# Preliminary Evaluation of Time Scales Based on Hydrogen Masers

Marc A. Weiss and Fred L. Walls, *Senior Member, IEEE*

***Abstract*—Two experimental time scales based on ensembles of hydrogen masers were generated and compared with a variety of other references both internal and external to NIST. The masers all had some type of active cavity control to reduce frequency drift due to cavity changes. The first experimental time scale, TA2M, was generated over the interval from MJD 49050 to 49190 (March 4, 1993–July 22, 1993) using measurements between the masers every 6 hours. TA2M was compared to the NIST AT1 time scale, primarily based on commercial cesium frequency standards, and via once-per-day GPS common-view measurements to the USNO unsteered master clock and to UTC. The linear frequency drift of TA2M relative to UTC was small compared to the uncertainty of $2 \cdot 10^{-16}$/d in estimating linear frequency drift.**

**The masers at NIST were then linked to two masers at USNO using GPS common-view time transfer to study the performance of a maser ensemble, TA2M1, at measurement times from a few days to a few months. Again we found that the frequency drift of TA2M1 relative to UTC was smaller than the estimation uncertainty of $0.6 \cdot 10^{-16}$/d. The stability of the ensemble at 1 month appeared to be about 1–2 parts in $10^{15}$.**

**From this work we see that the use of autotuned hydrogen masers in time scales is very promising. A single such maser rivals the 1993 stability of UTC. The typical fractional frequency stability of the masers showed flicker frequency modulation at $4 \cdot 10^{-15}$ or less for measurement times of 10 d–20 d. Linear frequency drift was measured for the masers against the TA2M1 time scale at levels under $1 \cdot 10^{-16}$/d with uncertainties of $0.3 \cdot 10^{-16}$/d. Measuring frequency drift against the definition of the SI second at this level pushes the limits of current technology.**

## I. INTRODUCTION

ACTIVE hydrogen masers (H-masers) with cavity tuning now maintain fractional frequency stability with little drift. In the past, active H-masers were not useful for time scale generation because of linear drifts of fractional frequency of the order of $10^{-14}$ to $10^{-15}$/d. Frequency drift in itself does not perturb a time scale, but rather the error in estimation of this deterministic effect [1]. The problem with estimation is due to the presence of flicker and random walk frequency modulation in the reference cesium standards. In the time it takes to attain the needed precision, the true value for the drift will have changed. The fractional frequency stability of several devices is shown in Fig. 1. The stability of the masers shown is without any estimation of frequency drift. Clearly, the drift must be estimated very precisely for these devices to compete with the best cesium standards in the long term.

A technique for interrogating the resonant cavity to measure its electrical size in real time was first developed at NIST in 1976 [2]. Our studies confirmed that the change in the cavity size was the major factor causing frequency drift in masers [3]. Somewhat different techniques are now used in a number of active H-masers. Frequency drift is now controlled in such masers to a level that is difficult to measure. We will see here that measuring the drift of these devices requires integrating over about 1 year against some of the world's best current cesium frequency standards.

We have combined a number of these devices at NIST and at USNO into experimental time scales to illustrate their potential. We find that the stability of a single maser is greater for measurement times out to several months, than most national time scales based on cesium beam frequency standards. Further, the drift at one year is competitive with some of the best ensembles in the world. This experiment was designed to document this performance and to better characterize the masers. H-masers with cavity autotuning can never be a primary reference for producing accurate frequency, due to limits in evaluating the wall shift of $5 \cdot 10^{-13}$ at best [4]. However, they can maintain a frequency once they have been evaluated.

## II. SHORT-TERM STABILITY

The short-term frequency stability of the masers at NIST was measured using a dual-mixer system [5]. The noise floor of the standard system is far above that of active masers for measurement times below about 1000 s. To circumvent this problem, data between three masers were taken on two channels each. Cross correlation was then used to reduce the contribution of both the reference masers and the measurement system to the estimation of the Allan variance for the maser under test [6]. In this way the measurement noise was reduced to approximately 1 part in $10^{14}$ for measurement times from 0.5 to 20 s. For measurement times from 10 s to $10^5$ s, we also used the modified Allan variance to estimate frequency stability. The fractional frequency stabilities of three masers at NIST, CH-A, CH-B, and H1, operating in a cavity autotuned mode are shown in Fig. 2. The stabilities estimated using these improved techniques are plotted for the three masers for integration times from 1 s to 10 240 s. The data points for longer times are measurements against the TA2M scale as described below using the NIST dual-mixer time scale system which records time differences between clocks every 2 hours.

