ISSN 1 746-7233, England, UK
World Journal of Modelling and Simulation
Vol. 14 (2018) No. 3, pp. 192-198

Modelling and Estimation of Climatic Variable using Time Series Trigonometric Analysis

T .O. Olatayo,A. I. Taiwo , A. A. Oyewole*

Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye

(*Received April 25 2016, Accepted November 02 2017*)

Abstract. Climate change is a phenomenon that must be carefully and vividly discuss with respect to climatic series fluctuation as this can lead to extreme situations. This article focus on the effect of cyclical or sinusoidal movement in some climatic variables on rainfall as factor that maybe responsible for climate change in south western Nigeria.

The methodology used was Time Series Trigonometric Analysis (TSTA) and the estimation method was based on Ordinary Least Square and Maximum Likelihood. The result revealed that the models has good fit and high predictive power. The Coefficient of determination confirmed the result. The periodic change in evaporation, humidity and temperature contributed to the rapid increase in rainfall over the years which may lead to climate change in south western Nigeria.

In conclusion, this study reveals a continuous cyclical and periodic movement in some climatic variables which may lead to sudden rises in rainfall over the years and proper monitoring of climatic variables must be put in place to checkmate extreme climatic events that can leads to flood, loss of life and properties as a result of climate change.

Keywords: time series trigonometric analysis, periodic and cyclical movement, climatic variables; estimation methods and climatic change.

1 Introduction

In statistics, the traditional or conventional technique of discussing the relationship between two or more variables is the classical regression analysis but when the variables are taken sequentially over a period of time (t) then the classical regression analysis may not be applicable. Hence, in order to analysis such series, we consider a time series regression analysis based on the linear regression model where the assumptions of ordinary least square still hold.

Time series regression model is a technique in time series analysis and it can be a single equation models like autoregressive, autoregressive integrated moving average, seasonal autoregressive integrated moving average models and so on (Box and Jenkins, 1982)[4] while it can be a multiple time series regression equation like vector autoregressive model, time series with lagged explanatory variable and many more (Wei, 2006)[12].

Time series regression models have been widely used in almost all spheres and many applicable areas where the time series observations exhibit any of the four components of time series. When a series exhibit a cyclical variation, the proper model is time series regression and this will be our interest in this article. Hence, we will make use of time series trigonometric analysis model to analyse series with cyclical, sinusoidal or periodic variations with the intention of unveiling hidden periodicities in a given series along with its frequencies. Essentially, different time series regression approaches has been used to discuss and analysis over the years and this can be seen in Adrian and Jessica (2001)[1], Nicholas, (2006)[8], Bhaskaren et al. (2013)[3], Maria et al. (2014)[7], Olatayo et al. (2014)[9] and Spelman et al. (2015)[11]. This research work was intended to be used to

* Corresponding author. *E-mail address*: otimtoy@yahoo.com and ab.com1982@yahoo.com

discuss and analysis the recent climate change in our environment that can be attributed to climate and weather variables.

The theoretical background of time series trigonometric analysis can be seen in Bliss (1958), where a sine wave model was proposed and analysis of variance was used to determine how many terms to be retained in the sine wave model. Harvey et al. (1987)[6] went ahead to propose a periodic regression that allowed the cyclical components to be modelled explicitly and stochastically. The work of Rawlings et al. (1998)[10] brought to light the comprehensive theoretical background of trigonometric regression. This method was discussed as a special case of linear regression analysis and it included a dependent variable, trigonometric functions such as sin(t) and cos(t) are periodic over time with a period of 2/, independent variable and trigonometric coefficients were given as $\beta_j sin(2\pi t/\omega)$ and $\beta_j^* cos(2\pi t/\omega)$ where $j = 1, 2,$ and $t = 1, 2, t-1$ and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is the error term. The coefficients were estimated based on ordinary least square approach and the interpretation of the coefficients β_j and β_j^* in terms of the phase angle of the trend and the period were discussed in Anderson (1971)[2].

In this article, we propose a Time Series Trigonometric Analysis with multiple independent variables as an improvement on the work of Rawlings et al. (1998)[10]. Coefficients estimation will be done using maximum likelihood and ordinary least estimation methods. The hypothesis testing of the trigonometric coefficients will be based on t and F-statistics.

The proposed method will be applied to climatic variables as a way of studying and understanding the climate change in south western, Nigeria. The climatic variables considered were rainfall, temperature, evaporation and humidity and their yearly series were obtained from Nigerian Meteorological Agency (NMA), for South West, Nigeria from 1984 to 2015 and article organization were in sections; section is an introduction, section two is theory and methods, section three is results and discussions while section four is conclusion of the research article respectively.

