Total Variance, an Estimator of Long-Term Frequency Stability

Charles A. Greenhall, *Member, IEEE*, Dave A. Howe, and Donald B. Percival

***Abstract*—Total variance is a statistical tool developed for improved estimates of frequency stability at averaging times up to one-half the test duration. As a descriptive statistic, total variance performs an exact decomposition of the sample variance of the frequency residuals into components associated with increasing averaging times. As an estimator of Allan variance, total variance has greater equivalent degrees of freedom and lesser mean square error than the standard unbiased estimator has.**

I. Introduction

ALMOST by definition, there can never be enough data when making long-term stability measurements of clocks and frequency standards. Having collected data during a time period T, we have to accept a tradeoff between averaging time τ and confidence in the estimate $\hat{\sigma}_y(\tau, T)$ of Allan deviation $\sigma_y(\tau)$. To improve this tradeoff, Howe *et al.* [1] introduced the practice of incorporating all of the available overlapping samples of the increment of τ-average frequency into the estimate. Of course, for the largest averaging time ($\tau = T/2$), there is only one such sample, the change in average frequency from the first half of the run to the second. The resulting estimate $\hat{\sigma}_y(T/2, T)$ often appears to be unrealistically low; an example can be seen in Fig. 1, the results of a test run of a pair of hydrogen masers.

Two reasons for the droop at the right end can be given. First, if the differences of the frequency residuals are modeled as Gaussian random variables with mean zero, implying no overall linear frequency drift, then $\hat{\sigma}_y^2(T/2, T)$ is proportional to a chi-squared random variable χ_1^2 with one degree of freedom. The distribution of such a random variable is heavily skewed toward values lower than its mean value $\sigma_y^2(T/2)$. Fig. 2 shows the probability density of the random variable $Q = \log_{10}[\hat{\sigma}_y(T/2, T)/\sigma_y(T/2)]$. The probability that $Q < 0$ is 0.68, more than twice the probability that $Q > 0$, and the left tail is much heavier than the right tail. Second, to prevent frequency drift from masking the long-term fluctuations, it is common practice to remove an estimate of overall linear drift from the data; in this case, $\hat{\sigma}_y^2(T/2, T)$ is likely to be reduced because drift removal tends to match the earlier and later frequencies. After drift removal, $\hat{\sigma}_y^2(T/2, T)$ still has one degree of freedom; so it is subject to both effects.

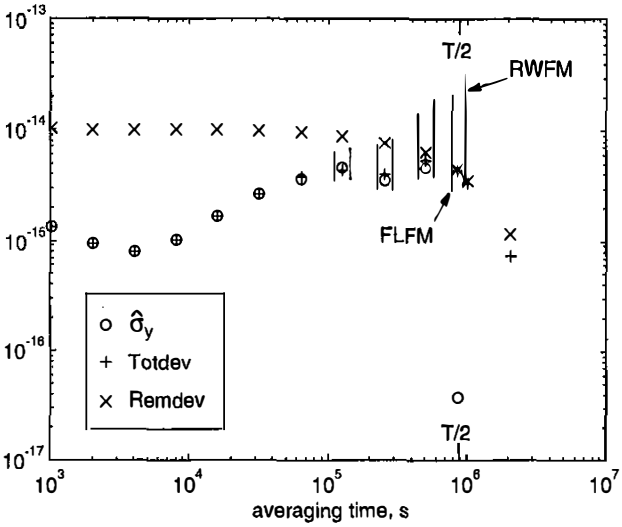

Fig. 1. Sigma-tau plot of $\hat{\sigma}_y$ (standard estimate of Allan deviation), total deviation, and remainder deviation for a pair of hydrogen masers. The error bars (offset vertical lines) are 90% confidence intervals for Allan deviation based upon total deviation.

In an effort to reduce these effects on the measurement of $\sigma_y^2(\tau)$ for large τ, the notion of total variance was developed over the last few years in a sequence of papers [2]–[5]. Fig. 3 illustrates how the idea of total variance was initially conceived. The top plot shows the frequency sampling function for the estimated Allan variance at $\tau = T/2$. By sampling function we mean a function $h(t)$ by which the frequency residuals $y(t)$ are to be multiplied; then, $\int_0^T h(t)\, y(t)\, dt$ is called the functional associated with the sampling function. (For sampled data, we mean an analogous summation.) Except for a scale factor, the absolute value of the functional associated with the top sampling function is just $\hat{\sigma}_y(T/2, T)$. Because this sampling function is odd about $T/2$, its functional rejects the even part of $y(t)$. If by chance or design (from the two effects discussed previously) it should happen that $y(t)$ tends to be even about $T/2$, then the functional could produce a value much smaller than a practical notion of the size of the

Manuscript received August 17, 1998; accepted March 5, 1999.

C. A. Greenhall is with the Jet Propulsion Laboratory (CIT), Pasadena, CA 91109 (e-mail: charles.greenhall@jpl.nasa.gov). The work of this author was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

D. A. Howe is with the Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80303.

D. B. Percival is with the Applied Physics Laboratory, University of Washington, Seattle, WA, and MathSoft, Inc., Seattle, WA 98195. Supported by an ONR ARL Grant (APL, UW) and an NSF Phase II SBIR contract (MathSoft).

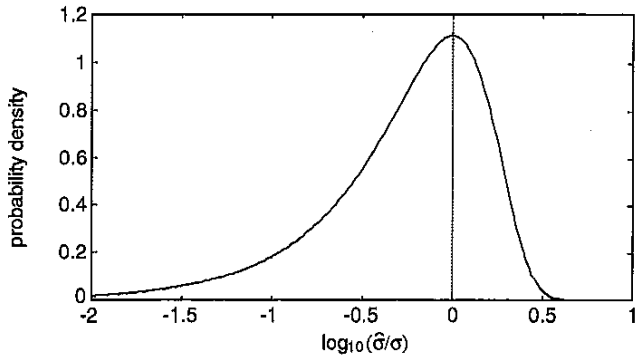

Fig. 2. Probability density function of the logarithm (base 10) of the standard estimator $\hat{\sigma}_y(T/2,T)$ [normalized by $\sigma_y(T/2)$], whose square has one degree of freedom.

