Metrologia

Towards an optimal weighting scheme for TAI computation

To cite this article: Gérard Petit 2003 *Metrologia* **40** S252

Metrologia 40 (2003) S252–S256 PII: S0026-1394(03)61403-9

Towards an optimal weighting scheme for TAI computation

Gérard Petit

BIPM, Pavillon de Breteuil, F-92312 Sèvres Cedex, France

E-mail: gpetit@bipm.org

Received 29 April 2002
Published 5 June 2003
Online at stacks.iop.org/Met/40/S252

Abstract
International atomic time (TAI) is an ensemble timescale based on a weighted average of clocks. The basic properties of the algorithm have been fixed since its inception: that is, the weight of a clock is inversely proportional to a variance measuring the instability of the clock, and it cannot exceed a maximum value. On the other hand, the procedure used to set the maximum weight has been subject to several changes over the years. While the most stable timescale would ideally be computed without an upper limit of weights, such an upper limit is introduced to bring reliability. Up to now, however, there has been no adopted measure of the reliability. In this paper, a quantitative estimator of the reliability is proposed, which therefore helps in choosing a weighting scheme. Different procedures for setting the maximum weight are examined in light of the application of the reliability criterion defined here. Tests using simulated and real data are presented. A weighting scheme to be used for TAI computation is proposed in which the reliability estimator is optimized.

1. Introduction

The weighting scheme for the algorithm used in the computation of international atomic time (TAI) [1] has been the subject of several studies over the past years. In 1996, a first study [2] resulted in the adoption of a scheme in which the maximum relative weight was fixed to a given value (0.7%) with a reminder that this value should be reconsidered in view of the evolution of the clock ensemble. A recent study for the working group on TAI of the Consultative Committee for Time and Frequency (CCTF) [3] outlined the need to increase the maximum relative weight attributed to a clock in the TAI computation because it is observed that a very large fraction of the clocks reach the maximum weight so that there is no discrimination between the very stable clocks and those that are only fairly stable. It was proposed to express the maximum relative weight as a fraction A/N where N is the total number of clocks. Such a procedure was effectively recommended by the working group on TAI and it was implemented, with $A = 2$, from January 2001.

While the conclusions of these studies seem to be quite reasonable and have been adopted by the members of the working group on TAI, it appears that they are based on intuition rather than on quantitative estimates. Indeed, while the most stable timescale would ideally be obtained without using an upper limit of weights, such an upper limit is introduced to bring reliability, but no quantitative measurement of the reliability has been proposed so far. One of the goals of this paper is to introduce such a measurement. First, in section 2, we recall the basics of the algorithm used for the computation of TAI. We then introduce in section 3 estimators to quantify the stability and the reliability of a timescale produced by such an algorithm. In section 4, after recalling different procedures that may be used to set a maximum weight for a clock, we present the results of tests carried out with simulated data or with the data-set of the TAI clock ensemble. We conclude in section 5 by suggesting a weighting scheme that aims to optimize the proposed estimators of reliability and stability.

2. The stability algorithm of TAI

In the computation of TAI, an ensemble timescale EAL is first computed with the algorithm ALGOS [4]. The next step, in which TAI is computed by steering the EAL frequency to primary frequency standards, is not considered here (see e.g. [5]). In the following we often explicitly consider the case of ALGOS and EAL, but more general statements could be made for similar algorithms, with obvious changes. The basic principles of such an algorithm are the following:

