International Comparison of Atomic Frequency Standards Via VLF Radio Signals[1]

A. H. Morgan, E. L. Crow, and B. E. Blair

National Bureau of Standards, Boulder, Colo.

A study was made of data obtained over an 18-month period (July 1961 to December 1962, inclusive) on the comparison of atomic frequency standards located in seven laboratories in the United States, Europe, and Canada, using the VLF signals of GBR (16 kc/s), Rugby, England, and NBA (18kc/s), Balboa, Canal Zone. Each laboratory observes the accumulated difference in phase over a 24-hr period (the same for all laboratories, or nearly so) between its own standard (either laboratory or commercially constructed) and the received VLF signal. A statistical analysis was designed to separate the observations at each laboratory into three components: (a) long-term mean differences among the atomic standards; (b) estimates of the standard deviations, $\hat{\alpha}_i$, at each receiving station; and (c) estimates of the transmitter standard deviations, $\hat{\tau}$. Each $\hat{\alpha}_i$ includes receiver fluctuations, propagation effects perculiar to the path, and measurement uncertainties; $\hat{\tau}$ includes the transmitter fluctuations and propagation effects common to all paths.

The study shows that $\hat{\alpha}_i$ at each receiver varied from a low of 0.39×10^{-10} units of fractional frequency (that is, 0.39 parts in 10^{10}) (GBR data) at LSRH to a high of 1.97×10^{-10} (GBR data) at NRC with an average for all stations of 1.01×10^{-10} measured against GBR and 0.99×10^{-10} when measured against NBA. Also, the average $\hat{\tau}$ for GBR is 1.26×10^{-10} and for NBA is 0.68×10^{-10}. Finally, it is shown that: (1) the means of the frequencies of the seven individual laboratories agreed with the grand mean of these seven laboratories to within ± 2 parts in 10^{10} for the 18-month period, and (2) the laboratory-type standards agreed with their grand mean to within ± 1 part in 10^{10}.

1. Introduction

The agreement between atomic frequency standards of varied construction in many laboratories distributed widely over the world is noteworthy. This agreement establishes confidence in such atomic devices as standards of time and provides a basis for defining an atomic second [NBS, 1964]. This present paper tests the agreement between atomic frequency standards through the medium of VLF radio signals and derives measures of their individual precisions.

The atomic frequency standards of seven laboratories located in the United States, Europe, and Canada were compared by means of a fairly complete statistical analysis of data gathered and exchanged over an 18-month period, July 1961 through December 1962. Each laboratory operates one or more atomic standards, which may be either laboratory designed and constructed or commercially constructed, and makes daily measurements of the received phase of stabilized VLF signals from stations GBR and NBA. Each then reports the deviations of the frequencies of the received signals from nominal (i.e., the rated frequency as defined by its own atomic standard). The reported values are 24-hr averages centered at about the same value of UT.

Six of the laboratories involved are: Centre National d'Études des Télécommunications (CNET), Bagneux, Seine, France; Cruft Laboratories (CRUFT), Harvard University, Cambridge, Mass.; Laboratoire Suisse de Recherches Horlogères (LSRH), Neuchatel, Switzerland; National Bureau of Standards (NBS), Boulder, Colo.; National Physical Laboratory (NPL), Teddington, Middlesex, England; National Research Council (NRC), Ottawa, Ontario, Canada. Also included are the data obtained by the U.S. Naval Observatory (NOB), Washington, D.C., which obtains a "mean atomic standard" by using weighted results from nine laboratories including the above six [Markowitz, 1962].

A number of comparisons of atomic frequency standards by radio transmissions, VLF transmissions in particular, have already been published [e.g., Pierce, 1957; Essen, Parry, and Pierce, 1957; Holloway, Mainberger, Reder, Winkler, Essen, and Parry, 1959; Pierce, Winkler, and Corke, 1960; Markowitz, 1961; Richardson, Beehler, Mockler, and Fey, 1961; Essen and Steele, 1962; Mitchell, 1963; Markowitz, 1964]. The summary statistics of these papers have been primarily mean differences and standard deviations of differences. In addition, Mitchell analyzed data from pairs of laboratories to separate out standard deviations associated with each laboratory and with the transmitter. The present paper provides an analysis similar to Mitchell's applicable to any number of laboratories simultaneously.

2. Stabilized Very-Low-Frequency Transmissions Used

The stabilized VLF transmissions of GBR (16 kc/s), Rugby, England, and NBA (18 kc/s), Balboa, Panama Canal Zone, were used by all the laboratories concerned to obtain the comparison data. The VLF transmitters are directly controlled by oscillators, which require regular periodic calibration and adjustment, and give transmitted stabilities (standard deviations of fractional frequency) of the order of 1 or 2 parts in 10^{10}. The transmitted carrier frequencies are held as constant as possible with reference to an atomic standard but are offset in frequency so that the transmitted time pulses may be kept in closer agreement with the UT–2 scale. The amount of fractional frequency offset is determined in advance for each year by the Bureau International d l'Heure, Paris, France, through cooperative efforts of several astronomical observatories throughout the world, and is based on the assumed value of the cesium resonance of 9,192,631,770 c/s.

3. Brief Description of Measurement Techniques

Figure 1 shows the locations of the laboratories participating in these studies and the radio transmis-

sion paths to each from GBR and NBA. Each laboratory maintains one or more atomic frequency standards, makes daily measurements of GBR and NBA, and reports the deviation of the received VLF signals from nominal (as indicated by the local atomic frequency standards). Table 1 lists the location, type, and stated accuracy of such standards at the receiving laboratories. Usually the received phase of each transmission is recorded in terms of a local, reference quartz oscillator which, in turn, is calibrated periodically by an atomic frequency standard.
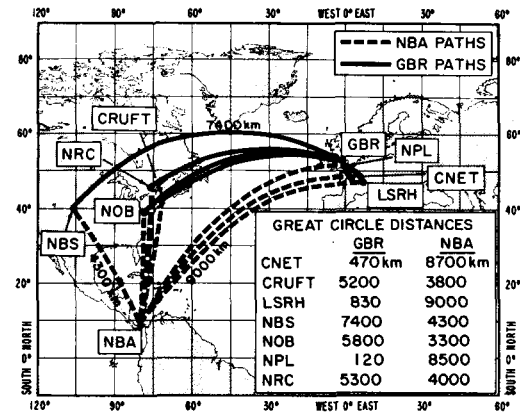

FIGURE 1. *Location of transmitters and receiving laboratories.*

Table 1. Characteristics of Laboratory Atomic Standards

(These are not complete. In particular, the averaging time with which the accuracy or precision should be associated is not listed.)

