

XXVI SIMPÓSIO BRASILEIRO DE RECURSOS HIDRÍCOS

ANÁLISE COMPARATIVA DE BATIMETRIA UTILIZANDO SENSORIAMENTO REMOTO E MÉTODOS CONVENCIONAIS

Raphael Gil Soares Ferreira¹; Igor da Silva Rocha Paz²; Raquel Aparecida Abrahão Costa e Oliveira³

ABSTRACT

This study aims to estimate the bathymetry of the Port of Pecém (CE) using Sentinel-2 satellite imagery and to compare the results with conventional multibeam bathymetric surveys. Three predictive models were tested: linear regression, nonlinear regression, and Random Forest. These approaches were used to correlate spectral data with measured water depths. The Sentinel-2 images were selected due to their free availability and 10-meter spatial resolution, suitable for shallow coastal waters. Different band combinations were applied to improve correlation and accuracy. Data were split for training and testing, using cross-validation to enhance model robustness and generalization. Results showed that linear and nonlinear regressions had low performance, with R^2 values of 0.198 and 0.178, respectively. In contrast, the Random Forest model achieved a significantly better result, with an R^2 of approximately 0.78, demonstrating high reliability in depth estimation using spectral reflectance alone. The findings indicate that machine learning techniques, especially Random Forest, can serve as effective, low-cost alternatives to traditional methods, providing broader spatial coverage and accurate bathymetric mapping.

Keywords – Bathymetry. Remote Sensing. Satellite.

RESUMO

Este estudo tem como foco a estimativa da batimetria do Porto do Pecém (CE) por meio de imagens de satélite Sentinel-2, comparando os resultados com os obtidos pelo método convencional de levantamento batimétrico multifeixe. Foram testados três modelos preditivos: regressão linear, regressão não linear e Random Forest. A análise utilizou diferentes combinações de bandas espectrais, com parte dos dados destinada ao treinamento e outra parte à validação cruzada dos modelos. Os modelos estatísticos foram avaliados com base na capacidade de correlacionar os dados espectrais com profundidades reais. Os resultados indicaram baixo desempenho para os métodos de regressão linear e não linear, com coeficientes de determinação (R^2) de 0,198 e 0,178, respectivamente. Em contrapartida, o modelo Random Forest apresentou desempenho superior, com R^2 de aproximadamente 0,78, revelando alta confiabilidade na estimativa das profundidades. Dessa forma, o estudo demonstra que o uso de sensoriamento remoto aliado a técnicas de aprendizado de máquina pode representar uma alternativa viável e eficiente para mapeamento batimétrico, oferecendo cobertura espacial ampliada e redução de custos operacionais.

Palavras-chave – Batimetria. Sensoriamento Remoto. Satélite.

INTRODUÇÃO

A batimetria desempenha um papel fundamental na segurança da navegação, na gestão portuária, no planejamento de infraestrutura costeira e em estudos ambientais relacionados à dinâmica sedimentar e à erosão costeira (LI *et al.*, 2023). Tradicionalmente, levantamentos batimétricos são realizados por meio de ecobatímetros multifeixe (MBES – Multibeam Echo Sounder), que oferecem alta precisão e resolução (ŠILJEG *et al.*, 2022). No entanto, esses métodos envolvem elevados custos

¹) Mestrando em Engenharia de Defesa, Instituto Militar de Engenharia (IME), Rio de Janeiro, RJ, soares.raaphael@ime.eb.br

²) Professor no Instituto Militar de Engenharia (IME), Rio de Janeiro, RJ, igorpaz@ime.eb.br

³) Professora no Instituto Militar de Engenharia (IME), Rio de Janeiro, RJ, raquel.oliveira@ime.eb.br

operacionais, demanda logística complexa e são limitados a áreas previamente acessadas por embarcações especializadas (GAGG, 2016).

