ULTRA LOW FREQUENCY ESTIMATION: A NEGLECTED APPORTION

Alfred C.H.Tan, A. Choudhury, Y.S. Ong & S.M. Veres

Faculty of Engineering and Applied Sciences
University of Southampton
Highfield, Southampton, SO17 1BJ, UK

ABSTRACT

Most frequency estimation techniques exhibit larger estimation errors in the ultra low frequency (ULF) region (0–10 Hz asymptotically) as compared to higher frequency ranges. This paper examines this situation and proposes to model the estimation errors through function approximation thereby making a priori error estimates available for new signals frequencies. The technique is shown to perform well on simulation data under varying noise conditions. Results show that it is possible to estimate frequencies in the ULF region with accuracies similar to those encountered in higher frequency ranges.

1. INTRODUCTION

Frequencies between 0.01Hz to 10Hz fall in the ultra low frequency (ULF) range. Such frequencies occur regularly in earthquakes measurements, onshore wave ripples, electromagnetic waves [1] and structural building vibrations [2]. A frequency domain analysis of such phenomena often provides useful information which can be used for a variety of purposes such as controlling excessive vibrations or frequency pattern recognition. Inaccuracies in frequency measurements can generally be attributed to: (i) numerical errors in the computer, called finite wordlength effect, (ii) precision constraints on the frequency content in the Discrete Fourier Transform (DFT) and (iii) noisy measurements of the signals. Most analog measurement instruments do not have these shortcomings as they are built upon transistors and banks of parallel filters, performing calculations to display the spectrum in the frequency domain. Most modern digital instruments are packaged with software programs to facilitate the building of spectrum analyzers for frequency analysis. Being solely mathematical in operation, they are more stable, not prone to temperature and humidity drift, demonstrate good performance-to-cost ratio and do not suffer from manufacturing variations or aging. Therefore DSP-based techniques have quickly replaced electronics-based spectrum analyzers in recent times.

A general input signal can be characterized by sampling its direct current (d.c.) content, amplitude, phase shift and the noise sequence. For simplicity, we shall focus on single (dominant) frequency sinusoidal signals, of the form:

$$y_t = A\sin(\omega_0 t + \phi) + \epsilon_t, \quad t = 0, 1, \ldots, T-1 \qquad (1)$$

where, A denotes the amplitude of the sinusoid, ω_0 its frequency, ϕ its phase. The signal is sampled at T equidistant time instances and ϵ_t is a noise term, assumed drawn from a stationary ergodic noise process.

In digital instruments, the input signal is typically transformed into the frequency domain by means of a numerical frequency estimation method, usually fast Fourier transform (FFT). In practice, the spectrum analyst would usually read the frequency off the monitor to get an estimate $\hat{\omega}$ of the true frequency content (ω_0) of the input signal. The difference between the estimated and the true frequencies , i.e. $g(\hat{\omega}, \omega_0) = \hat{\omega} - \omega_0$, is the *estimation error* incurred in the process. As discussed in the next section, the problem of estimation errors becomes especially severe in the ULF region, our domain of interest. This paper addresses this issue through the use of ideas from the function approximation literature for finite data. In particular, we utilize simulation data to generate an error model using the technique of support vector regression (introduced later). A subsequent application of this model generates a priori error estimates for new (unknown) signal frequencies.

Perhaps the most important decision to be taken is the choice of the frequency estimation technique itself. Some of the most commonly used schemes are the nonparametric FFT, parametric minimum entropy, ARMA model based schemes [3] and subspace methods such as multiple signal classification (MUSIC) method [4]. Other techniques proposed by Rife et. al. [5], Pisarenko [6] and Mackisack et. al. [7] are either computationally intensive or low in estimation accuracy. Figure 1 shows a comparison plot of estimation errors for some of the common frequency estimation techniques mentioned above, in the ULF region for a simulated noise free time series data. While different frequency estimation techniques produce quite different error profiles, all of them can incur significant estimation errors (relative to the true frequency) in this region.

It was shown that MUSIC [4], Rife [5] and Pisarenko [6]

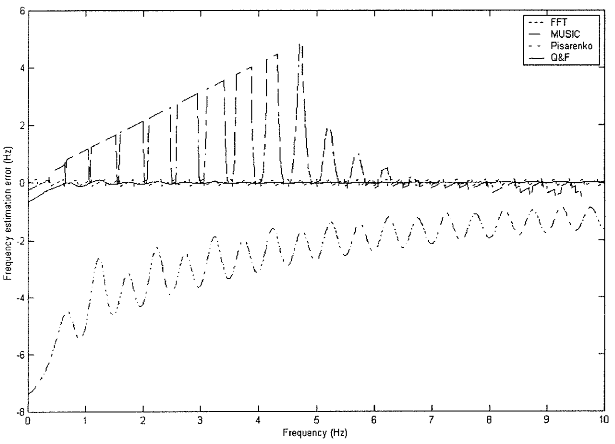

Figure 1: Different frequency estimation methods in the ULF region.

