Open Access
Original paper

Budownictwo i Architektura – Civil and Architectural Engineering 25(1) (2026) 26012
https://doi.org/10.35784/bud-arch.8789

Evaluating levelling adjustment methods in deformation analysis

Vasil Cvetkov[1,*]

[1] *Department of Applied Geodesy; Faculty of Geodesy; University of Architecture, Civil Engineering and Geodesy; 1 Hristo Smirnenski Blvd., 1046 Sofia, Bulgaria; tzvetkov_fgs@uacg.bg*

[*] *Corresponding Author*
Received: 25.11.2025; Revised: 07.02.2026; Accepted: 10.02.2026; Available online: 31.03.2026

Abstract:
The article evaluates the reliability of geometric levelling network adjustments used in deformation monitoring of buildings and infrastructure. It compares classical adjustment, adjustments using all combinations of measured height differences, and robust approaches such as M-, S-, and MM-Estimation. Simulated levelling data were generated so that forward and backward elevation observations shared the same expected value but had different variances. The results indicate that the classical method, which uses the mean of the two elevation measurements, is optimal only when both observation groups have identical variances, meaning that all data strictly follow a single Gaussian distribution. When variances differ, averaging the paired observations is no longer advisable. In such situations, adjusting all possible combinations of observations, including their mean, provides the most reliable results. The simulations show that this strategy offers higher empirical efficiency and greater robustness than commonly used robust estimators, including Huber, Tukey, Andrew, IGG III, L1-Norm, the Danish Method, S-Estimation, and MM-Estimation, even when variance differences are small.

Keywords:
deformation analysis, geometric levelling, M-estimation, S-Estimation, reliability of results

1. Introduction

Owing to its exceptional accuracy [1-4], precise geometric levelling is the preferred technique for deformation analysis involving vertical movements [1,5,6]. Height differences between reference and object benchmarks are measured in two opposite directions to eliminate gross errors and reduce certain systematic effects. Consequently, each elevation in the levelling network typically has two measured values. These are commonly averaged and used as the initial input for subsequent processing [7]. However, this practice, combined with the use of an unsuitable weighting scheme [8-10], can compromise the final adjustment results and lead to unreliable or incorrect conclusions about vertical displacements and velocities.

Research [7,10] indicates that, in precise levelling networks, the quality of the initial data has a stronger influence on adjustment accuracy than the choice of weights. Furthermore, selecting appropriate initial adjustment data significantly reduces the results' sensitivity to the weighting scheme [7,10].

According to studies [7,10,11], the mean of two random numbers drawn from a standard normal distribution has a smaller true error than either of its parent values in only about 30% of cases. When the two numbers follow a uniform distribution, this probability increases to roughly 33%, representing the most favourable scenario for using the mean. For other parent distributions, the likelihood that the mean has the smallest true error drops below 30%. Therefore, relying solely on the mean of two observations, rather than considering all available data, is neither reasonable nor correct, a conclusion supported by previous research [7,10,12].

Furthermore, the findings in [12] show that selecting one of the original levelling observations as initial adjustment data yields more accurate results than the traditional practice of averaging double-measured line elevations. In this context, advanced approaches such as M-split estimation [5,6,13-15] appear highly promising and will be explored in future work.

The primary objective of this study is to assess the reliability of the adjustment method proposed in [7] and to compare its performance with that of ordinary least squares (OLS), commonly used M-estimators [17-27], S-estimation [25-28], and MM-estimation [25,29] through Monte Carlo simulations. OLS is included as a reference method because it is the most widely applied technique in geodetic data processing and provides optimal results when the observation errors follow a normal distribution [16-19,23,26-29]. However, when this assumption is violated, OLS estimates become highly sensitive to outliers, which can significantly degrade the accuracy and reliability of the results [20,21]. In such cases, robust estimation methods such as M-, S-, and MM-estimation are expected to provide more reliable solutions [16-19,24-29].

The secondary objective of this study is to determine the conditions under which each estimation technique is most suitable for deformation analysis, particularly for applications involving vertical displacement measurements obtained from precise geometric levelling.

2. OLS and robust estimation methods used in research

Empirical testing of robust estimation methods is commonly applied in engineering and industrial surveying [16-27]. For this study, several well-established robust techniques were selected for evaluation, including Huber, Tukey, Andrew, IGG III, and L1-Norm M-Estimation, as well as the Danish Method, S-Estimation, and MM-Estimation.

2.1. Ordinary least squares (OLS)

Let **y** be an observation vector, **A** be a known matrix that defines the levelling network configuration, X be an unknown parameter vector, and **v** be a vector of random measurement error. Then, the Gauss-Markov functional model between them can be given by Eq. 1.

$$v = AX - y \tag{1}$$

Applying the objective function (2), the parameter vector X can be found by Eq. 3, and the adjusted observations $\hat{y}$ by Eq. 4.

$$[wvv] = min \tag{2}$$

$$X = (A^TWA)^{-1}A^TWy \tag{3}$$

$$\hat{y} = y + v \tag{4}$$

If all observations are independent and have the same accuracy, the weight matrix **W** will be equal to the identity matrix **I**. In our simulations, we set all levelling lines with equal lengths to avoid the problems with the choice of weighting models [8,9]. In the real-data example, we applied adjustments with equal weights (w = 1) and with weights inversely proportional to levelling distance (w = L^{-1}).

2.2. M-Estimation

The idea behind M-Estimation (maximum likelihood estimation) was introduced by Huber [20] and is based on selecting the objective function **ρ(v)** that increases less rapidly than the objective function given by Eq. 2 [13,16]. The conditions that must satisfy an objective function ρ(v) are widely discussed in many articles [14-17,21,23,24,28], so we can skip their mention. Based on the selected objective function **ρ(v)**, two other important functions can be derived, i.e., the influence function **ψ(v)** and the weighting function **w(v)**. The relationship between **ρ(v)**, **ψ(v)**, and **w(v)** can be given by Eq. 5.

$$w(v) = v^{-1}\psi(v) = v^{-1}.\delta\rho(v)/\delta v \tag{5}$$

To apply M-estimation using the algorithms described in studies [25,26], we present the weighting functions used in this study: Huber, Tukey, Andrew, IGG III, L1-norm, and the Danish method. In all these functions, the notation $\overline{v_\iota}$ is a standardized error of the i^{th} observation calculated by Eq. 6. In those cases, in equations below, when $\overline{v_\iota}$ = 0 in their denominators, we set $\overline{v_\iota}$ = 0.0000001.

$$\overline{v_\iota} = v_i/\hat{\sigma} \tag{6}$$

$$\hat{\sigma} = 1.4826.med|v_i - med(v_i)| \tag{7}$$

2.2.1. Huber method

The weighting function of Huber's method is given by Eq. 8.

$$w(\overline{v_\iota}) = \begin{cases} 1 & for\ |\overline{v_\iota}| \le k \\ \frac{k}{|\overline{v_\iota}|} & for\ |\overline{v_\iota}| > k \end{cases} \tag{8}$$

In our study, we use two values of *k*. To achieve an estimator efficiency of 95% [16,17, 26], we used *k* = 1.345. According to [14], the coefficient *k* is usually equal to 2, 2.5, or 3. Therefore, in the second estimation by Huber's method, we applied *k* = 2.5.

2.2.2. Tukey method

The weighting function of Tukey's method is given by Eq. 9.

$$w(\overline{v_\iota}) = \begin{cases} \left[1 - \left(\frac{\overline{v_\iota}}{k}\right)^2\right]^2 & for\ |\overline{v_\iota}| \le k \\ 0 & for\ |\overline{v_\iota}| > k \end{cases} \tag{9}$$

In our study, we used two values of *k*. The studies [15,25,26] recommended *k* = 4.685 or *k* = 6. Since applying the coefficient *k* = 6 will result in lower robustness [20,21], we used *k* = 4.685 in our experiments. Second, we decided to set the coefficient *k* = 1.547.

