

AVALIAÇÃO DO MÉTODO DATA SNOOPING POR SIMULAÇÕES MONTE CARLO

ALEXANDRE FLEURY DE SIQUEIRA E GOMES¹

MARCELO TOMIO MATSUOKA¹

VINÍCIUS FRANCISCO ROFATTO¹

IVANDO KLEIN²

RESUMO:

O teste Data Snooping (DS) é um teste estatístico utilizado para detecção de um erro grosseiro por iteração, sendo o método mais estabelecido para identificação de outliers em redes geodésicas. Desta forma, para avaliar esse método, duas redes de nivelamento altímetro numericamente simuladas foram contaminadas com outliers entre 3σ a 12σ de magnitude, positivos ou negativos, gerados randomicamente e adicionados as observações da rede de nivelamento que contém observações independentes. Para cada rede, 4000 cenários são realizados, com diferentes intervalos de magnitude de erros grosseiros, utilizando simulações Monte Carlo.

Palavras-chave: Data Snooping, Monte Carlo, Rede de nivelamento

1 INTRODUÇÃO

O critério mais adotado nas ciências geodésicas é o método dos mínimos quadrados (MMQ), ou seja, aquela solução para os parâmetros incógnitos que minimiza a soma do quadrado dos resíduos, ponderados pelos respectivos pesos das observações (GHILANI, 2010). Além disso, quando a ponderação das observações é tomada em função de suas variâncias e covariâncias, corretamente determinadas, o MMQ é dito o melhor estimador linear imparcial, e ainda, quando os erros aleatórios seguem distribuição normal multivariada, a solução pelo MMQ coincide com a solução de máxima verossimilhança (TEUNISSEN, 2003).

Entretanto, o MMQ é sensível à presença de erros não aleatórios (*outliers*) nas observações, ou seja, estes podem exercer influência na solução do ajustamento (ou seja, nos

1 UNIVERSIDADE FEDERAL DE UBERLÂNDIA, ENGENHARIA DE AGRIMENSURA E CARTOGRÁFICA, alexandrefleury.sg@hotmail.com¹; tomio@ufu.br²; vinicius.rofatto@ufu.br³

2 INSTITUTO FEDERAL DE SANTA CATARINA, CURSO TÉCNICO DE AGRIMENSURA, ivandroklein@gmail.com

parâmetros estimados), e assim as propriedades do MMQ não são satisfeitas (GHILANI, 2010).

Uma forma de contornar esse problema do MMQ é adotar os procedimentos convencionais de testes de detecção de *outliers* (por exemplo, Baarda, 1968; Pope, 1976; Kok, 1984). Dentro do contexto do MMQ um teste estatístico para identificação de outliers é baseado na hipótese de um único outlier ou de múltiplos outliers dentro das observações. O teste Data Snooping (DS) é um teste estatístico utilizado para detecção de um erro grosseiro por iteração, sendo o método mais estabelecido para identificação de outliers em Geodésia (LEHMANN, 2012).

Em um teste estatístico o espaço amostral, que consiste no conjunto de todos os resultados de um experimento, é dividido em duas partes sendo a região de aceitação e a região de rejeição da hipótese nula. Se a estatística de teste calculada em função da amostra se situar dentro da região crítica, a hipótese nula é rejeitada, caso contrário, não existem evidências para rejeitar a hipótese em questão (LARSON, 1974).

Dentro do escopo de testes estatísticos existem dois tipos de erros, sendo o primeiro denominado erro do tipo I que consiste em rejeitar a hipótese quando esta é verdadeira e a segunda analogamente chamada de erro do tipo II que se resume em não rejeitar a hipótese quando está de fato é falsa.

Segundo Larson (1974), a probabilidade de cometer o erro do tipo I é dada por α (nível de significância do teste) e a probabilidade de cometer o erro do tipo II é dada por β . Adaptando o procedimento do teste estatístico para o DS, a hipótese nula (H_0) é a ausência de erros grosseiros nas observações e para cada i -ésima observação do conjunto de dados se define apenas uma hipótese alternativa (H_i), com a presença de erro grosseiro, ou seja, o DS considera somente uma observação suspeita de estar contaminada por erro grosseiro por vez. O erro tipo I refere-se a rejeição de H_0 quando esta é verdadeira e o erro tipo II refere-se a não rejeição de H_0 quando esta é falsa.

Como o DS possui “ n ” hipóteses alternativas, pois cada observação é testada individualmente, tem-se ainda o erro do tipo III que consiste em identificar uma observação sendo que o erro grosseiro está presente em outra, ou seja, aceitar uma hipótese alternativa errada ao invés da hipótese alternativa correta (Förstner, 1983).

