AVERAGING SATELLITE TIMING DATA FOR NATIONAL AND INTERNATIONAL TIME COORDINATION

Judah Levine
Time and Frequency Division, National Institute of Standards and Technology, and
JILA, University of Colorado
Boulder, CO 80305, USA

Abstract

The International Bureau of Weights and Measures (BIPM) calculates International Atomic Time (TAI) and Coordinated Universal Time (UTC) using data from many national metrology institutes and timing laboratories. An important part of these data are the measured time differences between GPS (or GLONASS) time and the realization of UTC by each of the laboratories, UTC (lab). These time differences are acquired using tracking schedules published by the BIPM and are based on 1-second measurements averaged as specified in the corresponding technical directives.

The concept of a tracking schedule and the algorithms that are used for averaging the data were designed many years ago when all of the contributing laboratories used single-channel receivers with relatively slow internal processors. Although these receivers are still in use, many laboratories also use multi-channel receivers with much greater processing power. In addition, the speed of the network that links the contributing timing centers continues to increase and the cost of storage devices continues to decrease. Both of these developments make it feasible to acquire and store more data.

Given these advances, it is appropriate to reconsider the design of the averaging algorithms. In particular, I will show that the current 13-minute averaging scheme is not optimum in general, and that a shorter and simpler averaging scheme would provide a better means of handling the effects of multipath reflections and similar problems, which are not attenuated by common-view subtraction. In addition to remaining compatible with the method used in the existing receivers, it is desirable to design an averaging algorithm that could be compatible with data acquired by geodetic (carrier-phase) receivers, which typically report measurements every 30 s. In principle, these data cannot be made compatible with the algorithms currently specified for the 13-minute tracks; the incompatibilities will be most serious at sites with large multipath reflections or other noise sources that are not attenuated by common-view subtraction.

INTRODUCTION

In the common-view method, several stations passively receive a signal from the same transmitter. This method cancels or attenuates common mode offsets and errors, and it has been widely used to distribute time and frequency information for this reason. For example, the computations of International Atomic Time (TAI) and Coordinated Universal Time (UTC) depend on clock data submitted to the International Bureau of Weights and Measures (BIPM) by national metrology institutes and timing laboratories, and common-view comparisons of the time scales of these laboratories play a central role in supporting these data transfers. Short-baseline common-view observations are also used to calibrate the effective delay of a receiver relative to a standard unit by measuring the difference in the time-difference data recorded by the two devices when they are driven by a common clock. The common-view method is also used by NIST and by other standards laboratories to provide time and frequency information to users that is directly traceable to national standards. Although a number of different signals have been observed in common view over the years (LORAN, GLONASS, TV line 10, …), the signals from the GPS satellites are currently the most widely used choice for applications requiring the highest levels of accuracy.

The techniques for observing the signals from the GPS satellites in common view were developed in the 1970s and 1980s, when only relatively primitive (by current standards), single-channel receivers were available. In addition, communications channels were relatively slow, and memory was expensive. The data acquisition algorithms and the formats used to transmit and store time difference data were designed with these limitations in mind. Specifically, the method recognized that the receiver could track only one satellite at any time, that it might require up to 2 minutes to lock onto the satellite signal, and that only a relatively small amount of data could be stored and transmitted. These limitations are far less severe today, and it is no longer necessary to continue to use an algorithm whose design was determined to a great extent by the hardware and processing environment of 25 years ago. In particular, the purpose of this paper is to suggest that the decision to use a 13-minute track length, the design of the averaging algorithm used to estimate the time difference from the raw measurements, and the principles of the tracking schedule itself all should be reconsidered to take advantage of the increases in receiver performance and in the storage capacity of present-day, ordinary computers.

THE CURRENT ALGORITHM

In this section I will briefly describe the important features of the current algorithm that is used to acquire time difference data. The details of the method have been discussed in the literature **[1]**.

In code-based receivers, the fundamental datum is the time difference between the local clock connected to the receiver and the time transmitted by a GPS satellite. (The satellite time is derived from the code correlation processor with additional information from the transmitted data stream.) Since most timing laboratories use 1 Hz pulses to drive clock displays and to control other hardware, receivers generally provide a measurement of this time difference once per second using the 1 Hz ticks that are derived from the local frequency reference.

This basic time difference must be corrected for a number of effects, such as the time it took the signal to travel from the satellite to the receiver. In order to compute these corrections, the receiver parses the ephemeris message transmitted by the satellite. Since it takes 12.5 minutes to transmit the full ephemeris message, the length of the track was set to 13 minutes to be sure that the receiver had a complete copy of the latest satellite ephemeris, and also to ensure that all of the receivers that were participating in any common-view campaign were using the same ephemeris, including the most recent estimate of the correction for the ionosphere.

