Assessing benthic habitat distribution in Tunda Island, Banten, Indonesia using Sentinel-2A imagery

Fakhrurrozi[1*], *Idris*[1], Mohammad *Bagus* Rizan Ardiansyah[2], Andik Dwi *Muttaqin*[2], *Mauludiyah*[2]

[1]Indonesian Coral Reef Foundation (Yayasan TERANGI), Jalan Asyibaniyah No. 105 – 106 Pondok Jaya Cipayung Depok Jawa Barat 16438
[2]Department of Marine Science, Faculty of Science and Technology State Islamic, Universitas Islam Negeri Sunan Ampel Surabaya, East Java, Indonesia

Abstract. Benthic habitats in coastal areas are composed of diverse organisms and substrates such as seaweed, seagrass, algae, hard corals, dead corals, rock, and rubble. Understanding their distribution is essential for effective marine spatial planning and conservation. This study aimed to assess benthic habitat distribution around Tunda Island using Sentinel-2A satellite imagery and evaluate the classification accuracy. Field data were collected from March 5 to 7, 2024, using the photo transect method to identify key benthic classes. Image classification was performed using a Random Forest classifier, and accuracy was evaluated through a confusion matrix. The analysis identified four main benthic habitat classes: rock, rubble, seagrass, and hard coral. Habitat distribution generally followed a pattern across all stations—rock and rubble nearshore, transitioning to seagrass and hard coral further offshore. The areal coverage for each class was estimated as rock (31.89 ha), hard coral (31.29 ha), rubble (21.34 ha), and seagrass (11.99 ha). The classification achieved an overall accuracy of 69.86%, indicating that Sentinel-2A imagery combined with Random Forest classification holds promise for mapping benthic habitats in coastal environments like Tunda Island.

1 Introduction

Benthic habitats are composed of various types of organisms, such as seaweed, seagrass, algae, hard corals, and dead corals, with diverse substrate types, such as sand, silt, and rubble [1]. These benthic habitat ecosystems have many ecological and socioeconomic roles, such as a place to find food, spawn marine biota, coastal protection from waves, stabilization of sediments, water purification, carbon sequestration, and tourism [2]. Ecosystems in shallow waters have now become very important because they describe the condition of the ecosystem of a small island area, both temporally and spatially, by utilizing satellite imagery [3].

[*]Corresponding author: fakhrurrozi@terangi.or.id

Satellite technology can provide images or data related to Earth's surface information. This satellite image can be used to determine the distribution of benthic habitats and has the advantage of being able to analyze an object on a wide scale with a more affordable level of cost effectiveness compared to other methods. The use of this satellite image depends on the requirements of the data to be processed. The Sentinel-2A image has the advantage of being easily accessible and open source, in addition to having a spatial resolution of 10×10 m²/pixel in multispectral imagery consisting of three visible (blue 0.49µm, green 0.56µm, and red 0.665µm) and near-infrared (0.842µm) channels.. In a previous study [5], visualizing benthic habitats by utilizing images with Landsat 8 resulted in an overall accuracy rate of approximately 68.29% and a kappa accuracy of approximately 35.43%, while the utilization of Sentinel-2A imagery increased accuracy with an increase in overall accuracy of 80% and an increase in kappa accuracy of 59.27%.

The influence of the water column on bottom reflectance is one of the problems in the application of benthic habitat mapping [4]; therefore, the main factor in optimizing the ability of images to produce accurate benthic habitat maps is the selection of appropriate methods and image processing [1]. Google Earth Engine (GEE) is a reliable cloud-based computing platform for fast and efficient mapping in the digital era. GEE can access various satellite image archives with different spatial and temporal resolutions, which can be used for benthic habitat mapping through satellite images [6]. Sentinel 2 is one of the satellite image archives with a spatial resolution of 10 m, which can be used for free. In research conducted by [7] discussing object-based benthic habitat classification using machine learning algorithms, the best accuracy results were obtained in the random forest algorithm (83.33%) compared to other algorithms such as Support Vector Machine (SVM) and decision tree (DT). Research by [13] stated that the classification process using Random Forest has been able to produce higher accuracy and can handle a large number of training data samples more quickly and efficiently.

