A Frequency-Drift Estimator and Its Removal from Modified Allan Variance

C. A. Greenhall*
Jet Propulsion Laboratory, MS 298-100, 4800 Oak Grove Dr., Pasadena, CA 91109 USA

Abstract

A drift-rate estimator constructed from four values of the cumulative sum of clock residuals is shown to have good error performance in the presence of the five standard power-law noises. A comparison table of several drift estimators is given. The bias and variance (or equivalent degrees of freedom) of a modified Allan variance estimator incorporating drift removal is calculated.

1 Introduction

The confidence of estimates of modified Allan variance (mvar) can be derived from previously published formulas and algorithms [1, 2, 3], but only for situations in which mvar is not dominated by linear frequency drift. For such a situation to hold, either the actual drift rate must be negligible for a given span of clock data, or the drift rate must be removed after being estimated from a longer span of data or by another method, such as hydrogen-maser cavity tuning. The present investigation has two goals: 1) design of a drift estimator with satisfactory error performance in the presence of the five standard power-law phase noise models; 2) finding out how removal of drift, as estimated from the *current* data, affects the bias and variance of the estimated mvar of the residuals, and thereby designing an automatic procedure for assigning mvar confidence intervals.

The first goal is achieved by a linear combination of four values chosen from the sequence of cumulative sums of the time residuals. The variance of the chosen drift estimator, for the five standard noise models, is compared to that of several other drift estimators. Table 1 gives a concise presentation of their variances in a uniform notation, along with a "discreteness" classification of drift estimators.

The second goal is partly achieved, in that the required mean and variance computations were successfully carried out; results are presented below (Figs. 2 and 3). Unfortunately (and not unexpectedly), the bias of the "net" (drift removed) mvar estimator depends so heavily on the noise type that the author does not know how to compensate for the bias without human judgment of the dominant noise type and a risky extrapolation of the sigma-tau curve to an unobservable region.

2 Drift Estimator Design

The design is based on continous-time power-law models of phase noise. Let $x(t), 0 \le t \le T$, be the time departure of a clock, with $y(t) = dx(t)/dt$ the normalized frequency departure, and let $w(t) = \int x(t)\,dt$, the continuous-time analog of the sequence $w_n = \sum_{j=1}^{n} x(n\tau_0)$, whose third differences can be used for computing modified Allan variance [2]. The idea is to make an unbiased estimator of frequency-drift rate from discrete values of $w(t)$ instead of values of $x(t)$, thus gaining the advantage of an integration over the noise in $x(t)$. A quadratic component $\frac{1}{2}ct^2$ of $x(t)$ appears as a cubic component $\frac{1}{6}ct^3$ of $w(t)$; consequently, at least four values of $w(t)$ are needed.

Consider the one-parameter family of estimators

$$\hat{c}(r) = \frac{6}{T^3 r(1-r)} \times \left[w(T) - w(0) - \frac{w(T-rT) - w(rT)}{1-2r}\right], \quad (1)$$

where $0 < r < 1/2$. If $w(t)$ were a cubic polynomial with leading term $\frac{1}{6}ct^3$, then $\hat{c}(r)$ would equal c. Selection of r is based on the behavior of the variance of $\hat{c}(r)$ under the five standard power-law x noises: white PM, flicker PM, white FM, flicker FM, and random-walk FM, with spectral densities $S_x(f) \propto f^{\beta}$,

*Email: cgreen@pop.jpl.nasa.gov. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

$\beta = 0, -1, -2, -3, -4$. Since $S_w(f) \propto f^{\beta-2}$, which is integrable over high frequencies, one can do without a high-frequency cutoff.

