

COMPARING AND EVALUATING THE PERFORMANCE OF PRIMARY FREQUENCY STANDARDS: IMPACT OF DEAD TIME

T. E. Parker

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

Abstract

Comparing high-performance frequency standards often requires making comparisons over displacements in time and/or space. The uncertainties introduced into comparisons made with dead time or with displacements in time are discussed. The estimates of uncertainty are based on the noise characteristics of an ensemble of cavity-tuned hydrogen masers used as a stable frequency reference. Specific examples are given.

Key Words: Ensemble, frequency comparisons, hydrogen masers, uncertainties.

Introduction

Comparisons with other standards, and comparisons with the same standard operated at different times, are necessary steps in evaluating the performance of a primary frequency standard (PFS). Thus, one often has to deal with both displacements in space and time. Standards in different laboratories can be widely separated around the world. This, in many ways, is a desirable feature since it minimizes common-mode environmental effects, but it also requires the use of long-distance, high-stability time or frequency transfer techniques. The issues of long distance comparisons are discussed in another paper in these proceedings [1]. In the present paper the techniques for handling displacements in time are discussed.

Displacements in time necessitate a different set of tools than those for displacements in space. Specifically, comparisons over time require a stable (but not necessarily accurate) frequency reference. In the Time and Frequency Division of the National Institute of Standards and Technology (NIST) we use an ensemble of five cavity-tuned hydrogen masers and four high-performance, commercial cesium standards for this purpose. The ensemble uses a post-processed time-scale algorithm similar to our real-time time scale AT1. Details of this ensemble have been previously discussed [2], [3].

The post-processed scale as calculated is identified as TP171. It has a large frequency offset since it was started

from the state of AT1 in April of 1997. When a nominal offset of 483×10^{-15} is removed to bring TP171 close to the rate of International Atomic Time (TAI) the new scale is referred to as AT1E. The name AT1ES is used when a frequency offset *and* a linear slope are removed from TP171.

The availability of a highly stable frequency reference has become particularly important at NIST with the development of the new frequency standards that use lasers for optical pumping and/or atom cooling. It has generally not been possible to operate these standards continuously because they are inherently complicated, and the lasers are not yet fully reliable. Consequently the standards are difficult to operate on demand, and evaluations are commonly interrupted by loss of laser lock. As a result it is difficult to have a long continuous run, or to have two standards operate at exactly the same time. However, with a sufficiently stable frequency reference significant dead time can be tolerated with no major impact on the uncertainty of an evaluation. Also, when comparing two different standards it is not necessary that the runs overlap perfectly. The impact on the comparison uncertainty is generally significantly larger for time offsets than for dead time, but meaningful comparisons can still be made. The same approach can also be used to compare different runs on the same standard. This is very helpful in evaluating the run-to-run stability of a standard, which is important in determining whether a standard is operating properly.

The stability characteristics of AT1E are reviewed here, followed by quantitative examples of the impact of dead time and measurement time offsets using AT1E as a reference. Finally a discussion of how AT1E is used to help evaluate the performance of several primary frequency standards is presented.

Frequency Stability of AT1E

The short-term stability of TP171 (AT1E) has been discussed in previous papers [2]-[4]. It is better than 1×10^{-15} at one day, and often approaches 3×10^{-16} at ten days. Currently, these stability levels can be determined only by internal measurements among the masers at NIST, but

future improvements in fountain stability should provide an independent evaluation of the short-term stability of the ensemble.

The long-term stabilities (frequency drift and random-walk FM) of this maser based scale have not previously been discussed in detail, so this issue will be addressed here. The problem with using hydrogen masers (even with cavity tuning) in a time scale is that their frequencies drift at relatively high rates, on the order of $\pm 1 \times 10^{-16}$ /day. This necessitates some means of modeling drift in the algorithm. If the frequency drift were constant the problem would be fairly simple to handle, but in hydrogen masers the drift is continuously changing. A procedure had to be developed to detect changes in the drift rate. In principle a PFS could be used to estimate drift, but when TP171 was first started only thermal-beam devices were in regular operation. Unfortunately, the uncertainties and instabilities in thermal beam standards are sufficiently large that it takes several years to detect changes in maser drift rates at the 1×10^{-17} /day level (10 % of a typical maser drift rate). However, changes of this magnitude may occur several times per year in some masers. Thus, another approach had to be used.

