A Reappraisal of Frequency Domain Techniques for Assessing Frequency Stability Measurements

by

Donald B. Percival
Applied Physics Laboratory
Department of Statistics
University of Washington
Seattle, WA 98105

Abstract

In contrast to common practise in many other physical sciences, the statistical analysis of PTTI data is often based directly on time domain techniques rather than on frequency domain (spectral analysis) techniques. The predominant analysis technique in the PTTI community, namely, the two-sample (or Allan) variance, is often used to indirectly infer frequency domain properties under the assumption of a power-law spectrum. Here we argue that direct use and estimation of the spectrum of PTTI data have a number of potential advantages. First, spectral estimators are typically scaled independent chi-square random variables with a known number of degrees of freedom. These properties allow easy computation of the variance of estimators of various quantities that are direct functions of the spectrum. Second, the effect of detrending data can be quantified more easily in the frequency domain than in the time domain. Third, the variance of estimators of the two-sample variance can be expressed in terms of readily estimated spectral density functions. This allows one to generate confidence intervals for the two-sample variance without explicitly assuming a statistical model. Fourth, there exist tractable statistical techniques for estimating the spectrum from data sampled on an unequally spaced grid or from data corrupted by a small proportion of additive outliers. The two-sample variance cannot be readily generalized to these situations.

1. Introduction

Statistical techniques for the analysis of data indexed by time fall roughly into two different categories, namely, time domain techniques and frequency domain techniques. Although the correlation structure of Gaussian stationary processes can be completely characterized in either the time domain through the autocovariance function (acvf) or the frequency domain through the spectrum (or, equivalently, the spectral density function (sdf) when it exists), the preferred characterization for statistical analysis is the spectrum for three reasons. First, the spectrum is much easier to interpret physically than the acvf since the former can be related simply to power output from a narrow band-pass filter. Second, the statistical properties of estimators of the spectrum are much more tractable than those of the acvf. Third, the effect of linear operations is more easily expressed for the spectrum than for the acvf.

Precise time and time interval (PTTI) data is often analyzed using a specialized time domain technique called the two-sample (or Allan) variance. Part of the appeal of this technique lies in the fact that it can be used to infer the sdf for processes with a power-law sdf. It is thus a time domain technique with a frequency domain orientation which allows it to be physically interpreted. However, the sampling properties of the standard estimators of the two-sample variance are as undesirable as those of the acvf: the variance and covariance of the estimators depend in a complicated way upon the true sdf, the specific sampling times involved, and the number of data points available. This hampers the ability of data analysts to make meaningful statistical statements about certain quantities of interest. In addition, the effect of linear operations on data are difficult to express with the two-sample variance.

The central theme of this paper is that, since the two-sample variance is closely related to the frequency domain, use of direct frequency domain techniques (spectral analysis) both complements and extends the usefulness of the two-sample variance in a number of areas where sole reliance on the latter can lead to difficult statistical problems. After we establish some notation and review the relationship of the two-sample variance to the sdf in Section 2, we give examples of the usefulness and complementary nature of frequency domain techniques in the sections that follow.

2. The Two-Sample Variance and the Spectral Density Function

Let us assume that we observe a portion $y_1, y_2, \ldots, y_N$ of length N of $\{y_t\}$, a real-valued stationary process with sdf given by $S_y(\cdot)$. Here y_t represents the value of the process at time t. The sampling time between observations is assumed for convenience to be 1, which sets the Nyquist frequency at 1/2. By definition, the two-sample variance for sampling time τ is given by

$$\sigma_y^2(2;\tau) \equiv \tfrac{1}{2}E\{[\bar{y}_t(\tau) - \bar{y}_{t-\tau}(\tau)]^2\},$$

where

$$\bar{y}_t(\tau) \equiv \frac{1}{\tau}\sum_{t=0}^{\tau-1} y_{t-\tau}.$$

Thus $\sigma_y^2(2;\tau)$ is simply half the variance of

$$\bar{z}_t(\tau) \equiv \bar{y}_t(\tau) - \bar{y}_{t-\tau}(\tau),$$

a process whose sdf $S_{\bar{z}(\tau)}(\cdot)$ can be readily related to $S_y(\cdot)$ by using the theory of linear filters:

$$S_{\bar{z}(\tau)}(f) = \frac{4\sin^4 \pi f \tau}{\tau^2 \sin^2 \pi f} S_y(f) \equiv G_\tau(f) S_y(f).$$

Figure 1 shows $G_\tau(\cdot)$ for $\tau = 1$, 4, and 16. Since

$$\sigma_y^2(2;\tau) = \frac{1}{2}\int_{-1/2}^{1/2} S_{\bar{z}(\tau)}(f)\,df = \frac{1}{2}\int_{-1/2}^{1/2} G_\tau(f) S_y(f)\,df,$$

