# **Correlation in Polish precise levelling network**

A. ŁYSZKOWICZ<sup>1</sup>, A. JACKIEWICZ<sup>2</sup>

#### Abstract

When a levelling network is adjusted, the height differences  $\Delta H$  refer to individual levelling lines are given weights that are inversely proportional to the length *L* of these lines. This weighting scheme is acceptable when the height differences  $\delta H$  of the segments within each line are statistically independent. This condition has always been assumed within geodetic agencies while research groups express doubts e.g. (Lucht, 1972,), (Remmer, 1975) and others.

The main purpose of this paper is determination of the correlation between neighboring segments of a levelling lines, estimation lines weights and estimation how the new weighting scheme influence on the results of the adjustment of the levelling network measured in 1999-2002 in Poland.

#### **1** Introduction

In Poland precise levelling network was measured four times. The first precise levelling campaign began in 1926 and was finished in 1937. The network consists of 5907 sections, 121 lines and 36 loops. Total length of levelling line is 10 046 km (Wyrzykowski, 1988).

The second levelling campaign was carried out in two stages. The first measurements were done in 1947-1950, and the second measurements in 1953-1955. The second version of network comprises 4500 sections, 60 levelling lines and 12 loops. Total length of levelling lines is 5 778 km (ibid.).

The third levelling campaign was in 1974-1982. The network consists of 15827 sections, 371 lines and 135 loops. The total length of levelling lines is 17 015 km. The observed height difference was corrected due to temperature, rod scale and tides (ibid.).

The fourth precise levelling campaign started in 1999 and was finished in 2003. The network consists of 16 150 sections with average length 1.1 km, 382 lines with average length about 46 km, 135 loops, and 245 nodal points. Total length of levelling lines is 17 516 km (Łyszkowicz and Leończyk,2006).

All campaigns were adjusted by the least squares method assuming that the observations, i.e. height differences of the leveling lines, are not correlated. Weights of observations were computed from the formula

$$p = \frac{l}{\sigma_{_{AH}}^2} \tag{1-1}$$

where  $\sigma_{\Delta H}$  is a standard deviation of the differences  $\Delta H$  of the levelling line long *L* km.

In a case of independent observations a standard deviation of height differences of a line is express by the known relation

$$\sigma_{\Delta H} = \sigma_{I} \sqrt{L} \qquad (1-2)$$

where  $\sigma_i$  is a standard deviation of a height difference of a section long 1 km.

In a case of dependent observation (correlated) equation (1-2) have the following form

$$\sigma_{\rm AH} = \sigma_{\rm I} L^{\alpha} \tag{1-3}$$

where  $\alpha$  is a parameter defined in a range from 0.5 up to 1 ((Vaniček and Grafarend, 1980).

There are some works e.g. (Lucht, 1972), (Dymowski, 1973) in which authors investigated the correlation in small levelling networks. They showed that exist the correlation between the observed height differences of the sections. The comparatively small observation data set is the principal defect of quoted works and conclusions.

The results of the fourth campaign in Poland delivered numerous and good data set to investigate the correlation qualitatively. Therefore it was decided to count correlations in the levelling network this time on the basis of the large and reliable observational material.

#### 2 Covariance matrix

The observation equation of the levelling line connecting two benchmarks A and B (Fig. 2-1) have the form

$$v_{AB}^{\Delta H} = H_B - H_A - \Delta H_{AB} \tag{2-1}$$

<sup>&</sup>lt;sup>1</sup>Adam Łyszkowicz, University of Warmia and Mazury, 12 Heweliusza St, PL 10-724 Olsztyn, Fax: +48 89 5234878, Tel.: +48 89 5234579, e-mail: adam@moskit.uwm.edu.pl,

<sup>&</sup>lt;sup>2</sup> Anna Jackiewicz, Koszalin University of Technology, 2 Śniadeckich St, PL 2 75-453 Koszalin, Fax +48-94 34-27-652, Tel. +48 94 3478500, e-mail: agnos@wp.pl

where  $\Delta H_{AB}$  is observed height difference of a line AB in any arbitrary height system corrected due to systematic errors.



