Evaluating the Frequency and Time Uncertainty of GPS Disciplined Oscillators and Clocks

Michael A. Lombardi

Time and Frequency Division, National Institute of Standards and Technology, Boulder, Colorado, USA

Abstract: Global Positioning System (GPS) disciplined oscillators and clocks serve as standards of frequency and time in numerous calibration and metrology laboratories. They also serve as frequency and time references in many industries, perhaps most notably in the telecommunication, electric power, transportation, and financial sectors. These devices are inherently accurate sources of both frequency and time because they are adjusted via the GPS satellites to agree with the Coordinated Universal Time (UTC) time scale maintained by the United States Naval Observatory (USNO). Despite their excellent performance, it can be difficult to evaluate their uncertainty, and even more difficult for metrologists to prove their claims of uncertainty and traceability to skeptical laboratory assessors. This article is written for metrologists and laboratory assessors who work with GPS disciplined oscillators (GPSDOs) or GPS disciplined clocks (GPSDCs) and need to assess their uncertainty. It describes the relationship between GPS time and Coordinated Universal Time (UTC), explains why GPS time is traceable to the International System (SI), and provides methods for evaluating the frequency and time uncertainty of signals produced by a GPSDO or GPSDC.

1. Introduction

Accurate measurements of frequency and time involve comparisons of electrical signals generated by oscillators and clocks. For the purposes of this article, an oscillator is an instrument that generates electrical signals at a specific frequency, typically in the form of a sine wave. A clock is an instrument that generates an electrical pulse, typically at a frequency of one pulse per second (pps), that is synchronized as closely as possible to Coordinated Universal Time (UTC). There are numerous types of oscillators and clocks, but by definition, only those whose frequency is referenced to the cesium atom are currently defined as primary frequency and time standards. Since 1967, the base unit of time, the International System (SI) second (s), has been defined as the duration of 9,192,631,770 energy transitions of the cesium atom [1]. The SI unit of frequency, the hertz (Hz), is defined as one cycle per second.

Cesium standards reside in many calibration and metrology laboratories and are necessary for applications that require an autonomous source of frequency and time, which means a source that can generate accurate and stable signals without receiving input from another source. Oscillators and clocks that are disciplined by signals from Global Positioning System (GPS) satellites are not autonomous sources of frequency or time; they require a receiver and antenna to receive signals from the GPS satellites and their accuracy will quickly degrade if GPS cannot be received. However, for both economic and technical reasons, the use of GPS disciplined devices as both reference standards and working standards is a sensible choice for most laboratories. Cesium devices are far more expensive than GPS devices, typically by a factor of 10, and have a relatively short life expectancy [2], often less than 10 years. It is simply not practical to install a cesium standard at every location where accurate frequency and time are required.

GPS disciplined devices also have performance advantages over cesium devices. For example, when measured over long periods such as multiple days and weeks, a GPS disciplined oscillator (GPSDO) is typically more stable and accurate than a cesium oscillator because its frequency is continuously adjusted by signals broadcast from satellites [3]. In addition, a GPS disciplined clock (GPSDC) can synchronize itself by decoding messages transmitted by the satellites. Cesium clocks cannot synchronize themselves; their 1 pps output must be synchronized to another source. In fact, the synchronization of a cesium clock is usually performed by using a GPS clock as a reference.

The widespread use of GPSDOs and GPSDCs as frequency and time standards has led to some concern among laboratory assessors, who often view these devices as untraceable and uncalibrated "black boxes." These concerns are understandable because before an assessor determines whether or not a laboratory's measurements are traceable, they must have information available about the uncertainty of the laboratory's calibration chain back to the SI. If the laboratory operates a GPS device as its standard, there may be no historical measurement data for laboratory personnel or an assessor to analyze, other than the manufacturer's specifications. The lack of calibration data, as well as a general lack of understanding of how GPS

Address correspondence to Michael A. Lombardi, Time and Frequency Division, National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80305, USA. E-mail: lombardi@nist.gov

DOI:10.1080/19315775.2017.1316696

devices work, sometimes results in a laboratory's calibration and measurement capabilities being called into question, or either over or underestimated.

This article was written to address the concerns of both metrologists and laboratory assessors by describing methods to evaluate the frequency and time uncertainty of GPSDOs and GPSDCs. It begins by describing the relationship between the signals transmitted by GPS satellites and three time scales; Coordinated Universal Time (UTC); UTC(USNO), the UTC time scale maintained at the United States Naval Observatory; and UTC(NIST), the UTC time scale maintained at the National Institute of Standards and Technology.

2. GPS Time and Its Relationship to UTC, UTC(USNO), and UTC(NIST)

UTC, computed monthly by the Bureau International des Poids et Mesures (BIPM) in France, is the official world time scale. As of 2016, about 70 laboratories in more than 50 nations participate in the calculation of UTC by sending data from their time scales and clocks to the BIPM. The data are collected from local clock time difference measurements made in the various laboratories, and from international time scale comparisons conducted via satellite links. Each month, the BIPM publishes their measurement results in a document called the *Circular T* (available at www.bipm.org). This document lists the time differences between each participating time scale; known as UTC(k) where k is the acronym that designates the laboratory, and UTC itself [4]. Because one month is a long period to wait for updated UTC measurements, an unofficial version of UTC, known as "Rapid UTC" or UTCr, has been published weekly since 2013 [5].

The UTC calculations are extensive and require the BIPM to collect and average data from hundreds of clocks that contribute to the various UTC(k) time scales, and also from the primary frequency standards maintained by a few national metrology institutes. The calculations result in the best available realization of the SI second, making UTC the ultimate reference for all frequency and time measurements. However, calibration laboratories cannot directly access UTC because it is a post processed, virtual time scale that does not generate physical signals. Fortunately, the UTC(k) time scales listed on the *Circular T* operate in real time and do generate physical signals that closely agree with the UTC calculation, in many cases to within a few nanoseconds. These laboratories are participants in a BIPM key comparison, currently designated as CTF-K001. UTC, that serves as an internationally accepted method for establishing traceability to the SI. There is no separate key comparison for frequency, thus the key comparison participants have established traceability to both the second and the hertz [6].

GPS time is referenced to UTC(USNO), a time scale that not only appears on the *Circular T* [6], but that also, due to the large number of clocks that the USNO maintains, is the single largest contributor to the weighted average used by the UTC calculation. Thus, a traceability chain from GPS time to the SI second is always intact. The difference between GPS time and UTC(USNO) consists of two corrections, a large integer second correction and a small nanosecond-level correction. The integer second correction is needed to correct GPS time for leap seconds and is equal to the number of leap seconds that have been inserted into UTC since January 6, 1980, the day that the GPS time scale originated. The nanosecond-level correction compensates for the most recent measured difference between GPS time and UTC(USNO). The corrections are included in subframe 4 of the navigation message broadcast by the satellites [7] and nearly all GPS receivers apply both corrections by default. Thus, every GPSDO and GPSDC produces signals that agree with UTC(USNO) [8].

The UTC(USNO) time scale is a very close approximation of UTC. This is illustrated in Fig. 1, which shows the differences between the two time scales during the 10-year period from 2006–2015, which were nearly always within ±10 ns. These measurements were obtained from the *Circular T* which provides data at five day intervals. The average time difference during the 10-year period was just 0.5 ns. The uncertainty reported by the BIPM for the comparison never exceeded 4.1 ns ($k = 1$) during the 10-year period and was as small as 1 ns ($k = 1$), thus the $k = 2$ uncertainty (worst case) is about 8 ns. The slope of the blue line can be used to estimate the average daily frequency offset between UTC and UTC(USNO), which was negligible during the 10-year period, just $+1.1 \times 10^{-17}$ (details about how the frequency offset can be obtained from time difference measurements are provided in Section 3).

The UTC(NIST) time scale is also a close approximation of UTC. Figure 2 shows the difference between and UTC and UTC(NIST) from 2006–2015, as reported by the *Circular T*. The time difference between UTC and UTC(NIST) was always within ±20 ns and the average time difference was 4.1 ns. The uncertainty reported by the BIPM for the comparison was near 5 ns ($k = 1$) throughout the entire 10-year period, thus the $k = 2$ uncertainty is approximately 10 ns. The average daily frequency difference between UTC and UTC(NIST) during the 10-year period was also negligible, just $+5.7 \times 10^{-17}$.

The UTC(USNO) and UTC(NIST) time scales are continuously compared to each other using a variety of satellite time transfer methods, and can be considered as equivalent for all metrological and calibration purposes [9]. Figure 3, obtained by subtracting the *Circular T* data sets shown in Figs. 1 and 2, shows the time difference between USNO and NIST for the period from 2006–2015. The time differences always remained within ±25 ns and the average time difference was 2.2 ns. The average daily frequency difference

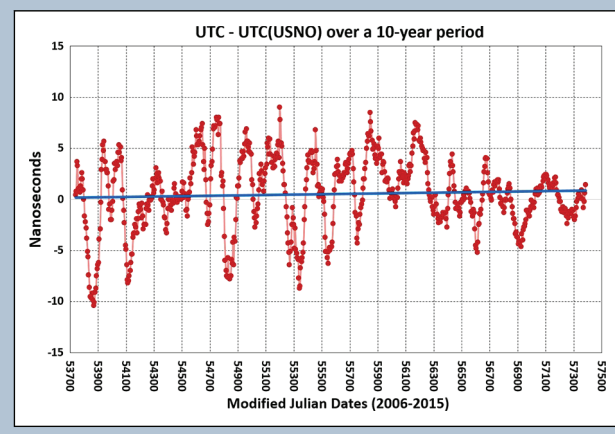

Figure 1. Time differences between UTC and UTC(USNO) from 2006–2015.

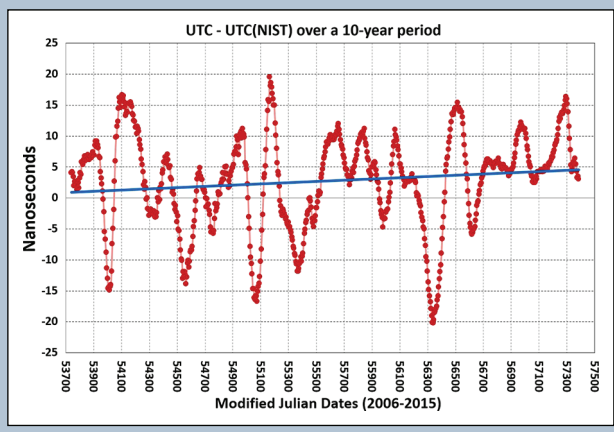

Figure 2. Time differences between UTC and UTC(NIST) from 2006–2015.

