

MÉTODO MONTE CARLO APLICADO À ANÁLISE DE CONFIABILIDADE EM GEODÉSIA

V. F. Rofatto^{1,2}, M. T. Matsuoka^{1,2}, I. Klein³

¹Universidade Federal de Uberlândia, Curso de Engenharia de Agrimensura e Cartográfica, Monte Carmelo, Minas Gerais, Brasil

²Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Sensoriamento Remoto, Rio Grande do Sul, Porto Alegre, Brasil

³Instituto Federal de Santa Catarina, Curso de Agrimensura, Florianópolis, Santa Catarina, Brasil

Comissão II - Geodésia, Astronomia, Topografia e Agrimensura

RESUMO

Com o crescente aumento da automatização dos processos de aquisição de dados geodésicos, é quase impossível que o conjunto de observações esteja livre de erros grosseiros (*outliers*). Uma das formas frequentemente empregada para o tratamento das observações contaminadas por *outliers* é baseada em testes estatísticos. Por ser uma estratégia formulada com base em hipóteses estatísticas, ou seja, com níveis de probabilidade associada, pode-se cometer erros na tomada de decisão. Embora, em teoria, seja possível descrever as funções densidade de probabilidade dos tipos de erros envolvidos nas decisões estatísticas, na prática, o cálculo algébrico de tais funções é altamente complexo. Hoje, no entanto, temos a nossa disposição computadores poderosos e rápidos que permitem a obtenção dos níveis de probabilidades associados ao procedimento de teste. Aqui, apresentamos brevemente o método Monte Carlo de simulação computadorizada. Para clarificar o método Monte Carlo de simulação, uma discussão objetiva do tema encontrado na literatura, acompanhada de um simples exemplo numérico é apresentado.

Palavras chave: Ajustamento, Detecção de Outliers, Teste de Hipóteses, Poder de um Teste, Método Monte Carlo.

ABSTRACT

With the increasing automation of the processes of geodetic data acquisition, it is nearly impossible that the set of observations is free of gross errors (*outliers*). One of the ways often employed for the treatment of observations contaminated by outliers is based on statistical hypothesis testing. Because it is a strategy formulated on the basis of statistical hypothesis, it may lead to a false decision. Although in theory it is possible to describe the probability density functions of the wrong decisions, in practice the algebraic computation of such functions is highly complex. However, today we have fast and power computers at our disposal that allow to accomplish this. Here, we briefly introduce the computer simulation based on Monte Carlo method. An objective discussion of the issue found in the literature followed by simple numerical example is presented in order to clarify the Monte Carlo simulation method.

Keywords: Geodetic Adjustment, Outlier Detection, Hypothesis Testing, Power of a Test, Monte Carlo Method.

1- INTRODUÇÃO

A teoria convencional de confiabilidade avalia se um determinado *outlier* em uma observação é detectável pelo procedimento de teste utilizado, e avalia também a influência deste nos resultados do ajustamento por mínimos quadrados, quando não detectado, segundo os níveis de probabilidade que foram estipulados (Baarda, 1968).

A qualidade da(s) estimativa(s) do(s) parâmetro(s) usando o método clássico dos mínimos

quadrados é afetada por *outliers* (veja por exemplo, Baarda, 1968).

Duas categorias empregadas para o tratamento das observações contaminadas por *outliers* são frequentemente usadas: métodos de estimação robusta e procedimentos baseados em testes estatísticos (Baarda, 1968; Pope, 1976; Lehmann and Lösler, 2016; Klein et al., 2016). Os métodos de estimação robusta não fazem parte do escopo deste trabalho (uma revisão é apresentada em Wilcox, 2012; Klein et al., 2015a).

O método baseado em testes estatísticos mais bem estabelecido para tratar *outliers* em dados geodésicos, com uma determinada probabilidade, é o procedimento estatístico proposto por Baarda, comumente chamado na literatura internacional de *Baarda's test* (Baarda, 1968). Este método consiste de três etapas (Teunissen 2000, 2006): detecção (também conhecido como teste global do ajustamento, na literatura inglesa *overall model test*), identificação (também conhecido como *Data Snooping*) e adaptação (geralmente, eliminação da observação identificada como *outlier*). Aqui, restringimos à *identificação de outliers*, ou seja, ao procedimento *Data Snooping* (DS) em sua forma iterativa (Teunissen, 2006).

