EFTF 2008

A Comparison of Cesium Fountain Primary Frequency Standards*

Thomas E. Parker
National Institute of Standards and Technology
Time and Frequency Division
325 Broadway
Boulder, CO 80305 USA
tparker@boulder.nist.gov

ABSTRACT

There are currently nine cesium fountain primary frequency standards reporting calibrations of TAI to the Bureau International des Poids et Measures (BIPM). An investigation has been carried out using data from the BIPM publication Circular T to evaluate the frequency offsets among these standards and to determine whether these offsets are consistent with the stated uncertainties. The results of this investigation show that, overall, the standards agree very well with each other.

INTRODUCTION

The first formal evaluation of a cesium fountain primary frequency standard (PFS) was reported to the Bureau International des Poids et Measures (BIPM) by the Observatoire de Paris in September 1995. However, only since November 1999 have laboratories with Cs fountain primary frequency standards been regularly reporting evaluation results to the BIPM. Currently there are nine fountains reporting to the BIPM, and since 1999 there have been 124 formal reports published in Circular T. The body of data is now large enough that a meaningful comparison can be made among the standards. There have been a few direct fountain comparisons reported in the literature [1-4], but the Circular T data now provide an overall better comparison.

In this study the comparison is made by using individual pairs of reports in Circular T that occurred close together in time. Since the report periods did not occur over exactly the same intervals (dead time was present) the stability of the reference flywheel must be taken into account in calculating the comparison uncertainty. Two different, independent, flywheel frequency references are used. One is International Atomic Time (TAI) and the other is the internal, post processed, maser based time scale, AT1E, at the National Institute of Standards and Technology (NIST). In addition to dead time uncertainties, the uncertainty introduced by frequency transfer must also be included. For any two standards a number of data pairs are available over time and these can be averaged to give an overall fractional frequency difference and a total uncertainty of comparison.

FOUNTAINS REPORTING TO THE BIPM

LPTF-FO1 (now SYRTE-FO1) was the first Cs fountain primary frequency standard to report to the BIPM in September 1995 with a fountain uncertainty of 3×10^{-15}. There was also about an equivalent amount of frequency transfer uncertainty. Twelve reports were submitted by the Laboratoire Primaire du Temps et des Frequences (now Bureau National de Métrologie Systèmes de Référence Temps Espace (BNM-SYRTE)) in France from September 1995 to November 1997, then there were no additional reports from this standard until November 2006. NIST-F1 from the National Institute of Standards and Technology (NIST) in the USA started reporting in November 1999 with a fountain uncertainty of 1.8×10^{-15}, and a frequency transfer uncertainty of 1.5×10^{-15}. NIST-F1 is still reporting regularly. In August 2000 PTB-CSF1 from Physikalisch Technische Bundesanstalt (PTB) made its first report. From this date on there have been at least two fountains reporting into Circular T several times per year. SYRTE-FO2, SYRTE-FOM and IT-CsF1 (from Istituto Nazionale di Ricerca Metrologica, INRIM, in Italy) all joined the fountain PFS community near the end of 2002. NPL-CsF1, from the National Physical Laboratory (NPL) in the United Kingdom started in 2004, NMIJ-F1 from the National Metrology Institute of Japan (NMIJ) started in 2005, and NICT-CsF1 from the National Institute of Information and Communication Technology (NICT) in Japan started in 2006. Currently nine fountains from seven laboratories are reporting on a regular basis. Almost every month now there is at least one fountain reporting into Circular T. Data from August 2000 until the present are used in this study. All uncertainties presented in this paper are 1 sigma.

Table 1 gives a list of the nine fountains and typical uncertainties from recent evaluation reports sent to the BIPM. The fountains are listed in order of total combined uncertainty, u. The Type A and B uncertainties, u_A and u_B, are “in laboratory” uncertainties and do not contain dead time or frequency transfer uncertainties. These fractional frequency uncertainties are in units of 10^{-15}. The run length is also given since the Type A (statistical) uncertainty is dependent on the run time. The Type B uncertainties may vary somewhat from run to run, but in general they tend to decrease slowly with time as more is learned about each standard. The last column in the

---

Table 1. Typical recent uncertainties of fountains currently reporting to the BIPM.

