

**LAURA COELHO DE ANDRADE**

**AVALIAÇÃO DE SENSORES ORBITAIS NA DETERMINAÇÃO DA BATIMETRIA  
DERIVADA DE SATÉLITE: ABORDAGEM COM MÉTODOS EMPÍRICOS**

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

Orientador: Italo Oliveira Ferreira

**VIÇOSA - MINAS GERAIS  
2023**

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

T

A553a  
2023  
Andrade, Laura Coelho de, 1998-  
Avaliação do sensores orbitais na determinação da  
batimetria derivada de satélite: abordagem com métodos  
empíricos / Laura Coelho de Andrade. – Viçosa, MG, 2023.  
1 dissertação eletrônica (82 f.): il. (algumas color.).

Inclui apêndices.

Orientador: Ítalo Oliveira Ferreira.

Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Engenharia Civil, 2023.

Referências bibliográficas: f. 46-51.

DOI: <https://doi.org/10.47328/ufvbbt.2023.198>

Modo de acesso: World Wide Web.

1. Mapeamento multifeixe. 2. Mapas batimétricos.  
3. Sensoriamento remoto. 4. Aprendizado do computador.  
I. Ferreira, Ítalo Oliveira, 1988-. II. Universidade Federal de  
Viçosa. Departamento de Engenharia Civil. Programa de  
Pós-Graduação em Engenharia Civil. III. Título.

CDD 22. ed. 551.46


**LAURA COELHO DE ANDRADE**

**AVALIAÇÃO DE SENSORES ORBITAIS NA DETERMINAÇÃO DA BATIMETRIA  
DERIVADA DE SATÉLITE: ABORDAGEM COM MÉTODOS EMPÍRICOS**

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


APROVADA: 13 de fevereiro de 2023

Assentimento:

Documento assinado digitalmente  
 LAURA COELHO DE ANDRADE  
Data: 24/04/2023 16:21:37-0300  
Verifique em <https://validar.iti.gov.br>

---

Laura Coelho de Andrade  
Autora

Documento assinado digitalmente  
 ITALO OLIVEIRA FERREIRA  
Data: 25/04/2023 16:40:03-0300  
Verifique em <https://validar.iti.gov.br>

---

Italo Oliveira Ferreira  
Orientador

## AGRADECIMENTOS

Agradeço primeiramente a Deus, pela possibilidade de viver esse sonho, ter resiliência, persistência e coragem para seguir em frente;

Aos meus pais, Maria Lúcia e Cláudio, e ao meu irmão Guilherme por todo amor, incentivo e por me proporcionarem todos os meios para que eu chegasse até aqui;

Ao Igor que além de ser meu ponto de paz, equilíbrio e inspiração, sempre foi meu companheiro e nunca deixou de me dar forças nos dias mais difíceis;

Ao professor Italo Oliveira Ferreira, pela orientação, disponibilidade e apoio prestados ao longo da graduação e do mestrado, além de ter me apresentado essa área de pesquisa que atualmente sou apaixonada, sempre me motivou e confiou em mim. Sem o seu suporte não teria chegado até aqui, muito obrigada!

Ao GEPLH por ser minha segunda família na UFV e me ensinar tanto durante esses cinco anos. Em especial à Victória que esteve comigo desde os primeiros trabalhos e nunca deixou de me apoiar. Ao Lucas e a Larissa pelo auxílio com as cartas náuticas e edição dos dados;

À professora Maria Lúcia Calijuri por ser exemplo de responsabilidade, integridade e coragem para enfrentar novos desafios. Ao Arthur pelas contribuições e tempo disponibilizado para o auxílio no desenvolvimento desse trabalho e ao SigeoNPA pelos aprendizados nos projetos de extensão, que contribuíram de forma direta e indireta para minha formação, além dos momentos diários de descontração durante as fases mais difíceis do mestrado.

À professora Nilcilene pela disposição e assistência com as ideias e discussões necessárias para a concretização dessa pesquisa;

Ao professor Júlio, pela paciência e pelos ensinamentos nesse período de treinamento;

À todos os professores e funcionários da Universidade Federal de Viçosa e, principalmente do Departamento de Engenharia Civil;

À Marinha do Brasil por disponibilizar os dados e permitir o progresso do trabalho;

À CAPES - Coordenadoria de Aperfeiçoamento de Pessoal de Ensino Superior, pela concessão de bolsa de auxílio aos estudos – Código de Financiamento 001.

A todos que não foram citados, nominalmente, mas que contribuíram para que o meu objetivo fosse alcançado.

“You can never cross the ocean until you have the courage to lose sight of the shore”.

(Colombus)

“Victory belongs to the most tenacious”  
(Roland Garros)

## RESUMO

ANDRADE, Laura Coelho de, M.Sc., Universidade Federal de Viçosa, fevereiro de 2023. **Avaliação de sensores orbitais na determinação da batimetria derivada de satélite: abordagem com métodos empíricos.** Orientador: Italo Oliveira Ferreira.

No âmbito hidrográfico, sabe-se que a principal técnica empregada para o levantamento batimétrico de um corpo hídrico consiste no sensoriamento remoto acústico, com o emprego de ecobatímetros, uma vez que esses fornecem informações acuradas mesmo em grandes profundidades. A aquisição de dados batimétricos, com o foco em atualizar e editar cartas náuticas, é o principal objetivo de um Levantamento Hidrográfico. No entanto, o procedimento para aquisição de dados empregando métodos acústicos de sondagem são de alto custo, além de demandarem um longo período de execução. Nesse sentido, a utilização de técnicas para obtenção da batimetria derivada de satélite está trazendo resultados cada vez mais precisos com o avanço tecnológico e aprimoramento dos sensores, uma vez que esse método permite uma produtividade maior se comparado aos sensores acústicos e podem contribuir para a segurança da navegação, detectando perigos que podem ser representados em cartas e, portanto, evitado por navegadores. Assim, este trabalho busca apontar, por meio de uma avaliação dos sistemas Landsat 8, CBERS 4 e Sentinel 2A, em conjunto com métodos empíricos de Machine Learning, que são abordagens mais efetivas para estimativa da batimetria derivada de satélites em diferentes áreas de estudo. A partir resultados foi possível perceber a superioridade do sistema Landsat 8 em relação aos demais, principalmente quando empregado junto com o algoritmo Random Forest, o qual possibilitou o alcance de valores de RMSE inferiores que as demais abordagens, até mesmo analisando as discrepâncias em relação a profundidades dos rios. Além disso, foi possível observar que a imagem CBERS 4 é capaz de estimar a profundidade com apenas 0,48 m de incerteza para corpos hídricos com uma menor taxa de sedimentos em suspensão.

Palavras-chave: Batimetria. Satellite Derived Bathymetry. Métodos Empíricos. Machine Learning.

## ABSTRACT

ANDRADE, Laura Coelho de, M.Sc., Universidade Federal de Viçosa, February, 2023. **Evaluation of orbital sensors in the determination of satellite-derived bathymetry: approach with empirical methods.** Adviser: Italo Oliveira Ferreira.

In the hydrographic scope, it is known that the main technique used for the bathymetric survey of a water body consists of acoustic remote sensing, with the use of echo sounders, since these provide accurate information even at great depths. The acquisition of bathymetric data, with a focus on updating and editing nautical charts, is the main objective of a Hydrographic Survey. However, the procedure for data acquisition using acoustic probing methods is expensive, in addition to requiring a long execution period. In this sense, the use of techniques to obtain satellite-derived bathymetry is bringing more and more accurate results with technological advancement and sensor improvement, since this method allows for greater productivity when compared to acoustic sensors and can contribute to safety of navigation, detecting dangers that can be represented in charts and, therefore, avoided by navigators. Thus, this work seeks to point out, through an evaluation of the Landsat 8, CBERS 4 and Sentinel 2A systems, together with empirical methods of Machine Learning, which are more effective approaches for estimating the bathymetry derived from satellites in different areas of study. Based on the results, it was possible to perceive the superiority of the Landsat 8 system in relation to the others, especially when used together with the Random Forest algorithm, which allowed the achievement of lower RMSE values than the other approaches, even analyzing the discrepancies in relation to river depths. In addition, it was possible to observe that the CBERS 4 image is capable of estimating the depth with only 0.48 m of uncertainty for water bodies with a lower rate of suspended sediments.

Keywords: Bathymetry. Satellite Derived Bathymetry. Empirical Methods. Machine Learning.

## LISTA DE FIGURAS

Figura 1 - Ilustração dos sistemas monofeixe (a) e multifeixe (b).....	15
Figura 2 - Processo de reamostragem e interpolação na imagem original.....	18
Figura 3 - Trajetória da luz entre a fonte, alvo e sensor orbital, com o processo de espalhamento e absorção.....	18
Figura 4 - Localização da área de estudo I.....	25
Figura 5 - Localização da área de estudo II .....	26
Figura 6 - Fluxograma metodológico .....	27
Figura 7 - Gráficos do RMSE, MAE e $R^2$ para o SDB no rio Madeira e Paraguai .....	34
Figura 8 - MDPs para o rio Madeira.....	38
Figura 9 - MDPs para o rio Paraguai.....	39

## LISTA DE TABELAS

Tabela 1 - Principais características dos sistemas empregados.....	24
Tabela 2 – Análise exploratória das discrepâncias para o rio Madeira.....	32
Tabela 3 – Análise exploratória das discrepâncias para o rio Paraguai .....	32
Tabela 4 – Valores obtidos para RMSE, MAE, $\Phi$ Robusta e $R^2$ do SBD para o rio Madeira e Paraguai .....	33
Tabela 5 – Análise da discrepância por profundidade – rio Madeira .....	35
Tabela 6 - Análise da discrepância por profundidade – rio Paraguai .....	36

## LISTA DE SIGLAS E ABREVIATURAS

ASV	Autonomous Surface Vehicle
AUV	Autonomous Underwater Vehicle
BAG	Bagging
DHN	Diretoria de Hidrografia e Navegação
DOS	Dark Object Subtraction
GBM	Gradient Boosting Machine
GNSS	Global Navigation Satellite System
IHO	International Hydrographic Organization
IOP	Inherent Optical Properties
LH	Levantamento Hidrográfico
LiDAR	Light Detection And Ranging
LSB	Least Square Boosting
MAE	Mean Average Error
MBES	Multi Beam Echo Sounder
MDP	Modelo Digital de Profundidade
MDT	Modelo Digital do Terreno
ML	Machine Learning
NDWI	Normalized Difference Water Index
NIR	Near Infrared
RF	Random Forest
RMSE	Root Mean Square Error
RNA	Redes Neurais Artificiais
RPA	Remotely Pilot Aircraft
RTK	Real Time Kinematic
SBES	Single Beam Echo Sounder
SDB	Satellite Derived Bathymetry
SE	Semi Empírico
SVM	Support Vector Machine
SWIR	Short Wave Infrared
USV	Unmanned Surface Vehicle
UTM	Universal Transversa de Mercator

## SUMÁRIO

1. INTRODUÇÃO.....	11
2. OBJETIVOS.....	13
3. JUSTIFICATIVA.....	14
4. REVISÃO DE LITERATURA.....	15
4.1. Levantamento Batimétrico.....	15
4.2. Imagens orbitais.....	16
4.3. Correção Geométrica e Radiométrica.....	17
4.4. Métodos empíricos para estimativa da batimetria por resposta espectral.....	20
4.4.1. <i>Machine Learning</i> .....	20
4.4.2. Redes Neurais Artificiais.....	20
4.4.3. <i>Random Forest</i> .....	21
4.4.4. <i>Gradient Boosting Machine</i> .....	22
4.4.5. <i>Support Vector Machine</i> .....	22
5. METODOLOGIA.....	23
5.1 Materiais.....	23
5.2. Área de estudo I – Rio Paraguai.....	25
5.3 Área de estudo II – Rio Madeira.....	26
5.4 Métodos.....	27
6. RESULTADOS.....	31
7.1 DISCUSSÕES.....	40
7.1. Parâmetros da incerteza vertical.....	40
7.2. Intervalos de profundidade.....	42
7.3. Modelos Digitais de Profundidade.....	43
8. CONCLUSÕES.....	44
REFERÊNCIAS.....	46
APÊNDICE A – Algoritmos empregados.....	52
APÊNDICE B – Histogramas das discrepâncias e Gráficos de dispersão.....	63
APÊNDICE C – Modelos digitais das discrepâncias.....	81

## 1. INTRODUÇÃO

Os rios e reservatórios são de essencial importância para a disponibilidade hídrica nos setores doméstico, industrial, de irrigação, hidrelétrica e de navegação. O aprimoramento da gestão desses locais em relação ao controle do volume de água, cálculo do nível de assoreamento e erosão pode contribuir de forma significativa para o desenvolvimento de todo meio ambiente local, podendo até mesmo facilitar o entendimento acerca dos impactos das mudanças climáticas (BIRKETT,1995; CRÉTAUX & BIRKETT, 2006; CRÉTAUX et al., 2011).

Para o conhecimento da profundidade de corpos hídricos e conseqüentemente seu volume, normalmente são utilizadas técnicas de levantamento batimétrico com recurso a ecobatímetros monofeixe ou multifeixe associadas a sistemas GNSS (Global Navigation Satellite System) e sensores inerciais para realização do posicionamento da embarcação (FERREIRA, RODRIGUES; SANTOS, 2015). Esse procedimento tem como intuito final a geração de cartas náuticas interiores e exteriores ao continente, em que a DHN (Diretoria de Hidrografia e Navegação) é o órgão responsável por gerar, editar e publicar esse documento, tendo com base algumas normas internacionais, como a S-44.

Com o crescente progresso da tecnologia, sistemas autônomos e não tripulados começaram a surgir nesse âmbito, podendo citar como exemplo o USV (*Unmanned Surface Vehicle*) ou ASV (*Autonomus Surface Vehicle*). Essas ferramentas possibilitam a coleta de dados em locais que oferecem perigo aos tripulantes e até mesmo locais rasos, onde embarcações com maiores calados não conseguem alcançar. Autores como Manley (2008), Giordano et al., (2015) e Ferreira et al., (2016a) estudaram esses veículos autônomos e puderam constatar que essas tecnologias também reduzem o custo e tempo de aquisição em campo.

O emprego do sensoriamento remoto na batimetria também é uma realidade existente. Essa técnica, conhecida na literatura internacional como SDB (*Satellite Derived Bathymetry*), ou batimetria por resposta espectral, é uma metodologia alternativa e de relativo baixo custo, se comparado aos métodos acústicos (FERREIRA et al., 2016b). O princípio básico dessa técnica se baseia na penetração da radiação eletromagnética na água em diferentes comprimentos de onda, tendo a profundidade do local proporcional à porção da radiação que atravessa a coluna d'água.

Avanços nessa área estão ocorrendo com certa rapidez, uma vez que o desenvolvimento dos sistemas orbitais com melhores resoluções espaciais e radiométricas, bem como a atualização de sensores acoplados a RPA's (*Remotely Pilot Aircraft*) têm permitido

encontrar uma profundidade estimada com uma incerteza vertical menor que 1 (um) metro (FERREIRA et al., 2016b; AGRAFIOTS et al., 2019; CAHALANE et al., 2019; ANDRADE et al., 2021).

Além dos diferentes sistemas orbitais e aéreos existentes, destaca-se também a vasta quantidade de métodos abordados na literatura para execução da predição da batimetria. De acordo com Ashphaq, Srivastava e Mitra (2021) existem duas abordagens principais para o SDB, sendo a primeira baseada em conceitos físicos, com métodos semi-empíricos e analíticos, e a segunda com métodos empíricos, que são de fácil aplicação e se baseiam puramente em abordagens estatísticas, necessitando assim de dados coletados in situ para a estimativa da profundidade.

A abordagem física, apesar de possuir inúmeras vantagens como a de levar em consideração as IOP's (*Inherent Optical Properties*) e a reflectância espectral do fundo submerso, se constitui como um método complexo, uma vez que se os parâmetros do corpo hídrico forem estimados incorretamente, ou até mesmo se não for aplicada uma correção atmosférica e radiométrica adequada nas cenas, o algoritmo não irá performar da melhor forma, informando valores incorretos para a profundidade estimada (ILORI & KNUDBY, 2020).

Por outra perspectiva, os métodos empíricos são de fácil implementação e empregam principalmente ferramentas de regressão para análise de dados usando os valores das bandas com dados in situ para calcular os coeficientes da equação de regressão. Lyzenga (1978) modificou a técnica da razão para desenvolver um método mais genérico com a utilização de uma banda única para a radiância de águas rasas e com isso pode identificar a profundidade do local, bem como as características de fundo, empregando a técnica da transformação log linear. Warne (1978) usou da técnica de banda única com imagens Landsat para profundidades de até 30 metros e conseguiu uma acurácia de 10% em profundidades de até 20 metros. Mishra et al., (2018) estudou a batimetria de locais distintos utilizando PCA (*Principal Component Analysis*) e imagens IKONOS, encontrando um  $R^2$  de 0.90 e um erro padrão de 0.64 m.

Com o advento do *Machine Learning* (ML), o emprego de métodos como Redes Neurais Artificiais (RNA), bem como algoritmos de árvores de decisão, flexibilizou a técnica de processamento de uma elevada quantidade de dados em um curto espaço de tempo. Ribeiro et al., (2008) empregaram uma RNA com multicamadas para obter um Modelo Digital de Profundidade (MDP), encontrando incertezas na faixa de 0,50 metros. Sagawa et al. (2019) utilizaram do *Random Forest* (RF) em cinco áreas de estudo para profundidades de até 20 metros e encontraram um RMSE (*Root Mean Square Error*) de 1,41 metros.

No entanto, é necessário destacar que a escolha de um sistema orbital pode interferir de forma significativa na predição do SDB, do mesmo modo que as correções efetuadas na cena da imagem (ex. correção geométrica, radiométrica, do brilho solar, etc.) e as técnicas de predição utilizadas.

A avaliação de diferentes sensores em combinação com técnicas comumente empregadas na literatura pode facilitar a escolha do usuário que deseja realizar a predição da batimetria por resposta espectral, além de trazer diferentes conclusões a respeito de um local estudado, de forma rápida, efetiva e por um custo relativamente menor se comparado aos métodos batimétricos tradicionais.

## **2. OBJETIVOS**

### **2.1 Objetivo geral**

O objetivo principal do trabalho consiste em efetuar um estudo comparativo das técnicas empíricas de obtenção da batimetria por resposta espectral utilizando diferentes sistemas orbitais.

### **2.2 Objetivos específicos**

Dentre os objetivos específicos, tem-se:

- Avaliar a incerteza vertical da batimetria estimada para três sistemas orbitais gratuitos utilizando diferentes técnicas empíricas para a predição da profundidade;
- Analisar a aplicabilidade das imagens e dos métodos para corpos hídricos distintos, com diferentes morfologias de fundo;
- Avaliar por intervalos de profundidade a incerteza vertical da batimetria estimada utilizando métodos empíricos, como o GBM, RF, RNA e SVM;
- Analisar a influência da resolução radiométrica e espacial na qualidade dos dados batimétricos.

### 3. JUSTIFICATIVA

A batimetria por resposta espectral ainda consiste em uma técnica de mapeamento em ascensão e com grandes perspectivas de investimentos com o desenvolvimento de novas tecnologias, concomitante ao avanço de pesquisas nesse âmbito.

No que concerne a realização da predição da profundidade, encontra-se na literatura várias opções para realizar a estimativa, com sensores multiespectrais e hiperespectrais associados a diferentes abordagens para a estimativa da batimetria (LYZENGA, 1978; LYZENGA et al. 2006; MOHAMED & NADAOKA, 2017; POURSANIDIS et al. 2019b). Além disso, é importante ressaltar que como apontado por Laporte et al., (2020), a IHO (International Hydrographic Organization) iniciou uma revisão da publicação da norma S-44, referente aos padrões a serem obedecidos para os Levantamentos Hidrográficos (LH's), com o principal intuito de inserir as informações coletadas com sistemas orbitais para servir de instrumento aos hidrógrafos. Tal episódio poderá contribuir sobremaneira para a difusão dessa técnica no mundo todo, visto que a IHO é uma organização intergovernamental.

Atualmente têm-se aplicado em larga escala métodos empíricos, semi-empíricos e semi-analíticos, sendo esses pontuais e focados em sistemas orbitais que fornecem maior resolução espacial e radiométrica, como World-View-2 e 3 e imagens do sistema Landsat (GABR; AHMED; MARMOUSH, 2020, ASHPHAQ; SRIVASTAVA; MITRA, 2021).

Além disso, como colocado por Ashphaq, Srivastava e Mitra (2021), poucos estudos relacionados aos métodos empíricos têm considerado fatores que podem afetar a determinação da SDB, como realizado por Tripathi e Rao (2002) que estudou a influência da turbidez na batimetria utilizando a Banda 1 dos sistemas IRS-1D LISS-III.

Isto posto, estudos que comparam sistemas e metodologias são de grande importância na área, uma vez que essa ainda necessita de mais pesquisas em ambientes aquáticos rasos com morfologias de fundo distintas, considerando parâmetros como nível de turbidez presente na água, especialmente com imagens que são disponibilizadas de forma gratuita, propiciando um custo consideravelmente baixo na aquisição da batimetria por resposta espectral.

Sendo assim, a contribuição principal deste trabalho é discutir a importância do uso de imagens orbitais gratuitas, sendo essas, CBERS 4, Sentinel 2A e Landsat 8 em conjunto com técnicas empíricas para obtenção da batimetria por resposta espectral, sobretudo a aplicabilidade em corpos d'água interiores, favorecendo assim a periodicidade do mapeamento do corpo hídrico, com a possível mitigação e controle de riscos de erosão e assoreamento.

## 4. REVISÃO DE LITERATURA

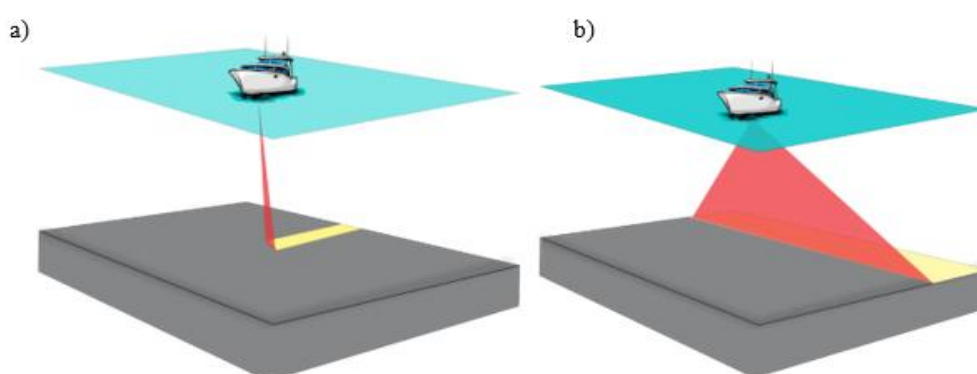
### 4.1. Levantamento Batimétrico

O levantamento batimétrico é considerado um dos principais afazeres de um LH (FERREIRA,2018). Os dados advindos desse levantamento são obtidos, tradicionalmente, por meio de um conjunto de instrumentos instalados em plataformas de levantamentos, que podem ser submergíveis, como os submarinos e os AUV (*Autonomous Underwater Vehicle*), de superfície, como as tradicionais embarcações e navios tripulados, e também plataformas aéreas. Embarcam-se nesses vários instrumentos, como sensores inerciais, receptores GNSS (*Global Navigation Satellite System*), perfiladores de velocidade do som, sonares de varredura lateral, perfiladores de subfundo e ecobatímetros.

Em resumo, os ecobatímetros são constituídos mais comumente por um transdutor piezoelétrico que gera e recebe sinais acústicos (MAYER,2006) e, após, retornam para a superfície do nível da água, fornecendo uma profundidade. Esses equipamentos podem ser SBES (*Single Beam Echo Sounder*), também chamado de monofeixe, ou MBES (*Multi Beam Echo Sounder*), isto é, multifeixe.

O SBES tem como produto final uma linha de pontos, uma vez que este realiza apenas um registro da profundidade a cada ping ou pulso transmitido. Já o MBES efetua a coleta de várias profundidades com um mesmo pulso, realizando a aquisição da batimetria por faixa, conforme evidenciado na Figura 1.

Figura 1 - Ilustração dos sistemas monofeixe (a) e multifeixe (b)



Fonte: Ferreira et al. (2022).

Ressalta-se que além da maior produtividade, os MBES apresentam um elevado ganho em resolução e acurácia, tanto em termos planimétricos quanto altimétricos (profundidade), e também um grande adensamento de dados, descrevendo quase que por completo o fundo

submerso. Por este motivo, os ecobatímetros multifeixe são utilizados obrigatoriamente na produção e/ou atualizações de cartas náuticas.

Por outro lado, em rios e reservatórios, ainda prevalece o emprego de sistemas monofeixe, sobretudo em levantamentos no âmbito da resolução conjunta ANA/ANEEL número 003 de 2010, conforme é descrito por (ESTIGON et al., 2015).

#### **4.2. Imagens orbitais**

As imagens de satélite são usadas há mais de trinta anos na avaliação de recursos hídricos e na estimativa da batimetria por resposta espectral (LAPORTE et al., 2020), devido a sua facilidade para garantir uma periodicidade entre os dados que necessitam de monitoramento e atualizações constantes. Desta forma, dados com informações sobre a morfologia de fundo, profundidade e volume são adquiridos de forma rápida e precisa, além de proporcionarem relativamente um baixo custo.

A aplicação dessas técnicas envolve, dentre outras funções, o delineamento de corpos d'água utilizando ferramentas de extração de informações temáticas. Além disso, existem inúmeros métodos para aquisição de informações da água por meio do sensoriamento remoto, que, de acordo com o número de bandas usadas, geralmente são divididas em duas categorias, sendo elas de banda única e multibanda.

