A KALMAN FILTER FOR ATOMIC CLOCKS AND TIMESCALES

Lee A. Breakiron
U.S. Naval Observatory
Washington, DC 20392-5420, USA

Abstract

In a test of whether a Kalman filter could provide viable frequencies for atomic clocks, as well phases for a mean timescale, both simulated and actual hydrogen maser frequency data were processed with a two-state Kalman filter involving the parameters frequency and frequency drift and incorporating noise measurements based on Hadamard variances. Separate clock solutions were then combined and integrated to generate a mean timescale for the ensemble. Results for sampling times of up to at least several days indicate that such a timescale is even more stable than the one being used to steer the USNO Master Clock.

INTRODUCTION

Tryon and Jones [1] proposed the first atomic timescale algorithm based on a Kalman filter in order to provide an ensemble time that is more uniform than the time kept by any one of the constituent clocks and whose estimated states are optimal in the minimum squared error sense. Their state equations contained as clock parameters the phase, frequency, and drift relative to another weighted clock. A problem with this type of filter is that, since time is unobservable, elements of the covariance matrix grow without bound [2].

Barnes and Allan [3] proposed a Kalman filter that utilized only frequency and drift, pointing out that only these two parameters are physically meaningful for a clock, and indeed showed by simulations that such a two-parameter filter performed better than the three-parameter one. Though they studied cesiums, and cesiums generally have no significant drift, a drift parameter was necessitated by the need to allow for random walk FM noise. Elimination of the phase variable properly requires the use of measurement data without significant phase noise.

Consequently, a reasonable model for our masers (which generally have inherent frequency drift) is:

\[ f(t + \Delta t) = f(t) + \Delta t \cdot d(t) + \varepsilon(t + \Delta t) \]

\[ d(t + \Delta t) = d(t) + \eta(t + \Delta t) \]

where \( f(t) \) is a clock's frequency at time \( t = 1, 2, \ldots \); \( d(t) \) is its frequency drift; \( \varepsilon(t) \) and \( \eta(t) \) are independent random variables with zero mean and normal distributions (i.e. white noise processes) uncorrelated in time (i.e. zero autocorrelation); and \( \Delta t \) is the time step (1 hour in our case, which is in the white FM noise regime of masers).

In order to minimize process noise, one must choose a reference that is stable as possible. The Mean, at least as derived from the clocks under consideration, cannot so serve because it is not yet available at the
A Kalman Filter for Atomic Clocks and Timescales

U.S. Naval Observatory, 3450 Massachusetts Avenue NW, Washington, DC 20392-5420

Approved for public release; distribution unlimited

See also ADM001482. 33rd Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 27-29 Nov 2001, Long Beach, CA

See report
first step of a recursion. Even if it were, the problem could not be solved because it would be underdetermined. One could refer all the frequencies and drifts to one of the weighted clocks or to a mean predicted from previous data. However, it would be easier and more error- and correlation-free to refer one type of clocks, say the masers, to the Mean of the other type of clocks (cesiums). In USNO's operational timescale, the maser rates and drifts are calibrated against the cesium Mean anyway, which is always available. And the fact that in practice (though not in this study) the maser states would not be measured against those of another maser or their mean limits covariance growth with time.

THE KALMAN FILTER

The state transition equations are:

\[
\begin{bmatrix}
    f_1(t + \Delta t) \\
    d_1(t + \Delta t) \\
    f_2(t + \Delta t) \\
    d_2(t + \Delta t) \\
    \vdots
\end{bmatrix} =
\begin{bmatrix}
    1 & \Delta t & 0 & 0 & \cdots \\
    0 & 1 & 0 & 0 & \cdots \\
    0 & 0 & 1 & \Delta t & \cdots \\
    0 & 0 & 0 & 1 & \cdots \\
    \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\begin{bmatrix}
    f_1(t) \\
    d_1(t) \\
    f_2(t) \\
    d_2(t) \\
    \vdots
\end{bmatrix} +
\begin{bmatrix}
    \varepsilon_1(t + \Delta t) \\
    \eta_1(t + \Delta t) \\
    \varepsilon_2(t + \Delta t) \\
    \eta_2(t + \Delta t) \\
    \vdots
\end{bmatrix}
\]

or:

\[ X(t + \Delta t) = \Phi(t) X(t) + W(t + \Delta t) \]

where the subscripts denote clocks and \( \Phi(t) \) is the state transition matrix.

