Research Article

Vasil Cvetkov*

Two adjustments of the second levelling of Finland by using nonconventional weights

https://doi.org/10.1515/jogs-2022-0148
received May 19, 2022; accepted December 7, 2022

Abstract: Despite being in use for more than 150 years, the error accumulation in precise levelling has not yet been completely clarified. It is believed that the error accumulation in this method is proportional to the square root of the levelling length. The first goal of this article is to demonstrate that this belief is not always scientifically proven. The second aim is to show that it is likely that a better adjustment decision will be missed if inverse distance weighting with a power parameter equal to one is automatically applied. Using linear regression analysis the measuring data of the Second Levelling of Finland is analysed. An inadequacy of the relationship between the absolute values of the differences between both measurements of the elevations in the levelling lines and their length is shown, which is due to heteroscedasticity. In order to obtain a homoscedastic model, the other two models are constructed. Based on the regression analysis results, the network is adjusted using three types of weights. The adjustment with traditional weights has produced significantly greater mean errors of the nodal benchmarks than both variants based on weights, which are functions of the absolute values of the line elevations.

Keywords: accuracy, elevation differences, error propagation, regression analysis

1 Introduction

There are different approaches for determining the vertical position of a single point or a group of points in geodetic practice. One of them is precise levelling. Due to its high accuracy, this method has been used since the last decades of the nineteenth century for investigation of the recent land uplifts of the Earth's crust (Kääriäinen 1953; Kowalczyk and Rapinski 2013). Other geodetic activities where precise levelling takes place are setting of regional or continental vertical reference systems (Adam et al. 1999; Sacher et al. 2008), estimating vertical datum offsets from Global Geopotential Models (Hayden et al. 2012; Kotsakis and Katsambalos 2010; Li 2018), intercontinental height datum relations (Gruber et al. 2012), and search of active geodynamic activities (Lehmuskoski 1996). Thus, this method is essential for various scientific studies. Precise levelling is also one of the most preferred methods in applied geodesy for registration of vertical movements of buildings, dam walls, *etc.* (Angelov 2017).

Despite the development of digital levels, modern rods and rod comparators (Takalo 1997, 1999) and prior optimization of the measurement process (Takalo 1978), it seems that precise levelling has reached the end of its development as a measuring method. For example, the total mean error of the Third Levelling in Finland, calculated from the closing errors of the levelling loops, is $\pm 0.86\,\mathrm{mm}/\sqrt{\mathrm{km}}$ (Saaranen et al. 2021). The corresponding accuracy of the Second Levelling is $\pm 0.60\,\mathrm{mm}/\sqrt{\mathrm{km}}$ (Kääriäinen 1966).

Identical results have been announced for the Polish precise levelling campaigns (Kowalczyk and Rapinski 2013).

Analysing the mean errors obtained from the adjustments of the precise levelling networks of the participants of the United European Levelling Network (UELN) (Sacher et al. 2008), one can see that only the Hungarian network has less mean error per kilometre than the Second Levelling of Finland, $\pm 0.47\,\mathrm{kgal\,mm}/\sqrt{\mathrm{km}}$ against $\pm 0.63\,\mathrm{kgal\,mm}/\sqrt{\mathrm{km}}$ (Kääriäinen 1966), respectively. All of these networks were measured and adjusted decades after the Second Levelling of Finland. Therefore, if we want to develop precise levelling as a method, we should ask ourselves whether we have missed something about the accumulation of errors in this measuring approach.

Looking at the study by Sacher et al. (2008, Table 2), one can divide presented countries by the yielded mean errors of their networks into two groups. In the first one,

* **Corresponding author: Vasil Cvetkov,** Department of Geodesy, Tochnost Ltd, Stephan Stambolov 50, Bankya, Sofia-city, 1320, Bulgaria, e-mail: tzvetkov_vasil@abv.bg

where the mean errors are less than 1.00 kgal mm/$\sqrt{\text{km}}$, we can include Austria, Finland, Hungary, Poland, Germany, Denmark, Netherlands, Lithuania and the common net of Croatia, Slovenia, Bosnia and Hercegovina. Then, the second group includes Belgium, Switzerland, Spain, France, Italy, Portugal, Great Britain, Norway, Sweden, Czech Republic, Slovakia, Romania, Estonia, Latvia and Bulgaria.

It is obvious that both groups contain countries from different European regions and subregions which differ from each other based on geographical, cultural, financial or historical criteria. Thus, the variation in the accuracy between both groups cannot be explained by the above-mentioned factors. There are some differences in the measuring methods applied in each country but they are not able to explain fully the gap in the accuracy between both groups. Moreover, they cannot be used as weights in an adjustment.