The parameter r is chosen according to a minimax criterion. By the method of the generalized autocovariance (gacv), closed-form expressions for $\operatorname{var}\hat{c}(r)$ as a function of r can be derived for the five noise types. Figure 1 shows a plot of the ratio of $\operatorname{var}\hat{c}(r)$ to its minimum value for each β. Since the upper envelope of the five functions has a minimum at $r = 0.0958\ldots$ (the intersection of the curves for white PM and white FM), it is reasonable to choose $r = 1/10$ for simplicity. Doing so gives a drift estimator

$$\hat{c}_{w4} = \frac{50}{3T^3} \times \left[4w(T) - 4w(0) - 5w\left(\frac{9T}{10}\right) + 5w\left(\frac{T}{10}\right)\right], \tag{2}$$

henceforth called the *four-point* w estimator, abbreviated as $w4$. It is also apparent from Fig. 1 that the performance of the estimator could be improved by eliminating white PM from the noise set; the corresponding minimax value of r would be about 0.0337.

In practice, one uses a discrete-time version of $w4$: Given phase data $x_n = x(n\tau_0)$ for $n = 1, \ldots, N$, form the sequence w_n, where w_0 is arbitrary (usually 0), $w_n = w_0 + \sum_{j=1}^{n} x_j$. Choosing an integer n_1 close to $N/10$, let $r_1 = n_1/N$. The drift estimator is given by

$$\hat{c}_{w4d} = \frac{6}{N^3\tau_0^2 r_1(1-r_1)}\left[w_N - w_0 - \frac{w_{N-n_1} - w_{n_1}}{1-2r_1}\right]. \tag{3}$$

Theoretical formulas for the variance of $\hat{c}_{w4}$ were checked by simulations of $\hat{c}_{w4d}$ for all five noise types, with 1000 runs of $N = 100$ points each. Excellent agreement was observed. Below, a possible method for estimating the variance of $\hat{c}_{w4}$ from the data is given.

3 Comparison with Other Estimators

Recent papers of Logachev and Pashev [4] and of Wei [5] give variance tables for other unbiased drift estimators under all or some of the standard noise types. Following are names and abbreviations for these estimators, and formulas for the continuous-time analogs that were used for verifying the previous results and consolidating them into a uniform notation.

Least-squares quadratic fit to x (LSx), optimal for white PM:

$$\hat{c}_{\mathrm{LS}x} = \frac{60}{T^5}\int_0^T (6t^2 - 6Tt + T^2)\,x(t)\,dt \tag{4}$$

$$= \frac{60}{T^3}[w(T) - w(0)] - \frac{360}{T^5}\int_0^T (2t - T)\,w(t)\,dt. \tag{5}$$

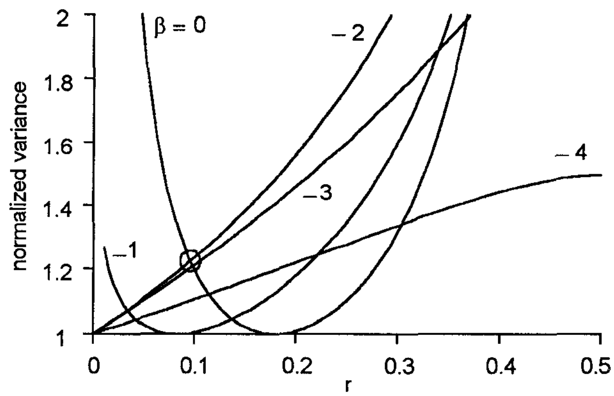

Figure 1: Parameter selection for the 4-point w drift estimator. The minimax point of the five normalized variance curves is circled.

Three-point x ($x3$), also called overall second-difference:

$$\hat{c}_{x3} = \frac{4}{T^2}\left[x(0) - 2x\left(\frac{T}{2}\right) + x(T)\right]. \tag{6}$$

Least-squares linear fit to y (LSy), optimal for white FM:

$$\hat{c}_{\mathrm{LS}y} = \frac{6}{T^3}\int_0^T (2t - T)\,y(t)\,dt \tag{7}$$

$$= \frac{6}{T^2}[x(0) + x(T)] - \frac{12}{T^3}\int_0^T x(t)\,dt. \tag{8}$$

Least-squares constant fit to $z = dy/dt$ (LSz), also called two-point y, optimal for random walk FM:

$$\hat{c}_{\mathrm{LS}z} = \frac{1}{T}\int_0^T z(t)\,dt = \frac{1}{T}[y(T) - y(0)]. \tag{9}$$

