Accuracy assessment of relative and absolute water column correction methods for benthic habitat mapping in Parang Island

A A Hadi¹ and P Wicaksono¹
¹Department of Geography Information Science, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia

*Corresponding author: afifalhadi@mail.ugm.ac.id, afifalhadi13@gmail.com

Abstract. The condition of benthic habitats in optically shallow sea waters becomes important information in the inventory and processing of coastal resources. Remote sensing is effective and efficient in mapping benthic habitats. This study aims to apply absolute and relative water column correction methods in order to map benthic habitats on Parang Island using PlanetScope image. The benthic habitat classification scheme used consists of coral reefs, seagrass, macroalgae, and substrate. We compared the accuracy of benthic habitat map based on absolute and relative water column correction methods. The classification methods used are the Maximum Likelihood (ML) algorithm and Support Vector Machine (SVM). The results showed that benthic habitat map with the highest accuracy was obtained by a combination of Lyzenga-ML at 61.63% followed by Purkis-SVM at 59.18%, Lyzenga-SVM at 41.90%, and Purkis-ML 16.87%. The results show that the Lyzenga water column correction method is the best choice in mapping benthic habitats.

1. Introduction
The existence of natural resources in coastal areas and small islands is added value to Indonesia. The exploitation of natural resources can be carried out to increase the development of coastal areas. The potential for development is not only in tourism and its economic potential, but also the ecological potential, especially for benthic habitats within this marine ecosystem. Therefore, monitoring is needed for the management of these areas. The best solution to monitoring benthic habitat is using multispectral remote sensing [1]. Monitoring and mapping of benthic habitat using multispectral remote sensing data requires several stages. One of the most important stages is water column correction.

The effect of the water column on the accuracy of benthic habitat mapping is quite significant [1]. There are two types of water column correction methods commonly used in benthic habitat mapping: absolute water column correction and relative water column correction. The significant difference between the two methods is the inputs required in the form of depth data and the value of the water column attenuation coefficient.

The relative water column correction most often used is the Lyzenga water depth invariant bottom index algorithm [2]. Previous studies have been conducted by [3], [4], and [5]. Lyzenga water column corrections have a good ability to normalize depth effect without an actual attenuation coefficient. The absolute water column correction method rarely used in the mapping of benthic habitats is the
developed by [6]. The Purkis and Pasterkamp water column corrections method have a good ability to calculate water column effects. This is influenced by the use of depth data and attenuation coefficients as input in the process. However, it is necessary to study the efficiency of each water column correction method in mapping benthic habitats in tropical areas. This study aims to map the benthic habitat of Parang Island using PlanetScope image and to compare the accuracy of the mapping of benthic habitats on Parang Island as a result of relative and absolute water column correction.

2. Materials and Methods

2.1. Study Area

This research was conducted in Parang Island, Karimunjawa Islands, Jepara District, Central Java Province, which is located at UTM Zone 49M 414095 – 418254mE and 9367229 – 9361615mN. The geomorphology of Parang is in the form of a fringing reef, where there is no lagoon and the reef flats are large. This is indicated by the distribution of coral reefs that are more than 50 meters from the coastline.

2.2. Materials

This research was conducted using the PlanetScope image acquired on April 30, 2018. The PlanetScope image has a spatial resolution of 3 m and includes visible and near-infrared bands. Other data used are depth data and transect photos. Depth data were obtained from a bathymetric survey using an echosounder mounted on a ship. Meanwhile, transect photos were obtained using a transect photo survey.

2.3. Methods

Sample preparation. Field data collection was carried out based on the determination of sample points during pre-field work using the photo-transect method [7]. The benthic habitat mapping survey in this study used the underwater photo-transect method. Photo-transect was used to describe the structure of the coral reefs community by looking at coral cover, substrate types (sand, mud), macroalgae, and the presence of biota. The transect line placement is based on coastal landscape characteristics on Parang Island. In addition to photo-transect, an echosounder was used to obtain bathymetric data whose values would be used for making bathymetric maps and absolute water column correction.

Bathymetry model. The equation used to create a bathymetric image was developed by [6] as an alternative method that uses a simpler formula. This method is an empirical method built through a regression equation between the image pixel value and the actual depth in the field. Therefore, the bathymetric image results from this modeling are highly dependent on the maximum and minimum depth of field samples.