Tunda Island is a small island located in the middle of the sea waters of Serang, Banten on Java Island. The intertidal area of Tunda Island has three important coastal ecosystems: mangroves, seagrass areas, and coral reefs. In a study conducted by [8], sand-mining activities caused environmental degradation and destruction. When marine sand mining becomes a current problem, Tunda Island is vulnerable to environmental changes. Sand mining in the sea also causes damage to the marine environment, especially in the coastal ecosystem of Tunda Island [9]. Therefore, the process of benthic habitat mapping by utilizing Sentinel - 2A imagery and using a classification algorithm, Random Forest classifier, needs to be performed in order to determine the level of accuracy produced. The purpose of this study was to assess benthic habitat conditions caused by sand mining activities on Tunda Island.

2 Method

2.1 Study Site

Field data collection was completed on March 5-7, 2024, on Tunda Island, Serang, Banten, Indonesia (Fig. 1).

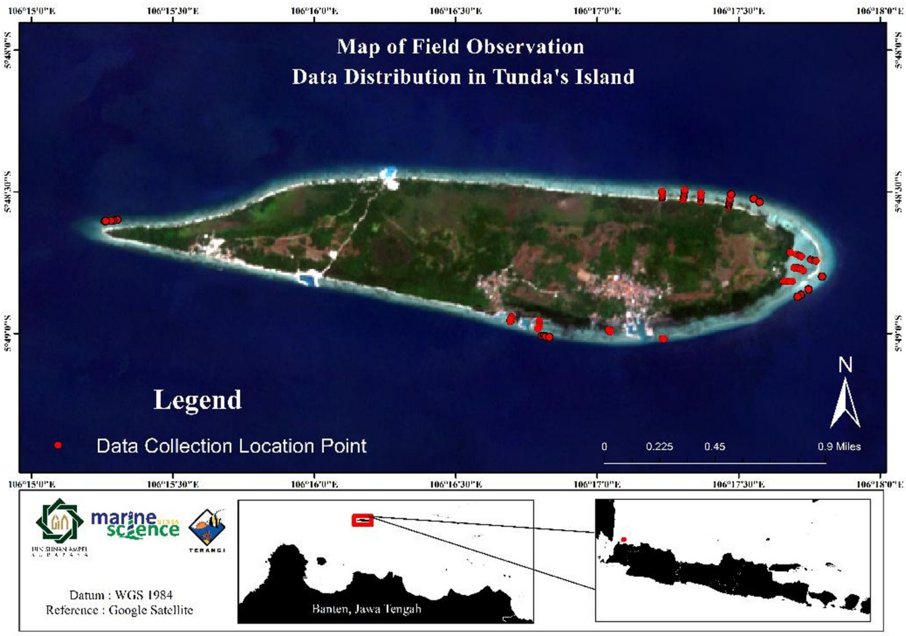

Fig. 1. The Study Area is Located on Tunda Island, Serang, Banten

The data used in this study were as follows: Sentinel - 2A image acquisition on March 10, 2023 – February 10, 2024, which can be downloaded from USGS and field data in the form of identification of the type and percentage of seagrass, coral, and open rock cover.

2.2 Data Collection

2.2.1 Transect

Field observations were conducted to obtain data and information on benthic habitats at the study site. Data collection using the georeferenced photo transect method (Fig. 2) and benthic habitat data collection using drone imagery were used, and the classification was 50 × 50 cm at every 1 m distance [10].

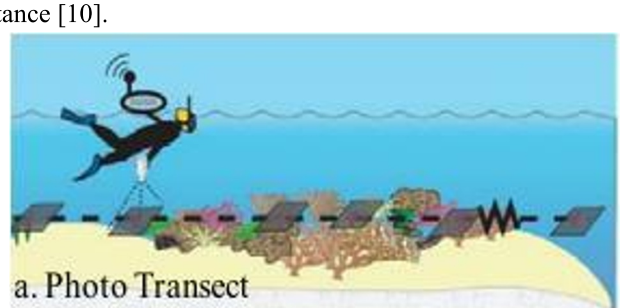

Fig. 2. The Georeferenced Photo Transect Method [10]

