Reduced Kalman Filters for Clock Ensembles

Charles A. Greenhall
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, CA 91109 USA
E-mail: cgreenhall@jpl.nasa.gov

Abstract— **This paper summarizes the author's work on timescales based on Kalman filters that act upon the clock comparisons. The natural Kalman timescale algorithm tends to optimize long-term timescale stability at the expense of short-term stability. By subjecting each post-measurement error covariance matrix to a non-transparent reduction operation, one obtains corrected clocks with improved short-term stability and little sacrifice of long-term stability.**

I. Free-Running Timescales

This paper is a summary of work on the formation of timescales, or composite clocks, that the author has carried out over the last ten years, with the goal of surpassing the performance of a previous timescale algorithm [1] that uses a Kalman filter to estimate what the clocks are actually doing.

Let us describe briefly the notion of a timescale derived from an "ensemble" of clocks. One has a set of n free-running clocks with "phase" (i.e, time) residuals $x_1^1(t), \ldots, x_1^n(t)$ at the nominal time t. One cannot read the values $x_1^i(t)$ directly (to the desired precision) but can only measure the phase differences $x_1^{ij}(t) := x_1^i(t) - x_1^j(t)$ at a set of measurement times t. A timescale algorithm, which uses these data and a mathematical model for the clocks, calculates numerical phase estimates $\hat{x}_1^i(t)$ to be subtracted from the clock phases, whether logically or physically, to give corrected clock phases

$$\tilde{x}_1^i(t) = x_1^i(t) - \hat{x}_1^i(t). \tag{1}$$

One hopes to produce corrected clocks with better stability[1] than the physical clocks over a range of averaging times. Since a motion common to all the clock phases is transparent to their differences, one cannot expect the corrected clocks to stay close to a physical clock. Improved timekeeping accuracy is not to be expected from combining clock readings in this way.

Section II describes the most common clock model. Section III sets up an ensemble of these clocks and describes the natural Kalman filter algorithm for calculating the phase estimates. The existing results motivate the author's work to reduce the short-term noise of the corrected clocks. Sections IV and V describe the author's "reduced" Kalman filter algorithms that apply to noiseless measurements, as in a local ensemble. Section VI describes Davis's improvements. Section VII extends the method to remote sets of clocks with significant time-transfer noise. Section VIII gives a recent result on the reduced Kalman covariance matrices of a clock ensemble.

[1]Stability of a phase $x_1(t)$ for an averaging time τ is defined by the Allan deviation $\sigma_A(\tau)$ or Hadamard deviation $\sigma_H(\tau)$, where $\sigma_A^2(\tau) = (2\tau^2)^{-1} \mathrm{E}\left[\Delta_\tau^2 x_1(t)\right]^2$, $\sigma_H^2(\tau) = (6\tau^2)^{-1} \mathrm{E}\left[\Delta_\tau^3 x_1(t)\right]^2$.

II. Clock Ensemble Model

Much work in this area has been based on a clock model [2][3] with three independent process-noise components, called white FM (WFM), random-walk FM (RWFM), and random-run FM (RRFM), with phase spectra $S(f)$ proportional to f^{-2}, f^{-4}, and f^{-6}, and Hadamard deviations $\sigma_H(\tau)$ proportional to $\tau^{-1/2}$, $\tau^{1/2}$, and $\tau^{3/2}$, respectively. Missing are the odd-powered models, in particular flicker FM (FFM), for which $S(f) \propto f^{-3}$, $\sigma_H(\tau) = \text{const.}$

The state of one clock at time t is represented as a column vector, $x(t) = \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) \end{bmatrix}^{\mathrm{T}}$, whose components are called phase, frequency state, and drift state. The state evolves according to the stochastic differential equations

$$\frac{dx_1}{dt} = x_2 + w_1, \ \frac{dx_2}{dt} = x_3 + w_2, \ \frac{dx_3}{dt} = w_3 \tag{2}$$

where $w_i(t)$ are independent mean-zero white noises with known two-sided spectral densities q_i. One may set $q_3 = 0$, for example, to simulate a constant drift $x_3(t) = c$, in which case $x_1(t)$ contains a quadratic component $\frac{1}{2}ct^2$.