Let \( z_i(t) \) be the phase of clock \( i \) at time \( t \) relative to the cesium Mean. The observation equations are:

\[
\begin{bmatrix}
    (z_1(t) - z_1(t - \Delta t))/\Delta t \\
    (z_2(t) - z_2(t - \Delta t))/\Delta t \\
    \vdots
\end{bmatrix} =
\begin{bmatrix}
    1 & 0 & 0 & 0 & \cdots \\
    0 & 0 & 1 & 0 & \cdots \\
    \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\begin{bmatrix}
    f_1(t) \\
    d_1(t) \\
    f_2(t) \\
    d_2(t) \\
    \vdots
\end{bmatrix} +
\begin{bmatrix}
    v_1(t) \\
    v_2(t) \\
    \vdots
\end{bmatrix}
\]

or:

\[ Z(t, \Delta t) = H(t) X(t) + V(t) \]

where \( Z(t, \Delta t) \) is the vector of measurements, \( H(t) \) is the observation matrix (time-dependent in that the number of weighted clocks with available data may vary), and \( V(t) \) is the vector of measurement errors \( v_i(t) \).

For uncorrelated parameters, the covariance matrix \( Q(\Delta t) \) of the process errors \( W(t) \), for sufficiently small \( \Delta t \), is such that:
\[
\frac{dQ}{dt} = \begin{bmatrix}
    s_{e^2} & 0 & 0 & 0 & \cdots \\
    0 & s_{n^2} & 0 & 0 & \cdots \\
    0 & 0 & s_{f^2} & 0 & \cdots \\
    0 & 0 & 0 & s_{n^2} & \cdots \\
    \cdots & \cdots & \cdots & \cdots & \cdots
\end{bmatrix}
\]

where \( s_{e^2} \) is the noise spectral density (sometimes called variance rate) of random-walk FM and \( s_{n^2} \) is the noise spectral density of random-run FM (random walk of frequency drift). The effect of these errors on the system for any interval \( \Delta t \) is:

\[
Q(\Delta t) = \int \Phi(t) \frac{dQ}{dt} \Phi^T(t) \, dt
\]

\[
\begin{bmatrix}
    s_{e^2} \Delta t + s_{n^2}(\Delta t)^3/3 & s_{n^2}(\Delta t)^2/2 & \cdots \\
    s_{n^2}(\Delta t)^2/2 & s_{n^2} \Delta t & \cdots \\
    \cdots & \cdots & \cdots
\end{bmatrix}
\]

where \( \Phi^T \) is the transpose of matrix \( \Phi \).

If the measurement errors are uncorrelated, the measurement error vector \( V(t + \Delta t) \) has the covariance matrix:

\[
R(\Delta t) = \begin{bmatrix}
    s_{v^2} \Delta t & 0 & \cdots \\
    0 & s_{v^2} \Delta t & \cdots \\
    \cdots & \cdots & \cdots
\end{bmatrix}
\]

where \( s_{v^2} \) is the noise spectral density of random walk of phase, or white FM. The assumption of zero autocorrelation does require prefiltering for such outliers as erroneous phase measurements shared by consecutive first differences.

So far we have assumed that there are data every time step. When this is not the case, \( \Delta t \) in our state equations and \( Q(t) \) formulation can be the actual interval between measurements, but not in our formulation for \( R(t) \), where \( \Delta t \) must remain the time step.

Rather than noise spectral densities, let us instead solve for the variances \( \sigma_v^2 = s_v^2 \Delta t \), \( \sigma_e^2 = s_e^2 \Delta t \), and \( \sigma_n^2 = s_n^2 \Delta t \) as our noise parameters. The Hadamard variance, \( \sigma_H \), when the noise sources consist only of white noise FM, random-walk FM, and random-run FM, can be expressed as [4]:

\[
\sigma_H^2 = \sigma_v^2 \tau^{-1} + (1/6) \sigma_e^2 \tau + (11/120) \sigma_n^2 \tau^3
\]

(1)
where $\tau$ is the sampling time. Our noise parameters may be solved by least squares from this equation of condition. Hadamard variance has the beneficial property of being insensitive to drift.

