Article

Which Are the Effects of Hard and Soft Equality Constraints on the Iterative Outlier Elimination Procedure?

Vinicius Francisco Rofatto [1,2] [ID], **Marcelo Tomio Matsuoka** [1,2,3,4], **Ivandro Klein** [5,6], **Maurício Roberto Veronez** [4] **and Luiz Gonzaga da Silveira Jr.** [4]

[1] Graduate Program in Remote Sensing, Federal University of Rio Grande do Sul, RS, Brazil; vfrofatto@gmail.com (V.F.R.);
[2] Institute of Geography, Federal University of Uberlandia, Monte Carmelo, 38500-000, MG, Brazil
[3] Graduate Program in Agriculture and Geospatial Information, Federal University of Uberlândia, 38500-000, Monte Carmelo, MG, Brazil; tomio@ufu.br (M.T.M.)
[4] Graduate Program in Applied Computing, Unisinos University, Av. Unisinos, 950, São Leopoldo, 93022-000, RS, Brazil; veronez@unisinos.br (M.R.V.); lgonzagajr@gmail.com (L.G.d.S.J.)
[5] Department of Civil Construction, Federal Institute of Santa Catarina, Florianopolis, 88020-300, SC, Brazil; ivandroklein@gmail.com (I.K.)
[6] Graduate Program in Geodetic Sciences, Federal University of Paraná, Curitiba, 81531-990, PR, Brazil
* Correspondence: vinicius.rofatto@ufu.br

Abstract: In this paper we evaluate the effects of hard and soft constraints on the Iterative Data Snooping (*IDS*), an iterative outlier elimination procedure. Here, the measurements of a levelling geodetic network were classified according to the local redundancy and maximum absolute correlation between the outlier test statistics, referred to as clusters. We highlight that the larger the relaxation of the constraints, the higher the sensitivity indicators MDB (Minimal Detectable Bias) and MIB (Minimal Identifiable Bias) for both the clustering of measurements and the clustering of constraints. There are circumstances that increase the *family-wise error rate* (FWE) of the test statistics, increase the performance of the *IDS*. Under a scenario of soft constraints, one should set out at least three soft constraints in order to identify an outlier in the constraints. In general, hard constraints should be used in the stage of pre-processing data for the purpose of identifying and removing possible outlying measurements. In that process, one should opt to set out the redundant hard constraints. After identifying and removing possible outliers, the soft constraints should be employed to propagate their uncertainties to the model parameters during the process of least-squares estimation.

Keywords: Constraints; Hypothesis Testing; Outlier Detection; Monte Carlo; Quality Control; Geodesy.

---

1. Introduction

It is very common to build models (i.e. the equation systems) based on some initial knowledge about a given problem. In other words, models are often set up in a way that the model parameters need to fulfil certain constraints. Such constraints are a prior knowledge embedded into a model to avoid a trivial solution; to guarantee the stability of estimates; to improve the precision and accuracy of the results by reducing the number of unknown parameters or accordingly by increasing the redundancy of the system; and to mitigate (or even estimate) a possible systematic effect [1–3]. For example, [4] adopted constraints to determine the transponder coordinates in a problem of combining

satellite positioning (GNSS) of a surface platform with acoustic ranging to seafloor transponders; [5–10] have used constraints to model the atmospheric effects on GNSS signals; and [11] have imposed the constraints of predicted satellite clocks to improve the precise orbit determination (POD) processing during maneuvers.

The models are usually formulated with minimal constraint or extra (redundant) constraints. In that case, we are referring to the so-called *equality constraints* which are usually incorporated into a system of equations in order to try to have a model well posed [12].

For the most part, minimal constraints are introduced to solve to the problem of rank deficiency in linear systems. The rank deficiency is often caused by the lack (or insufficient) information about a problem. In the field of geodesy, for example, minimal constraints are external information whose primary role is to specify the coordinate system to which the network station positions will be estimated by least-squares method (LS). This problem is known as *datum definition* (or also zero-order design or datum choice problem) [13–18]. Several works have been investigated in the sense of minimum-constrained adjustment and the datum choice problem in the geodetic literature, focusing on topics like free-adjustment and the role of inner constraints (see e.g. [19–22]).

If the number of constraints exceeds the minimum needed to solve the rank deficiency of the equation systems, we say that we have redundant (or extra) constraints. Extra constraints have also been used to check the stability of points in geodetic deformation analysis (see e.g [23–25]), to test the compatibility of constraints with the observations and the rest of the constraints [1,14,26,27]. So far we have only distinguished the constraints in terms of numerical quantity.

The model can also be subject to a *hard* and *soft* (or *weighted*) constraints. Hard constraints can often represent a case in which there exist known functional relations between the unknown parameters. Soft constraints (or *looser* constraints) are, however, for which such functional relations can slightly be violated depending on their uncertainty [3,27]. Soft constraints may also be referred to as a pseudo-observation model [28].

The well-known least-squares (LS) has been widely used as a standard method of model parameters estimation in geodetic applications and many others branches of modern science [29–42]. This is due to the flexibility of the LS, since no concepts from probability theory are used in formulating the least-squares minimisation problem.

LS has the property of being a linear unbiased estimator (LUE), and some special cases it coincides with the best linear unbiased estimator (BLUE). The estimator which has the smallest variance of all LUEs is called the best linear unbiased estimator (BLUE). If we have the full knowledge of the probability density function (PDF) of the measurements, the method of maximum likelihood estimation (MLE) can also be applied. In case of normally distributed measurements (Gauss-Markov model), the MLE estimators are identical to the BLUE ones, and therefore the LS and MLE principles provide identical results [32,43].

However, the presence of undesirable outliers in dataset makes LS no longer *unbiased* and it does not coincides with MLE. Here, we assume that an outlier is a measurement that is so probably caused by a blunder that it is better not used or not used as it is [44]. Failure to identify an outlier can jeopardise the reliability level of a system. Due to its importance, outliers must be appropriately treated to ensure the quality of data analysis [45].

In this paper, we employ the iterative data snooping (*IDS*), which is hypothesis test-based outlier. Important to mention that *IDS* is not restrict to the field of geodetic statistics, but it is a generally applicable method [46,47].

IDS is an iterative outlier elimination procedure, which combines estimation, testing and a corrective action [45,48]. Parameter estimation is often conducted in the sense of the LS. Then, hypothesis testing is performed with the aim to identify any outlier that may be present in dataset. After identification, the suspected outlier is then excluded from the dataset as a corrective action (i.e. adaptation), and the LS is restarted without the rejected measurement. If the model redundancy permits, this procedure is repeated until no more (possible) outliers can be identified (see e.g. [31], pp.

135). Although here we restrict ourselves to the case of one outlier at a time, *IDS* can also be applied for the case of multiples (simultaneous) outliers [49]. For more details about multiples (simultaneous) outlier refers to [50–52].

There are chances of correct and false decisions of IDS, because these procedure is based on statistical hypothesis testing. Recently, Rofatto et al. [45] have provided an algorithm based on Monte Carlo to determine the probability levels associated with *IDS*. In that case, they described six classes of decisions for *IDS*, namely probability of correct identification ($\mathcal{P}_{CI}$), probability of missed detection ($\mathcal{P}_{MD}$), probability of wrong exclusion ($\mathcal{P}_{WE}$), probability of over-identification positive ($\mathcal{P}_{over+}$), probability of over-identification negative ($\mathcal{P}_{over-}$) and probability of statistical overlap ($\mathcal{P}_{ol}$), as follows:

- $\mathcal{P}_{CI}$: Probability of identifying and removing correctly an outlying measurement;
- $\mathcal{P}_{MD}$: Probability of not detecting the outlier (i.e. Type II decision error for *IDS*);
- $\mathcal{P}_{WE}$: Probability of identifying and removing a non-outlying measurement while the '*true*' outlier remains in the dataset (i.e. Type III decision error [53] for *IDS*);
- $\mathcal{P}_{over+}$: Probability of identifying and removing correctly the outlying measurement and others;
- $\mathcal{P}_{over-}$: Probability of identifying and removing more than one non-outlying measurement, whereas the 'true outlier' remains in the dataset;
- $\mathcal{P}_{ol}$: occurs in cases where one alternative hypothesis has the same distribution as the another one. These hypotheses cannot be distinguished, because their test statistics are numerically the same, violating the *IDS* rule of one outlier at a time. In that case, they are nonseparable and an outlier cannot be identified. In other words, it corresponds to the probability of flagging simultaneously two (or more) measurements as outliers.

Based on the probability of correct detection ($\mathcal{P}_{CD}$) and probability of correct identification ($\mathcal{P}_{CI}$), the minimal biases, MDB (Minimal Detectable Bias) and MIB (Minimal Identifiable Bias), can be computed as sensitivity indicators for outlier detection and identification, respectively. "Outlier detection" only informs whether or not there might have been at least one outlier. However, the detection does not tell us which measurement is an outlier. The localization of the outlier is a problem of "outlier identification", i.e. "outlier identification" implies the execution of a search among the measurements for the most likely outlier [45]. Therefore, the smallest value of an outlier that can be detected given a certain $\mathcal{P}_{CD}$ defines the MDB. On the other hand, the smallest value of an outlier that can be identified given a certain $\mathcal{P}_{CI}$ defines the MIB.

In this paper, we investigate the effects of models subject to constraints (minimum, redundant, hard and soft) on the probability levels associated with *IDS*. It is important to emphasised that if a standard-deviation of a constraint (or a set of a constraint) is changed from zero to a nonzero value, it is called "relaxation" of the constraint [28]. Here, we also evaluate the effect of relaxing constraints on the MIB and MDB. This kind of assessment is a kind of sensitivity analysis. We also highlight that the task of clustering a set of geodetic measurements is applied for the first time in this study. We show that the clusters can be defined according to two deterministic parameters: local redundancy and correlation between the outlier test statistics. Moreover, critical values optimized by Monte Carlo method were used here [45,46] in order to compute the decision classes associated with IDS, i.e. $\mathcal{P}_{CI}$, $\mathcal{P}_{MD}$, $\mathcal{P}_{WE}$, $\mathcal{P}_{over+}$, $\mathcal{P}_{over-}$ and $\mathcal{P}_{ol}$.

2. Theoretical Framework

The IDS is an important case of multiple hypothesis testing. In this section, therefore, we briefly present the principal elements involved for the case of multiple testing.

First, the null hypothesis, denoted by $\mathcal{H}_0$, is formulated under the condition that random errors are normally distributed with expectation zero, i.e. in the absence of outliers. Thus, the null hypothesis $\mathcal{H}_0$ of the standard Gauss–Markov model in the linear or linearised form is given by [33]:

$$\mathcal{H}_0 : \mathbb{E}\{\boldsymbol{y}\} = \boldsymbol{Ax} + \mathbb{E}\{\boldsymbol{e}\} = \boldsymbol{Ax}; \ \mathbb{D}\{\boldsymbol{y}\} = \boldsymbol{Q}_e \qquad (1)$$

where $\mathbb{E}\{.\}$ is the expectation operator, $\mathbb{D}\{.\}$ is the dispersion operator, $y \in \mathbb{R}^{n \times 1}$ is the vector of measurements, $A \in \mathbb{R}^{n \times u}$ is the coefficient matrix, $x \in \mathbb{R}^{u \times 1}$ is the unknown parameter vector, $e \in \mathbb{R}^{n \times 1}$ is the unknown vector of measurement errors and $Q_e \in \mathbb{R}^{n \times n}$ is the positive-definite covariance matrix of the measurements y.

Under normal working conditions (i.e., $\mathcal{H}_0$), the measurement error model is then given by

$$e \sim N(0, Q_e), \tag{2}$$

First, we assume that the coefficient matrix A suffers from a rank deficiency, i.e. $u - rank(A) > 0$. In that case, minimal constraints are added to solve the problem of the rank-deficient system. An example of how to handle the problem of the rank-deficient is given in the section 3.1, where a minimal constraint in Equation 4 was added to the matrix A in 3. This means that the columns of the design matrix A in Equation 1 become linearly independent, i.e. the matrix A become full rank, such that $u - rank(A) = 0$.

The best linear unbiased estimator (BLUE) of e under $\mathcal{H}_0$ is the well-known estimated least-squares residual vector $\hat{e} \in \mathbb{R}^{n \times 1}$, which is given by

$$
\begin{aligned}
\hat{e} &= y - A\hat{x} \\
&= y - A(A^T W A)^{-1}(A^T W y) \\
&= Ax + e - A(A^T W A)^{-1}(A^T W(Ax + e)) \\
&= e - A(A^T W A)^{-1}(A^T W e) \\
&= (I - A(A^T W A)^{-1} A^T W)e \\
&= Re,
\end{aligned}
\tag{3}
$$

with $\hat{x} \in \mathbb{R}^{u \times 1}$ being the BLUE of x under $\mathcal{H}_0$; $W \in \mathbb{R}^{n \times n}$ is the known matrix of weights, taken as $W = \sigma_0^2 Q_e^{-1}$, where σ_0^2 is the variance factor, $I \in \mathbb{R}^{n \times n}$ is the identity matrix and $R \in \mathbb{R}^{n \times n}$ is known as the redundancy matrix. The R matrix is an orthogonal projector that projects onto the orthogonal complement of the range space of A. The main diagonal elements of matrix R are known as local redundancy numbers (denoted by r_i of the the system. The larger the number of local redundancy for a given measurement, the larger the degree of importance of that measurement for the model, i.e. the larger the absorption of a possible error of that measurement into their corresponding least-squares residual.

The degrees of freedom r (i.e. the overall redundancy) of the model under $\mathcal{H}_0$ (Equation (1)) is

$$r = rank(Q_{\hat{e}}) = n - rank(A) = n - u, where \tag{4}$$

$$Q_{\hat{e}} = Q_e - \sigma_0^2 A(A^T W A)^{-1} A^T \tag{5}$$

On the other hand, an alternative model is proposed when there are doubts about the reliability level of the model under $\mathcal{H}_0$. Here, we assume that the validity of the null hypothesis $\mathcal{H}_0$ in Equation (1) can be violated if the dataset is contaminated by outliers. The model in an alternative hypothesis, denoted by $\mathcal{H}_A$, is to oppose Equation (1) by an extended model that includes the unknown vector $\nabla \in \mathbb{R}^{q \times 1}$ of deterministic bias parameters as follows ([29,31]):

$$\mathcal{H}_A : y = Ax + C\nabla + e = \begin{pmatrix} A & C \end{pmatrix} \begin{pmatrix} x \\ \nabla \end{pmatrix} + e, \tag{6}$$

where $C \in \mathbb{R}^{n \times q}$ is the matrix that relates bias parameters, i.e., the values of the outliers to observations. We restrict ourselves to the matrix $\begin{pmatrix} A & C \end{pmatrix}$ having full column rank, such that

$$r = rank\begin{pmatrix} A & C \end{pmatrix} = u + q \leq n \tag{7}$$

One of the most used procedures based on hypothesis testing for outliers in linear (or linearised) models is the well-known data snooping method [29,54]. This procedure consists of screening each individual measurement for the presence of an outlier [49]. In that case, data snooping is based on a local model test, such that $q = 1$, and therefore, the n alternative hypothesis is expressed as

$$\mathcal{H}_A^{(i)} : \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{c_i}\nabla_i + \boldsymbol{e} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{c_i} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \nabla_i \end{pmatrix} + \boldsymbol{e}, \forall i = 1, \cdots, n \tag{8}$$

Now, matrix $\boldsymbol{C}$ in Equation (6) is reduced to a canonical unit vector $\boldsymbol{c_i}$, which consists exclusively of elements with values of 0 and 1, where 1 means that the ith bias parameter of magnitude ∇_i affects the ith measurement, and 0 means otherwise. In that case, the rank of $(\boldsymbol{A} \quad \boldsymbol{c_i}) \in \mathbb{R}^{n \times (u+1)}$ and the vector ∇ in Equation (6) reduces to a scalar ∇_i in Equation (42), i.e., $\boldsymbol{c_i} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 1^{i_{th}} & 0 & \cdots & 0 \end{pmatrix}^T$. When $q = n - u$, an overall model test is performed. For more details about the overall model test, see, for example, [55,56].

Note that the alternative hypothesis $\mathcal{H}_A^{(i)}$ in Equation (42) is formulated under the condition that the outlier acts as a systematic effect by shifting the random error distribution under $\mathcal{H}_0$ by its own value [44]. In other words, the presence of an outlier in a dataset can cause a shift of the expectation under $\mathcal{H}_0$ to a nonzero value. Therefore, hypothesis testing is often employed to check whether the possible shifting of the random error distribution under $\mathcal{H}_0$ by an outlier is, in fact, a systematic effect (bias) or merely a random effect. This hypothesis test-based approach is called the *mean-shift model* (see, e.g., [18,23,29,46,47,52,54,57–65].

In the context of the mean-shift model, the test statistic involved in data snooping is given by the normalised least-squares residual, denoted by $\boldsymbol{w_i}$. This test statistic, also known as Baarda's w-test, is given as follows:

$$w_i = \frac{c_i^T Q_e^{-1} \hat{e}}{\sqrt{c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_i}}, \forall i = 1, \cdots, n \tag{9}$$

The alternative hypothesis in Equation (42) is formulated in the sense that "*There is at least one outlier in the vector of measurements y_i*" [46]. In that case, we are interested in knowing which of the alternative hypotheses may lead to the rejection of the null hypothesis with a certain probability. This means testing $\mathcal{H}_0$ against $\mathcal{H}_A^{(1)}, \mathcal{H}_A^{(2)}, \mathcal{H}_A^{(3)}, \ldots, \mathcal{H}_A^{(n)}$. This is known as multiple hypothesis testing (see, e.g., [47,48,53,55,57,58,66–70]). In that case, the test statistic coming into effect is the maximum absolute Baarda's w-test value (denoted by max-w), which is computed as [47]

$$\text{max-}w = \max_{i \in \{1, \cdots, n\}} |\boldsymbol{w_i}| \tag{10}$$

The decision rule for this case is given by

$$\begin{aligned} &\textit{Accept } \mathcal{H}_0 \textit{ if } \text{max-}w \leq \hat{k} \\ &\textit{Otherwise,} \\ &\textit{Accept } \mathcal{H}_A^{(i)} \textit{ if } \text{max-}w > \hat{k} \end{aligned} \tag{11}$$

The decision rule in Equation 45 says that if none of the n w-tests get rejected, then we accept the null hypothesis $\mathcal{H}_0$. If the null hypothesis $\mathcal{H}_0$ is rejected in any of the n tests, then one can only assume that detection occurred. In other words, if the max-w is larger than some percentile of its probability distribution (i.e., some critical value $\hat{k}$), then there is evidence that there is an outlier in the dataset. Therefore, "outlier detection" only informs us whether the null hypothesis $\mathcal{H}_0$ is accepted or not [45].

However, the detection does not tell us which alternative hypothesis $\mathcal{H}_A^{(i)}$ would have led to the rejection of the null hypothesis $\mathcal{H}_0$. The localisation of the alternative hypothesis, which would have rejected the null hypothesis, is a problem of "outlier identification". Outlier identification implies the

execution of a search among the measurements for the most likely outlier. In other words, one seeks to find which of Baarda's w-test is the maximum absolute value max-w and if that max-w is greater than some critical value $\hat{k}$.

Therefore, the data snooping procedure of screening measurements for possible outliers is actually an important case of multiple hypothesis testing and not single hypothesis testing. Moreover, note that outlier identification only happens when outlier detection necessarily exists; i.e., "outlier identification" only occurs when the null hypothesis $\mathcal{H}_0$ is rejected. However, correct detection does not necessarily imply correct identification [47,55,68].