The state of n clocks is a $3n$-vector $X(t)$, a stack of n clock vectors $x^i(t)$. The measurements performed on the state consist of noisy phase comparisons, $x_1^{ij}(t) + v_{ij}(t)$, at a sequence of times $t = t_k$, where the measurement noise variances, $\mathrm{E}\, v_{ij}^2(t) = r_{ij}$, are known. Even without noise, however, the measurements do not determine the clock states.

III. Natural Kalman Timescale

To set up a Kalman filter to estimate the state from the measurements, the differential equations (2) for each clock are integrated between two measurement times, call them $t-\delta$ and t, to give a stochastic difference equation. For one clock this takes the form

$$x(t) = \phi(\delta)\, x(t-\delta) + w(t, \delta) \tag{3}$$

where

$$\phi(\delta) = \begin{bmatrix} 1 & \delta & \delta^2/2 \\ 0 & 1 & \delta \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

and the process noise $w(t,\delta)$ is a mean-zero 3-vector with covariance matrix

$$\mathrm{E}\, ww^{\mathrm{T}} = q(\delta) = \begin{bmatrix} q_1\delta + q_2\frac{\delta^3}{3} + q_3\frac{\delta^5}{20} & * & * \\ q_2\frac{\delta^2}{2} + q_3\frac{\delta^4}{8} & q_2\delta + q_3\frac{\delta^3}{3} & * \\ q_3\frac{\delta^3}{6} & q_3\frac{\delta^2}{2} & q_3\delta \end{bmatrix} \tag{5}$$

where the $*$ entries are determined by symmetry. The difference equation for the whole ensemble is written

$$X(t) = \Phi(\delta)\, X(t-\delta) + W(t,\delta)$$

where Φ has n copies of ϕ down the diagonal and W is a stack of n uncorrelated w^i's; consequently $\mathrm{E}\, WW^{\mathrm{T}} = Q(\delta)$, which has n 3×3 matrices $q^i(\delta)$ down the diagonal.

The phase comparison measurements at time t take the form $y(t) = H(t)\, X(t) + v(t)$, where H is a matrix of ± 1's and 0's such that each row takes the difference of two phase components of X. The mean-zero measurement noise v satisfies $\mathrm{E}\, vv^{\mathrm{T}} = R(t)$, a known matrix, usually diagonal. For a local ensemble one sets $R = 0$.

Let $\hat{X}(t)$ be an unbiased estimate of $X(t)$; its quality is specified by an error covariance matrix

$$C(t) = \mathrm{E}\left[X(t) - \hat{X}(t)\right]\left[X(t) - \hat{X}(t)\right]^{\mathrm{T}}.$$

Given $\hat{X}(t-\delta)$, $C(t-\delta)$, and $y(t)$, the Kalman filter equations produce $\hat{X}(t)$ and $C(t)$ as follows. For clarity, some dependencies on δ and t are suppressed.

$$\begin{aligned} \hat{X}^-(t) &= \Phi\hat{X}(t-\delta) \quad \text{(prediction of } X(t)\text{)} \\ C^- &= \Phi C(t-\delta)\,\Phi^{\mathrm{T}} + Q \quad \text{(error cov. of } \hat{X}^-(t)\text{)} \\ K &= C^- H^{\mathrm{T}}\left(HC^-H^{\mathrm{T}} + R\right)^{-1} \quad \text{(Kalman gain)} \\ \hat{X}(t) &= \hat{X}^-(t) + K\left[y(t) - H\hat{X}^-(t)\right] \\ C(t) &= (I - KH)\, C^- \end{aligned}$$

Evidently $\hat{X}(t)$ is a linear function of $\hat{X}(t-\delta)$ and $y(t)$. In fact, $\hat{X}(t)$ is unbiased for $X(t)$ and has the minimal error covariance matrix $C(t)$ of any unbiased linear function in the sense that $C_1 - C(t)$ is non-negative definite if C_1 is the error covariance matrix of any other unbiased linear function.

