https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions
Open Access

Minimal variance-based outlier detection method using forward search model error in a leveling network

Utkan M. Durdağ[1]

[1]Geomatics engineering, Artvin Coruh University, Artvin, 08100, Türkiye

Correspondence to: Utkan M. Durdağ (umdurdag@artvin.edu.tr)

Abstract. Conventional and robust methods are based on the additive bias model, which may cause type-I and type-II errors. However, outliers can be regarded as additional unknown parameters in the Gauss-Markov Model. It is based on modeling the outliers as unknown parameters, considering as many combinations as possible outliers selected from the observation set. In addition, this method is expected to be more effective than conventional methods as it is based on the principle of minimal variance and removes dependency in iterations. The primary purpose of this study is to seek the novel outlier detection approach efficiency in the geodetic networks. The efficiency of the proposed model was measured and compared with the robust and conventional methods by the Mean Success Rate (MSR) indicator for different types and magnitudes of outliers. Thereby, this approach enhances the MSR by almost 40-45% compared to the Baarda and Danish (with the variance unknown case) method for multiple outliers (i.e., $1<m<4$). Besides, the Forward Search of Model Error (FSME) is 20-30% more successful than the others in the low controllability observations of the leveling network.

1 Introduction

Conventional tests for outliers and robust M-estimation are based on the Least Squares Estimation (LSE). If an observation contains an outlier, the LSE method ceases to be the optimal estimation method in terms of a minimum variance unbiased estimator. Once gross errors are detected and isolated, the LSE can be called an efficient estimation. Otherwise, an undetected outlier has a slight deviation from the normality assumption that may cause a smearing effect on all estimation parameters regardless of whether using LSE directly or indirectly which may be named the local influence function of LSE (Gao et al. 1992; Hekimoglu et al. 2010; Nowel 2020). For different bias intervals, the smearing effect of LSE that behaves systematically as a function of the partial redundancy has been proven by Durdag et al, (2022). Normalized residuals, which would be exposed to the smearing effect, are investigated to identify and isolate outliers by conventional tests for outliers and some robust methods such as M-estimation (Zienkiewicz and Dąbrowski 2023; Wang et al. 2021; Batilović et al. 2020). Thereby the falsified test result may induce type error-I. In addition to the unreliability of LSE results, the low success of the F-test shown by Hekimoglu led researchers to seek a more reliable and effective method such as the univariate method, original observations, etc. (Hekimoglu 1999; Erdogan 2014; Hekimoglu et al. 2014). Although these novel methods boost the reliability of the conventional methods, the identification of outliers in these approaches is based on the same procedure

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

as the conventional and robust methods. If the normalized residual exceeds three times its standard deviation (SD), also called the 3-sigma rule, an observation is flagged as an outlier (Lehmann 2013). However, tests for outliers can be dealt with a single outlier sufficiently since the LSE has an unbounded IF (Duchnowski 2011; Maronna et al. 2006; Huber 1981; Durdag et al. 2022). Studies show that the reliability of these techniques, established with the additive bias model, decreases significantly as the number of outliers increases. In the decision stage, the outliers that mask or swamp other observations can produce a type-I error (false negative) and type-II error (false positive). Multiple outliers can be identified at most the number of possible outliers by repetitive test procedures. However, the efficiency of conventional tests is rather small when the outlier value is close to the critical value named as small outliers lies between 3-6σ.

If the rate of successful detection of an outlier using conventional and robust methods is 50%, and one outlier is determined incorrectly, the probability of correctly determining two outliers remains below 50%. This condition is based on the interdependence of each iteration. Incorrect determination at each step also reduces the possibility of identifying more than one outlier in the next step. Therefore, besides modeling the outliers as unknown, the proposed method is based on two essential factors: the principle of the slightest variance and the assumption of looking at all points with suspicion in each iteration. It has been proven by Hekimoglu et al. (2015) for linear regression that the method in which outlier is modeled as an additional unknown gives more successful results than the conventional method. The method suggests carrying outlier detection until all possible combinations are investigated. In the C_k^n combination, observation(s) is included as an additional unknown parameter in the proposed model. Then, observations are viewed with suspicion considering combinations of k elements (groups of two, three, and so forth) selected from a set of n elements (C_k^n), where n is the number of observations, and k denotes the number of outliers. The observation with the smallest variance among C_1^n combinations is determined. Considering the C_2^n combinations, the pair of observations were regarded as a model error, and the two observations with the smallest variance were flagged as candidates. All possible combinations will be regarded until as much as the maximum number of burdened observations that would occur up to one-half of the degrees of freedom ($\cong f/2$) for the geodetic network. The potential observations are clustered separately and compared with the specified critical values for each combination step. The model errors of the potential outlier(s) exceeding the critical value were flagged as suspicious for each combination pace. The test values of all potential outliers must exceed the critical value for each combination step, and if not, the previous candidates are detected as outliers.

The primary purpose of this study is to apply seek the proposed outlier detection method efficiency in geodetic networks. The suggested model was compared with the robust methods by the Mean Success Rate (MSR) indicator for different types and magnitude of outliers. As in the classic approaches, the number of outliers is inversely proportional to the success of the presented method. When outliers have various magnitudes (e.g., small, large, gross, and extreme outliers) and specific observations are not available in the network (observations with low controllability), it has been found that the proposed method is quite successful compared to the conventional and robust methods.