Since the processors in the early receivers were not fast enough to process the corrections from the ephemeris message and to compute the code cross-correlation simultaneously, these receivers fit quadratic functions of the time to 52 consecutive, non-overlapping 15 s blocks of the 1 s raw data using standard least-squares. Each one of these functions was evaluated at a time corresponding to the mid-point of the data block used to compute it, and only the resulting 52 midpoint values were carried forward to the next step.

The geometric delay and any other corrections (ionosphere, eccentricity, …) were applied to these 52 values. Since these corrections can be approximated by a quadratic function of the time for short time intervals, the rationale for this procedure was that the initial quadratic fit would approximate these corrections over each 15 s interval, and the more exact corrections would be applied to a smaller, more manageable data set of only 52 points instead of to the full data set of 780 points.

The final step is to compute a linear fit to the 52 values computed in the previous step and to output the estimate of the fitting procedure at the midpoint of this full 13-minute time span. The primary rationale for incorporating this fit is that it would compensate for the time dispersion in the data caused by the frequency offset of the local clock acting during the 13-minute track time. Both fits also provided some isolation against the effects of short-term glitches in the measurements, although this was not their primary function. (Simple least-squares fits only attenuate glitches.)

THE TRACKING SCHEDULE

Since early receivers could track only a single satellite at any one time, the cooperation that is implicit in the common-view method could be realized only if all stations followed a common tracking schedule, and the BIPM has been publishing such schedules for many years. Each schedule is optimized for a given region, and lists the observation times for each satellite to be used by all of the stations in that region. The observation times advance by 4 minutes every day so that the geometrical relationship between each satellite and each receiving station is almost the same from one day to the next.

The single-channel receivers require up to 2 minutes to lock onto the signal from the satellite, and some additional time after the track is finished for housekeeping. The tracking schedules are implemented on a 16-minute grid for this reason (2 minutes to lock onto the signal + 13 minutes of track time + 1 minute for processing). Since a GPS “day” is only 23h 56min long, it can accommodate only 89 full-length tracks. If all of the possible 89 tracks are assigned, the combination of the 16-minute grid and the 4-minute advance every day results in a tracking schedule that has a simple, 4-day periodicity.

Tracking schedules are less important for multi-channel receivers that can usually track all of the satellites that are in view at any instant. However, as I will show below, an important reason for reconsidering the design of the tracking schedule is that the 4-minute advance results in an undesirable side effect in the way that it converts a time-varying multipath offset to a nearly constant, difficult to evaluate, systematic error. This design may have been unavoidable for a single-channel receiver, but we should consider alternatives that can remain compatible with the design of older receivers, while at the same time providing better handling of these multipath effects.

ADVANTAGES OF NEWER RECEIVERS

In addition to being able to track several satellites at once, newer receivers are fast enough to be able to compute and apply the various corrections derived from the ephemeris message to the raw time difference data in near real time. Once these corrections have been applied, the data are normally very well characterized as white phase noise over a time interval of tens of seconds, which means that there is no reason and no advantage to computing a quadratic fit to consecutive 15 s blocks of data. A simple average would be just as good, and adding a median filter to reject single-point outliers could be even more robust. Furthermore, to the extent to which the data have flicker or random-walk components that cannot be characterized as white phase noise over a 15 s block, the interaction of the quadratic fit with the actual fluctuations in the data results in an estimate that has an unknown, time-varying bias, which might or might not be common to all of the stations participating in the common-view measurements. For example, the current algorithm was effective in attenuating the flicker-like clock dither of selective availability only because all receivers processed these fluctuations in exactly the same way, so that the bias resulting from the interaction between the processing algorithm and the clock dither canceled in the common-view time differences. If the receiver did not record a complete track (not an "exact common view" in BIPM parlance), the estimate from that track could not be used because the bias due to the clock dither no longer canceled in the common-view difference.

The same argument holds for the second linear fit over the entire 13-minute track. It will always produce a result in a formal sense, but the result will be unbiased only if the data can in fact be characterized as having a simple, constant frequency offset over the 13 minute track time. This is a more serious problem in practice, since multi-path effects can be significant over the 13-minute track time, and they are rarely well-characterized by a simple constant frequency offset.

