Monte Carlo Simulation for Outlier Identification Studies in Geodetic Network: An Example in A Levelling Network Using Iterative Data Snooping

M.T. Matsuoka [a,c] iD **, V.F. Rofatto** [a,c]**, I. Klein** [b]**, A. F. S. Gomes** [a] **, M.P. Guzatto** [b]

[a] *Institute of Geography, Federal University of Uberlândia (UFU), Monte Carmelo, Brazil*
[b] *Landing Surveying Program, Federal Institute of Santa Catarina (IFSC), Florianopolis, Brazil*
[c] *Graduate Program of Remote Sensing, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil*

Abstract: *Today with the fast and powerful computers, large data storage systems and modern softwares, the probabilities distribution and efficiency of statistical testing algorithms can be estimated using computerized simulation. Here, we use Monte Carlo simulation (MCS) to investigate the power of the test and error probabilities of the Baarda's iterative data snooping procedure as test statistic for outlier identification in the Gauss-Markov model. The MCS discards the use of the observation vector of Gauss-Markov model. In fact, to perform the analysis, the only needs are the Jacobian matrix; the uncertainty of the observations; and the magnitude intervals of the outliers. The random errors (or residuals) are generated artificially from the normal statistical distribution, while the size of outliers is randomly selected using standard uniform distribution. Results for simulated closed leveling network reveal that data snooping can locate an outlier in the order of magnitude 5σ with high success rate. The lower the magnitude of the outliers, the lower is the efficiency of data snooping in the simulated network. In general, considering the network simulated, the data snooping procedure was more efficient for α=0.01 (1%) with 82.8% success rate.*

How to Cite (APA 6 th Style):
Matsuoka, M.T., Rofatto, V.F., Klein, I., Gomes. A.F.S, Guzatto, M.P. (2019). Monte Carlo Simulation for Outlier Identification Studies in Geodetic Network: An Example in a Levelling Network Using Iterative Data Snooping. *Geoplanning: Journal of Geomatics and Planning, 6*(1), 21-30. doi:10.14710/geoplanning.6.1.21-30.

1. INTRODUCTION

Data snooping is the best established method for identification of gross errors in geodetic data analysis. This method is due to (Baarda, 1968). Here, it is assumed that outliers are observations contaminated by gross errors (blunders), following the statement of Lehmann (2012) that in Geodesy, 'outliers are most often caused by gross errors and gross errors most often cause outliers'. In practice, Data Snooping procedure is applied iteratively, identifying and removing an outlier at a time. The method is applied until no observations are identified. Here, this procedure will be called Iterative Data Snooping. Since data snooping is based on a statistical hypothesis testing, it may lead to a false decision as follows:

- Type I error or false alert (probability level α) – Probability of detecting an outlier when there is none;
- Type II error or missed detection (probability level β) – Probability of non-detecting an outlier when there is at least one; and
- Type III error or wrong exclusion (probability level κ) – Probability of misidentifying a non-outlying observation as an outlier, instead of the outlying one.

Lehmann & Voß-Böhme (2017) mention that while the rate of type I decision error can be selected by the user, the rate of type II decision error cannot. They also point out that a test statistic with a low rate of type II is said to be powerful. However, without considering the Type III error, there is a high risk of over-estimating the successful identification probability. Besides that, we highlight that the Iterative Data Snooping procedure can identify more observations than real number of outliers (we call here "over-

identification"). Thus, we consider a powerful statistical test when the rates of type II and type III errors as well as the over-identification are simultaneously minimized for a given probability level α.

From this point, we posed the following problem: how to compute the probabilities levels above? Unlike Baarda, we have fast computers at our disposal. In this paper we show that the statistical quantities can be determined by frequency distributions of computer random experiments performed using random numbers. This is known as Monte Carlo Simulation (MCS). MCS methods are used whenever the functional relationships are analytically not tractable, as is the case for Iterative Data Snooping procedure (Rüdiger Lehmann, 2012b). The MCS has already been applied in outlier detection (Lehmann & Scheffler, 2011; Lehmann, 2012; Klein et al., 2012; Klein et al., 2015; Erdogan, 2014; Niemeier & Tengen, 2017)

The studies presented in this paper are a continuation of the first experiments presented by Rofatto et al. (2017). However, unlike Rofatto et al., (2017), here in this paper we evaluate the proposed method in a geodetic network with uncorrelated observations and also we analyze the power of the test of Iterative Data Snooping procedure when outliers of magnitude equal to the MDB (Minimal Detectable Bias) are inserted into the geodetic network.