(a) The computation is carried out for a given computation interval (one month in ALGOS), considering all clocks that have data for the whole computation interval. The effect of entering and exiting clocks on the ensemble scale is handled by a time continuity condition and a frequency prediction [1].
(b) The weight of a clock is inversely proportional to a variance computed with the rates of the clock with respect to the ensemble scale over a number of computation intervals. The relative frequency instability of the clock is measured by the square root of the variance (standard deviation). In ALGOS the classical variance is used and it is computed over 11 past intervals plus the current interval, that is, a total of 12 months.
(c) A given clock cannot obtain a weight larger than a maximum value $w_{\max}$. Different ways to specify a maximum weight, including those used in ALGOS, are described in section 4.
(d) An abnormal behaviour is said to occur for a clock if the difference between its rate over the current interval and its past rate is larger than a given threshold. When this happens, the clock weight is set to zero. In ALGOS, the abnormal behaviour condition is that the difference between the current rate and its mean over the last 11 intervals is larger (in absolute value) than either three times its computed standard deviation over the past 11 intervals, or three times the instability of the least stable clock at maximum weight, whichever is larger.

One can see that, among these basic principles, item (b) is intended to provide stability to the ensemble timescale while items (c) and (d) are intended to provide reliability to it. In the following section, we describe estimators to measure these quantities.

3. Estimators for the stability and the reliability of a TAI-type timescale

Given an ensemble timescale that has been computed with an algorithm as described in section 2 over a continuous period of time, we define two estimators to measure its stability and its reliability. These estimators will then be used to compare different weighting schemes in order to determine 'optimal' ones.

3.1. The stability estimator $\mathcal{S}$

From the results of past computations (performed over successive one-month intervals), one may estimate the relative frequency instability of each clock with respect to EAL, for an averaging duration of one month. Because EAL is more stable than each individual clock, this provides a good estimate of the intrinsic relative frequency instability of each clock σ_i, with i taking values from 1 to N. If another timescale j is computed with the same clock data and a different weighting scheme, in which clock i obtains the normalized weight w_{ij}, assuming no correlation between the clocks, we may estimate the stability $\mathcal{S}_j$ of the resulting timescale by $\mathcal{S}_j^2 = \sum_{i=1}^{N}(w_{ij}^2\sigma_i^2)$. We choose $\mathcal{S}_j$ as a quantitative criterion to discuss the estimated stability of the timescale, with the smaller $\mathcal{S}_j$ corresponding to the most stable timescale. Note that this procedure assumes no correlation between each clock and the ensemble timescale. Because we are interested in ensembles of a large number of clocks (typically $N = 200$ for TAI) and because all weighting schemes studied here imply a maximum relative weight of the order of a few per cent at most, this approach has been considered to be satisfactory [1]. If this was not the case, an unbiased clock variance could be estimated [6].

3.2. The reliability estimator $\mathcal{R}$

Because the reliability is realized by imposing a maximum value on the weight of a given clock and by implementing a detection of 'abnormal behaviour' (see section 2), it is possible to quantify the reliability by estimating the maximum frequency change $\mathcal{R}_j$ that may be caused to the timescale j by one clock suffering a frequency step. This corresponds to the effect of one clock at maximum weight suffering a frequency step whose absolute value is just below the limit that would correspond to abnormal behaviour, that is, in our case three times the instability of the least stable clock at maximum weight, denoted here by σ_j. Therefore, $\mathcal{R}_j = 3\times\sigma_j$. We choose this estimator to quantify the reliability of the timescale, with the smaller $\mathcal{R}_j$ corresponding to the most reliable timescale. It may be proved (see appendix) that $\mathcal{R}_j$ is minimal when a simple condition on the maximum weight is verified: this is when the combined weight of the clocks at maximum weight is the half of the total weight, that is (using normalized weights), if N_m clocks have the maximum relative weight $w_{\max}$, the condition is $N_m \times w_{\max} = 0.5$. In the following section we verify this property for all test cases.