Field data collection activities were used to determine collection points and produce five classes in shallow waters, namely hard coral, rocks, algae, rubble, and seagrass, which can later be reclassified according to the classification results using calculations. Data collection was carried out in the southern, northern, and eastern parts of Tunda Island to obtain a representation of the benthic habitat classes. In this study, the number of samples taken by the photo transect was 250 points, which were distributed from the south, north, to the east, as shown by the red dots in Fig. 1. In addition, data from drone imagery were also used to complement areas that were not directly sampled in the field, totalling 352 points.

2.3 Analysis

2.3.1 Benthic Habitat Identification

Identification data were processed to determine the conditions of the benthic habitats on Tunda Island, Banten. The photos were sorted according to benthic habitat category and analyzed using Coral Point Count with Excel extension (CPCe) software. CPCe is a software that helps to identify and determine the conditions of benthic habitats in an area. Data processing in this CPCe program produces data in the form of an Excel file that shows the results of identification analysis [11]. The resulting data were used to determine each benthic habitat class in the classification using Sentinel 2-a imagery.

2.3.2 Data Integration

The next data-processing step is to calibrate each photo transect with the existing coordinate points. This was performed using DNRGPS software and adjusted to the local time in each transect photo file with GPS coordinates. Then, the data of each photo transect that already had coordinates were combined with the results of the calculation of percentage cover from CPCe using Microsoft Excel to produce files in csv format to be used as shp data at each observation point.

The data of each photo that has been shp were later processed with the aggregation calculation in the aggregate feature in QGIS 3.34.1. This aggregation process is used to combine several points into one by adjusting the sentinel image pixel size to 10×10 m²/pixel. The last process involves centroids, which are used to create shp points located in the center of the image resolution. This centroid process is used to determine the class by calculating the percentage of the largest class cover at each point. This resulted in four classes with 51 points each for rock, three points for rubble, 18 points for seagrass, and 15 points for Hard Coral.

2.3.3 Incorporacting Drone Imagery for Additional Data

Data from drone imagery processing were used to supplement the classification process, adding 352 points across different habitat classes. This approach was particularly useful for increasing the number of training and validation points, which were initially limited by field sampling constraints in the random forest classification. Additional drone-derived points were distributed as follows: 49, 129, 103, and 71 points for rock, rubble, seagrass, and hard coral, respectively. Additionally. Deep-sea data were obtained by manually identifying features from Sentinel-2A imagery in Google Earth Engine, resulting in 243 points. In this study, the deep sea was defined as the area where the bottom substrate was not visible. Incorporating deep-sea data allows for more comprehensive and detailed benthic habitat classification. In total, 682 data points were used for benthic habitat classification. These were divided into two datasets, 46% (317 points) for training and 54% (365 points) for validation.

2.3.4 Benthic Classification Process

The Sentinel - 2A image processing stage was run entirely on the Google Earth Engine (GEE). Atmospheric correction was applied to enhance image quality by reducing the disturbances caused by clouds or haze dispersion in the atmosphere. Image cropping was performed using a script to clip the imagery to the specified study area, thereby improving the computational efficiency of the GEE. Next, masking was applied to ensure that only pixels corresponding to water areas were retained, thereby preventing land pixels from interfering with the classification process. After masking, the training and validation data were extracted and merged for classification purposes.

For benthic habitat mapping, a red-green-blue (RGB) composite was used, incorporating bands 4 (red), 3 (green), and 2 (blue), along with band 8 (Near-Infrared/NIR) from Sentinel-2A. RGB composite imaging combines three spectral bands to enhance object visualization, whereas the addition of the NIR band improves the mapping accuracy in shallow marine environments [12]. The classification process was conducted using a Random Forest (RF) classifier, which assigns pixel values to predefined classes. Random Forest is an ensemble learning algorithm that consists of multiple decision trees, which significantly improves the classification accuracy and pattern recognition [13]. The classification system generates multiple decision trees based on random subsets of training data, making it robust against overfitting and effective for habitat classification in shallow waters.