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions

2 Gauss-Markov model

Let $\mathbf{A}_{\mathrm{nxu}}$ be a design matrix and has full column rank, i.e. rank (A)=u and $\mathbf{P}$ a positive definite weight matrix of the observations, $\mathbf{x}_{\mathrm{ux1}}$ a vector of the unknown parameter, $\mathbf{l}_{\mathrm{nx1}}$ an observation vector, $\mathbf{C}_{ll_{nxn}}$ an a priori covariance matrix of observations, $\mathbf{Q}_{ll_{nxn}}$ a weighted coefficient matrix of observations and σ_0^2 an a priori variance factor, where n and u a number of observation and number of unknowns, respectively. By adding $\mathbf{v}_{\mathrm{nx1}}$ a residual vector, one can get $\hat{\mathbf{x}}$ an estimated vector of unknown parameters presented in the following Gauss-Markov model (Koch, 1999)

$$\mathbf{l} + \mathbf{v} = \mathbf{A}\hat{\mathbf{x}}\,; \quad \mathbf{C}_{ll} = \sigma_0^2\mathbf{P}^{-1} = \sigma_0^2\mathbf{Q}_{ll}. \tag{1}$$

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l} \tag{2}$$

$$\mathbf{Q}_{xx} = (\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A})^{+} \tag{3}$$

$$\mathbf{Q}_{\mathrm{vv}} = \mathbf{P}^{-1} - \mathbf{A}\mathbf{Q}_{\mathrm{xx}}\mathbf{A}^{\mathrm{T}} \tag{4}$$

Where $\mathbf{Q}_{xx}$ denotes a cofactor matrix of the unknown parameters, $\mathbf{Q}_{vv}$ implies the cofactor matrix of the residuals.

2.1 Test for outliers

In Geodesy, procedures for the outlier detection were developed by Baarda (1968) and Pope (1976). If the observations come from the normal distribution, it is called good observations whereas the burdened observations that contains outlier originate from another distribution. Let l_i be a burdened observation has δl_i an outlier, the following hypothesis

$$H_0\colon\ \delta l_i = 0 \qquad \text{against} \qquad H_1\colon\ \delta l_i \neq 0 \tag{5}$$

is tested. If the observations are uncorrelated and the variance σ_0^2 is known, the normalized residuals can be written as

$$\tau_{i,B} = \frac{|v_i|}{\sigma_0\sqrt{q_{v_iv_i}}} \tag{6}$$

where τ_i is the test value and q_{vv} is the cofactor of the residual for i=1…n. This is known as Baarda's method (i.e. data-snooping test). A posteriori variance (m_0^2) is calculated in Pope's method given by

$$\tau_{i,P} = \frac{|v_i|}{m_0\sqrt{q_{v_iv_i}}}. \tag{7}$$

The observation with the biggest normalized or studentized residual is tested in one loop of the iterations. Test for outliers are used iteratively if the observations contain more than one outlier. The flagged observation is removed when H_0 is rejected. The remaining observations are adjusted once more. Until no more outliers are found, this process is repeated. However the multiple outliers cause swamping or masking effects that make it impossible to distinguish the burdened

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions
Open Access

observations from the good ones. In the following sections; the robust and the proposed methods will be demonstrated to prevent the smearing effect of LSE.

2.2 Robust methods

M-estimation (Huber, 1964) is a generalized form of maximum likelihood estimation**.** In this paper M-Estimation of Huber and Danish methods, commonly chosen to handle outliers in robust statistics, were used to compare the results of the proposed method.

2.2.1 M-estimation

Re-weighted LSE is applied iteratively to the non-linear normal equation of the M-estimation as follows:

$$\hat{\mathbf{x}}^r = (\mathbf{A}^{\mathbf{T}}\overline{\mathbf{W}}^r\mathbf{A})^+\mathbf{A}^{\mathbf{T}}\overline{\mathbf{W}}^r\mathbf{l} \tag{8}$$

$$\overline{\mathbf{W}}^r = \mathbf{P}\mathbf{W}_{(\bar{\mathbf{v}}^{r-1})}, \tag{9}$$

$$\mathbf{W}_{(\bar{\mathbf{v}}^0)} = \mathbf{E} \tag{10}$$

$$\bar{\mathbf{v}}^r = \mathbf{A}\hat{\mathbf{x}}^r - \mathbf{l} \tag{11}$$

$$\mathbf{W}(\bar{\mathbf{v}}) = diag(\,(W(\bar{v}_1), W(\bar{v}_2), \dots, W(\bar{v}_n)) \tag{12}$$

where $\hat{\mathbf{x}}^k$ equals the $\hat{\mathbf{x}}$ from the Eq. 2 for the first iteration, $\mathbf{E}$ stands for a unit matrix. r implies a number of iterations and is chosen as 5 in this paper. The weight function of Huber's M-estimation is given as follows

$$W(\bar{v}_i^r) = \begin{cases} 1 & |\bar{v}_i^r| \le c \\ \frac{c}{|\bar{v}_i^r|} & |\bar{v}_i^r| > c \end{cases} \quad ; \quad \text{i=1…n} \tag{13}$$

and the weight function of the Danish method is given by

$$W(\bar{v}_i^r) = \begin{cases} 1 & |\bar{v}_i^r| < c \\ \exp\left(-\frac{|\bar{v}_i^r|}{c}\right) & |\bar{v}_i^r| \ge c \end{cases} \quad ; \quad \text{i=1…n} \tag{14}$$

where $\bar{v}_i$ is the residual and c is taken as $1.5\sigma_0$. After the diagonal elements of the $\overline{\mathbf{W}}$ weight matrix are determined, $\bar{\mathbf{v}}^r$ and $\hat{\mathbf{x}}^r$ are recalculated for each iteration. The residual that is computed at the final iteration is detected as an outlier if it exceeds 3σ.

