Vietnam Journal of Earth Sciences, 48(2), 167−193, https://doi.org/10.15625/2615-9783/24020

Vietnam Journal of Earth Sciences
https://vjs.ac.vn/index.php/jse

Machine learning approaches for satellite-derived bathymetry in tropical coastal waters: A comparative study from Nha Trang marine protected area, Vietnam

Nguyen Trinh Duc Hieu[1*], Nguyen Hao Quang[2,3], Tri Nguyen-Quang[4], Tran Duc Dien[1], Vo Thi Ha[1], Nguyen Phuong Lien[1], Phuong Lan Nguyen[5], Nguyen Dang Huyen Tran[6], Gorin Sergey[7], Le Van Dan[8], Dang Ngoc Thi Giang[8], Ha Nam Thang[8*]

[1]*Coastal Branch of Joint Vietnam-Russia Tropical Science and Technology Research Center, Khanh Hoa, Vietnam*
[2]*Laboratory of Environmental Sciences and Climate Change, Institute for Computational Science and Artificial Intelligence, Van Lang University, Ho Chi Minh City, Vietnam*
[3]*Faculty of Environment, School of Technology, Van Lang University, Ho Chi Minh City, Vietnam*
[4]*Biofluids and Biosystems Modelling Laboratory (BBML), Dalhousie University, DAC, Truro, NS B2N 5E3, Canada*
[5]*Faculty of Information Technology and Semiconductor, Thai Binh Duong University, Khanh Hoa, Vietmam*
[6]*Department of Academic Affairs, Thai Binh Duong University, Khanh Hoa, Vietnam*
[7]*Lomonosov Moscow State University Marine Research Center, Moscow, Russia*
[8]*University of Agriculture and Forestry, Hue University, Hue City, Vietnam*

Received 05 September 2025; Received in revised form 12 December 2025; Accepted 29 December 2025

ABSTRACT

Bathymetry mapping plays a critical role in coastal zone management, marine conservation, and navigation safety. With the increasing availability of high-resolution satellite imagery, such as PlanetScope (3−5 m), remote sensing-based bathymetry retrieval offers a cost-effective and scalable alternative to traditional *in-situ* surveys. This study explores the capability of PlanetScope imagery to retrieve a wide range of bathymetry (-0.5 − ~ -40 m) in the southern area of the Nha Trang Marine Protected Area (MPA), Vietnam - an ecologically significant and dynamic coastal region. We conduct a comprehensive comparison between traditional approaches, including the Stumpf ratio model and Multiple Linear Regression (MLR), and a suite of advanced machine learning (ML) algorithms, including Random Forest (RF), Support Vector Machine (SVM), Light Gradient Boosting Machine (LGBM), Extreme Gradient Boosting (XGB), CatBoost (CB), and Gradient Boosting (GB). Among these, RF achieved the highest performance with an R^2 of 0.85, RMSE of 2.66 m, and MAE of 1.85 m, significantly outperforming the Stumpf model (R^2 = 0.29) and MLR (R^2 = 0.57). This study represents one of the most extensive model comparisons to date for satellite-derived bathymetry using PlanetScope data, offering a benchmark for future applications in tropical coastal environments. Results underscore the potential of machine learning to advance spatially detailed and accurate bathymetric mapping from space.

Keywords: PlanetScope, Nha Trang MPA, bathymetry, machine learning.

1. Introduction

Bathymetric mapping is a fundamental tool for coastal zone management, marine habitat conservation, and navigational safety. Traditional depth measurement methods, whether shipborne echo sounder or airborne LiDAR, can achieve high positional and depth

*Corresponding author, Email: nguyentrinhduchieu@gmail.com; hanamthang@hueuni.edu.vn

accuracy. LiDAR-based bathymetry, as an emerging technology, provides high-resolution, precise depth measurements using airborne laser pulses that can penetrate clear coastal waters, making it particularly valuable for mapping shallow and nearshore zones with complex bottom structures (Kujawa & Remondino, 2025). However, their spatial coverage is limited, and their application over large or remote areas is often constrained by high operational costs, safety risks, restricted survey extents, and logistical challenges (Gabr et al., 2020; Kalybekova, 2025; Liu et al., 2024). Many shallow coastal environments still rely on outdated or sparsely distributed bathymetric information from previously published nautical charts. Shipborne acoustic techniques, in the form of single- or multi-beam echo sounders, require specialized equipment and trained personnel and are further restricted by vessel draft and safe navigation zones (Gülher & Alganci, 2023b).

Recent advances in satellite remote sensing have demonstrated significant potential to address these challenges by providing spatially extensive and temporally frequent observations of shallow coastal waters. Satellite-derived bathymetry (SDB), in particular, has emerged as a rapid, cost-effective, and globally scalable approach for generating high-resolution seafloor topography. Satellite-derived bathymetry (SDB) benefits from both multispectral and hyperspectral imagery, each contributing unique strengths to depth retrieval. Multispectral sensors (e.g., Sentinel-2, PlanetScope) offer frequent coverage and suitable spectral bands for optical depth estimation, making them effective for large-area, operational mapping (He et al., 2024). Hyperspectral sensors, in addition, provide hundreds of narrow bands that improve water-column correction and bottom-type discrimination, thereby increasing depth-retrieval accuracy in an optically complex environment (Ye et al., 2025). These data sources allow SDB to complement traditional hydrographic surveys by expanding spatial coverage and increasing temporal update capacity (Hodúl et al., 2018). With its advantages and growing maturity, SDB is increasingly recognized as a promising tool for modern bathymetric mapping and coastal zone management (Gülher & Alganci, 2023b). Various satellite-derived bathymetry methods are currently available, with their accuracy influenced by factors including algorithm performance, spectral band selection, and sample distribution (Chu et al., 2023). A chronological review of empirical SDB studies reveals a clear evolution in methodological approaches. Early work in the late 20th century relied on foundational techniques, including the physics-based models of Lyzenga (1978) and Lyzenga et al. (2006), as well as semi-empirical approaches derived from the Depth of Penetration (DOP) concept, which established the optical basis for relating radiance attenuation to depth. Subsequent studies increasingly adopted simplified empirical methods, most notably linear or logarithmic band-ratio algorithms (Stumpf et al., 2003), to estimate depth from multispectral imagery, thereby improving robustness and reducing calibration requirements. In recent years, however, there has been a marked transition toward machine learning and deep learning techniques, driven by their superior capability to model complex data relationships and their potential to deliver higher accuracy (Liu et al., 2024).

Machine learning (ML), a branch of artificial intelligence, utilizes historical data to improve prediction and decision-making. Its strength in handling large datasets from satellite and remote sensing imagery has led to widespread use in classification and regression tasks. In recent years, machine learning techniques have gained increasing attention for SDB applications due to their potential for high accuracy and broad applicability. Moeinkhah et al. (2019) applied

the Random Forest (RF) model to Landsat-8 data and achieved low error (RMSE = 1.253 m, MAE = 0.766 m) in shallow waters (< 5 m), though errors increased significantly beyond 10 m depth. Eugenio et al. (2021) conducted a comprehensive comparison of nine regression and machine learning models using WorldView-2/3 imagery, including linear Stumpf, quadratic Stumpf, Sigmoid, linear Support Vector Machine (SVM), Gaussian SVM, K-Nearest Neighbors (KNN), Decision Tree (DT), Bagged Tree, and Subspace KNN. Their results showed that ensemble methods such as Bagged Tree performed best, especially in deeper waters up to 35 m, with an average RMSE of around 2 m. Liu et al. (2024) compared the performance of machine learning models (Gradient Boosting Machine (GBM) and Neural Network) with traditional methods (log-linear and Stumpf) using Sentinel-2 imagery over Hainan Island, China. The GBM achieved the highest accuracy (RMSE = 0.59 m), though errors increased in deeper waters (> 9 m) compared to shallower areas (< 3 m). Ashphaq et al. (2024) evaluated various regression techniques for SDB using Sentinel-2 imagery, including traditional multiple linear regression and ML models such as SVM, Gaussian Process (GP), DT, KNN, k-fold DT, and RF. Their findings demonstrated that machine learning models - especially GP and RF - outperformed traditional approaches, with R^2 values up to 0.97 and RMSE of 1.23 m. Tran et al. (2024) tested six ML models (RF, SVM, CatBoost (CB), Extreme Gradient Boosting (XGB), Light Gradient Boosting Machine (LGBM), KTBoost) on Landsat 9 imagery over a shallow, turbid lagoon. These models were evaluated both with and without meta-heuristic optimization (Dragon Fly, Particle Swarm Optimization, Grey Wolf Optimization). LGBM with Particle Swarm Optimization achieved the highest performance (R^2 = 0.908, RMSE = 0.31 m), demonstrating the effectiveness of combining feature selection with machine learning in complex environments. Nguyen et al. (2025) compared empirical models (Stumpf, MLR) with five nonlinear ML models (RF, SVM, CB, XGB, LGBM) to model the water depth from Sentinel-2 imagery in the highly dynamic coastal zone of Hasaki, Japan. Their analysis revealed that CB achieved the best results (R^2 = 0.84–0.92, RMSE < 0.5 m in shallow waters < 7 m), with ML models generally outperforming empirical models, especially in turbid and dynamic coastal conditions. In Vietnam, machine learning has been increasingly applied across various research domains, including investigations into the relationship between landslide susceptibility and land-cover change (Huu et al., 2024) and multi-step-ahead water-level prediction to support water-resources management in the Mekong Delta (Duc et al., 2024).

Open-access datasets from Landsat (30 m resolution) and Sentinel-2 (10 m resolution) have been widely adopted in SDB research (Ashphaq et al., 2021). By comparison, PlanetScope satellites deliver significantly finer spatial resolution of 3 m, along with eight spectral bands and an almost daily revisit cycle. This combination enables detailed detection of small-scale underwater features far beyond the capabilities of Landsat's 30 m resolution (Kalybekova, 2025). The ability to capture near-daily imagery makes PlanetScope especially valuable for monitoring dynamic coastal environments where conditions can change rapidly. PlanetScope satellite imagery has proven effective for shallow-water bathymetry estimation. In Egypt, Gabr et al. (2020) demonstrated that the Stumpf algorithm performed well with PlanetScope imagery, achieving an RMSE of 0.38 m. Li et al. (2019) developed an adaptive algorithm applied to Planet Dove imagery. They validated it across five coral reef areas using 61,025 *in-situ* depth records, yielding RMSE values of

1.22–1.86 m, with the highest accuracy observed within the 4–10 m depth range. Sesama et al. (2020) employed PlanetScope data over the Karimunjawa Islands and reported an R^2 of 0.6952 and RMSE of 2.85 m for depths up to 20 m. Wulandari and Wicaksono (2021) confirmed the high accuracy of the blue/green band ratio method in waters shallower than 10 m. Tatsuyuki Sagawa et al. (2023) emphasized that, while SDB is a valuable tool for coastal research and management, accuracy remains a significant challenge. Using PlanetScope imagery with machine learning algorithms, including RF and Deep Neural Networks, they demonstrated that SDB performance improves by incorporating neighboring-pixel information, increasing the number of spectral bands, and selecting suitable training data. Their model achieved RMSEs of 0.5–0.9 m (0–10 m depth) and 1–1.4 m (0–20 m depth), highlighting the strong potential of SDB in complex coastal environments. More recently, Khakhim et al. (2024) assessed the potential of newly available spectral bands, such as Coastal Blue and Coastal Yellow, achieving an R^2 of 0.47. Downes et al. (2025) demonstrated that selecting an optimal spectral band combination can significantly improve depth-retrieval accuracy from PlanetScope imagery. Their results showed that a multi-band linear regression model substantially outperformed the conventional band-ratio method, yielding a high coefficient of determination (R^2 = 0.94) and a low root-mean-square error (RMSE = 0.41 m). The model was most effective within the 0.5–5 m depth range, with slightly increased error at greater depths. These findings underscore the potential of combining high-resolution optical sensors with machine learning techniques to enhance coastal bathymetric mapping in complex shallow-water habitats. Overall, these studies demonstrate that PlanetScope imagery represents a reliable and effective data source for shallow-water bathymetric mapping.