Suppose we exclude Austria and the common network of Croatia, Slovenia, Bosnia and Hercegovina from the first group. Let us do the same with the networks of Belgium, Estonia and Latvia from the second group. As a result, one can easily see that the terrain of the countries in the first group is flatter than that of the countries in the second group, whose territories are mountains and rougher terrain areas. Consequently, it is likely that there is a relationship between the accuracy of the networks and the relief of the territory, which they cover. In fact, it has been shown (Cvetkov 2022b) that the accumulation of the absolute values of the discrepancies of the Balkan Mountain crossing lines in the Third Levelling of Bulgaria is more strongly related to the sum of the absolute values of the section elevations along the lines than to the length of the lines. Using simple scatter plots of the discrepancies in the sections, lines and loops against the square root of their length has revealed a presence of heteroscedasticity (Cvetkov 2022a). This fact implies that the popular assumption of the low accumulation of errors in precise levelling, which supposes direct proportion with the square root of levelling distances (Kääriäinen 1953, 1966; Lyszkowicz and Leonczyk 2006), among others, was not met in the Third Levelling of Bulgaria.

The same conclusions were made in the study by Gekov and Tzvetkov (2005a) where the Second Levelling of Bulgaria was analysed and adjusted using different weights. It was shown that the weights inversely proportional to the line lengths produce greater mean errors of the nodal benchmarks than the weights inversely proportional to the square of the sum of the absolute values of the section elevations in the lines.

It will be interesting to investigate whether or not the classic low accumulation of errors in precise levelling is met in a state with a flat area. For this reason, the results of the Second Levelling of Finland will be analysed. Additional reasons for choosing this network are the accessibility of the data (Kääriäinen 1966, Table IX) and the extremely high accuracy of this network.

The questions we will discuss in this article are the following:

- How adequate is the proportional relationship between the absolute differences between both measurements of the elevations between the endpoints in the lines and the line lengths in the Second Levelling of Finland?
- Are there any other weights that better fit the analysed network?

2 Experimental procedures

This section includes detailed information about some models used to explain the accumulation of discrepancies in the levelling lines. Also, the methodology of the adjustment of the network is explained, and two nonconventional weights are presented.

Let us have a levelling line, whose length is s km. If the mean error of the levelling per km is ε, the mean error of the difference in height between the terminals of this line is ε_s then, a relation among them can be given by equation (1) (Kääriäinen 1953, p. 39).

$$\varepsilon_s = \pm\varepsilon\sqrt{s}\,. \tag{1}$$

Thus, the weight of the line in an adjustment is

$$P = p_0/s \cdot \varepsilon^2. \tag{2}$$

Let p_0 be a constant suitably selected. It is better to take ε as a constant in the entire network so $p_0 = \varepsilon^2$. Let $s = L$. Thus, (2) turns in (3).

$$P = 1/L. \tag{3}$$

The mean error ε_s can also be expressed by H_1 and H_2 which are the values of the elevation between the terminals of the line yielded in both measurements.

$$\varepsilon_s = \frac{H_1 + H_2}{2} = \pm\frac{D}{2}. \tag{4}$$

After putting (4) in (1), replacing $2\varepsilon = \text{const.}$ and $s = L$, we derive model (5), which will be examined by regression analysis in Chapter 3.

$$|D| = \text{const}\sqrt{L}\,. \tag{5}$$

Equations (3) and (5) represent the popular assumption of the most appropriate weights in an adjustment of levelling networks and the most correct low of error accumulation in levelling, respectively. The relevance of (5) will be checked by regression analysis. Residuals derived from the analysis will be examined whether they meet the basic assumptions of the ordinary least-squared regression analysis. The relevance of (3) will be estimated by using paired t-tests with two samples for means. The first sample will contain the mean errors of the nodal benchmarks yielded by the use of weights (3), and the other samples will be generated from the mean errors of the correspondent nodal benchmarks derived from adjustments using nonconventional weights.

2.1 Other models

It is a known fact that the levelling results obtained on routes with long and high inclinations are more affected by different kinds of systematic errors, e.g. refraction, inclination of the rods during the measurements, errors in the rod meter due to calibration or temperature changes (Enman and Enman 1984), and so on. The levelling along inclined routes also increases the number of setups. As a result, the effect of vertical movements of the tripod and rods is getting bigger. Also, the influence of the errors due to vertical movements of the common station rod is amplified. Consequently, as was said in Chapter 1, the accumulation of the discrepancies in precise levelling should be better explained as a function of measured elevations. So, the relationship between the absolute values of the differences $|D|$ and the absolute values of the elevations between the terminal points of the lines $|H|$ seems to be reasonable. Thus, we can define the second model of our investigation (6).

$$|D| = \mathrm{const}|H|. \tag{6}$$

The relation given by (7) is also going to be analysed.

$$|D| = \mathrm{const}\sqrt{|H|}. \tag{7}$$

In order to increase the statistical power of our analysis, all lines included in (Kääriäinen 1966, Table IX) will be used.

