how the model fits the dataset, which is the proportion of variation of the predicted value in a unit change of time (Mulugeta et al, 2020). This measure is calculated as

$$R^2 = r^2$$

Where r is the correlation coefficient between the actual and the predicted observation by the model which is calculated as follows

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x}) * (\hat{x}_i - \bar{\hat{x}})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (\hat{x}_i - \bar{\hat{x}})^2}} \quad (21)$$

A model with the highest R^2 is the best model for the prediction of the population (Mulugeta et al, 2020).

III. Result

The unknown values for the models, that they were explained in section 2, were estimated using Ethiopia federal democratic republic (EFDR) population data. Using measure of accuracy criteria's, the best model was selected which has been used for the projection of Ethiopia population in the future.

a. Malthus growth model

The model has one parameter (growth rate) of the population for the country and the value estimated using method of least square with the result given in table 1 below. The growth rate of the model was 0.0276 and the parameter is statistically significant at 0.1% ($Pvalue < 0.001$) with a minimum standard error for the estimate (table 1). The significance of the parameter implies that the Malthus growth model is applicable and important to model the population of Ethiopia.

Table 1: Summary for the estimate of the growth rate of Malthus model

Coefficient	Estimate	Std. Error	t value	Pr(> t)
β	2.760e-02	7.979e-05	345.9	<2e-16 ***

*** significant at 0.1%

Therefore, the Malthus growth model of the population growth is given as

$$P(t) = 36036457 * e^{0.0276*t}$$

b. Logistic growth model

The Carrying Capacity of population of Ethiopia, which is deterministic value for the logistic models and considered as a constant value for the model, can be estimated at the points $(P_0, P_T, P_{2T}) = (36,036,457, 62,891,069, 108,113,150)$. Then the value for K can be estimated as follows

$$K = \frac{P_T(P_0 * P_T - 2P_0 * P_{2T} + P_T * P_{2T})}{(P_T)^2 - P_0 * P_{2T}}$$

By substitute the values on the equation, the carry capacity of population of Ethiopia estimated as

$$K = 1,351,470,438.$$

This indicates the maximum carry capacity of the country is 1,351,470,438. This can be written as

$$\lim_{t \rightarrow \infty} P(t) = \lim_{t \rightarrow \infty} \left(\frac{K}{1 + \left(\frac{K}{P_0} - 1\right)e^{-rt}} \right) = 1,351,470,438$$

And hence, by applied a least square technique for the estimation of growth rate (r) in logistic growth model. As mentioned in the table 2 below, the growth rate the population of Ethiopia is 2.88% which is statistically significant estimate of equation (2.0) (*Pvalue* < 0.001) and had a minimum standard error of the estimate.

Table 2: summary result of growth rate of logistic model

Coefficient	Estimate	Std. Error	t value	Pr(> t)
<i>r</i>	2.881e-02	7.968e-05	361.6	<2e-16 ***

*** significant at 0.1%

From this, we can get the formula for equation (2.0)

$$v(t) = e^{0.0288*t}$$

With some modification of equation (2.0), Logistic growth model is

$$P(t) = \frac{1,351,470,438}{1 + 36.503 * e^{-0.0288*t}}$$

c. General growth Model

We estimating β and γ after transforming the general growth model into linear model. We used R software to estimate the parameters by using least square method. The results of the estimates given in table 3 below. As indicated in table 3, the estimates were statistically significant at 1% (*pvalue*<0.001).

Table 3: summary of the estimates of the parameters for general growth model

Coefficients	Estimates	Std. Error	t value	Pr(> t)
γ	1.741e+01	3.429e-03	5076.6	<2e-16***
β	2.728e-02	1.476e-04	184.9	<2e-16***

*** significant at 0.1%

Then we transformed the linear model into exponential model $\gamma = 17.41$ and $\beta = 0.0273$ such that $\ln(\alpha) = \gamma$, $\alpha = \exp(\gamma) = \exp(17.41) = 36397112$, Therefore, the general equation for population growth of Ethiopia is given as

$$P(t) = \alpha * \exp(\beta * t) = 36397112 * \exp(0.0273 * t).$$

As a candidate, we have three models for the prediction of Ethiopia population. However, we need to select the best model to predict the population of Ethiopia.

As indicated in plot 1 below, the models were performing in almost the same in prediction of the actual population of Ethiopia. The performance of the models almost the same however, their errors from the actual models may not identical that the model has to be selected as a best model if it has the least error on average from the actual population observation. Using table 4 and the figure below the general growth model performs best to project Ethiopia population and it is suggested to the general growth model for predicting the Ethiopia population.