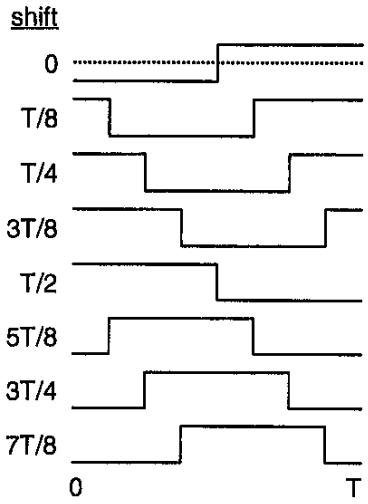

Fig. 3. Illustrating the first concept of total variance: cyclic shifts of the frequency sampling function (top plot) for $\hat{\sigma}_y^2(T/2,T)$.

long-term frequency variations. In this situation, it makes sense to apply also the functional associated with the even sampling function labeled by $T/4$, which is obtained by a cyclic shift (modulo T) of the top sampling function. Here, the odd part of $y(t)$ is rejected. One might suppose, then, that the square root of the sum of the squares of these two functionals is a better measure of long-term stability than either functional by itself.

Let us take this idea further. Having admitted one T-cyclic shift of the top sampling function, we might as well admit all of the shifts, seven of which are shown in Fig. 3. The goal is to improve on the fully overlapped unbiased estimator $\hat{\sigma}_y^2(\tau,T)$ [1], henceforth called the standard estimator, for averaging times $\tau \leq T/2$. Consider the sampling function $h_\tau(t) = -1$ for $0 \leq t \leq \tau$, 1 for $\tau < t \leq 2\tau$. Its support is the interval $0 \leq t \leq 2\tau$. The standard estimator is the scaled mean square output of the linear functionals associated with all of the available noncyclic time shifts of $h_\tau(t)$ whose supports fit within the data interval $0 \leq t \leq T$. [See (5) for a formula that applies to discrete-time data.] The initial version of total variance for τ is the scaled mean square output of the linear functionals associated with all of the possible T-cyclic shifts of $h_\tau(t)$. (For $\tau = T/2$, one-half of the sampling functions are redundant because they are the negatives of the other half.)

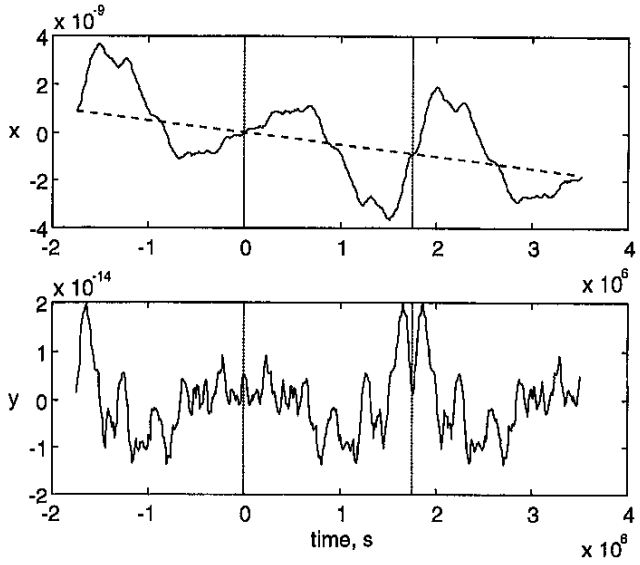

Fig. 4. Extension by reflection of time and normalized frequency residuals for computation of total variance. The original data are between the dotted lines.

This version of total variance enjoyed some success as an estimator of Allan variance with reduced variability and sensitivity to drift removal [2], [3], although it seemed to have a problem of increased variability for data dominated by random-walk FM. The same estimator can be obtained by fixing the sampling function $h_\tau(t)$ and shifting the data cyclically modulo T, or, what is the same, by applying $h_\tau(t)$ as a linear time-invariant filter to an input obtained by extending the original data $y(t)$ periodically with period T. If $y(t)$, $0 \leq t \leq T$, is viewed as a finite piece of an ergodic process, then its T-periodic extension can sometimes be regarded as a substitute for lack of knowledge of the data outside $[0,T]$ [6]. On the other hand, a piece of a random walk, if continued periodically, has a large random discontinuity at the data interfaces 0 and T, untypical of the process as a whole; its effect on the $h_\tau(t)$ functionals cannot be neglected, even for small τ. This problem of mismatched endpoints was solved by the technique of reflecting the data about both endpoints, resulting in a virtual dataset $y^{\#}(t)$ that can be extended to a $2T$-periodic sequence consisting of alternating forward and backward copies of $y(t)$. The lower plot of Fig. 4 shows a portion of $y^{\#}(t)$ about three times as long as the original $y(t)$ in the middle section. The current version of total variance is defined as the scaled mean square output of the $h_\tau(t)$ filter acting on this new sequence.

The intent of the present paper is to give a precise definition of total variance and an account of some of its properties. We abbreviate total variance for τ and T as Totvar(τ,T) or Totvar(τ) (pronounced tōt′-vär). The

square root of total variance is called total deviation (Totdev). The results given here fall into two categories.

A. Total Variance as a Descriptive Statistic

Both Totvar (τ, T) and $\hat{\sigma}_y^2(\tau, T)$ are statistics; that is, they are functions only of the data at hand. By a descriptive statistic, we mean a statistic that has something valuable to say about these data, regardless of any stochastic model that might be fitted or any assumptions about how the data might have evolved outside the interval of observation. Simple examples are sample mean and sample variance. In Section III, we show that total variance can be used to carry out an analysis of variance, an exact decomposition of the sample variance s_y^2 of the frequency residuals $y_n = y(n\tau_0)$, where τ_0 is the sample period. In particular, Totvar $\left(2^j\tau_0\right)$ (when rescaled) can be regarded as the portion of s_y^2 to be associated with the averaging time $2^j\tau_0$ or, equivalently, the octave frequency band $2^{-j-2}/\tau_0 < \nu < 2^{-j-1}/\tau_0$. Thus, after evaluating Totvar (τ) for $\tau = \tau_0, 2\tau_0, \ldots, 2^j\tau_0$, one can tell how much of the sample variance is yet unaccounted for, and one can associate the low frequency band $0 < \nu < 2^{-j-2}/\tau_0$ with the remainder. Analysis of sample variance is a central theme in statistics; an exact decomposition is highly desirable because it accounts for all of the observed variance in the data. The periodogram does such, as do most spectrum estimators and other decompositions, but the standard estimator $\hat{\sigma}_y^2(\tau, T)$ of Allan variance does not [7].

B. Total Variance as an Estimator of Allan Variance

Presented below are results for the bias and variance of Totvar (τ, T), regarded as an estimator of $\sigma_y^2(\tau)$, in the presence of three power-law FM noises: white FM, flicker FM, and random-walk FM. For white FM, we find that Totvar (τ, T) is unbiased for $0 < \tau \leq T$. At $\tau = T/2$, the bias is -24% for flicker FM and -37.5% for random-walk FM. For $0 < \tau \leq T/2$, the normalized bias has the simple form $-a\tau/T$, where a is a constant that depends on the noise type. (These biases apply to σ^2, not σ.)

