Two aspects on L1-norm adjustment of leveling networks

Dois aspectos sobre o ajustamento pela norma L1 de redes de nivelamento

Stefano Sampaio Suraci [1]
Leonardo Castro de Oliveira [1]
Ivandro Klein [2]

ABSTRACT

L1-norm adjustment corresponds to the minimization of the sum of weighted absolute residuals. Unlike Least Squares, it is a robust estimator, i.e., insensitive to outliers. In geodetic networks, the main application of L1-norm refers to the identification of outliers. There is no general analytical expression for its solution. Linear programming is the usual strategy, but it demands decorrelated observations. In the context of Least Squares, it is well known that the application of Cholesky factorization decorrelates observations without changing the results of the adjustment. However, there is no mathematical proof that this is valid for L1-norm. Besides that, another aspect on L1-norm is that equal weights may guarantee maximum robustness in practice. Therefore, it is expected to also provide a better effectiveness in the identification of outliers. This work presents contributions on two aspects concerning L1-norm adjustment of leveling networks, being them: the validity of Cholesky factorization for decorrelation of observations and the effectiveness for identification of outliers of a stochastic model with equal weights for observations. Two experiments were conducted in leveling networks simulated by the Monte Carlo method. In the first one, results indicate that the application of the factorization as previously performed in the literature seems inappropriate and needs further investigation. In the second experiment, comparisons were made between L1 with equal weights and L1 with weights proportional to the inverse of the length of the leveling line. Results show that the first approach was more effective for the identification of outliers. Therefore, it is an interesting alternative for the stochastic model in L1-norm adjustment. Besides providing a better performance in the identification of outliers, the need for observation decorrelation becomes irrelevant if equal weights are adopted.

[1]Seção de Engenharia Cartográfica. Instituto Militar de Engenharia (IME), Rio de Janeiro, Brasil. E-mail: stefanosampaio@hotmail.com; leonardo@ime.eb.br
[2]Departamento Acadêmico de Construção Civil. Curso Técnico de Agrimensura. Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina (IFSC), Santa Catarina, Brasil. E-mail: ivandro.klein@ifsc.edu.br

RESUMO

O ajustamento pela norma L1 corresponde à minimização da soma dos resíduos ponderados. Diferentemente do Método dos Mínimos Quadrados, ele é robusto, ou seja, insensível a *outliers*. Em redes geodésicas, a principal aplicação da norma L1 refere-se à identificação de *outliers*. Não há uma expressão analítica geral para sua solução. Programação Linear é a estratégia usual, mas isso demanda observações decorrelacionadas. No contexto dos Mínimos Quadrados, sabe-se que a aplicação da fatoração de Cholesky decorrelaciona observações sem modificar os resultados do ajustamento. Porém, não há demonstração matemática de que o mesmo vale para a norma L1. Além disso, outro aspecto sobre a norma L1 é que pesos iguais podem garantir a máxima robustez na prática. Assim, se espera que também provenha maior efetividade na identificação de *outliers*. Este trabalho apresentou contribuições com relação a dois aspectos do ajuste pela norma L1 de redes de nivelamento: a validade da fatoração de Cholesky para decorrelacionamento de observações e a efetividade de identificação de *outliers* com um modelo estocástico de pesos iguais para as observações. Dois experimentos foram conduzidos em redes de nivelamento simuladas pelo método de Monte Carlo. No primeiro, verificamos que a aplicação da fatoração como previamente realizada na literatura parece inadequada e necessita de maior investigação. No segundo experimento, comparações foram feitas entre ajustamentos pela norma L1 com pesos iguais e com pesos inversamente proporcionais ao comprimento das linhas de nivelamento. Os resultados mostram que a primeira abordagem foi mais efetiva para identificação de *outliers*. Assim, essa é uma alternativa interessante para o modelo estocástico no ajuste da norma L1. Além de proporcionar um melhor desempenho na identificação de *outliers*, a necessidade de decorrelacionamento das observações torna-se irrelevante se pesos iguais forem adotados.