Diante dessas limitações, o uso de sensoriamento remoto tem se consolidado como uma alternativa viável para estimar a profundidade de águas costeiras rasas, com custo reduzido e cobertura espacial ampliada (VOJINOVIC *et al.*, 2013). As imagens orbitais, em especial aquelas captadas por sensores com resolução espacial adequada e múltiplas bandas espectrais, como o Sentinel-2, permitem explorar relações empíricas entre a reflectância da coluna d'água e as profundidades reais do fundo (JENSEN; EPIPHANIO, 2009).

Neste contexto, o presente estudo tem como objetivo realizar uma análise comparativa entre métodos convencionais de levantamento batimétrico e estimativas obtidas por sensoriamento remoto, a partir de imagens do satélite Sentinel-2, aplicadas ao Porto do Pecém (CE). São testados três modelos – regressão linear, regressão não linear e Random Forest – cujos resultados são avaliados com base em sua capacidade preditiva e correlação com os dados obtidos em campo fornecidos por de Andrade Neto *et al.* (2024).

MATERIAIS E MÉTODOS

Local do Estudo

A pesquisa foi realizada na região do Porto do Pecém, no litoral do Ceará, devido à disponibilidade de dados batimétricos convencionais e imagens de satélite da área. Tais dados batimétricos de referência foram obtidos por ecobatímetro multifeixe, classificados pela marinha do Brasil como Classe A.

Processamento de Imagens e Composição Espectral

O processamento das imagens Sentinel-2 foi conduzido no ambiente de geoprocessamento QGIS, com suporte ao ArcGIS Pro e Google Colab para manipulação de dados espectrais. Foram utilizadas imagens com resolução de 10 metros das bandas B2 (azul), B3 (verde), B4 (vermelho) e B8 (infravermelho próximo – NIR).

A metodologia baseou-se na criação de composições multiespectrais e no cálculo de índices espectrais com relevância hidrológica, em especial o NDWI (*Normalized Difference Water Index*), definido pela fórmula (eq. 1):

$$NDWI = \frac{B3 - B8}{B3 + B8} \quad (1)$$

O índice NDWI foi usado para isolar áreas de água e melhorar a correlação entre reflectância espectral e profundidade (RODRIGUES *et al.*, 2018).

Cada pixel da imagem foi associado à profundidade correspondente do levantamento multifeixe, utilizando uma grade regular (*grid*) com amostragem espacial compatível com a resolução da imagem. A média da reflectância de cada banda por pixel foi extraída para compor a matriz de treinamento dos modelos.

Modelagem Preditiva de Profundidade

A modelagem preditiva consistiu na construção de modelos estatísticos para estimar a profundidade a partir da reflectância espectral das imagens. Três métodos distintos foram aplicados:

Segundo Chein (2019), Regressão Linear utiliza uma equação (eq. 2) do tipo:

$$Y = A + BX + \varepsilon \quad (2)$$

Onde Y representa a profundidade e Xn são os valores de reflectância das bandas espectrais. A regressão linear avalia a relação direta entre os valores espectrais e as profundidades medidas, assumindo independência entre as variáveis.

De acordo com Facure (2017), a regressão não linear permite utilizar diversos tipos de equações, modelos polinomiais, exponenciais e logarítmicos foram utilizados para capturar possíveis curvaturas na relação entre profundidade e reflectância. Foram testadas 3 funções não lineares, uma equação polinomial de 2ª ordem (eq. 3), uma equação logarítmica (eq. 4), e por fim uma equação exponencial (eq. 5):

Polinomial de 2ª ordem

$$Y = A + B1 * X + B2 * X^2 + B3 * X^3 + \dots + \varepsilon \quad (3)$$

Logarítmica

$$Y = A + B * \ln(X) + \varepsilon \quad (4)$$

Exponencial

$$Y = A + B * \ln(X) + \varepsilon \quad (5)$$

A escolha dos modelos foi baseada na aderência visual aos dados e no erro quadrático dos resíduos, buscando representar variações não lineares típicas de ambientes aquáticos.

Já o modelo Random Forest desenvolvido por Breiman (2001), foi implementado com 500 árvores, usando o pacote scikit-learn. Este método de aprendizado de máquina do tipo ensemble combina diversas árvores de decisão, treinadas com subconjuntos aleatórios dos dados (*bootstrap*), e seleciona subconjuntos aleatórios de variáveis para divisão em cada nó.