In a special case of having only one single alternative hypothesis, one should decide between the null hypothesis $\mathcal{H}_0$ and only one single alternative hypothesis $\mathcal{H}_A^{(i)}$ of Equation (42). In that case, the false decisions are restricted to Type I error and Type II error. The probability of a Type I Error α_0 is the probability of rejecting the null hypothesis $\mathcal{H}_0$ when it is true, whereas the probability of a Type II error β_0 is the probability of failing to reject the null hypothesis $\mathcal{H}_0$ when it is false (note: the index '0' represents the case in which a single hypothesis is tested). Instead of α_0 and β_0, there is the confidence level $CL = 1 - \alpha_0$ and the power of the test $\gamma_0 = 1 - \beta_0$, respectively. The first deals with the probability of accepting a true null hypothesis $\mathcal{H}_0$; the second addresses the probability of correctly accepting the alternative hypothesis $\mathcal{H}_A^{(i)}$. In that case, given a probability of a Type I decision error α_0, we find the critical value k_0 as follows:

$$k_0 = \Phi^{-1}\left(1 - \frac{\alpha_0}{2}\right) \tag{12}$$

where Φ^{-1} denotes the inverse of the cumulative distribution function (cdf) of the two-tailed standard normal distribution $N(0,1)$.

The normalised least-squares residual w_i follows a standard normal distribution with the expectation that $\mathbb{E}\{w_i\} = 0$ if $\mathcal{H}_0$ holds true (there is no outlier). On the other hand, if the system is contaminated with a single outlier at the ith location of the dataset (i.e., under $\mathcal{H}_A^{(i)}$), then the expectation of w_i is

$$\mathbb{E}\{w_i\} = \sqrt{\lambda_0} = \sqrt{c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_i \nabla_i^2} \tag{13}$$

where λ_0 is the non-centrality parameter for $q = 1$. Note, therefore, that there is an outlier that causes the expectation of w_i to become $\sqrt{\lambda_0}$. The square-root of the non-centrality parameter $\sqrt{\lambda_0}$ in Equation (47) represents the expected mean shift of a specific w-test. In such a case, the term $c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_i$ in Equation (47) is a scalar, and therefore, it can be rewritten as follows [71]:

$$|\nabla_i| = MDB_{0_{(i)}} = \sqrt{\frac{\lambda_0}{c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_i}}, \ \forall i = 1, \cdots, n \tag{14}$$

where $|\nabla_i|$ is the Minimal Detectable Bias ($MDB_{0_{(i)}}$) for the case in which there is only one single alternative hypothesis, which can be computed for each individual alternative hypothesis according to Equation (42).

For a single outlier, the variance of an estimated outlier, denoted by $\sigma_{\nabla_i}^2$, is

$$\sigma_{\nabla_i}^2 = \left(c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_i\right)^{-1}, \ \forall i = 1, \cdots, n \tag{15}$$

Thus, the MDB can also be written as

$$MDB_{0_{(i)}} = \sigma_{\nabla_i} \sqrt{\lambda_0}, \ \forall i = 1, \cdots, n \tag{16}$$

where $\sigma_{\nabla_i} = \sqrt{\sigma_{\nabla_i}^2}$ is the standard deviation of estimated outlier ∇_i.

The MDB in Equations (48) or (16) of an alternative hypothesis is the smallest-magnitude outlier that can lead to the rejection of the null hypothesis $\mathcal{H}_0$ for a given α_0 and β_0. Thus, for each model of the alternative hypothesis $\mathcal{H}_A^{(i)}$, the corresponding MDB can be computed [47,72,73]. The limitation of this MDB is that it was initially developed for the binary hypothesis testing case. In that case, the MDB is a sensitivity indicator of Baarda's w-test when only one single alternative hypothesis is taken into account. In this article, we are confined to multiple alternative hypotheses. Therefore, both the MDB and MIB are computed by considering the case of multiple hypothesis testing.

For a scenario coinciding with the null hypothesis $\mathcal{H}_0$ under multiple testing hypothesis, there is the probability of incorrectly identifying at least one alternative hypothesis. This type of wrong decision is known as the *family-wise error rate* (*FWE*). The *FWE* is defined as

$$FWE = \alpha' = \leq 1 - (1 - \alpha_0)^n \tag{17}$$

which is approximately

$$FWE = \alpha' \leq n \times \alpha_0 \tag{18}$$

where α_0 is the significance level for an individual test. The quantity in Equation (18) is just equal to the upper bound of the Bonferroni inequality, i.e., $\alpha' \leq n\alpha$ [74]. For example, if the *FWE* level (α') is 0.05 and one is running 5 tests, then each test will have an α_0 of 0.05/5 = 0.01. In other words, one uses a global Type I Error rate α' that combines all tests under consideration instead of an individual error rate α_0 that only considers one test at a time time [70]. In that case, the critical value k_{bonf} is computed as

$$k_{bonf} = \Phi^{-1}\left(1 - \frac{\alpha'}{2n}\right) \tag{19}$$

The Bonferroni in Equation (18) is a good approximation for the case in which alternative hypotheses are independent. In practice, however, the test results always depend on each other to some degree because we always have a correlation between w-tests. The correlation coefficient between any Baarda's w-test statistic (denoted by ρ_{w_i,w_j}), such as w_i and w_j, is given by [57]

$$\rho_{w_i,w_j} = \frac{c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_j}{\sqrt{c_i^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_i}\sqrt{c_j^T Q_e^{-1} Q_{\hat{e}} Q_e^{-1} c_j}}, \forall (i \neq j) \tag{20}$$

The correlation coefficient ρ_{w_i,w_j} can assume values within the range $[-1,1]$.

Here, the extreme normalised residuals max-w (i.e., maximum absolute) in Equation (44) are treated directly as a test statistic. Note that when using Equation (44) as a test statistic, the decision rule is based on a one-sided test of the form max-$w \leq \hat{k}$. However, the distributions of max-w cannot be derived from well-known test distributions (e.g., normal distribution). The procedure to compute the critical value of max-w is given step-by-step by Rofatto et al. [45].

The other side of the multiple testing problem is the situation in which there is an outlier in the dataset. In that case, apart from Type I and Type II errors, there is a third type of wrong decision associated with Baarda's w-test. Baarda's w-test can also flag a non-outlying observation while the '*true*' outlier remains in the dataset. We are referring to the Type III error [53], also referred to as the probability of wrong identification ($\mathcal{P}_{WI}$). The description of the Type III error involves a separability analysis between alternative hypotheses [57,66,68,69]. Therefore, we are now interested in the identification of the correct alternative hypothesis. In that case, the non-centrality parameter in Equation (47) is not only related to the sizes of Type I and Type II decision errors but also dependent on the correlation coefficient ρ_{w_i,w_j} given by Equation (20).

On the basis of the assumption that one outlier is in the ith position of the dataset (i.e., $\mathcal{H}_A^{(i)}$ is *'true'*), the probability of a Type II error (also referenced as the probability of "*missed detection*", denoted by $\mathcal{P}_{MD}$) for multiple testing is

$$\mathcal{P}_{MD} = \mathcal{P}\left(\bigcap_{i=1}^{n} |\boldsymbol{w_i}| \le \hat{k} \,\Big|\, \mathcal{H}_A^{(i)} : true\right), \tag{21}$$

and the size of a Type III wrong decision (also called "*misidentification*", denoted by $\mathcal{P}_{WI}$) is given by

$$\mathcal{P}_{WI} = \sum_{i=1}^{n} \mathcal{P}\left(|\boldsymbol{w_j}| > |\boldsymbol{w_i}| \,\forall i, \, |\boldsymbol{w_j}| > \hat{k}(i \ne j) \,\Big|\, \mathcal{H}_A^{(i)} : true\right) \tag{22}$$

On the other hand, the probability of correct identification (denoted by $\mathcal{P}_{CI}$) is

$$\mathcal{P}_{CI} = \mathcal{P}\left(|\boldsymbol{w_i}| > |\boldsymbol{w_j}| \,\forall j, \, |\boldsymbol{w_i}| > \hat{k}(i \ne j) \,\Big|\, \mathcal{H}_A^{(i)} : true\right) \tag{23}$$

with

$$1 - \mathcal{P}_{CI} = 1 - \mathcal{P}_{CD} + \mathcal{P}_{WI} = \mathcal{P}_{MD} + \mathcal{P}_{WI} \tag{24}$$

Note that the three probabilities of missed detection $\mathcal{P}_{MD}$, wrong identification $\mathcal{P}_{WI}$ and correct identification $\mathcal{P}_{CI}$ sum up to unity: i.e., $\mathcal{P}_{MD} + \mathcal{P}_{WI} + \mathcal{P}_{CI} = 1$.

The probability of correct detection $\mathcal{P}_{CD}$ is the sum of the probability of correct identification $\mathcal{P}_{CI}$ (selecting a correct alternative hypothesis) and the probability of misidentification $\mathcal{P}_{WI}$ (selecting one of the n-1 other hypotheses), i.e.,

$$\mathcal{P}_{CD} = \mathcal{P}_{CI} + \mathcal{P}_{WI} \tag{25}$$

The probability of wrong identification $\mathcal{P}_{WI}$ is identically zero, $\mathcal{P}_{WI} = 0$, when the correlation coefficient is exactly zero, $\rho_{w_i,w_j} = 0$. In that case, we have

$$\mathcal{P}_{CD} = \mathcal{P}_{CI} = 1 - \mathcal{P}_{MD} \tag{26}$$

The relationship given in Equation (26) would only happen if one neglected the nature of the dependence between alternative hypotheses. In other words, this relationship is valid for the special case of testing the null hypothesis $\mathcal{H}_0$ against only one single alternative hypothesis $\mathcal{H}_A^{(i)}$.

Since the critical region in multiple hypothesis testing is larger than that in single hypothesis testing, the Type II decision error (i.e., $\mathcal{P}_{MD}$) for the multiple test becomes smaller [47]. This means that the correct detection in binary hypothesis testing (γ_0) is smaller than the correct detection $\mathcal{P}_{CD}$ under multiple hypothesis testing, i.e.,

$$\mathcal{P}_{CD} > \gamma_0 \tag{27}$$

Detection is easier in the case of multiple hypothesis testing than single hypothesis testing. However, the probability of correct detection $\mathcal{P}_{CD}$ under multiple testing is spread out over all alternative hypotheses, and therefore, identifying is harder than detecting. From Equation (25), it is also noted that detection does not depend on identification. However, outlier identification depends on correct outlier detection. Therefore, we have the following inequality:

$$\mathcal{P}_{CI} \le \mathcal{P}_{CD} \tag{28}$$

Note that the probability of correct identification $\mathcal{P}_{CI}$ depends on the probability of missed detection $\mathcal{P}_{MD}$ and wrong identification $\mathcal{P}_{WI}$ for the case in which data snooping is run only once, i.e., a single round of estimation and testing. However, in this paper, we deal with data snooping in its iterative form (i.e., *IDS*), and therefore, the probability of correct identification $\mathcal{P}_{CI}$ depends on other decision rules.

In contrast to the data snooping single run, the success rate of correct detection $\mathcal{P}_{CD}$ for *IDS* depends on the sum of the probabilities of correct identification $\mathcal{P}_{CI}$, wrong exclusion ($\mathcal{P}_{WE}$), over-identification cases ($\mathcal{P}_{over+}$ and $\mathcal{P}_{over-}$), and statistical overlap ($\mathcal{P}_{ol}$), i.e.,

$$\mathcal{P}_{CD} = 1 - \mathcal{P}_{MD} = \mathcal{P}_{CI} + \mathcal{P}_{WE} + \mathcal{P}_{over+} + \mathcal{P}_{over-} + \mathcal{P}_{ol} \tag{29}$$

It is important to mention that the probability of correct detection is the complement of the probability of missed detection. Note from Equation (29) that the probability of correct detection $\mathcal{P}_{CD}$ is available even for cases in which the identification rate is null, $\mathcal{P}_{CI} = 0$. However, the probability of correct identification ($\mathcal{P}_{CI}$) necessarily requires that the probability of correct detection $\mathcal{P}_{CD}$ be greater than zero. For the same reasons given for the data snooping single run in the previous section, detecting is easier than identifying. In that case, we have the following relationship for the success rate of correct outlier identification $\mathcal{P}_{CI}$:

$$\mathcal{P}_{CI} = \mathcal{P}_{CD} - (\mathcal{P}_{WE} + \mathcal{P}_{over+} + \mathcal{P}_{over-} + \mathcal{P}_{ol}), \tag{30}$$

such as

$$\exists(\mathcal{P}_{CI}) \in [0,1] \iff (\mathcal{P}_{CD}) > 0 \tag{31}$$

It is important to mention that the wrong exclusion $\mathcal{P}_{WE}$ describes the probability of identifying and removing a non-outlying measurement while the '*true*' outlier remains in the dataset. In other words, $\mathcal{P}_{WE}$ is the Type III decision error for *IDS*). The overall wrong exclusion $\mathcal{P}_{WE}$ is the result of the sum of each individual contribution to $\mathcal{P}_{WE}$, i.e.,

$$\mathcal{P}_{WE} = \sum_{i=1}^{n-1} \mathcal{P}_{WE(i)} \tag{32}$$

On the basis of the probability levels of correct detection $\mathcal{P}_{CD}$ and correct identification $\mathcal{P}_{CI}$, the sensitivity indicators of minimal biases—Minimal Detectable Bias (MDB) and Minimal Identifiable Bias (MIB)—for a given α' can be computed as follows:

$$MDB = \arg \min_{\nabla_i} \mathcal{P}_{CD}(\nabla_i) > \tilde{\mathcal{P}}_{CD}, \forall i = 1, \cdots, n \tag{33}$$

$$MIB = \arg \min_{\nabla_i} \mathcal{P}_{CI}(\nabla_i) > \tilde{\mathcal{P}}_{CI}, \forall i = 1, \cdots, n \tag{34}$$

Equation (33) gives the smallest outlier ∇_i that leads to its detection for a user-defined correct detection rate $\tilde{\mathcal{P}}_{CD}$, whereas (34) provides the smallest outlier ∇_i that leads to its identification for a user-defined correct identification rate $\tilde{\mathcal{P}}_{CI}$.

As a consequence of the inequality in (28), the MIB will be larger than MDB, i.e., $MIB \geq MDB$. For the special case of having only one single alternative hypothesis, there is no difference between the MDB and MIB [55]. The computation of MDB_0 is easily performed by Equations (48) or (16), whereas the computation of the MDB in Equation (33) and the MIB in Equation (34) must be computed using Monte Carlo because the acceptance region (as well as the critical region) for the case of multiple alternative hypotheses is analytically intractable.

3. Case study of Levelling Geodetic Network

The analyse of the constraints effects on the probability levels of *IDS* are performed by an example of a levelling geodetic network.

A levelling geodetic network is a set of points located on the Earth's surface or near it and interconnected by height difference measurements. Some of these points are associated with a vertical height reference system (i.e. some points have known height), whereas others are parameters (unknown heights). The term "known" here means that those points with known heights are constraints necessary

to ensure that parameters (unknown heights) are sufficiently estimable. In the sense of modern geodetic reference systems, realisation and unification of a vertical height system consists of a combination of GNSS (Global Navigation Satellite System) and spirit levelling with geoid models. We will not go into the details about height geodetic reference system and frame. (Readers who wish to have further details on that issue please see e.g. [75]).

The mathematical model associated with a geodetic levelling network is linear (i.e. the levelling measurements bear a known linear relationship with the unknown heights). Geodetic networks usually has more measurements than parameters (i.e. $n > u$), i.e. we have a redundant measurement system. However, due to intrinsic random errors in a system, redundant measurements often lead to an inconsistent system of equations. To make the system consistent we have to introduce the information about the random measurement errors. The stochastical properties of the measurement errors are directly associated with the assumption of the probability distribution of these errors. In geodesy and many other scientific branches the well-known normal distribution is one of the most used as measurement error model [76]. Because of this, the model ceases to be purely mathematical and becomes a statistical model with functional and stochastic part.

In the absence of outliers, i.e. under null hypothesis $\mathcal{H}_0$ in Equation (1), the model for levelling geodetic network can be written as follows:

$$\Delta h_{i-j} + e_{\Delta h_{i-j}} = h_j - h_i, \tag{35}$$

where Δh_{i-j} is the height difference measured from point i to j and $e_{\Delta h_{i-j}}$ is the random error associated with the levelling measurement. For a only one single levelling line, one of these points is the constraint (known height), from which the height of another point is determined. The point with known height is also referred to as control point or benchmark. In a geodetic network, on the other hand, we have a set of levelling lines that connect both from points of unknown height and from points of known height (constraints). Under normal working conditions (i.e. $\mathcal{H}_0$), the measurement errors model is then given by Equation (2).

In this case study, we demonstrate the application of the proposed algorithm by [45] subject to a different constraints scenarios for that levelling geodetic network.

Recent work has been focused on non-stochastic constraints (hard constraints). Here, we also consider cases where constraints are subject to a random errors (i.e. soft constraints). That soft constraints are non-deterministic, and therefore they are measured model elements that can be designated as a priori information. The approach of this example can be applied in the design stage of geodetic network analysis [62,65]. Furthermore, the experiments performed here can also be extended to geodetic deformation analysis when the deformation effects are unmodelled in the sense of deterministic form [28].

3.1. Problem description

To analyse the constraints effects on the *IDS*, an example is taken from a geodetic levelling network with 12 height differences between the points. The equipment used to measure the level difference can be an electronic digital level. In that case, the levelling measurement system comprises of a special bar-coded staff (also called barcode rod) and a digital level (instrument). A digital level is basically a telescope that enables a horizontal line of sight. Digital levels consist of additional electronic image processing components to automatically read and analyse digital (bar coded) levelling staffs, where the graduation is replaced by a manufacturer dependent code pattern. Generally, the result is automatically stored in the data collector of the digital level. An example of a "*digital level − bar-code staff*" system is displayed in the Figure 1. For more details about digital level see e.g. [77–80].

Figure 1. Example of a *digital level − bar-code staff* system [45].

The standard-deviation of the uncorrelated measurements were the same and taken equal to $\sigma = 1mm$. The points are indicated as A to G. The eight network configuration are displayed in Figure 2(a, b, c, d, and e) and their details are given as follows:

1. Figure 2(a): Network with 1 hard constraint (i.e. network minimally constrained). Since the dimension of the network is 1D, the minimum information necessary to estimate the unknown heights is one. The height of G was fixed as control point (hard constraint), and 6 unknown heights (A,B,C,D,E,F) were minimally constrained. Therefore, the redundancy of the system (or overall degrees of freedom) was $r = n - rank(A) = n - u = 12 - 6 = 6$.

2. Figure 2(b): Network with one extra hard constraint (i.e. two hard constraints). The heights A and D were taken as hard constraints (i.e. heights A and D were fixed). The redundancy of the system in that case was $r = 12 - 5 = 7$ with 5 unknown heights (B,C,E,F,G) over-constrained.

3. Figure 2(c): Network with two extra hard constraints (i.e. three hard constraints).The heights A, D and G were taken as hard constraints. In that case, the redundancy of the system was $r = 12 - 4 = 8$.

4. Figure 2(d): Network with two soft constraints (A and D). In that case, a standard-deviation larger than zero was assigned to both constraints (i.e. $\sigma_c > 0$. In other words, A and D were processed as being both observations and unknown parameters, i.e. A and D were pseudo-observations. Here, the both constraints were simultaneously *relaxed* by considering their uncertainties 10 times worse than measurements (i.e. $\sigma_c = 10 \times \sigma = 10mm$); 10 times better than measurements (i.e. $\sigma_c = 0.1mm$); and their uncertainties equal to the measurements ($\sigma_c = 1mm$). In that case, the redundancy of the system was $r = 14 - 7 = 7$.

