

## Utilização de técnica de Machine Learning associada a imagem Sentinel 2A para análise da morfologia submersa do Rio Paraguai

Coelho de Andrade, L. (UNIVERSIDADE FEDERAL DE VIÇOSA) ; Oliveira Ferreira, I. (UNIVERSIDADE FEDERAL DE VIÇOSA) ; Amaral e Silva, A. (UNIVERSIDADE FEDERAL DE VIÇOSA) ; Teixeira Gibrim, V. (UNIVERSIDADE FEDERAL DE VIÇOSA) ; Mesquita Santos, F.C. (UNIVERSIDADE FEDERAL DE VIÇOSA)

### RESUMO

Para determinação da morfologia de rios e reservatórios, a principal técnica utilizada consiste no levantamento batimétrico com recurso a sensores acústicos. Todavia, esses procedimentos são de alto custo e demandam tempo para execução. Assim, a utilização do aprendizado de máquina para obtenção da batimetria derivada de satélite está trazendo resultados precisos com o avanço tecnológico e aprimoramento dos sensores, uma vez que esses métodos permitem uma maior produtividade se comparado aos sensores acústicos e podem colaborar para o mapeamento submerso e análise da morfologia de fundo dos corpos hídricos. Nesse sentido, o trabalho em questão utiliza do algoritmo GBM associado a imagem Sentinel 2A para análise da morfologia submersa do rio Paraguai, por meio da aquisição da batimetria desse corpo hídrico. A partir dos resultados encontrados, o GBM obteve uma incerteza vertical na faixa de 0.44 m, indicando o potencial de utilização para análise da morfologia de corpos hídricos.

### PALAVRAS CHAVES

*Batimetria; Batimetria derivada de satélite; Aprendizado de máquina; Gradient Boosting Machine; Sentinel 2A*

### ABSTRACT

To determine the morphology of rivers and reservoirs, the main technique used is the bathymetric survey using acoustic sensors. However, these procedures are costly and time consuming. Thus, the use of machine learning to obtain satellite-derived bathymetry is bringing accurate results with technological advancement and improvement of sensors, since these methods allow greater productivity compared to acoustic sensors and can collaborate for submerged mapping and analysis of bottom morphology of water bodies. In this sense, the work in question uses the GBM algorithm associated with the Sentinel 2A image to analyze the submerged morphology of the Paraguay River, through the acquisition of the bathymetry of this water body. From the results found, the GBM obtained a vertical uncertainty in the range of 0.44 m, indicating the potential use for the analysis of the morphology of water bodies.

### INTRODUÇÃO

A análise da morfologia de rios e reservatórios configura-se como uma tarefa essencial para o aprimoramento da gestão desses locais, podendo contribuir de forma indireta para o desenvolvimento de todo meio ambiente (BIRKETT, 1995). Para isso, são comumente empregados levantamentos batimétricos com recurso a ecobatímetros monofeixe ou multifeixe associados a sistemas GNSS (Global Navigation Satellite System) e sensores inerciais para realização do posicionamento da embarcação (FERREIRA et al. 2015). Com o avanço tecnológico, além da melhoria no âmbito hidrográfico com os sistemas acústicos não tripulados e submergíveis, o emprego do sensoriamento remoto na batimetria também é uma realidade existente na atualidade (FERREIRA et al., 2016). Essa técnica, também chamada de SDB (Satellite Derived Bathymetry), consiste no princípio de penetração da radiação eletromagnética na água em diferentes comprimentos de onda, tendo a profundidade proporcional à radiação que penetra na coluna de água. Destaca-se que além dos diferentes sistemas orbitais e aéreos, existe também uma grande quantidade de métodos abordados na literatura para a predição e análise da batimetria. Autores como Ashphaq et al., (2021) demonstram que existem atualmente abordagens físicas e empíricas para a determinação do SDB, sendo a primeira mais complexa, de difícil aplicação e a segunda mais

simples, baseada puramente em conceitos estatísticos. Os métodos empíricos empregam, principalmente, ferramentas de regressão para avaliação das informações, utilizando também os valores das bandas com dados in situ para os demais cálculos de predição. 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. Com o advento do aprendizado de máquina ou 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. 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 existem poucos estudos nesse âmbito, sobretudo de caráter nacional, empregando áreas de estudo que possuem grande impacto social e econômico pro país, com ferramentas e softwares livres e imagens gratuitas. Nesse sentido, o trabalho teve como objetivo principal a análise da morfologia do rio Paraguai, na região da Isla San Carlos e Isla Piquete-Cambá, com o emprego da batimetria derivada da imagem Sentinel 2A associada ao algoritmo Gradient Boosting Machine (GBM) e a carta náutica disponibilizada pela Marinha do Brasil.

