A.M. WASSEF
Survey of Egypt
Giza

ON THE SEARCH FOR RELIABLE CRITERIA OF THE ACCURACY OF PRECISE LEVELLING BASED ON STATISTICAL CONSIDERATIONS OF THE DISCREPANCIES

1. Introduction

In 1867 the European Association of Geodesy decided that a levelling could be qualified as precise when the probable error of the difference in height of two points distant one kilometre apart generally did not exceed 3 millimetres and was never more than 5 millimetres.

At that time, and for nearly half a century thereafter, the only type of error that was envisaged in this specification was the accidental error. For it was only in 1912 that Lallemand submitted at the Hamburg meeting of the International Association of Geodesy the hypothesis that levelling was affected by two categories of errors (i) accidental errors that followed the law of Gauss and could be assessed by a probable accidental error per kilometre η so that the probable error of the difference in height of two bench marks distant D kilometres apart was $\eta \sqrt{D}$, and (ii) systematic errors acting along the full extent of a sector of length L so that the probable error of height difference was proportional to D, say σD. The coefficient of proportionality σ represented the systematic errors and was called the probable systematic error per kilometre. Lallemand regarded the variation of the systematic error per kilometre from one sector to the other as purely accidental, even if the sectors were consecutive, and, hence, admitted that the totality of the values of the probable systematic error per kilometre would add up as accidental errors did. The I.A.G. adopted Lallemand's hypothesis and the computing formulae based on it, and established a category of levelling, called Levelling of High Precision, comprising lines which would be levelled once in each direction, and in which η did not exceed one millimetre and σ did not exceed 0.2 millimetre. The probable systematic error per kilometre σ was to be calculated from the cumulative discrepancy between the two independent measurements plotted against the distance D for sectors extending some 200 kilometres. A straight line was fitted through each set of points departing as little as possible from the broken line joining them while balancing the areas on both sides. Denoting by S the difference of ordinates of the two terminals of a fitted line, the estimate of the discrepancy per kilometre was S/L. The probable value of the estimate of the systematic error thus obtained was calculated from the weighted mean of S^2/L^2 taking L for the weight, thus

$$\sigma^2 = \frac{1}{q\,\Sigma L}\,\Sigma L\,.\,\frac{S^2}{L^2} = \frac{1}{q\,\Sigma L}\,\Sigma\,\frac{S^2}{L}$$

The most serious objection to this system of apparaisal of levelling was raised at the Stockholm meeting of the I.A.G. in 1930 by de Cifuentes and Gil Lasantas and also by Rune. They argued that, for the same method of levelling, the value of the probable systematic error per kilometre changed with the average length of the sectors into which the network was divided for its estimation. It appeared that the larger was the average length, the smaller was the calculated value of σ. This was considered inconsistent with the definition of σ as a probable error per kilometre. To evade this difficulty, Vignal [4], in 1936, proposed a different classification of the errors of levelling. The class of accidental errors was to be retained but they were re–named "erreurs pantophanes", because they acted everywhere. They could be represented by a probable error per kilometre η so that the probable error of the difference in height of two bench marks distant D kilometres apart was $\eta\sqrt{D}$. The systematic errors were to be replaced by the "erreurs apophanes". These errors could also be represented by a probable error per kilometre ζ but its propagation depended on whether the distance D was greater or smaller than a certain minimum length Z. If D was greater than Z the probable error of height difference was given by $\zeta\sqrt{D}$. If D was smaller than Z the influence of the apophane errors was still proportional to $\sqrt{D}$ but the coefficient of proportionality would gradually fall from the value ζ at $D = Z$ to zero as D approaches zero. The minimum length Z was the distance at which the value of ζ calculated from the cumulative discrepancies ceased to depend on D. Vignal also suggested that the combined influence of all the errors could be assessed by a probable total accidental error per kilometre

$$\epsilon = \sqrt{\eta^2 + \zeta^2}$$

for lines the average length of which exceeded the minimum Z, and that the total probable error ϵ should not exceed 2 millimetres per kilometre for levellings of high precision or 6 millimetres per kilometre for levellings of precision. The pantophane discrepancies over a distance of r kilometres were not to exceed ± 6 to $8\,\eta\sqrt{r}$. The apophane discrepancies were not to exceed ± 2 to $2.5\,\zeta\,L^{\frac{3}{4}}$ when L was smaller than "several tens of kilometres" or ± 6 to $8\,\zeta\sqrt{L}$ if L was longer.

