

UNIVERSIDADE FEDERAL DE UBERLÂNDIA *CAMPUS* MONTE CARMELO
INSTITUTO DE GEOGRAFIA – IG
CURSO DE ENGENHARIA DE AGRIMENSURA E CARTOGRÁFICA

JOÃO BATISTA SILVA NETO

**UMA ABORDAGEM BASEADA NO TESTE DE RAZÃO DE MÁXIMA
VEROSSIMILHANÇA PARA DETECÇÃO DE DESLOCAMENTOS VERTICAIS
POR GNSS**

MONTE CARMELO

2021

JOÃO BATISTA SILVA NETO

**UMA ABORDAGEM BASEADA NO TESTE DE RAZÃO DE MÁXIMA
VEROSSIMILHANÇA PARA DETECÇÃO DE DESLOCAMENTOS VERTICAIS
POR GNSS**

Trabalho de Conclusão de Curso, como exigência parcial para obtenção do Título de bacharel em Engenharia de Agrimensura e Cartográfica da Universidade Federal de Uberlândia, *campus* Monte Carmelo.

Orientador: Prof. Dr. Vinicius Francisco Rofatto

MONTE CARMELO

2021

JOÃO BATISTA SILVA NETO

**UMA ABORDAGEM BASEADA NO TESTE DE RAZÃO DE MÁXIMA
VEROSSIMILHANÇA PARA DETECÇÃO DE DESLOCAMENTOS VERTICAIS
POR GNSS**

Trabalho de Conclusão de Curso como exigência parcial para obtenção do Título de bacharel em Engenharia de Agrimensura e Cartográfica da Universidade Federal de Uberlândia, *campus* Monte Carmelo.

Orientador: Prof. Dr. Vinicius Francisco Rofatto

Aprovado em: ____/____/____

BANCA EXAMINADORA

Prof. Dr. Vinicius Francisco Rofatto (Orientador)
Universidade Federal de Uberlândia (IG/UFU)

Prof. Dr. Marcelo Tomio Matsuoka
Universidade Federal de Uberlândia (IG/UFU)

Prof. Dr. Wagner Carrupt Machado
Universidade Federal de Uberlândia (FECIV/UFU)

MONTE CARMELO

2021

“O mundo não está interessado nos temporais que você encontrou. Ele quer saber se você trouxe o navio. ”

(William McFee)

RESUMO

Deslocamentos referem-se à variação das coordenadas que um corpo sofre em relação a sua posição inicial. Assim, qualquer objeto natural ou artificial está sujeito a variações no espaço e no tempo. Neste trabalho propomos um método alternativo baseado no teste de razão de máxima verossimilhança para a detecção de deslocamentos entre campanhas de observações geodésicas. O método proposto foi aplicado em um problema de deslocamento vertical simulado em campo. Os experimentos foram realizados usando uma série de campanhas de desníveis estimados entre dois pontos por meio do posicionamento relativo. O desenho experimental foi desenvolvido considerando uma linha-base de 148 m e outra de 4,5 Km. Para cada um desses casos, testou-se também o efeito de usar receptor de simples frequência (L1) e dupla-frequência (L1 e L2). Os resultados sugerem que deve se evitar para a análise de deslocamento o uso de receptores de dupla frequência para linha de base curta. Além disso, o aumento da linha de base no caso de utilizar dados somente L1, torna-se mais difícil a detecção de deslocamentos.

Palavras chave: Teste de Razão de Máxima Verossimilhança, Deslocamento, Método dos mínimos quadrados, Teste de Hipótese.

SUMÁRIO

1 INTRODUÇÃO	6
1.2 OBJETIVOS	7
1.3 JUSTIFICATIVA	7
2 AJUSTAMENTO DE OBSERVAÇÕES GEODÉSICAS	7
2.1 Método dos mínimos quadrados (MMQ) em sistemas lineares	7
2.2 MMQ e solução de máxima verossimilhança.....	9
3 CONTROLE DE QUALIDADE DE OBSERVAÇÕES GEODÉSICAS	10
3.1 Teste de hipóteses.....	11
3.2 Teste global do ajustamento ($q = n - u$).....	14
3.3 <i>Data snooping</i> ($q = 1$)	15
4 MÉTODO PROPOSTO	17
5 MATERIAL E MÉTODOS	21
5.1 Área de estudo	21
5.2 Desenho Experimental	22
7 CONSIDERAÇÕES FINAIS.....	26
REFERÊNCIAS BIBLIOGRÁFICAS	28

1 INTRODUÇÃO

Deslocamentos referem-se à variação das coordenadas que um corpo sofre em relação a sua posição inicial. Assim, qualquer objeto natural ou artificial está sujeito a estas variações no espaço e no tempo. A determinação, quantificação e interpretação destas variações são os principais objetivos do monitoramento de estruturas, uma vez que, fornecem uma gama variada de dados, que podem prevenir desastres, sendo úteis à segurança de vidas humanas e de estruturas de engenharia.

Fazan (2010), define o termo deslocamento como a translação ou rotação de um ponto ou corpo rígido que podem ser determinados em relação a um referencial fixo. Para que seja possível avaliar a detecção de deslocamentos, as observações de uma mesma época e a solução final de cada época devem ser obtidas de forma independente, traduzindo em soluções não correlacionadas.

As indicações de deslocamentos são, a princípio, as diferenças de coordenadas de uma época com relação à outra, as quais devem ser analisadas estatisticamente para decidir se as diferenças de coordenadas representam dentro de um determinado nível de significância, os deslocamentos propriamente ditos ou as flutuações probabilísticas das observações.

Savšek-safić et al (2006), destaca que para a identificação do deslocamento de um ponto por meio de observações geodésicas, é necessário escolher os pontos de referência. Os pontos característicos do objeto são testados quanto aos deslocamentos. De acordo com a precisão exigida na determinação do deslocamento do ponto, a execução das observações deve ser realizada com cuidado e com as ferramentas adequadas, seguindo as abordagens observacionais padrão.

O processo constante da evolução tecnológica, tem colocado à disposição dos profissionais de geociências, novas técnicas e ferramentas para controle e para o monitoramento de estruturas, sejam elas naturais ou artificiais (TEIXEIRA, 2001). Scaioni et al (2018), ressaltam que além disso, as metodologias para processamento e integração de dados foram aprimoradas para analisar diferentes observações em um contexto espacial e temporal mais amplo.

Neste trabalho, propõe-se um método alternativo baseado no teste de razão de máxima verossimilhança para a detecção de deslocamento entre campanhas de observações geodésicas.

1.2 OBJETIVOS

O objetivo geral deste trabalho é propor um método alternativo para detecção de deslocamentos. O método proposto é baseado no teste de razão de máxima verossimilhança.

Deste modo, para alcançar o objetivo geral apresentado, os objetivos específicos são:

- Investigar a capacidade do método proposto em detectar deslocamentos;
- Avaliar o método proposto em problemas de deslocamentos verticais.

1.3 JUSTIFICATIVA

O monitoramento de deslocamentos em estruturas artificiais ou naturais tem um papel fundamental para garantir suas condições de segurança e/ou manter suas funções operacionais. Portanto, este monitoramento quando mal projetado e/ou mal executado, ou mesmo a ausência deste, pode comprometer a segurança das pessoas e gerar prejuízos econômicos e socioambientais. Como exemplo mais recente, pode-se citar a barragem em Brumadinho/MG, que entrou em colapso e se rompeu deixando um rastro de destruição, contaminação e mortes. Como consequência, é necessária uma vigilância cuidadosa com o objetivo de detectar possíveis situações críticas das estruturas.

