GLOBAL JOURNAL OF MATHEMATICAL SCIENNCES VOL. 7, NO. 1, 2008: 5 - 13
COPYRIGHT (C) BACHUDO SCIENCE CO. LTD. PRINTED IN NIGERIA. ISSN 1596-6208

APPLICATION OF FOURIER SERIES ANALYSIS TO TEMPERATURE DATA

E. J. EKPENYONG and C. O. OMEKARA

(Received 17 October 2006. Revision Accepted 16 July 2007)

ABSTRACT

This Paper seeks to model a periodic time series using Fourier Series Analysis Method and to use such model to forcast future values of such data The mean monthly temperature of Uyo Metropolis consisting of 180 data points (1991 – 2006) are collected for the study The parameter estimates of the Fourier series model are obtained by ordinary least squares method in multiple regression The test of significance of the general model and parameters indicate that the model is statistically significant and the significant parameters provide a Fourier series model of the form $\hat{Y}_t = 26.82 - 1.163\cos\omega t - 0.169\sin\omega t + 0.133\cos 2\omega t + 0.164\sin 2\omega t - 0.116\sin 4\omega t + 0.255e_{t-1}$. The P - P plot is also used to test for the overall goodness-of-fit and it is found out that the model fits well to the data and can be used to forcast the future values of the data

KEY WORDS: Fourier Series Analysis, Periodic Time Series, Forcasting.

INTRODUCTION

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)

Intuitively, when we speak of a wavelike structure we usually have in mind something like the pattern of ocean waves, that is a pattern, which more or less repeats itself after certain intervals This idea may be expressed more precisely in terms of what is called a "periodic function"

Suppose that a pattern of ocean waves happened to repeat itself perfectly at intervals of say, p feet or more precisely that the section of the surface repeated itself perfectly at intervals of p feet then, if we measure the distance along a horizontal line in the vertical plane, and let f(x) denote the height of the surface (measured from some fixed level) at a point whose distance is x feet (from some fixed origin), we express the repetitive nature of the pattern by means of the equation

$$f(x) = f(x + kp), \text{ all } x \qquad (i)$$

where k may take any integral values $0, \pm 1, \pm 2, \ldots\ldots$

Generally if a function f(x) satisfies an equation of the above form, it is said to be periodic, and if p is the smallest number such that equation (i) holds for all x, p is called the period of the function. The most familiar periodic function which we encounter are the sine and cosine functions, since of course $A\sin\omega x$ and $A\cos\omega x$ are both periodic, each with period $(2\pi/\omega)$. The quantity $\omega = 2\pi/p$ is called the angular frequency of $\sin\omega x$ (or $\cos\omega x$) and A is called the amplitude Any well-behaved periodic function can be expressed as (possibly infinite) sum of Sine and Cosine functions Hence a periodic function can be expressed as a sum of Cosine and Sine terms over a discrete set of frequencies. ω_1, ω_2

Thus

$$f(x) = \sum_{r=0}^{\infty} [a_r \cos\omega x + b_r \sin\omega x]. \qquad (2)$$

(Priestly 1981)

Therefore, a function f(x) is said to be periodic, if its function values repeat at regular interval of the independent variable. The regular interval between repetitions is the period of the oscillation (Stroud 1996)

Chatfield (1975) stated that Fourier Series Analysis is basically concerned with approximating a periodic function by a sum of Sine and Cosine terms called the Fourier series representation

Fourier Series Analysis (FSA) was developed in about 1822 by Joseph Fourier a French Mathematician and Physicist. Fourier showed that any periodic observations could be represented by a series of trigonometric functions of Sines and Cosines (Priestly 1981)

Chatfield (1975) further stated that suppose that a function f(t) is defined on $(-\pi, \pi)$ and satisfies the Dirichlet condition (i e. that it is periodic, absolutely integrable over this range, and has a finite number of maxima and minima), then f(t) may be approximated by the Fourier Series

