JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 91, NO. B9, PAGES 9045-9055, AUGUST 10, 1986

Further Analysis of the 1981 Southern California Field Test for Leveling Refraction

M. R. CRAYMER[1] AND P. VANÍČEK

Survey Science, University of Toronto, Mississauga, Ontario, Canada

Application of least squares spectral analysis and multiple linear regression techniques to the 1981 southern California field test for leveling refraction has revealed that in addition to differential refraction, rod/instrument settlement and an effect attributed to differential rod miscalibration are also detectable. The object of our analysis was the discrepancy between the forward and backward runnings of a section as this quantity properly reflects the direction of running thereby allowing for the detection of direction dependent effects. A multiple regression model using three arguments representing differential refraction, rod/instrument settlement, and differential rod miscalibration reduced the observed variation of the discrepancies by 61% as opposed to 53% when only the National Geodetic Survey (NGS) computed refraction correction is applied. It was found that of the original 23-mm accumulated discrepancy, 14 mm was attributed to differential refraction, 20 mm to settlement, and -14 mm to differential rod miscalibration. Analyses with the NGS computed refraction corrections applied (based on Kukkamäki's single sight equation with observed temperatures) gave similar results. It is also shown that the settlement effect is always present in any discrepancy and accumulates in the discrepancies between the forward and backward runnings while it cancels and is thus hidden in the accumulation of the NGS-derived discrepancies between the short and long sight length runnings.

INTRODUCTION

In May and June 1981 a joint U.S. Geological Survey and National Geodetic Survey field leveling experiment was carried out along a 50-km line from Saugus to Palmdale, California [cf. *Adams*, 1981]. The purposes of the experiment were [*Whalen and Strange*, 1983] (1) to measure the magnitude of the differences between heights determined using two different sight lengths along the same leveling line, (2) to determine if standard refraction models, in conjunction with measured vertical temperature gradients, would explain possible differences in observed heights, and (3) to determine how well the temperature model developed by *Holdahl* [1981] reproduces observed temperature differences.

For this experiment a single line of double run leveling over uniformly sloping terrain was observed. One running employed short sights (SSL) of an average length of 24.3 m and the other long sights (LSL) of an average length of 42.6 m. It was expected that the uniform slope and significant sight length difference would amplify the differential refraction effect on the discrepancies between the SSL and LSL runnings. It was also intended to frequently alter the direction of running of both the short and long sight levelings in order to minimize the rod settlement effect. Unfortunately, this procedure was not strictly followed as 12 of the 60 section runnings (20%) were not properly "balanced." In fact, all of the imbalance occurs over the last 42 runnings (70% of the line) corresponding to a 29% imbalance over this part.

A number of analyses have been performed on the collected data [e.g., *Stein et al.*, 1982, this issue; *Stein and Thatcher*, 1982; *Holdahl*, 1982; *Whalen and Strange*, 1983; *Castle et al.*, 1983]. While all of these studies agree that differential refraction effects contribute to the observed discrepancies between the SSL and LSL runnings, only *Castle et al.* [1983] consider an additional source of error (rod settlement). As will be shown below, the combined rod/instrument settlement effect is likely to systematically affect the F-B discrepancies of any leveling line.

This paper is also concerned with the investigation of refraction effects. In addition, it also investigates other sources of error. Our analysis attempts first to diagnose possible sources of systematic effects and then to model (simultaneously) and remove these effects from the discrepancies. The object of our attention, however, is not the SSL-LSL discrepancy used in the previous investigations. *Whalen and Strange* [1983] and *Stein et al.* [this issue] define this discrepancy as

$$d = |dH_{SSL}| - |dH_{LSL}| \tag{1}$$

where dH_{SSL} and dH_{LSL} are the observed section height differences for the SSL and LSL runnings. This quantity was chosen in order to maximize the refraction effect. However, it also effectively destroys the functional relation of d with other effects that depend on the sense of running. We use instead the discrepancy between the forward (F) and backward (B) runnings, defined by

$$d = dH_F + dH_B \tag{2}$$

where dH_F and dH_B are the observed height differences for the forward and backward runnings of a section, respectively. This discrepancy properly reflects the sense of running through the signs of the height differences (dH_F and dH_B are taken to have opposite signs). It therefore allows for the investigation of other possible "contaminations" of the discrepancies in addition to refraction (e.g., direction dependent effects). Furthermore, this discrepancy is generally the only one that is available in practical first-order leveling were the number of setups in both runnings are balanced. Throughout the remainder of the paper, the discrepancy d will be taken as the quantity defined by (2).

[1]Now at Geodetic Research Laboratory, University of New Brunswick, P.O. Box 4400, Fredericton, New Brunswick, E3B 5A3.

Paper number 5B5595.
0148-0227/86/005B-5596$05.00

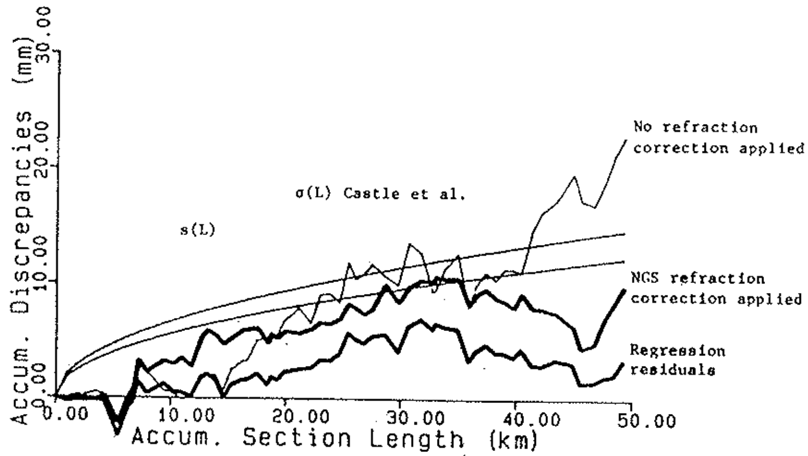

Fig. 1. Comparison of accumulated discrepancies (with and without NGS refraction correction) with propagation of uncorrelated random discrepancies based on the work by *Castle et al.* [1983] and sample estimates of standard deviation.

DATA

The data used in this analysis were kindly provided by the U.S. National Geodetic Survey (NGS) (R. J. Lee, personal communication, 1984) in their standard REDUC4 format and includes the following information for each section: starting and ending bench mark numbers, date and time at start and end of running, air temperature at start and end of running, number of setups in the run, length of section, observed (uncorrected) height difference, level collimation correction, rod scale calibration correction (from detailed laser calibrations of each graduation), rod scale temperature correction, astronomic correction, and NGS refraction correction (based on Kukkamäki's single sight correction with observed temperatures). Descriptions of the last five items (corrections) may be found in the work by *Balazs and Young* [1982]. *Whalen and Strange* [1983] also describe the refraction correction in more detail.