Water depth invariant bottom index algorithm. The water column correction method by [2] can be formulated as follows:

\[ D\ln I = \ln(B_i) - \left( \frac{k_i}{k_j} \right) \ln(B_j), \]  

where the value \( B_i \) and \( B_j \) are visible bands, \( k_i/k_j \) is the ratio of water column attenuation coefficient between two bands. The value \( k_i/k_j \) is obtained from Equation (3):

\[ k_i/k_j = a + \sqrt{a^2 + 1}. \]
where the value of $\alpha$ is the result of reducing the variance between bands divided by two times the covariance between bands with Equation (4):

$$
\alpha = \frac{\text{var}B_i - \text{var}B_j}{2 \times \text{Cov}B_i B_j},
$$

(Purkis and Pasterkamp algorithm). The water column correction method of [6] focuses on depth information with the assumption that the attenuation of energy in the water column in each water is homogeneous. Depth correction can be done with Equation (5):

$$
D = \frac{\ln(R_b - R_d) - \ln(R_b - R_s)}{-k(1+\tan(E))},
$$

where $D$ is the depth of each pixel resulting from the difference between the deepest pixel value and the shallowest pixel value and divided by the attenuation coefficient and zenith angle. $R_b$ is the overall pixel value or single visible band, $R_d$ is the deepest pixel value, $R_s$ is the shallowest pixel value, $k$ is the attenuation coefficient and is assumed to be homogeneous across the scene, and $E$ is the zenith angle. Next, the depth value ($D$) will be entered into the water column correction with Equation (6):

$$
R_b = \left(\frac{1}{0.646}\right) \exp\left(-\frac{k}{1+2\tan(E)}\right) x R_w,
$$

where $R_b$ is the water column corrected band, $k$ is the water column attenuation coefficient, $R$ is the visible band to be corrected, $D$ is the depth of each pixel that is corrected by depth, and $R_w$ is the value of the reflectance from optically deep water.

(Multispectral classification). Image classification was done using a pixel-based approach using water column corrected bands as input. The benthic habitat class used in the classification refers to the major classes in the form of (1) macroalgae, (2) bare substrate, (3) coral reefs, and (4) seagrass. Image classification uses the maximum likelihood (ML) and support vector machine (SVM), which is classified as supervised classification. The Region of Interest (ROI) were obtained from photo-transect field surveys.

3. Result

The classification scheme in the study was constructed based on the variation of major classification benthic habitats in the field. Benthic habitat classes at a major level are used because they can provide information related to monitoring activities [8]. The variation of benthic habitats in the field provides an overall view of the presence and distribution of existing benthic habitat. The benthic habitat mapping in this study uses two types of classification algorithms—maximum likelihood (ML) and SVM—with a combination of relative and absolute water column correction. The use of different water column corrections gave different results, which will indicate a different distribution of the object. The dominant class classified in ML-Lyzenga is coral reef with the composition of other classes almost evenly distributed, which contrasts with other results where the dominance of one class is very contrasting. The dominant class classified in ML-Purkis is the substrate. Meanwhile, the dominant class classified in SVM-Lyzenga and SVM-Purkis is coral reefs. This is due to the condition of the sand and macroalgae, which have similar reflectance characteristics seagrass at particular percent cover [9]. Likewise, coral reef objects are misclassified into macroalgae because the spectral response of macroalgae are very similar to the spectral response of coral reef [9].

The distribution of objects in ML-Lyzenga classification are relatively in accordance with field conditions; coral reefs are evenly distributed throughout the area, sand is clustered in the southern part, overestimated seagrass is distributed on the deeper parts, and macroalgae are distributed randomly. However, the distribution of class in the ML-Purkis classification result is not as good as the relative
ML-Lyzenga result. Coral reef and substrate in the ML-Purkis result are highly overestimated, where the objects that should be classified as macroalgae and seagrass are classified as coral reefs or substrate. For more details related to misclassification, see Tables 1 and 2, and Figures 1 and 2.

| Class       | Reference | Total | UA (%) |
|-------------|-----------|-------|--------|
| Coral       | 212       | 274   | 77.37  |
| Seagrass    | 2         | 12    | 33.33  |
| Macroalgae  | 83        | 98    | 14.29  |
| Substrate   | 91        | 178   | 29.78  |
| Total       | 388       | 735   |        |

Table 1. Confusion matrix from ML-Lyzenga classification result

| Class       | Reference | Total | UA (%) |
|-------------|-----------|-------|--------|
| Coral       | 51        | 59    | 86.44  |
| Seagrass    | 203       | 306   | 9.15   |
| Macroalgae  | 90        | 115   | 7.72   |
| Substrate   | 44        | 72    | 30.56  |
| Total       | 388       | 735   |        |

Table 2. Confusion matrix from ML-Purkis classification result

Figure 1. Benthic habitat classification using ML-Lyzenga.
Figure 2. Benthic habitat classification using ML-Purkis.