<p style="text-align:center">* * *</p>

Introduction

It is well known that the Least Squares (LS) adjustment causes incorrect results when outliers are present in an observation set (ROFATTO, MATSUOKA and KLEIN, 2017). Because the L1-Norm Minimization (L1NM) is a robust method, resistant to outliers, several authors have already tried it in the adjustment of geodetic networks, such as Marshal and Bethel (1996), Amiri-Simkooei (2003), and Yetkin and Inal

(2011). The main application of L1NM corresponds to outlier identification in the network, for later adjustment by LS.

In linear adjustment computations, Equation 1 expresses the functional model, in which $\mathbf{A_{mxn}}$ is the coefficient matrix of the $\mathbf{x_{nx1}}$ vector of unknowns, $\mathbf{L_{mx1}}$ is the vector of the observed values, $\mathbf{v_{mx1}}$ is the residual vector, $\mathbf{m}$ is the number of observations, and $\mathbf{n}$ is the number of unknowns. For the stochastic model, it is usual to assume that the *a priori* variance factor $\mathbf{\sigma_0^2}$ is unitary. Thus, the matrix of weights $\mathbf{P_{mxm}}$ is given by the inverse of the covariance matrix of the observations $\mathbf{\Sigma_{mxm}}$ (Equation 2). Different ways of adding datum information to the system of equations are shown in Ghilani (2010).

$$Ax = L + v \tag{1}$$

$$P = \sigma_0^2 * \Sigma^{-1} = \Sigma^{-1} \tag{2}$$

Under the assumption that observations follow the normal distribution, the most probable solution of the adjustment is usually obtained by the LS method, also referred to as L2-norm minimization (AMIRI-SIMKOOEI, 2003), which minimizes the sum of the squared elements of $\mathbf{v}$ weighted by $\mathbf{P}$ (Equation 3). As a consequence, it can be demonstrated that$\mathbf{x}$ is given by Equation 4. Further information on LS adjustment of geodetic networks can be found in Ghilani (2010).

$$LS: min(v^T P v) \tag{3}$$

$$x = (A^T P A)^{-1} A^T P L \tag{4}$$

Independent observations have no correlation with each other (GEMAEL, MACHADO and WANDRESEN, 2015). Thus, the matrix of weights $\mathbf{P_{mxm}}$ is diagonal. For the specific case of leveling networks, it is usual to

adopt weights of the observations as inversely proportional to the length of the respective leveling lines. For mathematical proof of this approach to the stochastic model we also recommend Ghilani (2010).

On the other hand, L1NM corresponds to the minimization of the sum of the absolute residuals weighted by the vector of weights of observations **p**mx1 (Equation 5), comprised of the elements of the main diagonal of **P** (AMIRI-SIMKOOEI, 2003). Therefore, since Equation 5 does not take correlations into account, observations should be independent (YETKIN and INAL, 2011).

$$L1NM: min\ (p^T|v|) \tag{5}$$

Unlike LS, there is no general analytical expression for the solution of L1NM. Linear programming is the usual strategy for L1NM (AMIRI-SIMKOOEI, 2018). Based on the basic concepts of Marshal and Bethel (1996), Amiri-Simkooei (2003) has presented a detailed L1NM formulation for implementation by the *simplex* method of linear programming. Readers interested in linear programming and *simplex* method should refer to Dantzig (1963).

Other approaches for L1NM solution can be found in Koch (1999), Junhuan (2005), Baselga (2007), Khodabandeh and Amiri-Simkooei (2011), and Amiri-Simkooei (2018). On the whole, it is essential to highlight that it is common in L1NM to consider the premise of independent observations in the network. Therefore, with **P** being diagonal, the obtainment of vector **p** is trivial, which allows the application of Equation 5.