When TP171 was started we were fortunate to have had a long history with a relatively stable maser. Thus we could estimate its drift rate fairly accurately. This provided

a means to estimate the drift for all of the masers as the scale was started. From that point on the primary method for detecting changes in drift rate has been to compare the frequencies of individual masers to the frequency of TP171. However, these masers are all in the scale so that an apparent change in drift rate is always underestimated. (The maximum weight of a maser in the scale is limited to 30 %.) A simple iterative procedure was developed to handle this problem – the drift parameter was changed until the observed drift relative to the scale agreed with the modeled drift. The main weakness in this procedure is that for all of the masers the drift tends to become less positive, or more negative. This common-mode change is not detectable from inside the scale and can be seen only relative to independent references. Nevertheless the technique has worked fairly well. Over almost four years of operation TP171 has exhibited a nearly linear downward drift of less than 1×10^{-17} /day. With more access to regular fountain data it is expected that this downward drift can be reduced.

Figure 1 shows the fractional frequency difference between TP171 and seven major primary frequency standards from three national laboratories – NIST, Physikalisch-Technische Bundesanstalt (PTB), and Laboratoire Primaire du Temps et des Fréquences (LPTF). These standards include three fountains (LPTF-FO1, NIST-

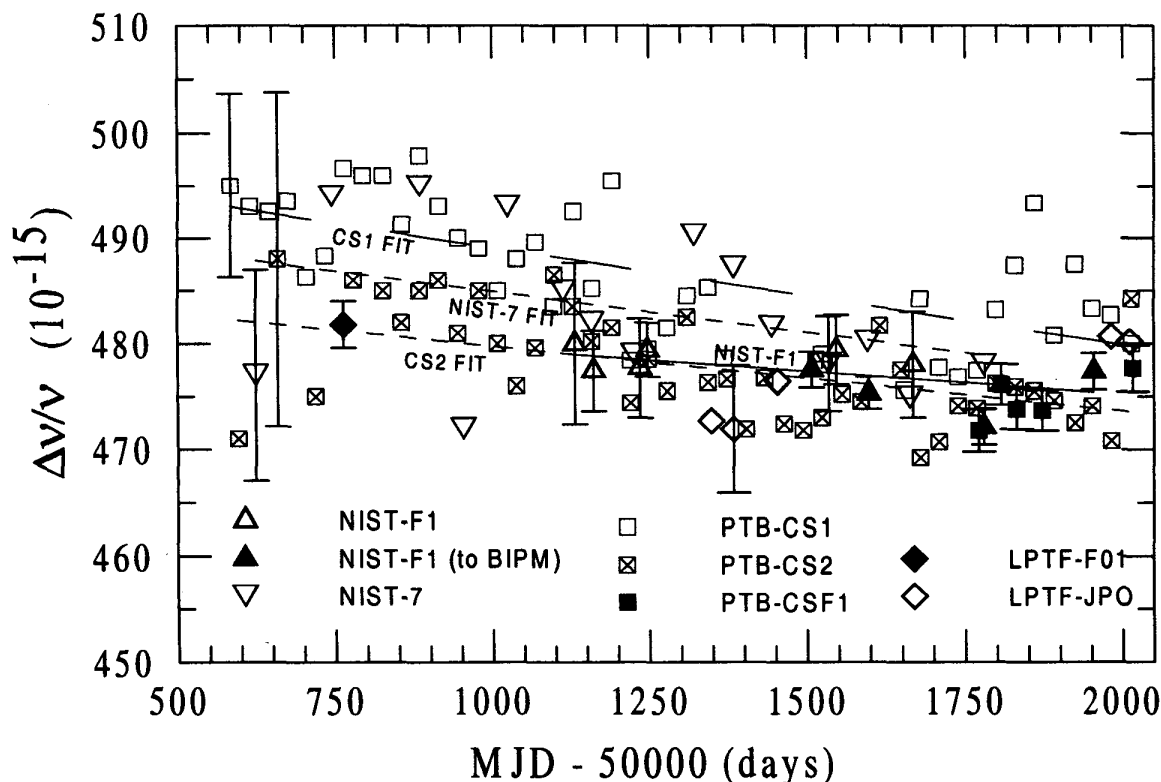


Figure 1. Fractional frequency offset of TP171 relative to seven primary frequency standards.

