

MARCELO TOMIO MATSUOKA

**Desenvolvimento de Novos Métodos de Avaliação de Risco para
Sistemas Geodésicos de Monitoramento de Estruturas**

**Relatório de Pós-doutorado realizado na
Universidade Estadual Paulista (UNESP),
Faculdade de Ciências e Tecnologia,
Programa de Pós-Graduação em Ciências
Cartográficas, Presidente Prudente.**

**Supervisor: Prof. Dr.
Paulo de Oliveira Camargo**

CNPq – Proc. N. 102498/2022-9

Presidente Prudente

2023

IDENTIFICAÇÃO

Título do Projeto: Desenvolvimento de Novos Métodos de Avaliação de Risco para Sistemas Geodésicos de Monitoramento de Estruturas

Instituições Executoras:

Principal: Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista (FCT/UNESP), Programa de Pós-Graduação em Ciências Cartográficas (PPGCC), Presidente Prudente, São Paulo, Brasil.

Em parceria: Universidade Federal de Uberlândia, Instituto de Geografia, Grupo de Pesquisa: Controle de Qualidade em Geodésia, Monte Carmelo, Minas Gerais.

Bolsista Pós-Doutorado Sênior: Professor Doutor Marcelo Tomio Matsuoka

e-mail: tomiomatsuoka@gmail.com

Pesquisador Supervisor: Professor Doutor Paulo de Oliveira Camargo

e-mail: paulo.camargo@unesp.br

PERÍODO: 01/Setembro/2022 até 31/Agosto/2023

SUMÁRIO

RESUMO	4
1. INTRODUÇÃO	6
2. MÉTODO PROPOSTO (SLRTUPI) – TESTE SEQUENCIAL DE RAZÃO DE VEROSSIMILHANÇA PARA O MONITORAMENTO DE DESLOCAMENTOS DE PONTOS	11
2.1 HIPÓTESE NULA CONTRA UMA HIPÓTESE ALTERNATIVA: TESTE BINÁRIO	11
2.2 HIPÓTESE NULA CONTRA MÚLTIPLAS HIPÓTESES ALTERNATIVAS	20
2.3 MÉTODO MONTE CARLO PARA CONTROLAR A TAXA DE FALSA DETECÇÃO	28
2.4 DETERMINAÇÃO DO NÚMERO MÁXIMO DE PONTOS POSSÍVEIS DE INSPEÇÃO -	31
3. EXPERIMENTOS E RESULTADOS OBTIDOS	33
4. CONSIDERAÇÕES FINAIS	43
REFERÊNCIAS	46
AGRADECIMENTOS	52

RESUMO

O presente relatório de atividades de estágio Pós-Doutorado Sênior – PDS apresenta a descrição da pesquisa que foi realizada para alcançar os objetivos indicados no plano de trabalho da proposta submetida. Tendo como principal objetivo do projeto proposto desenvolver novas ideias/métodos para avaliação de risco para sistemas geodésicos de monitoramento de estruturas, a execução da pesquisa cumpriu com êxito esta missão, conforme será visto neste relatório. Um novo método de monitoramento de pontos em estruturas geodésicas foi desenvolvido e dado o nome de SLRTUPI e descrita resumidamente a seguir. Um dos principais desafios na análise de deformação geodésica é inferir se a geometria de algumas estruturas de engenharia ou zonas de perigos naturais mudou ou não em relação ao seu estado inicial. O pilar que sustenta tal análise baseia-se firmemente nos fundamentos dos testes de hipóteses estatísticas. O modelo de hipótese nula indica que nenhum deslocamento ocorreu. É testado contra uma classe de modelos de hipóteses alternativas, que estipulam diferentes padrões de deslocamento. Nesta contribuição, apresentamos uma detecção inovadora de deslocamento geodésico que integra análise combinatória e testes de razão de verossimilhança em um procedimento sequencial para o caso em que as diferenças entre observações de duas épocas estão em análise. Esta estrutura é aplicada e investigada em dois cenários de teste: uma rede de trilateração sintética e uma rede de trilateração que foi projetada, materializada e com medidas de campo. Do ponto de vista estatístico, nossa abordagem é rigorosa, pois o modelo alternativo pode identificar simultaneamente mais de um ponto instável. Além disso, a relação entre os parâmetros desconhecidos e as observações é sempre linear, mesmo que o problema se manifeste como não linear. Consequentemente, evitamos um potencial de perda de poder do teste estatístico devido à linearização do modelo. Além disso, o método proposto controla a taxa de falsos positivos de forma eficiente e adota também uma metodologia que desenvolvemos para determinar o número de pontos máximo simultâneos que podem ser inspecionados baseado em critérios matemáticos. Um conjunto de scripts em Matlab foram implementados para a execução do método SLRTUPI. Tanto esses scripts quanto os dados utilizados nesta pesquisa são disponibilizados aos interessados em: <https://data.mendeley.com/datasets/msg783rh2y/draft?a=607de244-7c1d-417f-a9c8-81281d8a6056>.

ABSTRACT

This Senior Postdoctoral activity report presents a description of the research carried out to achieve the objectives outlined in the submitted work plan. With the main goal of the proposed project being the development of new ideas/methods for risk assessment in geodetic monitoring systems of structures, the research was successfully conducted, as will be detailed in this report. A new method for monitoring points in geodetic structures was developed, named SLRTUPI, and is briefly described below.

One of the main challenges in geodetic deformation analysis is to infer whether the geometry of certain engineering structures or natural hazard zones has changed compared to its initial state. The foundation of such analysis relies firmly on the principles of statistical hypothesis testing. The null hypothesis model states that no displacement has occurred and is tested against a class of alternative hypothesis models that stipulate different displacement patterns.

In this contribution, we present an innovative geodetic displacement detection method that integrates combinatorial analysis and likelihood ratio tests in a sequential procedure for the case where differences between two-epoch observations are under analysis.

This framework is applied and investigated in two test scenarios: a synthetic trilateration network and a trilateration network that was designed, materialized, and field-measured.

From a statistical standpoint, our approach is rigorous because the alternative model can simultaneously identify more than one unstable point. Moreover, the relationship between unknown parameters and observations remains linear at all times, even if the problem appears nonlinear. Consequently, we avoid a potential loss of statistical test power due to model linearization.

Furthermore, the proposed method efficiently controls the false positive rate and also adopts a methodology we developed to determine the maximum number of simultaneous points that can be inspected based on mathematical criteria.

A set of Matlab scripts was implemented to execute the SLRTUPI method. Both these scripts and the data used in this research are available at:

<https://data.mendeley.com/datasets/msg783rh2y/draft?a=607de244-7c1d-417f-a9c8-81281d8a6056>.

1. INTRODUÇÃO

Um dos principais desafios atuais da análise de dados geodésicos é detectar alterações geométricas de objetos ou áreas sujeitas a deslocamentos e/ou deformações, tais como: estruturas artificiais como barragens, diques, pontes, turbinas eólicas ou torres altas, bem como estruturas naturais como vulcões, áreas de mineração, etc. A análise das medições de monitoramento pode ser categorizada em quatro modelos de deformação: modelo de congruência, modelo cinemático, modelo estatístico e modelo dinâmico. O modelo de congruência descreve as deformações por meio de vetores de deslocamento sem especificar o tempo e nenhum fator relacionado às forças atuantes e às cargas internas e externas. Os modelos cinemáticos descrevem as mudanças geométricas em termos de variações temporais (velocidades e acelerações), considerando que o estado do objeto está permanentemente em movimento, mas também não há preocupações com as causas das deformações. Se houver interesse em investigar a relação funcional entre forças causais e reações geométricas do objeto, então o modelo estático será mais adequado. Neste último, as deformações são descritas a partir de propriedades físicas do objeto (por exemplo, coeficientes de expansão, temperatura e comprimentos), de modo que os aspectos temporais não ficam explícitos no modelo. Por fim, o modelo dinâmico é uma combinação de modelos estáticos e cinemáticos, ou seja, as deformações estão ligadas aos seus fatores de influência (forças causais, cargas internas e externas) e às propriedades físicas do objeto [1]. Devido ao seu uso comum, restringimo-nos ao modelo de congruência que apenas nos diz se o objeto se moveu ou não.

Na análise de congruência, a estrutura sob investigação é frequentemente monitorada por uma rede geodésica que é medida em pelo menos duas épocas no tempo, e são então analisadas estatisticamente. A rede geodésica funciona como um sistema de monitoramento de

deslocamentos. O teste estatístico é uma das abordagens mais utilizadas para a especificação de modelos de congruência de deformações [2-7]. A análise a partir dos chamados métodos robustos é outra também bastante utilizada e teve avanços importantes nos últimos anos [5,8,9,10], mas não faz parte do escopo desta pesquisa.

Normalmente, os dados de entrada são um vetor composto pelas diferenças entre as coordenadas dos pontos estimados por mínimos quadrados em uma época inicial (digamos, Época I) e uma época atual (digamos, Época 2). A hipótese nula, denotada por \mathcal{H}_0 , é formulada sob a condição de que todos os pontos sejam estáveis (pontos que possuem uma estrutura geométrica congruente/rígida em ambas as épocas consideradas). Por outro lado, as hipóteses alternativas são estipuladas a partir do pressuposto de que existe pelo menos um ponto instável. Como não se sabe qual ponto ou grupo de pontos é instável, um teste de hipótese iterativo é frequentemente aplicado para identificar um ponto instável após o outro [11]. Esse procedimento de teste é semelhante ao famoso método de identificação de outlier em um conjunto de medidas chamado de Data Snooping Iterativo [12]. No entanto, os testes de hipóteses iterativos não são rigorosos porque as hipóteses alternativas são restritas a apenas um único ponto instável [11]. O ponto sinalizado como mais suspeito de ser instável em uma determinada etapa não é inspecionado mais na próxima etapa da iteração. Além disso, o resultado de uma identificação incorreta em uma determinada iteração condiciona o resultado da próxima iteração [13]. A limitação dos testes de hipóteses iterativos para o caso em que múltiplos deslocamentos estão em análise foi relatada por vários autores [13-16].

Para superar o problema dos procedimentos iterativos, surge o procedimento combinatório não iterativo para o caso em que todas as combinações possíveis de pontos deslocados são consideradas [11,17,18]. Tal procedimento consiste em comparar todos os possíveis candidatos a pontos estáveis em uma mesma etapa, e conseqüentemente não é necessário especificar consecutivamente ponto a ponto o modelo de congruência. Este método

foi aplicado em alguns exemplos numéricos e discutido em detalhes por Velsink [17, 18]. Velsink propôs a razão entre a estatística de teste e seu valor crítico como regra de decisão. O ponto ou grupo de pontos com maior razão é sinalizado como instável, caso ultrapasse o valor crítico. Outro método combinatório interessante é discutido por Lehmann e Lösler [11]. Eles usam vários critérios de informação e, em seguida, selecionam entre todos os modelos candidatos possíveis aquele que fornece o menor valor de critério de informação como o melhor modelo. A ideia de utilizar o método de critérios de informação está fora do escopo do trabalho, mas será considerada em trabalhos futuros.

Infelizmente, o método apenas combinatório é frequentemente feito a partir do conjunto de modelos de dimensões diferentes [13]. A comparação entre modelos de diferentes dimensões é complicada neste caso. Por exemplo, quanto mais pontos são modelados como instáveis (maior dimensão de congruência do modelo – mais complexo é o modelo), maior é a ocorrência de *overfitting* (um modelo é sempre melhor ajustado a observações com maior número de parâmetros). A complexidade do modelo pode ser contornada por meio da aplicação de penalidades. Conforme destacado por Nowel [13], a qualidade do ajuste do modelo e um termo de penalidade constituem um critério de identificação. No entanto, alerta que existem muitos critérios e ainda não está claro qual deles adotar. Nowel usou as possibilidades de testes combinatórios e de razão de verossimilhança generalizada realizados em uma etapa iterativa para superar os pontos fracos relacionados à especificação do modelo iterativo consecutivo ponto a ponto e ao método apenas combinatório. Embora tenha ocorrido progressos neste novo campo da combinatória, ainda existem desafios que abrem novas perspectivas de investigação [19-26].

Nesta pesquisa de pós-doutoramento apresentamos um método alternativo e sofisticado que integra análise combinatória e testes de razão de verossimilhança em um procedimento sequencial, que chamamos de SLRTUPI (*Sequential Likelihood Ratio Tests for Unstable Points*

Identification – Testes de Razão de Verossimilhança Sequencial para Identificação de Pontos Instáveis). Parte do procedimento é uma extensão do método sequencial para detecção e identificação de múltiplas observações com outliers proposto por Klein et al. [27] e uma adaptação para o problema de detecção/identificação de pontos deslocados/instáveis.