5. Figure 2(e): Network processed with A, D and G as pseudo-observations. Those three constraints were simultaneously *relaxed* by considering their standard-deviations equal to $\sigma_c = 10mm$ (10 times worse than measurements); $\sigma_c = 0.1mm$ (10 times better than measurements); and $\sigma_c = 1mm$ (the same as the measurements). In that case, the redundancy of the system was $r = 15 - 7 = 8$.
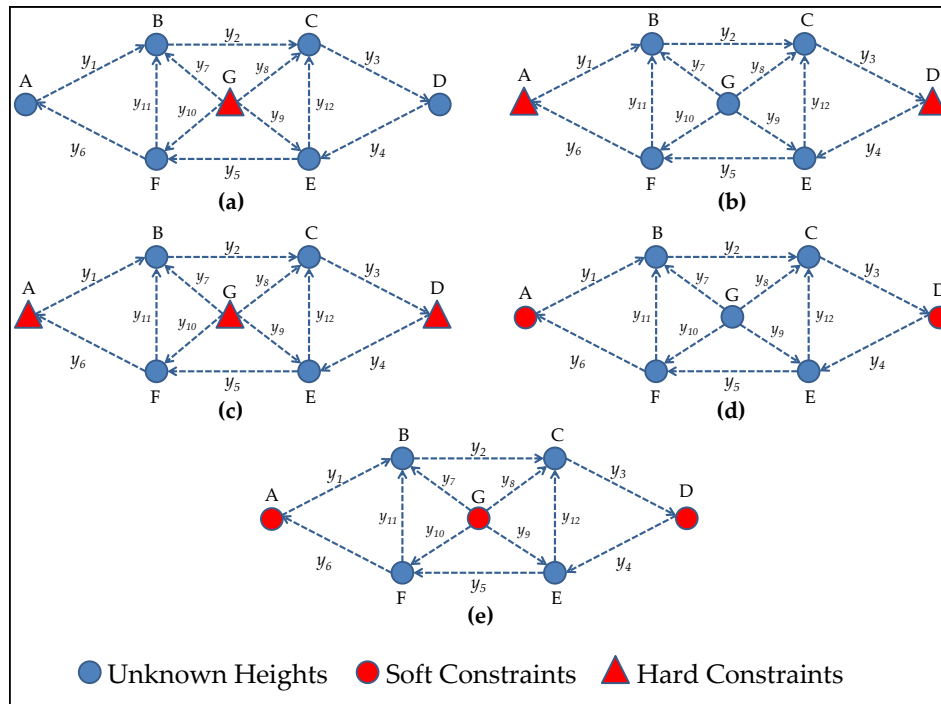
Figure 2. Levelling Geodetic Network subject to different constraint scenarios.

The following system of equations for that problem is given by:

$$y_1 + e_1 = h_B - h_A$$
$$y_2 + e_2 = h_C - h_B$$
$$y_3 + e_3 = h_D - h_C$$

$$\vdots$$

$$y_7 + e_7 = h_B - h_G$$
$$y_8 + e_8 = h_C - h_G$$

$$\vdots$$

$$y_{11} + e_{11} = h_B - h_F$$
$$y_{12} + e_{12} = h_C - h_E$$

(36)

The functional model (A) for the system of equations in 36 is given by:

$$
A = \begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 & 0
\end{bmatrix}
$$

(37)

Note that the rank defect of the matrix A is $u - rank(A) = 7 - 6 = 1$. In that case, there is needed to add at least one constraint in order to avoid rank deficiency of the matrix A. This is guaranteed when we take one height as known. For example, from the network in Figure 2(a), we have added the height G as known (i.e. as a hard constraint). In that case, the constraint equation should be added into the system in 36, i.e.:

$$y_{13} = h_G \ with \ \sigma_{y_{13}} = 0, \tag{38}$$

noticing that because the standard deviation is zero, the observation is non-stochastic (hard constraint) and the residual $e_{y_{13}} = 0$. This can generate problems in the inversion of the covariance matrix of the observations Q_e for the calculation of the weight matrix W, because the weight for that constraint would be $\frac{1}{0} = \infty$. In order to avoid that problem, we have eliminated the rank deficiency of matrix A by removing the seventh column of matrix A in 37 associated with the height G. Now, we have $u - rank(A) = 6 - 6 = 0$.

The constraint defines the geodetic datum, i.e. the S-system ([81],p.41). Another approach to solving the system of equations in 2 could be based on generalised (pseudo) inverses (see e.g. [82]).

The location of the constraints can be chosen in some circumstances, for example during the design stage of a geodetic network. For the special case of having a minimally constrained system, the location of the constraint will not influence the w-test statistics and the sensitivity indicators (MIB and MDB) [18]. However, more constraints than the minimum necessary to have a solution (i.e. extra constraints or redundant constraints) can change the least-squares residuals and hence w-test statistics and the minimal biases.

From the network with one extra constraint (2 constraints) in Figure 2(b), for example, the both first (height A) and fourth column (height D) of matrix A in 37 were eliminated in the case of having the two heights as hard constraints. For the case where these two heights (A and D) were taken as soft constraints, however, two observation equations were added to Equation 36, i.e.:

$$\begin{aligned} y_{13} + e_{13} &= h_A, \ \sigma_{y_{13}} > 0 \\ y_{14} + e_{14} &= h_D, \ \sigma_{y_{14}} > 0 \end{aligned} \tag{39}$$

In the case of soft constraints in Equation 39, two lines were added in matrix A. In other words, A and D were taken as pseudo-observations. In that case, the rank deficiency was also null (i.e. $u - rank(A) = 7 - 7 = 0$), the redundancy of the system was $r = n - rank(A) = n - u = 7$ and the matrix A was given as follows:

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{40}$$

For this example of two soft constraints, and by considering the both soft constraints with standard-deviation $\sigma_c = 10mm$, the symmetric and positive semi-definite covariance matrix of the observations (Q_e) was given as follows:

$$Q_e = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 100 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 100 \end{bmatrix} \tag{41}$$

The last two rows and columns of the matrix Q_e in Equation 41 refer to the variances ($\sigma_c^2 = (10mm)^2 = 100mm^2$) of the heights constraints A and D, respectively.

Similarly, matrices A and Q_e were constructed for the other cases studied here.

Although other measurements are able to identify an outlier for the case of having only one single soft constraint, the pseudo-observation (constraint) is not. In that case, the defect configuration is associated with the additional parameter in the constraint (i.e. the presence of an outlier in the constraint). In other words, an additional parameter on the soft constraint will not estimable. For example, if the height point G was taken as a soft constraint, the presence of an outlier in pseudo-observation G would lead to rank deficiency of matrix A, i.e. $u - rank(A) = 8 - 7 = 1$. Therefore, the case of having only one single soft constraint was not considered here.

3.2. Result of the Hard Constraint Effects on the Iterative Outlier Elimination Procedure

In this section, we present the results of the effects of the hard constraint on the *IDS*. The scenarios in Figure 2(a) (network minimally constrained), Figure 2(b) (two hard constraints) and Figure 2(c) (three hard constraints) were considered for the analysis.

The correlation coefficient between w-test statistics (ρ_{w_i,w_j}) were computed for the three scenarios, according to Equation(20). Table 1 provides that correlation for that network minimally constrained, Table 2 for that network over-constrained with two hard constraints and Table 3 for that network over-constrained with three hard constraints.

Table 1. Correlation matrix of w-test statistics for the levelling network minimally constrained in Figure 2(a).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$
w_{y_1}	1.0	0.2	0.1	0.1	0.2	1.0	-0.2	0.0	0.0	0.2	-0.4	-0.1
w_{y_2}		1.0	0.2	0.2	0.3	0.2	0.5	-0.5	-0.2	0.2	0.3	-0.3
w_{y_3}			1.0	1.0	0.2	0.1	0.0	0.2	-0.2	0.0	0.1	0.4
w_{y_4}				1.0	0.2	0.1	0.0	0.2	-0.2	0.0	0.1	0.4
w_{y_5}					1.0	0.2	-0.2	0.2	0.5	-0.5	0.3	-0.3
w_{y_6}						1.0	-0.2	0.0	0.0	0.2	-0.4	-0.1
w_{y_7}							1.0	-0.3	-0.3	-0.4	-0.4	-0.1
w_{y_8}								1.0	-0.4	-0.3	-0.1	-0.4
w_{y_9}									1.0	-0.3	0.1	0.4
$w_{y_{10}}$										1.0	0.4	0.1
$w_{y_{11}}$											1.0	-0.1
$w_{y_{12}}$												1.00

Table 2. Correlation matrix of w-test statistics for the levelling network with two hard constraints in Figure 2(b).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$
w_{y_1}	1.0	0.4	0.4	-0.3	-0.1	0.4	-0.3	0.1	0.1	0.1	-0.4	-0.1
w_{y_2}		1.0	0.4	-0.1	0.1	-0.1	0.4	-0.4	-0.1	0.1	0.3	-0.3
w_{y_3}			1.0	0.4	-0.1	-0.3	-0.1	0.3	-0.1	-0.1	0.1	0.4
w_{y_4}				1.0	0.4	0.4	0.1	0.1	-0.3	0.1	0.1	0.4
w_{y_5}					1.0	0.4	-0.1	0.1	0.4	-0.4	0.3	-0.3
w_{y_6}						1.0	-0.1	-0.1	-0.1	0.3	-0.4	-0.1
w_{y_7}							1.0	-0.4	-0.3	-0.4	-0.4	-0.1
w_{y_8}								1.0	-0.4	-0.3	-0.1	-0.4
w_{y_9}									1.0	-0.4	0.1	0.4
$w_{y_{10}}$										1.0	0.4	0.1
$w_{y_{11}}$											1.0	-0.1
$w_{y_{12}}$												1.0

Table 3. Correlation matrix of w-test statistics for the levelling network with three hard constraints in Figure 2(c).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$
w_{y_1}	1.0	0.3	0.1	-0.1	-0.1	0.1	-0.4	-0.1	-0.1	-0.1	-0.3	-0.1
w_{y_2}		1.0	0.3	-0.1	0.1	-0.1	0.3	-0.3	-0.1	0.1	0.3	-0.3
w_{y_3}			1.0	0.1	-0.1	-0.1	0.1	0.4	0.1	0.1	0.1	0.3
w_{y_4}				1.0	0.3	0.1	-0.1	-0.1	-0.4	-0.1	0.1	0.3
w_{y_5}					1.0	0.3	-0.1	0.1	0.3	-0.3	0.3	-0.3
w_{y_6}						1.0	0.1	0.1	0.1	0.4	-0.3	-0.1
w_{y_7}							1.0	-0.1	-0.1	-0.1	-0.3	-0.1
w_{y_8}								1.0	-0.1	-0.1	-0.1	-0.3
w_{y_9}									1.0	-0.1	0.1	0.3
$w_{y_{10}}$										1.0	0.3	0.1
$w_{y_{11}}$											1.0	-0.1
$w_{y_{12}}$												1.0

Table 4 gives the local redundancy (r_i), the standard-deviation of the LS-estimated outlier σ_{∇_i} and the maximum absolute correlation ($max_{\rho_{w_i,w_j}}$) for each scenario of hard constraint set out of in this study, i.e. Figure 2(a,b and c).

Table 4. Local redundancy (r_i), standard-deviation of the LS-estimated outlier σ_{∇_i} and the maximum absolute correlation ($max_{\rho_{w_i,w_j}}$) for each scenario of hard constraint.

Measurement	1 hard constraint			2 hard constraints			3 hard constraints		
	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$
y_1	0.396	1.589	1.00	0.583	1.309	0.36	0.708	1.188	0.41
y_2	0.500	1.414	0.47	0.583	1.309	0.36	0.583	1.309	0.32
y_3	0.396	1.589	1.00	0.583	1.309	0.36	0.708	1.188	0.41
y_4	0.396	1.589	1.00	0.583	1.309	0.36	0.708	1.188	0.41
y_5	0.500	1.414	0.47	0.583	1.309	0.36	0.583	1.309	0.32
y_6	0.396	1.589	1.00	0.583	1.309	0.36	0.708	1.188	0.41
y_7	0.563	1.333	0.47	0.583	1.309	0.36	0.708	1.188	0.41
y_8	0.563	1.333	0.47	0.583	1.309	0.36	0.708	1.188	0.41
y_9	0.563	1.333	0.47	0.583	1.309	0.36	0.708	1.188	0.41
y_{10}	0.563	1.333	0.47	0.583	1.309	0.36	0.708	1.188	0.41
y_{11}	0.583	1.309	0.43	0.583	1.309	0.36	0.583	1.309	0.32
y_{12}	0.583	1.309	0.43	0.583	1.309	0.36	0.583	1.309	0.32

Here, the twelve levelling measurements were clustered into four clusters. The clustering was defined according to the local redundancy (r_i) and the maximum absolute correlation ($max_{\rho_{w_i,w_j}}$) in Table 4. This is the first time that a clustering technique based on the similarity of local redundancy and the maximum absolute correlation between w-test statistics is applied to a problem of geodetic networks. Similarly, this has been done in [45]. The four cluster were defined as follows:

- Cluster 1: y_1, y_3, y_4 and y_6.
- Cluster 2: y_2 and y_5.
- Cluster 3: y_7, y_8, y_9 and y_{10}.
- Cluster 4: y_{11} and y_{12}.

The probability levels associated with *IDS* were averaged for each of these clusters. Here, we compute the probability levels of the *IDS* based on the procedure in Section (5) [45]. The critical values were $\hat{k} = 3.89$, $\hat{k} = 3.93$ and $\hat{k} = 3.93$ for 1 hard constraint, 2 hard constraints and 3 hard constraints, respectively. These critical values were found for $\alpha' = 0.001$ according to the procedure described in Appendix A. The probability of correct identification ($\mathcal{P}_{CI}$) and correct detection ($\mathcal{P}_{CD} = 1 - \mathcal{P}_{MD}$) are displayed in Figure 3 for each number of hard constraint (denoted by **h.c.**).
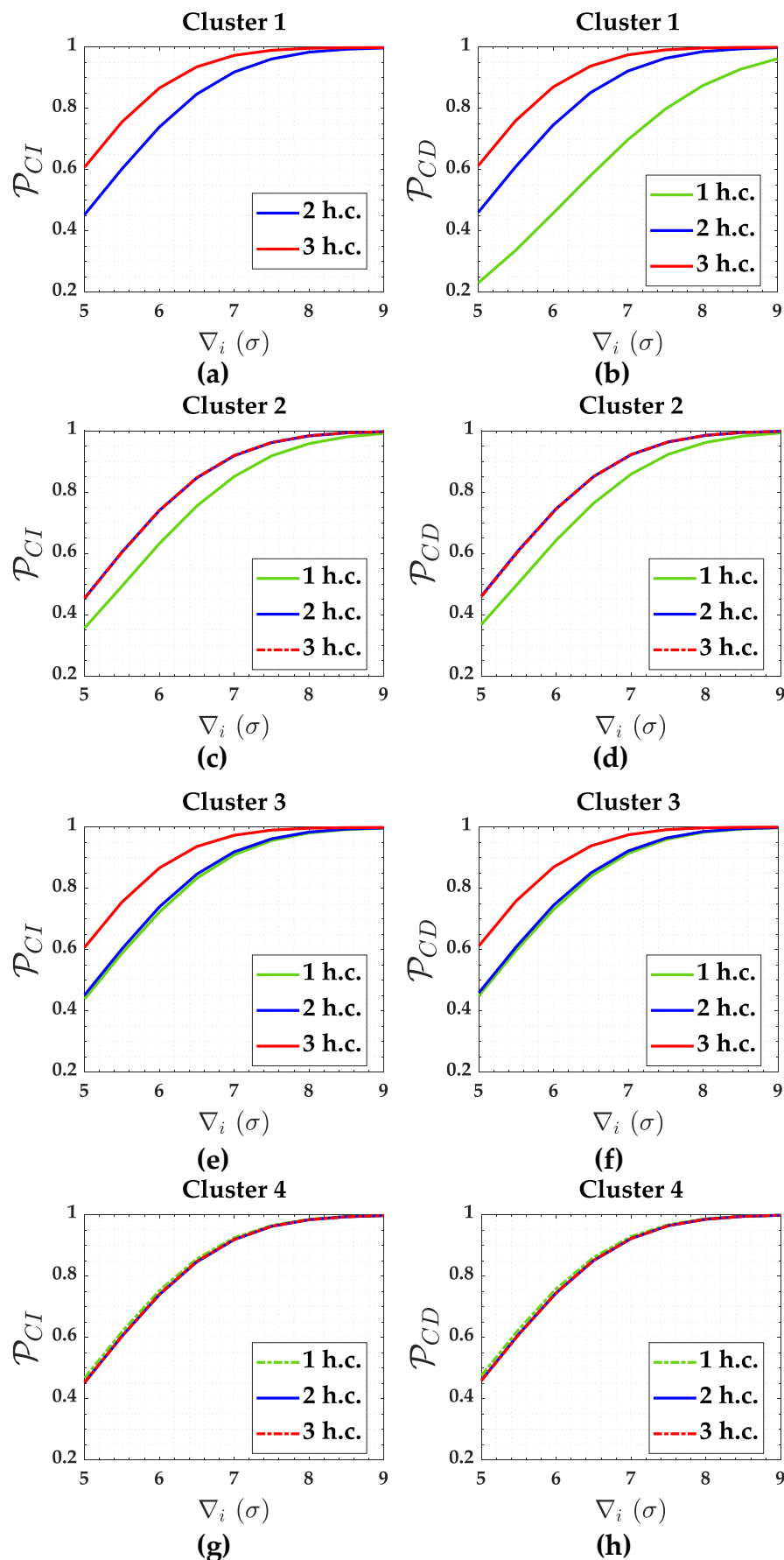
Figure 3. Probability of correct identification ($\mathcal{P}_{CI}$) and Probability of correct detection ($\mathcal{P}_{CD}$) for the case of hard constraints and for $\alpha' = 0.001$: Cluster 1(**a**,**b**), Cluster 2(**c**,**d**), Cluster 3(**e**,**f**) and Cluster 4(**g**,**h**).

The outlier magnitude were defined from $|5\sigma|$ to $|9\sigma|$. The outlier of $|5\sigma|$ was chosen because it is approximately the lowest $MDB_{0(i)}$ of the network when a single hypothesis testing is in play, as can be seen in Equation (48) or Equation (16). That $MDB_{0(i)}$ of $|5\sigma|$ was computed for significance level of $\alpha' = 0.001$ and a power of the test $\gamma_0 = 0.8$. This strategy reduces the search space for a MIB (Minimal Identifiable Bias), because we will always have the following inequality $MIB \geq MDB_{0(i)}$ [47,55]. Remember that the *IDS* procedure is an example of multiple hypothesis testing (see 2).

The sensitivity indicators MDB and MIB for *IDS* were also computed according to Equation (33) and (34), respectively. The success rate for outlier detection and outlier identification were taken as being $\tilde{\mathcal{P}}_{CD} = \tilde{\mathcal{P}}_{CI} = 0.8$, respectively. Table 5 provides the values of MDB and MIB for that case of hard constraints.

Table 5. MDB and MIB for the case of hard constraints based on $\alpha' = 0.001$ and $\tilde{\mathcal{P}}_{CD} = \tilde{\mathcal{P}}_{CI} = 0.8$.