Variance results are given in terms of the equivalent degrees of freedom [edf; see (24) below]. For $0 < \tau \leq T/2$, the edf of Totvar (τ, T) is always greater than that of $\hat{\sigma}_y^2(\tau, T)$; the edf of Totvar $(T/2, T)$ is 3 for white FM, 2.1 for flicker FM, and 1.5 for random-walk FM; the edf of $\hat{\sigma}_y^2(T/2, T)$ is 1. Moreover, the edf of Totvar (τ, T), $0 < \tau \leq T/2$, can be well approximated by first-degree polynomials in T/τ for each noise type. The mean square error of Totvar (τ, T) is less than that of $\hat{\sigma}_y^2(\tau, T)$, even though the former is biased and the latter is not. Confidence intervals for $\sigma_y^2(\tau)$ based on a chi-squared assumption for Totvar (τ, T) can easily be constructed; these will be tighter than those based on $\hat{\sigma}_y^2(\tau, T)$, and there is evidence that such confidence intervals are conservative.

In summary, total variance is presented as a tool for squeezing a modest amount of extra information about long-term stability from a set of clock residuals, information that is often obscured by the standard Allan variance estimator for τ at or near $T/2$. Analyzing frequency stability accurately in the long term has been problematic even for experienced users. The properties of total variance presented here suggest that it uses the available data more efficiently than the standard estimator for long-term characterizations. Confident of these properties, the authors expect to see wider usage of this tool.

II. Definition of Total Variance

The purpose of this section is to give a precise definition of Totvar (τ, T) for an N_x-point time-residual sequence with sample period τ_0. In the following description, the indices m, n, and N_x are related to time by $\tau = m\tau_0$, $t = t_0 + n\tau_0$, and $T = (N_x - 1)\tau_0$, where t_0 is the time origin and, without loss, may be made equal to 0.

We start with time-residual data $x_1, \ldots, x_{N_x}$, with normalized frequency residuals $y_n = (x_{n+1} - x_n)/\tau_0$, $1 \leq n \leq N_y = N_x - 1$. Extend the sequence y_n to a new, longer virtual sequence $y_n^\#$ by reflection as follows: for $1 \leq n \leq N_y$, let $y_n^\# = y_n$; for $1 \leq l \leq N_y - 1$ let

$$y_{1-l}^\# = y_l, \quad y_{N_y+l}^\# = y_{N_y+1-l}. \tag{1}$$

An equivalent operation can be performed on the original time-residual sequence x_n to produce an extended virtual sequence $x_n^\#$ as follows: for $1 \leq n \leq N_x$, let $x_n^\# = x_n$; for $1 \leq l \leq N_x - 2$, let

$$x_{1-l}^\# = 2x_1 - x_{1+l}, \quad x_{N_x+l}^\# = 2x_{N_x} - x_{N_x-l}. \tag{2}$$

This operation, illustrated in Fig. 4 by the data used for Fig. 1, is called extension by reflection about both endpoints. The result of this extension is a virtual data sequence $x_n^\#$, $3 - N_x \leq n \leq 2N_x - 2$, having length $3N_x - 4$ and satisfying $y_n^\# = \left(x_{n+1}^\# - x_n^\#\right)/\tau_0$, $3 - N_x \leq n \leq 2N_x - 3$.

We now define

$$\text{Totvar}\,(m, N_x, \tau_0) = \frac{1}{2(m\tau_0)^2(N_x - 2)} \times \sum_{n=2}^{N_x-1} \left(x_{n-m}^\# - 2x_n^\# + x_{n+m}^\#\right)^2, \tag{3}$$

for $1 \leq m \leq N_x - 1$. Note that τ is allowed to go to $(N_x - 1)\tau_0$ instead of the usual limit of $\lfloor (N_x - 1)/2 \rfloor \tau_0$. Because of the symmetry of the extended data, the number of summands in (3) does not depend on m. Total variance can also be represented in terms of extended normalized frequency averages by

$$\text{Totvar}\,(m, N_y + 1, \tau_0) = \frac{1}{2(N_y - 1)} \times \sum_{n=2}^{N_y} \left[\bar{y}_n^\#(m) - \bar{y}_{n-m}^\#(m)\right]^2 \tag{4}$$

where $\bar{y}_n^{\#}(m) = \left(x_{n+m}^{\#} - x_n^{\#}\right) / (m\tau_0)$.

The notations Totvar (τ, T) and Totvar (τ) are to be regarded as abbreviations for Totvar (m, N_x, τ_0).

A. Remarks

- For comparison, the standard Allan variance estimator, which we have been abbreviating as $\hat{\sigma}_y^2(\tau, T)$, is actually given by

$$\hat{\sigma}_y^2(m, N_x, \tau_0) = \frac{1}{2(N_x - 2m)} \times \sum_{n=1}^{N_x - 2m} \left[\bar{y}_{n+m}(m) - \bar{y}_n(m)\right]^2 \quad (5)$$

  where $1 \le m < N_x/2$, $\bar{y}_n(m) = (x_{n+m} - x_n) / (m\tau_0)$.
- Total variance, like Allan variance and its conventional estimators, is invariant to an overall shift in phase and frequency; that is, if a first-degree polynomial $c_0 + c_1 n$ is added to the original sequence x_n, then total variance does not change.
- It is possible to program total variance without creating an extended data array in memory [5, eq. (9)].
- To simplify the scaling of total variance relative to the decomposition of sample variance that it generates [see (19)–(21)], one might wish to use $N_x - 1$ in place of the denominator $N_x - 2$ in (3), and N_y in place of the denominator $N_y - 1$ in (4).
- In Section III D, the definition of total variance is extended to arbitrarily large m.

III. Analysis of Variances

A. FIR Filters

To prepare for this section, we review some notation regarding finite-impulse-response (FIR) filters acting on discrete-time signals x_n with a sample period τ_0, where n ranges over all integers. A FIR filter is an operator of form

$$Fx_n = \sum_{l=-\infty}^{\infty} f_l x_{n-l} \quad (6)$$

where f_n, the impulse response of F, is nonzero for only a finite set of indices n. For causal filters ($f_n = 0$ for $n < 0$), we can write the impulse response as a list $[f_0, f_1, f_2, \ldots, f_L]$. The transfer function of F is defined by

$$F(\nu) = \sum_{n=-\infty}^{\infty} f_n e^{-i2\pi\nu n\tau_0} \quad (7)$$

with F doing double duty as an operator on sequences in (6) and a function of frequency ν in (7). The squared frequency response is $|F(\nu)|^2$. Operator composition of filters corresponds to convolution of impulse responses and multiplication of transfer functions.