The outline of the paper is as follows: first the paper show a theoretical background of Iterative Data Snooping procedure in the Gauss Markov model. Next, the MCS approach is introduced as tool to analyze the power of the test and the probabilities of decision errors (type II, type III and over-identification) of the Iterative Data Snooping procedure. Then, the efficiency of the Iterative Data Snooping is demonstrated by means of the Monte Carlo method on the example of simulated closed leveling network. The mathematical model generally adopted in geodetic data analysis is the linear(ized) Gauss-Markov model, given by Koch (1999):

$$e = y - A\,x \; ....(1)$$

where e is the n x 1 random error vector, A is the n x u design (or Jacobian) matrix with full rank column, x is the u x 1 unknown parameters vector and y is the n x 1 observations vector. The most employed solution for a redundant system of equations ($n > u$) is the weighted least squares estimator (WLSE) for the vector of unknowns ($\hat{x}$):

$$\hat{x} = (A^{T}WA)^{-1}(A^{T}Wy) \; ....(2)$$

In which W is the n x n weight matrix of the observations, taken as $W = \sigma_0^2 \Sigma_y^{-1}$, where σ_0^2 is the variance factor (here it is assumed as known) and Σ_y is the covariance matrix of the observations; if Σ_y is diagonal, one speaks of weighted LSE (WLSE); if it is full, generalized LSE (GLSE). More details about LSE estimation in Ghilani (2017). A geometric interpretation of the LSE can be found in Teunissen (2003) and Klein et al. (2011).

The least-squares method is the *Best Linear Unbiased Estimator* (BLUE) for the unknown parameters and it is also a maximum likelihood solution when the observation errors follow a central Gaussian distribution (Teunissen, 2003). However, the least squares is no longer optimal in the presence of grossly erroneous observations (Baarda, 1968). In other words, despite optimal properties for least square, they lack robustness or insensitivity to outliers in observations (Huber, 1992; Rousseeuw & Leroy, 1987; Lehmann, 2013). In recent years, two categories of advanced techniques for the treatment of observations contaminated by outliers have been developed: robust adjustment procedures (Wilcox, 2011; Klein et al., 2015) and outlier detection based on statistical tests (Klein et al., 2016) . The first one is outside the scope of this paper. Besides the undoubted advantages of robust adjustment, the outlier tests are also used. The following advantages of outlier analysis are mentioned by Lehmann (2013);

- Detected outliers provide the opportunity to investigate causes of gross measurement errors;
- Detected outliers can be re-measured; and
- If the outliers were discarded from the observations then the standard adjustment software, which operates according to the least squares principle, can be used.

Data snooping procedure is a particular case of maximum likelihood ratio test when only one outlier (i.e. q = 1) is present in the data set at a time (Baarda, 1968; Berber & Hekimoglu, 2003; Lehmann, 2012) Thus, it is formulated by the following test hypotheses (Baarda, 1968; Teunissen, 2006):

$$H_0 : E\{y\} = Ax \qquad vs \qquad H_A : E\{y\} = Ax + c_y \nabla; \nabla \neq 0 \dots\dots\dots (3)$$

Where c_y is outlier model for q=1, i.e. the n x 1 unit vector with 1 in its ith entry and zeros in the remaining (e.g. $c_{y_{nx1}} = \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 & 0 \end{bmatrix}^T$), and ∇ is a scalar value with the gross error (*outlier*) at i^{th} observation being tested. Therefore, in the null hypothesis (H_0), it is assumed that there are no outliers in the observations, while in the alternative hypothesis (H_A), is it assumed that the ith observation being tested (for i $= 1,\dots,n$) is contaminated by gross error of magnitude ∇.

If we consider one outlying observation in at certain known locations (q = 1), then the likelihood ratio test for data snooping (T_q = 1) is given by Teunissen (2006):

$$T_{q=1} = \hat{e}_0^T \sum\nolimits_y^{-1} c_y (c_y^T \sum\nolimits_y^{-1} \sum\nolimits_{\hat{e}_0} \sum\nolimits_y^{-1} c_y)^{-1} c_y^T \sum\nolimits_y^{-1} \hat{e}_0 \dots\dots\dots (4)$$

where $\hat{e}_0$ and $\sum_{\hat{e}_0}$ is the estimated random error vector and a posteriori covariance matrix of the estimated random error computed by LSE into H_0, respectively. Under H_0, observation errors are zero-mean (multivariate) normally distributed. The null hypothesis is rejected if the following test statistic (T_q = 1) of the ith observation being tested exceeds a given critical value K_α , i.e.:

$$\begin{cases} \text{Reject } H_0 \text{ if: } T_{q=1} > K_\alpha \\ H_0 : T_{q=1} \sim \chi^2_{(1,0)}; \ H_A : T_{q=1} \sim \chi^2_{(1,\delta)}, \ \text{with } \delta = c_y^T \sum\nolimits_y^{-1} \sum\nolimits_{\hat{e}_0} \sum\nolimits_y^{-1} c_y \nabla^2 \end{cases} \dots\dots (5)$$

Important to mention that the critical value follows from a chi-squared distribution with one degree freedom at a significance level of α in a one-tailed test. Baarda (1968) and Teunissen (2006) demonstrate that if q = 1, then the test statistics (equation 4) can also be formulated based on a standard normal distribution in a two-tailed test (so-called *w-test*). Both the chi-squared and normal distribution tests are equivalent. Usually in geodesy, the value of α is set between 0.1% and 1% (Kavouras, 1982; Aydin & Demirel, 2004; Lehmann, 2013). Furthermore, data snooping contains multiple alternative hypotheses, as each observation is individually tested. Therefore, the only observation considered contaminated by outlier is the one whose test statistic satisfies the inequalities $T_{q=1} > K_\alpha$. In the case that two or more observations exceed the critical value K_α only the observation with the largest $T_{q=1}$ is flagged as an outlier. After having identified the observation most suspected of being an outlier (at given α), it is excluded usually from the model, and the WLSE and data snooping procedure are applied iteratively until there are no further outliers identified in the observations (Berber & Hekimoglu, 2003).

