Estimation of Parameters of Partially Sinusoidal Frequency Model

Swagata Nandi
Debasis Kundu

1

ESTIMATION OF PARAMETERS OF PARTIALLY SINUSOIDAL FREQUENCY MODEL

SWAGATA NANDI[1] AND DEBASIS KUNDU[2]

ABSTRACT. In this paper, we propose a modification of the multiple sinusoidal model such that periodic data observed with the presence of a trend component can be analyzed. In particular, we work with a linear trend model. But it can be developed similarly for the polynomial trend and the frequency model also. We consider the problem of estimation of frequency and amplitude parameters observed with a linear trend and a stationary linear process. We apply the usual least squares method to the differenced data. It has been proved that the proposed estimators are strongly consistent and asymptotically normally distributed. Extensive simulations have been reported to justify the suitability of the model and the proposed method. One real dataset, the monthly airline passenger data, has been analyzed using the model.

1. INTRODUCTION

The sinusoidal frequency model embedded in noise is an extensively important model because of its wide applicability in many areas of science and engineering. For around last forty years, researchers have approached the model from different point of view mainly motivated by real life problems. It has been observed in many applications that the data exhibit the presence of a trend component as well as a combination of some periodic components. The problem of analyzing such datasets leads to our present work. In this paper, we consider the following model to capture the periodic phenomenon observed with a linear trend;

$$y(t) = a + bt + \sum_{k=1}^{p} [A_k \cos(\omega_k t) + B_k \sin(\omega_k t)] + x(t), \qquad t = 1, \ldots, n+1. \quad (1)$$

1

Here $y(t)$ is the response variable and we observe $\{y(t), t = 1, \dots, n + 1\}$. Here a and b are unknown real numbers and parameters of the linear trend component. The parameters of the sinusoidal frequency component are noted as follows; $A_k, B_k \in \mathcal{R}$ are unknown amplitudes, $\omega_k \in (0, \pi)$ are the unknown frequencies. The number of sinusoidal components present is p and it is assumed to be known in advance. We have taken the initial sample size as $n + 1$ instead of the usual convention as n. The reason behind this assumption will be clear in section 2. The sequence of error random variables $\{x(t)\}$ satisfies the following assumption;

Assumption 1. *The sequence of random variables $\{x(t)\}$ has the following linear structure*

$$x(t) = \sum_{k=-\infty}^{\infty} \alpha(k)e(t - k), \tag{2}$$

where $\{e(t)\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and finite variance σ^2. The arbitrary real valued sequence $\{\alpha(k)\}$ is absolutely summable, that is, it satisfies the following condition:

$$\sum_{k=-\infty}^{\infty} |\alpha(k)| < \infty. \tag{3}$$

In model (1), if b is zero, it is nothing but the usual multiple sinusoidal model (Hannan [2]); Walker [7]). A more general model than model (1) could include a polynomial of degree q, where $2 < q << n$ instead of the linear part $a + bt$. The proposed method (will be discussed in section 2) is also applicable for such a general model, although the analysis and notation become very messy. So the above model belongs to the class of partially nonlinear models. The problem, we address here, is motivated by some real data, where it is evident that a number of periodic components superimposed with a trend (linear) component is present. With the aim of analyzing such datasets, we propose model (1) in this article.

Recently, partially nonlinear models were proposed by Li and Nie [5]. They studied general nonlinear models under the assumption of i.i.d. errors and proposed a weighted least squares approach applied to the differenced data. In Li and Nie [6], the approach is based on profile nonlinear likelihood and linear approximation with local linear regression. We observe that model (1) does not satisfy all the regularity conditions required in Li and Nie [5] and therefore their results can not be directly applied here.

The rest of the paper is organized as follows. We discuss about the estimation method in section 2. The asymptotic properties are provided in section 3. In section 4, we present the simulation results. For illustrative purposes, one real data set has been analyzed in section 5 and we conclude the paper in section 6. All the proofs are provided in the Appendix.

2. ESTIMATING THE UNKNOWN PARAMETERS

In this section, we discuss about the estimation procedure of the unknown parameters. We basically consider the usual least squares estimation method applied to the differenced observations. For simplicity, at this point we assume that $p = 1$, i.e. only one frequency is present and we write the model as

$$y(t) = a + bt + A\cos(\omega t) + B\sin(\omega t) + x(t), \quad t = 1, \ldots, n+1. \tag{4}$$

We work with the first difference series, $y(t+1) - y(t) = z(t)$, say for $t = 1, \ldots n$;

$$
\begin{aligned}
z(t) &= y(t+1) - y(t) \\
&= b + A[\cos(\omega t + \omega) - \cos(\omega t)] + B[\sin(\omega t + \omega) - \sin(\omega t)] + x(t+1) - x(t) \\
&= b - 2A\sin\left(\omega t + \frac{\omega}{2}\right)\sin\left(\frac{\omega}{2}\right) + 2B\cos\left(\omega t + \frac{\omega}{2}\right)\sin\left(\frac{\omega}{2}\right) + x_d(t) \\
&= b - A\sin(\omega)\sin(\omega t) - 2A\sin^2\left(\frac{\omega}{2}\right)\cos(\omega t) + B\sin(\omega)\cos(\omega t) \\
&\qquad\qquad\qquad -2B\sin^2\left(\frac{\omega}{2}\right)\sin(\omega t) + x_d(t). \tag{5}
\end{aligned}
$$

Here $x_d(t) = x(t+1) - x(t)$ is the first difference of $\{x(t)\}$ and satisfies Assumption 1, because

$$
\begin{aligned}
x_d(t) = x(t+1) - x(t) &= \sum_{k=-\infty}^{\infty} \alpha(k)e(t+1-k) - \sum_{k=-\infty}^{\infty} \alpha(k)e(t-k) \\
&= \sum_{k=-\infty}^{\infty} \left(\alpha(k+1) - \alpha(k)\right) e(t-k) = \sum_{k=-\infty}^{\infty} \beta(k)e(t-k),
\end{aligned}
$$

where $\{\alpha(k+1) - \alpha(k)\} = \{\beta(k)\}$. Then $\sum_{k=-\infty}^{\infty} |\beta(k)| < \infty$ and $\{\beta(k)\}$ is absolutely summable as $\{\alpha(k)\}$ is. The unknown a, being a constant in the original model (1) is not going to contribute in the differenced series. As we work with the differenced series $z(t)$ instead of the original observations $y(t)$'s, we start with $n+1$ observations instead of n.

In matrix notation, equation (5) is written as;

$$
\mathbf{Z} = b\mathbf{1} + \mathbf{X}(\omega)\mathbf{D}(\omega)\boldsymbol{\eta} + \mathbf{E}, \tag{6}
$$

where $\mathbf{Z} = (z(1), z(2), \ldots, z(n))^T$, $\mathbf{1} = (1, 1, \ldots, 1)^T$, $\mathbf{E} = (x_d(1), \ldots, x_d(n))^T$, $\boldsymbol{\eta} = (A, B)^T$ and

$$
\mathbf{X}(\omega) = \begin{bmatrix} \cos(\omega) & \sin(\omega) \\ \vdots & \vdots \\ \cos(n\omega) & \sin(n\omega) \end{bmatrix}, \qquad
\mathbf{D}(\omega) = \begin{bmatrix} -2\sin^2(\frac{\omega}{2}) & \sin(\omega) \\ -\sin(\omega) & -2\sin^2(\frac{\omega}{2}) \end{bmatrix}.
$$

In equation (5), we take average over t and then for large n, it is approximated as

$$
\frac{1}{n}\sum_{t=1}^{n} z(t) = b + O(n^{-1}).
$$

Thus for large n, we can estimate b as $\widehat{b} = \frac{1}{n}\sum_{t=1}^{n} z(t)$, which is a consistent estimator of b. Now, plugging $\widehat{b}$ in (6), we denote

$$
\mathbf{Z}^* = \mathbf{Z} - \widehat{b}\mathbf{1} = \mathbf{X}(\omega)\mathbf{D}(\omega)\boldsymbol{\eta} + \mathbf{E}. \tag{7}
$$

