Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

Article

Control Points Selection based on Maximum External Reliability for Designing Geodetic Networks

Marcelo Tomio Matsuoka [1,2,3,6], Vinicius Francisco Rofatto [1,2], Ivandro Klein [4,5], Maurício Roberto Veronez [6], Luiz Gonzaga da Silveira Jr. [6], João Batista Silva Neto [1] and Ana Cristina Ramos Alves [1]

1 Geography Institute/Quality Control Lab, Federal University of Uberlandia, 38500-000, Monte Carmelo, MG, Brazil; tomio@ufu.br; vfrofatto@gmail.com
2 Graduate Program in Remote Sensing, Federal University of Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil
3 Graduate Program in Agriculture and Geospatial Information, Federal University of Uberlândia, 38500-000, Monte Carmelo, MG, Brazil
4 Land Surveying Program, Federal Institute of Santa Catarina, 88020-300, Florianópolis, SC, Brazil; ivandroklein@gmail.com
5 Graduate Program in Geodetic Sciences, Federal University of Parana, 81531-990, Curitiba, Brazil
6 Graduate Program in Applied Computing/Advanced Visualization Laboratory (Vizlab), 93022-750, Unisinos University, São Leopoldo, RS, Brazil; veronez@unisinos.br
* tomio@ufu.br

Abstract: Geodetic networks are essential for most geodetic, geodynamics and civil projects, such as monitoring the position and deformation of man-made structures, monitoring the crustal deformation of the Earth, establishing and maintaining a geospatial reference frame, mapping, civil engineering projects and so on. Before the installation of geodetic marks and gathering of survey data, geodetic networks need to be designed according to the pre-established quality criteria. In this study, we present a method for designing geodetic networks based on the concept of reliability. We highlight that the method discards the use of the observation vector of Gauss-Markov model. In fact, the only needs are the geometrical network configuration and the uncertainties of the observations. The aim of the proposed method is to find the optimum configuration of the geodetic control points so that the maximum influence of an outlier on the coordinates of the network is minimum. Here, the concept of Minimal Detectable Bias defines the size of the outlier and its propagation on the parameters is used to describe the external reliability. The proposed method is demonstrated by practical application of one simulated levelling network. We highlight that the method can be applied not only for geodetic network problems, but also in any branch of modern science.

Keywords:geodetic network; outlier; reliability; reference points; surveying; quality control

1. Introduction

A geodetic network consists of a set of points located on the Earth's surface or near it. Their spatial positions are related to a reference system. The coordinates are derived from angles, distances and height differences between points and/or space-based geodetic techniques, such as Global Navigation Satellite Systems (GNSS). The geodetic networks provide the basic positional reference structure for mapping and civil engineering projects, such as deformations analysis of physical and man-made objects on the Earth surface, implementation of an urban and rural land register, establishment and maintenance of geospatial reference frame, etc.

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

Before the data acquisition, the geodetic network needs to be designed. The design problem have been widely developed and investigated since the pioneering work of [1]. Thenceforth, a series of papers have been published on the development of the new algorithms by simulated examples and real applications (see e.g. [2–13]). Although the design of geodetic networks is a widely investigated topic, there are open issues requiring further research.

A design of geodetic network depends on the pre-established quality criteria, such as precision, reliability and costs. The precision is related to the covariance matrix of the coordinates of the network points. The ability of the measurements scheme to detect outliers in the observations as well as to describe their effects on estimated parameters are associated with the measures of reliability. Finally, the cost is related to the effort required to implement the design and related expenses. In this study, we purpose a method based on the theory of reliability to design a geodetic network.

The reliability theory becomes a fundamental part of measurement analysis [13–17]. This is due to [18,19]. He proposed a procedure based on hypothesis testing for the detection of a single error in linear(ised) models, which he called data snooping. Most of conventional geodetic studies have a chapter on Baarda's data snooping procedure, e.g. [20,21]. Furthermore, this procedure has become very popular and is routinely used in adjustment computations [22]. Although data snooping was introduced as a testing procedure for use in geodetic networks, it is a generally applicable method [23].

Baarda's data snooping consists of screening each individual observation for a possible error. Baarda's w-test statistic for his data snooping is given by a normalised least-squares residual. This test, which is based on a linear mean-shift model, can also be derived as a particular case of the generalised likelihood ratio test [24]. Baarda's w-test makes a decision between the null $\mathcal{H}_0$ and a single alternative hypothesis $\mathcal{H}_A$ In that case, rejection of $\mathcal{H}_0$ automatically implies acceptance of $\mathcal{H}_A$, and vice-versa [15,16,25].

Based on the probability of rejecting a true null hypothesis $\mathcal{H}_0$ (type I error - 'false alarm', denoted by "α") and the probability of rejecting a true alternative hypothesis $\mathcal{H}_A$ (type II error - 'missed detection', denoted by "β"), Baarda [18,19] also derived the concept of the Minimal Detectable Bias (MDB) – the term given by [26]. The MDB is the additional bias (or are the additional biases) in the parameters vector that can be detected by the w-test with a certain probability of $1-\beta$. The MDB can be computed before actual measurements have been carried out, using only a functional model and the expected stochastic properties of the data [21]. In addition, it is possible to describe the influences of the MDBs on the geodetic network coordinates (i.e. on the parameters). The set of MDBs describes the internal reliability, whereas their propagation on the parameters is said to describe the external reliability [27]. The measures of internal and external reliability, therefore, are very useful tool to assess the magnitude of possible errors that can be detected during the pre-processing of the data. For this reason the concept of the internal (quantified by MDB) and external reliability can be applied during the design stage of geodetic network. In this context, the aim of this study is to apply the concept of reliability for designing geodetic network. The quality criterion considered here is based on the external reliability. The position of the control point of geodetic network is selected so that the maximum influence of an MDB on the coordinates of the network points is minimum. A simulated levelling network is used as an example of application of the proposed method. Here, we consider a scenario where the observations of the network have the same uncertainties and another with different uncertainties. We also consider the case of minimally constrained and over-constrained network.