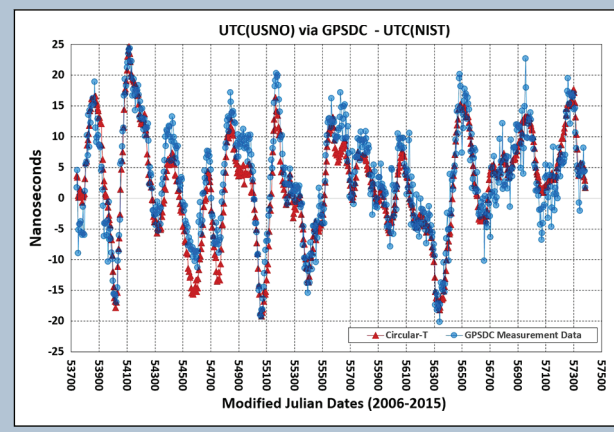

Figure 4. Time differences between UTC(USNO) via GPSDC and UTC(NIST), 2006–2015.

between UTC(USNO) and UTC(NIST) during the 10-year period was again negligible, just $+4.6 \times 10^{-17}$.

When the 1 pps signal from a calibrated GPSDC is compared to UTC(NIST), the results are essentially the same as directly comparing UTC(USNO) to UTC(NIST). Figure 4 overlays a comparison of a GPSDC and UTC(NIST) onto the *Circular T* data shown in Fig. 3. The GPSDC ran continuously at NIST in Boulder, Colorado during the entire 10-year period. To match the interval of the *Circular T*, only one value (a 24-hr average) is shown every five days. The GPSDC measurement has more outliers, but the structure of the data varies only slightly. The time differences remained within ±25 ns and the average time difference was 3.2 ns, differing by only 1 ns from the *Circular T* average. The GPSDC measurement data has an even gentler trend than the *Circular T* comparison and thus the average daily frequency difference between UTC(USNO) via a GPSDC and UTC(USNO) during the 10-year period was just $+1.1 \times 10^{-17}$. These results indicate that the output of a calibrated GPSDC is essentially equivalent to UTC(USNO).

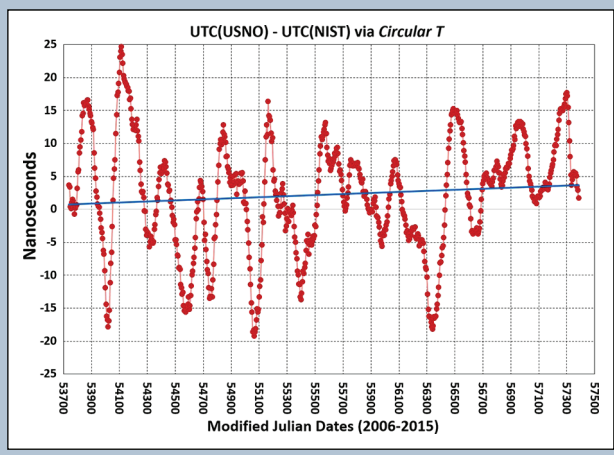

Figure 3. Time differences between UTC(USNO) and UTC(NIST) via *Circular T* from 2006–2015.

3. Estimating Frequency Offset

The graphs shown in Figs. 1–4 are known as phase or time difference graphs and are presented in the standard Cartesian *x*/*y* format. The *x*-values indicate when the measurements were recorded, and the *x*-axis units indicate the measurement period. The *y*-values represent the change in phase, ΔΦ, between the two electrical signals that are being compared. The phase changes are usually measured with an instrument, in this case a time interval counter, that records measurements in units of time rather than in radians or degrees. Thus, the *y*-axes in Figs. 1–4 are labeled to show the change in time, or Δt.

The difference in frequency, or frequency offset, between two signals can be estimated in the frequency domain as

$$f_{off} = \frac{f_{meas} - f_{nom}}{f_{nom}}, \tag{1}$$

where f_{off} is the frequency offset, f_{meas} is the frequency reported by the measurement (for example, a reading obtained from a frequency counter), and f_{nom} is the nominal frequency in hertz that the oscillator would ideally produce. The nominal frequency (10 MHz, for example) is either listed in the oscillator's specifications or found on a label next to the output connector. Unlike f_{meas}, which is a measured value with an associated measurement uncertainty, f_{nom} has no uncertainty. It is a defined (theoretical) value that indicates the frequency the oscillator would ideally generate.

The nominal frequency is included in both the numerator and the denominator of Eq. (1). Thus, the unit (hertz) cancels and f_{off} is a unitless value. It can be converted to units of hertz by multiplying by the nominal frequency. For example, if f_{off} is 1×10^{-7} the frequency offset in hertz for an oscillator with a nominal frequency of 10 MHz is (1×10^{-7}) (1×10^{7}), or 1 Hz. Equation (1) is often simplified as

$$f_{off} = \frac{\Delta f}{f}, \tag{2}$$

where f_{off} is the unitless frequency offset, Δf is the difference between the measured and nominal frequency in hertz, and f is the nominal frequency in hertz.

The data in Figs. 1–4 were not obtained in the frequency domain, but instead were collected in the time domain by measuring the time difference between two clocks. Even so, they can still be used to obtain the frequency offset, f_{off}. This is because frequency is the reciprocal of period, which is expressed as a time interval. Frequency is defined as

$$f = \frac{1}{T}, \tag{3}$$

where T is the period of a signal in seconds, and f is the frequency in hertz. This definition can also be expressed as

$$f = T^{-1}. \tag{4}$$

By performing mathematical differentiation on the frequency expression with respect to time and substituting in the result, we can show that an estimate of frequency offset obtained in the frequency domain is equivalent to an estimate of frequency offset obtained in the time domain, or that $\Delta f / f$ is equivalent to $-\Delta t / T$ [10]. For example,

$$\Delta f = -T^{-2}\Delta t = -\frac{\Delta t}{T^2} = -\frac{\Delta t}{T} f; \tag{5}$$

therefore,

$$f_{off} = \frac{\Delta f}{f} = -\frac{\Delta t}{T}. \tag{6}$$

Measuring frequency in the time domain requires collecting at least two time interval measurements (TI_1 and TI_2), typically from a time interval counter. The change in time interval, Δt, is equal to $TI_2 - TI_1$ and T is the period that elapsed between the two readings. Thus, f_{off} is obtained with

$$f_{off} = \frac{TI_2 - TI_1}{T} = -\frac{\Delta t}{T}. \tag{7}$$

To keep the sign of f_{off} in agreement with the slope of the phase graph, note that the first reading is subtracted from the second reading.

In practice, multiple time interval readings are usually recorded to estimate the average frequency offset. In that case, T simply becomes the period that elapsed between the first and the last reading. However, Δt is usually not obtained by subtracting the first and last readings; instead it is common practice to fit a linear least squares line to the phase data (as shown in Figs. 1–3). The slope of the least squares line is then used to estimate Δt. Therefore, it is common for metrologists to state that frequency is obtained from the slope of the phase.

Table 1 shows the time offset and frequency offset, f_{off}, of UTC (USNO) and UTC(NIST) with respect to UTC, for UTC(USNO) with respect to UTC(NIST) as obtained from *Circular T* data, and for UTC (USNO) with respect to UTC(NIST) as obtained from a GPSDC measurement. The frequency offset values were obtained from Figs. 1–4 by estimating Δt with a linear least squares line. The average daily frequency offsets between the three time scales are negligible, parts in 10^{17}, when measured over a long interval. These results confirm that GPSDOs and GPSDCs are inherently accurate with respect to UTC, approximately three orders of magnitude more accurate in the long term than a well maintained commercial cesium standard. The frequency accuracy of a free running cesium oscillator with respect to UTC is usually no better than a few parts in 10^{14}, partially because of their inherent stability limitations, but mostly because they are not being corrected by GPS to agree with UTC.

Comparison	Time Offset (ns)	Frequency Offset ($\times 10^{-17}$)	Data Source
UTC – UTC(USNO) from *Circular T*	0.5	+1.1	Fig. 1
UTC – UTC(NIST) from *Circular T*	4.1	+5.7	Fig. 2
UTC(USNO) – UTC (NIST) via *Circular T*	2.2	+4.6	Fig. 3
UTC(USNO) – UTC (NIST) via GPSDC	3.2	+1.1	Fig. 4

Table 1. Frequency and time comparisons between UTC, UTC (USNO), and UTC(NIST) averaged over a 10-year period.

4. Basic Principles of a GPS Disciplined Oscillator

Disciplined oscillators allow accurate frequency signals, controlled by a shared external reference, to be simultaneously generated at multiple sites. A disciplined oscillator has at least three parts: a local oscillator (LO), a receiver that collects data transmitted from a reference source, and a frequency or phase comparator. The comparator measures the phase or time difference between the LO and the reference and converts this difference to a frequency correction that is periodically applied to the LO. By continuously repeating this process, the LO is locked to the reference and can replicate its performance. No manual adjustment of the LO is ever necessary.

GPSDO manufacturers rarely disclose exactly how their products work, but a few basic elements are present in most designs. The LO is usually a quartz oscillator, but more expensive models include an atomic rubidium oscillator. The GPS receiver is nearly always a single-frequency (L1 band, 1575.42 MHz) instrument that decodes the coarse acquisition (C/A) code broadcast by the satellites. The receiver is connected to a small antenna and typically outputs 1 pps or a similar low frequency signal. Various types of phase comparators are used to measure the difference between the signal from the GPS receiver and a signal from the LO. The LO typically has a nominal frequency of 10 MHz, so its signal is divided to a lower frequency (often 1 pps) prior to this phase comparison. The output of the phase comparator is read by a microcontroller (MCU) whose firmware executes a control loop, which is often some variation of a proportional-integral-derivative (PID) controller [11]. The control loop keeps the LO locked to GPS by continually issuing frequency corrections [12, 13]. The correction interval can be either fixed or variable, depending upon the design and time constant of the control loop. If the GPSDO is properly designed, the corrections are applied at intervals that are shorter than the period when the uncorrected LO becomes less stable than GPS. Even if a stable rubidium device is used as the LO, the correction interval is usually less than one hour. If the LO is an inexpensive quartz oscillator the correction interval could be a few seconds or less.