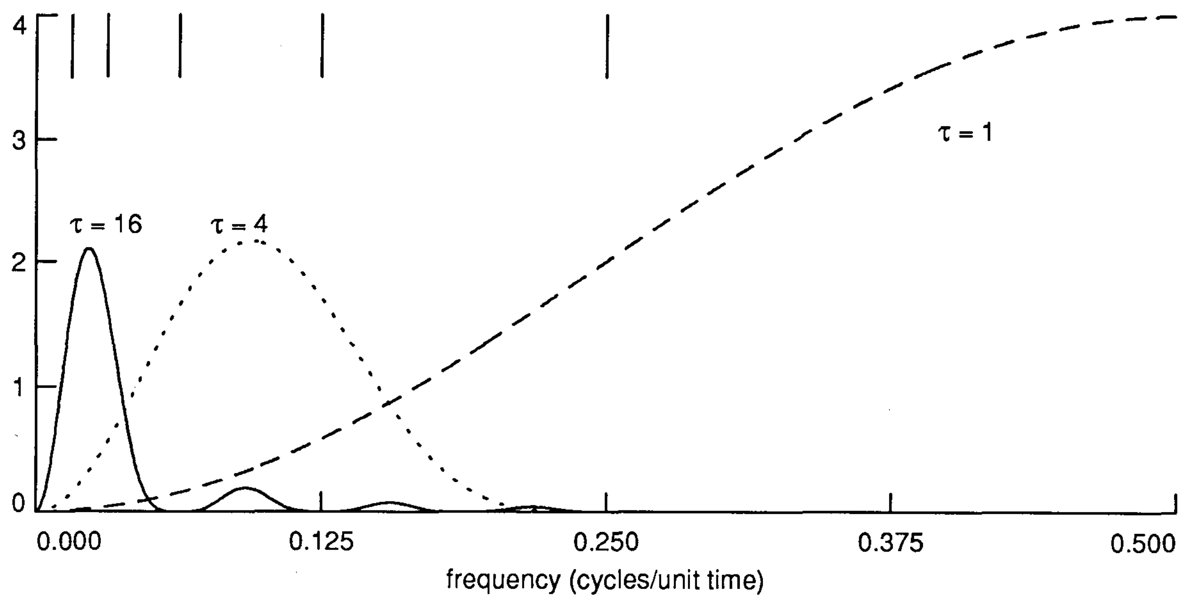

Figure 1. $G_\tau(\cdot)$ for $\tau = 1$ (dashed line), $\tau = 4$ (dotted line), and $\tau = 16$ (solid line). The five vertical lines in the upper left portion of the plot mark the positions of the frequencies 1/64, 1/32, 1/16, 1/8, and 1/4. Note that $G_1(\cdot)$ is concentrated mainly in the frequency interval [1/4, 1/2]; $G_4(\cdot)$, in [1/16, 1/8]; and $G_{16}(\cdot)$, in [1/64, 1/32].

we see that $\sigma_y^2(2;1)$ is related to the output power of a board-band high-pass filter applied to $\{y_t\}$. For larger values of τ, $\sigma_y^2(2;\tau)$ is related approximately to the output power of a band-pass filter over the frequency range from $1/4\tau$ to $1/2\tau$. The width and central frequency of the filter, namely, $1/4\tau$ and $3/8\tau$, both decrease as τ increases.

One simple estimator of the two-sample variance is related to an sdf estimation scheme called a *pilot analysis* (see Section 7.3.2 of Jenkins and Watts (1968)). This estimator is defined by

$$\tilde{\sigma}_y^2(2;\tau) \equiv \frac{1}{2M}\sum_{k=1}^{M}[\bar{y}_{2k\tau}(\tau) - \bar{y}_{(2k-1)\tau}(\tau)]^2 ,$$

where M is the largest integer less than or equal to $N/2\tau$. If $N = 2^p$ for some integer p, it can be shown that the sample variance can be decomposed in terms of $\tilde{\sigma}_y^2(2;\tau)$ as follows:

$$\frac{1}{N}\sum_{t=1}^{N}(y_t - \bar{y})^2 = \frac{1}{2}\sum_{k=0}^{p-1}\tilde{\sigma}_y^2(2;2^k), \qquad \text{where} \qquad \bar{y} \equiv \frac{1}{N}\sum_{t=1}^{N} y_t . \tag{1}$$

A pilot analysis consists of using $\frac{1}{2}\tilde{\sigma}_y^2(2;2^k)$ to estimate the sdf in the frequency range $[1/2^{k+2}, 1/2^{k+1}]$.