Improved values could be obtained from fits of the individual masers to the cesium Mean and recomputed occasionally, say at times of significant frequency and drift changes. These quantities introduce time constants into the noise terms that allow for the different noise processes. As with $H(t)$, the only dependence of $Q(\Delta t)$ and $R(\Delta t)$ on time is on the particular clocks involved.

The first step in the version of the Kalman recursion in which no knowledge of the process noise is required is to calculate the inverse of the transition covariance:

$$P^{-1}(t) = [P^{-}(t)]^{-1} + H^T(t)R^{-1}(t)H(t)$$

where the transition covariance matrix predicted for the next step $P^{-}(t)$ (whose superscript denotes "predicted") can be assumed initially to contain huge covariances.

Second, we compute the Kalman gain $K$:

$$K(t) = P(t)H^T(t)R^{-1}(t)$$

where $P(t)$ is the transition covariance matrix of the current step. Next, we update the parameter estimate thusly (where $\hat{\cdot}$ denotes "estimated"):

$$\hat{X}(t) = \hat{X}^{-}(t) + K(t) [Z(t) - H(t)\hat{X}^{-}(t)]$$

Then we make the following predictions for the next step

$$\hat{X}^{-}(t + \Delta t) = \Phi(t)\hat{X}(t)$$

$$P^{-}(t + \Delta t) = \Phi(t)P(t)\Phi^T(t) + Q(t)$$

and so on. Each application of the recursion yields an estimate of the system state that is a function of the elapsed time since the last filter update, which can occur any time, i.e. $\Delta t$ is not necessarily constant, and the data need not be equally spaced. Note that there are two matrix inversions instead of one as in the more common formulation of the Kalman filter. We will see that this is of no consequence. Further, note that our clock model assumes the frequencies to be real-time, so their initial estimates must be those applying at time $t = 0$.

While it may appear to be necessary to invert a matrix $2n$ on a side, where $n$ is the number of clocks, note that $\Phi(t)$ and $Q(t)$ are block-diagonal and $R(t)$ is diagonal, implying, respectively, that: (1) maser frequencies relative to the cesium Mean are independent of one another; (2) maser frequency errors are independent of one another; and (3) measurement errors are independent. These assumptions may not be strictly true due to, e.g., environmental influences, so there may be some mismodelling, but the diagonal structure of these matrices implies the same structure for $P(t)$ and, hence, that the problem may be solved as a series of $2 \times 2$ matrices, greatly reducing computational time and errors. Thus, the two inversions required by the above filter formulation is not a problem. And all correlations with the reference would be allowed for, while many Kalman-based algorithms neglect these.
If prefilter analysis of the data indicates the presence of a significant frequency step, the covariances in $P(t)$ could be reset to huge values in order to reinitialize the filter.

THE KALMAN MEAN

The frequencies and drifts produced by this two-state Kalman filter could be averaged and integrated to generate a Kalman-based mean timescale. The relative sizes of the clock covariances $P(t)$ would be inversely related to the individual clock weights. Robustness would require the imposition of an upper limit on any one clock's relative weight. If the filter and model operate correctly, the clock variances should slowly converge to steady-state values.

If, in forming the mean timescale, each clock's frequency is weighted according to the inverse of its variance, the timescale would be biased by the “clock-ensemble effect.” The prediction error is always too small because the frequency of a clock is correlated with the frequency of the ensemble, since the ensemble includes a contribution from each clock. The weights, which are proportional to the inverse of the variances, are therefore systematically too large, causing a positive feedback that increasingly biases the timescale toward the ensemble's best clock [5,6]. This can be avoided if the weight of each clock is based on its stability relative to a mean of which it is not a part, either by: (1) using another mean entirely, say one based on cesium clocks rather than masers; or (2) computing a mean consisting of the rest of the clocks in the ensemble and referring the clock in question to that mean. (2) can be accomplished by using the relation:

$$\sigma_u^2 = \sigma^2 / (1 - w)$$

where $\sigma^2$ is the uncorrected variance, $\sigma_u^2$ is the corrected (unbiased) variance, and weight is the relative weight [7]. In practice, an upper limit would also have to be placed on any one clock’s weight.

USNO's operational timescale assumes constant drifts and, aside from drift, constant frequencies. Proper tuning of the Kalman filter through the system noise $Q$ matrix should limit the steady-state parameters to the expected white FM and random-walk FM noises, without the necessity of exponential filters such as those employed by Stein to derive his parameters [8] and used by NIST in their AT1 and AT2 algorithms [2].

