Stochastic Models for Atomic Clocks

J. A. Barnes, R. H. Jones, P. V. Tryon*, and D. W. Allan

ABSTRACT

Most workers in the field of atomic clocks encounter frequency and time instabilities which can be characterized (or modelled) as random fluctuations. These random fluctuations typically display a power spectral density which varies as a power-law over some significant range of (Fourier) frequencies (e.q., $S_y(f) = h_2 f^2$, where Y denotes the normalized, instantaneous frequency and f denotes the Fourier frequency). Typical oscillators and/or clocks may have regions where one specific power-law predominates and other regions where other power-laws predominate. In general, various combinations of five different power-laws seem to be adequate to describe almost all observed random behavior in atomic clocks. The five types are:

White phase modulation	$S_y(f) = h_2 f^2$
Flicker phase modulation	$S_y(f) = h_1 f^1$
White frequency modulation	$S_y(f) = h_0 f^0$
Flicker frequency modulation	$S_y(f) = h_{-1} f^{-1}$
Random Walk frequency modulation	$S_y(f) = h_{-2} f^{-2}$

In addition to the random components, oscillators and clocks often show systematic, (i.e., deterministic) trends such as offsets in frequency and time, as well as linear drifts in frequency.

For the atomic clocks used in the NBS Time Scales, an adequate model is the superposition of white FM, random walk FM, and linear frequency drift for times longer than about one minute. The model has been tested on several clocks using maximum likelihood techniques for parameter estimation and the residuals have been "acceptably random." Conventional diagnostics indicate that additional model elements contribute no significant improvement to the model even at the expense of the added model complexity.

I. INTRODUCTION

Many authors (1, 2, 3) have documented the fact that most precision oscillators and clocks exhibit both random and systematic variations in their output signals. The typical random parts may include white noise

---

*Deceased

phase modulation (PM), flicker noise PM, white noise frequency modulation (FM), flicker FM, and random walk FM. A subset of these five noises is usually adequate. In addition, most oscillators also exhibit a linear drift in frequency, which is often difficult to measure.

Experimenters often diagnose the various noises using the two-sample variance (or "Allan Variance") (4,5). On occasion, they will use an estimate of the power spectral density of the frequency fluctuations (4, 5). Of course, one cannot adequately observe the fluctuations of a single clock or oscillator by itself -- one must look at the difference between two clocks. The allocation of noise levels to individual clocks requires three or more clocks of comparable quality. This allocation process does not always provide reasonable results. In fact, often the process yields negative values for the variance -- an undesirable artifact of the estimation procedure.

The Allan Variance is defined (1) as the infinite time average of sample variances based on a sample size of only two adjacent values of frequency. That is,

$$\sigma_y^2(\tau) = \lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{n} \frac{(\bar{y}_n - \bar{y}_{n-1})^2}{2} \qquad (1)$$

where $\bar{y}$ is the average frequency departure from nominal, averaged over the time interval and divided by the nominal frequency. An equivalent form of Eq. (1) is:

$$\sigma_y^2(\tau) = \lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{n} \frac{(X_n - 2X_{n-1} + X_{n-2})^2}{2\tau^2} \qquad (2)$$

where x(t) and y(t) are related by

$$y(t) = \frac{d}{dt} X(t) \qquad (3)$$

and X(t), the instantaneous time error, is related to the phase error of the oscillator by the relation:

$$X(t) = \frac{\phi(t)}{2\pi\nu_0} \qquad (4)$$

where $\phi(t)$ is the phase error and ν_0 is the nominal frequency (e.g., 5MHz).

The Allan Variance is normally computed from finite data sets of the time difference, X_n , where

$$X_n = X(n\tau_o) \qquad (5)$$

and the estimated Allan Variance is

$$\hat{\sigma}_y^2(\tau) = \frac{1}{N-2m} \sum_{n=1}^{N-2} \frac{(X_{n+2m} - 2X_{n+m} + X_n)^2}{2m^2\tau_o^2} \qquad (6)$$

where $\tau = m\tau_o$.

