Transactions of the Nigerian Association of Mathematical Physics
***Volume 13, (October – December, 2020 Issue*), pp101 –106**
© Trans. of NAMP

SEASONAL MODELLING OF AKWA IBOM STATE RELATIVE HUMIDITY USING FOURIER SERIES ANALYSIS

Udoh Grace Monday

Department of Statistics, Akwa Ibom State Polytechnic

Abstract

The focus of this work was to model a periodic time series function. The pure Fourier series analysis was used to model the data. Seasonal period with the highest intensity was obtained using periodogram, then Fourier method to robust regression was used to fit sine and cosine coefficients at each Fourier transform of that period. A stepwise regression was used to select the significant orthogonal trigonometric functions. This was applied to a relative humidity data. In this representation, the time series Y_t is expressed as a combination of the intercept and the orthogonal trigonometric functions. The Diagnostic checks showed that the model is adequate. This is similar to what Ekpenyong E.J., (2018) obtained without the stepwise regression, but with the introduction of the stepwise repression in this study the residual variance reduced and gave a better fit. This model can be used to make future forecast of relative humidity of the state but with effect of human activity on climate. The model can only be used for short term forecast.

Keywords: Fourier series, periodogram, seasonal period, spectral density, robust regression and diagnostic checks

INTRODUCTION

Relative humidity (RH) is the ratio of the current absolute humidity to the highest possible absolute humidity (which depends on the current air temperature) or we can say is the ratio of actual vapour pressure to saturated vapour pressure multiplied by 100. A reading of 100 percent RH means that the air is totally saturated with water vapour and cannot hold any more, creating the possibility of rain. This doesn't mean that the RH must be 100 percent in order for it to rain — it must be 100 percent where the clouds are forming, but the RH near the ground could be much less. RH is strongly proportional to temperature and highly sensitive to temperature changes. This means if you have a stable temperature in your system, your RH will also be stable. Humans are very sensitive to humidity, as the skin relies on the air to get rid of moisture. The process of sweating is your body's attempt to keep cool and maintain its current temperature. If the air is at 100 percent RH, sweat will not evaporate into the air. As a result, we feel much hotter than the actual temperature when the relative humidity is high. If the RH is low, we can feel much cooler than the actual temperature because our sweat evaporates easily, cooling us off. Humidity affects your health, home, and your sanity, you need insights on the types of humidity and this calls for forecast record. The time series analysis is used in this article to model and to obtain a short term forecast of the RH of Akwa Ibom state.

There are two approaches to time series analysis; however these approaches are separate but mutually exclusive they are: the time domain approach and the frequency domain approach (Hamilton, J. D.1994). The usual autoregressive integrated moving average (ARIMA) models developed by Box and Jenkins (1970) has been extensively used in modeling linear time series. The ARIMA models assume that the current observation depends on weighted previous observations, weighted previous random shocks and the current shock. However, most time series arising in nature do not assume linearity but rather, periodic or seasonal with linear trend. Time domain approach, studies temporal properties of the realizations of a variable, recorded at a predetermined frequency and does not relate any information concerning frequency components of the variable. On the other hand, the frequency approach sees time series analysis as involving the periodic or systematic sinusoidal variation which is found naturally in most data, with the breakdown of the various kinds of periodic variation usually achieved through periodogram. Many business, economic and weather time series are discrete data that contain a seasonal phenomenon that repeats itself after a regular period of time. The smallest time period for this repetitive phenomenon is called the seasonal period (Wei 2006). A periodic time series fluctuates in some form but maintains steady values and is not obviously steadily increasing or decreasing. This means that the series repeats itself after certain intervals (Priestly 1981).