2 AJUSTAMENTO DE OBSERVAÇÕES GEODÉSICAS

Segundo DALMOLIN (2002) o ajustamento de observações é um ramo da matemática aplicada que tem por objetivo apresentar solução única para problemas onde o número de observações é superabundante e o sistema de equações lineares é inconsistente. A inconsistência do sistema de equações é devido as flutuações probabilísticas das observações e faz com que um determinado subconjunto de dados proporcione valores diferentes de outro subconjunto. A solução única para estes problemas é dada, em geral, pelo Método dos Mínimos Quadrados (MMQ) (GEMAEL, 1994).

2.1 Método dos mínimos quadrados (MMQ) em sistemas lineares

Geralmente, temos a nossa disposição uma quantidade de medidas (ou observações) que são maiores do que o mínimo necessário para resolver um sistema de equações que descreve

um certo problema de interesse (ou seja, o número de observações é maior que o número de parâmetros incógnitos do modelo considerado). Com efeito, o sistema não permite uma solução única e exata, ou seja, o modelo que envolve os sistemas de equações é inconsistente. Porém, ao adicionar um vetor de erros no sistema de equações, ou seja, admitindo que as observações possuem erros, este sistema passa a ser matematicamente consistente, porém, com infinitas soluções (DALMOLIN, 2002). Para a solução única desse sistema redundante e agora consistente, o critério mais usual é o método dos mínimos quadrados (MMQ), ou seja, aquela solução para os parâmetros incógnitos que minimiza a soma dos quadrados dos erros, ponderados pelos respectivos pesos das observações (KLEIN, 2012).

Para o caso em que somente erros aleatórios estão presentes, o modelo matemático pode ser descrito da seguinte forma:

$$y_{nx1} = A_{nx1} x_{nxu} + e_{nx1} \quad (1)$$

Onde:

y_{nx1} é o vetor das n observações ou medidas realizadas (dados originais);

A_{nx1} é a matriz dos coeficientes das equações lineares (matriz *design*);

x_{nxu} é o vetor dos u parâmetros incógnitos do modelo a serem estimados;

e_{nx1} é o vetor dos n erros aleatórios, admitindo a inconsistência do sistema.

O vetor dos erros aleatórios e_{nx1} , por meio de uma operação algébrica em (1), pode ser expresso como (KLEIN, 2012):

$$e_{nx1} = y_{nx1} - A_{nx1} x_{nxu} \quad (2)$$

No método dos mínimos quadrados busca-se como estimativa para o vetor dos parâmetros x_{nxu} aquela cuja soma do quadrado dos erros aleatórios apresenta valor mínimo, ou seja:

$$(y_{nx1} - A_{nxu} x_{ux1})_{1 \times n}^T (y_{nx1} - A_{nxu} x_{ux1})_{n \times 1} = \text{Mínimo} \quad (3)$$

onde o sobrescrito T denota a operação de transposição matricial, e o produto interno entre um vetor e o seu transposto expressa o quadrado de sua norma (ou módulo).

Considerando agora uma matriz peso para as observações, simétrica e positiva-definida, designada por $W_{n \times n}$, o critério expresso em (3) se torna:

$$(y_{nx1} - A_{nxu}x_{ux1})_{1xn}^T W_{nzn} (y_{nx1} - A_{nxu}x_{ux1})_{nx1} = \text{Mínimo} \quad (4)$$

onde os elementos da matriz W_{nzn} podem ser escolhidos para ponderar a influência de observações específicas sobre a estimativa do vetor dos parâmetros x_{ux1} .

A solução encontrada para a expressão (4) é dada por:

$$\hat{x}_{ux1} = (A_{uxn}^T W_{nzn} A_{nxu})^{-1} A_{uxn}^T W_{nzn} y_{nx1} \quad (5)$$

com os erros estimados dados por:

$$\hat{e}_{ux1} = y_{nx1} - A_{nxu} \hat{x}_{ux1} \quad (6)$$

Maiores detalhes podem ser obtidos em (Leick, 2004), (Klein, 2012) e (Gemaël; Machado; Wandresen, 2015).

2.2 MMQ e solução de máxima verossimilhança

Segundo Papoulis (1991), dado o resultado de um experimento (como por exemplo, o vetor das observações y), o objetivo do método de máxima verossimilhança é estimar o vetor dos parâmetros x com base no vetor das observações y , ou seja, estimar o valor mais provável de que resultou no observado.

Denotando por $p(y|x)$ a função densidade de probabilidade (fdp) multivariada, e considerando o vetor y fixo (dado ou observado), é função apenas de x e é chamada de função de verossimilhança de x (BREIMAN, 1973).

Denotando esta função de verossimilhança do vetor dos parâmetros por $L(x) = p(y|x)$, o princípio da estimação por máxima verossimilhança é maximizá-la, ou seja (LARSON, 1974):

$$\max_x L(x) \quad (7)$$

Assim sendo, no método de máxima verossimilhança, é escolhido como estimativa para o vetor dos parâmetros x , aquele que maximiza a probabilidade $L(x) = p(y|x)$, para um dado vetor das observações y . Desta forma, deve-se conhecer ou estipular a função densidade de probabilidade $p(y|x)$ de y (LARSON, 1974).

Considerando o caso em que a fdp $p(y|x)$ segue distribuição normal multivariada, e que o vetor das observações y é normalmente distribuído com média $E = \{y\}$ e com matriz de covariância Σ_y , a fdp torna-se (TEUNISSEN, 2003):

$$p(y|x) = (2\pi)^{\frac{-n}{2}} |\Sigma_y|^{-\frac{1}{2}} e^{\left\{\frac{-1}{2} (y-Ax)^T \Sigma_y^{-1} (y-Ax)\right\}} \quad (8)$$

Onde n é a dimensão do vetor das observações y e $|\Sigma_y|$ é o determinante de sua matriz de covariância. Aplicando o logaritmo natural na expressão (15), resulta em:

$$\ln p(y|x) = -\frac{n}{2} \ln |2\pi| - \frac{1}{2} \ln |\Sigma_y| - \frac{1}{2} (y - Ax)^T \Sigma_y^{-1} (y - Ax) \quad (9)$$

Como o vetor das observações y é fixo, os termos $-\frac{n}{2} \ln |2\pi|$ e $-\frac{1}{2} \ln |\Sigma_y|$ são constantes para qualquer vetor dos parâmetros x e podem ser negligenciados.

Deste modo, maximizar a função $\ln p(y|x)$ consiste em maximizar o termo $-\frac{1}{2} (y - Ax)^T \Sigma_y^{-1} (y - Ax)$, o que é equivalente a minimizar apenas o termo quadrático $(y - Ax)^T \Sigma_y^{-1} (y - Ax)$, desconsiderando o sinal negativo e o escalar $\frac{1}{2}$.

Este termo corresponde a norma quadrática que se busca minimizar no MMQ. Portanto, nos casos em que o vetor das observações y segue distribuição normal multivariada, com média $E\{y\} = Ax$ e matriz de covariância Σ_y , e a matriz peso é escolhida como Σ_y^{-1} , o estimador de máxima verossimilhança para o vetor dos parâmetros é o mesmo estimador do método dos mínimos quadrados.