Ressaltamos que originalmente o DS proposto por Baarda (Baarda, 1967; Baarda, 1968) é baseado na decisão entre a hipótese nula e uma única hipótese alternativa. Yang et al., 2013 estendeu a teoria original de Baarda para duas hipóteses alternativas a hipótese nula. No entanto, o atual estado da arte em confiabilidade é devido à Yang et al., 2017, no qual estende o teste DS para múltiplas hipóteses alternativas a hipótese nula, uma para cada observação (Yang et al., 2017).

Independente das extensões do DS, todos são baseados em hipóteses estatísticas, ou seja, com níveis de probabilidade associado e, portanto, pode-se cometer erros na tomada de decisão (detalhes do método DS iterativo será abordado na próxima seção). Embora, em teoria, seja possível descrever as funções densidade de probabilidade dos tipos de erros envolvidos nas decisões estatísticas do DS iterativo, na prática, o cálculo algébrico de tais funções é altamente complexo. As funções densidades de probabilidade envolvidas são praticamente impossíveis de serem obtidas por uma “expressão analítica fechada”. Entende-se por expressão analítica fechada um conjunto de equações que, fornecidos os valores de entrada, fornecem resultados sobre o comportamento do modelo a partir de uma fórmula analítica direta.

Hoje, no entanto, é possível obter os níveis de probabilidade associado ao DS iterativo por meio de simulação computadorizada. Para clarificar o método Monte Carlo de simulação, uma discussão objetiva do tema encontrado na literatura, acompanhada de um simples exemplo numérico é apresentada.

2- PROCEDIMENTO DATA SNOOPING ITERATIVO PARA IDENTIFICAÇÃO DE OUTLIERS

No caso do procedimento de teste *Data Snooping*, proposto em Baarda (1968), inicialmente, considera-se o seguinte modelo linear (ou linearizado) de Gauss Markov (Koch, 1999):

$$e = y - Ax \quad (1)$$

onde e é o vetor de $n \times 1$ de erros aleatórios, A é a matriz jacobiana (ou *design*) de $n \times u$, x é o vetor de $u \times 1$ dos parâmetros desconhecidos, e y é o vetor de

$n \times 1$ das observações. Na equação 1, a solução mais empregada para o vetor dos parâmetros desconhecidos, considerando um sistema de equações redundante ($n > \text{posto}(A)$), é o estimador dos mínimos quadrados ponderados (MMQ), dado por:

$$\hat{x} = (A^T W A)^{-1} (A^T W y) \quad (2)$$

W é a matriz $n \times n$ dos pesos das observações, tomada como $W = \sigma_0^{-2} (\Sigma_y)^{-1}$, onde σ_0^2 é o fator de variância a priori e Σ_y é a matriz $n \times n$ das covariâncias das observações.

O MMQ é o melhor estimador linear imparcial e também uma solução de máxima verossimilhança quando os erros aleatórios seguem uma distribuição normal (Teunissen, 2003). O termo “melhor” é no sentido de ser uma solução de variância mínima. Porém, as propriedades da solução por MMQ não são mais garantidas quando existem *outliers* nas observações. Se os *outliers* não forem identificados, estes possivelmente influenciarão a estimativa do vetor dos parâmetros.

O procedimento *data snooping* (DS) é um método estabelecido para tratar *outliers*. Ele é um caso particular do teste de razão de máxima verossimilhança quando somente um *outliers* está presente no conjunto de observações (veja por exemplo, Baarda, 1968; Pope, 1976; Berber e Hekimoglu, 2003; Lehmann, 2012). O DS é baseado nas seguintes hipóteses (Baarda, 1968; Teunissen, 2006, Yang et al., 2013):

$$\begin{aligned} H_0 : E\{y\} &= Ax \\ H_i : E\{y\} &= Ax + c_i \nabla; \nabla \neq 0 \text{ e } i = 1, \dots \end{aligned} \quad (3)$$

onde $E\{y\}$ corresponde o operador de esperança matemática, c_i é um vetor unitário contendo a unidade na linha da i -ésima observação testada e zero nas demais linhas, ou seja $c_i = \begin{bmatrix} 0 & 0 & \dots & \dots & 1 \end{bmatrix}^{-T}$, e ∇ é um valor escalar de um (possível) *outlier* na i -ésima observação testada.