2. Conventional Reliability Theory

The null hypothesis (denoted by $\mathcal{H}_0$), which is also called the working hypothesis, corresponds to a supposedly full specification model. This model is used to estimate the unknown parameters, typically in a least-squares approach. Thus, the null hypothesis of the standard Gauss–Markov model in linear or linearised form is given by [20]:

$$\mathcal{H}_0 : \mathbb{E}\{\boldsymbol{y}\} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e}, \tag{1}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

where $\mathbb{E}\{.\}$ is the expectation operator, $\boldsymbol{y} \in \mathbb{R}^n$ the vector of measurements, $\boldsymbol{A} \in \mathbb{R}^{n \times u}$ the Jacobian matrix (also called design matrix) of full rank u, $\boldsymbol{x} \in \mathbb{R}^u$ the unknown parameter vector, and $\boldsymbol{e} \in \mathbb{R}^n$ the unknown vector of measurement errors.

Typically, it is assumed that the errors of the good measurements are normally distributed with expectation zero, i.e.:

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

with a known positive definite symmetric covariance matrix $\boldsymbol{Q_e} \in \mathbb{R}^{n \times n}$. Here, we confine ourselves to the case that $\boldsymbol{A}$ and $\boldsymbol{Q_e}$ have full column rank.

The redundancy (or degrees of freedom) of the model in Equation 1 is $r = n - u$. However, any model is only an approximation to the truth. This implies that we inevitably encounter misspecified models. In contrast to the $\mathcal{H}_0$, Baarda [19] introduced a mean shift model that defines the alternative hypothesis $\mathcal{H}_A$, also referred to as model misspecification, as follows:

$$\mathcal{H}_A : \mathbb{E}\{\boldsymbol{y}\} = \boldsymbol{Ax} + \boldsymbol{c_i}\nabla_i + \boldsymbol{e} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{c_i} \end{pmatrix} + \begin{pmatrix} \boldsymbol{x} \\ \nabla_i \end{pmatrix} + \boldsymbol{e}, \forall i = 1, \cdots, n \tag{3}$$

with $\boldsymbol{c_i}$ a canonical unit vector, which consists exclusively of elements with values of 0 and 1, where 1 means that an ith bias parameter of magnitude ∇_i affects an ith measurement and 0 otherwise. We have, for instance, $\boldsymbol{c_i}$=$\begin{pmatrix} 0 & 0 & 0 & \cdots & 1^{ith} & 0 & \cdots & 0 \end{pmatrix}^T$. In other words, $\boldsymbol{c_i}$ specifies the type of model error and ∇_i the size of the model error, or outlier.

The likelihood ratio test to test $\mathcal{H}_0$ against $\mathcal{H}_A$ is given by:

$$Accept\ \ \mathcal{H}_0\ \ if\ \ |\boldsymbol{w_i}| \leq k = \Phi^{-1}(1 - \frac{\alpha}{2}),\ reject\ otherwise\ in\ favour\ of\ \mathcal{H}_A \tag{4}$$

and the test statistic (known as w-test) is given by a normalised least-squares residual as follows [19]:

$$\boldsymbol{w_i} = \frac{\boldsymbol{c_i}^T \boldsymbol{Q_e}^{-1} \hat{\boldsymbol{e}}}{\sqrt{\boldsymbol{c_i}^T \boldsymbol{Q_e}^{-1} \boldsymbol{Q_{\hat{e}}} \boldsymbol{Q_e}^{-1} \boldsymbol{c_i}}} \tag{5}$$

According to 4 and 5, we have:

- k is the critical value. The critical value k is the the tabular value from the cumulative distribution function (cdf) of the standard normal $\boldsymbol{N}(0, 1)$ based on the chosen of a significance level α. Because we perform a two-sided test of the form $|\boldsymbol{w_i}| \leq k$ we have $\alpha/2$. For example, for $\alpha = 0.01$, we obtain $k = 2.576$. In this case, if $|\boldsymbol{w_i}| > 2.576$ for some y_i one may reject $\mathcal{H}_0$.
- Φ^{-1} denotes the inverse of the normal cumulative distribution function.
- $\hat{\boldsymbol{e}}$ is the least-squares residuals vector under $\mathcal{H}_0$ and $\boldsymbol{Q_{\hat{e}}}$ the covariance matrix of the best linear unbiased estimator of $\hat{\boldsymbol{e}}$ under $\mathcal{H}_0$.

The decision rule in 4 says that if the test statistic $|w_i|$ in 5 is larger than some critical value k, i.e. a percentile of its probability distribution, then we reject the null hypothesis $\mathcal{H}_0$ in favour of the alternative hypothesis $\mathcal{H}_A$.

Because w-test in its essence is based on binary hypothesis testing, in which one decides between the null hypothesis $\mathcal{H}_0$ and a unique alternative hypothesis $\mathcal{H}_A$, it may lead to type I decision error (α) and type II decision error (β). The probability level α is the probability of rejecting the null hypothesis when it is true, whereas β is the probability of failing to reject the null hypothesis when it is false. Instead of α and β, there is the confidence level (CL) and power of the test (γ). The first deals with the probability of accepting a true null hypothesis; the second, with the probability of correctly accepting the alternative hypothesis. The power of the test is a complement of type II decision error β, i.e. $= 1 - \beta$. Similarly, the confidence level (CL) is given by $CL = 1 - \alpha$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

The normalised least-squares residual w_i follows a standard normal distribution with the expectation that $\mu = 0$ if $\mathcal{H}_0$ holds true. On the other hand, if the system is contaminated with a single error at the ith position of the dataset, there is an outlier that causes the expectation of w_i to become $\mu > 0$. The effect can be best understood using the non-central chi-squared distribution with one degree of freedom (i.e. for one single outlier). Under the alternative hypothesis $\mathcal{H}_A$, the expectation of w_i is the square-root of the non-centrality parameter $\lambda_{q=1}$ from the chi-square distribution with one degree of freedom (q=1), which is given by:

$$\mathbb{E}\{w_i\} = \lambda_{q=1} = \mathbf{c_i}^T \mathbf{Q}_e^{-1} \mathbf{Q}_{\hat{e}} \mathbf{Q}_e^{-1} \mathbf{c_i} \nabla_i^2 \tag{6}$$

where $\lambda_{q=1}$ is the non-centrality parameter for one degree of freedom $q = 1$. Note that there is an outlier that causes the expectation of w_i to become $\lambda_{q=1}$.