3 Forward search of model error

The Gauss-Markoff model (1) is now expanded by the u x 1 vector $\boldsymbol{\epsilon}$ of additional unknown parameters also with the n x u design matrix $\mathbf{M}$

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development
Discussions

$$\mathbf{l}+\mathbf{v}=[\mathbf{A}\;\mathbf{M}]\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\boldsymbol{\epsilon}}\end{bmatrix};\quad \mathbf{C}_{ll}=\sigma^2\mathbf{P}^{-1}=\sigma^2\mathbf{Q}_{ll} \tag{15}$$

where the variance σ^2 stands for the unit weight of the augmented model and the vector $\boldsymbol{\epsilon}$ contains the outliers which are subtracted from the observations. If only the outlier Δ_j is present in the observation l_j, then one should define $\boldsymbol{\epsilon}=\Delta_j$ and $\mathbf{M}=\mathbf{e}_j$ where $\mathbf{e}_j=[0,\ldots,0,1,0,\ldots,0]$ for $j=1\ldots n$. The j_{th} component of $\mathbf{e}_j$ gets the value one. For the j_{th} observation with $\mathbf{A}=[\mathbf{A_1},\ldots,\mathbf{A}_j,\ldots]^{\boldsymbol{T}}$ the observation equation given as

$$\mathbf{l}_j+\mathbf{v}_j=\mathbf{A}_j^T\hat{\mathbf{x}}+\hat{\Delta}_j \tag{16}$$

where $\mathbf{A}_j^T$ is the j_{th} row vector of $\mathbf{A}$ and for the remainder of the observations $\mathbf{l}_k+\mathbf{v}_k=\mathbf{A}_k^T\hat{\mathbf{x}}$ $(k=1,2,\ldots,n)$, $k\neq j$. If the outliers exist in the observations $\boldsymbol{\epsilon}$ and $\mathbf{M}$ are rewritten as follows

$$\boldsymbol{\epsilon}=\left|\hat{\Delta}_j,\hat{\Delta}_{j+1}\ldots,\hat{\Delta}_t\right|^T \text{ and } \mathbf{M}=\left[\mathbf{e}_j,\mathbf{e}_{j+1},\ldots,\mathbf{e}_t\right]^T. \tag{17}$$

The estimated of unknown parameters of the augmented model can, therefore, be expressed as follows (Koch, 1999):

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\boldsymbol{\epsilon}}\end{bmatrix}=\begin{bmatrix}\mathbf{A^TPA} & \mathbf{A^TPM}\\ \mathbf{M^TPA} & \mathbf{M^TPM}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A^TPl}\\ \mathbf{M^TPl}\end{bmatrix} \tag{18}$$

where

$$\begin{bmatrix}\mathbf{A^TPA} & \mathbf{A^TPM}\\ \mathbf{M^TPA} & \mathbf{M^TPM}\end{bmatrix}^{-1}=$$

$$\begin{bmatrix}(\mathbf{A^TPA})^{-1}(\mathbf{E}+\mathbf{A^TPMSM^TPA}(\mathbf{A^TPA})^{-1} & -(\mathbf{A^TPA})^{-1}\mathbf{A^TPMS}\\ -\mathbf{SM^TPA}(\mathbf{A^TPA})^{-1} & \mathbf{S}\end{bmatrix} \tag{19}$$

$$\mathbf{S}=[\mathbf{M^T}(\mathbf{P}-\mathbf{PA}(\mathbf{A^TPA})^{-1}\mathbf{A^TP})\mathbf{M}]^{-1}=(\mathbf{M^TPQ}_{vv}\mathbf{PM})^{-1} \tag{20}$$

$$\hat{\boldsymbol{\epsilon}}=\mathbf{SM^TP}(\mathbf{E}-\mathbf{A}(\mathbf{A^TPA})^{-1}\mathbf{A^TP})\mathbf{l} \tag{21}$$

The residuals are expressed for the Gauss-Markov model in Eq.(1) by

$$\mathbf{v}=\mathbf{A}\hat{\mathbf{x}}-\mathbf{l}=\mathbf{A}(\mathbf{A^TPA})^{+}\mathbf{A^TPl}-\mathbf{l}=(\mathbf{E}-\mathbf{A}(\mathbf{A^TPA})^{+}\mathbf{A^TP})(-\mathbf{l}) \tag{22}$$

whose right-hand side can be replaced in Eq. 21 as follows

$$\hat{\boldsymbol{\epsilon}}=\mathbf{SM^TP}(\mathbf{E}-\mathbf{A}(\mathbf{A^TPA})^{-1}\mathbf{A^TP})\mathbf{l}=-\mathbf{SM^TPv} \tag{23}$$

and considering Eq. 20 the following equation yields

$$\hat{\boldsymbol{\epsilon}}=-(\mathbf{M^TPQ}_{vv}\mathbf{PM})^{-1}\mathbf{M^TPv}. \tag{24}$$

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

3.1 Testing Procedure

The alternative hypothesis, in the case presence of outliers, takes the form against the null hypothesis as follows:

$$H_0 : E\{l\} = \mathbf{A}\hat{\mathbf{x}} \tag{25a}$$

$$H_A : E\{l\} = [\mathbf{A}\ \mathbf{M}]\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\boldsymbol{\epsilon}}\end{bmatrix}. \tag{25b}$$

One should consider all possible combinations of potentially burdened observations for the correct specification of the alternative hypothesis (Teunissen 2006). All potential alternative hypotheses C_b^n, where n is the number of observations, and b is the number of potential outliers, are considered in the detection step. Firstly, the observations are assumed to be unknown one by one in the model. The additional unknowns of the model $\hat{\boldsymbol{\epsilon}}$ are calculated by rewriting the relevant rows for each observation in the coefficient matrix iteratively. The design matrix can be rewritten as follows by including a dimension in the model as an unknown:

$$\mathbf{A}^{\mathbf{1,1}}_{C_1^n} = [\mathbf{A}\ \mathbf{M}] = \left[\begin{array}{ccccccc:c} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1\\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0\\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \end{array}\right], \mathbf{A}^{\mathbf{1,2}}_{C_1^n} = \left[\begin{array}{ccccccc:c} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1\\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \end{array}\right] \tag{26}$$

where $\mathbf{A}^{b,i}$ denotes the matrix of coefficients for b=1,..,f/2 and i=1,…,n.

