Benthic Habitat Mapping Using Sentinel-2A Imagery in the Waters of Pengudang Village, Bintan Regency

Jumi Safitri[1] Esty Kurniawati[2] Asep Ma'mun[3] Try Febrianto[4]
Marine and Fisheries Science of Department, Faculty of Marine and Fisheries Sciences, Raja Ali Haji Maritime Univesity, Tanjung Pinang Regency, Riau Islands Province, Indonesia[1,2,3,4]
Email: jumisafitri410@gmail.com[1] estykurniawati@umrah.ac.id[2] asepmamun@umrah.ac.id[3] try.febrianto@umrah.ac.id[4]

Abstract

Pengudang Village has a diverse shallow-water ecosystem that forms the seabed and not much research has been done to map benthic habitats in these waters. Thus, the availability of spatial data regarding benthic habitats in the area is very limited. This study aims to map shallow water benthic habitat using Sentinel-2A imagery and calculate the accuracy of benthic habitat classification in Pengudang Village waters. Benthic habitat classification is carried out in 4 cover classes, namely seagrass (LM), sand (PS), sand with dead coral (KMP), and living coral with sand (PKH). In this study, we used supervised pixel-based image classification using two algorithms, namely Support Vector Machine (SVM) and Minimum Distanse Classification (MDC) with thematic layer input from field data. The results showed that the SVM classification algorithm model had a better accuracy of 76% while the MDC algorithm model had an accuracy of 66%. Based on the results of the study, the difference between the two classification algorithms is 10%, and the best classification algorithm in the waters of Pengudang Village is SVM.
Keywords: Benthic Habitat, Mapping, Sentinel-2A Imagery, Pengudang Village

INTRODUCTION

Indonesia is a country that has approximately 17,504 islands consisting of large and small islands (Inounu et al., 2017). Pengudang Village is one of the villages on Bintan Island that is visited by many tourists from within the country and abroad. This area has a diversity of unspoiled ecosystems consisting of seagrasses, algae, coral reefs, and mangroves, so management of the area is needed so that the ecosystem is maintained. In addition, this area has not done much spatial mapping research. Thus, researchers are interested in mapping benthic habitats in shallow waters of Pengudang Village using Sentinel-2A imagery with the application of SVM and MDC algorithms. Benthic habitat mapping is an effort to map the state of objects on the bottom of waters such as coral reefs, seagrasses, coral fragments, algae and sand (Talitha, 2016). This mapping is done by utilizing remote sensing techniques that provide spatial and temporal information. This technology works without direct contact with objects and utilizes a wide array of optical sensors ranging from low to high spatial resolution (Dekker et al., 2005; Hedley et al., 2009).

Sentinel-2A image is one of the medium satellite images that has a spatial resolution of 10 meters. Various studies on mapping benthic habitats in shallow waters have been carried out using this image with several algorithm models used. Mukrimin et al., (2021) have mapped benthic habitats using Sentinel-2A imagery in 2020 and 2016 by applying the SVM algorithm, on the Coast of Pulau Tiga (Tiworo Strait) which resulted in accuracy rates of 88% and 30% respectively from 6 classes. Another study was also conducted in the Kei Islands, North Maluku using the same image by applying the Maximum Likelihood Classification (MLC), Minimum Distance Classification (MDC), and Mahalanobis Distance Classification

(MaDC) algorithms which resulted in accuracy rates of 74.33%, 73.45%, and 78.35% respectively (La Ode Aliftari, 2022). Therefore, this study aims to map benthic habitats using Sentinel-2A minds and compare yhe accuracy level of the application of SCM and MDC algorithms in yhe waters of Pengudang Village, Bintan Regency.

