Baltic J. Modern Computing, Vol. 14 (2026), No. 1, 190-201
https://doi.org/10.22364/bjmc.2026.14.1.10

Uncertainty of a State Vertical Reference System below 1 mgpu – Is It Possible?

Vasil CVETKOV, Christina MICKRENSKA

University of Architecture, Civil Engineering and Geodesy,
Hristo Smirnenski Blvd. 1, Sofia, Bulgaria

tzvetkov_vasil@abv.bg, kmikrenska_fgs@uacg.bg

ORCID 0000-0001-9628-6768, ORCID 0000-0002-7054-1799

Abstract. Precise geometric levelling has been the most accurate method for measuring height differences on Earth for over 150 years. Yet, its accuracy has seen little improvement in the past 80 years. This stagnation is mainly due to the rise of GNSS technologies and a lack of progress in levelling data processing. This article presents new statistical insights into how levelling uncertainties accumulate and introduces a revised algorithm for adjusting levelling networks. Testing the method with data from Finland's Third Levelling (1978–2006) under 66% showed standard deviations of benchmark geopotential numbers below 0.7 mgpu (or 0.7 mm). These results prove that, by rethinking outdated approaches, a fifteenfold increase in accuracy is achievable.

Keywords: Probability Theory, Reference Systems, Geometric Levelling, Vertical Measurements, Adjustment Algorithms

1. Introduction

1.1. Precise levelling classical assumptions

According to Kääriäinen (Kääriäinen, 1966, pp. 52), "the errors are divided into two classes, accidental and systematic, which are assumed to be independent of each other. The accidental errors are caused by sources that are independent in all successive observations and obey the laws of Gauss. The systematic errors are due to causes acting similarly on successive or adjacent levelling observations; these do not obey the laws of Gauss. They become accidental only for distances exceeding a certain limit Z of the order of several tens of kilometres. The total error is the combined influence of both types of error and is calculated for distances longer than the above limit as the quadratic sum of the accidental and systematic errors". This popular belief about the accumulation of levelling uncertainties dominates in many scientific studies (Gerlach and Rummel, 2024; Izvoltova et al., 2024; Kääriäinen, 1966; Lyszkowicz and Leonczyk, 2006; Saaranen et al., 2021; among others) where levelling data is processed. However, it is not supported by some, and likely not by any, of the levelling data.

1.2. Statistical relationship between line discrepancies and their length

According to (Cvetkov, 2024g), the coefficients of correlation *R* of line descripencies $|D|$ to the square root of line length $\sqrt{L}$ are 0.530, 0.456, 0.530, and 0.444, respectively for the Second Levelling of Bulgaria /1953-1957/, the Third Levelling of Bulgaria /1975-1984/, The Second Levelling of Finland /1935-1955/, the Third Levelling of Finland /1978-2006/. According to (Lyszkowicz and Leonczyk, 2006, Figure 7), the coefficient of correlation between line length L and line discrepancies $|D|$ in the case of the Polish Fourth Levelling data (1999-2002) is $R = 0.41$. Taking into account that the correlation between $\sqrt{L}$ and L is usually higher than 0.99, it follows that the coefficient of correlation between the line discrepancies $|D|$ and the square root of line length , for the Polish Fourth network will be in the same magnitude, or $R \approx 0.41$. Thus, for the coefficients of determination R^2 of descripencies $|D|$ to the square root of line length $\sqrt{L}$ for the Second Levelling of Bulgaria /1953-1957/, the Third Levelling of Bulgaria /1975-1984/, The Second Levelling of Finland /1935-1955/, the Third Levelling of Finland /1978-2006/, and the Fourth Levelling of Poland are 0.281, 0.208, 0.281, 0.197 and 0.17. This simply shows that the variations in the line discrepancies *D* can be explained up to 30% by the square root of the levelling line length $\sqrt{L}$. This conclusion is also valid for the explanation of *D* by *L*. A graphical evidence of this fact is the dispersion of line discrepancies |D| concerning $\sqrt{L}$ for the data of the Third Levelling of Finland.

