Measurement Methods and Algorithms for Comparison of Local and Remote Clocks*

Judah Levine
Time and Frequency Division
and
Joint Institute for Laboratory Astrophysics
National Institute of Standards and Technology
Boulder, Colorado 80303

Abstract

We will discuss several methods for characterizing the performance of clocks with special emphasis on using calibration information that is acquired via an unreliable or noisy channel. We will discuss time–domain variance estimators and frequency–domain techniques such as cross–spectral analysis. Each of these methods has advantages and limitations that we will illustrate using data obtained via GPS, ACTS and other methods. No one technique will be optimum for all of these analyses, and some of these problems cannot be completely characterized by any of the techniques we will discuss.

We will also discuss the inverse problem of communicating frequency and time corrections to a real–time steered clock. Methods have been developed to mitigate the disastrous problems of data corruption and loss of computer control.

Introduction

Measuring the time or frequency of a clock inevitably involves transmitting the clock signal through a channel of some sort. The channel may consist of nothing more than a measurement system if the clock is nearby, while the channel for a remote clock is likely to be much more complex. In either case it is important to characterize the performance of the channel and to remove its effects if possible. This is quite difficult to do in general; we will discuss methods that are useful in several important special cases. None of these methods is optimal in all situations.

Differential Comparisons

One of the simplest methods of separating the contributions of the channel and the clock is to observe the same clock through two nominally independent channels. Figure 1 shows the difference between two channels, each of which is measuring the time of a single cesium standard with respect to a reference oscillator. Since the input signals are identical, the difference should consist primarily of a constant value that depends only on the differential delays in the cables and the measurement hardware, and a mean value has been removed from the data set to account for this. The residual fluctuations are channel noise. If the channels are independent and identical, we can estimate the channel performance by assigning 50% of the remaining variance to each channel.

We may characterize the data of Fig. 1 using the standard two–sample (Allan) variance. The magnitude of the Allan variance is 6×10^{-16} at 2 hours, and it decreases approximately as $1/\sqrt{\tau}$ for longer times (Fig. 2), suggesting that the difference between the two measurement systems can be modeled as white frequency noise. Although this is not impossible, the response of the hardware is more likely to be approximated by white phase noise, which would fall off more rapidly with averaging time. If we assume a white phase noise model for the channel, then the observed excess power at long periods must be the response of the channel to some other signal. The more detailed analysis below shows that the measurement systems are in fact responding to fluctuations in ambient temperature.

The air temperature in the vicinity of the measurement hardware is shown in Fig. 3; there is a clear qualitative correlation between these data and the data of Fig 1. This temperature sensitivity may be quantitatively estimated using correlation analysis in either the time or the frequency domains.

Analysis in the Time Domain

The simplest assumption is that the measurement hardware responds linearly to fluctuations in the ambient temperature, possibly with some time lag. If $R(t)$ is the perturbation in the measurement at time t when the ambient temperature differs from its long–term average value by $T(t)$, then we estimate $R(t)$ by

$$\mathbf{\hat{R}(t) = C_0 T(t) + C_1 T(t-\tau) + \ldots + C_k T(t-k\tau).} \quad (1)$$

The parameter τ is the time interval between measurements of both R and T. The C_k coefficients in eq. 1 are usually called admittances. They are generally not linearly independent of each other since the temperature series itself has a non–zero auto–correlation at finite lag. As a result, it is usually sufficient to use only a single term on the right–hand side of eq. 1. This choice may also be necessary to achieve numerical stability in the solution if the auto–correlation function of the temperature varies only slowly with lag time (which is the usual situation). We can estimate both k and C_k in the simple 1–term case by finding the value of k for which the cross–correlation function $< R(t)T(t-k) >$ is an extremum. The value of the cross-correlation for this value of k divided by $< T(t)T(t) >$, the variance in the temperature series itself, is then an estimate of C_k. The sharpness of this extremum depends on $< T(t)T(t-k) >$, the auto–correlation of the temperature function, and on how well eq. 1 models the variance in the measurement channel.