Durante a predição, o valor final da profundidade é dado pela média das predições de todas as árvores. Essa abordagem é robusta a multicolinearidade e outliers, e não requer normalização dos dados.

Avaliação da Precisão dos Modelos

A performance dos modelos foi avaliada com base em três métricas estatísticas:

- **Erro Médio Absoluto (MAE) (Eq. 6)**

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}_i| \quad (6)$$

Segundo Filho (2023), tal erro é responsável por medir a média das diferenças absolutas entre valores reais e estimados.

- **Raiz do Erro Quadrático Médio (RMSE) (Eq. 7)**

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

De acordo com Filho (2022), o erro quadrático médio é responsável por penalizar mais fortemente grandes erros de predição.

Coefficiente de Determinação (R^2) (Eq. 8)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (8)$$

Segundo Chein (2019), o coeficiente de determinação é responsável por expressar a proporção da variância explicada pelo modelo. Valores próximos a 1 indicam boa aderência.

Onde cada termo tem um significado para o resultado final.

y_i : valor real

\hat{y}_i : valor previsto

n : número total de observações

\bar{y} : média dos valores reais

A avaliação foi feita com validação cruzada (80% para treino, 20% para teste), repetida em cinco iterações com randomização dos dados, garantindo robustez estatística e generalização dos resultados.

RESULTADOS

Os resultados apresentados nesta seção refletem o desempenho dos três métodos de estimativa batimétrica aplicados à área do Porto do Pecém-CE, a partir da análise multiespectral das imagens Sentinel-2 e dos dados de batimetria convencional multifeixe utilizados como referência.

A comparação entre os modelos foi realizada com base nas métricas MAE, RMSE e R^2 , bem como por meio de mapas temáticos e gráficos de dispersão que ilustram visualmente o ajuste dos modelos às profundidades reais.

Regressão Linear

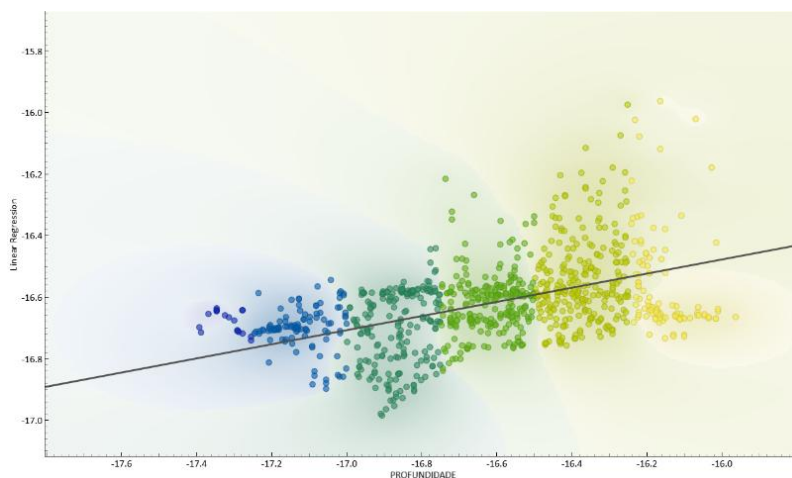
O modelo de regressão linear apresentou desempenho modesto. A relação entre os valores espectrais das bandas e a profundidade foi fraca, resultando em um ajuste de baixa qualidade. Os resultados estatísticos para o conjunto preditos serão exibidos na Tabela 1:

Tabela 1 – Dados estatísticos Regressão Linear - Preditos

Modelo	RMSE	MAE	R ²
Regressão Linear	0,274	0,217	0,198

O gráfico de dispersão entre profundidade observada e estimada mostrou elevada dispersão, com tendência a subestimar profundidades profundas e superestimar profundidades muito rasas, como mostra a Figura 1.

