

## Random Forests Algorithm for Two Levels of Coral Reef Ecosystem Mapping Using PlanetScope Image in Malalayang Beach, Manado

Fela Pritian Cera<sup>1</sup>, Projo Danoedoro<sup>2</sup> , Pramaditya Wicaksono<sup>2</sup> , Moh. Yasir<sup>3</sup>

<sup>1</sup>Post-Graduated Student of Remote Sensing, Faculty of Geography, Gadjah Mada University, Indonesia

<sup>2</sup>Geographic Information Science Department, Faculty of Geography, Gadjah Mada University, Indonesia

<sup>3</sup>Makassar Coastal and Marine Resources Management Office, Directorate General of Marine Spatial Management, Ministry of Maritime Affairs and Fisheries, Indonesia

### ARTICLE INFO

#### Article History:

Received: December 30, 2021

Revision: March 09, 2023

Accepted: March 28, 2023

#### Keywords:

Coral Reef Ecosystem

PlanetScope

Random Forests

#### Corresponding Author

E-mail:

[felapritiancera@mail.ugm.ac.id](mailto:felapritiancera@mail.ugm.ac.id)

### ABSTRACT

The coral reef ecosystem has a significant physical and biological function and is also one of the coastal ecosystem components apart from the seagrass and mangrove ecosystem. Besides their ecological function, the coral reef also has an economic function. The condition of the coral reef ecosystem in Malalayang Beach has been changing for years. The utilization of remote sensing images can monitor current conditions. This research aims to map the coral reef ecosystem mapping in Malalayang Beach, Manado and conduct a test for the accuracy of coral reef ecosystem mapping using field survey data as a classification and validation sample. PlanetScope multispectral image has four channels to detect underwater objects: red, green, blue and near infrared. PlanetScope level 3B image for the research has a surface reflectance value for its pixel. The image processing stages of this research consist of sunglint correction, water column correction, and then continue to classify the coral reef ecosystem using random forests algorithm. Classification and accuracy training sample data were obtained using the photo transect technique. The sunglint correction regression equation is between 0.27 – 0.38. The coefficient of attenuation ratio in B1 is 0.927797938, B2 is 0.168841585, and B3 is 0.29033029. This value then becomes the input for the Lyzenga formula. The classification accuracy for level one using random forests is 72,54%, and the accuracy for level two mapping is 37,61%.

### INTRODUCTION

Coral reef existence as part of the coastal ecosystem is essential to ocean life sustainability. Based on satellite imagery, the area of coral reefs in Indonesia is 2.5 million hectares (Giyanto et al., 2017). Coral reefs biologically function as habitat, nursery ground and feeding ground for living organisms or marine biota (Dean and Kleine, 2012). Besides, coral reefs have an ecological function and economic function. Coral reefs benefit by providing natural services in the form of habitat for various kinds of marine organisms, contributing

significantly to capturing fisheries in Indonesia. The excellent condition of coral reefs related to the food chain system can increase the productivity of fisheries (Rusandi, 2016). The need for sunlight in macroalgae zooxanthellae photosynthetic process that is symbiotic with coral in reef formation is a limiting factor for coral reefs distribution at the bottom of waters (Miththapala, 2008).

The coral reef's condition in Malalayang Beach has changed yearly. One of the community and government efforts is to restore the coral reefs' ecosystem by doing

coral reefs transplantation. Based on Subhan et al. (2014), coral reefs transplantation is a technique to multiply coral colonies by utilizing coral asexual reproduction in fragmentation (split into fragments).

PlanetScope has temporally data from June 22, 2016, up to now with an altitude of 400 km with  $\pm 52^\circ$  ISS (International Space Station) orbit and  $\pm 81,5^\circ$  sun-synchronous orbit. PlanetScope imagery is a multispectral high spatial resolution image in 3-meter size in ISS orbit recorded. The spectral channel consists of four channels which are blue, green, red and near infrared. At one time PlanetScope satellite can record an area with a maximum size of 20 km x 12 km in the multispectral channel. PlanetScope image has been proven capable of detecting underwater object existence based on research conducted by Wicaksono and Zhafarinia (2018) that uses PlanetScope image to do benthic mapping with an overall accuracy of 71,88%.