3 CONTROLE DE QUALIDADE DE OBSERVAÇÕES GEODÉSICAS

Considerando que as observações podem estar contaminadas não apenas pelos inevitáveis erros aleatórios, inerentes ao processo de mensuração, mas também por erros sistemáticos e/ou grosseiros, a etapa de detecção e identificação de erros não aleatórios nas observações pode ser realizada antes ou depois do ajustamento por mínimos quadrados ser realizado (KAVOURAS, 1982).

Do ponto de vista estatístico, observações com erros grosseiros não podem ser consideradas. Dessa forma, é importante ter técnicas para detecção e identificação desses erros.

A redundância de dados possibilita a aplicação de controle de qualidade nos resultados do ajustamento.

3.1 Teste de hipóteses

O teste de uma hipótese consiste em amostrar uma população, e, com base nestas amostras decidir entre rejeitar ou não a hipótese estabelecida sobre a função densidade de probabilidade (fdp) da população em estudo (KLEIN, 2012).

Como o teste é baseado em probabilidades, sempre há uma possibilidade de chegar a uma conclusão equivocada. Ao realizar um teste de hipóteses, dois tipos de erros são possíveis: tipo I e tipo II. Os riscos desses dois erros estão inversamente relacionados e são determinados pelo nível de significância e o poder do teste.

- Erro tipo I: probabilidade de rejeitar a hipótese nula (H_0) quando esta é de fato verdadeira.
- Erro tipo II: probabilidade de aceitar a hipótese nula (H_0) quando na verdade ela é falsa.

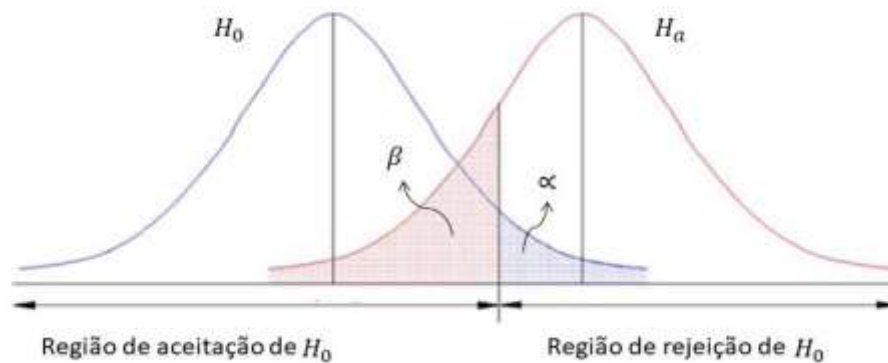
A probabilidade de cometer o erro tipo I é dado por α (nível de significância do teste), a probabilidade de cometer o erro do tipo II é dado por β . E considerando H_0 (hipótese nula) a hipótese que se deseja testar, tem-se a definição na equação (17) de (MAGALHÃES; LIMA, 2005):

$$\begin{cases} \alpha = P(\text{erro tipo I}) = P(\text{rejeitar } H_0 / H_0 \text{ verdadeira}) \\ \beta = P(\text{erro tipo II}) = P(\text{não rejeitar } H_0 / H_0 \text{ falsa}) \end{cases} \quad (10)$$

Pela definição do erro tipo II (não rejeitar a hipótese nula quando de fato ela é falsa), para se calcular a probabilidade β de cometer o erro tipo II, é necessário ainda definir uma segunda hipótese, que contradiz as suposições formuladas na hipótese nula, sendo esta última denominada de hipótese alternativa (LARSON, 1974).

Magalhães e Lima (2005), ressaltam que para avaliar a eficácia do teste deve-se considerar as duas hipóteses: H_0 (hipótese nula) e H_a (hipótese alternativa). De modo que o espaço amostral é dividido em uma região de aceitação de H_0 e em uma região de rejeição de H_0 , ou de aceitação de H_a (KLEIN, 2012a). As probabilidades do erro tipo I e do erro tipo II, a partir da região de aceitação e região de rejeição de H_0 , são apresentadas na Figura 1.

Figura 1 – Interpretação gráfica dos erros tipo I e tipo II



Fonte: Adaptado de Ghilani e Wolf (2006).

A Figura 1 ilustra as probabilidades de aceitação e rejeição de (H_0) , e demonstra que aumentando a região de aceitação de H_0 , diminui-se α (probabilidade do erro tipo I), entretanto, aumenta-se β (probabilidade do erro tipo II).

Considerando uma hipótese nula (H_0) e uma hipótese alternativa (H_a) para os parâmetros o critério de teste utilizado será a razão de verossimilhança, definida por (BREIMAN, 1973):

$$\lambda(y) = \frac{\text{máx } p(y|H_0)}{\text{máx } p(y|H_a)} \quad (11)$$

Onde:

$\text{máx } p(y|H_0)$ - corresponde ao máximo da função densidade de probabilidade do vetor das observações y sobre (H_0) ;

$\text{máx } p(y|H_a)$ - corresponde ao máximo da função densidade de probabilidade do vetor das observações y sobre (H_a) .

Como a hipótese nula deve ser formulada de forma que seu espaço amostral esteja contido no espaço amostral da hipótese alternativa, a razão de verossimilhança nunca apresentará valor superior a 1, ou seja, $0 \leq \lambda(y) \leq 1$ (BREIMAN, 1973).

Considerando que somente erros de natureza aleatória contaminam as observações, a hipótese nula pode ser formulada conforme equação (19), dada por (TEUNISSEN, 2006):

$$H_0: E\{y\} = AX \quad (12)$$

Onde:

$E\{. \}$: operador da esperança matemática;

y : vetor das observações;

A : matriz Jacobiana;

X : vetor dos parâmetros desconhecidos.

Sobre H_0 assume-se que somente erros aleatórios contaminam as observações, e assim o método de ajustamento por mínimos quadrados se torna o melhor estimador linear imparcial dos parâmetros incógnitos. No entanto, as observações podem estar contaminadas por erros não aleatórios, assim as propriedades do MMQ não são satisfeitas. Desta forma, introduz uma hipótese alternativa que contradiz a hipótese nula (TEUNISSEN, 2006):

$$H_a: E\{y\} = AX + C\nabla \quad (13)$$

Onde:

C : define o modelo de erro adotado e deve ser estipulado de acordo com a matriz *design*;

∇ : vetor contendo as observações desconhecidas contaminadas por *outliers*.

O vetor dos erros (∇) deve estar compreendido entre $1 \leq q \leq n - u$, pois não se pode detectar um número de erros q maior do que o número de graus de liberdade. (KLEIN, 2012).

Assim, para H_0 e H_a , o teste de razão de verossimilhança é indicado pela equação (14), onde se compara a razão de máxima verossimilhança com uma constante não negativa (c), dada por (TEUNISSEN, 2006):

$$\begin{cases} \text{rejeitar } H_0 & \text{se } \frac{\max p(y|x)}{\max p(y|x, \nabla)} \leq c \\ \text{não rejeitar } H_0 & \text{se } \frac{\max p(y|x)}{\max p(y|x, \nabla)} > c \end{cases} \quad (14)$$

Klein (2012) destaca que o princípio do teste da razão de verossimilhança é intuitivo: a razão de verossimilhança, definida em (14), tende a ser pequena quando a hipótese nula não é verdadeira, pois o denominador $\max p(y|x, \nabla)$ tende a ser maior que o numerador $\max p(y|x)$, e desta forma, é provável que a rejeição de H_0 em favor de H_a seja uma decisão correta.