Aqui, o método que foi desenvolvido utiliza diferenças das observações de duas épocas em vez de coordenadas estimadas, como proposto por alguns autores [9]. A ideia de usar diferenças de observação não ajustadas (originais) já existe há algum tempo, como pode ser visto em [28, 29]. Erdogan et al. [30] apresentaram uma metodologia para identificação de pontos instáveis baseada nesta abordagem de análise das diferenças entre medições feitas em duas épocas, que chamaram de abordagem univariada. Como resultado, não precisamos nos preocupar se se trata de uma rede geodésica linear (por exemplo, rede vetorial de nivelamento ou GNSS) ou não linear (por exemplo, trilateração), uma vez que o modelo univariado é sempre linear. Além disso, a abordagem univariada tem o benefício de reduzir o efeito de “espalhamento” e “mascaramento” dos deslocamentos. “Espalhar” significa que um ponto instável faz com que outro ponto estável pareça instável, e “mascarar” esse ponto instável impede que outro seja identificado. Do nosso ponto de vista, isso se deve ao fato de que a abordagem univariada reduz as correlações entre os erros estimados por mínimos quadrados (resíduos) e, conseqüentemente, as correlações entre as estatísticas de teste das hipóteses alternativas também são diminuídas. O efeito da geometria da rede também é reduzido e não requer soluções para o problema de *Datum* da rede por meio de transformações (por exemplo, transformação S [31-33]).

No entanto, Erdogan et al. [30] não controlam taxas de falsos positivos (taxa de erro tipo I – detectar deslocamentos quando na realidade não há nenhum). Isso porque geralmente consideram que o teste envolve apenas uma única hipótese alternativa, quando na realidade são testes múltiplos, ou seja, envolve múltiplas hipóteses alternativas. Conseqüentemente, é

necessária uma abordagem que permita controlar o Erro Tipo I. Aqui, por exemplo, usamos um método de Monte Carlo para que a taxa de erro Tipo I definida pelo usuário seja controlada de forma eficiente [34]. Além disso, os trabalhos citados acima não deixam claro como escolher o número máximo de pontos a serem investigados como instáveis (deslocados). Esta escolha é um tanto arbitrária. Para evitar uma escolha subjetiva, foi também proposto nesta pesquisa um método para definir o número máximo de pontos possíveis de serem inspecionados como deslocados utilizando o SLRTUPI.

A sequência do relatório está organizada da seguinte forma. A seção 2 descreve em detalhes o método SLRTUPI que foi desenvolvido nesta pesquisa para o monitoramento do deslocamento de pontos. A seção 3 apresenta um resumo dos experimentos e resultados da aplicação do método SLRTUPI baseados em simulações computacionais e conjuntos de dados reais para demonstrar a confiabilidade e o desempenho do SLRTUPI em diversos cenários de padrão de deslocamento dos pontos. O detalhamento completo dos experimentos e resultados podem ser vistos no manuscrito (artigo em fase de *preprint*) que foi elaborado nesta pesquisa e que será submetido em breve para uma revista de alto fator de impacto internacional (provavelmente ao *Journal of Geodesy*) e que pode ser acessado em: <https://drive.google.com/file/d/1mkPFV19mbdXbN84B701cCrlohNYpR827/view>. A seção 6 finaliza com as considerações finais e um resumo das contribuições do estudo dessa pesquisa de Pós-Doutorado.

2. MÉTODO PROPOSTO (SLRTUPI) – TESTE SEQUENCIAL DE RAZÃO DE VEROSSIMILHANÇA PARA O MONITORAMENTO DE DESLOCAMENTOS DE PONTOS

2.1 HIPÓTESE NULA CONTRA UMA HIPÓTESE ALTERNATIVA: TESTE BINÁRIO

Para iniciar, o modelo univariado para a identificação de pontos instáveis em uma rede geodésica é dada pelas seguintes equações:

$$\begin{aligned} \mathbf{y}^{\{1\}} &= \mathbf{y}_0^{\{1\}} + \mathbf{e}^{\{1\}} \text{ para Época 1} \\ \mathbf{y}^{\{2\}} &= \mathbf{y}_0^{\{2\}} + \mathbf{e}^{\{2\}} \text{ para Época 2} \end{aligned} \quad (1)$$

onde $\mathbf{y}^{\{i\}} \in \mathbb{R}^{n \times 1}$ são os vetores das medidas geodésicas, $\mathbf{y}_0^{\{i\}} \in \mathbb{R}^{n \times 1}$ são os vetores dos valores verdadeiros (desconhecidos) das grandezas medidas e $\mathbf{e}^{\{i\}} \in \mathbb{R}^{n \times 1}$ são os vetores dos erros das medidas (desconhecidos) (note: os índices ‘1’ e ‘2’ representam as quantidades relacionadas a primeira e a segunda época de medição, respectivamente).

Ao subtrair a equação da Época 2 da equação da Época 1, o modelo univariado pode ser formulado como um modelo linear de Gauss-Markov, como segue:

$$\begin{aligned} \mathbf{y}^{\{2\}} - \mathbf{y}^{\{1\}} &= \mathbf{y}_0^{\{2\}} - \mathbf{y}_0^{\{1\}} + \mathbf{e}^{\{2\}} - \mathbf{e}^{\{1\}}, \text{ fazendo} \\ \Delta_{\mathbf{y}} &= \mathbf{y}^{\{2\}} - \mathbf{y}^{\{1\}}, \\ \mathbf{e}_{\Delta_{\mathbf{y}}} &= \mathbf{e}^{\{2\}} - \mathbf{e}^{\{1\}}, \\ \mathbf{x} &= \mathbf{y}_0^{\{2\}} - \mathbf{y}_0^{\{1\}} \\ \therefore \Delta_{\mathbf{y}} &= \mathbf{A}\mathbf{x} + \mathbf{e}_{\Delta_{\mathbf{y}}} \end{aligned} \quad (2)$$

com $\Delta_{\mathbf{y}} \in \mathbb{R}^{n \times 1}$ sendo o vetor com as diferenças dos valores de medidas entre as duas épocas, $\mathbf{e}_{\Delta_{\mathbf{y}}} \in \mathbb{R}^{n \times 1}$ o vetor desconhecido dos valores de diferença entre os erros das duas

épocas, $\mathbf{A} \in \mathbb{R}^{n \times 1}$ a matriz Jacobiana (também chamada de matriz design) de posto completo $u = 1$, no qual neste caso é um vetor de unitários (i.e., $\mathbf{A} = [1 \ 1 \ 1 \ \dots \ 1]^T$), e $x \in \mathbb{R}^{1 \times 1}$ o parâmetro desconhecido estimável, sendo neste caso um escalar que representa a diferença entre os valores verdadeiros desconhecidos entre as duas épocas. É importante destacar que agora o vetor de observações corresponde às diferenças das medições geodésicas em duas épocas (ou seja, em momentos diferentes). Além disso, essas diferenças são calculadas a partir das observações idênticas de cada época, ou seja, em um caso hipotético de erros idênticos nas medidas e sem qualquer deslocamento, essas diferenças seriam iguais a zero.

Após ter definido o modelo univariado de acordo com a Eq. (2), agora é necessário recorrer à teoria do teste de hipóteses para inferir se um subconjunto de pontos medidos em duas épocas é estável/congruente ou não. Normalmente, na detecção de deslocamento, o modelo sob hipótese nula, denotado por \mathcal{H}_0 , é estipulado para o caso em que todos os pontos a serem analisados são tratados como pontos estáveis. Em outras palavras, a hipótese nula afirma que o modelo nulo na Eq. (2) explica as medições. Assumindo erros de observação normalmente distribuídos com expectativa zero, ou seja:

$$\mathbf{e}_{\Delta_y(\mathcal{H}_0)} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\Delta_y}) \quad (3)$$

a hipótese nula \mathcal{H}_0 do modelo padrão de Gauss–Markov na forma linear da Eq. (2) é então dada por:

$$\mathcal{H}_0: E(\Delta_y) = \mathbf{A}x, D(\Delta_y) = \boldsymbol{\Sigma}_{\Delta_y} = \boldsymbol{\Sigma}_y^{\{1\}} + \boldsymbol{\Sigma}_y^{\{2\}} \quad (4)$$

com $E(\cdot)$ sendo o operador de esperança matemática, $D(\cdot)$ o operador de dispersão, e $\boldsymbol{\Sigma}_{\Delta_y} \in \mathbb{R}^{n \times n}$ a matriz variância positiva-definida de Δ_y . A matriz variância $\boldsymbol{\Sigma}_{\Delta_y}$ é obtida pela aplicação

da lei de propagação de variância em Δ_y , que neste caso será o resultado da soma das matrizes variâncias das medidas geodésicas da Época 1 ($\Sigma_y^{\{1\}} \in \mathbb{R}^{n \times n}$) e da Época 2 ($\Sigma_y^{\{2\}} \in \mathbb{R}^{n \times n}$).

Quando o modelo nulo na Eq. (4) for considerado verdadeiro, o parâmetro escalar x pode ser estimado pelo cálculo simples do método dos mínimos quadrados como:

$$\hat{x}_{(\mathcal{H}_0)} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \Delta_y \quad (5)$$

e os erros estimados como sendo:

$$\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_0)} = \mathbf{A} \hat{x} - \Delta_y \quad (6)$$

onde $\mathbf{W} \in \mathbb{R}^{n \times n}$ é a conhecida matriz dos pesos, sendo $\mathbf{W} = \sigma_0^2 \Sigma_{\Delta_y}^{-1}$, onde σ_0^2 é o fator de variância a priori. O grau de liberdade global r (redundância) do modelo sob \mathcal{H}_0 é dado por:

$$r = \text{rank} \left(\Sigma_{\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_0)}} \right) = n - \text{rank}(\mathbf{A}) = n - 1, \text{ onde} \quad (7)$$

$$\Sigma_{\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_0)}} = \sigma_0^2 \mathbf{W}^{-1} - \sigma_0^2 \mathbf{A} (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \quad (8)$$

onde $\Sigma_{\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_0)}} \in \mathbb{R}^{n \times n}$ é a matriz variância dos erros estimados.

Por outro lado, um modelo alternativo pode ser proposto quando houver dúvidas sobre a estabilidade dos pontos da rede. Aqui, para simplificar as análises, será considerado que as medições não são contaminadas por outliers (obviamente que em caso real, testes estatísticos prévios devem ser conduzidos para identificação de possíveis outliers com certo nível de probabilidade). Então, para o caso do procedimento univariado, o modelo com relação a hipótese alternativa, denotado por \mathcal{H}_A , é para se opor a Eq. (4) a partir de um modelo estendido que inclui um parâmetro extra $\nabla \in \mathbb{R}^{p \times 1}$ que descreve a perturbação nas medidas em função do deslocamento de um subconjunto “ p ” de pontos da rede, conforme segue:

$$\mathcal{H}_A: E(\Delta_y) = \mathbf{A}x + \mathbf{G}\mathbf{v} = (\mathbf{A} \quad \mathbf{G}) \begin{pmatrix} x \\ \mathbf{v} \end{pmatrix}, D(\Delta_y) = \Sigma_{\Delta_y} = \Sigma_y^{\{1\}} + \Sigma_y^{\{2\}} \quad (9)$$

com $\mathbf{G} \in \mathbb{R}^{n \times p}$ sendo a matriz que representa a relação geométrica entre os “ p ” pontos deslocados e as mudanças das medidas relacionadas a esses pontos (doravante, esta matriz será chamada de “matriz design dos deslocamentos”). Se o modelo alternativo na Eq. (9) for verdadeiro, então os parâmetros desconhecidos serão obtidos como:

$$\begin{pmatrix} \hat{x} \\ \hat{\mathbf{v}} \end{pmatrix}_{(\mathcal{H}_A)} = [(\mathbf{A} \quad \mathbf{G})^T \mathbf{W} (\mathbf{A} \quad \mathbf{G})]^{-1} (\mathbf{A} \quad \mathbf{G})^T \mathbf{W} \Delta_y \quad (10)$$

A redundância do modelo sob \mathcal{H}_A é $n - \text{rank}(\mathbf{A} \quad \mathbf{G})$, com o vetor dos erros estimados $\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_A)}$ e sua respectiva matriz variância $\Sigma_{\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_A)}}$ sob a hipótese alternativa \mathcal{H}_A dadas respectivamente por:

$$\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_A)} = (\mathbf{A} \quad \mathbf{G}) \begin{pmatrix} \hat{x} \\ \hat{\mathbf{v}} \end{pmatrix} - \Delta_y \quad (11)$$

$$\Sigma_{\hat{\mathbf{e}}_{\Delta_y(\mathcal{H}_A)}} = \sigma_0^2 \mathbf{W}^{-1} - \sigma_0^2 (\mathbf{A} \quad \mathbf{G}) [(\mathbf{A} \quad \mathbf{G})^T \mathbf{W} (\mathbf{A} \quad \mathbf{G})]^{-1} (\mathbf{A} \quad \mathbf{G})^T \quad (12)$$

Como um exemplo simples, uma hipótese alternativa pode então ser formulada para o caso de ter um ponto específico como deslocado ($p = 1$), como pode ser visto na ilustração de uma rede de trilateração na Fig. (1). Presume-se que apenas o ponto F foi deslocado para uma nova posição F'. Conseqüentemente, isso causa mudanças apenas nas medidas relacionadas a este ponto. Isso significa que tais distâncias medidas na segunda época sofrem deformações no sentido de aumentar e/ou diminuir em relação aos seus estados iniciais, enquanto as demais medidas relacionadas com os outros pontos permanecem estáveis. O estado inicial refere-se à primeira época em que as medições foram coletadas, de modo que todos os pontos da rede são considerados livres de deslocamentos nesta primeira época. Considerando então o ponto F deslocado na época 2, por exemplo, a matriz $(\mathbf{A} \quad \mathbf{G})$ é dada por:

$$(\mathbf{A} \quad \mathbf{G})^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

O sinal atribuído aos elementos da matriz design dos deslocamentos \mathbf{G} depende do sinal das diferenças de observações geodésicas entre as duas épocas. Um sinal positivo significa que na época 2 a medida foi distorcida positivamente e, portanto, o parâmetro atua positivamente (+ ∇). O mesmo ocorre para o caso em que a medida é reduzida em relação ao seu estado inicial e, portanto, o parâmetro atua com sinal negativo (- ∇). Neste exemplo (Fig. 1), o ponto F deslocado para sua nova posição F' torna o segmento AF' menor que seu estado inicial AF. Conseqüentemente, o parâmetro de perturbação naquela medição em função do deslocamento do ponto F atua com sinal negativo (- ∇). O mesmo ocorre para BF, ou seja, o segmento BF' fica menor que o segmento BF (- ∇), enquanto para o caso de CF o sinal é positivo, pois o segmento CF' fica maior que CF (+ ∇).