EVIDENCE OF SYSTEMATIC ERRORS

The most striking display of the presence of systematic effects in the discrepancies is the plot of their accumulated values against the accumulated section length (see Figure 1). The discrepancies steadily accumulate to 23 mm over the 50 km length.

In order to assess the significance of this accumulation, a comparison was made with the expected accumulated random effects. *Castle et al.* [1983] have estimated the expected standard deviation of the random effect on the discrepancy to be $\sigma(L) = 2.1$ mm $\sqrt{(L \text{ km})}$, where L is the cumulative line length. This assumes that the short sight running was of first-order (class II) and the long sight running of second-order (class II). Using this value, the standard deviation of the total random discrepancy accumulated over 50 km would be 14.8 mm. Clearly, the total accumulated discrepancy (with all corrections except for refraction applied) falls outside the expected range for uncorrelated random error propagation (to be satisfied 68% of the time). On the other hand, 73% of the partial sums of discrepancies $(\Sigma d)_i$ along the line fall within the bounds for random, uncorrelated error propagation, thereby satisfying the expected 68% probability level.

An examination of the discrepancies reveals that the sample estimate of their standard deviation $s(L)$ is 1.75 mm$\sqrt{(L \text{ km})}$; less than the value implied by *Castle et al.* [1983] even though these discrepancies may be burdened with systematic errors. Using this estimate, we find that the total expected accumulated discrepancy is 12.3 mm and the frequency of $(\Sigma d)_i$ falling within the expected limits for $s(L_i)$ is only 58%. This is significantly smaller than the 68% expected for uncorrelated random error propagation and clearly indicates the presence of systematic effects.

Application of the NGS refraction correction does not reduce the accumulation noticably over most of the line but does reduce the excessive values near the end of the line. The total accumulation is reduced by about half (to 10 mm) and is now smaller than the value predicted using either of the above standard deviations. The variation of the discrepancies is also reduced by about 53%. From these results, one may be led therefore to conclude that the original discrepancies had been contaminated only by refraction and that the NGS correction effectively eliminates the effect. However, $s(L)$ is now 1.22 mm$\sqrt{(L \text{ km})}$ and the frequency of $(\Sigma d)_i$ falling within this random error limit is only 35%; much less than for the discrepancies uncorrected for refraction. This suggests that there may be other effects present. The remainder of this paper gives more evidence to support this.

MODELS FOR KNOWN SYSTEMATIC EFFECTS

The basis of the method employed here involves the determination of linear correlations between the discrepancies and quantities (arguments) characterizing the various systematic effects. For some known systematic effects we develop the mathematical models that relate the measured arguments to the discrepancies.

Differential Refraction

The refraction model used here is based on balanced sight equation developed by *Kukkamäki* [1939]:

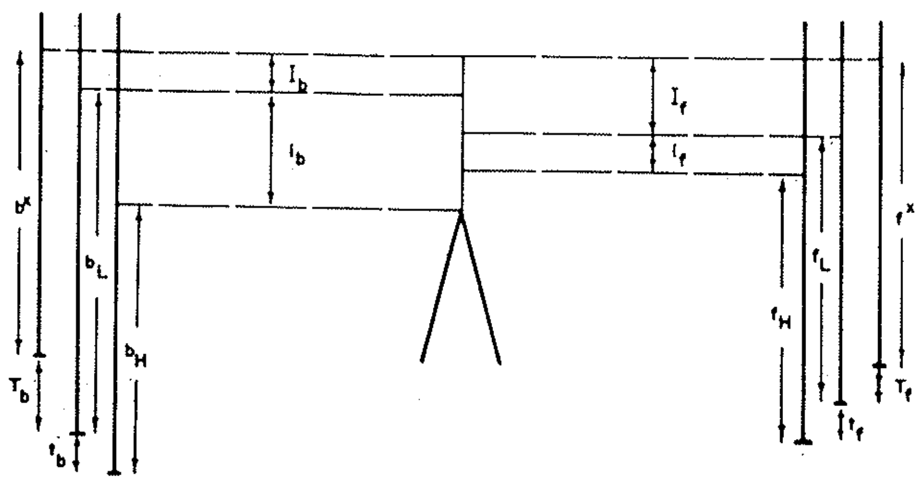

Fig. 2. Rod and instrument settlement.

$$e_R = A\ dt\ dh\ s^2 \tag{3}$$

where e_R is the error due to refraction in the observed setup height difference dh, s (not to be confused with the standard deviation in the previous section) is the sight length (assumed equal for foresight and backsight), dt is the temperature difference at heights 0.5 and 2.5 m above ground, and A is a coefficient depending on the form of the temperature gradient and other meteorological parameters.

A temperature function of the form

$$t = a + b\,z^c \tag{4}$$

is used, where t is the temperature at a height z above ground and a, b, and c are constants. For positive heat flux, c equals to about -1/3, and if average meteorological conditions in England are applicable, one obtains $A = 6.45 \times 10^{-8}$ mm/(°C m^3) [*Kukkamäki*, 1939].

Assuming A to be constant for a section of length S and n setups, the refraction error in an entire section is given by [*Craymer*, 1984]

$$R = \Sigma(A\ dt\ dh\ s^2) = A\ \overline{dt}\ dH\ \bar{s}^2 \tag{5}$$

where $\overline{dt}$ is the average temperature difference for the section and $\bar{s}$ is the average sight length for the section; that is,

$$\bar{s} = S\ /\ 2n \tag{6}$$

The resulting refraction contribution (d_R) to the discrepancy d is, from (2),

$$d_R = R_F + R_B \tag{7}$$

where R_F and R_B are the refraction effects on dH_F and dH_B, respectively. Expanding this equation, one obtains

$$dR = A\ [(\overline{dt}\ dH\ \bar{s}^2)_F + (\overline{dt}\ dH\ \bar{s}^2)_B] \tag{8}$$

Letting $a_R = [(\overline{dt}\ dH\ \bar{s}^{\,2})_F) + (\overline{dt}\ dH\ \bar{s}^2)_B]$, the linear model for the refraction effect on the discrepancies may be written as

$$d_R = c_R\ a_R \tag{9}$$

where the argument a_R is computed from the given auxiliary data and the coefficient c_R=A is to be estimated.

Rod and Instrument Settlement

The combined effects of rod and instrument settlement are illustrated in Figure 2 for the back-rod-low-scale, fore-rod-low-scale, fore-rod-hight-scale, back-rod-high-scale (bffb) observing procedure. Here, the following notation is used:

- f^*, b^* fore and back rod readings unaffected by settlement;
- f_L, b_L actual fore and back rod low scale readings;
- f_H, b_H actual fore and back rod high scale readings;
- T_f, T_b amounts of settlement of the fore and back rods (i.e., turning points) during the time elapsed between the last observation (f_H) of the previous setup and the first observation (b_L) of the current setup;
- t_f, t_b amounts of settlement of the fore and back rods during the time elapsed between the b_L and b_H rod readings;
- I_f, I_b amounts of instrument settlement during the time elapsed between the setting of the instrument and the back and fore low scale rod readings; respectively,
- i_f, i_b amounts of instrument settlement during the time elapsed between high and low scale rod readings on the fore and back rods, respectively.