#	Fountain	u_A (10^{-15})	u_B (10^{-15})	u (10^{-15})	Run Length (days)	Number of Formal Reports
1	NIST-F1	0.26	0.31	0.40	25	26
2	SYRTE-FO1	0.27	0.42	0.50	30	4*
3	SYRTE-FO2	0.26	0.46	0.53	30	24
4	SYRTE-FOM	0.2	0.7	0.7	25	17
5	IT-CsF1	0.8	0.5	0.9	20	14
6	PTB-CSF1	1.0	1.1	1.5	15	18
7	NPL-CsF1	0.5	1.8	1.9	35	8
8	NICT-CsF1	1.0	1.9	2.1	15	5
9	NMIJ-F1	0.9	3.9	4.0	15	8

* Since year 2006

table gives the number of reports from each standard that have been submitted to the BIPM as of March, 2008. Note for SYRTE-FO1 that this is only since its return in 2006.

As shown in Table 1, the combined uncertainties range over an order of magnitude among the fountains, but even the largest uncertainty is still lower than that of the best thermal beam standard. The 3 fountains with the lowest combined uncertainties are NIST-F1, SYRTE-FO1 and SYRTE-FO2, all with uncertainties in the mid 10^{-16} range. Given that these uncertainties vary a little from run to run, these 3 standards should be considered as essentially equivalent. There have been a total of 124 fountain reports to the BIPM from November, 1999 when NIST-F1 came on line to February, 2008.

Figure 1 shows the fractional frequency (rate) offset of TAI as measured by each fountain PFS since November 1999 (124 data points). The reported uncertainty, including frequency transfer uncertainty, is also shown for each fountain. This data from Circular T is plotted as a function of Modified Julian Date (MJD) and covers a period of over eight years. The long-term variations are in the rate of TAI, but the short-term fluctuations are from both the noise in TAI and variations in the fountain frequencies. There are a few apparent outliers near MJDs 52800 and 53600. Since MJD 54000 there have been so many fountains reporting that it is difficult to resolve individual data points. For this investigation we will be interested in pairs of fountain measurements that occur within 100 days of each other. Using data with time offsets greater than that is not recommended since the dead time uncertainty becomes quite large, and, more seriously, the long-term frequencies of TAI and AT1E are not independent of the fountain frequencies.

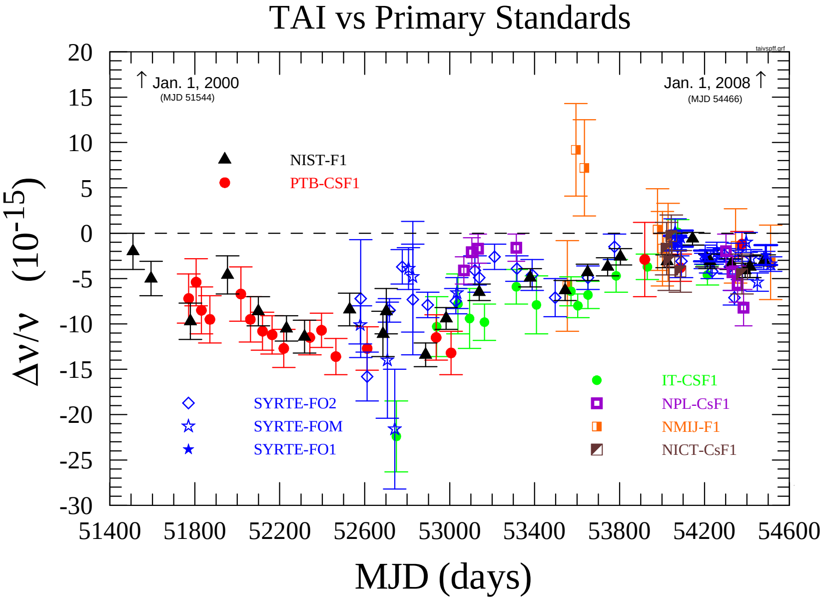

Figure 1. Fractional frequency offset of TAI relative to Cs fountain primary frequency standards as reported in Circular T since November, 1999.

