

LÍGIA DA SILVA BARBOSA

**AS DISTÂNCIAS DE MANHATTAN E CHEBYSHEV NA AVALIAÇÃO DA
ACURÁCIA POSICIONAL COM FEIÇÕES LINEARES EM PRODUTOS
CARTOGRÁFICOS**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Magister Scientiae*.

Orientador: Afonso de Paula dos Santos

Coorientadores: Júlio César de Oliveira
Nilcilene das Graças Medeiros

**VIÇOSA – MINAS GERAIS
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

B238d
2022
Barbosa, Lígia da Silva, 1996-
As distâncias de Manhattan e Chebyshev na avaliação da
acurácia posicional com feições lineares em produtos
cartográficos / Lígia da Silva Barbosa. – Viçosa, MG, 2022.
1 dissertação eletrônica (120 f.): il. (algumas color.).

Inclui apêndices.

Orientador: Afonso de Paula dos Santos.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Engenharia Civil, 2022.

Referências bibliográficas: f. 103-109.

DOI: <https://doi.org/10.47328/ufvbbt.2022.254>

Modo de acesso: World Wide Web.

1. Cartografia - Processamento de dados. 2. Dados
geoespaciais - Controle de qualidade. 3. Dados geoespaciais -
Normas. 4. Distâncias - Medição. I. Santos, Afonso de Paula dos,
1984-. II. Universidade Federal de Viçosa. Departamento de
Engenharia Civil. Programa de Pós-Graduação em Engenharia
Civil. III. Título.

CDD 22. ed. 526

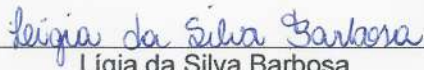
LÍGIA DA SILVA BARBOSA

**AS DISTÂNCIAS DE MANHATTAN E CHEBYSHEV NA AVALIAÇÃO DA
ACURÁCIA POSICIONAL COM FEIÇÕES LINEARES EM PRODUTOS
CARTOGRÁFICOS**

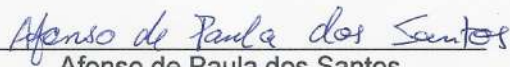
Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Magister Scientiae*.

APROVADA: 14 de fevereiro de 2022.

Assentimento:



Lígia da Silva Barbosa
Autora



Afonso de Paula dos Santos
Orientador

A Deus,
aos meus pais Neusa e José,
aos meus familiares,
e aos meus amigos
dedico.

AGRADECIMENTOS

A Deus em primeiro lugar por me possibilitar trilhar todo meu caminho até aqui apesar das dificuldades.

A minha família, que são meus pilares, que sempre acreditaram em mim mesmo quando eu mais duvidei e fizeram de tudo para que eu conseguisse ser o melhor de mim e conquistasse meus objetivos.

Ao grupo de pesquisa GeoTEC, em especial ao professor Gerson que me deu a oportunidade de crescer e aprender muito profissionalmente durante esse período.

Ao professor Afonso de Paula dos Santos por todo apoio e incentivo desde a graduação, sendo este o responsável por fazer me interessar pela pesquisa, além de toda dedicação nessa orientação, o que tornou essa jornada mais leve.

Aos professores Júlio e Nilcilene pelas sugestões que foram valiosas nessa pesquisa.

A Universidade Federal de Viçosa, por ter se tornado minha segunda casa, que com toda sua estrutura me possibilitou chegar aonde jamais imaginei, conseguindo superar os desafios.

A todos em geral que acreditaram e torceram por mim.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de estudos.

“Se o dinheiro for a sua esperança de independência, você jamais a terá. A única segurança verdadeira consiste numa reserva de sabedoria, de experiência e de competência.”

Henry Ford

RESUMO

BARBOSA, Lígia da Silva, M.Sc., Universidade Federal de Viçosa, fevereiro de 2022. **As distâncias de Manhattan e Chebyshev na avaliação da acurácia posicional com feições lineares em produtos cartográficos.** Orientador: Afonso de Paula dos Santos. Coorientadores: Nilcilene das Graças Medeiros e Júlio César de Oliveira.

Diversos países têm buscado atualizar e/ou editar suas normas e padrões referentes à qualidade dos dados espaciais, incluindo o Brasil através da ET-CQDG (Especificação Técnica para Controle de Qualidade de Dados Geoespaciais) elaborada em 2016 pela INDE (Infraestrutura de Dados Espaciais), visto que uma grande quantidade de dados é gerada a todo instante. Esta especificação utiliza como base o Decreto nº 89.817 de 1.984, a qual estabelece tolerâncias para avaliar e classificar produtos cartográficos em relação à sua qualidade posicional, utilizando feições pontuais. No entanto, o uso de feições lineares tem aumentado, especialmente por estarem presentes em maior quantidade em uma base cartográfica, além de possuírem uma boa distribuição espacial. Assim sendo, este trabalho propôs o uso das métricas Distância de Manhattan e Distância de *Chebyshev* para a avaliação da acurácia posicional de dados com feições lineares, através da adaptação do método Influência do Vértice e Distância de *Hausdorff*. Sendo assim, foram utilizados um conjunto de dados simulados com valores de discrepância média conhecidas para análise da eficiência dos métodos,. Além destes dados utilizou-se o conjunto de dados denominado “*MatchingLand*” desenvolvido por Xavier et al. (2017) que simulam dados reais, bem como dados reais na região da Universidade Federal de Viçosa em Viçosa – MG. Em relação aos dois últimos conjuntos citados, para fins de comparação dos resultados, utilizou-se o método *Buffer Duplo*. Com base nos resultados obtidos, verificou-se que em geral os métodos propostos que utilizam a distância de Manhattan possuem uma maior magnitude na determinação da discrepância posicional, enquanto os métodos que utilizam a distância de *Chebyshev* possuem uma menor magnitude. Pôde-se perceber que, aplicando as análises estatísticas para comparação das amostras de discrepâncias dos diferentes métodos aplicados, os métodos são similares estatisticamente independente da métrica utilizada, exceto o

método *Buffer Duplo*. Sendo assim, não é viável o uso das Distâncias de *Manhattan* e *Chebyshev*, já que produz o mesmo efeito prático do uso da distância Euclidiana e esta é bastante difundida e tradicional na Cartografia.

Palavras-chave: Controle de Qualidade Cartográfica. Distância de *Manhattan*. Distância de *Chebyshev*. Feições lineares. Acurácia Posicional. Decreto nº 89.817. ET-CQDG.

ABSTRACT

BARBOSA, Lígia da Silva, M.Sc., Universidade Federal de Viçosa, February, 2022.
Manhattan and Chebyshev distances in the assessment of positional accuracy with linear features in cartographic products. Adviser: Afonso de Paula dos Santos.
Co-advisers: Nilcilene das Graças Medeiros and Júlio César de Oliveira.

Several countries have sought to update and/or edit their norms and standards regarding the quality of spatial data, including Brazil through the ET-CQDG (Technical Specification for Quality Control of Geospatial Data) prepared in 2016 by INDE (Spatial Data Infrastructure), since a large amount of data is generated at all times. This specification is based on Decree No. 89,817 of 1984, which establishes tolerances for evaluating and classifying cartographic products in relation to their positional quality, using punctual features. However, the use of linear features has increased, especially because they are present in greater quantities in a cartographic base, in addition to having a good spatial distribution. Therefore, this work proposed the use of the Manhattan Distance and Chebyshev Distance metrics to evaluate the positional accuracy of data with linear features, through the adaptation of the Vertex Influence and Hausdorff Distance method. Therefore, a set of simulated data with known mean discrepancy values were used to analyze the efficiency of the methods. In addition to these data, the dataset called “MatchingLand” developed by Xavier et al. (2017) that simulate real data as well as real data in the region of Universidade Federal de Viçosa in Viçosa – MG. In relation to the last two sets mentioned, for the purpose of comparing the results, the Double Buffer method was used. Based on the results obtained, it was verified that, in general, the proposed methods that use the Manhattan distance have a greater magnitude in the determination of positional discrepancy, while the methods that use the Chebyshev distance have a smaller magnitude. It could be seen that, applying the statistical analysis to compare the discrepancy samples of the different methods applied, the methods are statistically similar regardless of the metric used, except for the Double Buffer method. Therefore, the use of Manhattan and Chebyshev Distances is not viable, since it produces the same practical effect as the use of Euclidean distance and this is quite widespread and traditional in Cartography.

Keywords: Cartographic Quality Control. Distance from Manhattan. Distance from Chebyshev. Linear features. Positional Accuracy. Decree No. 89,817. ET-CQDG.

LISTA DE ILUSTRAÇÕES

Figura 1 - Método Influência do Vértice considerando a linha de teste e a linha de referência.	28
Figura 2 - Distância de <i>Hausdorff</i> entre a linha de teste e a linha de referência.	29
Figura 3 - Método <i>Buffer</i> Duplo.	31
Figura 4 - Representação da distância Euclidiana entre dois pontos.....	33
Figura 5 - Distância de <i>Manhattan</i> tanto em preto, quanto em verde entre os pontos A e B. Em azul a Distância Euclidiana entre os pontos A e B.	35
Figura 6 - Possibilidades de caminho entre A e B.....	35
Figura 7 - Diferentes caminhos entre os pontos A e B.	36
Figura 8 - Distância de <i>Chebyshev</i> em verde e azul entre os pontos A e B. Em vermelho a Distância Euclidiana entre os pontos A e B.	39
Figura 9 - Exemplificação de diferentes tipos de distâncias: Dc - Distância de <i>Chebyshev</i> , Dm - Distância de <i>Manhattan</i> e De - Distância Euclidiana.	40
Figura 10 - Amostra da região de <i>MatchingLand</i>	43
Figura 11 - Amostra de morfologia das áreas.	44
Figura 12 - Amostra das classes de morfologia das linhas.	44
Figura 13 - Fluxograma da metodologia utilizada.....	47
Figura 14 – Ilustração das menores distâncias obtidas com as métricas Euclidiana, <i>Manhattan</i> e <i>Chebyshev</i> considerando o vértice superior da linha de referência.	51
Figura 15 - Ilustração das menores distâncias obtidas com as métricas Euclidiana, <i>Manhattan</i> e <i>Chebyshev</i> considerando o vértice central da linha de referência.....	51
Figura 16 - Feições lineares simuladas nº 1. Linha de referência em vermelho e linha de teste em azul.	60
Figura 17 - Feições lineares simuladas nº 2. Linha de referência em vermelho e linha de teste em azul.	61

Figura 18 - Feições lineares simuladas nº 3. Linha de referência em vermelho e linha de teste em azul.	61
Figura 19 - Feições lineares simuladas nº 4. Linha de referência em vermelho e linha de teste em azul.	62
Figura 20 - Feições lineares simuladas nº 5. Linha de referência em vermelho e linha de teste em azul.	63
Figura 21 - Feições lineares simuladas nº 6. Linha de referência em vermelho e linha de teste em azul.	64
Figura 22 - Feições lineares simuladas nº 7. Linha de referência em vermelho e linha de teste em azul.	65
Figura 23 - Feições lineares simuladas nº 8. Linha de referência em vermelho e linha de teste em azul.	65
Figura 24 - Feições lineares simuladas nº 9. Linha de referência em vermelho e linha de teste em azul.	66
Figura 25 - Feições lineares simuladas nº 10. Linha de referência em vermelho e linha de teste em azul.	67
Figura 26 - Feições lineares simuladas nº 11. Linha de referência em vermelho e linha de teste em azul.	67
Figura 27 - Feições lineares simuladas nº 12. Linha de referência em vermelho e linha de teste em azul.	68
Figura 28 – Feições lineares simuladas nº 13. Linha de referência em vermelho e linha de teste em azul.	69
Figura 29 - Feições lineares simuladas nº 14. Linha de referência em vermelho e linha de teste em azul.	69
Figura 30 - Feições lineares simuladas nº 15. Linha de referência em vermelho e linha de teste em azul.	70
Figura 31 - Conjunto de linhas utilizadas extraídas da base da Espanha.	70
Figura 32 - Pares de linhas muito suaves (em vermelho a linha de referência e em azul a linha de teste).	71
Figura 33 - Pares de linhas suaves.	72

Figura 34 - Pares de linhas suaves com direcionalidade estável.	72
Figura 35 - Pares de linhas suaves com direcionalidade variável.	73
Figura 36 - Pares de linhas sinuosas.	73
Figura 37 - Área de estudo da pesquisa.	74
Figura 38 - Gráfico das discrepâncias obtidas com o Influência do Vértice e os métodos sugeridos.	79
Figura 39 - Gráfico das discrepâncias utilizando o Influência do Vértice, os métodos sugeridos, Distância de <i>Hausdorff</i> , as adaptações propostas e o <i>Buffer Duplo</i> para os dados da Espanha.	85
Figura 40 - Gráfico das discrepâncias utilizando diversos métodos para os dados reais.	94
Figura 41 - Densificação e quantidade de vértices para o cálculo das distâncias...	110
Figura 42 - Adaptação do Influência do Vértice com a Distância de <i>Manhattan</i> (VI – M).	114
Figura 43 - Adaptação do Influência do Vértice com a Distância de <i>Chebyshev</i> (VI – C).	115
Figura 44 - Adaptação do influência do Vértice com a distância de <i>Manhattan</i> e <i>Chebyshev</i> (VI – EMC).	115
Figura 45 - Novo método HD - M2, calculando apenas as distâncias da linha de referência para a linha teste.	115
Figura 46 - Novo método HD – C2, calculando apenas as distâncias da linha de referência para a linha teste.	115
Figura 47 - Novo método HD - M, calculando apenas as distâncias da linha de referência para a linha teste e vice-versa.	115
Figura 48 - Novo método HD - C, calculando apenas as distâncias da linha de referência para a linha teste e vice-versa.	115

LISTA DE TABELAS

Tabela 1- Padrão de Exatidão Cartográfica Planimétrica (Decreto nº 89.817 aliada à ET-CQDG).....	25
Tabela 2 - Informações das linhas de teste e das linhas de referência dos dados simulados.	57
Tabela 3 - Informações dos pares de linhas da base <i>MatchingLand</i>	71
Tabela 4 - Características geométricas do conjunto de dados reais.	75
Tabela 5 - Discrepâncias obtidas com os dados simulados gerados.	78
Tabela 6 - Diferença das discrepâncias médias em porcentagem tendo como referência as discrepâncias médias conhecidas. Em vermelho estão as maiores diferenças de cada par de linha simulada e em azul os menores valores.	80
Tabela 7 - Discrepâncias obtidas com os dados simulados da Espanha.	84
Tabela 8 - Diferença das discrepâncias em porcentagem tendo como referência o Influência do Vértice para os dados da Espanha em valor absoluto.	87
Tabela 9 - Estatísticas Descritivas da amostra de dados da Espanha e a classificação.	90
Tabela 10 - Discrepâncias do conjunto de dados em cada método para os dados reais.	93
Tabela 11 - Diferença em porcentagem das discrepâncias obtidas para os dados reais.	96
Tabela 12 - Estatísticas descritivas da amostra de dados reais e a classificação.....	98

SUMÁRIO

1. INTRODUÇÃO	16
2. OBJETIVO.....	21
3. JUSTIFICATIVA	22
4. FUNDAMENTAÇÃO TEÓRICA.....	23
4.1. Controle de Qualidade Cartográfica	23
4.1.1. Acurácia Posicional	24
4.1.1.1. Influência do Vértice	27
4.1.1.2. Distância de Hausdorff	29
4.1.1.3. Buffer Duplo.....	30
4.2. Métricas.....	32
4.2.1. Distância Euclidiana	32
4.2.2. Distância de Manhattan	34
4.2.3. Distância de Chebyshev	38
4.3. Base MatchingLand.....	42
5. METODOLOGIA.....	47
5.1. Métodos propostos	49
5.1.1. Método “VI - M” (Influência do Vértice com a Distância de Manhattan) e Método “VI - C” (Influência do Vértice com a Distância de Chebyshev)	49
5.1.2. Método “VI - EMC” (Influência do Vértice com Distância Euclidiana, Manhattan e Chebyshev)	52
5.1.3. Método “HD - M” (Distância Hausdorff com a métrica de Manhattan) e Método “HD - C” (Distância Hausdorff com a métrica de Chebyshev)	53
5.1.4. Método “HD - M2” (Distância Hausdorff com a métrica de Manhattan adaptado) e Método “HD - C2” (Distância Hausdorff com a métrica de Chebyshev adaptado)	55
5.2. Experimentos	56
5.2.1. Dados Simulados	56
5.2.2. Base MatchingLand.....	70

5.2.3. Dados Reais: Ortoimagem	73
5.2.3.1. Área de Estudo	73
5.2.4. Materiais Utilizados	75
6. RESULTADOS E DISCUSSÕES	77
6.1. Dados simulados gerados	77
6.2. Dados simulados Espanha	83
6.3. Dados Reais	92
7. CONCLUSÕES E RECOMENDAÇÕES	100
REFERÊNCIAS BIBLIOGRÁFICAS	103
APÊNDICE A	110
APÊNDICE B	111
APÊNDICE C	114

1. INTRODUÇÃO

Vivencia-se um período em que praticamente tudo está em constante modificação e transformação, assim como o espaço geográfico que tem se tornado cada vez mais dinâmico. Devido a isso, há a necessidade em manter todas as informações espaciais atualizadas, para que os dados obtidos sejam utilizados de forma correta em tomadas de decisões (MEDEIROS, 1999; BRAVO & SLUTER, 2015).

Conforme Lisboa (2019), nos últimos anos houve um aumento considerável no interesse em utilizar dados espaciais nas mais variadas áreas de estudo, e dessa forma um grande número de dados são gerados a todo instante. Santos et al. (2015) ressaltam que tal fato é explicado principalmente pelo desenvolvimento acelerado das tecnologias utilizadas no mapeamento, softwares e dados disponíveis a qualquer usuário e equipamentos para medição cada vez mais simples de se utilizar.

Diante de tal fato, Zanetti et al. (2018) diz que há a necessidade de analisar e verificar a qualidade dos produtos cartográficos, especificando suas limitações e para quais finalidades podem ser utilizados, sua acurácia, além de explicitar as discrepâncias, omissões e incertezas. Ademais, Lunardi et al. (2012) também ressaltam que é importante que haja a padronização dos produtos gerados, facilitando a interoperabilidade entre os mesmos, além de avaliar sua qualidade.

No Brasil, conforme é dito em Santos et al. (2015) e Menezes et al. (2019), o controle de qualidade posicional dos produtos cartográficos é regulamentado de acordo com as normas do Decreto nº 89.817 de 20 de junho de 1984. Entretanto, este decreto não define uma metodologia específica para realizar a avaliação.

Visando resolver esta problemática, bem como propor novas medidas para os diversos elementos do controle de qualidade cartográfica, a Diretoria de Serviço Geográfico do Exército Brasileiro (DSG) em 2016 elaborou a Especificação Técnica para o Controle de Qualidade em Dados Geoespaciais (ET-CQDG) que define os procedimentos de qualidade dos produtos da cartografia terrestre.

Já os padrões internacionais mais conhecidos são: a ISO 19.157 – Informação Geográfica - Qualidade dos dados, o padrão EMAS/ASPRS (*Engineering Map Accuracy/American Society of Photogrammetry and Remote Sensing*), a NSSDA (*National Standard for Spatial Data Accuracy*), a STANAG (Standardization Agreement) da OTAN (Organização do Tratado do Atlântico Norte) e o padrão do serviço geológico americano de acurácia NMAS (*National Map Accuracy Standards*) (ARIZA & ATKINSON, 2006; ARIZA & ATKINSON, 2008; MIGUEL CASTRO, 2014; ZANETTI et al., 2018).

Como pode ser visto, há uma variedade de padrões e critérios que podem apresentar classificações distintas para um mesmo produto, o que torna necessário um estudo aprofundado nessa área (ZANETTI et al., 2018). Vale ressaltar que a utilização de padrões tem se tornado essencial devido principalmente ao fato de que a produção de dados espaciais envolve métodos específicos, demandando profissionais cada vez mais qualificados (DALMOLIN; LEAL, 2001; FREITAS, 2005; NOGUEIRA Jr., 2003; SANTOS et al., 2015).

Segundo a ISO 19.157 (2013), na avaliação da qualidade de dados espaciais é necessário considerar elementos como Completude, Consistência Lógica, Acurácia Temática, Acurácia Temporal, Acurácia Posicional e a Usabilidade. Um dos elementos desse grupo mais estudado e avaliado é a acurácia posicional, a qual verifica a qualidade da componente espacial (MOZAS et al., 2013; SAITO et al., 2019). Além disso, na análise da acurácia posicional podem ser utilizadas feições como pontos e feições lineares (linhas, polilinhas, segmentos de retas, polígonos e áreas).

Santos et al. (2015) verifica que em diversos países, as normas e leis utilizam metodologias baseadas na comparação entre pontos homólogos, devido à simplicidade dessa comparação, sendo que nas últimas décadas, várias pesquisas têm analisado o uso de feições lineares na avaliação da acurácia posicional, devido principalmente à quantidade considerável de informação geométrica, como pode ser visto em Ariza et al. (2011), Mozas e Ariza (2011), Santos et al. (2016), Cunha et al. (2019) e Saito et al. (2019).

Ainda de acordo com Santos et al. (2015), a avaliação utilizando feições lineares fundamenta-se na comparação entre linhas homólogas, analisando

elementos como a razão entre comprimentos e quantidade de vértices, angularidade, razão de segmentos curvilíneos, distância entre os vértices, percentual de inclusão em uma área de influência, dentre outros.

Grande parte dos estudos relacionados ao uso de feições lineares no controle de qualidade cartográfica tem como exemplo a norma atual da Espanha, UNE-148002, visto que esta prevê a utilização deste tipo de primitiva nas avaliações. Mozas & Ariza (2015) ressaltam que alguns autores recomendam a utilização dessa primitiva no controle de qualidade, pois este tipo de feição possui um grande número de vértices e alta frequência em um banco de dados espaciais. Ademais, conforme Santos (2015), as linhas possuem diversas informações geométricas que podem auxiliar na análise da acurácia em dados espaciais.

Já existem métodos para a avaliação da acurácia posicional de dados espaciais em 2D empregando feições lineares, sendo que na década de 1980 foram propostos os métodos Banda de Erro Épsilon (*Epsilon Band Method*) (MASRY et al., 1980; BLAKEMORE, 1984), Ponto Gerado (*Points Intermediate Method*) (MASRY et al., 1980) e Método da Distância Média (*Average Distance Method*) (McMASTER, 1986). Na década de 1990 criaram-se os métodos Distância de Hausdorff (*Hausdorff's Distance Method*) (ABBAS et al., 1995), Completude (*Completeness*) (HEIPKE et al., 1997), Corretude (*Correctness*) (HEIPKE et al., 1997), Buffer Simples (*Simple Buffer Method*) (GOODCHILD e HUNTER, 1997), Buffer Duplo (*Double Buffer Method*) (TVEITE e LANGAAS, 1999) e Retângulo Envolvente (FERREIRA e CINTRA, 1999); Já na década de 2000 foram criados Banda-G (*G-band*) (SHI e LIU, 2000) e Aproximações por Spline (*Spline Approaches*) (GALO et al., 2001).

E por fim, na década atual de 2010, foram elaborados alguns métodos, como por exemplo a Influência do Vértice (*Vertex Influence Method*) (MOZAS e ARIZA, 2011), Modelo de Erro Estatístico de Simulação (*Statistical Simulation Error Model*) (TONG et al., 2013), Modelo Unificado (*Unified Model*) (CAI et al., 2014), Influência do Vértice por Vértices (*Vertex Influence Method by Vertexes*) que é uma modificação no método Influência do Vértice (MOZAS e ARIZA, 2018). Vale ressaltar que estes são apenas alguns exemplos dos métodos existentes.

Na década de 2010 pesquisadores propuseram a utilização de feições lineares no processo de avaliação em 3D, como pode ser visto em Mozas et al. (2013), Mozas e Ariza (2015), Mozas et al. (2017), Fonseca (2018) e Lisboa (2019). Sendo assim, é possível observar que a comunidade científica tem buscado ao longo de quatro décadas o desenvolvimento de metodologias para avaliação da acurácia posicional utilizando feições lineares.

A distância é uma das métricas possíveis de serem utilizadas no processo de avaliação da acurácia posicional de feições lineares, sendo tradicionalmente a mais utilizada a distância Euclidiana (CRUZ, 2020). Entre as distâncias empregadas cita-se a Distância de *Manhattan*, que é eficiente para aplicações em tempo real, e a Distância de *Chebyshev*, a qual considera o máximo valor da distância em uma dimensão (KUGLER, 2003; LEI et al., 2017).

A adaptação de métodos existentes e criação de novos métodos utilizando estas métricas é uma alternativa para os casos de necessidade de avaliação conjunta utilizando feições lineares e pontuais, considerando, por exemplo, locais que possuem grande parte da área de estudo com matas fechadas, em que não é possível obter pontos nesses locais, sendo necessário o uso de feições lineares para que se tenha uma boa distribuição dos dados na área de estudo. Ainda vale ressaltar que no Brasil há uma pequena quantidade de estudos voltados ao uso de feições lineares na avaliação da acurácia posicional, o que torna ainda mais importante pesquisas relacionadas a esse assunto.

Dessa forma torna-se relevante cada vez mais trabalhos nessa área de avaliação da acurácia posicional com feições lineares, buscando-se outras métricas que possam ser aplicadas ao controle de qualidade posicional e sejam implementadas em *softwares* de SIG (Sistema de Informação Geográfica), já que as linhas permitem investigar aspectos que não são possíveis através da avaliação com feições pontuais.

Sendo assim, o uso das métricas de distâncias de *Manhattan* e distância de *Chebyshev* podem ser alternativas para obtenção de discrepâncias posicionais entre feições lineares no processo de controle de qualidade cartográfica.

É importante que os novos métodos a serem desenvolvidos ou adaptações de métodos já criados utilizando tais métricas sejam eficientes e de simples aplicação. Com isso, da mesma forma que é acessível a produção dos dados, o controle de qualidade também será acessível e popularizado.

2. OBJETIVO

O objetivo geral deste trabalho é compreender o processo de avaliação da acurácia posicional de produtos cartográficos, baseados em feições lineares.

Especificamente buscou-se:

- a) Analisar (comparar) a adaptação do método Influência do Vértice e do método Distância de Hausdorff utilizando as métricas Distância de *Manhattan* e Distância de *Chebyshev* no processo de avaliação da acurácia posicional com feições lineares;
- b) Validar a utilização destes métodos junto ao padrão de acurácia brasileiro, definido pelo Decreto nº 89.817 / ET-CQDG.

3. JUSTIFICATIVA

O processo de controle de qualidade cartográfica no Brasil tem como base o Decreto nº 89.817 (1984) aliado à Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG, 2016), elaborada pela Diretoria de Serviço Geográfico do Exército Brasileiro. Um dos elementos utilizados para representar a qualidade cartográfica de um produto é a acurácia posicional, que na maioria das vezes, utiliza a relação entre pontos homólogos para classificar o produto.

A justificativa deste trabalho está relacionada à necessidade de estudo e viabilidade de adaptações de métodos existentes, que consigam mensurar as discrepâncias posicionais entre feições lineares, bem como sua implementação em softwares livres e/ou proprietários.

Além disso, de acordo com Saito et al. (2019) e Santos et al. (2015), no processo de controle de qualidade em dados espaciais, considerando o elemento acurácia posicional, a utilização de feições lineares ao invés de feições pontuais é priorizada visto que:

- A feição mais utilizada na representação cartográfica é a linha, sendo que em uma base cartográfica cerca de 80% das feições representadas são lineares;
- Normalmente as feições lineares possuem boa distribuição espacial nas bases cartográficas;
- As linhas contêm uma quantidade considerável de informação geométrica com vértices, segmentos, orientação, dentre outros.

Dentre as métricas possíveis, a presente pesquisa utiliza a Distância de *Manhattan* e a Distância de *Chebyshev*, por apresentarem como diferencial a simplicidade e facilidade de implementação computacional, sendo importante verificar a viabilidade do uso destas no controle de qualidade cartográfica.

4. FUNDAMENTAÇÃO TEÓRICA

Neste tópico será abordado uma revisão sobre o controle de qualidade cartográfica, com ênfase na acurácia posicional, detalhando os métodos utilizados neste trabalho como: Influência do Vértice, Distância de *Hausdorff* e *Buffer* Duplo. Além disso, serão descritas as métricas que são o objeto dessa pesquisa, distância de *Manhattan* e *Chebyshev*, bem como a descrição de uma das bases de dados utilizada, a *MatchingLand*.

4.1. Controle de Qualidade Cartográfica

O conceito de qualidade é subjetivo, sendo que pode ser definido em geral como as características do produto que atendem necessidades implícitas ou explícitas dos usuários, proporcionando sua satisfação (JURAN, 1991; ARIZA, 2002).

Considerando a área da cartografia, é imprescindível que haja o controle de qualidade adequado dos produtos obtidos, já que esta é responsável por produzir modelos da realidade, e estes são utilizados para tomar decisões. Dessa forma, quanto maior o nível de acerto na tomada de decisões tendo como referência os dados espaciais, maior deve ser o nível de controle de qualidade destes dados (SANTOS et al., 2015). Ressalta-se que mesmo possuindo o conhecimento sobre as questões relativas ao processo de obtenção dos dados cartográficos, tal fato não é foco dessa pesquisa.

Vale ressaltar que o controle de qualidade em produtos cartográficos é um tema relevante, já que no Brasil há uma legislação específica que trata desse assunto, o Decreto nº 89.817 de 1984, o qual aborda especificamente sobre a acurácia posicional. Entretanto, de acordo com Silva et al. (2020), nem todos os dados produzidos possuem um indicador da qualidade, sendo necessário avaliar não apenas aspectos como a acurácia posicional, que é o foco deste trabalho, mas também a acurácia temática, a consistência lógica, a completude, fidelidade semântica, além de aspectos temporais e de linhagem, tornando a análise complexa.

