33rd Annual Precise Time and Time Interval (PTTI) Meeting

STUDIES OF AN OPTIMALLY UNBIASED MA FILTER INTENDED FOR GPS-BASED TIMEKEEPING

Y. Shmaliy, O. Ibarra-Manzano, R. Rojas-Laguna, and R. Vazguez-Bautista Electronics Dept., FIMEE, Guanajuato University, Salamanca 36730, Mexico E-mail: *shmaliy@salamanca.ugto.mx*

Abstract

In this paper, we examine one of the possible approaches to design an optimally unbiased moving average (MA) filter intended for the time error online estimating and synchronization in timekeeping. The approach is based on the slowly changing nature of a time error generated by the local clock and has led to a simple formula for the optimally unbiased MA filter, which does not require any a priori knowledge about the time error process. When filtering the various GPS-based time error processes, we found out that, of the filters with the same time constant, the designed filter demonstrates an error intermediate between the three-state and two-state Kalman filters. The filter has been examined for numerically simulated data and some practical situations of the GPS-based time errors generated with the crystal and rubidium oscillators.

INTRODUCTION

It is well known that timekeeping mainly solves two principal tasks, namely: estimating the time error of a local clock using the reference timing signals (of Global Positioning System (GPS) usually) and its elimination with the synchronization loop. In both cases, the "on-line" estimating and steering plays a key role, since one wants to use accurate time right now without a time delay. Thus, both the time error estimating and clock control must inevitably be provided in real time. Certainly, optimal linear stochastic estimation theory, named Kalman filtering theory [1], solves the both tasks straightforwardly. Even so, the answer is not so explicit in terms of the mean-error (bias) and variance (noise). The alternative approach, called moving averaging (MA), is still used in test and measurement [2], since of all the possible linear filters that could be used, including the Kalman, a simple MA produces the lowest noise [3], which is the input noise reduced by the square-root of the number of points N in the average. The problem here is the easily visible estimate bias caused by nonstationarity, unlike the Kalman case [4], the absence of which would make the simple MA the best.

The average bias is readily reduced in nonparametric regression estimation with the use of the weighting function, called a kernel, in statistics [5,6] and series analysis [7]. The kernel is normally a non-negative-valued function symmetric about zero, and its smooth is simply a weighted average of all data points [3]. The point is, however, the weight relates to the center of the averaging process; the result appears usually with a delay of half an averaging interval, so this is not real-time estimation. The identification theory [8] solves this problem through the proper coefficients utilized in the Auto-Regressive-Moving-Average model of the process. The coefficients may be readily identified based on the Kalman identification algorithm.

Normally, to know the tradeoff, we must examine each filter for the time error model identified by the finite polynomial [9]

$$x_n = x_0 + y_0 \Delta n + \frac{D}{2} \Delta^2 n^2 + w_{xn}, \qquad (1)$$

where $n = 0, 1, ...; \Delta = t_n - t_{n-1}$ is sample time; t_n is discrete time, x_0 is initial time error, y_0 is initial fractional frequency offset of a local clock from the reference frequency, D is linear fractional frequency drift rate, and w_{xn} is random time error deviation component. The results of such an examination for the simple MA, three-state Kalman, and second-order Wiener filters have recently been discussed in [10]. Based on those results [10], it has been deduced in [11] that the task may be solved alternatively by the bias elimination of a simple MA with the aid of stochastic approximation, taking into account that the time error (1) is ordinarily changing slowly and rather linearly during the averaging interval $\theta = \Delta(N-1)$.

In this paper we present and examine the designed filter [11] for the different time error processes. First, we examine the filter for the simulated "stationary" ($y_0 = 0$, D = 0) and "nonstationary" time error processes, then provide filtering of the real GPS-based processes generated with the rubidium and crystal clocks. To know the tradeoff, we employ four filters, namely, a simple MA, the optimally unbiased MA, and the two-state and three-state Kalman.

