REPORT ON THE TIMESCALE ALGORITHM TEST BED AT USNO

S. R. Stein *
Ball Communication Systems Division
Broomfield, CO 80038-1235

G. A. Gifford
Naval Research Laboratory
Washington, DC 20375-5000

L. A. Breakiron
U. S. Naval Observatory
Washington, DC 20392-5100

Abstract

A test bed is described for the evaluation of timescale algorithms. Test results for two algorithms (**AT1** and **KAS-1**) using simulated data are presented. Both algorithms are optimum in the sense that, with the appropriate selection of clock weights, they are able to produce an output with minimum squared time prediction error in steady state for ensembles of clocks whose noises are all proportional to the noise of one of the members. The KAS-1 algorithm is also optimum during known transients such as startup, utilizes robust outlier detection, and has the property of time continuity. As a result of these improvements, KAS-1 is more suited to automated operation. Neither algorithm is optimum at all sampling times for ensembles of clocks having radically different noise processes. The KAS-1 algorithm and a clock simulator, both representing a new Kalman–filter approach, are documented herein.

INTRODUCTION

The Master Clock Upgrade Program at USNO is a SPAWAR– and USNO–funded project which tasks NRL and USNO to procure, monitor, and evaluate advanced frequency standards (trapped mercury ion devices and hydrogen masers) currently under development by others; devise and test a low–noise time–difference measurement system for all the cesium clocks and hydrogen masers at USNO; and develop a test bed for algorithms designed to generate an optimum timescale[1]. In order to accomplish the latter, a software package is being developed by NRL that will be an integral part of the data collection and management system at USNO. The current version of the software that comprises the test bed is written in "Rocky Mountain Basic" for the HP (Hewlett Packard) series 300 computer. The next version of the test bed package will be in the standard "C" language and will be less machine dependent. In addition to the standard clock analysis tools, the test bed will allow two algorithms to run simultaneously using the same data, which may be real or simulated. In the case of the latter, the output of an algorithm is compared to truth, where truth is known by definition.

As a joint project, USNO and NRL are making complementary contributions to the algorithm test bed. NRL's contribution is the experience in developing automated clock systems. USNO's contribution is the experience of and knowledge about maintaining a timescale.

NRL is developing a number of automated clock systems for field use designed to have the following features:

1. the systems are fully automated clock ensembles with real–time output;

2. they can be steered to UTC (USNO) automatically via GPS;

3. their timescale algorithms permit the easy addition or deletion of clocks; and

4. the algorithms generate a timescale more uniform than any single clock in the ensemble.

For our purposes, we will define a timescale as a multifaceted procedure defining the scale of time by means of an algorithm that usually entails filtering or smoothing the data from two or more clocks or clock ensembles, and possibly steering one or more of the clocks toward the ensemble or steering them or the ensemble itself toward a more stable reference. We define a clock ensemble to be a procedure by which a single output is obtained from two or more clocks. Generally an ensemble performs better than its individual members. For a set of N clocks with equal noise processes, an output can be obtained that is $\sqrt{N}$ more stable than a single clock, assuming the absence of intercorrelations and systematic errors. For a set of clocks whose noises are not proportional, the output is more complicated, as we shall see.

RESOURCES

As mentioned above, the test bed uses both simulated and real data. One source of the latter is the 2–picosecond–noise time measurement system developed by NRL for USNO that is connected to at least twenty high–performance cesium clocks, twelve hydrogen masers, and three mercury stored–ion devices (SIDs). The measurement system incorporates a phase–difference measurement subsystem developed by the first author while at NBS[2]. The cesiums and masers provide phase-difference measurements, while the SIDs give frequency measurements. This mix of measurement types, a challenge to most ensembling procedures, is handled quite naturally by a Kalman–filter approach. In addition to these new data, there is available the old USNO database of several years from two masers and about fifty cesiums.

NRL has available for this effort a database of phase measurements from three masers, at least ten (some high–performance and some GPS–type) cesiums, and several rubidium clocks. In addition, NRL has a means of measuring the USNO Master Clock to the 50–picosecond level via the carrier frequency of a local television station using a system similar to one devised in France[3]. This precise method of remote measurement will permit clocks several miles apart to be compared and allow the potential ensembling of clocks at both NRL and USNO.

The clock simulator employed in this test bed is a Kalman–filter type introduced in the next section. Simulated data have several advantages over real data:

1. any amount of data over any length of time can be generated;

2. data free of systematic errors or containing specific types of errors can be produced;

3. any number of clocks or combination of clock types can be modelled; and

4. most importantly, truth is known by definition.

To illustrate the latter, consider the simulation of a single clock, which involves the generation of a series of numbers that represent the accumulation of phase, usually over equal time intervals. The

numbers are produced with a clock model in mind and against which they may be verified statistically. To generate phase–difference data, two or more clocks are simulated simultaneously; one clock is then selected as the reference and its phase value at each measurement step is subtracted from the other clocks.

THE ALGORITHMS

The first algorithm tested was AT1, developed at NIST[4,5]. The source code was provided by NBS (now NIST) in 1987. It is an ad hoc procedure designed for a steady-state ensemble of clocks manifesting the types of noise typical of cesium clocks. Exponential filters are used in the frequency estimation and weight calculation.

The second algorithm, called the Kalman Aiding Sources Algorithm (KAS-1), developed in 1987 by Ball Aerospace for NRL, uses Kalman–filter methodology. The KAS-1 algorithm was first employed in an automated clock ensemble at the Master Control Station of the Global Positioning System (GPS) at the Consolidated Space Operations Center (CSOC), Falcon AFB and was presented at the Third International Time Scale Algorithm Symposium held in Torino in 1988[6]. This paper provides a more complete description as well as operational and simulated performance data. The KAS-1 approach is now being tested by NIST for use with AT1 and their results are described in the paper by Weiss and Weissert in these same proceedings. The KAS-1 algorithm was the first successful use of a Kalman filter in connection with clock ensembles, and this approach has been enhanced for use at USNO.

KALMAN APPROACH TO SIMULATING AND ENSEMBLING CLOCKS

Background

The ensembling method employed in KAS-1 uses one clock to form an initial estimate of time and uses each additional clock to improve or 'aid' the initial estimate. The same approach is used in the NIST AT1 timescale. The work described here improves and extends the AT1 algorithm in several areas: optimum filtering of all clock types, automatic startup and clock addition, robust outlier processing, time and frequency step detection, real time clock parameter estimation, and adaptive modelling.

The KAS-1 algorithm utilizes Kalman filters to provide the necessary state estimation and forecasting functions, resulting in estimates which are optimum in the minimum squared–error sense both in steady state and during transient conditions such as turn–on. Any mix of clocks can be included provided that the correct models and covariances are used. The use of the Kalman filter is a computational device and any other method of minimum squared–error state estimation based on the same assumptions and system dynamics would produce an identical result.

In contrast, the NIST AT1 algorithm uses fixed length exponential filters to perform frequency estimation. It is optimum only in steady state and requires external clock calibration for startup and addition of clocks to the ensemble.

Definition of the KAS–1 Ensemble

Each of the N clocks provides an independent forecast of the time of clock r, the reference clock, with respect to the ensemble at time $t + \delta$. These individual forecasts are called 'primitive' forecasts of the reference clock time. The primitive forecast of the time of the reference clock at $t + \delta$ using clock i is

$$u_{re}^i(t + \delta | t + \delta) = u_{ie}(t + \delta | t) + u_{ri}(t + \delta) \tag{1}$$

271

where $u_{ie}(t + \delta|t)$ is the forecast of the time of clock i with respect to the ensemble at time $t + \delta$ based upon the true state of clock i with respect to the ensemble at time t and $u_{ri}(t + \delta)$ is the clock r vs. clock i time difference.