The power of the test (γ) is the probability of correctly identifying the outliers. In the case of a round of Data Snooping, the power of the test depends on the type II and type III errors, for a given level of significance (α) (i.e. γ = 1 − (β + κ)). Considering the "Iterative Data Snooping", the power of the test also depends on *over-identification* error, and it is given by γ = 1 − (β + κ + over-identification). Baarda's conventional reliability theory considers only a single alternative hypothesis (Baarda, 1968), and therefore, it is based only on type I and II errors. Type III error is addressed by Förstner (1983) considering two alternative hypotheses. Yang et al. (2013) extended the solution given by Förstner (1983), and presented an analytical solution for type III error considering multiple alternative hypotheses and the presence of an outlier (i.e. for a round of Data Snooping). Examples of the efficiency of the analytical solution presented by Yang et al. (2013) can be found in Klein et al. (2015).

The focus of this paper is the Iterative Data Snooping. An analytical solution to the probabilities of decision error and power of the test for Iterative Data Snooping has not yet been developed and is of rather difficult solution. A well-established procedure to compute the probabilities levels is the Monte Carlo Simulation (MCS). As pointed out by Lehmann (2012), in essence the MCS replaces random variables by computer generated pseudo random numbers, probabilities by relative frequencies and expectations by arithmetic means over large sets of such numbers. A computation of one set of pseudo random numbers is a Monte Carlo experiment. In Geodesy, Monte Carlo Simulation has been applied in some studies since the

pioneering work of (Hekimoglu & Koch, 1999). For example, Lehmann & Scheffler (2011) have already applied MCS in data snooping to determine the optimal level of error probability α (type I error). Here, on the other hand, we proposed to use MCS in the "Iterative Data Snooping" to compute the follows probabilities levels: power of the test; type II error and type III error. In addition to these probabilities, we also compute the rate of experiments where the Iterative Data Snooping procedure identified more outliers than simulated (we call *"over-identification"*- i.e., q > 1). In the next section we show how to obtain these probabilities levels experimentally. Thus, we can analyze the efficiency of the Iterative Data Snooping testing procedure based on MCS as promised by the title of this paper.

2. DATA AND METHODS

In order to analyze the Iterative Data Snooping procedure, the MCS was applied to compute the probabilities levels. To do so, a sequence of m random errors vector e_K, $k = 1, \ldots, m$ of a desired statistical distribution is generated. The "m" is known as the number of Monte Carlo experiments. Usually, assume that the random errors of the good measurements are normally distributed with expectation zero. Thus, we generate the random errors using multivariate normal distribution, since the assumed stochastic model for random errors is based on covariance matrix of the observations, i.e. $e \sim N(0, \sigma_0^2 \Sigma_y)$.

On the other hand, an outlier (q=1) is selected based on magnitude intervals of the outliers for each m Monte Carlo experiments. Positive and negative outliers are clipped between 3σ and 3.5σ, 3.5σ and 4σ; 4σ and 4.5σ; 4.5σ and 5σ; 5 and 5.5σ; 5.5σ and 6σ; 6σ and 6.5σ; 6.5σ and 7σ; 7σ and 7.5σ; 7.5σ and 8σ; 8σ and 8.5σ; 8.5σ and 9σ in each experiment (σ is the standard deviation of the observation). Here, we use the standard uniform distribution to select the outlier magnitude. The uniform distribution is a rectangular distribution with constant probability and implies the fact that each range of values that has the same length on the distributions support has equal probability of occurrence (Lehmann & Scheffler, 2011). For example, for 10,000 Monte Carlo experiments, if the one choices a magnitude interval of the outliers of |3σ to 9σ|, the probability of a 3σ error occurring is virtually the same as -3σ, and so on. At each iteration of the simulation, a specific observation is chosen to receive a gross error based on the discrete uniform distribution (i.e., all observations have the same probability of being selected). Random and gross errors are assumed to be independent (by definition) and both are combined to the total error as follow (Kavouras, 1982):

$$\varepsilon = e + c_y \nabla, \ \nabla \neq 0 \ \ldots\ldots (6)$$

where ε is the n x 1 total error vector, e is n x 1 random errors vector and c_y is outlier model for q=1 (see expression 3), and ∇ is a scalar value with the outlier at ith observation being tested. We assume that ∇ >e. Before computing statistical test T$_q$=1 (expression 4) it is necessary to relate the random error vector e and total error vector ε, since this statistical test depends on the estimated random error vector $\hat{e}_0$. In the sense of LSE, this relationship is given by Kavouras (1982): in which R is the n x n redundancy matrix and I is the n x n identity matrix.

$$\hat{e}_0 = R\varepsilon, \ \ldots\ldots (7)$$

$$R = I - A(A^T W A)^{-1} A^T W \ \ldots\ldots (8)$$

In the equation 7 the reader should be informed that the multiplication of the redundancy matrix (R) and the total error ε provides the estimated random error vector $\hat{e}_0$. Now, the $\hat{e}_0$ is not only composed by random errors, but also it has one of its elements contaminated by an outlier. Now it becomes possible to compute the test statistic T$_q$=1 considering the relation given by equation 4.