## MATERIAL E MÉTODOS

1. Área de estudo A área de estudo possui 16,37 km<sup>2</sup> e está localizada no rio Paraguai, entre a Isla San Carlos e a Isla Piquete-Cambá, em território Paraguai. Salienta-se que, de acordo com a carta náutica, as profundidades chegam aproximadamente a 8,00 metros. 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. 2. Materiais e métodos Para o desenvolvimento dessa pesquisa foi empregado o sistema Sentinel 2A. Destaca-se que a cena foi adquirida no Earth Explorer de forma gratuita no ano de 2017 para o período de seca do rio Paraguai (entre os meses de abril a outubro), dado que na carta náutica as profundidades são referidas ao nível de redução, que consiste na menor profundidade que o corpo hídrico poderia atingir em um possível período de seca. No que concerne a ortorretificação, sabe-se que os produtos disponibilizados pelo sistema Sentinel 2A (USGS, 2015) já são ortorretificados com o uso de um MDT (Modelo Digital do Terreno). Para correção radiométrica, foi utilizada a técnica DOS (Dark Object Subtraction), com o plug in SCP (Semi-automatic Classification) no software QGIS 3.0.0. Essa técnica 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). Em um momento concomitante à correção da cena, foi realizada a digitalização da carta náutica adquirida 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 o corpo hídrico de forma contínua. Nesse sentido, foi gerado um Modelo Digital de Profundidade (MDP) 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 MDP foi transformado para um arquivo de pontos, o qual foi inserido no software R para o processamento, em conjunto com os valores de radiância para cada banda da imagem Sentinel 2A, bem como as coordenadas UTM X e Y do centróide de cada pixel da imagem. Assim, utilizando o GBM, com o pacote denominado "gbm" foram selecionados, de forma aleatória, 70% da amostra de pontos para o treinamento e os 30% restantes para o teste da predição. Posteriormente, para a avaliação dos resultados, foi efetuado uma análise exploratória das discrepâncias encontradas a partir do valor da profundidade observada e estimada, após cinquenta repetições do algoritmo e foram armazenados os valores das discrepâncias por meio dos estimadores RMSE, MAE e o valor do R<sup>2</sup>.  $MAE = 1/n \sum |z_i - Z_{predi}|$   $RMSE = \sqrt{(1/n) \sum (z_i - z_{predi})^2}$   $R^2 = (\sum (z_{predi} - z_{media})^2) / (\sum (z_i - z_{media})^2)$  Em que z<sub>i</sub> corresponde aos valores