Figura 1 – Gráfico de dispersão da regressão linear em relação a profundidade real encontrada



Além de todos os dados estatísticos obtidos foram criados mapas para uma interpretação visual, facilitando o entendimento do modelo e sua qualidade. Foram criados 2 mapas para ambos os modelos. Mapa 1 (Figura 2) é o mapa das profundidades previstas pelo modelo, já o Mapa 2 (Figura 3) apresenta as diferenças entre as profundidades encontradas pelo modelo e as encontradas pela batimetria convencional.

Figura 2 – Mapa de profundidades geradas pelo modelo de Regressão Linear

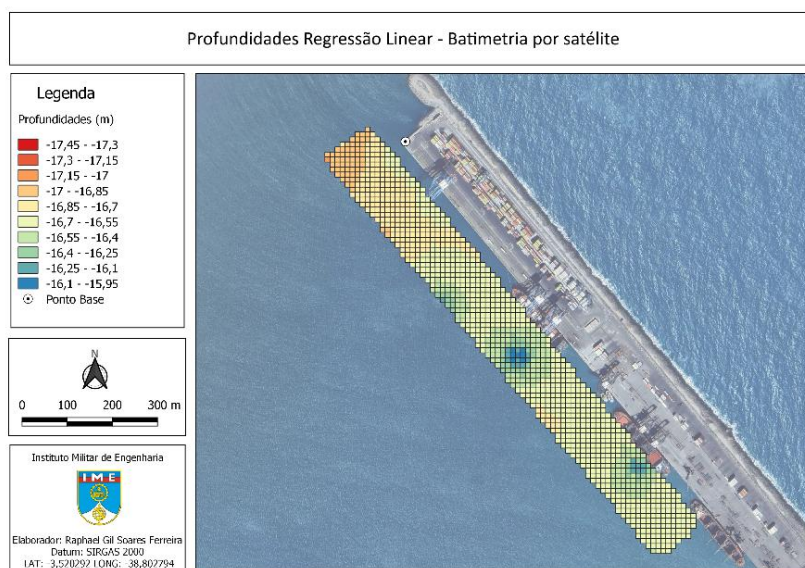
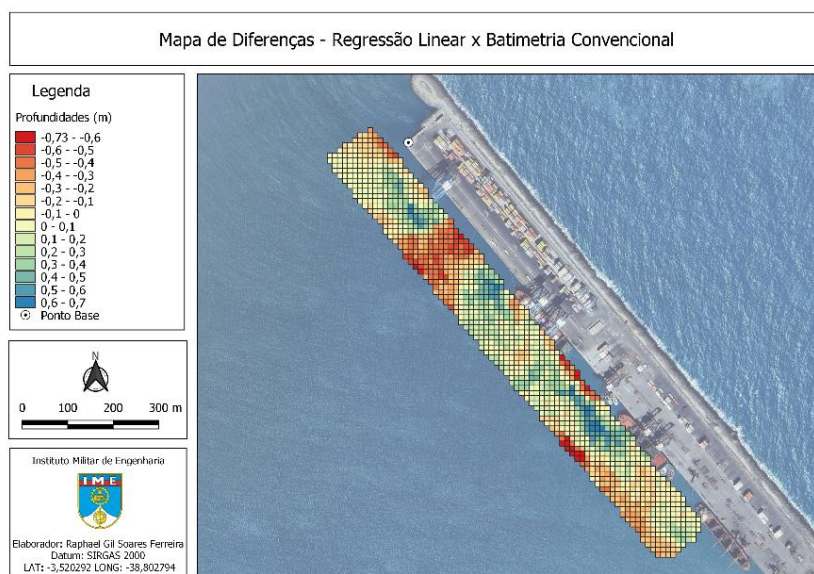


Figura 3 – Mapa de diferenças entre as profundidades encontradas na regressão linear e as profundidades reais