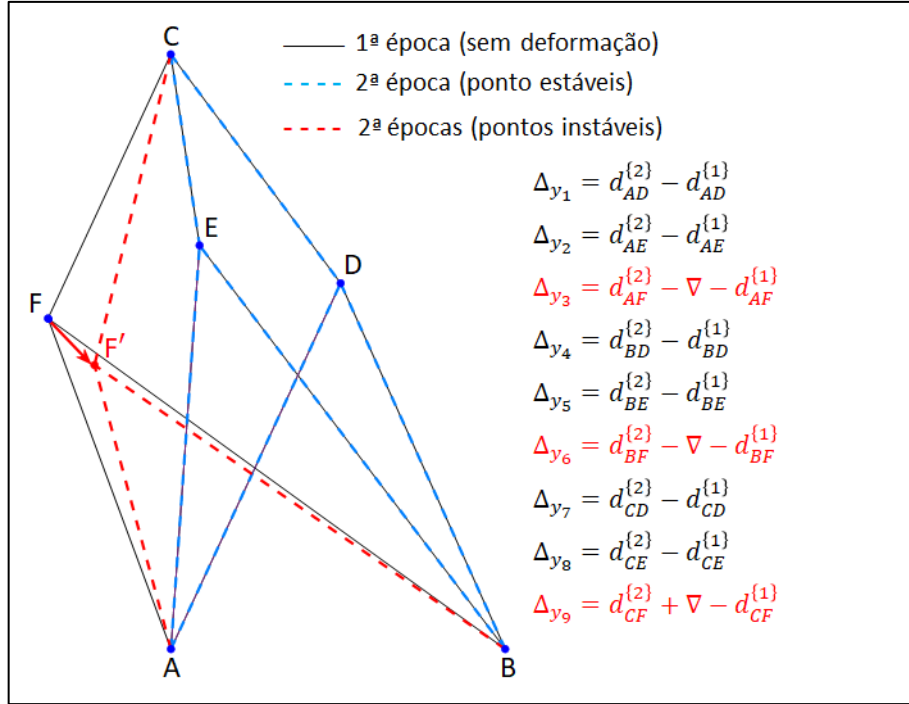


Figura 1 - Rede de trilateração horizontal deformada devido ao deslocamento do ponto F para uma nova posição F'.

A matriz de design dos deslocamentos \mathbf{G} na expressão (10) foi construída para o caso de assumir apenas um único ponto deslocado, que em nosso exemplo foi o ponto F. Nesse caso, o vetor de parâmetro desconhecido ∇ tornou-se um escalar ∇ , e a matriz \mathbf{G} foi reduzida a um vetor unitário, cujos elementos são formados exclusivamente pelos valores de 0, 1 e -1, onde 1 ou -1 significa que o i -ésimo parâmetro de magnitude $+\nabla$ ou $-\nabla$, respectivamente, afeta as medidas conectadas ao ponto e 0 significa o contrário. No entanto, como dado pela Eq. (9), pode-se projetar uma matriz global de design dos deslocamentos \mathbf{G} , de modo que cada coluna represente o deslocamento de um determinado ponto. Para isso, primeiro pode-se fornecer uma matriz auxiliar $\mathbf{C} \in \mathbb{R}^{n \times p}$ que descreve o caso geral da relação entre as medidas (linhas da matriz) e os pontos da rede (colunas da matriz), mas sem levar em conta o sinal do vetor ∇ . Chamando esta matriz \mathbf{C} de Matriz de Conexão dos Pontos, e considerando a rede da Fig. (1) como exemplo, ela seria construída da seguinte forma:

$$\mathbf{C} = \begin{pmatrix} \text{A} & \text{B} & \text{C} & \text{D} & \text{E} & \text{F} \\ \tilde{1} & \tilde{0} & \tilde{0} & \tilde{1} & \tilde{0} & \tilde{0} \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad (14)$$

Então, os sinais dos coeficientes da matriz design dos deslocamentos \mathbf{G} podem ser obtidos genericamente a partir da aplicação de uma função “sinal” como:

$$\text{sinal}(\Delta y_{(l)}) := \begin{cases} +1, & \text{se } \frac{\Delta y_{(l)}}{|\Delta y_{(l)}|} > 0 \\ -1, & \text{se } \frac{\Delta y_{(l)}}{|\Delta y_{(l)}|} < 0, \text{ para } l = 1, \dots, n \\ 0, & \text{se } \Delta y_{(l)} = 0 \end{cases} \quad (15)$$

onde $\text{sinal}(\cdot)$ é a função sinal e $|\Delta y_{(l)}|$ fornece um valor não negativo (valor absoluto) para cada diferença de observação geodésica de duas épocas. O resultado da Eq. (15) pode ser organizado em um vetor como $\Delta \mathbf{y}_{sgn} = \text{sinal}(\Delta y_{(l)})$. Assim, a matriz design dos deslocamentos para o caso geral torna-se:

$$\mathbf{G}_p = \text{diag}(\Delta \mathbf{y}_{sgn}) \mathbf{C} \quad (16)$$

onde $\text{diag}(\cdot)$ é o operador de diagonalização que converte um vetor em uma matriz diagonal.

Consequentemente, a hipótese alternativa é agora dada por:

$$\mathcal{H}_A: E(\Delta_y) = Ax + G_p \nabla_p = (A \quad G_p) \begin{pmatrix} x \\ \nabla_p \end{pmatrix}, D(\Delta_y) = \Sigma_{\Delta_y} = \Sigma_y^{\{1\}} + \Sigma_y^{\{2\}} \quad (17)$$

A matriz design dos deslocamentos G_p pode ser desenvolvida de várias maneiras, dependendo de qual ponto ou grupo de pontos será escolhido para ser monitorado. . Tomando outro exemplo na Fig. (2), se $n = 10$ e $p = 2$, então uma possível matriz design dos deslocamentos G_p poderia ser $G_{p=2} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}^T$ e $\nabla_{p=2} = \begin{pmatrix} \nabla_A \\ \nabla_B \end{pmatrix}$ para o caso onde os pontos A e B foram considerados simultaneamente como pontos suspeitos de terem deslocamentos. É importante observar que neste exemplo foram selecionadas a primeira e a segunda colunas da matriz C , que se referem aos pontos A e B, respectivamente.

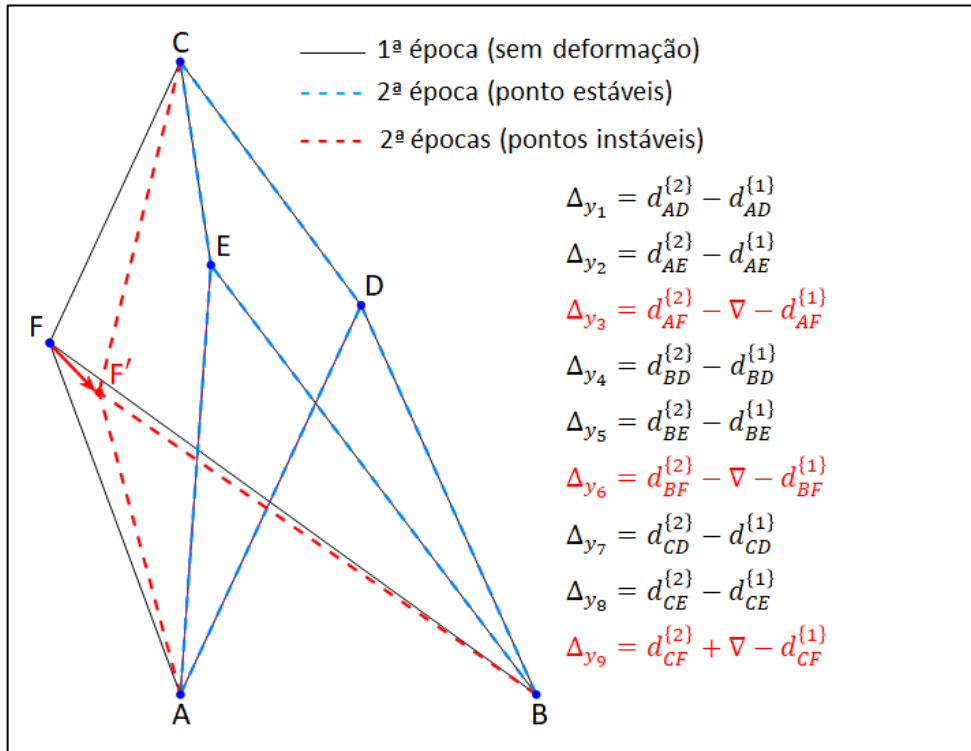


Figura 2 - Rede de trilateração horizontal deformada devido ao deslocamento simultâneo (degrau) do ponto A para uma nova posição A' e B para B'.

Depois de postular as hipóteses nula e alternativa, a estatística de teste advinda do conceito do teste de razão de verossimilhança máxima deve ser calculada da seguinte forma:

$$T_p = \hat{\mathbf{e}}_{\Delta_y}^T \mathbf{W} \mathbf{G}_p \left(\mathbf{G}_p^T \mathbf{W} \boldsymbol{\Sigma}_{\hat{\mathbf{e}}_{\Delta_y}} \mathbf{W} \mathbf{G}_p \right)^{-1} \mathbf{G}_p^T \mathbf{W} \hat{\mathbf{e}}_{\Delta_y} \quad (18)$$

Em seguida, uma decisão de teste é então realizada como:

$$\text{Aceitar } \mathcal{H}_0 \text{ se } T_p \leq c_0, \text{ Rejeitar caso contrário em favor de } \mathcal{H}_A \quad (19)$$

Sob a hipótese nula \mathcal{H}_0 , a estatística de teste T_p segue a distribuição qui-quadrada central com p graus de liberdade; sob a hipótese alternativa \mathcal{H}_A , a estatística de teste T_p segue a distribuição qui-quadrada não central com p graus de liberdade e parâmetro de não centralidade λ_0 . Se a estatística de teste T_p é maior do que um valor crítico ($T_p > c_0$), então uma combinação particular de pontos testados é sinalizado como deslocado, onde c_0 é o valor crítico da distribuição qui-quadrada considerando um determinado nível de significância α_0 com p graus de liberdade, i.e., $c_0 = \chi_{(p, \alpha_0)}^2$. O nível de significância α_0 é frequentemente chamado de “probabilidade de falso alarme”. Portanto, o modelo da hipótese nula é avaliada pela especificação do nível de significância α_0 .

O modelo da hipótese alternativa na Eq. (17) foi formulado sob a suposição de que um grupo específico de pontos sofreu deslocamento. Este grupo de pontos pode ser formulado para os casos em que temos um único ponto deslocado (Fig. 1) ou um subconjunto desses pontos da rede geodésica (Fig. 2). Ou seja, o número de pontos a serem monitorados e suas localizações na rede foram totalmente especificados para definir a hipótese alternativa. Este é um caso especial de testar a hipótese nula contra apenas uma única hipótese alternativa e, portanto, a

rejeição do modelo nulo implica automaticamente na aceitação do modelo alternativo e vice-versa.