Despite significant advances in SDB, studies that integrate PlanetScope imagery with ML techniques to estimate depths beyond 20 m remain limited. Most remote sensing research in marine protected areas has concentrated on clear-water atolls (Bertin et al., 2022; Li et al., 2023), with few validations in sediment-influenced environments such as Nha Trang Bay. The bay is affected by continental sediment discharge (Tkachenko, 2015) and ongoing construction on nearby islands (Tkachenko, 2023), emphasizing the urgent need for focused bathymetric investigations in this area. To address this gap, the present study validates a straightforward, practical yet reliable approach for estimating bathymetry in the southern area of the Nha Trang Marine Protected Area by integrating PlanetScope imagery and state-of-the-art learning methods.

Our work serves as the first advanced comparison of multiple ML models (RF, SVM, CatBoost, XGB, LGBM, GB) and traditional methods (MLR, Stumpf) using 8-band PlanetScope imagery for high-resolution bathymetry in a coral-rich tropical setting and optically complex Case-2 turbid waters. The sensor's meter-scale resolution and additional mid-visible bands (Green-2, Yellow) offer improved sensitivity to bottom-water interactions that are not well captured by standard RGB-type systems. Our results also provide initial operational evidence that depth retrieval can extend to 30–40 m, exceeding the ~25 m theoretical optical limit typically cited for such waters. While uncertainties naturally increase with depth, this highlights the combined value of expanded spectral information and modern ML algorithms. The demonstrated performance suggests strong potential for applying this framework to other complex coastal systems globally.

2. Material and methodology

2.1. Study site

Nha Trang Bay Marine Protected Area (MPA), located in Khanh Hoa Province along Vietnam's central coast, is the country's largest designated multi-use coastal conservation zone (Fig. 1). Encompassing around 160 km², the MPA was established in 2002 with the primary goal of preserving coral reefs and benthic habitats that define the region's ecological richness. The bay opens into the East Sea and includes a cluster of nine islands, seven of which fall within the scope of this study. The study area is located in the southern part of the Nha Trang Bay MPA. A diverse and complex seafloor morphology characterizes Nha Trang Bay; in particular, the southern sector is strongly influenced by the spatial fragmentation created by a system of islands, including Hon Mun, Hon Tam, Hon Mot, Hon Mieu, and Hon Tre. Among them, Hon Tre Island - the largest - acts as a natural barrier dividing the bay into northern and southern zones. Historically, up until the 1990s, Nha Trang Bay was recognized as one of Vietnam's most biologically diverse marine systems, harboring over 250 hard coral species across 60 genera (Britaev et al., 2017). Its ecological significance links it to the Western Pacific Coral Triangle, a globally renowned hotspot of coral biodiversity and evolutionary origin for Indo-Pacific coral species.

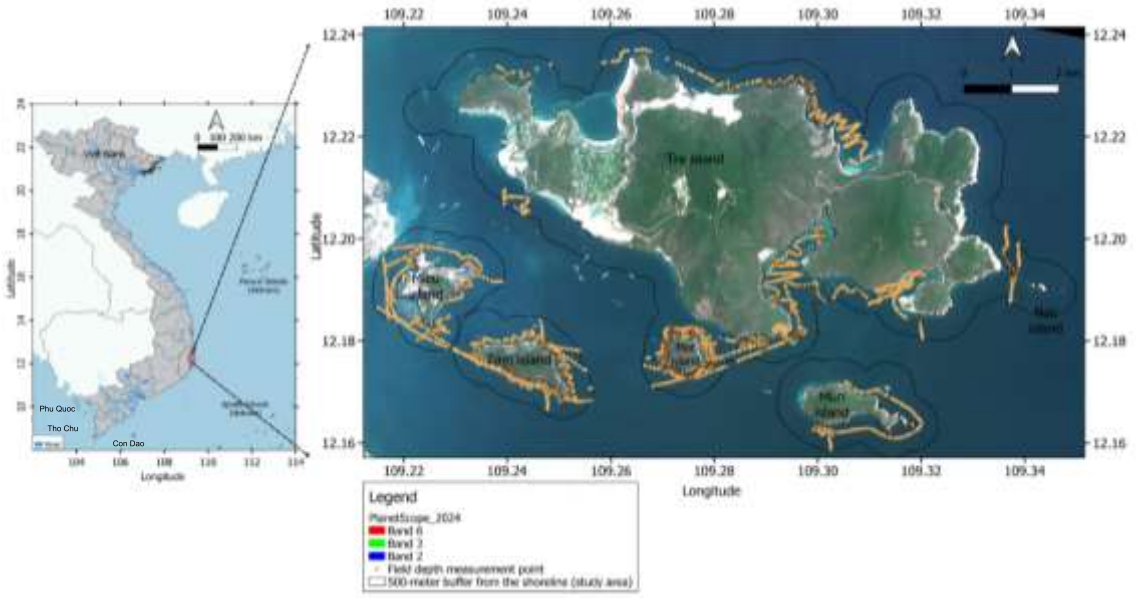

Figure 1. The southern area of the Nha Trang Marine Protected Area

The tides in Nha Trang Bay are irregular diurnal, with low tides varying seasonally (October-March in the morning, April-September in the afternoon, September-October at noon, and March-April at midnight), with the strongest tides in June-July and November-December (An et al., 2007). The southern Nha Trang MPA is influenced by both reef systems and continental sediment inputs from small rivers and coastal runoff, producing spatially heterogeneous optical conditions (i.e., a mixture of optically complex waters). In bio-optical terms, the study area is best characterized as dominated by Case-2 (optically complex) waters, with localized Case-1 characteristics in clearer, reef-protected patches; this classification reflects

significant contributions of inorganic particles and turbidity locally (Mobley et al., 2004). The implication for model design is that blue-green spectral bands are expected to carry strong depth signals in clearer patches, and the green 2-yellow spectrum for dominant suspended sediment areas.

Despite its importance, no large-scale or high-resolution bathymetric mapping has been conducted in the MPA to date. This lack of foundational data presents a major gap in effective marine spatial planning and habitat management. The water depth of the Nha Trang MPA spans from very shallow reef flats at 0–2 m to deeper channels reaching 40 m, establishing the full target detection range for Satellite-Derived Bathymetry (SDB) in this study.

Classical optical theory holds that the blue band is typically the deepest-penetrating, reaching 25 m, followed by green (15 m), red (5 m), and NIR (< 1 m) (Jupp, 1988). Therefore, 25 m penetration depth of the blue band serves as the theoretical maximum physical attenuation threshold (potential H_{max}) for passive optical SDB. While the 25 m limit is a remarkable reference, there is a strong rationale for testing the operational limit at 40 m using the deployed PlanetScope imagery and machine learning (ML) models in this study. Local water conditions in Nha Trang MPA may exceed the generalized clarity assumptions of classic SDB studies, allowing for marginally deeper signal penetration. In addition, modern satellite radiometry (e.g., PlanetScope's 12-bit data) offers enhanced sensitivity, enabling the detection of subtle, residual spectral signals returning from depths beyond the classic limit. Lastly, the ML algorithms employed (e.g., Random Forest) are nonlinear and non-parametric, enabling them to leverage complex relationships and subtle spectral signatures to extrapolate the depth-reflectance relationship more effectively into the deeper, near-attenuation range than traditional linear methods.

The study site was spatially limited by a 500 m buffer from the shoreline, which was sufficient to cover the distribution of coral reefs and other marine habitats in the southern part of the Nha Trang MPA (Fig. 1).

2.2. Working process, image acquisition, and field data collection

2.2.1. Working process

Bathymetric map generation from satellite data begins with the acquisition and preprocessing of planetary imagery. Raw top-of-atmosphere (TOA) data are corrected in ACOLITE to remove atmospheric effects and derive reliable surface reflectance. After preprocessing, depth estimation is carried out using models such as Stumpf, MLR, RF, or CB. Model performance is evaluated against ground-truth points (GTPs) using metrics like R^2, RMSE, and MAE. If accuracy is sufficient (e.g, $R^2 > 0.80$), a final bathymetric map is produced. If not, model choice and training are revisited until the required performance is achieved (Appendix, Figure A1).

2.2.2. Satellite image acquisition

To support high-resolution bathymetric mapping across the Nha Trang Bay MPA, we utilized PlanetScope imagery acquired in 2024. PlanetScope provides multispectral data at a 3-meter spatial resolution, which is particularly well-suited to capturing the fine-scale variability of seabed features in complex coastal environments. The selected image, retrieved under the Planet's Education and Research Program and contains eight spectral bands of Coastal Blue (≈431–452 nm), Blue (≈465–515 nm), Green I (≈513–549 nm) and Green II (≈547–583 nm), as well as Yellow (≈600–620 nm), Red (≈650–680 nm), Red-edge (≈697–713 nm) and NIR (≈845–885 nm) (Appendix, Table A1).

This spectral richness supports the detection of water column variations and benthic reflectance, both critical for depth retrieval. For spatial analysis and modeling,

the imagery was reprojected to the WGS 84 UTM Zone 49N coordinate system.

2.2.3. Field data collection

Field measurements of water depth were conducted under optimal environmental conditions (i.e., clear skies, low wind intensity, and minimal wave disturbance) across the coastal waters surrounding the Nha Trang MPA. We carried out these surveys over two distinct periods: from 21 March to 23 March and from 22 April to 24 April, 2024. Notably, these *in-situ* measurements were deliberately aligned with the dates of satellite image acquisition to ensure consistency between ground-truth data and remotely sensed observations (Fig. 1). Bathymetric data was collected using a Garmin GPSMAP 585 Plus single-beam echo sounder system (points logged roughly every ~ 1−5 s depending on vessel speed), which integrates ClearVü scanning sonar technology for enhanced subsurface imaging (Garmin, https://www.garmin.com.my/products/onthewater/gpsmap-585-plus/, accessed on 20 May 2023). This equipment allowed the precise recording of geographic coordinates (longitude and latitude) along with corresponding water depth at each surveyed location. The resulting dataset - comprising approximately 9,549 georeferenced points - spans a bathymetric gradient from -0.17 m in very shallow areas to depths of approximately -40.21 m (Table 1).

Table 1. Water depth statistics

No. of Observation	Minimum (m)	Mean (m)	Standard deviation (m)	Maximum (m)
9,549	-40.21	-12.75	6.83	-0.17

Observations were binned into 5-meter depth strata. This stratified summary shows the density of samples in the shallow (0–5 m), intermediate (5–20 m), and deep (> 20 m) depth ranges and supports the interpretation of model performance across depths (Appendix, Table A2).

The survey was completed by two trained technicians, responsible for navigating the survey vessel and operating the sonar equipment. It is necessary to convert the measured sea depth to values relative to mean sea level (MSL). To achieve this, tidal data from the study site were calibrated to MSL = 0 using reference information from the "Tidal Table" published by the Hydro-Meteorological Forecasting Center (Nha Trang station).

For each survey point, we computed:

depth_MSL = depth_raw + (tide_height_at_measurement - mean tidal level/2) (if neap tide) and

depth_MSL = depth_raw − (tide_height_at_measurement − mean tidal level/2) (if spring tide)

The satellite-based depth model was trained using these MSL-referenced depths to ensure consistency across survey dates. After model prediction, the final depth raster was converted from MSL to the instantaneous water level at the time of satellite image acquisition using the corresponding tide height.

2.3. Bathymetry mapping

2.3.1. Atmospheric correction

We applied the Dark Spectrum Fitting (DSF) method embedded in the ACOLITE software, a robust tool tailored for atmospheric correction over aquatic environments, to effectively remove the atmospheric influence on water-surface pixel values (Quinten, 2024). Dark Spectrum Fitting (DSF) works by identifying the darkest pixels in an image, or regions where the surface should reflect almost no light; thus, any signal the satellite detects there is primarily from the atmosphere. By matching these dark-pixel spectra to modeled

atmospheric effects, DSF estimates and removes atmospheric scattering, giving a clean surface reflectance.

This is a standard step to minimize atmospheric scattering and absorption, which can distort spectral signatures in coastal and marine satellite imagery. In our workflow, the DSF algorithm was executed via ACOLITE's command-line interface to process high-resolution PlanetScope (2024) imagery.