Segundo Klein (2012), a partir do caso geral de teste de hipóteses em modelos lineares ($1 \leq q \leq n - u$), serão derivados os testes para a detecção e identificação de *outliers*. Considerando para detecção o teste global do ajustamento ($q = n - u$) e para identificação de *outliers* o procedimento de teste *Data Snooping* ($q = 1$). Sendo, n o número de observações e u o número de parâmetros incógnitos do ajustamento.

3.2 Teste global do ajustamento ($q = n - u$)

Pode-se estimar qualquer valor para o fator de variância *a priori* (σ_0^2) antes de realizar o ajustamento para o cálculo da matriz dos pesos das observações, sendo que este valor não irá acarretar efeito no resultado do ajustamento. O fator de variância *a posteriori* ($\hat{\sigma}_0^2$) em função dos resíduos pode ser estimado após o ajustamento.

A comparação entre os valores de σ_0^2 e $\hat{\sigma}_0^2$ pode ser um indicador da qualidade do ajustamento. Primeiramente, se houver discrepância entre σ_0^2 e $\hat{\sigma}_0^2$ ($\sigma_0^2 \neq \hat{\sigma}_0^2$) aplica-se o teste global do ajustamento baseado na distribuição χ^2 para constatar se a discrepância é significativa a um certo nível de confiança (GEMAEL, 1994).

O teste global do ajustamento testa a hipótese nula (H_0), onde os fatores de variância *a posteriori* e *a priori* são estatisticamente iguais, e na hipótese alternativa (H_a), acontece o contrário:

$$H_0: \sigma_0^2 = \hat{\sigma}_0^2 \quad \text{contra} \quad H_a: \sigma_0^2 \neq \hat{\sigma}_0^2 \quad (15)$$

Para testar se a hipótese nula é rejeitada ou não, a um nível de significância α , compara-se o valor calculado por:

$$T = \left(\frac{\hat{\sigma}_0^2}{\sigma_0^2} \right) (n - u) = \frac{V^T P V}{\sigma_0^2} \quad (16)$$

com o valor teórico da distribuição $\chi_{(n-u, \alpha)}^2$.

Onde:

n : é o número de observações;

u : é o número de parâmetros;

V : é o vetor ($n \times 1$) dos resíduos das observações e

P : é a matriz peso das observações ($P = \sigma_0^2 \Sigma_{Lb}^{-1}$, onde Σ_{Lb}^{-1} é a MVC das observações).

A hipótese nula “não” é rejeitada a um nível de significância α , se:

$$T < \chi_{(n-u, \alpha)}^2 \quad (17)$$

A estatística $w_{lb(i)}$ tem como distribuição a raiz quadrada da distribuição F , com graus de liberdade igual a 1 e ∞ (infinito). A raiz quadrada da distribuição F tem distribuição normal padrão, isto é, $\sqrt{F_{a(1,\infty,\infty)}} = N_{\frac{a_0}{2}}(0,1)$. Então (TEUNISSEN, 2000):

$$w_{lb(i)} \sim N_{\frac{a_0}{2}}(0,1) \quad (21)$$

Portanto na hipótese nula a i -ésima observação não contém *outlier* ($\nabla_{Lb(i)}$) e na hipótese alternativa contém *outlier*, então, testa-se:

$$H_0: \nabla_{Lb(i)} = 0 \quad \text{contra} \quad H_a: \nabla_{Lb(i)} \neq 0 \quad (22)$$

Não se rejeita a hipótese básica a um determinado nível de significância α_0 se (TEUNISSEN, 2000):

$$|w_{lb(i)}| \leq N_{\frac{a_0}{2}}(0,1) \quad (23)$$

Caso contrário, suspeita-se que a i -ésima observação possui *outlier*. Por ser um teste que verifica apenas uma observação por vez e dependendo da magnitude do erro pode acontecer de mais de uma ou até todas as observações excederem o valor crítico, sendo assim a observação suspeita será aquela que possuir o maior valor em módulo para $W_{lb(i)}$. A j -ésima observação será suspeita de possuir *outlier* se (TEUNISSEN, 2000):

$$|w_{lb(j)}| > N_{\frac{a_0}{2}}(0,1) \quad e \quad |w_{lb(j)}| \geq |w_{lb(i)}|, \forall i \quad (24)$$

Por fim, retira-se a observação suspeita, refaz o ajustamento e aplica o teste novamente até que todas as observações suspeitas sejam eliminadas. Porém, ao aplicar o teste *Data Snooping* pode-se cometer erros do tipo I e II.

4 MÉTODO PROPOSTO

Como ponto de partida, caracterizamos as hipóteses nula e alternativa, denotadas por H_0 e H_A , respectivamente. Normalmente, na detecção de deslocamentos, a hipótese nula H_0 constitui no modelo livre de deslocamentos, ou seja, os pontos são considerados totalmente estáveis, sem movimentos, que, aqui, é assumido como:

$$H_0: E\{y\} = Ax; D\{y\} = \Sigma_y \quad (25)$$

com $E\{.\}$ sendo o operador de esperança matemática, $D\{.\}$ o operador de dispersão, $y \in \mathbb{R}^n$ o vetor das observações contendo as medidas coletadas em diferentes épocas, $x \in \mathbb{R}^u$ os parâmetros desconhecidos estimáveis, $A \in \mathbb{R}^{n \times u}$ a matriz *design* de posto igual a “ u ”, e $\Sigma_y \in \mathbb{R}^{n \times n}$ a matriz covariância positiva-definida de y . Nesse caso, a redundância do modelo de equações de observação sob H_0 é $r = n - u$.

Como hipótese alternativa, consideramos aquela que descreve um comportamento dinâmico sistemático da estrutura em consideração. Nesta contribuição, nos limitamos - por simplicidade das análises - a movimentos que podem ser caracterizados por um escalar comum em todos os pontos. O modelo observacional para esse caso é dado por:

$$H_A: E\{y\} = Ax + c_y \nabla; D\{y\} = \Sigma_y \quad (26)$$

em que $c_y \in \mathbb{R}^n$ descreve o modelo de deslocamento, e $\nabla \in \mathbb{R}$ é a magnitude do deslocamento. Nota-se que quando $\nabla = 0$, recaímos no modelo dado em (25). Aqui, restringimos para o caso em que $[A \ c_y] \in \mathbb{R}^{n \times u+1}$ tem posto completo igual à $u + 1$, e o escalar ∇ é um parâmetro desconhecido.

Com base na Equação (14), após algumas operações matemáticas, chega-se ao seguinte teste:

$$se \ Tq = \hat{v}_o^T P \hat{v}_o - \hat{v}_a^T P \hat{v}_a > k_{(\alpha, q=1)}, \text{ então rejeita-se a hipótese nula } H_0 \quad (27)$$

em que $v_o \in \mathbb{R}^n$ é o vetor dos erros estimados sob a hipótese nula H_0 em (25) e $v_a \in \mathbb{R}^n$ é o vetor de erros estimados sob a hipótese alternativa H_A em (26), e k é o valor crítico para a estatística calculada Tq . O valor crítico é obtido por meio da tabela da distribuição Qui-quadrado (χ^2) para um nível de significância estipulado e $q = 1$ grau de liberdade.