Correction technique selection is needed to change the digital value of the satellite recording to object reflectance value. Based on Hafizt and Danoedoro (2015) research, in addition to the influence of the atmosphere and the recording angle on the sun, there is another thing that affects the pixel value error for objects at the bottom of the water, which is the attenuation of energy due to the depth increasing of the water column. The impact of that incident can be reduced by applying water column correction. The spectral reflection value of coral reefs is also influenced by several things like water depth, water quality and geomorphology in addition to pigment composition (Xu, 2014)

Remote sensing is detecting and monitoring the process of the physical characteristic in an area by measuring the reflected radiation at a certain distance to the targeted area. According to Koch and Mather (2011), remote sensing analyzes and interprets electromagnetic wave radiation

measurement that is reflected or emitted by the object under study or recorded at a certain height as an observation point or observation instrument where there is no direct contact with the object. Algorithms in image processing, namely equations or several equations, are applied to each pixel in the image to identify characteristics or estimate environmental parameters (Phinn, 2013).

Benthic habitat mapping accuracy is limited due to the low number of water penetration bands in multispectral imagers Wicaksono (2016). therefore, machine learning algorithm will be applied to obtain higher accuracy. The image classification process to obtain a map of the coral reef ecosystem is an advanced stage after the image correction process—classification in mapping coral reef ecosystems using machine learning with the Random Forest algorithm. The Random Forests classification method is an ensemble method that makes several decision trees the basis for determining results. In recent years the random forest method has been applied as a tool for general classification and pattern recognition applications (Sun et al., 2015).

Utilizing remote sensing images for underwater object research, such as coral reefs multitemporal, needs several image correction processes that aim to equalize the atmospheric conditions of each image and eliminate noise in the form of sunglint and the influence of the water column. Random forest is a scheme proposed by Leo Breiman in 2000 to create an ensemble of predictors with a single set of decision trees that can develop on selected random data (Biau, 2012).

Data collection during the field survey is intended as a training area during the classification process and to test the accuracy of the image classification results. Transects for training areas and accuracy tests were determined before the survey by visually observing variations and

distributions of underwater objects on PlanetScope images. The length of the transect line starts from the shoreline to the seashore (known as a darker water colour). Researchers/divers take data by taking photos of objects with underwater cameras that have been equalized with GPS (Global Positioning System).

The condition of the coral reef ecosystem on Malalayang Beach continues to change from year to year. Malalayang Beach has water characteristics suitable for the growth of coral reefs, so this potential can be developed. This potential was studied by [Kambey \(2013\)](#) regarding the growth of *Acropora* sp. transplanted in the waters of Malalayang Beach. Reclamation activities around Malalayang Beach caused the initial damage experienced. Reclamation activities produce sediment carried away by the currents, blocking sunlight from entering the water and causing zooxanthellae microalgae to be unable to photosynthesize. One of the community and government efforts to repair damaged coral reef ecosystems is coral reef transplantation. According to [Subhan et al. \(2014\)](#), coral transplant is a coral colony propagation technique that utilizes coral asexual reproduction by fragmentation (breaking into fragments).

Successful coral reef transplantation efforts require ongoing or long-term care. The heaviest competitor in the growth of coral reefs when transplants have been implanted in the bottom waters is macroalgae. Transplantation activities at Malalayang Beach are carried out regularly by various agencies, including the Makassar Coastal and Marine Resource Management Center, with routine and maximum maintenance resulting in a reasonably high success rate of coral reef transplants at several transplant points. Overall, transplant results covering all locations that have been transplanted can be known by conducting research using remote sensing systems.

An accuracy test is a way to test the methods used in classification. Accuracy measurement uses the confusion matrix method, also called the error matrix, a table containing comparisons of actual classes based on classification samples and accuracy samples ([Kushardono, 2017](#)). The confusion matrix method produces several indicators of accuracy and error in the classification results: overall accuracy, user accuracy and producer's accuracy. Overall accuracy or overall accuracy considers the number of correct pixels with the total number of pixels in the overall confusion matrix ([Arif et al., 2016](#)).

## RESEARCH METHODS

Malalayang Beach is part of Manado Bay which has biodiversity and is one of the favourite tourist locations of the people of Manado City. Geographically, the location of Malalayang Beach is 1°27'46.498" LU, 124°47'24.936" BT and is in the 51 N zone, as shown in Figure 1.