The ACOLITE transformed top-of-atmosphere radiance to surface reflectance at the water surface, with particular emphasis on the blue-to-red spectral range, which is a key input for depth estimation. Since the selected images exhibited no signs of cloud cover, no cloud masking was applied (Appendix, Table A3).

Sun-glint correction was not applied to the PlanetScope scenes. The solar-sensor geometry of the PlanetScope scene indicates negligible sun-glint effects. The off-nadir viewing angle was only 4.8°, while the azimuthal separation between the Sun (84.8°) and sensor (98.3°) was approximately 13.5°, placing the sensor well outside the specular reflection corridor. Together with the high solar elevation (57.7°), these conditions inhibit mirror-reflected sunlight from reaching the sensor. Visual inspection of the NIR band shows no haze-like brightening, further confirming the absence of glint contamination. In addition, visual inspection of the imagery showed no localized specular hotspots or haze-like brightening, and the near-infrared (NIR) band likewise lacked the elevated diffuse reflectance patterns typically associated with glint. Given the absence of detectable glint and because many glint-correction routines can substantially alter per-pixel radiometry when applied to scenes with negligible glint, we preserved the original surface reflectance to avoid introducing corrective artefacts that could bias subsequent ML training and depth retrieval.

2.3.2. Traditional models and machine learning models

In this section, we focus on regression models that are widely used in remote sensing and environmental science. In this paper, we proceed to implement the Random Forest (RF), Support Vector Machine (SVM), Light Gradient Boosting Machine (LGBM), Extreme Gradient Boosting (XGB), CatBoost (CB), Gradient Boosting (GB) models, which cover a diverse range of bagging (RF), boosting (XGB, CB, GB, LGBM), traditional ML (SVM), and traditional learning models (MLR, Stumpf) to validate the performance of bathymetry retrieval using 8-band PS imagery.

2.3.2.1. Traditional models

Multiple Linear Regression model (MLR)

Multiple Linear Regression (MLR) is a statistical modeling technique used to predict the value of a single dependent variable from linear relationships with two or more independent variables. MLR describes the linear impacts of multiple inputs on the dependent variable using linear equations. This relationship is typically expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon \quad (1)$$

where:

Y is the dependent variable,

X_i are the independent variables,

β_i are the regression coefficients representing the change in Y for a one-unit change in X_i,

β_0 is the intercept, and ϵ is the error term.

The coefficients can be estimated using Ordinary Least Squares (OLS), which minimizes the sum of squared residuals between observed and predicted values. MLR is widely applied for prediction, forecasting, and for deriving relationships among variables across various research domains.

Stumpf model

The Stumpf method (i.e., the Log-Ratio) (Stumpf et al., 2003), is a widely used empirical approach for estimating shallow water depth from multispectral satellite imagery. It leverages the differential attenuation of light in water across various wavelengths. The principle is straightforward: calculate the logarithm of the ratio of reflectances from a blue band and a green band in the electromagnetic spectrum. This ratio is designed to normalize for variations in illumination and seafloor type, assuming a near-linear relationship between the log-ratio value and water depth. *In-situ* measured depths are required to calibrate specific coefficients to transform the log-ratio values across the image into a comprehensive bathymetric map. The method is generally reliable for clear, shallow waters, offering a cost-effective alternative to traditional hydrographic surveys. The technique can be expressed as the following equation:

$$D = m * ln(\frac{R_{blue}}{R_{green}}) - d_o \quad (2)$$

where:

D is the estimated water depth,

R_{blue} is the reflectance in the blue spectral band,

R_{green} is the reflectance in the green spectral band,

m is the slope coefficient derived from the linear regression during calibration, representing the rate of change in depth per unit change in the log-ratio,

d_0 is the intercept coefficient (or offset) derived from the linear regression, representing the reference depth when the log-ratio is zero.

This process establishes a linear relationship between the natural logarithm of the ratio of blue-to-green band reflectances and water depth, with coefficients m and d0 determined through calibration using *in-situ* depth measurements. The Stumpf method is included as a traditional empirical baseline because it remains widely used for shallow, clear waters. To ensure a fair comparison, we report model performance both across the full depth range and stratified by depth classes (*e.g.*, 0–5, 5–10 m), where Stumpf is typically applicable.

2.3.2.2. Machine learning (ML) models

Random Forest (RF)

Random Forest represents an ensemble-based machine learning approach that begins by creating multiple bootstrap samples from the original training dataset (Breiman, 2001). Within each subsample, the algorithm randomly selects a subset of input variables to construct individual decision trees, thereby ensuring diversity across the ensemble. The final prediction is obtained by aggregating the outputs from all constituent trees, typically by averaging for regression tasks, thereby significantly improving predictive accuracy and reducing overfitting compared to single decision trees. This methodology excels at modeling complex, nonlinear relationships among predictor variables and has recently shown exceptional capabilities in satellite-derived bathymetry research.

Support Vector Machine (SVM)

Support Vector Machine is a well-established and widely used machine learning technique, known for its ability to identify complex, nonlinear relationships between spectral reflectance values and shallow-water depth measurements (Gholami & Fakhari, 2017). The algorithm achieves this by applying a nonlinear mapping that projects the input data into a higher-dimensional feature space via kernel functions. During the model implementation, support vectors are identified by systematically dividing available depth measurements into balanced training and validation subsets. The optimal regression boundary is subsequently established using kernel transformations that enable the algorithm to handle nonlinear patterns in the data (Zhang et al., 2022). A distinguishing

characteristic of SVM is its optimization strategy, which prioritizes minimizing the upper bound on generalization error rather than merely reducing training-set error, thereby promoting robust model performance on unseen data.

Light Gradient Boosting Machine (LGBM)

Light Gradient Boosting Machine is an advanced gradient boosting framework developed by Microsoft, optimized for both high performance and computational efficiency (Li et al., 2018). Among its most distinguishing features is the leaf-wise tree growth algorithm, which contrasts with the level-wise approach used in traditional gradient-boosting methods. Instead of expanding all nodes at each level, LGBM grows the tree by continuously splitting the leaf with the highest potential for loss reduction. This strategy enables faster convergence and often yields superior predictive accuracy, particularly when applied to large-scale, high-dimensional datasets.

Another notable innovation of LGBM is its use of histogram-based decision rules, which bucket continuous feature values into discrete bins. This not only accelerates the training process but also significantly reduces memory usage without compromising model performance. In addition, LGBM supports parallel and GPU training, handles categorical features, and includes built-in regularization mechanisms, making it particularly effective for complex data distributions in modern machine learning tasks.

Extreme Gradient Boosting (XGB)

Extreme Gradient Boosting is an advanced implementation of the gradient boosting framework that leverages parallel processing, constituting a highly scalable machine learning system specifically engineered to enhance the efficiency and performance of tree-based boosting algorithms (Chen & Guestrin, 2016). The XGB methodology constructs predictive models through an iterative process that uses gradient descent to minimize a carefully designed loss function across an ensemble of weak learners, typically shallow decision trees. Each successive tree is trained to correct the residual errors from the previous iteration, creating a powerful additive model that progressively improves prediction accuracy. The algorithm incorporates several sophisticated features, including regularization terms to prevent overfitting, advanced tree-pruning strategies, and efficient memory management systems that enable the processing of large datasets. This ensemble approach has attracted considerable attention in remote sensing applications. It has recently been successfully implemented in satellite-derived bathymetry studies, where it has demonstrated exceptional performance in capturing the complex spectral-depth relationships inherent in shallow-water environments (Gülher & Alganci, 2023a; Tran et al., 2024).

CatBoost (CB)

CatBoost is a gradient boosting algorithm developed by Yandex (Prokhorenkova et al., 2018) that is specifically optimized for handling categorical features without extensive preprocessing. Its most distinctive innovation lies in the use of ordered boosting and target statistics with permutation-driven techniques, which effectively prevent target leakage and overfitting commonly associated with categorical data encoding. Unlike traditional methods that rely on one-hot or label encoding, CatBoost natively encodes categorical variables during training using efficient, data-driven schemes. This capability, combined with competitive accuracy, robust generalization, and ease of use, makes CatBoost particularly well-suited for structured datasets with high cardinality categorical features.

Gradient Boosting (GB)

Gradient Boosting is a robust and widely adopted ensemble learning method that builds predictive models by sequentially combining weak learners, most commonly decision trees,

to minimize a specified loss function (Belyadi & Haghighat, 2021). At its core, GB iteratively fits each new model to the negative gradient of the loss function with respect to the current ensemble's predictions, effectively correcting the residual errors from prior iterations. This stage-wise additive approach allows GB to progressively refine model accuracy, making it highly effective for both regression and classification tasks. A key advantage of GB is its flexibility in optimizing arbitrary differentiable loss functions, along with its capacity for fine-grained control through hyperparameters such as learning rate, number of iterations, and tree depth. While the method is powerful, it can be sensitive to overfitting if not adequately regularized, requiring careful parameter tuning and validation. Despite this, GB's strong theoretical foundation and high empirical performance have solidified its role as a cornerstone algorithm in modern machine learning, particularly for structured data applications.

2.3.3. Models implementation

The following methodology describes the implementation process for the proposed models to estimate water depth within the research framework.

Hyperparameter tuning

Machine learning algorithms include multiple hyperparameters that require optimization to achieve optimal model performance. An automated random search approach incorporating five-fold cross-validation was implemented using the scikit-learn (Pedregosa et al., 2011) framework to determine the most effective hyperparameter combinations (Appendix, Table A4). We used the mean squared error (MSE) as the stopping criterion, with the search terminating when the minimum MSE was reached.

Model execution

In this study, surface reflectance values from PlanetScope imagery were used directly as predictor variables without log transformation, despite the exponential decay of underwater light with depth. While log- or log-ratio transforms are often applied in traditional band-ratio or multi-linear regression approaches to linearize the depth-reflectance relationship, our modelling relies on non-parametric ML algorithms that do not require a priori linearity assumptions. These methods can learn nonlinear and exponential relationships directly from the raw data. Moreover, the low-magnitude surface reflectances observed in our study area (< 0.05) increase susceptibility to noise amplification when log-transformed, particularly in optically deep waters (> 20 m). By avoiding log transformation, we reduced the risk of introducing instability in the predictor variables while retaining the full dynamic range for model training. We retained variables for the MLR to validate its performance on similar datasets to those used for ML models.

A comprehensive dataset containing 9,549 observation points with 6 predictor variables (Rrs_coastal, Rrs_blue, Rrs_green, Rrs_green2, Rrs_yellow, and Rrs_red) was processed using six distinct ML models and two linear models to assess their ability to derive water depth estimates from PlanetScope satellite data. The dataset was randomly partitioned into training (70%, approximately 6,684 points) and test (30%, approximately 2,865 points) subsets using scikit-learn's train_test_split function. The Python programming language, integrated with the Anaconda package manager, was used to optimize hyperparameters and compare all selected models.

2.3.4. Model evaluation

To comprehensively evaluate the performance of the bathymetric models developed for the Nha Trang MPA, we

employed a set of widely recognized statistical evaluation metrics. R^2 quantifies the fraction of variance explained (higher is better). RMSE and MAE measure absolute predictive error (lower is better). AIC and BIC combine goodness-of-fit with a penalty for model complexity, with lower values indicating a better balance of fit and parsimony (*i.e.*, prefer models with lower AIC/BIC when comparing similarly-specified models). We therefore consider R^2, RMSE, and MAE as primary accuracy metrics, and AIC/BIC as secondary diagnostics for model complexity and relative model selection, as detailed in Equations (3) through (7). These metrics collectively provided a robust quantitative basis for assessing the accuracy, reliability, and complexity of the bathymetric estimations generated by the models.

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (3)$$

in which:

$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \varepsilon_i^2$ and $\overline{y} = \sum_{i=1}^{n} y_i$

ε : the error term

n: the total number of validation samples

$\hat{y}_i$: predicted value

y_i: corresponding true value

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (4)$$

$$MAE\ (y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

$$AIC = n \times \log\left(\frac{RSS}{n}\right) + 2 \times K \quad (6)$$

$$BIC = n \times \log\left(\frac{RSS}{n}\right) + K \times \log(n) \quad (7)$$

where:

n: number of observations

K: number of parameters (including intercept)

RSS: residuals sum of squares

3. Result

3.1. Advanced comparison of the performances of models for bathymetry estimation

The comparative analysis of machine learning models reveals substantial variations in predictive performance across different algorithmic approaches. RF emerged as the superior model, achieving an R^2 of 0.85, indicating it explains 85% of the variance in water depth measurements. The model achieved the lowest error metrics with RMSE and MAE values of 2.66 m and 1.85 m, respectively, while also exhibiting the most favorable information criteria scores (AIC = 5609.20, BIC = 5644.96), suggesting an optimal balance between model complexity and predictive accuracy (Table 2).