Como um exemplo simples, pode-se imaginar que a altitude de um único ponto é repetidamente observada em várias épocas e agrupadas no vetor y . A hipótese nula H_0 consiste de um modelo que considera o ponto totalmente estável, sem movimento. Para o caso simples de duas épocas, pode-se escrever as seguintes equações de observação:

$$\begin{aligned}
 h_{(1)}^{(1)} + v_{(1)}^{(1)} &= \hat{h} \\
 h_{(2)}^{(1)} + v_{(2)}^{(1)} &= \hat{h} \\
 h_{(3)}^{(1)} + v_{(3)}^{(1)} &= \hat{h} \\
 &\vdots \\
 h_{(n)}^{(1)} + v_{(n)}^{(1)} &= \hat{h} \\
 \\
 h_{(1)}^{(2)} + v_{(1)}^{(2)} &= \hat{h} \\
 h_{(2)}^{(2)} + v_{(2)}^{(2)} &= \hat{h} \\
 h_{(3)}^{(2)} + v_{(3)}^{(2)} &= \hat{h} \\
 &\vdots \\
 h_{(n)}^{(2)} + v_{(n)}^{(2)} &= \hat{h}
 \end{aligned} \tag{28}$$

em que o subscrito se refere a uma dada observação e o sobrescrito à época. Nesse caso, todas as observações se relacionam a um único parâmetro desconhecido, ou seja, a altitude do ponto em consideração $\hat{h}$, através da matriz de *design* A que é igual a um vetor em que seus elementos são todos iguais a “1”.

A hipótese alternativa H_a implica em uma mudança repentina na altitude do ponto, que é suposto, por exemplo, ocorrer na segunda época, ou seja:

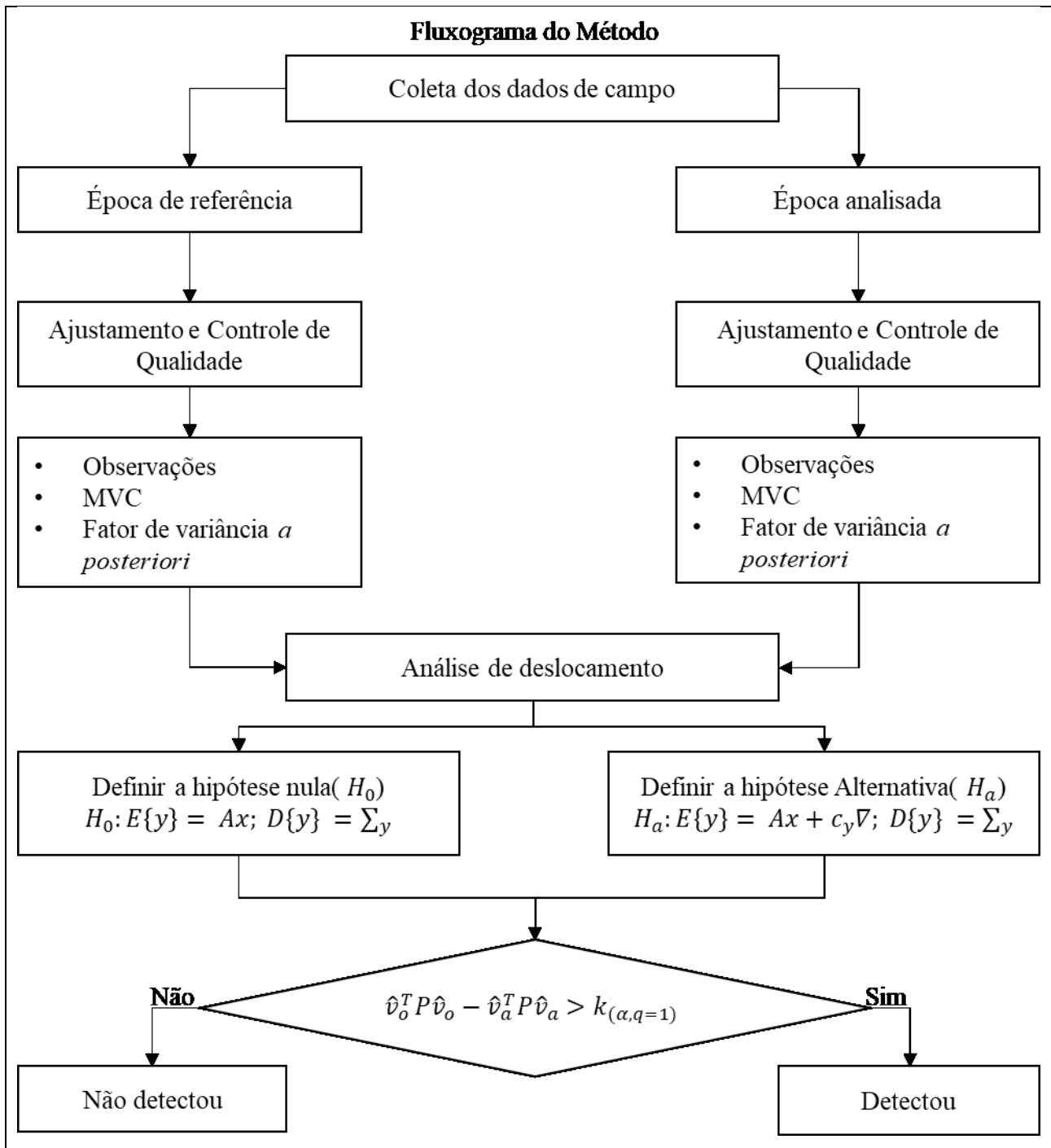
$$\begin{aligned}
 h_{(1)}^{(1)} + v_{(1)}^{(1)} &= \hat{h} \\
 h_{(2)}^{(1)} + v_{(2)}^{(1)} &= \hat{h} \\
 h_{(3)}^{(1)} + v_{(3)}^{(1)} &= \hat{h} \\
 &\vdots \\
 h_{(n)}^{(1)} + v_{(n)}^{(1)} &= \hat{h} \\
 \\
 h_{(1)}^{(2)} + v_{(1)}^{(2)} &= \hat{h} + \nabla \\
 \end{aligned} \tag{29}$$

$$\begin{aligned}
h_{(2)}^{(2)} + v_{(2)}^{(2)} &= \hat{h} + \nabla \\
h_{(3)}^{(2)} + v_{(3)}^{(2)} &= \hat{h} + \nabla \\
&\vdots \\
h_{(n)}^{(2)} + v_{(n)}^{(2)} &= \hat{h} + \nabla
\end{aligned}$$

O escalar ∇ representa o deslocamento desconhecido, e o vetor c_y constitui de valores “0” nas observações da primeira época e “1” nas épocas subsequentes. Assim, tem-se o ajustamento para o caso da hipótese nula e também para o caso da hipótese alternativa. Logo, aplica-se o teste dado por (27) para verificar se o parâmetro ∇ sob H_a é estatisticamente significativo. Caso a hipótese alternativa seja a mais provável, o vetor de parâmetros é estimado considerando não somente a altitude do ponto, mas também a magnitude e o sinal do deslocamento ∇ , bem como as incertezas associadas.

