STATISTICAL ANALYSIS OF TIME TRANSFER DATA FROM TIMATION II

J. McK. Luck and P. Morgan
Division of National Mapping
A.C.T., Australia

ABSTRACT

Between July 1973 and January 1974, three time transfer experiments using the Timation II satellite were conducted by the Division of National Mapping, A.C.T., Australia and the US Naval Research Laboratory, Washington, D.C., to measure time differences between the US Naval Observatory and Australia. Statistical tests showed that the results are unaffected by the satellite's position with respect to the sunrise/sunset line or by its closest approach azimuth at the Australian station. Further tests revealed that forward predictions of time scale differences, based on the measurements, can be made with high confidence.

Measurements Against the Satellite Clock

The results of the first two time transfer experiments between NRL and Australia have already been presented (Easton, Smith and Morgan, 1973, 1974).

The first series of statistical tests examined the residuals from a quadratic fit of the results TII-AUST, where TII represents the satellite on-board oscillator and AUST represents the local Australian time standard, in this case the National Mapping portable cesium standard DNM590 whose performance was linear with respect to a four component mean time scale. The measurements in July showed that the on-board oscillator had a constant aging rate during the experiment, and the residuals were normally distributed, while during the September run the residuals were not normally distributed. The oscillator was evidently much less stable in the January 1974 run, and a simple curve could not be fitted.

The curves fitted were:

July run: $\text{TII-AUST} = 5.371 + 4.7616(t-\bar{t}) + 0.25214(t-\bar{t})^2$ microseconds,

where t was the day of the year and $\bar{t}$ was the midpoint of the run (199.153). The standard deviation of the 38 residuals was 0.585 microseconds.

September run: TII-AUST = $38.940 + 3.9779(t-\bar{t}) - 0.00956(t-\bar{t})^2$ microseconds, where $\bar{t}$ was 265.544. The standard deviation of the 42 residuals was 1.653 microseconds.

For each run, the residuals were analyzed by a two-way analysis of variance with unequal cell sizes (Hamilton, 1964), the classifications being:

(1) Effect of sunlight, to see if it affected either the satellite clock through temperature variation, or the signal travel time.

This classification was sub-divided by the time of closest approach:

- (i) more than 2 hours before sunrise or more than 2 hours after sunset;
- (ii) 1 to 2 hours before sunrise or 1 to 2 hours after sunset;
- (iii) 0 to 1 hours before sunrise or 0 to 1 hours after sunset;
- (iv) 0 to 1 hours after sunrise or 0 to 1 hours before sunset;
- (v) 1 to 2 hours after sunrise or 1 to 2 hours before sunset;
- (vi) more than 2 hours after sunrise or more than 2 hours before sunset.

(2) Effect of azimuth, to see if the local surroundings had any effect. This classification was sub-divided by the azimuth at closest approach:

- (i) in the quadrant North to East;
- (ii) in the quadrant East to South;
- (iii) in the quadrant South to West;
- (iv) in the quadrant West to North.

The number of observations falling into each cell, and the row and column means, are given in Tables I and II. The residuals from the fitted curves are shown in Figure 1.

The analyses of variance presented show that there is no statistical evidence for effects due to the amount of sun illumination on the satellite-station path, or on the quadrant of observation; nor is there any significant interaction between these two factors. No data were available to the authors for testing more precisely the effects of temperature and proton bombardment on the satellite oscillator.

January 1974 Results

A third run was conducted in January 1974, principally to provide an interpolating line for calibrating the Timation system against a USNO flying clock which visited National Mapping on 7 December 1973. On this run, oscillator instability, the restriction to one transmission frequency, and turning the oscillator on and off during the run degraded the results of TII-AUST to the extent that comparisons with the satellite clock could not be statistically analyzed. The time transfer comparisons with USNO were also degraded, as can be seen from Figure 2, but the mean of the observations still proved useful. The results of fitting straight lines to the January measurements of USNO-AUST, together with combinations of all three runs, are given in Table III.

Inter-run Statistics

Three series of tests were conducted to establish whether the runs, and their combinations, were statistically equivalent.

In the following descriptions, the σ^2 are variances of residuals after fitting straight lines of form

$$\text{USNO} - \text{AUST} = \alpha + \beta(t-\bar{t}),$$

and the σ_β^2 are the variances of the observed rates β. The number of observations in each data set is denoted by n.

(1) The first series compared the results obtained, on the one hand, by subtracting direct Australian observations from points interpolated between NRL observations, and on the other hand, by subtracting interpolated Australian observations from direct NRL observations. The tests were:

(i) Equivalence of Sample Populations, i.e. whether NRL-interpolated samples were drawn from the same population as AUST-interpolated samples.