O método da banda única se limita a escolha de somente uma banda de uma imagem multiespectral e posteriormente um limiar é determinado para a discriminação de água e solo (RUNDQUIST et al., 1987). Já o multibanda utiliza das diferenças reflexivas de cada banda envolvida, através da análise de características de assinatura de cada alvo.

Nessa perspectiva, GAO (2009) chegou à conclusão de que o método passivo de determinação de batimetria, fundamentado nas bandas do infravermelho próximo e médio do sensor TM do sistema Landsat fornece um melhor desempenho quando comparado a demais sensores e pode mapear uma área considerável em um menor espaço de tempo. Ferreira et al., (2016b) concluíram que imagens do satélite RapidEYE, com uma resolução espacial de 5 metros podem fornecer insumos favoráveis em análises ambientais em reservatórios de águas rasas.

Collischonn e Clarke (2016) empregaram imagens CBERS e Landsat e dados de nível d'água para estimativa de curvas cota-volume de um reservatório em Minas Gerais e encontraram uma diferença de apenas 0,2% no volume encontrado com batimetria convencional.

Leder e Leder (2017) realizaram estudos no meio do Mar Adriático, na região de Kastela Bay com imagens do satélite Landsat 8 e obtiveram resultados promissores, concluindo que a batimetria derivada de satélites fornece um melhor desempenho quando se opera em águas rasas e claras (aproximadamente duas vezes a profundidade encontrada no Disco de Secchi). Poursanidis et al., (2019a) encontraram um erro que se encaixa nos requisitos propostos pela Organização Internacional de Hidrografia, utilizando de imagens PlanetScope para uma área com uma profundidade máxima de 10 metros na Ilha de Creta. Gabr, Ahmed e Marmoush (2020) também empregaram imagens do PlanetScope para a costa norte do Egito e conseguiram obter uma incerteza de 0,4 metros.

Casal et al. (2020) apresentaram o uso de imagens Sentinel 2 para a estimativa da batimetria em Dublin Bay, na costa leste da Irlanda, para profundidades de até 10 metros e encontraram uma incerteza de aproximadamente 1,4 metros.

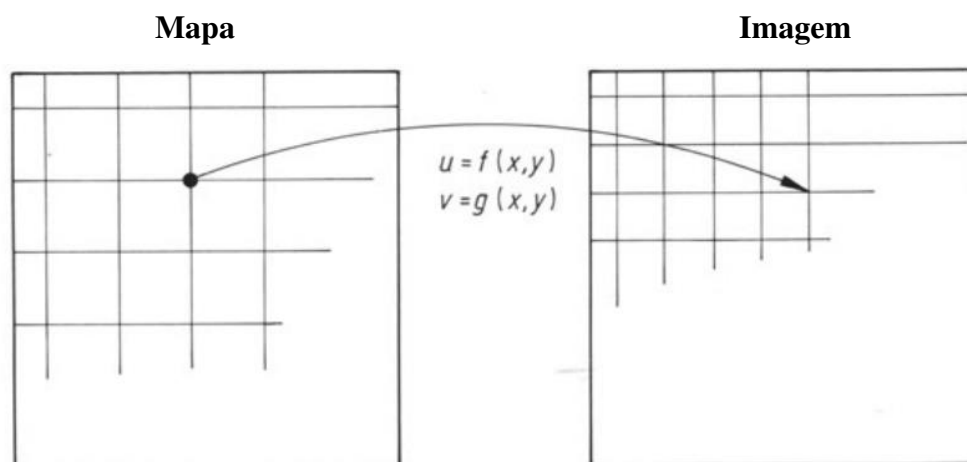
#### **4.3. Correção Geométrica e Radiométrica**

O pré-processamento de uma imagem é composto por correções de distorções geométricas e radiométrica que ocorrem no momento de aquisição dessa imagem.

Uma das correções fundamentais em imagens orbitais é a correção geométrica. Normalmente, essa correção está atrelada a efeitos sistemáticos de distorção existentes nas imagens, decorrentes da rotação da terra no momento da obtenção da imagem, fenômeno também chamado de *skew*, da instabilidade da plataforma (inclinações dos eixos), variações de velocidade e altitude da plataforma e distorções panorâmicas (MENESES & ALMEIDA, 2012). Além destas, quando se trata de imagens de alta resolução, há a necessidade também da correção referente ao deslocamento devido ao relevo.

Para a execução desse procedimento, em um primeiro passo, deve ser efetuado o mapeamento direto, que equivale a uma transformação geométrica com as coordenadas da imagem e as coordenadas geográficas. Após, necessita-se realizar uma inversão no mapeamento anterior, para que ocorra o estabelecimento dos níveis de cinza da imagem corrigida. O último processo, denominado de reamostragem, realiza essa definição dos níveis de cinza através de uma interpolação na imagem original (Figura 2) (SHLIEN,1979).

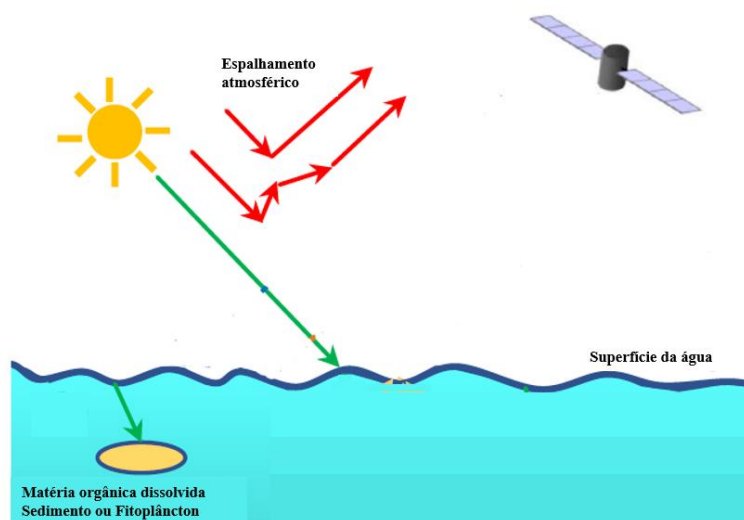
Figura 2 - Processo de reamostragem e interpolação na imagem original



Fonte: Richards (1993).

Além da realização da correção geométrica na imagem, deve-se também corrigir os valores de brilho da cena, em função, principalmente, dos efeitos atmosféricos, absorção e espalhamento (Figura 3).

Figura 3 - Trajetória da luz entre a fonte, alvo e sensor orbital, com o processo de espalhamento e absorção



Fonte: Adaptado de Barbosa et al. (2019).

De acordo com Zullo et al., (1996) o espalhamento consiste na mudança da direção de propagação da irradiação do sol, em razão da sua interação com partículas pequenas de aerossóis e gases em suspensão. Já absorção, tem um efeito de subtração, visto que esse absorve energia

de vários comprimentos de onda. Nesse sentido, é necessário efetuar uma correção atmosférica no objeto de estudo, para que as análises *a posteriori* não estejam eivadas com efeitos indesejáveis.

Autores como Liang et al., (2001) classificaram os métodos utilizados para efetuar tais correções radiométricas, mais especificamente atmosféricas, em quatro grupos, sendo esses: de objetos com reflectância constante, *matching* de histogramas, reflectância de objetos escuros e redução de contraste. Sendo o primeiro o mais simples, realizando apenas uma correção relativa por métodos estatísticos e tendo sua maior limitação na dificuldade de corrigir os ruídos causados por aerossóis heterogêneos.

O método comumente utilizado consiste no DOS (*Dark Object Subtraction*) (CHAVEZ,1988; ZHANG et al. 2010; CAHALANE et al. 2019). Sabe-se que nessa técnica é realizada a transformação dos níveis digitais para a subtração da radiância do objeto escuro, equivalente a 1% da reflectância (CONGEDO,2016). Assim, a radiância de subtração pode ser calculada por meio da Equação 1.

$$Lp = Lb - L1\% \quad (1)$$

Em que  $Lp$  consiste na radiância de subtração;  $Lb$  a radiância da banda e  $L1\%$  a radiância do objeto escuro.

A imagem final corrigida possui valores de reflectância na superfície terrestre (Equação 2)

$$\rho = \frac{\pi (Lb-Lp) x d^2}{ESUN x \cos Z} \quad (2)$$

Em que  $d$  consiste na distância entre a terra e o sol,  $ESUN$  a irradiância solar e  $\rho$  a reflectância na superfície terrestre.

No caso da absorção, métodos de correção mais complexos de transferência radiativa podem ser utilizados a fim de estabelecer a atmosfera no momento de aquisição das imagens. O algoritmo denominado FLAASH (*Fast Line of Sight Atmospheric Analysis of Spectral Hypercubes*) corresponde a uma ferramenta que utiliza como modelo de transferência radiativa a transmissão atmosférica de resolução moderada, conhecida como MODTRAN. Esse método efetua a correção de comprimentos de onda no visível por meio das bandas do NIR e SWIR e

emprega a taxa de refletância do pixel escuro para extrair a quantidade de aerossol e posteriormente estimar a visibilidade média da cena (MATTHEW et al.,2002).

#### **4.4. Métodos empíricos para estimativa da batimetria por resposta espectral**

##### **4.4.1. *Machine Learning***

O *Machine Learning* (ML) ou aprendizado por máquina compõe-se em uma técnica que obtém conclusões genéricas com base em um conjunto particular. De acordo com Samuel (1959) é possível classificar essa ferramenta em três tipos de aprendizado, sendo essas, o aprendizado supervisionado, não supervisionado e por reforço. O aprendizado supervisionado tem como intuito a predição com base em uma função que foi aprendida, mapeando uma saída de acordo com a respectiva entrada (RUSSEL & NORVIG,1995). Já o não supervisionado, segundo Hinton e Sejnowski (1999), é realizado com subcategorias que são capazes de entender e assimilar padrões sobre amostras que não constam previamente no treinamento. O último aprendizado supracitado, por reforço, efetua ações com o objetivo de maximizar a recompensa cumulativa (HU et al.,2020), todavia não é um método amplamente empregado.

Assim sendo, salienta-se que as técnicas de ML são utilizadas, em sua totalidade, em problemas que tem como objetivo encontrar tendências ou padrões para facilitar o entendimento dos dados. Essa ferramenta também possui alta flexibilidade para trabalhar com grandes amostras de dados, além de ser de fácil aplicabilidade (ASHPHAQ; SRIVASTAVA; MITRA, 2021). Autores como Ceyhun e Yalcin (2010) utilizaram de Redes Neurais em imagens do sensor Aster e Quickbird e encontraram um  $R^2$  de 0.8. Hassan et al., (2017) compararam a batimetria por resposta espectral empregando cinco métodos distintos de ML, sendo esses, *Bagging* (BAG), *Least Square Boosting* (LSB), *Support Vector Regression* (SVR), Redes Neurais Artificiais (RNA) e *Lyzenga Generalised Linear Model* (GLM) em conjunto com dados dos sistemas Landsat 8 e SPOT-6 e descobriram que o algoritmo BAG produziu resultados mais acurados que os demais.

Dessa forma, acredita-se que o ML pode gerar produtos com baixa incerteza vertical empregado em conjunto com sensores orbitais, além de ser uma ferramenta emergente no âmbito do SDB, com grande potencial para ascensão.

##### **4.4.2. Redes Neurais Artificiais**

As Redes Neurais Artificiais (RNA's) tem seu comportamento inspirado na estrutura do cérebro humano, que consiste em várias unidades de processamento bem estruturadas e

interligadas por canais de comunicação, sendo que cada canal tem seu peso, correspondente a um valor numérico. Desde a década de 90 inúmeros autores como Tagliarini et al. (1991), Møller (1993) e Hagan e Menhaj (1994) já discursavam sobre a utilização de redes neurais através da aprendizagem computacional e por reconhecimento de padrões, com a intenção de aproximar ainda mais o processamento das máquinas ao cérebro humano.

Assim, diferentes métodos e algoritmos foram propostos para efetuar o aprendizado dessas redes (ANASTASIADIS et al., 2015). Møller (1993) criou o método de gradientes conjugados para o aprendizado supervisionado da rede, além desse o algoritmo Broyden-Fletcher-Goldfarb-Shanno (GILL et al., 1981) e Levenberg-Marquardt (HAGAN; MENHAJ, 1994) também são algumas alternativas para os treinamentos de RNA's. De acordo com Khadse et al., (2016), o método denominado *backpropagation* passou a ser comumente utilizado para o treinamento de redes do tipo MLP (*MultiLayer Perceptron*).

Durante o processo de treinamento das redes, ou também chamado de aprendizagem, os pesos aleatoriamente definidos, vão se ajustando para atingir o objetivo pretendido. Em um neurônio artificial, as entradas são definidas por valores numéricos, que são multiplicados pesos e somados para compor um único sinal de entrada (Equação 3) (JAIN et al., 1996)

$$S = a + \sum x_i * w_i \quad (3)$$

Em que  $a$  é um valor constante do neurônio, chamado de “bias”,  $x_i$  representa os valores numéricos de entrada e  $w_i$ , os pesos de cada entrada. Ressalta-se que a soma ponderada é modelada com uma função de transferência, que pode ser binária, linear ou exponencial.

Segundo Haykin (1999), para solução de determinado problema, a RNA deve ser configurada com o tipo de rede, arquitetura adequada e ajuste de pesos adequados, fazendo com que o vetor de saída coincida com certo valor desejado.

No processo de aprendizado (treinamento), quando a saída da rede é calculada essa resposta é comparada com o vetor de saída esperado e assim é possível encontrar um erro. De acordo com esse mesmo erro, os pesos são ajustados com o intuito de diminuir essa discrepância. O processo de treinamento é repetido até que o erro para o conjunto encontre um valor mínimo ou previamente determinado.

#### 4.4.3. *Random Forest*

O *Random Forest* (RF) é um método de *Machine Learning* baseado em árvores de decisão. Em seu modo de regressão, as árvores são criadas e cada uma delas gera um conjunto

de dados em novos terminais. Usualmente geram árvores grandes e independentes, além disso, essa técnica possui baixo viés e alta variância (BREIMAN, 2001).

Para realizar o treinamento no RF são necessários três hiperparâmetros, sendo esses o “*ntree*”, que representa o número de árvores; “*nodesize*”, que é o número mínimo de pontos de dados em cada nó terminal e por último o “*mtry*”, que consiste no valor de variáveis a serem escolhidas de forma aleatória para cada nó da árvore. Dessa forma, para cada amostra é obtido o mesmo número de resultados para cada árvore criada (HOUBORG & MCCABE, 2018). O valor de saída de um modelo de regressão do RF é a média dos *outputs* de todas as árvores independentes.

#### **4.4.4. Gradient Boosting Machine**

O *Gradient Boosting Machine* (GBM), assim como o RF, é um algoritmo de *machine learning* que utiliza de árvores de decisão para a execução da predição de novas amostras. Segundo Natekin e Knoll (2013), esse método se fundamenta na confecção de novos aprendizados, com o intuito de se correlacionar com gradientes negativos de uma função de perda arbitrária e por fim vincular com o conjunto de dados.

O GBM se diferencia do RF uma vez que emprega árvores pequenas e sequenciais. Ademais, essas visam a minimização do erro, ou seja, cada árvore tem peso proporcional ao nível de acerto e com isso, árvores com resultados ruins possuem pesos menores.

Outro ponto a salientar é que o GBM tende a ter um alto viés se comparado ao RF. Autores como Bissacco et al. (2007) utilizaram do GBM principalmente devido a praticidade de adaptação do algoritmo aos dados e também uma alta flexibilidade. Já Touzani et al. (2018) empregaram o GBM para modelar o consumo de energia em prédios comerciais e obtiveram predições mais acuradas em 80% dos casos se comparados com modelos tradicionais que empregam a regressão linear ou até mesmo o RF.

#### **4.4.5. Support Vector Machine**

O *Support Vector Machine* (SVM) tem como base princípios simples que foram originados da teoria estatística, por meio da aplicação de um método linear em um espaço multi-dimensional e complexo (VAPNIK, 1998). Essa ferramenta utiliza de uma função, podendo ser polinomial, kernel quadrática, radial, sigmóide, dentre outras, com o intuito de mapear com precisão os dados e assim executar um treinamento eficaz.

No SVM as novas predições de entrada dependem de uma avaliação prévia da função determinada no estágio de treinamento. O algoritmo normalmente é formulado para resolver um projeto de otimização convexa do tipo programação quadrática (QP), efetuada utilizando a função de perda Vapnik's  $\epsilon$ -insensitive (DUAN & KEERTHI, 2005).

Vojinovic et al., (2013) aplicaram o SVM para modelar a estimativa da batimetria por resposta espectral em águas rasas e conseguiram atingir uma MAE (*Mean Average Error*) de 1,8m para profundidades de até 15 metros. Misra et al. (2018) empregaram o SVM para águas rasas de 1 a 15 metros na Holanda, com imagens Landsat 7 e 8, com isso puderam afirmar que esse método de ML oferece melhores resultados de  $R^2$  para águas rasas que são mais turvas, encontrando valores de 0,73.

## **5. METODOLOGIA**

### **5.1 Materiais**

Para o desenvolvimento deste trabalho foram utilizadas as imagens descritas na Tabela 1, com as respectivas bandas e resoluções.

Tabela 1 – Principais Características dos sistemas empregados

<b>Sistemas</b>	<b>Bandas Espectrais (<math>\mu\text{m}</math>)</b>	<b>Resolução Espacial</b>	<b>Resolução Radiométrica</b>
<b>LANDSAT 8</b>	0,43-0,45 (Coastal Aerosol) 0,45 - 0,51 (Blue) 0,53 - 0,59 (Green) 0,64 - 0,67 (Red) 0,85 - 0,88 (NIR) 1,56 – 1,65 (SWIR 1) 2,11-2,30 (SWIR 2)	30m (Multiespectral)	16 bits
<b>CBERS 4 (MUX)</b>	0,46 - 0,52 (Blue) 0,50 - 0,59 (Green) 0,59 - 0,67 (Red) 0,78 - 0,86 (NIR)	10 m (Multiespectral)	8 bits
<b>SENTINEL 2A</b>	0,45 – 0,52 (Blue) 0,54 – 0,57 (Green) 0,65 - 0,68 (Red) 0,69 – 0,71 (Red Edge 1) 0,73 – 0,74 (Red Edge 2) 0,77 – 0,79 (Red Edge 3) 0,78 -0,90 (NIR 1) 0,85-0,87 (NIR 2) 1,56 – 1,65 (SWIR 1) 2,10 – 2,28 (SWIR 2)	10 m e 20m (Multiespectral)	12 bits

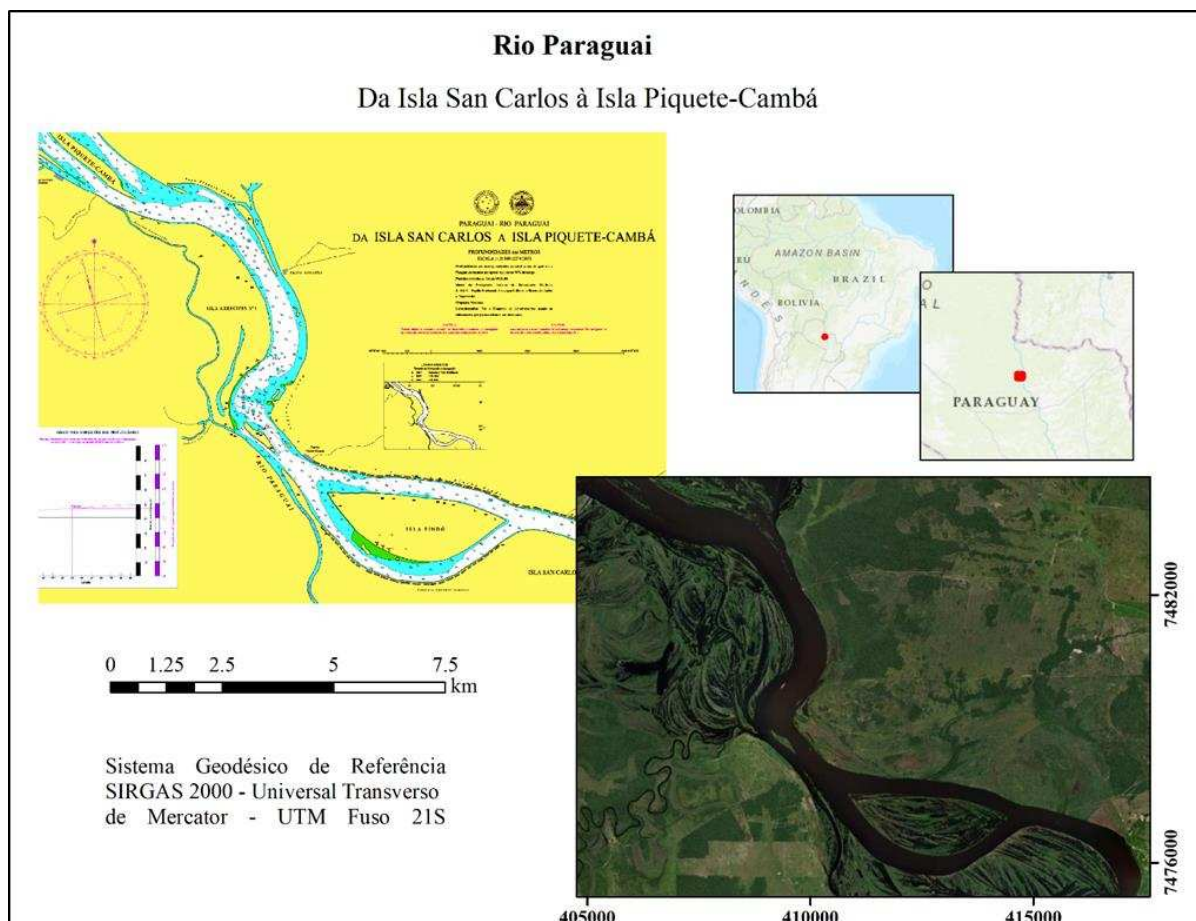
Além disso, também foram empregados os softwares:

- ArcGIS 10.5
- Idrisi Selva 17.0
- QGIS 3.0.0
- R (4.1.0)
- Excel

## 5.2. Área de estudo I – Rio Paraguai

A primeira área de estudo possui aproximadamente 16,37 km<sup>2</sup> e está localizada no rio Paraguai, entre a Isla San Carlos e a Isla Piquete-Cambá, situada em território Paraguai (Figura 4).

Figura 4 - Localização da área de estudo I



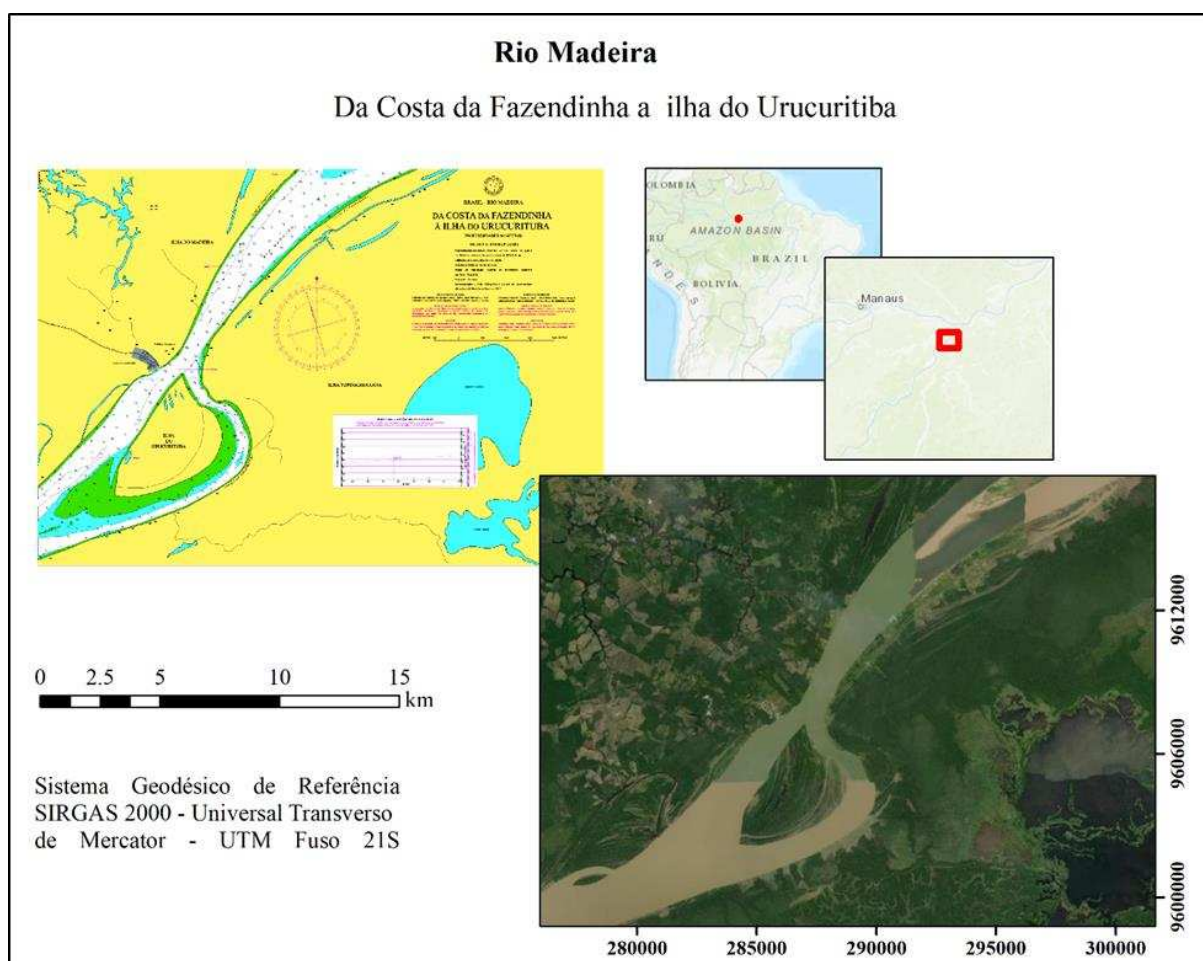
Salienta-se que, de acordo com a carta náutica disponibilizada pela Marinha, as profundidades chegam aproximadamente a 8,00 metros. Em termos gerais, o rio Paraguai é um corpo hídrico comumente utilizado para navegação, dado que é considerado um corpo hídrico de planície, com poucos acidentes geográficos (REYNALDO, 2007).

