

ACCURACY OF UNSUPERVISED CLASSIFICATION TO DETERMINE CORAL HEALTH USING SPOT-6 AND SENTINEL-2A

Nurjannah Nurdin^{1,3}, Supriadi¹, Mahatma Lanuru¹, M. Akbar AS⁴, Indah Kartika¹, Terushisa Komatsu^{2,3}

¹ Marine Science Department, Hasanuddin University, Makassar, 90245. Indonesia. nurj_din@yahoo.com

² Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa 277-8564, Japan

³ Present affiliation: Yokohama College of Commerce, Yokohama 230-8577, Japan

⁴ Research and Development Center for Marine, Coast, and Small Island, Hasanuddin University, Jl.Perintis Kemerdekaan km.10, Makassar, 90245. Indonesia. nurj_din@yahoo.com

KEY WORDS: Pulau Barrang Caddi, Spectral, IsoData, K-Means, Spermonde Archipelago

ABSTRACT:

Characteristics of corals spectral from different species are expected to have optically different characters. The aims of this research are to compare unsupervised classification between IsoData and K-Means methods with Lyzenga application, and to analyze the precision of SPOT-6 and Sentinel-2A satellite imagery in classifying shallow water habitat. The image processing are atmospheric correction, cropping, masking, Depth Invariant Index, Unsupervised classification, ground truthing, reclassify, accuracy assessment, and shallow water habitat spectral reflectance analysis. Rubble and dead coral with algae were indicating as coral death due to either damaging human activity or natural death such as bleaching. The accuracy of unsupervised classification IsoData and K-Means method have the same accuracy 62.50%. The IsoData method is better detected live coral and algae. Rubble were dominant detected in K-Means method.

1. INTRODUCTION

Coral reefs in Indonesia covers a total of 39,583 km² area, which are 18% of the world's coral reefs cover and 45.7% of the coral reefs triangle area extend (Giyanto et al., 2014). Indonesia's coral reefs area has around 590 species of coral reefs and 2,200 species of reef fish, which is the peak of coral species diversity in the coral triangle area. Spatial information about the composition, conditions and dynamics of coral reefs at the right spatial scale is a fundamental prerequisite in understanding and managing shallow water ecosystems (Phinn et al. 2011). A new approach has been available to obtain spatial information on the area of coral reefs ecosystem by using remote sensing technology. Remote sensing method can be done repeatedly and is considered able to increase the speed of observation on a relatively wide area, so that it becomes an important complementary tool for conventional methods (Mumby et al. 2004).

Mapping of shallow water habitats is a basis to understand and limit anthropogenic pressure on the marine environment, it has become a significant information in supporting environmental management and its application (Henriques et al., 2014). Total extent of shallow water areas in the Spermonde Islands were calculated using mapping techniques in order for the monitoring to be done effectively. This monitoring technique is carried out spatially by utilizing remote sensing data and geographic information system (GIS). This approach has been used in several studies, including mapping of coral habitats in New Caledonia with conservation approaches (Deas et al., 2014), studies related to habitat changes in coral reefs ecosystems in the Red Sea, Hurgada, Egypt (El-Askary et al., 2014) and on Bunaken Island in North Sulawesi, where for the first time detailed classification of > 150 habitat types using high resolution (2m) image data, Geoeye-1, were carried out in

Indonesia (Ampou et al., 2017), All of which had provide good results to support management of coral reefs area. Barrang Caddi Island has a variety of coral reefs cover and other seabed objects which are assumed to have different spectral reflectance values. Therefore, the study aims to analyze the spectral reflectance values of various coral covers and other seabed objects using SPOT-6 satellite imagery, Sentinel-2A and Landsat-8.

Characteristics of spectral reflectance of healthy corals from different species are expected to have optically different spectral characters. This spectral difference is based on distinction of the color produced due to differentiation in density and presence or absence of pigments in corals which results in higher reflectance, and also due to the loss of some spectral characteristics (Nurjannah, 2006). The aims of this research are to compare unsupervised classification between IsoData and K-Means methods with Lyzenga application, and to analyze the precision of SPOT-6 and Sentinel-2A satellite imagery in classifying shallow water habitat.

