

Time and Frequency Technology Related to the Operation of a Cesium Fountain Primary Frequency Standard*

Thomas E. Parker

National Institute of Standards and Technology
Time and Frequency Division
325 Broadway
Boulder, CO 80305 USA

Abstract

The operation of a cesium fountain primary frequency standard is greatly facilitated by the presence of two other important capabilities. The first is a stable frequency reference and the second is one or more high-stability frequency-transfer systems. A stable frequency reference such as a hydrogen maser is virtually a necessity since essentially no fountain dead time can be tolerated without it. Also, some systematic biases in the fountain become very difficult to evaluate without a stable reference. State-of-the-art frequency transfer technology also is necessary if the fountain is intended to contribute to TAI, or to be compared to other remote frequency standards without excessive degradation of stated uncertainties. This paper discusses the facilities available at the National Institute of Standards and Technology and how they impact the operation of NIST-F1, the primary frequency standard at NIST.

1. Introduction

A systems approach is used in the operation of NIST-F1, the cesium-fountain primary frequency standard at the National Institute of Standards and Technology (NIST) [1]. The fountain is of course the heart of the operation, but two other important technologies play a significant role in the details of how the fountain is operated. One is the high-stability frequency standard that is used by the fountain as a reference. This includes not only the individual frequency references but also the time-difference measurement equipment that allows different frequency references at NIST to be related to each other and to create a time scale. The stability of this frequency reference system, which at NIST is based primarily on an ensemble of cavity-tuned, active hydrogen masers [2], impacts the amount of fountain dead time that can be tolerated without significantly degrading the uncertainty of the fountain measurements. It also influences how certain systematic biases in the fountain are evaluated. In this paper we will look in some detail at the procedures we use to evaluate the fountain spin-exchange bias and its uncertainty.

The other important technologies are long-distance time and frequency comparison techniques such as GPS (Global Positioning System) common view, GPS carrier phase, and Two-Way Satellite Time and Frequency Transfer (TWSTFT). The stability of these techniques contributes to the uncertainty of the comparison of remote fountains, or to the uncertainty of a comparison to TAI (International Atomic Time). Uncertainty in frequency transfer generally decreases in a manner inversely proportional to the length of the time interval of the comparison, which makes long comparisons desirable [3]. Long fountain runs consequently mean that high atom densities are not required to reach an acceptable statistical (type A) uncertainty. Using low atom densities results in a smaller spin-exchange bias and therefore a smaller spin-exchange uncertainty.

Consequently, the choice of operating parameters in a cesium fountain depends on the stability of the available frequency reference and the frequency transfer techniques being used.

*U.S. government work, not protected by U.S. copyright.

This is illustrated in Table 1, which summarizes the fractional frequency uncertainty values for a 60-day NIST-F1 evaluation that was submitted to the Bureau International des Poids et Mesures (BIPM) in June 2004 and published in the BIPM publication Circular T. Rows 1, 2 and 3 in the table give respectively the type B, type A and quadrature sum of the fountain uncertainties. These uncertainties have been discussed elsewhere [4] and will not be covered in detail here. Row 4 is the dead-time uncertainty and row 5 is the quadrature sum of rows 1, 2, and 4. The report interval for this evaluation was 60 days, but the fountain was in operation for only 34.7 days. Of the 25.3 days of dead time, 91 % was intentional, including 7 days added before and 11 after the fountain run in order to increase the length of the report period. This in turn decreased the frequency transfer uncertainty. As seen in the sixth row the frequency-transfer uncertainty was 5×10^{-16} , and a 60-day run was required to achieve this (see Section 3 below). However, the added dead time increased the uncertainty due to dead time. A roughly equal level of dead-time uncertainty and frequency-transfer uncertainty was chosen in order to give a total uncertainty (row 7) of 8.8×10^{-16} . This is the first report of a primary frequency standard into TAI with a total uncertainty of less than 1×10^{-15} .

2. Stable Frequency Reference

NIST operates an ensemble of five commercial hydrogen masers and four commercial cesium thermal beam standards to generate a real time timescale, AT1, which is used to produce UTC(NIST). This ensemble is also used to generate a post-processed time scale identified generically as TP171, or as AT1E, which is offset in fractional frequency by -483×10^{-15} from TP171 [5]. Though the cesium standards help improve the long-term stability of the ensemble, it is by far the masers that determine the ensemble performance. All of the clocks are maintained in environmentally controlled chambers. Optimum maser performance can be obtained only if both temperature and humidity are controlled, and the standards must also be kept in an environment with a relatively stable magnetic field [6]. Figure 1 shows the Allan deviation, $\sigma_y(\tau)$, with a linear frequency drift removed, for a better than average maser. Note that the stability of the maser is in the 10^{-16} range from about 1 day to 10 days. The post-processed maser ensemble, which performs better than any single maser, plays an important role in the operation of NIST-F1.