No Paraguai, esse rio é fonte de pesca, irrigação e cultivo de terras, permitindo a criação de gado, plantação de milho e algodão (FACETTI; DEKOV; VAN GRIEKEN, 1998). Dessa forma, com o intuito de produzir estudos mais detalhados sobre o corpo d'água na região, a área de estudo em questão foi escolhida após uma análise minuciosa desses critérios e da profundidade.

### 5.3 Área de estudo II – Rio Madeira

A segunda área de estudo, também consiste em uma carta náutica da marinha e está localizada no rio Madeira, da costa da fazendinha a ilha do Urucuritiba, com uma área de aproximadamente 61,33 km<sup>2</sup>. Em síntese destaca-se que as profundidades situadas nessa carta vão até aproximadamente 32,00 metros (Figura 5).

Figura 5 - Localização da área de estudo II



O rio Madeira possui uma extensão total de aproximadamente 3400 km, sendo um dos principais afluentes do rio Amazonas (BERNARDI et al., 2012). Esse curso d'água é responsável pela pesca e transporte hidroviário em várias regiões do Brasil, Bolívia e Peru, sobretudo entre Porto Velho e a foz do rio Amazonas, onde se situa a área navegável de 1340 metros de extensão (GOMES et al., 2019).

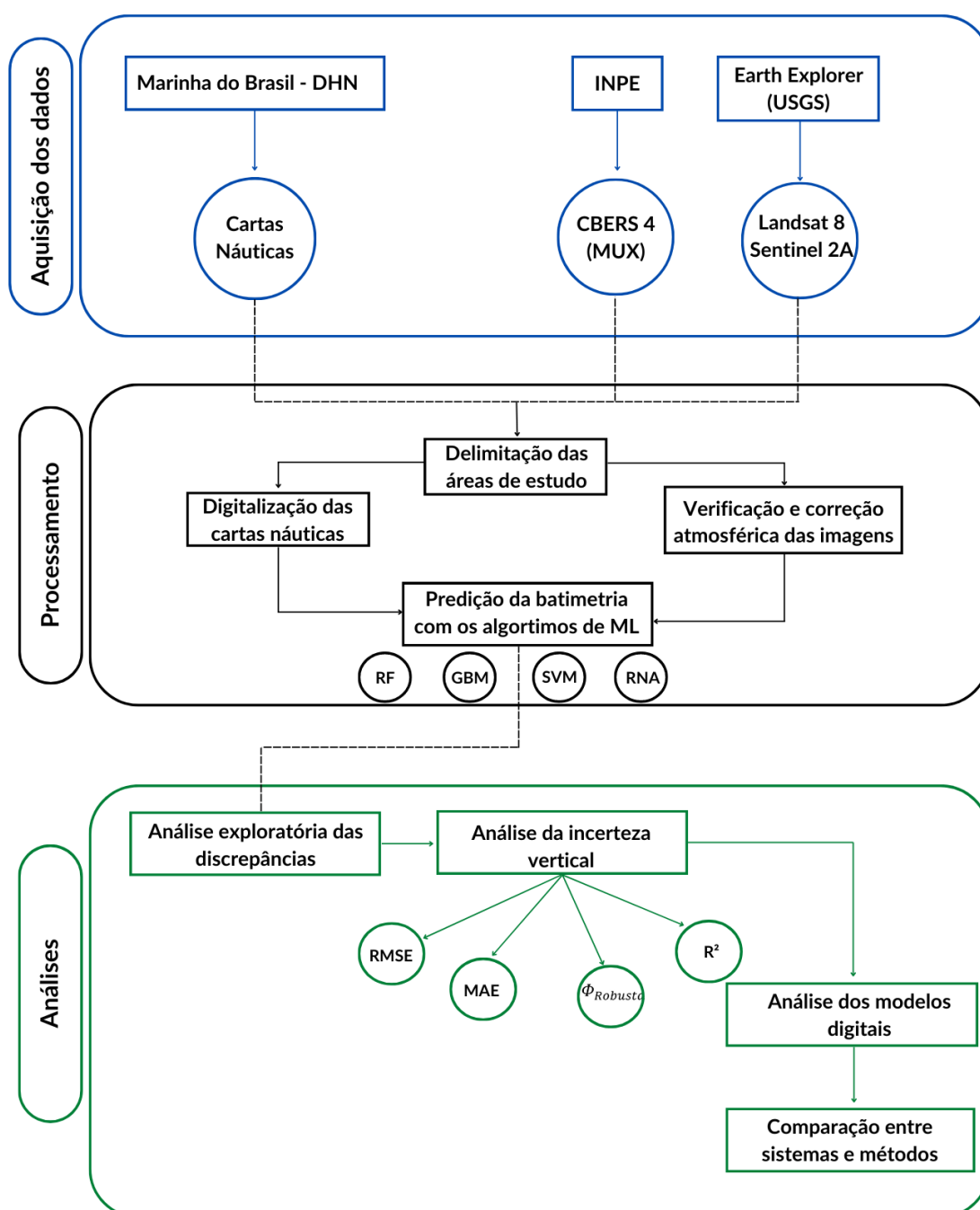
Tal como justificado para área de estudo I, o trecho do rio Madeira em questão foi escolhido após, primeiramente, a verificação da profundidade, bem como dos possíveis

impactos positivos que estudos mais detalhados podem gerar para a população local, sabendo-se que essa área é diretamente afetada por alterações no corpo hídrico, além de atingir de forma indireta os estados do Norte e Centro-Oeste do país.

## 5.4 Métodos

A metodologia adotada neste trabalho pode ser observada na Figura 6.

Figura 6 – Fluxograma metodológico



Para o desenvolvimento deste estudo foram utilizadas imagens orbitais dos sistemas Sentinel 2A, Landsat 8 e CBERS 4 (sensor MUX). Ressalta-se que tais sistemas disponibilizam as cenas adquiridas em sites como Earth Explorer (Landsat 8 e Sentinel 2A) e INPE (Instituto Nacional de Pesquisas Espaciais) (CBERS 4). A escolha dessas imagens se deu com base na data de aquisição das cartas náuticas dos locais empregados como área de estudo, que foram do ano de 2017.

É importante evidenciar que para cada sistema foram escolhidas cenas que possuem uma quantidade de nuvem abaixo de 20%. A data de aquisição foi selecionada para o período de seca nas áreas de estudo, visto que na carta náutica as profundidades são referidas em relação ao nível de redução, que consiste na menor profundidade que o corpo hídrico poderia atingir em um possível período de seca.

Nesse sentido, para o rio Madeira, foram selecionadas as imagens entre os meses de junho e novembro (GOMES et al.2019). Já para o rio Paraguai, entre os meses de abril a outubro (SANTOS et al. 2021).

No que concerne a ortorretificação das imagens, sabe-se que os produtos disponibilizados pelos sistemas Landsat 8 (USGS, 2013), Sentinel 2A (USGS,2015) e CBERS 4 (INPE,2014) já são ortorretificados com o uso de um MDT (Modelo Digital do Terreno).

Para correção radiométrica, foi utilizada a técnica DOS (Dark Object Subtraction), que é baseada na premissa de que objetos escuros não refletem luz e qualquer valor maior que zero deve resultar do espalhamento atmosférico (CHAVEZ, 1988). Nesse sentido, esse espalhamento deve ser subtraído de cada valor de pixel na banda (CAHALANE et al., 2019). Essa técnica foi empregada no software Idrisi Selva 17.2 para as imagens CBERS 4A. Para as imagens Landsat 8 e Sentinel 2A, foi utilizado o software QGIS 3.0.0, com o plug in SCP (Semi-automatic Classification).

Em um momento concomitante à correção das cenas mencionadas, foi realizada a digitalização das cartas náuticas adquiridas no site oficial da marinha brasileira. Esse procedimento foi efetuado no software QGIS 3.0.0 para determinação dos pontos batimétricos no formato vetorial, com as respectivas profundidades georreferenciadas. Dado que a quantidade de pontos vetorizados foi muito pequena, os mesmos não representavam os corpos hídricos de forma contínua. Nesse sentido, foi gerado um Modelo Digital de Profundidade dos dois rios empregando a Krigagem Simples, visto que essa comumente é utilizada quando não se observa tendência nos dados (Santos et al. 2011).

Em seguida, o raster do Modelo Digital de Profundidade foi transformado para um arquivo de pontos, o qual foi inserido no software R para o processamento utilizando os métodos Random Forest, Boosting GBM, SVM e RNA, para as duas áreas de estudo.

Nesse sentido, em todos os algoritmos foram selecionados, de forma aleatória, 70% da amostra de pontos para o treinamento dos métodos empíricos a serem avaliados, e os 30% restantes para o teste da predição. Autores como Parameswaran e Weinberger (2010), Verrelst et al. (2012), Cahalane et al. (2019), Mateo-Pérez et al. (2020) e Andrade et al. (2022), utilizaram desse método de separação dos dados para treinamento e teste e obtiveram resultados sem tendência, com a precisão menor que 1,00 metro para a estimativa da profundidade em ambientes rasos e límpidos.

Destaca-se que, para um melhor treinamento e predição da variável, além dos valores de radiância das bandas dos satélites empregados, também foram considerados nos métodos os índices NDWI (*Normalized Difference Water Index*) (Equação 4) e o NDTI (*Normalized Difference Turbidity Index*) (Equação 5), que foi desenvolvido por Lacaux et al., (2007) e utiliza dos valores de banda do vermelho e do verde.

$$NDWI \sim \frac{(\rho_{green}) - (\rho_{NIR})}{(\rho_{green}) + (\rho_{NIR})} \quad (4)$$

$$NDTI \sim \frac{(\rho_{red}) - (\rho_{green})}{(\rho_{red}) + (\rho_{green})} \quad (5)$$

Em que  $\rho_{green}$  representa o valor de radiância correspondente para a banda do verde;  $\rho_{NIR}$  o valor de radiância para a banda do infravermelho próximo e  $\rho_{red}$  o valor do pixel para a banda do vermelho.

O NDWI, criado por McFetters (1996), identifica corpos d'água. Além da representação e realce de feições existentes no ambiente aquático. Segundo Fisher et al. (2016), Soltanian et al. (2019), Ali et al. (2019) e Yang et al. (2020), esse índice se sobressai no monitoramento do comportamento hídrico visto que a utilização dele, além de possibilitar uma melhora significativa na visualização das áreas com presença de água e um alto teor de umidade, também apresenta uma variação quase linear se comparado a alteração na lâmina d'água.

Além do índice NDWI, aplicou-se o NDTI, que se trata de um índice criado por Lacaux et al., (2007) e empregado por autores como Cahalane et al., (2019) para predição de SDB com imagens Landsat 8, Pleiades e Rapideye. Os autores observaram que a turbidez da água pode resultar em valores maiores de radiância para a faixa do visível e também para o NIR,

subestimando profundidades em locais mais profundos (6 m a 10 m) (CASAL et al.,2019). Lacaux et al., (2007) ainda afirma que, em locais de águas turvas, a radiação eletromagnética correspondente à luz vermelha pode chegar a ter uma resposta maior se comparada com a luz verde.

Ressalta-se ainda que para uma avaliação coerente de todos os algoritmos com as imagens, e para as duas áreas de estudo, foi definida uma semente, denominada “seed” no software R, para os algoritmos selecionarem as amostras sempre na mesma ordem, ou seja, sempre pegando o mesmo conjunto de dados, além de uma repetição de cinco vezes para a análise.

Para o SVM foi utilizado como método de treinamento o “svmRadial” visto que esse forneceu melhores resultados para a amostra. Autores como Hong et al. (2016) e Harimoorthy e Thangavelu (2021) também empregaram esse método para obtenção de melhores previsões.

Já para os algoritmos GBM e RF, foram utilizados os métodos “gbm” e “rf” respectivamente, visto que é o padrão para fazer a previsão por esses métodos de *Machine Learning*. A RNA contou com uma camada escondida, sete camadas intermediárias, correspondentes ao valor de radiância para cada banda das imagens e uma camada de saída, sendo que essa última consiste na profundidade. Ademais, a RNA foi treinada e avaliada com o pacote “neuralnet”, que emprega a ferramenta do “backpropagation” no processo de aprendizado.

Nesse sentido, para execução de uma avaliação dos resultados, a priori foi efetuado uma análise exploratória das discrepâncias encontradas a partir do valor da profundidade observada daquela estimada com os algoritmos supracitados.

Para posterior análise e comparação de todos os métodos, foram armazenados os valores das discrepâncias abordadas neste trabalho por meio do RMSE, MAE e Incerteza Robusta, sendo esse último proposto por Ferreira (2018) (Equações 6, 7 e 8), bem como o valor do R<sup>2</sup> (coeficiente de correlação) (Equação 9).

$$MAE = \frac{1}{n} \sum_{j=1}^n |z_i - z_{predi}| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (z_i - z_{predi})^2} \quad (7)$$

$$\Phi_{Robusta} = \sqrt{(Q2)^2 + (NMAD)^2} \quad (8)$$

$$R^2 = \frac{\sum_{i=1}^n (z_{predi} - z_{media})^2}{\sum_{i=1}^n (z_i - z_{media})^2} \quad (9)$$

Em que  $z_i$  corresponde aos valores das amostras observadas;  $z_{predi}$ , o valor das amostras preditas;  $z_{media}$ , média dos valores observados e  $n$  o número de amostras.

O RMSE e MAE são indicadores da precisão do valor predito, dessa forma, quando a profundidade estimada e a profundidade predita ou interpolada possuem valores próximos, esses indicadores se aproximaram de zero, ou seja, quanto menor os valores de RMSE e MAE, maior a precisão encontrada pelo método (YANG et al., 2020).

A  $\Phi$ Robusta é calculada por meio de Q2 que corresponde a mediana e o NMAD (*Normalized Mean Absolute Deviation*). Ferreira (2018) propôs esse indicador com o intuito de estimar a incerteza vertical sem o efeito de possíveis *outliers* contidos na amostra. Enfatiza-se que os algoritmos desenvolvidos se encontram no Apêndice A.

Ademais, para espacialização e melhor visualização dos dados, foram gerados modelos digitais de profundidade para a batimetria estimada utilizando o interpolador krigagem simples, no software ArcGIS 10.5.

## 6. RESULTADOS

Em um primeiro momento, as cartas náuticas foram adquiridas no site oficial da Marinha, ambas são disponibilizadas no formato GeoTiff, de forma gratuita pela DHN/CHM. Posteriormente, foram obtidas as imagens orbitais dos sistemas CBERS 4, Sentinel 2A e Landsat 8.

Em seguida, com o treinamento e predição pelos algoritmos de ML, considerando os índices NDWI e NDTI, foi possível obter os valores das discrepâncias entre as profundidades observadas e preditas por cada algoritmo. Assim, foi efetuada uma análise estatística exploratória (Tabelas 2 e 3).

Tabela 2 - Análise exploratória das discrepâncias para o Rio Madeira

Discrepâncias - Rio Madeira								
Satélite	Algoritmo	Média (m)	Erro padrão	Mediana (m)	Desvio padrão (m)	Variância da amostra (m <sup>2</sup> )	Curtose	Assimetria
Sentinel 2A	RF	-0,070	0,064	-0,218	3,739	13,983	2,444	0,485
	GBM	-0,011	0,071	-0,425	4,144	17,173	1,465	0,591
	SVM	0,467	0,071	-0,043	4,146	17,187	2,010	0,840
	RNA	0,019	0,077	-0,280	4,477	20,044	0,524	0,220
CBERS 4	RF	-0,080	0,099	-0,717	5,780	33,404	0,223	0,472
	GBM	-0,090	0,098	-0,599	5,741	32,962	0,203	0,382
	SVM	0,566	0,104	-0,119	6,068	36,821	0,677	0,341
	RNA	-0,057	0,088	-0,546	5,130	26,321	0,418	0,373
Landsat 8	RF	-0,133	0,060	-0,173	3,482	12,124	3,567	0,002
	GBM	-0,081	0,078	-0,383	4,570	20,882	1,086	0,274
	SVM	0,230	0,082	-0,049	4,785	22,893	1,828	0,319
	RNA	-0,055	0,092	-0,750	5,412	29,288	0,217	0,471

Tabela 3 - Análise exploratória das discrepâncias para o Rio Paraguai

Discrepâncias - Rio Paraguai								
Satélite	Algoritmo	Média (m)	Erro padrão	Mediana (m)	Desvio padrão (m)	Variância da amostra (m <sup>2</sup> )	Curtose	Assimetria
Sentinel 2A	RF	-0,013	0,007	0,002	0,386	0,149	9,670	0,394
	GBM	-0,009	0,008	-0,001	0,420	0,176	8,340	0,201
	SVM	0,017	0,008	0,028	0,457	0,209	10,030	-0,252
	RNA	-0,004	0,009	-0,025	0,479	0,229	6,608	0,471
CBERS 4	RF	-0,003	0,009	-0,010	0,480	0,230	6,823	0,180
	GBM	-0,004	0,009	-0,009	0,475	0,225	6,897	0,187
	SVM	0,037	0,009	0,014	0,502	0,252	7,329	0,145
	RNA	0,096	0,009	0,067	0,494	0,244	6,758	0,357
Landsat 8	RF	0,005	0,006	0,000	0,346	0,120	10,039	1,025
	GBM	0,007	0,007	0,000	0,407	0,166	7,073	0,869
	SVM	0,031	0,008	0,014	0,440	0,193	8,281	0,655
	RNA	0,003	0,009	-0,036	0,494	0,244	6,525	0,529

Essa análise evidencia de forma geral, valores mais altos para a variância e para a média das discrepâncias considerando os testes no rio Madeira, para todos os algoritmos. Analisando os resultados para o rio Paraguai, foi possível constatar que o algoritmo Random Forest apresentou uma menor variância e também um menor erro padrão, sobretudo com a imagem Landsat.

Em relação aos valores para a assimetria e curtose, analisando os corpos hídricos de forma separada, nota-se que os valores não se diferenciaram de forma expressiva. No entanto, observa-se que a assimetria apresentou um valor negativo no Rio Paraguai para o Sentinel utilizando o algoritmo SVM. Nesse caso, como uma assimetria negativa indica que a predição da profundidade foi superestimada, esses casos devem ser analisados com mais cautela para produção de cartas náuticas.

De forma a complementar as análises exploratórias, foram confeccionados histogramas das discrepâncias encontradas para os métodos e sensores utilizados, bem como os gráficos de dispersão para os valores preditos e observados. Esses se encontram nos Apêndice B.

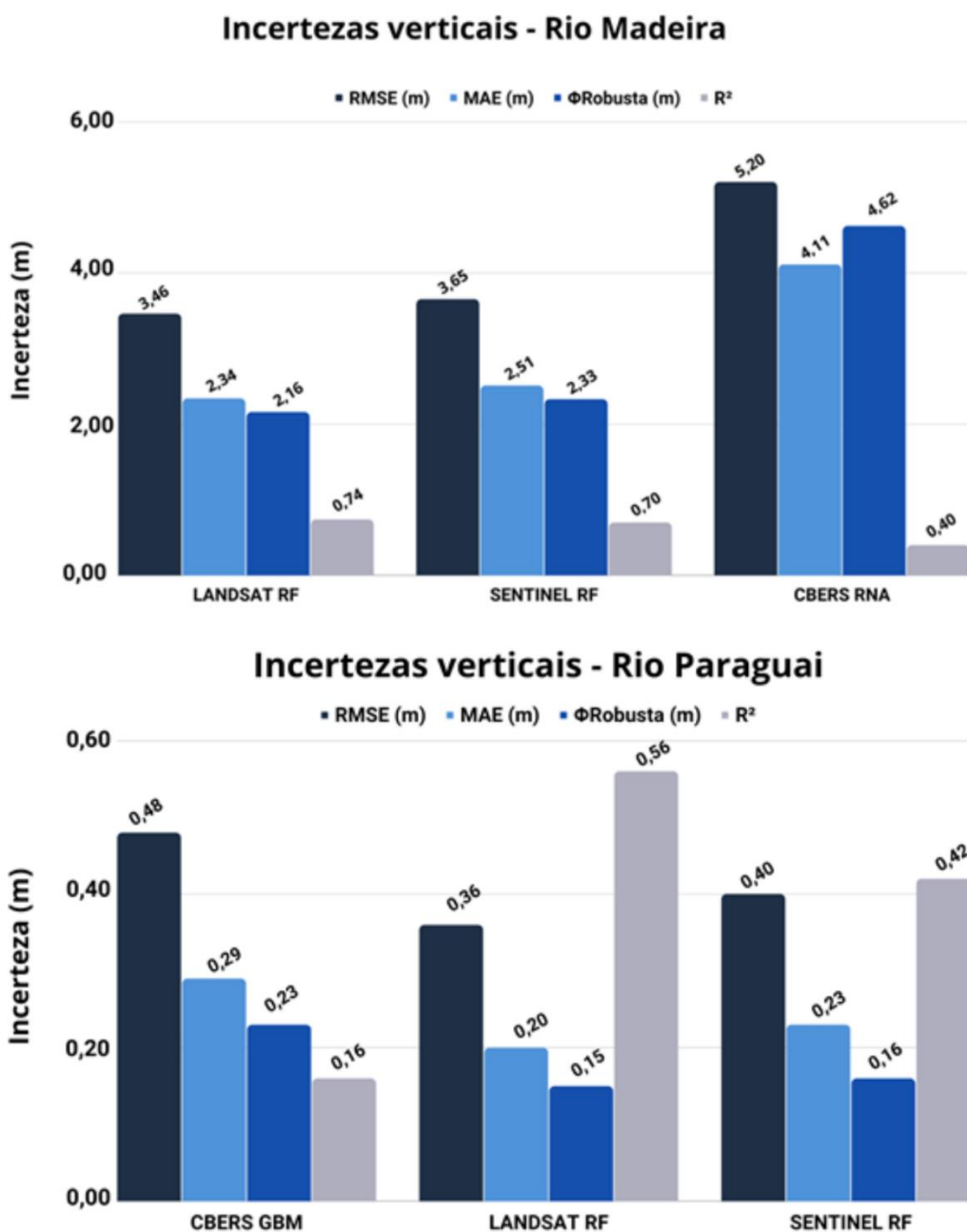
O cálculo da incerteza vertical foi realizado por meio do RMSE, MAE e  $\Phi$ Robusta, além do  $R^2$  para avaliação da correlação. Após as cinco repetições do algoritmo, foi calculado a média para os estimadores (Tabela 4).

Tabela 4 - Valores obtidos para RMSE, MAE,  $\Phi$ Robusta e  $R^2$  do SBD para o rio Madeira e Paraguai

	Rio Madeira				Rio Paraguai			
	RMSE	MAE	$\Phi$ Robusta	$R^2$	RMSE	MAE	$\Phi$ Robusta	$R^2$
<b>CBERS GBM</b>	5,66	4,40	4,93	0,28	0,48	0,29	0,23	0,16
<b>CBERS RF</b>	5,68	4,41	5,01	0,28	0,49	0,29	0,22	0,14
<b>CBERS SVM</b>	5,98	4,38	4,42	0,23	0,51	0,30	0,19	0,07
<b>CBERS RNA</b>	5,20	4,11	4,62	0,40	0,50	0,31	0,27	0,10
<b>LANDSAT GBM</b>	4,48	3,35	3,60	0,56	0,41	0,25	0,20	0,40
<b>LANDSAT RF</b>	3,46	2,34	2,16	0,74	0,36	0,20	0,15	0,56
<b>LANDSAT SVM</b>	4,72	3,29	3,06	0,51	0,45	0,26	0,19	0,30
<b>LANDSAT RNA</b>	5,25	4,13	4,89	0,38	0,50	0,32	0,28	0,12
<b>SENTINEL GBM</b>	4,06	2,98	3,16	0,63	0,43	0,25	0,19	0,34
<b>SENTINEL RF</b>	3,65	2,51	2,33	0,70	0,40	0,23	0,16	0,42
<b>SENTINEL SVM</b>	4,12	2,84	2,68	0,63	0,47	0,27	0,19	0,20
<b>SENTINEL RNA</b>	4,85	3,79	4,48	0,47	0,49	0,31	0,29	0,14

A Figura 7 evidencia os melhores desempenhos obtidos para cada imagem e área de estudo.

Figura 7 - Gráficos do RMSE, MAE e  $R^2$  para o SDB no rio Madeira e Paraguai



Avaliando as duas áreas de estudo em conjunto, é notória a maior incerteza vertical para o rio Madeira, visto que os valores para essa área foram 10x (RMSE), 12x (MAE) e 17x ( $\Phi$ Robusta) superiores aos valores do rio Paraguai.

Para o rio Madeira, a utilização da imagem CBERS com o algoritmo SVM, forneceu o valor mais alto para RMSE, sendo este de 5,98 m. No caso do MAE, a mesma imagem, utilizando o RF, alcançou uma incerteza de 4,41 m e uma  $\Phi$ Robusta de 5,01 m. Para o rio

Paraguai, assim como afirmado anteriormente para o Rio Madeira, a imagem CBERS, em conjunto com o algoritmo SVM, forneceu o valor mais alto do RMSE. Já para o MAE, a imagem Landsat com a RNA apresentou uma incerteza de 0,32 m. Com o  $\Phi$ Robusta, a imagem CBERS 4 com o SVM chegou a uma incerteza de 4,42 m.