The distribution of objects in the classification results of benthic habitats using the SVM-Lyzenga and SVM-Purkis schemes has a distribution pattern that is very different from the results obtained using the ML-Lyzenga and ML-Purkis schemes. The SVM-Purkis classification results show that coral reef objects dominate the entire area, with other objects randomly distributed. The coral reef object became highly underestimated, resulting in a class that seemed irrelevant. The sand objects are not suitable with the field condition, especially in the southern part of the island, where sand is misclassified as coral reefs. Meanwhile, the distribution of macroalgae objects in the SVM-Purkis result is more in accordance with the field condition although it is slightly misclassified into coral reefs. Details of misclassified objects can be seen in Tables 3 and 4, and Figures 3 and 4.

| Class     | Reference | Total | UA (%) |
|-----------|-----------|-------|--------|
| Coral     | 350       | 13    | 10     | 60 | 433 | 80.83 |
| Seagrass  | 9         | 12    | 5      | 16 | 42  | 28.75 |
| Macroalgae| 9         | 7     | 4      | 8  | 28  | 14.29 |
| Substrate | 84        | 46    | 15     | 87 | 232 | 37.50 |
| Total     | 452       | 78    | 34     | 171| 735 |      |
| PA (%)    | 77.43     | 15.38 | 11.76  | 50.88| |
| OA (%)    | 61.63     |       |        |     |     |     |

Table 4. Confusion matrix SVM-Purkis.

| Class     | Reference | Total | UA (%) |
|-----------|-----------|-------|--------|
| Coral     | 384       | 34    | 18     | 81  | 517 | 74.27 |
| Seagrass  | 7         | 12    | 0      | 24  | 43  | 27.91 |
| Macroalgae| 11        | 7     | 8      | 35  | 61  | 13.11 |
| Substrate | 50        | 25    | 8      | 31  | 114 | 27.19 |
| Total     | 452       | 78    | 34     | 171 | 735 |      |
| PA (%)    | 84.96     | 15.38 | 23.53  | 18.13| |
| OA (%)    | 59.18     |       |        |     |     |     |
The resulting accuracy calculation shows that PlanetScope image can map benthic habitats. Based on these results, it is known that the accuracy of the relative water column correction using Lyenga [2] method is 61.63% and 41.90% for SVM and ML, respectively. The accuracy of the absolute water column correction method using Purkis [6] method is 59.18% and 16.87% SVM and ML, respectively. Generally, high-resolution images can provide a high-accuracy result. However, this is not the case
with the result obtained where there is an accuracy of 16.87% in absolute water column correction with SVM algorithm. This is due to the inconsistent reflectance values, especially on coral reef objects. The consistency of the image itself can be determined based on the range of pixel values in the same object with different depths. Furthermore, this is indicated by a mismatch in the k-band modeling or attenuation coefficient band at the Purkis water column correction stage. The attenuation coefficient value that should represent the entire pixel value is not representative. The existence of a negative coefficient value is the cause of the decrease in the accuracy of the resulting benthic habitat map.

On the other hand, the difference in accuracy between the two water column correction methods is also caused by the depth data collection process. Field conditions wherein it is quite difficult for ships to pass have resulted in the inaccessibility of data in several areas. This is a limitation in the research conducted as it affects the bathymetry sample determination. The determination of the sample is related to the distribution and representation of samples to field conditions. The sample is divided into two: the bathymetric model sample and the classification model sample. The bathymetric model sample used in the absolute water column correction stage as a reference has not fully represented the existing depth values. It is known that the accuracy of the bathymetry model used is 57.67% with a standard error of 1.58 m. Referring to a study conducted by [10] with an SE value of 4.31 m, the bathymetry model can be used for classification. This can be done with the assumption that the optically shallow water has a relatively variable depth range from 0 to 22 m.