2.3.5 Accuracy Test

An accuracy test was applied to evaluate the random forest algorithm for correctly classifying and mapping the distribution of benthic habitats. The classification accuracy derived from the remote sensing analysis was validated using field data collected during ground-truth activities. The accuracy test method used was the confusion matrix table (Table 1). The confusion matrix table was created by linking the classification results with the data obtained from field activities [14].

Table 1. Calculation Table of Accuracy Test [14]

Classification	Training Area			Total	UA (%)	CE (%)
	1	2	3			
1	X_{11}	X_{12}	X_{13}	X_{1+}	X_{11}/X_{1+}	100-PA_1
2	X_{21}	X_{22}	X_{23}	X_{2+}	X_{22}/X_{2+}	100-PA_2
3	X_{31}	X_{32}	X_{33}	X_{3+}	X_{33}/X_{3+}	100-PA_3
Total	X_{+1}	X_{+2}	X_{+3}	N		
PA (%)	X_{11}/X_{+1}	X_{22}/X_{+2}	X_{33}/X_{+3}	OA (%)	$(X_{11}+X_{22}+X_{33})/N$	
OE (%)	100-UA_1	100-UA_2	100-UA_3			

Three categories of accuracy were generated from the confusion matrix table: Producer Accuracy (PA), User Accuracy (UA), and Overall Accuracy (OA). According to the Indonesian National Standard (7716:2011) for benthic habitat mapping on a 1:20,000 scale, the minimum acceptable overall accuracy was ≥60%..

2.3.6 Flowchart

The stages of Sentinel 2A Imagery data processing for benthic habitat mapping can be seen in Fig. 3.