F1 and PTB-CSF1), and four thermal beam standards (PTB-CS1, PTB-CS2, NIST-7 and LPTF-JPO). The error bars indicate representative uncertainties. All of the data points are from formal evaluations that have been reported to the Bureau International des Poids et Mesures (BIPM), except for six informal runs of the NIST fountain, NIST-F1, shown as hollow triangles. The dashed or solid lines are linear fits to the various standards. The -0.5×10^{-17} /day frequency drift of TP171 is clearly visible in Fig. 1.

For comparison Fig. 2 shows the fractional frequency difference between TP171 and TAI, while Fig. 3 shows the fractional frequency difference between TP171 and the BIPM's free atomic scale, EAL. Differences between the curves in Figs. 2 and 3 are due to the steering of TAI. In Fig. 2 some of the downward slope between MJD 50500 and 51000 was caused by steering of TAI to remove the blackbody correction. The downward frequency drift of TP171 apparent in Fig. 1 is not as obvious in Figs. 2 and 3.

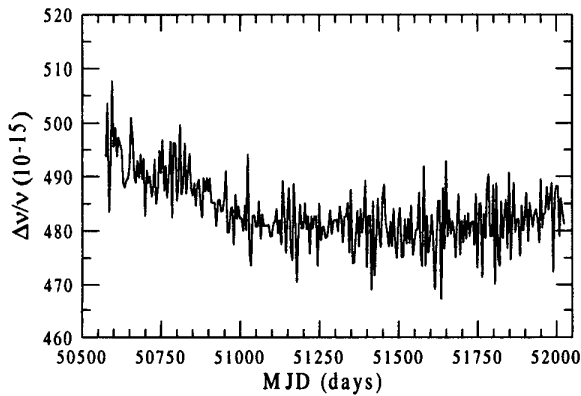


Figure 2. Fractional frequency offset of TP171 relative to TAI.

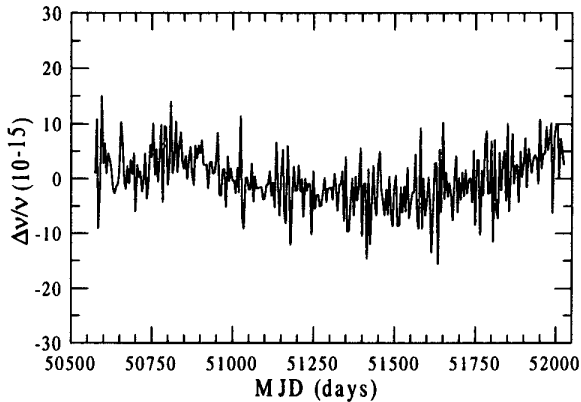


Figure 3. Fractional frequency offset of TP171 relative to EAL. An arbitrary frequency offset has been removed.

The random-walk FM noise level of AT1E was estimated from PTB-CS2 data using the Hadamard variance so that frequency drift did not have to be removed. PTB-CS2 was used since it is the best continuously operated frequency standard and has operated over the entire history of AT1E. AT1E was also compared with EAL as a second check of the random-walk noise level. Table 1 summarizes the results of a fairly conservative estimate of the noise characteristics, $\sigma_y(\tau)$, of AT1E in terms of the three basic noise types (white FM, flicker FM and random-walk FM). τ is in units of 1 day. The stability characteristics of EAL as published by the BIPM in Circular T are also shown for comparison. The main advantage of AT1E is a much lower level of white FM noise. This is to be expected since EAL is strongly influenced by the noise in GPS common-view time transfer.