PROCEDURES FOR COMPARING FOUNTAINS

Clearly, it is highly desirable to have direct fountain comparisons in which the start and stop times are coordinated and the frequency transfer techniques are optimized. However, operating fountains on demand has been very difficult, and relatively few such comparisons have been accomplished [1-4]. Another approach is to use the data in available Circular T. Wolf and Petit [5] used an approach in which variations of a TAI type timescale were calculated excluding one fountain at a time. For the investigation reported in this paper a different approach is being used in which pairs of fountain data points from Circular T are compared. This gives a large number of data points but they are not all well aligned in time. Each data point from Circular T gives the rate of TAI relative to a particular PFS. Differencing two data points that are closely spaced in time gives an estimate of the frequency difference between the two standards. However, this introduces a dead time uncertainty that depends on how closely aligned the two runs are and on the frequency stability of the frequency reference [6]. To help reduce this dead time uncertainty, the fountain frequency values are also referenced to a post processed, maser based ensemble at NIST referred to as AT1E [7]. TAI and AT1E are essentially independent time scales. By averaging the results using both time scale references (flywheels) the dead time uncertainty is slightly reduced.

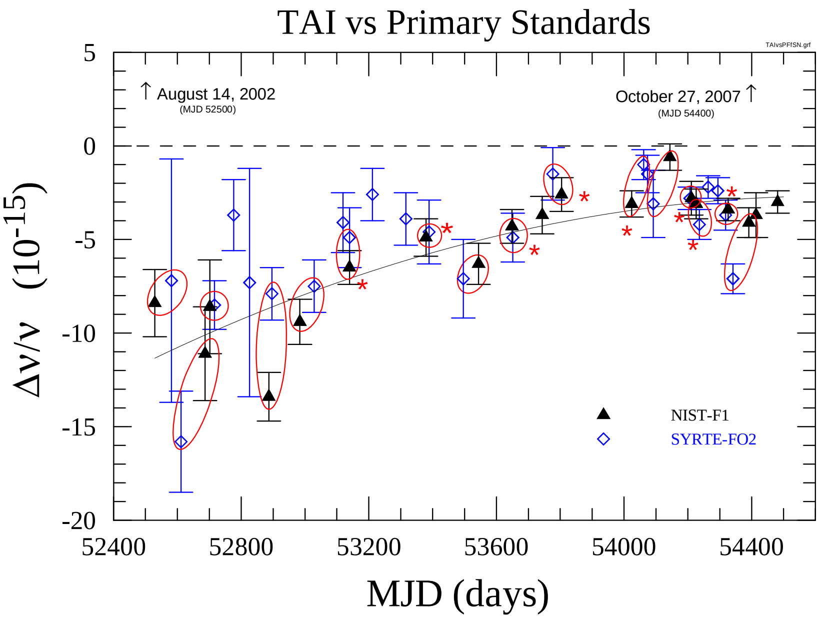

Figure 2. Pairs of data points for NIST-F1 and SYRTE-FO2 used for the frequency comparison. The 8 pairs marked with the red asterisks are used for a subset with a lower uncertainty.

Figure 2 shows a set of Circular T data pairs for NIST-F1 and SYRTE-FO2. The blue diamonds show all the data points of TAI versus SYRTE-FO2 that have appeared in Circular T to date. Also plotted as black triangles are the data points for NIST-F1 that were reported over the same time interval. Each pair of data points used in this analysis is circled in red, and only one data point from each standard was used for each comparison pair. There are 24 SYRTE-FO2 data points and 19 NIST-F1 points over this time interval. From these data we get 16 suitable pairs, with no pair having more than 75 days of center to center time offset. For these 16 points the mean center to center offset (including sign) is -5.2 days, with a mean magnitude of center to center offset of 28.0 days. 56% of the runs had at least some overlap in run time. The Type B uncertainties tend to decrease over time, so a subset of more recent pairs was also selected that are indicated with red asterisks. These pairs also have a lower frequency transfer uncertainty since transfer technology has been slowly improving. In addition a tighter requirement on overlap was used such that the mean center to center offset (including sign) is -2.6 days, with a mean magnitude of center to center offset of 12.5 days. 87% of the runs had at least some overlap in run time. This smaller subset has a slightly lower comparison uncertainty. The same procedure as that shown in Figure 2 was also used with AT1E as the frequency reference.