MEASUREMENTS AND DISCUSSION

In order to estimate the significance of these effects, we operated a number of multi-channel receivers at NIST in Boulder, Colorado and at the NIST radio station in Fort Collins, Colorado. The receivers could track eight satellites simultaneously, and we recorded the time differences every second between a local clock and GPS time using data from all of the satellites that were in view at that time. In all cases, we used the parameters in the broadcast ephemeris to correct the data for the geometrical path delay, for the offset of the satellite clock from GPS system time, and for the ionosphere.

In the following figures, we display the data in a "stacked" format. That is, each plot shows the data from the complete pass of a single satellite on several consecutive days as a function of UTC. The data for consecutive days are vertically offset for clarity, and the time tags are advanced by 4 minutes each day to mimic the corresponding advance in the tracking schedules. The symbol "!--!" on each plot shows the duration of a 13-minute track using the appropriate horizontal scale for that figure.

Figures 1 and 2 show these stacked time differences between UTC (NIST) and GPS time recorded using data from SV 11 and 19, respectively. The data obtained using other satellites are qualitatively the same, although different in detail. The data in the figures show that much of the structure repeats from day to day either with no change at all or with a slow evolution. These fluctuations are almost certainly due to multipath reflections and similar effects; although the receivers are sensitive to temperature, these fluctuations are too large and too rapid to be due to this cause.

Since the time tags for the data for each day are advanced by 4 minutes relative to the data from the previous day, the BIPM track schedule algorithm will select a vertical, 13-minute-wide slice through these data. The UTC time of the slice is the same at all stations, but the position of the satellite in the sky relative to the antenna at the site will vary from station to station, and stations that are far apart will tend to find a common view using only tracks that are near the left and right edges of the plot. If I apply the BIPM algorithm to consecutive 13-minute vertical slices of these data, the estimate that I get depends on which portion of the data that I use, and the answers can differ by up to 52 ns. If I exclude the first and last three tracks, which have larger fluctuations than the rest of the data, the variation is reduced to 35 ns.

Since the 4-minute advance used to stack the data from consecutive days mirrors the BIPM track advance, the BIPM algorithm takes a snapshot of these fluctuations and converts them into a nearly constant offset that is a characteristic of exactly which part of the total satellite pass is being used. This procedure results in four problems: (1) Although a multipath reflection always arrives after the direct signal, the effect of a multipath reflection on the measured time difference can have either sign depending on the design of the receiver. Therefore, even when the full data set has been recorded, it is not clear what the "right" time difference is. (2) Since the fluctuations in the stacked data evolve slowly in time, the effect is to insert a time-varying bias into the data that can be difficult to distinguish from the flicker noise of the reference clocks themselves. (3) Different common-view pairs of stations will use different parts of the data set, and common-view pairs that use data near the edges of the track will have an effective delay that is different from common-view pairs near the center of the track. (4) Both of these effects are nearly constant in time, and there is no way to estimate the impact of these fluctuations on any particular track using the data from a standard receiver.

Figure 3 shows the data acquired at the NIST radio station in Fort Collins, Colorado from the same type of receiver. The site has fewer reflectors near the antenna, and the variation, omitting the first and last three tracks, is only 17 ns.

Finally, Figure 4 shows a short-baseline common-view experiment between two identical receivers connected to a common clock. The receivers used independent antennas that were about 1 m apart on the roof of the building. This is the type of experiment that is used at NIST (and elsewhere) to calibrate the effective delay of a receiver in terms of a standard device. Even though the antennas are quite close together, the “calibration constant” derived from these data will vary by nearly 10 ns depending on which part of the data we choose to analyze, and the same concerns that I mentioned above are still relevant. To further complicate this problem, the receiver that has just been calibrated in the multipath environment of NIST usually will be operated at another site with a very different multipath environment.

We can improve matters somewhat by using choke-ring antennas, which attenuate multipath signals to some extent, but the improvement is not more than a decrease in the amplitude of the fluctuations by a factor of about 3 or 4.

THE PROPOSED NEW ALGORITHM

No averaging algorithm can completely remove the effects of temperature fluctuations or multipath reflections; only better receiver hardware will fix these problems. Although we cannot completely eliminate these effects from our time-difference data, the first step is to develop techniques that can provide robust estimates of the sizes of these problems, and the current methods do not satisfy this requirement. In particular, it is extremely unlikely that the current algorithm is consistent with the precision (and with the implied accuracy) of the current BIPM Circular T, which reports time differences between laboratories with a resolution of 0.1 ns.

Any new algorithm must be as compatible as possible with the firmware in existing receivers. In addition, the new algorithm must be a balance between providing a diagnostic estimate of these slowly varying systematic effects, which would favor as little processing as possible in the receiver, and minimizing the amount of data that must be transmitted and stored, which would favor as much pre-processing as possible. I propose computing common-view time differences using simple 15 s averages of the basic 1 s time-difference measurements. These averages might (but need not) include an additional median filter as an outlier detector. In the following sections I will discuss the advantages and potential difficulties of changing to this type of data.