Table 1. Frequency Stability Characteristics of AT1E and EAL

Scale	White FM	Flicker FM	RW FM
AT1E	$4 \times 10^{-16} (\tau^{-1/2})$	4×10^{-16}	$1.3 \times 10^{-16} (\tau^{1/2})$
EAL	$60 \times 10^{-16} (\tau^{-1/2})$	6×10^{-16}	$1.6 \times 10^{-16} (\tau^{1/2})$

Dead Time and Time Offsets

The noise characteristics in Table 1 along with the method of Douglas and Boulanger [5] provide a means for estimating the one-sigma uncertainty of a frequency comparison with dead time, or with an offset in the measurement intervals [4]. Figure 4 shows the uncertainty introduced into a 30-day measurement of the frequency of AT1E if the PFS has some dead time (that is the standard

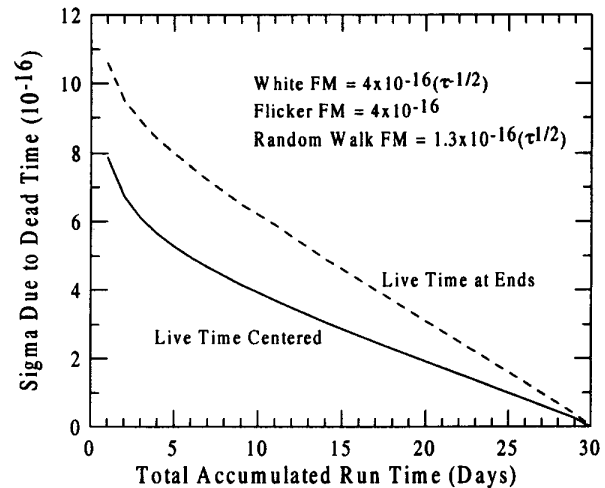


Figure 4. Fractional frequency uncertainty due to dead time in a 30 day evaluation interval.

does not run continuously throughout the 30-day period.). The solid line represents the situation where all of the live time is centered in the middle of the 30-day measurement interval, while the dashed line is for all of the live time located at either end. The noise characteristics used for the calculation are also shown in the figure. As expected if the dead time is near zero the uncertainty due to dead time is also near zero. With increasing dead time the uncertainty increases and its value depends on whether the live time is in the center of the interval or at the ends. If the noise characteristics were only white FM the location of the live time would not matter, but with flicker and random-walk noise the location is important. Measurements near the ends of the interval introduce more uncertainty than those in the middle.

At this time there is no rigorous method for calculating the uncertainty due to flicker or random walk noise in the presence of distributed dead time (that is the live time does not occur in just one interval, but in many smaller intervals). However, estimates have shown that distributed dead time results in uncertainties that are generally smaller than those for live time at the center [4]. Figure 4 shows that for 15 days of operation in a 30-day interval (50 % dead time) the uncertainty due to dead time is less than 5×10^{-16} . For a PFS with a combined uncertainty of 1.8×10^{-15} the dead time would make a nearly negligible contribution (the total uncertainty would be no worse than $\sim 1.87 \times 10^{-15}$). Thus a good stable frequency reference such as AT1E, or even a well-characterized and well-behaved single maser, will allow the presence of up to 50 % dead time in a 30-day period without significantly increasing the uncertainty of a measurement.

