Genetic algorithms based robust frequency estimation of sinusoidal signals with stationary errors

Amit Mitra *, Debasis Kundu

Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, Kanpur-208016, India

ABSTRACT

In this paper, we consider the fundamental problem of frequency estimation of multiple sinusoidal signals with stationary errors. We propose genetic algorithm and outlier-insensitive criterion function based technique for the frequency estimation problem. In the simulation studies and real life data analysis, it is observed that the proposed genetic algorithm based robust frequency estimators are able to resolve frequencies of the sinusoidal model with high degree of accuracy. Among the proposed methods, the genetic algorithm based least squares estimator, in the no-outlier scenario, provides efficient estimates, in the sense that their mean square errors attain the corresponding Cramér-Rao lower bounds. In the presence of outliers, the proposed robust methods perform quite well and seem to have a fairly high breakdown point with respect to level of outlier contamination. The proposed methods significantly do not depend on the initial guess values required for other iterative frequency estimation methods.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Estimating the parameters of multiple sinusoidal signals model embedded in additive noise is a fundamental problem in signal processing and in time series analysis. In several applications in signal processing (Kay, 1988; Stoica, 1993; Quinn and Hannan, 2001; Stoica and Moses, 2005) and time series analysis (Brillinger, 1987), the signals dealt with can be described by the following multiple sinusoidal model:

$$y(t) = \sum_{k=1}^{M} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) + \varepsilon(t). \tag{1}$$

Here $y(t)$s are observed at the equidistant time points, namely, $t=1,2,\dots,N$. The unknown parameters of the model are the frequencies $(\omega_1,\dots,\omega_M)$ and the corresponding amplitudes $(A_1,\dots,A_M)$ and $(B_1,\dots,B_M)$. A_ks and B_ks are arbitrary real numbers and ω_ks are distinct real numbers lying between $(0,\pi)$. The real valued additive white noise sequence $\{\varepsilon(t)\}$ is assumed to be stationary with finite variance σ^2 and it has the following form:

$$\varepsilon(t) = \sum_{i=1}^{P} \rho_i \varepsilon(t-i) + \delta(t), \tag{2}$$

$\{\delta(t)\}$ being a sequence of independent and identically distributed (i.i.d.) normal random variable with mean 0 and variance σ^2.

Furthermore, $\{\rho_i\}_{i=1}^{P}$ are such that the sequence $\{\varepsilon(t)\}$ is stationary. The particular case of $\rho_i=0$ for all i, corresponds to the i.i.d. noise case. M, the number of sinusoidal components is assumed to be known. Given a sample of size N, $\{y(1),\dots,y(N)\}$, the problem is to estimate the unknown frequencies and the corresponding amplitudes.

The sinusoidal model (1), is used to describe and model many real life applications where periodic phenomena is present. The extraction of frequencies of the sinusoidal signals model from time series data is a classical problem of ongoing interest in the literature of statistical signal processing (Mackisack et al., 1994; Kundu and Mitra, 1996; Kundu, 1997; Mitra and Kundu, 1997; Smyth and Hawkins, 2000; Nandi et al., 2002; Chan and So, 2004; Trapero et al., 2007; Bonaventura et al., 2007; Coluccio et al., 2008) and indeed has created interests among scientists, from various diverse fields. There exists a vast amount of literature addressing the computational aspect of the frequencies of the sinusoidal model as well as focusing on theoretical behavior of the estimators. The most intuitive and natural approach is the least squares approach. A closely related approach is the approximate least squares estimators (ALSEs) approach, which is asymptotically equivalent to the least squares estimators (LSEs). Asymptotic properties of the ALSE and LSE are studied in detail in Walker (1971), Hannan (1971), Kundu (1993, 1997) and Kundu and Mitra (1996). It is well known that, although the LSEs are the most desired estimators from theoretical point of point of view, obtaining the LSEs is numerically a very difficult problem (Kahn et al., 1993). It is observed that the least squares surface has local

* Corresponding author. Tel.: +91 512 2596064; fax: +91 512 2597500.
  E-mail address: amitra@iitk.ac.in (A. Mitra).

minima spaced $O(N^{-1})$ apart, making the gradient based search methods of general non-linear optimization ineffective without excellent starting values. Several methods are available in the literature to obtain the LSEs efficiently, but unfortunately all the methods are quite sensitive to the initial value chosen. It is further observed in Rice and Rosenblatt (1988), that unless the frequencies are resolved at the first step with order $O(N^{-1})$, the failure to converge to global minima may give a very poor estimate of the amplitudes. Thus, fitting these multiple sinusoidal models can involve daunting computational difficulties. The problem can be further complicated in case outliers are present in the dataset.

In this paper, we develop genetic algorithm based frequency estimation methods optimizing outlier-insensitive criterion functions. The aim here is to find an algorithm, under the assumption of stationary additive noise random variable, whose performance significantly does not depend on initial guess values (or intervals), and also have a high breakdown point with respect to outliers present in the data. Recently, Smyth and Hawkins (2000) proposed an algorithm based on elemental sets for robust frequency estimation, under the assumption of independently and identically distributed (i.i.d.) normal random variables. Contrary to the remarks made in Smyth and Hawkins (2000) that the genetic algorithms does not seem to be a suitable approach for this problem (under i.i.d. setup), especially with the presence of outlier, we observe that the proposed genetic algorithm based methods perform quite satisfactorily even in the dependent noise structure.

The rest of the paper is organized as follows. In Section 2, we give the least squares and the L1-norm formulation of the frequency estimation problem. In Section 3, we will give a brief review of the outlier-insensitive criterion functions. Section 4 presents the proposed genetic search based iterative algorithms for robust frequency estimation. The empirical studies, implementing the proposed algorithms, will be presented in Sections 5. Finally, the conclusions will be discussed in Section 6.

2. Least squares and L1-norm estimators

The least squares estimators of the parameters for the model (1) are the minimizers of the criterion function:

$$\psi(\underset{\sim}{\omega}, \underset{\sim}{\mathbf{A}}, \underset{\sim}{\mathbf{B}}) = \sum_{t=1}^{N} \left[y(t) - \sum_{k=1}^{M} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) \right]^2, \quad (3)$$

where $\underset{\sim}{\omega} = (\omega_1, \ldots, \omega_M)^T$ is the vector of frequencies and $\underset{\sim}{\mathbf{A}} = (A_1, \ldots, A_M)^T$ and $\underset{\sim}{\mathbf{B}} = (B_1, \ldots, B_M)^T$ are the amplitude vectors. Here 'T', denotes transpose of a vector or of a matrix. The sinusoidal model parameters estimated through minimization of (3) has the smallest least squares distance to the observed data. $\underset{\sim}{\omega}$, $\mathbf{A}$ and $\mathbf{B}$ obtained by minimizing (3) are called the non-linear least squares (NLS) estimators. When the noise $\varepsilon(t)$ is white Gaussian, the NLS estimators are same as the maximum likelihood estimators.

