

Análise gráfica das variáveis do controle de qualidade de dados geodésicos por meio de testes estatísticos

RESUMO

Na Geodésia, o controle de qualidade (CQ) é uma das etapas de maior importância, onde é decidido se os resultados obtidos serão validados ou não. Dentro desse contexto, os testes estatísticos ainda são as ferramentas mais utilizadas para detecção e identificação de *outliers* (erros não aleatórios), dentre eles o famoso *Data Snooping* proposto por Baarda. Logo, o entendimento das relações matemáticas entre as variáveis envolvidas nos testes estatísticos é de suma importância não só na aplicação do CQ, mas também em outros campos da Geodésia como a análise de robustez e o planejamento de redes. Dentro desse contexto, o objetivo desse trabalho é apresentar uma análise gráfica das diversas variáveis envolvidas no CQ de dados geodésicos por meio de testes estatísticos. Para atingir este objetivo, foram apresentados diversos gráficos envolvendo as variáveis consideradas, sendo estas: nível de confiança (ou de significância); poder do teste; Erros Tipo I, II e III; coeficientes de correlação entre as estatísticas de teste; graus de liberdade (número de *outliers* considerados) e parâmetro de não centralidade do modelo. Com base nos resultados obtidos, apresentam-se algumas considerações, conclusões e recomendações para trabalhos futuros.

PALAVRAS-CHAVE: Controle de qualidade. Testes estatísticos. Níveis de probabilidade.

Matheus Pereira Guzzatto

matheus.guzzatto@ifsc.edu.br
orcid.org/0000-0003-1645-2338
Instituto Federal de Santa Catarina (IFSC). Florianópolis, Santa Catarina, Brasil.

Ivandro Klein

ivandroklein@gmail.com
orcid.org/0000-0003-4296-592X
Instituto Federal de Santa Catarina (IFSC). Florianópolis, Santa Catarina, Brasil.

Marcelo Tomio Matsuoka

tomiomatsuoka@gmail.com
orcid.org/0000-0002-2630-522X
Universidade Federal de Uberlândia (UFU). Monte Carmelo, Minas Gerais, Brasil.

Vinicius Francisco Rofatto

rofattaum@gmail.com
orcid.org/0000-0003-1453-7530
Universidade Federal de Uberlândia (UFU). Monte Carmelo, Minas Gerais, Brasil.

Mauricio Roberto Veronez

veronez@unisinis.br
orcid.org/0000-0002-5914-3546
Universidade do Vale do Rio dos Sinos (UNISINOS). São Leopoldo, Rio Grande do Sul, Brasil.

INTRODUÇÃO

O projeto completo de uma rede geodésica consiste em quatro etapas, sendo estas: planejamento, levantamento de campo, ajustamento das observações e controle de qualidade (CQ). No CQ, é decidido se os resultados obtidos são validados ou não, sendo feito com a aplicação de métodos que detectam e identificam possíveis erros cometidos durante as etapas anteriores, como por exemplo, a presença de erros grosseiros (*outliers*) nas observações (KLEIN, 2014a). Os principais métodos existentes para o CQ de dados geodésicos, são os de estimação robusta (ver, por exemplo, KOCH, 1999); de inferência bayesiana (ver, por exemplo, GUI *et al.*, 2011); o QUAD (ver, por exemplo, GUO *et al.*, 2007); o teste Tau (ver, por exemplo, POPE, 1976); testes estatísticos para múltiplos (simultâneos) *outliers* (ver, por exemplo, KNIGHT *et al.*, 2010 e KLEIN *et al.* 2016), e o *Data Snooping* (ver, por exemplo, BAARDA, 1968). Uma revisão sobre os principais métodos utilizados para o CQ de dados geodésicos pode ser obtido em Klein (2012) e Klein *et al.* (2015).

Dentre os métodos citados anteriormente, o *Data Snooping* (DS) é o mais bem estabelecido para identificar *outliers* em dados geodésicos (LEHMANN, 2012; 2017). O DS pertence a uma família geral de testes estatísticos baseados na razão de máxima verossimilhança, onde o parâmetro de não centralidade do modelo (λ) da distribuição χ^2 (qui-quadrado) não central expressa a separação entre a hipótese nula (H_0 - ausência de *outliers*) e a hipótese alternativa (H_A - possível existência de *outlier(s)* em uma ou mais observações) (TEUNISSEN, 2006).

Como a separação entre as hipóteses H_0 e H_A depende da magnitude do(s) *outlier(s)*, e esta é desconhecida na prática, em geral, deve-se arbitrar um valor para o nível de significância (α), isto é, para a probabilidade de ocorrência de “falsos positivos”, e para o poder do teste (γ), ou seja, para a probabilidade de identificação correta do teste, em função dos quais se obtém o parâmetro de não centralidade do modelo correspondente (KLEIN, 2014a). Para mais detalhes sobre essa estratégia, ver, por exemplo, Baarda (1968), Kavouras (1982), Aydin e Demirel (2005), Teunissen (2006), Knight *et al.* (2010) e Klein (2014b).

Aydin e Demirel (2005) apresentam uma rotina de cálculo para encontrar o valor do parâmetro de não centralidade do modelo para distribuição χ^2 não central em função de γ , α e dos graus de liberdade do teste (h - número de *outliers* considerados na hipótese alternativa). No caso do procedimento DS, a hipótese nula (H_0) é a ausência de erros grosseiros nas observações, e para cada i -ésima observação do conjunto de dados, define-se uma hipótese alternativa (H_i) como sendo a presença de erro grosseiro nesta i -ésima observação testada. Logo, $h = 1$ para cada hipótese alternativa do DS, havendo então múltiplas hipóteses H_A , uma para cada observação testada (BAARDA, 1968; TEUNISSEN, 2006).

Na aplicação de um teste estatístico que possui apenas uma hipótese alternativa, pode-se cometer os erros do Tipo I (falsos positivos) e do Tipo II (falsos negativos). Porém, para o caso de múltiplas hipóteses alternativas, como no procedimento DS, pode ocorrer ainda o Erro Tipo III, isto é, identificar erroneamente a j -ésima observação como contendo *outlier*, quando na realidade o *outlier* está presente em outra (FÖRSTNER, 1983). Neste caso, deve-se considerar outra variável no problema: o coeficiente de correlação entre as estatísticas de teste das observações.

