Journal of Modern Applied Statistical Methods

Volume 9 | Issue 2 Article 14

11-1-2010

Nonlinear Trigonometric Transformation Time Series Modeling

K. A. Bashiru
Osun State University, Osogbo, Osun State, kehindeadekunle2@yahoo.com

O. E. Olowofeso
Federal University Of Technology, Akure Ondo State, Nigeria, olowofeso@yahoo.com

S. A. Owabumoye
Federal University Of Technology, Akure Ondo State, Nigeria, saowabumoye@yahoo.com

Journal of Modern Applied Statistical Methods
November 2010, Vol. 9, No. 2, 470-479

1538 – 9472/10/$95.00

Nonlinear Trigonometric Transformation Time Series Modeling

K. A. Bashiru
Osun State University,
Osogbo, Osun State

O. E. Olowofeso S. A. Owabumoye
Federal University Of Technology,
Akure Ondo State, Nigeria

The nonlinear trigonometric transformation and augmented nonlinear trigonometric transformation with a polynomial of order two was examined. The two models were tested and compared using daily mean temperatures for 6 major towns in Nigeria with different rates of missing values. The results were used to determine the consistency and efficiency of the models formulated.

Key words: Nonlinear time series, polynomial, consistency, efficiency, missing value, model and forecasting.

Introduction

Time series analysis is an important technique used in many disciplines, including physics, engineering, finance, economics, meteorology, biology, medicine, hydrology, oceanography and geomorphology (Terasvirta & Anderson, 1992). This technique is primarily used to infer properties of a system by the analysis of a measured time record (data) (Priestley, 1988); this is accomplished by fitting a representative model to the data with an aim of discovering the underlying structure as closely as possible.

Traditional time series analysis is based on assumptions of linearity and stationarity. However, time series analysis (Box & Jenkins, 1970; Brock & Potter, 1993) has nonlinear features such as cycles, asymmetries, bursts, jumps, chaos, thresholds and heteroscedasticity, and mixtures of these must also be taken into account. Thus, a problem arises regarding a suitable definition of a nonlinear model because not every time series analysis is purely linear: the nonlinear class clearly encompasses a large number of possible choices. For these reasons, non-linear time series analysis is a rapidly developing area and there have been major developments in model building and forecasting (De Gooijer & Kumar, 1992).

The growing interest in studying nonlinear and non-stationary time series models in many practical problems stems from the inherently non-linear nature of many phenomena in physics, engineering, meteorology, medicine, hydrology, oceanography, economics and finance, that is, many real world problems do not satisfy the assumptions of linearity and/or stationarity (Bates & Watts, 1988; DeGooijer & Kumar, 1992; Sugihara & May, 1990). Therefore, for many real time series data, nonlinear models are more appropriate than linear models for accurately describing the dynamic of the series and making multi-step-ahead forecast (Tsay, 1986; Barnett, Powell & Tauchen, 1991; Olowofeso, 2006). For example, financial markets and trends are influenced by climatic factors like daily temperature, amount of rainfall and intensity of sun, these are areas where a need exists to explain behaviors that are far from being even approximately linear. Nonlinear models would be more appropriate for forecasting and accurately describing returns and

K. A. Bashiru is a lecturer in the Department of Mathematical and Physical Sciences. His areas of interest are Geostatistics and Time series Econometrics. Email him at: kehindeadekunle2@yahoo.com. O. E. Olowofeso is an Associate Professor of Statistics in the Mathematical Sciences Department. He is also currently working in the Central Bank of Nigeria, Abuja. Email: olowofeso@yahoo.com. S. A. Owabumoye is a Master's student under the supervision of O. E Olowofeso in the Mathematical Sciences Department. Email: saowabumoye@yahoo.com.

volatility. Thus, the need for the further development of the theory and applications for nonlinear models is essential, and, because there are an enormous number of nonlinear models available for modeling and forecasting economic time series, research should help provide guidance for choosing the best model for a particular application (Robinson, 1983).

Methodology

The model proposed by Gallant (1981) called the Augmented Nonlinear Parametric Time Series Model (ANPTSM) was used in this study and a second model was formulated based on the Least Square Method Modified Nonlinear Trigonometric Transformation Time Series Model (MNTTTSM).

Data

Data used in this study were daily mean of temperatures from 1987 to 1996 for Ikeja, Ibadan, Ilorin, Minna and Zaria. The data were collected from the Meteorological Centre-Oshodi Lagos.

Model Formulation

Consider the format shown in Table 1. In this model, up to 9 years were considered and the model is formulated based on the data as shown in Table 2.

