



Serviço Público Federal
Ministério da Educação
Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Engenharias, Arquitetura e Urbanismo e Geografia - FAENG



MAURICIO DE SOUZA

**AVALIAÇÃO E GERAÇÃO DE MODELO DIGITAL DE
TERRENO COMBINANDO FOTOGRAMETRIA E VISÃO
COMPUTACIONAL**

PPCTA
TECNOLOGIAS AMBIENTAIS

**Campo Grande, MS.
2025**

**UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
FACULDADE DE ENGENHARIAS E ARQUITETURA E URBANISMO E GEOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS AMBIENTAIS**

MAURICIO DE SOUZA

**AVALIAÇÃO E GERAÇÃO DE MODELO DIGITAL DE
TERRENO COMBINANDO FOTOGRAMETRIA E VISÃO
COMPUTACIONAL**

Tese submetida ao Programa de Pós-graduação em Tecnologias Ambientais da Universidade Federal do Mato Grosso do Sul e apresentada à banca examinadora como requisito para obtenção do título de Doutor em Tecnologias Ambientais.

Orientador: Prof. Dr. José Marcato Junior
Coorientador: Prof. Dr. Wesley Nunes Gonçalves

Aprovada em:

Campo Grande, 10 de setembro de 2025.

Banca Examinadora:

Prof. Dr. José Marcato Junior
Orientador PGTA-UFMS

Profª. Dra. Ana Paula Marques Ramos
Departamento de Cartografia, UNESP

Prof. Dr. Lucas Prado Osco
Faculdade de Engenharia Arquitetura e
Urbanismo FEPP-UNOESTE

Prof. Dr. Jonathan de Andrade Silva
Faculdade de Computação
FACOM-UFMS

Prof. Dr. Robson Soares Silva
Faculdade de Computação
FACOM-UFMS

**Campo Grande, MS.
2025**

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus, por me conceder a força necessária para concluir esta etapa. Dedico esta conquista à minha esposa, Camila, por seu apoio inestimável, e ao meu filho, Bernardo, fonte de alegria e perseverança.

Expresso minha mais profunda gratidão ao meu orientador e coorientador, cuja paciência e sabedoria foram essenciais para o meu desenvolvimento acadêmico e para a finalização deste trabalho. Sou grato aos meus colegas de laboratório, pelas conversas produtivas, pelo ambiente colaborativo e pelo café, assim como a todos os amigos e familiares que me apoiaram.

Agradeço ao Programa de Pós-Graduação em Tecnologias Ambientais (PPGTA) da Faculdade de Engenharias, Arquitetura e Urbanismo (FAENG) e à Fundação Universidade Federal de Mato Grosso do Sul (UFMS/MEC) pelo suporte institucional.

SUMÁRIO

RESUMO GERAL	iv
GENERAL ABSTRACT	v
INTRODUÇÃO GERAL	1
OBJETIVO GERAL	3
Objetivos específicos	4
ORGANIZAÇÃO	4
REFERÊNCIAS	5
PRIMEIRO CAPÍTULO: EVALUATION OF ALTIMETRIC ACCURACY OF DIGITAL SURFACE MODELS IN THE URBAN AREA OF CAMPO GRANDE/BRAZIL	7
1.1. INTRODUCTION	9
1.2. MATERIALS AND METHODS	10
1.2.1 Study area and reference data	11
1.2.2 Digital Surface Models	12
1.2.3 Validation of the DSM	13
1.3. RESULTS	14
1.3.1 Results dataset 1	14
1.3.2 Results dataset 2	17
1.4. DISCUSSION	19
1.5. CONCLUSIONS	20
1.6. REFERENCES	21
SEGUNDO CAPÍTULO: ACCURACY OF HIGH RESOLUTION DIGITAL CARTOGRAPHIC PRODUCTS WITH ELEVATION CONTROL POINTS	23
2.1. INTRODUCTION	25
2.2. METHODOLOGY AND DATA	26
2.2.1. Area of study	26
2.2.2. Data Acquisition	27
2.2.3. Quantity of Control Points	28
2.2.4. Control Points at Different Elevations	30
2.2.5. Generation of Cartographic Products	31
2.2.6. Validation of Digital Cartographic Products	32
2.3. RESULTS	34
2.3.1 Number of GCPs	34
2.3.2. GCPs in Elevation	37
2.4 DISCUSSION	40
2.5 CONCLUSIONS	42
2.6 REFERENCES	43

TERCEIRO CAPÍTULO: FILTRAGEM DA NUVEM DE PONTOS GERADA POR UAV RGB COMBINANDO DEEP LEARNING E FOTOGRAMETRIA	44
3.1. INTRODUÇÃO	47
3.2. MATERIAL E MÉTODOS	49
3.2.1. Área de Estudo e aquisição das imagens RGB	49
3.2.2. Segmentação Semântica das árvores nas imagens RGB	51
3.2.3. Geração do MDT	53
3.2.4. Avaliação dos MDT	55
3.3. RESULTADOS	57
3.3.1. Avaliação das arquiteturas DL	58
3.3.2. Avaliação Altimétrica	62
3.3.2.1 Avaliação visual da nuvem densa de pontos	63
3.3.2.2 Avaliação visual dos MDTs	68
3.3.2.3 Avaliação das discrepâncias dos MDTs em relação aos pontos de checagem	73
3.3.3. Análise visual Ortofotos	74
3.4. DISCUSSÃO	77
3.5. CONCLUSÃO	80
3.6. REFERÊNCIAS	81
CONCLUSÃO GERAL	84

RESUMO GERAL

SOUZA, M. (2025). Avaliação e geração de modelo digital de terreno combinando fotogrametria e visão computacional. 2025. p. 85. Tese (Doutorado) - Programa de Pós-Graduação em Tecnologias Ambientais. Universidade Federal de Mato Grosso do Sul, Brasil.

Os Modelos Digitais de Terreno (MDT) são uma importante fonte de informação para estudos de inundações, enxurradas e dinâmica do escoamento superficial, uma vez que auxiliam na compreensão do percurso das águas. Apesar da ampla disponibilidade de Modelos Digitais de Superfície (MDS) globais e dos avanços no uso de veículos aéreos não tripulados (VANT) para reconstrução 3D, ainda são escassos os estudos que avaliam sistematicamente a acurácia desses produtos em áreas urbanas e, principalmente, que propõem estratégias automatizadas para a remoção de elementos acima do solo, como vegetação e edificações. Esta tese busca preencher essa lacuna ao integrar métodos tradicionais de fotogrametria com técnicas de aprendizado profundo para a geração direta de MDTs urbanos. No primeiro capítulo, foi realizada uma avaliação da acurácia vertical de MDS globais de acesso aberto, além do modelo pago TanDEM-X, na área urbana de Campo Grande-MS. Para isso, utilizaram-se como referência pontos com acurácia planialtimétrica centimétrica obtidos via GNSS, avaliando em que escala os modelos podem ser utilizados conforme o Padrão de Exatidão Cartográfica para Produtos Cartográficos Digitais (PEC-PCD). Os resultados indicaram que os modelos TanDEM-X e ALOS AW3D30 são adequados para aplicações na escala 1:50.000, classe A da PEC-PCD. No segundo capítulo, foi gerado um MDS a partir de imagens capturadas por VANT, e analisou-se a influência da quantidade e da distribuição dos pontos de controle em solo (Ground Control Points – GCP) na acurácia do modelo. Foram testadas diferentes quantidades de GCP (2 a 6) e comparadas duas configurações com seis pontos: uma com todos os GCPs no plano e outra com alguns em áreas elevadas. Verificou-se que a distribuição altimétrica dos GCPs resulta em melhoria significativa na qualidade do MDS gerado. No terceiro capítulo, foram aplicadas cinco arquiteturas de deep learning (FCN, GCNet, PointRend, PSANet e SegFormer) para a segmentação de árvores em imagens obtidas por VANT, com o objetivo de gerar MDTs diretamente a partir de dados SfM-MVS, utilizando apenas os pontos classificados como solo nas regiões de máscara. Os MDTs resultantes foram comparados com modelos gerados por meio do filtro tradicional Cloth Simulation Filtering (CSF). Os resultados demonstraram que o uso de DL produz MDTs com acurácia semelhante ao CSF, mas com maior potencial de automação e escalabilidade. Conclui-se que a integração de dados VANT com segmentação semântica por aprendizado profundo constitui uma abordagem viável e inovadora para a geração eficiente de MDTs urbanos, contribuindo com aplicações práticas em modelagem hidrológica, planejamento urbano e estudos ambientais.

Palavras-chave: Modelagem do terreno; Padrão de Exatidão Cartográfica; PEC-PCD; Acurácia Vertical

GENERAL ABSTRACT

SOUZA, M. (2025). Evaluation and generation of digital terrain models combining photogrammetry and computer vision. 2025. p. 85. Thesis (Doctorate) - Postgraduate Program in Environmental Technologies. Federal University of Mato Grosso do Sul, Brazil.

Digital Terrain Models (DTMs) are an important source of information for studies on flooding, flash floods, and surface runoff dynamics, as they help to understand the flow paths of water. Despite the wide availability of global Digital Surface Models (DSMs) and the advances in the use of Unmanned Aerial Vehicles (UAVs) for 3D reconstruction, few studies have systematically evaluated the vertical accuracy of these products in urban areas and, more importantly, proposed automated strategies for removing above-ground elements such as vegetation and buildings. This thesis aims to fill this gap by integrating traditional photogrammetric methods with deep learning techniques to enable the direct generation of urban DTMs. In the first chapter, a vertical accuracy assessment was conducted for open-access global DSMs, as well as the commercial TanDEM-X model, in the urban area of Campo Grande, MS, Brazil. Ground control data with centimeter-level planialtimetric accuracy, acquired via GNSS, were used as a reference to evaluate the suitability of these models at different map scales according to the Cartographic Accuracy Standard for Digital Cartographic Products (PEC-PCD). Results indicated that both the TanDEM-X and ALOS AW3D30 models meet the requirements for 1:50,000 mapping, class A of PEC-PCD. In the second chapter, a DSM was generated from UAV imagery, and the influence of the number and distribution of Ground Control Points (GCPs) on the accuracy of the model was analyzed. Different numbers of GCPs (2 to 6) were tested, along with two configurations using six GCPs: one with all points on a flat surface and another including elevated points. It was found that including GCPs in elevated positions significantly improved the quality of the DSM. In the third chapter, five deep learning architectures (FCN, GCNet, PointRend, PSANet, and SegFormer) were applied for tree segmentation in UAV imagery, aiming to directly generate DTMs from SfM-MVS data using only points classified as ground in the mask regions. The resulting DTMs were compared with those obtained using the traditional Cloth Simulation Filtering (CSF) method. Results showed that the DL-based approach yields DTMs with accuracy comparable to CSF, while offering greater potential for automation and scalability. It is concluded that the integration of UAV data with semantic segmentation through deep learning represents a viable and innovative approach for efficient urban DTM generation, with practical applications in hydrological modeling, urban planning, and environmental studies.

Keywords: Terrain Modeling; Cartographic Accuracy Standard; PEC-PCD; Vertical Accuracy

INTRODUÇÃO GERAL

As mudanças climáticas resultarão em uma maior frequência de eventos extremos (IPCC, 2012). Para mitigar os impactos socioeconômicos de tais eventos, é necessário o desenvolvimento de tecnologias de monitoramento e modelos preditivos capazes de indicar áreas potencialmente afetadas, permitindo a adoção de medidas preventivas (Trepekli et al., 2022). Nas simulações de enchentes e alagamentos, os Modelos Digitais de Superfície (MDS) constituem um ponto de partida amplamente adotado na literatura, uma vez que a topografia define o fluxo da água.

Atualmente, existem MDS globais de livre acesso com diferentes resoluções espaciais, utilizados em variadas aplicações, tais como: mapeamento de risco de inundação (Ho et al., 2010), modelagem de propagação de inundações (Yan et al., 2015), modelagem hidráulica de inundação de lagos glaciais (Wang et al., 2012), simulação de inundação (Xu et al., 2021) e mapeamento de vulnerabilidade a inundações costeiras (Zhang et al., 2019). Estudos como os de Wang & Zheng (2005) e Cook & Merwade (2009) evidenciam que resoluções espaciais menos detalhadas (maiores valores de pixel) tendem a superestimar a área prevista de inundação, enquanto resoluções mais detalhadas resultam em delimitações mais precisas.

Para a realização de simulações hidrológicas, é essencial a obtenção de um Modelo Digital do Terreno (MDT), que representa a superfície terrestre sem considerar objetos acima do solo. Nesse sentido, O’Loughlin et al. (2016) propuseram uma correção do *Shuttle Radar Topography Mission* (SRTM) utilizando dados de porcentagem de cobertura arbórea do sensor *Moderate Resolution Imaging Spectroradiometer* (MODIS) para remover o viés da vegetação, gerando um MDS global de superfície exposta (*bare-earth*). Entre os principais MDS globais disponíveis destacam-se o SRTM (30 m) e o *Advanced Spaceborne Thermal Emission and Reflection Radiometer* (ASTER) (30 m). Apesar de já existirem diversas avaliações desses modelos, há uma escassez de estudos sobre a acurácia altimétrica do MDS TanDEM-X em áreas urbanas brasileiras.

O nível de detalhamento e a acurácia dos MDS globais nem sempre permitem a geração de produtos cartográficos em escalas grandes (1:25.000 ou maiores). Nesse contexto, veículos aéreos não tripulados (VANTs) constituem uma alternativa para obtenção de MDS de maior resolução, seja por meio de sistemas *Light Detection and*

Ranging (LiDAR), seja pela aplicação de fotogrametria combinada a técnicas de visão computacional, como *Structure from Motion e Multi-View Stereo* (SfM-MVS) (Smith et al., 2016). O SfM-MVS possibilita a geração de nuvens de pontos densas em 3D a partir de um conjunto de imagens 2D, desde que sejam observados parâmetros específicos na aquisição. Embora VANTs disponham de receptores *Global Navigation Satellite System* (GNSS) embarcados, estes geralmente apresentam precisão posicional limitada (erro entre 5 e 15 metros). Por isso, o uso de pontos de controle em solo (Ground Control Points – GCPs), obtidos com receptores GNSS de acurácia centimétrica, é fundamental para a produção de MDS de alta precisão, possibilitando medições lineares e volumétricas confiáveis em ortoimagens.

Diversos estudos têm investigado o impacto da distribuição espacial dos GCPs na acurácia de modelos fotogramétricos obtidos com VANTs (James & Robson, 2014; Harwin & Lucieer, 2012; Tonkin et al., 2014). Essas pesquisas concentram-se principalmente na disposição horizontal e na densidade dos pontos, mostrando que arranjos bem distribuídos reduzem erros sistemáticos e aumentam a precisão posicional. Entretanto, poucos trabalhos avaliaram de forma detalhada a combinação estratégica entre GCPs no plano e em elevação — uma lacuna relevante para a otimização de levantamentos fotogramétricos.

Ao processar imagens de VANT via SfM-MVS, gera-se um MDS que inclui tanto pontos do terreno quanto elementos acima dele (vegetação e edificações). Para simulações hidrológicas urbanas, pode ser mais apropriado utilizar um Modelo Digital sem Vegetação (MDnV), no qual a vegetação arbórea é removida, mas as edificações são mantidas como barreiras físicas que influenciam a microtopografia e alteram a dinâmica hídrica. Essa abordagem, também denominada *bare-earth with buildings*, apresenta maior realismo na representação do escoamento superficial em ambientes urbanos (Schubert et al., 2008; Fewtrell et al., 2011; Neal et al., 2011), evitando trajetórias de fluxo irreais decorrentes da remoção indiscriminada das construções.

Segundo Gevaert et al. (2018), existem diversas técnicas para conversão de MDS em MDT, que podem ser agrupadas em cinco categorias: filtros morfológicos, densificação progressiva, métodos baseados em superfície, baseados em segmentos e métodos de *deep learning* (DL). Entre os métodos de DL, Na et al. (2020) utilizaram dados combinados de ortomosaico e MDS como entrada para arquiteturas de DL, identificando áreas de vegetação e extraíndo o MDT. Gevaert et al. (2018) também propuseram o uso dessas entradas em uma *fully convolutional network* (FCN) para

classificar pontos de terreno e não terreno. Em ambos os casos, o fluxo de trabalho envolve gerar inicialmente uma nuvem de pontos densa com os pontos de superfície e o ortomosaico, para depois aplicar a filtragem via DL. Jakovljevic et al. (2019) desenvolveram um método de aprendizado profundo para classificação de nuvens de pontos, conseguindo extrair o MDT diretamente a partir desses dados.

Visando acelerar esse processo, esta tese parte da hipótese de que métodos de DL podem ser integrados ao fluxo SfM-MVS para gerar diretamente o MDT, aplicando a filtragem nas imagens individuais antes mesmo da geração da nuvem de pontos densa. Assim, o produto final do SfM-MVS já seria uma nuvem de pontos densa já filtrada, eliminando a necessidade de filtragem posterior.

Apesar dos avanços em sensores e algoritmos, ainda persistem lacunas na avaliação de modelos globais em áreas urbanas brasileiras, no estudo do impacto da distribuição mista de GCPs em elevação e no plano, e na integração de DL diretamente no fluxo SfM-MVS para geração de MDT. Esta tese busca preencher essas lacunas, contribuindo para metodologias mais precisas e eficientes em modelagem hidrológica urbana.

OBJETIVO GERAL

Aprimorar a geração de Modelos Digitais do Terreno em áreas urbanas por meio da integração de deep learning diretamente no fluxo fotogramétrico.

Objetivos específicos

- Determinar a adequação de Modelos Digitais de Superfície globais para mapeamento em ambientes urbanos.
- Avaliar o impacto da quantidade, distribuição espacial e qualidade altimétrica de pontos de controle em elevação na acurácia posicional de produtos cartográficos gerados por VANT em escala local.
- Demonstrar a efetividade da integração de deep learning no fluxo fotogramétrico para obtenção direta de MDT urbano.

ORGANIZAÇÃO

Esta tese apresenta-se estruturada em três capítulos principais. O primeiro capítulo apresenta uma avaliação dos MDS globais no contexto urbano em Campo Grande/MS. O segundo capítulo trata do efeito de utilizar pontos de controle mesclados no plano e em elevação, nos produtos cartográficos gerados por VANT, além de indicar a escala em que podem ser usados. O terceiro capítulo apresenta uma técnica inovadora de filtragem para auxiliar na extração do Modelo Digital do Terreno (MDT), combinando deep learning com fotogrametria SfM-MVS, utilizando dados obtidos por VANT.

REFERÊNCIAS

- COOK, Aaron; MERWADE, Venkatesh. Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. *Journal of Hydrology*, v. 377, n. 1-2, p. 131-142, 2009.
- FEWTRELL, T. J.; DUNCAN, A.; SAMPSON, C. C.; NEAL, J. C.; BATES, P. D. Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. *Physics and Chemistry of the Earth*, v. 36, n. 7-8, p. 281-291, 2011.
- GEVAERT, C. M. et al. A deep learning approach to DTM extraction from imagery using rule-based training labels. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 142, p. 106-123, 2018.
- HARWIN, S.; LUCIEER, A. Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. *Remote Sensing*, v. 4, n. 6, p. 1573-1599, 2012.
- HO, L. T. K.; UMITSU, M.; YAMAGUCHI, Y. Flood hazard mapping by satellite images and SRTM DEM in the Vu Gia–Thu Bon alluvial plain, Central Vietnam. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, v. 38, parte 8, p. 275-280, 2010.
- IPCC. *Managing the risks of extreme events and disasters to advance climate change adaptation*. In: FIELD, C. B.; BARROS, V.; STOCKER, T. F.; QIN, D. et al. (ed.). Cambridge: Cambridge University Press, 2012. p. 582.
- JAKOVLJEVIC, Gordana et al. Accuracy assessment of deep learning based classification of LiDAR and UAV points clouds for DTM creation and flood risk mapping. *Geosciences*, v. 9, n. 7, p. 323, 2019.
- JAMES, M. R.; ROBSON, S. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surface Processes and Landforms*, v. 39, n. 10, p. 1413-1420, 2014.

- NA, Jiaming et al. UAV-based terrain modeling under vegetation in the Chinese Loess Plateau: A deep learning and terrain correction ensemble framework. *Remote Sensing*, v. 12, n. 20, p. 3318, 2020.
- NEAL, J.; SCHUMANN, G.; FEWTRELL, T.; BUDIMIR, M.; BATES, P.; MASON, D. Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. *Journal of Flood Risk Management*, v. 4, n. 4, p. 295–310, 2011.
- O'LOUGHLIN, Fiachra E. et al. A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sensing of Environment*, v. 182, p. 49-59, 2016.
- SCHUBERT, J. E.; SANDERS, B. F.; SMITH, M. J.; WRIGHT, N. G. Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. *Advances in Water Resources*, v. 31, n. 12, p. 1603–1621, 2008.
- SMITH, Mark William; CARRIVICK, Jonathan L.; QUINCEY, Duncan J. Structure from motion photogrammetry in physical geography. *Progress in Physical Geography*, v. 40, n. 2, p. 247-275, 2016.
- TONKIN, Toby N. et al. The potential of small unmanned aircraft systems and structure-from-motion for topographic surveys: A test of emerging integrated approaches at Cwm Idwal, North Wales. *Geomorphology*, v. 226, p. 35-43, 2014.
- TREPEKLI, Katerina et al. UAV-borne, LiDAR-based elevation modelling: a method for improving local-scale urban flood risk assessment. *Natural Hazards*, p. 1-29, 2022.
- WANG, Weicai; YANG, Xiaoxin; YAO, Tandong. Evaluation of ASTER GDEM and SRTM and their suitability in hydraulic modelling of a glacial lake outburst flood in southeast Tibet. *Hydrological Processes*, v. 26, n. 2, p. 213-225, 2012.

- WANG, Yong; ZHENG, Tao. Comparison of light detection and ranging and national elevation dataset digital elevation model on floodplains of North Carolina. *Natural Hazards Review*, v. 6, n. 1, p. 34-40, 2005.
- XU, Kepeng et al. The importance of digital elevation model selection in flood simulation and a proposed method to reduce DEM errors: A case study in Shanghai. *International Journal of Disaster Risk Science*, v. 12, n. 6, p. 890-902, 2021.
- YAN, Kun et al. Exploring the potential of SRTM topography and radar altimetry to support flood propagation modeling: Danube case study. *Journal of Hydrologic Engineering*, v. 20, n. 2, p. 04014048, 2015.
- ZHANG, Keqi et al. Accuracy assessment of ASTER, SRTM, ALOS, and TDX DEMs for Hispaniola and implications for mapping vulnerability to coastal flooding. *Remote Sensing of Environment*, v. 225, p. 290-306, 2019.

PRIMEIRO CAPÍTULO: EVALUATION OF ALTIMETRIC ACCURACY OF DIGITAL SURFACE MODELS IN THE URBAN AREA OF CAMPO GRANDE/BRAZIL

Resumo: Os Modelos Digitais de Superfície (MDSs) são utilizados em diversas aplicações, sendo essenciais para modelagem hidrológica. De acordo com a finalidade de interesse, um requisito antes de usar esses conjuntos de dados é verificar sua precisão na área geográfica de estudo. Aqui, descrevemos a precisão altimétrica de diferentes MDSs na área urbana de Campo Grande, Mato Grosso do Sul, Brasil. Como conjunto de dados de referência, usamos dois conjuntos de dados distintos com coordenadas GNSS 3D. Os modelos de MDSs testados foram TanDEM-X DEM, ALOS AW3D30, SRTMc, TOPODATA, SRTM v3 e ASTER GDEM. Estimamos as discrepâncias verticais em relação ao conjunto de dados de referência, usando as seguintes métricas: discrepância vertical mínima e máxima; média das discrepâncias verticais; desvio padrão (DP); raiz do erro quadrático médio (RMSE); e o Padrão de Exatidão Cartográfica dos Produtos Cartográficos Digitais (PEC-PCD). O ALOS AW3D30 apresentou os melhores resultados em termos de RMSE (1,77 m), seguido pelo TanDEM-X DEM (RMSE de 2,80 m). Os demais MDSs apresentaram RMSE variando de 3,6 a 7,5 m. O ALOS AW3D30 e o TanDEM-X DEM mostraram compatibilidade com a escala 1:50.000, enquanto os demais modelos com a escala 1:100.000, conforme PEC-PCD classe A. Demonstramos que tanto o ALOS AW3D30 quanto o TanDEM-X DEM podem ser utilizados para realizar projetos relacionados à modelagem hidrológica em escalas menores ou iguais a 1:50.000. No entanto, MDSs como os modelos SRTMc, TOPODATA, SRTM e ASTER GDEM devem ser adotados em projetos que envolvam escala menor ou igual a 1:100.000.

Palavras-chave: MDS; Padrão de Exatidão Cartográfica; PEC-PCD; Acurácia Vertical

Abstract: Digital Surface Models (DSMs) are used in several applications, being essential in hydrological modeling. According to the task of interest, a requirement before using these datasets is to verify their accuracy in the geographical area of study. Herein, we describe the altimetric accuracy of different DSMs in the urban area of Campo Grande, Mato Grosso do Sul, Brazil. As a reference dataset, we used two distinct datasets with 3D GNSS coordinates. The tested DSMs models were TanDEM-X DEM, ALOS AW3D30, SRTMc, TOPODATA, SRTM v3, and ASTER GDEM. We estimated the vertical discrepancies regarding the reference dataset, using the following metrics: vertical discrepancies minimum and maximum; mean and Standard Deviation (SD) vertical discrepancies; Root Mean Square Error (RMSE); and, the Brazilian standard of cartographic accuracy for digital cartographic products (PEC-PCD). The ALOS AW3D30 presented the better results in terms of RMSE (1.77 m), followed by the TanDEM-X DEM (RMSE of 2.80 m). The others DSMs presented RMSE ranging from 3.6 to 7.5 m. The ALOS AW3D30 and the TanDEM-X DEM showed compatibility with the 1:50,000 scale, while the other models with the 1:100,000 scale, according to PEC-PCD class A. We demonstrate that both ALOS AW3D30 and TanDEM-X DEM can be used to perform projects related to hydrological modeling at scales less or equal to 1:50,000. However, DSMs like models SRTMc, TOPODATA, SRTM, and ASTER GDEM should be adopted in projects involving a scale less or equal to 1:100,000.

Keywords: DSM; Cartographic Accuracy Standard; PEC-PCD; Vertical Accuracy

1.1. INTRODUCTION

The Digital Surface Models (DSMs) are altimetric representations in digital terrain form and are used for several works, such as evaluating landslide susceptibility (Rabby et al., 2020; Arabameri et al., 2019), mapping rainforest environments (Rennó et al., 2008), mapping the groundwater potentiality (Gaber et al., 2020), analysis of hydrological sensitivity for flood risk assessment (Sharma et al., 2018) and, three-dimensional representation of the landscape (Marini et al., 2017). Although the DSM has several practical applications, these data are not always compatible with certain types of tasks, requiring the assessment of their accuracy before its usage.