The solid black line in Fig. 2 is a second order fit line to the NIST-F1 data. Its only function is to illustrate the long-term frequency drift of TAI. It does not represent the short-term stability of TAI.

The procedure used for combining uncertainties is as follows. All Type A uncertainties are treated as uncorrelated and are therefore added in quadrature when comparing two standards. Over time they will average down. The Type B uncertainties are treated as uncorrelated between standards, but correlated over time. Neither of these statements regarding Type B uncertainties is strictly true, but for this study we will assume that they are reasonable assumptions. A rigorous process of combining the Type B uncertainties would require a detailed analysis of the Type B biases and uncertainties for each standard, and how these biases vary as a function of time. Such an analysis is beyond the scope of this study. Therefore, the Type B uncertainties of different standards will be combined in quadrature for a pair, but over time a weighted average of the Type B uncertainties will be used. Thus the Type B uncertainty will not average down and will never be smaller than the smallest individual Type B uncertainty of a single pair.

Table 2 shows the details of how a pair of data points is handled. This example is for a NIST-F1 and SYRTE-FO2 pair. The first column shows the start and stop dates (in MJD) for the runs for each standard. The second column gives the midpoints of each run and the third column is the duration in days. The overlap in days is given in the last column of the middle (offset) row. As is typical, these two runs were made with no knowledge that the other standard was being operated. Columns 4 through 8 in the top and bottom rows are respectively, (4) the fractional frequency difference between TAI and the particular fountain, (5) the reported Type A uncertainty, u_A, of the fountain, (6) the Type B uncertainty, u_B, of the fountain, (7) the uncertainty in the link between the PFS and the local clock contributing to TAI, u_l, usually

Table 2. Example of a fountain comparison data pair. Fractional frequencies in units of 10^{-15}.

NIST-F1

MJD	MJD	Duration	y(TAI-F1)	u_A	u_B	u_l	u_{TAI}
54314-54339	54326.5	25d	-3.4	0.3	0.3	0.1	0.4

u_{dead}	y(FO2-F1)	u_C	u_{CA}	u_{CB}	overlap
0.7	0.3	1.2	1.05	0.58	15d

SYRTE-FO2

MJD	MJD	Duration	y(TAI-FO2)	u_A	u_B	u_l	u_{TAI}
54309-54329	54319	20d	-3.7	0.3	0.5	0.1	0.5

dominated by fountain dead time, and (8) the frequency transfer uncertainty in the link to TAI, u_{TAI}.

The middle (offset) row contains the result of the frequency difference calculation. Here u_{dead} is a dead time uncertainty introduced by the run misalignment and the noise of the frequency reference, in this case TAI [6]. If the start and stop times of each fountain run were exactly the same this uncertainty would go to zero. Columns 2 through 5 in this middle row are respectively, (2) the calculated fractional frequency difference between SYRTE-FO2 and NIST-F1, (3) the total combined uncertainty of the comparison, u_C, (4) the Type A uncertainty of the comparison, u_{CA}, and (5) the Type B uncertainty of the comparison, u_{CB}.

Note that the two u_{TAI}s and u_{dead} are large compared to the other uncertainties. u_A, u_l, u_{TAI} (top and bottom rows) and u_{dead} are all Type A uncertainties and are combined in quadrature to give the u_{CA} of the comparison in the middle row. The two individual Type B uncertainties in the top and bottom rows are combined in quadrature to give the comparison Type B uncertainty, u_{CB}. u_{CA} and u_{CB} in the middle row are combined in quadrature to give the total uncertainty of the comparison, u_C.

For NIST-F1 and SYRTE-FO2 there are 16 data pairs like that in Table 2. The average fractional frequency offset for all 16 is determined by calculating the weighted average of y(FO2-F1) using u_C to calculate the weights. The Type A uncertainty for all 16 pairs averages down as

$$\frac{1}{u_A^2} = \sum_{i=1}^{n} \frac{1}{u_{Ai}^2} \quad . \tag{1}$$

The Type B uncertainty for the 16 data pair comparison is calculated as the weighted average of the individual Type B uncertainties using u_C to calculate the weights.