Laboratory	Abbreviation	Approximate Geographic Location	Type of Atomic Standard	Reported Accuracy x 10^{-10}	Reported Precision [*] x 10^{-11}	Reference for Accuracy and Precision	Method of Distributing Results by Laboratory Concerned
Centre National d'Études des Télécommunications Bagneux, Seine, France	CNET	48° 48' N 2° 19' E	Cs Beam (Atomichron #107)	+2.2 (Sys. Dev. from Avg. of 6 labs)	2-3	Decaux [1963]	L'ONDE Electrique and monthly circulation
Cruft Laboratories Harvard University Cambridge, Mass., USA	CRUFT	42° 22' N 71° 6' W	Cs Beam (Atomichrons #202; #112)	5 (Mfg.'s guarantee)	2	Pierce [1963]	Monthly circulation
Laboratoire Suisse de Recherche Horlogère Neuchatel, Switzerland	LSRH	47° 00' N 6° 57' E	Cs Beam	0.3	2.7	Kartaschoff [1962]	Observatoire de Neuchatel Bulletin Série c - monthly
National Bureau of Standards Boulder, Colorado, USA	NBS (USFS) (U. S. Frequency Standard)	40° 00' N 105° 16' W	Cs Beam	0.1	0.2	Beehler et al. [1962]	NBS - FM monthly report [***] from Frequency-Time Dissemination Research Section
National Physical Laboratory Teddington, Middlesex England	NPL	51° 25' N 0° 20' W	Cs Beam	0.3	?	Essen & Steele [1962],NPL [1963; 1964]	Electronic Technology and monthly circulation
National Research Council Ottawa, Ontario, Canada	NRC	45° 25' N 75° 43' W	Cs Beam	3 [**]	-	Kalra [1961]	Monthly circulation
U. S. Naval Observatory Washington, D. C., USA	NOB	38° 55' N 77° 39' W	Cs Beam (weighted average of 9 Cs resonators including Atomichrons)	0.2	-	Markowitz [1962]	Time Service Bulletins and weekly circulation

[*]Standard deviation is often but not always specified.

[**]The term "stability" was used, so that this result perhaps should be in the "precision" column.

[***]About 10 per cent more days of data were used in the present analysis than was reported in the FM monthly reports.
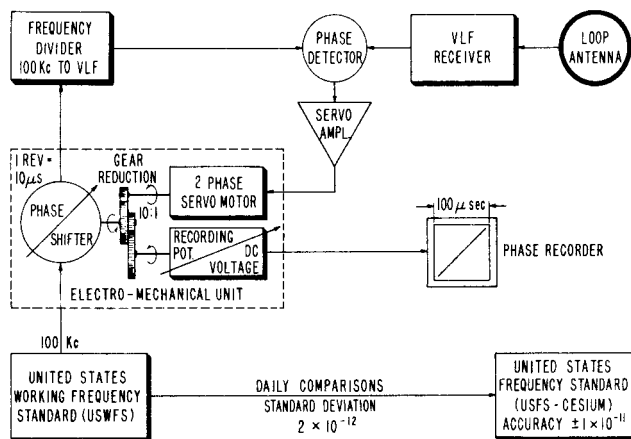
FIGURE 2. *Block diagram of typical VLF phase comparison at NBS, Boulder, Colo.*

One prevalent VLF measuring system employs a phase-lock receiver such as shown in the simplified diagram of figure 2 [Morgan and Andrews, 1961]. In such a system the servodriven phase-shifter continuously phase locks a synthesized signal from the local standard to the received VLF signal. A linear potentiometer, connected to a constant direct voltage, generates the voltage analog of the phase-shifter position. The sensitivity of the voltage analog recorder is determined by the gear ratio between the potentiometer shaft and the phase-shifter shaft. The overall frequency bandwidth of a typical phase-lock servo-system is in the range of 0.01 to 0.001 c/s. The maximum fractional frequency difference between an incoming signal and the local standard that may be tracked is usually near 5 parts in 10^8 (or a change of about 3 μsec per minute).

The measuring system of figure 2 (located at NBS) produces phase records of width 4.5 in. for the full scale sensitivity of 100 μsec, with coordinate lines at intervals of 2 μsec. The time scale, controlled by the recorder speed, normally provides ¼ in. per 20 min, which is the interval between coordinate lines. Measurements on the phase records are made during the time the propagation path is sunlit and phase fluctuations are minimal. The duration of such a quiet period varies with the seasons; it ranges, however, from about 2 hr to 8 or 10 hr. (The standard deviation of phase fluctuations in NBA recordings at NBS during the daytime was found to be several tenths of a μsecond for a series of 20-min measurements taken over a 7-hr period.)

The error of observing the accumulated phase at each laboratory has not been completely evaluated; any measurement system, however, introduces measurement error. The measuring system may also introduce a smoothing or averaging effect, especially if it produces a continuous record. Such smoothing may reduce fluctuations of interest, from frequency standards and propagation in the present case, as well as measurement error. Fluctuations of periods as short as 5 min are visible on the NBS records;

this implies that the averaging time of the system is of the order of 2 min or less. Consequently, only a negligible proportion of the accumulated phase difference over 24 hr could be averaged out.

The phase records from the system of figure 2 are read to the nearest μsec. (They could be read more closely, but the improvement might be marginal relative to the other errors.) The resulting maximum reading error of 0.5 μsec corresponds to 6 parts in 10^{12} over a 24-hr period. Taking account of the two independent readings, initial and final, needed to produce the 24-hr increment in phase and assuming a uniform distribution of errors up to the maximum, leads to the figure of 0.05 parts in 10^{10} for the standard deviation of the fractional frequency due to reading error. This is a contribution to the estimated standard deviation $\hat{\alpha}_i$ of the local atomic standard defined in section 4 and evaluated in figures 5 and 6, but it is an order of magnitude less than the total. There may be appreciable measurement error beyond the reading error. No attempt has been made to survey the measurement errors of the laboratories other than NBS.

The reported frequency values, obtained from phase measurements at each laboratory, represent 24-hr averages centered at the same value of UT, 0300 UT, except for one set of values—those reported by NBS for NBA transmissions. These were centered at 0600 UT to avoid diurnal side effects during the winter sunrise periods.