Tools and materials

1. PlanetScope image level 3B acquisition December 4, 2020.
2. Field data in the form of validation and accuracy samples
3. Underwater camera Gopro 7
4. GPS is used to take coordinates Garmin 64s
5. Laptop with ENVI, EnMAP Box and ArcGIS software
6. Basic diving equipment

Data Processing Procedure

This research requires materials such as PlanetScope level 3B images acquired on December 4, 2020, field data as a classification sample and accuracy samples obtained from field surveys. The sequence in the data processing is sunglint correction, water column correction, field survey, guided classification, and accuracy test. The research flow chart is shown in Figure 2. [Adji et al. \(2016\)](#) studied coral reef cover on

Pombo Island using the underwater transect photo method as image classification sample data. This method requires a GPS linked to a snorkeller who takes photos along a predetermined transect. Sunlight correction was made by using the Hedley algorithm, which is  $L_i(VIS)' = L_i(VIS) - b_i[L(NIR) - L_{min}(NIR)]$ . Then the water column correction was made by using which is

$$Y = \ln(L_i) - [(k_{ikj}) \times \ln(L_j)].$$

The image mapping process using the random forest algorithm using the EnMAP Box software goes through several stages. The final step was the accuracy test which is overall accuracy using the total number of correctly classified pixels divided by the total number of reference pixels and the producer accuracy using the number of pixels that are correctly classified divided by the total number of reference pixels based on [Danoedoro \(2012\)](#).

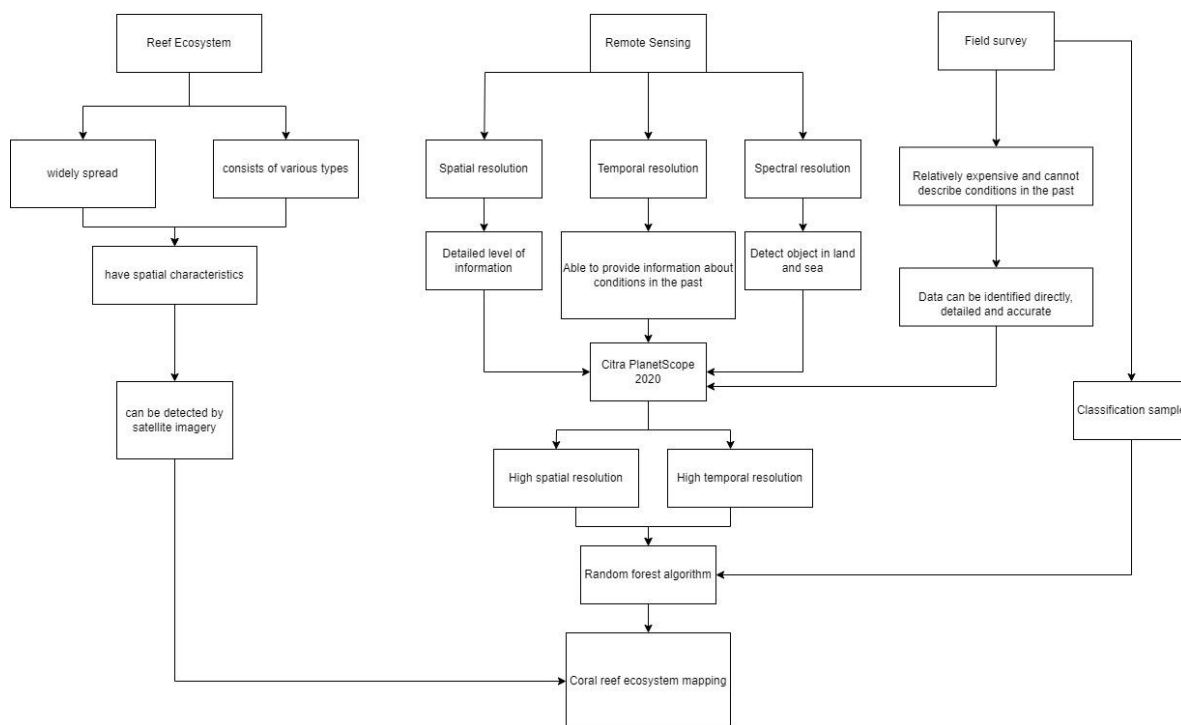


Figure 1. Research Flowchart

## RESULTS AND DISCUSSION

The correction and classification process were applied to the PlanetScope image acquired on December 4, 2020. PlanetScope imagery has been operating since 2017 with a spatial resolution of 3 meters for the multispectral channel [\(Planet Labs, 2018\)](#). The spectral channel consists of four channels: blue, green, red and near infrared [\(Collison and Wilson 2017\)](#). The PlanetScope level 3B image is an orthorectified image at the surface reflectance level, so it does not require

atmospheric correction [\(Wicaksono and Lazuardi 2019\)](#). PlanetScope image processing begins with changing the Surface Reflectance (SR) value, which was previously an integer, to a float. The purpose of this process is to ease the following image processing process because the value of each pixel becomes smaller. The pre-processing stage aims to minimize distortion or errors in the image that interfere with the classification stage and extract or sharpen the essential information to make the classification process easier [\(Khorram, 2012\)](#).