As Jenkins and Watts point out, this sdf estimation scheme is quite crude and should not be regarded as a replacement for more serious estimators. The original motivation for using pilot analyses was that, to quote Jenkins and Watts, " ... [they] are easily carried out without using an automatic computer ...," a feature that was important at the time the authors wrote their book in the 1960's but is of limited value today. From a statistical point of view, a pilot analysis is a poor estimate of the sdf both because of its inherent lack of resolution and because of the significant correlation between, say, $\tilde{\sigma}_y^2(2;2^k)$ and $\tilde{\sigma}_y^2(2;2^{k+1})$. This correlation arises because of the significant overlap of the regions of the sdf that determine $\sigma_y^2(2;2^k)$ and $\sigma_y^2(2;2^{k+1})$. The overlap reflects the fact that the transfer function associated with the two-sample variance is only a crude approximation to that of a band-pass filter (see Figure 1). One of the main reasons for the popularity of spectral analysis is that good sdf estimators are approximately uncorrelated for estimates separated in frequency by typically a multiple of $1/N$. This allows a data analyst to make statements about the confidence of quantities calculated from spectral estimates without overly restrictive assumptions. As we argue in subsequent sections of this paper, this does not hold for the two-sample variance.

Nonetheless, the two-sample variance is important because it tells us which portions of the sdf are important for measuring frequency stability in the time domain for various sampling times. Accordingly, we follow Rutman (1978) and define a *band-pass variance*:

$$\beta_y^2(\tau) \equiv 2\int_{1/4\tau}^{1/2\tau} S_y(f)\,df$$

(the factor of 2 above is due to the fact that $S_y(\cdot)$ is a two-sided sdf). The rationale behind considering this quantity is that, whereas $\sigma_y^2(2;\tau)$ has an associated transfer function that is *approximately* that of a band-pass filter for the interval $[1/4\tau, 1/2\tau]$, the transfer function for $\beta_y^2(\tau)$ is *exactly* so. We may derive estimators for $\beta_y^2(\tau)$ by appropriate integration of a good quality sdf estimator. In contrast to estimators of $\sigma_y^2(2;\tau)$, the statistical properties of estimators of $\beta_y^2(\tau)$ are tractable under rather mild assumptions. This is due simply to the fact that the latter can be estimated directly in terms of sdf estimators, which are approximately uncorrelated on a known grid of frequencies.

As pointed out by Rutman (1978), the two-sample variance and the band-pass variance are closely related. Thus, for a power-law sdf of the form

$$S_y(f) = h_\alpha f^\alpha\,,$$

a quick calculation shows that the band-pass variance mimics the two-sample variance in that

$$\beta_y^2(\tau) = \frac{C_\alpha}{\tau^{\alpha+1}}\,,$$

where C_α depends only on α and not τ. In contrast to the two-sample variance, the band-pass variance is well-defined for *all* values of α. Moreover, in consideration of Equation (1) and the fact that the variance σ_y^2 of $\{y_t\}$ can be expressed as