4. Tests of different weighting schemes and different clock ensembles

4.1. Some possible weighting schemes

As outlined in section 2, the general weighting scheme is that the weight w of a clock is inversely proportional to a variance V. Here, we consider normalized weights (the sum of the weights is 1). Several approaches have been (or may be) proposed to set the maximum weight attributable to a clock. We first mention a method that uses a threshold for the clock variance. If V for a clock is smaller than a given threshold $V_{\min}$, just replace V by $V_{\min}$ for the computation of the weight. This procedure was used in TAI computation until 1997 (with $V_{\min} = 5.4 \times 10^{-28}$ from May 1995 to 1997). Though simple, this procedure is only applicable when one knows the overall instability distribution of the clocks and the threshold should be updated whenever necessary—it is presented here only for completeness. All other procedures described below (P1 to P5) aim at defining a maximum relative weight, $w_{\max}$, more or less independently of the typical value of the instability of the clocks. They require iterations to obtain the final distribution of weights: at each iteration step, all weights w larger than $w_{\max}$ are replaced by $w_{\max}$ and the set of values is re-normalized.

The five procedures differ in the way $w_{\max}$ is chosen:

(P1) Set $w_{\max}$ as a given numerical value. This procedure was used in TAI computation between 1998 and 2000 (with $w_{\max} = 0.7\%$).

(P2) Set $w_{\max} = A/N$ where A is a constant and N is the number of clocks considered for weighting. This procedure has been used in TAI computation since January 2001 (with $A = 2$).

(P3) Set $w_{\max}$ as the smallest value for which the number of clocks reaching maximum weight, N_m, is larger than a certain threshold.

(P4) Set $w_{\max}$ as the smallest value for which the percentage of clocks reaching maximum weight, P_m, is larger than a certain threshold.

(P5) Set $w_{\max}$ as the smallest value for which the combined weight of clocks reaching maximum weight, W_m, is larger than a certain threshold.

We can make some general remarks on the procedures above, which help us in interpreting the results of the tests reported in the following sub-sections.

1. Procedures (P1) and (P3) obviously require updates to reflect changes in the number of clocks, while the others do not suffer this drawback. However, all procedures may require an update if the stability distribution of the ensemble of clocks changes. Note that, for a given number of clocks, procedures (P1) and (P2) on the one hand, and (P3) and (P4) on the other hand, can be considered equivalent; that is, there is a trivial relation between them.

2. Note also that, for a given type of stability distribution, the procedures (P2) and (P4) can also be considered to be equivalent, that is, there exists some monotonic (decreasing) function of $w_{\max}$ representing P_m.

3. Procedure (P5) is obviously particularly adapted to the use of the reliability estimator $\mathcal{R}$. However, it does not seem advisable to use (P5) alone because it does not provide a constraint on the number of clocks that actually receive the maximum weight, so that this number may, in principle, become very small, which is not satisfactory for reliability. This appears to be a second criterion for reliability, linked to the effect on the ensemble scale of the loss of one clock at maximum weight. This criterion deserves more study but this is beyond the scope of this paper. Instead, we shall empirically introduce a minimum percentage of clocks at maximum weight (see section 5). Because this cannot be directly introduced for (P5), it is not considered in the tests below.

4.2. Tests on simulated data

Several hundreds of random normal distributions were generated to simulate the logarithms of the instabilities of ensembles of clocks. As mentioned above, because we generate homogeneous series of distributions of a given number of clocks, there exists a direct relationship between the results of all the procedures above, so it is not necessary to test each of them. It is sufficient to test only a number of values of the maximum relative weight, that is, in effect to test procedure (P1). The estimators for stability and reliability are computed and examples are presented in figures 1 and 2 for two series of distributions of instability whose logarithm has the same average (-14.0). For one series (called case S) the standard deviation of the logarithm is small (0.2), that is, all clocks have a similar instability, and for the other one (called case L) the standard deviation is larger (0.33), that is, some clocks are markedly better or worse than average. Comparing cases S and L, the main observations are the following:

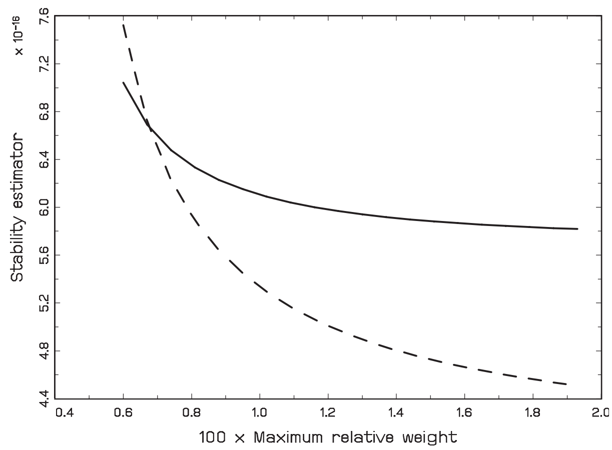

Figure 1. Stability estimator $\mathcal{S}$ as a function of the maximum weight for two normal distributions of clock instabilities (case S: ——, case L: - - - -, see text).

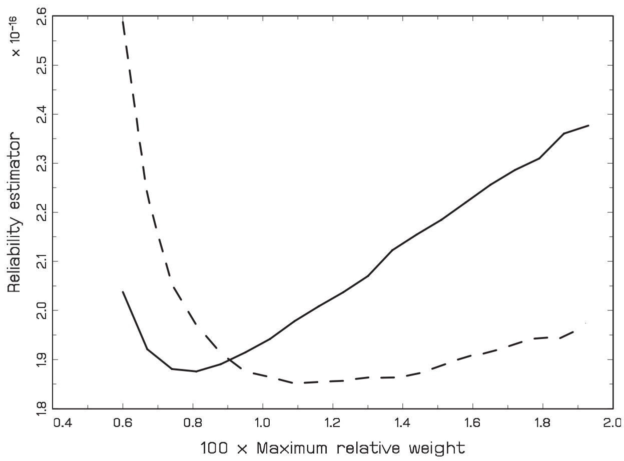

Figure 2. Reliability estimator $\mathcal{R}$ as a function of the maximum weight for two normal distributions of clock instabilities (case S: ——, case L: - - - -, see text).

(a) the value $w_{\max}$ corresponding to optimal reliability is smaller in case S than in case L;

(b) for this particular choice of $w_{\max}$, the relative loss of stability (as measured by $\mathcal{S}$) with respect to the case of free weights (i.e. the limit for large values of $w_{\max}$) is smaller in case S than in case L;

(c) on the other hand, for this particular choice of $w_{\max}$, the relative gain in reliability (as measured by $\mathcal{R}$) with respect to the case of free weights is larger in case S than in case L.

We conclude that, for ensembles of ideal clocks, using the reliability criterion is specially important when the clocks have a similar instability (case S), because unduly increasing the maximum weight then yields a significantly larger sensitivity to undetected frequency steps but a rather small gain in theoretical stability. This may be important for

TAI, because it is presently mostly based on an ensemble of clocks of the same model (see next section).

In addition, the reliability estimator is indeed minimal when the combined weight of clocks at maximum weight is half of the total weight (figure 3).

4.3. *Tests on the TAI clock ensemble*

We have at our disposal the set of clock rates with respect to EAL resulting from the regular TAI computation (available at ftp://62.161.69.5/pub/tai/publication). First, we note (see figure 4) that the observed distribution function of instability of the TAI clocks (average over 1999–2001) is close to a normal distribution with some significant tail in the large instability end (this is not very disturbing because these clocks will receive infinitesimal weight). We use this set of clock instabilities to compute the weight distribution that would be obtained for each interval of one month using the ALGOS algorithm with different procedures for setting the maximum relative weight. For each computation, we obtain a set of weights w_{ij} from which we estimate $\mathcal{S}_j$ and $\mathcal{R}_j$. At the same time we obtain the subset of clocks that reach the maximum weight, characterized by its number N_m or by the fraction P_m of the total number of clocks. Such a computation has been performed for 36 months (January 1999 to December 2001), for a number of values of the maximum relative weight. Results are presented for the 36 months for three values of the maximum weight ($w_{\max} = 1.5/N$, $2.5/N$ or $4/N$) following procedure (P2), along with the results corresponding to the actual computation of TAI.