1 Observation decorrelation and equal weights in L1NM

What if observations are not independent? Although not in the L1NM context, this issue has already been addressed in geodetic literature by Yang (1994), Yang, Song and Xu (2002), and Klein et al. (2015), for example.

Regarding L1NM, Yetkin and Inal (2011) have brought an attempt to circumvent this limitation. They have adopted the exact same formulation as Amiri-Simkooei (2003), with only one more detail: in the presence of correlation between observations of the network (**P** not diagonal), the authors have applied a solution for the observation decorrelation via Cholesky factorization of the matrix **P**. As a consequence, the authors have indicated that the "new" matrix **P'** is not only diagonal but also equal to the identity matrix **I** of the same dimensions; that is, all weights become unitary.

However, we emphasize that Strang and Borre (1997), who were mentioned by Yetkin and Inal (2011) on this issue, have demonstrated Cholesky factorization non-influence only in the results of the LS adjustment. Despite this, Cholesky factorization was applied in L1NM (not LS) by the latter authors.

Let **W** be the lower triangular matrix computed in the Cholesky factorization of **P** (P=W^{T}W). In the context of LS adjustment, adopting **P'=I**, transformed matrix of coefficients **A'=WA**, and transformed vector of observed values **L'=WL**, the vector **x** of unknowns remains unchanged in relation to the original system (Equation 6). Thus, it becomes a system with the same solution for the unknowns of the original but now with uncorrelated (independent) observations. The result for the residual vector **v'=Wv** is also related to the residual vector **v** of the original system.

$$x = ((WA)^T IWA)^{-1}(WA)^T IWL = (A^T W^T IWA)^{-1} A^T W^T IWL = (A^T PA)^{-1} A^T PL \quad (6)$$

Formal proof of Cholesky factorization validity in the context of L1NM was not found either in Strang and Borre (1997) or in any other work in the related literature. As it seems inconsistent, a contribution of our work is the investigation of the consequences of using Cholesky factorization for L1NM results.

In addition, another point seems to deserve further investigation about the stochastic model in L1NM: the potential effectiveness of an adjustment with equal weights for observations in the identification of outliers. The numerical results of Kampmann and Krause (1997) suggest that equal weights provide the maximum robustness in practice for L1NM. This issue was also addressed in our experiments.

In the experiments, we simulated leveling networks using Monte Carlo Simulation. It is a type of simulation that is based on repeated random samples of the input variables, with known probability distribution, to characterize the variation of the results of the model to be studied (RAYCHAUDHURI, 2008). The greater the number of random samples simulated, the more accurate the characterization tends to be. Its application is usual in many fields. It has also become common in the simulation of geodetic networks, as performed by Hekimoglu and Erenoglu (2007) and Rofatto et al. (2018), for example. Since it has never been done before in geodetic literature, the analysis of L1NM with Monte Carlo Simulation is also a contribution of our work.

2 Methodology

For the investigation of the elements discussed, two experiments were conducted in leveling network scenarios with independent observations. The first experiment (Experiment 1) aimed to evaluate the validity of the application of the Cholesky factorization for observation decorrelation for L1NM in the terms of Yetkin and Inal (2011). The solution of this adjustment was compared with and without decorrelation for the simulated scenarios. Since the solution for $\mathbf{x}$ (and $\mathbf{v}$) may not be unique in L1NM (ABDELMALEK and MALEK, 2008), we compared the L1 norm itself of the vector of absolute weighted residuals, here called "L1 norm" for simplicity (equation 7), which must be equivalent for equally correct solutions.

$$L1\ norm = p^T |v| \tag{7}$$

It is important to highlight that Yetkin and Inal (2011) dealt with correlated observations (non-diagonal **P** matrix) in their experiments. However, if we had done this, we would not have been able to verify if Cholesky factorization changed (or not) the L1NM results, because covariances cannot be insert in L1NM by the formulation of Equation 5. Thus, we used a network with a diagonal **P** matrix. Therefore, we could solve L1NM without applying Cholesky factorization (value considered correct for reference), as well as with Cholesky factorization, to check if there is (or not) equality between these adjustments.