Fig. 2-1 Levelling line and a section

Height difference  $\Delta H_{AB}$  is a sum of measured differences  $\delta H_{AB}$  of each section of a line (Fig. 2-1)

$$\Delta H = \sum_{i=1}^{n} \delta H_{i} = \mathbf{u} \, \delta \mathbf{H} \tag{2-2}$$

where **u** is a unit vector.

$$\mathbf{u} = \begin{bmatrix} l & l & \dots & l \end{bmatrix}$$
(2-3)

In order to adjust the network we have to know the covariance matrix  $C_{\mathcal{SH}}$  of observation  $\mathcal{\delta H}_i$ . In a case, when the height differences  $\mathcal{\delta H}_i$  are independent, then the variance of observation  $\Delta H$  is

$$\sigma_{\Delta H}^2 = \mathbf{u} \, \mathbf{C}_{\delta H} \mathbf{u}^T \tag{2-4}$$

and  $\mathbf{C}_{\delta H}$  is covariance matrix of height differences  $\delta H$  of a successive sections of a line. Covariance matrix  $\mathbf{C}_{\delta H}$  can be express by the correlation coefficients in the form

$$\mathbf{C}_{\partial H} = \sigma_{I} \begin{bmatrix} I & r_{I} & r_{2} & \dots & r_{n} \\ r_{I} & I & r_{I} & r_{2} \dots & r_{n-I} \\ r_{2} & r_{I} & I & \ddots & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \\ r_{n} & r_{n-I} & \vdots & r_{I} & I \end{bmatrix}$$
(2-5)

where correlation coefficient  $r_i$  is defined in the following way

$$r_i = \frac{\sigma_{kl}}{\sigma_k \sigma_l} \operatorname{dla} i = |k-l|.$$

If all differences  $\delta H_i$  of the same line are measured with the same accuracy, what mean the same variance  $\sigma_i^2$ , and if there are no correlation between observations, then covariance matrix  $C_{\delta H}$  is unit matrix and variance of height difference of a line is defined by the formula (1-2).

In a case when observations  $\delta H_i$  are dependent, then the variance  $\sigma_{\Delta H}^2$  of height difference  $\Delta H$  of a levelling line is inside the following compartment

$$\sigma_I^2 L \le \sigma_{AH}^2 \le \sigma_I^2 L^2 \tag{2-6}$$

Or more generally in the form

$$\frac{\sigma_{_{\Delta H}}^2}{\sigma_{_{I}}^2} = L^{\alpha} \tag{2-7}$$

where  $0.5 < \alpha < 1$ . The graphical illustration of the equation (2-7) is show on the Fig. 2-2.



Fig. 2-2 Law of propagation of variances of correlated observations

# 3 Calculation of correlation from observations

If we assume, that we have *n* random v  $\alpha = 1.0$  $X_2, \dots, X_n$  which have normal distribution by  $S_T$  the sum of the random vector such that

$$S_T = X_i + X_k$$
 for  $T = (k-i) = 1, 2, 3....n-1$  (3-1)

Then the variance of the sum of the correlated random variables  $X_i$ ,  $X_k$  is

$$\sigma_{s_{\tau}}^{2} = \sigma_{x_{i}}^{2} + 2\sigma_{x_{i}x_{k}} + \sigma_{x_{k}}^{2} \quad (3-2)$$