These attempts to define and evaluate levelling errors did not escape critical comment. Remarking on the hypothesis that the error which was systematic throughout each line of the network had random values in different lines, Bomford [1], in 1952, wrote "There is no particular reason why this should be so, although there is equally no reason why the systematic error should remain constant indefinitely, particularly in the mean of the fore– and back–levelling, and the theory behind the formulae is not satisfactory". Recently, in 1968, Müller and Schneider [3] declared that they found no room for the "so–called systematic error of classical theory". In their own words "Die Diskrepanz zwischen der Bewertung eines Netzes aus Strecken und Linien verschwindet bei Verwendung geeigneter Gewichte, wie sie aus dem Exponentialgesetz gewonnen wurden, so dass kein Platz mehr für den sogenannten systematischen Fehler der klassischen Theorie vorhanden ist".

In the present paper the problem of evaluating the accuracy of high–precision levelling is investigated in the light of researches on the application of mathematical statistics to the analysis of levelling errors [3, 6–13]. It is hoped to show that, while this venture should continue along all conceivable approaches, the progress made so far is enough to enable us to define without ambiguity the types of error to be reckoned with and the method of estimating each, leading to the development of more reliable indices of accuracy and satisfactory tolerance specifications. These are needed to refine the methods of assessing technical improvements and to increase the utility of levelling data for geodetic and geophysical researches.

2. Error model and the estimation of the random component

Let us set out by supposing that for any position of the levelling instrument the discrepancy δ_{ijk} between the direct and the reverse measurements of the difference of elevation is composed of a constant c', an error κ_i' which is somehow associated with the levelling lines, changing from one line to the other but being constant for one and the same line, and a random component ξ_{ijk} of zero mean and standard deviation τ'. This model of the discrepancy is quite useful. For if we could prove that κ_i' is in fact nil we would at once adopt the estimate of the random error component as representative of the quality of the levelling network. If, on the other hand, κ_i' would prove statistically significant we should push the analysis further to estimate this component and look for its cause. The mathematical model for the analysis of δ according to this mode of classification of errors is

$$\delta_{ijk} = c' + \kappa_i' + \xi_{ijk}'$$

where k denotes the position of the instrument on the j th section of the i th line of the network, a section being the stretch between two consecutive bench marks. Assuming that the successive instrument positions are equally spaced at, say, r metres we have for the discrepancy ρ_{ij} between the direct and the reverse levelling of the j th section between two bench marks distant R_{ij} on the line i

$$\rho_{ij} = \sum_k \delta_{ijk} = (R_{ij}/r)(c' + \kappa_i') + \sum_k \xi'_{ijk}$$

$$= R_{ij}(c + \kappa_i) + \sum_k \xi'_{ijk}$$

Denoting the discrepancy per kilometre of the length of the section by w_{ij} we have

$$w_{ij} = \rho_{ij}/R_{ij} = c + \kappa_i + (1/R_{ij}) \sum_k \xi'_{ijk}$$

$\Sigma \xi'_{ijk}$ has zero mean and standard deviation $(R_{ij}/r)^{\frac{1}{2}} . \tau'$. Hence, the last term in the expression for w_{ij} has zero mean and standard deviation $(1/R_{ij})$ $(R_{ij}/r)^{\frac{1}{2}} . \tau'$, or $\tau/(R_{ij})^{\frac{1}{2}}$. The model for w_{ij}, therefore, takes the form

$$w_{ij} = c + \kappa_i + \xi_{ij}(0, \tau/R_{ij}^{\frac{1}{2}})$$

Levelling networks are accordingly expected to feature the following :

(i) a significant linear regression of ρ_{ij} on R_{ij},

(ii) no significant regression of w_{ij} on R_{ij} and

(iii) an inequality of the within–line variances of w_{ij} that can be eliminated or effectively reduced by weighting w_{ij} proportionally to R_{ij}.

Let us examine the evidence of the levelling networks which have been subject of analysis by statistical methods. The analysis of the levelling network of the Nile delta [6] and the analysis of the Second and the Third Precise Levellings of the Netherlands [11, 13] give ample data for a discussion of the regression of ρ on R. In each case, the Sum of the Squares of the differences of w_{ij} from the respective line–means was broken up into two sums of squares, one is due to linear regression and the other is due to the deviations from the line of regression. Writing

$$B' = \sum_i \sum_j (w_{ij} - w_i)^2$$

$$C' = \Sigma \; \Sigma \; (w_{ij} - w_i)(R_{ij} - R_i) \tag{2}$$

$$A' = \Sigma \; \Sigma \; (R_{ij} - R_i)^2$$

we have

S.S. due to linear regression $= C'^2 / A'$

S.S. of deviations from regression line $= B' - C'^2 / A'$ (3)

The statistical significance of the regression is indicated by the relative values of the mean square due to linear regression and the mean square of the deviations from the regression lines. The mean square is obtained by dividing the sum of the squares by its degrees of freedom. In the present case the sum of the squares due to regression has one degree of freedom and the sum of the squares of the deviations from the regression lines has $N-m-1$ degrees of freedom, where N is the total number of sections and m is the number of lines carrying them. Table 1 gives the relevant quantities.