Regressão não linear

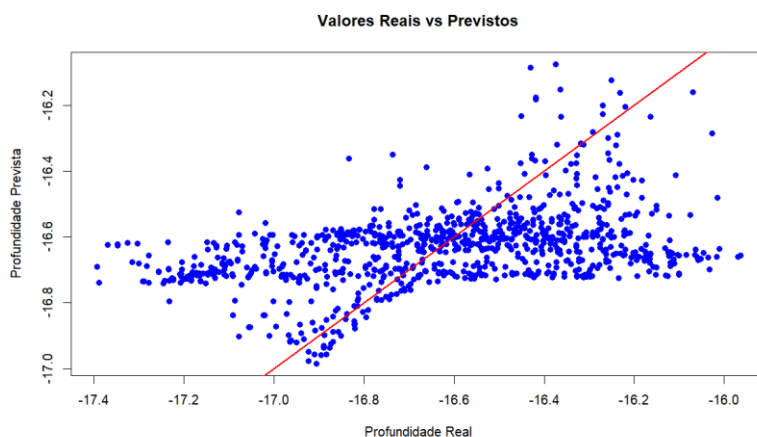
A regressão não linear, utilizando modelo polinomial, buscou capturar relações curvas entre reflectância e profundidade. Ainda assim, o desempenho foi bem similar ao da regressão linear, indicando que a relação espectral é complexa e não facilmente modelável com funções simples. Os parâmetros estatísticos serão apresentados na Tabela 2 a seguir.

Tabela 2 – Dados estatísticos Regressão não Linear - Preditos

Modelo	RMSE	MAE	R ²
Regressão não linear	0,277	0,220	0,178

Observou-se que o modelo não linear tendeu a gerar ajustes instáveis nas extremidades dos dados (profundidades muito rasas ou profundas), demonstrando uma correlação fraca, como mostra o gráfico de dispersão apresentado a seguir na Figura 4.

Figura 4 – Gráfico de dispersão da regressão não linear em relação a profundidade real encontrada



Foram desenvolvidos mapas comparativos das profundidades, apresentando diferentes análises dos resultados. O primeiro mapa criado para esse modelo (Figura 5) exibe as profundidades estimadas pela regressão não linear, enquanto o segundo (Figura 6) ilustra as diferenças entre os valores obtidos pela regressão e os dados da batimetria convencional.

Figura 5 – Mapa de profundidades geradas pelo modelo de Regressão não Linear

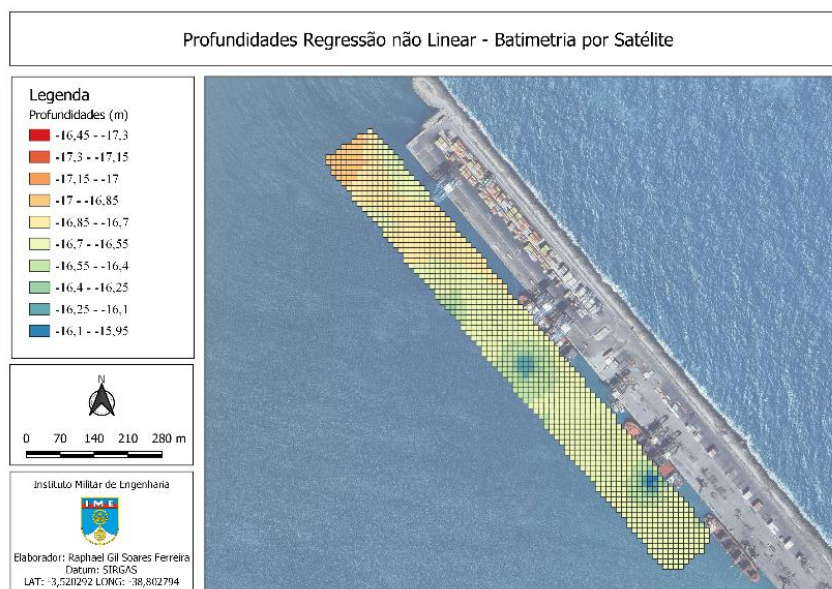
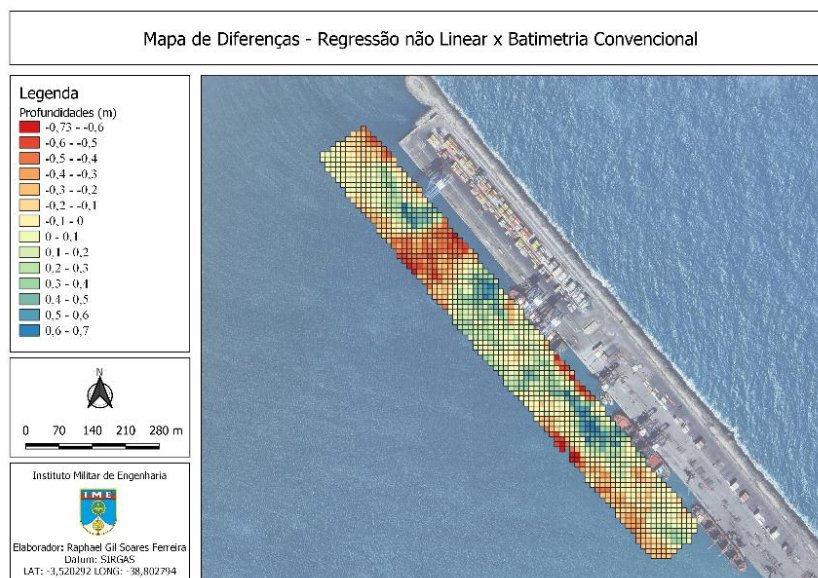