Cluster	1 hard constraint		2 hard constraints		3 hard constraints	
	MDB (σ)	MIB (σ)	MDB (σ)	MIB (σ)	MDB (σ)	MIB (σ)
1	7.5	-	6.3	6.3	5.7	5.7
2	6.7	6.8	6.3	6.4	6.3	6.4
3	6.4	6.4	6.3	6.3	5.8	5.8
4	6.4	6.4	6.4	6.4	6.4	6.4

Figure 4 shows the probability of wrong exclusion ($\mathcal{P}_{WE}$). The over-identification cases ($\mathcal{P}_{over+}$ and $\mathcal{P}_{over-}$) were smaller than 0.001 (i.e. they were practically null). There were not statistical overlap ($\mathcal{P}_{ol}$) for clusters 2, 3 and 4. We will discuss more about statistical overlap ($\mathcal{P}_{ol}$) later.
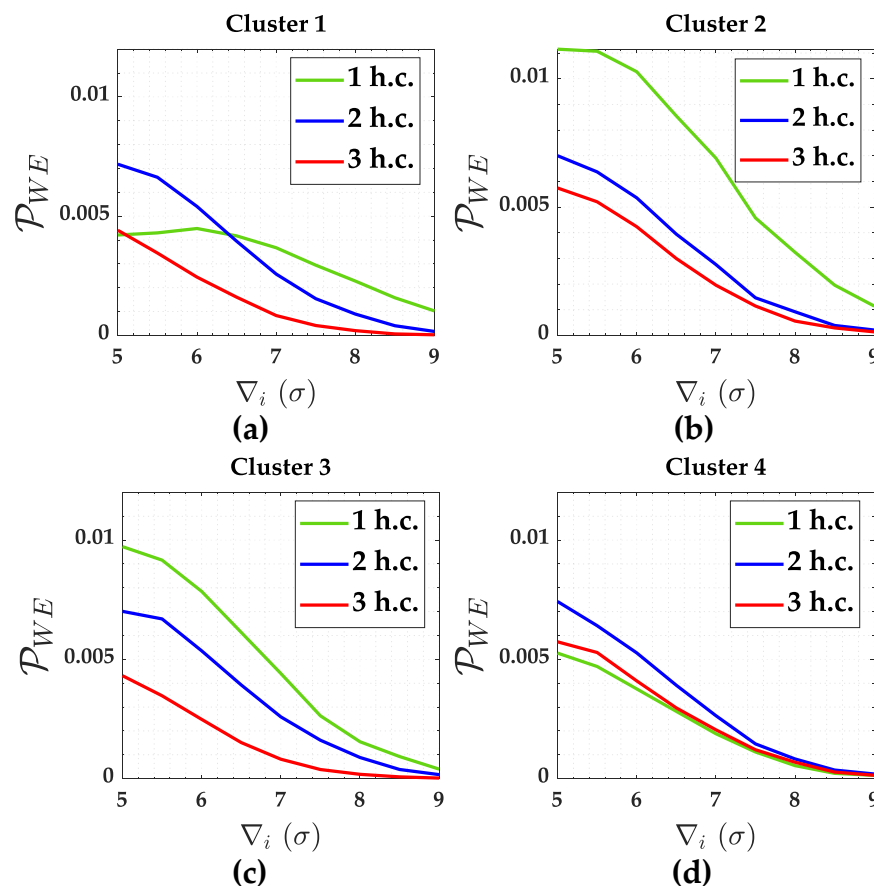

Figure 4. Probability of wrong exclusion ($\mathcal{P}_{WE}$) for the case of hard constraints and for $\alpha' = 0.001$: Cluster 1(**a**), Cluster 2(**b**), Cluster 3(**c**) and Cluster 4(**d**).

3.3. Result of the Soft Constraint Effects on the Iterative Outlier Elimination Procedure

The both configuration in Figure 2(d) and Figure 2(e) were analysed in terms of soft constraints. In that case, the critical values were $\hat{k} = 3.95$, $\hat{k} = 3.95$ and $\hat{k} = 3.92$ for two soft constraints with $\sigma_c = 0.1mm$, $\sigma_c = 1mm$ and $\sigma_c = 10mm$, respectively. In the case of three soft constraints, the critical values found were $\hat{k} = 3.99$, $\hat{k} = 3.99$ and $\hat{k} = 3.96$ for $\sigma_c = 0.1mm$, $\sigma_c = 1mm$ and $\sigma_c = 10mm$, respectively. All these critical values were computed for $\alpha' = 0.001$.

The correlation coefficient between w-test statistics (ρ_{w_i,w_j}) are displayed in Table 6, Table 7 and Table 8 for two soft constraints with standard-deviation 10 times larger than measurements (i.e. $\sigma_c = 10 \times \sigma = 10mm$); 10 times better than measurements (i.e. $\sigma_c = 0.1mm$); and equal to the measurements ($\sigma_c = 1mm$), respectively.

Table 6. Correlation matrix of w-test statistics for the levelling network with two soft constraints with $\sigma_c = 10mm$ in Figure 2(d).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$	$w_{y_{13}}$	$w_{y_{14}}$
w_{y_1}	1.0	0.2	0.1	0.1	0.2	1.0	-0.2	0.0	0.0	0.2	-0.4	-0.1	0.06	-0.1
w_{y_2}		1.0	0.2	0.2	0.3	0.2	0.5	-0.5	-0.2	0.2	0.3	-0.3	0.03	0.0
w_{y_3}			1.0	1.0	0.2	0.1	0.0	0.2	-0.2	0.0	0.1	0.4	0.1	-0.1
w_{y_4}				1.0	0.2	0.1	0.0	0.2	-0.2	0.0	0.1	0.4	-0.1	0.1
w_{y_5}					1.0	0.2	-0.2	0.2	0.5	-0.5	0.3	-0.3	0.0	0.0
w_{y_6}						1.0	-0.2	0.0	0.0	0.2	-0.4	-0.1	-0.1	0.1
w_{y_7}							1.0	-0.3	-0.3	-0.4	-0.4	-0.1	0.0	0.0
w_{y_8}								1.0	-0.4	-0.3	-0.1	-0.4	0.0	0.0
w_{y_9}									1.0	-0.3	0.1	0.4	0.0	0.0
$w_{y_{10}}$										1.0	0.4	0.1	0.0	0.0
$w_{y_{11}}$											1.0	-0.1	0.0	0.0
$w_{y_{12}}$												1.0	0.0	0.0
$w_{y_{13}}$													1.0	-1.0
$w_{y_{14}}$														1.0

Table 7. Correlation matrix of w-test statistics for the levelling network with two soft constraints with $\sigma_c = 0.1mm$ in Figure 2(d).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$	$w_{y_{13}}$	$w_{y_{14}}$
w_{y_1}	1.0	0.4	0.4	-0.3	-0.1	0.4	-0.3	0.1	0.1	0.1	-0.4	-0.1	0.6	-0.6
w_{y_2}		1.0	0.4	-0.1	0.2	-0.1	0.4	-0.4	-0.1	0.1	0.3	-0.3	0.4	-0.4
w_{y_3}			1.0	0.4	-0.1	-0.3	-0.1	0.3	-0.1	-0.1	0.1	0.4	0.6	-0.6
w_{y_4}				1.0	0.4	0.4	0.1	0.1	-0.3	0.1	0.1	0.4	-0.6	0.6
w_{y_5}					1.0	0.4	-0.1	0.1	0.4	-0.4	0.3	-0.3	-0.4	0.4
w_{y_6}						1.0	-0.1	-0.1	-0.1	0.3	-0.4	-0.1	-0.6	0.6
w_{y_7}							1.0	-0.4	-0.3	-0.4	-0.4	-0.1	-0.2	0.2
w_{y_8}								1.0	-0.4	-0.3	-0.1	-0.4	0.2	-0.2
w_{y_9}									1.0	-0.4	0.1	0.4	0.2	-0.2
$w_{y_{10}}$										1.0	0.4	0.1	-0.2	0.2
$w_{y_{11}}$											1.0	-0.1	0.0	0.0
$w_{y_{12}}$												1.0	0.0	0.0
$w_{y_{13}}$													1.0	-1.0
$w_{y_{14}}$														1.0

Table 8. Correlation matrix of w-test statistics for the levelling network with two soft constraints with $\sigma_c = 1mm$ in Figure 2(d).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$	$w_{y_{13}}$	$w_{y_{14}}$
w_{y_1}	1.0	0.3	0.2	-0.1	0.1	0.7	-0.3	0.0	0.1	0.2	-0.4	-0.1	0.4	-0.4
w_{y_2}		1.0	0.3	0.1	0.3	0.1	0.4	-0.4	-0.1	0.1	0.3	-0.3	0.3	-0.3
w_{y_3}			1.0	0.7	0.1	-0.1	0.0	0.3	-0.2	-0.1	0.1	0.4	0.4	-0.4
w_{y_4}				1.0	0.3	0.2	0.1	0.2	-0.3	0.0	0.1	0.4	-0.4	0.4
w_{y_5}					1.0	0.3	-0.1	0.1	0.4	-0.4	0.3	-0.3	-0.3	0.3
w_{y_6}						1.0	-0.2	-0.1	0.0	0.3	-0.4	-0.1	-0.4	0.4
w_{y_7}							1.0	-0.3	-0.3	-0.4	-0.4	-0.1	-0.1	0.1
w_{y_8}								1.0	-0.4	-0.3	-0.1	-0.4	0.1	-0.1
w_{y_9}									1.0	-0.3	0.1	0.4	0.1	-0.1
$w_{y_{10}}$										1.0	0.4	0.1	-0.1	0.1
$w_{y_{11}}$											1.0	-0.1	0.0	0.0
$w_{y_{12}}$												1.0	0.0	0.0
$w_{y_{13}}$													1.0	-1.0
$w_{y_{14}}$														1.0

Table 9 gives the local redundancy (r_i), the standard-deviation of the LS-estimated outlier σ_{∇_i} and the maximum absolute correlation ($max_{\rho_{w_i,w_j}}$) for the scenarios of two constraints.

Table 9. Local redundancy (r_i), standard-deviation of the LS-estimated outlier σ_{∇_i}(mm) and the maximum absolute correlation ($max_{\rho_{w_i,w_j}}$) for each scenario of two soft constraints.

Measurement	$\sigma_c = 0.1mm$			$\sigma_c = 1mm$			$\sigma_c = 10mm$		
	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$
y_1	0.581	1.312	0.564	0.471	1.457	0.681	0.397	1.587	0.994
y_2	0.582	1.311	0.376	0.533	1.369	0.423	0.501	1.413	0.471
y_3	0.581	1.312	0.564	0.471	1.457	0.681	0.397	1.587	0.994
y_4	0.581	1.312	0.564	0.471	1.457	0.681	0.397	1.587	0.994
y_5	0.582	1.311	0.376	0.533	1.369	0.423	0.501	1.413	0.471
y_6	0.581	1.312	0.564	0.471	1.457	0.681	0.397	1.587	0.994
y_7	0.583	1.310	0.359	0.571	1.324	0.423	0.563	1.333	0.471
y_8	0.583	1.310	0.359	0.571	1.324	0.423	0.563	1.333	0.471
y_9	0.583	1.310	0.359	0.571	1.324	0.423	0.563	1.333	0.471
y_{10}	0.583	1.310	0.359	0.571	1.324	0.423	0.563	1.333	0.471
y_{11}	0.583	1.309	0.358	0.583	1.309	0.398	0.583	1.309	0.433
y_{12}	0.583	1.309	0.358	0.583	1.309	0.398	0.583	1.309	0.433
y_{13}	0.007	1.163	1.000	0.300	1.826	1.000	0.497	14.189	1.000
y_{14}	0.007	1.163	1.000	0.300	1.826	1.000	0.497	14.189	1.000

From Table 9, five clusters were defined for each case of two soft constraints, i.e. for the case where heights A and D were given as soft constraints in Figure 2(d), as follows:

- Cluster 1: y_1, y_3, y_4 and y_6.
- Cluster 2: y_2 and y_5.
- Cluster 3: y_7, y_8, y_9 and y_{10}.
- Cluster 4: y_{11} and y_{12}.
- Cluster 5: y_{13} and y_{14}.

The probabilities of correct identification ($\mathcal{P}_{CI}$) and correct detection ($\mathcal{P}_{CD}$) for the measurements (Cluster 1 to Cluster 4) subject to the scenarios of two soft constraints (heights A and D) are displayed in Figure 5.
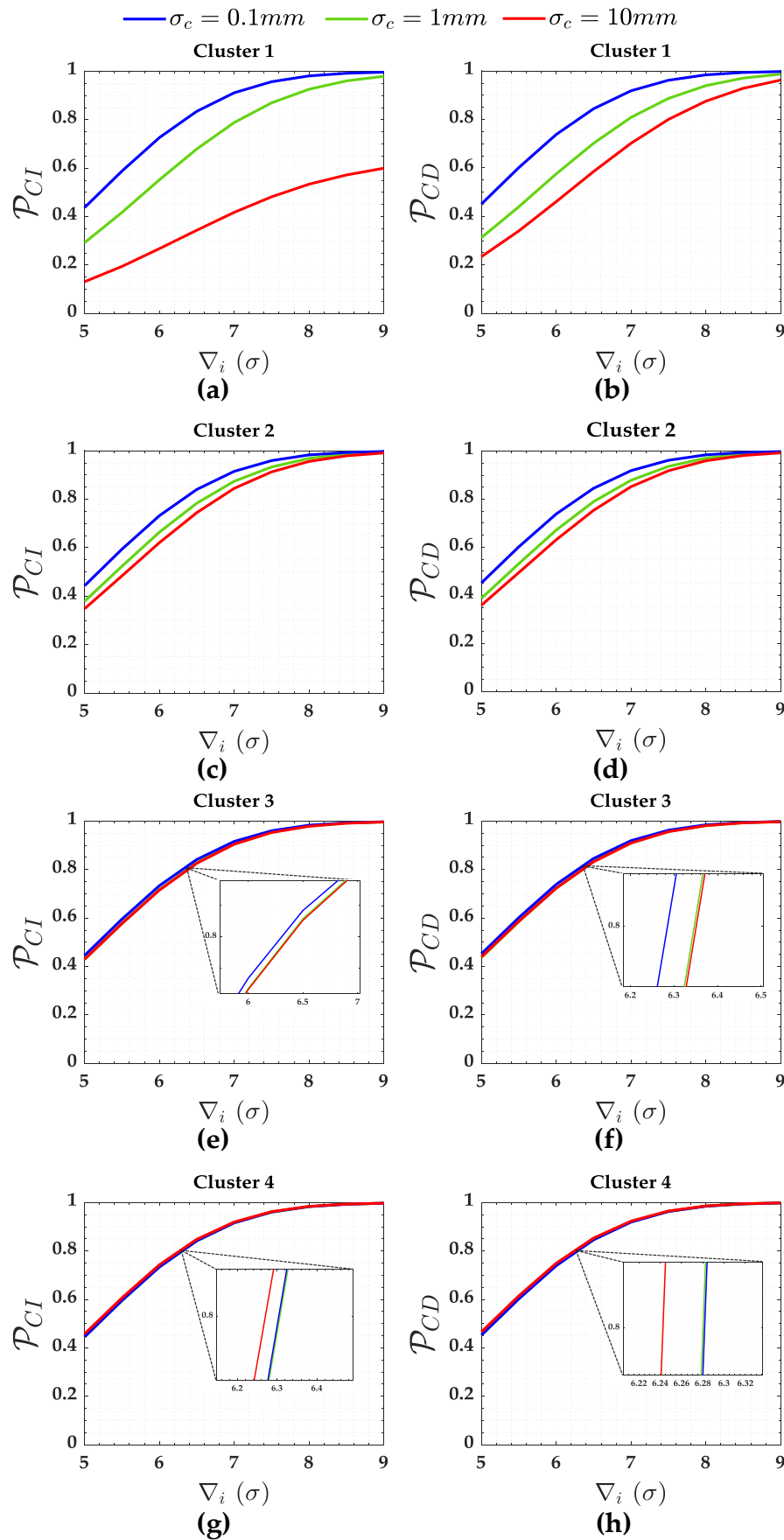
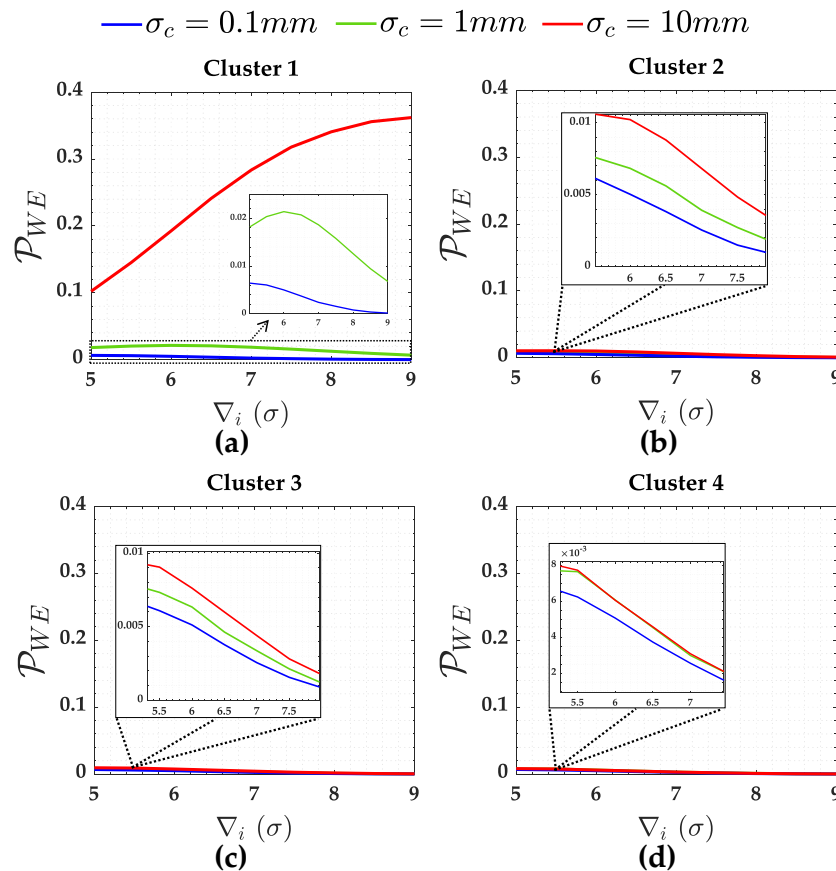**Figure 5.** Probability of correct identification ($\mathcal{P}_{CI}$) and Probability of correct detection ($\mathcal{P}_{CD}$) for the measurements subject to the scenarios of two soft constraints for $\alpha' = 0.001$: Cluster 1(**a**,**b**), Cluster 2(**c**,**d**), Cluster 3(**e**,**f**) and Cluster 4(**g**,**h**).