The non-centrality parameter $\lambda_{q=1}$ in Equation 6 represents the expected mean shift of a specific w-test. In such case, the term $\mathbf{c_i}^T \mathbf{Q}_e^{-1} \mathbf{Q}_{\hat{e}} \mathbf{Q}_e^{-1} \mathbf{c_i}$ in Equation 6 is a scalar and therefore it can be rewritten as follows [26]:

$$|\nabla_i| = MDB_{(i)} = \sqrt{\frac{\lambda_{q=1}}{\mathbf{c_i}^T \mathbf{Q}_e^{-1} \mathbf{Q}_{\hat{e}} \mathbf{Q}_e^{-1} \mathbf{c_i}}}, \ \forall i = 1, \cdots, n \tag{7}$$

where $|\nabla_i|$ is the minimal detectable bias $MDB_{(i)}$, which can be computed for each of the n alternative hypotheses according to Equation 3.

For a single outlier, the variance of estimated outlier, denoted by $\sigma^2_{\nabla_i}$, is:

$$\sigma^2_{\nabla_i} = \left(\mathbf{c_i}^T \mathbf{Q}_e^{-1} \mathbf{Q}_{\hat{e}} \mathbf{Q}_e^{-1} \mathbf{c_i}\right)^{-1}, \ \forall i = 1, \cdots, n \tag{8}$$

Thus, the MDB can also be written as:

$$MDB_{(i)} = \sigma_{\nabla_i} \sqrt{\lambda_{q=1}}, \ \forall i = 1, \cdots, n \tag{9}$$

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

The MDB in Equation 7 or 9 of an alternative hypothesis is the smallest magnitude outlier that can lead to rejection of a null hypothesis for a given α and β. Thus, for each model of the alternative hypothesis $\mathcal{H}_A$, the corresponding MDB can be computed. The key point of MDB is that it can work as a tool for designing systems capable of withstanding outlier with a certain degree of probability.

The non-centrality parameter $\lambda_{q=1}$ can be computed as a function of type 1 decision error α, type 2 decision error β and the degrees of freedom of the test q. Here, we use the recursive algorithm based on the work by [28], namely bisection algorithm, in order to obtain the non-centrality parameter $\lambda_{q=1}$. With the non-centrality parameter, and knowing the uncertainty of the sensor and the architecture of the model, it is possible to compute the MDB according to Equation 7 or 9. The MDB was further investigated for a single outlier in a singular Gauss–Markov model [29]. There are also studies covering either independent or correlated measurements [30–34]. For the case of one assumes that more than one outlier is present in the dataset. In other words, it is possible to set up for the case of multiple outliers. The readers who are interested in multiple outliers issue can refer to [35–38]. For more details about alternative models, refer to [39,40].

In order to quantify the external reliability, one should propagate each MDB on the parameters. In other words, the external reliability measures the influence of an undetected outlier on the estimation of coordinates of the geodetic network, and it is given by:

$$\nabla \mathbf{X} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{W} \mathbf{c_i} |\nabla_i|) \tag{10}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at *Appl. Sci.* **2020**, *10*, 687; doi:10.3390/app10020687

where $\nabla X \in \mathbb{R}^u$ is the influence of an undetectable outlier ∇_i located at a given position according to the vector $\boldsymbol{c_i}$ in 3 and $\boldsymbol{W} \in \mathbb{R}^{n \times n}$ the known matrix of weights, taken as $\boldsymbol{W} = {\sigma_0}^2 \boldsymbol{Q_e}^{-1}$, where σ_0^2 is the variance factor.

Here, we compute the maximum external reliability (max $(\nabla \boldsymbol{X})$) as follows:

$$\max\{\nabla \boldsymbol{X}\} = \max\{(\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A})^{-1} (\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{c_i} MDB)\} \tag{11}$$

Important to mention that the maximum external reliability max $\{\nabla \boldsymbol{X}\}$ can be a positive or a negative value. According to Equation 1, we consider the maximum influence of an undetectable outlier $\nabla_i = MDB$ on the parameters.

In the next section, we present an automatic method for geodetic network design that was computationally developed based on reliability theory. Specifically, we apply reliability theory to automatically define the best location of control points. We apply the proposed method in order to design a levelling geodetic network. Although the method was applied in a specific network, it is a generally applicable method. For example, the reliability theory has been used to measure the integrity of the receivers for civil aviation, which is a main tool for safety-of-life applications, see e.g. [41].

3. Automatic Procedure to Design the Location of Control Points in the Geodetic Network

For the establishment of a geodetic network, we must define which points of the network will have their coordinates previously determined in the desired reference system. These points are called control points, or constraint points. These points that allow the other points of the geodetic network to be linked to a reference system. Therefore, it is essential to define the location of these control points at the design stage of a geodetic network. The proposed automatic method here focuses on designing of the geodetic networks in terms of high reliability. Under the present proposal, the quality criterion to be considered during the design stage is based on the lowest value of the maximum influence of an outlier on the coordinates of the network (i.e., maximum external reliability). The method does not depend on the real measurements values but only on the model design, i.e. the network geometry and covariance matrix. The computations can be performed as follows:

1. Defining a significance level α and the type II error β in order to compute the non-centrality parameter. Here, we use the recursive algorithm based on the work by Aydin and Demirel [28], namely bisection algorithm, in order to obtain the non-centrality parameter for one degree of freedom,i.e. $\lambda_{q=1}$. Typically a value of the level $alpha = 0.001$ and $\beta = 0.2$ is adopted (see, e.g. [19]).