As defined, this estimator can be applied to random walk FM but not to infinite-bandwidth white and flicker FM (let alone the PM noises), for which point values of $y(t)$ are not defined. In its place one uses a discrete-time version called *two-point* $\bar{y}$ ($\bar{y}2$) or mean second-difference:

$$\hat{c}_{\bar{y}2} = \frac{N^2}{(N-1)T^2} \times \left[x(T) - x\left(T - \frac{T}{N}\right) - x\left(\frac{T}{N}\right) + x(0)\right], \tag{10}$$

for which the sample period of $x(t)$ is presumed to be T/N.

These drift estimators fall into a natural classification that determines the noise types over which they are effective. The $w4$ and LSx estimators are called *w-discrete* because they contain discrete values of $w(t)$, and perhaps also integrals over $w(t)$. Likewise, $x3$, LSy, and $\bar{y}2$ are x-discrete, and LSz is y-discrete.

Table 1 gives the variance of all these drift estimators over the five noise types, scaled according to the convention $S_y^+(f) = h_{\beta+2} f^{\beta+2}$ for the one-sided spectral density of $y(t)$. For these results to apply to the actual discrete-time estimators, the high-frequency cutoff f_h of the noise must satisfy the Nyquist criterion for the sample period τ_0, i.e., $2f_h\tau_0 \leq 1$ [6]. The results are asymptotic relative to the assumptions $2\pi f_h T \gg 1$, $N \gg 1$. With minor changes in logarithmic expressions, the results agree with the cited references. The numbers in brackets are the rankings of the estimators over those noises for which the variance is independent of f_h and data size N (in the range of the assumptions).

The similarity of the estimators in the same discreteness class is apparent. The w-discrete estimators are bandwidth-independent for all the noises, the x-discrete estimators only for the FM noises. If all the noise types are included, then $w4$ is the best overall drift estimator. If only the FM noises are included, then LSy is best; even so, for random walk FM the $w4$ variance is only 10% more than the LSy variance.

4 Gross and Net Mvar

Assume that the time deviation process $x(t)$ has stationary second differences. Then it has a constant frequency drift rate c_x, which, if nonzero, gives rise to an mvar component $c_x^2\tau^2/2$ that dominates mvar for long averaging times. In terms of the time residuals $x_n = x(n\tau_0)$ and their cumulative sums w_n, we have

$$E\left(\Delta_m^2 x_n\right) = c_x\tau^2, \quad E\left(\tau_0\Delta_m^3 w_n\right) = c_x\tau^3,$$

where $\tau = m\tau_0$, E denotes mathematical expectation, and Δ_m is the backward difference operator with stride m. According to the third-difference formulation of mvar [2],

$$\operatorname{mod}\sigma_y^2(\tau) = \frac{1}{2\tau^4} E\left[\left(\tau_0\Delta_m^3 w_n\right)^2\right].$$

Because it includes drift, this is called gross mvar. To define net mvar, replace the expected square by the variance:

$$\begin{aligned}\operatorname{mod}\sigma_{y0}^2(\tau) &= \frac{1}{2\tau^4}\operatorname{var}\left(\tau_0\Delta_m^3 w_n\right)\\ &= \frac{1}{2\tau^4}E\left[\left(\tau_0\Delta_m^3 w_n - c_x\tau^3\right)^2\right]\\ &= \operatorname{mod}\sigma_y^2(\tau) - \frac{c_x^2\tau^2}{2}.\end{aligned}$$

Net mvar, which is invariant to the value of c_x, can also be defined as mvar of the reduced time residual process $x(t) - c_x t^2/2$.