2.2 Nonconventional weights and the adjustment procedure

Our aim is to be compared the results derived from adjustments of the Second Levelling of Finland network using weights given by equation (3) and both nontraditional weights (8) and (9).

$$P = 1/|H|^2, \tag{8}$$

$$P = 1/|H|. \tag{9}$$

The configuration of the network can be found in the study by Kääriäinen (1966, p. 11). There are some slight differences between this study and the original one, namely:

- Loops IV and V are not included in the adjustment because of the fact that the entire information for lines 13, 14 and 17 was not published (Kääriäinen 1966, Table IX).
- In order to obtain directly the weight coefficients Q_{ii} of the nodal benchmarks, the parametric adjustments are used here. The condition taken by Kääriäinen (1953, p. 54) and the treatment of the weight coefficients described by (Kääriäinen 1953, p. 55–56) are also included in our calculations.
- The connections to the tide gauges are not included in our adjustment. For this reason, a fundamental benchmark is chosen the benchmark in Qulunkylä instead of the Fundamental Bench Mark in Helsinki.
- Because of the simplicity, some lines between nodal bench marks are combined. For example, lines 31, 32 and 33, which take part in loop X, are joined in a common line. Lines 35, 36 and 37 are treated in the same manner and so on. As a result, our network contains 26 nodal bench marks, where the nodal benchmark in Qulunkylä has a known height and other bench marks have unknown heights. Finally, the result network is consisted of 41 levelling lines.

In order to clarify utterly the adjustment procedure, a brief explanation is given below. Let v_{ij} are the corrections in the measured heights h_{ij} between bench marks i and j, and their initial heights are H_i and H_j, respectively. Let x_i and x_j are the corrections of H_i and H_j. Then, our correction equations can be written as (10).

$$\begin{aligned} V_{ij} &= (H_i + x_i) - (H_j + x_j) - h_{ij} \\ &= x_i - x_j + (H_i - H_j - h_{ij}) \\ &= x_i - x_j + f_{ij}. \end{aligned} \tag{10}$$

Thus, in matrix form (10) can be presented by (11).

$$V = AX + f. \tag{11}$$

In equation (11), A is an information matrix, which contains the coefficients forward the unknowns x_i, X is a vector of the unknowns x and f is a vector of the free members in (10). Using the above symbols, we can write our additional condition (12) and in matrix form (13).

$$\sum x_i = 0, \tag{12}$$

$$BX = 0. \tag{13}$$

In (13), B is a size 25 vector of ones. Now, our aim is to obtain the unknowns x_i in accordance with conditions (13) and (14).

$$V^T PV \rightarrow \min. \tag{14}$$

In (14), P is a matrix of the weights. All members of the P matrix are zeros except the members in the main diagonal, which are values of the weights of each levelling line calculated by (3), (8), or (9). To obtain the corrections v_{ij} and x_i, we use equation (15), where K is a vector of correlates. In our adjustment, we have only one correlate.

$$Q = V^T PV + 2K^T BX. \tag{15}$$

After substitution into Lagrange's equation $\mathrm{d}Q/\mathrm{d}V = 0$ and some matrix manipulations, we yield (16).

$$A^T PAX + B^T K + A^T Pf = 0. \tag{16}$$

Let $N = A^T PA$ is a normal matrix and $A^T Pf = F$. Thus, we yield our extended normal system (17) or (18), and after its solution, we can obtain the unknown corrections x_i.

$$\begin{vmatrix} N & B^T \\ B & 0 \end{vmatrix} . \begin{vmatrix} X \\ K \end{vmatrix} + \begin{vmatrix} F \\ 0 \end{vmatrix} = 0, \tag{17}$$

$$N_e . X_e + F_e = 0. \tag{18}$$

Using equation (11), one can yield the corrections v_{ij}. The mean error of the weight unit m can be calculated by (19), where $r = n - k = 41 - 25 = 16$.

$$m^2 = \frac{[PVV]}{r}. \tag{19}$$

The matrix $Q = N_e^{-1}$ is an extended covariance matrix. The first $k = 25$ values in the main diagonal of Q are the inverse values of the nodal benchmark weights. These values are treated as described by Kääriäinen (1953, p. 55–56) in order to receive the final values Q'_{ii}. Finally, the mean errors of the nodal benchmarks can be computed by (20).

$$m_i = m \cdot \sqrt{Q'_{ij}}. \tag{20}$$

All heights, which are used in the adjustment, are corrected with a rod meter correction, a refraction correction and a land uplift correction to the epoch 1944.