Masers CH-A and CH-B are cavity tuned by using a magnetic field gradient to modulate the atomic linewidth, and hence to modulate frequency pulling due to cavity mistuning. In this mode the atomic line width of one maser is modulated,

Manuscript received May 17, 1994; revised May 5, 1995.

The authors are with NIST Time and Frequency Division, Boulder, CO 80303 USA.

Publisher Item Identifier S 0018-9456(96)00076-9.

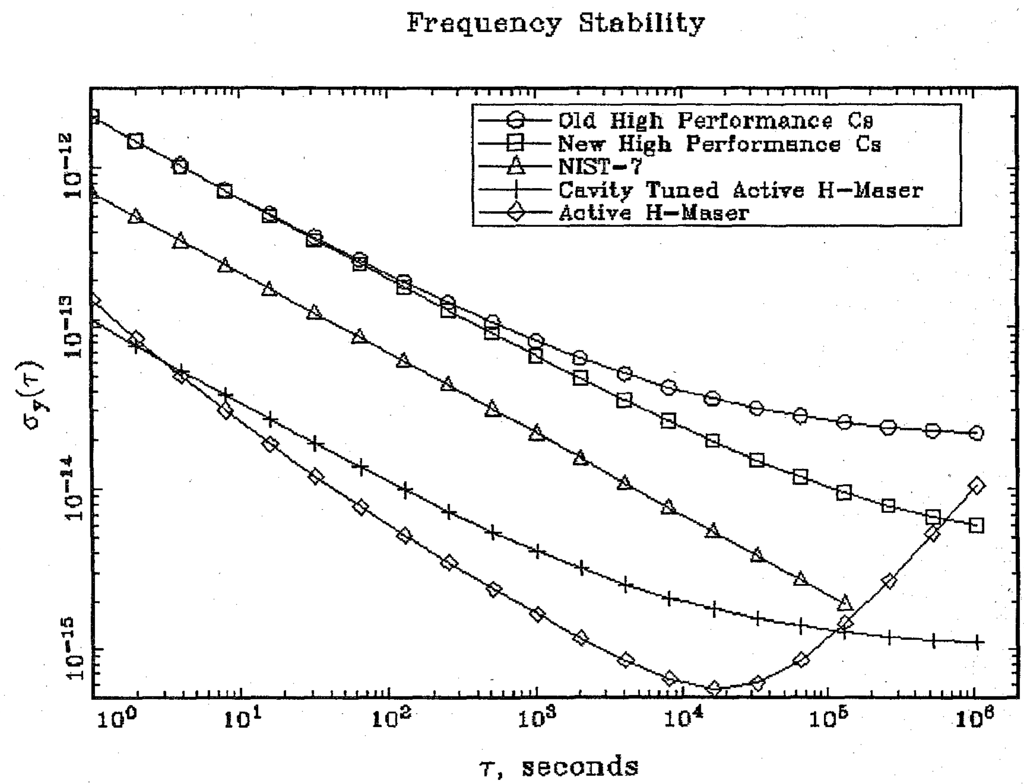



Fig. 1. The frequency stability of current and recent state-of-the-art clocks. The stability of the autotuned masers now exceeds the long-term performance of commercial cesium beam standards.

and the cavity is frequency-adjusted under servo control until its oscillation frequency as measured against the other maser is independent of atomic line width. The reference and tuned masers are then switched so that the cavity frequencies of both masers are autotuned. In principle this should remove the effect of cavity mistuning which is the major contribution to frequency drift in conventional masers [3]. Maser H1 uses a different scheme for cavity autotuning in which the cavity frequency is modulated. A servo system then tunes the cavity frequency until the oscillation frequency is independent of small mistunings. This approach does not require an external frequency reference.

## III. Description of the Time Scales

We report here the generation of two experimental time scales. We used the TA2 time scale algorithm to generate both these scales. TA2 is a smoothing algorithm, in which a time scale is run both forwards and backwards in time. The results from the two directions are then combined to obtain a post-processed scale optimized for frequency stability [7], [8]. The one-directional time scale algorithm used, AT2, is an extension of AT1, the algorithm used to generate the real-time unsteered time scale at NIST, with the addition of an estimate of confidence of the frequency estimate. This confidence estimate allows the forward and backward AT2 estimates of frequency to be combined into a smoothed estimate of frequency. A third pass in the forward direction uses this smoothed estimate of the frequency offset of each clock to estimate the time offsets.

