JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 100, NO. B8, PAGES 15129-15145, AUGUST 10, 1995

Estimation of rod scale errors in geodetic leveling

Michael R. Craymer
Geodetic Survey Division, Geomatics Canada, 615 Booth Street, Ottawa, Ontario

Petr Vaníček
Department of Surveying Engineering, University of New Brunswick, Fredericton, Canada

Robert O. Castle
U.S. Geological Survey, Menlo Park, California

Abstract. Comparisons among repeated geodetic levelings have often been used for detecting and estimating residual rod scale errors in leveled heights. Individual rod-pair scale errors are estimated by a two-step procedure using a model based on either differences in heights, differences in section height differences, or differences in section tilts. It is shown that the estimated rod-pair scale errors derived from each model are identical only when the data are correctly weighted, and the mathematical correlations are accounted for in the model based on heights. Analyses based on simple regressions of changes in height versus height can easily lead to incorrect conclusions. We also show that the statistically estimated scale errors are not a simple function of height, height difference, or tilt. The models are valid only when terrain slope is constant over adjacent pairs of setups (i.e., smoothly varying terrain). In order to discriminate between rod scale errors and vertical displacements due to crustal motion, the individual rod-pairs should be used in more than one leveling, preferably in areas of contrasting tectonic activity. From an analysis of 37 separately calibrated rod-pairs used in 55 levelings in southern California, we found eight statistically significant coefficients that could be reasonably attributed to rod scale errors, only one of which was larger than the expected random error in the applied calibration-based scale correction. However, significant differences with other independent checks indicate that caution should be exercised before accepting these results as evidence of scale error. Further refinements of the technique are clearly needed if the results are to be routinely applied in practice.

Introduction

During the past 15 years there has been much debate regarding the accuracy of geodetic leveling and its suitability as a tool for assessing vertical displacements due to crustal motion. Although it is generally agreed that comparisons among repeated levelings may provide very precise determinations of vertical movements of the Earth's surface, a number of geodesists and geophysicists have argued that many of these determinations are merely artifacts attributable to systematic errors in the leveling procedure. Of particular concern are errors proportional to height differences, such as rod scale calibration errors and refraction effects (see, e.g., *Vaníček et al.* [1980]).

One of the most intense debates has focused on the existence of the so-called Palmdale Bulge in southern California. From an analysis of more than 10,000 km of leveling, *Castle et al*. [1976; 1984] and *Castle and Gilmore* [1992] concluded that an approximately 80,000 km^2 area had sustained uplift of as much as 45 cm between 1955 and 1974, followed by partial collapse. On the other hand, *Jackson and Lee* [1979], *Jackson et al*. [1980], and *Reilinger and Brown* [1981] argued that this apparent uplift is largely the product of a height-dependent systematic error in some or all of the leveling measurements; specifically, scale errors due to errors in the laboratory calibrations of the graduated invar strips housed in the leveling rod frames (referred to here as simply rod scale errors). *Mark et al*. [1981] examined the calibration measurements and procedures in some detail and concluded that rod scale errors could have accounted for only a “trivial” part of the measured uplift.

In order to provide a better foundation for these analyses, *Stein* [1981] undertook a statistical regression analysis of selected parts of the southern California leveling record. He concluded that levelings between 1953 and 1979 contain a height-dependent average linear error of 3 ± 46 ppm (95% confidence interval). Even though it is statistically insignificant, this error was attributed to errors in the applied rod scale corrections, rather than changes in calibration and field procedures. Stein also concluded that this effect would tend to randomize over lines longer than about 80 km and could thus be treated as a random error. Nevertheless, he identified rod-pairs (312) 268/274 and (316) 132180/87849 as ones characterized by relatively large scale errors of about +120 ± 24 ppm and –90 ± 40 ppm (95% confidence intervals), respectively. Stein also identified another unspecified rod-pair used in 1971 that contained a scale error of about 40 ppm.

After examining both his procedure and his data we concluded that Stein's results are somewhat questionable. For example, it has since been determined that the leveling with rodpair (316) 132180/87849 used an improperly computed rod scale correction [*Mark et al*., 1981] and was based on a Zeiss Ni 1 level, an instrument now known to be subject to a potentially large magnetic error [*Rumpf and Meurisch*, 1981]. Of particular concern was the subjective omission of as much as 20% of the data. That is, data were often discarded simply because their omission strengthened the regressions [*Stein*, 1981, pp. 443, 445]. Moreover, Stein's solution for the individual rod-pair scale errors was further subjective because the system of normal equations to be solved was singular (i.e., an infinite number of solutions exist) and no justification was given for the particular solution selected.

Our purpose here is to present a more objective and rigorous statistical analysis based on a refinement of Stein's technique for estimating height-dependent errors in historic leveling; specifically, linear rod-pair scale errors developed from a broader spectrum of leveling data. Reliable estimates of this error are important in a wide range of activities including: (1) monitoring of critically engineered structures or sites such as dams, nuclear power plants, and high-level radioactive waste disposal sites, (2) measurements of compaction-induced subsidence, (3) geophysical studies related to earthquake prediction and analysis, and (4) geologic framework investigations.

Methodology

Our analysis is based on comparisons of repeated levelings over the same line segments. It is essentially a two-step process based on the procedure suggested by *Stein* [1981]. In the first step, models describing rod-pair scale errors are fitted to the differences between two levelings. This step provides estimates of differences in rod-pair scale errors between levelings. These results are then combined in the second step where individual rod-pair scale errors are determined. This approach is analogous to methods employed in processing GPS data for differential positioning, where coordinate differences (baseline vectors) are separately estimated and subsequently combined in a network adjustment for estimating individual position coordinates.

Before describing the details of this technique we first consider the rod-pair scale models used in our analyses. In particular, we assess the validity of the widely accepted assumption that rod scale errors are linearly proportional to height, height difference, and slope. Details of the estimation method used in the first step are then followed by a description of the method used to estimate the individual rod-pair scale errors in the second step.

Models for Rod-Pair Scale Errors

It is often assumed that rod scale errors are proportional to topographic height or height differences (e.g., *Bomford* [1971, p. 240], *Jackson and Lee* [1979], *Jackson et al*. [1981], *Mark et al*. [1981], *Reilinger and Brown* [1981], *Stein* [1981] and *Strange* [1980; 1981]). This intuitive assumption, however, is less supportable than one might think, because the rods are alternated between foresight and backsight in first-order leveling. Consider an individual leveling setup (Figure 1) where rod A is used for the backsight and rod B for the foresight. The observed height difference Δh^{obs} is simply the backsight reading b^{obs} minus the foresight reading f^{obs}. If scale errors λ_A and λ_B are present in each rod and uniform (i.e., constant) along the entire length of the rod, the rod readings will be

$$b^{obs} = b\,(1+\lambda_A) \tag{1a}$$

$$f^{obs} = f\,(1+\lambda_B)\,, \tag{1b}$$

where b and f are the correct (error-free) rod readings we would obtain in the absence of any scale errors or other systematic effects. Note that we have not included index errors because they cancel in alternate setups. The resulting observed height difference for an individual setup is then

$$\begin{aligned} \Delta h^{obs} &= b^{obs} - f^{obs} \\ &= (b-f) + (b\lambda_A - f\lambda_B) \\ &= \Delta h + \varepsilon_{\Delta h}\,, \end{aligned} \tag{2}$$

where $\Delta h = b - f$ is the correct height difference and $\varepsilon_{\Delta h} = (b\lambda_A - f\lambda_B)$ is the effect of rod scale errors on Δh.

When considering the entire leveled section, the situation becomes more complex due to the alternating use of the rods for foresight and backsight in adjacent setups. The observed height difference ΔH^{obs} over a section is simply the summation of the individual setup height differences; that is,

$$\Delta H^{obs} = \sum \Delta h^{obs} + \sum \varepsilon_{\Delta h} = \Delta H + \varepsilon_{\Delta H}\,, \tag{3}$$

where ΔH is the correct section height difference, and $\varepsilon_{\Delta H}$ is the section effect of rod scale errors. Expanding the error $\varepsilon_{\Delta H}$ in terms of the individual rod scale errors gives

$$\begin{aligned} \varepsilon_{\Delta H} &= (b\lambda_A - f\lambda_B)_1 + (b\lambda_B - f\lambda_A)_2 + (b\lambda_A - f\lambda_B)_3 + (b\lambda_B - f\lambda_A)_4 + \ldots \\ &= \lambda_A\,(b_1 - f_2 + b_3 - f_4 + \ldots) + \lambda_B\,(-f_1 + b_2 - f_3 + b_4 - \ldots) \\ &= \lambda_A\,T_A + \lambda_B\,T_B\,, \end{aligned} \tag{4}$$

where the subscripts denote the setup sequence number, and T_A and T_B are the collection of b and f terms for rods A and B, respectively. Note that rods A and B are alternated as foresight and backsight rods and that $T_A + T_B = \Delta H$. Clearly, the error in this case is not a simple linear function of the section height difference, as many investigators have assumed.

In those cases where the setup height differences are the same and the sight lengths are constant (i.e., constant slope along the leveling route), $T_A = T_B = \Delta H/2$. Letting $\bar{\lambda} = (\lambda_A + \lambda_B)/2$ denote the average scale error for the rod-pair, the error in the section height difference then becomes a linear function of height difference

$$\varepsilon_{\Delta H} = (\lambda_A + \lambda_B)\,\Delta H/2 = \bar{\lambda}\,\Delta H. \tag{5}$$

The slope of the terrain need not be constant over the entire length of the section but only over adjacent (not overlapping) pairs of setups so that in (4),

$$(b_1 - f) = (-f_1 + b_2)\ ,\ \ (b_3 - f_4) = (-f_3 + b_4)\ ,\ \ \ldots \tag{6}$$

In other words, the wavelength of the terrain must be greater than twice the setup length. This condition is likely to be satisfied in practice, since levelings are usually carried out along smooth railway grades and roads whose topographic wavelength is significantly greater than the length of an adjacent pair of setups (about 100 to 150 m). Nevertheless, this assumption should be checked whenever possible; for example, by inspecting the variation of section slopes β about their mean (i.e., the standard deviation of βs).