Assumption and Notation for the Models

Let:

$X_{t,i,k}$ = value of occurrence for day t of Month i in the year k;

$X_{t,k}$ = mean occurrence for day t of year k;

$X^{*}_{i,k}$ = mean occurrence for month i of year k;

X^{*y}_{K} = overall yearly mean for the sampled month;

X^{*m}_{i} = overall monthly mean for the sampled year;

t = the position of the day from the first day of the Month. $1 \leq t \leq 31$;

t_i = the sum of days in month i for $1 \leq i \leq 12$;

t_{ik} = the sum of days from the initial sampled month of initial sampled year to month i of year k;

t_i* = the sum of days from the initial sampled month to month I;

k = the position of a particular year from an initial sample year for $-\infty \leq k \leq \infty$;

n = the number of sampled years;

m = the number of sampled months; and

X_* = Grand Mean occurrence for k year(s) examined.

The first model was reviewed based on the assumption that the sum of the occurrences were presented monthly, where i^{th} month represents the month i for $1 \leq i \leq 12$ which is to be modeled using the number of days in each month (see Table 3).

Table 1: Model Formulation for a Particular Year

t/i	1	2	3	4	5	6	7	8	9	10	11	12	$\Sigma x_i/12$
1	$x_{1,1,k}$	$x_{1,2,k}$	$x_{1,3,k}$	$x_{1,4,k}$	$x_{1,5,k}$	$x_{1,6,k}$	$x_{1,7,k}$	$x_{1,8,k}$	$x_{1,9,k}$	$x_{1,10,k}$	$x_{1,11,k}$	$x_{1,12,k}$	$\bar{X}_1$
2	$x_{2,1,k}$	$x_{2,2,k}$	$x_{2,3,k}$	$x_{2,4,k}$	$x_{2,5,k}$	$x_{2,6,k}$	$x_{2,7,k}$	$x_{2,8,k}$	$x_{2,9,k}$	$x_{2,10,k}$	$x_{2,11,k}$	$x_{2,12,k}$	$\bar{X}_2$
t	$x_{t,1,k}$	$x_{t,2,k}$	$x_{t,3,k}$	$x_{t,4,k}$	$x_{t,5,k}$	$x_{t,6,k}$	$x_{t,7,k}$	$x_{t,8,k}$	$x_{t,9,k}$	$x_{t,10,k}$	$x_{t,11,k}$	$x_{t,12,k}$	$\bar{X}_t$
	$^{*}_{1,k}$	$x^{*}_{2,k}$	$x^{*}_{3,k}$	$x^{*}_{4,k}$	$x^{*}_{5,k}$	$x^{*}_{6,k}$	$x^{*}_{7,k}$	$x^{*}_{8,k}$	$x^{*}_{9,k}$	$x^{*}_{10,k}$	$x^{*}_{11,k}$	$x^{*}_{12,k}$	$\Sigma x^{*}_{i,k}/12$

Table 2: Model Data Formulation

t/i_k	1	2	3	…	i_k	$\Sigma x_i/ik$
1	$x_{1,1,1}$	$x_{1,2,1}$	$x_{1,3,1}$	…	$x_{1,i,k}$	$\bar{X}_{1,k}$
2	$x_{2,1,1}$	$x_{2,2,1}$	$x_{2,3,1}$	…	$x_{2,i,k}$	$\bar{X}_{2,k}$
t	$x_{t,1,1}$	$x_{t,2,1}$	$x_{t,3,1}$	…	$x_{t,i,k}$	$\bar{X}_{t,k}$
	$x^{*}_{1,k}$	$x^{*}_{2,k}$	$x^{*}_{3,k}$	…	$x^{*}_{i,k}$	$\Sigma x^{*}_{i,k}/ik = X$

Table 3: Months and Sums of Occurrences Modeled

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan		
i	1	2	3	4	5	6	7	8	9	10	11	12	13	…	i_k
t	31	$28^1/_4$	31	30	31	30	31	31	30	31	30	31	31	…	t_i
t_i	31	$59^1/_4$	$90^1/_4$	$120^1/_4$	$151^1/_4$	$181^1/_4$	$212^1/_4$	$243^1/_4$	$273^1/_4$	$304^1/_4$	$334^1/_4$	$365^1/_4$	$396^1/_4$	…	t

Augmented Nonlinear Parametric Time Series Model (ANPTSM)

Trigonometric (sine and cosine) transformation augmented with polynomial of order two was applied to formulate the model across the year, that is, the monthly mean sample and the least square methods were used for estimating the model's parameters as follows. Let the equation be of the form

$$X_{t,i,k} = a_1 + a_2 t\mathrm{Sin}(t_{ik}) + a_3 t^2 \mathrm{Cos}(t_{ik}) + \varepsilon_i \qquad 1 \le i \le 12 \tag{3.0}$$