2 Theory and methods

2.1 Time series trigonometric analysis and estimation methods

Given a simple deterministic model as

$$y = \rho \cos(\omega t - \theta), \tag{1}$$

where ρ is the amplitude, ω is the frequency and θ is the phase.

Using the compound angle formula

$$\cos(A - B) = cosAcosB + sinAsinB.$$

We rewrite (1) as

$$\begin{aligned} y &= \rho \cos\theta cos(\omega t) + \rho \sin sin(\omega t), \\ y &= \beta \cos(\omega t) + \beta^* \sin(\omega t), \end{aligned} \tag{2}$$

where $\beta = \rho cos\theta$, $\beta^* = \rho sin\theta$ and $\beta^2 + \beta^{*2} = \rho^2$

Based on (2) a time series trigonometric model can be expressed

$$y_t = \beta_0 + \sum_{j=0}^{p} \beta_j \cos(\omega_j t) + \sum_{j=0}^{p} \beta_j^* \sin(\omega_j t) + \varepsilon_t, \quad \begin{matrix} j = 1, \ldots, p \\ t = 1, \ldots, T \end{matrix} \tag{3}$$

where $sin\omega$ and $cos\omega$ are the trigonometric function, $\omega = 2\pi k/n$, X_t are the predictors, y_t is the dependent variable, $\beta_j's$ are the trigonometric coefficients and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is the uncorrelated error term.

By substituting $t = X_t$ in (3), the multiple form of trigonometric regression becomes

$$y_t = \beta_0 + \sum_{j=1}^{p} \beta_j cos\omega X_t + \sum_{j=1}^{p} \beta_j^* sin\omega X_t + \varepsilon_t, \quad \begin{array}{l} j = 1, \dots, p, \\ t = 1, \dots, T, \end{array} \tag{4}$$

where X_t are the predictors, y_t is the dependent variable and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ is the error term

2.1.1 Ordinary least squares estimation method

Based on the trigonometric regression model in (4)

$$\begin{gathered} y_t = \beta_0 + \beta_j cos\omega X_t + \beta_j^* sin\omega X_t + \varepsilon_t, \\ \varepsilon_t = y_t - \beta_0 - \beta_j cos\omega X_t - \beta_j^* sin\omega X_t. \end{gathered}$$

The residual sum of square becomes

$$\sum^{(\varepsilon_t)^2} = \sum^{(y_t - \beta_0 - \beta_j \cos\omega X_t - \beta_j^* \sin\omega X_t)^2} \tag{5}$$

Differentiating (5) with respect to β_0, β_j and β_j^* gives

$$\left. \begin{array}{c} \Sigma y_t = n\beta_0 + \beta_j \Sigma \cos\omega X_t + \beta_j^* \Sigma \sin\omega X_t \\ \Sigma\left(\cos\omega X_t y_t\right) = \beta_0 \Sigma\left(\cos\omega X_t\right) + \beta_j \Sigma\left(cos^2\omega X_t\right) + \beta_j^* \Sigma(\cos\omega X_t \sin\omega X_t) \\ \Sigma\left(\sin\omega X_t y_t\right) = \beta_0 \Sigma\left(\sin\omega X_t\right) + \beta_j \Sigma\left(\cos\omega X_t \sin\omega X_t\right) + \beta_j^* \Sigma sin^2\omega X_t) \end{array} \right\} \tag{6}$$

Transforming (6) into matrix form gives

$$\begin{pmatrix} \sum^(y_t) \\ \sum^(\cos\omega X_t y_t) \\ \sum^(\sin\omega X_t y_t) \end{pmatrix} = \begin{pmatrix} n & \sum^(\cos\omega X_t) & \sum^(\sin\omega X_t) \\ \sum^(\cos\omega X_t) & \sum^(\cos^2\omega X_t) & \sum^(\cos\omega t \sin\omega X_t) \\ \sum^(\sin\omega X_t) & \sum^(\cos\omega X_t \sin\omega) X_t & \sum^(\sin^2\omega X_t) \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_j \\ \beta_j^* \end{pmatrix}. \tag{7}$$

2.1.2 Maximum likelihood estimation method

If the error random variables (ε_t) are assumed to be independent and normally distributed with mean zero, then the maximum likelihood estimates is obtained as follows