We have computed the cross correlation between R and T (data in Figs. 1 and 3), and the result is shown in Fig. 4 as a function of lag $k\tau$, where $\tau = 2$ hours. We find a clear extremum at about $k = 2$; the normalized admittance at that lag is -6.59 ps/°C. As we should expect, the extremum is rather broad because the temperature changes slowly in time and peaks at a non–zero lag because of the thermal inertia of the measuring hardware. We can use eq. 1 to model the temperature–induced variance in Fig. 1 and to remove the temperature–dependence from the data. This operation reduces the Allan variance by a factor of 3 and whitens the time difference data by removing much of the longer–period structure. The Allan variance now decreases more rapidly than $1/\sqrt{\tau}$, suggesting that we are approaching the underlying white phase noise of the channel. There is still some residual long–period structure in the measurements, however, and we can do better.

Analysis in the Frequency Domain

The constant time–delay used in eq. 1 is phenomenologically useful, but does not accurately model the actual system. We inserted the delay to model the thermal inertia of the hardware, but this inertia does more than just introduce a time delay—it also acts as an integrator of the fluctuations in the ambient temperature. This integration acts as a low–pass filter. In addition, the correlation analysis does not recognize that both of the time series have high–frequency uncorrelated noise which nevertheless contributes to the computation. Both of these considerations imply that the actual admittances are likely to decrease at higher Fourier frequencies (shorter period perturbations). This dependence is not incorporated into eq. 1, which estimates a frequency–independent average admittance. The admittance estimated using that model is therefore likely to be too small at long periods where the temperature fluctuations are significant and too large at short periods where the data are largely noise due to other causes.

The simplest way of incorporating these considerations into the model is to assume a linear frequency–dependent admittance. If R' and T' are the Fourier Transforms of R and T, respectively, then the admittances will be estimated to satisfy

$$\mathbf{R}'(\mathbf{f}) = \mathbf{C}'(\mathbf{f})\mathbf{T}'(\mathbf{f}), \tag{2}$$

where $C'(f)$ is the admittance as a function of frequency. The admittance at each frequency may be complex to incorporate both a magnitude and a phase shift; the phase shift is the frequency–domain analog of the time delay in eq. 1.

If R' and T' are computed using standard Fast Fourier Transform methods, then each will have n degrees of freedom, where n is the number of points in the time–domain functions R and T. (These frequency–domain degrees of freedom are normally assigned to $n/2$ amplitudes and $n/2$ phases, but other assignments are possible.) Equation 2 can then be solved for $n/2$ complex admittances, all of which will be approximately linearly independent of each other:

$$\mathbf{C}'(\mathbf{f}) = \frac{\mathbf{R}'(\mathbf{f})}{\mathbf{T}'(\mathbf{f})} \tag{3}$$

This process would reduce eq. 2 to an identity, but it would result in an admittance estimate that was not physically reasonable, since $C'(f)$ should not vary rapidly with frequency. We can introduce this constraint by averaging $C'(f)$ both in frequency and in time. The frequency averaging is motivated by the fact that the hardware and its surroundings are a non–resonant thermal system and cannot have a response that is a rapidly–varying function of frequency. The time–averaging recognizes that the admittance to temperature should depend on the mechanical and electrical design of the system and its surroundings and should therefore be time–independent. (This latter condition may not be true if the system is also sensitive to the spatial gradient of the temperature. This quantity may vary in time even if the temperature does not.)

It is possible to satisfy both of these averaging criteria simultaneously by breaking the time series into blocks and by averaging the admittances computed in each block. The blocks are usually chosen to be consecutive, non-overlapping subsets of the data. The frequency resolution of a Fourier transform is inversely proportional to the length of the input time series, so that shorter blocks implies a wider bandwidth for each estimate and hence greater frequency averaging. In addition, as the length of each block decreases the number of blocks in a given data set increases, thereby increasing the number of admittance estimates that are averaged to yield the final value. The admittances that make up each average are computed using data from different times, resulting in an averaging of the admittances over the time period of the full data set.