Parameter and error estimates would be available from the filter in near real time, but practical implementation would require robust detection and rejection of outliers (measurements whose errors are unlikely to have originated in the clock model's process noises), as well as prompt recognition of time and frequency steps. Kalman filters provide both a forecast of the next data point and an estimate of its uncertainty, so it is possible to implement robust outlier detection, though one must allow for the possibility that discrepant points are indications of a step in time or frequency. Stein proposed an altered gain function to smoothly deweight outliers and an adaptive filter to recognize time and frequency steps [9].

Initially the filter's knowledge of the parameters will be crude, but in postprocessing one can make use of “future data” by running forward and backward filters through the data and averaging the corresponding parameters weighted by the number of time steps contributing to each using weights based on filter variances (“smoothing”) [10]. In order to avoid using the same datum twice, the state at any one time from one of the filters has to be averaged with the predicted state at that time from the other filter. This kind of processing of time data has only been done by NIST a posteriori for their AT2 timescale [11].
This could be done in near real time every time step for some interval of the latest data, though retroactive recomputation would be involved.

TESTS ON SIMULATED DATA

In order to try the Kalman filter algorithm on data where the truth is already known, tests were made on 10,000 units of hourly simulated data for 10 masers with white FM and random-walk FM noises characteristic of the auto-tuned hydrogen masers at USNO. Randomly chosen frequency drifts were injected, as was measurement phase noise of 2 ps (all kindly provided by Paul Koppang). The frequencies of these simulated masers are displayed in Figure 1 (though several are obscured by overlap) and the instabilities of their frequencies are shown by their Hadamard deviations as plotted in Figure 2.

Figure 1. Simulated frequencies
(x \times 10^{-15}) for 10 masers

![Graph showing simulated frequencies for 10 masers with varying drifts.](image)

The two-state Kalman filter described above, including the forward-and-backward averaging and a correction for the clock-ensemble effect, was applied to all 10 simulated masers. The results agreed well with the a priori frequencies and were similar to those shown in Figure 3 for maser 1 (the one with the greatest positive drift), which only shows the first 20 days for clarity. Note the large initial variations in the forward-filter data as the filter strives to learn the true frequency; the small variations in the backward-filter data, whose filter knows the frequency very well by this time; and the favoring of the average for the backward-filter data.

The frequencies of the average filter data were then integrated to obtain a mean timescale, weighting the clocks according to their inverse frequency variances and correcting for the clock-ensemble effect. The frequencies of this mean timescale are plotted in Figure 1 and its Hadamard deviations are plotted in
TESTS ON REAL DATA

In order to test the filter on real data, 13141 units of hourly data (from MJD 51644 to 52192) were selected for four Datum (Sigma Tau model) masers at USNO, as produced by the Timing Solutions Corporation's (formerly “the Steintec”) measurement system [12]. The dominant noise of this system is a phase noise of 2 ps. Though this system is still experimental, its data were chosen in preference to those of the operational Data Acquisition System (DAS) of universal time-interval counters and multiplexed switches because of the latter system’s significant phase noise (~35 ps), which would have violated the two-state model assumed in the filter. The data were referenced to the mean of the four masers as computed by the currently operational timescale algorithm [13], which removed constant frequencies and drifts from each steady-state segment of every maser by modelling it against the cesium mean computed by the same algorithm, but using DAS data. All data had been robustly pre-filtered.

The Kalman filter yielded very satisfactory results for each clock, as shown for maser NAV2 in Figure 4, where the Kalman frequencies are compared to the hourly first differences of phase. The Kalman frequencies were not very sensitive to choice of noise parameters (tuning of the filter), e.g. any of the portions in Figure 4 gave sufficiently accurate noise parameters through their Hadamard variances and Equation (1). Implementation of such method, in preference to the current clock modelling by least-squares fits to constant-rate-and-drift data segments [13], would eliminate a significant source of subjectivity and labor.
Next, the frequencies and drifts produced by the Kalman filter were averaged and integrated with the basic timescale equation (Equation (5) in [13]) to generate a Kalman-based mean timescale, which was then compared with the maser mean timescale generated by the currently operational algorithm. The clocks were weighted, in the case of the Kalman timescale, according to their inverse frequency variances corrected for the clock-ensemble effect and, in the case of the operational timescale: (a) at unity among each class of clock, i.e. either cesiums or masers; and (2) according to the inverse Allan variances averaged over each class, which varied with look-back time ("dynamic weighting") [13]. All data had been prefiltered for significant outliers. The differences between the two mean timescales never exceeded 123 ps (see Figure 5). The excursions at MJD 52034 and MJD 52135 are due to the different methods each algorithm used to extrapolate over data gaps. Since the clock weights in the Kalman timescale are not dynamic (i.e. the timescale is never revised retroactively), it would serve as a stabler target toward which to steer a Master Clock than the operational timescale if their inherent frequencies stabilities were comparable.