Then the LSEs of $\boldsymbol{\eta}$ and ω minimizes the residual sum of squares

$$
Q(\boldsymbol{\eta}, \omega) = \mathbf{E}^T\mathbf{E} = (\mathbf{Z}^* - \mathbf{X}(\omega)\mathbf{D}(\omega)\boldsymbol{\eta})^T(\mathbf{Z}^* - \mathbf{X}(\omega)\mathbf{D}(\omega)\boldsymbol{\eta}). \tag{8}
$$

So for a given ω, the LSE of $\boldsymbol{\eta}$, as a function of ω is

$$\widehat{\boldsymbol{\eta}}(\omega) = \left[\mathbf{D}(\omega)^T\mathbf{X}(\omega)^T\mathbf{X}(\omega)\mathbf{D}(\omega)\right]^{-1}\mathbf{D}(\omega)^T\mathbf{X}(\omega)^T\mathbf{Z}^*. \tag{9}$$

Now using (9) in (8), we have

$$\begin{aligned}
R(\omega) &= Q(\widehat{\boldsymbol{\eta}}(\omega), \omega) \\
&= \mathbf{Z}^{*T}\left(\mathbf{I} - \mathbf{X}(\omega)\mathbf{D}(\omega)\left[\mathbf{D}(\omega)^T\mathbf{X}(\omega)^T\mathbf{X}(\omega)\mathbf{D}(\omega)\right]^{-1}\mathbf{D}(\omega)^T\mathbf{X}(\omega)^T\right)\mathbf{Z}^* \\
&= \mathbf{Z}^{*T}\left(\mathbf{I} - \mathbf{P_{XD}}(\omega)\right)\mathbf{Z}^*, \qquad \text{(say)}. \tag{10}
\end{aligned}$$

The matrix $\mathbf{P_{XD}}(\omega)$ as a function of ω, is the projection matrix of $\mathbf{X}(\omega)\mathbf{D}(\omega)$ and hence is idempotent and symmetric. Therefore, $R(\omega)$ takes the form given in equation (10). The LSE of ω, $\widehat{\omega}$ is obtained by maximizing $\mathbf{Z}^{*T}\mathbf{P_{XD}}(\omega)\mathbf{Z}^*$ and then plugging in $\widehat{\omega}$ into (9), $\boldsymbol{\eta}$ is estimated as $\widehat{\boldsymbol{\eta}}(\widehat{\omega})$.

Remark 1. We note that for efficient estimation of the sinusoidal frequency parameters, the maximization of $\mathbf{Z}^{*T}\mathbf{P_{XD}}(\omega)\mathbf{Z}^*$ works, which is based on the mean corrected first differenced series. This is due to the reason that the rate of convergence of $\widehat{b}$ is $o(n^{-1})$.

Remark 2. In Introduction, we have mentioned that the proposed method can be extended in the general model when the linear trend part is replaced by a higher degree polynomial of unknown coefficients. Suppose we consider a model of p degree polynomial superimposed with a single sinusoid. Then to define $z^*(t)$, one is required to difference the observed data p times. In such a scenario, ω can be obtained by maximizing a similar criterion function as $R(\omega)$, defined in (10). Then the entries of $\mathbf{D}(\omega)$ matrix will be more complicated functions of ω.

3. Asymptotic Properties

In this section, we discuss about the theoretical properties of the LSEs of A, B and ω. We denote $\boldsymbol{\theta} = (A, B, \omega) = (\boldsymbol{\eta}, \omega)$ and let $\boldsymbol{\theta}^0$ be the true value of $\boldsymbol{\theta}$ and $\widehat{\boldsymbol{\theta}}$ be the LSE of

$\boldsymbol{\theta}^0$. In the following, we state the consistency property of $\widehat{\boldsymbol{\theta}}$; the proof is provided in the appendix.

THEOREM 3.1. Let the true parameter vector $\boldsymbol{\theta}^0 = (A^0, B^0, \omega^0)$ be an interior point of the parameter space $(-\infty, \infty) \times (-\infty, \infty) \times (0, \pi)$ and $A^{0^2} + B^{0^2} > 0$. If the error random variables $x(t)$ satisfy Assumption 1, then $\widehat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

Now we compute the asymptotic joint distribution of the LSE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}^0$. We use $Q'(\boldsymbol{\theta})$ and $Q''(\boldsymbol{\theta})$ to denote respectively the 1×3 vector of first derivatives and the 3×3 matrix of second derivatives of $Q(\boldsymbol{\theta})$, defined in (8). Expanding $Q'(\widehat{\boldsymbol{\theta}})$ around the true parameter value $\boldsymbol{\theta}^0$, in a Taylor series, we obtain

$$Q'(\widehat{\boldsymbol{\theta}}) - Q'(\boldsymbol{\theta}^0) = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)Q''(\bar{\boldsymbol{\theta}}), \tag{11}$$

where $\bar{\boldsymbol{\theta}}$ is a point on the line joining the points $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Suppose $\mathbf{D_A}$ is a 3×3 diagonal matrix $\mathbf{D_A} = diag\{n^{-1/2}, n^{-1/2}, n^{-3/2}\}$. The diagonal entries of $\mathbf{D_A}$ correspond to the rates of convergence of the LSEs of A, B and ω. Since $Q'(\widehat{\boldsymbol{\theta}}) = 0$, (11) can be written as

$$(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D_A}^{-1} = -\left[\mathbf{Q'}(\boldsymbol{\theta}^0)\mathbf{D_A}\right]\left[\mathbf{D_A}\mathbf{Q''}(\bar{\boldsymbol{\theta}})\mathbf{D_A}\right]^{-1}, \tag{12}$$

as $\left[\mathbf{D_A}\mathbf{Q''}(\bar{\boldsymbol{\theta}})\mathbf{D_A}\right]$ is an invertible matrix *a.e.* for large n. Using Theorem 1, it follows that $\widehat{\boldsymbol{\theta}}$ converges *a.e.* to $\boldsymbol{\theta}^0$ and so $\bar{\boldsymbol{\theta}} \longrightarrow \boldsymbol{\theta}^0$ *a.e.* Since each element of $Q''(\boldsymbol{\theta})$ is a continuous function of $\boldsymbol{\theta}$, we have

$$\lim_{n \to \infty}[\mathbf{D_A}Q''(\bar{\boldsymbol{\theta}})\mathbf{D_A}] = \lim_{n \to \infty}[\mathbf{D_A}Q''(\boldsymbol{\theta}^0)\mathbf{D_A}] = \boldsymbol{\Sigma}(\boldsymbol{\theta}^0) \; (say). \tag{13}$$