2.2.3. Andrew method

Andrew's weighting function is given by Eq. 10 [22].

$$w(\overline{v_\iota}) = \begin{cases} \frac{sin\left(\frac{|\overline{v_\iota}|}{k}\right)}{\frac{|\overline{v_\iota}|}{k}} & for\ |\overline{v_\iota}| \le \pi k \\ 0 & for\ |\overline{v_\iota}| > \pi k \end{cases} \tag{10}$$

In our study, we used two values of *k*. The first one is *k* = 4.206 [16]. The second one is *k* = 1.339 [15]. The first choice of the coefficient *k* is related to greater effectiveness, and the second choice is related to a more robust decision.

2.2.4. IGG III method

The weighting function for the IGG III method is given by Eq. 11.

$$w(\overline{v_\iota}) = \begin{cases} p_i & for\ |\overline{v_\iota}| \le k_1 \\ \frac{p_i k_1}{|\overline{v_\iota}|}\left(\frac{k_2 - |\overline{v_\iota}|}{k_2 - k_1}\right)^2 & for\ k_1 < |\overline{v_\iota}| \le k_2 \\ 0 & for\ |\overline{v_\iota}| > k_2 \end{cases} \tag{11}$$

In our study, we used two sets for k_1 and k_2. The studies [15,18] recommended k_1 = 1.5 and k_2 = 3. To achieve more efficiency when two samples have approximately equal variances, we also used coefficients k_1 = 3 and k_2 = 6. The quantities p_i are the observation weights in the previous iteration. The first choice of the coefficients k_1 and k_2 prioritises higher efficiency, whereas the second choice emphasises greater robustness.

2.2.5. L1-Norm

The weighting function for the L1-norm method, also known as the least absolute deviation method, is given by Eq. 12.

$$w(\overline{v_\iota}) = \left|\frac{1}{v_i}\right| \tag{12}$$

If v_i = 0, then we set v_i = 0.0000001.

2.2.6. The Danish method

The weighting function of the Danish method is given by Eq. 13.

$$w(\bar{v}_i) = \begin{cases} 1 & for\ |\bar{v}_i| \le 2\sigma \\ \exp\,(k\bar{v}_i^2) & for\ |\bar{v}_i| > 2\sigma \end{cases} \tag{13}$$

In our study, we used $k = 2$. In Eq. 13, σ is the standard deviation, calculated by Eq. 7 in each iteration.

2.3. S-Estimation

S-estimators, or estimators of scale, are proposed by Rousseeuw and Yohai to deal with large fractions of contaminated data [28]. They have high robustness, and their breakdown point can reach 50%. In parallel, they can have high asymptotic efficiency. In this manner, S-estimators are flexible competitors of M-estimators. S-etimators' objective function is

$$\rho(\bar{v}_i) = \begin{cases} \frac{\bar{v}_i^{\,2}}{2} - \frac{\bar{v}_i^{\,4}}{2c^2} + \frac{\bar{v}_i^{\,6}}{6c^4} & for\ |\bar{v}_i| \le c \\ \frac{c^2}{6} & for\ |\bar{v}_i| > c \end{cases} \tag{14}$$

the derivative of which is Tukey's biweighted function (15) [28].

$$\psi(\bar{v}_i) = \begin{cases} \bar{v}_i\left[1 - \left(\frac{\bar{v}_i}{c}\right)^2\right]^2 & for\ |\bar{v}_i| \le c \\ 0 & for\ |\bar{v}_i| > c \end{cases} \tag{15}$$

The constant $c > 0$ is such that the objective function, given by Eq. 14, is strictly increasing on [0, c] and is a constant on $[c, \infty]$. Therefore, the weight function of S-estimators can be given by Eq. 16.

$$w(\bar{v}_i) = \begin{cases} \left[1 - \left(\frac{\bar{v}_i}{c}\right)^2\right]^2 & for\ |\bar{v}_i| \le c \\ 0 & for\ |\bar{v}_i| > c \end{cases} \tag{16}$$

Thus, the robustness of S-estimators depends on the value of c. In our study, we applied three values of c = 1.547, 2.937, and 5.182, corresponding to breakdown points of 50%, 25%, and 10%, respectively [27,28]. The S-estimation algorithm is given in [25,26,28].

2.4. MM-Estimation

MM-estimates are a class of robust estimates that simultaneously possess the following properties: they are highly efficient when the errors follow a normal distribution, and their breakdown point is 0.5 under strongly contaminated data [29]. The idea of MM-estimates was introduced in the study [28] and consists of a two-stage S-estimation procedure. At the first stage, an S-estimation with a c = 1.547 coefficient is recommended because of its 50% breakdown point. At the second stage, the process continues with M-estimation with a more efficient influence function ψ to get higher asymptotic efficiency [28]. In our research, at the first stage, we applied S-estimation with a coefficient c = 1.547. At the second stage, we used Tukey's M-estimation method with a coefficient k = 4.685. The MM-estimation algorithm is given in [25,26,29].

2.5. 3ⁿ combinations

The 3^n combination adjustment method for precise levelling networks was introduced by Cvetkov [7]. The principle of this approach is to identify those observations, or mean observations, of line elevations within the network that are most mutually consistent, thereby minimising the objective function *[pvv] = min* because of their minimal true errors [11]. Although the method is computationally demanding, its use in monitoring vertical displacements during deformation analysis is both feasible and advantageous, particularly since such networks typically consist of fewer than 15 levelling lines [1,5,6,15,23,30].

3. Empirical and real data tests

3.1. Empirical tests

To avoid gross errors and to increase the accuracy, each elevation between benchmarks in precise levelling networks is usually measured twice. In our simulated levelling network, presented in Fig. 1, the forward observations are shown as black arrays, and the backward observations as white arrays. The common practice is to use its mean values in the adjustment. For example, the elevation between benchmarks 1 and 3 is measured twice – $h_{1,3}$ and $h_{3,1}$. As an initial data point in the adjustment, we usually use their average $h^A_{1,3} = (h_{1,3} - h_{3,1})\,/\,2$. According to [11], such an approach is incorrect because there is more than a 66% probability that the true error of either h1,3 or h3,1 is smaller than the true error of their average, $h^A_{1,3}$.

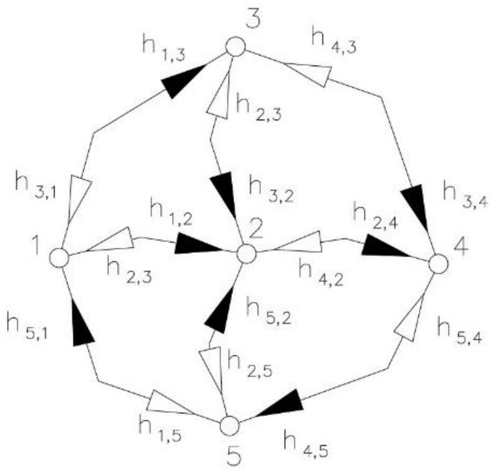

Fig. 1. Simulated levelling network

In the case of two observations of a quantity, there are possible situations where both observations have positive signs of true errors, i.e., their values are greater than the true value of the measured quantity. Also, there are situations in which both observations have negative signs of true error, i.e., their values are smaller than the true value of the measured quantity. In addition, there are situations in which one observation value is larger, and the other is smaller than the true quantity. This case is a valuable example of Gaussian data, where observations with opposite signs are compensated in the average as the number of observations tends to infinity. However, in precise levelling networks, we have only two or, at most, a few observations for each elevation. Since measurements of a quantity are often taken under varying conditions, they may originate from different distributions or from the same distribution with differing parameters. Consequently, the assumption of homogeneity and equal data accuracy is questionable. A simple illustration of these considerations is shown in Fig. 2. The number of paired observations is set to eight, consistent with Fig. 1.