The KAS–1 algorithm utilizes the natural definition

$$u_{re}(t + \delta) \equiv \sum_{i=1}^{N} a_i(t + \delta) u_{re}^i(t + \delta|t + \delta). \tag{2}$$

The weights, $a_i(t)$, may be chosen in any way subject only to the restriction that the sum of the weights is one. In the absence of noise, the ensemble definition is deterministic. However, in the presence of clock and measurement noise, the computational problem is to find the minimum squared–error estimate of the time of the reference clock with respect to the ensemble.

Computational Methodology for the KAS–1 Algorithm

One method of estimating ensemble time is to use the definition of Eq. 2 directly. Taking the estimate of both sides, one obtains

$$\hat{u}_{re}(t + \delta|t + \delta) = \sum_{i=1}^{N} a_i(t + \delta)[\hat{u}_{ie}(t + \delta|t) + \hat{u}_{ri}(t + \delta)]. \tag{3}$$

The estimate of the time of clock i with respect to the ensemble at time $t + \delta$ based on measurements through time t, the first term in the square brackets on the right side of Eq. 3, is calculated from the results of the last computation of the ensemble. Thus

$$\hat{u}_{ie}(t + \delta|t) = \hat{x}_{ie}(t|t) + \delta \hat{y}_{ie}(t|t) + \frac{\delta^2}{2} \hat{w}_{ie}(t|t) \tag{4}$$

where x is the phase before perturbation by white phase noise, y is the frequency and w is the frequency aging. The second term in the square brackets on the right side of Eq. 3 is the minimum square–error estimate of the clock pair difference state. This estimate is easily computed using the appropriate Kalman filter.

Kalman Filter for Computing Clock Difference States

The n parameters which are to be estimated are formed into an n dimensional state vector $\vec{x}(t)$. The system evolves from time t to time $t + \delta$ according to

$$\vec{x}(t + \delta) = \mathbf{\Phi}(\delta)\vec{x}(t) + \mathbf{\Gamma}\vec{s}(t + \delta|t) + \mathbf{\Phi}(\delta)\vec{p}(t), \tag{5}$$

where the $n \times n$ dimensional state transition matrix $\mathbf{\Phi}(\delta)$ embodies the system model, the n dimensional vector $\mathbf{\Gamma}\vec{s}(t + \delta|t)$ contains the noise inputs to the system during the interval from t to $t + \delta$, and the n dimensional vector $\vec{p}(t)$ contains the control inputs made at time t. The state transition matrix is assumed to depend on the length of the interval, but not on the origin. Each element of $\vec{s}(t + \delta|t)$ is Normally distributed with zero mean and is uncorrelated in time, thereby generating a random walk in the elements of the state vector. The observation vector $\vec{z}(t)$ is described by the measurement equation

$$\vec{z}(t) = \mathbf{H}(t)\vec{x}(t) + \vec{v}(t). \tag{6}$$

The r observations made at time t are related linearly to the n elements of the state vector by the $r \times n$ dimensional measurement matrix $\mathbf{H}(t)$ and the r dimensional white noise vector $\vec{v}(t)$. Kalman and

Bucy defined a recursive procedure for estimating the next state, which requires the mean squared-error of the estimates from the true state to be minimum.

The error in the estimate of the state vector after the measurement at time t is $\hat{\vec{x}}(t|t) - \vec{x}(t)$ and the error covariance matrix is defined to be

$$\mathbf{P}(t|t) = E\left\{ \left[\hat{\vec{x}}(t|t) - \vec{x}(t)\right] \left[\hat{\vec{x}}(t|t) - \vec{x}(t)\right]^T \right\}. \tag{7}$$

The diagonal elements of this $n \times n$ matrix are the variances of the estimates of the components of $\vec{x}(t)$ after the measurement at time t. Next, the error covariance matrix just prior to the measurement at time $t + \delta$ is defined as

$$\mathbf{P}(t + \delta|t) = E\left\{ \left[\hat{\vec{x}}(t + \delta|t) - \vec{x}(t + \delta)\right] \left[\hat{\vec{x}}(t + \delta|t) - \vec{x}(t + \delta)\right]^T \right\}. \tag{8}$$

Finally, $\mathbf{R}(t)$ is the covariance matrix of the measurement noise and $\mathbf{Q}(t + \delta|t)$ is the covariance matrix of the system noise generated during the interval from t to $t + \delta$

$$\mathbf{R}(t) = E\left[\vec{v}(t)\vec{v}(t)^T \right] \tag{9}$$

$$\mathbf{Q}(t + \delta|t) = E\left[\vec{s}(t + \delta|t)\vec{s}(t + \delta|t)^T \right]. \tag{10}$$

The error covariance matrix evolves according to the system model

$$\mathbf{P}(t + \delta|t) = \mathbf{\Phi}(\delta)\mathbf{P}(t|t)\mathbf{\Phi}(\delta)^T + \mathbf{\Gamma}\mathbf{Q}(t + \delta|t)\mathbf{\Gamma}^T. \tag{11}$$

The new estimate of the state vector depends on the previous estimate and the current measurement

$$\begin{aligned}
\hat{\vec{x}}(t + \delta|t + \delta) &= \mathbf{\Phi}(\delta)\hat{\vec{x}}(t|t) + \mathbf{\Phi}(\delta)\vec{p}(t) \\
&\quad + \mathbf{K}(t + \delta)\left[\vec{z}(t + \delta) - \mathbf{H}(t + \delta)\mathbf{\Phi}(\delta)\hat{\vec{x}}(t|t) - \mathbf{H}(t + \delta)\mathbf{\Phi}(\delta)\vec{p}(t) \right],
\end{aligned} \tag{12}$$

where the gain matrix, $\mathbf{K}(t + \delta)$, determines how heavily the new measurements are weighted. The optimum or Kalman gain, $\mathbf{K}_{opt}$, is determined by minimizing the "square of the length of the error vector," i.e., the sum of the diagonal elements (the trace) of the error covariance matrix

$$\mathbf{K}_{opt}(t + \delta) = \mathbf{P}(t + \delta|t)\mathbf{H}(t + \delta)^T \left[\mathbf{H}(t + \delta)\mathbf{P}(t + \delta|t)\mathbf{H}(t + \delta)^T + \mathbf{R}(t + \delta) \right]^{-1}. \tag{13}$$

Finally, the updated error covariance matrix is

$$\begin{aligned}
\mathbf{P}(t + \delta|t + \delta) &= \left[\mathbf{I} - \mathbf{K}(t + \delta)\mathbf{H}(t + \delta)\right]\mathbf{P}(t + \delta|t)\left[\mathbf{I} - \mathbf{K}(t + \delta)\mathbf{H}(t + \delta)\right]^T \\
&\quad + \mathbf{K}(t + \delta)\mathbf{R}(t + \delta)\mathbf{K}(t + \delta)^T
\end{aligned} \tag{14}$$

where $\mathbf{I}$ is the identity matrix. If the optimum filter gain is used, Eq. 14 reduces to a simpler form:

$$\mathbf{P}(t + \delta|t + \delta) = \left[\mathbf{I} - \mathbf{K}_{opt}(t + \delta)\mathbf{H}(t + \delta)\right]\mathbf{P}(t + \delta|t). \tag{15}$$

Equations 11 through 15 define the Kalman filter, and so defined it is an optimal estimator in the minimum squared error sense. Each application of the Kalman recursion yields an estimate of the state of the system which is a function of the elapsed time since the last filter update. Updates may occur at any time. In the absence of observations, the updates are called forecasts. Simultaneous observations may be processed in parallel or serially, in any order, if the measurement noise is uncorrelated with

the process noise. The interval between updates, δ, is arbitrary and is specifically not assumed to be constant.