3 Results

Using Ordinary Least Squared Regression, the presence of heteroscedasticity concerning the classic model (5) is demonstrated in this chapter. It is shown that both nonconventional models given by equations (6) and (7) lead to homoscedastic results. In the second part of this section, the results derived from the adjustments with classic and nonconventional weights are presented.

3.1 Regression analysis results

The results derived from the regression analysis of the models defined by equations (5), (6) and (7) are given in Table 1.

The residual plots of the investigated models (5), (6) and (7) are pictured in Figures 1–3, respectively.

3.2 Adjustment results

Figure 4 shows the mean errors of the nodal benchmarks yielded from the adjustments of the Second

Table 1: Regression analysis results

Results	$\|D\| = \text{const}\sqrt{L}$	$\|D\| = \text{const}\|H\|$	$\|D\| = \text{const}\sqrt{\|H\|}$
R	0.764	0.617	0.667
R^2	0.584	0.381	0.445
Adjusted R^2	0.577	0.374	0.438
Standard error	5.137 mm	6.264 mm	5.930 mm
Observations	146	146	146
F	203.269	89.192	116.324
Significance	2.57×10^{-29}	8.97×10^{-17}	3.01×10^{-20}
Constant	$0.9594\ \text{mm}/\sqrt{\text{km}}$	0.1398 mm/m	$1.0437\ \text{mm}/\sqrt{\text{m}}$

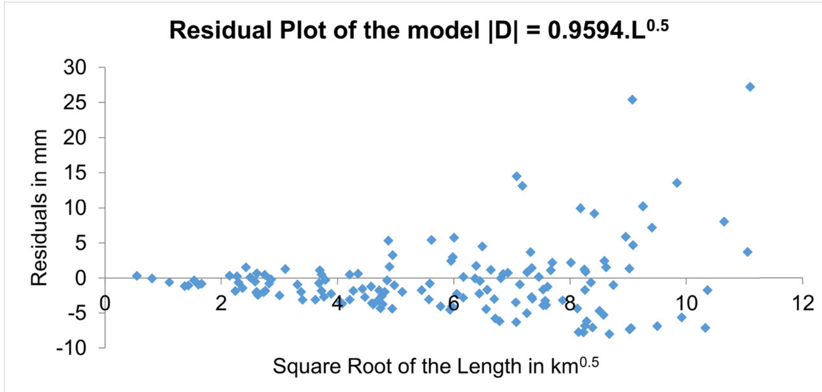

Figure 1: Heteroscedastic pattern of the residuals of the regression model $|D| = 0.9594{\cdot}\sqrt{L}$, where $|D|$ denotes the absolute difference between both measurements of the height between the terminals of a line with length L. The regression model is based on 146 lines given by Kääriäinen (1966, Table IX).

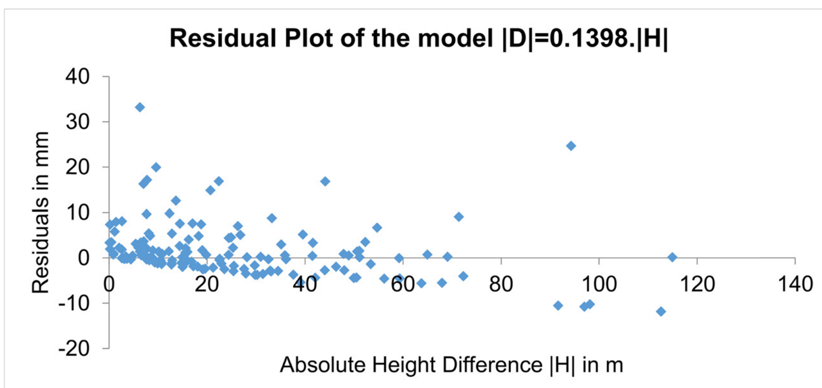

Figure 2: Pattern of the residuals of the regression model $|D| = 0.1398|H|$, where $|D|$ denotes the absolute difference between both measurements of the absolute value of the height $|H|$ between the terminals of a line. The regression model is based on 146 lines given by Kääriäinen (1966, Table IX).

