

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE PESQUISAS HIDRÁULICAS**

Termo de Execução Descentralizada

**COOPERAÇÃO EM TECNOLOGIAS PARA ANÁLISES HIDROLÓGICAS EM
ESCALA NACIONAL**

entre

Agência Nacional das Águas e Saneamento Básico (ANA)

e

Instituto de Pesquisas Hidráulicas

Universidade Federal do Rio Grande do Sul (IPH/UFRGS)

Subprojeto

Modelo Digital do Terreno para América do Sul

Equipe:

Anderson Ruhoff

Bruno César Comini de Andrade

Leonardo Laipelt

Antonia Homrich da Silva

Frederico Carlos Reichert

Metodologia de estimativas de MDTs baseados em sensoriamento remoto

Resumo: Os modelos digitais de superfície (MDS) desempenham um papel crucial em diversas aplicações, incluindo estudos hidrológicos e hidrodinâmicos. Contudo, os principais conjuntos de dados globais de elevação disponíveis apresentam valores de elevações influenciados pela altura e cobertura da vegetação, o que compromete a acurácia dos modelos na representação da superfície do terreno. O presente relatório apresenta um levantamento das principais metodologias e informações utilizadas para a remoção do viés da vegetação em produtos globais de MDS. Além disso, é apresentada uma metodologia para a remoção do viés da vegetação no MDS Copernicus DEM GLO-30, utilizando *machine learning* e produtos de sensoriamento remoto. A validação foi realizada para área de alta influência da densidade do dossel, comparando com os principais modelos que apresentam a remoção do efeito da vegetação. Os resultados demonstraram boa capacidade da metodologia proposta na remoção do efeito da vegetação, obtendo erros condizentes com os demais produtos disponíveis. O método utilizado tem grande potencial para aplicação em escalas regionais e globais, permitindo a elaboração de um modelo de elevação acurado e de alta resolução espacial.

Palavras-chave: Modelo digital de elevação; GEDI; ICESat.

1. Introdução

Modelos digitais de superfície (MDS) são essenciais em diferentes aplicações, sendo inclusive utilizados em simulações hidrológicas e hidrodinâmicas (Liu et al., 2021; Baugh et al., 2013; Shurmann et al., 2008), análises de vulnerabilidade do terreno, como em casos de inundações e deslizamentos (Zhang et al., 2019; Kulp e Strauss, 2016), entre outros. A acurácia dos MDS é diretamente influenciada pela presença da vegetação na superfície, visto que grande parte dos sensores utilizados para a medição da elevação do terreno geram resultados com influência da altura e densidade do dossel (Carabajal e Harding, 2005; Liu et al., 2021; Mantelli et al., 2009; O’Loughlin et al., 2016; Schumann e Bates, 2018).

Nesse sentido, diferentes estudos buscaram soluções para a remoção do viés ocasionado pela vegetação (O’Loughlin et al., 2016; Baugh et al., 2013; Yamazaki et al., 2017). Baugh et al. (2013) constataram que a remoção de 50-60% da altura estimada da vegetação no modelo SRTM (*Shuttle Radar Topography Mission*) permite uma modelagem hidrodinâmica com maior precisão. Já O’Loughlin et al. (2016) utilizaram produtos de sensoriamento remoto, inclusive informações de densidade da vegetação e de altura da vegetação, e testaram diferentes equacionamentos para a obtenção do

melhor resultado de remoção do viés da vegetação no produto SRTM. Também utilizando os dados de elevação do SRTM, como também do AW3D (*ALOS: Advanced Land Observing Satellite, World 3D*) o modelo MERIT (*Multi-Error-Removed Improved-Terrain*) (Yamazaki et al., 2017) utilizou banco de dados globais de características da vegetação e informações do produto ICESat (*Ice, Cloud, and land Elevation Satellite*) para a remoção do viés da vegetação. Além disso, também propôs remover outras inconsistências identificadas nos MDS utilizados, como a existência de diferentes ruídos devido ao processo de aquisição dos dados.

Recentemente, metodologias de *machine learning* possibilitaram o uso de um número expressivo de informações para estimar o viés presente em MDS, como os ocasionados pela vegetação. Liu et al. (2021) utilizaram o modelo *Random Forest* (RF) e aumentaram a acurácia do produto MERIT em simulações de inundações em 18% em relação ao original. Já o FABDEM (*Forest and Buildings removed Copernicus DEM*) (Hawker et al., 2022), utilizando dados do produto obtidos pela missão Tandem-X, removeram o efeito da vegetação e erros ocasionados por perímetros urbanos através da aplicação do modelo RF. Nesse caso, as informações de elevação utilizadas foram obtidas através de sensores LIDAR (*Light Detection and Ranging*) do produto GEDI (*Global Ecosystem Dynamics Investigation*) e diferentes bases de dados referentes a densidade da população e iluminação noturna.

O presente relatório apresenta a revisão de metodologias utilizadas para a remoção do viés ocasionado pela vegetação em MDS, como também resultados preliminares da metodologia adotada.

Os itens apresentados no relatório incluem:

- revisão sobre os principais bancos de dados de altimetria utilizados para a remoção do efeito da vegetação.
- revisão das principais informações baseadas em sensoriamento remoto.
- revisão sobre as metodologias propostas pelos principais modelos digitais do terreno disponíveis.
- proposta de metodologia a ser utilizada para remoção do viés da vegetação.
- resultados preliminares.

2. Base de dados de altimetria do terreno.

a. ICESat

O satélite ICESat (*Ice, Cloud and Land Elevation Satellite*) com o sensor GLAS (*Geoscience Laser Altimeter System*), desenvolvido pela NASA (*National Aeronautics and Space Administration*), foi o primeiro a obter medições de altimetria a laser do tipo LIDAR (*Light detection and ranging*) de forma global (Schutz et al., 2005). Entre 2003 e 2009, o ICESat-1 obteve observações globais da superfície a partir de um *footprint* de aproximadamente 65 metros, para cada 172 metros, no traçado realizado pela sua órbita (600 km). Os dados da missão ICESat-1 visam auxiliar nos estudos relacionados a variações espaciais e temporais da espessura das camadas de gelos, contribuir para estudos em relação às estruturas de nuvens e aerossóis, além de permitir estimativas da vegetação e da superfície terrestre.

A continuidade da missão ICESat veio em 2018 com o lançamento do segundo satélite (ICESat-2) (Figura 2.1) e a utilização do sensor ATLAS (*Advanced Topographic Laser Altimeter System*) (Abdalati et al., 2010). Entre os avanços apresentados pelo sensor ATLAS, está o aprimoramento na obtenção das medições, e o aumento significativo no número de medições ao longo da órbita do satélite. Ainda, o *footprint* passou para aproximadamente 14 metros, enquanto o intervalo de medição foi reduzido para cada 90 metros.

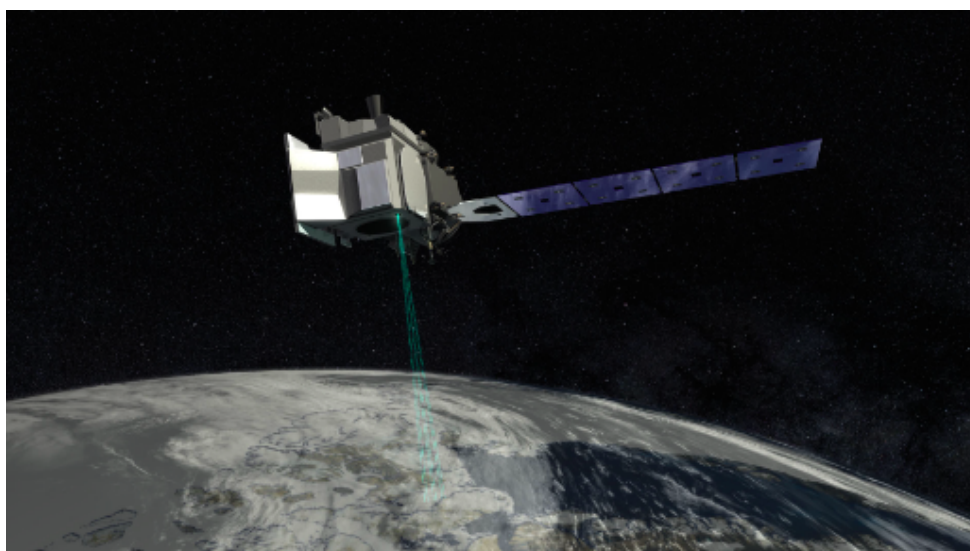


Figura 2.1 - ICESat-2 (*Ice, Cloud and land Elevation Satellite-2*) lançado em Setembro de 2018 pela NASA. Fonte: NASA.

b. GEDI

O instrumento GEDI (*Global Ecosystem Dynamics Investigation*) (Coyle et al., 2015), também de responsabilidade da NASA, possui um sistema LIDAR composto por 3 lasers que permitem a obtenção de dados em diferentes trajetórias. Cada medição possui um *footprint* de 25 m, intercalados entre 60 m ao longo da passagem. Assim como outros sensores LIDAR, o GEDI obtém informações da altura da vegetação, para diferentes patamares (Figura 2.2), de acordo com o retorno da energia refletida em diferentes níveis acima da superfície terrestre.