For the sinusoidal model (1), the criterion function (3) can conveniently be concentrated with respect to the conditionally linear parameters $\mathbf{A}$ and $\mathbf{B}$. Introducing the notations:

$$\underset{\sim}{Y} = [y(1), y(2), \ldots, y(N)]^T, \quad (4)$$

$$A(\underset{\sim}{\omega}) = \begin{bmatrix} \cos(\omega_1) & \sin(\omega_1) & \cdots & \cos(\omega_M) & \sin(\omega_M) \\ \vdots & \vdots & & \vdots & \\ \cos(\omega_1 N) & \sin(\omega_1 N) & \cdots & \cos(\omega_M N) & \sin(\omega_M N) \end{bmatrix}, \quad (5)$$

$$\underset{\sim}{\alpha} = [A_1, B_1, \ldots, A_M, B_M]^T, \quad (6)$$

we can write $\psi(\underset{\sim}{\omega}, \underset{\sim}{\mathbf{A}}, \underset{\sim}{\mathbf{B}})$ as

$$\psi(\underset{\sim}{\omega}, \underset{\sim}{\mathbf{A}}, \underset{\sim}{\mathbf{B}}) = (\underset{\sim}{Y} - A(\underset{\sim}{\omega}) \underset{\sim}{\alpha})^T (\underset{\sim}{Y} - A(\underset{\sim}{\omega}) \underset{\sim}{\alpha}). \quad (7)$$

With distinct frequencies, if $N \geq 2M$ the Vandermonde matrix $A(\underset{\sim}{\omega})$ is of rank $2M$ and $(A(\underset{\sim}{\omega})^T A(\underset{\sim}{\omega}))^{-1}$ exists. We thus observe that the vectors $\underset{\sim}{\omega}$ and $\underset{\sim}{\alpha}$ which minimize (3) are given by

$$\hat{\underset{\sim}{\omega}}_{LSE} = \arg\max_{\underset{\sim}{\omega}} [\underset{\sim}{Y}^T A(\underset{\sim}{\omega}) (A(\underset{\sim}{\omega})^T A(\underset{\sim}{\omega}))^{-1} A(\underset{\sim}{\omega})^T \underset{\sim}{Y}], \quad (8)$$

$$\hat{\underset{\sim}{\alpha}}_{LSE} = (A(\underset{\sim}{\omega})^T A(\underset{\sim}{\omega}))^{-1} A(\underset{\sim}{\omega})^T \underset{\sim}{Y} |_{\underset{\sim}{\omega} = \hat{\underset{\sim}{\omega}}_{LSE}}. \quad (9)$$

It is well known that the least square estimators for this problem are optimal under various considerations on the noise sequence. It is observed that, under the i.i.d. assumption on the noise sequence, the estimators are strongly consistent (Kundu and Mitra, 1996), asymptotically normal with a covariance matrix that coincides with the Cramér-Rao bound under the normality assumption. It is further observed that the LSEs under the dependent error structure are also strongly consistent and asymptotic normal (Kundu, 1993).

As an alternate to the above LSE formulation of the problem, we can use L1-norm formulation, which is often used in the literature of robust regression. The L1-norm estimates of the parameters of the multiple sinusoidal model (1) are obtained by minimizing

$$\phi(\omega, \mathbf{A}, \mathbf{B}) = \sum_{t=1}^{N} \left| y(t) - \sum_{k=1}^{M} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) \right|. \quad (10)$$

The non-linear optimization problem (10) is solved using standard non-linear optimization routines in order to get the L1-norm estimates. The L1-norm estimators, also called the least absolute deviation (LAD) estimators, correspond to the maximum likelihood estimators under the assumption that noise are i.i.d. with double exponential distribution. In recent literatures of signal processing, use of modified simplex algorithm is proposed for obtaining the L1-norm estimates. The estimates are then computed using the Barrodale–Roberts modified simplex algorithm (Barrodale and Roberts, 1973, 1974). For a more detailed review of L1-norm techniques, readers are referred to Bloomfield and Steiger (1983).

3. Outlier-insensitive criterion functions

The conventional estimates that are found by the least squares criterion, i.e. minimizing the sum of squares of all the N residuals, are motivated by the ideas of statistical efficiency. However, the estimates are inappropriate if some of the observations are contaminated. The L1-norm estimates are potentially better options in situations where the dataset contains outliers. Deviating from the use of usual sum of square errors or sum of absolute errors criterion functions, literature of robust regression provides us with alternate criterion functions that are relatively insensitive to the presence of outliers in the data. The primary aim of these outlier-insensitive criteria is to protect the estimate from such outlier contamination.

Among the most widely used specialized outlier-sensitive criterion functions, are the least trimmed (LT) sum and the least median (LM) criteria. We now formulate the criteria to be used in the robust frequency estimation methods proposed in this paper.

3.1. Least trimmed criterion

Let $e_{(1)}^2 < e_{(2)}^2 <, \ldots, < e_{(N)}^2$ be N ordered estimated squared residuals, for an estimated value of A, B, ω. The unordered $e(i)^2$s

for the model (1) are given by

$$e(t)^2_{\underset{\sim}{\mathbf{A}}, \underset{\sim}{\mathbf{B}}, \underset{\sim}{\omega}} = \left(y(t) - \sum_{k=1}^{M} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) \right)^2 ; t = 1, 2, \ldots, N. \tag{11}$$

The least trimmed (sum of) squares (LTS) estimator, proposed in Rousseeuw (1984), is found by finding the parameters that satisfy

$$\underset{\underset{\sim}{\omega}, \underset{\sim}{\mathbf{A}}, \underset{\sim}{\mathbf{B}}}{Min} \sum_{t=1}^{h} e^2_{(t)}. \tag{12}$$

It is well known (Rousseeuw, 1984) that the best robustness property is obtained when $h=N/2$, approximately. In this case a breakdown point of 50% is attained. Higher efficiency of the estimates is obtained with lower trimming proportions. We consider in the present paper, a 50% trimming.

Alternatively, under L1-norm estimation setup, we can similarly define a least trimmed (sum of) absolute (LTA) deviation estimator. The LTA deviation estimator is found by finding the parameters that satisfy

$$\underset{\omega, \mathbf{A}, \mathbf{B}}{Min} \sum_{t=1}^{h} |e_{(t)}|, \tag{13}$$

where $|e_{(1)}| < |e_{(2)}| <, \ldots, < |e_{(N)}|$ be N ordered absolute residuals, the unordered $|e(t)|$s for model (1) are given by

$$|e(t)| = \left| y(t) - \sum_{k=1}^{M} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) \right| ; t = 1, 2, \ldots, N. \tag{14}$$

Once again we consider a 50% trimming for the LTA based estimators in the present paper.

3.2. Least median criterion

The least median squares (LMS) estimator is obtained by finding the model parameters that minimizes the hth-ordered squared residual, i.e. $e^2_{(h)}$, where h is usually taken as $h=[N/2]+[(p+1)/2]$, p denotes the number of parameters in the model. This estimator was introduced in Rousseeuw (1984) (see also Rousseeuw, 1988; Rousseeuw and Leroy, 1987).

Similar to the LMS estimator, we can define the least median absolute (LMA) estimator as the estimator that is obtained by finding the parameters that minimize the hth-ordered absolute residual, i.e. $|e_{(h)}|$, with appropriate choice of h.

4. Proposed robust frequency estimation methods

In this section, we present the proposed genetic algorithm based frequency estimation techniques for the multiple sinusoidal model.

We propose six different estimators based on genetic algorithm and different criterion functions. These estimators are: (i) genetic algorithm based least square estimator (GA-LS), (ii) genetic algorithm based least trimmed square estimator (GA-LTS), (iii) genetic algorithm based least median square estimator (GA-LMS), (iv) genetic algorithm based L1-norm estimator (GA-L1), (v) genetic algorithm based least trimmed absolute deviation estimator (GA-LTA), (vi) genetic algorithm based least median absolute deviation estimator (GA-LMA).