For the Nile delta network the mean square due to regression, $3.0/1 = 3.0$, is almost equal to the M.S. residual from the regression line. Linear regression is therefore not indicated. For the Second Levelling of the Netherlands the M.S. due to linear regression is larger than the M.S. deviation from the regression line, but the ratio

$$\frac{C'^2/A'}{1} \Big/ \frac{B'-C'^2/A'}{N-m-1} = 9.7/1.9 = 5.1$$

with 1 , 4248 degrees of freedom is too small to indicate significance of linear regression at the 1 °/o level. The Third Levelling of the Netherlands gave the ratio $3.9/1.9 = 2.0$ with 1 , 4717 degrees of freedom. This value is well below the 10 °/o level. We may therefore conclude that the precise levellings in Egypt and in the Netherlands support the hypothesis that the variate $w_{ij} = \rho_{ij}/R_{ij}$ is not associated with R_{ij} within the lines of the networks.

Let us now study the evidence given by the analysis of the Precise Levelling of the German Democratic Republic by Müller and Schneider [3] in 1968. This network comprised 7138 sections on 107 lines. Müller and Schneider set out by calculating the correlation coefficient between the discrepancy ρ (regardless of the sign) and R within each of the lines, using the well–known expression of the product–moment coefficient of correlation

$$\frac{\Sigma(|\rho| - |\bar{\rho}|)(R-\bar{R})}{\sqrt{\Sigma(|\rho| - |\bar{\rho}|)^2 \, \Sigma(R-\bar{R})^2}}$$

The value of the correlation coefficient for each line is quoted in Table 2 together with the value which would indicate significance at the 1 °/o level.

The authors noted that 75 °/o of the correlation coefficients were significantly different from zero at the probability level 1 °/o and that the correlation coefficient for all the $(|\rho|, R)$ was 0.45 which is highly significant since, for 7136 degrees of freedom, a correlation coefficient equal to 0.09 would be significant at the 1 °/o level. The authors further remarked that "taking into consideration that the correlation coefficient calculated from the observations is considerably suppressed from its true value by the chance deviations of the observations, it must be supposed that the correlation between $|\rho|$ and R is in fact much larger than 0 45". All this is good evidence that the association of the two variables is in fact linear or can be effectively represented by a linear relationship. For when we fit a straight line of regression to data, the sum of the squares of the deviations from the line of regression of either variable on the other is equal to $N\sigma^2(1-k^2)$, where k is the correlation coefficient and σ is the standard deviation of the variable whose regression is sought. If the correlation coefficient is 1 or -1 the sum of the squares of the deviations from the regression line is zero. Consequently, each deviation is zero, and all the points lie on both lines of regression. These two lines then coincide ; and there is a linear functional relationship between the variables, giving perfect correlation. The nearer k^2 is to unity, the closer are the points to the lines of regression, and the nearer are these two lines to coincide. Thus the magnitude of the correlation coefficient may be taken as a measure of the degree to which the association between the variables approaches a linear functional relationship.

The equation $|\rho| = 0.47\ R^{0\ 6}$ which Müller and Schneider obtained from their analysis of the regression of $|\rho|$ on R should not distract our attention from the evidence of linearity given by the correlation coefficient, because the regression analysis which led to the index 0.6 had not been designed to examine the association between ρ and R within the lines of the network. Let us see how the regression was calculated. The 7138 values of $|\rho|$ were arranged according to increasing values of R and divided into 23 classes in such a way that each class contained only closely succeeding values of R and roughly equal numbers of $|\rho|$ values, about 360 per class. In this way, to each R_j of a class there belonged n_j values of $|\rho|$. The weighted mean $\overline{R}$ was calculated for each class, $\overline{R} = \Sigma\, n_j R_j / \Sigma\, n_j$, and a representative value of $|\rho|$ was calculated, viz. $|\overline{\rho}| = \Sigma$ (all $|\rho|$ of the class) $/ \Sigma\, n_j$. A curve of the form $|\rho| = a R^{\beta}$ was fitted to the $(|\overline{\rho}|, \overline{R})$ values both graphically and numerically using weights, giving $|\overline{\rho}| = 0.47\ R^{0.6}$. This, however, does not describe the behaviour of ρ_{ij} within the lines, because each of the twenty three classes was compiled from elements selected from the 107 lines of the network. The values of ρ in a class thus constituted a mixture from sub–populations that are likely to be of different means. In other words, the process of classification mixed the within–line variation with the line effect and hence the resulting regression function describes neither and does not confute the argument for the representation of the discrepancy per kilometre within the lines by the expression (1).