Assuming, that  $\sigma_{x_i}^2 = \sigma_{x_k}^2 = \sigma^2$  we have

$$\sigma_{s_T}^2 = 2\sigma^2 + 2\sigma_{x_i x_k} \tag{3-3}$$

From the definition of the correlation coefficient we have

$$r_{ik} = \frac{\sigma_{X_i X_k}}{\sigma^2} \tag{3-4}$$

It means that from equation (3-3) and (3-4) for each sub diagonal of the matrix  $C_{_{SH}}$  we have

$$r_{T} = r_{ik} = \frac{\sigma_{s_{T}}^{2}}{2\sigma^{2}} - 1$$
(3-5)

In a case when we have particular realization  $x_1$ ,  $x_2$ , ..., $x_n$  of the random vector  $X_1$ ,  $X_2$ , ..., $X_n$  then estimation of a variance  $\sigma_{s_r}^2$  can be computed from the formula (Lucht, 1972), (Dymowski, 1973)

$$\sigma_s^2 = \frac{\sum s_T^2}{n_s}$$
(3-6)

and correlation coefficients can be computed easily from (3-5).

The computed coefficients are more or less reliable. It depends on the size of population. The bigger population is, the more credible are values of these coefficients. The investigation of the reliability of the computed coefficients consists in checking the hypothesis H<sub>0</sub>: r = 0, again the H<sub>1</sub>:  $r \neq 0$ . The hypothesis H<sub>0</sub> is true, when the computed statistics *Z* is smaller than the boundary value  $Z_{\underline{\alpha}}$  (Harvey,

1990, p. 72) given by

$$\left|\frac{\sqrt{n-3}}{2}\ln\left(\frac{l+r}{l-r}\right)\right| < Z_{\frac{\alpha}{2}}$$
(3-7)

where *n* is the number of a data in population, *r* is correlation coefficient and statistic  $Z_{\underline{\alpha}}$  have normal

distribution at a significance level  $\alpha$ .

#### 3.1 Computation of neighboring correlation

In order to know if and how the coefficients of the correlation change in the levelling network, three test areas were chosen (Fig. 3-1) and they were denoted as the test nets # 1, # 2, # 3. Finally the correlation coefficients were computed for the whole network.



Fig. 3-1 Levelling network of Poland and the location of three test sub networks

The test net # 1 consists of 55 lines and 2226 sections. The computation of the correlation  $r_l$ ,  $r_2$ ,... $r_n$  based on normalized misclosures  $\rho^*$  of neighboring observations  $\delta H$ 

$$\rho^* = \frac{\rho}{\sqrt{r}} \tag{3-8}$$

where  $\rho$  is a misclosures from double leveling of leveling section *r* km long. Then for T = 1, 2, 3...i the sums were formed for every line

$$s_{T} = \rho_{i}^{*} + \rho_{k}^{*} \qquad (3-9)$$

and next their empirical variances were estimated.

$$\sigma_T^2 = \frac{l}{n_s} \sum s_T^2 \qquad (3-10)$$

where  $n_s$ , for each given *T*, is a number of summarizing components. Variance  $\sigma^2$  was estimated from

$$\hat{\sigma}^2 = \frac{l}{n} \sum \left( \rho^* \right)^2 \tag{3-11}$$

Calculation was done for the first three coefficients. The significance of the correlation coefficients was computed according to the formula (3-7). The level of significance was assumed equal  $\alpha = 0.05$ . The results of calculations are given in Table 3-1.

Table 3-1

Correlation coefficient	Computed value	u Statistic Z		Z <sub>0.025</sub>	Significant ?
$\mathbf{r}_1$	0.145	2172	6.796	1.96	yes
$r_2$	0.122	2116	5.626	1.96	yes
r <sub>3</sub>	0.062	2061	2.807	1.96	yes

From results of calculation shown in Table 3-1 we can noticed that computed first three coefficients are essential and that the degree of correlation is rather weak.

The test net # 2 (northern) consists of 36 line and 1796 sections. First three correlation coefficients and their significance were computed in the identical way as for the net # 1. The results of the calculations are given in Table 3-2.