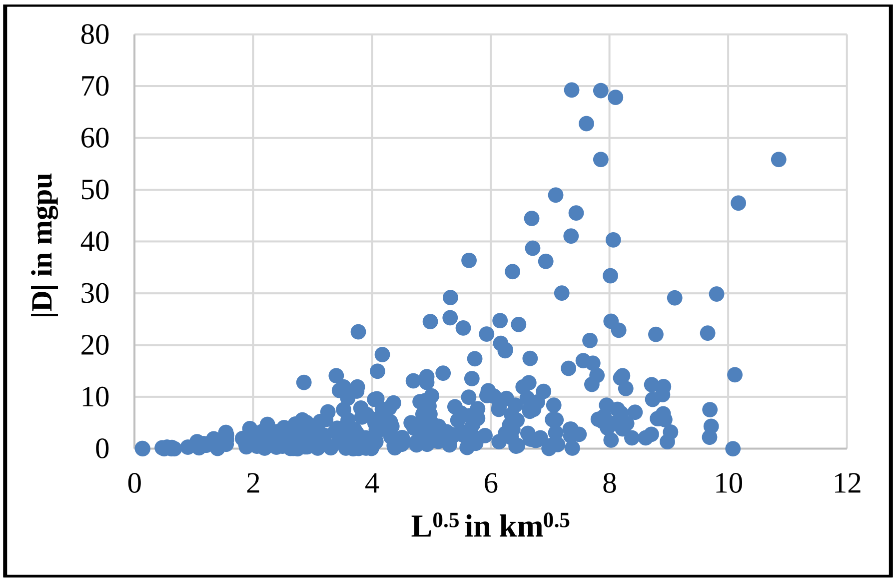

Fig. 1. Dispersion of line discrepancies |D| concerning $\sqrt{L}$ for the data of the Third Levelling of Finland. The data are taken from (Saaranen et al., 2021)

A similar pattern of spread of |D| in respect of $\sqrt{L}$ (*L*) is presented in (Cvetkov, 2023; Cvetkov, 2022; Lyszkowicz and Leonczyk, 2006) for the Second Levelling of Finland, the Third Levelling of Bulgaria, and the Fourth Levelling of Poland. Taking into account the differences in measurement epochs, levelling equipment, methodology of performance, terrain and climate conditions in Finland, Poland and Bulgaria, it will be naive to say that this common pattern of the spread of |D| in respect of $\sqrt{L}$ (L) is incidental. On the contrary, the pattern is natural for precise geometric levelling, and its physical explanation will be given in a future publication.

1.3. Statistical relationship between loop closing errors and loop circumferences

Let us explore the relationship between loop closing errors $|W|$ and loop circumferences L or its square root $\sqrt{L}$. According to (Cvetkov, 2024a; Cvetkov, 2024g), the coefficients of correlation R of closing errors $|W|$ to the square root of line length $\sqrt{L}$ are 0.447, 0.348, 0.234, and -0.006, respectively for the Second Levelling of Bulgaria /1953-1957/, the Third Levelling of Bulgaria /1975-1984/, The Second Levelling of Finland /1935-1955/, the Third Levelling of Finland /1978-2006/. According to (Lyszkowicz and Leonczyk, 2006, Figure 8), the coefficient of correlation between loop circumference L and loop closing errors $|W|$ in the case of the Polish Fourth Levelling data (1999-2002) is R = 0.12. Thus, the coefficient of correlation between the loop closing errors $|W|$ to the square root of the loop circumference, for the Polish Fourth network, will be R ≈ 0.12. Thus, for the coefficients of determination R^2 of $|W|$ to the square root of loop circumference $\sqrt{L}$ for the Second Levelling of Bulgaria /1953-1957/, the Third Levelling of Bulgaria /1975-1984/, The Second Levelling of Finland /1935-1955/, the Third Levelling of Finland /1978-2006/, and the Fourth Levelling of Poland are 0.200, 0.121, 0.055, 0.000 and 0.014. This simply shows that the variations in the closing errors W can be explained up to 20% by the square root of the loop length $\sqrt{L}$. This conclusion is also valid for the explanation of W by L. A graphical evidence of this fact is the dispersion of the $|W|$ concerning $\sqrt{L}$ for the data of the Second and the Third Levelling of Finland given by Figures 2 and 3.