Let us now proceed to examine whether the estimates

$$s_i^2 = \frac{1}{n_i - 1} \Sigma (w_{ij} - w_i)^2 \tag{4}$$

of the variance within the lines may be combined to give an estimate of the random or accidental error. This would be the case if they formed a homogeneous set. To test homogeneity, in the sense of the equality of the estimates of variance, the Barlett function

$$M = (\sum_1^m \nu_i) . \log_e \left\{ \Sigma (\nu_i s_i^2) / \Sigma \nu_i \right\} - \Sigma \nu_i \log_e s_i^2 \tag{5}$$

was calculated for each net, noting that $\nu_i = n_i - 1$ is the number of degrees of freedom on which the estimation of s_i^2 was based. The function M is distributed as χ^2 with $m-1$ degrees of freedom. Table 3 shows the size of M compared with the $1\ ^\circ/\circ$ value of χ^2. The within–line variances are glaringly heterogeneous in all three networks. When weights R_{ij} were attached to w_{ij}, however, the inequality of the within–line variance for one–kilometre sections given by the equation

$$\bar{s}_i^2 = \frac{1}{n_i - 1} \Sigma R_{ij} (w_{ij} - \bar{w}_i)^2 \tag{6}$$

becomes much less significant (Table 4). This shows that weighting reduced the inequality of the within–line variances in the Egyptian network to insignificance, and it did much to stabilize the variance within the Dutch lines although it could not quite eliminate the significant heterogeneity.

At this point of the analysis it repays to scrutinize the data in search of outlying values. Let us, for example, examine the set of variances of the Third Levelling of the Netherlands (Table 5). Sixteen values exceed 2.0 and two values are less than 0.25. When these values are set aside, the value of M sharply falls from 369 to 202 which does not reach the $1\ ^\circ/\circ$ level of significance. To demonstrate that the sorting out of the outlying values was not carried out too far let us calculate the probability of occurrence of an outlying value and compare it with the observed frequency. Noting that the casting away of the outlying values leads to a reduction of the estimate of the variance to 1.18 with 159 degrees of freedom, and taking 20 for the average number of degrees of freedom that are available for the estimation of a within–line variance, we should expect no more than $1\ ^\circ/\circ$ of the values to exceed $2 \times 1.18 = 2.36$. We may thus expect to find one value, possibly two, whereas the actual number of values exceeding 2.36 in Table 5 is 12. We thus come to the conclusion that weighting w_{ij} according to R_{ij}, possibly followed by the setting aside of a small number of outlying values on probability basis, has a very good chance to lead to a set of estimates of the within–line variance that shows no excessive heterogeneity.

Opinion may, however, be divided on the level of significance at which the hypothesis of equality of the within–line variances should be rejected. In this respect it is worthwhile to consider the consequences of accepting the hypothesis of equality when it should be rejected, and the other way round. Two points merit mention :

(1) A large value of M , indicating inequality, may ensue from the presence of a few bad observations or from observations that are too good to belong to the main bulk of the data. In such cases the sorting out of a small number of outlying values is legitimate provided it is done on probability basis. The variances of the main bulk will display moderate differences and the estimation of a common variance becomes feasible. The common variance is the measure of the accidental or random part of the error and the gauge which can be used to test the significance of non–random effects.

(2) In case the large M can not be attributed to a small number of outlying values we may then conclude that the network is essentially heterogeneous and proceed to study the feasibility of breaking it up into two or more homogeneous blocks.

To sum up, the within–line variances

$$s_i^2 = \frac{1}{n_i - 1} \Sigma (w_{ij} - w_i)^2 \quad , \; i = 1 , 2 , \ldots , m$$

are not expected to be estimates of a common variance unless all the sections are equal in length. The estimate of the within–line variance for one–kilometre section, denoted by $\bar{s}_i^2$, should be more stable. They are obtained by weighting w_{ij} proportionally to R_{ij} , thus

$$\bar{s}_i^2 = \frac{1}{n_i - 1} \Sigma R_{ij} (w_{ij} - \bar{w}_i)^2$$

The expected equality of these estimates may however be upset by a small number of outlying values. When these are sorted out on probability basis, the rest, which constitutes the main bulk of the data, should be homogeneous enough for the purpose of estimating a common within–line variance for one–kilometre section by pooling the weighted sum of squares. In this way we obtain

$$\bar{s}_R^2 = \frac{1}{N - m} \Sigma R_{ij} (w_{ij} - \bar{w}_i)^2 \qquad (7)$$

where m is the number of lines and N is the number of sections on them. $\bar{s}_R^2$ is based on $N - m$ degrees of freedom.