The 1×3 random vector $\left[\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D_A}\right]$ is

$$
\begin{bmatrix}
\frac{2}{\sqrt{n}} \sum_{t=1}^{n} x_d(t) \left[2\sin^2(\frac{\omega^0}{2})\cos(\omega^0 t) + \sin(\omega^0)\sin(\omega^0 t) \right] \\[2mm]
\frac{2}{\sqrt{n}} \sum_{t=1}^{n} x_d(t) \left[2\sin^2(\frac{\omega^0}{2})\sin(\omega^0 t) - \sin(\omega^0)\cos(\omega^0 t) \right] \\[2mm]
\frac{2}{n^{\frac{3}{2}}} \sum_{t=1}^{n} t x_d(t) \left[-2A^0\sin^2(\frac{\omega^0}{2}) + B^0\sin(\omega^0)\sin(\omega^0 t) + \right. \\
\left. A^0\sin(\omega^0)\cos(\omega^0 t) + 2B^0\sin^2(\frac{\omega^0}{2})\cos(\omega^0 t) \right]
\end{bmatrix},
$$

where $x_d(t) = x(t+1) - x(t) = \sum\limits_{k=-\infty}^{\infty} \beta(k)e(t-k)$, $\beta(k) = \alpha(k+1) - \alpha(k)$ as defined in section 2. Using a central limit theorem (see Fuller[1], page 251) for stochastic processes, it follows that;

$$
Q'(\boldsymbol{\theta}^0)\mathbf{D_A} \xrightarrow{d} \mathcal{N}_3(\mathbf{0}, \mathbf{G}(\boldsymbol{\theta}^0)),
$$

as $n \to \infty$. It can be shown that

$$
\boldsymbol{\Sigma}(\boldsymbol{\theta}^0) = 2(1-\cos(\omega^0))
\begin{bmatrix}
1 & 0 & B^0/2 \\
0 & 1 & -A^0/2 \\
B^0/2 & -A^0/2 & \frac{1}{3}(A^{0^2}+B^{0^2})
\end{bmatrix},
$$

and then the matrix $\mathbf{G}(\boldsymbol{\theta}^0)$ takes the following form;

$$
\mathbf{G}(\boldsymbol{\theta}^0) = 2\sigma^2 c_\beta(\omega^0)\boldsymbol{\Sigma}(\boldsymbol{\theta}^0)
$$

with

$$
c_\beta(\omega) = \left| \sum_{k=-\infty}^{\infty} \beta(k)e^{-i\omega k} \right|^2. \tag{14}
$$

Therefore,

$$
(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D_A}^{-1} \xrightarrow{d} \mathcal{N}_3(\mathbf{0}, \Sigma(\boldsymbol{\theta}^0)^{-1}\mathbf{G}(\boldsymbol{\theta}^0)\Sigma(\boldsymbol{\theta}^0)^{-1}), \tag{15}
$$

where

$$
\begin{aligned}
\Sigma(\boldsymbol{\theta}^0)^{-1}\mathbf{G}(\boldsymbol{\theta}^0)\Sigma(\boldsymbol{\theta}^0)^{-1} &= 2\sigma^2 c_\beta(\omega^0)\boldsymbol{\Sigma}(\boldsymbol{\theta}^0)^{-1} \\
&= \frac{\sigma^2 c_\beta(\omega^0)}{(1-\cos(\omega^0))(A^{0^2}+B^{0^2})}
\begin{bmatrix}
A^{0^2}+4B^{0^2} & -3A^0 B^0 & -6B^0 \\
-3A^0 B^0 & 4A^{0^2}+B^{0^2} & 6A^0 \\
-6B^0 & 6A^0 & 12
\end{bmatrix}.
\end{aligned}
$$

Remark 3. We note that $c_\beta(.)$ as a function of ω is related to the spectral density function of the stationary linear process $x_d(t) = \sum\limits_{k=-\infty}^{\infty} \beta(k)e(t-k)$. We write the spectrum of $\{x_d(t)\}$ as $f_{x_d}(\omega)$, then $\dfrac{\sigma^2}{2\pi}c_\beta(\omega) = f_{x_d}(\omega)$.

Remark 4. We would like to compare the asymptotic variances of the LSEs of unknown parameters of sinusoidal model observed in presence of a linear trend (model (4)) and the asymptotic variances of LSEs of the same model when observed without any trend component (i.e. $a = b = 0$ in model (4)). In the former case it is $\widehat{\boldsymbol{\theta}}$ considered in this paper and we denote $\widehat{\boldsymbol{\theta}}^*$ in the later case. Then $\mathrm{Var}(\widehat{\boldsymbol{\theta}}) = \dfrac{c_\beta(\omega^0)}{2(1 - \cos(\omega^0))c_\alpha(\omega^0)}\mathrm{Var}(\widehat{\boldsymbol{\theta}}^*)$, asymptotically, where $c_\alpha(\omega)$ is the same function of ω defined in (14), using $\{\alpha(k)\}$ instead of $\{\beta(k)\}$.

4. Multiple Sinusoids

In this section, we first provide the least squares method of estimation for the unknown parameters of the model (1) and then provide the asymptotic properties of the LSEs. Without loss of generality, we assume that

$$A_1^{0^2} + B_1^{0^2} > A_2^{0^2} + B_2^{0^2} > \cdots > A_p^{0^2} + B_p^{0^2} > 0. \tag{16}$$

The method is basically the one discussed in section 2 applied sequentially. We take the first difference and define $z(t) = y(t + 1) - y(t)$. Following the same techniques as the single component model (4), the matrix equation (6) takes the form

$$\mathbf{Z} = b\mathbf{1} + \mathbf{X_p}(\boldsymbol{\omega})\mathbf{D_p}(\boldsymbol{\omega})\boldsymbol{\psi} + \mathbf{E}. \tag{17}$$

The first difference vector $\mathbf{Z}$, $\mathbf{1}$ and the error vector $\mathbf{E}$ are same as defined in section 2; $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_p)$, $\boldsymbol{\psi} = (\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_p)^T$, $\boldsymbol{\eta}_j = (A_j, B_j)^T$ and $\mathbf{X_p}(\boldsymbol{\omega})$ and $\mathbf{D_p}(\boldsymbol{\omega})$ take the

following form;

$$\mathbf{X_p}(\boldsymbol{\omega}) = \begin{bmatrix} \mathbf{X}(\omega_1) & \mathbf{X}(\omega_2) & \cdots & \mathbf{X}(\omega_p) \end{bmatrix}, \qquad \mathbf{D_p}(\boldsymbol{\omega}) = \begin{bmatrix} \mathbf{D}(\omega_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}(\omega_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D}(\omega_p) \end{bmatrix}.$$

The matrices $\mathbf{X}(\omega_j)$ and $\mathbf{D}(\omega_j)$ are same as those of $\mathbf{X}(\omega)$ and $\mathbf{D}(\omega)$, defined in section 2, replacing ω by ω_j. We note that (17) can also be written as

$$\mathbf{Z} = b\mathbf{1} + \sum_{j=1}^{p} \mathbf{X}(\omega_j)\mathbf{D}(\omega_j)\boldsymbol{\eta}_j + \mathbf{E}. \tag{18}$$

We estimate b as before and obtain $\mathbf{Z}^* = \mathbf{Z} - \widehat{b}\mathbf{1}$. Then the LSEs of $\boldsymbol{\omega}$ and $\boldsymbol{\eta}$ are obtained by minimizing the residual sum of squares;

$$U(\boldsymbol{\eta}, \boldsymbol{\omega}) = \left(\mathbf{Z}^* - \sum_{j=1}^{p} \mathbf{X}(\omega_j)\mathbf{D}(\omega_j)\boldsymbol{\eta}_j \right)^T \left(\mathbf{Z}^* - \sum_{j=1}^{p} \mathbf{X}(\omega_j)\mathbf{D}(\omega_j)\boldsymbol{\eta}_j \right). \tag{19}$$

Let $\boldsymbol{\psi}^0$ and $\boldsymbol{\omega}^0$ be the true values of $\boldsymbol{\psi}$ and $\boldsymbol{\omega}$ respectively and $\widehat{\boldsymbol{\psi}}$ and $\widehat{\boldsymbol{\omega}}$ denote the LSEs of $\boldsymbol{\psi}^0$ and $\boldsymbol{\omega}^0$ respectively. Then for a given $\boldsymbol{\omega}$, $\widehat{\boldsymbol{\psi}}$ can be directly written from (17) as

$$\widehat{\boldsymbol{\psi}}(\boldsymbol{\omega}) = \left[\mathbf{D_p}(\boldsymbol{\omega})^T \mathbf{X_p}(\boldsymbol{\omega})^T \mathbf{X_p}(\boldsymbol{\omega}) \mathbf{D_p}(\boldsymbol{\omega}) \right]^{-1} \mathbf{D_p}(\boldsymbol{\omega})^T \mathbf{X_p}(\boldsymbol{\omega})^T \mathbf{Z}^*. \tag{20}$$