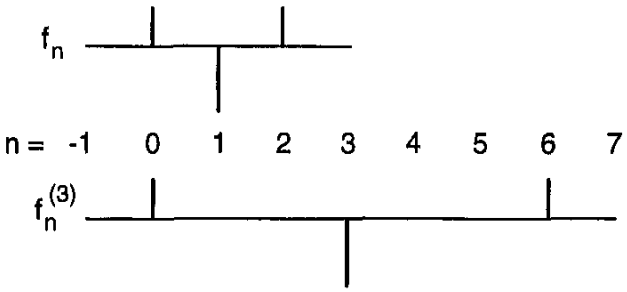

Fig. 5. The effect of 3-upsampling on an impulse response.

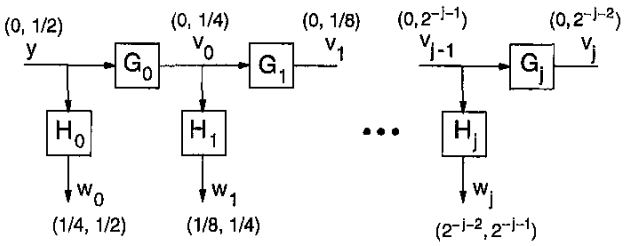

Fig. 6. Multiresolution scheme that leads to variance decompositions by Allan variance and total variance. The frequency bands associated with the signals are shown for unit sample period.

If r is a positive integer, the r-upsampled version of F is defined as the FIR filter $F^{(r)}$ with impulse response $f_n^{(r)} = f_{n/r}$, if n is a multiple of r, and $f_n^{(r)} = 0$ otherwise. In other words, $r-1$ zeros are inserted between successive entries of the original impulse response. (See Fig. 5 for an example with $r = 3$.) Letting $z = e^{-i2\pi\nu\tau_0}$, we see that the upsampled filter has transfer function $F^{(r)}(\nu) = \sum_l f_{rl}^{(r)} z^{rl} = \sum_l f_l z^{rl} = F(r\nu)$.

B. Multiresolution Scheme

The variance decomposition properties of Allan variance and total variance can be derived from an overlapped Haar wavelet transform [8]. The scheme consists of a ladder of FIR filters (Fig. 6), which, acting on an input sequence y_n with sample period τ_0, decomposes the original frequency range $0 < \nu < 2^{-1}/\tau_0$ into successively lower octave bands; each stage leaves a smoothed version of the input for further analysis. The ladder is built from two simple filters: a lowpass filter G_0 with impulse response $\frac{1}{2}[1, 1]$, and a highpass filter H_0 with impulse response $\frac{1}{2}[1, -1]$. The corresponding transfer functions

$$G_0(\nu) = \frac{1}{2}\left(1 + e^{-i2\pi\nu\tau_0}\right) = e^{-i\pi\nu\tau_0}\cos(\pi\nu\tau_0)$$

$$H_0(\nu) = \frac{1}{2}\left(1 - e^{-i2\pi\nu\tau_0}\right) = ie^{-i\pi\nu\tau_0}\sin(\pi\nu\tau_0)$$

satisfy

$$|G_0(\nu)|^2 + |H_0(\nu)|^2 = 1. \quad (8)$$

Write $G_j = G_0^{(2^j)}$, $H_j = H_0^{(2^j)}$, the 2^j-upsampled ver-

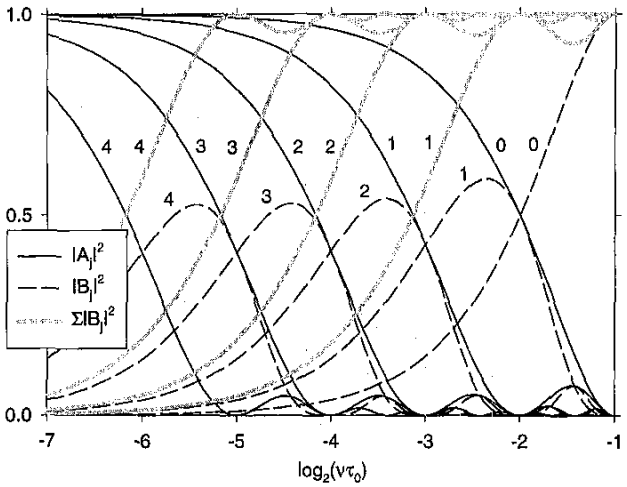

Fig. 7. Squared frequency responses of the multiresolution filters. The j indices are lettered alongside the curves. The label $\sum |B_j|^2$ means $|B_0|^2 + \cdots + |B_j|^2$.

sions of G_0 and H_0. These filters are applied to y_n according to the scheme shown in Fig. 6. Its jth stage has output sequences $v_{j,n} = A_j y_n$ and $w_{j,n} = B_j y_n$, where $A_0 = G_0$, $B_0 = H_0$, and

$$A_j = G_j G_{j-1} \cdots G_0, \quad B_j = H_j G_{j-1} \cdots G_0 = H_j A_{j-1} \tag{9}$$

for $j \geq 1$. One can show by induction that A_j is a moving-average filter with impulse response $2^{-j-1}[1, \ldots, 1]$ (2^{j+1} ones), a lowpass filter for the band $0 < \nu < 2^{-j-2}/\tau_0$. Then, by (9), B_j has impulse response $2^{-j-1}[1, \ldots, 1, -1, \ldots, -1]$ (2^j ones, then 2^j minus-ones), which is $2^{-1/2}$ times the filter associated with $\sigma_y^2(2^j\tau_0)$. This filter is an approximate bandpass filter for the octave band $2^{-j-2}/\tau_0 < \nu < 2^{-j-1}/\tau_0$ [9]: the low frequency skirt falls off smoothly at 6 dB per octave; the high frequency skirt is a sequence of sidelobes alternating with deep nulls. The squared frequency responses of A_j and B_j,

$$|A_j(\nu)|^2 = \frac{\sin^2(2^{j+1}\pi\nu\tau_0)}{4^{j+1}\sin^2(\pi\nu\tau_0)}, \quad |B_j(\nu)|^2 = \frac{\sin^4(2^j\pi\nu\tau_0)}{4^j\sin^2(\pi\nu\tau_0)},$$

are plotted in Fig. 7 against $\log_2(\nu\tau_0)$ for $0 \leq j \leq 4$.