3. Assessment of the non–random errors

Let us now proceed to examine the variation of the m line–means of w_{ij} . Dividing by $m-1$ the weighted sum of the squares of the differences between the line means and their weighted mean we obtain the Mean Squares between line–means

$$\frac{1}{m-1} \Sigma L_i (\bar{w}_i - \bar{w})^2 \tag{8}$$

where $L_i = \Sigma R_{ij}$ is the length of the i th line and

$$\bar{w}_i = \Sigma R_{ij} w_{ij} / L_i$$

$$\bar{w} = \Sigma L_i \bar{w}_i / \Sigma L_i$$

If only the common random error existed the M.S. between–line–means would be an unbiased estimate of the variance of the random error and should not differ significantly from the estimate $\bar{s}_R^2$ derived from the within–line variation. If, on the other hand, the M.S. between line–means proves to be significantly larger than $\bar{s}_R^2$ we can no longer assume that all κ_i are nil nor that the all variation in the network is accountable by the random error, and we should try to calculate an estimate of the variation of κ_i .

The equality of the M.S. between line–means and $\bar{s}_R^2$ is tested by calculating the ratio

$$\frac{1}{m-1} \Sigma L_i (\bar{w}_i - \bar{w})^2 / \bar{s}_R^2$$

and comparing the ratio with the tabulated F at the desired probability level using the right number of degrees of freedom. Table 6 gives the relevant quantities for the three networks mentioned above. The M.S. between line–means is evidently much larger than the estimate $\bar{s}_R^2$ of the variance of the random error. To calculate the variance of κ_i , which we denote by $\bar{s}_L^2$, it suffices to work out the mathematical expectation of the M.S. between line–means. First, we note that

$$\Sigma L_i (\bar{w}_i - \bar{w})^2 = \Sigma L_i (\bar{w}_i - w_c)^2 - \Sigma L_i (\bar{w} - w_c)^2 \tag{10}$$

We also have

$$E(\bar{w}_i - w_c)^2 = \bar{s}_L^2 + \frac{\bar{s}_R^2}{L_i}$$

$$E(\bar{w} - w_c)^2 = \left\{ \Sigma L_i^2 / (\Sigma L_i)^2 \right\} . \bar{s}_L^2 + (1/\Sigma L_i) . \bar{s}_R^2$$

Hence,

$$E \; \Sigma L_i (\bar{w}_i - \bar{w}) = (\Sigma L_i - \Sigma L_i^2 / \Sigma L_i) . \bar{s}_L^2 + (m-1) . \bar{s}_R^2$$

It follows that the estimate of the variance of the non–random error, or the line effect, is given by the equation

$$\bar{s}_L^2 = \left\{ \sum^{m} L_i (\bar{w}_i - \bar{w})^2 - (m-1) . \bar{s}_R^2 \right\} / (\Sigma L_i - \Sigma L_i^2 / \Sigma L_i) \quad (11)$$

with $m-1$ degrees of freedom.

We may wish to ascertain that the significance of the non–random effect has not resulted from the casting away of outlying values. This is very unlikely unless the lines which have outlying within–line variances happen to have outlying line–means. To illustrate this point, the contributions of the sixteen lines of the Third Levelling whose $s_i^2 \geqslant 2.0$ are given in Table 7. The total contribution is 43.8 which leaves almost unchanged the value of the M.S. between line–means, now $472/161 = 2.93$. Since the within–line M.S. is now 1.18 in place of 1.30 the statistical significance of the between–lines effect is accentuated.

4. Types of non–random error

Let us now consider the types of error which constitute the line–effect estimated by $\bar{s}_L^2$, in an attempt to tie it up with the idea of systematic error and the concept of internal correlations. Let us first cite one of the classical examples which showed that the breaking up of the variation into a constant systematic error and a random error was grossly insufficient. The experiment, which was conducted by Karl Pearson about 1900, consisted of bisecting a line by eye, the accuracy being afterwards checked by measurement. There were three observers, who each made about 500 observations. When the observations were taken in groups of 25 to 30 it was found that the means fluctuated, not by the amounts that would be expected of the means of 25 to 30 random errors, but by as much as the means of 2 to 15 independent observations should. The non–random error was not constant but reversed its sign at irregular intervals.