Many studies have evaluated the altimetric accuracy of DSMs. Falorni et al. (2005) assessed the altimetric accuracy of the DSM SRTM (Shuttle Radar Topography Mission) for two areas in the USA and concluded that the relief has a strong effect on the accuracy of these data. Previous work also assessed the DSMs' accuracy for all of Australia (Hirt et al., 2010), for Northern Greece (Mouratidis et al., 2010), for areas in Peru (Sánchez et al., 2012), for the whole world (Mukulet al., 2016), (Mukulet al., 2017), for a city in India (Patel al., 2016), Istanbul, and Turkey (Alganci al., 2018). Those are examples of work carried out in different parts of the world on the same thematic of assessing the vertical accuracy of DSMs, aiming to check the reliability of these data. In the Brazilian context, examples of this type of work are those carried out by Morais et al. (2017), which evaluated the SRTM, and ASTER GDEM (Global Digital Elevation Model) for the area of Belo Horizonte, Minas Gerais. Franca et al. (2019) evaluated the models' ASTER GDEM and SRTM for the region west of the State of Bahia. Marini et al. (2017) evaluated the vertical accuracy of the SRTM, ASTER GDEM, and TOPODATA models for application in the 3D representation of the Nhicolândia Pantanal.

The DSM TanDEM-X DEM is a product of a 2010 global radar mission, with high spatial resolution, 12 meters. Grohmann (2018) evaluated the TanDEM-X DEM in seven areas, including different morphological contexts, vegetation cover, and land use, in Brazil, performing a visual analysis for comparing it with the SRTM, ASTER GDEM, and ALOS AW3D30. The author concluded that TanDEM-X DEM has a high level of detail and consistency. Moreover, they pointed out that the effective horizontal

resolution of the SRTM is greater than the nominal 30 m, and highlights the errors in ASTER GDEM and ALOS AW3D30 due to the incompatibility between adjacent scenes in the photogrammetric process.

Avtar et al. (2015) evaluated the TanDEM-X DEM in Tokyo. They compared it with global (SRTM, ASTER GDEM, and TanDEM-X DEM) and local (LiDAR and GSI) models, obtaining the lowest value of RMSE (Root Mean Square Error) among the global DSMs. However, to the best of our knowledge, there are no studies that have evaluated the altimetric accuracy of the DSM TanDEM-X DEM urban areas in Brazil. To fulfill this gap, this work evaluates the vertical accuracy of the DSMs, including TanDEM-X DEM, in the urban area of the municipality of Campo Grande, Mato Grosso do Sul, Brazil. The following questions are addressed in this work: “What is the largest cartographic scale that can be adopted in the urban area mapping the municipality of Campo Grande with the DSMs data?”; “What are the magnitude of the discrepancies in the vertical accuracy between the TanDEM-X DEM data and the other DSM (ALOS AW3D30, SRTMc, TOPODATA, SRTM, and ASTER GDEM) in the urban area of the municipality of Campo Grande?”. In urban areas, the DSM is an important source of data for planning; it helps in the identification of areas prone to landslides and flooding, also contributing to hydrological modeling, among other applications.

1.2. MATERIALS AND METHODS

The vertical accuracy of DSMs was evaluated in the urban area of Campo Grande, Mato Grosso do Sul, Brazil considering two datasets as reference collected by 3D GNSS (Global Navigation Satellite System) surveying (section 1.2.1). We evaluated the DSM (section 1.2.2) based on the vertical discrepancies between the 3D GNSS points, and DSMs value, and finally classified the DSMs based on the Brazilian standard of cartographic accuracy for digital cartographic products (PEC-PCD) (section 1.2.3).

1.2.1 Study area and reference data

The study area comprises the urban area of Campo Grande (Figure 1), which is the capital of Mato Grosso do Sul state and has a total area of 8,092.95 km² and 359.03 km² of the urban area (PLANURB 2017).

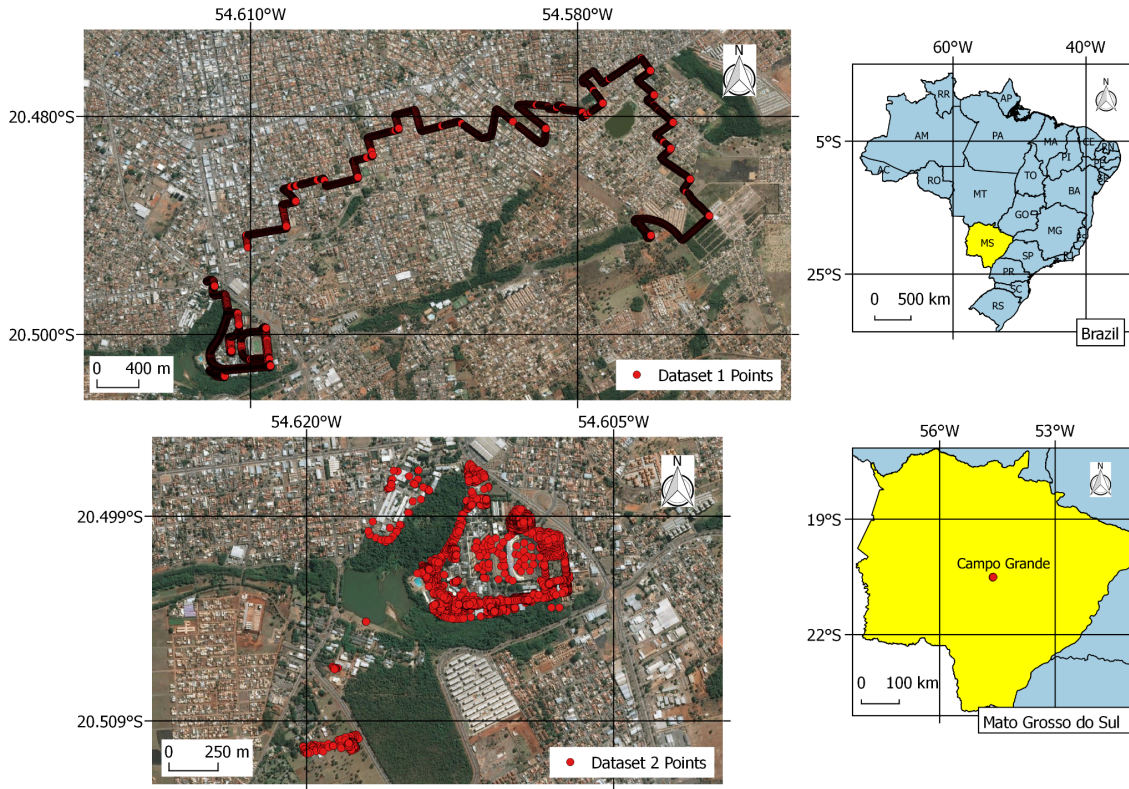


Figure 1. Illustration of the portions of the study area, the Campo Grande city, in Mato Grosso do Sul state, Brazil, and spatial distribution of the GNSS points (in red color) adopted as reference datasets to evaluate the accuracy of the DSMs in this area.

To validate the DSMs, two datasets were used as references. The first dataset with 3D data was collected with a dual-frequency GNSS receiver in kinematic mode (1 second), adapted in a vehicle that traveled the streets of the urban area of Campo Grande. A total of 21,631 reference points were collected, which were obtained using the software Inertial Explorer Version 8.30.1007. The mean standard deviation of planimetry (latitude and longitude) and altimetry (altitude) of the GNSS points was approximately 0.10 and 0.16 m, respectively. After preliminary studies, we verified that in some regions, due to the vehicle stop, the points were concentrated, and this was

removed applying a square filter of 1m, ensuring only one GNSS point per square meter. Finally, 14,907 points were used in the evaluation.

The second dataset has 3D data obtained by surveying with GNSS receivers, using the Real-Time Kinematic (RTK) method, through Trimble R8s equipment. This dataset has distributed points in different environments: vegetation, buildings, and near a lake. The accuracy of these points is 0.05 m for latitude and longitude and 0.10 m for altitude. 3,159 points were raised.

1.2.2 Digital Surface Models

The following models have been validated (Table 1): TanDEM-X DEM, ALOS AW3D30, Bare Earth SRTM (O’Loughlin et al., 2016) - SRTMc; TOPODATA (DSM generated with kriging based on SRTM, for the Brazilian territory) (Valeriano 2005), SRTM v3, and the ASTER GDEM.

Table 1. Main characteristics of the DSMs and links to download the data.

DSM	Spatial Resolution	Altimetric Referential	Link to download
TanDEM-X DEM	12 (m)	Ellipsoid (GRS80)	http://www.eoweb.dlr.de/
ALOS AW3D30	30 (m)	Ellipsoid (GRS80)	http://www.eorc.jaxa.jp/
SRTMc	30 (m)	Geoid (EGM96)	https://data.bris.ac.uk/data/dataset/9d090aba1b9b3fe56dbb597b4b161e9f
TOPODATA	30 (m)	Geoid (EGM96)	http://www.webmapit.com.br/inpe/topodata/
SRTM v3	30 (m)	Geoid (EGM96)	https://earthexplorer.usgs.gov/
ASTER GDEM	30 (m)	Geoid (EGM96)	https://earthexplorer.usgs.gov/

The surface reference for the TanDEM-X DEM and AW3D30 data is the GRS80 ellipsoid, while the others consider the EGM96 geoid. Therefore, a pre-processing stage was necessary to allow the comparison among the models, as to be explained in the following sub-section (1.2.3).

1.2.3 Validation of the DSM

The altitude obtained with the GNSS is referred to as the GRS80 ellipsoid. Thus, the altitudes (vertical coordinate) represented in the TanDEM-X DEM and AW3D30 could be compared directly with the altitude of the reference points (GNSS points). For the other models, however, first, the GNSS altitudes were converted to orthometric altitudes. For this, we used the official Brazilian software - MAPGEO 2015 (Blitzkow et al., 2016), available on the website of the Brazilian Institute of Geography and Statistics (IBGE), which is compatible with the orthometric altitude with altitude represented in the DSMs models, i.e., Geoid (EGM96).

To obtain the altitude of the DSM to each GNSS point, the Point sampling tool of the Open Source GIS QGIS (QGIS 2015) was used, which allows the collection of model data in the same planimetric coordinate (E, N) of the GNSS data. The altitude discrepancy was estimated using the GNSS altitude (H) and the DSM altitude (H_{DSM}). The maximum and minimum discrepancies were verified, while the RMSE (Eq. 1), the mean (Eq. 2), and the standard deviation - SD (Eq. 3) were estimated.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_{DSM_i} - H_i)^2} \quad (1)$$

$$mean = \frac{1}{n} \sum_{i=1}^n (H_{DSM_i} - H_i) \quad (2)$$

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n ((H_{DSM_i} - H_i) - mean)^2} \quad (3)$$

Where i varies from 1 to the number of points (n).

The models were evaluated using the Brazilian cartographic accuracy standard for digital cartographic products (PEC-PCD) (DSG 2011). The DSMs are classified in different scales (1:100,000; 1:50,000; 1:25,000; etc.) and classes A, B, C, and D, according to Table 2.

Table 2. Brazilian cartographic accuracy standard for digital cartographic products (PEC-PCD) values for the scales 1:25,000, 1:50,000 and 1:100,000. Standard Error (SE).

PEC-PCD	1:25,000		1:50,000		1:100,000	
	PEC (m)	SE (m)	PEC (m)	SE (m)	PEC (m)	SE (m)
A	2.70	1.67	5.50	3.33	13.70	8.33
B	5.00	3.33	10.00	6.66	25.00	16.66
C	6.00	4.00	12.00	8.00	30.00	20.00
D	7.50	5.00	15.00	10.00	37.50	25.00

For the DSM to be classified as class A on the 1:50,000 scale, for example, it is necessary to present an RMSE less than or equal to 3.33 m, and 90% of the discrepancies must be less than or equal to 5.50 m. It should be noted that this approach to assess the vertical accuracy of DSM based on the PEC-PCD is consolidated in the literature, considering as reference data points obtained by field survey (Marini et al., 2017; Morais et al., 2017; Franca et al., 2019; Franca and Silva, 2018).

1.3. RESULTS

The results were divided into two sections one with the results of the first reference dataset and another with the results of the second reference dataset.

1.3.1 Results dataset 1

The results of the statistical analyzes applied for the evaluation of the DSMs, with the first dataset, are shown in Table 3 and Figure 2.

Table 3. Minimum and maximum vertical discrepancy, RMSE, Mean and Standard Deviation (SD) for the urban area of Campo Grande, Brazil.

	TanDEM -X DEM	ALOS AW3D30	SRTMc	TOPO- DATA	SRTM v3	ASTER GDEM
Minimum vertical Discrepancy (m)	-25.27	-7.92	-6.25	-4.92	-5.92	-9.96
Maximum vertical Discrepancy (m)	13.25	12.56	15.83	17.16	17.12	29.15
RMSE (m)	2.80	1.77	3.59	4.34	4.35	7.53
Mean (m)	1.69	0.89	2.49	3.82	3.55	5.68
SD (m)	1.63	1.08	2.02	1.62	1.86	3.82

ALOS AW3D30 showed the best result (RMSE 1.77 m), followed by TanDEM-X DEM with an RMSE of 2.80 m. However, it should be considered that the difference between the RMSE of the ALOS AW3D30 and the RMSE of the other DSM was always at least twice the value of the RMSE, except for the TanDEM-X DEM.

The SRTMc showed a result (RMSE of 3.59 m) lower than the SRTM v3 (RMSE of 4.35 m), which was expected since it was generated with the proposal to be a DTM (Digital Terrain Model) (O’Loughlin et al., 2016). SRTM v3 and TOPODATA presented RMSE very close to each other (4.35 and 4.34 m, respectively). It is worth remembering that TOPODATA is also generated from SRTM data (Valeriano 2005). ASTER GDEM had an RMSE of 7.53 m, which corroborates with (Marini et al., 2017; Graf et al., 2018) that also evaluated the DSM ASTER GDEM, which proved to be the model least accurate vertically. Additionally, when we analyze the Boxplot (Figure 2), the ASTER GDEM model presents the highest variability compared to other models. While, the ALOS AW3D30 model showed the least variation between discrepancies, with a median close to zero, and with most of the discrepancies between 10 m and -10 m.

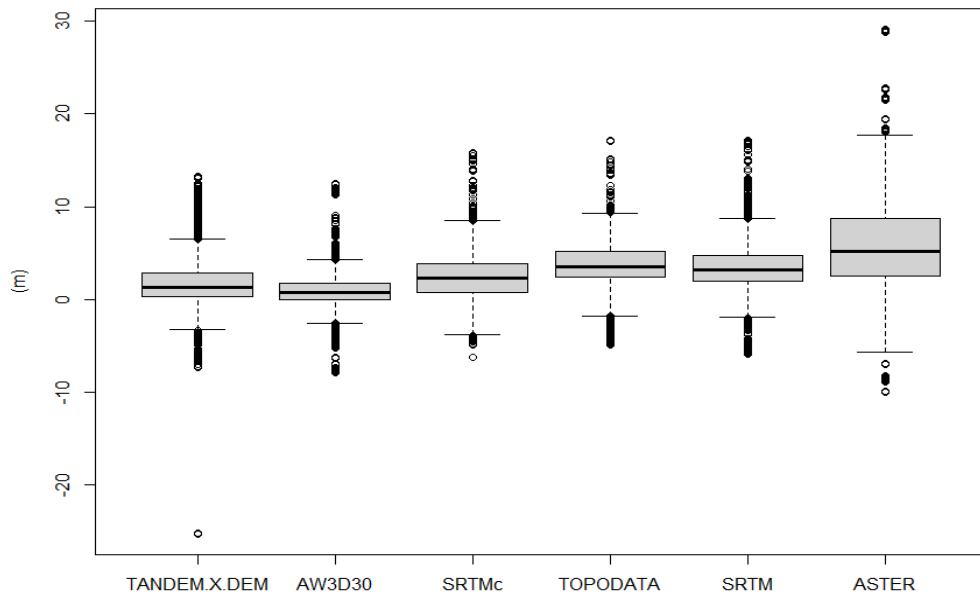


Figure 2. Boxplot of DSM discrepancies regarding the 3D GNSS points for the urban area of Campo Grande, Brazil.

We verified that the TanDEM-X DEM presented an isolated outlier, which was presented in Table 3 (minimum discrepancy of -25 m). This outlier is in a single-pixel (Figure 3), which had one point collected with GNSS. The same situation was not observed for the others DSM as the presented minimum discrepancies were between -10 and -5 m, approximately (Figure 2). SRTMc, TOPODATA, and SRTM v3 obtained similar boxes, SRTMc with a lower median but with more distribution of discrepancies. ASTER is the median furthest from zero and with the greatest dispersion of results.



Figure 3. Location of the minimum discrepancy point (yellow) for the DSM TanDEM-X DEM, (a) in an orthorectified image, (b) in the TanDEM-X DEM model.

For the maximum discrepancies (Table 3), the ALOS AW3D30 had the lowest value, 12.56 m, followed by the TanDEM-X DEM (13.25 m) and the SRTMc (15.84 m). The ASTER GDEM model has presented the worst results being the maximum discrepancy of 29 m approximately. The results of the altimetric assessment of the DSMs for the urban area of Campo Grande based on the PEC-PCD are presented in Tables 4 and 5.

Table 4. Classification according to the PEC-PCD, considering the Standard Error for class A.

Scale	Standard Error (m)	RMSE (m)					
		TanDEM-X DEM	ALOS AW3D30	SRTMc	TOPODATA	SRTM	ASTER GDEM
1:25,000	1.67						
1:50,000	3.33	2.80	1.77				
1:100,000	8.33			3.59	4.34	4.35	7.53

Table 5. Percentage of points with discrepancy below the PEC-PCD, for class A.

Scale	PEC - PCD (m)	TanDEM- X DEM (%)	ALOS AW3D30 (%)	SRTMc (%)	TOPOD ATA (%)	SRTM (%)	ASTER GDEM (%)
1:25,000	2.70	71.81	90.70	55.58	29.72	38.74	23.51
1:50,000	5.50	93.88	98.93	86.40	79.04	80.23	51.61
1:100,000	13.70	99.96	100	99.89	99.93	99.88	94.56

We observed that the TanDEM-X DEM and ALOS AW3D30 were compatible with the scale 1:50,000, class A, according to the standard error criteria (RMSE 2.80 and 1.77) and percentage of points below the PEC-PCD (93.88 and 98.93), respectively. By the percentage criterion, the ALOS AW3D30 could be classified on a larger scale; however, the RMSE, although close, prevents this classification. The other evaluated DSMs are classified on the scale of 1:100,000, class A, for the area under study.

1.3.2 Results dataset 2

The results of the statistical analyzes applied for the evaluation of the DSMs, with the second reference dataset, are shown in Table 6 and Figure 4.

Table 6. Minimum and maximum vertical discrepancy, RMSE, Mean and Standard Deviation (SD) for the urban area of Campo Grande, Brazil.

	TanDEM -X DEM	ALOS AW3D30	SRTMc	TOPO- DATA	SRTM v3	ASTER GDEM
Minimum vertical Discrepancy (m)	-25.40	-4.34	-7.89	-3.85	-7.49	-5.70
Maximum vertical Discrepancy (m)	11.77	11.13	9.13	10.86	13.94	29.02
RMSE (m)	2.68	1.57	2.99	4.47	4.20	7.75
Mean (m)	1.31	0.33	1.61	3.73	3.03	6.04
SD (m)	2.34	1.54	2.51	2.46	2.91	4.86

As shown in Table 6, the RMSE of the ALOS AW3D30 is the smallest (1.57 m), followed by the TanDEM-X DEM (2.68 m). For this Dataset the ALOS AW3D30 has RMSE twice as low compared to the models except for the TanDEM-X DEM and SRTMc (2.99 m).

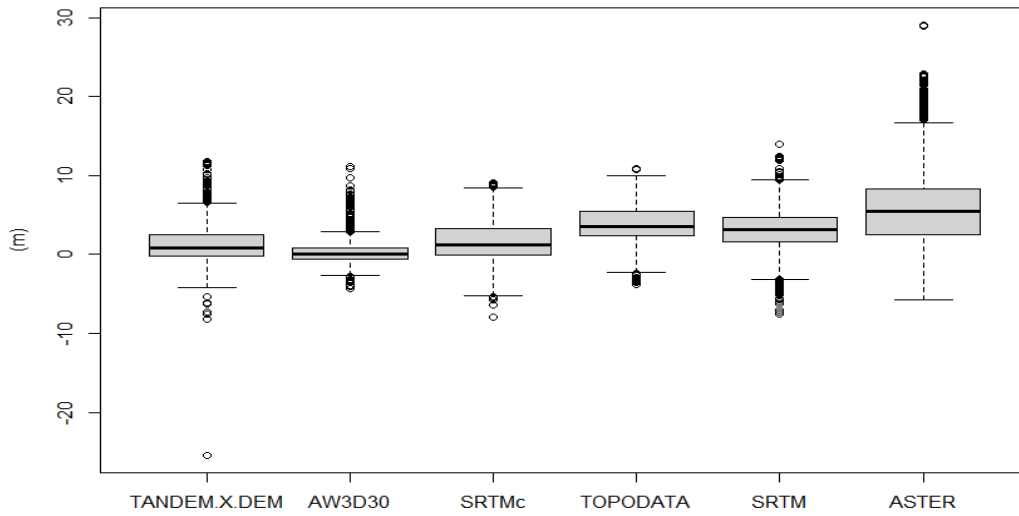


Figure 4. Boxplot of DSM discrepancies regarding the 3D GNSS RTK points for the urban area of Campo Grande, Brazil.

The distribution of discrepancies was observed in figure 4, with the ALOS AW3D30 with the smallest box, indicating that the discrepancies are uniform and close to zero, as well as the TanDEM-X DEM, whose box is shifted to values above zero, this is observed with the median above zero and the box shifted to values above the median. For the SRTMc, TOPODATA, and SRTM models, the boxes have similar dimensions, although the SRTMc with the Box is closer to zero and the SRTM with the more compact box.

Table 7. Classification according to the PEC-PCD, considering the Standard Error for class A.

Scale	Standard Error (m)	RMSE (m)					
		TanDEM-X DEM	ALOS AW3D30	SRTMc	TOPODATA	SRTM	ASTER GDEM
1:25,000	1.67		1.57				
1:50,000	3.33	2.68		2.99			
1:100,000	8.33				4.47	4.20	7.75

Table 8. Percentage of points with discrepancy below the PEC-PCD, for class A.

Scale	PEC - PCD (m)	TanDEM-X DEM (%)	ALOS AW3D30 (%)	SRTMc (%)	TOPODATA (%)	SRTM (%)	ASTER GDEM (%)
1:25,000	2.70	76.95	92.66	65.37	31.97	39.92	26.12
1:50,000	5.50	94.62	98.39	92.85	75.75	81.96	50.97
1:100,000	13.70	99.97	100	100	100	99.94	92.91

According to tables 7 and 8, the ALOS AW3D30 can be applied at 1:25,000 scale according to the PEC-PCD, since it presented RMSE (1.57 m) lower than the standard error limit for the class (1.67 m) and more than 90% of the data are below the error limit of the PEC-PCD of 2.7 m, of the evaluated models the only one with such classification.

TanDEM-X DEM and SRTMc were rated for the scale of 1:50,000 according to the PEC-PCD class A. With RMSE 2.68 and 2.99 m respectively being the standard error accepted for class up to 3.33 m and more than 90% of data below the limit of 5.50 m. The other models fit the 1:100,000 scale according to the PEC-PCD class A.

1.4. DISCUSSION

A DSMs is used as input data to attend different applications, such as in the studies involving relief mapping in urban areas, which requires data on a larger cartographic scale. Evaluating then a DSM's altimetric accuracy, before its adoption, is a project requirement. Here we characterized the altimetric accuracy of different DSMs in the urban area of Campo Grande using two datasets. Additionally, we identified the recommended scale based on the PEC-PCD criterion. ALOS AW3D30 and TanDEM-X DEM are the best free options, being indicated for supporting studies at a scale of 1:50,000.

Using two distinct reference datasets, one with a more linear points distribution and the other with a circular distribution, we noted that the RMSE values and discrepancies in height are similar to the ALOS AW3D30 model and the TanDEM-X DEM. However, the values of the PEC-PCD for the second reference dataset, are even better, indicating that the ALOS AW3D30 can be used to map on the scale of 1:25,000. Another model that improved the result was the SRTMc, which was classified on the scale of 1:50,000. We believe this is due to the lower number of points 3,159 in the second dataset against 14,907 in the first dataset.

Kramm and Hoffmeister (2019), and Grohmann (2019) argue that, depending on the topography of the study area, the ALOS AW3D30 presents just as well as the TanDEM-X DEM. To demonstrate the accuracy of the TanDEM-X DEM and ALOS AW3D30, Kramm and Hoffmeister (2019) concluded that they could represent a landscape with the same precision as very high-resolution DSMs with a 5 m GSD.

According to Avtar et al. (2015), the estimated RMSE for TanDEM-X is considerably lower than SRTM and ASTER GDEM.

Most of the vertical accuracy assessment works use data from LiDAR (Moraes al., 2017; Grohmann et al., 2019; Kramm and Hoffmeister 2019), as well as high-resolution photogrammetry data (Alganci al., 2018; Kramm and Hoffmeister 2019). In the absence of this type of data, we have works that use GNSS data (Franca et al., 2019). Due to the number of GNSS points, with good precision and the two datasets presents similar results, we believe that this has been compensated.

1.5. CONCLUSIONS

A comparative analysis was developed on the altimetric accuracy of different DSMs for the urban area of the municipality of Campo Grande, Brazil, using more than 18 thousand GNSS points. The TanDEM-X DEM has higher altimetric accuracy than the traditional DSM (SRTM, ASTER GDEM). In this study, the freely available DSM ALOS AW3D30 was the one with the highest altimetric accuracy. However, its results are still similar to those obtained for the TanDEM-X DEM in terms of PEC-PCD classification.

For the municipality of Campo Grande, it is concluded that the TanDEM-X DEM and ALOS AW3D30, are compatible with the scale 1:50,000, and the models SRTMc, TOPODATA, SRTM v3, and ASTER GDEM with the 1:100,000, when considering the PEC-PCD, class A. It is important to highlight that in this work, it was used a large number of GNSS points were collected for the validation of the DSMs. The quantity used is much higher than that adopted in many studies for the validation of digital surface models.