## Sunglint correction

Table 1. Sunglint correction regression equation on 2020 image

Band	Formula	R <sup>2</sup>
B1	$(B1-(0.5805*(B4-0.0191))$	0.3859
B2	$(B2-(0.5148*(B4-0.0191))$	0.2723
B3	$(B3-(0.5021*(B4-0.0191))$	0.3728

The presence of sunglint in images with RGB composites can be recognized on objects that form long lines and are brightly coloured (white) in deep waters. Sunglint in the image is caused when recording the position of the water surface in such a way as to the sun, and then the reflection of sunlight is directly received by the sensor (Kay et al., 2009). The sunglint correction utilizes the regression equation formula's smallest reflection value in the near-infrared channel. The sunglint correction result is shown in Figure 2 below.

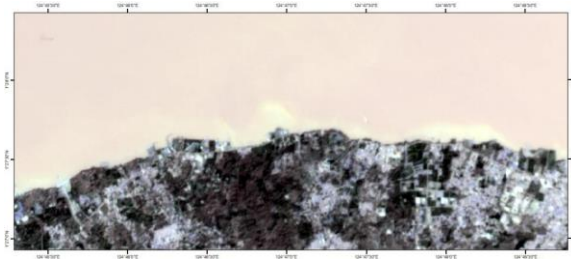


Figure 2. The result after sunglint correction

## Water column correction

The purpose of the water column correction is to reduce the effect of the water column to get the value of the spectral reflection of objects below the water surface. The Lyzenga algorithm requires the attenuation coefficient ratio value obtained from the ROI pixel value of the sample from each channel. ROI sampling in the form of sand objects at various depths. The sand object was chosen because it is easily recognizable in the image at the surface reflectance level, the composite appearance of an actual colour image based on its white or light colour and smooth texture. This method produces a bottom index unaffected by the depth and successfully applies to clear waters such as coral reef ecosystems (Jaelani et al., 2015).

Table 2. The visible channel variance and covariance values in the image

Kanal	Varian	Kanal	Covar	Koefisien pelemahan	Rasio koefisien pelemahan		
B1	0.026476	B12	0.023963	a12	-0.075011476	k1/k2	0.927797938
B2	0.030071	B13	0.034303	a13	-2.876934962	k1/k3	0.168841585
B3	0.223851	B23	0.061439	a23	-1.577011345	k2/k3	0.29033029

A sampling of the sand ROI in the 2020 image, as much as 70, was carried out using surface reflectance images at different depths. Regression analysis of the sand ROI sample shows that the determinant coefficient values range from 0.1986 to 0.7213.

The determinant coefficient between B1 (green) and B3 (red) has the lowest value,

indicating that the correlation of the spectral values of the sunglint-corrected image in the two bands is very low. The coefficient of attenuation ratio in B1 is 0.927797938, B2 is 0.168841585, and B3 is 0.29033029. This value becomes the input for the Lyzenga formula listed in Table 3. And the result after water column correction was applied to the image is shown in Figure 3.

Table 3. Equation of water depth invariant index for image

Saluran	Persamaan
B1 (k1/k2)	$\text{alog}(B1) - (0.927797938 * (\text{alog}(B2)))$
B2 (k1/k3)	$\text{alog}(B1) - (0.168841585 * (\text{alog}(B3)))$
B3 (k2/k3)	$\text{alog}(B2) - (0.29033029 * (\text{alog}(B3)))$

Source: Data Processing (2022).