E. J. Ekpenyong, Department of Statistics Akwa Ibom State Polytechnic Ikot Osurua Akwa Ibom State Nigeria
C. O. Omekara, Dept of Maths Statistics and Computer Sci Michael Okpara Univ of Agriculture Umudike Abia State Nigeria

$$F(t) = \frac{a_0}{2} + \sum_{r=1}^{k} (a_r Cos rt + b_r Sin r t) \qquad 3$$

where $a_0 = \frac{1}{\pi} \int_{\pi}^{\pi} f(t)\, dt$

Roerink et al (2000) used Fourier Series Analysis or Harmonic Analysis of time series to screen and remove cloud affected observations and temporarily interpolate the remaining observations to reconstruct gapless images at a prescribed time. They applied Fourier Series Analysis to show that cloud affected data are recognized and replaced successfully. Moreover Hannan, (1974) modeled Noise generated by a stationary, egordic and purely non-deterministic process using the methods of Fast Fourier Transform.

According to Stroud (1996), one important advantage of Fourier Series is that it can represent a function containing discontinuities. Also Priestly (1981) added that one of the most important advantages of Fourier Series Analysis is its simple way of modeling a series with seasonality or cyclicalness. Fourier Series can be used to model seasonal data using several seasonal peaks per year. Based on statistics generated during modeling process, the unimportant seasonal patterns of the model can be dropped without re-estimating the remaining model. Most other methods require a refitting of the model when one or more terms are dropped. However, Fourier Series Analysis is more methodical in identifying significant and insignificant peaks and eliminating those that are not significant. Another advantage he (Priestly) stated is that Fourier Series Analysis gives parameter estimates that are statistically independent of each other (i.e. they are orthogonal).

METHOD OF ANALYSIS

The Method of Analysis is the Fourier Series Analysis used in modeling a periodic data. Since temperature data are seasonal, they are therefore periodic in nature.

The general model for this is : $Y_t = \alpha_0 + \sum_{l=1}^{k} [a_i Cos(i\omega t) + \beta_i Sin(i\omega t)] + e_t$

and the estimated model is given as:

$$\hat{Y}_t = \alpha_0 + \sum [\alpha_i \cos(i\omega t) + \beta_i Sin(i\omega t)] + \hat{e}_t$$

where Y_t = fitted or forecasted value at time t

α_0 = constant used to set the level of the series

β_0 = trend estimate of the series

$\alpha_i, \beta_i\ (i = 1,2,\ldots k)$ = coefficients defining the amplitudes and phased (parameter estimates)

$\omega = \frac{2\pi f}{n}$ which is the Fourier frequency.

k = highest harmonic of ω

e_t = the random error

$\hat{e}_t$ = estimated error term

It is noteworthy here that the highest harmonic, k in Fourier Series Analysis model is the number of observations per season divided by two (2) for an even number of observation and $\frac{(n\ \ 1)}{2}$ for an odd number of observations. When n is an even number the last sine term is zero and should be omitted. The method used in estimating the parameters of the model is the method of least squares in multiple regression analysis.

A statistical test of significance of the parameter estimates is to determine if a parameter estimate is to be included in the model. The test statistic that would be used in this study is the t-test for the significance of the parameters. If the computed t-test is greater than the tabulated t-test, the parameter will be statistically significant.

In testing for the existence of autocorrelation in the error term, Durbin-Watson test will be used. Koutsoysannis (1983) stated that when there is autocorrelation among the error values in a Regression equation, an autoregressive model of order one is usually used to estimate the error component, so that the error term will meet the assumption of no serial correlation.

The autoregressive model of order one is usually given as:

$e_t = \Phi e_{t-1} + z_t$

where z_t is a purely random process and Φ is the parameter estimate.

The P-P plots shall be used to test for the overall goodness-of-fit. In P-P plots the observed cumulative distribution function is plotted against the theoretical cumulative distribution function. The ith observation is plotted against one axis as i/n (i.e the observed distribution function) and against the other axis as $F(x_{(i)})$, where $F(x_{(i)})$, stands for the value of the theoretical cumulative distribution function for the respective observation $x_{(i)}$. If the theoretical cumulative distribution approximates the observed distribution so well, then all points in this plot should fall onto the diagonal line. In other words, the straighter the line formed by the P-P Plot, the more the variable's distribution conforms to the selected test distribution.

DATA ANALYSIS

Testing for significance of Parameters

The Fourier Series Analysis is performed using a statistical package called MINITAB.
The Table 1 below represents the Analysis of variance test for the significance of the general model which indicates that the Fourier Series model is statistically significant. Also Table II shows that the significant parameter estimates in the model are constant, coefficients of cosωt, sinωt, cos2ωt and sin4ωt.

