Equivalent: Jurnal Ilmiah Sosial Teknik
Volume 7, Issue 2, 142-155
e_ISSN: 2775-0833
https://jurnalequivalent.id/index.php/jequi/index
DOI: doi.org/10.59261/jequi.v7i2.252

Shallow Water Bathymetry Mapping Using Sentinel-2 and Machine Learning in Raja Ampat Coastal Waters

***Muhammad Ichsan**
Universitas Dr. Soetomo, Indonesia

Septa Erik Prabawa
Universitas Dr. Soetomo, Indonesia

Reynalda Anindia Mawarni
PT. Anable Hexagon Performa, Indonesia, Indonesia

***Corresponding author:**
Muhammad Ichsan, Universitas Dr. Soetomo, Indonesia
✉ info@unitomo.ac.id

Article Info :
Received: October 15, 2025
Revised: November 4, 2025
Accepted: December 17, 2025

Keywords:
satellite-derived bathymetry; sentinel-2 imagery; random forest; shallow water bathymetry; raja ampat waters.

Abstract

Background: Satellite-Derived Bathymetry (SDB) has emerged as a cost-effective alternative for shallow-water depth mapping, particularly in remote and data-scarce marine regions. Raja Ampat, Southwest Papua, presents a challenging environment for bathymetric mapping due to its complex seafloor morphology and high water clarity, requiring robust analytical approaches to improve depth estimation accuracy.
Objective: This study aims to evaluate the effectiveness of Sentinel-2–based SDB by comparing empirical regression models and a Random Forest machine learning approach for shallow-water bathymetry mapping in the waters of Raja Ampat.
Methods: The research integrates Sentinel-2 multispectral imagery with three empirical regression methods—green band power regression, blue band power regression, and the Stumpf logarithmic ratio method—and a Random Forest algorithm. Field bathymetric data obtained from fishfinder measurements were used for model calibration and validation. All processing and analysis were conducted using the Google Earth Engine (GEE) platform.
Results: The empirical regression analysis shows that the green band model (Band 3) achieved the highest accuracy (R^2 = 0.7097; RMSE = 1.80–14.12 m across depth classes), followed by the blue band model (R^2 = 0.6194) and the Stumpf method (R^2 = 0.5693). The Random Forest model outperformed all empirical approaches by effectively capturing non-linear relationships between spectral reflectance and water depth, particularly over heterogeneous seabed substrates. The green band model performed optimally at depths of 0–20 m, while the Stumpf method demonstrated greater stability beyond 20 m.
Conclusion: The integration of Sentinel-2 imagery with machine learning provides an accurate, scalable, and cost-efficient solution for shallow-water bathymetric mapping in tropical archipelagic environments. This approach supports marine conservation planning and coastal resource management in Raja Ampat and similar regions.

To cite this article: Ichsan, M., Prabawa, S. E., & Mawarni, R. A. (2025). Shallow Water Bathymetry Mapping Using Sentinel-2 and Machine Learning in Raja Ampat Coastal Waters. *Equivalent: Jurnal Ilmiah Sosial Teknik,* 7(2), 142-155. https://doi.org/10.59261/jequi.v7i2.252

INTRODUCTION

Raja Ampat is a Coral Triangle region with a complex and rich marine ecosystem, making bathymetric mapping crucial for conservation. Shallow marine bathymetric information is essential for various fields, including marine navigation, coastal ecosystem conservation, marine tourism planning, and disaster mitigation such as tsunamis and abrasion (Asaad et al., 2018; Huang & Coelho, 2017; White et al., 2014). Raja Ampat is known for having the highest marine biodiversity in the world, making shallow bathymetric mapping crucial for conservation area management. Shallow Bathymetry Shallow marine depth mapping is traditionally conducted using echosounders. However, for large and remote areas, this method is not always efficient.

In-situ survey methods such as single beam and multibeam echosounders can produce high-precision data, but are constrained by cost, time, and limited access in archipelagic regions (Bekiashev, 1981). Traditionally, echosounders are used to map shallow water depths, but they are less efficient in large, remote areas (Lyzenga, 1978). Therefore, the Shallow Depth Bathymetry (SDB) method, which utilizes remote sensing data, offers an alternative solution (Misra et al., 2018). Shallow Depth Bathymetry (SDB) is a technique that utilizes remote sensing imagery, particularly optical satellite data, to map shallow water depths. SDB utilizes the empirical relationship between the intensity of light reflection (reflectance) in the satellite's spectral channel and water depth (Lyzenga, 1978).

This method works based on the relationship between the spectral reflectance measured by the satellite sensor and the water depth below. This technique is effective in clear waters, where sunlight can penetrate the water column to the seabed and be reflected back to the sensor (Caballero & Stumpf, 2020). Shallow Depth Bathymetry (SDB) The SDB method was first introduced by Lyzenga, utilizing the logarithmic ratio of blue and green reflectance bands. SDB is widely applied with Landsat, Sentinel, and WorldView data (Stumpf et al., 2003). In SDB, the reflectance of light (R) at a specific wavelength measured by a satellite sensor is the result of complex interactions between water, the seabed substrate, and the water column. The basic model that explains this relationship is namely:

$$R(\lambda)=R0(\lambda)\ e^{(-kd(\lambda)Z)}$$

Where:

R(λ) = Reflectance at a certain wavelength, with units of %

R0(λ)e = Reflectance of the seabed substrate at that wavelength, with unit %

kd(λ) = Light attenuation coefficient in the water column, depending on the length waves with their coefficient values:

1. For very clear sea water (e.g. type I Jerlov water):
   a. Kd (490 nm) (green-blue) ≈ 0.02 – 0.05 m −1
   b. Kd (550 nm) (green) ≈ 0.05 – 0.1 m −1
   c. Kd (600–700 nm) (red) can be > 0.2 m −1 because red is quickly damped
2. For coastal waters (type II/III Jerlov water):
   a. It can rise to 0.1 – 0.5 m −1 or even more, depending on the turbidity
3. For turbid waters/ estuaries:
   a. Can be > 1 m −1 at certain wavelengths