The first scale we generated is called TA2M. It used measurements among masers at NIST over the interval from MJD 49050 to 49190, (March 4, 1993–July 22, 1993). Measurements taken every 6 hours among the three masers, CH-A, CH-B, and H1, discussed in the previous section at NIST were combined to generate TA2M. This experimental scale was compared to the NIST AT1 time scale primarily based on commercial cesium frequency standards, and via GPS to the USNO unsteered master clock and to the international UTC, as generated by the BIPM and published in Circular T, [9], [10].

Measurements taken once per day at NIST were linked to once-per-day measurements at USNO using the GPS common-view time transfer technique to generate the TA2M1 scale, with scale estimates once per day. The CH-A, CH-B, and H1 masers at NIST were then linked to the N4 and N5 masers at USNO to study the performance of a maser ensemble at longer measurement times. This ensemble was run from MJD 49050-49349, (March 4, 1993–December 28, 1993), though maser CH-A at NIST left after 49190, and maser CH-B was not cavity-tuned for most of this run. Thus, after the first 140 d, for about half of the time interval of TA2M1, the scale was based on the maser H1 at NIST and masers N4 and N5 at USNO. These masers all use the same type of

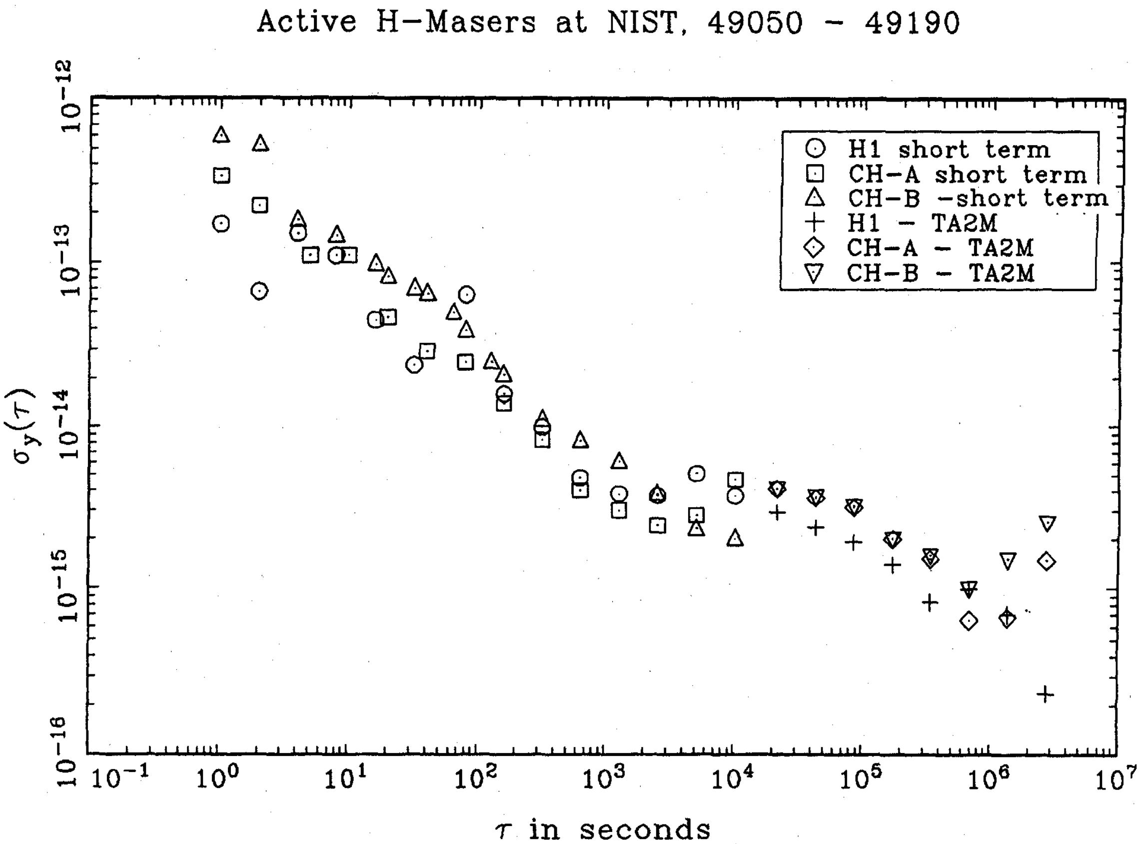



Fig. 2. The short-term stability of three autotuned H-masers at NIST. The stabilities from 1 to 10 240 s are from the enhanced measurement system and analysis technique; the stabilities from 21 600 s = 1/4 d to 16 d are against TA2M, corrected for correlation.

cavity autotuning mentioned above which does not require an external frequency reference.