Null hypothesis H_o : $\sigma_1^2 = \sigma_2^2$

Alternative H_1 : $\sigma_1^2 \neq \sigma_2^2$

Test statistic : $f = \sigma_1^2/\sigma_2^2$ (Fisher's F)

Evaluation : Accept H_o if $0.60 < f < 1.67$.

The f-column of Table IV shows that the populations were statistically equivalent at the 95% level for all except the September runs, which were nearly equivalent.

(ii) Non-zero Significance of Rates, i.e. whether the rates of each run were statistically equal to zero.

Null hypothesis $H_o : \beta = 0$

Alternative $H_1 : \beta \neq 0$

Test statistic $: t = [\beta]/(\sigma_\beta^2)^{\frac{1}{2}}$ (Student's t)

$:$ Accept H_o if $t < 1.96$.

The t-columns of Table IV show that the rates were very different from zero in all runs except January. This is due in part to the large standard error and small data set; but reference to Table III shows that the rate was indeed small, which is possibly explainable by the vagaries of the satellite oscillator which made the interpolation scheme unstable. It will be shown in a later test that the January rates were different from the rates determined from the other runs.

(iii) Equivalence of Rates, i.e. whether rates obtained by interpolating NRL observations equalled AUST-interpolated rates.

Null hypothesis $H_o : \beta_1 = \beta_2$

Alternative $H_1 : \beta_1 \neq \beta_2$

Test statistic $: T = (\beta_1 - \beta_2)/S$ (Student's t)

where $S = \{[(n_1-2)\sigma_1^2 + (n_2-2)\sigma_2^2][\sigma_{\beta 1}^2/\sigma_1^2 + \sigma_{\beta 2}^2/\sigma_2^2]/[n_1+n_2-4]\}^{\frac{1}{2}}$

Evaluation $:$ Accept H_o if $T < 1.96$.

The T-column of Table IV shows that the two interpolation schemes gave the same rates.

(2) The second series of tests evaluated whether the midpoint of one run (α at time $\bar{t}$ in Table III) coincided with the value at $\bar{t}$ on the line fitted through another run, i.e. whether runs gave consistent values when extrapolated.

Null hypothesis: $\alpha_1 = \alpha_2 + \beta_2(t-\bar{t})$

Alternative : $\alpha_1 \neq \alpha_2 + \beta_2(t-\bar{t})$

Test statistic : $t = \alpha_1 - [\alpha_2+\beta_2(t-\bar{t})]/S$ with

d-2 degrees of freedom, where $S = [\sigma_1^2/n_1 + \sigma_2^2/n_2]^{\frac{1}{2}}$,

$$d = S^4/[\sigma_1^4/n_1^2(n_1+1) + \sigma_2^4/n_2^2(n_2+1)].$$

This statistic is approximately distributed as Student's t - it is the incomplete Fisher-Behrens statistic (Welch, 1937; Hamilton 1964).

Evaluation : Accept H_o if $t < 1.96$.

Table V shows that forward extrapolation of the NRL-interpolated samples is valid - even extrapolating from the July run into January is satisfactory at the 2% level. The fact that the AUST-interpolated samples do not give such good extrapolation characteristics is attributed to the sparser Australian data sets - otherwise it is a little puzzling.

(3) The third series tested the hypotheses that, for the NRL-interpolated samples, the sample populations and rates from each run and combination were equivalent; and similarly for the AUST-interpolated samples. All single runs satisfied a χ^2 goodness-of-fit test for normality.

(i) Equivalence of population variances.

Null hypothesis $H_o : \sigma_1^2 = \sigma_2^2$

Alternative $H_1 : \sigma_1^2 \neq \sigma_2^2$

Test statistic : $f = \sigma_1^2/\sigma_2^2$ (Fisher's F)

Evaluation : Accept H_o if $0.60 < f < 1.67$.

The f-column of Table VI shows clearly that the July run had significantly lower variance than any other run or combination, but that the other runs were, by and large, from the same population.

(ii) Equality of rates.

Null hypothesis $H_o : \beta_1 = \beta_2$

Alternative $H_1 : \beta_1 = \beta_2$

Test Statistic : $T = [\beta_1 - \beta_2]/S$ (Student's t)

where $S = \{[(n_1-2)\sigma_1^2 + (n_2-2)\sigma_2^2][\sigma_{\beta_1}^2/\sigma_1^2 + \sigma_{\beta_2}^2/\sigma_2^2]/[n_1+n_2-4]\}^{\frac{1}{2}}$

Evaluation : Accept H_o if $t < 1.96$.