Desta forma, na hipótese nula (H_0), assume-se que não existe *outliers* nas observações, enquanto na hipótese alternativa (H_i), assume-se que a i -ésima observação testada está contaminada por um *outlier* de magnitude ∇ .

A estatística de teste do DS para i -ésima observação (w_i) pode ser calculada por (Baarda, 1968; Kok, 1984):

$$w_i = \frac{c_i^T W \hat{e}_0}{\sqrt{c_i^T \Sigma_y^{-1} \Sigma_{\hat{e}_0} \Sigma_y c_i}}; \quad (4)$$

onde \hat{e}_0 é o vetor de erros aleatórios e $\Sigma_{\hat{e}_0}$ sua respectiva matriz covariância estimados por meio do

MMQ sob a hipótese nula. Nota-se que para calcular a estatística de teste na equação 4 não é necessário realizar o ajustamento sob a hipótese alternativa, basta apenas formular o modelo de *outlier* (dado pelo vetor c_i).

Como cada observação é testada individualmente, a (única) observação considerada contaminada por *outlier* será aquela que apresenta valor máximo absoluto da estatística de teste maior que um dado valor crítico ($K_{\alpha_0/2}$). Em outras palavras, a hipótese nula é rejeitada quando (Baarda, 1968; Teunissen, 2006):

$$\max_{i=1, \dots} |w_i| > K_{\alpha_0/2} \quad (5)$$

$K_{\alpha_0/2}$ é o valor crítico teórico na distribuição normal padrão, para um dado nível de significância α_0 (teste bi-lateral).

Por exemplo, uma j -ésima observação é considerada um *outlier* quando satisfazer as seguintes condições (Baarda, 1968; Teunissen, 2006):

$$|w_j| > |w_i| \quad \forall_i, \quad |w_j| > K_{\alpha_0/2} \quad (6)$$

Identificada a observação suspeita de estar contaminada por erro grosseiro, segundo o nível de significância estipulado (α_0), usualmente, exclui-se a mesma do conjunto de dados e repete-se o ajustamento e o procedimento de teste DS até que todas as observações suspeitas sejam devidamente identificadas, em um processo iterativo de ajustamento, identificação e remoção de *outliers* (um por vez). Para mais detalhes sobre este procedimento, ver, por exemplo, Baarda, 1968; Kavouras, 1982; Koch, 1999; Teunissen, 2006; Lehmann, 2012 e Klein et al., 2015b.

No entanto, ao se aplicar o procedimento de teste DS, podem ocorrer três tipos de decisões falsas (considerando o caso bidimensional, ou seja, baseada na decisão entre a hipótese nula e duas hipóteses alternativas):

- Erro tipo I ou “*false alert*” – probabilidade rejeitar H_0 (ou aceitar uma hipótese alternativa), quando na realidade H_0 é verdadeira. Em outras palavras, identificar um *outlier* quando na realidade não existe.

- Erro tipo II ou “*missed detection*” – probabilidade de rejeitar uma específica hipótese alternativa quando na realidade ela é verdadeira, ou seja, é a probabilidade de não identificar um *outlier* quando na realidade existe pelo menos um; e

- Erro tipo III ou “*wrong exclusion*” – probabilidade de rejeitar a hipótese nula corretamente, porém aceitar a hipótese alternativa errada. Em outras palavras, é a probabilidade de identificar (erroneamente) uma observação como sendo *outlier*, enquanto o *outlier* não foi identificado.

A probabilidade de cometer o Erro Tipo I é chamada de nível de significância do teste e é dada por α ; a probabilidade de cometer o Erro Tipo II é dada por β ; e a probabilidade de cometer o Erro Tipo III é dada por κ . Em contrapartida, o intervalo de confiança (IC) é o complemento do Erro Tipo I, ou seja, é a probabilidade de não identificar um *outlier* corretamente, ($IC = 1 - \alpha$), enquanto o poder do teste, designado por γ , é o complemento da soma das probabilidades do Erro Tipo II e do Erro Tipo III ($\gamma = 1 - (\beta + \kappa)$), ou seja, é a probabilidade de identificar (corretamente) um *outlier*. Portanto, o intervalo de confiança e o poder do teste são as probabilidades do resultado do teste conduzir a decisões corretas, ao contrário da ocorrência dos erros tipo I, II e III (para mais detalhes, ver, por exemplo, Förstner, 1983; Teunissen, 2006). Quando o valor para a probabilidade do Erro Tipo I, do Erro Tipo II ou do poder do teste é pré-definido/estipulado, este valor é designado por α_0 , β_0 ou γ_0 , respectivamente.