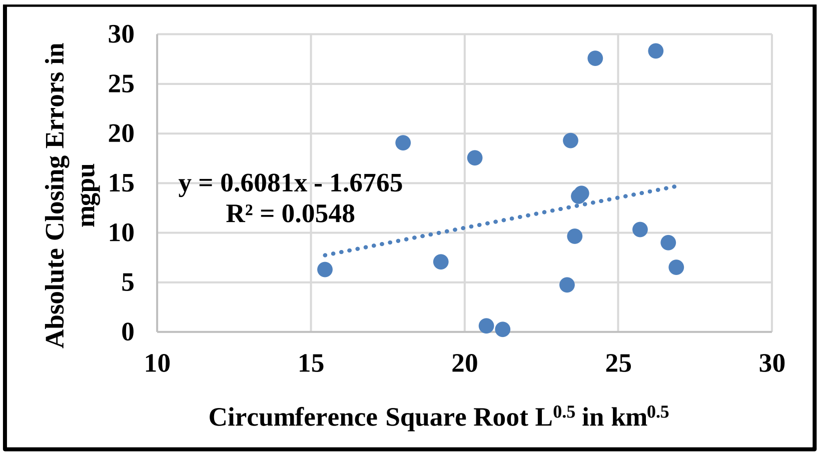

Fig. 2. Closing Errors in the Second Levelling of Finland.
The data are taken from (Kääriäinen, 1966).

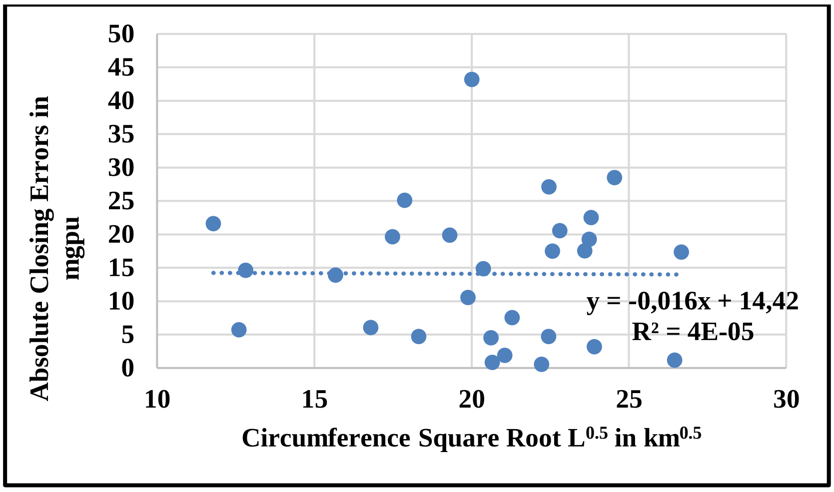

Fig. 3. Closing Errors in the Third Levelling of Finland.
The data are taken from (Saaranen et al., 2021)

According to Figures 2 and 3, which present the determination R^2 and the trend of the increase of absolute closing errors |W| with respect to the squared loop length $\sqrt{L}$ in km, there is a weak (almost nothing) relationship between |W| and $\sqrt{L}$. A similar pattern of the spread of $|W|$ with respect to $\sqrt{L}$ (L) is presented in (Cvetkov, 2022) for the Third Levelling of Bulgaria. Again, given the differences in measurement epochs, levelling equipment, procedures, and the terrain and climate conditions across Finland, Poland, and Bulgaria, it would be naive to claim that the common pattern in the spread of $|W|$ concerning $\sqrt{L}$ (or L) is merely coincidental. On the contrary, this pattern is inherent to precise geometric levelling, and its physical explanation will be presented in a future publication.

1.4. Observations true error vs. observation average true error

The probability *Cv(n)* that the average of *n* independent random variables is closer to the distribution expectation than some of the independent variables is a function of the entropy of the standard Normal distribution and the number of variables *n*. The probability *Cv(n)* can be given by equation (1), where the coefficients *a, b,* and *c* depend on the distribution and the number of variables *n* (Cvetkov, 2024b, 2024c).

$$Cv(n) = 0.25. log_b(2\pi e). (n^{-c} - n^{-a}) \quad (1)$$

Based on (1), it can be calculated that in the case of two measurements, i.e., *n = 2*, the probability *Cv(2)* tends to 30% if both measurements are normally distributed. In our little experiment, we obtained *Cv(2)* = 32%. If one generates their random sample, the result will differ, but if their sample size is n → ∞, then their *Cv(2) → 30%.*

Suppose the situation where both measurements derive from the distribution with the highest entropy, i.e., the Uniform distribution, then $Cv(2) \rightarrow 33\%$. Thus, in more than 66% of the cases, the value of one of the two measurements is closer to the true value of the measured quantity in comparison to the mean. Figure 5 shows how often the first observation, the second observation, or their average comes closest to a known expected value across different distributions (Cvetkov, 2024e).

