Comparison between frequency standards in Europe and the USA at the 10^{-15} uncertainty level

A. Bauch[(1)], J. Achkar[(2)], S. Bize[(2)], D. Calonico[(3)], R. Dach[(4)], R. Hlavać[(5)], L. Lorini[(3)], T. Parker[(6)], G. Petit[(7)], D. Piester[(1)], K. Szymaniec[(5)], and P. Uhrich[(2)]

(1) Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany

(2) Laboratoire National de Métrologie et d'Essais – Observatoire de Paris / Systèmes de Référence Temps Espace (OP), Paris, France

(3) Istituto Elettrotecnico Nazionale Galileo Ferraris (IEN), Torino, Italy

(4) Astronomical Institute, University of Bern (AIUB), Switzerland

(5) National Physical Laboratory (NPL), Teddington, UK

(6) National Institute of Standards and Technology (NIST), Boulder, Colorado, USA

(7) Bureau International des Poids et Mesures (BIPM), Sèvres, France

ABSTRACT

Istituto Elettrotecnico Nazionale Galileo Ferraris (IEN), National Institute of Standards and Technology (NIST), National Physical Laboratory (NPL), Laboratoire National de Métrologie et d'Essais – Observatoire de Paris / Systèmes de Référence Temps Espace (OP), and Physikalisch-Technische Bundesanstalt (PTB) operate cold-atom based primary frequency standards which are capable to realize the SI-second with a relative uncertainty of 1×10^{-15} or even below. These institutes performed an intense comparison campaign of selected frequency references maintained in their laboratories during about 25 days in October/November 2004. Active hydrogen maser reference standards served as frequency references for the institutes' fountain frequency standards. Three techniques of frequency (and time) comparisons were employed. Two-way satellite time and frequency transfer (TWSTFT) was performed

in an intensified measurement schedule of 12 equally spaced measurements per day. The data of dual-frequency geodetic GPS receivers were processed to yield an ionosphere-free linear combination of the code observations from both GPS frequencies, typically referred to as GPS TAI P3 analysis. Last but not least, the same GPS raw data were separately processed allowing GPS carrier-phase based (GPS CP) frequency comparisons to be made. These showed the lowest relative frequency instability at short averaging times of all the methods. The instability was at the level of 1 part in 10^{15} at one-day averaging time using TWSTFT and GPS CP. The GPS TAI P3 analysis is capable of giving a similar quality of data after averaging over two days or longer. All techniques provided the same mean frequency difference between the standards involved within the 1-σ measurement uncertainty of few parts in 10^{16}. The frequency differences between the three fountains of IEN (IEN-CsF1), NPL (NPL-CsF1), and OP (OP-FO2) were evaluated. Differences lower than the 1-σ measurement uncertainty were observed between NPL and OP, whereas the IEN fountain deviated by about 2 σ from the other two fountains.

1. Introduction

The study of time and frequency transfer is an important sector of time and frequency metrology in general since it is essential for the wide application of state-of-the-art frequency standards. In recent years, research into primary frequency standards has led to several devices whose uncertainty to realize the SI second is at the level of one part in 10^{15} or below and whose frequency instability is low enough to verify this accuracy during averaging times well below one day [1 - 6]. Data from these primary frequency standards and from other highly accurate clocks have been regularly used by the Bureau International des Poids et Mesures (BIPM) as an input for the realization of International Atomic Time (TAI) and of Coordinated Universal Time

(UTC) [7]. Two different methods of time transfer have been traditionally used by BIPM for sustaining the network of participating laboratories, Global Positioning System (GPS) common view (CV) time transfer and two-way satellite time and frequency transfer (TWSTFT) via geostationary satellites [8, 9]. Classical GPS coarse acquisition (C/A) code analysis is not discussed in this article since it does not provide the required measurement resolution. A step forward has been the use of geodetic GPS receivers providing dual-frequency code observables which allow the so-called GPS TAI P3 analysis to be made [10, 11]. This technique was used – in parallel with TWSTFT – to compare the caesium fountain frequency standards of OP and PTB during 10 days in 2003 [12].

The additional use of GPS carrier-phase observables (GPS CP) for frequency transfer was proposed for similar purposes [13-15]. It was previously employed in parallel with TWSTFT to compare the fountain clocks CSF1 of the Physikalisch-Technische Bundesanstalt (PTB), Germany, and F1 of the National Institute of Standards and Technology (NIST), USA [16]. Later in this paper we will compare the recent findings with some of the previously obtained ones. The products provided by the International GNSS service (IGS)[1] rely also on the processing of all kinds of GPS observables. IGS clock products could in principle also be used directly for the determination of the properties of the frequency sources at IGS sites [15, 17, 18], but this is beyond the scope of this paper.

Three institutes, the UK National Physical Laboratory (NPL), the Italian Istituto Elettrotecnico Nazionale Galileo Ferraris (IEN), and the French Laboratoire National de Métrologie et d'Essais – Observatoire de Paris / Systèmes de Référence Temps Espace (abbreviated as OP throughout the paper), agreed with NIST and PTB on a

[1] In March 2005 the „International GPS Service“ (IGS) renamed himself as "International GNSS Service" to denote that the activity of the IGS has been and will be further extended from GPS to other Global Navigation Satellite Systems (GNSS), e. g., to GLONASS and Galileo.

campaign of comparisons among their fountain type primary frequency standards, to be effected during 20 to 25 days in October and November 2004, at the Modified Julian Days (MJD) between 53304 and 53329. The five institutes are part of a regularly operating TWSTFT network, and for the duration of the campaign an intensified measurement schedule was agreed upon. All institutes operate geodetic GPS receivers and their stations are acknowledged as IGS sites. The institutes obtained the support of the Astronomical Institute of the University of Bern (AIUB), Switzerland, which processed the GPS data with an innovative software, providing long-term monotonous series of clock solutions for the stations involved with extremely high resolution, based entirely on the GPS CP [19], which represents an extension of the well-known Bernese GPS software [20]. Last not least, the BIPM provided the GPS TAI P3 analysis of the GPS pseudo-range code data [10]. All GPS raw data were obtained from the same geodetic GPS receivers.

This paper is an extended version of previous conference reports [21, 22]. Again a selection of results to be presented had to be made, and we aimed at showing the achievable performance as well as pointing to some observations that were not yet understood. During the campaign it turned out that the fountains of NIST and PTB were not ready to provide data so that fountain comparisons were only carried out between the IEN-CsF1, the NPL-CsF1, and the OP-FO2. The fountain comparison results are reported in Section 5. In the following sections we will detail the equipment and time transfer techniques involved (Section 2) before presenting results for each technique (Section 3) and for inter-comparisons among techniques (Section 4).

2. Description of hardware and time transfer techniques

2.1 Station equipment

In Table 1, we enlist the equipment involved in the participating stations. In all cases, the active hydrogen masers (HM) served as the fly-wheel oscillators for comparisons with the fountain frequency standards, and the maser signals were connected, either directly or through a transfer clock, to the time transfer equipment. The most straightforward configuration, depicted in Figure 1, applied to NIST, NPL, and OP, indicating the three different time transfer techniques connecting the masers at the remote sites. In the cases of NIST and NPL, the maser output represented at the same time the local UTC scale, UTC(NIST) [23], and UTC(NPL), respectively. In the cases of IEN and PTB, however, the GPS receivers were connected to UTC(IEN) and UTC(PTB), respectively, which are derived from caesium clocks, a commercial model at IEN, and the primary clock CS2 at PTB [24], respectively. As a result, further steps in the data analysis were required, because all techniques should be referenced to the same frequency standard. The time differences UTC(PTB) – HM(PTB) were available only on an hourly basis and this dictated the minimum averaging time for GPS CP and GPS TAI P3 data presented below. At IEN the corresponding measurements were taken more frequently so that a more complete data analysis could be performed. This has, however, not been fully included in this paper.

In the sections dealing with results the respective links have been conveniently denoted by the station acronyms, e. g., by PTB-NPL. In the graphs, connecting lines serve the purpose to guide the eye, and the symbols given in brackets have been assigned to the institutes as follows, IEN (∇), NIST (∆), NPL (□), OP (O), and PTB (♦).

2.2 TWSTFT analysis

The stations involved performed TWSTFT co-ordinated by the CCTF Working Group on TWSTFT following standard procedures which shall be briefly described. TWSTFT operation requires scheduled operation of pairs of stations, named A and B for now, which transmit and receive signals simultaneously. The signals consist of characteristic pseudo-random noise (PRN) spread-spectrum signals which are designated by their Mitrex code (see Table 2) and which are synchronous with the local 1 PPS time references $T(\mathrm{Ref_A})$ and $T(\mathrm{Ref_B})$, respectively. They are generated as a phase modulation of the 70 MHz (intermediate) frequency provided by the modem. This signal is up-converted to the RF region (Ku-band, up link at about 14 GHz) and transmitted to a geostationary telecommunication satellite, here the IS 903 of Intelsat Corp., located at 325.5° east. At the satellite, the signal is translated to the downlink frequency (about 12 GHz) and sent back to Earth. The time of arrival (TOA) of the PRN coded signal from station A (B) is determined at station B (A) by cross-correlation of the received and down-converted PRN signal with a local replica of the same PRN code, which is synchronous with the local reference clock. TOA measurements are provided each second during sessions of two minutes per station pair. At the end of a session, a quadratic function is fitted to the data and the TOA midpoint values derived from the fit are exchanged among stations A and B via file transfer through the internet. This allows the quantity of interest, $T(\mathrm{Ref_A}) - T(\mathrm{Ref_B})$ to be calculated, as explained by Kirchner [9] and in an ITU-R Recommendation [25]. During the campaign, the quadratic fit residuals exhibited 1-σ standard deviations of below 400 ps in favourable cases, but going up to almost 1.5 ns for some links. This fact is discussed further below.

For the purpose of an intense comparison, TWSTFT sessions were performed nominally once every two hours. The data sets contained very few gaps. On some links about 315 out of 320 possible data points were recorded, while in a few cases about 20 data points were missing. The TWSTFT measurement schedule including the Mitrex codes assigned to the institutes is reproduced in Table 2. One can recognize that measurements between different pairs of institutes cannot be taken perfectly simultaneously.

