

HISTORICITY, STRENGTHS, AND WEAKNESSES OF ALLAN VARIANCES AND THEIR GENERAL APPLICATIONS

David W. Allan

President, Allan's TIME; Fountain Green, Utah 84632-0066

E-mail: david@allanstime.com

Web sites: www.allanstime.com www and ItsAboutTimeBook.com

Keywords: Allan variances, Time series analysis, Atomic clocks, Precision analysis, Non-stationary processes

Summary

Over the past 50 years, variances have been developed for characterizing the instabilities in precision clocks and oscillators. These instabilities are often modeled by non-stationary processes, and these variances have been shown to be well-behaved and to be unbiased, efficient descriptors of these processes. The time-domain and frequency-domain relationships are shown along with the strengths and weaknesses of these characterization metrics. These variances are also shown to be useful elsewhere, as in navigation.

Introduction

Nature gives us many non-stationary and chaotic processes. If we can properly characterize these processes, then we can use optimal procedures for estimation, smoothing, and prediction. During the 1960s through the 1980s, the Allan variance, the modified Allan variance, and the Time variance were developed to this end for the timing and the telecommunication communities. Since that time, useful refining techniques have been developed. This activity has been a learning endeavor, and the strengths and weaknesses of these variances will be enumerated herein. The applicability of these variances has been recognized in other areas of metrology as well, because the above processes are ubiquitous. Knowing the strengths and weaknesses is important not only in time and frequency but so that these variances may be properly utilized in other application areas, such as navigation.

Prior to the 1960s and before atomic clocks were commercially available, quartz-crystal oscillators were used for timekeeping. The greatest long-term-frequency instabilities in these oscillators were their frequency drifts. Also, it was commonly recognized that their long-term performance seemed to be modeled by what is commonly called flicker-noise frequency modulation (FM), which model is a non-stationary process, because this noise has a power-spectral-density proportional to $1/f$, where f is the Fourier frequency. In integrating this kind of noise to determine the classical variance, one observes that the integral is non-convergent.

In 1964, James A. Barnes developed a generalized auto-correlation function that was well behaved for flicker noise. I was fortunate to have him for my mentor at the National Bureau of Standards (NBS) in Boulder, Colorado. That same year, the IEEE and NASA held a special conference at NASA, Goddard, in Beltsville, Maryland, addressing the problem of how to characterize clocks with these non-stationary behaviors. Jim and I presented a paper at this conference, and it was well received. His work was the basis for his Ph.D. thesis, and it also gave me critical information that I needed for my master's thesis. We both finished our theses the following year. In addition to Jim's work, I relied heavily on the book that Jim had shown me by Sir James Michael Lighthill, *Fourier Analysis and Generalized Functions*. Along with Jim's work, this book was invaluable.

In my thesis I studied the effects on the classical variance as a function of how long the frequency was averaged (the averaging time, τ), how many samples were included in the variance, N , how much dead-time there was between frequency averages, $T-\tau$ (in those days it took time for a frequency counter to reset after a frequency had been measured over some interval τ ; so T was the time between the beginning of one measurement to the beginning of the next), and how it depended on the measurement system bandwidth, f_h . We developed a set of spectral-density, power-law noise models that covered the characterization of the different kinds of instabilities we were observing in clocks – resulting from the noise of the measurement systems, the clocks, and from environmental influences. Since then, we have observed that these noise models are much more general than we'd originally thought and have a broad application in metrology.

Both Jim's and my theses were published, along with several other papers from the 1964 IEEE/NASA conference, in a February 1966 special issue of the Proceedings of the IEEE on "Frequency Stability."

Modeling nature with power-law noise processes

The pioneering work of Mandelbrot and Voss introducing "fractals" shows the importance of these self-similar and non-stationary processes in modeling nature. Flicker noise is in that class. We found that five different kinds of noise were useful in modeling clocks. Many of these may be used as good models in other natural processes – including errors in navigation systems.

Modeling the noise processes in nature is revealing. The better we can model nature, the better we can use optimization to know more about the underlying processes masked by nature's noise.

We have been able to use the variances I will share in this paper in characterizing and modeling many different processes in nature. As I look back over the 50 years we have been doing this work, it has been rewarding to see the insights into nature that have been gained. I will show some exciting examples of these insights later in this paper.

For clocks, if the free-running frequency of a clock is $v(t)$ and we denote its nominal frequency as v_o , then we may write the normalized frequency deviation of a clock as $y(t) = (v(t) - v_o) / v_o$. The time-deviation of a clock may be written as $x(t)$, which is the integral of $y(t)$. Studying the time-domain and frequency-domain characteristics of $x(t)$ and $y(t)$ opens the opportunity to model the clock's behavior and then to perform optimum estimation, smoothing, and prediction of its "true" behavior in the midst of noise – even when the noise is non-stationary.

We symbolize the frequency-domain measures using spectral densities – denoted by $S_y(f)$ and $S_x(f)$. In the time domain we have found useful the Allan variance (AVAR), the modified Allan variance (MVAR), and the Time variance (TVAR). Other variances have been found useful as well. Often shown are the square-root of these variances:

• <u>SQUARE</u>	<u>SQUARE ROOT</u>
• $\sigma_y^2(\tau) = \text{AVAR}$	$\sigma_y(\tau) = \text{ADEV}$
• $\text{mod. } \sigma_y^2(\tau) = \text{MVAR}$	$\text{mod. } \sigma_y(\tau) = \text{MDEV}$
• $\sigma_x^2(\tau) = \text{TVAR}$	$\sigma_x(\tau) = \text{TDEV}$

Figure 1. Common nomenclature for the variances and their square-roots as used at the National Bureau of Standards (now National Institute of Standards and Technology) in the United States of America as well as in international scientific literature and as IEEE standards.

The power-law spectral densities may be represented as $S_y(f) \sim f^\alpha$ and $S_x(f) \sim f^\beta$, and because x is the integral of y , one may show that $\alpha = \beta + 2$. The models for the random variations for clocks, their measurement systems, and for their distribution systems that work well have values of alpha as follows: $\alpha = -2, -1, 0, +1, \text{ and } +2$. These models seem to reasonably fit the random frequency variations observed. These models seem to fit in many other areas of metrology as well. Flicker noise has been shown to be ubiquitous in nature. In the case of time and frequency, we have observed both flicker-noise FM ($\alpha = -1$) and flicker-noise PM ($\beta = -1$).

Figure 2 demonstrates how these models apply for different kinds of clocks. Typically, the noise model changes from short-term averaging times to long-term – almost always moving toward more negative values of α . Included in the following chart is the value $\alpha = -3$, as this is the long-term model for earth-rotation noise for Fourier frequencies below one cycle per year after subtracting all the systematic terms from the data.

**Given: $S_y(f) \sim f^\alpha$; $S_x(f) \sim f^\beta$;
and $\beta = \alpha - 2$**

Power-law spectra models

α	β	Earth	Qu	H-m	Cs	Rb	Optical
+2	0		•	•			
+1	-1		•				
0	-2		•	•	•	•	•
-1	-3	•	•	•	•	•	•
-2	-4		•	•	•	•	•
-3	-5	•					

Figure 2 Matrix showing the usefulness of power-law, spectral-density models for Earth = noise in the earth's rotation rate (after removing all systematics), in Qu = quartz-crystal oscillators, H-m = hydrogen masers, Cs = cesium-beam and cesium-fountain frequency standards, Rb = rubidium-gas-cell frequency standards, and in the new and most stable atomic clocks using frequencies in the optical region of the electromagnetic spectrum.

As one can see in the next figure, the visual appearance of these power-law spectra are very different, and the eye, in some sense, can be a good spectrum analyzer. One of the many reasons why in data analysis one should always visually look at the data is that the brain is an amazing and miraculous processor.

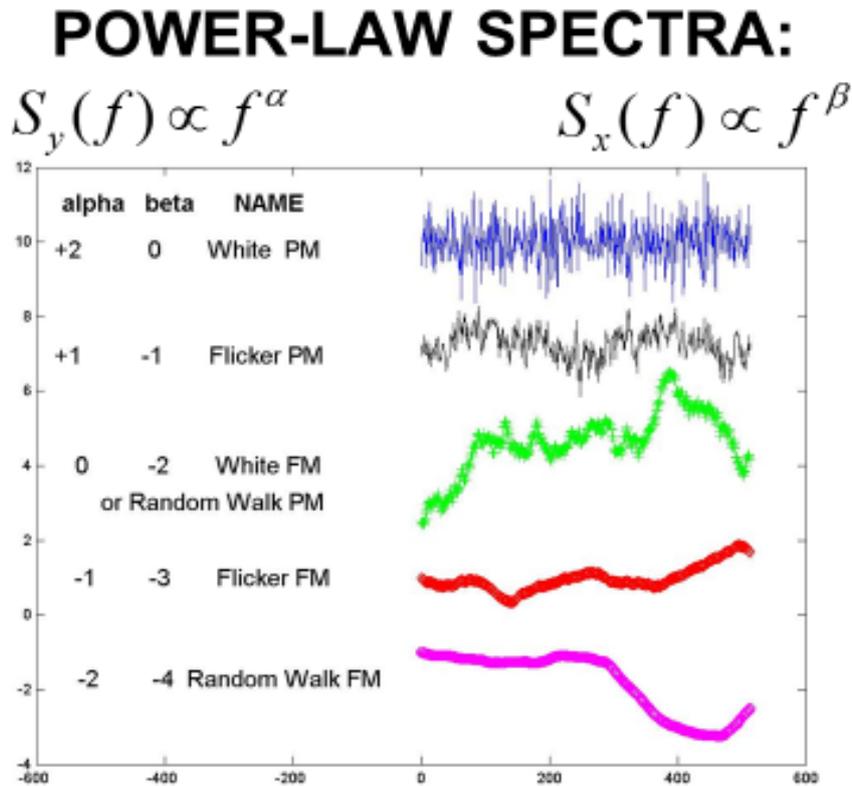


Figure 3. Illustration of visual difference for different power-law, spectral-density models.