Fig. 5 shows the coherence between the two data sets as a function of frequency. The coherence is simply the average of $C'(f)$ computed as above and then normalized by the average power in the time series at frequency f. Each block contained 64 measurements (128 hours). If $C'(f) = 1$ then the two series are perfectly coherent at that frequency, which means that there is a constant relationship between the two amplitudes and phases. A measurement of the amplitude and phase in one series can be used to construct a perfect prediction of the corresponding parameters in the other. Note that $C'(f) = 0.82$ at very low frequencies, so that most of the observed variance in the time difference data at low frequencies is due to temperature fluctuations and could be removed using the temperature observations. The rather surprising peak in the coherence at 1.25 c/day is not a resonance — it is due to the fact that the temperature spectrum itself has a peak in its power spectrum at that frequency, perhaps because of harmonics generated by the interaction of the underlying diurnal temperature cycle with the slower weekday/weekend cycle.

The admittance analysis shows that the effects of temperature are largest at periods of several hours and longer, and it is possible to remove these effects from the data by low-pass filtering the temperature data and then subtracting them from the time-difference measurements with a time shift of a few hours as indicated in the previous analysis. The coherence analysis is used to estimate the characteristics of the filter function and the scale factor that should be applied to the low-passed data.

Discussion

Each of the preceding analyses modeled the data differently. A two-sample Allan variance computation, for example, is most useful if the power spectrum of the variance in the data set can be approximated by a polynomial in frequency. The correlation analysis models the relationship using a constant time-delay independent of frequency, and is usually appropriate only if the data to be examined are band-limited signals with appreciable broad-band coherence. The frequency-domain analysis is the most general of the three approaches, but it may produce numerically unstable estimates without lots of averaging. In addition, the coherence estimates are not normally distributed if the "true" coherence is small — the estimates tend to be too large when the actual coherence is less than about 0.55. While all describe the same data set, the interpretations are quite different in each case.

A common assumption in all of the analyses is that the characteristics of the data set are time invariant, so that the entire series can be described by a single set of global parameters. These methods will not be useful for modeling isolated transmission errors, for example, and problems of this type must be treated with some form of time–dependent analysis. Kalman techniques and moving–average models are well–known methods of detecting these problems, but all of these methods tend to require surprisingly large quantities of data to function reliably. Since there is no reliable means of predicting when a glitch will occur, the measurement interval is usually driven by how long a large glitch can be allowed to remain undetected rather than by the more usual statistical considerations derived from an Allan variance computation. This problem is more difficult than the choice between Type I and Type II errors in conventional statistical analyses because the magnitude of a channel error is not governed by a statistical distribution and is at least potentially unbounded.

Another situation that is difficult to model using the techniques we have described is a channel that has significant multi–path effects, such as a radio channel or a wide–area computer network. There is a lower bound but essentially no upper bound to the channel delay in both of these situations, so that it is not the mean but the minimum value of a group of measurements that is an invariant of the system and hence represents the "correct" delay. The mean is unbounded in principle, but it will always be too large even if some artificial bound is enforced. The channel delay can be estimated by comparing measurement data received via several independent routes, but it is often quite difficult to construct several independent channels to the same system since the local hardware tends to be common to all of these channels.

Application to a Steered Clock

Time–difference data are often used to steer a remote clock so that it is kept on time or on frequency with respect to some other system. Although the channel is "inside" this servo loop in principle, the loop gain does not attenuate offsets and phase shifts of the type we have been discussing. These offsets in the measurement channel are transmitted to the control system without attenuation. In addition, the control loop must be designed to minimize the effects of a channel failure.