Therefore the Fourier Series model consisting of the significant parameter is $\hat{Y}$ = 26.82 – 1.163cosωt - 0.169 sinωt +0.133cos2ωt + 0.164sin2ωt - 0.116sin4ωt

Table 1: Summary Result for Testing the Significance of the General Model.
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	11	130.983	11.908	42 59	0.000
Error	168	46.967	0.280		
Total	179	177.950			

Table II: Summary Result for Testing the Significance of the Parameter Estimates.

Predictor	Coef	StDev	T	P
Constant	26.8167	0.0394	680.45	0.000
coswt	-1.16309	0.05573	-20.87	0.000
sinwt	-0.16912	0.05573	-3.03	0.003
cos2wt	0.13328	0.05573	2.39	0.018
sin2wt	0.16362	0.05573	2.94	0.004
cos3wt	-0.01114	0.05573	-0.20	0.842
sin3wt	0.06666	0.05573	1.20	0.233
cos4wt	-0.04992	0.05573	-0.90	0.372
sin4wt	-0.11550	0.05573	-2.07	0.040
cos5wt	-0.07580	0.05573	-1.36	0.176
sin5wt	0.05230	0.05573	0.94	0.349
cos6wt	0.01667	0.03941	0.42	0.673

S = 0.5287 R-Sq = 73.6% R-Sq(adj) = 71.9%

Test for Autocorrelation

To test for the existence of autocorrelation, Durbin Watson test is used.

d^{*} = 1.48

therefore since $0 < d^{*} < 2$; there exists some autocorrelation among the error values. In view of this, an autoregressive model of order one [AR (1)] is fitted to the error values. Thus

$\hat{e}_t = 0255\, e_{t-1} + z_t$

Hence the general Fourier Series model is given by: $\hat{Y}_t$ = 26.82 – 1.163cosωt - 0.169 sinωt + 0.133cos2ωt +0.164sin2ωt - 0.116sin4ωt + 0.255e_{t-1}

Testing for the Overall Goodness-of-fit

The P-P plot below (Fig 1) indicates that the model fits well to the data although some of the points falls outside the diagonal. The points falling outside are so close to the line.

Moreover, a comparison between the actual and estimated values shows that the model estimates the actual values well. This is illustrated in plotting the actual and fitted values together in Fig II below.

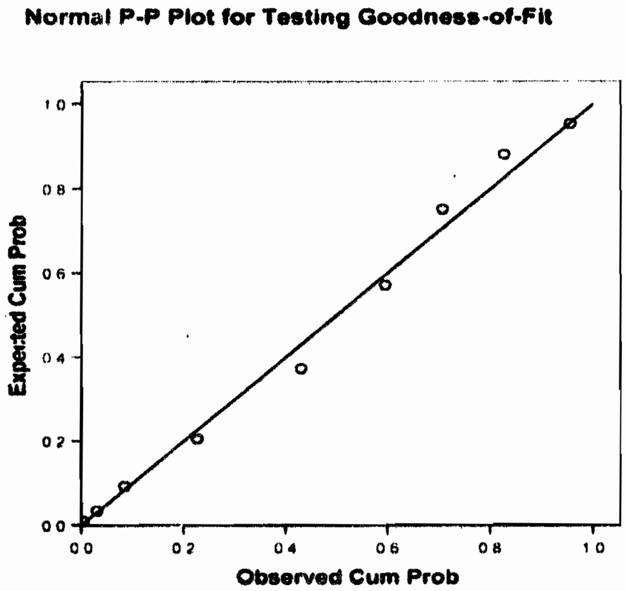

Fig 1: P-P plot for Testing Goodness-of-fit.