COMPATIBILITY WITH SINGLE-CHANNEL RECEIVERS

Although single-channel “NBS-type” receivers are old and are increasingly difficult to maintain, many timing laboratories still use them, and any new proposal must be consistent with the averaging algorithms and data formats that are incorporated into the firmware of these receivers. A scheme that was based on transmitting and storing 15 s data blocks from those receivers that could support this capability would be nearly completely compatible with the data sets from the older receivers, since the 15 s data blocks could be combined to form a 13-minute track using the procedure outlined in the technical directives. The result would not be identical to the procedure specified in the technical directives, since the rapid data would use a simple average to compute each 15 s datum, while the older receivers used a quadratic least-squares fit. The difference is likely to be quite small over this short time interval, since the data are pretty well characterized as white phase noise over periods of a few seconds.

At least in principle, it is probably not too difficult to modify the firmware of the older receivers to transmit the intermediate 15-second averages, since these values are already present inside the receiver and no new processing would be required. This would be desirable, but need not be a requirement for adopting this proposal.

COMPATIBILITY WITH OTHER TYPES OF RECEIVERS

There have been a number of proposals and experiments to include data from other types of receivers into the BIPM data sets. The most common proposal would add data from carrier-phase receivers, which have been postprocessed to convert the observations to the standard format. Since these receivers do not produce a time-difference measurement every second, it is impossible in principle to process these data in a manner that is completely consistent with the technical directives. Most of these receivers report a measurement every 30 s, and the value that is reported is often an average value over the measurement interval. The difference between this simple 30 s average and the quadratic least-squares fit to two 15 s blocks will depend on a detailed evaluation of the underlying noise of the data. The difference is likely to be negligibly small if the raw data can be characterized as consisting predominantly of white phase noise, but the difference will be a slowly varying offset in the cases we have discussed, since the two methods average the multipath variations in different ways. (Short-term experiments may not be sensitive to this problem, since the offset varies only slowly.) To further complicate this problem, code-based receivers and those that measure the phase of the carrier may have very different multipath characteristics.

This time-varying offset could be reduced if we had 15 s averages from the code-based receivers, since it would then be a simple matter to use them to match the output of the carrier-phase receivers. This is an important advantage, since data from carrier-phase receivers are likely to be become more common in the near future.

STORAGE AND TRANSMISSION REQUIREMENTS

The bandwidth required to transmit these more rapid estimates and the memory capacity required to store them are both quite modest by current standards. For example, suppose that we could characterize each 15 s time-difference observation using 10 characters. Since there are four measurements per minute and 1,440 minutes in a day, a single receiver would produce 57,600 characters per day. If the communication link that transmitted these data operated at an effective speed of only 1,000 characters per second (about 9,600 baud), it would take less than 1 minute to transmit these data. This time is less than 0.1% of the capacity of this channel.

The storage capacity required for these data is equally modest. For example, suppose that the BIPM received these more rapid time-difference data from 50 receivers every day. The total data from all of the receivers would be about 2.8 megabytes. Even if housekeeping and ancillary data increased this size by 10%, the result would still be small by current standards. Even a modest-sized PC has a 30-gigabyte disk, and, at this rate, a disk of this capacity could hold the time-difference data from all of these receivers for almost 30 years. The size of the data set is trivial on this scale, even if the simple estimate I have presented is wrong by a factor of 2.

SUMMARY AND CONCLUSIONS

The averaging algorithm specified by the BIPM for acquiring common-view GPS data, and the design of the tracking schedule used by timing laboratories convert non-random fluctuations with periods of a few minutes into slowly varying systematic offsets that are difficult to estimate from the data obtained using the standard 13-minute tracks. Shorter, more frequent tracks cannot completely fix this problem, but they can provide a much more realistic estimate of its magnitude.

In this paper, I have suggested that transmitting and storing simple averages of 15 consecutive 1 s time-difference measurements is not difficult to realize using standard communications channels and computer disks. These data would provide a much more realistic estimate of the effects of systematic errors such as multipath reflections and similar effects that have a diurnal (or nearly diurnal) variation. Finally, these more rapid averages would simplify including data from carrier-phase receivers, which usually output a measurement every 30 s. In principle, these data cannot be processed according to the technical directives at present, with the result that data from carrier-phase receivers are likely to have a slowly varying time offset compared to data acquired using the typical code-based hardware.

