METHODS OF IMPROVING THE ESTIMATION OF LONG-TERM FREQUENCY VARIANCE

D. A. Howe

Time & Frequency Div., National Institute of Standards & Technology
Boulder, Colorado 80303

1. Abstract

I discuss a novel technique for periodically extending by reflection an ordered series of phase-difference measurements $\{x_{k'}\}$ to yield estimation of new frequency and time variances having improved long-term confidence and the same mean as the Allan variance. In addition, I describe a correction to a negative bias in the sample variance based on the actual number of observations in the measurements at long term.

Keywords: *two-sample frequency variance; time variance; Allan variance; TOTAL variance*

2. Background

This paper assumes some familiarity with the Allan variance and five common noise models (white phase modulation, WHPM; flicker of phase modulation, FLPM; white frequency modulation, WHFM; flicker of frequency modulation, FLFM; random walk of frequency modulation, RWFM) present in frequency standards, clocks, and synchronization systems [Refs. 3,21]. It is widely recognized that a trend (given by a slope in log coordinates) in the autocorrelation function and hence its Fourier transform $\hat{S}_{\bar{y}}(f) \propto f^{\alpha}$ has a correspondence to a trend (given by μ) in the two-sample frequency variance [Refs. 3,10,21]. This power-law correspondence between α and μ means that, in general, procedures for better estimation of the power-law type in the f-domain have a parallel in the τ-domain [Ref. 2]. For those in the business of operating extremely precise clocks and oscillators, a principal task is characterizing the frequency stability of the devices relative to other devices and for comparison with the device's own history. Thus pertinent characterizations often refer to changes over relatively long segments of time.

One of the recommended ways to estimate the stability has been the two-sample frequency variance known as the Allan variance and corresponding statistic given by ("< >" denote infinite time average)

$$\sigma^2_{\bar{y}}(\tau) = \frac{1}{2}<(\bar{y}_{i+1}-\bar{y}_i)^2)> ; \sigma^2_{\bar{y}} = \frac{1}{2(M-1)}\sum_{k=1}^{M-1}(\bar{y}_{k+1}-\bar{y}_k)^2 \qquad (1)$$

where if $\{\bar{y}_{k'}\}$, $k' = 1,2,3,\ldots,N-1$ are fractional frequency differences averaged over τ_o derived from N consecutive phase differences $\{x_{k'}\}$("primed" indexes means τ_o-spacing); then $\{y_k\}$, $k = 1,2,3,\ldots,M$ are fractional frequency differences averaged over interval $m\tau_o \equiv \tau$. Hence $\sigma^2_{\bar{y}}$ is implicitly dependent on dimensionless quantity m, a scale parameter which for efficiency can be limited to rational powers of 2, i.e., $2^i=m$, $i=0,1,2,3,\ldots$; (see for example, [Ref. 1]).

A record of residual fractional frequency fluctuations implies that we have administered some form of trend removal (detrending). The removal of a trend such as drift is done electronically (as a voltage steering correction in a clock or oscillator synchronizing servo, for example), computationally (as a regression to an internal estimate of a continuous polynomial), or even mechanically (as a thermally-compensated cavity, for example). The response of the two-sample Allan variance eq (1) at long term is highly variable with forms of detrending and exhibits a negative bias [Refs. 4,8,13,17,23,24]. In cases where the last point is obviously false as judged by the rest of $\sigma_y(\tau)$ plot, it is thrown out. Worse, however, if it seems plausible and is judged somehow as "okay," it is retained and can be used to conform to some expected or suitable long-term behavior.

Realizing this, I introduce a new variance which has the same mean as the Allan variance eq (1) but re-expresses deviates in terms of an averaged combination of "in-phase" and "phase-shifted" sampling functions [Ref. 11] The new variance takes the traditional Allan variance and its single 2 sample function and combines it with an orthogonal 3 sample function proportioned so that its averaging time τ is taken out of the 2τ sampling-interval's middle (not just the first half and second half). The comparison is shown in fig. 1a and 1b. This increases the effective number of independent observations from one to two in the sampling interval thus increasing the number of degrees of freedom which I'll discuss in a moment. Even more crucial however is that sensitivity to removing drift from a segment of data in process (detrending) is eliminated in this variance. The new variance is given by