Seguindo esta abordagem, Yang *et al.* (2013) propõem o cálculo do poder do teste mínimo para o *Data Snooping* considerando a ocorrência do Erro Tipo III para múltiplas (n) hipóteses alternativas, ou seja, considerando a ocorrência do Erro Tipo III para todas as (n) observações testadas. De forma geral, pode-se dizer que a relação matemática entre as variáveis envolvidas na aplicação de testes estatísticos para identificação de *outliers* é bastante complexa, pois depende de diversos fatores (KLEIN, 2012; 2014a).

Assim, conforme contextualização supracitada, o objetivo do presente trabalho foi apresentar uma análise gráfica das variáveis envolvidas no CQ de dados geodésicos por meio de testes estatísticos, sendo estas: o parâmetro de não centralidade do modelo (λ); o nível de significância (α) ou analogamente o nível de confiança (NC); o poder do teste (γ) ou analogamente a probabilidade do Erro Tipo II (β); e os graus de liberdade ou número de *outliers* considerados (h). Além disso, analisar o comportamento das variáveis citadas anteriormente para o caso bidimensional, ou seja, considerando a probabilidade do Erro Tipo III (κ) entre duas hipóteses alternativas, quando o coeficiente de correlação entre estas estatísticas de teste (ρ_{ij}) deve ser considerado. Com a apresentação e análise dos gráficos, busca-se um melhor entendimento da relação matemática existente entre as variáveis envolvidas no CQ de dados geodésicos por meio de testes estatísticos.

É importante destacar que, além da notória aplicação no CQ, as variáveis estudadas neste trabalho são utilizadas em diversas outras aplicações geodésicas, como no cálculo das medidas de confiabilidade (ver, por exemplo, KNIGHT *et al.*, 2010; TEUNISSEN, 2006); no planejamento de redes geodésicas (ver, por exemplo, KUANG, 1991; SIMKOOEI, 2001; KLEIN, 2012, 2014a, 2014b); na análise de robustez da rede (ver, por exemplo, VANÍČEK *et al.*, 1990; 2001; BERBER, 2006); na análise de separabilidade das observações (ver, por exemplo, FÖRSTNER, 1983; YANG *et al.*, 2013; WANG *et al.*, 2012), na análise de covariância de vértices geodésicos (ver, por exemplo, VANÍČEK *et al.*, 1986), dentre outras.

REVISÃO TEÓRICA

Para facilitar o entendimento das variáveis estudadas neste trabalho, aqui é apresentada uma pequena revisão dos principais temas abordados, sendo estes: nível de significância; nível de confiança; poder do teste, Erro Tipo I; Erro Tipo II; Erro Tipo III; graus de liberdade do teste; parâmetro de não centralidade do modelo e coeficientes de correlação entre as estatísticas de teste.

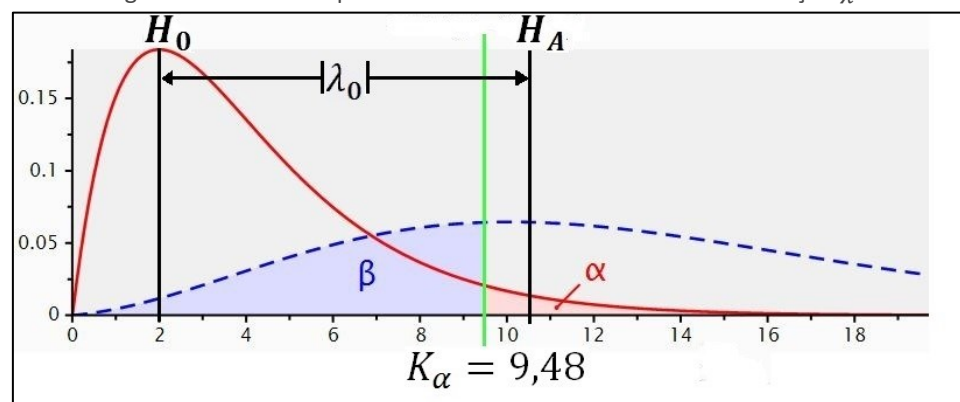
O nível de significância ou α (normalmente arbitrado como sendo $\alpha \leq 0,01$), é a probabilidade de rejeitar a hipótese nula quando esta de fato é verdadeira, ou analogamente, aceitar a hipótese alternativa quando esta de fato é falsa, ou seja, é a probabilidade de cometer o Erro Tipo I (ocorrência de falso positivo). Para o procedimento de teste DS, por exemplo, é a probabilidade de identificar um *outlier* quando na verdade não há erros grosseiros em nenhuma observação (TEUNISSEN, 2006). Em função do nível de significância, se obtém o valor crítico de teste (K_α) na distribuição de probabilidade correspondente. A hipótese nula é rejeitada quando o valor da estatística de teste considerada (w_i) excede o respectivo valor crítico.

O nível de confiança (NC) é o complemento do nível de significância ($NC = 1 - \alpha$), ou seja, é a probabilidade de não rejeitar a hipótese nula quando de fato esta é verdadeira (LARSON, 1974).

O Erro Tipo II consiste em não rejeitar a hipótese nula quando esta é falsa. No caso do DS, corresponde a não detectar erro grosseiro em nenhuma observação quando de fato pelo menos uma está contaminada (ver Figura 1). A probabilidade do Erro Tipo II é dada por β .

No caso de múltiplas hipóteses alternativas, têm-se ainda a ocorrência do Erro Tipo III, que consiste em aceitar a hipótese alternativa falsa ao invés da hipótese alternativa verdadeira. No caso do DS, corresponde em identificar uma observação errada ao invés da observação de fato contaminada. A probabilidade do Erro Tipo III é dada por κ .

Figura 1 – Teste de hipóteses entre H_0 e H_A utilizando a distribuição χ^2



Fonte: Adaptado do Software GPower 3.1.

O poder do teste (γ) é a probabilidade de o resultado do teste aceitar a hipótese alternativa correta. Para o caso de apenas uma hipótese alternativa, é dado por: $\gamma = 1 - \beta$. Para o caso de duas hipóteses alternativas ou mais, é dado por: $\gamma = 1 - \beta - \kappa$.