Z = Water depth in meters

Light with shorter wavelengths (such as blue) penetrates deeper than longer wavelengths (such as red), so the blue and green bands in satellite imagery are often used for depth estimation (Fan et al., 2022). The blue and green bands (such as B2 and B3 in Sentinel-2) are the most relevant in SDB because they have better penetration ability in the water column. Meanwhile, the red or infrared band is used to distinguish between water and land areas (Hewageegana, S., & Canestrelli, A. 2022). Accurate ocean depth data collection, designed to support environmental conservation and coastal resource management in Raja Ampat, is extremely limited. Therefore, it is necessary to apply the SDB method to provide information on water depth along the Raja Ampat coast and to evaluate the accuracy of the SDB method by comparing it with in-situ bathymetry data.

Despite extensive SDB applications globally, a critical research gap persists in integrating traditional empirical regression approaches with advanced machine learning techniques for bathymetric mapping in complex tropical archipelagic environments. While previous studies have predominantly relied on either empirical models (Stumpf et al., 2003; Lyzenga, 1978) or machine learning methods separately (Al Najar et al., 2022; Hewageegana & Canestrelli, 2022), few have systematically compared their performance in regions with exceptionally high water clarity and diverse substrate compositions such as Raja Ampat.

Research on Satellite-Derived Bathymetry (SDB) has expanded considerably in recent years, particularly with the increasing availability of medium- and high-resolution satellite imagery. Caballero and Stumpf (2020) demonstrated that empirical spectral-ratio-based

methods, such as the logarithmic band ratio, can produce relatively stable depth estimates in optically clear waters; however, their performance deteriorates in environments characterized by heterogeneous seabed substrates and complex bathymetric features due to simplifying assumptions regarding water column homogeneity. In contrast, Hewageegana and Canestrelli (2022) reported that machine learning approaches offer superior capability in capturing non-linear relationships between spectral reflectance and water depth, yet their study primarily focused on algorithmic comparison without a systematic evaluation against conventional empirical models across different depth classes. Although both studies provide important contributions, a critical research gap remains in the comprehensive integration and comparative assessment of empirical regression techniques and machine learning models within complex tropical archipelagic environments exhibiting exceptionally high water clarity and diverse benthic conditions, such as Raja Ampat.

The novelty of this study lies in: (1) the hybrid methodological framework combining power regression models with Random Forest machine learning implemented through Google Earth Engine (GEE) for operational efficiency; (2) comprehensive validation across multiple depth classes (0–5 m, 5–10 m, 10–20 m, and >20 m) to establish depth-specific model applicability; and (3) application in Raja Ampat's unique hydrographic conditions, characterized by pristine water quality (Jerlov Type I waters) and complex coral reef topography. This integrated approach provides actionable insights for selecting optimal SDB methodologies based on depth characteristics and environmental conditions, thereby advancing both the theoretical understanding and practical implementation of satellite-derived bathymetry in remote marine protected areas.

This study addresses that gap by implementing a hybrid methodological framework that combines empirical regression models (blue band, green band, and Stumpf methods) with a Random Forest algorithm using the Google Earth Engine platform, accompanied by rigorous validation across multiple depth classes. The primary objective of this research is to evaluate the relative performance and depth-specific applicability of each SDB approach under unique hydrographic conditions. The findings are expected to contribute academically by advancing the theoretical understanding of SDB model behavior in tropical marine environments, and practically by providing a cost-effective, scalable, and reliable bathymetric mapping solution to support marine conservation planning, coastal resource management, and sustainable governance of remote marine protected areas such as Raja Ampat.

METHOD

This research was conducted in the shallow waters of the Raja Ampat Coast, Southwest Papua using the shallow water depth estimation method (Shallow Depth Bathymetry/SDB) on June 10, 2023. The tools and materials used in this research are:

Tools and materials

Table 1. Tools and materials

Component	Information
Satellite Imagery	Sentinel-2 Level-2A (atmospherically corrected), Main bands: B2 (Blue), B3 (Green), B4 (Red), B8 (NIR) 10 meter resolution 5 day revision cycle
Field Data	Bathymetry from Fishfinder (acoustic sonar) is used as reference data for calibration & validation.
Processing Platform	Google Earth Engine (GEE) Image processing, masking, overlay, regression, depth prediction and visualization
Additional Software	Microsoft Excel (for regression analysis of CSV data)

Source: Personal document

Table 1 presents the tools and materials utilized in this study for Satellite-Derived Bathymetry (SDB) analysis in Raja Ampat waters. Sentinel-2 Level-2A satellite imagery was employed as the primary remote sensing data source due to its high spatial resolution (10 m), frequent revisit cycle (5 days), and atmospheric correction, which enhances water column reflectance accuracy. The blue (B2) and green (B3) bands were specifically selected for bathymetric estimation because of their superior penetration capabilities in clear water environments, while the red (B4) and near-infrared (B8) bands were used for land–water masking and noise reduction.

Satellite Imagery Data and Resolution

Sentinel-2 Level-2A imagery was chosen because it has Blue Bands (B2) and Green Bands (B3) that can penetrate the water column to a depth of 20-30 meters (Wilson et al. 2022). Therefore, the 10-meter spatial resolution is suitable for coastal area mapping. Sentinel-2 allows for rapid analysis of temporal changes because the revision cycle only takes 5 days under optimal conditions. This is supported by research by Al Najar et al. (2022) regarding Sentinel-2 having a normal revision cycle of 5 days that allows for rapid and efficient monitoring of temporal changes in coastal areas. The combination of Blue Bands (B2), Green Bands (B3), and Red Bands (B4) produced by Sentinel-2 imagery produces RGB images for visualization. Meanwhile, the NIR Band (B8) is used to separate land and water where land reflectance is high and water is low (Casal et al. 2019).