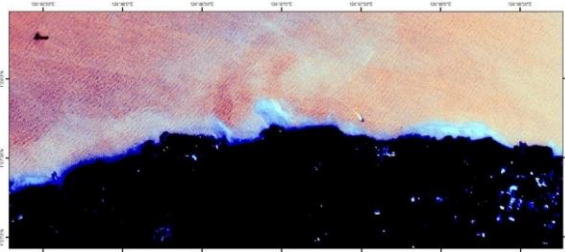


Figure 3. Planetscope image after water column correction

### Coral Reef Ecosystem Map Level One

Data on remote sensing images with the high spatial resolution has a more extensive data volume than low spatial resolution images, so an algorithm is needed to store and recognize the characteristics of images with large pixel dimensions. The machine learning approach is often used to analyze more diverse data, including the random forest algorithm. Based on previous research, the random forest algorithm produces high mapping accuracy. The random forest algorithm has several important characteristics, such as diversity, where not all attributes and features are considered when making a tree, so there is less feature space. Each tree is created

independently based on different data attributes and features. Random forests are more stable because they consider the average values that appear most frequently in determining classes.

Level one is a classification of 3 classes consisting of coral, rubble, and sand. The coral class consists of various coral reef life forms and DCA, mapped to the second-level classification. The corrected image of the water column is input to the classification using the random forests algorithm. The classification sample used is 316 data from a field survey conducted on December 10, 2020, in the study area. Training sample data in photos and coordinates are converted into shapefiles. The Random Forests algorithm based on Dewi et al. (2011) shows an error in the classification of random forests suspected of using the OOB (Out-of-Bag) error. The classification process uses EnMAP-Box software with a total of 100 tree scenarios. Non-parametric machine learning algorithm currently is adapted and shows better accuracy based on [Wicaksono \(2019\)](#).

Table 4. Area of coral reef ecosystem composition level one in 2020

No.	Composition	Area (Ha)
1	Coral	58,30
2	Rubble	1,47
3	Sand	22,67
Total		82,51

Random forests are a combination of tree predictors where each tree depends on the values in a random vector sampled independently. Each tree has the same distribution for all trees in the forest. Based

on the classification results, the coral class area on Malalayang Beach is 58.30 Ha, the rubble class is 1.47 Ha, and the sand class is 22.67. The total study area is 82.51 Ha.

The accuracy value for level one mapping is 72.54%, with a kappa value of 18.16 for the 2020 DII level image. The overall accuracy value shows the comparison of the correct number of samples and the total number of samples. The kappa value of 12.49% from the level

one classification with a composition of 3 classes could avoid 12.49% of errors in the random classification. The highest user accuracy value in the coral class indicates that 94.96% of objects in the form of corals classified in the image are coral objects in actual conditions.