A Tabela 1 mostra um exemplo para as decisões associadas a duas hipóteses alternativas no caso de uma única rodada do DS, sem considerar o processo de remoção de *outliers* (Yang et al., 2013). Desta forma, α_{00} é o nível de significância “global” do DS (enquanto α_{0i} é o nível de significância apenas para o teste da i -ésima observação, e α_{0j} é o nível de significância apenas para o teste da j -ésima observação). Além disso, β_{i0} e κ_{ij} são as probabilidades de cometer, respectivamente, o Erro Tipo II e o Erro Tipo III para a i -ésima observação, enquanto β_{j0} e κ_j são as probabilidades de cometer, respectivamente, o Erro Tipo II e o Erro Tipo III para a j -ésima observação. Finalmente, β e γ_{ii} são, respectivamente, a soma de β_{i0} e κ_j e o poder do teste para a i -ésima observação, enquanto β_{jj} e γ_{jj} são, respectivamente, a soma de β_{j0} e κ_{ji} e o poder do teste para a j -ésima observação. As estatísticas de teste w_i e w_j de duas observações quaisquer possuem um coeficiente de correlação (ρ_{ij}), dado pela seguinte expressão (Förstner, 1983):

$$\rho_{ij} = \frac{c_i^T W \Sigma_{\hat{e}_i} W c_j}{\sqrt{c_i^T W \Sigma_{\hat{e}_i} W c_i} \sqrt{c_j^T W \Sigma_{\hat{e}_j} W c_j}} \quad (7)$$

Além do coeficiente de correlação entre duas estatísticas de teste do DS, tem-se ainda o parâmetro de não centralidade do modelo (δ_0), que expressa a separação entre a hipótese nula e a hipótese alternativa.

No caso bidimensional aqui considerado, o parâmetro de não centralidade do modelo (δ_0) depende do nível de significância do teste (α_0), ou analogamente, do valor crítico teórico para este ($K_{\alpha_0/2}$), do poder do teste (γ_0), ou analogamente, da soma das probabilidades do Erro Tipo II (β_{0i}) e do Erro Tipo III (κ_{ij}), e também do coeficiente de correlação entre as estatísticas de teste consideradas (ρ_{ij}). Desta forma,

matematicamente, as probabilidades de cometer o Erro Tipo I (α_{0i}), o Erro Tipo II (β_{0i}) e o Erro Tipo III (κ_{ij}) no teste de uma i -ésima observação qualquer, considerando o coeficiente de correlação (ρ_{ij}) da estatística de teste desta (w_i) com a de uma j -ésima observação qualquer (w_j), bem como o parâmetro de não centralidade do modelo correspondente (δ_0), são dadas, respectivamente, por (Förstner, 1983):

$$f_0 = \frac{1}{2\pi\sqrt{1-\rho^2}} - \frac{1}{2(1-\rho^2)} [w_i^2 - 2\rho w_i w_j + w_j^2] \quad (8)$$

$$f_1 = \frac{1}{2\pi\sqrt{1-\rho^2}} - \frac{1}{2(1-\rho^2)} [(w_i - u_i)^2 - 2\rho(w_i - u_i)(w_j - u_j) + (w_j - u_j)^2] \quad (9)$$

$$\alpha_{0i} = \iint_{|w_i| > K_{\alpha_0/2}, |w_j| > |w_i|} f_0 dw_i dw_j \quad (10)$$

$$\beta_{0i} = \iint_{|w_i| \leq K_{\alpha_0/2}, |w_j| > K_{\alpha_0/2}} f_1 dw_i dw_j \quad (11)$$

$$\kappa_{ij} = \iint_{|w_i| > K_{\alpha_0/2}, |w_j| > |w_i|} f_1 dw_i dw_j \quad (12)$$

Na Expressão 10, o termo dentro da integral corresponde à função densidade de probabilidade (fdp) conjunta de w_i e w_j na hipótese nula H_0 (ausência de outliers nas duas observações), enquanto nas expressões 11 e 12, o termo dentro das integrais corresponde a fdp conjunta de w_i e w_j na hipótese alternativa H (existência de outlier na i -ésima observação, com $\mu = \delta_0$ e $\mu_j = \rho_j \delta_0$). Como as distribuições de probabilidade de w_i e w_j em H_i e H_j são simétricas, devido ao coeficiente de correlação entre estas, decorre que $\alpha_{0i} = \alpha_{0j}$, $\beta_{0i} = \beta_{0j}$, e $\kappa_{ij} = \kappa_{ji}$ (para mais detalhes, ver Förstner, 1983 e Yang et al., 2013).