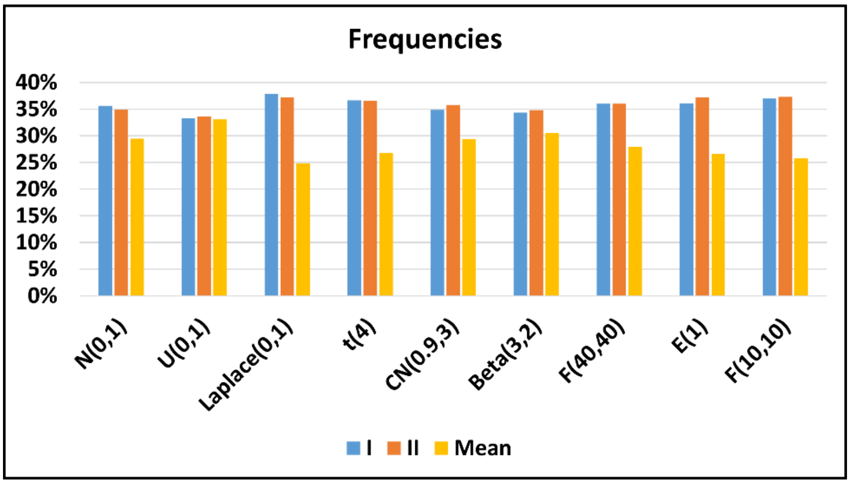

Fig. 5. The frequencies of occurrence of the first, the second observation, or their mean are closest to a known expectation of the applied distributions.

Figure 5 shows that for skewed (asymmetrical) distributions, the average of two observations is more accurate than each observation in less than 30% of the cases. We should remember that with only two real observations, we can not determine their distribution, so assuming normality is just a convenient simplification.

1.5. Research objectives

Building on the statistical findings and probability insights presented above, the primary objective of this article is to expand on Cvetkov's study (Cvetkov, 2024d) by applying newly proposed weighting methods and reconfiguring the levelling network as a free network, with the datum point set at its centroid. A secondary goal is to assess how each step of the proposed algorithm contributes to improved accuracy and to identify the causes behind systematic tilts observed in certain reference systems (e.g., Gerlach and Rummel, 2024).

2. Data and data processing

2.1. Data processing

In this study, we used data from the Third Levelling of Finland (Saaranen et al., 2021), previously analysed by Cvetkov (Cvetkov, 2024d). The network layout is shown in Figure 6, and the data used are listed in Table 1.