Figure 5 shows the stability estimator $\mathcal{S}_j$ for the resulting timescale and figure 6 shows the reliability estimator $\mathcal{R}_j$. In addition the combined weight of the clocks at maximum weight is shown in figure 7. The main conclusions are the following:

(a) The stability of the timescale continuously improves (figure 5) as the maximum relative weight increases, with fewer clocks reaching the maximum weight. This is, of course, expected. We note that the gain from $w_{\max} = 2.5/N$ to $w_{\max} = 4/N$ is marginal (<10%).

(b) Comparing figures 6 and 7, we see that the reliability estimator is indeed minimal when the combined weight of clocks at maximum weight is half of the total weight (which is about the situation for $w_{\max} = 2.5/N$).

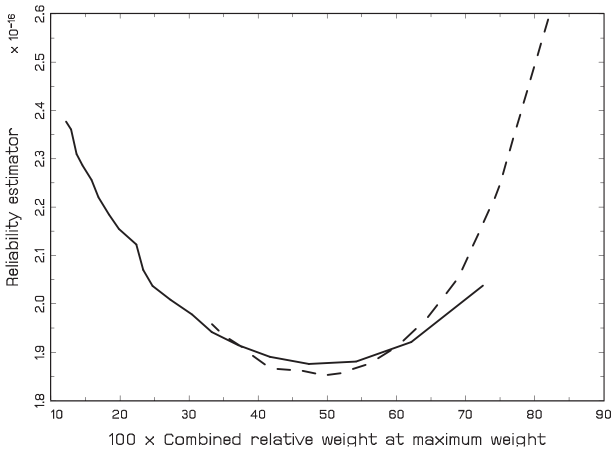

Figure 3. Reliability estimator $\mathcal{R}$ as a function of the combined weight of clocks at maximum weight for two normal distributions of clock instabilities (case S: ——, case L: - - - -, see text).

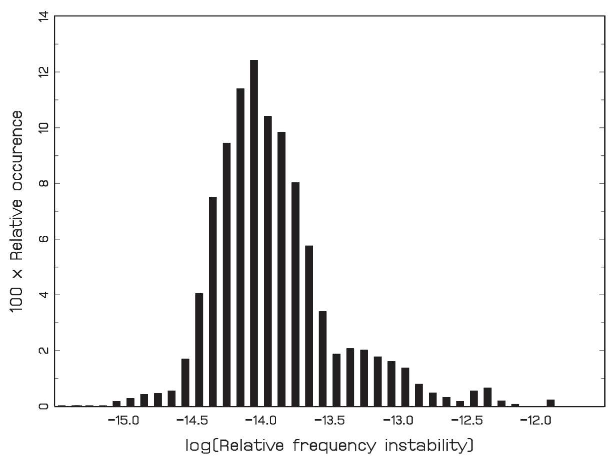

Figure 4. Average distribution of the 30-day relative frequency instability with respect to EAL for the TAI clocks over 1999–2001.

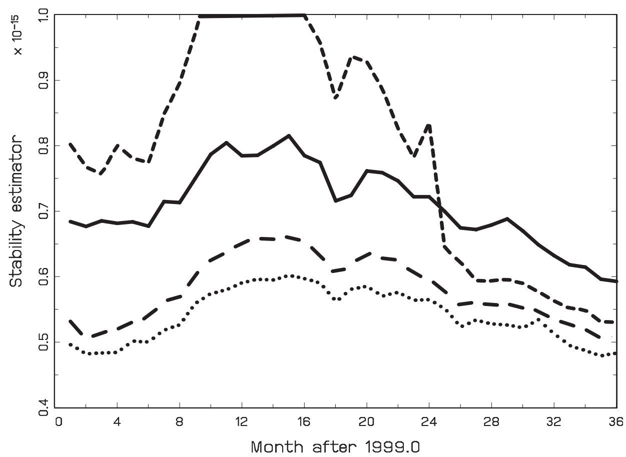

Figure 5. Stability estimator $\mathcal{S}$ for the ensemble scale computed with TAI clocks over 1999–2001 with a maximum weight of $1.5/N$ (——), $2.5/N$ (– – –), $4/N$ (· · · · · ·) and for the real TAI (- - - -, note that some points are outside the plot).