OPTIMALLY UNBIASED MA FILTER

Suppose we are given the discrete-time GPS-based noisy time error ξ_n (observation) on the time interval of the discrete points n - N + 1, ..., n. The observation is assumed to be an additive sum of a time error (1) and a zero-mean white Gaussian noise w_n of the GPS timing signals [12] with constant variance σ_w^2 , that is,

$$\xi_n = x_n + w_n, \tag{2}$$

where x_n is assumed change linearly on the averaging interval n - N + 1, ..., n and the noise $w_{xn}(t)$ in (1) is inherently negligible as compared to w_n in (2). Estimate the time error with a simple MA

$$\hat{x}'_{n} = \frac{1}{N} \sum_{i=0}^{N-1} \xi_{n-i}$$
(3)

Since x_n changes linearly, the filter (3) inevitably produces the bias

$$\Delta \hat{x}_n = \hat{x}'_n - x_n \cong x_{n-N+1} - \hat{x}'_n,$$
(4)

which looks negligible being related to the center of the averaging interval. To eliminate the bias (4), obtain an additional weighting function $W'_i(N)$, i = 0, 1, ..., N-1 for the simple MA (3) in the way of

$$\hat{x}_n = \frac{1}{N} \sum_{i=0}^{N-1} W'_i(N) \xi_{n-i}$$
 and provide the unbiased estimate in form of the MA model

$$\hat{x}_n = \sum_{i=0}^{N-1} W_i(N) \xi_{n-i} , \qquad (5)$$

where $W_{i}(N) = W'_{i}(N) / N$ is the required weight.

WEIGHTING FUNCTION

Since the time error function is assumed to be linear while averaging, the bias may be formally compensated for (3) in the following way:

$$\hat{x}_n = \hat{x}'_n - \Delta \hat{x}_n \cong \hat{x}'_n - (x_{n-N+1} - \hat{x}'_n) = 2\hat{x}'_n - x_{n-N+1}.$$
(6)

where the unknown variable x_{n-N+1} is time error at the start point of the averaging interval. Supposing that x_{n-N+1} is known explicitly, the estimate (6) will be unbiased with almost the same smallest noise as that produced by a simple MA. To evaluate x_{n-N+1} , use nonparametric linear regression [13], which serves in statistics as an optimal stochastic approximation in a sense of least mean squares (LMS) [14]. As a result, the continuous linear regression function appears for the interval n - N + 1, ..., n

$$\lambda(t) = a_n + b_n(t - c_n), \qquad (7)$$

where a_n , b_n , and c_n are coefficients given as

$$a_n = \frac{1}{N} \sum_{i=0}^{N-1} \xi_{n-i} = \hat{x}'_n, \qquad (8)$$

$$b_n = \frac{Cov(\xi_n, t_n)}{\sigma_m^2},$$
(9)

$$c_n = \frac{1}{N} \sum_{i=0}^{N-1} t_{n-i} , \qquad (10)$$

where $Cov(\xi_n, t_n)$ is sample covariance of ξ and t, and σ_m^2 is sample variance of time. Since (7) stochastically approximates the process in the optimal way, its value $\lambda(t_{n-N+1})$ may be treated as the most accurate evidence of x_{n-N+1} . Then substituting $\lambda(t_{n-N+1})$ for (6) should yield the optimally unbiased estimate.

OPTIMALLY UNBIASED WEIGHT

Now examine the above-mentioned opportunity, putting down

$$x_{n-N+1} \cong \lambda(t_{n-N+1}). \tag{11}$$

Substituting (8) and (11) for (6) yields the formula

$$\hat{x}_n = \hat{x}'_n - b_n (t_{n-N+1} - c_n), \qquad (12)$$

for which, first, calculate (10)

$$c_n = \frac{1}{N} \sum_{i=0}^{N-1} t_{n-i} = \frac{t_{n-N+1} + t_n}{2} = \frac{(n-N+1)\Delta + n\Delta}{2} = \Delta \left(n - \frac{N-1}{2} \right), \tag{13}$$

then a routine transform produces the variance

$$\sigma_m^2 = \frac{1}{N} \sum_{i=0}^{N-1} (t_{n-i} - c_n)^2 = \Delta^2 \frac{N^2 - k}{12}, \qquad (14)$$

where an intentionally introduced coefficient is k = 1 here, and leads to the covariance of ξ and t, that is,

$$Cov(\xi_n, t_n) = \frac{1}{N} \sum_{i=0}^{N-1} (\xi_{n-i} - a_n)(t_{n-i} - c_n) = \frac{\Delta}{N} \sum_{i=0}^{N-1} \left(\frac{N-1}{2} - i\right) \xi_{n-i}$$
(15)