Using Lighthill's book, we can transform these spectra to the time domain. In doing so we obtain figure 4.

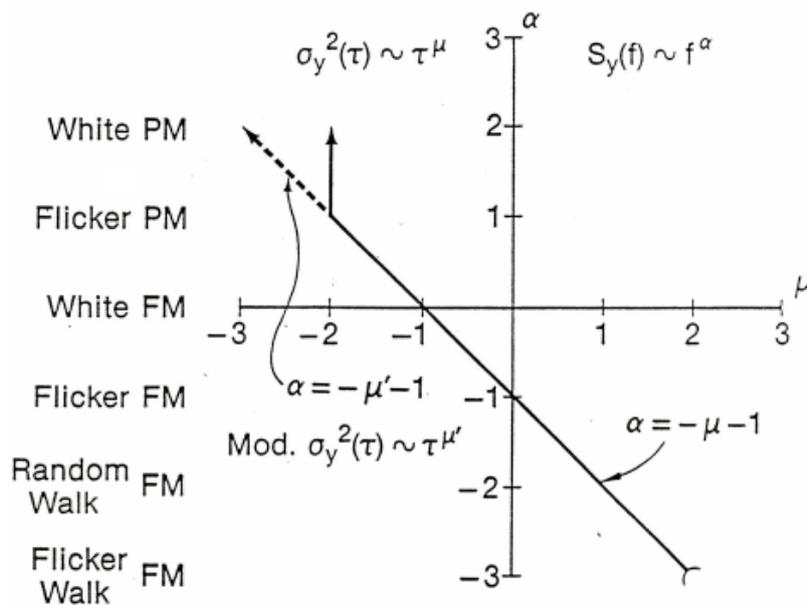


Figure 4. We have α as the ordinate and μ as the abscissa, where μ is the exponent on τ showing the time-domain dependence, and where $AVAR = \sigma_y^2(\tau)$ and $MVAR = \text{mod. } \sigma_y^2(\tau)$. We have an elegant Fourier transform relationship in the simple equation $\alpha = -\mu - 1$; we jokingly call it the super-fast Fourier transform, because the AVAR can be computed very quickly from an equally spaced set of data.

Since $\sigma_y^2(\tau) \sim \tau^\mu$, by plotting $\log \sigma_y(\tau)$ versus $\log \tau$, the slope will be $\mu/2$; hence, we can ascertain both the kind of noise as well as its level from such a plot. This sigma-tau plotting technique has been used literally thousands of times to great advantage – giving a quick “super-fast Fourier transform” of the data.

In Figure 4, we notice an ambiguity problem for AVAR at $\mu = -2$. The simple equation no longer applies, and we cannot tell the difference in the time domain between white-noise phase or time modulation (PM) and flicker-noise PM. This problem was a significant limitation in clock characterization for the time and frequency community for 16 years after AVAR was developed. Even though there was ambiguity in the τ dependence in this region, we knew that it could be resolved because there remained a measurement bandwidth sensitivity. Since it was inconvenient to modulate the measurement system bandwidth, this approach never became useful. But in 1981 we discovered a way to modulate the bandwidth in the software, and this was the breakthrough we needed. This gave birth to MVAR, and the concept is illustrated in the following figure.

One can think of software bandwidth modulation in the following way. There is always a finite measurement system bandwidth. We call it the hardware bandwidth, f_h . Let $\tau_h = 1/f_h$. Then every time we take a phase or time reading from the data, it inherently has a τ_h sample-time window. If we average n of these samples, we have increased the sample-time window using software by n , $\tau_s = n\tau_h$. Let $\tau_s = 1/f_s$, then if we increase the number of samples averaged as we increase τ , then one can show that we are decreasing the software bandwidth by $1/n$. We were able to show that by modulating the bandwidth in this way we removed the above ambiguity and maintained validity for our simple super-fast Fourier transform equation over all the power-law noise processes of interest; $\alpha = -\mu' - 1$. There is an unknown proportionality constant between the f_s shown below and the f_s in the above equations, but fortunately we don't need to know it to characterize the data.

Figure 5 is an illustration of this software bandwidth modulation for $n = 4$; in principle, n can take on any integer value from 1 to $N/3$.

In 1981 we learned how to modulate the bandwidth in the software;
fundamental breakthrough

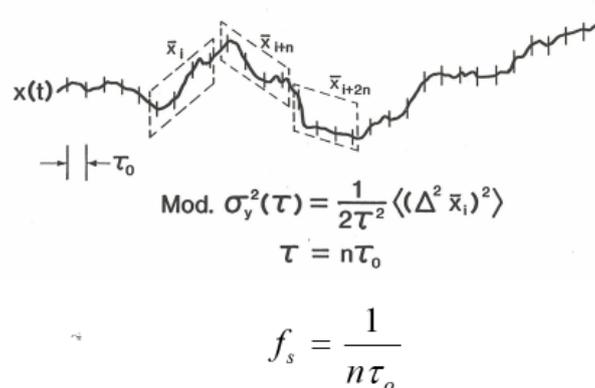


Figure 5. A pictorial of the software-bandwidth modulation technique used in the modified Allan variance to resolve the ambiguity problem at $\mu = -2$; Hence, this software modulation technique allows us to characterize all of the power-law spectral density models from $\alpha = -3$ to $\alpha = +2$. This covers the range of useful noise models for most clocks. Illustrated in this figure is the case for $n = 4$; n may take on values from 1 to $N/3$, where N is the total number of data points in the data set with a spacing of τ_0 .

Data length dependent variances are not useful

Going back to 1964, Dr. Barnes had shown that the second and third finite-difference operators on the time variations of a clock gave a convergent statistic in the presence of flicker noise FM. This was the basis of his PhD thesis in helping to use a quartz-crystal oscillator ensemble calibrated by the National Bureau of Standards primary cesium-beam-frequency standard to construct a time scale for generating time for NBS and hence for the USA civil sector; the USNO is the official time reference for the USA defense sector.

I had shown in my master's thesis the divergence of the classical variance or lack thereof for the above power-law noise processes as a function of the number of data points taken. The degree of divergence depends upon both the number of data points in the set as well as upon μ the kind of noise. In other words, the classical

variance was data-length dependent for all of the power-law noise models we were using to characterize clocks except for classical-white noise FM. Hence, the classical variance was deemed not to be useful in characterizing atomic clocks because other than white-noise FM models were needed. This divergence problem seems to exist in all areas of metrology as a result of nature’s natural processes and environmental influences on whatever we are measuring.

I used the two-sample variance as a normalizing factor because I knew from Lighthill and from Barnes’ work that it was convergent and well behaved for all of the interesting power-law spectral density processes that are useful in modeling clocks and measurement systems. The two-sample variance I used may be written as follows:

$$\sigma_y^2(\tau) = \frac{1}{2} \langle (\Delta y)^2 \rangle = \frac{1}{2\tau^2} \langle (\Delta^2 x)^2 \rangle,$$

where the brackets $\langle \rangle$ denote the expectation value, or ensemble average and the “2” in the denominator normalizes it to be equal to the classical variance in the case of classical white-noise FM. Don Halford, my Section chief at the time, named this the Allan variance, and the name persists. I don’t mind; jokingly, some ask if I am at variance with the world? When one takes the square root and it becomes the Allan deviation, I cringed a bit, but then as I thought about it, I said to myself, “I am not a deviant!” Deviation is the measure of performance – the change in a clock’s rate – the smaller the better. If I can help these be smaller and smaller, that is good and will help society, and I am all for that.

The ratio of the N-sample variance to the Allan variance as a function of N is shown in the figure 6. Realizing that the N-sample variance is the classical variance for N samples, one sees why it is not useful for characterizing these different kinds of noise, as it is not convergent in many cases and is biased as a function of N in all cases except for classical-white noise. One can turn this dependence to an advantage and use it to characterize the kind of noise using the B1 bias function: $B_1(N) = \sigma^2(N) / \sigma_y^2(\tau_0)$.