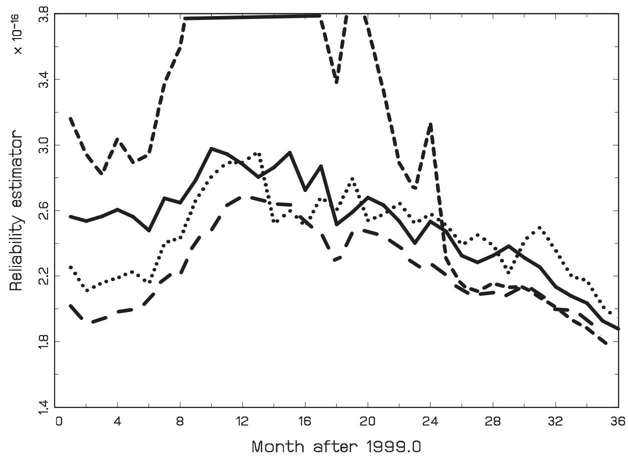

Figure 6. Reliability estimator $\mathcal{R}$ for the ensemble scale computed with TAI clocks over 1999–2001 with a maximum weight of $1.5/N$ (——), $2.5/N$ (– – –), $4/N$ (· · · · · ·) and for the real TAI (- - - -, note that some points are outside the plot).

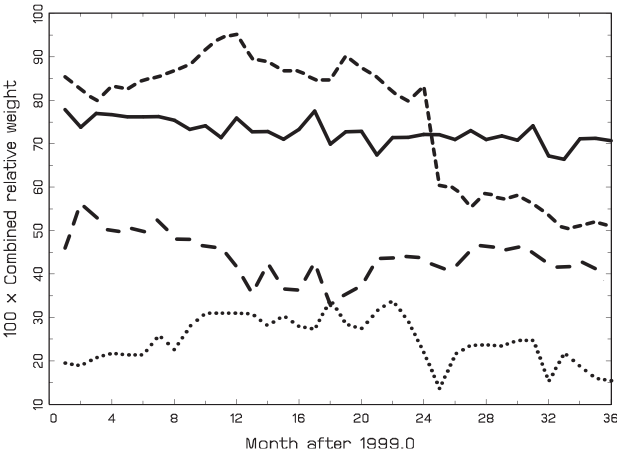

Figure 7. Combined weight of clocks at maximum weight for the ensemble scale computed with TAI clocks over 1999–2001 with a maximum weight of $1.5/N$ (——), $2.5/N$ (– – –), $4/N$ (· · · · · ·) and for the real TAI (- - - -).

(c) The actual situation (line with short dashes in figures 5 and 6) was far from optimal in 1999/2000, both in stability and in reliability. This was corrected by the change of the weighting scheme in 2001 but the optimal reliability is only reached around the very end of the period (figure 7), therefore leaving room to improve the stability without much loss in reliability.

(d) It is to be noted (figures 5 and 6) that the general situation is slightly worse between October 1999 and May 2000. This is possibly due, at least in part, to the smaller number of clocks available in that period (170 on average) compared to a general average of 181 over the three years. Also the situation improves starting August 2001, a period where the average number of clocks is 193. This indicates that, as obviously expected, it is no less important to try to improve the number and quality of the clocks than to use the proper algorithm. Nevertheless both actions are required for optimal results.