As we have seen, the approximate passbands of the filters $B_0, \ldots, B_j, A_j$ partition the original frequency domain $0 < \nu < 2^{-1}/\tau_0$ into octaves, shown in Fig. 7 as the intervals between the x-axis tick marks. All of the variance decompositions discussed subsequently follow from a counterpart of this statement for the squared frequency responses, which satisfy the frequency-domain decomposition equation

$$\sum_{j=0}^{J} |B_j(\nu)|^2 + |A_J(\nu)|^2 = 1. \tag{10}$$

Fig. 7 also shows $\sum_{j=0}^{J} |B_j(\nu)|^2$ for $0 \leq J \leq 4$.

Equation (10) can be proved by induction on J from (8) and (9) (or from the identity $\sin^4 x = \sin^2 x - \frac{1}{4}\sin^2 2x$). If $\nu\tau_0$ is not an integer, then $|A_j(\nu)|^2 \to 0$ as $j \to \infty$, and it follows from (10) that

$$\sum_{j=0}^{\infty} |B_j(\nu)|^2 = 1 \tag{11}$$

$$\sum_{j=J+1}^{\infty} |B_j(\nu)|^2 = |A_J(\nu)|^2. \tag{12}$$

Equation (11) says that the squared frequency responses associated with Allan variance for $\tau = 2^j\tau_0$ sum to 2, except at zero frequency and its aliases.

C. Ensemble Variance

Before we derive the sample variance decomposition property of total variance, it is useful to understand how an analogous property of Allan variance follows from the frequency-domain decompositions. Let y_n be a stationary random process with variance $\operatorname{var} y_n$ and one-sided spectral density $S_y(\nu)$. Then, the stationary processes $v_{j,n}$ and $w_{j,n}$ have spectra $|A_j(\nu)|^2 S_y(\nu)$ and $|B_j(\nu)|^2 S_y(\nu)$, respectively, and $\operatorname{var} w_{j,n} = \frac{1}{2}\sigma_y^2(2^j\tau_0)$. Accordingly, if we multiply (10)–(12) by $S_y(\nu)$ and integrate over ν from 0 to $2^{-1}/\tau_0$, we obtain

$$\operatorname{var} y_n = \frac{1}{2}\sum_{j=0}^{J} \sigma_y^2(2^j\tau_0) + \operatorname{var} v_{J,n}$$

$$= \frac{1}{2}\sum_{j=0}^{\infty} \sigma_y^2(2^j\tau_0) \tag{13}$$

$$\operatorname{var} v_{J,n} = \frac{1}{2}\sum_{j=J+1}^{\infty} \sigma_y^2(2^j\tau_0). \tag{14}$$

If y_n has stationary first increments but is not stationary, like flicker FM and random-walk FM, then $w_{j,n}$ is stationary, $v_{j,n}$ is not stationary, and the frequency-domain integrals giving $\operatorname{var} y_n$ and $\operatorname{var} v_{J,n}$ are infinite, as are the infinite series in (13) and (14).

D. Sample Variance

From the random-process setting, we return to the consideration of a finite data sequence $y_1, \ldots, y_N$ with sample period τ_0. Before invoking the extension procedure of total variance, let us consider temporarily a simpler periodic extension with period N; that is, we agree that $y_{n+N} = y_n$ for all integers n. The sample mean m_y and sample variance s_y^2 of y_n are conveniently expressed in terms of its discrete Fourier transform (DFT), given by

$$Y_k = \sum_{n=1}^{N} y_n e^{-i2\pi kn/N}$$

and also indicated by the notation $y_n \leftrightarrow Y_k$. We have $m_y = Y_0/N$, and

$$s_y^2 = \frac{1}{N}\sum_{n=1}^{N} y_n^2 - m_y^2 = \frac{1}{N^2}\sum_{k=1}^{N-1} |Y_k|^2 \tag{15}$$

by Parseval's theorem in the DFT setting.

Let F be a FIR filter with impulse response f_n and transfer function $F(\nu)$. Define the periodized impulse response $f_n(N)$ to be the sum of f_l over all l such that $\mathrm{mod}(l, N) = n$. Two facts about the periodic sequence $f_n(N)$ must now be set down. First, if y_n is N periodic, then so is Fy_n, and

$$Fy_n = \sum_{l=-\infty}^{\infty} f_l y_{n-l} = \sum_{l=1}^{N} f_l(N)\, y_{n-l}.$$

The second summation, which expresses a cyclic convolution, can be taken over any period. Second, we have $f_n(N) \leftrightarrow F(\nu_k)$, where $\nu_k = k/(N\tau_0)$. Because the DFT maps cyclic convolutions to products, we conclude that $Fy_n \leftrightarrow F(\nu_k) Y_k$.

Let the input to our multiresolution scheme be an N-periodic sequence y_n. According to the previous paragraph, all of the output sequences are periodic, and $v_{j,n} = A_j y_n \leftrightarrow A_j(\nu_k) Y_k$, $w_{j,n} = B_j y_n \leftrightarrow B_j(\nu_k) Y_k$. By (15),

$$\begin{aligned} s_{v_j}^2 &= \frac{1}{N^2}\sum_{k=1}^{N-1} |A_j(\nu_k)|^2 |Y_k|^2, \\ s_{w_j}^2 &= \frac{1}{N^2}\sum_{k=1}^{N-1} |B_j(\nu_k)|^2 |Y_k|^2. \end{aligned} \tag{16}$$

Combining (15) and (16) with the frequency-domain partitions (10)–(12) yields the analogs of (13)–(14) for sample variances, namely,

$$s_y^2 = \sum_{j=0}^{J} s_{w_j}^2 + s_{v_J}^2 = \sum_{j=0}^{\infty} s_{w_j}^2, \tag{17}$$

$$s_{v_J}^2 = \sum_{j=J+1}^{\infty} s_{w_j}^2. \tag{18}$$

Observe also that $m_{v_j} = m_y$, $m_{w_j} = 0$ because $A_j(0) = 1$, $B_j(0) = 0$.

In connection with the definition of total variance (Section II), we now let the role of $y_1, \ldots, y_N$ in the previous discussion be taken by $y_1, \ldots, y_{N_y}, y_{N_y}, \ldots, y_1$. Let $y_n^{\#}$ be the extension of this finite sequence to a $2N_y$-periodic sequence; then $y_n^{\#}$ agrees with the definition given in Section II. We feed $y_n^{\#}$ to the multiresolution ladder and proceed to interpret the meaning of (17)–(18). Because of symmetry, the terms of the sample variance sums occur in pairs; consequently, we have $s_{y^{\#}}^2 = s_y^2$ (the sample variance of $y_1, \ldots, y_{N_y}$), and

$$s_{w_j}^2 = \frac{N_y - 1}{2N_y}\,\mathrm{Totvar}\left(2^j\tau_0\right). \tag{19}$$

Although total variance was previously defined only for $m < N_y$, (4) retains meaning for all m if $y_n^{\#}$ is extended $2N_y$ periodically as far as needed.