1.6. REFERENCES

- Agência Municipal de Meio Ambiente e Planejamento Urbano – PLANURB. **Perfil Socioeconômico de Campo Grande**. 24. ed. rev. Campo Grande, Brazil, 2017, pp. 1 -100.
- Alganci, U., Besol, B., & Sertel, E. (2018). Accuracy assessment of different digital surface models. **ISPRS International Journal of Geo-Information**, 7(3), 114.
- Arabameri, A., Pradhan, B., Rezaei, K., & Lee, C. W. (2019). Assessment of landslide susceptibility using statistical-and artificial intelligence-based FR–RF integrated model and multiresolution DEMs. **Remote Sensing**, 11(9), 999.
- Avtar, R., Yunus, A. P., Kraines, S., & Yamamuro, M. (2015). Evaluation of DEM generation based on Interferometric SAR using TanDEM-X data in Tokyo. **Physics and Chemistry of the Earth, Parts A/B/C**, 83, 166-177.
- Blitzkow, D., de Matos, A. C. O. C., Xavier, E. M. L., & Fortes, L. P. S. (2016). MAPGEO2015: O novo modelo de ondulação geoidal do brasil. **Revista Brasileira de Cartografia**, (68/10), 1873-1884.
- DSG - Diretoria de Serviço Geográfico. Especificação Técnica para a Aquisição de Dados Geoespaciais Vetoriais - ET-ADGV. 2 ed. 2011. 254p. Disponível em <http://www.geoportal.eb.mil.br/images/PDF/ET_ADGV_Vs_2_1_3.pdf>. (accessed on 18 jun. 2018).
- Falorni, G., Teles, V., Vivoni, E. R., Bras, R. L., & Amaratunga, K. S. (2005). Analysis and characterization of the vertical accuracy of digital elevation models from the Shuttle Radar Topography Mission. **Journal of Geophysical Research: Earth Surface**, 110(F2).
- França, L. L. S. D., & Silva, L. F. C. F. D. (2018). Comparison between the Double Buffer Method and the Equivalent Rectangle Method for the quantification of discrepancies between linear features. **Boletim de Ciências Geodésicas**, 24, 300-317.
- França, L. L. S. D., Penha, A. D. L. T. D., & Carvalho, J. A. B. D. (2019). Comparison between Absolute and Relative positional accuracy assessment-a case study applied to Digital Elevation Models. **Boletim de Ciências Geodésicas**, 25.
- Gaber, A., Mohamed, A. K., ElGalladi, A., Abdelkareem, M., Beshr, A. M., & Koch, M. (2020). Mapping the groundwater potentiality of West Qena Area, Egypt, using integrated remote sensing and hydro-geophysical techniques. **Remote Sensing**, 12(10), 1559.
- Graf, L., Moreno-de-las-Heras, M., Ruiz, M., Calsamiglia, A., García-Comendador, J., Fortesa, J., ... & Estrany, J. (2018). Accuracy assessment of digital terrain model dataset sources for hydrogeomorphological modelling in small mediterranean catchments. **Remote Sensing**, 10(12), 2014.

- Grohmann, C. H. (2018). Evaluation of TanDEM-X DEMs on selected Brazilian sites: Comparison with SRTM, ASTER GDEM and ALOS AW3D30. **Remote Sensing of Environment**, 212, 121-133.
- Hirt, C., Filmer, M. S., & Featherstone, W. E. (2010). Comparison and validation of the recent freely available ASTER-GDEM ver1, SRTM ver4. 1 and GEODATA DEM-9S ver3 digital elevation models over Australia. **Australian Journal of Earth Sciences**, 57(3), 337-347.
- Kramm, T., & Hoffmeister, D. (2019). A relief dependent evaluation of digital elevation models on different scales for northern Chile. **ISPRS International Journal of Geo-Information**, 8(10), 430.
- Marini, L. B., Junior, J. M., Ramos, A. P. M., Paranhos Filho, A. C., Barros, W. M., & Higa, L. T. (2017). Análise da acurácia altimétrica dos modelos digitais de superfície SRTM, ASTER e TOPODATA e aplicação na representação 3D do Pantanal da Nhecolândia. **Anuário do Instituto de Geociências**, 40(3), 48-54.
- Morais, J. D., Faria, T. S., Elmiro, M. A. T., Nero, M. A., Silva, A. D. A., & Nobrega, R. A. D. A. (2017). Altimetry assessment of ASTER GDEM v2 and SRTM v3 digital elevation models: a case study in urban area of Belo Horizonte, MG, Brazil. **Boletim de Ciências Geodésicas**, 23, 654-668.
- Mouratidis, A., Briole, P., & Katsambalos, K. (2010). SRTM 3 "DEM (versions 1, 2, 3, 4) validation by means of extensive kinematic GPS measurements: a case study from North Greece. **International Journal of Remote Sensing**, 31(23), 6205-6222.
- Mukul, M., Srivastava, V., & Mukul, M. (2016). Accuracy analysis of the 2014–2015 Global Shuttle Radar Topography mission (SRTM) 1 arc-sec C-Band height model using international Global Navigation Satellite System Service (IGS) network. **Journal of Earth System Science**, 125(5), 909-917.
- Mukul, M., Srivastava, V., Jade, S., & Mukul, M. (2017). Uncertainties in the shuttle radar topography mission (SRTM) Heights: Insights from the indian Himalaya and Peninsula. **Scientific reports**, 7(1), 1-10.
- O'Loughlin, F. E., Paiva, R. C., Durand, M., Alsdorf, D. E., & Bates, P. D. (2016). A multi-sensor approach towards a global vegetation corrected SRTM DEM product. **Remote Sensing of Environment**, 182, 49-59.
- Patel, A., Katiyar, S. K., & Prasad, V. (2016). Performances evaluation of different open source DEM using Differential Global Positioning System (DGPS). **The Egyptian Journal of Remote Sensing and Space Science**, 19(1), 7-16.
- QGIS Development Team, 2015. QGIS Geographic Information System. Versão 2.8 Wien. Open Source Geospatial Foundation Project. <<http://qgis.osgeo.org>>.
- Rabby, Y. W., Ishtiaque, A., & Rahman, M. (2020). Evaluating the effects of digital elevation models in landslide susceptibility mapping in rangamati district, bangladesh. **Remote Sensing**, 12(17), 2718.

- Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., & Tomasella, J. (2008). HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. **Remote Sensing of Environment**, 112(9), 3469-3481.
- Sánchez, E. P., & de Villarán, R. F. (2012). SRTM 3" comparison with local information: Two examples at national level in Peru. **Journal of Applied Geodesy**, 6(2), 75-81.
- Sharma, S. K., Kwak, Y. J., Kumar, R., & Sarma, B. (2018). Analysis of hydrological sensitivity for flood risk assessment. **ISPRS International Journal of Geo-Information**, 7(2), 51.
- Valeriano, M. D. M. (2005, April). Modelo digital de variáveis morfométricas com dados SRTM para o território nacional: o projeto TOPODATA. In **XII Simpósio Brasileiro de Sensoriamento Remoto** (Vol. 12, pp. 1-8). São José dos Campos: INPE.

SEGUNDO CAPÍTULO: ACCURACY OF HIGH RESOLUTION DIGITAL CARTOGRAPHIC PRODUCTS WITH ELEVATION CONTROL POINTS

Resumo: O uso de Veículos Aéreos Não Tripulados (VANTs) como ferramenta para aquisição de imagens tem sido aplicado em diversas áreas, algumas dessas aplicações requerem produtos cartográficos com alta acurácia. Com isso, surge a necessidade de planejamento para a aquisição das imagens e para a distribuição de pontos de controle em solo (GCP), os quais serão utilizados na correção geométricas destas, visando conferir às cenas o nível de acurácia exigido. Este trabalho investiga se a quantidade de pontos de controle tal como sua distribuição em diferentes planos de altitude em terreno elevado interfere na acurácia dos produtos cartográficos gerados. Foram utilizadas imagens RGB registradas por uma câmera com resolução de 20 MP embarcada em um VANT multirrotor. As imagens foram capturadas com uma sobreposição de 80% (longitudinal e lateral) e GSD estimado em 0,017 m. A área sobrevoada de 5,5 ha contou com 31 alvos levantados com GNSS RTK, sendo 21 definidos como pontos de checagem (CP) e 12 como pontos de controle em solo (GCP), os quais foram utilizados no processamento de imagens para geração dos ortomosaicos. Avaliamos a acurácia dos produtos com base na PEC-PCD. Os resultados evidenciaram que ao utilizar somente 2 GCPs os erros altimétricos são altos, sendo a única configuração que não se enquadrou na PEC-PCD escala 1:1.000 classe A. Com 5 GCPs obtivemos o melhor RMSE em altimetria (0,026 m). Com 6 GCPs obtivemos o melhor RMSE em planimetria (0,046 m). A altimetria se mostra a parte mais sensível na geração de produtos cartográficos e a utilização de GCPs em elevação melhora a acurácia altimétrica.

Palavras-chave: VANT; GCP; PEC-PCD

Abstract: The use of Unmanned aerial vehicles (UAVs) as a tool for image acquisition has been applied in several fields, some applications require cartographic products with high accuracy. With this comes the need for planning the acquisition of images and distribution of control points (GCP) so that digital products meet the required level of accuracy. The aim of this work was to investigate whether the quantity of control points as well as their distribution in different altitude planes in elevated ground can improve the accuracy of the generated cartographic products. RGB images captured by an onboard camera with a resolution of 20 MP were used. Images were captured by a multirotor UAV with an overlap of 80% (front and side) and estimated GSD of 0.017 m. The surveyed area of 5.5 ha overflowed area had 31 targets surveyed with GNSS RTK, 21 defined as checkpoints (CP) and 12 as ground control points (GCP), which were used in image processing to generate orthomosaic. We evaluated the accuracy of the generated products based on the PEC-PCD. The results showed that when using only 2 GCPs the altimetric errors are high, being the single configuration that did not fit the PEC-PCD scale 1:1,000 class A. With 5 GCPs we obtained the best RMSE in altimetry (0.026 m). With 6 GCPs we obtained the best RMSE in planimetry (0.046 m). Altimetry is the most sensitive aspect in generating cartographic products, and the use of GCPs in elevation improves altimetric accuracy.

Keywords: UAV; GCP; PEC-PCD

2.1. INTRODUCTION

Unmanned aerial vehicles (UAVs) have become popular in recent decades because of their relatively low cost in data acquisition compared to aerial surveys performed with a metric camera, or even by high spatial resolution satellites. There has been considerable advancement in technology in this sector, both in terms of hardware, such as more robust multicopter, fixed and mixed wing UAVs, lighter and higher resolution cameras, and in the software set, such as flight plan applications and photogrammetric programs using computer vision, enabling the generation of high spatial resolution products at a lower cost.

The images captured by sensors embedded in UAVs are being used for studies in various areas, such as plant count and detection of planting lines (Osco et al. 2021), estimation of pasture biomass (Batistoti et al. 2019), analysis of glacier dynamics (Lewińska et al. 2021), 3D mapping of open pit mines (Le Van et al. 2020), quantification of soil erosion (Meinen & Robinson 2020), among others. The 3D position of the camera at the time of photo capture is an important factor for georeferencing the products generated with UAV. With the absolute position and orientation of the camera, accurate products can be obtained without the need for ground control points (GCPs) (Le Van et al. 2020). However, most UAVs are equipped with low accuracy Global Navigation Satellite System (GNSS) receivers (Han et al. 2020). There are studies with high-accuracy embedded GNSS that recommend using at least one GCP to achieve centimetric accuracy (Le Van et al. 2020, Padró et al. 2019, Revuelto et al. 2021).

There are several flight planning factors that influence the quality of products generated with UAV images. Garcia & Oliveira (2021) performed an experiment evaluating some of these factors, concluding that the flight height (21, 31 and 40 m) did not show any difference in the quality of the products generated. Regarding the image overlap, if it falls below 80% in the lateral direction and 60% in the longitudinal direction, the orthomosaic may exhibit noise and lower quality details. The adjustments of the camera parameters and the amount of GCPs affected the generated products accuracy, performing manually the previous adjustment of the camera parameters it is possible to use only 5 GCPs, while letting the software calculate the parameters is needed than 8 GCPs.

The issue of the quantity and distribution of GCPs has been addressed by several works. LeVan et al. (2020) researched open pit mine mapping with DJI Phantom 4 RTK using 0, 1 and 2 GCPs and reported a Root Mean Squared Error (RMSE) variation in the altimetric component of 0.844 m, 0.043 m and 0.031 m, respectively. Calou et al. (2021) verified the quality of the orthomosaics using 3, 5 and 8 GCPs, obtaining RMSE values for the altitude of 0.033 m, 0.026 m and 0.0044 m, respectively. Yu et al. (2020) evaluated, in a small (7 ha), medium (39 ha) and large (342 ha) area, the GPCs variation of 3, 6, 9, 12, 15 and 18, and recommended for altimetry the use of at least 12 GCPs for small, 12 GCPs for medium areas, and more than 18 GCPs for large areas, with the RMSE for the altitude of these recommendations being 0.045 m, 0.038 m and 0.067 m.

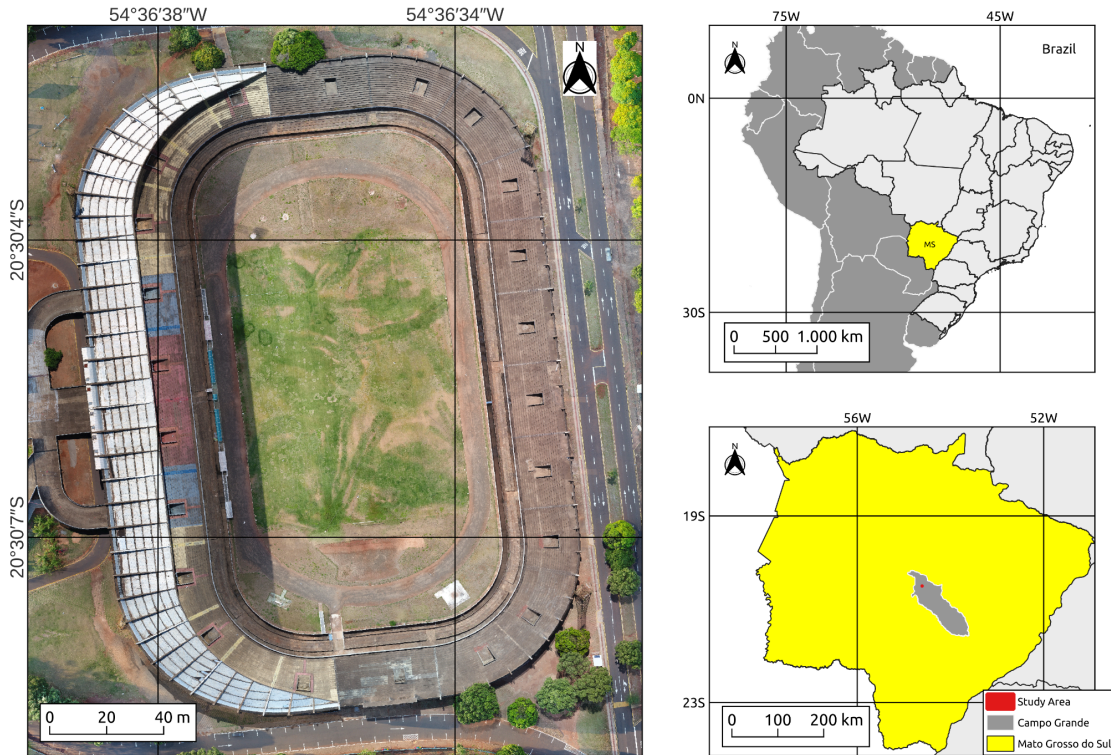
The aforementioned works, as well as that of James et al. (2017), reveal that altitude is the most sensitive variable to the amount of GCPs. However, there are few studies evaluating the mixture of control points in elevation and in the plane, and this is the gap that we intend to fill in this study. That said, the objective of this work is to investigate whether the quantity and spatial arrangement of control points located on flat and elevated terrain interfere with the accuracy of cartographic products generated from images collected by sensors embedded in UAVs. The cartographic products generated are orthomosaic and the Digital Surface Model (DSM).

2.2. METHODOLOGY AND DATA

2.2.1. Area of study

A set of 468 images, with resolution of 4864 x 3648 pixels, from a region of approximately 5.5 ha of the University stadium Pedro Pedrossian de Campo Grande, Mato Grosso do Sul, Brazil was used to fulfill the objective of this work (Figure 1). The images have a Ground Sample Distance (GSD) of 0.0172 m. The study area was selected due to the ease of access and the possibility of using GCPs with different heights, that is, points located in the stadium stands.

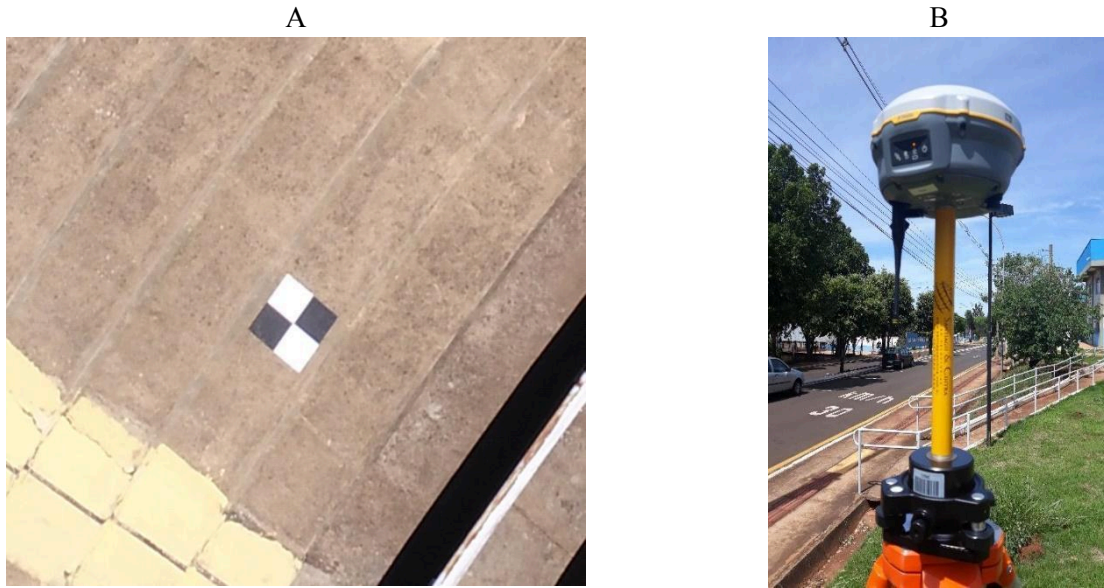
Figure 1 – Location of study area, Pedro Pedrossian University Stadium, Campo Grande/MS, Brazil.



2.2.2. Data Acquisition

Thirty-one targets were distributed in the study area (Figure 2A): twenty-one were used for checkpoints, and ten were applied as ground control points in different arrangements, from which the three-dimensional coordinates were collected, using a Trimble RTK GNSS R8s receiver. (Figure 2B). The 3D accuracy of this type of survey is approximately 30 mm, with the horizontal RTK survey in the order of 8 mm + 0.1 ppm, and vertical 15 mm + 1 ppm (Trimble, 2013). With the targets still on the ground, the area was overflited with a DJI Phantom 4 – Advanced UAV, which has a 20 MP on-board camera. With the help of the flight plan application Pix4Dcapture, the overflight area, the lateral and longitudinal overlaps of 80%, and the flight height of 70 m were determined, with the camera positioned towards the nadir.

Figure 2 – Equipment used to obtain three-dimensional coordinates: A Targets used to materialize control (GCP) and check (CP) points; B Trimble GNSS RTK R8s.



2.2.3. Quantity of Control Points

With the collected points, the ones to be used as GCPs were selected. For this purpose, the premise was to have well-distributed points in the study area, as Zanetti et al. (2017) found that clustered control points result in less accurate cartographic products. The quantity of distributed points was evaluated, ranging from 2 to 6 GCPs, as shown in Figure 3.

Figure 3 – Configurations considered regarding the distribution of control points on the ground, with 2, 3, 4, 5, and 6 GCPs.



2.2.4. Control Points at Different Elevations

To investigate whether the significant difference in elevation between points, for example, points on the ground (soccer field of the stadium) and points at higher elevations (in the stadium stands), interferes with the geometric quality of the generated orthomosaic, a comparison of the 3D accuracy of the orthomosaic was conducted using different GCPs located in flat areas and at higher elevations (Figure 4). It's worth noting that the GCPs in flat areas were placed directly on the ground, while the GCPs at higher elevations were a mixture of GCPs located in both flat areas and elevated sections (bleachers). The distribution can be seen in Figure 5. The leave-one-out methodology was used, where one GCP is removed and the processing is performed, repeating this process until all GCPs have been excluded. Since 6 GCPs were used, the images were processed 6 times for the GCPs in flat areas and 6 times for the GCPs with some at higher elevations, with one GCP being left out in each processing. This methodology ensures that there are no gross errors in the reference data (GCPs), which increases the reliability of the final result.

Figure 4 – Difference in elevation between GCPs on the ground (blue) and at higher elevations (red).

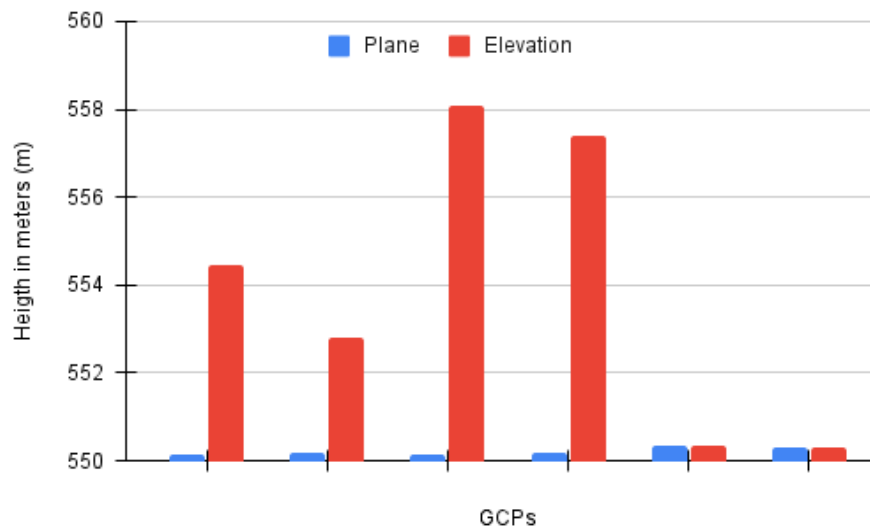
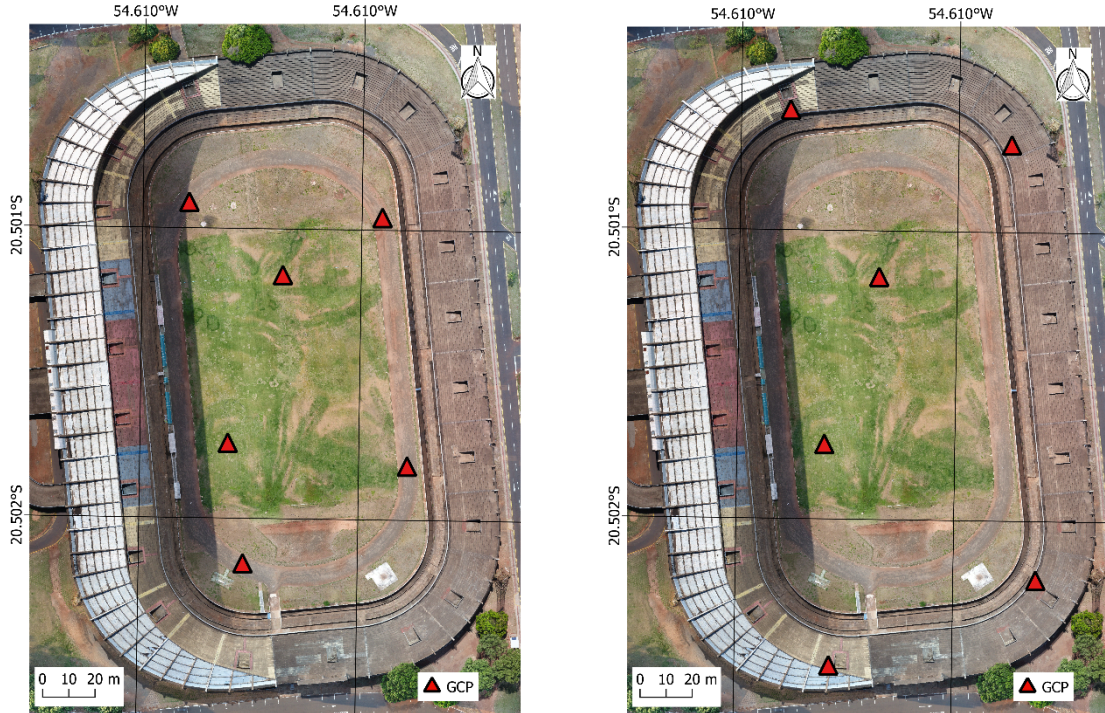


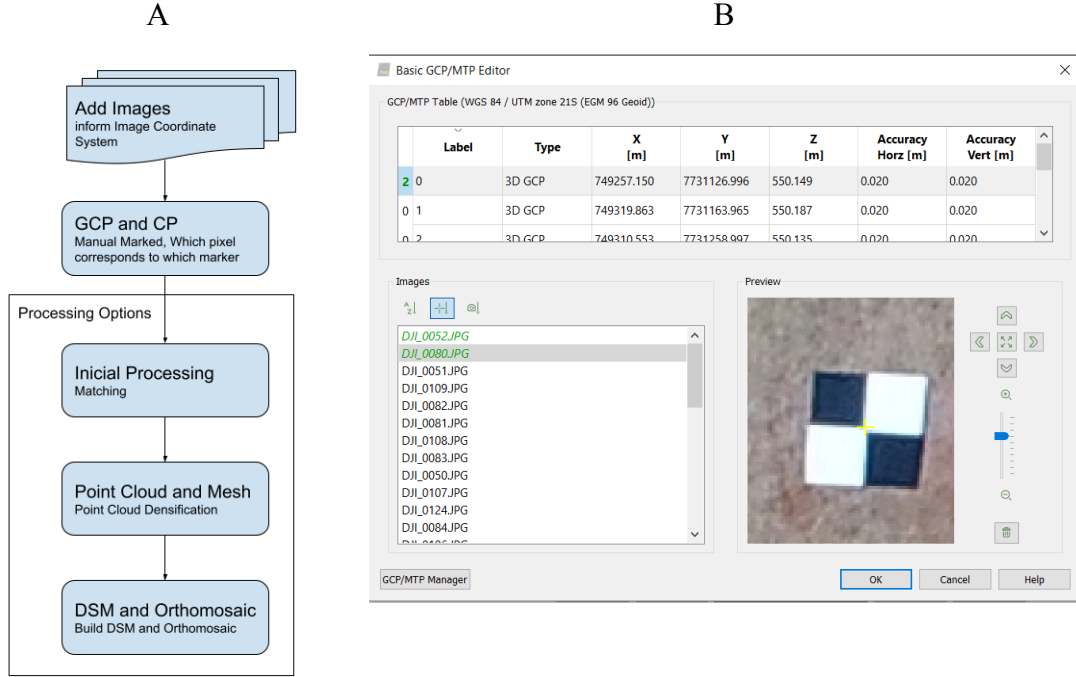
Figure 5 – Distribution of control points: A: In flat terrain; B: With points at higher elevations.



2.2.5. Generation of Cartographic Products

The processing in Pix4Dmapper (v.4.1.25) for the generation of cartographic products (DSM and Orthomosaic) followed the flowchart outlined in Figure 6A. The coordinate system used in the project was WGS 84 / UTM zone 21S (EGM96). In Figure 6B, the process of indicating which pixel in the image corresponds to which control point can be observed. This marking is done manually, and all control points, both GCPs and CPs, should be marked. Processing options: Pix4D allows for changing some parameters for processing. For this experiment, we kept the default values for all processing steps. It separates the processing into three stages: Initial processing (basically matching the images, i.e., finding common points between the images), point cloud and mesh (based on the homologous points found in the previous stage, it densifies the point cloud, i.e., generates coordinates for each point in the image), and DSM and Orthoimage (In this stage, the DSM and orthoimage are generated from the data generated in the previous stage).

Figure 6 – Pix4D Processing: A. Process Flowchart; B. Manual process of marking control points (GCP and CP). Window that is used to indicate which pixel in the image corresponds to the GCP.