O GEDI esteve acoplado na estação espacial internacional (*International Space Station* - ISS) desde 2018 até início de 2023 (4 anos de aquisição de dados), encontrando-se atualmente em manutenção, e deve retornar novamente à atividade a partir de 2024. Seu desenvolvimento ocorreu com o propósito de auxiliar em estudos da vegetação e biodiversidade (Liu et al., 2021), com a disponibilização de dados precisos de altura da vegetação e de elevação terrestre, e permitir análises de processos do ciclo do carbono, como estimativas de biomassa (Albinet et al., 2019).

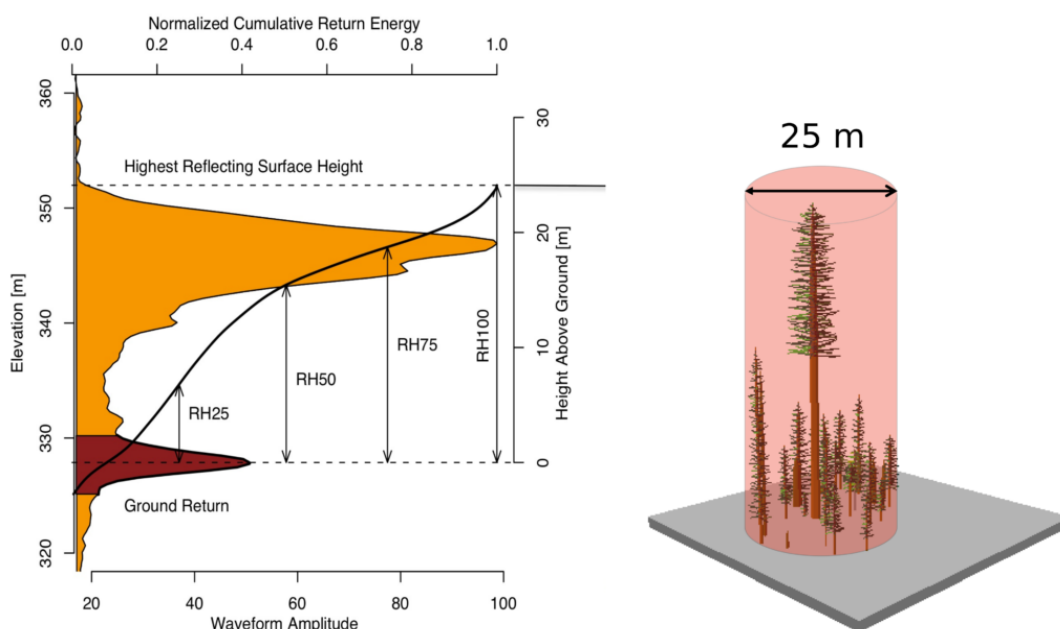


Figura 2.2 - Resultados obtidos para cada footprint (25 m) do GEDI. As informações de altura da vegetação e ponto de referência da altura do terreno são obtidas a partir do retorno da energia refletida em diferentes níveis.

3. Produtos de sensoriamento remoto

a. Índices

i. Simple Ratio (SR)

O índice Simple Ratio (SR) é a razão entre as bandas de reflectância no infravermelho próximo (NIR) e a banda vermelha (Red) (Equação 3.1). O SR é utilizado para indicar a saúde da vegetação, sendo valores próximos de 1 indicando uma vegetação saudável, enquanto que valores próximos de zero indicam áreas com ausência de vegetação.

$$SR = \frac{NIR}{RED} \quad (3.1)$$

ii. Modified Simple Ratio (MSR)

O *Modified Simple Ratio* (MSR) foi proposto por [Chen \(1996\)](#) com objetivo de representar os parâmetros biofísicos de florestas boreais através de dados de sensoriamento remoto, utilizando equacionamento similar ao apresentado pelo SR (Equação 3.2):

$$MSR = \frac{\left(\frac{NIR}{RED} - 1\right)}{\left(\sqrt{\frac{NIR}{RED}} + 1\right)} \quad (3.2)$$

iii. Enhanced Vegetation Index (EVI)

O *Enhanced Vegetation Index* (EVI) (Equação 3.3) é um índice com propósito similar ao *Normalized Difference Vegetation Index* (NDVI), sendo utilizado também para identificação da saúde da vegetação, porém corrigindo efeitos atmosféricos e ruídos da vegetação observados no NDVI. As correções são realizadas através de coeficientes que representam o fator de perda (G), de ajuste da resposta espectral do solo (L) e de correção dos aerossóis atmosféricos (C1 e C2). Os valores comumente utilizados para os coeficientes são 2,5 (G), 6 (C1), 7,5 (C2) e 1 (L).

$$EVI = G((NIR - R)/(NI + C1 * R - C2 * B + L)) \quad (3.3)$$

iv. Normalized Difference Moisture Index (NDMI)

O *Normalized Difference Moisture Index* (NDMI) é utilizado para a detecção de umidade na vegetação através da combinação das bandas espectrais do NIR e do infravermelho de onda curta (SWIR) (**Equação 3.4**). O NDMI é comumente aplicado para a identificação de estresse hídrico na vegetação, com valores próximos de 1 indicando uma boa condição hídrica, enquanto valores próximos de zero representam ambiente com estresse hídrico, como em períodos de estiagem.

$$NDMI = \frac{(NIR - SWIR)}{(NIR + SWIR)} \quad (3.4)$$

v. Modified Soil Adjusted Vegetation Index (MSAVI)

O *Modified Soil Adjusted Vegetation Index* (MSAVI) (**Equação 3.5**) é uma versão modificada do índice SAVI, eliminando a necessidade de obtenção do fator de correção do brilho do solo (L), necessário na aplicação do SAVI. O MSAVI é aplicado como alternativa ao NDVI na identificação da saúde da vegetação, principalmente em áreas em que a grande presença de solo superficial exposto.

$$MSAVI = \frac{2 * NIR - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)}}{2} \quad (3.5)$$

4. Principais metodologias utilizadas para o desenvolvimento de MDE

a. Modelos digitais de superfície (MDS)

i. SRTM

A missão SRTM (*Shuttle Radar Topography Mission*) ([Werner et al., 2001](#)) foi responsável pelo primeiro levantamento em escala global da superfície terrestre através de sensores a bordo da nave espacial Endeavour (**Figura 4.1**). O levantamento foi realizado em um período de 11 dias, a partir de fevereiro de 2000, utilizando o método de interferometria de radar, em que duas antenas alinhadas e distanciadas entre si coletaram dados no intervalo do espectro magnético da banda C, sendo posteriormente processados para a obtenção do modelo da superfície.

A resolução espacial inicialmente disponibilizada para o modelo global foi de aproximadamente 90 m, enquanto que uma versão revisada foi disponibilizada

posteriormente em 30 m através da reamostragem dos dados adquiridos, corrigindo também uma série de inconsistências verificadas na primeira versão.



Figura 4.1 - A missão SRTM realizou o primeiro levantamento em escala global da superfície terrestre, utilizando sensores acoplados na nave espacial Endeavour. Fonte: NASA.

ii. **NASADEM**

O modelo NASADEM (Crippen et al., 2016) é um reprocessamento dos dados adquiridos pela missão SRTM. Para isso, foram utilizados novos algoritmos de processamento, buscando a melhoria da acurácia e correções de falhas existentes na versão original do SRTM (Figura 4.2). Entre as correções realizadas, foram utilizados dados do modelo ASTER para preenchimento de valores ausentes no modelo original. Ainda, o modelo também realizou a reamostragem das informações obtidas e disponibilizou o produto final com resolução espacial de 30 m.