The significance level is varied, taken as α = 0.001 (0.1%), α = 0.01 (1%), α = 0.025 (2.5%), α = 0.05 (5%) and α = 0.1 (10%). Each simulation has a unique combination of significance level and interval of magnitude of outliers. We ran 10,000 experiments for each simulation and compute the probabilities levels of type II error, type III error, the power of the test and the number of over-identification (more outliers identified than simulated) in Iterative Data Snooping, totaling 12 x 5 x 10,000 = 600,000 Monte Carlo simulations. It is

important to emphasize that the proposed method does not depend on the unknown parameters vector or the vector of observations.

3. RESULTS AND DISCUSSION

In order to demonstrate the analysis of the efficiency of data snooping, we simulated a closed leveling network. The goal is to illustrate how to use MCS approach to compute statistical quantities numerically; further considerations about levelling networks are outside the scope of this study.

We consider a closed levelling network, with one control station (benchmark), and 4 points with unknown heights (A, B, C and D), totaling four minimally constrained points as shown in Figure 1. The benchmark is fixed, and the distances of the adjacent and non-adjacent stations are approximately 240 m and 400 m, respectively. The equipment used is a spirit level with nominal standard deviation for a single staff reading of 0.02 mm/m. Lines of sight distances are kept at 40 m. Thus, each total height difference Δh_i between adjacent or non-adjacent stations is made of, respectively, three or five partial height differences (p). Each partial height difference, in turn, involves one instrument setup and two sightings: forward and back. The standard deviation for each Δh_i equals to $\sigma_i = \pm\sqrt{2p} \times 40 \times 0.02 \text{ mm/m} \therefore \sigma_i = \pm\sqrt{2p} \times 0.8 \text{ mm}$, where p is 3 or 5. The readings are assumed uncorrelated and $\sigma_0^2 = 1$.
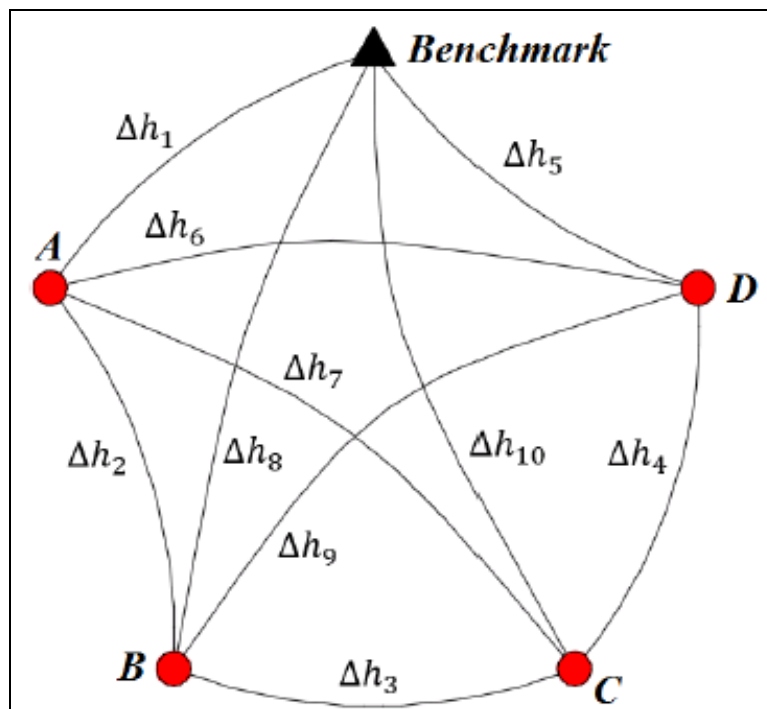

Figure 1. Simulated leveling network

For each unknown point, there are four height difference measurements. Thus, there are $n = 10$ observations, $u = 4$ unknowns, and $n - u = 6$ degrees of freedom in this simulation. The design matrix (A) has dimension 10 x 4 and the covariance matrix of observations has dimension $10 \text{ x } 10$. Each station is involved in four height differences, so there are three redundant observations for the determination of each unknown.