Ressalta-se ainda, que o coeficiente de correlação encontrado para todas imagens e algoritmos foi relativamente baixo. O maior valor para esse parâmetro foi para Landsat 8 com a utilização do RF, de 0,74, seguido do valor de 0,70 para a imagem Sentinel 2A, com o mesmo algoritmo no rio Madeira. O RF com a Landsat 8 também forneceu os valores mais satisfatórios de RMSE, MAE e  $\Phi$ Robusta para o rio Madeira e Paraguai. Destaca-se ainda, que os valores de  $R^2$  observados para o rio Paraguai foram inferiores aos do rio Madeira (Tabela 4).

No que tange a imagem CBERS 4 para o rio Paraguai, é importante destacar o fato de que o menor RMSE foi encontrado com a utilização do GBM, já o menor MAE, com o RF e, com o SVM, a menor  $\Phi$ Robusta.

Uma análise das discrepâncias em relação a profundidade também foi realizada com o objetivo de avaliar de forma distinta os algoritmos empregados e as imagens com as duas áreas de estudo (Tabela 5 e Tabela 6).

Tabela 5 - Análise das discrepâncias por profundidade – rio Madeira

<b>Rio Madeira - Média das discrepâncias (m)</b>				
<b>Profundidade (m)</b>	<b>Landsat 8</b>			
	<b>RNA</b>	<b>SVM</b>	<b>RF</b>	<b>GBM</b>
0,00 - 5,00	-3,33	-2,17	-1,89	-2,76
5,00 - 15,00	-0,17	0,10	-0,07	0,03
15,00 -25,00	6,95	5,80	3,43	5,50
25,00 – 32,28	14,31	12,76	6,38	11,64
<b>Profundidade (m)</b>	<b>Sentinel 2A</b>			
	<b>RNA</b>	<b>SVM</b>	<b>RF</b>	<b>GBM</b>
0,00 - 5,00	-2,76	-1,40	-1,71	-2,04
5,00 - 15,00	-0,15	0,34	-0,18	-0,10
15,00 -25,00	6,04	4,76	3,80	4,55
25,00 – 32,28	11,86	9,67	6,87	9,30
<b>Profundidade (m)</b>	<b>CBERS 4</b>			
	<b>RNA</b>	<b>SVM</b>	<b>RF</b>	<b>GBM</b>
0,00 - 5,00	-3,14	-3,22	-4,15	-4,05
5,00 - 15,00	-0,14	0,85	0,17	0,24
15,00 -25,00	6,42	8,58	8,01	7,78
25,00 – 32,28	12,20	13,65	14,93	14,19

Na Tabela 5, observa-se que as discrepâncias foram maiores nas maiores profundidades (25,00 a 32,28 m), chegando a 14,93 m para o RF com a CBERS 4. Quanto as menores profundidades (0,00 a 5,00 m), o SVM com a imagem Sentinel 2A foi o que apresentou o melhor desempenho, seguido do RF com o mesmo sistema e também o RF com a Landsat 8. Avaliando as profundidades médias para o rio Madeira (5,00 a 15,00 m), a imagem Landsat 8 com o GBM evidenciou uma discrepância de 0,03 m, seguido do RF com -0,07 e o SVM de 0,10 m. Para as profundidades de 15,00 a 25,00 m, o RF com a Landsat 8 foi significativamente inferior que as demais, sendo a Sentinel 2A, com o mesmo algoritmo, a mais próxima de 3,80m.

Tabela 6 - Análise das discrepâncias por profundidade – rio Paraguai

<b>Rio Paraguai - Média das discrepâncias (m)</b>				
<b>Profundidade (m)</b>	<b>Landsat 8</b>			
	<b>RNA</b>	<b>SVM</b>	<b>RF</b>	<b>GBM</b>
0,00 -2,50	-1,06	-0,92	-0,56	-0,77
2,50 - 5,00	0,05	0,07	0,02	0,03
5,00 – 7,87	2,31	2,19	1,68	1,95
<b>Profundidade (m)</b>	<b>Sentinel 2A</b>			
	<b>RNA</b>	<b>SVM</b>	<b>RF</b>	<b>GBM</b>
0,00 -2,50	-1,02	-1,04	-0,80	-0,90
2,50 - 5,00	0,04	0,06	0,02	0,03
5,00 – 7,87	2,26	2,20	1,86	1,98
<b>Profundidade (m)</b>	<b>CBERS 4</b>			
	<b>RNA</b>	<b>SVM</b>	<b>RF</b>	<b>GBM</b>
0,00 -2,50	-1,00	-1,10	-1,07	-1,02
2,50 - 5,00	0,15	0,09	0,04	0,04
5,00 – 7,87	2,36	2,30	2,13	2,10

Para o rio Paraguai, a Tabela 6 demonstra que o menor valor para as discrepâncias está com a Landsat 8 em conjunto com o RF. Nas menores profundidades, para Landsat 8, esse algoritmo apresentou discrepâncias inferiores que os demais, com diferenças de até 0,50 m. Para CBERS 4 (até 2,50 m), a RNA mostrou-se com uma discrepância média inferior que as demais, com até 0,10 m de diferença para o SVM. No entanto, em profundidades médias (até 5,00 m), o GBM e RF, com a imagem CBERS 4, apresentaram uma discrepância média de 0,04 m, inferior os demais algoritmos com essa mesma imagem, tal como para as profundidades até 7,87 m.

Destaca-se, ainda, o fato das discrepâncias com a utilização do GBM, para imagens Sentinel 2A e Landsat 8, se mostrarem próximas, com menos de 0,30 m de diferença para as menores profundidades.

Posteriormente, para gerar os modelos digitais (Figura 8 e Figura 9), em um primeiro instante, foi analisada a tendência nos dados para utilização da krigagem universal (FERREIRA et al. 2013). Visto que não foi observada tendência, como indicado por Santos et al. (2011), foi empregada a krigagem simples.

Figura 8 - MDPs para o rio Madeira

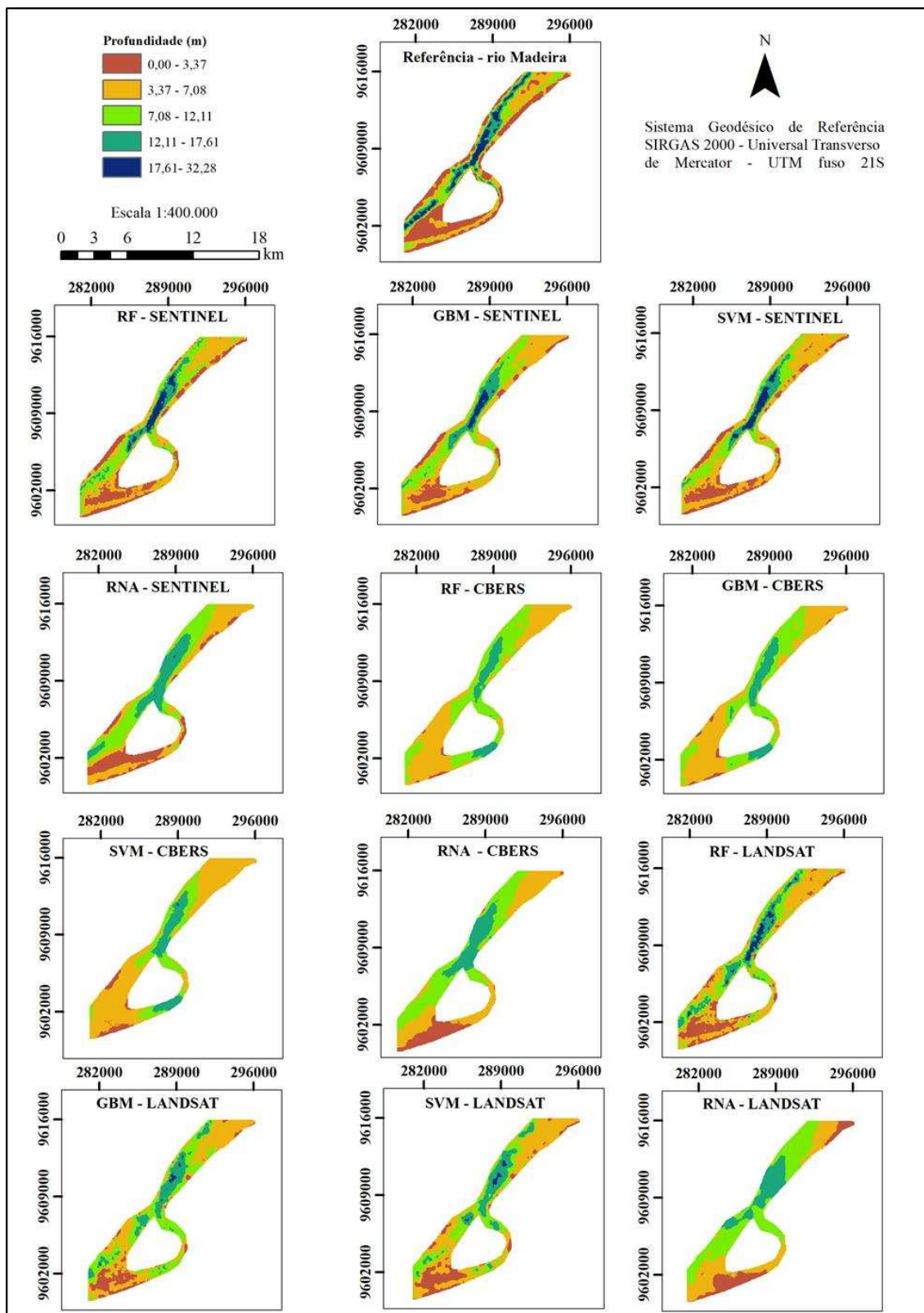
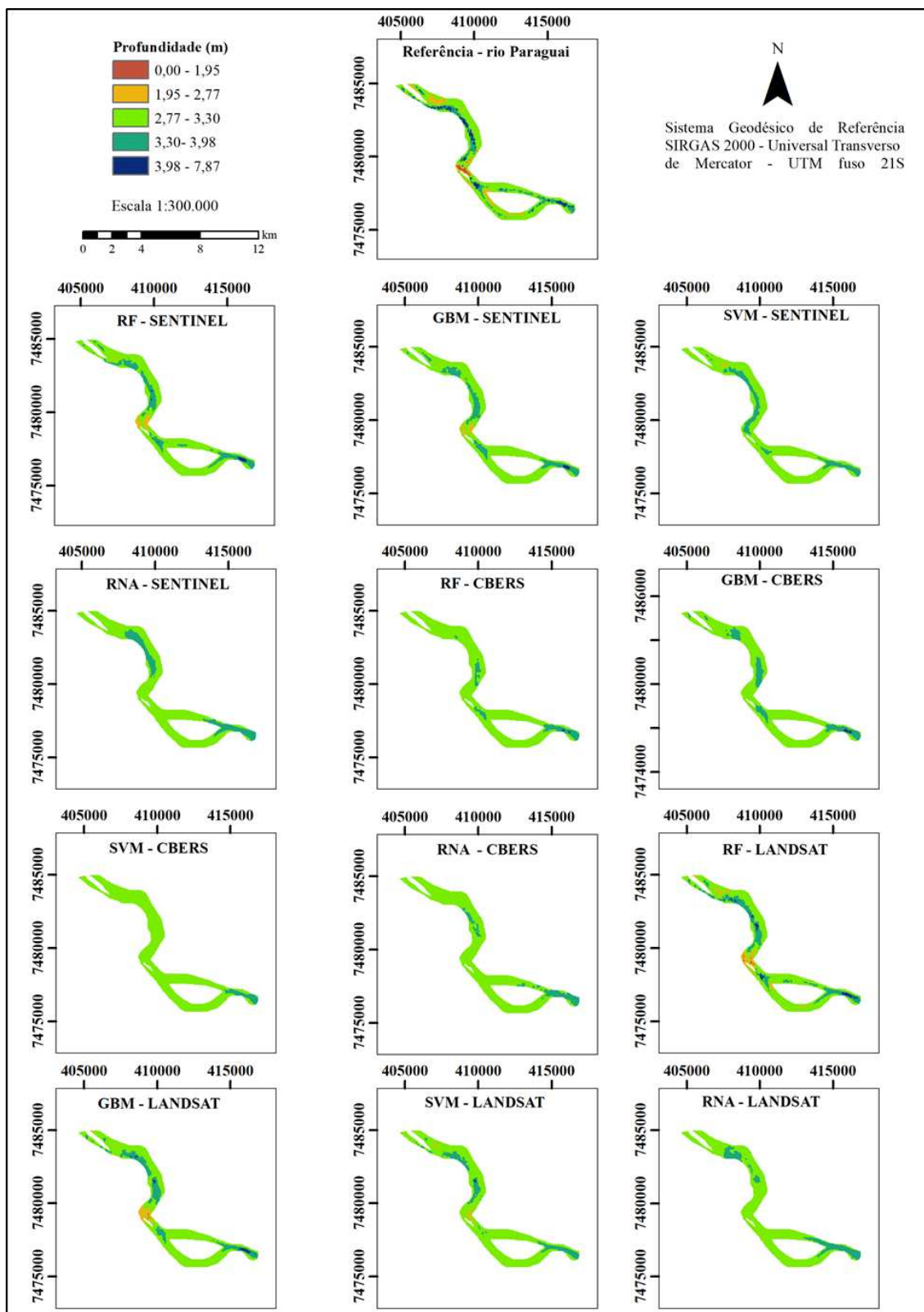


Figura 9 - MDPs para o rio Paraguai



Por meio da análise dos MDPs, de forma geral, observa-se uma maior diferença na profundidade com o emprego da imagem CBERS 4 para ambos rios.

Em relação ao rio Madeira, para a imagem Sentinel 2, os algoritmos RF e SVM foram superiores aos demais. Já para a Landsat 8, além do RF, o GBM apresentou um modelo próximo ao de referência. Para a CBERS 4, a RNA apresentou uma menor diferença nas profundidades mais rasas, se comparado com os demais algoritmos.

Para o rio Paraguai, o CBERS 4 com o GBM foi o que apresentou uma maior semelhança com a morfologia submersa de referência. Para o Sentinel 2, tal como para o Landsat 8, o RF foi o algoritmo que melhor representou a profundidade do corpo hídrico.

Ademais, também foram confeccionados os modelos digitais das discrepâncias dos algoritmos que apresentaram melhores valores para os parâmetros analisados na Tabela 4, para cada imagem, considerando as duas áreas de estudo. Em todos os modelos, foi executado a subtração do modelo de referência em relação ao estudado (Apêndice C).

## **7.1 DISCUSSÕES**

### **7.1. Parâmetros da incerteza vertical**

Analisando os valores encontrados para os parâmetros da incerteza vertical nas duas áreas de estudo, as imagens Landsat 8 e Sentinel 2A forneceram os melhores valores para os parâmetros avaliados, principalmente com o emprego do algoritmo RF (Tabela 4). Tal como apresentado por Hedley et al. (2018), os quais usaram estas imagens para estimar a profundidade da costa da Austrália, em locais de até 30,00 m e obtiveram uma boa estimativa de até 15 m, com um  $R^2$  de 0,89. O estudo em questão evidenciou um  $R^2$  semelhante para o rio Madeira, visto que esse possui condições semelhantes ao local estudado, diferentemente do rio Paraguai, o qual possui uma profundidade máxima de aproximadamente 8,00 metros.

Moeinkhan et al., (2019) e Sagawa et al. (2019) também utilizaram de imagens Landsat e encontraram um RMSE de aproximadamente 1,25 m para profundidades de até 20 m. Essas pesquisas demonstram aderência aos resultados obtidos nesse trabalho, dado que, para o rio Madeira (com uma profundidade máxima de aproximadamente 30,00 m), foi encontrado um RMSE de 3,46 m, posto que esse corpo hídrico possui um maior nível de sedimentos em suspensão e uma coloração mais turva. Já para o rio Paraguai, o RMSE foi de 0,36 m, sendo que esse é menos profundo e com água mais límpida, sem a presença notória de plantas aquáticas.

Além disso, os resultados encontrados para o parâmetro  $\Phi$ Robusta foram de 2,16 m para o rio Madeira e 0,15 m para o rio Paraguai, considerando a imagem Landsat 8 com o RF. O valor encontrado para o rio Paraguai é inferior ao evidenciado na literatura, em que foi alcançado um  $\Phi$ Robusta de 0,73 metros para um reservatório com profundidade de 4,00 m na cidade de Viçosa – MG, utilizando imagens de RPA (ANDRADE et al., 2021). A diferença nesses valores é justificada no uso do NDWI como método para predição da batimetria e também no fato de que o reservatório possui muitos sedimentos em suspensão e plantas aquáticas. Ressalta-se que não existem estudos utilizando o parâmetro da  $\Phi$ Robusta para imagens Landsat 8 e Sentinel 2A.

Em relação aos valores do coeficiente de correlação ( $R^2$ ) obtidos, evidencia-se na Tabela 4 que, para o rio Paraguai, esses foram significativamente inferiores se comparados ao rio Madeira. Essa ocorrência pode ser devido ao número de pontos que foram empregados para treinamento e teste nos algoritmos. Como existiam mais pontos no rio Madeira, a correlação linear para esse foi melhor, ou seja, apresentou uma menor dispersão se comparado com o rio Paraguai, com o mesmo algoritmo e mesma imagem orbital. Autores como Najar et al. (2022) também apresentaram um  $R^2$  de 0,71 para uma área de estudo com profundidade de até 50,00 m, empregando imagens Sentinel 2A. Sagawa et al. (2019) obteve menores  $R^2$  quando utilizou menos pontos para o treinamento da predição com o RF e imagens Landsat 8.

Comparando os algoritmos em relação a eficácia de predição da profundidade, é possível afirmar que o RF se destacou em relação aos demais, principalmente empregando as imagens Landsat 8 e Sentinel 2A. Tal fato também foi constatado por Zhou et al. (2023), que avaliou os algoritmos RF, SVM, RNA e o método de Stumpf na predição da batimetria empregando imagens Sentinel 2A, encontrando menores RMSE e MAE para o RF, de 0,82 m e 0,55 m, respectivamente. Dessa forma, pode-se afirmar que o RF foi capaz de criar modelos mais flexíveis e acurados baseados nos dados que foram fornecidos pelo modelo, com baixo viés se comparado aos demais algoritmos empregados.

O GBM, de forma semelhante ao RF, consiste em um algoritmo de árvore de decisão. No entanto, esse se utiliza de árvores sequenciais e pequenas que tendem a ter um maior viés, se comparado ao RF (AYYADEVARA, 2018). Analisando o estudo em questão, o GBM apresentou o segundo melhor valor para RMSE, MAE e  $R^2$ . Além disso, com a imagem CBERS 4 no rio Paraguai, o algoritmo alcançou um RMSE de 0,48 m, sendo superior que o RF. Alguns estudos como o de Abdul et al. (2022) mostram que, o emprego do GBM pode fornecer uma melhoria na predição da batimetria de até 30% se comparado a modelos lineares de regressão.

Para o rio Madeira, com o emprego da imagem CBERS 4, o algoritmo que apresentou os melhores valores para os parâmetros da incerteza vertical, foi a RNA, com um RMSE de 5,20 m; MAE de 4,11 m;  $\Phi$ Robusta de 4,62 m, e  $R^2$  de 0,40, evidenciando uma menor dispersão na predição dos dados. Andrade et al. (2022) também utilizou a RNA em um reservatório de acumulação de água e obteve uma incerteza vertical de 0,73 m para a  $\Phi$ Robusta. No entanto, diferente do rio Madeira, esse reservatório possuía uma profundidade máxima de aproximadamente 4,00 m. É importante enfatizar que o elevado número de informações fornecidas para o algoritmo, como as bandas existentes em cada satélite e os índices NDTI e NDWI, podem ter dificultado a interpretação da RNA, além da ocorrência do excesso de ajuste, ou *overfitting* (KROGH, 2008), o que acarretou em uma menor precisão na estimativa da profundidade para imagens Landsat 8 e Sentinel 2A.

Para o SVM, os valores de RMSE, MAE,  $\Phi$ Robusta e  $R^2$  com imagens Landsat 8 e Sentinel 2A, se aproximaram daqueles obtidos com o GBM, para as duas áreas de estudo. Autores como Misra et al. (2018), afirmam que para águas turvas, o SVM pode fornecer um  $R^2$  de 0,73, semelhante ao estudo em questão com imagens Landsat 8 no rio Madeira, em que esse algoritmo foi capaz de obter um  $R^2$  de 0,51. No entanto, para imagens CBERS 4 esse algoritmo apresentou uma diferença superior a 30 cm no rio Madeira e de 3 cm no rio Paraguai, muito provavelmente devido à baixa resolução radiométrica da imagem, associada a elevada quantidade de sedimentos em suspensão, o que levou a uma limitação na interpretação dos valores de radiância pelo SVM (TONG et al. 2009).

## 7.2. Intervalos de profundidade

Em locais de baixas profundidades (< 5,00 m para o rio Madeira e <2,50 m para o rio Paraguai), de modo geral, as discrepâncias foram maiores do que para os locais de médias profundidades (< 5,00 m para o rio Paraguai e < 15,00 m para o rio Madeira). Esse acontecimento é consequência do espalhamento por flutuações de densidade e por partículas, uma vez que em menores profundidades podem ocorrer movimentos moleculares aleatórios das partículas em um dado volume de água, causando flutuações no índice de refração ao longo da trajetória da luz, além da interação dos fótons de luz com as matérias orgânicas e inorgânicas em suspensão no meio aquático, que implicam na alteração da direção do feixe de luz (BARBOSA et al. 2019). Autores como Su et al. (2008) também observaram essa situação utilizando algoritmos não lineares com imagens IKONOS na Ilha Molokai, no estado do Havaí nos Estados Unidos. Casal et al. (2018) encontraram valores maiores para as incertezas em

profundidades de até 6,00 m quando comparadas às profundidades de até 10,00 m na baía de Dublin na Irlanda.

Para as profundidades médias ( $< 15,00$  m para o rio Madeira e  $< 5,00$  m para o rio Paraguai) e altas ( $< 32,28$  m para o rio Madeira e  $< 7,78$  m para o rio Paraguai) os valores das discrepâncias aumentaram com o aumento da profundidade, como apontado também por Chénier et al. (2018) e Zhou et al. (2023).

Ademais, os resultados da análise por profundidade foram semelhantes a aqueles obtidos com os parâmetros da incerteza vertical, em que o RF com a imagem Landsat 8 forneceu os melhores valores de forma geral, comprovando sua superioridade em relação aos demais para todas as profundidades avaliadas. Destaca-se que, para as imagens CBERS 4, no rio Madeira a RNA mostrou-se mais eficaz, já no rio Paraguai, o GBM apresentou os menores valores para as discrepâncias em quase todos intervalos de profundidade.

### 7.3. Modelos Digitais de Profundidade

A construção dos modelos digitais mostrados nas Figuras 8 e 9 proporciona uma análise da predição da batimetria, dado que a avaliação da incerteza vertical com os parâmetros, e também por meio dos intervalos de profundidade, possibilitam um diagnóstico quantitativo. Ferreira et al. (2016b), Andrade et al. (2021) e Andrade et al. (2022) utilizaram de modelos digitais de profundidade e discrepância para análise da predição da batimetria utilizando imagens RapidEye e de RPA's, respectivamente, o que permitiu evidenciar de forma mais simples as áreas mais propícias a erros.

Nos MDP's para as duas áreas de estudo, evidencia-se a semelhança daquele criado com o algoritmo RF e a imagem Landsat 8 com o MDP de referência, como apontado nas análises quantitativas. Autores como Elshazly et al. (2021) e Zhou et al. (2023) também constataram a superioridade desse algoritmo com imagens Landsat 8 e Sentinel 2A, respectivamente. Além disso, o RF com imagens Sentinel 2A também forneceu um modelo similar ao gerado com os dados de sonar.

O algoritmo GBM, para imagens Landsat 8 e Sentinel 2A no rio Paraguai, mostrou mais conformidade com a referência quando comparado com o SVM e a RNA. No entanto, com imagens CBERS 4 no rio Madeira, a RNA foi capaz de produzir um MDP mais semelhante ao de referência quando comparado com outros algoritmos. Acredita-se que tal ocorrência deve-se ao fato de que as imagens CBERS 4 possuem menos bandas e, por conseguinte, menos informação a ser interpretada pela RNA, facilitando o processo de predição, sem a probabilidade de causar *overfitting* (KROGH, 2008).

## 8. CONCLUSÕES

A partir dos estudos realizados, pode-se concluir, de forma geral, que o emprego de imagens Landsat 8, associadas ao algoritmo RF, pode oferecer melhores resultados para águas turvas e também para águas mais límpidas. Sobretudo quando comparado com imagens com uma menor resolução radiométrica, como a CBERS 4 (MUX) e Sentinel 2A. Além disso, foi possível evidenciar que para a determinação da batimetria derivada de satélite, não basta apenas a avaliação da resolução espacial, essa deve ser analisada em conjunto com outros diversos fatores, como por exemplo, a própria resolução radiométrica.

É importante destacar que a imagem CBERS 4 forneceu resultados interessantes para o rio Paraguai, local que apresenta menor quantidade de sedimentos em suspensão e com água mais límpida. O RMSE de 0,48 m com o GBM, associado aos valores de 0,29 para MAE, 0,23 para o  $\Phi$ Robusta, indica que esse satélite tem potencial para ser utilizado no SDB, visto que os valores dos parâmetros foram próximos daqueles considerados mais robustos. Nesse sentido, é importante realizar a difusão da utilização do sistema CBERS nesse âmbito, uma vez que não possuem estudos que tem como finalidade a estimativa da batimetria com essas imagens.