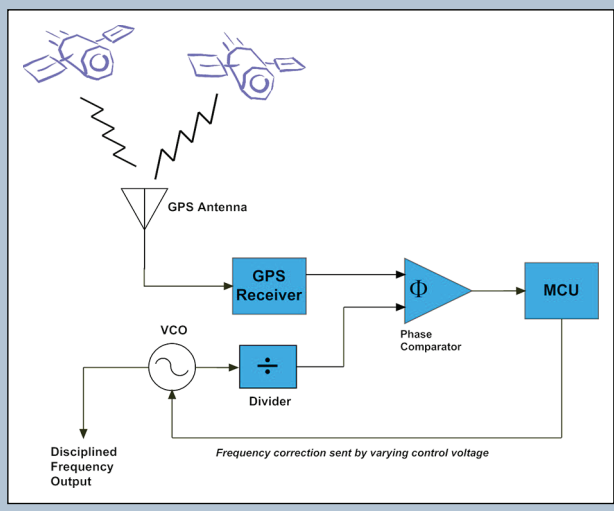

Figure 5. A GPSDO where the local oscillator is adjusted by varying the LO control voltage.

In a simple GPSDO design, the LO might be a voltage controlled oscillator (VCO) and frequency corrections may be sent by varying the control voltage (Fig. 5). If the LO can be digitally controlled, a digital frequency correction is sent either directly to the LO or to a direct digital synthesizer (DDS). When a DDS is utilized, the LO is allowed to free run and the corrections are applied to the output of the synthesizer (Fig. 6). This method has the advantage of allowing very small frequency corrections to be applied without disturbing the LO stability. For example, a DDS with 1 µHz resolution at 10 MHz can instantly apply frequency corrections as small as 1×10^{-13}.

In all GPSDO designs, the device is locked when the phase of the LO has a constant offset relative to the phase of GPS. Ideally, the control loop must be loose enough to ignore the short-term fluctuations of the GPS signals, to reduce the amount of phase noise and to allow the LO to provide reasonably good short-term stability. However, the control loop must also be tight enough to quickly respond to conditions when the device is unlocked, and to track GPS closely enough so that the LO is always accurate and stable in the long term.

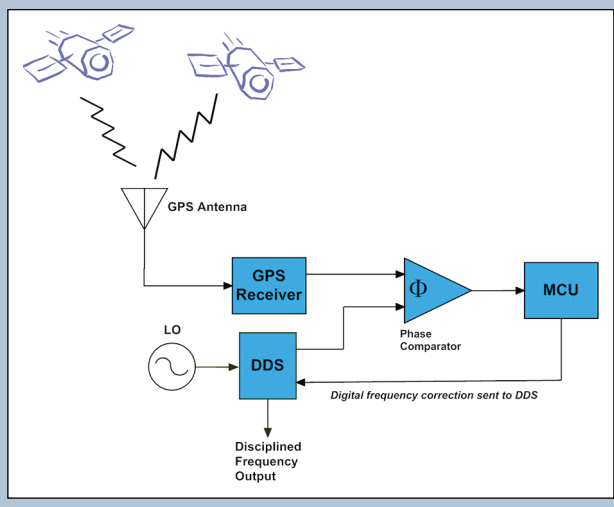

Figure 6. A GPSDO where the output frequency is synthesized.

5. Estimating Frequency Stability

As shown in Sections 2 and 3, the frequency generated by GPS disciplined devices is inherently accurate and any offset in frequency with respect to UTC will be negligible when measured over a long interval. However, metrologists and laboratory assessors are typically only interested in the uncertainty of a GPSDO during the period when it serves as the reference for a calibration. For example, if a calibration lasts for 24 hr, they need to know the uncertainty of a GPSDO over a 24-hr period. Evaluating this uncertainty requires knowing the frequency stability of a GPSDO over that same period. This is because the accuracy of an oscillator over a given interval can never be better than its stability during that same interval.

Frequency stability differs from accuracy. It simply indicates how well an oscillator can produce the same frequency over a given period. Oscillators that can produce the same frequency over the period or interest are stable, regardless of whether their frequency is "right" or "wrong" with respect to its nominal value.

The most common metric for expressing frequency stability is the Allan deviation (ADEV), expressed mathematically as $\sigma_y(\tau)$. ADEV is similar to the classical standard deviation but is better suited for frequency metrology because it works with non-stationary data, meaning data that has a trend and that does not converge to a mean value. Time difference measurements are non-stationary because they have a trend contributed by the frequency offset. The Allan deviation has two other important advantages over standard deviation; it can estimate stability over different periods from a single data set, and it can identify most types of oscillator noise [14–16]. The equation for ADEV using frequency domain measurements and non-overlapping samples is

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(M-1)} \sum_{i=1}^{M-1} (\bar{y}_{i+1} - \bar{y}_i)^2}, \tag{8}$$

where y_i is the ith in a series of M unitless frequency offset measurements averaged over a measurement or sampling interval that is designated as τ. Note that while standard deviation subtracts the mean from each measurement before squaring their summation, ADEV subtracts the previous data point. Stability is a measure of frequency fluctuations and not of frequency offset, and the differencing of successive data points is done to remove the trend contributed by the frequency offset [15]. Also, note that the $\bar{y}$ values in the equation do not refer to the average or mean of the entire data set. They instead imply that it is acceptable for the individual measurements in the data set to be obtained by averaging.

The equation for ADEV using phase (time domain) measurements and non-overlapping samples is

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(N-2)\tau^2} \sum_{i=1}^{N-2} (x_{i+2} - 2x_{i+1} + x_i)^2}, \tag{9}$$

where x_i is the ith in a set of N phase measurements spaced by the measurement interval τ.

Equations (8) and (9) show the original, basic forms of ADEV. However, in practice ADEV is normally implemented with overlapping samples, which not only improves the confidence of a stability estimate but also allows stability to be estimated with all possible combinations of the data set. In the frequency domain, ADEV with overlapping samples is calculated as

$$\sigma_y(\tau) = \sqrt{\frac{1}{2m^2(M-2m+1)} \sum_{j=1}^{M-2m+1} \left\{ \sum_{i=j}^{j+m-1} [y_{i+m} - y_i] \right\}^2}. \quad (10)$$

The equation for ADEV using time domain (phase) measurements and overlapping samples is

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(N-2m)\tau^2} \sum_{i=1}^{N-2m} (x_{i+2m} - 2x_{i+m} + x_i)^2}. \quad (11)$$

The modified Allan deviation, known as MDEV, is also commonly used to estimate frequency stability. It includes some additional phase averaging that improve the results slightly at longer averaging periods. The equation for MDEV using time domain (phase) measurements and overlapping samples is

$$Mod\,\sigma_y(\tau) = \sqrt{\frac{1}{2m^2\tau^2(N-3m+1)} \sum_{j=1}^{N-3m+1} \left\{ \sum_{i=j}^{j+m-1} [x_{i+2m} - 2x_{i+m} + x_i] \right\}^2}. \quad (12)$$

The overlapping versions of ADEV and MDEV add an averaging factor, m, that was not found in the original Allan deviation as shown in Eqs. (8) and (9), but that is found in Eqs. (10), (11), and (12). To understand the averaging factor, consider that τ_0 is the shortest interval at which data are taken. For example, if the frequency or phase of the oscillator was measured every second, then $\tau_0 = 1$ s. To obtain stability estimates for longer intervals, τ_0 is simply multiplied by m, thus $\tau = m\tau_0$. Even though the overlapping samples are not statistically independent, the number of degrees of freedom still increases, thus improving the confidence in the stability estimate [14, 16].

An ADEV graph plots log τ on the x-axis to indicate the averaging period, and log $\sigma_y(\tau)$ on the y-axis to indicate frequency stability. Historically, most ADEV graphs were generated using the octave method, where each successive value of τ is twice as long as the previous value. This was once required to save computational time, but with modern computers it has become common to estimate ADEV for all possible values of τ. This allows estimating the frequency stability at any interval that is a multiple of τ_0, often making it possible to exactly match the period of the calibration where the GPSDO was used as the reference.

Figure 7 shows an "all tau" ADEV graph from a GPSDO that was calibrated at NIST. The device under test was stable to less than 1×10^{-11} at all averaging periods ($\tau_0 = 1$ minute). This indicates the presence of a stable LO, which in this case was an oven controlled quartz oscillator (OCXO). The stability rapidly improves when the stability of GPS surpasses the stability of the OCXO, a condition that is clearly evident when τ exceeds 1,000 s. The bump in the red line when τ is between 100 s and 1,000 s indicates that the correction period when the LO is adjusted to agree with GPS lies somewhere within this range. A close examination of the graph reveals that smaller bumps appear at periods where τ is a multiple of the correction interval, a structure that is common in ADEV plots of GPSDOs. The frequency stability at $\tau = 1$ day, as indicated by the red arrow, is 1.4×10^{-13}.

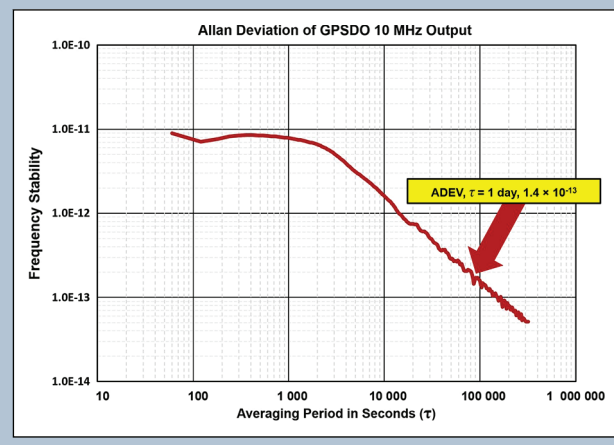

Figure 7. Overlapping ADEV plot ("all-tau") from a GPSDO calibration at NIST.