As I have shown, the design of the tracking schedule converts relatively rapid time-difference fluctuations into slowly varying time offsets, and this can degrade the long-term stability of TAI. In addition, it can also have an impact on short-baseline common views, which are widely used to calibrate timing receivers in terms of a standard device. We propose to evaluate these effects in the coming months, and we invite other laboratories to join us in this work.

The exact format used to transmit these extra data values is not a critical aspect of the proposal. One possibility might be to append these 52 additional values to the end of the text that reports the values for the 13-minute track parameters. The current format specifications already contain provisions for additional data at the end of the required parameters, so that adding these extra values would not cause a serious problem for software that is not prepared to accept them.

REFERENCES

[1] D. W. Allan and C. Thomas, 1994, *"International Report: Technical Directives for Standardization of GPS Time Receiver Software to be Implemented for Improving the Accuracy of GPS Common-view Time Transfer,"* **Metrologia, 31,** 69-79.

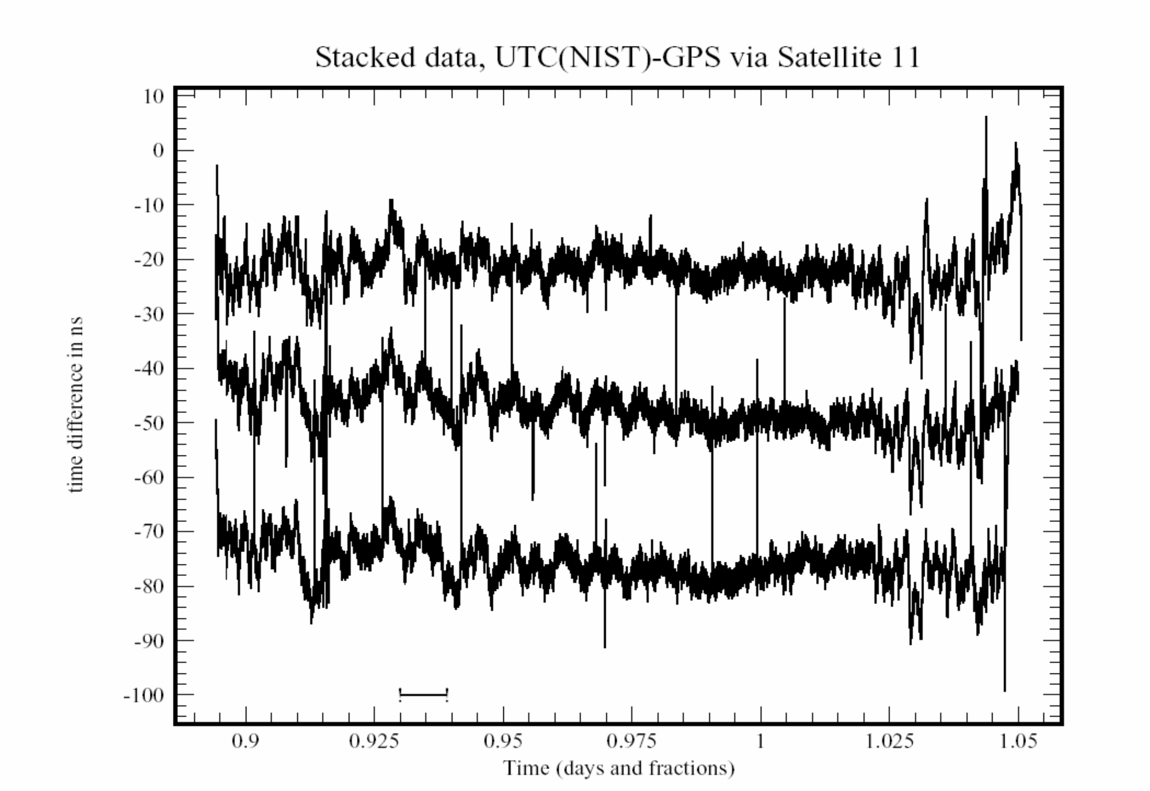

Figure 1. UTC (NIST) – GPS time using 1 s data from Satellite 11 only. Each trace shows the data obtained during a pass of the satellite from horizon to horizon. The three traces show data from 3 consecutive days. Each trace is offset vertically by 30 ns for clarity and the time tags are advanced 4 minutes relative to the data from the previous day. The symbol !--! corresponds to a time interval of 13 minutes.