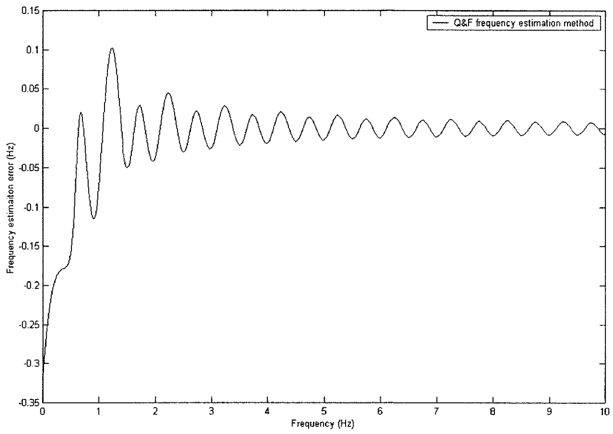

Figure 2: ULF response for Q&F frequency estimation method.

methods have a standard error of $T^{-1/2}$, while Mackisack's method [7] is accurate to the order of $T^{-5/4}$. The Quinn and Fernandes method (Q&F) [3] gives an improved level of accuracy of the order T^{-2} and is relatively computationally efficient. The superior performance of this scheme is also brought out in Figure 1. Accordingly, this technique shall form the core estimation technique in this paper.

Section 2 provides the relevant theoretical background for frequency estimation beginning with a brief description of the Q&F method. This is followed by an overview of the support vector technique for regression in section 3. Section 4 explores how these two schemes can be brought together to reduce estimation errors in the ULF region, while section 5 presents experimental investigations using the scheme proposed in this work. Finally, section 6 concludes the paper by summarizing the research presented in this work and discussing some directions of future research.

2. RELEVANT BACKGROUND

Autoregressive moving average models (ARMA) have been traditionally used to model time series observational data arising in many areas [8]. They also provide a convenient tool for estimating frequencies of sinusoidal time series data. Let $Y_t = \sum_{i=1}^{t} y_i$ and $U_t = \sum_{i=1}^{t} \epsilon_i$, then the ARMA model corresponding to the sinusoid in Equation 1 is conveniently written as:

$$Y_t - (2\cos\omega_0)Y_{t-1} + Y_{t-2} = U_t - (2\cos\omega_0)U_{t-1} + U_{t-2} \tag{2}$$

Following Quinn and Fernandes [3], we can iteratively determine ω_0 from this model in an efficient fashion. In general, their scheme (henceforth Q&F) involves iteratively estimating the parameters α and β of a second order ARMA model of the following form:

$$x(t) - \beta x(t-1) + x(t-2) = \varepsilon(t) - \alpha\varepsilon(t-1) + \varepsilon(t-2), \tag{3}$$

subject to the condition $\alpha = \beta$. For our case (see Equation 2), determining α (β) this way is tantamount to finding an estimate of the true frequency. The iteration begins by setting $\alpha_1 = 2\cos\hat{\lambda}_1$, where $\hat{\lambda}_1$ is an initial estimator of ω_0. The corresponding β is estimated as a regression coefficient of $\xi(t) + \xi(t-2)$ on $\xi(t-1)$ as follows, where $\xi(t) = y(t) + \alpha_j\xi(t-1) - \xi(t-2), j \geq 1$:

$$\beta_j = \alpha_j + 2\frac{\sum_{t=0}^{T-1} x(t)\xi(t-1)}{\sum_{t=0}^{T-1} \xi^2(t-1)} \tag{4}$$

If the unsigned difference $|\beta_j - \alpha_j|$ is suitably small, the iteration is terminated, otherwise, we set $\alpha_{j+1} = \beta_j$, and estimate β_{j+1} according to Equation 4. If the scheme terminates at the nth iteration, the estimated frequency, $\hat{\omega}_T$, is given by $cos^{-1}(\beta_n/2)$. Figure 2 shows the result of applying the Q&F technique to simulated noise free time series datasets generated from various sinusoids in the ultra low frequency region. Once again, we note that the estimation errors at lower frequencies are relatively larger than those at higher frequencies. Finally, we have observed the Q&F method to be more appropriate than the other techniques mentioned earlier for low signal to noise ratio (SNR) with a fixed sampling period T.