Note that the Cluster 5 is associated with the two soft constraints (i.e. y_{13} and y_{14}). The probability of correct identification ($\mathcal{P}_{CI}$) for these both soft constraints were null. However, the probability of correct detection were not. Figure 6 shows the probability of correct detection ($\mathcal{P}_{CD}$) for these two soft constraints (i.e. heights A and D).
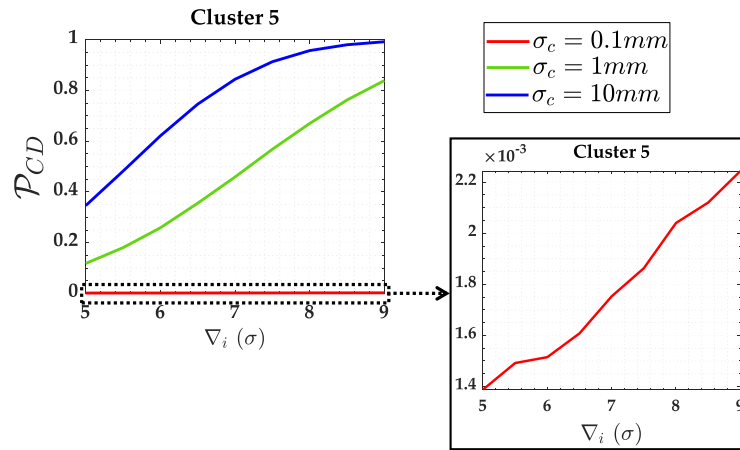
Figure 6. Probability of correct detection $\mathcal{P}_{CD}$ for the two soft constraints (Cluster 5: heights A and D) and for $\alpha' = 0.001$.

The probability of wrong exclusion $\mathcal{P}_{WE}$ for the measurements (Cluster 1 to Cluster 4) subject to the scenarios of two soft constraints (heights A and D) are displayed in Figure 7. Figure 8 gives the the probability of wrong exclusion $\mathcal{P}_{WE}$ for that two constraints (i.e. heights A and D).

Figure 7. Probability of wrong exclusion $\mathcal{P}_{WE}$ for the measurements subject to the scenarios of two soft constraints for $\alpha' = 0.001$: Cluster 1(**a**), Cluster 2(**b**), Cluster 3(**c**) and Cluster 4(**d**).
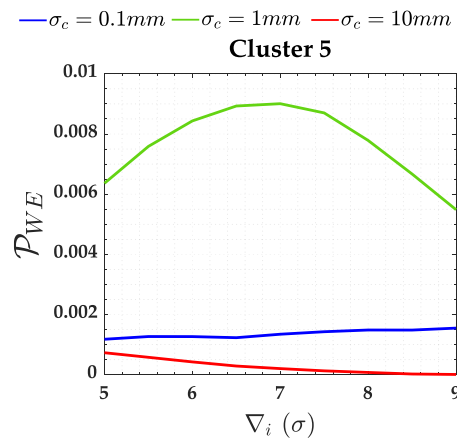
Figure 8. Probability of wrong exclusion $\mathcal{P}_{WE}$ for the two soft constraints (Cluster 5: heights A and D) and for $\alpha' = 0.001$.

The over-identification cases ($\mathcal{P}_{over+}$ and $\mathcal{P}_{over-}$) and the statistical overlap ($\mathcal{P}_{ol}$) were practically null for that case.

The sensitivity indicators (MDB and MIB) for each scenario of two soft constraints are displayed in Table 10.

Table 10. MDB and MIB for the case of two soft constraints based on $\alpha' = 0.001$ and $\tilde{\mathcal{P}}_{CD} = \tilde{\mathcal{P}}_{CI} = 0.8$.

Cluster	$\sigma_c = 10mm$ MDB (σ)	MIB (σ)	$\sigma_c = 1mm$ MDB (σ)	MIB (σ)	$\sigma_c = 0.1mm$ MDB (σ)	MIB (σ)
1	7.5	25	7	7.1	6.3	6.3
2	6.8	6.8	6.6	6.6	6.3	6.3
3	6.4	6.4	6.4	6.4	6.3	6.3
4	6.3	6.3	6.3	6.3	6.3	6.3
5	6.8	-	8.8	-	57	-

As with the case of two soft constraints, Tables 11,12 and 13 show the correlations (ρ_{w_i,w_j}) for the case where there were three soft constraints, i.e. for the network configuration in Figure 2(e).

Table 11. Correlation matrix of w-test statistics for the levelling network with three soft constraints with $\sigma_c = 10mm$ in Figure 2(e).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$	$w_{y_{13}}$	$w_{y_{14}}$	$w_{y_{15}}$
w_{y_1}	1.0	0.2	0.1	0.1	0.2	1.0	-0.2	0.0	0.0	0.2	-0.4	-0.1	0.1	0.0	0.0
w_{y_2}		1.0	0.2	0.2	0.3	0.2	0.5	-0.5	-0.2	0.2	0.3	-0.3	0.0	0.0	0.0
w_{y_3}			1.0	1.0	0.2	0.1	0.0	0.2	-0.2	0.0	0.1	0.4	0.0	-0.1	0.0
w_{y_4}				1.0	0.2	0.1	0.0	0.2	-0.2	0.0	0.1	0.4	0.0	0.1	0.0
w_{y_5}					1.0	0.2	-0.2	0.2	0.5	-0.5	0.3	-0.3	0.0	0.0	0.0
w_{y_6}						1.0	-0.2	0.0	0.0	0.2	-0.4	-0.1	-0.1	0.0	0.0
w_{y_7}							1.0	-0.3	-0.3	-0.4	-0.4	-0.1	0.0	0.0	0.0
w_{y_8}								1.0	-0.4	-0.3	-0.1	-0.4	0.0	0.0	0.0
w_{y_9}									1.0	-0.3	0.1	0.4	0.0	0.0	0.0
$w_{y_{10}}$										1.0	0.4	0.1	0.0	0.0	0.0
$w_{y_{11}}$											1.0	-0.1	0.0	0.0	0.0
$w_{y_{12}}$												1.0	0.0	0.0	0.0
$w_{y_{13}}$													1.0	-0.5	-0.5
$w_{y_{14}}$														1.0	-0.5
$w_{y_{15}}$															1.0

Table 12. Correlation matrix of w-test statistics for the levelling network with three soft constraints with $\sigma_c = 0.1mm$ in Figure 2(e).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$	$w_{y_{13}}$	$w_{y_{14}}$	$w_{y_{15}}$
w_{y_1}	1.0	0.3	0.1	-0.1	-0.1	0.1	-0.4	-0.1	-0.1	-0.1	-0.3	-0.1	0.7	-0.1	-0.4
w_{y_2}		1.0	0.3	-0.1	0.1	-0.1	0.3	-0.3	-0.1	0.1	0.3	-0.3	0.3	-0.3	0.0
w_{y_3}			1.0	0.1	-0.1	-0.1	0.1	0.4	0.1	0.1	0.1	0.3	0.1	-0.7	0.4
w_{y_4}				1.0	0.3	0.1	-0.1	-0.1	-0.4	-0.1	0.1	0.3	-0.1	0.7	-0.4
w_{y_5}					1.0	0.3	-0.1	0.1	0.3	-0.3	0.3	-0.3	-0.3	0.3	0.0
w_{y_6}						1.0	0.1	0.1	0.1	0.4	-0.3	-0.1	-0.7	0.1	0.4
w_{y_7}							1.0	-0.1	-0.1	-0.1	-0.3	-0.1	-0.4	-0.1	0.4
w_{y_8}								1.0	-0.1	-0.1	-0.1	-0.3	-0.1	-0.4	0.4
w_{y_9}									1.0	-0.1	0.1	0.3	-0.1	-0.4	0.4
$w_{y_{10}}$										1.0	0.3	0.1	-0.4	-0.1	0.4
$w_{y_{11}}$											1.0	-0.1	0.0	0.0	0.0
$w_{y_{12}}$												1.0	0.0	0.0	0.0
$w_{y_{13}}$													1.0	-0.2	-0.6
$w_{y_{14}}$														1.0	-0.6
$w_{y_{15}}$															1.0

Table 13. Correlation matrix of w-test statistics for the levelling network with three soft constraints with $\sigma_c = 1mm$ in Figure 2(e).

	w_{y_1}	w_{y_2}	w_{y_3}	w_{y_4}	w_{y_5}	w_{y_6}	w_{y_7}	w_{y_8}	w_{y_9}	$w_{y_{10}}$	$w_{y_{11}}$	$w_{y_{12}}$	$w_{y_{13}}$	$w_{y_{14}}$	$w_{y_{15}}$
w_{y_1}	1.0	0.3	0.1	0.0	0.1	0.6	-0.3	0.0	0.0	0.1	-0.4	-0.1	0.5	-0.2	-0.2
w_{y_2}		1.0	0.3	0.1	0.3	0.1	0.4	-0.4	-0.1	0.1	0.3	-0.3	0.2	-0.2	0.0
w_{y_3}			1.0	0.6	0.1	0.0	0.0	0.3	-0.1	0.0	0.1	0.4	0.2	-0.5	0.2
w_{y_4}				1.0	0.3	0.1	0.0	0.1	-0.3	0.0	0.1	0.4	-0.2	0.5	-0.2
w_{y_5}					1.0	0.3	-0.1	0.1	0.4	-0.4	0.3	-0.3	-0.2	0.2	0.0
w_{y_6}						1.0	-0.1	0.0	0.0	0.3	-0.4	-0.1	-0.5	0.2	0.2
w_{y_7}							1.0	-0.3	-0.2	-0.3	-0.4	-0.1	-0.2	0.0	0.2
w_{y_8}								1.0	-0.3	-0.2	-0.1	-0.4	0.0	-0.2	0.2
w_{y_9}									1.0	-0.3	0.1	0.4	0.0	-0.2	0.2
$w_{y_{10}}$										1.0	0.4	0.1	-0.2	0.0	0.2
$w_{y_{11}}$											1.0	-0.1	0.0	0.0	0.0
$w_{y_{12}}$												1.0	0.0	0.0	0.0
$w_{y_{13}}$													1.0	-0.4	-0.5
$w_{y_{14}}$														1.0	-0.5
$w_{y_{15}}$															1.0

Table 14. Local redundancy (r_i), standard-deviation of the LS-estimated outlier σ_{∇_i} (mm) and the maximum absolute correlation ($max_{\rho_{w_i,w_j}}$) for each scenario of three soft constraints.

Measurement	$\sigma_c = 0.1mm$			$\sigma_c = 1mm$			$\sigma_c = 10mm$		
	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$	r_i	σ_{∇_i}	$max_{\rho_{w_i,w_j}}$
y_1	0.702	1.194	0.660	0.502	1.411	0.577	0.398	1.586	0.992
y_2	0.582	1.311	0.326	0.533	1.369	0.412	0.501	1.413	0.470
y_3	0.702	1.194	0.660	0.502	1.411	0.577	0.398	1.586	0.992
y_4	0.702	1.194	0.660	0.502	1.411	0.577	0.398	1.586	0.992
y_5	0.582	1.311	0.326	0.533	1.369	0.412	0.501	1.413	0.470
y_6	0.702	1.194	0.660	0.502	1.411	0.577	0.398	1.586	0.992
y_7	0.704	1.192	0.415	0.602	1.289	0.412	0.563	1.333	0.470
y_8	0.704	1.192	0.415	0.602	1.289	0.412	0.563	1.333	0.470
y_9	0.704	1.192	0.415	0.602	1.289	0.412	0.563	1.333	0.470
y_{10}	0.704	1.192	0.415	0.602	1.289	0.412	0.563	1.333	0.470
y_{11}	0.583	1.309	0.326	0.583	1.309	0.385	0.583	1.309	0.433
y_{12}	0.583	1.309	0.326	0.583	1.309	0.385	0.583	1.309	0.433
y_{13}	0.012	0.904	0.660	0.425	1.534	0.542	0.663	12.283	0.501
y_{14}	0.012	0.904	0.660	0.425	1.534	0.542	0.663	12.283	0.501
y_{15}	0.019	0.718	0.63	0.5	1.414	0.542	0.665	12.268	0.501

The probabilities of correct identification ($\mathcal{P}_{CI}$) and correct detection $\mathcal{P}_{CD}$ for the measurements (Cluster 1 to Cluster 4) subject to the scenarios of three soft constraints (heights A, D and G) are displayed in Figure 9.
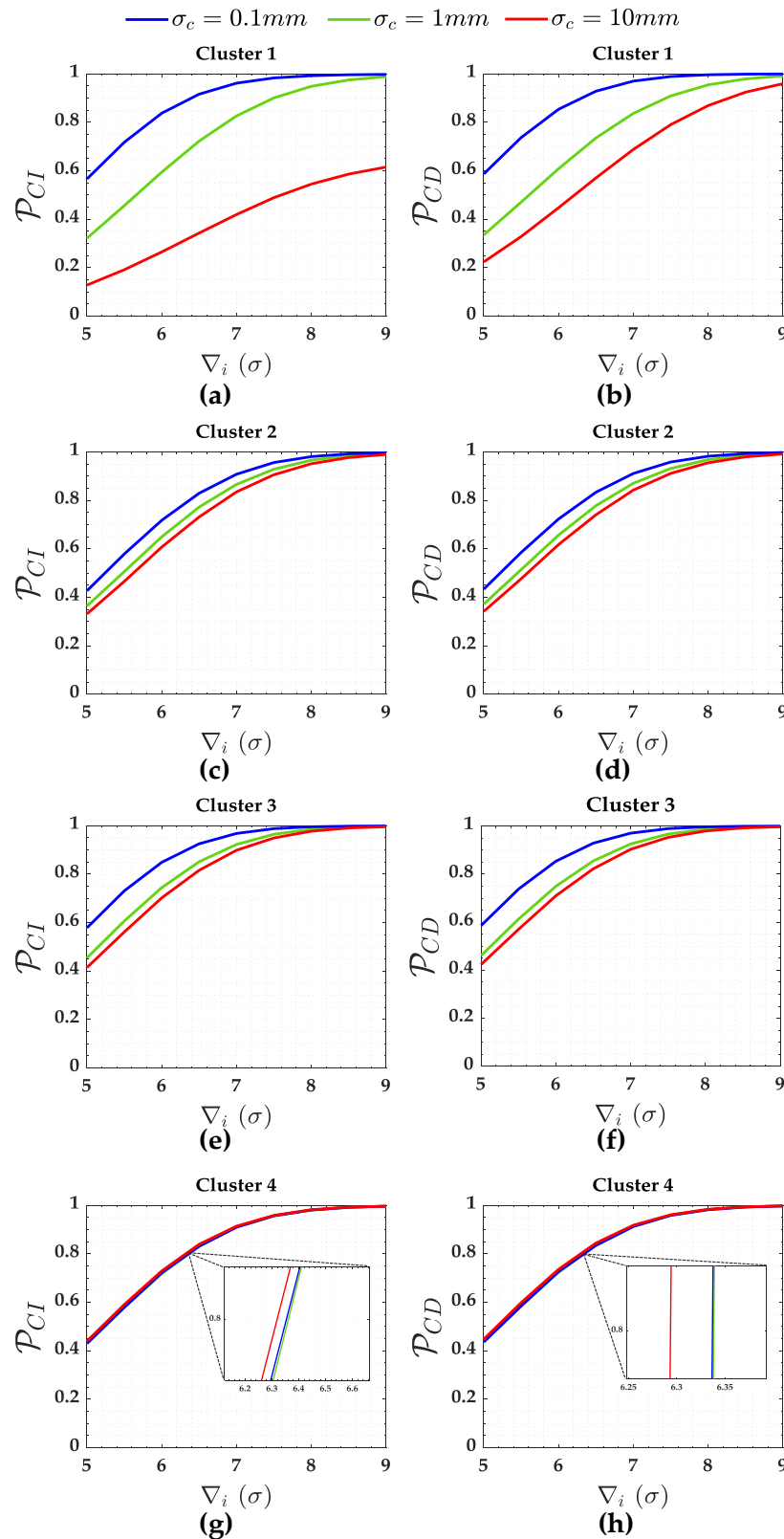
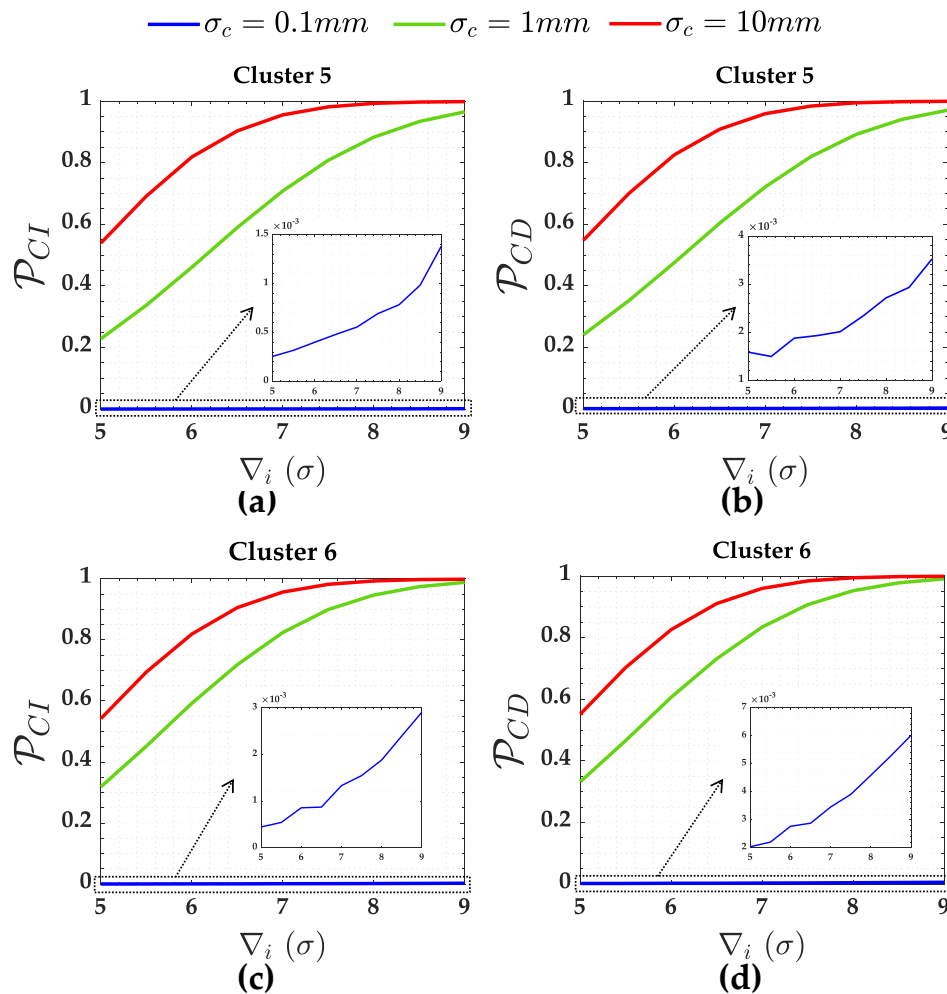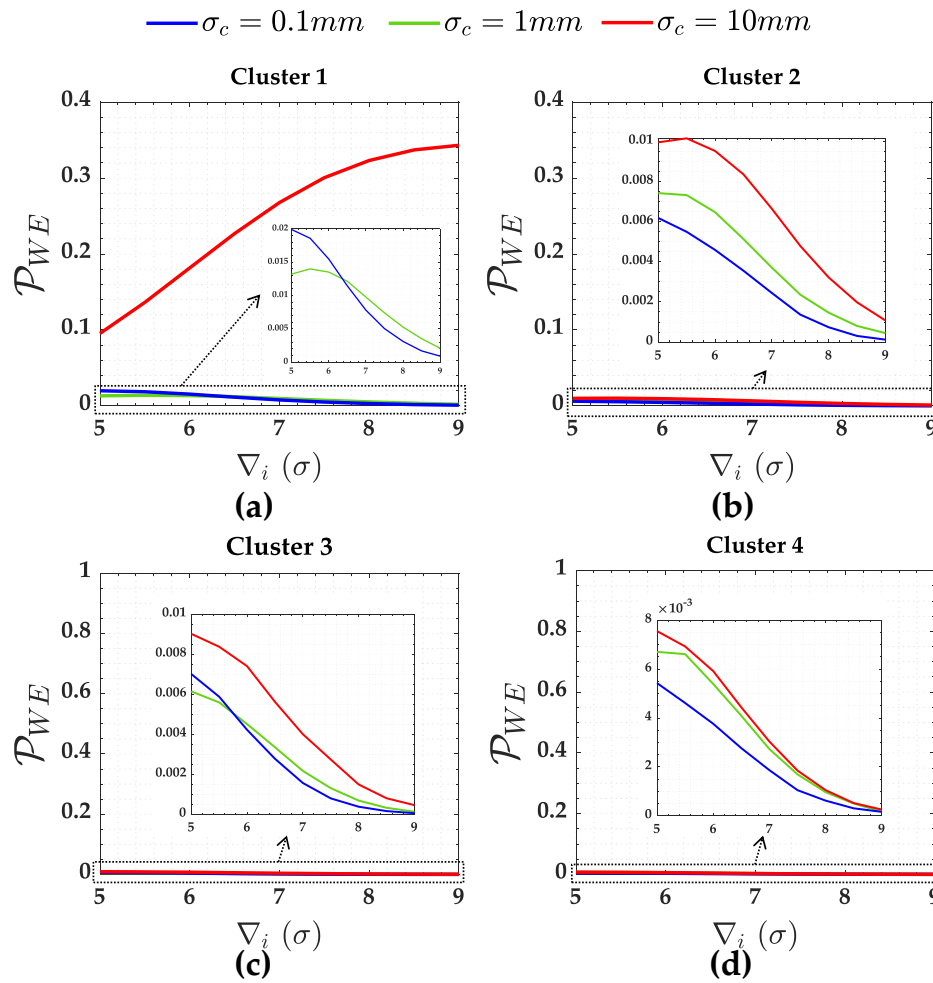**Figure 9.** Probability of correct identification ($\mathcal{P}_{CI}$) and Probability of correct detection ($\mathcal{P}_{CD}$) for the measurements subject to the scenarios of three soft constraints for $\alpha' = 0.001$: Cluster 1(**a**,**b**), Cluster 2(**c**,**d**), Cluster 3(**e**,**f**) and Cluster 4(**g**,**h**).