Devido ao surgimento de novas tecnologias para a coleta e aquisição de dados espaciais, tornou-se importante a integração adequada entre os mesmos. Dessa

forma, para avaliar determinado produto em relação à sua qualidade, definindo sua finalidade, além de explicitar suas discrepâncias e incertezas, é fundamental utilizar uma determinada norma ou padrão (ZANETTI et al., 2018).

Diante disso, foi elaborada a ET-CQDG, sendo esta aliada à norma ISO 19.157 (Organização Internacional de Padronização) que estabelece os elementos utilizados para avaliar um conjunto de dados, sendo divididos em seis categorias:

- Acurácia posicional: avalia a proximidade da posição de um dado espacial em relação a sua realidade no terreno.
- Acurácia temática: é a correta interpretação das feições e atributos de um dado espacial em relação à realidade.
- Acurácia temporal: refere-se à exatidão encontrada na componente temporal dos dados geoespaciais.
- Completude: indica presença ou ausência de feições na informação geográfica.
- Consistência lógica: é o cumprimento das regras lógicas da estrutura dos dados, dos atributos e das relações.
- Usabilidade: está ligada aos requisitos do usuário e/ou o universo de discussão.

Dentre esses elementos, um dos mais importantes e que tem mais se destacado, é a acurácia posicional.

4.1.1. Acurácia Posicional

Conforme a ISO 19.157, a acurácia posicional pode ser dividida em três subelementos, sendo eles:

- Acurácia absoluta ou externa: é o quão próximo estão os valores das coordenadas do dado espacial em análise com as coordenadas tidas como referência no terreno;
- Acurácia relativa ou interna: está relacionada com a proximidade entre as posições relativas de uma feição e suas respectivas posições relativas tidas como referência em um produto de maior acurácia;

- Acurácia posicional em dados matriciais: refere-se à proximidade entre as posições dos dados de uma estrutura quadricular regular e suas respectivas posições tidas como referência.

No Brasil, a avaliação da qualidade posicional dos dados espaciais utiliza-se das tolerâncias Padrão de Exatidão Cartográfica (PEC) e Erro-Padrão (EP), contidas no Decreto nº 89.817 de 1984 (BRASIL, 1984). Em 2010, a Infraestrutura Nacional de Dados espaciais (INDE) elaborou a Especificação Técnica de Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV), que além de regular e padronizar a aquisição da geometria dos dados geoespaciais vetoriais e atributos correlacionados, complementa o Decreto.

A ET-ADGV criou uma classe mais restritiva para produtos cartográficos digitais, intitulada de Padrão de Exatidão Cartográfica Para Produtos Cartográficos Digitais (PEC-PCD) (Tabela 1). Já em 2016, a DSG lançou a ET-CQDG, para o auxílio da determinação da acurácia posicional pelo Decreto nº 89.817 aliada à ET-ADGV.

Tabela 1- Padrão de Exatidão Cartográfica Planimétrica (Decreto nº 89.817 aliada à ET-CQDG).

Classe PEC	Classe PEC-PCD	Planimetria	
		PEC	EP
-	A	0,28 mm x escala	0,17 mm x escala
A	B	0,50 mm x escala	0,30 mm x escala
B	C	0,80 mm x escala	0,50 mm x escala
C	D	1,00 mm x escala	0,60 mm x escala

Fonte: Adaptado da ET-CQDG (2016).

Dessa forma, para que um produto cartográfico obtenha a classificação conforme o padrão de acurácia posicional do Decreto nº 89.817/ET-CQDG, é fundamental que satisfaça a duas condições (DSG, 2016; SANTOS et al. 2015; BRASIL, 1984):

i. Noventa por cento (90%) dos pontos testados devem apresentar valores de discrepâncias (Equação 1) iguais ou inferiores ao valor do PEC-PCD em relação à escala e a classe testada, e;

ii. O RMS (*Root Mean square*) (Equação 2) da amostra de discrepâncias posicionais deve ser igual ou inferior à tolerância EP definido pela norma, para a escala e classe testada.

$$d_i = \sqrt{(X_{test} - X_{ref})^2 + (Y_{test} - Y_{ref})^2} \quad (1)$$

$$RMS = \sqrt{\frac{\sum_{i=1}^n (d_i)^2}{n}} \quad (2)$$

Sendo:

d_i : discrepância posicional da feição i ;

n : número de feições de checagem;

X_{test}, Y_{test} : Coordenadas de teste; e

X_{ref}, Y_{ref} : Coordenadas de referência.

De acordo com Santos et al. (2015) e Zanetti et al. (2018), grande parte das normas de controle de qualidade posicional fazem o uso de metodologias em que os pontos são bem definidos e distribuídos espacialmente. Contudo, segundo esses autores, pode-se utilizar não somente feições pontuais, mas também feições lineares no processo de avaliação da acurácia posicional. Além disso, em áreas em que os pontos não estejam bem distribuídos por toda área de estudo, mas que se tenha uma certa quantidade desses pontos em boa parte da área de estudo, pode-se fazer o uso da feição linear em conjunto.

Tong et al. (2013) afirma que a utilização de métodos de avaliação da acurácia posicional baseada em feições lineares tem sido bastante pesquisada visando o uso em Sistemas de Informações Geográficas (SIG). Isso acarretou o esforço da comunidade científica durante as últimas décadas objetivando a compreensão desse assunto e o desenvolvimento de diversos métodos de avaliação.

De acordo com Santos et al. (2015), os métodos que utilizam feições lineares podem ser divididos em três grupos principais:

- Grupo de faixa de incerteza que fazem o uso de uma abordagem determinística, tais como o método da Banda *Épsilon* ou Método das Áreas, *Buffer Simple* e *Buffer Duplo*;

- Grupo de faixa de incerteza com uma abordagem estocástica, como a Banda Genérica e Modelo de Erros Estatístico por Simulação;
- Grupo de métodos determinísticos que fazem análises geométricas, como a Distância de *Hausdorff*, Influência do Vértice, Aproximação por *Spline* e Ponto Gerado.

A seguir serão detalhados os métodos da Influência do Vértice e da Distância de Hausdorff que foram escolhidos para fins de adaptação utilizando as métricas desta pesquisa (Distância de *Manhattan* e Distância de *Chebyshev*). Também é apresentado o método do Buffer Duplo que foi utilizado para fins de comparação com os resultados obtidos pelos métodos propostos neste trabalho.

4.1.1.1. Influência do Vértice

Conforme Mozas e Ariza (2011), neste método, cada vértice de uma linha determina uma posição envolvendo os segmentos adjacentes, sendo assim qualquer deslocamento em um determinado vértice afeta apenas os dois segmentos adjacentes a ele.

Ainda de acordo com esses autores, o método Influência do Vértice baseia-se na medida de distância dos vértices da linha de referência (L_R) a linha de teste (L_T). Tais distâncias são ponderadas em relação ao comprimento dos segmentos adjacentes ao vértice analisado ($l_i^{(i-1)}$ e $l_i^{(i+1)}$). No final do processo, é calculada a média aritmética ponderada utilizando o comprimento da linha de referência (CL_R), obtendo-se a discrepância média (Di) (Equação 3) entre as linhas, de acordo com a Figura 1.

$$Di = \frac{\left[\sum_{k=1}^m (d_{h_i} (l_i^{(i-1)} + l_i^{(i+1)})) \right]}{2 \cdot CL_{R_i}} \quad (3)$$

sendo:

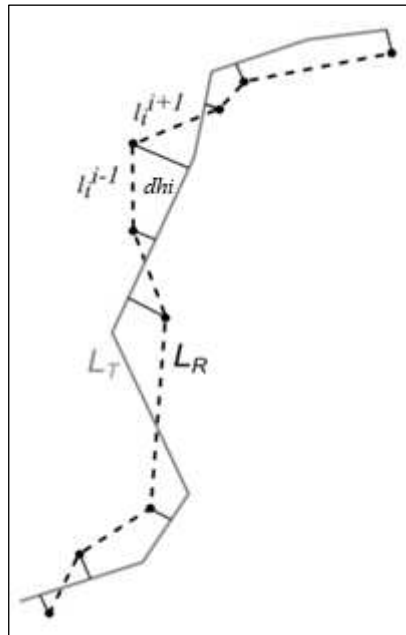
m : número de vértices da linha de referência;

dh_i : distância horizontal mínima entre o vértice k de L_R para L_T ;

$l_i^{(i-1)}$ e $l_i^{(i+1)}$: comprimento dos segmentos adjacentes ao vértice k em L_R ;

CLR: comprimento da feição linear de referência.

Figura 1 - Método Influência do Vértice considerando a linha de teste e a linha de referência.



Fonte: Adaptado de Mozas e Ariza (2011).

De acordo com Santos et al. (2015) a utilização do método Influência do Vértice no processo de avaliação da acurácia posicional baseada em feições lineares conforme o Decreto nº 89.817/ET-CQDG, consiste em:

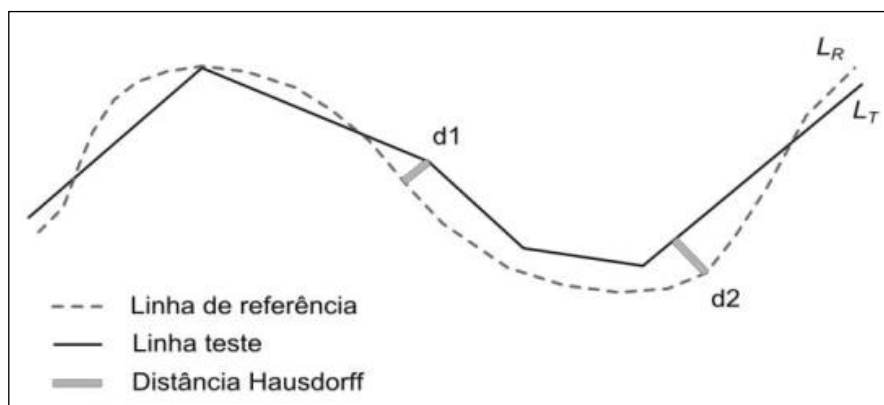
1. Selecionar uma amostra de n linhas homólogas de teste (L_T) e de referência (L_R);
2. Calcular o comprimento (CL_R) e extrair os vértices de cada linha i de referência;
3. Calcular a menor distância (dh_k) para cada vértice k da linha referência em relação a qualquer segmento da linha teste;
4. Obter os valores dos comprimentos dos segmentos de retas anteriores ($l_k^{(k-1)}$) e posteriores ($l_k^{(k+1)}$) aos vértices k da linha i de referência;
5. Calcular a discrepância média (D) para cada linha i de referência (Equação 4);

6. Para o dado espacial ser classificado de acordo com o Decreto nº 89.817, , 90% das linhas testadas devem apresentar valores de D menores ou iguais ao valor do PEC e o RMS da amostra de discrepâncias médias D tem que ser menor ou igual ao valor do EP, para a classe e escala empregadas no processo de avaliação da acurácia posicional.

4.1.1.2. Distância de Hausdorff

De acordo com Abbas (1995), este método consiste em determinar a Distância de *Hausdorff* sobre pares de linhas homólogas. Esta métrica consiste em obter o máximo valor das menores distâncias euclidianas entre duas linhas (Figura 2).

Figura 2 - Distância de *Hausdorff* entre a linha de teste e a linha de referência.



Fonte: Santos et al. (2015).

Conforme Ariza (2002) para a aplicação deste método é necessário obter o maior valor das menores distâncias de cada vértice da linha teste em relação à linha de referência ($d1$), e o maior valor das menores distâncias de cada vértice da linha de referência em relação à linha teste ($d2$). No final, o maior valor entre $d1$ e $d2$ será considerada a Distância de *Hausdorff* (dh) (Equação 4).

$$dh = \max [d1, d2] \quad (4)$$

A Distância de *Hausdorff* já teve diversas modificações, visto que esta é sensível a erros grosseiros ou ruídos. De acordo com Ariza (2002), os valores de distâncias obtidos são impróprios para adotá-los como discrepâncias entre as linhas

devido principalmente às diferenças de tamanhos, forma, complexidade, entre outros fatores.

Devido a isso Mozas (2017) sugeriu utilizar como valor de $d1$ e $d2$ a média das menores distâncias entre os vértices da linha teste com relação à linha de referência, e vice-versa. Posteriormente, o maior valor entre $d1$ e $d2$ será utilizado como a Distância de *Hausdorff*.

Dessa forma, segundo Santos et al. (2015) a utilização do método Distância de *Hausdorff* na avaliação da acurácia posicional baseada em feições lineares conforme o Decreto nº 89.817/ET-CQDG, consiste em:

1. Selecionar uma amostra de n linhas homólogas de teste (L_T) e de referência (L_R);
2. Extrair os vértices das linhas teste e da linha de referência;
3. Calcular a menor distância de cada vértice da linha teste em relação à linha de referência. Obter a média de todas as menores distâncias calculadas, atribuindo o valor a $d1$;
4. Calcular a menor distância de cada vértice da linha de referência em relação à linha teste. Obter a média de todas as menores distâncias calculadas, atribuindo o valor a $d2$;
5. A Distância de *Hausdorff* (dh) é definida pelo maior valor entre $d1$ e $d2$;
6. Para a análise deste método junto ao Decreto-lei nº 89.817, utiliza-se o valor de dh como medida de discrepância. Portanto, para ser classificado deve-se atender duas condições: 90% das linhas testadas devem apresentar valores de dh menores ou iguais ao valor do PEC e; o RMS da amostra de dh deve ser menor ou igual ao valor do EP para a classe e escala empregadas.

4.1.1.3. **Buffer Duplo**

O método do *Buffer Duplo* foi proposto por Tveite e Langaas (1999) com o nome de *Buffer Overlay Statistics*. Este método foi desenvolvido tendo como referência o método do *Buffer Simples*, o qual foi elaborado por Goodchild e Hunter (1997).

O princípio deste método consiste em utilizar as medidas de *Buffer* visando obter a discrepância entre a linha de teste e a linha de referência. É aplicado o *Buffer*

em cada uma das linhas, sendo formados polígonos e realizadas operações de interseção entre eles (SANTOS et al., 2015).

Tveite e Langaas (1999) afirmam que as vantagens de se utilizar o *Buffer Duplo* é que ao aplicar um *Buffer* tanto na linha testada quanto na linha de referência, obtêm-se uma ponderação das incertezas, sendo possível obter uma discrepância média (dm_i) (Equação 5).

$$dm_i = \pi \cdot \chi \cdot \frac{(\sum A_1 - A_3)}{A_2} \quad (5)$$

onde:

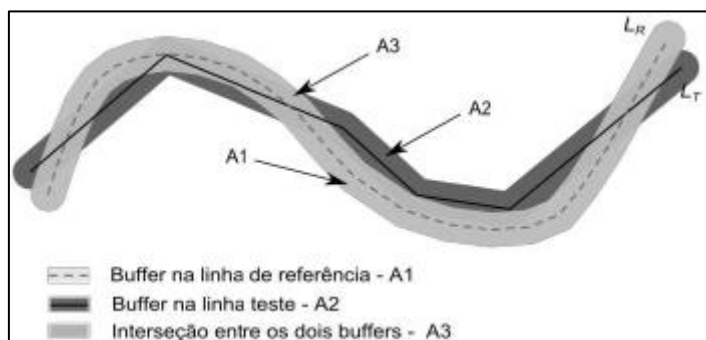
A_1 é a área do *Buffer* da linha de referência que não faz interseção com o *buffer* da linha de teste;

A_2 é a área formada pelo *Buffer* da linha de teste;

A_3 é a área do *Buffer* referente à interseção do *buffer* da linha de referência com o *buffer* da linha de teste.

A Figura 3 representa a técnica do *Buffer Duplo*.

Figura 3 - Método *Buffer Duplo*.



Fonte: Santos et al. (2015).

Segundo Santos et al. (2015) para utilização do *Buffer Duplo* na avaliação da acurácia posicional em feições lineares utilizando o padrão brasileiro, deve-se proceder os seguintes passos:

1. Selecionar uma amostra de n linhas homólogas de teste (LT) e de referência (LR);

2. Aplicar um *buffer* de tamanho x em cada linha i da amostra de linhas homólogas de teste e referência. Para o emprego do Decreto nº 89.817, utiliza-se como tamanho do *buffer* o valor do PEC para a escala e a classe utilizada no processo de avaliação;
3. Calcular a área A_2 gerada pela operação de *buffer* em cada linha de teste;
4. Fazer uma sobreposição entre os polígonos gerados pelos *buffers* da linha teste e da linha de referência e calcular o valor de A_1 , referente a área do polígono de referência que não faz interseção ao polígono teste;
5. Calcular a discrepância média (dm), para cada linha teste i (equação 3);
6. Para análise do Decreto nº 89.817, faz a verificação se 90% das linhas testadas apresentam discrepância média (dm) menor ou igual ao valor do PEC (da classe e escala utilizados para gerar o *buffer* x), e ainda, o RMS das discrepâncias dm ser menor ou igual ao valor do EP (idem ao PEC). Se as duas condições forem aceitas, classifica-se o dado espacial de acordo com a classe e escala trabalhada. Caso não obtenha classificação na classe trabalhada, deve-se testar as tolerâncias da classe posterior, utilizando um novo valor para o *buffer* de acordo com a classe e assim sucessivamente.

4.2. Métricas

4.2.1. Distância Euclidiana

A distância Euclidiana que é a medida de similaridade mais conhecida, e também a mais utilizada na prática, foi definida pelo matemático Euclides e representa a menor distância entre dois objetos no plano multidimensional. Segundo Prado (2015), esta métrica considera a raiz quadrada do somatório dos quadrados das diferenças entre os valores de atributos para cada par de variáveis.

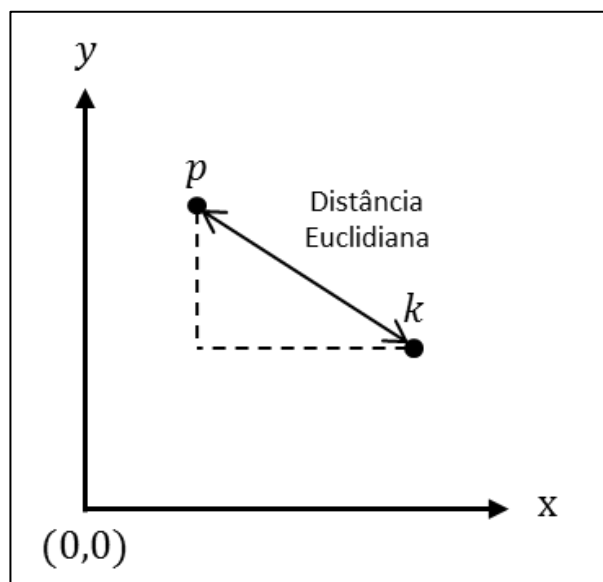
Através da Equação 6, pode ser vista como esta distância é calculada.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i^2 - y_i^2)} \quad (6)$$

sendo x_i e y_i são, respectivamente, os valores do i -ésimo atributo para os dados x e y .

De acordo com Machado (2011), a distância Euclidiana é mais utilizada para aplicações do algoritmo de agrupamento k-médias. Na Figura 4 é ilustrada a distância Euclidiana, a qual sempre será a menor distância em linha reta entre dois objetos.

Figura 4 - Representação da distância Euclidiana entre dois pontos.



Fonte: Adaptado de Pereira (2004).

De acordo com Cordeiro (2012), a distância euclidiana é considerada a métrica ideal pelo fato de adotar as medidas reais dos espaços euclidianos, apresentando ótimos resultados na maioria das aplicações.

Carvalho Júnior et al. (2009) avaliaram a utilização de assinaturas temporais para verificar os tipos de fisionomias de Savanas no Campo de Instrução Militar de Formosa. Para isso, os autores utilizaram na sua metodologia uma etapa de classificação utilizando a Mínima Distância Euclidiana.

Outro trabalho relevante é o de Galate et al. (2014), o qual teve como objetivo realizar uma estimativa preliminar da variabilidade de parâmetros genéticos e avaliar a divergência fenotípica de matrizes de açaizeiro. Esta divergência foi estimada pela distância euclidiana média padronizada com os métodos de agrupamento *Unweighted Pair-Group Method using Arithmetic Averages* (UPGMA) e de *Tocher*.

Já Prado (2015) utilizou a distância Euclidiana para a obtenção de matrizes de distância para analisar as taxas de incidência de dengue nos 27 estados brasileiros,

no período de 2000 a 2014. Através disso, o autor conseguiu identificar e descrever o perfil dos estados com comportamentos similares a fim de contribuir para o redirecionamento das estratégias dos programas de controle dessa endemia e fomentar ações de prevenção.

Todos esses trabalhos e demais outros já realizados, demonstram o quanto a distância Euclidiana é popularizada e utilizada em diversas áreas buscando resolução de problemas.

Além disso, vale ressaltar que os *softwares* de SIG utilizam bastante o conceito da distância Euclidiana na implementação de algumas de suas ferramentas para classificação, análise de agrupamentos, análises espaciais, dentre outras, visto a sua ampla utilização e eficiência computacional. Além disso, dois métodos já citados que serão utilizados nessa pesquisa também utilizam a métrica Euclidiana em suas formulações, que é o Influência do Vértice e a Distância de *Hausdorff*.

4.2.2. Distância de *Manhattan*

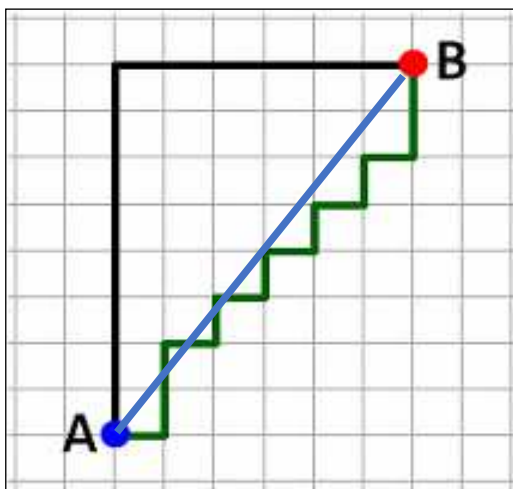
De acordo com Lei et al. (2017), a distância de *Manhattan*, também chamada *City Block*, ou Geometria do Táxi, é uma medida de cálculo de distância que ao contrário da distância Euclidiana não se eleva as diferenças ao quadrado e nem obtém-se a raiz, permitindo que uma considerável distância entre dois pontos únicos não tenha uma influência elevada no valor de distância final. A fórmula da distância *Manhattan* é apresentada pela Equação 7.

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (7)$$

Onde x_i e y_i são, respectivamente, os valores d i-ésimo atributo para os dados x e y.

Souza et al. (2012) diz que essa métrica possui este nome em referência às ruas da ilha de *Manhattan*, em virtude da disposição quadriculada. Sendo assim, a distância de *Manhattan* foi utilizada para resolver uma situação que envolve qual o percurso a ser realizado por um carro entre dois pontos da cidade. Considerando que o carro deve andar somente em ruas de pavimento asfáltico. Tal situação é exemplificada na Figura 5.

Figura 5 - Distância de *Manhattan* tanto em preto, quanto em verde entre os pontos A e B. Em azul a Distância Euclidiana entre os pontos A e B.

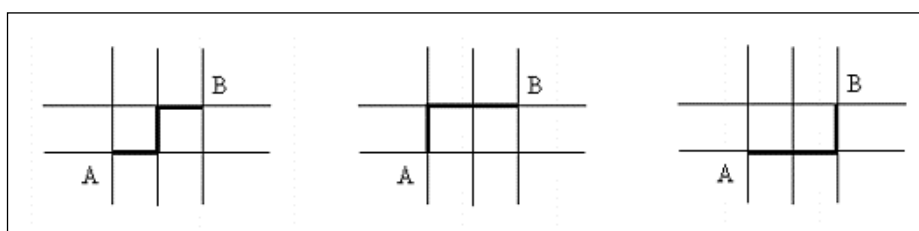


Fonte: Souza et al. (2012).

É possível notar que tanto a distância de A para B em preto como a em verde são iguais, visto que elas sobem 8 quadras e vão a direita 6 quadras, o que totaliza um deslocamento de 14 quadras.

De acordo com Gardner (1997) na geometria euclidiana, a distância mínima entre dois pontos é uma linha reta única. Já na distância de *Manhattan* pode-se ter vários caminhos, todos igualmente mínimos, que juntam dois pontos, conforme é exemplificado na Figura 6.

Figura 6 - Possibilidades de caminho entre A e B.



Fonte: Miranda (1999).

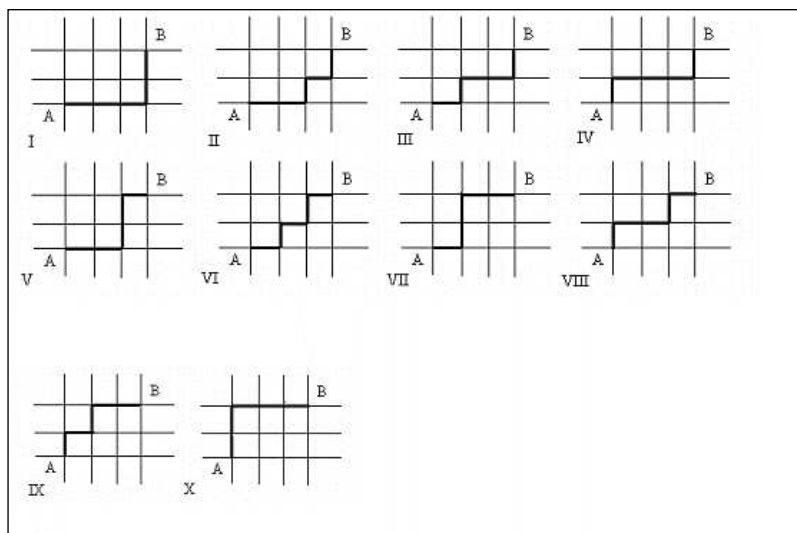
A Figura 6 ilustra três caminhos diferentes para sair do ponto A e chegar ao ponto B. No primeiro caminha-se uma unidade na horizontal, uma unidade na vertical, e uma unidade na horizontal totalizando 3 unidades para todo o percurso. No segundo caminho inicia-se uma unidade na vertical e duas unidades na horizontal, perfazendo também três unidades. No último, da mesma forma, a distância entre os pontos A e B

é de três unidades, só que nessa situação anda-se duas unidades na horizontal e uma unidade na vertical.

No último, da mesma forma, a distância entre os pontos A e B é de três unidades, só que nessa situação anda-se duas unidades na horizontal e uma unidade na vertical.

Observa-se nesse caso então, que o número de unidades na vertical é sempre igual a um e o número de unidades na horizontal é sempre igual a dois. À medida que os pontos se distanciam, a quantidade de caminhos diretos irá aumentar, conforme a Figura 7.

Figura 7 - Diferentes caminhos entre os pontos A e B.



Fonte: Miranda (1999).

Sendo assim, César (2010) afirma que de forma genérica, pode-se calcular a quantidade de caminhos diretos através da Equação 8:

$$Q = \frac{(m+n)!}{m!n!} \quad (8)$$

Onde m é a distância horizontal entre os pontos A e B e n , a distância vertical.

Essa medida pertence a um conjunto de espaços métricos criados pelo matemático Minkowski (1865–1909) (MIRANDA, 1999). Conforme César (2010) em sua pesquisa, o enfoque principal dessa métrica deve ser no uso em questões geográficas, visto que sua forma de mensurar a distância entre dois pontos é mais

próxima do mundo real das cidades, do que a forma habitual através da geometria Euclidiana.

Como já foi dito, a distância de *Manhattan* é denominada por alguns autores como Geometria do Táxi. Um exemplo é que em 1975 Eugene F. Krause em seu livro “*Táxicab Geometry: An adventure in non-euclidean geometry*” descreveu essa métrica sob dois pontos de vista: uma perspectiva educacional, como um instrumento didático, tendo como objetivo despertar o interesse daqueles que tem seu estudo focado na geometria. Já o seu segundo ponto de vista foi visando uma geometria de importantes aplicações práticas.

Esse estudo serviu como base para a tese de doutorado em 1978 na Universidade de Michigan, em Ann Arbor, onde J. Shun C. Lau estudou as propriedades básicas desta medida de similaridade e realizou aplicações, determinando áreas de influência econômica e estudou caminhos urbanos preferidos pela população da cidade de Ann Arbor, a qual a malha é retangular. Segundo Gusmão et al. (2017), essa métrica com o decorrer dos anos se difundiu no meio acadêmico, onde foi criticada e apreciada por diversos estudiosos.

Kugler (2003) utilizou essa medida de similaridade para tornar a estrutura de uma rede neural do tipo LVQ (*Learning Vector Quantization* - Quantização de Vetores por Aprendizagem) simples de ser implementada em hardware. Caballero (2006) explica e demonstra que existem diferenças significativas entre a distância Euclidiana e a distância de *Manhattan* através da elaboração de um mapa de distorção que simula a propagação de um incêndio.

Vadivel et al. (2003) utilizaram a distância de *Manhattan*, distância Euclidiana, distância da Interseção do Histograma, entre outras, para vários histogramas de cores em um grande banco de dados de imagens e os resultados mostraram que a distância de *Manhattan* tem um desempenho melhor do que as outras métricas de distância para todos os tipos de histogramas adotados.

Há outros trabalhos que envolvem o emprego da métrica de *Manhattan*, como por exemplo Jorge et al. (2011), que analisaram a influência de medidas de similaridade para previsão de dados temporais e concluíram que a mais adequada é

a distância de *Manhattan*, pois foi a medida que apresentou menor custo computacional em comparação com as demais utilizadas (distância Euclidiana e métrica L3). Além disso, foi a distância que não demonstrou diferença estatisticamente significativa entre os erros de previsão das medidas de similaridade adotadas.

Já Kugler (2003) afirma que a distância de *Manhattan* é considerada uma métrica simples, eficiente e de fácil implementação computacional, além de ser mais eficiente para aplicações em tempo real em função da sua simplicidade.

Ponnmoli e Selvamuthukumaran (2014) utilizaram as métricas de *Manhattan*, distância Euclidiana, distância de Chebyshev, dentre outras, para analisar o reconhecimento facial através de uma imagem segmentada. Esses autores verificaram que a taxa de reconhecimento da segmentação da imagem mostra o resultado com precisão de 97% utilizando a distância de *Manhattan* em comparação, por exemplo com a distância Euclidiana.