2.3 GPS CP analysis

All five participating institutes operate registered IGS stations. The observations from the respective receivers are provided in receiver independent exchange format (RINEX) that contains code and carrier-phase measurements from the two GPS frequencies L1 and L2 for all tracked satellites with a sampling interval of 30 seconds. The GPS CP analysis was performed at AIUB using the Bernese GPS Software, version 5.1 [20]. The solutions are based on the contributions to the IGS of the Analysis Center CODE (Center for Orbit Determination in Europe) hosted at AIUB. CODE is a collaboration between AIUB, the Federal Office of Topography (Swisstopo, Switzerland), the Bundesamt für Kartographie und Geodäsie (BKG, Germany), and the Institut Géographique National (IGN, France). The satellite orbits, Earth orientation parameters, station coordinates, and troposphere delays are introduced from the (geodetic) three-day, double-difference solution calculated for the IGS at CODE. In the clock estimation procedure the receiver and satellite clock parameters are treated as unknowns, along with other parameters (e. g., phase ambiguities). The result is a network solution with a set of consistent receiver and satellite clock corrections for each epoch, every 30 seconds. All receiver clock comparisons (baselines) are extracted from this network solution, hence all loops of

extracted baselines give zero. The GPS CP solution used in the current campaign contains no day boundary discontinuities[2] because the phase ambiguities from the daily data processing are reconnected before estimating the clock parameters. As in addition frequency transfer has been the main interest, the code observations have been discarded.

Such an approach, using only phase observations and neglecting code measurements, has never been made before for a several-weeks long campaign by the AIUB and, as far as we know, has also not been published by any other group. A detailed discussion is given in [19]. Recently, reports of comparisons between frequency standards inside the USA and between NIST and PTB processing the GPS CP data in 1-day batches were published [17, 26]. Some results therein will be mentioned in the relevant sections of this article, but typically the data base available in the earlier studies was scarce compared to that discussed here.

2.4 GPS TAI P3 analysis

The GPS TAI P3 technique has been used since 2002 by the BIPM, with about 15 participating laboratories equipped with geodetic GPS receivers [10, 11]. The dual-frequency P1 and P2 observables delivered by the receivers are linearly combined to form the ionosphere-free P3 observable. Starting with RINEX observation files the data are processed using locally recorded GPS broadcast messages to generate the CGGTTS[3] format used in standard GPS time transfer [10, 27] using locally recorded broadcast GPS parameters. This format dictates in particular that the sampling

[2] Typically, GPS CP data are processed in daily batches, and time differences obtained are not monotonous at day boundaries, see [15, 17] for discussion.
[3]CGGTTS stands for CCTF Working Group on GNSS time transfer standards, CCTF is the Consultative Committee for Time and Frequency, and GNSS stands for Global Navigation Satellite System.

rhythm for the time differences is 16 minutes and that the tracking schedule shifts in time from day to day by four minutes. The CGGTTS data files are gathered by BIPM and used to compute time links after applying different corrections: precise satellite orbit and satellite clock corrections provided by the IGS, and station displacement due to solid Earth tides [11]. The six time links (no data from NIST were available) were computed using the common-view technique. For each 16-minute interval, all available common-view differences were formed and averaged with a weighting scheme based on the satellite elevation, after a first screening for big outliers. BIPM provided unsmoothed and Vondrak smoothed time series. The instability analysis is based on unsmoothed data, all rates presented in the tables below in Sections 4 and 5 are based on smoothed data.

3. Link performance analysis

3.1 General observations

To give an example, data obtained on the link NPL-OP during one day are depicted in Figure 2. The TWSTFT measurements were available once every two hours (full dots, one data point missing). From the GPS CP solutions available every 30 seconds those at full hour epochs were extracted for assessing the achieved instability (open circles in Figure 2). For comparisons between the two techniques (e. g., in Section 4.1) the GPS CP solutions closest to the respective epochs of the TWSTFT measurements (within $\pm$15 seconds) were chosen. In this way the impact of clock instabilities on the comparison of the two frequency transfer methods was minimized. The TAI P3 data which were available at 16-minute intervals (black crosses in Figure 2) are offset by 2 ns for clarity.

A few data points were in general missing for all links, and because of the applied tracking schedule even complete GPS TAI P3 data would not represent a continuous

time series. In all stability plots, however, the minimum τ was chosen as the nominal value irrespective of the actual mean separation in time between measurements as the availability was quite high and the impact of missing data on the plots was considered as minor in the current study.

In Figure 3 the general performance of the hydrogen masers involved is illustrated. Here UTC(NIST) was chosen as the common reference, and GPS CP time differences to all other masers at an hourly separation were evaluated. The residuals to linear fitted functions were plotted. It can be seen that the NPL maser performed very well except for a kink towards the end of the campaign, and the OP maser had a significant non-linear frequency change during the period under study.

3.2 TWSTFT results

The extended TWSTFT schedule provided valuable insight into the achievable performance of this time transfer technique. Compared to earlier work the instability, e. g., in the PTB-NIST link, was improved and could be assessed for shorter averaging times. Plumb and Larson [26] deduced a value of 6×10^{-15} at τ = 2.5 days, Parker et al. [16] found 2.4×10^{-15} at τ = 2.5 days. Both previous studies were based on about six months of data taken at a time when TWSTFT between PTB and NIST was performed only three times per week so that the average τ_0 was about 2.5 days.

It turned out that during the current campaign the links to NIST were the more noisy ones. A different transponder on the satellite than for the intra-European links had to be utilized. Figure 4, showing PTB-NPL and PTB-NIST data, may serve as an illustration. Here we find $\sigma_x(\tau$ = 2 h, PTB-NPL) = 0.12 ns, and $\sigma_x(\tau$ = 2 h, PTB-NIST) = 0.28 ns. These findings are seemingly in contradiction to the observed standard deviation of the individual one second measurement results around the quadratic

regression function obtained at the respective sites which are reported in Figure 5. Apparently, the short-term measurement noise whose origin is not yet understood is efficiently averaged, and longer term variations dominate the links to the USA. Presumably, these are in general due to the longer baseline which may have an impact due to unaccounted satellite motion (respectively incorrect time tagging of the measurements) or due to uncorrelated environmental effects. Errors due to unaccounted satellite motion are distinctively sinusoidal in shape (with a one day period), and this was not observed. The temperature sensitivity of the NIST ground station as well as the unaccounted ionospheric delay (small at Ku-band anyway) can be ruled out as causes. What remains as a potential cause is the instability of the transponder on the satellite which was also noted by the United States Naval Observatory (which is routinely part of the TWSTFT network, but which did not participate in the extended measurement schedule).

The frequency instability obtained from the TWSTFT comparisons of all institutes with respect to the masers at OP and NPL is depicted in Figure 6. The slightly excessive noise in the links to IEN is probably due to the older TWSTFT modem of Mitrex type still used at IEN (see Table 1). At an averaging time of about one day the measurement noise (white phase noise) tends to become insignificant compared to the instability of the masers involved. The lowest measurement noise at averaging times below one day is observed in the links between NPL, OP, and PTB. We thus took this triple of stations to demonstrate the so-called closure (OP-NPL) + (NPL-PTB)+ (PTB-OP), the results are depicted in Figure 7. A 0.34 ns deviation from zero is observed, which can be partially explained by the mean clock rate of the OP maser of about −28 ns/day and the fact that the two measurements involving OP are separated in time by about 9 minutes (see Table 2). A real closure error would point to the fact that the signal delay in the receive path of at least one of the stations

depends on the received PRN code. The closure data exhibited a standard deviation around the mean value of 0.14 ns which represents the combined measurement noise of three TWSTFT measurements plus the time stability of the involved frequency standards over 9 minutes.

3.3 GPS CP results

The GPS CP data analysis revealed a very low level of measurement noise and thus an assessment of the instability of the frequency standards involved was possible for the shortest averaging times. As explained before, the original data provided by AIUB had UTC(PTB) and UTC(IEN), respectively, as the frequency references. In order to get a meaningful comparison of the three frequency transfer techniques, local measurements were used to provide the link to the masers at PTB and IEN, and only hourly data were dealt with. The results of GPS CP frequency comparisons were previously used as illustration in Figure 3. In Figure 8 we provide the frequency instability derived from these data. One plot was added which gives further proof of the merits of this technique. When the frequency step of the NPL maser (see Figure 3) after MJD 53328 is removed from the data, the NPL-NIST instability is further lowered. The frequency instability of UTC(NIST) has been estimated based on an analysis with respect to the group of active hydrogen masers in operation at NIST. Subtracting this noise contribution from the previously mentioned data should allow the combined instability of the frequency transfer and the NPL maser to be assessed (grey squares).

It is difficult to compare the findings of earlier studies [16, 17, 26] with those presented here since either the maser involved at PTB was not the same [26] or during the period of the study only few TWSTFT data points per week were recorded.

In general, the instability values reported here are equal or better by up to a factor of two than those demonstrated earlier.

3.4 TAI P3 results

In the GPS TAI P3 analysis the same raw data provided by the geodetic GPS receivers are employed as in the GPS CP analysis discussed before, but the code measurements rather than the phase data are employed. Thereby the analysis is simplified, and in fact, BIPM does the analysis on several links routinely as part of the monthly work of generating TAI. To fully assess the quality of the TAI P3 data it would have been necessary to extract GPS CP data for the epochs of the midpoints of the GPS TAI P3 16-minute average values. This was actually left out. In Figure 9, the frequency instability results obtained in the three links connecting to NPL are depicted. The links to OP and IEN each connect two hydrogen masers and are suited to assess fully the merits of the TAI P3 analysis. The PTB data are characterized by white frequency noise (slope $-1/2$) at averaging times around one day, and the observed instability is in almost perfect agreement with the expectation value for the primary clock CS2 of PTB [24], $\sigma_y(\tau = 1\ \text{day}) = 12\times10^{-15}$.