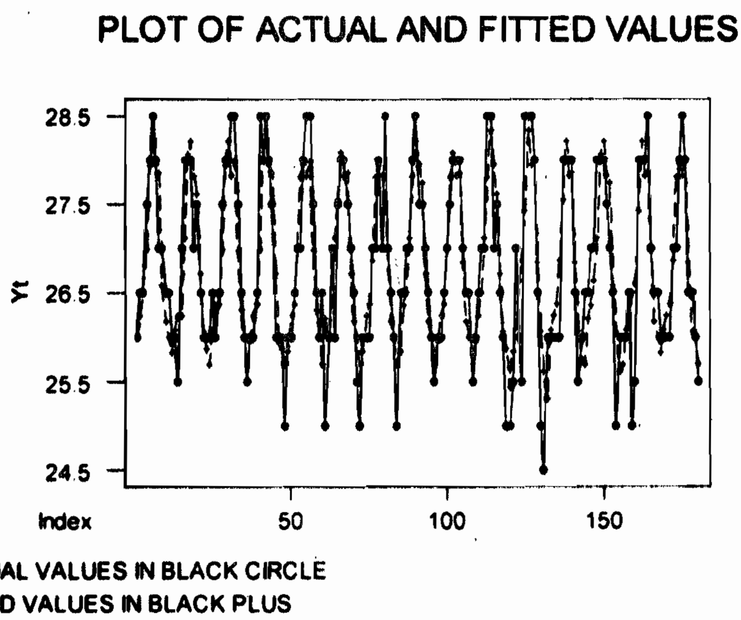

Fig II: Plot of Actual and Fitted Value

CONCLUSION

Analysis of seasonal and cyclical (periodic) data has attracted attention because of its enormous importance in academic and other sector of life. This work has analysed a periodic set of data using the Fourier Series Analysis technique. The results indicate that the Fourier Series model $\hat{Y}_t = 26.82 - 1.163\cos\omega t - 0.169\sin\omega t + 0.133\cos2\omega t$

$+0.164\sin 2\omega t - 0.116\sin 4\omega t + 0.255e_{t-1}$ fits well to the data and could be used to predict the mean monthly temperature; other things being equal.

REFERENCES

Chatfield, C., 1975. Analysis of Time Series: Theory and Practice; Chapman and Hall, London.

Hannan, E., 1974. Time Series Analysis of Noise Properties Automatic Control, IEEE Transactions 19 (6): 706 – 715 Australian National University, Canberra

Koutsoyiannis, A., 1983. Theory of Econometrics: Macmillan Publishers. Hampshire

Priestly, M. B., 1981. Spectral Analysis and Time Series. Academic Press, London.

Roerink et al., 2000. Harmonic Analysis of Time Series; International Journal of Remote Sensing 21 (9): 1911–1917 Taylor and Francis Group, United Kingdom.

Stroud, K. A., 1996. Further Engineering Mathematics: Palgrave Publishers New York

APPENDIX 1

FITTED VALUES ($\hat{Y}$)

1	26.0
2	26.2
3	26.5
4	27.0
5	27.9
6	28.2
7	28.0
8	27.9
9	26.6
10	26.2
11	26.0
12	25.8
13	26.1
14	26.2
15	26.2
16	27.1
17	28.1
18	28.2
19	27.8
20	27.6
21	26.7
22	26.1
23	25.9
24	25.7
25	26.1
26	26.4
27	26.4
28	27.0
29	27.9
30	28.2
31	27.8
32	28.0
33	27.0
34	26.3
35	26.0
36	25.7
37	26.0
38	26.2
39	26.4
40	27.0
41	28.2
42	28.2
43	28.0
44	27.9
45	26.7
46	25.9
47	25.9
48	25.7
49	25.8
50	26.2
51	26.4
52	27.0
53	27.8
54	28.0
55	27.8
56	28.0
57	26.7
58	26.3
59	26.0
60	25.7
61	26.2
62	26.0
63	26.4
64	27.1
65	27.6
66	28.1
67	27.8
68	27.9
69	26.7
70	26.2
71	26.0
72	25.6
73	25.8
74	26.2
75	26.4
76	26.9
77	27.8
78	28.0
79	27.8
80	27.6
81	27.0
82	26.2
83	26.0
84	25.7
85	25.8
86	26.4
87	26.5
88	27.1
89	27.8
90	28.2
91	28.0
92	27.7
93	26.7
94	26.2
95	26.0
96	25.7
97	26.0
98	26.2
99	26.4
100	27.0
101	27.8
102	28.1
103	27.8
104	27.9
105	26.8
106	26.2
107	26.0
108	25.8
109	26.0
110	26.2
111	26.5
112	27.1
113	27.8
114	28.3
115	28.0
116	27.6
117	26.7
118	26.1
119	25.9
120	25.4
121	25.8
122	26.1
123	26.6
124	26.9
125	27.4