Now suppose that one has time data $x_1, \ldots, x_N$ with sample period τ_0, and let $T = N\tau_0$. For any constant c, form the quantity

$$V_x(\tau, T, c) = \frac{1}{2\tau^4 M}\sum_{n=3m}^{N}\left(\tau_0\Delta_m^3 w_n - c\tau^3\right)^2, \quad (11)$$

where $M = N-3m+1$. Then $V_x(\tau, T, 0)$ is an unbiased estimator of gross mvar (and also gives (11) for any c if x_n is replaced by $x_n - c\tau_0^2 n^2/2$ or w_n by $w_n - c\tau_0^2 n^3/6$). If c_x is known, then $V_x(\tau, T, c_x)$ is an unbiased estimator of net mvar. More often, one has some unbiased estimate $\hat{c}$ that depends only on the data at hand. In this case, the corresponding estimator $V_x(\tau, T, \hat{c})$, while nonnegative and invariant to the true value of c_x, is biased for net mvar because subtracting an estimated drift tends to cut into the long-term random fluctuations.

For theoretical computations of the mean and variance of these estimators, it is convenient to approximate the above setting by a continuous-time formulation that uses the asymptotic modified Allan variance of Bernier [6] and a continuous-time analog of (11) in which the sum becomes an integral. This approximation is valid provided $\tau/\tau_0 \gg 1$; simulations indicate that $\tau/\tau_0 \geq 8$ is adequate. The $w4$ drift estimator $\hat{c}_{w4}$ is used for forming the biased net mvar estimator. Because $V_x(\tau, T, c_x)$ and $V_x(\tau, T, \hat{c}_{w4})$ are invariant to true drift rate c_x, one can assume $c_x = 0$; then the third w-differences have mean zero. Using the gacv method, one can compare the mean of $V_x(\tau, T, \hat{c}_{w4})$ to the true net mvar; assuming also that the third w-differences form a Gaussian process, one can compute the variance of $V_x(\tau, T, c_x)$ and $V_x(\tau, T, \hat{c}_{w4})$. The computations, similar to those for conventional Allan variance [7], are not given here.

Figure 2 shows the bias of the net mvar estimator $V_x(\tau, T, \hat{c}_{w4})$ in terms of mdev (square root of mvar) as a function of τ/T for the five standard noise types. *Caution:* Plotted is the square root of the normalized mean of the mvar estimator, *not* the normalized mean of the square root of the mvar estimator. As an example, take the most extreme case, random walk FM and $\tau = T/3$, for which $EV_x(\tau, T, \hat{c}_{w4}) = 0.06352 \operatorname{mod}\sigma_{y0}^2(\tau)$; the plotted value is $\sqrt{0.06352} = 0.252$. Simulation results ($N = 1152$, 10000 trials),

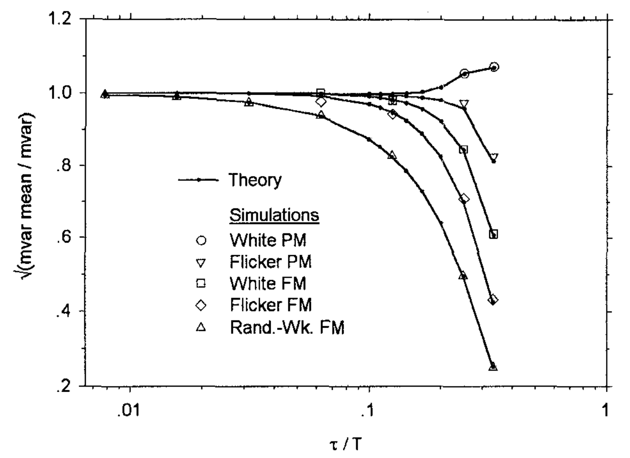

Figure 2: Bias of net mvar estimator with drift removed, expressed as square root(estimator mean / true mvar).