2. MATERIAL AND METHODS

2.1 Study area

Field survey was conducted in Barrang Caddi island as a part of Spermonde archipelago (Figure 1). This research was conducted in November 2018 until April 2019, comprising literature studies, imagery analysis, field data collection, processing, and research data analysis. Retrieval of field data was carried out at the end of February 2019 in Barrang Caddi Island, Spermonde Islands, Makassar Strait, South Sulawesi (Figure 1).

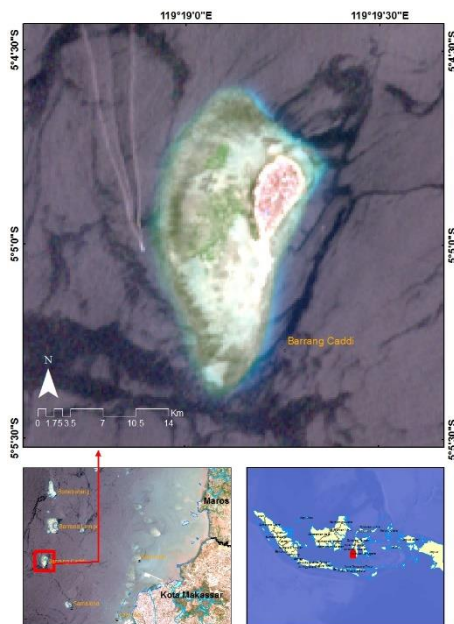


Figure 1. Research location, Badi Island, Spermonde Islands, Indonesia.

2.2 Data Collection

SPOT-6 is an optical imaging satellite capable of imaging the Earth with a resolution of 1.5 meter panchromatic and 6 meter multispectral (blue, green, red, near-IR), Sentinel-2A with 10 m spatial resolution, and Landsat-8 with 15 m spatial resolution panchromatic, and in situ data. The Sentinel-2A imagery was an acquisition from 17 October 2018, and Spot-6 from 15 May 2018. Image data can be obtained from: (<http://earthexplorer.usgs.gov/>)

2.3 Methods

2.3.1 Pre processing image

Image pre-processing stage includes atmospheric correction, image cropping, and image masking, where the atmospheric correction process is carried out to eliminate atmospheric influences / disturbances on images captured by satellites. After that, the process of cropping (cutting) was done to cut certain area of the image until it only displays the focus area of the study. Later on, the masking process was done, where the boundary area between land and water is made, so the digital value of the land becomes 0 (zero) and only shows the digital value of the water area.

2.3.2 Depth invariant index

The application of the Lyzenga algorithm is done by combining two channels of visible light, namely the blue channel and the green channel which is channel 2 with a wavelength of 490 nm and channel 3 with a wavelength of 560 nm, respectively. Two channels of visible light at the wavelength are obtained as follows:

$$K_i/K_j = a + \sqrt{a^2 + 1} \quad (1)$$

Where:

$$a = \frac{\sigma_{ii} - \sigma_{jj}}{2\sigma_{ij}} \quad (2)$$

$$\sigma_{ij} = \overline{X_i X_j} - \overline{X_i} \overline{X_j} \quad (3)$$

K_i/K_j , attenuation coefficient, σ_{ij} is the variance measurements of X_i , σ_{jj} is the variance measurements of X_j and σ_{ij} is X_i and X_j covariance. The final step is to calculate the depth invariant index (DII) with the equation:

$$DII = \log(x_i) + [(K_i/K_j) * \log(x_j)] \quad (4)$$

Where :

Y: Extraction of basic information

i : Blue channel value

j : Green channel value

ki/kj : attenuation coefficient

2.3.3 Unsupervised Classification

Unsupervised classification is a classification process in which the process is automated. Its work principle is to classifies its classes based on the spectral value of an object which has similarities to a particular class or cluster. (Jaya, 2007; 2010). The unsupervised classification method used in this study was IsoData and K-Means.