4. Comparisons between different techniques

4.1 Discussion TWSTFT – GPS

During this campaign, the offsets in the time differences provided by the three techniques are not of relevance since frequency comparisons, not true time comparisons, are the major concern. In this case only the variations of such offsets matter since they could cause an erroneous frequency measurement result, which would in turn be detrimental for the success of fountain comparisons [22], see Section 5. TWSTFT data and GPS CP data obtained for the same epoch - within

±15 s caused by the 30 s separation of the GPS CP data - were compared. In Figure 10 two examples are depicted. In both cases the mean offset in the data was subtracted. Whereas in the NPL-OP data all points are scattered almost within a 1-ns interval, a significant drift of about 2 ns is obvious in the NIST-OP data. The links among NPL, OP, and PTB seem not to be affected by any significant drift, whereas also other links between European stations and NIST show this drift. It was in the meantime found out in a joint effort of one of the Authors (R. D.) and Marc Weiss, NIST, that the apparent drift can be explained as a particularity of the receiver type in use at NIST, either caused by the receiver firmware or caused by the conversion to the RINEX format.

Other patterns were found in the comparisons between TWSTFT and GPS CP results, e. g., in the link IEN-OP a 1-ns drift during the first half of the comparison period is noticeable, and in the link IEN-NPL a 0.8-ns drift persists during the entire comparison period. In [17] it was discussed whether the choice of the algorithm for calculating average frequency values from the daily independent GPS CP processing batches would influence the result. Two algorithms furnished differences in the low 10^{-16} range. Further studies would be needed to assure that the new AIUB algorithm [19] provides perfectly unbiased results. More traditionally, changes in the signal delay in the involved equipment are suspected when two comparison techniques give different results for the quantity frequency. The earlier studies of this kind [11, 16, 26] had fewer data points spread over longer observation times at their disposal. In [16], the double differences TWSTFT minus GPS CP (NIST-PTB) stayed within 3 ns during more than 100 days, but exhibited a trend of about -3 ns/60 days towards the end of the study period (fall 2000). More than two years later, the same equipment was still in use [26], and the double differences stayed within ±2 ns over 200 days with no distinct signature. Little further information on long-term performance of the

equipment involved in the current study could be retrieved from the available data bases. Note that the receiver NISU (see Table 1) is not the same as the one used earlier at NIST. The GPS CP data analysis was only made for the purpose of this study, and IGS products have not been analyzed yet.

When the stability of the current double-difference data TWSTFT minus GPS CP for the links between OP and all four stations is analyzed, the results depicted in Figure 11 are obtained. Since the data represent the combined uncertainty of two techniques, the European link data prove the capability of frequency comparisons with a measurement uncertainty of close to 1 part in 10^{15} at averaging times of 1 day to be made utilizing the hardware configuration and measurement schedule prevailing during the current study period. The potential to do even better once a longer averaging time can be accepted is appealing. But the delay variations of all involved equipment deserve careful monitoring for a correct assessment of the measurement uncertainty in such cases. A look at long term comparisons between TWSTFT and GPS TAI P3 confirms this statement.

As the GPS TAI P3 data are based on the same GPS equipment, the double differences TAI P3 minus TWSTFT are also worth being analyzed. In [12] the instability of the TAI P3 frequency comparisons between OP and PTB was estimated based on double differences with simultaneous TWSTFT measurements. Only 35 TWSTFT data were, however, available from the 62-day period, and mod$\sigma_y(\tau)$ = $8.5\text{x}10^{-15}$ was estimated at τ = 42 h. As mod$\sigma_y(\tau)$ is sensitive to τ_0, this value is not directly comparable with those obtained during the present experiment. In [11, 15] and references therein extensive comparisons of TWSTFT and GPS TAI P3 were analyzed. Overall, it was estimated that mod$\sigma_y(\tau)$ is of order 2×10^{-15} or below at τ = 3-

4 days. It was shown that that over extended periods (several months) the differences between the two techniques are characterized by a standard deviation below 1 ns under favorable conditions. Some cases show a somewhat poorer behavior, e. g., for the link IEN-PTB an extended continuous set of data which is depicted in Figure 12 was recorded. One can concede that periods of stable operation of the equipment alternate with significant changes. Therefore the excellent results shown in Figure 10, upper graph, cannot be immediately considered valid for extended periods and for all links. It seems that the results discussed in this paper were obtained under quite favorable conditions.

4.2 Link comparisons and discussion of the frequency transfer uncertainty

As said before, the links between IEN, NPL, and OP, where fountains were operating during the study period, are of greatest interest. The graphs shown in Figure 13 represent in a certain way the summary of the finding presented in different graphs before. We demonstrate the instability of frequency comparisons for the links NPL-OP, NPL-IEN, and IEN-OP using the three techniques. The same frequency references were used in all cases (in particular the maser at IEN). In addition, the graphs contain the calculated instability for the double differences TWSTFT minus GPS CP. From this analysis the instability of a single frequency transfer method cannot be separated but clearly it must be lower than the (combined) instability of the double differences. We notice that the instability of GPS CP frequency transfer reaches the clocks' instability at about half a day, in case of TWSTFT (12 measurements per day) a full day is needed. About 2 days are needed in case of TAI P3. In the case of longer averaging times there is no significant distinction between the three techniques visible from the graphs because of the clocks' instability. The instability of double differences reflects probably the long-term delay variations of the

involved equipment. Although the equipment is the same for GPS CP and GPS TAI P3 some systematic differences in the long term cannot be rules out since GPS CP is based on phase measurements only whereas GPS TAI P3 is based on code measurements. Different sensitivities of the two types of observations on environmental changes were occasionally reported, so the matter deserves further long term studies.

In conclusion of this section we demonstrate the mean frequency differences obtained in comparisons of the hydrogen masers along selected links using the different techniques. The mean values were calculated in a straightforward manner, (ΔT(end) minus ΔT(start)) divided by the measurement interval. We chose the triplets of stations IEN-NPL-OP and NPL-OP-PTB. GPS TAI P3 data with reference to the PTB maser were calculated based on its average frequency with respect to UTC(PTB) during the full period. The results are compiled in Tables 3 and 4. No closure values are given for the GPS CP analysis because they are zero in consequence of the consistent network solution. The maximum deviation within one column is less than 1×10^{-15} for each baseline, but it is typically below 0.5×10^{-15}.

5 Fountain comparisons

The institutes OP, IEN, and NPL provided data of their fountains during a 20 days long period (MJD 53304-53324), allowing, for the first time, a comparison between three primary frequency standards of this kind in simultaneous operation to be made. Results were reported earlier [22] and subsequently used by BIPM for the determination of the TAI scale unit. The detailed uncertainty budgets for the fountains valid during the comparisons were described in [22] and in the reports submitted to BIPM by each laboratory [28]. Table 5 summarizes the data contained in the BIPM

reports. The methods used to evaluate the accuracy of the three fountains are not identical, detailed descriptions were reported in [5], [1] and [4] for IEN-CsF1, OP-FO2 and NPL-CsF1, respectively.

The usual method adopted for fountain frequency comparisons, referred to as "average" in this paper, consists in a first step of measuring the average frequency of the local maser with respect to the fountain during the evaluation period. Then in a second step the average frequencies measured locally between masers and fountains are combined with the average of the maser frequency differences obtained by remote frequency transfer. Following this approach, the estimate of the measurement uncertainty of the average frequency between the fountains requires that a model for the maser frequency fluctuations during the time intervals when the fountain is not running (dead time) is known [29]. This could be a critical point, especially when the maser fluctuations are not as easy to model as a linear frequency drift.

During the comparison experiment reported here another technique for data analysis was tested, which is referred to as "synchronous" in this paper. Following the TWSTFT measurement schedule, the comparison period was divided in 2-hour intervals which began and ended mid time between each individual TWSTFT session (see Table 2). The aim was to synchronize the time tags of all measurements – local and remote ones – within a $\pm$15 minute period.

For each 2 hours interval, measurements of the maser frequency with respect to the fountains (local) and of the frequency difference between the masers via the transfer methods (remote) were performed and later analysed. Figure 14 reports, for example, the frequency data of the HM at IEN as measured by IEN-CsF1. Each data point represents the frequency averaged over 2 hours and the error bars represents

the associated statistical uncertainty, limited in this case by the fountain frequency instability.

The frequency differences between pairs of fountains (IEN-CsF1 – NPL-CsF1), (IEN-CsF1 – OP-FO2), and (NPL-CsF1 – OP-FO2) were evaluated. This ensemble of values was averaged at the end of the campaign. Thereby only those measurement results were considered for which the respective pair of fountains was in operation at the same time (with 2 hour resolution). This approach enabled the potential limitations to such comparisons set by the dead time in the fountain operation to be overcome, at the cost of reducing the total available measurement time.

The results of the comparison are shown in Table 6. Each column reports the differences of the fountains operated in the institutes as given in the header. The first two lines report the results using the TWSTFT and the GPS-CP transfer methods in the synchronous mode. The last three lines report the results of a standard data analysis based on the fountains vs. maser data average over the whole period of the comparison.

The uncertainty of each comparison between couples of fountains (reported in brackets in the columns) is calculated combining the uncertainty contributions u_A and u_B of each fountain as reported in Table 5 and the link uncertainty contribution which can be conservatively taken as 1×10^{-15} according to the analysis about the link instability and the closure errors reported in the previous paragraphs. It turned out that the two evaluation techniques, “average” and “synchronous” in some cases entail differences which are not negligible compared to the estimated combined uncertainty of the comparison.

6 Conclusion

We have reported on the results of a campaign originally stimulated by the intention to compare caesium fountain frequency standards in five institutes. GPS CP based analysis was used for the first time in such a campaign comprising a network of institutes. The campaign has provided much data not all of which could be presented in detail here. The following results are very encouraging:

1. TWSTFT in an intensified schedule and GPS CP analysis have allowed a frequency comparison between remote standards with a statistical uncertainty of 1×10^{-15} or only marginally larger at averaging times of one day. The GPS CP requires extra computational efforts whereas TWSTFT is done routinely and only an extension of the standard measurement schedule needed to be agreed upon. GPS TAI P3, on the other hand, requires averaging times of several days to reach a similar statistical uncertainty.
2. All studied techniques seem to be equally suited for the comparison of fountains; the associated measurement noise does not contribute significantly to the combined uncertainty of such comparisons, provided that a few days of averaging are allowed.
3. The analysis presented here (see Tables 3 and 4) shows that all techniques provide the same mean frequency difference between the standards involved within the 1-σ measurement uncertainty of few parts in 10^{16}. Any potential systematic effects associated with the comparison techniques have a non-negligible contribution when combined with the intrinsic uncertainty of the fountains.