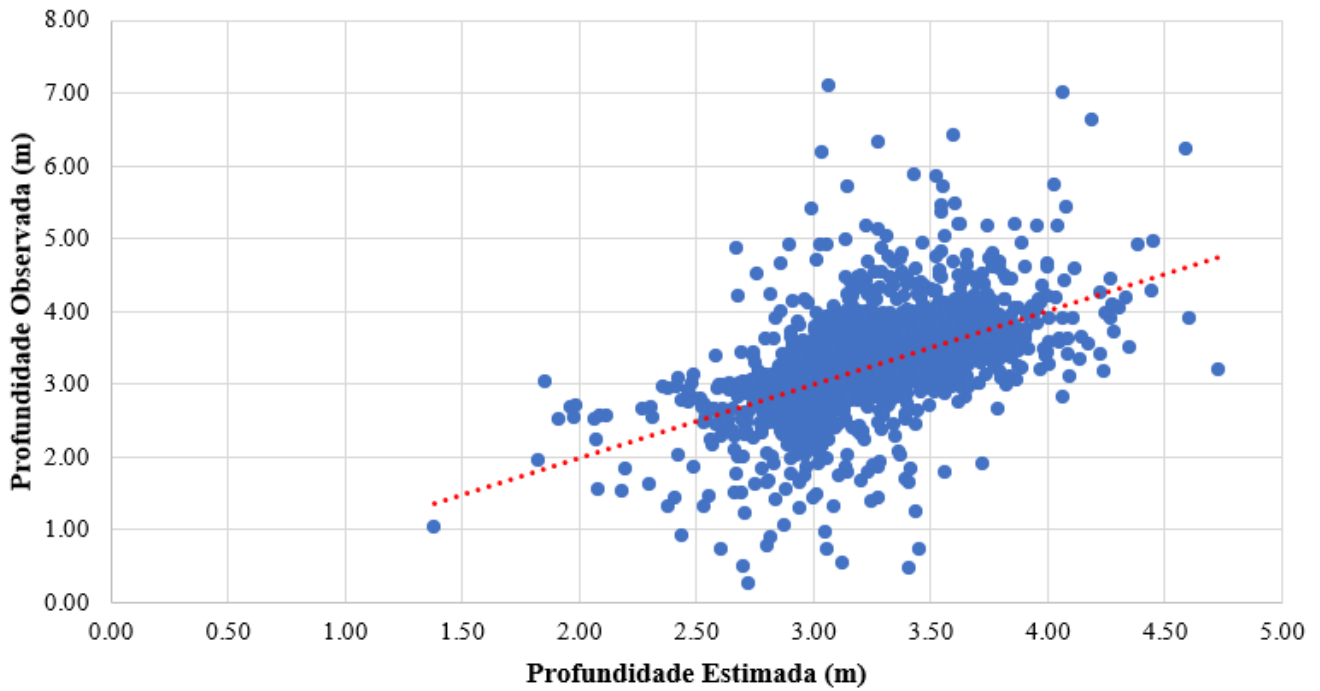
das amostras observadas;  $z_{predi}$ , o valor das amostras preditas;  $z_{media}$ , média dos valores observados e  $n$  o número de amostras. Ademais, para espacialização e melhor visualização dos dados, foram gerados modelos digitais de profundidade para a batimetria estimada utilizando o interpolador krigagem simples, no software ArcGIS 10.5.

## RESULTADOS E DISCUSSÃO

Em um primeiro momento, com o treinamento e predição pelo GBM, foi possível obter os valores das discrepâncias entre as profundidades observadas e preditas e posteriormente a execução de uma análise estatística exploratória das mesmas (Tabela 1). Tabela 1. Análise exploratória das discrepâncias. Satélite Algoritmo Média (m) Erro Padrão Desvio Padrão (m) Curtose Assimetria Sentinel 2A GBM -0.21 0.05 0.34 1.05 -1.02 Essa análise evidencia valores relativamente baixos para as discrepâncias, com baixo desvio e erro padrão. A assimetria negativa indica que a predição da profundidade foi superestimada, dado que a profundidade estimada foi maior que aquela observada, já a curtose, medida que caracteriza o achatamento da curva da função de distribuição de probabilidade, indica que a amostra em questão possui uma distribuição levemente platicúrtica. Como forma de complementar a análise das discrepâncias, foi construído um gráfico de dispersão dos valores observados e preditos, de acordo com a profundidade (Figura 1). A partir do gráfico, pode-se perceber que as maiores dispersões encontram-se nas profundidades observadas menores (até 2,00 m) e maiores (entre 5,00 e 8,00 m), enquanto que para as profundidades médias (2,00 a 5,00 m), a predição apresentou uma menor dispersão. Tal fato pode ser consequência do espalhamento por flutuações de densidade e por partículas que se deslocam do fundo submerso para a camada mais rasa do corpo hídrico. 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. O cálculo da incerteza vertical foi realizado por meio do RMSE, MAE e  $R^2$  para avaliação da correlação. Após as cinquenta repetições do algoritmo, foi calculado a média para os estimadores (Tabela 2). Tabela 2. Valores obtidos para as incertezas verticais. RMSE (m) 0.44 MAE (m) 0.26  $R^2$  0.35 Os valores obtidos para RMSE e MAE demonstram a eficácia da metodologia apresentada. Hedley et al. (2018) também utilizaram de imagens orbitais para a predição da batimetria na costa da Austrália e obtiveram uma predição com uma incerteza semelhante para profundidades de até 15 m. Ressalta-se ainda, que autores como Andrade et al. (2022) utilizaram de redes neurais para a estimativa da profundidade de um reservatório na cidade de Viçosa – MG de aproximadamente 4.00 m e obtiveram uma incerteza vertical de 0.73 m. Sabe-se que o GBM consiste em um algoritmo de árvore de decisão, em que se utiliza de árvores sequenciais e pequenas que tendem a ter maior viés quando comparados a demais algoritmos como o RF (Random Forest) (Zhou et al, 2023). No entanto, o GBM configura-se como uma ferramenta mais rápida e simples para implementação, necessitando de menos tempo para o treinamento e predição. Já o coeficiente de correlação ( $R^2$ ) de 0.35, muito provavelmente é devido ao número de amostras empregada. Autores como Najjar et al. (2022) apresentaram um coeficiente de correlação maior quando empregaram mais pontos no treinamento e predição do algoritmo e Sagawa et al. (2019) obteve menores  $R^2$  quando utilizou menos pontos para o treinamento da predição com o RF e imagens Landsat 8. Posteriormente, para gerar o modelo digital (Figura 2), 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 2. A) MDPs de referência B) MDP da profundidade estimada com GBM e C) Modelo Digital das Discrepâncias. A partir dos MDPs, pode-se perceber que a imagem orbital, associada ao algoritmo GBM foi capaz de identificar os principais pontos de profundidade baixa, média e alta, de forma geral, seguindo parcialmente a morfologia submersa do modelo de referência. De acordo com o Modelo Digital das Discrepâncias em C), nota-se as maiores discrepâncias nos locais mais profundos (~6.00 m), enquanto que nas profundidades baixas e médias (até 3.00 m) as discrepâncias foram consideravelmente baixas.