Isodata classification is one of the unsupervised classification processes where the principle is to classify pixel values based on mean values into certain groups / clusters. Isodata algorithm works by combining the same cluster and data in a cluster so that they can divide based on maximum standard deviations (Wang, 2016). Here are the parameters used in the Isodata Algorithm.

K_0 = Number of initial clusters

n_{min} = Minimum cluster data

σ_{2max} = Maximum variant (merge)

d_{min} = minimum (split) pair distance

K-Means classification is one of the unsupervised classification processes where the principle is to classify pixel values based on K values which are the desired number of class constants (clusters). Calculate the cluster center using the equation below:

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (5)$$

Where:

C_{kj} = Center the cluster to k on feature j

X_{ij} = Feature j in data i

P = Amount of data

Calculate the distance of each data to the cluster center with the equation below:

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - C_{kj})(x_{ij} - C_{kj})^2} \quad (6)$$

Where:

d_{ik} = Distance of data I to center of cluster k

m = Number of features

X_{ij} = Feature on data i

C_{kj} = Feature on the cluster center k

2.3.4 Ground truthing

Determination of ground truthing points is based on the result of unsupervised classification analysis on the image. Coordinates were matched using a GPS (Global Positioning System) receiver in the field to collect data of shallow water habitat on Barrang Caddi Island. Survey points position acquisition with GPS is expected to be able to give accurate measurement.

Data collection of shallow water habitat was carried out using the RRA (Rapid Reef Assessment) method to determine the location/ representative point. Identifications were done in the area of 10 x 10 meters with 4 pixels repetition for each category of shallow water habitat, based on the results of unsupervised classification.

2.3.5 Reclassify

Image reclassification is done by overlaying 10 classes classification results with ground truthing points data which contained information of shallow water habitat. The same shallow water habitat will be combined together into one class, so that the number of classes formed after the reclassification process become smaller in order to facilitate the accuracy test.

2.3.6 Accuracy assessment

Accuracy tests were carried out using the confusion matrix. The matrix compared habitat classes in the image classification result to the actual classes of field survey result. Column data is the result of classifications that represent the calculation of producer accuracy (PA) while row data is the result of observations by observers and were used in calculating user accuracy (UA). The more the classification results are aligned with the results of observation, the overall accuracy will be higher (Agus, 2012).

The value of accuracy is expected to meet the requirements greater than 60% (Green et al. 2000), so that the results of the percentage of OA values obtained can be used as proof of the accuracy of image classification. The calculation of each accuracy is done by the equation as follows (Congalton and

Green 2009).

2.3.7 Extraction of Spectral Reflectance Values of Shallow water habitat

Analysis of the spectral reflectance pattern of shallow water habitat was carried out to find out how are the spectral reflectance patterns of some shallow water habitat based on the results of the unsupervised classification, which are the IsoData and K-Means method with the application of the lyzenga algorithm. The average spectral reflectance of each shallow water habitat from the unsupervised classification result, IsoData and K-Means, were tabulated based on the wavelength used, and were analyzed using the line graph feature in a data processing application.

3. RESULTS AND DISCUSSION

Barrang Caddi Island is one of the small islands in the inner zone of the Spermonde Islands, South Sulawesi. Administratively, Barrang Caddi Island is included in the Makassar City area, Ujung Tanah District. The island has an area of ± 4 Ha with 2 RWs, each of which has 4 RTs. Barrang Caddi Island's population reaches 1,532 people, including 333 families, consisting of 812 women and 720 men. The morphology of the Barrang Caddi reef is fringing reef that follows the coastline. In the eastern part of Barrang Caddi Island is an area that has the lowest live coral cover and is generally covered by a stretch of sand, while the south side is an area that has a fairly high covers of live coral with quite diverse species composition at a depth of 3 m, further north - southwest, reef flat is wider between 200 - 500 m.

The unsupervised classification results with the IsoData method and the K-Means method on SPOT-6 image produce 16 classes and in Sentinel-2 image produce 10 classes with the same iterative number of 100 and the threshold number 3. Unsupervised classification on images produce different colors formed based on different spectral values. The difference in the results of reclassification of IsoData and K-Means with the application of the lyzenga algorithm can be seen in Figures 2 & 3.