Both of these problems tend to be more serious if frequency steering is used, since the measurement error or channel failure is converted to a control signal whose effect grows linearly with time. Although time steering does not have this problem, it introduces step–like discontinuities into the output of the steered device which complicate the analysis of its performance and degrade its spectral purity.

Our steered clocks are currently controlled using a combination of time and frequency steering. The frequency component of the steering is designed to control the long–term performance of the clock while the time corrections have a limited range and are only intended to compensate for the short–term fluctuations in the output. The steering system that realizes UTC(NIST), for example, is steered in this way. Using a measurement interval of 12 minutes, the average time correction that must be applied is 200 ps. The frequency steering is designed to realize UTC(NIST) in long term, so that if the system fails, the divergence of the output from UTC(NIST) is on the order of ns/day and is governed by the free–running stability of a cesium standard rather than by a frequency steering command that was intended to be applied for 12 minutes but in fact remained in force because the channel failed. The performance of this servo is nearly optimum in the sense that the short–term ($<$ 2 hours) RMS deviation in the steered output is essentially the same as the frequency noise in the cesium standard itself while the long–period performance is better than this and is limited by the performance of the AT1 time scale that is used as the reference for estimating UTC(NIST).

References

Julius S. Bendat and Allan G. Piersol, Measurement and Analysis of Random Data, New York, John Wiley and Sons, Inc., 1966, Chapter 9.

Alan V. Oppenheim and Ronald W. Schafer, Digital Signal Processing, Englewood Cliffs, New Jersey, Prentice–Hall, Inc., 1975, Chapter 11.