Using data in Circular T and internal data at NIST, the value y(TAI-FO2) can be transformed to y(AT1E-FO2). With these values the same procedures used for TAI can be applied to AT1E to give another set of data for y(FO2-F1), but by using AT1E as a flywheel. The data sets from TAI and AT1E are then averaged to obtain the final results, which are discussed in the next section. Averaging the TAI and AT1E results reduces the u_{dead} contribution by about a factor of $1/\sqrt{2}$ and this gives a modest reduction in comparison uncertainty if the other Type A and Type B uncertainties are not significantly larger than u_{dead}.

COMPARISON RESULTS

Table 3 shows the results of the comparisons of NIST-F1, SYRTE-FO1 and SYRTE-FO2, the three standards with the lowest total individual uncertainties (see Table 1). Column 1 lists the two fountains being compared and Column 8 shows the number of pairs averaged. Columns 2, through 5 are respectively, (2) the average fractional frequency offset, y_{avg}, (3) the total comparison uncertainty, U_C, (4) the Type A comparison uncertainty, U_{CA}, and (5) the Type B comparison uncertainty, U_{CB}. All are in units of 10^{-15}. U_{CA} and U_{CB} are added in quadrature to get U_C. Columns 6 and 7 are Birge ratios [8] as defined in Eq. 2.

$$R_B = \sqrt{\frac{\sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{u_i^2}}{(n-1)}} \tag{2}$$

The term $(y_i - \bar{y})$ is the deviation of each data pair from the mean, and u_i is either u_{CAi} for R_{BA} or u_{Ci}, for R_{BC}. n is the number of data points. The Birge ratio can be used as a measure of whether the actual scatter about the mean in a data set is consistent with the uncertainty associated with each point. If the uncertainties are correct, R_B will be close to 1. If the uncertainties are overstated, R_B will be less than 1, and if they are understated, R_B will be greater than 1. In a PFS the scatter in the data may come from just the Type A frequency fluctuations, or from both Type A and B offsets. If the Type B biases are constant and don't change over time, the scatter about the mean in the data will be consistent with just the Type A

Table 3. Comparisons of NIST-F1, SYRTE-FO1 and SYRTE-FO2. Fractional frequencies in units of 10^{-15}.

Fountains	y_{avg}	U_C	U_{CA}	U_{CB}	R_{BA}	R_{BC}	# pairs
SYRTE FO2 vs NIST-F1	-0.19	0.75	0.35	0.66	1.10	1.02	16
SYRTE FO2 vs NIST-F1*	-0.16	0.73	0.36	0.64	1.00	0.86	8
SYRTE FO1 vs SYRTE FO2	-0.10	0.61	0.24	0.57	0.76	0.55	3
SYRTE FO1 vs NIST-F1	-0.37	0.85	0.69	0.51	0.65	0.62	4

* Selected data

uncertainty. If both Type A and B fluctuations contribute to the scatter of the data, then the Birge ratio is best calculated using the total uncertainty. Therefore Table 3 lists R_B calculated with both u_A and u_C. R_{BA} will always be larger than R_{BC}. A reasonable scenario is for R_{BA} to be a little larger than 1 and R_{BC} to be somewhat smaller than 1. There is cause for concern if R_{BC} is found to be significantly larger than 1, or if R_{BA} is significantly smaller than 1.

The first row in Table 3 shows the results for the 16 data point set for y(FO2-F1). As can be seen the two fountains are in excellent agreement, with an average frequency offset of only -0.19×10^{-15} and a total comparison uncertainty of 0.75×10^{-15}. The average frequency offset is smaller than the Type A uncertainty, U_{CA}, of 0.35×10^{-15}, so the offset is not statistically significant. The offset is also well within the Type B uncertainty U_{CB}. Both Birge ratios are close to 1 indicating consistent data. The second row (fountain names in red) shows the result for the selected set of 8 pairs discussed earlier. By using the smaller data set (red asterisks in Fig. 2) a slightly smaller total uncertainty is obtained because of the smaller u_B. Again NIST-F1 and SYRTE-FO2 show excellent agreement, with an average frequency offset of only -0.16×10^{-15} and a comparison uncertainty of 0.73×10^{-15}. The Birge ratios are slightly smaller, but not in a statistically significant way. The uncertainty of 0.73×10^{-15} is almost as good as a single, well coordinated 30 day direct comparison with no dead time uncertainty and optimized frequency transfer uncertainty, which would give a comparison uncertainty of about 0.68×10^{-15}. With the current fountains, even if the total Type A uncertainty could be reduced to a negligible level, the comparison uncertainty would be no smaller than about 0.55×10^{-15} due to the Type B uncertainties.