NIST-F1 is not operated as a clock. In normal operation NIST-F1 measures the frequency of one of the hydrogen masers (usually a maser with low drift and better than average short-term stability), and the output from NIST-F1 is a series of frequency-offset measurements made at intervals of every 2 to 3 seconds. Barring unexpected interruptions this data is usually reduced to a 24-hour average starting and ending at 0 UTC. The average maser frequency is then related to the frequency of TP171 through internal time-difference measurements. Our ability to do this depends, in part, on the stability of the instrumentation that measures the time (phase) difference of the 5 MHz signals coming from each maser and cesium standard. This data is used to create the time scale. We use a dual mixer system and have recently upgraded the equipment. Figure 2 shows the time deviation, $\sigma_x(\tau)$, (right axis) and Allan deviation, $\sigma_y(\tau)$, (left axis) for this new measurement system. Note that the time deviation is less than 1 ps out to about 10 days. At 1 day the Allan deviation is almost 1×10^{-17} . This is insignificant compared to the uncertainty of the fountain at 1 day. Even our older system, which is worse by about a factor of 5 to 10, does not contribute significantly to the measurement uncertainty.

Figure 3 shows the Allan deviation of NIST-F1 versus the reference maser in both a normal low atom-density mode, and also at a density about a factor of five higher. The stability of the maser itself is also shown by the dashed line. At low density the Allan deviation of NIST-F1 is well above that of the maser for all values of τ shown. In the high-density mode the fountain approaches the stability of the maser only at large τ values. This enables us to characterize the

stability of the fountain with very little impact from the masers. The white FM noise level obtained from the Allan deviation plot for each individual 24-hour run is used to determine the statistical uncertainty of that run, which is the Allan deviation at the run length.

Most systematic biases and their associated uncertainties in NIST-F1 are measured before and/or after a formal evaluation [1, 4], but some can be measured during the evaluation. The high stability of our local frequency reference gives us a unique flexibility in how we measure these biases. This is particularly true of the spin exchange bias which will be discussed in detail below as an example of a bias measured during the formal evaluation.

2.1 Spin Exchange Bias There are several possible approaches to measuring the spin exchange bias and its uncertainty. One is to measure the bias before and/or after the formal evaluation period and to use these numbers to correct the results of the evaluation. The assumption here is that nothing has changed to affect the bias during the evaluation. Another approach is to measure the spin exchange bias and uncertainty during the formal evaluation period, which is what we do. This minimizes the risk that something has changed, and also is an efficient use of fountain run time. There are two ways to implement this approach. One way is to quickly alternate, with a period of minutes to hours, between different atom densities. This minimizes the impact of any maser noise because it averages down. However, the mechanics of quickly changing atom densities usually limits the change in density to about a factor of two on NIST-F1. The other approach is to use large density changes, but these changes require adjustments to the laser power, cesium oven temperature, and molasses time, which may take 1 or more hours to accomplish. It has been verified that these changes do not affect the atom cloud profile [1]. In principal, we could always run in the high density configuration and use the state selection cavity to quickly reduce the atom density. However, this would require a large change in microwave power that may change the atom profile. To minimize the dead time from the large (slow) changes, the density cycle rate has to be on the order of days rather than hours. A larger density variation gives a smaller uncertainty in the slope but the slower cycle time puts more demand on the stability of the maser frequency reference. Our own independent estimates of the stability of the maser ensemble [2] indicate that it is sufficiently stable for this purpose. However, we now have enough fountain data to quantify the impact of the maser instabilities on the calculation of the spin-exchange uncertainty.

Figure 4 shows in chronological order the average fractional frequency values relative to TP171 for the 38 runs that made up the June evaluation of NIST-F1. Most runs are 24 hours in length, but a few were shorter due to intentional or unintentional interruptions. Excluding intentional interruptions, the fountain ran 94 % of the time. There were a total of 40 runs, but two were not used. One was excluded because of equipment problems and the other because it was made at a high microwave power. The evaluation was started at low density (1.0 in laboratory units) and 12 runs were made. Next we made 3 runs at medium density (2.25). This was followed by four 24 hour runs at high density (5.2). Finally we returned to medium density for 5 runs and low density for 14 runs. The error bars represent the statistical uncertainty for each run, which was determined from Allan deviation data. Note the smaller error bars on the higher-density runs. Not all runs lasted a full 24 hours and this accounts for some larger than normal error bars. A reasonable degree of time symmetry was used for the various densities in order to minimize the impact of any linear frequency drift in TP171. No post processing is performed on TP171 during an evaluation unless one of the clocks malfunctions. In the case of the June evaluation no post processing was required. Note that the large frequency offset between TP171 and NIST-F1 has no significance.