Further studies should address the following subjects

1. the cause of the excessive noise observed in some TWSTFT links and of the relatively large closure error when the average frequency OP-NPL-IEN is calculated;
2. a more elaborate analysis of the differences between synchronous and non-synchronous ("average") techniques for frequency comparisons of the fountains which might explain the differences within the columns of Table 6.

The low measurement noise provided by GPS CP at averaging times of less than one day makes it the first choice for use in comparisons among remote optical frequency standards since it is currently very challenging to operate such devices for extended periods. In the existing network such a comparison could be organized between NIST, NPL, and PTB, whose standards would be almost immediately ready for such a purposes, and could immediately draw on the experiences provided by the current study.

Acknowledgement

This work would not have been possible without the support of several individuals in the participating stations who supported the operation of equipment, data retrieval, and analysis: Jürgen Becker, John Davis, Philippe Merck, Diego Orgiazzi, Valerio Pettiti, Jean-Yves Richard, and David Valat. The transponder on the IS-903 was provided free of charge by Intelsat Corp.

References

[1] Bize S et al. 2004 Advances in atomic fountains *C. R. Physique* **5** 829-43

[2] Jefferts S R et al. 2002 Accuracy evaluation of NIST-F1 *Metrologia* **39** 321-36

[3] Weyers S, Bauch A, Schröder R and Tamm Chr 2001 The Atomic caesium fountain CSF1 of PTB *Proc. 6th Symposium on Frequency Standards and Metrology (St. Andrews, Scotland, Sept. 2001)* pp 64-71

[4] Szymaniec K, Chalupczak W, Whibberley P B, Lea S N and Henderson D 2005 Evaluation of the primary frequency standard NPL-CsF1 *Metrologia* **42** 49-57

[5] Levi F, Lorini L, Calonico D and Godone A 2004 IEN-CsF1 accuracy evaluation and two-way frequency comparison *IEEE Trans. Ultrason. Ferroelec. and Freq. Control* **51** 1216–24

[6] Wynands R and Weyers S 2005 Atomic fountain clocks *Metrologia* **42** S64-S79

[7] Arias F and Guinot B 2005 Atomic timekeeping from 1955 to the present *Metrologia* **42 S**20-S30

[8] Levine J 2002 Time and frequency distribution using satellites *Rep. Progr. Phys.* **65** 1119–64

[9] Kirchner D 1999 Two-way Satellite Time and Frequency Transfer (TWSTFT): Principle, implementation, and current performance *Review of Radio Sciences 1996-1999, Oxford University Press* 27-44

[10] Defraigne P and Petit G 2003 Time transfer to TAI using geodetic receivers *Metrologia* **40** 184-88

[11] Petit G and Jiang Z 2005 Stability of geodetic time links and their comparison to two-way time transfer *Proc. 36th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting (Washington DC 2004)* pp 31-40

[12] Richard J-Y et al. 2004 Comparison of remote cesium fountains using GPS P3 and TWSTFT links *Proc. 18th European Frequency and Time Forum (Guildford UK 2004),* on CD-ROM

[13] Schildknecht T, Beutler G, Gurtner W and Rothacher M 1990 Towards subnanosecond GPS time transfer using geodetic processing techniques *Proc. 4th European Frequency and Time Forum (Neuchâtel, Switzerland, 1990)* pp 335-46

[14] Larson K M and Levine J 1999 Carrier-phase time transfer *IEEE Trans. Ultrason. Ferroelec. and Freq. Control* **46** 1001-12

[15] Ray J and Senior K 2005 Geodetic techniques for time and frequency comparisons using GPS phase and code measurements *Metrologia* **42** 215-32

[16] Parker T, Hetzel P, Jefferts S, Weyers S, Nelson L, Bauch A and Levine J 2001 First comparison of remote cesium fountains *Proc. IEEE Int. Freq. Control Symp. and PDA exhibition (Seattle, Washington, 2001)* pp 63-8

[17] Hackman C, Levin J, Parker T, Piester D and Becker J 2005 A new technique for estimating frequency from GPS carrier-phase time transfer data *Proc. 2004 IEEE Int. Ultrasonics, Ferroelectrics, and Frequency Control Joint 50th Anniversary Conference (Montreal, Canada 2004)* on CD-Rom 10005500.pdf and pp 341-9

[18] Beutler G, Rothacher M, Schaer S, Springer T A, Kouba J and Neilan N E, 1999 The International GPS Service (IGS): An interdisciplinary service in support of Earth sciences *Adv. Space Res.* **23** 631-5, internet address http://igscb.jpl.nasa.gov, see also Meindl M ed. 2005 *Proc. Celebrating a decade of the International GPS Service – workshop and symposium (Bern Switzerland 2004)*

[19] Dach R, Bernier L-G, Dudle G, Schildknecht T and Hugentobler U 2005 Geodetic frequency transfer – actual developments in the AIUB-METAS collaboration, *Proc. 19th European Frequency and Time Forum (Besançon, France, 2005)* in print

[20] Hugentobler U, Schaer S and Fridez P, eds. 2001 The Bernese GPS Software 4.2., Astronomical Institute of the University of Bern

[21] Bauch A et al 2005 Time and frequency comparisons between four European timing institutes and NIST using multiple techniques *Proc. 19th European Frequency and Time Forum (Besançon, France, 2005)* in print

[22] Calonico D et al. 2005 Comparison between remote Cs fountain primary frequency standards *Proc. 19th European Frequency and Time Forum (Besançon, France, 2005)*, in print

[23] Parker T 1999 Hydrogen maser ensemble performance and characterization of frequency standards *Proc. Joint meeting of the European Frequency and Time Forum and the IEEE Int. Freq. Control Symp. (Besançon, France, 1999)* pp 173-76

[24] Bauch A 2005 The PTB primary clocks CS1 and CS2 *Metrologia* **42** S43-S54

[25] Recommendation ITU-R TF.1153-1 The operational use of two-way satellite time and frequency transfer employing PN time codes, ITU Radiocommunication Study Group (Geneva) last update 2003

[26] Plumb J E and Larson K M 2005 Long-term comparisons between two way satellite and geodetic time transfer systems *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control* accepted for publication

[27] Allan D W and Thomas C 1994 Technical directives for standardization of GPS time receiver software *Metrologia* **31** 69-79

[28] ftp://ftp2.bipm.fr/pub/tai/data/PFS_reports/

[29] Douglas R J and Boulanger J S 1997 Standard uncertainty for average frequency traceability *Proc. 11th European Frequency and Time Forum (Neuchâtel, Switzerland,1997)* pp 345-9

Tables and Table Captions

Table1. Institutes involved and equipment in use

Institute	Clock or frequency reference	TWSTFT modem	IGS station code	GPS receiver antenna
IEN	HM: Symmetricom, UTC(IEN): Agilent 5071	MITREX	IENG	ASHTECH Z-XII3T, ASH701945C_M
NPL	HM: Symmetricom, representing UTC(NPL)	SATRE	NPLD	ASHTECH Z-XII3T, AOAD/M_T
NIST	HM: Symmetricom, representing UTC(NIST)	SATRE	NISU	NOV EURO4-1.00-222, NOV600
OP	HM: Symmetricom	SATRE	OPMT	ASHTECH Z-XII3T, 3S-02-TSADM
PTB	HM: VREMYA-CH UTC(PTB): primary clock CS2	SATRE	PTBB	ASHTECH Z-XII3T, ASH700936E

Table 2. TWSTFT measurement schedule and Mitrex codes assigned to the participating stations. The receive (RX)and transmit (TX) frequencies in Europe differ between the intra-European link and the USA-Europe link.

Measurements during each second hour of a day in UTC			OP		NPL		PTB		IEN		NIST	
First	Last	Duration										
hh:mm:ss	hh:mm:ss	(s)	TX	RX	TX	RX	TX	RX	TX	RX	TX	RX
00:10:00	00:11:59	120	0	1	1	0						
00:13:00	00:14:59	120					4	6	6	4		
00:19:00	00:20:59	120	0	4			4	0				
00:22:00	00:23:59	120			1	4	4	1				
00:25:00	00:26:59	120	0	6					6	0		
00:28:00	00:29:59	120			1	6			6	1		
00:37:00	00:38:59	120	0	6							6	0
00:40:00	00:41:59	120			1	6					6	1
00:49:00	00:50:29	120					4	6			6	4
00:52:00	00:53:59	120							6	6	6	6

Table 3 Mean relative frequency differences in parts in 10^{15} during the interval 53304 to 53323 (inclusive, 20 days) between hydrogen masers at IEN, NPL, and OP, calculated using the three frequency comparison techniques, and the closure results.

	IEN – OP	IEN – NPL	NPL – OP	Closure
TWSTFT	-430.8	-95.3	-336.1	0.6
GPS CP	-430.8	-94.5	-336.3	
TAI P3	-431.6	-95.1	-336.5	0.0

Table 4 Mean relative frequency differences in parts in 10^{15} during the interval 53304 to 53324 (inclusive, 21 days) between hydrogen masers at NPL, OP, and PTB and the closure results.

	PTB – NPL	PTB – OP	NPL – OP	Closure
TWSTFT	-74.2	-410.2	-336.1	0.1
GPS CP	-73.9	-410.1	-336.2	
TAI P3	-73.7	-410.1	-336.5	0.1

Table 5 Uncertainties of IEN-CsF1, OP-FO2, NPL-CsF1 during the fountain comparison experiment (MJD 53304 – 53324), as declared by each laboratory in its report to BIPM [28]. Here u_A is the uncertainty originating from the instability of the devices, u_B is the combined uncertainty from systematic effects.

	u_A	u_B	u_{tot}
IEN-CsF1	0.3	1.0	1.1
OP-FO2	0.2	0.7	0.7
NPL-CsF1	0.6	1.0	1.2

Table 6 Mean relative frequency differences in parts in 10^{15} between the fountain frequency standards operated in the institutes as given in the header during the interval 53304 to 53324 (last day not included, 20 days total measurement time). The values in brackets reflect the combined uncertainty contribution (combination of the uncertainty of the two fountains and of the involved link).