Outros exemplos em que se utilizou a métrica de *Manhattan* como medida de similaridade podem ser vistos em Eidenberger (2006), Balan (2007), Buggati et al. (2008), Nguyen (2008) e Bueno (2009).

A partir do exposto acima, verifica-se que a distância de *Manhattan* já foi utilizada em diversas áreas, mostrando-se como uma alternativa eficiente. Dessa forma, uma nova área que pode ser verificada a possibilidade de utilização dessa métrica é o controle de qualidade posicional através da adaptação de métodos já existentes ou por meio de novos métodos.

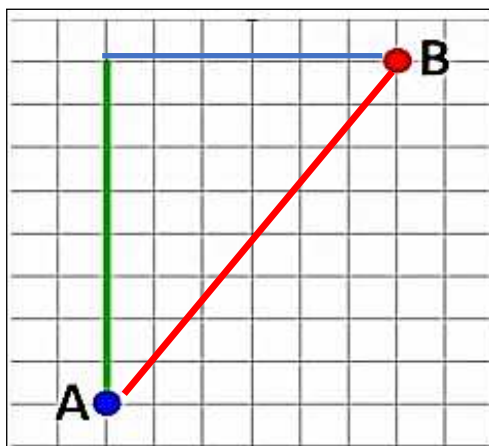
4.2.3. Distância de *Chebyshev*

A distância de *Chebyshev* leva em consideração a maior distância entre um par de variáveis (Figura 8). Com isso, para cada par de amostra de dados, compara-se a diferença entre cada par de atributos, e o maior valor de diferença entre estes elementos é tomado como a distância entre as amostras de dados (LEI et al., 2017). A Equação 9 representa como é calculada esta métrica de distância.

$$d(x, y) = \max_i |x_i - y_i| \quad (9)$$

Onde x_i e y_i são, respectivamente, os valores d i-ésimo atributo para os dados x e y .

Figura 8 - Distância de *Chebyshev* em verde e azul entre os pontos A e B.
Em vermelho a Distância Euclidiana entre os pontos A e B.



Fonte: Souza et al. (2012).

Considerando a Figura 8, a Distância de *Chebyshev*, representada pela linha verde da figura, será:

$$d(x, y) = \max |8, 6| = 8,$$

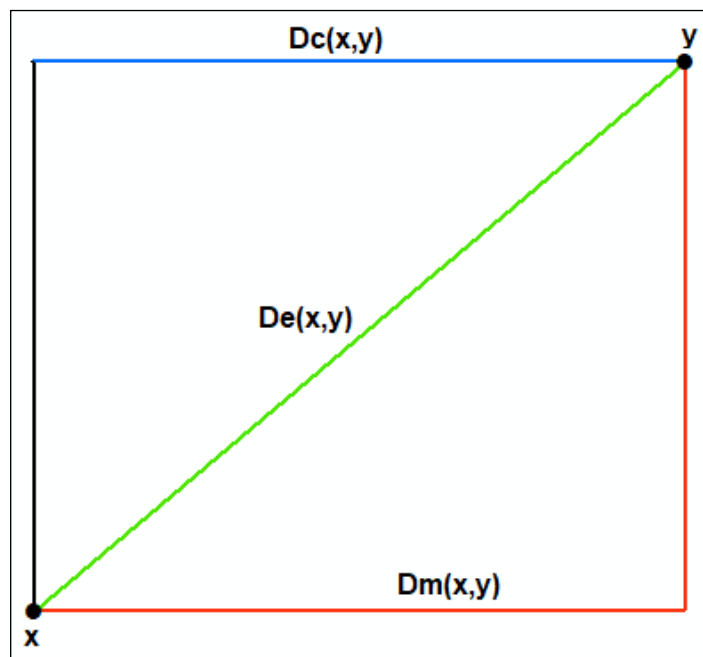
De acordo com Lei et al. (2017), a distância de *Chebyshev* é também conhecida como Distância do Tabuleiro de Xadrez, visto que a distância entre dois espaços do tabuleiro refere-se a mínima quantidade de movimentos para que o rei possa se deslocar entre eles. Ou seja, o movimento do rei é igual a distância de *Chebyshev* entre os pontos centrais de cada quadrado.

Essa métrica também pertence ao conjunto de espaços métricos elaborados pelo matemático Minkowski, onde de acordo com Lima (2015) pode ser expressa pela Equação 10.

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^{1/p} \right)^p \quad (10)$$

Sendo que quando p é 1, obtêm-se a fórmula da distância de *Manhattan*. Já quando $p=2$, tem-se a fórmula da distância Euclidiana, e se p tende a infinito, obtêm-se a distância de *Chebyshev*. Palma (2018) exemplifica essa situação conforme pode ser visto na Figura 9.

Figura 9 - Exemplificação de diferentes tipos de distâncias: D_c - Distância de *Chebyshev*, D_m - Distância de *Manhattan* e D_e - Distância Euclidiana.



Fonte: Adaptado de Palma (2018).

De acordo com a Figura 9, dados dois pontos x e y verifica-se que em azul está representada a distância de *Chebyshev* (Equação 11), que é a maior componente em cada eixo x e y, ou seja:

$$D_c = \max(\Delta x, \Delta y) \quad (11)$$

Já a distância Euclidiana (Equação 12) é ilustrada pela linha na cor verde, referindo-se como a menor distância conhecida entre os pontos.

$$D_e = \sqrt{\Delta x^2 + \Delta y^2} \quad (12)$$

Por último, em vermelho, está ilustrada a distância de *Manhattan* (Equação 13), onde a distância entre dois pontos é obtida pela soma das diferenças absolutas de suas coordenadas.

$$Dm = \Delta x + \Delta y \quad (13)$$

Vicini e Souza (2005) afirmam que entre as medidas mais utilizadas, para estabelecer o conceito de distância entre dois objetos, pode-se adotar além de outras métricas a distância de *Chebyshev*. De acordo com esses autores um exemplo de aplicação dessa medida é em uma classificação hierárquica, onde os objetos são agrupados de acordo com sua semelhança, a qual é baseada em uma classificação taxonômica e é representada em um gráfico com estrutura de árvore, conhecida como dendograma. E para ser feita esta classificação, é necessário definir matematicamente o que será caracterizado como proximidade, o que pode ser feito através da distância entre dois objetos, determinando o critério de agrupamento de duas classes.

Cha (2007) em seu trabalho afirma que as medidas de distância ou similaridade são cruciais para resolver muitos problemas de reconhecimento de padrões, como classificação, clusters, e problemas de recuperação. Dentre as métricas que podem ser utilizadas não somente nessa área, mas em outras também, o autor cita a distância de *Chebyshev*, que é uma das medidas de similaridade que difere de uma das mais utilizadas, a Euclidiana, sendo uma nova opção de se adotar.

Já Oliveira (2017) realizou um estudo do processo de soldagem para investigar a integridade da solda no tubo de aço, visando a localização e dimensionamento do defeito em um intervalo de milímetros, que é utilizado para aprovar ou rejeitar os tubos. Para isso esse autor utilizou uma matriz de correlação para quantificar o grau de homogeneidade decrescente da estrutura do aço base, da solda de referência e do tubo teste. Posteriormente aplicou a métrica de *Chebyshev* nos dados e observou que os resultados foram consistentes com os apresentados pela matriz de correlação. Além disso, a partir do dendograma da distância de *Chebyshev* versus o número de clusters, foi possível localizar o defeito da solda.

Gomes (2018) realizou uma análise de agrupamento hierárquico para seleção de genótipo em mandiocas visando a formação de grupos divergentes de genótipos, em que a melhor combinação desses métodos hierárquicos foi a distância de

Chebyshev em comparação com a outra medida de similaridade utilizada, que foi a distância de *Mahalanobis*.

Aragão et al. (2019) utilizou a mineração de dados por meio de instâncias para prever volume de um povoamento do gênero *Eucalyptus* e através do algoritmo k-vizinhos mais próximos foi feita a estimativa. Os autores testaram as medidas de dispersão com a distância de *Chebyshev* e a distância euclidiana, considerando um, cinco e dez vizinhos para cada. Concluiu-se que a técnica analisada foi eficiente para estimar o volume das árvores de eucalipto e entre os modelos testados, a distância de *Chebyshev* com 10 vizinhos denotou desempenho superior aos demais.

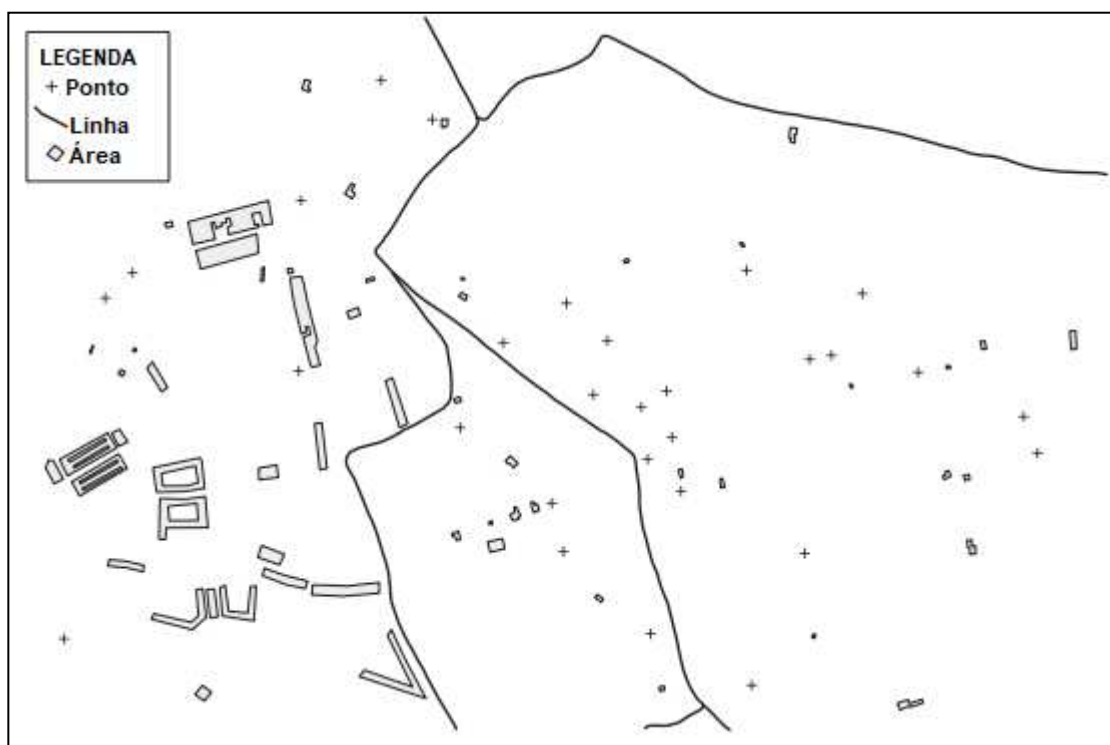
Mais exemplos da aplicação da Distância de *Chebyshev* pode ser visto em Marques (2006), Balan (2007), Bueno (2009), Ponciano-Silva (2009), Dai et al. (2011) e Sharma e Batra (2014).

Diante disso, é possível perceber que a distância de *Chebyshev* pode ser analisada para verificar a viabilidade de sua utilização no controle de qualidade posicional por meio de adaptação de métodos já existentes ou através de novos métodos, já que foi empregada como uma medida de similaridade em outros trabalhos.

4.3. Base MatchingLand

Xavier et al. (2017) elaboraram um conjunto de dados com o objetivo de auxiliar na avaliação de métodos geoespaciais utilizando dados vetoriais, incluindo a análise da qualidade posicional. Tal conjunto foi construído a partir de dados de mapeamento produzidos por agências cartográficas espanholas (Figura 10).

Figura 10 - Amostra da região de *MatchingLand*.



Fonte: Xavier et al. (2017).

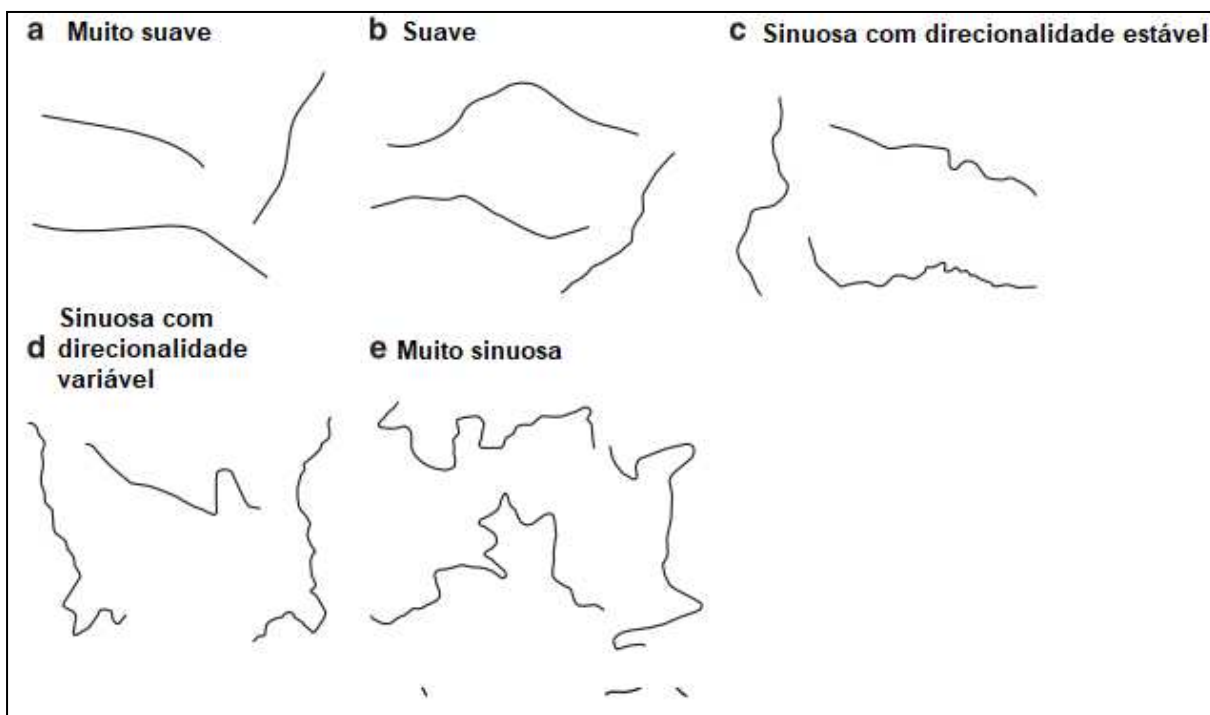
Os dados iniciais foram submetidos a transformações geométricas para gerar um conjunto de dados sintéticos. As transformações utilizadas simbolizam fatores que podem influenciar o desempenho de métodos de avaliação geoespacial, sendo estes a morfologia de características lineares ou de área, transformações sistemáticas e distúrbios aleatórios sobre os dados iniciais.

Xavier et al. (2017) separaram os dados em quatro conjuntos: dados iniciais; morfologia modificada; perturbação sistemática e perturbação aleatória. Os dados iniciais foram obtidos nas escalas 1: 25.000 e 1: 10.000 através das agências espanholas. Os outros três grupos foram derivados do conjunto de dados iniciais na escala 1: 25.000.

O conjunto de dados de teste foram divididos em nove regiões (S1-S9), bem como também foram elaboradas as regiões correspondentes nos conjuntos de dados de referência (B1-B9). Como a morfologia de objetos lineares e de área pode afetar o desempenho do processo de avaliação geoespacial, os autores adotaram uma

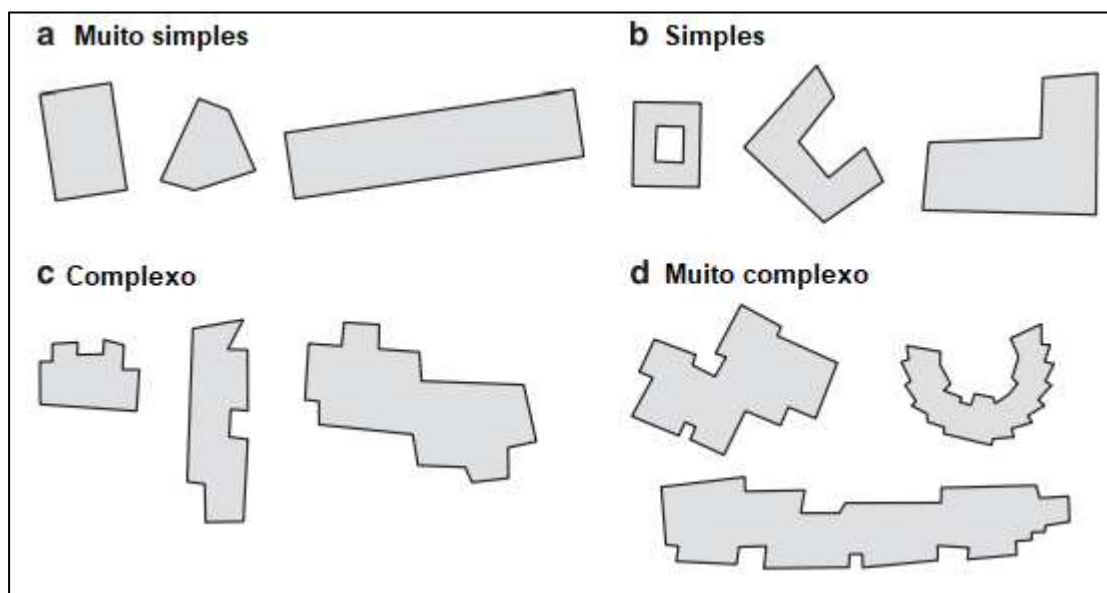
classificação de rugosidade para linhas (Figura 11) e desenvolveram uma classificação de complexidade para áreas (Figura 12).

Figura 12 - Amostra das classes de morfologia das linhas.



Fonte: Xavier et al. (2017).

Figura 11 - Amostra de morfologia das áreas.



Fonte: Xavier et al. (2017).

Já o objetivo do conjunto de dados com perturbação sistemática inicial, foi identificar a influência de perturbações sistemáticas intencionais sobre os procedimentos de correspondência.

Considerando esse caso, os autores simularam vários deslocamentos sobre os dados originais, como translações, rotações e escala, gerando os dados sintéticos com a transformação afim. Em cada deslocamento foram gerados um conjunto de distúrbios sistemáticos, sendo: translações em oito direções; rotações no sentido anti-horário e horário em três pivôs diferentes e; dois fatores de escala (dilatação e cisalhamento).

Já considerando a perturbação sistemática, os autores utilizaram quatro deslocamentos: 5 - 12,5 - 25 e 50 metros. Tais valores foram escolhidos levando em consideração o padrão brasileiro de qualidade dos dados espaciais, em que segundo os autores, o erro posicional máximo aceito para dados de 1: 25.000, variando de 7,0 a 25 metros. Sendo assim, foi escolhido um valor abaixo deste limite (5 m), dois valores dentro do limite (12,5 e 25 m) e outro valor acima (50 m). Esses quatro deslocamentos originaram 7.425 combinações diferentes para cada tipo de geometria (ponto, linha e área).

Para a validação dos dados foram feitos os seguintes tipos de análises: análise da qualidade dos dados espaciais, verificação de pares correspondentes e teste de randomização.

Na avaliação da qualidade dos dados espaciais, foi englobada os procedimentos que garantem a qualidade dos conjuntos de dados espaciais no *Shapefile*. Utilizou-se a estrutura de avaliação publicada na ISO 19157: 2013. Conforme este padrão, existem cinco categorias de elementos de qualidade para dados espaciais: integridade, precisão posicional, precisão temática, qualidade temporal e consistência lógica. Como trata-se de dados sintéticos, apenas a consistência lógica foi avaliada, sendo que nesta categoria avalia-se dois elementos de qualidade: consistência de formato e consistência topológica.

Em relação à consistência de formato, nenhum erro foi encontrado. A consistência topológica foi avaliada adotando como medida de correção a

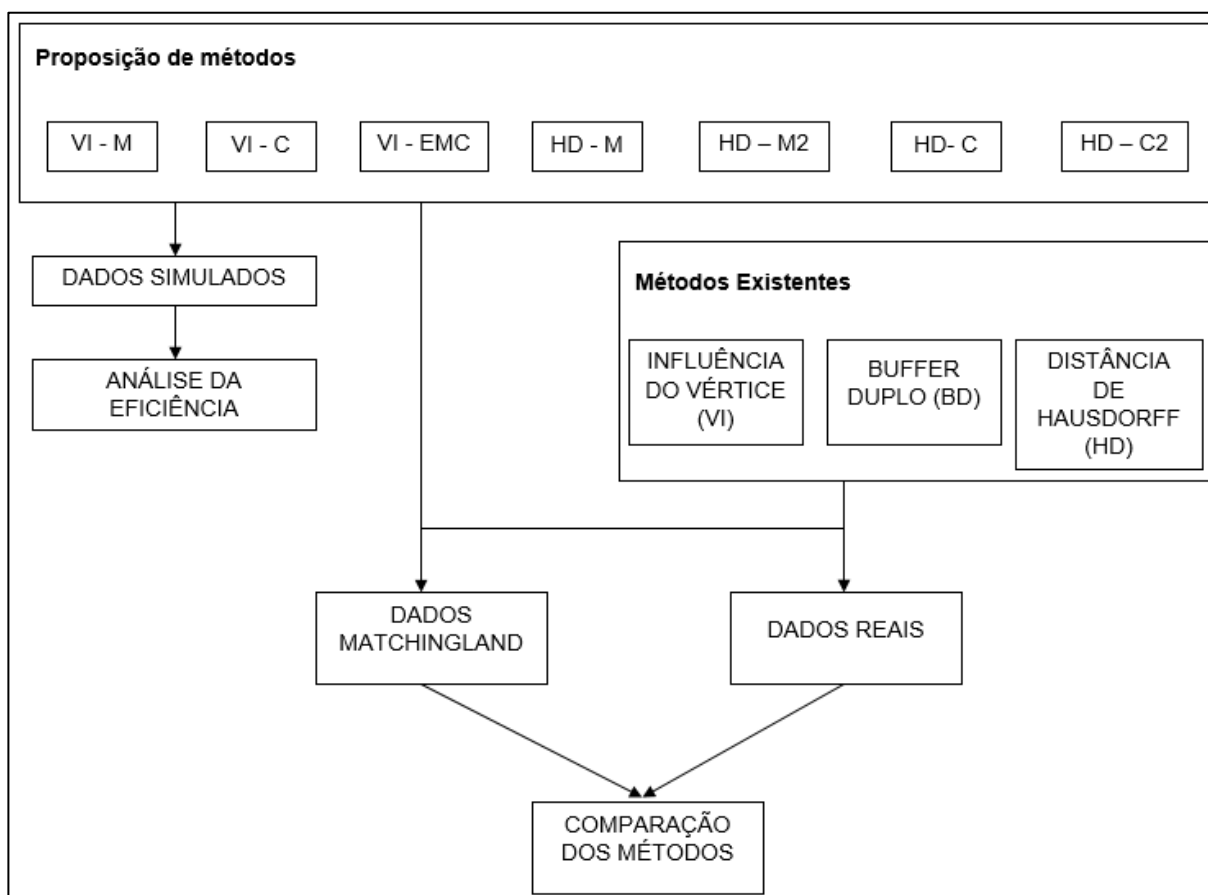
conformidade das geometrias com a Especificação de Características Simples publicado pela *Open Geospatial Consortium* (OGC), e todas as geometrias avaliadas passaram no teste.

Em relação à qualidade dos dados, verificou-se que as discrepâncias encontradas foram abaixo de 5%. A última análise de qualidade refere-se à aleatoriedade do vetor de deslocamento usado para criar o grupo de perturbação aleatória do conjunto de dados. Foi utilizado o variograma para avaliar qualquer dependência espacial de seus vetores. Usando o pacote *geoR* verificou-se que não há dependência espacial no vetor de deslocamento de campo.

5. METODOLOGIA

A metodologia adotada para o desenvolvimento deste trabalho encontra-se ilustrada no fluxograma da Figura 13.

Figura 13 - Fluxograma da metodologia utilizada.



Conforme o fluxograma (Figura 13), o objetivo é utilizar a Distância de *Manhattan* e a Distância de *Chebyshev*, adaptando o método já existente “Influência do Vértice” e “Distância de Hausdorff”, para a avaliação da acurácia posicional utilizando feições lineares. Este métodos foram denominados de: Método “VI-M” (Influência do Vértice com a Distância de *Manhattan*), Método “VI-C” (Influência do Vértice com a Distância de *Chebyshev*), Método “VI-EMC” (Influência do Vértice com Distância Euclidiana, *Manhattan* e *Chebyshev*), Método “HD-M” (Distância *Hausdorff* com a métrica de *Manhattan*), Método “HD-M2” (Distância *Hausdorff* com a métrica de *Manhattan* adaptado), Método “HD-C” (Distância *Hausdorff* com a métrica de

Chebyshev) e Método “HD-C2” (Distância *Hausdorff* com a métrica de *Chebyshev* adaptado).

Dessa forma, foram utilizados três conjuntos de dados, sendo eles: (i) dados simulados, para simular situações que podem ocorrer no mundo real; (ii) dados da base da Espanha denominados de “*MatchingLand*”, que são dados que simulam o comportamento de dados reais, e; (iii) um conjunto de dados reais. Os dados simulados foram utilizados com o objetivo de testar a eficiência dos novos métodos propostos tendo como base os valores de discrepâncias conhecidos *a priori*. Essa eficiência é avaliada somente nos dados simulados, já que se cria linhas com determinadas condições e visa-se verificar se o método está apresentando um bom comportamento ou não. Já nos dados reais, muitas vezes não se possui o controle ou a possibilidade de ter determinadas situações para verificar se o método é eficiente ou não em reportar a incerteza posicional.

Já os dados *MatchingLand* e os dados reais foram avaliados também através dos métodos já existentes da Influência do Vértice, *Buffer Duplo* e Distância de *Hausdorff*, conforme metodologia de Santos et al. (2005), para fins de comparação com os métodos sugeridos.

Visando uma comparação entre os métodos, foi realizado a Análise de Variância (ANOVA), que trata-se de um método estatístico que permite realizar comparações simultâneas entre duas ou mais médias, ou seja, permite testar hipóteses sobre médias de distintas populações, baseado nas análises de variâncias amostrais.

A Análise de Variância também se estrutura sob algumas suposições ou pressupostos para que seja aplicável, sendo estes:

- Todas as observações devem ser independentes;
- As observações em cada grupo devem possuir uma distribuição, aproximadamente normal;
- As variâncias em cada grupo devem ser aproximadamente iguais.

Dentro dos procedimentos estatísticos da ANOVA, fez-se a análise da normalidade das observações a partir do teste do Shapiro-Wilk, já para a análise de

quais médias são significativamente distintas entre si, utilizou-se o teste de Tukey, com nível de significância de 5%.

A descrição dos métodos de feições lineares propostos, bem como o uso de cada conjunto de dados, será exposta em sequência.

5.1. Métodos propostos

5.1.1. Método “VI - M” (Influência do Vértice com a Distância de *Manhattan*) e Método “VI - C” (Influência do Vértice com a Distância de *Chebyshev*)

Conforme descrito no tópico 4.1.1.1., o método Influência do Vértice utiliza a métrica Euclidiana em sua definição. Uma das proposições dessa pesquisa é utilizar outras métricas, como a Distância de *Manhattan* e a Distância de *Chebyshev* no método Influência do Vértice, verificando a viabilidade do uso das mesmas.

Entretanto, um problema encontrado durante a execução da pesquisa, foi em relação de como encontrar a menor distância de *Manhattan* ou *Chebyshev* entre um ponto e uma linha em um ambiente computacional, no caso SIG, conforme é feito utilizando a distância Euclidiana.

Devido a isso, optou-se por densificar a linha de teste considerando sua escala, o erro de graficismo de 0,2 mm (ABNT, 1994) e a eficiência computacional. Dessa forma, foram realizados testes com a base *MatchingLand*, visto que é uma base de dados que simula dados reais, considerando diversos valores de densificação (Dens). Sendo assim, através dos testes realizados que constam no Apêndice, estabeleceu-se uma equação empírica para o valor da densificação (Equação 10).

$$Dens (cm) = (0,2 * esc)/5000 \quad (10)$$

Onde *esc*, é o *denominador* da escala de teste.

Esse artifício utilizado foi necessário devido à falta de ferramentas com estas métricas implementas em SIG. Assim, em vez de calcular a distância entre um ponto da linha de referência e a linha de teste, utilizou-se a distância entre pontos da linha de referência e pontos da linha de teste densificada. Para cada vértice da linha de referência foi gerada uma tabela com n pontos da linha teste mais próximos do vértice da linha de referência (de acordo com a distância Euclidiana). Com o vértice de

referência e os n vértices testes, foram calculadas as n distância de *Manhattan* e n distância de *Chebyshev*. Dessa forma, o vértice da linha de teste que apresentou o menor valor de distância seria considerado para o cálculo.

Devido a essa densificação, outra questão a ser resolvida foi em relação à quantidade de vértices da linha de teste a serem considerados para obter a menor Distância de *Manhattan* ou *Chebyshev* da linha de referência para essa linha teste. Isso foi necessário ser considerado pelo fato da linha teste densificada possuir um número alto de vértices, o que poderia comprometer a eficiência computacional.

Dessa forma, para resolver essa questão, foi elaborada uma ferramenta no *ArcToolbox* do software ArcGIS 10.5, onde verifica-se o número de vértices da linha de referência e da linha teste antes de densificar, e define-se a quantidade de vértices (n) a ser considerada para o cálculo das distâncias, igual à quantidade de vértices da linha com menor valor (com menor quantidade de vértices) (Apêndice A).