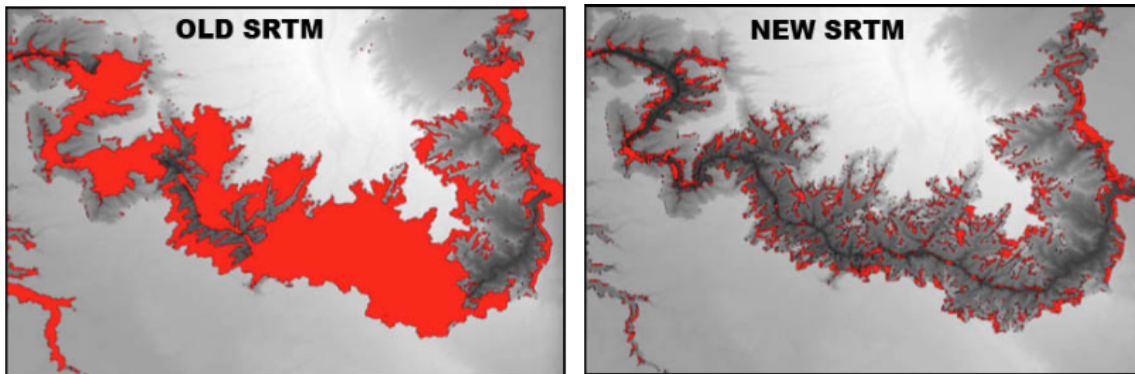


Figura 4.2 -Dados ausentes no produto original obtidos pela missão SRTM e no NASADEM. Fonte: [Crippen et al. \(2016\)](#)

iii. ALOS AW3D

A partir dos dados obtidos pelo satélite ALOS (*Advanced Land Observing Satellite*) foi elaborado o MDS AW3D (*ALOS World 3D*) (**Figura 4.3**). O ALOS, desenvolvido pela agência japonesa de exploração espacial (JAXA) utiliza o sensor pancromático PRISM (*Panchromatic Remote-sensing Instrument for Stereo Mapping*) para aquisição dos dados. Para a obtenção do modelo AW3D ([Tabono et al., 2015](#)), foi aplicado técnicas de estereoscopia com informações adquiridas entre 2006 e 2011 para diferentes direções (*backward - nadir - forward*) ao longo da passagem do satélite. A combinação das imagens adquiridas pelo satélite possibilita a geração de um produto tridimensional (latitude, longitude e elevação) de elevada acurácia, com resolução espacial de 30 m.

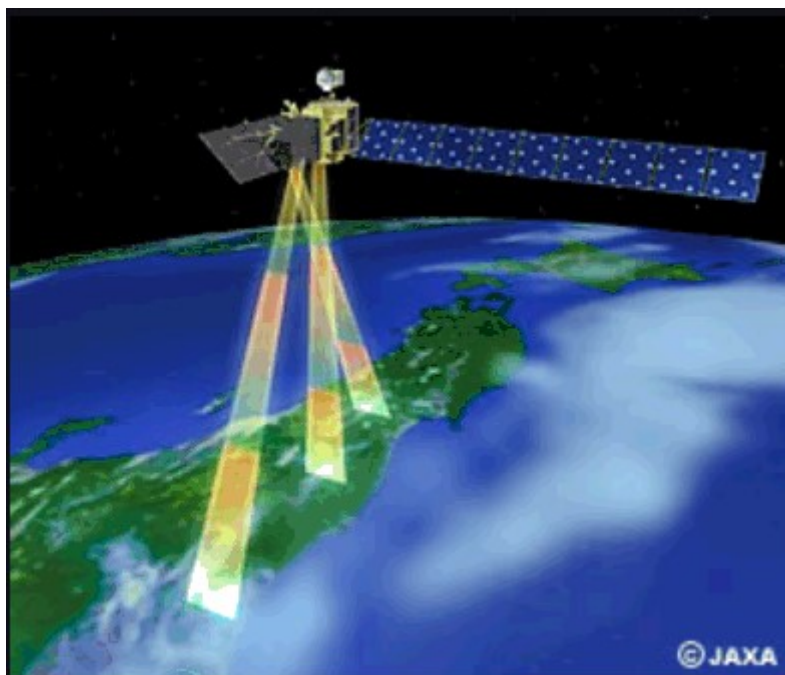


Figura 4.3 -Dados do satélite ALOS (*Advanced Land Observing Satellite*) são utilizados para a produção do modelo AW3D a partir de técnicas de estereoscopia. Fonte: JAXA.

iv. COPDEM

O MDS Copernicus GLO-30 (COPDEM) é uma versão derivada do MDS WorldDEM, o qual foi desenvolvido a partir de informações adquiridas da parceria entre a Airbus e a agência espacial europeia (ESA) através da missão Tandem-X, entre 2011 e 2015, utilizando técnica de interferometria. As diferenças do COPDEM em relação às informações disponibilizadas pelo WorldDEM ocorrem em processamentos adicionais de consistência dos valores de elevação de corpos hídricos, além de correções de elevações anômalas presentes no modelo WorldDEM.

A versão global e de acesso livre é disponibilizada pela ESA em resolução de 30 m (GLO-30), havendo também uma versão global de 90 m (GLO-90), e uma versão de 10 m restrita apenas para o continente europeu (**Figura 4.4**).

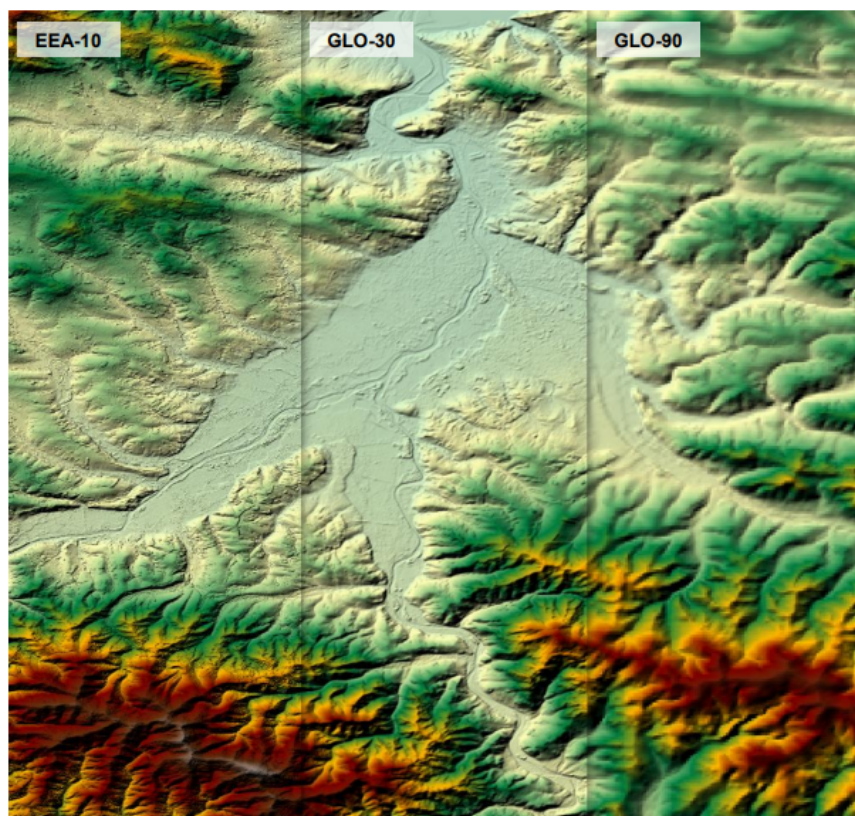


Figura 4.4 - Comparação entre as diferentes versões do COPDEM disponíveis. As versões globais são disponibilizadas com resoluções de 30 e 90 m (GLO-30 e GLO-90), sendo também disponibilizado uma versão restrita à europa, com 10 m de resolução (EEA-10). Fonte: ESA.

b. Remoção do viés da altura da vegetação

A remoção do viés ocasionado pela altura da vegetação em MDS é proposta em diversos estudos a partir do uso de banco de dados de altura da vegetação, subtraindo-a dos valores de elevação do MDS (O'Loughlin et al., 2016; Pinel et al., 2015; Baugh et al., 2013). Os MDS baseados em interferometria utilizam sensores de radar para a aquisição de dados, os quais emitem um pulso de radiação eletromagnética e medem o sinal refletido pela superfície. Em áreas de densa cobertura vegetal, o pulso emitido pelo sensor tende a penetrar apenas uma parcela no dossel, retornando uma elevação acima da altura do terreno, porém não condizente com a diferença total da altura da vegetação. Já no caso de produtos de MDS baseados em estereoscopia, por serem elaborados a partir de sensores passivos e imagens multiespectrais, o viés ocasionado pela vegetação tende a ser próximo ao da altura da vegetação (Hu et al., 2017; Yap. et al, 2018).