The estimates of the clock states relative to clock r are obtained from $N - 1$ independent Kalman filters. The four dimensional state vectors are:

$$\vec{x}_{ri}(t) = \begin{bmatrix} u_{ri}(t) \\ x_{ri}(t) \\ y_{ri}(t) \\ w_{ri}(t) \end{bmatrix}. \tag{16}$$

Every clock pair has the same state transition matrix

$$\Phi(\delta) = \begin{bmatrix} 0 & 1 & \delta & \delta^2/2 \\ 0 & 1 & \delta & \delta^2/2 \\ 0 & 0 & 1 & \delta \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{17}$$

and Γ matrix

$$\Gamma(\delta) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{18}$$

The system covariance matrices are:

$$\mathbf{Q}^{ir}(t+\delta|t) = \tag{19}$$

$$\begin{bmatrix} S_{\beta'}^{ir}(t)f_h & 0 & 0 & 0 \\ 0 & S_{\xi}^{ir}(t)\delta + S_{\mu}^{ir}(t)\delta^3/3 + S_{\varsigma}^{ir}(t)\delta^5/20 & S_{\mu}^{ir}(t)\delta^2/2 + S_{\varsigma}^{ir}(t)\delta^4/8 & S_{\varsigma}^{ir}(t)\delta^3/6 \\ 0 & S_{\mu}^{ir}(t)\delta^2/2 + S_{\varsigma}^{ir}(t)\delta^4/8 & S_{\mu}^{ir}(t)\delta + S_{\varsigma}^{ir}(t)\delta^3/3 & S_{\varsigma}^{ir}(t)\delta^2/2 \\ 0 & S_{\varsigma}^{ir}(t)\delta^3/6 & S_{\varsigma}^{ir}(t)\delta^2/2 & S_{\varsigma}^{ir}(t)\delta \end{bmatrix}$$

where the clock pair spectral densities are the sum of the individual contributions from each of the clocks:

$$S^{ir} = S^i + S^r. \tag{20}$$

The white phase measurement noise is given by the measurement model:

$$\vec{z}_{ri} = \mathbf{H}\vec{x}_{ri} + \vec{v}_{ri}, \tag{21}$$

where each time measurement is described by the same 4×1 row matrix

$$\mathbf{H}_{ri} = (\ 1 \quad 0 \quad 0 \quad 0\ ). \tag{22}$$

The updated difference states are given by Eq. 12, but only the time state is used in the following steps of the ensemble calculation.

Calculation of the Weighted Average of the Times of the Clocks

Although it is not necessary, a Kalman filter is used to calculate the 'weighted average' of the N primitive estimates of the time of clock r. Kalman filters are normally applied to dynamic systems, but they are equally relevant to the static problem of 'averaging' the N estimates of the time of clock

r at sample time t. The choice of the Kalman filter methodology is motivated by the fact that it deals with the clock weights in an easy and natural manner and automatically provides scale factors for use in outlier deweighting. The deweighting scheme is similar to one developed by Percival[7] for maximum likelihood estimation, but it has been adapted for use with the Kalman methodology. It is desirable for the resulting estimator (with outlier deweighting) to be both efficient and robust. For certain types of outliers, the method used here is both more efficient and more robust than rejection rule methods.

The state transition matrix is a scalar and is equal to one since the system does not evolve between 'observations'. The 'measurement' model is

$$u_{re}(t) = u_{re}^i(t|t) + v_i(t) \tag{23}$$

where the random shocks $v_i(t)$ have variances $\sigma_i^2(t)$ and are used to set the weights of the clocks. For example, $v_i(t) = v(t)$ independent of the clock number produces equal weighting. On the other hand, setting $v_i(t)$ equal to the estimated time prediction error of clock i since the last ensemble calculation produces an ensemble with minimum time prediction error. Note that the latter operating condition is not necessarily the best since the clocks may not be perfectly modeled.

It is possible to solve for the Kalman gain in closed form:

$$K_i(t) = \frac{\frac{1}{\sigma_i^2(t)}}{\sum_{j=1}^i \frac{1}{\sigma_j^2(t)}}. \tag{24}$$

However, the full Kalman recursion is more suited to the deweighting of outliers. Since a scalar is being estimated, $\mathbf{\Phi} = \mathbf{H} = 1$, $\mathbf{Q} = 0$, and $\mathbf{R} = \sigma_j^2(t)$. In order to deweight outliers, a robust initial guess is required for the time of clock r with respect to the ensemble after the measurement at time $t + \delta$. The median[1] of the available estimates is a good choice since it is not unduly affected by outliers.

$$\hat{u}_{re}(t + \delta|t + \delta)_1 = \text{median}\left\{u_{re}^i(t + \delta|t + \delta)\right\} \tag{25}$$

This estimate is subsequently refined by using the information in the remaining primitive estimates. The primitive estimates are sorted according to their deviation from the median so that clock $I(k)$ is the kth nearest the median and $I(1)$ is the median. To start the Kalman recursion one needs to initialize the state and the covariance matrices. Having just completed the estimate of the ensemble time for the measurements taken at time t, the ensemble calculation for time $t + \delta$ commences by setting the initial estimate of the time of clock r vs. the ensemble equal to the 'primitive estimate' using clock $I(1)$. The starting point is important only for outlier deweighting. A 'robust' initial estimate for the time of clock r is required and using one of the primitive estimates avoids biasing the final estimate. The first estimate must be robust because it cannot be deweighted. This implies no loss of generality since it is not possible to detect an outlier with only one clock and it is not possible to assign responsibility for an outlier time difference measurement with only two clocks. That is, there is one more cycle in the Kalman recursion than is needed for outlier deweighting. The recursion begins by setting the initial state estimate:

$$\hat{u}_{re}(t + \delta|t + \delta)_1 = \hat{u}_{re}^{I(1)}(t + \delta|t + \delta).$$

---

[1] When the number of points is odd, the number of observations larger than the median equals the number of observations smaller than the median. When the number of observations is even, the median is one of the two central observations. Other definitions of the median are sometimes used.