Using the fact that $\mathbf{X}(\omega_j)$ and $\mathbf{X}(\omega_k)$, $k \neq j$, are orthogonal matrices, we have $\frac{1}{n}\mathbf{D}(\omega_j)^T\mathbf{X}(\omega_j)^T\mathbf{X}(\omega_k)\mathbf{D}(\omega_k) = \mathbf{0}$ for large n. Therefore, (20) reduces to

$$\widehat{\boldsymbol{\psi}}(\boldsymbol{\omega}) = \begin{bmatrix} \widehat{\boldsymbol{\eta}}_1(\omega_1) \\ \vdots \\ \widehat{\boldsymbol{\eta}}_p(\omega_p) \end{bmatrix} = \begin{bmatrix} \left[\mathbf{D}(\omega_1)^T\mathbf{X}(\omega_1)^T\mathbf{X}(\omega_1)\mathbf{D}(\omega_1)\right]^{-1}\mathbf{D}(\omega_1)^T\mathbf{X}(\omega_1)^T\mathbf{Z}^* \\ \vdots \\ \left[\mathbf{D}(\omega_p)^T\mathbf{X}(\omega_p)^T\mathbf{X}(\omega_p)\mathbf{D}(\omega_p)\right]^{-1}\mathbf{D}(\omega_p)^T\mathbf{X}(\omega_p)^T\mathbf{Z}^* \end{bmatrix}, \tag{21}$$

which is the same as deduced in (9), using $\boldsymbol{\eta}_j$ and ω_j instead of $\boldsymbol{\eta}$ and ω respectively. Now plugging in $\widehat{\boldsymbol{\psi}}(\boldsymbol{\omega})$ in (19), the residual sum of squares can be written as

$$\begin{aligned} S(\boldsymbol{\omega}) = U(\widehat{\boldsymbol{\eta}}(\omega), \omega) &= \left[\mathbf{Z}^* - \sum_{j=1}^{p} \mathbf{P_{XD}}(\omega_j)\,\mathbf{Z}^* \right]^T \left[\mathbf{Z}^* - \sum_{j=1}^{p} \mathbf{P_{XD}}(\omega_j)\,\mathbf{Z}^* \right] \\ &= \mathbf{Z}^{*T}\mathbf{Z}^* - \sum_{j=1}^{p} \mathbf{Z}^{*T}\mathbf{P_{XD}}(\omega_j)\mathbf{Z}^*, \end{aligned}$$

where $\mathbf{P_{XD}}(\omega)$ has been defined in section 2. Then LSE of $\omega_1, \omega_2, \ldots$ and ω_p are computed by maximizing $\mathbf{Z}^{*T}\mathbf{P_{XD}}(\omega_j)\mathbf{Z}^*$ sequentially.

Now we discuss the theoretical properties of the LSEs of the unknown parameters $\boldsymbol{\psi}$ and $\boldsymbol{\omega}$. We write $\boldsymbol{\xi} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p)$. $\boldsymbol{\theta}_j = (A_j, B_j, \omega_j)$. The least squares estimators of the parameters are obtained by minimizing the objective function, $U(\boldsymbol{\xi}) = U(\boldsymbol{\eta}, \boldsymbol{\omega})$, defined in (19). Let $\widehat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}^0$ denote the least squares estimator and the true value of $\boldsymbol{\xi}$. The consistency of $\widehat{\boldsymbol{\xi}}$ follows as did the consistency of $\widehat{\boldsymbol{\theta}}$ in section 3, considering the parameter vector $\boldsymbol{\xi}$. We state the asymptotic distribution of $\widehat{\boldsymbol{\xi}}$. The proof involves use of multiple Taylor series expansion and routine calculations along the same line provided in section 3.

We define a diagonal matrix $\mathbf{D_{Ap}}$ of order $3p$ corresponding to the rates of convergences of $\widehat{\boldsymbol{\xi}}$;

$$
\mathbf{D_{Ap}} = \begin{bmatrix} \mathbf{D_A} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D_A} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D_A} \end{bmatrix},
$$

where $\mathbf{D_A}$ is defined in section 3. Then under Assumption 1 and condition (16), as $n \to \infty$

$$
(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0)\mathbf{D_p}^{-1} \xrightarrow{d} \mathcal{N}_{3p}\left(\mathbf{0}, \boldsymbol{\Gamma}(\boldsymbol{\xi}^0)^{-1}\mathbf{H}(\boldsymbol{\xi}^0)\boldsymbol{\Gamma}(\boldsymbol{\xi}^0)^{-1}\right). \tag{22}
$$

The $3p \times 3p$ matrices $\boldsymbol{\Gamma}(\boldsymbol{\xi}^0)$ and $\mathbf{H}(\boldsymbol{\xi}^0)$ take the following forms;

$$
\boldsymbol{\Gamma}(\boldsymbol{\xi}) = \begin{bmatrix} \boldsymbol{\Sigma}(\boldsymbol{\theta}_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}(\boldsymbol{\theta}_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma}(\boldsymbol{\theta}_p) \end{bmatrix}, \quad \mathbf{H}(\boldsymbol{\xi}) = \begin{bmatrix} c_\beta(\omega_1)\boldsymbol{\Sigma}(\boldsymbol{\theta}_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & c_\beta(\omega_2)\boldsymbol{\Sigma}(\boldsymbol{\theta}_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & c_\beta(\omega_p)\boldsymbol{\Sigma}(\boldsymbol{\theta}_p) \end{bmatrix}.
$$

The matrix $\boldsymbol{\Sigma}(.)$ and $c_\beta(.)$ as functions of ω have been defined in previous section. Then (22) reduces to

$$
(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k^0)\mathbf{D_A}^{-1} \xrightarrow{d} \mathcal{N}_3\left(\mathbf{0}, 2\sigma^2 c_\beta(\omega_j^0)\boldsymbol{\Sigma}(\boldsymbol{\theta}_k^0)^{-1}\right),
$$

and $\widehat{\boldsymbol{\theta}}_k$ and $\widehat{\boldsymbol{\theta}}_j$, $j \neq k$ are asymptotically independently distributed.

5. Numerical Experiment

In this section, we present the numerical experiment results based on simulation involving synthesized data vector. We consider model (1) with $p = 1$ (Model 1) and $p = 2$ (Model 2). Data are generated using the following true parameter values:

Model 1: $a = 3.0$, $b = .8$, $A_1 = 1$, $B_1 = 1$, $\omega_1 = 2.5$.

Model 2: $a = 3.0$, $b = .8$, $A_1 = 2$, $B_1 = 2$, $\omega_1 = .5$, $A_2 = 1$, $B_2 = 1$,

$\omega_2 = 2.5$.