2. Defining a geodetic network configuration as well as the uncertainties of the observations, i.e. the design matrix $\boldsymbol{A}$ and the covariance matrix of the observations $\boldsymbol{Q_e}$, respectively. The covariance matrix of the observations $\boldsymbol{Q_e}$ may consist of random effects and the uncertainties associated with the correction of systematic effects. The latter follows from the instrument precision, measurement techniques and field condition. In this step, the design matrix and covariance matrix are conditioned to the position of the control point (or by the combination of control points) in the network. It is important to mention that the design matrix defined must have a minimum configuration to avoid rank deficiency [42].

3. Computing the MDB for each observation according to Equation 7 or 9.

4. Computing the external reliability according to Equation 10.

5. Computing and store the maximum external reliability according to Equation 11.

6. Checking whether all the points (or all combination of points) of the network were configured as control point. If not, select a new control point (or new combination of control points) and return

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

to step 3. Otherwise, the algorithm selects the configuration of the network that has the lowest value of the maximum external reliability. Important to mention that matrix $\boldsymbol{A}$ is modified when a new point (or a new combination of points) is selected as the control.

The proposed method is summarised as a flowchart in Figure 1.

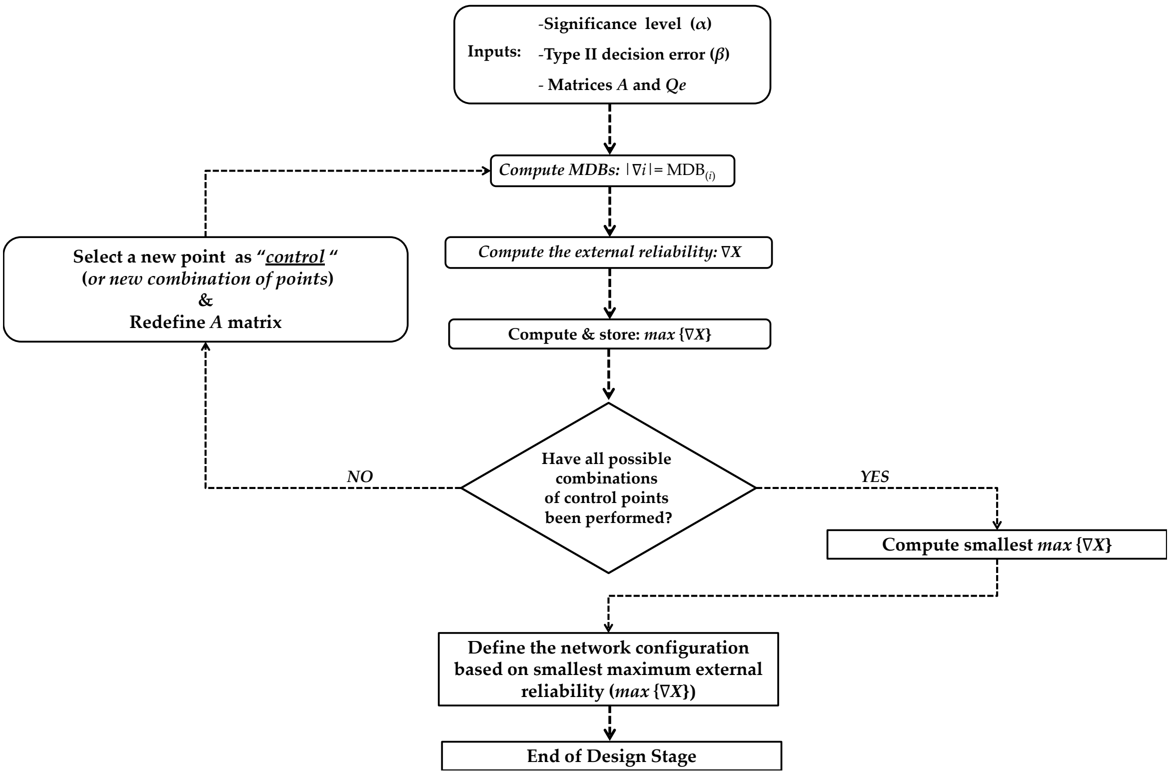

Figure 1. Flowchart of the automatic method proposed to design the location of control points in the geodetic network.

4. Results and Discussion

In order to demonstrate the design method in practice, in this section we apply it to a simulated closed levelling network. The network is displayed in Figure 2. The goal is to illustrate the design method; further considerations about levelling networks are outside the scope of this study. The results of this paper are presented for $\gamma = 0.8$ and $\alpha = 0.001$, which gives $\lambda_{q=1} = 17.075$.

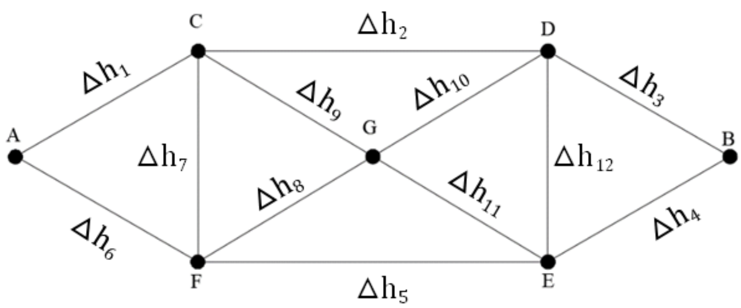

Figure 2. Simulated levelling geodetic network.