Table 2. Model performance for water depth retrieval in the Nha Trang MPA

Model	R^2	RMSE	MAE	AIC	BIC
RF	0.85	2.66	1.85	5609.20	5644.96
GB	0.84	2.69	1.91	5677.09	5712.85
XGB	0.84	2.69	1.84	5687.71	5723.47
CB	0.84	2.69	1.94	5689.37	5725.13
LGBM	0.84	2.76	1.99	5831.56	5867.33
SVM	0.82	2.89	2.05	6086.63	6122.39
MLR	0.57	4.46	3.27	8575.46	8611.22
Stumpf	0.29	5.72	4.45	9997.69	10009.22

The ensemble tree-based methods, including GB, XGB, and CB, demonstrated remarkably similar and competitive performance metrics, with R^2 values of 0.84 and comparable error statistics, indicating their robust ability to capture the complex nonlinear relationships between spectral characteristics and bathymetric properties in coastal waters.

Conversely, the Stumpf empirical model underperformed across all evaluation metrics, achieving the lowest R^2 (0.29) and the highest error metrics (RMSE = 5.72 m, MAE = 4.45 m),

along with substantially elevated information criteria (AIC = 9997.69, BIC = 10009.22). The MLR model (depth = -212.52 × $Rrs_{coastal}$ + 27.7 × Rrs_{blue} + 577.92 × Rrs_{green} + 1842.82 × Rrs_{green2} − 664.75 × Rrs_{yellow} − 606.29 × Rrs_{red}) also demonstrated limited effectiveness with an R^2 of 0.57, while SVM and LGBM showed moderate to good performance with R^2 values of 0.82 and 0.84, respectively.

Ensemble-based machine learning models, specifically RF, GB, XGB, CB, and LGBM, demonstrated exceptional accuracy in predicting water depths. These models achieved high correlation coefficients (approximately 0.91−0.92), standard deviations that closely matched the observed data, and minimal RMSD values. This strong performance indicates their superior ability to accurately reproduce both the magnitudes and spatial patterns of actual water-depth measurements. Notably, the RF model performed slightly worse on the performance metrics, suggesting greater confidence in its depth predictions (Appendix, Figure A2).

The SVM model also performed relatively well, with a correlation coefficient of about 0.91 and a slightly lower standard deviation than the observations. However, its greater distance from the reference point (indicating less accuracy) compared to the ensemble models suggests a marginal reduction in its predictive capability. In contrast, the MLR and Stumpf models underperformed significantly. The MLR model had a correlation coefficient below 0.8, whereas the Stumpf model's was below 0.6. This indicates that both models substantially underestimated data variability and deviated considerably from observed patterns. The Stumpf model, in particular, failed to accurately capture either the variability or the spatial structure of the observed bathymetry, as evidenced by its significantly underestimated standard deviation (~2.5) and the highest RMSD (Appendix, Figure A2).

We further assessed the performance of our water depth estimation models using scatter density plots (Appendix, Figure A3(a-h)), which illustrate prediction accuracy and distribution across the entire depth range. The RF model (Appendix, Figure A3c) delivered the highest predictive accuracy, achieving an R^2 of 0.85 and the lowest RMSE of 2.66 m. Its predictions clustered tightly around the 1:1 reference line, demonstrating excellent agreement with measured values across all depths. The SVM model (Appendix, Figure A3d) also performed strongly (R^2 = 0.82; RMSE = 2.89 m), though with slightly greater dispersion, particularly at deeper ranges.

The ensemble gradient-boosting models LGBM (Appendix, Figure A3e), XGB (Appendix, Figure A3f), CB (Appendix, Figure A3g), and GB (Appendix, Figure A3h) showed comparable robust performance. While minor scattering was observed at deeper ranges across these four boosting algorithms, it did not significantly diminish their overall performance. In contrast, both the MLR (Appendix, Figure A3a) and Stumpf models (Appendix, Figure A3b) underperformed at all scenarios. Their plots showed significant deviation from the 1:1 line, particularly at greater depths, where these models consistently underestimated actual depths. Rather than implying a theoretical shortcoming, this result indicates that linear models such as MLR and simple log-ratio approaches fail to capture complex, spatially varying nonlinear relationships between multispectral reflectance and depth in optically heterogeneous coastal waters.

3.2. Feature rank across ML models

The feature importance analysis reveals distinct patterns in the utilization of spectral features for satellite-derived bathymetry. Rrs_green2 emerges as the dominant predictor across all models, though with substantial variation in relative importance. XGB and SVM demonstrate a concentration at

~70–80%, while RF, CB, and GB show moderate dependence (~35–40%), and LGBM exhibits intermediate values (~20%). This universal dominance of green wavelength reflectance aligns with established optical principles, which state that green light provides optimal water-column penetration and depth sensitivity in coastal environments.

Model-specific feature preferences reveal significant algorithmic differences. A balance in feature selection was found for LGBM, with Rrs_yellow, Rrs_red, Rrs_coastal, and Rrs_blue each contributing 15–18% alongside the primary green features. This distribution differs from XGB's concentrated strategy, in which secondary features contribute little. Tree-based ensemble methods (RF, GB, CB) show convergent patterns, with Rrs_green consistently ranking second (~15 – <30%) and other spectral bands contributing modestly (~5–15%). SVM occupies an intermediate position, showing moderate feature diversity while maintaining strong Rrs_green2 dependence. Rrs_green maintains consistent secondary importance across most models; Rrs_blue shows variable importance, ranking moderately in RF and GB (~15%) but receiving less weight in XGB and SVM. The coastal (Rrs_coastal), yellow (Rrs_yellow), and red (Rrs_red) bands identify relative magnitudes across all models (~3–15%), reflecting their limited but consistent bathymetric utility (Appendix, Figure A4).

Previous studies consistently highlight the green band as the most influential predictor for bathymetry estimation in machine learning models, particularly because it has the highest feature importance in the RF algorithm (Xie et al., 2023). The blue band is generally ranked second in importance, with both bands being essential for bathymetric modeling because their reflectance signals can penetrate deeper into the water column (Xie et al., 2023). This finding aligns with Saeidi et al. (2023), who emphasized that the green and blue spectral ranges are the most critical for accurate bathymetric mapping. Similarly, Ashphaq et al. (2021) concluded that the optimal spectral bands for shallow-water bathymetry lie within the blue-green region.

Given the optically complex, sediment-rich coastal waters at our study site, depth-dependent reflectance arises from the interaction of multiple processes, including particle backscattering in the green-yellow domain, absorption in the blue, and surface or sediment contributions in the coastal band. The feature-importance analyses (Appendix, Figure A4) illustrate that tree-based models differ in how they exploit these signals. While XGB concentrates heavily on Rrs_green2, likely because this band carries the strongest first-order relationship with bathymetry in turbid environments, RF distributes importance more evenly across the spectrum. This balanced reliance is physically plausible: the blue and coastal bands help constrain attenuation in clearer portions of the transects; the yellow band captures suspended particulate effects; and the red band contributes to shallow-water discrimination where bottom influence weakens. Such multi-band integration prevents overemphasis on a single wavelength whose response may saturate or become unstable under varying sediment loads. Consequently, RF's more uniform feature usage reflects a modelling strategy that aligns with the optical oceanography of the region and may partially explain its more stable performance across depth ranges and turbidity regimes.

The consistent prioritization of green wavelengths across all approaches validates the physical basis of optical bathymetry. It underscores the critical importance of high-quality green-band measurements in satellite-derived bathymetric applications.

3.3. Satellite-derived seafloor topography

The RF-derived bathymetric map of the Nha Trang MPA (Apendix, Figure A5) reveals a complex underwater landscape with significant variations in water depth and

diverse seafloor features. Our bathymetry analysis shows a total depth of 35 m. The shallowest areas (< 0 to -5 m) appear in yellow, while the deepest zones (-30 to -35 m) are depicted in dark blue to purple. Intermediate depths (-5 to -25 m) transition from green to teal, indicating a gradual depth change across the study area.

Shallow-water habitats (< 0 to -10 m deep) are mainly found along the edges of coral reef systems and near the coast. These areas have irregular and patchy boundaries, suggesting the presence of reef flats, lagoons, and nearshore environments. The yellow and light green zones indicate ideal conditions for coral growth and high biodiversity, typically due to sufficient light for photosynthesis.

The deeper sections (> 25 m) are primarily located in the central and southeastern parts of the mapped area, forming distinct basins and channels. These dark blue to purple zones likely correspond to channels between reefs, fore-reef slopes, and sandy seafloor habitats. Sharp depth changes in several locations suggest the presence of reef walls or steep underwater inclines.

The bathymetric map highlights significant spatial heterogeneity, with numerous isolated coral patches and intricate reef formations creating a mosaic of different depth zones. This varied distribution is particularly noticeable in the central region, where alternating shallow and deep areas form a complex three-dimensional habitat. The fragmented nature of the reef systems suggests both natural geological processes and potential human impacts.

4. Discussions

4.1. Bathymetry model accuracy across depth ranges

To reveal higher distinct performance patterns of different approaches for bathymetry estimation, we highlight the RMSE values across depth ranges. All models achieve optimal performance in shallow water conditions and experience progressive accuracy deterioration as water depth increases. In the shallow-water zones (< 0 to -15 m), most models exhibit remarkably similar and satisfactory performance, with RMSE values consistently below 3 m across this depth range. The ML algorithms (RF, GB, CB, XGB, SVM, LGBM) and the traditional MLR approach show comparable accuracy levels in these shallow conditions, where water clarity is typically optimal and bottom reflectance provides strong signal return, while the Stumpf model begins to exhibit slightly elevated error rates around 3–4 m in the deeper areas.

The transition to moderate depths reveals a notable shift in model performance, with accuracy varying beyond the -20 m threshold. The Stumpf model performs relatively well in the shallow range (0– ~20 m) but degrades rapidly with depth. Compared to the ML models, which maintain relatively stable RMSE values around 4–5 m, the MLR approach occupies an intermediate position with error rates of approximately 5 m. The accuracy decline becomes substantially more severe in deeper waters (-25 to -35 m), where the Stumpf model becomes progressively unreliable (RMSE ranging from 12 to 17 m), while the ML models continue to demonstrate superior performance but still exhibit increasing error rates of 5 to 7 m. In the deepest surveyed range (-35 to -41 m), all models exhibit significant limitations, with the Stumpf model becoming essentially unusable (RMSE exceeding 22 m), while ML models show substantial errors of 8−9 m (Appendix, Figure A6).

Previous studies consistently show that the accuracy of satellite-derived bathymetry decreases with increasing depth, regardless of sensor type or modelling framework. Moeinkhah et al. (2019) reported low errors in very shallow waters (< 5 m, RMSE = 1.253 m), but accuracy dropped

markedly beyond 10 m. Duan et al. (2022) similarly demonstrated that once depths exceeded ~15 m, estimation errors increased substantially, and most models approached practical limits near 20–25 m. Liu et al. (2024) further confirmed this depth-dependent decline, noting that although Gradient Boosting achieved high shallow-water accuracy (RMSE = 0.59 m), all methods tested (Stumpf, Log-Linear, Neural Networks, and Gradient Boosting) exhibited escalating RMSE values as ddepth increased. PlanetScope-based studies show the same pattern, of which Sagawa et al. (2023) reported very low errors (0.5–0.9 m for 0 to 10 m; 1.0–1.4 m for 0 to 20 m), with accuracy decreasing in deeper classes, while Caballero and Stumpf (2023) demonstrated consistently lower errors in shallower intervals across multiple sites. Most recently, Downes et al. (2025) emphasized that model performance was strongest within the 0.5–5 m range, with modest increases in error at greater depths. Altogether, these studies highlight a robust inverse relationship between water depth and SDB accuracy.