A weighted linear least-mean-square fit of frequency versus atom density is used with data such as that in Fig. 4 to determine the frequency and its uncertainty at zero atom density. The weights for each point are determined from the statistical uncertainties (error bars) for each run. Figure 5 shows fractional frequency versus atom density for both the June 2004 and the

December 2003 NIST-F1 evaluations. For clarity of presentation all the data for each density in the individual evaluations has been combined into one point with an appropriate error bar. The fit results are shown by the two solid lines, and the slopes agree within their uncertainties. Note that the frequency shift from a density of 1 to 0 is only about 5×10^{-16} . The uncertainties for the intercept and slope were 5.6×10^{-16} and 1.6×10^{-16} respectively for the December run, and 5.1×10^{-16} and 1.5×10^{-16} for the June run.

In the fitting routine that we use, the uncertainties of the intercept and the slope are based only on the uncertainties for each run, and are not determined by the actual scatter of the data [7]. If the white FM noise characteristic seen in Fig. 3 extends out to the duration of the evaluation these calculated uncertainties will be correct. However, if the noise of the maser ensemble begins to degrade the long-term stability of the measurements in Fig. 4, the uncertainties from the weighted fit will be underestimated. The uncertainty of the intercept is the statistical (type A) uncertainty of the evaluation and includes the uncertainty of the spin exchange bias (which is also type A) [4].

Figures 6 and 7 show (in chronological order) the residuals of the fits to atom density for the June and December evaluations. Each data point is a one-day average, and some short runs have been combined. These data now constitute a time series of frequency data covering a total of nearly 56 days, from which the spin-exchange bias has been removed. Though there is some dead time in the data for each evaluation, an Allan deviation calculation can be performed since we are dealing primarily with white FM noise. Figure 8 is a composite Allan deviation plot using 14 days of 2-second (measurement cycle by measurement cycle) low-density data from the June evaluation (solid circles), and also using the daily averages from both the June and December runs. The 24-hour averages include all densities (in order to minimize the dead time), and this biases the Allan deviation values slightly low. However, 72 % of the 24-hour data is from low-density data so the bias is only about 12 %, which is negligible on the scale of Fig. 8. The Allan deviation values from the December and June runs were averaged together to give better confidence levels (the runs were not concatenated because of the long interval between them). TOTAL and Theol deviations [8] were also used for the larger τ values to further improve the confidence levels. For comparison an estimate of the stability of AT1E (or TP171) is also shown. Note that the stability of AT1E is better in the long term than an individual maser because the maser frequency drift is accounted for in AT1E. The important observations to be made from Figure 8 are that, (1) the noise of the fountain at low density (where most of the data is collected) is white FM all the way out to time intervals on the order of the length of an evaluation; and (2) that the fountain noise level is well above that of the maser ensemble except at the largest τ values. The first observation above verifies that our assumptions about the stability of the maser ensemble are correct.

A more quantitative estimate of the impact of maser noise on our estimates of the uncertainty of the intercept can be obtained with a chi-squared analysis [7] of the data in Figs. 6 and 7. The reduced chi squared is given by

$$\chi_r^2 = \frac{1}{d} \left[\sum_{i=1}^n \frac{(y_i - (sx_i + b))^2}{u_i^2} \right] , \quad (1)$$

where n is the number of data points, $d = n - 2$ is the number of degrees of freedom, y_i is the i^{th} 24-hour frequency average, with a frequency uncertainty of u_i , for atom density x_i . s and b are respectively the slope and intercept of the weighted linear least-mean-square fit. The term in the numerator in the brackets is the square of the residuals of the fit. We can define R as

$$R = \sqrt{\chi_r^2} , \quad (2)$$

and it represents the rms ratio of the standard deviation of the fit (many days of data) to the uncertainty expected from the fountain white FM noise level measured each day. If the white fountain noise is the dominant noise source over the course of an evaluation the average value of R should be 1. For the December and June evaluations the R values were respectively 0.94 and 1.19. The weighted mean of these two is 1.10, where the weighting is based on the duration of the evaluation run times. The weighted average of R from the eight most recent NIST-F1 evaluations is 1.06. This indicates that, on average, we have been underestimating the uncertainty of the intercept by at most 6 %. Given that the uncertainty of the intercept typically makes up about one quarter (added in quadrature) of the total uncertainty reported into TAI, the additional uncertainty due to the maser noise is a negligible contribution. By using a large range of atom densities (which necessitates a slow cycle time between different densities) we ultimately reduce the uncertainty due to the spin-exchange bias, but this can be done only because of the exceptionally high stability of our maser ensemble. For the June 2004 evaluation the uncertainty of the intercept would have been 70 % larger if the four high density runs were replaced by medium density runs of equivalent length. The uncertainty of the slope would be more than three times larger.