Field Bathymetry Data

Field bathymetric data was obtained from direct measurements using a Fishfinder (aquatic sonar) to detect seawater depth. It works by emitting sound waves that bounce off the seabed. The time it takes for the sound waves to return is used to calculate depth, thus obtaining an estimated depth. Data obtained from Fishfinder has a high level of accuracy, so it is often used as reference data to calibrate and validate the depth estimation results from the Shallow Depth Bathymetry (SDB) method.

Processing Platform

The use of Google Earth Engine (GEE) in Shallow Depth Bathymetry (SDB) analysis is very useful for processing Sentinel-2 satellite imagery, which has high resolution and relevant spectral bands for shallow water column analysis (Kurniawan et al., 2024; Cahyono & Sihombing, 2025). As in the study of Traganos (2018) explained that through GEE, spectral bands such as blue (B2) and green (B3) can be easily accessed and processed to establish a relationship between spectral reflectance and water depth. This process includes image import, land masking, field data overlay, statistical regression, depth prediction using Fishfinder, and thematic map visualization to facilitate evaluation.

GEE's key advantages include time and cost efficiency, large-scale analysis capabilities, and easy access to geospatial data and analysis tools. However, the platform has limitations, such as its reliance on an internet connection and the need for programming knowledge. Nevertheless, GEE remains a highly effective tool in SDB research, particularly in processing satellite data and integrating field data to produce accurate and efficient depth prediction models.

Data Processing Stages

Data processing in this study consists of 13 stages, including:

1. Determine the study area through area boundary variables.
2. Import Sentinel-2 satellite imagery taken from the Google Earth Engine (GEE) catalog.
3. Band Selection and Location Information Addition. The selected bands are RGB bands (B2, B3, B4) relevant for SDB analysis, selected using the select method. In addition, geographic coordinates (latitude, longitude) and UTM projection coordinates are added as additional bands.
4. RGB images are visualized on a map using specific color scale parameters (rgbVis). As shown in Figure 1.

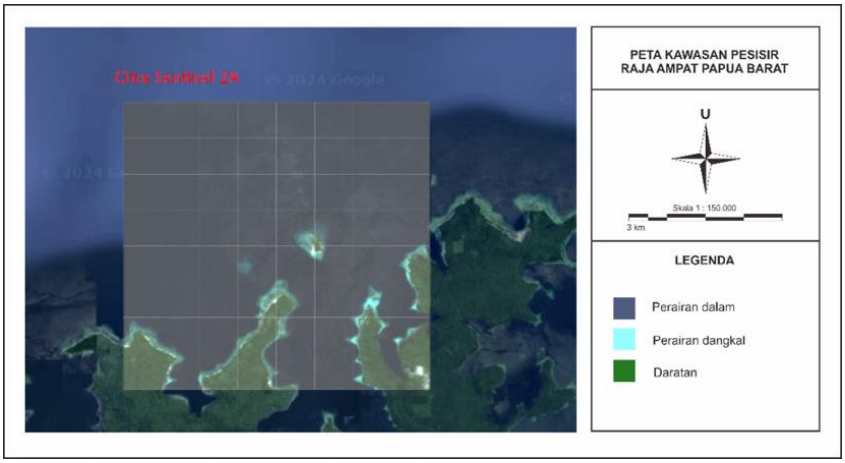

Figure 1. Satellite imagery used/selected
Source: Personal document

5. The masking process to separate water areas from land uses the near-infrared band (B8). This band was chosen because land reflectance tends to be higher than water. A threshold value (NIR_thres) is determined to distinguish land and water pixels. Pixels with reflectance values above the threshold are considered land, while those below the threshold are considered water.
6. Masking verification involves visualizing the masking results to verify the spatial separation of land and water. As shown in Figure 2.

Figure 2. Results of masking land and sea
Source: Personal document

7. The division of field data, namely field bathymetry data (batipoin data) in the form of depth measurement points, is randomly separated into two subsets, namely (70%) Training dataset to train the regression model, and (30%) Validation dataset which is used to evaluate accuracy, as in Figure 3.

system:index	depth	B2	B3	B4	x	y	longitude	latitude
00000000000000000d1_0	51.029	1,224	1,119	1,078	694,215	-84,635	130.745	-0.765
000000000000000158_0	66.834	1,232	1,142	1,082	695,345	-85,975	130.755	-0.777
00000000000000003e_0	5.478	1,398	1,360	1,054	693,925	-84,375	130.743	-0.763
00000000000000007f_0	29.586	1,168	1,099	1,024	695,795	-86,065	130.759	-0.778
000000000000000092_0	34.665	1,368	1,288	1,225	696,135	-83,775	130.762	-0.758
000000000000000184_0	70.851	1,235	1,171	1,095	695,495	-85,035	130.757	-0.769
00000000000000016f_0	68.72	1,191	1,103	1,061	694,205	-83,635	130.745	-0.756
0000000000000001a6_0	72.931	1,150	1,068	1,009	692,025	-84,365	130.726	-0.763
000000000000000169_0	68.274	1,149	1,083	1,021	695,035	-86,005	130.753	-0.778
0000000000000001a0_0	72.64	1,169	1,073	1,013	692,095	-84,255	130.726	-0.762