**This ratio is called the bias function, $B_1(N)$;
can be used to ascertain kind of noise:**

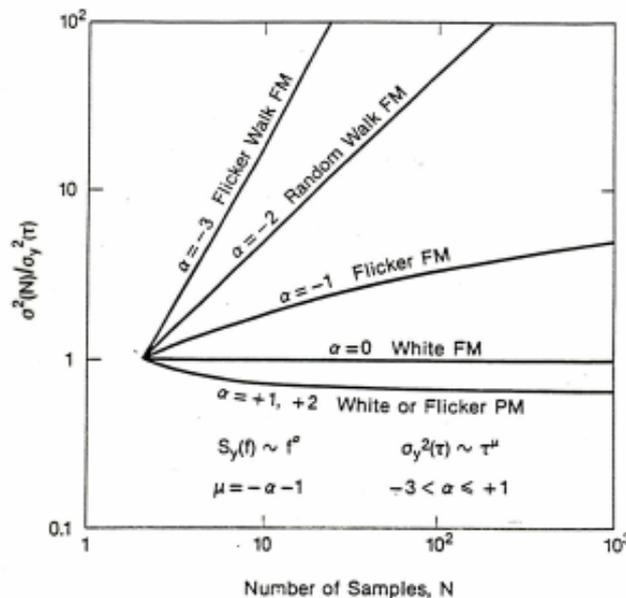


Figure 6. Illustration of the data-length dependence of the classical variance for the different kinds of power-law noise processes used in modeling precision oscillators and atomic clocks as a function of the data length.

Following the 1966 IEEE special issue on “Frequency Stability,” the IEEE asked Dr. Barnes to chair a panel of experts and to prepare a special paper on “Characterization of Frequency Stability.” That paper was published in 1971 in which they recommended the spectral density $S_y(f)$ and the two-sample variance as the recommended measures of frequency stability. They also called it the “Allan variance.” This paper is available on the NIST Time and Frequency Divisions web site: <http://tf.boulder.nist.gov/general/pdf/118.pdf> Dr. Leonard S. Cutler, who was one of these experts, was the first to write the equation for the time-domain variances in terms of the spectral density, and this is developed in this paper for, equation 23.

As a point of interest, many years ago I was asked to write a paper entitled, “Should the Classical Variance Be Used as a Basic Measure in Standards Metrology?” <http://tf.boulder.nist.gov/general/pdf/776.pdf> . I researched voltage standards and gage-blocks, and I found flicker-noise behavior in their long-term performance. A fundamental statement that came out of that research was that if the bias function $B_1(N)$ is not 1 (one) within some reasonable confidence limits, then the classical variance is not a good measure. That advice was not followed by the BIPM standard’s committee even though it has a solid scientific basis. Traditions seem too strong many times even when these traditions are not the best for progress when these flawed traditions continue to be followed.

Note also that the two-sample or Allan variance is without dead-time. In other words, the frequency measurements are sequentially adjacent. For example, the i^{th} frequency deviation taken over an averaging time, τ , may be derived from the time deviations as follows: $y_i = (x_i - x_{i-1})/\tau$. This equation gives us the true average frequency deviation over that interval; it may not be the optimum estimate of frequency. One notices that if the average is taken over the whole data set, then all the intermediate values cancel, and one is left with the true average frequency deviation over the data set: $y_{\text{avg}} = (x_N - x_0)/N\tau$. This is one of the benefits of no-dead-time data. Another is that for classical white-noise FM, as has been found to be the fundamental performance limitation in most atomic clocks, then $\sigma_y^2(\tau)$ is an optimum-variance estimator of the change of frequency over any averaging time, τ , and is equal to the classical variance for – the minimum data-spacing variance.

Dr. Barnes has also shown that $\sigma_y^2(\tau)$ is an unbiased estimator for the level of the power-law noise process of interest in modeling atomic clocks and that it is Chi-squared distributed. The value of τ in the software analysis can take on values for all $\tau = n \tau_0$ for any integer $n = 1$ to $N/2$. The confidence of the estimate is best at $\tau = \tau_0$ decreasing to $\tau = N/2$, where there is only one-degree of freedom for the confidence of the estimate and the Chi-squared-distribution function has a most probable value of zero for one degree of freedom. Even though it is unbiased, the probability of small values is significant. In a $\sigma_y^2(\tau)$ versus τ plot, one often observes too-low of values for $\sigma_y^2(\tau)$ as the value of τ approaches half the data length; then the degrees of freedom are too small for a good confidence on the estimate. David A. Howe and his group at NIST have addressed this problem and have come up with some elegant solutions by adding degrees of freedom to the long-term data; that work is still in progress and is extremely useful. <http://tf.boulder.nist.gov/tf-cgi/showpubs.pl>

In other areas of metrology, one needs to pay attention to this dead-time issue if the noise is not white (random and uncorrelated). As shown in my thesis, the dead-time has an impact on the resulting variance if the noise is not white. The dead-time problem was studied and subsequent papers written. The following link in Chapter 8 of Monograph 140 covers this issue for both the N dependence and dead-time with the bias functions B_1 and B_2 , respectively: <http://tf.boulder.nist.gov/general/pdf/771.pdf> Later, I will show this need not be a significant problem in general metrology applications; it seems to be a unique problem in time and frequency.

The time variance

In the later part of the 1980s the telecom industry in the United States came to me and asked if I would help them with a metric for characterizing telecommunication networks. I asked Dr. Marc Weiss, who was in my group at NIST at the time, to help me with this project. It was a fascinating work, as we analyzed a lot of their data to find the best metric. Out of this work we developed the time variance, TVAR. It is defined as follows: $\text{TVAR} = \tau^2 \text{MVAR}/3$. The “3” in the denominator normalizes it to be equal to the classical variance in the case of classical white-noise PM. Like as for AVAR for FM, one can show that for white-noise PM, TVAR is an optimum estimator of change in the phase or time residuals in a variance sense.

The work in the United States caught on and these three variances became international IEEE time-domain measurement standards in 1988. Interestingly, in their historic application we see that these three variances had three general regions of applicability:

1. AVAR for characterizing the performance of frequency standards and clocks.
2. MVAR for characterizing the performance of time and frequency distribution systems.
3. TVAR for characterizing the timing errors in telecommunication networks.

Following the development of each of these three variances, many other areas of applicability have arisen. Conveniently, TDEV (the square-root of TVAR) has no dead-time issues and has become a standard metric in the international telecommunications industry. All three have application capability in many other areas of

metrology. Navigation system errors and gyro errors are some examples. If you search Google for “Allan variance,” you will find about 50,000 results.

Equations and their transforms

The equations for computing AVAR, MVAR, and TVAR from the time-deviations and for N data points are respectively:

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2(N-2n)} \sum_{i=1}^{N-2n} (x_{i+2n} - 2x_{i+n} + x_i)^2,$$

$$\text{mod.}\sigma_y^2(\tau) = \frac{1}{2\tau^2n^2(N-3n+1)} \sum_{j=1}^{N-3n+1} \left(\sum_{i=j}^{n+j-1} (x_{1+2m} - 2x_{i+n} + x_i) \right)^2,$$

$$\sigma_x^2(\tau) = \frac{1}{6n^2(N-3n+1)} \sum_{j=1}^{N-3n+1} \left(\sum_{i=j}^{n+j-1} (x_{1+2m} - 2x_{i+n} + x_i) \right)^2,$$

where the x_i are the time deviation data separated by a time interval, τ_0 , and $\tau = n\tau_0$.

For MVAR and TVAR, the computation involves a double sum. One may think that this could cause the computation time to increase as N^2 , but one can employ some computation tricks, such as simple drop-add averaging, to make it linear. Otherwise this could be a problem for large data sets. Such tricks have been successfully implemented, and the software references cited later include these computation techniques.

The following equations show how the three time-domain variances may be derived from frequency-domain information. One cannot do the reverse – derive the spectral densities from time-domain analysis. When possible, it is often very useful to analyze the data in both the frequency and time domains. Below we see the frequency-domain view of these variances.