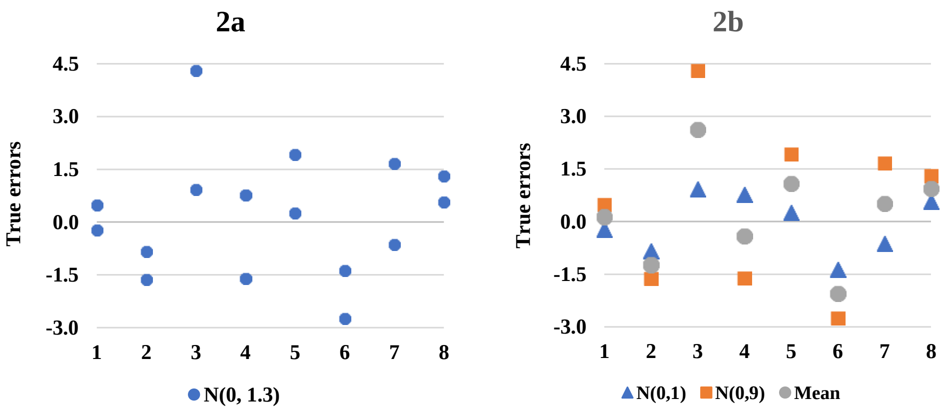

Fig. 2. Random normal true errors of line elevations considering the network in Fig. 1. 2a) A general sample from the N(0, 3) distribution. 2b) Separated samples of the more accurate observation (blue triangles), the worst observation (orange squares), and the averages of both observations in each line

Figure 2a presents random pairs derived from a normal distribution with expectation μ=0 and scale σ=1.69. Apparently, there are no outliers or gross errors. All values are within μ±3σ. However, the data can be viewed from another perspective. Since observations in pairs often have different true errors, we can select the observations with the least true errors. The result of such selection is presented in Fig. 2b, where the data set from Fig. 2a is separated into two sets, which can be related to normal distribution with different parameters, e.g., N(μ=0, σ^2=1) and N(μ=0, σ^2=9). Thus, the ratio between the variances of the samples denoted by the blue triangles and the orange squares in Fig. 2b can be written as $(\sigma_{worse}/\sigma_{better})^2 = 9$. As shown in Fig. 2a, even such a large ratio in the sample variances may not be distinguished when data are prepared for adjustment. Obviously, from Fig. 2b, the spread of the means of the pairs is greater than the spread of the selected data with distribution N(μ=0, σ^2=1). One can calculate the ratio $(\sigma_{mean}/\sigma_{better})^2 \approx 3$.

To investigate the relevance of different robust methods, we simulated observations with random errors derived from a normal distribution with different variances. The first group of observations, let us call them "the forward observations", comes from N (μ=0, σ^2=1). The second group's observation, let they be named "the backward" observations, are derived from N (μ=0, σ^2=1), N (μ=0, σ^2=2), N (μ=0, σ^2=3), N (μ=0, σ^2=4), N (μ=0, σ^2=5), N (μ=0, σ^2=10), N (μ=0, σ^2=25), N (μ=0, σ^2=50), and N (μ=0, σ^2=100). Using simulated data, we performed 10,000 parametric adjustments of the levelling network shown in Fig. 1 for each data pair, treating it as a free levelling network and applying the robust estimation methods described in Section 2. To avoid assumptions about the choice of the weighting model [7,8], we accept that the length of all levelling lines is equal to 1 km. Additionally, all nodal benchmarks were assigned equal heights, such that $H_1 = H_2 = H_3 = H_4 = H_5 = 0$. Thus, the line elevation values are, in fact, their simulated true errors, and the adjusted elevations are deviations from the true elevations. To evaluate the different estimators discussed in Section 2, we computed the root-mean-squared error (RMSE) θ for each adjustment using Eq. 17.

$$\theta_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(h_i^{adjusted}\right)^2} \tag{17}$$

Finally, we assessed each estimator for each „forward"–„backward" sample pair using Eq. 18.

$$\theta_{mean} = \frac{\sum_{j=1}^{10000}\theta_j}{10000} \tag{18}$$

The obtained values $\theta mean$ for all the analysed methods are given in Tables 1-3. On this basis, the empirical relative efficiency of each method, referenced to the OLS θ_{mean}, is calculated according to Eq. 19, and the results are presented in Tables 4-6.

$$e(OLS, Robust\ etimator) = \frac{\theta^2_{mean\ (OLS)}}{\theta^2_{mean\ (Robust\ estimator)}} \tag{19}$$

3.2. Real data test

To demonstrate the performance of the proposed all-combinations adjustment method, we used levelling network data from Zrinjski et al. [30], Table 1. The observations, summarised in Table 1, include average benchmark distances as reported in the original study. Forward and backward measurements were ensured by using data from both epochs. The network configuration is illustrated in Fig. 1.

Table 1. Initial data in the test

Levelling line		Distance L (m)	Height differences (m)		
From	To		I	II	Mean
R1	R2	32.70	0.06696	0.06718	0.06707
R1	R3	46.95	0.02623	0.02653	0.02638
R1	R4	92.25	0.02294	0.02352	0.02323
R1	R5	81.55	0.00728	0.00752	0.00740
R1	R6	72.95	0.05175	0.05227	0.05201
R2	R3	20.20	-0.04052	-0.04031	-0.04042
R2	R4	79.45	-0.04429	-0.04382	-0.04406
R2	R5	68.60	-0.05972	-0.05932	-0.05952
R2	R6	69.65	-0.01536	-0.01482	-0.01509
R3	R4	87.85	-0.00321	-0.00359	-0.00340
R3	R5	76.75	-0.01914	-0.01920	-0.01917
R3	R6	69.40	0.02546	0.02537	0.02542
R4	R5	16.95	-0.01539	-0.01540	-0.01540
R4	R6	20.00	0.02906	0.02912	0.02909
R5	R6	14.20	0.04448	0.04427	0.04438

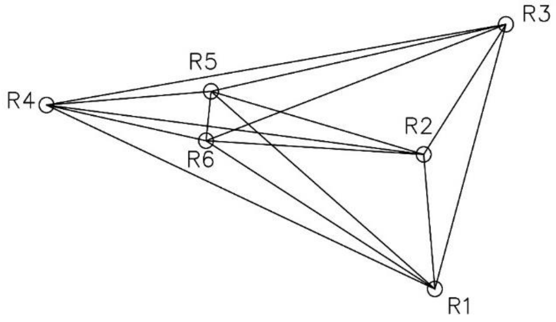

Fig. 3. Scheme of the levelling network used in the real data example

4. Results

4.1. Empirical test results

Tables 2-4 present the root-mean-squared errors obtained using Eq. 18 for each estimation method. As shown, the adjustment that incorporates all combinations of measured heights yields the most stable solution, particularly when considering changes in the variance ratio between the two virtual samples constructed from the measurements with the smallest and largest true errors for each elevation.

Table 2. Accuracy of adjustments based on the OLS, all combinations of observations and their means, Tukey's, Huber's, and Andrew's M-Estimation with different parameters

$(\sigma_m/\sigma_n)^2$	OLS	3^n	Tukey k=4.685	Tukey k=1.547	Huber k=1.345	Huber k=2.5	Andrew k=1.339	Andrew k=4.206
1	0.589	0.644	0.658	0.751	0.616	0.593	0.614	0.596
2	0.714	0.722	0.795	0.895	0.735	0.716	0.740	0.715
3	0.813	0.775	0.874	0.991	0.815	0.812	0.823	0.811
4	0.901	0.818	0.946	1.070	0.886	0.890	0.885	0.898
5	0.974	0.847	1.007	1.121	0.941	0.963	0.944	0.980
10	1.280	0.965	1.215	1.313	1.175	1.244	1.174	1.280
25	1.877	1.136	1.619	1.620	1.646	1.805	1.642	1.878
50	2.567	1.275	2.120	1.916	2.186	2.457	2.173	2.551
100	3.567	1.363	2.827	2.321	2.956	3.342	2.955	3.526

Table 3. Accuracy of adjustments based on the OLS, all combinations of observations and their means, IGG III, L1-Norm M-Estimation with different parameters, and the Danish Method