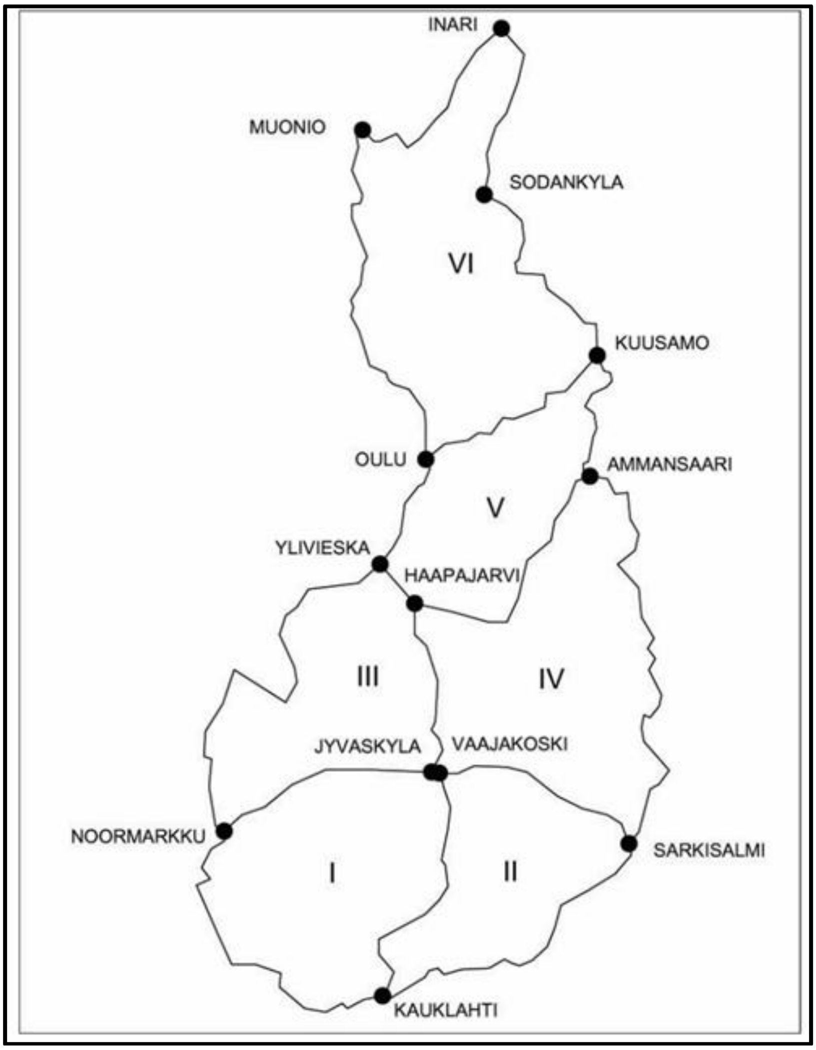

Fig. 6. Scheme of the analysed network, part of the Third Level of Finland network (Cvetkov, 2024d; Saaranen et al., 2021).

Table 1. Summarised data about levelling lines, their length, differences *I* and *II* between geopotential numbers of the start and the end line benchmarks, and the averages of both measurements *I* and *II*.

Line	**Distance**	**Height differences***		**Mean**
		I	***II***	$(I + II)/2$
	(km)	**(mgpu)**	**(mgpu)**	**(mgpu)**
Kauklahti-Noormarkku	363.778	35284.13	35231.27	**35257.70**
Noormarkku-Jyväskylä	257.280	42172.85	**42164.91**	42168.88
Kauklahti-Vaajakoski	337.837	**80998.47**	81017.91	81008.19
Jyväskylä-Vaajakoski	4.836	**3580.65**	3581.17	3580.91
Vaajakoski-Särkisalmi	239.680	-16433.88	**-16399.66**	-16416.77
Kauklahti-Särkisalmi	397.821	64603.73	64598.05	**64600.89**
Noormarkku-Ylivieska	490.659	21623.15	**21533.81**	21578.48
Haapajärvi-Ylivieska	64.685	-40398.99	-40415.51	**-40407.25**
Jyväskylä-Haapajärvi	201.790	**19778.81**	19749.75	19764.28
Haapajärvi-Ammänsaari	302.491	94304.50	**94322.86**	94313.68
Sarkisalmi-Ammänsaari	513.336	126924.65	**126920.11**	126922.38
Ylivieska-Oulu	129.545	**-57206.60**	-57204.58	-57205.59
Oulu-Kuusamo	254.097	**257863.35**	257827.31	257845.33
Ammänsaari-Kuusamo	173.309	**65926.36**	65930.36	65928.36
Oulu-Muonio	434.562	237884.45	237875.21	**237879.83**
Kuusamo-Sodankylä	266.298	-91600.12	-91656.04	**-91628.08**
Sodankylä-Inari	210.199	-54031.64	-54184.14	**-54107.89**
Muonio-Inari	255.329	**-125752.04**	-125707.40	-125729.72

*The values of the *I* and the *II* measurements we calculated based on the information given by Appendix C, columns 5-11 in the study by Saaranen et al. (Saaranen et al., 2021). All height differences contain rod metre, refraction, and temporal tidal and land uplift corrections.

2.2. Data processing

The basic steps of the proposed adjustment of the analysed part of the Third Levelling of Finland network (1978–2006) are as follows:

- Step 1 – In our network, we have 18 lines and three elevation values for each line, i.e., from I and II measurements and their average. Thus, in our case, we have $3^{18} = 387{,}420{,}489$ combinations between line elevations, or 3^{18} independent adjustments. Using those 3^{18} independent adjustments, we found those line elevations that minimised the loop closing errors in the network shown in Figure 6. These selected values are highlighted in bold in Table 1.