The settlement effect on the height difference (average of high and low scale) at a single setup is then (see Figure 2)

$$e_{TP} = (T_b\text{-}T_f) + (t_b\text{-}t_f)/2 - (I_b\text{-}I_f) + (i_b\text{-}i_f)/2 \tag{10}$$

Note that to avoid conflicting notation later in the paper, the subscript TP (refering to turning point) has been used even though it is actually describing the entire settlement effect.

The resulting effect on the section discrepancy (d_{TP}) is obtained by assuming the settlement e_{TP} (i.e., ground conditions and leveling procedure) to be constant for all setups. Note that for the first setup of a section running no back rod settlement is assumed because the starting bench mark is used for the "turning point." Here, the settlement of the forward rod and instrument are ignored since these effects are much less than the the settlement of the back rod (i.e., turning point). There being thus n_F-1 turning points in the forward running and n_B-1 turning points in the backward running, the error can be simplified to

$$d_{TP} = e_{TP}\,(n_F+n_B-2) = e_{TP}\,n_{TP} \tag{11}$$

where n_{TP} is the total number of turning points in both runnings. The linear model for the analysis may then be written as

$$d_{TP} = c_{TP}\,a_{TP} \tag{12}$$

where $a_{TP}=n_{TP}$ is the computed argument and $c_{TP}=e_{TP}$ is the coefficient to be estimated.

It should be realized that unlike the refraction effect (where dH_F and dH_B are of opposite signs), rod settlement does not even partially cancel in the discrepancy. Thus any rod settlement present in either the forward or backward runnings will definitely accumulate in the discrepancy. On the other hand, the effects on the average section height difference cancel each other if the numbers of forward and backward setups are equal and all field operations (i.e., timing, instrument operation, etc.) are the same for both runnings.

For comparative purposes, estimates of the expected magnitude of $c_{TP}=e_{TP}$ may also be obtained from the empirical expressions for rod and instrument settlement derived by *Anderson* [1983] from experiments. For hard packed gravel the function relating rod settlement h (positive for sinking) to the time t (seconds) elapsed since the setting of the rod is

$$h = 0.02\exp[-0.009\,t] - 0.03 \quad \text{(mm)} \tag{13}$$

and for asphalt,

$$h = 0.06\exp[-0.008\,t] - 0.05 \quad \text{(mm)} \tag{14}$$

Using in these equations, typical time intervals for the bffb observing procedure (M. R. Elliott, personal communication, 1983), the expected settlement effect is, from (10),

$$e_{TP} = 0.02 \quad \text{mm/turning point} \tag{15}$$

for hard packed gravel and

$$e_{TP} = 0.05 \quad \text{mm/turning point} \tag{16}$$

for asphalt.

It is interesting to note the effect of settlement on the NGS discrepancy defined by (1). Let ε_S and ε_L denote the errors in the short and long sight running, respectively. Also, without any loss of generality, let dH (forward direction) be positive. Then for $dH_S=dH$ in the forward direction, the NGS discrepancy becomes

$$d_{NGS} = |dH + \varepsilon_S| - |-dH + \varepsilon_L| = \varepsilon_S + \varepsilon_L \tag{17}$$

On the other hand, for $dH_L=dH$ in the forward direction,

$$d_{NGS} = |-dH + \varepsilon_S| - |dH + \varepsilon_L| = -(\varepsilon_S + \varepsilon_L) \tag{18}$$

For negative dH the signs for d_{NGS} are reversed.

Clearly, the settlement effect (assumed constant along the line) cancels when accumulated if the direction of the short and long sight runnings are periodically reversed so that the frequency in both directions are balanced. This assumption was largely met in the Palmdale experiment. We note that the effect on individual discrepancies is masked by random noise which is generally larger in magnitude. Furthermore, when the sense of direction of leveling is not respected, further masking of the settlement effect takes place. Thus, although this effect is always present, it will be all but obliterated in the discrepancy used by the NGS.

Rod Miscalibration

The rod scale error (λ) is defined by

$$dH = (1+\lambda)\,dH^* \tag{19}$$

where $dH^* = dH^*_F = -dH^*_B$ is the correct height difference and dH is the observed height difference. The effect on the discrepancy is then

$$d_\lambda = [(1+\lambda)\,dH^*]_F + [(1+\lambda)\,dH^*]_B = (\lambda_F - \lambda_B)\,dH^* \tag{20}$$

Replacing dH^* with the average observed height difference $d\overline{H}$, the contribution of rod miscalibration becomes

$$d_\lambda = (\lambda_F - \lambda_B)\,d\overline{H} = d\lambda\,d\overline{H} \tag{21}$$

where $d\lambda$ is the difference in the average scale error of the forward and backward runnings. The linear model for differential rod miscalibration may then be given by

$$d_\lambda = c_\lambda\,a_\lambda \tag{22}$$

where the argument $a_\lambda = d\overline{H}$ is known and $c_\lambda = d\lambda$ is to be estimated through regression.

The average scale error $\overline{\lambda}$ of the rod graduations clearly does not affect the discrepancy d when both runnings use the same rods and is therefore not detectable. However, the differential effect (i.e., variations $d\lambda$ in scale errors along the length of the rod) may be identified providing the same parts of the rod were consistently observed during each running, which would be the case when leveling along a constant grade.

Because of the constant grade along this leveling line and the relatively constant sight lengths during each running, it may be expected that scale error plays an important role in this experiment. That is, the long sight runnings would typically observed the ends of the rods, while the short sight runnings would observe closer to the middle of the rods. Thus the differential scale error $d\lambda$ represents the difference in average scale errors at the ends and middle of the rods.

The differential effect should not be larger than that implied by the limiting precision of the rod calibration. For the Palmdale experiment, both rods were calibrated at every graduation using a laser interferometer [*Whalen and Strange*, 1983]. *Balázs and Young* [1982] report the accuracy of this calibration to be better than 50 µm, while *R. S. Stein* (personal communication, 1984) claims it is closer to 10 µm. This latter level of accuracy has also been quoted by others [cf. *Heister et al.*, 1983].

Nevertheless, it should be noted that the calibration was performed with the rods in a horizontal position, whereas in practical use they are in a vertical position. Several investigators have shown that this causes significant changes in the scale [e.g., *Gottwald and Witte*, 1983; *Maurer and Schnadelbach*, 1983]. *Gottwald and Witte* [1983] have reported distortions of up to 35 μm: thus even though the rods may be perfectly calibrated in the horizontal position, significant scale errors up to 35 μm may be present when the rod is in actual use.

Other Arguments

Virtually any quantity characterizing the conditions under which the leveling was performed can be used to search for a systematic behavior of the section discrepancies. The existence of a significant correlation (i.e., regression) may imply the presence of some type of systematic effect dependent on the argument and would warrant further investigation.