The covariance matrix prior to processing the measurement of clock $I(2)$ is

$$P(I(2)|I(1)) = \sigma^2_{I(1)}$$

and the optimum Kalman gain is

$$K_{I(2)} = \frac{P(I(2)|I(1))}{P(I(2)|I(1)) + \sigma^2_{I(2)}}.$$

At this point in the calculation, the standard Kalman methodology must be modified to account for the possible presence of outlier observations. If primitive estimate $I(2)$ does not come from the distribution included in the physical model, then it should not be used to determine the Kalman state estimate. In order to preserve the continuity of the state estimates, measurements should not be rejected precipitously, rather the Kalman gain should be reduced as the measurement departs from expectation. A non–optimum Kalman gain is calculated from

$$K'_{I(2)} = \frac{\psi_{HA}[q_{I(2)}]}{q_{I(2)}} K_{I(2)} \tag{26}$$

where

$$\psi_{HA}(q) = \begin{cases} q & \text{if} & |q| \leq a \\ \frac{a}{b-a}\left[\frac{bq}{|q|} - q\right] & \text{if} & a < |q| \leq b \\ 0 & \text{if} & |q| > b \end{cases} \tag{27}$$

is Hampel's ψ function and

$$q_{I(2)} = \frac{\hat{u}_{re}^{I(2)}(t+\delta|t+\delta) - \hat{u}_{re}(t+\delta|t+\delta)_1}{s_2}. \tag{28}$$

The scale factor, s_2, is discussed later. The second step in the Kalman recursion is completed by using the modified Kalman gain to calculate the updated covariance matrix according to the formula for non–optimum Kalman gain

$$P[I(2)|I(2)] = \left[1 - K'_{I(2)}\right]^2 P[I(2)|I(1)] + \left[K'_{I(2)}\right]^2 \sigma^2_{I(2)} \tag{29}$$

and to calculate the new estimate of the time of clock 1 with respect to the ensemble

$$\hat{u}_{re}(t+\delta|t+\delta)_2 = \left(1 - K'_{I(2)}\right)\hat{u}_{re}(t+\delta|t+\delta)_1 + K'_{I(2)}\hat{u}_{re}^{I(2)}(t+\delta|t+\delta). \tag{30}$$

The recursion continues in this way until all the measurements are processed. The computation for clock $I(j)$ starts by setting

$$P[I(j)|I(j-1)] = P[I(j-1)|I(j-1)].$$

The first time the recursion is performed, each of the constants s_j is set equal to the maximum of the σ_j. This prevents the high noise level of the median from causing a very quiet clock to be downweighted. Next, the computation is repeated setting the initial estimate of the time of the reference with respect to the ensemble equal to the primitive estimate closest to the result of the previous computation and setting s_j equal to the second largest σ_j or its own prediction error, whichever is larger. The recursion is repeated until each estimate is processed using the σ_i of the corresponding clock. If any of the clocks were downweighted in this process, then the new estimate of the time of clock r with respect

to the ensemble is compared with the previous estimate and the process is repeated until the change in the estimate between iterations is negligible. When there are no outliers, this recursive solution is identical to Eq. 3. In the presence of outliers, it is still a weighted average of the estimates from the individual clocks. The weights of the clocks are given by

$$a_{I(j)} = K'_{I(j)} \prod_{i=j+1}^{N} (1 - K'_{I(i)})$$

(31)

where $K'_{I(1)} \equiv 1$. To preserve the reliability of the ensemble, one usually limits the weights of each of the clocks to some maximum value, a_{max}. A deweighting factor, κ, may be calculated to accomplish this goal by starting at index $I(N)$ and proceeding to index $I(1)$ so that $\kappa K'$ are new weights which do not exceed the limit. Each time the Kalman recursion is used to estimate the time of clock r, the previously calculated κ is used. The process rapidly converges to a stable value.

Finally, the times of the remaining clocks are computed from the time of clock r and the estimated time differences. Thus

$$\hat{u}_{ie}(t + \delta | t + \delta) = \hat{u}_{re}(t + \delta | t + \delta) - \hat{u}_{ri}(t + \delta | t + \delta).$$

(32)

Calculation of the Frequency and Frequency Aging States

The time of clock i relative to the ensemble is used as input to a Kalman filter that estimates the frequency and frequency aging of clock i relative to the ensemble. This filter uses the state transition matrix of Eq. 17. It is assumed that the ensemble is so large that its contribution to the noise can be accounted for to first order only. Based on this assumption of a large ensemble, the system covariance matrix is:

$$\mathbf{Q}^{ie}(t + \delta | t) = \mathbf{Q}^{i}(t + \delta | t).$$

(33)

and

$$\Gamma(t)^{ie} = \begin{bmatrix} 1 - a_i(t) & 1 - a_i(t) & 0 & 0 \\ 0 & 1 - a_i(t) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(34)

The change in Γ from the case of a single clock or clock pair reflects the fact that each clock is a member of the ensemble with respect to which it is measured.

The measurement model is noiseless since the calculated times of the N clocks are used as pseudomeasurements. This procedure guarantees that the Kalman filter reproduces the phase estimates obtained directly from the ensemble definition while simultaneously estimating consistent values of the frequency and frequency aging.

Adaptive Modelling
Parameter Estimation

A variance analysis technique compatible with irregular observations has been developed[8]. The variance of the innovation sequence of the Kalman filter is analyzed to provide estimates of the parameters of the filter. The result is

$$E\left[\vec{v}(t + \delta)\vec{v}(t + \delta)^T\right] = \mathbf{H}(t + \delta)\mathbf{P}(t + \delta | t)\mathbf{H}(t + \delta)^T + \mathbf{R}(t + \delta).$$

(35)

Like the Allan variance analysis, which is performed on the unprocessed measurements, the innovation variance analysis requires only a limited memory of past data. However, the forecasts provided by the Kalman filter allow the computation to be performed at arbitrary intervals once the algebraic form of the innovation variance has been calculated. Adaptive modelling begins with an approximate Kalman gain, $\hat{\mathbf{K}}$. As the state estimates are computed, the variance of the innovations on the left side of Eq. 35 is also computed. The right–hand side of this equation is written in terms of the actual filter element values (covariance matrix elements) and the theoretical parameters. Finally, the equations are inverted to produce improved estimates for the parameters.

Using the autocovariance function to solve for the parameters is inappropriate here, because the autocovariance function is highly correlated from one lag to the next and the efficiency of data utilization is therefore small. Instead, only the autocovariance of the innovations for zero lags, *i.e.*, the variance of the innovations, is used. The variance is given by

$$
\begin{aligned}
E\left[\vec{\nu}_{ie}(t+\delta)\vec{\nu}_{ie}(t+\delta)^T\right] \;=\; & \mathbf{P}^i_{11}(t|t) + 2\delta\mathbf{P}^i_{12}(t|t) + \delta^2\mathbf{P}^i_{13}(t|t) \\
& + \delta^3\mathbf{P}^i_{23}(t|t) + \delta^2\mathbf{P}^i_{22}(t|t) + \frac{\delta^4}{4}\mathbf{P}^i_{33}(t|t) \\
& + [S^i_\xi(t)\delta + S^i_{\beta'}(t)f_h][1 - a_i(t)]^2 + S^i_\mu(t)\frac{\delta^3}{3} + S^i_\varsigma(t)\frac{\delta^5}{20} + \sigma^2_{v_{xi}}(t).
\end{aligned}
\tag{36}
$$

It is assumed that the oscillator model contains no hidden noise processes. This means that each noise in the model is dominant over some region of Fourier frequency space. The principle of parsimony encourages this approach to modelling. Inspection of Eq. 36 leads to the conclusion that each of the parameters dominates the variance of the innovations in a unique region of prediction time interval, δ, making it possible to obtain high–quality estimates for each of the parameters through a process of bootstrapping.