2.2 HIPÓTESE NULA CONTRA MÚTIPLAS HIPÓTESES ALTERNATIVAS

Uma das questões que surgem é como definir a quantidade de pontos e suas posições na estrutura a ser monitorada. Claramente, os geodestas precisam trabalhar de forma interdisciplinar com outros profissionais de outras áreas para que um modelo alternativo seja adequadamente formulado. Porém, mesmo com tais informações, ainda se pode ter dúvidas se existem pontos que a priori eram considerados estáveis, quando na verdade podem estar se deslocando. Do ponto de vista prático, portanto, muitas vezes não se sabe bem quantos e quais pontos são ou não estáveis. Assim, uma hipótese alternativa mais conservadora seria assumir que existe pelo menos um ponto que se move entre todos os pontos envolvidos na rede geodésica. Conseqüentemente, ter-se-ia n_p modelos alternativos para $p = 1$ a serem testados contra a hipótese nula, ou seja:

$$\mathcal{H}_A^{(g_i)}: E \left(\Delta_y^{(g_i)} \right) = Ax + G_{p=1}^{(g_i)} \nabla_{p=1}^{(g_i)}, \quad i = 1, \dots, n_p \quad (20)$$

Note que $G_{p=1}^{(g_i)}$ representa o i -ésimo vetor coluna da matriz design dos deslocamentos $G_{p=1}$. Para o caso da rede ilustrada na Fig. 1 e Fig. 2, poderia ter 6 hipóteses alternativas, pois tem-se um total de $n_p = 6$ pontos, tal que $g_1 := [A]$; $g_2 := [B]$; $g_3 := [C]$; $g_4 := [D]$; $g_5 := [E]$; $g_6 := [F]$. Isto significa que ter-se-ia que testar \mathcal{H}_0 contra $\mathcal{H}_A^{(A)}, \mathcal{H}_A^{(B)}, \mathcal{H}_A^{(C)}, \mathcal{H}_A^{(D)}, \mathcal{H}_A^{(E)}, \mathcal{H}_A^{(F)}$, i.e.:

$$\begin{aligned}
\mathcal{H}_0 \text{ vs } \mathcal{H}_A^{(A)}: E(\Delta_y^{(A)}) &= Ax + G_{p=1}^{(A)} \nabla_{p=1}^{(A)} \\
\mathcal{H}_0 \text{ vs } \mathcal{H}_A^{(B)}: E(\Delta_y^{(B)}) &= Ax + G_{p=1}^{(B)} \nabla_{p=1}^{(B)} \\
\mathcal{H}_0 \text{ vs } \mathcal{H}_A^{(C)}: E(\Delta_y^{(C)}) &= Ax + G_{p=1}^{(C)} \nabla_{p=1}^{(C)} \\
\mathcal{H}_0 \text{ vs } \mathcal{H}_A^{(D)}: E(\Delta_y^{(D)}) &= Ax + G_{p=1}^{(D)} \nabla_{p=1}^{(D)} \\
\mathcal{H}_0 \text{ vs } \mathcal{H}_A^{(E)}: E(\Delta_y^{(E)}) &= Ax + G_{p=1}^{(E)} \nabla_{p=1}^{(E)} \\
\mathcal{H}_0 \text{ vs } \mathcal{H}_A^{(F)}: E(\Delta_y^{(F)}) &= Ax + G_{p=1}^{(F)} \nabla_{p=1}^{(F)}
\end{aligned} \tag{21}$$

Agora lida-se com testes envolvendo múltiplas hipóteses alternativas. Agora, o interesse é em saber qual das hipóteses alternativas pode levar à rejeição da hipótese nula com uma certa probabilidade. Para responder a isso, a estatística de teste na Eq. (18) agora é calculada para cada um desses testes $(T_{p=1}^{(gi)})$, tal que um vetor de estatísticas de teste poderia ser construído, denotado por $T_{p=1}$. Nesse caso, a estatística de teste que será destacada será o valor máximo entre todos os valores, que é obtido como:

$$\max(T_{p=1}) = \max_{i \in \{1, \dots, n_p\}} T_{p=1}^{(gi)} \tag{22}$$

Então, obtida a estatística de teste que possui o valor máximo e a hipótese alternativa associada, a regra de decisão para este caso é dada por:

$$\text{Aceitar } \mathcal{H}_0 \text{ se } \max(T_{p=1}) \leq c,$$

$$\text{caso contrário rejeitá – la em favor da } \mathcal{H}_A^{(gi)} \text{ relacionada ao } \max(T_{p=1}) \tag{23}$$

A regra de decisão em (23) afirma que se nenhum dos n_p testes for rejeitado, então se aceita a hipótese nula \mathcal{H}_0 . Se $\max(T_{p=1})$ for maior que o valor crítico c pré-estipulado (com certo nível de significância), então há evidências de que há um ponto deslocado na estrutura. Neste caso, só se pode assumir que a detecção ocorreu. A identificação, no entanto, não é simples. A identificação (ou localização) do ponto deslocado consiste em buscar o ponto que

produziu a estatística de teste máxima - $\max(\mathbf{T}_{p=1})$ - e cujo valor é maior que o valor crítico pré-estipulado c . Assim, a identificação do ponto só acontece quando necessariamente existe a detecção do ponto, ou seja, a “identificação do ponto” só ocorre quando a hipótese nula \mathcal{H}_0 é rejeitada. Isto significa que a detecção correta não implica necessariamente uma identificação correta. Por exemplo, podemos detectar corretamente a ocorrência de deformação, mas identificar erroneamente um ponto como deslocado, quando na verdade é outro ponto que se deslocou. A descrição das taxas de sucesso e fracasso do método será abordada nas próximas seções.

É importante destacar que a estatística de teste máxima ($\max(\mathbf{T}_p)$) é tratada diretamente como uma estatística de teste. Assim, a distribuição de $\max(\mathbf{T}_p)$ não pode ser derivada de distribuições de teste bem conhecidas (por exemplo, distribuição qui-quadrado). Portanto, os valores críticos não podem ser retirados de uma tabela estatística, mas devem ser calculados numericamente. Aqui, o valor crítico c é calculado por Monte Carlo de modo que um erro de decisão Tipo I definido pelo usuário “ α_t ” para o procedimento proposto seja garantido. Na próxima seção, será fornecido detalhes sobre como calcular o valor crítico c para o procedimento sequencial proposto.

O procedimento descrito até agora permite identificar apenas um único ponto. No entanto, pode-se ter vários pontos deslocados simultaneamente. Uma abordagem mais apropriada seria aplicar um teste sequencial para decidir o número e a localização dos pontos deslocados para o caso em que o modelo nulo é rejeitado. Se o valor máximo da estatística de teste ultrapassar o valor crítico no nível de significância adotado, ou seja, se $\max(\mathbf{T}_{p=1}) > c$, a estatística de teste prossegue para $p = 2$ (dois possíveis pontos instáveis). Como não se sabe qual grupo de dois pontos pode estar deslocando, tem-se que calcular as estatísticas de teste pela Eq. (18) a partir das combinações possíveis para dois pontos instáveis ($\mathbf{T}_{p=2}$), e então

identificando o grupo candidato correspondente para $p = 2$ por meio do valor máximo da estatística de teste $\max(\mathbf{T}_{p=2})$. Observe que a estatística de teste extremo $\max(\mathbf{T}_p)$ retorna apenas um grupo de pontos ou ponto. Em geral, o número de grupos possíveis de pontos, denotado por $K_{n_p}^p$ é dado por:

$$K_{n_p}^p = \binom{n_p}{p} = \frac{n_p!}{(n_p-p)!p!} \quad (24)$$

Por exemplo, para uma rede geodésica com $n_p = 6$ pontos e assumindo $p = 2$ grupo de dois pontos a ser testado, ter-se-ia um total de $K_6^2 = 15$ combinações possíveis. Este seria o caso da rede geodésica aqui apresentada, como pode ser visto tanto pela Fig. 1 quanto pela Fig. 2, que resulta em 15 modelos alternativos para o exemplo citado, ou seja:

$$(\mathcal{H}_A^{(A,B)}, \mathcal{H}_A^{(A,C)}, \mathcal{H}_A^{(A,D)}, \mathcal{H}_A^{(A,E)}, \mathcal{H}_A^{(A,F)}, \mathcal{H}_A^{(B,C)}, \mathcal{H}_A^{(B,D)}, \mathcal{H}_A^{(B,E)}, \mathcal{H}_A^{(B,F)}, \mathcal{H}_A^{(C,D)}, \mathcal{H}_A^{(C,E)}, \mathcal{H}_A^{(C,F)}, \mathcal{H}_A^{(D,E)}, \mathcal{H}_A^{(D,F)}, \mathcal{H}_A^{(E,F)}).$$

Observe que para $p = 1$ é um caso particular de ter $K_{n_p}^{p=1} = n_p$, ou seja, igual o número de pontos da rede, conforme apresentado anteriormente. Assim, os casos gerais das hipóteses alternativas podem ser expressos como:

$$\mathcal{H}_A^{(g_i)}: E(\Delta_y^{(g_i)}) = \mathbf{A}x + \mathbf{G}_p^{(g_i)}\mathbf{V}_p^{(g_i)}, \quad i = 1, \dots, K_{n_p}^p \quad (25)$$

Se o ponto identificado com $\max(\mathbf{T}_{p=1})$ também estiver contido no caso de $\max(\mathbf{T}_{p=2})$, ou seja, se o ponto sinalizado inicialmente estiver entre um dos pares sinalizados posteriormente, então a hipótese nula passa a ser especificada de acordo com o modelo identificado em $\max(\mathbf{T}_{p=1})$ e a hipótese alternativa ao modelo definido em $\max(\mathbf{T}_{p=2})$, como pode ser visto no exemplo a seguir:

$$\begin{aligned}\mathcal{H}_0^{(g_i)}: E(\Delta_{\mathbf{y}}^{(g_i)}) &= \mathbf{Ax} + \mathbf{G}_{p=1}^{(g_i)} \nabla_{p=1}^{(g_i)} \\ \mathcal{H}_A^{(g_j)}: E(\Delta_{\mathbf{y}}^{(g_j)}) &= \mathbf{Ax} + \mathbf{G}_{p=2}^{(g_j)} \nabla_{p=2}^{(g_j)}, \quad g_i \subset g_j\end{aligned}\quad (26)$$

Por exemplo, se $\max(\mathbf{T}_{p=1}) > c$ e o ponto A na rede geodésia ilustrada na Fig. 1 (ou Fig. 2) foi identificado pelo $\max(\mathbf{T}_{p=1})$, na sequência, o modelo sob a hipótese nula se tornaria $\mathcal{H}_0^{(A)}: E(\Delta_{\mathbf{y}}^{(A)}) = \mathbf{Ax} + \mathbf{G}_{p=1}^{(A)} \nabla_{p=1}^{(A)}$. Se no próximo passo para $\mathbf{T}_{p=2}$, os pontos A e C forem identificados pelo $\max(\mathbf{T}_{p=2})$, então na sequência a hipótese alternativa seria definida como sendo $\mathcal{H}_A^{(A,C)}: E(\Delta_{\mathbf{y}}^{(A,C)}) = \mathbf{Ax} + \mathbf{G}_{p=2}^{(A,C)} \nabla_{p=2}^{(A,C)}$. Note que $\mathcal{H}_0^{(A)}$ é um subconjunto de $\mathcal{H}_A^{(A,C)}$. Para decidir qual desses modelos selecionar, calcula-se a estatística de teste baseada na razão de máxima verossimilhança (*maximum likelihood ratio* - MLR) entre \mathcal{H}_0 e \mathcal{H}_A , dada por $\Lambda_{\text{MLR}}(\Delta_{\mathbf{y}})$:

$$\Lambda_{\text{MLR}}(\Delta_{\mathbf{y}}) = \hat{\mathbf{e}}_{\Delta_{\mathbf{y}}(\mathcal{H}_0)}^T \mathbf{W} \hat{\mathbf{e}}_{\Delta_{\mathbf{y}}(\mathcal{H}_0)} - \hat{\mathbf{e}}_{\Delta_{\mathbf{y}}(\mathcal{H}_A)}^T \mathbf{W} \hat{\mathbf{e}}_{\Delta_{\mathbf{y}}(\mathcal{H}_A)} \quad (27)$$

O resultado da Eq. (27) vem da razão entre o máximo da função densidade de probabilidade de $\Delta_{\mathbf{y}}$ sob \mathcal{H}_0 e \mathcal{H}_A e o fato de que a hipótese nula é (e deve ser) um subconjunto da hipótese alternativa. Assim, a seleção da hipótese mais provável é baseada na seguinte decisão:

$$\text{Aceitar } \mathcal{H}_0 \text{ se } \Lambda_{\text{MLR}}(\Delta_{\mathbf{y}}) \leq c, \text{ caso contrário rejeitá-la em favor de } \mathcal{H}_A \quad (28)$$

onde $\hat{e}_{\mathcal{H}_0}$ e $\hat{e}_{\mathcal{H}_A}$ são os erros de observação estimados por mínimos quadrados para o modelo sob as hipóteses nula e alternativa, respectivamente. Para o exemplo em questão, seria $\hat{e}_{\mathcal{H}_0} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{G}_{p=1}^{(A)}\hat{\mathbf{v}}_{p=1}^{(A)} - \Delta_{\mathbf{y}}$ para a hipótese nula e $\hat{e}_{\mathcal{H}_A} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{G}_{p=2}^{(A,C)}\hat{\mathbf{v}}_{p=2}^{(A,C)} - \Delta_{\mathbf{y}}$ para a alternativa.