The results of the internal reliability are shown based on the relationship between MDB and the standard-deviation of the observation. As an example of that relationship, for $MDB = 5mm$ and $\sigma = 2.5mm$, the ratio is $\frac{MDB}{\sigma} = \frac{5mm}{2.5mm} = 2\sigma$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at *Appl. Sci.* **2020**, *10*, 687; doi:10.3390/app10020687

Two typical cases were considered here: (**a**) a minimally constrained and (**b**) a over-constrained least squares adjustment. The variances of the height difference (denoted by $\sigma^2_{\Delta h_i}$) are assumed normally distributed and uncorrelated. The variances were based on the relation between the differential levelling lines and their lengths. In other words, the variances of differential levelling lines are proportional to their lengths, i.e. the larger the lengths, the larger the variances of differential levelling lines.

We consider that the equipment used here is a spirit level with nominal standard deviation of $\pm 1mm/km$ for a double run levelling. In each scenario two variants are considered here:

1. all lengths of the differential levelling with 1 km, and therefore the variances equal to $1mm^2$; and
2. lines with diversified lengths, and therefore levelling lines with different variances, whose values are given in Table 1.

Table 1. Levelling lines with different lengths and variances for the case 2.

Observation	Length of line (km)	$\sigma^2_{\Delta h_i}$ (mm^2)
Δ_{h1}	1.000	1.00
Δ_{h3}	1.000	1.00
Δ_{h4}	1.000	1.00
Δ_{h6}	1.000	1.00
Δ_{h2}	1.414	2.00
Δ_{h5}	1.414	2.00
Δ_{h8}	1.732	3.00
Δ_{h9}	1.732	3.00
Δ_{h10}	1.732	3.00
Δ_{h11}	1.732	3.00
Δ_{h7}	2.000	4.00
Δ_{h12}	2.000	4.00

In the first scenario (**a**), we consider the closed levelling network in Figure 2 with availability of one control station, and 6 points with unknown heights, totalling six minimally constrained points and 7 possible cases of control point configuration. In that case, there are $n = 12$ observations and $u = 6$ unknowns, which lead to $n - u = 6$ degrees of freedom.

Moreover, the design matrix $\boldsymbol{A}$ has dimension 12×6 and the covariance matrix of observations $\boldsymbol{Q_e}$ has dimension 12×12. The stations C, D, E, F and G are involved in 4 height differences, so there are three redundant observations for the determination of these heights. On the other hand, there is one redundant observation for the determination of heights of the stations A and B.

The MDBs computed for each observation of the network and for each case of variances configuration are displayed in Figure 3. Important to mention that MDBs were invariant with regard to the position of a single control point in the network. It can be noted that the observations Δh_7 and Δh_{12} are more resistant to outlier than others, because theirs MDBs were the smallest on the network.

Table 2 shows the maximum external reliability of the network. It can be noted that the smallest value of the maximum influence of an MDB on the heights occurred when the station G was taken as control point, i.e. when the control point was set to the centre of the network (3.28 mm, marked in bold). The $\pm$ sign in Table 2 means that the maximum influence of an outlier on the network occurs in two directions (*up* and *down*). The best network configuration obtained based on the optimal position of the control point is shown in Figure 4.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

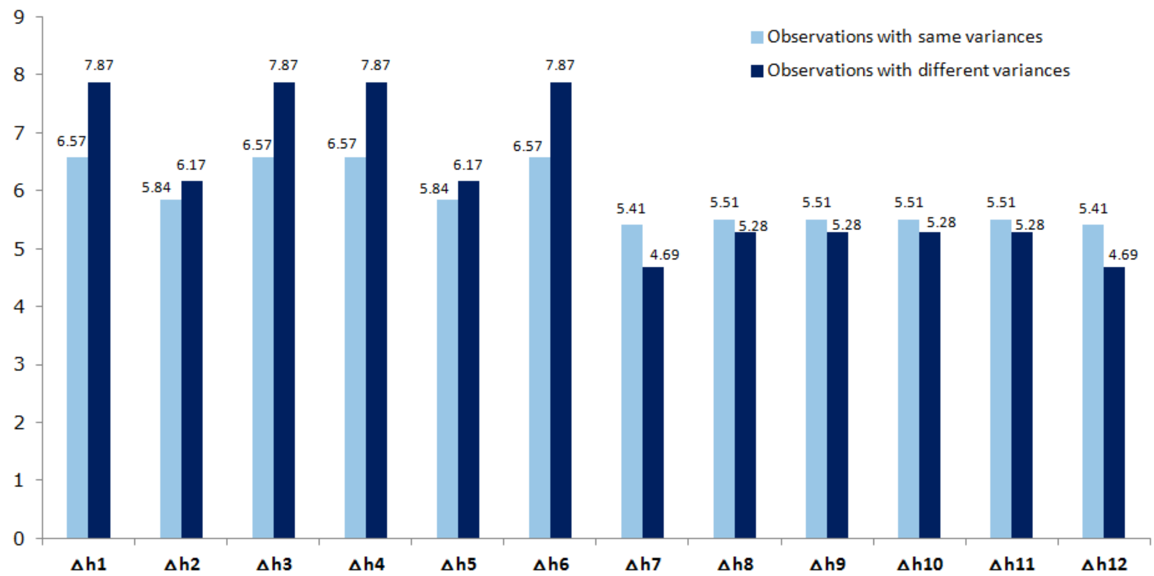

Figure 3. MDB (σ) for a single control point - minimally constrained scenario (**a**) - for the cases of observations with same variances and different variances.

Table 2. Maximum external reliability for a single control point for both observations with equal variances and with different variances.