Dentro deste contexto, o objetivo deste trabalho é analisar a confiabilidade de uma rede de nivelamento simulada, analisando a detecção de erros grosseiros por meio do procedimento DS, utilizando simulações numéricas por meio do método Monte Carlo.

2 REVISÃO BIBLIOGRÁFICA

Proposto inicialmente por Baarda (1968), o DS assume que somente uma observação está contaminada por outlier por iteração. Portanto, na prática, o ajustamento pelo MMQ e o procedimento de teste continua até que não existam mais outliers identificados. No caso do procedimento DS se considera o seguinte modelo linear de ajustamento:

$$Lb = (Ax + e) \quad (1)$$

Na expressão 1, Lb é o vetor das n observações, x é o vetor dos u parâmetros incógnitos do modelo (no caso de redes de nivelamento, as altitudes/cotas dos vértices), A é a matriz *design* do ajustamento (ou matriz jacobiana), e e é o vetor dos n erros aleatórios das observações.

Para observações correlacionadas, ou seja, observações dependentes, para cada observação (i) é calculada a seguinte estatística de teste (TEUNISSEN, 2006)

$$Wi = \frac{c_i^T P v}{\sqrt{c_i^T P \Sigma v P c_i}}; |w_i| > c_{\alpha/2} \quad (2)$$

Na Expressão 2, P é a matriz peso do ajustamento, v é o vetor dos resíduos ajustados, Σv é a matriz de covariância dos resíduos ajustados, e $c_{\alpha/2}$ é o valor crítico teórico na distribuição normal padrão, para um dado nível de significância α (teste bilateral).

Para observações não correlacionadas, ou seja, observações independentes, é utilizado uma equação simplificada para a estatística de teste de cada observação, dada por (KLEIN, 2012)

$$Wi = Vi / \sigma_{Vi} \quad (3)$$

Na Expressão 3, Vi é a i -ésima componente do vetor dos resíduos ajustados e σ_{Vi} é a i -ésima componente do desvio padrão dos resíduos ajustados. Como cada observação é testada individualmente, a observação considerada eivada de erro grosseiro será aquela

cujas estatísticas de teste W_i satisfazer as seguintes condições (BAARDA,1968; TEUNISSEN,2006):

$$\begin{cases} |W_i| > C_{\alpha/2} \\ |W_i| > |W_j|, \forall i \neq j \end{cases} \quad (4)$$

Ao se identificar a observação suspeita de estar contaminada por outlier conforme o nível de significância estipulado, a mesma é removida do conjunto de dados e se repete o ajustamento e o teste DS até que todas as observações suspeitas estejam identificadas e removidas, se caracterizando um processo iterativo de ajustamento, identificação e remoção de erros grosseiros.

3 RESULTADOS E EXPERIMENTOS

Neste estudo é analisado um exemplo numérico envolvendo uma rede de nivelamento simulada com um ponto fixado e quatro pontos de altitude desconhecida ($u=4$), totalizando cinco pontos pertencentes a rede. Para cada ponto de altitude desconhecida existe no mínimo três desníveis ligados ao mesmo, totalizando assim 8 desníveis ($n = 8$), contabilizando $n-u = 8-4=4$ observações redundantes nesta simulação. Todas as observações possuem uma precisão nominal de $\sigma = \pm 0,02$ m e as observações são consideradas independentes (ver a Figura 1).

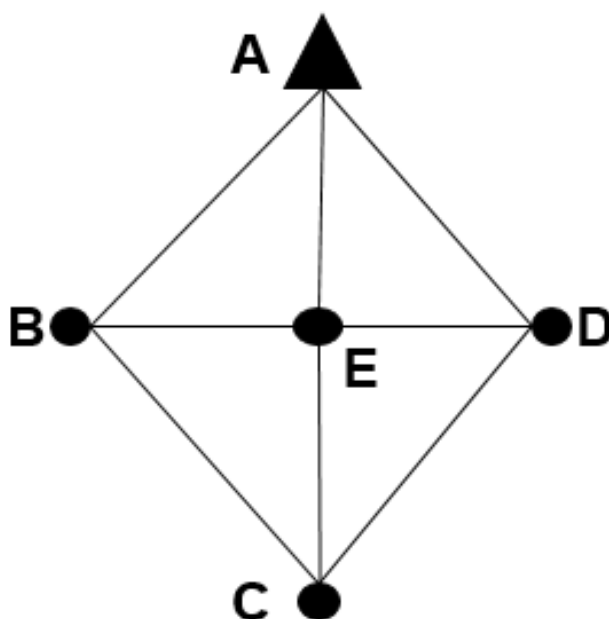


FIGURA 1 – Rede de nivelamento A.