Sendo assim, para cada vértice da linha de referência foi gerada uma tabela com n pontos da linha teste mais próximos do vértice da linha de referência (de acordo com a distância Euclidiana). Com o vértice de referência e os n vértices testes, foram calculadas as n distâncias de *Manhattan* e de *Chebyshev*. Dessa forma, o vértice da linha de teste que apresentou o menor valor de distância seria considerado para o cálculo.

A descrição dessa ferramenta, bem como os métodos elaborados estão contidos nos Apêndices A, B e C.

Para fins de exemplificação, tem-se representado as três métricas (Figuras 14 e 15), considerando uma densificação de 5 cm em um par de linhas, considerando um vértice superior e o central da linha de referência. Em preto a distância Euclidiana, em azul a *Manhattan* e em roxo a distância de *Chebyshev*. A linha de referência em preto, o vértice em verde e a linha teste em vermelho.

Figura 14 – Ilustração das menores distâncias obtidas com as métricas Euclidiana, *Manhattan* e *Chebyshev* considerando o vértice superior da linha de referência.

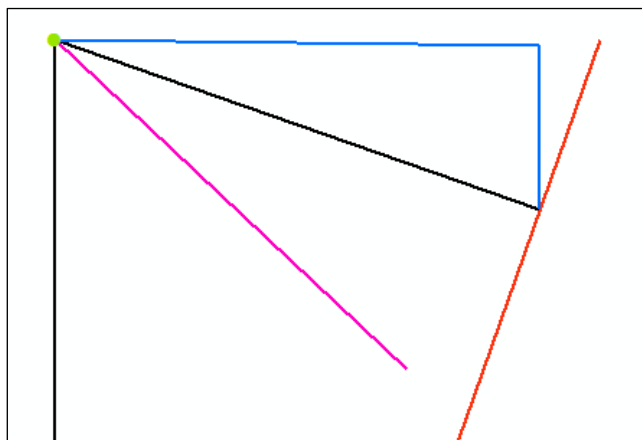
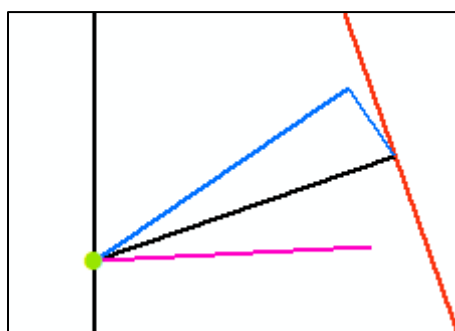


Figura 15 - Ilustração das menores distâncias obtidas com as métricas Euclidiana, *Manhattan* e *Chebyshev* considerando o vértice central da linha de referência.



Ressalta-se que a distância de *Chebyshev* não chega na linha teste devido à sua definição de ser sempre menor ou igual à distância Euclidiana. Dessa forma, ela só chegará na linha teste se tiver o mesmo valor da distância Euclidiana. A distância de *Manhattan* sempre chega na linha teste devido à mesma ser sempre maior ou igual à distância Euclidiana.

Sendo assim, essa adaptação do método Influência do Vértice no processo de avaliação da acurácia posicional, baseada em feições lineares consistiria em (Apêndice C):

1. Selecionar uma amostra de n linhas homólogas de teste (L_T) e de referência (L_R);
2. Calcular o comprimento (CL_R) e extrair os vértices de cada linha i de referência;

3. Verificar a quantidade de vértices da linha de referência e da linha teste (antes de densificar). Considerar o valor de “p” igual à menor quantidade de vértice da linha de referência ou da linha de teste;
4. Realizar a densificação da linha teste, de acordo com a Equação 10;
5. Calcular a menor distância (dh_k) de *Manhattan* ou *Chebyshev* (deve-se escolher qual métrica irá utilizar) para cada vértice k da linha teste densificada;
6. Obter os valores dos comprimentos dos segmentos de retas anteriores ($l_k^{(k-1)}$) e posteriores ($l_k^{(k+1)}$) aos vértices k da linha i de referência;
7. Calcular a discrepância média (D_i) para cada linha i de referência (Equação 3).

Para o caso em que se deseja aplicar o Decreto/ET-CQDG, a metodologia deve seguir os passos citados, além de:

- 90% das linhas testadas devem apresentar valores de D menores ou iguais ao valor do PEC e;
- O RMS da amostra de discrepâncias médias D tem que ser menor ou igual ao valor do EP, para a classe e escala empregadas no processo de avaliação da acurácia posicional.

5.1.2. Método “VI - EMC” (Influência do Vértice com Distância Euclidiana, *Manhattan* e *Chebyshev*)

Outra adaptação do Influência do Vértice proposta consiste em utilizar a distância Euclidiana, a distância de *Manhattan* e a distância de *Chebyshev* em conjunto, realizando uma média entre essas três métricas, para verificar qual é o comportamento destas em conjunto.

Dessa forma, essa adaptação do método Influência do Vértice para feições lineares consistiria em (Apêndice C):

1. Selecionar uma amostra de n linhas homólogas de teste (L_T) e de referência (L_R);
2. Calcular o comprimento (CL_R) e extrair os vértices de cada linha i de referência;

3. Verificar a quantidade de vértices da linha de referência e da linha teste (antes de densificar). Considerar o valor de “p” igual à menor quantidade de vértice da linha de referência ou da linha de teste;
4. Realizar a densificação da linha teste, de acordo com a Equação 10;
5. Calcular a menor distância (dh_i) Euclidiana, de Manhattan e Chebyshev para cada vértice k da linha referência em relação aos vértices da linha teste densificada. Posteriormente obter a média das três métricas;
6. Obter os valores dos comprimentos dos segmentos de retas anteriores ($l_k^{(k-1)}$) e posteriores ($l_k^{(k+1)}$) aos vértices k da linha i de referência;
7. Calcular a discrepância média (D_i) para cada linha i de referência (Equação 3).

Para o caso em que se deseja aplicar o Decreto/ET-CQDG, a metodologia deve seguir os passos citados, além de:

- 90% das linhas testadas devem apresentar valores de D menores ou iguais ao valor do PEC e;
- O RMS da amostra de discrepâncias médias D tem que ser menor ou igual ao valor do EP, para a classe e escala empregadas no processo de avaliação da acurácia posicional.

5.1.3. Método “HD - M” (Distância *Hausdorff* com a métrica de *Manhattan*) e Método “HD - C” (Distância *Hausdorff* com a métrica de *Chebyshev*)

Além da adaptação do método Influência do Vértice, objetiva-se definir novos métodos utilizando a Distância de *Manhattan* e a Distância de *Chebyshev*. Esses novos métodos têm como base o método Distância de *Hausdorff*, sendo sugeridos dois novos métodos para verificar qual apresenta melhor desempenho, levando em consideração características como o número de vértices das linhas, sinuosidade, comprimento, dentre outros.

O método HD – M e HD – C, consiste em calcular a menor Distância de *Manhattan* ou *Chebyshev* de cada vértice da linha de referência para a linha teste, sendo feita uma média posteriormente com as menores distâncias obtidas em cada par de linhas. Em seguida seria calculada também a menor Distância de *Manhattan*

ou *Chebyshev* de cada vértice da linha teste para a linha de referência, também realizando a média.

Sendo assim, a discrepância do par de linhas seria a média das médias das distâncias obtidas. Ou seja:

1. $D1$ = média das menores distâncias de *Manhattan* ou *Chebyshev* (deve-se escolher qual métrica irá utilizar) dos vértices da linha de referência para a linha teste;
2. $D2$ = média das menores distâncias de *Manhattan* ou *Chebyshev* (também deve-se escolher qual métrica irá utilizar) dos vértices da linha teste para linha de referência;
3. Onde a discrepância final de cada par de linha i seria (Equação 11):

$$D_{m(i)} = \text{média} (D1, D2) \quad (11)$$

Nesse caso também é utilizado o artifício de densificar as linhas para obter as menores distâncias.

Sendo assim, a utilização deste método no processo de avaliação da acurácia posicional, baseada em feições lineares, consistiria em (Apêndice C):

1. Selecionar uma amostra de n linhas homólogas de teste (L_T) e de referência (L_R);
2. Verificar a quantidade de vértices da linha de referência e da linha teste antes de densificar. Dessa forma, a quantidade de vértices da linha teste a ser considerada para o cálculo da distância de *Manhattan* ou *Chebyshev* será igual ao número de vértices da linha com menor quantidade;
3. Calcular a menor distância ($D1$) de *Manhattan* ou *Chebyshev* (deve-se escolher qual métrica irá utilizar) para cada vértice k da linha referência em relação aos vértices da linha teste densificada;
4. Calcular a média das menores distâncias obtidas em cada par de linha do item 2;
5. Calcular a menor distância ($D2$) de *Manhattan* ou *Chebyshev* (deve-se escolher qual métrica irá utilizar) para cada vértice k da linha teste em relação aos vértices da linha de referência densificada;

6. Calcular a média das menores distâncias obtidas em cada par de linha do item 5;
7. Calcular a discrepância média ($D_{m(i)}$) para cada par de linha i (Equação 11).

Para o caso em que se deseja aplicar o Decreto/ET-CQDG, a metodologia deve seguir os passos descritos e analisar se:

- 90% das linhas testadas devem apresentar valores de D menores ou iguais ao valor do PEC e;
- O RMS da amostra de discrepâncias médias D tem que ser menor ou igual ao valor do EP, para a classe e escala empregadas no processo de avaliação da acurácia posicional.

5.1.4. Método “HD - M2” (Distância *Hausdorff* com a métrica de *Manhattan* adaptado) e Método “HD - C2” (Distância *Hausdorff* com a métrica de *Chebyshev* adaptado)

Já este método é uma adaptação do método HD – M e HD - C, em que ao invés de se calcular as menores distâncias de *Manhattan* ou *Chebyshev* da linha de referência para a linha teste e vice-versa, é calculada apenas as menores distâncias da linha de referência para a linha teste.

Sendo assim, a discrepância final de cada par de linhas i seria a média das distâncias obtidas. Ou seja:

$D_{(i)}$ = média das menores distâncias de *Manhattan* ou *Chebyshev* (deve-se escolher qual métrica irá utilizar) dos vértices da linha de referência para a linha teste.

Dessa forma, a utilização deste método no processo de avaliação da acurácia posicional, baseada em feições lineares, consistiria em (Apêndice C):

1. Selecionar uma amostra de n linhas homólogas de teste (L_T) e de referência (L_R);
2. Calcular a menor distância de *Manhattan* ou *Chebyshev* (deve-se escolher qual métrica irá utilizar) para cada vértice k da linha referência em relação aos vértices da linha teste densificada. Deve-se verificar a quantidade de vértices da linha de referência e da linha teste antes de densificar. Dessa forma, a quantidade de vértices da linha teste a ser considerada para o

cálculo da distância de *Manhattan* ou *Chebyshev* será igual ao número de vértices da linha com menor quantidade;

3. Calcular a discrepância média ($D_{(i)}$) para cada par de linha i .

Para o caso em que se deseja aplicar o Decreto/ET-CQDG, a metodologia deve analisar os passos descritos e analisar se:

- 90% das linhas testadas devem apresentar valores de D menores ou iguais ao valor do PEC e;
- O RMS da amostra de discrepâncias médias D tem que ser menor ou igual ao valor do EP, para a classe e escala empregadas no processo de avaliação da acurácia posicional.

Com base nos resultados que serão obtidos a partir dos testes que serão executados, será possível analisar se os métodos propostos são eficientes para serem utilizadas no processo de avaliação da acurácia posicional com feições lineares.

5.2. Experimentos

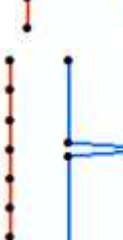
5.2.1. Dados Simulados

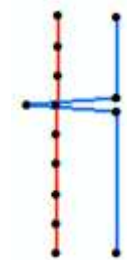
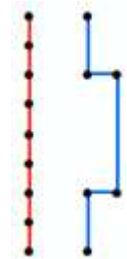
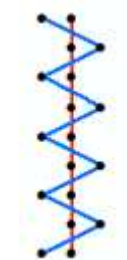
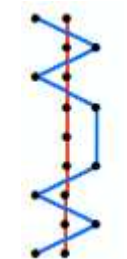
O conjunto de dados utilizado contém 15 pares de linhas artificiais, elaboradas no software ArcGIS 10.5 baseadas em Mozas & Ariza (2008) e Cruz (2020). Conforme Mozas & Ariza (2008), esses dados simulam erros que podem acontecer na realidade como variações ocasionadas por ruídos aleatórios e deslocamentos sistemáticos.

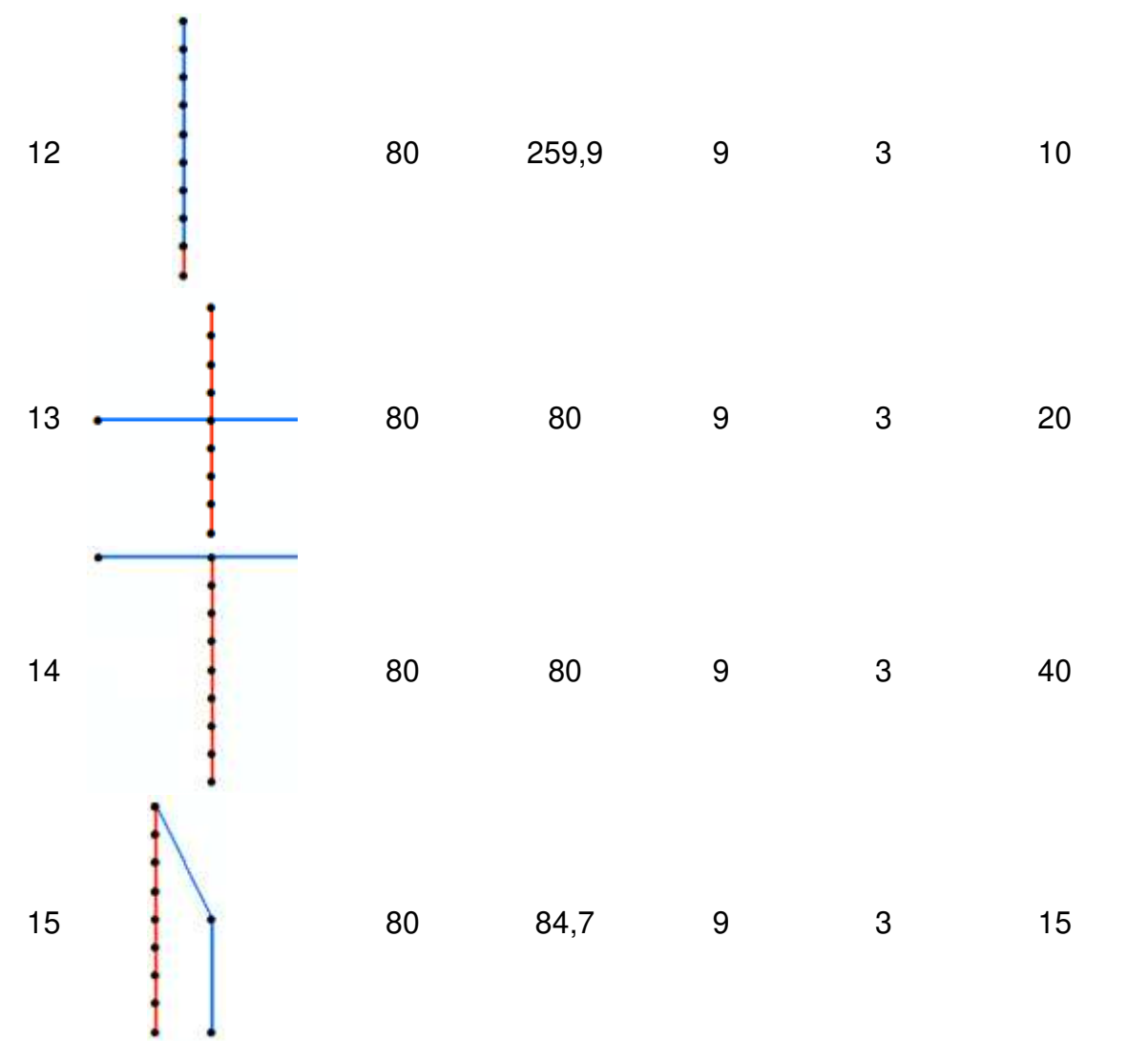
As linhas de referência estão representadas pela cor vermelha, e as linhas de teste pela cor azul. Os vértices de cada feição linear foram caracterizados pela cor preta.

Todas as linhas de referência possuem 9 vértices e 80 metros de comprimento, sendo esses valores estabelecidos de forma empírica. Já as linhas de teste variam tanto em relação à quantidade de vértices, bem como o seu comprimento (Tabela 2).

Tabela 2 - Informações das linhas de teste e das linhas de referência dos dados simulados.

ID	Feição	Comprimento (m)		Número de vértices		Discrepância média baseada na distância Euclidiana (m)
		Referência	Teste	Referência	Teste	
1		80	80	9	3	20
2		80	60	9	3	20,5
3		80	89,4	9	3	10
4		80	82,5	9	3	30
5		80	135,2	9	5	20

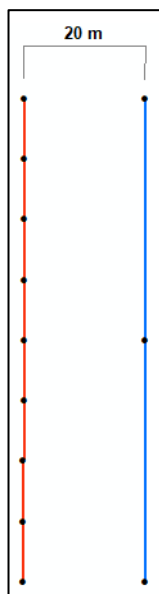
6		80	135,2	9	5	0
7		80	135,2	9	5	20
8		80	100	9	6	23,3
9		80	178,9	9	9	5
10		80	153,9	9	8	5,5
11		80	155,2	9	9	20



Fonte: Elaborado pelo autor (2021).

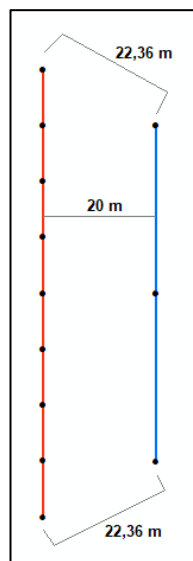
A feição linear simulada n° 1 consiste em duas linhas idênticas, paralelas entre si, reproduzindo um erro de translação (Figura 16). A linha de teste sofreu um deslocamento de 20 metros e por conta do paralelismo, essa distância é constante entre as linhas, sendo esse o valor da discrepância média desse par de linhas. Além disso, a linha de teste possui somente três vértices, no início, meio e fim da linha simulando uma generalização cartográfica, enquanto a linha de referência possui nove.

Figura 16 - Feições lineares simuladas nº 1. Linha de referência em vermelho e linha de teste em azul.



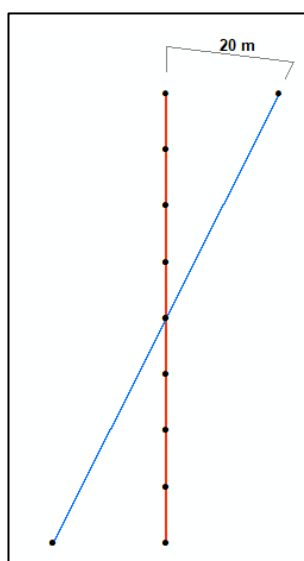
A feição linear simulada nº 2 possui uma linha de teste menor, com a dimensão de 60 metros. Foi utilizada essa diferença para simular um erro de escala e uma incerteza na determinação dos pontos finais e iniciais. É um caso parecido com a simulação 1, onde as duas feições em análise eram paralelas com uma distância de 20 metros entre si. Entretanto, neste teste 2, os vértices iniciais e finais de cada linha estiveram afastados em 22,36 m (Figura 17). Além disso, a discrepância média desse par de linha é de 20,5 metros. Essa discrepância média é calculada obtendo a distância euclidiana de cada vértice da linha de referência para a linha teste, sendo calculada uma média com os valores obtidos.

Figura 17 - Feições lineares simuladas nº 2. Linha de referência em vermelho e linha de teste em azul.



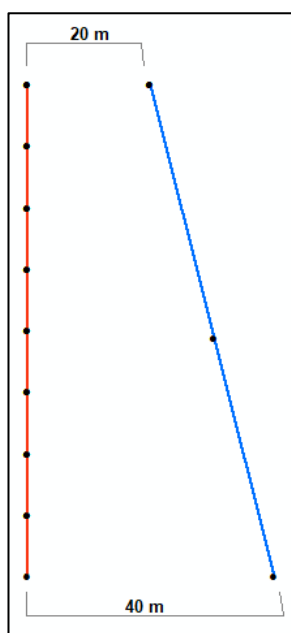
Em relação à feição linear simulada nº 3, este refere-se ao caso em que a linha de teste intercepta a linha de referência em seu vértice central, devido à uma rotação. Tanto os vértices iniciais, quanto finais de ambas as linhas estão afastados em 20 metros. Nesse caso também a linha de teste possui menos vértices que a linha de referência (Figura 18). O valor de discrepância média para esse par de linha equivale a 10 metros.

Figura 18 - Feições lineares simuladas nº 3. Linha de referência em vermelho e linha de teste em azul.



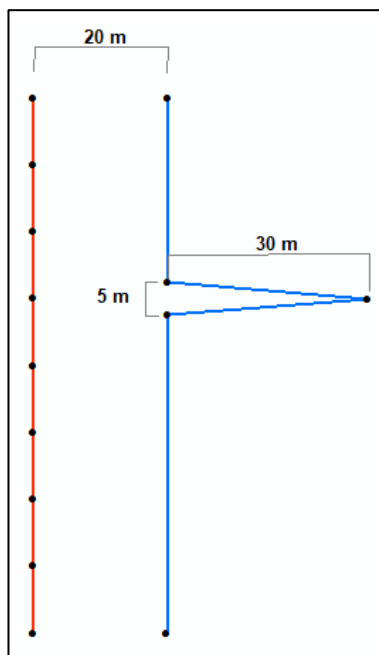
Já a feição linear simulada nº 4 representa uma situação em que a linha de teste também sofreu uma rotação se distanciando da linha de referência, onde os vértices iniciais de ambas as linhas estão afastados em 20 metros. Já os vértices finais estão distanciados em 40 metros, e a linha de referência possui apenas três vértices (Figura 19). Nesse caso, a discrepância média entre as linhas é de 30 metros.

Figura 19 - Feições lineares simuladas nº 4. Linha de referência em vermelho e linha de teste em azul.



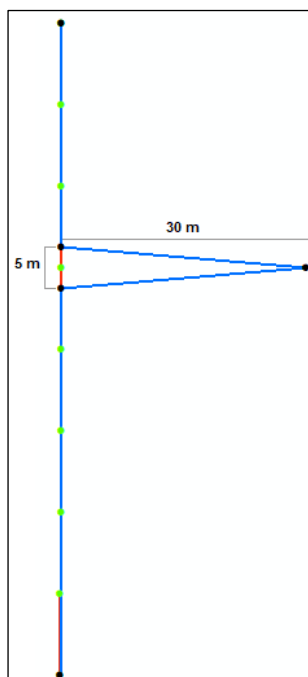
Na feição linear simulada nº 5, as linhas estão afastadas de si em 20 metros e apresenta um erro pontual (*outlier*) de 50 metros de distanciamento da linha teste para a linha de referência, fazendo com que a linha de teste possuisse 5 vértices e 135,2 metros de comprimento (Figura 20). A discrepância média entre as linhas é de 20 metros.

Figura 20 - Feições lineares simuladas nº 5. Linha de referência em vermelho e linha de teste em azul.



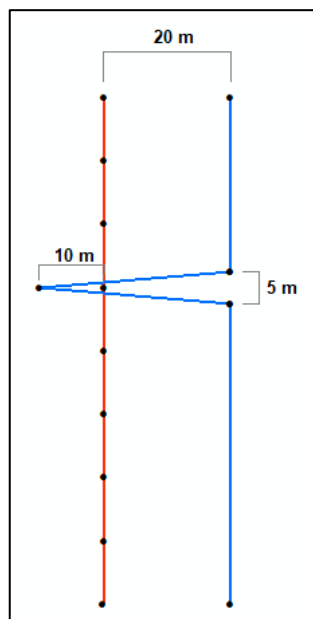
A feição linear simulada nº 6 é parecida com o anterior e representa um erro pontual (*outlier*) de 30 metros de afastamento da linha teste para a linha de referência. A linha teste coincide em grande parte com a linha de referência. Além disso, esse erro fez com que a linha de teste possuisse 5 vértices e 135,2 metros de comprimento (Figura 21). Para esse par de linha a discrepância média equivale a aproximadamente 0 metros.

Figura 21 - Feições lineares simuladas nº 6. Linha de referência em vermelho e linha de teste em azul.



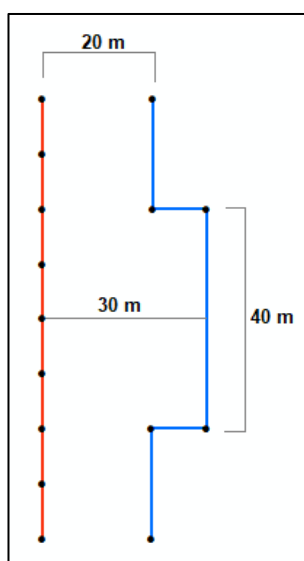
O teste da feição linear simulada nº 7 é parecido com o teste 6, onde as linhas estão afastadas de si em 20 metros. Além disso, o erro pontual possui 10 metros de distanciamento da linha teste para a linha de referência. Esse erro fez com que a linha de teste possuísse também 5 vértices e 132,5 metros de comprimento (Figura 22). A discrepância média considerada nesse caso é de 20 metros.

Figura 22 - Feições lineares simuladas nº 7. Linha de referência em vermelho e linha de teste em azul.



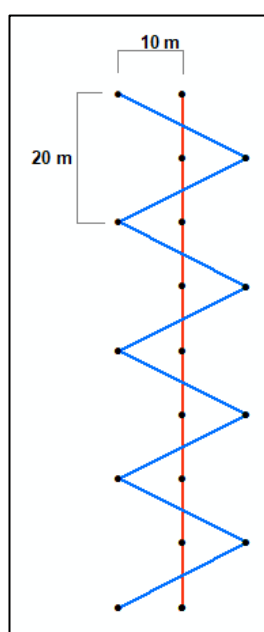
Em relação à feição linear simulada nº 8, nesta foi replicada um erro pontual que ocasionou em certa região, um afastamento de 30 metros entre as linhas, e permaneceu por uma extensão de 40 metros, deixando a linha de teste com 6 vértices e 100 metros de comprimento (Figura 23). Além disso, a discrepância média entre as linhas é de 23,3 metros.

Figura 23 - Feições lineares simuladas nº 8. Linha de referência em vermelho e linha de teste em azul.



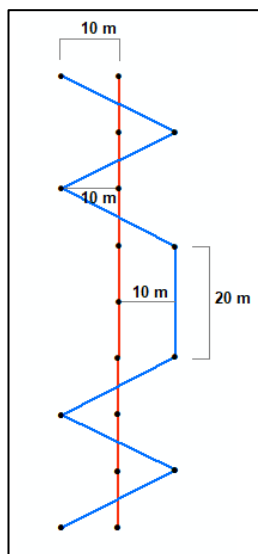
A feição linear simulada nº 9 simula um erro do tipo zigue-zague entre as linhas de teste e referência, onde os vértices de cada linha foram afastados em 10 metros. Além disso, a linha teste possui 178,9 metros de comprimento, 9 vértices e discrepância média de 5 metros (Figura 24).

Figura 24 - Feições lineares simuladas nº 9. Linha de referência em vermelho e linha de teste em azul.



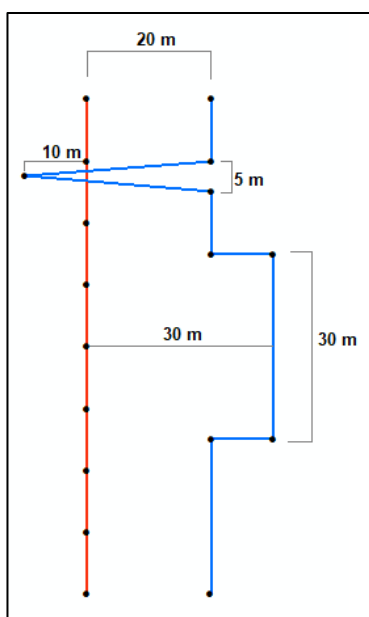
Já a feição linear simulada nº 10 representa um erro do tipo zigue-zague também de 10 metros na linha teste, além de um erro pontual de 20 metros. Nesse caso, ambos os vértices iniciais e finais das linhas estão afastados em 10 metros e a linha de teste possui 8 vértices e 153,9 metros de comprimento (Figura 25), com discrepância média de 5,5 metros.

Figura 25 - Feições lineares simuladas nº 10. Linha de referência em vermelho e linha de teste em azul.



A feição linear simulada nº 11 representou uma situação onde ocorreu a combinação dos erros presentes nos últimos exemplos: erros pontuais e zigue-zague. Dessa forma, a feição linear de teste apresentou um comprimento de 155,2 metros e 9 vértices (Figura 26), com discrepância média de 20 metros.

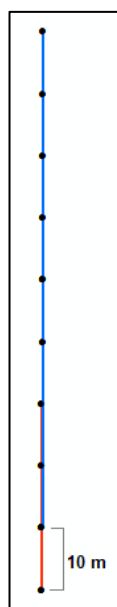
Figura 26 - Feições lineares simuladas nº 11. Linha de referência em vermelho e linha de teste em azul.