$$\sigma^2_{\bar{y}*}(\tau) = \frac{1}{4}\left[<\left[\frac{\bar{y}_{k+3}+\bar{y}_{k+2})}{2} - \frac{(\bar{y}_{k+1}+\bar{y}_k)}{2}\right]^2> + <\left[\bar{y}_{k+3} - \frac{(\bar{y}_{k+2}+\bar{y}_{k+1})}{2} + \bar{y}_k\right]^2>\right], \qquad (2)$$

where the average frequency is taken over $\tau/2$ rather than τ as in eq (1). All possible $\tau/\tau_o = m$ time shifts (i.e., meaning deviates in the statistic are maximally overlapped) in addition to two shifts (separated by $\tau/2$

only as in eq (2)) result in smoothed estimates of the new variance at long term. The sampling functions are shown in fig. 1c. The smoothed estimate to eq (2), called TOTALVAR, uses a novel data manipulation which simplifies the procedure and which maximally uses data at hand.

To begin, an estimate to eq (2) has an equivalent description in signal processing as an "in-phase" and "phase-quadrature" discrete functional component separation. The variance component having the 3 sample function is computed by shifting the process's observation window by $\tau/2(\propto m/2)$, for the "phase-quadrature" variance and adding this to the Allan component or "in-phase" variance [Ref. 11]. This yields a combination of sample statistics given by

$$\sigma^2_{\bar{y}'}(m\tau_o) = \frac{1}{4(M-1)}\left[\sum_{k=1}^{M-1}(\nabla\bar{y}_{k+1,m})^2 + \sum_{k=1}^{M-1}(\nabla\bar{y}^{*}_{k+1,m})^2\right], \qquad (3)$$

where, in terms of phase data $\{x_{k'}\}$ spaced by τ_o, hence the "primed-k" index *(t is the running time in seconds)*

$$\bar{y}_{k,m} = \frac{x_{k'-m}(t_{k'}-m\tau_o) - x_{k'}(t_{k'})}{m\tau_o},$$

and

$$\bar{y}^{*}_{k,m} = \frac{x_{k'-m/2}(t_{k'}-\frac{m\tau_o}{2}) - x_{k'+m/2}(t_{k'}+\frac{m\tau_o}{2})}{m\tau_o},$$

$x_1, x_2, \ldots x_N$ is wrapped such that $x_\varsigma = x_{\varsigma \bmod N}$ for $\varsigma < 1$ and $\varsigma > N$, i.e., $x_1 = x_{N+1}$ which reindexes to x_N. It is important to point out that shifting the data and using a wrap (circularizing $\{x_{k'}\}$) is equivalent to changing the sampling function and merely simplifies the form of the sample variance eq (3) corresponding to eq (2) [Ref. 6]. There is no assumed extension of the original observed data; the wrap (or implied extension) is a computational procedure.

The treatment of changes in systematics (i.e., non-stationary second increments of $\{x_{k'}\}$) has been addressed [Refs. 12,13]. This has resulted in a wrap of series $\{x_{k'}\}$ which reflects about the last value of the series rather than simply wraps the data about an axis. This pictorially is shown in fig. 2 and is discussed in **4** below. The final procedure is as follows with notation change $x_{k'} = x(k')$ to distinguish a different (extended) set $\{x_o\}$:

For *x(1),...,x(N),* remove a slope and constant (endmatching procedure) to produce *$x_0(1)$,..., $x_0(N)$,* where *$x_0(1) = x_0(N) = 0$.* Adjoin *$x_0(N+1)$,...,$x_0(2N-1)$,* where *$x_0(N+j) = -x_0(N-j)$, j=1 to N-1.* For *$\tau = m\tau_0$,* TOTALVAR (with square-root as TOTALDEV) is given by

TOTAL FREQUENCY VARIANCE:

$$\sigma^2_{TOTAL}(\tau)_{\bar{y}} = \left(\frac{1}{\tau^2}\frac{1}{2(N-m-1)}\right) \times \sum_{n=1}^{N-m-1}[x_0(n) - 2x_0(n+m) + x_0(n+2m)]^2. \qquad (4)$$

Notice that n+m stops just short of N; that avoids an identically zero second difference.