- Step 2 - Next, the weights of the selected line elevations were obtained based on the real impact of each one on the final network accuracy. The network was readjusted 18 times, each time omitting a different levelling line. For each case, the network was adjusted using the remaining 17 lines, and the standard errors of the adjusted benchmark heights were summed. If a levelling line degrades the network accuracy, omitting it leads to a smaller sum of nodal benchmark standard errors, implying a lower weight for that line. Because the number of levelling lines is the same in each independent adjustment, the average of the nodal benchmark standard errors (ASE) was used instead of their sum. From the 18 obtained ASE values, their mean value (ASE_{MEAN}) was calculated, and the weight w_{L_i} for line *i* was then computed using equation (2):

$$w_{L_i} = (ASE_i/ASE_{MEAN})^2 \qquad (2)$$

- Step 3 - Finally, the network was adjusted as a free network using the selected height differences for each line together with the assumption-free weights defined by equation (2).

3. Results

Figure 7 compares the accuracy of the classical adjustment method (Saaranen et al., 2021) with the individual steps outlined in Section 2.2. Figure 8 presents the differences between the official benchmark geopotential numbers and those derived from our adjustment approach.

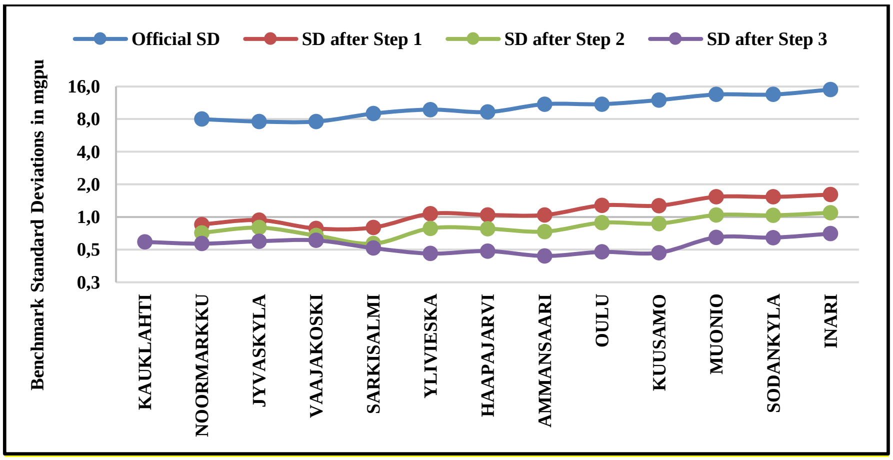

Fig. 7. Standard deviations of the adjusted benchmark geopotential numbers in mgpu – official results (blue), after data processing Step 1 (red), Step 2 (green), and Step 3 (violet).

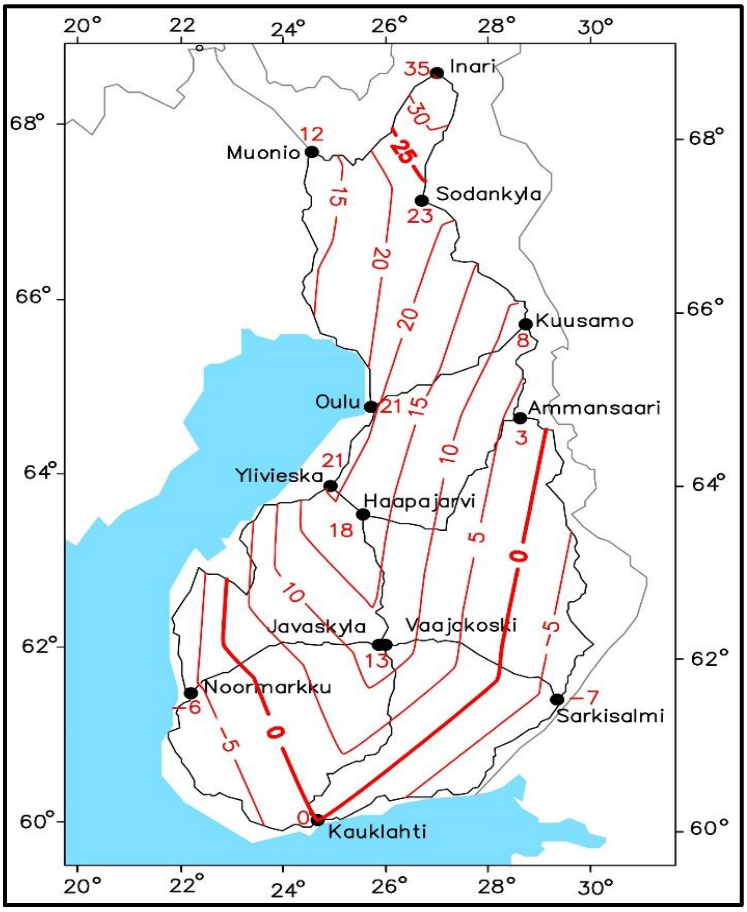

Fig. 8. Differences in mgpu between the official benchmark geopotential numbers (Saaranen et al., 2021) and those calculated using our adjustment method.