The sequence of noise random variables $\{x(t)\}$ is an moving average process of order one such that

$$x(t) = \epsilon(t-1) + .75\ \epsilon(t), \qquad \epsilon(t) \sim \mathcal{N}(0, \sigma^2).$$

The parameter values in case of Model 2 are selected according to condition (16). We have reported the results for different values of the error variance $\sigma^2 = .5$, 1.0, 1.5 and 2.0 and two sample sizes $n = 50$ and 100. To begin with, we have generated a sample of size $(n+1)$ with a predefined set of model parameters as mentioned above. As we are mainly interested in estimating the parameters of the periodic component, we remove the trend part $a + bt$ by considering the mean-corrected first order differenced series $\mathbf{Z}^*$, as described in section 2. Then we estimate the non-linear parameter (ω_1 in case of Model 1) by minimizing $R(\omega)$, defined in (10) or by maximizing $\mathbf{Z}^{*\mathbf{T}}\mathbf{P_{XD}}(\omega)\mathbf{Z}^*$ with respect to ω. Once we have the estimate of the non-linear parameter ω, the linear parameters A and B are estimated according to (9). In section 3, we have developed the joint asymptotic distribution of the LSEs of the unknown parameters of the periodic component, which can be used to obtain approximate confidence bounds of the LSE of each parameter. Therefore, we use (15) to obtain approximate confidence bounds of each parameter. In case of Model 2, we have to use the sequential estimation technique as described in section 4. As $A_1^2 + B_1^2 > A_2^2 + B_2^2$, we first estimate ω_1 exactly in the same

way as described in case of Model 1. Then we remove the effect of the first component from the mean corrected first order differenced observations and again we apply the same procedure step by step on the remaining quantity to estimate the LSEs of the second component. We replicate the complete procedure for 1000 times in case of both Model 1 and Model 2 and in each case we estimate the average estimate (AVEST) and the mean squared error (MSE). Now we would like to compute the confidence intervals for different parameter estimators and then it is required to estimate σ^2 as well as $c_\beta(\omega)$. Although, we can not estimate them separately, it is possible to estimate $\sigma^2 c_\beta(\omega)$, which is actually needed to estimate the confidence intervals. It can be shown that

$$\sigma^2 c_\beta(\omega) = E\left(\frac{1}{n}\left|\sum_{t=1}^{n} x_d(t) e^{-i\omega t}\right|^2\right),$$

that is, the expected value of the periodogram function (defined in section 6) of $x_d(t)$, the error random variables present in $z^*(t)$, at the respective frequency. For numerical calculations, we use the method of averaging the periodogram of the estimated noise over a window $(-L, L)$. We have reported results with $L = 10$. Now we estimate the 95% coverage probability (COVP) by calculating the proportion covering the true parameter value in case of each parameter estimate and the average length of the confidence intervals (AVLEN) for each parameter estimator over these 1000 replications. We also report the asymptotic variance (ASYMV) and the expected length of interval (EXPLEN) at 95% nominal level to compare the MSE and AVLEN for each parameter estimators of the experiment considered here .

We observe that average estimates are quite good. This is reflected in the fact that the average biases are small in absolute value. The MSEs are reasonably small and quite close to the asymptotic variances. According to the asymptotic distributions of LSEs, the non-linear parameters ω's have faster rate of convergence than the linear parameters, which is clearly observed in simulation results. Following it, the MSEs are in decreasing order of $(A, B), \omega$. Similar findings are found in average lengths of confidence intervals.

TABLE 1. The average estimates, mean squared errors, asymptotic variances, average and expected lengths of the confidence intervals and coverage probabilities of the LSEs of parameters of Model 1 when the sample size $n + 1 = 51$.

σ^2 ↓		AVEST	MSE	ASYMV	AVLEN	EXPLEN	COV
.5	A	.993687	2.23006e-2	1.49600e-2	.544393	.479460	.923
	B	.995867	1.80843e-2	1.49600e-2	.542833	.479460	.946
	ω	2.500097	1.86690e-5	1.38040e-5	1.65696e-2	1.45642e-2	.936
1.0	A	.989012	4.46874e-2	2.99201e-2	.768153	.678059	.921
	B	.990792	3.62924e-2	2.99201e-2	.765307	.678059	.945
	ω	2.500134	3.77115e-5	2.76079e-5	2.34353e-2	2.05970e-2	.934
1.5	A	.984408	6.71821e-2	4.48801e-2	.938760	.830449	.918
	B	.985632	5.46430e-2	4.48801e-2	.934788	.830449	.944
	ω	2.500175	5.71561e-5	4.14119e-5	2.87032e-2	2.52260e-2	.932
2.0	A	.979827	8.97619e-2	5.98402e-2	1.081784	.958920	.918
	B	.980374	7.32059e-2	5.98402e-2	1.076914	.958920	.943
	ω	2.500205	7.70549e-5	5.52159e-5	3.31425e-2	2.91285e-2	.931

The average length is close to the expected length of the interval in case of each parameter estimators. Moreover, the biases, MSEs and average lengths of intervals increase as the error variance increases and they decrease as the sample size increases for all parameter estimates. The coverage probabilities are quite close to the nominal level in case of Model 1. In case of Model 2, when the sample size is 51, the coverage probabilities of LSEs, specifically for linear parameters in some occasions, are not that good. However, this is not observed if the sample size increases to 101. Therefore, the numerical experiment conducted here, suggests that the asymptotic results can be used in small and moderate size samples.

6. DATA ANALYSIS

In this section, we present the analysis of a classical real data. The data represent the monthly international airline passengers during January 1953 to December 1960 and are collected from the Time Series Data Library of Hyndman[3]. The data is plotted in Figure

TABLE 2. The average estimates, mean squared errors, asymptotic variances, average and expected lengths of the confidence intervals and coverage probabilities of the LSEs of parameters of Model 1 when the sample size $n + 1 = 101$.

σ^2 ↓		AVEST	MSE	ASYMV	AVLEN	EXPLEN	COV
0.5	A	.997993	9.78627e-3	7.55408e-3	.364015	.340704	.928
	B	.997368	8.61015e-3	7.55408e-3	.364095	0.340704	.940
	ω	2.500019	2.13123e-6	1.77726e-6	5.59446e-3	5.22590e-3	.936
1.0	A	.996272	1.96033e-2	1.51082e-2	.514085	.481828	.928
	B	.994929	1.72620e-2	1.51082e-2	.514281	.481828	.941
	ω	2.500020	4.28425e-6	3.55451e-6	7.91345e-3	7.39054e-3	.936
1.5	A	.994496	2.94643e-2	2.26622e-2	.628811	.590116	.926
	B	.992572	2.59580e-2	2.26622e-2	.629101	.590116	.941
	ω	2.500028	6.46187e-6	5.33177e-6	9.69365e-3	9.05153e-3	.936
2.0	A	.992694	3.93894e-2	3.02163e-2	.725188	.681408	.925
	B	.990215	3.47182e-2	3.02163e-2	.725569	.681408	.941
	ω	2.500034	8.67335e-6	7.10902e-6	1.11949e-2	1.04518e-2	.935

TABLE 3. The average estimates, mean squared errors, asymptotic variances, average and expected lengths of the confidence intervals and coverage probabilities of the LSEs of parameters of Model 2 when the sample size $n + 1 = 51$ and the error variance $\sigma^2 = 0.5$ & 1.0.

σ^2 ↓		AVEST	MSE	ASYMV	AVLEN	EXPLEN	COV
0.5	A_1	1.017310	2.20221e-2	1.49600e-2	.536868	.479460	.913
	B_1	.978581	1.90324e-2	1.49600e-2	.548365	.479460	.945
	ω_1	2.499309	1.87460e-5	1.38040e-5	1.64852e-2	1.45642e-2	.932
	A_2	1.993197	.171865	.185048	1.94738	1.68627	.971
	B_2	1.964031	.183684	.185048	1.96977	1.68627	.970
	ω_2	.499926	4.75845e-5	4.26869e-5	2.98993e-2	2.56114e-2	.958
1.0	A_1	1.012461	4.37051e-2	2.99201e-2	.757671	.678059	.913
	B_1	.973674	3.78916e-2	2.99201e-2	.772731	.678059	.944
	ω_1	2.499351	3.73501e-5	2.76079e-5	2.33127e-2	2.05970e-2	.932
	A_2	1.975675	.340850	.370095	2.736113	2.384749	.970
	B_2	1.942525	.367370	.370095	2.775661	2.384749	.968
	ω_2	.499966	9.46144e-5	8.53737e-5	4.21998e-2	3.62200e-2	.957

TABLE 4. The average estimates, mean squared errors, asymptotic variances, average and expected lengths of the confidence intervals and coverage probabilities of the LSEs of parameters of Model 2 when the sample size $n + 1 = 51$ and the error variance $\sigma^2 = 1.5$ & 2.0.