Logo, o nível de confiança e o poder do teste são as probabilidades de o teste conduzir a decisões corretas, ao contrário dos Erros Tipo I, II e III. É importante ressaltar que não é possível minimizar as probabilidades dos Erros Tipo I e II simultaneamente, ou seja, maximizar o poder do teste e o nível de confiança ao mesmo tempo, em função de variações no valor crítico de teste (ver a Figura 1).

Para um teste com apenas uma hipótese alternativa, o parâmetro de não centralidade do modelo, ou seja, a separação entre a hipótese nula e a hipótese alternativa (ver a Figura 1), é dada por Teunissen (2006):

$$\lambda = \nabla^T C_h^T \Sigma_y^{-1} \Sigma_v \Sigma_y^{-1} C_h \nabla \quad (1)$$

onde λ é o parâmetro de não centralidade do modelo; ∇ é o vetor de h erros grosseiros; C_h é uma matriz de dimensão $n \times h$ (onde “ n ” é o número de observações e “ h ” é o número de outliers considerados no teste); Σ_y é a matriz de covariância das observações e Σ_v é a matriz de covariância dos resíduos ajustados. Como na prática, a quantidade (dada por C_h) e a magnitude dos possíveis outliers (dada por ∇) são desconhecidas, se arbitra um valor para h , para o nível de significância e para o poder do teste, em função dos quais se obtém o parâmetro de não centralidade do modelo. Para mais detalhes sobre

esta estratégia de obtenção de λ em função de h , α e γ , ver, por exemplo, Baarda (1968), Kavouras (1982), Aydin & Demirel (2005), Teunissen (2006), Knight *et al.* (2010), Klein (2012, 2014). Em suma, o parâmetro de não centralidade expressa o nível de separação entre a hipótese nula e a hipótese alternativa, ou seja, quanto maior o valor calculado maiores as chances de o teste estatístico conduzir a decisão correta.

Förstner (1983), considerando a ocorrência do Erro Tipo III no caso de duas hipóteses alternativas, apresenta a “análise de separabilidade”. No referido trabalho, o autor apresenta o cálculo dos coeficientes de correlação entre as estatísticas de teste para o caso bidimensional, ou seja, para cada par de observações (ou par de hipóteses alternativas) considerado. No caso do DS, estes coeficientes de correlação entre as estatísticas de teste são dados por:

$$\rho_{ij} = \frac{c_i^T W \Sigma_v W c_j}{\sqrt{c_i^T W \Sigma_v W c_i} \sqrt{c_j^T W \Sigma_v W c_j}}; \quad -1 \leq \rho_{ij} \leq 1 \quad (2)$$

sendo ρ_{ij} o coeficiente de correlação entre as estatísticas de teste da i -ésima e da j -ésima observação; c_i o vetor canônico de dimensão $n \times 1$ contendo 1 no i -ésimo elemento e zero nos demais; c_j o vetor canônico de dimensão $n \times 1$ contendo 1 no j -ésimo elemento e zero nos demais; W a matriz peso das observações e Σ_v a matriz de covariância dos resíduos ajustados.

Dentro desse contexto, Förstner (1983) propõe o cálculo do poder do teste para o caso bidimensional, ou seja, o poder do teste de uma observação qualquer “ i ” considerando as probabilidades do Erro Tipo II e do Erro Tipo III para outra hipótese alternativa (referente a outra observação qualquer “ j ”), conforme as expressões (3) e (4):

$$\gamma_{ij} = 1 - \tau_{ij} = \iint_{|w_i| > c_{\alpha_0/2}, |w_i| > |w_j|} \frac{1}{2\pi \sqrt{1 - \rho_{ij}^2}} e^{-m} dw_i dw_j \quad (3)$$

$$m = \frac{1}{2(1 - \rho_{ij}^2)} \left[(w_i - \mu_i)^2 - 2\rho_{ij}(w_i - \mu_i)(w_j - \mu_j) + (w_j - \mu_j)^2 \right] \quad (4)$$

Nas expressões (3) e (4), $\tau_{ii} = \beta_{ij} + \kappa_{ij}$, isto é, a soma dos Erros Tipo II (β_{ij}) e Tipo III (κ_{ij}); $\mu_i = \delta_{0ij}$, isto é, o parâmetro de não centralidade do modelo para a distribuição normal: $\delta_0 = \sqrt{\lambda}$; $\mu_j = \rho_{ij} \delta_{0ij}$; w_i é a estatística de teste da i -ésima observação testada e w_j é a estatística de teste da j -ésima observação testada; e $c_{\alpha_0/2}$ é o valor crítico correspondente na distribuição normal para o nível de significância α_0 (teste bi-caudal). Para mais detalhes sobre esse procedimento de cálculo, ver, por exemplo, Yang *et al.* (2013); Wang *et al.* (2012); Klein *et al.* (2014b).

Os graus de liberdade considerados para o cálculo do parâmetro de não centralidade correspondem ao número de “ h ” outliers considerados na hipótese alternativa. O procedimento de teste *Data Snooping* pertence a uma única família geral de testes, formulada com base na razão de verossimilhança, denominado aqui de T_h . Para o caso do DS: $h = 1$, mas pode-se aplicar outros testes, por exemplo, $T_{h=2}$, $T_{h=3}$, $T_{h=4}$, ..., $T_{h=n-u}$ (quando $h = n - u$ é chamado de Teste Global do Ajustamento, onde “ n ” é o número de observações e “ u ” é o número de incógnitas do modelo).

Ressalta-se que os mesmos conceitos e relações apresentados anteriormente para o caso do DS são válidos para qualquer teste com “*h*” outliers simultâneos, como nível de confiança, poder do teste, Erros Tipo I, II e III. Para mais detalhes sobre este caso geral do teste da razão de máxima verossimilhança para *h* outliers, ver, por exemplo, Teunissen (2006); Knight *et al.* (2010) e Klein (2012). Uma vez que o número de variáveis envolvidas no CQ por meio de testes estatísticos é relativamente alto, na próxima sessão são apresentados alguns gráficos visando uma melhor compreensão destes conceitos por meio de análises visuais.