Translation Between Frequency and Time Domains

$$\tau = n\tau_0$$

AVAR:

$$\sigma_y^2(\tau) = \int_0^\infty 2 \left[\frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2} \right] S_y(f) df$$

MVAR:

$$\text{Mod } \sigma_y^2(\tau) = \int_0^\infty 2 \left[\frac{\sin^3(\pi f \tau)}{(n\pi f \tau) \sin(\pi f \tau_0)} \right]^2 S_y(f) df$$

TVAR:

$$\sigma_x^2(\tau) = \frac{8}{3n^2} \int_0^\infty \left[\frac{\sin^3(\pi f \tau)}{\sin(\pi f \tau_0)} \right]^2 S_x(f) df$$

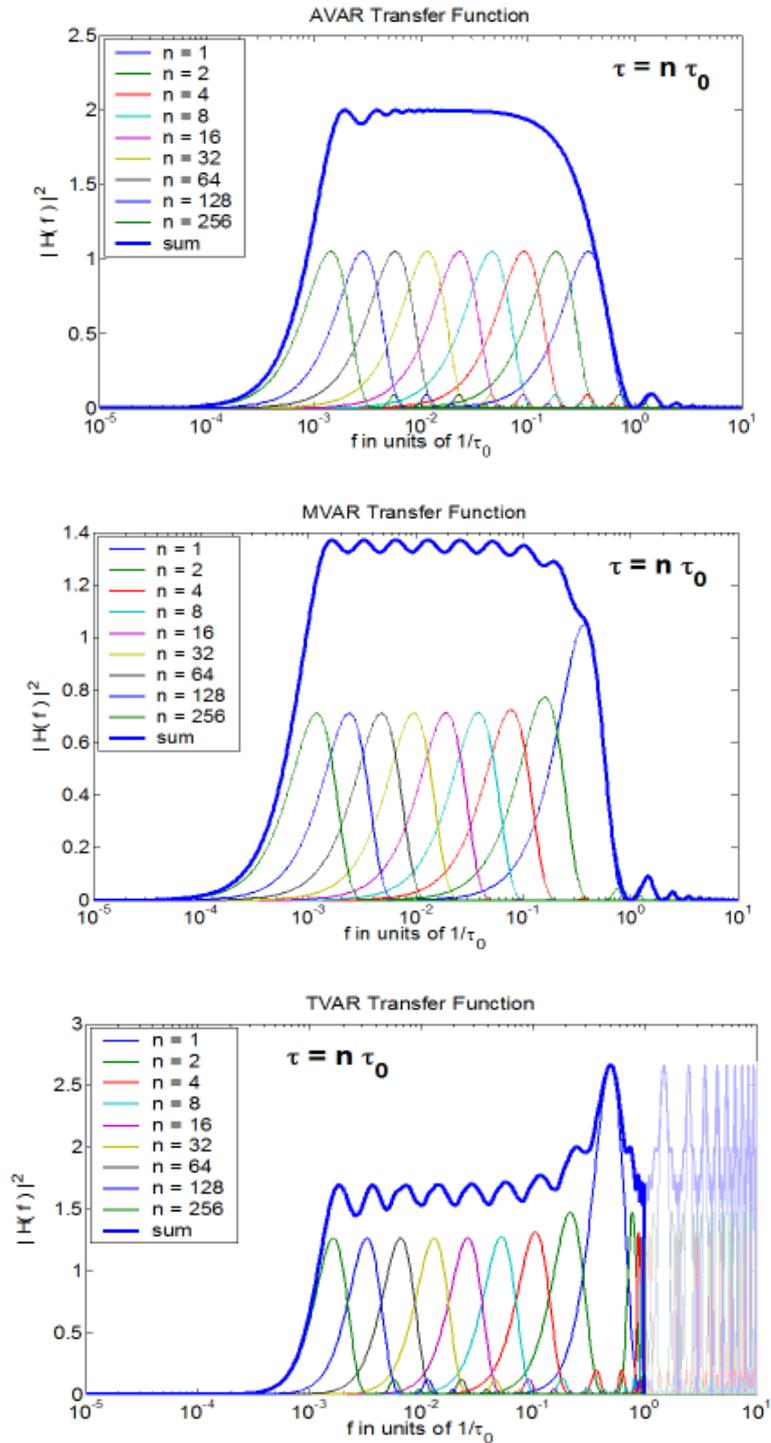


Figure 7 a, b, c, and d. Figure 7a shows the three time-domain equations as derived from spectral densities. Figures b, c, and d show the effective Fourier windows using the transfer functions of each of these three variances for $n = 1, 2, 4, 8, 16, 32, 64, 128, \text{ and } 256$.

Three years ago, I was asked to write a paper on “Conversion of Frequency Stability Measures from the Time domain to the Frequency domain, vice-versa and Power-law Spectral Densities.” This paper is available on our web site, and has a lot more detail about these conversion processes: http://www.allanstime.com/Publications/DWA/Conversion_from_Allan_variance_to_Spectral_Densities.pdf The conversion relationships are shown in the following table for the five noise types:

NOISE TYPE	$S_y(f)$	$S_x(f)$
White PM	$\frac{(2\pi)^2}{3f_h} [\tau^2 \sigma_y^2(\tau)] f^2$	$\frac{1}{\tau_0 f_h} [\tau \sigma_x^2(\tau)] f^0$
Flicker PM	$\frac{(2\pi)^2}{A} [\tau^2 \sigma_y^2(\tau)] f^1$	$\frac{3}{3.37} [\tau^0 \sigma_x^2(\tau)] f^{-1}$
White FM	$2 [\tau^1 \sigma_y^2(\tau)] f^0$	$\frac{12}{(2\pi)^2} [\tau^{-1} \sigma_x^2(\tau)] f^{-2}$
Flicker FM	$\frac{1}{2\theta\tau^2} [\tau^0 \sigma_y^2(\tau)] f^{-1}$	$\frac{20}{(2\pi)^2 9\theta\tau^2} [\tau^{-2} \sigma_x^2(\tau)] f^{-3}$
Random Walk FM	$\frac{6}{(2\pi)^2} [\tau^{-1} \sigma_y^2(\tau)] f^{-2}$	$\frac{240}{(2\pi)^4 11} [\tau^{-3} \sigma_x^2(\tau)] f^{-4}$

$$A = 1.038 + 3 \ln(2\pi f_h \tau)$$

Estimation, smoothing, and prediction

■ Estimation and smoothing

Box and Jenkins in their book, *Time Series Analysis*, using the ARIMA process, do a great work on how to estimate and smooth for various kinds of random processes. I will not review their paramount work here.

There is a simple, powerful and useful statistical theorem that I will use for estimation, smoothing, and prediction. It is that the optimum estimate of the mean of a process with a white-noise spectrum is the simple mean. As examples, if we have white-noise PM, then the optimum estimate of the phase or the time is the simple mean of the independent phase or time residual readings added to the systmatics.

If we have white-noise FM, then the optimum estimate of the frequency is the simple mean of the independent frequency readings, which is equivalent to the last time reading minus the first time reading divided by the data length, if there is no dead-time between the frequency measurements. As we have shown before, the true average frequency is given by: $y_{avg} = (x_N - x_0)/N\tau$.

■ Prediction

Using the above theorem for optimum prediction, if we take the current time as “t,” and we wish to predict ahead an interval τ , then the optimum time prediction, for a clock having white-noise FM and an average offset frequency y_{avg} given by the above equation, is given by: $\hat{x}(t + \tau) = x_N + \tau y_{avg}$. A simple pictorial for the optimum time prediction using this theorem for the five different noise processes is shown in the figure 8; for white-noise FM, the y_{avg} is assumed to be zero.

The even-powered exponents are directly amenable to this theorem, but the flicker-noise (odd exponents) are more complicated. However, there is a simple prediction algorithm for flicker FM using what we call the second difference predictor. It is very close to optimum and is simple. If you desire to predict τ into the future then this prediction can be obtained by the following equation: $\hat{x}(t + \tau) = 2x(t) - x(t - \tau)$, where t is the current time. I have seen this equation used on the stock market, which is often flicker-like in its performance.

Knowing the stability, $\sigma_y(\tau)$, of a clock allows us to calculate its time predictability capability. As an approximate rule of thumb, the predictability is given by $\tau \sigma_y(\tau)$, Using this equation the figure 9 shows the time predictability of a variety of timing devices that have been used over human history.

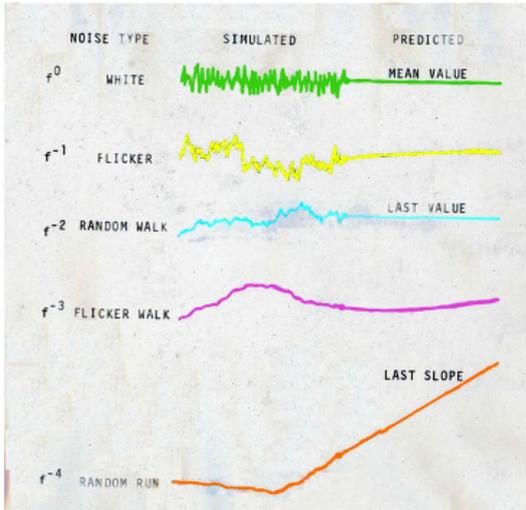


Figure 8. A pictorial illustrating optimum prediction for the five different power-law noise processes used in modeling the time deviations in precision clocks. These prediction algorithms have general application.