As funções de densidade de probabilidade apresentadas nas expressões 10, 11 e 12 ficam mais complexas quando se considera o cenário tridimensional, tetradimensional, e até mesmo, estendendo até o caso n -dimensional, ou seja, considerando todas as n observações testadas (individualmente) – o caso de múltiplas hipóteses alternativas (Yang et al., 2017). E, no caso do DS iterativo, ou seja, ajustamento, identificação e remoção de *outliers* (um por vez), é ainda mais complicado, uma vez que envolve, além de múltiplas hipóteses, o problema de considerar a remoção de *outliers* em cada iteração, o que já altera toda a região de probabilidade supracitado nesse trabalho.

Nesse sentido, simulações computadorizadas, tal como o método Monte Carlo, são necessárias a fim de obter os níveis de probabilidade do DS iterativo de forma mais realista. Esse método é factível, uma vez que a probabilidade pode sempre ser escrita como uma esperança matemática e esta esperança pode ser aproximada tomando a média (teorema central do limite) de um número suficiente de amostras (lei dos grandes números) geradas computacionalmente a partir de uma distribuição. Detalhes do método Monte Carlo são apresentadas no próximo tópico.

3- INTRODUÇÃO À SIMULAÇÃO DE MONTE CARLO

Esse tópico é introdutório e mais detalhes sobre o(s) método(s) Monte Carlo podem ser obtidos, por exemplo Robert e Casella, (2013), e Gamerman e Lopes, (2006).

O método Monte Carlo de simulação (MCS) é uma ferramenta fundamental em muitas áreas da ciência moderna (variando desde a física teórica até a ciência política). Foi inicialmente usado para solucionar problemas de difusão de nêutrons no desenvolvimento da bomba atômica em 1944. Hoje, é usado na biologia, física médica, física quântica, estatística aplicada, e muitas outras áreas. MCS também é usado para solucionar problemas que envolvem um grande número de variáveis, vários tipos de distribuições estatísticas e modelos não-lineares complexos (Robert e Casella, 2013).

O Teorema Central do Limite e a Lei Forte dos Grandes Números são os ingredientes principais do MCS. A ideia é gerar uma amostragem aleatória e conduzir um grande número de experimentos. Em cada experimento, os possíveis valores das variáveis aleatórias de entrada $X=(X_1, X_2, \dots, X_n)$ são amostradas (geradas) de acordo com suas distribuições. Em seguida, os valores da variável de saída Y são calculados por meio de uma função, por exemplo $Y=F(X)$. Com uma série de experimentos realizados desta maneira, um conjunto de amostras da variável de saída está disponível para a análise estatística.

Para clarificar o método MCS, suponhamos que teríamos que resolver a seguinte integral de uma função $f(x)$ no intervalo $[a, b]$, ou seja:

$$I = \int_a^b f(x) dx \quad (13)$$

Essa integral pode ser escrita como:

$$I = \int_a^b (b-a) f(x) \frac{1}{b-a} dx = (b-a) E\{f(x)\} \quad (14)$$

Identificando x como uma variável aleatória com distribuição uniforme $U[a, b]$. Assim, transformamos o problema de avaliar a integral no problema estatístico de encontrar a esperança matemática (média), $E\{f(x)\}$. Se dispomos de uma amostra aleatória de tamanho n (x_1, x_2, \dots, x_n) da distribuição uniforme no intervalo $[a, b]$, teremos também uma amostra de valores $f(x_1), f(x_2), \dots, f(x_n)$ da função $f(x)$, e a integral acima pode ser estimada pela média amostral, ou seja:

$$\hat{I} = (b-a) \frac{1}{n} \sum_{i=1}^n f(x_i) \quad (15)$$

A estimativa dada pela expressão 15 é não enviesada. Assim podemos usar o seguinte algoritmo:

1. gerar x_1, \dots, x_n a partir de $U[a, b]$;
2. Calcular $f(x_1), f(x_2), \dots, f(x_n)$;