In the analyses presented here, the following arguments were investigated in addition to those described earlier:

- a_H height of end bench mark of the section relative to the beginning of the line;
- a_S section length;
- a_L accumulated section length;
- a_{slope} average slope of section;
- a_{dT} difference (T_F-T_B) between average air temperatures of the forward and backward runnings;
- a_{NGS} difference (R_F-R_B) between the NGS computed refraction corrections for the forward and backward runnings. The corrections were computed for each setup using Kukkamäki's single sight refraction correction with observed temperatures. This is different from that described in (8) which is based on the average balance sight correction for the entire section [cf. *Whalen and Strange*, 1983].

It should be noted that any of these arguments may also be correlated with other effects that we are not aware of.

IDENTIFICATION OF SYSTEMATIC EFFECTS

The premise upon which this study is based is that the presence of systematic effects in the discrepancies may be revealed through the presence of trends in the data series (only linear trends are estimated here) and/or autocorrelation among the discrepancies.

The approach used here is that developed by the authors and it involves an application of time series (or, in this context, data series) analysis techniques [cf. *Vaníček and Craymer*, 1983*a*,*b*; *Craymer*, 1984; *Vaníček et al.*, 1985]. Various data series are constructed from the discrepancies by simply ordering them with respect to the argument of concern. Complete descriptions of the techniques employed are given in the above quoted references. Briefly, the steps involve (1) creation of the discrepancy series, ordered with respect to the monotonically increasing argument of interest, (2) removal of bias and trend from the series (i.e., transformation to a stationary series), (3) least squares spectral analysis of series residuals, (4) transformation of least squares spectrum to autocorrelation function (ACF) using a standard cosine transform, (5) high-frequency smoothing of the computed ACF using a Gaussian weighted moving average filter as described by *Vaníček et al.* [1985]. This is done since we are only interested in the general shape of the ACF.

As pointed out already, statistically significant trends and autocorrelations in the discrepancy series indicate the presence of systematic effects. In addition, the ACF may be used to help in the construction of proper linear models that account for the various systematic effects. Residuals from suitably developed linear models should exhibit no autocorrelation.

Analyses of the original discrepancies, corrected for level collimation, rod scale calibration, rod scale temperature, and astronomic effects but not for refraction, are summarized in column 3 of Table 1. These results are used solely as diagnostic tools to indicate the arguments that should be used in the multiple linear regression as described in the next section. In Table 1 the critical uncertainty level (1-α) for the null hypothesis that "the tested quantity equals to zero" is enclosed in parentheses following the estimated value. Thus the larger 1-α, the greater the misgiving in accepting the null hypothesis and the greater the probability that the tested quantity does not equal to zero.

As expected, the series for the refraction arguments (a_R and a_{NGS}) display highly significant trends when no refraction correction is applied to the discrepancies (column 1). The trend for a_R is 4.39 x 10^{-5}mm/(°C m^3) and agrees well with Kukkamäki's estimate of A (6.45 x 10^{-5}mm/(°C m^3)). The series for the NGS refraction correction argument (a_{NGS}) also has a very significant trend of 82% thereby indicating that this correction apparently overcompensates for the actual effect by about 18%. Furthermore, the ACF of the residual discrepancies for this series displays a long-period trend (see Figure 3) characteristic of the case when linear modelling is inappropriate. For an explanation of this, see *Vaníček and Craymer* [1983*a*].

The apparent absence (disguised by low significance levels) of other effects is likely a result of the refraction effect masking the others. For typical leveling lines, where the refraction effect is not purposely amplified, other effects are more easily identified [cf. *Vaníček et al.*, 1985]. The presence of a significant absolute term (i.e., the intercept or bias) is puzzling, and we have no explanation for it.

Analyses of the discrepancies, with the NGS refraction correction (a_{NGS}) applied, are summarized in column 4 of Table 1. These results support the above conclusions that the NGS refraction correction somewhat overcompensates for the effect and the residuals from the series for the NGS refraction argument still display autocorrelation (see Figure 4).

It is interesting to see how other effects are now becoming significant (e.g., $d\overline{H}$, slope). The model linear in $d\overline{H}$ is now inadequate; the corresponding ACF has a clear exponential trend (see Figure 5). We also note that the application of the NGS refraction correction accounts for the lowering of the absolute bias by almost 60% and the reduction of its uncertainty level to about 75%.

ESTIMATION OF SYSTEMATIC EFFECTS

The presence of more than one type of systematic effect in a data series limits the usefulness of the linear trends computed from the discrepancy series. In this case a multiple linear regression approach should be used to estimate simultaneously the linear trends for the multiple arguments involved.

The general regression model for the discrepancies as a

TABLE 1. Results of Discrepancy Series Analysis With and Without NGS Refraction Correction

Argument	Attribute	Discrepancy Series (Without NGS Correctio)	Discrepancy series (With NGS Correction)
$a_s{\approx}dH$ (m)	s_v (mm)	1.59	1.03
	bias (mm)	0.38 (94%)	0.16 (78%)
	trend (ppm)	-12 (48%)	-30 (99%)
	LSS shape	few medium peaks	few small/medium peaks
	ACF shape	small, wavy	exponential trend
a_H (m)	s_v (mm)	1.58	1.08
	bias (mm)	0.38 (94%)	0.16 (76%)
	trend (ppm)	10 (62%)	-0.01 (52%)
	LSS shape	few medium peaks	several medium peaks
	ACF shape	small, wavy	small wavy
a_S (km)	s_v (mm)	1.58	1.07
	bias (mm)	0.38 (94%)	0.16 (77%)
	trend (mm/km)	-0.785 (69%)	-0.798 (86%)
	LSS shape	few medium peaks	few medium peaks
	ACF shape	long-period trend	small, wavy
a_L (km)	s_v (mm)	1.58	1.09
	bias (mm)	0.38 (94%)	0.16 (76%)
	trend (mm/km)	0.014 (68%)	-0.004 (32%)
	LSS shape	few medium peaks	several medium peaks
	ACF shape	small, periodic	small, wavy
a_{slope} (mm/mm)	s_v (mm)	1.58	1.05
	bias (mm)	0.38 (94%)	0.16 (78%)
	trend (mm)	-1.04×10^{-4} (46%)	-2.43×10^{-4} (97%)
	LSS shape	long-period peak	few medium peaks
	ACF shape	negative linear trend	long-period trend
a_{dT} (°C)	s_v (mm)	1.58	1.07
	bias (mm)	0.38 (94%)	0.16 (77%)
	trend (mm/°C)	0.713 (68%)	-0.686 (85%)
	LSS shape	few medium peaks	few medium/large peaks
	ACF shape	small, wavy	small, periodic
a_{TP} (turn pt)	s_v (mm)	1.59	1.09
	bias (mm)	0.38 (94%)	0.16 (76%)
	trend (mm/tp)	-0.011 (35%)	-0.007 (35%)
	LSS shape	flat	many small peaks
	ACF shape	flat	small exponential trend
a_R (°C m^3)	s_v (mm)	1.05	1.07
	bias (mm)	0.38 (99.5%)	0.16 (77%)
	trend (mm/(°C m^3))	4.39×10^{-5} (>99.9%)	-0.75×10^{-5} (85%)
	LSS shape	many small peaks	several medium peaks
	ACF shape	small exponential trend	small, wavy
a_{NGS} (mm)	s_v (mm)	1.06	1.06
	bias (mm)	0.38 (99.5%)	0.16 (77%)
	trend (mm/mm)	-0.82 (>99.9%)	0.180 (94%)
	LSS shape	many small peaks	several medium/large peaks
	ACF shape	small negative trend	