We first present the algorithm for the GA-LS estimator. In the genetic search formulation of the GA-LS estimator of the parameters of the model (1), we take the objective function (8) as the fitness function in the genetic search setup and aim to find the optimum member through repeated applications of the three genetic operators of selection, crossover and mutation, over the successive generations. The parameter space, Ω^{freq} for the frequency vector, $\underset{\sim}{\omega} = (\omega_1, \ldots, \omega_M)^T$, for the sinusoidal model (1) is given by

$$\Omega^{freq} = (0, 1) \times (0, 1) \times \cdots \times (0, 1) \subset \mathfrak{R}^M. \tag{15}$$

We first obtain the binary chromosomal representation of the parameter space. We form, for any possible solution belonging to the original parameter space Ω^{freq}, a binary string of length $M \times p$. Where, p denotes the length of the binary bit representation of any component of the parameter vector ω, i.e. for each of the unknown frequencies, we obtain a p-bit coded binary representation. It is however well known that ordinary binary coding can result in search process being deceived, i.e. unable to efficiently locate the global minima, due to large hamming distances in the representational mapping between adjacent values (Hollstien, 1971). A hamming distance, between two binary strings is defined as minimum number of bits that must be changed in order to convert one bit string into another. In order to avoid the above-mentioned problem, a Gray coding approach of the original binary strings is adopted. The literature of GA and its applications report that Gray coding exhibits accelerated convergence rate of the objective function, and provides better accuracy than the binary coded GA (Caruana and Schaffer, 1988; Yokose et al., 2000). Superior performance of a Gray coded GA is mainly attributed to the fact that Gray codes do not bias the searching direction, as in the case of ordinary binary coding, having a large hamming distance between adjacent values. A Gray code represents each number in the sequence of integers $\{0, 1, \ldots, 2^K - 1\}$ as a binary string of length K in an order such that adjacent integers have Gray code representations that differ in only one bit position. Use of Gray code thus allows, going through the integer sequence requiring flipping just one bit at a time. This is called the adjacency property of Gray codes. Gray code takes a binary sequence and shuffles it to form some new sequence with the adjacency property. We use here a Gray coding derived from the initial binary coding.

To initialize the genetic search, we populate an initial population of a pre-determined size. Each member of this initial population is a randomly chosen parameter vector $\omega^0 \in \Omega^{freq}$, coded to get the chromosomal string representation of bit length $M \times p$. The ranking based fitness of each of the members of this initial population is evaluated according to the criterion (8). For a detailed discussion on various selection procedures, see for example Goldberg (1989). Using a stochastic sampling with replacement approach, we next populate fit parents pool, size of the pool depending on the generation gap. From the selected parent pool, we select pairs in order and apply a two-point crossover (with a pre-assigned crossover probability), exchanging genetic material of parents to obtain new chromosomes. Crossover produces new individuals that have some parts of both the parent's genetic material. An example of a multipoint crossover is illustrated in Fig. 1.

Mutation is applied on the mated chromosome strings with a low pre-assigned mutation probability. Mutation is considered to be the genetic operator that ensures that the probability of searching any given string will never be zero and thus has the effect of tending to inhibit the possibility of convergence of the GA to a local optimum. Mutation changes the genetic representation of the chromosomes according to a probabilistic rule. In the binary string representation, mutation will cause a single bit to change its state, i.e. $0 \Rightarrow 1$ or $1 \Rightarrow 0$.

An elitist strategy is used to fill the generation gap. An elitist strategy (De Jong, 1975; Thierens, 1997) is adopted while populating a new generation. Elitism encourages the inclusion of highly fit chromosome strings, from earlier generations, in the subsequent generations. The fractional difference between the
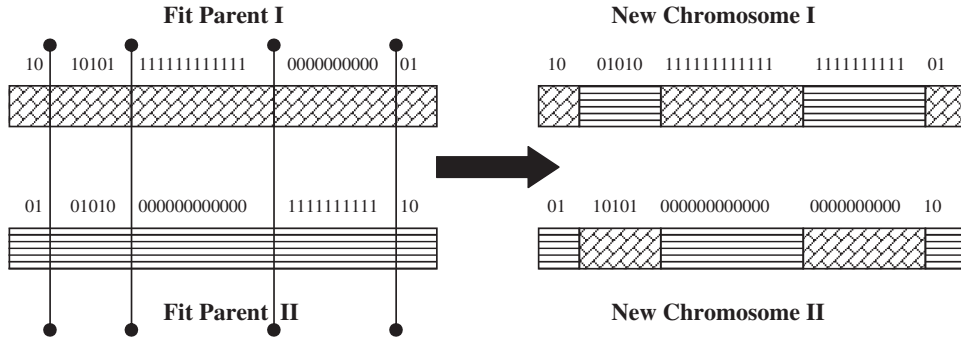
Fig. 1. Example of a 4-point crossover.

number of chromosomes in the old population and the number of new chromosomes produced by selection and recombination is termed as the generation gap. Under the elitist approach, a fraction (based on the value of the pre-determined generation gap) of the most-fit individuals is deterministically allowed to propagate through successive generations.

Since GA is a stochastic optimization algorithm, the application of conventional termination criteria becomes problematic in GA based optimization procedure. We follow here the most commonly adopted practice, where the cycles of selection, crossover and mutation is carried on until a pre-determined number of generations have been completed or no better solution is found after a pre-determined number of successive generations have evolved, whichever is earlier. We walk through the GA steps repeatedly, until the termination criterion is reached.

After completion of each generation, we preserve the information regarding the most fit, i.e. the parameter vector that is the best solution for the optimization of (3) ((8) for frequency estimation), in that generation. The GA based least square (GA-LS) solution of $\underset{\sim}{\omega}$, say $\hat{\underset{\sim}{\omega}}_{GA\text{-}LSE}$, is the most-fit individual evolving among all the generations, at the point when termination criterion is reached. Once we obtain $\hat{\underset{\sim}{\omega}}_{GA\text{-}LSE}$, the estimates of the conditionally linear parameters, the amplitudes, $\underset{\sim}{\alpha}$, may be obtained using (9).

The algorithmic steps for the proposed procedure are given below:

Step 1: Randomly initialize initial population generation (of a pre-determined size) of chromosomes of Gray coded binary strings of length $M \times p$, each of these chromosomes is the coded binary representation of a possible solution for the least square frequency estimation problem.

Step 2: Decode the Gray coded binary strings using a linear scaling.

Step 3: Evaluate the objective function (8) for each of the decoded strings and obtain their fitness values using a ranking based approach. Preserve the information about the string with highest fitness value.

Step 4: Using a stochastic sampling with replacement approach, populate fit parents pool, size of the pool depending on the generation gap.

Step 5: From the selected parents pool, we select pairs in order and apply a two-point crossover (with a pre-assigned crossover probability), exchanging genetic material of parents to obtain new chromosomes.

Step 6: Apply mutation on the mated chromosome strings with small pre-assigned mutation probability.

Step 7: Use elitist strategy to fill the generation gap.

Step 8: Repeat the steps 2–7 till maximum number of generations is reached or no better solution is found after the pre-determined maximum number of generations is reached.

Step 9: $\hat{\underset{\sim}{\omega}}_{GA\text{-}LSE}$ is the most-fit decoded string found among all the generations.

Step 10: Calculate the estimates of the conditionally linear parameters $\underset{\sim}{\alpha}$ through (9).

For the GA-LTS estimator, we consider the $3M$ dimensional model parameter vector as $\underset{\sim}{\eta} = (A_1, B_1, \omega_1, \ldots, A_M, B_M, \omega_M)$ and consider the objective function as

$$\underset{\underset{\sim}{\omega},\underset{\sim}{\mathbf{A}},\underset{\sim}{\mathbf{B}}}{Min} \sum_{t=1}^{[N/2]} e_{(t)}^2. \tag{16}$$

where $e_{(1)}^2 < e_{(2)}^2 < \cdots < e_{([N/2])}^2$ are $[N/2]$ smallest squared residuals, the unordered squared residuals, $e(t)^2$s, are given by (11). The parameter space, Ω for the present setup for the model (1) is given by

$$\Omega = (-\infty, \infty) \times (-\infty, \infty) \times (0,1) \cdots (-\infty, \infty) \times (-\infty, \infty) \times (0,1) \subset \Re^{3M}. \tag{17}$$

Similar to the GA-LS method, we first obtain the binary chromosomal representation of the parameter space Ω. We form, for any possible solution belonging to the original parameter space Ω, a binary string of length $3Mp$. p denotes the length of the binary bit representation of any component of the parameter vector $\underset{\sim}{\eta}$.

The algorithmic steps for obtaining the GA-LTS estimator $\underset{\sim}{\tilde{}}$ is similar to the steps followed to obtain the GA-LS estimates with the difference that in Step 1 we initialize the initial population now with binary strings of length $3Mp$ and in Step 9 we obtain the GA-LTS estimates of the entire parameter vector.