1 2 10 30 35

(a)

system:index	depth	B2	B3	B4	x	y	longitude	latitude
0000000000000000b7_0	45.828	1,145	1,066	1,029	693,595	-85,625	130.74	-0.774
000000000000000004b_0	9.574	1,478	1,469	1,138	694,675	-85,245	130.749	-0.771
000000000000000000f_0	0.95	1,712	1,758	1,225	694,245	-85,055	130.746	-0.769
0000000000000000194_0	71.834	1,132	1,075	1,012	693,035	-83,595	130.735	-0.756
0000000000000001d1_0	76.62	1,162	1,086	1,026	692,565	-83,965	130.73	-0.759
0000000000000000c4_0	48.686	1,244	1,136	1,094	694,255	-84,595	130.746	-0.765
0000000000000001bd_0	75.31	1,160	1,080	1,020	692,865	-84,235	130.733	-0.762
0000000000000000a7_0	41.755	1,182	1,090	1,039	694,825	-83,585	130.751	-0.756
000000000000000018a_0	71.292	1,179	1,113	1,046	694,065	-82,935	130.744	-0.75
000000000000000086_0	31.152	1,190	1,091	1,035	694,145	-84,945	130.745	-0.768

1 2 10 13

(b)

Figure 3. Image of field bathymetric measurement data
Source: Personal document

8. Visualization of bathymetric points, namely measurement points are determined as in Figure 4.

Figure 4. Bathymetric data mapping
Source: Personal document

9. Field Data Overlay Process with Imagery. The overlay process is performed to associate the reflectance values of spectral bands (B2, B3, B4) with the depth data at each sampling point. This operation is performed using the sample Regions function.
10. The resulting overlay dataset is exported to CSV format using the Export.table.toDrive function. This export facilitates further regression analysis using statistical software such as Excel.
11. The application of the regression model carried out in Excel software is the green power regression band (B3), blue power regression band (B2) and Stumpf (B2/B3 ratio) to obtain the regression equation.

12. Prediction and visualization of thematic depth maps based on models applied to water imagery.
13. The Random Forest model is applied as a machine learning approach to predict depth. This model is trained using training data with input in the form of spectral band reflectance (B2, B3, B4) and a target depth. The masked water imagery (marine imagery) is classified using the Random Forest model. The classification result is a depth prediction map, which is visualized separately.

Depth Prediction Model

Application to water areas after the masking process is calculated using the following equation:

Power regression (B2)

$$Y_1 = B2^b \; x \; a$$

Power regression (B3)

$$Y_2 = B3^b \; x \; a$$

Stumpf Model

$$Y_3 = m_1 \; x \ln\left(\frac{B2}{B3}\right) + m_0$$

In this case,

Y : predicted depth value (meters)
a & b : coefficients
m 1 : slope coefficient
m_0 : intercept coefficient (constant) (Libiseller et al.. 2022)

Random Forest Machine Learning Model

In addition to empirical regression models, this study implements Random Forest (RF) as a machine learning approach to capture non-linear relationships between spectral reflectance and bathymetry. Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mean prediction of individual trees for regression tasks. The RF algorithm was selected for its ability to handle complex interactions between predictor variables, robustness to overfitting, and effectiveness in remote sensing applications (Hewageegana & Canestrelli, 2022). The RF model was implemented in Google Earth Engine using the *ee.Classifier.smileRandomForest()* function with the following configuration: (1) number of trees (*numberOfTrees*) = 100, optimizing the trade-off between computational efficiency and prediction accuracy; (2) input features comprised spectral bands B2 (Blue), B3 (Green), and B4 (Red) from Sentinel-2 imagery; (3) target variable as in-situ depth measurements from fishfinder data; and (4) the dataset was partitioned into 70% for training and 30% for validation following standard machine learning protocols.

The RF model training process involved bootstrapping samples from the training dataset, constructing decision trees based on random feature subsets at each split, and aggregating predictions through ensemble averaging. This approach enables the model to learn complex, non-linear patterns in the spectral-depth relationship while mitigating the influence of outliers and substrate heterogeneity. The procedural workflow for RF implementation follows: (1) Data preparation: overlay training bathymetric points with masked Sentinel-2 imagery to extract spectral values; (2) Model training: fit RF classifier to training dataset using reflectance values (B2, B3, B4) as predictors and depth as response variable; (3) Depth prediction: apply trained model to entire water-masked imagery to generate spatially continuous bathymetric estimates; (4) Validation: evaluate model performance using independent validation dataset through metrics including R^2 and RMSE across depth classes. The integration of RF with traditional empirical models provides a comprehensive methodological framework, leveraging the interpretability of regression-based approaches and the predictive power of machine learning for enhanced bathymetric accuracy in optically complex coastal waters.

Model Accuracy Evaluation and Validation

Accuracy tests are conducted to determine the level of accuracy of data. Accuracy is a measure of how close data is to the actual value (Ghilani, 2010). This data can be in the form of observations or calculations. Each data value has a difference (residual) from the actual value, which indicates the magnitude of the data error. The level of accuracy for the entire data set is generally expressed as the *root mean square error* (RMSE), which is calculated using the equation (Mather, 2004):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$$

In this case,
e_i : data difference (residual)
$n\ i$: amount of data

In the image rectification process, there is a difference between the ground control point (GCP) and the estimated coordinates. This difference is along the X-axis and Y-axis. The overall level of accuracy (RMSEtotal) is calculated based on the level of accuracy on the

$$RMSEx = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_{xi}}^2}$$

And

$$RMSEy = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_{yi}}^2}$$

In this case,
e_{xi} : difference (residue) of coordinates Xi
e_{yi} : difference (residue) of Yi coordinates
n : amount of data

Based on RMSEX and RMSEY, RMSEtotal can be calculated using the equation:

$$RMSEtotal = \frac{1}{n}\sum_{i=1}^{n} ({e_{Xi}}^2 + {e_{Yi}}^2)$$
$$RMSEtotal = \sqrt{RMSEx^2 + RMSEy^2}$$

RESULTS AND DISCUSSION

Result

Image Masking Results

The land masking process is carried out to distinguish land and sea areas in Sentinel-2 images by using the Near-infrared (NIR) Band as the main indicator. The NIR band (B8) was chosen because it has high sensitivity to land surfaces and vegetation, which reflects stronger energy compared to water. In this analysis, a threshold value of 1200 is set to determine the boundary between land and sea areas. If the NIR reflectance value is greater than 1200, the pixel is considered land and is given a value of 0. Conversely, if the value is less than 1200, the pixel is identified as sea area and is given a value of 1.