Na feição linear simulada nº 12, a feição linear de referência esteve sobreposta à de teste, cuja qual se deslocou em 10 metros. Entretanto, tal deslocamento ocorreu

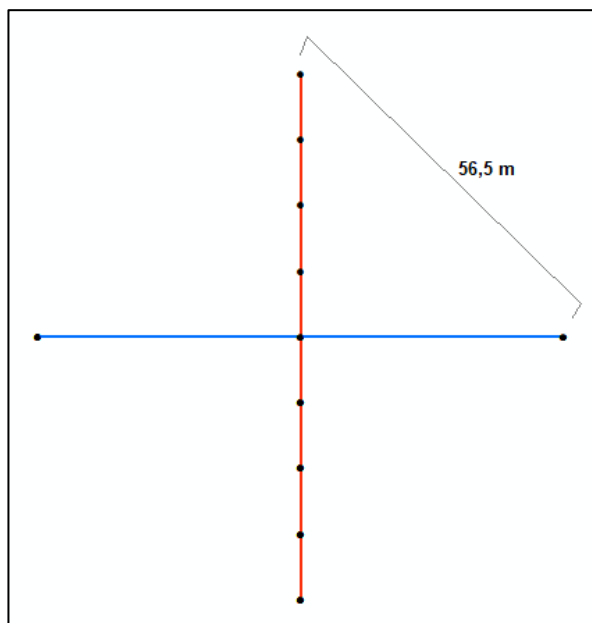
na direção em que as linhas se encontram, ou seja, na vertical (Figura 27). Neste exemplo, ambas as linhas possuem 80 metros de comprimento e a linha teste possui 3 vértices: inicial, final e o central, com discrepância média de 10 metros.

Figura 27 - Feições lineares simuladas nº 12. Linha de referência em vermelho e linha de teste em azul.



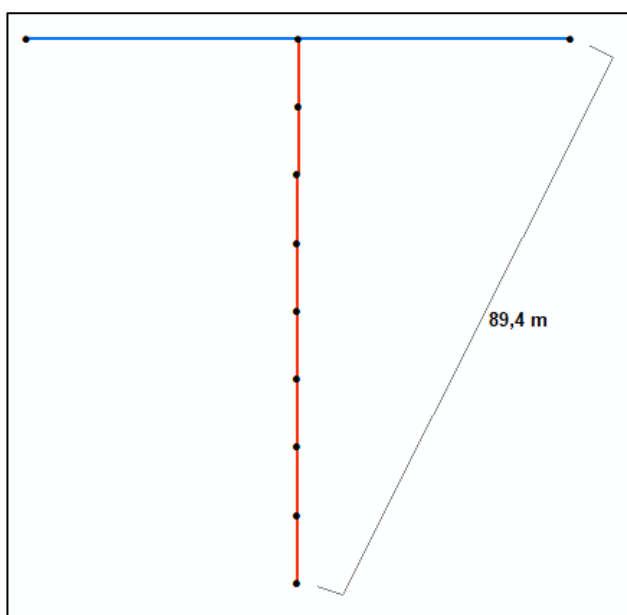
A feição linear simulada nº 13 considerou duas linhas perpendiculares entre si com a interseção ocorrendo exatamente em seus centros, simulando um erro de rotação de 90° (Figura 28). A distância entre os vértices iniciais das linhas é de 56,5 metros, com 3 vértices distribuídos na linha teste, e discrepância média de 20 metros entre as linhas.

Figura 28 – Feições lineares simuladas n° 13. Linha de referência em vermelho e linha de teste em azul.



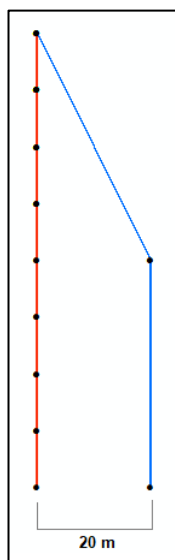
A feição linear simulada n° 14 foram concebidas de forma a garantir a perpendicularidade entre as elas, entretanto a interseção entre as linhas ocorreu exatamente no centro da linha de teste com o vértice inicial da linha de referência, simulando também um erro de rotação de 90°, com discrepância média de 40 metros entre as linhas (Figura 29).

Figura 29 - Feições lineares simuladas n° 14. Linha de referência em vermelho e linha de teste em azul.



A feição linear simulada nº 15 representa um caso em que os vértices iniciais da linha testem e de referência coincidem, entretanto os vértices finais estão afastados em 20 metros. Nesse caso a feição de teste possui 3 vértices e 84,7 metros de comprimento, apresentando uma discrepância média entre as linhas de 15 metros. (Figura 30).

Figura 30 - Feições lineares simuladas nº 15. Linha de referência em vermelho e linha de teste em azul.



5.2.2. Base *MatchingLand*

O conjunto de dados utilizado foi escolhido devido ao fato de serem dados sintéticos que podem simular dados reais. Este conjunto contém 10 pares de linhas elaboradas por Xavier et al. (2017), sendo escolhido dois pares para cada grupo de sinuosidade, além de variar o comprimento. A Figura 31 apresenta o conjunto de linhas escolhidos.

Figura 31 - Conjunto de linhas utilizadas extraídas da base da Espanha.



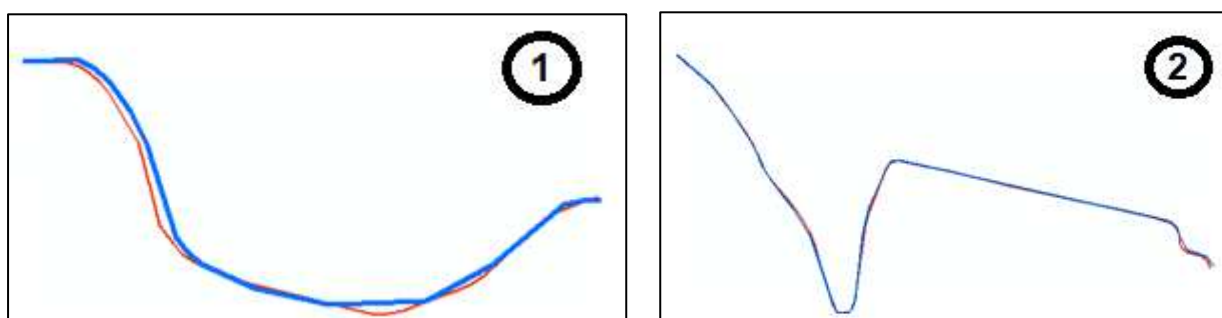
Em relação às informações de cada par de linha, essas podem ser vistas na Tabela 3.

Tabela 3 - Informações dos pares de linhas da base *MatchingLand*.

ID	Classe de morfologia	Comprimento (m)		Número de Vértices	
		Referência	Teste	Referência	Teste
1	Muito Suave	311,7	312,8	30	18
2		1104,2	1101,2	71	39
3	Suave	490,1	508,7	23	25
4		1068,1	1076,1	66	58
5	Suave com direcionalidade estável	734,1	812,1	39	36
6		228,9	233,9	13	6
7	Suave com direcionalidade variável	311,7	312,8	30	18
8		700,1	699,1	55	55
9	Muito sinuosa	1774,9	2162,6	102	181
10		2779,2	3199,8	259	247

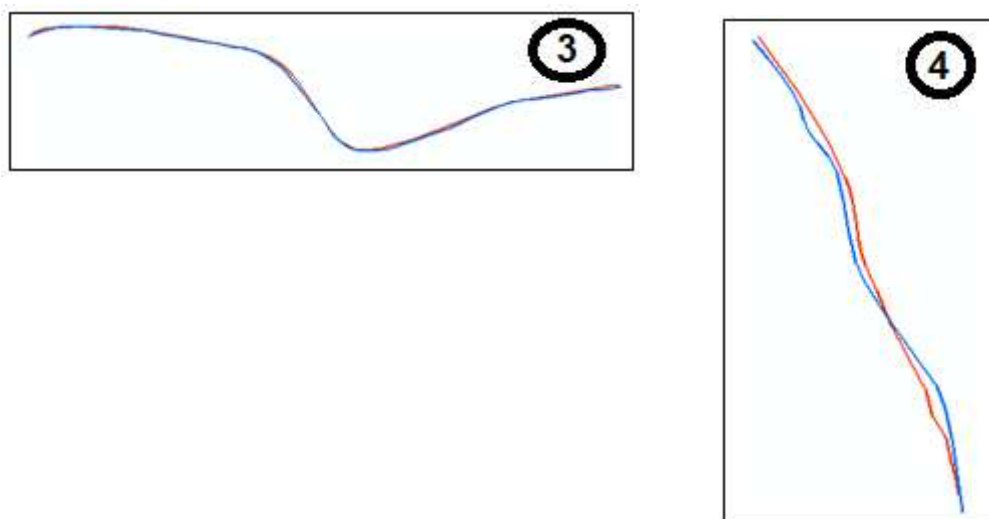
Os dois primeiros pares de linhas, correspondem à linhas muito suaves, que não apresentam muita complexidade, além das linhas teste e de referência terem quase o mesmo comprimento (Figura 32).

Figura 32 - Pares de linhas muito suaves (em vermelho a linha de referência e em azul a linha de teste).



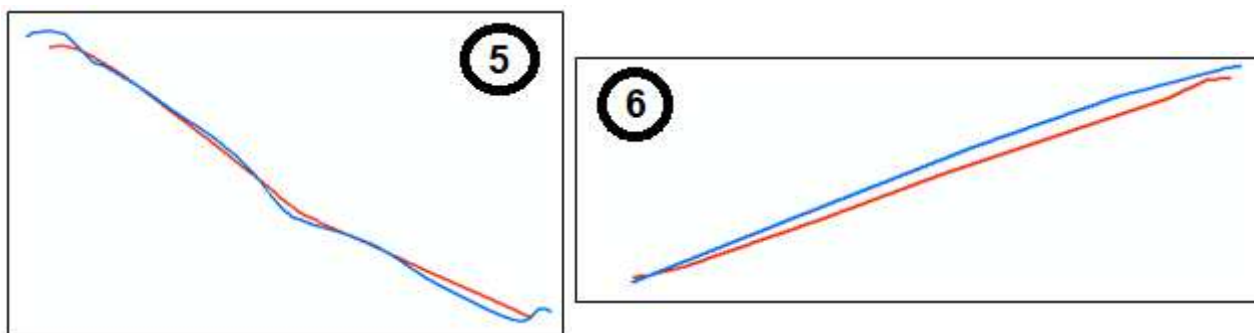
Os pares três e quatro referem à linhas suaves, em que foram escolhidos pares de comprimentos diferentes para analisar as diferenças entre os resultados obtidos (Figura 33).

Figura 33 - Pares de linhas suaves.



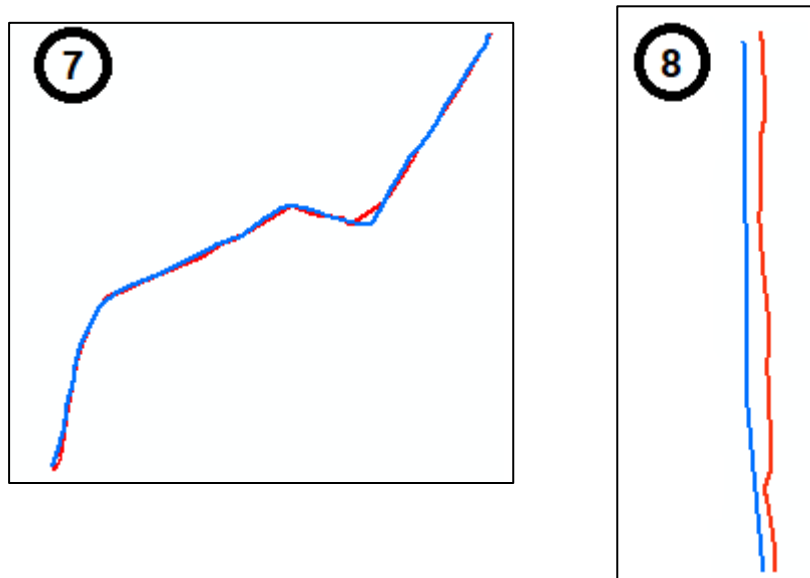
Os pares cinco e seis representam linhas suaves, mas que já apresentam uma certa direcionalidade que é estável (Figura 34).

Figura 34 - Pares de linhas suaves com direcionalidade estável.



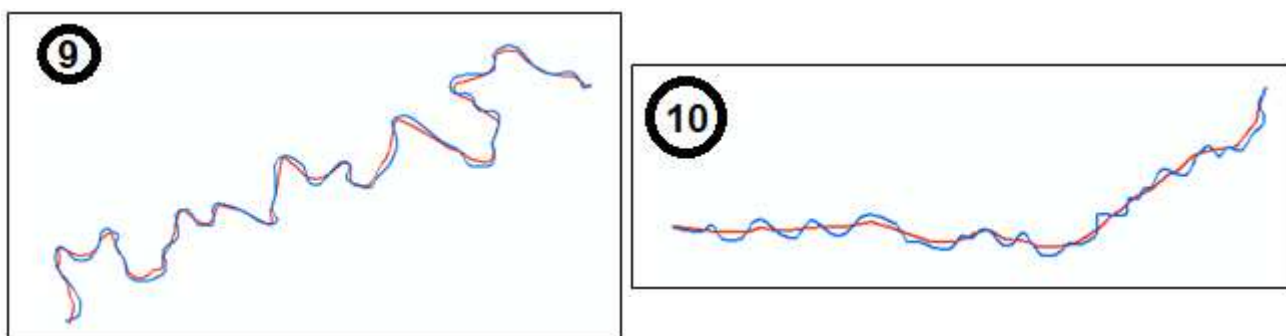
Já os pares sete e oito correspondem à linhas que são suaves e com uma certa direcionalidade, porém essa direcionalidade varia (Figura 35).

Figura 35 - Pares de linhas suaves com direcionalidade variável.



Os dois últimos pares de linhas, são linhas bastante sinuosas, com uma grande quantidade de vértices, além dos comprimentos da linha de referência ter uma diferença com relação ao comprimento da linha de teste (Figura 36).

Figura 36 - Pares de linhas sinuosas.



5.2.3. Dados Reais: Ortoimagem

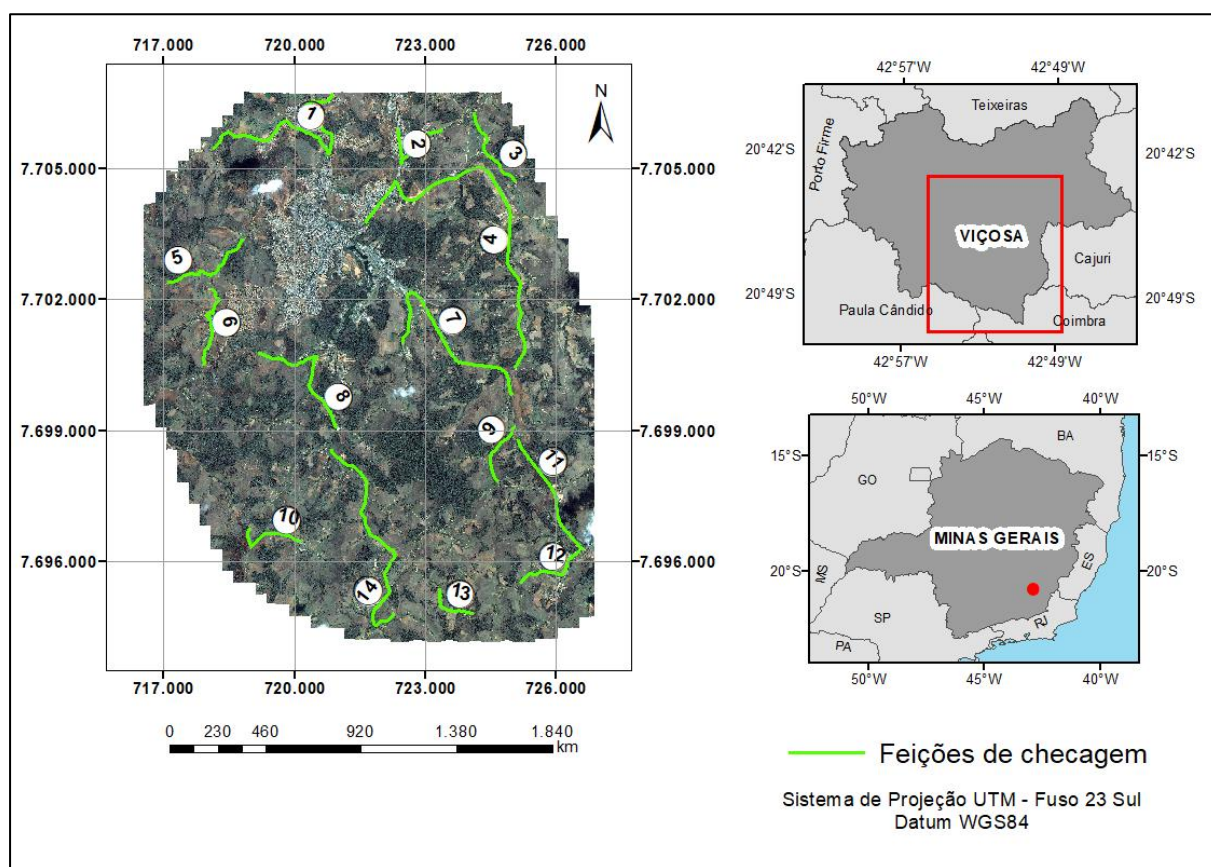
5.2.3.1. Área de Estudo

Visando a aplicação e posterior comparação dos métodos desenvolvidos para a avaliação da acurácia posicional baseados no uso de feições lineares com dados

reais, avaliou-se a acurácia posicional planimétrica de uma ortoimagem Ikonos que encontra-se disponível para estudos no acervo da Universidade Federal de Viçosa.

A área de estudo compreende parte do município de Viçosa, no estado de Minas Gerais, contendo a sub-bacia do ribeirão São Bartolomeu, a qual faz parte da bacia hidrográfica do Rio Doce. Possui uma área aproximada de 134 km², delimitada pelas latitudes 20,73° S e 20,84°S, e longitudes 42,82°W a 42,92°W (Figura 37).

Figura 37 - Área de estudo da pesquisa.



Fonte: Elaborado pelo autor (2021).

Na Tabela 4 estão contidas as informações geométricas de cada linha.

Tabela 4 - Características geométricas do conjunto de dados reais.

ID	COMPRIMENTO (m)		NÚMERO DE VÉRTICES	
	REFERÊNCIA	TESTE	REFERÊNCIA	TESTE
1	5637,2	5640,9	639	754
2	2088,5	2078,0	307	189
3	2090,7	2095,6	232	223
4	8654,2	8642,6	763	472
5	2280,2	2324,5	376	369
6	2131,0	2138,1	451	390
7	4978,1	4974,5	955	364
8	3412,8	3414,7	659	391
9	1562,5	1573,0	338	225
10	1770,1	1757,0	328	166
11	3013,0	3016,9	471	222
12	1806,0	1812,2	473	135
13	1205,9	1207,9	258	63
14	5618,9	5619,9	1044	352

5.2.4. Materiais Utilizados

- Ortoimagem Ikonos com imagens multiespectrais de resolução de 4 metros e imagem pancromática de 1 metro de resolução, obtida por Santos et al. (2015), no datum *World Geodetic System* (WGS) de 1984, na projeção Universal Transversa de Mercator (UTM) zona 23 Sul.
- Base de dados de rodovias levantadas com sistema GNSS da marca Leica modelo GPS 900CS no formato *shapefile* relativo ao município de Viçosa-MG, obtida de Santos et al. (2015), no datum *World Geodetic System* (WGS) de 1984, na projeção Universal Transversa de Mercator (UTM) zona 23 Sul disponibilizado por Santos et al. (2015).

A imagem Ikonos é composta por imagens multiespectrais (resolução espacial de 4 metros) e a imagem pancromática (resolução espacial de 1 metro) que foram

fusionadas. Esta foi submetida a um processo de ortorretificação, através do software PCI Geomatics 10, com modelos matemáticos baseados nas transformações racionais polinomiais, com o uso de RPCs (*Rational Polynomial Coefficients*), sem o emprego de pontos de controle. Na ortorretificação, Santos et al. (2015) utilizaram um MDE proveniente de dados das cartas topográficas na escala de 1:50.000 do mapeamento sistemático brasileiro.

Posteriormente, os autores coletaram 14 estradas, cerca de 46,2 km visando a aplicação e comparação das adaptações desenvolvidas. Para esse levantamento utilizou-se receptores GNSS de dupla frequência através do método relativo cinemático. Em seguida, realizaram o processo de monorestituição das linhas correspondentes aos trechos de rodovias através do *software* ArcGIS.

Através dos trechos de estradas coletados em campo e seus homólogos na ortoimagem, estes foram utilizados para aplicar os métodos propostos de avaliação da acurácia posicional planimétrica com feições lineares.

6. RESULTADOS E DISCUSSÕES

Abaixo serão descritos os resultados obtidos com cada conjunto de dados utilizado nessa pesquisa.

6.1. Dados simulados gerados

A Tabela 5 e a Figura 38 mostram os resultados obtidos aplicando o Influência do Vértice, suas adaptações e os novos métodos. Como trata-se de dados simulados, optou-se por utilizar um valor de densificação de 10 cm.

Vale ressaltar que de acordo com a definição da Distância de *Manhattan* no tópico 4.2.2., esta é uma simplificação da Distância Euclidiana. Sendo assim, a métrica de *Manhattan* sempre é maior ou igual à métrica Euclidiana.

Dessa forma, conforme é esperado, as discrepâncias obtidas com a adaptação do Influência do Vértice considerando a Distância de *Manhattan* foram maiores que os valores encontrados com o método na forma original.

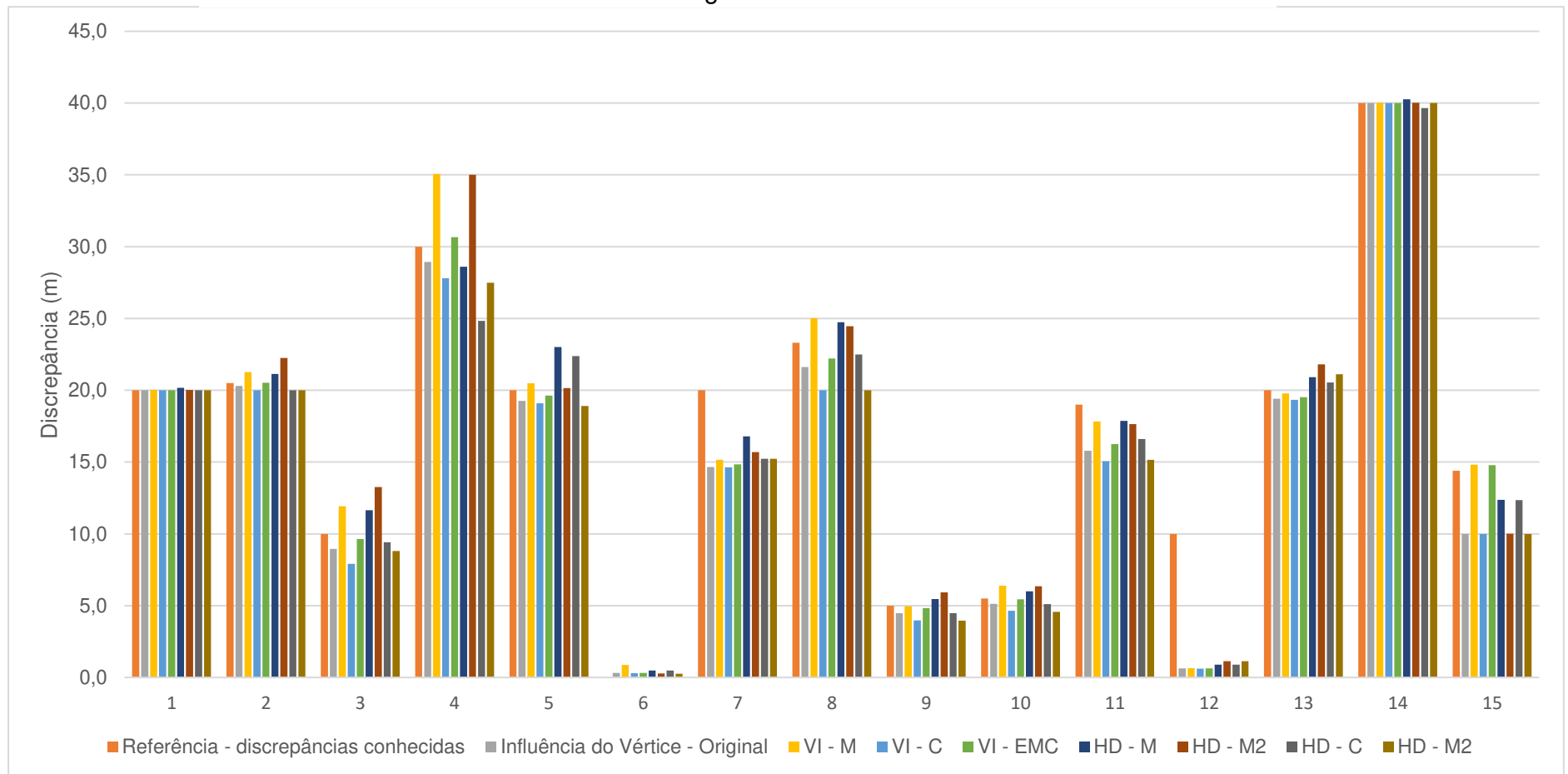
Já em relação à definição da Distância de *Chebyshev* no tópico 4.2.3., esta leva em consideração a maior distância em cada dimensão entre um par de variáveis, o que implica em sempre possuir um valor menor ou igual à Distância Euclidiana.

Sendo assim, em relação ao método VI - C, as discrepâncias foram menores que os valores encontrados com o método em sua forma original, que utiliza a distância Euclidiana. Na adaptação utilizando as três métricas em conjunto (VI – EMC), os valores das discrepâncias de cada par de linhas ficaram entre os valores que foram obtidos com o Influência do Vértice, em sua forma original, e o Influência do Vértice adaptado com a Distância de *Manhattan* (VI – M).

Tabela 5 - Discrepâncias obtidas com os dados simulados gerados.

ID	Referência - discrepâncias conhecidas	Influência do Vértice - Original	Discrepância (m)						
			VI - M	VI - C	VI - EMC	HD - M	HD - M2	HD - C	HD - M2
1	20,0	20,000	20,021	20,000	20,000	20,167	20,021	20,000	20,000
2	20,5	20,295	21,273	20,000	20,520	21,127	22,242	20,000	20,000
3	10,0	8,944	11,923	7,920	9,648	11,641	13,258	9,408	8,805
4	30,0	28,945	35,058	27,795	30,659	28,604	35,015	24,831	27,482
5	20,0	19,254	20,486	19,088	19,625	23,007	20,156	22,378	18,912
6	0,0	0,311	0,866	0,302	0,319	0,481	0,277	0,476	0,265
7	20,0	14,662	15,151	14,628	14,846	16,792	15,692	15,225	15,225
8	23,3	21,625	25,015	20,000	22,213	24,733	24,460	22,496	20,002
9	5,0	4,472	4,955	3,966	4,824	5,466	5,922	4,481	3,958
10	5,5	5,135	6,398	4,647	5,441	5,997	6,346	5,103	4,569
11	19,0	15,793	17,837	15,067	16,258	17,863	17,652	16,604	15,147
12	10,0	0,625	0,649	0,608	0,625	0,888	1,134	0,892	1,133
13	20,0	19,409	19,781	19,337	19,515	20,905	21,806	20,549	21,109
14	40,0	40,000	40,029	40,000	40,012	40,264	40,027	39,657	40,000
15	14,4	10,000	14,812	10,000	14,783	12,371	10,012	12,358	10,000

Figura 38 - Gráfico das discrepâncias obtidas com o Influência do Vértice e os métodos sugeridos.



Verifica-se que em todos os casos o valor da discrepância de cada linha foi variável. Sendo assim, visando uma melhor visualização dos resultados obtidos, elaborou-se uma tabela (Tabela 6), que contém as diferenças de discrepâncias obtidas em porcentagem, tendo como base as discrepâncias médias das linhas conhecidas a priori, sendo em vermelho representado a maior diferença em porcentagem e em valor absoluto, e em azul a menor diferença para cada par de linha considerando todos os métodos adotados.

Tabela 6 - Diferença das discrepâncias médias em porcentagem tendo como referência as discrepâncias médias conhecidas. Em vermelho estão as maiores diferenças de cada par de linha simulada e em azul os menores valores.

Diferença em valor absoluto das discrepâncias obtidas tendo como referência as discrepâncias conhecidas de cada par de linha (%)								
ID	Influência do Vértice original	VI - M	VI - C	VI - EMC	HD - M	HD - M2	HD - C	HD - C2
1	0,00	0,11	0,00	0,00	0,84	0,11	0,00	0,00
2	1,00	3,77	2,44	0,10	3,06	8,50	2,44	2,44
3	10,56	19,23	20,80	3,52	16,41	32,58	5,92	11,95
4	3,52	16,86	7,35	2,20	4,65	16,72	17,23	8,39
5	3,73	2,43	4,56	1,88	15,04	0,78	11,89	5,44
6	31,10	86,60	30,20	31,90	48,10	27,70	47,60	26,50
7	26,69	24,25	26,86	25,77	16,04	21,54	23,88	23,88
8	7,19	7,36	14,16	4,67	6,15	4,98	3,45	14,15
9	10,56	0,90	20,68	3,52	9,32	18,44	10,38	20,84
10	6,64	16,33	15,51	1,07	9,04	15,38	7,22	16,93
11	16,88	6,12	20,70	14,43	5,98	7,09	12,61	20,28
12	93,75	93,51	93,92	93,75	91,12	88,66	91,08	88,67
13	2,96	1,10	3,32	2,43	4,53	9,03	2,75	5,55
14	0,00	0,07	0,00	0,03	0,66	0,07	0,86	0,00
15	30,56	2,86	30,56	2,66	14,09	30,47	14,18	30,56

Como pode ser observado, o primeiro par por ser o mais simples de todos os casos, não houve uma grande diferença entre os valores obtidos.

Já em relação ao segundo par, o método HD – M2 apresentou a maior diferença (8,50%). Quando utilizou-se o HD - M, essa diferença foi de apenas 3,06%. Entretanto, a adaptação do Influência do Vértice com as três métricas (VI – EMC) foi mais eficiente nesse caso, já que a diferença foi de apenas 0,10%.

O terceiro par também teve uma maior diferença em relação a um dos métodos sugeridos (32,58%), e novamente a menor diferença foi em relação ao VI - EMC (3,52%). O mesmo ocorreu com o quarto par, onde a maior diferença de discrepância foi em relação ao método HD – C (17,23%).