2.1. Noiseless case

For the case of no noise ($\epsilon_t = 0$), the estimation problem in the ARMA model in Equation 2 is reduced to a one–

dimensional minimization problem, viz.,

$$\underset{\alpha\in\mathbb{R}}{\text{Minimize}}\ (1/T)\sum_{t=0}^{T-1}(Y_t - 2\alpha Y_{t-1} + Y_{t-2})^2 \qquad (5)$$

whence the frequency estimate is given by $\hat{\omega}_0 = cos^{-1}(\alpha^*)$, where α^* denotes the optimal solution to the minimization problem 5. In practice however, one rarely encounters noise free datasets. It is therefore more desirable to build models which can accommodate varying noise levels in the data. It is for this reason that we shall be more interested in the noisy case (discussed below) in this work and shall not pursue the noiseless scenario any further.

2.2. Noisy case

In practice, the presence of noise in the data (or signal) is almost inevitable thus making it necessary to accommodate its effects in the design of estimation techniques. As mentioned earlier, the problem of estimating ω_0 in Equation 2 can then be solved according to the recipe in [3], which is known to converge to the true solution in the limiting case $(T \to \infty)$. The following result [3] relates the limiting distribution of estimation errors to the spectral density of the noise process:

Theorem. 1 (Quinn and Fernandes(1991)) *If $\hat{\omega}_T$ is the T sample estimate of the true frequency ω_0, then $T^{3/2}(\hat{\omega}_T - \omega_0)$ has a distribution converging to the normal distribution with mean zero and variance $48\pi f_\epsilon(\omega_0)/A^2$, where $f_\epsilon(\cdot)$ denotes the spectral density of the noise process.*

It is known that the Q&F estimation technique in [3] amounts to finding a local maximizer of a *smoothed periodogram* of the form $\kappa_T(\omega) = \int_{-\pi}^{\pi} I_y(\lambda)\mu_T(\omega-\lambda)d\lambda$, where $I_y(\lambda) = (2/T)|\sum_{t=0}^{T-1} y(t)\exp -it\lambda|^2$, is the periodogram, and $\mu_T(x)$ denotes a smoothing function with the property that it is roughly of the order of $\log T$ for x of order T^{-1} and tends to zero otherwise. A consequence of this fact is that the Q&F technique (despite its excellent convergence properties) is not adequate for low frequency estimation, more so in the presence of noise. This is because the assumption that the smoothed periodogram is maximized in an $\mathcal{O}(T^{-1})$ neighborhood of ω_0, need no longer hold. [1]

The main motivation of the current work is to rectify this situation and make low frequency estimation more accurate. In particular, we wish to identify frequency estimation errors through the use of a popular function approximation technique (viz., SVM), while still retaining Quinn and Fernandes's recipe as the core estimation technique. In other words, we train SVMs to indirectly estimate the noise spectral density $f_\epsilon(\cdot)$ at ω_0 and hence suggest corrections to the Q&F estimate accordingly.

[1] We mention that working with the unsmoothed periodogram is inconvenient in any case because of the presence of multiple local maxima within the $\mathcal{O}(T^{-1})$ neighborhood of ω_0, and the persistence of sidelobes asymptotically.

3. SUPPORT VECTOR MACHINES FOR REGRESSION

The Support Vector Machine (SVM) [9], derived from Vapnik's statistical learning theory has become a popular technique for learning models from data. These algorithms create a sparse decision function expansion by choosing only a select number of training points, the so-called 'support vectors'. Through the use of the so called 'kernel trick', linear function approximation algorithms involving explicit inner products between data points in an input space can be conveniently and efficiently transformed into their nonlinear generalizations. SVMs approximately implement Vapnik's *structural risk minimization principle* through a balanced tradeoff between empirical error (risk) and model complexity (measured through the VC dimension).