Many French mathematicians derived famous set of orthogonal basis functions in vector analysis so that every vector in the space can be written as linear combinations of the basis, but Fourier noticed that the cosines and sine made up a set of

Correspondence Author: Udoh G.M., Email: graceudoh14@gmail.com, Tel: +2348089107079

orthogonal functions capable of representing periodic functions, no wonder Ekpeyong (2008) used pseudo additive(mixed) Fourier series analysis to model rainfall data, Ekpeyong and Omekara (2007) applied fourier series analysis to model temperature data. Also, Nachane and Clavel (2014) modeled interest rate using the mixed Fourier series, significant Fourier series model comprising the trend, seasonal and error components termed outliers were fitted to the data. Konarasinghe and Abeynayake (2015) decided to focus on applying Fourier transformation on the model fitting for Stock Returns of Sri Lanke share market and compared the forecasting ability of the above model with the Autoregressive Integrated Moving Average (ARIMA) model. Masset, P. (2008) also presented a set of tools, which allows gathering information about the frequency components. He discussed Spectral Analysis and Filtering methods and pointed out that Spectral and Filter methods have two drawbacks as mentioned in the introduction and that Wavelet analysis which take its root from Spectral and Filter methods can be used to overcome this drawbacks. In the same vein Rahman, M. M., Harun-or-Roshid, Mozid pk., M. A. and Mamun, M. A. A. (2011) carried out a comparative analysis of Fourier transform and Wavelet transform. Perhaps it was such finding which encouraged Mukhopadhyay, S., Dash, D., Mitra, A. and Bhattaacharya, P. (2013) to adopt Fourier and Wavelet transform to study wind speed data. Ribeiro, M. K., Braga Júnior, R. A., Sáfafi, T. and Horgan, G. (2013), in their study attempted to compare the constituents parts of the speckle signal according to Fourier and Wavelet transform for numerical analysis.

In this light this paper seeks to model and give a short term forecast of the relative humidity of Akwa Ibom state using the Fourier series analysis. The aim is to completely reconstruct the relative data humidity of Akwa Ibom state with little error using a small number of sine and cosine functions.

METHODOLOGY.

Periodogram

The notion of any time series, from a heavily patterned process to a random white-noise process, represented by series of weighted, orthogonal, sinusoidal (i.e, sine and cosine) function $\{Y_t\}$ is refered to as a frequency domain time series. Where t = 0, 1, 2, ---, (N-1) is expressible as:

$$Y(t)=\sum_{i=0}^{n/2}(a_i\cos\omega_i t+b_i\sin\omega_i t) \qquad (1)$$

The expressions a_i and b_i are found by using the robust regression to fit sine and cosine coefficient at each Fourier transform, Y_t are taken at equally spaced intervals t. The off diagonal terms become zero, where as the diagonal terms become a single function of t, the coefficient β_0 does not exist , α_0 is found to be $\alpha_0=\frac{\sum Y_i}{N}$.

$\omega_i=\frac{2\pi i}{N}$ $\quad n=0,1,2,\dots,\left(\frac{N-1}{2}\right)$ are Fourier frequencies, $\mathrm{f_i}=\frac{n}{\mathrm{N}}$ and frequency correspond to a period or season, $s=n/i$

$$\hat{a}_i=\frac{2}{N}\sum_{i=0}^{n}Y_t\cos\omega_i t \qquad (2)$$

$$\hat{b}_i=\frac{2}{N}\sum_{i=0}^{n}Y_t\sin\omega_i t \qquad (3)$$

$a,\ b\ and\ \omega$ are sets of sinusoids' amplitudes and frequencies weights respectively, the sets of discrete Fourier transform (DFT) weights obtained by using equation (2) and (3) are set of orthogonal functions, the sum of these N-1 sine and cosine coefficients represents the data so that the number of functions is totally equal to the number of actual data, so the aim is to completely reconstruct the data with little error using a small number of functions. The amplitudes or amplitude squared indicates the strength, of a given pattern in the data, and this strength is known as the power, energy or intensity of the series.