σ^2 ↓		AVEST	MSE	ASYMV	AVLEN	EXPLEN	COV
1.5	A_1	1.007751	6.55065e-2	4.48801e-2	.926239	.830449	.911
	B_1	.968628	5.68631e-2	4.48801e-2	.943545	.830449	.944
	ω_1	2.499388	5.633889e-5	4.14119e-5	2.85530e-2	2.52260e-2	.932
	A_2	1.957266	.509014	.555143	3.335896	2.920709	.971
	B_2	1.922550	.549989	.555143	3.389328	2.920709	.967
	ω_2	.500018	1.41471e-4	1.28061e-4	5.15889e-2	4.43602e-2	.957
2.0	A_1	1.002998	8.74218e-2	5.98402e-2	1.067751	.958920	.910
	B_1	.963560	7.59975e-2	5.98401e-2	1.086628	.958920	.942
	ω_1	2.499417	7.57711e-5	5.52159e-5	3.29702e-2	2.91285e-2	.932
	A_2	1.938177	.678415	.740190	3.839300	3.372545	.970
	B_2	1.902366	.735248	.740190	3.904061	3.372545	.965
	ω_2	.500091	1.90176e-4	1.70747e-4	5.94594e-2	5.12228e-2	.957

TABLE 5. The average estimates, mean squared errors, asymptotic variances, average and expected lengths of the confidence intervals and coverage probabilities of the LSEs of parameters of Model 2 when the sample size $n + 1 = 101$ and the error variance $\sigma^2 = 0.5$ & 1.0.

σ^2 ↓		AVEST	MSE	ASYMV	AVLEN	EXPLEN	COV
0.5	A_1	1.008961	9.77201e-3	7.55408e-3	.362109	.340704	.922
	B_1	.991908	8.759064e-3	7.554083e-3	.365729	.340704	.941
	ω_1	2.499879	2.11115e-6	1.77726e-6	5.57687e-3	5.22590e-3	.940
	A_2	1.991753	7.71350e-2	9.34398e-2	1.156615	1.198263	.948
	B_2	1.993350	8.08795e-2	9.34398e-2	1.157121	1.198263	.945
	ω_2	.500009	4.98537e-6	5.49592e-6	8.88821e-3	9.18981e-3	.943
1.0	A_1	1.00721	1.94695e-2	1.51082e-2	.511389	.481828	.923
	B_1	.989493	1.75315e-2	1.51082e-2	.516527	.481828	.940
	ω_1	2.499879	4.22751e-6	3.55451e-6	7.88806e-3	7.39054e-3	.938
	A_2	1.981957	.154165	.186880	1.631052	1.694600	.947
	B_2	1.987423	.161701	.186880	1.631096	1.694600	.944
	ω_2	.500029	9.98189e-6	1.09918e-5	1.25544e-2	1.29963e-2	.943

TABLE 6. The average estimates, mean squared errors, asymptotic variances, average and expected lengths of the confidence intervals and coverage probabilities of the LSEs of parameters of Model 2 when the sample size $n + 1 = 101$ and the error variance $\sigma^2 = 1.5$ & 2.0.

$\sigma^2 \downarrow$		AVEST	MSE	ASYMV	AVLEN	EXPLEN	COV
1.5	A_1	1.005392	2.92124e-2	2.26622e-2	.625545	.590116	.921
	B_1	.987164	2.63547e-2	2.26622e-2	.631792	.590116	.942
	ω_1	2.499884	6.37116e-6	5.33177e-6	9.66238e-3	9.05153e-3	.936
	A_2	1.972361	.231457	.280320	1.99293	2.075452	.946
	B_2	1.981038	.242808	.280320	1.992312	2.07545	.942
	ω_2	.500047	1.50249e-5	1.64878e-5	1.53623e-2	1.59172e-2	.942
2.0	A_1	1.003557	3.90002e-2	3.02163e-2	.721470	.681408	.920
	B_1	.984851	3.52095e-2	3.02163e-2	.728625	.681408	.940
	ω_1	2.499895	8.53656e-6	7.10903e-6	1.11588e-2	1.04518e-2	.935
	A_2	1.962656	.309024	.373759	2.29668	2.39653	.945
	B_2	1.974593	.324063	.373760	2.29491	2.39653	.939
	ω_2	.500065	2.01116e-5	2.19837e-5	1.77249e-2	1.83796e-2	.941

1. We notice that the variance is not constant. In the aim of stabilizing the variance, we apply the log transform to the data which is plotted in Figure 2. Now we observe that the variance is approximately constant and a prominent linear trend is present with several periodic components. Our aim is to estimate the periodic components. So we take $z^*(t)$, the mean corrected first difference series of the log transform data. We present it in Figure 3. Now to estimate the frequency components, or more specifically, the frequency parameters ω_k^0 and the linear parameters A_k^0 and B_k^0, first we need to initialize the frequencies. So we plot the periodogram function of $z^*(t)$, $t = 1, \ldots n$ over a fine grid of $(0, \pi)$ in Figure 4 to estimate the number of frequencies present and their initial estimates. The periodogram function is defined by

$$I(\omega) = \frac{1}{n} \left| \sum_{t=1}^{n} z^*(t) e^{-i\omega t} \right|^2 .$$

The periodogram function is asymptotically equal to $\mathbf{Z}^{*T} \mathbf{P_{XD}}(\omega) \mathbf{Z}^*$. There are six peaks corresponding to six dominating frequencies in the periodogram plot and we obtain the initial estimates of $\omega_1, \omega_2 \ldots$ one by one using the sequential procedure. Once the six
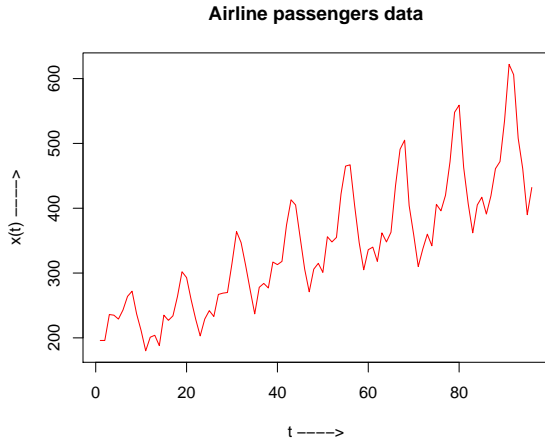
FIGURE 1. The observed airline passengers data.
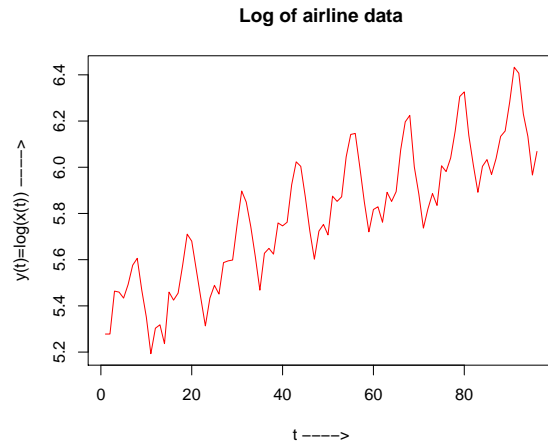

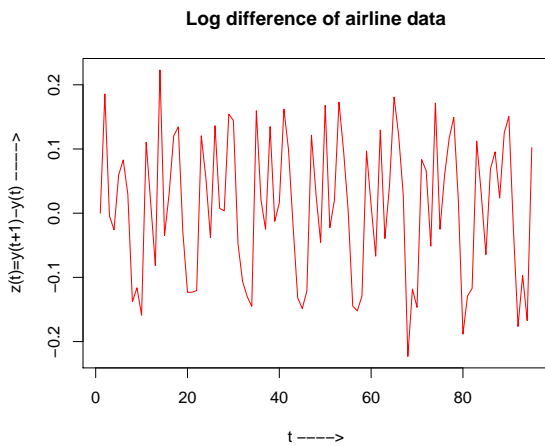FIGURE 2. The logarithm of the observed data.