2.2.6. Validation of Digital Cartographic Products

To validate the experiments, 21 CP were used, although these points were not used in the processing (Figure 7). After processing the images in the Pix4Dmapper application (v.4.1.25), we exported the orthophoto and digital surface model on the QGIS application (QGIS DEVELOPMENT TEAM, 2018), where we collected the coordinates where the orthophoto indicates the target. Then, we collected the altitude of the same point in the DSM. With these three-dimensional coordinates from the orthophoto and DSM, we compared them with the coordinates collected in the field using RTK. The planimetric and altimetric discrepancies were calculated according to Equation 1 and Equation 2. Using the discrepancies, we calculated the RMSE (Equation 3), mean, and standard deviation (SD).

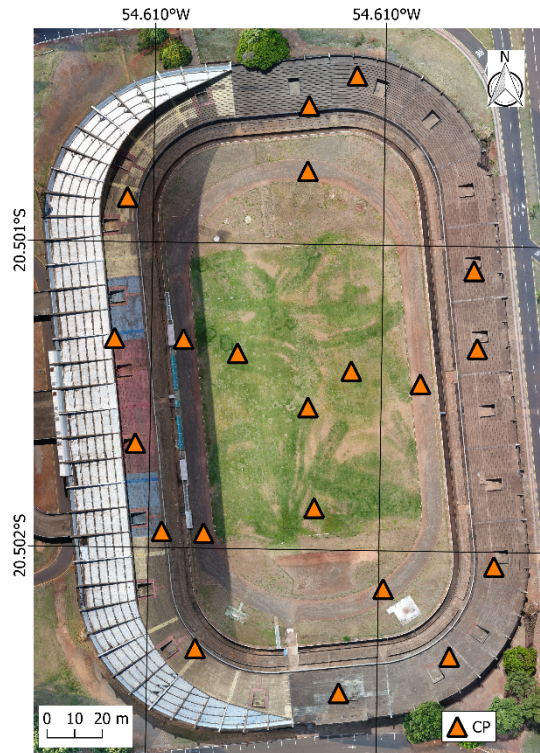
$$\Delta S = \sqrt{(X_i - X_{RTK_i})^2 + (Y_i - Y_{RTK_i})^2} \quad (1)$$

$$\Delta H = Z_i - ZRTK_i \quad (2)$$

$$RMSE = \sqrt{\frac{\sum Discrepância^2}{n}}$$

Where ΔS represents the planimetric discrepancies, ΔH represents the altimetric discrepancies, X_i , Y_i are the UTM coordinates (Easting and Northing) obtained from the orthophoto, Z_i represents the altitudes obtained from the DSM, $XRTK_i$, $YRTK_i$, $ZRTK_i$ are the UTM coordinates (Easting, Northing, and altitude) obtained from the RTK GNSS survey, n is the number of CP evaluated.

Figure 7 – Distribution of checkpoints (CP).



The digital products were classified according to the Cartographic Accuracy Standard for Digital Cartographic Products (PEC-PCD) (DSG, 2016). The orthophoto was classified according to the PEC-PCD for planimetry, and the DSM was classified according to the PEC-PCD for altimetry. These classifications categorize the digital products into different scales (1:1,000, 1:2,000, 1:5,000) and classes (A, B, C, and D).

Since the generated products have high spatial resolution, they were evaluated to determine if they fit into the 1:1,000 scale, Class A, as specified in Table 1.

Table 1 - Tolerances for assessing the planimetric and altimetric accuracy of PEC-PCD for Class A.

PEC-PCD Classe A	Planimetry		Altimetry	
	PEC (m)	EP (m)	PEC (m)	EP (m)
1:1,000	0.28	0.17	0.27	0.17
1:2,000	0.56	0.34	0.27	0.17
1:5,000	1.40	0.85	0.54	0.34

Source: Adapted from (DSG, 2016).

For a cartographic product to be classified as Class A at a scale of 1:1,000, for example, it must have an RMSE lower than the Standard Error (EP) of 0.17 m, and 90% of the errors (discrepancies between the reference coordinates and those represented in the cartographic products) must be less than 0.28 m for planimetry and 0.27 m for altimetry.

2.3. RESULTS

2.3.1 Number of GCPs

The data dispersion can be seen in Figure 8. It were observed that the planimetry with two GCPs had the highest dispersion, while the altimetry values were even more contrasting. Due to this, the data for two GCPs were removed from Figure 8 to allow for better observation of the differences for other numbers of GCPs. With 3 and 4 GCPs, similar results were obtained in terms of dispersion for both planimetry and altimetry. The dispersion decreased with 5 and 6 GCPs, and outlines began to appear in the altimetry results.

Figure 8 - Boxplots of the evaluation of the number of control points in A planimetry and B altimetry.

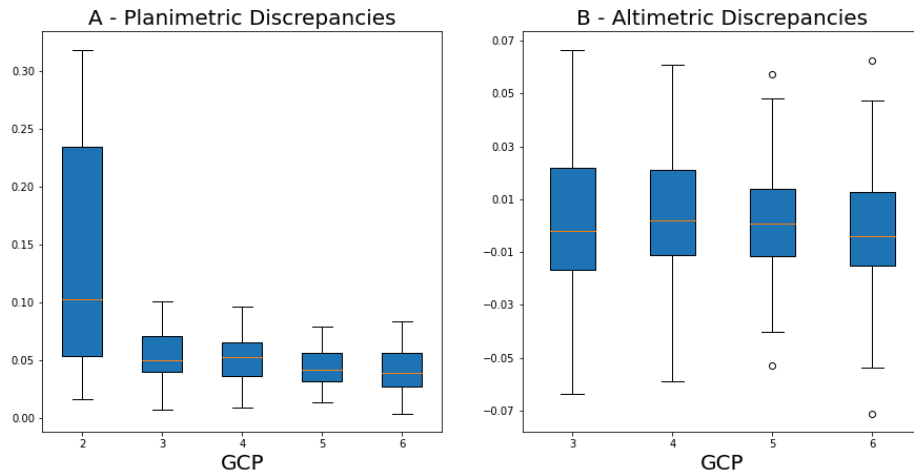
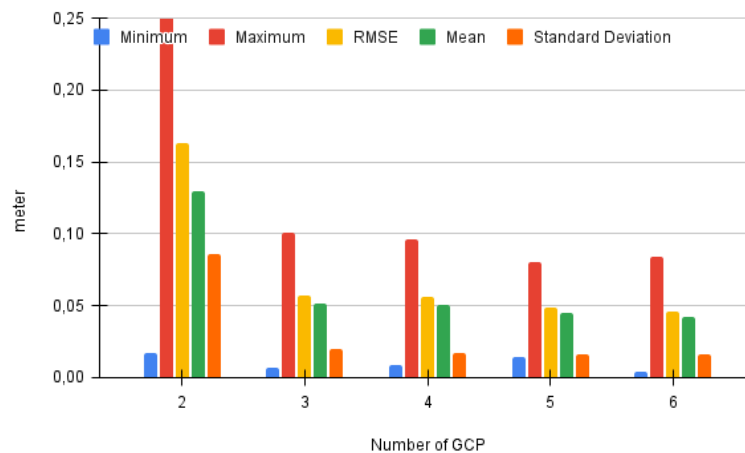


Figure 9 presents the statistical data for planimetry concerning the number of GCPs. The minimum discrepancy values should tend to zero since planimetry measures distances. For the processing with 2 and 5 GCPs, the minimum values deviated from zero at 0.017 m and 0.014 m, respectively, while the others had similar values ranging from 0.004 m to 0.009 m. Regarding the maximum discrepancy, 2 GCPs had a higher value of 0.319 m, while the others ranged from 0.101 m to 0.08 m.

As for RMSE, it decreased as the number of GCPs increased. A similar pattern was observed for the standard deviation (SD). Both metrics measure the dispersion of the data, with RMSE representing accuracy and SD representing precision. Lower values are better for both metrics. With 6 GCPs, RMSE and SD were 0.046 m and 0.016 m, respectively, while with 5 GCPs, they were 0.049 m and 0.016 m, respectively.

Figure 9 - Data analysis of planimetric discrepancies in meters for different numbers of control points.



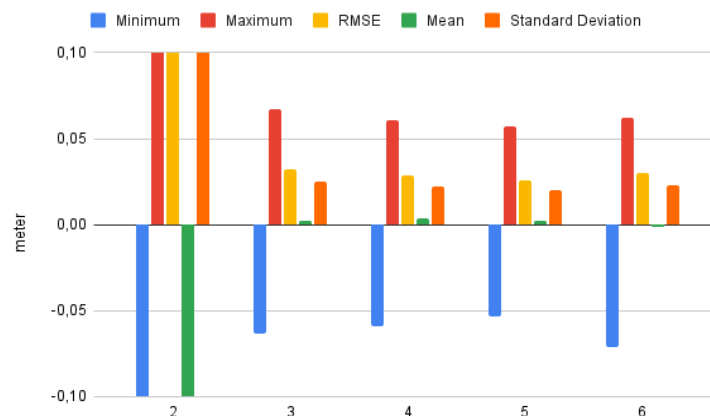
The orthophoto classification according to the PEC-PCD for planimetry (Table 2) shows that for all numbers of GCPs, the classification was at a 1:1,000 scale, Class A. It is worth noting that with 2 GCPs, only 91% of the control points had discrepancies smaller than 0.28 m.

Table 2 - Verification of data according to the tolerances of PEC-PCD for planimetry at a 1:1,000 scale, Class A.

PEC-PCD	(m)	2	3	4	5	6
EP	0.17	0.163	0.057	0.056	0.049	0.046
PEC	0.28	91%	100%	100%	100%	100%

Figure 10 presents the statistical data for altimetry based on the number of GCPs. The minimum altimetric discrepancy values were similar for tests with 3 to 6 GCPs, ranging from -0.071 m to -0.053 m. When using only 2 GCPs, the minimum value was -3.3 m. The same trend is observed for the maximum values, where the 2 GCPs case had a difference of 2.61 m compared to the others, which ranged from 0.057 m to 0.067 m. Regarding the RMSE (Root Mean Square Error) and DP (Standard Deviation), there is a pattern of decreasing values as the number of GCPs increases. However, the results stabilize after 5 GCPs. Figure 9 shows that with 5 GCPs, the RMSE and DP values were 0.026 m and 0.02 m, respectively, while with 6 GCPs, the values were 0.030 m and 0.023 m, respectively.

Figure 10 - Data analysis of altimetric discrepancies in meters for different quantities of ground control points (GCPs). The y-axis of the graph is limited to +/- 0.1 m to avoid distortions caused by values obtained with only 2 GCPs (that obtained values in meters of: -3.3 minimum, 2.6 maximum, RMSE 1.77, mean -0.48, and standard deviation of 1.46), which would make it impossible to compare with the other GCP quantities used.



The classification of altimetry according to the PEC-PCD (Brazilian Cartographic Accuracy Standard) is presented in Table 3. It is noteworthy that only with 2 GCPs, it was not possible to achieve classification in the 1:1,000 scale class A. This is observed both in terms of the RMSE (Root Mean Square Error) validation against the standard error, which is greater than 1.77 m compared to 0.17 m, and in terms of 90% of the evaluations being below the PEC (Positional Error of the Cartographic Database) of 0.27 m, where only 9% met this requirement.

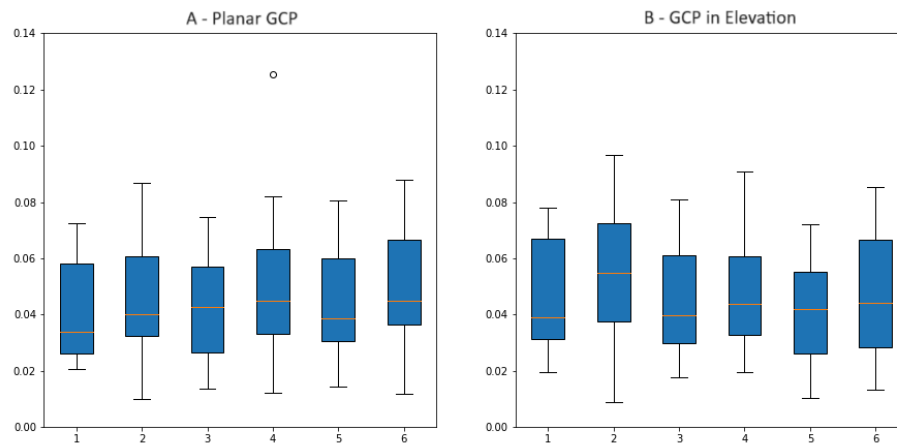
Table 3 - Verification of data regarding tolerances of altimetry according to the PEC-PCD in the 1:1,000 scale class A.

PEC-PCD	(m)	2	3	4	5	6
EP	0.17	1.77	0.032	0.029	0.026	0.030
PEC	0.27	9%	100%	100%	100%	100%

2.3.2. GCPs in Elevation

The results are shown in Figure 11, which displays the boxplot graph for each processing, evaluating the discrepancy in planimetry. It can be observed that the distribution of discrepancies ranged from 0.01 m to 0.10 m, with an outline of 0.13 m for the data with planar GCPs.

Figure 11 - Boxplot of the leave-one-out experiment for each of the 6 processings with A planar GCPs and B GCPs in elevation, for planimetry.



As for the planimetric PEC-PCD at a 1:1,000 scale, in all processings, both criteria were met, with values lower than the threshold values (Table 4).

Table 4 - Verification of planar and elevation processings regarding the tolerances of the planimetric PEC-PCD at a 1:1,000 scale, class A.

	PEC-PCD	(m)	1	2	3	4	5	6
Plane	EP	0.17	0.044	0.048	0.045	0.056	0.047	0.053
	PEC	0.28	100%	100%	100%	100%	100%	100%
Elevation	EP	0.17	0.048	0.058	0.048	0.049	0.045	0.050
	PEC	0.28	100%	100%	100%	100%	100%	100%

For the evaluation in altimetry (Figure 12), there are some differences. The distribution of altitude discrepancies was smaller and had fewer outliers in the dataset with GCPs in elevation, indicating a subtle improvement when using these points. Regarding the PEC-PCD for altimetry at the 1:1,000 scale, both criteria were met in all processing, with values lower than the specified thresholds (Table 5).

Figure 12 - Boxplot of the leave-one-out experiment for each of the 6 processings with A GCPs in the planar and B elevation datasets for altimetry.

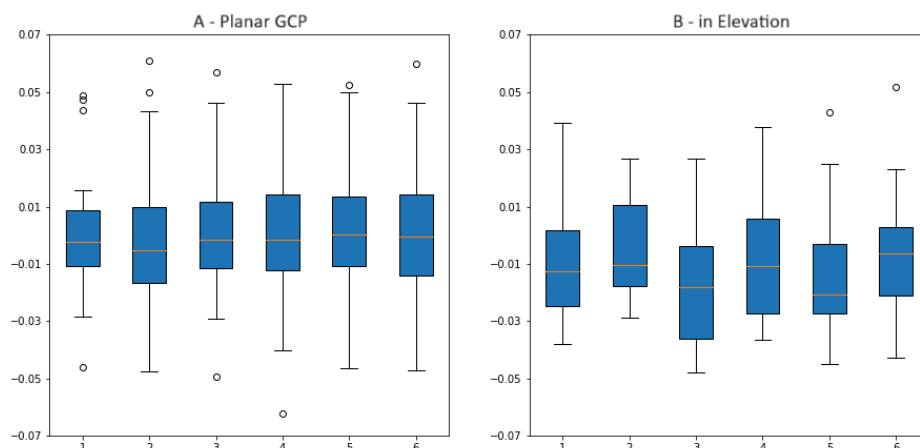


Table 5 - Verification of the planar and elevation processings regarding the tolerances of the PEC-PCD for altimetry at the 1:1,000 scale, class A.

	PEC-PCD	(m)	1	2	3	4	5	6
Plane	EP	0.17	0.027	0.028	0.029	0.030	0.027	0.029
	PEC	0.27	100%	100%	100%	100%	100%	100%
Elevation	EP	0.17	0.023	0.021	0.028	0.023	0.028	0.024
	PEC	0.27	100%	100%	100%	100%	100%	100%

In general, there was no difference in using GCPs in elevation for planimetry, as the dispersion of discrepancies was similar. For altimetry data, there were slight changes, with a small reduction in data dispersion and RMSE values. It is important to highlight that the best result in terms of arrangement was obtained with 5 GCPs, which represents an improvement compared to the results of this study. All processings were classified as class A in the PEC-PCD for both planimetry and altimetry at the 1:1,000 scale.

2.4 DISCUSSION

The objective of this work is to investigate whether the quantity and spatial arrangement of control points located on flat and elevated terrain interfere with the accuracy of cartographic products generated from images collected by sensors embedded in UAVs. The surveyed area of 5.5 ha overflowed area had 31 targets surveyed with GNSS RTK, 21 CP and 12 GCP.

Regarding the methodology employed in this study, its suitability is evident, as it aligns with approaches used in various scientific works. For instance, Ferrer-González et al. (2020) employed a similar methodology when assessing corridor mapping, testing the distribution of 3 to 18 Ground Control Points (GCPs) in four distinct ways. This study involved the collection of 47 points (comprising GCPs and Control Points - CP) using GNSS equipment in a 40-hectare area.

Hastaoglu et al. (2022) investigated the difference in height and geometry of GCPs in a larger area of 160 hectares. In this case, 122 CPs and 46 GCPs were surveyed, utilizing GNSS RTK technology. Liu et al. (2022) addressed the impact of GCP configuration in a 50-hectare area, collecting 16 points (including GCPs and CP) for planimetric evaluation and 120 CPs for vertical assessment, all obtained through GNSS RTK.

It is evident that these studies employed both GCPs and CPs acquired through GNSS equipment. The variation in the quantity of GCPs and CPs used is noteworthy and directly linked to the size of the study areas, demonstrating an adaptation of the methodology to the specific scale of each research.

However, regarding the quantity of GCPs, it is observed that 2 GCPs present different results compared to the others, with a significant discrepancy in altimetry, as also observed by Siqueira et al. (2020), who found RMSE values of 0.3 m for planimetry and 1.7 m for altimetry with 2 GCPs. With 3 and 4 GCPs, the results were good for both altimetry and planimetry. For this quantity of GCPs, Siqueira et al. (2020) reported RMSE values of 0.13 m and 0.19 m for planimetry and altimetry, respectively. Yu et al. (2020) obtained RMSE values of 0.05 m and 0.11 m for a small area (7 ha), but these results with RMSE up to 0.20 m for altimetry can only be considered for small areas (in this experiment, 5.5 ha). Yu et al. (2020) reported RMSE values of 0.12 m and 1.148 m for planimetry and altimetry, respectively, for a medium-sized area (39 ha), and 0.91 m and 4.1 m for planimetry and altimetry, respectively, for a large area (342 ha)

with 3 GCPs. With 5 and 6 GCPs, similar results were obtained for planimetry, but for altimetry, 6 GCPs showed slightly worse results, indicating that the accuracy had already been achieved. As presented by Siqueira et al. (2020), above 6 GCPs, the variation tends to be small, and using 6 and 8 GCPs did not increase accuracy.

Regarding the planimetric PEC-PCD (Cartographic Accuracy Standard), a class A classification was obtained for the 1:1,000 scale for the GCP distributions, which is consistent with the findings of Barbosa et al. (2021), who also classified UAV-derived orthophotos as class A in the 1:1,000 scale. However, Siqueira et al. (2020) classified the orthophoto with 2 GCPs as class A in the 1:5,000 scale. One possible explanation for this difference is the area size; they used a larger area of 51 ha compared to this experiment. In larger areas, the use of a greater number of control points becomes necessary to georeference the orthophoto. It should be noted that only 91% of the control points had values smaller than the threshold, so 2 GCPs approached the limit values.

For the altimetric PEC-PCD, except for 2 GCPs, all data points meet the class A criteria for the 1:1,000 scale. The results from Siqueira et al. (2020) were similar, with significant discrepancies observed with 2 GCPs, not meeting the criteria for the 1:1,000 scale, except with 4 GCPs, where the scale was 1:5,000 class A. However, the criterion for classification in the 1:1,000 scale, with 90% of the data having discrepancies below 0.27 m, was met.

In general, there was no difference in using GCPs in elevation for planimetry, as the dispersion of discrepancies was similar. For altimetry data, there were slight changes, with a small reduction in data dispersion and RMSE values. It is important to highlight that the best result in terms of arrangement was obtained with 5 GCPs, which represents an improvement compared to the results of this study. All processings were classified as class A in the PEC-PCD for both planimetry and altimetry at the 1:1,000 scale.

Ferrer-Gonzalez et al. (2020) noted that the influence of GCPs distribution is more evident for altimetric accuracy. Liu et al. (2022) observed that, for the same area, the quantity of GCPs needed to stabilize accuracy differs between planimetry and altimetry, requiring more GCPs for altimetry. It is worth noting that the authors also found that altimetry is more sensitive, corroborating with the findings of the present study.

2.5 CONCLUSIONS

The necessary quantity of GCPs and the use of GCPs in elevation to improve accuracy in UAV surveys were evaluated. It was concluded that using only 2 GCPs results in low accuracy for the orthophoto and generated MDS in image processing. As the number of GCPs increases, accuracy improves, but only up to a certain limit. Beyond a certain quantity of GCPs, accuracy stabilizes (5 GCPs). In the case study, it was found that elevation is the most sensitive variable, and the quantity and arrangement of GCPs significantly influence elevation accuracy. The use of GCPs for elevation improves elevation precision.

2.6 REFERENCES

- Barbosa, L. S., de Souza, L. M., da Cunha, M. J. P., & dos Santos, A. D. P. (2021). Análise Comparativa das Normas de Controle de Qualidade Posicional de Produtos Cartográficos do Brasil, do INCRA e da ASPRS. **Rev. Bras. Cartogr**, vol. 73, no. 3.
- Batistoti, J., Marcato Junior, J., Ítavo, L., Matsubara, E., Gomes, E., Oliveira, B., & Dias, A. (2019). Estimating pasture biomass and canopy height in Brazilian Savanna using UAV photogrammetry. **Remote Sensing**, vol. 11, no. 20, pp. 2447.
- Calou, V. B. C., Teixeira, A. D. S., Silva, J. A. D., Oliveira, M. R. R. D., & Nascimento, Í. V. D. (2021). Statistical process control and mapping accuracy standards applied to aerial surveys. **Revista Ciência Agronômica**, vol. 52.
- DSG - Diretoria de Serviço Geográfico. Normas da Especificação Técnica para a Aquisição de Dados Geoespaciais Vetoriais de Defesa da Força Terrestre - ET-ADGV. 2 ed. 85p. 2016. viewed 22 Out. 2021, <https://docs.ufpr.br/~deni_ern/CD2020/A1/ET_ADGV_2a_Edicao_2016_Textual_Anexo_A_Assinado.pdf>.
- Ferrer-González, E., Agüera-Vega, F., Carvajal-Ramírez, F., & Martínez-Carricondo, P. (2020). UAV photogrammetry accuracy assessment for corridor mapping based on the number and distribution of ground control points. **Remote Sensing**, vol. 12 no. 15, pp. 2447.
- Garcia, M. V. Y., & Oliveira, H. C. D. (2021). The influence of flight configuration, camera calibration, and ground control points for digital terrain model and orthomosaic generation using unmanned aerial vehicles imagery. **Boletim de Ciências Geodésicas**, vol. 27.
- Han, X., Thomasson, J. A., Wang, T., & Swaminathan, V. (2020). Autonomous mobile ground control point improves accuracy of agricultural remote sensing through collaboration with UAV. **Inventions**, vol. 5, no. 1, pp. 12.
- Hastaoglu, K. O., Kapicioglu, H. S., Gül, Y., & Poyraz, F. (2023). Investigation of the effect of height difference and geometry of GCP on position accuracy of point cloud in UAV photogrammetry. **Survey Review**, vol. 55, no. 391, pp. 325-337.
- Le VAN, C. A. N. H., CUONG, C. X., Nguyen, Q. U. O. C., ANH, T. T., & Xuan-Nam, B. U. I. (2020). Experimental Investigation on the Performance of DJI Phantom 4 RTK in the PPK Mode for 3D Mapping Open-Pit Mines. **Inżynieria Mineralna**, vol. 1, no. 2, pp. 65-74.
- Lewińska, P., Głowacki, O., Moskalik, M., & Smith, W. A. (2021). Evaluation of structure-from-motion for analysis of small-scale glacier dynamics. **Measurement**, vol. 168, pp. 108327.

- Liu, X., Lian, X., Yang, W., Wang, F., Han, Y., & Zhang, Y. (2022). Accuracy assessment of a UAV direct georeferencing method and impact of the configuration of ground control points. **Drones**, vol. 6, no. 2, pp. 30.
- Meinen, B. U., & Robinson, D. T. (2020). Where did the soil go? Quantifying one year of soil erosion on a steep tile-drained agricultural field. **Science of The Total Environment**, vol. 729, pp. 138320.
- Osco, L. P., de Arruda, M. D. S., Gonçalves, D. N., Dias, A., Batistoti, J., de Souza, M., & Gonçalves, W. N. (2021). A CNN approach to simultaneously count plants and detect plantation-rows from UAV imagery. **ISPRS Journal of Photogrammetry and Remote Sensing**, vol. 174, pp. 1-17.
- Padró, J. C., Muñoz, F. J., Planas, J., & Pons, X. (2019). Comparison of four UAV georeferencing methods for environmental monitoring purposes focusing on the combined use with airborne and satellite remote sensing platforms. **International journal of applied earth observation and geoinformation**, vol. 75, pp. 130-140.
- PIX4D. Pix4Dmapper. Versão 4.1.25. Pix4D, 2018.
- QGIS Development Team. QGIS. Versão 2.18.24 [Las Palmas]. QGIS Development Team, 2019.
- Revuelto, J., López-Moreno, J. I., & Alonso-González, E. (2021). Light and shadow in mapping alpine snowpack with unmanned aerial vehicles in the absence of ground control points. **Water Resources Research**, vol. 57, no. 6, pp. e2020WR028980.
- Siqueira, H. L., Junior, J. M., Matsubara, E. T., Colares, R. A., & Santos, F. M. (2020). Acurácia de Produtos Fotogramétricos Gerados com Aeronave Remotamente Pilotada em Relevo Acidentado. **Rev. Bras. Cartogr**, vol. 72, no. 3, pp. 48413.
- TRIMBLE (2013) - Trimble R8, R6 and R4 User Guide Version 4.80 Revision A. viewed 22 Out. 2022, <https://www.trimble.com/support_trl.aspx?Nav=Collection-65944&pt=Trimble%20R4>.
- Yu, J. J., Kim, D. W., Lee, E. J., & Son, S. W. (2020). Determining the optimal number of ground control points for varying study sites through accuracy evaluation of unmanned aerial system-based 3D point clouds and digital surface models. **Drones**, vol. 4, no. 3, pp. 49.
- Zanetti, J., Junior, J. G. & Dos Santos, A. De P. (2017). Influência do número e distribuição de pontos de controle em ortofotos geradas a partir de um levantamento por VANT. **Revista Brasileira de Cartografia**, vol. 69, no. 2.