Based on

$$\varepsilon_t = y_t - \beta_0 - \beta_j cos\omega X_t - \beta_j^* sin\omega X_t$$

the likelihood function with respect to the error term becomes

$$\begin{aligned} & L = (\sigma\sqrt{2\pi})^{-1} e^{-\frac{1}{2}\left(\frac{y_1 - \beta_o - \beta_j \cos\omega X_1 - \beta_j^* \sin\omega X_1}{\sigma}\right)^2} \\ & \times(\sigma\sqrt{2\pi})^{-1} e^{-\frac{1}{2}\left(\frac{y_2 - \beta_o - \beta_j \cos\omega X_2 - \beta_j^* \sin\omega X_2}{\sigma}\right)^2} \\ & \times \cdots \times (\sigma\sqrt{2\pi})^{-1} e^{-\frac{1}{2}\left(\frac{y_m - \beta_o - \beta_j \cos\omega X_m - \beta_j^* \sin\omega X_m}{\sigma}\right)^2}. \end{aligned} \tag{8}$$

Take the log, linearize and differentiate (8) with respect to β_0, β_j and β_j^* gives

$$\hat{\beta}_0 = {}^{\Sigma y_t}\!/_{m} - \hat{\beta}_j{}^{\Sigma \cos! X_t}\!/_{m} - \hat{\beta}_j^{*}{}^{\Sigma \sin! X_t}\!/_{m}. \tag{9}$$

$$\hat{\beta}_j = \sum{}^{(} cos\omega X_t) y_t - \hat{\beta}_0 \sum{}^{c} os\omega X_t - \hat{\beta}_j^* \sum{}^{c} os\omega X_t sin\omega X_t \Big/ {\textstyle\sum}^{(cos\omega X_t)^2}. \tag{10}$$

$$\hat{\beta}_j^* = \sum{}^{(} sin\omega X_t) y_t - \hat{\beta}_0 \sum{}^{s} in\omega X_t - \hat{\beta}_j^* \sum{}^{c} os\omega X_t sin\omega X_t \Big/ {\textstyle\sum}^{(sin\omega X_t)^2}. \tag{11}$$

The results in equations (9), (10), and (11) gives the parameters estimates using Maximum likelihood estimation method.

2.1.3 Error term performance based on durbin watson statistic

If ε_t is the residual associated with the observation at time t, then the test statistic is

$$d = \frac{\sum_{t=2}^{T} \left(\varepsilon_t - \varepsilon_{t-1}\right)^2}{\sum_{t=1}^{T} \varepsilon_t^2}, \tag{12}$$

where T is the number of observations. Note that if the sample is lengthy, then this can be linearly mapped to the Pearson correlation of the time-series data with its lags (Durbin and Watson, 1950)[5].

2.1.4 Test of hypothesis

In order to perform test of hypothesis regarding the significance of individual parameters that is to test the hypothesis $H_0 : \beta_0 = 0$ vs $H_1 : \beta_0 0$ and others, we will used the t-statistic given as

$$t = \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\frac{\hat{\sigma}^2}{n}}}, \tag{13}$$

and reject $H_0 : \beta_0 = 0$ if $|t| > t(\alpha/2, n-1)$

Similarly to test the null hypothesis$H_0 : \theta_i = 0$ vs $H_1 : \theta_i 0$ where $i = 1, 2, 3, .$ We will use F-statistic given as

$$F = \frac{MS(Reg)}{MS(RES)}, \tag{14}$$

and reject H_0 if $f_{cal} > f_{\alpha, r_1, r_2}$ where r_1 = degree of freedom of regression and r_2 = degree of freedom of residual

3 Result and discussion

3.1 Time plot

The climatic data used in this research work was obtained from Nigerian Meteorological Agency, south western, Nigeria from 1984 to 2015. The climatic variables are rainfall, temperature, humidity and evaporation. The time plot of the climatic series in Fig. 1 below exhibit a cyclical movement or variation over the years for all the series under consideration. Hence, the behaviour of the climatic series is periodic and this can be attributed to several factors like natural occurrence, technology and many more.

3.2 Time series trigonometric analysis estimation

The time series trigonometric model fitted for the climatic variables is

$$y_t = \beta_0 + \sum_{j=1}^{3} \alpha_j cos\frac{\pi}{12}X_t + \sum_{j=1}^{3} \beta_j sin\frac{\pi}{12}X_t + \varepsilon_t, \qquad (15)$$

where the period is $2\Pi/24 = \Pi/12$.