Figure 5 shows the impact of an offset in the measurement intervals on the uncertainty of 30-day frequency measurements, again using the procedure of [5]. Here two 30-day measurements of AT1E that are offset in

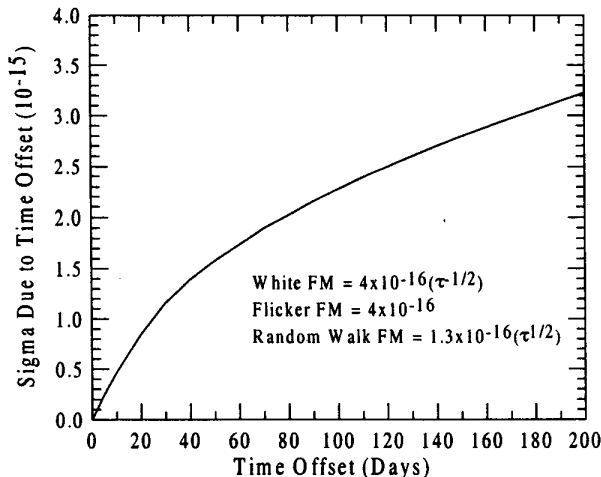


Figure 5. Fractional frequency uncertainty due to a time offset for a 30-day evaluation interval.

time are being considered. The measurements may be made by two different primary frequency standards or may consist of two measurements with the same standard. With two standards the measurements may overlap either partially or completely. If the overlap is perfect (measurement offset is zero) there is no additional uncertainty since both standards are measuring the same thing at the same time. However, if there is an offset in measurement interval the instability of AT1E becomes an important factor in how well one PFS can be compared to the other. Figure 5 shows that for the noise characteristics of AT1E the offset must be less than about 12 days to keep the comparison uncertainty to less than 5×10^{-16} . However, the uncertainty increases roughly as the square root of the offset time so it does not grow rapidly. By 200 days offset the uncertainty is just a little over 3×10^{-15} . Thus the maser ensemble can be used to compare standards at the mid-to-low 10^{-15} level over intervals of hundreds of days.

In making frequency estimates with either dead time or measurement interval offsets the frequency drift of the reference must also be properly accounted for.

Comparing Primary Frequency Standards

The primary frequency standards at NIST are not operated as clocks. Formal evaluations are performed several times per year in which the frequency of one of the masers is measured, typically over a 30-day period. The maser is one of the clocks normally reported to the BIPM, and therefore provides the means for relating the frequency of the PFS to TAI. The frequency of the reference maser can also be related to the frequency of TP171, with negligible uncertainty.

Figure 6 shows the fractional frequency offset of TP171 since its inception as determined by the two NIST primary frequency standards. All of the NIST-7 measurements shown were reported to the BIPM, as well as the measurements for the cesium fountain, NIST-F1, indicated by solid triangles. The hollow triangle points

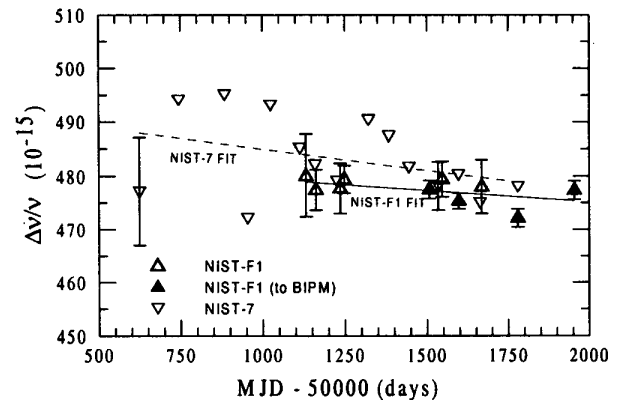


Figure 6. Fractional frequency offset of TP171 relative to NIST primary frequency standards.

represent informal NIST-F1 runs. The larger uncertainties for these informal measurements are due to the short duration of the runs. Figure 6 provides a means to compare the two standards, and also helps to illustrate the run-to-run repeatability of both standards. Such a comparison using TP171 is better than with TAI or EAL because of the lower white FM noise in TP171.

Theoretically NIST-7 may be capable of accuracy better than 5×10^{-15} [6], but many of the runs prior to MJD 51100 in Fig. 6 show peak-to-peak instabilities at the 2×10^{-14} level. The repeated occurrence of instabilities outside of the range of known uncertainties clearly indicates the presence of unknown (and unstable) biases. After a number of improvements [6] were made to the standard a more stable performance was eventually achieved. NIST-7 has generally shown good agreement with the fountain over the last 600 days. The two NIST-7 runs near MJD 51300 that show large deviations were made with a single beam direction and therefore had larger uncertainties. The NIST-F1 runs have all been reasonably stable, and are consistent with the stated one-sigma uncertainties of the standard and the uncertainty due to the time offset between runs.