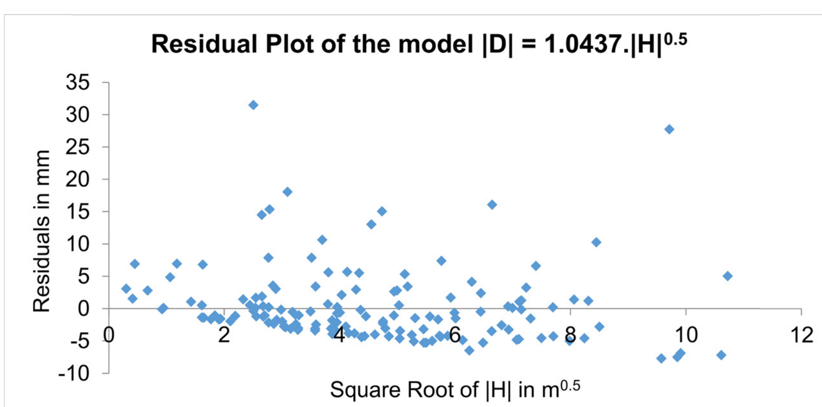

Figure 3: Pattern of the residuals of the regression model $|D| = 1.0437{\cdot}\sqrt{|H|}$, where $|D|$ denotes the absolute difference between both measurements of the absolute value of the height $|H|$ between the terminals of a line. The regression model is based on 146 lines given by Kääriäinen (1966, Table IX).

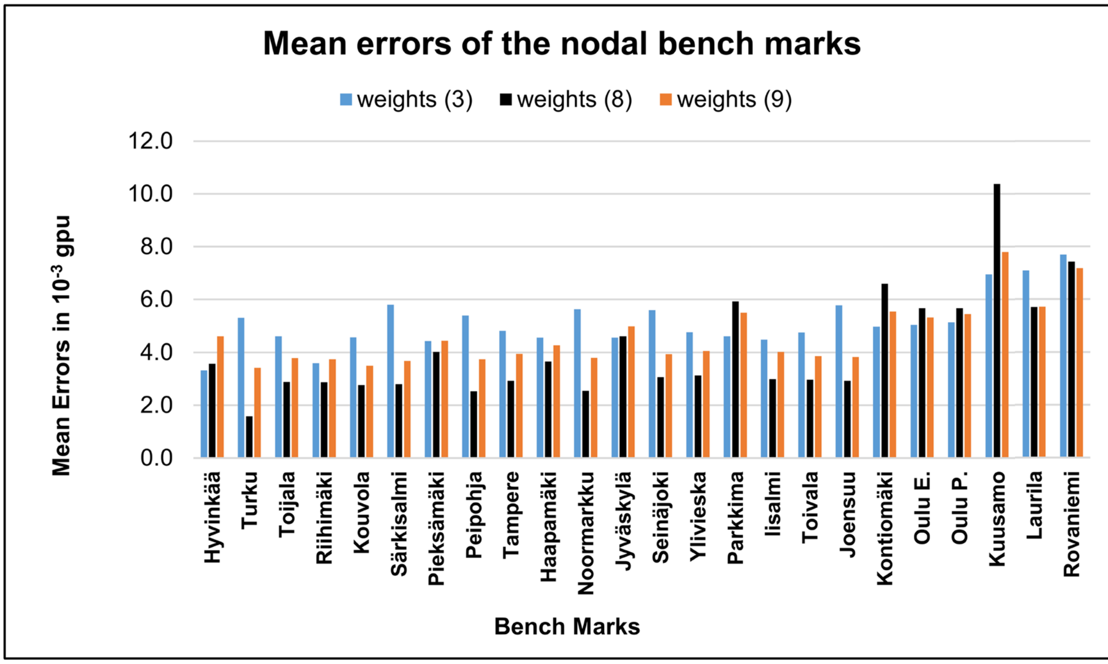

Figure 4: Mean errors of the nodal benchmarks derived from the adjustments with weights (3) – $P = 1/L$, weights (8) – $P = 1/|H|^2$, and weights (9) – $P = 1/|H|$.

Levelling of Finland by using the classic and nonconventional weights.