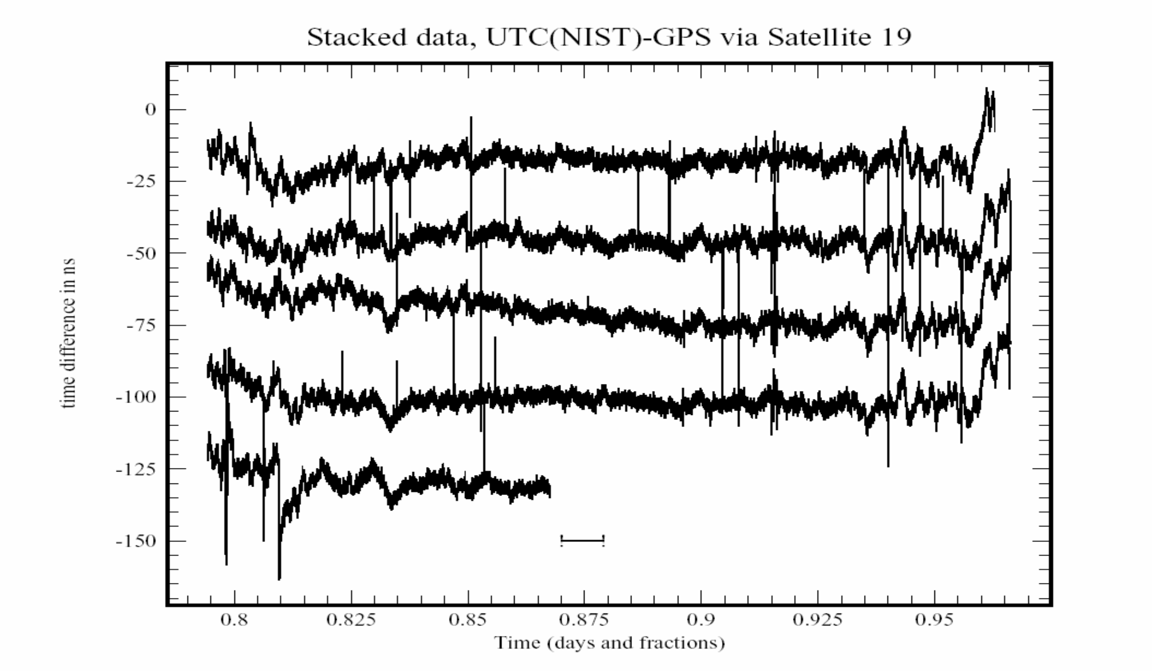

Figure 2. Same as Figure 1, except that the plot shows the data from SV 19.

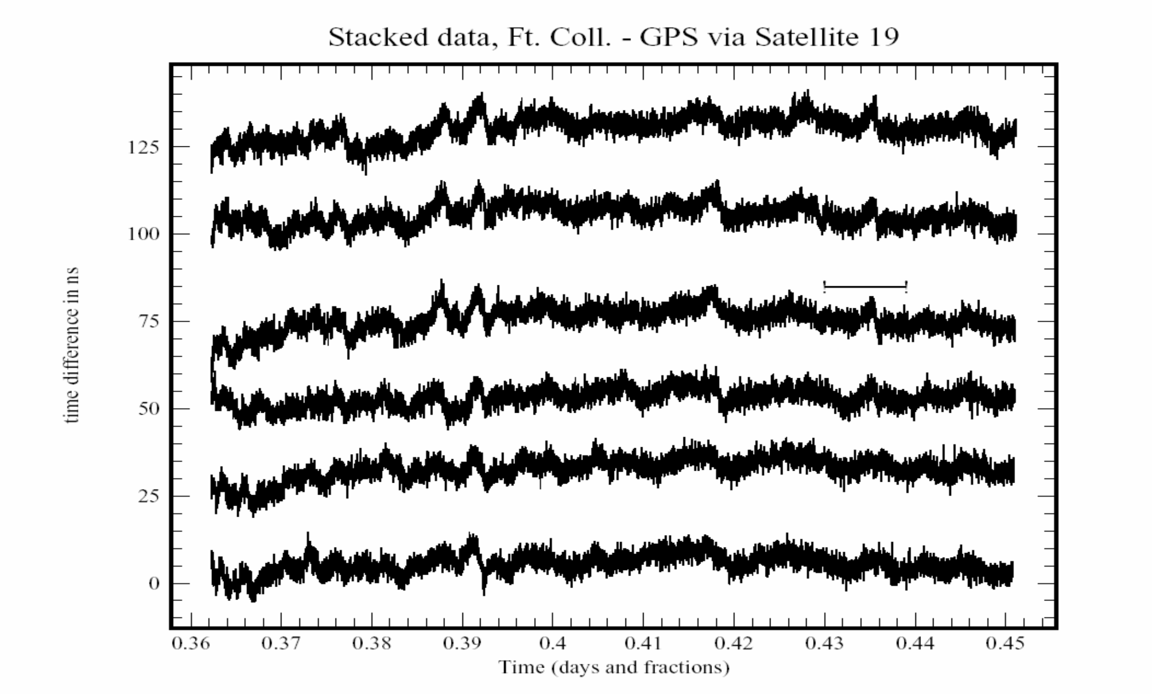

Figure 3. Same data as in Figure 2 except that the data were acquired at the NIST radio station in Fort Collins, Colorado, and the reference is a commercial cesium standard and not UTC (NIST).