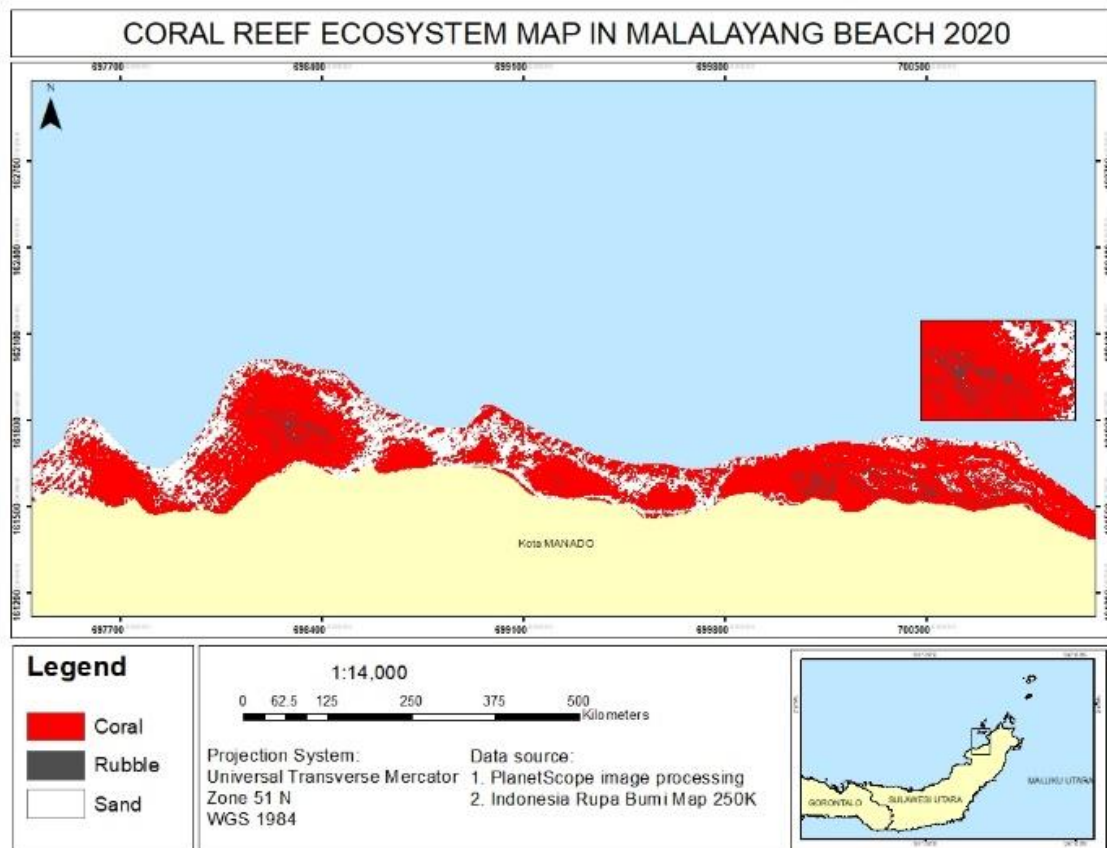


Figure 4. Map of coral reef ecosystem level one

Tabel 5. Confusion matrix for accuracy test of level one classification

Composition	Coral	Rubble	Sand	Total	User accuracy (%)
Coral	226	8	4	238	94,96
Rubble	23	0	0	23	0,00
Sand	57	0	17	74	22,97
Total	306	8	21	335	
Producer accuracy	73,86	0,00	7,43		OA 72.54% Kappa 18,16

Source: Data Processing (2022).

### Coral Reef Ecosystem Map Level Two

Classification at level two details coral classes divided into DCA, foliose, massive and mix coral classes. The input data for

training samples consisted of 201 samples of DCA class, seven samples of foliose class, 234 samples of massive and mixed coral

class, 65 samples of rubble class and 61 samples of sand class. Class determination on field data uses CPCe software which is then analyzed by ROI index separability using ENVI software. The coral reef ecosystem classification map is shown in Figure 3. The classification results show that

the DCA class has an area of 24.96 Ha, the foliose class is 1.84 Ha, and the most dominating class is the massive and mixed coral class of 30.73 Ha. Most foliose objects were found in the northeastern part of the study area.