TERCEIRO CAPÍTULO: FILTRAGEM DA NUVEM DE PONTOS GERADA POR UAV RGB COMBINANDO DEEP LEARNING E FOTOGRAMETRIA

Resumo: Os Modelos Digitais do Terreno (MDT) são fundamentais para estudos hidrológicos, especialmente em simulações de escoamento, enchentes e inundações, exigindo alta resolução espacial e acurácia altimétrica. Veículos Aéreos Não Tripulados (VANTs) equipados com câmeras RGB representam uma solução acessível para aquisição de dados, uma vez que esses sensores são mais econômicos que os multiespectrais ou LiDAR. As imagens RGB, quando processadas via *Structure from Motion e Multi-View Stereo* (SfM-MVS), permitem gerar uma nuvem densa de pontos representando tanto o solo quanto os objetos acima do terreno, resultando em um Modelo Digital de Superfície (MDS). A filtragem dessa nuvem, necessária para isolar os pontos do solo e obter o MDT, é uma etapa custosa e pouco automatizada. Este trabalho propõe uma abordagem inovadora ao integrar segmentação semântica por *deep learning* (DL) com o processo SfM-MVS para gerar MDTs diretamente, sem a necessidade de filtragem posterior. Foram avaliadas cinco arquiteturas de DL (FCN, GCNet, PointRend, PSANet e SegFormer) na segmentação de árvores em imagens RGB obtidas por VANT. As máscaras geradas foram utilizadas para excluir automaticamente as regiões com vegetação antes da reconstrução da nuvem de pontos, de forma que o SfM-MVS operasse apenas sobre áreas classificadas como solo. Para comparação, também foi gerado um MDT convencional, a partir da nuvem densa de pontos processado com o algoritmo de filtragem *Cloth Simulation Filtering* (CSF). Os resultados demonstraram que a arquitetura SegFormer apresentou o melhor desempenho quali-quantitativo na segmentação de árvores, e que os MDTs gerados pela abordagem proposta apresentaram acurácia compatível aos obtidos com o método tradicional. Conclui-se que a integração entre DL e fotogrametria representa uma estratégia promissora e escalável para extração automatizada de MDTs em ambientes urbanos.

Palavras-chave: Segmentação semântica; imagens RGB; Fotogrametria Digital; mapeamento do terreno

Abstract: Digital Terrain Models (DTMs) are essential for hydrological studies, particularly in simulations of runoff, floods, and inundations, as they require high spatial resolution and vertical accuracy. Unmanned Aerial Vehicles (UAVs) equipped with RGB cameras represent a cost-effective solution for data acquisition, given that such sensors are more affordable than multispectral or LiDAR systems. RGB images, when processed through Structure from Motion and Multi-View Stereo (SfM-MVS), allow the generation of dense point clouds representing both ground and above-ground objects, resulting in a Digital Surface Model (DSM). The filtering of this cloud, necessary to isolate ground points and derive the DTM, is a labor-intensive and poorly automated step. This study proposes an innovative approach by integrating semantic segmentation via deep learning (DL) with the SfM-MVS workflow to generate DTMs directly, without the need for post-processing filters. Five DL architectures (FCN, GCNet, PointNet, PSANet, and SegFormer) were evaluated for the segmentation of trees in RGB images captured by UAVs. The resulting masks were used to automatically exclude vegetated areas prior to point cloud reconstruction, ensuring that the SfM-MVS pipeline processed only regions classified as ground. For comparison, a conventional DTM was also generated from the full dense cloud, followed by filtering using the Cloth Simulation Filtering (CSF) algorithm. The results showed that the SegFormer architecture achieved the best qualitative and quantitative performance for tree segmentation. Moreover, the DTMs generated using the proposed approach demonstrated accuracy comparable to those produced by traditional filtering. It is concluded that the integration of deep learning with photogrammetry represents a promising and scalable strategy for the automated extraction of DTMs in urban environments.

Keywords: Semantic segmentation; RGB imagery; Digital photogrammetry; Terrain mapping

3.1. INTRODUÇÃO

O Modelo Digital de Superfície (MDS) e o Modelo Digital de Terreno (MDT) são usados em diversas aplicações, como na estimativa da biomassa de pastagem ou altura do dossel de plantas (Batistoti et al., 2019), na detecção de árvore individual (Belcore et al., 2020), entre outras. O MDT é uma importante fonte de dados para estudos hidrológicos como em: simulações de enchentes (Annis et al., 2020; Jeon & Choi, 2020), alagamento de áreas costeiras (Leal-Alves et al., 2020, Madvedev et al., 2020, Anzidei et al., 2018), erosão do solo (Meinen & Robinson, 2020; Naranjo et al., 2021), entre outros. Anders et al. (2019), Naranjo et al., (2021) destacam a importância de ter modelos de alta resolução espacial e acurados para essas análises. Segundo Na et al. (2020), vários métodos de aquisição de dados remotamente são aplicados para modelar o terreno em escalas locais - grande escala cartográfica -, como o *Radio Detection and Ranging* (RADAR), *Light Detection and Ranging* (LiDAR) e a estereofotogrametria. O LiDAR é amplamente utilizado para a geração simultânea de Modelos Digitais de Superfície (MDS) e de Terreno (MDT), uma vez que sua capacidade de registrar múltiplos retornos permite identificar pontos no solo mesmo em áreas densamente vegetadas (Meng, Currit & Zhao, 2010). No entanto, esse equipamento possui custos maiores, se comparados aos sensores RGB.

No que se refere à plataforma de coleta, o veículo aéreo não tripulado (*Unmanned Aerial Vehicle* - UAV) pode ser embarcar diversos sensores que permitem coletar dados com alta resolução temporal e espacial, além de ser uma tecnologia de baixo custo se comparado a aeronaves pilotadas (Batistoti et al., 2019, Na et al., 2020, Belcore et al., 2020). Em geral, os UAV possuem câmeras RGB, que capturam imagens sequenciais utilizadas em processos de fotogrametria para a geração de produtos cartográficos.

A partir dessas imagens, é possível gerar uma nuvem de pontos, fundamental para a construção de ortofotos—representações corrigidas das distorções de escala, que permitem medições precisas de áreas e distâncias. No entanto, a qualidade da nuvem de pontos influencia diretamente a precisão das ortofotos, especialmente em terrenos com superfícies de referência muito rugosas, onde as distorções tendem a ser mais acentuadas. Para minimizar esse efeito, podem ser aplicadas técnicas de filtragem da

nuvem de pontos, que suavizam a superfície de referência e resultam em ortofotos menos distorcidas.

Segundo Gevaert et al., (2018), existem diversas técnicas de filtragens do MDS para a geração do MDT, podendo ser divididas em cinco grupos: filtros morfológicos, densificação progressiva, baseado na superfície, baseada em segmentos e métodos de *Deep Learning* (DL).

Na et al. (2020) propuseram a aplicação de DL para segmentar áreas de vegetação, usando como dado de entrada imagens RGB e dados de elevação (nuvem de pontos). Os autores mostram que os segmentos pequenos de vegetação podem ser removidos e que para segmentos grandes são estimadas alturas da vegetação, resultando na obtenção do MDT. Ademais, os autores demonstram uma melhora na acurácia da ortofoto gerada com o MDT. Gevaert et al., (2018) propuseram utilizar o ortomosaico+MDS como dado de entrada em uma *Fully convolutional network* (FCN), visando classificar pontos de solo e pontos acima do solo, para, então, gerar o MDT. Jakovlevic et al. (2019) propuseram um método baseado em aprendizado profundo para classificação de nuvem de pontos, conseguindo assim extrair o MDT. Nos três casos é usado a nuvem de pontos para a classificação, sendo necessário primeiro gerar a nuvem de pontos, por exemplo, com a técnica *Structure from Motion* e *Multi-View Stereo* (SfM-MVS), visando extrair o MDT.

No processo fotogramétrico convencional, a obtenção do MDT envolve as etapas: (1) aquisição das imagens; (2) geração da nuvem de pontos densa via SfM-MVS; (3) aplicação de filtragem para remoção de elementos acima do solo; (4) geração do MDT.

No método proposto neste trabalho, a filtragem é antecipada por meio de segmentação semântica aplicada diretamente nas imagens RGB, gerando segmentações que direcionam o SfM-MVS a reconstruir apenas regiões sem vegetação arbórea. Dessa forma, o MDT pode ser obtido diretamente, eliminando a necessidade de filtragem posterior.

A segmentação semântica consiste em atribuir a cada pixel da imagem uma classe específica (ex.: vegetação, solo, edificação). Entre as arquiteturas de rede neural profunda mais relevantes para essa tarefa, destacam-se:

Fully Convolutional Network (FCN) – Rede que realiza convoluções sucessivas, produzindo um mapa de calor com a mesma resolução da entrada, onde pixels da classe de interesse se destacam (Long et al., 2015). GCNet – Introduz atenção global para melhorar a contextualização e classificação em áreas com padrões complexos (Cao et al., 2019). PointRend – Baseia-se em amostragem adaptativa de pontos incertos nas bordas das classes, refinando os limites (Kirillov et al., 2020). PSANet – Conecta cada pixel a todos os outros para capturar dependências globais e manter consistência espacial (Zhao et al., 2018). SegFormer – Utiliza arquitetura baseada em transformadores, combinando informações globais e locais sem convoluções tradicionais (Xie et al., 2021).

Visando dar celeridade nesse processo, este trabalho aplica arquitetura de DL ao longo do SfM-MVS de modo a gerar diretamente uma nuvem de pontos densa já filtrada, possibilitando a geração do MDT, ao invés do MDS. Como hipótese, tem-se que os modelos de DL podem ser aplicados em cada imagem individualmente antes da geração da nuvem de pontos, e assim remover os pontos de vegetação arbórea, gerando-se uma nuvem de pontos já filtrada para a posteriormente gerar diretamente o MDT. Por fim, como resultado do SfM-MVS tem-se o MDT, não sendo necessário aplicar posteriormente uma filtragem na Nuvem densa de pontos.

3.2. MATERIAL E MÉTODOS

3.2.1. Área de Estudo e aquisição das imagens RGB

O estudo foi realizado em áreas urbanas da Universidade Federal do Mato Grosso Do Sul (UFMS), localizada no município de Campo Grande/MS (Figura 1), com vegetação arbórea espaçada. Esse padrão de distribuição foi escolhido porque é o que permite o modelo proposto extrair o MDT com maior precisão.

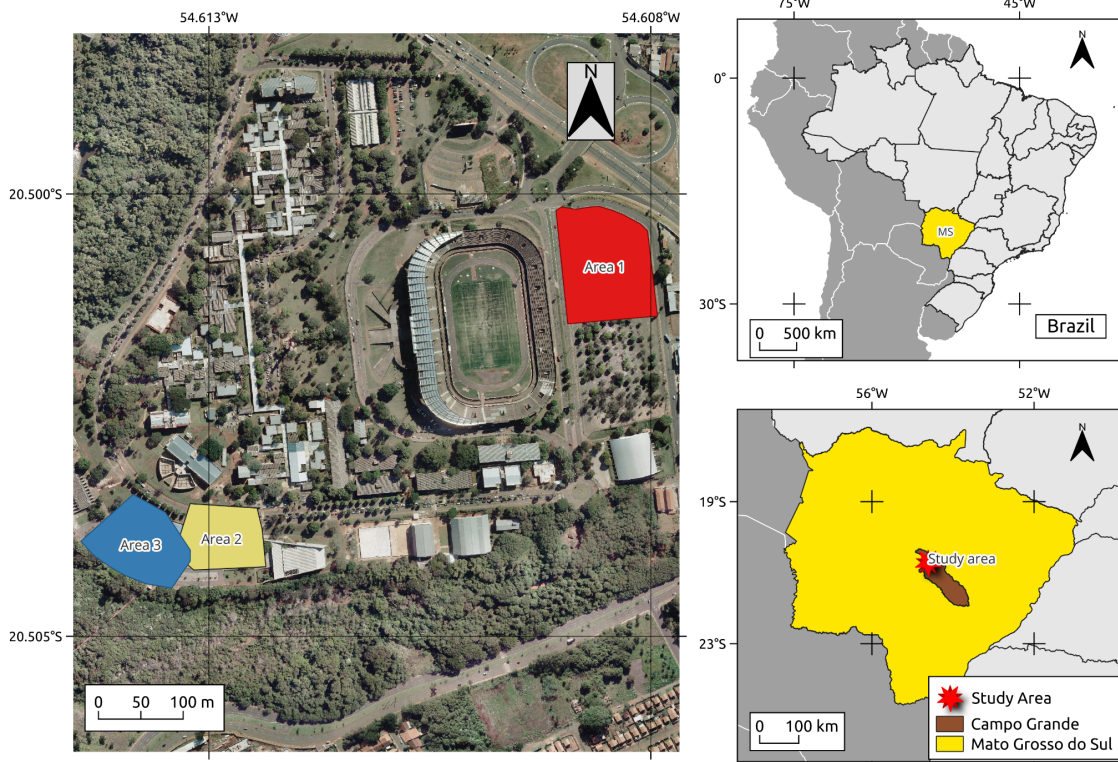


Figura 1. Áreas de estudo.

A coleta das imagens aéreas foi efetuada com a plataforma UAV *Phantom 4 Advanced* (Figura 2) multi-rotor, que possui câmera RGB de 20 Megapixel, sistema inercial (gimbal) e sistema de GNSS integrado. Os 3 voos, um para cada área, foram realizados com altura de voo de 50 m, câmera na posição vertical, sobreposição frontal de 80% e sobreposição lateral de 60%. As imagens possuem 1,36 cm/px de *Ground Sample Distance* (GSD).



Figura 2. UAV Phantom 4 Advanced.

Para o refinamento geoespacial de imagens coletadas pelo UAV foram usados alvos no tamanho de 50x50 cm semelhantes ao da Figura 3.c, os quais foram usados

como pontos de controle no solo (GCP) ou/e pontos de verificação (CP) durante a etapa de processamento.

Tabela 1. Descrição do número de fotos tiradas pelo UAV, ground control points e check points para cada área levantada.

	Imagens	GCP	CP
Área 1	132	8	75
Área 2	39	7	77
Área 3	35	8	80

Em campo, realizou-se a distribuição dos alvos de forma homogênea ao longo de toda a área de estudo. O levantamento geodésico destes pontos foi realizado com método de posicionamento *real time kinematic* (RTK) com o GNSS R8s (Figura 3.a e 3.b) da Trimble, cuja precisão 3D é de aproximadamente 30 mm, sendo no levantamento RTK horizontal na ordem de 8 mm + 0,1 ppm, na vertical de 15 mm + 1 ppm (Trimble, 2013). Realizou-se o levantamento das coordenadas em UTM (Universal Transversa de Mercator) no centro dos alvos (Figura 3.c).

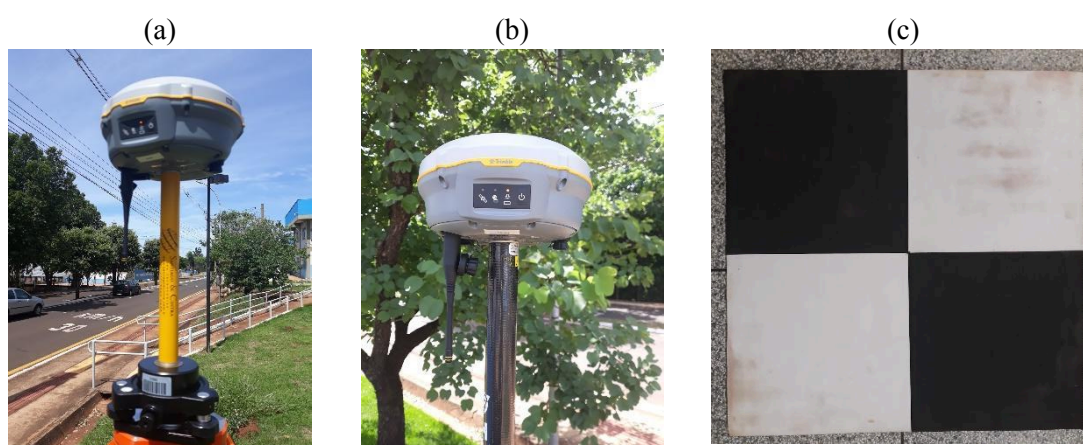


Figura 3. Equipamento Trimble R8s RTK (a) e (b) utilizado para levantamento dos pontos em campo GCP e CP (c).

3.2.2. Segmentação Semântica das árvores nas imagens RGB

Na Figura 4, esquematiza-se o processo para a aplicação do DL. Os dados de entrada foram as imagens RGB obtidas pelo UAV (Figura 5.a), e as máscaras de árvores

manualmente vetorizadas (Figura 5.b), utilizando a ferramenta ImageJ (Schneider et al. 2012).

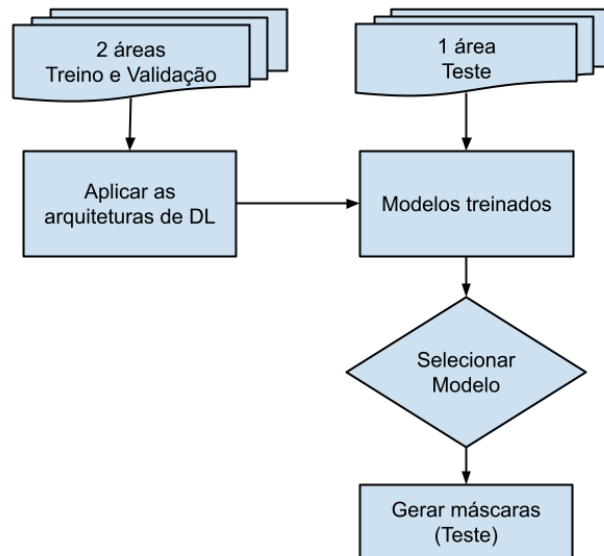


Figura 4. Esquema para avaliar a arquitetura de DL, geração das máscaras de segmentação de áreas arbóreas.



Figura 5. Imagens obtidas pelo UAV (a), Segmentação manualmente, sendo branco as regiões de copas de árvores, e preto o fundo da imagem (b).

Separou-se o conjunto de imagens em dois grupos: treinamento e validação, e teste das redes de DL. Realizou-se três experimentos, onde cada experimento conteve 2 áreas experimental para treino e validação, e uma para teste (Tabela 2).

Tabela 2. Descrição das áreas usadas para treino e validação, e teste em cada experimento.

Experimentos	Treino + Validação	Teste
1	Área 2 e 3	Área 1
2	Área 1 e 3	Área 2
3	Área 1 e 2	Área 3

O desempenho das redes na segmentação das áreas arbóreas foi analisada pela métrica de Interseção sobre a União (IoU) Eq. (1), tempo de inferência em cada foto e número de parâmetros. A métrica foi calculada considerando a classe de interesse como sendo “árvore” e o restante da imagem como “fundo”. Com isso, calcula-se: os valores de Verdadeiros Positivos (VP) - classificação correta da classe árvore; Falso Negativos (FN) - erro em que o modelo previu a classe fundo quando o valor real era classe árvore; Falso Positivo (FP) - erro em que o modelo previu a classe árvore quando o valor real era classe fundo; e Verdadeiro Negativo (VN) - quando o modelo previu classe fundo e realmente é classe fundo.

$$IoU = \frac{VP}{FP+FN+VP} \quad (1)$$

As máscaras destacando as áreas de interesse (árvores) foram geradas com a arquitetura que apresentou melhor resultado na tarefa de segmentação semântica. Para executar as redes de DL, foi utilizado o computador com sistema operacional Ubuntu 18.04 version, uma *Central Processing Unit* (CPU) Intel (R) Xeon (R) E3-1270@3.8Ghz, *Random Access Memory* (RAM) de 64 GB, uma *Graphics Processing Unit* (GPU) Titan V de memória gráfica de 12 GB com 5120 núcleos de *Compute Unified Device Architecture* (CUDA). Os algoritmos foram codificados utilizando MMSegmentation (MMSeg, 2020), um projeto de código aberto com ferramentas de segmentação baseado em PyTorch.

3.2.3. Geração do MDT

As imagens foram processadas no software Agisoft Metashape (versão 1.5.1), utilizando a técnica de Structure from Motion (SfM) na primeira etapa e Multi-View Stereo (MVS) na segunda.

Na etapa 1, os algoritmos de SfM estimam um modelo tridimensional da cena, gerando uma nuvem de pontos esparsa e calculando os parâmetros intrínsecos da câmera (como comprimento focal e centro de projeção) e os parâmetros extrínsecos (pose 3D, incluindo rotação e translação). Esse método utiliza um conjunto de imagens com sobreposição lateral e longitudinal para obter correspondências entre os pontos (Lewinska et al., 2020). Como resultado, essa etapa gera os parâmetros de calibração da

câmera e uma nuvem de pontos esparsa, onde cada ponto 3D corresponde a uma coordenada 2D nas imagens de entrada, permitindo o alinhamento das imagens.

Na etapa seguinte, o algoritmo MVS é aplicado para refinar a reconstrução tridimensional. Diferente do SfM, que opera com características extraídas das imagens, o MVS trabalha diretamente com os valores de pixel, permitindo a construção de uma malha poligonal 3D detalhada e a preservação de detalhes finos da geometria da cena (Ryan et al., 2015).

A Figura 6 apresenta um esquema comparativo entre a geração do MDT pelo processo fotogramétrico convencional (SfM-MVS) e o método proposto neste trabalho.

No método proposto, as imagens segmentadas por uma rede de Deep Learning (DL) são utilizadas como máscaras para o processamento no Metashape, sendo aplicadas após a etapa 1. Segundo Agisoft (2020), essas máscaras podem reduzir a complexidade da nuvem densa, eliminando áreas das fotos que não são de interesse. Com isso, obtém-se uma nuvem densa de pontos já filtrada, ou seja, sem pontos na copa das árvores.

No processo tradicional, essa filtragem ocorre fora do Metashape, exigindo a exportação da nuvem de pontos densa para outro software. Neste estudo, utilizamos o CloudCompare, onde a filtragem foi realizada pelo método Cloth Simulation Filtering (CSF). Esse método simula um tecido estendido sobre os dados invertidos, classificando e separando pontos do solo e da vegetação (Zhang et al., 2016).

Após a filtragem, a nuvem de pontos correspondente ao terreno é utilizada para gerar o MDT. Essa etapa consiste na conversão da nuvem de pontos densa em uma matriz regular de elevação, utilizando técnicas de interpolação para definir os valores das células da matriz com base nos pontos irregularmente distribuídos. O MDT resultante serve como dado de entrada para a geração do ortomosaico, garantindo a continuidade do fluxo de processamento.

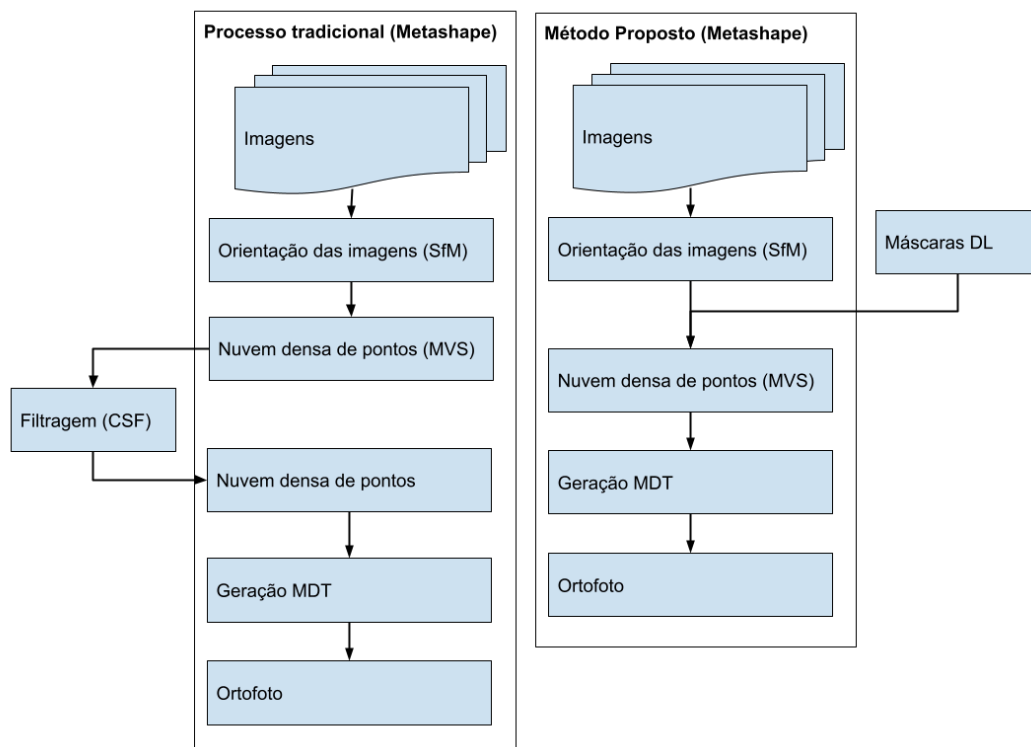


Figura 6. Esquema de processamento para a geração do MDT e Ortofoto, a esquerda método tradicional com a filtragem da nuvem densa, e a direita no esquema método proposto com a entrada das máscaras geradas pelo DL.

3.2.4. Avaliação dos MDT

A avaliação da acurácia vertical dos Modelos Digitais de Terreno (MDT) foi realizada para cada área de estudo, comparando o método tradicional e o modelo proposto. Foram utilizados pontos de checagem (CP) coletados com posicionamento GNSS RTK, cujos valores altimétricos foram comparados aos valores extraídos dos MDTs gerados.

A Figura 7 apresenta a distribuição de 75 CP na Área 1, a Figura 8 mostra 77 CP na Área 2, e a Figura 9 exhibe 80 CP na Área 3. Esses pontos foram distribuídos para avaliar a precisão dos MDTs em diferentes condições de terreno.



Figura 7. Distribuição dos pontos de checagem (CP) e pontos de controle em solo (GCP) na área 1.

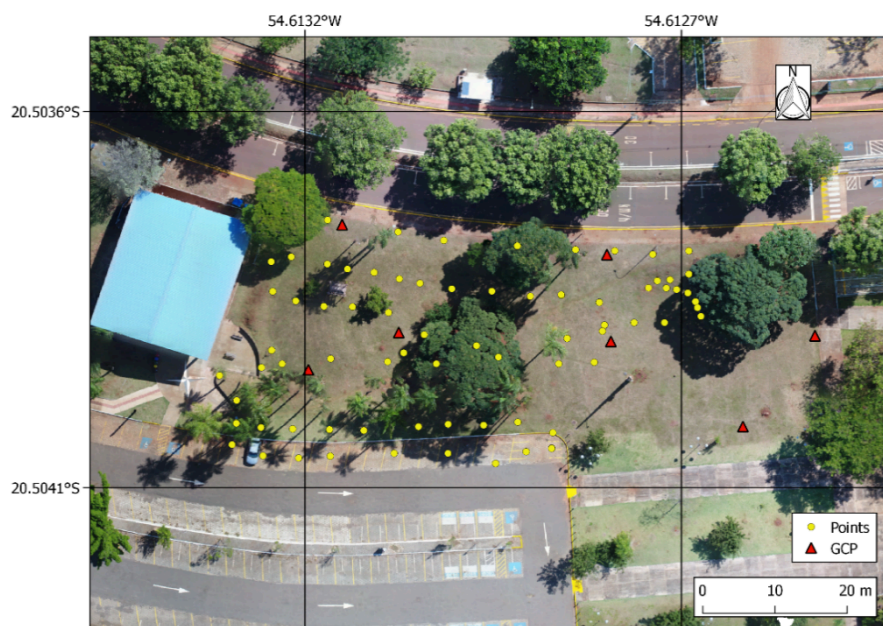


Figura 8. Distribuição dos pontos de checagem (CP) e pontos de controle em solo (GCP) na área 2.