3.1.1 Calculation steps for model error

The rows of the additional column vector are rewritten iteratively for each observation, and the corresponding one is modeled as an unknown using calculation steps given below.

1. After calculating the cofactor matrix, the unknowns matrix is obtained:

$$\boldsymbol{Q}_{xx}^{\mathrm{b}} = (\mathbf{A}^{\mathrm{b}^{\boldsymbol{T}}}\mathbf{P}\mathbf{A}^{\mathrm{b}})^{+} \tag{27}$$

$$\hat{\mathbf{x}}^{\mathrm{b}} = \left(\mathbf{A}^{\mathrm{b}^{\mathbf{T}}}\mathbf{P}\mathbf{A}^{\mathrm{b}}\right)^{+}\mathbf{A}^{\mathrm{b}^{\mathbf{T}}}\mathbf{P}\mathbf{l}. \tag{28}$$

2. To determine the observation that gives the smallest variance value, the step of calculating the residuals is given by

$$\mathbf{v}^{\mathrm{b}} = \mathbf{A}^{\mathrm{b}}\hat{\mathbf{x}}^{\mathrm{b}} - \mathbf{l}. \tag{29}$$

3. The posteriori variance is calculated as

$$(\mathrm{s}^{\mathrm{b}})^2 = \sqrt{\mathbf{v}^{\mathrm{b}^{\mathbf{T}}}\mathbf{P}\mathbf{v}^{\mathrm{b}} / f^{\mathrm{b}}}\,. \tag{30}$$

4. Determining the observation with minimum variance

$$\mathrm{j} = \min(\mathrm{s}^{\mathrm{b}})^2 \tag{31}$$

5. After the relevant observation is determined, the test value is calculated as given by

$$T = \hat{\Delta}_j / (\mathrm{s}_{\mathbf{0}}\sqrt{\boldsymbol{q_{jj}}}). \tag{32}$$

Thus, the unit-weighted posteriori variances for each additional unknown parameter are calculated given by

$$\hat{s}^2 = \frac{\mathbf{v}^{\mathbf{T}}\mathbf{P}\mathbf{v}}{n-u_k}\ ; i = 1 \dots n \tag{33}$$

where $u_k = u + 1$ represents the number of the unknowns calculated for the model given in Eq. 15. The number of elements in the set of the posterior variances calculated for each observation appears as C_1^n. After the acceptance or rejection of the H_0 hypothesis is evaluated in the identification phase mentioned below, the decision is made to rewrite the model, where the unknowns are expanded for the observations two by two for the C_2^n combination. The smallest variance value $min\ \{\hat{\sigma}_i^2\}$ belongs to which observation is identified and the unknown of the relevant observation compared with the critical value. When $min\ \{\hat{\sigma}_i^2\} = \hat{\sigma}_k^2$, the absolute value of Δ_k is compared with the t-test. If $|\Delta_k| \geq t_{f-1,1-\alpha}$, H_0 is rejected and k_{th} observation is flagged as an outlier. If the null hypothesis is accepted, the process ends. The model is expanded for another alternative hypothesis which assume two potential blunder in case H_0 is rejected. The coefficients matrix is rewritten for each combination of C_2^n given by

$$\mathbf{A}_{C_2^n}^{\mathbf{2,1}} = \left[\begin{array}{ccccccc:cc} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \end{array}\right], \mathbf{A}_{C_2^n}^{\mathbf{2,2}} = \left[\begin{array}{ccccccc:cc} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \end{array}\right]. \tag{34}$$

An important point to be emphasized here is; that all combinations are taken into account independently of the previous result (i.e. regardless of the biased observation flagged in the previous step). For example; all potential C_2^n combinations are considered, neglecting the previous result where the k_{th} observation was flagged. The absolute value of model errors (Δ_i, Δ_j) for i=1…n and j=1…n and $i \neq j$, which have the smallest variance, are compared with the $t_{f-1,1-\alpha}$ threshold value where $\alpha = 0.05$. If both are greater than the critical value, the relevant observations are flagged as outliers It is sought for C_3^n possible combinations, and the coefficient matrix rewritten as follows:

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development
Discussions

$$\mathbf{A}_{C_3^n}^{3,1} = \left[\begin{array}{ccccccc:ccc} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \end{array}\right], \quad \mathbf{A}_{C_3^n}^{3,2} = \left[\begin{array}{ccccccc:ccc} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \end{array}\right], \qquad (35)$$

whether the model errors of the observations that give the smallest variance value are higher than the critical value or not. If all three values of unknowns exceed the critical value, they are flagged as outliers. This process is repeated for four or more combinations until all the combinations of potentially burdened observations have been considered. The $\hat{\boldsymbol{\epsilon}}$ vector of the observations corresponding to the minimum variance value calculated for each combination step is compared with the critical value. If at least one of the relevant unknowns of the observations does not exceed the critical value, the H_0 hypothesis is accepted and the observations flagged in the previous step (i.e. the latest rejected H_0) are approved as outliers. The flowchart of the FSME (Forward Search of Model Error) approach is presented in Fig.1.