Given the effect of rod scale errors on a section height difference over constant slope, and in the absence of any other systematic effects in the leveling or any vertical displacements, the difference $\delta\Delta H$ in the height differences between the two levelings of a section using rod-pairs k and l is due entirely to random and rod scale errors. That is, for any section,

$$\begin{aligned} \delta\Delta H^{obs} &= \Delta H_k^{obs} - \Delta H_l^{obs} \\ &= \varepsilon_{\Delta H_k} - \varepsilon_{\Delta H_l} + \eta_{\delta\Delta H} \\ &= \Delta\bar{\lambda}_{kl}\, \Delta H + \eta_{\delta\Delta H}\ , \end{aligned} \tag{7}$$

where $\eta_{\delta\Delta H}$ is the random error in $\delta\Delta H$ and $\Delta\bar{\lambda}_{kl} = (\bar{\lambda}_k - \bar{\lambda}_l)$ is the difference between the average scale errors of the rod-pairs k and l.

Similarly, the difference δH between the height of a point derived from two levelings is simply the accumulation of values given by (7) for each section along the line. For the i^{th} point we get

$$\delta H_i^{obs} = \sum_{j=1}^{i} \delta\Delta H_j + \eta_{\delta H} = \Delta\bar{\lambda}_{kl}\, H_i + \eta_{\delta H}\ , \tag{8}$$

where H_i is the height relative to the beginning of the line and $\eta_{\delta H}$ is the random error in δH. Again, the difference between average scale errors for the rod-pairs used for the two levelings appears on the right-hand side.

Finally, the difference $\delta\beta$ between the section slopes β derived from two levelings are obtained by dividing (7) by the section length S. This results in

$$\delta\beta^{obs} = \Delta\bar{\lambda}_{kl}\, \frac{\Delta H}{S} + \eta_{\delta\beta} = \Delta\bar{\lambda}_{kl}\, \beta + \eta_{\delta\beta}\ , \tag{9}$$

where $\eta_{\delta\beta}$ is the random error in $\delta\beta$. Because successive levelings may follow slightly different routes, the section length in the above equation may not be the same in both levelings. In calculating the slope we always use the shortest section length. This has the effect of maximizing the slope.

The variances that should be used for weighting the observed differences δH, $\delta\Delta H$, and $\delta\beta$ are derived from the expected random errors for leveling, the weight being the inverse of the variance. For the section height difference ΔH the random error is generally assumed to be proportional to the square root of the section length S [*Vaníček et al.*, 1980]. The variance is then

$$\sigma^2_{\Delta H} = \kappa^2 S\ , \tag{10}$$

where κ is the random error per square root of distance in kilometers, a typical value for first-order leveling being about 1 mm km$^{-1/2}$. For two successive levelings of the same accuracy the variance of the difference $\delta\Delta H$ between the levelings is twice as large; that is,

$$\sigma^2_{\delta\Delta H} = 2\, \sigma^2_{\Delta H} = 2\kappa^2 S\ . \tag{11}$$

This assumes the two levelings are uncorrelated (any correlations, if present, will tend to be positive and thus reduce the error in $\delta\Delta H$). The weights are then computed as the inverse of these variances.

The weights for δH are obtained from the variances of $\delta\Delta H$. The difference δH is simply the accumulation of the section differences $\delta\Delta H$ along the leveling line. Its variance $\sigma^2_{\delta H}$ is therefore the cumulative summation of the variances of $\delta\Delta H$. For point P_i,

$$\sigma^2_{\delta H_i} = 2\, \kappa^2 \sum_{j=1}^{i} S_j\ . \tag{12}$$

The weight is the inverse of this variance; that is, the weight is inversely proportional to the cumulative length along the leveling line. However, the individual δH_i are now mathematically correlated because they are derived by accumulating the same $\delta\Delta H$ observations along the leveling line. Thus any δH is correlated with all the spatially preceding δHs, since it is derived from the same data. By the law of propagation of errors the covariance matrix $\boldsymbol{C_H}$ for the vector of heights $\boldsymbol{H}$ is given as

$$\boldsymbol{C_H} = \boldsymbol{J}\, \boldsymbol{C_{\Delta H}}\, \boldsymbol{J}^{\mathrm{T}}\ , \tag{13}$$

where $\boldsymbol{C_{\Delta H}}$ is the covariance matrix for section height differences $\boldsymbol{\Delta H}$, and $\boldsymbol{J}$ is the Jacobian matrix of transformation. For $\boldsymbol{\Delta H}$ ordered with respect to their location along the line,

$$\boldsymbol{J} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{bmatrix}. \tag{14}$$

For uncorrelated and equally weighted $\boldsymbol{\Delta H}$ the covariance matrix will then have the form

$$C_H = \sigma^2_{\Delta H} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & \dots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ : & : & : & & : \\ 1 & 2 & 3 & \dots & n \end{bmatrix}, \qquad (15)$$

where $\sigma^2_{\Delta H}$ is the common variance for all ΔH_i and n is the number of points. The covariance matrix for unequally weighted ΔH_i will be more complicated but easily computed from (13).

For the slope $\beta = \Delta H/S$ the propagation of errors using (10) gives

$$\sigma^2_\beta = \frac{1}{S^2} \sigma^2_{\Delta H} = \frac{\kappa^2}{S} . \qquad (16)$$

The variance for the temporal difference in slopes is then

$$\sigma^2_{\delta\beta} = 2\, \sigma^2_\beta = \frac{2\kappa^2}{S} , \qquad (17)$$

and the weight (the inverse of (17)) is directly proportional to the section length.

Because the weighting is essentially a standardization (or normalization) procedure, these weighted error models are mathematically equivalent to each other in that they provide exactly the same estimates of $\Delta\bar{\lambda}$. This is not the case, however, for the unweighted/uncorrelated models and may explain some discrepancies between our results and those of previous investigations. For example, *Mark et al.* [1981] and *Stein* [1981] both argued that any long wavelength height-dependent error will alias as a short wavelength error as a result of the interaction between the long wavelength component and short wavelength variations in benchmark (BM) spacing. However, these problems occur only when unweighted models are used; the effect of variations in BM spacing is removed by the weighting scheme. Furthermore, visual comparisons of terrain and height change profiles (plotted with respect to distance along the leveling route), or simple (uncorrelated) regressions of height change on height are often incorrectly used as evidence of rod scale errors (see Figures 2 and 3 for an example). Only when the mathematical correlations between the changes in height are incorporated are the results correct and in agreement with the other models. Throughout our investigations we elect to use only the ΔH model because of its simpler form.

It is important to remember that rod scale corrections have already been applied to the leveling data used here (see next section). Therefore we are estimating the residual average scale error for a rod-pair, that is, any scale error remaining after applying the calibration-based correction. Such residual errors may be due to the nonlinearity of the scales, changes in scale tension between levelings, or errors in the rod scale calibrations. Because the residual calibration errors may be time dependent, calibrations of the same rod-pair carried out at different times are considered to have different residual scale errors. In other words, with each calibration and resulting rod correction the same rod-pair is treated as if it were a different rod-pair with a different scale error. In order to ensure that the regression coefficients correctly refer to the individual residual scale errors for each rod-pair, the leveling lines generally had to be decomposed into shorter lines, each involving a single rod-pair.

Regression models usually include a constant offset or intercept as well as a linear trend. A constant term would make sense in our models only if it accounted for index errors in the rod scales. Because corrections for index errors had already been applied to our data, a constant term would represent only residual index errors. Moreover, any remaining residual index errors would tend to cancel in the accumulation of alternating setups. For an even number of setups the error would completely cancel. For an odd number of setups the resulting error would be equal to the difference in the residual index error for each rod pair in a single setup, a negligible quantity. Accordingly, we have not included the intercept in our regressions, assuming both that the computed index error correction was properly applied and that any residual rod-pair index error is negligible. This assumption is supported by *Stein* [1981] who found that the intercept (referred to as the "mean residual tilt") generally was statistically insignificant.

Other Systematic Effects

The occurrence of vertical displacements and systematic errors other than rod scale error may also contribute to differences in H, ΔH, and β between levelings over the same line. Vertical displacements proportional to topographic height in tectonically active regimes, such as in the Transverse Ranges and the Great Basin, present the most serious problem. Any resulting correlations will be indistinguishable from those attributable to rod scale errors.

Among the other known systematic errors, only differential refraction is proportional to ΔH [*Kukkamäki*, 1938; *Strange*, 1980]. This error also depends on other factors such as the temperature gradient, which is generally unavailable from historic leveling records but can, in theory, be modeled [*Holdahl*, 1981]. Both theory and experimental observations suggest that differential refraction errors may accumulate to very large values over gently sloping terrain [*Whalen*, 1980; *Strange*, 1980; 1981; *Stein et al.*, 1986]. However, evidence of the accumulation of this error in procedurally constrained geodetic levelings is generally nonexistent [*Castle et al.*, 1985; *Mark et al.*, 1987]. The effectiveness of these procedural constraints ultimately is attributable to the limitations that they impose on acceptable levels of atmospheric scintillation and hence refraction [*Castle et al.*, 1994].

In view of the continuing debate over the impact of the differential refraction error in geodetic leveling, we have decided at this time not to apply refraction corrections to the data used in our analysis. We believe this is justified primarily because this error will tend to randomize (or average out) when differencing leveling segments over various terrain profiles and under changing meteorological conditions.

Estimation of Rod-Pair Scale Errors

Estimates of the average scale error $\bar{\lambda}$ for each separately calibrated rod-pair are obtained using a two step process. In the first step, estimates of the differences $\Delta\bar{\lambda}$ between two levelings over the same line or line segment are obtained using a least squares fit of one of the linear models to the observed differences (we chose the ΔH model because of its simpler form). These solutions are referred to here as the "regression solutions." In the second step estimates of $\Delta\bar{\lambda}$ are assembled into a system of equations which are then solved for the individual average rod-pair scale errors $\bar{\lambda}$. The presence of suspected rod-pair scale errors are disclosed as statistically significant scale parameters.