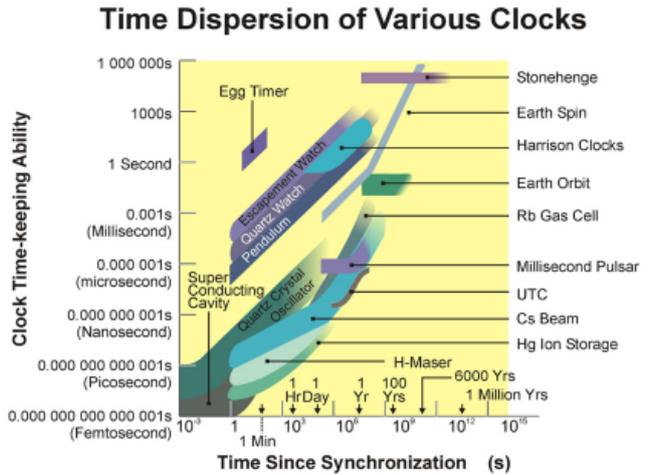


Figure 9. This chart was made back in 1997. If we were to include the ytterbium clock on this graph, it would be represented by a $\tau^{-1/2}$ line crossing through 40 femtoseconds at seven hours; or it would be about 100 times better than the best clocks shown here.

A similar plot could be made for the navigation community showing the position dispersion rate for various navigation devices. This may be a useful tool to see which technologies could be brought together in combination to make a significant improvement in both the short-term and long-term performance.

Systematics

A good model for time deviations in a clock is: $x(t) = x_0 + y_0 t + \frac{1}{2} D t^2 + \epsilon(t)$, where x_0 and y_0 are, respectively, the synchronization error and syntonization error at $t = 0$, D is the frequency drift, and $\epsilon(t)$ represents the remaining random errors on top of the first three systematic error terms. It is good to subtract the systematics from the data so that the random effects can be viewed visually and then analyzed with better insights. Much can often be learned by this approach.

If frequency drift is present in a clock, and it usually is, then it affects AVAR, MVAR, and TVAR in the following way:

$$\sigma_y(\tau) = \text{mod.} \sigma_y(\tau) = \frac{D\tau}{\sqrt{2}} \quad \text{and} \quad \sigma_x(\tau) = \frac{D\tau^2}{\sqrt{6}} .$$

An example of the effect of frequency drift on an ADEV plot is shown in the next figure.

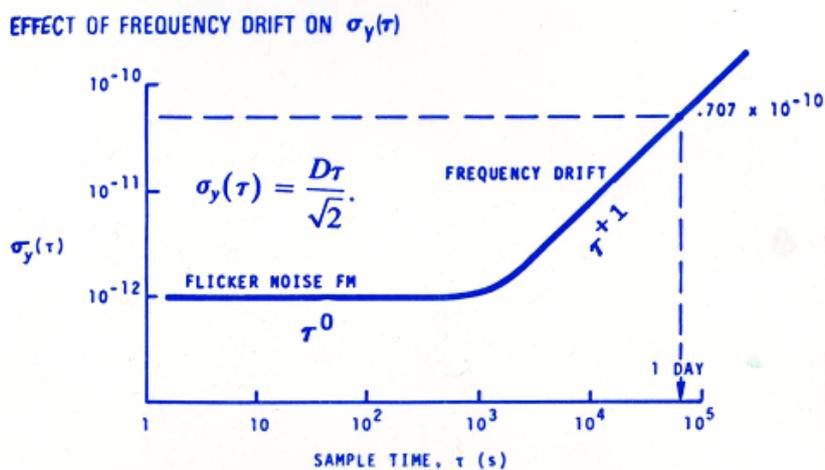


Figure 10. Illustration of the effects of frequency drift on an ADEV plot.

If there is frequency drift, the values of $\sigma_y(\tau)$, in that region where the drift is affecting the plot, will lie tightly on the τ^{-1} line. If there is random noise present then the values will not fit tightly to this line. If there is a frequency modulation, f_m , present in the data then it also has a systematic effect on the analysis in the following way for ADEV: $\sigma_y(\tau) = \frac{x_{pp}}{\tau} \sin^2(\pi f_m \tau)$, where x_{pp} is the peak-to-peak amplitude of the modulation. The following figure shows the effect on an ADEV plot.

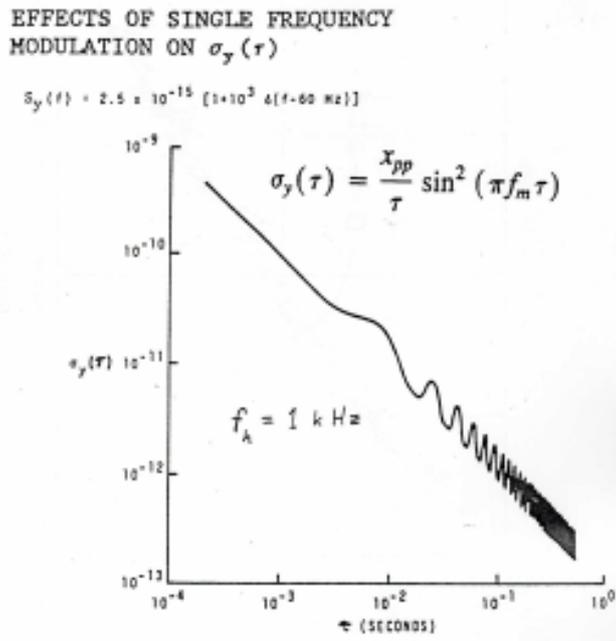


Figure 11. ADEV with frequency modulation, f_m , present on the data.

Both MVAR and TVAR are affected as well. A plot of the effect on TDEV is shown in the next figure.

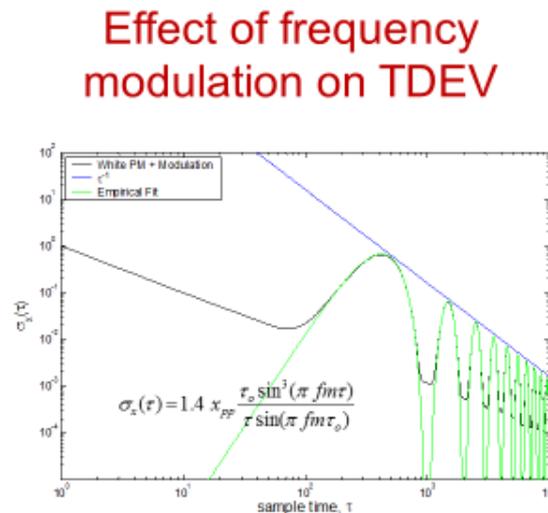


Figure 12. The effect of frequency modulation on a TDEV plot when that modulation is on top of the signal and noise. The white-noise PM causes the $\tau^{-1/2}$ behavior in the plot. Notice the modulation averages down as τ^{-1} . The equation fitting the effects of the modulation is empirical.

At $\tau = n/f_m$, the effect of the modulation is aliased away, where n is any positive integer. Recognizing this null effect allows these three variances to be used as low-frequency spectrum analysis techniques for bright Fourier-frequency lines in the data. It is my experience that low-frequency spectral lines can often be observed using this null approach in the time domain better than can be observed in the frequency domain. We will find this effect very useful in the EXAMPLES section of the paper, which is next.

Examples and some application opportunities of these variances

■ Clocks of the 1960s and 70s

In the following figure we have a sigma-tau plot of the frequency instabilities between a precision free-running, quartz-crystal oscillator and a commercial cesium-beam atomic clock. One sees for sample times, τ , shorter than one second a τ^{-1} behavior due to the measurement noise. This plot was made before MDEV was developed, so we are not sure of the noise type because of the ambiguity problem with ADEV for this slope. I have observed the same ambiguity problem in some of the navigation stability plots that I have seen. Whenever a τ^{-1} behavior occurs in an ADEV plot, one should then analyze the data using MDEV to hopefully resolve the ambiguity regarding the kind of noise modulation present in the data.

The rise in the value of $\sigma_y(\tau)$ as the sample or averaging time approaches 10 seconds is due to the attack time of the cesium-beam locking its quartz-crystal-slave-oscillator to the cesium resonance. Over the next decade we see a $\tau^{-1/2}$ behavior; or $\mu = -1$ which means $\alpha = 0$ from our simple "super-fast-Fourier" transform relationship, and this is classical white-noise being measured for this cesium-beam atomic clock. For the longest averaging times we see a τ^0 behavior, which then corresponds to $\alpha = -1$, and this is due to the flicker-noise FM of the precision, quartz-crystal oscillator. Even with the ADEV ambiguity problem, we were delighted in the 1960s and 70s to be able to characterize so easily the noise type and level of the clocks then being used for timekeeping for the USA.

