FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 6 No. 1, March, 2022, pp54 - 59
DOI: https://doi.org/10.33003/fjs-2022-0601-881

AN APPLICATION OF TIME INDEPENDENT FOURIER AMPLITUDE MODEL ON FORECASTING THE UNITED STATE POPULATION

***Sameer, A. S., Yusuf M. and Okafor, U. L.**

Department of Mathematical Sciences, Federal University Dutsin-Ma, Katsina State Nigeria

*Corresponding authors' email: sameer.sas.sa@gmail.com

ABSTRACT

This study applied the Time Independent Fourier Amplitude Model Approach to forecast the Population of the United States of America from 1790 to 2020 and beyond on a 10-year interval using Number Crunches Statistical software (NCSS). Results obtained using this methodology was compared with the results obtained in the other models: Malthusian, Logistics, and Logistics (Least Squares) Model. These models were compared using the goodness of fit (the coefficient of determination (R2) and the sum of square error (SSE)), the Akaike information criterion (AIC), Bayesian information criterion (BIC), Mean Absolute Deviation (MAD), Mean Error (ME), and Mean Sum of square Error (MSSE), Results displays that the Time Independent Fourier Amplitude Model and also is a suitable model for predicting the United States population.

Keywords: Forecasting, Fourier Coefficients, Model, Population, Time

INTRODUCTION

In understanding of the world population unceasingly developing in the last century, population forecast becomes more and more significant in policy making, economic development, education, and so on Lassila, Jukka, *et al* (2014). Population is pretentious by numerous issues such as policy, economy, and culture. Therefore, it is tough for demographic researchers to explore each factor. Among all philosophies, historic data are the significant basis in forecasting. By evaluating the essential tendency inside the historical data, a comprehensive and reasonable forecast can be made without deliberating each factor that affects the population. In order to achieve satisfying forecast performance reliable past data are required. The Malthusian, Logistic and Logistics (Least Squares) models are frequently used approaches for population forecast and have publicized good performance. The Logistic model was initially presented in 1837 by the Dutch bio mathematician Pierre Verhulst (1845, 1847). Later advanced the Logistic model which has showed to be well suitable for population forecasting Miranda *et al* (2010).

This paper reviews four population growth models and tries to find out an appropriate way to clarify and predict population growth. It shows that the Malthusian, Logistics, and Logistics (least squares) model should not be used to predict population growth. The Malthusian model predicts that the population would grow without certain, but this cannot possibly happen indeterminately. Most populations are forced by limitations on resources even in the short run and none is unrestrained forever. Therefore, the conclusion cannot be applied to the continuing population growth since no weak dependent stable relationships exist in these models. Hence, it would be tough to make population predictions using these models. This ultimately goes into capacity and converges to it carrying capacity as seen in the application and results. Obviously, these models are limited to a range of dates where the growth rate remains relatively constant and is highly questionable to predict a population over long periods of time where growth rate varies because of the nature of these models and can rapidly diverge from the actual population.

Population forecasting is an essential effort to understand population growth, which affects various features of a country's society and economy, including future demand for food, water, energy, and services. Mathematical models are frequently used to understand the relationship of the migration, birth and death rates on population growth. Mathematical models support population forecasting by capturing statistical predispositions from past data sets. However, these need to be cautiously compared to understand the allegations of different model formulations in predicting future population, which the models have not seen or were qualified on. Professionals in this field deliberate over the finest ways to use these models to make dependable forecasts. The Malthusian, Logistics, and Logistics (least squares) model is one of three traditional population growth models; the population growth calculated by these models is limited, but the equilibrium between the population and social resources is concerned after a period of unrestrained population growth, with a decline in a population, the aging of the population is a significant problem.

We theorized that the Time Independent Fourier Amplitude Model is more accurate and genuine. To test this hypothesis, we obtained the US population data from 1970 to 2020 from the US Census Bureau. Then, we employed the data for numerical verification and to compare the accuracy of the mathematical models of the Malthusian, logistic, Logistics (least Squares) and the Time Independent Fourier Amplitude law for population predictions. Moreover, to regulate whether the model gives a statistically significant perfection in the fit as compared to the former laws, we showed a T-test, Sum of Square of Error (SSE), Mean Error (ME), Mean Absolute Deviation (MAD), R-Squared (R^2), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mean Sum of Square of Error (MSSE) assessment. In order to verify and strengthen the conclusion that it can be widely used. Regression permits us to calculate how much of that difference in outcome is due to the predictors we are to test it using the Time Independent Fourier Amplitude Model and also runs an actual explanation of the population's growth. As exposed in this paper, predictions attained via the Time Independent Fourier Amplitude Model are in relatively acceptable agreement with present certified predictions from the US Census Bureau. In conclusion, it is clear that predicting population tendencies is not just a mathematical problem but a complex set of interrogations that can generate distressed deliberations about the value of our life and future.