$(\sigma_m/\sigma_n)^2$	OLS	3^n	IGG III $k_1 = 1.5$, $k_2 = 3$	IGG III $k_1 = 3$, $k2 = 6$	L1 – Norm	Danish Method
1	0.589	0.644	0.686	0.600	0.645	0.651
2	0.714	0.722	0.815	0.724	0.773	0.771
3	0.813	0.775	0.903	0.822	0.846	0.846
4	0.901	0.818	0.968	0.894	0.931	0.918
5	0.974	0.847	1.022	0.962	0.981	0.975
10	1.280	0.965	1.198	1.231	1.181	1.222
25	1.877	1.136	1.511	1.762	1.554	1.812
50	2.567	1.275	1.901	2.354	1.996	2.542
100	3.567	1.363	2.466	3.233	2.626	3.500

Table 4. Accuracy of adjustments based on the OLS, all combinations of observations and their means, S-Estimation with different breakdown points (BDP) of 50%, 25%, and 10%, and MM-Estimation

$(\sigma_m/\sigma_n)^2$	OLS	3^n	S BDP = 50%	S BDP = 25%	S BDP = 10%	MM
1	0.589	0.644	0.751	0.731	0.635	0.685
2	0.714	0.722	0.882	0.863	0.759	0.819
3	0.813	0.775	0.976	0.950	0.851	0.901
4	0.901	0.818	1.040	1.001	0.917	0.969
5	0.974	0.847	1.101	1.055	0.975	1.020
10	1.280	0.965	1.287	1.222	1.206	1.223
25	1.877	1.136	1.585	1.524	1.652	1.631
50	2.567	1.275	1.892	1.864	2.169	2.106
100	3.567	1.363	2.347	2.380	2.920	2.810

Tables 5-7 present the empirical relative efficiency of each estimation method with respect to the classical least squares adjustment. As shown, the ordinary least squares is reliable only when both observation groups have identical variances and follow a single Gaussian distribution. When variances differ, even slightly, this averaging approach becomes suboptimal. In such cases, the most appropriate method is the adjustment in all combinations (3^n adjustment).

Table 5. Empirical relative efficiency in percents of adjustments based on the OLS, all combinations of observations and their means,Tukey's, Huber's, and Andrew's M-Estimation with different parameters

$(\sigma_m/\sigma_n)^2$	OLS	3^n	Tukey k=4.685	Tukey k=1.547	Huber k=1.345	Huber k=2.5	Andrew k=1.339	Andrew k=4.206
1	100.0	83.6	80.1	61.6	91.4	98.6	92.0	97.5
2	100.0	98.0	80.7	63.7	94.5	99.6	93.2	99.9
3	100.0	110.0	86.5	67.3	99.6	100.2	97.7	100.5
4	100.0	121.2	90.8	70.9	103.4	102.5	103.6	100.8
5	100.0	132.1	93.5	75.4	107.0	102.1	106.4	98.8
10	100.0	175.9	111.0	95.1	118.8	105.9	118.9	100.1
25	100.0	272.9	134.4	134.3	130.1	108.2	130.7	99.9
50	100.0	405.2	146.6	179.5	137.9	109.1	139.6	101.3
100	100.0	684.5	159.2	236.1	145.6	113.9	145.7	102.3

Table 6. Empirical relative efficiency in percents of adjustments based on the OLS, all combinations of observations and their means, IGG III, L1-Norm M-Estimation with different parameters, and the Danish Method

$(\sigma_m/\sigma_n)^2$	OLS	3^n	IGG III k_1=1.5, k_2=3	IGG III k_1=3, k_2=6	L1 – Norm	Danish Method
1	100.0	83.6	73.8	96.2	83.3	81.8
2	100.0	98.0	76.9	97.5	85.5	85.9
3	100.0	110.0	81.0	97.9	92.5	92.3
4	100.0	121.2	86.6	101.6	93.6	96.4
5	100.0	132.1	90.8	102.4	98.6	99.7
10	100.0	175.9	114.2	108.1	117.6	109.7
25	100.0	272.9	154.4	113.6	145.9	107.3
50	100.0	405.2	182.4	118.9	165.5	102.0
100	100.0	684.5	209.3	121.7	184.5	103.9

Table 7. Empirical relative efficiency in percents of adjustments based on the OLS, all combinations of observations and their means, S-Estimation with different breakdown points (BDP) of 50%, 25%, and 10%, and MM-Estimation

$(\sigma_m/\sigma_n)^2$	OLS	3^n	S BDP = 50%	S BDP = 25%	S BDP = 10%	MM
1	100.00	83.59	61.41	64.82	85.92	73.95
2	100.00	98.02	65.69	68.56	88.67	76.18
3	100.00	110.03	69.49	73.24	91.40	81.46
4	100.00	121.20	75.07	81.02	96.61	86.42
5	100.00	132.13	78.22	85.21	99.73	91.08
10	100.00	175.90	98.99	109.85	112.62	109.66
25	100.00	272.90	140.34	151.71	129.19	132.48
50	100.00	405.19	184.15	189.72	140.09	148.51
100	100.00	684.52	230.85	224.61	149.16	161.06

4.2. Real data test results

Figures 4-9 present the standard errors of the adjusted benchmarks in the network shown in Fig. 3, produced by different adjustment methods: classical least squares, M-, S-, and MM-estimation, and the All Combinations (3^n combinations) method. Figures 10-13 illustrate the difference in adjusted elevations between Tukey (k = 1.457) M-Estimation and the other discussed approaches.

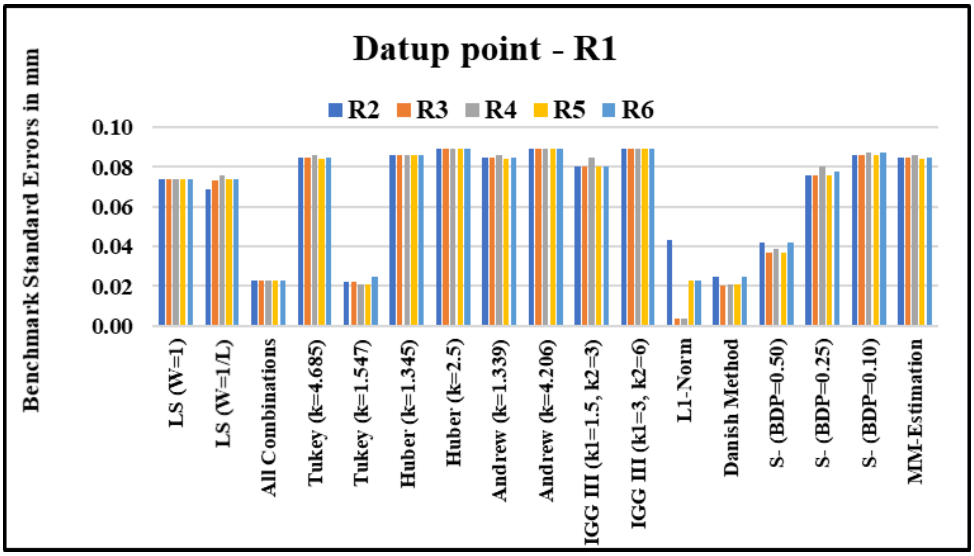

Fig. 4. Standard errors of the adjusted benchmark heights in the network pictured in Fig. 3. The datum point was set in benchmark R1