## IV. Results from the TA2M Scale

Since there are only three clocks in the scale and they are all similar in the long term, they are each weighted with 1/3 of the scale. There is a brief period where maser H1 is unavailable, because adjustments were being made. For this period, the scale becomes a simple average of the other two masers. There is no measurable shift in the scale during this period. Fig. 3 shows the fractional frequency of the maser H1 versus TA2M, and the scale TA2M versus AT1 and USNO A1r on the same plot. Frequency smoothing and offsets were applied to separate them on the plot. We can see the period MJD 49136–49150 when the maser H1 is absent, and there is no apparent shift in TA2M against either AT1 or USNO A1r at these times.

Fig. 4 shows the frequency of TA2M versus the international UTC. The drift of the maser scale was negligible compared to the GPS common-view noise and frequency stability of the scale over the period of the measurements. The confidence on an overall drift estimate due to these factors was $1.5 \cdot 10^{-16}$/d. The drift of the ensemble relative to UTC was small compared to this uncertainty.

The stability of the masers against the TA2M scale is shown in Fig. 2, extending the short- dterm measurements. The values against TA2M begin at measurement times of 6 h = 21 600 s.

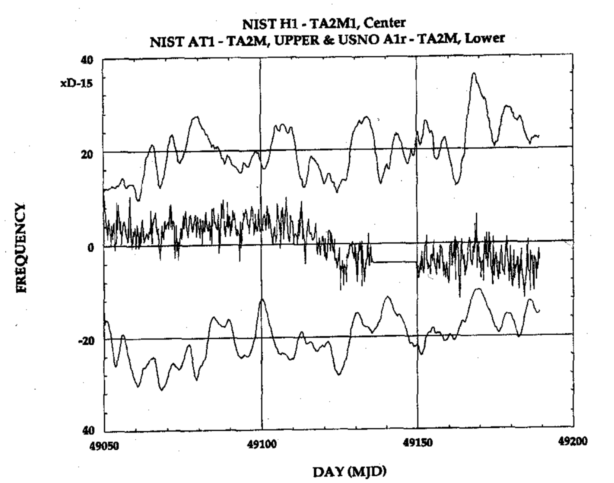



Fig. 3. The residual fractional frequencies of the NIST H1 maser, the NIST AT1 scale and the unsteered USNO A1r scale all against the TA2M scale. Mean frequencies have been removed, AT1-TA2M has been offset by $2.0 \cdot 10^{-14}$, and A1r-TA2M has been offset by $-2.0 \cdot 10^{-14}$. In addition, AT1-TA2M and A1r-TA2M have been smoothed to show slower frequency changes.

These stabilities have been adjusted for correlation with the scale by the following formula relating the true variance to the measured variance [11]

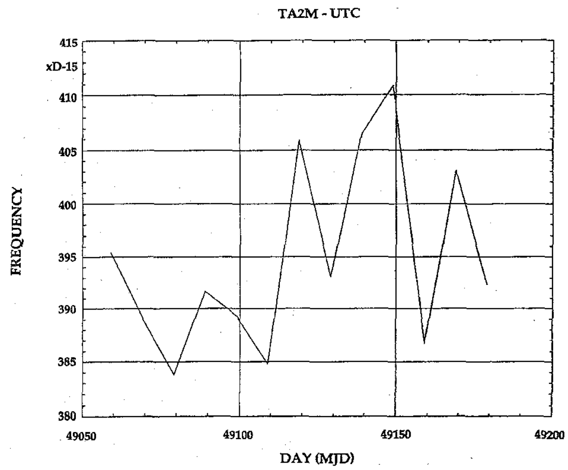



Fig. 4. The frequency of TA2M versus the international UTC in 10 d frequency averages. The drift of the maser scale was negligible compared to the GPS common-view noise and frequency stability of the scale over the period of the measurements. The confidence on an overall drift estimate due to these factors was $1.5 \cdot 10^{-16}$/d. The drift of the ensemble was small compared to this.