The expected value of $X_{t,i,k}$ is $X^*_{i,k}$ then the equation can be reformed as below to estimate the parameters; a_1, a_2 and a_3 using Least Square Method.

$$X^*_{i,k} = a_1 + a_2 t_i \mathrm{Sin}(t_{ik}) + a_3 t_i^2 \mathrm{Cos}(t_{ik}) + \varepsilon_i \qquad 1 \le i \le 12 \tag{3.1}$$

$$\therefore \varepsilon_i = X^*_{i,k} - \left(a_1 + a_2 t_i \mathrm{Sin}(t_{ik}) + a_3 t_i^2 \mathrm{Cos}(t_{ik})\right) \tag{3.2}$$

Let $\boldsymbol{\Sigma}\varepsilon_i^2 = S$

$$S = \Sigma(X^*_{i,k} - \left(a_1 + a_2 t_i \mathrm{Sin}(t_{ik}) + a_3 t_i^2 \mathrm{Cos}(t_{ik})\right))^2 \tag{3.3}$$

Differentiating 3.3 with respect to a_1, a_2, a_3, a_3, as results in

$$\Sigma X^*_{i,k} = ma_1 + a_2 \Sigma t_i \mathrm{Sin}(t_{ik}) + a_3 \Sigma t_i^2 \mathrm{Cos}(t_{ik}) \tag{3.4}$$

where m is the number of the monthly sample mean examined. Similarly, as

$$\frac{\partial S}{\partial a_2} \to 0$$

then

$$\Sigma t_i \mathrm{Sin}(t_{ik}) X^*_{i,k} = a_1 \Sigma t_i \mathrm{Sin}(t_{ik}) + a_2 \Sigma t_i^2 \mathrm{Sin}^2(t_{ik}) + a_3 \Sigma t_i^3 \mathrm{Sin}(t_{ik}) \mathrm{Cos}(t_{ik}) \tag{3.5}$$

and as

$$\frac{\partial S}{\partial a_3} \to 0$$

then

$$\begin{aligned}\Sigma t_i^2 \mathrm{Cos}(t_{ik}) X^*_{i,k} = {} & a_1 \Sigma t_i^2 \mathrm{Cos}(t_{ik}) \\ & + a_2 \Sigma t_i^3 \mathrm{Sin}(t_{ik}) \mathrm{Cos}(t_{ik}) \\ & + a_3 \Sigma t_i^4 \mathrm{Cos}^2(t_{ik})\end{aligned} \tag{3.6}$$

Simultaneously solving equations 3.4, 3.5 and 3.6 using Cramer's Rule results in equations 3.7-3.10.

$$\Delta_0 = m\{\Sigma t_i^2 Sin^2(t_{ik})\Sigma t_i^4 Cos^2(t_{ik}) - (\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik}))^2\}$$
$$-\Sigma t_i Sin(t_{ik})\{\Sigma t_i Sin(t_{ik})\Sigma t_i^4 Cos^2(t_{ik}) - \Sigma t_i^2 Cos(t_{ik})\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik})\} \quad (3.7)$$
$$+\Sigma t_i^2 Cos(t_{ik})\{\Sigma t_i Sin(t_{ik})\Sigma t^3 Sin(t_i)Cos(t_{ik}) - \Sigma t_i^2 Cos(t_i)\Sigma t_i^2 Sin^2(t_{ik})\}$$

$$\Delta_1 = \Sigma X^*_{i,k}\{\Sigma t_i^2 Sin^2(t_{ik})\Sigma t_i^4 Cos^2(t_{ik}) - (\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik}))^2\}$$
$$-\Sigma t_i Sin(t_{ik})\{\Sigma t_i Sin(t_{ik})X^*_{i,k}\Sigma t^4 Cos^2(t_{ik}) - \Sigma t_i^2 Cos(t_{ik})X^*_{i,k}\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik})\} \quad (3.8)$$
$$+\Sigma t_i^2 Cos(t_{ik})\{\Sigma t_i Sin(t_{ik})\ X^*_{i,k}\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik}) - \Sigma t_i^2 Cos(t_{ik})\ X^*_{i,k}\Sigma t_i^2 Sin^2(t_{ik})\}$$

$$\Delta_2 = m\{\Sigma t_i Sin(t_{ik})X^*_{i,k}\Sigma t_i^4 Cos^2(t_{ik}) - (\Sigma t_i^2 Cos(t_{ik})X^*_{i,k}\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik})\}$$
$$-\Sigma X^*_{i,k}\{\Sigma t_i Sin(t_{ik})\Sigma t_i^4 Cos^2(t_{ik}) - \Sigma t_i^2 Cos(t_{ik})\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik})\} \quad (3.9)$$
$$+\Sigma t_i^2 Cos(t_{ik})\{\Sigma t_i Sin(t_{ik})\Sigma t_i^2 Cos(t_{ik})X^*_{i,k} - \Sigma t_i^2 Cos(t_{ik})\Sigma t_i Sin(t_{ik})X^*_{i,k}\}$$