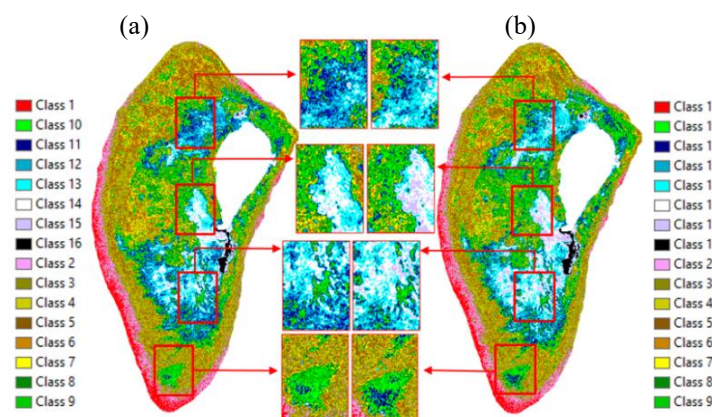


Figure 2. Unsupervised Classification Results of IsoData and K-Means Methods in SPOT-6 images with the application of the Lyzenga Algorithm (a) the classification results of the IsoData Method, (b) the results of the K-Means method classification

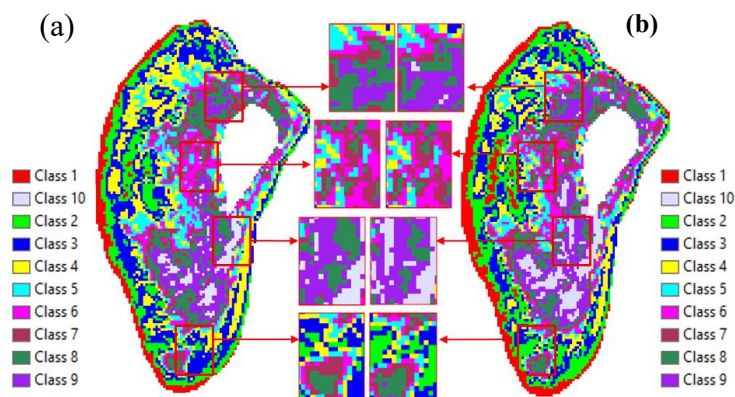


Figure 3. Results of Unsupervised Classification of IsoData and K-Means Methods in Sentinel-2 images with the application of the Lyzenga Algorithm (a) the classification results of the IsoData Method, (b) the results of the K-Means method classification

Unsupervised classification IsoData method results shows that there are 12 classes of shallow water objects and K-

Means method results shows 13 classes after reclassification process.