3. Calcular a média amostral $\bar{f} = \sum_{i=1}^n f(x_i) / n$

4. Calcular $\hat{I} = (b-a)\bar{f}$

Executando o algoritmo descrito acima, por exemplo m vezes para uma certa amostra n e no intervalo $[a,b]$, existirá uma variação considerável na estimativa da integral. Uma vez que as amostras são independentes, pela Lei Forte dos Grandes Números segue que \hat{I} converge certamente para I . O Teorema Central do Limite também se aplica aqui de modo que:

$$\frac{\bar{f} - E\{f(x)\}}{\sqrt{v}} \sim N(0,1) \quad (16)$$

$$v = \frac{1}{n^2} \sum_{i=1}^n (f(x_i) - \bar{f})^2 \quad (17)$$

onde v é a variância do estimador.

A expressão 16 tem distribuição próxima da normal padrão $N(0,1)$. Pode-se usar este resultado para testar a convergência e construir intervalos de confiança. A expressão 17 mostra que a aproximação pode ser tão acurada quanto se deseje bastando aumentar n . É importante notar que n é controlado, e não se trata do tamanho da amostra de dados.

Portanto, podemos usar o método MCS para avaliar o poder do teste DS, os erros tipo I, II e III, entre outras possibilidades. Em geodesia, o método MCS tem sido amplamente discutido e usado por diversos autores, por exemplo, Klein et al., 2012; Klein et al., 2015b; Knight et al., 2010; Lehmann e Scheffler, 2011; Lehmann, 2012; Lehmann, 2013; Erdogan, 2014; Koch, 2017; Yang et al., 2013; Yang et al., 2016; Klein et al., 2016; Yang et al., 2017; Lehmann e Voß-Böhme, 2017; Teunissen, 2017; Prószyński, 2015. Na próxima seção, um exemplo numérico simples é apresentado com intuito de estimar o poder do teste DS iterativo.

4- EXEMPLO NUMÉRICO SIMPLES

Aqui, apresentamos um exemplos simples do método MCS com a finalidade de estimar os níveis de probabilidade associado ao DS iterativo. Para isso, consideramos o exemplo de uma rede GNSS (*Global Navigation Satellite System*) usada por Klein, 2014. A rede envolve um ponto de controle (fixa) e cinco estações com coordenadas 3D (X,Y,Z) desconhecidas (Veja Figura 1). Para cada par de estações, existem quatro a cinco linhas base ($\Delta X, \Delta Y, \Delta Z$). Assim, existem 39 observações ($n = 13 \times 3$), 15 parâmetros desconhecidos ($u = 5 \times 3$), e $n - u = 39 - 15 = 24$ observações redundantes. As estações pertencem a Rede Brasileira de Monitoramento Contínuos dos sinais GNSS (RBMC). Os vetores linhas base, livre de outliers, consistem nas diferenças entre as coordenadas oficiais no referencial SIRGAS2000. A matriz covariância foi obtida por meio do processamento de

sessões de 6-horas. Para cada vetor linha base foi obtida uma matriz covariância completa de 3×3 , resultando no total de $13 \times (3 \times 3) = 39 \times 39$ de bloco diagonal. Mais detalhes sobre a rede pode ser encontrada em Klein, 2014.

O MCS é aplicado para estimar o poder do teste do procedimento DS iterativo para a rede considerada. Para isso, 30,000 experimentos foram realizados. Em cada experimento foi gerado um vetor de erros aleatórios segundo a distribuição normal multivariada $e \sim N(0, \Sigma_y)$ e um outlier foi inserido propositalmente em um dada observação. A magnitude do outlier foi estabelecida no intervalo entre 3σ a 5σ , sendo “ σ ” o desvio-padrão da observação correspondente. O número de experimentos que o procedimento *data snooping* detecta corretamente um outlier é quantificado por meio do método MCS.

Foi aplicado o procedimento de teste DS iterativo com um nível de significância estipulado em $\alpha_0 = 1\%$ (0.01), ou seja, com um valor crítico tabelado correspondente de $K_{\alpha_0/2} = 2,5758293$. O poder do teste DS iterativo para toda a rede (“poder do teste global”) obtido experimentalmente (empiricamente) por meio do MCS foi de 0,67. Isso significa que na presença um outlier de magnitude entre 3σ a 5σ , considerando ainda a geometria da rede (modelo funcional) e a incerteza das observações (modelo estocástico), o teste DS iterativo consegue identificar um outlier corretamente em 67% dos casos à um nível de significância de 0.01 (Erro tipo I). Diante do exposto, nota-se a possibilidade de usar essa poderosa ferramenta para análise de confiabilidade em geodesia.