Assim, a modelagem do viés ocasionado pela vegetação pode ser realizada através da remoção de um percentual correspondente da altura da vegetação (Baugh et al., 2013; O’Loughlin et al., 2016). Baugh et al. (2013) demonstraram uma melhora na acurácia em simulações hidrológicas subtraindo do produto SRTM um percentual de 50% da altura correspondente da vegetação. Já O’Loughlin et al. (2016) utilizaram informações de altura e densidade da vegetação, além de informações de altimetria do ICESat, para estimar o percentual da vegetação a ser removido através de modelagem numérica.

Para a modelagem do viés ocasionados pela vegetação, estudos utilizaram, além da informação de altura da vegetação, dados de densidade percentual da vegetação como forma de compensação do efeito de penetração do sinal emitido pelo sensor. Com as informações obtidas por sensores LIDAR, é possível obter uma estimativa adequada referente a altura da vegetação (**Figura 2.2**), e por serem comumente medições pontuais, outras bases de dados de sensoriamento remoto são utilizadas para a espacialização da informação de altura da vegetação.

A remoção de inconsistências em produtos de MDS também é realizada através da aplicação de técnicas de *machine learning* (Kulp e Strauss et al., 2018; Hawker et al., 2022; Li et al., 2023; Yue et al., 2017), os quais permitem espacializar as informações de altimetria de produtos LIDAR através de diferentes modelos de regressão. Hawker et al. (2022) utilizaram o modelo RF para a estimativa do viés da vegetação a partir dos dados GEDI e informações de altura da vegetação do produto *Global Forest Canopy Height 2019*. Já Li et al., 2023 corrigiram o efeito da vegetação em três produtos de elevação (SRTM, OCPDEM e AW3D30) utilizando modelo de rede neural artificial, melhorando a acurácia em 55-78% no comparativo com as medições de altura do terreno do ICESat.

Na próxima seção, são apresentadas as metodologias para remoção do viés da vegetação dos principais modelos de elevação com correção do viés disponível.

c. Modelos digitais de elevação (terreno)

i. Bare-Earth

O modelo Bare-Earth foi apresentado por O’Loughlin et al. (2016) como uma proposta de modelo de elevação, removendo efeitos da cobertura da vegetação do

produto SRTM. Para a remoção do efeito da vegetação, os autores utilizaram uma abordagem envolvendo diferentes dados de sensoriamento remoto disponíveis.

Foram utilizados dados de altura da vegetação produzidos por [Simard et al. \(2011\)](#), a partir de medições da altura total da vegetação (RH100) do ICESat GLA. O RH100 é determinado no ICESat e em outros produtos LIDAR como a diferença entre o primeiro sinal de retorno (início do dossel) até o pico do sinal referente a elevação correspondente do solo. O dado de altura da vegetação foi especializado utilizando o método de regressão *Random Forest*, inserindo no modelo informações de temperatura, precipitação, elevação, e de cobertura vegetal do MODIS (*Moderate Resolution Imaging Spectroradiometer*).

Além disso, para representação da vegetação, foi utilizado também dados referentes à densidade do dossel, oriundos do MODIS (*Vegetation Continuous Field - VCF*). Para a regionalização da correção do efeito da vegetação, utilizou-se informações das classificações climáticas (Koppen-Geiger) e mapa de uso e cobertura da terra do produto MODIS *Land Cover Type* (MCD12Q1).

Para a estimativa do erro oriundo da vegetação no produto SRTM, os autores assumiram as seguintes hipóteses:

- A parcela da vegetação a ser subtraída do SRTM está relacionada com a densidade do dossel.
- O produto VCF representa com acurácia a densidade do dossel.
- As medições do ICESat GLA representam as elevações no solo.
- A parcela a ser removida do SRTM pode ser representada como **(Equação 4.1)**:

$$Veg_{remoção} = f(VCF) * H_{veg} \quad (4.1)$$

A remoção da vegetação no SRTM pode ser feita através da **Equação 4.2**:

$$SRTM_{bare\ earth} = SRTM - Veg_{remoção} \quad (4.2)$$

Para determinar o equacionamento da função $f(VCF)$, a **Equação 4.1** e **Equação 4.2** foram combinadas, substituindo os valores do $SRTM_{bare\ earth}$ pelas medições do ICESat (**Equação 4.3**):

$$f(VCF) = \frac{SRTM-ICESat}{H_{veg}} \quad (4.3)$$

Por fim, foram testadas cinco alternativas para representação de $f(VCF)$:

- (1) $f(VCF) = a * VCF + b$
- (2) $f(VCF) = a * VCF^2 + b * VCF + c$
- (3) $f(VCF) = a * VCF^b$
- (4) $f(VCF) = a * VCF^b + c$
- (5) $f(VCF) = a * exp^{b*VCF}$

Apesar de representarem cinco alternativas de funções para a remoção do viés da vegetação, os resultados apresentados por [O'Loughlin et al. \(2016\)](#) demonstraram que todas as alternativas de $f(VCF)$ resultaram em valores de VCF similares, para valores entre 20% e 80% de densidade da vegetação. Porém, apenas a função (3) apresentou curva com comportamento esperado, indicando valores próximos de zero em condições que a densidade relativa do dossel também se aproxima a zero. Além disso, a função (3) foi a que apresentou os melhores resultados na performance estatística, com RMSE (*Root mean squared error*) de 6,11 m, seguido da função (1) (6,17 m) e função (3) (6,20m) (**Figura 4.5**)

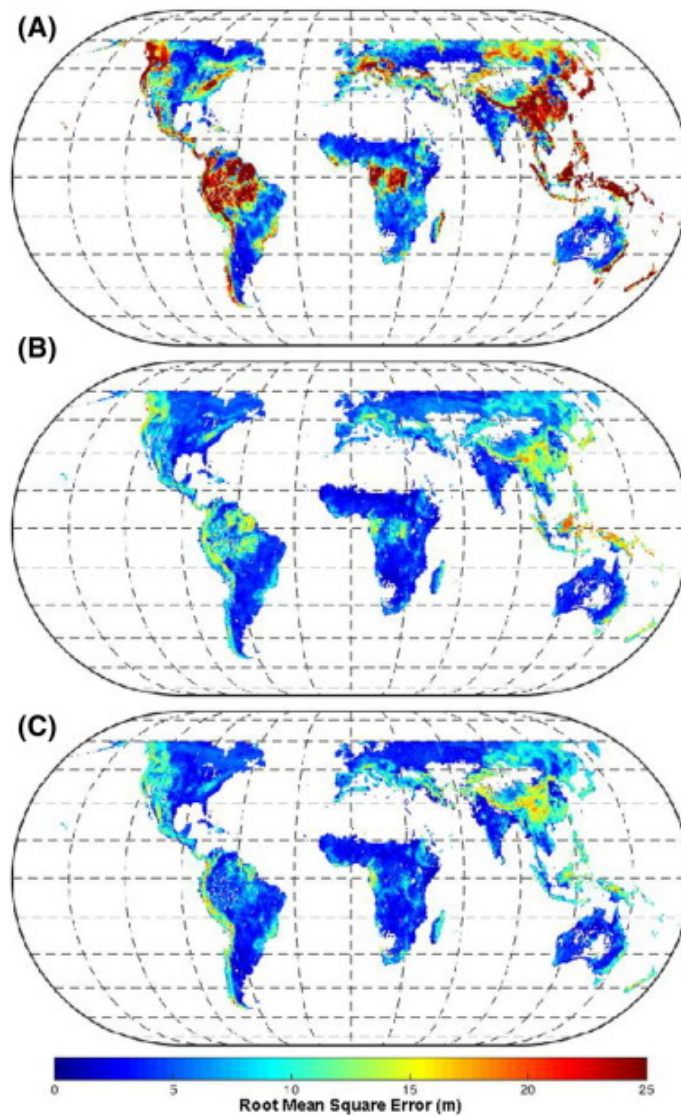


Figura 4.5 - Mapa do resultado de raiz do erro médio quadrático (RMSE) em áreas vegetadas para o (a) SRTM, (b) SRTM com correção e (c) Bare Earth aplicando a correção com a função (3). Fonte: [O'Loughlin et al. \(2016\)](#).

ii. MERIT DEM

Como uma proposta de modelo global de elevação de terreno de alta acurácia, o MERIT DEM (*Multi-Error-Removed Improved-Terrain DEM*) ([Yamazaki et al., 2017](#)) apresentou uma série de correções para a remoção do viés ocasionado pela vegetação, como também erros oriundos de ruídos na aquisição dos dados (*stripe and speckle noises*), e outros existentes nos produtos SRTM e AW3D.

Para a elaboração do MERIT DEM, utilizou-se, além do SRTM e AW3D, informações do VFP-DEM (*Viewfinder Panoramas DEM*) para o preenchimento de falhas em áreas em que os demais modelos não possuíam informações. Além disso,

dados de altimetria do ICESat foram utilizados como referencial da elevação correspondente do terreno para estimativa do erro atribuído pela vegetação.