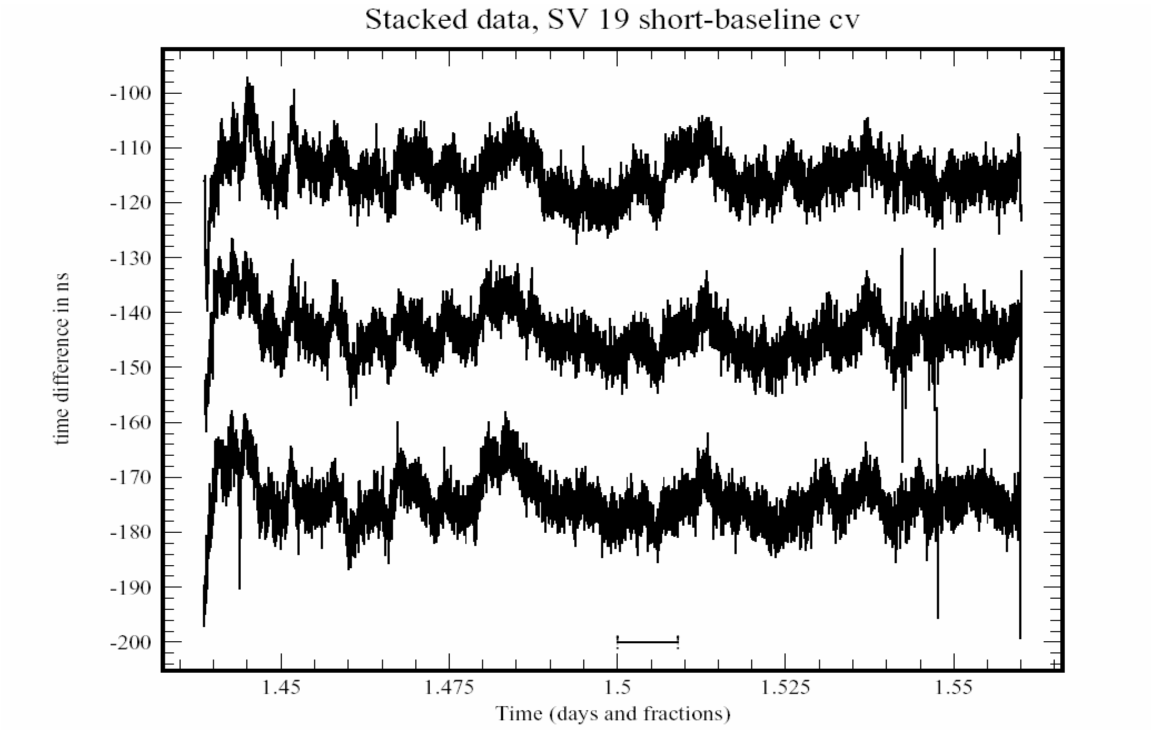

Figure 4. Short-base line common-view time differences between two receivers whose antennas are about 1 m apart. The reference for both receivers was derived from UTC (NIST). The plots are offset vertically for clarity and each trace has time tags that are advanced 4 minutes with respect to the previous trace as in the previous figures.

QUESTIONS AND ANSWERS

DAVID HOWE (National Institute of Standards and Technology): That is a very revealing talk, Judah. Let me just add a couple of things. One, that what a quadratic does is it summarizes all the data over that interval in a couple of coefficients. Said differently, though, the bandwidth, if you will, of the original data is reduced to one over that period, which means that the spectrum at high frequency, which is due to multi-path, which is the phase and amplitude, is lost in that process. Would it be possible to pick up the phase and amplitude terms at a higher frequency and somehow use that? Because it introduces a bias in the result.

JUDAH LEVINE: Yes. What you are implicitly pushing towards is recording data even faster. Now, if I were working in a room without any windows, I might suggest that. But my guess is if I posed that to the time laboratories, they would say "Oh, yeah? Never. We've got enough trouble dealing with what we've got enough trouble with." So my proposal was the 15 second, which I saw as something of a compromise.