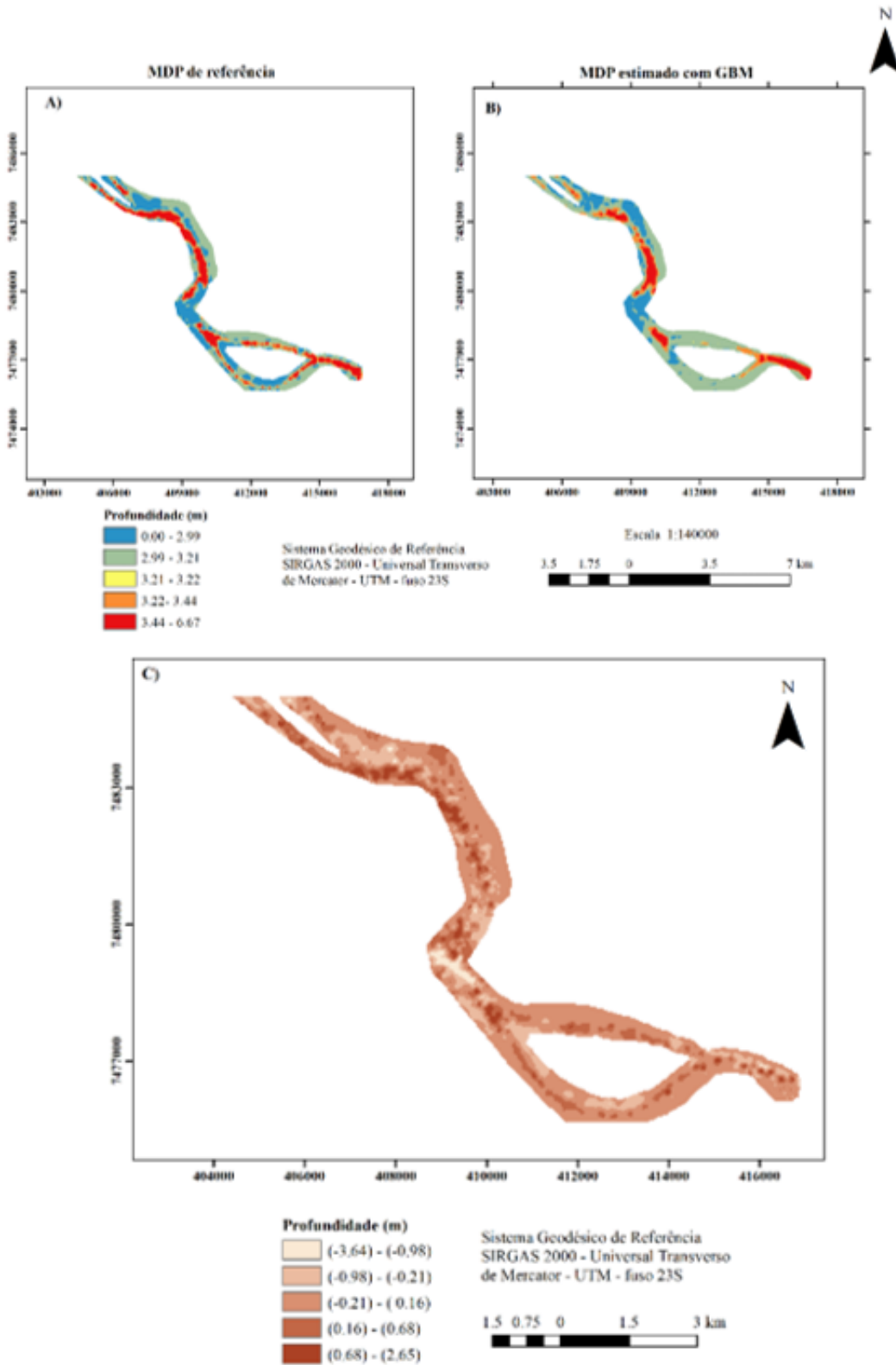
Gráfico de dispersão entre as profundidades observadas e estimadas.