$$\Delta_3 = m\{\Sigma t_i^2 Sin^2(t_{ik})\Sigma t_i^2 Cos(t_{ik})X^*_{i,k} - (\Sigma t_i^3 Sin(t_{ik})Cos(t_i)\Sigma t_i Sin(t_{ik})X^*_{i,k}\}$$
$$-\Sigma t_i Sin(t_{ik})\{\Sigma t_i Sin(t_{ik})\Sigma t_i^2 Cos(t_{ik})X^*_{i,k} - \Sigma t_i^2 Cos(t_{ik})\Sigma t_i Sin(t_{ik})X^*_{i,k}\} \quad (3.10)$$
$$+\Sigma X^*_{i,k}\{\Sigma t_i Sin(t_{ik})\Sigma t_i^3 Sin(t_{ik})Cos(t_{ik}) - \Sigma t_i^2 Cos(t_{ik})\Sigma t_i^2 Sin^2(t_{ik})$$

Therefore, from equations 3.7, 3.8, 3.9 and 3.10, the following result:

$$a_1 = \frac{\Delta_1}{\Delta_0} \quad (3.11)$$

$$a_2 = \frac{\Delta_2}{\Delta_0} \quad (3.12)$$

$$a_3 = \frac{\Delta_3}{\Delta_0} \quad (3.13)$$

Next, substituting 3.11, 3.12 and 3.13 into 3.1 gives:

$$\hat{X}^*_{i,k} = \frac{\Delta_1}{\Delta_0} + \frac{\Delta_2}{\Delta_0} tSin(t_{ik}) + \frac{\Delta_3}{\Delta_0} t^2 Cos(t_{ik}). \quad (3.14)$$

Because $X^*_{i,k}$ is the expected value of $X_{t,i,k}$, equation 3.14 can be rewritten as

$$\hat{X}_{t,i,k} = \frac{\Delta_1}{\Delta_0} + \frac{\Delta_2}{\Delta_0} tSin(t_{ik}) + \frac{\Delta_3}{\Delta_0} t^2 Cos(t_{ik}). \quad (3.15)$$

Models 3.14 and 3.15 would only be visible provided there is an occurrence within a month of any sampled year.

Modified Nonlinear Trigonometric Transformation Time Series Model (MNTTTSM)

In a situation where a whole month of data is missing, the above model may be difficult to apply and a different model would be needed. The model for such occurrence is formulated as follows. If the data in 3.2 are reformed such that the monthly means are those shown in Table 4. Consider:

$$X^*_{i,k} = a + b\sin(t_i *) + \varepsilon_i \quad (3.16)$$

where

$$1 \le i \le 12\ ,\ 1 \le t_i \le 365\frac{1}{4}$$

If the expected value of $X^*_{i,k}$ is X^{*m}_i, then equation 3.16 can take the form

$$X^{*m}_i = a + b\sin(t_i *) + \varepsilon_i \qquad (3.17)$$

where

$$1 \le i \le 12\ ,\ 1 \le t_i * \le 365\frac{1}{4}$$

An ordinary least square method was used in estimating the parameters a and b. If $S_m = \varepsilon_i^2 = \mathbf{\Sigma}$ $(X^{*m}_i$ $-(a+b\sin(t_i^*))^2$, then differentiating with respect to a and b

$$\frac{\partial S_m}{\partial a} = -2\Sigma\left(X^{*m}_i - a + b\sin\left(t_i^*\right)\right)$$

as

$$\frac{\partial S_m}{\partial a} \to 0$$

$$\Rightarrow\ ð\Sigma X^{*m}_i = 12a + b\Sigma\sin\left(t_i^*\right) \qquad (3.18)$$

Also,

$$\frac{\partial S_m}{\partial b} = -2\Sigma(\sin\left(t_i^*\right)\left(X^{*m}_i - \left(a + b\sin\left(t_i^*\right)\right)\right)$$

as

$$\frac{\partial S_m}{\partial b} \to 0$$

$$\Rightarrow\ \Sigma\sin(t_i *)\ X^{*m}_i = a\Sigma\sin\left(t_i^*\right) + b\Sigma\sin^2\left(t_i^*\right) \qquad (3.19)$$