For the GA-LMS estimator, the objective function in the GA-LTS setup is replaced by

$$\underset{\underset{\sim}{\omega},\underset{\sim}{\mathbf{A}},\underset{\sim}{\mathbf{B}}}{Min} \quad e_{([N/2]+[(p+1)/2])}^2. \tag{18}$$

The parameter space and the algorithmic steps remain same as that of GA-LTS estimator. For the L1-norm based estimators, namely the GA-L1 estimator, the GA-LTA estimator and the GA-LMA estimator, the parmeter space remains (17). The objective function for the GA-L1 estimator is given by

$$\underset{\underset{\sim}{\mathbf{A}},\underset{\sim}{\mathbf{B}},\underset{\sim}{\omega}}{Min} \sum_{t=1}^{N} \left| y(t) - \sum_{k=1}^{M} (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)) \right|. \tag{19}$$

The algorithmic steps remain the same as the steps for obtaining GA-LTS estimator. The objective function for the

GA-LTA estimator is given by

$$\underset{\underset{\sim}{\omega}, \mathbf{A}, \mathbf{B}}{Min} \sum_{t=1}^{[N/2]} |e_{(t)}|. \tag{20}$$

where $|e_{(1)}| < |e_{(2)}| < \cdots < |e_{([N/2])}|$ are $[N/2]$ smallest absolute deviations, the unordered $|e(t)|$s for model (1) are given by (14). Finally, for the GA-LMA estimator, the objective function for the GA procedure is given by

$$\underset{\underset{\sim}{\omega}, \mathbf{A}, \mathbf{B}}{Min} |e_{([N/2]+[(p+1)/2])}|. \tag{21}$$

5. Simulation studies and real life data analysis

In this section, we will apply the proposed procedures of GA based frequency estimation techniques for frequency estimation of various simulated sinusoidal models. We will also perform extensive simulation studies to investigate the possible effect of outliers present in the data. In the simulation studies, we consider both dependent error as well as independent error structures. We report here the performance of the following estimators: (i) genetic algorithm based least square estimator (GA-LS), (ii) genetic algorithm based least trimmed square estimator (GA-LTS), (iii) genetic algorithm based least median square estimator (GA-LMS), (iv) genetic algorithm based L1-norm estimator (GA-L1), (v) genetic algorithm based least trimmed absolute deviation estimator (GA-LTA), (vi) genetic algorithm based least median absolute deviation estimator (GA-LMA). Real life data analysis using the proposed methods will also be presented.

5.1. Simulation results for independent error structure

In this subsection, we present the empirical studies for 1-component and 2-component simulated sinusoidal models with independent error structure. For the purpose of comparing the performance of the proposed robust methods with the elemental set based robust frequency estimates of Smyth and Hawkins (2000), we consider the same models as reported therein. We report the average estimates, the root mean square errors (RMSE) and the standard deviations (St. Dev.) over 100 simulation runs. The random numbers are generated using MATLAB random number generator.

5.1.1. One sinusoid

We consider the following one-component sinusoidal model

$$y(t) = \cos(\omega t + \phi) + \varepsilon(t); \quad t = 1, 2, \ldots, N. \tag{22}$$

The true value of the frequency of the simulation model is $\omega = 0.5$ and that of ϕ is 0.1. $\varepsilon(t)$ is taken as i.i.d. normal noise sequence, with mean zero and standard deviation $\sigma = 0.2$. The sample size is taken as 100. The Cramér-Rao bound, which is same as the asymptotic variance of the LSE (Kundu and Mitra, 1996), for the frequency parameter is $9.6E-07$. For each of the simulated datasets, we estimated the frequency using the methods described in Section 4. The particular choice of the genetic parameters for the genetic formulation setup for the simulation model is given in Table 1.

The trimmed proportions for GA-LTS are taken as 50% and 80% and for the GA-LTA it is taken as 50%. The root mean square errors (RMSE), the average estimates and the standard deviation over 100 simulations for the frequency is computed for all the proposed methods. We also report, for comparison, the corresponding result of the best performing robust frequency estimate of Smyth and Hawkins (2000). We also investigate the performance of the proposed estimators when data contains 30%

Table 1
Choice of genetic parameters for the simulations for model (22).

Genetic parameter	Values
Number of chromosomes in one population	200
Bits of precision	80
Coding	Gray coding
Scaling	Linear
Range of parameters for initial population	$\omega \in [0, \pi]$
Crossover probability	0.70
Crossover method	2-point
Mutation probability	0.01
Elitism	Top 10%
Maximum number of generation	200

Table 2
Simulation results for one cosine signal model (22).

Method	No outlier			30% outlier		
	Average	St. Dev.	RMSE	Average	St. Dev.	RMSE
GA-LS	0.5001	0.00099	0.00099	1.4750	0.79855	1.26025
GA-LTS	0.5001	0.00104	0.00105	0.5004	0.00116	0.00126
GA-LMS	0.5002	0.00127	0.00128	0.5008	0.00153	0.00172
GA-L1	0.5002	0.00101	0.00102	0.5004	0.00153	0.00158
GA-LTA (50)	0.5004	0.00121	0.00124	0.5004	0.00134	0.00141
GA-LTA (80)	0.5002	0.00103	0.00104	0.5002	0.00119	0.00120
GA-LMA	0.5004	0.00105	0.00109	0.5004	0.00112	0.00118

outliers. Outliers were generated to have standard deviations 100 times that of the good observations. The outliers were associated to a randomly selected subset of 30 observations. The results for the no outlier and the outlier scenarios are presented in Table 2.

From the results for the non-outlier case, we observe that the proposed estimators perform quite well, even for the trimmed cases. Among the proposed estimators, GA-LS and GA-L1 performs the best. The performance of the GA-LS is almost fully efficient (98%) and better than the best performing estimators ELS-LS, LTS-LI1-MM and LMS-LI1-MM (efficiency 94%) reported in Smyth and Hawkins (2000). We further observe that the best performing method in the non-outlier case, the GA-LS method fails completely in the presence of outliers. The performance of the GA-L1 method is still quite promising. However, much better results are obtained with genetic algorithm based trimmed and least median criterion functions. The best results are obtained for the GA-LMA method. The performance of the GA-LTS with 50% trimming and the GA-LTA (80) method is also quite encouraging. For the outlier scenario, the performances of the proposed GA-LMA and GA-LTA (80) are better than the best performing method LTS-L1-MM (with St. Dev. and RMSE 0.00127) of Smyth and Hawkins (2000).

5.1.2. Two sinusoids

We consider the following two-component sinusoidal model

$$y(t) = \cos(\omega_1 t + \phi_1) + \cos(\omega_2 t + \phi_2) + \varepsilon(t); \quad t = 1, 2, \ldots N, \tag{23}$$

where we take the true values of the frequencies of the simulation model as $\omega_1 = 0.3$ and $\omega_2 = 0.7$ and that of ϕ_1 as 0.2 and ϕ_2 as 0.1. $\varepsilon(t)$ is taken as i.i.d. normal noise sequence, with mean zero and standard deviation $\sigma = 0.2$. The sample size is taken as 100. The Cramér-Rao bounds for the frequency parameters are same and equal to $9.6E-07$. For each of the simulated dataset, we estimated the frequencies using the methods described in Section 3. The choice of the genetic parameters for the two-component sinusoidal model is similar to the ones mentioned for the one-component model (Table 1). However, to accommodate for higher-dimensional

Table 3
Simulation results for the two cosine signals model (23).