4 Discussion

One of the goals of this article was to demonstrate the inadequacy of the classic assumption concerning the accumulation of the discrepancies in precise levelling and consequently to raise a question about the choice of the most suitable weights used in the adjustment of precise levelling networks. Since Lallemand's times, it has been believed that the accumulation of errors in levelling is a function of the levelling distance. This belief supposes that the longer the levelling line, the greater the elevation error. The higher the circumference of the loop, the greater the closing error. Is this statement always true? Of course not. If we have a look at the study by Kääriäinen (1966, Table VI, p. 84), we will see that there are loops in the First and the Second Levellings in Finland whose closing errors are greater than the closing error of the circumference loop even though their lengths are 7–9 times less than the length of the circumference loop. According to Cvetkov (2022a), the closing error of the ring loop in the Third Levelling of Bulgaria is only 2.85 mm but the length of the loop is 2550.18 km. Only one loop in the Bulgarian network has less closing error even though the average circumference of the loops is approximately 300 km. A slight correlation between the closing errors and the loop lengths in the Fourth Levelling in Poland is also illustrated in Figure 8 (Lyszkowicz and Leonczyk 2006). It is likely that there will be the same situation in many other regional or national precise levelling networks in the world. If closing errors, which are true errors, are slightly correlated with the perimeters of loops why we have been explaining the accuracy of precise levelling as a function of the levelling length?

Analysing the residuals, derived from the regression analysis of the model defined by equation (5), one can see a presence of heteroscedasticity. The residuals pictured in Figure 1 are fan-shaped, which is evidence of unequal variance of the residuals. This fact shows inadequacy of the classic model (Rawlings et al. 1998) in the case of the Second Levelling in Finland. The inadequacy of this model is also shown for the Polish Fourth Campaign (Lyszkowicz and Leonczyk 2006, Figure 7), the Bulgarian Third Levelling (Cvetkov 2022a, Figures 1–4) and the Second Levelling in Bulgaria (Gekov and Tzvetkov 2005b, Application). Perhaps, surprisingly, the residuals pictured in Figure 6 (Kotsakis and Katsambalos 2010) also reveal a presence of heteroscedasticity when differences in elevations are tried to be described by the square root of the levelling length.

Figures 2 and 3 show that the residuals of models (6) and (7) are more randomly distributed, and the presence of heteroscedasticity is more difficultly detected. Figures 2 and 3 are closer to Figure 11.2 than to Figure 11.3 in the study by Rawlings et al. (1998) where the expected pattern on regression residuals is illustrated. In spite of the fact that the standard error of model (5) is less than the standard errors of the other two models, this model should be revised due to the fact that the heteroscedasticity invalidates its results.

Figure 4 illustrates the mean errors of the nodal bench marks. Comparison between the results derived from using weights (3) and the published mean errors of the corresponding benchmarks (Kääriäinen 1966, Table VIII, p. 88), one can see that they are almost equal. The differences in the values can be explained by the differences between both adjustments the current and the original one, which are described in Chapter 2.2. What is more important is the fact that both adjustments with the use of the nonconventional weights produce less mean errors of the nodal benchmarks. This fact is another evidence of the irrelevance of model (5) and weights (3).

Owing to the fact that the mean error of unit weight in our adjustments is dimensionally different and for this reason incomparable, the samples of the mean errors derived in the separate adjustments are compared using *t*-test: Paired Samples for Means. The results are given in Tables 2–4. According to these tables, the mean errors yielded from the adjustment with the weights given by equation (3) are significantly greater than the mean errors obtained from the adjustments with both weights (8) and (9). Both tests are significant at a 99% confidence level. One can compute that weights (8) produce on average 27% less mean errors than the classical ones. Also, weights (9) lead to 22% smaller mean errors than those produced with weights (3). According to Table 4, there is a significant difference between the means of the mean errors of the nodal benchmarks yielded by the use of weights (8) and (9). Thus, weights (8) produce the least mean errors referring to our three adjustments.

It is important to be remarked that the median of the mean errors obtained by the weights (8) is 3 mm and only one mean error slightly exceeds 10 mm. Analysing Figure 4, one can count that 18 out of 25 nodal benchmarks have mean errors less than 5 mm in the case of the adjustment with weights (8).

The median standard deviation derived from the adjustment with weights (9) is approximately 4 mm, and only 7 of 25 nodal benchmarks have mean errors greater than 5 mm. The greatest mean error is below 10 mm.

Table 2: *t*-Test: Paired Two Sample for Means using the samples of the mean errors of the benchmark derived from the adjustments based on weights (3) and (8)

Description	$P = 1/L$	$P = 1/\lvert H\rvert^2$
Mean	5.101×10^{-3} gpu	4.014×10^{-3} gpu
Median	4.810×10^{-3} gpu	3.070×10^{-3} gpu
Variance	1.024×10^{-3} gpu^2	4.240×10^{-3} gpu^2
Observations	25	25
Pearson correlation	0.535	
df	24	
t Stat	3.122	
$P(T \le t)$ one-tail	0.002	
t Critical one-tail	1.711	
$P(T \le t)$ two-tail	0.005	
t Critical two-tail	2.064	