Errors of observations in levelling are expected to cause internal correlations of the type just described. But one can cite many other causes, for practically every fluctuation of the physical or meteorological conditions under which the levelling is carried out is a likely cause of internal correlations, provided that it lasts the comparatively short time needed to level a line or the larger part of a line. To these should be added the systematic errors proper, a systematic error in levelling being appropriately defined as a quantity associated with an observation of a height difference which, if its value could be accurately known for one observation of a series, would be calculable for all others.

The collective of systematic errors [2, 5] and other causes of internal correlations may be called the non–random error and can be assessed by the standard error $\bar{s}_L$.

5. Propagation of the random and the non–random errors in a levelling network

To obtain the standard value of the random error of the height difference over a line of length L_i kilometres obtained by double levelling we proceed as follows : The variance of a section of length R_{ij} kilometres is $\bar{s}_R^2 / R_{ij}$ and, hence, the variance of ρ_{ij} is $(\bar{s}_R^2 / R_{ij}) \cdot R_{ij}^2 = R_{ij}\,\bar{s}_R^2$. Since ρ_{ij} is the difference between the direct and the reverse measurements of the height difference, the variance of the mean of the two measurements is one quarter the variance of ρ_{ij} , or $\frac{1}{4}\,\bar{s}_R^2 \cdot R_{ij}$. Adding up the variances of the sections of the i th line we find that the variance of the mean of the direct and the reverse measurements for the line is

$$\frac{1}{4}\,\bar{s}_R^2 \sum_j R_{ij} = \frac{1}{4}\,\bar{s}_R^2 \cdot L_i \tag{12}$$

When the line effect is statistically established, the variance is calculated as follows : The quantity κ_i is presumably common to all w_{ij} of the i th line, hence it contributes $\kappa_i \sum_j R_{ij}$ to the total discrepancy ρ_{ij} on this line and

$$\frac{1}{4}\,\bar{s}_L^2 \left(\Sigma R_{ij}\right)^2 = \frac{1}{4}\,\bar{s}_L^2 \cdot L_i^2 \tag{13}$$

to the variance of the mean of the direct and the reverse measurements.

Summing over i we see that the variance of the mean is given by

$$\frac{1}{4}\left(\bar{s}_L^2 \cdot \sum^{p} L_i^2 + \bar{s}_R^2 \cdot \sum^{p} L_i\right) \tag{14}$$

It follows that the total error, represented by the square root of this variance, increases neither with the distance nor with its square root, nor indeed with any particular function of the distance. It does not seem presumptuous at the moment to submit that the weight function which may rightly be assigned to the heights of bench marks in the process of network adjustment is

$$1 / (\bar{s}_L^2 \sum^{p} L_i^2 + \bar{s}_R^2 \sum^{p} L_i) .$$

o

o o

REFERENCES

[1] G. BOMFORD : Geodesy. Oxford (1952).

[2] M. KNEISSL : Nachweis systematischer Fehler beim Feinnivellement. Abh. Bayer. Akad. Wiss. N.F. 68 (1955).

[3] K. MÜLLER & E. SCHNEIDER : Nivellementswidersprüche in statistischer Sicht. Vermessungstech. 16 (1968).

[4] J. VIGNAL : Evaluation de la précision d'une méthode de Nivellement. Bulletin Géodésique (1936) No. 49.

[5] Karl–Heinz RODIG : Zum Einfluss systematischer Fehler beim Nivellement, insbesondere beim motorisierten Prazisionsnivellement. Geod. Geoph. Veroff. R. III H. 22 (1971).

[6] A.M. WASSEF : Statistical analysis of discrepancies in levelling with applications to the first–order levelling of the Nile delta. Bull. Géodé–sique (1955) No. 36.

[7] A.M. WASSEF & F.Z.A. MESSIH : Study of the normality and homogeneity of levelling errors. Bull. Géodésique (1959) No. 52.

[8] A.M. WASSEF : Note on the application of mathematical statistics to the analysis of levelling errors. Bull. Géodésique (1959) No. 52.

[9] A.M. WASSEF & F.Z.A. MESSIH : On the statistical distribution of levelling errors. Bull. Géodésique (1960) No. 56.

[10] A.M. WASSEF : Principles and methods of statistical analysis of levelling errors. Trav. Assoc. Intern. Géodésie (1962) T. 21.

[11] A.M. WASSEF & F.Z.A MESSIH : Introductory statistical analysis of the second and third levellings of the Netherlands. Trav. Ass. Intern. Géodésie (1968). T. 23.

[12] A.M. WASSEF & F.Z.A. MESSIH : Statistical analysis of the precise levelling network of Egypt and the Netherlands. Proceedings of the 6th Arab Science Congress. Damascus (1969).

[13] A.M. WASSEF & F.Z.A. MESSIH : Introductory statistical analysis of the second and third levellings of the Netherlands (Part 2). I.A.G., Moscow (1971).