For each parameter to be estimated, a Kalman filter is computed using a subset of the data chosen to maximize the number of predictions in the interval for which that parameter makes the dominant contribution to the innovations. The filters are designated 0 through 4, starting with 0 for the main state estimation filter which runs as often as possible. Each innovation is used to compute a single–point estimate of the variance of the innovations for the corresponding δ. Substituting the estimated values of the remaining parameters, Eq. 36 is solved for the dominant parameter, and the estimate of that parameter is updated in an exponential filter of appropriate length. If the minimum sampling interval is too long, it may not be possible to estimate one or more of the parameters. However, there is no deleterious consequence of this situation, since a parameter that cannot be estimated is not contributing appreciably to the prediction errors. It is not necessary to use separation of variances to estimate individual clock parameters since the Γ matrix takes into account the first order clock–ensemble correlations.

The KAS–1 algorithm is adaptive in two ways. First, the clock noise parameters are updated after each cycle of the recursion and used during the next cycle of the recursion. Second, the clock weights are updated after each cycle based on the real time parameter estimates or the measured time prediction errors. For minimum time prediction errors, the variances (inverse weights) needed for the calculation of the time of the reference clock with respect to the ensemble are

$$
\sigma^2_i = S^i_{\beta'}(t)f_h + S^i_\xi(t)\delta + S^i_\mu(t)\delta^3/3 + S^i_\varsigma(t)\delta^5/20.
\tag{37}
$$

Kalman Clock Simulation Algorithm

The algorithm presented here may be used to compute the simulated state of a precision clock at any time in the future based on the current state of the clock and appropriate noise inputs. White phase measurement noise, white phase additive noise, white frequency noise, random walk frequency noise, and random walk frequency aging noise may all be included. Although the algorithm is quite simple, it has several advantages over the ARIMA technique described by Barnes[9]. The time interval between observations is included explicitly so that a time series of observations separated by varying intervals may be created. In addition, the parameters are easily calculated from the power spectral densities of the noise. These physical parameters are the same ones required for the Kalman analysis of the clock. The final advantage is that the simulated clock noise for each state is appropriately correlated with the noise on other states, as should be the case since the physical model for a clock is an integrator.

The simulator is based on the same equation of motion for a clock, Eq. 5, used in Kalman filter analysis. Each element of $\vec{s}(t + \delta|t)$ is Normally distributed with zero mean and is uncorrelated in time. The problem is to generate simulated noise vectors, $s'(t + \delta|t)$, having the desired covariance, given in Eq. 19. The spectral densities which appear in the system covariance matrix may be written in terms of the standard oscillator noise coefficients. The values and units are shown in Table 1.

Table 1: Relationship between spectral densities and h coefficients

Spectral density	h coefficient	units
$S_{\beta'}(t)$	$(2\pi)^{-2}h_2(t)$	seconds^3
$S_{\xi}(t)$	$h_0(t)$	seconds
$S_{\mu}(t)$	$(2\pi)^2 h_{-2}(t)$	seconds^{-1}
$S_{\varsigma}(t)$	$(2\pi)^4 h_{-4}(t)$	seconds^{-3}

The goal is to derive a solution for the noise vector, $\vec{s}$, in the form

$$\vec{s} = \mathbf{\Lambda}\vec{r},\tag{38}$$

where $\vec{r}$ is a four dimensional vector of Normal deviates with unity variance whose elements are uncorrelated with each other and uncorrelated in time. Multiplying this equation by its transpose and taking the expectation value, one finds

$$Q(t + \delta|t) = \mathbf{\Lambda}\mathbf{\Lambda}^T.\tag{39}$$

Since the elements of the state vector are successive integrals, $\mathbf{\Lambda}$ is an upper right half triangular matrix. Equating the corresponding elements of the previous equation, one finds

$$\mathbf{\Lambda}(t + \delta|t) = \tag{40}$$

$$
\begin{bmatrix}
\sqrt{S_{\beta'}(t)f_h} & 0 & 0 & 0 \\
0 & \sqrt{S_{\xi}(t)\delta + \frac{S_{\mu}(t)\delta^3}{12} + \frac{S_{\varsigma}(t)\delta^5}{720}} & \frac{\delta}{2}\sqrt{S_{\mu}(t)\delta + \frac{S_{\varsigma}(t)\delta^3}{12}} & \frac{\delta^2}{6}\sqrt{S_{\varsigma}(t)\delta} \\
0 & 0 & \sqrt{S_{\mu}(t)\delta + \frac{S_{\varsigma}(t)\delta^3}{12}} & \frac{\delta}{2}\sqrt{S_{\varsigma}(t)\delta} \\
0 & 0 & 0 & \sqrt{S_{\varsigma}(t)\delta}
\end{bmatrix}.
$$

The algorithm is easily mechanized. A random number generator function normally produces deviates uniformly distributed over the unit interval. They may be converted to Normally distributed deviates using one of the standard procedures. Four Normal deviates are then combined to form the noise vector $\vec{r}$ and multiplied by Λ to produce the $\vec{s}$ vector needed to evaluate Eq. 5 for the next clock state. The element $u(t)$ is the desired simulated clock time state. The simulated clock measurement is obtained by adding white measurement noise, $v(t)$, to $u(t)$.

$$z(t) = u(t) + v(t) \tag{41}$$

TEST RESULTS

The two algorithms were run simultaneously in order to intercompare them. Only the 1987 version of the KAS–1 algorithm (i.e. without automatic parameter estimation and mixed–mode measurement capability) was available for testing.

Figure 1 displays the difference between the two algorithms, in the sense AT1 - KAS-1, for an ensemble of eleven equally weighted clocks with no data rejection. The two algorithms are seen to be essentially identical in steady state on the long term. However, as mentioned above, AT1 is not designed for the natural handling of transients, while KAS–1 is still optimal in such situations.

In Figure 2, AT1 was allowed to exercise its 3–standard–deviation rejection criterion, while the KAS–1 algorithm was forced to accept all the data. As a result, on the short term AT1 is shown to suffer from small discontinuities. This effect is avoided with the KAS–1 algorithm (in its usual form) by its use of progressive deweighting of outlying data based on Hampel's ψ function[7].

Figure 3 is a plot of Allan deviation vs. sampling time τ for one of the eleven clocks (no data rejection), assuming it to have the white FM and random–walk FM noises characteristic of a conventional (HP5061A model) high–performance cesium.

Figure 4 is a similar plot for the ensemble when the KAS–1 algorithm and its rejection criterion are employed. Each of the clocks has equal white FM and equal random–walk FM noises; hence, they are equally weighted. There is a $\sqrt{N}$ improvement in the ensemble's collective stability, and the τ of minimum Allan variance is constant.

Figures 5 and 6 show simulated data of contrasting clock types. Figure 5 shows the frequency stability of VLG11–type hydrogen maser without a long–term frequency drift and fig. 6 is a similar plot of a SID or very stable (e.g. a De Marchi–tuned) cesium (no random walk of frequency). The clock types represented in figs. 3, 5, and 6 are used in an ensemble of unequal clocks similar to that which will soon comprise the USNO ensemble (60 percent cesiums, 30 percent masers, 10 percent SID's). This unequal clock set is used as the input to two algorithms with contrasting weighting philosophies.

Figure 7 is the Allan deviation of an algorithm (either AT1 or KAS–1) which automatically determines the weights of individual clocks in the ensemble according to the inverse time prediction errors at the sampling time; whereas fig. 8 is the deviation of an algorithm that forces all clocks to be weighted equally. Comparing the two figures, representing eight years of data, the weighted algorithm is significantly better in the short–term, whereas the equally weighted algorithm is better in the long–term. Future versions of the KAS algorithm will produce optimum results at all sampling times by completely accounting for the clock–ensemble correlations.