What you are suggesting is that all things being equal, the faster the better. I agree with that, that is true. It is simply a logistical compromise.

WLODZIMIERZ LEWANDOWSKI (Bureau International des Poids et Mesures): I have a couple comments. Concerning the systematic effects that you have shown, at least a part of it is due to temperature somewhere. So what we advocate, at least for some cases, is to stabilize some hardware with temperature. That can reduce at least some of the effects you are showing for calibration issues mainly.

LEVINE: Remember that the picture that I have showed you has many, many wiggles in it, and it is only a few hours long. So it is not temperature. There may be some longer-term temperature effect.

LEWANDOWSKI: One could see some diurnal effects. But I would not argue on this.

LEVINE: The temperature is going to be one cycle per day or two cycles per day, or something like that. But, this is many, many wiggles in a few hours. So I would bet you a nickel that it is multi-path. I would bet you a dime that it is multi-path. It is not temperature.

LEWANDOWSKI: But we observed some of these journeys, or some of our systematics, also from one day to another, with other sets of data, and we have shown that it was due to temperature. I am sure it was another set of data.

But the most important comment that I would like to make is concerning the resolution of Circular T, and you are right that BIPM is producing Circular T with 0.1-nanosecond resolution for UTC minus UTC (k), and with this kind of data how it can be. So I will explain why we switched from one nanosecond to 0.1. This is due to Two-Way technique; this was one reason, because we had a sudden resolution wash away. And the second reason was the excellent quality of UTC (USNO). For some Circular T's, we had differences between UTC and UTC (USNO), almost stable during the period of 1 month; and it means that it was 3 nanosecond, 3 nanosecond – we could not distinguish values from the period of 5 days. So we had to move to smaller resolution to see the variation in UTC (USNO). So it was a second motivation.

And, at BIPM, we are discussing what to do with this, because we are aware that we cannot provide 0.1 to

Civil Code GPS data, because it is moving by nanoseconds. One point of view was that we should concentrate just to put 0.1 to laboratories that could do Two-Way. But the point of view prevailed that we should not make this distinction and we should apply to all laboratories. So this is the way we are doing this.

LEVINE: Right. And I was not really suggesting that we go back to 1-nanosecond resolution. I wasn't suggesting that. We could talk about USNO, but I think we perhaps should do it privately because I think the stability that you see with respect to USNO is misleadingly good. And the reason is that USNO has such a heavy weight in the calculation of UTC that they are going to go up and down together. Because USNO contributes like 40 percent of UTC, so there is a strong correlation there. And you would expect – in fact, if UTC and UTC USNO move apart, then that is a very strange effect because of the large weight that USNO has in UTC. But that is something that has nothing to do with the current discussion.

DEMETRIOS MATSAKIS (U.S. Naval Observatory): UTC (NIST) is also very stable. We did have a period I call the "Forlorn Period" when all our Circular T values were coming out minus four, no matter what we did. But the truth is that USNO in Washington now contributes about 30 percent of the weight to TAI, so any fluctuation of the average of our clocks will be forgiven at the rate of 30 percent. In other words, thirty percent of any fluctuation of the average will be absorbed into TAI; the rest will show in UTC – UTC (USNO). That is absolutely the case.

However, we look at our clocks in comparison with each other, and it does not seem that time transfer noise is dominating, even though all the clocks are in Washington. Every laboratory, and all of France, too, has a certain fraction of their average fluctuations forgiven because of this, even if their clocks are not physically collocated.

That was not what I was going to talk about when I was going to speak earlier, though. I was just going to point out also that the USNO had a problem just like you showed with our GPS receivers jumping 10 to 20 nanoseconds relative to each other, in worse case common view, as a function of sky angle. This was many years ago. And our solution was to build a structure that was about 12 feet tall so that our antennas were above everything else on our roof. And once we did that, a lot of the multi-path issues went down.

I do have one question for you. I noticed you did not produce any data estimating how different the results would be with a fit over a 15-minute track from just a straight average. Do you have any numbers for that?

LEVINE: I don't have those numbers, but it is an easy thing to do. The answer is that I did not do it. I did not do it because I didn't think of it. But I got 100 megabytes of data, I have lots of data on it, I can always do it. If anybody is interested, just ask for it and I will do it when I get home. But I do not know the number now.

MATSAKIS: Two-Way people do the same thing. They do a parabolic fit when they want to compare their 2-minute integrations. There was an old controversy, should they just average 1-second points and forget about the fit. At the USNO, we truly do it both ways and difference them, and the result is negligible.

LEVINE: The answer is that I have not done it, but all the stuff is there.