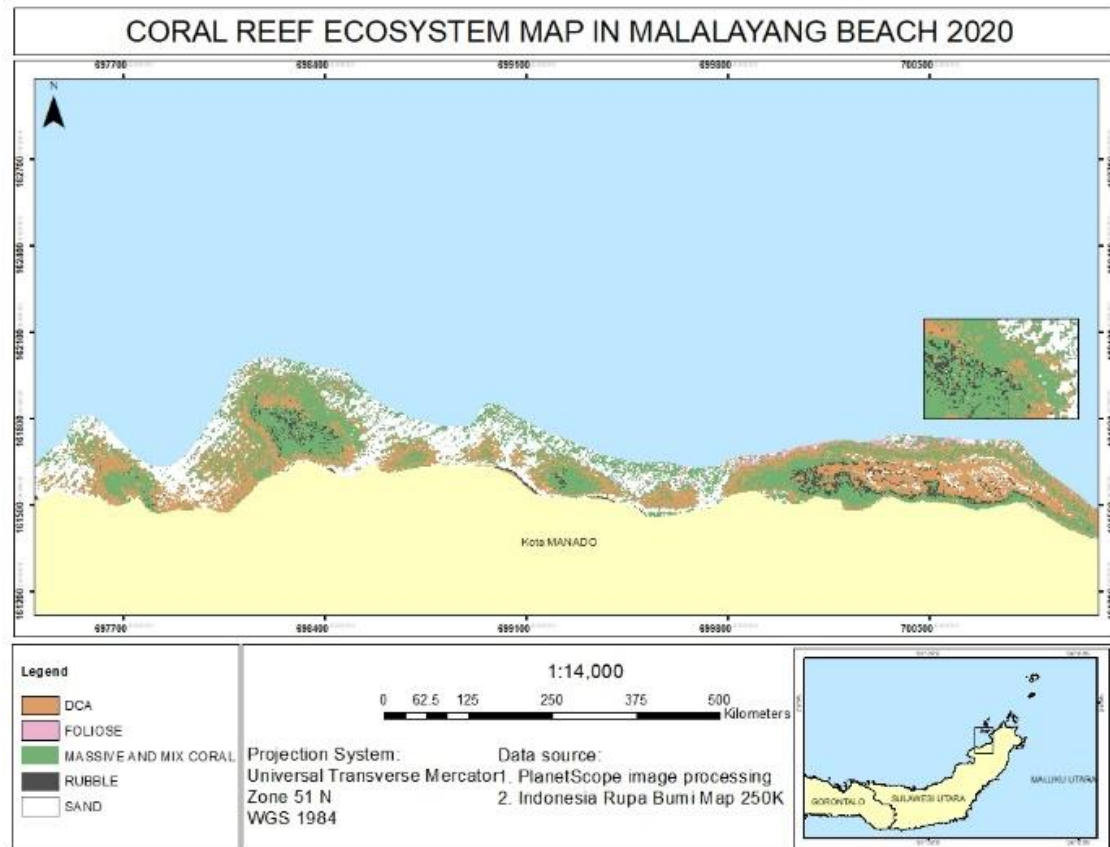


Figure 5. Map of coral reef ecosystem level two

Table 6. Area of coral reef ecosystem composition level two in 2020

No.	Composition	Area (Ha)
1	DCA	24,96
2	Foliose	1,84
3	Massive and Mix Coral	30,73
4	Rubble	2,65
5	Sand	22,26
Total		82,51

Source: Data Processing (2022).



Table 7. Confusion matrix for accuracy test of level two classification

Composition	DCA	Foliose	Massive and Mix Coral	Rubble	Sand	Total	User accuracy (%)
DCA	23	2	61	1	1	88	26,14
Foliose	0	0	0	0	1	1	0,00
Massive and Mix Coral	30	5	96	7	3	141	68,09
Rubble	4	0	22	0	0	26	0,00
Sand	8	3	52	0	16	79	20,25
Total	65	10	231	8	21	335	
Producer accuracy	35,38	0,00	1,90	0,00	8,33		OA 40.30% Kappa 7.02%

Source: Data Processing (2022).

The accuracy test on the second-level classification resulted in an overall accuracy value of 40.30% with a kappa value of 7.02%. Overall the class that is classified correctly on the image is 40.30%. The highest UA value is the massive and mixed coral class of 68.09% which shows that 68.09% of the pixels classified as massive and mixed coral are the object in their actual condition.

## CONCLUSION

This research resulted in two different levels of mapping classification using the random forests algorithm. Level one consists of three classification classes with an area of 58.30 Ha for the coral class, 1.47 Ha for the rubble class and 22.67 Ha for the sand class. The result of level two classification consists of five classes which are 24.96 Ha for the DCA class, 1.84 Ha for the foliose class, 30.73 Ha for the massive and mixed coral class, 2.65 Ha for the rubble class and 22.26 Ha for sand class. The accuracy test results at level one classification are 72.54%, and level two classification is 40.30%.

## REFERENCE LIST

Adji, A. S., Indrabudi, T., dan Alik, R. Penerapan Metode Foto Transek Bawah Air untuk Mengetahui Tutupan Terumbu Karang di Pulau Pombo, Maluku. *Jurnal Ilmu dan Teknologi Kelautan Tropis*. 8(2):633-643

- Arie, N. Wahyuni, F. S. 2016. Penggunaan Metode Machine Learning Untuk Pengenalan Pola Tutupan Lahan Pada Citra Satelit. *Seminar Nasional Teknologi Informasi dan Multimedia 2016*
- Biau, G. Analysis of a Random Forests Model. *Journal of Machine Learning Research*. 13: 1063 - 1095
- Collison, A., and N. Wilson. 2017. *Planet Surface Reflectance Product*. San Francisco, CA: Planet Labs.
- Danoedoro, P. 2012. *Pengantar Penginderaan Jauh Digital*. Yogyakarta: CV. Andi Offset
- Dean, A. Dan Kleine, D. 2012. *Coral Reefs and Climate Change*. The University of Queensland: Queensland
- Dewi, N. K., Syafitri, U. D., Mulyadi, S. Y. 2011. Penerapan Metode Random Forest dalam Driver Analysis. *Forum Statistika dan Komputasi*. 16(1):35-43
- Giyanto, Abrar, M., Hadi, T. A., Budiyo, A., Hafizt, M., Salatalohy, A., & Iswari, M. Y. (2017). Status Terumbu Karang Indonesia 2017. *Giyanto Muhammad Abrar Tri Aryono Hadi Agus Budiyo Muhammad Hafizt Abdullah Salatalohy Marindah Yulia Iswari COREMAP-CTI Pusat Penelitian Oseanografi - LIPI (Issue June)*.
- Hafizt, M. dan Danoedoro, P. 2015. Kajian Pengaruh Koreksi Kolom Air pada Citra Multispektral Worldview-2 untuk Pemetaan Habitat Bentik di Pulau Kemujan Kepulauan Karimunjawa