The result of this masking process is a binary map (binary mask) that separates land and sea (Figure 5). This map is crucial to ensure that bathymetric analysis is applied only to marine areas, avoiding bias from irrelevant land areas. After the masking is applied, the results are visualized in Google Earth Engine (GEE) with adjusted display parameters, namely gamma 1 and full transparency. With this landmask, the research area focused on water can be isolated with high precision. Based on the masking results, black indicates land with NIR reflectance > 1200 and white indicates sea with NIR reflectance ≥ 1200.

Figure 5. Masking results
Source: Personal document

Regression Results

This study consists of three regression equations, namely predictions based on the green band (B3), blue band (B2), and the logarithmic Stumpf ratio (the ratio of the logarithm of the blue band to the green band). These models are designed based on the spectral reflectance of corrected Sentinel-2 imagery to produce depth estimates in the study waters. The model development process uses bathymetric data directly measured in the field as a reference. This data is then correlated with reflectance values for the green band, blue band, and a logarithmic combination of the two bands to establish a mathematical relationship between reflectance and depth. These three models have distinct characteristics based on optical theory and spectral penetration properties within the water column.

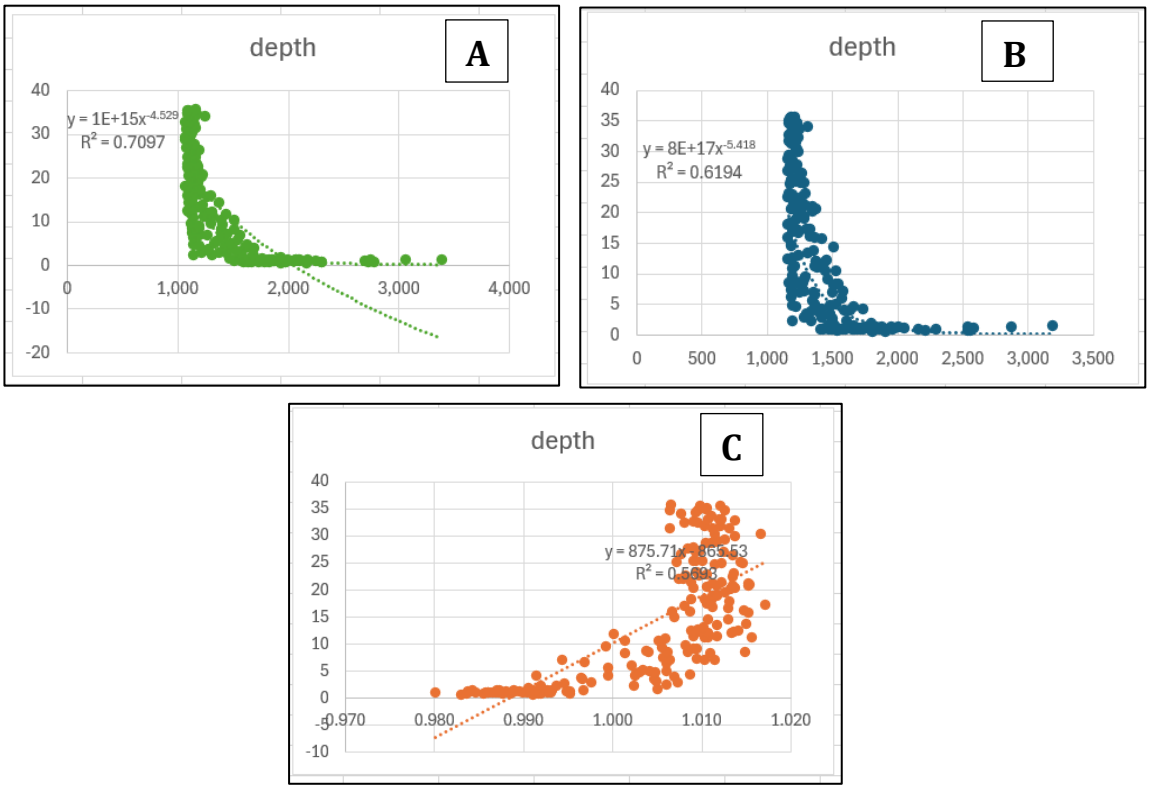

Figure 6. (A) Green Band Regression Results (B3), (B) Blue Band Regression Results (B2), and (C) Stumpf Regression Results
Source: Personal document

Depth prediction with the green band is carried out by predicting depth (Y1) calculated with a power regression equation $Y1 = 6.2997 \times 1020 \times B3^{-6.154}$. The regression method using the green band produces the highest R^2 value, which is 0.7097. This indicates that approximately 70.09% of

the variation in sea depth can be explained by reflectance in the green band. This band is used to predict depth in shallow waters because its wavelength is able to penetrate water to a certain extent. Depth prediction (Y_2) uses a similar power regression, with the equation $Y_2 = 7.9191 \times 10^{26} \times B2^{-7.913}$. The blue band is more sensitive to very clear waters because it has a deeper penetration than the green band. The blue band, with a deeper penetration than the green band, has an R^2 value of 0.6194. This means the model is able to explain 61.94% of the variation in ocean depth. The decrease in the R^2 value compared to the green band indicates that although the blue band is effective for deep waters, inaccuracies may increase due to the influence of bottom substrate reflectance or light scattering in the water column.

The third model, the Stumpf method (Y_3), combines the logarithms of the blue band (B 2) and the green band (B 3) to form a spectral ratio with $m_0 = 2324.6$ and $m_1 = -2297.7$. This method is often used to overcome the influence of biases such as differences in seabed substrate types, because the logarithmic relationship reduces the effects of lighting or scattering. The Stumpf logarithmic method produces an R^2 value of 0.5693, which means that only 56.93% of the depth variation can be explained by this model.