Plots such as Fig. 6 are very useful in identifying problems. On both NIST standards there have been instances where the observations of large frequency offsets have led to the discovery of previously unknown biases. In

some cases these biases were caused by new and unexpected hardware failures, which sometimes can be very subtle. Any problems must be understood or fixed before a formal evaluation can be carried out. During a formal evaluation a single maser is normally used as the physical reference and the presence of ATIE is very helpful in determining whether a problem lies with the maser or with the PFS.

Figure 7 shows a plot of the fractional frequency offset of ATIES relative to seven of the world's best primary frequency standards (the same seven as in Fig. 1). Here a frequency offset (481.6×10^{-15}) and linear frequency drift ($-0.54 \times 10^{-17}/\text{day}$) that were obtained from a linear fit of TP171 to all of the available fountain data (LPTF-FO1, NIST-F1 and PTB-CSF1) have been removed. Therefore, by definition, the fountain data all lie scattered about the horizontal axis. The linear fit was done under the assumption that the fountain data provides the best estimate of the SI second. The fact that there is no obvious systematic deviation of the fountain data from the zero line (except possibly at the very end) indicates that the drift in TP171 has been very close to linear. The dashed lines are linear fits of other standards to ATIES. The error bars show representative uncertainties for some of the runs of the thermal beam standards. Individual uncertainties are shown for all of the fountain runs.