The T-column of Table VI shows that, for the NRL-interpolated samples, the rate determined from the January run was statistically different from the rates determined from all other runs and combinations, which were in turn statistically equal to each other. This confirms the result found in test (ii) of the first series. The poorer results obtained from the AUST-interpolated samples confirm the second series tests wherein extrapolation between some run combinations was not valid.

Comparison of Time Scales

The tests described above all used a single cesium standard, DNM590, for the time scale denoted AUST. To demonstrate that the out-of-character results of the January run were not due to a change of rate in this clock, a special Australian artificial time scale (AATS) was constructed, comprising the four cesium standards DNM590, the original Mount Stromlo standard DNM205, the newer standard NSL338 of the National Standards Laboratories, CSIRO, Sydney, and standard HP052 maintained by Hewlett Packard (Australia) Limited, Melbourne. These clocks are all compared daily by ABC television comparisons (Miller 1970) and were not stopped or adjusted in the period between 8 February 1973 and 23 May 1974. No other cesium standard in Australia satisfied both these conditions. The time scale was a simple unweighted mean of the four clocks, offset (in phase only) so that it agreed approximately with UTC (USNO) determined by flying clocks.

An extrapolating ephemeris for AATS was constructed, using least squares straight line fits, in which:

$$E[\text{USNO-AATS}] = \sum_{i=1}^{4} E[\text{USNO-Clock}_i]/4$$

and

$E[\text{USNO-Clock}_i] = E[\text{USNO-DNM590}]$ (by Timation) $+E[\text{DNM590-Clock}_i]$ (by Television).

Selected points on the graphs of the clocks against the artificial time scale are shown in Figure 3. It can be seen that no significant rate change occurred in DNM590. Assuming that no such rate change occurred in UTC (USNO), the poor results in rate from the January run would reflect deficiencies in the Timation II technique.

Table III includes the result of a visit by USNO flying clock PC572 in December 1973, and, in the column headed USNO-AUST, gives the values obtained by inserting t = 7 December 1973 in the various formulae obtained for different Timation II runs and combinations. For each interpolating system, the July-September combination gave the best agreement, which was 0.31 microseconds for NRL-interpolations and 0.17 microseconds for UAST-interpolations. The 95% confidence interval at 7 December for the NRL-interpolated July-September combination was ± 0.27 microseconds, and for the AUST-interpolated combination ± 0.20 microseconds. Thus, on the assumption that no error was attached to the flying clock result, the former set gave almost statistically correct results, while the latter set showed excellent agreement. When the quoted error of ± 0.2 microseconds from the USNO certification was taken into account, the NRL-interpolated result also became acceptable.

Extrapolating the July-September combinations backwards to the date of the previous USNO flying clock visit on 8 February 1973, the differences were 3.03 microseconds with ± 0.05 microseconds 95% confidence interval for USNO-interpolates, and 3.12 microseconds with 0.37 microseconds 95% confidence interval for AUST-interpolates. When the errors of the flying clock measurements are taken into account, these results are not unsatisfactory. On the other hand, the agreement between USNO-AATS by the extrapolating ephemeris and USNO-AATS by the flying clock measurement was 0.8 microseconds, which is regarded as very satisfactory over such an interval. It is unfortunate that no subsequent definitive flying clock trip has been made, as a further test of the predictive power of Timation would have been very beneficial, especially as time keeping in Australia during 1974 has been plagued with breakdowns both in a number of cesium standards and in the television network system.

On the basis of the consistency of the NRL-interpolated July-September combination in both the extrapolative and interpolative senses, its agreement with USNO flying clock measurements, and its consistency with a selected Australian time scale, the formula 4 in Table III was adopted as the definitive comparison between Australian clocks and UTC(USNO MC), and was accordingly made the sole interpolating link between UTC(USNO MC) and the regular television-compared Australian mean time scale UTC(Aus).

Conclusions

The foregoing statistical analysis shows quite clearly the value of the Timation satellite for intercontinental time transfer at the sub-microsecond level. The major areas requiring particular attention are:

(i) Long runs are required to establish rates reliably. The durations of the three runs were thirteen, seventeen and twelve days in July, September and January respectively, and even then the sparser results in January produced anomalies.

(ii) The stability of the satellite oscillator must be good enough to carry interpolations over several hours. The poor January results establish this point forcibly.

(iii) The superior results from the July run show that dual frequency transmissions (150 MHz and 400 MHz) do indeed reduce errors.

(iv) Every effort should be made to have the data set as dense as possible.