4. Methods of Statistical Analysis

The purpose of the statistical analysis is to attempt to separate the relative observations at each laboratory into components associated with: (a) the long-term mean differences between the atomic standards; (b) effects of the fluctuations of the receiving system, propagation effects peculiar to the particular radio path, and measurement errors; and (c) fluctuations of the transmitter signals and propagation effects common to all the radio paths.

The daily values of fractional frequency differences were placed into a matrix with the columns representing laboratories and the rows representing days. Each matrix contained one quarter of a year of data, so that there were six matrices included in the 18-month period of study. Each matrix is regarded in the analysis as a sample of observations from an infinite ensemble or population of daily values from the given laboratories. The method required that no data be missing in any cell of the matrix; therefore, when any laboratory omitted a daily value, the complete row of data for all laboratories was discarded. This admittedly reduced the amount of usable data, but the average number of days remaining per quarter, about 40, is believed to be sufficient.

The analysis of the relative frequency observations at several receiving laboratories into components associated with each laboratory standard and the transmitter can be accomplished in several ways, of which a relatively simple one will be described. The expectation of a random variable, or average over a

population, will be denoted by $E[\ \ ]$. This and other statistical concepts may be found in books such as that by Cramér [1946].

Let k be the number of receiving laboratories and n the number of days that observations are made by each laboratory. Let x_{ij} be the observed fractional difference in frequency of received and local standard signals, in parts in 10^{10}, at the ith receiving laboratory on the jth day ($i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n$). The observations x_{ij} include (1) any systematic differences between the ith local standard and the oscillator used by the transmitter; (2) fluctuations associated only with the ith standard; (3) fluctuations associated only with the transmitter oscillator, and (4) radio propagation fluctuations. Thus we can represent x_{ij} by the equation

$$x_{ij} = \mu_i + a_{ij} + t_j (i = 1, 2, \ldots, k; \quad (1)$$

$$j = 1, 2, \ldots, n),$$

where

$\mu_i = $ systematic (long-term) mean fractional frequency difference at the ith receiver (i.e., $\mu_i = E[x_{ij}]$;

$a_{ij} = $ daily fluctuations of fractional frequency difference associated with the ith receiver on the jth day, in particular fluctuation of its standard but including also propagation effects peculiar to the ith path and measurement errors;

$t_j = $ daily fluctuation of the transmitted signal, including any effects, propagation in particular, common to all received signals.

Thus $E[a_{ij}] = E[t_j] = 0$, and we assume that a_{ij} and t_l are uncorrelated, so that $E[a_{ij}t_l] = 0$. Likewise we assume that $E[a_{ij}a_{hl}] = 0$, $i \neq h$ or $j \neq l$, and $E[t_jt_l] = 0$, $j \neq l$.

The systematic mean frequency difference μ_i is easily estimated by the sample mean at the ith receiver,

$$\hat{\mu}_i = x_{i.} = \frac{1}{n} \sum_{j=1}^{n} x_{ij} (i = 1, 2, \ldots, k), \quad (2)$$

where the circumflex accent denotes "estimate of." It is known that $\hat{\mu}_i$ is a best estimate in the sense that it is an unbiased estimate of μ_i and has variance less than that of any other unbiased estimate that is a linear combination of the observations.

Following in part the notation of Mitchell [1963], we let α_i be the "true" (long-term) standard (root-mean-square) deviation of the a_{ij}, associated with the ith receiver, and τ the true standard deviation of the t_j, associated with the transmitter. Variance being defined as the square of a standard deviation, we may for brevity refer to α_i^2 as the ith receiver variance and τ^2 as the transmitter variance. Let σ_i be the true standard deviation of the observations x_{ij} at the ith receiver. Then it follows from (1) and the above assumptions that

$$\sigma_1^2 = \alpha_1^2 + \tau^2, \ \sigma_2^2 = \alpha_2^2 + \tau^2, \ \ldots, \ \sigma_k^2 = \alpha_k^2 + \tau^2. \quad (3)$$

The purpose of further analysis is to estimate α_1, $\alpha_2, \ldots \alpha_k$, and τ from the kn observations x_{ij}. Let the sample variance of the n observations from the ith receiver be denoted by

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} (x_{ij} - \overline{x}_{i.})^2. \quad (4)$$

Since s_i^2 is an unbiased estimate of σ_i^2, substitution in (3) gives k equations for the $k+1$ unknowns $\hat{\alpha}_1$, $\hat{\alpha}_2$, ..., $\hat{\alpha}_k$, $\hat{\tau}$, where $\hat{\alpha}_i^2$ and $\hat{\tau}^2$ are unbiased estimates of α_i^2 and τ^2. The one required additional equation is furnished by calculating the means

$$\overline{x}_{.j} = \frac{1}{k} \sum_{i=1}^{k} x_{ij} \ (j = 1, 2, \ldots, n) \quad (5)$$

over all k receivers for each day and then calculating the sample variance of these averages,

$$s_0^2 = \frac{1}{n-1} \sum_{j=1}^{n} (\overline{x}_{.j} - \overline{x})^2, \quad (6)$$

where $\overline{x}$ is the mean of all kn observations. It can be shown that s_0^2 is an unbiased estimate of

$$\sigma_0^2 = \tau^2 + \frac{1}{k^2} (\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_k^2). \quad (7)$$

Solving (3) and (7) and substituting estimates for true values, we obtain the unbiased estimates

$$\hat{\tau}^2 = \frac{k}{k-1} \left(s_0^2 - \frac{1}{k^2} \sum_{i=1}^{k} s_i^2 \right), \quad (8)$$

$$\hat{\alpha}_i^2 = s_i^2 - \hat{\tau}^2 (i = 1, 2, \ldots, k). \quad (9)$$

For $k = 2$, that is, two receivers, the estimates (8) and (9) reduce essentially to Mitchell's estimates [1963].

The theoretical precisions of the estimates (8) and (9) have been derived under the assumption of independence of the measurements from day to day; it is hoped to include formulas for these precisions in a further paper, along with approximations for the degree of dependence. The precision of $\hat{\alpha}_i$ (and of $\overline{x}_{i.}$) would appear in theory to be improved by using all observations available at the ith receiver (including those on days when other receivers provide no data) to calculate s_i^2, but s_0^2 can be obtained only by using days common to all receivers.

The above model for separating $\hat{\alpha}_i$ and $\hat{\tau}$ is in effect included in that of Grubbs [1948] for separating measurement errors of several instruments from the product variability which the instruments are measuring.