linear function of the arguments $a_{i1},\cdots,a_{iu}$ is given in matrix notation by [*Vaníček and Krakiwsky*, 1982]

$$\mathbf{d} = \mathbf{B}\,\mathbf{c} + \mathbf{e} \tag{23}$$

where **d** is a vector of the n observed discrepancies, **c** is a vector of u coefficients containing the parameters (i.e., trends) to be estimated, **e** is a vector of residuals to be somehow minimized by **c** and **B** is a $(n \times u)$ design matrix containing the values of the arguments as its columns

$$\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1u} \\ a_{21} & a_{22} & \cdots & a_{2u} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nu} \end{bmatrix} \tag{24}$$

i.e., for the ith section, a_{i1} is the value of the first argument, a_{i2} the value of the second, etc.

Note that the above **B** matrix does not provide for bias

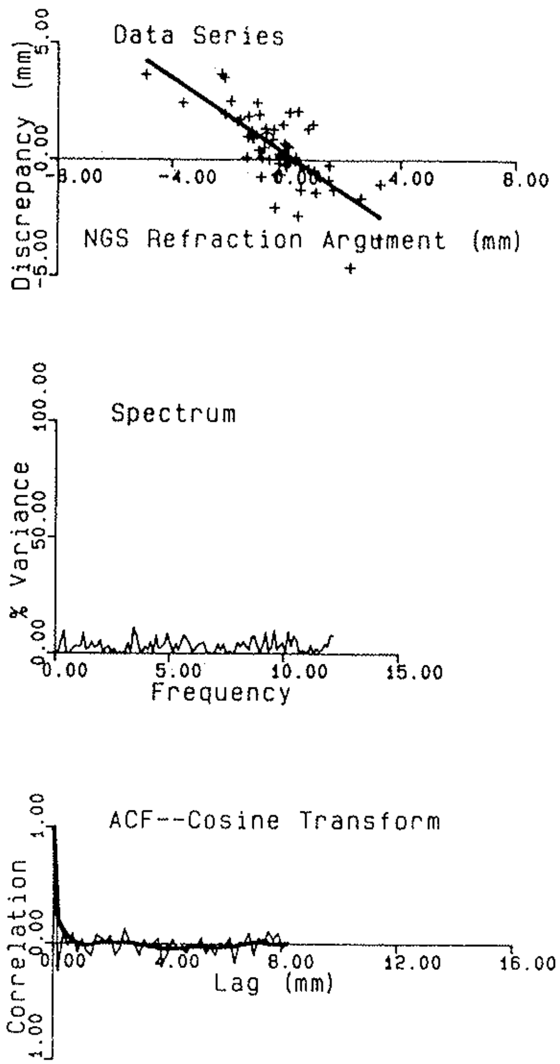

Fig. 3. Discrepancy series analysis (without refraction correction) for NGS refraction argument (a_{NGS}).

estimation even though it appears significant from the analyses of individual effects. If this were desired, there would be an extra column of ones. Because we have (as yet) no physical explanation for this constant effect, we have somewhat arbitrarily omitted it in our analyses. However, its lower magnitude and significance level after the application of the NGS refraction correction (see Table 1) seems to justify our decision to neglect it. This point clearly needs more investigation.

The regression coefficients may be solved for by minimizing the quadratic norm of the residuals. We assume the arguments to be statistically independent and the discrepancies equally weighted so that the least squares solution $\hat{c}$ and it's covariance matrix $C_{\hat{c}}$ are given by [*Vaníček and Krakiwsky*, 1982]

$$c = (B^T B)^{-1} B^T d \tag{25}$$

$$C_c = s_o^2 (B^T B)^{-1} \tag{26}$$

where

$$s_o^2 = (Bc - d)^T (Bc - d) / (n-u) \tag{27}$$

The best fitting model was selected by comparing results from the various combinations of the different kinds of arguments. In order to reduce the number of models to be compared, a backward, stepwise approach was used [cf. *Neter and Wasserman*, 1974] where the least statistically significant arguments are eliminated one-by-one in a stepwise fashion until only significant ones remain.

Initially, we included all arguments. However, correlation among the arguments can limit the usefulness of multiple regression analyses [*Neter and Wasserman*, 1974]. This is often referred to as multicollinearity. In such cases it becomes difficult to separate the effects of the the individual kinds of arguments. In our analyses we have avoided such problems by omitting from the pairs of correlated arguments the ones for which we have no physical explanation for its effect on the discrepancy.

The least significant coefficient for each model (i.e., step) is determined on the basis of the uncertainty of the hypothesis that "each c_i=0," using the Fisher F statistic [*Srivastava and Carter*, 1983]. The argument for which the uncertainty of the hypothesis is the lowest is rejected from the model. This is repeated for the subsequent model until all remaining coefficients c_i have uncertainty levels greater than 95%.

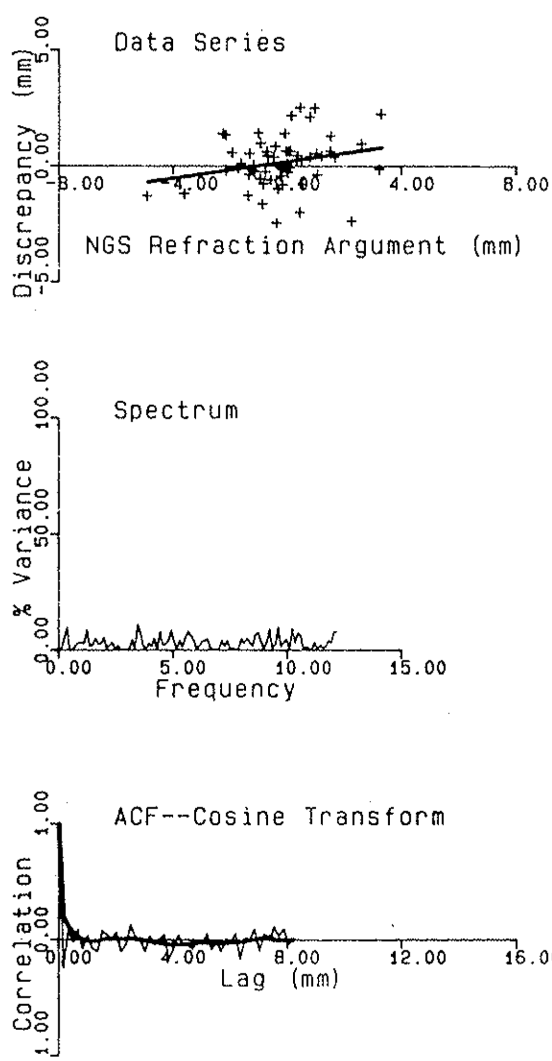

Fig. 4. Discrepancy series analysis (with NGS refraction correction) for NGS refraction argument (a_{NGS}).