Em relação ao quinto par a maior diferença chegou à 15,04% com o método HD - M. Já utilizando o método HD – M2, obteve-se a menor diferença correspondente a 0,72%.

O par 6 foi o que apresentou o maior valor de diferença entre as discrepâncias através do método VI - M (86,60%). Ao utilizar o método proposto HD – C2, obteve-se a menor diferença de 26,50%. Vale ressaltar que para o par de linha 6, todos os métodos apresentaram alta diferença em relação à referência, o que demonstra que não conseguiram detectar corretamente o *outlier* presente na linha de teste.

Já os pares 7 e 15, as maiores diferenças encontradas foram em relação ao método VI - C, sendo 26,86% e 30,56% respectivamente. Já o método HD – M foi responsável pela menor diferença das discrepâncias em relação à linhas 7 (16,04%), e já em relação ao par 15 essa menor diferença foi encontrada com o método VI - EMC (2,66%).

Ao utilizar o Influência do Vértice adaptado com a distância de *Chebyshev* (VI – C), este possibilitou obter a maior diferença em relação ao par de linhas 14 (0,86%), e a menor em relação ao par 8 (3,45%). Entretanto, com a adaptação do Influência do Vértice com as três métricas, obteve-se as menores diferenças de discrepâncias para as linhas 10 e 15 de 1,07% e 2,66% respectivamente.

A menor diferença entre as discrepâncias para o par de linha 11 foi através do novo método HD – M, equivalente a 5,98%. Já ao utilizar o método HD – M2, esse

forneceu a menor diferença para o par de linhas 12 (88,66%), e a maior diferença para o par de linhas 13 (9,03%). Em relação ao par de linhas 12, este apresentou um alto valor de discrepância em todos os métodos.

O método HD – C2 foi o que apresentou as maiores diferenças das discrepâncias para os pares 8, 10 e 15 de 14,15%, 16,93% e 30,56% respectivamente.

Tendo em vista os valores de diferenças de discrepâncias em valor absoluto, foi possível verificar que os métodos em que foram feitas adaptações do Influência do Vértice possibilitou obter diferenças menores entre as discrepâncias em comparação com os métodos adaptados a partir da Distância de *Hausdorff*, sendo mais eficientes. Em relação aos métodos sugeridos utilizando a distância de *Manhattan*, tanto em relação à adaptação do Influência do Vértice e a Distância de *Hausdorff*, estes foram responsáveis pelas maiores diferenças obtidas, apresentando uma maior magnitude na determinação da discrepância posicional em comparação com o método de referência de modo geral.

Analisando os resultados com os novos métodos sugeridos utilizando a distância de *Chebyshev*, percebeu-se que os resultados foram em geral menos discrepantes em comparação com os novos métodos que utilizam a distância de *Manhattan*.

Já a adaptação do Influência do Vértice com as três métricas, em geral foi o que se mostrou mais satisfatório e eficiente, visto que as diferenças entre as discrepâncias foram bem menores em comparação com os demais métodos.

Ao aplicar o teste estatístico da ANOVA, foram atendidos todos os requisitos para a aplicação do teste. No teste de Bartlett o p-valor (0,1) foi maior que 0,05, atestando a homogeneidade das variâncias. Além disso, o p-valor do teste do Shapiro-Wilk foi maior que 0,05 para todos os métodos, o que confirma a normalidade dos dados.

As hipóteses testadas foram: H_0 : os valores esperados de discrepância posicional dos métodos são similares e H_a : pelo menos um método é distinto dos demais. Como o p-valor estimado foi igual à 0,999, sendo maior que 0,05, então isto

indica que os métodos são todos similares, não apresentando diferenças significativas entre eles.

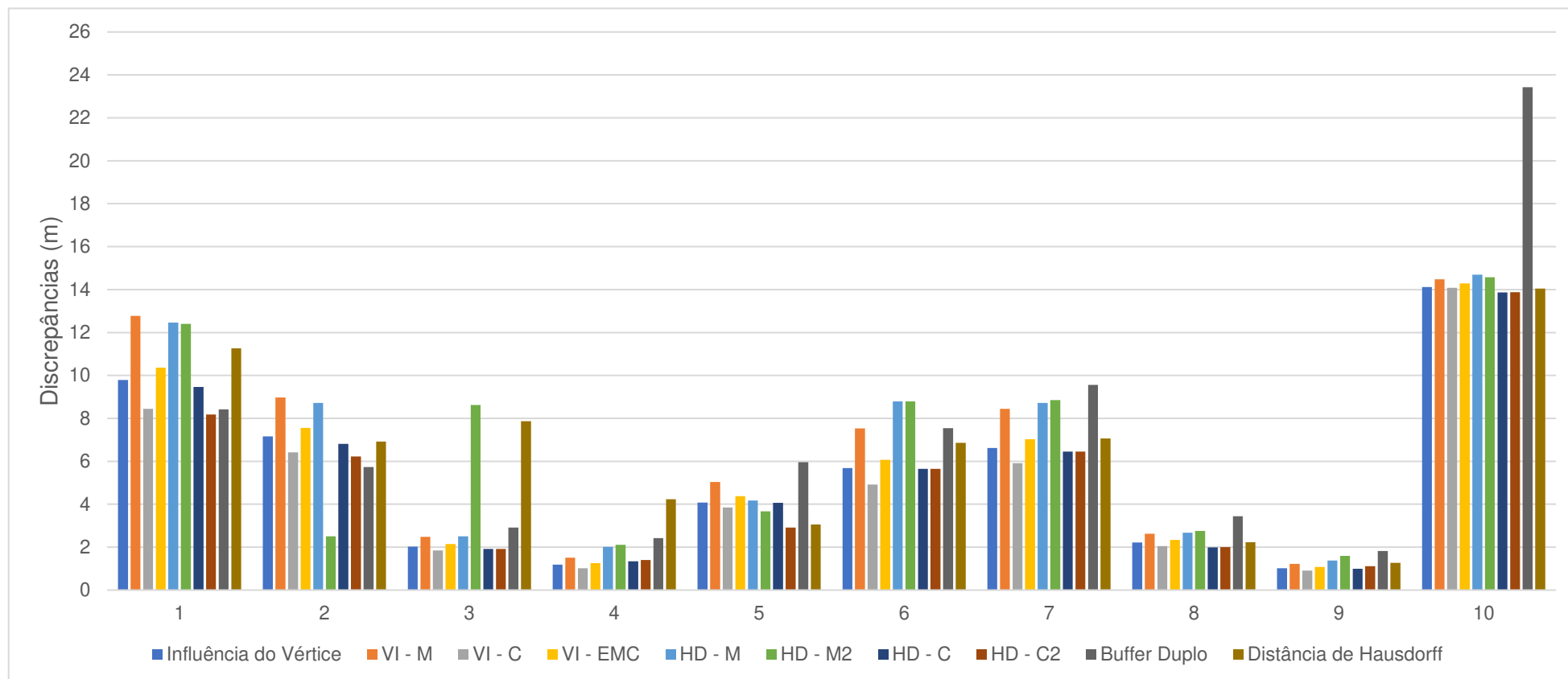
6.2. Dados simulados Espanha

A Tabela 7 e a Figura 39 mostram os resultados obtidos aplicando o Influência do Vértice e suas adaptações, a Distância de *Hausdorff* e suas adaptações, e o *Buffer* Duplo nos 10 pares de linhas considerados. O valor de densificação utilizado conforme a equação empírica desenvolvida foi de 1 cm, já que a escala é de 1: 25.000.

Tabela 7 - Discrepâncias obtidas com os dados simulados da Espanha.

ID	Discrepância (m)									
	Influência do Vértice	Buffer Duplo	Distância de Hausdorff	VI - M	VI - C	VI - EMC	HD - M	HD - M2	HD - C	HD - C2
1	9,785	8,416	11,259	12,772	8,446	10,366	12,467	12,401	9,463	8,177
2	7,160	5,734	6,924	8,975	6,412	7,557	8,719	2,505	6,810	6,228
3	2,024	2,918	7,863	2,477	1,844	2,141	2,511	8,627	1,914	1,915
4	1,187	2,422	4,229	1,515	1,012	1,261	2,009	2,115	1,345	1,399
5	4,082	5,958	3,055	5,032	3,853	4,373	4,178	3,666	4,061	2,912
6	5,679	7,548	6,865	7,533	4,916	6,070	8,793	8,789	5,646	5,646
7	6,616	9,565	7,062	8,448	5,918	7,033	8,714	8,846	6,457	6,457
8	2,217	3,438	2,231	2,624	2,053	2,337	2,672	2,755	1,993	2,000
9	1,012	1,824	1,266	1,219	0,913	1,078	1,374	1,595	0,989	1,113
10	14,114	23,429	14,050	14,480	14,085	14,287	14,689	14,572	13,862	13,876

Figura 39 - Gráfico das discrepâncias utilizando o Influência do Vértice, os métodos sugeridos, Distância de *Hausdorff* , as adaptações propostas e o *Buffer Duplo* para os dados da Espanha.



Como trata-se de um conjunto de linhas que foram modificadas por Xavier et al. (2017), essas podem ser consideradas uma simulação de um caso real. Sendo assim, foram utilizados os métodos já existentes *Buffer* duplo e Distância de *Hausdorff* para fins comparativos.

No caso anterior como eram linhas apenas criadas para simular erros sistemáticos mais comuns visando verificar a viabilidade dos novos métodos, verificou-se que não era necessário utilizar esses outros dois métodos.

Visando analisar os resultados em função das diferenças obtidas, tendo como referência o Influência do Vértice, elaborou-se a Tabela 8 que contém as diferenças de discrepâncias obtidas em porcentagem em valor absoluto. Em vermelho refere-se à maior discrepância obtida em cada par de linha, e em azul a menor discrepância.

Tabela 8 - Diferença das discrepâncias em porcentagem tendo como referência o Influência do Vértice para os dados da Espanha em valor absoluto.

Diferença das discrepâncias obtidas em valor absoluto tendo como referência o Influência do Vértice (%)									
ID	Buffer duplo	Distância de Hausdorff	VI - M	VI - C	VI - EMC	HD - M	HD - M2	HD - C	HD - C2
1	13,99	15,06	30,53	13,68	5,94	27,41	26,73	3,29	16,43
2	19,92	3,30	25,35	10,45	5,54	21,77	65,01	4,89	13,02
3	44,17	288,49	22,38	8,89	5,78	24,06	326,24	5,43	5,39
4	104,04	256,28	27,63	14,74	6,23	69,25	78,18	13,31	17,86
5	45,96	25,16	23,27	5,61	7,13	2,35	10,19	0,51	28,66
6	32,91	20,88	32,65	13,44	6,89	54,83	54,76	0,58	0,58
7	44,57	6,74	27,69	10,55	6,30	31,71	33,71	2,40	2,40
8	55,07	0,63	18,36	7,40	5,41	20,52	24,27	10,10	9,79
9	80,24	25,10	20,45	9,78	6,52	35,77	57,61	2,27	9,98
10	66,00	0,45	2,59	0,21	1,23	4,07	3,25	1,79	1,69

Os dois primeiros pares de linhas referem-se a linhas muito suaves, porém diferente do primeiro par de linhas simuladas criadas, como essas são linhas utilizadas para simular um caso real, houve uma diferença significativa entre os métodos. Sendo que essa diferença em valor absoluto foi maior no segundo par de linha, visto que essa é mais complexa que a primeira, apresentando um comprimento muito maior, bem como a quantidade de vértices.

Nesse caso, a menor diferença entre as discrepâncias (3,29%) para o primeiro par de linha encontrada foi considerando o novo método HD - C, O segundo par de linha a diferença entre as discrepâncias foi praticamente o mesmo, só que foi obtida utilizando a Distância de *Hausdorff*.

O terceiro e o quarto par de linhas, referem-se à linhas suaves, sendo estas possuindo as maiores diferenças de discrepâncias em comparação com as outras linhas, 326,24% e 256,28% respectivamente. Apesar de a maior diferença de discrepância do terceiro par de linha ter sido no método HD – M2, a Distância de *Hausdorff* também forneceu uma diferença alta de 288,49%.

As menores discrepâncias nesse caso, foi através do novo método HD - C, 5,39%, e a adaptação do Influência do Vértice com ambas as distâncias de *Manhattan* e *Chebyshev* (VI – EMC), com valor de 6,23%.

A quinta e sexta linha, são suaves com direcionalidade estável, sendo que em ambas, as menores diferenças de discrepâncias foram obtidas através do método HD – C2, 0,51 e 0,58% respectivamente. Já as maiores diferenças foram encontradas no *Buffer Duplo*, 45,96% para o par de linha 5, e 54,83% no novo método HD - M para o par de linha 6.

Já os últimos 4 pares de linhas que são mais complexas que as anteriores, as maiores diferenças entre as discrepâncias foram no *Buffer Duplo*, e as menores foram encontradas em métodos relacionados à distância de *Chebyshev*, com exceção do par de linha oito, o qual foi na distância de *Hausdorff*, 0,63%.

Vale ressaltar que as diferenças encontradas entre os métodos são esperadas, visto que cada um tem sua particularidade, definição e incertezas.

Com o objetivo de realizar uma análise mais aprofundada, foi elaborada uma tabela com as estatísticas descritivas dos dados e realizada a classificação do conjunto de dados considerando cada método, de acordo com os requisitos do Decreto nº 89.817/ET-CQDG, adotando a escala dos dados de teste 1:25.000. Tais resultados estão contidos na Tabela 9.

Tabela 9 - Estatísticas Descritivas da amostra de dados da Espanha e a classificação.

MÉTODOS	Estatísticas Descritivas					Classificação			
	Média (m)	Mínimo (m)	Máximo (m)	Desvio Padrão (m)	Variância (m ²)	RMS (m)	PEC (m)	EP (m)	CLASSIFICAÇÃO
Influência do Vértice	5,388	1,012	14,114	4,217	17,787	7,074	12,5	7,5	Classe B
<i>Buffer Duplo</i>	7,125	1,824	23,429	6,311	39,825	7,921	20	12,5	Classe C
Distância de <i>Hausdorff</i>	6,48	1,266	14,05	4,011	16,086	9,81	20	12,5	Classe C
VI - M	6,508	1,219	14,48	4,715	22,234	8,324	20	12,5	Classe C
VI - C	4,945	0,913	14,085	4,08	16,644	6,619	12,5	7,5	Classe B
VI - EMC	5,65	1,078	14,287	4,316	18,628	7,355	12,5	7,5	Classe B
HD - M	6,613	1,374	14,689	4,714	22,225	8,414	20	12,5	Classe C
HD – M2	6,587	1,592	14,572	4,672	21,83	8,369	20	12,5	Classe C
HD - C	5,254	0,989	13,862	4,117	16,947	6,901	12,5	7,5	Classe B
HD – C2	4,972	1,113	13,876	3,998	15,986	6,592	12,5	7,5	Classe B

Conforme os resultados acima, verifica-se que todos os métodos que incluem a distância de *Chebyshev* possuem menor magnitude na determinação das discrepâncias em comparação com as que incluem a distância de *Manhattan*.

Em relação às adaptações propostas nos métodos com a distância de *Chebyshev*, esses também possuem menor magnitude na determinação das discrepâncias até mesmo que o Influência do Vértice, o que pode ser justificado pelo fato da distância de *Chebyshev* sempre ser igual ou menor que a distância Euclidiana. O mesmo pode ser dito em relação à distância de *Manhattan*, a qual sempre é maior ou igual à distância Euclidiana, possuindo então maior magnitude.

Dentre os métodos utilizados, o *Buffer Duplo* apresentou o maior valor de RMS, 9,810 m, sendo seguido pelos métodos que utilizam a distância de *Manhattan*. Além disso, como pode ser visto na Tabela 9, em geral os dados foram classificados na Classe B e C em função das características do método empregado.

Vale ressaltar que o *Buffer Duplo* foi o método que apresentou os maiores valores das estatísticas descritivas da Tabela 9, enquanto o VI - C apresentou o menor valor de média (2,024 m) e valor mínimo (0,812 m), o que já era esperado. Já os novos métodos sugeridos da distância de *Chebyshev* (HD – C e HD - C2) proporcionaram os menores valores de máximo, desvio padrão e consequentemente a variância.

Neste caso também foi aplicado o teste estatístico da ANOVA, sendo atendidos todos os requisitos para a aplicação do teste. Visto que o p-valor (0,9463) do teste de Bartlett foi maior que 0,05, foi atestado a homogeneidade das variâncias. Além disso, o p-valor do teste do Shapiro-Wilk foi maior que 0,05 para todos os métodos, o que confirma a normalidade dos dados.

As hipóteses testadas foram: H_0 : os valores esperados de discrepância posicional dos métodos são similares e H_a : pelo menos um método é distinto dos demais. Como o p-valor estimado foi igual à 0,972, sendo maior que 0,05, então isto indica que os métodos são todos similares, não apresentando diferenças significativas entre eles, conforme anteriormente.

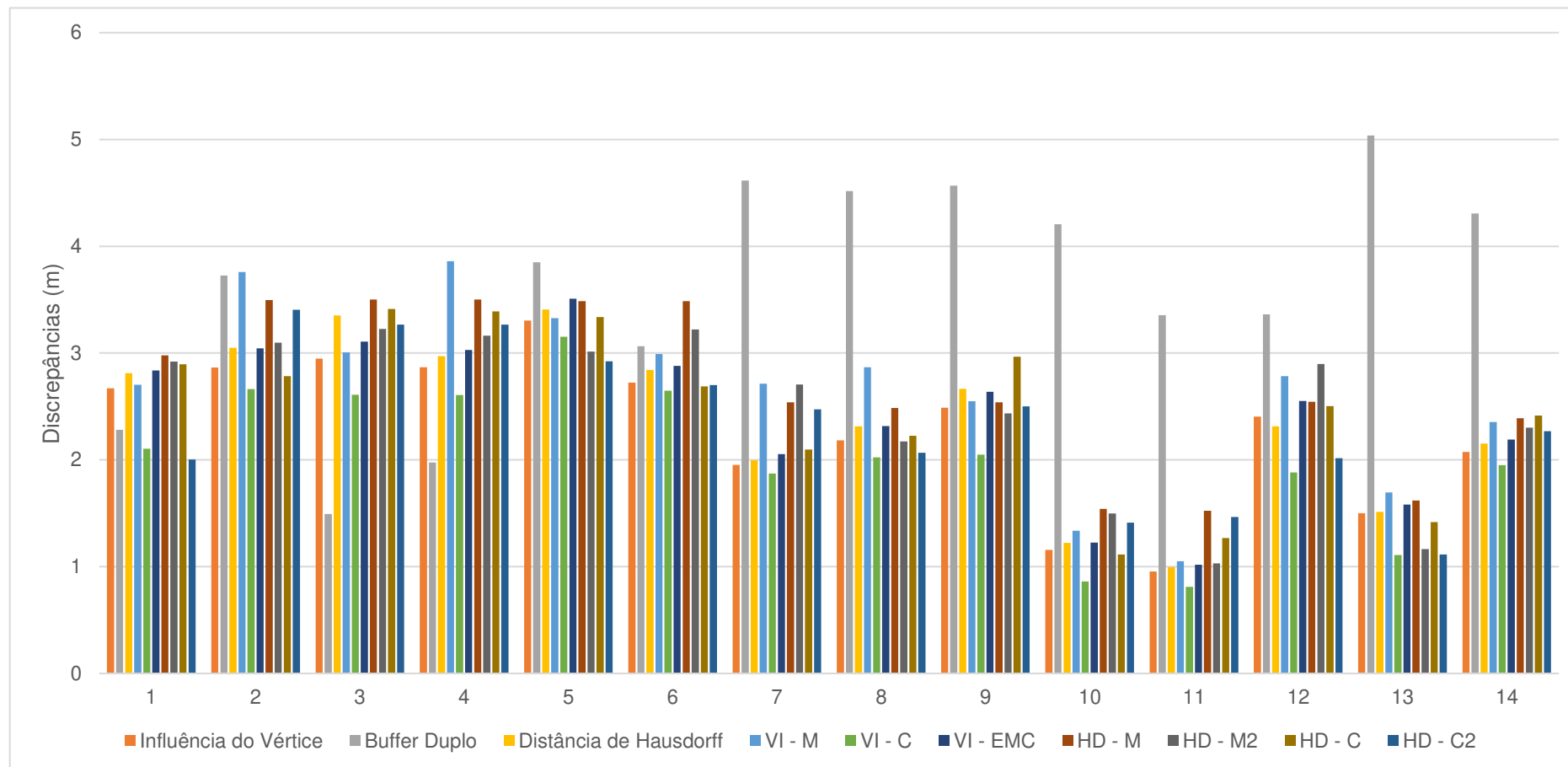
6.3. Dados Reais

A Tabela 10 e a Figura 40 mostram os resultados obtidos aplicando o Influência do Vértice, suas adaptações, O *Buffer* Duplo, a Distância de *Hausdorff* e os novos métodos nos 14 pares de linhas considerados. O valor de densificação utilizado conforme a equação empírica desenvolvida foi de 0,4 cm, já que a escala é de 1: 10.000.

Tabela 10 - Discrepâncias do conjunto de dados em cada método para os dados reais.

ID	Discrepância (m)									
	Influência do Vértice	Buffer Duplo	Distância de Hausdorff	VI - M	VI - C	VI - EMC	HD - M	HD - M2	HD - C	HD - C2
1	2,671	2,280	2,812	2,703	2,105	2,837	2,979	2,921	2,895	2,002
2	2,864	3,725	3,048	3,759	2,662	3,045	3,497	3,096	2,784	3,404
3	2,949	1,494	3,351	3,007	2,609	3,107	3,502	3,226	3,413	3,266
4	2,866	1,976	2,970	3,861	2,608	3,029	3,500	3,163	3,389	3,265
5	3,305	3,849	3,408	3,327	3,152	3,509	3,487	3,013	3,338	2,922
6	2,724	3,065	2,843	2,991	2,647	2,880	3,486	3,222	2,689	2,701
7	1,952	4,616	1,995	2,713	1,871	2,053	2,538	2,706	2,096	2,473
8	2,183	4,517	2,313	2,868	2,023	2,316	2,485	2,172	2,225	2,067
9	2,487	4,567	2,665	2,549	2,049	2,637	2,539	2,435	2,965	2,500
10	1,158	4,205	1,223	1,336	0,861	1,225	1,541	1,498	1,113	1,411
11	0,956	3,354	0,995	1,052	0,812	1,017	1,522	1,031	1,269	1,465
12	2,405	3,362	2,313	2,783	1,881	2,551	2,543	2,898	2,503	2,016
13	1,501	5,037	1,514	1,694	1,109	1,582	1,62	1,165	1,416	1,113
14	2,075	4,307	2,153	2,354	1,949	2,191	2,389	2,302	2,414	2,269

Figura 40 - Gráfico das discrepâncias utilizando diversos métodos para os dados reais.



Observa-se que em geral, o *Buffer Duplo* foi o método em que as discrepâncias foram maiores, demonstrando ser mais restritivo, enquanto os métodos que utilizam a distância de *Chebyshev* em geral apresentou as menores discrepâncias, bem como no caso dos dados da Espanha.

Os métodos incluindo a distância de *Manhattan* também apresentaram altos valores de discrepância em geral em comparação com as demais, confirmando que os métodos com a distância de *Manhattan* possuem maior magnitude na determinação das discrepâncias do que as que utilizam a distância de *Chebyshev*.

Já o método que utiliza em conjunto ambas as distâncias (VI – EMC), é um meio termo, onde as discrepâncias oscilam de um valor mais alto para um valor mais baixo dependendo do par de linha. Além disso, a Distância de *Hausdorff* demonstrou ser mais restritiva em geral em comparação com o Influência do Vértice.

Para uma análise mais aprofundada, conforme nos casos anteriores, elaborou-se a Tabela 11 que contém as diferenças de discrepâncias em valor absoluto obtidas em porcentagem. Em vermelho a maior discrepância obtida em cada par de linha, e em azul a menor discrepância.

Tabela 11 - Diferença em porcentagem das discrepâncias obtidas para os dados reais.

Diferença das discrepâncias em valor absoluto obtidas tendo como referência o Influência do Vértice (%)									
ID	Buffer duplo	Distância de Hausdorff	VI - M	VI - C	VI - EMC	HD - M	HD - M2	HD - C	HD - C2
1	14,64	5,28	1,20	21,19	6,21	11,53	9,36	8,39	25,05
2	30,06	6,42	31,26	7,05	6,32	22,10	8,10	2,79	18,85
3	49,34	13,63	1,97	11,53	5,36	18,75	9,39	15,73	10,75
4	31,05	3,63	34,72	9,00	5,69	22,12	10,36	18,25	13,92
5	16,46	3,12	0,67	4,63	6,17	5,51	8,84	1,00	11,59
6	12,52	4,37	9,80	2,83	5,73	27,97	18,28	1,28	0,84
7	136,48	2,20	38,99	4,15	5,17	30,02	38,63	7,38	26,69
8	106,92	5,96	31,38	7,33	6,09	13,83	0,50	1,92	5,31
9	83,63	7,16	2,49	17,61	6,03	2,09	2,09	19,22	0,52
10	263,13	5,61	15,37	25,65	5,79	33,07	29,36	3,89	21,85
11	250,84	4,08	10,04	15,06	6,38	59,21	7,85	32,74	53,24
12	39,79	3,83	15,72	21,79	6,07	5,74	20,50	4,07	16,17
13	235,58	0,87	12,86	26,12	5,40	7,93	22,39	5,66	25,85
14	107,57	3,76	13,45	6,07	5,59	15,13	10,94	16,34	9,35

Conforme analisado anteriormente, dos 14 pares de linhas, em 11 pares a maior discrepância foi obtida utilizando o *Buffer* Duplo, o que acarretou consequentemente nas maiores diferenças em porcentagem tendo como referência o método Influência do Vértice.

No par de linha um, a maior diferença foi em relação ao novo método HD – C2, enquanto os pares 6 e 9 com esse método as diferenças foram as menores, 0,84% e 0,52% respectivamente. Vale ressaltar que a maior diferença de discrepância do par 6, foi em relação ao novo método HD - M de 27,97%.

Em geral a Distância de *Hausdorff* apresentou as menores diferenças entre as discrepâncias. Já os pares 2 e 10, as menores diferenças foram com o novo método HD - C, sendo 2,79% e 3,89% respectivamente, enquanto os pares 1, 3 e 5 as menores diferenças foram em relação à adaptação do Influência do Vértice utilizando a distância de *Manhattan* (VI- M), 1,20% 1,97% 0,67% respectivamente. Entretanto, com o par de linha 2, a sua maior diferença de discrepância (31,26%) foi encontrada utilizando esse método.

O par de linha 8 foi o único em que a menor diferença foi encontrada utilizando o novo método HD – M2 (0,50%).

Nesse caso, também foi elaborada uma tabela com as estatísticas descritivas dos dados e realizada a classificação do conjunto de dados considerando cada método, de acordo com os requisitos do Decreto nº 89.817/ET-CQDG, adotando a escala dos dados de teste 1:10.000. Tais resultados estão contidos na Tabela 12.

Tabela 12 - Estatísticas descritivas da amostragem de dados reais e a classificação.

MÉTODOS	Estatísticas Descritivas					Classificação			
	Média (m)	Mínimo (m)	Máximo (m)	Desvio Padrão (m)	Variância (m²)	RMS (m)	PEC (m)	EP (m)	CLASSIFICAÇÃO
Influência do Vértice	2,293	0,956	3,305	0,700	0,490	2,479	5	3	Classe B
Buffer Duplo	3,597	1,494	5,037	1,075	1,157	2,604	5	3	Classe B
Distância de <i>Hausdorff</i>	2,400	0,995	3,408	0,759	0,576	3,884	8	5	Classe C
VI - M	2,643	1,052	3,861	0,821	0,674	2,863	5	3	Classe B
VI - C	2,024	0,812	3,152	0,705	0,497	2,216	5	3	Classe B
VI - EMC	2,427	1,017	3,509	0,742	0,551	2,626	5	3	Classe B
HD - M	2,688	1,522	3,502	0,752	0,565	2,889	5	3	Classe B
HD – M2	2,489	1,031	3,226	0,763	0,583	2,694	5	3	Classe B
HD - C	2,465	1,113	3,413	0,768	0,589	2,671	5	3	Classe B
HD – C2	2,348	1,113	3,404	0,723	0,522	2,542	5	3	Classe B

Conforme os resultados, verifica-se que em todas os métodos, a média das discrepâncias ficou na casa dos dois metros, com exceção do *Buffer Duplo* que chegou a mais de 3,5 metros, além de apresentar também o maior valor de máximo (5,037 m), desvio padrão (1,075 m) e variância (1,157 m²).

Isso confirma que esse método possui uma maior magnitude na determinação das discrepâncias que as demais. Além disso, na classificação de acordo com o Decreto nº 89.817/ET-CQDG foi o único método em que a classificação foi obtida na classe C, devido ao seu RMS de 3,884 metros.

Seguido pelo *Buffer Duplo*, a adaptação do Influência do Vértice com a distância de *Manhattan* (VI – M), apresentou os maiores valores de máximo (3,861 m), desvio padrão (0,821 m) e variância (0,674 m²), mostrando também ter uma maior magnitude na determinação das discrepâncias.

Já em relação às estatísticas descritivas da Distância de *Hausdorff* e da adaptação do Influência do Vértice com a distância de *Manhattan* e a distância de *Chebyshev* (VI – EMC), os valores foram intermediários em relação aos demais métodos.

Ao aplicar o teste estatístico da ANOVA, foram atendidos todos os requisitos para a aplicação do teste. O p-valor (0,9046) do teste de Bartlett foi maior que 0,05, atestando a homogeneidade das variâncias. Além disso, o p-valor do teste do Shapiro-Wilk foi maior que 0,05 para todos os métodos, o que confirma a normalidade dos dados.

As hipóteses testadas foram: Ho: os valores esperados de discrepância posicional dos métodos são similares e Ha: pelo menos um método é distinto dos demais. Como o p-valor estimado foi igual à 0,000211, sendo menor que 0,05, então isto indicou que os métodos não são todos similares, sendo verificado que tratava-se do *Buffer Duplo*, visto que ao realizar a comparação de método com método, o p-valor deste com qualquer um outro método era menor que 0,05.