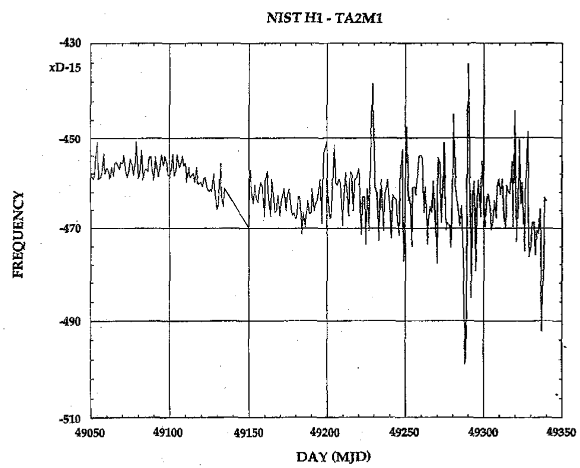



Fig. 5. The fractional frequency of the NIST H1 maser versus TA2M1.

$$\sigma^2_{\text{true}} = \frac{\sigma^2_{\text{measured}}}{1 - w} \tag{1}$$

where $w$ is the normalized weight of the clock in the ensemble. Since each clock had weight 1/3, the true variance is 1.5 times the measured variance. We see that the white FM level of the masers is about $1 \cdot 10^{-13}\tau^{-1/2}$, $\tau$ in s, reaching about $3 \cdot 10^{-15}$ at 1 d.

The drift of the three masers was measured against TA2M. The CH-B masers had a linear frequency drift value of $-0.9 \cdot 10^{-16}$/d with a confidence of $\pm 0.4 \cdot 10^{-16}$/d. The H1 and CH-A masers had negligible drift versus TA2M.

## V. Results from the TA2M1 Scale

Fig. 5 shows the fractional frequency of the maser H1 at Boulder, and Figs. 6 and 7 show the masers N4 and N5 at

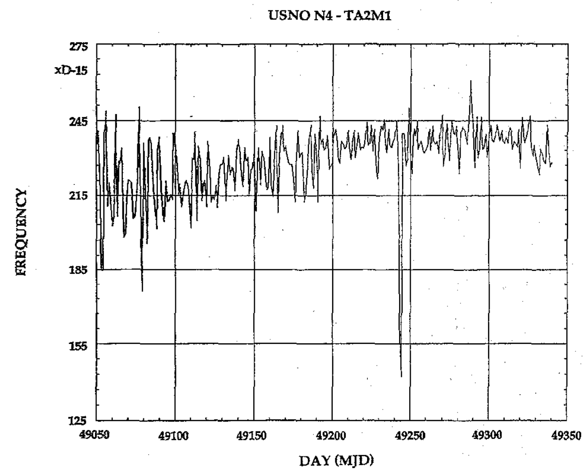



Fig. 6. The fractional frequency of the USNO N4 maser versus TA2M1.

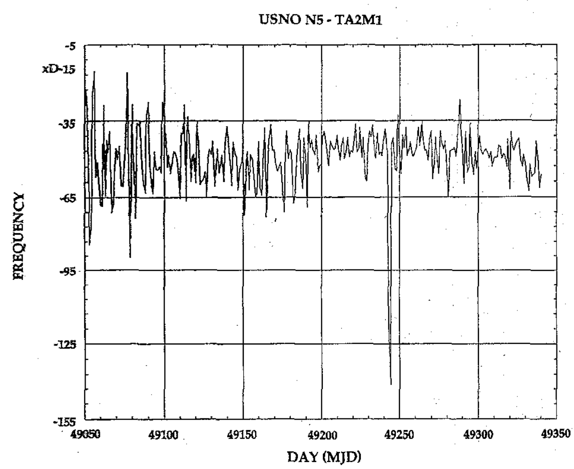



Fig. 7. The fractional frequency of the USNO N5 maser versus TA2M1.

USNO versus TA2M1. Since the masers CH-A and CH-B at Boulder were removed from the scale at MJD 49190, the GPS common-view noise became absorbed into the scale after that. Hence the apparent noise of H1 went up as the scale could no longer distinguish between maser noise and time transfer noise. Similarly, the noise of masers N4 and N5 went down at this time, as the GPS time transfer noise was no longer assigned to them. To minimize this effect we added another maser at NIST, H11, to the ensemble. Since this maser had no cavity autotuning we modelled its drift and weighted it very low in the long term. We obtained drift estimates by first including it in the scale with weight 0, allowing the scale to estimate its drift, then rerunning the scale with H11 included and drift modelled. In this way, when the drift values were used for predicting H11 in the TA2 algorithm, TA2M1 was not steered to any other scale, as could have occurred if the drift was originally measured against another scale.