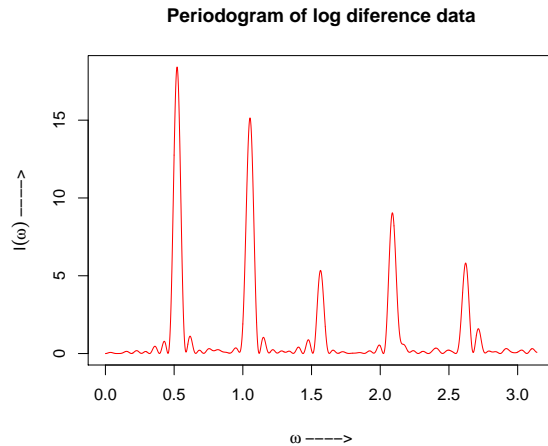FIGURE 3. The first difference values of the logarithmic data.

FIGURE 4. The periodogram function of the log-difference data.

periodic components corresponding to six visible peaks are removed, we again study the periodogram of the remaining series and there is still another significant frequency present. So finally we have estimated $\widehat{p}$ as seven and plugging in the estimated parameters, we extracted the deterministic periodic signal. This fitted series is presented in Figure 5 along with the mean corrected log difference data. They match quite well.
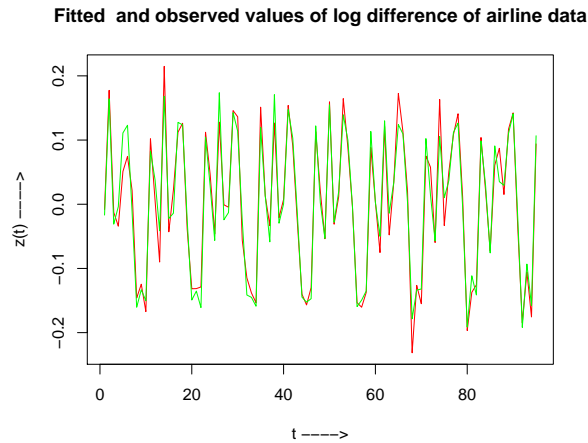
Fitted and observed values of log difference of airline data

FIGURE 5. The fitted values (green) along with the log difference data (red).

7. CONCLUSIONS

In this paper we have introduced a new model - the multiple frequency model observed with a linear trend and second-order stationary errors. We have mentioned that a similar method can be developed in case of a polynomial trend instead of a linear trend. We aim to estimate the unknown parameters involved in the periodic components and are not interested in estimating the trend part. We have used the usual least squares method to estimate the parameters of interest. We have shown that the proposed estimators of the frequency and the amplitude parameters are strongly consistent. The asymptotic distribution has come out as multivariate normal under the assumption that the error random variables satisfy the condition of a stationary linear process. We have conducted extensive simulations based on synthesized data to check whether the asymptotic results can be used for finite moderate size samples. A real dataset has been analyzed using the model and outcome is quite satisfactory.

REFERENCES

[1] Fuller W. A. (1976), *Introduction to Statistical Time Series*, John Wiley and Sons, New York.

[2] Hannan, E.J. (1973) "The estimation of frequency", *Journal of Applied Probability*, Vol. 10, 510-519.

[3] Hyndman, R.J. (n.d.) Time Series Data Library, http://www.robhyndman.info/TSDL. Accessed on 01.01.2008.

[4] Kundu, D. (1997), "Asymptotic properties of the least squares estimators of sinusoidal signals", *Statistics*, Vol. 30, 221-238.

[5] Li, R. and Nie, L. (2007), "A new estimation procedure for a partially nonlinear model via a mixed-effects approach", *The Canadian Journal of Statistics*, Vol. 35, No. 3, 399-411.

[6] Li, R. and Nie, L. (2008), "Efficient statistical inference procedures for partially nonlinear models and their applications", *Biometrics*, Vol. 64, 904-911.

[7] Walker, A.M. (1971), "On the estimation of harmonic components in a time series with stationary residuals", *Biometrika*, Vol. 58, 21-36.

[8] Wu, C.F.J. (1981), "Asymptotic theory of the nonlinear least squares estimation", *Annals of Statistics*, Vol. 9, 501-513.

Appendix

In the appendix, we provide the proof of the consistency result stated in Theorem 3.1. The following two lemmas are required to prove Theorem 3.1.

Lemma 1. *Let* $S_{\delta,M} = \left\{ \boldsymbol{\theta}; \boldsymbol{\theta} = (A, B, \omega), |\boldsymbol{\theta} - \boldsymbol{\theta}^0| \geq 3\delta, |A| \leq M, |B| \leq M \right\}$. *If for any* $\delta > 0$ *and for some* $M < \infty$, $\liminf\limits_{n \to \infty} \inf\limits_{\boldsymbol{\theta} \in S_{\delta,M}} \frac{1}{n} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right] > 0$ *a.s., then* $\widehat{\boldsymbol{\theta}}$ *is a strongly consistent estimator of* $\boldsymbol{\theta}^0$.

Lemma 2. *Let* $\{x(t)\}$ *be stationary sequence which satisfy Assumption 1, then*

$$\lim_{n \to \infty} \sup_{\omega} \left| \frac{1}{n} \sum_{t=1}^{n} x(t) e^{i\omega t} \right| = 0 \quad a.s.$$

Lemma 3. *Let* $\{x(t)\}$ *be the same as in Lemma 2, then as* $n \to \infty$,

$$\sup_{\omega} \left| h(\omega) \frac{1}{n} \sum_{t=1}^{n} x(t) e^{i\omega t} \right| \to 0 \quad a.s.$$

where $h(\omega) = \alpha \sin(\omega)$ *or* $2\alpha \sin^2 \left(\frac{\omega}{2} \right)$ *with* $\alpha = A$ *or* B.

Proof of Theorem 3.1: We note that using (5) and notation $\boldsymbol{\theta} = (A, B, \omega)$, $Q(\boldsymbol{\theta})$ defined in equation (8), can be written as

$$Q(\boldsymbol{\theta}) = \sum_{t=1}^{n} [z^*(t) - c_1(\boldsymbol{\theta}) \sin(\omega t) - c_2(\boldsymbol{\theta}) \cos(\omega t)]^2,$$

where $z^*(t) = z(t) - \widehat{b}$, $c_1(\boldsymbol{\theta}) = -A\sin(\omega) - 2B\sin^2\left(\dfrac{\omega}{2}\right)$ and $c_2(\boldsymbol{\theta}) = B\sin(\omega) - 2A\sin^2\left(\dfrac{\omega}{2}\right)$ and $\widehat{b}$ is defined in section 2.

Now we write $\widehat{\boldsymbol{\theta}}$ as $\widehat{\boldsymbol{\theta}}_n$ to emphasize that $\widehat{\boldsymbol{\theta}}$ depends on n. If $\widehat{\boldsymbol{\theta}}_n$ is not a consistent estimator of $\boldsymbol{\theta}$, then there exists a subsequence $\{n_k\}$ of $\{n\}$ such that $\widehat{\boldsymbol{\theta}}_{n_k}$ does not converge to $\boldsymbol{\theta}$.