A remoção do erro, de forma geral, foi realizada em quatro etapas (**Figura 4.6**): *speckle noise*; *stripe noise*; erros absolutos e viés da altura da vegetação. O erro associado ao efeito de *speckle noise* ocorre devido à variabilidade da superfície em áreas planas, influenciando aleatoriamente nos valores de um número reduzido de pixels e em comprimentos de ondas muito curtos. Já os erros ocasionados pelo efeito de *stripe noise* no produto SRTM ocorreram, principalmente, devido à inconsistências na aplicação da interferometria, enquanto no AW3D os erros de *stripe noise* estão associados com a inclinação das elevações. Os erros absolutos de elevação são identificados como mudanças repentinas de elevações dentro de um área de domínio (> 20 km), os quais podem ser observados em ambos os modelos de interferometria e estereometria. Por fim, o viés ocasionado nas elevações devido à vegetação ocorre em ambos os modelos (SRTM e AW3D), visto que ambas as metodologias não são capazes de obter informações da elevação do solo abaixo do dossel.

A ordem de remoção dos erros foi definida como forma de evitar interferências em cada aplicação. O primeiro erro a ser removido dos modelos SRTM e AW3D foi o ocasionado pelo efeito de *stripe noise*, aplicando um filtro 2D de transformada de Fourier. O erro seguinte removido foi o de erro absoluto, a partir do cálculo da diferença entre as elevações apresentadas pelos modelos e as elevações do ICESat. A remoção do viés da vegetação foi posteriormente estimada como uma função da densidade da vegetação e da altura da vegetação utilizando dados ICESat, banco de dados de cobertura global de florestas e comparando os resultados obtidos para ambos os modelos utilizados. Na última etapa, o efeito de *speckle noise* foi removido aplicando um filtro de suavização (*smoothing filter*). Por fim, foi realizada a junção entre o produto resultante da aplicação das quatro correções no SRTM e AW3D para obtenção do resultado final do MERIT DEM.

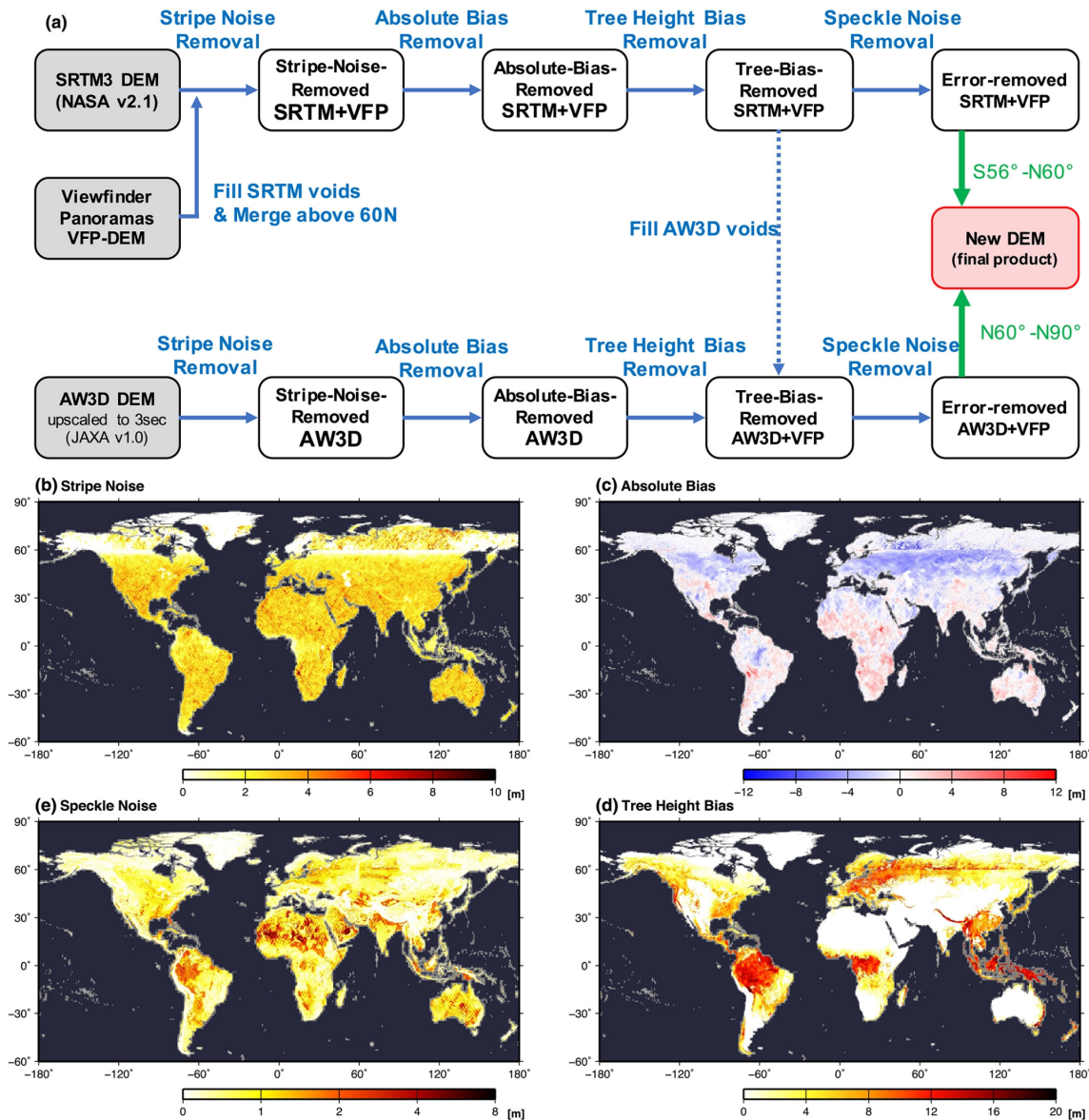


Figura 4.6 - Esquemática das etapas aplicadas para a remoção dos erros presentes no SRTM e no AW3D (Yamazaki et al. 2017).

iii. FABDEM

Utilizando técnicas de *machine learning* para a remoção do viés da vegetação e de perímetros urbanos, o FABDEM (*Forest And Buildings removed Copernicus DEM*) (Hawker et al., 2022) utiliza dados de elevação provenientes do produto Copernicus GLO-30 (COPDEM) (Figura 4.7). A estimativa do viés ocasionado pela vegetação e por construções em perímetros urbanos foi realizada através da aplicação do modelo RF, utilizando diferentes dados de sensoriamento remoto. Foram utilizados dois modelos de RF: um para estimar o viés da vegetação, e outro para estimar o viés de áreas urbanas. Em ambos os casos, os dados de altimetria LIDAR do produto GEDI

foram utilizados para determinar a altura referente ao solo. Além disso, foram utilizados dados de altura da vegetação do produto global *Forest Canopy Height* (2019) (Popatov et al., 2020), o qual foi gerado a partir de imagens Landsat e dados do GEDI.

A remoção do efeito da floresta ocorreu através da aplicação do RF para estimar a diferença entre a elevação do COPDEM com os dados de elevação no nível do solo. As correções foram realizadas para os pixels em que a altura da vegetação apresentava valores maiores do que 3 m e percentual de cobertura da vegetação maior do que 10%. Ainda, foi utilizado ferramentas disponíveis do software *WhiteboxTools* para aplicações de filtros em uma janela 3x3 para redução de ruído, assim como para proporcionar maior nitidez.

Para a remoção do viés de edificações, foi aplicado o RF utilizando informações das diferenças entre o COPDEM e dados de elevação do terreno do GEDI para áreas urbanas. O modelo RF também foi desenvolvido utilizando informações de densidade da população (WorldPop), luzes noturnas, *footprint* de edificações (*World Settlement Footprint*) e informações socioeconômicas, como o percentual de áreas construídas per capita.

Após a correção do viés da vegetação e de áreas urbanas, foi realizado um pós-processamento para reduzir valores inconsistentes, ocasionados principalmente por excesso de ruídos na aplicação do RF. O preenchimento de depressões foi realizado apenas em pixels modificados pela correção dos viés. Já filtros para a redução do ruído foram aplicados buscando a remoção de efeitos de *speckle noise* e também para a preservação de bordas, implementados com o auxílio da ferramenta *WhiteBoxTools*.