Table 1

Insignificance of the regression of w on R within the lines in three large levelling networks

	Nile Delta network (1927 – 1937)	Second Levelling of the Netherlands (1926 – 1940)	Third Levelling of the Netherlands (1950 – 1959)
B'	3529.4	8112.5	8877.4
A'	843.0	1330.1	658.6
C'	50.2	– 113.5	– 50.7
C'^2 / A'	3.0	9.7	3.9
d.f.	1	1	1
$B' - C'^2 / A'$	3526.4	8102.8	8873.5
d.f.	1097	4248	4717
$\frac{B' - C'^2 / A'}{N - m - 1}$	3.2	1.9	1.9

Table 2

Correlation of ρ on R

Line		1 °/o level	Line		1 °/o level	Line		1 °/o level	Line		1 °/o level
1	0.433	0.21	28	0.356	0.32	55	0.265	0.32	82	0.473	0.26
2	0.413	0.26	29	0.553	0.25	56	0.599	0.27	83	0.441	0.17
3	0.191	0.70	30	0.524	0.20	57	0.479	0.35	84	0.493	0.30
4	0.397	0.28	31	0.460	0.22	58	0.441	0.38	85	0.466	0.37
5	0.487	0.24	32	0.649	0.44	59	0.517	0.28	86	0.326	0.34
6	0.535	0.25	33	0.796	0.35	60	0.595	0.51	87	0.550	0.25
7	0.443	0.32	34	0.449	0.29	61	0.596	0.34	88	0.148	0,31
8	0.601	0.49	35	0.747	0.56	62	0.787	0.60	89	0.438	0.40
9	0.547	0.30	36	0.633	0.25	63	0.299	0.39	90	0.116	0.37
10	0.599	0.39	37	0.352	0.37	64	0.425	0.33	91	0.384	0.31
11	0.411	0.37	38	0.209	0.35	65	0.257	0.24	92	0.471	0.39
12	0.416	0.34	39	0.306	0.37	66	0.092	0.53	93	0.137	0.55
13	0.363	0.21	40	0.270	0.41	67	0.398	0.35	94	0.080	0.34
14	0.506	0.36	41	0.445	0.39	68	0.548	0.28	95	0.519	0.26
15	0.687	0.46	42	0.413	0.43	69	0.641	0.25	96	0.485	0.33
16	0.394	0.29	43	0.436	0.45	70	0.416	0.25	97	0.643	0.47
17	0.278	0.27	44	0.588	0.67	71	0.607	0.36	98	0.436	0.23
18	0.539	0.34	45	0.781	0.47	72	0.438	0.26	99	0.413	0.25
19	0.575	0.35	46	0.640	0.57	73	0.259	0.29	100	0.449	0.20
20	0.432	0.33	47	0.474	0.34	74	0.415	0.27	101	0.097	0.37
21	0.504	0.28	48	0.547	0.33	75	0.516	0.40	102	0.387	0.23
22	0.637	0.58	49	0.347	0.32	76	0.544	0.30	103	0.180	0.36
23	0.432	0.53	50	0.475	0.29	77	0.248	0.28	104	0.546	0.25
24	0.400	0.34	51	0.431	0.31	78	0.185	0.29	105	0.334	0.61
25	0.347	0.54	52	0.419	0.25	79	0.546	0.52	106	0.489	0.25
26	0.592	0.39	53	0.446	0.24	80	0.560	0.39	107	0.522	0.23
27	0.276	0.42	54	0.183	0.28	81	0.639	0.34			

Table 3

Inequality of the within–line variances of the unweighted W_{ij}

	Nile Delta Levelling	Second Levelling of the Netherlands	Third Levelling of the Netherlands
M	140	622	468
$\chi^2_{.01}$	118	289	223

Table 4

Stabilization of the within–line variance by weighting proportionally to R_{ij}

	Nile Delta Levelling	Second Levelling Netherlands	Third Levelling Netherlands
M for s_i^2	140	622	468
M for $\bar{s}_i^2$	102	323	369

Table 5

The within–line variances for one–kilometre section in the Third Levelling of the Netherlands