The third row shows the results of a comparison of SYRTE-FO1 with SYRTE-FO2 using 3 data pairs from Circular T. Obviously this comparison can be made in a better fashion within SYRTE, but one of the three pairs had exactly the same start and stop times. In this situation u_{dead} goes to zero and the u_{TAI}'s cancel out. Except for possible internal dead time in u_l, this one comparison is essentially an in-house comparison. The runs for the other 2 pairs were not perfectly synchronized and hence had larger uncertainties. In any case, the agreement between SYRTE-FO1 and SYRTE-FO2 is excellent, with an average frequency offset of only -0.10×10^{-15} and a comparison uncertainty of 0.61×10^{-15}. The Birge ratios are smaller than 1, but with only 3 points this is not statistically significant. Obviously, a well conducted direct in-house comparisons of SYRTE-FO1 and SYRTE-FO2 should give more reliable results than this data.

The forth row shows the comparison of SYRTE-FO1 with NIST-F1. Since there are only 4 pairs, and the fountain runs are not aligned, the statistical uncertainty is large relative to the SYRTE-FO1 vs SYRTE-FO2 comparison. However, again the average frequency difference is small compared to the comparison uncertainties. With only 4 points the small Birge ratios are not significant.

Since NIST-F1 and SYRTE-FO2 are in excellent agreement with each other, and have low uncertainties, both will be used as standards of comparison for the other fountains. Thus Table 4 shows the results of comparisons of SYRTE-FOM, IT-CsF1, PTB-CSF1, NPL-CsF1, NICT-CsF1 and NMIJ-F1 with either NIST-F1 or SYRTE-FO2. (As more data from SYRTE-FO1 becomes available, it can also be used as a comparison standard.) The procedures used to obtain this data were the same as those used to compare the standards in Table 3. The only difference in Table 4 is that the last column shows the number of pairs used along with the total number of possible data points for the particular fountain.

The first row in Table 4 shows the frequency offset of SYRTE-FOM. Better than 80% of the available runs could be used, with the comparison standards being equally divided between NIST-F1 and SYRTE-FO2. The analysis shows that the average frequency offset of SYRTE-FOM is within the stated total uncertainty, but that it does have a bias larger than the Type A comparison uncertainty. Thus the standard has a statistically significant bias, but it is still within the Type B uncertainty. Therefore, it is performing within its stated uncertainty. At 1.37, the Birge ratio R_{BA} is larger than 1 in a statistically significant manner, but R_{BC} is much closer to 1 at 1.04. This indicates that some of the variations in the fountain frequency probably come from variations in one or more of the biases responsible for the Type B uncertainty. This is

Table 4. Comparisons of SYRTE-FOM, IT-CsF1. PTB-CSF1, NPL-CsF1, NICT-CsF1 and NMIJ-F1 with either NIST-F1 or SYRTE-FO2. Fractional frequencies in units of 10^{-15}.

Fountains	y_{avg}	U_C	U_{CA}	U_{CB}	R_{BA}	R_{BC}	# pairs
SYRTE FOM vs NIST-F1 or SYRTE FO2	-0.96	1.09	0.28	1.05	1.37	1.04	14 of 17
SYRTE FOM vs NIST-F1 or SYRTE FO2*	-0.85	1.09	0.33	1.04	1.42	0.99	8 of 11
IT-CsF1 vs NIST-F1 or SYRTE FO2	+1.91	1.10	0.64	0.89	1.24	1.08	12 of 14
IT-CsF1 vs NIST-F1 or SYRTE FO2 @	+1.58	1.07	0.65	0.85	0.82	0.73	11 of 13
PTB-CSF1 vs NIST-F1 or SYRTE FO2	+0.30	1.50	0.78	1.28	1.13	0.98	10 of 18
NPL-CsF1 vs NIST-F1 or SYRTE FO2	-0.51	2.03	0.51	1.95	1.56	0.97	8 of 8
NICT-CsF1 vs NIST-F1 or SYRTE FO2	-0.51	2.20	0.98	1.97	0.72	0.51	4 of 5
NMIJ-F1 vs NIST-F1 or SYRTE FO2	-3.56	4.10	1.10	3.95	1.44	0.97	5 of 8