The accuracy resulting from this study ranges from 16% to 61%, which is quite relevant to several previous studies where an accuracy of 16.87% in ML-Purkis is an exception. Based on previous research, the accuracy of benthic habitat mapping using the Lyzenga water column correction method was 67.70% [4], 58.61% [3], and 89% [5]. The three studies were conducted in the same area of Kemujan Island and Karimunjawa Island. However, the three studies only used the multispectral maximum likelihood classification. Another study conducted by [11] resulted in an accuracy of 57.26% using the Lyzenga water column correction with the SVM classification algorithm. Referring to the four studies, the accuracy value for the Lyzenga water column correction can be categorized as proper to apply. Fewer errors were propagated due to only using statistics from the image itself from which is the reflectance of objects at different depths.

Meanwhile Purkis method requires information on the depth of each pixel and also the water column attenuation coefficient. The accuracy of the predicted-depth for each pixel and also the process of obtaining the attenuation coefficient value will greatly affect the success or failure of the Purkis method. Thus, if the two parameters are less accurate, then the Purkis correction results will also be less than optimal because errors in bathymetry and attenuation coefficient predictions propagate in the results of the water column correction. This is what makes the Lyzenga algorithm better, especially for PlanetScope images that contain a lot of noise.

4. Conclusion
Based on this research, it can be concluded that the distribution of benthic habitat on Parang Island is evenly distributed, with some parts dominated by coral reefs or bare substrate. Based on this research, benthic habitat mapping using the Lyzenga water column correction method is still the main choice. This is supported by an accuracy value of 61.63% and 41.90% using support vector machine and maximum likelihood, respectively. Meanwhile, the accuracy of the Purkis water column correction was 59.18% and 16.87% using support vector machine and maximum likelihood, respectively. Increasing the accuracy value using the water column absolute correction method is still possible by selecting a more selective and representative sample in both the bathymetry and classification model. The selection of images with a more consistent pixel value is recommended in the application of the absolute water column correction model.
References

[1] Green EP, Mumby PJ, Edwards AJ, Clark CD. Remote Sensing Handbook for Tropical Coastal Management. Coastal Management Sourcebooks. Paris: UNESCO; 2000.

[2] Lyzenga DR. Passive remote sensing techniques for mapping water depth and bottom features. Appl Opt. 1978;17(3) p 379.

[3] Bima DPS. Pemetaan Ekologi dan Morfologi Habitat Bentik Menggunakan Citra Penginderaan Jauh Resolusi Tinggi di Pulau Kemujan Karimunjawa Provinsi Jawa Tengah. Universitas Gadjah Mada; 2015.

[4] Putra FMG, Khakhim N. Pemetaan Habitat Bentik Menggunakan Citra Quickbird di Sebagian Pulau Kemujan, Kepulauan Karimunjawa. J Bumi Indones. 2014;3(2).

[5] Nurlidiasari M, Budhiman S. Mapping Coral Reef Habitat With and Without Water Coulumn Correction using Quickbird Image. Remote Sens Earth Sci. 2005;2(September) p 45–56.

[6] Purkis SJ, Pasterkamp R. Integrating in situ reef-top reflectance spectra with Landsat TM imagery to aid shallow-tropical benthic habitat mapping. Coral Reefs. 2004;23(1) p 5–20.

[7] Roelfsema C, Phinn S. Integrating field data with high spatial resolution multispectral satellite imagery for calibration and validation of coral reef benthic community maps. J Appl Remote Sens. 2010;4.

[8] Wicaksono P. The Use of Image Rotations on Multispectral-Based Benthic Habitats Mapping. In: 12th Biennial Conference of Pan Ocean Remote Sensing Conference. Bali; 2014.

[9] Wicaksono P. Improving the accuracy of multispectral-based benthic habitats mapping using image rotations: The application of principle component analysis and independent component analysis. Eur J Remote Sens. 2016;49 p 433–63.

[10] Poursanidis D, Traganos D, Reinartz P, Chrysoulakis N. On the use of Sentinel-2 for coastal habitat mapping and satellite-derived bathymetry estimation using downscaled coastal aerosol band. Int J Appl Earth Obs Geoinf. 2019;80(March) p 58–70.

[11] Lazuardi W. Kajian Citra Multiresolusi untuk Pemetaan Life-Form Terumbu Karang Studi Kasus Pulau Parang, Kepulauan Karimunjawa. Universitas Gadjah Mada; 2018.

Acknowledgement
This research was funded by "RTA Research Scheme" and the data used in this research was partially obtained from the research funded by “Direktorat Riset dan Pengabdian Masyarakat — Direktorat Jenderal Penguatan Riset dan Pengembangan – Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia” through “Penelitian Berbasis Kompetensi” Research Scheme, Grant No.1697/UN1/DITLIT/DIT-LIT/LT/2018.