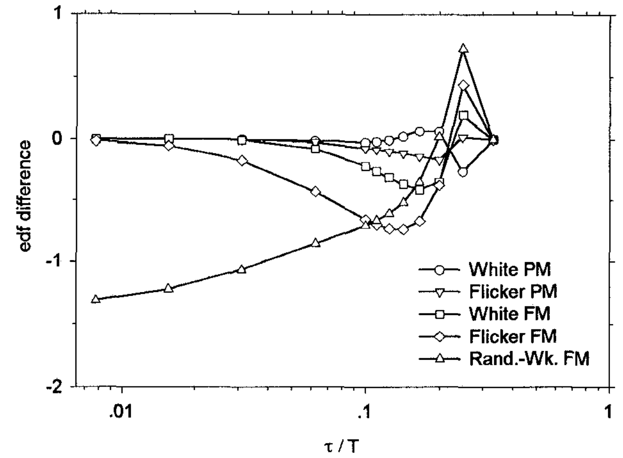

Figure 3: Change in equivalent degrees of freedom when removing drift from mvar estimate. Plotted is edf(biased) - edf(unbiased).

shown by the open symbols, agree well enough with theory to serve as curve labels. Especially in view of the persistent large negative bias for random walk FM (still -12.5% in terms of mdev for $\tau = T/10$), one needs to adjust measurement results on a model-dependent basis.

Figure 3 shows how removing the $w4$ estimated drift changes the confidence of the mvar estimator. Confidence is measured by equivalent degrees of freedom (edf), defined for a positive random variable X by $\operatorname{edf} X = 2\,(EX)^2 / \operatorname{var} X$. Computations and approximations for the edf of the unbiased estimator $V_x(\tau, T, c_x)$ have previously been given [2, 3]. Here, the continuous-time formulation was used for approximating those computations and computing the edf of the biased net mvar estimator $V_x(\tau, T, \hat{c}_{w4})$. Plotted is edf (biased) $-$ edf (unbiased) vs. τ/T for the standard noises. The relative difference is small because all the edfs are of order T/τ; a simple conservative approximation for edf (biased) is edf (unbiased) $-$ 0.75. At $\tau = T/3$, each edf is 1 because the estimator is the square of a single Gaussian random variable.

5 Estimating the Drift Estimator Variance

In their discussion of the $x3$ drift estimator, Weiss and Hackman [8] point out that its variance is simply $(8/T^2)\,\sigma^2_{y0}(T/2)$, where $\sigma^2_{y0}(\tau)$ is net conventional Allan variance, i.e., Allan variance with the true drift removed. In turn, $\sigma^2_{y0}(T/2)$ is to be estimated from the data by extrapolating the *estimated* $\sigma^2_{y0}(\tau)$ (using $\hat{c}_{x3}$ itself to remove drift) for lesser τ out to $\tau = T/2$. This requires human judgment of the behavior of the net sigma-tau curve in the face of increasing bias and variance as τ increases.

The variance of the $w4$ drift estimator can be estimated by a similar method using net mvar. One finds that

$$\operatorname{var} \hat{c}_{w4} = \frac{A^2_\beta}{T^2} \operatorname{mod} \sigma^2_{y0}(T/3),$$

where $A_\beta = 3.70, 3.14, 3.14, 3.41, 3.80$ for $\beta = 0$ to -4 (white PM to random walk FM). Therefore, a conservative estimate of the standard deviation of $\hat{c}_{w4}$ is $(3.8/T) \operatorname{mod} \sigma_{y0}(T/3)$. Again, this requires intelligent extrapolation of the curve for estimated net mdev out to $\tau = T/3$, where net mvar is essentially unobservable because its estimator has one degree of freedom and a bias as large as -93.6%.

6 Concluding Remarks

The four-point w drift estimator described above deserves consideration as a general-purpose method for estimating frequency drift rate. From Table 1 one can calculate the ratio of its standard deviation to those of the optimal estimators for the even-power noises: 1.242 for white PM, 1.111 for white FM, and 1.151 for random-walk FM. Although the random-walk FM case is important, its optimal drift estimator, mean second difference, performs poorly in the presence of other phase noises. Moreover, the standard deviation of the four-point w estimator is only 1.051 times that of the second-place estimator, least-squares linear fit to frequency.