We consider the problem of SVM regression modeling given observational data of the form $(x_i, y_i)_{i=1}^{\ell}$ where $x_i \in \mathbb{R}^p$, denotes a p dimensional input vector and $y_i \in \mathbb{R}$, is a real valued target. We seek to model the relationship between the inputs and the output. Assume that the functional form we seek is the familiar linear function, $f(x, w, b) = \langle w, x\rangle + b$, where $w \in \mathbb{R}^p$, denotes a p dimensional vector of unknown coefficients and $b \in \mathbb{R}$ is an unknown but constant bias term. Then we aspire to find w, b so as to minimize the empirical errors while keeping a check on the L_2 norm of the weight vector $w \in \mathbb{R}^p$ for model capacity control. Formally, we pose the following basic convex programming problem :

$$\text{Minimize } (1/2)\langle w, w\rangle \qquad (6)$$

subject to the constraints

$$\begin{cases} y_i - \langle w, x_i\rangle - b & \le & \epsilon_i \\ \langle w, x_i\rangle + b - y_i & \le & \epsilon_i \end{cases} \qquad (7)$$

Since a feasible solution may not exist satisfying the above optimization problem (or we may want to tolerate some noise), we need to introduce slack variables $\xi_i, i = 1,\ldots,l$ to relax the constraints in the original optimization problem. An equivalent optimization problem with quadratic penalization on ξ_is can be formulated as follows:

$$\text{Minimize } F(\xi) = (1/2)\langle w, w\rangle + (C/2)\sum_{i=1}^{\ell}(\xi_i)^2 \qquad (8)$$

subject to the constraints:

$$\begin{cases} y_i(\langle w, x_i\rangle + b) & \ge & 1 - \xi_i \\ \xi_i & \ge & 0 \end{cases} \qquad (9)$$

The desired weight vector has the form: $w = \sum_i^\ell (\alpha_i - \alpha_i^*) x_i$, where α_i, α_i^* are non–negative Lagrange multipliers required to solve the above optimization problem. The parameter C measure a *trade–off* between empirical error and model complexity and is usually set *a priori* (through cross validation, for example). A nonlinear generalization is effected by simply noting that the resulting solution $f(x)$ can be explicitly written in terms of inner products between data points; these inner products are then replaced by a Mercer kernel $k(x, x_i)$ and the resulting solution has the form

$$f(x) = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) k(x, x_i) + b \tag{10}$$

4. ERROR CORRECTION OF FREQUENCY ESTIMATES

The frequency estimation error function would seem to depend on the amplitude (A), the true frequency (ω_0) and sampling period (T) of the signal. Of these, the sampling period T is usually predetermined and known to the analyst. Further, since the true frequency is unknown, we can only use the Q&F estimate as its replacement. Hence the estimation error can be written as :

$$g = \tilde{g}(A, \hat{\omega}_T, T) + \eta \tag{11}$$

where η is a stationary noise process used to account for anomalies in the data. To establish a suitable error model of the form in Equation 11, we tabulate estimation errors at various signal frequencies (ULF), sampling periods and noise intensities and then apply the SVM regression technique discussed in section 3 to generate an appropriate mapping. Once ready, such a model can be used to generate error estimates for a new signal (with unknown ultra low frequency) which in turn can be used to improve the estimate obtained by the Q&F method.

5. EXPERIMENTAL INVESTIGATIONS

To investigate the veracity of the proposed system, we tested it on a range of ultra low frequencies (varying between 0–10 Hz) and over various noise intensity levels. In our experiments, we used the ϵ–SVR regression machine from the LIBSVM library [10] of support vector machine techniques. The tradeoff parameter in the SVM regression scheme which reflects the balance between model complexity and empirical errors was set by 10-fold cross validation for each noise level. The data itself was generated from sinusoids with amplitude (A), frequency (ω_0) and sampling period (T) set according to the latin hypercube sampling scheme. For each such set of parameters, a time series (of T) samples was generated and corrupted with a Gaussian noise, $\mathcal{N}(0, \sigma^2)$.

Error Measure	⟵Noise Level (σ)⟶ 0.25	0.5	1.5
$\bar{g}$	0.01	0.02	0.02
$\bar{g}_0$	0.020	0.05	0.04
MSE	2.59×10^{-2}	6.12×10^{-2}	3.57×10^{-2}
R^2	0.03	0.86	0.86

Table 1: Various test set statistics measuring the performance of the proposed scheme under varying noise conditions; refer text for details.

This data was used to generate a frequency estimate $\hat{\omega}_T$ using the Q&F technique discussed in section 2. For each noise level (σ), 3000 instances were created to train the SVM regression model and another 2000 instances kept aside for testing, where each instance had the following form: $(x = \{A, \hat{\omega}_T, T\}, g(x))$.

Table 5 presents various test set statistics to measure the performance of the estimation error modeling scheme proposed above, under various noise conditions. The first block presents a comparison of the mean absolute values of the SVM outputs (i.e., predicted estimation errors) over the testing data given by $\bar{g} = \sum_{i=1}^{2000} |\tilde{g}(x_i)|$, w.r.t the benchmark estimation error of $\bar{g}_0 = \sum_{i=1}^{2000} |g(x_i)|$, where x_i's belong to the test set. The second block presents the conventional mean square error (MSE) performance measure along with the associated square correlation coefficient (R^2).