To interpret $s_{v_j}^2$, it is convenient to define a remainder variance, [Remvar $(m\tau_0)$], such that

$$s_y^2 = \frac{N_y - 1}{2N_y}\,\mathrm{Remvar}\,(\tau_0) \tag{20}$$

$$s_{v_j}^2 = \frac{N_y - 1}{2N_y}\,\mathrm{Remvar}\left(2^{j+1}\tau_0\right). \tag{21}$$

The square root of remainder variance is called remainder deviation (Remdev). In this setting, the variance decompositions (17)–(18) become

$$\begin{aligned} \mathrm{Remvar}\,(\tau_0) &= \sum_{j=0}^{J} \mathrm{Totvar}\left(2^j\tau_0\right) + \mathrm{Remvar}\left(2^{J+1}\tau_0\right) \\ &= \sum_{j=0}^{\infty} \mathrm{Totvar}\left(2^j\tau_0\right), \end{aligned} \tag{22}$$

$$\mathrm{Remvar}\left(2^{J+1}\tau_0\right) = \sum_{j=J+1}^{\infty} \mathrm{Totvar}\left(2^j\tau_0\right). \tag{23}$$

In other words, the sum of all of the Totvar $\left(2^j\tau_0\right)$ equals the sample variance of y_n scaled by $2N_y/(N_y - 1)$, or $2s_y^2$ for practical purposes; Remvar $\left(2^{J+1}\tau_0\right)$ indicates how much of this rescaled sample variance has not yet been accounted for by Totvar $\left(2^j\tau_0\right)$, $0 \le j \le J$.

1. Remarks:

- Total variance is the first modern estimator of Allan variance to mimic its ensemble variance decomposition properties (13)–(14); moreover, the sample variance decompositions (22)–(23) apply to any finite data sequence.
- Because higher order Daubechies wavelet filters [10] also satisfy (8), the previous development extends easily to higher order wavelet variances. (Allan variance is essentially twice the Haar wavelet variance.) These higher order wavelet variances are suitable substitutes for some of the variations on Allan variance that have been proposed and studied in the literature (modified Allan variance, for example). For details, see [11].
- Even though one can evaluate Totvar (τ, T) for arbitrarily large τ without taking more data, its value for $\tau > T$ ought not to be regarded as an estimate of $\sigma_y^2(\tau)$; in particular, if $N_y = 2^K$ then Remvar $\left(2^j\tau_0\right)$ and Totvar $\left(2^j\tau_0\right)$ vanish for $j \ge K + 1$.
- See [7] for a discussion of analysis of sample variance by the nonoverlapped estimator of Allan variance.

2. Example: Fig. 1 shows $\hat{\sigma}_y(\tau, T)$, Totdev (τ, T), and Remdev (τ, T) for the hydrogen-maser data shown in Fig. 4, with $N_x = 1727$, $\tau_0 = 1024.1$ s, and $T = 1.77 \times 10^6$ s. The averaging times include $2^j\tau_0$, $0 \le j \le 11$, and $T/2 = 8.84 \times 10^5$ s. Relative frequency drift, -1.81×10^{-15} per day, was estimated from the data by the 4-point w

method [12] and removed. As a result, $\hat{\sigma}_y(T/2, T)$ becomes severely depressed, a common consequence of drift removal. (It would have been identically zero if the 3-point x drift estimate [13] had been used.) On the other hand, Totdev (τ) shows no depression until τ exceeds $T/2$. The flatness of remainder deviation at the lower τ indicates that the first several values of total variance contribute little to Remvar (τ_0), which is approximately twice the sample variance of the frequency data. At $\tau = 2^{10}\tau_0$, remainder deviation and total deviation are almost equal; by (23), this indicates that the values of Totvar $\left(2^j\tau_0\right)$ for $j \geq 11$ also contribute little to Remvar (τ_0).

The error bars, which are confidence intervals for $\sigma_y(\tau)$ based upon Totdev(τ, T), are discussed in Section IV B.

IV. Total Variance as an Estimator of Allan Variance

Although total variance can stand on its own as a descriptive statistic that performs an analysis of variance on a data sequence, its usefulness for time and frequency measurement is based mainly on its statistical properties as an estimator of Allan variance under assumptions about the underlying random noise processes. Because we are interested mostly in long-term frequency-stability statistics, and for mathematical convenience, we treat only the power-law FM noise processes that are likely to dominate long-term measurements: white FM, flicker FM, and random-walk FM. With these assumptions, the properties of Totvar $(m, N_y + 1, \tau_0)$ depend mainly on the ratio $m/N_y = \tau/T$ for large m and N_y. It is convenient, then, to approximate Totvar (τ, T) by a continuous-time analog in which sums involving x_n are replaced by integrals involving $x(t)$, a power-law process with spectral density proportional to $\nu^{\alpha-2}$ $(\alpha = 0, -1, -2)$ for $0 < \nu < \infty$ (no high-frequency cutoff) [5]. The theoretical computations assume that $x(t)$ is a process with stationary Gaussian zero-mean second differences.

A. Bias and Variance

Although total variance is most conveniently expressed as a function of the extended data $x_n^\#$ or $y_n^\#$, each term of (3) can also be thought of as the square of a linear functional of the original data x_n or y_n. These functionals, although complicated by the fold-over implicit in the extension by reflection, are still second-order functionals of the x_n, that is, they are invariant to time and frequency shifts. (See [5] for formulas and pictures of the sampling functions.) This property makes it possible to compute the mean and variance of total variance by means of manipulations on the generalized autocovariances of the three FM noise processes [14].