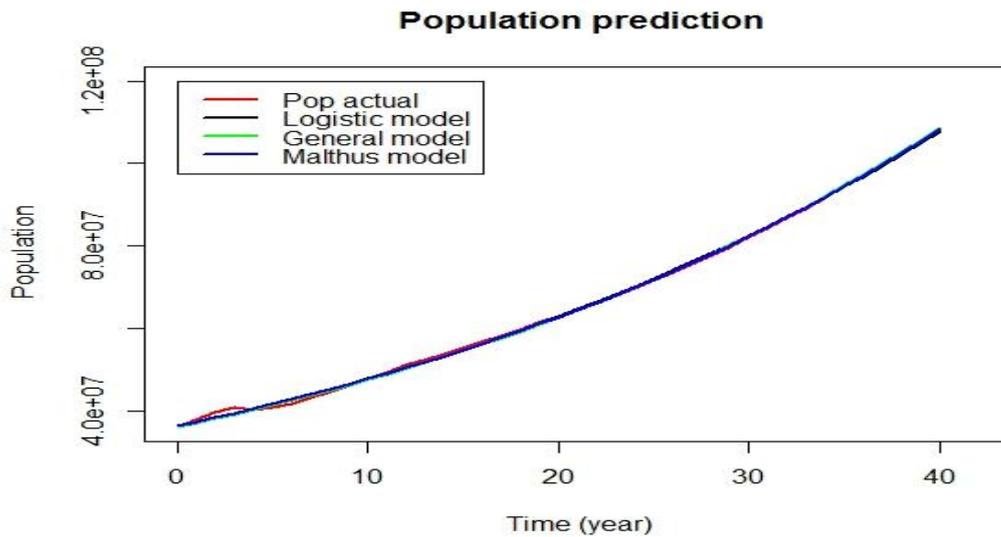


Figure 1: Time versus the predicted population

Table 4 indicates the numerical diagnosis of the models using by using SEE, MAD, MAPE and R^2 which are essential to select the best models to predict the population of Ethiopia. The model with the smallest SEE, MAD, MAPE and the model having the highest R^2 was the best model to predict the population of Ethiopia. As a result, the General growth model has the smallest SEE, MAD and MAPE and additionally the model has the higher R^2 as compared to the other models. Hence, $P(t) = \alpha * e^{\beta * t}$ is the best model to predict the population of Ethiopia.

Table 4: The diagnosis for the models

Growth Models	RSE	MAD	R2	MAPE(%)	Rank
Malthus	545,705	414,844	0.9995	0.7857	2
General	492,155	379,942	0.9997	0.7517	1
Logistic	522,925	417,086	0.9994	0.7784	3

Figure 2 indicates the residual plot of the modes which conveys the deviation of the actual and the predicted values using the models and indicates the distance from zero. The proposed model using the graph is that the general growth model because on average the model is best relative to the other model that is supported by table 4.

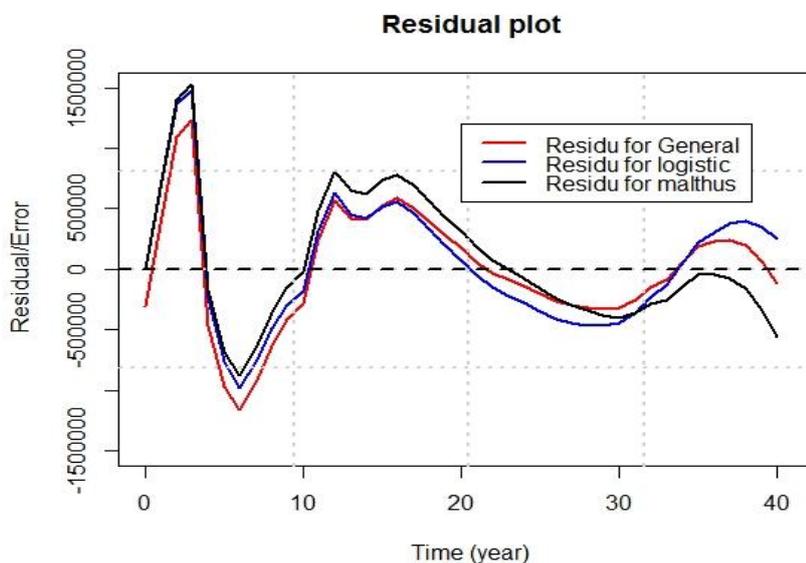


Figure 2: The plot for residuals of the models

IV. Discussion

Using the figures provided in the above the actual and predicted population were almost the same for Malthus, Logistic and General growth models. The study used 41 years population from 1980-2020 inclusively. It finds the models that fitted the dataset very well. The model parameters were statistically significant at 0.1% for all models which was the evidence to use the models to fit the data, that led the models were good to apply for the projection of population of Ethiopia.