The power represented by

$$\sum Y_i^2=\begin{cases} Na_0^2+\frac{N}{2}\sum_{n=1}^{\frac{N-1}{2}}(a_n^2+b_n^2)\ if\ N\ is\ odd \\ Na_0^2+\frac{N}{2}\sum_{n=1}^{\frac{N-1}{2}}(a_n^2+b_n^2)+Na_{N/2}^2 if\ N\ is\ even \end{cases} \qquad (4)$$

is known as parseval's relation, when Fourier series is plotted against the Fourier frequencies the plot is called a periodogram. Peaks in that periodogram indicates the most important pattern of different frequencies within the data.

Test for the Maximum Periodic Ordinates

The natural test statistic will be

$\mathbf{I}_{(\omega_1)}=\mathrm{Max}[\mathbf{I}_{(\omega_K)}]$

"Where I indicates the Fourier frequency with maximum periodogram ordinate"

The Fourier Series Model

Suppose kth ordinate was identified at frequencies N then the general model is

$$\mathrm{Y}(t)=\sum_{i=0}^{k/2}(\alpha_i\cos \mathrm{i}\omega\, t+\beta_i\,\mathrm{sini}\,\omega\mathrm{t})+\varepsilon_\mathrm{t} \qquad (5)$$

The estimated model is of the form

$$\hat{Y}(t) = \sum_{i=0}^{k}\left(\hat{\alpha}_i \cos i\omega t + \hat{\beta}_i \sin i\omega t\right) \quad (6)$$

$\hat{Y}(t)$= estimated values

$\hat{\alpha}_i, \hat{\beta}_i (i = 0,1,\ldots,k/2)$ = parameter estimates

$\omega = {}^{2\pi}/_{k}$

$k =$ is the significant harmonic of ω

ε_t is a random component.

Stepwise Selection of Terms

Stepwise regression is a way of selecting important variables to get a simple and easily interpretable model.
Forward stepwise is a variable selection method which

1. Begins with a model that contains no variable
2. Then starts adding the most significant variables one after the other
3. Until a pre-specified stopping rule is reached or until all the variables under consideration are included in the model.

Data Analysis

The raw data plot in Figure 1 clearly shows the existence of seasonality. This is indicated by the periodic pattern and upward and downward movement of the graph.

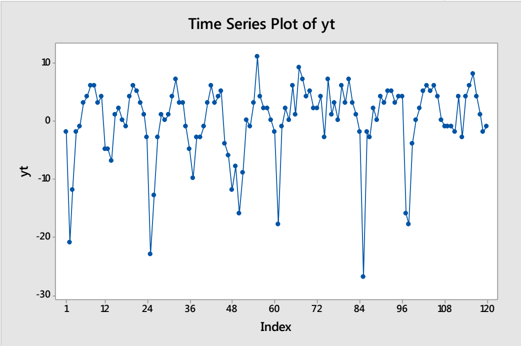

Figure 1. Raw data plot of the series Y_t

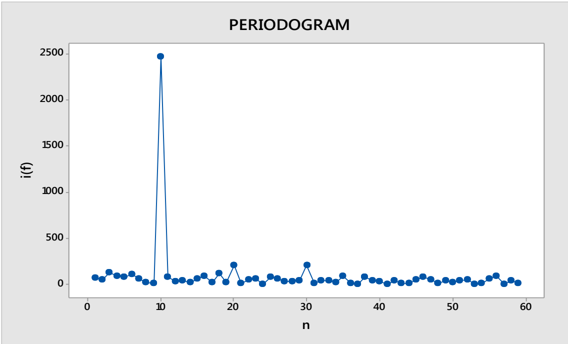

Fig.2 Periodogram Plot

Periodogram

The periodogram analysis was conducted using excel and the plot by Minitab. The periodogram is displayed in figure 2. The values for the spectral densities, periods and frequencies are equally displayed in Table 1. From the periodogram it was observed that the frequency and the period with the largest intensity are 0.0833 and 12 months respectively This period was used to fit the Fourier Series model. The expressions a_i and b_i are found by using the robust regression to fit sine and cosine coefficient at each Fourier transform, Y_t are taken at equally spaced intervals t. The off diagonal terms become zero, where as the diagonal terms become a single function of t, the coefficient β_0 does not exist ,