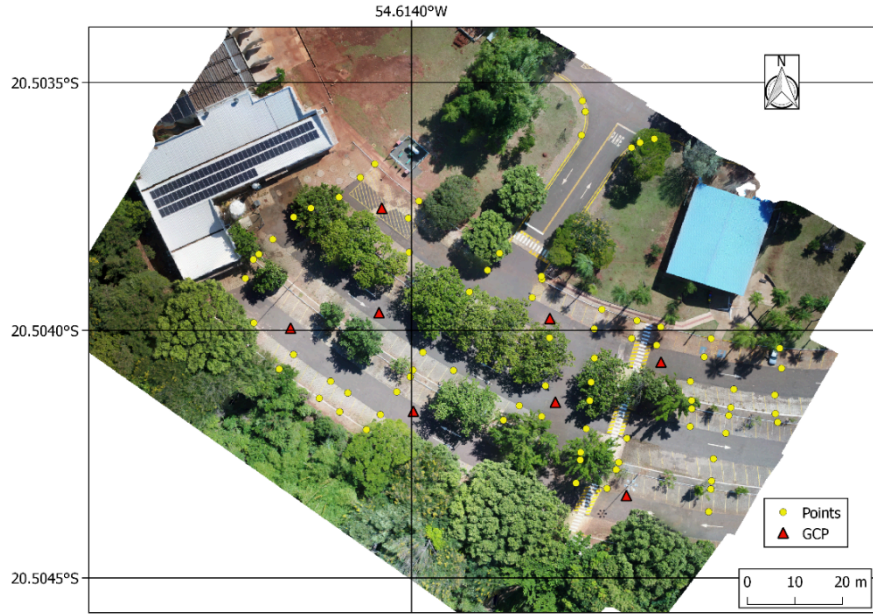


Figura 9. Distribuição dos pontos de checagem (CP) e pontos de controle em solo (GCP) na área 3.

As discrepâncias altimétricas foram calculadas de acordo com a Eq. (2). Com as discrepâncias calculou-se o *Root Mean Square Error* (RMSE) Eq. (3), média e desvio-padrão.

$$\Delta Z = Z_i - Z_{RTK_i} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum \Delta Z^2}{n}} \quad (3)$$

Onde ΔZ são as discrepâncias altimétricas, Z_i são as altitudes obtidas pelo MDT, Z_{RTK_i} são as altitudes obtidas pelo levantamento com o GNSS RTK, n é o número de CP avaliados.

3.3. RESULTADOS

Os resultados foram organizados por seções, onde DL é a seção com o resultado da aplicação das 5 arquiteturas, na seção Avaliação de MDT apresenta a comparação das nuvens de pontos e MDT gerados com a utilização da segmentação e com a aplicação de filtro, e por último, na seção Ortofoto é feita uma análise visual das

ortofotos geradas com MDT pelo método proposto com MDS sem nenhum ajuste, ou seja, sem o processo de filtragem, uma abordagem comumente utilizada na prática.

3.3.1. Avaliação das arquiteturas DL

Para comparar o desempenho das diferentes arquiteturas de DL utilizadas neste estudo, foi realizada a análise do Intersection over Union (IoU) e de seu desvio padrão para cada experimento. O IoU é um dos principais indicadores de acurácia em tarefas de segmentação, enquanto o desvio padrão reflete a variabilidade dos resultados obtidos.

Na Figura 10, cada barra representa o valor médio de IoU para uma arquitetura específica, com o respectivo desvio padrão indicado acima, permitindo a visualização das diferenças de desempenho entre os modelos e a consistência de suas previsões.

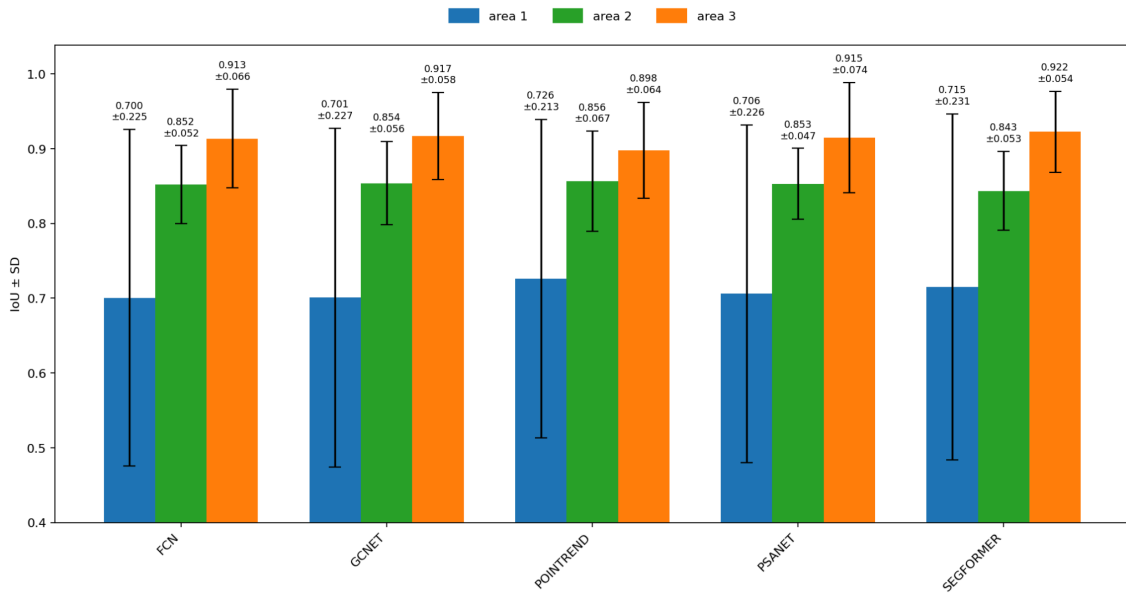


Figura 10. *IoU* (Intersection over Union) de cada arquitetura de DL, com cores indicando os diferentes experimentos/áreas. As linhas verticais mostram o desvio padrão ($\pm$ STD) e os rótulos exibem o valor médio seguido do desvio padrão.

A comparação entre as arquiteturas de *Deep Learning* revelou que os valores médios de IoU foram muito próximos entre si, indicando desempenho semelhante na segmentação, independentemente do modelo utilizado. Essa proximidade sugere que, nas condições deste estudo, a escolha da arquitetura teve influência limitada sobre a acurácia final, possivelmente devido à utilização do mesmo conjunto de dados base.

Entretanto, observou-se que a área 1 apresentou o menor valor de IoU e o maior desvio padrão (STD) em comparação com as demais áreas. Essa maior variabilidade e menor acurácia podem estar relacionadas à quantidade reduzida de imagens disponíveis para treinamento e avaliação nessa área, resultando em menor capacidade de generalização dos modelos. Assim, embora as arquiteturas tenham apresentado desempenhos semelhantes, a representatividade e o tamanho do conjunto de dados mostraram-se fatores determinantes para a consistência e a precisão dos resultados obtidos.

Nas Figuras 12, 13 e 14, foi montado uma imagem composta com a {foto + segmentação manual + predição dos modelos} comparado a anotação manual com as predições dos modelos FCN, GCNet, Pointrend, PSANet e SegFormer.

A análise visual não evidenciou diferenças significativas de desempenho entre as arquiteturas testadas.

As principais omissões ocorreram em árvores sem folhagem e troncos, como no caso de palmeiras, enquanto os falsos positivos estavam majoritariamente associados a sombras e canteiros gramados (Figuras 12 e 13). Esses padrões de erro foram consistentes entre as redes, indicando que as dificuldades de segmentação estão mais relacionadas às características visuais desses elementos do que à arquitetura utilizada.



Figura 12. Comparação entre a máscara de referência (anotação manual) e as predições dos modelos FCN, GCNet, Pointrend, PSANet e SegFormer para a área 1. As regiões em amarelo indicam áreas corretamente segmentadas (acertos), enquanto as áreas em vermelho representam regiões omitidas pelo modelo (*falsos negativos*) e as áreas em verde representam regiões segmentadas indevidamente (*falsos positivos*).



Figura 13. Comparação entre a máscara de referência (anotação manual) e as predições dos modelos FCN, GCNet, Pointrend, PSANet e SegFormer para a área 2. As regiões em amarelo indicam áreas corretamente segmentadas (acertos), enquanto as áreas em vermelho representam regiões omitidas pelo modelo (*falsos negativos*) e as áreas em verde representam regiões segmentadas indevidamente (*falsos positivos*).

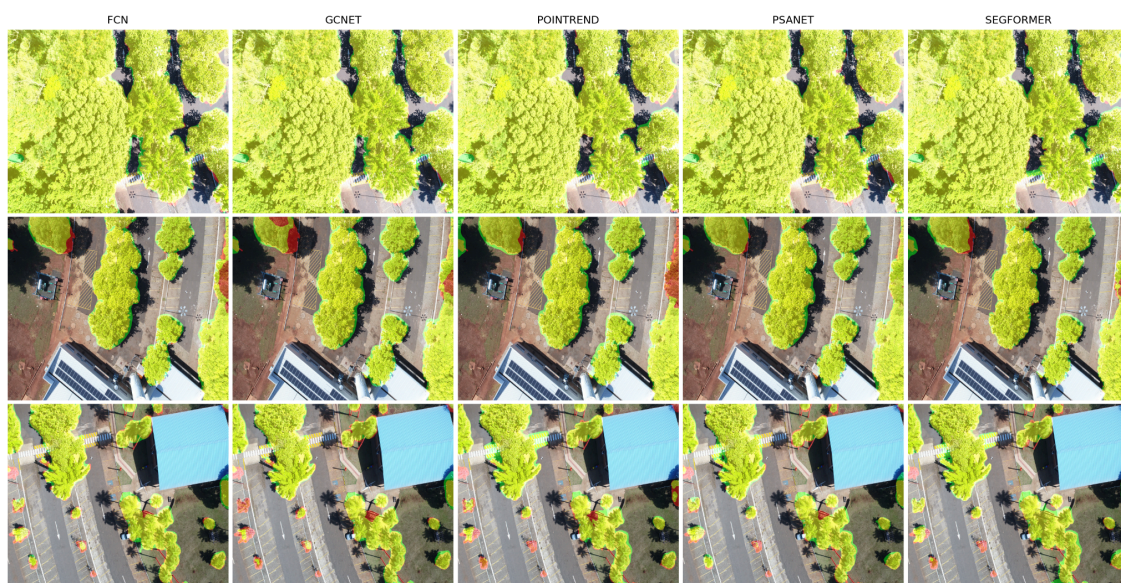


Figura 14. Comparação entre a máscara de referência (anotação manual) e as predições dos modelos FCN, GCNet, Pointrend, PSANet e SegFormer para a área 3. As regiões em amarelo indicam áreas corretamente segmentadas (acertos), enquanto as áreas em vermelho representam regiões omitidas pelo modelo (*falsos negativos*) e as áreas em verde representam regiões segmentadas indevidamente (*falsos positivos*).

Como não houve uma distinção clara quanto à superioridade de uma arquitetura específica, seja visualmente ou pelas métricas de IoU, ampliamos nossa análise para aspectos de eficiência, como o tempo médio de inferência por imagem e o número de parâmetros de cada modelo. A Tabela 3 apresenta os valores médios de IoU, desvio padrão (STD), tempo de inferência e número de parâmetros para as arquiteturas avaliadas.

Tabela 3. Valores de *Intersection over Union* (IoU), desvio padrão (STD), tempo de inferência e número de parâmetros obtidos para as arquiteturas de deep learning avaliadas.

		IoU	std IoU	mean time	std time	Parameters
area 1	FCN	0,7005	0,2250	0,1490	0,0147	49.484.644
	GCNET	0,7009	0,2265	0,1494	0,0165	49.617.125
	POINTREND	0,7259	0,2131	0,1258	0,0230	28.715.119
	PSANET	0,7060	0,2257	0,1639	0,0196	59.125.604
	SEGFORMER	0,7150	0,2315	0,1091	0,0171	13.678.786
area 2	FCN	0,8518	0,0523	0,1539	0,0299	49.484.644
	GCNET	0,8537	0,0557	0,1527	0,0373	49.617.125
	POINTREND	0,8562	0,0670	0,1281	0,0414	28.715.119
	PSANET	0,8529	0,0474	0,1680	0,0360	59.125.604
	SEGFORMER	0,8435	0,0529	0,1098	0,0270	13.678.786
area 3	FCN	0,9134	0,0662	0,1513	0,0265	49.484.644
	GCNET	0,9169	0,0579	0,1522	0,0304	49.617.125
	POINTREND	0,8978	0,0640	0,1282	0,0416	28.715.119
	PSANET	0,9149	0,0736	0,1684	0,0332	59.125.604
	SEGFORMER	0,9224	0,0541	0,1101	0,0282	13.678.786
Média	FCN	0,8219	0,1145	0,1514	0,0237	49.484.644
	GCNET	0,8238	0,1134	0,1514	0,0281	49.617.125
	POINTREND	0,8267	0,1147	0,1274	0,0353	28.715.119
	PSANET	0,8246	0,1156	0,1668	0,0296	59.125.604
	SEGFORMER	0,8270	0,1128	0,1097	0,0241	13.678.786

A Figura 15 ilustra a eficiência relativa dessas variáveis. Para o cálculo, foi atribuído o valor de 100% ao melhor desempenho médio em cada variável. No caso do IoU, por ser diretamente proporcional à qualidade da segmentação, o maior valor foi considerado o ideal. Para o tempo de inferência e o número de parâmetros, por serem inversamente proporcionais à eficiência, os menores valores foram considerados como referência (100%), e os demais expressos proporcionalmente em relação a eles.

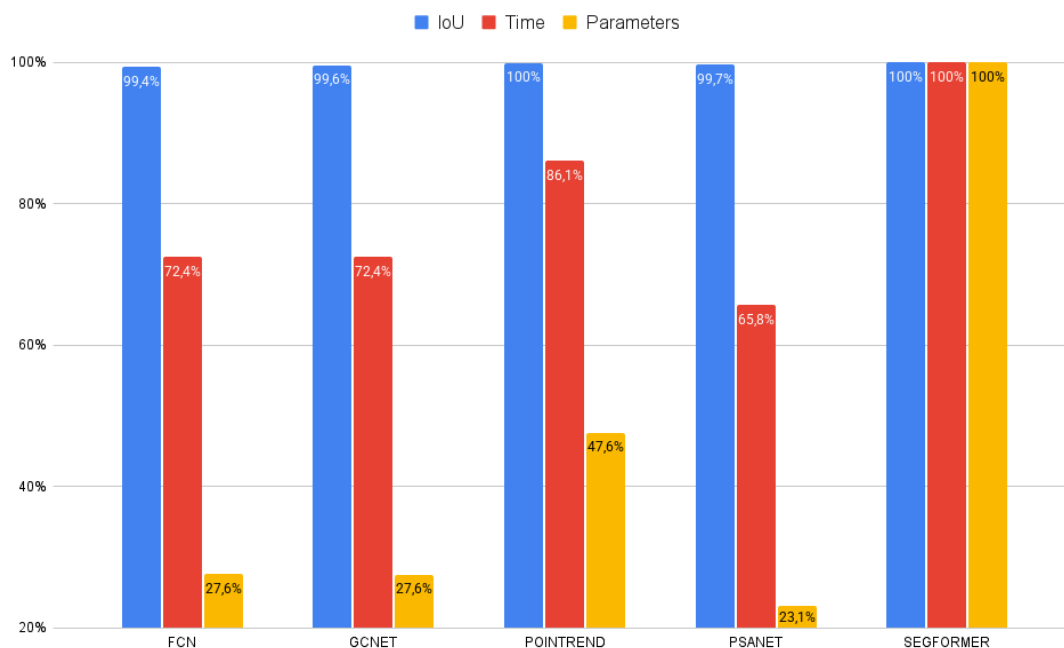


Figura 15. Comparação entre as arquiteturas de segmentação semanticamente avaliadas quanto ao desempenho (IoU), tempo de inferência e número de parâmetros. Os valores foram normalizados proporcionalmente ao melhor desempenho observado para cada métrica, atribuindo-se 100% à arquitetura com melhor valor médio em cada categoria.

Observa-se que a SegFormer apresentou os melhores resultados nas três variáveis, servindo como referência para as demais. Embora todas as arquiteturas tenham apresentado valores de IoU muito próximos entre si (acima de 99% em relação ao melhor), houve maior variação no tempo de inferência e, principalmente, na quantidade de parâmetros, evidenciando diferenças relevantes em termos de eficiência computacional. Isso reforça que, em contextos de aplicação prática, a escolha da arquitetura pode ser orientada não apenas pela acurácia, mas também por critérios como velocidade e demanda computacional.

3.3.2. Avaliação Altimétrica

A avaliação altimétrica foi conduzida em duas etapas complementares: a avaliação visual da nuvem densa de pontos (Seção 3.3.2.1) e a avaliação visual dos Modelos Digitais de Terreno (MDTs) (Seção 3.3.2.2). Na primeira, buscou-se verificar visualmente a eficiência dos métodos aplicados na remoção das copas de árvores, por meio do filtro tradicional CSF e da segmentação semântica via *deep learning*. Já na segunda, a análise concentrou-se nos MDTs gerados a partir dessas nuvens, permitindo

identificar discrepâncias altimétricas, avaliar a preservação da topografia real e comparar os resultados obtidos com pontos de checagem de referência. Dessa forma, os dois níveis de avaliação — nuvem de pontos e MDT — fornecem uma visão integrada da qualidade e confiabilidade dos métodos testados para representação do relevo.

3.3.2.1 Avaliação visual da nuvem densa de pontos

A aplicação do método *Cloth Simulation Filtering* (CSF) requer a definição de três parâmetros principais: *Cloth resolution*, que determina o espaçamento da malha de pontos (valores mais altos resultam em nuvens mais grosseiras); *Max iteration*, que define o número máximo de iterações (padrão de 500); e *Classification threshold*, limiar que classifica se um ponto pertence ou não ao terreno.

Na presente pesquisa, adotaram-se os parâmetros *Cloth resolution* = 0,4 (considerando a ortofoto com GSD de 1,36 cm/px), *Max iteration* = 500 (valor padrão) e *Classification threshold* variando entre 0,4; 1,4; e 2,4.

Conforme ilustrado nas Figuras 15, 16 e 17, a redução do *Classification threshold* de 2,4 para 0,4 resultou em melhoria visual da filtragem, com remoção mais eficiente dos pontos referentes às copas das árvores. Entretanto, na Figura 17, observa-se que o parâmetro mais restritivo também eliminou parte da estrada, que deveria ser preservada como superfície do terreno, possivelmente em razão da inclinação acentuada da área 3.

Assim, com base na análise visual, optou-se pelo valor *Classification threshold* = 0,4 para a geração do MDT pelo método tradicional (Figura 6).

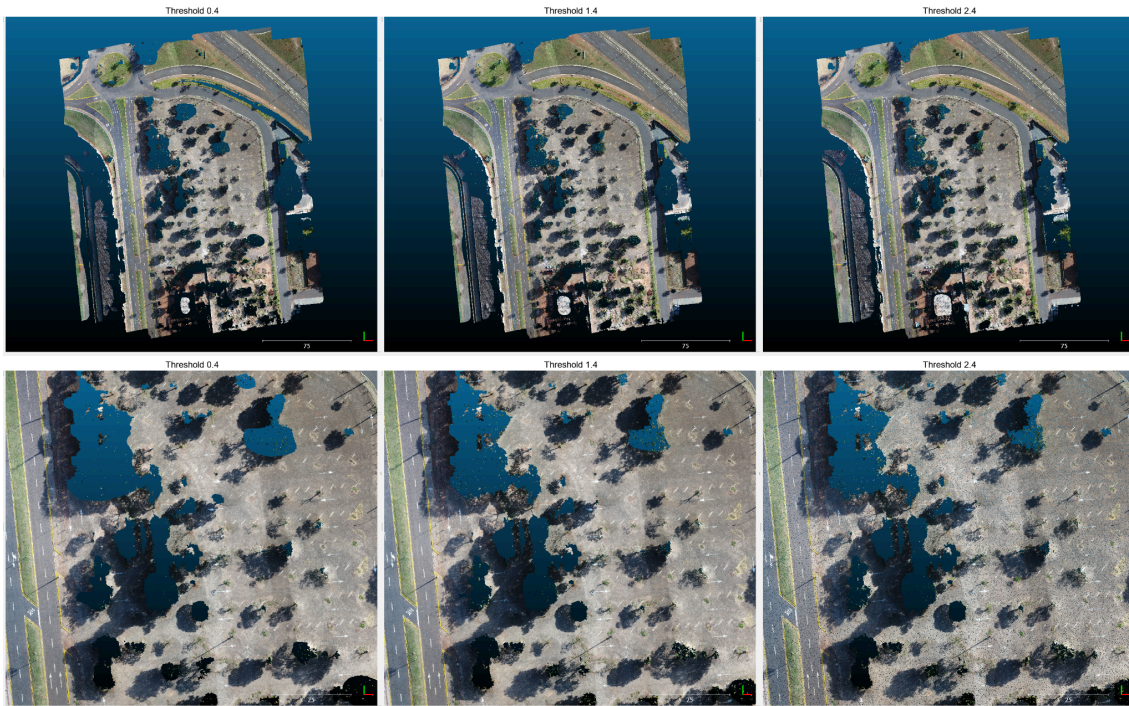


Figura 15. Nuvens de pontos filtradas no CloudCompare com o método CSF para a área 1. *Classification threshold* 0.4 na primeira coluna, 1.4 na segunda e 2.4 na terceira. A primeira linha mostra a área total e a segunda uma aproximação para melhor visualização.

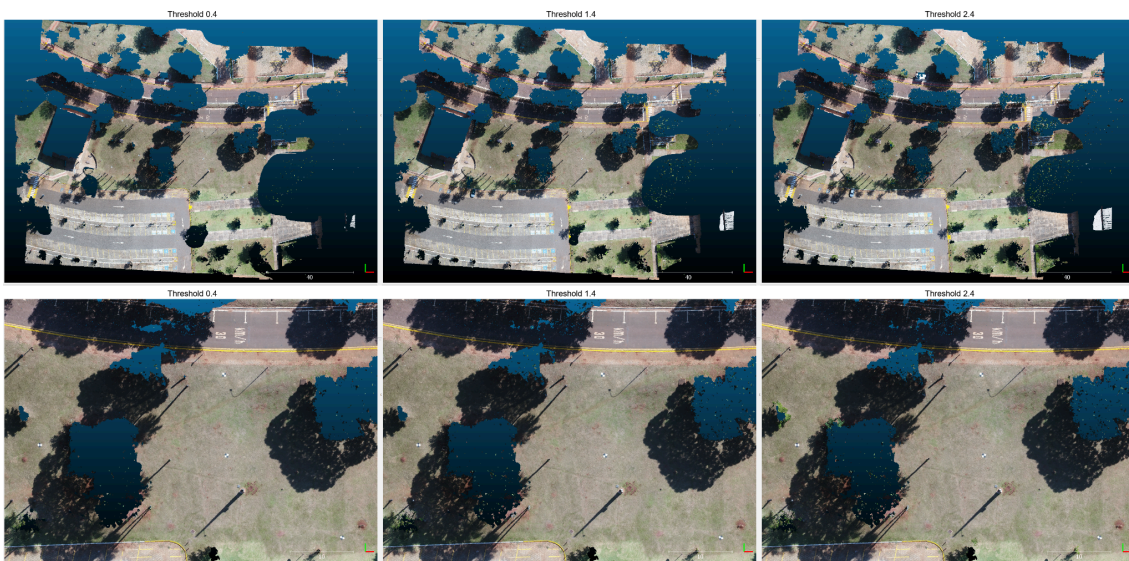


Figura 16. Nuvens de pontos filtradas no CloudCompare com o método CSF para a área 2. *Classification threshold* 0.4 na primeira coluna, 1.4 na segunda e 2.4 na terceira. A primeira linha mostra a área total e a segunda uma aproximação para melhor visualização.

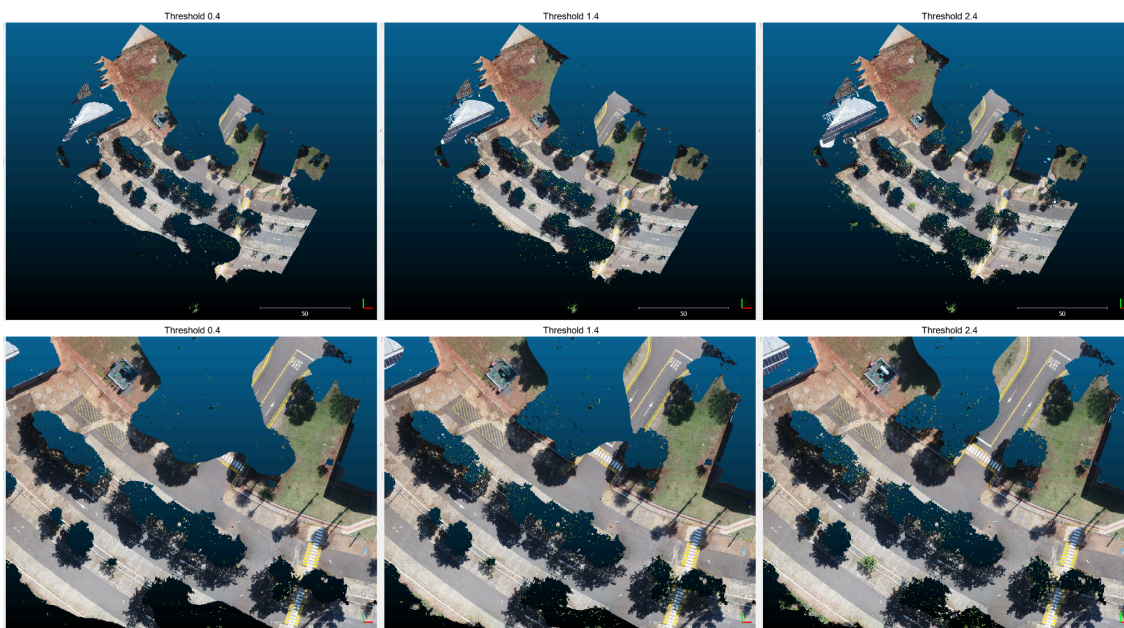


Figura 17. Nuvens de pontos filtradas no CloudCompare com o método CSF para a área 3. *Classification threshold* 0.4 na primeira coluna, 1.4 na segunda e 2.4 na terceira. A primeira linha mostra a área total e a segunda uma aproximação para melhor visualização.

As nuvens de pontos derivadas do método proposto, a partir da segmentação semântica com a arquitetura SegFormer (Figuras 18, 19 e 20), apresentaram remoção satisfatória dos pontos correspondentes às copas das árvores. Comparativamente, o método tradicional com o filtro CSF ($\text{threshold} = 0,4$) removeu edificações e árvores de forma eficaz, mas manteve alguns pontos residuais nas copas e, em contrapartida, eliminou também pontos legítimos do terreno.