Taking the slope of the line on ADEV graph can help identify the type of oscillator noise (Fig. 8). Five noise types are commonly discussed in the frequency metrology literature: white phase and flicker phase (both have a slope of τ^{-1} when used with standard ADEV but MDEV can distinguish between the two types of phase noise), white frequency (a slope of $\tau^{-1/2}$), flicker frequency (τ^0, no slope), and random walk frequency (a slope of $\tau^{1/2}$). For both types of phase noise, the stability improves at a rate proportional to the

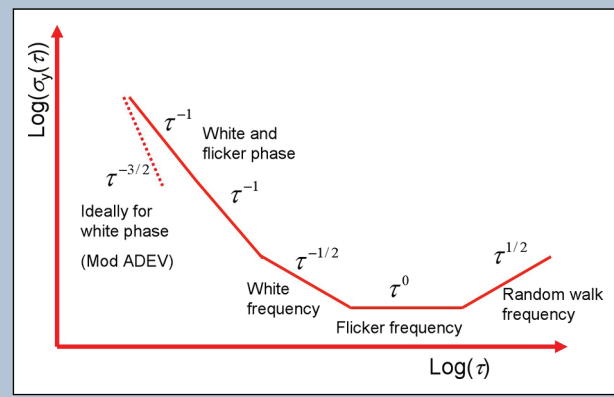

Figure 8. Determining the type of oscillator noise from a frequency stability graph.

GPSDO ID Code	Frequency Stability Estimated with Overlapping Allan Deviation, $\sigma_y(\tau)$				
	GPSDO Local Oscillator Type	1 second	1 minute	1 hour	1 day
SYX	Rubidium	NA	1×10^{-11}	7×10^{-13}	0.5×10^{-13}
ET6	Rubidium	NA	1×10^{-12}	4×10^{-13}	0.6×10^{-13}
ERT	Rubidium	6×10^{-12}	1×10^{-12}	6×10^{-13}	0.7×10^{-13}
AR1	Quartz	6×10^{-12}	1×10^{-11}	1×10^{-12}	0.9×10^{-13}
BR8	Quartz	5×10^{-10}	9×10^{-12}	4×10^{-12}	1×10^{-13}
GD3	Quartz	NA	6×10^{-12}	3×10^{-12}	1×10^{-13}
PT1	Quartz	4×10^{-12}	1×10^{-11}	4×10^{-12}	3×10^{-13}
HPZ	Quartz	1×10^{-12}	2×10^{-12}	1×10^{-12}	4×10^{-13}
TS6	Rubidium	6×10^{-12}	8×10^{-13}	6×10^{-12}	4×10^{-13}
FL9	Rubidium	7×10^{-12}	3×10^{-12}	3×10^{-13}	7×10^{-13}

Table 2. Frequency stability of various GPSDOs at intervals of one day or less.

averaging period. For white frequency noise, the stability is still improving, but the rate of improvement has slowed down and is now proportional to the square root of the averaging period. When the flicker frequency region of an ADEV graph is reached, the oscillator has reached a noise floor that shows its best possible stability (often called the "flicker" floor). When this point is reached, there is nothing to be gained by more averaging. In fact, continuing to average will degrade the stability because the noise type becomes random walk, meaning that the oscillator will appear to produce successive random steps in frequency when averaged over sufficiently long periods [14].

The noise identification graph shown in Fig. 8 is commonly found in the time and frequency literature, but note that ADEV was designed to estimate the frequency stability of free running oscillators, rather than disciplined oscillators. The flicker frequency region of an ADEV graph will never be reached with a GPSDO because its frequency is continuously being corrected to agree with UTC (USNO), which, as indicated in Table 1, is essentially equivalent to UTC and UTC(NIST). Notice, for example, that the stability in Fig. 7 has dropped below 1×10^{-13} after τ exceeds about two days and would continue to get smaller indefinitely if more data were collected. There are periods when this rate of improvement will slow down, and other periods where the stability will temporarily get worse due to frequency corrections or very low frequency noise sources (such as seasonal variations), but in theory a GPSDO will not reach a noise floor and both its stability and frequency uncertainty will continue to improve as the calibration period increases. This means that if a cesium oscillator were calibrated with a GPSDO over several months, for example, it should make little difference which make or model of GPSDO was used as the reference. All GPSDOs should have the accuracy and stability necessary to produce similar results, provided that they remain locked to the GPS satellites during the entire calibration.

There are, however, significant differences in the frequency stability of a GPSDO during periods of one day or less. The short-term stability of a GPSDO, at intervals shorter than the correction interval, should be identical to the short-term stability of its free running LO. Its medium-term stability, at intervals longer than the correction interval but shorter than one day, is design dependent and influenced by many factors. These factors include the quality of the receiver and antenna, the stability of the LO, the resolution of the comparator, the correction method, the correction uncertainty, and the correction interval.

Table 2 shows the frequency stability of ten GPSDOs calibrated by comparison to UTC(NIST), where τ is equal to one second (when data were available), one minute, one hour, and one day. The ten devices were each produced by different manufacturers. They are provided here only for purposes of example and represent only a very small sample of all of the available models.

As Table 2 indicates, a number of GPSDOs are stable to within parts in 10^{12} at $\tau = 1$ s. Even so, some GPSDOs have low cost LOs that are not particularly stable at short averaging periods, so it is best if a frequency calibration lasts for at least a few seconds, regardless of the measurement requirements. Each GPSDO reached a stability of at least 1×10^{-11} at $\tau = 1$ minute and at least 6×10^{-12} at $\tau = 1$ hr. Because calibrations of high performance oscillators typically last for one day or longer, a good metric to use when evaluating GPSDO performance is their frequency stability at $\tau = 1$ day. Stability of 1×10^{-13} or less at $\tau = 1$ day normally indicates an instrument of high quality and six of the ten devices tested reached or exceeded this specification. It is interesting to note that a few GPSDOs come close to matching the stability of the best commercially available cesium standards at $\tau = 1$ day, which, according to their manufacturer's specification, is about 0.3×10^{-13}.

6. A Method for Evaluating the Frequency Uncertainty of a GPSDO

The uncertainty analysis of frequency measurement referenced to GPS is simpler than the uncertainty analysis of a time measurement, which will be described in Section 8. This is mainly because equipment and signal path delays (the delays through cables, antennas,

GPS receivers, and so on) do not have to be calibrated or known if they can be assumed to be constant.

By international recommendation [17], expanded measurement uncertainty is reported in the form

$$Y = y \pm U, \tag{13}$$

where Y is the measurand or the quantity being measured (in this case frequency), y is the best estimate of the measurand (in this case the average frequency measured over a specified interval with respect to the SI), and U is the expanded measurement uncertainty.

The frequency accuracy of GPS is transferred to the LO with the techniques described in Section 4. These techniques "tune" the GPSDO frequency to agree with UTC and we can assume that a locked GPSDO is an accurate source of UTC. In fact, as summarized in Table 1, the frequency offset of a GPSDO with respect to UTC can be measured in parts in 10^{17} over a long interval and thus for all practical purposes y can be regarded as 0.

However, as noted previously, the accuracy of an oscillator over a given period can never be better than its stability over that same period. Therefore, Section 5 explained how to estimate frequency stability over the period of interest by use of ADEV. Going back to Eq. (12), the range of values from $y - U$ to $y + U$ is expected to "encompass a large fraction of the distribution of values that could reasonably be attributed to Y [17]." This is what ADEV does; it shows the distribution in values of frequency at a given period by including all of the various noise types that can cause the frequency to change. Thus, a convenient and robust way to estimate U for a GPSDO is simply to multiply its ADEV value by 2 to obtain a $k = 2$ coverage factor. The value chosen for ADEV should be at $\tau = f_{cd}$, where f_{cd} is the duration of the frequency calibration or measurement that utilized the GPSDO as a reference.

The frequency stability of a GPSDO for periods from of one day or longer can be obtained by having the device measured by a national metrology institute (NMI) such as NIST or by another laboratory with sufficient capability. If no measurement data are available, the only option is to use the specifications provided by the manufacturer.

We can now work through an example to evaluate the frequency uncertainty of a GPSDO during a calibration where $f_{cd} = 1$ day. Only two factors contribute to the frequency uncertainty of the GPSDO and both can be obtained through Type A (statistical) evaluations. The first factor is the stability of the NMI's time scale with respect to UTC at $\tau = 1$ day; the second factor is the stability of the GPSDO at $\tau = 1$ day with respect to the NMI.

Using NIST as our example NMI, Fig. 9 shows the stability of UTC(NIST) when compared to UTC, estimated by applying ADEV to the time difference data from *Circular T* that was graphed in Fig. 2. The stability of UTC(NIST) at $\tau = 5$ days is 1.8×10^{-15}. If we make the reasonable and conservative assumption that white phase noise is the dominant noise type in the *Circular-T* measurements at averaging periods shorter than five days (see the Fig. 8 model) we can extrapolate back to estimate ADEV at one day by applying a slope of τ^{-1}. This simply means that the stability estimate at one day will be five times larger than the stability estimate at five days, thus 9×10^{-15} is the number we will use in our calculations. The arrow on the previously shown Fig. 7 graph points to the stability of a GPSDO under test at $\tau = 1$ day with respect to UTC (NIST), which is 1.4×10^{-13}.

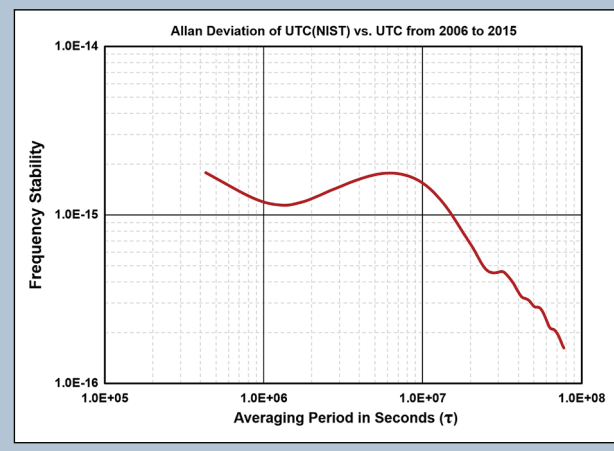

Figure 9. Frequency stability of UTC(NIST) with respect to UTC.