	IEN-OP	IEN-NPL	OP-NPL
TWSTFT Synchronized	3.2 (1.6)	5.4 (1.9)	0.9 (1.7)
GPS CP Synchronized	4.7 (1.6)	5.0 (1.9)	-0.1 (1.7)
TWSTFT Average	3.8 (1.6)	4.1 (1.9)	0.3 (1.7)
GPS CP Average	3.8 (1.6)	4.1 (1.9)	0.3(1.7)
GPS TAI P3 Average	3.0 (1.6)	3.5 (1.9)	0.5 (1.7)

Figures and Figure Captions follow

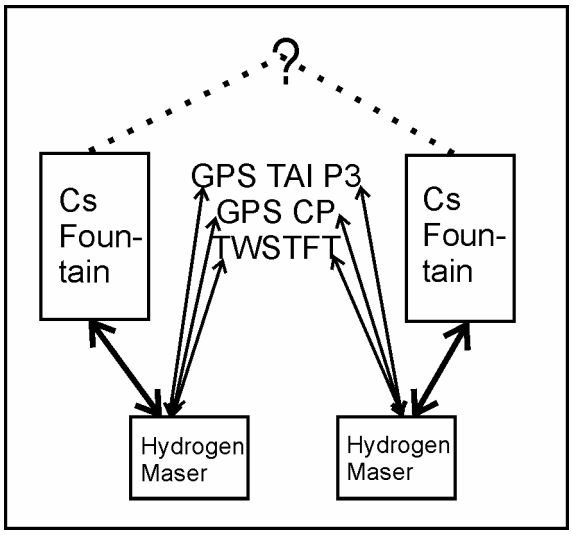

Figure 1. Configuration of frequency comparisons during the campaign; the frequency standards at left and at right are operated in different institutes.

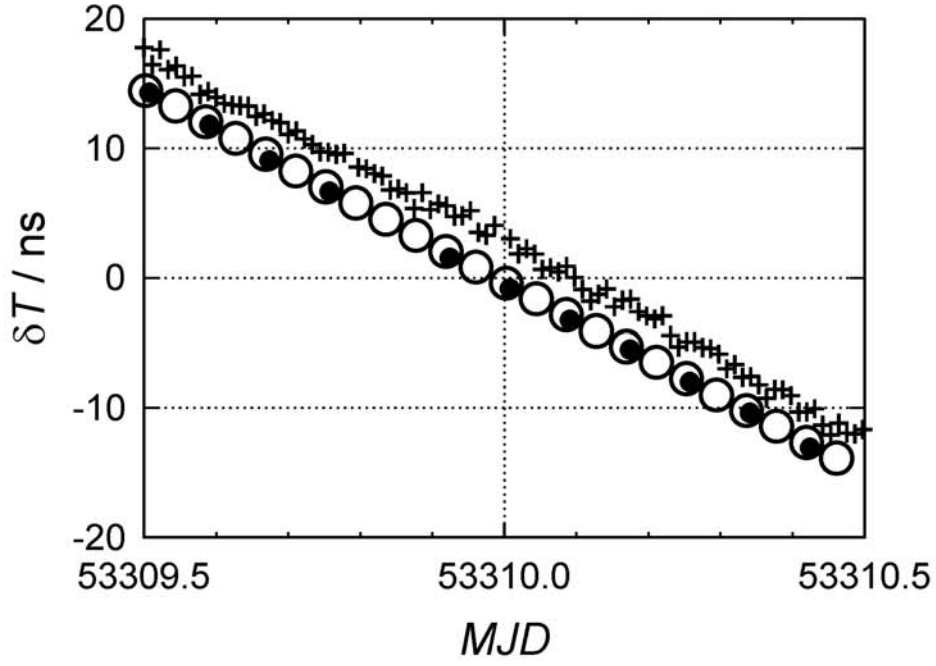

Figure 2. Example of data collected during one day on the NPL – OP link; GPS CP (open circles), TWSTFT (full dots), and GPS TAI P3 (crosses), the latter offset by 2 ns for clarity.

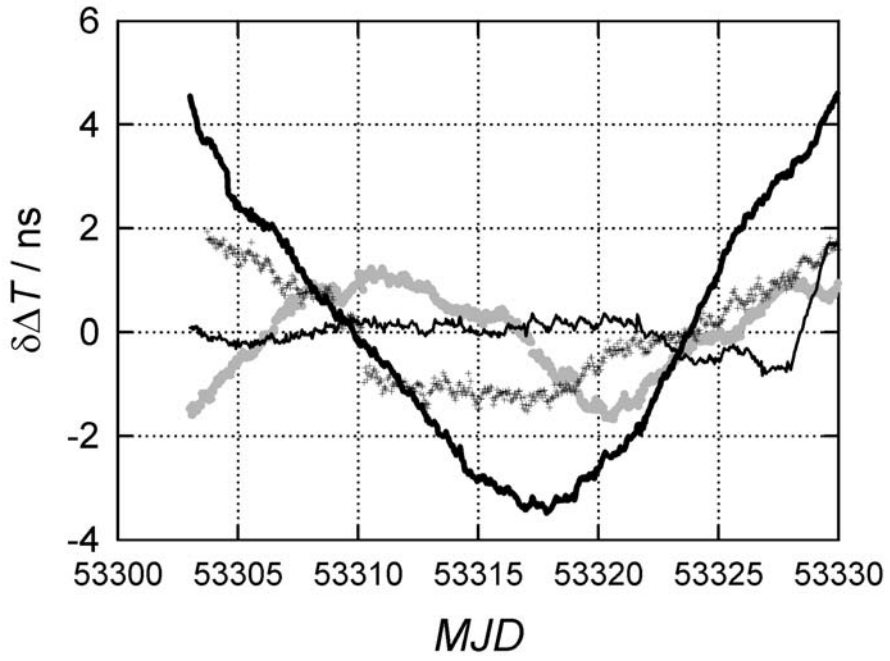

Figure 3. General characteristics of the frequency standards at NIST (common reference), IEN (crosses), NPL (thin line), OP (thick line), and PTB (grey line) during the study period, MJD 53303 – 53330; residuals to linear fits to time differences UTC(NIST) − local references obtained via GPS CP.

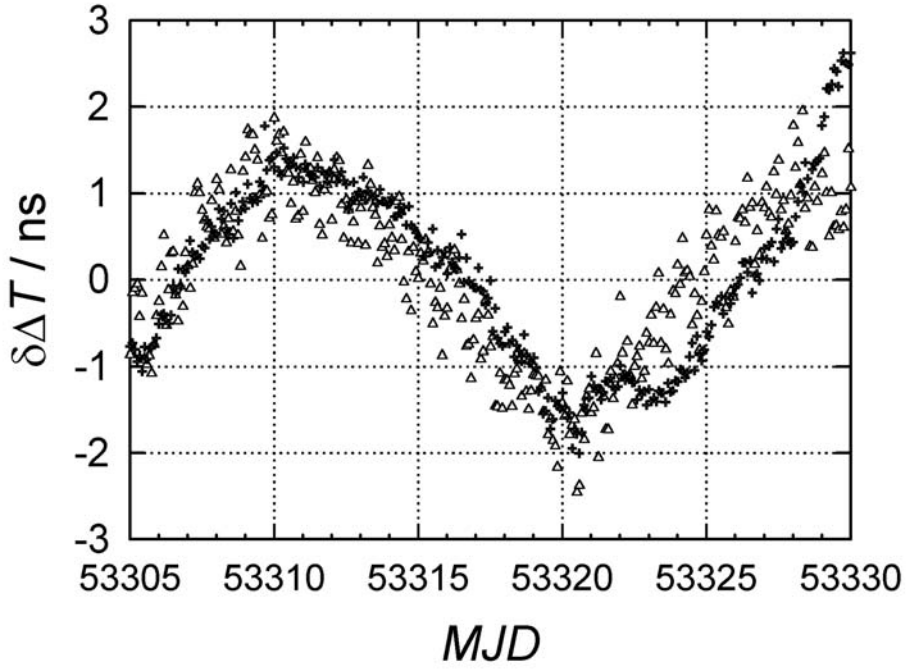

Figure 4. Time differences PTB-NPL (+), and PTB-NIST (Δ) obtained via TWSTFT; mean time offset and mean rate removed.