$$\sigma_y^2 = \sum_{k=0}^{\infty} \beta_y^2(2^k)\,,$$

it is plausible that, for certain power-law processes,

$$\beta_y^2(\tau) \approx \tfrac{1}{2}\sigma_y^2(2;\tau)\,. \tag{2}$$

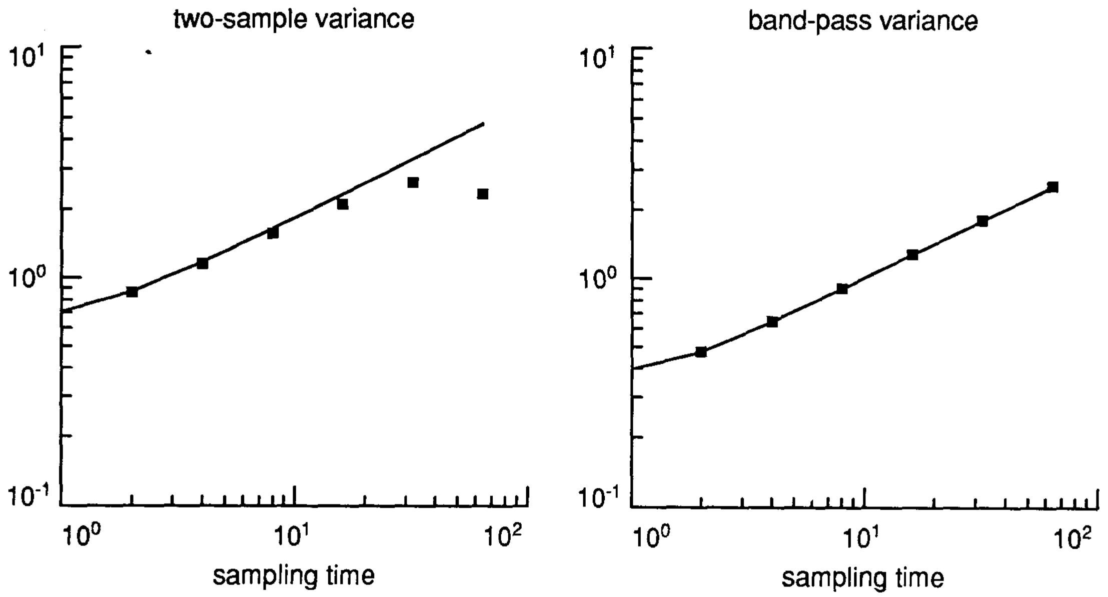

Figure 2. Comparison of the first moment (bias) properties of the two-sample variance and the band-pass variance after removal of linear drift. The solid lines show $\sigma_y(2;\tau)$ (left plot) and $\beta_y(\tau)$ (right plot) as functions of τ. The solid squares show the square roots of the expected values of estimates of $\sigma_y^2(2;\tau)$ (left plot) and $\beta_y^2(\tau)$ (right plot) for drift-corrected data.

3. The Frequency Drift Problem

A common problem in the analysis of PTTI data is the presence of linear (or quadratic) drifts in frequency. For example, suppose that we are interested in the stability properties of $\{y_t\}$, but that we actually observe

$$y'_t \equiv a + bt + y_t\,, \qquad t = 1, \ldots, N\,,$$

where a and b are unknown parameters. It is well known that, for b large enough, use of the original y'_t data yields a upwardly biased estimate of the two-sample variance $\sigma_y^2(2;\tau)$. The usual solution to this problem is to estimate b by, say,

$$\hat{b} \equiv \frac{y'_N - y'_1}{N-1}\,,$$

and to estimate $\sigma_y^2(2;\tau)$ using

$$\hat{y}_t \equiv y'_t - \hat{b}t$$

in place of the original data. It can be shown (see Percival (1983)) that this yields a downwardly biased estimator of $\sigma_y^2(2;\tau)$ for τ close to $N/2$. For example, the solid line in the left-hand plot of Figure 2 shows $\sigma_y(2;\tau)$ as a function of τ for a random walk process, whereas the solid squares show the square root of the expected value of the usual estimators of $\sigma_y^2(2;\tau)$ for $N = 128$ observations after a drift component has been removed from the data. For $\tau = 64$, the expected value of the drift-corrected estimator of the two-sample variance is about a factor of 2 below the true value. It can be shown that this bias increases when a quadratic term is removed instead of just a linear term. There appears to be no easy way to correct for this bias using time domain techniques.

This bias problem can be lessened by use of the band-pass variance. In this approach we consider the sdf $S_z(\cdot)$ of the first difference of our original data, namely,

$$z_t \equiv y'_t - y'_{t-1} = b + y_t - y_{t-1} \, .$$

This differencing operation reduces the linear drift term to a constant offset, yet it allows us to recover $S_y(\cdot)$ since the sdf's of $\{z_t\}$ and $\{y_t\}$ can be related using the theory of linear filters:

$$S_y(f) = \frac{S_z(f)}{4\sin^2 \pi f} \, .$$

We may estimate the band-pass variance by first estimating $S_z(\cdot)$ directly and then $S_y(\cdot)$ indirectly using the above equation (followed by an appropriate integration). To return to the example sited previously, the solid line in right-hand plot of Figure 2 shows $\beta_y(\tau)$ as a function of τ, and the solid square dots, the square root of the expected value of a frequency domain based estimator of $\beta_y^2(\tau)$ based upon $\{z_t\}$. We see that the estimator is essentially unbiased for this special important case of a random walk process. If we compare the solid lines in the two plots in Figure 2, we see that there is a systematic difference of about $\sqrt{2}$ between $\beta_y(\tau)$ and $\sigma_y(2;\tau)$ as suggested by Display (2). (This procedure has not been thoroughly tested for processes other than a random walk, but there are reasons to believe that the bias reduction will be quite good in other cases. Quadratic terms can be similarly dealt with by using second differences of the original data.).

Why do estimators of the band-pass variance have better first moment properties than those of the two-sample variance? Since the band-pass variance is more closely tied to the frequency domain, linear operations such as differencing are analytically tractable — a feature that does not hold for the two-sample variance. In addition to the better first moment properties, it can be shown that the band-pass variance also has tractable second moment properties in the drift removal problem (again in contrast to the two-sample variance).

4. Estimation of Parameters of Power-Law Processes

Suppose that $\{y_t\}$ is a power-law process with sdf

$$S_y(f) = h_\alpha f^\alpha \, , \qquad |f| \leq \tfrac{1}{2} \, , \tag{3}$$

where the exponent α and coefficient h_α are unknown parameters. These must be estimated from available data, say, $y_1, \ldots, y_N$, where, for notational convenience, we assume the $N = 2^p$ for some integer p. We compare here two different estimation schemes, one based on the two-sample variance, and the other, directly on the sdf.