(v) There are factors affecting the stability of the oscillator which we do not yet understand, since an analysis of variance failed to reveal two possible causative, or perhaps correlated, effects. It is significant here that the residuals AUST-TII were not normally distributed yet the residuals from USNO-AUST = (USNO-TII)-(AUST-TII) were normally distributed, thus indicating the presence of a perturbing influence in the region of the satellite. No data was available to test accurately the hypothesis that the temperature around the crystal caused it to fluctuate.

The Timation II results presented here have been incorporated into UTC(AUS), so that predictions of the relationship between Australian clocks and USNO can be made with confidence at the microsecond level. It is hoped that regular observations of Timation III can be carried out - its improved clock should improve the statistics considerably and enable our geographically isolated clocks to contribute to International Atomic Time.

Acknowledgements

The authors particulary thank Mr. R. Easton of the Naval Research Laboratory, Washington, D.C. for his enthusiastic collaboration in the project. They also thank Mr. R.J. Bryant and Mr. R. D. Craven of the Division of National Mapping for their assistance with the calculations for this paper, and Miss M. Dowhy, Miss M. Schussig and Miss J. Thurling of the Division for preparing the manuscript.

References

Easton, R.L., Smith, H.M., Morgan, P.: "Submicrosecond Time Transfer Between the United States, United Kingdom and Australia Via Satellite," Proc. 5th Precise Time and Time Interval Planning Meeting, Goddard Space Flight Center, Greenbelt, Md., December 1973. NASA publication X-814-74-225.

Easton, R.L., Morgan, P., Smith, H.M.: "Satellite Time Transfer from the U.S. Naval Observatory to the Royal Greenwich Observatory and to Australia," paper presented to conference on Precision Electromagnetic Measurements, London, U.K., July 1974.

Hamilton, W.C.: "Statistics in Physical Sciences," The Ronald Press Co., N.Y., 1964.

Hoel, P.G.: "Introduction to Mathematical Statistics," 3rd ed., John Wiley and Sons Inc., N.Y., 1966.

Miller, M.J.: "Synchronization of Distant Clocks by Television Pulse Comparisons," Proc. Astronomical of Australia, Vol 1 No 7 p 352, April 1970.

Welch, B.L.: "The Significance of the Difference Between Two Means when the Population Variances are Unequal," Biometrika, Vol 29 pp 350-362, 1937.