Table 3-2

Correlation coefficient	Computed value	n	Statistic Z	Z <sub>0.025</sub>	Significant?
r <sub>1</sub>	0.149	1760	6.284	1.96	yes
r <sub>2</sub>	0.143	1724	5.964	1.96	yes
r <sub>3</sub>	0.082	1688	3.359	1.96	yes

Computed values of the first three coefficients are of the same order as in the case of the net, # 1 but their credibility is considerably better, because they were computed from almost two times bigger population.

The test net # 3 (south) consists of 43 lines and 1874 sections. First three correlation coefficients o and their significance were estimated in the identical way as for the net # 1. The results of calculations are shown in Table 3-3.

Table 3-3						
Correlation coefficient	Computed value	n	Statistic Z	Z <sub>0.025</sub>	Significant?	
r <sub>1</sub>	0.140	1831	6.016	1.96	yes	
r <sub>2</sub>	0.113	1788	4.811	1.96	yes	
r <sub>3</sub>	0.130	1746	5.481	1.96	yes	

The whole network consists of 382 lines and 16150 sections. First three values of correlation coefficients and their significance were computed in the same way as for the net # 1. The results of calculations are shown in the Table 3-4.

Tal	hle	3	-4
1 a	uic	2	

Correlation coefficient	Computed value	n	Statistic Z	Z <sub>0.025</sub>	Significant ?
r <sub>1</sub>	0.132	15768	16,639	1.96	yes
r <sub>2</sub>	0.103	15386	12,780	1.96	yes
r <sub>3</sub>	0.095	15004	11,637	1.96	yes

From the conducted calculations implied, that the value of computed coefficients is almost the same for all studied nets and we have  $r_1 \approx 0.14$ ,  $r_2 \approx 0.12$  and  $r_3 \approx 0.09$  (see Fig. 3-2). H. Lucht received for the Lower Saxony the following values of coefficients;  $r_1 \approx 0.10$ ,  $r_2 \approx 0.09$  (Lucht, 1972).



Fig. 3-2 Comparison of the computed correlation coefficients.

Reliability of coefficients computed in the present study getting better with the increase of the size of the population.

### 3.2 Weight computation

Having computed for the given levelling network correlation coefficients one can calculate the variance of individual lines, what finally will let estimate the weights of these lines.

In the paper (Lucht, 1972) it was shown that variance  $\sigma_{_{\Delta H}}^2$  of dependent observations is

$$\sigma_{\Delta H}^{2} = \sigma_{l}^{2} L \left( l + \frac{2\sum r_{l}}{n} \right)$$
(3-12)

where  $r_t$  are the succeeded correlation coefficients of the neighboring observations in given line, n is the number of observations and L is the length of the line.

If we assume that the network of the precise levelling measured in fourth campaign is characterizes by the following coefficients of the correlation:  $r_1 = 0,132$ ,  $r_2 = 0,103$ ,  $r_3 = 0,095$  (see Table 3-4) then e.g. for the line Kuźnica - Sokółka consist of 19 sections and the total length 18.56 km and standard deviation  $\sigma_1 = 0.278$ , the expression  $\left(\sum_{i=1}^{n} \sum_{j=1}^{n} c_{j}\right)$ 

 $\begin{pmatrix} 1 + \sum_{n} r \\ n \end{pmatrix}$  is equal 1.017 which at last gives

$$\sigma_{AH} = 1.208 \text{ mm}/\sqrt{\text{km}}$$

# 4 The alternative calculation of the matrix $C_{su}$

It is possible calculation of the elements of the covariance matrix  $C_{sH}$  in the following way (Vaniček and Grafarend, 1980).