Although Eq. (6) is very close in form to the definition of the Allan Variance (see Eq. (1)), it is NOT an optimum estimator of the "true" Allan Variance. That is, there are other statistical techniques which provide more precise estimates of the frequency variability. These improved techniques, however, are usually valid only for very specific clock models. Fortunately, commercial cesium beam atomic clocks have been studied extensively, and good models are well documented.

II. Optimum Estimates

In the introduction, we identified two problems:

A. Statistically inefficient estimators of the level of oscillator noises and drift, and

B. Difficulties in separating individual clock performances.

While these two problems cannot be totally eliminated, they are amenable to optimal estimation techniques. That is, we can minimize their effects.

The means of estimating these parameters has been developed by R.H. Jones and P.V. Tryon (6, 7). Basically, the technique is that of maximum likelihood estimation. The technique requires an ensemble of comparable clocks ($M > 2$) and time difference data between clocks covering a significant duration (e.g., a year). With the assumptions that the perturbing noises are both independent and Gaussian, and that the basic model is adequate, then it is possible to form the likelihood function as a function of the oscillator parameters. The likelihood function is obtained from a Kalman Filter algorithm applied to the clock ensemble data.

Using essentialy the same notation as used by Gelb (8), the clock model and measurements can be expressed as follows:

$$\begin{pmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \cdot \\ \cdot \end{pmatrix}_n = \begin{pmatrix} 1 & 1 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 1 & \ldots \\ 0 & 0 & 0 & 1 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \cdot & \ldots \end{pmatrix} \begin{pmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \cdot \\ \cdot \end{pmatrix}_{n-1} + \begin{pmatrix} \varepsilon_1 \\ \eta_1 \\ \varepsilon_2 \\ \eta_2 \\ \cdot \\ \cdot \end{pmatrix}_n \qquad (7)$$

where the subscripts on the matrices denote the recursion number (i.e., time).

$$\underline{Q} = \begin{pmatrix} \sigma_{\varepsilon 1}^2 & 0 & 0 & 0 \ldots \\ 0 & \sigma_{\eta 1}^2 & 0 & 0 \ldots \\ 0 & 0 & 0 & \sigma_{\varepsilon 2}^2 \ldots \\ \cdot & \cdot & \cdot & \cdot \quad \ldots \\ \cdot & \cdot & \cdot & \cdot \quad \ldots \end{pmatrix} \qquad (8)$$

$$\underline{H} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & 0 & -1 & 0 & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ldots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ldots \end{pmatrix} \qquad (9)$$

$$\underline{R} = 0 \qquad (10)$$

where the the number of clocks is M, the state vector, $\underline{X}$, is a 2M column vector, Φ and $\underline{Q}$ are 2M by 2M square matrices, and the measurement matrix, $\underline{H}$, is M-1 by 2M, since there are only M-1 independent clock differences.

In matrix form the equations become:

$$\underline{X}_n = \underline{\Phi} \; \underline{X}_{n-1} + \underline{U}_n \qquad (11)$$

and the measurements, $\underline{Z}_n$, are:

$$\underline{Z}_n = \underline{H} \cdot \underline{X}_n \qquad (12)$$

The forecasts of $\underline{X}_n$ and $\underline{Z}_n$ to step n+1 based on data up to and including step n are:

$$\hat{\underline{X}}^-_{n+1} = \underline{\Phi} \cdot \underline{X}^+_n \qquad (13)$$

$$\hat{\underline{Z}}_{n+1} = \underline{H} \cdot \hat{\underline{X}}^-_{n+1} \qquad (14)$$

Of interest are the innovations at step n+1. The innovations are given by

$$\tilde{\underline{Z}}_{n+1} = \underline{Z}_{n+1} - \hat{\underline{Z}}_{n+1} \qquad (15)$$

with the covariance matrix $\underline{V}_{n+1}$

$$\underline{V}_{n+1} = \underline{H}' \cdot \underline{P}^-_{n+1} \cdot \underline{H} \qquad (16)$$

where $\underline{P}^-_{n+1}$ is the error covariance matrix for the state vector (see Appendix A for a brief summary of the Kalman filter relations).