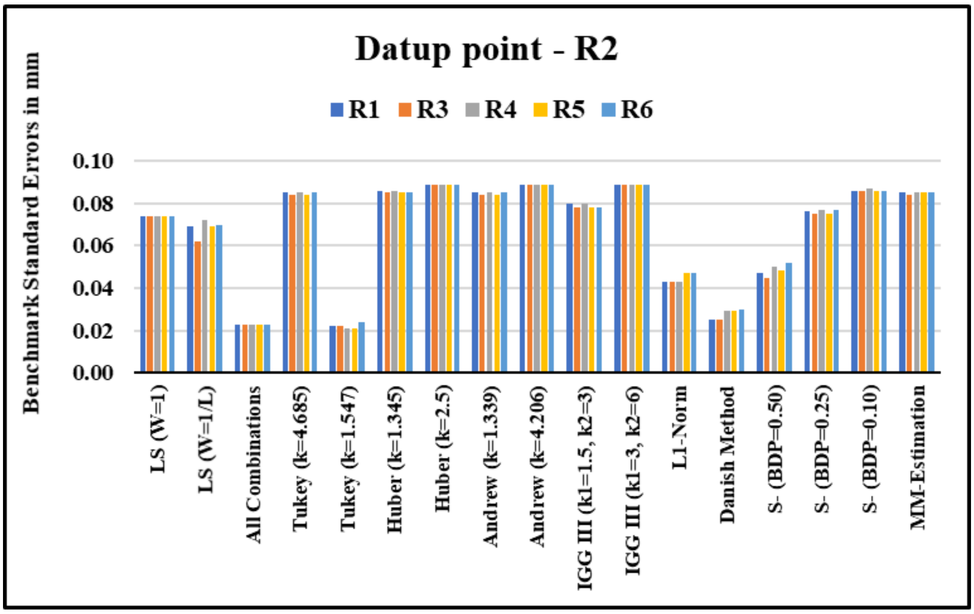

Fig. 5. Standard errors of the adjusted benchmark heights in the network pictured in Fig. 3. The datum point was set in benchmark R2

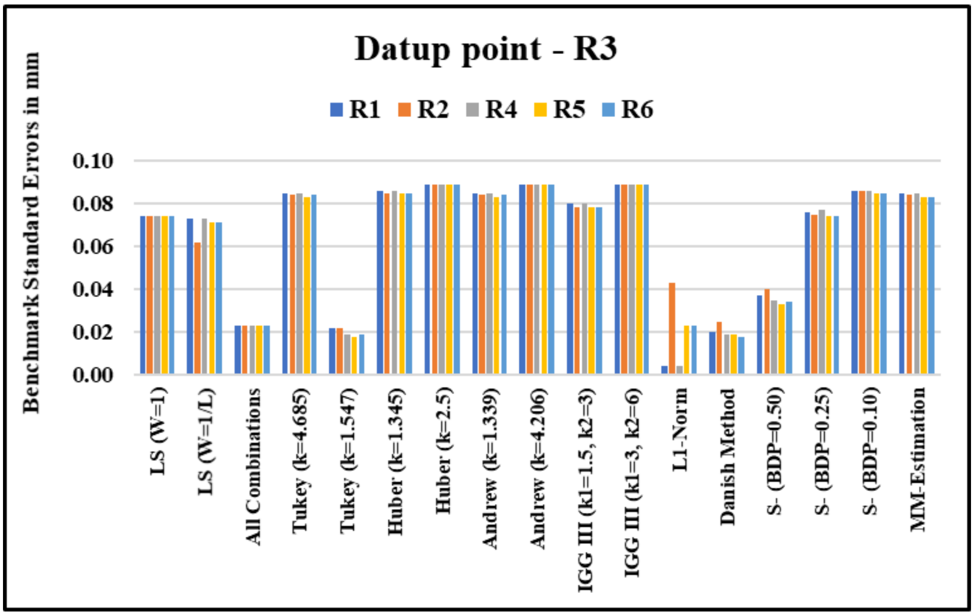

Fig. 6. Standard errors of the adjusted benchmark heights in the network pictured in Fig. 3. The datum point was set in benchmark R3

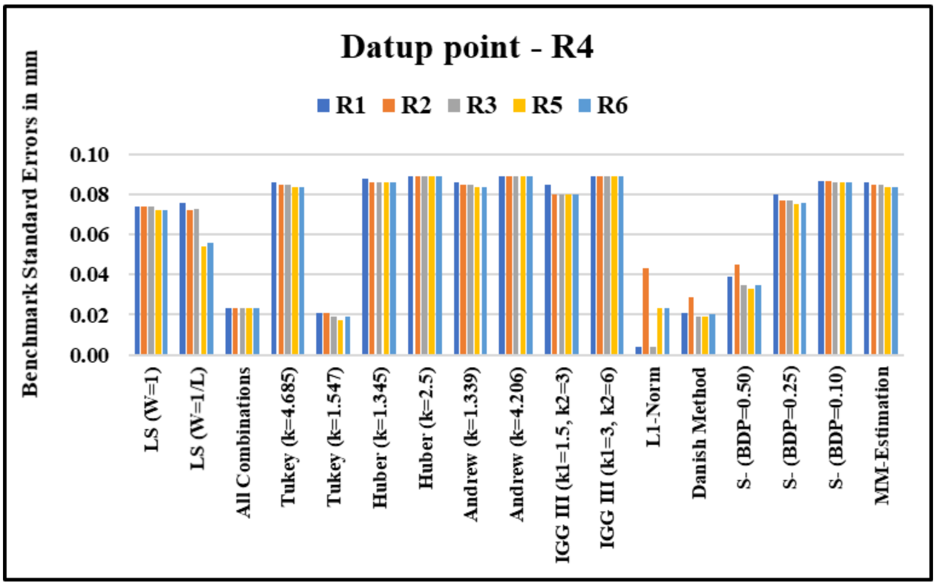

Fig. 7. Standard errors of the adjusted benchmark heights in the network pictured in Fig. 3. The datum point was set in benchmark R4

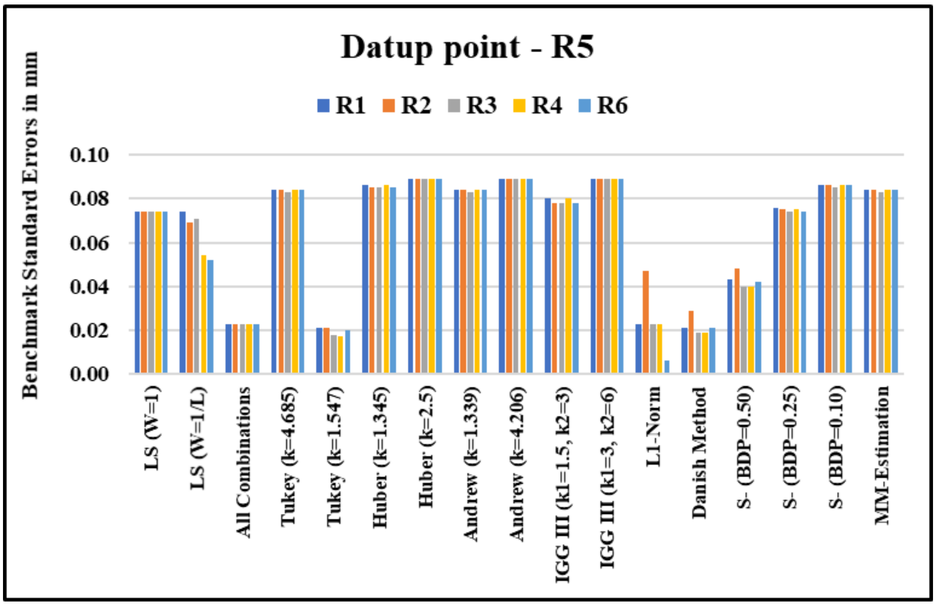

Fig. 8. Standard errors of the adjusted benchmark heights in the network pictured in Fig. 3. The datum point was set in benchmark R5