We again found here that the linear frequency drift of the ensemble was too small to be measured against any available external frequency standard. The value of frequency drift for

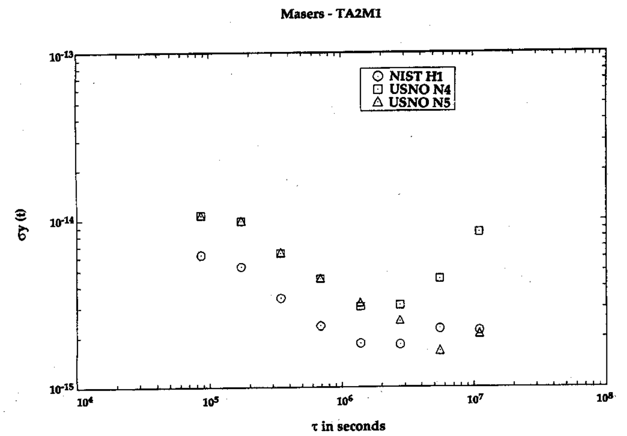



Fig. 8. The fractional frequency stability of the H1 maser at Boulder and the masers N4 and N5 at USNO versus TA2M1. The noise of GPS conceals the maser performance out to about 10 days. The N4 maser had a linear frequency drift of $1.0 \pm 0.3 \cdot 10^{-16}$/d.

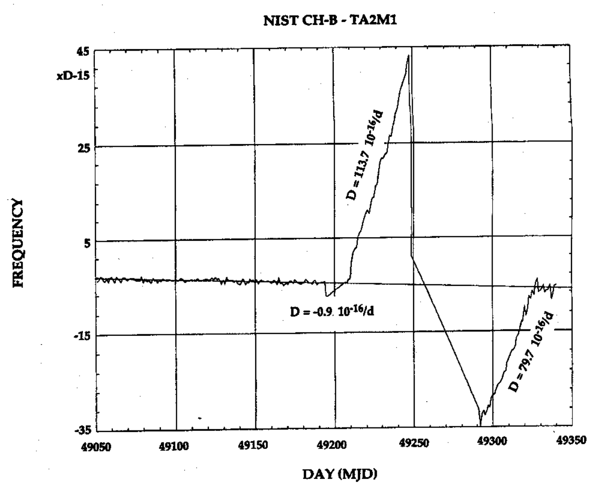



Fig. 9. The fractional frequency of the NIST CH-B maser versus TA2M1. Cavity tuning was turned off at 49190 and back on at 49336 with a different reference and outside its environmental cavity. In spite of this the standard returned to its predicted frequency within the uncertainty, as illustrated by the drawn line.

TA2M1-UTC over the 280 d interval MJD 49059–49339 for which data were available for the scales, was $-0.1 \cdot 10^{-16}$/d from a three-point second-difference estimator. If we assume the model of linear frequency drift in the presence of a flicker FM stochastic process at a level of $6 \cdot 10^{-15}$ for this period for UTC, the confidence on this drift estimate, comparing 150 d frequency averages is [12]

$$\frac{\sqrt{2}}{\tau}\sigma_y(\tau)|_{\tau=150\text{ d}} = 0.6 \cdot 10^{-16}/\text{d}. \qquad (2)$$

Thus, any linear frequency drift of TA2M1 against UTC was smaller than the uncertainty due to stochastic fluctuations according to this model.

Fig. 8 shows the stability of the masers against the ensemble. Comparing to Fig. 2 we see that the GPS time transfer noise dominates out to about 10 d. The N4 was the only maser with any apparent drift. It was estimated to be $1.0 \pm 0.3 \cdot 10^{-16}$/d using a three-point overall second-difference, with the confidence estimated as above. We want to point out that, although there is statistical significance to the magnitude of the N4 drift against UTC, we may not have a frequency standard accurate enough to determine whether the maser or UTC was drifting. The overall second-difference estimator compares average frequencies over two successive 150 d intervals. To determine average frequency offset from an SI accurate frequency over 150 d at a level of $(3/\sqrt{2}) \cdot 10^{-16}$ we need a frequency standard accurate to about $3 \cdot 10^{-14}$. This approaches the limits of the best available primary frequency standards today [13]–[15]. Thus we can only measure the change in frequency between the masers and UTC.