In the sense of reliability, the minimum and maximum redundancy numbers of the observations in the network are 0.46 and 0.75, respectively. This means that the ability of the outlier detection is not uniform in every part of the network. We also compute the Minimal Detectable Bias (MDB) as an indicator of system internal reliability. The MDB is derived from a local test proposed by Baarda (1968), which makes a decision between the null and a unique alternative hypothesis. By definition the MDB is based on Type I (false alert) and Type II (missed detection) error. The conventional MDBs are ranged from 4.7σ to 6σ for α=0.001; 3.9σ to 5σ for α=0.01; 3.5σ to 4.5σ for α=0.025; 3.2σ to 4σ for α=0.05; finally, 2.8σ and 3.6σ for

α=0.1. The computations of these MDBs were based on probability of Type II error of 0.2 (Baarda, 1968). In addition, the maximum positive and negative correlation between the test statistics are 0.61 (between Δh_4 and Δh_5) and -0.58 (between Δh_2 and Δh_3), respectively. The correlation coefficient is presented by Förstner (1983).

Applying the method presented in section 3, the success and error probabilities of Iterative Data Snooping were estimated. Figure 2 shows the success rate (number of experiments that only outlying observation was identified), i.e. the power of the Iterative Data Snooping testing procedure for one simulated outlier (γ). The misidentifications rates are showed in the Figures 3-4. The misidentifications are divided in two types of classes are counted in the simulations: number of experiments where the procedure yielded none observation identification (type II error - β); number of experiments in which the procedure detected a single observation but wrong identification (type III error - κ). In addition to these classes, we consider "over-identification", i.e. the number of experiments where the procedure detects more outliers than simulated.
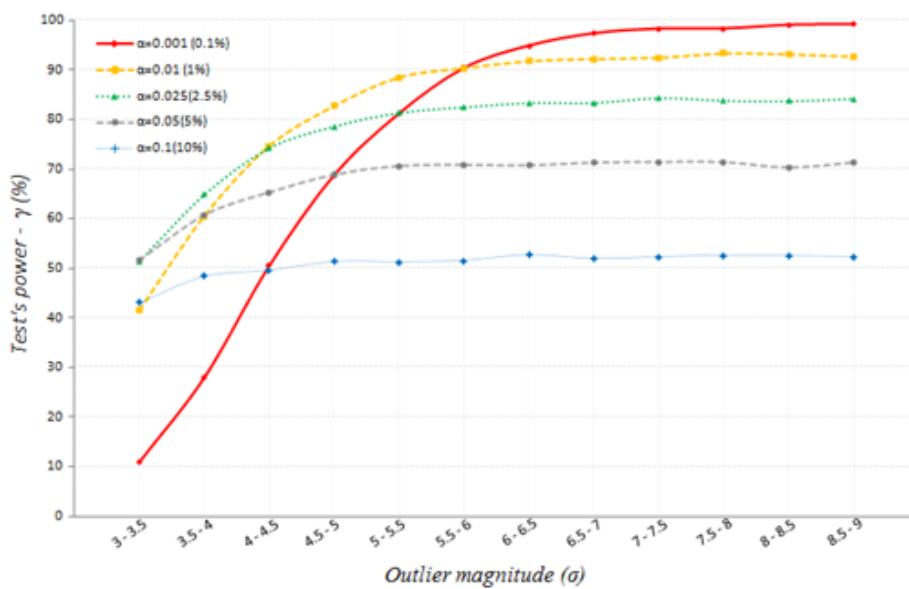

Figure 2. Success rate (power of the test) of the iterative data snooping testing procedure for simulated leveling network vs. magnitude intervals of the outliers for each probability level α.
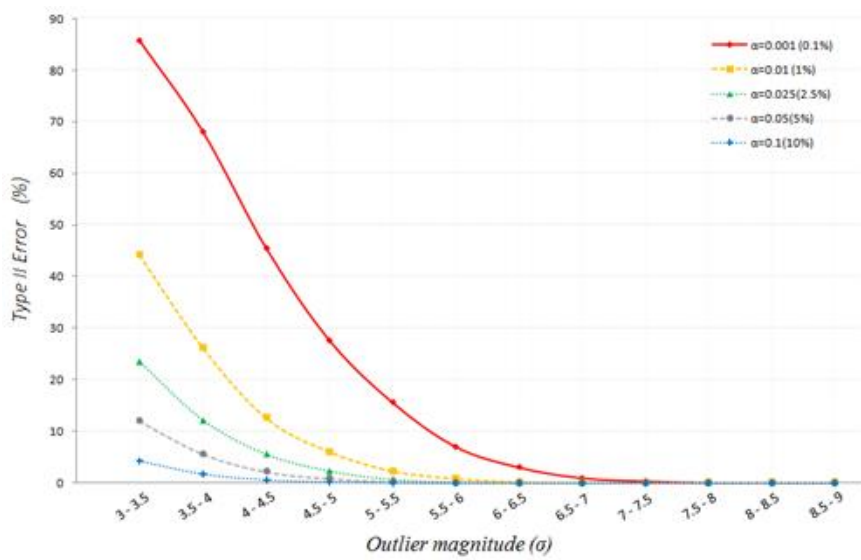