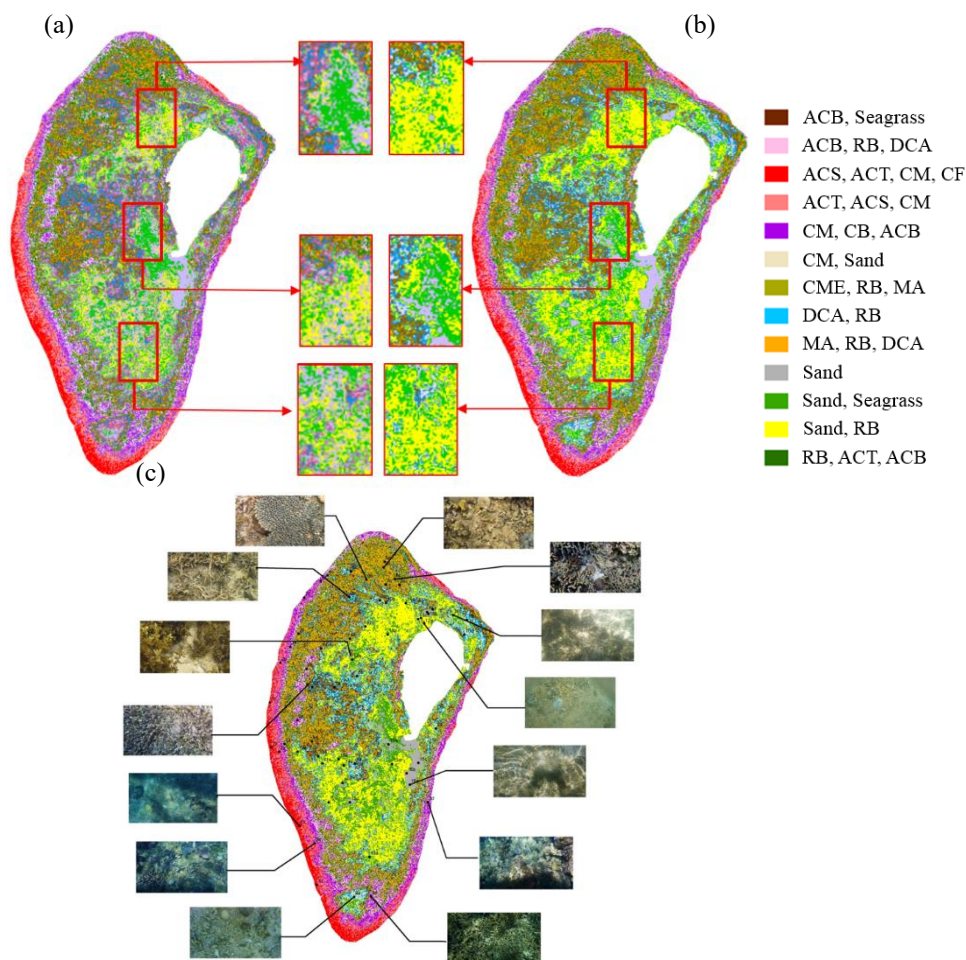


Figure 4. Differences in reclassification results with the IsoData and K-Means methods on SPOT-4 images. (a) The results of reclassification using the IsoData method, (b) Results of reclassification using the K-Means method, and (c) field survey

Living coral classes represent reef areas that are still growing and developing with good closure. Dead coral classes represent areas that are mostly dead coral reefs and dead corals overgrown with algae (Dead Coral Alga). The class of coral fragments represents an area consisting of crushed corals, in the form of coral fragments. The sand class represents an area that is still dominated by small coral fragments and fine sand in shallow waters. Seagrass classes represent sandy areas dominated by seagrass plants. In accordance with the objectives of the study, the land area was separated beforehand to avoid classes that were mixed with waters.

The difference in the results of reclassification of IsoData and K-Means with the application of the lyzenga algorithm can be seen in Figure 15. Area calculation in the Sentinel-2 image classified in 2019 by the IsoData method produces a live coral class of 260,200 m², algae 63,800 m², rubble class 149,400 m², sand classes covering 161,600 m² and seagrass classes covering 85,800 m². The cover of the seabed substrate waters of Barrang Caddi P. in 2019 is more dominated by living coral and coral fragments. While the area calculation in the Sentinel-2 image is classified in 2019 using the K-Means method produces a living coral class of 193,400 m², an algae of 57,900

m², a class of rubble covering 120,900 m², a sand class of 186,800 m² and seagrass classes of 130,900 m². The base substrate cover of Barrang Caddi waters in 2018 is more dominated by living coral and sand.

Based on the comparison of the results of the unsupervised classification between IsoData and K-Means, it shows that on the west of Barrang Caddi island, there is an increase in the area of deep sea, coral reefs and seagrass on the K-Means method. Meanwhile, in the IsoData method, it was more dominantly detected as live coral and algae. The ground truthing point used as a reference in the reclassify process showed that on the north and south of Barrang Caddi Island, rubble objects that were more dominant in the K-Means method were detected as live corals in the IsoData method, so that in the IsoData method, live coral objects were more dominant than in the K-Means method. The results of field verification indicate that there are variations in the structure and morphology of various shallow water habitat, each of which has a pattern of electromagnetic response.