5. Discussion of Results

Mean values. The mean fractional differences in frequency recorded at each receiver from both GBR and NBA were calculated for each month and for the entire 18-month period, using all the daily observations.

Likewise a grand mean over all months and receivers was found for GBR to be − 130.06 parts in 10^{10} relative to the assumed cesium resonance frequency of 9,192,631,770 c/s. The corresponding grand mean for NBA was − 129.78. The 18-month means, $\bar{x}_{i.}$, deviations from grand mean, $\bar{x}_{i.} − \bar{x}$, and standard deviations, s_i, of daily recorded values over the 18-month period are listed in table 2 for GBR and in table 3 for NBA.

The mean deviations of individual laboratories from the grand mean range from about + 2 to − 1 parts in 10^{10} measured against either transmission and in general appear to be statistically significantly different from 0 due to the large number of observations. The statistical significance cannot be greater than that based on regarding each daily observation as statistically independent of those for other days (whereas in fact there is some autocorrelation). Based on this assumption and the standard deviations and numbers of daily observations in tables 2 and 3, the standard deviations of the means range from 0.04 to 0.19 (parts in 10^{10}), and it would follow that all values of $\bar{x}_{i.} − \bar{x}$ except those for NPL in tables 2 and 3 would be significantly different from zero at the 5 percent probability level.

However, a more refined test of significance of differences of means of the "common" data is possible by means of a two-factor analysis of variance with standards and days as factors, which eliminates the transmitter and common propagation variation (and hence also any autocorrelation due to these) [Cramér, 1946, pp. 543–546]. This more refined test still neglects the effect of any autocorrelation in individual

standards. If this neglect is kept in mind, it is still impressive that all such tests result in significance beyond the 0.1 percent (0.001) probability level. Hence it is believed that even if any autocorrelation in individual standards is accounted for, the values of $\bar{x}_{i.} − \bar{x}$ will prove to be significantly different from zero. In other words, the observed average differences of standards from their grand mean are concluded to be real systematic differences, though a few standards differ little in pairs, like NOB and NPL.

The reality of the differences between laboratories is further shown by the mean deviations from the grand mean over the years 1959–60 calculated from the data of Essen and Steele [1962] and given in the "ES" column of tables 2 and 3. These indicate that systematic differences between many of the laboratories continued over the years 1959 or 1960 through 1962.

The difference in grand mean frequencies received from GBR and NBA from all data reported in the 18-month period, − 0.28 parts in 10^{10}, is replaced by − 0.15 parts in 10^{10} when only the "common" (C) data are considered. If a formal Student t test of the difference of two means is made using standard errors derived from (6), the difference − 0.28 is found to be significant at the 5 percent probability level, whereas the difference − 0.15 is not. Since significance of the differences would be lessened by taking account of the effect of autocorrelation, no significant average difference between the transmitters GBR and NBA is claimed.

The monthly means of "common" data for each laboratory are shown in figures 3 and 4; these figures

Table 2. Means and Standard Deviations of Frequencies Compared with GBR

July 1961 - December 1962

(unit, 1 part in 10^{10})

Atomic Standard	No. of Days		$\bar{x}_{i.}$ Mean		$\bar{x}_{i.} − \bar{x}$ Deviation from Grand Mean			s_i Standard Deviation of Daily Values	
	T	C	T	C	T	C	ES	T	C
CNET	478	278	-128.27	-128.14	+1.79	+1.68	+1.52	1.73	1.77
CRUFT	388	136	-131.14	-130.60	-1.06	-0.78	-1.91	1.87	1.39
LSRH	422	244	-130.99	-130.86	-0.93	-1.04	-1.22	1.48	1.49
NBS	469	244	-130.70	-130.50	-0.64	-0.68	-	1.56	1.55
NOB	486	278	-130.28	-130.16	-0.22	-0.34	-	1.49	1.51
NPL	476	278	-129.93	-129.77	+0.13	+0.05	+0.41	1.59	1.66
NRC	458	222	-129.37	-129.17	+0.69	+0.65	+1.39	2.19	2.35
GRAND MEAN, $\bar{x}$			-130.06	-129.82					

Note: 1. In applying the notation $\bar{x}_{i.}$, s_i, and $\bar{x}$ to all data, the definitions in equations (2), (4), and that for $\bar{x}$ under equation (6) are understood to be extended to unequal numbers of days, n_i, for each receiver.

2. T indicates total data reported; C indicates data common to stations within quarters. ES denotes mean deviations from grand mean (with signs appropriately reversed) calculated from Table 2 of Essen and Steele (1962) for the years 1959-60, except that only last half of 1960 is given for LSRH and NRC.

Table 3. Means and Standard Deviations of Frequencies Compared with NBA

July 1961 - December 1962

(unit, 1 part in 10^{10})

Atomic Standard	No. of Days		$\bar{x}_{i.}$ Mean		$\bar{x}_{i.} - \bar{x}$ Deviation from Grand Mean			s_i Standard Deviation of Daily Values	
	T	C	T	C	T	C	ES	T	C
CNET	427	197	-127.90	-127.78	+1.88	+1.89	+1.72	1.43	1.38
CRUFT	385	131	-131.00	-130.72	-1.22	-1.05	-1.45	1.66	1.66
LSRH	437	197	-130.74	-130.66	-0.96	-0.99	-1.05	1.14	1.17
NBS	369	161	-130.26	-130.33	-0.48	-0.66	-	0.99	0.96
NOB	474	197	-129.89	-129.94	-0.11	-0.27	+0.53	0.89	0.88
NPL	375	102	-129.98	-129.95	-0.20	-0.28	-0.17	1.43	1.45
NRC	363	113	-128.75	-128.37	+1.03	+1.30	-	1.96	1.97
GRAND MEAN, $\bar{x}$			-129.78	-129.67					

Note: 1. In applying the notation $\bar{x}_{i.}$, s_i, and $\bar{x}$ to all data, the definitions in equations (2), (4), and that for $\bar{x}$ under equation (6) are understood to be extended to unequal numbers of days, n_i, for each receiver.

2. T indicates total data reported; C indicates data common to stations within quarters. ES denotes mean deviations from grand mean (with signs appropriately reversed) calculated from Table 2 of Essen and Steele (1962) for July-December 1960.

show that not only are observations for individual days missing for particular laboratories but even series for entire months or more. The variability of monthly means is visibly greater than that of the 18-month means reported above, but variations in laboratory data common to all laboratories for a given transmitter are apparent. Similar results were given by Mitchell [1963, fig. 4] for the same period for LSRH, NBS, and NPL.