Figura 6 – Mapa de diferenças entre as profundidades encontradas na regressão não linear e as profundidades reais



Random Forest

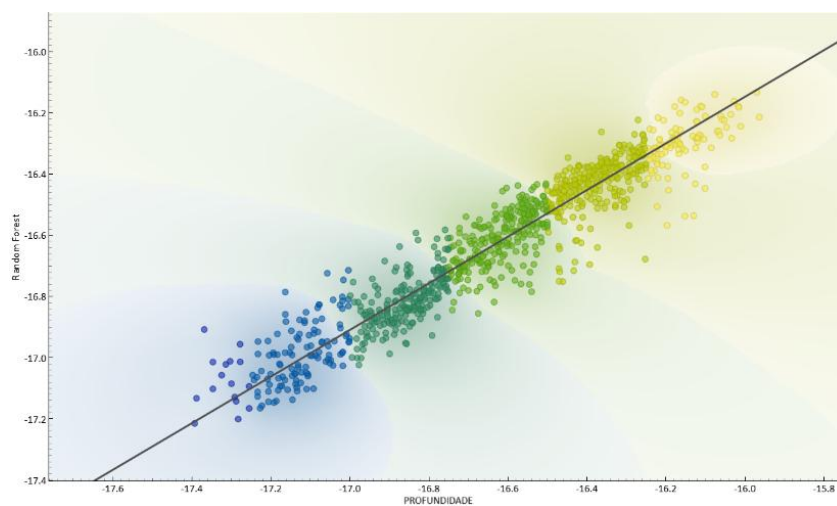
O modelo Random Forest apresentou **desempenho significativamente superior** em todos os indicadores de qualidade. A capacidade do algoritmo de capturar relações não lineares e interações complexas entre as bandas espectrais contribuiu para uma predição mais precisa e consistente, como pode ser vista pelos parâmetros estatísticos presentes na Tabela 3.

Tabela 3 – Dados estatísticos Random Forest- Preditos

Modelo	RMSE	MAE	R ²
Random Forest	0,143	0,104	0,780

O gráfico de dispersão (Figura 7) demonstrou boa aderência dos valores preditos às profundidades reais, com baixo viés, como mostra a reta de correlação.

Figura 7 – Gráfico de dispersão da regressão não linear em relação a profundidade real encontrada



Com base nos resultados obtidos, foi possível realizar uma comparação visual por meio da criação de mapas, facilitando a interpretação mais precisa dos dados. Foram gerados: um mapa com as profundidades estimadas pelo modelo Random Forest (Figura 8) e um mapa com as diferenças entre as profundidades estimadas e os valores reais (Figura 9).