The mean $E[\text{Totvar}(\tau, T)]$ is compared to $\sigma_y^2(\tau)$ to give normalized bias (nbias); the variance is communicated through the eqluivalent degrees of freedom (edf), defined

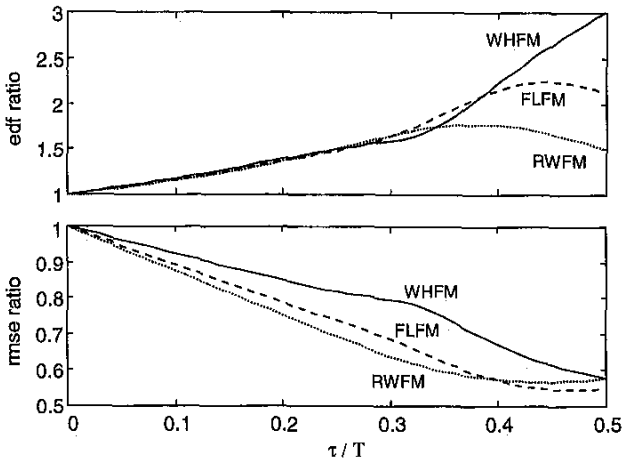

Fig. 8. Ratios of edf and root mean square error for total variance over those of the standard estimator of Allan variance.

for a random variable V by

$$\text{edf}\, V = \frac{2(EV)^2}{\text{var}\, V}. \tag{24}$$

The results can be expressed by the formulas:

$$\text{nbias}(\tau) := \frac{E[\text{Totvar}(\tau, T)]}{\sigma_y^2(\tau)} - 1 = -a\frac{\tau}{T}, \tag{25}$$

$$\text{edf}(\tau) := \text{edf}[\text{Totvar}(\tau, T)] \approx b\frac{T}{\tau} - c, \quad 0 < \tau \leq \frac{T}{2} \tag{26}$$

where a, b, and c are given in Table I. The values of nbias and edf for the important longest term case $\tau = T/2$ are also tabulated. The edf formula (26) is empirical, having an observed error below 1.2% of numerically computed exact values; the tabulated values of edf $(T/2)$ are the exact ones. As is obvious from their form, a and b were derived from theory; in particular, $b = \lim_{T/\tau \to \infty}(\tau/T)\,\text{edf}\left[\hat{\sigma}_y^2(\tau, T)\right]$ [15]. The only coefficient that had to be chosen empirically was c. These results were checked by simulations of Totvar (m, N_x, τ_0), with $N_x = 101$. The simplicity, accuracy, and range of applicability of (26) are striking in view of existing approximations for edf $\left[\hat{\sigma}_y^2(\tau, T)\right]$ [1], [15], [16]; although total variance is more complicated than the standard estimator, some of its statistical properties are simpler.

Fig. 8 compares Totvar (τ, T) to the standard unbiased Allan variance estimator $\hat{\sigma}_y^2(\tau, T)$ in two different ways: the upper plot shows the ratio of the edf of the two estimators for the three FM noises; the lower plot shows the ratio of their root mean square errors $(\sqrt{\text{bias}^2 + \text{variance}})$ as estimators of $\sigma_y^2(\tau)$.

1. Remarks:

- Because of the continuous-time analog used for the theoretical calculations, (26) should be used only if $\tau \geq 8\tau_0$ for white FM, $\tau \geq 3\tau_0$ for flicker FM.

TABLE I
COEFFICIENTS FOR COMPUTING NBIAS AND EDF OF TOTAL VARIANCE IN THE PRESENCE OF FM NOISES. TABULATED ALSO ARE THE EXACT QUANTITIES FOR $\tau = T/2$.

Noise[1]	a	b	c	nbias($T/2$)	edf($T/2$)
WHFM	0	3/2	0.000	0	3.000
FLFM	$(3\ln 2)^{-1}$	$24(\ln 2)^2\pi^{-2}$	0.222	−0.240	2.097
RWFM	3/4	140/151	0.358	−3/8	1.514

[1] WHFM = white FM, FLFM = flicker FM, RWFM = random walk FM.

- For white FM, Totvar (τ, T) is an unbiased estimator of $\sigma_y^2(\tau)$ for $\tau \leq T$. This fact appeared as an outcome of algebraic manipulations; unfortunately, the authors cannot give a simple reason why it is so. The result that edf $(\tau) = \frac{3}{2}T/\tau$ for white FM, although obtained numerically, seems to be exact for $\tau \leq T/2$ and for $\tau = T$. No calculations of edf were performed for $T/2 < \tau < T$.

B. Confidence Intervals

In the tradition of time and frequency statistics, it is customary to derive confidence intervals for frequency stability by assuming that the probability distribution of a frequency stability estimator V, when scaled appropriately, follows the chi-squared distribution with the same edf as V [1]. Fix τ, and write $V = \text{Totvar}(\tau, T)$, $\sigma^2 = \sigma_y^2(\tau)$. Let $r = 1 + \text{nbias} = E(V)/\sigma^2$, $q = \text{edf}(V)$. Then, the random variable

$$X = \frac{qV}{r\sigma^2} \tag{27}$$

has the same mean and edf, namely q, as a χ_q^2 variable has. Presume for the moment that X has a χ_q^2 distribution. For $0 \leq p_1 < p_2 < 1$, let ξ_1 and ξ_2 be the corresponding levels of this distribution. (A simple approximation algorithm for χ_q^2 levels can be found in [17].) Then, $\xi_1 < X < \xi_2$ with probability $p = p_2 - p_1$. Substituting (27) and rearranging, we have the confidence statement that

$$\frac{qV}{r\xi_2} < \sigma^2 < \frac{qV}{r\xi_1} \tag{28}$$

with probability p. A negative bias ($r < 1$) pushes the confidence interval upward.

The error bars in Fig. 1, shifted horizontally for visibility, are 90% confidence intervals for $\sigma_y(\tau)$ as computed by this method ($p_1 = 0.05$; $p_2 = 0.95$) under the assumption of flicker FM and random-walk FM noise models. Both sets of error bars suggest the hypothesis that random-walk FM is the dominant noise type for $10^5\,\text{s} < \tau < 10^6\,\text{s}$, although a flicker FM hypothesis is not ruled out. A longer test run ($T = 4.23 \times 10^6$ s) of the same pair of standards supports the random-walk hypothesis, with $\hat{\sigma}_y(\tau, T)$ increasing like $\tau^{1/2}$. On the other hand, the longer run has a sharp frequency step of about 4×10^{-14}, untypical of the shorter run (Fig. 4), so that the authors hesitate to declare a successful characterization.

We note that (25)–(26) have not been shown to be accurate when estimated frequency drift is removed, as was done for Fig. 1; the authors have not carried out the required theoretical computations, which depend on the method of drift estimation and are more intricate than before. It seems clear, though, that the effect of drift removal on total variance is less than its effect on conventional $\sigma_y^2(\tau)$ estimators, which tend to be severely depressed for τ near $T/2$ [3].

The χ_q^2 assumption for total variance has been investigated, for $\tau = T/2$ only, by simulation of the three FM noise types [5]. The empirical distributions of X as defined by (27) were observed to have heavier left tails than those of the corresponding χ_q^2 distributions. If this turns out to be true in general, it means that the upper ends of confidence intervals (28) based on the χ_q^2 distribution are pessimistic. For now, use of the χ_q^2 distribution for this purpose seems to be a conservative policy.