### Gráfico de dispersão



A figura contém o gráfico de dispersão entre as profundidades observadas e estimadas com o algoritmo GBM

A) MDPs de referência B) MDP da profundidade estimada com GBM e C) Mod



A figura contém em A) o MDP de referência da profundidade, B) o MDP estimado com o GBM e em C) o Modelo das discrepâncias entre os dois modelos A e B.

## CONSIDERAÇÕES FINAIS

A partir dos trabalhos realizados, pode-se concluir que o emprego de imagens Sentinel 2A, associadas ao algoritmo GBM, pode oferecer um bom desempenho para determinação da batimetria derivada de satélite, sobretudo para análises ambientais preliminares acerca da morfologia submersa. Destaca-se ainda, que o estudo em questão também pode facilitar de forma significativa a localização de perigos para a navegação. O emprego de imagens orbitais como a Sentinel 2A para a predição da batimetria é algo amplamente estudado e os métodos utilizados nesse trabalho evidenciam sua aplicabilidade também para o estudo de rios navegáveis. O emprego dessa metodologia de forma periódica, além da utilização de mais imagens orbitais, associadas a índices como o de turbidez da água, material particulado em suspensão e até mesmo o NDCI (Normalized Difference Chlorophyll Index) podem permitir avanços significativos no contexto da batimetria derivada de satélites, bem como contribuir em estudos de análises morfológicas.

## REFERÊNCIAS BIBLIOGRÁFICA

1. 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.
2. 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.
3. 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.
4. 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.
5. 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.
6. CHAVEZ, P.S.; JR. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* 1988, 24, 459-479.
7. 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.
8. 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.
9. FERREIRA, Í. O.; RODRIGUES, D. D.; SANTOS, G. R. Coleta, processamento e análise de dados batimétricos. 1ª ed. Saarbrucken: Novas Edições Acadêmicas, v. 1, 100f. Res. 27, 7-17. 2015. doi: 10.1007/s11001-005-0267-x
10. 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).2016.
11. 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>
12. LYZENGA, D. R. Passive remote sensing techniques for mapping water depth and bottom features. *Applied Optics*, 17(3), 379-383.1978.
13. NAJAR, M. A.; BENSILHA, R.; BENNIQUI, Y. E.; THOUMYRE, G.; ALMAR, R.; BERGSMAN, 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.
14. 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.
15. 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.

16. 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.
17. 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.
18. 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.
19. 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.* 2023, 15, 393. <https://doi.org/10.3390/rs15020393>.