FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 7 No. 4, August, 2023, pp 133 - 138
DOI: https://doi.org/10.33003/fjs-2023-0704-1908

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

***Sameer A. S., Yusuf Isah, Abdullahi Bello**

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

*Corresponding authors' email: Ssameer@fudutsinma.edu.ng

ABSTRACT

This study smeared the Time Dependent Fourier Amplitude Model Approach to forecast the Population of the United States of America from 1790 to 2020 on a 10-year interval using Number Crunches Statistical software (NCSS). Results obtained using this procedure was matched with the results obtained in the other models: Malthusian, Logistics, and Logistics (Least Squares) Model. These models were matched using the goodness of fit (the coefficient of determination (R^2), 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 Dependent Fourier Amplitude Model has the highest R^2 and has the lowest SSE, AIC, BIC, ME, MAD, and MSSE. The normal probability plot of residual also forms a lined pattern. The Time Dependent Fourier Amplitude Model gives a statistically significant development in the data sets as compared to the earlier models and also is a suitable model for forecasting the United States population.

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

INTRODUCTION

In accepting of the world population endlessly developing in the last century, population forecast becomes more and more significant in policy making, economic growth, education, and so on Lassila *et al* (2014). Population is affected by numerous issues such as procedure, economy, and culture. Therefore, it is tough for demographic researchers to explore each factor. Among all philosophies, important data are the significant basis in forecasting. By estimating 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 Square models are frequently used approaches for population forecast and have publicized good performance. The Logistic model was originally 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 analyses 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 congregates to it carrying capacity as seen in the application and results. Apparently, 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.

Sameer et al (2022), Population forecasting is a critical effort to appreciate 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 projecting by taking statistical tendencies from older 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 possessions is concerned after a period of unrestricted population growth, with a decline in a population, the aged of the population is a significant delinquent.

We hypothesized that the Time Dependent Fourier Amplitude ruling is more accurate and genuine. To test this proposition, we obtained the U.S population data from 1970 to 2020 from the US Census Bureau. Then, we engaged the data for numerical verification and to compare the accuracy of the mathematical models of the Malthusian, logistic, Logistics least Square and the Time Dependent Fourier Amplitude law for population predictions. Furthermore, to normalize whether the Time Dependent Fourier Amplitude law gives a statistically significant exactness in the fit as compared to the earlier regulations, 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 demand to verify and support the conclusion that it can be widely used. Regression permits us to calculate how much of that transformation in outcome is due to the predictors we are to investigate it using the Time Dependent Fourier Amplitude model and also runs an actual explanation of the population's growth. As wide-open in this paper, predictions attained via the Time Dependent 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 composite set of questionings that can generate upset deliberations about the value of our life and future.

MATERIALS AND METHODS

A logical function could be represented by means of a series of sine's and cosine's unambiguously by the series. For the Standard Fourier series, the synthesis equation is given by:

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

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

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

If we integrate both sides of Equation (1.0) and device that it's suitable to integrate the series term by term and we obtain.

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

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

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

To decide $\alpha_o$for n≥1 we obtain the product of both sides of equation (1) by $Cos\, mt$ (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 fluctuating 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 distinct valued function which is continuous and has a fixed 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 emerging 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)

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

But the Time Dependent Fourier Amplitude Model equation can be written as:

$$Y_t = A + B \times t + C \times t^2 + \frac{t}{N}\sum_{N=1}^{m}\left(D\, cos\frac{2\pi}{L} + E\, sin\frac{2\pi}{L}\right) \quad (11)$$

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

RESULT AND DISCUSSION

In this work, we used the Time Dependent Fourier Amplitude Model to forecast the population of the United States of America from 1790 to 2020 and beyond. In actual, we shall be connecting these results acquired by the Time Dependent Fourier Amplitude Model and the forecast values of the Malthusian, Logistics and Logistics (Least Squares) Models, which was predicted by aforementioned authors.

Table 1: A comparison of the Malthusian and Logistics Models with U.S Census Data (population is given in millions Nagel *et al* (2012).

Year	U.S Census (millions)	Malthusian	Logistics	Logistics (least squares)
1790	3.93	3.93	3.93	4.11
1800	5.31	5.19	5.30	5.42
1810	7.24	6.84	7.13	7.14
1820	9.64	9.03	9.58	9.39
1830	12.87	11.92	12.82	12.33
1840	17.07	15.73	17.07	16.14
1850	23.19	20.76	22.60	21.05
1860	31.44	27.40	29.70	27.33
1870	39.82	36.16	38.66	35.28
1880	50.19	47.72	49.7`	45.21
1890	62.98	62.98	62.98	57.41
1900	76.21	83.12	78.42	72.11
1910	92.23	109.69	95.73	89.37
1920	106.02	144.76	114.34	109.10
1930	123.20	191.05	133.48	130.92
1940	132.16	252.13	152.26	154.20
1950	151.33	333.74	169.90	178.12
1960	179.32	439.12	185.76	201.75
1970	203.30	579.52	199.50	224.21
1980	266.54	764.80	211.00	244.79
1990	248.71	1009.33	220.38	263.01
2000	281.42	1332.03	227.84	278.68
2010	308.75	1757.91	233.68	281.80
2020	?	2319.95	238.17	302.66