The natural Kalman corrected clock states are defined by $\tilde{X}(t) = X(t) - \hat{X}(t)$. The corrected phase $\tilde{x}_1^i(t)$ represents the timescale at clock i. The corrected phases need not coincide for remote clocks with noisy phase comparisons, but they do coincide for a set of local clocks with negligible measurement noises ($R = 0$). In this case the common corrected phase constitutes the unique timescale, which may be called $\tilde{x}_1(t)$.

A natural Kalman timescale was implemented at NIST during the 1980's as the local timescale TA(NIST). The GPS Kalman filter [4] also incorporates a natural Kalman timescale. When applied to a simulated ensemble of dissimilar clocks, the NIST timescale was observed to be quiet in long term but noisier in short term than some of the clocks in the ensemble. Weiss and Weissert [5] suggested that the Kalman filter is optimizing accuracy at the expense of stability. This author confirmed their observations with his own simulations and set out to find an improved method of using a Kalman filter to make a timescale that is more stable than any clock in its ensemble over a wide range of averaging times. Stein [6][7] had already accomplished this goal with the KAS-2 algorithm and had described some of its principles, but a full description remains proprietary.

In searching for improved Kalman timescales, the author has held to one fixed principle: the natural Kalman filter does a good job of estimating the underlying frequency and drift states, which, in the author's opinion, reflect the physics of the frequency standards driving the clocks. The author's filter modifications preserve the frequency and drift state estimates but modify the phase estimates to improve the stability of the corrected phases.

IV. Kalman Plus Weights

The author's first successful timescale algorithm [8][9] applies to a local ensemble with negligible measurement noise. The natural Kalman phase estimates are not used at all, but the Kalman frequency and drift estimates, $\hat{x}_2^i$ and $\hat{x}_3^i$, are used in a version of the "basic timescale" iteration. The propagation of the timescale, call it x_1^e, from $t-\delta$ to t can be written most simply as

$$x_1^e(t) - x_1^e(t-\delta) = \sum_{i=1}^{n} \lambda_i(t) \cdot\left[x_1^i(t) - x_1^i(t-\delta) - \delta\hat{x}_2^i(t-\delta) - \tfrac{1}{2}\delta^2\hat{x}_3^i(t-\delta)\right], \tag{6}$$

which says that the increment of x_1^e is a weighted average of the true phases $x_1^i(t)$ minus their predicted values. The weights λ_i are to be chosen subject to $\sum\lambda_i = 1$. The true phases are unknown, of course. To turn this into something that depends only on known quantities, subtract $x_1^j(t) - x_1^j(t-\delta)$ from both sides to give

$$x_1^{ej}(t) - x_1^{ej}(t-\delta) = \sum_{i=1}^{n} \lambda_i(t) \cdot\left[x_1^{ij}(t) - x_1^{ij}(t-\delta) - \delta\hat{x}_2^i(t-\delta) - \tfrac{1}{2}\delta^2\hat{x}_3^i(t-\delta)\right], \tag{7}$$

which is now a computable iteration for $x_1^{ej}(t) := x_1^e(t) - x_1^j(t)$, the offset of the timescale from physical clock j.

To set the weights λ_i one tries to minimize the variance of $x_1^e(t) - x_1^e(t-\delta)$ in (7). From (3) and (5) one can argue for the approximation that $1/\lambda_i$ should be made proportional to the matrix element $q_{11}^i(\delta)$. As a further approximation, one can make the weights inversely proportional to the white FM noise levels q_1^i of the clocks.

Figure 1 [9] shows the performance of the natural Kalman and KPW timescales on a simulated ensemble of 8 clocks. The odd-numbered clocks have a combination of the three noise types WFM, RWFM, and RRFM. The KPW scale is much quieter than the natural Kalman scale at short averaging times; the natural Kalman scale is slightly quieter around $\tau = 10^6$ s.