Scale error differences ($\Delta\lambda$). Each of the three error models can be written in the same symbolic form

$$\boldsymbol{y} = \boldsymbol{x}\,\Delta\lambda + \eta\,, \tag{18}$$

where $\boldsymbol{y}$ is the vector of the observed differences (δH, $\delta\Delta H$, or $\delta\beta$), $\boldsymbol{x}$ is the vector of H, ΔH, or β, $\Delta\lambda$ is the parameter to be estimated (the difference between the average rod-pair scale errors for the repeated levelings), and η is a vector of random errors, assumed to be normally distributed. Note that the bar over λ is dropped here for notational convenience. The least squares solution of this model is given by [*Vaníček and Krakiwsky*, 1986]

$$\Delta\hat{\lambda} = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{x})^{-1}\,\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{y}\,, \tag{19a}$$

$$\sigma^2_{\Delta\hat{\lambda}} = \hat{\sigma}^2_{\mathrm{o}}\,(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{x})^{-1}\,, \tag{19b}$$

$$\hat{\sigma}^2_{\mathrm{o}} = \frac{\hat{\boldsymbol{r}}^{\mathrm{T}}\,\boldsymbol{P}\,\hat{\boldsymbol{r}}}{n-1}\,, \tag{19c}$$

$$\hat{\boldsymbol{r}} = \boldsymbol{A}\,\Delta\hat{\lambda} - \boldsymbol{y}\,, \tag{19d}$$

where $\boldsymbol{P}$ is the weight matrix for the vector of observed differences $\boldsymbol{y}$, $\sigma^2_{\Delta\hat{\lambda}}$ is the estimated variance of $\Delta\hat{\lambda}$, $\hat{\sigma}^2_{\mathrm{o}}$ is the estimated variance factor and $\hat{\boldsymbol{r}}$ is the vector of estimated observation residuals.

Individual scale errors (λ). In the second step, all the regression estimates of the differences in individual rod-pair scale errors $\Delta\hat{\lambda}$ are combined into a system of equations, which are solved for the individual residual scale errors $\hat{\lambda}$ for individually calibrated rod-pairs. This is referred to as the "global solution." The regression parameters $\Delta\hat{\lambda}$ are treated essentially as "observations" of differences in average rod-pair scale errors and the individual rod-pair scale errors are the parameters to be estimated. The normal equations for the least squares solution can be written as

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{A})\,\hat{\lambda} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\Delta\hat{\lambda}\,, \tag{20a}$$

$$\boldsymbol{C}_{\hat{\lambda}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{A})^{-1}\,, \tag{20b}$$

where $\boldsymbol{P}$ is the (diagonal) weight matrix of the observations (regression parameters); that is, the inverse of $\boldsymbol{C}_{\Delta\hat{\lambda}} = \mathrm{diag}(\sigma^2_{\Delta\hat{\lambda}_i})$, where the variance factor is considered to be equal to one in this second adjustment, because separate variance factors have already been computed and applied in the individual regression solutions.

Although unmentioned by *Stein* [1981] in his analysis, the matrix $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{A}$ is singular and thus an infinite number of solutions exist. This means that before the system of equations in (20a) is solved, a "reference datum" for the rod scale errors must be defined. The preferred approach, and the one used by us, is to use an inner constraint solution which minimizes the sum of weighted squares of all the scale errors. This is equivalent to fixing the average of all the individual scale errors to zero. This constraint is automatically enforced by using the Moore-Penrose pseudoinverse, denoted by a superscript plus sign, to invert the normal equation matrix $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{A}$. Thus the solution $\hat{\lambda}$ and its associated covariance matrix $\boldsymbol{C}_{\hat{\lambda}}$ are

$$\hat{\lambda} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{A})^{+}\,\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\Delta\hat{\lambda}\,, \tag{21a}$$

$$\boldsymbol{C}_{\hat{\lambda}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\,\boldsymbol{A})^{+}\,. \tag{21b}$$

An alternative approach, and the one used by *Stein* [1981], is to define the reference datum by fixing one or more of the rod-pair scale errors to a value of, say, zero. The estimated scale differences are then determined only with respect to the fixed rod-pair(s). Any uncertainty in the assumed value(s) for the fixed scale(s) directly affects the other scales. This approach is clearly more subjective than the pseudoinverse solution.

Residuals from the preceding solution were also computed and tested for outliers. These residuals represent the discrepancies between the individual regression solutions for the rod-pair scale differences (the observations in the second step) and the estimates from the global solution. The residuals $\hat{\boldsymbol{r}}$ and their covariance matrix $\boldsymbol{C}_{\hat{r}}$ are computed from (cf. *Vaníček and Krakiwsky* [1986], pp. 207, 210)

$$\hat{\boldsymbol{r}} = \boldsymbol{A}\,\hat{\lambda} - \Delta\hat{\lambda}\,, \tag{22a}$$

$$\boldsymbol{C}_{\hat{r}} = \boldsymbol{C}_{\Delta\hat{\lambda}} - \boldsymbol{A}\,\boldsymbol{C}_{\hat{\lambda}}\,\boldsymbol{A}^{\mathrm{T}}\,. \tag{22b}$$

We consider as outliers any residual larger than 3.5 times its estimated standard deviation, which corresponds to an in-context test at the 95% confidence level (cf. *Vaníček and Krakiwsky* [1986], pp. 229-231). Such outliers are omitted from the second step and a new solution computed. This cycle is repeated until no outliers remain. These outliers may be due to vertical displacements, systematic effects other than rod scale errors, or departures from reality of the stated assumptions (e.g., uniform slope).

It is important to emphasize that it is not possible to separate the effects due to residual rod scale errors and those due to other effects proportional to height difference. Both will contribute in the same way to the regression coefficient in the error models. Thus care must be exercised in interpreting the results for rod-pairs of specific calibration dates where the rod-pairs have not been used in more than one leveling and in different tectonic regimes. These restrictions severely limit the usefulness of this technique, since a large number of leveling lines must be used to obtain results that can be confidently interpreted as scale errors.

Leveling Data

Two data sets, referred to here as A and B, were used in these analyses. Data set A (Table 1) consists of 25 leveling lines in southern California and is nearly identical to that used by *Stein* [1981]. All of these data were obtained from the predatabase summary sheets of the National Geodetic Survey (NGS). In addition to those used by Stein, we have also included data for several other southern California lines obtained from the summary sheet listings; namely, L14778, L18299, L19781, L20279, L21589, and L21962. On the other hand, we have rejected a few lines used by Stein, either because they were contaminated by an improperly computed rod scale correction (see Mark et al. [1981]) or because they were based on Ni 1 levels.

Data set B (Table 2) consists of 31 other leveling lines selected from NGS database that involve rod-pairs identified by *Stein* [1981], *Reilinger and Brown* [1981] and our own preliminary regression analyses as possibly contaminated by significant rod scale errors. These rod-pairs are (312) 268/274 (calibrated in 1932), (312) 301/304 (calibrated in 1933), and (315) 9/19 (calibrated in 1969).

All of the lines in data set B meet the following criteria: (1) at least 100 m of topographic relief in order to increase the probability of detecting a height-dependent rod scale error, (2) no Ni 1 levels used because of possible magnetic errors identified with these instruments [see Rumpf and Meurisch, 1981], (3) no reasonable likelihood of compaction-induced subsidence along the selected lines, since these displacements may be relatively large and tend to overwhelm any correlations attributable to rod scale errors, (4) no reasonable likelihood of major coseismic deformation between levelings for the same reasons as in (3), and (5) double-run leveling procedure used in order to minimize the presence of systematic errors such as rod and instrument settlement. The same criteria are generally applicable to data set A as well. The exceptions are lines L22391a and L23675, both of which traverse an area that includes the northern edge of the coseismic field associated with the 1971 San Fernando earthquake.

Very few releveled lines could be found which included one of the suspect rod-pairs in at least one of the levelings and still met our stipulated criteria (see Table 2). Five different levelings were found for rod-pair (312) 268/274, based on a 1932 calibration; seven levelings were found for rod-pair (312) 301/304, based on a 1933 calibration; and one leveling was found for rod-pair (315) 9/19, based on a 1969 calibration.

Note that in both data sets, letter codes (a, b, c, etc.) were added to the usual level line identifier where the leveling file had to be split into separate segments such that each was based on a single rod-pair. This gave a total of 34 leveling segments in data set A and 39 in data set B, for a total of 73 leveling segments based on single rod-pairs. From these we were able to construct 137 releveled segments in data set A and 49 in data set B, resulting in a total of 186 releveled segments. Because each separately calibrated rod-pair is treated as a different pair in our analysis, there are a combined total of 37 such “rod-pairs.”

Applied Corrections

Corrections for temperature expansion and rod scale error (or rod excess) were applied to all the leveling data. As stated above, refraction corrections were not included. The rod scale corrections for data set A were based on the pre-1975 algorithm, where the rod excess was computed by simply dividing the total cumulative excess length by the nominal length of the rod. Rod scale corrections for data set B were based on the post-1981 algorithm which used linear interpolation between all of the calibration points. The rod scale corrections in both cases have been determined from the most recent calibration of the rods prior to their use. The only exception is the 1955 calibration for rod-pair (312) 268/274 used in data set A (see Table 1), which is actually a mix. That is, the 1955 calibration for the first meter of these rods was combined with the next preceding calibration of the upper 1.8 m to produce a scale correction for this rod-pair of 0.0 ppm. It is this value that was used to correct those levelings given in Table 1, based on this rod pair. The 1932 calibrations for the same rod pair used in data set B produced a rod correction of -45.2 ppm, based on the post-1981 algorithm. It is this value that was used in generating the rod corrections for those levelings based on rod-pair (312) 268/274 in data set B (Table 2). The difference between the estimated rod-pair scale errors for these calibrations was fixed at 45.2 ppm in our estimation of individual rod scale errors.

A Priori Elimination of Data

A major criticism of the analysis by *Stein* [1981] is the large amount of leveling data subjectively omitted as outliers. As much as 20% of the entire sample was removed, often simply because the their removal improved the statistical significance of the model [*Stein*, 1981, pp. 443 and 445]. We have previously argued against such wholly subjective procedures [see *Craymer and Vaníček*, 1989]. In our opinion, one should add more data, rather than selectively and subjectively deleting existing data in order to increase the statistical confidence in the analysis.

In our analysis we have omitted only those points (benchmarks) for which differences in the section height differences between levelings exceed 5 cm, a value large enough that it could not be reasonably attributed to either rod scale or random errors. We consider such large differences to be due to blunders, benchmark disturbance, or localized crustal deformation. Among the 1480 comparisons of leveling segments examined here a total of only 53 such points were found and thus omitted.