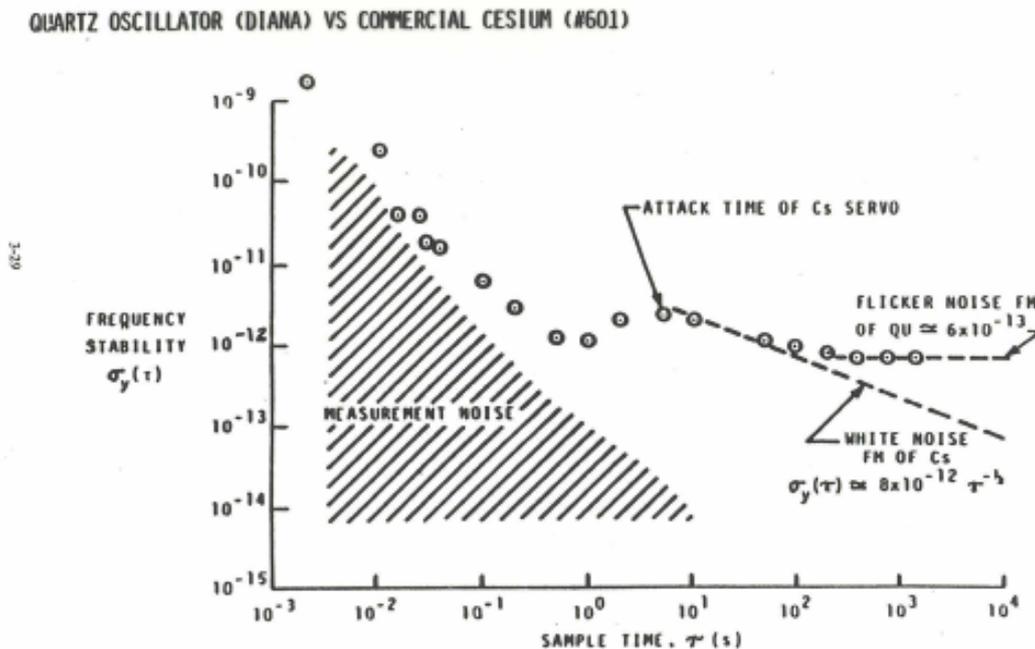


Figure 13. An ADEV plot for a precision, quartz-crystal oscillator versus a commercial cesium-beam atomic clock.

In 1965, we had a very interesting clock comparison at NBS in Boulder, Colorado. Bob Vessot brought his hydrogen maser from Boston, Massachusetts. Harry Peters brought his hydrogen maser from NASA Goddard, Beltsville, Maryland. Len Cutler brought his Hewlett Packard commercial-cesium-beam, atomic clock from Palo Alto, CA. We had the NBS primary frequency standard and data acquisition systems. This grand-clock-comparison effort resulted in an interesting 12 author paper. <http://tf.boulder.nist.gov/general/pdf/172.pdf>.

It was my responsibility at that time to provide the NBS reference time-scale for comparing all of these clocks. Up to this point, Jim's algorithm had been generating time for NBS and for the civil-sector of the USA. With Jim's ever-present help, I wrote a new time-scale algorithm, AT-1. With several refinements by Tom Parker and Judah Levine since that time, that algorithm is still generating time for the USA today. This time-scale algorithm was a major application of the "Allan variance" and generated a near real-time software clock with the following optimization features:

- Its software-clock output can be shown to be better than the best clock providing input;
- Even the worst clock enhances the output;
- If a clock misbehaves, it is rejected and not used – avoiding unnecessary perturbations;
- Each clock gets an optimum weighting factor for inclusion in the time computation;

- The weights are adaptive so that if a clock improves over time, its weight increases and vice versa;
- The optimum time of each clock as well as the optimum estimate of the frequency of each clock are estimated at each measurement cycle;
- Both the short-term as well as the long-term stability of the software ensemble output are optimized;
- And it is able to deal with white-noise FM, flicker-noise FM, and random-walk FM, which are the kinds of noise processes that well model the atomic clocks being used;

Originally, I used a PDP-8 computer and had eight clocks in the ensemble; AT-1 had 94 lines of code and provided error messages. I had to use some variables three times to not exceed the available logic limit. It is interesting to watch this algorithm's performance, because it is almost as if it is alive as it breathes with each clock's behavior. AT-1 has been generating the official time for the USA for nearly 50 years.

■ New optical clock stabilities using total ADEV

As we look at some of the exciting new optical clocks, the following ADEV plot is from data taken at NIST in Boulder, Colorado comparing two ytterbium optical lattice clocks in 2013. This plot utilizes the "Total ADEV" approach developed by David A. Howe and his group, which gives optimum confidence on the long-term stability estimates for ADEV. Long-term data are extremely valuable, so this "Total ADEV" technique adds greatly to the information one is able to learn from the data.

Here we see the best stability ever observed to that date of $\sigma_y(\tau = 25,000 \text{ seconds}) = 1.6 \times 10^{-18}$. This is like an error of 50 ps in a year. A picosecond is a million-millionth of a second (10^{-12} s); 50 ps is the time it takes light to travel 1.5 cm. This is 20 times better than the nanosecond accuracy that GPS needs and they have to upload their GPS corrections at least once a day. In this plot we see the nearly ideal atomic-clock ($\tau^{-1/2}$) white-noise FM behavior over about four decades of averaging time at a remarkable level of 3.2×10^{-16} at 1 second.

World record for stability NIST ytterbium optical lattice clock

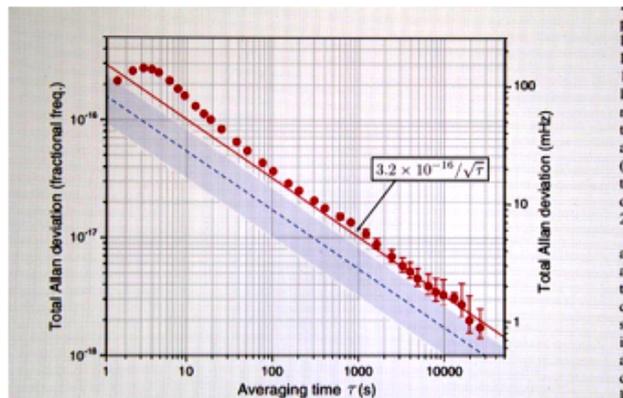


Figure 14. Comparison of two ytterbium optical-lattice atomic clocks operating at 518 295 836 591 600 Hz

Millisecond pulsar timing using MDEV

Going back to 1982, the first millisecond pulsar was discovered by Donald Backer, Shri Kulkarni, Carl Heiles, Michael Davis, and Miller Goss. Its name PSR B1937+21 is derived from the word "pulsar" and the declination and right ascension at which it is located, with the "B" indicating that the coordinates are for the 1950.0 epoch. This pulsar had the best astronomical timing performance of anything ever observed. I read their paper and was intrigued. I could see some ways we could help them, so I made contact with Dr. Michael Davis, who was the scientist in charge at the Arecibo Observatory where the data were being taken. Mike invited me down, and in 1984 I installed a GPS common-view receiver to tie their clock, which was making the pulsar measurements, to the world's best atomic clocks.

One can see in the next figure, the very complicated system for this millisecond pulsar measurement and the nominal behavior in each link of the measurement system chain.

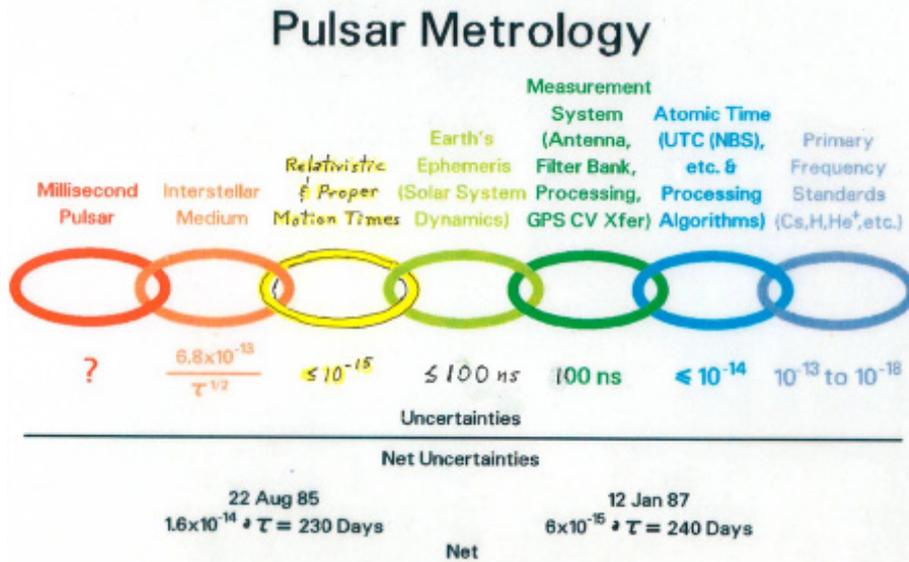


Figure 15. Measurement system for the millisecond pulsar. Note the improvement of nearly a factor of three in the stability measurements from 1985 to 1987 indicated at the bottom of the figure due to the help we were able to give them.

The next figure is an excellent example of the advantages of the modified Allan variance. As I studied the data, I was able to observe random-walk (f^2 spectrum) in the delay between two different observation frequencies for the pulsar. They had assumed that the electron content along the path was constant. This result showed that it was not, and when the $1/f^2$ ion content correction was applied the random-walk effect was suppressed leaving white-noise PM residuals as is shown in the next figure using MDEV to show this benefit.