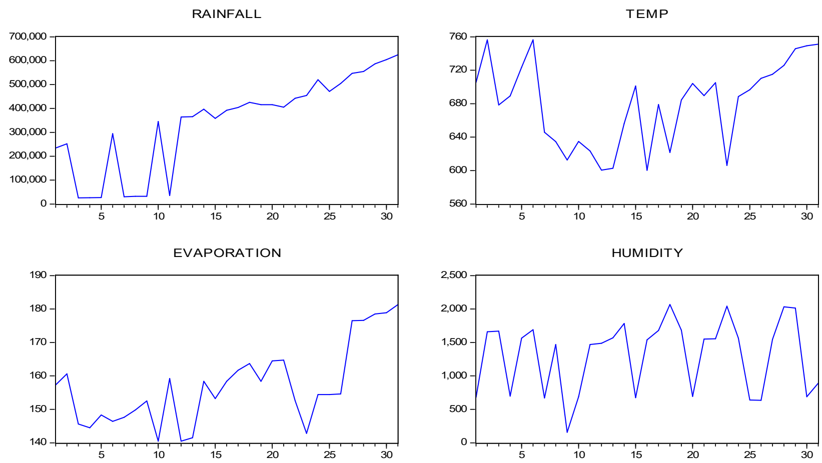

Fig. 1: Yearly rainfall, teperature,evaporation and humidity series from 1984-2015

The time series trigonometric analysis equation using OLS obtained was: Rainfall = 77.23982 + $8492.3cos\frac{\pi}{12}t_{temp} + 0.6sin\frac{\pi}{12}t_{temp} + 2209.5cos\frac{\pi}{12}t_{evap} + 254.4cos\frac{\pi}{12}t_{evap} + 426.1sin\frac{\pi}{12}t_{hum} + \varepsilon_t$

with $R^2 = 0.9584$, Adjusted $R^2 = 0.9447$, MSE= 3.114 and Durbin Watson= 1.8564

The time series trigonometric analysis equation using MLE obtained was:

Rainfall = $42.38982 - 15.57117cos\frac{\pi}{12}t_{temp} + 0.92457sin\frac{\pi}{12}t_{temp} + 7.62201cos\frac{\pi}{12}t_{evap} - 2.29061sin\frac{\pi}{12}t_{evap} - 4.00436cos\frac{\pi}{12}t_{hum} + 3.65274sin\frac{\pi}{12}t_{evap} - 4.00436cos\frac{\pi}{12}t_{hum} + \varepsilon_t$

with $R^2 = 0.9624$, Adjusted $R^2 = 0.9547$, MSE= 0.9672 and Durbin Watson= 1.8564.

3.2.1 Interpretation of the time series trigonometric analysis using ols and ml

The values of the Durbin Watson statistics for both models showed the error term were not serially correlated. The time series trigonometric analysis equation for OLS showed that the coefficient of $cost_{temp}, sint_{temp}, sint_{evap}, cost_{hum}, cost_{evap}$ and $sint_{hum}$ indicated that for every increase in temperature, evaporation and humidity there is an increase in the millimeter of rainfall yearly. The values of coefficient of determination (R^2) revealed that temperature, evaporation and humidity explained the variations in rainfall up to 95%. The value of adjusted coefficient of determination ($\bar{R}^2$) also revealed that the model is a good fit and has a high predictive power. While the time series trigonometric analysis equation for MLE revealed that the coefficient of $sint_{temp}, cost_{evap}, sint_{evap}$ and $sint_{hum}$ indicated that for every increase in temperature, evaporation and humidity there is an increase in the millimeter of rainfall over the years while the coefficients of $cost_{temp}$ and $cost_{hum}$ revealed negative and a unit change will reduce rainfall over the years. The values of coefficient of determination (R^2) revealed that temperature, evaporation and humidity explained the variation in rainfall up to 96%. Adjusted coefficient of determination ($\bar{R}^2$) value also revealed that the model is a good fit and has a high predictive power. Based on the mean square error value of the maximum likelihood estimation, the method is better than ordinary least square estimation method for analysis periodic climatic data.

3.3 Test of hypothesis

$H_0 : \beta_0 = 0$ vs$H_1 : \beta_0 0$

Using 0.05 level of significance

$t_{(0.025,30)} = 2.042$

Since all $t_{cal} > t_{cal}$, we reject the null hypothesis and concluded that all β_0 are statistically different to zero at 0.05 level of significance.