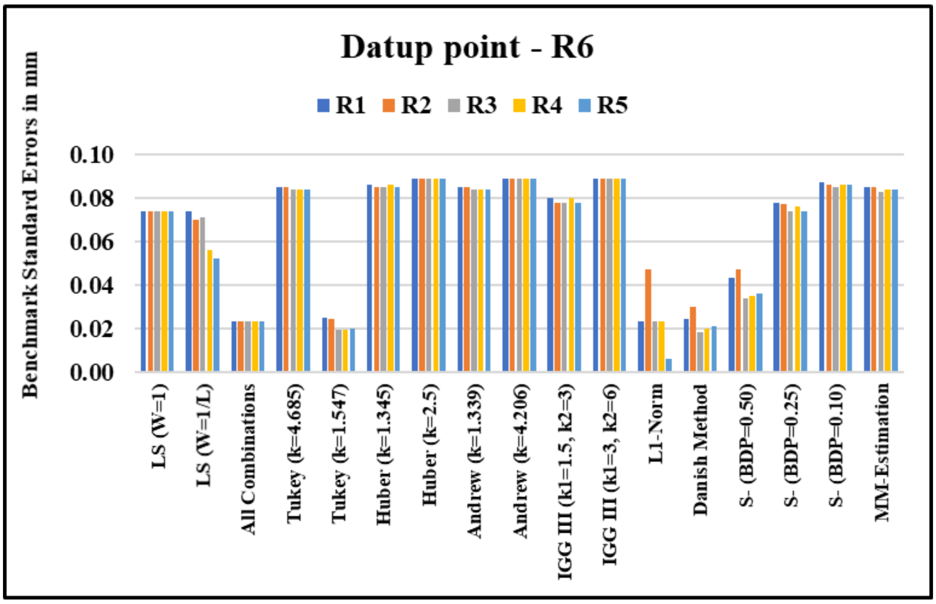

Fig. 9. Standard errors of the adjusted benchmark heights in the network pictured in Fig. 3. The datum point was set in benchmark R6

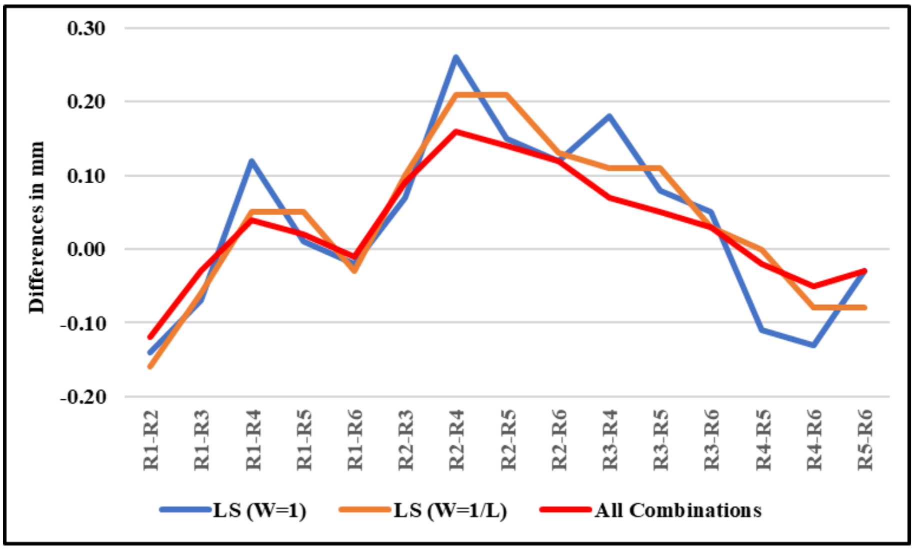

Fig. 10. Differences between the adjusted heights by Tukey (k=1.457) and Least Squares adjustments with equal weights (W=1), classical weights in levelling (W=1/L), and the All Combinations (3^n combinations) method

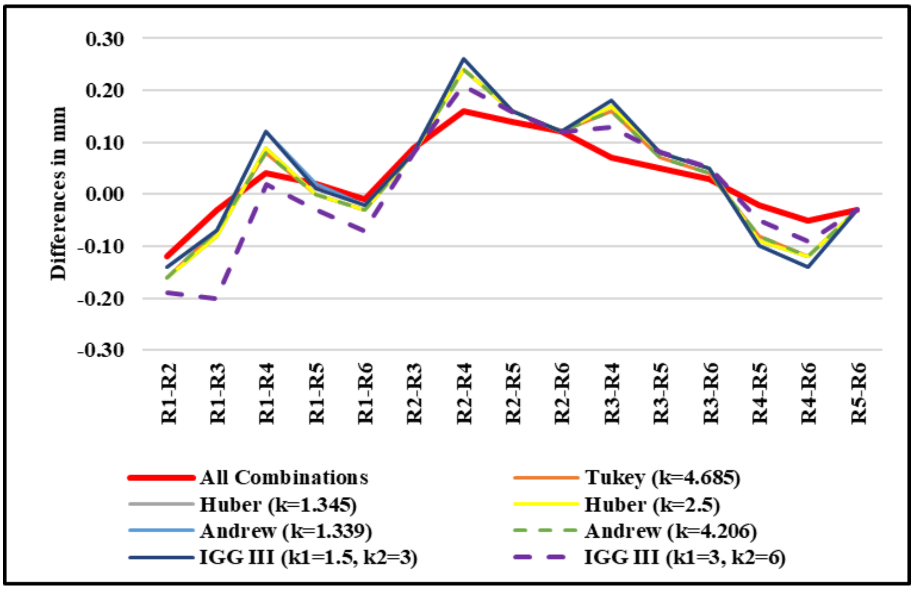

Fig. 11. Differences between the adjusted heights by Tukey (k=1.457), the All Combinations (3^n combinations) method, and various M-Estimators with different parameters

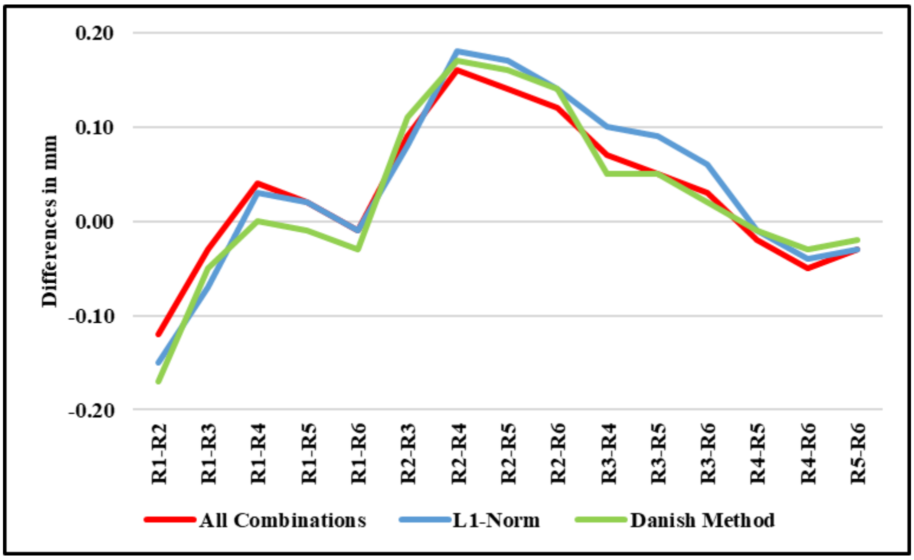

Fig. 12. Differences between the adjusted heights by Tukey (k=1.457), L1-Norm, the Danish method, and the All Combinations (3^n combinations) method

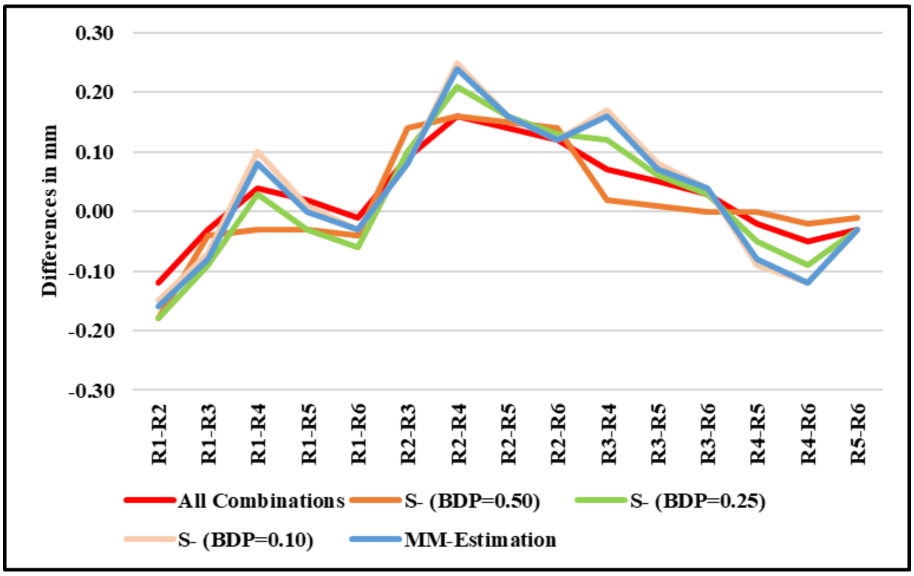

Fig. 13. Differences between the adjusted heights by Tukey (k=1.457), the All Combinations (3^n combinations) method, S-Estimation with different breakdown points (BDP), and MM-Estimation

5. Discussion

5.1. Discussion of empirical test results

Looking across the results in Tables 2-7, several patterns become clear. First, when both sets of measured elevations in the levelling sections originate from the same distribution, with equal variances, the ordinary least squares (OLS) method remains the most reliable option. This aligns well with earlier findings in the literature [7,15-18]. In such homogeneous conditions, only a few alternatives come close: Huber's and Andrew's M-estimators, specifically with coefficients $k = 2.5$ and $k = 4.206$. Other M-estimators, such as Tukey, the L1-norm, IGG III, or the Danish method, consistently underperform relative to OLS when the data behaves approximately Gaussian.