The two-sample variance scheme is based on the well-known result that, to a good approximation,

$$\sigma_y^2(2;\tau) \approx \frac{A_\alpha}{\tau^{\alpha+1}} \, , \tag{4}$$

where A_α depends only on α and h_α but not τ. If we estimate $\sigma_y^2(2;\tau)$ by $\hat{\sigma}_y^2(2;\tau)$, we may use the following regression model to estimate α and h_α indirectly:

$$u_k = \delta + \beta v_k + \eta_k\,, \qquad k = 0, 1, \ldots, p-1\,, \tag{5}$$

where

$$u_k \equiv \log \hat{\sigma}_y^2(2; 2^k)\,; \quad \delta \equiv \log A_\alpha\,; \quad \beta \equiv -(\alpha+1)\,; \quad v_k \equiv k \log 2\,;$$

and $\{\eta_k\}$ is a sequence of error terms. Unfortunately the statistical properties of the error terms do not match those of classical regression models: both $E\{\eta_k\}$ and $var\{\eta_k\}$ depend upon k and the unknown exponent α, and $cov\{\eta_j, \eta_k\} \neq 0$ (particularly for $j = k \pm 1$). Nonetheless, it is still possible to obtain ordinary least squares estimates of δ and β (and hence A_α and α) from the above model; it is not possible to obtain meaningful measures of the statistical variability of these estimates directly from the model.

The sdf estimation method is based upon unsmoothed (but possibly tapered) direct estimates $\hat{S}_y(f_k)$ of the sdf over a grid of frequencies $\{f_k\}$ (typically $f_k = k/N$, but certain data tapers may require the use of a slightly coarser grid). From Equation (cc) we can formulate the following regression model:

$$w_k = \gamma + \alpha x_k + \nu_k\,, \qquad k = 1, \ldots, M\,,$$

where

$$w_k \equiv \log \hat{S}_y(f_k)\,; \quad \gamma \equiv \log h_\alpha\,; \quad v_k \equiv \log f_k\,;$$

$\{\nu_k\}$ is a sequence of error terms; and M is the number of frequencies in the grid (usually $M = N/2$). Because spectral estimators are typically independently distributed scaled chi-square random variables with a known number of degrees of freedom, the statistical properties of the error sequence are close to those of classical regression models: $E\{\nu_k\} \neq 0$, but it is a constant that depends only on the number of degrees of freedom of the spectral estimator; $var\{\nu_k\}$ is a *known* constant; and $cov\{\nu_j, \nu_k\} \approx 0$ for $j \neq k$. Thus, it is possible to obtain not only ordinary least squares estimates of γ (and hence A_α) and α, but also meaningful internal and external measures of the statistical variability of these estimates. (Further details on this estimation technique can be found in a thesis by Mohr (1981).)

The relative merits of these two estimation schemes were investigated by generating a thousand different realizations of length $N = 128$ of a Gaussian white noise process with variance 1. For this special case,

$$\alpha = 0\,; \quad S_y(f) = 1\,; \qquad \text{and} \qquad \sigma_y^2(2;\tau) = \tau^{-1}\,,$$

so Equation (dd) holds exactly. The average of the thousand different estimates of α for the two methods indicated that, while the sdf method was essentially unbiased ($\hat{\alpha} = -0.002$), the two-sample variance method was significantly biased ($\hat{\alpha} = 0.24$ instead of 0). This bias can be attributed to the fact that $E\{\eta_{p-1}\}$ in Model (5) is quite different from $E\{\eta_k\}$ for $k < p-1$. When this term is dropped from the regression model, the properties of the two-sample variance estimate improved considerably ($\hat{\alpha} = 0.01$). The real advantage of the sdf approach, however, is that, in contrast to the two-sample variance, one can readily calculate the variance of the estimated parameters directly from the regression model.

5. Estimation of Variance of Two-Sample Variance Estimators

The usual approach to estimating the variance of two-sample variance estimators requires one to specify a particular model for the data (see Lesage and Audoin (1977) and Yoshimura (1978)). This is unsatisfactory from both a data analytic and an operational point of view. The usual procedure seems to be to, first, plot the estimated two-sample variance as a function of τ; second, make a judgement about what pure power-law model seems appropriate for various sampling times; and third, calculate the variance estimate based upon this assumed model. The data analytic problem here is that the theoretical works referenced above assume an *a prior* known pure power-law model and not a composite power-law model determined from the data; the operational problem is the difficulty in automating this procedure for use on a digital computer.