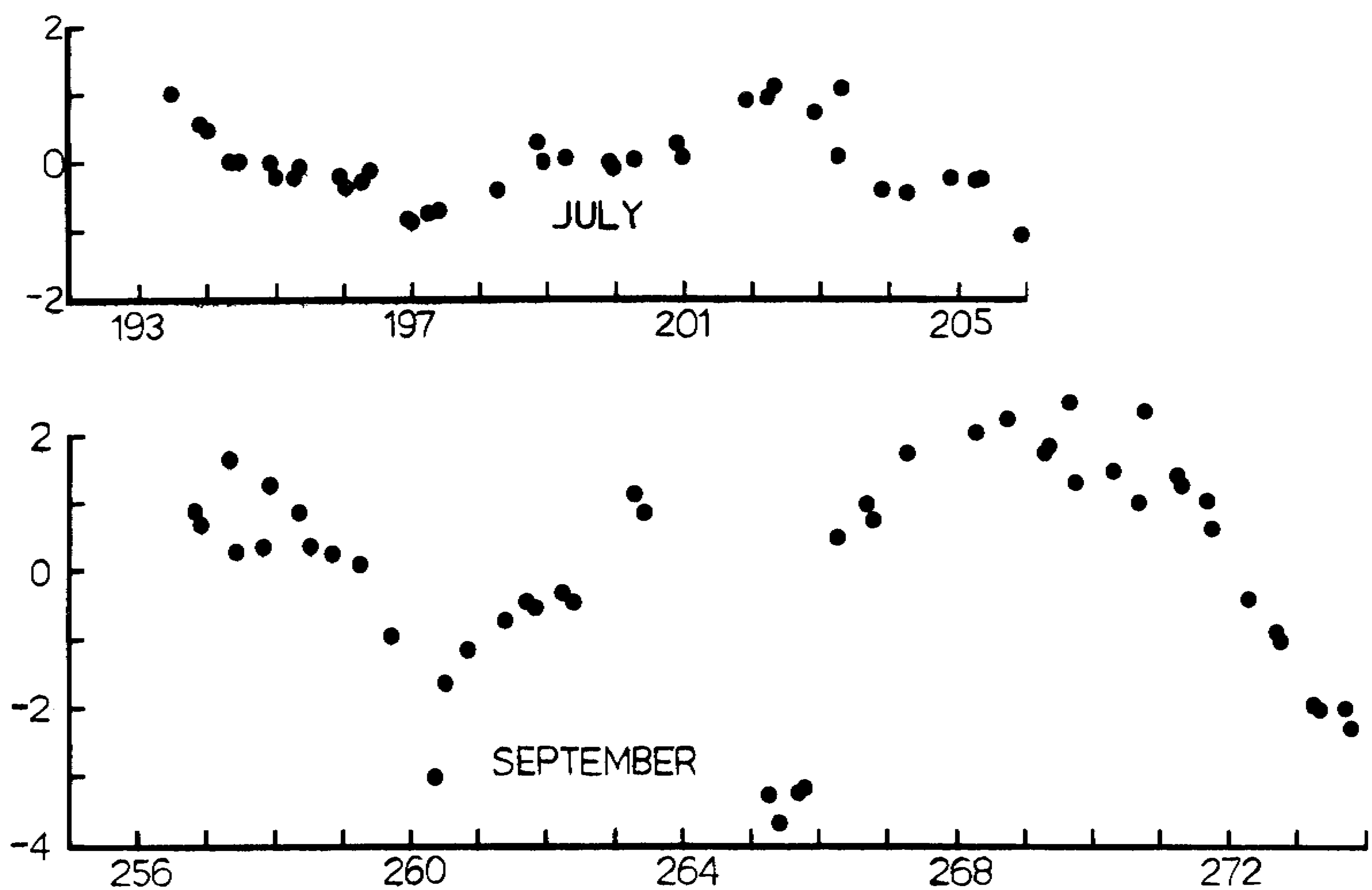

Figure 1: Residuals from Quadratic Fits of TII-AUST (NRL-Interpolated), July and September 1973. The Horizontal Scale is Marked in UT Days. The Vertical Scale is Marked in Microseconds.