Using Cramer's Rule to solve equations 3.18 and 3.19 simultaneously, results in:

$$\Delta_4 = 12\Sigma\mathrm{Sin}^2(t_i^*) - \left(\Sigma\mathrm{Sin}(t_i^*)\right)^2$$

$$\Delta_5 = \Sigma X^{*m}_i\Sigma\mathrm{Sin}^2(t_i^*) - \Sigma\mathrm{Sin}(t_i^*)X^{*m}_i\Sigma\mathrm{Sin}(t_i^*)$$

$$\Delta_6 = 12\Sigma\mathrm{Sin}(t_i^*)X^{*m}_i - \Sigma X^{*m}_i\Sigma\mathrm{Sin}(t_i^*)$$

where parameters

$$a = \frac{\Delta_5}{\Delta_4} = \frac{\Sigma X^{*m}_i\Sigma\mathrm{Sin}^2(t_i^*) - \Sigma\mathrm{Sin}(t_i^*)X^{*m}_i\Sigma\mathrm{Sin}(t_i^*)}{12\Sigma\mathrm{Sin}^2(t_i^*) - \left(\Sigma\mathrm{Sin}(t_i^*)\right)^2} \qquad (3.20)$$

and

$$b = \frac{\Delta_6}{\Delta_4} = \frac{12\Sigma\mathrm{Sin}(t_i^*)X^{*m}_i - \Sigma X^{*m}_i\Sigma\mathrm{Sin}(t_i^*)}{12\Sigma\mathrm{Sin}^2(t_i^*) - \left(\Sigma\mathrm{Sin}(t_i^*)\right)^2} \qquad (3.21)$$

Therefore, the model for monthly occurrence is

Table 4: Modified Nonlinear Trigonometric Transformation Time Series Model Data

t/i	1	2	3	4	5	6	7	8	9	10	11	12	$\Sigma x_i/12$
1	$X^*_{1,1}$	$X^*_{2,1}$	$X^*_{3,1}$	$X^*_{4,1}$	$X^*_{5,1}$	$X^*_{6,1}$	$X^*_{7,1}$	$X^*_{8,1}$	$X^*_{9,1}$	$X^*_{10,1}$	$X^*_{11,1}$	$X^*_{12,1}$	X^{*y}_1
2	$X^*_{1,2}$	$X^*_{2,2}$	$X^*_{3,2}$	$X^*_{4,2}$	$X^*_{5,2}$	$X^*_{6,2}$	$X^*_{7,2}$	$X^*_{8,2}$	$X^*_{9,2}$	$X^*_{10,2}$	$X^*_{11,2}$	$X^*_{12,2}$	X^{*y}_2
k	$X^*_{1,k}$	$X^*_{2,k}$	$X^*_{3,k}$	$X^*_{4,k}$	$X^*_{5,k}$	$X^*_{6,k}$	$X^*_{7,k}$	$X^*_{8,k}$	$X^*_{9,k}$	$X^*_{10,k}$	$X^*_{11,k}$	$X^*_{12,k}$	X^{*y}_k
	x^{*m}_1	x^{*m}_2	x^{*m}_3	x^{*m}_4	x^{*m}_5	x^{*m}_6	x^{*m}_7	x^{*m}_8	x^{*m}_9	x^{*m}_{10}	x^{*m}_{11}	x^{*m}	$\Sigma x^{*m}_{,i} = X_{*12}$

$$X^{*m}{}_{i} = \frac{\Delta_5}{\Delta_4} + \frac{\Delta_6}{\Delta_4} Sin(t_i^*) \qquad (3.22)$$

Because $X^{*m}{}_i$ is an expected value for $X^{*}{}_{i,k}$ then equation 3.22 can be rewritten as

$$X^{*}{}_{i,k} = \frac{\Delta_5}{\Delta_4} + \frac{\Delta_6}{\Delta_4} Sin(t_i^*) \qquad (3.23)$$

Similarly, along the sampled year $X^{*y}{}_K = c + d \, Sin(\lambda k)$ for $-\infty \le k \le +\infty$, $15 \le \lambda \le 75$. The λ must be chosen such that $\Sigma \varepsilon_i = 0$, $\Sigma \varepsilon_i^2$ is as small as possible.

If $S_y = \varepsilon_i^2 = \Sigma (X^{*y}{}_K - (c + d\sin(\lambda k))^2$ then

$$\frac{\partial S_y}{\partial c} = -2\Sigma(X^{*y}{}_K - (c + d\sin(\lambda k))$$

as

$$\frac{\partial S_y}{\partial c} \to 0$$

$$\Rightarrow \; ð\Sigma X^{*y}{}_K = nc + d\Sigma \sin(\lambda k) \qquad (3.24)$$

Also,

$$\frac{\partial S_y}{\partial d} = -2\Sigma(\sin(\lambda k) \; X^{*y}{}_K - (c + d\sin(\lambda k)))$$

as

$$\frac{\partial S_y}{\partial d} \to 0$$

$$\Rightarrow \; ð\Sigma \sin(\lambda k) \; X^{*y}{}_K = c\Sigma \sin(\lambda k) + d\Sigma \sin^2(\lambda k)$$