For $H_0 : \beta_j = 0$ vs $H_1 : \beta_j 0$

$t_{(0.025,30)} = 2.042$

Since all $t_{cal} > t_{tal}$, we reject the null hypothesis and concluded that all β_j are individually statistically different to zero at 0.05 level of significance

For $H_0 : \beta*_j = 0$ vs $H_1 : \beta *_j 0$

$t_{(0.025,30)} = 2.042$

Since all $t_{cal} > t_{tal}$, we reject the null hypothesis and concluded that all $\beta*_j$ are individually statistically different to zero at 0.05 level of significance

$H_0 : \beta_i = 0$ vs $H_1 : \beta_i 0$

$$F = \frac{MS(Reg)}{MS(RES)},$$

$f_{(0.05,6,24)} = 2.99$

$$f_{cal} = \frac{MS(Reg)}{MS(RES)},$$

$f_{cal} = 3.114$

$f_{cal} = 3.114 > f_{tab} = 2.99$

Since $f_{cal} > f_{cal}, werejectH_0$ and concluded that all the coefficients are jointly statistically different from zero at 0.05 percent level of significance.

4 Conclusion

In conclusion, climate and rainfall are highly non-linear and complicated phenomena which require detailed models that will capture its sinusoidal or periodic variations with the intention of unveiling hidden periodicities in a given series along with its frequencies. Rainfall is a climate parameter that affects the way and manner men lives and their ecological systems. This study was used to propose a Time Series Trigonometric model which is capable of analysing the effects of two or more periodic independent variables on a periodic dependent variable. Estimation methods were ordinary least square and maximum likelihood. The steps for test of hypothesis was as well derived. Time Series Trigonometric model efficiency was tested by analysing yearly climatic data series (Temperature, Evaporation, Humidity and Rainfall) of South Western, Nigeria from 1984 C 2015. Based on the overall results, the time series trigonometric model obtained revealed that the coefficients of $sint_{temp}, cost_{Evap}$, and $sint_{hum}$ indicated that for every increase in temperature, evaporation and humidity there is an increase in the millimeter of rainfall annually, while the coefficients of $cost_{temp}, sint_{evap}$ and $cost_{hum}$ revealed a negative and a unit change will reduce the annual amount of rainfall. The value of R^2 showed that 96% of the variation in rainfall were jointly explained by temperature, evaporation and humidity. The value of Adjusted R^2 shows the model is a good fit and has a high predictive power.

Finally, this article revealed the relationship between rainfall and other variables of climate change which could help in understanding the current and future climatic variables complexity which could lead to climatic events such as flood, loss of life and properties.

References

[1] Adrian M. H., Jessica J.M. Detecting lunar cycles in marine ecology: periodic regression versus categorical anova. *Marine Ecology Progress*, 2001, **214**(1): 307–310.

[2] Anderson, T. W. *The Statistical Analysis of Time Series. by T. W. Anderson*. Wiley, 1971.

[3] Bhaskaran K., Gasparrini A. et al. Time series regression studies in environmental epidemiology. *International Journal of Epidemiology*, 2013, **42**(4): 1187–1195.

[4] Box G.E.P., Jenkins G.M. *Time Series Analysis: Forecasting and Control*. Holden day, 1982.

[5] Durbin J., Watson G.S. Testing for serial correlation in least squares regression: I. *Biometrika*, 1950, **37**(3-4): 409–428.

[6] Harvey, A. C., Souza, R.C. Assessing and modeling the cyclical behavior of rainfall in northeast brazil. *Journal of Applied Meteorology*, 2010, **26**(10): 1339–1344.

[7] Maria, P. F., Alexander, V. I. et al. Detecting hidden periodicities for models with cyclical errors. *Statistics & Its Interface*, 2014, **10**(1): 1–15.

[8] Nicholas, J. C. In praise of trigonometric predictors. *The Stata Journal*, 2006, **6**(4): 561–579.

[9] Olatayo T.O., Taiwo A. I., Lasisi T. A. Population prediction modelling with autoregressive integrated moving average (arima) model. *Journal of the Nigerian Association of Mathematical Physics*, 2014, **27**(1): 189–200.

[10] Rawlings J. O., Pantula S. G., Dickey D. A. *Applied Regression Analysis*. Springer-Verlag, 1998.

[11] Spelman T., Gray O. et al. Method of trigonometric modeling of seasonal variation demostrated with multiple sclerosis relapse data. *Journal of Visualized Experiment*, 2015, 1–20.

[12] Wei, W. S. *Time Series Analysis Univariate and Multivariate Methods*. Pearson Addison Welsley, 2006.