Table 3: *t*-Test: Paired Two Sample for Means using the samples of the mean errors of the benchmark derived from the adjustments based on weights (3) and (9)

Description	$P = 1/L$	$P = 1/\lvert H\rvert$
Mean	5.101×10^{-3} gpu	4.514×10^{-3} gpu
Median	4.810×10^{-3} gpu	4.020×10^{-3} gpu
Variance	1.024×10^{-3} gpu^2	1.393×10^{-3} gpu^2
Observations	25	25
Pearson correlation	0.590	
df	24	
t Stat	2.925	
$P(T \le t)$ one-tail	0.0037	
t Critical one-tail	1.711	
$P(T \le t)$ two-tail	0.0074	
t Critical two-tail	2.064	

Table 4: *t*-Test: Paired Two Sample for Means using the samples of the mean errors of the benchmark derived from the adjustments based on weights (8) and (9)

Description	$P = 1/\lvert H\rvert^2$	$P = 1/\lvert H\rvert$
Mean	4.014×10^{-3} gpu	4.514×10^{-3} gpu
Median	3.070×10^{-3} gpu	4.020×10^{-3} gpu
Variance	4.240×10^{-3} gpu^2	1.393×10^{-3} gpu^2
Observations	25	25
Pearson correlation	0.980	
df	24	
t Stat	−2.680	
$P(T \le t)$ one-tail	0.0065	
t Critical one-tail	1.711	
$P(T \le t)$ two-tail	0.0131	
t Critical two-tail	2.064	

5 Conclusions

In this article, linear regression analyses based on various predefined models were used in order to analyse the low of the accumulation of the discrepancies in the levelling lines in the Second Levelling of Finland. Based on the results of these analyses, two nonconventional types of weights were proposed. The levelling network was adjusted using these nontraditional weights and results derived from the adjustments were compared with the results obtained from the network adjustment using traditional weights (3). On the basis of the results, the following conclusions can be made:

- An inadequacy of the relationship between the absolute differences $|D|$ produced from both measurements of the heights between the terminals of the lines on the one hand and the square root of the length of lines L, on the other hand, is detected in the case of the Second Levelling of Finland. Such inadequacy was also illustrated in the case of the Third Levelling of Bulgaria (Cvetkov 2022c). This inadequacy could explain the differences among the levelling accuracies calculated on the basis of the discrepancies in sections, lines and loops.
- The models used to explain the accumulation of the absolute differences $|D|$ by the absolute value of the height between the terminals of the lines lead to more homoscedastic results.
- The correlation between $|D|$ and $\sqrt{L}$ is greater than the correlation between $|D|$ and $|H|$, but this fact does not automatically lead to a better fitting in an adjustment with weights inversely proportional to lengths L in comparison with another adjustment of the same networks with the use of weights inversely proportional to the absolute heights $|H|$.
- Nonconventional weight models (8) and (9) lead to approximately 20% smaller mean errors of the nodal benchmarks in the network of the Second Levelling of Finland than the most popular weights (3). This difference in accuracy is significant at a 99% confidence level. The last fact implies that it is always possible to find a better weighting model than the classic one.
- It is likely that independent analyses of the highest-order levelling networks of other states will confirm the above reveals.

Acknowledgments: The author thanks reviewer 6 (Dr. Xiaopeng Li at Nation Oceanic and Atmospheric Administration) and the other anonymous reviewers for their useful comments and suggestions, which resulted in significant improvements in this paper.

Conflict of interest: Author states no conflict of interest.

Data availability statement: The data were published in publication N61 of the Finnish Geodetic Institute – “The Second Levelling of Finland in 1935–1955” by Erkki Kaariainen and is available on http://hdl.handle.net/10138/347072.

References

Adam, J., W. Augath, F. Brouwer, G. Engelhardt, W. Gurtner, B. G. Harsson, et al. 1999. “Status and development of the european height systems.” In *Geodesy beyond 2000: The challenges of the first decade*. IAG General Assembly, Birmingam. July 19 – 30 1999.

Angelov, A. 2017. Geodetic methods in the study for deformation process of high buildings and engineering facilities, Monographic, 1 (uacg.bg), (in Bulgarian)

Cvetkov, V. 2022a. *Discrepancies and misclosures in the third levelling of Bulgaria*, Vol. 55, No. 1, p. 195–201. Annual of the University of Architecture, Civil Engineering and Geodesy, Sofia. 14-G-3.pdf (uacg.bg).

Cvetkov, V. 2022b. *Accumulation of the discrepancies [D] along the Balkan Mountains crossing lines 26, 31, 35, 37 in the Third Levelling of Bulgaria*, Vol. 55, No. 1, p. 203–11. Annual of the University of Architecture, Civil Engineering and Geodesy, Sofia. 15-G-4.pdf (uacg.bg).