Assuming that the driving noises, ε_n and η_n, are normal random deviates with zero mean, then the multivariate probability distribution can be written in the form

$$f(z_1, z_2 \ldots) = [(2\pi)^{m/2} \left| \underline{V} \right|^{\frac{1}{2}}]^{-1} \quad \exp [-\tfrac{1}{2} \sum_{n=0}^{N} \tilde{\underline{Z}}'_n \cdot \underline{V}_n^{-1} \cdot \tilde{\underline{Z}}_n] \qquad (17)$$

The function, ℓ, given by -2 times the log of the likelihood function, is

$$\ell = \sum_{n=1}^{N} \ln \underline{V}_n + \sum_{n=1}^{N} \tilde{\underline{Z}}'_n \cdot \underline{V}^{-1} \cdot \tilde{Z}_n \qquad (18)$$

Now, ℓ is an implicit function of the parameters σ^2, because both the innovations and the error covariance matrix, $\underline{P}^-_n$, are dependent on these model parameters. The estimation procedure finds that set of parameters (σ^2's) which minimizes ℓ (that is, maximizes the likelihood function). Unfortunately, ℓ is a non-linear function of the parameters and must be calculated by a complete pass through the data for each trial set of the 2M parameters. For example, if one has M=10 clocks and daily time difference data for a year, then one has 365 x (M-1) = 3285 independent measurements and 2M = 20 parameters to adjust in order to maximize the likelihood function. There exist standard computer algorithms to perform such calculations.

Three additional concerns are (a) the estimates of confidence intervals for the parameters, (b) the diagnostics to test the adequacy of the basic model assumptions, and (c) the extension of the maximum likelihood estimates to include a frequency drift parameter for each clock (9). The model adequacy can be tested by testing of the residuals ($\tilde{Z}_n$) for "whiteness" (i.e., randomness); and by comparing results to more complex model assumptions. References 6,7 include a discussion of the methods used to estimate the confidence intervals of the parameter estimates.

III. Experimental Results

For many years, the National Bureau of Standards (NBS) has accumulated large quantities of clock comparison data on the commercial cesium clocks used in the NBS time scale. We used a recent sample of time comparisons on a dozen clocks over about two months sampled every two hours. We also used another set of daily data on seven clocks over a period of one year.

The basic model assumption was that of white FM noise plus random walk FM noise plus linear frequency drift. Thus, for each clock in a data set we estimated σ_ε, σ_η, and D the drift parameter. Also estimated were the corresponding confidence intervals. The three parameters can be related to the more conventional Allan Variance through the equation (see Appendix B):

$$\hat{\sigma}_y^2(n\tau_0) = \frac{\sigma_\varepsilon^2}{n\tau_0^2} + \frac{\sigma_\eta^2(2n^2+1)}{6n\tau_0^2} + \frac{(Dn\tau_0)^2}{2} \qquad (19)$$

Figure 1 displays plots of the Allan variance obtained from the use of Eq. 19, above and the estimated parameters. Figure 2 displays a cumulative periodogram of residuals for one of the clocks. A periodogram of pure "white" noise would fall within the boundaries shown 90% of the time. On the shorter data run, (~ 2 mos.) linear frequency drift was not statistically significant. In fact, even on the longer run (1 year), only infant clocks or older clocks approaching end of life showed significant drift. (Of course, the algorithm could only detect relative drifts between clocks, not a common drift shared by all clocks.) Tests were made using more complex models, but any improvement was found to be statistically insignificant.

IV. Conclusions

A viable clock model for commercial cesium beam clocks consists of three elements:

(1) White FM
(2) Random walk FM
(3) Linear frequency drift

Maximum likelihood estimation techniques yield reasonable results and confidence intervals also. Conventional tests show the model to adequately describe observed clock behavior. Further, the technique allows one to estimate the individual performance of each clock. As pointed out by Jones, one can avoid the problem of negative variances by using a log transformation, $y = \ell n(\sigma^2)$.