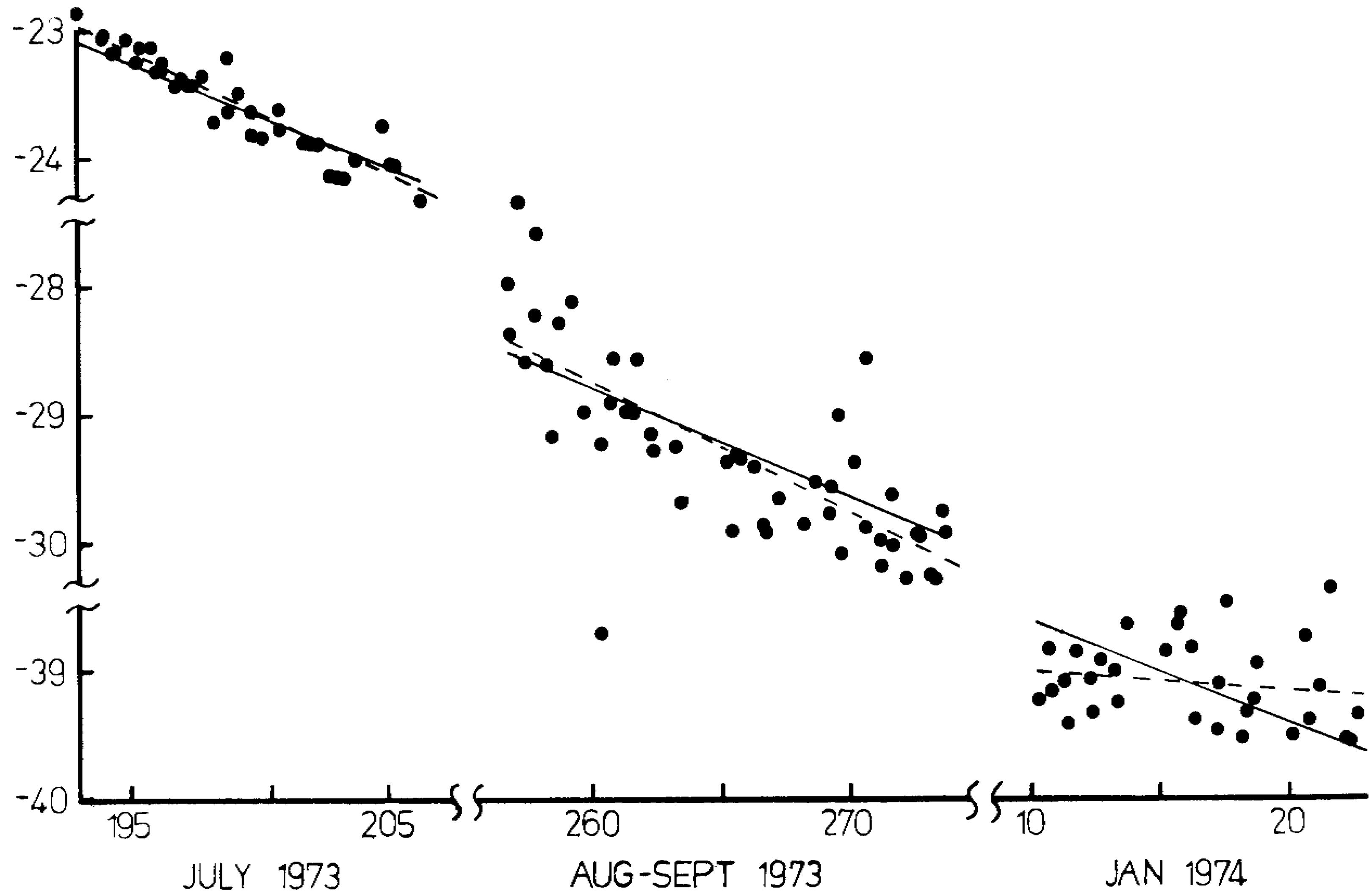

Figure 2: USNO–Aust (NRL–Interpolated).
The Full Line is Line 6 of Table III, i.e. Combination of All Three Runs.
The Dashed Lines are Lines 1,2,3, of Table III, i.e. Individual Runs.
The Horizontal Scale is Marked in UT Days. The Vertical Scale is Marked in Microseconds.