2.2 Dead Time In addition to influencing the procedures used to evaluate frequency biases, the stability of the reference masers also impacts on how much dead time can be tolerated [5]. Though we have made great progress in reducing unintentional dead time in NIST-F1, we can still use intentional dead time to reduce the overall uncertainty of a long-distance frequency comparison when the uncertainty of frequency transfer is taken into consideration. As will be discussed in Section 3, the frequency-transfer uncertainty decreases as the time interval increases. Consequently a longer evaluation gives a smaller transfer uncertainty. On the other hand, increasing dead time results in a larger value for the uncertainty of the maser frequency measurement. Figure 9 shows how the uncertainty of a 30-day frequency measurement interval increases as the amount of live time decreases (dead time increases) [5]. This curve is based on typical maser noise characteristics shown in the figure. Note that it makes a difference where the live time is located in the 30 day interval.

Since the uncertainties due to frequency transfer and dead time are both based primarily on previous evaluations of instabilities, and are of similar quality, it is reasonable to trade off one for the other to obtain the lowest total uncertainty. This occurs when the report period of a fountain evaluation is arbitrarily increased symmetrically with dead time to where the uncertainties from dead time and frequency transfer are approximately equal. This was done in the June 2004 evaluation discussed earlier and to a lesser extent on all previous NIST-F1 evaluations. As a practical matter, no matter how long the fountain is run, it is almost always desirable for us to add dead time to obtain the minimum overall uncertainty.

3. Frequency Transfer

As mentioned earlier, the uncertainty of frequency transfer plays a significant role in the operation of a fountain. In this section we will present a review of some recent data on the stability of various time/frequency transfer techniques. Figure 10 shows the time deviation of a comparison between the maser ensembles at NIST and the United States Naval Observatory (USNO) using three different time (frequency) transfer techniques [9]. The three techniques are common-view GPS, two-way satellite time and frequency transfer (TWSTFT), and carrier-phase GPS. In all cases the value of $\sigma_x(\tau)$ represents the noise of the time transfer system at τ 's smaller than a few days because the maser ensembles are very quiet. The noise in common view dominates out to about 20 days, but even TWSTFT and carrier phase, which are more stable, dominate over the clock noise out to about 5 days. In this figure we can see that the instabilities in common view are flicker PM in nature (no dependence on τ) at a level of about 1

ns from less than a day out to beyond 10 days. TWSTFT is intermediate in stability at about 200 ps, with evidence of a diurnal (daily) cycle at τ close to half a day. As with common view, the time deviation of two-way is predominantly flicker PM until the clock noise sets in. GPS carrier phase is the most stable technique at short time intervals, but its noise increases with τ at a rate approximating random-walk PM (white FM). Though this resembles clock noise the level is too high to be caused by the maser ensembles for τ less than about 5 days. Figure 10 is a good example of the stability characteristics of these three time transfer techniques for τ less than 5 to 10 days (beyond that clock noise dominates).

To eliminate clock noise we can take the difference between two independent time-transfer techniques using the same pair of clocks. Figure 11 shows the time difference between common-view and TWSTFT (the line with high short-term noise), and carrier phase and TWSTFT for hydrogen masers at NIST and Physikalisch Technische Bundesanstalt (PTB) in Germany over a period of about 1000 days. Figure 12 shows the time deviation for these data, and Figure 13 shows the Allan deviation. Though clock noise has been eliminated, we now have the combined noise of two time/frequency transfer techniques. Assuming that common-view noise is larger than TWSTFT noise (as Fig. 10 indicates) we again see that the time deviation of common view is flicker PM in nature at a level of about 1 ns, but now out to beyond 200 days. The carrier phase minus TWSTFT time deviation starts at a level of about 300 ps (probably dominated by the two-way noise) but rises to above 2 ns at 100 days. The high level at 100 days is clearly due to carrier phase. The long-term stability of carrier phase and other time transfer techniques is currently a subject of investigation and is not yet well understood. Better long-term carrier-phase stability than that shown in Figs. 12 and 13 has been observed in other studies [10].