The evolution of a time error over τ (TIME VARIANCE) between clocks is defined as the frequency variance times (τ^2·constant) and is calculated using the same procedure as eq (4) with a slight modification which distinguishes levels of phase coherence, namely WHPM and FLPM [Ref. 21]. Determining phase coherence is important particularly in synchronous networks which need to characterize time coherence in the long term in addition to the more relaxed but less meaningful requirement of frequency coherence. Phase coherence, in a statistical sense, is inferred by the trend in the variability (convergence property) as a function of τ-averaged phase. Thus we construct TOTAL-TVAR (with square-root as TOTAL-TDEV) as

TOTAL TIME VARIANCE:

$$\sigma^2_{TOTAL}(\tau)_{\bar{x}} = \left(\frac{1}{3}\frac{1}{2(N-m-1)}\right) \times \sum_{n=1}^{N-m-1}[\bar{x}_0(n) - 2\bar{x}_0(n+m) + \bar{x}_0(n+2m)]^2. \qquad (5)$$

A summary of methods of improved estimation applied to eqs (4) and (5) are presented next.

3. Endpoint Matching

Often we remove at least a regressed linear slope in an ordered set of time-difference measurements given by $\{x_{k'}\}$. In practice this removes an overall frequency difference but instead of removing a regressed linear slope, there is a case for removing a linear slope designed to match the endpoints of time series $\{x_{k'}\}$. This has other advantages in the context of the data extension procedure of eqs (4) and (5).

A series $\{x_{k'}\}$ represents the finite observation of ordered random variables. All observations are made through an observation window function which (unless otherwise noted) is rectangular in shape. Treating the series as a function over a finite interval, one commonly determines rate and drift using regression analysis of x(t) or its derivative y(t). Subsequent removal of a chosen model of trends can significantly alter an essential characteristic of the original series of measurements particularly at long-term for non-white noise types [Refs. 4,5]. By using a circular representation of a finite time series, one can arbitrarily time shift any chosen observation window or sampling function such as shown in fig.1 [Refs. 6,11,14]. In the moved-window case, the beginning and endpoints of the series $\{x_{k'}\}$ are somewhat matched given by the trend removal, correlation, finite physical boundaries, and measurement-system bandwidth of the measured series. However, the turn-on/turn-off transient of the window, becomes an artifact which is not representative of the functional "roughness" or "smoothness" of the series and as such should be eliminated by a removal of linear slope which matches the endpoints prior to calculating a measure of such quantities as variance. If the incremental differences are stationary this methodology does not change the linear combinations of incremental differences contained in eqs (4) and (5).

4. Reflected Time Series

The wrap procedure implies an extension of the data by the original series which of course has identically distributed statistics, that is, the circularized series has the same variance as the original series. To reduce endpoint turn-off transients, a source of what is called in signal processing "leakage" [Refs. 15,18], we apply an endpoint match. We also desire smooth derivatives in the extension to properly estimate low-frequency noise components in whatever statistic we choose. This is particularly important in handling RWFM and drift. Frequency variances are not affected by reversing the direction of the series of measurements. Therefore, reflecting and wrapping the time series about the first and last points (both made to equal 0 as shown in 3 above) implies that we have made the most prudent assumptions and practical extensions of the time series in order to reduce transients, and hence reduce leakage. Figure 3 which shows series $\{x_{k'}\}$ reflected about the last point, represents simulated RWPM (or equivalently WHFM). We construct a new sequence of numbers which reflect about the last point (zero) however again I clarify that this construction is not a real extension; it is a convenience in calculating a maximally-overlapped estimate of the new 3-sample/2-sample variance suggested here.

5. Normalization of Sample Variance and Effective Degrees of Freedom (edf)

5.1 Illustration: We start with the sample mean of a set of discrete variables indexed by k' as given by

$$\bar{x} = \frac{1}{M}\sum_{k'=1}^{M} x_{k'}, \qquad k'=1,2,3,...M. \tag{6}$$

The mathematical form of the standard sample variance of the mean looks very much like the sample mean, an average, but in fact it is not an average, i.e., a sum of numbers divided by the total number of numbers as in eq (6). The sample variance is sometimes described as a sum of deviates divided by (or "scaled" by) the assumed number of degrees of freedom, often believed to be M, the total number of deviates. But the individual deviates are always some "difference" quantity derived from the grand mean of the set of random variables (recall the basic measure is "variance") and there can never be any more than M-1 differences that are "independent observations" (M-1 degrees of freedom). Calculating the variance implies that M (the number of samples) is at least two. Random variables may be differenced relative to a non-zero mean (or an assumed or actual zero mean), relative to a starting, ending, preceding, or following random variable, or for that matter any quantity derived from the set of M variables. Starting with independent random variables, any variance which properly preserves independence should have a scaling or normalization factor given by M-1. Division by this scaling factor yields a sample variance often called an "unbiased estimate," and the scaling factor is the number of independent observations.