The second experiment (Experiment 2) aimed to investigate whether L1NM with equal weights for observations is more effective than with usual weights of the observations as inversely proportional to the length of the respective leveling line in the identification of outliers. We considered outliers the observations with a L1NM residual higher than **3*σ$_i$**, in which **σ$_i$** is the standard deviation of the respective *i-th* observation. Probability of correct identification (PCI) and the probabilities of Type II error, Type III error, Over-identification+, and Over-identification- were computed by adopting both stochastic models for comparison. A description of these indices is presented in Table 1.

Table 1 – PCI and probabilities of errors.

Probability	Description
PCI	probability of correctly identifying the outlying observation with no other observations
Type II error	probability of non-identifying an outlier when there is at least one outlier
Type III error	probability of misidentifying a non-outlying observation as an outlier, instead of the outlying observation
Over-identification+	probability of correctly identifying the outlying observation and other observations
Over-identification-	probability of identifying more than one non-outlying observation, whereas the "true outlier" remains in the dataset

Source: Adapted from Rofatto, Matsuoka and Klein (2017) and Rofatto et al. (2018).

In order to evaluate the higher robustness of L1NM with equal weights, we also compared the "rate of null residual" and the "rate of highest residual" for the outlier. The occurrence of null residuals is common in L1NM (ABDELMALEK and MALEK, 2008). If it happens with the outlier, we can infer that its identification will not be viable, as it will be completely masked in the adjustment. This may indicate that the adjustment was not robust, not resistant to the outlier. In this context, it is expected that the outlier will have the highest residual in a robust adjustment.

The experiments were conducted using the software *Octave*, version 4.4.1. The solution of L1NM was performed by *simplex* method of linear programming, using the *glpk* routine of the same software. The reader can contact the authors to obtain the *Octave* codes. The experiments are unprecedented in geodesic literature, justifying the scientific relevance of this work.

2.1 Monte Carlo simulations

For the experiments, leveling network scenarios with independent observations were simulated by the Monte Carlo method. "Observations without outliers" and "bad observations" (outliers) with random sign for the scenarios were obtained as in the heteroscedasticity case (different weights for observations) of Hekimoglu and Erenoglu (2007).

Following the strategy of these authors, the simulated leveling network used as the basis for the scenarios by the Monte Carlo method is shown in Figure 1 and Table 2. The observations **Li** were computed directly from the heights of the vertices of the network; that is, they represent the exact elevation difference between them, without any measurement errors.

Intervals of magnitude of the generated outliers (from 3σ to 6σ, and from 6σ to 12σ, in which **σ** is the standard deviation of the respective observation) used in the reference cited were also used in this work. In

addition, one more interval (from 12σ to 25σ) for magnitudes of the generated outliers was added. For each experiment and each interval of magnitude of outliers, we simulated 200,000 scenarios with one outlier. A result accuracy analysis as a function of the number of simulations was performed by Rofatto et al. (2018).

Figure 1 – Simulated leveling network
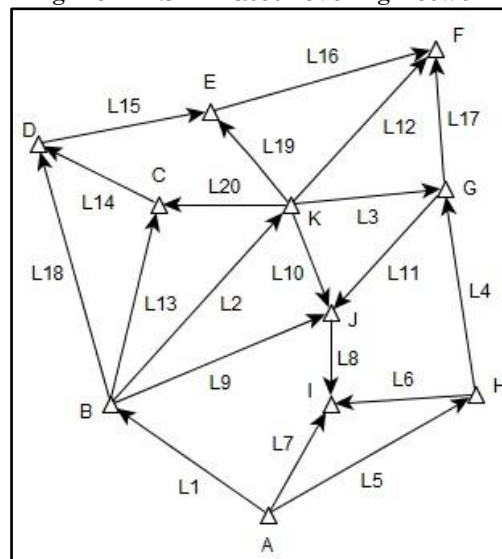
Source: Prepared by the authors.