Note in Figure 13 that the Allan deviation of common-view minus TWSTFT (probably dominated by common-view) decreases approximately as $1/\tau$. Since the noise characteristic is predominately flicker PM the decrease is actually slightly slower than $1/\tau$. If we assume that the ratio of the stability of TWSTFT to common-view is the same in the long term as it is at a few days (this has not yet been verified) then we would expect the Allan deviation of two-way instabilities to also have an approximate $1/\tau$ dependence, but at a level 3 to 4 times lower. Note that the carrier phase minus TWSTFT does not decrease nearly as rapidly. The Allan deviation is a good estimate of the uncertainty of a frequency comparison over the interval τ . When directly comparing two remote fountains it is clearly desirable to use two-way or carrier phase (for smaller τ), or preferably both, rather than common view.

Figure 14 shows the time and Allan deviations of NIST's AT1E versus TAI (data on TAI is obtained from the BIPM publication Circular T). This provides information on time/frequency-transfer instabilities into TAI. Clock noise dominates beyond about 30 days, but for smaller τ time/frequency-transfer noise dominates. The time-transfer links into TAI are made up of both TWSTFT and common-view GPS. For small τ we again see that the time deviation is flicker in nature at a level of about 1 ns. The solid line in Fig. 14 represents the expression $u_{\text{link/TAI}} = 3 \times 10^{-14}/\tau$ (for τ in days) currently used by the BIPM for the uncertainty of frequency comparisons with TAI. We see that it is in reasonably good agreement with the AT1E minus TAI Allan deviation in the region where time/frequency-transfer noise dominates. A report duration of 30 days is required to reduce the frequency-transfer uncertainty to 1×10^{-15} .

4. Summary

The operation of a cesium fountain primary frequency standard requires a significant amount of time and frequency infrastructure, and the nature of this infrastructure impacts on the details of how the fountain is operated. When frequency comparisons are made between remote fountains, or when the fountain is compared to TAI, one must take into account the stability of

the local frequency reference and the stability of the frequency transfer technique if a minimum uncertainty in the comparison is to be achieved. This requires a thorough understanding of the instabilities in these systems. A high-stability local frequency reference allows improved evaluation of systematic biases, and a careful use of intentional dead time can be used to minimize the total uncertainty of a long-distance comparison.

Acknowledgements

The author gratefully acknowledges the many people who contribute to the operation of NIST-F1: including Steve Jefferts, Tom Heavner and Elizabeth Donley of the NIST-F1 team; and Jim Gray, Judah Levine, Trudi Pepler, Marc Weiss, and Victor Zhang who operate the time scale and the time and frequency dissemination systems. The cooperation of the groups at USNO and PTB in the time transfer experiments is also gratefully acknowledged.

References

1. S.R. Jefferts, J.H. Shirley, T.E. Parker, T.P. Heavner, D.M. Meekhof, C.W. Nelson, F. Levi, G. Costanzo, A. DeMarchi, R.E. Drullinger, L. Hollberg, W.D. Lee, and F.L. Walls, "Accuracy evaluation of NIST-F1," *Metrologia*, vol. 39, pp. 321-336, 2002.
2. T.E. Parker, "Hydrogen maser ensemble performance and characterization of frequency standards," in *Proc. of the 1999 Joint Meeting of the IEEE International Frequency Control Symp. and the European Frequency and Time Forum*, pp. 173-176, 1999.
3. T.E. Parker, P. Hetzel, S.R. Jefferts, S. Weyers, L.M. Nelson, A. Bauch, and J. Levine, "First Comparison of Remote Cesium Fountains," in *Proc. of the 2001 IEEE International Frequency Control Symposium*, pp. 63-69, 2001.
4. T.P. Heavner, S.R. Jefferts, E.A. Donley, J.H. Shirley, and T.E. Parker, "Recent improvements in NIST-F1 and a resulting accuracy of $\delta f/f = 0.61 \times 10^{-15}$," *IEEE Trans. on Instrum. and Meas.*, to be published.
5. T.E. Parker, "Comparing and Evaluating the Performance of Primary Frequency Standards: Impact of Dead Time," in *Proc. of the 2001 IEEE International Frequency Control Symposium*, pp.57-62, 2001.
6. T.E. Parker, "Environmental Factors and Hydrogen Maser Frequency Stability," *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 46, no. 3, pp.745-751, 1999.
7. P.R. Bevington, "Data Reduction and Error Analysis for the Physical Sciences," McGraw-Hill Inc. 1969.
8. D.A. Howe and T.K. Pepler, "Very Long-term Frequency Stability Estimation using a Special-purpose Statistic," in *Proc. of the 2003 Joint Meeting of the IEEE International Frequency Control Symp. and the European Frequency and Time Forum*, pp. 233-238, 2003.
9. T.E. Parker, V.S. Zhang, A. McKinley, L.M. Nelson, J. Rohde, and D. Matsakis, "Investigation of instabilities in two-way time transfer," in *Proc. of the 2002 Precise Time and Time Interval Meeting*, pp. 381-390, 2002.
10. K.M. Larson, J. Levine, L.M. Nelson, and T.E. Parker, "Assessment of GPS Carrier-Phase Stability for Time-Transfer Applications," *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 47, pp 484-494, 2000.