Se a hipótese nula não for rejeitada, o teste termina e apenas o ponto correspondente a $\max(\mathbf{T}_{p=1})$ é identificado como instável. Por outro lado, se a hipótese nula for rejeitada novamente, então uma nova etapa é iniciada tomando o modelo sob a hipótese alternativa identificada na etapa anterior como a hipótese nula (e.g., $\mathcal{H}_0 \equiv \mathcal{H}_A^{(gi)}$). Então, o procedimento de teste prossegue calculando as estatísticas de teste para $p = 3$ de acordo com a Eq. (18) para todas as combinações possíveis dadas pela Eq. (24), identificando o grupo candidato correspondente para $p = 3$ por meio do valor máximo da estatística de teste $\max(\mathbf{T}_{p=3})$, e então verificando se um dos pontos sinalizados na etapa anterior (ou seja, para $p = 2$) está entre aqueles sinalizados na etapa atual para $p = 3$. Se for esse o caso, então o teste é aplicado por meio da Eq. (26) para decidir entre o modelo nulo definido pelo modelo identificado para $p = 2$ e o modelo alternativo para $p = 3$. Este procedimento sequencial é repetido até que qualquer uma das seguintes condições seja atendida:

- (i) A hipótese nula atual não pode ser rejeitada;
- (ii) Mais do que um grupo de pontos da rede geodésica é identificada pela estatística extrema $\max(\mathbf{T}_p)$ para um dado p , i.e., para o caso onde as hipóteses não podem ser separáveis (sobreposição estatística - *statistical overlap*), e portanto a identificação não pode ser realizada; ou
- (iii) a iteração atinge o limiar p_{max} (i.e., o número máximo de pontos a ser avaliado foi totalmente inspecionado). A definição do número máximo de pontos que pode ser inspecionado será detalhada na próxima seção.

Em geral, portanto, o procedimento de teste sequencial aqui proposto é baseado na razão de verossimilhança, que agora será chamado pela sigla **SLRTUPI** (*Sequential Likelihood Ratio*

Tests for Unstable Points Identification - Testes de Razão de Verossimilhança Sequencial para Identificação de Pontos Instáveis). Consiste em determinar se o subconjunto adicional de pontos deslocados, considerado em cada nova hipótese alternativa, é estatisticamente significativo ou não, em termos do seu impacto na forma quadrática dos erros de observação estimados, semelhante à sua forma para detecção de outliers [27]. Consequentemente, podemos identificar o número e a localização dos pontos da rede geodésica deslocados. A Figura 3 mostra o passo a passo de como executar o procedimento SLRTUPI.

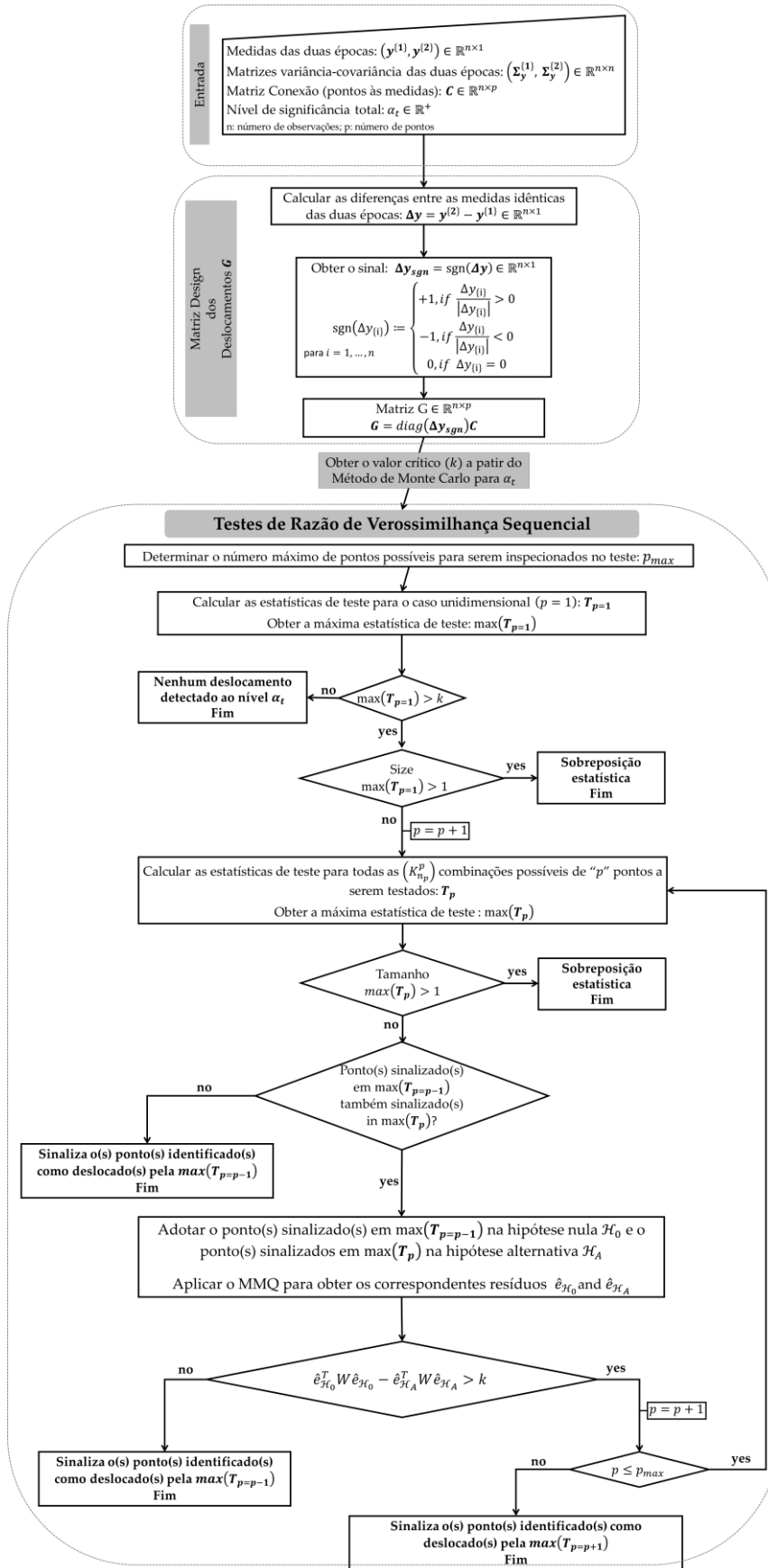


Figura 3 – Fluxograma do método proposto – SLRTUPI.

2.3 MÉTODO MONTE CARLO PARA CONTROLAR A TAXA DE FALSA DETECÇÃO

Ao realizar o procedimento aqui proposto (denominado SLRTUPI) existe a probabilidade de cometer pelo menos uma falsa detecção, também chamado de erro Tipo I, ao se trabalhar com os testes de múltiplas hipóteses alternativas. Assim, especificamente a taxa de erro tipo I para o caso em que o SLRTUPI está sendo aplicado corresponde à probabilidade de detectar incorretamente pelo menos um ponto como deslocado quando na verdade não há nenhum (ou seja, aceitar pelo menos uma hipótese alternativa, quando, na verdade, a hipótese que é verdadeira). Isso significa que o erro de decisão Tipo I do procedimento SLRTUPI não depende de todas as etapas de teste subsequentes, mas apenas da estatística de teste máxima calculada em sua primeira etapa, ou seja, apenas de $\max(\mathbf{T}_{p=1})$ na Eq. (22). Portanto, o risco de rejeitar a hipótese nula (sendo ela verdadeira) pode ocorrer em qualquer uma das $K_{n_p}^{p=1} = n_p$ hipóteses alternativas. Continuando, seja a probabilidade de rejeitar uma hipótese nula \mathcal{H}_0 verdadeira em um teste i dada por α_i (a chamada “taxa de erro experimental) e que $\alpha_i \ll 1$. Além disso, assumindo que os eventos aleatórios “rejeitar uma \mathcal{H}_0 verdadeira em um teste i ” sejam aproximadamente estatisticamente independentes. Então, a probabilidade total de rejeitar uma \mathcal{H}_0 verdadeira em um múltiplo teste de hipóteses (a chamada “taxa de erro familiar”, denotada aqui por α_t) é dada por [11]:

$$\alpha_t \approx \sum_{i=1}^{n_p} \alpha_i \quad (29)$$

Um procedimento clássico e bem conhecido para controlar o α_t seria aplicar a equação de Bonferroni [37]:

$$\alpha_i := \alpha_t/n_p \quad (30)$$

Infelizmente, as estatísticas de teste $T_{p=1}^{(g_i)}$ e consequentemente os eventos aleatórios “rejeição de uma \mathcal{H}_0 verdadeira no teste i ” são estatisticamente dependentes. A estatística extrema $\max(\mathbf{T}_{p=1})$ captura tais dependências, conforme é extraída de $T_{p=1}^{(g_i)}$. Se tais dependências entre as estatísticas de teste forem negligenciadas e usar as Eq. (29) e (30), então os valores críticos calculados com base nesses níveis de significância estarão errados e não se terá garantia do controle do erro tipo I (falsa detecção) desejado.

Aqui, o valor máximo $\max(\mathbf{T}_{p=1})$ na Eq. (22) é tratado diretamente como uma estatística de teste [38]. Note que quando se usa a Eq. (22) como uma estatística de teste, a regra de decisão é baseada em um teste unilateral da forma $\max(\mathbf{T}_{p=1}) \leq c$, como pode ser vista na expressão (23). No entanto, a distribuição de $\max(\mathbf{T}_{p=1})$ não pode ser derivada de alguma distribuição de probabilidade bem definida e conhecida (e.g., distribuição χ^2). Portanto, os valores críticos não podem ser retirados de uma tabela estatística, mas devem ser calculados numericamente. Um cálculo rigoroso de valores críticos requer uma técnica de Monte Carlo.

O procedimento para calcular os valores críticos para $\max(\mathbf{T}_{p=1})$ é da seguinte forma:

1. Gerar uma sequência de m vetores randômicos de erros de medição para ambas as épocas 1 $\mathbf{e}_k^{\{1\}}$ e 2 $\mathbf{e}_k^{\{2\}}$, $k = 1, \dots, m$ da distribuição desejada, e.g.:

$$\begin{aligned} \mathbf{e}_k^{\{1\}} &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_y^{\{1\}}) \text{ para Época 1} \\ \mathbf{e}_k^{\{2\}} &\sim N(\mathbf{0}, \boldsymbol{\Sigma}_y^{\{1\}}) \text{ para Época 2} \end{aligned}, k = 1, \dots, m \quad (31)$$

onde m é conhecido como o número de experimentos Monte Carlo. Por exemplo, o comando "mvnrnd" do Matlab pode ser usado nesta etapa.

2. Para cada par $\mathbf{e}_k^{\{1\}}$ e $\mathbf{e}_k^{\{2\}}$, $k = 1, \dots, m$ calcule as diferenças dos erros entre as duas épocas, i.e.:

$$\Delta_{e_k} = \mathbf{e}_k^{\{2\}} - \mathbf{e}_k^{\{1\}}, k = 1, \dots, m \quad (32)$$

3. Aplique os mínimos quadrados para estimar o vetor das diferenças dos erros de observação, como segue:

$$\widehat{\Delta}_{e_k} = \mathbf{R}\Delta_{e_k}, k = 1, \dots, m \quad (33)$$

onde $\mathbf{R} \in \mathbb{R}^{n \times n}$ é conhecida como z matriz redundância e é dada por:

$$\mathbf{R} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \quad (34)$$

com $\mathbf{I} \in \mathbb{R}^{n \times n}$ sendo a matriz identidade.

4. Monte a matriz de deslocamentos $\mathbf{G}_{p=1}^{(g_i)}, i = 1, \dots, n_p$ para cada experiment Monte Carlo, de acordo com a Eq. (16), considerando os sinais dos seus coeficientes dados por $\mathbf{sgn}(\Delta_{e_k}), k = 1, \dots, m$, e calcule a estatística de teste $\max(\mathbf{T}_{p=1(k)})$ pelas Eq. (18) e (22). A distribuição de frequência de $\max(\mathbf{T}_{p=1(k)})$ aproxima a distribuição de probabilidade de $\max(\mathbf{T}_{p=1})$.

5. Classifique em ordem ascendente os m valores de $\max(\mathbf{T}_{p=1(k)})$, obtendo um vetor classificado $\mathbf{T}\mathbf{s}$, tal que:

$$\mathbf{T}\mathbf{s} = T_S^{(1)} \leq T_S^{(2)} \leq \dots \leq T_S^{(m)} \quad (35)$$

Os valores classificados em $\mathbf{T}\mathbf{s}$ fornecem uma representação discreta da função de densidade cumulativa da estatística máxima de teste $\max(\mathbf{T}_{p=1})$.

6. Determina o valor crítico c :

$$c = \frac{1}{2} (\mathbf{T}\mathbf{s}_{[(1-\alpha_t)m]} + \mathbf{T}\mathbf{s}_{[(1-\alpha_t)m]+1}) \quad (36)$$

onde $[\cdot]$ denota arredondamento para o próximo número inteiro que indica a posição dos elementos selecionados na ordem ascendente de $\mathbf{T}\mathbf{s}$. Esta posição corresponde

a um valor crítico para uma probabilidade geral estipulada de falsa detecção α_t . Isto pode ser feito para uma sequência de valores α_t em paralelo [34].