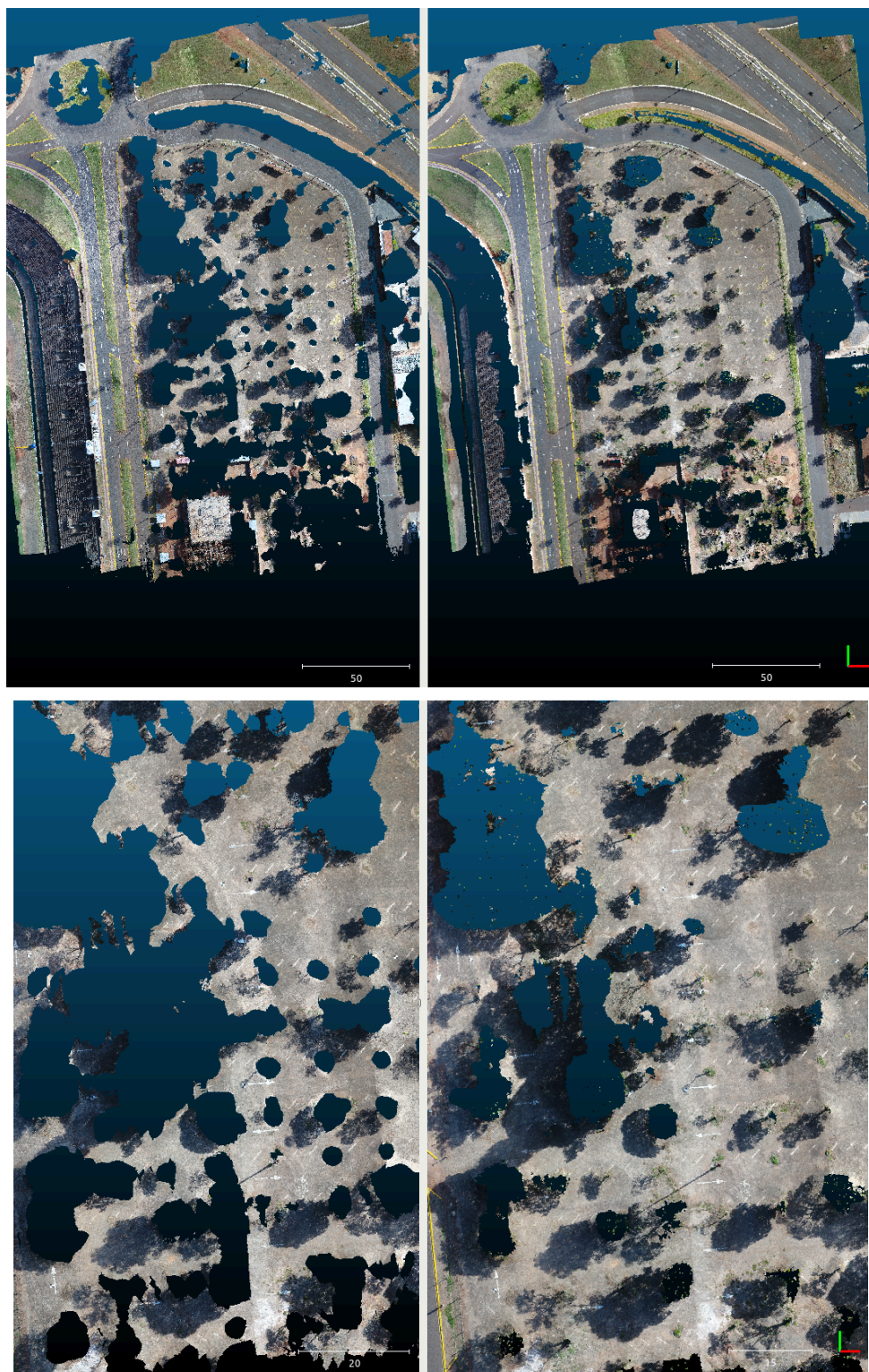


Figura 18. Nuvens de pontos aplicando a segmentação pela DL e aplicando o filtro CSF, para a área 1. Primeira linha a área toda, na segunda linha um recorte mais detalhado.

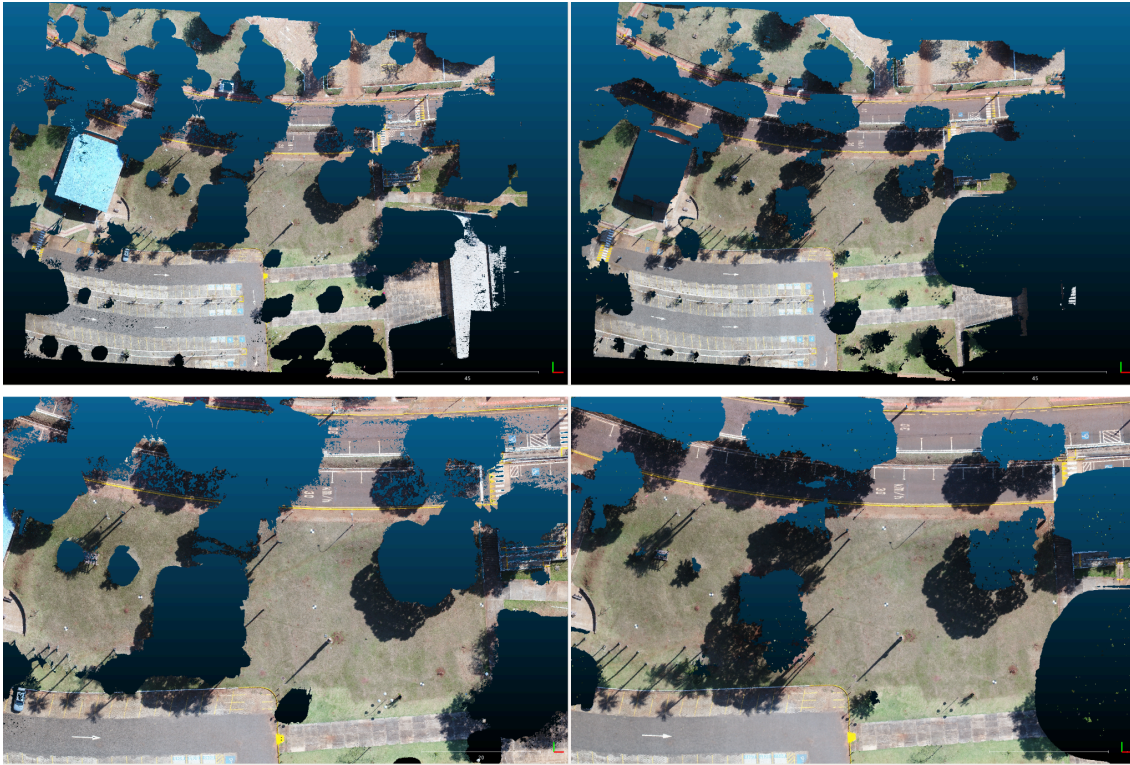


Figura 19. Nuvens de pontos aplicando a segmentação pela DL e aplicando o filtro CSF, para a área 2. Primeira linha a área toda, na segunda linha um recorte mais detalhado.

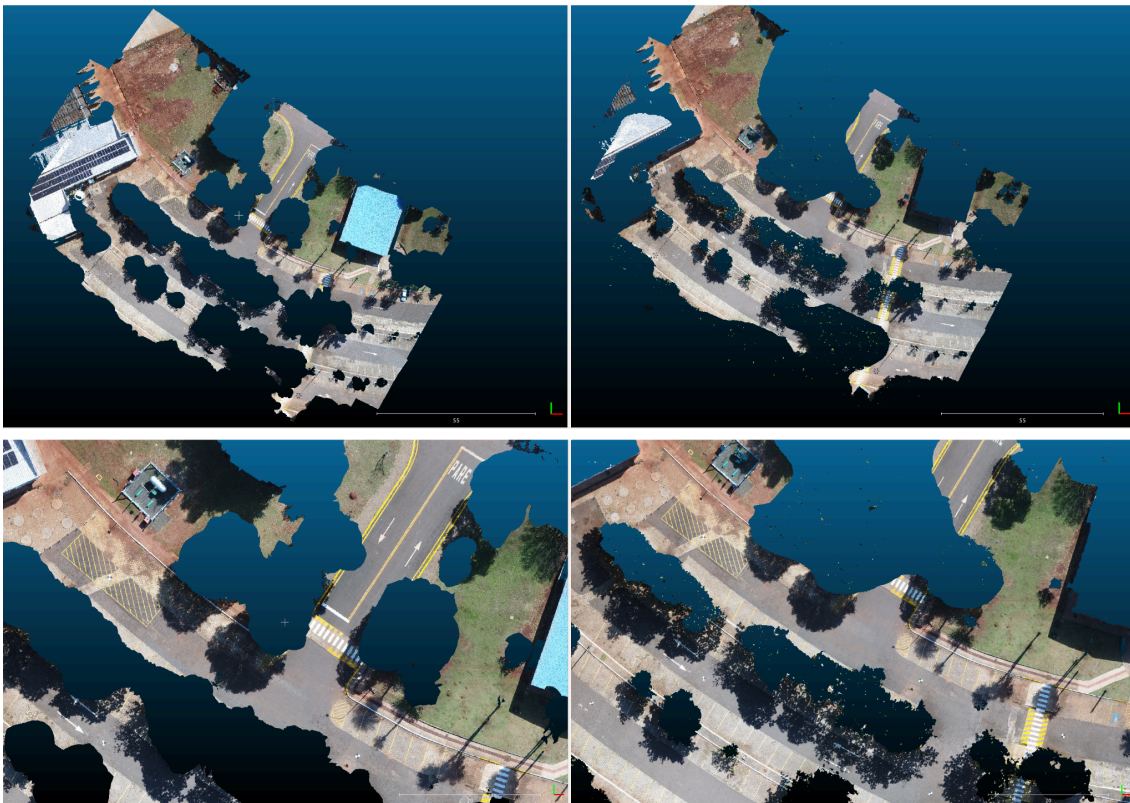
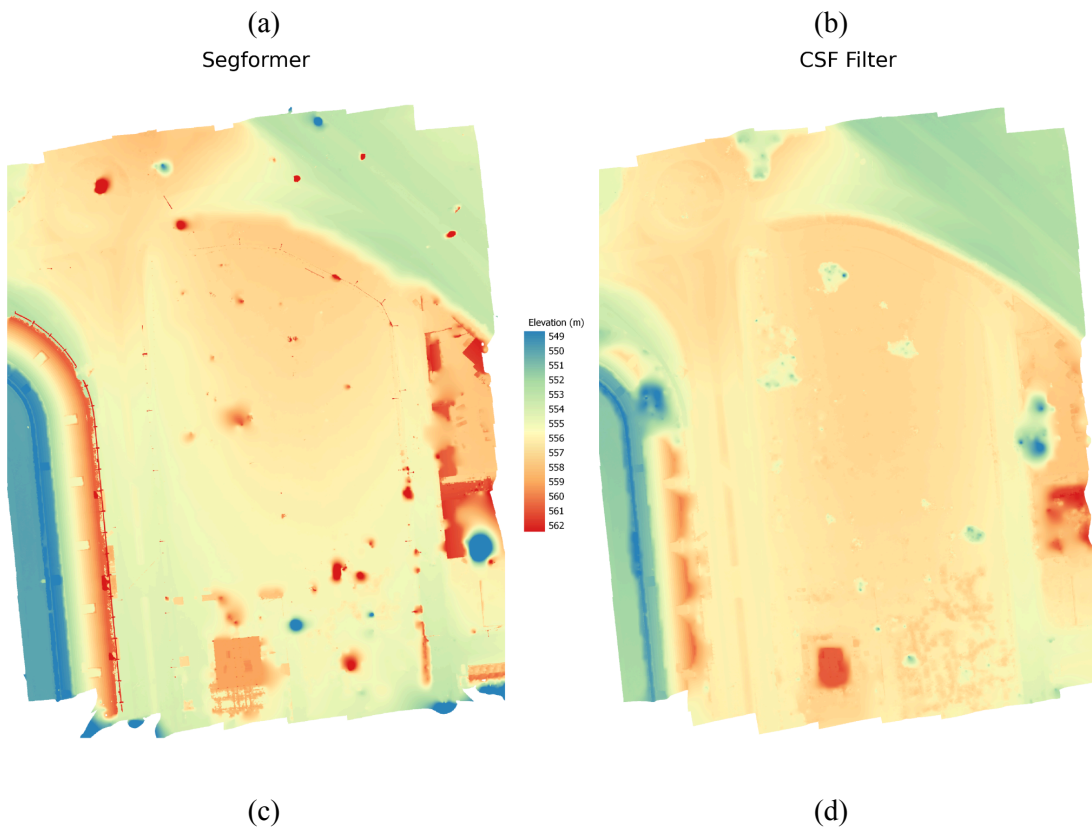


Figura 20. Nuvens de pontos aplicando a segmentação pela DL e aplicando o filtro CSF, para a área 3. Primeira linha a área toda, na segunda linha um recorte mais detalhado.

3.3.2.2 Avaliação visual dos MDTs

Na Figura 21 (a, c, e) apresentam-se os MDTs obtidos pelo método proposto. Verifica-se que as maiores altitudes correspondem às edificações, que não foram eliminadas pelas máscaras aplicadas. Na região do estacionamento, caracterizada por árvores esparsas e relevo plano (Figura 18), observam-se áreas pontuais em vermelho, indicativas de maiores altitudes. Essas discrepâncias configuram erros, uma vez que não são compatíveis com a topografia local. Parte desses pontos decorre da permanência de copas de árvores, de objetos com altura distinta de vegetação (p. ex., postes) ou de falhas no processo de segmentação.

Na Figura 21 (b, d, f) observam-se inconsistências em azul, correspondentes a altitudes inferiores ao terreno. Inicialmente, esses pontos — também visíveis nas Figuras 18, 19 e 20 — foram interpretados como copas de árvores. No entanto, por apresentarem valores mais baixos que a própria superfície do terreno, conclui-se que se trata, na realidade, de ruídos resultantes da classificação incorreta pelo filtro.



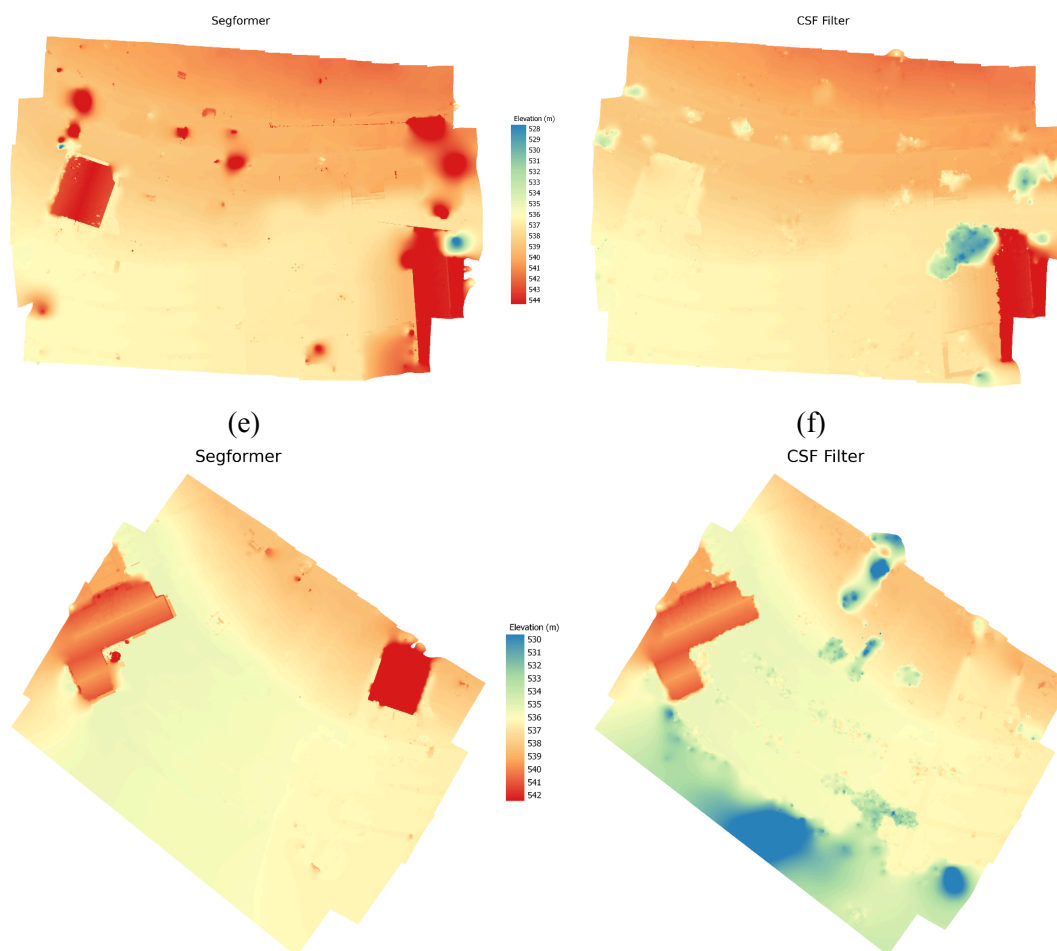


Figura 21. MDTs das áreas 1 (a, b), 2 (c, d) e 3 (e, f).

Nas Figuras 22, 23 e 24 são apresentadas as principais diferenças entre os MDTs. Para facilitar a interpretação, a primeira coluna mostra o recorte da ortoimagem, a segunda o MDT segmentado pelo SegFormer, e a terceira o MDT obtido pelo método tradicional com o filtro CSF.

De modo geral, o MDT segmentado preserva alguns pontos mais elevados (em vermelho nos recortes), provavelmente devido a resquícios de vegetação arbórea mantidos na nuvem de pontos. Já o método tradicional também apresenta inconsistências, mas em áreas com valores mais baixos (em azul nos recortes), provavelmente ruídos.

Na área 1 (Figura 22), o detalhe 1 mostra que o MDT segmentado representou corretamente o terreno plano do estacionamento, enquanto o filtrado indicou variações como se houvesse um buraco. Nos detalhes 2 e 3, ambos os métodos foram impactados por árvores sem ou com poucas folhas: no segmentado permaneceram pontos altos, e no tradicional surgiram depressões artificiais.

Na área 2 (Figura 23), os detalhes 1 e 2 mostram que o MDT segmentado não conseguiu representar fielmente a superfície do terreno, apresentando variações artificiais de altitude. O método tradicional também apresentou distorções, com valores mais baixos indicando depressões inexistentes. Já no detalhe 3, o método proposto conseguiu representar de forma adequada o relevo, enquanto o tradicional gerou um buraco na posição da árvore.

Na área 3 (Figura 24), nos três detalhes o método proposto representou satisfatoriamente o relevo, ao passo que o tradicional produziu novamente depressões artificiais no local das árvores.

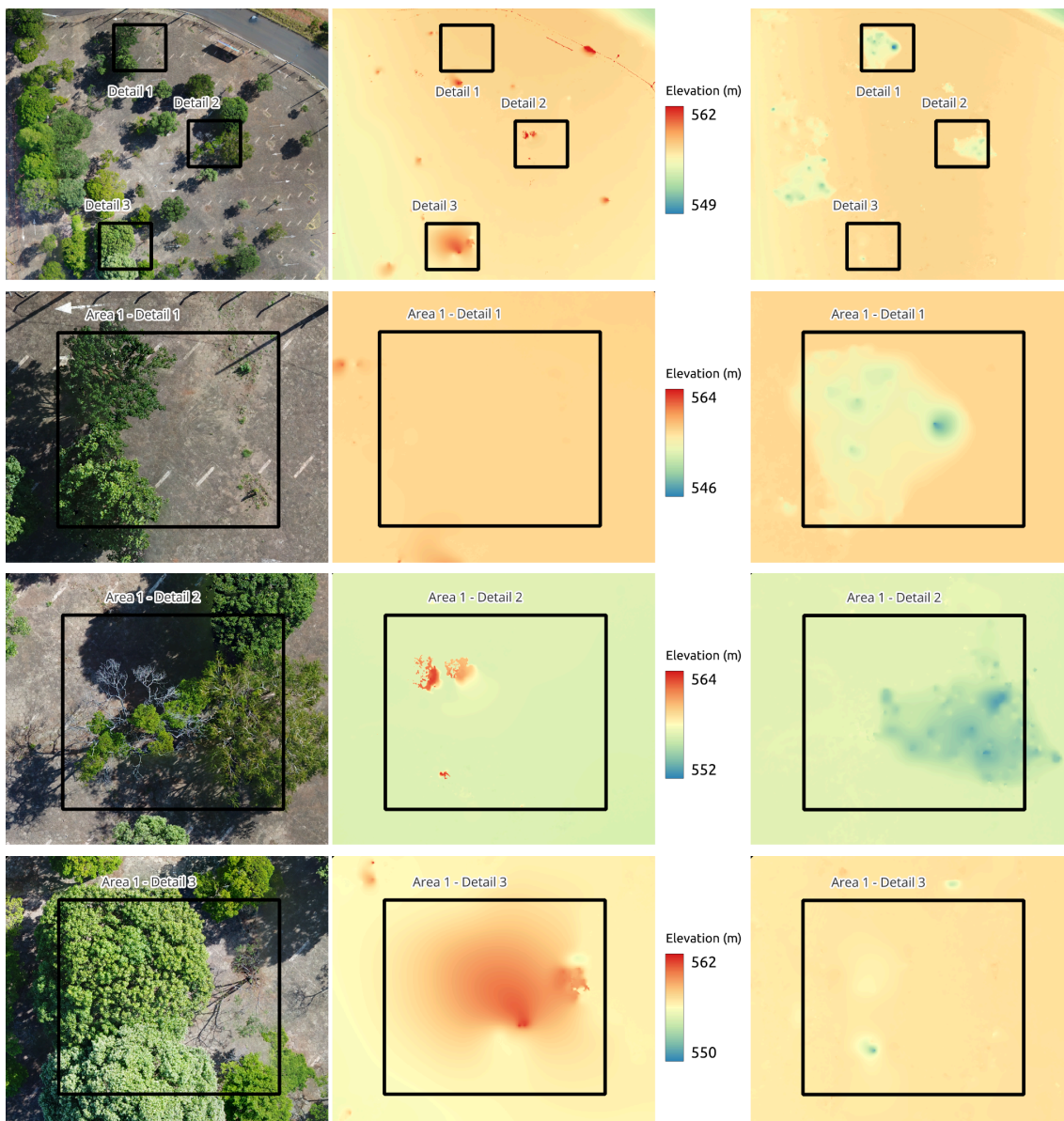


Figura 22. Detalhamento entre as principais diferenças dos MDTs, na área 1.

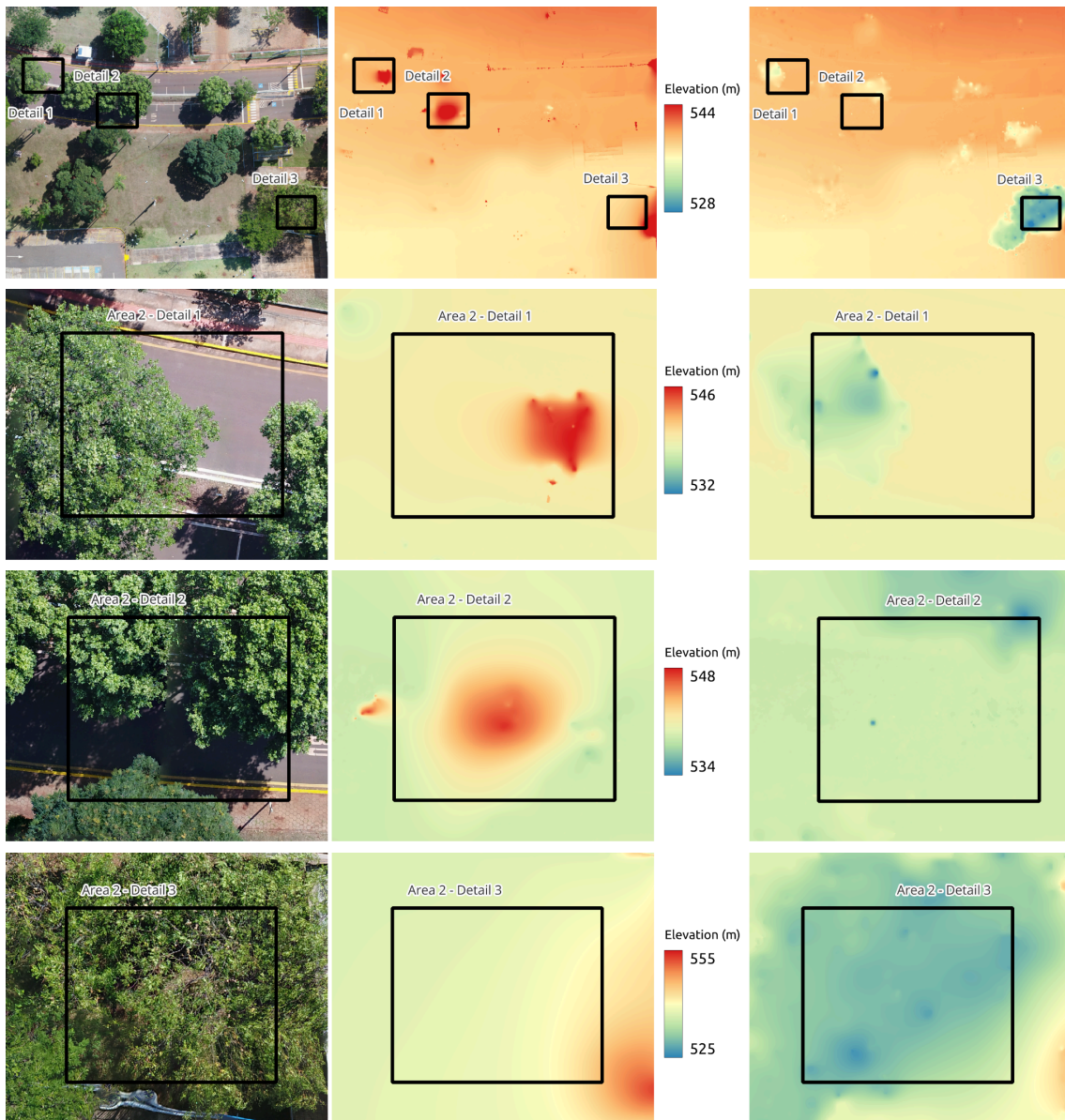


Figura 23. Detalhamento entre as principais diferenças dos MDTs, na área 2.