REFERENCES

[1] D. A. Howe, D. W. Allan, and J. A. Barnes, "Properties of signal sources and measurement methods," in *Proc. 35th Annu. Symp. Freq. Contr.*, Philadelphia, PA, 1981, pp. 1–47.

[2] D. A. Howe, "An extension of the Allan variance with increased confidence at long term," in *Proc. IEEE Int. Freq. Contr. Symp.*, 1995, San Francisco, CA, pp. 321–329.

[3] D. A. Howe and K. J. Lainson, "Effect of drift on TOTALDEV," in *Proc. IEEE Int. Freq. Contr. Symp.*, Honolulu, HI, 1996, pp. 883–889.

[4] D. A. Howe, "Methods of improving the estimation of long-term frequency variance," in *Proc. Eur. Freq. Time Forum*, Neuchâtel, Switzerland, 1997, pp. 91–99.

[5] D. A. Howe and C. A. Greenhall, "Total variance: a progress report on a new frequency stability characterization," in *Proc. 29th Ann. PTTI Syst. Appl. Mtg.*, Long Beach, CA, 1997, pp. 39–48.

[6] D. A. Howe, "Circular representation of infinitely extended sequences," in *Proc. IEEE Int. Freq. Contr. Symp.*, San Francisco, CA, 1995, pp. 337–345.

[7] D. B. Percival and D. A. Howe, "Total variance as an exact analysis of the sample variance," *Proc. 29th Ann. PTTI Syst. Appl. Mtg.*, Long Beach, CA, 1997, pp. 97–105.

[8] D. B. Percival and H. O. Mofjeld, "Analysis of subtidal coastal sea level fluctuations using wavelets," *J. Amer. Stat. Assoc.*, vol. 92, pp. 868–880, 1997.

[9] D. B. Percival, "Characterization of frequency stability: frequency-domain estimation of stability measures," *Proc. IEEE*, vol. 79, pp. 961–972, 1991.

[10] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM, 1992.

[11] D. A. Howe and D. B. Percival, "Wavelet variance, Allan variance, and leakage," *IEEE Trans. Instrum. Meas.*, vol. 44, pp. 94–97, 1995.

[12] C. A. Greenhall, "A frequency-drift estimator and its removal from modified Allan variance," in *Proc. IEEE Int. Freq. Contr. Symp.*, Orlando, FL, 1997, pp. 428–432.
[13] M. A. Weiss and C. Hackman, "Confidence on the three-point estimator of frequency drift," in *Proc. 24th Ann. PTTI Appl. Planning Mtg.*, McLean, VA, 1992, pp. 451–455.
[14] C. A. Greenhall, "A structure function representation theorem with applications to frequency stability estimation," *IEEE Trans. Instrum. Meas.*, vol. 32, pp. 364–370, 1983.
[15] ——, "Recipes for degrees of freedom of frequency stability estimators," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 994–999, 1991.
[16] K. Yoshimura, "Degrees of freedom of the estimate of the two-sample variance in the continuous-sampling method," *IEEE Trans. Instrum. Meas.*, vol. 38, pp. 1044–1049, 1989.
[17] M. A. Weiss and C. A. Greenhall, "A simple algorithm for approximating confidence on the modified Allan variance and the time variance," in *Proc. 28th Ann. PTTI Syst. Appl. Mtg.*, Reston, VA, 1996, pp. 215–224.

Charles A. Greehall (M'92) was born in New York City in 1939. He graduated from Pomona College in 1961 with a major in physics and received the Ph.D. degree in mathematics at Caltech in 1966 with a thesis on Fourier series. For five years, he was an assistant professor of Mathematics at the University of Southern California. Since 1981, he has worked in the field of clock statistics and measurement at Jet Propulsion Laboratory. He developed the concept of generalized autocovariance as a theoretical and computational tool for studying nonstationary clock-noise processes, thereby deriving results for bias and confidence of frequency-stability and drift estimators. He also takes part in the design and programming of frequency-stability measurement systems. He invented the picket-fence method for collecting zero-dead-time data from interval counters and wrote the signal-processing routines for a real-time CW stability analyzer. He is now working on the design of a dual-mixer stability measurement system based on a time-stamp counter.

David A. Howe has been a physicist with the Time and Frequency Division of NIST, Boulder Laboratories since 1970 having graduated with a B.A. from the University of Colorado that year. From 1970 to 1973, he was with the Dissemination Research Section, where he helped coordinate the first lunar-ranging and spacecraft time-synchronization experiment jointly with Jet Propulsion Laboratory. He joined the staff of the Atomic Standards Section from 1973 to 1984, doing advanced research on cesium and hydrogen maser standards and ruggedized, compact rubidium and ammonia standards. His main activities included spectral estimation using novel digital processing techniques, digital servo design, automated accuracy evaluation of a primary cesium standard, atomic beam analysis, and reduction of oscillator acceleration sensitivity for special applications. He returned to the Dissemination Research Section in 1984 to lead and implement several global high-accuracy satellite-based time-synchronization experiments with other national laboratories. For this contribution, he was awarded the Commerce Department's Gold Medal in 1990 for advancements in time calibrations among standards laboratories who participate in the maintenance of UTC. Since 1994, he has been a statistical analyst for the Time Scale Section, which maintains UTC(NIST) from an ensemble of laboratory atomic frequency standards. He is a member of Sigma Pi Sigma and Phi Beta Kappa societies.

Donald B. Percival received the B.A. degree in astronomy from the University of Pennsylvania, Philadelphia, in 1986; the M.A. degree in mathematical statistics from George Washington University, Washington, DC, in 1976; and a Ph.D. degree in statistics from the University of Washington, Seattle, in 1982.

From 1968 to 1978, he was an astronomer with the Time Service Division, U.S. Naval Observatory, Washington, DC, where he worked on the generation of atomic clock time scales and on the analysis of the frequency stability of high-performance oscillators. Since 1982, he has been a senior mathematician at the Applied Physics Laboratory, University of Washington, with a joint appointment in the Department of Statistics. He is one of the founding members of the National Research Center for Statistics and the Environment at the University of Washington. In 1995, he joined MathSoft, Inc., Seattle, as a part-time research scientist. His current research interests include spectral analysis, wavelets, and use of statistical methodology in the physical sciences. He is coauthor with A. T. Walden of *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge: Cambridge Univ. Press, 1993) and the forthcoming *Wavelet Methods for Time Series Analysis*.

Dr. Percival is a member of the American Statistical Association.