Substitute (9), (13)—(15) for (12), make the transformation, and get the desired unbiased estimate

$$\hat{x}_n = \sum_{i=0}^{N-1} \frac{2(2N-1) - 6i}{N(N+1)} \xi_{n-i}$$
(16)

Then write the required weighting function for (5) as

$$W_{i}(N) = \begin{cases} \frac{2(2N-1)-6i}{N(N+1)}, & 0 \le i \le N-1, \\ 0, & otherwise \end{cases}$$
(17)

which evidently is nonzero on the averaging interval only. It seems important to note that the same formula (16) appears if one just converts the linear regression value $\lambda(t_n)$ into the MA model (5).

Let us now note that (17) is linear (Fig. 1) with its maximum $W_0(N) = \frac{2(2N-1)}{N(N+1)}$ and its minimum

 $W_{N-1}(N) = -\frac{2(N-2)}{N(N+1)}$, and with zero-point $n_0 = (2N-1)/3$. Two special features of the weight (17) may

be observed, namely: 1) its square is unity, thus, this is nothing more than impulse response of the Finite Impulse Response (FIR) filter [14]; 2) the maximum-to-minimum ratio of the weight tends to

$$\lim_{N \to \infty} \frac{W_0(N)}{W_{N-1}(N)} = -2.$$
(18)

It was shown in [11] that the filter (16) is *optimally unbiased*, since the bias is totally compensated for $2 \le N$, this is $\Delta \hat{x} = 0$. Its estimate minimal RMSD $\sigma_{OMA}|_{N=2} = \sqrt{2}\sigma_{wN}$ (Appendix) corresponds to the minimal number N = 2 and, since N increases, the noise asymptotically tends to $\sigma_{OMA}|_{1 \le N} = 2\sigma_{wN}$.

Now note that even though the bias is totally compensated, the filter (16) produces noise bigger than a simple MA; this is $\sqrt{2}\sigma_{wN} \le \sigma_{OMA} < 2\sigma_{wN}$. Moreover, because bias is negligible, both the RMSE and the maximal error (A1) have for N = 2 almost the same minimal magnitude; this is $\approx \sqrt{2}\sigma_{wN}$.

IMPROVED WEIGHT

Minimization of the global filtering error (A1) leads to the improved weight provided in [11] rather in a heuristic way based on simulation. Here, since the function slope (9) strongly influences the compensation efficiency, one can just change the variance in the denominator of (9) by the integer k in (14) and minimize the global error for the smallest number N = 2. Then k = -6 provides appreciable diminishing of RMSD, so that the filter exhibits $\sigma_{WN} < \sigma_{IMA} < 2\sigma_{WN}$ for $2 \le N$ and both the RMSE and the maximal error start approximately with the same magnitude $\approx \sigma_{WN}$ as well. The *improved optimally unbiased filter* for k = -6 then becomes

$$\hat{x}_{n} = \sum_{i=0}^{N-1} \frac{2N(2N-3) + 9 - 6i(N-1)}{N(N^{2}+6)} \xi_{n-i},$$
(19)

with its weighting function

$$W_{i}(N) = \begin{cases} \frac{2N(2N-3)+9-6i(N-1)}{N(N^{2}+6)}, & 0 \le i \le N-1, \\ 0, & otherwise \end{cases}$$
(20)

The special feature of the filter (19) is for N = 2 all its particular errors (A1) are almost equal to those of the simple MA and then, since N increases, tend to those of the optimal filter (16). However, the improvement is reached at expense of bias, which in a case of (16) is totally compensated. This is the reason why the filter (19) is in fact asymptotically optimally unbiased, since its bias tends toward zero as $N \rightarrow \infty$. With $20 \le N$ both filters, (16) and (19), demonstrate almost the same performance and the improved filter looks better only for N < 20.

FILTER ANALYSIS

Let us now observe major properties of the filters (16) and (19) aiming to measure their performance by the factors such as the accuracy of the obtained solutions, their convergence speed, their tracking ability for the time error process, and their computational complexity.