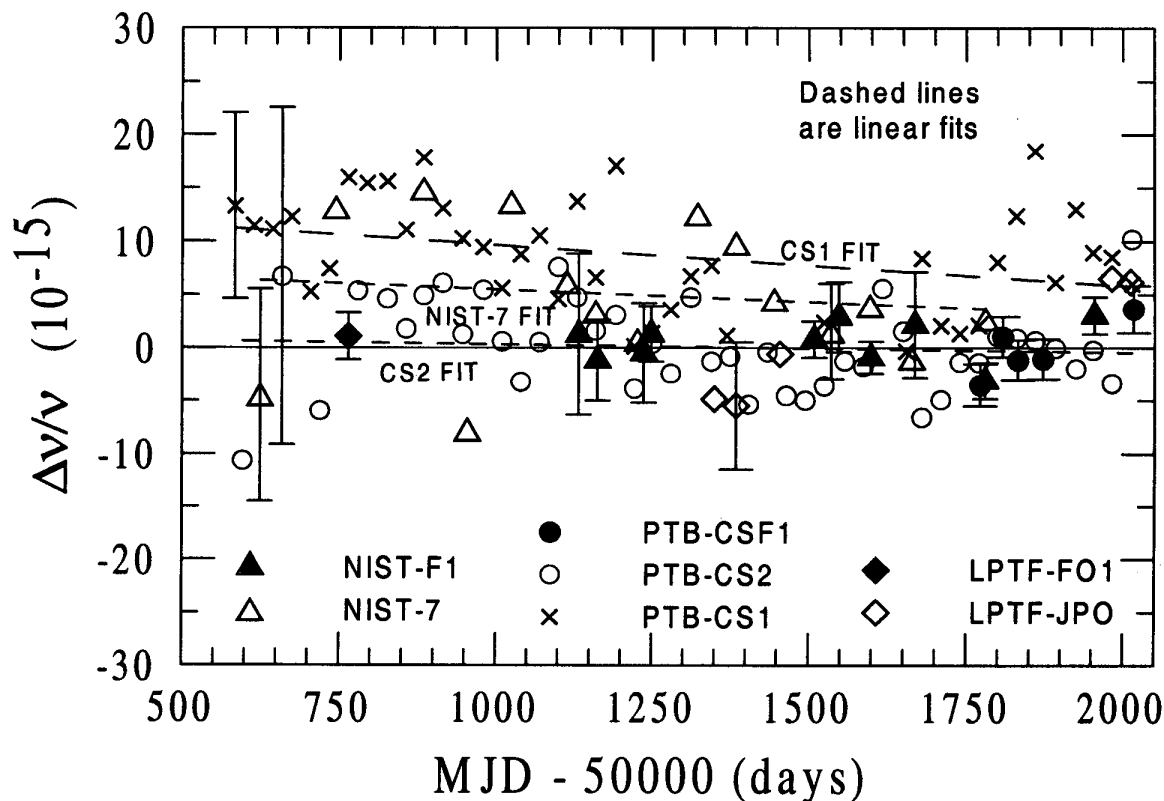


Figure 7. Fractional frequency of offset of ATIES relative to seven primary frequency standards.

Figure 7 provides a convenient means of comparing the seven frequency standards (a similar comparison could also be made using TAI as the reference, but with increased scatter in the data). A very interesting observation is that PTB-CS2 has performed much better than its stated uncertainty of 1.2×10^{-14} . Careful inspection of the figure shows that the linear fit of the PTB-CS2 data (dashed line) is nearly indistinguishable from the zero line and thus is in excellent overall agreement with the fountain data. The stated uncertainty of PTB-CS1 is smaller than that of PTB-CS2, but its performance has not been as good as that of CS2, as indicated by the CS1 fit line. Overall NIST-7 has also done well, even though, as mentioned earlier, it exhibited some relatively large instabilities prior to about MJD 51100. Its average offset from the fountains is approximately 5×10^{-15} . The LPTF standards all appear to be consistent within their stated uncertainties. Data such as those in Fig. 7 indicate that a clear picture of how well a standard is actually performing may not be available until the next generation of standards is operational.

The combination of ATIES and PTB-CS2 serves as a frequency transfer standard with excellent short and long-term stability. This provides a means of comparing all three fountains in a somewhat qualitative sense and they all appear to be in good agreement. A more precise and detailed comparison of NIST-F1 and PTB-CSF1 is presented in [1].

Summary

A stable frequency reference such as ATIE allows for significant dead time or time offsets in the operation and performance assessment of a primary frequency standard. Dead times of up to 50 % in a 30-day evaluation run can be present with only a negligible impact on the uncertainty of the evaluation. Run-to-run comparisons or comparisons between different standards can also be made with uncertainties in the low 10^{-15} range even though the runs may be separated by hundreds of days. Such a capability has proven to be very useful in the development and operation of the NIST primary frequency standards, as well as in comparisons with standards in other laboratories.

References

- 1 T. Parker, P. Hetzel, S. Jefferts, S. Weyers, L. Nelson, A. Bauch, J. Levine, "First Comparison of Remote Cesium Fountains," in these proceedings.
- 2 T. E. Parker and J. Levine, "Impact of New High Stability Frequency Standards on the Performance of the NIST AT1 Time Scale," in *IEEE Trans. on Ultrason., Ferroelect., and Freq. Contr.*, vol. 44, no. 6, pp. 1239-1244, 1997.
- 3 T. E. Parker, "Hydrogen Maser Ensemble Performance and Characterization of Frequency Standards," in *Proc. 1999 Joint Meeting of the European Freq. and Time Forum and the IEEE International Freq. Control Symp.*, pp. 173-176, 1999.
- 4 T. E. Parker, D. A. Howe and M. Weiss, "Accurate Frequency Comparisons at the 1×10^{-15} Level," in *Proc. 1998 IEEE International Freq. Control Symp.*, pp. 265-272, 1998.
- 5 R. J. Douglas and J. S. Boulanger, "Standard Uncertainty for Average Frequency Traceability," in *Proc. 11th European Freq. and Time Forum*, pp. 345-349, 1997.
- 6 J. H. Shirley, W. D. Lee and R. E. Drullinger, "The Accuracy Evaluation of NIST-7," to be published in *Metrologia*.