Against this backdrop, the RF model developed in this study achieved R^2 = 0.85, RMSE = 2.66 m, and MAE = 1.85 m across an extensive range of retrieval intervals (0.5–40 m). This performance aligns with the expected behaviour of SDB systems while demonstrating stability over a markedly larger depth range than most empirical or ML-based studies. Recent work using WorldView-2, PlanetScope, and Sentinel-2 typically reports low vertical errors but is generally confined to much shallower environments. For example, Çelik et al. (2023) achieved $R^2 \approx 0.85$, RMSE 0.93–2.41 m, and MAE 0.65–1.86 m within 6–9 m depth intervals. Comparable shallow-water error levels are reported in other WorldView-2 coral-reef studies (RMSE ≈ 0.82–0.87 m) (Wicaksono et al., 2024), while PlanetScope and Landsat-8 comparisons show RMSE ≈ 0.38–0.43 m for shallow coastal zones (Gabr et al., 2020). These results are fully consistent with ours, but they are derived under far less optically attenuated conditions and over much smaller depth windows.

Studies using freely available multispectral sensors reinforce the idea that SDB accuracy is determined more by optical water properties, substrate, and the representativeness of the training data than by the specific ML regressor. Kwon et al. (2024) demonstrated stable RF performance on Sentinel-2 across Korean coastal sites, while Ye et al. (2024) showed that integrating Sentinel-2 with ICESat-2 may extend shallow-water retrievals but remains most reliable within the upper ~15 m. Together, these findings confirm that (i) ML regressors reliably outperform simple empirical models in optically complex waters; (ii) RMSE scales with accessible depth range; and (iii) the contribution of this study lies in showing that PlanetScope + RF can retain competitive accuracy (R^2 = 0.85; RMSE = 2.66 m) while mapping depths from nearshore to ~40 m, a range that exceeds the typical operational limits documented in the SDB literature. This capability is particularly valuable for coastal management, habitat monitoring, and applications that require consistent depth retrieval across both shallow and moderately deep tropical waters.

Our results show that MLR and Stumpf models do not necessarily achieve their highest accuracy in the very shallow zone (0–5 m). This pattern contrasts with the standard expectation for empirical bathymetric models. This behaviour can be attributed to the interaction between discharged sediment from the shoreline and seabed morphology in the Nha Trang MPA. While CR1 (please refer to Fig. 2) presents a smooth depth gradient, other transects (CR3-CR5) exhibit abrupt depth changes within the

0–20 m range, reducing model stability in the shallowest waters. In contrast, the more gradual transitions observed between -12 and -18 m at CR4 and CR5 correspond to the best performance of both models. Similar depth-dependent behaviour has been reported elsewhere; for example, Wu et al. (2021) found higher RMSEs for machine-learning models in the 0–3 m zone than at slightly more profound depths. Additionally, elevated turbidity near the shoreline - likely influenced by ongoing construction and dredging - may weaken the bottom signal, contributing to reduced accuracy in these models.

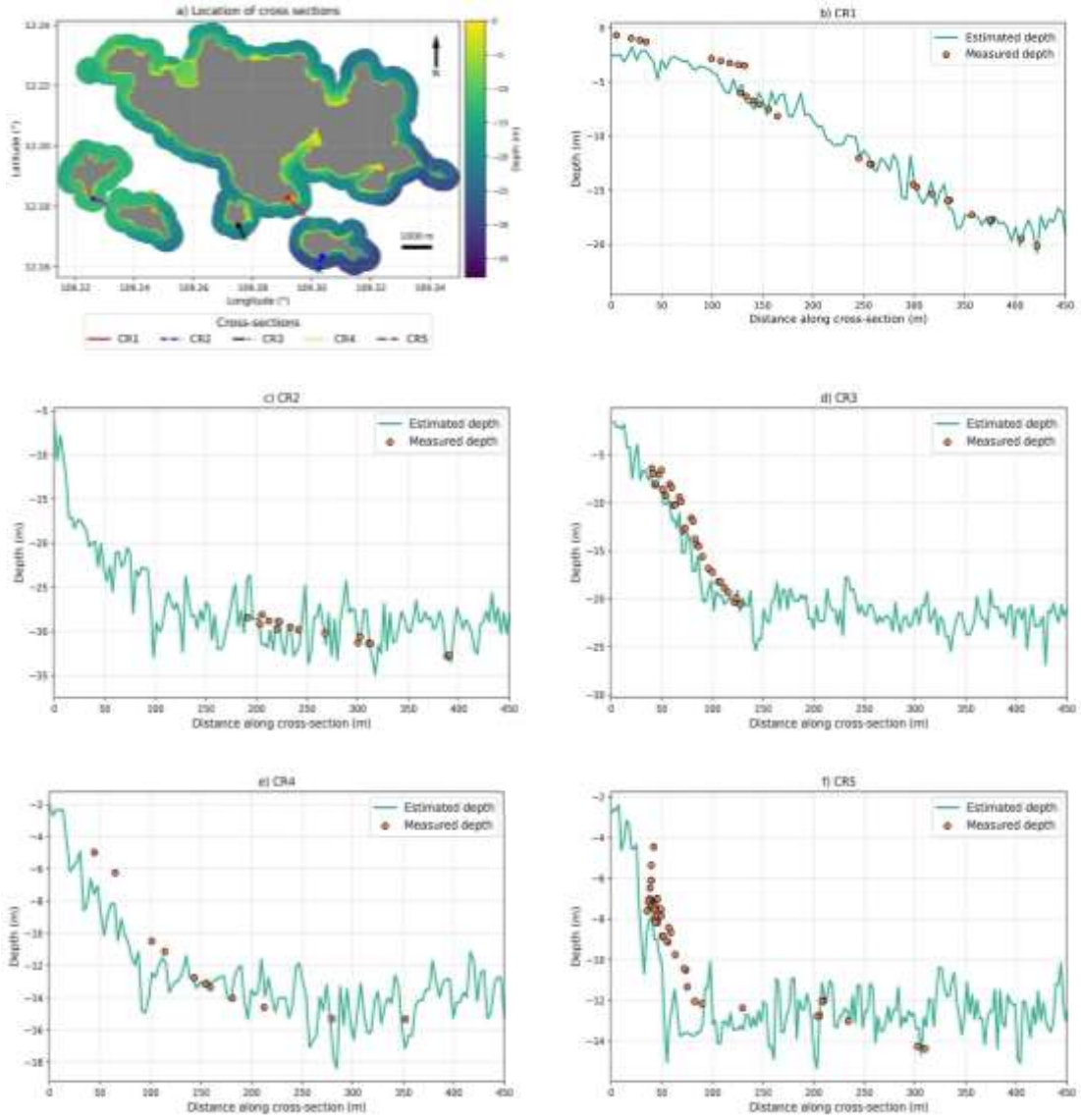

Figure 2. Bathymetric variation in different cross sections

Based on the depth-stratified performance already reported, the RF model maintains stable accuracy (R^2 > 0.8) down to an operationally reliable range of 25–30 m, after

which RMSE gradually increases. This pattern is consistent with theoretical penetration thresholds for blue-green wavelengths (Jupp, 1988). Reported RMSE values at 30–40 m reflect known physical limitations of light penetration, and therefore, we interpret $H_{max} \approx 25$ m as the effective operational limit of the model in Nha Trang MPA. Depths > 25 m were still mapped, but should be considered beyond the physically reliable range of passive multispectral SDB in this environment.

Notably, the machine learning approaches (XGB, RF, CB) consistently outperformed the multi-linear regression model without the need for log-transforming spectral bands. This is because tree-based algorithms can inherently model the nonlinear and exponential relationships between surface reflectance and depth, eliminating the need for pre-linearization that linear models require. In addition, given the very low surface reflectances observed in our dataset (< 0.05) and the inclusion of optically deep waters (> 20 m), applying a log transform would likely amplify sensor noise and atmospheric residuals, especially in the blue-green wavelengths, potentially reducing model stability. Using surface reflectance values, the ML models preserved the true dynamic range of the input data while flexibly adapting to variations in water clarity, depth, and bottom type across the study area. The green-2 (Green II, ~547–583 nm) band consistently ranks highest in importance in our models because, in tropical coastal waters like Nha Trang, green wavelengths often maximize bottom-signal penetration while limiting scattering relative to shorter blue wavelengths under sediment-rich (Case-2) conditions. In clearer reef channels where scattering is reduced, blue wavelengths also contribute strongly; this explains the observed dominance of green and blue bands across models.

This comprehensive evaluation clearly demonstrates that ensemble ML methods (RF, GB, CB, XGB) consistently outperform both traditional empirical approaches and conventional regression methods across all depth ranges, with this performance advantage becoming more pronounced as environmental conditions become more challenging (e.g., deeper area, sedimentation, complex seafloor). However, the steady increase in estimation errors with depth reflects the inherent physical limitations of optical remote sensing for bathymetric applications, where factors such as light attenuation, reduced bottom contrast, and diminished signal-to-noise ratios progressively reduce depth estimation accuracy, suggesting that satellite-derived bathymetry techniques are most reliable for shallow coastal mapping applications up to approximately 25–30 m depth in the Nha Trang MPA.

Despite the models' promising performance, a few limitations remain. Depth retrieval beyond approximately 30–40 m remains challenging due to reduced signal penetration and increased noise in optically complex waters. Additionally, producing habitat-specific depth profiles (e.g., for coral or rocky reef zones) was not feasible because validated benthic-habitat maps are currently unavailable for the study area; therefore, cross-sections were kept generic to avoid speculative interpretation. Future research should prioritize the development or acquisition of reliable habitat maps and explore integrating recently developed depth-invariant indices as synthetic spectral inputs for ML-based SDB. Such indices could further decouple bottom and water-column effects and potentially enhance depth-estimation accuracy, particularly when leveraging PlanetScope's 8-band sensor capabilities in coral-reef environments.

4.2. Bathymetric variation across depth sections

The assessment of estimated versus measured depths across transects CR1-CR5

provides important insights into the depth retrieval model's performance and reliability under diverse coastal bathymetric conditions. The estimated depth profiles (continuous lines) show substantial agreement with in situ measurements (discrete points), particularly within transects exhibiting pronounced, well-delineated depth gradients. This concordance is most evident in CR1, CR3, and CR5, where the model successfully resolves the abrupt transition from shallow nearshore environments to deeper offshore waters within the initial 100–150 m of each profile. The proximity of continuous and discrete data in these regions demonstrates the model's efficacy under optimal optical conditions, characterized by strong benthic reflectance and enhanced water transparency that facilitate signal penetration (Fig. 2).

Conversely, CR2 and CR4 exhibit more intricate and heterogeneous bathymetric configurations, accompanied by increased variability in both estimated and observed depth values. These transects are characterized by relatively moderate slopes interspersed with undulating or flat seafloor segments, particularly beyond 200 m, where divergences between estimated and measured depths become increasingly noticeable. The estimation profiles in these areas exhibit greater fluctuations, while the dispersion of measured data points increases, suggesting a decrease in model reliability in areas with reduced contrast in bottom features, potentially due to sediment interference or weaker optical signals. Although the overall bathymetric trend is reasonably well represented, the model-data agreement is substantially weaker than for steeper but flatter transects.

The collective analysis across all transects demonstrates that the model achieves optimal performance in environments featuring steep bathymetric gradients and optically transparent waters, where seafloor features are distinct and spatially continuous. Under more complex benthic conditions - particularly where depth variations are subtle or optical signal attenuation occurs - the model still provides valid estimates but with reduced precision. The tight clustering of measured points around estimation lines throughout most profile sections nevertheless validates the model's applicability for coastal bathymetric mapping, while identifying opportunities for improvement through enhanced data smoothing algorithms and error-correction procedures in regions of increased seafloor complexity or environmental interference.

5. Conclusions

This study demonstrates the substantial potential of high-resolution PlanetScope imagery for satellite-derived bathymetry (SDB) in shallow, optically complex tropical waters, exemplified by the Nha Trang MPA, Vietnam. Through a rigorous comparative analysis between conventional empirical models (Stumpf and MLR) and a suite of advanced ML algorithms, we provide one of the most comprehensive evaluations to date for PlanetScope-based bathymetric mapping. The RF algorithm produced the most accurate depth estimates, achieving an R^2 of 0.85, an RMSE of 2.66 m, and an MAE of 1.85 m. At the same time, other gradient-boosting models (GB, XGB, CB, and LGBM) also performed strongly ($R^2 \approx 0.84$), substantially outperforming traditional approaches. Feature importance analysis consistently highlighted the contributions of Rrs_green2, Rrs_green, and Rrs_blue, underscoring their sensitivity to water-column depth.