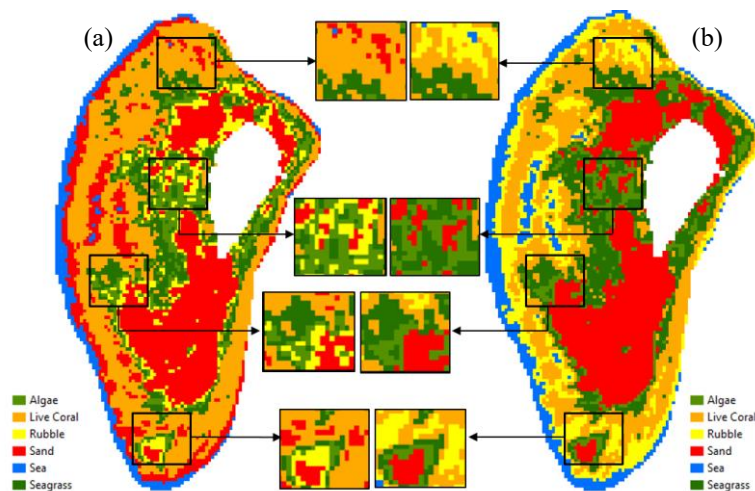


Figure 5. Differences in reclassification results with the IsoData and K-Means method on Sentinel-2 images. (a) The results of reclassification using the IsoData method, (b) The results of reclassification using the K-Means method.

There are 40 points scattered on Barrang Caddi Island where these points represent each classes with each of the 4 replications (a, b, c, d) formed from the results of unguided image classification. Accuracy tests are carried out on each classification method, IsoData and K-Means methods. Based on the results of the accuracy test, the value of accuracy using the unsupervised classification IsoData method is 62.50% (25 out of 40 pixels), which has the same accuracy as the unsupervised classification of the K-Means method which is 62.50% (25 out of 40 pixels). This accuracy value only shows the accuracy of each method in performing classification based on differentiation and compatibility with the ground-truthing results. Accuracy value similarity between the two methods does not mean that the objects being classified are the same. Real condition on the field shows 10 objects are different from the classification result. Those objects are different in the Iso Data method and K Means method, as shown in Figure 5.

From the observation result of coral habitats damage in Barrang Caddi, most of which are a result of human activities, this can be seen from the number of rubble caused by fishing activities using explosives / bombs and chemicals by fishermen to catch fish (Nurjannah., et al 2015) in Barrang Caddi's water. Another phenomenon that causes a decrease in the area of live coral cover and the increasing area of dead coral cover in Barrang Caddi island is the coral bleaching phenomenon.

The graph of the SPOT-6 image spectral reflectance pattern from the classification of unsupervised IsoData and K-Means with the application of the lyzenga algorithm is formed based on the tabulation of the average spectral reflectance of shallow water habitat (Figure 6 & 7). Spectral reflectance pattern of shallow water habitat based on unsupervised K-Means classification with the application of the lyzenga algorithm can be seen in Figure 6 & 7.

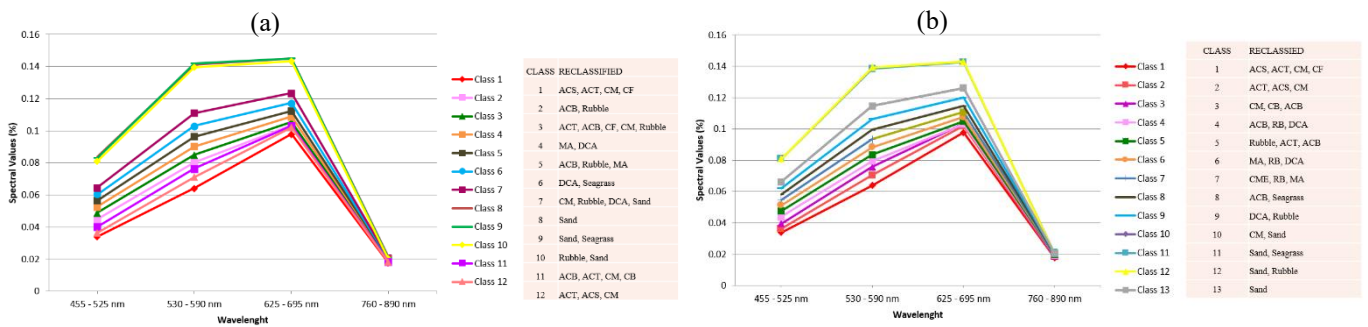


Figure 6. Spectral reflectance pattern of water base objects based on IsoData and K-Means method on SPOT-6 image, (a) Spectral reflectance pattern with IsoData method, (b) Spectral reflectance pattern using K-Means method.