The models fitted the population strongly as indicated in figure 1 with the fitted versus the actual population. The models were compared using the accuracy measures to identify the best model out the three models. Mean absolute percentage error (MAPE) of Malthus, Logistic and General growth model were 0.79%, 0.78% and 0.75% respectively (table 4). The best model for the population projection was found to be General growth model because the model had the least values in MAD (379,942), MAPE (0.75%) and RSE (492,155) and had the highest R^2 (99.97%). This result is identical with the result of Mulugeta et al (2020) and it is contradicted with the result in Ethiopia population whereas Malthus growth model was selected as best model for the projection of Ethiopian population (Sintayehu A., 2016) and Logistic growth model was found the best model for the projection of United republic of Tanzania population (Mussa A. et al 2019).

The 2nd best model for the study were Malthus growth model as indicated in table 4. The population growth rate of Ethiopia based on the information in International Data Base (IDB) and world meter were approximately 2.74% per year in 2017 and 2.62% in the years 2018 and 2.60% in year 2019 which founds the population growth rate 2.76% for Malthus model, 2.88% for logistic growth model and 2.73% per year for General growth model. The study supposed that the population growth rate of the general growth model and the approximated growth rate.

V. Conclusion

The study has been explained the population of Ethiopia using Malthus, Logistic and General growth model. On the bases of the result that was obtained in the analysis part. The estimation of the model parameters was statistically significant which means that all the models were important for the projection of population of Ethiopia. The estimation method for the models parameters of the models was least square.

The general growth model was the best model for the projection of Ethiopian population because it is the 1st in ranking using the measure of accuracy techniques. Malthus growth model was 2nd model for the projection population of Ethiopia which it had the 2nd performance of the prediction of the population of Ethiopia. The study proposed that General growth model is the best model that is capable to project the future population of Ethiopia.

As recommendation, the model parameters of the models may be estimated various techniques of estimation out of which one can have a power to estimate the model parameters (growth rate). It is better to work on the estimation of the model parameters using different estimation methods rather than using least square method of estimation.

Reference

- [1]. Alemayehu B and Yihunie L, 2014, Projecting Ethiopian demographics from 2012-2050 Using the spectrum suite of models, Ethiopian public health association.
- [2]. Ali LE, Khan BR and Sams IS, 2015, Brief study on census and predicted population of Bangladesh using logistic population model. *Ann. Pure Appl. Math.* 10(1): 41-47.
- [3]. Al-Eideh BM and Al-Omar HO, 2019, Population projection model using exponential growth function with a birth and death diffusion growth rate processes. *Eur. J. Sci. Res.* 151(3): 271-276
- [4]. Amare Wubishet Ayele and Mulugeta Aklilu Zewdie, 2017, Modeling and Forecasting Ethiopian Human Population Size and its Pattern, *International Journal of Social Sciences Arts and Humanities* Vol. 4 No. 3. 2017. Pp. 71-82
- [5]. Andongwisye J Mwakisisile1 and Allen R Mushi, 2019, Mathematical Model for Tanzania Population Growth, *Tanzania Journal of Science* 45(3): 346-354, ISSN 0856-1761, e-ISSN 2507-7961.
- [6]. Ashraf Q and Galor O, 2008, Malthusian population dynamics: Theory and evidence. *Working paper, Brown University, Department of Economics*, No. 2008-6
- [7]. Augustus Wali, 2011, Mathematical Modeling of Rwanda's Population Growth, *Applied Mathematical Sciences*, Vol. 5, no. 53, 2617 – 2628
- [8]. Central Statistical Agency (CSA). 2007. *2007 National Population and Housing Census of Ethiopia*. Addis Ababa: CSA.
- [9]. Ethiopian Demographic and Health Survey (2011). Central Statistics Agency, Addis Ababa, Ethiopia
- [10]. Mulugeta T. et al, 2020, Review on Estimation and Prediction of Population Using Mathematical Models in Tanzania. *IOSR Journal of Mathematics (IOSR-JM)*, e-ISSN: 2278-5728, p-ISSN: 2319-765X, 16(3): pp. 33-40.
- [11]. Mussa A. S and Jung I.H., 2019, Estimation and Prediction of Population Using Mathematical Models in Tanzania, *IOSR Journal of Mathematics (IOSR-JM)*, e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 15, Issue 2 Ser. III, PP 73-82 SU
- [12]. Jehan I et al, 2009, Neonatal mortality, risk factors and causes: A prospective population-based cohort study in urban Pakistan, *world health organ.* 87(2):130–138 IITE
- [13]. Sintayehu, A. M., 2016, *Mathematical Model of Ethiopia's Population growth*. *Journal of Natural Sciences Research*. Vol. 6 (17). ISSN 2224-3186