* After FOM rebuilt. @ One outlier removed.

entirely reasonable in standards with relatively large Type B uncertainties. SYRTE-FOM has recently been rebuilt so the second row (fountain name in red) shows the results for 8 data points acquired since the fountain came back on line. Five of the eight points are compared against NIST-F1. The results are nearly the same as in row 1, but with a slightly smaller offset. Note that the Type B uncertainty for SYRTE-FOM shown in Table 1 is from its most recent report (February 2008), and that in prior reports the Type B uncertainty was 0.9×10^{-15}. Thus the data in Table 4 reflect a larger Type B uncertainty for this standard than that shown in Table 1.

The third row shows the results for IT-CsF1 using 12 of 14 data points, again equally divided between NIST-F1 and SYRTE-FO2. This fountain exhibits a systematic frequency offset larger than the combined comparison uncertainty. It is a statistically significant offset that is larger than the Type B uncertainty. The Birge ratios are both larger than 1. Even if the first report from this fountain, which looks like an outlier near MJD 52750, is deleted, a frequency bias larger than the uncertainty is still present, as shown in row 4. However, as expected the Birge, ratios are reduced. Though they are both smaller than 1, the deviation is not large enough to be of concern with only 11 points.

The results for PTB-CSF1 are shown in row 5 with 10 out of 18 possible data points. 9 of the 10 data points are comparisons with NIST-F1 since most of the PTB-CSF1 data is prior to MJD 53014. Since that date there have been only 3 new reports, with 1 aligning best with SYRTE-FO2 and 1 with NIST-F1. Thus most of this data is relatively old. However, the results in Table 4 show that this standard is behaving in a manner consistent with its stated uncertainties. R_{BA} and R_{BC} both have reasonable values.

Row 6 shows the results for NPL-CsF1. All 8 of the reports from this standard were used, though only one aligned best with NIST-F1. These 8 data points appear in two groups, with 4 data points occurring between MJD 53049 and 53329, and the remaining 4 between MJD 54284 and 54399. The frequency offset of this standard is on the edge of being statically significant, though it is well within its Type B uncertainty. A large R_{BA} value indicates that Type B biases are likely fluctuating.

The results for NICT-CsF1, the newest PFS, are shown in row 7. It is behaving in a completely consistent manner though there are only a small number of points (equally divided between NIST-F1 and SYRTE-FO2). The Birge ratios are low, but with only 4 points this is not significant. Finally, row 8 shows the results for NMIJ-F1 with 5 out of 8 points (4 using NIST-F1). This standard exhibits the largest offset, but it is within its uncertainty. R_{BA} is again relatively large which indicates fluctuations in a Type B bias.

As a final step in the comparison process all of the data in Table 4 (excluding rows 1 and 3) can be averaged to determine how these fountains compare as a group with SYRTE-FO2 and NIST-F1. The results are

$$y_{wtdavg} = 0.13 \times 10^{-15} \quad U_{Cavg} = 0.61 \times 10^{-15} \quad R_B = 0.85,$$

where y_{wtdavg} is the weighted average of the offsets, U_{Cavg} is the combined uncertainty and R_B is the Birge ratio (using U_C for the individual uncertainties). The combined uncertainties of the individual comparisons were treated as being uncorrelated. Thus we see that these 6 fountains as a group agree with NIST-F1 and STRTE-FO2 at a level well within the comparison uncertainty. This small frequency offset is consistent with the assumption that the Type B biases among the standards are largely uncorrelated. The Birge ratio also shows the data to be reasonably consistent with the stated uncertainties, though it is a little on the low side.

Using rows 1 and 3 in place of rows 2 and 4 makes only a small difference, as shown below.