Standard deviations. The standard deviations s_i of tables 2 and 3 are estimates of the σ_i of (3). They tend not to vary greatly for a given transmitter because each contains the common variation of the transmitter and common propagation effects. The standard deviation of the common effects, $\hat{\tau}$, is plotted for each quarter and each transmitter in figures 5 and 6. The component standard deviation associated with each receiver, $\hat{\alpha}_i$, is also similarly plotted. A comparison of $\hat{\alpha}_i$ for each receiver as determined from both the GBR and NBA transmissions is plotted in figure 7.

The quarterly receiver (local standard) and transmitter (NBA and GBR) standard deviations are all roughly of the same order of magnitude, 1 part in 10^{10}, but there is substantial variation, between transmitters, from receiver to receiver, and from quarter to quarter. Thus the quarterly $\hat{\alpha}_i$ vary from a low of zero to a high of 2.24 (parts in 10^{10}), and the quarterly $\hat{\tau}$ vary from 0.28 to 1.01 for NBA and from 0.73 to 1.48 for GBR.

Whether these variations are significant, that is, reflect differences among the true values α_i and τ, or are only to be expected statistically in estimates from small samples, can be judged from the standard deviations of the estimates, which have as yet been evaluated only under the assumption of independence from day to day. Even under this assumption each α_i and τ is estimated from all 18 months of data with 95 percent confidence limits roughly 15 percent below and 20 percent above the estimates $\hat{\alpha}_i$ and $\hat{\tau}$, except that the four smaller $\hat{\alpha}_i$ obtained from GBR data (Cruft, NBS, LSRH, NOB) have larger percentage errors, with 95 percent confidence limits up to 100 percent above or below the estimate $\hat{\alpha}_i$. The effect of day-to-day dependence is to extend these limits even farther. However, it is evident that some receivers have consistently, and hence probably significantly, lower values of α_i than others.

Since the above confidence limits show that much of the fluctuating behavior of $\hat{\alpha}_i$ and $\hat{\tau}$ is not significant, it is reasonable to compute weighted root-mean-square average values of $\hat{\alpha}_i$ and $\hat{\tau}$ over all quarters; these averages are tabulated in figures 5 and 6. The smallest average $\hat{\alpha}_i$ are 0.51 when measured against NBA transmissions, obtained for the Naval Observatory, and 0.39 when measured against GBR, obtained for LSRH. The corresponding largest average $\hat{\alpha}_i$ are 1.82 and 1.97, both obtained for the National Research Council of Canada.

It is of interest to compare these results with those obtained by Mitchell [1963] over the same six quarters of 1961 and 1962. He compared three receivers, NBS (USFS), NPL, and LSRH, in pairs using both NBA and GBR transmissions. The comparative re-
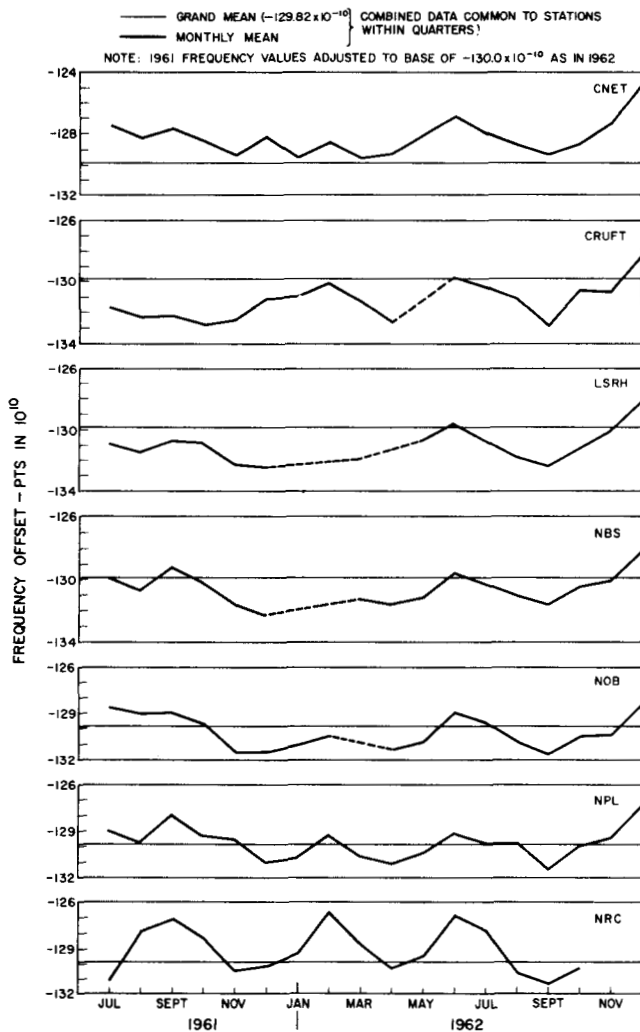
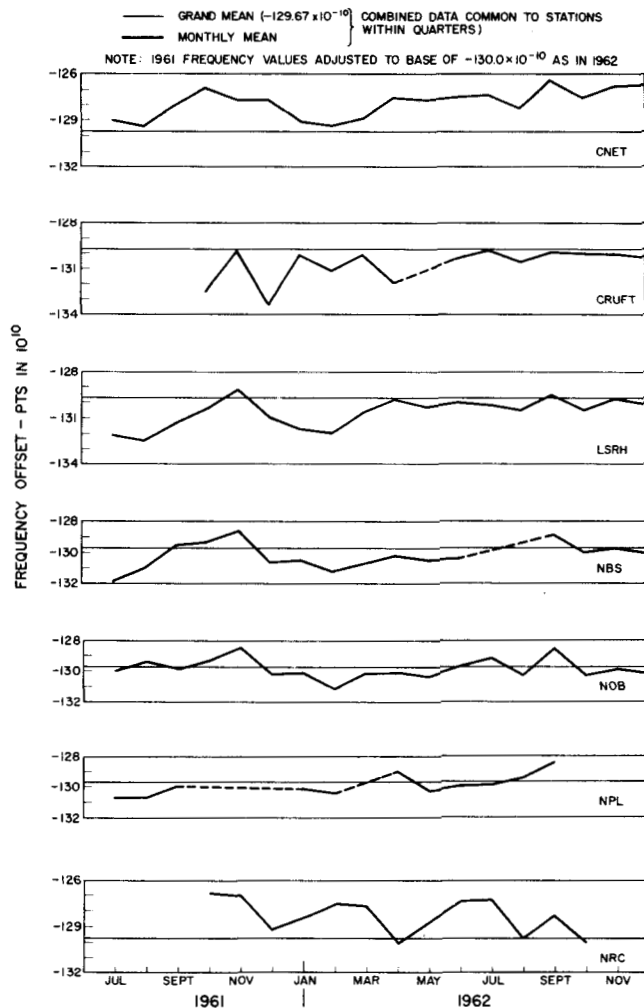FIGURE 3. *GBR monthly means versus grand mean of all laboratory standards.*