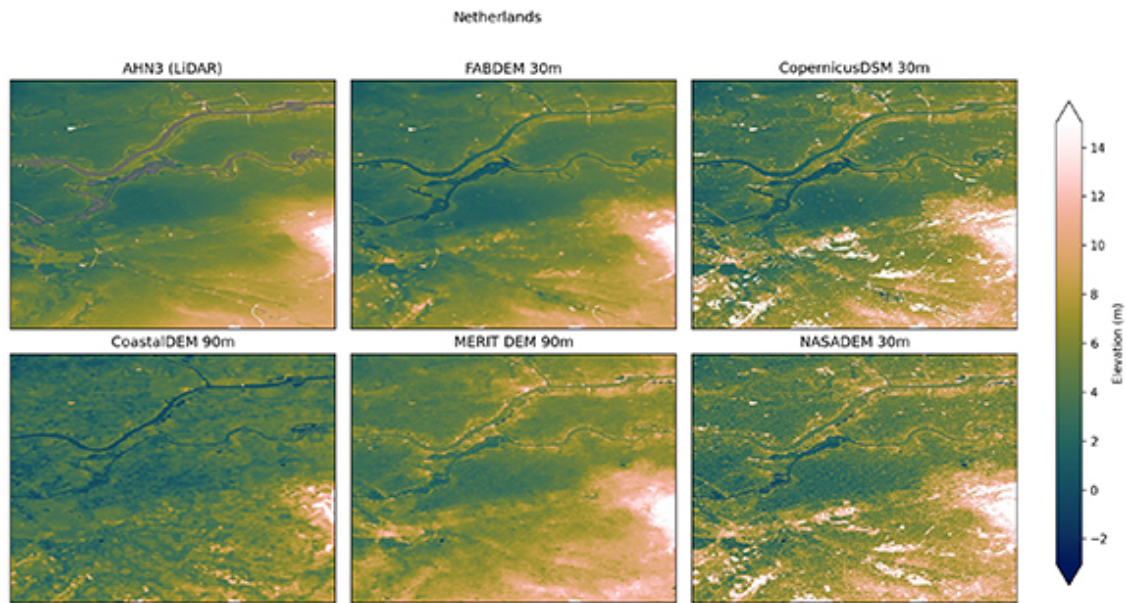


Figura 4.7 - Comparação entre diferentes modelos digitais de elevação e resultados obtidos com o modelo FABDEM.

5. Metodologia preliminar para remoção do viés da vegetação

a. Dados de sensoriamento remoto

Para a predição do erro associado à vegetação no produto COPDEM foi utilizado diferentes bancos de dados de sensoriamento remoto. Imagens multiespectrais dos satélites Landsat 8 OLI/TIRS e do Sentinel 2 MSI foram obtidas para o cálculo de índices que evidenciam as respostas espectrais entre áreas com intensa densidade do dossel e solo exposto. Os índices inseridos no modelo foram selecionados após revisão bibliográfica e análises preliminares, apresentando uma melhor representação na discretização em relação ao viés da vegetação em produtos de elevação. Além disso, foi aplicado um filtro interquartil nos índices selecionados, visando a remoção de ruídos. A **Tabela 5.1** apresenta os índices utilizados e as respectivas entradas.

Tabela 5.1 - Índices utilizados pelo modelo para estimar o viés da altura da vegetação no produto COPDEM.

Índice	Equação	Dados
EVI	$G * ((NIR - R)/(NI + C1 * R - C2 * B + L))$	Landsat
NDMI	$\frac{(NIR - SWIR)}{(NIR + SWIR)}$	Landsat/Sentinel
MSR	$\frac{(\frac{NIR}{RED} - 1)}{(\sqrt{\frac{NIR}{RED} + 1})}$	Landsat
MSAVI	$\frac{2 * NIR - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)}}{2}$	Landsat

A informação de altimetria e de altura da vegetação foi obtida a partir dos dados disponíveis do sensor GEDI (*Global Ecosystem Dynamics Investigation*) para a coleção L2A (https://doi.org/10.5067/GEDI/GEDI02_A.002). A estimativa do viés gerado pela influência da vegetação foi realizada utilizando a informação de altura relativa da vegetação (RH) referente ao percentual de 80%.

b. Remoção do viés da vegetação e pós-processamento

Para a remoção do viés da vegetação, foi utilizado o algoritmo de decisão de *TreeBoost*, modelo disponibilizado na plataforma *Google Earth Engine* (Gorelick et al., 2017), e baseado no modelo de mesmo nome disponível na plataforma *open source* de *machine learning* SMILE (*Statistical Machine Intelligence and Learning Engine*). A aplicação do modelo *TreeBoost* ocorreu para cada cena definida para a área do estudo de caso, utilizando 50 mil amostras, e 250 árvores de decisão.

Após a obtenção do viés ocasionado pela vegetação ($BIAS_{vegetação}$), foi realizada uma etapa de correção do viés de acordo com a fração de cobertura vegetal (fvc), assumindo que o erro no MDS de referência ($MDS_{copernicus}$) é proporcional à fvc . Dessa forma, o modelo foi obtido através da **Equação 5.1**:

$$MDE = MDS_{\text{copernicus}} - (BIAS_{\text{vegetação}} * fvc) \quad (5.1)$$

c. Validação

Para a validação do modelo sugerido foram utilizadas medições de altimetria do produto ICESat-2, comparando informações de elevação disponíveis referentes ao terreno ('*h_te_best_fit*') com as elevações do produto gerado, nomeado provisoriamente de IPHDEM, com outros modelos que também apresentam a remoção do efeito da vegetação (FABDEM e MERIT), e o COPDEM por ser a informação inicial de elevação utilizada. Foi realizada uma amostragem aleatória de 50 mil pontos ICESat, extraíndo os valores correspondentes do pixel de cada produto de elevação, para a região do estado de Rondônia, definida como estudo de caso.

A performance estatística dos produtos de elevação foi obtida através da análise do *Root Mean Square Error* (RMSE) (**Equação 5.2**), *Mean Bias Error* (MBE) (**Equação 5.3**) e o coeficiente de determinação (r^2) (**Equação 5.4**).

$$RMSE = \sqrt{\frac{\sum(O_i - P_i)^2}{n}} \quad (5.2)$$

$$MBE = \frac{\sum(P_i - O_i)}{n} \quad (5.3)$$

$$r^2 = \left(\frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\left[\sum_{i=1}^n (O_i - \bar{O}) \right] \left[\sum_{i=1}^n (P_i - \bar{P}) \right]}} \right)^2 \quad (5.4)$$

6. Resultados preliminares

Os resultados da validação com dados ICESat são apresentados na **Tabela 5.2**. De forma geral, o modelo IPHDEM apresentou RMSE de 6,55 m, demonstrando resultados similares ao apresentado pelo modelo FABDEM (6,45 m), e menor erro em relação ao observado com o modelo MERIT (7,18 m). A diferença observada entre os modelos pode ser compreendida através dos resultados obtidos de viés entre os modelos e as elevações do ICESat. Enquanto o modelo MERIT apresentou uma

superestimativa de 4,03 m, o IPHDEM e o FABDEM apresentaram uma superestimativa de 1,4 m e 0,22 m, respectivamente. Já o coeficiente de determinação foi similar (0,991 - 0,992) para todos os modelos com o viés da vegetação corrigida. Por fim, a performance estatística apresentada pelo COPDEM representa a influência direta nos valores de elevação devido a altura da vegetação no modelo, apresentando RMSE e MBE de 15,66 e 13,2 m, respectivamente. A **Figura 5.1** apresenta o gráfico de dispersão das amostras utilizadas para o cálculo da performance estatística dos modelos em relação aos dados ICESat, enquanto a **Figura 5.2** apresenta o erro em relação aos dados ICESat e os modelos analisados.

Tabela 5.2 - Resultados preliminares da performance estatística dos modelos obtidos no comparativo com medições ICESat.