Therefore, if $\bar{x} = 0$, the normalized variance (or variation) about a zero mean is given by the mean-square deviates as:

$$\sigma^2_{\bar{x}=0} = \frac{1}{M-1}\sum_{k'=1}^{M}(x_{k'})^2. \tag{7}$$

Recall that the actual number of observations is M in our quest to find some "true" mean which we don't know but which the finite grand or sample mean estimates. Therefore we can assume M independent observations (M degrees of freedom) to the extent that the sample mean estimates the true mean. For a white (WH) process, the sample mean whose variability decreases as M is the optimum estimate of the true mean, thus M degrees of freedom yields an unbiased estimate of its variance. For a random walk (RW) process, there is no true mean, nevertheless there is a sample (moving) mean, the bias on its variance is readily calculable and turns out to be small (an error by a worst case factor of 1.5 at $\tau=T/2$ for the two-sample frequency variance). If we can judge the noise type, hence correct for bias at long term, then we have essentially M degrees of freedom. Noise type can be determined from empirical data by its normalized autocorrelation function discussed next.

5.2 Effective Number of Independent Observations (*edf*): Barnes [Refs. 3,10,21] did extensive work on estimation of the two-sample frequency variance, and introduced bias functions and fractional degrees of freedom. Small-sample statistics in which $\tau \to T/2$ was not treated as a special case is the main subject at hand. In order to treat this case, I very briefly introduce concepts of self-similarity, or the autocorrelation function. Where $\bar{y}_t$ are frequency deviations averaged over τ, we can relate an autocorrelation to the two sample variance [Ref. 25]. We have in simplified form

$$\Sigma(\bar{y}_t-\bar{y}_{t+\tau})^2 = \Sigma(\bar{y}_t^2+\bar{y}_{t+\tau}^2-2\bar{y}_t\bar{y}_{t+\tau}) = \tag{8}$$

$$\Sigma\bar{y}_t^2 + \Sigma(\bar{y}_{t+\tau}^2 - 2\bar{y}_t\bar{y}_{t+\tau}) = \Sigma\bar{y}_t^2 + (\Sigma\bar{y}_{t+\tau}^2)(1-\frac{r_\tau}{2})$$

$$where\ r_\tau = \frac{\Sigma\bar{y}_t\bar{y}_{t+\tau}}{\Sigma\bar{y}_{t+\tau}^2}.$$

The coefficient r_τ is a good approximation to the normalized sample autocovariance (autocorrelation) given by

$$r_L = \Sigma x_i x_{i+L} / \Sigma x_i^2 \qquad (9)$$

where the x_i are deviations from the mean of the series considered, and L denotes the lag between the values of the product. It has often been found adequate to assume that the correlation between more distant values arises solely from that between the directly neighboring values; if that correlation is R, the correlation coefficient r_L of values L time units apart becomes equal to R^L [Ref. 16].

The assumption for WHFM is that the normalized autocorrelation of average frequency deviations yields $r_\tau = 0$ everywhere except at $\tau = 0$ where $r_0 = 1$. A non-zero autocorrelation in a series "reddens" its spectrum of deviations $S_y(f) = \sum h_\alpha f^\alpha$, giving a greater share of the total variability to longer periods and a smaller share to shorter ones [Ref. 7]. Processes where $\alpha < 0$ contain memory in the sense that correlation between long time intervals arises from values at the shortest interval denoted as τ_0 [Ref. 5]. As a consequence, the variability of sample averages of frequency deviations with memory (approaching RWFM) decreases more slowly with increasing sample size than does that of averages from a white series without memory [Ref.22]. Thus the number of independent observations is expressed as an *effective* number of degrees of freedom *(edf)* smaller than the actual number used in the summand of the sample variance [Ref. 20]. This effective number of independent observations *(edf)* is essentially the number of equivalent degrees of freedom based on the autocorrelation function in the samples themselves. We have

$$(edf) = (M - \epsilon_\gamma(M)) \qquad (10)$$

where ϵ depends on M and autocorrelation properties, subscript γ designates the type of sample frequency variance, and *(edf)* can be fractional, i.e., non-integer. To be concise, I limit the discussion to M=2 which turns out to be the most interesting case (longest term) representing $\tau = T/2$.