Frequency	Method	No outlier			30% outlier		
		Average	St. Dev.	RMSE	Average	St. Dev.	RMSE
$\omega = 0.7$	GA-LS	0.6999	0.00075	0.00075	1.9237	0.77294	1.44734
	GA-LTS	0.6998	0.00123	0.00124	0.7000	0.00137	0.00137
	GA-LMS	0.6999	0.00108	0.00108	0.7000	0.00121	0.00121
	GA-L1	0.7001	0.00092	0.00092	0.6999	0.00151	0.00151
	GA-LTA	0.6999	0.00127	0.00127	0.6998	0.00138	0.00139
	GA-LMA	0.7001	0.00116	0.00116	0.6999	0.00128	0.00128
$\omega = 0.3$	GA-LS	0.3002	0.00067	0.00069	0.9250	0.70135	0.93941
	GA-LTS	0.3004	0.00103	0.00106	0.3005	0.00113	0.00125
	GA-LMS	0.3006	0.00102	0.00107	0.3009	0.00114	0.00147
	GA-L1	0.3002	0.00081	0.00097	0.3008	0.00124	0.00148
	GA-LTA	0.3004	0.00123	0.00127	0.3008	0.00146	0.00169
	GA-LMA	0.3007	0.00110	0.00117	0.3009	0.00124	0.00153

parameter space, we form a larger chromosome pool (350) for each population. The root mean square errors (RMSE), the average estimates and the standard deviations over 100 simulations for all the frequencies are computed for all the proposed methods. Once again we also report, for comparison, the corresponding results of the best performing robust frequency estimate of Smyth and Hawkins (2000). Similar to the one-component model, we also investigate the performance of the proposed estimators when data contains 30% outliers. Once again outliers were generated to have standard deviations 100 times that of the good observations. The outliers were associated to a randomly selected subset of 30 observations. The results for the outlier as well as the non-outlier cases are presented in Table 3.

From the results of the two-component model, we observe that for the no-outlier scenario, GA-LS and GA-L1 methods perform the best. These estimators give super efficient estimates in the sense that their MSEs are lower than the corresponding Cramér-Rao bounds. The performances of GA-LS (efficiency of 131% for the higher frequency and 124% for the lower frequency) and GA-L1 (efficiency 113% for the higher frequency and 102% for the lower frequency) are much better than Smyth and Hawkins (2000) elemental set based methods (the reported maximum efficiency of 109% is reported for the higher frequency and 104% for the lower frequency). The performances of the genetic algorithm based trimmed criterion function estimators are also reasonably good. The results for the two-component sinusoidal model with 30% outliers are qualitatively same as the results of the one sinusoid. The results once again indicate satisfactory performance of the proposed robust estimators.

5.2. Simulation results for dependent error structure

In this subsection, we present the simulation studies for sinusoidal models with dependent error structure.

5.2.1. One sinusoid
We consider the following one-component sinusoidal model

$$y(t) = 5.0\cos(0.4t) + 4.0\sin(0.4t) + \varepsilon(t); \quad t = 1, 2, \ldots, N. \tag{24}$$

The error structure of $\varepsilon(t)$ is taken as

$$\varepsilon(t) = 0.3\varepsilon(t-1) + \delta(t), \tag{25}$$

where $\delta(t)$s are i.i.d. normal noise sequence, with mean zero and standard deviation σ. We consider two different values of σ, 0.01 and 0.05. The sample size is taken as 100. For each of the simulated datasets, we estimated the frequency using the
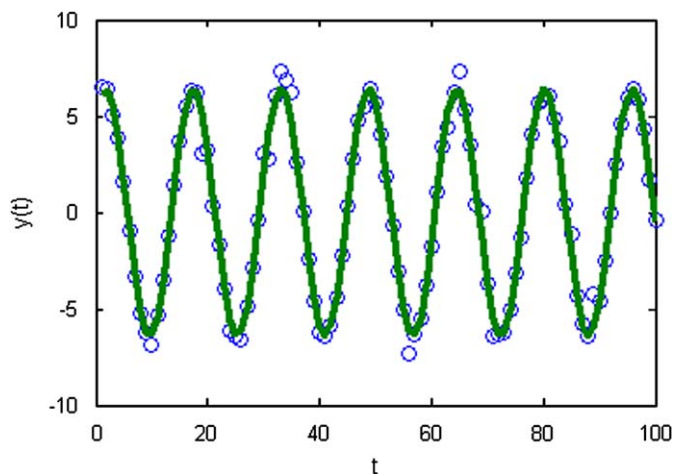

Fig. 2. A representative dataset of one-component dependent error ($\sigma = 0.01$) sinusoidal model containing 30% outlier and the corresponding GA-L1 fit for model (24).

methods described in Section 4. The choices of the genetic parameters for the genetic formulation setup for the simulation model are as in Table 1. The trimming proportions for GA-LTS and GA-LTA are taken as 80%. The root mean square errors (RMSE), the average frequency estimates and the associated standard deviations over 100 simulations are computed for all the proposed methods. The theoretical asymptotic standard deviation of the least squares estimator for $\sigma = 0.01$ is $1.044E-5$ and that for $\sigma = 0.05$ is $5.218E-5$. We next investigate the performance of the proposed estimators when data contains 30% outliers, under the correlated error structure. Outliers were generated to have standard deviations 100 times that of the good observations and associated to a randomly selected subset of 30 observations. Two representative plots of 30% outlier dataset in the dependent error setup and the corresponding GA-L1 fit of the data are given in Figs. 2 and 3. The results for the non-outlier as well as the outlier scenarios are presented in Table 4.

We observe that the proposed methods are able to resolve the unknown frequency with high level of accuracy for dependent error structure as well. For the non-outlier scenario, GA-LS performs the best closely followed by GA-L1. Even the performances of the least trimmed and least median based approaches provide fairly accurate estimates. From the simulations of the outlier study, we observe that the proposed robust frequency estimation methods perform quite well. While the performance of

the GA-LS deteriorates significantly as compared to non-outlier case, the performance of GA-L1 and the least trimmed and least median approaches remain fairly stable even with 30% outlier contamination in the data. We had observed similar pattern for independent errors also. The GA-L1 estimator performs the best in this situation. A further investigation reveals that for $\sigma \geq 0.1$, the GA-LS totally breaks down, the robust frequency estimators still continue to give reasonably good results.
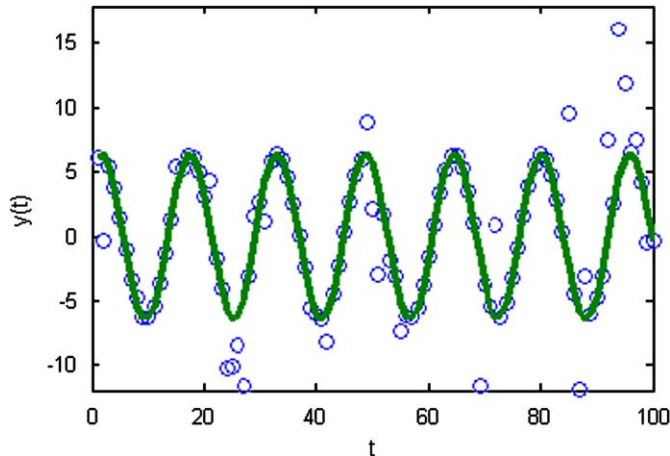

Fig. 3. A representative dataset of one-component dependent error ($\sigma=0.05$) sinusoidal model containing 30% outlier and the corresponding GA-L1 fit for model (24).

5.2.2. Two sinusoids

We consider the following two-component sinusoidal model

$$y(t) = 1.0 \cos(0.3t) + 1.5 \sin(0.3t) + 2.5 \cos(0.8t)$$
$$+ 2.0 \sin(0.8t) + \varepsilon(t); \quad t = 1, 2, \ldots, n. \qquad (26)$$

The error structure of $\varepsilon(t)$ is taken as $\varepsilon(t) = 0.3\varepsilon(t-1) + \delta(t)$, where $\delta(t)$s are i.i.d. normal random variables, with mean zero and standard deviation σ. The sample size is taken as 75. We have considered two different values of σ, 0.01 and 0.1. Performances of the proposed estimators with 30% outlier contamination are also investigated. As in the previous cases, outliers were generated to have standard deviations 100 times that of the good observations and are associated to a randomly selected subset. For each of the simulated datasets, we estimated the frequencies using the methods described in Section 4. The choice of the parameters for the genetic formulation setup for the two-component dependent error simulation model is similar to that of the two-component independent error model. The results for the lower of the two sinusoids are presented in Table 5 and the results for the higher of the two sinusoids are presented in Table 6. A representative plot of 30% outlier dataset in the two-component dependent error setup is given in Fig. 4 and the data along with the fit corresponding to GA-L1 solution is given in Fig. 5.