CASE I: Suppose $|\widehat{A}_{n_k}| + |\widehat{B}_{n_k}|$ is not bounded, then at least $|\widehat{A}_{n_k}|$ or $|\widehat{B}_{n_k}|$ tends to ∞. This implies that $\dfrac{1}{n_k}Q(\widehat{\boldsymbol{\theta}}_{n_k}) \to \infty$. Since $\dfrac{1}{n_k}Q(\boldsymbol{\theta}^0) < \infty$, $\dfrac{1}{n_k}[Q(\widehat{\boldsymbol{\theta}}_{n_k}) - Q(\boldsymbol{\theta}^0)] \to \infty$. But as $\widehat{\boldsymbol{\theta}}_{n_k}$ is the LSE of $\boldsymbol{\theta}^0$, $Q(\widehat{\boldsymbol{\theta}}_{n_k}) - Q(\boldsymbol{\theta}^0) < 0$. This leads to a contradiction.

CASE II: Suppose $|\widehat{A}_{n_k}| + |\widehat{B}_{n_k}|$ is bounded. Then if $\widehat{\boldsymbol{\theta}}_{n_k}$ is not a consistent estimator of $\boldsymbol{\theta}^0$, then for some $\delta > 0$ and an $0 < M < \infty$, there exists a set $S_{\delta,M}$ (as defined in Lemma 1) such that $\widehat{\boldsymbol{\theta}}_{n_k} \in S_{\delta,M}$. Now we write

$$
\begin{aligned}
\frac{1}{n}[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] &= \frac{1}{n}\sum_{t=1}^{n} [z^*(t) - c_1(\boldsymbol{\theta})\sin(\omega t) - c_2(\boldsymbol{\theta})\cos(\omega t)]^2 \\
&= f(\boldsymbol{\theta}) + g(\boldsymbol{\theta}) \quad \text{(say)},
\end{aligned}
$$

where

$$
f(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}\left[c_1(\boldsymbol{\theta}^0)\sin(\omega^0 t) - c_1(\boldsymbol{\theta})\sin(\omega t) + c_2(\boldsymbol{\theta}^0)\cos(\omega^0 t) - c_2(\boldsymbol{\theta})\cos(\omega t)\right]^2,
$$

$$
g(\boldsymbol{\theta}) = \frac{2}{n}\sum_{t=1}^{n} x_d(t)\left[c_1(\boldsymbol{\theta}^0)\sin(\omega^0 t) - c_1(\boldsymbol{\theta})\sin(\omega t) + c_2(\boldsymbol{\theta}^0)\cos(\omega^0 t) - c_2(\boldsymbol{\theta})\cos(\omega t)\right].
$$

Note that $x_d(t) = x(t+1) - x(t) = \displaystyle\sum_{k=-\infty}^{\infty} \beta(k)e(t-k)$ where $\{\beta(k)\} = \{\alpha(k+1) - \alpha(k)\}$ are absolutely summable. Using Lemma 3, it follows that

$$
\lim_{n \to \infty} \sup_{\boldsymbol{\theta} \in S_{\delta,M}} g(\boldsymbol{\theta}) = 0 \quad a.s. \tag{23}
$$

Now we define the following sets

$$S_{\delta,M,1} = \left\{ \boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \omega), |A - A^0| \geq \delta, |A| \leq M, |B| \leq M \right\},$$

$$S_{\delta,M,2} = \left\{ \boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \omega), |B - B^0| \geq \delta, |A| \leq M, |B| \leq M \right\},$$

$$S_{\delta,M,3} = \left\{ \boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, \omega), |\omega - \omega^0| \geq \delta, |A| \leq M, |B| \leq M \right\}.$$

Then $S_{\delta,M} \subset (S_{\delta,M,1} \cup S_{\delta,M,2} \cup S_{\delta,M,3}) = S$ (say). Hence,

$$\underline{\lim} \inf_{\boldsymbol{\theta} \in S_{\delta,M}} \frac{1}{n} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right] \geq \underline{\lim} \inf_{\boldsymbol{\theta} \in S} \frac{1}{n} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right]. \tag{24}$$

Now we have to show

$$\underline{\lim} \inf_{\boldsymbol{\theta} \in S_{\delta,M,j}} \frac{1}{n} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right] > 0 \quad a.s. \tag{25}$$

for $j = 1, \ldots 3$ and then (24) leads to $\underline{\lim} \inf_{\boldsymbol{\theta} \in S_{\delta,M}} \frac{1}{n} \left[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0) \right] > 0$ $a.s.$ Therefore, using (23), the lemma is proved provided we show (25) for $j = 1, \ldots, 3$. For $j = 1$,

$$\varliminf \inf_{\boldsymbol{\theta} \in S_{\delta,M,1}} \frac{1}{n}[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] = \varliminf \inf_{\boldsymbol{\theta} \in S_{\delta,M,1}} f(\boldsymbol{\theta})$$

$$= \varliminf \inf_{|A-A^0|>\delta} \frac{1}{n} \sum_{t=1}^{n} \left[c_1(\boldsymbol{\theta}^0) \sin(\omega^0 t) - c_1(\boldsymbol{\theta}) \sin(\omega t) + c_2(\boldsymbol{\theta}^0) \cos(\omega^0 t) - c_2(\boldsymbol{\theta}) \cos(\omega t) \right]^2$$

$$= \varliminf \inf_{|A-A^0|>\delta} \frac{1}{n} \sum_{t=1}^{n} \left[- \left\{ A^0 \sin(\omega^0) + 2B^0 \sin^2(\frac{\omega^0}{2}) \right\} \sin(\omega^0 t) \right.$$

$$- \left\{ A \sin(\omega) + 2B \sin^2\left(\frac{\omega}{2}\right) \right\} \sin(\omega t) + \left\{ B^0 \sin(\omega^0) - 2A^0 \sin^2\left(\frac{\omega^0}{2}\right) \right\} \cos(\omega^0 t)$$

$$\left. - \left\{ B^0 \sin(\omega) - 2A \sin^2\left(\frac{\omega^0}{2}\right) \right\} \cos(\omega t) \right]^2$$

$$= \lim_{n\to\infty} \inf_{|A-A^0|>\delta} \frac{1}{n} \sum_{t=1}^{n} \left[(A - A^0) \left\{ \sin(\omega^0) \sin(\omega^0 t) + 2 \sin^2\left(\frac{\omega^0}{2}\right) \cos(\omega^0 t) \right\} \right]^2$$

$$= \lim_{n\to\infty} \inf_{|A-A^0|>\delta} (A - A^0)^2 \frac{1}{n} \sum_{t=1}^{n} \left\{ \sin(\omega^0) \sin(\omega^0 t) + 2 \sin^2\left(\frac{\omega^0}{2}\right) \cos(\omega^0 t) \right\}^2$$

$$\geq \delta^2 \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \left[\sin^2(\omega^0) \sin^2(\omega^0 t) + 4 \sin^4\left(\frac{\omega^0}{2}\right) cos^2(\omega^0 t) \right]$$

$$= \frac{\delta^2}{2} \left[\sin^2(\omega^0) + 4 \sin^4\left(\frac{\omega^0}{2}\right) \right] = \delta^2(1 - \cos(\omega^0)) > 0 \quad a.s. \quad \forall \quad \omega^0 \in (0, \pi).$$

For other j, proceeding along the same line, we show (25) and that proves the theorem.

[1]THEORETICAL STATISTICS AND MATHEMATICS UNIT, INDIAN STATISTICAL INSTITUTE, 7, S.J.S. SANSANWAL MARG, NEW DELHI - 110016, INDIA, *E-mail*: NANDI@ISID.AC.IN

[2]DEPARTMENT OF MATHEMATICS AND STATISTICS, INDIAN INSTITUTE OF TECHNOLOGY KANPUR, PIN 208016, INDIA, *E-mail*: KUNDU@IITK.AC.IN