Table 1. Summary of Uncertainties for the June 2004 Evaluation of NIST-F1

NIST-F1 Systematic Uncertainty (Type B)	3.3×10^{-16}
NIST-F1 Statistical Uncertainty (Type A), includes spin exchange uncertainty	5.1×10^{-16}
Combined Fountain Uncertainty	6.1×10^{-16}
Dead Time Uncertainty	4.0×10^{-16}
Combined Uncertainty including dead time	7.3×10^{-16}
Frequency Transfer Uncertainty	5.0×10^{-16}
Total Uncertainty into TAI	8.8×10^{-16}

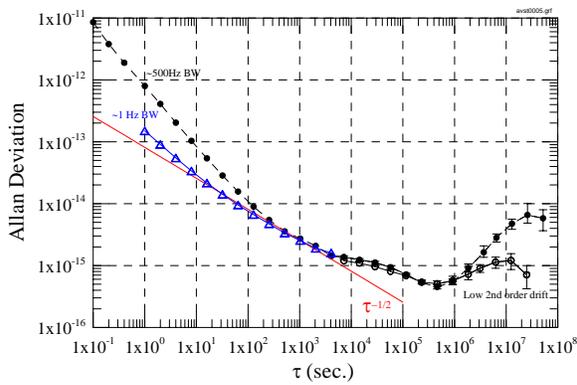


Figure 1. Allan deviation of a hydrogen maser.

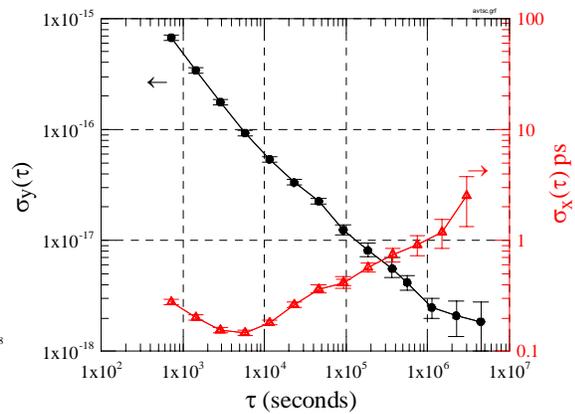


Figure 2. Stability of measurement system.

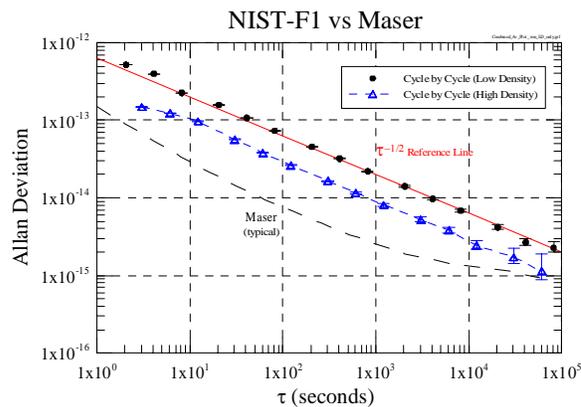


Figure 3. Allan deviation of NIST-F1.