Figure 3. Type II error for simulated leveling network vs. magnitude intervals of the outliers for each probability level α.
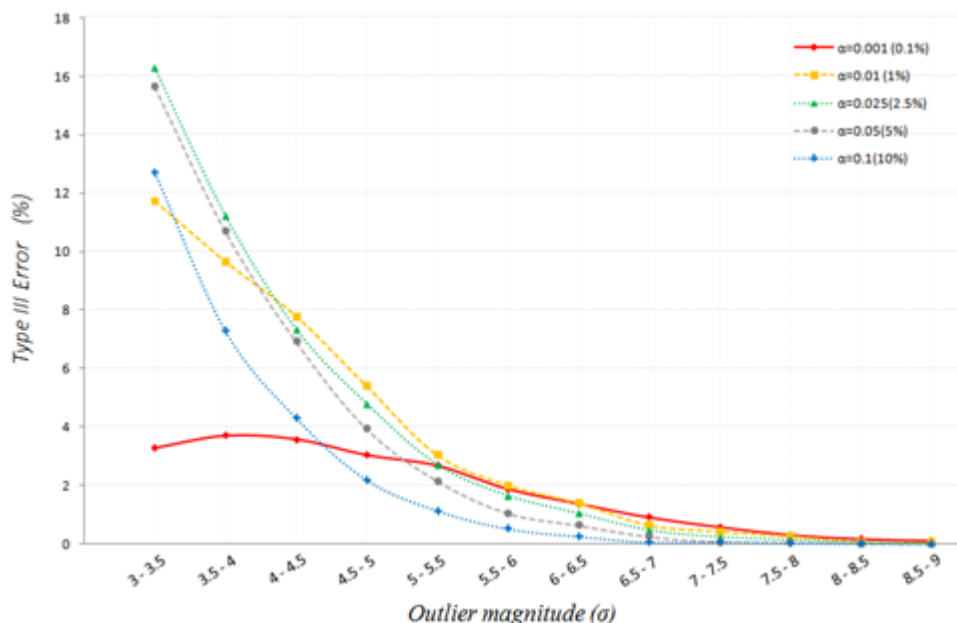
Figure 4. Type III error for simulated leveling network vs. magnitude intervals of the outliers for each probability level α.
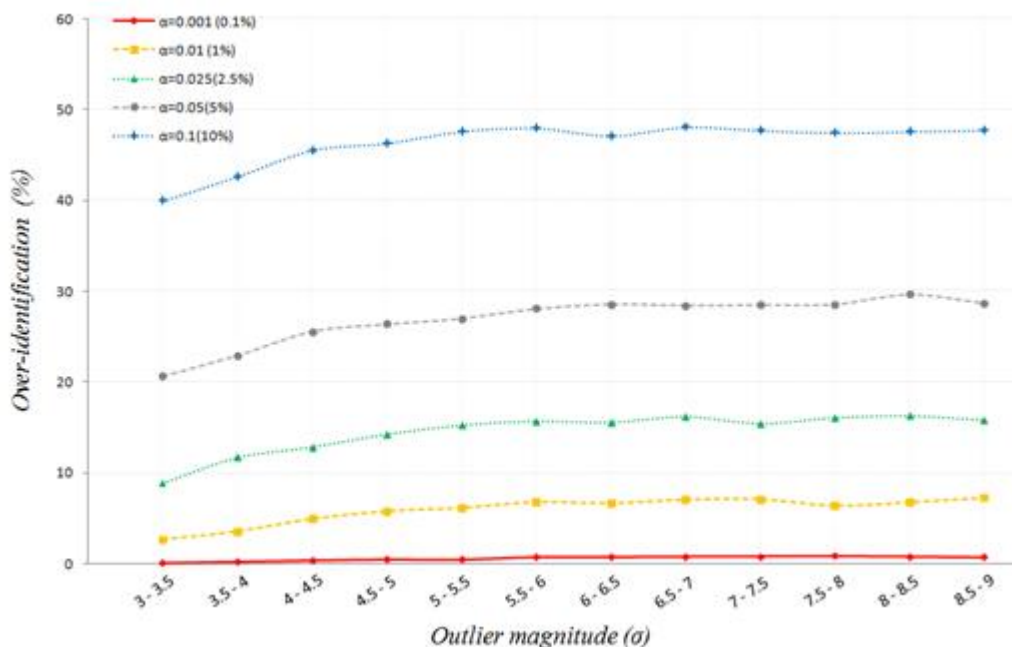

Figure 5. Over-identification vs. magnitude intervals of the outliers for each probability level α.

Figures 3–5 show that, in general, the lower the magnitude of the outliers, the lower is the efficiency of data snooping in the simulated network. This is expected. However, there is not direct relation between the power of the test and significance level (α) for the network analyzed. Here, we do not recommend to use α=0.1(10%), because many good observations are eliminated. In this case, as shown in the Figure 5, the over-identification rate stands out in relation to the other types of errors. Furthermore, in this case (α=0.1), the power of the test is virtually independent of the outlier size (see Figure 1).

It appears that higher values for α are not recommended for outliers of greater magnitude and that lower values for α are not recommended for outliers of smaller magnitude. Therefore, these results show the importance of a correct choice of α, as pointed out by Lehmann (2012); it also highlights the challenges in controlling the error rate in multiple hypotheses tests. Regarding the three classes of misidentification rates, in general, an increase in the magnitude interval of outliers, leads to a slight increase in the over-

identification rate (more outliers being identified than simulated) and a cutback in the type II error. This fact is due to the error propagation among all residuals. The rate of cases with correct number of outliers but with wrong identification (type III error) also decreases when increasing the magnitude interval of outliers.