The situation changes considerably when the variances of the observation groups diverge. Once the ratio of the smaller to the larger true-error variances exceeds 2, the adjustment method based on all observation combinations clearly stands out. The RMSE values in Tables 2-4 show that this method's error grows more slowly than that of the competing approaches. Consequently, its empirical efficiency relative to OLS increases rapidly. This is strong evidence of its robustness under non-homogeneous conditions. Importantly, this approach does not rely on assumptions about contamination levels, nor does it depend on tuning breakdown points or selecting method coefficients, practical advantages that make it broadly applicable.

In contrast, the overall use of M-estimators for monitoring vertical displacements with precise geometric levelling appears impractical. Their limitations are twofold: one must choose both the specific M-estimation method and appropriate coefficients, and even then, performance gains are limited. When the data is fully homogeneous and Gaussian, M-estimators cannot surpass OLS. When the data is contaminated, they fail to outperform S-estimators [24,27,28]. This makes their application rather unattractive in real deformation-monitoring settings.

S-estimators may be useful only under extreme conditions, specifically, when the variance ratio between the two observation groups exceeds 25. Even in this highly hypothetical case, however, their robustness still falls short of the all-observation combination method.

Finally, the MM-estimation approach tested here (defined in Section 2.4 and applied in the experiment of Section 3) also shows limited relevance. The chosen MM variant yields low empirical efficiency under homogeneous conditions and demonstrates weaker robustness than several other methods, including S-estimation with breakdown points above 0.25, the L1-norm, IGG III ($k_1 = 1.5$, $k_2 = 3$), Tukey ($k = 1.547$), and, again, the all-observation combination adjustment, which consistently emerges as the most robust option in the scenarios tested.

5.2. Discussion of real data test results

An examination of the results presented in Figs 4-13, together with the data summarised in Table 3 and their processing using the discussed adjustment methods, allows several observations regarding the performance of the applied techniques.

The 3^n combinations method, Tukey M-estimation with k=1.547, the L1-norm method, and the Danish method yielded the smallest standard errors for the adjusted benchmark heights in the levelling network shown in Fig. 3. The benchmark accuracies obtained by these methods are nearly identical. Among them, the 3^n combinations method produced the most homogeneous distribution of benchmark standard errors across different network datums. In contrast, the L1-norm method exhibited the highest sensitivity to the choice of datum in the real-data example. Nevertheless, the differences in the adjusted heights obtained by these four methods remain below 0.05 mm.

Ordinary Least Squares (OLS) and Weighted Least Squares (WLS) adjustment variants produced almost identical results for both benchmark accuracy and adjusted heights, even though the ratio of the longest levelling distance (R1–R4) to the shortest (R5–R6) is approximately 6.5. The equal-weight OLS variant resulted in a more homogeneous distribution of standard errors compared to the weighted variant. The observed differences between OLS and WLS results are mainly attributable to the choice of the datum benchmark; in the weighted variant, benchmarks closer to the datum exhibit smaller standard errors.

When compared to the 3^n combinations method, Tukey M-estimation (k = 1.547), the L1-norm method, and the Danish method, both OLS and WLS produced standard errors of adjusted benchmark heights that are more than three times larger. Similar results, in terms of both accuracy and adjusted heights, were also obtained using S-estimation with a breakdown point of 25%.

The M-estimation variants of Tukey (k = 4.685), Huber (both variants), Andrew (both variants), and IGG III (both variants), as well as S-estimation with a breakdown point of 10% and MM-estimation, produced mutually consistent results regarding benchmark accuracy and adjusted heights. However, the accuracy achieved by these methods is approximately four times lower than that of the best-performing methods.

S-estimation with a breakdown point of 50% demonstrates improved performance compared to the M- and S-estimation variants. Nevertheless, its accuracy remains inferior to that achieved by the 3^n combinations method and Tukey M-estimation with k = 1.547.

Overall, the real-data example supports the conclusions drawn from the simulation results discussed in Section 5.1. It should be emphasised that real data sets exhibit specific properties, making the identification of a universally optimal processing method difficult, if not impossible. Consequently, it is advisable to employ methods that are more likely to yield reliable results across a wide range of practical situations. Furthermore, adjustment methods should ideally be minimally influenced by subjective parameters, such as the choice of breakdown point or tuning constants in M-estimation.

Based on the results presented in Tables 2-7 and Figs 4-13, as well as the arguments discussed above, general criteria for selecting the most reliable adjustment method for precise levelling networks applied in deformation analyses can be formulated. These criteria are summarised in Table 8.

Table 8. Criteria for selecting the most reliable adjustment approach of precise levelling networks applied for monitoring of object vertical deformations

Criterion	Adjustment method
$(\sigma_{worse}/\sigma_{better})^2 \leq 2$	OLS, Alternatives: Huber (k=2.5), Andrew (k=4.206)
$(\sigma_{worse}/\sigma_{better})^2 \geq 2$	3^n Combinations

The procedure of applying the above criterion is simple and, in fact, a three-step process:

1. Apply the 3^n combinations adjustment and get these values from "the forward", "the backward", or their mean for each line elevation in the adjusted levelling line, which are supposed to have the least true errors. Let the MSE of this adjustment be σ_{better}.
2. Apply an OLS adjustment with the opposite of their observations. If the observation $h_{1,3}$ is estimated to have a smaller true error than $h_{3,1}$, in this step, use $h_{3,1}$. Let the MSE of this adjustment be σ_{worse}.
3. Calculate the ratio $(\sigma_{worse}/\sigma_{better})^2 = k$. If $k \leq 2$, readjust the network using as initial data the means of "the forward" and "the backward" by OLS, or perform either Huber (k=2.5) or Andrew (k=4.206) M-Estimation. If $k > 2$, it is likely that the 3^n combinations adjustment from step 1 is the most reliable decision.

6. Conclusion

This study provides a comprehensive comparison of classical and robust adjustment methods for precise geometric levelling under both simulated and real-data conditions. The results demonstrate that the relative performance of adjustment techniques is strongly governed by the homogeneity of observation variances and the degree of deviation from Gaussian assumptions.

When observations originate from the same distribution with equal variances, ordinary least squares remains the most reliable and efficient solution. This finding is consistent with established theory and prior studies and confirms that robust alternatives offer no tangible benefit under strictly homogeneous, Gaussian conditions. Among the robust methods, only selected M-estimators (Huber's and Andrew's with specific tuning constants) approach the performance of OLS, while most others exhibit a clear loss of efficiency.

In contrast, when variance heterogeneity is present, the advantages of the adjustment method based on all observation combinations become evident. Both the simulation experiments and the real-data test show that this method exhibits superior robustness, as reflected in slower growth of RMSE and a rapid increase in empirical efficiency relative to OLS as variance ratios increase. A key practical advantage of this approach is its independence from subjective parameters, such as tuning constants or assumed contamination levels, making it particularly attractive for real deformation-monitoring applications where such information is rarely known a priori.