7. CONCLUSÕES E RECOMENDAÇÕES

Em suma, este estudo apresentou a adaptação do método Influência do Vértice e do método Distância de *Hausdorff* utilizando as métricas da distância de *Manhattan* e a distância de *Chebyshev* no processo de avaliação da acurácia posicional utilizando feições lineares.

Diante dos resultados encontrados nos experimentos realizados, percebe-se que com a metodologia desenvolvida neste trabalho, o uso das métricas de distância de *Manhattan* e *Chebyshev* não apresentaram viabilidade para o uso na estimação da discrepância posicional no controle de qualidade cartográfica.

Vale ressaltar que as adaptações dos métodos têm uma certa sensibilidade em relação à quantidade de vértices presentes nas feições lineares avaliadas, bem como em relação à sua sinuosidade. No experimento com os dados da Espanha e com os dados reais, foi utilizado o método *Buffer Duplo* para fins de comparação, o que permitiu verificar que o *Buffer Duplo* em geral tem uma maior magnitude na discrepância posicional estimada em relação aos demais métodos. Já o Distância de *Hausdorff*, a sua magnitude mostrou ser menor que a do *Buffer Duplo* por meio dos resultados obtidos.

Em geral, verificou-se que dentre as adaptações dos métodos propostos, aqueles que utilizam a distância de *Manhattan* possuem uma maior magnitude na determinação da acurácia posicional, devido ao fato dessa sempre ser maior ou igual à distância Euclidiana, sendo esta utilizada como base no Influência do Vértice.

Já as adaptações com a distância de *Chebyshev* possuem menor magnitude na determinação da discrepância posicional, já que essa é sempre menor ou igual que a distância Euclidiana. Em relação à adaptação do Influência do Vértice com ambas as distâncias em conjunto, *Manhattan* e *Chebyshev* (VI – EMC), esta apresentou resultados intermediários entre os métodos que utilizam essas distâncias de forma separada.

Em relação às adaptações, optou-se por testar quando se calcula somente a distância dos vértices da linha de referência para os vértices da linha de teste, bem como quando é calculada ambas as distâncias. Pôde-se notar que ao calcular tanto a

distância dos vértices da linha de referência para a linha de teste e a distância dos vértices da linha de teste para a linha de referência, o resultado em geral apresentou uma menor magnitude na determinação da discrepância posicional, do que quando calculou-se somente a distância dos vértices da linha de referência para a linha de teste.

Além disso, pode-se verificar que a adaptação do Influência do Vértice com as três métricas, Euclidiana, *Manhattan* e *Chebyshev* (VI – EMC), foi o método que demonstrou ser mais viável de se utilizar, visto que em geral os valores de discrepâncias dos dados simulados gerados se aproximaram dos valores de discrepâncias médias utilizadas como referência.

Nesse trabalho foi desenvolvida uma equação empírica para a densificação em função da escala dos dados, onde foi possível obter resultados satisfatórios, bem como realizar as análises. Vale ressaltar que ao utilizar esse artifício de densificar as linhas, deve-se atentar à eficiência computacional.

Dessa forma, aplicando as análises estatísticas para comparação das amostras de discrepâncias dos diferentes métodos aplicados, verificou-se que os métodos são similares estatisticamente independente da métrica utilizada, exceto o método *Buffer Duplo*. Logo, salienta-se, com vistas aos resultados obtidos, que não é viável o uso das Distâncias de *Manhattan* e *Chebyshev*, já que são similares à distância Euclidiana e esta é bastante difundida e tradicional na Cartografia.

Devido à ausência de normas que dizem respeito à avaliação para feições lineares, no caso dos experimentos com dados da base da Espanha e os dados reais, utilizou-se como base para obter classificação o Decreto nº 89.817/ET-CQDG, porém vale lembrar que as tolerâncias e condições definidas neste decreto não foram desenvolvidas para essa primitiva geométrica. Dessa forma, para trabalhos futuros é necessário mais estudo a respeito de normas e padrões voltados à avaliação da acurácia posicional utilizando feições lineares.

Além disso, recomenda-se também o desenvolvimento de técnicas para que seja possível calcular a distância de *Manhattan* ou *Chebyshev* dos vértices da linha de referência para a linha de teste ou vice-versa, sem a necessidade de ter que

densificar as linhas, o que afeta também nos resultados dependendo do quanto uma linha é densificada.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABBAS, I.; GRUSSENMEYER, P.; HOTTIER, P. **Contrôle de la planimétrie d'une base de données vectorielles: une nouvelle méthode basée sur la distance de Hausdorff: la méthode du contrôle linéaire**. Bul. S.F.P.T., (137), p. 6–11, 1995.
- ARAGÃO, M. A.; SANTOS, J. S.; SILVA, M. L. M. Técnica de mineração de dados aplicada a estimação de volume de árvores de Eucalyptus. In: **V SEEFLO-BA Semana de Engenharia Florestal da Bahia, II mostra de pós-graduação em Ciências Florestais da UESB**. Universidade Federal do Oeste do Pará, Santarém, Pará, Brasil, 2019.
- ARIZA, F. J. **Calidad en la Producción Cartográfica**. Ra-Ma, 2002.
- ARIZA, F. J.; ATKINSON, A. D. **Sample Size and Confidence when Applying the NSSDA**. Grupo de Investigación en Ingeniería Cartográfica, Universidad de Jaén, Jaén, España. 2006.
- ARIZA, F. J., and ATKINSON, A. D. **Analysis of Some Positional Accuracy Assessment Methodologies**. *Journal of Surveying Engineering* 134 (2). American Society of Civil Engineers: 45–54. 2008.
- BALAN, A. G. R. **Métodos Adaptativos de Segmentação Aplicados à Recuperação de Imagens por Conteúdo**. Tese (Doutorado), ICMC, Universidade de São Paulo, São Paulo, Brasil, 2007.
- BLAKEMORE, M. Generalization and error in spatial databases. In: **Cartographica**. p. 131- 139, 1984.
- BRASIL. Decreto nº 89.817 de 20 de Junho de 1984. **Normas Técnicas da Cartografia Nacional**. Brasil, 1984. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto/1980-1989/D89817.htm>. Acesso em Outubro de 2020.
- BRAVO, J. V. M; SLUTER, C. R. O problema da qualidade de dados espaciais na era das informações geográficas voluntárias. **Boletim de Ciências Geodésicas**, v. 21, p. 56-73. 2015.
- BUENO, R. **Tratamento do Tempo e Dinamicidade em Dados Representados em Espaços Métricos**. 2009. Tese (Doutorado), ICMC, Universidade de São Paulo, São Paulo, Brasil, 2009.
- BUGGATI, P. H. **Análise da Influência de funções de distância para o processamento de consultas por similaridade em recuperação de imagens por conteúdo**. Dissertação de Mestrado. Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos. 2008.
- CABALLERO, D. Taxicab Geometry: some problems and solutions for square grid-based fire spread simulation. In: **V International Conference on Forest Fire Research**. 2006.

CAI, J.; XU, P.; LI, D.; ZHU, G. A multi-scale positional uncertainty model that incorporates multi-scale modeling errors. In: **Spatial Accuracy 2014**. p. 107–13, 2014, Michigan.

CARVALHO JÚNIOR, O.; COUTO JÚNIOR, A. F.; SILVA, N. C.; MARTINS, E. S.; CARVALHO, A. P. F.; GOMES, R. A. T. Avaliação dos classificadores espectrais de mínima distância euclidiana e Spectral Correlation Mapper em séries temporais NDVI-MODIS no campo de instrução militar de formosa (GO). **Revista Brasileira de Cartografia**, v. 61, p. 399-412. 2009.

CESAR, S. **Geometria Táxi – Uma exploração através de atividades didáticas**. Dissertação de Mestrado. Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, Brasil. 2010.

CHA, S. H., Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions. In: **International Journal Of Mathematical Models And Methods In Applied Sciences**. Vol. 1, 8 p., 2007.

CORDEIRO, D. M. **Modelo conceitual de transporte integrado à rede de hospitais como atributo de otimização de um sistema urbano de atendimento pré-hospitalar móvel**. Tese de Doutorado. Programa de Pós-Graduação em Engenharia Civil. UFPE. Recife. 2012.

CRUZ, D. T. **Distância de Fréchet aplicada no controle de qualidade cartográfica**. Dissertação (mestrado) - Universidade Federal de Viçosa. Viçosa, MG, Brasil, 2019.

CUNHA, M. M.; SECATTO, G. Z.; GALINDO, J. R. F.; SANTOS, A. P. **Proposta de um Método de Avaliação da Acurácia Posicional Baseado na Modificação do Buffer Simples**. *Revista Brasileira de Cartografia*, vol. 71, n. 04, p. 1.193-1.218, out-dez, 2019. Disponível em: <<http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/49301/27587>>. Acessado em Outubro de 2020.

DAI, S.; PEI, D.; WANG, S. Perturbation of fuzzy sets and fuzzy reasoning based on normalized Minkowski distances. In: **Science Direct**. 2011.

DALMOLIN, Q.; LEAL, E. M. Análise da qualidade posicional em bases cartográficas geradas em CAD. In: **Boletim de Ciências Geodésicas**. v. 7, n. 1, p. 21-40, 2001.

DSG. **Especificação Técnica Para Aquisição de Dados Geoespaciais Vetoriais (ET-ADGV)**. Brasil, 2011. Disponível em: <http://www.geoportal.eb.mil.br/images/PDF/ET_ADGV_Vs_2_1_3.pdf>. Acessado em Outubro de 2020.

DSG. **Especificação Técnica para Controle de Qualidade de Dados Geoespaciais (ET-CQDG)**. Brasil, 2016. Disponível em: <http://www.geoportal.eb.mil.br/images/PDF/ET_CQDG_1a_edicao_2016.pdf>. Acesso em Outubro de 2020.

EIDENBERGER, H. Evaluation and Analysis of Similarity Measures for Content-based Visual Information Retrieval. **ACM Multimedia Systems Journal**. 2006.

FERREIRA, L. F.; CINTRA, J. P.: Quantificação de discrepâncias entre feições lineares por retângulos equivalentes. In: **Revista Brasileira de Cartografia**. nº51, p.1-8, 1999.

FONSECA, I. G. R. **Avaliação da acurácia posicional tridimensional utilizando feições lineares**. 2018. 84 f. Dissertação de Mestrado - Curso de Engenharia de Agrimensura e Cartográfica, Engenharia Civil, Universidade Federal de Viçosa.

FREITAS, A. L. B. **Catálogo de metadados de dados cartográficos como suporte para a implementação de clearinghouse nacional**. Dissertação (mestrado) – Instituto Militar de Engenharia. Rio de Janeiro, RJ, Brasil, 2005.

GALATE, R. S.; MOTA, M. G. C.; GAIA, J. M. D.; COSTA, M. S. S. **Distância fenotípica entre matrizes de açaizeiro (*Euterpe oleracea* Mart.) procedentes do nordeste do Pará**. Semina: Ciências Agrárias, Londrina, v. 35, n. 4, p. 1667-1682. 2014.

GALO, M.; DAL POZ, A. P.; FERREIRA, F. M. O uso de feições no controle de qualidade em cartografia. In: **XIX Congresso Brasileiro de Cartografia**. Porto Alegre, RS, Brasil, 2001.

GARDNER, M. **The Last Recreations**. Nova York: Springer-Verlag, p 159-175, 1997.

GOMES, D. A. **Utilização de análise multivariada para seleção de grupos divergentes em genótipos de mandioca**. Dissertação (mestrado), Universidade Federal Rural de Pernambuco, Recife, PE, Brasil, 2018.

GOODCHILD, M. F.; HUNTER, G. J. A simple positional accuracy measure for linear features. In: **International Journal of Geographical Information Science**. v. 11, n. 3, p. 299-306, 1997.

GOODCHILD, M. F. Foreword. In: **SHI, W. Principles of Modeling Uncertainties in Spatial Data and Spatial Analyses**. CRC Press, 2010.

GUSMÃO, N. L.; SKAGUTI, F. Y.; PIRES, L. A. A geometria do táxi: uma proposta da geometria não euclidiana na educação básica. In: **Educação, Matemática, Pesquisa**. São Paulo. v.19, n.2, p. 211-235, 2017.

HEIPKE, C; MAYER, H.; WIEDEMANN, C.; JAMET, O. Evaluation of Automatic Road Extraction. In: **International Archives of Photogrammetry and Remote Sensing**. p. 47–56, 1997.

ISO 19.157. **Geographic information - Data quality**. 2013.

JORGE, A. J.; LEE, H. D.; FERRERO, C. A.; WU, F. C. Estudo da influência de medidas de similaridade da norma lp no algoritmo knn-tsp para previsão de dados temporais. In: **10a Conferência de Dinâmica, Controle e Aplicações**. São Paulo, 28 de agosto a 1º de setembro, 2011.

JURAN, J. M. A função qualidade. In: JURAN, J. M.; GRZYNA, F. M. (Orgs.). **Controle da qualidade: handbook**. São Paulo: Makron Books, 1991. v. 1, p. 10-31.

KRAUSE, E. **Taxicab Geometry: an adventure in non-Euclidean Geometry**. Nova York: Dover, 1975.

KUGLER, M. **Uma contribuição ao desenvolvimento de interfaces cérebro-computador utilizando potenciais visualmente evocados**. Dissertação (mestrado) – Centro Federal de Educação Tecnológica do Paraná. Curitiba, PR, Brasil, 2003.

LAU, J. S. C. **On Taxicab Geometry and Geography**. Ann Arbor, University of Michigan, 1978.

LEI, B.; XU, G.; FENG, M.; ZOU, Y.; VAN der HEIJDEN, F.; RIDDER, D.; TAX, D. M. J. **Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB**. John Wiley & Sons, 2017.

LIMA, E. L. **Espaços Métricos**. 5ª Edição, 2015.

LISBOA, M. H. M. et al (2019). **Avaliação da Acurácia Posicional Tridimensional de Produtos Cartográficos Utilizando um Elipsoide de Incertezas**. Revista Brasileira de Cartografia, v. 71, n. 4, p.598-618, out-nov, 2019. Disponível em: <<http://200.19.146.79/index.php/revistabrasileiracartografia/article/view/49488/27585>>. Acesso em Outubro de 2020.

LUNARDI, O. A.; PENHA, A. L. T.; CERQUEIRA, R. W. O exército brasileiro e os padrões de dados geoespaciais para a INDE. **IV Simpósio Brasileiro de Ciências Geodésicas e Tecnologia da Geoinformação**. Recife, PE, Brasil. p. 8, 2012.

MARQUES, J. **Realimentação de Relevância para Recuperação por Conteúdo de Imagens Médicas Visando Diminuir a Descontinuidade Semântica**. Dissertação (Mestrado), ICMC, Universidade de São Paulo, São Paulo, Brasil, 2006.

MASRY, S. E.; GAUTHIER, J. R. R.; LEE, Y. C. Accuracy and time comparisons of digital maps. In: **Congress of the International Society of Photogrammetry**. Hamburgo, Alemanha. 1980.

McMASTER, R. B. **A statistical analysis of mathematical measures for linear simplification**. The American Cartographer, 13(2), p. 103–16, 1986.

MEDEIROS, J. S. Banco de Dados Geográficos e Redes Neurais Artificiais: Tecnologias de Apoio à Gestão do Território. Tese (doutorado) – Universidade de São Paulo, Faculdade de Filosofia, Letras e Ciências Humanas, Departamento de Geografia. São Paulo, SP, Brasil, 1999.

MENEZES, R. R. V.; LISBOA, M. H. M.; SANTOS, A. P.; DIAS, J. S. **Avaliação da acurácia planimétrica das imagens do Google Earth para produção de base cartográfica**. Revista Brasileira de Cartografia, vol. 71, n. 02, p. 367-391, abr-jun, 2019. Disponível em: <<http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/46327/26303>>. Acessado em Outubro de 2020.

MIGUEL CASTRO, O. 2014. **La evaluación de la exactitud posicional de la información**

geográfica según los estándares en uso. Revista Cartográfica, Nº 90, pp. 81-96.

MIRANDA, D. F. **Geometria Táxi, uma métrica para os espaços geográficos e urbanos uma análise exploratória.** Dissertação (mestrado), PUC-MG, Belo Horizonte, 1999.

MOZAS, A. T.; ARIZA, F. J. New method for positional quality control in cartography based on lines. A comparative study of methodologies. In: **International Journal of Geographical Information Science**. v. 25, n. 10, p. 1681-1695, 2011.

MOZAS, A. T. C.; URENA, M. A.; PÉREZ, J. L. Accuracy of Contour Lines Using 3D Bands. In: **International Journal of Geographical Information Science**. Taylor & Francis: 2362–74, 2013.

MOZAS, A. T.; ARIZA, F. J. Adapting 2D positional control methodologies based on linear elements to 3D. In: **Survey Review**, 47(342), p. 195–201, 2015.

MOZAS, A.T.; UREÑA, M.A.; ARIZA-L, F.J. Determination of 3D Displacements of Drainage Networks Extracted from Digital Elevation Models (DEMs) Using Linear-Based Methods. In: **International Journal of Geographical Information Science**. 2017.

MOZAS, A. T.; ARIZA, F. J. Assessment of Displacements of Linestrings Based on Homologous Vertices. In: **International Journal of Geographical Information Science**. 14 p., 2018.

MOZAS, A. T.; ARIZA, F. J. **Methodology for Positional Quality Control in Cartography Using Linear Features**. v. 47, n. 4, p. 371-378, 2010. Disponível em: <<https://www.ingentaconnect.com/content/maney/caj/2010/00000047/00000004/art00010>>. Acesso em Outubro de 2020.

NETO, I. F. **Um novo conceito de distância: a distância do táxi e aplicações.** Dissertação (mestrado), UNESP, São Paulo, 2013.

NGUYEN, N. T. Inconsistency of Knowledge and Collective Intelligence. **Cybernetics and Systems**, v. 39, p. 542-562. 2008.

NOGUEIRA Jr., J. B.; MALDONADO, V. C. **Controle de Qualidade Posicional em Cartografia, um estudo de caso: município de Paulínia – SP.** Trabalho de Graduação do Curso de Engenharia Cartográfica, UNESP, Faculdade de Ciências e Tecnologia. Presidente Prudente, SP, Brasil. 99 p. 2000.

NOGUEIRA Jr., J. B. **Controle de qualidade de produtos cartográficos: uma proposta metodológica.** 2003. 131 f. Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia, 2003. Disponível em: <<http://hdl.handle.net/11449/88556>>. Acesso em Outubro de 2020.

OLIVEIRA, K. M. **Tomografia Gama Computadorizada e Simulações com MCNPX Aplicados para Estudo da Solda.** Dissertação (mestrado), Universidade Federal de Pernambuco, Recife, PE, Brasil, 2017.

PALMA, L. F. **Agrupamento de dados: K- médias**. Trabalho de Conclusão de Curso, Universidade Federal do Recôncavo da Bahia, Cruz das Almas, 2018.

Pereira, Júlio Cesar Rodrigues. 2004. **Análise de Dados Qualitativos: Estratégias Metodológicas para as Ciências da Saúde, Humanas e Sociais**. 3 ed. São Paulo: Editora da Universidade de São Paulo.

PONCIANO-SILVA, M.; TRAINA, A. J. M.; AZEVEDO-MARQUES, P. M.; FELIPE, J. C.; TRAINA Jr., C. Including the Perceptual Parameter to Tune the Retrieval Ability of Pulmonary CBIR Systems. In: **Proceedings of the 22nd IEEE International Symposium on Computer-Based Medical Systems**. 2009.

PONNMOLI, K. M; SELVAMUTHUKUMARAN, Dr. S. Analysis of Face Recognition using Manhattan Distance Algorithm with Image Segmentation. In: **International Journal of Computer Science and Mobile Computing**. vol.3 Issue.7, p. 18-27, 2014.

PRADO, B. Q. M. **Análise de Agupamentos das taxas de incidência de dengue nos estados brasileiros**. Trabalho de Conclusão de Curso. Universidade Federal de Uberlândia, MG. 2015.

SAITO, Y. K.; VIANA, L. J. F.; SILVA, S. R.; FONSECA, I. G. R.; SANTOS, A. P. **Influência da densidade de vértices nos métodos Distância de Hausdorff e Influência do Vértice**. Revista Brasileira de Cartografia, v. 71, n. 2, p.598-618, abr-jun, 2019. Disponível em: <<http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/44126/26345>>. Acessado em Outubro 2020.

SANTOS, A. P. **Avaliação do Padrão de Exatidão Cartográfica em Imagens Ikonos e Cbers-2b da Bacia do Ribeirão São Bartolomeu, em Viçosa-MG**. Monografia, Universidade Federal de Viçosa. Viçosa, MG, Brasil, 2008.

SANTOS, A. P. **Avaliação Da Acurácia Posicional Em Dados Espaciais Com O Uso de Estatística Espacial**. Dissertação (mestrado) - Universidade Federal de Viçosa. Viçosa, MG, Brasil, 2010.

SANTOS, A. P. **Controle de Qualidade Cartográfica: Metodologias para avaliação da Acurácia Posicional em dados Espaciais**. Tese (doutorado) – Universidade Federal de Viçosa. Viçosa, MG, Brasil, 2015.

SANTOS, A. P.; MEDEIROS, N. das G.; SANTOS, G. R.; RODRIGUES, D. D. Controle de Qualidade Posicional em Dados Espaciais Utilizando Feições. In: **Boletim de Ciências Geodésicas**. v. 21, n. 2, p. 233, 2015.

SANTOS, A. P.; MEDEIROS, N. das G.; SANTOS, G. R.; RODRIGUES, D. D. Avaliação da acurácia posicional planimétrica em Modelos Digitais de Superfície com uso de feições lineares. In: **Boletim de Ciências Geodésicas**. vol. 22, n. 1, 2016b. pp. 157-174.

SHARMA, M.; BATRA, A. Analysis of Distance Measures in Content Based Image Retrieval. In: **Global Journal of Computer Science and Technology: Ginterdisciplinary**. Vol. 14, n. 1, 2014.

SHI, W.; LIU, W. A stochastic process-based model for the positional error of line segments in GIS. In: **International Journal of Geographical Information Science**. v. 14, n. 1, p. 51-66, 2000.

SILVA, P. L.; SANTOS, A. P.; FILHO, J. L. 2020. Proposal of an Academic Spatial Data Infrastructure for the Federal University of Viçosa. In: **International Journal of Spatial Data Infrastructures Research**, v. 15, p. 88-109, 2020.

SILVA, R. P. **Análise da qualidade posicional de produtos gerados por LST e ARP em diferentes alturas de voo**. Dissertação (mestrado) - Universidade Federal de Viçosa. Viçosa, MG, Brasil, 2019.

SOUZA, W. A.; LOTUFO, R.A.; RITTNER, L. **Análise comportamental da Optimum Path Forest em diferentes funções métricas**. Monografia, Universidade Estadual de Campinas. Campinas, SP, Brasil, 2012.

TONG, X.; SUN, T.; FANA, J.; GOODCHILD, M. F.; SHI, W. A statistical simulation model for positional error of line features. In: **International Journal of Applied Earth Observation and Geoinformation**. v. 21, p. 136-148, 2013.

TVEITE, H.; LANGAAS, S. An accuracy assessment method for geographical line data sets based on buffering. In: **International Journal of Geographical Information Science**. v. 13, n. 1, p. 27-47, 1999.

VADIVEL, A.; MAJUMDAR, A. K.; SHAMIK, S. **Performance comparison of distance metrics in content-based Image retrieval applications**. Bhubaneswar, India, 2003.

VICINI, L.; SOUZA, A. M. **Análise multivariada da teoria à prática**. Biblioteca Central da Universidade Federal de Santa Maria, 215 p., Santa Maria, RS, Brasil, 2005.

XAVIER, E. M. A. ARIZA-LÓPEZ, F. J. URENA-CAMARA, M. A. **MatchingLand, geoespacial data testbed for the assessment of matching methods**. 2017.

ZANETTI, J.; BRAGA, F. L. S.; A. P. S. **Comparativo das normas de controle de qualidade posicional de produtos cartográficos do Brasil, da ASPRS e da OTAN**. Revista Brasileira de Cartografia, vol. 70, n. 01, p. 359-390, jan-mar, 2018.

APÊNDICE A

A.1: Valor da densificação da linha teste e/ou referência e a quantidade de vértices a serem considerados no cálculo.

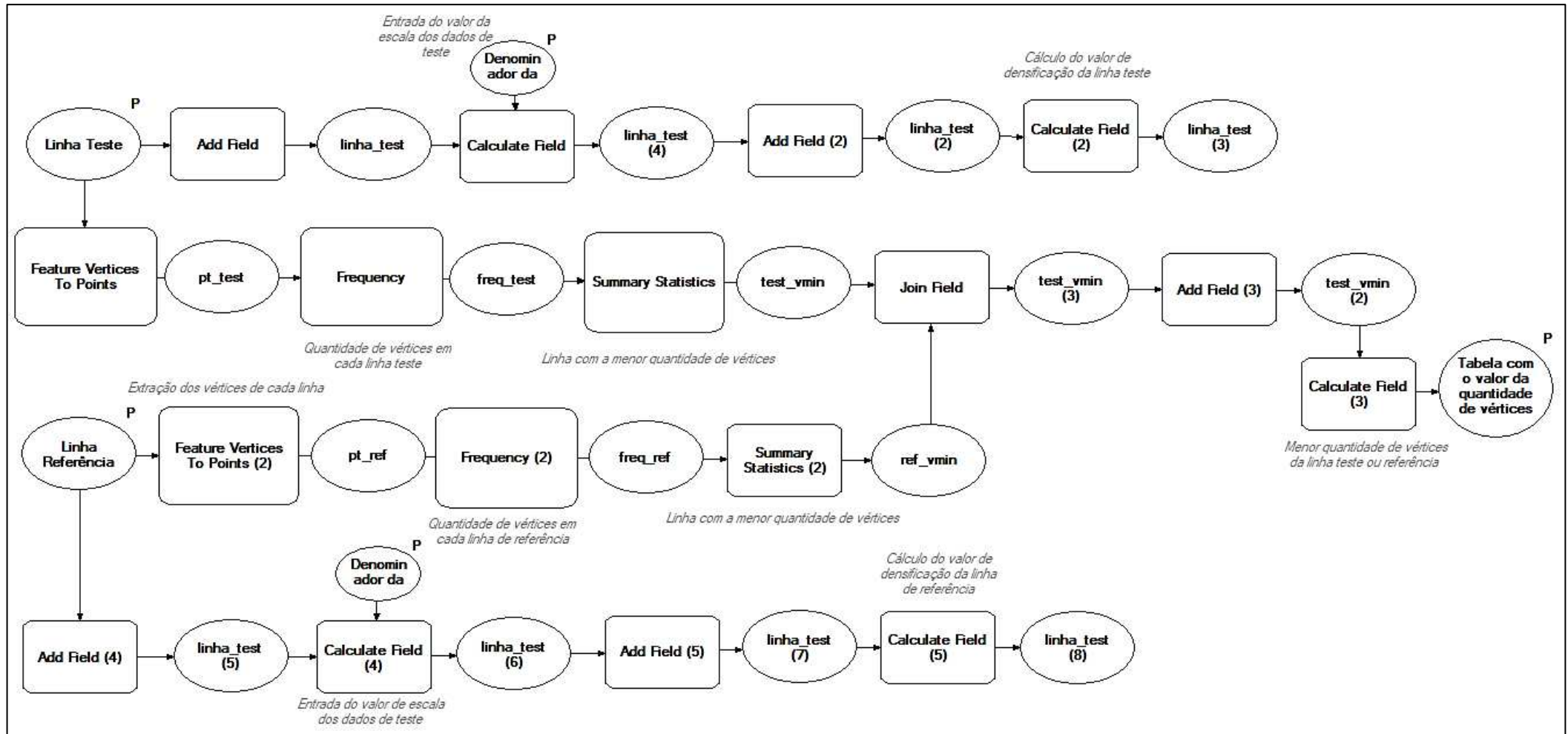


Figura 41 - Densificação e quantidade de vértices para o cálculo das distâncias.

APÊNDICE B

Foram executados alguns testes utilizando dados da base *MatchingLand* com diferentes valores para o espaçamento da densificação considerando o erro de graficismo (0,2 mm) e a escala 1:25.000, para o caso do método Influência do Vértice com a distância de Manhattan (VI – M). Inicialmente foi feito em um par de linha, onde o número máximo de vértices considerado no *generate near table* (comando do ArcGIS) foi de 19, visto que era o valor correspondente à linha de referência com menor números de vértices. Então obteve-se os resultados abaixo, onde eg corresponde à abreviação do erro de graficismo.

1 par de linha		
Tolerância	Espaçamento	Discrepância (m)
eg*esc /1000	0,5 cm	1,397
eg*esc /500	1 cm	1,385
eg*esc 100	5 cm	1,317
eg*esc /50	10 cm	1,298
eg*esc /25	20 cm	1,276
eg*esc /20	25 cm	1,277
eg*esc /15	33 cm	1,272
eg*esc /10	50 cm	1,312
eg*esc /6	84 cm	1,361
eg*esc /4	1,25 m	1,428
eg*esc /2	2,50 m	1,462

*eg = erro de graficismo

*ec = denominador da escala

Como pode observar, considerando somente a metade do erro de graficismo, foi obtido o maior valor de discrepância em relação aos outros valores. Com 33 cm, o valor da discrepância foi o menor, mas ressalta-se que aqui não se procurou um valor de espaçamento para a densificação que resulte na menor discrepância possível. Esse valor de densificação deve ser utilizado para obter o valor de discrepância mais próximo do valor real possível, que seria aquele em que não precisasse densificar a linha e tivesse como calcular a distância pra qualquer local do segmento, como é feito considerando a distância Euclidiana.