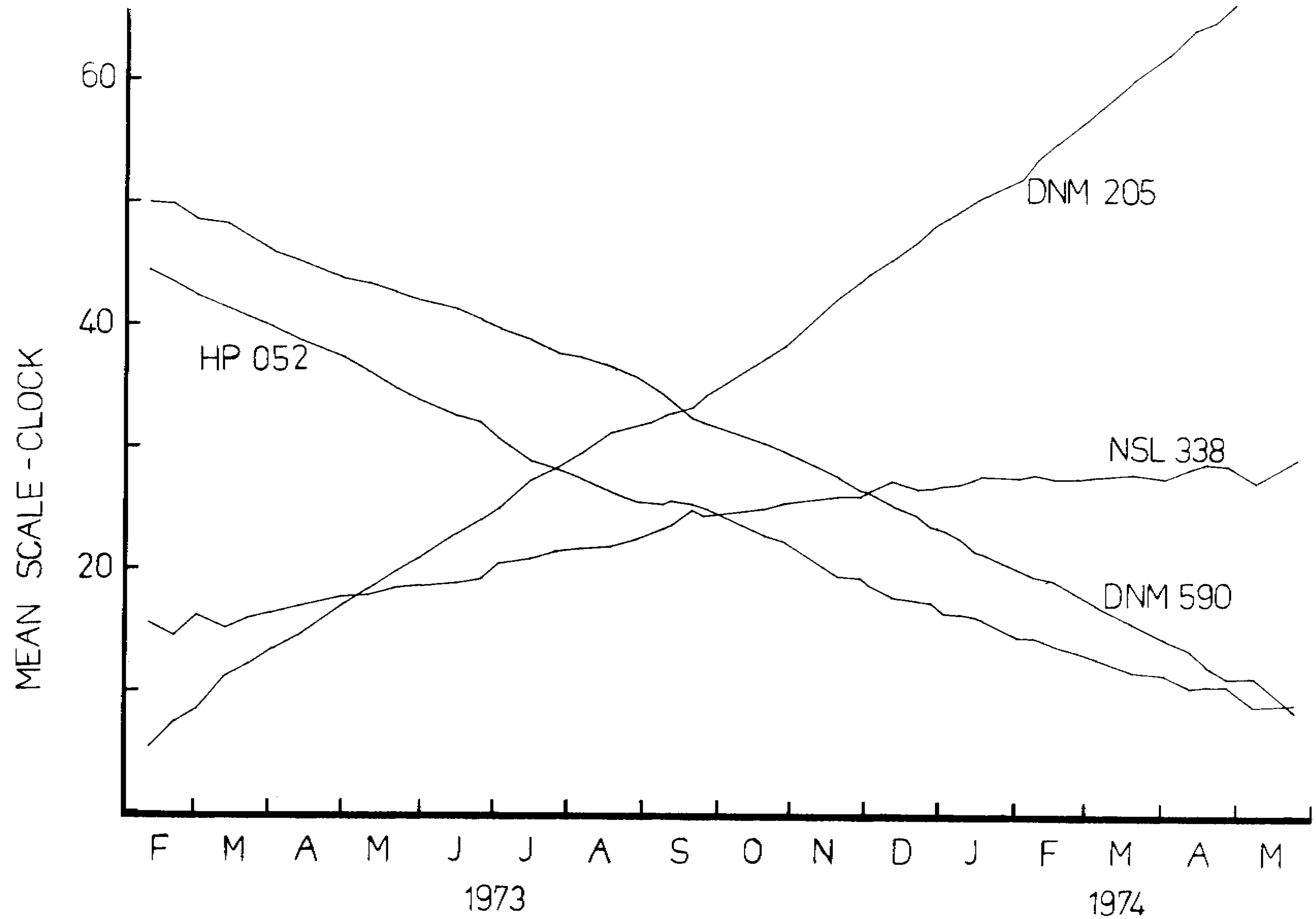

Figure 3: Australian Artificial Mean Time Scale (AATS)-Clock for Each Component Clock. The Vertical Scale is Marked in Microseconds; its Origin is Arbitrary.

TABLE I

Analysis of Variance, July 1973 Run

Number of Observations and Cell Sums

Illumination at t.c.a.	Quadrant of c.a.				Row Sums
	NE	SE	SW	NW	
1. $>2^h$ before sunrise or $>2^h$ after sunset (dark)	0 -	0 -	0 -	5 0.215	0.215
2. 1^h to 2^h before sunrise or after sunset	1 0.176	1 -0.846	1 0.027	2 2.279	1.636
3. 0^h to 1^h before sunrise or after sunset (twilight)	1 0.117	4 -1.720	2 -0.682	0 -	-2.285
4. 0^h to 1^h after sunrise or before sunset	3 0.466	2 1.114	1 0.734	0 -	2.314
5. 1^h to 2^h after sunrise or before sunset	3 0.349	1 -0.283	4 -1.887	0 -	-1.821
6. $>2^h$ after sunrise or before sunset (daylight)	0 -	0 -	7 -0.057	0 -	-0.057
Column Sums	1.108	-1.735	-1.865	2.494	0.002
Total Sum of Squares					11.974034

Residual Mean Square:	0.3851 with 14 degrees of freedom
Illumination Mean Square:	0.5196 with 5 degrees of freedom
F-statistic (calculated):	1.349
Critical Value for rejection:	2.96 for $F_{5,14}$ (5%)
Quadrant Mean Square:	0.5500
F-statistic (calculated):	1.428
Critical Value for rejection:	3.34 for $F_{3,14}$ (5%)

TABLE II

Analysis of Variance, Aug/Sept Run

Number of Observations and Cell Sums

Illumination at t.c.a.	Quadrant of c.a. NE	SE	SW	NW	Row Sums
1. $>2^h$ before sunrise or $>2^h$ after sunset (dark)	1 -2.095	6 1.385	4 -4.709	0 -	-5.419
2. 1^h to 2^h before sunrise or after sunset	3 2.580	3 -4.481	5 3.737	4 3.297	5.133
3. 0^h to 1^h before sunrise or after sunset (twilight)	3 -2.459	2 1.313	0 -	3 -4.772	-5.918
4. 0^h to 1^h after sunrise or before sunset	5 7.283	0 -	3 -1.217	3 -1.241	4.825
5. $>1^h$ after sunrise or before sunset (daylight)	2 -0.551	0 -	0 -	2 1.929	1.378
Column Sums	4.758	-1.783	-2.189	-0.787	-0.001
Total Sum of Squares					132.343029

Residual Mean Square: 2.881 with 29 degrees of freedom
Illumination Mean Square: 2.8490 with 4 degrees of freedom
F-statistic (calculated): 0.989
Critical Value for rejection: 2.70 for $F_{4,29}$ (5%)
Quadrant Mean Square: 0.786 with 3 degrees of freedom
F-statistic (calculated): 0.273
Critical Value for rejection: 2.93 for $F_{3,29}$ (5%)

TABLE III

USNO-AUST for each run, and combinations of runs.

The time scale designated AUST is the portable caesium standard DNM590. Units are in microseconds or microseconds/day. t is measured in UT days from 1970 Jan 0; $\bar{t}$ is the mean date of observation in each run. n is the number of observations in each run.