Properties of the distribution of the usual Allan-deviation estimate have been studied using fractional degrees of freedom on the confidence interval [Refs. 9,10]. Similar studies can also be applied to the actual time series, since there is a correspondence between these distribution properties and the underlying noise process of the data. Keep in mind that the two-sample frequency variance is a time-averaged, standard-sample variance against a previously measured mean. Its square root defines a relative uncertainty on this τ-averaged mean. Thus the sample Allan deviation is an uncertainty of a sample mean frequency decomposed by τ averaging times. The underlying noise type defines the trend in this uncertainty. Moreover the number of degrees of freedom in the underlying noise type has a correspondence to the degrees of freedom in the uncertainty. In this form for the Allan variance, $\epsilon_{AVAR}(2)$ can only range from 1 to 1.5 corresponding to WHFM to RWFM respectively since RWFM as an integral of WHFM. This is because the means of segments of data will converge to the grand mean only half as fast for RWFM vs. WHFM.

5.3 Actual Number of Observations (h): The sample Allan variance is useful as a power-law (octave-band) spectral estimator but is time-shift (phase) sensitive and depends on where we start the calculation with respect to data in process. For large data sets and small scale values of m, the odd and even values of index k overlap and average together in eq (1) for a fairly accurate estimation of a broadband spectral distribution of variance of first differences of average frequency. The division by 2(M-1) which corresponds to 2(N-m-1) in eqs (4) and (5) is arguably due to overlapping two sets of deviates and has constant "2" only for WHPM, FLPM, and WHFM but ought to approach M-1 (or N-m-1) for an accurate estimate in the statistic as $m\tau_0 \rightarrow T/2$, since the first and last deviates do not overlap. However, at the largest scale, the estimate is negatively biased for non-WHFM (i.e., FLFM and RWFM) because there is only 1 (not 2) "two-sample" sample. This reason causes an estimation error at long term in virtually all cases in which the noise is no longer WHFM but is FLFM or RWFM even though the estimation is supposedly unbiased even at these large τ values (or equivalently large values of m).

For the statistic TOTALVAR given in eq (4), there is considerable overlap in the summation. It is widely known that this effectively smooths the estimate [Refs. 10-12]. But the number h of actual observations (hence the scaling factor) has not changed and remains a straightforward calculation giving

$$h = \frac{[2(\log_2(\frac{N-1}{m})) - 1]}{(\log_2 \frac{N-1}{m})} (N-m-1). \qquad (11)$$

We find in simulation studies that eq (11) should be applied for RWFM, and not at all for WHFM [Refs. 12,19]. This suggests that the division by 2(M-1) ought to approach M-1 because of the effects of correlation and not necessarily because of a connection with the actual number of observations as discussed above. Nevertheless, if h is the actual number of observations, note that for large M values, $(edf) \doteq h$, but that for small M, *(edf)* becomes increasingly sensitive to the value of ϵ. A table of values of M, and corresponding ϵ_{AVAR} for various noise types is planned for future work. For the purpose here, we find that in the presence of RWFM, substituting 2(N-m-1) in eqs (4) and (5) by h in eq (11) above removes a negative bias in estimates of sample frequency variance at long averaging times.

6. Conclusion

I have introduced new statistics of frequency and time variances which yield improved estimation of both frequency stability and noise type, particularly at long averaging times. An initial procedure involves regressing to global basis functions such as orthogonal polynomials. This procedure of detrending assumes the residuals are white phase noise. Unfortunately frequency variations usually exhibit systematic effects that eventually change and are interpreted as arising from a divergent noise type hence non-stationary sample frequency variance. Hence in the case of the two-sample Allan variance the effective number of observations (equivalent degrees of freedom) is slightly reduced by factor ϵ which ranges from 1 in the presence of WHFM to 1.5 for RWFM for the largest τ-value at T/2.

7. References

[1] Allan, D.W. and Barnes, J.A., A modified "Allan Variance" with increased oscillator characterization ability, Proc. 35th Freq. Control Symp., 470-476 (1981).

[2] Allan, D.W., Weiss, M.A., and Jespersen, J.L., A frequency-domain view of time-domain characterization of clocks and time and frequency distribution systems, Proc. 45th Freq. Control Symp., 667-678 (1991).