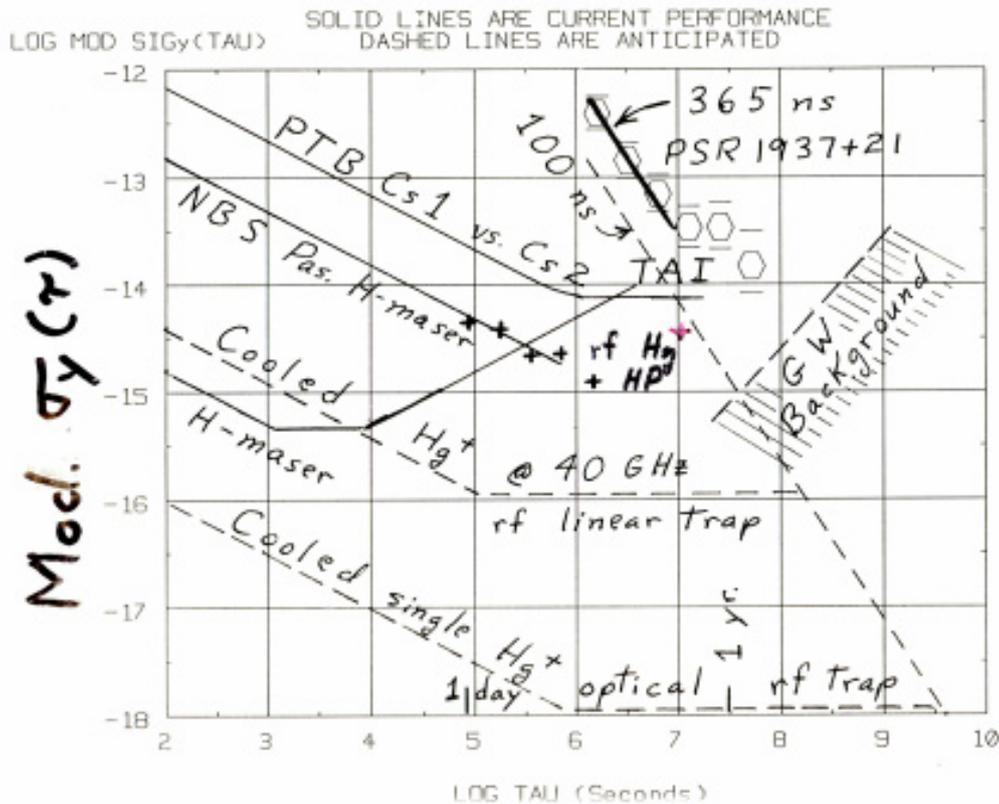


Figure 16. The PTB is the primary frequency standards laboratory for Germany. "Pas" stands for passive hydrogen maser. H-maser is for an active hydrogen maser. For shorter than 100-second averaging times, active hydrogen masers often exhibit white-noise PM, which is a $\tau^{-3/2}$ slope like the slope of the measurement noise of the pulsar. PSR 1937+21 is located about $1/7^{\text{th}}$ of the way across our galaxy, and the hope then was that we might see gravitational-wave perturbations in the path indicated by "GW Background."

Over the next several years, they did a 25-million dollar upgrade on the Arecibo telescope to bring about several major improvements and hoping to move the measurement noise down to the indicated 100 ns level.

In about 1990, I shared the following frequency stability plot at a UC Berkley, California, millisecond pulsar workshop. As these new fast spinning pulsars were thought to be competitive with atomic clocks, there is a very fundamental message in this plot. Even if they achieve the 100 ns white-noise PM measurement noise level with their upgrade, the data need to be averaged for about 200 years to reach a stability level of 10^{-18} , which is about where the best clocks are now. In other words, you would get one data point every 200 years at the 10^{-18} level – clearly not a competitive clock. I had no comments from my expert millisecond pulsar colleagues in the audience! One may further note that the ytterbium stability has, as of two years ago, surpassed by a factor of three that shown in the figure for the best anticipated performance of the cooled single Hg ion.

■ Opportunity for improving GPS accuracy

The GPS satellites (SVs) orbit the earth at about 4.2 earth radii. This distance creates a significant geometry problem for determining the vertical distance of the satellites from the center of the earth because the tracking stations' vectors are too close to parallel. GLONASS solves this problem by using retro-reflectors on the satellites and by doing round-trip-laser ranging from a ground station of known position to each of the satellites. This they can do to about 5-cm accuracy, which is about 12 times better than GPS.

Kepler's third-law has built into it the needed orthonality to solve this vertical-distance problem:

$T^2 = \frac{4\pi^2}{GM} r^3$, where T is the orbit period, G is the universal gravitational constant, M is the mass of the earth, and r is the radius from the center of the earth to the satellite. Since the orbit is tangential to the radius vector, if we can determine the point of closest approach to a tracking station of known coordinates, then we have the orthogonal information we need to determine the radius vector. The Doppler shift of an SV's clock will go to zero at its point of closest approach with respect to its tracking station. This zero-Doppler shift gives us a precise marker in its orbit period, T, with an uncertainty δT . From the above equation, we can derive the uncertainty in the radius vector: $\delta r = \frac{2}{3} r \frac{\delta T}{T}$. With current high-performance atomic clocks and using MDEV to assure that

the residuals are white-noise PM, so that all the systematics have been properly removed, then our calculations indicate that δr can be made to be less than a centimeter. The value of δT can be made very small because with white-noise PM being the limiting measurement noise, its value decreases as the data-length to the minus 3/2s power for the viewing-time of the satellite's pass.

There are some important contingencies associated with making this equation work properly. Professor Neil Ashby, who did the relativity equations for GPS, and I worked on modeling this approach in the 1990s and got some excellent results. Clocks have gotten significantly better since then, and the requirement for a zero-g environment as was done for Gravity Probe-B is now more readily available. There are some other contingencies, but the advantages are enormous; being an all-weather system is one. One of the biggest disadvantages is that this approach is a major change in system architecture, but these changes could be done in a meaningful step-wise process.

I felt that this-high accuracy technique was far enough along that I sent a letter to the GPS Headquarters folks for their consideration. I describe this some in Chapter 20 of my book, and the details may be found in Appendix K of the book's web site, www.ItsAboutTimeBook.com.

■ New unified field theory results validated using ADEV

Starting in 1999, we were working to understand a new concept in relationship to the UFT. This concept is explained in papers on our web site: www.AllansTIME.com/UFT_private and in Chapter 21 of my book. The book has exciting new information that has never been published before. The concept is that what we call diallel-field lines can carry all four of the force fields plus much more and connect everything to everything. The book by Lynne McTaggart, *The Field*, describes many experiments consistent with this new UFT.

We first did experiments to show the existence of these diallel-field-lines, and then that they had quantum states. Whereas the quantum states in atoms or molecules are generally thought of as being spherical or elliptical, they are nominally cylindrical in the diallel-field structure. In the fall of 2000 we observed for the first time quantum-transition emissions from these diallel-field-lines. This and some of the other experiments were performed in the laser physics lab at BYU. They kindly let me use it as an alumnus. All these experiments are described in Chapter 21 of my book. We have done seven experiments to date validating this new UFT diallel-field-line theory.

The following figure illustrates the diallel-field-line coupling of the planets to the sun and their effect on the sun-spot activity. I used 100 years of sunspot data and I analyzed it using ADEV as shown in the following figure. As predicted by this new UFT, we were delighted to find the periods of all the major planets present in the sun-spot data except for Uranus, where our 100 year data length was insufficient. I also used the masses, the magnetic fields, and the planet's orientations in space to see if I could nominally duplicate the sun-spot activity over the last 100 years. I had about an 80% correlation coefficient fit to the data. There is still a lot to be learned. We feel like infants exploring the biggest forest in the world.

Diallel-field line coupling of the planets to the sun shown in sunspot activity

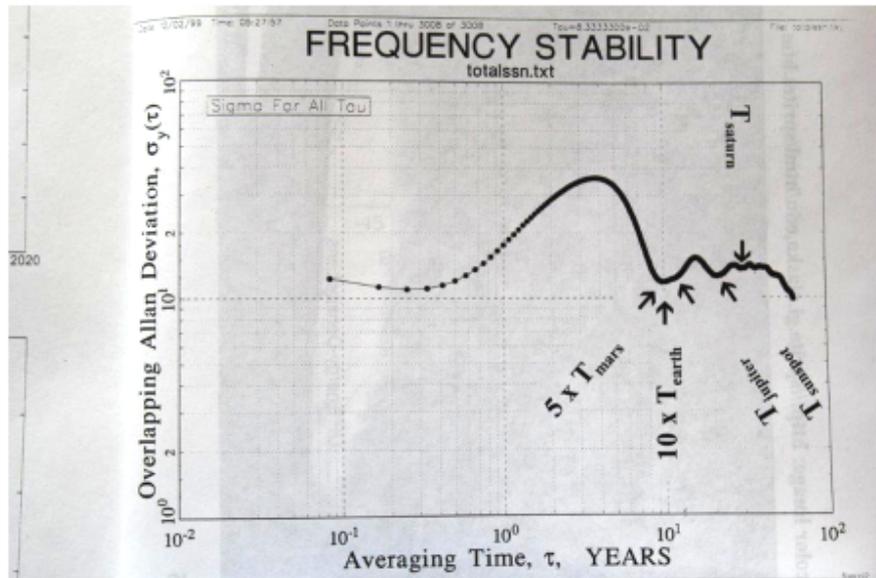


Figure 17. ADEV plot of 100 years of sun-spot data showing the periods of all the major planets except for Uranus, where our data length was too short to resolve its period.