There is an alternative approach to this problem. It can be shown that, for large N, the variance of commonly used two-sample variance estimators can be expressed in terms of an integral involving the sdf of $\{y_t\}$. For example, suppose that we consider the fully overlapped estimator of the two sample variance:

$$\hat{\sigma}_y^2(2;\tau) \equiv \frac{1}{2(N-2\tau+1)} \sum_{t=2\tau}^{N} \bar{z}_t^2(\tau)$$

(using the notation of Section 2). If $\{y_t\}$ is a Gaussian process, it can be shown (see Percival (1983)) that $\hat{\sigma}_y^2(2;\tau)$ is asymptotically normally distributed with mean $\sigma_y^2(2;\tau)$ and asymptotic variance

$$avar\{\hat{\sigma}_y^2(2;\tau)\} = \frac{1}{2(N-2\tau+1)} \int_{-1/2}^{1/2} \frac{16\sin^8(\tau\pi f)}{\tau^4 \sin^4(\pi f)} S_y^2(f)\, df\,. \tag{6}$$

Suppose that we estimate $S_y(f)$ by

$$\hat{S}_y(f) \equiv \frac{1}{N} |\sum_{t=1}^{N} h_t y_t e^{-i2\pi f t}|^2\,,$$

where $\{h_t\}$ is a data taper normalized such that $\sum h_t^2 = N$. We may then estimate $avar\{\hat{\sigma}_y^2(2;\tau)\}$ by replacing $S_y(f)$ with $\hat{S}_y(f)$ in Equation (6). The resulting integral may be computed either by an exact technique (using Parseval's theorem) or by numerical integration. Further work is needed to assess the usefulness of this technique (particularly for cases where N is small), but it is a promising automatic non-parametric approach for estimating $var\{\hat{\sigma}_y^2(2;\tau)\}$.

6. Detection of Periodic Components

One of the chief uses of spectral analysis is in the detection of narrow-band enhancements of power. In PTTI data, these enhancements might be an indication of an undesirable environmental influence on an oscillator. Here we show with an artificial example