Table 2– Simulated network – elevation differences.

Li	Elevation difference (mm)	Distance (km)	Li	Elevation difference (mm)	Distance (km)
L1	163854.9	49	L11	110227.2	62
L2	6446.2	41	L12	155928.2	50
L3	57037.0	38	L13	52875.0	35
L4	126209.5	34	L14	62904.2	43
L5	101128.6	22	L15	3889.5	20
L6	296885.8	13	L16	42705.7	28
L7	398014.4	23	L17	98891.2	19
L8	60449.1	48	L18	115779.2	39
L9	173710.4	15	L19	113222.5	27
L10	167264.2	24	L20	46428.8	21

Source: Prepared by the authors.

The standard deviation of the observations adopted in the simulations is given by $\sigma_i = 1.0 (mm) * \sqrt{K_i}$, where **K** (in km) is the length of the respective leveling section. In each simulated scenario adjustment, the

height of point A was considered fixed and with value $h_A=0$. Thus, all Monte Carlo scenarios contained 20 observations and 10 points of heights to be determined (unknowns), so the number of degrees of freedom was 10.

3 Results and Discussion

3.1   Experiment 1

At first, all simulated scenarios were adjusted by LS with the usual stochastic model of weights of the observations as inversely proportional to the length of the respective leveling lines. We adjusted each scenario with the application of Cholesky factorization for observation decorrelation and without decorrelation. Adjustment results in each scenario were always the same (as expected) for the "L1 norm" (Equation 7), considering the tolerance of $10^{-10}mm^{-1}$ ($mm^{-1} = 1/mm$ is the unit of "L1 norm").

Then, all adjustments with and without the application of the Cholesky factorization to decorrelate observations were also performed by L1NM. Table 3 shows the percentage of scenarios in which the results for "L1 norm" (Equation 7) were equal, that is, the decorrelation succeed, considering different levels of tolerance. In order to highlight differences between L1NM and LS, Table 3 also presents the aforementioned LS results.

Table 3– Success rate of decorrelation – Experiment 1.

Adjustment method	Tolerance (mm^{-1})	Magnitude of outliers		
		3-6σ	6-12σ	12-25σ
L1NM	10^{-1}	83.46%	80.48%	78.54%
L1NM	10^{-10}	21.68%	20.96%	20.89%
LS	10^{-10}	100.00%	100.00%	100.00%

Source: Prepared by the authors.

Despite the equality of results with and without Cholesky factorization in 100% of LS adjustments, it happened in only approximately

20% of the cases for L1NM, considering the tolerance of 10^{-10}mm^{-1}. Even considering the tolerance of 10^{-1}mm^{-1}, equality for results in L1NM was not always achieved. Therefore, we verified that coherence between the results of L1NM with and without decorrelation through Cholesky factorization in the terms of Yetkin and Inal (2011) is not guaranteed.

Moreover, we found that the higher the magnitude of outliers in the data, the greater the probability of non-coherence. Thus, although previously performed in the literature, we verified that the application of the factorization in this context seems inappropriate and still needs further investigation.

In addition, Experiment 1 further corroborated the theory on the subject, in the case of decorrelation for LS adjustment, for which the results with and without Cholesky factorization were always consistent.

3.2   Experiment 2

Table 4 shows the PCI and the probabilities of Type II error, Type III error, Over-identification+, and Over-identification- for 200,000 scenarios simulated with one outlier for each different interval of magnitude of the outlier.

Table 4– PCI and probabilities of errors – Experiment 2.