The heavily model-dependent biases shown in Fig. 2 lead to an unsatisfactory situation in which guesses about the long-term noise type have to be made in order to compensate for the bias of the net mvar estimator. Other methods for uncoupling deterministic and random aspects of clock data are already being investigated or used. Higher-order variances, such as Hadamard variance (mean-square third difference of x) and wavelet variances, automatically kill the quadratic component of $x(t)$. The "totalvar" processing method, which augments a data sequence with a reflected copy of itself, has been found to reduce the bias of drift removal from conventional Allan variance in a specific case [9]. Perhaps a combination of frequency-domain techniques could be useful: one might perform a spectral estimation procedure to characterize the random noise, while estimating the drift rate as the mean of the second phase differences by applying a data taper with low sidelobes, to reject all but the lowest-frequency components. One could hope to assign confidence intervals to the results in a model-free way.

References

[1] T. Walter, "Characterizing frequency stability: a continuous power-law model with discrete sampling", *IEEE Trans. Instrum. Meas.*, vol. 43, pp. 69–79, 1994.

[2] C. Greenhall, "Estimating the modified Allan variance", *Proc. 1995 IEEE Internat. Freq. Control Symp.*, pp. 346–353, 1995.

[3] M. Weiss and C. Greenhall, "A simple algorithm for approximating confidence on the modified Allan variance and the time variance", *24th Ann. PTTI Meeting*, 1996.

[4] V. A. Logachev and G. P. Pashev, "Estimation of linear frequency drift coefficient of frequency standards", *Proc. 1996 IEEE International Freq. Control Symp.*, pp. 960–963, 1996.

[5] G. Wei, "Estimations of frequency and its drift rate", *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 79–82, 1997.

[6] L. G. Bernier, "Theoretical analysis of the modified Allan variance", *Proc. 41st Ann. Freq. Control Symp.*, pp. 116–121, 1987.

[7] C. Greenhall, "The fundamental structure function of oscillator noise models", *Proc. 14th Ann. PTTI Meeting*, pp. 281–294, 1982.

[8] M. Weiss and C. Hackman, "Confidence on the three-point estimator of frequency drift", *Proc. 24th Ann. PTTI Meeting*, pp. 451–455, 1992.

[9] D. Howe and K. Lainson, "Effect of drift on TOTALDEV", *Proc. 1996 IEEE Internat. Freq. Control Symp.*, pp. 883–889, 1996.

Table 1: Variance of frequency drift estimators. Names: $w4$ = 4-point w, LSx = least-squares quadratic fit to x, $x3$ = 3-point x, LSy = least-squares linear fit to y, LSz = least-squares constant fit to z, $\bar{y}2$ = 2-point $\bar{y}$. Each entry is to be multiplied by the factor on the right. The numbers in brackets are rankings within each noise type.

	w-discrete		x-discrete			
noise type	$w4$	LSx	$x3$	LSy	LSz or $\bar{y}2$	factor
white PM	$\frac{1250}{9}$[2]	90[1]	$24 f_h T$	$18 f_h T$		$h_2 \pi^{-2} T^{-5}$
flicker PM	74.84[1]	75[2]	$24 \ln(4.441 f_h T)$	$18 \ln(4.117 f_h T)$		$h_1 \pi^{-2} T^{-4}$
white FM	$\frac{200}{27}$[2]	$\frac{60}{7}$[4]	8[3]	6[1]	N	$h_0 T^{-3}$
flicker FM	10.900[2]	$\frac{25}{2}$[4]	$16 \ln 2$[3]	9[1]	$3 + 2 \ln N$	$h_{-1} T^{-2}$
rand.-wk. FM	$\frac{358}{135}$[3]	$\frac{20}{7}$[5]	$\frac{8}{3}$[4]	$\frac{12}{5}$[2]	2[1]	$h_{-2} \pi^2 T^{-1}$