We can observe (Figure 2) from the interval of magnitude outlier (5σ - 5.5σ) the value of the power of the test (success rate) is practically stable for all levels of significance (except for α=0.001): approximately, 50% (α=0.1), 70% (α=0.05), 83% (α=0.025) and 92% (α=0.01). For α=0.001, the success rate is greater than 90% from the interval of magnitude outlier (5.5σ - 6σ).

In general, considering the network simulated, the Iterative Data Snooping procedure is efficient for outliers greater than 5σ, with a mean success rate of 76.4% for α=0.001(0.1%); 82.8% for α=0.01 (1%); 78% for α=0.025 (2.5%); and 67.9% for α=0.05 (5%). Therefore, with an appropriate choice of α, results show that data snooping can locate an outlier in the order of magnitude 5σ with high success rate. However, the number of outliers to be considered, that also affects the efficiency of Iterative Data Snooping, requires further investigation.

In order to compare the power of the test of Iterative Data Snooping and the conventional power of the test (80%), the method described in section 3 was also applied considering the outliers with the size of its MDBs for each significance level. As pointed out, the MDBs ranged from 4.7σ to 6σ for α=0.001; 3.9σ to 5σ for α=0.01; 3.5σ to 4.5σ for α=0.025; 3.2σ to 4σ for α=0.05; finally, 2.8σ and 3.6σ for α=0.1. Such MDBs were based on the conventional power of the test of 0.8 (80%). The probabilities of committing different types of errors and the power of the Iterative Data Snooping considering these MDBs for each significance level are showed in the Table 1.

Table 1. Probabilities of Iterative Data Snooping (%) considering the size of conventional MDBs

Significance levels α	Power of the test %	Type II Error %	Type III error %	Over-identification %
0.001	77.09	19.97	2.5	0.45
0.01	70.72	17.66	6.76	4.86
0.025	62.94	15.7	10.35	11.01
0.05	53.05	12.66	12.82	21.47
0.1	38.41	9.15	15.52	36.92

In Table 1, it is noticeable that the higher the significance level, the greater the divergence between the power of the Iterative Data Snooping and the power used to compute the MDB (i.e. 80%). The explanation for this difference is that the computation of the MDB depends only on Type I and Type II error, while the Iterative Data Snooping also considers the probability of Type III error and over-identification. In future research, it is intended to investigate a function that relates the power of the test considering the MDB to the power of Iterative Data Snooping.

To conclude, it is important to mention that outlying observation can be presented among the detected observations in the over-identification case. If all detected observations are wrong, the over-identification case could be classified as type III error. The over-identification case will be investigated in more details in future studies.

4. CONCLUSIONS

Many methods of quality control for geodetic data analysis have been developed and investigated since the pioneering work of Baarda (1968). However, these methods still deserve further investigation. Thus, the goal of this paper was to analyze the data snooping testing procedure to locate an outlier by means of the MCS. The MCS discards the use of the observation vector of Gauss-Markov model. In fact, to perform the analysis, the only needs are the geometrical network configuration (given by Jacobian matrix); the uncertainty of the observations (given by nominal standard deviation of the equipment); and the

magnitude intervals of the outliers. The random errors (or residues) are generated artificially from the normal statistical distribution, while the size of outliers is selected using standard uniform distribution.

Iterative Data Snooping shows high success rates in the experiments of a simulated levelling network for single outlier randomly generated between four and five standard deviations. However, the efficiency of Iterative Data Snooping significantly decreases for outlier smaller than five standard deviations. The efficiency of the data snooping also depends on the significance level (α). Here, the optimal value for the significance level was 0.01 (1%) for the simulated network. When we insert the MDB as an outlier in the geodetic network, we verified that the higher the significance level, the greater the difference between the power of the test of Iterative Data Snooping and the power of the test used for the MDB computation. In future research, it is intended to investigate a function that relates the power of the test considering the MDB to the power of Iterative Data Snooping.

Finally, we show that Monte Carlo Simulation is a feasible method to compute the probabilities level associated to a statistical testing procedure regardless of the statistical tables. Future studies should consider various issues: the performance of the data snooping in cases of linearized (originally non-linear) models; it should consider geodetic networks in the sense of multiple outliers; the development of reliability measures; and the method performance in different networks with various geometry and varying redundancy. There are others approaches to identify multiple outliers in the observations, such as the recent proposal of Lehmann & Lösler (2016) using the p-value concept and the Sequential Likelihood Ratio Tests for Multiple Outliers (SLRTMO) presented by Klein et al. (2016). A suggestion for future work is to increase the power of the test (success rate) of Iterative Data Snooping procedure by means of a unifying testing procedure relating the iterative Data Snooping (a single outlier at time) with approaches for multiple outliers identification such as those presented in Lehmann & Lösler (2016) and Klein (2016).