5- OBSERVAÇÕES FINAIS

Nós mostramos que o MCS é um método factível para estimar os níveis de probabilidade associado ao procedimento de teste estatístico. O método não faz uso do vetor de observações reais coletadas em campo, dependendo apenas da configuração geométrica da rede (modelo funcional), das incertezas das observações (modelo estocástico) e da magnitude do outlier. Estudos futuros deverão considerar cenários com a presença de múltiplos outliers; considerar a problemática do DS iterativo exposta nesse trabalho.

Diante do exposto, nota-se a possibilidade de usar o método MCS, uma poderosa ferramenta para análise de confiabilidade em geodesia, podendo ser estendida em estudos de planejamento e otimização de redes geodésicas; aplicar em análises de propagação de incertezas; combinar MCS com técnicas inteligentes, tais como algoritmos genéticos, enxame de partículas (*Particle Swarm Optimization*), *Simulated Annealing* e até mesmo com Redes Neurais e *Support Vector Machine*; dentre outras possibilidades.

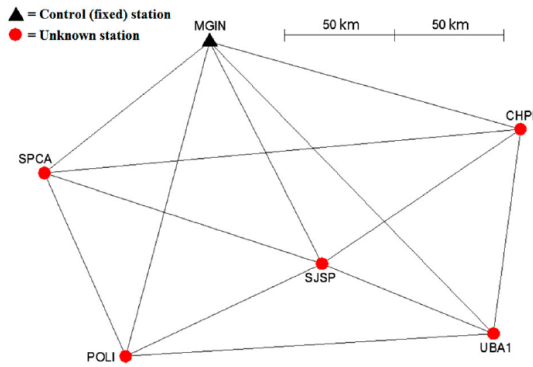


Fig. 1 – Exemplo de Rede GNSS (Klein, 2014).

TABELA 1 – Decisões associadas a duas hipóteses alternativas no DS (Fonte: Adaptado de Yang et al., 2013).

Realidade ("desconhecida")	Resultado do teste		
	H_0	H_i	H_j
$ w_i \leq \alpha_0/2$ $ w_j \leq \alpha_0/2$	Decisão Correta $1 - \alpha_0$	Erro tipo I α_{0i}	Erro tipo I α_{0j}
$ w_i > \alpha_0/2$ $ w_i > $	Erro tipo II β_{i0}	Decisão correta $\gamma_{ii} = 1 - \beta_{ii}$	Erro tipo III κ_{ij}
$ w_j > \alpha_0/2$ $ w_j > $	Erro tipo II β_{j0}	Erro tipo III κ_{ji}	Decisão correta $\gamma_{jj} = 1 - \beta_{jj}$

AGRADECIMENTOS

Esse trabalho é fruto dos esforços do Grupo de Pesquisa em "Controle de Qualidade em Geodesia" (dgp.cnpq.br/dgp/espelhogrupo/3674873915161650). Os autores também agradecem o Instituto de Geografia (IG), da Universidade Federal de Uberlândia pelo apoio à divulgação desse trabalho. Os autores agradecem também ao CNPq pela Bolsa de Produtividade em Pesquisa (Proc.n.305599/2015-1) do segundo autor.

REFERÊNCIAS BIBLIOGRÁFICAS

Baarda, W., 1967. Statistical Concepts in Geodesy. Publications on Geodesy. Delft, Netherlands, Vol. 2, Nr. 4.

Baarda, W., 1968. A testing procedure for use in geodetic networks. Publications on Geodesy, Netherlands Geodetic Commission, Vol. 2, Nr. 5.

Berber, M. e Hekimoglu, S. 2003. What is the reliability of conventional outlier detection and robust estimation in trilateration networks? Survey Review,37(290), pp.308-318.

Erdogan, B., 2014. An outlier detection method in geodetic networks based on the original observations. Boletim de Ciências Geodésicas. 20(3), pp.578-589.