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions

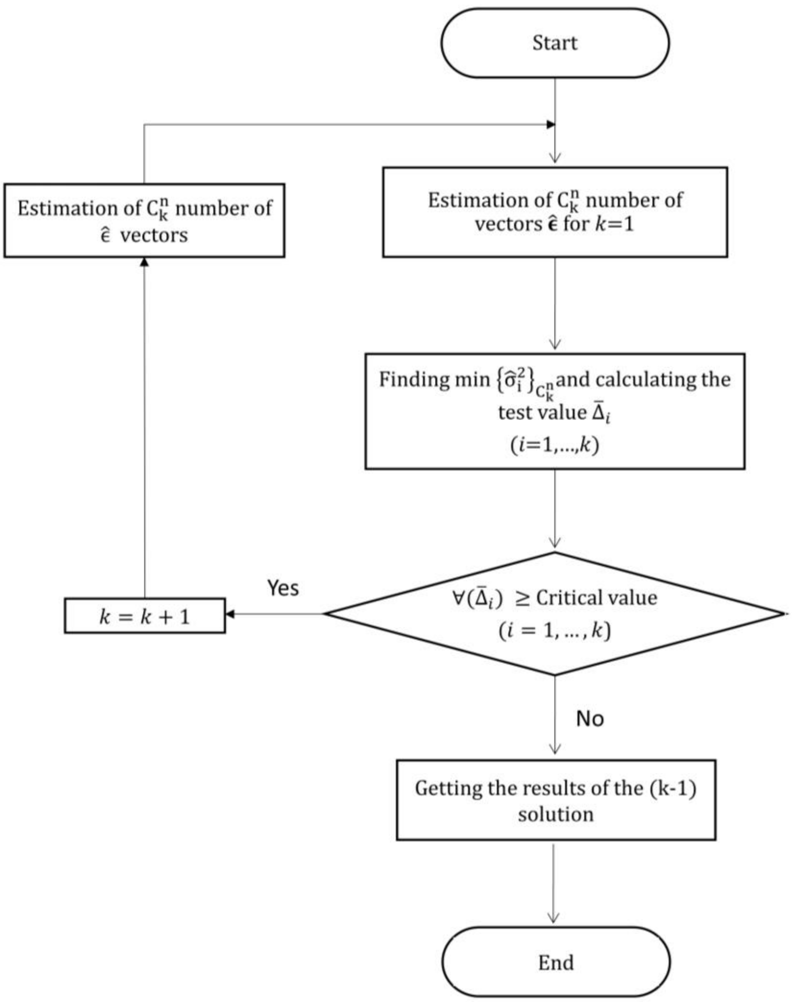


Figure 1: Flowchart of the forward search of model error

4 Leveling network

In statistics, there are different indicators to measure the reliability of tests and estimators. Hekimoglu and Koch (2000) have shown that the global reliability of an estimator and a test procedure were determined by finite-sample breakdown point. Using the power function of the global test, a capacity in deformation networks is explored as suggested by Niemeier (1985). Also, it has been shown that the MSR results of the two testing procedures (chi-square and f-test) are identical to their respective test powers known beforehand (Aydın, 2012). MSR depends on the number of outliers, the magnitude of an outlier, the number of unknowns, the number of observations, and the type of outliers. Since it considers these different cases, MSR is more reliable, whereas the power of the test is the same for all disparate conditions. Also, Erdogan et al.

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions
Open Access

(2019) have proven that the MSR is the empirical estimation of the power of the test in outlier detection. In this study, therefore, MSR is used to specify the ability of the conventional, robust, and proposed approaches. By this purpose three different leveling networks have been simulated. The random errors ε_i for i=1…n, were generated using a normal distribution N(0, σ^2) with a mean of zero and a variance of σ^2. Also, the good and biased observations were acquired by simulation technique as described in detail by Hekimoglu and Erenoglu (2007). Since the outliers are produced through simulation, it is easy to determine whether an observation is burdened before analyzing. The method is deemed successful if the observation recognized as an outlier matches the really burdened observation. The process is considered unsuccessful if it fails. When the simulated observation is chosen randomly, the successful rate indicates the global MSR and the local MSR can be computed for each particular observation in the leveling network for 10 000 samples. The same samples were subjected to conventional, robust, and novel methods to compare their MSRs with different scenarios. This study simulated outliers randomly chosen from small and large magnitudes outliers (variously described gross and influential outliers) for three leveling networks. A leveling network used for the simulation has 7 points and 15 observations as seen in Fig. 2. The precision is considered to be $\sigma_i^2 = \sigma_0/\sqrt{S}$ where S is the length of the leveling line in km and $\sigma_0 = 1mm/\sqrt{1km}$. MSRs for 10 000 samples were calculated for each method when there were different magnitudes and different numbers of outliers in the network. The small and large outliers were generated in the intervals of [3-6σ] and [6-12σ], respectively.

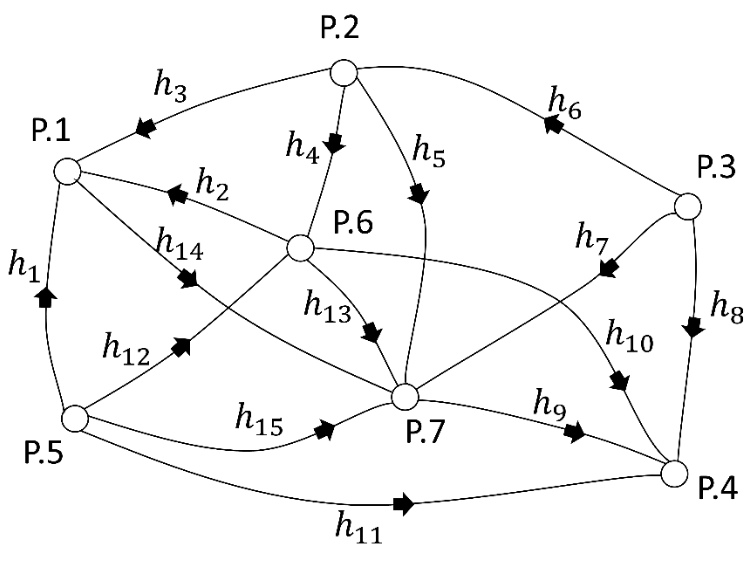


Figure 2: Leveling network

As Table I shows, even if the number of outlier changes, the MSRs of the proposed method increases significantly compared to the conventional and robust methods. In cases where there is no outlier (e.g. m=0), the results, in which the H_0 hypothesis is rejected, are also seen in Table 1. The proposed method generated type-2 error at the rate of 5%, where the significance level was at 0.05.