2.4 DETERMINAÇÃO DO NÚMERO MÁXIMO DE PONTOS POSSÍVEIS DE INSPEÇÃO -

Nesta seção é demonstrada uma forma objetiva que desenvolvemos nesta pesquisa para calcular o número máximo simultâneo de pontos “ p_{max} ” que pode ser verificado pelo procedimento SLRTUPI.

O procedimento para definir o número máximo de pontos “ p_{max} ” consiste em encontrar um modelo regular (i.e., uma matriz $(\mathbf{A} \quad \mathbf{G}_p)$ com posto completo) e, claro, que esse modelo tenha redundância suficiente para a identificação de pontos instáveis, conforme explicado abaixo.

Primeiro, se a matriz design dos deslocamentos \mathbf{G}_p é especificada considerando o número total de pontos na rede, ou seja, $p = n_p$ de acordo com a equação (16), gera-se um problema de deficiência de posto da matriz $(\mathbf{A} \quad \mathbf{G}_{p=n_p})$, ou seja, não é possível inspecionar os n_p simultaneamente pelo SLRTUPI. Ao reduzir o número de pontos em uma unidade, ou seja, $p = n_p - 1$, esperamos encontrar modelos regulares (sem deficiência de posto). Para isso, deve-se verificar novamente se há ou não deficiência de posto. Mas agora tem-se a disposição não apenas um grupo de pontos, mas $K_{n_p}^{n_p-1} = n_p$ grupos possíveis de $n_p - 1$ pontos. Se ao menos um modelo for encontrado como irregular (ou seja, com deficiência de posto), então deve-se reduzir novamente p para $p - 1$, i.e., $p = n_p - 2$. Caso contrário, caso não tenha nenhum modelo com deficiência de posto, o próximo passo é avaliar se as estatísticas de teste calculadas pela Equação (18) são todas diferentes entre si. Em outras palavras, mesmo no caso de

termos todas as matrizes com suas matrizes normais como não singulares, ainda devemos verificar se as estatísticas de teste são distinguíveis umas das outras pela Equação (18). Se houver pelo menos duas estatísticas com valores iguais, será sinalizada uma sobreposição estatística. Neste caso, teríamos que reduzir novamente p para $p - 1$, no qual resultaria em $p = n_p - 3$. Então, todo o procedimento deve ser repetido para o novo valor de p . Quando o procedimento encontrar o p que satisfaz tanto a questão de não ter deficiência de posto e de não ter sobreposição de estatísticas de teste, é encontrado o número de pontos simultâneos máximo (p_{max}) que o método SLRTUPI pode verificar se deslocaram ou não. Com o p_{max} encontrado o procedimento SLRTUPI pode ser executado conforme o fluxograma da Figura 3.

3. EXPERIMENTOS E RESULTADOS OBTIDOS

O volume de resultados gerados nos experimentos foi grande e impossível de colocar tudo aqui neste relatório (tem a questão da limitação do tamanho do arquivo na Plataforma Carlos Chagas). Como já comentado no final da seção de Introdução, enfatizamos aqui que acessem este link para terem todos os resultados que foram gerados nesta pesquisa (<https://drive.google.com/file/d/1mkPFV19mbdXbN84B701cCrlohNYpR827/view>). Na sequência um resumo geral será apresentado.

Os experimentos foram realizados em dois cenários, a saber: (i) avaliação da eficiência do método proposto SLRTUPI por meio de simulações computacionais (ou seja, em ambiente controlado); (ii) deslocamentos simulados em campo (ambiente semicontrolado mas executando medidas reais em campo). No primeiro cenário, considerou-se os seguintes casos: deslocamento de pontos individuais envolvendo apenas um único ponto deslocado por vez; dois pontos mutuamente deslocados; e três pontos mutuamente deslocados. Nos experimentos tanto simulados quanto de campo, as medidas envolvidas foram somente de distâncias horizontais entre os pontos. Obviamente, o método SLRTUPI pode também ser empregado envolvendo outros tipos de medidas como angulares, de desníveis, etc. Em trabalhos futuros pretende-se fazer novos experimentos envolvendo vários tipos de medidas. A Figura 4 apresenta a rede geodésicas envolvida no experimento simulado e que também foi materializada em campo para os experimentos com dados reais.

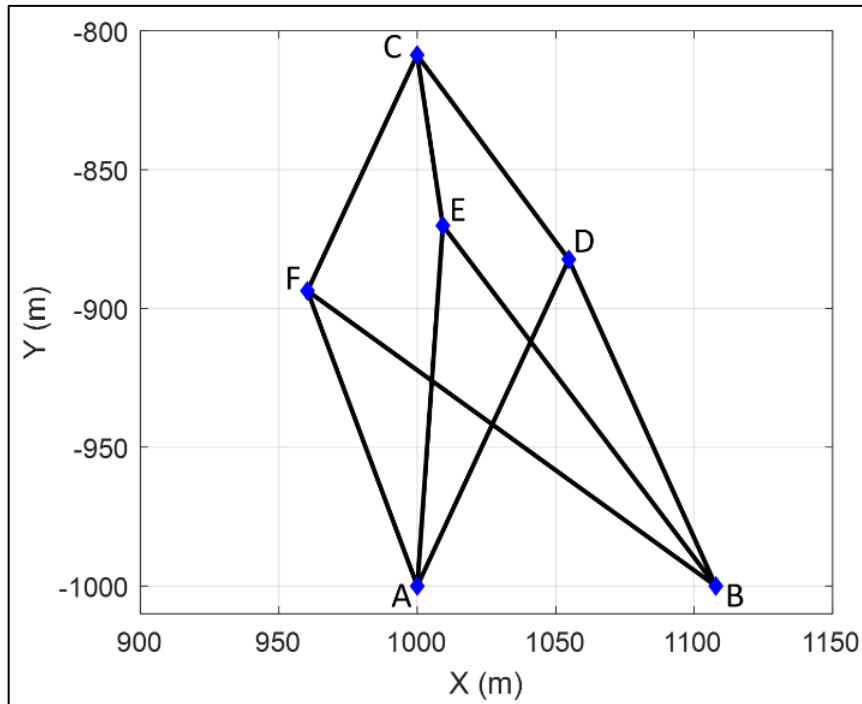


Figura 4 – Rede de Trilateração envolvida nos experimentos

Na primeira análise das simulações, consideramos o caso de haver apenas um ponto deslocado por vez. Nesse caso, tivemos então 6 grupos simulados individualmente, a saber: $g_1 := [A]$, $g_2 := [B]$, $g_3 := [C]$, $g_4 := [D]$, $g_5 := [E]$, $g_6 := [F]$. A magnitude do deslocamento para cada grupo foi definida dentro do intervalo de $[1,10]\sigma$ com incrementos de 1σ , onde σ é o desvio padrão da observação de distância, tomado como $\sigma = 2mm$. Os deslocamentos foram simulados para atuar em diferentes orientações. A orientação do deslocamento θ foi simulada para a faixa de $[0^\circ, 355^\circ]$ com incremento de 5° . Como resultado, tivemos dez magnitudes de deslocamento para cada uma das 72 orientações, totalizando 720 deslocamentos simulados para cada ponto. Consideramos 200.000 experimentos Monte Carlo para cada um desses 720 deslocamentos simulados para cada ponto, e então calculamos as taxas de correta identificação e correta detecção. As taxas são calculadas como a razão entre a ocorrência de um evento específico – correta detecção ou correta identificação – e o número de

casos possíveis (ou seja, número total de experimentos Monte Carlo $m = 200.000$), que são dados respectivamente por [38]:

$$\mathcal{P}_{CD} = \frac{n_{CD}}{m} \quad (37)$$

$$\mathcal{P}_{CI} = \frac{n_{CI}}{m} \quad (38)$$

onde:

n_{CD} é o número de vezes que ocorre a correta detecção, ou seja, é o número de experimentos em que o procedimento SLRTUPI detecta corretamente a presença de deslocamentos.

n_{CI} é o número de vezes que ocorre a correta identificação, ou seja, é o número de experimentos em que o procedimento SLRTUPI identifica corretamente o deslocamento.

É importante ressaltar que a “detecção de deslocamento” apenas nos informa a existência de pontos instáveis. Para que exista a identificação, a detecção deve ter ocorrido. A detecção não garante a correta identificação, pois, pode identificar o ponto errado. Isso é explicado detalhadamente por Rofatto et al. [38]. Ao se falar de monitoramento de deslocamentos, a detecção, mesmo se ocorrer posterior identificação errada, ela já tem grande importância: pensemos em um monitoramento de uma barragem; só o fato de detectar que existe deslocamentos já é sucesso do ponto de vista da segurança.

Continuando, a Figura 5 mostra o resultado da taxa de identificação correta em cada ponto da rede. Os círculos pretos representam a faixa radial de deslocamentos de magnitude de 4mm, 1cm e 2cm. É importante notar que a incerteza das diferenças nas observações geodésicas de duas épocas é $\sigma_{\Delta y} \sim 2,82mm$. Portanto, essas magnitudes de 4mm, 1cm e 2cm correspondem a aproximadamente $1,4\sigma_{\Delta y}$, $2,8\sigma_{\Delta y}$ e $7\sigma_{\Delta y}$. Claramente, observa-se que a probabilidade de

identificação correta depende da geometria da rede. Observe que as maiores taxas de sucesso ocorrem na linha de visada da medição das distâncias. Por outro lado, é mais difícil identificar quando os deslocamentos ocorrem nas direções perpendiculares às visadas, principalmente para magnitudes próximas às incertezas das diferenças das observações entre as duas épocas. ($\nabla \leq 4mm$). O aumento do nível de significância e, conseqüentemente, uma região de rejeição de hipótese nula maior (valores críticos menores) favorece a identificação de deslocamentos de baixa magnitude. Por outro lado, reduz (ligeiramente) a probabilidade de identificar corretamente deslocamentos de maiores magnitudes.

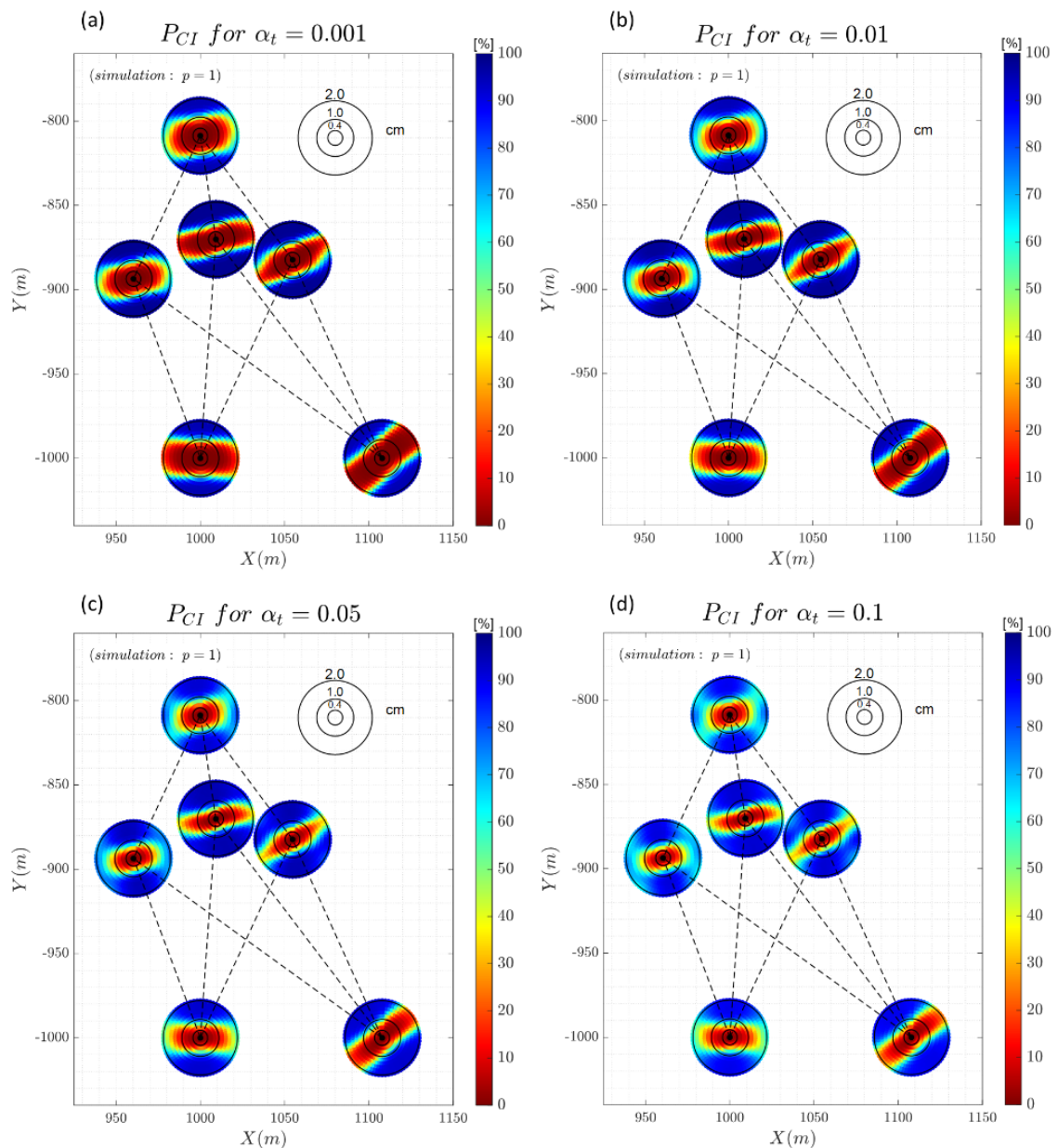


Figura 5 – Taxa de correta identificação \mathcal{P}_{CI} de deslocamentos dos pontos para diferentes níveis de significância

A Figura 6 mostra o resultado da taxa de correta detecção dos deslocamentos dos pontos. A detecção segue o mesmo comportamento da identificação em termos da influência da geometria sob os deslocamentos. Não há diferenças significativas em termos de detecção e identificação para o caso em que é adotado um nível de significância baixo ($\alpha_t = 0,001$). Quanto menor o valor crítico (maior nível de significância), maior foi a taxa de detecção.