Control Point	Maximum External Reliability (mm) Observations with same variance	Maximum External Reliability (mm) Observations with different variances
A	± 3.97	± 5.70
B	± 3.97	± 5.70
C	-3.97	-5.70
D	3.97	5.70
E	-3.97	-5.70
F	3.97	5.70
G	**$\pm$3.28**	**$\pm$3.94**

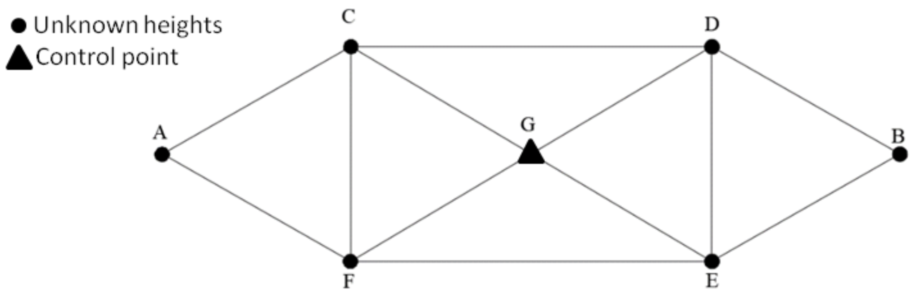

Figure 4. Optimum Configuration of the network for a single control point - minimally constrained scenario (**a**).

In the second case, we consider the closed levelling network in Figure 2 over-constrained with two control station, totalling 21 possible combinations of control points. For example, taking A and B fixed, we have $u = 5$ unknown heights (C, D, E, F, G), $n = 12$ observations and $n - u = 7$ redundant observations. In that case, the design matrix $\boldsymbol{A}$ has dimension 12×5 and the covariance matrix of observations $\boldsymbol{Q_e}$ has dimension 12×12. On the other hand, we have $u = 5$ unknown heights (B, D, E, F, G), $n = 11$ observations and $n - u = 16$ redundant observations in the case of selecting control points A and C. In that case, when the control points are adjacent, the respective levelling line are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at Appl. Sci. 2020, 10, 687; doi:10.3390/app10020687

not observed and, therefore, the design matrix $\boldsymbol{A}$ has dimension 11×5 and the covariance matrix of observations $\boldsymbol{Q}_e$ has dimension 11×11. Figure 5 shows an example when the control points are adjacent.

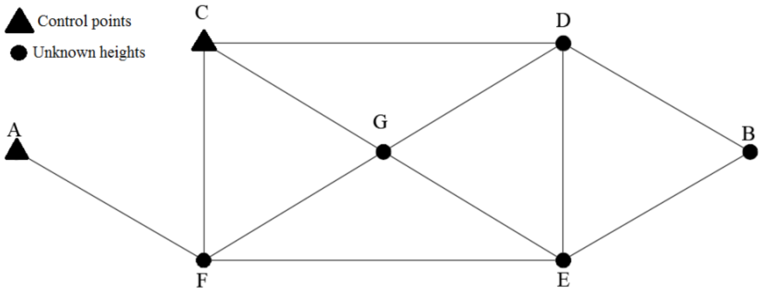

Figure 5. Example of the network configuration for adjacent control points A and C.

For the second case, we produce and analyse 21 graphs showing the MDB values. Due to the large number of graphs, we chose to show here a table with the summary of the results. Table 3 shows the overall statistics of the MDBs (average, maximum, minimum and standard-deviation) for each possible combination, two by two, of control points and for each scenario of variance. For the case of same variances, it can be noted that when the points A and B were taken as control points (AB), the observations presented a good level of homogeneity (homogeneous redundancy), i.e. all of them had the same internal reliability. It means that, in the presence of an outlier, all observations have the same ability to detect it. The search for homogeneous redundancy in all the observations has already been investigated in [7,8].

Table 3. Statistics of MDB for two control points − over-constrained scenario (**b**).

	MDB (in σ *unit*)							
	Observations with same variances				Observations with different variances			
Control point	Average	Max.	Min.	Std. Dev.	Average	Max.	Min.	Std. Dev.
AB	**5.41**	**5.41**	**5.41**	**0.00**	**5.52**	**6.39**	**4.69**	**0.67**
AC, AF, DB or BE	5.64	6.56	5.1	0.50	5.86	7.83	4.49	1.16
CD or EF	5.69	6.45	4.98	0.63	5.93	7.5	4.6	1.28
CG, DG, EG or FG	5.69	6.56	5.00	0.65	5.97	7.85	4.57	1.35
AD, AE, CB or BF	5.46	6.47	4.99	0.50	5.61	7.43	4.59	1.00
AG or BG	5.48	6.57	4.95	0.57	5.69	7.87	4.57	1.19
DE or CF	5.65	6.53	5.06	0.50	5.8	7.73	4.66	1.03
CE or DF	5.51	6.36	4.90	0.66	5.68	7.16	4.53	1.12

Figure 6 shows the maximum external reliability for the over-constrained network. The maximum external reliability of the network for combinations of control points AB, AG, BG, CD, CF, DE and EF had positive and negative signals. This is represented by the $\pm$ sign on the Figure 6. It means that the maximum influence of an outlier on the network occurs in two directions. It can be noted that both cases of observations with same variances and different variances, the smallest value of the maximum influence of an MDB on the heights occurred when the stations A and B were fixed as control, with $\max(\nabla \boldsymbol{X}) = 2.3mm$. In general, the inflation of the variances in the network inflated the maximum external reliability, i.e. it increased the maximum influence of a possible outlier on the network. The best configuration of the network according to our algorithm is showed in the Figure 7.