4. Discussion

According to Figure 7, the standard deviations of all adjusted benchmark geopotential numbers, obtained after Step 3, are below 0.75 mgpu. The adjusted geopotential number of the Ammansaari benchmark has a minimal standard deviation of 0.44 mgpu. This benchmark is close to the network mass centre. The benchmarks in the network's periphery, e.g., Inari, Sodankyla, and Mounio, have greater standard deviations of the adjusted geopotential numbers, respectively 0.70, 0.65, and 0.65 mgpu. The mean of the standard deviations is 0.55 mgpu. Comparison between the standard deviations of the adjusted geopotential numbers pictured in Figure 7 and the standard deviations of the adjusted geopotential number of the same benchmarks taken from the Saaranen et al.'s study (Saaranen et al., 2021, Figure 6.3) shows that the uncertainty of a levelling network, adjusted in the manner presented here, can be reduced more than 15-20 times.

Additionally, Figure 7 illustrates a stepwise improvement in the accuracy and homogeneity of the adjusted network. A two-tailed paired T-test comparing the mean standard deviations of the benchmark-adjusted geopotential numbers yielded p-values of 1.28×10^{-8}, 1.23×10^{-5}, and 0.015 for the comparisons between the official results and Step 1, Step 2, and Step 3 given in Section 2.2, respectively. These results confirm that each step produced a statistically significant accuracy gain at a confidence level higher than 95%. Notably, the greatest improvement in adjustment accuracy comes from Step 1.

A key aspect of levelling network adjustment is the reliability of the final results. While Least Squares Estimates (LSE) are known to be the Best Linear Unbiased Estimates (BLUE) or Minimum Variance Unbiased Estimates (MVUE) under the assumption of normally distributed measurement errors, they lack robustness against outliers and systematic errors. Instead of isolating such anomalies, LSE spreads its influence across all observations, leading to potential degradation of results (Duchnowski and Wyszkowska, 2019; Javadi, 2017; among others).

One source of this degradation is the routine use of the mean of two-line observations in the adjustment, rather than selecting the observation that minimizes loop closure errors. The consequences of this approach are illustrated in Figure 8. To generate the figure, we fixed the geopotential number of the Kauklahti benchmark (based on Saaranen et al., 2021) and recalculated the adjusted geopotential numbers for the rest of the network using our proposed method. We then computed the differences between these adjusted values and the official results.

Figure 8 reveals a clear systematic tilt in the South–North direction, averaging about 4 mgpu per degree of latitude. This suggests that one potential cause of such tilting in levelling networks (see Gerlach and Rummel, 2024) is the systematic bias introduced by averaging two observations rather than selecting the most consistent one. A secondary, though smaller, contributor to the tilt is the use of weights that are inversely proportional to levelling line lengths.

5. Conclusion

The primary objective of this research was to expand on Cvetkov’s study (Cvetkov, 2024d) by introducing new weighting strategies and reconfiguring the analyzed levelling network as a free network, with the datum positioned at its mass centre. The results show that the adjustment algorithm presented here reduces the uncertainty of adjusted benchmark geopotential numbers by a factor of 15 to 20. The key step in the algorithm is selecting observations - or their means - that minimise levelling loop closure errors.

The accuracy achieved confirms that precise levelling remains the most reliable method for measuring height differences and investigations. Investigations similar to Izvoltona et al. (Izvoltova et al., 2024) are necessary to evaluate the method. GNSS-based alternatives, while valuable, cannot yet match this level of precision (Apollo et al., 2021; Borowski et al., 2025; Celms et al. 2024; Kurtz et al., 2024). However, precise levelling also has its limitations (Borowski et al., 2025). For this reason, combining terrestrial and satellite-based methods offers the most practical path forward in developing modern vertical reference systems (Borowski et al., 2025; Gerlach and Rummel, 2024).

Leveraging the strengths of both approaches, geoid-based vertical reference frames with uncertainties below 10 mgpu are likely within reach.