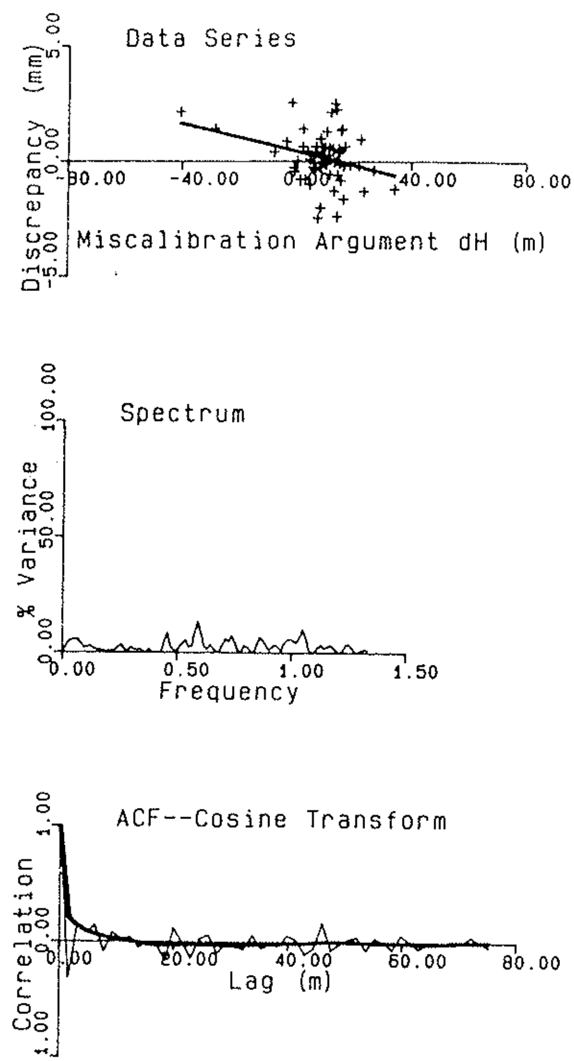

Fig. 5. Discrepancy series analysis (with NGS refraction correction) for argument $a_s = dH$.

A number of other criteria were also utilized to assess the suitability of the multilinear models. These were (1) the critical uncertainty level (1-α) for accepting the hypothesis that c=0, using the Fisher *F* statistic [*Srivastava and Carter*, 1983], (2) the coefficient of determination (i.e., the percentage of the total sum of squares of the discrepancies explained by the model [*Neter and Wasserman*, 1974]), (3) the sum of the residual discrepancies Σv along with the critical uncertainty level for accepting the hypothesis that $\Sigma v=0$, using the student-*t* statistic [*Srivastava and Carter*, 1983], (4) the estimated standard deviation s_v of the residual discrepancies, (5) the estimated standard deviation s_v(1 km) of the residual discrepancies for 1 km of leveling, and (6) the percentage of the partial sums of residual discrepancies $(\Sigma v)_i$ falling outside the limits for uncorrelated random error propagation based on s_v(1 km) (68% of the $(\Sigma v)_i$ should fall within these bounds if their probability density function is normal).

The results of such an analysis of the discrepancies are summarized in Table 2. Here, only the final model from the stepwise regression analysis (containing arguments a_λ, a_{TP}, a_R) of the discrepancies uncorrected for refraction is given in column 3. In order to avoid the above mentioned problems of multicollinearity, the initial model included all of the arguments used in the data series analysis except a_L which was highly correlated with a_H.

From Table 2 it can be seen that the critical significance levels for the arguments a_R, a_{TP} and a_λ are greater than the 95% rejection limit. This three-parameter model reduces the variation in the discrepancies by 61% thereby reducing the standard deviation of the discrepancies from 1.6 mm to 1.0 mm. The residual discrepancies now accumulate to only 3.39 mm. The frequency of partial sums of residuals $(\Sigma v)_i$ falling within the limits for random error propagation based on the sample estimate s_v(1 km) is 93%, well outside the 68% limit for normal Gaussian behavior.

As expected, the coefficient $\hat{c}_R = A = 4.6\times10^{-5}$ mm/(°C m³) for the refraction argument a_R is highly significant and again agrees reasonably well with the estimate for $A = 6.5\times10^{-5}$ mm/(°C m³) given by *Kukkamäki* [1939]. The difference is likely a result of specific environmental conditions in California. Note that the contribution of the refraction effect to the total accumulated discrepancy is the least (13.56 mm) of the three modeled effects.

The settlement coefficient $\hat{c}_{TP} = 0.014$ mm/turning point is expected as the effect does not cancel in the discrepancy. The magnitude of the effect agrees remarkably well with Anderson's estimate (0.02 mm/turning point) and contributes the most (20.26 mm) to the total accumulated discrepancy.

The dependence on $d\bar{H}$ has been attributed by us to differential rod miscalibration where the SSL and LSL readings deal consistently with different parts of the rods. As explained earlier, this may be a direct consequence of the distortion of the vertical rod scale relative to its calibrated horizontal position. The coefficient $\hat{c}_s = -26$ ppm agrees well with the effects measured by *Gottwald and Witte* [1983] and contributes -14.15 mm to the total accumulated discrepancy. This effect is actually a by-product of the design of this leveling experiment. Using typical leveling procedures, the same areas of the rods would generally be observed during both forward and backward runnings and the effect would tend to cancel or randomize.

The high degree of dependence of the above coefficient estimates on specific "influential" observations has been questioned by R. S. Stein (personal communication, 1985). However, inspection of the *t* statistic for the difference in the regression coefficients (Δc) when an observation is removed from the sample reveals no significant influences from any of the observations. All of the corresponding *t* statistics are less than 0.6. Thus, the hypothesis that Δc=0 when the ith observation is removed passes for all observations at any reasonable significance level.

Finally, an analysis was also performed on the series of residuals from the three-parameter regression model. The results display no statistically significant trends or autocorrelations thereby indicating the success of the model to account for all systematic effects present. Figure 6 gives the residual series for the refraction argument. Clearly, no refraction effect appears to remain as witnessed by the absence of trend and autocorrelation.

In order to assess the validity of our approach to modeling the refraction effect, analyses were also performed using the NGS computed refraction argument a_{NGS}. By applying this refraction correction prior to the analyses, one

TABLE 2. Results of Multiple Linear Regression Analyses.