Based on observed performance and optical constraints, the proposed ML-based SDB framework provides reliable estimates to approximately 25-30 m depth in shallow areas, offering clear practical value for habitat mapping and coastal management. Nonetheless, a few limitations should be acknowledged, including the inherent depth ceiling imposed by light attenuation in optically complex waters, potential temporal mismatches between satellite and *in-situ*

measurements, and uncertainties in transferring trained models to areas with different environmental conditions.

This research establishes a solid methodological foundation for operational SDB applications in data-limited tropical regions. It highlights the advantages of combining accessible, high-revisit satellite data with modern analytical techniques. Future work may explore physics-informed or hybrid ML-radiative-transfer approaches, assess model generalizability across seasons and water types, and integrate multisensor observations to enhance depth-retrieval accuracy and spatial coverage further.

Acknowledgements

This research was conducted under project code VB.Đ1.11/24 as part of a larger initiative by the Joint Vietnam-Russia Tropical Science and Technology Research Center. The authors would like to express their deep appreciation to colleagues at the Coastal Branch of the Joint Vietnam-Russia Tropical Science and Technology Research Center for their valuable support and contributions to the successful completion of this study.

References

Ashphaq M., Srivastava P.K., Mitra D., 2021. Review of nearshore satellite derived bathymetry: Classification and account of five decades of coastal bathymetry research. Journal of Ocean Engineering and Science, 6(4), 340−359. https://doi.org/10.1016/j.joes.2021.02.006.

Ashphaq M., Srivastava P.K., Mitra D., 2024. Satellite-derived bathymetry in dynamic coastal geomorphological environments through machine learning algorithms. Earth and Space Science, 11(7). https://doi.org/10.1029/2024EA003554.

Belyadi H., Haghighat A., 2021. Machine learning guide for oil and gas using Python: A step-by-step breakdown with data, algorithms, codes, and applications. Gulf Professional Publishing. https://doi.org/10.1016/C2019-0-03617-5.

Bertin S., Floc'h F., Le Dantec N., Jaud M., Cancouët R., Franzetti M., Cuq V., Prunier C., Ammann J., Augereau E., 2022. A long-term dataset of topography and nearshore bathymetry at the macrotidal pocket beach of Porsmilin, France. Scientific Data, 9(1), 79. https://doi.org/10.1038/s41597-022-01170-3.

Breiman L., 2001. Random forests. Machine learning, 45(1), 5−32. https://doi.org/10.1023/A:1010933404324.

Britaev T.A., Tkachenko K.S., Nguyen T.H., 2017. Modern threats to the coral communities of Nha Trang Bay (Central Vietnam). Journal of Tropical Science and Technology, 14, 3−15. https://doi.org/10.58334/vrtc.jtst.n14.01.

Caballero I., Stumpf R.P., 2023. Confronting turbidity, the major challenge for satellite-derived coastal bathymetry. Science of the Total Environment, 870. https://doi.org/10.1016/j.scitotenv.2023.161898.

Çelik O.İ., Büyüksalih G., Gazioğlu C., 2023. Improving the accuracy of satellite-derived bathymetry using multi-layer perceptron and random forest regression methods: A case study of Tavşan Island. Journal of Marine Science and Engineering, 11(11), 2090. https://doi.org/10.3390/jmse11112090.

Chen T., Guestrin C., 2016. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 785–794. https://doi.org/10.1145/2939672.2939785.

Chu S., Cheng L., Cheng J., Zhang X., Liu J., 2023. Comparison of six empirical methods for multispectral satellite-derived bathymetry. Marine Geodesy, 46(2), 149−174. https://doi.org/10.1080/01490419.2022.2132327.

Downes J., Bruce D., Miot da Silva G., Hesp P.A., 2025. Optimising Satellite-Derived Bathymetry Using Optical Imagery over the Adelaide Metropolitan Coast. Remote Sensing, 17(5), 849. https://doi.org/10.20944/preprints202502.0544.v1.

Duan Z., Chu S., Cheng L., Ji C., Li M., Shen W., 2022. Satellite-derived bathymetry using Landsat-8 and Sentinel-2A images: assessment of atmospheric correction algorithms and depth derivation models in shallow waters. Optics Express, 30(3), 3238−3261. https://doi.org/10.1364/oe.444557.

Duc H.N., Tien G.N., Le Xuan H., Ngoc V.T., Huu D.N., 2024. Multi-step-ahead prediction of water levels using machine learning: A comparative analysis in the Vietnamese Mekong Delta. Vietnam Journal of Earth Sciences, 46(4), 468−488. https://doi.org/10.15625/2615-9783/21067.

Eugenio F., Marcello J., Mederos-Barrera A., Marqués F., 2021. High-resolution satellite bathymetry mapping: Regression and machine learning-based approaches. IEEE Transactions on Geoscience and Remote Sensing, 60, 1−14. https://doi.org/10.1109/tgrs.2021.3135462.

Gabr B., Ahmed M., Marmoush Y., 2020. PlanetScope and Landsat 8 imageries for bathymetry mapping. Journal of Marine Science and Engineering, 8(2), 143. https://doi.org/10.3390/jmse8020143.

Gholami R., Fakhari N., 2017. Support vector machine: principles, parameters, and applications. In Handbook of neural computation. Elsevier, 515−535. https://doi.org/10.1016/b978-0-12-811318-9.00027-2.

Gülher E., Alganci U., 2023a. Satellite-derived bathymetry mapping on horseshoe island, Antarctic peninsula, with open-source satellite images: evaluation of atmospheric correction methods and empirical models. Remote Sensing, 15(10), 2568. https://doi.org/10.3390/rs15102568.

Gülher E., Alganci U., 2023b. Satellite-derived bathymetry in shallow waters: evaluation of Gokturk-1 satellite and a novel approach. Remote Sensing, 15(21), 5220. https://doi.org/10.3390/rs15215220.

He J., Zhang S., Cui X., Feng W., 2024. Remote sensing for shallow bathymetry: A systematic review. Earth-Science Reviews, 258, 104957. https://doi.org/10.1016/j.earscirev.2024.104957.

Hodúl M., Bird S., Knudby A., Chénier R., 2018. Satellite derived photogrammetric bathymetry. ISPRS Journal of Photogrammetry and Remote Sensing, 142, 268−277. https://doi.org/10.1016/j.isprsjprs.2018.06.015.

Huu D.N., Cong T.V., Bretcan P., Petrisor A.-I., 2024. Assessing the relationship between landslide susceptibility and land cover change using machine learning. Vietnam Journal of Earth Sciences, 46(3), 339−359. https://doi.org/10.15625/2615-9783/20706.

Jupp D.L.P., 1988. Background and extensions to depth of penetration (DOP) mapping in shallow coastal waters. Proceedings of the symposium on remote sensing of the coastal zone, Gold Coast, IV(2), 1−19.

Kalybekova A., 2025. A Review of Advancements and Applications of Satellite-Derived Bathymetry. Engineered Science, 35, 1541. https://doi.org/10.30919/es1541.

Khakhim N., Kurniawan A., Wicaksono P., Hasrul A., 2024. Assessment of Empirical Nearshore Bathymetry Model Using New Emerged PlanetScope Instrument and Sentinel-2 Data in Coastal Shallow Waters. International Journal of Geoinformatics, 20(2), 95−105. https://doi.org/10.52939/ijg.v20i2.3071.

Kujawa P., Remondino F., 2025. A Review of Image- and LiDAR-Based Mapping of Shallow Water Scenarios. Remote Sensing, 17(12), 2086. https://doi.org/10.3390/rs17122086.

Kwon J.-Y., Shin H.-Y., Kim D.-H., Lee H.-G., Bouk J.-K., Kim J.-H., Kim T.-H., 2024. Estimation of shallow bathymetry using Sentinel-2 satellite data and random forest machine learning: a case study for Cheonsuman, Hallim, and Samcheok Coastal Seas. Journal of Applied Remote Sensing, 18(1), 014522−014522. https://doi.org/10.1117/1.jrs.18.014522.

Li F., Zhang L., Chen B., Gao D., Cheng Y., Zhang X., Yang Y., Gao K., Huang Z., Peng J., 2018. A light gradient boosting machine for remainning useful life estimation of aircraft engines. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 3562−3567. https://doi.org/10.1109/itsc.2018.8569801.

Li J., Knapp D.E., Schill S.R., Roelfsema C., Phinn S., Silman M., Mascaro J., Asner G.P., 2019. Adaptive bathymetry estimation for shallow coastal waters using Planet Dove satellites. Remote Sensing of Environment, 232, 111302. https://doi.org/10.1016/j.rse.2019.111302.

Li Z., Peng Z., Zhang Z., Chu Y., Xu C., Yao S., García-Fernández Á.F., Zhu X., Yue Y., Levers A., 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. https://doi.org/10.3389/fmars.2023.1178845.

Liu Y., Wu S., Wu Z., Zhou S., 2024. Application of gradient boosting machine in satellite-derived bathymetry using Sentinel-2 data for accurate water depth estimation in coastal environments. Journal of Sea Research, 201, 102538. https://doi.org/10.1016/j.seares.2024.102538.

Liu Z., Liu H., Ma Y., Ma X., Yang J., Jiang Y., Li S., 2024. Exploring the most effective information for satellite-derived bathymetry models in different water qualities. Remote Sensing, 16(13), 2371. https://doi.org/10.3390/rs16132371.

Lyzenga D.R., 1978. Passive remote sensing techniques for mapping water depth and bottom features. Applied optics, 17(3), 379−383. https://doi.org/10.1364/ao.17.000379.

Lyzenga D.R., Malinas N.P., Tanis F.J., 2006. Multispectral bathymetry using a simple physically based algorithm. IEEE Transactions on Geoscience and Remote Sensing, 44(8), 2251−2259. https://doi.org/10.1109/tgrs.2006.872909.

Mobley C.D., Stramski D., Paul Bissett W., Boss E., 2004. Optical modeling of ocean waters: Is the case 1-case 2 classification still useful? Oceanography, 17(2), 60−67. https://doi.org/10.5670/oceanog.2004.48.

Moeinkhah A., Shakiba A., Azarakhsh Z., 2019. Assessment of regression and classification methods using remote sensing technology for detection of coastal depth (case study of Bushehr Port and Kharg Island). Journal of the Indian Society of Remote Sensing, 47(6), 1019−1029. https://doi.org/10.1007/s12524-019-00959-x.

Nguyen H.Q., Banno M., Ha N.-T., 2025. Satellite derived bathymetry using empirical and machine learning approaches: a case study in the highly dynamic coastal water. Coastal Engineering Journal, 67(2), 232−251. https://doi.org/10.1080/21664250.2024.2445418.

Nguyen Tac An, Le Dinh Mau, Pavlov D., 2007. Assessment of hydro-dynamical processes for ecosystem studies in Nha Trang bay, Vietnam. Collection of Reports of the National Conference "East Sea-2007", Nha Trang, 379−390.

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 12, 2825−2830. https://doi.org/10.48550/arXiv.1201.0490.

Prokhorenkova L., Gusev G., Vorobev A., Dorogush A.V., Gulin A., 2018. CatBoost: unbiased boosting with categorical features. Advances in Neural Information Processing Systems, 31, 6639–6649.
https://doi.org/10.48550/arXiv.1706.09516.

Quinten V., 2024. Acolite. Retrieved 15/4/2025 from https://github.com/acolite/acolite.

Saeidi V., Seydi S.T., Kalantar B., Ueda N., Tajfirooz B., Shabani F., 2023. Water depth estimation from Sentinel-2 imagery using advanced machine learning methods and explainable artificial intelligence. Geomatics, Natural Hazards and Risk, 14(1), 2225691.
https://doi.org/10.1080/19475705.2023.2225691.