Outra conclusão importante, refere-se à análise dos algoritmos de forma separada, em que se pode perceber certa similaridade nos valores obtidos com o RF e GBM, especialmente no rio Paraguai. Com a imagem Sentinel 2A, a diferença entre esses dois foi menor que 0,30 m. Já para imagem Landsat 8, essa diferença foi menor que 0,50 m, considerando o parâmetro RMSE, indicando uma potencialidade de utilização do GBM para a determinação da batimetria derivada de satélite.

A utilização de cartas náuticas para o SDB configura-se como uma inovação no âmbito, sobretudo na esfera nacional. A oportunidade de desenvolver trabalhos com tecnologias mais avançadas, que não exigem de muitos gastos quando comparado ao levantamento batimétrico tradicional, possibilita várias descobertas que podem facilitar a determinação da profundidade de inúmeros locais. Projetos como o Seabed 2030, que buscam o mapeamento completo de todo oceano até o ano de 2030, evidenciam os avanços na área do sensoriamento remoto associado a batimetria.

Para trabalhos futuros, recomenda-se o emprego e avaliação de outras bandas dos sistemas com resolução espacial distinta, como a Banda 1 do Sentinel 2A, com a utilização de bandas com menores resoluções espaciais, como a PAN do CBERS 4 e Landsat 8. Ademais, também poderá ser analisado outras formas de correção atmosférica para essas imagens, com o emprego de técnicas mais robustas que exigem um maior esforço computacional. Novos testes

com imagens CBERS 4A, que possuem uma melhor resolução espacial também se configuram como uma análise importante a ser efetuada.

Enfatiza-se ainda, a possibilidade de efetuar otimizações nos algoritmos utilizados, como por exemplo, a RNA, em que podem ser testados diferentes números de camadas escondidas, repetições e métodos de validação. Por outra perspectiva, recomenda-se também a automatização do processo para estimativa da profundidade em um software de SIG (Sistemas de Informações Geográficas), visto que as técnicas abordadas no SDB não são de simples aplicação e necessitam de vários softwares para a concretização.

## REFERÊNCIAS

- ABDUL G. F.; AL-SHEHHI, M. R.; CHO, C. S.; GHEDIRA, H. Gradient Boosting and Linear Regression for Estimating Coastal Bathymetry Based on Sentinel-2 Images. *Remote Sensing*, 14(19), 5037.2022.
- AGRAFIOTIS, P.; SKARLATOS, D.; GEORGOPOULOS, A.; KARANTZALOS, K. Shallow water bathymetry mapping from uav imagery based on machine learning. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII. 2019.
- ANDRADE, L.C.; FERREIRA, Í.O.; MEDEIROS, N.G.; TEIXEIRA, V.G.; SANTOS, F.C.M. Avaliação de imagens obtidas com o sensor Micasense na estimativa da batimetria de corpos d'água opticamente rasos. *Revista Brasileira de Cartografia*. 2021.
- ANDRADE, L. C.; FERREIRA, I. O.; SILVA, A. A.; GIBRIM, V. T.; SANTOS, F. C. M. On the use of artificial neural networks in remotely piloted aircraft acquired images for estimating reservoir's bathymetry. *Boletim de Ciências Geodésicas*, 28.2022.
- ASHPHAQ, M.; SRIVASTAVA, P.K.; MITRA, D. Review of near-shore satellite derived bathymetry: Classification and account of five decades of coastal bathymetry research. *Journal of Ocean Engineering and Science*, v. 6, n. 4, p. 340-359, 2021.
- AYYADEVARA, V. K. Gradient boosting machine. In: *Pro machine learning algorithms*. Apress, Berkeley, CA, 2018. p. 117-134.
- BARBOSA, C. C. F.; NOVO, E. M. L.M.; MARTINS, V. S. Introdução ao sensoriamento remoto de sistemas aquáticos: princípios e aplicações (Vol. 1). Instituto Nacional de Pesquisas Espaciais.2019.
- BERNARDI, J. V. E. et al. Aplicação da análise das componentes principais na ordenação dos parâmetros físico-químicos no alto Rio Madeira e afluentes, Amazônia Ocidental. *Geochimica Brasiliensis*, v. 23, n. 1, p. 79-90, 2012.
- BIRKETT, C.M. The contribution of TOPEX/POSEIDON to the global monitoring of climatically sensitive lakes. *Journal of Geophysical Research-Oceans*, 100.1995. pp.25179-25204.
- BREIMAN, L. Random forests. *Machine Learning*. V.45.1, p.27,2001.
- CAHALANE, C.; MAGEE, A.; MONTEYS, X.; CASAL, G.; HANAFIN, J.; HARRIS, P. A comparison of Landsat 8, RapidEye and Pleiades products for improving empirical predictions of satellite-derived bathymetry. *Remote sensing of environment*, v. 233, p. 111414, 2019.
- CASAL, G.; MONTEYS, X.; HEDLEY, J.; HARRIS, P.; CAHALANE, C.; MCCARTHY, T. Assessment of empirical algorithms for bathymetry extraction using Sentinel-2 data. *International journal of remote sensing*, 40(8), 2855-2879.2019.
- CASAL, G., HEDLEY, J. D., MONTEYS, X., HARRIS, P., CAHALANE, C., & MCCARTHY, T. Satellite-derived bathymetry in optically complex waters using a model inversion approach and Sentinel-2 data. *Estuarine, Coastal and Shelf Science*, 241, 106814.2020.

- CEYHUN, O.; YALÇIN, A. "Remote sensing of water depths in shallow waters via artificial neural networks," *Estuarine, Coastal Shelf Sci.*, vol. 89, no. 1, pp. 89–96, Sep. 2010, doi: 10.1016/j.ecss.2010.05.015.
- CHAVEZ, P.S.; JR. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* 1988, 24, 459–479.
- CHÉNIER, R.; FAUCHER, M.; AHOLA, R. Satellite-derived bathymetry for improving Canadian Hydrographic Service charts. *ISPRS International Journal of Geo-Information*, v. 7, n. 8, p. 306, 2018.
- COLLISCHONN, B.; CLARKE, R. T. Estimativa e incerteza de curvas cota-volume por meio de sensoriamento remoto. *RBRH*, 21, 719-727.2016.
- CONGEDO, L. "Semi-automatic classification plugin documentation: release 4.8.0.1", 2016. Disponível em: <<https://media.readthedocs.org/pdf/semiautomaticclassificationmanual-v5/latest/semiautomaticclassificationmanual-v5.pdf>>.
- CRÉTAUX, J.F.; BIRKETT, C. Lake studies from satellite radar altimetry. *Compts Rendus Geosciences*, 338. 2006. pp.1098-1112.
- CRÉTAUX, J.F.; JELINKSKI, W. ; CALMANT, S. KOURAEV; A. VUGLINSKI, V. BERGENGUYEN, M. et al. SOLS: A lake database to monitor in the near real time water level and storage variations from remote sensing data. *Advances in Space Research*, 47. 2011. pp.1497-1507.
- DUAN, K.; KEERTHI, S. S. Which is the best multiclass SVM method? An empirical study. In: *International workshop on multiple classifier systems*. Springer, Berlin, Heidelberg, 2005. p. 278-285.
- EASTMAN, J.R. *IDRISI Selva - GIS and Image Processing Software 17.2*. Worcester-MA/USA: Clark Labs, 2012.
- ELSHAZLY, R. E.; ARMANUOS, A. M.; ZEIDAN, B. A.; ELSHEMY, M. Evaluating remote sensing approaches for mapping the bathymetry of Lake Manzala, Egypt. *Euro-Mediterranean Journal for Environmental Integration*, 6(3), 1-12.2021.
- ESTIGON, M. V.; MATOS, A. J. S.; MAUAD, F. F. Análise da orientação da ANA para a atualização das curvas cota-área-volume de reservatórios. *X Encontro Nacional de Engenharia Sanitária*, Foz do Iguaçu, PR, Brasil. Anais... Foz do Iguaçu, PR, Brasil, 3.2015.
- FACETTI, J.; DEKOV, V. M.; VAN GRIEKEN, R. Heavy metals in sediments from the Paraguay River: a preliminary study. *Science of the total environment*, v. 209, n. 1, p. 79-86, 1998.
- FERREIRA, Í. O.; SANTOS, G. R. RODRIGUES, D. D. Estudo sobre a utilização adequada da krigagem na representação computacional de superfícies batimétricas. *Revista Brasileira de Cartografia*, Rio de Janeiro, v. 65, n.5, p. 831-842, 2013.
- FERREIRA, Í. O.; RODRIGUES, D. D.; SANTOS, G. R. Coleta, processamento e análise de dados batimétricos. 1ª ed. Saarbrücken: Novas Edições Acadêmicas, v. 1, 100f. Res. 27, 7–17. 2015. doi: 10.1007/s11001-005-0267-x

FERREIRA, I. O.; NETO, A. A.; MONTEIRO, C. S. O uso de embarcações não tripuladas em levantamentos batimétricos. *Revista Brasileira de Cartografia*, v. 68, n. 10, 2016a.

FERREIRA, I. O.; ZANETTI, J.; GRIPP, J. S.; MEDEIROS, N.G. Viabilidade do uso de imagens do sistema Rapideye na determinação da batimetria de águas rasas. *Revista Brasileira de Cartografia*, 68(7).2016b.

FERREIRA, I. O. Controle de qualidade em levantamentos hidrográficos. Tese de Doutorado. Universidade Federal de Viçosa. Programa de Pós Graduação em Engenharia Civil. Viçosa. Brasil. 2018.

FERREIRA, I.O., ANDRADE, L.C., TEIXEIRA, V.G., SANTOS, F.C.M. State of art of Bathymetric surveys. *Boletim de Ciências Geodésicas*, 28. 2022.

GABR, B.; AHMED, M.; MARMOUSH, Y. PlanetScope and Landsat 8 Imageries for Bathymetry Mapping. *Journal of Marine Science and Engineering*, 8(143), 2–17.2020.

GAO, J. Bathymetric mapping by means of remote sensing: methods, accuracy and limitations. *Physical Geography*, v. 33, n. 1, p. 103-116, 2009.

GIORDANO, F.; MATTEI, G.; PARENTE, C.; PELUSO, F.; SANTAMARIA, R. Integrating Sensors Into. A Marine Drone For Bathymetric 3d Surveys In Shallow Waters. *Sensors*, V. 16, N. 1, P. 41, 2015.

GOMES, D. C.; FERREIRA, N. S.; LIMA, A. M. M.; DE OLIVEIRA SERRÃO, E. A. Variabilidade espaço-temporal da precipitação: Bacia Hidrográfica do rio Madeira. *Revista Geográfica Acadêmica*, 13(1), 90-104. 2019.

HARIMOORTHY, K.; THANGAVELU, M. Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system. *Journal of Ambient Intelligence and Humanized Computing*, v. 12, n. 3, p. 3715-3723, 2021.

HEDLEY, J. D., ROELFSEMA, C., BRANDO, V., GIARDINO, C., KUTSER, T., PHINN, S., ... KOETZ, B. Remote Sensing of Environment Coral reef applications of Sentinel-2: Coverage , characteristics , bathymetry and benthic mapping with comparison to Landsat 8. *Remote Sensing of Environment*, 216(April), 598–614.2018. <https://doi.org/10.1016/j.rse.2018.07.014>

HINTON, G.; SEJNOWSKI, T. Unsupervised Learning: Foundations of Neural Computation. MIT Press. 398 p. 1999.

HONG, H.; PRADHAN, B.; JEBUR, M. N.; BUI, D. T.; XU, C. & AKGUN, A. Spatial prediction of landslide hazard at the Luxi area (China) using support vector machines. *Environmental Earth Sciences*, 75(1), 1-14.2016.

HU, J.; NIU, H.; CARRASCO, J.; LENNOX, B.; ARVIN, F. Voronoi-Based Multi-Robot Autonomous Exploration in Unknown Environments via Deep Reinforcement Learning. *IEEE Transactions on Vehicular Technology*. v. 69. n. 12. p. 14413-14423. Dez. 2020.

ILORI, C. O.; KNUDBY, A. An approach to minimize atmospheric correction error and improve physics-based satellite-derived bathymetry in a coastal environment. *Remote Sensing*, v. 12, n. 17, p. 2752, 2020.

INPE – Instituto Nacional de Pesquisas Espaciais. Câmeras Imageadoras CBERS - 3 e 4. 2014. Disponível em: [INPE/CBERS](http://INPE/CBERS). Acesso em 29 de setembro de 2021.

JAIN, A. K.; MAO, J.; MOHIUDDIN, K. M. Artificial neural networks: A tutorial. *Computer*, v. 29, n. 3, p. 31-44, 1996.

KROGH, A. What are artificial neural networks?. *Nature biotechnology*, v. 26, n. 2, p. 195-197, 2008.

LACAUX, J. P.;TOURRE, Y. M.; VIGNOLLES, C.; NDIONE, J. A.; LAFAYE, M. Classification of ponds from high-spatial resolution remote sensing: Application to Rift Valley Fever epidemics in Senegal. *Remote Sensing of Environment*, 106(1), 66-74.2007.

LAPORTE, J.; DOLOU, H.; AVIS, J.; ARINO, O. Thirty years of Satellite Derived Bathymetry: The charting tool that Hydrographers can no longer ignore. *The International Hydrographic Review*. International Hydrographic Organization. Monaco. Publication P-1. pp. 129-154. 2020.

LEDER, N.; LEDER, T. D. Satellite Derived Bathymetry–Low cost survey systems. *Book of*, p. 516, 2017.

LIANG, S.; FANG, H; CHEN, M. Atmospheric correction of Landsat ETM+ Land surface imagery—part I: methods. *IEEE Transactions on Geoscience and Remote Sensing*, v. 39, n. 11, p. 2490–2498, 2001.

LYZENGA, D. R. Passive remote sensing techniques for mapping water depth and bottom features. *Applied Optics*, 17(3), 379–383.1978.

LYZENGA, D. R., MALINAS, N. P., & TANIS, F. J. Multispectral Bathymetry Using a Simple Physically Based Algorithm. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 44(8), 2251–2259.2006.

MANLEY, J. E. Unmanned surface vehicles, 15 years of development. In: *OCEANS 2008*, IEEE, p.1-4,2008.

MATEO-PÉREZ, V.; CORRAL-BOBADILLA, M.;ORTEGA-FERNÁNDEZ, F.; VERGARA-GONZÁLEZ, E. P. Port bathymetry mapping using support vector machine technique and sentinel-2 satellite imagery. *Remote sensing*, 12(13), 2069.2020.

MATTHEW, M.W.; ADLER-GOLDEN, S.M.; BERK, A.; FELDE, G.; ANDERSON, G.P.; GORODETZKY, D.; PASWATERS, S.; SHIPPERT, M. Atmospheric correction of spectral imagery: Evaluation of the FLAASH algorithm with AVIRIS data. In *Proceedings of the 31st Applied Imagery Pattern Recognition Workshop*, Washington, DC, USA, 16–17 October 2002.

MAYER, L. *Frontiers in seafloor mapping and visualization*. Mar. Geophys.2006.

MENESES, P; ALMEIDA, T. *Introdução ao processamento de imagens de sensoriamento remoto*. 1 ed. Brasília: UnB/CNPq, v.1, 266p., 2012.

MISHRA, A., VOJINOVIC, Z., RAMAKRISHNAN, B., & LUIJENDIJK, A. Shallow water bathymetry mapping using Support Vector Machine ( SVM ) technique and multispectral imagery Shallow water bathymetry mapping using Support Vector Machine ( SVM ) technique

and multispectral imagery. *International Journal of Remote Sensing*, 00(00), 1–20.2018. <https://doi.org/10.1080/01431161.2017.1421796>

MOHAMED, H. & NADAOKA, K. Assessment of machine learning approaches for Bathymetry Mapping in shallow water environment using MS images. *International Journal of GeoInformatics*, 13(2). 2017.

NAJAR, M. A.; BENSILA, R.; BENNIQUI, Y. E.; THOUMYRE, G.; ALMAR, R.; BERGSMA, E. W.; WILSON, D. G. Coastal Bathymetry Estimation from Sentinel-2 Satellite Imagery: Comparing Deep Learning and Physics-Based Approaches. *Remote Sensing*, 14(5), 1196.2022.

PARAMESWARAN, S.; WEINBERGER, K.Q. Large margin multi-task metric learning. In *Advances in Neural Information Processing Systems*; Curran Associates Inc.57: Red Hook, NY, USA, 2010; pp. 1867–1875.2010.

POURSANIDIS, D.; TRAGANOS, D.;CHRYSOULAKIS, N.; REINARTZ, P. Cubesats allow high spatiotemporal estimates of satellite-derived bathymetry. *Remote Sensing*, 11(11), 129.2019a.

POURSANIDIS, D.; TRAGANOS, D.;REINARTZ, P.; CHRYSOULAKIS, N. On the use of Sentinel-2 for coastal habitat mapping and satellite-derived bathymetry estimation using downscaled coastal aerosol band. *Int J Appl Earth Obs Geoinformation*, 80(March), 58–70. 2019b. <https://doi.org/10.1016/j.jag.2019.03.012>.

REYNALDO, N. I. Comércio e navegação no rio Paraguai (1870-1940). XI Jornadas Interescuelas. Universidad de Tucumán, San Miguel de Tucumán-Argentina, p. 1-47, 2007.

RIBEIRO, S.R.A.; CENTENO, J.A.S.; KRUEGER, C. P. Estimativa de profundidade a partir de levantamento batimétrico e dados IKONOS II mediante redes neurais artificiais. *Boletim de Ciências Geodésicas*, v. 14, n. 2, p. 171-185, 2008.

RICHARDS, J. A. Error correction and registration of image data. In: *Remote Sensing Digital Image Analysis: an introduction*. 2.ed. Germany: Springer-Verlag Berlin Heidelberg, 1993. 339p. p.39-63.

RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. Prentice Hall. Series in Artificial Intelligence. New Jersey. 946 p. 1995.

SAGAWA, T., YAMASHITA, Y., OKUMURA, T., & YAMANOKUCHI, T. Satellite Derived Bathymetry Using Machine Learning and Multi-Temporal Satellite Images. *Remote Sensing*, 11(1155), 2–19. 2019.

SAMUEL, A. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*. v. 3. 3 ed. p. 535-554. 1959.

SANTOS, G.R.; OLIVEIRA, M.S.; LOUZADA, J.M.; SANTOS, A.M.R.T. Krigagem simples vs krigagem universal: qual o preditor mais preciso?. *Energia na agricultura (UNESP. Botucatu. CD-Rom)*, v.26, p.49-55,2011.

SANTOS, M.S. HENRIQUES, M.P. MATOS, A.J.S. Relatório de Atividades. Departamento de Hidrologia. Sistema de alerta hidrológico da bacia do rio Paraguai. Relatório Anual de operação 2021. Serviço Geológico do Brasil – CPRM. Brasília. 2021.

SHLIEN, S. Geometric correction, registration, and resampling of Landsat imagery. *Canadian journal of remote Sensing*, v. 5, n. 1, p. 74-89, 1979.

SU, H.; LIU, H.; HEYMAN, W. D. Automated derivation of bathymetric information from multi-spectral satellite imagery using a non-linear inversion model. *Marine Geodesy*, v. 31, n. 4, p. 281-298, 2008.

TONG, H.; CHEN, D. R.; PENG, L. Analysis of support vector machines regression. *Foundations of Computational Mathematics*, 9(2), 243-257. 2009.

USGS EROS. U.S.Geological Survey. Sentinel Missions. Sentinel 2A . 2015. Disponível em : U.S. Geological Survey Distribution of European Space Agency’s Sentinel-2 Data (usgs.gov). Acesso em : 20 de set. de 2021.

USGS. U.S.Geological Survey. Landsat Missions. Landsat 8. 2013. Disponível em : Landsat Missions (usgs.gov). Acesso em : 20 de set. de 2021.

VAPNIK V. *Statistical Learning Theory*. Wiley, New York.1998.

VERRELST, J.; MUÑOZ, J.; ALONSO, L.; DELEGIDO, J.; RIVERA, J.P.; CAMPS-VALLS, G.; MORENO, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* 2012, 118, 127–139.

VOJINOVIC, Z. ET AL., “A machine learning approach for estimation of shallow water depths from optical satellite images and sonar measurements,” *J. Hydroinform.*, vol. 15, no. 4, pp. 1408–1424, Oct. 2013, doi: 10.2166/hydro. 234.2013.

WARNE, D. K. Landsat as an aid in the preparation of hydrographic charts. *Photogrammetric engineering and remote sensing*, v. 44, n. 8, p. 1011-1016, 1978.

YANG, W.; ZHAO, Y.; WANG, D.; WU, H.; LIN, A.; HE, L. Using Principal Components Analysis and IDW Interpolation to Determine Spatial and Temporal Changes of Surface Water Quality of Xin’anjiang River in Huangshan, China. *Int. J. Environ. Res. Public Health*, 17, 2942.2020.<https://doi.org/10.3390/ijerph17082942>

ZHANG, Z.; HE, G.; WANG, X. A practical DOS model-based atmospheric correction algorithm. *Int. J. Remote Sens.* 31 (11), 2837–2852.2010.

ZHOU, W.; TANG, Y.; JING, W.; LI, Y.; YANG, J.; DENG, Y.; ZHANG, Y. A Comparison of Machine Learning and Empirical Approaches for Deriving Bathymetry from Multispectral Imagery. *Remote Sens.* 15, 393. 2023. <https://doi.org/10.3390/rs15020393>

ZULLO, J.; X. F. GU, R.; LAMPARELLI, G.; GUYOUT, P. BEZERRA. Préprocessamentos das imagens de satélites. *CIG – Caderno de Informações Georeferenciadas*, 1(1).1996. Disponível em: <http://www.cpa.unicamp.br/revista/cigv1n1a7.html>

## APÊNDICE A – Algoritmos empregados

### ##ALGORITMO RNA

##Parte inicial - Carregando os pacotes utilizados

```
library(nnet)
```

```
library(neuralnet)
```

```
library(ggplot2)
```

```
library(ggrepel)
```

```
library(caret)
```

```
library (rpart)
```

```
library("gstat")
```

```
library("sp")
```

```
library("rgdal")
```

```
library(dplyr)
```

```
library(readr)
```

```
library(GGally)
```

```
library(dplyr)
```

```
library(mlbench)
```

#Definindo o número de repetições e semente aleatória

```
nrep = 5
```

```
set.seed(310)
```

```
vrandom = sample(1:1000, nrep)
```

#Criando o dataframe dos resultados finais

```
df = data.frame(id = integer(nrep), rmse = numeric(nrep), MAE = numeric(nrep),
```

```
Rsquared = numeric(nrep),Inc_Robust = numeric(nrep))
```

#Leitura dos dados

```
dados<-read.table("./pts_sent_mad_AGO.txt", sep = ";", header = TRUE)
```

```
glimpse(dados)
```

##Criação das funções

#MAE

```
fmae = function(ypred,yobs){
```

```
  ae = abs(ypred - yobs)
```

```
  mae = mean(ae)
```

```
  return(mae)
```

```

}
# RMSE
frmse = function(ypred,yobs){
  ae = mean((ypred - yobs)^2)
  rmse = sqrt(ae)
  return(rmse)
}
#Incerteza Robusta
finc_robust = function(ypred,yobs){
  ae = mad(ypred - yobs)^2
  ae2 = median(ypred - yobs)^2
  Inc_Robust = sqrt(ae+ae2)
  return(Inc_Robust)
}
# Normalização do conjunto de dados de acordo com o valor máximo e mínimo
preproc2 <- preProcess(dados, method=c("range"))
norm2 <- predict(preproc2, dados)
summary(norm2)
#####
#Início da repetição
for (i in 1:nrep) {
  print(i) #Mostra para o usuário em qual repetição está
  set.seed(vrandom[i]) #Semente definida de acordo com a repetição
  vf = createDataPartition(norm2$Z, times = 1, p = 0.7, list = FALSE)
  treino = norm2[vf,]
  teste = norm2[-vf,]

  #Treinamento da Rede Neural
  print(RNA_Sentinel_Madeira <- neuralnet(Z ~. , data=treino, hidden=7, threshold = 1,
rep=10, stepmax = 1e+06))
  R = plot(RNA_Sentinel_Madeira)
  R[i] = paste("Estrutura_RNA_sent_MAD",i,".png")
  savePlot(filename = R[i], type = "png")
  #Fazendo a predição para a mostra de teste

```

```

amostra_teste<- predict(RNA_Sentinel_Madeira,teste)
teste$predict = amostra_teste
teste$dif = teste$Z - teste$predict
#Desnormalizando os dados
teste$X <- ((max(dados$X)-min(dados$X))*teste$X) + min(dados$X)
teste$Y <- ((max(dados$Y)-min(dados$Y))*teste$Y) + min(dados$Y)
teste$Z <- ((max(dados$Z)-min(dados$Z))*teste$Z) + min(dados$Z)
teste$predict <- ((max(dados$Z)-min(dados$Z))*teste$predict) + min(dados$Z)
teste$dif <- ((max(dados$Z)-min(dados$Z))*teste$dif) + min(dados$Z)
pr_RNA = postResample(teste$predict, teste$Z)
pr_RNA = c(pr_RNA, Min = min(teste$Z), Média = mean(teste$Z), Max = max(teste$Z))
kableExtra::kable(t(pr_RNA), digits = 2)
print(pr_RNA)
#Gerando e salvando o gráfico (predito x observado)
g = ggplot(data = teste, mapping = aes(x = Z, y = predict)) +
geom_point() + labs( title = "Comparação entre valores observados e preditos vom RNA",x =
'Z observado', y = 'Z predito') + geom_abline() + theme_classic() + theme(axis.text.x =
element_text(angle = 0))
print(g)
g[i] = paste("g_RNA_sent_MAD",i,".png")
ggsave(g[i])
#Colocar no Dataframe "df" os valores de RMSE, MAE e Incerteza Robusta
df$id[i] = i
df$rmse[i] = frmse(teste$Z, teste$predict)
df$MAE[i] = fmae(teste$Z, teste$predict)
df$Rsquared[i]= pr_RNA[2]
df$Inc_Robust[i] = finc_robust(teste$Z, teste$predict)

write.table(teste, file = "RNA_Sent_MAD.txt", append = TRUE, row.names = TRUE,
sep = ";", dec = ",")
}

#Exportando o resultado do R