Na rede numericamente simulada, foi considerado a cota do vértice A sendo fixa (injunção absoluta) e igual a 100 m, podendo assim ser considerada isenta de erros. Posteriormente foram geradas oito observações (desníveis), considerando seus respectivos desvios padrões, em seguida foi sorteado utilizando a distribuição uniforme uma observação para ser contaminada com um erro grosseiro de magnitude entre 3σ a 5σ , 5σ a 7σ , 7σ a 9σ e 9σ a 12σ , (σ é o respectivo desvio padrão da observação), em cada experimento, somente após a inserção dos erros aleatórios e grosseiros foi realizada o ajustamento das observações utilizando o MMQ.

Para melhor entendimento do procedimento adotado nos experimentos feitos com a rede de nivelamento, quatro cenários foram utilizados para testar a rede A, a diferença entre os cenários é a magnitude do erro grosseiro, sendo cada cenário testado mil vezes, totalizando assim quatro mil simulações realizadas pelo método Monte Carlo.

A Tabela 1 mostra a taxa de sucesso do DS, que corresponde as frações dos outlier corretamente identificados, bem como as taxas de falhas de identificação do outlier. As taxas de falhas de identificação são divididas em duas classes: detectou errado, que significa que o teste identificou o outlier na observação errada (erro tipo III), ou detectou nada, significando que o teste não identificou nenhum outlier nas observações (erro tipo II).

TABELA 1 – Resultados da rede de nivelamento A

REDE A	Sucesso	Detectou errado	Detectou nada
3σ a 5σ	35,6%	2,6%	61,8%
5σ a 7σ	76,9%	3,7%	19,4%
7σ a 9σ	96,9%	0,8%	2,3%
9σ a 12σ	99,5%	0,5%	0%

Ao se fazer uma análise dos resultados da Tabela 1 é possível notar que quanto maior a magnitude do outlier, maior a taxa de sucesso na detecção utilizando o teste DS. Para analisar a correlação entre redundância de observações e o aumento da taxa de sucesso foi feito uma segunda rede de nivelamento semelhante a primeira, possuindo a mesma geometria, porém, com nivelamento e contranivelamento.

A rede de nivelamento B possui um ponto fixado e quatro pontos de altitude desconhecida ($u=4$), totalizando cinco pontos pertencentes a rede. Para cada ponto de altitude desconhecida existe no mínimo seis desníveis ligados ao mesmo, totalizando assim 16 desníveis ($n = 16$), contabilizando $n-u = 16-4=12$ observações redundantes nesta simulação. Todas as observações possuem uma precisão nominal de $\sigma = \pm 0,02$ m e as observações são independentes (ver a Figura 2).

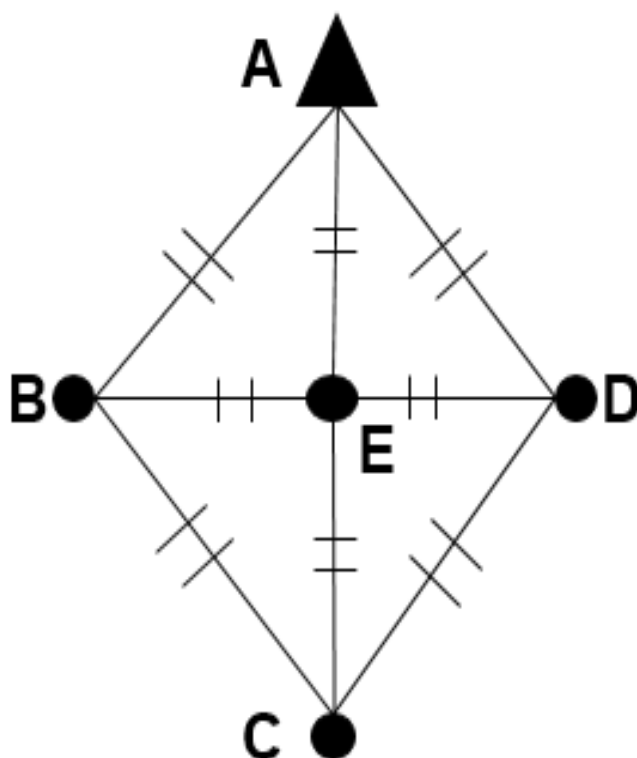


FIGURA 2 – Rede de nivelamento B.

Novos experimentos foram realizados exatamente da mesma maneira para a rede A, sendo os resultados apresentados na Tabela 2.