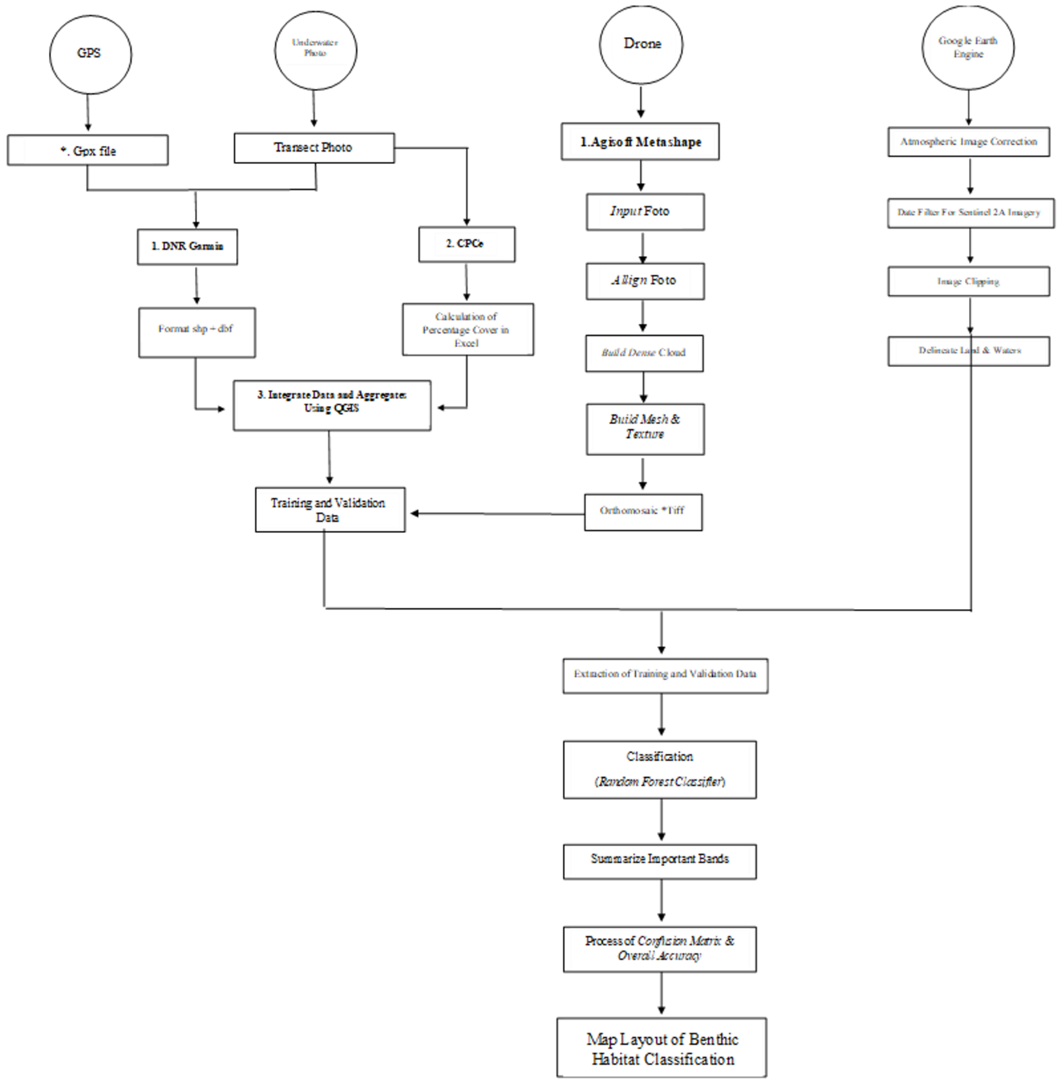

Fig. 3. Flow of Data Processing

3 Result and Discussion

3.1 Ecological Condition

Based on the identification results, it was found that the benthic components that make up the coastal ecosystem on Tunda Island include rocks, seagrass, rubble, and hard coral. Based on the results of the calculation of the percentage of cover using CPCe, the largest benthic habitat distribution in the southern part of Tunda Island is seagrass (51%), followed by hard corals (55 %) in the eastern part of Tunda Island, and rubble (72 %) in the north of Tunda Island (Fig. 4).

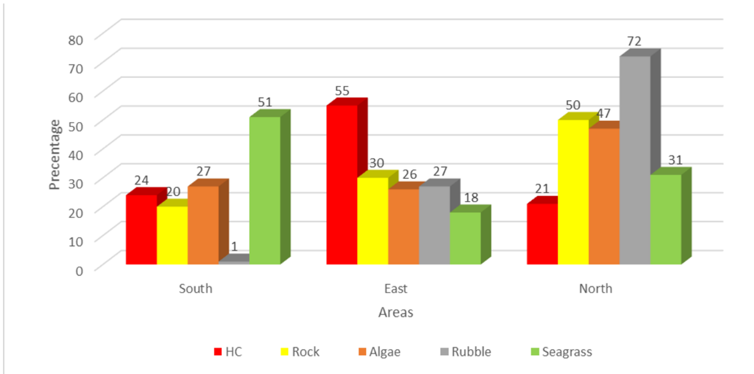

Fig. 4. Percentage Cover of Benthic Habitat on Tunda Island

Seagrass is distributed more in areas close to settlements and harbors, especially in the southern and northern parts of Tunda Island. This is due to the high anthropogenic activity in the area, which causes an increase in the amount of nutrients in the water and makes the area fertile. This causes some parts of the north and south of Tunda Island to have strata or sequences starting from seagrass. In addition, a study conducted by [18] showed that changes in the area of benthic habitat in Puerto Rico for more than 50 years resulted in a significant increase in the area of seagrass, which was previously sand and rubble. In a similar research conducted by [7] showed almost the same results on the distribution of seagrasses on Pari Island, Thousand Islands, the level of seagrass distribution was found in areas adjacent to settlements and community activities.