Given the estimation of the average value  $\sigma_{M}/\sigma_{I}$  for the levelling network of a certain region

and together with the average length of the levelling line  $L_s$ . This estimation reflects general properties of this region e.g. the climate. From the equation (1-3) one can compute the value of the parameter  $\alpha$ .

$$\alpha = \frac{ln\left(\frac{\sigma_{_{AH}}}{\sigma_{_{I}}}\right)}{lnL_{_{s}}}$$
(4-1)

After that from the same equation (1-3) the variance of the height difference of any line  $L_i$  can be computed from the formula

$$\sigma_i^2 = \left(\sigma_I S_i^{\alpha}\right)^2 \tag{4-2}$$

# 4.1 Accuracy estimation of the levelling network

Usually the accuracy of the large precise levelling networks is characterized on the basis of misclosures  $\rho$  from the double levelling of the section (Jordan at al., 1956, pp. 223-255)

$$m_I^2 = \frac{1}{4n_r} \sum \frac{\rho^2}{r}$$
(4-3)

and on the basis of misclosures  $\lambda$  of the double levelling of a line (ibid.).

$$m_2^2 \frac{l}{4n_L} \sum \frac{\lambda^2}{L}$$
(4-4)

where *r* is a length of a section, *L* is a length of a line,  $n_r$  is a number of the sections,  $n_L$  is a number of the lines. One can also estimate the accuracy of the levelling network on the basis of the misclosures  $\varphi$  of the levelling loops from the formula

$$m_{3}^{2} = \frac{l}{n_{F}} \sum \frac{\varphi^{2}}{F}$$

$$(4-5)$$

where F is circumference of the loop in km. These errors can be estimated before network adjustment.

In a case of the precise levelling network measured in 1999 - 2003 the mentioned mean errors are as follows:  $m_1 = \pm 0.278 \text{ mm}/\sqrt{\text{km}}$ ,  $m_2 = \pm 0.518 \text{ mm}/\sqrt{\text{km}}$ ,  $m_3 = \pm 0.826 \text{ mm}/\sqrt{\text{km}}$ .

In order to estimate the parameter  $\alpha$  in the equation (1-3) it was assumed that  $\sigma_1 = 0.278 \text{ mm}/\sqrt{\text{km}}$  and for the average length of the line 46 km, it was computed that  $\sigma_{\Delta H} = 0.518\sqrt{46} = 3.515 \text{ mm}/\sqrt{\text{km}}$ 

what gives value of the fraction  $\sigma_{\Delta H} / \sigma_{I}$  equal 12.656. On the basic of this value and assuming the average length of the line 46 km from the equation

(4-1) parameter  $\alpha$  equal 0.663 was computed.

For example for the line Kuźnica - Sokółka 18.56 km long the variance of the height difference is  $\sigma_{_{AH}}$  =0.278 18.56<sup>0.663</sup> =1.928 mm/ $\sqrt{\text{km}}$ , while in the case of traditional weights (lack of the correlation) we have  $\sigma_{_{AH}}$  =0.278 18.56<sup>0.5</sup>=1.198 mm/ $\sqrt{\text{km}}$ .

### **5** Network adjustment

The precise levelling network (Fig. 3-1) contains 382 observations (height differences of the line) and 244 unknowns (nodal benchmark heights). It was assumed that the height of one nodal benchmark (Warsaw) is known. The network was adjusted in three variants. The traditional way of weights computation was accepted in the first case

$$p_i = \frac{l}{L_i} \tag{5-1}$$

In the second case the weights were computed according to the formula

$$p_i = \frac{l}{L_i \left( l + \frac{2\sum r}{n} \right)}$$
(5-2)

where n jest is a number of section in a given line. In the third case the weights were computed from

$$p_{i} = \frac{I}{L_{i}^{2\times 0.663}}$$
(5-3)



Fig. 5-1 Comparison of weights computed in different way for the levelling line Kuźnica - Sokółka.

In order to compare the weights computed in three different ways the weights for a test line Kuźnica - Sokółka were carried out and displayed on Fig. 5-1.

From comparison shown on this figure is seen that the first approach (traditional) and the approach given by the formula (5-2) give almost the same value. Weight computed by formula (5-3) gives significantly higher values than previous one.