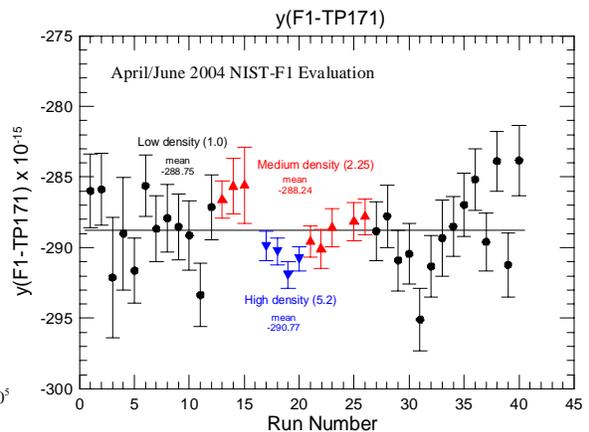


Figure 4. Series of NIST-F1 runs for the June evaluation

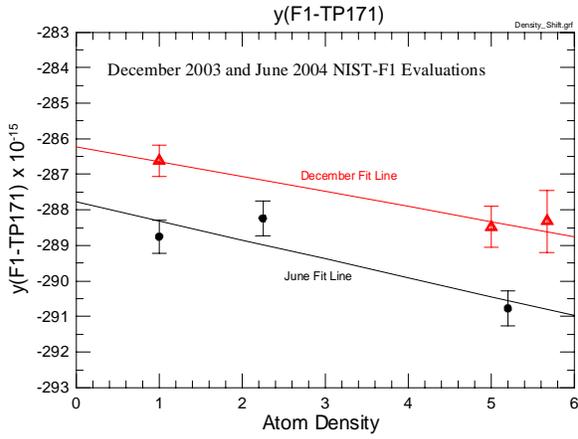


Figure 5. Least-mean-square fit to atom density.

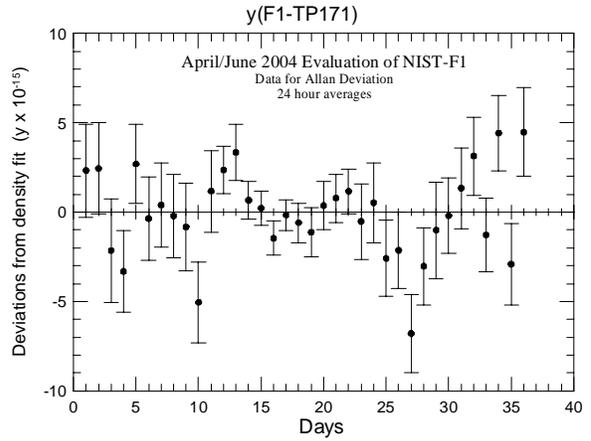


Figure 6. Residuals of June evaluation.

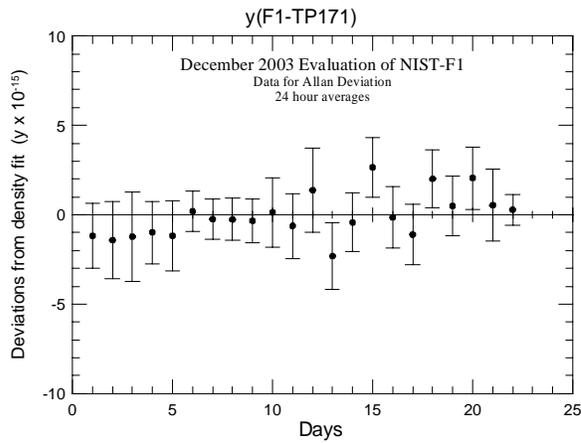


Figure 7. Residuals of December evaluation.

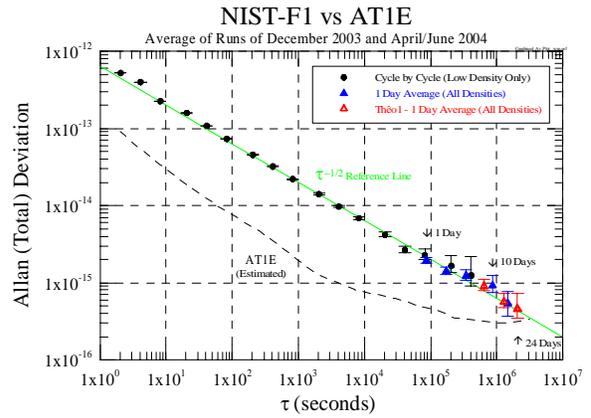


Figure 8. Allan deviation of NIST-F1.

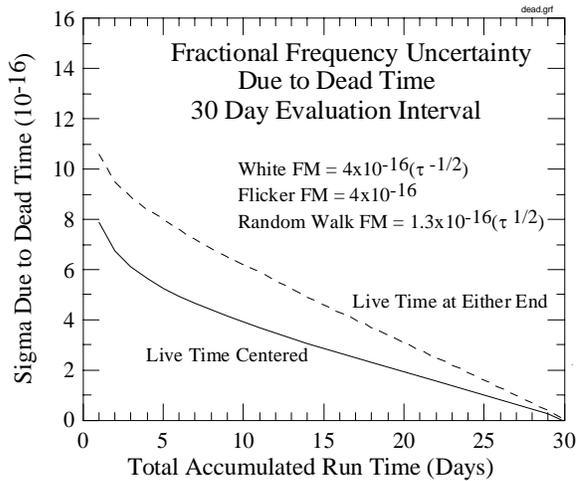


Figure 9. Uncertainty due to dead time.