- [14]. T. R. Malthus. 1798. An Essay on The Principle Of Population (1st edition, plus excerpts 1803 2nd edition), Introduction by Philip Appleman, and assorted commentary on Malthus edited by Appleman, Norton Critical Editions, ISBN 0-393-09202-X.
- [15]. World meter, 2020, United Nations, Department of Economic and Social Affairs, Population Division, ([Online] Available: <http://www.worldometers.info/worldpopulation/ethiopia-population/> Accessed: 2020)
- [16]. UNDP, 1992, Human Development Report 1992 -The Real Wealth of Nations: Pathways to Human Development. New York: Human Development Report Office.

Appendix

A. Logistic growth model

pt=62891069; p0=36036457; pl=108113150

$k = \frac{pt * (p0 * pt - 2 * p0 * pl + pt * pl)}{(pt)^2 - p0 * pl}$

b=k/p0

> v=t\$pop*(1-b)/(t\$pop-k)

> t1<-data.frame(v,t)

> log.lm<-lm(log(v)~-1+t,data=t1)

> summary(log.lm)

Call:

lm(formula = log(v) ~ -1 + t, data = t1)

Residuals:

Min	1Q	Median	3Q	Max
-0.023963	-0.005729	0.000000	0.005642	0.038154

Coefficients:

Estimate Std. Error t value Pr(>|t|)
t 2.881e-02 7.968e-05 361.6 <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01186 on 40 degrees of freedom

Multiple R-squared: 0.9997, Adjusted R-squared: 0.9997

F-statistic: 1.307e+05 on 1 and 40 DF, p-value: < 2.2e-16

> pred.log<-k/(1+(b-1)*exp(-coef(log.lm)*t\$t))

B. General growth model

> mls<-lm(log(pop)~t,data=t)

> summary(mls)

Call:

lm(formula = log(pop) ~ t, data = t)

Residuals:

Min	1Q	Median	3Q	Max
-0.0275281	-0.0039193	-0.0004352	0.0048279	0.0307393

Coefficients:

Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.741e+01 3.429e-03 5076.6 <2e-16 ***

t 2.728e-02 1.476e-04 184.9 <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01118 on 39 degrees of freedom

Multiple R-squared: 0.9989, Adjusted R-squared: 0.9988

F-statistic: 3.418e+04 on 1 and 39 DF, p-value: < 2.2e-16

> pred.mls<-exp(coef(mls)[1]+coef(mls)[2]*t\$t)

C. Malthus growth model

> ex.mlt<-lm(log(scale)~-1+t,data=t)

> summary(ex.mlt)

Call:

lm(formula = log(scale) ~ -1 + t, data = t)

Residuals:

Min	1Q	Median	3Q	Max
-0.020938	-0.003893	-0.000733	0.009324	0.038271

Coefficients:

Estimate Std. Error t value Pr(>|t|)
t 2.760e-02 7.979e-05 345.9 <2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
 Residual standard error: 0.01187 on 40 degrees of freedom
 Multiple R-squared: 0.9997, Adjusted R-squared: 0.9997
 F-statistic: 1.196e+05 on 1 and 40 DF, p-value: < 2.2e-16

```
> pred.mlt<-p0*exp(coef(ex.mlt)*t$t)
> yt<-t$pop
> time<-t$t
> tim<-seq(1:51)
> for.mls<-exp(coef(mls)[1]+coef(mls)[2]*tim)
> dat<-data.frame(tim,for.mls)
> bets<-data.frame(yt,t,pred.log,pred.mls,pred.mlt)
> #par(mfrow=c(1,3))
```