The frequency stabilities of the two algorithms were investigated by deriving the Allan deviations of the two mean timescales relative to the operational maser Mean generated for an independent ensemble of three masers in another building (Bldg. 78 rather than Bldg. 52) for 8424 hours of data that were contemporaneous. Figure 6 shows that the Kalman mean is significantly more stable than the operational mean for sampling times shorter than about 5.8 days.

The three masers in Bldg. 78 were also connected to an independent TSC measurement system, so 8400 hours of their TSC data were utilized to generate a Kalman mean timescale, which was in turn compared to the operational maser Mean in Bldg. 52, with entirely similar results (see Figure 7).
CONCLUSIONS

It has been shown that, for an ensemble of hydrogen masers connected to a low-noise measurement system, a two-state Kalman filter with the parameters of frequency and frequency drift is quite capable of producing a mean timescale with a frequency stability that is not only viable, but at sampling times up to several days, actually superior to that produced by a conventional least-squares algorithm.

If the clocks are properly modelled and this two-state Kalman filter properly tuned, the predictions of the clock frequencies with respect to the ensemble will be optimal out to the sampling time at which random-walk FM obtains. If this is true for the individual clocks and since the relative clock weights should also be optimal (being based on unbiased frequency variances), it should be true for the mean of the ensemble. The frequency residuals should be white noise over a range of sampling times that includes a few days, the time constant of the frequency steering employed to synchronize the USNO Master Clock with the mean timescale of the rest of the ensemble. Consequently, such a near-real-time Kalman filter would produce optimal predictions of the Mean toward which a Master Clock could be steered.

If the Kalman-based maser Mean is stable enough in real time, its utilization to steer the Master Clocks would obviate the jumps in the real-time Mean due to the retroactive recomputations of the operational timescale that currently destabilizes the latter [14]. Such recomputations are done with the current algorithm both because of revised modelling of the clocks as more data accumulate, but also because of the dynamic weighting of the masers relative to the cesiums, which are combined in the mean to which the Master Clock is steered. While some recomputation would be entailed with the Kalman filter if smoothing were involved, this effect would be likely to be much smaller than that caused by significant changes in weight or frequency, as is currently the case. In short, such a filter would optimize the clock weights (hence, use of data), the stability of the Mean, and the stability of the Master Clock.
Figure 5. Phase differences (ns) between the Kalman and operational means

The current algorithm assumes each clock is being modelled against a Mean composed of statistically independent clocks. But clock parameter correlations can occur in practice due, say, to the environmental sensitivities of co-located clocks. A Kalman filter, however, would be able to remove such correlations on the fly, which if not removed might result in systematic errors. Furthermore, the presence of autocorrelated errors in the ordinary least squares currently used to derive the clock parameters is known to have deleterious effects: (1) though the correlation coefficients are unbiased, they no longer minimize the variance, and may be quite inefficient; and (2) the estimate of the variance may seriously underestimate the true errors of the parameters [15].

The insensitivity of the filter to choice of noise parameters bodes well for the use of data intervals on the scale of 60 days (found to be optimal for the current algorithm [16]) rather than the shorter time constant of an exponential filter (more data entailing greater stability), though this should be tested further. Generation of a real-time timescale with such a Kalman filter will be attempted and the results compared with those of the operational algorithm.

ACKNOWLEDGMENTS

I would like to thank Paul Koppang for valuable comments and providing simulated maser data.
REFERENCES


Figure 7. Frequency stabilities vs. Bldg. 52 operational Mean

-14.5
-14.6
-14.7
-14.8
-14.9
-15
-15.1

Log Allan Deviation

Log Tau (sec)

OPERATIONAL MEAN (BLDG. 78)
KALMAN MEAN (BLDG. 78)