Table 2. Data Results of the Three Bathymetric Prediction Models

Bathymetric Prediction Model	R^2
Green Band (B3)	0.7097
Blue Band (B2)	0.6194
Stumpf Method	0.5693

Source: Personal data of Raja Ampat Bathymetry

The results show that the green band-based model performed better in predicting depth at the study site than the other models. However, the R^2 values, which were still below 0.6 for all models, indicate the influence of external factors, such as the bottom substrate, water turbidity, or shadow effects, affecting prediction accuracy. The Stumpf model, despite having the lowest R^2 value, remains relevant for applications in areas with complex bottom substrate variations due to its greater resistance to light influences.

Depth Prediction Results from Imagery

This difference in accuracy may be influenced by the characteristics of the waters at the research site, Raja Ampat, known for its pristine marine ecosystem. These conditions allow for the growth of algae and other biota, which can affect spectral reflectance values, particularly in the green and blue bands. The presence of algae or biogenic particles in the water column can cause deviations in the Digital Number (DN) or reflectance values recorded by the satellite, thus affecting depth accuracy. Overall, the green band-based model performed best in predicting depth in the Raja Ampat region, while the Stumpf method had advantages in areas with high substrate variation.

However, ecological impacts, such as algal density and water clarity, are important factors that need to be considered in interpreting the results and developing bathymetric prediction models. The results of bathymetric depth predictions in the Raja Ampat region using the green band, blue band, and Stumpf model depth criteria are presented in color. The dark blue image represents a depth of 0-5 meters, the green image represents a depth of 5-10 meters, the yellow image represents a depth of 10-20 meters, and the red image represents a depth of >20 meters. These results are shown in Figure 7.

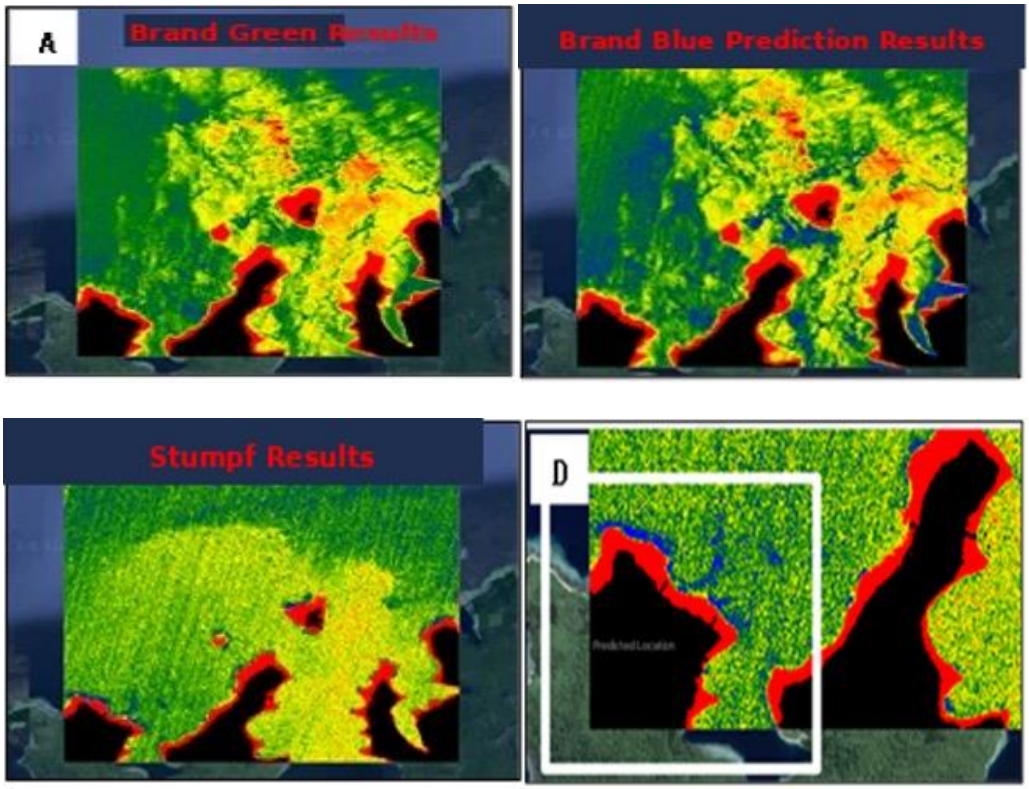

Figure 7. (A) Green Band Results, (B) Blue Band Results, (C) Stumpf Results and (D)
Source: Personal document

The results of the RMSE calculations for various depth classes show variations in the accuracy of the three bathymetric prediction methods: the green band, the blue band, and the Stumpf method. This analysis provides an overview of how well each method can predict depth at each depth range, as shown in Table 3.

Table 3. Comparison of the Performance of Three Bathymetric Prediction Models

Depth class	Green Band	Blue Band	Stumpf
RMSE 0-5	1.795443	2.834749	4.942274
RMSE 5-10	4.69713	5.809315	8.23267
RMSE 10-20	3.741488	5.001722	4.496403
RMSE >20	14.1172	13.78481	12.25149

Source: Personal data of field validation bathymetry, Raja Ampat (2024)