A. D. Whalen, Detection of Signals in Noise, New York, Academic Press, 1971, Chapter 9.

Fig. 1. Time Difference Between Two Channels

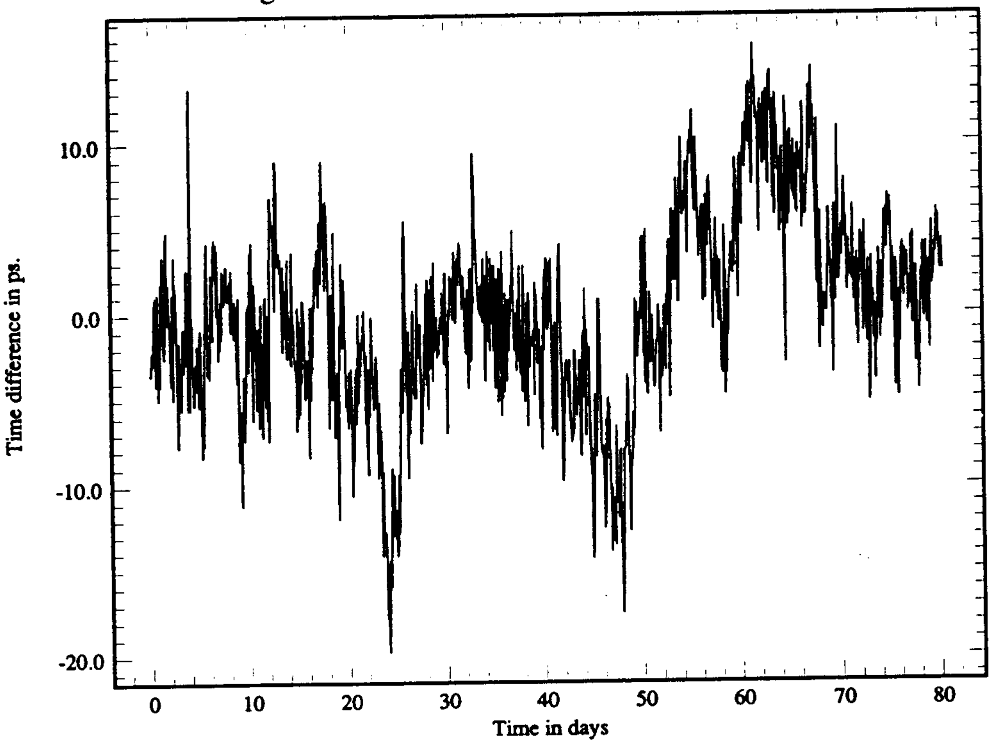

Fig. 2. Two-Sample Variance of Data in Fig. 1