The combined uncertainty ($k = 2$) of the GPSDO in this example is estimated with the root sum squares (RSS) method as

$$\begin{aligned} U_c &= k\sqrt{U_{NMI}^2 + U_{GPSDO}^2} \\ &= 2\sqrt{(9 \times 10^{-15})^2 + (1.4 \times 10^{-13})^2} = 2.8 \times 10^{-13}, \end{aligned} \tag{14}$$

where k is the coverage factor, U_{NMI} is the ADEV of the NMI frequency standard (in this case NIST) with respect to UTC, and U_{GPSDO} is the ADEV of the GPSDO with respect to the NMI frequency standard.

If a calibration from an NMI or from an accredited laboratory with sufficient capability is not available, the GPSDO manufacturer's specification for stability at $\tau = 1$ day can be used in place of the measured value. This method is not optimal or desired, because manufacturers often provide very few details when specifying stability, sometimes even publishing the same specifications for GPSDOs of varying quality. The data sheet for the GPSDO in our example conservatively lists its frequency stability at 1 day as 1×10^{-12}, about a factor of seven worse than the measurement results shown in Fig. 7. Using 1×10^{-12} as the value for U_{GPSDO} the combined uncertainty is

$$\begin{aligned} U_c &= k\sqrt{U_{NMI}^2 + U_{GPSDO}^2} \\ &= 2\sqrt{(9 \times 10^{-15})^2 + (1 \times 10^{-12})^2} = 2 \times 10^{-12}. \end{aligned} \tag{15}$$

Note that in both Eq. (14) and Eq. (15) that the stability of the NIST frequency standard with respect to UTC is inconsequential and has no influence on the result. Thus, the ADEV value for the GPSDO at $\tau = 1$ day, multiplied by two, is the frequency uncertainty. However, with a less stable NMI and a more stable GPSDO, the U_{NMI} term would become more significant.

Note also that it is important to be able to distinguish between reasonable and unreasonable frequency uncertainty claims. For example, if a laboratory performs frequency calibrations that last for 24 hr with a GPSDO as their reference, a measurement uncertainty claim of 1×10^{-14} is probably false and should be closely scrutinized. However, an uncertainty of 1×10^{-13} for a 24-hour

calibration is possible with some GPSDOs and uncertainties near 1×10^{-12} should be achievable with nearly any GPSDO, provided that it has been correctly installed and is working properly. Calibration and metrology laboratories must also have procedures in place to determine whether a GPSDO is locked and working properly, and be able to show those procedures to laboratory assessors if necessary [3].

This completes our discussion of GPSDOs. We can now move on to discuss the time uncertainty of a GPS disciplined clock by beginning with a brief overview of time transfer.

7. Basic Principles of Time Transfer and GPS Disciplined Clocks

Time transfer is the science of transferring time at high accuracies from one location to another. All time transfer systems have a reference clock at their source (point A). Information from the reference clock is encoded on a signal that is transmitted through a wired or wireless medium to its destination (point B), the site where the remote clock is located. In the simplest form of time transfer, known as the "one-way" method (Fig. 10), the remote clock is synchronized with the time from the reference clock, which preferably has been adjusted to compensate for the path delay through the medium, d_{ab}. Even if the reference clock is a perfect source of UTC, the accuracy of the time transferred to the remote clock cannot be better than the uncertainty of the path delay measurement and its associated compensation [18].

The GPS satellites fly in medium Earth orbit at a height of about 20,200 km, and the path delay, even though radio signals travel at the speed of light (299,792,458 m/s), is about 67 ms. Despite the large path delay, GPS allows the path delay between the reference and remote clocks, d_{ab}, to be accurately measured and compensated to within a few nanoseconds (which means that the uncertainty in the path delay compensation is less than 1×10^{-7}). A major advantage of GPS is that the radio signals originate from the sky and thus the path between the transmitter and receiver is unobstructed. This makes it much easier to measure and compensate for path delay variations than it is for signals that originate from terrestrial radio stations, even though the terrestrial signals travel across much shorter paths.

Compensating for path delay is also much easier because the position of both the satellites and the receiver are accurately known. The position and velocity of the satellites can be estimated by using the orbital data from the GPS broadcast, and the receiver, through a series of range measurements performed with multiple satellites, can calculate its own position. By estimating the travel time of the signal and the exact time when the signal left the satellite, time from the satellite clocks is transferred to the receiver clock. This time difference between the satellite and receiver clock is converted to distance by multiplying by the speed of light. Due to a number of factors that contribute to uncertainty, the range measurements are not the true geometric range. Instead, they are known as the pseudorange, p, and are calculated as

$$p = \rho + c(dt - dT) + d_{ion} + d_{trop} + \varepsilon_{mp} + \varepsilon_{rn}, \qquad (16)$$

where p is the pseudorange measurement, ρ is the true geometric range, c is the speed of light, dt and dT are the time offsets of the satellite and receiver clocks with respect to GPS time, d_{ion} is the delay added to the radio signal as it passes through the ionosphere, d_{trop} is the delay added to the radio signal as it passes through the troposphere, $\varepsilon_{\mathrm{mp}}$ is the delay added by multipath signal reflections, and ε_{rn} represents the effects of receiver and antenna noise and all other delays that affect performance [19–21]. Section 8 explores the factors that make up the d_{ion}, d_{trop}, ε_{mp}, and ε_{rn} terms in more detail and while doing so provides a method for evaluating the time uncertainty of a GPSDC.

8. A Method for Evaluating the Time Uncertainty of a GPSDC

As discussed in Section 2, GPSDCs are synchronized to UTC (USNO), which is a very close approximation of UTC. This means that a GPSDC is inherently accurate, but the uncertainty of its 1 pps output is influenced by numerous factors and can vary over a fairly wide range. Here we identify eight factors that contribute to the time measurement uncertainty over a one-day period. The first contributor we discuss is evaluated with the Type A method and the remaining seven are evaluated with the Type B method [17]. Table 3 provides a summary.

8.1 U_{AS}, Time Stability

Time stability, as indicated by the variation in the 1 pps output of the GPSDC over a given interval, is evaluated with the Type A method. This is done by use of the time deviation (TDEV) statistic [14, 16] at $\tau = 1$ day. The time deviation, expressed as $\sigma_x(\tau)$, is an established metric for estimating time stability and time transfer noise when

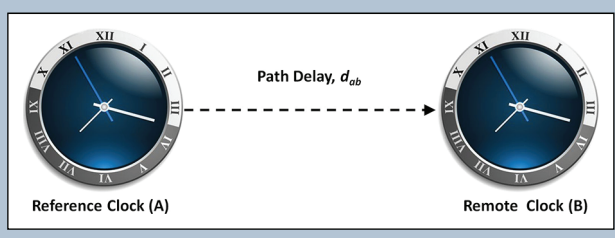

Figure 10. A one-way time transfer system.

Cable Type	Velocity Factor (%)	Delay per foot (ns)	Delay per meter (ns)
RG-58 (solid polyethylene)	66	1.54	5.05
RG-58 (foam polyethylene)	73	1.39	4.57
RG-8 (solid polyethylene)	66	1.54	5.05
RG-8 (foam polyethylene)	78	1.30	4.28
LMR-400	85	1.20	3.92

Table 3. Propagation delay through various types of GPSDC antenna cable.

the dominant type is white phase noise or flicker phase noise. It is closely related to MDEV and calculated as

$$\sigma_x(\tau) = \sqrt{\left(\frac{\tau^2}{3}\right) \cdot Mod\,\sigma_y(\tau)}, \qquad (17)$$

where $Mod\ \sigma_y(\tau)$ is calculated as shown in Eq. (12).

Figure 11 shows a TDEV graph of a GPSDC compared to UTC (NIST). In this example, TDEV with respect to UTC(NIST) is 1.4 ns at $\tau = 1$ day (indicated by arrow on graph). This result is fairly typical, and it would be unlikely for TDEV at $\tau = 1$ day to exceed 5 ns for any properly functioning GPSDC.

8.2 U_{BH}, Hardware Delays

Delays in the GPSDC hardware, specifically delays in the receiver, antenna, and antenna cable, can be significant contributors to the measurement uncertainty. The propagation delay through the antenna cable is usually the largest hardware delay. Nearly all GPSDCs allow the user to enter a value for the antenna cable delay that is used to correct the 1 pps output.

The cable delay can be estimated if the length and type of antenna cable are known. Cable manufacturers provide a velocity factor (VF) for their cables, also known as the velocity of propagation, that indicates how fast an electromagnetic signal will travel through a cable with respect to the speed of light in a vacuum. This number is usually expressed as a percentage of light speed; for example, a VF of 66 indicates that a signal would travel through the cable at 66% of the speed of light. Table 3 lists several types of antenna cables commonly used by GPSDCs, their velocity factors, and the time delay added by each foot or meter of cable. This information shows that the type of antenna cable used can have a significant impact on the delay. An RG-58 cable of 20 m in length, for example, would have a 101 ns delay, as opposed to about 78 ns for the same length of LMR-400.

It is preferable, of course, to measure, rather than estimate, the cable delay. Measurements made before the cable is installed not only eliminate errors in the cable length estimate, but also include the delays in cable connectors (which are typically sub-nanosecond).

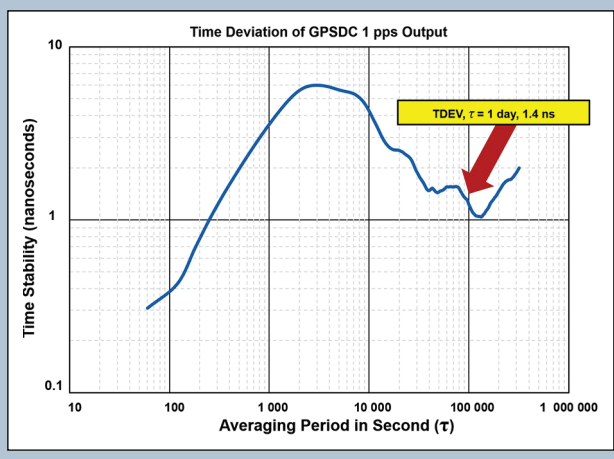

Figure 11. Time stability of a GPSDC with respect to UTC(NIST).

Several methods of accurately measuring cable delays are discussed in Jong [22] and Rovera et al. [23].