Cvetkov, V. 2022c. *An attempt at explaining of the absolute discrepancies |D| accumulated between the fundamental bench marks of the First and the Second order in the Third Levelling of Bulgaria by some regression models*, Vol. 55, No. 2, p. 275–82. Annual of the University of Architecture, Civil Engineering and Geodesy, Sofia. 4-G1.pdf (uacg.bg).

Enman, S. and V. Enman. 1984. “Systematic errors in leveling of mountainous areas.” *Bulletin Géodésique* 58, 475–93.

Gekov, D. and V. Tzvetkov. 2005a. “The problem about choice of weights in the adjustment of precise leveling networks.” In *Proceedings of the International Symposium on Modern Technologies, Education and Professional Practice in Geodesy and Related Fields*, p. 290–6, Sofia (in Bulgarian).

Gekov, D. and V. Tzvetkov. 2005b. “Limits for differences in values of forward and backward heights between nodal bench marks in levelling lines.” In *Proceedings of the International Symposium on Modern Technologies, Education and Professional Practice in Geodesy and Related Fields*, p. 297–304, Sofia (in Bulgarian).

Gruber, T., C. Gerlach, and R. Haagmans. 2012, “Intercontinental height datum connection with GOCE and GPS-levelling data.” *Journal of Geodetic Science* 2, 4, 270–80, doi: 10.2478/v10156-012-0001-y.

Hayden, T., B. Amjadiparvar, E. Rangelova, and M.G. Sideris. 2012, “Estimating Canadian vertical datum offsets using GNSS/levelling bench mark information and GOCE global geopotential model.” *Journal of Geodetic Science* 2, 4, 257–69. doi: 10.2478/v10156-012-0008-4.

Kääriäinen, E. 1953. *On the recent uplift of the Earth's crust in Finland*, Publications of the Finnish Geodetic Institute No. 42, Helsinki.

Kääriäinen, E. 1966. *The Second Levelling of Finland in 1935–1955*. Publications of the Finnish Geodetic Institute No. 61, Helsinki.

Kotsakis, C. and K. Katsambalos. 2010. "Quality analysis of global geopotential models at 1542 GPS/levelling bench marks over the Hellenic Mainland" *Survey Review* 42, 318, 327–44. doi: 10.1179/003962610X12747001420500.

Kowalczyk, K. and J. Rapinski. 2013. "Evaluation of levelling data for use in vertical crustal movements model in Poland." *Acta Geodynamica et Geomaterialia*, 10(4, 172), 401–10. doi: 10.13168/AGG.2013.0039.

Lehmuskoski, P. 1996. *Active fault line search in southern and central Finland with precise levellings*, Reports of the Finnish Geodetic Institute No. 96, 5.

Li, X. 2018. "Modeling the North American vertical datum of 1988 errors in the conterminous United States" *Journal of Geodetic Science* 8, 1, 1–13. doi: 10.1515/jogs-2018-0001.

Lyszkowicz, A. and M. Leonczyk. 2006. "The Fourth Precise Levelling Campaign of Poland in 1999–2003" In *Promoting Land Administration and Good Governance, 5th FIG Regional Conference Accra*, Ghana, 8–11 March 2006, Microsoft Word - ts19_04_lyskowicz_leonczyk.doc (fig.net).

Rawlings, J., S. Pantula and D. Dickey. 1998. *Applied Regression Analysis: A Research Tool*, 2nd edn. New York: Springer.

Saaranen, V., P. Lehmuskoski , M. Takalo, and P. Rouhiainen. 2021. *The Third Precise Levelling of Finland*, FGI Publications No. 161, Kirkkonummi, The third precise levelling of finland (helsinki.fi).

Sacher, M., J. Ihde, G. Liebsch, and J. Mäkinen. 2008. "*EVRF2007 as Realization of the European Vertical Reference System*." Presented at the Symposium of the IAG Sub-commission for Europe (EUREF) in Brussels, June 18–21 2008.

Takalo, M. 1978. *Measuring method for the Third Levelling of Finland*, Reports of the Finnish Geodetic Institute No. 78, 3, Helsinki.

Takalo, M. 1997. *Automated calibration of precise levelling rods*, Reports of the Finnish Geodetic Institute No. 97, 3, Kirkkonummi.

Takalo, M. 1999. *Verification of automated calibration of precise levelling rods in Finland*, Reports of the Finnish Geodetic Institute No. 99, 7, Kirkkonummi.