The accuracy strongly depends on the produced bias and noise. From this point of view, the filter (16) is optimal in terms of minimal bias, and the filter (19) is optimal in terms of the minimal global error.

All three filters, namely, a simple MA, (3), the optimally unbiased, (16), and the improved, (19), exhibit the same convergence of the algorithms, since all the data should be processed for the same time, and both weights, the rectangular for (3) and the triangle for (5) are restricted by a unit square. In contrast to a simple MA (3), the filter (16) exhibits much better tracking because the cause, the bias, is removed. The filter (19) also looks much better than (3), but slightly worse than (16).

Certainly, each algorithm, (16) and (19), takes more time for computation, because of an additional weight introduced for a simple MA. However, this argument cannot be seriously taken into account as a visible disadvantage if we recall that the sample time in the GPS-based timekeeping is usually $\Delta = 10...1000$ sec, so there is enough time for computation.

SIMULATION AND FILTERING

We now examine both filters for the different practical situations. First, we simulate and study the noisy processes with known time errors close to those obtained with use of the GPS timing signals. Then we investigate the filtering errors respecting the rubidium and crystal clocks. In this last case, the time error had been measured based on the GPS Timing Receiver of the Motorola Oncore UT+ type with the sample time $\Delta = 100$ sec. To obtain small filtering noise, we put down $80 \le N$, in which case both filters, (16) and (19), exhibit rather the same error, and, for this reason, we use only the improved formula (19). While studying it, we compare the results to those inherently provided by the two-state and three-state Kalman filters [15] and evaluate the filtering error by (A1). The results are obtained based on MATHCAD software.

SIMULATION

To test the filter for all the errors (A1), we simulate two time error processes named the "stationary" and the "nonstationary" cases. The process is assumed to be stationary if $x_0 = y_0 = D = 0$ and the nonstationary case is obtained with $x_0 = 0$, $y_0 = -5 \cdot 10^{-12}$ and D = 0. For both the cases we set N = 100 and simulate the white Gauss noise with an rms deviation of $\sigma_w = 25$ ns. Both a simple MA (3) and the unbiased filter (19) have been examined, and the filtering errors were calculated for the simulated process x_n taken as an accurate function of the time error; this is $x_n^o = x_n$ in (A1). As had been expected, noise of the filter (19) (and, strictly, of all other filters [2], if one were to extend) exceeds that of a simple MA in the stationary case (Fig. 2). In contrast, a simple MA demonstrates big bias in the nonstationary case (Fig. 3).

Based on Figs. 2 and 3, one may conclude that both filters, a simple MA and the unbiased, produce almost the same bias in the stationary case. The other conclusion is the noise produced by the unbiased filter in this particular case is about three times bigger than that of simple MA, though theoretically the limit equals 2 [11]. On the whole, we see that simple MA is best for the stationary case (Fig. 2) and the unbiased filter is best for the nonstationary case (Fig. 3) with the ratio of the RMSE of 0.43 and 4.93 for each case, respectively.

GPS-BASED TIME ERROR GENERATED WITH THE RUBIDIUM CLOCK (THE FIRST CASE)

Here we deal with the real data of the GPS-based time error generated by the rubidium clock with the offset of $y_0 \approx 2.3 \cdot 10^{-12}$ and $\Delta = 100$ sec. Following the tradition, we examine both a simple MA and the unbiased filter, and do the same for the two-state and three-state Kalman algorithms [15]; those are the most matched with the time error model (1). Since the transient is the important performance of the filter in timekeeping, it delays receiving of the first estimate, then, to obtain the same inertia for each filter, we put N = 80 and set the proper states noises in the signal noise matrices of the Kalman filters. In this way, we obtain the transient time at the level of 0.9 to be equal to the averaging interval $\theta = \Delta(N-1) \cong 2.22$ hours of a MA. Figure 4a shows the observation and all the four filtered curves. Because in this case we do not know the origin of the time error x_n^o in (A1), we assume the three-state Kalman estimate \hat{x}_n^{3Kal} to be the most accurate. Hereupon, substituting $x_n^o = \hat{x}_n^{3Kal}$ for (A1) produces Fig. 4b. The visible finding here is the unbiased filter is best as compared to the simple MA and even to the twostate Kalman. Just as in the simulated case (Fig. 3), we watch here for the good compensation of the bias, which is almost the same as that of the three-state Kalman filter. Though the noise of the simple MA is inherently small, the RMSD of its estimate is bigger by about 2.5 and 1.6 times that of the unbiased and two-state Kalman filters, respectively. With this, the unbiased filter exhibits about 1.5 times smaller error than that of the two-state Kalman filter. Yet, both the two-state Kalman and the unbiased filters yield rather the same maximal error.