The probabilities of correct identification ($\mathcal{P}_{CI}$) and correct detection ($\mathcal{P}_{CD}$) in Figure 8 were computed for the clusters based on Table 14, as follows:

- Cluster 1: y_1, y_3, y_4 and y_6.
- Cluster 2: y_2 and y_5.
- Cluster 3: y_7, y_8, y_9 and y_{10}.
- Cluster 4: y_{11} and y_{12}.
- Cluster 5: y_{13} and y_{14}.
- Cluster 6: y_{15}.

Figure 10 shows the probabilities of correct identification $\mathcal{P}_{CI}$ and correct detection $\mathcal{P}_{CD}$ for the three soft constraints, i.e. for Cluster 5 (heights A and D) and Cluster 6 (height G) in Figure 2(e).

Figure 10. Probability of correct identification ($\mathcal{P}_{CI}$) and Probability of correct detection ($\mathcal{P}_{CD}$) for the three constraints and for $\alpha' = 0.001$: Cluster 5(**a**,**b**) and Cluster 6(**c**,**d**).

The probability of wrong exclusion $\mathcal{P}_{WE}$ for the measurements (Cluster 1 to Cluster 4) subject to the scenarios of three soft constraints (heights A, D and G) are displayed in Figure 11. Figure 12 gives the the probability of wrong exclusion $\mathcal{P}_{WE}$ for that three constraints (i.e. heights A, D and G).

Figure 11. Probability of wrong exclusion $\mathcal{P}_{WE}$ for the measurements subject to the scenarios of three soft constraints and for $\alpha' = 0.001$: Cluster 1(**a**), Cluster 2(**b**), Cluster 3(**c**) and Cluster 4(**d**).
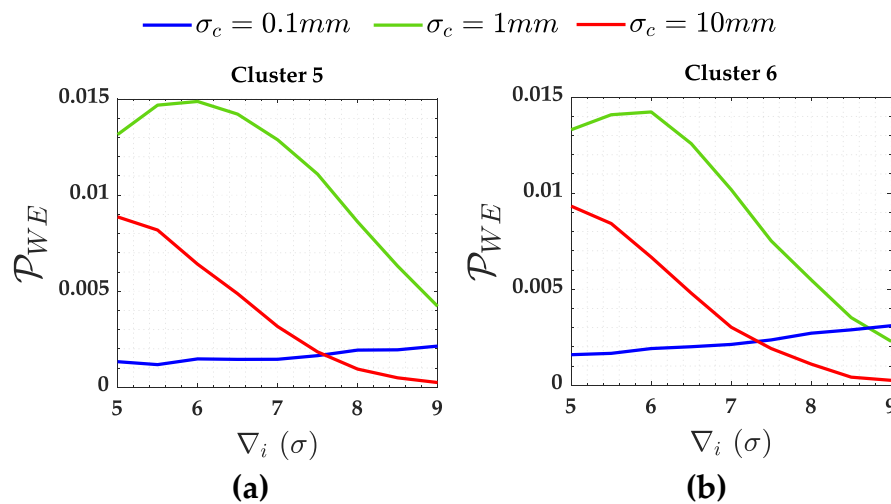
Figure 12. Probability of wrong exclusion ($\mathcal{P}_{WE}$) for the three constraints and for $\alpha' = 0.001$: Cluster 5(**a**) and Cluster 6(**b**).

The over-identification cases ($\mathcal{P}_{over+}$ and $\mathcal{P}_{over-}$) and the statistical overlap ($\mathcal{P}_{ol}$) were also practically null for that case of three soft constraints.

The sensitivity indicators (MDB and MIB) for each scenario of three soft constraints are displayed in Table 15.

Table 15. MDB and MIB for the case of three soft constraints based on $\alpha' = 0.001$ and $\tilde{\mathcal{P}}_{CD} = \tilde{\mathcal{P}}_{CI} = 0.8$.

Cluster	$\sigma_c = 10mm$		$\sigma_c = 1mm$		$\sigma_c = 0.1mm$	
	MDB (σ)	MIB (σ)	MDB (σ)	MIB (σ)	MDB (σ)	MIB (σ)
1	7.5	22	6.8	6.9	5.8	5.9
2	6.8	6.9	6.6	6.7	6.4	6.4
3	6.4	6.4	6.3	6.3	5.8	5.8
4	6.3	6.3	6.3	6.3	6.3	6.3
5	5.9	6.0	7.4	7.5	43.5	45
6	5.9	5.9	6.9	6.9	34.6	35.5

4. Discussion

We start from the scenario of one hard constraint in Figure 2(a). Table 4 reveal that the maximum correlation between w-test statistics for the measurements constituting Cluster 1 is exactly equal to 1.00 (i.e. $max_{\rho_{w_i,w_j}} = 1.00$). This means that the measurements belonging to Cluster 1 are those that are connected with unknown heights whose connections are limited to only two. The both unknown heights A and D are tied only to two measurements (i.e. y_1 and y_6 linked to A, and y_3 and y_4 linked to D), therefore if an outlier occurred in one of these measurements, we would only be able to analyse the consistency between them, but we would not be able to distinguish which of them was contaminated by an outlier. This means that we would only be able to detect, because the w-test statistics could be larger than a critical value $\hat{k}$, however, in that case, the values of w-test statistics would be the same, and therefore we would not have only one unique maximum w-test statistics, but actually would have four maximum w-test statistics. In other words, the equation systems associated with the measurements of Cluster 1 are linearly dependent [83]. Therefore, there is no reliability in terms of outlier identification for the Cluster 1, as can be seen in Figure 3(a).

From Figure 3(b), we can note that there is reliability in terms of outlier detection for Cluster 1, and it is caused by overlapping w-test statistics. The probability of statistics overlap $\mathcal{P}_{ol}$ for Cluster 1 in this scenario of minimally constrained network is displayed in Figure 13.
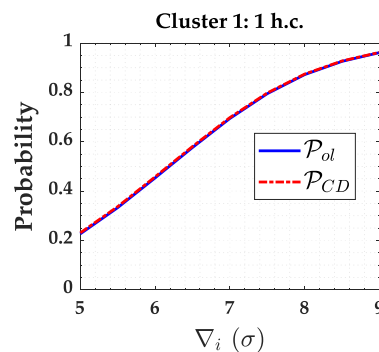
Figure 13. Probability of correct detection $\mathcal{P}_{CD}$ and statistical overlap ($\mathcal{P}_{ol}$) for the Cluster 1 subject to one hard constraint and for $\alpha' = 0.001$.

The problem of not having more connections (i.e. more measurements) for the unknown heights A and D in the case of one hard constraint with G fixed is overcome when these heights (A and D) are taken as hard constraints in 2(b) or when the heights A, D and G are hard constraints in figure 2(c). Figure 3(a, b) show that the measurements of Cluster 1 are now able to identify an outlier when two hard constraints (A and D fixed) are in play. It is also verified to the case of three hard constraints (A, D and G fixes) in Figure 3(e, f), i.e. there is reliability in terms of both outlier detection and identification for these measurements in those conditions.

One can argue that the more constraints, the larger the probability of correct identification ($\mathcal{P}_{CI}$) and correct detection ($\mathcal{P}_{CD}$). This claim is not entirely correct. This may be true in case where all points on the network have the same number of connections. However, in the example given, if both heights B and G were selected as hard constraints, the measurements of the Cluster 1 (i.e. y_1, y_6, y_3 and y_4) would still have their least-squares residuals fully correlated (i.e. $\rho_{w_{y_1}, w_{y_6}} = 1.00$ and $\rho_{w_{y_3}, w_{y_4}} = 1.00$), and therefore there would be no reliability in terms of outlier identification for these measurements.

From Table 5, we observe different behaviour for the clusters as follows:

- Cluster 1: there was no MIB for the case of having only one single hard constraint, whereas $MDB = MIB$ for the other cases. However, both MDB and MIB decrease significantly with the increase in the number of hard constraints.
- Cluster 2: MDB slightly smaller than MIB. Both MDB and MIB were practically the same for the case of having two or three hard constraints.
- Cluster 3: $MDB = MIB$ for all cases of hard constraints, however both MDB and MIB decrease significantly with the increase in the number of hard constraints.
- Cluster 4: MDB e MIB were equal for all cases.

In terms of outlier detection and identification, therefore, Cluster 1 was more sensitive to constraints, Cluster 3 relatively sensitive to constraints, whereas the Cluster 4 completely insensitive to constraints and Cluster 2 relatively insensitive to constraints. This is also can be seen in Figure 3. The reason for this is that the local redundancy (r_i) of the Cluster 1 increased with the increase of the number of hard constraints, whereas the Cluster 4 have remained the same, as can be seen in Table 4.
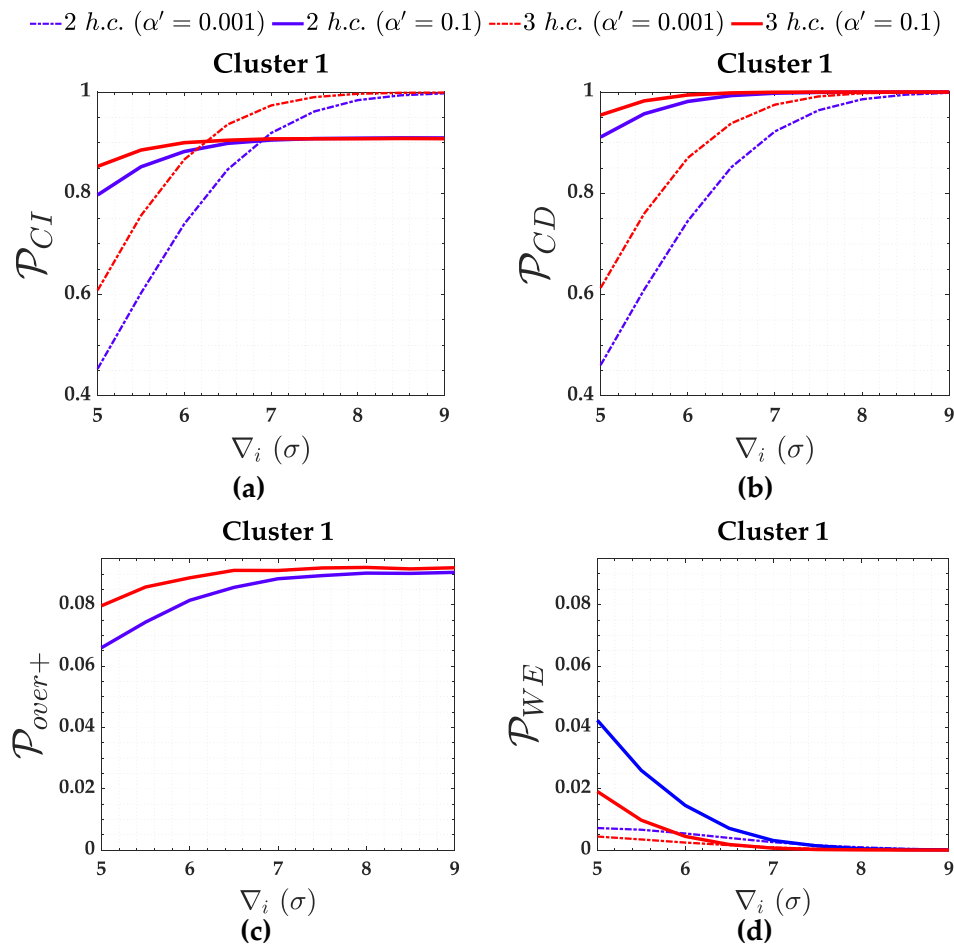
Figure 14. Probabilities of correct identification $\mathcal{P}_{CI}$ (**a**), correct detection $\mathcal{P}_{CD}$ (**b**), over-identification positive $\mathcal{P}_{over+}$ (**c**) and wrong exclusion $\mathcal{P}_{WE}$ for the Cluster 1 subject to two and three hard constraints and for $\alpha' = 0.001$ and $\alpha' = 0.1$.

Leaving aside the cases of statistical overlap ($\mathcal{P}_{ol}$), the network presents low least-squares residuals correlation ($\rho_{w_i,w_j} < 0.5$) and high local redundancy ($r_i > 0.5$). Because of this, the probabilities of wrong exclusion ($\mathcal{P}_{WE}$) were less than 1%, as can be seen in Figure 4. The over-identification cases ($\mathcal{P}_{over+}$ and $\mathcal{P}_{over-}$) were practically null. Consequently, $\mathcal{P}_{CI} \approx \mathcal{P}_{CD}$. Due to this fact, the family-wise error rate (α') should be increased in order to have more success rate in the outlier detection and identification [45].

From Figure 14, we observe that increasing the α' increases the both probabilities of correct identification ($\mathcal{P}_{CI}$) and correct detection ($\mathcal{P}_{CD}$) for outlier magnitude from 5σ to 6σ in the case of three hard constraints and from 5σ to 6.8σ in the case of two hard constraints. Although the rates of $\mathcal{P}_{over+}$ and $\mathcal{P}_{WE}$ also increase, they are not significant when compared to the improvement of correct identification ($\mathcal{P}_{CI}$) and detection ($\mathcal{P}_{CD}$). This same analysis can be done for the other clusters.

In terms of soft constraints for the cases of two constraints in Figure 2(d), we observe from Table 9 that the larger the relaxation of the constraint (i.e. the larger the standard-deviation of the constraint σ_c), the larger the residuals correlation (ρ_{w_i,w_j}) and the standard-deviation of the outlier σ_{∇_i}, and the smaller the local redundancy (r_i). Consequently, the probabilities of correct identification ($\mathcal{P}_{CI}$) and detection ($\mathcal{P}_{CD}$) get smaller and smaller with the relaxation of the constraints, whereas the probability of wrong exclusion gets larger ($\mathcal{P}_{WE}$). This can be more clearly verified in Figure 5(a, b) and Figure 7(a) for Cluster 1, whose measurements are connected with the constraints A and D (i.e. y_{13} and y_{14} in Table 9, respectively).

Note from Figure 7(a) that the probability of wrong exclusion ($\mathcal{P}_{WE}$) increases as the magnitude of the outlier (∇_i) increases. However, this is true up to a certain limit of outlier magnitude. The effect of residuals correlation ρ_{w_i,w_j} on the rates of wrong exclusion ($\mathcal{P}_{WE}$) and correct identification ($\mathcal{P}_{CI}$) tends to decrease with the increase in the magnitude of the outlier ∇_i. This effect is more clearly verified for Cluster 1 in case where the precision of the constraints are ten times worse than the measurements $\sigma_c = 10\sigma = 10mm$. Note from Figure 5 that identifying an outlier in the Cluster 1 (i.e. y_1, y_3, y_4 and y_6) when $\sigma_c = 10mm$ is more difficult than the other clusters. This is due to the fact the Cluster 1 has a higher residuals correlation $\rho_{w_i,w_j} = 0.994$ than other clusters. In general, therefore, we observe that the larger the relaxation of the constraints, the larger is the effect of the correlation ρ_{w_i,w_j} on the success rate of outlier identification ($\mathcal{P}_{CI}$). Consequently, the higher the sensitivity indicator for outlier identification (MIB). Table 10 reveals that the ratio between MIB and MDB for the Cluster 1 and for the scenario where the standard-deviations of that two soft constraints are $\sigma_c = 10mm$ is MIB/MDB=25/7.5=3.3. On the other hand, the relationship between MIB and MDB is practically one (i.e. MIB/MDB=1.0) for the others scenarios.
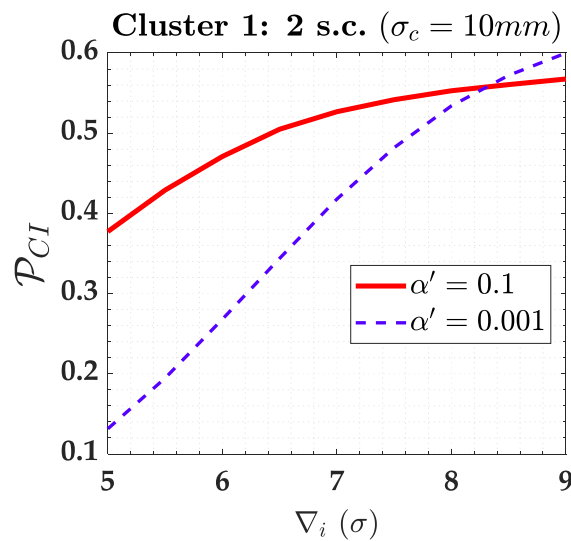
Figure 15. Probabilities of correct identification ($\mathcal{P}_{CI}$) for the Cluster 1 subject to two soft constraints (2 s.c.) A and D for $\alpha' = 0.001$ and $\alpha' = 0.1$.

If the FWE rate (α') were increased for the case where the two soft constraints of $\sigma_c = 10mm$ are in play, we would not have great advantages for Cluster 1, due to its high residuals correlation ($\rho_{w_i,w_j} = 99.4\%$). From Figure 15, we can observe that the probability of correct identification ($\mathcal{P}_{CI}$) for outlier magnitudes from 5σ to 8σ is effectively larger to a user-defined $\alpha' = 0.1$ than $\alpha' = 0.001$. However, the success rate is still less than 80%, i.e. $\mathcal{P}_{CI} < 0.8$. Note, for example, the correct identification rate is $\mathcal{P}_{CI} = 56\%$ for an outlier magnitude of $\nabla_i = 8\sigma$ and $\alpha' = 0.1$. For $\alpha' = 0.1$ the $MIB = 33.5\sigma = 33.5mm$, whereas for $\alpha' = 0.001$ is $MIB = 25\sigma = 25mm$. Therefore, in that case, the MIB for $\tilde{\mathcal{P}}_{CI} = 0.8(80\%)$ and $\alpha' = 0.1$ would be 34% larger than user-defined $\alpha' = 0.001$.