Results

Rod-Pair Scale Differences

Solutions for differences in rod-pair scale errors were obtained for all possible comparisons of levelings in both data sets A and B. The results are summarized in Table 3 for data set A and in Table 4 for data set B. Included for each solution are the number of benchmarks (BMs) used, the standard deviation of the terrain slope (σ_β), the estimated scale differences $\Delta\hat{\lambda}$ (i.e., the difference in the residual rod-pair scale errors between the levelings), its standard deviation $\sigma_{\Delta\hat{\lambda}}$, and the test statistic $t_{\Delta\hat{\lambda}}$ for testing the statistical significance of $\Delta\hat{\lambda}$. In the absence of complete setup data, the roughness of the topography (i.e., the degree to which the basic assumption of smooth slope is invalid) can be gauged to some extent from the standard deviation of the terrain slope (smaller values indicate smoother slopes). The statistical significance of each regression is checked using the following standard test for the null hypothesis H_o: $\Delta\hat{\lambda} = 0$:

$$\text{If } t_{\Delta\hat{\lambda}} = \frac{\Delta\hat{\lambda}}{\sigma_{\Delta\hat{\lambda}}} < t_{\upsilon;1-\alpha/2}, \text{ accept } H_o, \tag{23}$$

$$\text{otherwise reject } H_o .$$

Here $t_{\upsilon;1-\alpha/2}$ is the abscissa from the Student distribution, α is the significance level for the test (5%), and υ is the degrees of freedom for the regression (number of BMs minus 1).

Inspection of these tables shows that it is difficult to identify which of the estimated scale differences are due to rod scale errors, other systematic effects, or vertical displacements. Each solution involves a combination of two rod-pairs, either or both of which may contain residual scale errors. Thus these estimates represent differences in scale errors between two rod-pairs. The compatibility of different solutions for the same rod-pairs can be tested statistically. For the null hypothesis H_o: $\Delta\hat{\lambda}_i=\Delta\hat{\lambda}_j$, where the subscripts denote individual regressions, the following standard test is used:

$$\text{If } \frac{\Delta\hat{\lambda}_i - \Delta\hat{\lambda}_j}{(\sigma^2_{\Delta\hat{\lambda}_i} + \sigma^2_{\Delta\hat{\lambda}_j})^{1/2}} < t_{\upsilon;1-\alpha/2}, \text{ accept } H_o; \tag{24}$$

$$\text{otherwise reject } H_o.$$

Here, $t_{\upsilon;1-\alpha/2}$ is as defined above, and υ is the degrees of freedom for both regressions (sum of the degrees of freedom for the two individual regressions). Significant scale differences may be revealed by statistically compatible regression estimates from different levelings for the same combination of two rod-pairs. However, there are too few of these identical combinations available. Rather than try to decipher all possible compatibility tests among the different scale difference solutions, we have used instead the more powerful tests of the residuals from the global solution for the individual scale errors (see next section). Nevertheless, a few preliminary observations and analyses of the scale difference solutions were made.

The most statistically significant scale difference (largest $t_{\Delta\hat{\lambda}}$ statistic) in Tables 3 and 4 is obtained for the comparison between levelings L16254b (rod-pair (312) 422/438, calibrated in 1951) and L18242 (rod-pair (312) 378/383, calibrated in 1956). This comparison gives a 95% confidence interval for the regression coefficient of –129 ± 8 ppm (Table 3). Unfortunately, there are no other comparisons using the same set of rod-pairs in order to properly check for location dependency. In spite of the apparently strong possibility of a rod scale error this regression is based on only four sections and was identified as an outlier in comparison with the other individual scale solutions (see next section). Two other comparisons exhibit highly significant regressions, but they were also based on only a few sections and identified as outliers in the individual scale solutions.

An example of strong evidence supporting the existence of a residual rod-pair scale error can be seen in the comparison involving the early 1964 leveling L19752n (rod-pair (312) 268/274, calibrated in 1955); see Table 3. All but 4 of the 18 solutions using this leveling give statistically significant and statistically compatible scale difference estimates. Even the statistically insignificant estimates are compatible with the others at the 95% confidence level. Comparisons involving leveling L19781, which was also based on the 1955 calibration of this rod-pair but through a different part of the Transverse Ranges, also display similarly significant estimates, albeit with much larger uncertainties due to a smaller range in section height differences. Although these results suggest the presence of a large rod scale error in this rod-pair, terrain correlated vertical displacements could produce equally significant estimates. Note, however, the change in sign of $\Delta\hat{\lambda}$ with levelings before and after levelings L19752n and L19781 based on rod-pair (312) 268/274. If attributed to crustal motion, this pattern would imply episodic uplift (or tilt) along the routes of L19752n and L19781 sometime after 1961 or earlier, followed by a sharply episodic and localized collapse or tilt reversal. Although evidence of episodic, oscillatory, and localized aseismic vertical displacements (where the period of oscillation ranges from a year to two to a decade or more) is relatively common along the plate margin traversing southern California [*Castle et al.*, 1974; *Castle et al.*, 1984; *Castle and Gilmore*, 1992], the terrain-correlated nature of these oscillatory displacements is much less common.

The validity of the assumption of uniform terrain slope, as defined by (6), can be checked using the standard deviation σ_β of the terrain slope for each solution (see Tables 3 and 4). In all of those cases associated with large σ_β values (greater than 5% or 5 m/km variation in grade), the scale difference estimates were statistically insignificant and, hence, have no impact on our results.

Individual Rod-Pair Scale Estimates

Estimates of individual rod-pair scale errors were computed simultaneously from both data sets A and B as described earlier. Because the same rod-pair (312) 268/274 was used with two different calibration corrections (1932 and 1955), we referenced the estimated residual scale error to only the 1932 calibration, since this is the one now used in the NGS database. This was done by fixing the difference between the residual scale errors for the 1955 and 1932 calibrations to be 45.2 ppm, the difference in the applied calibration corrections. The estimated residual scale error can be referenced to the 1955 calibration by simply adding 45.2 ppm to the estimate for the 1932 calibration (as was done in Table 5).

The final global solution, with four outliers omitted, is summarized in Table 5. The estimated scale errors are expressed as 95% confidence intervals. All detected outliers (see Table 6) occur along the leveling route extending northward from south of Gorman to Grapevine. All of these outliers can be attributed to a lack of data for the scale difference solutions (<10 BMs) and, perhaps, crustal deformation. Nevertheless, the solutions with and without these outliers differ insignificantly, except for a larger variance factor and, consequently, slightly fewer statistically significant scale errors.

Seven rod-pairs were found to have statistically significant estimates of scale errors that were based on more than one leveling (denoted by asterisks in Table 5). None of these, however, was larger than the expected random error in the calibration corrections; that is, 83 ppm at the 95% confidence level (see next section). Only when the scale error estimate for rod-pair (312) 268/274 is referenced to the 1955 calibration does it become larger than the expected calibration error. In this case we obtain a residual scale error estimate of 109 ± 34 ppm (95%), relative to the 1955 calibration correction. This estimate is only 64 ± 34 ppm (95%) when referenced to the 1932 calibration. These results argue for a significant error in the 1955 calibration correction.

Four other rod pairs also display statistically significant residual scale errors larger than the expected error in the calibration correction. These are rod-pairs (312) 243/244 (calibrated in 1928), (312) 308/322 (calibrated in 1951), (312) 327/360 (calibrated in 1933), and (312) 420/421 (calibrated in 1900). However, these rod-pairs were used in only one leveling. As explained earlier, we are unable to separate effects due to terrain-correlated vertical displacements and those due to rod scale errors in cases where the indicated rod-pair is used in only one leveling. The greater the number of levelings, the more reliable the estimate of the scale error. Thus the results for rod-pair errors obtained from only one leveling should be considered suspect until more data can be acquired to verify the presence of a rod scale error in different levelings through different tectonic regimes. Because the estimated scale error for rod-pair (312) 268/274 was based on seven levelings through a variety of geologic domains, it stands out as one that is especially difficult to dismiss as something other than a residual rod scale error.

Discussion of Results

In his analysis of rod scale errors, *Stein* [1981] determined that rod-pair (312) 268/274, calibrated in 1955, had an unusually large error of 120 ± 24 ppm (95%), very close to our estimate of 109 ± 34 ppm. *Stein* [1981] also found that rod-pair (316) 132180/87849 had a rod-pair specific error of –90 ± 40 ppm (95%). However, as discussed earlier, this was based on an Ni 1 leveling and an incorrect rod scale correction, and thus was excluded in our analyses.

In order to check the reasonableness of our estimated residual scale errors, we compared them against the expected accuracy for the applied scale corrections determined from rod calibrations as reported by the National Geodetic Survey (NGS). The calibration procedures of the U.S. Coast and Geodetic Survey (the predecessor to NGS), upon which the rod scale corrections were based, are divided into three basic periods: 1930-1939, 1949-1963, and 1964-1968. Analyses by NGS indicate that the random errors in these calibrations were all about 30 ppm (root mean square error) and that systematic errors relative to the 1964-1968 period ranged from 20 ppm during the period 1949-1963 to 50 ppm during the period 1930-1939 [*Strange*, 1982]. On the basis of these estimates, the random scale errors in rod-pairs would be about $(30^2+30^2)^{1/2}$ = 42.4 ppm. The 95% confidence interval for such errors would be about ±1.96 x 42.4 = ±83 ppm. Adding the systematic error gives a total error of up to about 20 ± 83 ppm (95%) for the 1949-1963 calibrations and up to 50 ± 83 ppm (95%) for the 1930-1939 calibrations, all relative to the 1964-1968 calibrations. The scale error estimate for rod-pair (312) 268/274, calibrated in 1955, is the only statistically significant value based on more than one leveling which exceeds the expected random error. It is, however, statistically compatible with the expected total error for this period.

The validity of the statistically determined residual scale error for rod-pair (312) 268/274 is seemingly supported by the fact that this result is based on seven levelings through a variety of contrasting geologic domains. However, because this rod-pair was apparently used only in the westernmost states, all of these levelings are within tectonically active areas. Thus these differences in geologic domain may be less significant than suggested by the geographic diversity of the seven surveys.