We were able to observe another significant effect using the TA2M1 scale. The maser CH-B lost maser CH-A as its reference for cavity tuning on 49190. We decided to use the H1 maser as a reference on MJD 49250, and the autotuning came back on for a few days, but unfortunately came unlocked again due to a power outage. Some time later, after moving masers to another room, we again relocked the CH-B autotuning to maser H1. Fig. 9 shows the performance of maser CH-B against the TA2M1 scale over the entire 340 d interval. When cavity tuning was reactivated, the maser came back to a frequency very close to what we would have predicted by frequency linearly using the constant drift of $-0.9 \cdot 10^{-16}$/d from the period 49050–49190.

## VI. Conclusion

One implication of these results is that a time scale based on three active autotuned hydrogen masers would be significantly more stable for integration times less than 1 to 2 months than our present NIST cesium ensemble. Such a maser ensemble would compete with the stability of UTC as it was in 1993. The stability of UTC will also change as it includes more new standards. In the very long term AT1, TA2M1, and UTC seem to flicker at a few parts in $10^{15}$.

Another conclusion we can draw is that the cavity autotuning systems eliminate the major cause of frequency drift in active masers. There may be some residual drift due to other effects, but at least for the masers during this experiment the drift relative to UTC was smaller than about $10^{-16}$/d. This number is only a relative drift, not a measure of drift against an accurate standard. The primary frequency standards of the world can only barely measure frequency drift at this level over the interval of this experiment. It is particularly interesting that maser CH-B came back to the predicted frequency after 100 d without autotuning, after being moved to a new room outside of its environmental chamber, and being locked to a different frequency reference.

Further, to make good use of the stability of these devices, advances in time and frequency comparisons will be necessary. Here the GPS common-view time transfer noise dominated over the masers out to about 10 d. This measurement noise can be lessened by using precise ephemerides and ionospheric measurements [16], [17], but over the approximately 1000 km baseline between NIST and USNO there will not be more than a factor of 2 or 3 improvement. Since UTC is generated based on 10 d measurements, we can see these new frequency standards, but barely. Being able to use such frequency standards reliably in UTC at 10 d will require better frequency comparison techniques.


## ACKNOWLEDGMENT

The authors would like to particularly thank G. Winkler for supplying the hydrogen-maser and commercial primary cesium-beam clock data from the USNO. This data was invaluable to the results obtained in the paper. Several significant suggestions have been made by J. Levine, S. Stein, D. Sullivan, and M. Young.


## REFERENCES


[1] D. W. Allan, "Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. UFFC-34, no. 6, pp. 647–654, Nov. 1987.

[2] F. L. Walls and H. Hellwig, "A new kind of passively operating hydrogen frequency standard," in *Proc. 30th Annu. Symp. Frequency Control*, Ft. Monouth NJ, June 1976, pp. 473–480.

[3] F. L. Walls and D. A. Howe, "Timekeeping potential using passive hydrogen masers," *J. Physique*, vol. 42, no. 12, pp. C8-151–C8-158, Dec. 1981.

[4] F. L. Walls, "Frequency standards based on atomic hydrogen," *Proc. IEEE*, vol. 74, no. 1, pp. 142–146, Jan. 1986.

[5] S. Stein, D. Glaze, J. Levine, J. Gray, D. Hilliard, D. Howe, and L. A. Erb, "Automated high-accuracy phase measurement system," in *Proc. IEEE Trans. Instrm. Meas.*, vol. IM–32, no. 1, pp. 227–231, Mar. 1983.

[6] A. Lepek and F. L. Walls, "Cross correlation analysis improves time domain measurements," in *Proc. 47th Annu. Symp. Frequency Control*, Salt Lake City, UT, June, 1993, pp. 312–320.

[7] M. A. Weiss and T. Weissert, "TA2, a time scale algorithm for post-processing: AT1 plus frequency variance," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 496–501, 1991.

[8] ——, "Sifting through nine years of NIST clock data with TA2," *Metrologia*, vol. 31, pp. 9–19, 1994.

[9] B. Guinot and C. Thomas, "Establishment of international atomic time," BIPM Annual Report, Part D, pp. D1–D22, 1988.

[10] C. Thomas, P. Wolf, and P. Tavella, *Time Scales*, Monograph 94/1, Bureau International des Poids et Mesures, Nov. 1994.