Finalmente, mencionamos que, na prática, as hipóteses alternativas podem também envolver *outliers* incidentais e falhas nas medições do sistema de monitoramento. Essas hipóteses são tratadas em cada ajustamento individual por época, por meio do controle de qualidade descrito na seção (3). Além disso, importante ressaltar que no método proposto considera-se também a análise de variância entre épocas, com o objetivo de verificar se o grau de variabilidade entre as campanhas (épocas) é estatisticamente significativo. Nesse último caso, a ideia é a mesma do teste global apresentado na seção 3.2, com a diferença que agora a relação é entre os fatores de variância entre épocas, e, portanto, com valor crítico calculado em função da distribuição $F_{(\alpha, n-u, n-u)}$. A Figura 2 mostra um fluxograma simplificado do método proposto.

Figura 2 – Fluxograma do método proposto.



Fonte: O autor (2021).

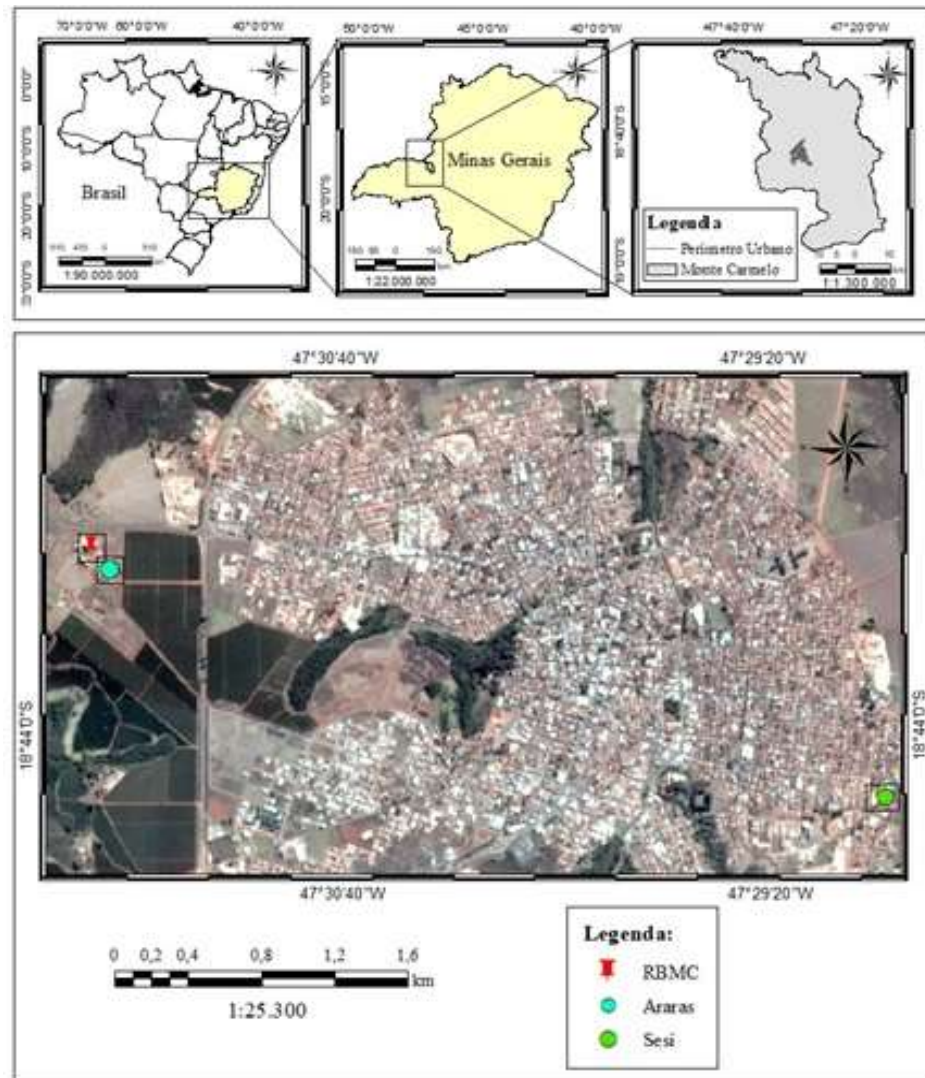
5 MATERIAL E MÉTODOS

5.1 Área de estudo

O município de Monte Carmelo está localizado no interior de Minas Gerais, na mesorregião do triângulo mineiro e alto Paranaíba. Com uma altitude de 890 metros e coordenadas geográficas $-18^{\circ} 44' 05''$ (latitude) e $-47^{\circ} 29' 47''$ (longitude).

A Figura 3 apresenta o mapa de localização do município de Monte Carmelo e a localização dos pontos utilizados para os estudos de caso apresentados nesse trabalho.

Figura 3 – Mapa de localização da área de estudo.



Sistema de Coordenadas e Sistema de Referência: UTM Zona 23 S – SIRGAS 2000
 Elaborado por: SILVA NETO J.B. (2021)
 Fonte: IBGE – Instituto Brasileiro de Geografia e Estatística (2015); Planet (2021).

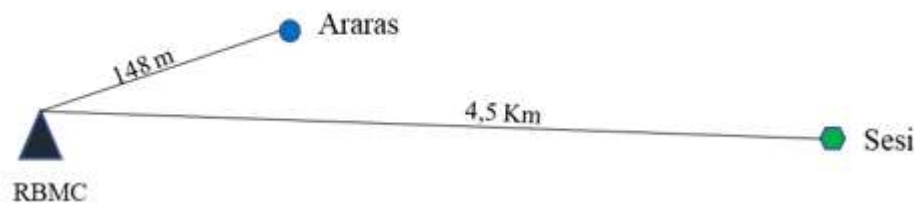
Fonte: O autor (2021)

5.2 Desenho Experimental

Para a análise dos deslocamentos verticais os dados foram coletados com receptor GNSS de dupla frequência, utilizando o método de posicionamento relativo estático, neste tipo de posicionamento, dois ou mais receptores rastreiam, simultaneamente, os satélites visíveis por um período de tempo que pode variar de 20 minutos (no mínimo) até algumas horas.

Para avaliar o impacto de se utilizar receptor de dupla frequência (L1 & L2) ou simples frequência (L1) na detecção de deslocamentos e função do comprimento da linha base, foram realizados dois experimentos, denominadas de Araras e Sesi, sendo as linhas base com comprimento aproximado de 148 metros e 4,5 quilômetros, respectivamente como mostra a Figura 4.

Figura 4 – Croqui dos experimentos



Fonte: O autor (2021).

Para os experimentos Araras e Sesi utilizou-se os mesmos procedimentos para aquisição dos dados, sendo o método de posicionamento relativo estático, com taxa de coleta de dados de 15 segundos (a mesma taxa de coleta de dados da estação base a MGMT – Monte Carmelo SAT 99680, pertencente à Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS (RBMC)) e duração de 3 horas cada seção de rastreamento. O levantamento foi realizado em dias diferentes escolhidos de forma aleatória no período da manhã e tarde (Tabela 1). Os deslocamentos foram simulados em campo por meio da variação da altura da antena. A altura da antena foi configurada para os seguintes valores 2,00 m, 2,005 m e 2,01m.

Tabela 1 – Datas da coleta dos dados para a análise de deslocamento vertical.