The soft constraints A and D were grouped in the Cluster 5 (i.e. A and D were treated as pseudo-observations in the model). There is no reliability in terms of outlier identification for the constraints, because the residual correlation between them is $\rho_{w_i,w_j} = 100\%$, as can be seen in Table 9 for y_{13} and y_{14}. However, these soft constraints are able to detect an outlier. In that case, the probability of correct detection $\mathcal{P}_{CD}$ in Figure 6 is mainly caused by the statistical overlap $\mathcal{P}_{ol}$, as can be seen in the example of the case of $\sigma_c = 10mm$ in Figure 16. From Table 10, we observe that the larger the relaxation of the constraints, the larger the MDB. Note that the values of MDB are given in σ, and therefore the MDB for $\sigma_c = 10mm$ is larger than $\sigma_c = 1mm$ and $\sigma_c = 0.1mm$, i.e. we had the following inequality: $MDB = 6.8\sigma_c = 6.8 \times 10mm = 68mm > MDB = 8.8\sigma_c = 8.8 \times 1mm = 8.8mm > MDB = 57\sigma_c = 57 \times 0.1mm = 5.7mm$. In that case, if the FWE (α') were increased, the rate of outlier detection by the Cluster 4 (i.e. by the soft constraints) would increase.
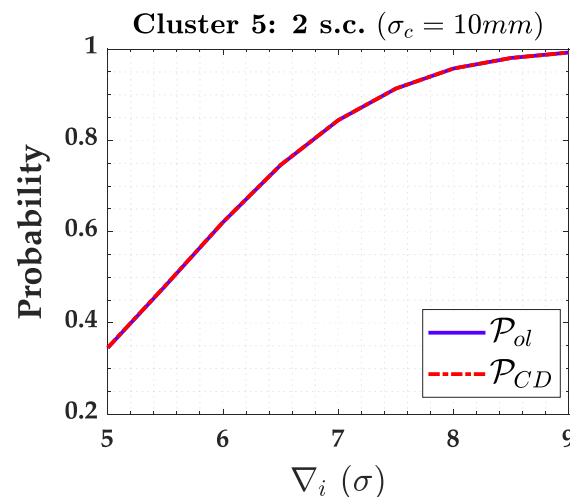
Figure 16. Probabilities of correct detection $\mathcal{P}_{CD}$ and statistical overlap $\mathcal{P}_{ol}$ for the two soft constraints A and D (Cluster 5) with $\sigma_c = 10mm$ and for $\alpha' = 0.001$.

Similar effects of the relaxation of the constraints on the performance of the *IDS* in case of two soft constraints are verified in case of three soft constraints, as can be seen in the Figures 9, 10, 11 and 12.

In case of having three soft constraints in Figure 2(e), there is reliability in terms of outlier identification for the three pseudo-observations y_{13}, y_{14} and y_{15} (i.e. for A,D and G), as can be seen in Figure 10 and Table 15. In that case, we also observe that the probabilities of correct detection $\mathcal{P}_{CD}$ of the soft constraints A and D (i.e. Cluster 5) were approximately 13% for $\sigma_c = 10mm$, 16% for $\sigma_c = 1mm$ and 24% for $\sigma_c = 0.1mm$ larger than the scenario of the network subject to two soft constraints. Table 15 reveals that the advantage of having three soft constraints instead of two constraints is that the constraints become identifiable in the presence of an outlier. The behaviour of the probabilities of correct detection ($\mathcal{P}_{CD}$), correct identification ($\mathcal{P}_{CI}$) and wrong exclusion ($\mathcal{P}_{WE}$) was similar to the case of the two soft constraints.

Furthermore, the larger the relaxation of the constraints, the smaller the residuals correlation between the measurements and the soft constraints and the larger the residuals correlation among the measurements. This effect can be verified in Tables 6, 7 and 8 for two soft constraints as well as in Tables 11, 12 and 13 for three soft constraints.

In general, we observe that the case of two soft constraints for $\sigma_c = 0.1mm$ was comparable with two hard constraints (see e.g. Table 5 and Table 10) in terms of the probability levels associated with *IDS* for the measurements (i.e. clusters 1, 2, 3 and 4). In the same way for the case of two soft constraints with $\sigma_c = 1mm$ or $\sigma_c = 10mm$, the probabilities levels were similar to the one hard constraint for that measurements, with the benefit of two soft constraints having reliability in terms of outlier identification for the Cluster 1. Finally, the three soft constraints with $\sigma_c = 1mm$ and $\sigma_c = 10mm$ were comparable to the two soft constraints for that scenario of constraints relaxation, wheres the three soft constraints for $\sigma_c = 0.1mm$ showed similar outcomes with three hard constraints for the measurements (see e.g. Tables 5 and 15). In that case, however, an advantage of the three soft constraints on the three hard constraints is the possibility of analysing the sensitivity of the constraints. It is important to emphasised that the stochastic models of the measurements and constraints were assumed well-known and defined for the analyzes performed here.

5. Material and Methods

Here, we use the procedure provided by Rofatto et al. [45] to compute the probability levels associated with *IDS*, as well as to estimate the minimal biases — Minimal Detectable Bias(MDB) and Minimal Identifiable Bias (MIB). The procedure is summarised in Figure 16.

After finding the critical value $\hat{k}$ by the process described in Appendix A, the procedure based on Monte Carlo is also applied to compute the probability levels of *IDS* when there is an outlier in the dataset. The overview of of the main elements involved with the *IDS* was detailed in Section 2. The steps displayed in Figure 17 are detailed as follows:

First, random error vectors are synthetically generated on the basis of a multivariate normal distribution because the assumed stochastic model for random errors is based on the matrix covariance of the observations. Here, we use the Mersenne Twister algorithm [84] to generate a sequence of random numbers and Box–Muller [85] to transform it into a normal distribution.

The magnitude intervals of simulated outliers are user-defined. The magnitude intervals are based on the standard deviation of the observation (σ), e.g., $|3\sigma|$ to $|6\sigma|$. Since the outlier can be positive or negative, the proposed algorithm randomly selects the signal of the outlier (for $q = 1$). Here, we use the discrete uniform distribution to select the signal of the outlier. Thus, the total error (ε) is a combination of random errors, and its corresponding outlier is as follows:

$$\varepsilon = e + c_i \nabla_i \tag{42}$$

In Equation (42), e is the random error generated from the normal distribution according to Equation (2), and the second part $c_i \nabla_i$ is the additional parameter that describes the alternative model according to Equation (42). Next, we compute the least-squares residuals vector according to Equation (3), but now we use the total error (ε) as follows:

$$\hat{e} = R\varepsilon \tag{43}$$

For *IDS*, the hypothesis of (42) for $q = 1$ (one outlier) is assumed, and the corresponding test statistic is computed according to (43). Then, the maximum test statistic value is computed according to Equation (44). Now, the decision rule is based on the critical value $\hat{k}$ computed by Monte Carlo (see the steps in Appendix A). After identifying the measurement suspected to be the most likely outlier, it is excluded from the model, and least-squares estimation (LSE) and data snooping are applied iteratively until there are no further outliers identified in the dataset. Every time that a measurement suspected to be the most likely outlier is removed from the model, we check whether the normal matrix $A^T W A$ is invertible or not. If the determinant of $A^T W A$ is 0, $det|A^T W A| = 0$, then there is a necessary and sufficient condition for a square matrix $A^T W A$ to be non-invertible. In other words, the *IDS* is ended when $det|A^T W A| = 0$.
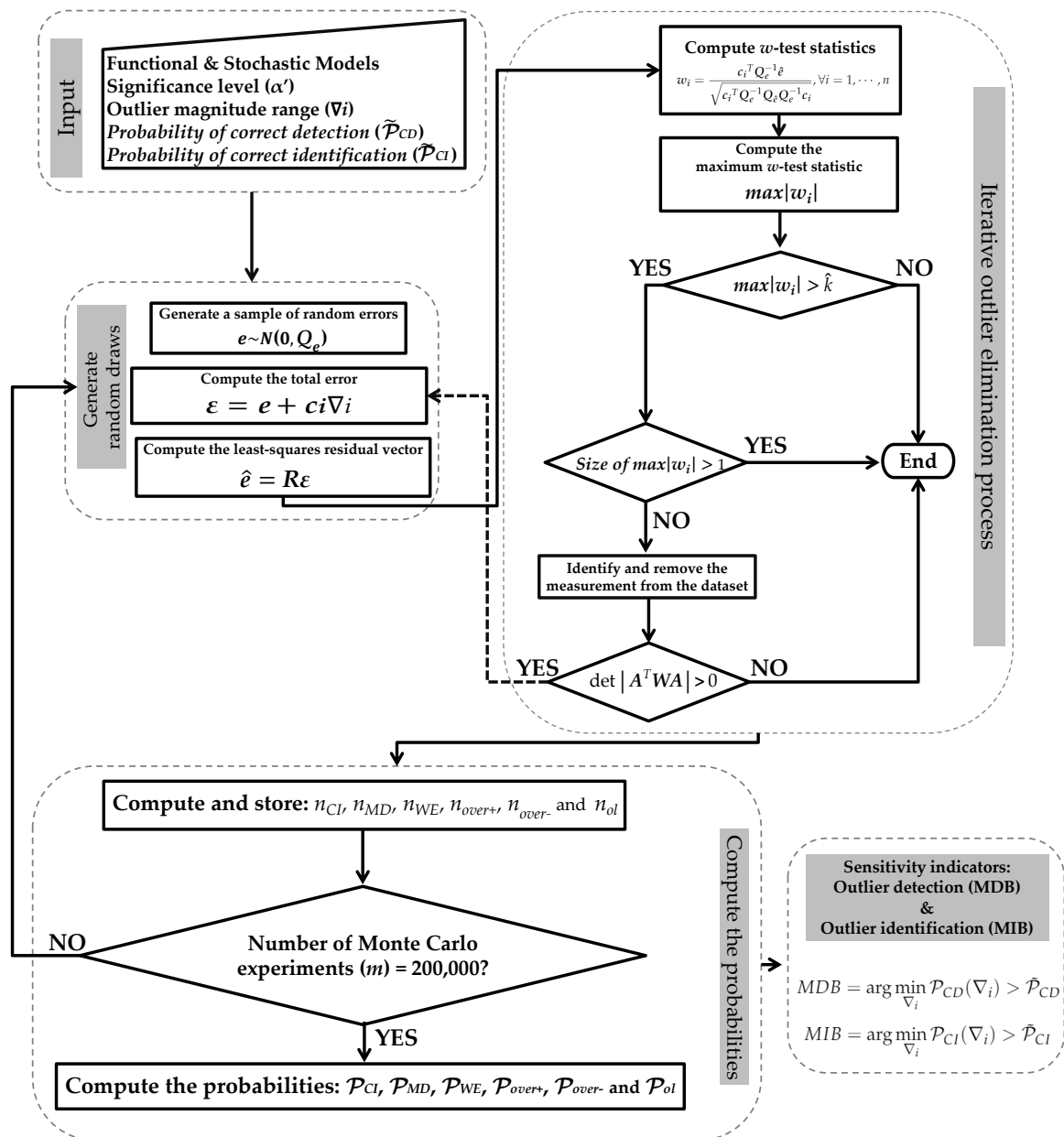
Figure 17. Flowchart of the algorithm to compute the probability levels of Iterative Data Snooping (IDS) for each measurement in the presence of an outlier [45].

The *IDS* procedure is performed for m experiments of random error vectors for each experiment contaminated by an outlier in the ith measurement. Therefore, for each measurement contaminated by an outlier, there are $v = 1, \ldots, m$ experiments.

The probability of correct identification ($\mathcal{P}_{CI}$) is obtained by the ratio between the correct identification cases and possible cases. Thus, if m is the total number of Monte Carlo experiments (possible cases), then we count the number of times that the outlier is correctly identified (denoted as n_{CI}) and then approximate the probability of correct identification $\mathcal{P}_{CI}$ as

$$\mathcal{P}_{CI} = \frac{n_{CI}}{m} \qquad (44)$$

Similar to Equation (44), false decisions are computed as

$$\mathcal{P}_{MD} = \frac{n_{MD}}{m} \tag{45}$$

$$\mathcal{P}_{WE} = \frac{n_{WE}}{m} \tag{46}$$

$$\mathcal{P}_{over_+} = \frac{n_{over_+}}{m} \tag{47}$$

$$\mathcal{P}_{over_-} = \frac{n_{over_-}}{m} \tag{48}$$

$$\mathcal{P}_{ol} = \frac{n_{ol}}{m} \tag{49}$$

where:

- n_{MD} is the number of experiments in which *IDS* does not detect the outlier ($\mathcal{P}_{MD}$ corresponds to the rate of missed detection);
- n_{WE} is the number of experiments in which the *IDS* procedure flags and removes only one single non-outlying measurement while the '*true*' outlier remains in the dataset ($\mathcal{P}_{WE}$ is the wrong exclusion rate);
- n_{over_+} is the number of experiments in which *IDS* correctly identifies and removes the outlying measurement and others, and $\mathcal{P}_{over_+}$ corresponds to its probability;
- n_{over_-} represents the number of experiments in which *IDS* identifies and removes more than one non-outlying measurement, whereas the '*true outlier*' remains in the dataset ($\mathcal{P}_{over_-}$ is the probability corresponding to this error probability class); and
- n_{ol} is the number of experiments in which the detector in Equation (44) flags two (or more) measurements simultaneously during a given iteration of *IDS*. Here, this is referred to as the number of statistical overlap n_{ol}, and $\mathcal{P}_{ol}$ corresponds to its probability.

In this paper, the probability levels associated with *IDS* were computed for each observation individually and for each outlier magnitude. However, they were grouped into clusters based on number of local redundancy (r_i) and maximum absolute correlation between the *w*-test statistics (ρ_{w_i,w_j}). Furthermore, we take care to control the *family-wise error rate*.

6. Conclusions

In general, the probability of correct identification ($\mathcal{P}_{CI}$) for the case of hard constraints is larger than soft constraints. It can be verified in Figure 3, Figure 5 and Figure 9. Therefore, hard constraints should be used in the stage of pre-processing data for the purpose of identifying and removing possible outlying measurements. In that process, one should opt to set out the redundant hard constraints at points in the network where the smallest connections exist. After identifying and removing possible outliers, the soft constraints should be employed to propagate the uncertainties of the constraints (pseudobservations) to the model parameters during the process of least-squares estimation. This recommendation is valid for outlier detection and identification purpose.

We highlight the main findings of this research as follows:

- Under a system of a high local redundancy $r_i > 0.5$ and low residuals correlation ($\rho_{w_i,w_j} < 0.5$), if one increase the *family-wise error rate* (FWE) of the test statistic, the performance of the procedure will be improved for both scenarios of hard constraints and soft constraints.
- The larger the relaxation of the constraints, the larger is the effect of the residuals correlation (ρ_{w_i,w_j}) on the success rate of outlier identification ($\mathcal{P}_{CI}$). Consequently, the higher the sensitivity indicator for outlier identification (MIB), and therefore more difficult it becomes to identify an outlier.
- Under a scenario of soft constraints, one should set out at least three soft constraints in order to identify an outlier in the constraints.

In other types of analysis, for example, deformation analysis of geodetic networks one should formulated the constraints as pseudo-observations with $\sigma_c > 0$ in an adjustment model of system equations. In other words, the relaxation of constraints makes it possible to estimate deformation effects that are unmodelled in the deterministic model matrix, for example, deformations which have a spatial and temporal variations. In that case, the deformation can be modelled by covariance functions that determine the cofactors between constraints [28]. This should be applied to control points located in the structure, but not to control points located outside the structure. The control points located outside the structure should be hard constraints and have their stability verified by an independent process in order to analyse the deformation of the structure. So, one will be also able to analyse the sensitivity in terms of minimal detectable deformations and/or minimal identifiable deformations [23]. Future researches will be addressed to sensitivity analysis for deformation monitoring networks.

Appendix A. Procedure to compute the critical value of *max*-w

The procedure to compute the critical value of max-w is given step-by-step as follows:

1. Specify the probability density function (pdf) of the w-test statistics. The pdf assigned to the w-test statistics under an $\mathcal{H}_0$-distribution is

$$(w_1, w_2, w_3, \cdots, w_n)^T \sim N\left(\mathbf{0}, \boldsymbol{\mathcal{R}_w}\right) \tag{A1}$$

   where $\boldsymbol{\mathcal{R}_w} \in \mathbb{R}^{n \times n}$ is the correlation matrix with the main diagonal elements equal to 1, and the off-diagonal elements are the correlation between the w-test statistics computed by Equation (20).

2. In order to have w-test statistics under $\mathcal{H}_0$, uniformly distributed random number sequences are produced by the Mersenne Twister algorithm, and then they are transformed into a normal distribution by using the Box–Muller transformation [85]. Box–Muller has already been used in geodesy for Monte Carlo experiments [46,86,87]. Therefore, a sequence of m random vectors from the pdf assigned to the w-test statistics is generated according to Equation (A1). In that case, we have a sequence of m vectors of the w-test statistics as follows:

$$\left[(w_1, w_2, w_3, \cdots, w_n)^{T^{(1)}}, (w_1, w_2, w_3, \cdots, w_n)^{T^{(2)}}, \cdots, (w_1, w_2, w_3, \cdots, w_n)^{T^{(m)}}\right] \tag{A2}$$

3. Compute the test statistic by Equation (44) for each sequence of w-test statistics. Thus, we have

$$\left(\max_{i \in \{1, \cdots, n\}} |\boldsymbol{w_i}|^{(1)}, \max_{i \in \{1, \cdots, n\}} |\boldsymbol{w_i}|^{(2)}, \cdots, \max_{i \in \{1, \cdots, n\}} |\boldsymbol{w_i}|^{(m)}\right) \tag{A3}$$

4. Sort in ascending order the maximum test statistic in Equation (A3), getting a sorted vector $\tilde{w}$, such that

$$\tilde{w}^{(1)} < \tilde{w}^{(2)}, \tilde{w}^{(3)}, \cdots, < \tilde{w}^{(m)} \tag{A4}$$

The sorted values $\tilde{w}$ in Equation (A4) provide a discrete representation of the cumulative density function (cdf) of the maximum test statistic max-w.

5. Determine the critical value $\hat{k}$ as follows:

$$\hat{k} = \tilde{w}_{[(1-\alpha') \times m]} \tag{A5}$$

where [.] denotes rounding down to the next integer that indicates the position of the selected elements in the ascending order of $\tilde{w}$. This position corresponds to a critical value for a stipulated overall false alarm probability α'. This can be done for a sequence of values α' in parallel.

It is important to mention that the probability of a Type I decision error for multiple testing α' is larger than that of Type I for single testing α_0. This is because the critical region in multiple testing is larger than that in single hypothesis testing.

References

1. Lehmann, R.; Neitzel, F. Testing the compatibility of constraints for parameters of a geodetic adjustment model. *Journal of Geodesy* **2013**, *87*, 555–566. doi:10.1007/s00190-013-0627-2.

2. Fang, X. Weighted total least-squares with constraints: a universal formula for geodetic symmetrical transformations. *Journal of Geodesy* **2015**, *89*, 459–469. doi:10.1007/s00190-015-0790-8.