Outlier	Weights	PCI	Type II	Type III	Over+	Over-
3-6σ	Equal	44.89%	27.70%	11.91%	12.85%	2.68%
	Usual	44.16%	28.97%	13.05%	9.91%	3.90%
6-12σ	Equal	70.53%	1.23%	1.19%	26.58%	0.47%
	Usual	66.89%	1.71%	2.91%	19.94%	8.55%
12-25σ	Equal	71.76%	0.00%	0.00%	28.24%	0.00%
	Usual	68.17%	0.00%	0.01%	21.83%	9.99%

Source: Prepared by the authors.

In Experiment 2, PCI was always higher in the adjustment by L1NM with equal weights, for all magnitude ranges of the inserted outlier. Therefore, this approach for the stochastic model was more effective for

L1NM in the identification of outliers than the usual one. Both stochastic models tended to provide higher PCI with the increase of the magnitude of outliers in the data.

The probabilities of Type II and Type III errors also "improved" with equal weights, as their values were lower. Only Type II errors were equally zero for 12-25σ outliers. In addition, they decreased as the magnitude of the outlier increased for both stochastic models. Probabilities of Over- were advantageous (lower) with equal weights as well. Moreover, they tended to decrease for higher magnitudes of outliers, while the usual weight ones tended to increase.

Probabilities of Over+ were always higher with equal weights, which was its only disadvantage. However, it is important to remember that, although among false positives, Over+ is the only error in which an existing outlier is at least identified (the main point in ever applying L1NM). Therefore, it is better to have Over+ than Over-. If we add the PCI and Over+ for 12-25σ outliers, for example, we can see that all outliers were identified with equal weights, while 90% were identified with usual weights, even though the difference of PCI was 3.59% (significantly less than 10%).

Finally, Table 5 presents the "rate of null residual" and the "rate of highest residual" for the outlier. Equal weights always provided less of the first, and more of the second, which indicates the higher resistance to the outlier in relation to usual weights in this experiment. Both stochastic models tended to provide a lower "rate of null residual" and a higher "rate of highest residual" with the increase of the magnitude of outliers in the data.

Table 5 – Rate of null and highest residual of the outlier – Experiment 2.

Outlier	Weights	Null residual	Highest residual
3-6σ	Equal	5.40%	63.92%
	Usual	15.50%	58.22%
6-12σ	Equal	0.15%	95.17%
	Usual	10.33%	84.29%
12-25σ	Equal	0.00%	99.94%
	Usual	9.97%	89.82%

Source: Prepared by the authors.

4 Conclusion

This work presented contributions regarding two points concerning L1-norm adjustment: the validity of the application of Cholesky factorization for observation decorrelation and the effectiveness in identifying outliers of a stochastic model with equal weights for observations. Two experiments were performed in leveling network scenarios with independent observations, simulated by the Monte Carlo method.

Regarding Cholesky factorization, although its application for observation decorrelation is appropriate in LS adjustment, it presented at least alarming results for L1NM. This indicates that the application of factorization in the latter context, as previously performed in the literature, seems inappropriate and still needs further investigation. Therefore, this result suggests that other methods for decorrelation should be tested, and that the formulation of Cholesky factorization for application in L1NM should be re-evaluated.

However, we also verified that L1NM with equal weights may be an interesting alternative, because of the advantage seen in relation to the usual stochastic model. Besides providing a better performance in the identification of outliers, the need for observation decorrelation becomes irrelevant if equal weights are adopted. Therefore, future work should be developed in order to obtain a better understanding of the influence of the stochastic model for the identification of outliers by L1NM.

Finally, in this work we considered the absolute residual over 3σ as the objective criterion for the identification of outliers in L1NM. Although this limit is common in LS approach, we believe further investigations must be conducted in order to analyze possible objective criteria specific for the results of L1NM.

Authors' Contributions

The experiments were designed by S.S.S and L.C.O; S.S.S. performed the experiments; S.S.S and L.C.O analyzed and discussed the numerical results; S.S.S. drafted the article; I.K. critically revised the article; All authors approved the manuscript.