An explicit test for the presence of terrain-correlated vertical displacements in specific areas can be made by comparing estimated scale differences (regressions) obtained from levelings based on the same rod-pairs. In such comparisons the residual rod scale errors should be the same in both levelings and would not be discernible in the differencing of the heights. A significant regression would indicate either location dependent systematic effects other than rod scale error or terrain-correlated vertical displacements. Unfortunately, only two possibilities for such comparisons were found; L19755 (late 1964) versus L20650x1 (1966) over the route between Lebec and Grapevine, and L21589 (1968) versus L21962b (1969) over one of the two routes between Saugus and Palmdale. Neither of these comparisons displayed any statistically significant regression thereby indicating neither location-dependent systematic errors nor terrain-correlated displacements during the specified time intervals. Other errors due to differences in the computation of the applied rod scale correction may also exist. However, *Mark et al*. [1981] found these differences to be negligible.

In spite of the persuasive statistical evidence of a large residual scale error in rod-pair (312) 268/274, *Castle and Gilmore* [1992] present one direct and four indirect (comparative) arguments that any scale error for this rod-pair is within the expected random error of the calibration correction. One of their comparative arguments is based on an apparent breakdown in the correlation between vertical displacements and terrain along the eastern end of the Saugus-Palmdale line (3219 USGS to V811), obtained from a comparison of levelings L18299b against L19781 (Table 3). However, this argument is unsupported by our regression analysis. The difference between the regressions for the two line segments is statistically insignificant because of the large uncertainties in the regressions (Table 7). Thus, no meaningful conclusion can be made about any difference between the regressions.

Although a level of uncertainty is associated with each of the *Castle and Gilmore* [1992] arguments, collectively, they cast doubt on the interpretation of our estimated solution ($\hat{\lambda}$) for rod-pair (312) 268/274 as residual rod scale error. The most compelling of these arguments is the direct comparison with the 1965 calibration for this rod-pair (National Bureau of Standards test G-35760), which leads to a rod-pair scale error of 17.6 ppm. Because the 1965 and later calibrations were thought to be more accurate than the pre-1965 calibrations [*Mark et al*., 1981; *Strange*, 1982], we should have expected a larger error that more closely agrees with our estimate of 109 ± 34 ppm (95%) relative to the 0.0 ppm error obtained from the 1955 calibration. Relative to the 1965 calibration, our estimate amounts to 91 ± 34 ppm, which is still greater than even the 95% confidence interval for the expected random error in the pre-1965 calibrations (i.e., 83 ppm).

One possible explanation for the large discrepancy between our estimated error and that based on the 1965 calibration may be due to the presence of some unidentified systematic error in the calibration. However, this seems unlikely owing to the presumably more accurate calibration procedures adopted in 1965. Moreover, the 1965 calibration agrees to within 30 ppm (well within the expected random error) of a subsequent calibration of the same rod-pair only 1 year later; the difference between the 1966 calibration and our estimate is 119 ppm. A second possible explanation may be related to a difference in the tension applied to the rod scales (invar strips) while in use and that during calibration. However, beginning in 1965 the invar strips were calibrated while still in the rod frame, under the design load of 11 kg [*Strange*, 1982]. Nevertheless, even a worst-case change in load (0 to 11 kg) would decrease our estimate of the scale error by no more than about 32 ppm [*Mark et al*., 1981]. Accordingly, it is unlikely that any change in load on the invar strips could account for the difference between our statistical estimate of the residual scale error and the 1965 calibration.

It has also been suggested that the invar strips for this rod-pair could have been replaced sometime after the 1964 levelings, but prior to the 1965 calibration (R. Stein, personal communication, December 1994). Thus the 1965 calibration could refer to new invar strips in the old rod frames. Unfortunately, no record of any changes in the invar strips was ever kept. Although it is possible that the strips could have been replaced following the 1964 levelings, this seems unlikely. First, the invar strips calibrated in 1965 had the same unusual 3.0 m (rather than the standard 3.2 m) scribe point as did the strips in the earlier complete calibrations of this rod-pair. It is doubtful that similar nonstandard strips would have been used as replacements. Second, the 1965 calibration is particularly unusual in that three separate calibrations were performed for this and one other rod-pair: as received, with tension released, and with tension reapplied (see National Bureau of Standards test G-35760). If the invar scales had been replaced, there would have been little reason for such an elaborate series of calibrations. On the other hand, this would have been a logical procedure if a large error was suspected in the earlier calibrations, for which a basis existed (see, e.g., *Castle and Gilmore* [1992], Figure 5). Third, rods of this type were being rapidly phased out for first-order leveling at least as early as 1964. This particular rod-pair was never again used for any first-order work following the 1964 levelings used in our analysis. Its subsequent use was confined to three 1970 second-order levelings in Florida, where the effect of any scale error would have been trivial owing to the lack of any significant relief.

The arguments presented by *Castle and Gilmore* [1992], however uncertain, clearly indicate that the regression estimates should not be attributed entirely to rod scale errors without first considering crustal deformation as a possible contributor, even if this deformation seems to be of an unusual nature. This is the main limitation of this technique and could restrict its application to areas of modest tectonic activity. Finally, the disagreement between a recent, relatively more accurate calibration and our estimate of the significant residual scale error for rod-pair (312) 268/274 indicates that estimates of errors in historic leveling deserve further attention and research.

Conclusions

We have shown that the assumption that rod scale errors are correlated with height, height difference, or terrain slope is valid only if the slope is constant over at least pairs of adjacent setups. Deviations from this requirement will reduce the validity of the estimated rod scale errors. Although not rigorously investigated in our analyses, the degree to which this assumption is satisfied can be assessed using the variation in terrain slope; large variations in slope will invalidate the assumption. We recommend that the slope variations be taken into consideration in developing an improved weighting scheme for the estimation of the individual rod-pair scale errors.

We have also shown that rod scale errors can be modeled in terms of heights, height differences, or terrain slope only when correct relative weighting, and all mathematical correlations are taken into account. Under these conditions, all three models give identical results, although the one based on height differences is easier to implement because it requires the least manipulation of the original data.

The results of our analysis of seven different levelings suggest the presence of a rod-pair dependent systematic error of about 109 ± 34 ppm (95% confidence interval) in rod-pair (312) 268/274, relative to the 1955 calibration. This value agrees closely with that given by *Stein* [1981]. Although this estimate may be reasonably attributed to residual rod scale errors in the 1955 calibration, disagreement with the 1965 calibration of this rod-pair, together with other independent evidence, leads us to treat this conclusion with caution. More research into this disagreement is needed. Statistically significant height-dependent rod-pair errors were also found for four other rod-pairs. However, these rod-pairs were used in only one or two levelings and the resulting estimates may be no more than artifacts attributable to other systematic effects or crustal deformation. Thus it is not always possible to distinguish between scale error and other systematic effects dependent on height, height difference, or slope. This limitation is diminished where the same rod-pairs are used in more than one leveling, since location-dependent systematic errors will generally appear as outliers when compared to estimates without such errors.

A number of improvements could be made to future applications of the technique which would increase the reliability of the results. Some of these are: (1) include more levelings for each rod-pair from different (tectonically inactive) areas, (2) investigate the effect of including the modeled refraction corrections of the National Geodetic Survey and the possibility of enhancing our error models by including parameters for other possible systematic errors, and (3) assess the estimation technique more objectively by performing a complete simulation of the leveling process, including rod scale errors and other typical systematic errors. This would enable one to fully control the errors and to judge the success of the estimation technique. These investigations should be pursued before considering this technique as a reliable tool for the actual correction of rod scale errors in historic leveling.

Acknowledgments. This study was funded by the U.S. Geological Survey under contract 133636-90. Financial assistance was also provided by the Canadian National Science and Engineering Research Council. We gratefully acknowledge the help provided by Ross Stein, who supplied us with much of the data used in our data set A, and Thomas Gilmore, who helped us identify, acquire, and assemble (often manually) the rest of the data used in this analysis.. Emery Balazs and Kathy Koepsell also provided much assistance in the acquisition of the data. Their generous help and fielding of many questions is very much appreciated. Finally, we also thank Art Sylvester and another (anonymous) reviewer for their helpful suggestions in improving the original manuscript.

References

Bomford, G., *Geodesy* (3rd ed.), Oxford University Press, New York, 1971.

Castle, R.O., and T.D. Gilmore, A revised configuration of the southern California uplift, *Geol.Soci. Am. Bull.*, *104*, 1577-1591, 1992.

Castle, R.O., J.N. Alt, J.C. Savage, and E.I. Balazs, Elevation changes preceding the San Fernando earthquake of February 9, 1971, *Geology*, *2*, 61-66, 1974.

Castle, R.O., J.P. Church, and M.R. Elliott, Aseismic uplift in Southern California, *Science*, *192*, 251-253, 1976.

Castle, R.O., M.R. Elliott, J.P. Church, and S.H. Wood, The evolution of the southern California uplift, 1955 through 1976, *U.S. Geol. Surv. Prof. Pap.*, *1342*, 1984.

Castle, R.O., T.D. Gilmore, R.K. Mark, and R.H. Shaw, Empirical estimates of cumulative refraction errors associated with procedurally constrained levelings based on the Gaithersburg–Tucson refraction tests of the National Geodetic Survey, *Geophys. Res. Lett.*, *12*(5), 239-242, 1985.

Castle, R.O., R.K. Mark, and R.H. Shaw. An empirical assessment of refraction error in leveling as a function of survey order and environment. *U.S. Geol. Surv. Bull.*, *2114*, 1994.

Craymer, M.R., and P. Vaníc̆ek, Comment on "Saugus-Palmdale, California, field test for refraction error in historical leveling surveys" by R.S. Stein, C.T. Whalen, S.R. Holdahl, W.E. Strange, and W. Thatcher, and reply to "Comment on 'Further analysis of the 1981 southern California field test for leveling refraction' by M.R. Craymer, and P. Vaníc̆ek" by R.S. Stein, C.T. Whalen, S.R. Holdahl, W.E. Strange, and W. Thatcher", *J.Geophys. Res.*, *94*(B6), 7667-7672, 1989.

Holdahl, S.R, A model of temperature stratification for correction of leveling refraction, *Bull. Geod.*, *55*, 231-249, 1981.

Jackson, D.D., and W.B. Lee, The Palmdale Bulge – An alternate interpretation (abstract), *Eos Trans.*, *AGU*, *60*, 810, 1979.

Jackson, D.D., W.B. Lee, and C.-C. Liu, Aseismic uplift in southern California: An alternate interpretation, *Science*, *210*, 534-536, 1980.