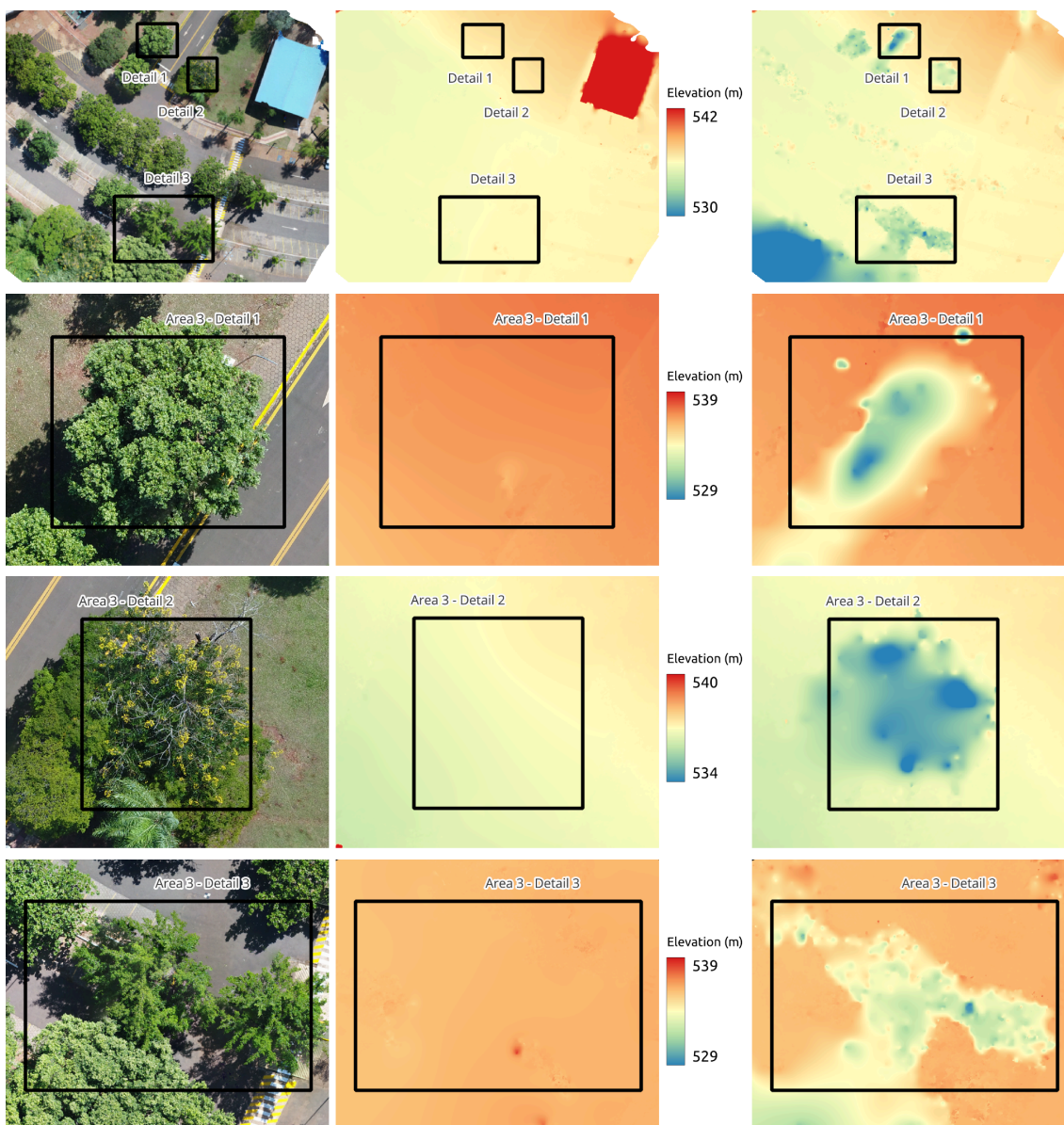


Figura 24. Detalhamento entre as principais diferenças dos MDTs, na área 3.

3.3.2.3 Avaliação das discrepâncias dos MDTs em relação aos pontos de checagem

Na Tabela 4, são apresentadas as discrepâncias dos MDTs nas áreas 1, 2 e 3, a partir de 232 pontos de checagem obtidos por GNSS RTK. O MDT filtrado apresentou os menores valores mínimos (-4,43 m na área 3), evidenciando a retenção de ruídos associados a altitudes subestimadas. Já o MDT segmentado mostrou discrepâncias mínimas mais próximas de zero (entre -0,82 e -0,10 m), indicando menor tendência a gerar depressões artificiais.

Por outro lado, a discrepância máxima foi mais alta no MDT segmentado (2,49 m na área 1), enquanto no filtrado os maiores desvios não ultrapassaram 1,43 m. Em termos de RMSE e desvio padrão, o filtrado apresentou maior dispersão em duas áreas (RMSE de 0,79 m na área 1 e 0,59 m na área 3), enquanto o segmentado manteve valores mais baixos e consistentes (0,44 m, 0,19 m e 0,13 m nas áreas 1, 2 e 3, respectivamente).

De forma geral, o MDT baseado em segmentação preservou melhor a superfície do terreno, reduzindo erros de depressão, mas apresentou tendência de superestimar a altitude em algumas regiões, possivelmente por falhas de segmentação de copas ou efeitos de borda. Já o MDT filtrado mostrou maior variabilidade nos extremos, sobretudo nas áreas de maior declividade, embora em alguns casos tenha registrado médias mais próximas de zero.

Tabela 4. Análise das discrepâncias do MDT das 03 áreas para os dados do GNSS RTK.

	MDT segmentação - Altitude			MDT filtro - Altitude		
Área	1	2	3	1	2	3
Mínimo	-0,818	-0,103	-0,246	-2,335	-0,174	-4,428
Máximo	2,494	1,590	0,468	1,434	0,778	0,267
RMSE	0,441	0,194	0,127	0,795	0,137	0,589
Média	0,065	0,038	0,081	0,412	0,027	-0,010
Des. P.	0,195	0,070	0,060	0,460	0,062	0,195

3.3.3. Análise visual Ortofotos

Na análise visual das ortofotos, pelo método proposto e utilizando o MDS, sem nenhum processo de filtragem, temos as Figuras 25, 26 e 27, demonstrando a definição das árvores individuais geradas com a segmentação e sem o uso dela, para cada área de estudo.

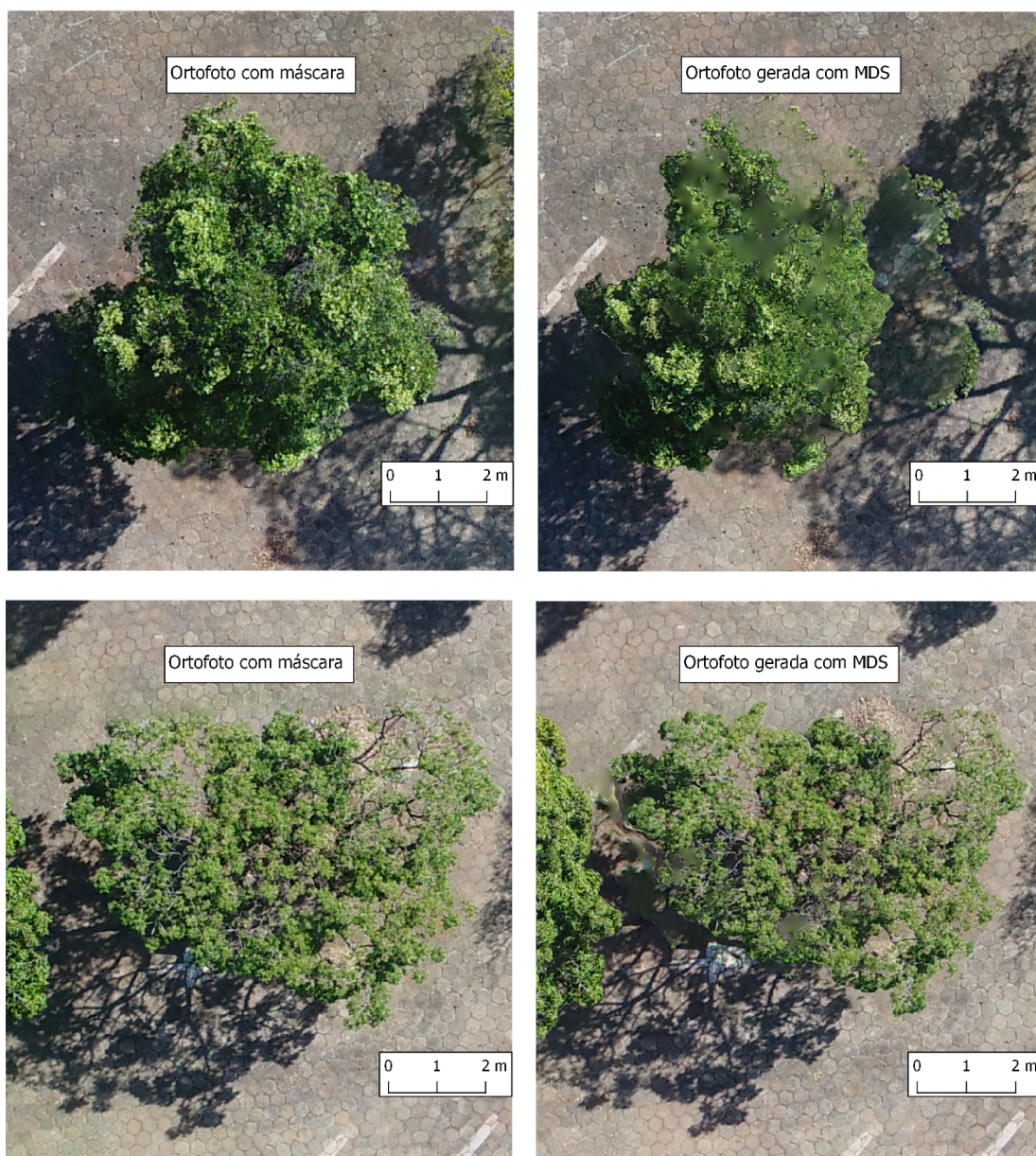


Figura 25. Recortes das ortofotos da área 1, a esquerda ortofoto gerada pelo metodo proposto, da direita ortofoto gerada com MDS.



Figura 26. Recorte da ortofoto da área 2, a esquerda ortofoto gerada pelo método proposto, da direita ortofoto gerada com MDS.



Figura 27. Recorte da ortofoto da área 3, a esquerda ortofoto gerada com pelo método proposto, da direita ortofoto gerada com MDS.

De forma geral, o uso do método proposto com a segmentação das árvores através da DL resultou em imagens de árvores individuais mais definidas. Para a área 1 (Figura 25), isso fica bem nítido, com a segmentação as bordas das árvores ficam mais definidas e sem manchas suavizadas. Na área 2 e 3 (Figuras 26 e 27), a melhora não fica tão nítida. Em campo, na área da Figura 26, tem-se três palmeiras, duas nas bordas que ficam nítidas com o uso de segmentação, e uma no centro da copa de outra árvore, como observa-se sem o uso de filtragem (Figura 28).



Figura 28. Recorte da foto da área 2, verificação das 3 palmeiras.

3.4. DISCUSSÃO

A segmentação semântica aplicada por meio da arquitetura SegFormer apresentou resultados eficientes em comparação com outras abordagens, corroborando os achados de Spasev et al. (2024), que avaliaram suas variantes (B0 a B5) no conjunto de dados UAVID. Nesse estudo, os autores demonstraram que o SegFormer não apenas supera arquiteturas convolucionais tradicionais, mas também equilibra precisão e eficiência, sendo capaz de lidar com classes complexas em ambientes urbanos, como ruído de fundo, edificações, vias, árvores, vegetação baixa, carros em movimento, carros estacionados e pedestres. Embora ainda persistam alguns ruídos associados à vegetação e a objetos elevados, a abordagem mostrou potencial superior ao método tradicional por evitar a eliminação indevida de superfícies reais do terreno.

Além disso, comparações sistemáticas entre arquiteturas de deep learning revelam diferenças relevantes. Formichini e Avizzano (2025) realizaram uma análise envolvendo 15 redes neurais convolucionais aplicadas à classificação de terrenos a partir de imagens aéreas (LandCover.ai), explorando classes como edificações, áreas florestais, corpos d'água e vias. O estudo evidenciou que, embora a acurácia global tenha atingido até 99%, a análise dos erros (outliers) revelou que modelos como PSPNet, FCN e ICNet apresentaram maior robustez e menor ocorrência de falhas críticas. Esse tipo de investigação demonstra que a escolha da arquitetura impacta não apenas o desempenho médio, mas também a consistência em cenários desafiadores.

No contexto agrícola, Zhu et al. (2023) propuseram o PD-SegNet (Powerful Decoder SegFormer Network), um aprimoramento do SegFormer voltado para alvos pequenos em ambientes complexos, como flores e frutos em pomares. A inovação central foi a introdução do Dynamic Kernel Head, capaz de atualizar dinamicamente os kernels durante a segmentação, e do Complex Points Head, que melhora a detecção de bordas em objetos de difícil distinção. Os autores validaram o modelo no dataset MinneApple, alcançando o estado da arte na segmentação de maçãs mesmo em condições adversas de iluminação, clima e oclusão. Esse avanço reforça o potencial das arquiteturas baseadas em transformadores para aplicações de agricultura de precisão, incluindo estimativas de florescimento, produtividade e colheita automatizada.

Dessa forma, observa-se que a evolução das arquiteturas de segmentação semântica vem ampliando não apenas a precisão em ambientes urbanos e naturais, mas também sua aplicabilidade em cenários agrícolas e de monitoramento ambiental, permitindo a criação de modelos mais especializados e robustos para diferentes objetos acima do terreno.

Por outro lado, o método tradicional, baseado no filtro Cloth Simulation Filtering (CSF), demonstrou limitações, sobretudo em áreas de maior declividade ou em superfícies mais heterogêneas. Apesar de eficiente na remoção de árvores e edificações, o filtro apresentou erros recorrentes, como a eliminação indevida de pontos legítimos do terreno (exemplo da estrada na área 3) e a introdução de depressões artificiais nos MDTs. Esses problemas decorrem da sensibilidade dos parâmetros do algoritmo, principalmente do classification threshold, que, quando ajustado para maior rigor, tende a comprometer superfícies planas e contínuas.

Essas limitações já haviam sido observadas por Zhang et al. (2016), que apresentaram o CSF como uma alternativa simples e eficiente para dados LiDAR

aéreos, mas destacaram sua dependência da calibração adequada dos parâmetros. Posteriormente, Zhou et al. (2020) compararam diferentes algoritmos de filtragem em dados UAV-LiDAR e constataram que, embora o CSF fosse robusto em florestas com diferentes densidades, ainda apresentava falhas em áreas de vegetação muito densa ou em terrenos íngremes. Da mesma forma, Wang et al. (2023), ao avaliar nuvens de pontos geradas por fotogrametria de UAV, confirmaram a versatilidade do CSF em diferentes cenários, mas também evidenciaram sua suscetibilidade a erros em ambientes complexos, recomendando métodos híbridos. Avanços mais recentes, como o CSF4Geomesh proposto por Yu et al. (2022), buscaram superar essas deficiências ao aplicar a filtragem em malhas 3D, utilizando detecção hierárquica de colisão e correções por linhas de varredura, o que ampliou sua robustez e reduziu erros associados à densidade irregular de pontos.

Na avaliação dos MDTs, verificaram-se discrepâncias tanto nos modelos obtidos via segmentação quanto nos derivados do filtro CSF. O método tradicional apresentou maior variabilidade, com ruídos de altitudes negativas e depressões artificiais. Já a segmentação preservou melhor a superfície do terreno, mas apresentou tendência a superestimar altitudes em alguns pontos, provavelmente devido a falhas na separação de copas ou efeitos de borda. Essas inconsistências indicam que ambos os métodos ainda carecem de ajustes, especialmente quando aplicados em áreas com topografia complexa ou objetos diversos sobre o solo.

Embora o método proposto não tenha gerado um MDT completo — já que a segmentação removeu apenas árvores e não outros elementos sobre o terreno —, as análises visuais e métricas realizadas em áreas com vegetação esparsa permitiram considerar os resultados como representativos de um MDT. Além disso, o procedimento permitiu derivar um Modelo Digital sem Vegetação (MDnV), categoria de dado digital específica que tem se mostrado importante em aplicações hidrológicas e de modelagem de inundações (SCHUBERT et al., 2008; FEWTRELL et al., 2011; NEAL et al., 2011). Essa limitação, contudo, evidencia a necessidade de evolução da técnica, incluindo treinamentos específicos para diferentes classes de objetos. Com o avanço das arquiteturas de deep learning e a maior disponibilidade de bases de dados para treinamento, torna-se factível o desenvolvimento de modelos mais abrangentes, capazes de diferenciar e remover múltiplos elementos acima do solo. Assim, a proposta apresenta-se como uma abordagem promissora, com potencial para superar as

limitações do filtro CSF e gerar MDTs mais precisos e consistentes em cenários complexos.

3.5. CONCLUSÃO

O estudo evidenciou que o MDT obtido por segmentação preserva melhor o relevo em comparação ao método filtrado, ainda que apresente superestimação da altitude em áreas com vegetação. Embora o experimento tenha focado apenas na classe "árvores", o método de segmentação pode ser expandido para outras classes, como edificações e veículos, o que amplia sua aplicabilidade. Além disso, a contínua evolução das arquiteturas e técnicas de deep learning tende a facilitar o processo e aprimorar os resultados, reforçando o potencial da segmentação como alternativa promissora para a geração de MDTs mais precisos e adaptáveis a diferentes cenários.

3.6. REFERÊNCIAS

- AGISOFT. **Agisoft Metashape User Manual**. Disponível em: <https://www.agisoft.com/pdf/metashape-pro_1_6_en.pdf> acesso em: 30/11/2020.
- ANDERS, Niels et al. Comparing filtering techniques for removing vegetation from UAV-based photogrammetric point clouds. **Drones**, v. 3, n. 3, p. 61, 2019.
- ANNIS, Antonio et al. UAV-DEMs for small-scale flood hazard mapping. **Water**, v. 12, n. 6, p. 1717, 2020.
- ANZIDEI, Marco et al. Coastal retreat and marine flooding scenario for 2100: A case study along the coast of Maddalena Peninsula (southeastern Sicily). **Geogr. Fis. Din. Quat**, v. 41, p. 5-16, 2018.
- BATISTOTI, J., et al. Estimating Pasture Biomass and Canopy Height in Brazilian Savanna Using UAV Photogrammetry. **Remote Sensing** 11.20 (2019): 2447.
- BELCORE, Elena et al. Individual Tree Detection from UAV Imagery Using Hölder Exponent. **Remote Sensing**, v. 12, n. 15, p. 2407, 2020.
- FEWTRELL, T. J.; DUNCAN, A.; SAMPSON, C. C.; NEAL, J. C.; BATES, P. D. Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. **Physics and Chemistry of the Earth**, v. 36, n. 7–8, p. 281–291, 2011.
- FORMICHINI, M.; AVIZZANO, C. A. A comparative analysis of deep learning-based segmentation techniques for terrain classification in aerial imagery. **AI**, v. 6, n. 7, p. 145, 2025. DOI: 10.3390/ai6070145.
- GEVAERT, C. M. et al. A deep learning approach to DTM extraction from imagery using rule-based training labels. **ISPRS journal of photogrammetry and remote sensing**, v. 142, p. 106-123, 2018.
- JAKOVLJEVIC, Gordana et al. Accuracy assessment of deep learning based classification of LiDAR and UAV points clouds for DTM creation and flood risk mapping. **Geosciences**, v. 9, n. 7, p. 323, 2019.
- JEON, Jeongbae; CHOI, Won. Prediction accuracy of reservoir break flood simulation model using finite volume method and UAV. **International Journal of Agricultural and Biological Engineering**, v. 13, n. 6, p. 7-15, 2020.
- KIRILLOV, A., WU, Y., HE, K., GIRSHICK, R. Pointrend: Image segmentation as rendering, in: **Proceedings of the IEEE/CVF conference on computer vision and pattern recognition**, 2020, pp. 9799–9808.50
- LEAL-ALVES, Deivid Cristian et al. Unmanned Aerial Vehicle and Structure from Motion Approach for Flood Assessment in Coastal Channels. **Journal of Coastal Research**, v. 95, n. SI, p. 1162-1166, 2020.

- LEWIŃSKA, Paulina et al. Evaluation of structure-from-motion for analysis of small-scale glacier dynamics. **Measurement**, v. 168, p. 108327, 2020.
- LONG, Jonathan; SHELHAMER, Evan; DARRELL, Trevor. Fully convolutional networks for semantic segmentation. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2015. p. 3431-3440.
- MEDVEDEV, Andrey et al. UAV-derived data application for environmental monitoring of the coastal area of Lake Sevan, Armenia with a changing water level. **Remote Sensing**, v. 12, n. 22, p. 3821, 2020.
- MEINEN, B. U., & ROBINSON, D. T. Where did the soil go? Quantifying one year of soil erosion on a steep tile-drained agricultural field. **Science of The Total Environment**, 729, 2020. p. 138320.
- MENG, Xuelian; CURRIT, Nate; ZHAO, Kaiguang. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. **Remote Sensing**, v. 2, n. 3, p. 833-860, 2010.
- MMSegmentation Contributors: **Openmmlab semantic segmentation toolbox and benchmark**, <https://github.com/open-mmlab/mmdetection> (2020).
- NA, Jiaming et al. UAV-Based Terrain Modeling under Vegetation in the Chinese Loess Plateau: A Deep Learning and Terrain Correction Ensemble Framework. **Remote Sensing**, v. 12, n. 20, p. 3318, 2020.
- NARANJO, Sergio et al. Effects of spatial resolution of terrain models on modelled discharge and soil loss in Oaxaca, Mexico. **Hydrology and Earth System Sciences**, v. 25, n. 10, p. 5561-5588, 2021.
- NEAL, J.; SCHUMANN, G.; FEWTRELL, T.; BUDIMIR, M.; BATES, P.; MASON, D. Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. **Journal of Flood Risk Management**, v. 4, n. 4, p. 295–310, 2011.
- OH, S.-J.; JUNG, M.-J.; LIM, C.; SHIN, S.-C. Automatic Detection of Welding Defects Using Faster R-CNN. **Appl. Sci.**, v. 10, n. 23, p. 8629, 2020.
- RYAN, Johnny C. et al. UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet. **The Cryosphere**. v. 9, p. 1 -11, 2015.
- SCHNEIDER, Caroline A.; RASBAND, Wayne S.; ELICEIRI, Kevin W. NIH Image to ImageJ: 25 years of image analysis. **Nature methods**, v. 9, n. 7, p. 671-675, 2012.
- SCHUBERT, J. E.; SANDERS, B. F.; SMITH, M. J.; WRIGHT, N. G. Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. **Advances in Water Resources**, v. 31, n. 12, p. 1603–1621, 2008.

- SPASEV, V.; DIMITROVSKI, I.; CHORBEV, I.; KITANOVSKI, I. Semantic segmentation of unmanned aerial vehicle remote sensing images using SegFormer. **arXiv preprint**, arXiv:2410.01092, 2024.
- TRIMBLE (2013) - **Trimble R8, R6 and R4 User Guide Version 4.80 Revision A**. Disponível em: <https://www.trimble.com/support_trl.aspx?Nav=Collection-65944&pt=Trimble%20R4>. Acesso em: 2 out. 2021.
- WANG, Y.; ZHANG, J.; ZHOU, G.; JIANG, W.; LU, X.; CHENG, Y. Comparative Analysis of Ground Filtering Algorithms for UAV Photogrammetry Point Clouds. **Forests**, v. 14, n. 7, p. 1520, 2023. DOI: 10.3390/f14071520.
- YU, H.; XU, Z.; ZHU, Q.; ZHANG, Y.; GUO, B.; WANG, W.; SHI, X. Unsupervised Ground Filtering of Airborne-Based 3D Meshes Using a Robust Cloth Simulation. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 188, p. 65-78, 2022. DOI: 10.1016/j.isprsjprs.2022.03.004.
- ZHANG, W. et al. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. **Remote Sensing**, v. 8, n. 6, p. 501, 2016.
- ZHAO, H., ZHANG, Y., LIU, S., SHI, J., CHANGE LOY, C., LIN, D., JIA, J. Psanet: Point-wise spatial attention network for scene parsing, in: **Proceedings of the European Conference on Computer Vision (ECCV)**, 2018, pp. 267–283.
- ZHOU, W.; ZHANG, C.; TANG, T.; LI, W.; YAN, K. Evaluation of ground filtering algorithms for UAV LiDAR data in forested areas. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 166, p. 333-343, 2020. DOI: 10.1016/j.isprsjprs.2020.06.011.
- ZHU, Z.; JIANG, M.; DONG, J.; WU, S.; MA, F. PD-SegNet: Semantic segmentation of small agricultural targets in complex environments. **IEEE Access**, v. 11, p. 90214-90230, 2023. DOI: 10.1109/ACCESS.2023.3284036.

CONCLUSÃO GERAL

Com base nos estudos realizados, esta tese alcançou seu objetivo geral de aprimorar a geração de Modelos Digitais de Terreno (MDT) em áreas urbanas, demonstrando a eficácia da integração de deep learning ao fluxo fotogramétrico. A pesquisa foi estruturada em três capítulos que, de forma progressiva, abordaram desde a avaliação de modelos de elevação globais até a proposição de uma metodologia inovadora para a extração de MDTs a partir de dados de Veículos Aéreos Não Tripulados (VANTs).

O primeiro capítulo avaliou a acurácia de Modelos Digitais de Superfície (MDS) globais para a área urbana de Campo Grande, MS, utilizando um conjunto de pontos de controle preexistentes. Os resultados indicaram que os modelos TanDEM-X e ALOS AW3D30 são adequados para aplicações na escala 1:50.000 (classe A, PEC-PCD), enquanto os demais se mostraram compatíveis com a escala 1:100.000. Para futuros trabalhos com o objetivo de avaliar a acurácia de produtos cartográficos, recomenda-se a utilização de uma distribuição de pontos de controle homogênea em toda a área de estudo para garantir uma análise mais robusta.

Uma vez estabelecida a limitação de escala dos modelos globais, o segundo capítulo investigou a geração de produtos cartográficos de alta resolução com VANTs. O estudo analisou a influência da quantidade e da distribuição altimétrica de Pontos de Controle no Solo (GCPs) na acurácia do modelo. Verificou-se que a utilização de apenas dois GCPs resulta em erros altimétricos elevados, mas a partir de cinco GCPs a acurácia se estabiliza, obtendo-se o melhor resultado em altimetria. Crucialmente, a inclusão de GCPs em diferentes elevações melhora a acurácia altimétrica dos produtos gerados.

Acredita-se que a quantidade de GCPs necessários varia de acordo com o tamanho e as características da área, sendo um possível trabalho futuro investigar essa relação para determinar uma métrica de quantos GCPs por hectare (GCPs/ha) seriam ideais.

O terceiro capítulo apresentou a contribuição central da tese: uma abordagem inovadora para a geração direta de MDTs. Integrando a segmentação semântica de árvores por deep learning ao fluxo Structure from Motion e Multi-View Stereo

(SfM-MVS), foi possível gerar uma nuvem de pontos já filtrada, eliminando a etapa de pós-processamento. A arquitetura SegFormer demonstrou o melhor desempenho qualitativo e quantitativo na segmentação. Os MDTs resultantes apresentaram acurácia compatível com os gerados pelo filtro tradicional Cloth Simulation Filtering (CSF), com a vantagem de preservar melhor a superfície do terreno e evitar a criação de depressões artificiais, embora ainda apresente tendência a superestimar a altitude em alguns pontos.

Como trabalhos futuros, recomenda-se testar modelos pré-treinados, como o *Segment Anything Model* (SAM), para facilitar a criação da segmentação. Outra possibilidade é a verificação deste método para simulações hidrológicas, uma vez que se obteve o Modelo Digital sem Vegetação (MDnV), que é recomendado para essa aplicação.

Em síntese, a tese demonstrou uma trajetória metodológica coesa, partindo da caracterização de dados de acesso global e avançando para a otimização de levantamentos locais com VANTs, culminando em uma proposta inovadora que automatiza e refina a geração de MDTs urbanos. A integração de dados de VANT com segmentação semântica por deep learning constitui uma abordagem viável e escalável, com aplicações diretas em modelagem hidrológica, planejamento urbano e estudos ambientais.