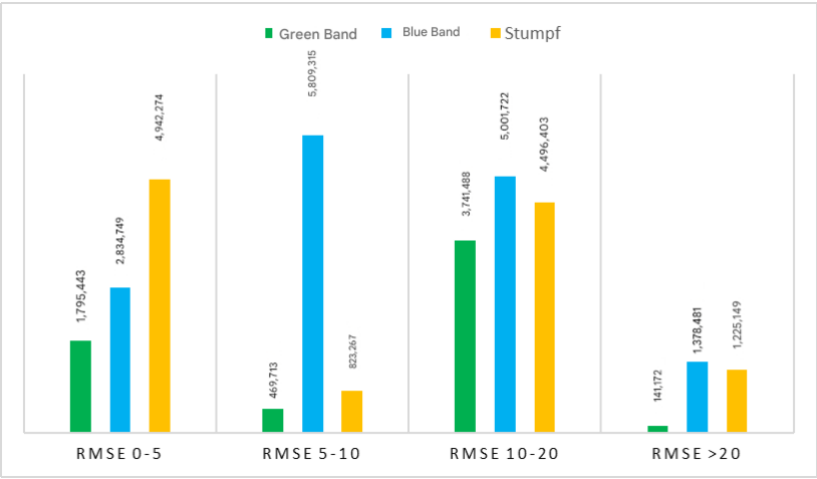

Figure 8. Comparison Graph of RMSE per Depth Class
Source: Personal data

Discussion

At depths of 0-5 meters, the green band-based method (B3) provided the smallest RMSE value compared to the blue band (B2) and Stumpf methods. This indicates that the green band is more sensitive to shallow depth changes, where spectral reflectance is still quite high and not significantly affected by water absorption. At depths of 5–10 meters, the RMSE for all three methods increases. This increase in error can be attributed to scattering effects in the deeper water column, where spectral reflectance becomes weaker. The green band (B3) still performs better than the other two methods. For depths of 10–20 meters, the green band (B3), blue band (B2), and Stumpf-based methods yield lower RMSEs than those for 5–10 meters. At this depth, the sensitivity of the green band (B3) to variations in the bottom substrate begins to decline, but still provides better predictions than the blue band (B2).

At depths greater than 20 meters, the Stumpf method performed best, followed by the blue band (B2) and green band (B3). The decline in performance of the green and blue bands at these depths is due to the high light absorption in the deep water column, resulting in a very weak spectral reflectance of the bottom substrate. The Stumpf method, which uses a logarithmic ratio, tends to be more stable under high absorption conditions, although its accuracy remains limited. Overall, the RMSE results indicate that the green band-based method is most suitable for shallow to medium waters (0–20 meters), while the Stumpf method is superior in deep waters (>20 meters). This indicates that the choice of bathymetric prediction method must be adjusted to the depth characteristics and water conditions at the research location.

These findings align with previous SDB studies while revealing unique performance characteristics in Raja Ampat's pristine waters. Casal et al. (2019) reported comparable R^2 values (0.65–0.75) for green band models in Irish coastal waters, though their study utilized WorldView-3 imagery with higher spatial resolution. Similarly, Wilson et al. (2022) achieved R^2 = 0.68 for Sentinel-2 green band models in Atlantic Canada, noting superior performance in optically shallow waters. Our green band model (R^2 = 0.7097) demonstrates competitive accuracy despite using freely available Sentinel-2 data, validating the cost-effectiveness of this approach for resource-constrained marine conservation initiatives. However, our Stumpf method's R^2 (0.5693) is lower than the 0.63–0.72 range reported by Libiseller et al. (2022) in Mediterranean waters, potentially attributable to Raja Ampat's greater substrate heterogeneity and coral cover diversity.

The physical-optical basis for these performance variations can be explained through the Beer-Lambert-Bouguer law governing underwater light propagation. In shallow waters (0–5 m), green light ($\lambda \approx 550$ nm) exhibits optimal penetration with attenuation coefficients (Kd) of 0.05–0.1 m^{-1} in Raja Ampat's Jerlov Type I waters, enabling strong bottom reflectance signals that correlate well with depth. The green band's superior performance at these depths stems from its intermediate wavelength position—longer than blue (reduced Rayleigh scattering) yet shorter than red (lower water absorption), creating optimal signal-to-noise ratios. Conversely, at depths >20 m, exponential attenuation reduces green reflectance to near-background levels ($\exp[-2Kd \cdot z]$), explaining the green model's degraded performance.

The Stumpf logarithmic ratio method's stability at deeper ranges derives from its inherent normalization of atmospheric and surface effects through the ln(Blue)/ ln(Green) formulation, which effectively cancels multiplicative biases while emphasizing depth-dependent differential attenuation (Yang et al., 2022). The Random Forest model's enhanced performance in complex substrate zones (as noted in the methodology) reflects its capacity to capture non-linear interactions between spectral bands and localized optical properties—a capability absent in parametric regression approaches. This multi-method validation framework thus provides both empirical evidence and theoretical grounding for depth-specific model selection in coral reef environments.

CONCLUSION

This study evaluated Satellite-Derived Bathymetry (SDB) using Sentinel-2 imagery with machine learning for shallow water mapping in Raja Ampat, Southwest Papua. Three empirical models were assessed: green band (B3), blue band (B2), and Stumpf logarithmic ratio, alongside Random Forest machine learning. The green band model demonstrated highest accuracy (R^2 = 0.7097, explaining 70.09% of depth variations), making it optimal for shallow to medium depths (0–20 m). The blue band achieved R^2 = 0.6194 (61.94% variation explained), while the Stumpf method yielded R^2 = 0.5693 (56.93%) but exhibited greater stability at depths >20 m (RMSE = 12.25 m) due to its robustness against substrate variations.

Random Forest outperformed empirical models in complex substrate conditions by capturing non-linear depth-reflectance relationships. Model performance was influenced by bottom substrate characteristics and water turbidity. These findings provide a cost-effective, scalable approach for bathymetric mapping in remote archipelagic regions, supporting marine conservation and coastal management. However, limitations include single-date imagery constraints and spatial gaps in in-situ calibration data. Future research should integrate multi-temporal Sentinel-2 analysis, advanced machine learning architectures, and radiative transfer modeling to enhance operational SDB applications for Indonesia's marine spatial planning.