$y_{wtdavg} = 0.18x10^{-15}$ $U_{Cavg} = 0.62x10^{-15}$ $R_B = 0.96$

Of the 9 standards, only one exhibits a frequency offset more than 1 sigma. Statistically, if all uncertainties (both Type A and Type B) were 1 sigma, one would expect 2 to 4 standards to exceed 1 sigma. The Type A uncertainties are virtually all 1 sigma, but it is questionable to assume that all the Type B uncertainties are. Many are 1 sigma, but in situations where it is difficult to rigorously determine the uncertainty of a Type B bias, the natural tendency is to be conservative. Thus some Type B uncertainties may in fact be closer to 2 or 3 sigma. The fact that only 1 of 9 standards has an average frequency offset from the mean that is more than 1 sigma suggests that some of the Type B uncertainties may be slightly over estimated. The R_B values above of 0.85 and 0.96 are consistent with this observation. However, the data set is small and more data will be needed before a firm conclusion can be drawn.

A weakness in using Circular T data is that there are significant contributions to the Type A uncertainties from dead time and frequency transfer (see Table 2). Thus many points need to be averaged to get a low uncertainty. This makes it difficult to resolve changes in the frequency offset of a PFS as a function of time. Such information would be very useful and would best be observed by a series of well coordinated fountain comparisons with no dead time and optimized frequency transfer.

SUMMARY

With 9 fountains from 7 laboratories reporting on a more or less regular basis, the uncertainty in TAI is now at an unprecedented low level near $0.5x10^{-15}$. The three Cs fountain primary frequency standards with the lowest uncertainties, NIST-F1, SYRTE-FO1 and SYRTE-FO2, all agree at a level below $0.4x10^{-15}$, which is well within the comparison uncertainty of $0.7x10^{-15}$. The best comparison in fountains from different laboratories is between NIST-F1 and SYRTE-FO2, which shows a fractional frequency offset of $-0.16x10^{-15}$. The six standards SYRTE-FOM, IT-CsF1, PTB-CSF1, NPL-CsF1, NICT-CsF1 and NMIJ-F1, when compared as a group to either NIST-F1 or SYRTE-FO2, have an average frequency offset of less than $0.2x10^{-15}$. Also, the Birge ratios do not show any significant inconsistencies in the stated uncertainties. Only one of the 9 standards has an average offset that is larger than 1 sigma.

Based on the data presented in this paper one has to conclude that the community of Cs fountain primary frequency standards is in a very healthy state.

ACKNOWLEDGEMENTS

Very useful discussions with Steve Jefferts, Tom Heavner and David Smith are gratefully acknowledged.

REFERENCES

[1] 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.

[2] J.Y Richard, P. Uhrich, A. Clairon, J. Achkar, S. Bize, H. Marion, P. Laurent, M. Abgrall, I. Maksimovc, J. Grunert, A. Bauch, S. Weyers, and R. Wynands, “Comparison of Remote Cesium Fountains using GPS P3 and TWSTFT Links,” in *Proc. of the 18th European Frequency and Time Forum*, 2004.

[3] A. Bauch, J. Achkar, S. Bize, D. Calonico, R. Dach, R. Hlavac, L. Lorini, T. Parker, G. Petit, D. Piester, K. Szymaniec, and P. Uhrich, “Comparison between frequency standards in Europe and the USA at the 10^{-15} uncertainty level,” *Metrologia*, **43**, pp. 109-120, 2006.

[4] M. Fujieda, T. Gotoh, D. Piester, M. Kumagai, S. Weyers, A. Bauch, R. Wynands,M. Hosokawa, “First Comparison of Primary Frequency Standards Between Europe and Asia,” to be published in the *Proc. of the Joint Meeting of the European Frequency and Time Forum and the IEEE International Frequency Control Symposium*, 2007.

[5] P. Wolf and G. Petit, “Comparing high accuracy frequency standards via TAI,” in *Proc. of the 20th European Frequency and Time Forum*, pp. 476-485, 2006.

[6] 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.

[7] 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.

[8] R. Kacker, R. Datla and A. Parr, “Combined results and associated uncertainty from interlaboratoy evaluations based on the ISO Guide,” *Metrologia*, **39**, pp. 279-293, 2002.