(3.25)

Solving equations 3.24 and 3.25 simultaneously using Cramer's Rule results in

$$\Delta_7 = n\Sigma Sin^2(\lambda k) - (\Sigma Sin(\lambda k))^2$$
$$\Delta_8 = \Sigma X^{*y}{}_k \Sigma Sin^2(\lambda k) - \Sigma Sin(\lambda k) X^{*y}{}_k \Sigma Sin(\lambda k)$$
$$\Delta_9 = n\Sigma Sin(\lambda k) X^{*y}{}_k - \Sigma X^{*y}{}_k \Sigma Sin(\lambda k)$$

Where the parameters

$$c = \frac{\Delta_8}{\Delta_7} = \frac{\Sigma X^{*y}{}_k \Sigma Sin^2(\lambda k) - \Sigma Sin(\lambda k) X^{*y}{}_k \Sigma Sin(\lambda k)}{n\Sigma Sin^2(\lambda k) - (\Sigma Sin(\lambda k))^2} \qquad (3.26)$$

and

$$d = \frac{\Delta_9}{\Delta_7} = \frac{n\Sigma Sin(\lambda k) X^{*y}{}_k - \Sigma X^{*y}{}_k \Sigma Sin(\lambda k)}{n\Sigma Sin^2(\lambda k) - (\Sigma Sin(\lambda k))^2} \qquad (3.27)$$

$$\therefore X^{*y}{}_k = \frac{\Delta_8}{\Delta_7} + \frac{\Delta_9}{\Delta_7} Sin(\lambda k) \qquad (3.28)$$

The method of placing expected occurrences in a contingency table of a Chi-square was applied using equations 3.23 and 3.28 to obtain the model to find the daily occurrences for a particular month of a particular year. Therefore, the model for expected daily occurrences is

$$X_{t,i,k} = \frac{n(X^{*m}{}_i)(X^{*y}{}_k)}{\Sigma X^{*y}{}_k} \qquad (3.29)$$

Substituting 3.23 and 3.28 into 3.29, results in

$$X_{t,i,k} = \frac{n\left(\frac{\Delta_5}{\Delta_4} + \frac{\Delta_6}{\Delta_4} Sin(t_i^*)\right)\left(\frac{\Delta_8}{\Delta_7} + \frac{\Delta_9}{\Delta_7} Sin(\lambda k)\right)}{\sum\left(\frac{\Delta_8}{\Delta_7} + \frac{\Delta_9}{\Delta_7} Sin(\lambda k)\right)} \qquad (3.30)$$

Results

Model Analysis and Discussion

The data on the daily mean temperature for Ikeja, Ibadan, Ilorin, Minna and Zaria collected from the Meteorological Centre-Oshodi Lagos were used. The parameters of the models were estimated and the fitted models for each zone are shown in Table 5 for Ikeja, Ibadan, Ilorin and Minna for ANPTSM. Data for the daily mean temperature was used to estimate the parameters. The fitted model for Zaria could

Table 5: The Fitted Models for ANPTSM

Zones	Augmented Nonlinear Parametric Time Series Model (ANPTSM)
IKEJA	$26.88642582 + 0.047971536t\text{Sin}(t_{ik}) - 0.000143793t^2\text{Cos}(t_{ik})$
IBADAN	$26.36612286 + 0.054847742t\text{Sin}(t_{ik}) - 0.0000344912t^2\text{Cos}\,t_{ik}$
ILORIN	$26.2476883 + 0.048115874t\text{Sin}(t_{ik}) - 0.000833551t^2\,Cos(t_{ik})$
MINNA	$25.72428 + 0.062853tSin(t_{ik}) - 0.00073t^2\,Cos(t_{ik})$
ZARIA	–

Table 6: The Fitted Models for MNTTTSM

Zones	Modified Nonlinear Trigonometric Transformation Time Series Model (MNTTTSM)
IKEJA	$\dfrac{10(26.87226 + 1.420072Sin\,t_i^*)(26.88996 + 0.13116Sin60k)}{\sum_{K=1}^{10}(26.88996 + 0.1311Sin60k)}$
IBADAN	$\dfrac{10(26.36749 + 1.591834Sin\,t_i^*)(26.36761 + 0.13535Sin90k)}{\sum_{K=1}^{10}(26.36761 + 0.1311Sin90k)}$
ILORIN	$\dfrac{10(26.45708 + 1.816182Sin\,t_i^*)(26.40106 + 0.409024Sin45k)}{\sum_{K=1}^{10}(26.40106 + 0.409024Sin45k)}$
MINNA	$\dfrac{10(27.56143 + 2.508736Sin\,t_i^*)(27.67047 + 0.112148Sin90k)}{\sum_{K=1}^{10}(27.67047 + 0.112148Sin90k)}$
ZARIA	$\dfrac{10(24.98532 + 1.210108Sin\,t_i^*)(25.00445 + 0.222282Sin90k)}{\sum_{K=1}^{10}(25.00445 + 0.222282Sin90k)}$

not be formulated due to the fact that many months of data were missing.