APRESENTAÇÃO DOS GRÁFICOS

Adaptando o algoritmo cedido por Aydin e Demirel (2005), foram gerados os gráficos contidos nas Figura 2, 3, 4 e 5 com o auxílio do *software Scilab* 5.5.2. Foram realizados dois estudos distintos, o primeiro deles é para a avaliação do comportamento do poder do teste (γ) para uma hipótese alternativa e do parâmetro de não centralidade (λ) do modelo para a distribuição χ^2 contidos nas Figura 2 e 3. Em cada gráfico gerado o nível de confiança é fixo e as curvas representam diferentes graus de liberdade (*h*) do teste estatístico (isto é, número de outliers considerados), conforme as cores indicadas nas legendas.

Figura 2 – Gráficos de λ versus γ com o NC (nível de confiança) fixo em 0,80 e 0,90

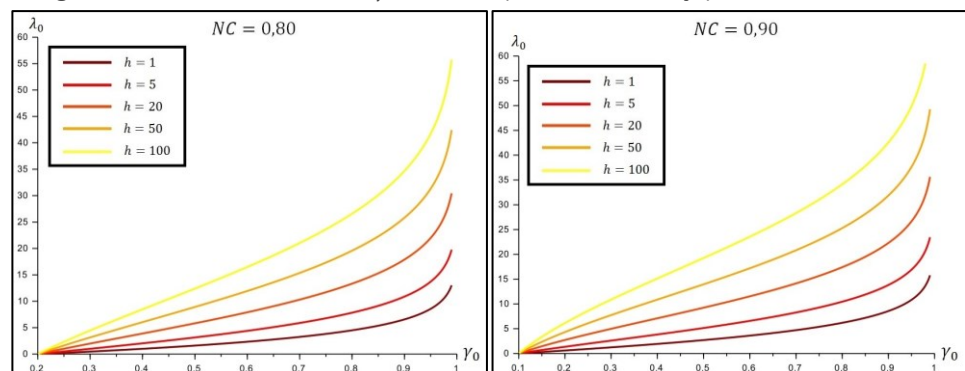
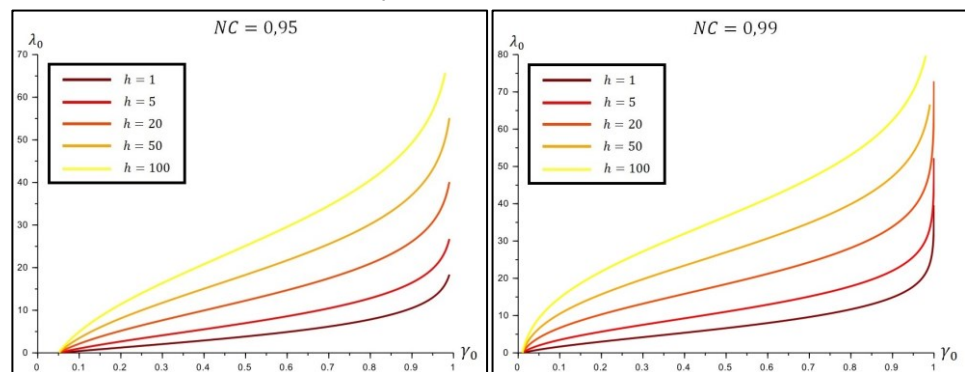


Figura 3 – Gráficos de λ versus γ com os graus de liberdade (*h*) fixos em 50 e 100



Analisando os gráficos contidos na Figura 2, percebe-se que quanto maior o número de graus de liberdade, maior é a inclinação da curva para um mesmo nível de confiança. Além disto, quanto maior a inclinação da curva, maior é o parâmetro de não centralidade, considerando o mesmo poder do teste. Por

exemplo, para $\gamma = 0,90$, $NC = 0,80$ e $h = 1$, o valor do parâmetro de não centralidade é aproximadamente $\lambda = 6,4$; enquanto para $\gamma = 0,90$, $NC = 0,80$, e $h = 100$, o valor do parâmetro de não centralidade é aproximadamente $\lambda = 34,6$. Ou seja, é mais fácil de o teste conduzir a decisão correta quando $h=100$, pois o valor para o parâmetro de não centralidade é maior.

Analisando e comparando os gráficos nas Figura 2 e 3 Figura 4 nota-se que as inclinações das curvas aumentam conforme o nível de confiança aumenta, ou seja, quanto maior o NC considerado no teste, maior deve ser o parâmetro de não centralidade para atingir o mesmo poder do teste. Por exemplo, para $\gamma = 0,90$, $h = 1$ e $NC = 0,80$, o valor do parâmetro de não centralidade é de aproximadamente $\lambda = 6,4$, enquanto para $\gamma = 0,90$, $h = 1$ e $NC = 0,99$, o valor correspondente é $\lambda \sim 15$. Em outras palavras, quando o NC aumenta o valor crítico (K_α) também aumenta, em contra partida o erro tipo I (α) diminui. A consequência disso é que o erro tipo II (β) aumentará (ver Figura 1), logo o poder do teste diminui (considerando λ fixo). Para manter o mesmo poder do teste o parâmetro de não centralidade deve ser maior para que diminua o valor de β .

Além disto, também nota-se que a curva se torna cada vez mais acentuada a partir de $\gamma > 0,8$, sendo praticamente uma reta de inclinação constante quando $\gamma < 0,80$ em todos os casos apresentados. Ressalta-se que na prática, em geral é desejado $\gamma \geq 0,8$ (intervalo de valor mais encontrado na Literatura), o que demonstra a importância deste tipo de análise. O poder do teste deve ser alto para que o CQ se justifique, em contra partida, redes com γ muito alto ($\gamma \geq 0,95$, por exemplo) são economicamente inviáveis, pois precisam de uma grande quantidade de observações para garantir tal nível de assertividade.

Em uma segunda análise, obtida a partir das Figura 4 e 5 Figura 5, os graus de liberdade foram fixados para cada gráfico enquanto nas curvas são representados os diferentes NC, conforme o esquema de cores representado nas legendas.