GPS-BASED TIME ERROR GENERATED WITH THE RUBIDIUM CLOCK (THE SECOND CASE)

Consider the other case of the time error of the rubidium clock with $y_0 \approx -1.44 \cdot 10^{-12}$ and the same conditions as those provided for the first case. Again, we see (Fig. 5) the visible smallest error of the unbiased filter. Calculation shows as well the best compensation of the bias, which is smallest among those of a simple MA and the two-state Kalman filter. The unbiased filter demonstrates the RMSD, which is about 3.0 and 1.8 times lower than that of a simple MA and the two-state Kalman filter, respectively. Both the RMSE and the maximal errors are also smaller by more than two times in this case.

GPS-BASED TIME ERROR GENERATED WITH THE CRYSTAL CLOCK

Let us now investigate the final example of the time error filtering. This case corresponds to the ovenized crystal oscillator (OCXO), in which case the time error is inherently bigger than that of the rubidium clock. The time error function and the filtered results are sketched in the Fig. 6. On the whole, we come here to the same conclusions, namely: the optimally unbiased filter demonstrates negligible bias and exhibits the RMSD, the RMS, and the maximal errors by 4.1 to 4.4 times smaller as compared to a simple averaging, and by 2.7 to 3.2 times smaller than those provided by the two-state Kalman filter. Recall we use the estimate of the three-state Kalman filter as a reference.

CONCLUSIONS

We have examined in this paper the optimally unbiased MA filter [11] designed especially for the tasks of "on-line" estimation and synchronization in timekeeping. The special features of the filters, (16) and (19), are

- The filter (16) is optimally unbiased and the filter (19) is asymptotically optimally unbiased with the maximum produced noise by 2 times bigger with $1 \ll N$ than that of a simple MA.
- Both filters yield the same result, since $20 \le N$ and the filter (19) produces noise smaller, since N < 20.
- The filters do not require any *a priori* knowledge about the GPS-based time error process.

Based on results of the numerical simulation and the filtering of the real GPS-based time error processes generated by the rubidium and crystal clocks, we come to the following conclusions:

- In practice, the noise of the optimally unbiased filters, (16) and (19), may be more than by 2 times bigger that that of a simple MA. This holds true for a small number of samples *n*. Since *n* increases, the ratio of RMSD tends to the theoretical limit of 2.
- The filters exhibited the intermediate error between those provided by the three-state and twostate Kalman filters, since all the filters were tuned for the same time constant.
- In fact, both filters may be used in timekeeping to provide an unbiased estimate of time error with small noise.

APPENDIX : FILTERING ERRORS

Find out the estimate error as

$$\varepsilon_n = x_n^o - \hat{x}_n, \qquad (A1)$$

where \hat{x}_n is estimate of a time error, x_n^o is assumed to be accurate value of a time error. Evaluate (A1) by the particular sample errors for an arbitrary number *M* of estimates; those are, bias $\Delta \hat{x} = E[\varepsilon_n]$, variance $\sigma_{\varepsilon}^2 = E[(\varepsilon_n - \Delta \hat{x})^2]$, root-mean-square deviation (RMSD) $\sigma_{\varepsilon} = \sqrt{\sigma_{\varepsilon}^2}$, rms error (RMSE) $\varepsilon_{\text{RMS}} = \sqrt{E[\varepsilon^2]} = \sqrt{\Delta \hat{x}^2 + \sigma_{\varepsilon}^2}$, maximal error $\varepsilon_{\text{max}} = \max |\varepsilon_n|$, and global error, which we would like to evaluate as $\varepsilon_{\text{global}} = 0.5(\varepsilon_{\text{RMS}} + \varepsilon_{\text{max}})$.