For the two-component model non-outlier cases, we observe that the GA-LS performs the best closely followed by GA-L1. The theoretical asymptotic standard deviation of the LSE of $\omega=0.8$ at $\sigma=0.01$ is $9.738E-6$ and that for $\sigma=0.1$ is $9.738E-5$ and the asymptotic standard deviation of the LSE of $\omega=0.3$ at $\sigma=0.01$ is

Table 4
Simulation results for one sinusoid dependent error model (24).

σ	Method	No outlier			30% outlier		
		Average	St. Dev.	RMSE	Average	St. Dev.	RMSE
0.01	**GA-LS**	0.4000	$1.009E-5$	$1.011E-5$	0.4000	$3.218E-4$	$3.227E-4$
	GA-LTS	0.4000	$4.363E-5$	$4.364E-5$	0.4000	$4.828E-5$	$4.842E-5$
	GA-LMS	0.4000	$5.433E-5$	$5.455E-5$	0.4000	$6.705E-5$	$6.709E-5$
	GA-L1	0.4000	$1.336E-5$	$1.339E-5$	0.4000	$1.849E-5$	$1.875E-5$
	GA-LTA	0.4000	$6.937E-5$	$6.999E-5$	0.4000	$7.003E-5$	$7.011E-5$
	GA-LMA	0.4000	$7.444E-5$	$7.447E-5$	0.4000	$7.460E-5$	$7.476E-5$
0.05	**GA-LS**	0.4000	$5.174E-5$	$5.282E-5$	0.4000	$1.355E-3$	$1.356E-3$
	GA-LTS	0.4000	$1.179E-4$	$1.179E-4$	0.4000	$1.556E-4$	$1.557E-5$
	GA-LMS	0.4000	$1.330E-4$	$1.331E-4$	0.4000	$1.432E-4$	$1.464E-4$
	GA-L1	0.4000	$5.543E-5$	$5.735E-5$	0.4000	$9.303E-5$	$9.309E-5$
	GA-LTA	0.4000	$1.086E-4$	$1.094E-4$	0.4000	$1.105E-4$	$1.112E-4$
	GA-LMA	0.4000	$1.219E-4$	$1.229E-5$	0.4000	$1.237E-4$	$1.249E-4$

Table 5
Simulation results for the lower of the two sinusoids for model (26).

σ	Method	No outlier			30% outlier		
		Average	St. Dev.	RMSE	Average	St. Dev.	RMSE
0.01	**GA-LS**	0.3000	$7.082E-5$	$7.117E-5$	0.3000	$1.296E-3$	$1.296E-3$
	GA-LTS	0.3000	$1.890E-4$	$1.918E-4$	0.3000	$4.922E-4$	$4.930E-4$
	GA-LMS	0.2999	$6.064E-4$	$6.161E-4$	0.3000	$7.171E-4$	$7.183E-4$
	GA-L1	0.3000	$9.723E-5$	$9.753E-5$	0.3000	$1.265E-4$	$1.290E-4$
	GA-LTA	0.3000	$2.703E-4$	$2.717E-4$	0.3000	$2.989E-4$	$2.997E-4$
	GA-LMA	0.3000	$6.398E-4$	$6.398E-4$	0.3000	$6.652E-4$	$6.658E-4$
0.1	**GA-LS**	0.3000	$6.021E-4$	$6.036E-4$	0.3366	$1.057E-1$	$1.119E-1$
	GA-LTS	0.3000	$6.930E-4$	$6.934E-4$	0.3000	$8.987E-4$	$8.987E-4$
	GA-LMS	0.2999	$8.027E-4$	$8.030E-4$	0.3000	$8.991E-4$	$8.996E-4$
	GA-L1	0.3000	$6.151E-4$	$6.157E-4$	0.3000	$7.816E-4$	$7.832E-4$
	GA-LTA	0.3001	$6.703E-4$	$6.800E-4$	0.3001	$7.870E-4$	$7.884E-4$
	GA-LMA	0.3000	$8.834E-4$	$8.365E-4$	0.3000	$8.966E-4$	$8.969E-4$

Table 6
Simulation results for the higher of the two sinusoids for model (26).

σ	Method	No Outlier			30% Outlier		
		Average	St. Dev.	RMSE	Average	St. Dev.	RMSE
0.01	GA-LS	0.8000	3.650E−5	3.678E−5	0.8001	8.241E−4	8.312E−4
	GA-LTS	0.8000	9.397E−5	9.416E−5	0.8000	2.040E−4	2.044E−4
	GA-LMS	0.8001	2.153E−4	2.346E−4	0.8000	3.454E−4	3.485E−4
	GA-L1	0.8000	4.911E−5	4.911E−5	0.8000	6.465E−5	6.558E−5
	GA-LTA	0.8000	9.362E−5	9.363E−5	0.8000	9.448E−5	9.450E−5
	GA-LMA	0.8000	2.230E−4	2.236E−4	0.8000	3.909E−4	3.931E−4
0.1	GA-LS	0.8000	2.962E−4	2.963E−4	0.7992	1.256E−2	1.262E−2
	GA-LTS	0.8000	3.561E−4	3.571E−4	0.7999	5.707E−4	5.743E−4
	GA-LMS	0.8000	5.257E−4	5.265E−4	0.8001	6.121E−4	6.151E−4
	GA-L1	0.8000	3.440E−4	3.450E−4	0.8000	4.283E−4	4.298E−4
	GA-LTA	0.8000	3.710E−4	3.710E−4	0.8000	4.737E−4	4.738E−4
	GA-LMA	0.8000	5.117E−4	5.121E−4	0.8000	5.516E−4	5.519E−4
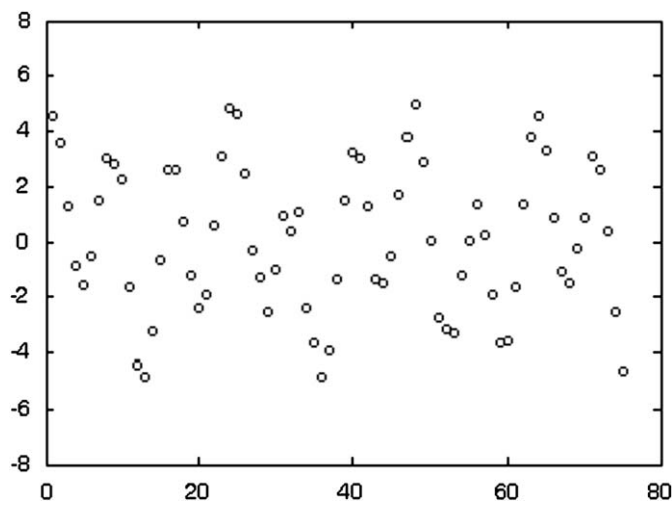
Fig. 4. A representative dataset of two-component, dependent error sinusoidal model (26) containing 30% outlier.
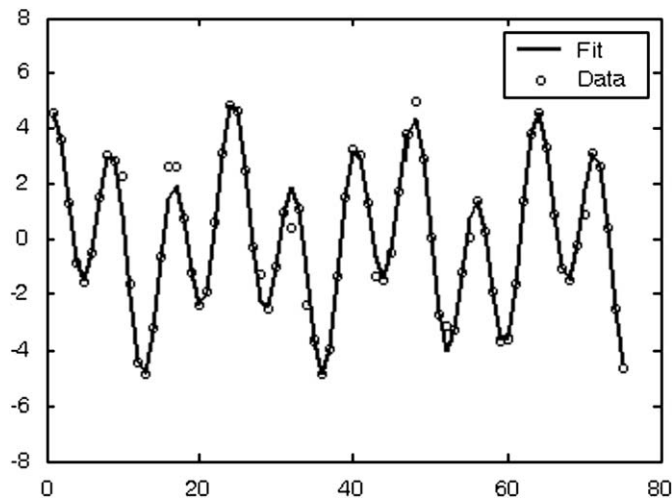

Fig. 5. A representative dataset of two-component dependent error sinusoidal model containing 30% outlier and the corresponding GA-L1fit for model (26).