Figures 9 and 10 are analogous to Figs. 7 and 8, respectively, but display 25 years of data. The same

effects are noted, with the long-term degradation of the weighted algorithm being even more evident. Clearly, neither the AT1 nor the KAS-1 algorithm when using a weighting scheme is optimal for sets of dissimilar clocks. A new algorithm, KAS-2, has been developed by Ball Aerospace that is optimum for all clock types and all sample times. It will be tested in the Algorithm Test Bed in 1990.

CONCLUSION

The use of simulated data is shown to be a powerful method of comparing different clock models and ensembling algorithms. A Kalman filter–based algorithm, such as KAS-1, can use unequally spaced data and can handle transients well, unlike an ad hoc procedure like the AT1 (1987) algorithm. The KAS-1 algorithm is able to handle a mix of time and frequency measurements quite naturally.

Data rejection with AT1 leads to small discontinuities in the timescale, unlike the case for KAS-1. While both algorithms are optimal in steady state for clocks having proportional noises, they are not optimal for sets of significantly different clocks. The stability of a real clock ensemble would be further degraded by clock intercorrelations and systematic errors, so that, if the clock weights are based only on Allan variances, the results may not be as good as those employed simply using equal weights[10].

As the project continues, more study will be made of simulated data; real clock data will be analyzed; and the improvements to the KAS-1 algorithm suggested herein will be implemented and tested.

References

[1] Gifford, G. A. and Wheeler, P., "Report on the master clock upgrade program at USNO," Proceedings of the 43rd Annual Symposium on Frequency Control, 31 May–2 June, 1989, Denver, CO, pp. 158–161.

[2] Stein, S. R., Glaze, D., Levine, J., Gray, J., Hilliard, D., Howe, D., and Erb, L., "Performance of an automated high accuracy phase measurement system," Proceedings of the 36th Annual Symposium on Frequency Control, 2–4 June, 1982, Philadelphia, PA, pp. 314–320.

[3] Gabry, A., Faucheron, G., and Dubois, B., "Distant comparison of stable frequency standards by means of the transmission of a beat note between the carrier of a TV broadcast signal and a frequency synthesized from the frequency standards," Proceedings of the 31st Annual Symposium on Frequency Control, 1–3 June, 1977, Atlantic City, NJ, pp. 499–502.

[4] Allan, D. W., Gray, J. E., and Machlan, H. E., "The National Bureau of Standards Atomic Time Scale: Generation, Stability, Accuracy, and Accessibility," chap. 9, Time and Frequency: Theory and Fundamentals (B. E. Blair, ed.), 1974, National Bureau of Standards Monograph No.140, pp. 214–229.

[5] Varnum, F. B., Brown, D. B., Allan, D. W., and Peppler, T. K., "Comparison of Time Scales Generated with the NBS Ensembling Algorithm," Proceedings of the 19th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 1–3 December, 1987, Redondo Beach, CA, pp. 13-23.

[6] Stein, S. R., "Kalman Ensembling Algorithm: Aiding Sources Approach," Proceedings of the Third International Time Scale Algorithm Symposium, 12–13 September, 1988, Turin, Italy, pp. 345–357.

[7] Percival, D. B., "Use of robust statistical techniques in time scale formulation," presented at the Second International Time Scale Algorithm Symposium, 23–25 June, 1982, Boulder, CO, report

on USNO contract N70092–82–M–0579.

[8] Stein, S. R., "Kalman Filter Analysis of Precision Clocks with Real–Time Parameter Estimation," Proceedings of the 43rd Annual Symposium on Frequency Control, 31 May –2 June, 1989, Denver, CO, pp. 232–236.

[9] Barnes, J. A., "Simulation of Oscillator Noise," Proceedings of the 38th Annual Symposium on Frequency Control, 29 May – 1 June, 1984, Philadelphia, PA, pp. 319–326.