[3] Barnes, J.A., Chi, A.R., Cutler, L.S., Healy, D.J., Leeson, D.B., McGunigal, T.E., Mullen Jr., J.A., Smith, W.L., Sydnor, R.L., Vessot, R.F.C., and Winkler, G.M.R., Characterization of frequency stability, IEEE Trans. Instrum. Meas., **IM-20**, 105-120 (1971).

[4] Barnes, J.A., The measurement of linear frequency drift in oscillators, Proc. 15th Ann. PTTI Meeting, (1983). Article is also contained in NIST Technical Note 1337 (1990).

[5] Beran, J., Statistical methods for data with long-range dependence, Statistical Science, 7, 404-427 (1992).

[6] Bloomfield, P., Fourier analysis of time series: an introduction, Wiley, New York, 1976. Pages 55-60 are a useful discussion of the "circular definition" of a noise (or roughness) coefficient, namely the variance, using "wrapped" data.

[7] Gilman, D.L., Fuglister, F.J., and Mitchell, J.M., Jr., On the Power Spectrum of "Red Noise", J. Atm. Sci. **20**, 182-184 (1963).

[8] Greenhall, C.A., A structure function representation theorem with applications to frequency stability estimation, IEEE Trans. Instrum. Meas., **IM-32**, pp. 364-370 (1983).

[9] Greenhall, C.A., Recipes for degrees of freedom of frequency stability estimators, IEEE Trans. Instrum. Meas., **40**, pp. 994-999 (1991).

[10] Howe, D.A., Allan, D.W., and Barnes, J.A., Properties of signal sources and measurement methods, Proc. 35th Ann. Symp. Freq. Control, A25-A35 (1981). Article is also contained in Natl. Inst. Stand. Technol. TN-1337 (1990).

[11] Howe, D.A., An extension of the Allan variance with increased confidence at long term, Proc. 1995 IEEE Int. Freq. Cont. Symp., pp. 321-329.

[12] Howe, D.A. and Lainson, K.J., Simulation study using a new type of sample variance, Proc. 1995 PTTI Meeting, November, pp. 279-290 (1995).

[13] Howe, D.A. and Lainson, K.J., Effect of drift on TOTALDEV, Proc. 1996 IEEE Int. Freq. Control Symp., pp. 883-889.

[14] Howe, D.A., Circular representation of infinitely extended sequences, Proc. 1995 IEEE Int. Freq. Cont. Symp., pp. 337-345 (1995).

[15] Howe, D.A. and Percival, D.B., Wavelet Variance, Allan Variance, and Leakage, IEEE Trans. Instrum. Meas., IM-44, 94-97 (1995). Also, Howe, D.A. and Percival, D.B., Wavelet Analysis for Synchronization and Timekeeping, Proc. IEEE Int. Freq. Control Symp., pp. 791-794 (1994).

[16] Kendall, M.G. and Stuart, A., *The Advanced Theory of Statistics*, **3**, Griffin, London (1966).

[17] Legachev, V.A. and Pashev, G.P., Estimation of linear frequency drift coefficient of frequency standards, Proc. 1996 IEEE Int. Freq. Control Symp., pp.

[18] Percival, D.B. and Walden, A., Spectral analysis for physical applications, Cambridge University Press, New York, 1992.

[19] Riley, W.J., Addendum to a test suite for the calculation of time domain stability, Proc. 1996 IEEE Int. Freq. Control Symp., pp. 890-893.

[20] Radok, U., Assessing climate persistence from climatic noise, *Climatic Change* (1992).

[21] Sullivan, D.B., Allan, D.W., Howe, D.A., and Walls, F.L., Characterization of Clocks and Oscillators, NIST Technical Note 1337 (1990).

[22] Vernot, F., and Vincent, M., Estimation of the uncertainty of a mean frequency measurement, Proc. 11th European Freq. and Time Forum, 1997 (to be published these proceedings).

[23] Weiss, M.A., Allan, D.W., Howe, D.A., Confidence on the second difference estimation of frequency drift, Proc. IEEE 46th Int. Freq. Control Symp., pp. 300-305 (1992).

[24] Weiss, M.A. and Hackman, C., Confidence on the three-point estimator of frequency drift, Proc. 24th Ann. PTTI Meeting, pp. 451-460 (1992).

[25] Yoshimura, K., Characterization of frequency stability: Uncertainty due to the autocorrelation of the frequency fluctuations, IEEE Trans. Instrum. Meas., **IM-27**, 1-7 (1978).