3.2 Benthic Habitat Classification

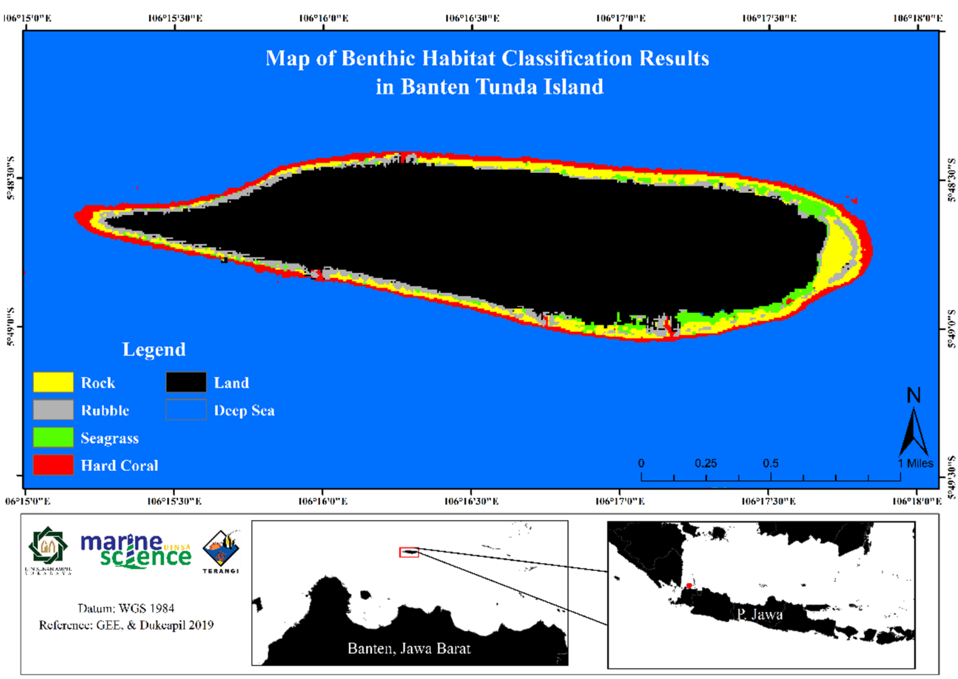

Fig. 5. Map of Benthic Habitat Classification in Banten Tunda Island

The results of benthic habitat classification are shown in Fig. 5. In this research, the area of the deep sea was not calculated because it was used only for the classification process and for easier visualization on the map. The total area of benthic habitat on Tunda Island was 96.64 ha. The largest areas are rock 31.89 Ha, hard coral 31.29 Ha, rubble 21.34 Ha, and seagrass 11.99 Ha. Based on direct field observations, the strata or sequence patterns were almost the same at each station. The composition of the benthic habitat at the north station, west station, and south station has a sequence after the shoreline, and rocks consist of sand, dead coral, silt, seagrass, rubble, and hard coral close to the reef slope. At the eastern station, the distribution of benthic habitats close to the shoreline was seagrass, and rocks, rubble, and hard corals were close to the reef slope.

In this study, the classes observed in the field, seagrass, and algae were combined into one class, seagrass, because seagrass dominates in that place and most of the algae that appear also grow near seagrass, so in capturing images, the pixels will have almost the same characteristics. Seagrass objects tend to be associated with substrate objects. Substrate objects were dominated by sand and rubble detected from a depth of 0–2 m, whereas macroalgae were found in the northern and eastern parts of Tunda Island with irregular distribution patterns and tended to be associated with objects such as sand, dead coral, and coral fragments. A comparison of field observations with classification results shows the appearance of the same objects between seagrass and algae, which is a challenge during the process of separating these objects.

The difficulty in distinguishing seagrass and macroalgae objects is not only influenced by the visual similarity in the image but also by the spatial resolution of the image, which may be insufficient to detect detailed differences between the two objects. In addition, the spectral variations of seagrass and macroalgae can overlap, especially under similar environmental conditions, such as sand or dead coral substrates. This suggests that the integration of additional data, such as multi-temporal spectral data or more detailed field data, can improve the accuracy of benthic habitat classification in the study area.

The condition of seagrass distribution on Tunda Island has decreased, as seen in the results of research conducted by [16] in 2023, where the seagrass area was 32 ha. This was caused by sand mining activities conducted on Tunda Island. The existence of excessive population and development activities without considering their impact on the environment will result in a decrease in the quality of surrounding waters. Activities that can damage marine ecosystems include excessive exploitation, fishing processes, population, and climate change [17]. Mentioned in research conducted by [9] states that the high potential of marine sand is due to Tunda Island being at the confluence of the Karimata Strait and Sunda Strait which carries sediments from the area around the strait.