Jackson, D.D., W.B. Lee, and C.-C. Liu, Height dependent errors in southern California leveling, in *Maurice Ewing Series for Earthquake Prediction and International Review*, edited by D.W. Simpson and P.G. Richards, AGU, Washington, D. C., 422-440, 1981.

Kukkamäki, T.J., *Über die nivellitische refraktion*, publication 25, Finnish Geodetic Institute, Helsinki, 1938.

Mark, R.K., J.C. Tinsley III, E.B. Newman, T.D. Gilmore, and R.O. Castle, An assessment of the accuracy of geodetic measurements that define the southern California uplift, *J. Geophys. Res.*, *86*(B4), 2783-2808, 1981.

Mark, R.K., T.D. Gilmore, and R.O. Castle, Evidence of the suppression of the unequal refraction error in geodetic leveling, *J. Geophys. Res.*, *92*(B3), 2767-2790, 1987.

Reilinger, R.E., and L.D. Brown, Neotectonic deformation, near-surface movements and systematic errors in U.S. releveling measurements: Implications for earthquake prediction, in *Maurice Ewing Series for Earthquake Prediction and International Review*, edited by D.W. Simpson and P.G. Richards, AGU, Washington, D. C., 422-440, 1981.

Rumpf, W.E., and H. Meurisch, Systematische anderungen der ziellinie eines prozisions kompensator niverlliers insbesondere des wechselfelder, paper presented at the XVI International Congress, Int. Fed. Surv., Montreux, Switzerland, 1981.

Stein, R.S., Discrimination of tectonic displacement from slope-dependent errors in geodetic leveling from southern California, 1953-1979, in *Maurice Ewing Series for Earthquake Prediction and International Review*, edited by D.W. Simpson and P.G. Richards, AGU, Washington, D. C., 441-456, 1981.

Stein, R.S., C.T. Whalen, S.R. Holdahl, W.E. Strange, and W. Thatcher, Saugus-Palmdale, California field test for refraction error in historic leveling surveys, *J. Geophys. Res.*, *91*(B1), 9031-9044, 1986.

Strange, W.E., The effect of systematic errors on geodynamic analysis, paper presented at Second International Symposium on Problems

Related to the Redefinition of North American Vertical Geodetic Networks, Can. Inst. Surv., Ottawa, Ont., 705-728, May 26-30, 1980.

Strange, W.E., The impact of refraction correction on leveling interpretations in southern California, *J. Geophys. Res.*, *86*(B4), 2809-2824, 1981.

Strange, W.E., Scribing, graduation and calibration of U.S. Coast and Geodetic Survey leveling rods from 1877 to 1968, *NOAA Tech. Rep. NOS 94 NGS 23*, Nat. Ocean Surv., Nat. Oceanic Atmos. Admin., Rockville, Md., 1982.

Vaníček, P., and E.J. Krakiwsky, *Geodesy: The Concepts* (2nd ed.), North-Holland, New York, 1986.

Vaníček, P., R.O. Castle, and E.I. Balazs, Geodetic leveling and its applications, *Rev. Geophys. Space Phys.*, *18*, 505-524, 1980.

Whalen, C.T., Refraction errors in leveling – NGS test results, paper presented at Second International Symposium on Problems Related to the Redefinition of North American Vertical Geodetic Networks, Can. Inst. Surv., Ottawa, Ont., 757-782, May 26-30, 1980.

Robert O. Castle, U.S. Geological Survey, MS 975, 345 Middlefield Road, Menlo Park, CA 94025. (e-mail: rcastle@isdmnl.wr.usgs.gov)

Michael R. Craymer, Geodetic Survey Division, Geomatics Canada, 615 Booth Street, Ottawa, ON K1A 0E9, Canada. (e-mail: craymer@emr.ca)

Petr Vaníček, Department of Geodesy and Geomatics Engineering, University of New Brunswick, P.O. Box 4400, Fredericton, NB E3B 5A3, Canada. (e-mail: vanicek@unb.ca)

(Received May 26, 1994; revised February 7, 1995; accepted February 16, 1995.)

Paper number 95JB00614.
0148-0227/95/95JB-00614$05.00

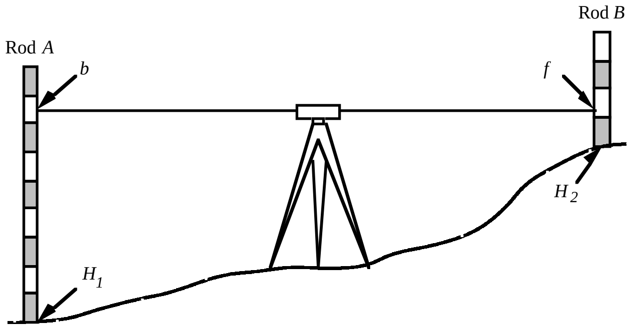

Figure 1. Typical leveling setup showing the foresight *f* and backsight *b* rod readings. The height difference for the setup is $\Delta h = H_2 - H_1 = b - f$.

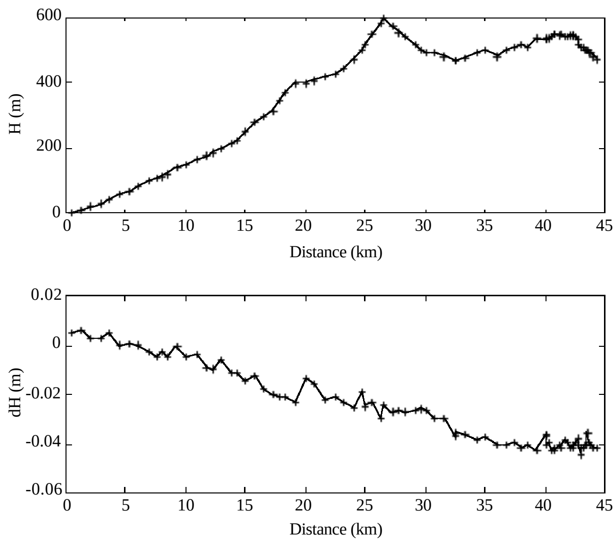

Figure 2. Profiles of topography (H) and changes in observed heights (dH) between levelings L21589 and L21962b from Saugus to Palmdale along the Mint Canyon (northern route). Note the apparent strong correlation between dH and H indicated by these plots.

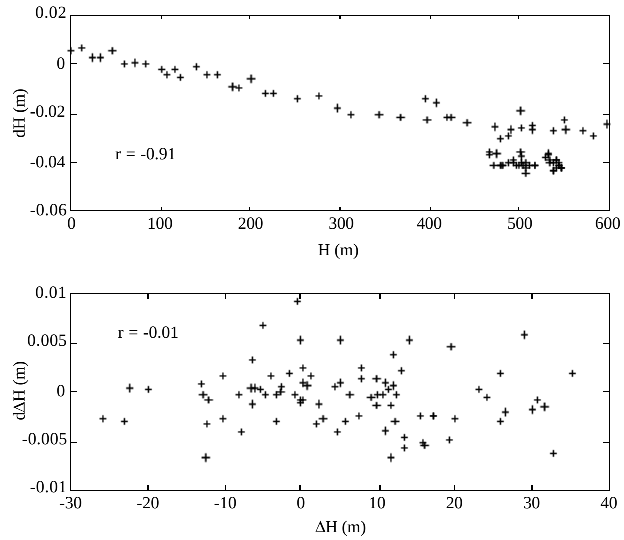

Figure 3. Plots of apparent correlation between changes in height (dH) and height (H), and between changes in section height difference ($d\Delta H$) and section height difference (ΔH) for the levelings in Figure 2. The top plot shows an apparent strong correlation (r=–0.91) between dH and H which falsely indicates the presence of possible rod scale error. The bottom plot correctly shows, however, that no correlation is actually present (r=–0.01). The significant correlation coefficient between dH and H is caused by the large mathematical correlations among the dH data, which are derived from the uncorrelated ΔH data (see text). The mathematical correlations artificially amplify greatly even the slightest insignificant relation between dH and H. Thus plots of dH and H should never be used to gauge the presence of rod scale error in leveling. Plots of $d\Delta H$ and ΔH, or $d\beta$ and β are more reliable indicators.

Table 1. Leveling Segments and Supplemental Information Used in Data Set A

Line	Date	Rod Pair	Calibration	Location, Southern California
L14778	1953	(312) 398/409	1944	Bakersfield to Mojave
L14799n	1953	(312) 398/409	1944	Mettler to Lebec
L14799s	1953	(312) 257/289	1952	Lebec to Saugus
L16241x6a	1957	(312) 302/348	1945	Wheeler Ridge
L16241x6b	1957	(312) 422/438	1951	Wheeler Ridge
L16254a	1956-1957	(312) 302/348	1945	Lebec to Wheeler Ridge
L16254b	1956-1957	(312) 422/438	1951	Lebec to Wheeler Ridge
L17206x1	1959	(312) 251/310	1953	Wheeler Ridge
L18242	early 1961	(312) 378/383	1956	Gorman-Lebec area
L18299b	1961	(312) 308/322	1951	Saugus to Palmdale
L18529	late 1961	(312) 391/459	1945	Grapevine area
L18529x10	1961-1962	(312) 391/459	1945	Grapevine area
L19752n	early 1964	(312) 268/274	1955	Reservoir to Grapevine
L19752s	early 1964	(312) 248/254	1945	Saugus to Reservoir
L19755	late 1964	(317) 0163/0263	1964	Lebec to Reservoir
L19781	1964	(312) 268/274	1955	Saugus to Palmdale
L20130x10	1965	(316) 87815/87859	1965	Wheeler Ridge
L20145	1965	(317) 0163/0263	1964	Saugus to Lang
L20169n	1965	(316) 87815/87859	1965	Sandberg to Lebec
L20169s	1965	(317) 0163/0263	1964	Saugus to Sandberg
L20279	1965	(316) 87815/87859	1965	Caliente to Mojave
L20298	1965	(317) 0163/0263	1964	Lang to Palmdale
L20650x1	1966	(317) 0163/0263	1964	Lebec to Grapevine
L21366	1968	(316) 119358/119362	1966	Lebec area
L21589	1968	(315) 95/96	1969	Saugus to Palmdale
L21962	1969	(315) 9/19	1969	Sandberg to Palmdale
L21962a	1969	(315) 9/19	1969	Saugus area
L21962b	1969	(315) 95/96	1969	Saugus to Palmdale
L22024x32	1970	(316) 124734/124735	1967	Wheeler Ridge
L22391a	1971	(316) 124734/124735	1967	Castaic area
L22391b	1971	(316) 87849/121178	1966	Castaic-Sandberg area
L22391c	1971	(316) 124734/124735	1967	Sandberg area
L22391d	1971	(316) 121178/87849	1966	Sandberg to Grapevine
L23673	1973-1974	(315) 2139A/2139B	1973	Reservoir to Lebec
L23675	1973-1974	(315) 2139A/2139B	1973	Saugus to Reservoir

Calibration denotes the date of the last calibration of the rods prior to the date of the leveling.