Table 6 shows the fitted models for Ikeja, Ibadan,Ilorin, Minna and Zaria for MNTTTSM using the daily mean temperature data to estimate their parameters. The fitted model for Zaria was formulated because MNTTTSM has the strength of addressing the problem of missing values. Thus, although many months' data were missing from Zaria's daily mean temperature, MNTTTSM parameters could still be estimated. This is one of the advantages of MNTTTSM over ANPTSM.

Table 7 shows that the results of the Pearson Product Moment Correlation coefficients and Spearman Brown's rank Order Correlation coefficients for Ikeja, Ibadan, Ilorin and Minna are highly and positively correlated, indicating a strong relationship between the actual data and estimated data for the daily mean temperature. In Zaria the correlation coefficient for MNTTTSM is positive but low which may indicate a weak relationship between the actual and estimated daily mean temperatures.

Apart from Ibadan, in which the correlation coefficient in ANPTSM is greater than MNTTTSM and Ikeja which has equal correlation coefficients, all other Zones, the correlation coefficient in MNTTTSM is greater than ANPTSM. This indicates that MNTTTSM shows a stronger relationship between the actual and estimated values than does ANPTSM. Although the relationship between actual and estimated values of MNTTTSM in Zaria is weak but positive, that of ANPTSM could not be estimated due to the large number of missing values in the data. Also, all of the correlations are significant at the 0.01 level (2-tailed).

As shown in Table 8, the mean of the actual and estimated values for each zones of all models are almost equal; differences are due to approximation (truncation error) during calculation. Also, the mean of the actual and estimated values of MNTTTSM are closer than those of ANPTSM, which implies that MNTTTSM estimates better than ANPTSM. It was also discovered from results in Table 8 that the more missing values in the data, the weaker the ANPTSM is in estimating, while in MNTTTSM, the model maintains its precision.

Table 7: Correlation Coefficients

Zones	Types	ANPTSM		MNTTTSM	
		Coefficients	Sig.	Coefficients	Sig.
IKEJA	Pearson's r	0.607	.000	0.607	.000
	Spearman's Rho	0.620	.000	0.620	.000
IBADAN	Pearson's r	0.594	.000	0.575	.000
	Spearman's Rho	0.622	.000	0.584	.000
ILORIN	Pearson's r	0.503	.000	0.589	.000
	Spearman's Rho	0.560	.000	0.612	.000
MINNA	Pearson's r	0.596	.000	0.676	.000
	Spearman's Rho	0.656	.000	0.686	.000
ZARIA	Pearson's r	-	-	0.419	.000
	Spearman's Rho	-	-	0.445	.000

Table 8: Comparison of ANPTSM and MNTTSM Means

Zones	N	ANPTSM		MNTTTSM	
		Actual	Estimated	Actual	Estimated
IKEJA	3,660	26.9077	26.8759	26.9077	26.8759
IBADAN	3,601	26.3749	26.3756	26.3749	26.3791
ILORIN	3,580	26.4558	26.2443	26.4558	26.4593
MINNA	3,362	27.5489	26.3611	27.5489	27.5559
ZARIA	3,588	-	-	25.0514	25.0172

Table 9 shows that the standard deviations for MNTTTSM are less than those of ANPTSM which indicates that MNTTTSM is better in estimating and forecasting than ANPTSM. Similarly, apart from the standard error of ANPTSM and MNTTTSM of Ikeja, which are equal, it may be observed that the standard errors for MNTTTSM were also smaller than those of ANPTSM, which indicates that MNTTTSM is better in estimating and forecasting than ANPTSM for time series data with missing values.

Table 10 shows that at Ikeja, there is a 95% chance that the differences between the actual and estimated daily mean temperature would lie between -0.00749 and 0.07119 in ANPTSM and -0.00748 and 0.07118 in MNTTTSM. Similarly, at Ibadan; -0.0462 and 0.04473 in ANPTSM and -0.0496 and 0.04114 in MNTTTSM, at Ilorin; 0.1505 and 0.2723 in ANPTSM and -0.0592 and 0.05218 in MNTTTSM, at Minna; 1.1155 and 1.2601 in model I and -0.0689 and 0.05482 in MNTTTSM while in Zaria is between -0.0546 and 0.12310.