Table 1: MSR of models (small outliers)

m	Baarda	Pope	Danish		Huber		FSME
			*	**	*	**	
0	99.99	95.96	85.00	94.97	96.99	99.00	95.00
1	56.71	36.70	69.76	72.44	63.41	52.93	88.78
2	24.48	2.32	49.26	27.21	38.45	19.92	70.40
3	7.86	0.04	29.58	6.32	20.25	10.36	46.15
4	1.26	0.00	15.27	2.22	8.93	5.20	21.17

Two cases which the variance is known and unknown were considered for Robust methods as follows:

* ***The A priori variance is known:*** For the robust techniques, c was taken to be 1.5. When the residual from the robust techniques exceeded the 3σ threshold value, it was regarded as an outlier. In the case where the A priori variance is known, it can be seen in Table 1 that the MSRs of the robust methods are higher than the Baarda test with α considered by 0.001. Pope's test had a lower MSR than Baarda's did. However, the MSRs of the FSME (Forward Search of Model Error) are higher than the robust methods in both cases where the a priori variance is known and unknown.

** ***The A priori variance is unknown:*** The standard deviation from the first iteration (LSE) was obtained for robust methods. The α was chosen as 0.05 for the Pope's test, which had a lower MSR than Baarda's. Except for the Danish* method, all other approaches identified an excellent observation as an outlier with a risk ranging from 0.01% to 5% if there was no outlier in the observations. The a posteriori variance negatively impacted the robust method's results, and the outlier's spoilt variance significantly contributed to the false detection. The a priori variance significantly impacts how reliable the procedures are.

Additionally, the a posteriori variance is easily influenced by outliers in the data set, which harms the abilities of methods that use the a posteriori variance. The a posteriori variance from LSE is typically utilized as a threshold value instead of the a priori variance if the a priori variance is unknown. Therefore, the MSRs of the robust techniques of the former case are higher than the latter. As a result of these findings, only the case where the variance is known, which is less affected by an outlier, is taken into account in the results shown in the tables (Tables 2-8) to compare with FSME hereafter.

5 Results

Extensive experiments have been done, for comparing the proposed method with Robust methods, such as Danish and Huber methods, besides the conventional outlier detection procedures (i.e. Baarda and Pope). The redundancies are an important indicator to recognize the most vulnerable observations to bias (Durdag, 2020). The redundancy matrix is calculated from $\mathbf{R} = \mathbf{H} - \mathbf{E}$ where $\mathbf{H} = \mathbf{A}(\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^{\mathbf{T}}\mathbf{P}$ is a hat matrix. The local MSRs have been calculated for the specific observations

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions

with the highest and lowest redundancy in the leveling network. Among the observations, those with the two largest redundancies are h_{13} and h_9, and the three lowest are h_1, h_7 and h_8. As can be seen from the table below MSRs increase as the redundancy does.

Table 2: Local MSRs (small outliers)

h_i	Baarda	Pope	Danish	Huber	FSME	Redundancy
h_1	45.47	27.78	54.16	45.84	84.14	0.50
h_7	43.53	30.66	51.66	42.65	80.07	0.50
h_8	41.76	27.48	55.02	39.46	80.28	0.48
h_9	66.52	42.23	80.04	77.40	93.33	0.69
h_{13}	68.08	45.02	81.97	80.47	94.69	0.71

It is apparent from Table 2 that the highest MSRs for the biased observation h_1 amongst the conventional and robust methods is Danish 54%. In addition, the MSR of the proposed method is higher than the Danish by 30%. The highest MSR has been obtained by FSME as %94 for the observation with highest redundancy h_{13}. As the redundancy gets smaller, the difference in MSR between the proposed method and other methods increases.

Table 3: MSR of models (large outliers)

m	Baarda	Pope	Danish	Huber	FSME
1	99.50	90.97	91.46	94.69	99.92
2	92.66	19.64	82.95	77.77	94.11
3	74.57	0.27	68.44	51.76	78.22
4	44.31	0.01	48.60	29.96	50.16

As shown in Table 3, the highest MSRs are obtained by FSME in contrast with other techniques for different numbers of outliers. When Tables 1 and 3 are compared, the MSRs increase with the enlargement in the magnitude of outlier.

The smearing effect of LSE, almost equivalent to its SC (Sensitivity Curve), behaves systematically as a function of the partial redundancy (Durdag, 2021). For this reason, the MSRs have been calculated for the pair of observations with the lowest and largest partial redundancy with small outliers in Table 4. The neighboring observations, especially the point that has three leveling lines, are one of the most vulnerable to bias (e.g. h_6, h_7 and h_8, h_7) in the leveling network (Fig. 2). The local MSRs are lower than the global MSRs, in case m=2 in Table 1, for ones with both lowest and highest redundancies as shown in Table 4.

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development
Discussions

Table 4: The effect of large and low partial redundancies on MSRs for pair of observations

$m = 2$	Baarda	Pope	Danish	Huber	FSME	Redundancy
h_6, h_7	0.67	4.45	20.95	27.39	30.47	0.21
h_8, h_7	0.16	2.68	16.44	24.02	25.51	0.30
h_{11}, h_{15}	12.73	3.22	28.28	28.54	54.37	0.15
h_{10}, h_3	33.01	0.76	54.30	37.98	80.17	0.00
h_5, h_{11}	28.20	0.88	41.77	29.23	75.80	0.00
h_{13}, h_1	30.78	1.53	50.94	34.97	80.53	0.00

The results, as shown in Table 4, indicate that the MSRs of observations h_6, h_7 with the highest partial redundancies increases compared to h_{13}, h_1 with the lowest ones by almost 30% for Baarda and Danish approaches, and 50% for FSME approach.