FIGURE 4. *NBA monthly means versus grand mean of all laboratory standards.*

sults are as follows (standard deviations in parts in 10^{10}):

		Present analysis		Mitchell (average)
		GBR	NBA	
NBS (USFS)	$\hat{\alpha}_1$	0.66	0.63	0.7
LSRH	$\hat{\alpha}_2$	.39	.63	.6
NPL	$\hat{\alpha}_3$	1.00	1.24	1.1
Transmitter	$\hat{\tau}$	1.26	0.68	1.1 (GBR), 0.7 (NBA)

The check between the two sets of results is almost complete, to the expected precision. This is not surprising, since essentially the same data for these receivers are used. However, the present analysis should tend to yield a more precise estimate of τ by virtue of using 5 to 7 receivers for its estimation,

whereas Mitchell's method should tend to yield more precise values of s_i by virtue of having fewer missing data in the pair-by-pair treatment (although a general treatment is possible using all data). The average number of daily pairs of observations per quarter used by Mitchell was about 60, whereas the number of daily sets of observations per quarter used herein is about 40. The principal justification of the present method is the simultaneous analysis of data from any number of receiving laboratories.

Since the $\hat{\alpha}_i$ include propagation as well as local standard fluctuations and $\hat{\tau}$ includes possibly fluctuations from sources other than the transmitter oscillator, it is desirable to attempt further separation. This should be possible from internal estimates of precision of local standards but has not yet been carried through. Mitchell [1963] estimated the standard deviation of transmission time fluctuations in transatlantic comparisons of cesium-controlled oscillators at 0.2×10^{-10}, which would leave the standard devia-
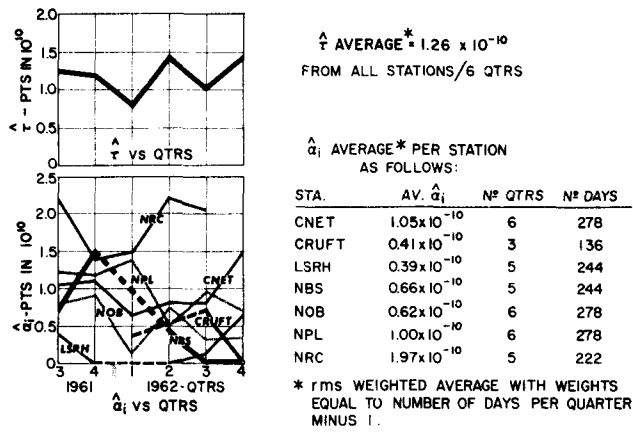
319-911

FIGURE 5. *Estimated standard deviations, $\hat{\alpha}_i$, of receiver variations (including noncommon propagation variations) and estimated standard deviation, $\hat{\tau}$, of variations of transmitter GBR (including common propagation variations).*

For Figure 5:

$\hat{\tau}$ AVERAGE* = 1.26 x 10⁻¹⁰
FROM ALL STATIONS/6 QTRS

$\hat{\alpha}_i$ AVERAGE* PER STATION AS FOLLOWS:

STA.	AV. $\hat{\alpha}_i$	Nº QTRS	Nº DAYS
CNET	1.05×10^{-10}	6	278
CRUFT	0.41×10^{-10}	3	136
LSRH	0.39×10^{-10}	5	244
NBS	0.66×10^{-10}	5	244
NOB	0.62×10^{-10}	6	278
NPL	1.00×10^{-10}	6	278
NRC	1.97×10^{-10}	5	222

* rms WEIGHTED AVERAGE WITH WEIGHTS EQUAL TO NUMBER OF DAYS PER QUARTER MINUS 1.
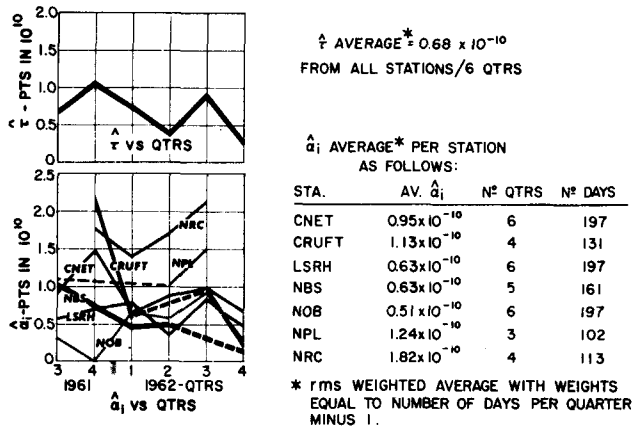

FIGURE 6. *Estimated standard deviations, $\hat{\alpha}_i$, of receiver variations (including noncommon propagation variations) and estimated standard deviation, $\hat{\tau}$, of variations of transmitter NBA (including common propagation variations).*

For Figure 6:

$\hat{\tau}$ AVERAGE* = 0.68 x 10⁻¹⁰
FROM ALL STATIONS/6 QTRS

$\hat{\alpha}_i$ AVERAGE* PER STATION AS FOLLOWS:

STA.	AV. $\hat{\alpha}_i$	Nº QTRS	Nº DAYS
CNET	0.95×10^{-10}	6	197
CRUFT	1.13×10^{-10}	4	131
LSRH	0.63×10^{-10}	6	197
NBS	0.63×10^{-10}	5	161
NOB	0.51×10^{-10}	6	197
NPL	1.24×10^{-10}	3	102
NRC	1.82×10^{-10}	4	113

* rms WEIGHTED AVERAGE WITH WEIGHTS EQUAL TO NUMBER OF DAYS PER QUARTER MINUS 1.

tions estimated above essentially unaltered (due to the combination by squares).