```

```
write.csv(df, file = "Results_RNA_Sent_MAD_RMSE_MAE_IncRobust.txt", row.names =  
TRUE)
```

## ##ALGORITMO RF

```
##Parte inicial - Carregando os pacotes utilizados
```

```
library(caret)
```

```
library (rpart)
```

```
library(ggplot2)
```

```
library(ggrepel)
```

```
library("gstat")
```

```
library("sp")
```

```
library("rgdal")
```

```
library(readr)
```

```
library(sf)
```

```
library(dplyr)
```

```
library(randomForest)
```

```
#Definindo o número de repetições e semente aleatória
```

```
nrep = 5
```

```
set.seed(310)
```

```
vrandom = sample(1:1000, nrep)
```

```
df = data.frame(id = integer(nrep), rmse = numeric(nrep), MAE = numeric(nrep),  
Rsquared = numeric(nrep), Inc_Robust = numeric(nrep))
```

```
#Leitura dos dados
```

```
dados<-read.table("./pts_sent_mad_AGO.txt", sep = ";", header = TRUE)
```

```
### MAE
```

```
fmae = function(ypred,yobs){
```

```
  ae = abs(ypred - yobs)
```

```
  mae = mean(ae)
```

```
  return(mae)
```

```
}
```

```
### RMSE
```

```

frmse = function(ypred,yobs){
  ae = mean((ypred - yobs)^2)
  rmse = sqrt(ae)
  return(rmse)
}
### Incerteza Robusta
finc_robust = function(ypred,yobs){
  ae = mad(ypred - yobs)^2
  ae2 = median(ypred - yobs)^2
  Inc_Robust = sqrt(ae+ae2)
  return(Inc_Robust)
}
#Inicio da repeticao
for (i in 1:nrep) {
  print(i) #Mostra para o usu?rio em qual repeti??o est?
  set.seed(vrandom[i]) #Semente definida de acordo com a repeti??o
  vf = createDataPartition(dados$Z, times = 1, p = 0.7, list = FALSE) #parti??o
  #dos dados em treino e teste (70% para treino e 30% para teste)
  treino = dados[vf,]
  teste = dados[-vf,]
  ctrl <- trainControl(method = 'repeatedcv', number = 10, repeats = 3)
  model_rf <- train(Z ~., data = treino, metric = 'MAE', method = "rf", preProcess =
c('BoxCox', 'center','scale'), tuneLength = 5, verbose = FALSE, trControl = ctrl)

  #Teste
  v_Z_rf_teste = predict(model_rf, teste)
  teste$predict = v_Z_rf_teste
  pr_rf = postResample(v_Z_rf_teste, teste$Z)
  print('Desempenho no conjunto de teste: ')

  pr_rf = c(pr_rf, Min = min(teste$Z), Media = mean(teste$Z), Max = max(teste$Z))
  kableExtra::kable(t(pr_rf), digits = 2)
  print(pr_rf)
  teste$dif = teste$Z - teste$predict

```

```

#Gráfico de dispersão
g = ggplot(data = teste, mapping = aes(x = Z, y = predict)) + geom_point() + labs( title =
"Comparação entre valores observados e preditos com RF Sentinel Madeira",x = 'Z
observado', y = 'Z predito') + geom_abline() + theme_classic() + theme(axis.text.x =
element_text(angle = 0))

print(g)

g[i] = paste("g_RF_sentinel_MAD",i,".png")
ggsave(g[i])
df$id[i] = i
df$rmse[i] = frmse(teste$Z, teste$predict)
df$MAE[i] = fmae(teste$Z, teste$predict)
df$Rsqared[i]= pr_rf[2]
df$Inc_Robust[i] = finc_robust(teste$Z, teste$predict)
write.table(teste, file = "RF_Sent_MAD.txt", append = TRUE, row.names = TRUE, sep = ";",
dec = ",")
}
write.table(df, file = "Results_RF_Sentinel_mad_RMSE_MAE_INC_Robust.txt", row.names
= TRUE, sep = ";", dec = ",")

```

## ##ALGORITMO GBM

```
##Parte inicial - Carregando os pacotes utilizados
```

```

library(caret)
library (rpart)
library(gbm)
library(ggplot2)
library(ggrepel)
library("sp")
library("raster")
library("rgdal")
library(readr)
library(sf)
library(gstat)
library(dplyr)
nrep = 5

```

```

set.seed(310)
vrandom = sample(1:1000, nrep)
### Funcao para o cálculo do MAE
fmae = function(ypred,yobs){
  ae = abs(ypred - yobs)
  mae = mean(ae)
  return(mae)
}
### Funcao para o calculo do RMSE
frmse = function(ypred,yobs){
  ae = mean((ypred - yobs)^2)
  rmse = sqrt(ae)
  return(rmse)
}
### Funcao para o calculo da Incerteza Robusta
finc_robust = function(ypred,yobs){
  ae = mad(ypred - yobs)^2
  ae2 = median(ypred - yobs)^2
  Inc_Robust = sqrt(ae+ae2)
  return(Inc_Robust)
}
df = data.frame(id = integer(nrep), rmse = numeric(nrep), MAE = numeric(nrep),
               Rsquared = numeric(nrep),Inc_Robust = numeric(nrep))
#Leitura dos dados
dados<-read.table("./pts_sent_mad_AGO.txt", sep = ";", header = TRUE)
for (i in 1:nrep) {
  print(i)
  set.seed(vrandom[i])
  vf = createDataPartition(dados$Z, times = 1, p = 0.7, list = FALSE)
  #dos dados em treino e teste (70% para treino e 30% para teste)
  treino = dados[vf,]
  teste = dados[-vf,]
  ctrl <-trainControl(method = 'repeatedcv',
                     number = 10,

```

```

repeats = 3)
model_gbm <- train(Z ~., data = treino,
  metric = 'MAE',
  method = "gbm",
  preProcess = c('BoxCox', 'center','scale'),
  tuneLength = 5,
  verbose = FALSE,
  trControl = ctrl)
v_Z_gbm_teste = predict(model_gbm, teste)
teste$predict = v_Z_gbm_teste
pr_gbm = postResample(v_Z_gbm_teste, teste$Z)
print('Desempenho no conjunto de teste: ')
pr_gbm = c(pr_gbm, Min = min(teste$Z), Média = mean(teste$Z),
  Max = max(teste$Z))
kableExtra::kable(t(pr_gbm), digits = 2)
print(pr_gbm)
teste$dif = teste$Z - teste$predict
g = ggplot(data = teste, mapping = aes(x = Z, y = predict)) +
  geom_point() + labs( title = "Comparação entre valores observados e preditos
  com GBM-SENTINEL-MADEIRA",x = 'Z observado', y = 'Z predito') +
  geom_abline() + theme_classic() + theme(axis.text.x = element_text(angle = 0))

print(g)
g[i] = paste("g_RNA_sent_MAD",i,".png")
ggsave(g[i])
df$id[i] = i
df$rmse[i] = frmse(teste$Z, teste$predict)
df$MAE[i] = fmae(teste$Z, teste$predict)
df$Rsquared[i]= pr_gbm[2]
df$Inc_Robust[i] = finc_robust(teste$Z, teste$predict)

write.table(teste, file = "GBM_Sent_MAD.txt", append = TRUE, row.names = TRUE,
  sep = ";", dec = ",")

```

```

}
write.table(df, file = "Results_Sent_mad_RMSE_MAE_incRobust.txt", row.names = TRUE,
sep = ";", dec = ",")

```

## ##ALGORITMO SVM

##Parte inicial - Carregando os pacotes utilizados

```
library(caret)
```

```
library (rpart)
```

```
library(svmpath)
```

```
library(ggplot2)
```

```
library(ggrepel)
```

```
library("sp")
```

```
library("rgdal")
```

```
library(readr)
```

```
library(sf)
```

```
library(gstat)
```

```
library(dplyr)
```

```
nrep = 5
```

```
set.seed(310)
```

```
vrandom = sample(1:1000, nrep)
```

### Funcao para o cálculo do MAE

```
fmae = function(ypred,yobs){
```

```
  ae = abs(ypred - yobs)
```

```
  mae = mean(ae)
```

```
  return(mae)
```

```
}
```

### Funcao para o calculo do RMSE

```
frmse = function(ypred,yobs){
```

```
  ae = mean((ypred - yobs)^2)
```

```
  rmse = sqrt(ae)
```

```
  return(rmse)
```

```
}
```

### Funcao para o calculo da Incerteza Robusta

```
finc_robust = function(ypred,yobs){
```

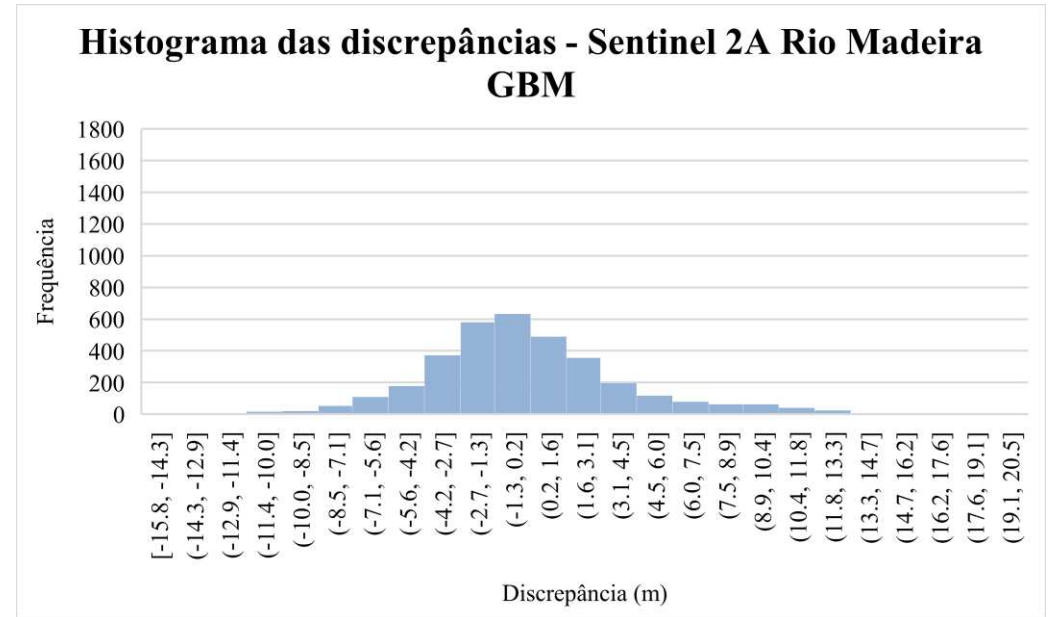
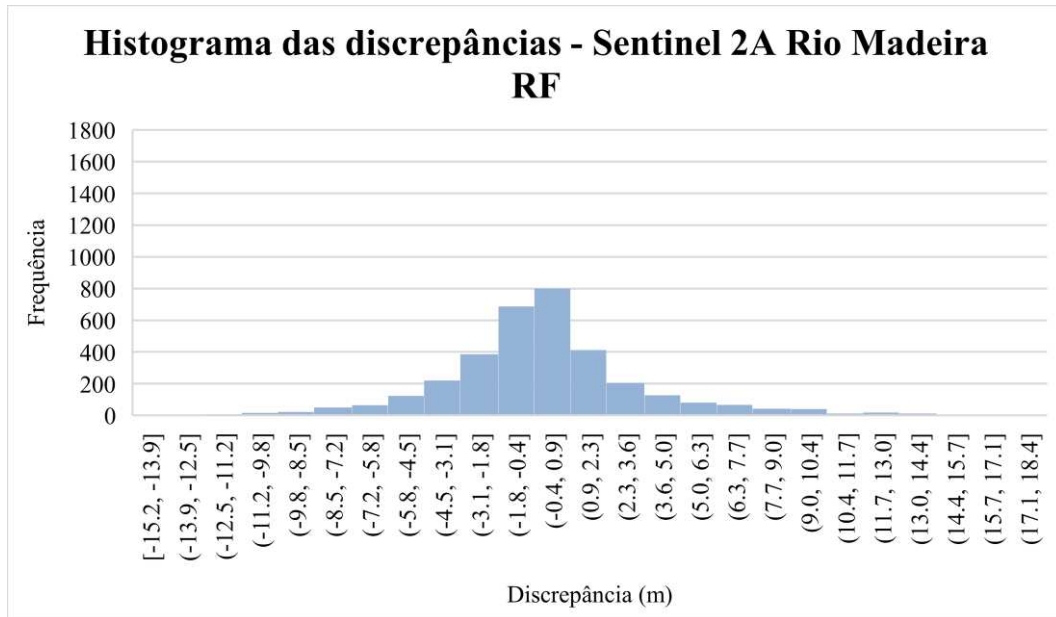
```

ae = mad(ypred - yobs)^2
ae2 = median(ypred - yobs)^2
Inc_Robust = sqrt(ae+ae2)
return(Inc_Robust)
}
df = data.frame(id = integer(nrep), rmse = numeric(nrep), MAE = numeric(nrep),
               Rsquared = numeric(nrep), Inc_Robust = numeric(nrep))
#Leitura dos dados
dados<-read.table("./pts_sent_mad_AGO.txt", sep = ";", header = TRUE)
#Inicio da repeticao
for (i in 1:nrep) {
  print(i)
  set.seed(vrandom[i])
  vf = createDataPartition(dados$Z, times = 1, p = 0.7, list = FALSE)
  treino = dados[vf,]
  teste = dados[-vf,]
  ctrl <-trainControl(method = 'repeatedcv', number = 10, repeats = 3)
  model_SVM <- train(Z ~., data = treino, metric = 'MAE', method = "svmRadial",
preProcess = c('BoxCox', 'center','scale'), tuneLength = 5, verbose = FALSE, trControl =
ctrl)
  v_Z_SVM_teste = predict(model_SVM, teste)
  teste$predict = v_Z_SVM_teste
  pr_SVM = postResample(v_Z_SVM_teste, teste$Z)
  print('Desempenho no conjunto de teste: ')
  pr_SVM = c(pr_SVM, Min = min(teste$Z), Média = mean(teste$Z), Max = max(teste$Z))
  kableExtra::kable(t(pr_SVM), digits = 2)
  print(pr_SVM)
  teste$dif = teste$Z - teste$predict
  g = ggplot(data = teste, mapping = aes(x = Z, y = predict)) +
    geom_point() + labs( title = "Comparação entre valores observados e preditos
                        com SVM- Sentinel Madeira",x = 'Z observado', y = 'Z predito') +
    geom_abline() + theme_classic() + theme(axis.text.x = element_text(angle = 0))
  print(g)
  ggsave("g.png")

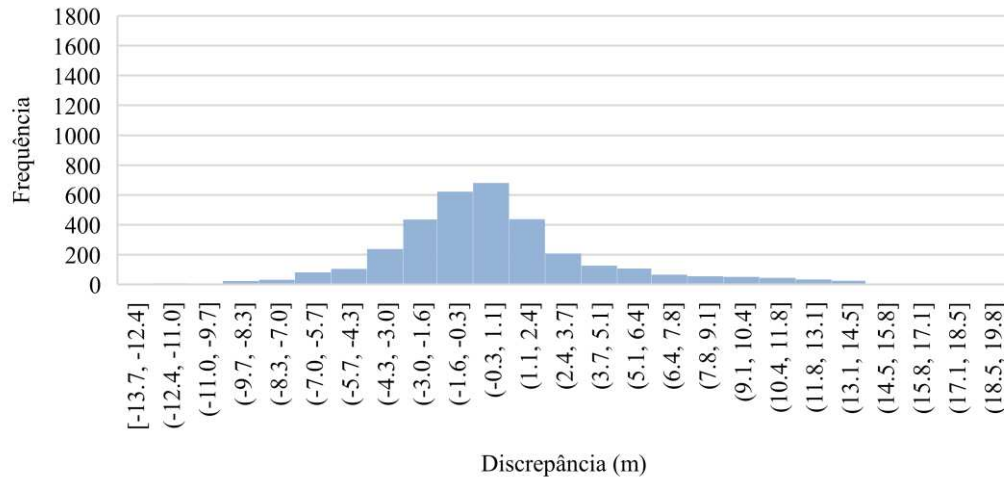
```

```
g[i] = paste("g_SVM_sent_MAD",i,".png")
ggsave(g[i])
df$Id[i] = i
df$rmse[i] = frmse(teste$Z, teste$predict)
df$MAE[i] = fmae(teste$Z, teste$predict)
df$Rsquared[i]= pr_SVM[2]
df$Inc_Robust[i] = finc_robust(teste$Z, teste$predict)
write.table(teste, file = "SVM_Sent_MAD.txt", append = TRUE, row.names = TRUE,
            sep = ";", dec = ",")
write.table(df, file = "Results_Sent_mad_RMSE_MAE_IncRobusta.txt", row.names = TRUE,
            sep = ";", dec = ",")
}
```

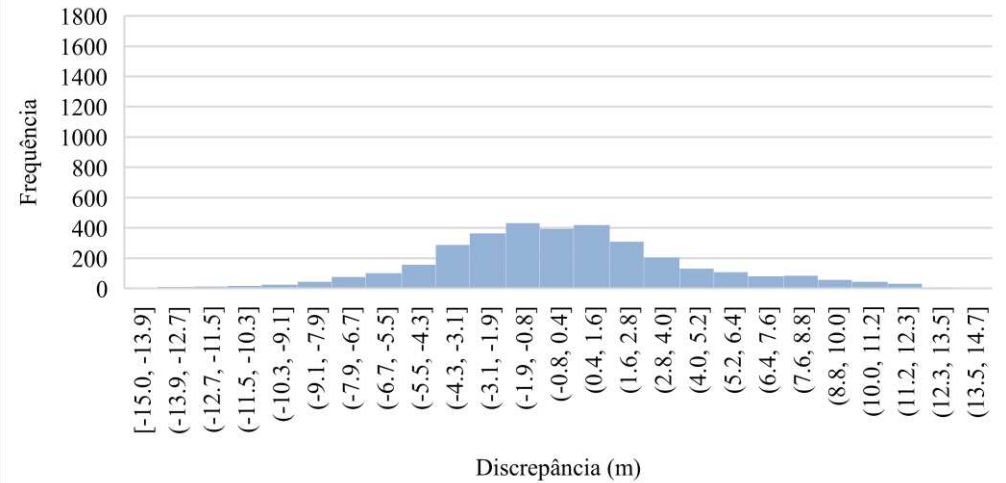
## APÊNDICE B – Histogramas das discrepâncias e Gráficos de dispersão



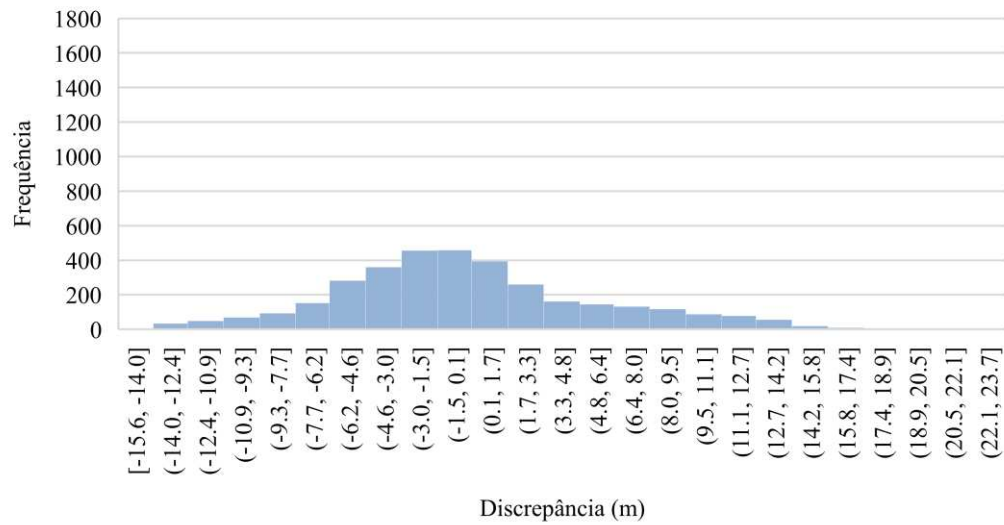
**Histograma das discrepâncias - Sentinel 2A Rio Madeira SVM**



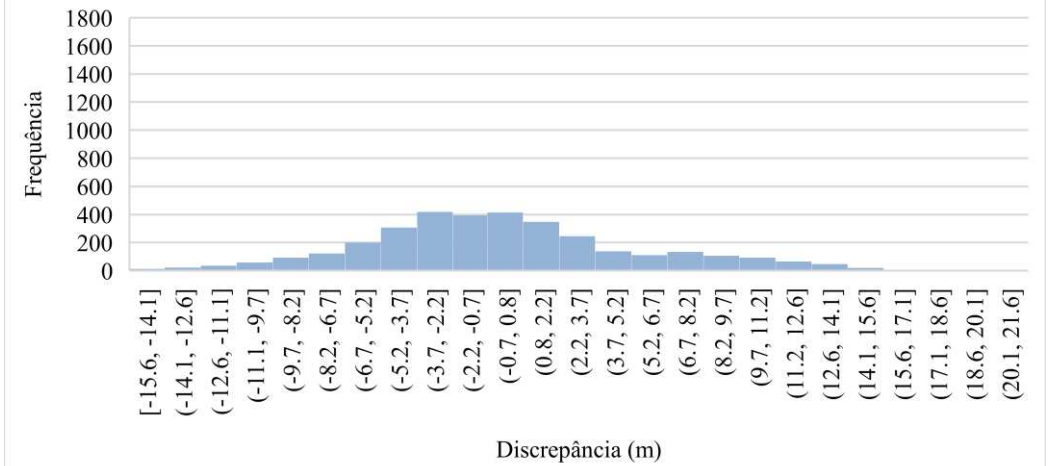
**Histograma das discrepâncias - Sentinel 2A Rio Madeira RNA**



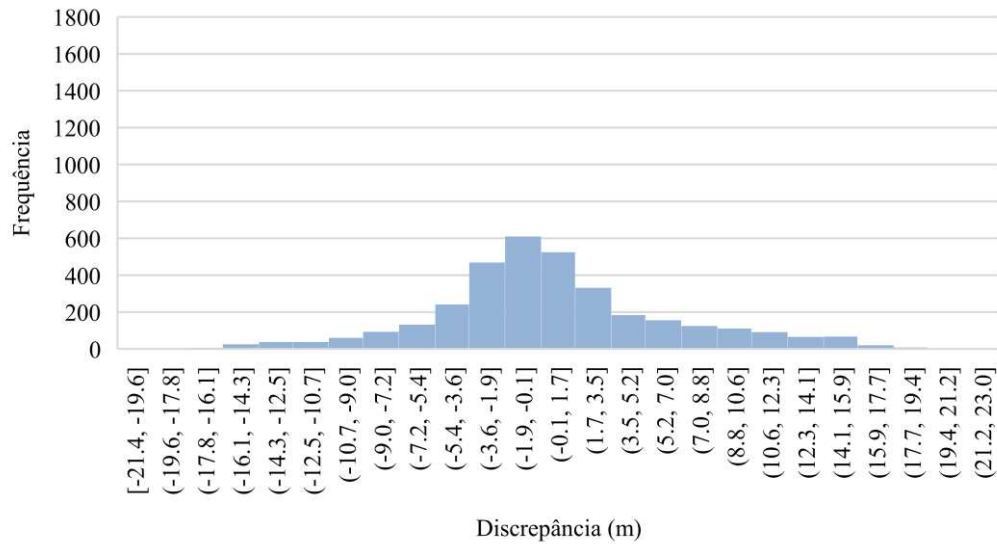
### Histograma das discrepâncias - CBERS Rio Madeira RF



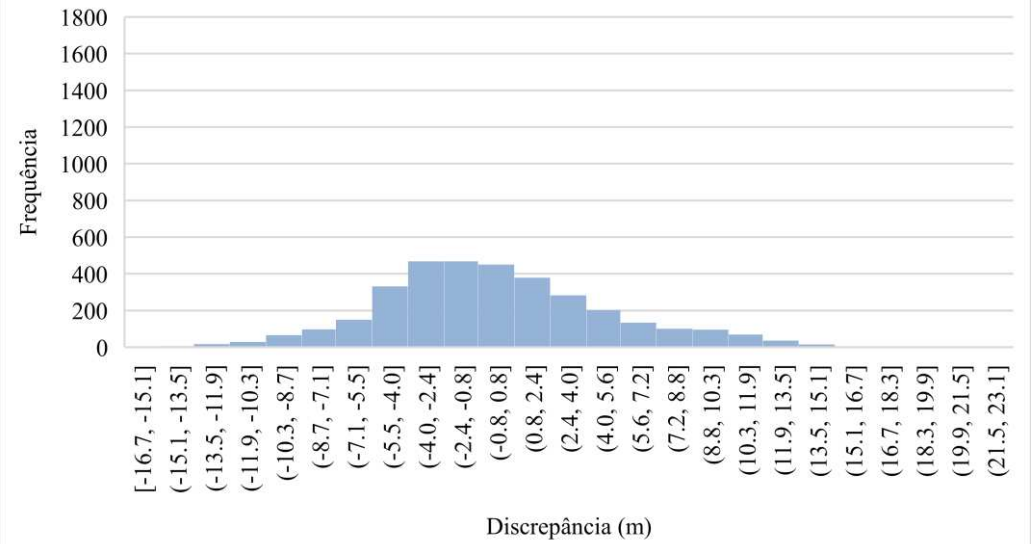
### Histograma das discrepâncias - CBERS Rio Madeira GBM