The Seasonal Fourier Representation of $\mathbf{Y}_t$

The seasonal component is then estimated from the time series data as:

$$k = \frac{n}{2} = \frac{12}{2} = 6$$

And

$$\omega = 2 \times \pi \times 0.0833$$

$$\therefore Y_t = \sum_{i=0}^{6} [\alpha_i \cos i\omega t + \beta_i sin i\omega t]$$

The parameter $\alpha_i{}'^s and \beta_i{}'^s$ are obtained by method of least squares as seen in table 1

Table 1 Test for Significance of the parameter Estimates in the Seasonal Components

Term	Coef	SE Coef	T-Value	P-Value
Constant	**80.858**	**0.418**	**193.63**	**0.000**
Cos2wt	-4.362	0.591	-7.38	0.000
Cos3wt	-0.692	0.591	-1.17	0.244
Cos4wt	1.159	0.591	1.96	0.052
Cos5wt	0.658	0.591	1.11	0.268
Cos6wt	0.500	0.590	0.85	0.399
Sinwt	-0.725	0.843	-0.86	0.392
Sin2wt	-4.737	0.590	-8.02	0.000
Sin3wt	-1.755	0.590	-2.97	0.004
Sin4wt	-1.424	0.591	-2.41	0.018
Sin5wt	-0.085	0.591	-0.14	0.886
Sin6wt	-0.357	0.593	-0.60	0.548

4.57387

Stepwise Selection of Terms

Coefficients

Term	Coef	SE Coef	T-Value	P-Value
Constant	80.865	0.414	195.21	0.000
Cos2wt	-4.347	0.586	-7.42	0.000
Cos4wt	1.174	0.586	2.00	0.047
Sin2wt	-4.733	0.586	-8.08	0.000
Sin3wt	-1.747	0.586	-2.98	0.003
Sin4wt	-1.410	0.586	-2.41	0.018

4.53772

Regression Equation

$\hat{Y}_t$ = 80.865 - 4.347cos2wt + 1.174cos4wt - 4.733sin2wt - 1.747sin3wt – 1.410sin4wt

5. Diagnosis

Two diagnostic checks are presented here to ensure that the model fitted to the data is adequate.

5.1 Actual and Estimate Plots

The overlaid plots of the actual values (Y_t) and the estimated values ($\hat{Y}_t$) are displayed in figure 3. The two superimposed plots move together in the same direction indicating closeness and strong correlation between the values of the two variables. This shows that the model is adequate.

Residual variance of the fitted model is 4.53772. This is smaller than 4.57387 obtained without the stepwise selection. The residual autocorrelation function is displayed in figure 4 and it shows that there is no significant autocorrelation of the residuals. That is, all autocorrelations at all lags are within the range $\pm 2/\sqrt{n}$ (as indicated by the two black lines in figure 4). This means the residuals of the fitted model are not serially correlated. In more precise terms, the residuals follow a white noise process. Hence, the model is adequate.

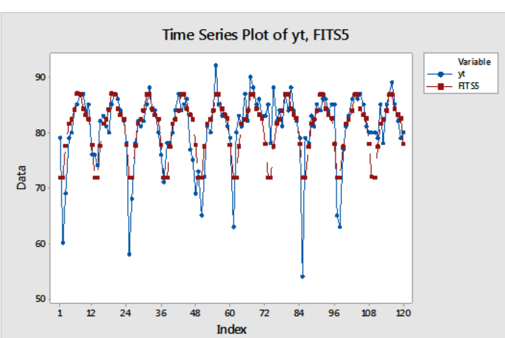

Fig. 3 Time Plot for Raw Data and Fits

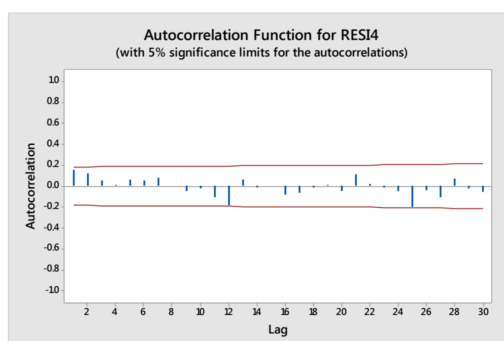

Fig. 4 Autocorrelation Function Plot for Residuals.