- Kabupaten Jepara. Prosiding Pertemuan Ilmiah Tahunan XX 2015. 566-575
- Jaelani, L., Laili, N., Marini, Y. (2015). Pengaruh Algoritma Lyzenga Dalam Pemetaan Terumbu Karang Menggunakan WorldView-2. *Jurnal Penginderaan Jauh* Vol. 12 No. 2.
- Kay, S., Hedley, J. D., & Lavender, S. (2009). Sun glint correction of high and low spatial resolution images of aquatic scenes: A review of methods for visible and near-infrared wavelengths. *Remote Sensing*, 1(4), 697-730. <https://doi.org/10.3390/rs1040697>
- Khorram, S. Koch, F H. Wiele, C F. Nelson, S A C. 2012. *Remote Sensing*. Springer
- Koch, M and Mather, M., 2011. *Computer Processing of Remotely Sensed Images An Introduction*. Fourth Edition. Wiley Blackwell.
- Kushardono, D. 2017. *Klasifikasi Digital Pada Penginderaan Jauh*. Bogor: IPB Press
- Miththapala, S. 2008. *Coral Reefs*. Coral Ecosystem Series (Vol 1) pp 1-36 + iii. Colombo, Sri Lanka: Ecosystem and Livelihoods Group Asia, IUCN
- Planet Labs. 2018. *Planet Imagery Product Specifications*.
- Phinn, S. R. Hochberg, E. M., dan Roelfsema, C. M. 2013. *Visible and Infrared Overview*. *Coral Reef Remote Sensing*. Springer Chapter 1: 3-25
- Rani, C. 2003. Perikanan dan Terumbu Karang yang Rusak: Bagaimana Mengelolanya?. *Jurnal Bionatura*. 5(2): 97-111
- Rusandi, A. 2016. *Rencana Aksi Nasiolan Konservasi Karang Periode I: 2017-2021*. Direktorat Konservasi dan Keanekaragaman Hayati Laut. Kementerian Kelautan dan Perikanan: Jakarta.
- Subhan, B. Madduppa, H. Arafat, D. dan Soedharma, D. 2015. Bisakah Transplantasi Karang Perbaiki Ekosistem Terumbu Karang? *Risalah Kebijakan Pertanian Dan Lingkungan Rumusan Kajian Strategis Bidang Pertanian Dan Lingkungan*, 1(3), 159-164.
- Sun, J. Zhong, G. Dong, J. dan Cai, Y. *Banzhaf Random Forests*. Lecture Notes in Computer Science: Authors Instructions.
- Wicaksono, P. (2016) Improving the accuracy of Multispectral-based benthic habitats mapping using image rotations: the application of Principle Component Analysis and Independent Component Analysis, *European Journal of Remote Sensing*, 49:1, 433-463, DOI: [10.5721/EuJRS20164924](https://doi.org/10.5721/EuJRS20164924).
- Wicaksono, P. Lazuardi, W. 2019. *Random Forest Classification Scenarios For Benthic Habitat Mapping Using PlanetScope Image*. *IEEE International Geoscience and Remote Sensing Symposium* 8245-8248
- Wicaksono, P. Aryaguna, P. Lazuardi, W. 2019. *Benthic Habitat Mapping Model and Cross Validation Using Machine Learning Classification Algorithm*. *Remote Sensing*. 11. [doi:10.3390/rs11111279](https://doi.org/10.3390/rs11111279)
- Xu, J. Dan Zhao, D. 2014. *Review of coral reef ecosystem remote sensing*. *National Marine Environmental Monitoring Center. Acta Ecologica Sinica* 34: 19-25
- Zhafarina, Z. Wicaksono, P. *Benthic habitat mapping on different coral reefs types using random forest and support vector machine algorithm*. *Proc. SPIE* 11372, Sixth International Symposium on LAPAN-IPB Satellite