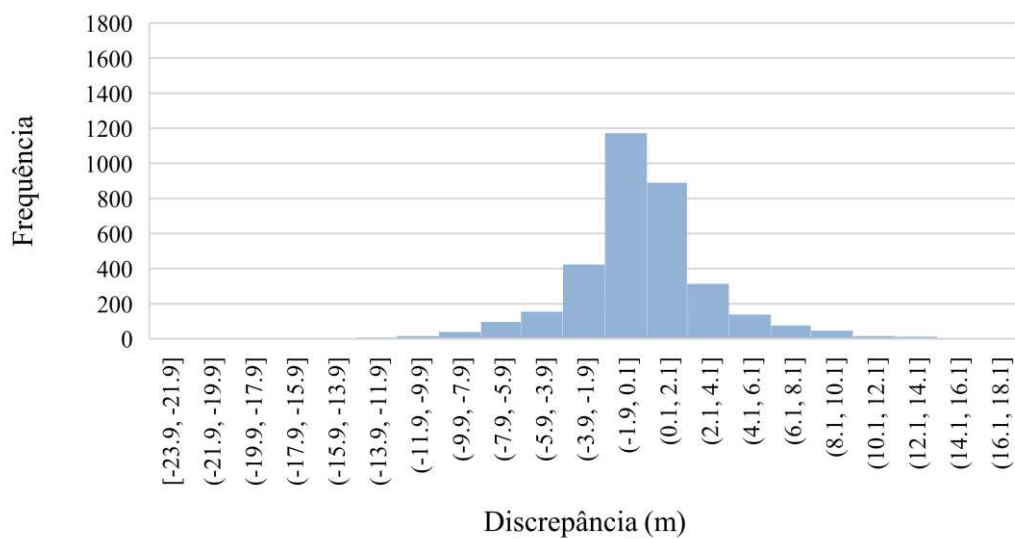
**Histograma das discrepâncias - CBERS Rio Madeira SVM**



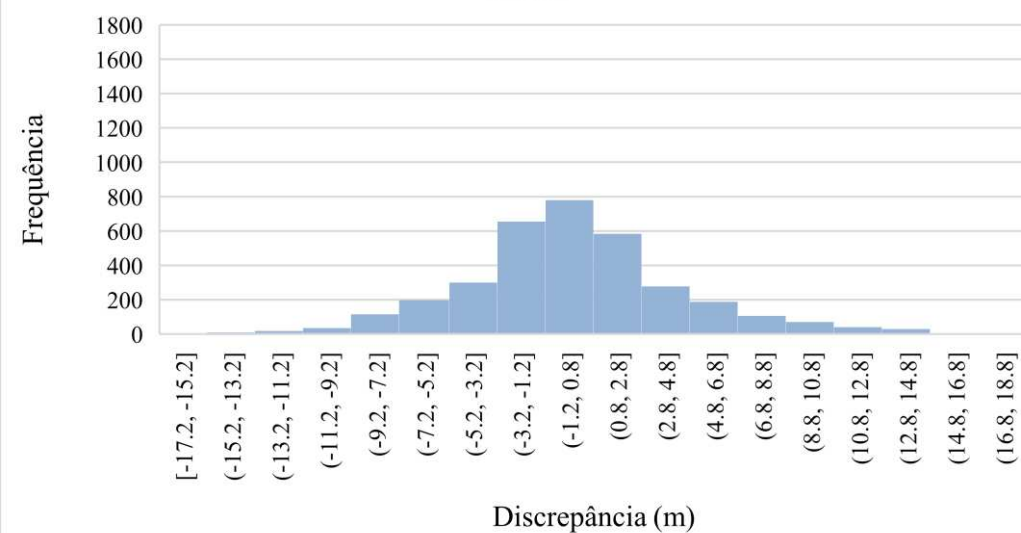
**Histograma das discrepâncias - CBERS Rio Madeira RNA**



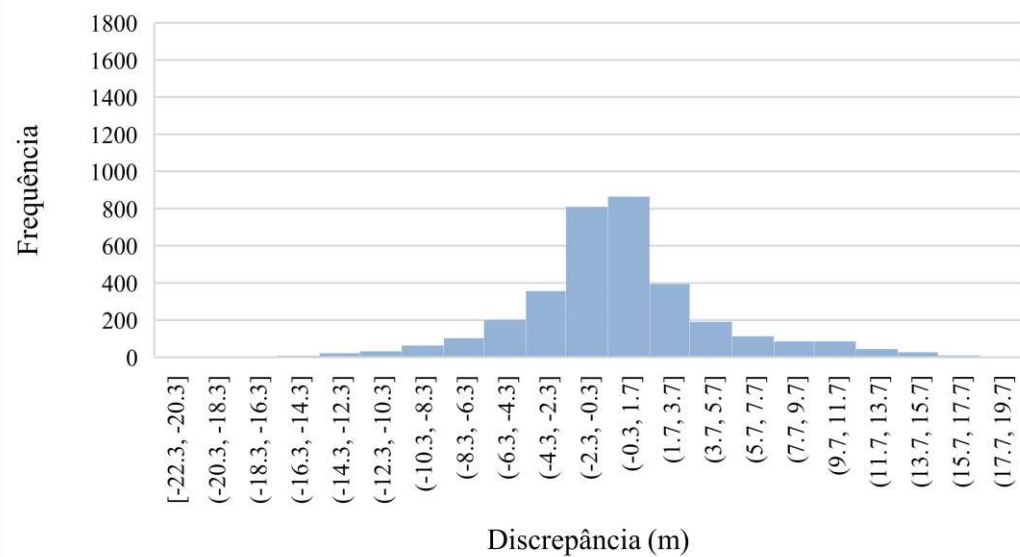
### Histograma das discrepâncias - Landsat Rio Madeira RF



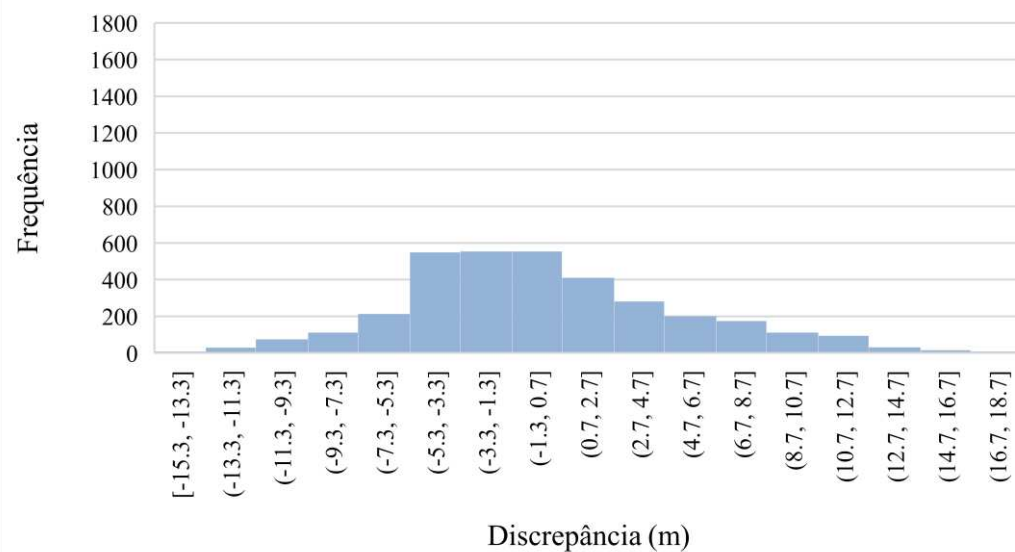
### Histograma das discrepâncias - Landsat Rio Madeira GBM



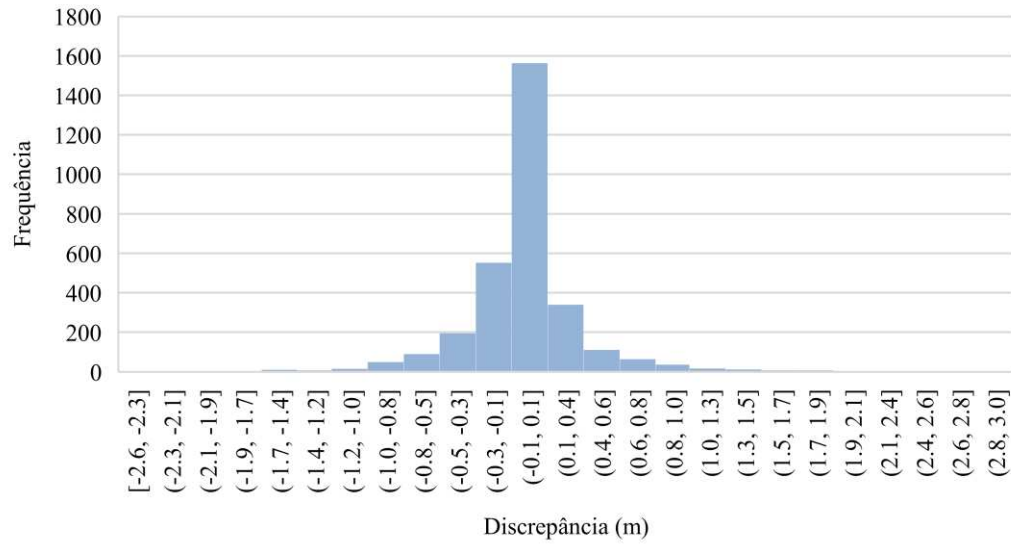
**Histograma das discrepâncias - Landsat Rio Madeira SVM**



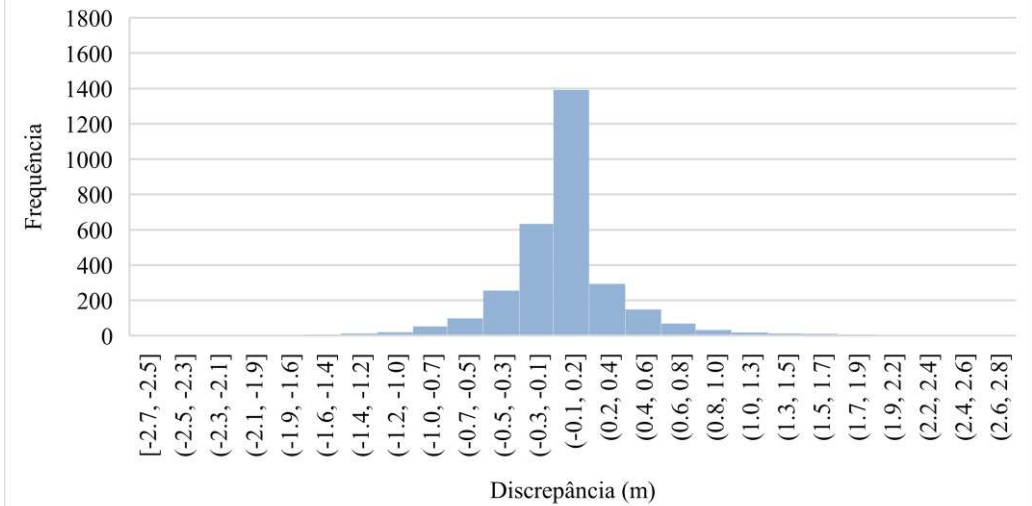
**Histograma das discrepâncias - Landsat Rio Madeira RNA**



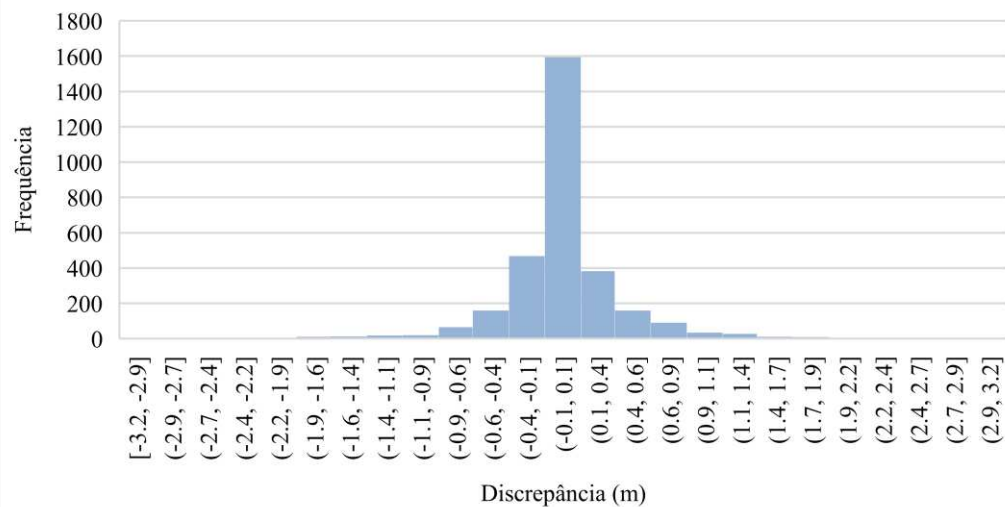
### Histograma das discrepâncias - Sentinel Rio Paraguai RF



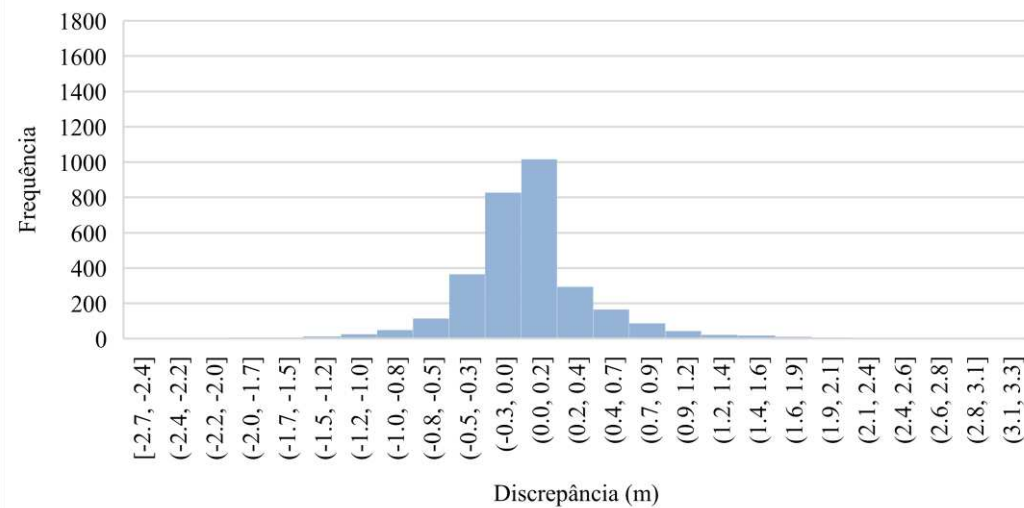
### Histograma das discrepâncias - Sentinel Rio Paraguai GBM



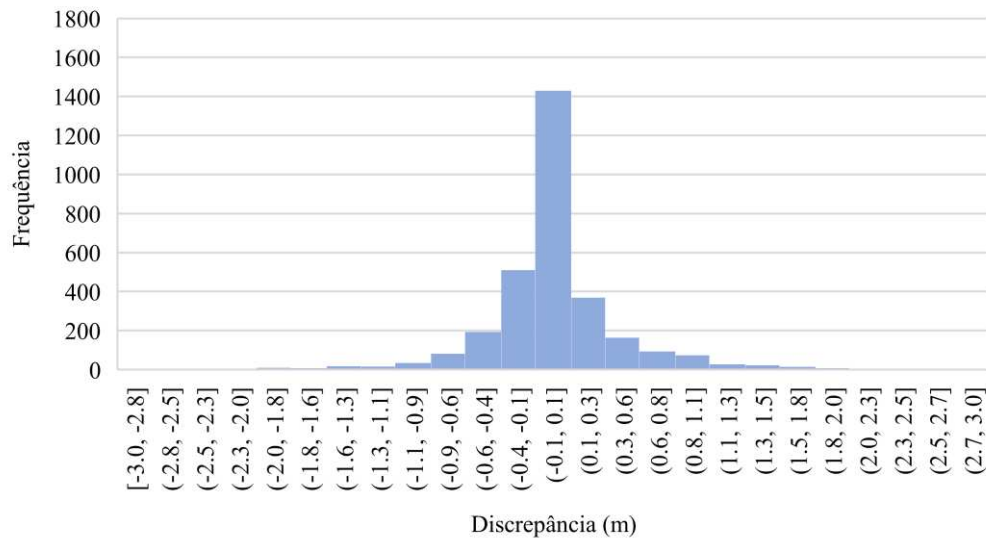
**Histograma das discrepâncias - Sentinel Rio Paraguai  
SVM**



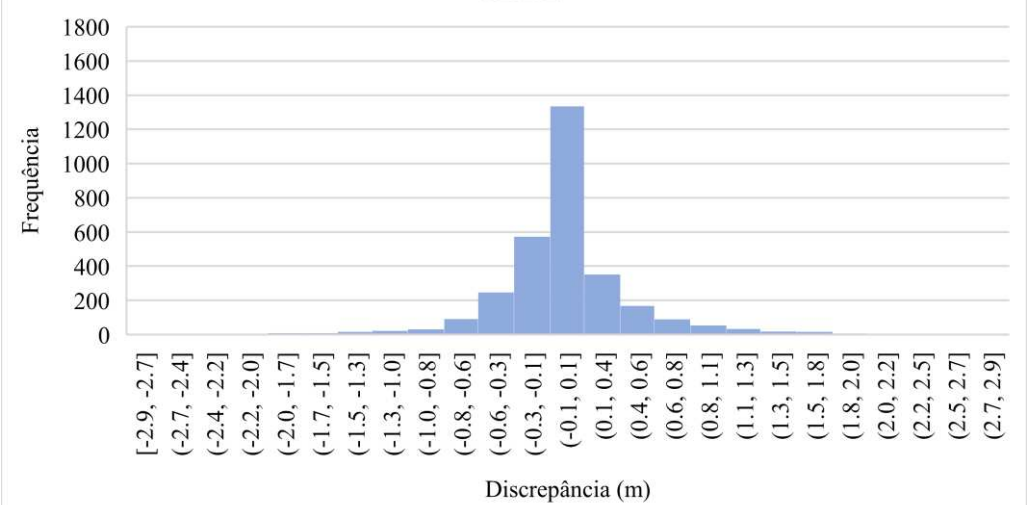
**Histograma das discrepâncias - Sentinel Rio Paraguai  
RNA**



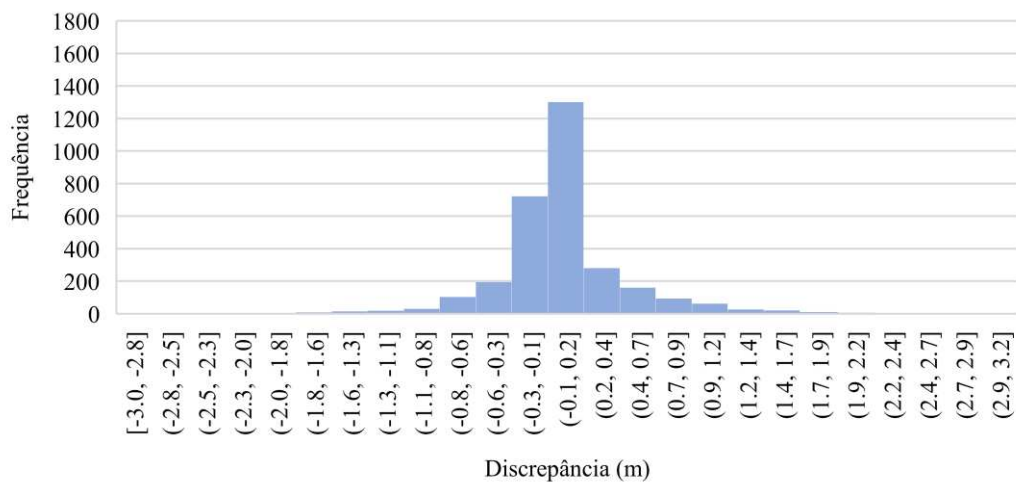
### Histograma das discrepâncias - CBERS Rio Paraguai RF



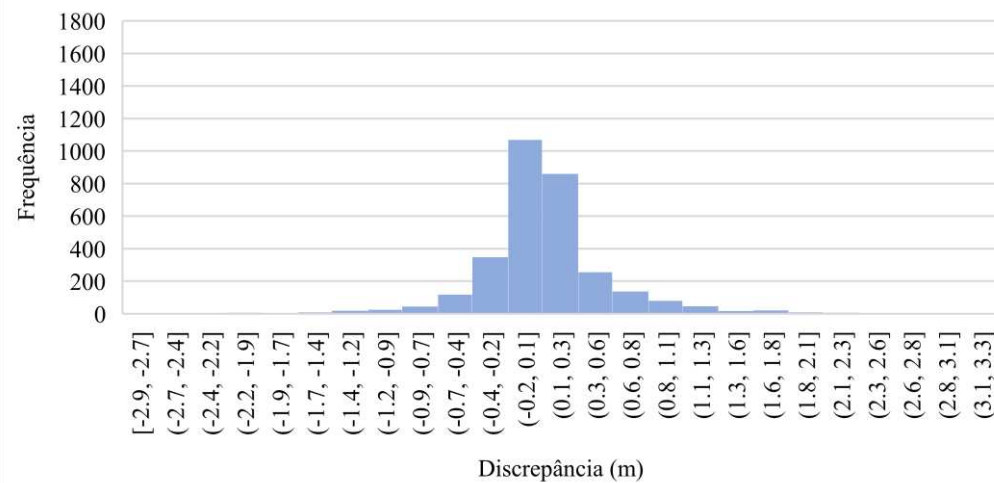
### Histograma das discrepâncias - CBERS Rio Paraguai GBM



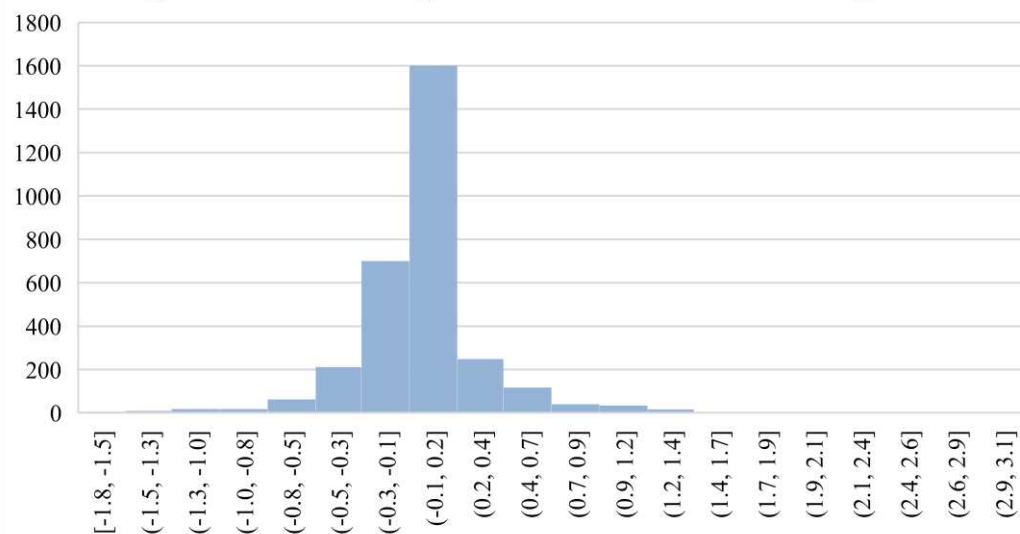
**Histograma das discrepâncias - CBERS Rio Paraguai  
SVM**



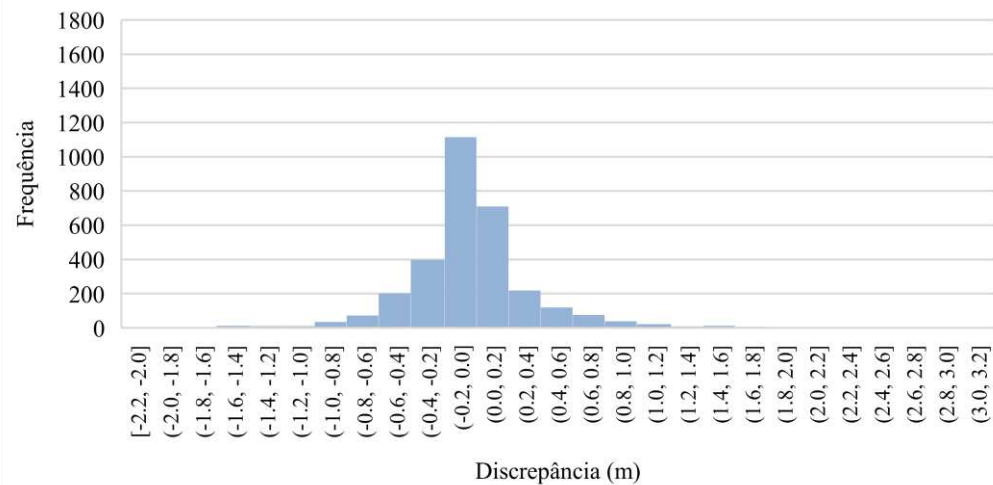
**Histograma das discrepâncias - CBERS Rio Paraguai  
RNA**