Förstner, W., 1983. Reliability and discernability of extended gauss-markov models. In: Deut. Geodact. Komm. Seminar on Math. Models of Geodetic Photogrammetric Point Determination with Regard to Outliers and Systematic Errors, pp. 79-104.

Gamerman, D. e Lopes, H. F., 2006. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. Chapman and Hall/CRC, 342 páginas.

Kavouras, M. 1982. On the detection of outliers and the determination of reliability in geodetic networks. PhD thesis, Fredericton: Department of Surveying Engineering, University of New Brunswick.

Klein, I. 2014. Proposal of a new method for geodetic networks design (in Portuguese). PhD thesis, Porto Alegre: Universidade Federal do Rio Grande do Sul.

Klein, I. et al. 2012. Design of geodetic networks reliable against multiple outliers (in Portuguese). Boletim de Ciências Geodésicas, 18(3), pp.480-507.

Klein, I. et al. 2015a. On evaluation of different methods for quality control of correlated observations. Survey Review. 47(340), pp.28-35.

- Klein, I. et al. 2015b. How to estimate the minimum power of the test and bound values for the confidence interval of data snooping procedure (in Portuguese). *Boletim de Ciências Geodésicas*. 21(1), pp.26-42.
- Klein, I. et al. 2016. An approach to identify multiple outliers based on sequential likelihood ratio tests, *Survey Review*. Latest articles - DOI: 10.1080/00396265.2016.1212970
- Knight, N. L. Wang, J. Rizos, C. 2010. Generalised measures of reliability for multiple outliers. *Journal of Geodesy*. 84(10), pp 625-635.
- Koch, K. R. 1999. *Parameter Estimation and Hypothesis Testing in Linear Models*. Berlin Heidelberg New York, Springer Verlag.
- Koch, K. R. 2017. Expectation Maximization algorithm and its minimal detectable outliers. *Stud. Geophys. Geod.*, 61, pp. 1-18.
- Lehmann, R. and Scheffler, T. 2011. Monte Carlo-based data snooping with application to a geodetic network. *Journal of Applied Geodesy*. 5(3-4), pp.123-134. DOI: 10.1515/JAG.2011.014
- Lehmann, R. 2012. Improved critical values for extreme normalized and studentized residuals in gauss-markov models. *Journal of Geodesy*. 86(12), pp.1137-1146.
- Lehmann, R. 2013. On the formulation of the alternative hypothesis for geodetic outlier detection. *Journal of Geodesy*, 87(4), pp.373-386.
- Lehmann, R. and Lösler, M. 2016. Multiple outlier detection: hypothesis tests versus model selection by information criteria. *Journal of Surveying Engineering*, 142(4). DOI: 10.1061/(ASCE)SU.1943-5428.0000189
- Lehmann, R. and Voß-Böhme, A. 2017. On the statistical power of Baarda's outlier test and some alternative. *Journal of Geodetic Science*, 7(1), pp.68-78.
- Pope, A. J., 1976. *The statistics of residuals and the detection of outliers*. NOAA Technical Report NOS65 NGS1, US Department of Commerce, National Geodetic Survey Rockville, Maryland.
- Prószyński, W. 2015. Revisiting Baarda's concept of minimal detectable bias with regard to outlier identifiability. *Journal of Geodesy*, 89(10), pp. 993-1003.
- Robert C, Casella G (2013) *Monte Carlo statistical methods*. Springer, Berlin, 580 páginas.
- Teunissen, P. J. G. 2000. *Testing theory: An Introduction*. Series on Mathematical Geodesy and Positioning, first edition. VSSD, Delft University of Technology, The Netherlands.
- Teunissen, P. J. G. 2006. *Testing theory: An Introduction*. Series on Mathematical Geodesy and Positioning. 2nd ed. VSSD, Delft University of Technology, The Netherlands.
- Teunissen, P. J. G. 2017. Distributional theory for the DIA method. *Journal of Geodesy*, pp.1-22. DOI: 10.1007/s00190-017-1045-7.
- Wilcox, R.R. 2012. *Introduction to robust estimation and hypothesis testing*. Academic Press, Waltham, MA.
- Yang, L et al., 2013. Outlier separability analysis with a multiple alternative hypotheses test. *Journal of Geodesy*, vol. 87, Issue 6, pp. 591–604.
- Yang, L et al., 2017. Extension of Internal Reliability Analysis Regarding Separability Analysis. *Journal of Surveying Engineering*, vol. 143, issue 3.