Table 2. Leveling Segments and Supplemental Information Used in Data Set B

Line	Date	Rod Pair	Calibration	Location
L12622x3	1948	(312) 301/304	1933	Wash., Kittitas to Vantage
L12638	1948	(312) 420/421	1900	S. Calif., Gaviota area and north
L13460	1950	(312) 244/254	1945	Nevada-Arizona, Boulder Dam
L13514x1	1950	(312) 244/254	1945	Nevada, Lake Mead area
L15908	1956	(312) 302/348	1945	S. Calif., Redlands to Big Bear
L16455	1957	(312) 288/353	1945	Wash., Kittatas to Vantage
L17206x5	1959	(312) 301/304	1933	S. Calif., Shiedeck area
L17789	1960	(312) 268/274	1966	Arizona, Tuscon area
L17856	1960	(312) 301/304	1933	S. Calif., Gaviota to Los Olivos
L18245	1961	(312) 301/304	1933	S. Calif., Wheeler ridge to Lebec
L18535	1961	(312) 301/304	1933	S. Calif., Azusa to Mid San Gabriels
L18544	1961	(312) 301/304	1933	S. Calif., Redlands to Victorville
L18721	1962	(312) 301/304	1933	Montana, Wickes to Cascade Mtn.
L19231	1963	(312) 268/274	1932	Nevada-Arizona, Boulder Dam
L19234	1963	(312) 268/274	1932	Nevada, Lake Mead area
L19377	1963	(312) 268/274	1932	Utah, Loa area
L19598x1	1963-1964	(312) 268/274	1932	Central Calif, Fairfield-Vacaville area
L1960	1934	(312) 327/360	1933	Utah, Loa area
L21067x1a	1966-1967	(312) 325/348	1965	Central Calif., Fairfield area
L21067x1b	1966-1967	(312) 411/421	1965	Central Calif., Vacaville area
L21366a	1968	(316) 119358/119362	1966	S. Calif., Shiedeck area
L21679	1968	(315) 9/19	1969	S. Calif., Topanga Canyon area
L22509	1971	(316) 121178/87849	1966	S. Calif., Wheeler ridge to Lebec
L22515	1971	(316) 124734/124735	1967	S. Calif., Azusa to Mid San Gabriels
L22560	1971	(315) 95/96	1970	S. Calif., Topanga Canyon area
L24301x15	1978	(315) 83/84	1978	S. Calif., Azusa to Mid San Gabriels
L24555x5	1979	(316) 270713/277930	1979	Arizona, Tuscon area
L24687	1982	(316) 270719/277924	1980	Nevada-Arizona, Boulder Dam
L6312	1935	(312) 320/321	1933	Arizona, Lake Mead area
L6314b	1935	(312) 391/405	1934	Nevada-Arizona, Boulder Dam area
L6314c	1935	(312) 344/345	1933	Nevada-Arizona, Boulder Dam area
L6314d	1935	(312) 391/405	1934	Nevada-Arizona, Boulder Dam area
L6314e	1935	(312) 344/345	1933	Nevada-Arizona, Boulder Dam area
L6314f	1935	(312) 320/321	1933	Nevada-Arizona, Boulder Dam area
L6314g	1935	(312) 336/337	1933	Nevada-Arizona, Boulder Dam area
L7684x1	1936	(312) 340/414	1934	Montana, Wickes to Cascade Mtn.
L7684x4	1936	(312) 340/414	1934	Montana, Wickes to Cascade Mtn.
L9079x1	1941	(312) 393/394	1934	Nevada, Lake Mead area
L9079x2	1941	(312) 243/244	1928	Arizona, Lake Mead area

Calibration denotes the date of the last calibration of the rods prior to the date of the leveling; Wash., Washington; S. Calif., southern California; Mtn., mountain.

Table 3. Regression Solutions for Data Set A

1st Leveling	2nd Leveling	Number of Benchmarks	σ_β, m/km	$\Delta\hat{\lambda} \pm \sigma_{\Delta\hat{\lambda}}$, ppm	$t_{\Delta\hat{\lambda}}$
L14778	L20279	47	17	31 ± 71	0.4
L14799n	L16254a	7	13	-35 ± 43	0.8
L14799n	L16254b	10	22	26 ± 28	0.9
L14799n	L18242	12	19	-40 ± 47	0.9
L14799n	L18529	9	16	154 ± 51	3.0 *
L14799n	L19752n	15	22	129 ± 42	3.1 *
L14799n	L19755	8	21	16 ± 68	0.2
L14799n	L20169n	8	21	35 ± 59	0.6
L14799n	L20650x1	7	22	13 ± 56	0.2
L14799n	L21366	8	20	20 ± 57	0.3
L14799n	L21589	4	17	289 ± 205	1.4
L14799n	L22391d	11	22	13 ± 53	0.2
L14799n	L23673	8	16	18 ± 65	0.3
L14799s	L18242	27	33	-110 ± 82	1.4
L14799s	L19752n	43	36	111 ± 46	2.4 *
L14799s	L19752s	25	27	24 ± 121	0.2
L14799s	L20169n	28	31	2 ± 68	0.0
L14799s	L20169s	37	30	49 ± 60	0.8
L14799s	L21589	56	31	57 ± 52	1.1
L14799s	L21962	34	33	11 ± 68	0.2
L14799s	L22391b	12	30	107 ± 100	1.1
L14799s	L22391c	11	34	-19 ± 87	0.2
L14799s	L22391d	20	28	15 ± 188	0.1
L14799s	L23673	16	24	13 ± 103	0.1
L14799s	L23675	26	36	55 ± 69	0.8
L16241x6a	L17206x1	7	30	10 ± 25	0.4
L16241x6a	L18529x10	7	30	-25 ± 59	0.4
L16241x6a	L20130x10	7	30	-50 ± 50	1.0
L16241x6a	L22024x32	4	33	-42 ± 49	0.9
L16241x6b	L17206x1	7	55	45 ± 89	0.5
L16241x6b	L18529x10	8	63	16 ± 72	0.2
L16241x6b	L20130x10	8	63	43 ± 99	0.4
L16241x6b	L22024x32	8	63	69 ± 82	0.8
L16254a	L18242	8	15	96 ± 108	0.9
L16254a	L19752n	7	12	126 ± 63	2.0
L16254a	L19755	4	12	-82 ± 159	0.5
L16254a	L20169n	5	16	150 ± 43	3.5 *
L16254a	L20650x1	4	12	-98 ± 81	1.2
L16254a	L21366	5	16	-59 ± 99	0.6
L16254a	L22391d	7	13	-66 ± 53	1.2
L16254a	L23673	5	10	-61 ± 85	0.7
L16254b	L18242	4	29	-129 ± 8	15.3 *
L16254b	L18529	9	38	19 ± 44	0.4
L16254b	L19752n	9	36	170 ± 40	4.3 *
L16254b	L19755	9	36	64 ± 46	1.4
L16254b	L20650x1	6	26	10 ± 45	0.2
L16254b	L21366	7	23	-4 ± 53	0.1
L16254b	L22391d	7	24	6 ± 63	0.1
L17206x1	L18529x10	16	69	-30 ± 20	1.5
L17206x1	L20130x10	16	69	-19 ± 33	0.6
L17206x1	L22024x32	13	76	4 ± 42	0.1
L18242	L19752n	40	36	119 ± 18	6.5 *
L18242	L19755	9	22	16 ± 58	0.3
L18242	L20169n	31	31	1 ± 39	0.0
L18242	L20650x1	9	22	6 ± 69	0.1
L18242	L21366	10	21	-4 ± 83	0.1
L18242	L21589	24	28	56 ± 55	1.0
L18242	L22391d	31	35	-8 ± 112	0.1
L18242	L23673	25	33	-1 ± 69	0.0

Table 3. (continued)