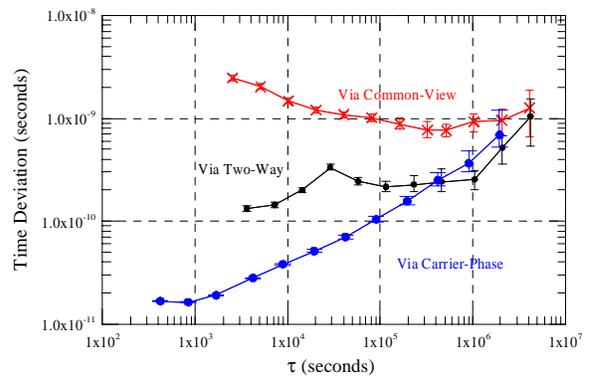


Figure 10. Time deviation of time transfer.

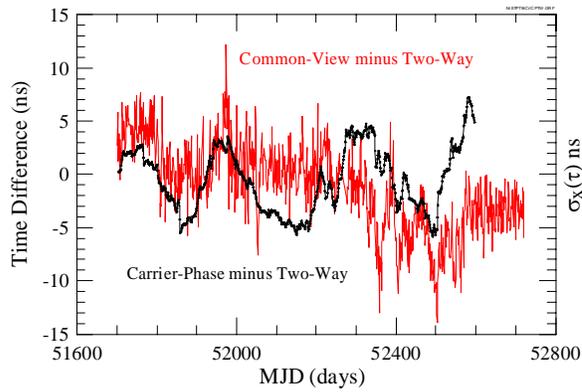


Figure 11. Time transfer from NIST to PTB.

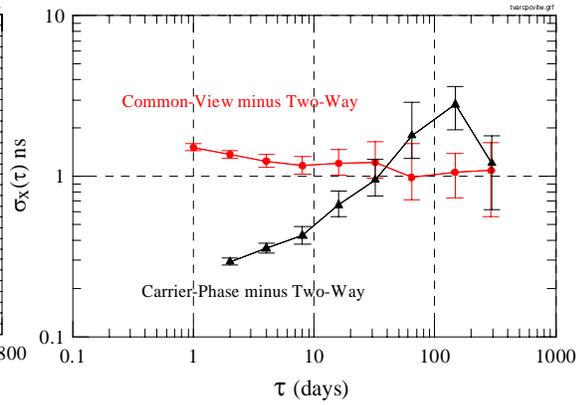


Figure 12. Time deviation of data in Fig. 11.

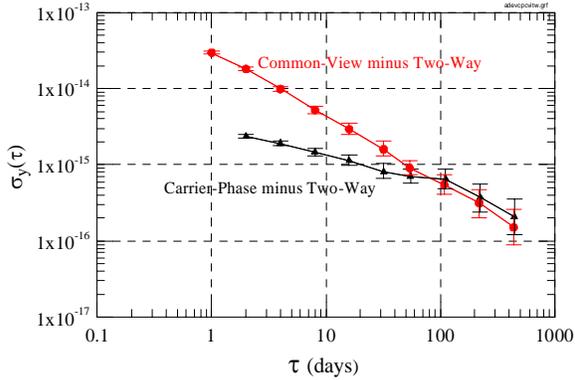


Figure 13. Allan deviation of data in Fig. 11.

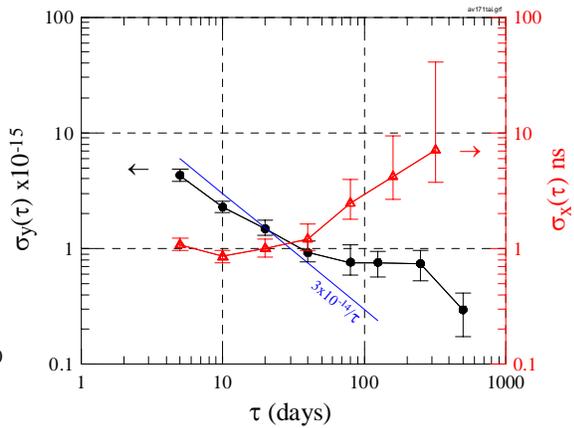


Figure 14. Stability of AT1E vs TAI.