Equation 19 allows one to express the results in the form of conventional Allan Variances.

The new NBS time scale algorithm (TA(NBS)) makes use of the parameter estimation routines covered in this paper. The technique is also used for NBS clock calibrations.

REFERENCES

1. Allan, D. W., "Statistics of Atomic Frequency Standards," Proc. IEEE 54, No. 2 pp. 221-236, February 1966.

2. Cutler, L. S., and Searle, C. L. "Some Aspects of the Theory and Measurement of Frequency Fluctuation in Frequency Standards," Proc. IEEE 54, No. 2 pp. 136-154, February 1966.

3. Barnes, J. A., et. al. "Characterization of Frequency Stability," IEEE Trans. on I + M Vol. IM-20, No. 2 pp. 105-120, May 1971 (Also published in NBS Tech. Note 394, April 1971.

4. Rutman, J., "Instabilité de Frequence des Oscillateurs," L'Onde Electrique 52, No. 11, pp. 480-487, Dec. 72.

5. Lesage, P., and Audoin, C., "Characterization and Measurement of Time and Frequency Stability," Radio Science, Vol. 14, No. 4 pp. 521-539, July-August, 1979.

6. Tryon, P. V. and Jones, R. H., "Estimation of Parameter in Models for Cesium Beam Atomic Clock," to be published in NBS Journal of Research, Vol. 88, 1983. Appears in the unofficial proceedings of the 2nd International Symposium on Atomic Time Scale Algorithms, National Bureau of Standards, Boulder, CO, 23-25 June 1982.

7. Jones, R. H., and Tryon, P. V., "Estimating Time from Atomic Clocks," to be published in NBS Journal of Research, Vol. 88, 1983. Appears in the unofficial proceedings of the 2nd International Symposium on Atomic Time Scale Algorithms, National Bureau of Standards, Boulder, CO, 23-25, June 1982.

8. Gelb, A., et. al., "Applied Optical Estimation," The MIT Press 1974.

9. Percival, D. B., "The U.S. Naval Observatory Clock Time Scales," IEEE Trans. on I + M, Vol. IM-27, No. 4, Dec. 1978.

APPENDIX A

SUMMARY OF DISCRETE KALMAN EQUATIONS

Model: $\underline{X}_n = \Phi \cdot \hat{\underline{X}}_{n-1} + \underline{U}_n$

Measurement: $\underline{Z}_n = \underline{H}\,\underline{X}_n + \underline{V}_n$

Forecast: $\hat{\underline{X}} = \underline{\Phi}\,\hat{\underline{X}}^{+}_{n-1}$

Error Covariance: $\underline{P}_n = \underline{\Phi} \cdot \underline{P}^{+}_{n-1} \cdot \Phi' + \underline{Q}$

Kalman Gain: $\underline{K}_n = \underline{P}_n \cdot \underline{H}' \cdot [\underline{H} \cdot \underline{P}_n \cdot \underline{H}' + \underline{R}]^{-1}$

Error Covariance: $\underline{P}_n^{+} = \underline{P}_n - \underline{K}_n \cdot \underline{H} \cdot \underline{P}_n$

State Update: $\hat{\underline{X}}_n = \hat{\underline{X}}_n + \underline{K}_n \cdot [\underline{Z}_n - \underline{H} \cdot \hat{\underline{X}}_n]$

Allan Variance

$$\sigma_y^2(m\tau_o) = E\left[\frac{(X_{n+m} - 2X_{n+m} + X_n)^2}{2m^2\tau_o^2}\right]$$

$$= E\left\{\frac{\left[-\sum_{i=n+1}^{n+m} \eta_i + \sum_{i=n+m+1}^{n+2m} \eta_i\right]^2}{2m^2\tau_o^2}\right\}$$

$$= \frac{\sigma_\eta^2}{2m^2\tau_o^2}\left[\sum_{i=1}^{m}(i-1)^2 + \sum_{i=1}^{m}(m-i+1)^2\right]$$

$$= \frac{\sigma_\eta^2}{2m^2\tau_o^2}\left[2\sum_{i=0}^{m-1}(i)^2 + m^2\right]$$

$$= \frac{\sigma_\eta^2}{2m^2\tau_o^2}\cdot\left[\frac{m(2m^2+1)}{3}\right]$$

Random Walk FM $\sigma_y^2(m\tau_o) = \sigma_\eta^2 \cdot \left(\frac{2m^2+1}{6m\tau_o^2}\right)$