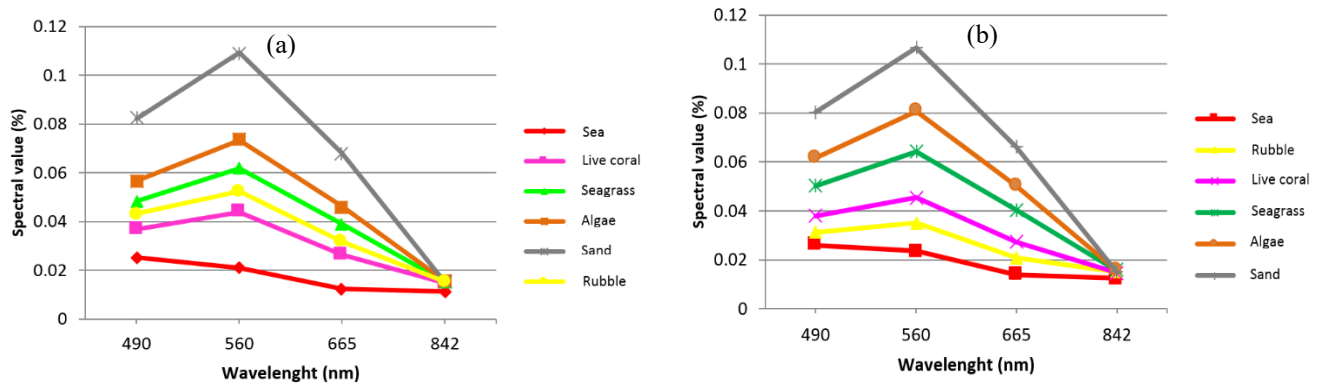


Figure 7. Spectral reflectance pattern of shallow water habitat based on IsoData and K-Means method on Sentinel-2 images, (a) Spectral reflectance pattern with IsoData method, (b) Spectral reflectance pattern using K-Means method.

The reflectance pattern graph shows that the two unsupervised classification methods IsoData and K-Means have an object reflectance pattern, where deep sea have the highest reflectance value on a blue channel with a length of 490 nm, this is because the green, red and near infrared wavelengths are absorbed more by the deep sea so the deep sea has a darker color appearance. Whereas other shallow water habitat, such as live coral, coral fault, sand, algae and seagrass shows reflectance values that tend to be high in green channels with a wavelength of 560 nm, this is because these objects tend to reflect more at green wavelengths and absorb blue and red wavelengths.

There are difference between the two unsupervised IsoData and K-Means classification methods in classifying objects based on their spectral reflectance values. In the results of the IsoData unsupervised classification, coral fracture objects have higher spectral reflectance values than live coral objects, while in unsupervised K-Means classification, live coral has higher spectral reflectance value than a coral fracture object. However, the highest spectral reflectance value of the IsoData method and the K-Means method is sand with the same values each.

5. CONCLUSIONS

Based on the comparison of the results of the unsupervised classification between IsoData and K-Means, it shows that there is an increase in the area of deep sea, coral reefs and seagrass on the K-Means method. The IsoData method, it is more dominantly detected as live coral and algae. Rubble

objects were more dominant in the K-Means method and they detected as live corals in the IsoData method, live coral objects is more dominant than the K-Means method result. The accuracy of unsupervised classification IsoData method is 62.50% (25 out of 40 pixels), which has the same accuracy as the unsupervised classification of the K-Means method which is 62.50% (25 out of 40 pixels). Based on environmental conditions in the Spermonde Islands, it is recommended to use Iso Data for detecting and mapping coral reefs, while K Means method is better to mapping seagrasses.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Research Technology and Higher Education of the Republic of Indonesia for the generous support in providing research funds. Our appreciation is further extended to the Ocean Remote Sensing Project of Subcommittee of Western Pacific Intergovernmental Oceanographic Commission/UNESCO supported by Japan Fund in Trust provided by Ministry of Education, Culture, Sports, Science and Technology, Japan. Thanks to Environmental Systems Research Institute (ESRI) Indonesia for software supporting