Table 2 Period and frequency

Col1	Col2	Col3	Col4	Col5	Col6
a	b	a^{**2}	b^{**2}	Intensity i(f)	Period
-6.50489	-64.1338	42.31366	4113.14	69.25756	120
28.59567	46.33261	817.7125	2146.711	49.40706	60
-4.84907	-85.5203	23.51349	7313.716	122.2872	40
-58.2172	41.82522	3389.237	1749.349	85.6431	30
-53.8265	-39.5065	2897.294	1560.763	74.30095	24
-31.0676	-72.8285	965.1983	5303.992	104.4865	20
-28.6988	48.48465	823.6201	2350.761	52.90635	17.14286
25.69955	-24.7932	660.467	614.7048	21.25286	15
21.12754	-7.70981	446.3728	59.44114	8.430232	13.33333
-256.208	-286.761	65642.54	82231.83	2464.573	12
36.53977	-59.009	1335.155	3482.057	80.28687	10.90909
3.305347	41.42103	10.92532	1715.702	28.77712	10
47.58283	-16.8284	2264.125	283.1936	42.45532	9.230769
-25.1533	-22.3685	632.6883	500.3513	18.88399	8.571429
34.44756	-45.2718	1186.634	2049.539	53.93622	8
11.42393	-69.6838	130.5061	4855.834	83.10567	7.5
-22.2063	18.19364	493.1192	331.0085	13.73546	7.058824
18.35188	82.60837	336.7913	6824.142	119.3489	6.666667
29.96969	1.891973	898.1823	3.579562	15.02937	6.315789
-37.1809	-105.398	1382.416	11108.65	208.1844	6
12.44475	-15.8593	154.8718	251.5177	6.773159	5.714286
45.95334	22.91653	2111.709	525.1672	43.94794	5.454545
32.87301	46.79166	1080.635	2189.46	54.50157	5.217391
14.86239	-1.5383	220.8906	2.366373	3.72095	5
66.40612	16.50098	4409.772	272.2822	78.03424	4.8
44.25417	-39.491	1958.431	1559.539	58.63284	4.615385
-9.2609	39.02484	85.76429	1522.938	26.81171	4.444444
-22.2063	30.52993	493.1203	932.0765	23.75328	4.285714
42.18928	25.99963	1779.935	675.981	40.93194	4.137931
75.08018	-81.259	5637.033	6603.03	204.0011	4
14.80732	20.17276	219.2567	406.9404	10.43662	3.870968
28.56236	39.25333	815.8087	1540.824	39.27721	3.75
-11.2377	47.64142	126.2857	2269.705	39.93318	3.636364
-4.80668	-35.3653	23.10414	1250.701	21.23008	3.529412
53.60292	50.06034	2873.273	2506.038	89.65518	3.428571
7.663914	-25.5924	58.73558	654.9732	11.89515	3.333333
-8.63552	4.218291	74.57212	17.79398	1.539435	3.243243
-47.643	47.99348	2269.86	2303.374	76.22056	3.157895
7.629243	48.20169	58.20535	2323.403	39.69347	3.076923
41.882	0.419915	1754.102	0.176329	29.23797	3
0.86546	5.978191	0.749021	35.73876	0.60813	2.926829
-13.9801	44.96324	195.4428	2021.693	36.95226	2.857143
-3.75795	22.99733	14.12219	528.877	9.049986	2.790698
14.37781	-23.0068	206.7215	529.3131	12.26724	2.727273
-19.4728	51.30114	379.1886	2631.807	50.18326	2.666667
69.64805	-3.44416	4850.851	11.86221	81.04521	2.608696
-49.0322	18.49682	2404.157	342.1324	45.77148	2.553191
-17.6129	-5.38821	310.2155	29.03277	5.654137	2.5
-33.1189	32.93302	1096.862	1084.584	36.35743	2.44898
33.45795	-13.2196	1119.434	174.7569	21.56986	2.4
-46.5104	-16.6299	2163.222	276.5543	40.66293	2.352941
-37.7001	41.46705	1421.297	1719.516	52.34689	2.307692
-6.83882	-12.9621	46.76949	168.0157	3.579753	2.264151
1.375865	-26.6335	1.893006	709.3453	11.85397	2.222222
-48.7809	35.57131	2379.572	1265.318	60.74817	2.181818
66.16245	31.67889	4377.47	1003.552	89.68371	2.142857
-6.79058	6.341471	46.11194	40.21426	1.43877	2.105263
-42.9476	19.77794	1844.497	391.167	37.26107	2.068966
9.192854	16.36049	84.50857	267.6656	5.869569	2.033898
-9704	3.078	94167600	9.474085	1569460	2