The mathematical model (1) implies the condition that all daily measurement intervals are simultaneous, whereas the NBS measurements of NBA transmissions were offset 3 hr from all the others, which were simultaneous. An upper bound to the effect of this offset can be derived by assuming that the transmitted fluctuations during the 3-hr offset are uncorrelated with the transmitted fluctuations for which they substitute at the other end of the 24-hr period. The maximum expected effect on $\hat{\tau}$ is to multiply it by the factor $[1 - 1/(8k^2)]^{1/2}$, where k is the number of receiver stations, so that to correct for the effect one would divide by this factor. Since k is 5 or 6, the maximum effect on $\hat{\tau}$ is 0.25 percent, or 0.002×10^{-10} for the $\hat{\tau}$ average, a negligible amount. The effect on $\hat{\alpha}_i$ varies somewhat with $\hat{\alpha}_i$ but is at most 0.43 percent for any $\hat{\alpha}_i$ average.

As noted earlier, the method used for estimating α_i and τ requires restriction of the data to those meas-

urements common to all stations compared within a given quarter. To estimate the effect of such restriction, a comparison was made of the column (station) variances for both restricted (C) and total reported data (T) of each station for the entire 18-month period for both the GBR and NBA transmissions. These data are plotted in figure 8, and the square roots or standard deviations are given in tables 2 and 3 for the GBR and NBA data, respectively. As can be seen there is close agreement in all instances except Cruft-GBR; this exception can be attributed to having only three quarters represented in the restricted data. A new estimate of α_i, say $\hat{\alpha}_i'$, was calculated on the basis of the relative increase or decrease of the restricted data in terms of the total reported data for the overall 18-month period. The previously derived transmitter variance estimate, $\hat{\tau}^2$, was employed in these calculations. For the NBA data the new $\hat{\alpha}_i'$ differed from -5 to $+6$ percent from the previous estimate, $\hat{\alpha}_i$. With the exception of the Cruft data, the new $\hat{\alpha}_i'$ for the GBR data differed from -11 to $+3$ percent from $\hat{\alpha}_i$. The $\hat{\alpha}_i$ calculated from the total reported data for the Cruft GBR data was about three times the $\hat{\alpha}_i$ computed from the quarterly restricted data. This new value, however, is much more in line with that computed from the Cruft NBA data. Thus, with the exception of the Cruft GBR data, the restriction to data days common within quarters has no disturbing effect on the estimation of the atomic standard deviations α_i.
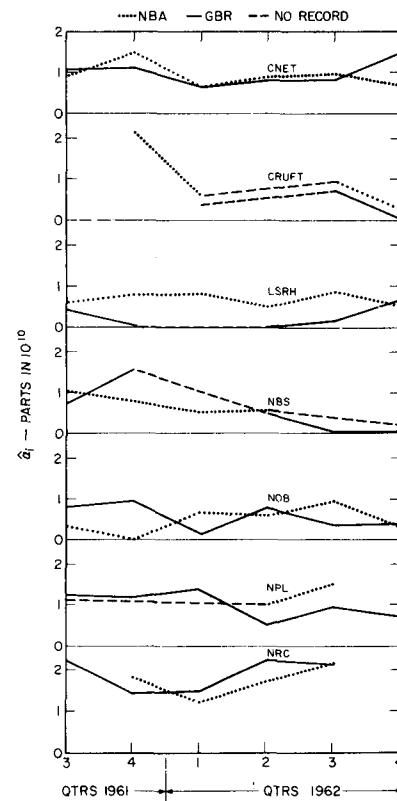

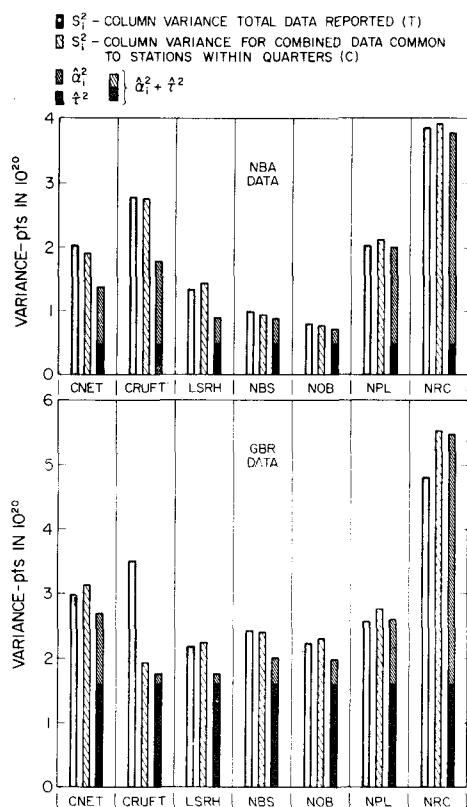FIGURE 7. *Comparison of each laboratory standard $\hat{\alpha}_i$ using either GBR or NBA signals.*

FIGURE 8. *Comparison of variances for GBR and NBA data groups.*

Another factor affecting the estimates of α_i and τ is the elimination of possible variation of mean values from quarter to quarter. Hence, figure 8 includes the $\hat{\alpha}_i^2 + \hat{\tau}^2$ values which were determined from the averaged quarterly values of $\hat{\alpha}_i^2$ and $\hat{\tau}^2$ in figures 5 and 6. This sum would equal s_i^2, the column variance for each station based on data days common to all stations within quarters, in accordance with (9), except for the fact that the column variances plotted in figure 8 were obtained for the overall 18-month period and include variability due to differences among the quarterly means. It is therefore seen that the quarterly means contribute appreciable variation in 7 out of the 14 comparisons; the greatest percentage excess of s_i^2 over $\hat{\alpha}_i^2 + \hat{\tau}^2$, occurring for LSRH-NBA, is 60 percent, which would correspond to 26 percent increase in both $\hat{\alpha}_i$ and $\hat{\tau}$ if their ratio is unchanged, that is, to 0.81 and 0.86 part in 10^{10}. The second greatest percentage increase in $\hat{\alpha}_i$ and τ would be 24 percent, for Cruft-NBA. The other estimated changes in $\hat{\alpha}_i$ and $\hat{\tau}$ due to variability of quarterly means are all increases but range down to a negligible 0.5 percent.

6. Possible Future Improvements in VLF Measurements

One improvement envisioned for VLF phase comparison measurements is the use of the higher-power WWVL (20 kc/s) transmissions (1 to 2 kW radiated power) which commenced from the new NBS Fort Collins Standard Radio Transmitting Site in August 1963 [Richardson, 1964; Blair and Morgan, 1965].