Além disso, a detecção é sempre maior que a identificação. Embora a identificação seja importante para localizar o deslocamento, a detecção desempenha um papel dominante para as atividades que envolvem avaliação de risco, como a análise de deformações. É preferível que a detecção ocorra mesmo que a identificação não ocorra da forma esperada. É crucial que um sistema emita um alarme quando são detectados deslocamentos, mesmo que a sua localização possa estar errada. Nesse sentido, observamos também que aumentar o nível de significância nem sempre melhora a taxa de identificação, mas sempre melhora a detecção.

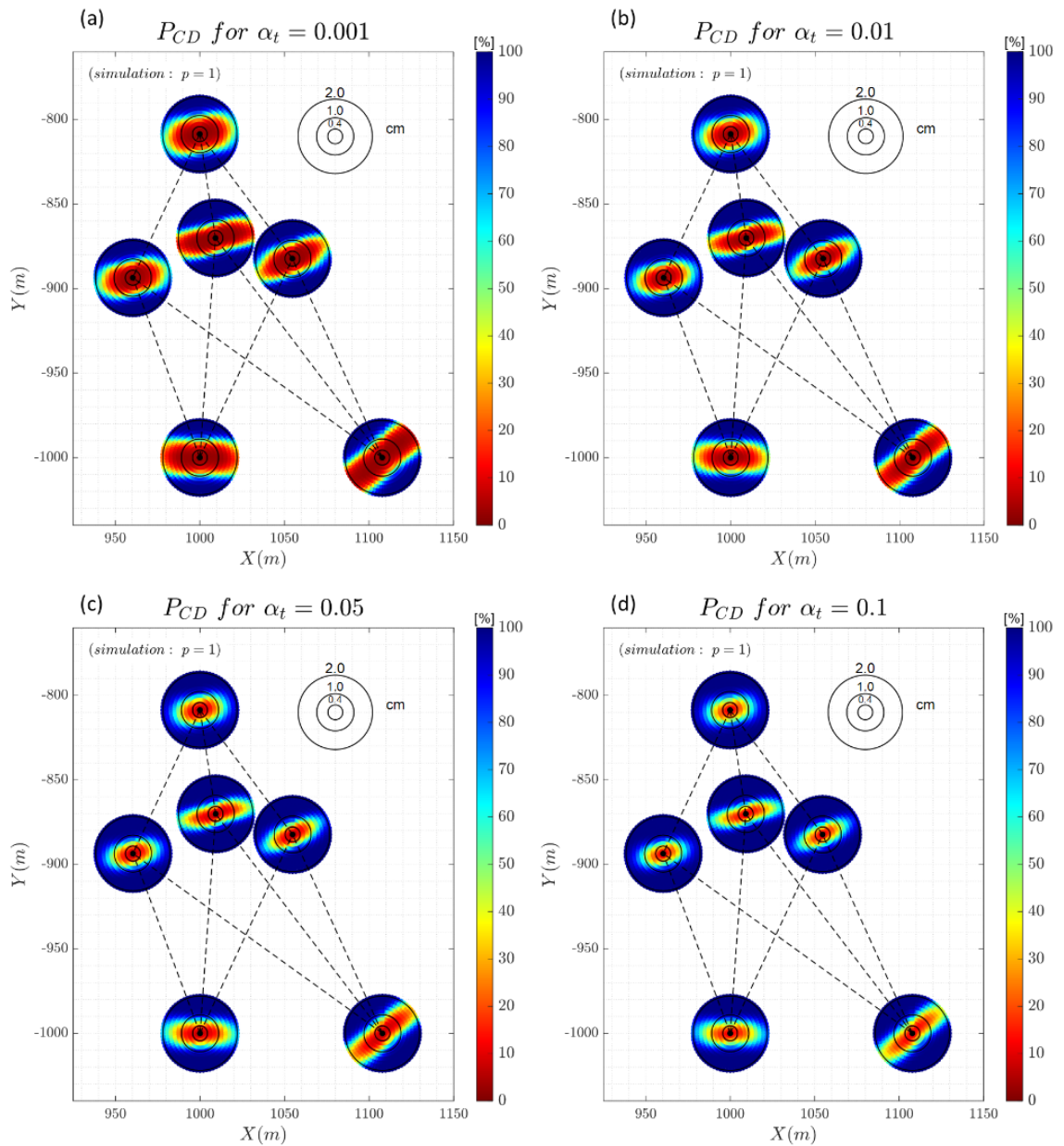


Figura 6 – Taxa de correta detecção \mathcal{P}_{CD} de deslocamentos dos pontos para diferentes níveis de significância.

Na segunda análise simulada foram criados os cenários para o caso de haver dois pontos simultaneamente instáveis. Para isso, simulamos todos os casos possíveis para dois pontos deslocados de cada vez, ou seja, tivemos então 15 grupos de 2 pontos deslocados. A magnitude do deslocamento foi constante para cada grupo, como sendo de $10\sigma = 2\text{cm}$. Os padrões de deslocamento dos grupos de dois pontos variaram em sentido e direção. Em termos de taxa de

sucesso de identificação, o SLRTUPI depende da interação entre o padrão em que os pontos se movem e a geometria da rede geodésica. Uma geometria pode ser melhor para um determinado padrão de deslocamento que se deseja monitorar. Para o caso da rede em questão, os pontos [A, C] seriam mais facilmente identificados se seus deslocamentos ocorressem para $\theta_A = \theta_B = 0^\circ$ ou $\theta_A = \theta_B = 180^\circ$ (ver Tabela 1). Para o caso em que tivéssemos os deslocamentos em orientações opostas, o melhor grupo de pontos a serem monitorados seria [D, E] e [A, E] para o padrão $\theta_D = 0^\circ$ e $\theta_E = 180^\circ$ ou $\theta_A = 0^\circ$ e $\theta_E = 180^\circ$, como pode ser visto na Tabela 2.

Tabela 1 – Máximo e mínimo dos níveis P_{CI} e P_{CD} para $\theta_A = \theta_B = 0^\circ$ e $\theta_A = \theta_B = 180^\circ$

Padrão de deslocamento	α_t	$\frac{\max}{\min}$	$\mathcal{P}_{CI}[\%]$	$\mathcal{P}_{CD}[\%]$	grupo: $\max(\mathcal{P}_{CI})$
$\theta_A = \theta_B = 0^\circ$ ou $\theta_A = \theta_B = 180^\circ$	0,1%	Max	78,79	100,00	[A, C]
		Min	0,93	80,71	
	1%	Max	80,33	100,00	[A, C]
		Min	3,05	86,74	
	5%	Max	78,96	100,00	[A, C]
		Min	3,49	93,10	
	10%	Max	76,99	100,00	[A, C]
		Min	3,30	95,66	

Tabela 2 – Máximo e mínimo dos níveis P_{CI} e P_{CD} para $\theta_D = 0^\circ$ e $\theta_E = 180^\circ$ ou $\theta_A = 0^\circ$ e $\theta_E = 180^\circ$

Padrão de deslocamento	α_t	$\frac{\max}{\min}$	$\mathcal{P}_{CI}[\%]$	$\mathcal{P}_{CD}[\%]$	$g_i: \max(\mathcal{P}_{CI})$
$\theta_D = 0^\circ$ e $\theta_E = 180^\circ$ ou $\theta_A = 0^\circ$ e $\theta_E = 180^\circ$	0.1%	Max	99.92	100.00	[D, E]
		Min	18.32	95.43	
	1%	Max	99.48	100.00	[D, E]
		Min	20.4	99.34	
	5%	Max	97.57	100.00	[A, E]
		Min	20.05	99.95	
	10%	Max	96.16	100.00	[A, E]
		Min	19.43	99.99	

Na última análise em ambiente controlado, foram criados os cenários para o caso de haver três pontos simultaneamente instáveis. Para isso, simulamos todos os casos possíveis para três pontos deslocados de cada vez, ou seja, tivemos então 20 grupos de 3 pontos deslocados. A magnitude do deslocamento também foi mantida constante para cada grupo como sendo de $10\sigma = 2cm$. Os padrões de deslocamento dos grupos de três pontos também variaram em sentido e direção. O padrão descrito como $\theta_1 = 315^\circ$, $\theta_2 = 135^\circ$ e $\theta_3 = 135^\circ$ foi o mais fácil de identificar (Tabela 3). Neste último, os grupos de pontos com maiores taxas de identificação foram [B, D, E] e [B, D, F] com $max(\mathcal{P}_{CI}) \cong 72\%$, e a taxa de sucesso na identificação para este caso não dependeu da escolha do nível de significância.

Tabela 3 – Máximo e mínimo dos níveis P_{CI} e P_{CD} para $\theta_1 = 315^\circ$, $\theta_2 = 135^\circ$ e $\theta_3 = 135^\circ$

Padrão de deslocamento	α_t	$\frac{\max}{\min}$	$\mathcal{P}_{CI}[\%]$	$\mathcal{P}_{CD}[\%]$	$g_i: max(\mathcal{P}_{CI})$
$\theta_1 = 315^\circ$, $\theta_2 = 135^\circ$ e $\theta_3 = 135^\circ$	0.1%	max	72.54	100.00	[B, D, E]
		min	0.00	69.88	
	1%	max	72.53	100.00	[B, D, F]
		min	0.00	83.39	
	5%	max	72.01	100.00	[B, D, F]
		min	0.02	91.13	
	10%	max	71.77	100.00	[B, D, F]
		min	0.06	93.86	

Por fim, no experimento com dados de campo, a Época 1 foi registrada como referência (sem deslocamentos), enquanto na Época 2 aplicou-se intencionalmente deslocamentos em campo. Os deslocamentos nos pontos foram aplicados radialmente. Os pontos A, B e C foram mantidos fixos, ou seja, não sofreram deslocamento, enquanto D, E e F foram considerados os pontos a serem monitorados. Os deslocamentos foram realizados em campo, deslocando o

prisma à 1 cm da posição inicial (Época 1) para as novas posições (Época 2). Foram aplicados deslocamentos para cada ponto individual, dois e três pontos simultâneos. O SLRTUPI foi aplicado em sua forma original, de modo que todos os pontos foram testados como sendo instáveis. Neste caso, tivemos uma taxa média de sucesso na identificação de ~74%. Além disso, a taxa média de correta detecção foi de 100%, ou seja, todos os casos ocorreram a correta detecção. Ao aplicar o SLRTUPI condicionado ao tomar os três pontos (A, B e C) como estáveis, a taxa média de correta identificação foi de ~96%.

4. CONSIDERAÇÕES FINAIS

A pesquisa de pós-doutorado foi concluída com êxito em 12 meses. Foi desenvolvido um novo método para detecção e identificação de pontos instáveis para análise de deformação de estruturas, que denominados de Teste de Razão de Verossimilhança Sequencial para Identificação de Pontos Instáveis (SLRTUPI). O método utiliza as diferenças das próprias medições feitas em duas épocas no tempo, em vez de usar a clássica diferença de coordenadas estimadas. Como vantagem teremos sempre um modelo linear e, portanto, o problema de não linearidade do modelo não afeta o poder do teste. O método SLRTUPI foi implementado em forma de scripts em Matlab, foi aplicado em cenários simulados e também com dados reais de campo.

O método SLRTUPI não se restringe aos pontos sujeitos previamente ao monitoramento, masse desejado também é aplicado de forma que todos os pontos possam ser inspecionados. Contudo, o poder da identificação de pontos instáveis melhora consideravelmente se soubermos antecipadamente quais pontos são estáveis. Os resultados sustentam que o sucesso do método proposto depende da interação entre o padrão de deslocamento dos pontos e a geometria da rede. No caso da rede de trilateração utilizada como estudo de caso nesta pesquisa, observamos que para o caso de ter apenas um único ponto deslocado, as maiores taxas de sucesso ocorrem no alinhamento dos segmentos que unem os pontos. Quando mais de um ponto é instável simultaneamente, as taxas de sucesso são geralmente maiores para os cenários onde os deslocamentos ocorrem em direções opostas. Mais experimentos baseados em diferentes tipos de redes geodésicas (como redes de nivelamento e GNSS) serão realizados em trabalhos futuros.