 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at *Appl. Sci.* **2020**, *10*, 687; doi:10.3390/app10020687

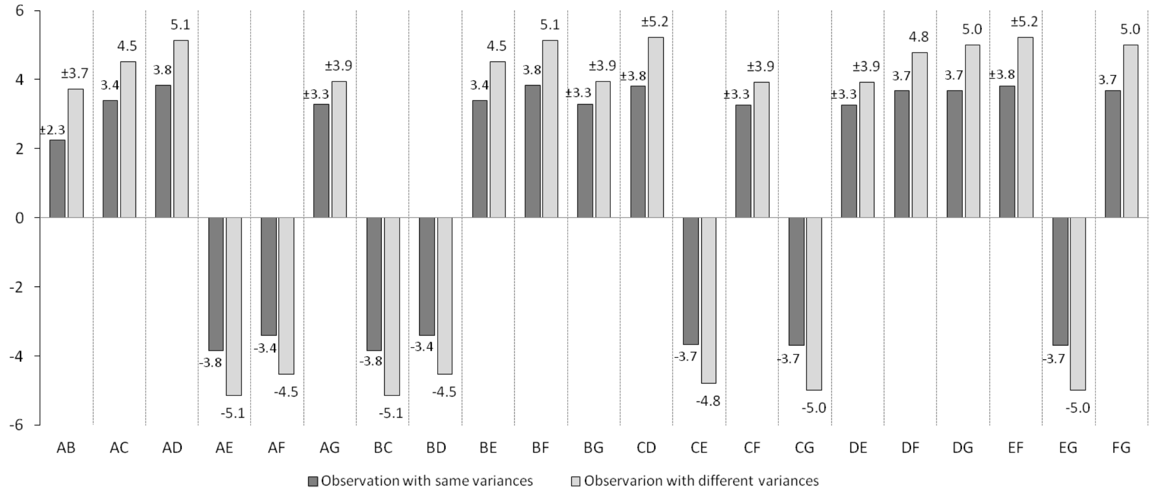

Figure 6. Maximum external reliability max (∇X) [mm] for two control points in the network.

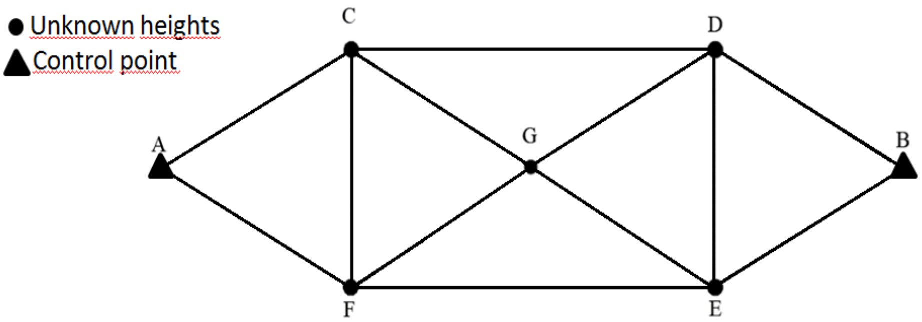

Figure 7. Optimum configuration of the network for two control points.

5. Conclusions

Within the context of applied sciences, we present an automatic procedure based on the concepts of reliability for designing geodetic networks for the definition of the best control point position. The conclusions are highlighted as follows:

1. The proposed method to design a geodetic network is based on the lowest value of the maximum external reliability. The size of the outlier is defined according to MDB for a given type I and type II errors probabilities. We highlight that the method discards the use of the observation vector of Gauss-Markov model. In fact, the only needs are the geometrical network configuration (given by Jacobian matrix) and the uncertainties of the observations (given by instrument precision, measurement techniques and/or field condition). The aim of the proposed method is to find the optimum configuration of the geodetic control points so that the maximum influence of an outlier on the coordinates of the network is minimum. Therefore, it can be applied for any kind of geodetic network.

2. Here, the proposed method was applied to a closed levelling network. The MDB was computed based on a power of the test of data snooping of 0.80 (80%) and the significance level of 0.001 (0.1%). In order to apply the concepts in practice, two scenarios were presented for a simulated levelling geodetic network: a minimally constrained network one and a over-constrained network. The observations were assumed normally distributed and uncorrelated, which usually happens in the practice of levelling network adjustment. In each scenario two variants were also considered: one in which the variances of the measurements were assumed equal and another in which the variances were different. In the case of minimally constrained, we highlight that

the centre of the simulated network was the optimum position to set the control point. In the over-constrained network, we highlight that among the 21 possibilities of configuring the control points, the stations with less line connections (i.e. with less redundant observations) provided the best configuration of geodetic network.

In future studies the idea will be to integrate reliability with precision. Therefore, it will be possible to make an analysis that considers not only the bias (given by the external reliability), but also the precision of the coordinates in the network (given by the covariance matrix of the parameters).

Author Contributions: conceptualization, M.T.M. and V.F.R.; methodology, M.T.M. and V.F.R.; software, V.F.R. and L.G.S.; validation, M.T.M., J.B.S.N. and A.C.R.A.; formal analysis, M.T.M., M.R.V. and V.F.R.; investigation, M.T.M., V.F.R., I.K., J.B.S.N. and A.C.R.A.; data curation, M.T.M., V.F.R.; writing–original draft preparation, M.T.M. and V.F.R.; writing–review and editing, M.T.M., V.F.R. and I.K.; supervision, M.T.M. and M.R.V.; project administration, M.T.M. and M.R.V.; funding acquisition, M.T.M., M.R.V. and L.G.S.

Funding: This research was funded by the CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil (proc. nº 103587/2019-5). This research and the APC were also funded by PETROBRAS (Grant Number 2018/00545-0).