Modelo	RMSE (m)	MBE (m)	r²
COPDEM	15,66	13,20	0,984
IPHDEM	6,55	1,40	0,991
MERIT	7,18	4,03	0,992
FABDEM	6,45	0,22	0,991

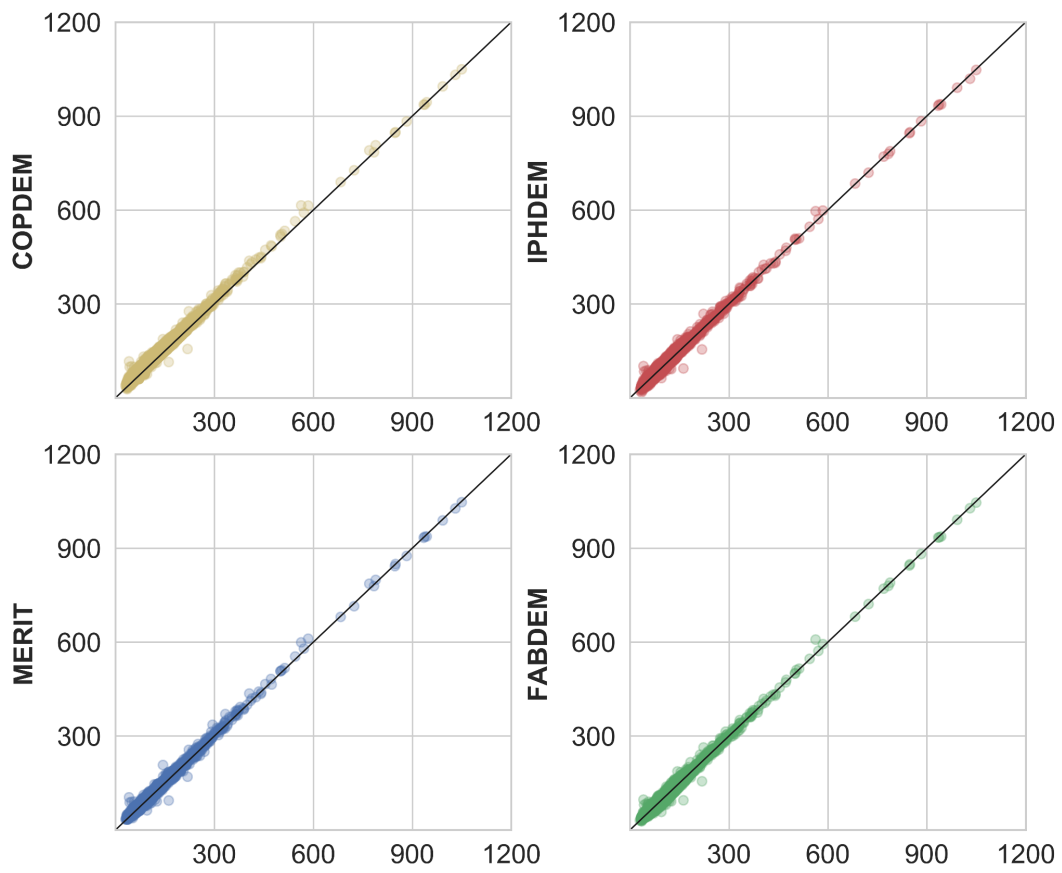


Figura 5.1 - Gráfico de dispersão para os quatro modelos avaliados em relação aos dados de elevação do ICESat. (a) COPDEM, (b) IPHDEM, (c) MERIT e (d) FABDEM.

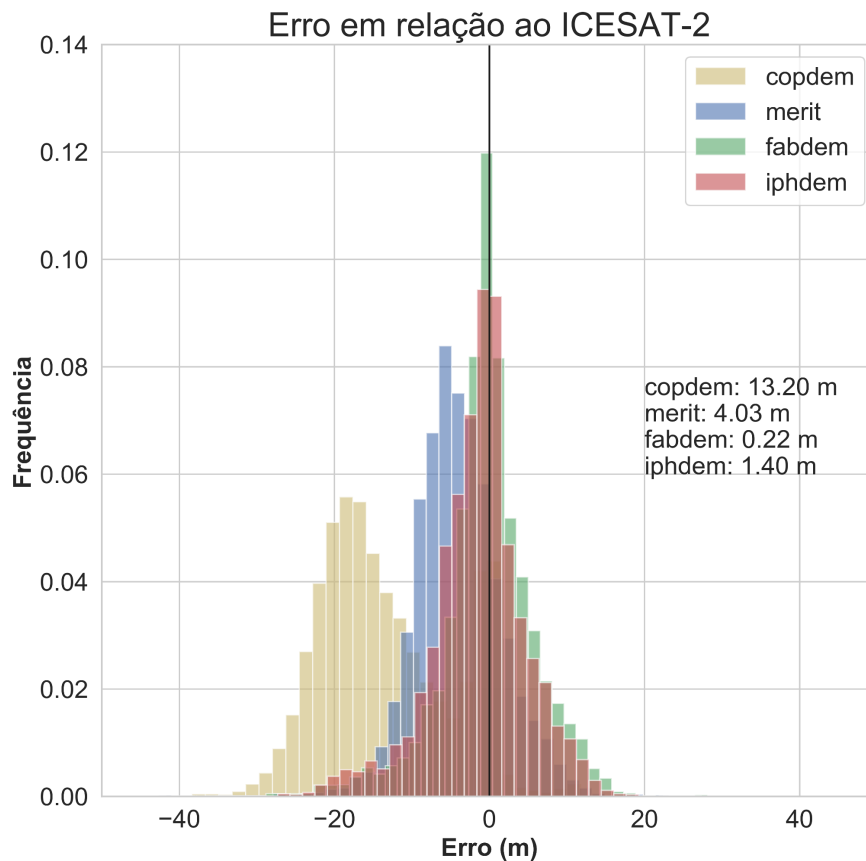


Figura 5.2 - Erro dos modelos (COPDEM, MERIT, IPHDEM e FABDEM) em relação aos dados de elevação do ICESat.

A comparação visual dos resultados obtidos para o IPHDEM e para os outros três modelos globais são apresentados na **Figura 5.3**. o comparativo foi realizado na região do estado de Rondônia, em trecho do Rio Madeira. Como é mostrado em **Figura 5.3a**, a região possui intensa mudança de uso e cobertura da terra, apresentando áreas de pastagens e de florestas. A remoção do viés do efeito da vegetação fica evidente para os três modelos apresentados em **Figura 5.3** (b), (c) e (d): IPHDEM, MERIT e FABDEM, respectivamente. Ainda é possível observar resquícios de feições de áreas de florestas nos três modelos, menos perceptível no FABDEM devido a aplicação de diferentes filtros para a suavização dos resultados de elevações (Hawker et al., 2022). A versão inicial do modelo proposto para a remoção do viés da vegetação (IPHDEM) apresenta menor suavização dos valores de elevação. Apenas um filtro “low-pass”, utilizando uma janela 5x5, foi aplicado no IPHDEM para suavização de ruídos ocasionados pela aplicação do modelo *TreeBoost*.

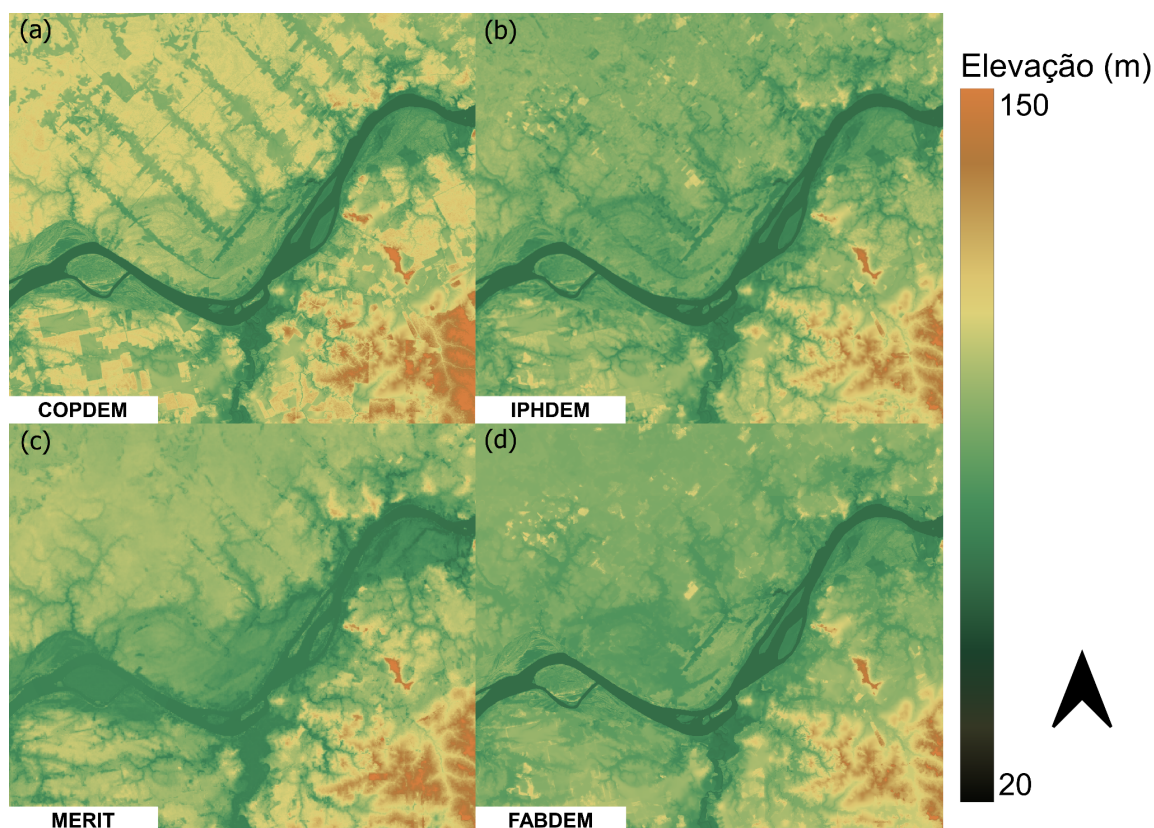


Figura 5.3 - Comparativo visual de elevação para o MDS COPDEM (a) em relação aos outros modelos de elevação com correção do viés da vegetação: (b) IPHDEM; (c) MERIT e (d) FABDEM.