3.3 Accuracy Value of Benthic Habitat Mapping

The results of the image accuracy test used in the confusion matrix included user accuracy, producer accuracy, and overall accuracy. The confusion matrix for benthic habitat distribution is presented in Table 2.

Table 2. Result of Confusion Matrix

Classification	Training Area					Total	UA (%)	CE(%)
	Rock	Rubble	Seagrass	Hard Coral	Deep Sea			
Rock	46	5	1	2	0	54	85	15
Rubble	21	69	5	2	0	97	71	29
Seagrass	16	19	44	3	0	82	54	46
Hard Coral	2	1	0	10	29	42	24	76
Deep Sea	0	0	1	3	86	90	96	4
Total	85	94	51	20	115	365		
PA (%)	54	73	86	50	75	OA (%)	69.86	
OE (%)	46	27	14	50	25			

User Accuracy is the accuracy between the class in the training area and the class used for validation in the field from the perspective of map users. The highest value of user accuracy in this research was 96% in the deep-sea class, whereas the lowest value was 24% in the hard coral class. The highest commission error is in the hard coral class at 76%, which means that more are included in other classes, including two samples in the rock class, one sample in the rubble class, and 33 samples in the deep-sea class. This is because of the similarity of the pixel values generated between the hard coral and deep-sea classes, causing errors or overlapping pixel values. The research conducted by [14] resulted in overlapping spectral patterns between seagrass classes and other classes. This could be caused by the similar pixel values between one class and another. Producer accuracy is the range of areas from the training area taken to represent a particular class and the accuracy of the map from the mapmaker's perspective. Based on the calculation, the highest value of producer accuracy in this research is 86%, in the seagrass class, while 14% of pixels that belong to the seagrass class are mapped as other classes (omission error). The lowest producer accuracy value was in the hard coral class, with a percentage of 50% and an omission error value of 50%.

The number of classes used can affect the level of accuracy of an image; the more classes used in a thematic map, the lower the level of accuracy; otherwise, if fewer classes are used, the higher the level of accuracy. This has been proven in other research [5] obtained the results of that compared the performance of Sentinel-2A images and Landsat 8 images in three classes of shallow water, where the overall accuracy of Sentinel - 2A was 80.00%, and that of Landsat 8 was 68.29%. Meanwhile, research conducted by [19] using six classes of shallow water using SPOT-7 imagery with true color composites and DII transformation resulted in Overall Accuracy values of 66.66% and 75.43%, respectively. It is also said in the research conducted by [14] that in addition to the number of classes used, the level of turbidity, light intensity, and water bottom material can also affect accuracy, especially in classes where rocks are easily lifted due to water movement.

The larger the number of classes displayed on the thematic map, the lower the accuracy. In addition, the accuracy of the mapping results can also be affected by several other factors, such as errors in determining the training area for each class of shallow water bottom habitat, errors in identifying shallow water bottom habitat objects based on the definitions used, and shifts in the location of observed objects owing to differences in position between imagery and GPS [19].

Based on the calculation results of the accuracy test in this study, the overall accuracy of the image classification results was 69.86%. The overall Accuracy is the accuracy value between the classification results and data collection points in the field. The level of accuracy of this research mapping is classified as sufficient because it has an overall accuracy value above 60%, according to SNI 7716:2011 in the Guidelines for Processing Remote Sensing Data for Coral Reef Ecosystems by LAPAN. The classification accuracy of the Sentinel-2A imagery can be used to map benthic habitats.

4 Conclusion

This study successfully mapped the benthic habitats of Tunda Island using Sentinel-2A imagery and Random Forest classification, and identified four major classes: rock, rubble, seagrass, and hard coral. The classification achieved an overall accuracy of 69.86%, meeting the SNI 7716:2011 standard. The results showed a clear zonation pattern, where seagrass was concentrated near anthropogenic activity zones, whereas hard corals dominated the eastern region. These spatial patterns suggest that coastal activities influence habitat distribution, thus emphasizing the need for further ecological monitoring and adaptive management strategies.