Run	USNO-AUST=$\alpha + \beta(t-\bar{t})$			USNO-AUST 1973 Dec 07	n	standard deviation	
	α	β	$\bar{t}$			σ_α	σ_β
A. Australian observations to NRL interpolates							
1. July 1973	-23.566	-0.09864	199.0	-37.57	38	.0221	.00600
2. Sept 1973	-29.320	-0.10195	265.4	-37.03	50	.0710	.01296
3. Jan 1974	-39.102	-0.01470	381.4	-38.51	32	.0590	.01528
4. July - Sept 1973	-26.835	-0.08686	236.7	-35.89	88	.0418	.00126
5. Sept 1973 - Jan 1974	-33.137	-0.08437	310.7	-35.69	82	.0528	.00093
6. July - Sept - Jan	-30.106	-0.08511	275.3	-35.70	120	.0369	.00053
B. NRL observations to Australian interpolates							
7. July 1973	-23.649	-0.10202	199.6	-38.07	43	.0223	.00620
8. Sept 1973	-29.403	-0.10021	265.5	-36.97	62	.0489	.01009
9. Jan 1974	-39.004	-0.02308	381.8	-38.06	47	.0500	.01363
10. July - Sept 1973	-27.046	-0.08762	238.5	-36.03	105	.0305	.00094
11. Sept 1973 - Jan 1974	-33.543	-0.08249	315.7	-35.63	109	.0380	.00066
12. July - Sept - Jan	-30.744	-0.08399	282.8	-35.63	152	.0292	.00041

UTC(USNO MC) - DNM590 (by USNO PC572): $-36.2 \pm 0.2\,\mu s$, 73 Dec 07.

TABLE IV

Comparison Between NRL-Interpolated and AUST-Interpolated Runs

Run	Hypothesis			
	1(i) Equal Variances	1(ii) Zero Rate		1(iii) Equal Rates
	f	t (NRL-Int)	t (AUST-Int)	T
July	0.868	16.440*	16.455*	0.390
Sept	1.700*	7.867*	9.932*	0.107
Jan	0.948	0.962	1.693	0.407
July-Sept	1.568	68.937*	93.213*	0.531
Sept-Jan	1.450	90.720*	124.985*	1.695
July-Sept-Jan	1.259	160.585*	204.854*	1.697

* Hypothesis rejected.

TABLE V

Validity of Extrapolating Between Runs

Run: Extrapolated from	Run: to Midpoint of	Diff (Extrap. - Mean)	Number of Points (n_2)	s.d. of Extrap. Value (σ_2)	t
NRL - Interpolated					
July	Sept	-0.800	38	0.400	1.969
July	Jan	-2.454	38	1.095	2.238 §
Sept	Jan	-2.035	50	1.504	1.352
July-Sept	Jan	-0.294	88	0.187	1.499
Sept-Jan	July	-0.146	82	0.117	1.137
AUST-Interpolated					
July	Sept	-0.964	43	0.409	2.340 *
July	Jan	-3.233	43	1.130	2.858 *
Sept	Jan	-2.059	62	1.175	1.751
July-Sept	Jan	-0.600	105	0.138	4.087 *
Sept-Jan	July	0.324	109	0.085	3.687 *

§ Hypothesis accepted at 2% level.

* Hypothesis rejected.

TABLE VI

Evaluation of Inter-Run Consistency

Run 1	Run 2	Hypothesis	
		3(i) Equal Variances f	3(ii) Equal Rates T
NRL - Interpolated			
July	Sept	0.073 *	0.166
July	Jan	0.166 *	5.356 *
July	Sept-Jan	0.081 *	0.800
Sept	Jan	2.259 *	3.733 *
Jan	July-Sept	0.726	4.164 *
July-Sept	July-Sept-Jan	0.941	1.263
July-Sept	Sept-Jan	0.673	1.521
Sept-Jan	July-Sept-Jan	1.400	0.724
AUST - Interpolated			
July	Sept	0.144 *	0.117
July	Jan	0.181 *	5.068 *
July	Sept-Jan	0.135 *	1.327
Sept	Jan	1.260	4.408 *
Jan	July-Sept	1.201	5.023 *
July-Sept	July-Sept-Jan	0.756	2.161 *
July-Sept	Sept-Jan	0.622	4.171 *
Sept-Jan	July-Sept-Jan	1.216	1.980 §

§ Hypothesis accepted at 2% level.

* Hypothesis rejected.

QUESTION AND ANSWER PERIOD

DR. WINKLER:

I think there are several impressions which I have gotten, particularly yesterday and today when systems applications were discussed and performances have been disputed.

One fact came out very strongly this morning was that apparently people talking about the same subject can claim directly opposite extremes and yet both may be right. One radio astronomer said that the time limitation of experimentation is the atmosphere and not the clocks. The second said it's the clock, not the atmosphere and I believe that both are right and points out the fact that it is not sufficient to specify simply the performance of an atomic clock, for instance. That there are so many parts, ten to 12th or 10 to the 13th is useless unless one also specifies environment or specifies the timing that is required and a lot of other additional things.