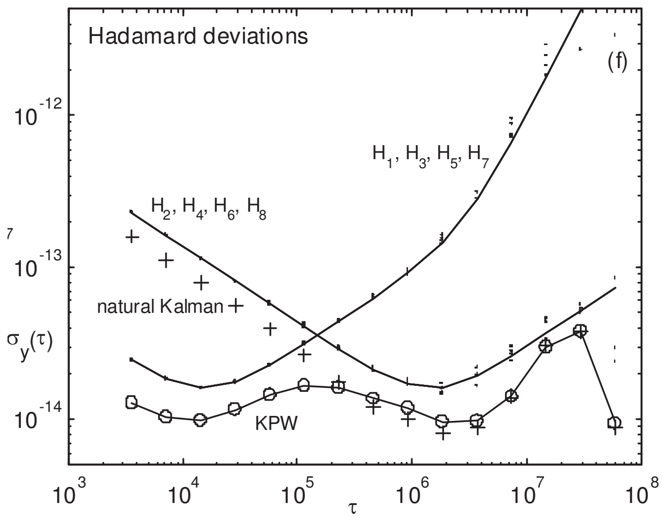

Fig. 1. Hadamard deviation of 8 simulated clocks and two timescales.

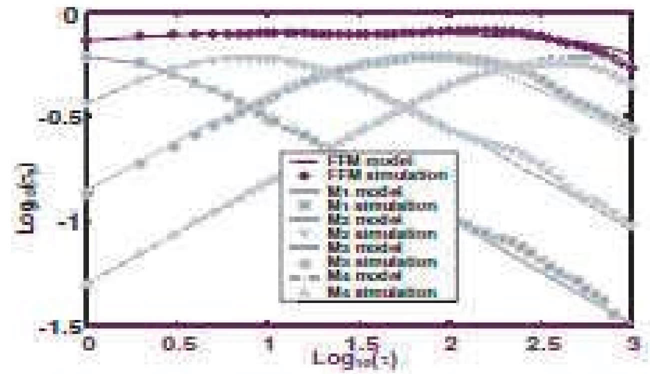

Fig. 2. Approximating flicker FM noise by a sum of independent Markov frequency processes.

V. Reduced Kalman Timescale

For the KPW scale the Kalman phase estimates are irrelevant, but those state components are still carried along by the Kalman filter. Because only the phase differences are measured, the phase rows and columns of the covariance matrix $C(t)$ grow fast. After a processing a set of noiseless measurements, however, C is singular; in fact, the phase rows and columns are all the same vector. The author proved that setting all these matrix elements to zero after each measurement leaves unchanged the future Kalman frequency and drift estimates and their covariances, thus giving a more stable calculation. The author was fortunate to discover, just before giving the paper [9] at BIPM, that this crude operation of covariance reduction, called the "guillotine", also causes the ignored phase estimates $\hat{x}_1^i(t)$, when used as in (1), to yield automatically a different "reduced" timescale $\tilde{x}_1(t)$ that equals a KPW timescale $x_1^e(t)$ with the exact weights λ_i that minimize the mean squared value of the timescale increment $x_1^e(t) - x_1^e(t-\delta)$ [10]. These implicit optimal weights, which can easily be obtained from the Kalman gain matrix, are usually not far from the explicit weights chosen previously for KPW, and the two timescales differ little in practice. The natural Kalman timescale is also a KPW scale with implicit weights [11], but these can differ greatly from the optimal weights.

A Kalman filter that modifies its future estimates by using a reduced covariance matrix is called a reduced Kalman filter. The covariance matrix becomes an operational tool instead of a measure of the actual uncertainty of the phase estimates.

VI. Work with J. Davis

During 2003–5 John Davis of NPL and the author worked together on the reduced Kalman timescale of a local ensemble. Two of his improvements of the art [12] are summarized here.