REFERENCES

Agus SB. 2012. *Kajian Konektivitas Habitat Ontogeni Ikan Terumbu Menggunakan Pemodelan Geospasial Di Perairan Kepulauan Seribu*. [disertasi]. Bogor (ID): Institut Pertanian Bogor

- Ampou, E.E., S. Ouillon, and S. Andrefouet. 2017. Challenges in rendering coral triangle habitat richness in remotely sensed habitat maps: The case of Bunaken island (Indonesia). *Marine Pollution Bulletin*, 131: 72-82. <http://dx.doi.org/10.1016/j.marpolbul.2017.10.026>
- Congalton, R.G. and K. Green. 1999. *Assessing the accuracy of remotely sensed data: principles and practices*. CRC Press, Inc., Florida. 130p
- Deas, M., Andréfouët, S., Léopold, M. & Guillemot, N. (2014). “Modulation of Habitat-Based Conservation Plans by Fishery Opportunity Costs: A New Caledonia Case Study Using Fine-Scale Catch Data”. *Plos One* 9 (5): e97409–e97409. <https://doi.org/10.1371/journal.pone.0097409>
- El-Askary, H., S.H. Abd El-Mawla, J. Li, M.M. El-Hattab, and M. El-Raey. 2014. Change detection of coral reef habitat using Landsat-5 TM, Landsat 7 ETM+ and Landsat 8 OLI data in the Red Sea (Hurghada, Egypt). *Int. J. Remote Sensing*, 35(6):2327-2346.
- Giyanto, Manuputty, A. E., Abrar, M. dan Siringoringo, R. M., 2014. Monitoring Terumbu Karang. In: *Panduan Monitoring Kesehatan Terumbu Karang*. Jakarta: COREMAP CTI LIPI, p. 63
- Green, E.P., P.J. Mumby. A.J. Edwards, and C.D. Clark. 2000. *Remote sensing handbook for tropical coastal management*. UNESCO, Paris
- Henriques, V., M.T. Guerra, B. Mendes, M.J. Gaudencio, and P. Fonseca. 2014. Benthic habitat mapping in a Portuguese marine protected area using EUNIS: An integrated approach. *J. of Sea Research*, 30: 14-28. <http://dx.doi.org/10.1016/j.seares.2014.10.007>
- Phinn, S.R., C.M. Roelfsema, and P.J. Mumby. 2011. Multi-scale. object-based image analysis for mapping geomorphic and ecological zones on coral reefs. *International J. of Remote Sensing*, 33:3768-379
- Jaya INS. 2010. *Analisis Citra Digital: Prespektif Penginderaan Jauh Untuk Pengelolaan Sumberdaya Alam*. Bogor (ID): Fakultas Kehutanan IPB
- Mumby PJ, Skirving W, Strong AE, Hardy JT, LeDrew E, Hochberg EJ, Stumpf RP, David LT. 2004. Remote sensing of coral reefs and their physical environment. *Marine Pollution Bulletin*. 48: 219-228
- Nurdin N, Komatsu T, Agus, Akbar AS M, Djalil AR, Amri K. 2015. Multisensor and multitemporal data from Landsat images to detect damage to coral reefs, small islands in the Spermonde Archipelago, Indonesia. *Ocean Science Journal* 50(2): 1-9
- Wang, R., 2016. *Classification : Unsupervised Classification-Clustering*. [Online] Available at: <http://fourier.eng.hmc.edu/e161/lectures/classification/node12.html> [Accessed 6 Juni 2016].
- Wilson, S.K., N.A.J. Graham, and N.V.C. Polunin. 2007. Appraisal of visual assessments of habitat complexity and benthic composition on coral reefs. *Mar Biol.*, 151:1069–1076 (<http://earthexplorer.usgs.gov/>)

Revised August 2019