Receiver and antenna delays are more difficult to measure than antenna cable delays. A common technique is to calibrate the GPSDC as a system that includes the receiver, antenna, and antenna cable. This is usually done by mounting the antenna in a position with known coordinates, allowing the receiver to acquire satellites, and then comparing its 1 pps output to a reference, either to a reference GPSDC (the reference is sent to the device under test site or vice versa) [24], or to a reference time scale such as UTC(NIST). Measurements are generally averaged for a period of one or more days to obtain a delay constant that is entered into the GPSDC to compensate for all hardware delays. When a GPSDC is calibrated as a system, the calibration is relative to the reference. Thus, any difference between the reference and UTC at the time of the calibration must be included in the delay.

An absolute calibration of the GPS receiver delays can be performed using a GPS simulator, an instrument that connects to the antenna input of the GPSDC under test and mimics the signals broadcasts by the satellites. The simulator has its own reference clock so the 1 pps signal input to the receiver can be compared to the 1 pps signal output by the receiver to determine the absolute delay in the receiver circuits [25, 26]. The biggest drawbacks to the method are the high cost of GPS simulators and the fact that the antenna cable and antenna delays must still be calibrated separately, and antenna delay measurements typically require a network analyzer and an anechoic chamber.

The uncertainty assigned to hardware delays can vary widely. If the GPS receiver, antenna, and antenna cable are all carefully calibrated using either the relative or absolute methods just described, it might be possible to reduce U_{BH} to about 2 ns, but that level of calibration may require a considerable effort. It is highly likely that the receiver and antenna delays will be less than 100 ns, so simply compensating for the antenna cable delay should reduce U_{BH} to less than 100 ns. Thus, calibrating the antenna cable is always recommended, even if measuring the other delays is not practical. In the worst case, if all hardware delays are ignored and left uncalibrated, and if the longest possible antenna cable is used (for example some GPSDCs might work with 100 m of LMR-400 if their antenna has sufficient gain), then U_{BH} could be as large as 500 ns.

8.3 U_{BA}, Antenna Coordinate Error

Most GPSDCs can survey their antenna position by averaging position fixes for an interval ranging from a few hours to one day (some models allow the user to specify the interval, others use a fixed interval). This method can usually determine horizontal position (latitude and longitude) with sub-meter uncertainties; 20 or 30 cm is typical. However, because most GPSDCs only receive the L1 band signal, the self-survey often does a poor job of determining vertical position (altitude). If the antenna position was independently surveyed with an uncertainty of less than 1 m (by use of a dual-frequency geodetic GPS receiver, for example), or if the GPSDC itself is dual-frequency, we can assign sub-meter uncertainties for both the horizontal and vertical positions and estimate U_{BA} as 1 ns.

In most cases, however, there will be an error in the vertical position that results in a bias in the 1 pps output. In the worst case, if a GPSDC was getting time from a single satellite at an elevation angle

of 90°, this bias will equal the vertical position error multiplied by the speed of light, or ~3.3 ns of time error for each meter of position error. However, the time error per meter will be smaller than that because GPSDCs get time from multiple satellites. The error can be estimated by multiplying the speed of light constant by the sine of the satellite's elevation angle, which will be 1 at 90° and less than 1 at lower angles. If we assume, for example, that the average elevation angle of the satellites received by the GPSDC is 45°, or halfway between the horizon and the highest point in the sky, then the time bias can be roughly estimated at about 2.3 ns per meter ($3.3 \times \sin(45°) = 3.3 \times 0.707 = 2.3$ ns).

NIST maintains numerous L1-band GPS receivers at various locations that originally have self-surveyed their antenna and then later have been resurveyed with a dual-frequency geodetic receiver and a precise point positioning (PPP) service that determines vertical position with an uncertainty of 15 cm. The results of these resurveys have shown that the error in the self-survey of vertical position is usually less than 10 m, but the largest recorded error was 23.4 m for a receiver located in California. Figure 12 shows the daily time offsets of that particular GPSDC for 10 days using the self-surveyed coordinates and an additional 10 days after the results of the PPP survey were entered. By taking the average of the two 10-day segments, we can determine that the time shifted by 49.4 ns or 2.1 ns per meter. Because this particular self-survey appears to be an extreme case with a larger than normal error in vertical position, it seems reasonable to assume that U_{BA} is not likely to exceed 50 ns. Even so, errors in the vertical position of the antenna are the largest contributor to uncertainty for many GPSDCs.

8.4 U_{BE}, Environmental Effects

The hardware delays of a GPSDC change as a function of temperature and other environmental factors. Even though they are located indoors, the GPSDC's receiver and internal oscillator can be more sensitive to temperature changes than either the outdoor antenna or antenna cable. For example, if sudden or gradual changes in the indoor temperature occur, the 1 pps output of a GPSDC might experience a time shift or step of several nanoseconds, especially if the LO is a quartz device without temperature compensation. In most cases, the time shift will reverse when the temperature returns to normal.

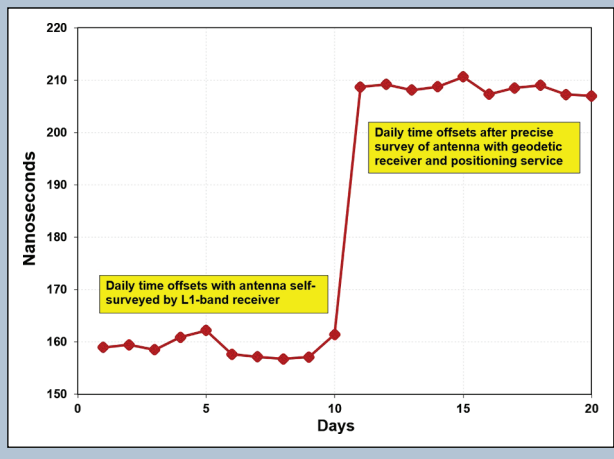

Figure 12. Graph of a GPSDC timing output before and after the vertical position of the antenna was precisely surveyed.

Antennas and antenna cable are less likely to experience sudden time shifts. The temperature coefficient of high quality antenna cable is usually much less than 1 ps (0.001 ns) per °C per meter [27]. Even at locations with a large range of outdoor temperatures and a long section of exposed cable, the delay variation in the antenna cable over the course of a year should be less than 1 ns. A similar range of delay variations is true of the antenna, where a temperature coefficient near 10 ps/°C is typical. The total contribution of U_{BE} to the measurement uncertainty, including delay variations in the receiver, antenna, and antenna cable; typically ranges from 2–5 ns, depending upon the sensitivity of the GPSDC hardware to temperature changes and the range of indoor and outdoor temperatures.

8.5 U_{BI}, Ionospheric Delay

The ionosphere is a region of the atmosphere that extends from about 60–1000 km above the Earth's surface. It is ionized by solar radiation and has a large influence on the propagation of radio signals. In the case of GPS signals, which originate from above the ionosphere, the radio signals are slightly refracted or bent as they pass through the ionospheric layers, leading to a physical phenomenon known as dispersion. The dispersion introduces a variable propagation delay that contributes to the time uncertainty.

Dual-frequency GPS receivers can accurately measure and compensate for ionospheric delay by differencing pseudo range measurements from both the L1 (1575.42 MHz) and L2 (1227.6 MHz) carrier frequencies [28, 29]. This works because the delay through the ionosphere due to dispersion is frequency dependent and can be accurately estimated from the ratio of the delays of two signals at different frequencies. However, as previously noted, most GPSDOs and GPSDCs are based on single-frequency (L1 only) receivers and there is no way to compare delays when only one frequency is received. Thus, in most cases the ionospheric delay must be estimated, typically by utilizing a model developed by Klobuchar [30, 31]. The Klobuchar model requires the latitude, longitude, elevation angle, and azimuth as its inputs, in addition to eight coefficients received from the GPS broadcast. The four "alpha" values, α_n, are the coefficients for a cubic equation that represents the amplitude of the vertical delay. The four "beta" values, β_n, are the coefficients of a cubic equation that represents the period of the model.

At all times of day, satellites at low elevation angles normally require a larger ionospheric delay correction than those at higher elevation angles. However, the magnitude of the correction is much larger during the daytime for satellites of all elevation angles than it is during the nighttime. The graph in Fig. 13 shows the modeled ionospheric correction for each satellite track collected in Boulder, Colorado for a one-day period in July. The shaded area indicates the period between sunset and sunrise when the corrections are the smallest. The blue markers indicate satellites where the elevation angle was between 10° and 20° and the red markers indicate satellites where the elevation angle was above 20°. The largest ionospheric corrections, for satellites below 20° during the daytime hours, sometimes exceeded 40 ns. The corrections from the Klobuchar model are expected to remove at least 50% of the actual ionospheric delay [30].