A large increase in coverage is anticipated over the lower-power broadcasts so that many more laboratories may be able to measure their standards directly in terms of the USFS.

The periodic adjustment of quartz crystal oscillators at VLF transmitters may cause undesired phase changes in the transmitted signals. Phase changes in the transmitting antenna may also occur due to the effects of the weather and cause changes [Watt-Carter and Corke, 1961] of at least 1 part in 10^{10}. By means of a VHF radio link from the USWFS (United States Working Frequency Standard) to the VLF transmitter and a continuously phase-locked servosystem [Milton et al., 1962] these sources of error have been eliminated from the WWVL broadcasts.

Another improvement in comparing atomic standards, and especially for separating out propagation errors, now appears possible through the use of portable quartz crystal clocks. Their advantages are their low power consumption, small weight and size, reliability, and excellent stability. A 2.5–Mc/s oscillator, with a stability (standard deviation) of a few parts in 10^{11}, is employed in an experimental portable clock at NBS. The effects of temperature, vibration, and environmental and voltage changes are being carefully evaluated, so that optimum results may be obtained. Preliminary results obtained with portable clocks are very promising, and further developmental work on them is under way at NBS.

7. References

Beehler, R. E., W. R. Atkinson, L. E. Heim, and C. S. Snider (Dec. 1962), A comparison of direct and servo methods for utilizing cesium beam resonators as frequency standards, IRE Trans. Instr. I-11, 231–238.

Blair, B. E., and A. H. Morgan (1965), Control of WWV and WWVH Standard Frequency Broadcasts by VLF and LF Signals, Radio Sci. J. Res NBS **69D**, No. 7, 915–928.

Cramér, H. (1946), Mathematical Methods of Statistics (Princeton University Press, Princeton, N.J.).

Decaux, B. (June 1963), private communication.

Essen, L., J. V. L. Parry, and J. A. Pierce (Sept. 14, 1957), Comparison of caesium resonators by transatlantic radio transmission, Nature **180**, 526.

Essen, L., and J. McA. Steele (Jan. 1962), The international comparison of atomic standards of time and frequency, Proc. IEE **43B**, No. 3752M, 41–47.

Grubbs, F. E. (1948), On estimating precision of measuring instruments and product variability, J. Am. Statist. Assn. **43**, 243–264.

Holloway, J., W. Mainberger, F. H. Reder, G. M. R. Winkler, L. Essen, and J. V. L. Parry (Oct. 1959), Comparison and evaluation of cesium atomic beam frequency standards, Proc. IRE **47**, 1730–1736.

Kalra, S. N. (Mar. 1961), Frequency measurement of standard frequency transmissions, Can. J. Phys. **39**, 477.

Kartaschoff, P. (Dec. 1962), Operation and improvements of a cesium beam standard having 4-meter interaction length, IRE Trans. on Instr. **I-11**, 224–230.

Markowitz, W. (April 1961), Définition, détermination et conservation de la seconde, Comité Consultatif pour la Définition de la Seconde auprès du Comité International des Poids et Mesures, 2e session, Annexe 2 (Gauthier-Villars & Cie, Paris), 45–51.

Markowitz, W. (December 1962), The atomic time scale, IRE Trans. on Instr. **I-11**, 239–242.

Markowitz, W. (July–August 1964), International Frequency and Clock Synchronization, Frequency, **2**, No. 4, 30–31.

Milton, J. B., R. L. Fey, and A. H. Morgan (March 17, 1962), Remote phase control of radio station WWVL, Nature **193**, No. 4820, 1063–1064.

Mitchell, A. M. J. (June 22, 1963), Frequency comparison of atomic standards by radio links, Nature **198**, No. 4886, 1155–1158.

Morgan, A. H., and D. H. Andrews (April 1961), Méthodes et Techniques de contrôle des ondes kilométriques et myriamétriques aux Boulder Laboratories, Comité Consultatif pour la Définition de la Seconde auprès du Comité International des Poids et Mesures, 2e session, Annexe 6 (Gauthier-Villars & Cie, Paris), 68–72.

Morgan, A. H., B. E. Blair, and E. L. Crow (1965), International comparison of atomic frequency standards via VLF radio signals, Progress in Radio Science 1960–1963, **1**, Radio Standards and Measurements, (Elsevier Publishing Company, Amsterdam/London/New York, 1965), 40.

NBS (Dec. 1964), World sets atomic definition of time, NBS Tech. News Bull. **48**, 209–210.

NPL (1963), Report for year 1962, standards division, electrical and frequency measurements, NPL (Teddington, England), 217.

NPL (1964), Report for year 1963, standards division, electrical and frequency measurements, NPL (Teddington, England), 216, 217.

Pierce, J. A. (June 1957), Intercontinental frequency comparisons by very low frequency radio transmission, Proc. IRE **45**, 794–803.

Pierce, J. A., G. M. R. Winkler, and R. L. Corke (September 10, 1960), The 'GBR Experiment': A trans-Atlantic frequency comparison between caesium-controlled oscillators, Nature **187**, No. 4741, 914–916.

Pierce, J. A. (July 1963), private communication.

Richardson, J. M., R. E. Beehler, R. C. Mockler, and R. L. Fey (April 1961), Les étalons atomiques de fréquence au N.B.S., Comité Consultatif pour la Définition de la Seconde auprès du Comité International des Poids et Mesures, 2e session, Annexe 5 (Gauthier-Villars & Cie, Paris), 57–67.

Richardson, J. M. (January 1964), Establishment of new facilities for WWVL and WWVB, Radio Sci. J. Res. NBS **68D**, No. 1, 135.

Watt-Carter, D. E., and R. L. Corke (September 23, 1961), VLF transmitting system phase transfer variations, Nature **191**, No. 4795, 1286.

Note added in proof. A mistake in the calculation of the rms weighted average $\hat{\alpha}_t^2$ for LSRH using GBR transmissions has been discovered; it should be 0.0156 rather than 0.156, resulting in an rms average $\hat{\alpha}_t$ of 0.12 rather than 0.39 (parts in 10^{10}). The change should be made in the abstract, the table on page 911, and the table within figure 5. The apparently considerable change is of doubtful significance because, as stated on page 910, the confidence limits on this α_t are about 100 percent above and below the estimate $\hat{\alpha}_t$. In fact the further data for 1963 and 1964 (502 days) yield 0.43, and when these are combined with the 244 days of data in 1961 and 1962, the overall rms average $\hat{\alpha}_t$ is 0.36.

(Paper 69D7–524)