[10] Breakiron, L. A., "The effects of data processing and environmental conditions on the accuracy of the USNO timescale," Proceedings of the 20th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 29 November–1 December, 1988, Tysons Corner/Vienna, VA, pp. 221–236.
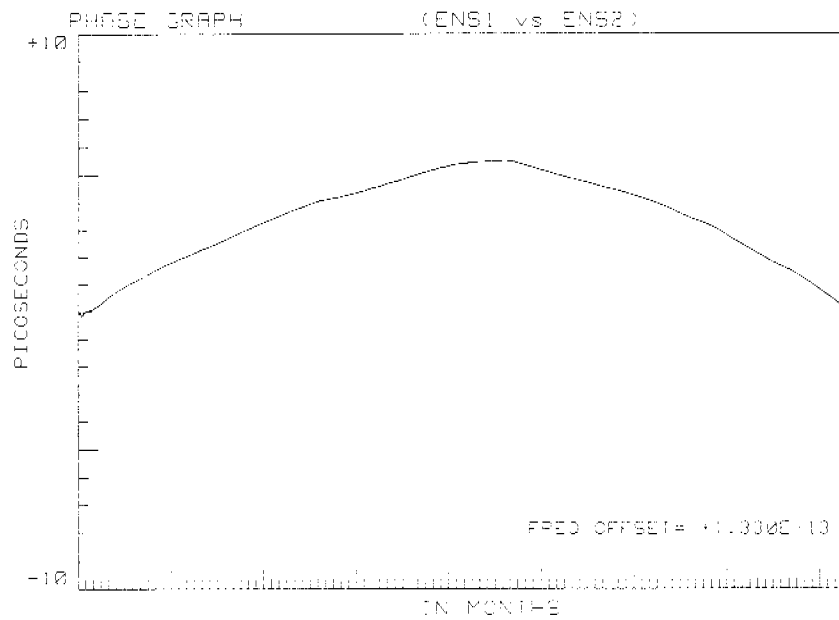
Figure 1: AT1 and KAS-1 are essentially identical in steady state.
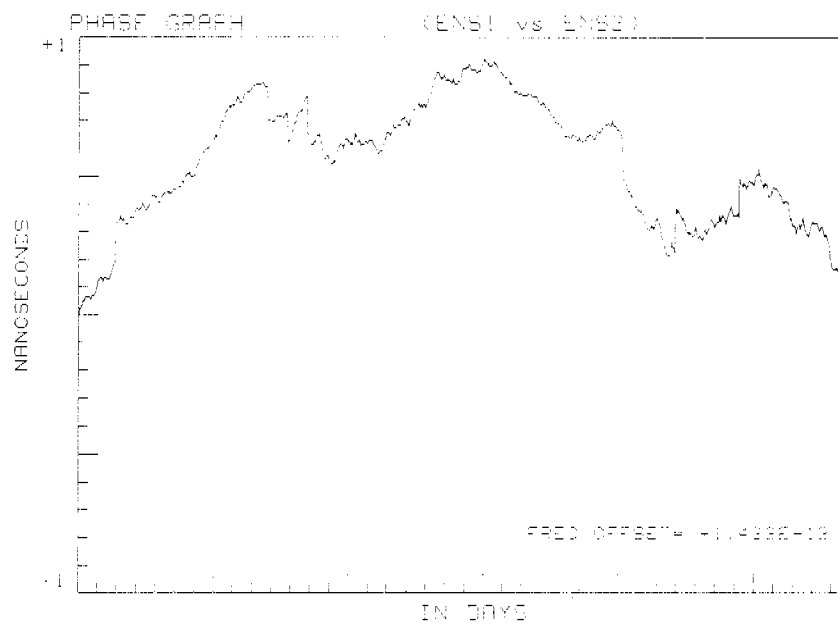
Figure 2: Small discontinuities in AT1 rejection method.
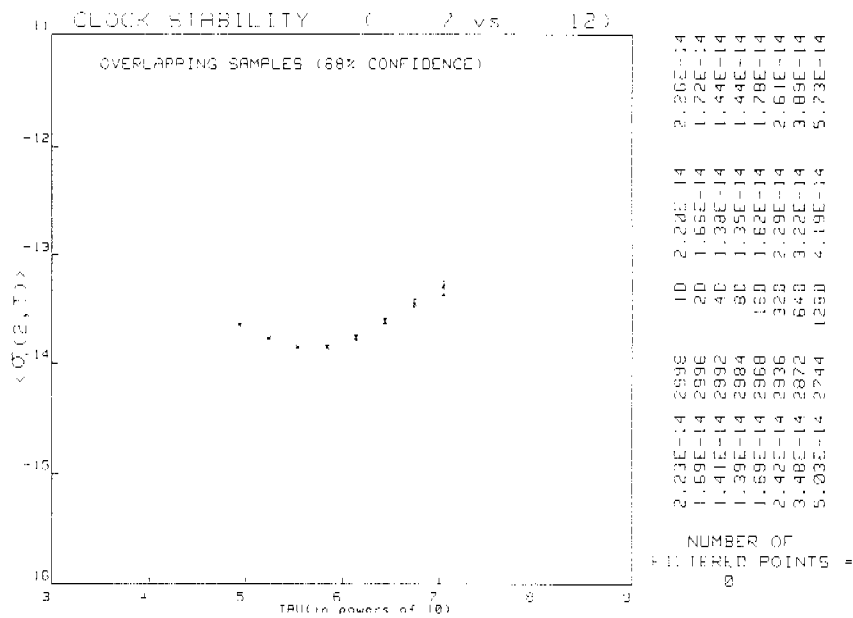
Figure 3: Simulated cesium clock data.
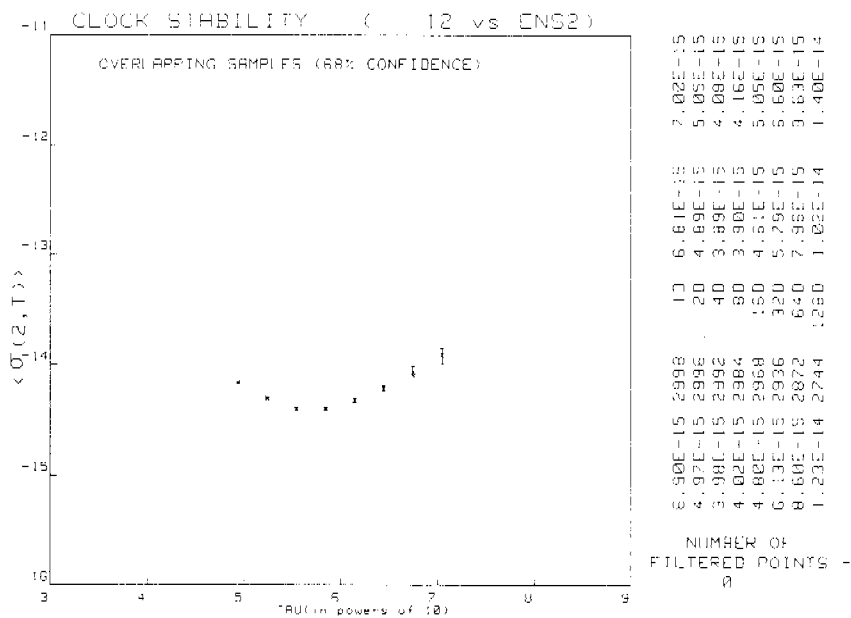
Figure 4: Square-root of N improvement with 11 equal clocks.
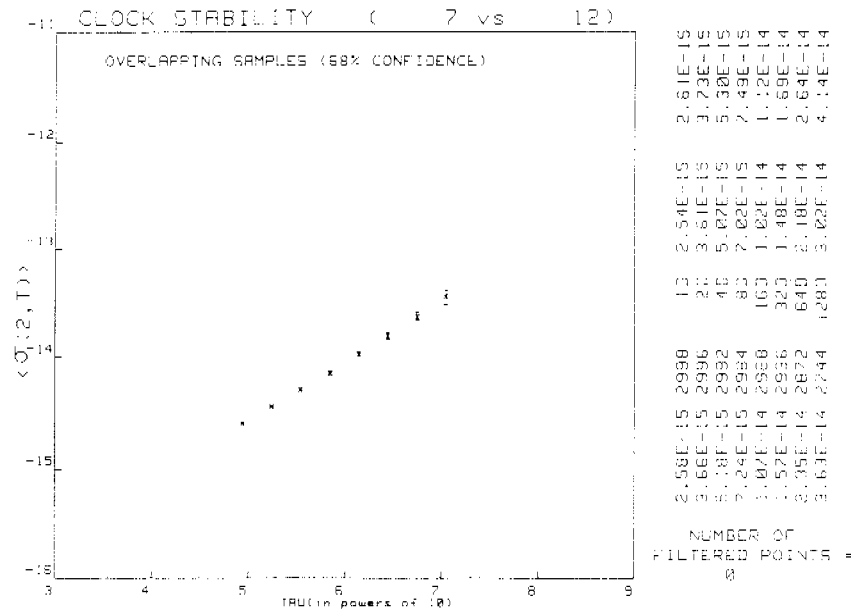
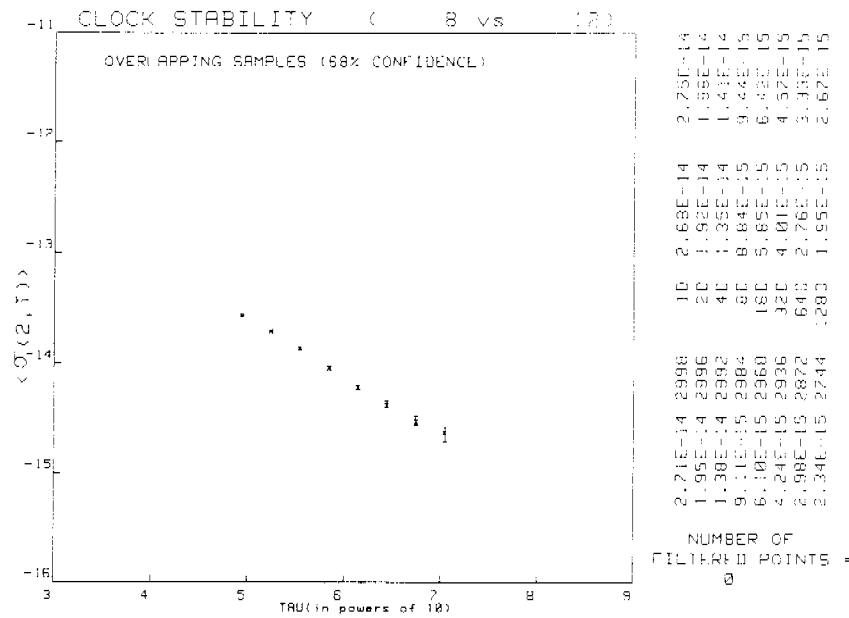Figure 5: Simulated hydrogen maser data.