Argument	Attribute	Discrepacies Without NGS Correction	Discrepancies With NGS Correction	
			Model 1	Model 2
a_R (°C m^3)	c_R (mm/(°C m^3))	4.60x10^{-5}		
	uncert (c_R=0)	> 99.9 %		
	effect on Σd (mm)	13.56		
a_{NGS} (mm)	c_{NGS} (%)		15 %	
	uncert (c_{NGS}=0)		87 %	
	effect on Σd (mm)		-1.92	
a_{TP} (turn pt)	c_{TP} (mm/turn pt)	0.014	0.014	0.014
	uncert (c_{TP}=0)	96 %	97 %	97 %
	effect on Σd (mm)	20.26	21.57	21.38
$a_s = dH$ (m)	c_s (ppm)	-26	-25	-27
	uncert (c_s=0)	97 %	95 %	97 %
	effect on Σd (mm)	-14.15	-13.23	-14.77
Model assessment	uncert (**c**=0)	> 99.9 %	96 %	94 %
	% of var explained	61 %	13 %	9 %
	s_d (mm)	1.58	1.08	1.08
	s_v (mm)	1.00	1.01	1.03
	s_d(1 km) (mm)	1.75	1.22	1.22
	s_v(1 km) (mm)	1.11	1.14	1.15
	Σd (mm)	23.05	9.87	9.87
	Σv (mm)	3.39	3.45	3.26
	uncert (Σv = 0)	33 %	34 %	32 %
	frequency of $\Sigma d < s(L)$	53 %	35 %	35 %
	frequency of $\Sigma v < s(L)$	93 %	100 %	90 %

can see if any other effects (including residual refraction) are present.

Initially, we included the arguments for rod miscalibration, turning point settlement, and the NGS refraction correction (i.e., a_λ, a_{TP}, and a_{NGS}). The results of the regression analysis are given in Table 2 under model 1. Clearly, the results are similar to those obtained with our refraction argument. Note that the NGS refraction argument gives a slight trend indicating that it overcompensates by about 15%. However, its uncertainty level is not very high (87%). This "borderline" level of uncertainty for such a small sample does not firmly prove or disprove the effectiveness of the NGS refraction correction. Further investigations are clearly needed to resolve this.

As a final test of the NGS refraction correction, a third model was analyzed using only the arguments for rod miscalibration and settlement (i.e., a_λ and a_{TP}). Again, similar results were obtained as for the previous two models. These results are listed in Table 2 under model 2.

The most obvious result of these analyses is that rod settlement and an effect dependent on $d\overline{H}$ also appear to affect the observed discrepancies. Thus any examination of these data should account for these effects. In fact, although the refraction effect is statistically the strongest, i.e., the most clearly identifiable, it contributes the least to the total accumulated discrepancy as compared with rod settlement and the differential rod miscalibration effects.

CONCLUSIONS

It has been shown, both physically and statistically, that the settlement effect is always present in the discrepancies. When analyzing the discrepancy between forward and backward runnings, the effect does not even partially cancel as these discrepancies are accumulated. Thus any examination of these discrepancies should properly account for this effect. It should be noted, however, that the settlement effect does cancel in the accumulation of the discrepancies between the short and long sight runnings as used in the NGS analysis of the experiment by *Whalen and Strange* [1983] and *Stein et al.* [this issue]. This is due to the randomization of the sign of the settlement error as a result of the balancing of the direction of the runnings.

The result of our analyses has been the discovery that rod and instrument settlement appear to have affected the discrepancies between forward and backward runnings even more than refraction. Although the settlement and miscalibration effects are not clearly visible when analyzed individually, their statistical significance and the good

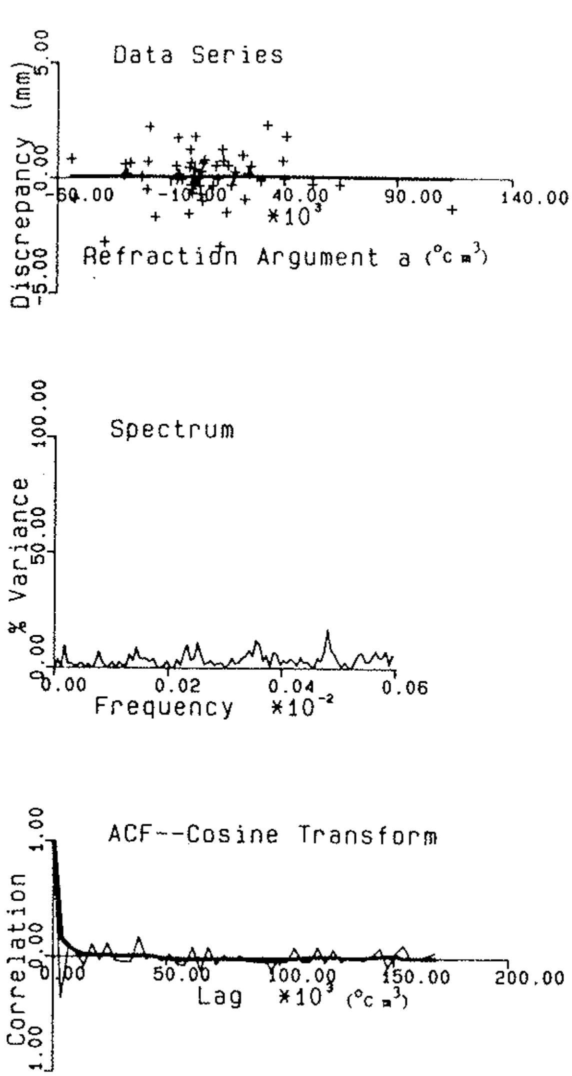

Fig. 6. Residual discrepancy analysis for refraction argument (a_R).

agreement of the values resulting from the multiple linear regression with independent estimates convinces us that these effects are real.

It should be borne in mind that the settlement effect has much less severe consequences on the average section height difference if the field work is performed adhering to the specifications [cf. *Federal Geodetic Control Committee*, 1980] as is done in routine leveling. If the magnitude of the effect as determined here is taken to be realistic, an abnormally large setup imbalance of the order of 10 setups would only produce an error of about 0.1—0.2 mm in the average section height difference.

On the other hand, the miscalibration effect may be dangerous when the leveling is carried out along a slope of constant gradient. Since the same parts of the rod would tend to be used for both forward and backward runnings, any miscalibration error will accumulate and affect both the discrepancies and average elevation differences.

Concerning refraction, the NGS correction seems to have worked reasonably well. Our results confirm those of *Whalen and Strange* [1983] and *Stein et al.* [this issue]. Nevertheless, it may also be worth considering the alternate regression approach we have used here where the value of Kukkamäki's *A* function would be evaluated separately for different environmentally similar regions of the continent from actual leveling data. It should be realized, however, that in practice the refraction effect will tend to cancel in the discrepancy between the forward and backward runnings. Thus some other quantity would have to be analyzed to estimate the refraction effect (e.g., loop misclosures). This approach has already been proposed by *Remmer* [1980].