The whole levelling network (Fig. 3-1) was adjusted with Geolab (version 2001.9.20.0), which created in 1985, and currently requires Windows system (Łyszkowicz and Jackiewicz, 2005).

Finally the first variant was compared with the second and third variant. Differences between them are show in Table 5-1, Fig. 5-2 and Fig. 5-3.

Table 5-1 Statistics between the succeed variants of network adjustment (mm)

	mean	Standard deviation	min	max
(1)-(2)	0.06	0.29	-1.00	1.00
(1)-(3)	-0.63	0.66	-2.00	1.00

In general it can be said that such small correlation of the order 0.1 does not produce any significant results in levelling network adjustment. Differences of the order of 2 mm occur very rarely.

Obtained results are quite different from the results obtained by H. Lucht, who having the same correlation found differenced even 50 mm (Lucht, 1983, p.324).



Fig. 5-2 Histogram of differences of adjusted height, variant 1 minus variant 2



Fig. 5-3 Histogram of differences of adjusted height, variant 1 minus variant 3

## **6** Conclusions

From conducted investigations it is seen that, the neighboring height differences  $\delta H$  in the levelling network are correlated. The degree of correlation is rather weak and the first three coefficients have value:  $r_1 = 0.132$ ,  $r_2 = 0.103$ ,  $r_3 = 0.095$  (see Table 3-4). The almost identical value of coefficients was received in works (Lucht, 1972), (Dymowski, 1978).

The influence of computed correlation coefficients on the results of the adjustment of the levelling network is negligible (see Figs. 5-2). This conclusion this does not agree with results obtained by H. Lucht (Lucht, 1972).

Alternative way of calculation of the correlation in levelling network proposed by (Vaniček and Grafarend, 1980) shows the weak influence of the correlation on the results of the network adjustment.

#### 7 Bibliography

- Dymowski K., 1973, Analiza dokładności sieci niwelacyjnych lokalnego znaczenia przy odrzuceniu założenia o niezależności obserwacji(Accuracy analysis of the local levelling network of correlated observations) , Ph.D. thesis, Kraków
- Harvey B.R., 1990, *Practical east squares and statistics for surveyors*, Monograph 13, The University of New South Wales, Kensington, Australia
- Jordan, Eggert, Kneissl, 1956, Handbuch der Vermessungskunde, Band III, Höhenmessung, Stuttgart
- Lucht H.,1972, Korrelation im Prazisionsnivellement, Wiss. Arb. Der Lehrstuhle fur Geodasie, Photogrammetrie und Kartographie an der Tech. Univ. Hannover, N 48
- Lucht H., 1983, *Neighboring correlations among observation in levelling networks*, In: H. Pelzer and W. Niemeier (Eds.): Precise Levelling, 38 Contribution to the Workshop on Precise Levelling, held in Hannover, Fed. Ref. of Germany, March 16 – 18,

- Łyszkowicz A., A. Jackiewicz, 2005, Adjustment of Polish precise levelling network using Geolab package, Proceedings of the 6th International Conference "Environmental Engineering", May 26-27, Vilnius, Lithuania, pp. 939-944
- Łyszkowicz A., M. Leończyk, 2006, The fourth precise levelling campaign of Poland in 1999 – 2003, Proceedings of the 5th FIG Regional Conference, Accra, Ghana, March 8-11, 2006, pp.TS19/1-14
- Remmer O., 1975, *Levelling errors in Stau* Nascendi, Geodaetisk Instytut, Kobenhavn
- Vaniček P., E.W. Grafarend, 1980, *On the weight estimation in levelling*, NOAA Technical Report NOS 86 NGS 17
- Wyrzykowski T.,1988, Monografia krajowych sieci niwelacji precyzyjnej I klasy(Monography of Polish precise levelling networks), Instytut Geodezji i Kartografii, Warszawa 1988