3. Amiri-Simkooei, A.R. Weighted Total Least Squares with Singular Covariance Matrices Subject to Weighted and Hard Constraints. *J. Surv. Eng.* **2017**, *143*, 04017018,

4. Sakic, P.; Ballu, V.; Royer, J.Y. A Multi-Observation Least-Squares Inversion for GNSS-Acoustic Seafloor Positioning. *Remote Sensing* **2020**, *12*. doi:10.3390/rs12030448.

5. Odijk, D. Stochastic modelling of the ionosphere for fast GPS ambiguity resolution. Geodesy Beyond 2000; Schwarz, K.P., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2000; pp. 387–392.

6. de Oliveira, P.S.; Morel, L.; Fund, F.; Legros, R.; Monico, J.F.G.; Durand, S.; Durand, F. Modeling tropospheric wet delays with dense and sparse network configurations for PPP-RTK. *GPS Solutions* **2017**, *21*, 237–250. doi:10.1007/s10291-016-0518-0.

7. Chen, B.; Wu, L.; Dai, W.; Luo, X.; Xu, Y. A new parameterized approach for ionospheric tomography. *GPS Solutions* **2019**, *23*, 96. doi:10.1007/s10291-019-0893-4.

8. Ahmed, E.; Ahmed, S.; Mohamed, D.; Mostafa, R., jag; 2019; Vol. 13, chapter Vertical ionospheric delay estimation for single-receiver operation, p. 81. 2, doi:10.1515/jag-2018-0041.

9. Wang, A.; Chen, J.; Zhang, Y.; Meng, L.; Wang, J. Performance of Selected Ionospheric Models in Multi-Global Navigation Satellite System Single-Frequency Positioning over China. *Remote Sensing* **2019**, *11*. doi:10.3390/rs11172070.

10. Deo, M.; El-Mowafy, A. Comparison of advanced troposphere models for aiding reduction of PPP convergence time in Australia. *Journal of Spatial Science* **2019**, *64*, 381–403. doi:10.1080/14498596.2018.1472046.

11. Dai, X.; Lou, Y.; Dai, Z.; Hu, C.; Peng, Y.; Qiao, J.; Shi, C. Precise Orbit Determination for GNSS Maneuvering Satellite with the Constraint of a Predicted Clock. *Remote Sensing* **2019**, *11*. doi:10.3390/rs11161949.

12. Courant, R.; Hilbert, D. *Methods of mathematical physics. Vol. 2, Vol. 2*,; Wiley: [New York], 1962.

13. Grafarend, E.W. Optimization of Geodetic Networks. *The Canadian Surveyor* **1974**, *28*, 716–723. doi:10.1139/tcs-1974-0120.

14. Teunissen, P. Zero Order Design: Generalized Inverses, Adjustment, the Datum Problem and S-Transformations. Optimization and Design of Geodetic Networks; Grafarend, E.W.; Sansò, F., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1985; pp. 11–55.

15. Dermanis, A. Free network solutions with the DLT method. *ISPRS Journal of Photogrammetry and Remote Sensing* **1994**, *49*, 2 – 12. doi:https://doi.org/10.1016/0924-2716(94)90061-2.

16. Kotsakis, C. Reference frame stability and nonlinear distortion in minimum-constrained network adjustment. *Journal of Geodesy* **2012**, *86*, 755–774. doi:10.1007/s00190-012-0555-6.

17. Kotsakis, C., Datum Definition and Minimal Constraints. In *Encyclopedia of Geodesy*; Grafarend, E., Ed.; Springer International Publishing: Cham, 2018; pp. 1–6. doi:10.1007/978-3-319-02370-0_157-1.

18.   Matsuoka, M.T.; Rofatto, V.F.; Klein, I.; Roberto Veronez, M.; da Silveira, L.G.; Neto, J.B.S.; Alves, A.C.R. Control Points Selection Based on Maximum External Reliability for Designing Geodetic Networks. *Applied Sciences* **2020**, *10*. doi:10.3390/app10020687.

19.   Baarda, W. S-transformations and criterion matrices. *Publ. on geodesy, New Series* **1973**, *5*.

20.   Schaffrin, B. Aspects of Network Design. Optimization and Design of Geodetic Networks; Grafarend, E.W.; Sansò, F., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1985; pp. 548–597.

21.   Xu, P.B. A general solution in geodetic nonlinear rank-defect models. 1997.

22.   Altamimi, Z.; Dermanis, A. The Choice of Reference System in ITRF Formulation. VII Hotine-Marussi Symposium on Mathematical Geodesy; Sneeuw, N.; Novák, P.; Crespi, M.; Sansò, F., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp. 329–334.

23.   Velsink, H. On the deformation analysis of point fields. *J. Geod.* **2015**, *89*, 1071–1087. doi:10.1007/s00190-015-0835-z.

24.   Velsink, H. Extendable linearised adjustment model for deformation analysis. *Survey Review* **2015**, *47*, 397–410. doi:10.1179/1752270614Y.0000000140.

25.   Hiddo, V., jag; 2016; Vol. 10, chapter Time Series Analysis of 3D Coordinates Using Nonstochastic Observations, p. 5. 1, doi:10.1515/jag-2015-0027.

26.   Rao, C.R. Markoff's Theorem with Linear Restrictions on Parameters. *SankhyÄ: The Indian Journal of Statistics (1933-1960)* **1945**, *7*, 16–19.

27.   Velsink, H. Testing Methods for Adjustment Models with Constraints. *Journal of Surveying Engineering* **2018**, *144*, 04018009. doi:10.1061/(ASCE)SU.1943-5428.0000260.

28.   Hiddo, V. Testing deformation hypotheses by constraints on a time series of geodetic observations. *J. Appl. Geod.* **2017**, *12*, 77–93. doi:10.1515/jag-2017-0028.

29.   Baarda, W. A testing procedure for use in geodetic networks. *Publ. on geodesy, New Series* **1968**, *2*.

30.   Teunissen, P. First and second moments of non-linear least-squares estimators. *Bull. Geodesique (Journal of Geodesy)* **1989**, *63*, 253–262. doi:https://doi.org/10.1007/BF02520475.

31.   Teunissen, P. *Testing Theory: an introduction*, 2 ed.; Delft University Press, 2006.

32.   Amiri-Simkooei, A. Least-squares variance component estimation: theory and GPS applications. Phd thesis, Delft University of Technology, 2007.

33.   Koch, K.R. *Parameter estimation and hypothesis testing in linear models.*, 2 ed.; Springer, 1999.

34.   Tao, H.; Feng, H.; Xu, L.; Miao, M.; Long, H.; Yue, J.; Li, Z.; Yang, G.; Yang, X.; Fan, L. Estimation of Crop Growth Parameters Using UAV-Based Hyperspectral Remote Sensing Data. *Sensors* **2020**, *20*. doi:10.3390/s20051296.

35.   Han, M.; Wang, Q.; Wen, Y.; He, M.; He, X. The Application of Robust Least Squares Method in Frequency Lock Loop Fusion for Global Navigation Satellite System Receivers. *Sensors* **2020**, *20*. doi:10.3390/s20041224.

36.   Wei, C. Estimation for the Discretely Observed Coxâ"Ingersollâ"Ross Model Driven by Small Symmetrical Stable Noises. *Symmetry* **2020**, *12*. doi:10.3390/sym12030327.

37.   Sakic, P.; Ballu, V.; Royer, J.Y. A Multi-Observation Least-Squares Inversion for GNSS-Acoustic Seafloor Positioning. *Remote Sensing* **2020**, *12*. doi:10.3390/rs12030448.

38.   Livadiotis, G. General Fitting Methods Based on Lq Norms and their Optimization. *Stats* **2020**, *3*, 16–31. doi:10.3390/stats3010002.

39.   Farooq, S.Z.; Yang, D.; Ada, E.N.J. A Cycle Slip Detection Framework for Reliable Single Frequency RTK Positioning. *Sensors* **2020**, *20*. doi:10.3390/s20010304.

40.   Araveeporn, A. Comparing Parameter Estimation of Random Coefficient Autoregressive Model by Frequentist Method. *Mathematics* **2020**, *8*. doi:10.3390/math8010062.

41.   Zhang, C.; Peng, T.; Zhou, J.; Ji, J.; Wang, X. An Improved Autoencoder and Partial Least Squares Regression-Based Extreme Learning Machine Model for Pump Turbine Characteristics. *Applied Sciences* **2019**, *9*. doi:10.3390/app9193987.

42.   Ji, J.; Yang, M.; Jiang, L.; He, J.; Teng, Z.; Liu, Y.; Song, H. Output-Only Parameters Identification of Earthquake-Excited Building Structures with Least Squares and Input Modification Process. *Applied Sciences* **2019**, *9*. doi:10.3390/app9040696.

43.   Kargoll, B. On the theory and application of model misspecification tests in geodesy. Doctoral thesis, University of Bonn, Landwirtschaftliche Fakultät, German, Bonn, 2007.

44.   Lehmann, R. On the formulation of the alternative hypothesis for geodetic outlier detection. *J. Geod.* **2013**, *87*, 373–386. doi:10.1007/s00190-012-0607-y.

45.   Rofatto, V.F.; Matsuoka, M.T.; Klein, I.; Roberto Veronez, M.; da Silveira, L.G. A Monte Carlo-Based Outlier Diagnosis Method for Sensitivity Analysis. *Remote Sensing* **2020**, *12*. doi:10.3390/rs12050860.

46.   Lehmann, R. Improved critical values for extreme normalized and studentized residuals in Gauss–Markov models. *J. Geod.* **2012**, *86*, 1137–1146. doi:10.1007/s00190-012-0569-0.

47.   Rofatto, V.F.; Matsuoka, M.T.; Klein, I.; Veronez, M.R.; Bonimani, M.L.; Lehmann, R. A half-century of Baarda's concept of reliability: a review, new perspectives, and applications. *Surv. Rev.* **2018**, *0*, 1–17, [https://doi.org/10.1080/00396265.2018.1548118]. doi:10.1080/00396265.2018.1548118.

48.   Zaminpardaz, S.; Teunissen, P. DIA-datasnooping and identifiability. *J. Geod.* **2019**, *93*, 85–101. doi:10.1007/s00190-018-1141-3.

49.   Kok, J.J.; States., U. *On data snooping and multiple outlier testing [microform] / Johan J. Kok*; U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Charting and Geodetic Services : For sale by the National Geodetic Information Center, NOAA Rockville, Md, 1984; p. 61 p. :.

50.   Knight, N.L.; Wang, J.; Rizos, C. Generalised measures of reliability for multiple outliers. *Journal of Geodesy* **2010**, *84*, 625–635. doi:10.1007/s00190-010-0392-4.

51.   Gui, Q.; Li, X.; Gong, Y.; Li, B.; Li, G. A Bayesian unmasking method for locating multiple gross errors based on posterior probabilities of classification variables. *J. Geod.* **2011**, *85*, 191–203. doi:10.1007/s00190-010-0429-8.

52.   Klein, I.; Matsuoka, M.T.; Guzatto, M.P.; Nievinski, F.G. An approach to identify multiple outliers based on sequential likelihood ratio tests. *Surv. Rev.* **2017**, *49*, 449–457, [https://doi.org/10.1080/00396265.2016.1212970]. doi:10.1080/00396265.2016.1212970.

53.   Hawkins, D.M. *Identification of Outliers*, 1 ed.; Springer Netherlands, 1980. https://doi.org/10.1007/978-94-015-3994-4.

54.   Baarda, W. Statistical concepts in geodesy. *Publ. on geodesy, New Series* **1967**, *2*.

55.   Imparato, D.; Teunissen, P.; Tiberius, C. Minimal Detectable and Identifiable Biases for quality control. *Surv. Rev.* **2019**, *51*, 289–299, [https://doi.org/10.1080/00396265.2018.1437947]. doi:10.1080/00396265.2018.1437947.

56.   Teunissen, P.J.G. Distributional theory for the DIA method. *Journal of Geodesy* **2018**, *92*, 59–80. doi:10.1007/s00190-017-1045-7.

57.   Förstner, W. Reliability and discernability of extended Gauss-Markov models. Seminar on mathematical models of Geodetic/Photogrammetric Point Determination with Regard to Outliers and Systematic Errors; , 1983; Vol. Series A, pp. 79–103.

58.   Wang, J.; Knight, N.L. New Outlier Separability Test and Its Application in GNSS Positioning. 2012.

59.   Lehmann, R.; Lösler, M. Multiple Outlier Detection: Hypothesis Tests versus Model Selection by Information Criteria. *J. Surv. Eng.* **2016**, *142*, 04016017,

60.   Teunissen, P.J.G.; Imparato, D.; Tiberius, C.C.J.M. Does RAIM with Correct Exclusion Produce Unbiased Positions? *Sensors* **2017**, *17*. doi:10.3390/s17071508.

61.   Lehmann, R.; Lösler, M. Congruence analysis of geodetic networks – hypothesis tests versus model selection by information criteria. *J. Appl. Geod.* **2017**, *11*, 271–283. doi:10.1515/jag-2016-0049.

62.   Rofatto, V.; Matsuoka, M.; Klein, I. DESIGN OF GEODETIC NETWORKS BASED ON OUTLIER IDENTIFICATION CRITERIA: AN EXAMPLE APPLIED TO THE LEVELING NETWORK. *Bull. Geod. Sci.* **2018**, *24*, 152–170. http://dx.doi.org/10.1590/s1982-21702018000200011.

63.   Nguyen, V.K.; Renault, ; Milocco, R. Environment Monitoring for Anomaly Detection System Using Smartphones. *Sensors* **2019**, *19*. doi:10.3390/s19183834.

64.   Nie, Y.; Yang, L.; Shen, Y. Specific Direction-Based Outlier Detection Approach for GNSS Vector Networks. *Sensors* **2019**, *19*. doi:10.3390/s19081836.

65.   Klein, I.; Matsuoka, M.T.; Guzatto, M.P.; Nievinski, F.G.; Veronez, M.R.; Rofatto, V.F. A new relationship between the quality criteria for geodetic networks. *Journal of Geodesy* **2019**, *93*, 529–544. doi:10.1007/s00190-018-1181-8.

66.   Mierlo, J.V. Statistical Analysis of Geodetic Measurements for the Investigation of Crustal Movements. In *Recent Crustal Movements, 1977*; Whitten, C.; Green, R.; Meade, B., Eds.; Elsevier, 1979; Vol. 13, *Developments in Geotectonics*, pp. 457 – 467. doi:https://doi.org/10.1016/B978-0-444-41783-1.50072-6.

67.   van der Marel, H.; Rösters, A.J.M. Statistical Testing and Quality Analysis in 3-D Networks (part II) Application to GPS. Global Positioning System: An Overview; Bock, Y.; Leppard, N., Eds.; Springer New York: New York, NY, 1990; pp. 290–297.

68.   Yang, L.; Wang, J.; Knight, N.L.; Shen, Y. Outlier separability analysis with a multiple alternative hypotheses test. *J. Geod.* **2013**, *87*, 591–604. doi:10.1007/s00190-013-0629-0.

69.   Prószyński, W. Revisiting Baarda's concept of minimal detectable bias with regard to outlier identifiability. *J. Geod.* **2015**, *89*, 993–1003. doi:10.1007/s00190-015-0828-y.

70.   Romano, J.P.; Wolf, M. Multiple Testing of One-Sided Hypotheses: Combining Bonferroni and the Bootstrap. Predictive Econometrics and Big Data; Kreinovich, V.; Sriboonchitta, S.; Chakpitak, N., Eds.; Springer International Publishing: Cham, 2018; pp. 78–94.

71.   Teunissen, P. An Integrity and Quality Control Procedure for use in Multi Sensor Integration. 1990.

72.   Aydin, C.; Demirel, H. Computation of Baarda's lower bound of the non-centrality parameter. *J. Geod.* **2004**, *78*, 437–441. doi:10.1007/s00190-004-0406-1.

73.   Aydin, C. Power of Global Test in Deformation Analysis. *J. Surv. Eng.* **2012**, *138*, 51–56. https://doi.org/10.1061/(ASCE)SU.1943-5428.0000064.

74.   Bonferroni, C. *Teoria statistica delle classi e calcolo delle probabilità*; Pubblicazioni del R. Istituto superiore di scienze economiche e commerciali di Firenze, Libreria internazionale Seeber, 1936.

75.   Ihde, J.; Sánchez, L.; Barzaghi, R.; Drewes, H.; Foerste, C.; Gruber, T.; Liebsch, G.; Marti, U.; Pail, R.; Sideris, M. Definition and Proposed Realization of the International Height Reference System (IHRS). *Surveys in Geophysics* **2017**, *38*, 549–570. doi:10.1007/s10712-017-9409-3.

76.   Lehmann, R. Observation error model selection by information criteria vs. normality testing. *Studia Geophysica et Geodaetica* **2015**, *59*, 489–504. doi:10.1007/s11200-015-0725-0.

77.   Algarni, D.A.; Ali, A.E. Heighting and Distance Accuracy with Electronic Digital Levels. *Journal of King Saud University - Engineering Sciences* **1998**, *10*, 229 – 239. doi:https://doi.org/10.1016/S1018-3639(18)30698-6.

78.   Gemin, A.R.S.; Matos, .A.S.; Faggion, P.L.A.s. APPLICATION OF CALIBRATION CERTIFICATE OF DIGITAL LEVELING SYSTEMS IN THE MONITORING OF STRUCTURES: A CASE STUDY AT THE GOVERNADOR JOSÃ‰■ RICHA HYDROELECTRIC POWER PLANT - PR. *Boletim de CiÃGeodÃ* **2018**, *24*, 235 – 249.

79.   Takalo, M.; Rouhiainen, P. Development of a System Calibration Comparator for Digital Levels in Finland. *Nordic Journal of Surveying and Real Estate Research* **1**, *1*.

80.   Wiedemann, W.; Wagner, A.; Wunderlich, T. Using IATS to Read and Analyze Digital Levelling Staffs. SIG 2016; Paar, R.; Marendić, A.; Zrinjski, M., Eds.; Croatian Geodetic Society, , 2016; pp. 515–526.

81.   P.J.G., T. The Geometry of Geodetic Inverse Linear Mapping and Non-linear Adjustment. *Publications on Geodesy, New Series* **1985**, *8*.

82.   Rao, C.R.; Mitra, S.K. Generalized inverse of a matrix and its applications. 1972.

83.   Hekimoglu, S.; Erenoglu, R.C.; Sanli, D.U.; Erdogan, B. Detecting Configuration Weaknesses in Geodetic Networks. *Survey Review* **2011**, *43*, 713–730. doi:10.1179/003962611X13117748892632.

84.   Matsumoto, M.; Nishimura, T. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation* **1998**, *8*, 3–30. https://dl.acm.org/citation.cfm?id=272991.

85.   Box, G.E.P.; Muller, M.E. A Note on the Generation of Random Normal Deviates. *The Annals of Mathematical Statistics* **1958**, *29*, 610–611. https://doi.org/10.1214/aoms/1177706645.

86.   Lemeshko, B.Y.; Lemeshko, S.B. Extending the Application of Grubbs-Type Tests in Rejecting Anomalous Measurements. *Measurement Techniques* **2005**, *48*, 536–547. doi:10.1007/s11018-005-0179-9.

87.   Lehmann, R.; Scheffler, T. Monte Carlo based data snooping with application to a geodetic network. *J. Appl. Geod.* **2011**, *5*, 123–134. https://doi.org/10.1515/JAG.2011.014.