Acknowledgments: Will be given after review.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Helmert, F.R. Studien über rationelle Vermessungen im Gebiet der höheren" Geodäsie". Master's thesis, 1868.
2. Baarda, W. S-transformations and criterion matrices. *Publ. on geodesy, New Series* **1973**, *5*.
3. Teunissen, P.J.G. Quality Control in Geodetic Networks. Optimization and Design of Geodetic Networks; Grafarend, E.W.; Sansò, F., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1985; pp. 526–547.
4. Vaníček, P.; Craymer, M.R.; Krakiwsky, E.J. Robustness analysis of geodetic horizontal networks. *J. Geod.* **2001**, *75*, 199–209. doi:10.1007/s001900100162.
5. Seemkooei, A.A. Comparison of reliability and geometrical strength criteria in geodetic networks. *Journal of Geodesy* **2001**, *75*, 227–233. doi:10.1007/s001900100170.
6. Seemkooei, A.A. Strategy for Designing Geodetic Network with High Reliability and Geometrical Strength. *Journal of Surveying Engineering* **2001**, *127*, 104–117. doi:10.1061/(ASCE)0733-9453(2001)127:3(104).
7. Amiri-Simkooei, A. A NEW METHOD FOR SECOND ORDER DESIGN OF GEODETIC NETWORKS: AIMING AT HIGH RELIABILITY. *Survey Review* **2004**, *37*, 552–560. doi:10.1179/sre.2004.37.293.552.
8. Amiri-Simkooei, A.; Sharifi, M.A. Approach for Equivalent Accuracy Design of Different Types of Observations. 2004.
9. 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.
10. Hekimoglu, S.; Erdogan, B. New Median Approach to Define Configuration Weakness of Deformation Networks. *Journal of Surveying Engineering* **2012**, *138*, 101–108. doi:10.1061/(ASCE)SU.1943-5428.0000080.
11. Klein, I.; Matsuoka, M.T.; Souza, S.F.d.; Collischonn, C. Planejamento de redes geodÃresistentes a mÃoutliers. *Boletim de CiÃGeodÃ* **2012**, *18*, 480 – 507.
12. 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.
13. 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.
14. 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.
15. 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.

16. 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.
17. Rofatto, V.F.; Matsuoka, M.T.; Klein, I.; Veronez, M.R. Monte-Carlo-based uncertainty propagation in the context of Gauss–Markov model: a case study in coordinate transformation. *Scientia Plena* **2019**, *15*, 1–17. doi:10.14808/sci.plena.2019.095401.
18. Baarda, W. Statistical concepts in geodesy. *Publ. on geodesy, New Series* **1967**, *2*.
19. Baarda, W. A testing procedure for use in geodetic networks. *Publ. on geodesy, New Series* **1968**, *2*.
20. Koch, K.R. *Parameter estimation and hypothesis testing in linear models.*, 2 ed.; Springer, 1999.
21. Teunissen, P. *Testing Theory: an introduction*, 2 ed.; Delft University Press, 2006.
22. Ghilani, C.D. *Adjustment Computations: Spatial Data Analysis*, 6 ed.; John Wiley Sons, Ltd, 2017.
23. 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.
24. Kargoll, B. On the theory and application of model misspecification tests in geodesy. Doctoral thesis, University of Bonn, Landwirtschaftliche Fakultät, German, Bonn, 2007.
25. 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.
26. Teunissen, P. An Integrity and Quality Control Procedure for use in Multi Sensor Integration. 1990.
27. Vaníček, P.; Krakiwsky, E.J. *Geodesy: The Concepts*, 2 ed.; Elsevier Science, 1987.
28. 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.
29. Wang, J.; Chen, Y. On the reliability measure of observations. *Acta Geodaetica et Cartographica Sinica* **1994**, pp. 42–51.
30. Pelzer, H. Detection of errors in the functional adjustmentmodel. Mathematical models of geodetic/photogrammetric point determination with regard to outliers and systematic errors; Ackermann, F., Ed.; , 1983; Vol. A 98.
31. Prószyński, W. Criteria for internal reliability of linear least squares models. *Bulletin géodésique* **1994**, *68*, 162–167. doi:10.1007/BF00808289.
32. Schaffrin, B. Reliability Measures for Correlated Observations. *Journal of Surveying Engineering* **1997**, *123*, 126–137. doi:10.1061/(ASCE)0733-9453(1997)123:3(126).
33. Teunissen, P.J.G. Minimal detectable biases of GPS data. *Journal of Geodesy* **1998**, *72*, 236–244. doi:10.1007/s001900050163.
34. Prószyński, W. Another approach to reliability measures for systems with correlated observations. *Journal of Geodesy* **2010**, *84*, 547–556. doi:10.1007/s00190-010-0394-2.
35. Ding, X.; Coleman, R. Multiple outlier detection by evaluating redundancy contributions of observations. *Journal of Geodesy* **1996**, *70*, 489–498. doi:10.1007/BF00863621.
36. 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.
37. Gökalp, E.; Güngör, O.; Boz, Y. Evaluation of Different Outlier Detection Methods for GPS Networks. *Sensors* **2008**, *8*, 7344–7358. doi:10.3390/s8117344.
38. Baselga, S. Nonexistence of Rigorous Tests for Multiple Outlier Detection in Least-Squares Adjustment. *Journal of Surveying Engineering* **2011**, *137*, 109–112. doi:10.1061/(ASCE)SU.1943-5428.0000048.
39. 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.
40. Lehmann, R.; Lösler, M. Multiple Outlier Detection: Hypothesis Tests versus Model Selection by Information Criteria. *J. Surv. Eng.* **2016**, *142*, 04016017,
41. Fu, L.; Zhang, J.; Li, R.; Cao, X.; Wang, J. Vision-Aided RAIM: A New Method for GPS Integrity Monitoring in Approach and Landing Phase. *Sensors* **2015**, *15*, 22854–22873. doi:10.3390/s150922854.
42. Xu, P. Sign-constrained robust least squares, subjective breakdown point and the effect of weights of observations on robustness. *Journal of Geodesy* **2005**, *79*, 146–159. doi:10.1007/s00190-005-0454-1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 December 2019 doi:10.20944/preprints201912.0232.v1

Peer-reviewed version available at *Appl. Sci.* **2020**, *10*, 687; doi:10.3390/app10020687