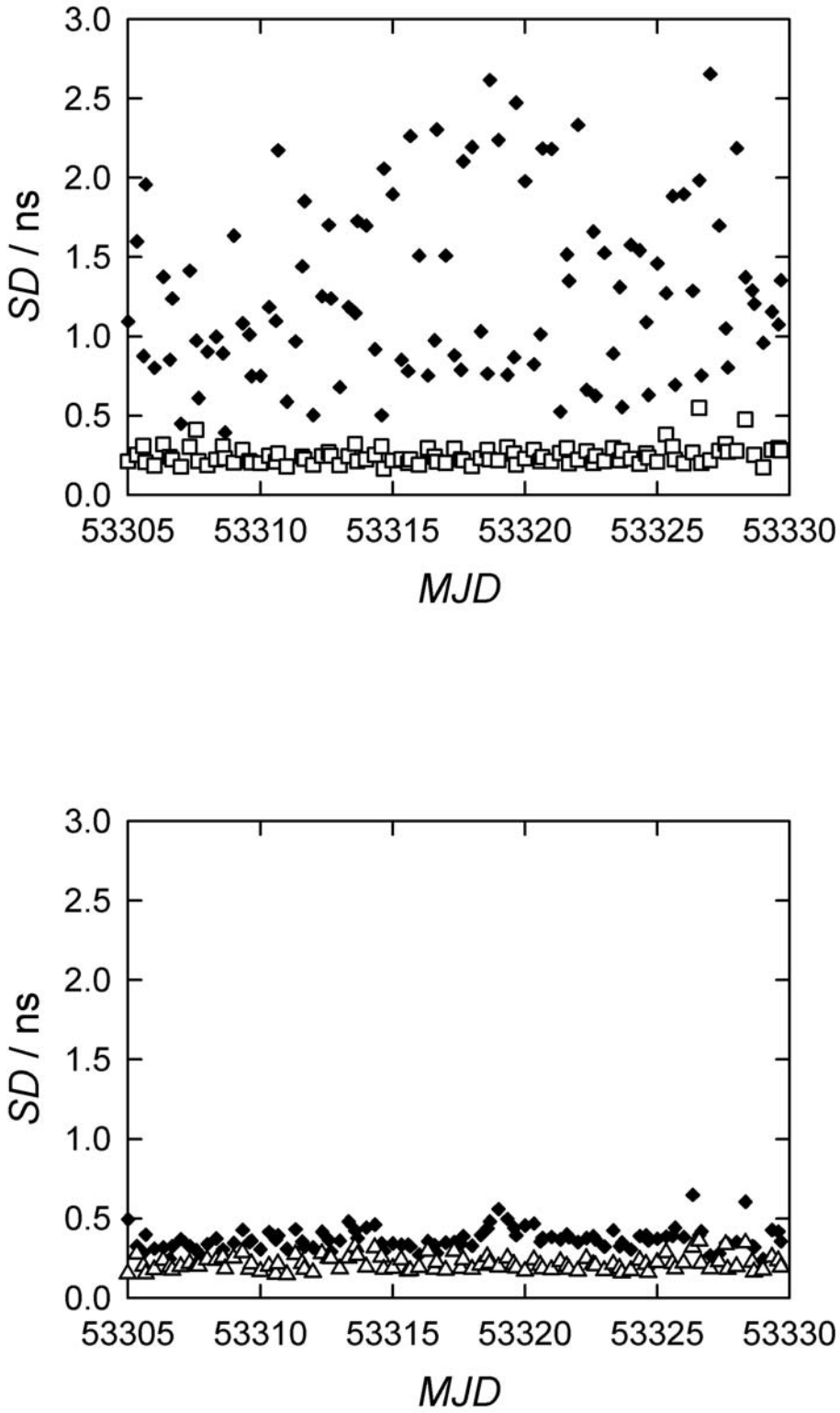

Figure 5 Observed instability of the TWSTFT data for the links PTB-NPL (upper) and PTB-NIST (lower); standard deviation *SD* in ns of the individual one second measurement data recorded at PTB (♦), NPL (□) and NIST (∆), respectively, around the quadratic regression function during the comparison campaign.

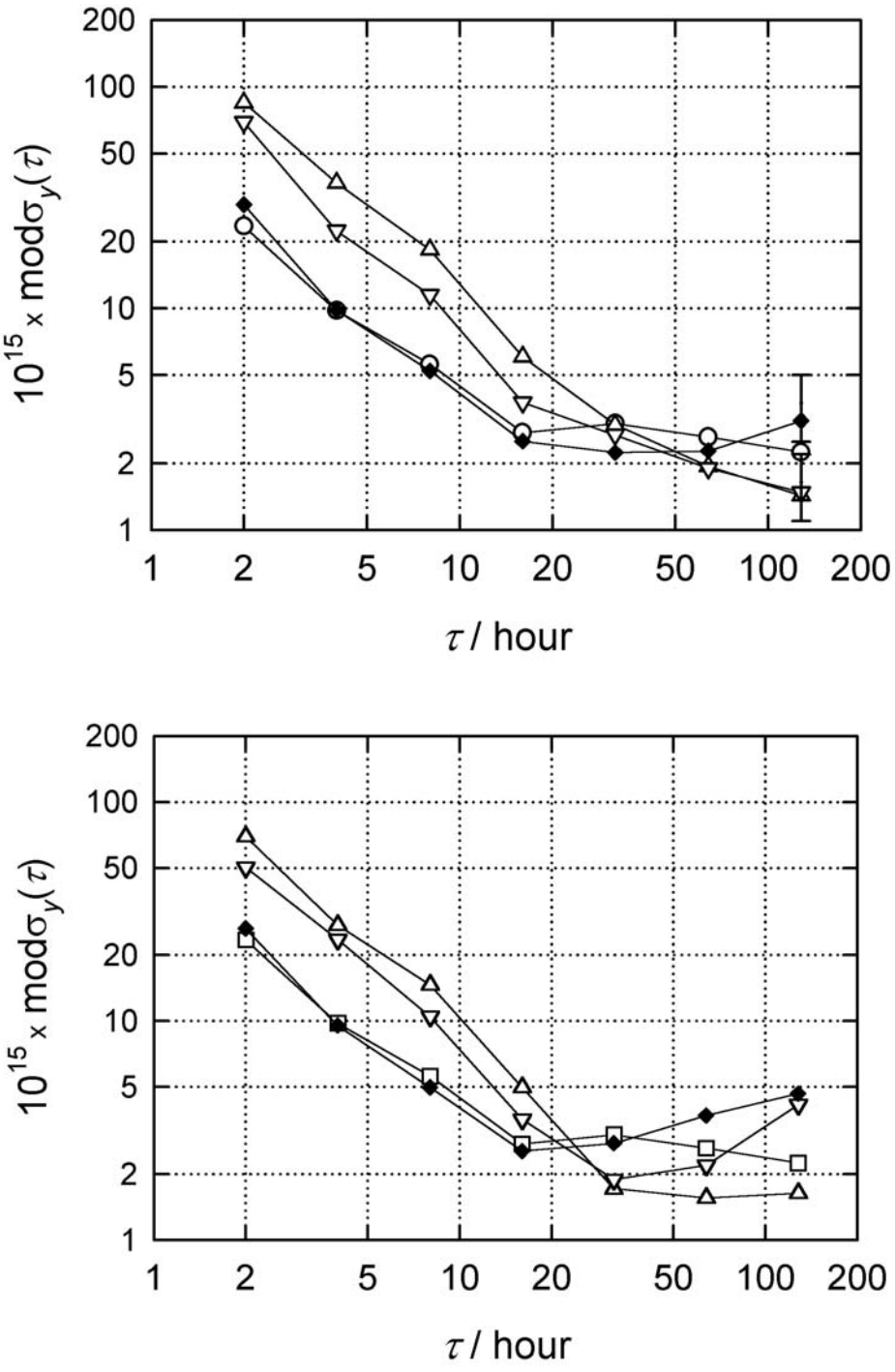

Figure 6. Relative frequency instability, expressed by mod$\sigma_y(\tau)$, in the TWSTFT comparison of masers at NPL (upper graph) and OP (lower graph) to the other participating stations. The symbols reflect the remote station for each link: IEN (∇), NIST (Δ), NPL (□), OP (O), and PTB (♦), respectively. For clarity of the presentation, we indicate only the size of uncertainty bars for the two last points in the upper graph. The bars are of the size of the symbols for averaging times of 32 hours or less.

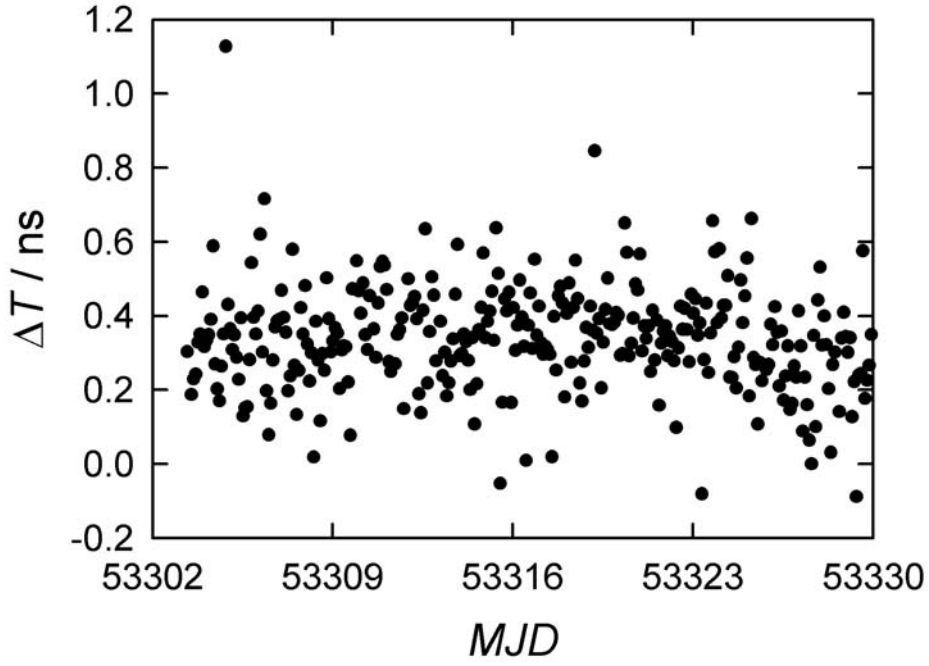

Figure 7. TWSTFT closure results – in total 312 data points - combining the three individual measurements OP-NPL, PTB-OP, and NPL-PTB collected during the campaign according to the schedule represented in Table 2.