5. ACKNOWLEDGMENTS

6. REFERENCES

Aydin, C., & Demirel, H. (2004). Computation of Baarda's lower bound of the non-centrality parameter. *Journal of Geodesy, 78*(7–8), 437–441. [Crossref]

Baarda, W. (1968). A testing procedijre for use in geodetic networks. *Netherlands Geodetic Commission, 2(*5). [Crossref]

Berber, M., & Hekimoglu, S. (2003). What is the reliability of conventional outlier detection and robust estimation in trilateration networks? *Survey Review, 37*(290), 308–318. [Crossref]

Erdogan, B. (2014). An outlier detection method in geodetic networks based on the original observations. *Boletim de Cincias Geodsicas, 20*(3), 578–589. [Crossref]

Förstner, W. (1983). Reliability and discernability of extended Gauss-Markov models. Deut. Geodact. Komm. Seminar on Math. *Models of Geodetic Photogrammetric Point Determination with Regard to Outliers and Systematic Errors* p 79-104 (SEE N 84-26069 16-43).

Ghilani, C. D. (2017). *Adjustment computations: spatial data analysis*. John Wiley & Sons.

Hekimoglu, S., & Koch, K. R. (1999). How can reliability of the robust methods be measured? *Proceedings of the Third Turkish-German Joint Geodetic Days, Istanbul*, 179–196.

Huber, P. J. (1992). Robust Estimation of a Location Parameter. *In Springer Series in Statistics* (pp. 492–518). [Crossref]

Kavouras, M. (1982). *On the Detection of Outliers and the Determination of Reliability in Geodetic Networks*. 1982. m. sc. e. Thesis-Department of Geodesy and Geomatics Engineering, University of New Brunswick.

Klein, I, Matsuoka, M. T., Guzatto, M. P., de Souza, S. F., & Veronez, M. R. (2014). On evaluation of different methods for quality control of correlated observations. *Survey Review, 47*(340), 28–35. [Crossref]

Klein, I, Matsuoka, M. T., Guzatto, M. P., & Nievinski, F. G. (2016). An approach to identify multiple outliers based on sequential likelihood ratio tests. *Survey Review, 49*(357), 449–457. [Crossref]

Klein, Ivandro, Matsuoka, M. T., & Guzzato, M. P. (2015). How to estimate the minimum power of the test and bound values for the confidence interval of data snooping procedure. *Boletim de Cincias Geodsicas, 21*(1), 26–42. [Crossref]

Klein, Ivandro, Matsuoka, M. T., Souza, S. F. de, & Collischonn, C. (2012). Design of geodetic networks reliable against multiple outliers. *Boletim de Ciências Geodésicas, 18*(3), 480–507.

Klein, Ivandro, Matsuoka, M. T., Souza, S. F. de, & Veronez, M. R. (2011). Adjustment of observations: a geometric interpretation for the least squares method. *Boletim de Ciências Geodésicas, 17*(2), 272–294.

Koch, K.-R. (1999). Parameter Estimation and Hypothesis Testing in Linear Models. [Crossref]

Lehmann, Rdiger, & Scheffler, T. (2011). Monte Carlo-based data snooping with application to a geodetic network. *Journal of Applied Geodesy, 5*(3–4). [Crossref]

Lehmann, Rüdiger. (2012). Improved critical values for extreme normalized and studentized residuals in Gauss Markov models. *Journal of Geodesy, 86*(12), 1137–1146. [Crossref]

Lehmann, Rüdiger. (2012). On the formulation of the alternative hypothesis for geodetic outlier detection. *Journal of Geodesy, 87*(4), 373–386. [Crossref]

Lehmann, Rüdiger, & Lösler, M. (2016). Multiple Outlier Detection: Hypothesis Tests versus Model Selection by Information Criteria. *Journal of Surveying Engineering, 142*(4), 4016017. [Crossref]

Lehmann, Rüdiger, & Voß-Böhme, A. (2017). On the statistical power of Baarda's outlier test and some alternative*. Journal of Geodetic Science, 7*(1), 68–78. [Crossref]

Niemeier, W., & Tengen, D. (2017). Uncertainty assessment in geodetic network adjustment by combining GUM and Monte-Carlo-simulations. *Journal of Applied Geodesy, 11*(2), 67–76. [Crossref]

Rofatto, V. F., Matsuoka, M. T., & Klein, I. (2017). An Attempt to Analyse Baarda's Iterative Data Snooping Procedure based on Monte Carlo Simulation. *South African Journal of Geomatics, 6*(3), 416. [Crossref]

Rousseeuw, P. J., & Leroy, A. M. (1987). Robust Regression and Outlier Detection. [Crossref]

Teunissen, P. J. G. (2003). Adjustment theory. VSSD.

Teunissen, P. J. G. (2006). Testing theory. VSSD.

Wilcox, R. R. (2011). Introduction to robust estimation and hypothesis testing. Academic press.

Yang, L., Wang, J., Knight, N. L., & Shen, Y. (2013). Outlier separability analysis with a multiple alternative hypotheses test. *Journal of Geodesy, 87*(6), 591–604. [Crossref]