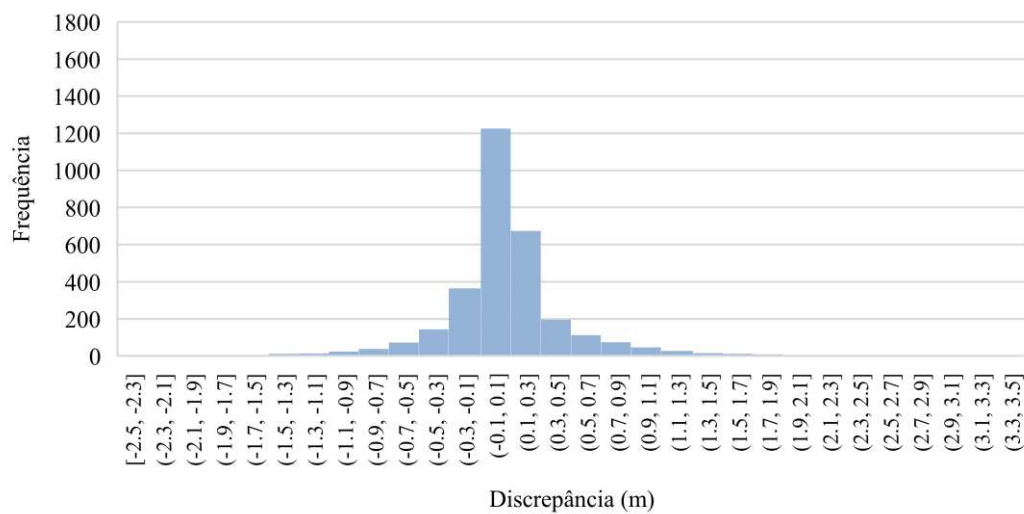
### Histograma das discrepâncias - Landsat Rio Paraguai RF



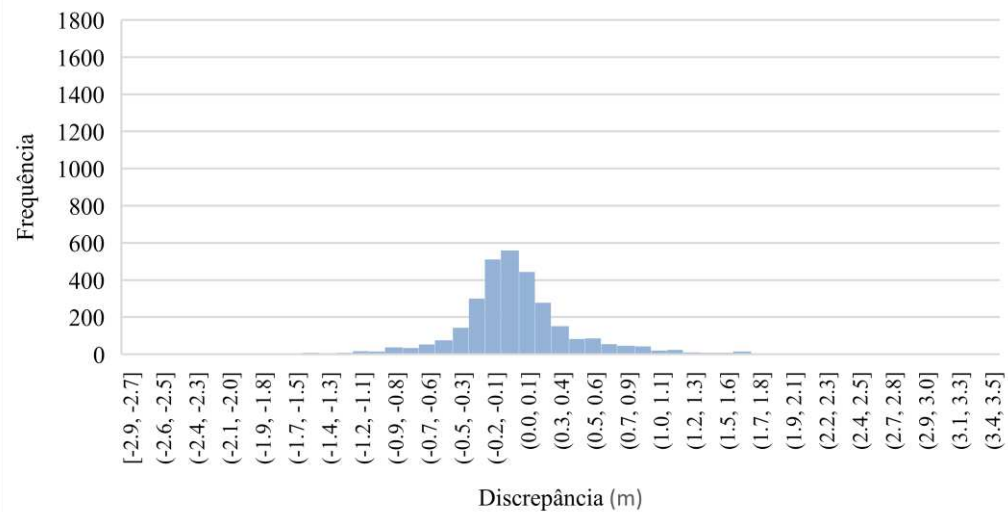
### Histograma das discrepâncias - Landsat Rio Paraguai GBM

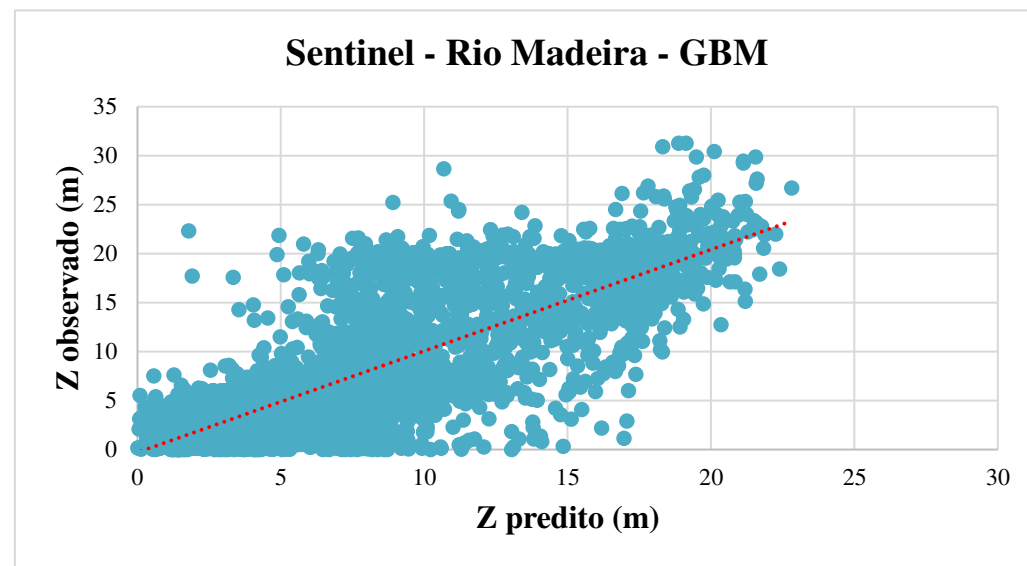
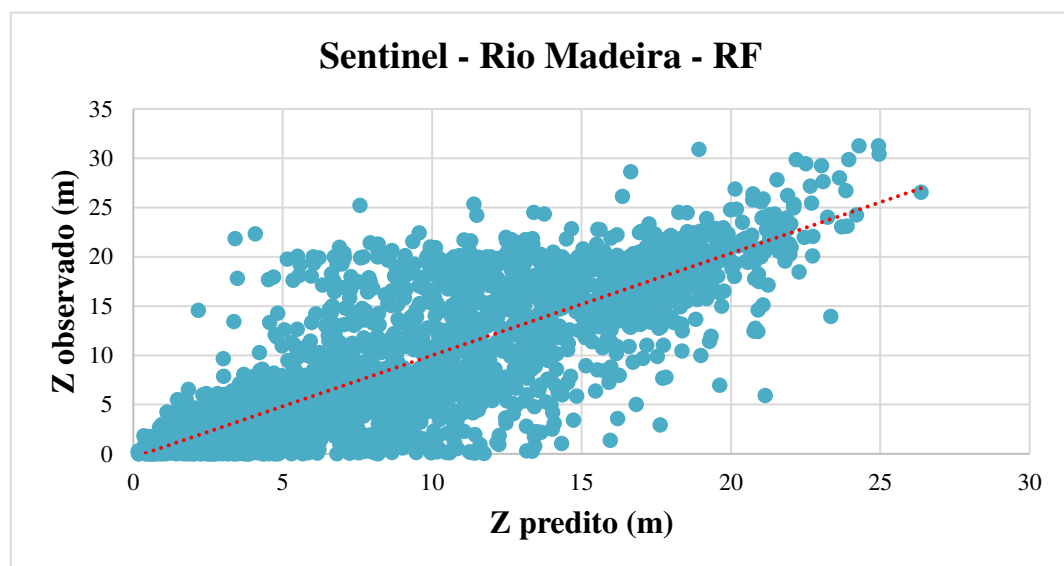


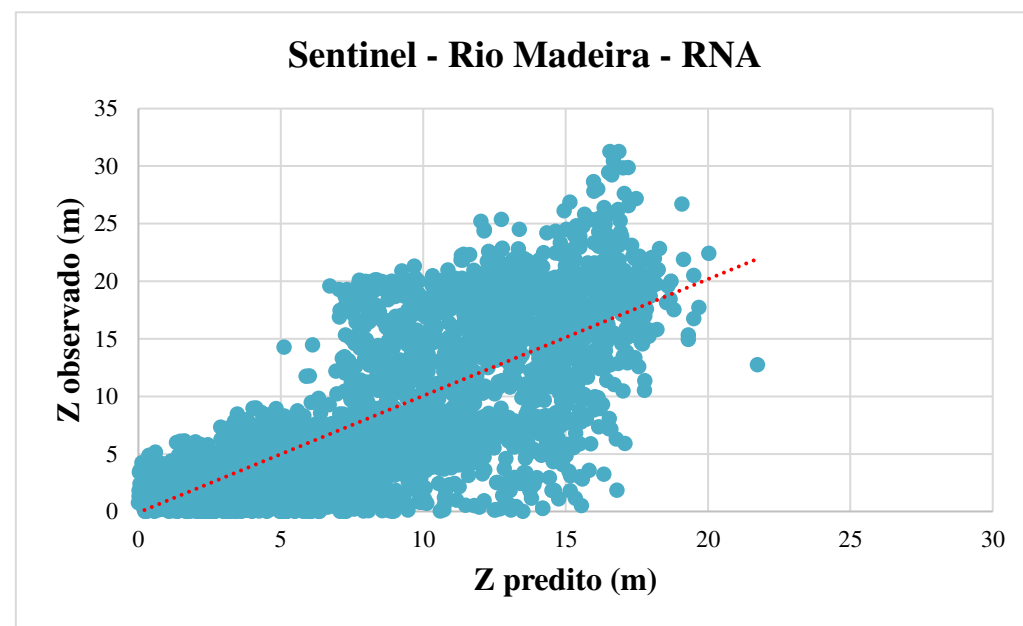
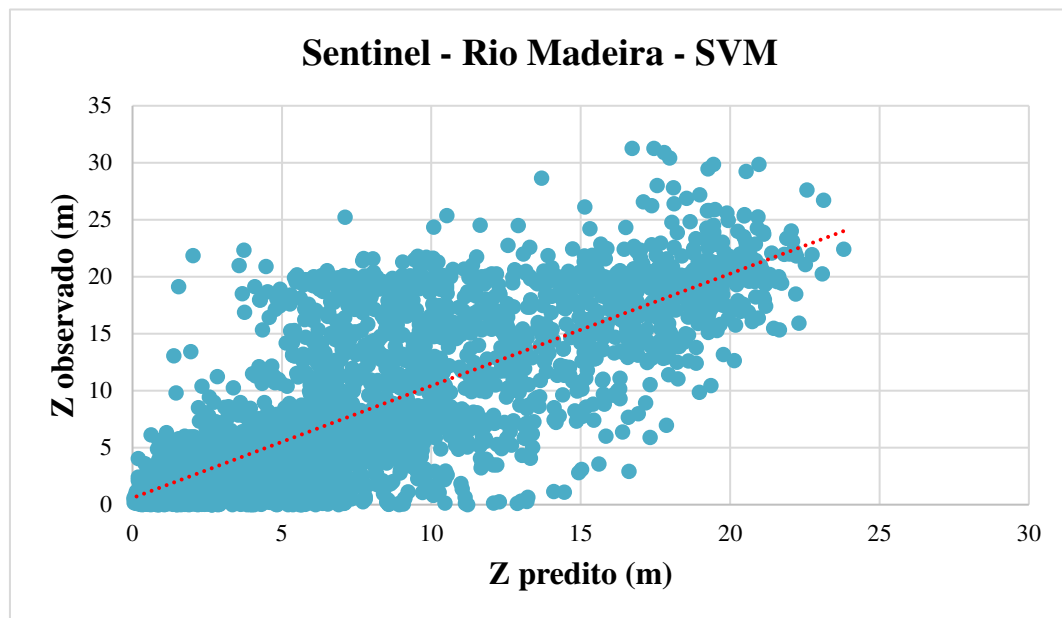
**Histograma das discrepâncias - Landsat Rio Paraguai SVM**

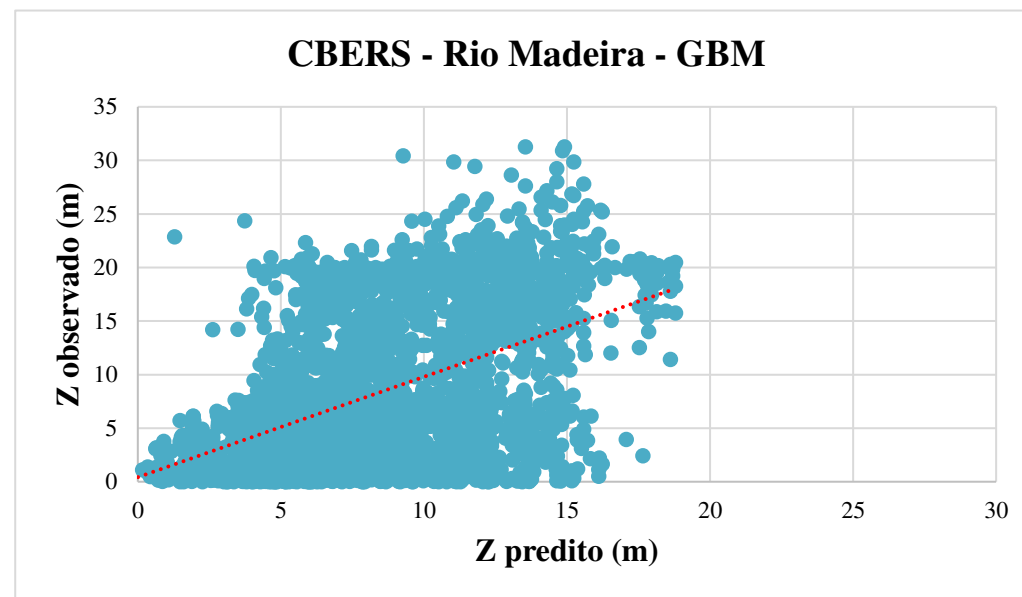
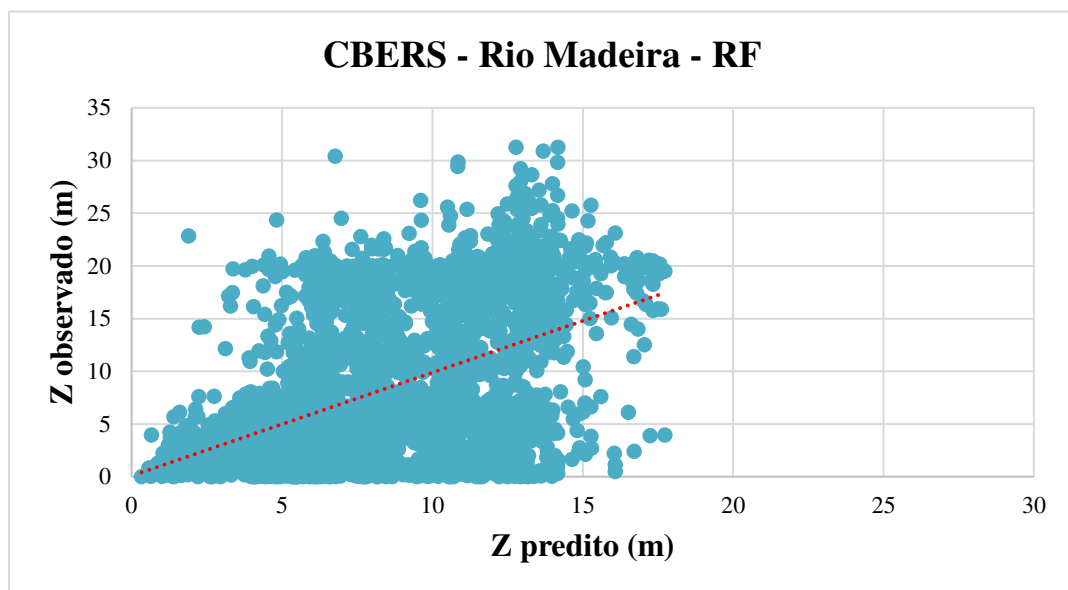


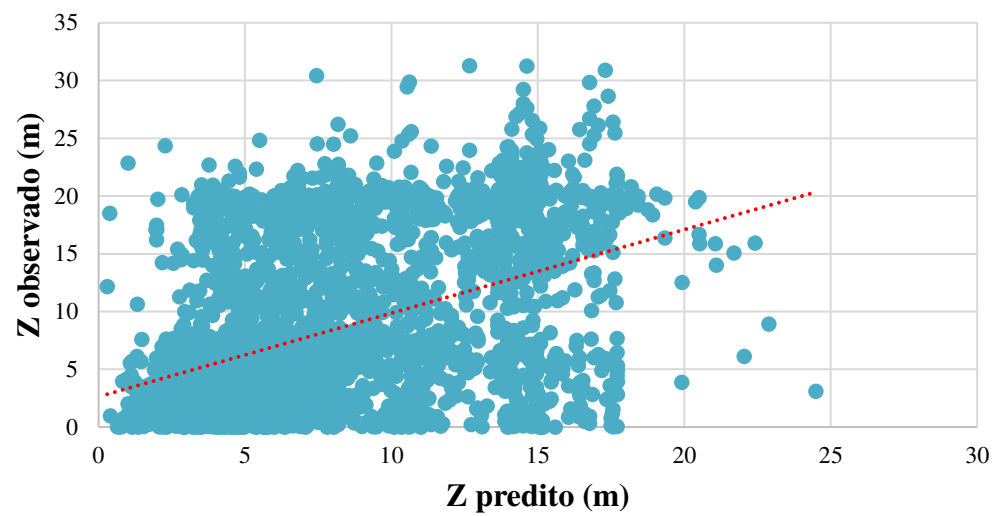
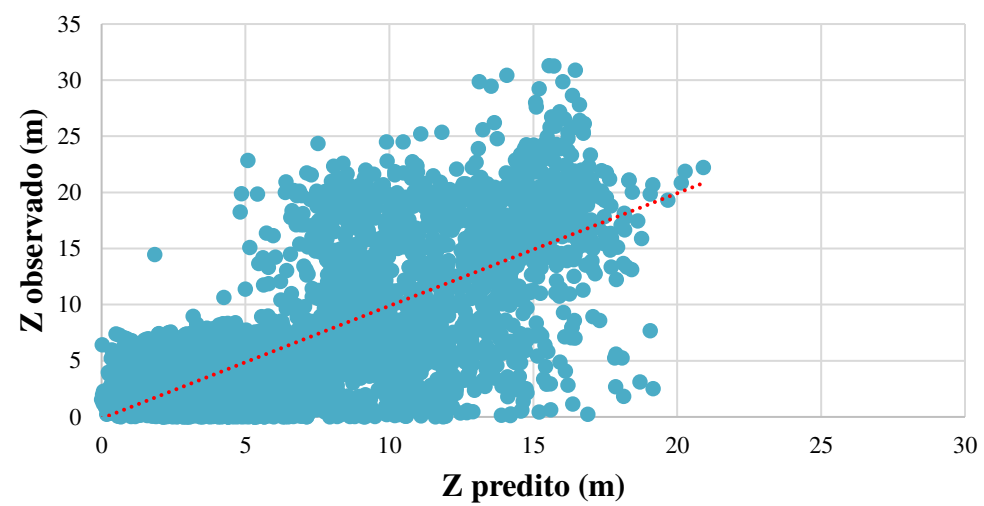
**Histograma das discrepâncias - Landsat Rio Paraguai RNA**

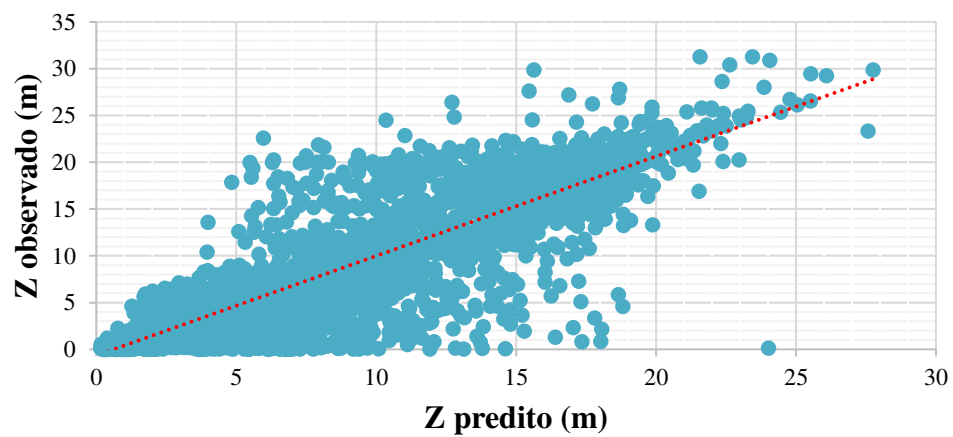
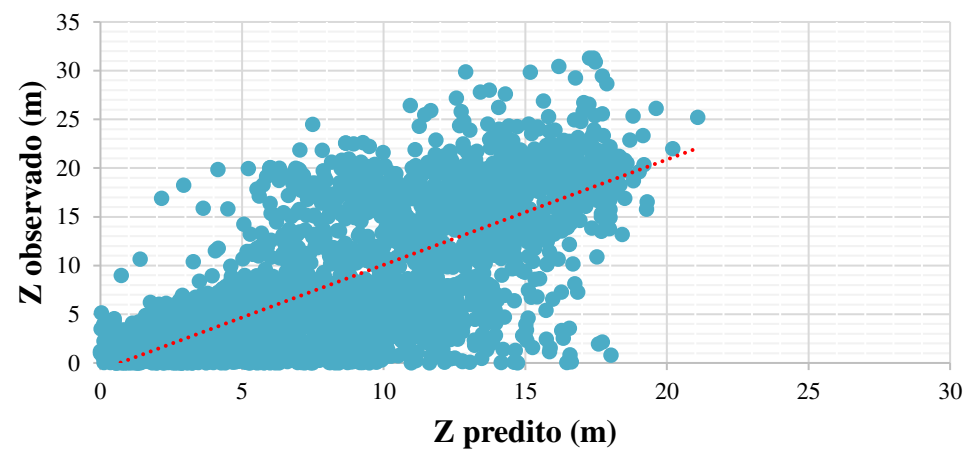


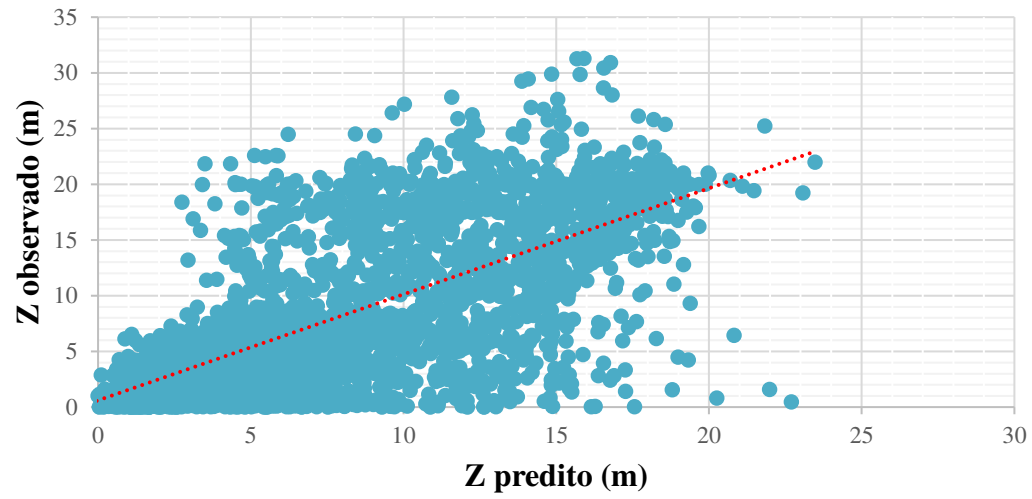
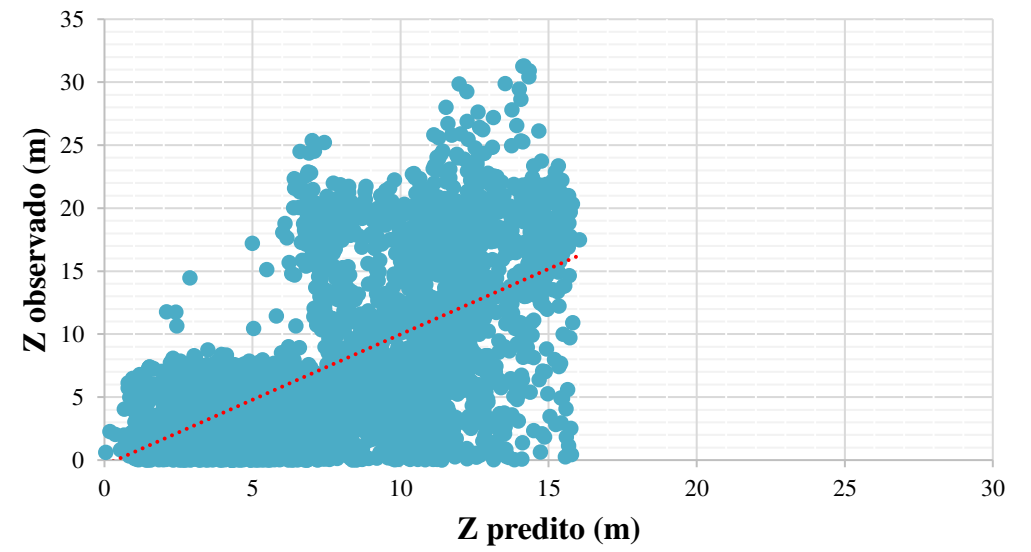






**CBERS - Rio Madeira - SVM****CBERS - Rio Madeira - RNA**

**Landsat - Rio Madeira - RF****Landsat - Rio Madeira - GBM**

**Landsat - Rio Madeira - SVM****Landsat - Rio Madeira - RNA**

## APÊNDICE C – Modelos digitais das discrepâncias