Line	ν_i	$\bar{s}_i^2$	Line	ν_i	$\bar{s}_i^2$	Line	ν_i	$\bar{s}_i^2$	Line	ν_i	$\bar{s}_i^2$	Line	ν_i	$\bar{s}_i^2$	Line	ν_i	$\bar{s}_i^2$
1	9	3.788	31	14	0.703	61	8	0.321	91	9	0.561	121	32	1.723	151	28	0.641
2	8	0.591	32	7	0.234	62	6	0.821	92	19	1.418	122	28	1.357	152	28	1.339
3	18	2.882	33	35	0.841	63	34	1.931	93	26	1.851	123	15	2.706	153	39	0.827
4	5	3.088	34	27	0.657	64	20	0.962	94	10	0.791	124	36	0.921	154	23	0.963
5	15	2.672	35	15	1.693	65	14	1.019	95	26	0.907	125	17	1.079	155	39	1.420
6	17	1.805	36	77	1.044	66	50	0.947	96	16	1.556	126	40	0.805	156	6	0.313
7	11	0.483	37	31	1.493	67	9	0.666	97	21	1.440	127	43	1.298	157	37	1.397
8	34	2.368	38	29	3.060	68	17	0.384	98	34	1.613	128	30	1.449	158	14	0.781
9	17	1.372	39	61	1.332	69	11	0.430	99	28	2.324	129	17	1.628	159	27	0.847
10	8	1.528	40	56	1.133	70	19	1.249	100	17	0.908	130	43	1.278	160	32	1.134
11	13	1.022	41	47	1.472	71	33	1.356	101	48	1.745	131	57	1.133	161	26	1.078
12	30	1.325	42	32	0.642	72	4	0.091	102	51	2.451	132	15	0.956	162	39	1.442
13	23	1.362	43	36	1.139	73	23	1.504	103	34	1.456	133	38	0.804	163	58	1.146
14	13	1.485	44	16	1.319	74	17	1.098	104	12	0.865	134	6	1.290	164	28	0.468
15	13	1.023	45	34	0.772	75	27	1.538	105	33	1.155	135	28	1.029	165	27	0.900
16	17	1.015	46	22	0.386	76	10	0.894	106	23	1.226	136	40	1.614	166	23	0.645
17	6	0.479	47	25	1.150	77	41	1.412	107	42	2.311	137	55	1.023	168	7	0.475
18	30	0.766	48	20	0.681	78	24	1.283	108	29	2.126	138	39	1.294	169	5	0.613
19	15	1.721	49	36	0.761	79	26	1.342	109	32	0.952	139	30	1.248	170	43	1.771
20	19	1.098	50	33	1.158	80	40	1.665	110	50	1.108	140	31	1.198	171	26	1.493
21	29	1.475	51	17	1.093	81	10	1.324	111	39	1.373	141	54	1.115	173	43	1.517
22	34	0.619	52	37	1.057	82	5	0.732	112	52	1.613	142	27	0.854	175	19	1.091
23	23	0.731	53	14	0.772	83	11	0.750	113	38	1.937	143	21	0.998	176	27	0.768
24	11	0.687	54	35	1.616	84	5	1.505	114	12	0.867	144	32	1.260	177	21	3.451
25	6	1.269	55	15	1.527	85	16	1.473	115	23	1.076	145	31	0.782	178	8	3.986
26	10	0.689	56	21	1.105	86	32	1.306	116	26	3.846	146	25	1.625	179	35	0.996
27	47	0.753	57	12	1.092	87	24	0.868	117	39	1.664	147	38	1.550	180	47	1.406
28	41	1.052	58	24	1.576	88	46	1.366	118	15	1.260	148	12	0.949	181	37	1.746
29	20	2.033	59	31	1.122	89	10	1.240	119	37	0.704	149	24	1.297			
30	22	2.805	60	42	1.014	90	45	1.122	120	49	1.355	150	30	1.210			

Table 6

Significance of the between–lines effect

	Nile Delta Levelling	Second Levelling Netherlands	Third Levelling Netherlands
Calculated ratio	1.71	1.47	2.24
d.f.	86, 1098	236, 4249	177, 4718
$F_{.01}$	1.41	1.24	1.27
$F_{.001}$	1.49	1.32	1.37

Table 7

Contribution, to the between–lines sum of squares, of lines of large outlying values.

Line	$\overline{w}_i$	L_i	$L_i(\overline{w}_i - \overline{w})^2$	Line	$\overline{w}_i$	L_i	$L_i(\overline{w}_i - \overline{w})^2$
1	– 0.442	14	2.4	99	+ 0.298	30	3.1
3	– 0.081	23	0.1	102	– 0.239	44	2.0
4	– 0.106	6	0	107	+ 0.146	45	1.3
5	+ 0.443	20	4.4	108	– 0.103	22	0.1
8	+ 0.524	42	12.7	116	– 0.412	28	4.2
29	+ 0.042	21	0.1	123	– 0.651	18	7.0
30	– 0.218	27	1.0	177	– 0.118	30	0.3
38	+ 0.163	33	1.2	178	– 0.687	9	3.9