MATERIALS AND METHOD

An analytic function could be represented by means of a series of sine's and cosines specifically by the series. For the Classical Fourier series, the synthesis equation is given by:

$$Y_t = \alpha_o + \sum_{n=1}^{\infty} \left(a_n \operatorname{Cos}\frac{n\pi t}{L} + b_n \operatorname{Sin}\frac{n\pi t}{L}\right), \text{ for } [-N \le t \le N] \quad (1)$$

The trigonometric series which converges and has a continuous function Y_t as its sum on the interval $[-N, N]$is given by;

$$Y_t = \alpha_o + \sum_{n=1}^{\infty}(a_n \operatorname{Cos} \mathrm{nt} + b_n \operatorname{Sin} \mathrm{nt}) \qquad [-N \le t \le N] \quad (2)$$

If we integrate both sides of Equation (1) and implement that it's acceptable to integrate the series term by term and we acquire.

$$\int_{-N}^{N} Y_t dt = \int_{-N}^{N} \alpha_o dt + \int_{-N}^{N} \sum_{n=1}^{\infty}[a_n Cos(nt) + b_n \operatorname{Sin}(nt)]dt$$

$$= N\alpha_o + \int_{-N}^{N} \sum_{n=1}^{\infty}(a_n \operatorname{Cos}(\mathrm{nt}))\, dt + \int_{-N}^{N} \sum_{n=1}^{\infty}(b_n \operatorname{Sin}(\mathrm{nt}))\, dt$$

And we obtain $\alpha_o = \frac{1}{N}\int_{-N}^{N} Y_t dt$ (3)

To determine $\alpha_o$for n≥1 we multiply both sides of equation (1) by $\operatorname{Cos} mt$ (where m is an integer and m ≥1) and integrate term by term from $[-N \le t \le N]$:

$$\int_{-N}^{N} Y_t Cosmt\, dt = a_m N$$

Solving for a_m and then changing m by n, we have

$$\alpha_n = \frac{1}{N}\int_{-N}^{N} Y_t\, Cos(nt)\, dt \qquad n = 1, 2, 3\ldots \quad (4)$$

Likewise, if we multiply both sides of equation (1) by $Sin\ mt$ and integrate from $[-N, N]$ we get

$$b_n = \frac{1}{N}\int_{-N}^{N} Y_t\, Sin(nt)\, dt \qquad n = 1, 2, 3\ldots \quad (5)$$

This is for a single valued function which is continuous and has a finite number but without discontinuities. Equation (4) and (5) will be used if the data is continuous.
But when the data is discrete, then equation (6) and (7) called the classical equation:

$$a_n = \frac{2}{N}\sum_{N=1}^{23} Y_t Cos(nt) \quad (6)$$

$$b_n = \frac{2}{N}\sum_{N=1}^{23} Y_t\, Sin(nt) \quad (7)$$

In a state where the data is developing or trending, we will have the form of equation

$$Y_t = A + B \times t^n + \sum_{n=1}^{m} \left(a_n \cos\frac{n\pi t}{L} + b_n \sin\frac{n\pi t}{L}\right) \quad (8)$$

If equation (8) is linear then we have the equation (9)

$$Y_t = A + B \times t + \sum_{n=1}^{m} \left(a_n \cos\frac{n\pi t}{L} + b_n \sin\frac{n\pi t}{L}\right) \quad (9)$$

If the equation (8) is quadratic then we have the equation (10)

$$Y_t = A + B \times t^2 + \sum_{n=1}^{m} \left(a_n \cos\frac{n\pi t}{L} + b_n \sin\frac{n\pi t}{L}\right) \quad (10)$$

The Time Independent Fourier Amplitude Model can be written as:

$$Y_t = A + B \times t + C \times t^2 + \frac{t}{N}\sum_{N=1}^{m} \left(\mathrm{D} \cos\frac{2\pi}{L} + \mathrm{E} \sin\frac{2\pi}{L}\right) \quad (10)$$

Where Y_t and $\hat{Y}_t$ is the actual and predicted values respectively A, B, C, D and E are the parameters used in the model, t is the time which is censored, N is the set of data points and L is the seasonal length.