It was also discovered that the range of the confidence interval for MNTTTSM is less than that of ANPTSM for Ikeja and Ibadan. In Ilorin and Minna, the lower confidence intervals of differences for ANPTSM are positive which indicates a 95% chance that the differences between their actual and estimated daily temperature (actual – estimate) are positive while those of MNTTTSM are not. This implies that the estimated daily temperatures for ANPTSM at Ilorin and Minna were under-estimated. Hence MNTTTSM is better in estimating and forecasting than ANPTSM when there are missing values in the time series.

Table 9: Comparison of ANPTSM and MNTTSM'S Standard Deviation and Standard Error of Differences

Zones	ANPTSM		MNTTTSM	
	Std. Dev.	Std. Error of the Mean	Std. Dev.	Std. Error of the Mean
IKEJA	1.2138	0.02006	1.2137	0.02006
IBADAN	1.3913	0.02319	1.3882	0.02313
ILORIN	1.8585	0.03106	1.6996	0.02841
MINNA	2.1381	0.03688	1.8293	0.03155
ZARIA	-	-	2.7152	0.04533

Table 10: Comparison of ANPTSM and MNTTTSM's 95 % Confidence Interval of the Difference

Zones	ANPTSM		MNTTTSM	
	Lower	Upper	Lower	Upper
IKEJA	-0.00749	0.07119	-0.00748	0.07118
IBADAN	-0.0462	0.04473	-0.0496	0.04114
ILORIN	0.1505	0.2723	-0.0592	0.05218
MINNA	1.1155	1.2601	-0.0689	0.05482
ZARIA	-	-	-0.0546	0.12310

Conclusion

The two models tested in this study were the Augmented Nonlinear Parametric Time Series Model (ANPTSM) and the Modified Nonlinear Trigonometric Transformation Time Series Model (MNTTTSM). Both models were tested using daily mean temperatures at Ikeja, Ibadan, Ilorin, Minna and Zaria, and the results were analyzed. It was discovered that ANPTSM could be used in forecasting provided the data is having few missing values. However MNTTTSM estimates forecasts better than ANPTSM in estimating missing values and forecasting. Based on results of this study, MNTTTSM is more efficient in estimating missing values and forecasts better than ANPTSM.

The beauty of a good model developed for nonlinear time series modeling is the ability to forecast better, the new method MNTTTTSM is therefore recommended for numerical solutions for a nonlinear model with missing values due to its higher capacity to address missing values. It was also noted that the mathematical derivative of MNTTTSM is simpler than ANPTSM which did not forecast better. Further research could be conducted by placing a condition in which data having a year or more of missing values is taken into consideration.

References

Barnett, W. A., Powell, J., & Tauchen, G. E. (1991). *Nonparametric and semi parametric methods in econometrics and statistics*. Proceedings of the 5th International Symposium in Economic Theory and Econometrics. Cambridge: Cambridge University Press.

Bates, D. M., & Watts, D. G. (1988). *Nonlinear regression analysis and its applications*. New York: Wiley.

Box, G. E. P., Jenkins, G. M. (1970). *Time series analysis, forecasting and control*. San Francisco: Holden-Day.

Brock, W. A., Potter, S. M. (1993). Nonlinear time series and macroeconometrics. In: G.S. Maddala, C. R. Rao & H. R. Vinod, Eds., *Handbook of Statistics*, *Vol. 11*, 195-229. Amsterdam: North-Holland.

De Gooijer, J. G., & Kumar, K. (1992). Some recent developments in non-linear time series modeling, testing and forecasting. *International Journal of Forecasting*, *8*, 135-156.

Friedman, J. H., & Stuetzle, W. (1981). Projection pursuit regression. *Journal of the American Statistical Association*, *76*, 817-823.

Gallant, A. R. (1987). *Nonlinear statistical models*. New York: Wiley.

Gallant, A. R., & Tauchen, G. (1989). Semi nonparametric estimation of conditionally constrained heterogeneous processes: asset pricing application. *Econometrica*, *57*, 1091-1120.

Granger, C. W. J., & Hallman, J. J. (1991a). Nonlinear transformations integrated time series. *Journal of Time Series Analysis*, *12*, 207-224.

Olowofeso, O. E. (2006). *Varying co-efficient regression models for non-linear time series: Estimation and testing*.13th Meeting of the Forum for Interdisciplinary Mathematical and Statistical Techniques.

Priestley, M. (1988). *Non-linear and non-stationary time series analysis*. London: Academic Press.

Robinson, P. M. (1983). Non-parametric estimation for time series models. *Journal of Time Series Analysis*, *4*, 185-208.

Sugihara, G., & May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, *344*, 734-741.

Terasvirta, T., & Anderson, H. M. (1992). Modelling nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics*, *7*, S119-S136.

Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika*, *73*, 461-466.