First, an approximation of FFM noise can be added to the 3-state clock model by including a sum of independent Markov frequency processes (white noise through a single-pole lowpass filter). One such process is described by

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} + Rx_2 = w$$

where x_1 is phase, x_2 is frequency, and w is a white noise with spectral density S_w. Integrating this from $t-\delta$ to t gives the stochastic difference equation

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{R}\left(1 - e^{-R\delta}\right) \\ 0 & e^{-R\delta} \end{bmatrix} \begin{bmatrix} x_1(t-\delta) \\ x_2(t-\delta) \end{bmatrix} + \begin{bmatrix} w_1(t,\delta) \\ w_2(t,\delta) \end{bmatrix}$$

where the covariance matrix of the w_i depends on S_w, R, and δ. One can include several such components in each clock with values of S_w and R chosen so that their sum approximates FFM noise over a range of time scales, as shown in figure 2, which plots the Allan deviation of the components and their sum. One can vary the parameters of the Markov components to produce a more complex frequency stability profile between regions of short and long averaging times.

Second, Davis devised a numerical calculation of the theoretical stability of the reduced Kalman timescale. This is done by calculating the true covariance matrix in a separate iteration that runs alongside the iteration for the operational reduced covariance matrix. One can also calculate the theoretical stability of each clock phase minus the timescale, and the results can be compared with the measured stability of the clock corrections $\hat{x}_1^i(t)$ calculated by the operational filter.

VII. Time-Transfer Noise and IEM

When noise is admitted to the clock phase comparisons, the corrected clocks need not coincide, but the ith corrected phase $\tilde{x}_1^i$ can be said to represent the timescale at clock i. In his implementation of the natural Kalman timescale, Brown [4] devised a so-called implicit ensemble mean (IEM), a 3-vector $\bar{x}^B = \begin{bmatrix} \bar{x}_1^B & \bar{x}_2^B & \bar{x}_3^B \end{bmatrix}^{\mathrm{T}}$. This is a weighted average $W_{IE}\tilde{X}$ of the n corrected clock states $\tilde{x}^i$, using a $3\times 3n$ weight matrix W_{IE} that is calculated by solving a least-squares problem whose noise covariance matrix is just the Kalman filter matrix C. The corrected states tend to cluster about the IEM, and the covariance matrix C_B of the n "representation errors" $\tilde{x}^i - \bar{x}^B$ is Brown's reduced form of C. The representation errors tend to be smaller than the corrected clock states, and C_B grows more slowly than C does. The use of C_B in the Kalman filter is transparent to the future state estimates, and the $\tilde{x}^i$ still represent the natural Kalman timescale. For the case of noiseless measurements, when all the corrected clocks coincide, Galleani and Tavella [13] showed that the IEM phase component $\bar{x}_1^B$ does not reduce to the common timescale phase $\tilde{x}_1$. This happens because $\bar{x}_1^B$ depends on the corrected frequency and drift states [15].

By modifying Brown's ideas, the author was able to devise an extension of the reduced Kalman timescale that works well in the presence of measurement noise [14]. The IEM is now a one-dimensional phase $\bar{x}_1^G$, a weighted average $w_{IE}\tilde{X}_1$ of the n corrected clock phases $\tilde{x}_1^i$, where w_{IE} is a row of weights calculated from a 1-parameter least-squares problem with noise covariance matrix C_{11}, the phase-phase submatrix of C. After processing the measurements, one replaces the error covariance matrix C by the reduced version $C_G := SCS^{\mathrm{T}}$, where

$$S = \begin{bmatrix} I_n - 1_n w_{IE} & 0 \\ 0 & I_{2n} \end{bmatrix}$$

and 1_n is a column vector of n 1's. (The state vector has been rearranged so that all n phases come first.) It can be shown that C_G is the true covariance matrix of the corrected states with each $\tilde{x}_1^i$ replaced by its representation error $\tilde{x}_1^i - \bar{x}_1^G$.

The new covariance reduction is not transparent to the future phase estimates. When it is carried out after each set of measurements, the mean squared values of the following increments are minimized:

- $\tilde{x}_1^i(t) - \bar{x}_1^G(t-\delta)$, $i = 1$ to n, the corrected phases minus the previous IEM;
- $\bar{x}_1^G(t) - \bar{x}_1^G(t-\delta)$, the IEM increment.