Sesama A.S., Setiawan K.T., Julzarika A., 2020. Bathymetric Extraction Using Planetscope Imagery (Case Study: Kemujan Island, Central Java). International Journal of Remote Sensing and Earth Sciences (IJReSES), 17(2), 209−216. https://doi.org/10.30536/j.ijreses.2020.v17.a3445.

Stumpf R.P., Holderied K., Sinclair M., 2003. Determination of water depth with high-resolution satellite imagery over variable bottom types. Limnology and oceanography, 48(1-2), 547−556. https://doi.org/10.4319/lo.2003.48.1_part_2.0547.

Tatsuyuki Sagawa, Masanao Sumiyoshi, Honami Watanabe, Yamashita Y., 2023. Bathymetry from PlanetScope imagery using random forest and deep learning. 2023 Asian Conference on Remote Sensing (ACRS2023), 1, 601−607.

Tkachenko K.S., 2015. Ecological status of coral communities in the island area of the Nha Trang Bay (Vietnam). Russian Journal of Ecology, 46(5), 456−462.
https://doi.org/10.1134/s1067413615050197.

Tkachenko K., 2023. Degradation of coral reefs under complex impact of natural and anthropogenic factors with Nha Trang Bay (Vietnam) as an example. Biology Bulletin Reviews, 13(5), 442−459. https://doi.org/10.1134/s2079086423050079.

Tran H.T.T., Nguyen Q.H., Pham T.H., Ngo G.T.H., Pham N.T.D., Pham T.G., Tran C.T.M., Ha T.N., 2024. Novel learning of bathymetry from landsat 9 imagery using machine learning, feature extraction and meta-heuristic optimization in a shallow Turbid Lagoon. Geosciences, 14(5), 130. https://doi.org/10.3390/geosciences14050130.

Wicaksono P., Harahap S.D., Hendriana R., 2024. Satellite-derived bathymetry from WorldView-2 based on linear and machine learning regression in the optically complex shallow water of the coral reef ecosystem of Kemujan island. Remote Sensing Applications: Society and Environment, 33, 101085. https://doi.org/10.1016/j.rsase.2023.101085.

Wu Z., Mao Z., Shen W., 2021. Integrating multiple datasets and machine learning

algorithms for satellite-based bathymetry in seaports. Remote Sensing, 13(21), 4328. https://doi.org/10.3390/rs13214328.

Wulandari S.A., Wicaksono P., 2021. Bathymetry mapping using PlanetScope imagery on Kemujan Island, Karimunjawa, Indonesia. IOP Conference Series: Earth and Environmental Science, 686(1), 012032. https://doi.org/10.1088/1755-1315/686/1/012032.

Xie C., Chen P., Zhang Z., Pan D., 2023. Satellite-derived bathymetry combined with Sentinel-2 and ICESat-2 datasets using machine learning. Frontiers in Earth Science, 11, 1111817. https://doi.org/10.3389/feart.2023.1111817.

Ye L., Chu S., Chen H., Cheng J., Xu R., Qu Z., 2025. Hyperspectral Satellite-Derived Bathymetry Considering Substrate Characteristics. Marine Geodesy, 1−25. https://doi.org/10.1080/01490419.2025.2460216.

Ye M., Yang C., Zhang X., Li S., Peng X., Li Y., Chen T., 2024. Shallow Water Bathymetry Inversion Based on Machine Learning Using ICESat-2 and Sentinel-2 Data. Remote Sensing, 16(23), 4603. https://doi.org/10.3390/rs16234603.

Zhang Y., Liu J., Shen W., 2022. A review of ensemble learning algorithms used in remote sensing applications. Applied Sciences, 12(17), 8654. https://doi.org/10.3390/app12178654.

APPENDIX

Table A1. PlanetScope image profile used for bathymetry mapping

Scene ID	Date of acquisition	Cloud coverage (%)	Spatial resolution (m)	Radiometric resolution (bit)	View angle (∘)	Scene swath (km)
20240424_022649_77_24d0	24 April 2024	0	3	12	4.8	16.1

Table A2. Number of observations per 5-m depth strata

Depth bin (m)	Number of observations
0−5	441
5−20	7,127
> 20	1,981

Table A3. ACOLITE parameters for atmospheric correction

Parameter	Value
dsf_interface_reflectance	True
min_tgas_aot	0.85
min_tgas_rho	0.70
dsf_residual_glint_correction	True
adjacency_correction	True
dsf_aot_estimate	fixed
Output	Surface reflectance (R_{rs})

Table A4. Hyperparameters used for ML models

RF	**max_features**	**sqrt**	**SVM**	**kernel**	**RBF**
	min_sample_leaf	3		C	80.84
	n_estimators	144		Epsilon	0.46
	min_sample_split	6		gamma	scale
XGB	subsample	0.88	LGBM	subsample	0.86
	colsample_bytree	0.94		learning_rate	0.03
	learning_rate	0.03		max_depth	None
	max_depth	20		n_estimators	170
	n_estimators	253		num_leaves	67
CB	depth	10	GB	min_child_samples	25
	iterations	487		learning_rate	0.02
	learning_rate	0.05		max_depth	9
	l2_leaf_reg	2		max_features	sqrt
				min_samples_leaf	4
				min_samples_split	6
				n_estimators	198
				subsample	0.84

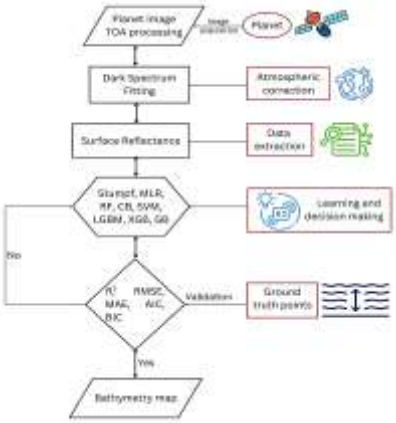

Figure A1. A straightforward workflow of bathymetric retrieval in the Nha Trang MPA

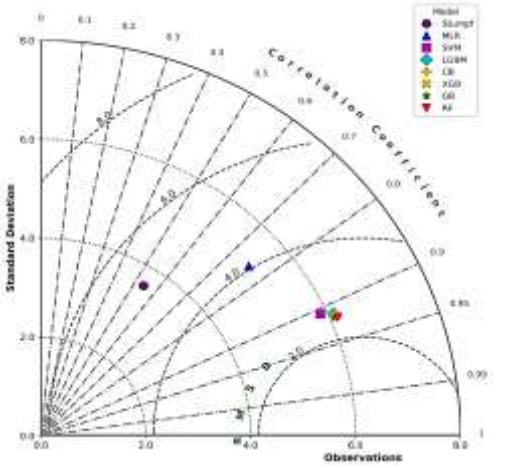

Figure A2. Model comparison using Taylor plot

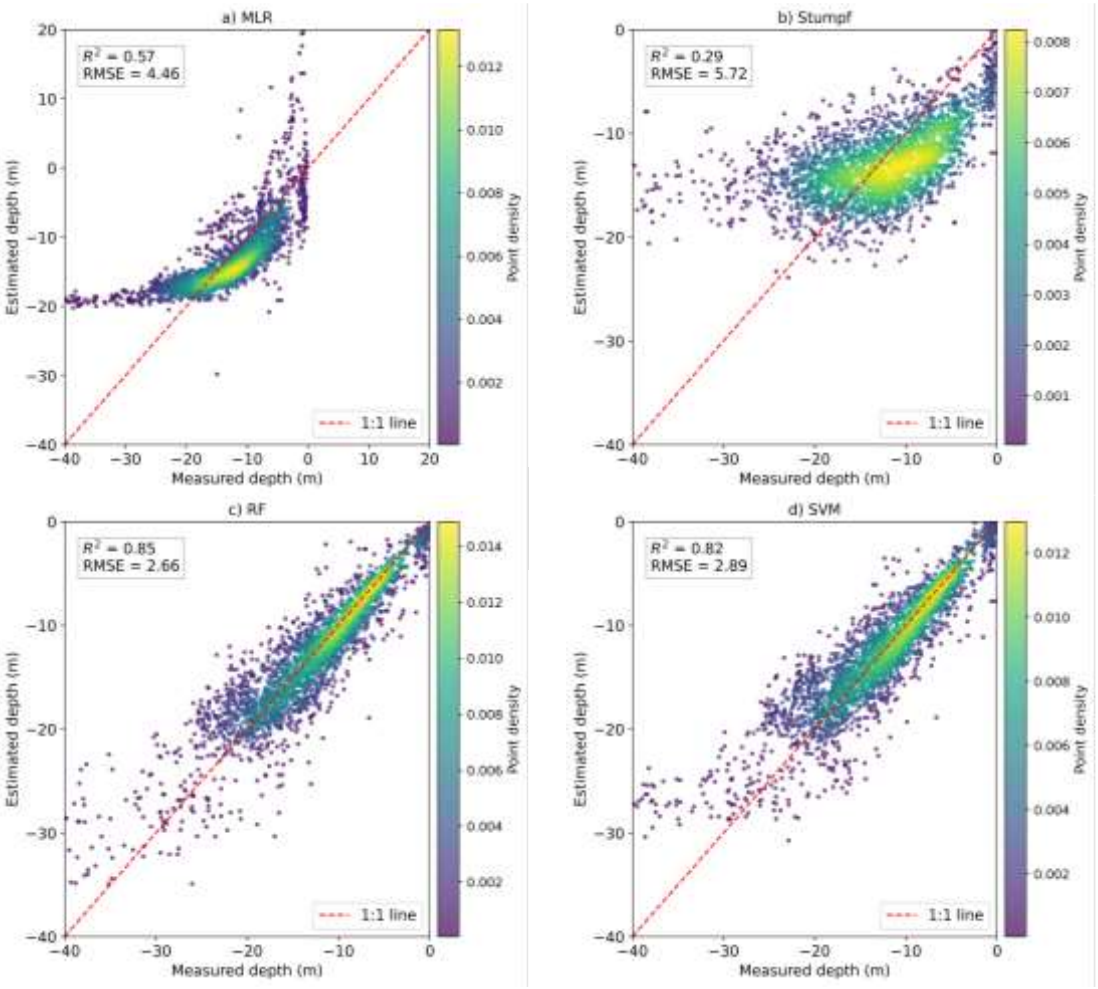

Figure A3. Model performance in different forms of scatter plot

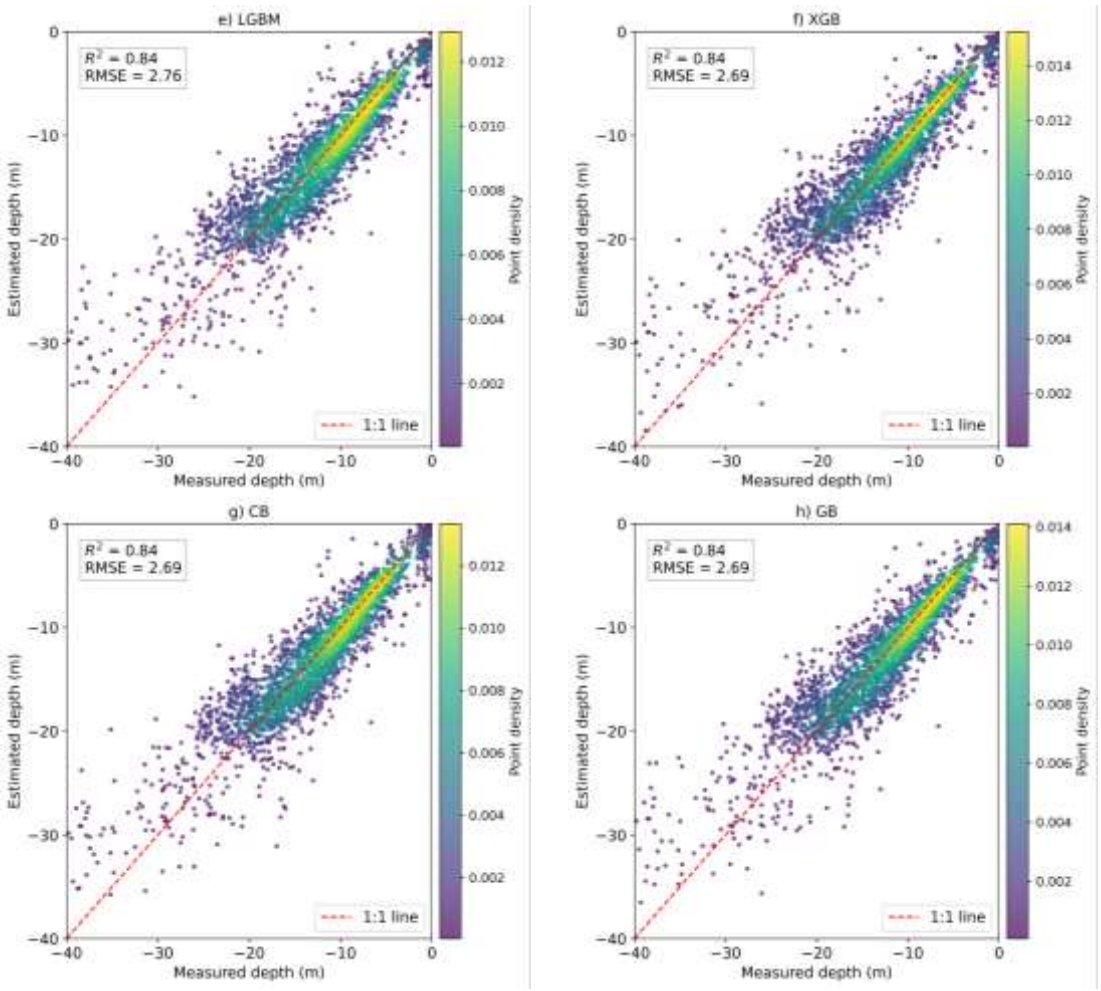

Figure A3. Cont.