126.	28.3
127.	28.0
128.	28.0
129.	26.8
130.	26.1
131.	25.6
132.	25.3
133.	26.1
134.	26.2
135.	26.1
136.	26.9
137	27.6
138.	28.2
139.	27.8
140.	27.9
141.	26.8
142.	26.1
143.	25.7
144.	25.7
145.	26.2
146.	26.4
147.	26.6
148.	27.1
149.	28.1
150.	28.2
151.	27.8
152.	27.7
153.	26.6
154.	26.1
155.	25.6
156.	25.7
157.	26.1
158.	26.2
159.	26.5
160.	26.6
161	27.4
162.	28.2
163.	27.8
164.	27.7
165.	27.0
166.	26.2
167.	26.0
168.	25.8
169.	26.1
170.	26.2
171	26.4
172.	27.0
173.	27.8
174.	28.0
175.	27.8
176.	28.0
177	26.8
178.	26.1
179.	26.0
180.	25.7

APPENDIX II

Actual Values (Yt)

1.	26.0
2.	26.5
3.	26.5
4.	27.5
5.	28.0
6.	28.5
7.	28.0
8.	27.0
9.	27.0
10.	26.5
11.	26.5
12.	26.0
13.	26.0
14.	25.5
15.	27.0
16.	28.0
17.	28.0
18.	28.0
19.	27.0
20.	27.5
21.	26.5
22.	26.0
23.	26.0
24.	26.0
25.	26.5
26.	26.0
27.	26.5
28.	27.5
29.	28.0
30.	28.0
31.	28.5
32.	28.5
33.	27.5
34.	26.5
35.	26.0
36.	25.5
37	26.0
38.	26.0
39.	26.5
40.	28.5
41.	28.0
42.	28.5
43.	28.0
44.	27.5
45.	26.0
46.	26.0
47.	26.0
48.	25.0
49.	26.0
50.	26.0
51.	26.5
52.	27.0
53.	27.0
54.	28.0
55.	28.5
56.	28.5
57.	27.5
58.	26.5
59.	26.0
60.	26.5
61.	25.0
62.	26.0
63.	27.0
64.	26.0
65.	27.5
66.	28.0
67.	28.0
68.	27.5
69.	27.0
70.	26.5
71.	25.5
72.	25.0
73.	26.0
74.	26.0
75.	26.0
76.	27.0
77.	27.0
78.	28.0
79.	27.0
80.	28.5
81.	27.0
82.	26.5
83.	26.0
84.	25.0
85.	26.5
86.	26.5
87.	27.0
88.	27.0
89.	28.0
90.	28.5
91.	27.5
92.	27.5
93.	27.0
94.	26.5
95.	26.0
96.	25.5
97.	26.0
98.	26.0
99.	26.5
100.	27.0
101	27.5
102.	28.0
103.	28.0
104.	28.0
105	27.0
106	26.5
107	26.5
108.	25.5
109.	26.0
110.	26.5
111.	27.0
112.	27.0
113.	28.5
114.	28.5
115.	27.0
116.	27.5
117.	26.5
118.	26.0
119.	25.0
120.	25.0
121.	25.5
122.	27.0
123.	26.0
124.	25.5

125	28.5	153	26.5
126	28.5	154	25.0
127	28.5	155	26.0
128	28.0	156	26.0
129	26.5	157	26.0
130	25.0	158	26.5
131	24.5	159	25.0
132	26.0	160	25.5
133	26.0	161	28.0
134	26.0	162	28.0
135	26.0	163	28.0
136	26.0	164	28.5
137	28.0	165	27.0
138	28.0	166	26.5
139	28.0	167	26.5
140	28.0	168	26.0
141	26.5	169	26.0
142	25.5	170	26.0
143	26.0	171	26.0
144	26.5	172	27.0
145	26.5	173	27.0
146	27.0	174	28.0
147	27.0	175	28.5
148	28.0	176	28.0
149	28.0	177	26.5
150	28.0	178	26.5
151	27.5	179	26.0
152	27.0	180	25.5