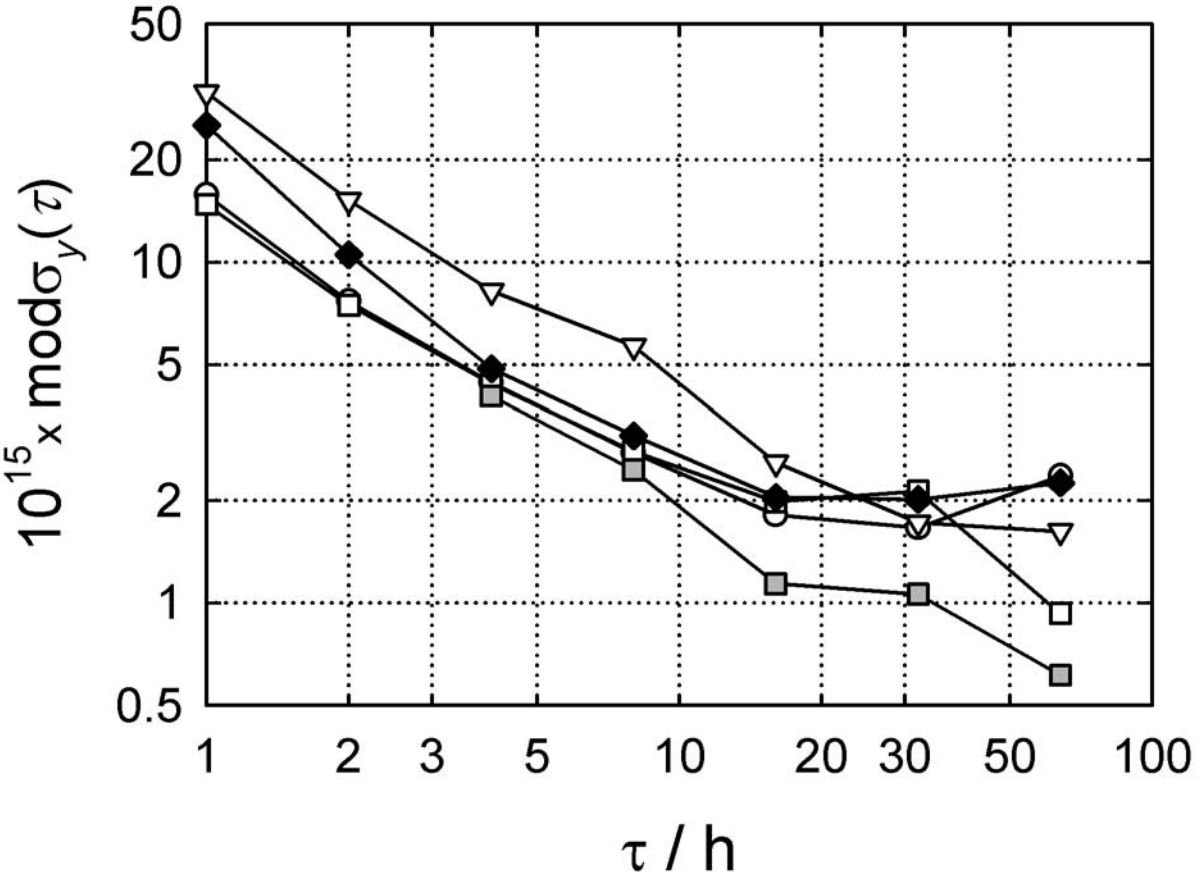

Figure 8. Relative frequency instability, expressed by mod$\sigma_y(\tau)$, in the GPS CP comparison of UTC(NIST) with masers at OP, NPL, PTB, and IEN, coding of symbols as before, IEN (▽), NPL (□), OP (O), and PTB (♦). Additional plot: instability after removal of the frequency step of the NPL maser and after subtracting thereof the estimated instability of UTC(NIST) (grey squares).

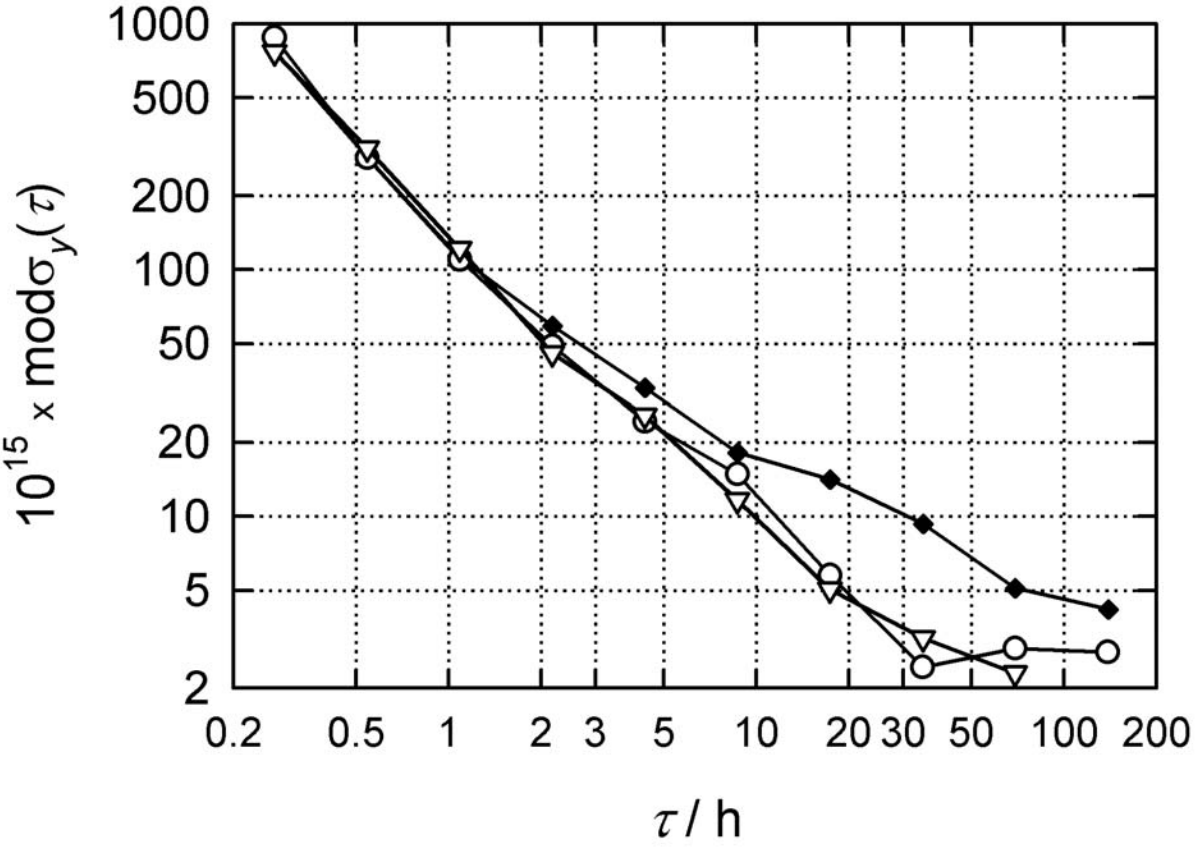

Figure 9. Relative frequency instability, expressed by mod$\sigma_y(\tau)$, in the GPS TAI P3 comparison of UTC(NPL) with the maser at OP (O), and IEN (∇), and UTC(PTB) (♦).

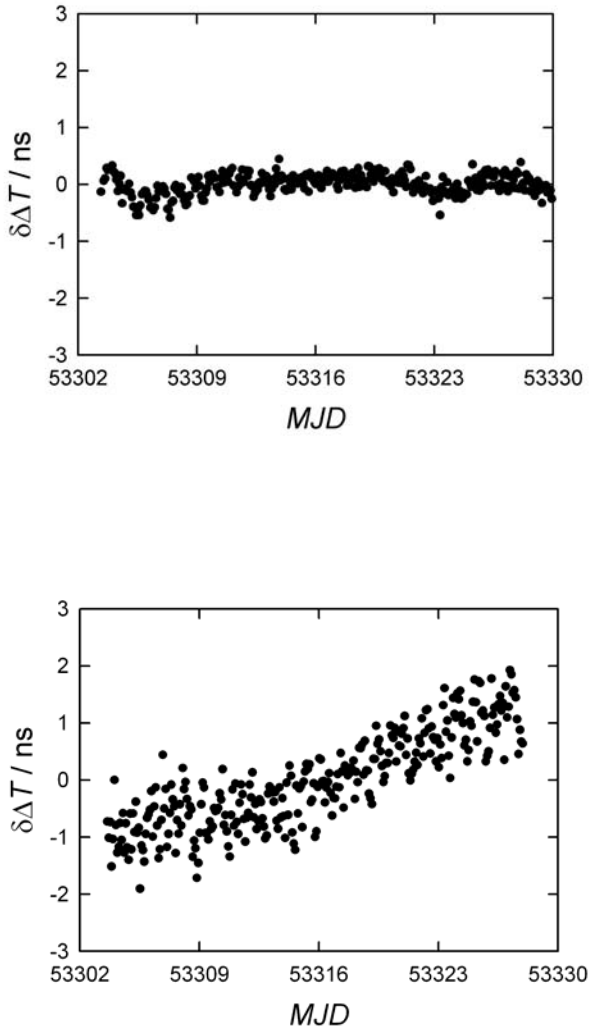

Figure 10. Double differences of TWSTFT − GPS CP results, mean offset subtracted, upper: NPL-OP, lower: NIST-OP.