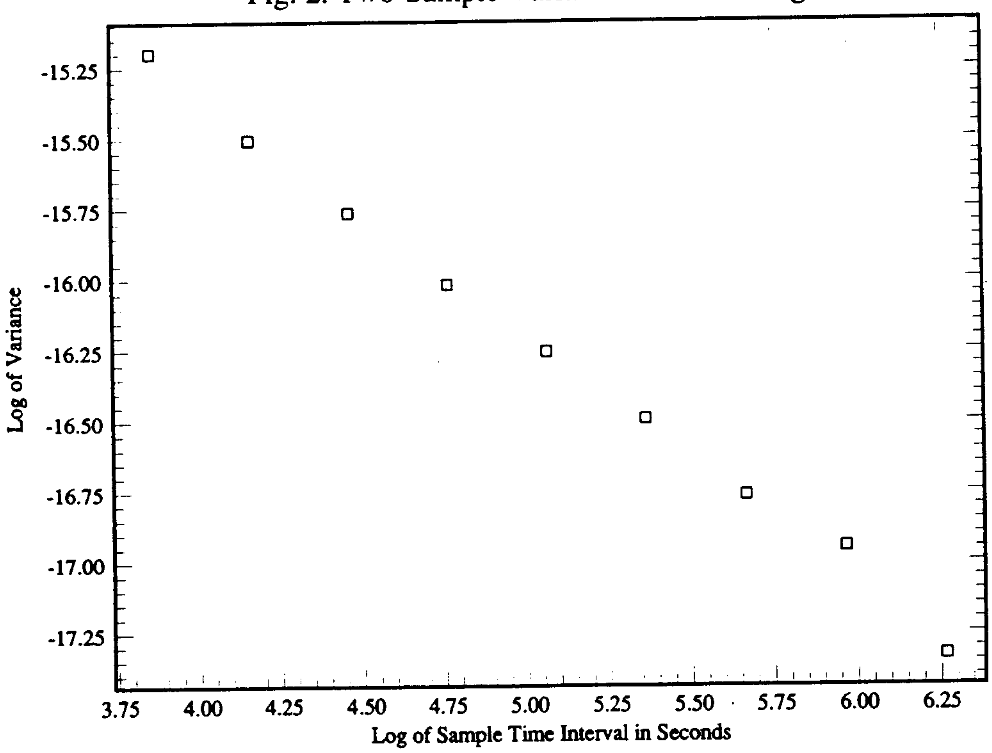

Fig. 3. Ambient Temperature

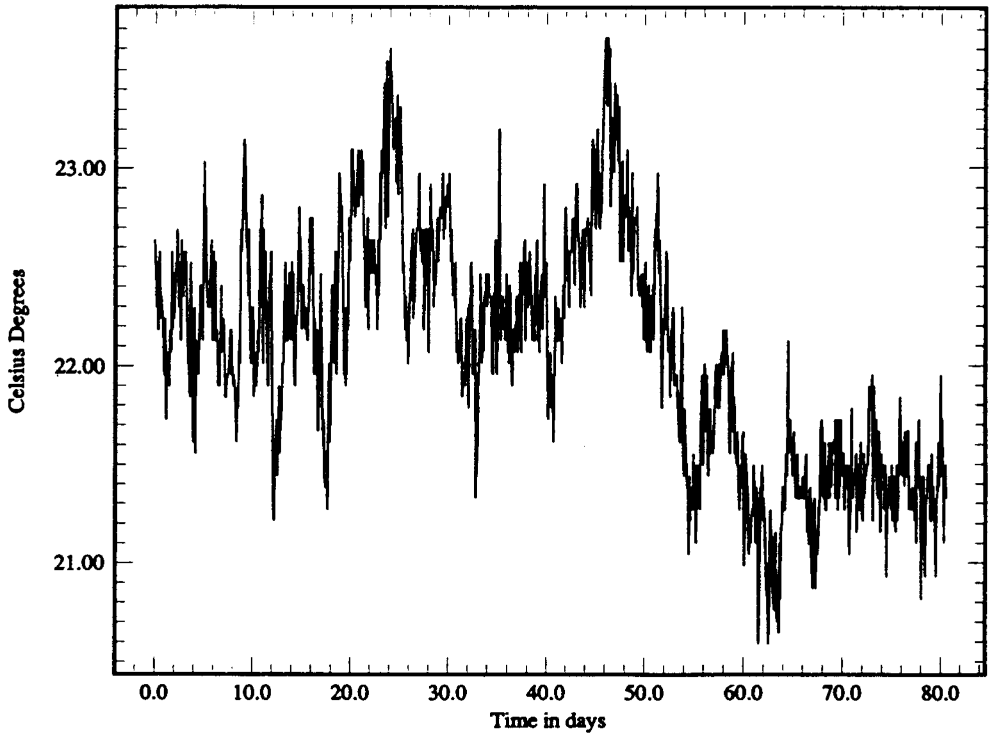

Fig. 4. Correlation as a Function of Lag

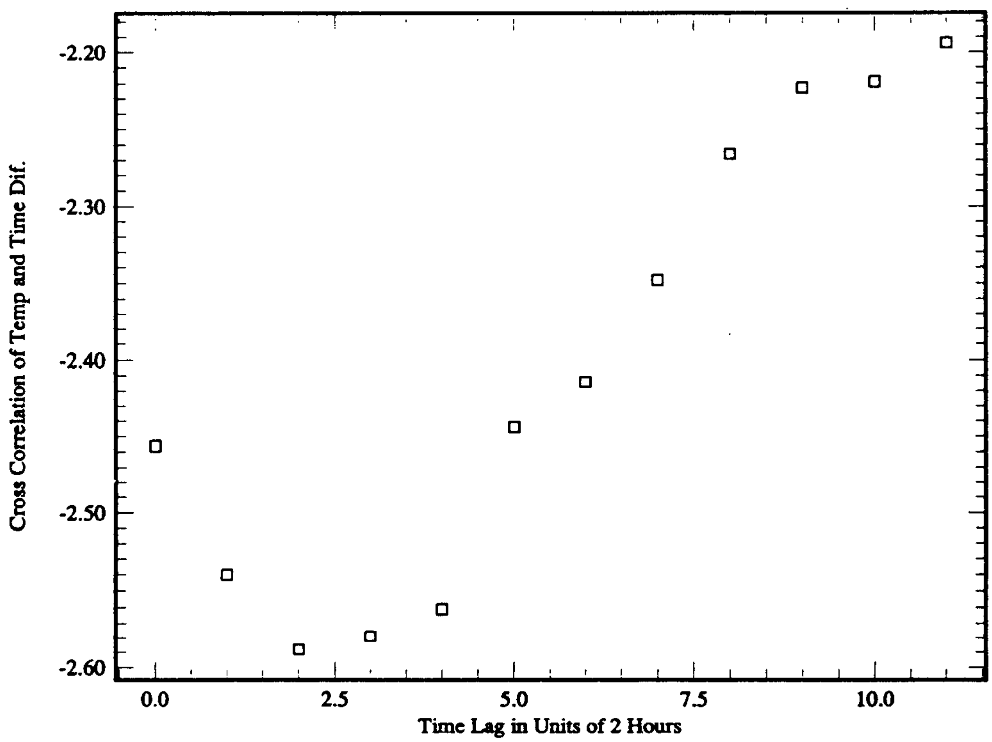

Fig. 5. Coherence as a Function of Frequency