Figura 8 – Mapa de profundidades geradas pelo modelo de Regressão não Linear

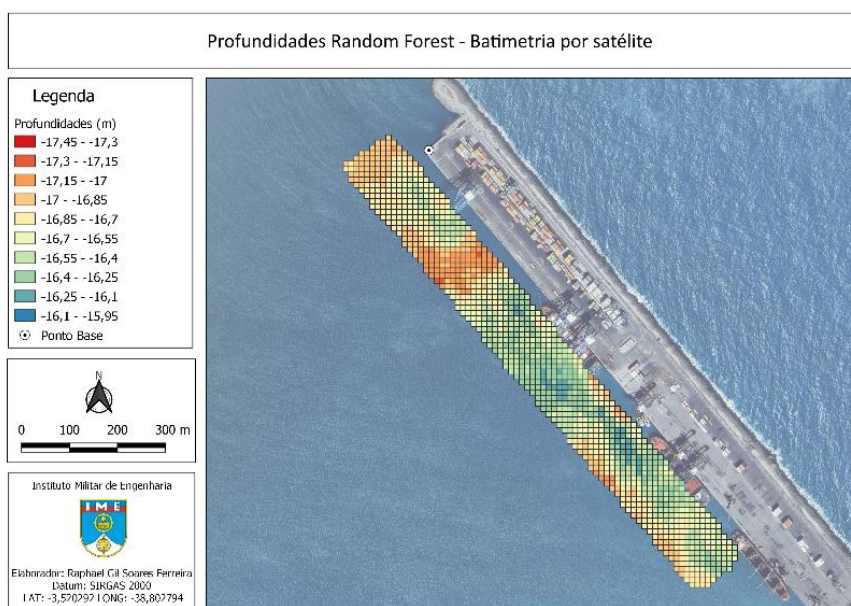
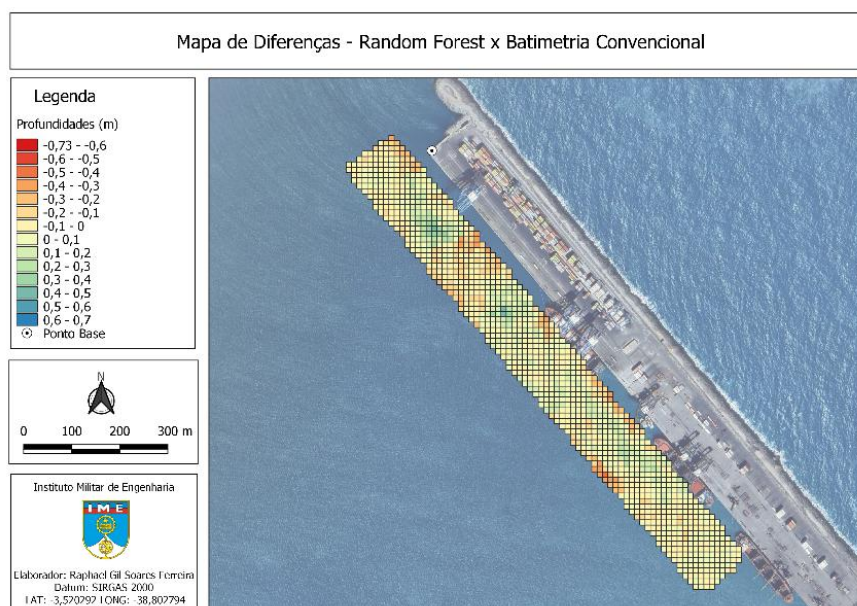


Figura 9 – Mapa de diferenças entre as profundidades encontradas na regressão não linear e as profundidades reais



CONCLUSÃO

O estudo avaliou a viabilidade do uso de sensoriamento remoto por imagens de satélite, especificamente do Sentinel-2, na estimativa de profundidades em ambientes costeiros, comparando três métodos de modelagem: regressão linear, regressão não linear e Random Forest. A regressão linear e a não linear apresentaram desempenhos insatisfatórios, com baixos coeficientes de determinação ($R^2 \approx 2\%$ e $1,8\%$, respectivamente), o que evidenciou limitações na capacidade desses modelos de capturar a complexa relação entre as bandas espectrais e as profundidades medidas in loco.