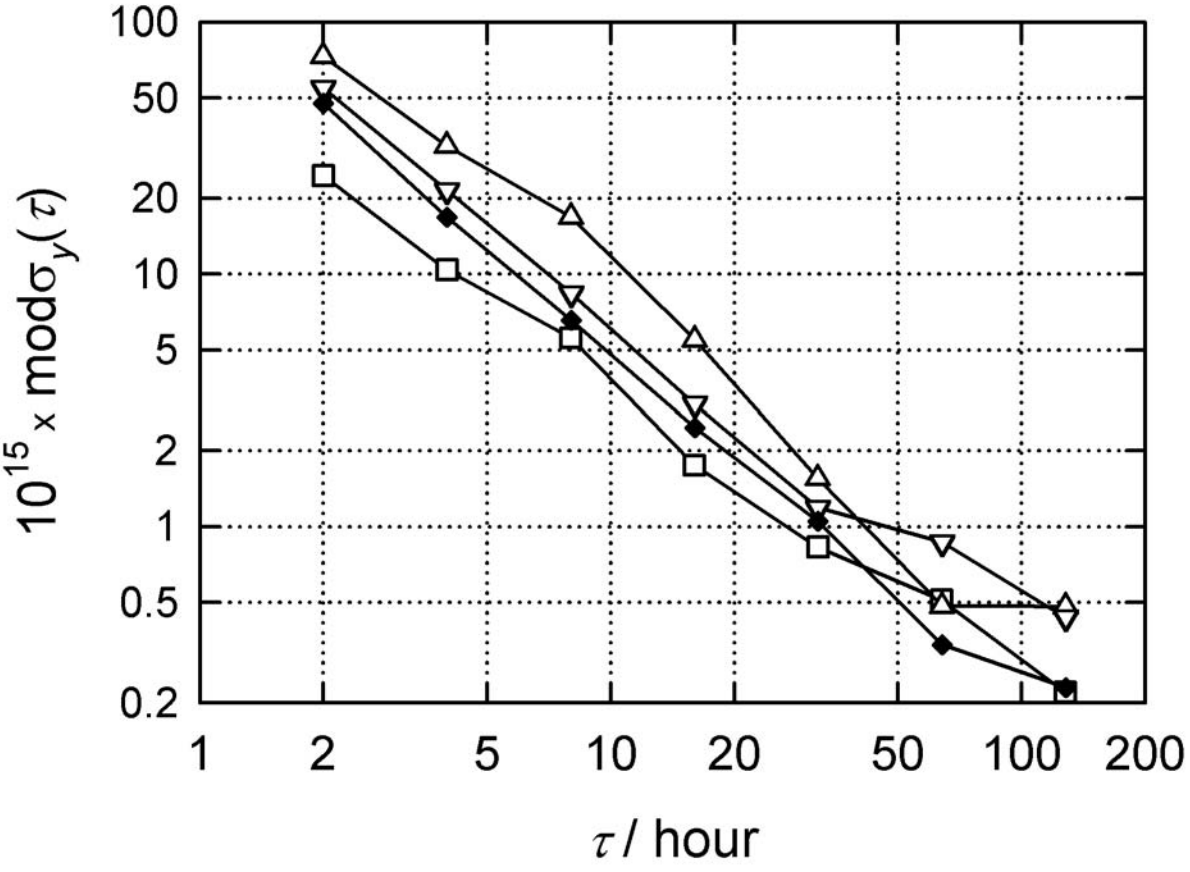

Figure 11. Stability analysis of the double difference TWSTFT – GPS CP for all links referenced to OP, coding of the symbols IEN (∇), NIST (Δ), NPL (□), and PTB (♦).

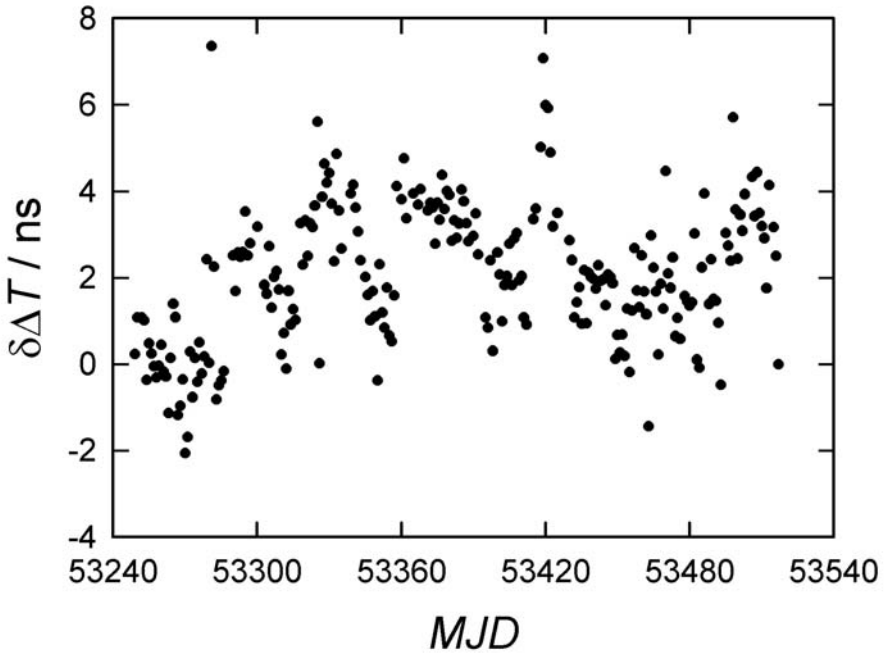

Figure 12. Double differences of TWSTFT − GPS TAI P3 for the link IEN-PTB during 10 months, mean offset subtracted.

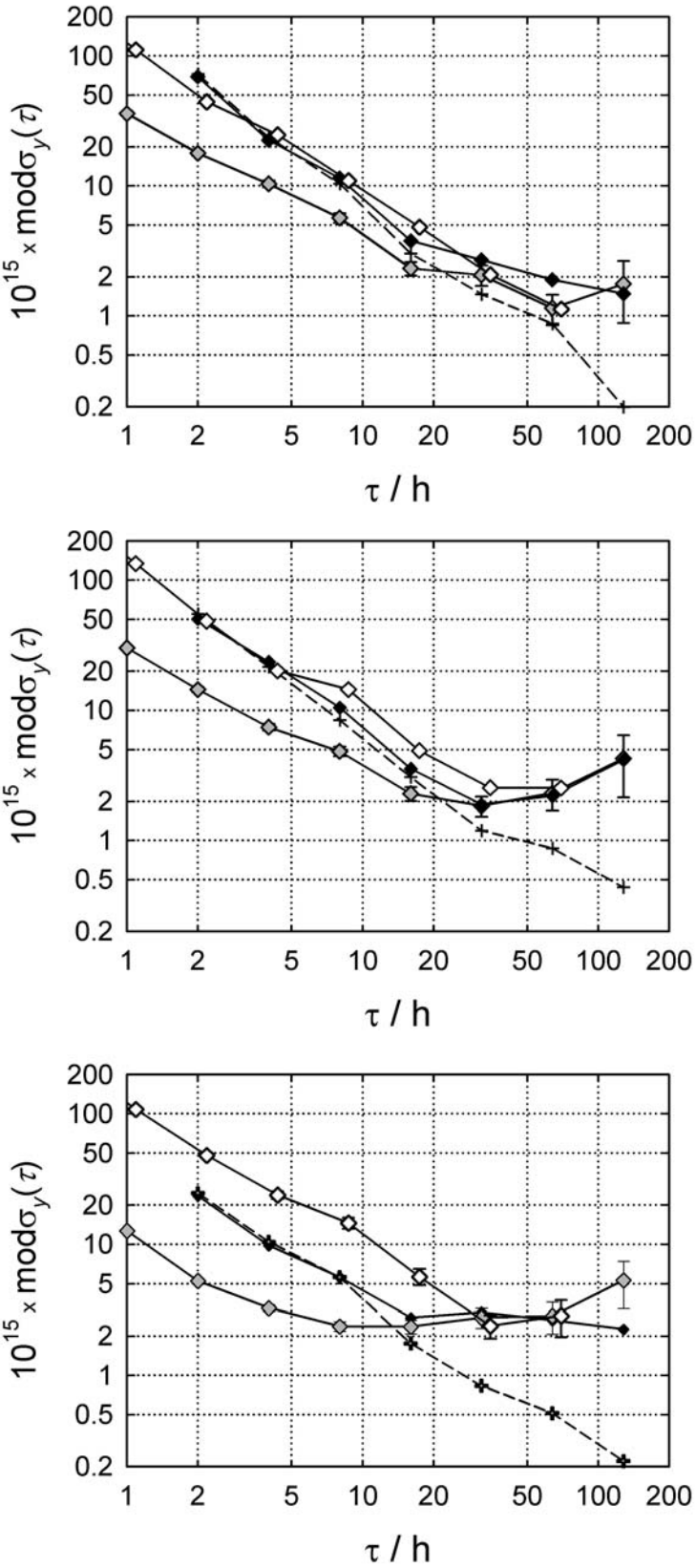

Figure 13. Relative frequency instability, expressed by mod$\sigma_y(\tau)$, obtained for the three techniques, GPS TAI P3 (white), GPS CP (grey), and TWSTFT (black);. upper graph: link IEN-NPL: middle graph: link IEN-OP; lower graph: link NPL-OP. In addition, the instability of the double-difference TWSTFT – GPS CP (crosses) has been illustrated for each link.

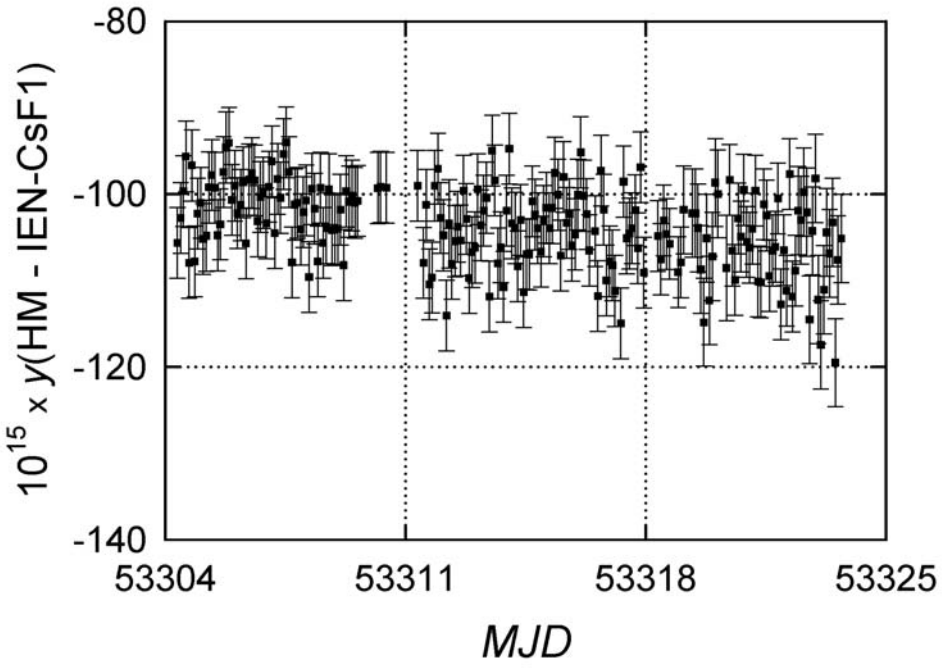

Figure 14. Frequency comparison between IEN-CsF1 and the local hydrogen maser. Points represent the frequency averaged over 2 hours and the error bars represent the associated uncertainty.