Measurements of a GPSDC will show a diurnal variation near sunrise and sunset, where the time changes (sometimes by more

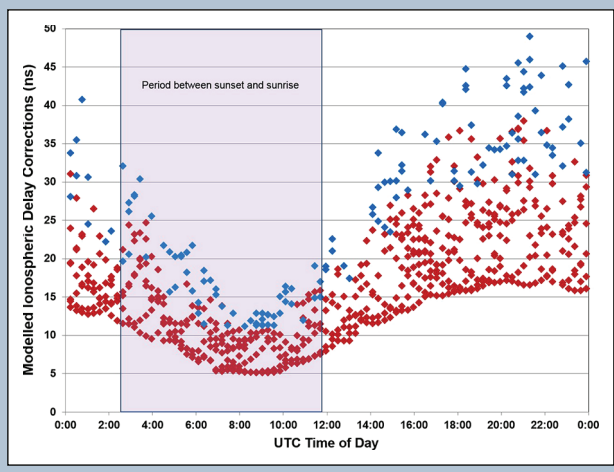

Figure 13. Modeled ionospheric delay corrections for all satellites received in Boulder, Colorado over a one-day period. The blue markers indicate satellites where the elevation angle was between 10° and 20°. The red markers indicate satellites where the elevation angle was above 20°.

than 10 ns) due to uncertainties in the modeled ionospheric delay corrections. Most of the effects of the diurnal variation are removed by averaging for one day, but a residual bias, typically of a few nanoseconds, will remain in the average daily time recorded by a GPSDC. To illustrate this, Fig. 14 shows the average daily time offset of a GPSDC located in Boulder, Colorado during the month of August 2016. The graph compares the results obtained when applying a real-time correction from the Klobuchar model to the results obtained when applying a post processed correction from an ionospheric delay measurement. The average daily difference between the model and the measurement represents the bias in the model, which was 2.1 ns in this example. This value will vary as a function of geographic location and the current amount of geomagnetic activity, which can cause the model to have larger

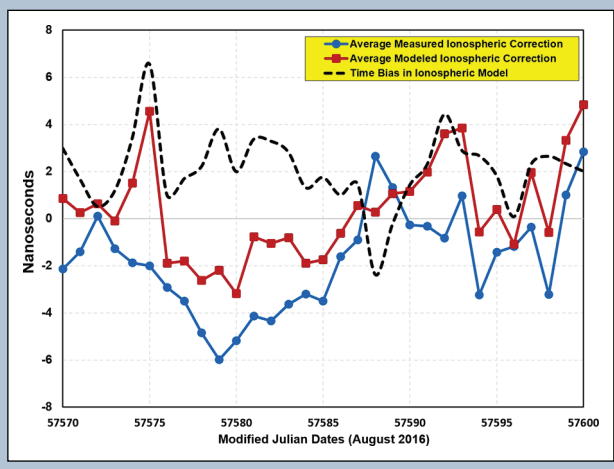

Figure 14. Differences between modeled and measured ionospheric delay corrections that are averaged for one day.

errors than usual. However, U_{BI} will normally not exceed 10 ns with values of less than 5 ns being typical.

8.6 U_{BT}, Tropospheric Delay

The troposphere is the lowest layer of the Earth's atmosphere, extending from the Earth's surface to a height ranging from about 7–20 km, depending upon the geographic location and time of year. The weather that we experience on Earth occurs in the tropospheric region. Another atmospheric region, the stratosphere, is above the troposphere and below the ionosphere. For the purposes of this discussion, the delays encountered by the GPS signals as they pass through the stratosphere (and other layers below the ionosphere) are considered to be part of the tropospheric delay.

Like the ionosphere, the troposphere adds delay to the GPS signals. However, unlike the ionosphere, the delay introduced is not frequency dependent and neither a single frequency or dual frequency GPS receiver can measure it. Also, unlike the coefficients broadcast by GPS for the Klobuchar model, there is nothing in the broadcast related to a tropospheric delay correction. Even so, numerous tropospheric models exist [32, 33] that have been incorporated by GPS receiver designers with good results. The main piece of information needed for these models is the elevation angle of the satellite, because the delay of a GPS signal increases as the elevation angle decreases. Roughly 90% of the tropospheric delay, known as the dry component, is due to signal refraction due to atmospheric pressure. This part can be modeled much more easily than the wet component, which depends on the amount of water vapor in the atmosphere. The wet component is very difficult to model but fortunately only accounts for about 10% of the total tropospheric delay [33].

A tropospheric delay model that is known to have been implemented by at least some GPSDCs is the NATO model, which gets its name because it was once published as a North Atlantic Treaty Organization (NATO) standard [32, 34]. This model, in a fashion similar to other models implemented by GPSDCs, focuses entirely on the dry component; ignoring the wet component because no input data is available. It estimates the delay through the troposphere in meters, d_{trop}, as

$$d_{trop} = d^z_{trop} \frac{1}{\sin(\varepsilon) + \frac{0.00143}{\tan(\varepsilon)+0.0455}}, \tag{18}$$

where ε is the elevation angle of the satellite in radians. The d^z_{trop} term (total zenith delay) is computed differently based on the orthometric height (altitude) of the antenna. For a height of less than 1 km

$$d^z_{trop} = \left\{1430 + 732 + \left\{SR \times (1-H) + 0.5\Delta N\left(1 - H^2\right)\right\}\right\} \times 10^{-3}, \tag{19}$$

and for a height greater than or equal to 1 km and not exceeding 9 km

$$d^z_{trop} = \left\{732 + \left\{\frac{N_1}{c} \times \exp(-c(H-1)) - \exp(-8c)\right\}\right\} \times 10^{-3}, \tag{20}$$

where

SR is the global mean sea level refractivity, a constant of 324.8 was recommended in [34];

H is the orthometric height in kilometers;

$$\Delta N = -7.32 \times \exp(0.005577 \times SR);$$
$$N_1 = SR + \Delta N; \text{ and}$$
$$c = \frac{1}{8} \times \log\left(\frac{N_1}{105}\right).$$

The delay correction, d_{trop}, is then converted from meters to nanoseconds using the speed of light constant. Figure 15 shows the tropospheric delay correction as a function of elevation angles between 10° and 90° at an antenna height of 1645 m (the altitude of the NIST laboratories in Boulder).

Unless revealed by the manufacturer, it is usually not possible to determine what tropospheric model a GPSDC is using. However, even simple tropospheric models, such as the NATO model, can compensate for approximately 90% of the delay introduced by the dry component. Note in Fig. 15 that the correction is less than 10 ns at an average elevation angle of 45°, suggesting that the uncertainty in the dry component can be reduced by a simple model to less than 1 ns. Even if the wet component is uncompensated in the model, it is only about 10% of the total tropospheric delay, and thus is likely to also be below or near 1 ns for a satellite at an average elevation angle. Therefore, a typical value for U_{BT} is ~2 ns, increasing or decreasing slightly due to the quality of the tropospheric model and current weather conditions.

8.7 U_{BM}, Multipath Signal Reflections

Not all GPS signals travel in a straight line from the satellite to the antenna. Instead, some are reflected from surfaces both below and above the antenna before being received. These multipath signal reflections can interfere with, or be mistaken for, the signals that travel a straight-line path from the satellite. This results in delay changes that affect the time uncertainty.

Multipath is one of the most discussed GPS problems, due to the adverse effect it can have on GPS positioning [35]. However, the uncertainty contributed by multipath to a GPSDC, where the antenna is stationary, is often not significant if the antenna is located at a site with a clear, unobstructed view of the sky. It can be made even less significant by choosing an antenna specifically designed to reduce multipath. For example, the inexpensive patch antennas supplied with most GPSDCs are more susceptible to multipath reflections than more expensive antennas with a quadrifilar helix, choke ring, or pinwheel design [36].

Multipath reflections can be identified by looking at the individual satellite tracks. If a multipath reflection causes a delay in a signal from a given satellite, a similar delay should be noticeable again on the following day. However, the delay will occur approximately 4 min earlier on the following day, because GPS orbits are based on the sidereal day, which is about 4 min shorter than the solar day. Figure 16 shows segments of GPS satellite tracks (1 min averages) recorded from the same satellite at the same time for five consecutive days. During each satellite flyover, its signal was briefly reflected by a metal cabinet located on the roof near the antenna. This caused the peaks shown in the graph, which occurred 4 min earlier each day. The magnitude of these time shifts is large (~50 ns), but they are quickly reduced by averaging because their duration is relatively short and because only satellites at specific positions in the sky are affected by the multipath reflection. The situation could, of course, be improved by moving the antenna away from the metal cabinet or by replacing the antenna with one that does a better job of rejecting reflected signals.

With most GPSDCs, there is no easy way to analyze the phase or the stability of individual satellite tracks, which makes it difficult to know if multipath, or another uncertainty source, is limiting performance. Normally, multipath is not a significant problem for GPSDC users if the antenna has an unobstructed sky view. The total contribution of U_{BM} to the time uncertainty is often insignificant when averaged for one day, near 1 ns if a multipath mitigating antenna is used and ~2 ns with a standard patch antenna. If a reasonable amount of caution is exercised when mounting the antenna; for example, if the antenna is placed on the highest part of the roof instead of on the side of the building, then U_{BM} should be less than 5 ns.

8.8 U_{BU}, UTC – USNO(USNO) Offset

The uncertainty of a GPSDC should be estimated with respect to UTC and the SI and, as noted in Section 2, GPSDCs distribute

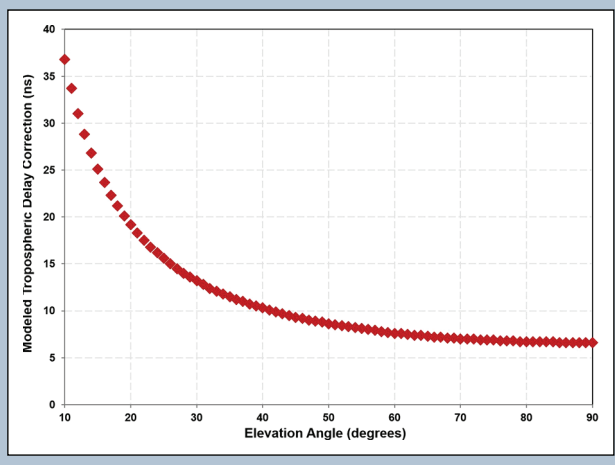

Figure 15. The tropospheric delay correction (NATO model) as a function of elevation angle in Boulder, Colorado.

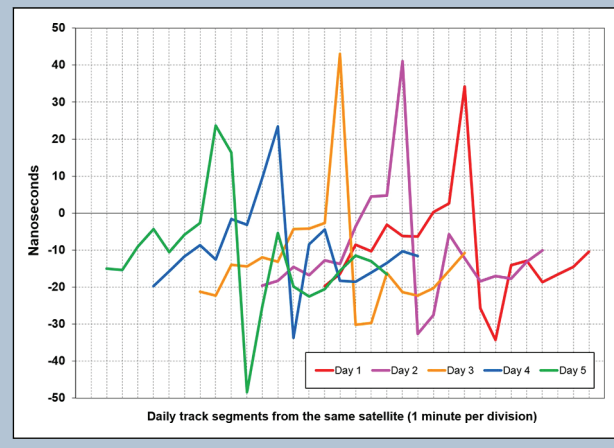

Figure 16. Multipath signal reflections from a GPS satellite recorded for five consecutive days.