RESULT AND DISCUSSION

The Time Independent Fourier Amplitude Model was used to run a nonlinear regression analysis under 0.05 level of significance using Number Crunches Statistical Software (NCSS). Results attained from this investigation used in this model displays the values of the parameters, value of coefficient of determination (R^2) the Sum of Square Error (SSE), the population prediction and forecast, and the normal Probability plot of error.
Table 2 shows how five (5) parameters were used in this model, there are four (4) values of these parameters that are statistically significant. The value attained from these parameters was compared with that of the T-table value of 1.96 under 0.05 level of significance. If the values of the discrete parameters are greater than the T-table value, this accomplishes that these parameters are statistically significant, but if otherwise they are not. The coefficient of determination (R^2) of the Time Independent Fourier Amplitude Model has value of 0.9993 showing that out of 100% it's accounted for 99.93% of the consistency of the population leaving only 0.07% to coincidental. Moreover, the higher the R^2 the better the model fits any data set.

Table 1: Model Estimation Section of the Time Independent Fourier Amplitude Method

Parameter Name	**Parameter Estimate** $\boldsymbol{\alpha}$	**Asymptotic Standard** $\boldsymbol{\beta}$	**Lower 95% Confidence limit**	**Upper 95% Confidence limit**	**T-Values** $\frac{\alpha}{\beta}$
A	6.059	2.046	1.743	10.376	2.961
B	-1.099	0.413	-1.971	-0.228	-2.661
C	0.667	0.017	0.630	0.703	39.235
D	1.641	0.871	-0.197	3.479	1.884
E	1.556	0.867	-0.273	3.385	1.795

Model: $Y_t = A + B \times t + C \times t^2 + \frac{t}{N}\sum_{N=1}^{m}\left(\mathrm{D}\cos\frac{2\pi}{L} + \mathrm{E}\sin\frac{2\pi}{L}\right)$

R-Squared: 0.9993

Iterations: 3

Estimated Model:

$\hat{Y}_t = 6.059 - 1.099 \times \mathrm{t} + 0.667 \times \mathrm{t} \times \mathrm{t} + \frac{t}{23} \times 1.641 \times \cos(0.6284) + \frac{t}{23} \times 1.556 \times \sin(0.6284)$

The values obtained in table 3 are the actual, predicted, forecast, residual, lower and upper 95% confidence limit. In column 2 (Actual), the prediction made by the United States census bureau stopped at the year 2010 but forecast made by the Time Independent Fourier Amplitude Model exceed the year 2010 and forecast to 2060. Showing it's a very reliable model when predicting population is involved. A confidence interval denotes the probability that a population parameter will descent between a set of values for a certain percentage of times. The values satisfying this interval $a < P < b$ indicates that there's a certainty that the confidence interval contains the true population parameter. Where a, b is the upper and lower 95.0% confidence limit and P is the population of either the actual or the predicted values. By establishing a 95.0% confidence interval using the sample's mean and standard deviation, and applying the normal distribution plot then we attain an upper and lower bound that covers the true mean of 95.0%.

Table 2: Predicted, Projected and Residuals Value Section of the Time Independent Fourier Amplitude Method

ROW NO.	**ACTUAL** Y_t	**PREDICTED** $\hat{Y}_t$	**LOWER 95% CONFIDENCE LIMIT**	**UPPER 95% CONFIDENCE LIMIT**	**RESIDUAL**
1790	3.93	5.450	1.023	12.664	-2.035
1800	5.31	7.868	1.1726	14.564	-2.559
1810	7.24	8.513	2.150	14.877	-1.274
1820	9.64	9.732	3.508	15.956	-0.093
1830	12.87	11.913	5.727	18.098	0.956
1840	17.07	15.584	9.411	21.758	1.485
1850	23.19	21.216	15.052	27.380	1.973
1860	31.44	29.037	22.868	35.207	2.402
1870	39.82	38.951	32.746	45.156	0.868
1880	50.19	50.569	44.302	56.836	-0.379
1890	62.98	63.362	57.031	69.692	-0.382
1900	76.21	76.860	70.490	83.230	-0.650
1910	92.23	90.835	84.465	97.205	1.394