The Time Dependent Fourier Amplitude Model remained used to run a nonlinear regression analysis under 0.05 level of significance using Number Crunches Statistical Software (NCSS). Results achieved from this analysis 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 exactly 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 Dependent Fourier Amplitude Model has value of 0.9997 showing that out of 100% it's accounted for 99.97% of the dependability of the population leaving only 0.13% to unpredicted. Moreover, the higher the R^2 the better the model fits any data set.

Table 2: Model Estimation Section of the Time Dependent Fourier Amplitude Method

PARAMETER NAME	PARAMETER ESTIMATE α	ASYMPTOTIC STANDARD ERROR β	LOWER 95.0% CONFIDENCE LIMIT	UPPER 95.0% CONFIDENCE LIMIT	T-VALUES $\frac{\alpha}{\beta}$
A	6.546	1.672	3.017	10.074	3.914
B	-1.089	0.346	-1.820	-0.358	-3.144
C	0.662	0.015	0.630	0.694	43.895
D	3.974	1.420	0.978	6.970	2.798
E	3.208	1.260	0.548	5.868	2.545

Model:

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

R-Squared: 0.9997

Iterations: 3

Estimated Model:

$$\hat{Y}_t = 6.546 - 1.089 \times t + 0.662 \times t \times t + \frac{t}{23} \times 3.974 \times cos(0.6284) + \frac{t}{23} \times 3.208 \times sin(0.6284)$$

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 Dependent 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 symbolises the probability that a population parameter will descent between a set of values for a certain percentage of times. The values satisfying this interval $a < P < b$ specifies that there's a certainty that the confidence interval contains the correct 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 3: Predicted, Predicted and Residuals Value Section of the Time Dependent Fourier Amplitude Method

ROW NO.	ACTUAL Y_t	PREDICTED x $\hat{Y}_t$	LOWER 95.0% CONFIDENCE LIMIT	UPPER 95.0% CONFIDENCE LIMIT	ERROR
1790	3.93	6.340	0.426	6.467	-0.045
1800	5.31	7.391	0.621	12.061	-1.031
1810	7.24	9.481	1.913	12.867	-0.150
1820	9.64	12.558	4.160	14.798	0.160
1830	12.87	16.799	7.333	17.783	0.311
1840	17.07	22.531	11.626	21.972	0.270
1850	23.19	30.083	17.381	27.681	0.658
1860	31.44	39.601	24.922	35.245	1.356
1870	39.82	50.930	34.378	44.824	0.218
1880	50.19	63.636	45.607	56.252	-0.740
1890	62.98	77.171	58.222	69.050	-0.656
1900	76.21	91.111	71.720	82.621	-0.961
1910	92.23	105.381	85.670	96.552	1.118
1920	106.02	120.340	99.943	110.820	0.638
1930	123.20	136.682	114.860	125.820	2.859
1940	132.16	155.177	131.143	142.221	-4.522
1950	151.33	176.339	149.610	160.744	-3.847
1960	179.32	200.178	170.778	181.901	2.980
1970	203.30	226.128	194.624	205.731	3.121
1980	226.54	253.233	220.581	231.675	0.411
1990	248.71	280.505	247.694	258.772	-4.523
2000	281.42	307.336	274.874	286.136	0.914
2010	308.75	333.786	301.283	313.388	1.414
2020	?	360.623	326.865	340.706	
2030		389.068	352.578	368.668	
2040		420.327	379.977	398.159	
2050		455.102	410.499	430.156	
2060		493.262	444.882	465.323	