Acknowledgements. This work has been partially funded by a University of Toronto applied research grant and a Geodetic Survey of Canada research contract (OST 83-00053). In addition, Peter Wall of the University of Toronto Computing Services helped to provide most of the computer time for the analyses. We would like to thank R.O. Castle and R.S. Stein of the U.S. Geological Survey, R.R. Steeves of the Geodetic Survey of Canada and G.H. Carrera of the University of Toronto for many useful discussions and suggestions. We also extend our sincere thanks to E.I. Balazs and R.J. Lee of the National Geodetic Survey for kindly providing the leveling data and for promptly answering our questions about it. Finally, we wish to acknowledge the Dept. of Surveying Engineering at the University of New Brunswick on whose time and equipment the final camera ready copy of this paper was produced.

REFERENCES

Adams, G. W., Joint refraction experiment USGS and NGS, NGS project report, Task 83622362, Natl. Geod. Surv., Rockville, Md., May 1 to June 16, 1981.

Anderson, E. G., The effect of vertical motion of the level and rods in precise geodetic levelling, report, Div. of Surv. Eng., Univ. of Calgary, Calgary, Alberta, 1983.

Balazs, E. I., and G. M. Young, Corrections applied by the National Geodetic Survey to precise levelling observations, *NOAA Tech. Memo. NOS NGS 34*, NOS/Natl. Geod. Inf. Cent., Rockville, Md., 1982.

Castle, R. O., T. D. Gilmore, R. K. Mark, B. W. Brown, Jr., and R. C. Wilson, An examination of the southern California field test for the systematic accumulation of the optical refraction error in geodetic levelling, *Geophys. Res. Lett.*, *10*(11), 1081—1084, 1983.

Craymer, M. R., Data series analysis and systematic effects in levelling, *Surv. Sci. Tech. Rep. 6*, Univ. of Toronto, Mississauga, Ont., 1984.

Federal Geodetic Control Committee, *Specifications to support classification, standards of accuracy, and general specifications of geodetic control surveys*, Natl. Ocean Surv., NOAA, U.S. Dep. of Commer., Washington, D.C., June 1980.

Gottwald, R., and B. Witte, A microcomputer-controlled comparator for visual and photoelectric calibration of levelling rods, in *Precise Levelling: 38 Contributions to the Workshop on Precise Levelling, Hannover, Fed. Rep. of Germany, 16—18 March 1983*, edited by H. Pelzer and W. Niemeier, pp. 127—140, Dummler Verlag, Bonn, 1983.

Heister H., W. Caspary, and R. Tille, Problems in reading the graduation lines on levelling staffs by photoelectric methods, in *Precise Levelling: 38 Contributions to the Workshop on Precise Levelling, Hannover, Fed. Rep. of Germany, 16—18 March 1983*, edited by H. Pelzer and W. Niemeier, pp. 141—153, Dummler Verlag, Bonn, 1983.

Holdahl, S. R., A model of temperature stratification for the correction of levelling refraction, *Bull. Geod.*, *55*(3), 231—249, 1981.

Holdahl, S. R., Recomputation of vertical crustal motions near Palmdale, California, 1959-1975, *J. Geophys. Res.*, *87*(B11), 9374—9388, 1982.

Kukkamäki, T. J., Formeln und tabellen zur berechnung der nivellischen refraktion, *Finn. Geod. Inst. Publ. 27*, Helsinki, 1939.

Maurer, W., and K. Schnadelbach, First experiences with a vertical comparator for the calibration of invar rods, in *Precise Levelling: 38 Contributions to the Workshop on Precise Levelling, Hannover, Fed. Rep. of Germany, 16—18 March 1983*, edited by

H. Pelzer and W. Niemeier, pp. 155—163, Dummler Verlag, Bonn, 1983.

Neter, J., and W. Wasserman, *Applied Linear Statistical Models*, Richard D. Irwin, Homewood, Ill., 1974.

Remmer, O., Role and evaluation of refraction for the processing of levelling networks, in *Proc. of the 2nd International Symposium on Problems Related to the Redefinition of North American Vertical Geodetic Networks, Ottawa, 26—30 May 1980*, pp. 623—646, Can. Inst. of Surv., Ottawa, 1980.

Srivastava, M. S., and E. M. Carter, *An Introduction to Applied Multivariate Statistics*, North-Holland, New York, 1983.

Stein, R. S., and W. Thatcher, Field test for refraction in levelling, southern California, *Eos Trans. AGU, 63*(45), 1106, 1982.

Stein, R. S., W. Thatcher, W. E. Strange, S. R. Holdahl, and C. T. Whalen, Field test for the accumulation of differential refraction in southern California (abstract), paper presented at Advances in Geodetic Refraction, General Meeting, Int. Assoc. of Geod., Tokyo, May 7—20, 1982.

Stein, R. S., C. T. Whalen, S. R. Holdahl, W. E. Strange, and W. Thatcher, Saugus-Palmdale, California, field test for refraction error in historical levelling surveys, *J. Geophys. Res.*, this issue.

Strange, W. E, Scribing, graduation and calibration of U.S. Coast and Geodetic Survey levelling rods from 1877 to 1968, *NOAA Tech. Rep. NOS 94 NGS 23*, NOS/Natl. Geod. Inf. Cent., Rockville, Md., 1982.

Vaníček, P., and M. R. Craymer, Autocorrelation functions as a diagnostic tool in levelling, in *Precise Levelling: 38 Contributions to the Workshop on Precise Levelling, Hannover, Fed. Rep. of Germany, 16—18 March 1983*, edited by H. Pelzer and W. Niemeier, pp. 327—334, Dummler Verlag, Bonn, 1983a.

Vaníček, P., and M. R. Craymer, Autocorrelation functions in the search for systematic errors in levelling, *Manuscripta Geodaetica, 8*, 321—341, 1983*b*.

Vaníček, P., and E. J. Krakiwsky, *Geodesy: The Concepts*, North-Holland, Amsterdam, 1982.

Vaníček, P., G. H. Carrera, and M. R. Craymer, Corrections for systematic errors in the Canadian levelling networks, *Can. Geod. Surv. Contract Rep. 83-00053*, Surv. and Mapp. Branch, Dep. of Energy, Mines and Resour., Ottawa, 1985.

Whalen, C. T., and W. E. Strange, The 1981 Saugus to Palmdale, California levelling refraction test, *NOAA Tech. Rep. NOS 98 NGS 27*, NOS/Natl. Geod. Inf. Cent., Rockville, Md., 1983.

M. R. Craymer and P. Vaníček, Survey Science, University of Toronto, Mississauga, Ontario, Canada L5L 1C6.

(Received April 16, 1985;
revised January 6, 1986;
accepted March 25, 1986.)