1.228E−5 and that for σ=0.1 is 1.228E−4. The GA-LS almost attains these above-mentioned asymptotic values for the respective frequencies. The GA estimators based on least trimmed (GA-LTS and GA-LTA) approaches also provide reasonably accurate estimates. For the outlier contaminated data, the performance of the GA-LS deteriorates significantly, especially for the higher σ value. The GA-L1 and the GA trimmed and median based approaches appear to be fairly robust with respect to outliers in the data. Among the robust methods GA-L1 performs the best.

5.3. Real life data analysis

In this subsection, we present the real life data analysis results. Two different datasets, the 'Circadian Rhythms' data and the 'Variable Star' data, are considered for analysis.

5.3.1. Fitting Circadian Rhythms data

We consider the 'Circadian Rhythm' dataset. The data was collected at the Princeton University in the late 1960s under the direction of Dr. C.S. Pittendrich. In order to observe the periodicities in the behavior of *Perognathus formosus* (also called long-tail pocket mouse), a nocturnal mammal, the animal was given 8 days of 12 hours light and 12 hours darkness as an adjustment period, which was followed by about 73 days of constant darkness (Andrews and Herzberg, 1985). The data are temperature recordings made at 2-min intervals over 3 months. It is known that problems occurred during the experiment associated with transient failures of the monitoring equipment and with imperfections in the data logging process. As a result of which the data contains a good proportion of outliers. The data have been downloaded from http://www.statsci.org/data/general/pformosu.html.

For the analysis of the Circadian Rhythms dataset, we analyze 20-min averages of the temperatures. We fit a one-component sinusoid model of the form

$$y(t) = K + A\cos(\omega t) + B\sin(\omega t)$$

to the Circadian data using GA-LTA (50% trimming) approach. The parameter initialization for genetic search is made in the following ranges:

Parameter	Initialization range
K	$[Median(y(1)\ldots y(n)) - 50, Median(y(1)\ldots y(n)) + 50]$
A	$[-100, 100]$
B	$[-100, 100]$
ω	$[0, \pi]$

The final fitted model, for the first 8 days data, arrived after 13 GA generations is given below:

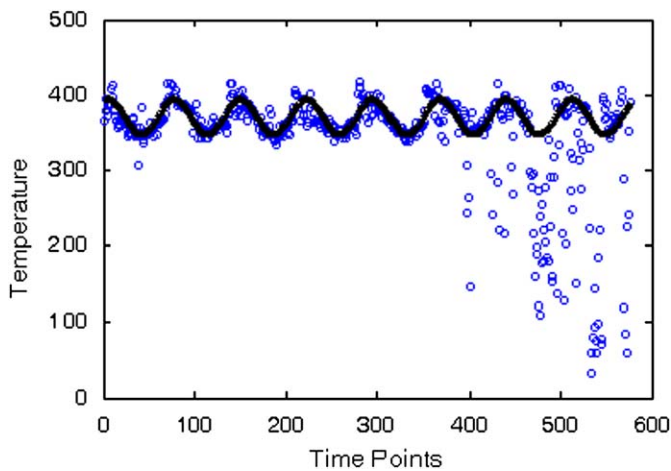$$y(t) = 370.49 + 20.69\cos(0.08672t) + 12.12\sin(0.08672t).$$

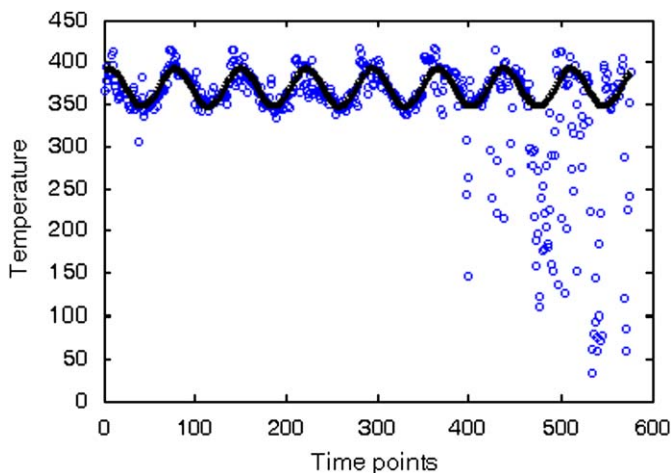**Fig. 6.** Circadian data fitting using GA-LTA.

Fig. 7. Circadian data fitting using GA-LMA.

The plot of the fitted and the observed data with outliers is given in Fig. 6.

It is obvious from the data plot that the dataset contains a large number of outliers, but the fitted curve successfully ignores them and follows nicely the periodic pattern. The less obvious outliers, closer to the fitted curve, also do not distort the data fit. We get similar results using other proposed outlier-insensitive robust frequency estimation techniques. Fig. 7 gives fit of the data using GA-LMA estimator. The fitted model under this approach is

$$y(t) = 369.84 + 19.01\cos(0.08711t) + 12.51\sin(0.0.08711\,t).$$

Similar fits are observed for other methods, except the GA-LS method, which fails completely. Considering the same dataset, it is reported in Smyth and Hawkins (2000) that the fitted frequency for the first 8 days to be 0.87273, which is very close to our frequency estimates.

5.3.2. Fitting Variable Star data

The variable star dataset is another important and very frequently used data. The determination of the periodicities of a variable star and the shape of its light curve is important in studies of stellar structure and evolution. The relationship between the period and magnitude is used to determine distances on a cosmic scale, for example. The data in this example gives observations on the magnitude of a variable star, made from the Mount Stromlo Observatory near Canberra in Australia over a period of about 250 days (Reimann, 1994). Magnitudes were
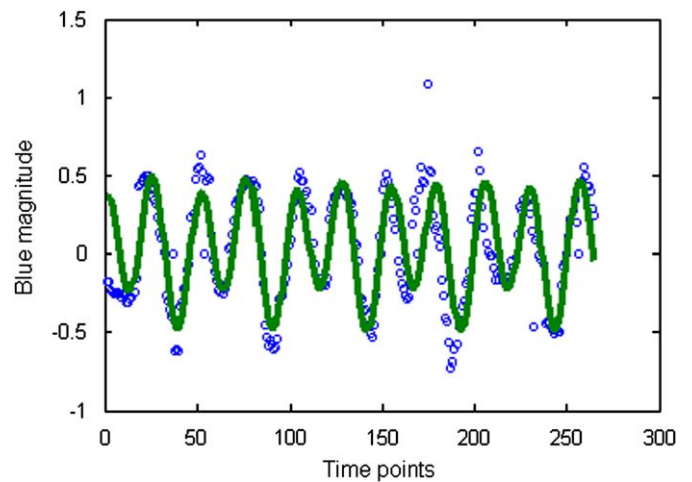

Fig. 8. Variable star (blue band) data fitting using GA-LTA.

recorded separately for the blue and red bands. Observation times were irregularly spaced depending on the conditions of sky and the observation schedule. We consider here the analysis of blue band measurements. A number of observations were considered to be unreliable due to observation conditions. The data have been downloaded from http://www.statsci.org/data/oz/ceph2.html.