Datas da coleta dos dados do experimento Araras				Datas da coleta dos dados do experimento Sesi			
Data	Hora (UTC)		Altura da antena (m)	Data	Hora (UTC)		Altura da antena (m)
	Início	Fim			Início	Fim	
02/04/2019	12:07:00	15:07:00	2,00	29/08/2019	11:55:00	14:55:00	2,00
02/04/2019	16:14:00	19:14:00	2,00	29/08/2019	16:00:00	19:00:00	2,00
05/04/2019	11:30:00	14:30:00	2,01	30/08/2019	11:25:00	14:25:00	2,01
05/04/2019	16:25:00	19:25:00	2,01	30/08/2019	15:55:00	18:55:00	2,01
23/08/2019	13:02:00	16:02:00	2,005	20/09/2019	11:40:00	14:40:00	2,005
23/08/2019	17:00:00	20:00:00	2,005	20/09/2019	15:55:00	18:55:00	2,005

Fonte: O autor (2021).

Após a realização das campanhas de campo, os dados obtidos em cada seção foram editados em 6 arquivos contendo dados referentes a 30 minutos de rastreo. No processamento dos dados, a altura foi fixada em 2,00 metros para todos os cenários em campo. O processamento dos dados foi realizado por meio do *software TopconTools v.8.5*, com máscara de elevação de 10 graus, para os dados de simples frequência (L1) realizou-se o processamento das duplas diferenças, para os dados de dupla frequência (L1 & L2) efetuou-se o processamento das duplas diferenças de L1&L2 de forma independente (isto é, não foi realizado nenhum tipo de combinação linear entre as observáveis). A estação de referência (base) utilizada foi a estação MGMT – Monte Carmelo SAT 99680, pertencente à Rede Brasileira de Monitoramento Contínuo dos Sistemas GNSS (RBMC).

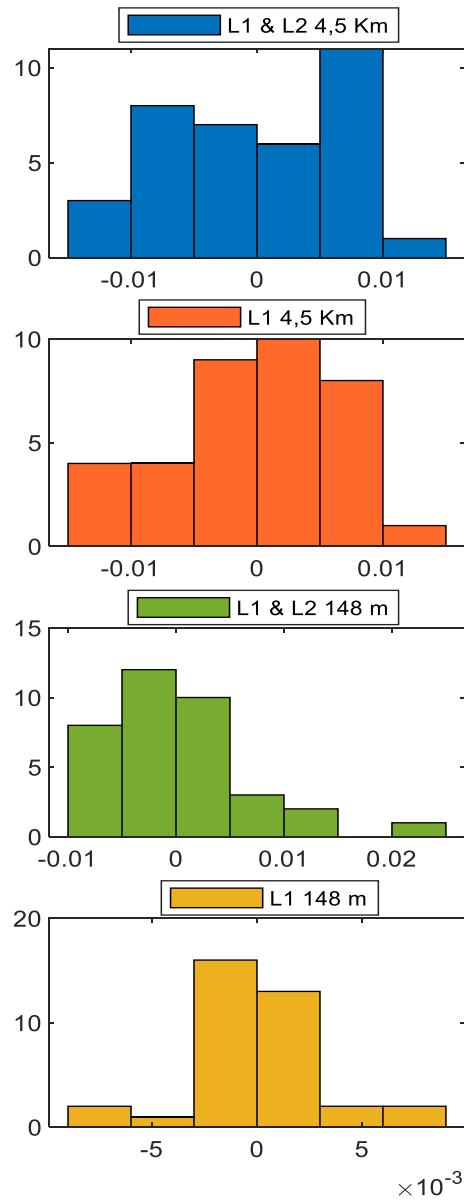
6 EXPERIMENTOS E RESULTADOS

A diferença entre os desníveis estimados e seu valor mais provável (média das estimativas), considerando o processamento dos dados com a informação da altura da antena medida em campo (ou seja, sem simular deslocamentos verticais), foi calculada para cada cenário experimental. Essa diferença pode ser entendida como sendo o resíduo.

As distribuições dos resíduos para os casos “L1 & L2 4,5 Km” e “L1 4,5 Km” se caracterizaram por serem menos propensas a *outliers* (curtoses menores que 3, como pode ser verificado na Tabela 2). Nesse caso, as distribuições são chamadas de platicúrticas. Por outro lado, as distribuições dos resíduos para os casos “L1 & L2 148 m” e “L1 148 m” seguiram uma distribuição leptocúrtica (curtoses maiores que 3), ou seja, mais propensas a *outliers*. Nota-se também que o cenário “L1 & L2 148 m” não segue distribuição normal, conforme o teste de

normalidade Shapiro-Wilk para $\alpha = 0,05$ (5%). As distribuições dos resíduos não foram perfeitamente simétricas em todos os casos, com especial atenção para o caso “L1 & L2 148 m” em que a distribuição teve uma assinatura assimétrica à direita (Figura 5).

Figura 5 – Distribuição dos resíduos



Fonte: O autor (2021).

Tabela 2 – Descrição estatística dos resíduos para cada cenário experimental.

Linha base	Kurt.	Max. (m)	Med. (m)	Min. (m)	Mod. (m)	Sim. (m)	SD (m)	Normal para $\alpha = 0,05$
L1 & L2 4,5 Km	1,94	0,015	0,006	< 0,001	< 0,001	-0,25	0,007	Sim (p-value = 0,104 > 0,05)
L1 4,5 Km	2,13	0,014	0,006	< 0,001	< 0,001	-0,4	0,007	Sim (p-value = 0,055 > 0,05)
L1 & L2 148 m	4,74	0,02	0,005	< 0,001	< 0,001	1,27	0,006	Não (p-value = 0,004 < 0,05)
L1 148 m	3,83	0,008	0,002	< 0,001	< 0,001	0,54	0,003	Sim (p-value = 0,178 > 0,05)

Fonte: O autor (2021).

Kurt. = curtose dos resíduos; Max. = máximo valor dos resíduos absolutos; Med. = média dos resíduos absolutos; Min. = mínimo valor os resíduos absolutos; Mod. = moda dos resíduos absolutos; SD = desvio-padrão dos resíduos; p-value = probabilidade de se obter uma estatística de teste igual ou mais extrema que aquela observada em uma amostra, sob a hipótese nula.

Para verificar a ocorrência de possíveis erros entre as campanhas, o método proposto foi aplicado para os casos em que a altura da antena medida em campo foi considerada no processamento, ou seja, desprezando as simulações dos deslocamentos em campo.

O deslocamento estimado foi significativo ($\alpha = 0,05$) somente para o cenário “L1 & L2 148 m” na comparação entre o cenário com altura da antena de 2 m e 2,005 m, com magnitude de $\nabla = 4,19 \text{ mm} \pm 1,9 \text{ mm}$ para 1σ (Tabela 3). Nesse caso, $T_q = 17,9 > k = 3,8$. Mesmo que fosse adotado um nível de significância menor que $\alpha = 0,05$ (por exemplo, $\alpha = 0,0001$, que corresponde à $k = 15,1$), essa tendência ainda seria estatisticamente diferente de zero, ou seja, significativa. Importante destacar que foi identificado e removido um *outlier* durante a etapa de controle de qualidade para cada um dos casos “L1 & L2 148 m”, considerando um nível de significância de $\alpha = 0,001$ (0,1%).