1920	106.02	105.384	99.053	111.715	0.635
1930	123.2	120.895	114.628	127.163	2.304
1940	132.16	137.898	131.693	144.104	-5.738
1950	151.33	156.862	150.693	163.031	-5.532
1960	179.32	178.016	171.852	184.180	1.303
1970	203.3	201.261	195.088	207.435	2.038
1980	226.54	226.210	220.025	232.396	0.329
1990	248.71	252.333	246.110	258.557	-3.623
2000	281.42	279.162	272.798	285.525	2.257
2010	308.75	306.466	299.770	313.163	2.283
2020	?	334.346	327.079	341.613	
2030		363.188	355.150	371.226	
2040		393.522	384.611	402.434	
2050		425.818	416.039	435.598	
2060		460.304	449.732	470.876	

Figure 1 shows the data which are plotted against a theoretic normal distribution in such a way that the points forms an approximately linear pattern, which specifies that Time Independent Fourier Amplitude Method is a good model for this data set. Deserting from this straight line indicate deserting from normality.

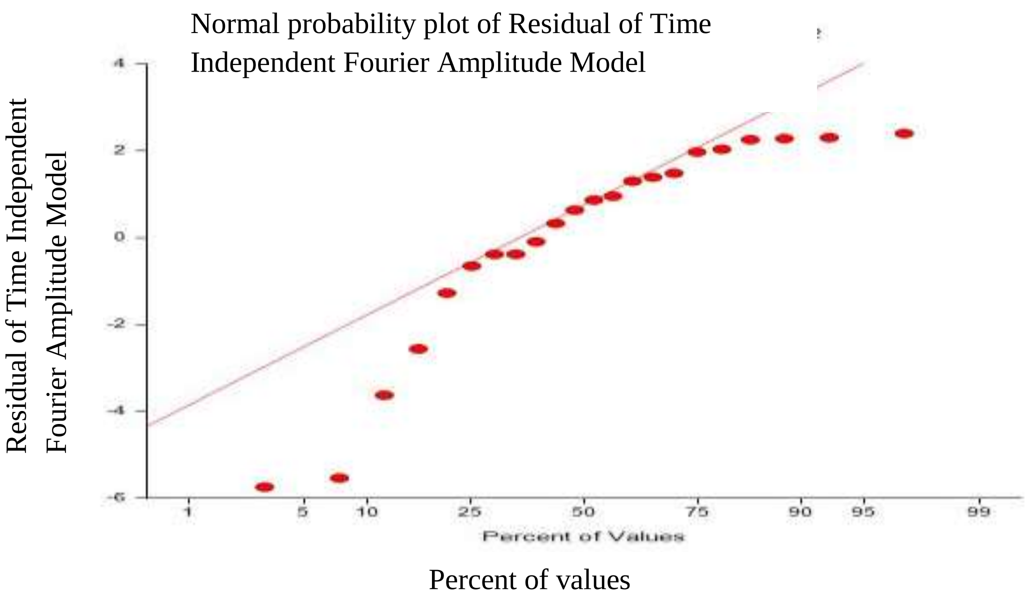

Figure 1: Normal Probability Plot of Error of U.S Population of the Time Independent Fourier Amplitude Method

Table 3 shows the United State census, the predicted values of the Malthusian, Logistics, logistics (Least Squares) and that of the Time Independent Fourier Amplitude model of the United States population. Observe in column 2 (US census (millions)), there was no census in the year 2020 predicted by the United States census bureau but Time Independent Fourier Amplitude, Malthusian, Logistics, Logistics (Least Squares) model predicted these data sets.

Table 3: Predicted values of the U.S. Population using Malthusian, Logistics, Logistics (Least Squares), and Time Independent Fourier Amplitude Model.