7. Próximos passos

De forma geral, os resultados da validação realizada através do uso de dados ICESat estão de acordo com estudos anteriores da correção do efeito da vegetação em modelos globais de elevação (Hawker et al. 2021; Li et al., 2023; Pinel et al., 2015), e diferenças na validação podem estar associadas a variações entre os dados LIDAR utilizados para a calibração do modelo (GEDI) e para a validação (ICESat). Estudos mostram erros distintos em relação às elevações de terreno obtidas pelo ICESat e GEDI (Liu et al., 2021; Musthafa et al., 2023), indicando a existência de incertezas em aplicações como a proposta neste estudo. Além disso, os resultados apresentados para o modelo proposto são preliminares e concentrados para uma área de aplicação (Rondônia), e novas calibrações e etapas de pós-processamentos serão aprimorados em busca de

melhoria na performance estatística, assim como também na validação para diferentes localidades e uso e cobertura da terra.

8. Referências bibliográficas

- Abdalati, W., Zwally, H.J., Bindschadler, R., Csatho, B., Farrell, S.L., Fricker, H.A., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., Webb, C., 2010. The ICESat-2 Laser Altimetry Mission. *Proc. IEEE* 98, 735–751.
<https://doi.org/10.1109/JPROC.2009.2034765>
- Albinet, C., Whitehurst, A.S., Jewell, L.A., Bugbee, K., Laur, H., Murphy, K.J., Frommknecht, B., Scipal, K., Costa, G., Jai, B., Ramachandran, R., Lavalley, M., Duncanson, L., 2019. A Joint ESA-NASA Multi-mission Algorithm and Analysis Platform (MAAP) for Biomass, NISAR, and GEDI. *Surv. Geophys.* 40, 1017–1027. <https://doi.org/10.1007/s10712-019-09541-z>
- Baugh, C.A., Bates, P.D., Schumann, G., Trigg, M.A., 2013. SRTM vegetation removal and hydrodynamic modeling accuracy. *Water Resour. Res.* 49, 5276–5289. <https://doi.org/https://doi.org/10.1002/wrcr.20412>
- Coyle, D.B., Stysley, P.R., Poullos, D., Clarke, G.B., Kay, R.B., 2015. Laser transmitter development for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar, in: *Proc.SPIE.* p. 961208.
<https://doi.org/10.1117/12.2191569>
- Crippen, R., Buckley, S., Agram, P., Belz, E., Gurrola, E., Hensley, S., Kobrick, M., Lavalley, M., Martin, J., Neumann, M., Nguyen, Q., Rosen, P., Shimada, J., Simard, M., Tung, W., 2016. NASADEM GLOBAL ELEVATION MODEL: METHODS AND PROGRESS. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLI-B4, 125–128.
<https://doi.org/10.5194/isprs-archives-XLI-B4-125-2016>

- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., Neal, J., 2022. A 30 m global map of elevation with forests and buildings removed. *Environ. Res. Lett.* 17, 24016. <https://doi.org/10.1088/1748-9326/ac4d4f>
- Hu, Z., Peng, J., Hou, Y., Shan, J., 2017. Evaluation of Recently Released Open Global Digital Elevation Models of Hubei, China. *Remote Sens.* <https://doi.org/10.3390/rs9030262>
- Kulp, S., Strauss, B.H., 2016. Global DEM Errors Underpredict Coastal Vulnerability to Sea Level Rise and Flooding . *Front. Earth Sci.* .
- Kulp, S.A., Strauss, B.H., 2018. CoastalDEM: A global coastal digital elevation model improved from SRTM using a neural network. *Remote Sens. Environ.* 206, 231–239. <https://doi.org/https://doi.org/10.1016/j.rse.2017.12.026>
- Li, Y., Li, L., Chen, C., Liu, Y., 2023. Correction of global digital elevation models in forested areas using an artificial neural network-based method with the consideration of spatial autocorrelation. *Int. J. Digit. Earth* 16, 1568–1588. <https://doi.org/10.1080/17538947.2023.2203953>
- Liu, A., Cheng, X., Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sens. Environ.* 264, 112571. <https://doi.org/https://doi.org/10.1016/j.rse.2021.112571>
- Liu, Y., Bates, P.D., Neal, J.C., Yamazaki, D., 2021. Bare-Earth DEM Generation in Urban Areas for Flood Inundation Simulation Using Global Digital Elevation Models. *Water Resour. Res.* 57, e2020WR028516. <https://doi.org/https://doi.org/10.1029/2020WR028516>
- Musthafa, M., Singh, G., Kumar, P., 2022. Comparison of forest stand height interpolation of GEDI and ICESat-2 LiDAR measurements over tropical

- and sub-tropical forests in India. *Environ. Monit. Assess.* 195, 71.
<https://doi.org/10.1007/s10661-022-10657-w>
- O'Loughlin, F.E., Paiva, R.C.D., Durand, M., Alsdorf, D.E., Bates, P.D., 2016. A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sens. Environ.* 182, 49–59.
<https://doi.org/https://doi.org/10.1016/j.rse.2016.04.018>
- Pinel, S., Bonnet, M.-P., Santos Da Silva, J., Moreira, D., Calmant, S., Satgé, F., Seyler, F., 2015. Correction of Interferometric and Vegetation Biases in the SRTMGL1 Spaceborne DEM with Hydrological Conditioning towards Improved Hydrodynamics Modeling in the Amazon Basin. *Remote Sens.*
<https://doi.org/10.3390/rs71215822>
- Schumann, G., Matgen, P., Cutler, M.E.J., Black, A., Hoffmann, L., Pfister, L., 2008. Comparison of remotely sensed water stages from LiDAR, topographic contours and SRTM. *ISPRS J. Photogramm. Remote Sens.* 63, 283–296. <https://doi.org/https://doi.org/10.1016/j.isprsjprs.2007.09.004>
- Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., DiMarzio, J.P., 2005. Overview of the ICESat Mission. *Geophys. Res. Lett.* 32.
<https://doi.org/https://doi.org/10.1029/2005GL024009>
- Tadono, T., Takaku, J., Tsutsui, K., Oda, F., Nagai, H., 2015. Status of “ALOS World 3D (AW3D)” global DSM generation, in: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 3822–3825.
<https://doi.org/10.1109/IGARSS.2015.7326657>
- Werner Marian, 2001. Shuttle Radar Topography Mission (SRTM) Mission Overview. *Frequenz* 55, 75–79.
<https://doi.org/doi:10.1515/FREQ.2001.55.3-4.75>
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations. *Geophys. Res. Lett.* 44, 5844–5853.
<https://doi.org/https://doi.org/10.1002/2017GL072874>

- Yap, L., Kandé, L.H., Nouayou, R., Kamguia, J., Ngouh, N.A., Makuate, M.B., 2019. Vertical accuracy evaluation of freely available latest high-resolution (30 m) global digital elevation models over Cameroon (Central Africa) with GPS/leveling ground control points. *Int. J. Digit. Earth* 12, 500–524.
<https://doi.org/10.1080/17538947.2018.1458163>
- Yue, L., Shen, H., Zhang, L., Zheng, X., Zhang, F., Yuan, Q., 2017. High-quality seamless DEM generation blending SRTM-1, ASTER GDEM v2 and ICESat/GLAS observations. *ISPRS J. Photogramm. Remote Sens.* 123, 20–34.
<https://doi.org/https://doi.org/10.1016/j.isprsjprs.2016.11.002>
- Zhang, K., Gann, D., Ross, M., Robertson, Q., Sarmiento, J., Santana, S., Rhome, J., Fritz, C., 2019. Accuracy assessment of ASTER, SRTM, ALOS, and TDX DEMs for Hispaniola and implications for mapping vulnerability to coastal flooding. *Remote Sens. Environ.* 225, 290–306.
<https://doi.org/https://doi.org/10.1016/j.rse.2019.02.028>