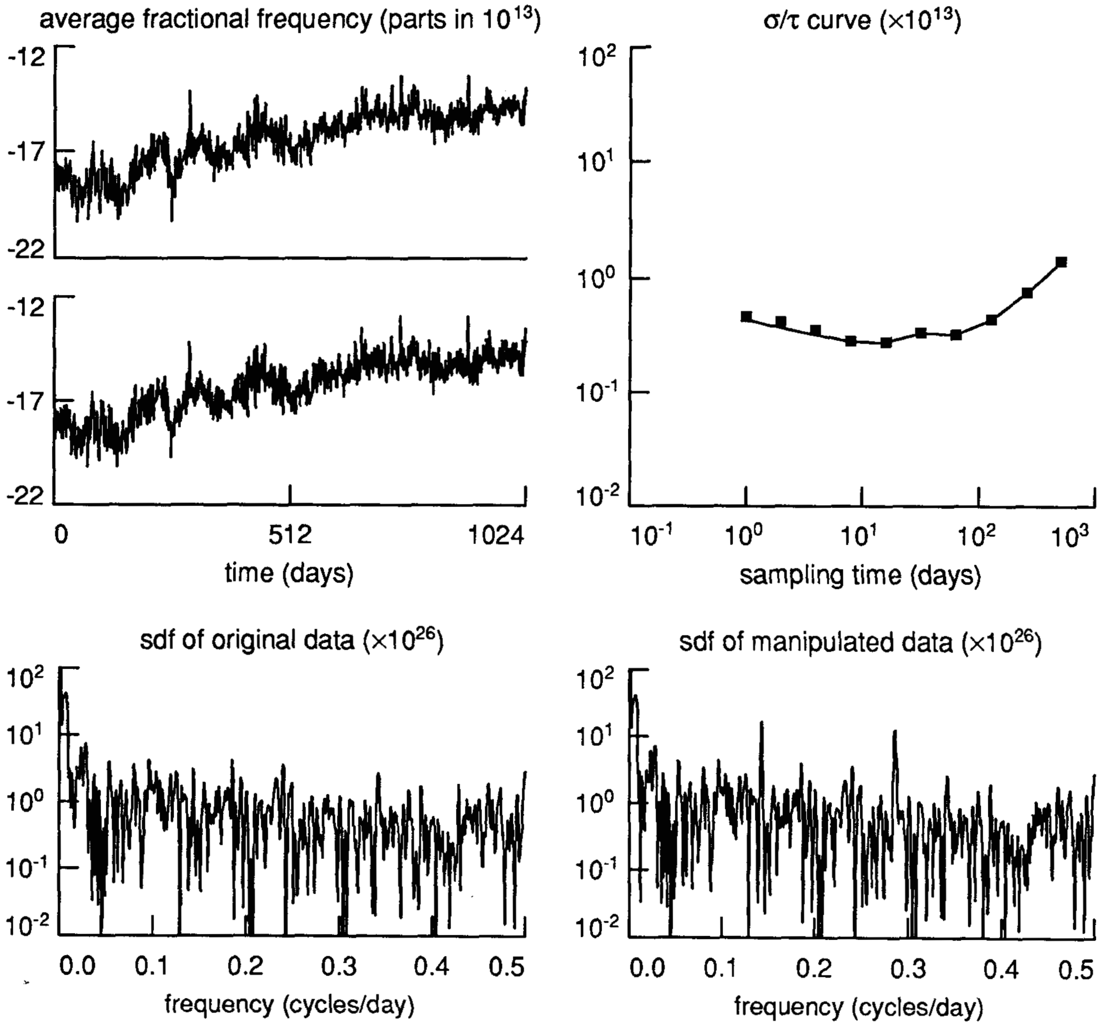

Figure 3. Effect of a narrow-band spectral component on the estimated two-sample variance and sdf. The two small plots in the upper left-hand corner show the original (top) and manipulated (bottom) clock data. The estimated two-sample variance for these series are shown, respectively, by the solid line and the solid squares in the upper right-hand plot; the corresponding estimated sdf's are shown, respectively, in the lower left-hand and right-hand plots.

that a plot of the estimated two-sample variance at sampling times that are powers of two (a common practice) can completely fail to give any indication of narrow-band features in the data. This failure points out one of the chief dangers in sole use of the two-sample variance and argues for routine use of spectral analysis (particularly for exploratory data analysis).

Our example concerns 1024 daily average fractional frequency deviates of a cesium beam atomic clock compared to a Naval Observatory clock time scale (upper left-hand plot of Figure 3). We have simulated a "weekend" environmental effect by adding 4×10^{-14} to each deviate occurring on Saturday or Sunday. The manipulated data are plotted below the original data in Figure 3. There is no important visual difference between the two series. The estimated two-sample variances for the original and manipulated data are shown, respectively, as a solid line and solid squares in the upper right-hand plot. These are quite similar to each other. Tapered (but unsmoothed) estimates of the sdf of the original and manipulated data are shown, respectively, on the lower left-hand and right-hand plots. The two additional peaks are prominent in the right-hand plot. These occur at the fundamental frequency corresponding to a period of one week and at harmonics associated with that frequency. The "weekend" effect stands out prominently in the sdf but not the two-sample variance.

It should be noted that, if the estimated two-sample variances were plotted for *all* possible values of τ (instead of just for a logarithmically spaced subset as is usually done), the narrow-band feature would manifest itself as an oscillation in the one portion of the plot. This would be a tip-off to an experienced analyst that a narrow-band feature was present in the data, but it would be extremely difficult to determine the exact nature of the feature without the aid of the estimated sdf.

7. Conclusions

We have argued in this paper that direct use of frequency domain techniques can be lead to a qualitative improvement in the analysis of PTTI data. From the point of view of statistical analysis, these improvements are mainly due, first, to the statistical nature of spectral estimators, which are (to a very good approximation) independent of each other on a known grid of frequencies, and, second, to the tractable response of the spectrum under *linear filtering — neither of which are shared by the two-sample variance*. Although the chief difference between the statistical properties of estimators based upon the spectrum and of those based upon the two-sample variance is that the former have more tractable second moment properties, there are some small (but important) improvements in first moment properties (see Sections 3 and 4).

There is, however, a qualitative improvement that can be expected from a second point of view, namely, that of exploratory data analysis (EDA). For our purposes here, EDA can be regarded as the search for interesting (and perhaps unexpected) features in data. This aspect of spectral analysis was touched upon in Section 6. In fact, spectral analysis is a prime example of an EDA tool: Tukey (1984) has stated "... it was my experience with the practice of spectrum analysis that led to my recognition of the importance of exploration in more general data analysis." Because statistics such as the two-sample variance and the band-pass variance are broad-band summaries of spectral properties, they cannot be a substitute for spectral analysis in EDA. In the view of this author, one does not know that the two-sample or band-pass variance is a meaningful measure of oscillator performance until *after* a spectral analysis has been done.

Let us close with a few remarks about the status of modern spectral analysis. A recent major advance in the subject of spectral estimation is the multiple orthogonal data taper approach due to Thomson (1982). This approach quantifies clearly the tradeoffs between resolution, bias and variance of spectral estimators. There is an extension of Thomson's approach (due to Bronez (1985)) that works for data collected at unequally spaced intervals. Chave, Thomson, and Anders (1987) give details on a robust spectrum estimation scheme which works well for data corrupted by a small portion (say, 10%) of additive outliers. Finally, with the advent of relatively low cost, yet powerful, personal computers (such as the Macintosh II and forthcoming versions of the IBM PS II), the computational cost of doing spectral analysis should no longer present any problem.