Figura 4 – Gráficos de λ versus γ com os graus de liberdade (h) fixos em 1 e 10

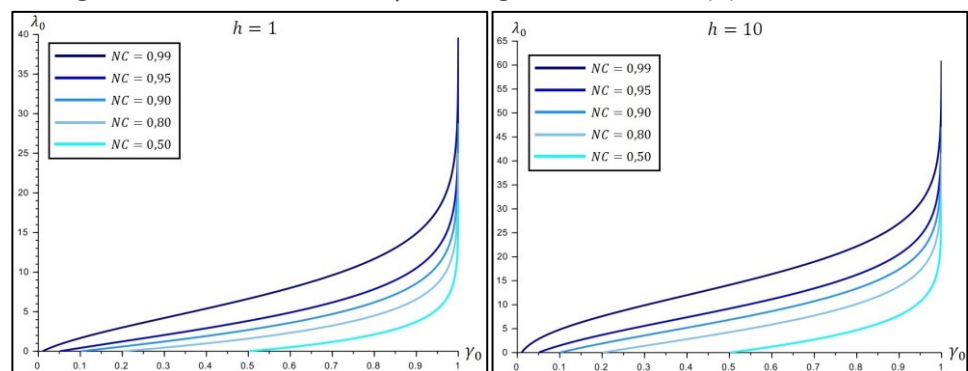
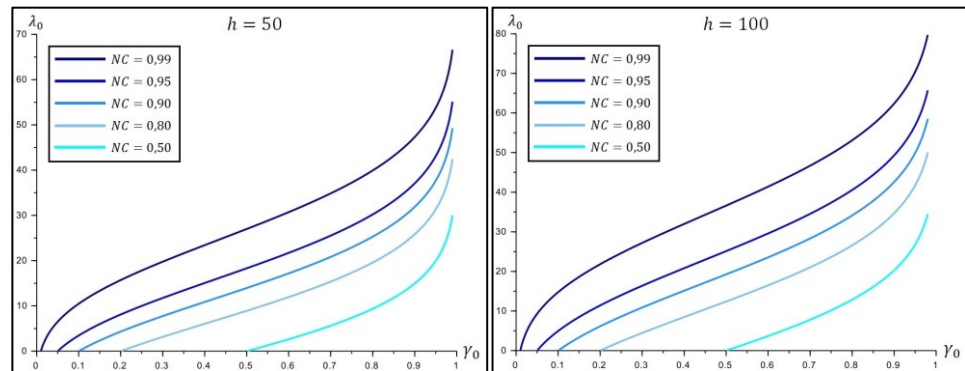


Figura 5 – Gráficos de λ versus γ com os graus de liberdade (h) fixos em 50 e 100.



Analisando os gráficos das Figura 4 e 5, é notório que o comportamento exponencial das curvas a partir do valor de $\gamma > 0,80$ se torna mais moderado conforme os graus de liberdade aumentam. Para apenas $h = 1$ outlier considerado no teste, um pequeno incremento no poder do teste a partir de $\gamma = 0,80$ gera um significativo aumento no parâmetro de não centralidade correspondente, sendo que, quanto mais próximo de 1 é o valor de γ , mais acentuado é o crescimento do parâmetro de não centralidade (λ), tendendo a uma reta vertical.

Ressalta-se que na prática, o número de outliers (simultâneos) considerados no teste será pequeno, em geral, varia de 0,1% a 1% do número total de observações, por exemplo, $h \leq 10$ quando $n = 1000$ (mais detalhes em BAARDA, 1968; KLEIN, 2014a). Enquanto o valor desejado para o poder do teste em geral será $\gamma \geq 0,80$, ou seja, justamente a região mais crítica para o parâmetro de não centralidade do modelo nas Figuras 4 e 5, o que também ressalta a importância deste tipo de análise. Novamente, para $\gamma < 0,80$, em geral, as curvas apresentadas podem ser aproximadas por uma reta.

Outro comportamento interessante é quando o NC se aproxima de 1 ($NC \geq 0,95$), pois as curvas começam a apresentar uma pequena exponencial no início do intervalo, para: $0 \leq \gamma \leq 0,15$. Este comportamento, entretanto, não é crítico, pois, embora na prática seja desejado $NC \geq 0,95$, valores para o poder do teste da ordem de $\gamma \leq 0,15$ são altamente não recomendados.

Nas Figura 6 e 7, são apresentados os resultados obtidos com a aplicação da Expressão (3) para o cálculo do poder do teste bidimensional, ou seja, considerando duas hipóteses alternativas relativas as observações "i" e "j" quaisquer. Os gráficos foram gerados com o auxílio do software Scilab 5.5.2, onde para cada gráfico é fixado o valor do nível de confiança e nas curvas são representados os diferentes valores para os coeficientes de correlação, conforme o esquema de cores representado nas legendas.

Figura 6 – Gráficos de λ versus γ com o nível confiança fixo em 0,80 e 0,90

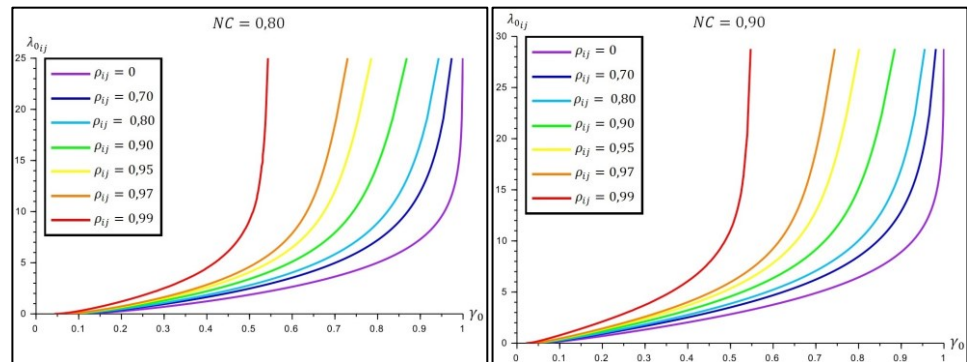
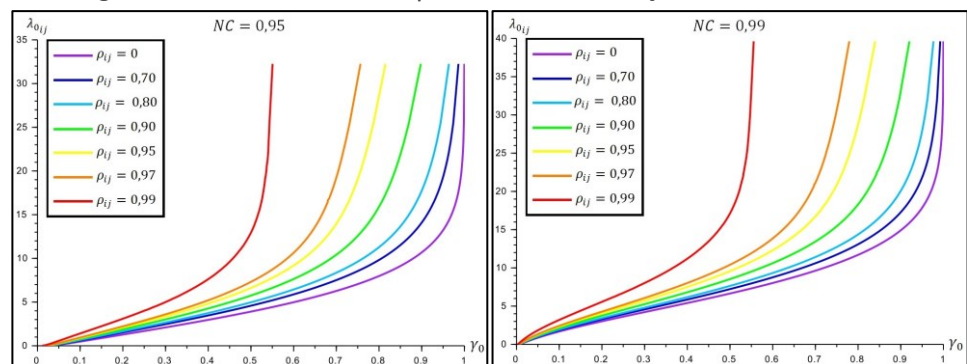