Em contrapartida, o modelo Random Forest demonstrou resultados significativamente superiores, atingindo um R^2 de aproximadamente 0,78. Essa abordagem se mostrou eficaz mesmo em áreas com grandes variações de profundidade, sendo capaz de fornecer estimativas precisas e consistentes, o que reforça o potencial do aprendizado de máquina na substituição parcial ou total dos métodos batimétricos convencionais em determinadas situações.

Apesar dos avanços, o estudo reconhece limitações, como a não consideração de fatores ambientais (e.g., turbidez, salinidade, maré) e a dependência da qualidade das imagens. Recomenda-se a incorporação de variáveis complementares e o teste de novos modelos, como redes neurais artificiais, em pesquisas futuras. Assim, este trabalho contribui para o avanço da batimetria remota, oferecendo uma alternativa viável, econômica e eficiente para apoio à gestão portuária e à segurança da navegação.

AGRADECIMENTOS

O autor Igor da Silva Rocha Paz agradece à Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) pelo financiamento da pesquisa realizada no âmbito do Programa Jovem Cientista do Nosso Estado (JCNE).

REFERÊNCIAS

- BREIMAN, L. (2001). Random Forests. *Machine Learning*, v 45, 5–32 p.
- CHEIN, F. (2019). Introdução aos modelos de regressão linear: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas. Enap, 77 p.
- DE ANDRADE NETO, W.P.; PAZ, I.d.S.R.; OLIVEIRA, R.A.A.C.e.; DE PAULO, M.C.M. Comparison of the vertical accuracy of satellite-based correction service and the PPK GNSS method for obtaining sensor positions on a multibeam bathymetric survey. *Sci. Rep.* 2024, 14, 11104.
- FACURE, M. (2017). Regressão Polinomial. Disponível em: <https://matheusfacure.github.io/2017/02/26/regr-poli/>
- FILHO, M. (2023). RMSE (Raiz Do Erro Quadrático Médio) Em Machine Learning. Disponível em: <https://mariofilho.com/rmse-raiz-do-erro-quadratico-medio-em-machine-learning/#:~:text=Quanto%20menor%20o%20RMSE%2C%20melhor,caso%20de%20previ s%C3%A3o%20de%20temperatura.>
- GAGG, G. (2016). Apostila de Levantamentos Hidrográfico - Noções Gerais. Rio Grande do Sul, Brasil: UFRGS – Universidade Federal do Rio Grande do Sul.
- JENSEN, J. R.; EPIPHANIO, J. C. N. (2009). Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres. Parêntese Editora. 598 p.
- LI, Z., PENG, Z., ZHANG, Z., CHU, Y., XU, C., YAO, S., GARCÍA-FERNANDÉZ, Á.F., ZHU, X., YUE, Y., LEVERS, A., ZHANG, J., MA, J. (2023). Exploring modern bathymetry: A comprehensive review of data acquisition devices, model accuracy, and interpolation techniques for enhanced underwater mapping. *Frontiers in Marine Science*, 10, 1178845.
- RODRIGUES, W. B.; MORAIS, F. de; PASCHOAL, L. G. (2008). Índice de diferença normalizada da água (ndwi) calculado para estações chuvosas e secas na bacia do córrego barreiro, lagoa da confusão – to. *Simpósio Brasileiro de Geografia Física Aplicada*, v. 1.
- ŠILJEG, A., MARIC, I., DOMAZETOVIC, F., CUKROV, N., LOVRIC, M., & PANDA, L. (2022). Bathymetric survey of the St. Anthony Channel (Croatia) using Multibeam Echosounders (MBES)—A new methodological semi-automatic approach of point cloud post-processing. *Journal of marine science and engineering*, 10(1), 101.
- VOJINOVIC, Z.; ABEBE, Y. A.; RANASINGHE, R.; VACHER, A.; MARTENS, P.; MANDL, D. J.; FRYE, S. W.; ETTINGER, E. van; ZEEUW, R. de. (2013). A machine learning approach for estimation of shallow water depths from optical satellite images and sonar measurements. *Journal of Hydroinformatics*, v. 15, n. 4.