Quanto maior o número de pontos deslocados simultaneamente, mais difícil será a identificação. No entanto, as taxas de sucesso dependem da geometria e da redundância da rede. Além disso, a detecção é sempre maior que a identificação. Embora a identificação seja importante para localizar o deslocamento, a detecção desempenha um papel dominante para as atividades que envolvem avaliação de risco, como a análise de deformações de grandes estruturas, como por exemplo, barragens. É preferível que a detecção ocorra mesmo que a identificação não ocorra da forma esperada. É crucial que um sistema emita um alarme quando são detectados deslocamentos, mesmo que a sua localização possa estar errada. Nesse sentido, observamos também que aumentar o nível de significância do procedimento SLRTUPI nem sempre melhora a taxa de identificação, mas sempre melhora a detecção.

É importante destacar que o SLRTUPI tem a taxa de erro tipo I (falsa detecção) efetivamente controlada, a partir do uso do método Monte Carlo. Em outras palavras, se o usuário deseja um determinado nível de significância global para o teste, de fato ele será garantido pelo procedimento SLRTUPI. Isto é importante de enfatizar, pois, muitos métodos não fazem esse controle rigoroso.

Outra contribuição importante é que o número máximo de pontos que podem ser inspecionados pelo procedimento SLRTUPI não é baseado em escolhas não objetivas, mas sim determinado a partir da análise do posto das matrizes de design de deslocamentos envolvidas e da ocorrência de sobreposição estatística (igualdade de estatísticas de teste).

Outro fato importante é que a metodologia de simulação que foi empregada nos experimentos também pode ser aplicada na fase de projeto de uma rede geodésica para fins de análise de deformações. Isto permite ao usuário saber a priori como o sistema de medição (neste caso, a rede geodésica) deve ser projetado de forma otimizada para que seja capaz de detectar e/ou identificar um determinado padrão de deslocamento para uma dada probabilidade. Como

resultado, medidas de confiabilidade podem ser facilmente extraídas, como Deslocamento Mínimo Identificável e Deslocamento Mínimo Detectável a partir de suas taxas de sucesso correspondentes.

Mais experimentos para diferentes redes geodésicas serão necessários para avaliar a capacidade do sistema de medição em identificar possíveis medidas com outliers (confiabilidade interna) e o efeito desses outliers não detectados na qualidade dos resultados da análise de deformação (confiabilidade externa).

Por fim, os algoritmos gerados nesta pesquisa e todo conjunto de dados estão disponíveis para os interessados em reproduzir os resultados e/ou aplicá-los a outras metodologias, no seguinte endereço eletrônico:

<https://data.mendeley.com/datasets/msg783rh2y/draft?a=607de244-7c1d-417f-a9c8-81281d8a6056> .

REFERÊNCIAS

[1] Welsch WM, Heunecke O. Models and terminology for the analysis of geodetic monitoring observations. (2001). In: Official report of the ad-hoc committee of FIG working group 6.1. X FIG international symposium on deformation measurements. Fig. Publication No. 25, ISSN 87-90907-10-8.

Retrieved from <https://www.fig.net/resources/publications/figpub/pub25/figpub25.asp>

[2] Pelzer H (1971) Zur Analyse geodätischer Deformationsmessungen, Ph.D. Thesis (in German). Deutsche Geodätische Kommission, Reihe C: Dissertationen - Heft Nr. 164, München, Germany.

[3] van Mierlo J (1978) A testing procedure for analysing geodetic deformation measurements. In: Proceedings of the II. International symposium of deformation measurements by geodetic methods. Bonn, Germany, September 25–28, 1978, Konrad Wittwer, Stuttgart, pp 321–353.

[4] Niemeier W (1981). Statistical tests for detecting movements in repeatedly measured geodetic networks. In: P. Vyskočil, R. Green and H. Mälzer (Editors), Recent Crustal Movements, 1979. Tectonophysics, 71: 335–351.

[5] Caspary WF (2000) Concepts of network and deformation analysis. The University of New South Wales, Kensington.

[6] Niemeier W (2008) Ausgleichsrechnung, statistische auswertemethoden, 2nd edn. de Gruyter, Berlin.

[7] Heunecke O, Kuhlmann H, Welsch WM, Eichhorn A, Neuner H (2013) Handbuch Ingenieurgeodäsie: Auswertung geodätischer Überwachungsmessungen. 2nd edn., Wichmann, Heidelberg.

[8] Chen YQ (1983) Analysis of deformation surveys—a generalized method. Technical Report No. 94. University of New Brunswick. Fredericton.

[9] Nowel K, Kamiński W (2014) Robust estimation of deformation from observation differences for free control networks. *J Geod* 88(8):749–764.

[10] Nowel K (2015) Robust M-estimation in analysis of control network deformations: classical and new method. *J Surv Eng* 141(4):04015002.

[11] Lehmann, Rüdiger and Lösler, Michael. "Congruence analysis of geodetic networks – hypothesis tests versus model selection by information criteria" *Journal of Applied Geodesy*, vol. 11, no. 4, 2017, pp. 271-283. <https://doi.org/10.1515/jag-2016-0049>

[12] Baarda W (1968) A testing procedure for use in geodetic networks, Vol. 2, Number 5, Netherlands Geodetic Commission, Publication on Geodesy, Delft, Netherlands.

Retrieve from <http://www.ncgeo.nl/phocadownload/09Baarda.pdf>

[13] Nowel, K. Specification of deformation congruence models using combinatorial iterative DIA testing procedure. *J Geod* 94, 118 (2020). <https://doi.org/10.1007/s00190-020-01446-9>

[14] Hekimoglu S, Erdogan B, Butterworth S (2010) Increasing the efficacy of the conventional deformation analysis methods: alternative strategy. *J Surv Eng* 136(2):53–62.

[15] Erdogan B, Hekimoglu S (2014) Effect of subnetwork configuration design on deformation analysis. *Surv Rev* 46(335):142–148.

- [16] Durdag UM, Hekimoglu S, Erdogan B (2018) Reliability of models in kinematic deformation analysis. *J Surv Eng* 144(3):04018004.
- [17] Velsink, H. On the deformation analysis of point fields. *J Geod* 89, 1071–1087 (2015).
<https://doi.org/10.1007/s00190-015-0835-z>
- [18] Velsink H (2018) Testing methods for adjustment models with constraints. *J Surv Eng* 144(4):04018009.
- [19] Baselga S (2011) Exhaustive search procedure for multiple outlier detection. *Acta Geod Geophys Hung* 46(4):401–416.
- [20] Biagi L, Caldera S (2013) An efficient leave one block out approach to identify outliers. *J Appl Geod* 7(1):11–19.
- [21] Wujanz D, Krueger D, Neitzel F (2016) Identification of stable areas in unreferenced laser scans for deformation measurement. *Photogram Rec* 31(155):261–280.
- [22] Zienkiewicz M H (2014) Application of Msplit estimation to determine control points displacements in networks with an unstable reference system. *Surv Rev* 47(342):174–180.
- [23] Zienkiewicz M H, Baryla R (2015) Determination of vertical indicators of ground deformation in the old and main city of Gdansk area by applying unconventional method of robust estimation. *Acta Geodyn Geomater* 12(3):249–257
- [24] Wiśniewski Z, ZienkiewiczMH(2016) Shift-M* split estimation in deformation analyses. *J Surv Eng* 142(4):1–13
- [25] Duchnowski R (2010) Median-based estimates and their application in controlling reference mark stability. *J Surv Eng* 136(2):47–52.

- [26] Duchnowski R (2013) Hodges-Lehmann estimates in deformation analyses. *J Geod* 87(10–12):873–884.
- [27] I. Klein, M. T. Matsuoka, M. P. Guzatto & F. G. Nievinski (2017) An approach to identify multiple outliers based on sequential likelihood ratio tests, *Survey Review*, 49:357, 449-457, DOI: 10.1080/00396265.2016.1212970.
- [28] Lazzarini T, Laudyn I, Chrzanowski A, Ga'zdziński J, Janusz W, Wiłun Z, Mayzel B, Mikucki Z (1977) Geodetic measurements of displacements of structures and their surroundings. PPWK, Warsaw (in Polish).
- [29] Chrzanowski A, Chen YQ (1990) Deformation monitoring, analysis and prediction-status report FIG XIX international congress. Helsinki 6(604.1):83–97
- [30] Baarda W., 1973. S-Transformation and Criterion Matrices. Publications on Geodesy, New Series, Vol. 5, No. 1, Netherland Geodetic Commission, Delft, The Netherlands.
- [31] Nowel K (2019) Squared Msplit(q) S-transformation of control network deformations. *J Geod* 93:1025–1044
- [32] Erdogan B., Hekimoglu S., Durdag U M. A new univariate deformation analysis approach considering displacements as model errors. *Stud. Geophys. Geod.*, 65 (2021), 1-14, DOI: 10.1007/s11200-020-1024-y.
- [33] Aydin C (2017) Effects of displaced reference points on deformation analysis. *J Surv Eng* 143(3):1–8.
- [34] Hekimoglu S, Erdogan B, Soycan M, Durdag U M (2014). Univariate Approach for Detecting Outliers in Geodetic Networks. *J Surv Eng* 140(2):1–8.

- [35] Lehmann, R. and Lösler, M. "Hypothesis Testing in Non-Linear Models Exemplified by the Planar Coordinate Transformations" *Journal of Geodetic Science*, vol. 8, no. 1, 2018, pp. 98-114. <https://doi.org/10.1515/jogs-2018-0009>
- [36] Lehmann R (2012) Improved critical values for extreme normalized and studentized residuals in Gauss–Markov models. *J Geod* (86)12: 1137 – 1146. DOI: 10.1007/s00190-012-0569-0.
- [37] Teunissen PJG (2000) *Testing theory: an introduction*. Series on mathematical geodesy and positioning. Delft University Press, Delft.
- [38] Lehmann, R. On the formulation of the alternative hypothesis for geodetic outlier detection. *J Geod* 87, 373–386 (2013). <https://doi.org/10.1007/s00190-012-0607-y>
- [39] Rofatto VF, Matsuoka MT, Klein I, Roberto Veronez M, da Silveira LG Jr. A Monte Carlo-Based Outlier Diagnosis Method for Sensitivity Analysis. *Remote Sensing*. 2020; 12(5):860.
- [40] Abdi H (2007) The Bonferonni and Šidák corrections for multiple comparisons. In: Neil Salkind (ed) *Encyclopedia of measurement and statistics*. Sage, Thousand Oaks.
- [41] Matsumoto, M.; Nishimura, T. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.* 1998, 8, 3–30.
- [42] Box, G.E.P.; Muller, M.E. A Note on the Generation of Random Normal Deviates. *Ann. Math. Stat.* 1958, 29, 610–611.
- [43] Vinicius Francisco Rofatto, M. T. Matsuoka, I. Klein, M. R. Veronez, M. L. Bonimani & R. Lehmann (2020) A half-century of Baarda's concept of reliability: a review, new perspectives, and applications, *Survey Review*, 52:372, 261-277, DOI: 10.1080/00396265.2018.1548118.

[44] Vinicius Francisco Rofatto, Marcelo Tomio Matsuoka, Ivandro Klein, Maria Luísa Silva Bonimani, Bruno Póvoa Rodrigues, Caio Cesar de Campos, Mauricio Roberto Veronez & Luiz Gonzaga da Silveira Jr. (2022) An artificial neural network-based critical values for multiple hypothesis testing: data-snooping case, *Survey Review*, 54:386, 440-455, DOI: 10.1080/00396265.2021.1968176

[45] Lehmann R, Lösler M (2016) Multiple Outlier Detection: Hypothesis Tests Versus Model Selection by Information Criteria. *J Surv Eng*, 142(4), DOI 10.1061/(ASCE)SU.1943-5428.0000189.

[46] Zaminpardaz, S., Teunissen, P.J.G. DIA-datasnooping and identifiability. *J Geod* 93, 85–101 (2019). <https://doi.org/10.1007/s00190-018-1141-3>

AGRADECIMENTOS

- Gostaria de agradecer ao CNPq pela oportunidade de desenvolver esta pesquisa científica por meio do apoio da Bolsa de Pós-Doutorado Sênior. Em tempos mais difíceis essa bolsa foi um estímulo a continuar tentando fazer Ciência em alto nível depois de anos na carreira do magistério superior federal.

- Agradeço muito especialmente ao meu supervisor Prof. Dr. Paulo de Oliveira Camargo. Depois de mais de 15 anos foi uma satisfação imensa ter sido novamente meu supervisor!! Voltamos a trabalhar juntos e agora em um tema bem diferente da nossa parceria do passado. Novos frutos vieram!!

- Na Ciência moderna não se faz mais pesquisa em alto nível sozinho. Quero agradecer aos colegas que mais uma vez toparam colaborar em mais uma empreitada: Prof. Dr. Vinícius Francisco Rofatto e Prof. Dr. Ivandro Klein. Aos colegas Jhonatta Willyan Miato Assunção e Valdiney José da Silva meu muito obrigado pelo apoio no experimento de campo.

Obrigado!!

Marcelo Tomio Matsuoka