References

ABDELMALEK, N.; MALEK, W. **Numerical linear approximation in C**. London: CRC Press, 2008, 928p.

AMIRI-SIMKOOEI, A. Formulation of L1 Norm Minimization in Gauss-Markov Models. **Journal of Surveying Engineering**, vol. 129, n. 1,2003, pp.37-43.

AMIRI-SIMKOOEI, A. On the use of two L1 norm minimization methods in geodetic networks. **Earth Observation and Geomatics Engineering**, vol. 2, n. 1, 2018, pp.1-8.

BASELGA, S. Global Optimization Solution of Robust Estimation. **Journal of Surveying Engineering**, vol. 133, n. 3, 2007, pp.123-128.

DANTZIG, G. **Linear Programming and Extensions**. Princeton: Princeton University Press, 1963.

GEMAEL, C.; MACHADO, A. M. L.; WANDRESEN, R. **Introdução ao ajustamento de observações: aplicações geodésicas**. 2. ed. Curitiba: Ed. UFPR, 2015. 428p.

GHILANI, C. **Adjustment Computations: Spatial Data Analysis**. 5. ed. Hoboken: John Wiley & Sons, Inc., 2010, 647p.

HEKIMOGLU, S.; ERENOGLU, R. Effect of heteroscedasticity and heterogeneousness on outlier detection for geodetic networks. **Journal of Geodesy**, vol. 81, 2007, pp.137–148.

JUNHUAN, P. The asymptotic variance–covariance matrix, Baarda test and the reliability of L1-norm estimates. **Journal of Geodesy**, vol. 78, 2005, pp.668-682.

KAMPMANN, G.; KRAUSE, B. A breakdown point analysis for the straight line fit based on balanced observations. **Bollettino di Geodesia e ScienzeAffini**, n. 3, 1997, pp. 293-303.

KHODABANDEH, A.; AMIRI-SIMKOOEI, A. Recursive Algorithm for L1 Norm Estimationin Linear Models. **Journal of Surveying Engineering**, vol. 137, n. 1, 2011, pp.1-8.

KLEIN, I. *et al.* On evaluation of different methods for quality control of correlated observations. **Survey Review**, vol. 47, n. 340, 2015, pp. 28-35.

KOCH, K. **Parameter Estimation and Hypothesis Testing in Linear Models**. 2. ed. Berlin: Springer, 1999, 333p.

MARSHALL, J.; BETHEL, J. Basic concepts of L1 norm minimization for surveying applications. **Journal of Surveying Engineering**, vol. 122, n. 4, 1996, pp. 168-179.

RAYCHAUDHURI, S. Introduction to Monte Carlo Simulation. **Proceedings of the 2008 Winter Simulation Conference**, Miami, 2008. pp. 91-100.

ROFATTO, V. F.; MATSUOKA, M. T.; KLEIN, I. An Attempt to Analyse Baarda's Iterative Data Snooping Procedure based on Monte Carlo Simulation. **South African Journal of Geomatics**, vol. 6, n. 3, 2017, pp. 416-435.

ROFATTO, V. F. *et al.* A half-century of Baarda's concept of reliability: a review, new perspectives, and applications. **Survey Review**, 2018, DOI: 10.1080/00396265.2018.1548118.

STRANG, G.; BORRE, K. **Linear algebra, geodesy, and GPS.** Wellesley: Cambridge Press, 1997, 642p.

YANG, Y. Robust estimation for dependent observations. **Manuscripta Geodaetica**, vol. 19, 1994, pp. 10–17.

YANG, Y.; SONG, L.; and XU, T. Robust estimation for correlated observations based on bifactor equivalent weights. **Journal of Geodesy**, vol. 76, n. 6-7, 2002, pp. 353–358.

YETKIN, M.; INAL, C. L1 norm minimization in GPS networks, **Survey Review**, vol. 43, n. 323, 2011, pp. 523–532.