D. Plot code

```
> plot ( yt ~ time , data = bets ,
+       xlab = " Time (year)", ylab = " Population",
+       main = " Population prediction", type = "n",ylim=c(3.6e+7,12e+7),xlim=c(0,42))
> lines ( yt ~ time , data = bets,col="red",lwd = 2)
plot ( yt ~ time , data = bets ,
       xlab = " Time (year)", ylab = " Population",
       main = " Population prediction", type = "n",ylim=c(3.6e+7,12e+7),xlim=c(0,42))
lines ( yt ~ time , data = bets,col="red",lwd = 2)
#points(time, yt, col="red", pch="1")
lines ( pred.log ~ time , data = bets,col="black",lwd = 2)
lines ( pred.mlt ~ time , data = bets, lty =1,col="green",lwd = 2)
lines (pred.mls ~ time , data = bets, lty =1,col="blue",lwd = 2)
#points(time, pred.log, col="blue", pch="*")
legend(0,1.2e+8,legend=c("Pop actual","Logistic model","General model","Malthus model"),
       col=c("red","black","green","blue"),lty=c(1,1,1,1),lwd=c(2,2,2),ncol=1)
#####
```

E. Comparison code

```
yg<-pred.mls
Res<-yt-yg
TD<-yt-mean(yt)
SR<- yg-mean(yt)
bias<- sum(Res)/length(Res)
SEEml<- sqrt(sum(Res^2)/(length(Res)-1))
MADml<-sum(abs(Res))/length(Res)
MAPml<-sum(abs(Res)/yt)*100/length(Res)
R2ml<-cor(yt,yhat)^2
D=1-(sum(Res^2)/sum((abs(yhat-mean(y))+abs(y-mean(y))))^2)
#####
ylog<-pred.log
rlog<-yt-ylog
TD<-yt-mean(yt)
SR<- ylog-mean(yt)
bias<- sum(rlog)/length(rlog)
SEEllog<- sqrt(sum(rlog^2)/(length(rlog)-1))
MADlog<-sum(abs(rlog))/length(rlog)
MAPlog<-sum(abs(rlog)/yt)*100/length(rlog)
R2log<-cor(yt,ylog)^2
D=1-(sum(rlog^2)/sum((abs(ylog-mean(y))+abs(y-mean(y))))^2)
#####
yex<-pred.mlt
Resex<-yt-yex
TD<-yt-mean(yt)
SR<- yex-mean(yt)
bias<- sum(Resex)/length(Resex)
SEEx<- sqrt(sum(Resex^2)/(length(Resex)-1))
MADex<-sum(abs(Resex))/length(Resex)
```

```
MAPex<-sum(abs(Resex)/yt)*100/length(Resex)
R2ex<-cor(yt,yex)^2
D=1-(sum(Resex^2)/sum((abs(yex-mean(yt))+abs(y-mean(yt))))^2)
#####
R2<-c(R2ex,R2ml,R2log)
SEE<-c(SEEex,SEEml,SEElog)
MAD<-c(MADex,MADml,MADlog)
MAP<-c(MAPex,MAPml,MAPlog)
name<-c("Expo","MLS","logistic")
stat<-data.frame(name,SEE,MAD,R2, MAP)
print(stat,digits = 5)
stat2<-data.frame(R2ml,SEEml,MADml)
print(stat2,digits = 5)
time<-t$t
rg<-Res
rex<-Resex
rl<-rlog
dat<-data.frame(time,rg,rl,rex)
plot ( rg ~ time , data = dat,
       xlab = " Time (year)", ylab = " Residual/Error",
       main = " Residual plot", type = "n",ylim=c(-1.5e+6,1.5e+6),xlim=c(0,41))
lines( rg ~ time , data = dat,col="red",lwd = 2)
lines ( rl ~ time , data = dat,col="blue",lwd = 2)
lines (rex ~ time , data = dat,col="black",lwd = 2)
grid(4, 4, lwd = 2) legend(20,1.2e+6,legend=c("Residu for General", "Residu for logistic", "Residu for malthus"),
       col=c("red", "blue", "black"), lty=c(1,1,1), ncol=1,lwd=c(2,2,2))
abline(h=0,lwd=2,lty=2)
```

Mulugeta Tesfa, et. al. "Estimation of Growth model for Population of Ethiopia using Least square method." *IOSR Journal of Mathematics (IOSR-JM)*, 16(4), (2020): pp. 01-11.