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

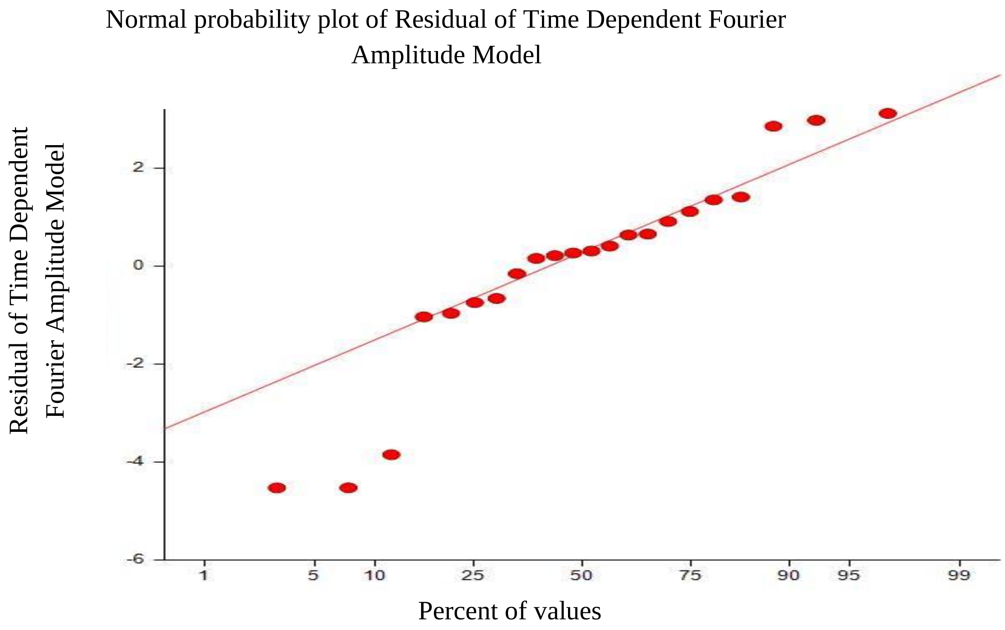

Figure 1: Normal Probability Plot of Error of U.S Population of the Time Dependent Fourier Amplitude Method.

Table 4 shows the United State census, the predicted values of the Malthusian, Logistics, logistics (Least Squares) and that of the Time Dependent 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 Dependent Fourier Amplitude, Malthusian, Logistics, Logistics (Least Squares) model predicted these data sets.

Table 4: Predicted values of the U.S. Population using Malthusian, Logistics, Logistics (Least Squares), and Time Dependent Fourier Amplitude Model.

Year	U.S Census (millions)	Malthusian	Logistics	Logistics (least squares)	Time Dependent Fourier Amplitude Model
1790	3.93	3.93	3.93	4.11	6.341
1800	5.31	5.19	5.30	5.42	7.390
1810	7.24	6.84	7.13	7.14	9.489
1820	9.64	9.03	9.58	9.39	12.568
1830	12.87	11.92	12.82	12.33	16.809
1840	17.07	15.73	17.07	16.14	22.531
1850	23.19	20.76	22.60	21.05	30.083
1860	31.44	27.40	29.70	27.33	39.601
1870	39.82	36.16	38.66	35.28	50.930
1880	50.19	47.72	49.7`	45.21	43.143
1890	62.98	62.98	62.98	57.41	63.631
1900	76.21	83.12	78.42	72.11	77.170
1910	92.23	109.69	95.73	89.37	91.116
1920	106.02	144.76	114.34	109.10	105.381
1930	123.20	191.05	133.48	130.92	120.341
1940	132.16	252.13	152.26	154.20	136.681
1950	151.33	333.74	169.90	178.12	155.180
1960	179.32	439.12	185.76	201.75	176.342
1970	203.30	579.52	199.50	224.21	200.187
1980	266.54	764.80	211.00	244.79	226.139

1990	248.71	1009.33	220.38	263.01	253.238
2000	281.42	1332.03	227.84	278.68	280.518
2010	308.75	1757.91	233.68	281.80	307.343
2020	?	2319.95	238.17	302.66	333.795

Table 5 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 Dependent Fourier Amplitude Model has a lowest of errors compared to the Malthusian, Logistics, Logistics (Least Squares) model, also the Time Dependent Fourier Amplitude Model has the highest value of coefficients of determination (R^2) of 0.9997 showing that out of 100% it's accounted for 99.97% of the consistency of the population leaving only 0.13% to coincidental and the higher the R^2 the better the model fits the data set. This determines that Time Dependent 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 5: Precipitate of Data Investigation (Expressive statistics of the results)

S/No.	Items	Malthusian	Logistics	Logistics (Least Squares)	Time Dependent Fourier Amplitude Model
1	**SSE**	3795	4335357	10572.77	92.69
2	**ME**	-220.54	5.39	3.66	2.0×10^{-6}
3	**MAD**	222.46	11.71	8.65	1.490
4	**R^2**	0.396	0.972	0.993	0.9997
5	**AIC**	165.712	89.039	83.649	53.274
6	**BIC**	162.521	85.847	80.503	52.050
7	**MSSE**	165	188493.78	459.69	4.03

CONCLUSION

In this paper, four models were used to inspect the United States Population. Via associating and exploring 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 Dependent Fourier Amplitude model were all expected to predict the United States Population. However, based on goodness of fit criteria; SSE, ME, MAD, R^2, AIC, BIC, MSSE values, the Time Dependent Fourier Amplitude model provides a better-quality result in forecasting 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.

Sameer et al..(2022). An Application of time independent Fourier amplitude model in forecasting the United States Population,., Vol. 6 No. 1, March, 2022, pp54 - 59

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..