5. Conclusions: an optimal weighting scheme

We recall that procedure (P5), which could be particularly adapted to the strict application of minimizing the reliability estimator, is not advisable because, although it sets the combined weight of clocks at maximum weight, there is no provision on the number N_m of clocks that actually receive the maximum weight, so that, conceivably, only one clock could receive it. There is a very low probability that N_m becomes very small with the ensemble of more than 200 clocks for TAI, but we consider that it is not advisable to allow this situation to happen. We therefore promote a weighting scheme that is based both on ensuring that a very minimum percentage of the clocks reach the maximum weight and ensuring that the combined weight of clocks at maximum weight is close to, but slightly smaller than, half the total weight. This ensures near optimal reliability (because $\mathcal{R}$ is nearly minimized) and slightly better stability (because $\mathcal{S}$ is smaller than if we strictly fulfil the reliability criterion). It is practical to demand

(a) that a minimum of, for example, 5% to 7% of the clocks be at maximum weight, and

(b) that the combined weight of clocks at maximum weight be at least 40% of the total weight.

These conditions are readily applicable now (2002) with the ensemble of clocks presently used for TAI. If it happens in the future that these two conditions cannot be fulfilled at the same time (i.e. that assigning the maximum weight to the 7% most stable clocks would provide more than 50% of the total weight), it would mean that a small number of very stable clocks have appeared. Then, the whole design of the TAI computation would need to be reconsidered.

Acknowledgments

I am very grateful to my colleague Peter Wolf for numerous discussions and suggestions, notably the second reliability criterion envisioned in section 4.1.

Appendix. Minimization of the reliability estimator

Let $n(x)$ be the distribution function of instability x of the clocks, and let x_t be a given value of instability.

We denote by $N(x_t)$ the number of clocks whose instability is smaller than x_t, so that $N(x_t) = \int_0^{x_t} n(x)\,\mathrm{d}x$.

We denote by $V(x_t)$ the sum of the free weights of the clocks whose instability is larger than x_t, so that $V(x_t) = \int_{x_t}^{\infty} (n(x)/x^2)\,\mathrm{d}x$. The sum of all free weights is V_0.

Let $p(x_t)$ be the maximum normalized weight when x_t is the instability threshold (i.e. all clocks with instability lower than x_t take the maximum weight). The normalization factor $K(x_t)$ is such that $p(x_t) = K(x_t)/x_t^2$ and the sum of all free weights is

$$K(x_t)\left(\frac{N(x_t)}{x_t^2} + V(x_t)\right) = V_0 \qquad (1)$$

which implies by differentiation with respect to x_t

$$\frac{K'(x_t)}{K(x_t)} + \frac{N'(x_t)/x_t^2 - 2N(x_t)/x_t^3 + V'(x_t)}{N(x_t)/x_t^2 + V(x_t)} = 0 \qquad (2)$$

or, because $N'(x_t) = n(x_t)$ and $V'(x_t) = -n(x_t)/x_t^2$,

$$\frac{K'(x_t)}{K(x_t)} - \frac{2N(x_t)}{x_t^3 \times (N(x_t)/x_t^2 + V(x_t))} = 0 \qquad (3)$$

The reliability estimator defined in section 3.2 is $\mathcal{R} = 3 \times K(x_t)/x_t$ and it is minimum ($\mathcal{R}' = 0$) when $K'(x_t) = K(x_t)/x_t$, that is, by (3) when $N(x_t)/x_t^2 = V(x_t)$ or $K(x_t)N(x_t)/x_t^2 = V_0/2$, that is, the combined weight of the clocks at maximum weight is half of the total weight.

References

[1] Thomas C, Wolf P and Tavella P 1994 *BIPM Monograph* **94/1**
[2] Thomas C and Azoubib J 1996 *Metrologia* **33** 227–40
[3] Azoubib J 2000 *Report to the CCTF Working Group on TAI*
[4] Guinot B and Thomas C 1988 *Annual Report of the BIPM Time Section* **1** D1–22
[5] Petit G 1999 *Proc. 31st PTTI meeting* pp 297–304
[6] Yoshimura K 1980 *Metrologia* **16** 133–9