In the absence of measurement noise, the covariance reduction reduces to the simple guillotine reduction of section V, the IEM $\bar{x}_1^G$ and all $\tilde{x}_1^i$ reduce to $\tilde{x}_1$, and the optimized increments reduce to $\tilde{x}_1(t) - \tilde{x}_1(t-\delta)$.

Figure 3 shows a simulated two-clock ensemble with measurement noise. For the natural Kalman scale, the corrected phases and Brown's IEM phase component all have the stability of the green clock, which is noisy in short term and quiet in long term. For the reduced Kalman scale, the IEM and corrected blue clock have essentially the same stabilities, while the corrected green clock overcomes the noise and joins the IEM for $\tau \geq 8$.

VIII. A Recent Result on Covariance Reduction

In a further comparison of the natural and reduced Kalman timescales [16], the author's colleagues M. Süß and D. Matsakis discovered that the covariance reductions of Brown (B) and the author (G) can be concatenated in either order to give a third, further reduced covariance matrix. Afterwards

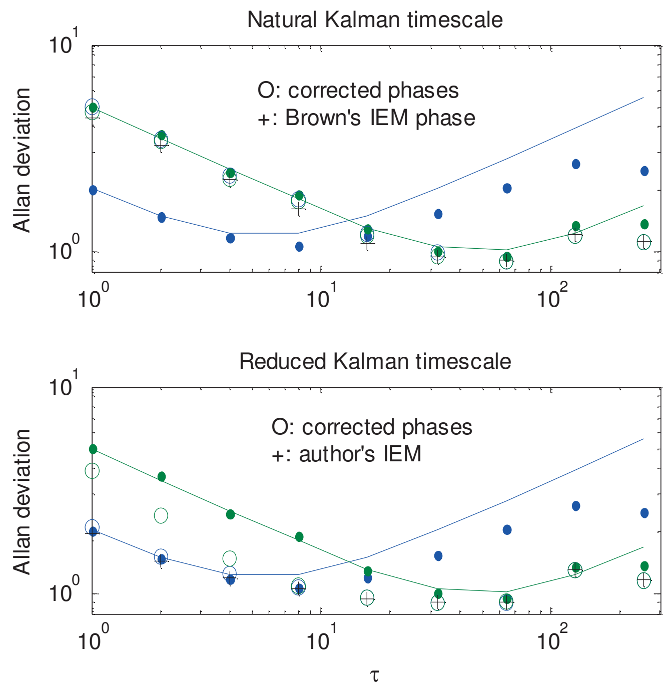

Fig. 3. A two-clock ensemble with measurement noise, showing the stability of the clocks, corrected clocks, Brown's IEM (phase component) and the author's IEM.

the author was able to prove this result mathematically. The reductions are illustrated in fig. 4, in which the matrices are rearranged so that all the phase components come first. For a starting covariance matrix C, the final matrix, $B(G(C)) = G(B(C))$ has the phase border of $G(C)$ and the frequency-drift submatrix of $B(C)$. The final matrix can be used in place of $G(C)$ for operating the reduced Kalman timescale.

IX. Concluding Remarks

By running simulated clock models, the author has found that that the natural method of constructing a timescale from a Kalman filter can be improved by a suitable reduction of the Kalman error covariance matrix, different from the transparent reduction of Brown. The natural Kalman timescale, when applied to dissimilar clocks, tends to minimize the long-term stability of the corrected clocks without regard to their short-term stability. The reduced Kalman timescale, which is designed to minimize the increments of the corrected clocks instead of their values at one time, seems to produce good results for stability over all available averaging times. So far, though, this is only an empirical observation from experiments in simulation playpens. One needs to gain experience with these algorithms on real clock ensembles while solving practical problems such as noise parameter estimation, clock dropouts, and changes of clock behavior.