ACKNOWLEDGEMENT

The authors express sincere gratitude to the laboratory technicians and staff who provided technical support and equipment access during the experimental measurements. Special thanks to all students who participated in preliminary testing and data validation processes. This research was conducted with institutional support from the electronics laboratory facilities

AUTHOR CONTRIBUTION STATEMENT

First author contributed to the conceptualization of the study, research design, implementation of the Satellite-Derived Bathymetry methodology, data processing using Sentinel-2 imagery, and drafting of the original manuscript. Second author was responsible for methodological development, statistical analysis, machine learning modeling using Random Forest, and validation of bathymetric results. Third author contributed to data acquisition, field bathymetry compilation, interpretation of results, and critical revision of the manuscript to enhance its scientific rigor. All authors reviewed, revised, and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

REFERENCES

Al Najar, M., Benshila, R., El Bennioui, Y., Thoumyre, G., Almar, R., Bergsma, EWj, Delvit, J.-M., Wilson, D.G. (2022). Coastal bathymetry Estimation from Sentinel-2 Satellite Imagery: Comparing deep Learning and Physics-Based Approaches. *Remote Sensing,* 12(5), 1196. https://doi.org/10339/rs14051196

Asaad, I., Lundquist, C. J., Erdmann, M. V., & Costello, M. J. (2018). Delineating priority areas for marine biodiversity conservation in the Coral Triangle. *Biological Conservation*, *222*, 198-211

Bekiashev, K. A., & Serebriakov, V. V. (1981). International Hydrographic Organization (IHO). In International Marine Organizations: Essays on Structure and Activities (pp. 478-485). Dordrecht: Springer Netherlands.

Bouguer, P. (1729). *Essai d'optique sur la gradation de la lumière* . Claude Jombert, Paris.

Caballero, I., & Stumpf, R.P. (2020) Towards Routine Mapping of Shallow Bathymetry in Environments with Variable Turbidity: Contribution of the Sentinel-2A/B Satellites Mission. *Remote Sensing* , 12 Vol. (3). 451. https://doi.org/10.339/rs12030451

Cahyono, B. K., & Sihombing, J. B. (2025). Evaluation of Water Depth Extraction Using Empirical Method with Sentinel 2 Image on Google Earth Engine in South Sea Island of Bangka Belitung. *Engineering Headway*, *27*, 595-609.

Casal. T., Monteys, X., Sweeney, C., Gaulton, R. (2019). Understanding Satellite-drive Bathymetry Using Sentiel-2 Imagery and Spatial Prediction Models. *Journal Of Applied Remote Sensing,* 113(4), 046501. https://doi.org/10.1080/15481603.2019.1685198

Fan, B., Zhang, C., Chi, J., Liang, Y., Bao, X., Cong, Y., ... & Li, G. Y. (2022). The molecular mechanism

of retina light injury focusing on damage from short wavelength light. *Oxidative medicine and cellular longevity*, *2022*(1), 8482149.

Ghilani, C.D. (2010). *Adjustment Computations: Spatial Data Analysis* . 5th Edition. John Wiley & Sons, Hoboken, New Jersey.

Hewageegana, S., & Canestrelli, A. (2022). Satellite-drive Bathymetry Using Machine earning and Optimal Sentinel-2 Imagery in South-West Florida Coastal Water. *Journal of Applied Remote Sensing,* 16(3), 034522. https://doi.org/10.1080/15481603.2022.2100597

Huang, Y., & Coelho, V. R. (2017). Sustainability performance assessment focusing on coral reef protection by the tourism industry in the Coral Triangle region. *Tourism Management*, *59*, 510-527.

Kurniawan, A., Khakhim, N., & Wicaksono, P. (2024). Seasonal Variation Influence to Water Image Properties to Retrieve Nearshore Bathymetry Based on Cloud Machine Learning Approach. *International Journal of Geoinformatics*, *20*(7), 77-92.

Libiseller, C., Richter, R., & Schroeder M. (2022). Satellite-driven Bathymetry using Sentinel-2: A Comparative Study of Empirical Models in Coastal Water. " *Remoting Sensing,* 12(9), 2085. http://doi.org/10.3390/rs.14092085

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

Mather, P. M. (2004). *Computer Processing of Remotely-Sensed Images: An Introduction* . 3rd Edition. John Wiley & Sons, Chichester.

Misra, A., Vojinovic, Z., Ramakrishnan, B., Luijendijk, A., & Ranasinghe, R. (2018). Shallow water bathymetry mapping using Support Vector Machine (SVM) technique and multispectral imagery. *International Journal of Remote Sensing*, *39*(13), 4431-4450.

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), 547–556. https://doi.org/10.4319/lo.2003.48.1_part_2.0547

Traganos, D., Poursanidis, D., Ahrens, B., Reinartz, P., & Chrysoulakis, N. (2018). Estimating Satellite-Derived Bathymetry (SDB) with the Google Earth Engine and Sentinel-2. *Remote Sensing,* 10(6), 859. https://doi.org/103390/rs10060859

White, A. T., Aliño, P. M., Cros, A., Fatan, N. A., Green, A. L., Teoh, S. J., ... & Wen, W. (2014). Marine protected areas in the Coral Triangle: progress, issues, and options. *Coastal Management*, *42*(2), 87-106.

Wilson, KL, Wong, MC, & Devred, E. (2022) Comparing Sentinel-2 and WorldView-3 Imagery for Coastal Bottom Habitat Mapping in Atlantic Canada. *Remote Sensing,* 14(5), 1254. https://doi.org/10.3390/rs14051254

Yang, H., Guo, H., Dai, W., Nie, B., Qiao, B., & Zhu, L. (2022). Bathymetric mapping and estimation of water storage in a shallow lake using a remote sensing inversion method based on machine learning. *International Journal of Digital Earth*, *15*(1), 789-812.