REFERENCES

- [1] B. Anderson and J. Moore, 1979, Optimal Filtering (Prentice-Hall, Englewood Cliffs, New Jersey).
- [2] S. Mitra and J. Kaiser (eds.), 1993, Handbook for Digital Signal Processing (John Wiley & Sons, New York).
- [3] S. Smith, 1997, **The Scientist and Engineer's Guide to Digital Signal Processing** (California Technical Publishing, San Diego).
- [4] R. Kalman, 1960, "A New Approach to Linear Filtering and Prediction Problems," Transactions of the ASME, Ser. D., Journal of Basic Engineering, 82, 35-45.
- [5] L. Devroye, 1981, "On the almost everywhere convergence of nonparametric regression function estimates," Annals of Statistics, 9, 1310-1319.
- [6] W. Greblinski, A. Krzyzak, and M. Pawlak, 1984, "Distribution-free pointwise consistency of kernel regression estimate," Annals of Statistics, 12, 1570-1575.
- [7] P. Robinson, 1983, "Nonparametric estimators for time series," Journal of Time Series Analysis, 4, 185-207.
- [8] T. Söderström and P. Stoica, 1989, System Identification (Prentice-Hall, Englewood Cliffs, New Jersey).
- [9] G.810. "Definitions and Terminology for Synchronization Networks" (ITU-T, Geneva), August 1996.
- [10] Y. Shmaliy, A. Marienko, M. Torres-Cisneros, and O. Ibarra-Manzano, 2001, "GPS-Based Time Error Estimates Provided by Smoothing, Wiener, and Kalman Filters: A Comparative Study," in Proceedings of the 32nd Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 28-30 November 2000, Reston, Virginia, USA (U.S. Naval Observatory, Washington, D.C.), 157-169.
- [11] Y. Shmaliy, "A New Simple Optimally Unbiased MA Filter For Timekeeping," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control (under revision).

- [12] D. Allan and J. Barnes, 1982, "Optimal Time and Frequency Transfer Using GPS Signals," in Proceedings of the 36th Annual Frequency Control Symposium, 2-4 June 1982, Philadelphia, Pennsylvania, USA (IEEE Publication AD-A130811), pp.378-387.
- [13] L. Ljung, G. Pflug, and H. Walk, 1992, Stochastic Approximation and Optimization of Random Systems (Birkhauser Verlag, Basel).
- [14] G. Glentis, K. Berberidis, and S. Theodoridis, 1999, "Efficient Least Squire Adaptive Algorithms for FIR Transversal Filtering," IEEE Signal Processing Magazine, July 1999, 13-41.
- [15] Y. Shmaliy, A. Marienko, et al., 2000, "GPS-Based Optimal Kalman Estimation of Time Error, Frequency Offset, and Aging," in Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, 7-9 December 1999, Dana Point, California, USA (U.S. Naval Observatory, Washington, D.C.), pp. 431-440.



Figure 1. The optimal weighting function.



Figure 2. Filtering of the simulated stationary time error process for $x_0 = 0$, $y_0 = 0$, D = 0, $\sigma_w = 25ns$, $\Delta = 100s$, and N = 100.



Figure 3. Filtering of the simulated nonstationary time error process for $x_0 = 0$, $y_0 = 5 \cdot 10^{-12}$, D = 0, $\sigma_w = 25ns$, $\Delta = 100s$, and N = 100.



Figure 4. Filtering of the GPS-based time error process generated by the rubidium standard for $\Delta = 100$ s, and N = 80: a) Estimates; b) Errors with respect to the three-state Kalman estimate.



Figure 6. Filtering of the GPS-based time error process generated by the rubidium standard for $\Delta = 100$ s, and N = 80: a) Estimates; b) Errors with respect to the three-state Kalman estimate.



Figure 7. Filtering of the GPS-based time error process generated by the OCXO for $\Delta = 100$ s, and N = 80: *a*) Estimates; *b*) Errors with respect to the three-state Kalman estimate.

QUESTIONS AND ANSWERS

JIM CAMPARO (The Aerospace Corporation): You had random walk frequency in the rubidium clock data. I am very sure time scales that random walk would look like just a linear frequency offset. And so if I understand this correctly, your averaging time for your moving average has to be on that time scale where random walk looks like some frequency offset. Is that correct?

YURIY SHMALIY: That's right. Yes.