Table 5: MSRs for gross and influential outliers

Scenario	m	Baarda	Pope	Danish	Huber	FSME
I	1	99.69	99.74	90.87	92.24	100
	2	92.79	10.77	85.92	61.39	90.52
II	1	99.69	93.25	91.53	6.91	100
	2	92.70	2.94	90.05	2.10	90.41

It is apparent from Table 5 that Baarda and Danish are the two most successful methods against gross and influential outliers among classical and robust methods. In the case of small outliers with gross or influential outliers, the robustness of the models has been tested. Different types of outliers have been generated to evaluate the MSRs of the models for various scenarios as follows: I. Gross outlier (50σ), II. Influential outliers (1000σ), III. A small outlier and a gross outlier, IV. A small outlier and an influential outlier, V. Two small outliers and a gross outlier VI. Two small outliers and an influential outlier.

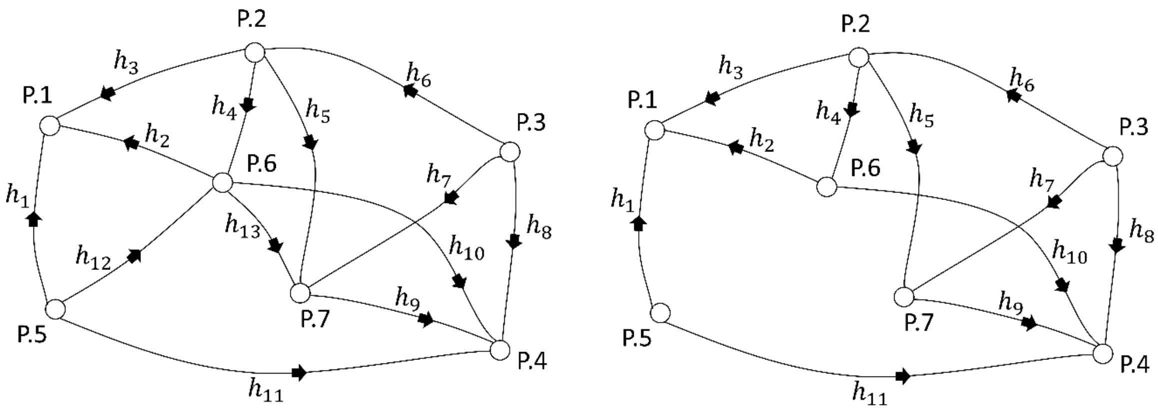
Table 6: MSRs for small outliers with a gross or an influential outliers

Scenario	Baarda	Pope	Danish	Huber	FSME
III	52.55	19.17	53.82	48.17	70.40
IV	52.55	19.17	56.7	4.5	70.40
V	20.17	0.58	24.85	22.87	63.34
VI	20.17	0.58	31.66	2.62	63.36

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions

Comparing Table 5 with Table 6, it was observed how MSRs of these two methods were affected in case of one or two small outliers. If a small outlier occur, the MSRs drop dramatically by about 40%. The MSRs drop to 20% with two small outliers in the network. Furthermore, this loss is around 35-40% for the proposed method. FSME, however, stands out as the approach with an MSR of 60-70% in scenarios involving small outliers.



2a) 2b)

Figure 3: Leveling networks 2a and 2b with low redundancy

When the redundancy of the observations decreases in the leveling network, difficulties arise in determining the outliers due to the swamping and masking effects. Two different leveling networks are considered to obtain the MSR of the methods in such cases. In the first of these, the MSR of the approaches has been compared by excluding an observation of the network. As seen in Table 7, the MSR decreased by 30% in all approaches when m=2 compared with the case m=1. Although the number of small outliers changes, the highest MSRs have been obtained by FSME for the leveling network 2a. The network is further weakened, so only two lines of the corner point P.5 remain in the leveling network 2b.

Table 7: MSR of models (small outliers) for leveling network 2a

m	Baarda	Pope	Danish	Huber	FSME
1	49.55	24.15	63.31	53.72	83.78
2	15.19	0.37	36.99	26.88	55.08
3	2.38	0.00	16.69	10.97	23.54

The results, as shown in Table 8, indicate that the FSME is the approach with the highest MSR for m=1. When m>1 compared with the case m=1 in Table 8, MSRs of the conventional and robust methods show more dramatic decrease than

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development
Discussions
Open Access
EGU

FSME approach. Comparing the estimated results for the network 2a and 2b reveals an approximate 15% drop in MSR values when m=1. MSRs decrease as the controllability of the observations in the network decreases.

Table 8. MSR of models (small outliers) for leveling network 2b

m	Baarda	Pope	Danish	Huber	FSME
1	36.59	10.13	52.93	37.82	68.32
2	5.43	0.04	23.69	14.31	24.57
3	0.13	0.00	8.07	4.90	3.89

6 Conclusion

The present study was designed to determine the usability of the presented method in geodetic leveling networks. The design of the FSME (Forward Search of the Model Error) approach was based on identifying the minimum variance from all possible combinations that assume observations as model errors in the Gauss-Markov model. This approach gives more reliable results by preventing the swamping and masking effect. The MSRs of the suggested method was obtained for various kinds of outliers in three different leveling network. The results of this investigation show that FSME is a more efficient approach than the robust and conventional methods. The proposed method enhanced the MSR by almost 40-45% compared to the Baarda and Danish (with the variance unknown case) method for multiple outliers (i.e., $1<m<4$). In cases where the leveling network-1 does not have specific observations at the corner point, the proposed method was 20-30% more successful than the others.

Authorship contribution statement: Conceptualization, Methodology, Software, Validation, Data curation, Formal analysis, Investigation, Writing – original draft, Visualization,

Competing interest: The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability: The source code and the data can be freely downloaded from https://github.com/Godesist/OutlierDetectionForGeodeticLevelingNetwork.git

References

Aydin, C.: Power of global test in deformation analysis. J. Surv. Eng., 138(2), 51–56, doi:10.1061/(ASCE)SU.1943-5428.0000064, 2012.