The results in the first block show that mean estimation error predictions of the model on the test set are fairly close to the true values. Similarly, the MSE values are low over all the noise levels considered in our experiments, indicating good overall performance. However, the squared correlation coefficient (R^2) values are significantly low for the low noise case. This indicates that it might not be a good idea to use our error prediction model for the low noise case (or noiseless case). Indeed, as pointed out in section 2.1, in such a case, one would be better off solving the one dimensional least squares problem in Equation 5 instead of using the $Q\&F$ technique for estimating the true frequency. Finally, we note that the high R^2 values (close to 1.0) for noise levels of $\sigma = 0.5, 1.5$ indicate the good quality of our error models under noisy scenarios. Thus we have demonstrated that it is possible to build effective error correction models which can complement frequency estimation techniques in the ULF region for noisy time series data.

6. CONCLUDING REMARKS

Most frequency estimation techniques incur relatively large errors in the ultra low frequency (ULF) range as compared to higher frequency ranges, especially in the presence of noise. This can generally be attributed to the failure of some underlying assumption(s) for the particular technique, thus reducing the efficiency of the corresponding estimator. This article proposed an error modeling approach for accurate ULF prediction to complement a core frequency estimation algorithm of choice (here the Q&F technique). The powerful function approximation technique of support vector regression was utilized to create a suitable predictive model for frequency error estimation. Given a new signal of unknown frequency, this model can supply an a priori error estimate which can be used to improve the frequency estimate obtained using the Q&F technique. Simulation results under various noise conditions show that the proposed scheme has merit and deserves further investigation.

While the present work was aimed at building frequency error estimation models for a signal, a case of more general interest is that of completely characterizing an unknown signal belonging to a given class of signals (e.g sinusoidal). Since any signal can be represented as a sum of sinusoids, it seems worthwhile to investigate such a characterization for this important class of signals. In particular, besides estimating the frequency, one needs to estimate the phase ϕ of the signal. As in the present case, one may proceed to use any reliable phase estimation technique and then build an estimation error model for it. From a modeling point of view, it would seem sensible to build a single model for predicting errors in both frequency and phase estimation. This might require examining the possible coupling effects between these two kinds of errors. While this issue has been ignored in the present article, it could play a dominant role in building integrated error models for a more complete characterization of sinusoidal signals.

Acknowledgement

The work of Alfred C.H. Tan under the supervision of Prof. S.M. Veres, is supported by an EPSRC grant number GR / N32297 / 01, while that of A. Choudhury and Y.S. Ong are supported by the BAe-RR University Technology Partnership for Design.

7. REFERENCES

[1] A.K., Banerjee, M.N. Alam, and A.A. Mamun, Ultra-low-frequency dust-electromagnetic modes in self-gravitating magnetized dusty plasmas, *Pramana - Journal of Physics*, Volume 56(5), pp. 643-656, May 2001.

[2] R.G. Shoulberg, and R.H. Tuft, Equipment considerations for ultra low frequency model tests, *Shock Vibration Bulletin*, U.S. Naval Research Laboratory, n. 42 Pt. 3, pp. 25-37, January 1972.

[3] B.G. Quinn, and J.M. Fernandes, A fast efficient technique for the estimation of frequency, *Biometrika*, Volume 78(3), pp. 489–497, 1991.

[4] R.O. Schmidt, Multiple Emitter Location and Signal Parameter Estimation, *IEEE Transaction on Antennas and Propagation*, Volume 34, pp. 276-280, 1986.

[5] D.C. Rife, and G. A.Vincent, Use of the discrete Fourier transform in the measurement of frequencies and levels of tones, *Bell Systems Technical Journal*, Volume 49, pp. 197–228, 1970.

[6] V.F. Pisarenko, The retrieval of harmonics from a covariance function, *Geophysical Journal of the Royal Astronomical Society*, Volume 33, pp. 347–366, 1973.

[7] Mackisack, M. S. and Poskitt, D. S., Autoregressive frequency estimation, *Biometrika*, Volume 76, pp. 565–575, 1989.

[8] D.J. Aigner, A compendium on estimation of the autoregressive moving average model from time series data, *International Economic Review*, Volume 12(3), pp. 348-371, October, 1971.

[9] Vapnik, V., *Statistical Learning Theory*, John Wiley & Sons, 1998.

[10] C.C. Chang, and C.J. Lin, LIBSVM : a library for support vector machines, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`