Figura 7 – Gráficos de λ versus γ com o nível confiança fixo em 0,95 e 0,99



Com a análise dos gráficos representados na Figura 6, nota-se que os maiores poderes do teste são obtidos para os menores coeficientes de correlação, ou seja, quanto maior é a correlação entre as estatísticas de teste (hipóteses alternativas) menor é o poder do teste resultante, portanto, maior é a probabilidade de ocorrência do Erro Tipo III.

Em todas as curvas representadas nos gráficos das Figura 6 e 7 o poder do teste tem um valor limite para cada coeficiente de correlação e para cada nível de confiança considerado, sendo que, a partir desses valores, o poder do teste permanecerá praticamente o mesmo, pois a curva se torna uma reta vertical. Em outras palavras, dependendo do coeficiente de correlação entre as estatísticas de teste de duas observações, pode não ser possível atingir os valores desejados para NC e γ , devendo, obrigatoriamente, reduzir os coeficientes de correlação ρ_{ij} , o que demonstra a importância deste tipo de análise na etapa de planejamento da rede.

Também é possível notar que, conforme o nível de confiança aumenta, os valores para o parâmetro de não centralidade também aumentam (fixando o poder do teste e o coeficiente de correlação), pois maior deve ser a separação entre as estatísticas de teste para alcançar o NC desejado. Por exemplo, para $NC = 0,80$, $\gamma = 0,90$ e $\rho_{ij} = 0,70$, o valor de λ é de 11,86; enquanto para $NC = 0,99$, $\gamma = 0,90$ e $\rho_{ij} = 0,70$, o valor de λ é de 17,25.

A separação entre as curvas para coeficientes de correlação entre 0 e 0,80 é pequena quando comparada com a separação entre as curvas para coeficientes de correlação a partir de 0,95 em todos os casos apresentados. Com isto, é possível afirmar que, para um valor fixo do NC, os poderes do teste resultantes para os diferentes coeficientes de correlação serão mais afetados (ou seja, menores) para os valores de $\rho_{ij} > 0,80$. Este tipo de análise também é muito útil

na etapa de planejamento da rede. Nota-se ainda que o comportamento das curvas é praticamente igual em todos os casos apresentados, variando apenas o intervalo para o parâmetro de não centralidade do modelo correspondente.

Uma quarta análise, envolvendo o Erro Tipo III para duas hipóteses alternativas, em função dos coeficientes de correlação das estatísticas de teste, do nível de confiança e do poder do teste, é apresentada nas Figuras 8 e 9. Ressalta-se que este tipo de gráfico é inédito na literatura relacionada ao CQ de dados geodésicos.

Figura 8 – Gráficos de κ_{ij} versus NC com o poder de teste fixo em 0,70 e 0,80

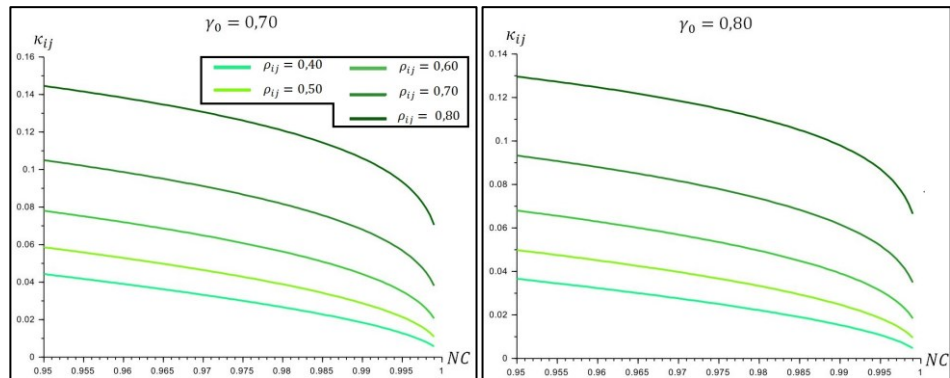
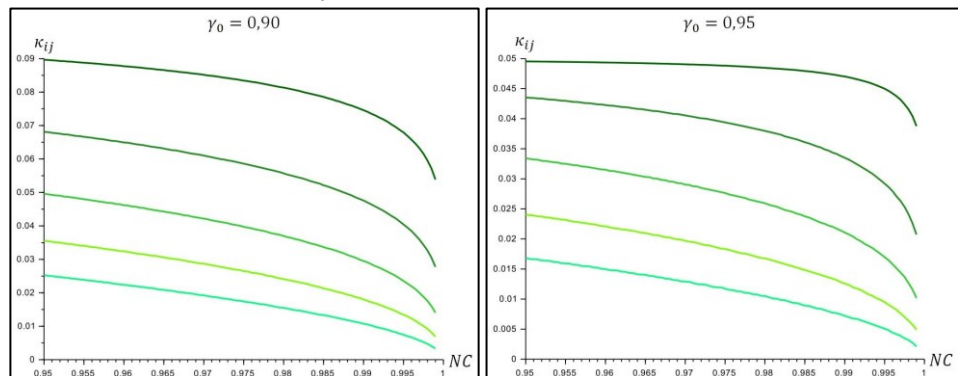


Figura 9 – Gráficos de κ_{ij} versus NC com o poder do teste fixo em 0,90 e 0,95



Analisando os gráficos das Figuras 8 e 9, nota-se que a probabilidade de ocorrência do Erro Tipo III para duas hipóteses alternativas é proporcionalmente maior conforme o poder do teste aumenta, mantendo o NC e ρ_{ij} fixos. Por exemplo, para $\gamma_0 = 0,95$, $\rho_{ij} = 0,80$ e $NC = 0,95$, a probabilidade do Erro Tipo III é de aproximadamente $\kappa_{ij} = 0,05$ (sendo este o valor máximo possível, pois $\kappa_{ij} = 1 - \gamma_0 - \beta_{ij}$, logo $\beta_{ij} \sim 0$). Para $\gamma_0 = 0,70$, $\rho_{ij} = 0,80$ e $NC = 0,95$; a probabilidade do Erro Tipo III é de aproximadamente $\kappa_{ij} = 0,145$, que corresponde a aproximadamente metade da soma dos Erros Tipo II e III, que neste caso deve ser igual a $1 - \gamma_0 = 0,30$.