Baarda, W.: A testing procedure for use in geodetic networks. Publications on Geodesy, 2=5, Netherlands Geodetic Commission. Delft, Netherlands, ISBN 90 6132 209 X, 1968.

Batilović, M., Sušić, Z., Kanović, Ž., Marković, M. Z., Vasić, D., and Bulatović, V.: Increasing efficiency of the robust deformation analysis methods using genetic algorithm and generalised particle swarm optimisation. Survey Review, 53(378), 193-205, doi:10.1080/00396265.2019.1706294, 2021.

Duchnowski, R.: Sensitivity of robust estimators applied in strategy for testing stability of reference points. EIF approach. Geodesy and Cartography, 60(2), 123–134, doi: 10.2478/v10277-012-0011-z, 2011.

Durdag, U. M., Hekimoglu, S., and Erdogan, B.: What is the relation between smearing effect of least squares estimation and its influence function? Surv. Rev., 54(385), 320-331, doi: 10.1080/00396265.2021.1939590, 2022.

Erdogan, B.: An outlier detection method in geodetic networks based on the original observations. Bol. Ciênc. Geod., 20 (3), 578–589, doi: 10.1590/S1982-21702014000300033, 2014.

Erdogan, B., Hekimoglu, S., Durdag, U. M., and Ocalan, T.: Empirical estimation of the power of test in outlier detection problem. Studia Geophys. et Geod., 63, 55-70, doi:10.1007/s11200-018-1144-9, 2019.

Gao, Y., Krakiwsky, E.J., and Czompo, J.: Robust testing procedure for detection of multiple blunders. J. Surv. Eng., 118 (1), 11–23, doi:10.1061/(ASCE)0733-9453(1992)118:1(11), 1992.

Hekimoglu, Ş.: Reliabilities of χ2-and F Tests in Gauss-Markov Model. J. Surv. Eng., 125(3), 109-135, doi:10.1061/(ASCE)0733-9453(1999)125:3(109), 1999.

Hekimoglu, S., and Koch, K. R.: How can reliability of the test for outliers be measured? Allgemeine Vermessungsnachrichten, 107(7), 247–254, doi: , 2000.

Hekimoglu, S., and Erenoglu, R. C.: Effect of heteroscedasticity and heterogeneousness on outlier detection for geodetic networks. J. Geod., 81(2), 137–148, doi:10.1007/s00190-006-0095-z, 2007.

Hekimoglu, S., Erdogan, B., and Butterworth, S.: 2010. Increasing the efficacy of the conventional deformation analysis methods: alternative strategy. J. Surv. Eng., 136 (2), 53–62, doi:10.1061/(ASCE)SU.1943-5428.0000018, 2010.

Hekimoglu, S. and Erdogan, B.: Application of median-equation approach for outlier detection in geodetic networks. Bol. Ciênc. Geod., 19 (4), 548–557, doi:10.1590/S1982-21702013000400002, 2013.

Hekimoglu, S., Erdogan, B., Soycan, M., and Durdag, U. M.: Univariate approach for detecting outliers in geodetic networks. J. Surv. Eng., 140 (2), 04014006, doi:10.1061/(ASCE)SU.1943-5428.0000123, 2014.

Hekimoglu, S., Erdogan, B., and Erenoglu, R. C.: A new outlier detection method considering outliers as model errors. Exp Techniques, 39, 57-68, doi:10.1111/j.1747-1567.2012.00876.x, 2015.

Huber,P.J.: Robust Statistics. New York, USA: John Wiley and Sons, 312 pp., ISBN:9780471418054, 1981.

Koch, K. R.: Parameter estimation and hypothesis testing in linear models, New York, USA: Springer, 334 pp., ISBN: 978-3-540-65257-1, 1999.

Lehmann, R.: 3 σ-rule for outlier detection from the viewpoint of geodetic adjustment. J. Surv. Eng., 139(4), 157-165, doi:10.1061/(ASCE)SU.1943-5428.0000112, 2013.

https://doi.org/10.5194/gmd-2023-210
Preprint. Discussion started: 5 December 2023

Geoscientific Model Development Discussions

Maronna, R. A., Martin, R. D., Yohai, V. J., and Salibián-Barrera, M.: Robust statistics: theory and methods (with R). Chichester, UK: Wiley, 430 pp., ISBN:9781119214687, 2019.

Niemeier, W.: Anlage von Überwachungsnetzen. Geodaetische netze in landes-und ingenieurvermessung II., H. Pelzer, ed., Verlag Konrad Wittwer, Stuttgart, Germany, 527–558 (in German), 1985.

Nowel, K.: Specification of deformation congruence models using combinatorial iterative DIA testing procedure. J. Geod., 94 (12), 1–23, doi: 10.1007/s00190-020-01446-9, 2020.

Pope, A. J.: The statistics of residuals and the detection of outliers. NOAA Technical Rep. NOS 65 NGS 1, U.S. Dept. of Commerce. Rockville, Maryland, USA, 98 pp., 1976.

Teunissen, P. J.: Testing theory; an introduction. VSSD Leeghwaterstraat 42, 2628 CA Delft, The Netherlands, 147 pp., ISBN:978-9040719752, 2006.

Wang J., Zhao J., Liu Z. and Kang. Z.: Location and estimation of multiple outliers in weighted total least squares. Measurement, 181, 109591, doi:10.1016/j.measurement.2021.109591, 2021.

Zienkiewicz, M. H., and Dąbrowski, P. S.: Matrix Strengthening the Identification of Observations with Split Functional Models in the Squared Msplit (q) Estimation Process. Measurement, 112950, doi:10.1016/j.measurement.2023.112950, 2023.