8. Acknowledgements

This work was sponsored by the Space Applications Branch, Aerospace Systems Division, of the Naval Research Laboratory. The author would like to particularly thank Jim Buisson and Tom McCaskill of that organization for their helpful critiques and comments on the subject matter of this paper.

9. References

[1] Bronez, T. P. (1985), *Nonparametric Spectral Estimation of Irregularly-Sampled Multidimensional Processes*, unpublished PhD dissertation, Arizona State University.

[2] Chave, A. D., D. J. Thomson, and M. E. Ander (1987), "On the Robust Estimation of Power Spectra, Coherences, and Transfer Functions," *Journal of Geophysical Research*, **92**(B1), pp. 633–648.

[3] Jenkins, G. M., and D. G. Watts (1968), *Spectral Analysis and its Applications*, Holden-Day, Inc., San Francisco, California.

[4] Lesage, P., and C. Audoin (1977), "Estimation of the Two-Sample Variance with a Limited Number of Data," *Proc. 31th Annual Symp. on Freq. Control*, Atlantic City, NJ, pp. 311–318.

[5] Mohr, D. L. (1981), *Modeling Data as a Fractional Gaussian Noise*, unpublished PhD dissertation, Princeton University.

[6] Percival, D. B. (1983), *The Statistics of Long Memory Processes*, unpublished PhD dissertation, University of Washington.

[7] Rutman, J. (1978), "Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress," *Proc. IEEE*, **66**(9), pp. 1048–1075.

[8] Thomson, D. J. (1982), "Spectrum Estimation and Harmonic Analysis," *Proc. IEEE*, **70**(9), pp. 1055–1096.

[9] Tukey, J. W. (1984), "Data Analysis: History and Prospects," in *Statistics: An Appraisal*, edited by H. A. David and H. T. David, Iowa State University Press, pp. 183–202.

[10] Yoshimura, K. (1978), "Characterization of Frequency Stability: Uncertainty due to the Autocorrelation of the Frequency Fluctuations," *IEEE Transactions on Instrumentation and Measurement*, **IM-27**(1), pp. 1–7.

QUESTIONS AND ANSWERS

Mark Weiss, National Bureau of Standards: Can you say something about the confidence of the estimate of the spectrum versus the sigma tau curve. Do you get a better confidence on the power spectrum? Are there ways of improving the confidence level on the Fourier spectrum?

Prof. Percival: Some of the work that Thompson did in '82 solved these questions. What he advocates is a multiple orthogonal data taper method. That sounds frightening but after you look at the work you ask why nobody thought of this before. Actually, somebody did think of it before—Lord Rayleigh, 100 years ago and his work has just now surfaced. It turns out that you can actually get very good estimates of the confidence on the spectrum and good bounds on what he terms the local bias and broad band bias. That is bias due to smearing out locally and bias due to components far away in frequency from the area that you are interested in. Spectral analysis has taken a real leap forward in the last five years due to this work.

Mr. Weiss: Are you saying that is as good, or better?

Prof. Percival: It would be hard to compare them because they are two different things, the sigma tau curve is one thing and the spectrum is another quantity altogether. Out at the $n/2$ case which is a X^2 random variable which has two degrees of freedom. That is kind the fundamental, unsmoothed, sampling properties of the spectrum. Then it can't be any worse than the spectrum and I think that they might be better. The point is that with the spectrum you have tractable statistical properties, but with the Allan variance, because of the correlation between estimators, things tend to get very sticky when you try to combine things.

Charles Greenhall, Jet Propulsion Laboratory: How does the Beta of tau depend on tau for white phase noise? Does it go as $\frac{1}{\tau^{3/2}}$?

Prof. Percival: It would be the slope in the spectrum. It would be $\tau^{-1+\alpha}$ so whatever the slope is in the spectrum, it directly translates into the band pass variance.

Mr. Greenhall: Yes, I was just thinking that for white phase and for flicker of phase, Beta of tau resembles more the modified Allan variance.

Prof. Percival: It could. It could in some sense alleviate the need to use the modified sigma tau in some cases and the ordinary sigma tau in others because the bandpass variance is convergent for all alphas and has a unique signature for all alphas. There is no mapping of various power laws onto each other.

Anthony Hewitt, General Electric: What is effect of outliers in the data in this kind of analysis?

Prof. Percival: Again, there are some recent results which would help quite a bit. There is a very nice article by Chase, Thompson and Andrews in the JGR last January on robust estimation of the spectrum. As long as your data is not contaminated too badly, say at the level of 10% to 20% outliers, you can get good estimates of the spectrum in the presence of additive outliers. The tradeoff is that you have to use a blocking scheme or Welch type estimator, so that there is a loss of resolution. The loss in resolution is probably not too important for typical PTTI data, so those methods might be very attractive