[11] P. Tavella, J. Azoubib, and C. Thomas, "Study of the clock-ensemble correlation in ALGOS using real data," in *Proc. 5th European Frequency and Time Forum*, Besancon, Mar. 1991, pp. 435–441.

[12] M. A. Weiss and C. Hackman, "Confidence on the three point estimator of frequency drift," in *Proc. 24th Annu. Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, McLean, VA, Nov. 1992, pp. 451–460.

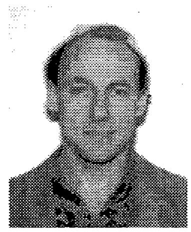
[13] R. E. Drullinger, J. H. Shirley, J. P. Lowe, and D. Glaze, "Error analysis of the NIST optically pumped primary frequency standard," *IEEE Trans. Instrum. Meas.*, vol. 42, pp. 453–456, 1993.

[14] A. Bauch *et al.*, "CS2: The PTB's new primary clock," *IEEE Trans. Instrum. Meas.*, vol. IM-36, pp. 613–616, June, 1987.

[15] A. Bauch *et al.*, "The new PTB new primary cesium clocks," *IEEE Trans. Instrum. Meas.*, vol. 42, pp. 444–447, Apr. 1993.

[16] W. Lewandowski, G. Petit and C. Thomas, "Precision and accuracy of GPS time transfer," *IEEE Trans. Instrum. Meas.*, vol. 42, no. 2, p. 474, 1993.

[17] W. Lewandowski and M. A. Weiss, "Precise ephemerides for GPS time transfer," in *Proc. 21st Precise Time and Time Interval Applications and Planning Meeting*, Redondo Beach, CA, 1989, pp. 95–106.

**Marc A. Weiss** received the B.S. degree from Valparaiso University, Valparaiso, IN, in 1973, and the M.S. degree in mathematics and the Ph.D. degree in mathematical-physics, from the University of Colorado, Boulder, in 1975 and 1981, respectively.

Since 1978, Dr. Weiss has worked at the National Institute of Standards and Technology (NIST, formerly the National Bureau of Standards, NBS), Boulder CO. He wrote the software for the NBS/GPS Time Transfer System for which he received the Applied Research Award of the NBS in 1983, along with the other principals. He also designed and wrote the software which operates the NIST Automated Computer Time Service system, and the NIST Ionospheric Measurement System. He has been active in studying and developing time transfer systems especially using the Global Positioning System. He has also specialized in new time scale algorithms, and in synchronization in telecommunications systems.

**Fred L. Walls** (SM'94) was born in Portland, OR, on October 29, 1940. He received the B.S., M.S. and Ph.D. degrees in physics from the University of Washington, Seattle, in 1962, 1964, and 1970, respectively. His Ph.D. thesis was on the development of long-term storage and nondestructive detection techniques for electrons stored in Penning traps and the first measurements of the anomalous magnetic $(g-2)$ moment of low energy electrons.

From 1970 to 1973, he was a Postdoctoral Fellow at the Joint Institute for Laboratory Astrophysics in Boulder, CO. This work focused on developing techniques for long-term storage and nondestructive detection of fragile atomic ions stored in Penning traps for low energy collision studies. Since 1973, he has been a Staff Member of the Time and Frequency Division of the National Institute of Standards and Technology, formerly the National Bureau of Standards, Boulder, CO. He is presently Leader of the Phase Noise Measurement Group and is engaged in research and development of ultra-stable clocks, crystal-controlled oscillators with improved short- and long-term stability, low-noise microwave oscillators, frequency synthesis from RF to infrared, low-noise frequency stability measurement systems, and accurate phase and amplitude noise metrology. He has published more than 110 scientific papers and holds 5 patents.

Dr. Walls received the 1995 European "Time and Frequency" award from the Societe Francaise des Microtechniques et de Chromometrie "for outstanding work in ion storage physics, design and development of passive hydrogen masers, measurements of phase noise in passive resonators, very low noise electronics and phase noise metrology". He is the recipient of the 1995 IEEE Rabi award for "major contributions to the characterization of noise and other instabilities of local oscillators and their effects on atomic frequency standards". He has also received two silver medals from the U.S. Department of Commerce for fundamental advances in high resolution spectroscopy and frequency standards, and the development of passive hydrogen masers. He is a member of the American Physical Society, a member of the Technical Program Committee of the IEEE Frequency Control Symposium, and a member of the Scientific Committee of the European Time and Frequency Forum.