Acknowledgements

The authors express gratitude to the Centre for Scientific Research and Design at the UACEG for funding and support for the completed research project under contract No BN315/2025 on the topic "Optimisation of levelling networks by studying the applicability of resampling methods."

References

Apollo, M., Mostowska, J., Maciuk, K., Wengel, Y., Jones, T. E., Cheer, J. M. (2021). Peak-bagging and cartographic misrepresentations: a call to correction. *Current Issues in Tourism*, 24(14), 1970–1975. https://doi.org/10.1080/13683500.2020.1812541

Borowski, L., Banasik, P., Maciuk, K. (2025). Application of GNSS-levelling for updating the base vertical network, *Journal of Applied Geodesy*, DOI: 10.1515/jag-2024-0096.

Celms, A., Grinberga, L., Lidumnieks, T., Luksa, J., Brinkmanis-Brimanis, M. (2024). Assessment of accuracy of GNSS measurement models using base station and radio module solutions. *Baltic J. Modern Computing*, 12(4):454–462. https://doi.org/10.22364/bjmc.2024.12.4.07

Cvetkov, V. (2022). Discrepancies and misclosures in the Third Levelling of Bulgaria, *Annual of the University of Architecture, Civil Engineering and Geodesy, Sofia*, 55(1):195-201.

Cvetkov V. (2023). Two adjustments of the second levelling of Finland by using nonconventional weights, *Journal of Geodetic Science*, 13(1):20220148, https://doi.org/10.1515/jogs-2022-0148.

Cvetkov, V. (2024). A myth concerning the geometric levelling: closing errors and loop circumferences. Proceedings of 24th International Multidisciplinary Scientific GeoConference SGEM 2024, 24(2.1):149-154, DOI: 10.5593/sgem2024/2.1/s09.19.

Cvetkov, V. (2024). Averages, entropy and observations, *Deutsche Internationale Zeitschrift für zeitgenössische Wissenschaft*, 93:52–7.

Cvetkov, V. (2024). Entropy, observations and some standard deviations, *Deutsche Internationale Zeitschrift für zeitgenössische Wissenschaft*, 94:13–8.

Cvetkov, V. (2024). On initial data in adjustments of the geometric levelling networks (on the mean of paired observations). *Journal of Geodetic Science*, 14(1):20220170. DOI: 10.1515/jogs-2022-0170.

Cvetkov, V. (2024). Geometric Levelling Data and Some Systematic Faults in Their Treatment, *Inżynieria Mineralna - Journal of the Polish Mineral Engineering Society*, 2024, 1(1), 293–300. https://doi.org/10.29227/IM-2024-01-32

Cvetkov, V. (2024). *Adjustment of the precise levelling networks*, Monography, Sofia, ISBN 978-619-92816-0-4. (in Bulgarian).

Duchnowski, R., Wyszkowska, P. (2019). Empirical influence functions of different robust estimation methods applied in displacement analysis. *4th Joint International Symposium on Deformation Monitoring (JISDM)*, 15-17 May 2019, Athens, Greece.

Gerlach, C., Rummel R. (2024). Benefit of classical leveling for geoid-based vertical reference frames. *Journal of Geodesy*, 98(64). DOI: 10.1007/s00190-024-01849-y.

Izvoltova, J., Chromcak, J., Bacova, D. (2024). The Effect of Critical Distance in Digital Levelling. *Computation*, 12(6), 111. https://doi.org/10.3390/computation12060111.

Javadi, P. (2017). Detecting the Unstable Points in Deformation Monitoring Geodetic Networks in Analysis Method of Subnetwork. Modern Applied Science, 11(3):61-75, ISSN 1913-1844, E-ISSN 1913-1852, http://dx.doi.org/10.5539/mas.v11n3p61.

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

Kurtz, B., Gómez D., Bevis, M. (2024). Characterization of the precision of PPP solutions as a function of latitude and session length, *Journal of Geodetic Science*, 14(1):20220176. DOI: 10.1515/jogs-2022-0176.

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

Saaranen, V., Lehmuskoski, P., Takalo M., Rouhiainen, P. (2021). The Third Precise Levelling of Finland, *FGI Publications No. 161*, Kirkkonummi, https://helda.helsinki.fi/items/921bf91e-e93d-4ccd-9f83-aac063ee014f.

Received December 2, 2025, revised January 12, 2026, accepted February 18, 2026