The author thanks P. Tavella for inviting him to give this paper. This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract

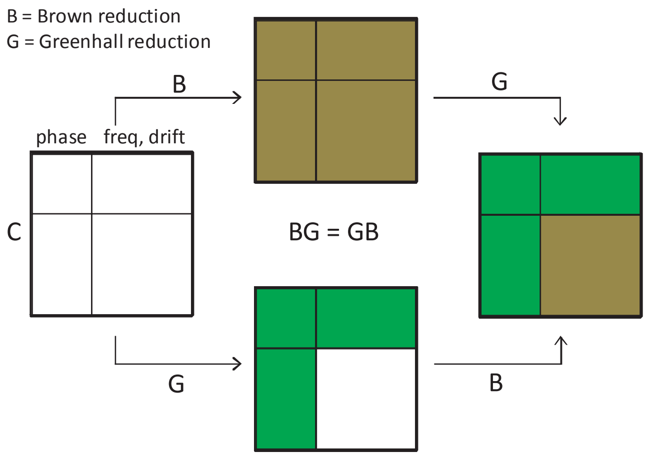

Fig. 4. Two commuting reduction operations, B and G, applied to a covariance matrix C.

with the National Aeronautics and Space Administration.

References

[1] R. H. Jones and P. V. Tryon, "Continuous time series models for unequally spaced data applied to modeling atomic clocks", SIAM J. Sci. Stat. Comput., vol. 8, pp. 71–81, 1987.

[2] J. W. Chaffee, "Relating the Allan variance to the diffusion coefficients of a linear stochastic differential equation model for precision oscillators", IEEE Trans. UFFC, vol. 34, pp. 655-658, Nov. 1987.

[3] C. Zucca and P. Tavella, "The clock model and its relationship with the Allan and related variances", IEEE Trans. UFFC, vol. 52, no. 2, pp. 289–296, Feb.. 2005.

[4] K. R. Brown Jr., "The theory of the GPS composite clock", Proc. ION GPS-91 Meeting, pp. 223–241, 1991.

[5] M. Weiss and T. Weissert, "AT2, a new time scale algorithm: AT1 plus frequency variance", Metrologia, vol. 28, pp. 65–74, 1991.

[6] S. R. Stein, "Advances in time-scale algorithms", Proc. 24th PTTI Meeting, pp. 289–302, 1992.

[7] S. R. Stein, "Time scales demystified", Proc. Joint IEEE IFCS and EFTF Meeting, pp. 223-227, 2003.

[8] C. A. Greenhall, "Kalman plus weights: a time scale algorithm", Proc. 33rd PTTI Meeting, pp. 445–454, 2001.

[9] C. A. Greenhall, "Addition of random run FM noise to the KPW time scale algorithm", IV Time Scale Algorithms Symposium, Sèvres, 2002.

[10] C. A. Greenhall, "Forming stable timescales from the Jones–Tryon Kalman filter", Metrologia, vol. 40, pp. S335–S341, 2003.

[11] M. A. Weiss, D. W. Allan, and T. K. Peppler, "A study of the NBS time scale algorithm", IEEE Trans. Instrum. Meas., vol. 38, pp. 631-635, 1989.

[12] J. A. Davis, C. A. Greenhall, and P. W. Stacey, "A Kalman filter clock algorithm for use in the presence of flicker frequency modulation noise", Metrologia, Vol. 42, pp. 1–10, 2005

[13] L. Galleani and P. Tavella, "On the use of the Kalman filter in timescales", Metrologia, vol. 40, pp. S326–S334, 2003.

[14] C. A. Greenhall, "A Kalman filter clock ensemble algorithm that admits measurement noise", Metrologia, vol. 43, pp. S311–S321, 2006.

[15] C. A. Greenhall, "A Kalman filter clock ensemble algorithm that admits measurement noise: corrections and update", Metrologia, vol. 44, pp. 491–494, 2007.

[16] M. Suess, D. Matsakis, and C. Greenhall, "Simulating future GPS clock scenarios with two composite clock algorithms", Proc. 42nd PTTI Meeting (to appear), 2010.