Conclusions

It has been fifty years since I finished my master's thesis. With fifty years of experience in the time and frequency community, the use and improvements upon the Allan variance have matured significantly. I observe a similar maturing in its application in navigation scenarios and in other areas of metrology.

As I review the literature in this regard, there are three technical areas where I will make suggestions that I feel will be of most help in this maturing process.

First, for 16 years we lived with the ambiguity problem with the Allan variance when its square-root (ADEV) behaves as τ^{-1} – not being able to distinguish between white-noise phase modulation (PM) and flicker-noise PM. That ambiguity problem was resolved in 1981 with the development of the Modified Allan variance, which allowed us to modulate the bandwidth in the software. The quantization errors in integrated navigation systems have a white-noise spectrum; the use of MDEV would be very useful here. I have looked at several navigation papers by some of the best experts in the field using the Allan variance, and I have found very few using MDEV. In the case of quantization errors, MDEV allows to average the noise down as $\tau^{-3/2}$, which then allows the observation of other noise types and instability problems more quickly. MDEV is also the optimum averaging technique for such errors. In this same regard, if one desires to know the rate with white-noise residuals present, then a linear regression on the slope improves the knowledge of the slope as $N^{-3/2}$, where N is the number of data points in the regression analysis and is the optimum estimate of the slope for exactly the same reason that MDEV improves as $\tau^{-3/2}$. I am happy to see that it is beginning to be introduced by some navigation researchers.

Second, after the quantization errors are averaged down, ADEV works well and is an efficient metric for characterizing the intermediate and longer-term instabilities. But because the Allan variance is Chi-squared distributed, when the degrees of freedom get too small for the longest averaging times available from the data, then the ADEV values are often too small. This problem has in large measure been solved by David A. Howe

and his group at NIST, Boulder, Colorado, with “Total ADEV” and useful variations thereto (see the references). Their work, in a clever way, adds the needed degrees of freedom. Their work could be applied usefully to navigation error analysis as well. Long-term data are often very expensive to acquire, so Total ADEV and its cousins give a more efficient use of the data.

And Third, systematic errors are often hard to deal with. We have found it generally useful over the years to subtract the systematics from the data, as much as is reasonable, before analyzing the residuals for their noise characteristics. This practice is usually done after the fact, but can be done in real-time with proper filter functions and prediction algorithms that will estimate and remove the systematics. Since optimum estimation procedures depend upon the kind of noise, this problem can be solved recursively or from some prior knowledge of the noise characteristics of a given system. The principle of parsimony dictates that we use the simplest and most efficient metric in our noise analysis. ADEV satisfies that requirement in many areas of metrology, and I believe that is the main reason for its becoming as widely used as it is and ever growing in different areas of metrology. If the systematics are not subtracted from the data before the noise analysis, these systematics often adversely affect the long-term values in an ADEV plot. For this reason, the GPS program people went from Allan variance to the Hadamard variance, which is a third difference operator on the time residuals and is not sensitive to the systematic-frequency drift that plagues the performance of rubidium-atomic clocks used in the GPS satellites. The Hadamard variance is not parsimonious, and they would have better noise analysis confidence estimates by estimating and removing the frequency drift, and then use ADEV and Total ADEV (or similar) to analyze the noise characteristics of the residuals – from which they deduce their Kalman filter parameters for optimizing the GPS performance.

References

There are a very large number of publications about the three variances developed in this paper. I refer you to the NIST Time and Frequency Division publication web site for many of these. And I refer you to the 1988 IEEE Standard 1139-1988: *Standard Terminology for Fundamental Frequency and Time Metrology*, to the 1990 NIST Technical Note 1337, *Characterization of Clocks and Oscillators*, to the 1997 ITU HANDBOOK: “Selection and Use of Precise Frequency and Time Systems,” to the 1997 Hewlett Packard Application Note 1289, *The Science of Timekeeping*, to the 2000-2014 additional variance work at NIST giving additional degrees of freedom and providing efficiency & tighter confidences on the variance estimates – <http://tf.nist.gov/general/publications.htm>, to the “Handbook of Frequency Stability Analysis,” by W. J. Riley, NIST Special Publication, SP 1065 (2007), also available at www.wiley.com, and to Chapter 20 of my 2014 book www.ItsAboutTimeBook.com; Appendix J at this link is also the booklet *The Science of Timekeeping*. I also refer you to “publications” on my other web site: www.allanstime.com. The above publications by Bill Riley provide analysis software, which include the advantages of Total ADEV, etc. For this reason and others, his software is in common use in the time and frequency community.

As mentioned before, if you search using Google for “Allan variance,” there are about 50 thousand results. If you add navigation to that, there are about 3 thousand results. In reviewing some of the 3 thousand I found some very interesting papers. Though I have been around the navigation community and am a Fellow of the Institute of Navigation, I do not consider myself an expert in the literature of this community. I suggest here a few papers, which I found in my search, which I thought were outstanding: *Analysis and Modeling of Inertial Sensors Using Allan Variance* by El-Sheimy, N., Calgary University; Haiying Hou; Xiaoji Niu; *Allan Variance Analysis on Error Characters of MEMS Inertial Sensors for an FPGA-based GPS/INS System* by Xin Zhang, Yong Li, Peter Mumford, Chris Rizos; *School of Surveying and Spatial Information Systems University of New South Wales, Australia*; at the following link <http://www.vectornav.com/support/library/gyroscope> is a fascinating paper on using ADEV to measure gyroscope instabilities; *Allan Variance Analysis on Error Characters of Lowcost MEMS Accelerometer MMA8451Q* by Marin Marinov*, Zhivo Petrov* (*Aviation Faculty, NVU), V. Levski”, and Dolna Mitropolia, Bulgaria; *Modeling Inertial Sensors Errors Using Allan Variance*, http://www.ucalgary.ca/engo_webdocs/NES/04.20201.HaiyingHou.pdf (URL: <http://www.geomatics.ucalgary.ca/links/GradTheses.html>) by Haiying Hou September 2004; Department of Geomatics Engineering; *A Comparison between Different Error Modeling of MEMS Applied to GPS/INS Integrated Systems* by Alex G. Quinchia (Barcelona, Spain), Gianluca Falco (Torino, Italy), Emanuela Falletti (Torino, Italy), Fabio Dosis (Torino, Italy), and Carles Ferrer (Barcelona, Spain); *Notes on Stochastic Errors of Low Cost MEMS Inertial Units*, Yigiter Yuksel & Huseyin Burak Kaygisiz; *Two Methods for the Determination of Inertial Sensor Parameters*, Vladimir Vukmirica*, Ivana Trajkovski*, Nada Asanović*, *Military Technical Institute (VTI), Ratka Resanovića, Belgrade, Serbia; and *Modified Allan Variance Analysis on Random Errors of MINS* by Bin Fang and Xiaoqi Guo, TELKOMNIKA, Vol.11, No.3, March 2013, pp. 1227 ~ 1235 e-ISSN: 2087-278X. Even though these references are excellent resources in my opinion several of them suffer from the

ambiguity problem in ADEV when it behaves as τ^{-1} for the quantization noise problem. MDEV is a better metric in this case, as I have cited before.

Because $1/f$ noise and fractals are so ubiquitous in modeling nature, we expect non-stationary analysis techniques – like in the family of Allan variances – to be useful as efficient time-series analysis metrics. The usage seems to be growing, but there are many areas where these metrics seem to be unknown statistical tools. In my own research, I have shown these variances to be useful in analyzing the stability of gage blocks and volt standards. Richard F. Voss has demonstrated $1/f$ noise in a large variety of music. Musha and Higuchi have identified $1/f$ noise in traffic flow. The height of the River Nile at flood stage over the last some thousands of years for which there are data has a $1/f$ spectral density. Such noise is found in economics, psychology, and in neurons. Pink noise is another name for $1/f$ or flicker-noise. You will find a fascinating article in Wikipedia on “Pink Noise” – showing its ubiquitous nature – and a large number of references are given there.

As a fun health example, since neuron noise is $1/f$, if you were to stand on one foot and then map the motion of the top of your head, the time series would be a flicker-noise process. If now you get on a bicycle and ride it to follow a straight line, since you have to integrate when riding a bike to maintain balance, the front tire deviations from the straight line will be an f^3 spectral-density process. With a controlled set of parameters, this bicycle balancing activity could be used – using ADEV to analyze the deviations – in a very simple way to assess improvement or degradation in your balance. Since I am an avid mountain bike rider, I am observing this phenomenon a lot – especially on a narrow deer trail on a steep slope in the mountains near our home.