Year	U.S Census (millions)	Malthusian	Logistics	Logistics (least squares)	Time Independent Fourier Amplitude Model
1790	3.93	3.93	3.93	4.11	5.450
1800	5.31	5.19	5.30	5.42	7.868
1810	7.24	6.84	7.13	7.14	8.513
1820	9.64	9.03	9.58	9.39	9.732
1830	12.87	11.92	12.82	12.33	11.913
1840	17.07	15.73	17.07	16.14	15.584
1850	23.19	20.76	22.60	21.05	21.216
1860	31.44	27.40	29.70	27.33	29.037
1870	39.82	36.16	38.66	35.28	38.951
1880	50.19	47.72	49.70`	45.21	50.569
1890	62.98	62.98	62.98	57.41	63.362
1900	76.21	83.12	78.42	72.11	76.860
1910	92.23	109.69	95.73	89.37	90.835
1920	106.02	144.76	114.34	109.10	105.384
1930	123.20	191.05	133.48	130.92	120.895
1940	132.16	252.13	152.26	154.20	137.898
1950	151.33	333.74	169.90	178.12	156.862
1960	179.32	439.12	185.76	201.75	178.016
1970	203.30	579.52	199.50	224.21	201.261
1980	266.54	764.80	211.00	244.79	226.210
1990	248.71	1009.33	220.38	263.01	252.333
2000	281.42	1332.03	227.84	278.68	279.162
2010	308.75	1757.91	233.68	281.80	306.466
2020	?	2319.95	238.17	302.66	334.346

Table 4 illustrates the assessments of summary of the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Sum of Square Error (SSE), Mean Error (ME), Mean Absolute Deviation (MAD), and Mean Sum of Square of Error (MSSE) was shown and results shows that the Time Independent Fourier Amplitude Model has a lowest of errors compared to the Malthusian, Logistics, Logistics (Least Squares) model, also the Time Independent Fourier Amplitude Model has the highest value of coefficients of determination (R^2) of 0.9993 showing that out of 100% it's accounted for 99.93% of the consistency of the population leaving only 0.07% to coincidental and the higher the R^2 the better the model fits the data set. This determines that Time Independent Fourier Amplitude Model outperforms the other models and it has an improved fit in the statistics and better accuracy in forecasting the United State population.

Table 4: Summary of Data Analysis (Descriptive statistics of the results)

S/No.	Items	Malthusian	Logistics	Logistics (Least Squares)	Time Independent Fourier Amplitude Model
1	**SSE**	3795	4335357	10572.77	123.04
2	**ME**	-220.54	5.39	3.66	5.45×10^{-6}
3	**MAD**	222.46	11.71	8.65	1.84
4	**R^2**	0.396	0.972	0.993	0.9993
5	**AIC**	165.712	89.039	83.649	55.98
6	**BIC**	162.521	85.847	80.503	58.071
7	**MSSE**	165	188493.78	459.69	5.35

CONCLUSION

In this study, four models were used to examine the United States Population. Via comparing and investigating the SSE, ME, MAD, R^2, AIC, BIC, MSSE values of each model, the four models comprising of the Malthusian, Logistic, Logistics (least squares) and the Time Independent Fourier Amplitude model were all assumed to predict the United States Population. Nevertheless, based on goodness of fit criteria; SSE, ME, MAD, R^2, AIC, BIC, MSSE values, the Time Independent Fourier Amplitude model provides an improved result in predicting the United States population.

REFERENCES

Lassila et al. . (2014,). “Demographic forecasts and fiscal policy rules.” . *International Journal of Forecasting,.*, vol.30, no.4, pp. 1098-1109.

Malthus. (1798). *An essay on the principle of population; or, a view of its past and present effects on human happiness, with an inquiry into our prospects respecting the future removal or mitigation of the evils which it occasions.* U.S.

Miranda et al (2010). “On the Logistic Modeling and Forecasting of Evolutionary Processes: Application to Human Population Dynamics.”. *Technological Forecasting and Social Change,*, vol. 77, no. 5, pp. 699–711.

Nagel et al (2012). s EIGHT EDITION. In *Fundamentals of Differential Equation* (p. 719). U.S: Pearson.

Verhulst., P. F. (, 1845). “Recherches mathématiques sur la loi d'accroissement de la population.”. *Nouvelle mémoire de l'Academie Royale de Sciences et Belle-Lettres de Bruxelles [i.e. Mémoire Series 2],* , vol. 18, pp. 1–42.

Verhulst., P. F. (, 1847,). “Deuxiéme mémoire sur la loi d'accroissement de la population, . *” Mémoire de l'Academie Royale des Sciences, des Lettres et de Beaux-Arts de Belgique,*, vol. 20 pp. 1–32..