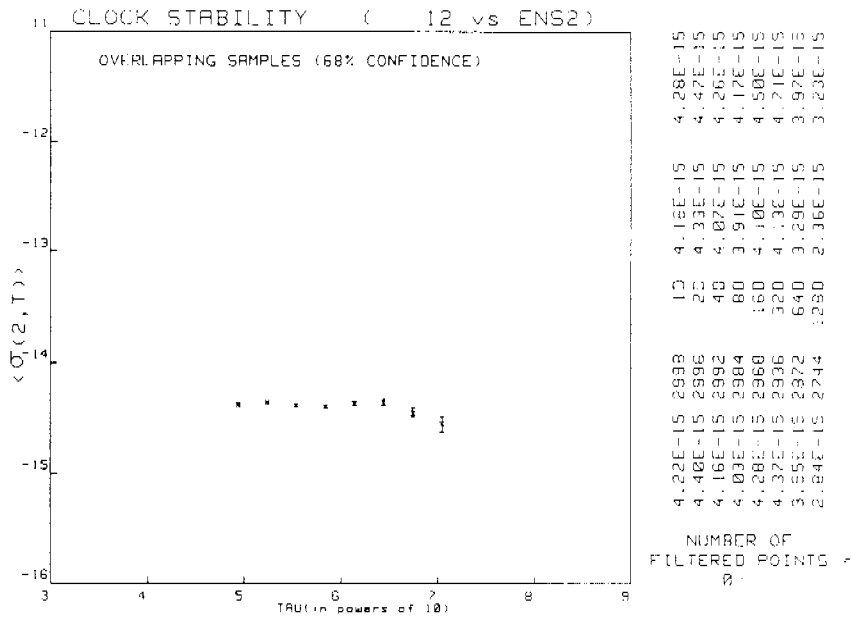Figure 6: Simulated clock with no random walk of frequency.

Figure 7: Output of algorithm using inverse variance weighting (8 years of simulated data).
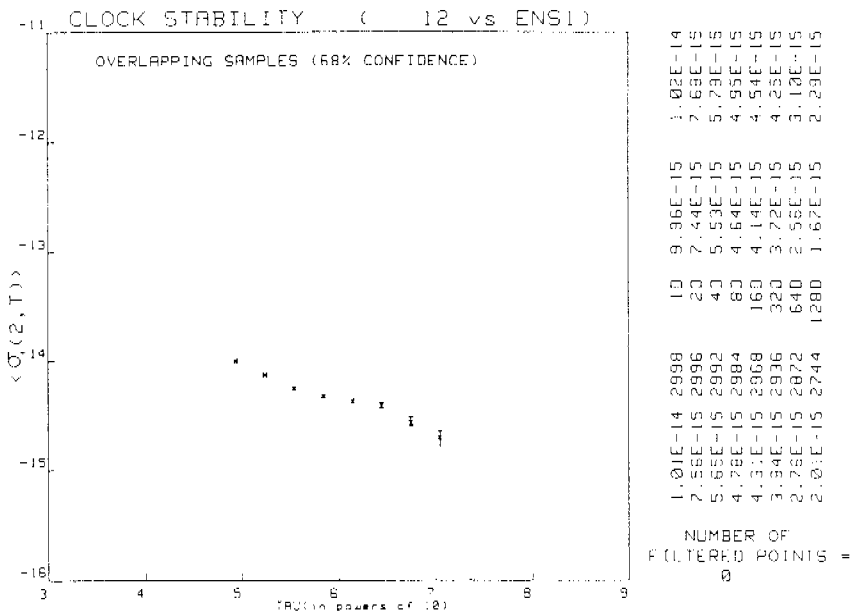

Figure 8: Output of algorithm using equal weights for all clocks (8 years of simulated data).
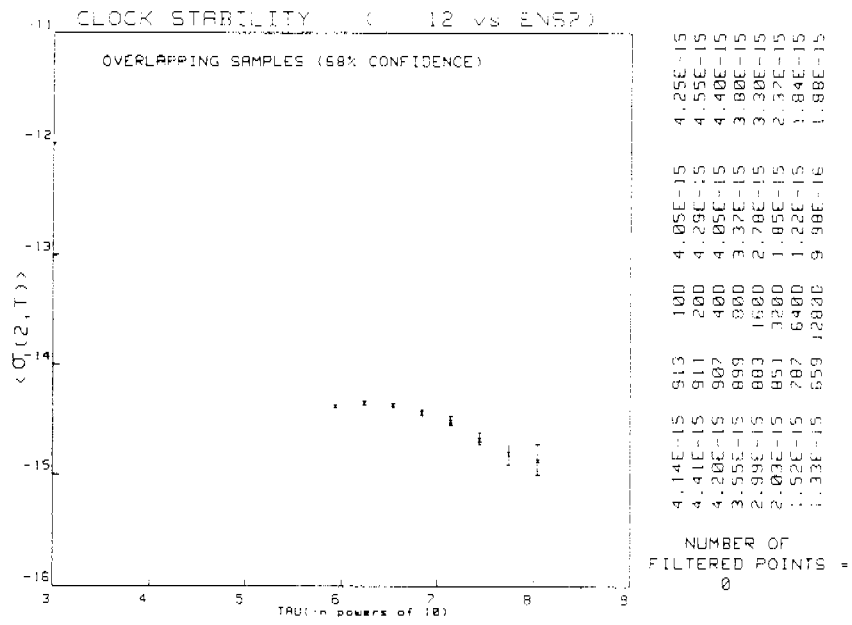
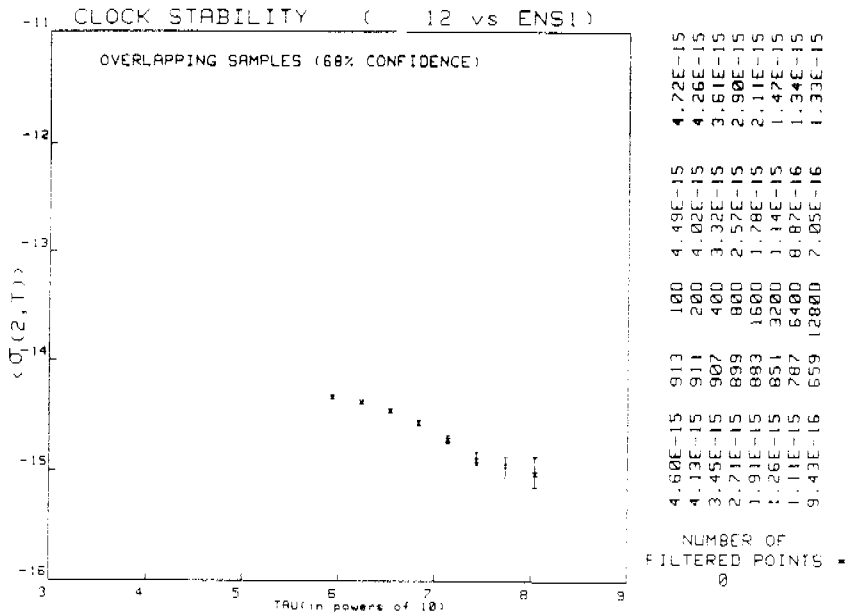Figure 9: Output of algorithm using inverse variance weighting (25 years of simulated data).

Figure 10: Output of algorithm using equal weights for all clocks (25 years of simulated data).