1st Leveling	2nd Leveling	Number of Benchmarks	σ_β, m/km	$\Delta\hat{\lambda} \pm \sigma_{\Delta\hat{\lambda}}$, ppm	$t_{\Delta\hat{\lambda}}$
L18299b	L19781	229	21	182 ± 88	2.1 *
L18299b	L20145	11	5	161 ± 122	1.3
L18299b	L20298	204	20	82 ± 91	0.9
L18299b	L21589	23	30	80 ± 298	0.3
L18299b	L21962b	24	31	276 ± 503	0.6
L18529	L19752n	11	33	161 ± 24	6.9 *
L18529	L19755	11	33	61 ± 35	1.8
L18529	L20650x1	6	27	58 ± 5	12.3 *
L18529	L21366	7	23	62 ± 48	1.3
L18529	L22391d	7	24	78 ± 56	1.4
L18529x10	L20130x10	17	67	9 ± 28	0.3
L18529x10	L22024x32	14	74	45 ± 33	1.4
L19752n	L19755	20	28	-106 ± 13	8.1 *
L19752n	L20169n	50	29	-105 ± 32	3.3 *
L19752n	L20169s	17	36	-77 ± 25	3.1 *
L19752n	L20650x1	15	21	-95 ± 17	5.7 *
L19752n	L21366	18	20	-106 ± 29	3.7 *
L19752n	L21589	54	32	-73 ± 97	0.8
L19752n	L21962	19	40	-147 ± 27	5.5 *
L19752n	L22391b	6	29	-33 ± 30	1.1
L19752n	L22391c	12	35	-78 ± 34	2.3 *
L19752n	L22391d	48	34	-97 ± 70	1.4
L19752n	L23673	29	33	-102 ± 49	2.1
L19752n	L23675	21	41	-26 ± 115	0.2
L19752s	L20169s	33	27	34 ± 32	1.1
L19752s	L21589	26	27	15 ± 99	0.2
L19752s	L21962	24	27	-45 ± 55	0.8
L19752s	L22391a	6	20	38 ± 113	0.3
L19752s	L22391b	12	32	-176 ± 77	2.3 *
L19752s	L23675	18	27	-47 ± 78	0.6
L19755	L20650x1	15	21	19 ± 21	0.9
L19755	L21366	16	21	12 ± 29	0.4
L19755	L22391d	16	22	26 ± 37	0.7
L19755	L23673	8	22	-3 ± 104	0.0
L19781	L20145	20	5	50 ± 123	0.4
L19781	L20298	204	20	-103 ± 46	2.3 *
L19781	L21589	25	30	-81 ± 201	0.4
L19781	L21962b	24	31	159 ± 511	0.3
L20130x10	L22024x32	14	74	34 ± 32	1.1
L20145	L21589	5	4	-20 ± 246	0.1
L20169n	L21366	4	35	-114 ± 10	11.4 *
L20169n	L21589	41	28	16 ± 19	0.9
L20169n	L21962	6	26	-112 ± 73	1.5
L20169n	L22391d	39	30	-34 ± 75	0.5
L20169n	L23673	23	27	9 ± 38	0.2
L20169n	L23675	6	26	-5 ± 27	0.2
L20169s	L21589	56	33	-3 ± 58	0.1
L20169s	L21962	55	33	-45 ± 60	0.7
L20169s	L22391a	6	20	-16 ± 119	0.1
L20169s	L22391b	21	33	-106 ± 71	1.5
L20169s	L22391c	21	39	8 ± 64	0.1
L20169s	L23675	44	34	-29 ± 50	0.6
L20298	L21589	28	29	20 ± 195	0.1
L20298	L21962b	32	52	83 ± 346	0.2
L20650x1	L21366	15	20	-4 ± 25	0.2
L20650x1	L22391d	15	21	0 ± 26	0.0
L20650x1	L23673	8	22	-2 ± 84	0.0
L21366	L22391d	24	26	-17 ± 24	0.7
L21366	L23673	10	19	7 ± 45	0.2

Table 3. (continued)

1st Leveling	2nd Leveling	Number of Benchmarks	σ_β, m/km	$\Delta\hat{\lambda} \pm \sigma_{\Delta\hat{\lambda}}$, ppm	$t_{\Delta\hat{\lambda}}$
L21589	L21962	96	37	-46 ± 28	1.6
L21589	L21962a	4	4	-1065 ± 399	2.7
L21589	L21962b	86	29	-6 ± 33	0.2
L21589	L22391a	16	18	95 ± 140	0.7
L21589	L22391b	31	29	-12 ± 32	0.4
L21589	L22391c	25	38	13 ± 35	0.4
L21589	L22391d	53	34	10 ± 63	0.2
L21589	L23673	33	33	71 ± 42	1.7
L21589	L23675	79	39	6 ± 31	0.2
L21962	L22391a	16	18	156 ± 70	2.2 *
L21962	L22391b	31	29	58 ± 47	1.2
L21962	L22391c	25	38	90 ± 31	2.9 *
L21962	L22391d	9	20	-5 ± 75	0.1
L21962	L23675	77	39	64 ± 32	2.0
L22391a	L23675	17	18	-10 ± 70	0.1
L22391b	L23675	33	34	2 ± 30	0.1
L22391c	L23675	23	37	-4 ± 23	0.2
L22391d	L23673	39	33	37 ± 22	1.7
L22391d	L23675	9	20	34 ± 34	1.0

* denotes statistically significant regressions at the 95% confidence level.

Table 4. Regression Solutions for Data Set B

1st Leveling	2nd Leveling	Number of Benchmarks	σ_β, m/km	$\Delta\hat{\lambda} \pm \sigma_{\Delta\hat{\lambda}}$, ppm	$t_{\Delta\hat{\lambda}}$
L12622x3	L16455	10	15	-7 ± 33	0.2
L12638	L17856	31	31	-176 ± 60	3.0 *
L13460	L19231	115	39	139 ± 82	1.7
L13460	L24687	73	31	40 ± 64	0.6
L13460	L6314b	10	36	79 ± 74	1.1
L13460	L6314c	5	17	32 ± 55	0.6
L13460	L6314d	9	30	-83 ± 142	0.6
L13460	L6314e	28	38	31 ± 172	0.2
L13460	L6314f	13	29	25 ± 25	1.0
L13460	L6314g	26	30	-10 ± 360	0.0
L13514x1	L19234	41	33	121 ± 80	1.5
L13514x1	L6312	36	34	0 ± 44	0.0
L13514x1	L9079x1	38	34	9 ± 39	0.2
L15908	L18544	78	36	-29 ± 29	1.0
L17206x5	L21366a	29	31	-10 ± 52	0.2
L17789	L24555x5	39	23	-394 ± 758	0.5
L18245	L21366a	6	45	54 ± 46	1.2
L18245	L22509	34	42	70 ± 21	3.3 *
L18535	L22515	27	27	66 ± 52	1.3
L18535	L24301x15	25	26	11 ± 56	0.2
L18721	L7684x1	52	11	-38 ± 42	0.9
L18721	L7684x4	9	10	135 ± 252	0.5
L19231	L24687	75	32	-97 ± 26	3.8 *
L19231	L6312	32	9	-104 ± 35	3.0 *
L19231	L6314b	9	37	-134 ± 70	1.9
L19231	L6314c	5	17	-53 ± 98	0.5
L19231	L6314d	8	30	-199 ± 233	0.9
L19231	L6314e	28	39	-72 ± 255	0.3
L19231	L6314f	13	29	-90 ± 30	3.0 *
L19231	L6314g	25	27	-113 ± 484	0.2
L19231	L9079x2	38	8	69 ± 51	1.4
L19234	L6312	33	35	-129 ± 33	3.9 *
L19234	L9079x1	35	34	-110 ± 30	3.7 *
L19377	L1960	7	13	-191 ± 33	5.8 *
L19598x1	L21067x1a	13	18	251 ± 367	0.7
L19598x1	L21067x1b	8	3	39 ± 235	0.2
L21679	L22560	46	38	-1 ± 168	0.0
L22515	L24301x15	107	47	-3 ± 10	0.3
L24687	L6314b	6	20	-50 ± 24	2.1
L24687	L6314c	5	16	107 ± 217	0.5
L24687	L6314d	8	30	-196 ± 360	0.5
L24687	L6314e	24	38	105 ± 276	0.4
L24687	L6314f	8	27	-1 ± 42	0.0
L24687	L6314g	7	12	157 ± 84	1.9
L6312	L9079x1	40	35	13 ± 19	0.7
L6312	L9079x2	33	9	174 ± 37	4.8 *

* denotes statistically significant regressions at the 95% confidence level.

Table 5. Global Solution for Individual Rod-Pair Scale Errors with Outliers Omitted

Rod Pair	Calibration Date	Number of Levelings	95% c.i. for $\hat{\lambda}$, ppm	
(317) 0163+0263	1964	5	16 ± 34	
(316) 119358+119362	1966	2	11 ± 37	
(316) 121178+87849	1966	3	8 ± 35	
(316) 124734+124735	1967	4	23 ± 36	
(315) 2139A+2139B	1973	2	20 ± 36	
(312) 243+244	1928	1	130 ± 54	
(312) 244+254	1945	2	-64 ± 42	*
(312) 248+254	1945	1	12 ± 46	
(312) 251+310	1953	1	6 ± 39	
(312) 257+289	1952	1	-16 ± 44	
(312) 268+274	1932	7	64 ± 34	*
(312) 268+274	1955	7	109 ± 34	**
(316) 270713+277930	1979	1	-330 ± 1032	
(316) 270719+277924	1980	1	-32 ± 42	
(312) 288+353	1945	1	-45 ± 59	
(312) 301+304	1933	7	-38 ± 39	
(312) 302+348	1945	3	9 ± 39	
(312) 308+322	1951	1	-87 ± 82	
(312) 320+321	1933	2	-46 ± 38	*
(312) 325+348	1965	1	315 ± 500	
(312) 327+360	1933	1	-127 ± 56	
(312) 336+337	1933	1	110 ± 118	
(312) 340+414	1934	2	-71 ± 68	*
(312) 344+345	1933	2	-15 ± 71	
(312) 378+383	1956	1	-13 ± 38	
(312) 391+405	1934	2	-75 ± 49	*
(312) 391+459	1945	2	-40 ± 35	*
(312) 393+394	1934	1	-39 ± 41	
(312) 398+409	1944	2	-13 ± 39	
(312) 411+421	1965	1	102 ± 321	
(312) 420+421	1900	1	138 ± 90	
(312) 422+438	1951	2	-21 ± 40	
(315) 83+84	1978	1	19 ± 39	
(316) 87815+87859	1965	3	-2 ± 36	
(315) 9+19	1969	3	-47 ± 38	*
(315) 95+96	1969	2	12 ± 37	
(315) 95+96	1970	1	-48 ± 232	

c.i. denotes confidence interval;

* denotes statistically significant estimates for rod-pairs used in more than one leveling;

** denotes statistically significant estimates larger than the expected random error in the rod calibration.

Table 6. Regression Solution Residual Outliers (Test Statistic $t_{\hat{r}}$ Larger Than 3.5) in the Global Solution for Individual Rod-Pair Scale Errors

1st Leveling	2nd Leveling	Number of Benchmarks	$\hat{r} \pm \sigma_{\hat{r}}$, ppm	$t_{\hat{r}}$
L14799n	L18529	9	–164 ± 39	4.2
L16254a	L20169n	5	–139 ± 32	4.3
L16254b	L18242	4	21 ± 3	6.6
L20169n	L21366	4	39 ± 6	6.2

Table 7. Location Dependency of Regression Solutions for Comparison of Levelings L18299b and L19781

From	To	Number of Benchmarks	$\Delta\hat{\lambda} \pm \sigma_{\Delta\hat{\lambda}}$, ppm	$t_{\Delta\hat{\lambda}}$
Y486	V811	198	156 ± 111	1.4
Y486	3219 USGS	19	262 ± 107	2.5 *
3219 USGS	V811	180	92 ± 146	0.6

* denotes statistically significant regression at the 95% confidence level.