This research was made possible with the supported by Yayasan TERANGI and UIN Sunan Ampel Surabaya (UINSA). We extend our gratitude to the research team, field assistants, and all individuals who contributed to the data collection, analysis, and manuscript preparation. We extend our special thanks to Bumi Resources for providing funding for this activity and to the local communities of Tunda Island for their cooperation during the fieldwork. Finally, we acknowledge the valuable feedback from the reviewers, which helped improve the quality of this study.

References

[1] C. A. Sari, F. S. Achmad, P. Bayu, and S. Abdullah, *J. Penginderaan Jauh*, vol. 17, no. 1, pp. 33–42, 2020.

[2] C. D. Prawoto and Hartono, *J. Mar. Res.*, vol. 1, no. 1, pp. 1–10, 2022.

[3] E. Rahmani, I. W. Gede, A. Karang, I. D. Nyoman, and N. Putra, *J. Mar. Res. Technol.*, vol. 5, no. 1, pp. 29–39, 2022.

[4] L. O. K. Mastu, N. Bisman, and P. James, *J. Ilmu dan Teknol. Kelaut. Trop.*, vol. 10, no. 2, pp. 381–396, 2018.

[5] I. W. G. A. Karang, M. K. P. A. I Dewa, M. D. N. I Wayan, and H. I Gede, *J. Ecotrophic*, vol. 132, no. 2, pp. 227–237, 2019, [Online]. Available: https://scihub.copernicus.eu.citra

[6] M. P. Jauh *et al.*, *Maj. Geogr. Indones.*, vol. 37, no. 2, pp. 168–185, 2023, doi: 10.22146/mgi.70636.

[7] A. D. Purwanto, A. Ibrahim, A. Ulfa, E. Parwati, and A. Supriyono, *J. Kelaut. Nas.*, vol. 17, no. 2, pp. 131–146, 2022.

[8] S. Yusri, in *Yayasan Terangi*, 2018, p. 19.

[9] W. Wahyudi, E. Riani, and S. Anwar, *J. Ilmu dan Teknol. Kelaut. Trop.*, vol. 10, no. 2, pp. 277–289, 2018.

[10] C. Roelfsema, *J. Appl. Remote Sens.*, vol. 4, no. 1, p. 043527, 2010, doi: 10.1117/1.3430107.

[11] A. Sugara, A. S. Citra, A. Ari, K. Esty, W. Ully, and S. Robin, *Maj. Ilm. Globe*, vol. 24, no. 2, pp. 73–80, 2022.

[12] K. Amrillah, W. Adi, and Kurniawan, *J. Trop. Mar. Sci.*, vol. 2, no. 2, pp. 59–70, 2019.

[13] A. D. Purwanto, A. Ibrahim, A. Ulfa, E. Parwati, and A. Supriyono, *J. Kelaut. Nas.*, vol. 17, no. 2, p. 131, 2022, doi: 10.15578/jkn.v17i2.10377.

[14] I. G. Putu, B. Arri, I. D. Nyoman, N. Putra, and I. N. Giri, *J. Mar. Aquat. Sci.*, vol. 9, no. 1, pp. 18–28, 2023.

[15] C. A. Sari and A. F. Syah, *J. Sci. Technol.*, vol. 14, no. 1, pp. 114–120, 2021.

[16] F. Fakhrurrozi *et al.*, *BIO Web Conf.*, vol. 70, pp. 1–11, 2023, doi: 10.1051/bioconf/20237001009.

[17] E. Kurniawati, S. Vincentius, and N. I Wayan, *J. Ilmu dan Teknol. Kelaut. Trop.*, vol. 12, no. 2, pp. 421–436, 2020.

[18] V. P. Siregar, S. B. Agus, S. Adriani, and S. Tarlan, *J. Ilmu dan Teknol. Kelaut. Trop.*, vol. 12, no. 1, pp. 37– 51, 2020.

[19] I. D. M. K. P. Astaman, I. W. G. A. Karang, I. G. Hendrawan, and K. T. Setiawan, *J. Mar. Aquat. Sci.*, vol. 7, no. 2, p. 184, 2021, doi: 10.24843/jmas.2021.v07.i02.p07.