Posteriormente foram selecionadas 7 pares de linhas aleatórias dos dados simulados da *MatchingLand*, e entre esses pares utilizou-se o par de linha em que foi feito teste acima. O OBJECTID referente a esse par de linha é o 2. Utilizou-se os mesmos espaçamentos do teste anterior para a densificação.

Entretanto no número de vértices para ser usado como tolerância no *generate near table*, foi utilizado o valor 12, pois nesse caso havia uma linha de referência com esse número de vértices, que era a que possuía o menor valor. Devido a isso observa-se que para o par de linha utilizado em ambos os testes, os resultados foram um pouco diferentes, o que é esperado devido ao próprio método. E vale ressaltar, que esse valor a ser utilizado no *generate near table*, deve ser correspondente mesmo à linha que possui o menor número de vértices, pois se utilizar um valor maior, um vértice de uma linha pode acabar calculando a distância considerando um vértice de uma linha que não seja homóloga à ele. Sendo assim, os resultados estão abaixo.

		7 pares de linha						
		Discrepâncias (m)						
OBJECTID		1	2	3	4	5	6	7
Espaçamento	0,50 cm	1,204	1,403	2,224	1,711	14,622	3,343	2,719
	1,00 cm	1,198	1,394	2,210	1,705	14,611	3,331	2,713
	5,00 cm	1,187	1,349	2,144	1,688	14,543	3,247	2,674
	10,00 cm	1,187	1,314	2,083	1,634	14,479	3,147	2,640
	20,00 cm	1,204	1,298	2,041	1,604	14,400	3,022	2,603
	25,00 cm	1,210	1,302	2,029	1,597	14,376	3,000	2,578
	33,00 cm	1,211	1,286	2,027	1,597	14,356	2,956	2,552
	50,00 cm	1,225	1,312	2,038	1,608	14,324	2,983	2,529
	84,00 cm	1,266	1,361	2,100	1,662	14,342	3,000	2,56
	1,25 m	1,348	1,369	2,214	1,713	14,397	3,103	2,605
	2,50 m	1,418	1,462	2,423	1,735	14,574	3,330	2,683

OBJECTID	Discrepância máxima (m)	Discrepância mínima (m)	Discrepância média (m)	Amplitude (m)
1	1,418	1,187	1,242	0,231
2	1,462	1,286	1,350	0,176
3	2,423	2,027	2,139	0,396
4	1,735	1,597	1,659	0,138
5	14,622	14,324	14,457	0,298
6	3,343	2,956	3,133	0,387
7	2,719	2,529	2,623	0,190

Da mesma forma, nesse conjunto de dados, o maior valor de discrepância foi considerando uma densificação de 2,5 m. Abaixo está o resultado do Influência do Vértice em sua forma original.

Influência do Vértice em sua forma original		
	OBJECTID	Discrepância (m)
1 par de linha	-	1,038
	1	0,92
	2	1,038
7 pares de linhas	3	1,815
	4	1,265
	5	11,853
	6	2,546
	7	1,997

APÊNDICE C

A.1: Ferramenta para a adaptação do Influência do Vértice com a Distância de *Manhattan* (VI – M) no ArcToolbox.

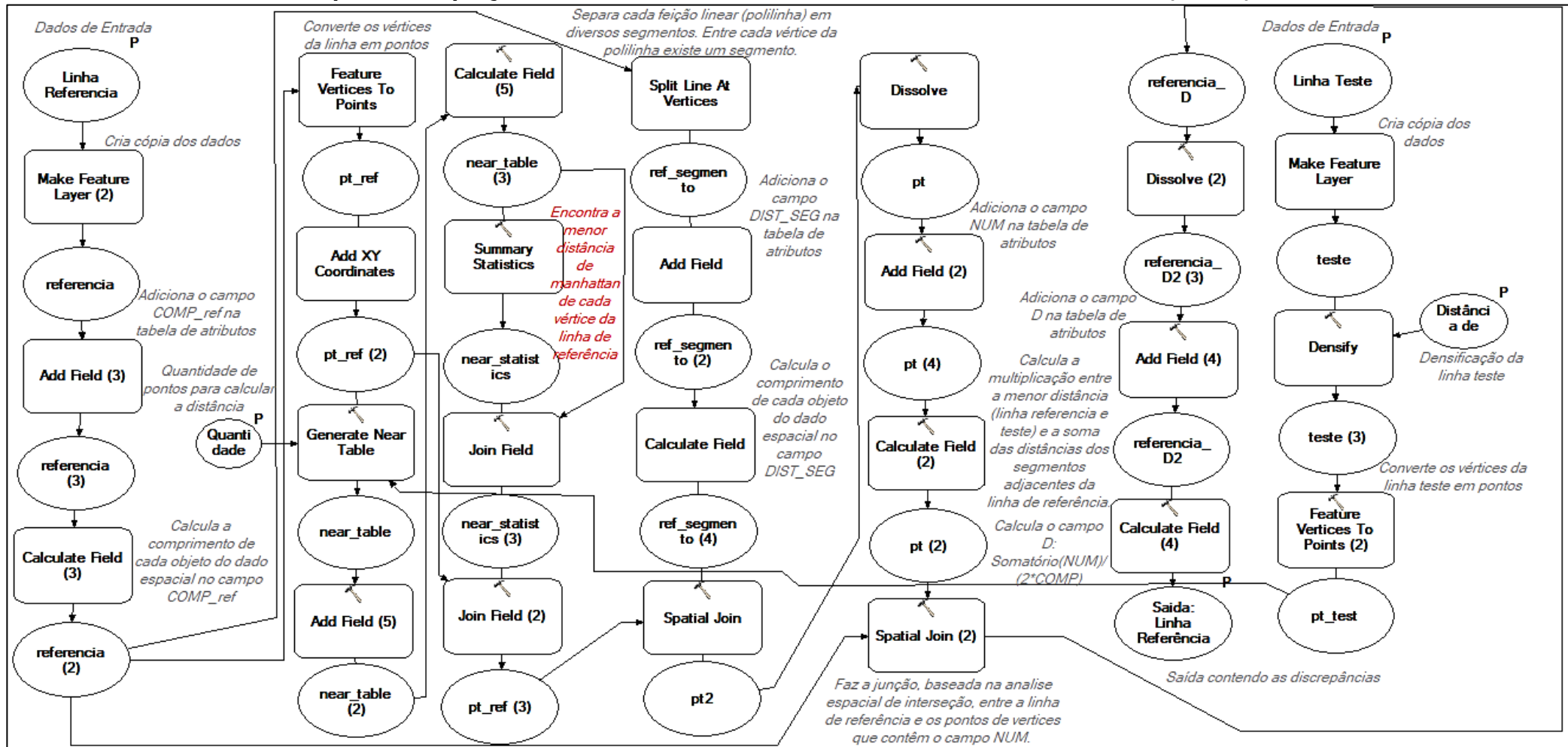


Figura 42 - Adaptação do Influência do Vértice com a Distância de *Manhattan* (VI – M).

A.2: Ferramenta para a adaptação do Influência do Vértice com a Distância de *Chebyshev* (VI – C) no ArcToolbox.

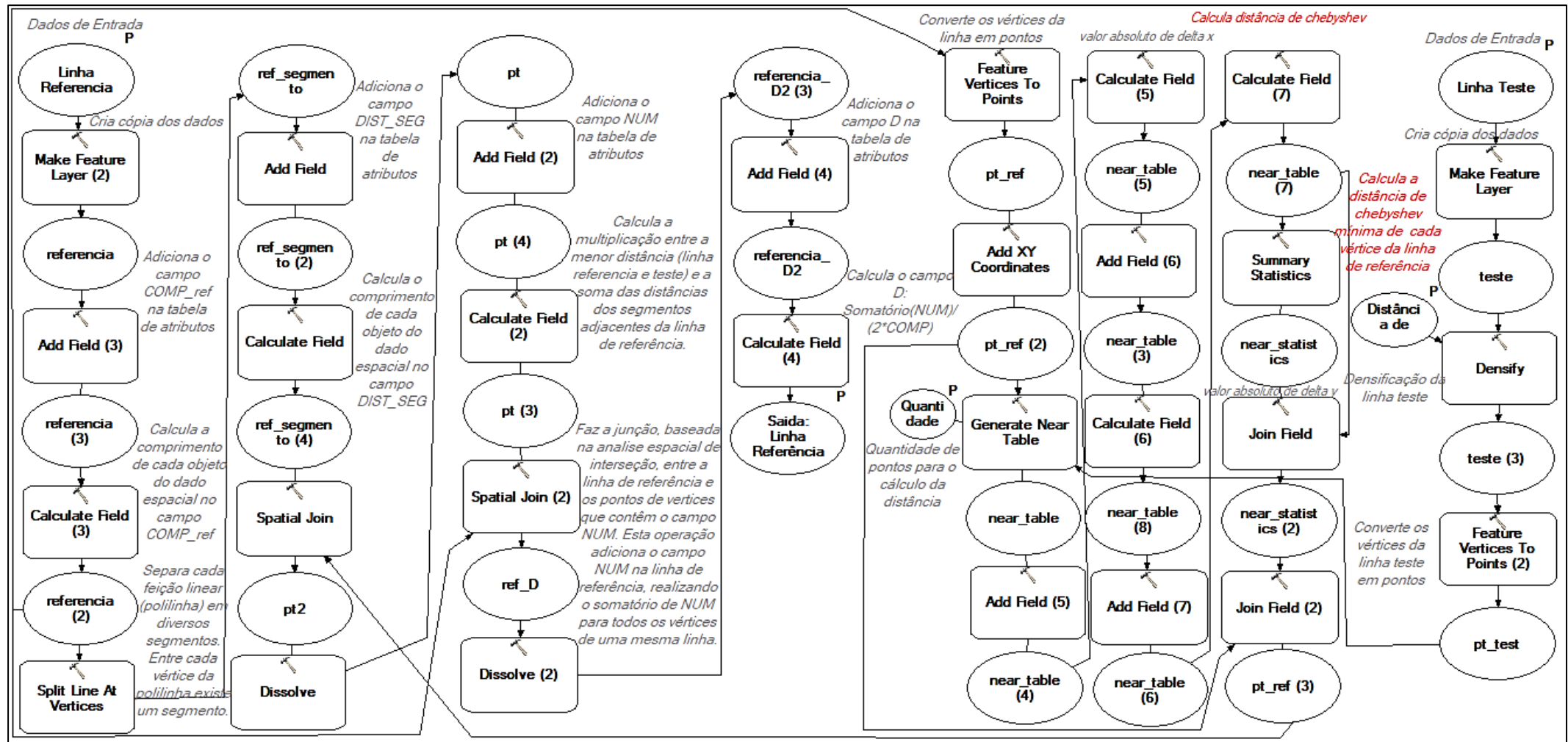


Figura 43 - Adaptação do Influência do Vértice com a Distância de *Chebyshev* (VI – C).

A.3: Ferramenta referente ao método VI - EMC.

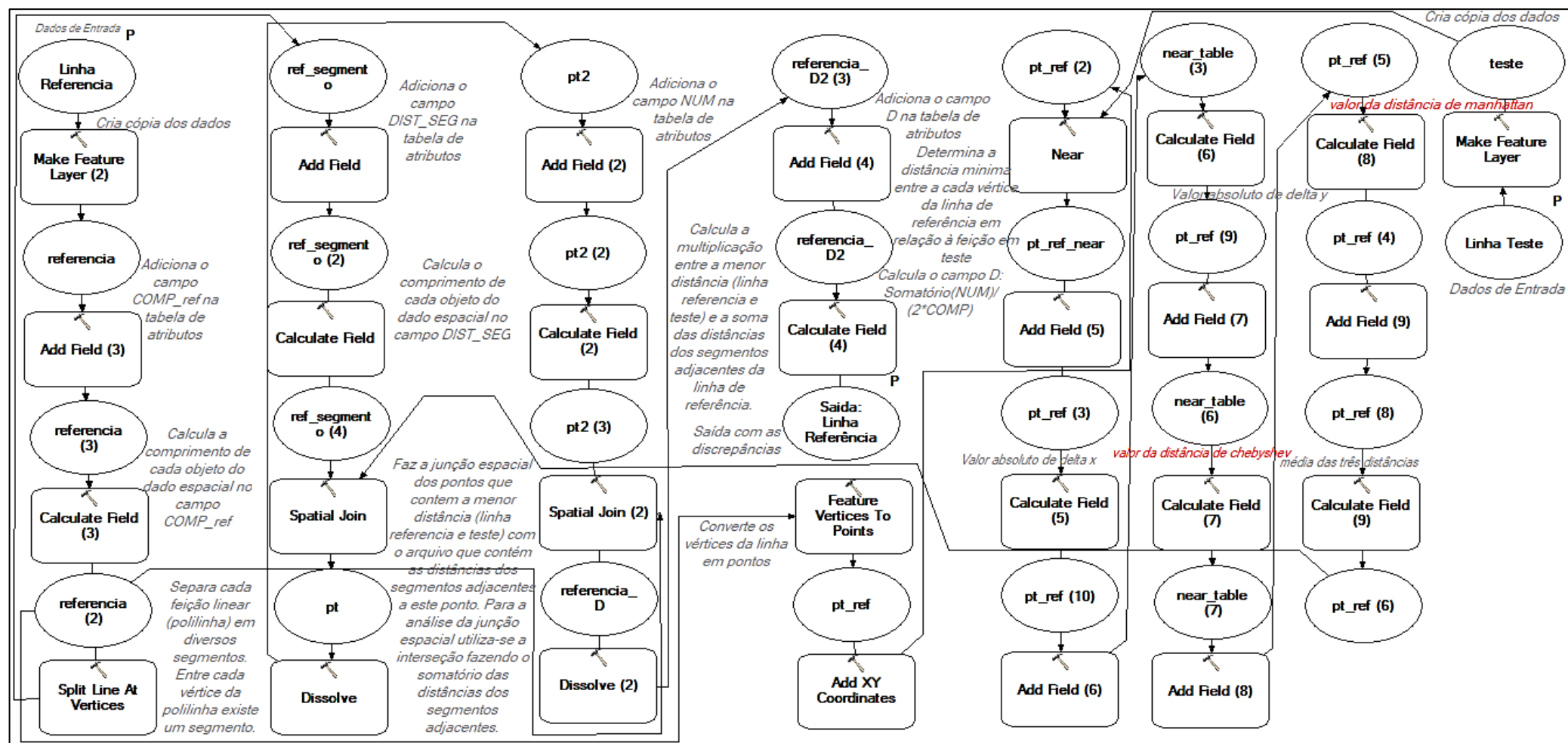


Figura 44 - Adaptação do influência do Vértice com a distância de *Manhattan* e *Chebyshev* (VI – EMC).

A.4: Distância de *Manhattan* calculando a distância dos vértices da linha de referência para os vértices da linha teste.

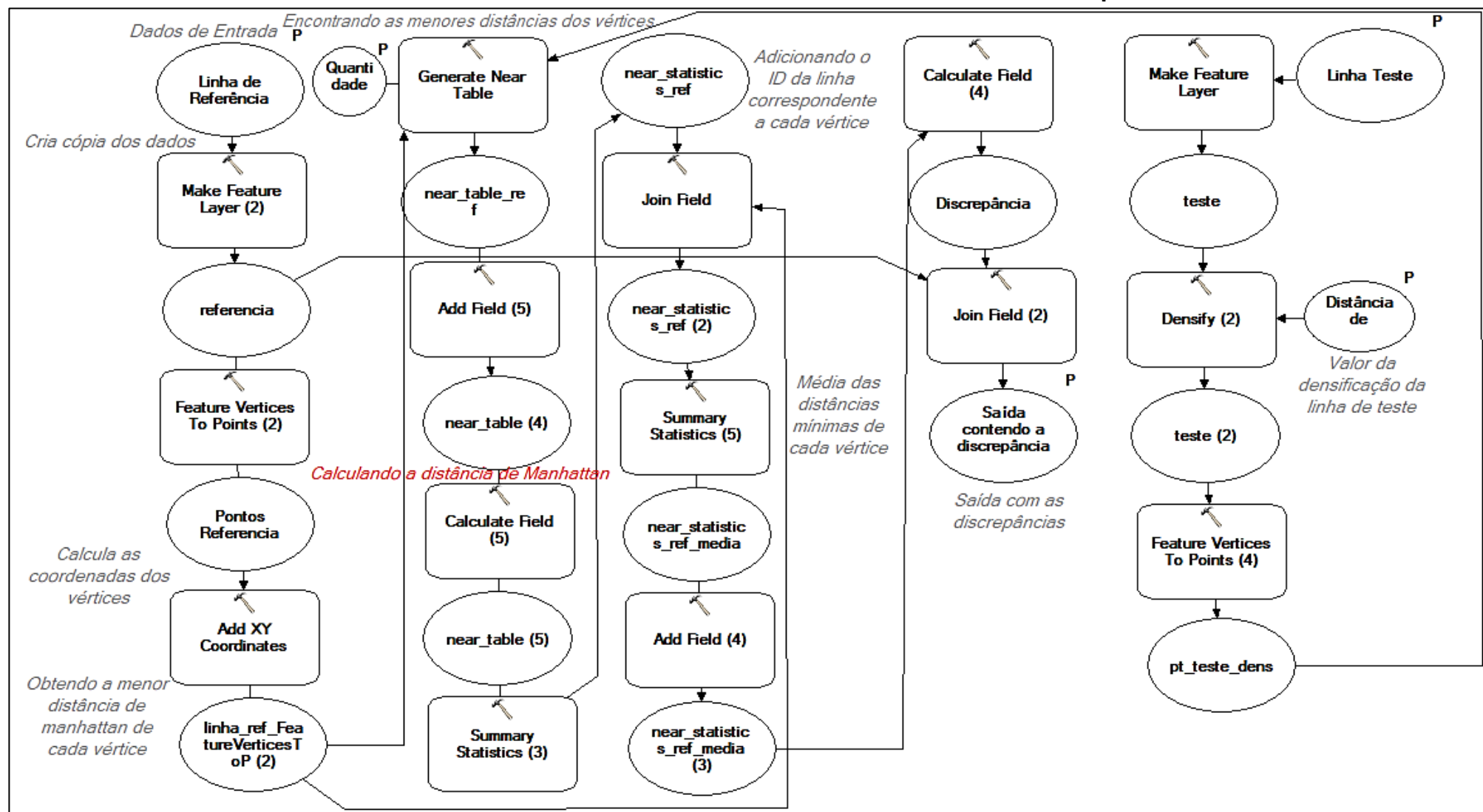


Figura 45 - Novo método HD - M2, calculando apenas as distâncias da linha de referência para a linha teste.

A.5: Distância de Chebyshev calculando a distância dos vértices da linha de referência para os vértices da linha teste.

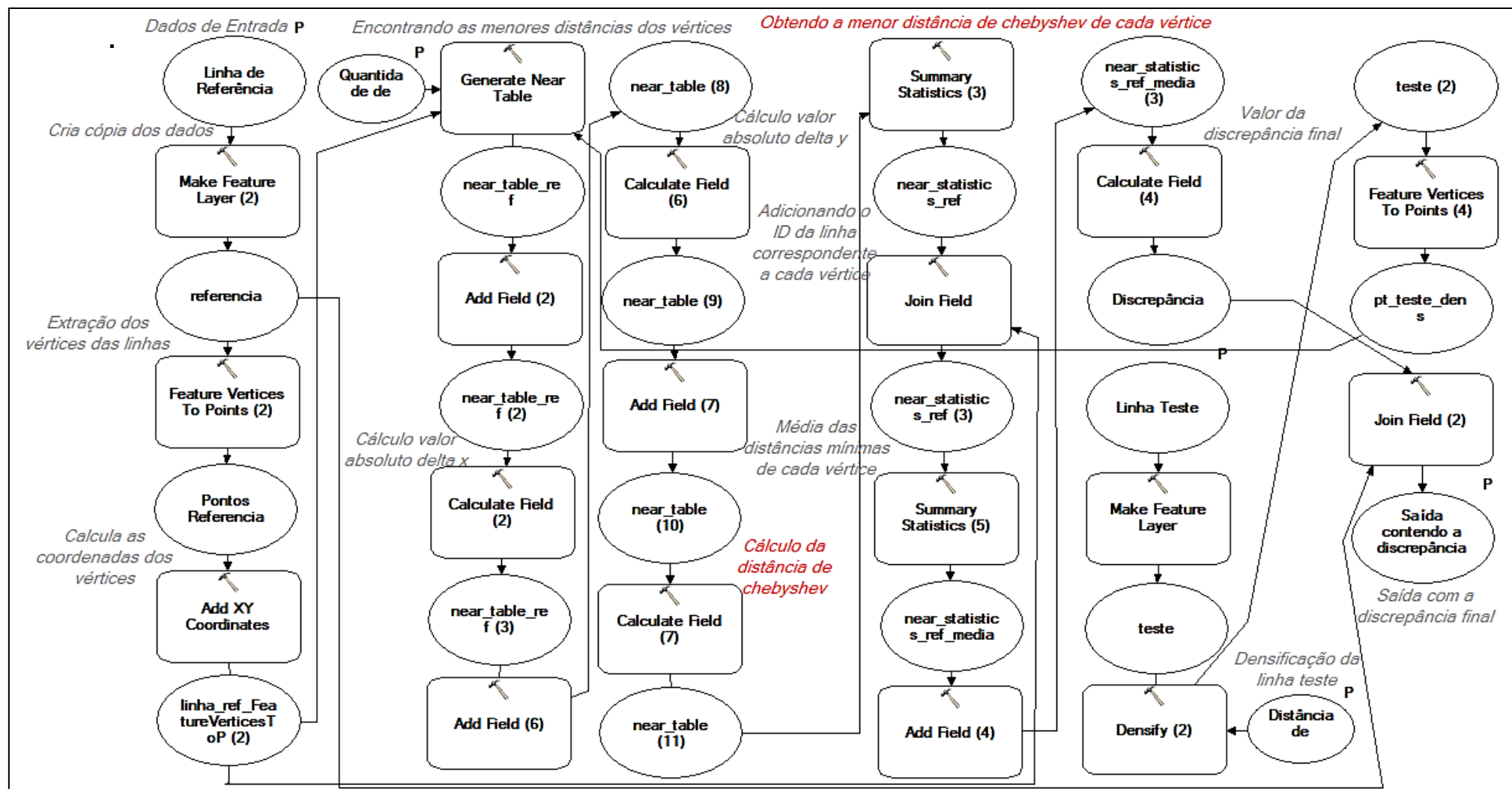


Figura 46 - Novo método HD – C2, calculando apenas as distâncias da linha de referência para a linha teste.

A.6: Distância de *Manhattan* com a distância dos vértices da linha de referência para a linha teste e vice-versa

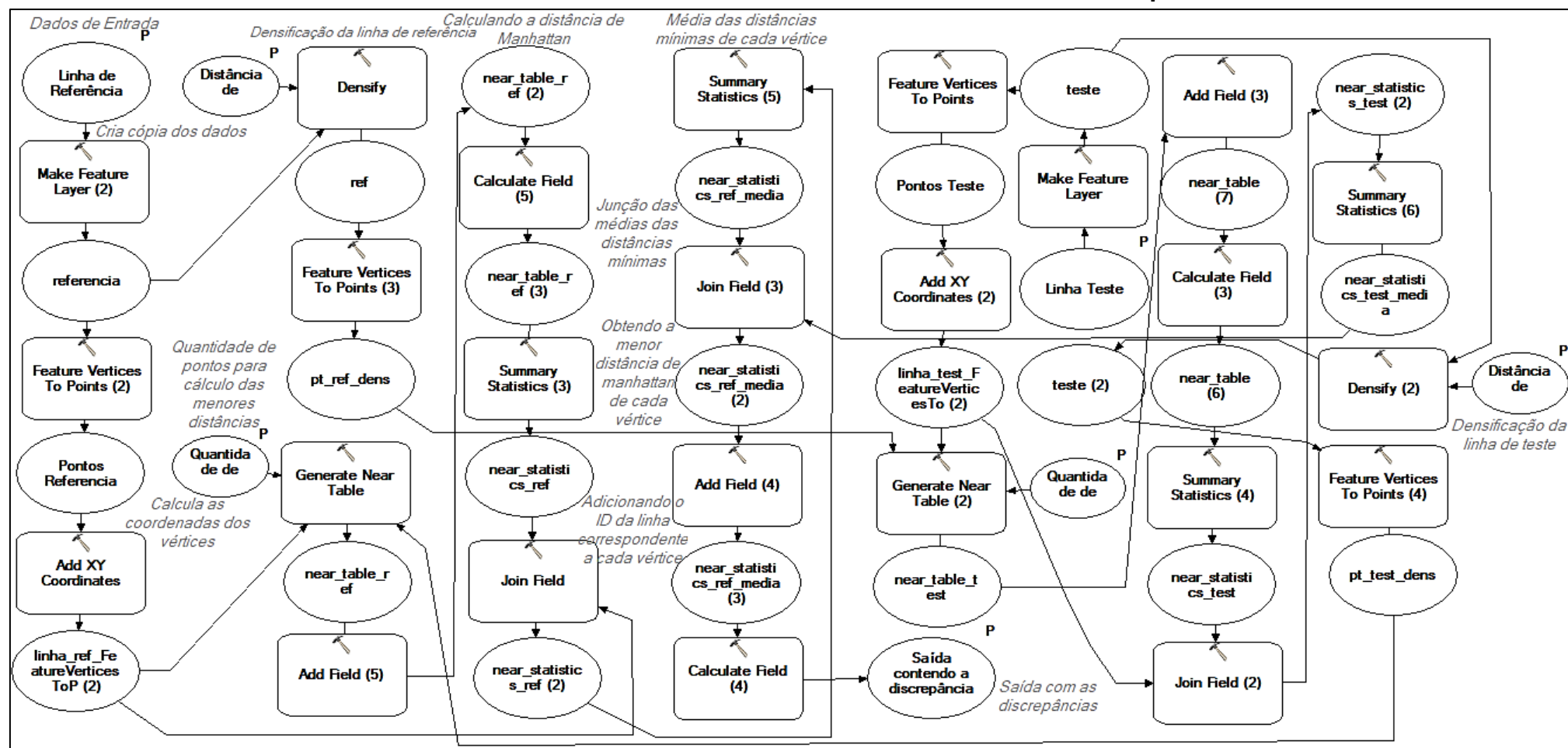


Figura 47 - Novo método HD - M, calculando apenas as distâncias da linha de referência para a linha teste e vice-versa.

A.7: Distância de *Chebyshev* com a distância dos vértices da linha de referência para a linha teste e vice-versa.

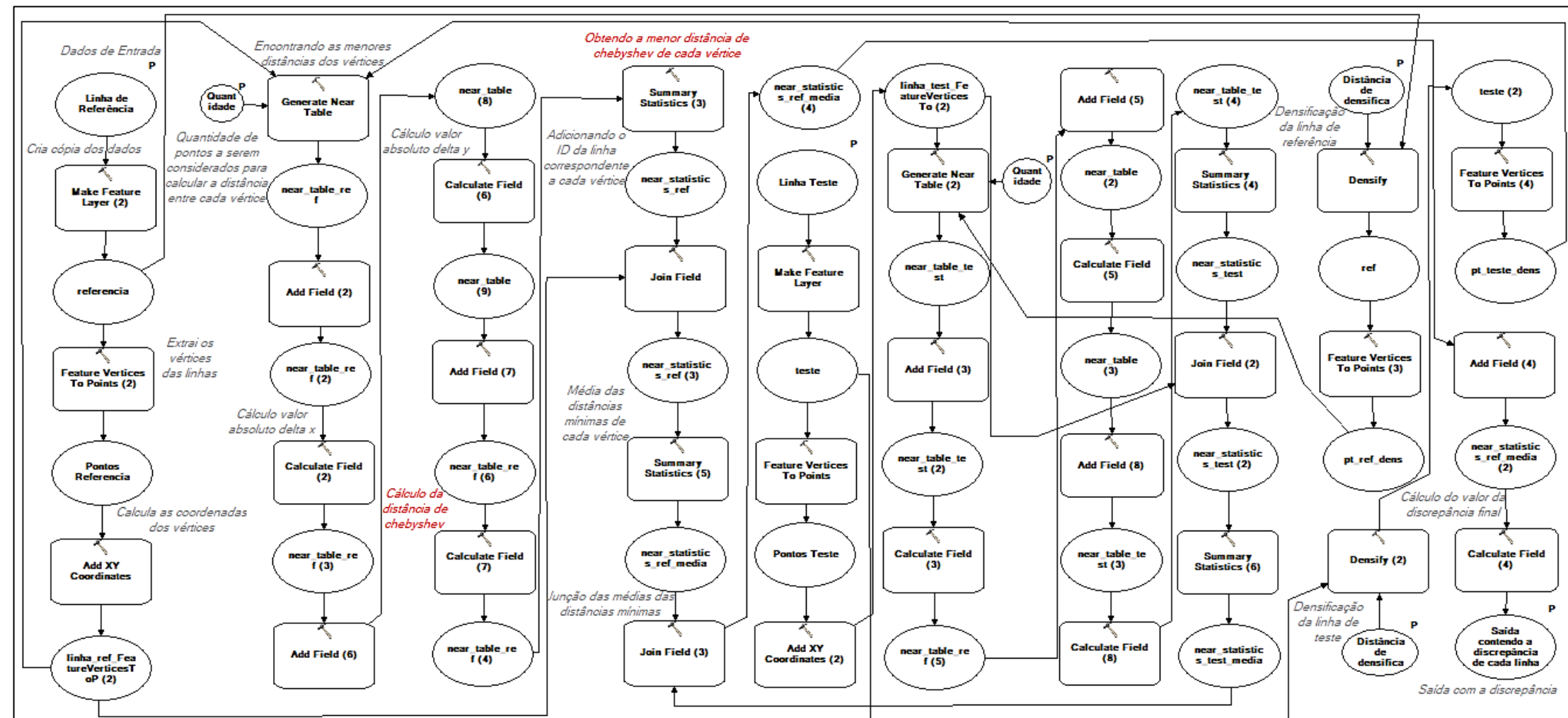


Figura 48 - Novo método HD - C, calculando apenas as distâncias da linha de referência para a linha teste e vice-versa.