Em todos os casos apresentados, conforme o NC aumenta (mantendo a correlação e o poder do teste fixos), o Erro Tipo III diminui, sendo esse comportamento mais acentuado a partir de $NC \geq 0,99$. Como na prática, em geral o valor adotado para o nível de confiança é $NC \geq 0,99$, ressalta-se a importância da escolha adequada do nível de confiança (ou de significância) do teste, uma vez que pequenas alterações a partir de $NC \geq 0,99$ podem resultar

em significativas reduções do Erro Tipo III (e, conseqüentemente, significativos aumentos para o Erro Tipo II, para um dado poder do teste fixo). A escolha correta do nível de significância é um tema de intenso estudo na comunidade afim, podendo-se citar, por exemplo, as discussões expostas em LEHMANN (2012).

Analisando ainda as Figuras 8 e 9, nota-se que as inclinações das curvas diminuem conforme se aumenta o poder do teste, ou seja, menor é a taxa de decaimento do Erro Tipo III conforme o valor de γ_0 aumenta.

CONCLUSÕES

Neste trabalho, foi apresentada uma breve revisão teórica a respeito das principais variáveis envolvidas no controle de qualidade de dados geodésicos por meio de testes estatísticos; sendo estas: nível de significância; nível de confiança; poder do teste; Erros Tipo I, II e III; graus de liberdade (número de *outliers* considerados); parâmetro de não centralidade do modelo e coeficientes de correlação entre as estatísticas de teste.

Visando um melhor entendimento das relações matemáticas entre estas variáveis, foram gerados gráficos considerando os cenários de uma e duas hipóteses alternativas, sendo alguns destes de caráter inédito dentro da literatura relacionada ao tema. Dentre as principais conclusões obtidas, ressalta-se que, em geral, as maiores variações ocorrem justamente nos valores usualmente adotados na prática, como $NC \geq 0,99$; $\gamma_0 \geq 0,80$ e $h \leq 10$.

Além disso, o poder do teste apresenta um valor limite para cada coeficiente de correlação e para cada nível de confiança considerado, tendendo a uma reta vertical a partir destes valores. Estes tipos de análises fornecem informações muito úteis não apenas no CQ como na etapa de planejamento da rede, como por exemplo, estipulando valores limites para os coeficientes de correlação entre as estatísticas de teste em função do poder do teste desejado e do NC arbitrado.

Outra conclusão relevante obtida foi que, pequenas alterações no nível de confiança, a partir de $NC \geq 0,99$, resultam em grandes reduções na ocorrência do Erro Tipo III, ou analogamente, grandes aumentos na ocorrência do Erro Tipo II, considerando um poder do teste fixo. Desta forma, a escolha do nível de confiança (ou de significância) não está intimamente relacionada somente com a probabilidade do Erro Tipo I, como também dos Erros Tipo II e III.

Como sugestões para trabalhos futuros, sugerem-se investigações destas variáveis em cenários mais complexos, envolvendo mais do que uma ou duas hipóteses alternativas, bem como, a sua influência em outros campos da Geodésia, como no cálculo das medidas de confiabilidade e na análise de robustez da rede. Além disso, estudos de como aumentar o parâmetro de não centralidade mudando a geometria e/ou o modelo estocástico da rede, adicionando novas observações, ou seja, estratégias para diminuir os coeficientes de correlação dos resíduos e das observações.

Graph analysis of the variables involved in the quality control of geodetic data through statistical testing

ABSTRACT

In geodesy, quality control (QC) is one of the most important tasks, where it's decided whether the results obtained are validated or not. Within this context, statistical tests are still the most used tool for the detection and identification of outliers, among them the famous Data Snooping proposed by Baarda. Therefore, the understanding of the mathematical relationships between the variables involved in the statistical tests is of utmost importance not only for QC applications, but also in other fields of Geodesy such as robustness analysis and network design. Within this context, the goal of this work is to present a graph analysis of the variables involved in the QC of geodetic data through statistical tests. To reach it, some graphs involving the considered variables were presented, being it: confidence (or significance) level; test power; Type I, II and III Errors; correlation coefficients between the test statistics; degrees of freedom and non-central model parameter. Based on the results obtained, some considerations, conclusions and recommendations for future works are addressed.

KEYWORDS: Quality control. Statistical testing. Probability levels.

AGRADECIMENTO

Este trabalho faz parte de um Projeto de Pesquisa com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (processo n. 305599 / 2015-1). O quinto autor também agradece ao CNPq pela Bolsa de Produtividade em Pesquisa (Processo 309399/2014-9).

REFERÊNCIAS

ALMAGBILE, A.; WANG, J.; DING, W.; KNIGHT, N. **Sensitivity analysis of multiple fault test and reliability measures in integrated GPS/INS systems.** In: 7th International Symposium on Mobile Mapping Technology, Cracow, Poland, 2011.

AYDIN, C.; DEMIREL, H. Computation of Baarda's lower bound of the non-centrality parameter. **Journal of Geodesy**, v. 78, n. 7, p. 437-441, 2005. <https://doi.org/10.1007/s00190-004-0406-1>.

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

BERBER, M. **Robustness Analysis of Geodetic Networks.** 2006. 121 p. Ph.D. Dissertation – Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada, 2006.

FÖRSTNER, W. **Reliability and discernability of extended Gauss-Markov models.** In: Seminar on mathematical models to outliers and systematic errors, Deutsche Geodätische Kommission, Series A, no. 98. Munich, Germany, p. 79–103, 1983.

GUI, Q.; LI, X.; GONG, Y.; LI, B.; LI, G. A Bayesian unmasking method for locating multiple gross errors based on posterior probabilities of classification variables. **Journal of Geodesy**, v. 85, n. 4, p. 191–203, 2011. <https://doi.org/10.1007/s00190-010-0429-8>.