The analysis further indicates that the practical relevance of M-estimators in precise levelling is limited. Their dependence on method selection and coefficient tuning, combined with modest performance gains at best, reduces their appeal in operational settings. S-estimators offer only a modest improvement under extreme variance disparities, yet even in these cases, they remain inferior to the all-observation combination method. The tested MM-estimation variant similarly fails to provide a competitive balance between efficiency and robustness.

The real-data experiment corroborates the simulation-based findings. Methods such as the 3^n combinations approach, Tukey M-estimation with a low tuning constant, the L1-norm, and the Danish method yield the highest benchmark accuracies, with the combinations method additionally providing the most homogeneous error distribution across different datum choices. Classical OLS and WLS adjustments, although mutually consistent, yield significantly larger standard errors, underscoring their sensitivity to variance heterogeneity in practice.

Overall, the results highlight that no single adjustment method can be considered universally optimal for all levelling scenarios. Nevertheless, methods that demonstrate stable performance across a wide range of conditions and that minimise reliance on subjective parameter choices are preferable. In this respect, the adjustment based on all observation combinations emerges as the most robust and practically reliable option for monitoring vertical displacements using precise geometric levelling.

References

[1] Gera O., Oleskiv R., Dorosh L., Hrynishak M., Mykhailyshyn V. Geodetic monitoring of the subsidence of a multi-storey building foundation. *Budownictwo i Architektura* 24(3) (2025) 005–014. https://doi.org/10.35784/bud-arch.6827

[2] Borowski L., Kudrys J., Kubicki B., Slámová M., Maciuk K. Phase centre corrections of GNSS antennas and their consistency with ATX catalogues. *Remote Sensing* 14(13) (2022) 3226. https://doi.org/10.3390/rs14133226

[3] Borowski L., Banasik P., Maciuk K. Application of GNSS-levelling for updating the base vertical network. *Journal of Applied Geodesy* 19(4) (2025) 595-606. https://doi.org/10.1515/jag-2024-0096

[4] Maciuk K., Szombara S. Annual crustal deformation based on GNSS observations between 1996 and 2016. *Arabian Journal of Geosciences* 11 (2018) 667. https://doi.org/10.1007/s12517-018-4022-4

[5] Pleterski Ž., Ambrožič T., Mulahusić A., Tuno N., Topoljak J., Hajdar A., Hamzić A., Điđelija M., Kulo N., Rak G., Marjetič A., Kregar K. Kvadratna metoda Msplit v deformacijski analizi na primeru 2D geodetske mreže. Squared Msplit Estimation in Deformation Analysis – 2D Geodetic Network Case Study. *Geodetski vestnik* 69(2) (2025) 115-147. https://doi.org/10.15292/geodetski-vestnik.2025.02.115-147

[6] Wyszkowska P., Duchnowski R. Systematic bias of selected estimates applied in vertical displacement analysis. *Journal of Geodetic Science* 10(1) (2020) 41-47. https://doi.org/10.1515/jogs-2020-0103

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

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

[9] Mickrenska C., Cvetkov V. Multiple regression in help of the precise levelling. World Summit: Civil Engineering-Architecture-Urban Planning Congress (CAUSummit 2024) ARPHA Proceedings 7, edited by Isik Yilmaz et al., Pensoft Publishing, Sofia, 2024, 65-71. https://doi.org/10.3897/ap.7.e0065

[10] Cvetkov V. An algorithm for adjustment of geometric levelling networks. *Inżynieria Mineralna - Journal of the Polish Mineral Engineering Society* 1(1) (2024) 187–194. https://doi.org/10.29227/IM-2024-01-20

[11] Cvetkov V. Averages, entropy and observations. *Deutsche Internationale Zeitschrift für Zeitgenössische Wissenschaft* 93 (2024) 52–57. https://doi.org/10.5281/zenodo.14289165

[12] Mickrenska C., Antova G. Adjustment of a precise levelling network without using the means of line elevations. *ARPHA Proceedings* 9 (2025) 133-138. https://doi.org/10.3897/ap.9.e0133

[13] Wiśniewski Z., Duchnowski R., Dumalski A. Efficacy of msplit estimation in displacement analysis. *Sensors* 19(22) (2019) 5047. https://doi.org/10.3390/s19225047

[14] Duchnowski R., Wyszkowska P. Robust procedures in processing measurements in geodesy and surveying: a review. *Measurement Science and Technology* 35(5) (2024) 1-16. https://doi.org/10.1088/1361-6501/ad28ec

[15] Duchnowski R., Wyszkowska P. Empirical influence functions of different robust estimation methods applied in displacement analysis", in *Proceedings - 4th Joint International Symposium on Deformation Monitoring* (JISDM), 15-17 May 2019, Athens, Greece. Available: https://www.fig.net/resources/proceedings/2019/04_JISDM2019/96.pdf [Accessed: 13 Nov 2025]

[16] Borowski Ł., Banaś M. Application of robust estimation in polynomial modelling", in *Proceedings - 2018 Baltic Geodetic Congress, BGC-Geomatics 2018*, 2018 62–66. https://doi.org/10.1109/BGC-Geomatics.2018.00018

[17] Banaś M., Ligas M. Empirical tests of performance of some M-estimators. *Geodesy and Cartography* 63(2) (2014) 127-146. https://doi.org/10.2478/geocart-2014-0010

[18] Yang Y. Robust estimation of geodetic datum transformation. *Journal of Geodesy* 73 (1999) 268–274. https://doi.org/10.1007/s001900050243

[19] Valero J., Moreno S. Robust estimation in geodetic networks. *Fisica de la Tiera* 17 (2005) 7-22.

[20] Huber P.J. Robust estimation of a location parameter. *The Annals of Mathematical Statistics* 35(1) (1964) 73-101. https://doi.org/10.1214/aoms/1177703732

[21] Hampel F., Ronchetti E., Rousseeuw P., Stahel W. *Robust statistics: The approach based on influence functions*. New York - Chichester - Brisbane - Toronto - Singapore, John Wiley & Sons, 1986, p. 502

[22] Andrews D.F. A robust method for multiple linear regression. *Technometrics* 16(4) (1974) 523–531. https://doi.org/10.1080/00401706.1974.10489233

[23] Kwaśniak M. Effectiveness of chosen robust estimation methods compared to the level of network reliability. *Geodesy and Cartography* 60(1) (2011) 3-19. https://doi.org/10.2478/v10277-012-0014-9

[24] Borowski Ł., Banaś M. Application of robust estimation in polynomial modelling. *Baltic Journal of Modern Computing* 7(4) (2018) 525–540. https://doi.org/10.22364/bjmc.2019.7.4.06

[25] Susanti L., Pratiwi H., Sulistijowati S., Liana T. M estimation, S estimation, and MM estimation in robust regression. *International Journal of Pure and Applied Mathematics* 91(3) (2014) 349-360. https://doi.org/10.12732/ijpam.v91i3.7

[26] Ismail M., Rasheed H. Robust regression methods / a comparison study. *Turkish Journal of Computer and Mathematics Education* 12(14) (2021) 2939-2949.

[27] Çetin M., Toka O. The comparing of S-estimator and M-estimators in linear regression. *Gazi University Journal of Science* 24(4) (2011) 747-752.

[28] Rousseeuw P., Yohai V. Robust regression by means of S-estimators. In: Franke, J., Härdle, W., Martin, D. (eds) Robust and nonlinear time series analysis. *Lecture Notes in Statistics* 26 (1984) 256-272, Springer, New York, NY. https://doi.org/10.1007/978-1-4615-7821-5_15

[29] Yohai V. High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics* 15(2) (1987) 642-656. https://doi.org/10.1214/aos/1176350366

[30] Zrinjski M., Tupek A., Barković Đ., Cesarec M. Stability analysis of a levelling network as a prelude to a residential building construction. *Journal of Faculty of Civil Engineering*, Conference proceedings of the 8th International Conference "Contemporary achievements in civil engineering 2021", 2021 79-89. https://doi.org/10.14415/konferencijaGFS2021.07