Uncertainty Component	Best Case (geodetic antenna survey, stable hardware with all delays calibrated)	Typical (self-survey of antenna, antenna cable delay calibrated)	Worst Case (self-survey of antenna, receiver and antenna delays uncalibrated)
U_{AS}, Time Stability	1	2	5
U_{BH}, Hardware Delays	2	20	500
U_{BA}, Antenna Coordinates	1	20	50
U_{BE}, Environmental Effects	2	3	5
U_{BI}, Ionospheric Delay	2	5	10
U_{BT}, Tropospheric Delay	1	2	3
U_{BM}, Multipath Reflections	1	2	5
U_{BU}, UTC – UTC(USNO) Offset	1	5	10
U_C, $k = 2$	8	60	1005

Table 4. Time uncertainties of GPSDCs for a one-day period (values are in nanoseconds).

UTC(USNO). Therefore, the current time offset between UTC and UTC(USNO) should be included in the uncertainty analysis. This number can be obtained from a current *Circular T* document, but as shown in Fig. 1, it was 0.5 ns on average for the 10-year period of 2006 through 2015. It seldom exceeds 5 ns and 10 ns can be considered the worst case based on historical data.

Table 4 shows the "best case" and "worst case" uncertainties of a GPSDC with respect to UTC. The combined uncertainty, U_c, is obtained with the root sum of squares method [17], where k is the coverage factor, as

$$U_c = k\sqrt{U_{AS}^2 + U_{BH}^2 + U_{BA}^2 + U_{BE}^2 + U_{BI}^2 + U_{BT}^2 + U_{BM}^2 + U_{BU}^2}. \quad (21)$$

As indicated in Table 4, the largest sources of GPSDC uncertainty are usually uncompensated hardware delays, mainly caused by the delay through the antenna cable which tends to be larger than the receiver and antenna delays, and by antenna coordinate errors. The other uncertainty sources are relatively small, and typically only attract attention after the hardware delays have been calibrated and the antenna has been accurately surveyed. The "best case" uncertainty ($k = 2$) for a high-quality single-frequency GPSDC that has been carefully calibrated is about 8 ns. A more typical uncertainty, assuming that the antenna was self-surveyed and that the antenna cable has been calibrated, is often in the 50–60 ns range. The "worst case" uncertainty for a GPSDC where all hardware delays are ignored and uncalibrated could reach 1 µs (1,000 ns), but uncertainties this large are unusual. A 1 µs specification meets the needs of critical infrastructure timing systems, and the GPSDCs deployed in these systems do not need to be calibrated to meet the specifications. They can be trusted to keep time within 1 µs of UTC as long as they are receiving satellite signals and working properly [37].

9. Summary

Many, if not most, calibration laboratories now operate GPS disciplined oscillators and clocks as their frequency and time standards, a decision that makes sense for both economic and technical reasons. Despite their widespread use, laboratory assessors are sometimes skeptical of the traceability and uncertainty claims made by calibration laboratories that rely on GPS devices. To help alleviate this skepticism, this article described how GPS produces frequency and time signals that are inherently accurate and traceable to the SI. It has also provided methods for evaluating the uncertainty of frequency and time measurements referenced to GPS disciplined oscillators and clocks.

References

[1] Resolution 1 of the *13th Conference Generale des Poids et Mesures (CGPM)*, 1967.

[2] J. Kusters, L. Cutler, and E. Powers, "Long-term experience with cesium beam frequency standards," *Proc. of the IEEE Frequency Control Symposium and European Frequency and Time Forum (EFTF)*, pp. 159–163, April 1999.

[3] M. Lombardi, "The use of GPS disciplined oscillators as primary frequency standards for calibration and metrology laboratories," *NCSLI Meas. J. Meas. Sci.* vol. 3, no. 3, pp. 56–65, September 2008.

[4] G. Panfilo, "The coordinated universal time," *IEEE Instrum. Meas. Mag.* vol. 19, no. 3, pp. 28–33, June 2016.

[5] G. Petit, F. Arias, A. Harmegnies, G. Panfilo, and L. Tisserand, "UTCr: a rapid realization of UTC," *Metrologia* vol. 51, no. 1, pp. 33–39, February 2014.

[6] P. Whibberly, J. Davis, and S. Shemar, "Local representations of UTC in national laboratories," *Metrologia* vol. 48, pp. S154–S164, July 2011.

[7] Global Positioning Systems Directorate, "Systems Engineering & Integration Interface Specification," *IS-GPS-200F*, September 2011.

[8] M. Miranian, "UTC dissemination to the real-time user: The role of USNO," *Proc. of the 27th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting*, San Diego, California, pp. 75–85, November 1995.

[9] V. Zhang, T. Parker, R. Bumgarner, J. Hirschauer, A. McKinley, S. Mitchell, E. Powers, J. Skinner, and D. Matsakis, "Recent Calibrations of UTC(NIST) – UTC(USNO)," *Proc. of the 44th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting*, Reston, Virginia, pp. 35–42, November 2012.

[10] G. Kamas and S. Howe, eds., "Time and frequency user's manual," *National Bureau of Standards Special Publication 559*, 256 pp., November 1979.

[11] K. Åström and T. Hägglund, *PID Controllers: Theory, Design, and Tuning*, 2nd ed., Instrument Society of America, Research Triangle Park, North Carolina, 1995.

[12] B. Shera, "A GPS-based frequency standard," *QST Mag.* vol. 82, pp. 37–44, July 1998.

[13] B. Cui, X. Hou, and D. Zhou, "Methodological Approach to GPS Disciplined OCXO Based on PID PLL," *Proc. of the 9th International Conference on Electronic Measurements and Instruments (ICEMI)*, Beijing, China, pp. 1–528 to 1–533, August 2009.

[14] W. Riley, *Handbook of Frequency Stability Analysis*, NIST Special Publication 1065, July 2008.

[15] J. Jespersen, "Introduction to the time domain characterization of frequency standards," *Proc. of the 23rd Annual Precise Time and Time Interval (PTTI) Systems and Application Meeting*, Pasadena, California, pp. 83–102, December 1991.

[16] IEEE, "Standard definitions of physical quantities for fundamental frequency and time metrology—Random instabilities," *IEEE Standard 1139*, 2008.

[17] JCGM, "Evaluation of measurement data—Guide to the expression of uncertainty in measurement," *JCGM 100*, 2008.

[18] J. Jespersen and L. Fey, "'Time-telling' techniques," *IEEE Spectrum* vol. 9, no. 5, pp. 51–58, May 1972.

[19] E. Kaplan and C. Hegerty, *Understanding GPS: Principles and Applications*, 2nd ed., Artech House Publishers, 2005.

[20] P. Misra and P. Enge, *Global Positioning System: Signals, Measurement, and Performance*, 2nd ed., Ganga-Jamuna Press, 2011.

[21] W. Lewandowski and C. Thomas, "GPS time transfer," *P. IEEE*, vol. 79, no. 7, pp. 991–1000, July 1991.

[22] G. de Jong, "Measuring the propagation time of coaxial cables used with GPS receivers," *Proc. of the 17th Annual Precise Time and Time Interval (PTTI) Systems and Application Meeting*, pp. 223–232, December 1985.

[23] D. Rovera, M. Abgrall, P. Uhrich, and M. Siccardi, "Techniques of antenna cable delay measurement for GPS time transfer," *Proc. of the 2015 Joint Conference of the IEEE International Frequency Control Symposium and the European Frequency and Time Forum*, Denver, Colorado, pp. 239–244, April 2015.

[24] W. Lewandowski, J. Azoubib, and W. Klepczynski, "GPS: Primary tool for time transfer," *P. IEEE*, vol. 87, no. 1, pp. 163–172, January 1999.

[25] G. Petit, Z. Jiang, J. White, R. Beard, and E. Powers, "Absolute calibration of an Ashtech Z12-T GPS receiver," *GPS Solut.* vol. 4, no. 4, pp. 41–46, April 2001.

[26] J. Plumb, K. Larson, J. White, and E. Powers, "Absolute calibration of a geodetic time transfer system," *IEEE T. Ultrason. Ferr.*, vol. 52, no. 11, pp. 1904–1911, November 2005.

[27] E. Powers, P. Wheeler, D. Judge, and D. Matsakis, "Hardware delay measurements and sensitivities in carrier phase time transfer," *Proc. of the 30th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting*, pp. 293–303, 1998.

[28] V. Zhang and Z. Li, "Measured ionospheric delay corrections for code-based GPS time transfer," *NCSLI Measure J. Meas. Sci.*, vol. 10, no. 3, pp. 66–71, September 2015.

[29] J. Rose, R. Watson, D. Allain, and C. Mitchell, "Ionospheric corrections for GPS time transfer," *Radio Sci.*, vol. 49, pp. 196–206, 2014.

[30] J. Klobuchar, "Ionospheric time-delay algorithm for single-frequency GPS users," *IEEE T. Aero. Elec. Sys.*, vol. AES-23, no. 3, pp. 325–331, 1987.

[31] ARINC Interface Control Document, "Navstar GPS Space Segment/Navigation User Interfaces," *ICDGPS-200*, Revision C, Arinc Research Corporation, IRN-200C-004, 2000.

[32] V. Mendes, "Modeling the neutral-atmosphere propagation delay in radiometric space techniques," Ph.D. dissertation, Dept. of Geodesy and Geomatics Engineering Technical Report No. 199, University of New Brunswick, Fredericton, New Brunswick, Canada, 353 pp., September 1998.

[33] J. Spilker, Jr., "Tropospheric Effects on GPS," Ch. 13 in *Global Positioning System: Theory and Applications*, vol. 1, pp. 517–545, American Institute of Aeronautics and Astronautics, Inc., 1996.

[34] North Atlantic Treaty Organization (NATO), "Navstar Global Positioning System (GPS) system characteristics," *NATO Standardization Agreement (STANAG) 4294*, 1993.

[35] J. Soubielle, I. Fijalkow, P. Duvaut, and A. Bibaut, "GPS positioning in a multipath environment," *IEEE T. Signal Process.*, vol. 50, no. 1, pp. 141–150, January 2002.

[36] M. Lombardi and A. Novick, "Effects of the rooftop environment on GPS time transfer," *Proc. of the 38th Annual Precise Time and Time Interval (PTTI) Systems and Application Meeting*, pp. 449–465, December 2006.

[37] M. Lombardi, "Microsecond timing at multiple sites: Is it possible without GPS?," *IEEE Instrum. Meas. Mag.*, vol. 15, no. 5, pp. 14–21, October 2012.