References

[1] Box, G. E. P., & Jenkins, G. M. (1970). Time Series Analysis; forecasting and control. Holden Day. San Francisco, California.

[2] Omekara, C. O.,Ekpenyong, E. J. and Ekerete, M. P. (2013): Modeling the Nigeria Inflation Rates using Periodogram and Fourier Series Analysis: *Cbn Journal of Applied Statistics* Vol. 4 No. 2.

[3] Rahman, M. M., Harun-Or-Roshid, Mozid PK, M. A. A. (2011): A comparative study of Wavelet Transform and Fourier Transform. *Journal of physical sciences*, vol.15, 2011,149-160 ISSN: 0972-8791,www. Vidyasagar.ac.in/journal.

[4] Ribeiro, K. M., Braga Junior, R.A., Safadi, T. and Horgan G. (2013). Comparison between Fourier and Wavelets Transforms in Biospeckle Signal. *Applied Mathematics*, 2013,4,11-22http//www.scirp.org/journal/am), http//dx.doi.org/10.4236/am 2013.411A.

[5] Mukhopadhyay, S., Dash, D., Mitra, A., Bhattacharya, P. (2013). A comparative study between seasonal wind speed by Fourier and Wavelet analysis. *India institute of Science Research Kolkata, Mohanpur campus*, nadia-741252, india.

[6] Masset, P. (2008) *Analysis of Financial Time-Series using Fourier and Wavelet Methods*. University of Fribourg, Department of Finance, Bd.Pérolles 90, CH-1700 Fribourg, philippe.masset@unifr.ch.

[7] Konarasinghe, W. G. S. and Abeynayake, N. R. (2015): Fourier Transformation on Model Fitting for Sri Lankan Share Market Returns: *Global Journal for research Analysis , ISSN No22-860,* Volume: iv,

[8] Ekpenyong, E.J. (2008): Pseudo-Additive (Mixed) Fourier Series Model of Time Series: *Asian Journal of Mathematics And Statistics* I(1): 63-68

[9] Ekpenyong, E. J. and Omekara, C. O. (2007): Application of Fourier Series Analysis to Modeling Temperature Data of Uyo Metropolis: *Global Journal of Mathematical Sciences,* 7(1): 5-13

[10] Hamilton, J. D. (1994). *Time Series Analysis. Princeton*: Princeton University Press.

[11] Iwok, I.A. and Udoh, G. M. (2016): a comparative study between the ARIMA-Fourier Model and the Wavelet Model: American Journal of Scientific and industrial Research, ISSN: 2153-649x, oi:10.5251/ajsir.2016.7.6.137.144

[12] Wei. W.S. W (2006): Time Series Analysis Univariate and Multivariate Methods second Edition. pg78