GUO, J. F.; OU, J. K.; WANG, H. Quasi-accurate detection of outliers for correlated observations. **Journal of Surveying Engineering**, v. 133, n. 3, p. 129–133, 2007. [https://doi.org/10.1061/\(ASCE\)0733-9453\(2007\)133:3\(129\)](https://doi.org/10.1061/(ASCE)0733-9453(2007)133:3(129)).

KAVOURAS, M. **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, Fredericton, Canada, 1982.

KLEIN, I. *et al.* An approach to identify multiple outliers based on sequential likelihood ratio tests. **Survey Review**, v. 49, n. 357, p. 449-457, 2017. <https://doi.org/10.1080/00396265.2016.1212970>.

KLEIN, I. **Proposta de um novo método para o planejamento de redes geodésicas**. Porto Alegre: UFRGS, 2014. 134f. Tese (Doutorado em Sensoriamento Remoto) – Programa de pós-graduação em sensoriamento remoto, Universidade Federal Do Rio Grande Do Sul, Porto Alegre, 2014a.

KLEIN, I.; MATSUOKA, M. T. ; GUZATTO, M. P. ; De SOUZA, S. F. ; VERONEZ, M. R. . On evaluation of different methods for quality control of correlated observations. **Survey Review**, v. 47, n. 340, p. 28-35, 2015. <http://dx.doi.org/10.1179/1752270614Y.0000000089>.

KLEIN, I.; MATSUOKA, M. T.; GUZATTO, M. P. Como estimar um poder do teste mínimo e valores limites para o intervalo de confiança do *Data Snooping*. **Boletim de ciências geodésicas**, v. 21, n. 1, p. 26-42, 2014b. <http://dx.doi.org/10.1590/S1982-21702015000100003>.

KLEIN, I.; MATSUOKA, M. T.; SOUZA, S. F.; COLLISCHONN, C. Planejamento de Redes Geodésicas Resistentes a Múltiplos Outliers. **Boletim de Ciências Geodésicas**, v. 18, n. 3, p. 480–507, 2012. <http://dx.doi.org/10.1590/S1982-21702012000300008>.

KNIGHT, N. L.; WANG, J.; RIZOS, C. Generalised Measures Of Reliability For Multiple Outliers. **Journal of Geodesy**, v. 84, n. 10, p. 625–635, 2010. <https://doi.org/10.1007/s00190-010-0392-4>.

KOCH, K. R. **Parameter estimation and hypothesis testing in linear models**. 2. ed. Berlin: Springer, 1999. 309p.

KUANG, S. L. **Optimization and design of deformation monitoring schemes**, Fredericton: UNB, 1991, 196p. Dissertation - Department of Surveying Engineering, , University of New Brunswick, Fredericton, 1991.

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. <https://doi.org/10.1007/s00190-012-0569-0>.

LEHMANN, R.; VOß-BÖHME, A. On the statistical power of Baarda's outlier test and some alternative. **Journal of Geodetic Science**, v. 7, n. 1, p. 68-78, 2017. <https://doi.org/10.1515/jogs-2017-0008>.

LEHMANN, Rüdiger. 3 σ -Rule for Outlier Detection from the Viewpoint of Geodetic Adjustment. **Journal of Surveying Engineering**, v. 139, n. 4, p. 157-165, 2013b. [https://doi.org/10.1061/\(ASCE\)SU.1943-5428.0000112](https://doi.org/10.1061/(ASCE)SU.1943-5428.0000112).

LEHMANN, Rüdiger. On the formulation of the alternative hypothesis for geodetic outlier detection. **Journal of geodesy**, v. 87, n. 4, p. 373-386, 2013a. <https://doi.org/10.1007/s00190-012-0607-y>.

SIMKOOEI, A. A. Strategy for designing geodetic networks with high reliability and geometrical strength. **Journal of Surveying Engineering**, v. 127 n. 3, p. 104–117, 2001. [https://doi.org/10.1061/\(ASCE\)0733-9453\(2001\)127:3\(104\)](https://doi.org/10.1061/(ASCE)0733-9453(2001)127:3(104)).

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

VANÍČEK, P.; CRAYMER, M. R.; KRAKIWSKY, E. J. **Geodesy: The Concepts**. 2. ed. Amsterdam: North Holland/Elsevier, 1986. 714p.

VANÍČEK, P.; CRAYMER, M. R.; KRAKIWSKY, E. J. Robustness analysis of geodetic horizontal networks. **Journal of Geodesy**, v. 75, n. 4, p. 199–209, 2001. <https://doi.org/10.1007/s001900100162>.

VANÍČEK, P.; KRAKIWSKY, E. J.; CRAYMER, M. R.; GAO, Y.; ONG, P. J. **Robustness analysis**. Final contract report, Department of Surveying Engineering Technical Report No. 156, University of New Brunswick, Fredericton, New Brunswick, Canada, 1990. 116p.

WANG, Jinling; KNIGHT, Nathan L. New outlier separability test and its application in GNSS positioning. **Journal of Global Positioning Systems**, v. 11, n. 1, p. 46-57, 2012. <https://doi.org/10.5081/jgps.11.1.46>.

YANG, L.; WANG, J.; KNIGHT, N.; SHEN, Y. Outlier separability analysis with a multiple alternative hypotheses test. **Journal of Geodesy**, v. 85, n. 6, p. 591–604, 2013. <https://doi.org/10.1007/s00190-013-0629-0>.

Recebido: 20 set. 2017

Aprovado: 04 jul. 2018

DOI: 10.3895/rbgeo.v6n3.7097

Como citar: GUZATTO, M. P.; KLEIN, I.; MATSOUKA, M. T.; ROFATTO, V. F.; VERONEZ, M. R. . Análise gráfica das variáveis do controle de qualidade de dados geodésicos por meio de testes estatísticos. R. bras. Geom., Curitiba, v. 6, n. 3, p. 194-209, jul/set. 2018. Disponível em: < <https://periodicos.utfpr.edu.br/rbgeo>>. Acesso em: XXX.

Correspondência:

Matheus Pereira Guzatto

Rua Bolonha, 115, AP02, CEP 88132-201, Palhoça, Santa Catarina, Brasil.

Direito autoral: Este artigo está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.