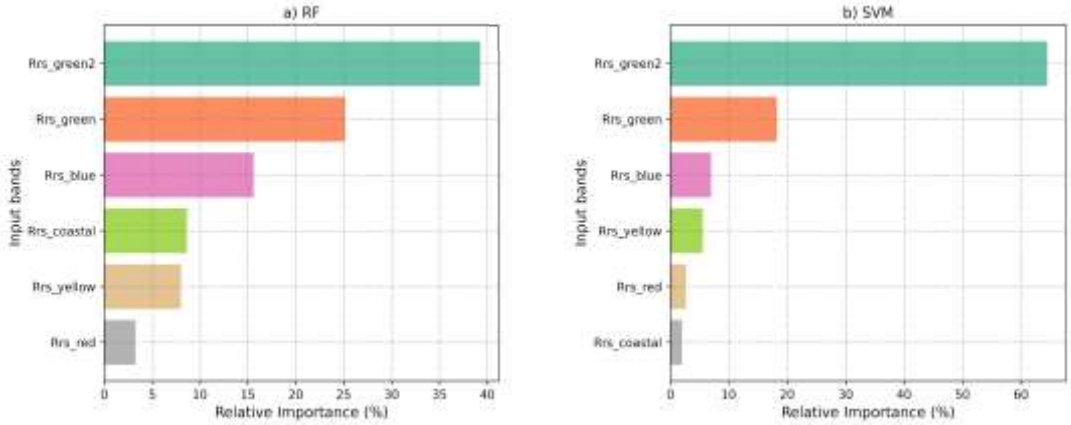

Figure A4. Feature importance derived from ML models

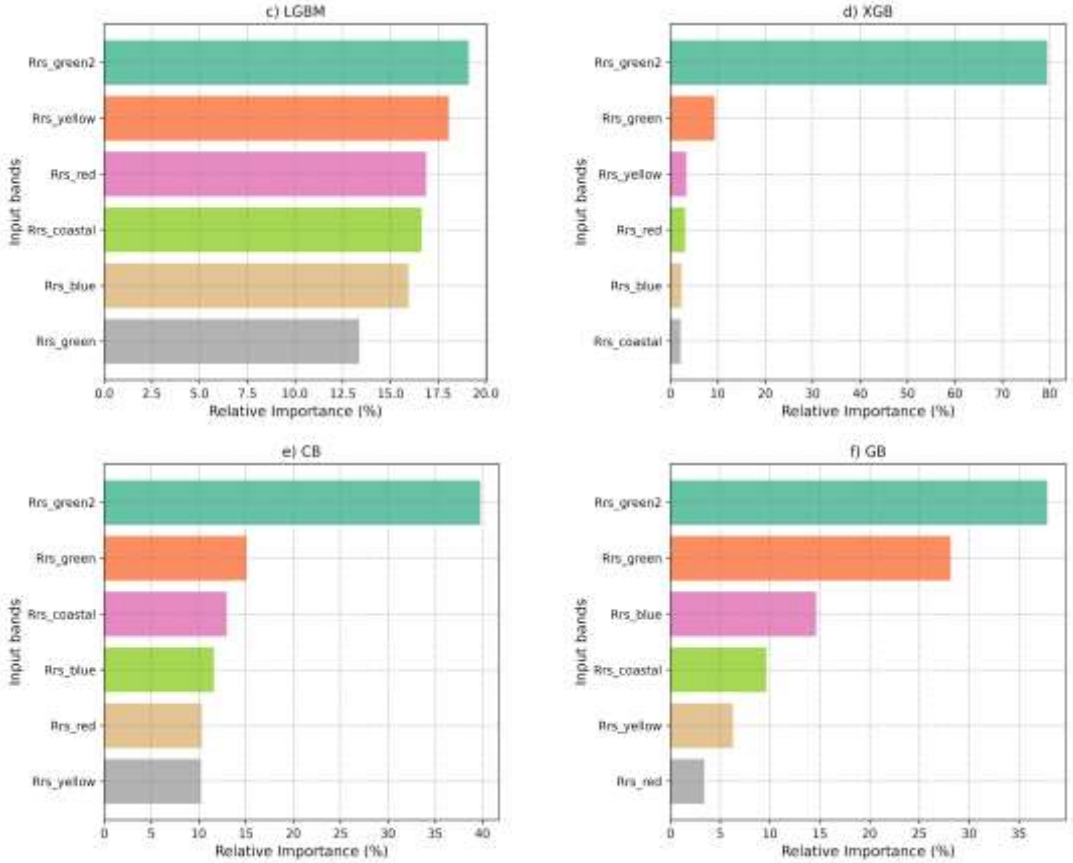

Figure A4. Cont.

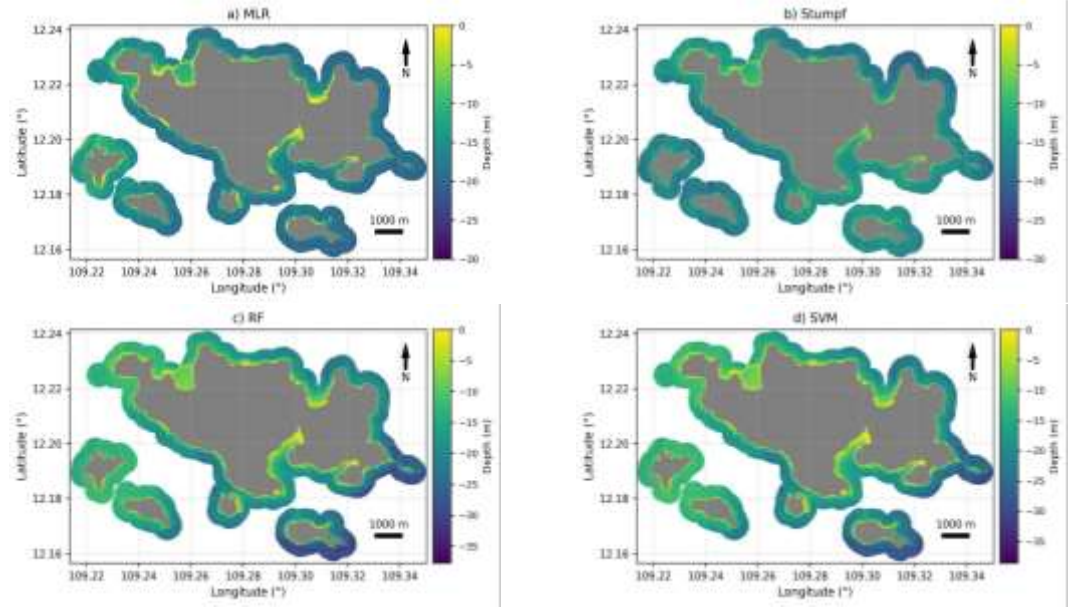

Figure A5. Bathymetry map derived from proposed models in the Nha Trang MPA

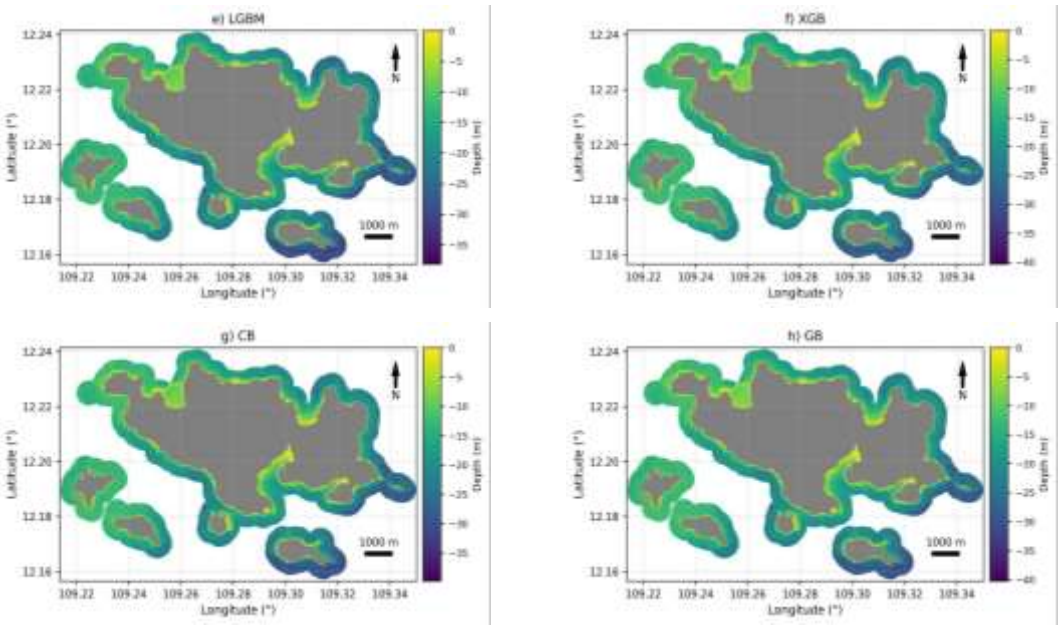

Figure A5. Cont.

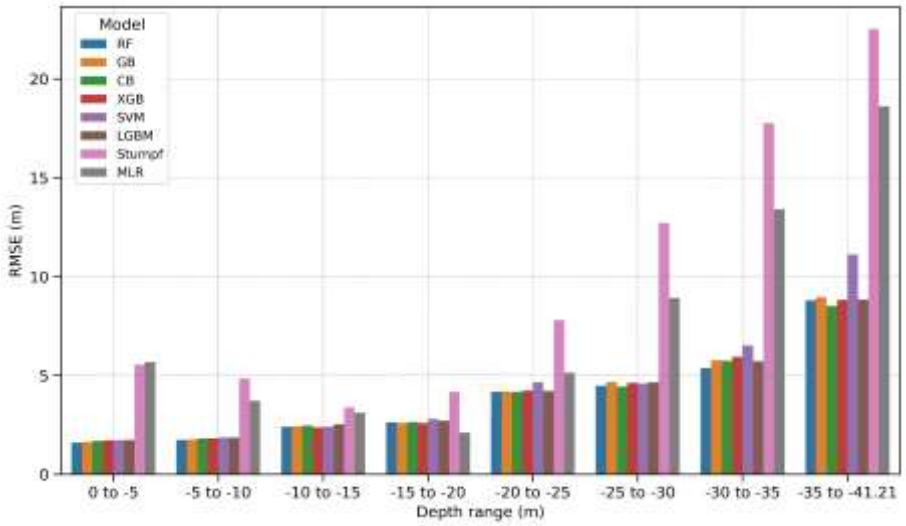

Figure A6. Bathymetry model accuracy across depth ranges