Linear Frequency Drift

$$X_n = \tfrac{1}{2}D(n\tau_o)^2 \qquad \text{(Deterministic)}$$

Allan Variance

$$\sigma_y^2(m\tau_o) = \frac{(X_{n+2m} - 2X_{n+m} + X_n)^2}{2m^2\tau_o^2}$$

Allan Variance

$$\sigma^2(m\tau_o) = \frac{\frac{1}{2}D\tau_o^2[(n+2m)^2 - 2(n+m)^2 + n^2]^2}{2m^2\tau_o^2}$$

$$= \frac{[\frac{1}{2}D\tau_o^2 \cdot (2m^2)]^2}{2m^2\tau_o^2}$$

$$= \tfrac{1}{2}(Dm\tau_o)^2$$

Composite: Assumes noises statistically independent.

$$\sigma_y^2(m\tau_o) = \frac{\sigma_\varepsilon^2}{m\tau_o^2} + \sigma_\eta^2 \frac{(2m^2+1)}{6m\tau_o^2} + \tfrac{1}{2}(Dm\tau_o)^2$$

If the time error, X_n, is sampled from a continuous process, then*

$$\sigma_y^2(m\tau_o) = \frac{\sigma_\varepsilon^2}{m\tau_o^2} + \frac{\sigma_\eta^2 m\tau_o}{3\tau_o^3} + \tfrac{1}{2}(Dm\tau_o)^2$$

* Private communication C. Greenhall.

QUESTIONS AND ANSWERS

DR. STEIN:

The model which you say seems to fit most of our clocks contains random walk frequency, no flicker frequency. Therefore, much much more pessimistic view of the clocks than we have usually adopted, and I was wondering if you could comment on, for instance, what happens if you include a flicker term, do the residuals get better, or, what is the reason for rejecting the flicker frequency model?

MR. BARNES:

OK. Everybody -- I guess from my reputation, if anybody used flicker noise it would have been me, but, if you had had flicker FM present, what would happen in these models would be that the minimum value would move up and tend to have a flatter region right in the center if you really had flicker noise.

I can't go and say that, unequivocably, there is no flicker noise in these clocks, but, I can say that if you looked at the residuals for the models without flicker noise and then added flicker noise, you would find that the improvement in the whiteness of the residuals was not statistically significant.

At least, that is our experience in this data.

We tried to test the model, to see if adding other noises would significantly reduce the magnitude of the residual, and for no other model did we find significant improvement. That doesn't mean that it does not exist, but it does mean that in this sample, we were unable to observe it.

DR. WINKLER:

I just want to take this beautiful opportunity to point out that your frequency standard is an excellent one for long-term.

MR. BARNES:

I guess that's what that things says, and it doesn't have error bars on this particular graph, but I would guess that this is not surprising, because the fact that it is so bad in short-term, means that it is harder to measure reliably the long-term performance, and hence, the confidence intervals along the random walk frequency, are so broad that it allows it to almost look too good. And that, I suspect, is an artifact of any kind of analysis. Thank you.

DR. STEIN:

In this case, I would have to say that our qualitative observations from many years indicate that this analysis is probably correct.

DR. WINKLER:

Nevertheless, I think that it is important to note that one should not use short-term stability parameters for rating of clocks long-term.

MR. ALLAN:

Johnson tried, independently, to assess the ability for the clock ensemble, and the assessment has been over a year period, and a year elapsed, and another year period, so there was a great deal of difference in the stream of data -- and the parameters on this particular clock were the same, two totally independent years.