TABELA 2 – Resultados da rede de nivelamento B

Rede Duplicada	Sucesso	Detectou errado	Detectou nada
3 σ a 5 σ	53,9%	1,8%	44,3%
5 σ a 7 σ	93,9%	0,4%	5,7%
7 σ a 9 σ	99,9%	0,1%	0%
9 σ a 12 σ	100%	0%	0%

No primeiro cenário, com o erro grosseiro com a magnitude de 3 σ a 5 σ , a taxa de sucesso foi maior que cinquenta por cento, e nos próximos cenários foi rapidamente convergindo para cem por cento, ao se fazer um comparativo com a Tabela 1 que corresponde a rede de nivelamento A, é evidente que o aumento de observações influencia diretamente a taxa de sucesso do DS. Os aumentos nos dois primeiros cenários são da ordem de vinte por cento, sendo assim dependendo da finalidade da rede em questão, para controle de qualidade é fundamental analisar o número de observações redundantes, especialmente para o caso de erros grosseiros de menor magnitude.

Considerando todos os resultados, a média de sucesso na rede A é de 77,2% e 86,9% na rede B. Em geral o Data Snooping pode ser considerado eficiente para outliers maiores que 5σ , com uma taxa de sucesso de 76,9% para rede A e 93,9% para rede B. Para outliers menores que 5σ a taxa de sucesso decresce para 35,6% na rede A e 53,9% na rede B. No entanto, como na rede B foi simulado o nivelamento e o contranivelamento a taxa de sucesso do menor intervalo de outliers já é superior a cinquenta por cento, evidenciando a importância de realizar o contranivelamento.

4 CONSIDERAÇÕES FINAIS

O teste Data Snooping mostra uma taxa de sucesso alta nos experimentos realizados em uma rede de nivelamento simulada para um único outlier gerado uniformemente entre 5σ e 12σ . Entretanto, a eficiência do DS decresce significativamente para outliers menores que 5σ .

A maior deficiência do método analisado foi a detecção de outliers menores que 5σ , sendo assim é de suma importância a realização do nivelamento e contranivelamento pelo fato de que a taxa de sucesso nesse intervalo aumentou quase vinte por cento e na prática a magnitude de possíveis erros grosseiros é sempre desconhecida.

Para futuros trabalhos pode ser considerado vários cenários: a performance do Data Snooping para múltiplos outliers, fazer várias iterações com o DS, analisar a geometria da rede e sua correlação com o controle de qualidade.

REFERÊNCIAS

BAARDA, W. A testing procedure for use in geodetic networks. **Publications on Geodesy, New Series**, v. 2, n. 5, Delft: Netherlands Geodetic Commission, 1968.

FÖSTNER, W. Reliability analysis of parameter estimation in linear models with applications to mensuration problems in computer vision. *Computer Vision, Graphics and Image processing*, v. 40, p. 273-310, 1987.

GHILANI, C. D.; WOLF, P. R. **Adjustment Computations: Spatial Data Analysis**. 4. Ed. New York: John Wiley & Sons, 2006, 611p.

KLEIN, Ivandro. **CONTROLE DE QUALIDADE APLICADO NO AJUSTAMENTO DE OBSERVAÇÕES PELO MÉTODO PARAMÉTRICO PARA APLICAÇÕES GEODÉSICAS**. 2011. 310 f. Dissertação (Mestrado) - Curso de Engenharia Cartográfica, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.

KLEIN, I. et al. On evaluation of different methods for quality control of correlated observations. **Survey Review**. Coldharbour Lane, p. 28-35. jan, 2015.



I Simpósio Regional de Agrimensura e Cartografia

“A Topografia e a Geodésia no cenário nacional”

Universidade Federal de Uberlândia – UFU / Campus Monte Carmelo

22 a 24 de novembro de 2016

KLEIN, I.; MATSUOKA, M. T.; GUZATTO, M. P.. Como estimar o poder do teste mínimo e valores limetes para o intervalo de confiança do data snooping. **Boletim de Ciências Geodésicas**. Curitiba, p. 26-42. jan. 2015.

LARSON, H. J. Introduction to probability theory and statistical inference. 2. ed. **New York: John Wiley & Sons**, 1974. 430p.

LEHMANN, R. Improved critical values for extreme normalized and studentized residuals in Gauss–Markov models. **Journal of Geodesy**, v. 86, n. 12, p. 1137– 1146, 2012.

TEUNISSEN, P. J. G. Testing theory: an introduction. 2. ed. **Delft: Ed. VSSD**, 2006. 147p.