We fit a two-component sinusoid model of the form

$$y(t) = K + A_1\cos(\omega_1 t) + B_1\sin(\omega_1 t) + A_2\cos(\omega_2 t) + B_2\sin(\omega_2 t)$$

using GA-LTA (50% trimming) approach. The initial population is populated from the following ranges of the respective parameters:

Parameter	Initialization range
K	$[Median(y(1)\ldots y(n)) - 0.5,$ $Median(y(1)\ldots y(n)) + 0.5]$
A_1, A_2	$[-0.5, 0.5]$
B_1, B_2	$[-0.5, 0.5]$
ω_1, ω_2	$[0, \pi]$

The final fitted model is

$$y(t) = 0.01992 - 0.0128\cos(0.1247t) + 0.2032\sin(0.1247t)$$
$$+ 0.4613\cos(0.2467t) - 0.0532\sin(0.2467t).$$

The plot of the observed data and the fitted model is given in Fig. 8. We observe from the plot that the fit ignores the outliers and is able to trace the correct sinusoidal pattern.

Smyth and Hawkins (2000) considered the same dataset for testing the usefulness of their robust frequency estimation technique. For implementation of their method, which requires time points to be equidistant, the data was first interpolated linearly onto an equally spaced grid of time points of the same length, no such preprocessing of the data is required for implementation of our methods. The estimated frequencies reported in Smyth and Hawkins (2000) are 0.126 and 0.253, which once again are close to our frequency estimates.

The GA-LTA estimates of the frequencies 0.1247 and 0.2461 correspond to periods of 50 and 25 days. The star is therefore determined to be periodic with period of about 50 days.

6. Conclusion

In this paper, we propose genetic algorithm based robust frequency estimation techniques for multiple sinusoidal models with correlated error structures. The proposed methods use genetic search technique for optimizing various outlier-insensitive

criterion functions. The methods do not require the data points to be equidistant or the noise sequence to be independent Gaussian structure, which is otherwise required in other robust frequency estimation techniques for this model (see for example Smyth and Hawkins, 2000). Furthermore, the GA based robust frequency estimation techniques search a population of possible optimal solutions in parallel and do not require derivative information or other auxiliary information, only the levels of fitness influence the direction of search. Another advantage of using the proposed methods is that since they are based on genetic algorithms they use probabilistic transition rules and have potentially high chance of converging to the optimal solution.

In the simulation studies and real life data analysis, it is observed that the proposed genetic algorithm based robust frequency estimators, optimizing outlier-insensitive criteria are able to resolve frequencies of the sinusoidal model with high degree of accuracy and provides reasonably high breakdown point robust estimates.

Acknowledgement

References

Andrews, D.F., Herzberg, A.M., 1985. In: Data: A Collection of Problems from Many Fields for the Student and Research Worker. Springer, New York.

Barrodale, I., Roberts, F.D.K., 1973. An improved algorithm for discrete L1 linear approximations. SIAM Journal of Numerical Analysis 10, 839–848.

Barrodale, I., Roberts, F.D.K., 1974. Solution of an overdetermined system of equations in the L1 norm. Communications of the ACM 17, 319–320.

Bloomfield, P., Steiger, W.L., 1983. Least Absolute Deviations: Theory, Applications, and Algorithms. Birkhauser, Boston, Mass.

Bonaventura, A., Coluccio, L., Fedele, G., 2007. Frequency estimation of multi-sinusoidal signal by multiple integrals. In: IEEE International Symposium on Signal Processing and Information Technology, pp. 564–569.

Brillinger, D.R., 1987. Fitting cosines: some procedures and some physical examples. In: MacNeill, B., Umphrey, G.J. (Eds.), Applied Probability and Stochastic Process and Sampling Theory. D. Reidel Publishing Company, USA, pp. 75–100.

Caruana, R.A., Schaffer, J.D., 1988. Representation and hidden bias: Gray vs. binary coding. In: Proceedings of the Sixth International Conference Machine Learning, pp. 153–161.

Chan, K.W., So, H.C., 2004. Accurate frequency estimation for real harmonic sinusoids. IEEE Signal Processing Letters 11 (7), 609–612.

Coluccio, L., Eisinberg, A., Fedele, G., 2008. A property of the elementary symmetric functions on the frequencies of sinusoidal signals. Issue Series Title: Signal Processing, doi:10.1016/j.sigpro.2008.10.021.

De Jong, K.A., 1975. An analysis of the behavior of a class of genetic adaptive systems. Ph.D. Thesis, University of Michigan.

Goldberg, D.E., 1989. In: Genetic Algorithms in Search, Optimization and Machine Learning. Pearson Education Inc..

Hannan, E.J., 1971. Non-linear time series regression. Journal of Applied Probability 8, 767–780.

Hollstien, R.B., 1971. Artificial genetic adaptation in computer control systems. Doctoral Dissertation, University of Michigan, Dissertation Abstracts International, 32(3), 1510B, University Microfilms No. 71-23,773.

Kahn, M., Osborne, M.R., Smyth, G.K., 1993. On the consistency of Prony's method and related algorithms. Journal of Computational and Graphical Statistics 1, 329–349.

Kay, S.M., 1988. Modern Spectral Estimation: Theory and Applications. Prentice-Hall, New York.

Kundu, D., 1993. Asymptotic theory of least-squares estimators of a particular non-linear regression model. Statistics & Probability Letters 18, 13–17.

Kundu, D., 1997. Asymptotic theory of least-squares estimators of sinusoidal signals. Statistics 30 (3), 221–238.

Kundu, D., Mitra, A., 1996. Asymptotic theory of the least-squares estimators of a nonlinear time series regression model. Communications in Statistics, Theory Methods 25, 133–141.

Mackisack, M.S., Osborne, M.R., Smyth, G.K., 1994. A modified Prony algorithm for estimating sinusoidal frequencies. Journal of Statistical Computation and Simulation 49, 111–124.

Mitra, A., Kundu, D., 1997. Consistent method of estimating sinusoidal frequencies: a non-iterative approach. Journal of Statistical Computation and Simulation 58, 171–194.

Nandi, S., Iyer, S., Kundu, D., 2002. Estimating the frequencies in presence of heavy tail errors. Statistics and Probability Letters 58 (3), 265–282.

Quinn, B.G., Hannan, E.J., 2001. In: The Estimation and Tracking of Frequency. Cambridge University Press, Cambridge.

Reimenn, J.D., 1994. Frequency estimation using unequally-spaced astronomical data. Ph.D. Thesis, University of California, Berkeley.

Rice, J.A., Rosenblatt, M., 1988. On frequency estimation. Biometrika 75, 477–484.

Rousseeuw, P.J., 1984. Least-median of squares regression. Journal of American Statistical Association 79, 871–880.

Rousseeuw, P.J., 1988. Robust estimation and identifying outliers. In: Wadsworth, H.M. (Ed.), Handbook of Statistical Methods for Engineers and Scientists. McGraw-Hill, New York (Chapter 17).

Rousseeuw, P.J., Leroy, A.M., 1987. In: Robust Regression and Outlier Detection. Wiley, New York.

Smyth, G.K., Hawkins, D.M., 2000. Robust frequency estimation using elemental sets. Journal of Computational and Graphical Statistics 9, 196–214.

Stoica, P., 1993. List of references on spectral line analysis. Signal Processing 31 (3), 329–340.

Stoica, P., Moses, R., 2005. Spectral Analysis of Signals. Prentice-Hall, Upper Saddle River, NJ.

Thierens, D., 1997. Selection schemes, elitist recombination, and selection intensity. In: Back, T. (Ed.), Proceedings of the Seventh International Conference on Genetic Algorithms. San Francisco, USA, pp. 152–159.

Trapero, J.R., Sira-Ramirez, H., Batlle, V.F., 2007. An algebraic frequency estimator for a biased and noisy sinusoidal signal. Signal Processing 87 (6), 1188–1201.

Walker, A.M., 1971. On the estimation of the Harmonic components in a time series with Stationary residuals. Biometrika 58, 21–26.

Yokose, Y., Cingoski, V., Kaneda, K., Yamashita, H., 2000. Performance comparison between gray coded and binary coded genetic algorithms for inverse shape optimization of magnetic devices. Applied Electromagnetics (pp. 115–120), 115–120.