Tabela 3 – Detecção de tendências para diferentes alturas de antena em diferentes épocas

Linha base	Deslocamento Estimado ∇ (mm)	
	2,0 m vs 2,005 m	2,0 m vs 2,01 m
L1 & L2 4,5 Km	-0,61	2,05
L1 4,5 Km	3,36	3,67
L1 & L2 148 m	4,19 $\pm$ 1,9 (detectado)	0,15
L1 148 m	-0,71	0,35

Fonte: O autor (2021).

Para a análise dos deslocamentos considerou-se as observações da época de referência sem deslocamento, já na época analisada admite-se que as observações contem deslocamentos, ou seja, existe um deslocamento sistemático em todas os desníveis. Adotando os valores de $\alpha =$

0.01 para o *Data Snooping* e $\alpha = 0.05$ para a detecção de deslocamentos, quando simulado o deslocamento em campo de 5 milímetros (Tabela 4) e 1 centímetro (Tabela 5) verificou-se que para o caso “L1 4,5 Km” não foi detectado deslocamento $T_q < 3.84$. Para o experimento Araras que possui a linha base de 148 m, os resultados para o deslocamento estimado são melhores quando utilizado dados somente de L1. Já para o experimento Sesi que possui a linha base de 4,5 Km os resultados para o deslocamento estimado são melhores quando utilizado dados L1 & L2.

Tabela 4 – Deslocamento simulado de 5 mm em campo

Linha base	Deslocamento estimado (mm)	SD Época de referência (mm)	SD Época analisada (mm)	T_q	$T_q > k$ para $\alpha = 0,05$
L1 & L2 4,5 Km	$-5,6 \pm 3,0$	1,4	1,7	10,1	Detectado ($T_q > 3.84$)
L1 4,5 Km	$-1,6 \pm 2,6$	0,3	0,4	1,8	Não detectado ($T_q < 3.84$)
L1 & L2 148 m	$11,2 \pm 2,9$	1,8	1,7	127,4	Detectado ($T_q > 3.84$)
L1 148 m	$6,3 \pm 1,7$	1,3	1,8	72,1	Detectado ($T_q > 3.84$)

Fonte: O autor (2021).

Tabela 5 – Deslocamento simulado de 1 cm em campo

Linha base	Deslocamento estimado (mm)	SD Época de referência (mm)	SD Época analisada (mm)	T_q	$T_q > k$ para $\alpha = 0,05$
L1 & L2 4,5 Km	$-8,0 \pm 2,8$	1,4	1,4	23,8	Detectado ($T_q > 3.84$)
L1 4,5 Km	$-6,3 \pm 2,7$	0,3	0,4	1,5	Não detectado ($T_q < 3.84$)
L1 & L2 148 m	$10,1 \pm 2,7$	1,8	1,7	119,7	Detectado ($T_q > 3.84$)
L1 148 m	$10,3 \pm 2,3$	1,3	1,1	191,6	Detectado ($T_q > 3.84$)

Fonte: O autor (2021).

7 CONSIDERAÇÕES FINAIS

Neste trabalho foi apresentado uma forma de descrever o deslocamento baseado no teste de razão de máxima verossimilhança, manipulando as equações de observação do ajustamento, como é apresentada em muitos casos de teste de detecção e identificação de erros. O método foi aplicado em um exemplo simples de análise de deslocamento vertical simulado em campo

utilizando a tecnologia GNSS e o método de posicionamento relativo estático. Os resultados sugerem que deve-se evitar para a análise de deslocamento o uso de receptores de dupla frequência L1 & L2 com processamento de forma independente em linhas de base curta. Além disso, o aumento da linha de base no caso de utilizar dados somente L1, torna-se mais difícil a detecção de deslocamentos.

Para trabalhos futuros recomenda-se realizar a aplicação do método proposto em um número maior de experimentos com diferentes variações no comprimento de linhas de base. Recomenda-se também testar outros valores de α para o *Data Snooping* e na detecção de deslocamentos, assim como a coleta dos dados de temperatura e pressão no período da realização dos levantamentos de campo.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAARDA, W. **A testing procedure for use in geodetic networks**. Netherlands Geodetic Commission –Publication on Geodesy – News Series, Delft, v.2, n.5, 1968.
- BREIMAN, L. **Statistics: with a view toward applications**. Boston: Houghton Mifflin Company, 1973. 399p.
- DALMOLIN, Q. **Ajustamento por Mínimos Quadrados**. Curitiba: Curso de Pós-Graduação em Ciências Geodésicas/UFPR, 2002. 175p.
- FAZAN, J. A. **Aplicação do Teste de Congruência Global e Análise Geométrica para Detecção de Deslocamentos em Redes Geodésicas: Estudo de Caso na UHE de Itaipu**. Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo, São Paulo, 2010. 249p.
- GEMAEL, C. **Introdução ao ajustamento de observações: aplicações geodésicas**. Curitiba: Ed. UFPR, 1994. 319p.
- GEMAEL, C.; MACHADO, A. M. L.; WANDRESEN, R. **Introdução ao Ajustamento de Observações: Aplicações Geodésicas**. 2. ed. Curitiba: Ed. UFPR, 2015. 430 p. (Pesquisa; n.281).
- IBGE. **Cidades**. Disponível em: <<https://cidades.ibge.gov.br/brasil/mg/monte-carmelo/panorama>>. Acessado em: out. 2019.
- 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. **Controle de Qualidade no Ajustamento de Observações Geodésicas**. Dissertação (Mestrado) – Programa de Pós-Graduação em Sensoriamento Remoto, UFRGS, 2012. 322p.
- LARSON, H. J. **Introduction to probability theory and statistical inference**. 2. ed. New York: John Wiley & Sons, 1974. 430p.
- LEICK, A. **GPS satellite surveying**. 3. ed. New York: John Wiley & Sons, 2004.
- MAGALHÃES, M. N.; LIMA, A. C. P. **Noções de probabilidade e estatística**. 6. ed. São Paulo: Ed. USP, 2005. 392p.
- OLIVEIRA, R.; DALMOLIN, Q. **A Influência da Redundância da Observação sobre a Precisão dos Parâmetros**. Boletim de Ciências Geodésicas, v. 14, n. 3, p. 295-315, 2008.
- PAPOULIS, A. **Probability, random variables and stochastic processes**. New York: McGraw-Hill Book Company, 1991.

SAVŠEK, S.; AMBROŽIČ, T.; STOPAR, B.; TURK, G. **Determination of Point Displacements in the Geodetic Network.** Journal of Surveying Engineering, 2006.

SCAIONI, M.; MARSELLA, M.; CROSETTO, M.; TORNATORE, V.; WANG, J. **Geodetic and Remote-Sensing Sensors for Dam Deformation Monitoring.** *Sensors (Basel, Switzerland)*, 18(11), 3682, 2018. DOI: 10.3390/S18113682.

TEIXEIRA, N. N. **Detecção e análise de deslocamento vertical utilizando o Sistema NAVSTAR-GPS.** Dissertação (Mestrado) - Curso de Pós-Graduação em Ciências Geodésicas, Universidade Federal do Paraná, Curitiba, 2001. 112 p.

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

TEUNISSEN, P. J. G. **Adjustment Theory: an introduction.** Delft: Ed. Delft University Press, 2003. 193p.

TEUNISSEN, P.J.G. **Testing Theory: an introduction.** Delft: Ed. Delft University Press, 2000. 193p.