RESEARCH METHODS
Time and Place

The research data collection will take place on June 10 - 11, 2023 in Pengudang Village, Bintan Island, Riau Islands. The sampling method used is systematic random sampling, where the determination of data collection points is carried out randomly, with a distance of each observation point of 15 meters to other points. Data processing starts on June 14 – September 8, 2023 which is carried out at the OCM (Oceanography, Computation, Modelling) Laboratory, Faculty of Marine and Fisheries Sciences, Raja Ali Haji Maritime University, Tanjung Pinang City. Peta the location of the study is shown in Figure 1.
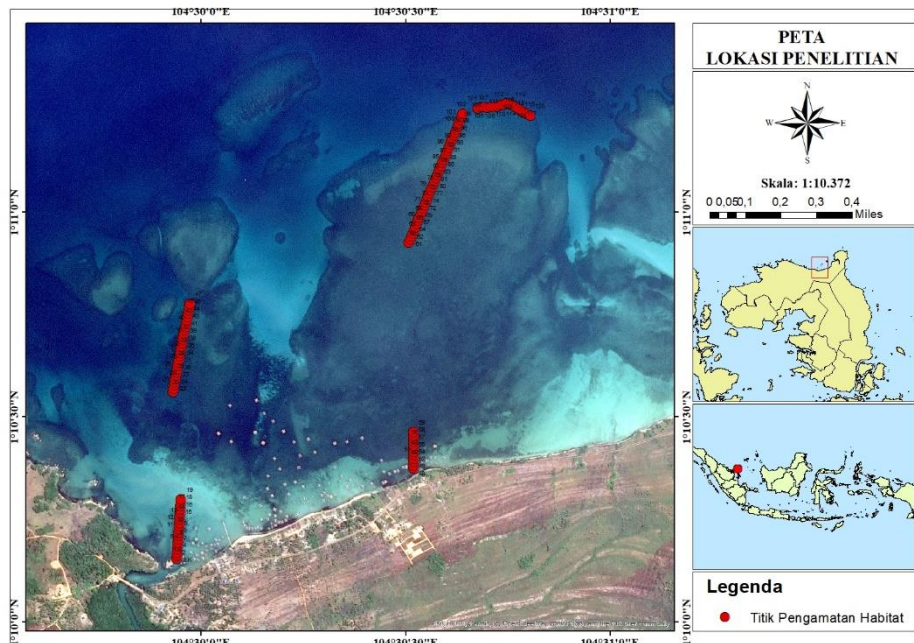

Figure 1. Map of Research Location

Tools and Materials

The equipment used in this study is a field survey device and software for data analysis. The following are some of the tools and materials and software used in this study presented in Table 1 and 2.

Table 1. Field Survey Equipment

No	Tools and Materials	Uses
1	GPS	Determine coordinate points
2	Underwater Camera	Habitat class photo capture
3	Diving equipment	Habitat class validation and photo capture
4	Transect squared 1 m x 1 m	Observation of benthic habitat
5	Laptop	Data processing and analysis hardware
6	Pencil and newtop paper	Underwater writing aids
7	Sentinel 2A imagery	Image data to be processed

Table 2. Software Type

No	Software	Role
1	Coral Point Count With Extensions (CPCe) 4.1	Analysis of benthic habitat cover
2	ArcGIS 10.8	Map layeout creation
3	Microsoft Excel 2013	Data tabulation, diagram presentation
4	XLSTAT	Statistical analysis
5	ENVI 5.3	Image preprocessing

Data Collection Techniques

Field data collection includes direct observation (GTH) at the research site and taking photos using underwater cameras. Direct observations were made using a 1 m x 1 m square transect (English et al., 1997; Roelfsema & Phinn 2008) and given a striking color to be seen while in the water. Transect data measuring 1 m x 1 m is adjusted to the spatial resolution of the image used, namely the Sentinel-2A image.
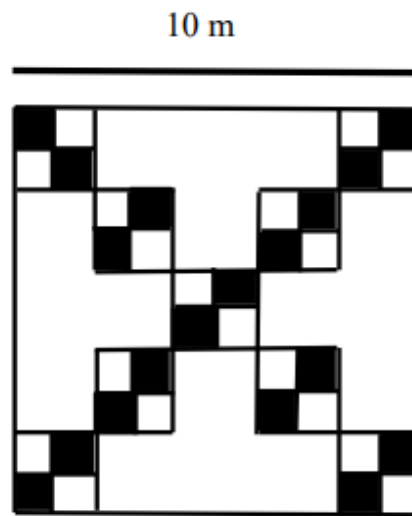

Figure 2. Sentinel-2A Image Data Capture Transect

The method of collecting benthic habitat data uses systematic random sampling, where points are determined randomly with each point 15 meters away from the other. The number of field observation points is 120 points, where the sample points will be divided into 2, namely 70 points used as a reference in the calcification process and 50 points used as accuracy tests. The data of each cover photo was analyzed using Coral Point Count with Excel extensions (CPCe) software (Kohler & Gill 2006). Furthermore, analyzed by Agglomerative Hierarchical Clustering (AHC) method with XLSTAT software was used for bray-curtis dissimilarity analysis and produced a dendogram of benthic habitat classification scheme based on field survey data. Based on the results of the field survey of the water area of Pengudang Village which has been carried out, AHC analysis has produced four classes, namely sand (PS), seagrass (LM), dead coral sand (KMP), and live coral sand (PKH).

Data Processing Techniques

The satellite image used in this study is Sentinel-2A imagery, this image is processed using ENVI 5.3 software which begins with the process of compossite band, cropping, and data masking.
1. Layer Stacking and Composite Band. Stacking or Layer stacking is the process of combining bands into one layer. Composite bands are used to unite several image bands so that recorded objects are easy to recognize and distinguish. Compossite bands provide RGB

(Red-Green-Blue) colors like color camera photos. The compossite band used in benthic habitat analysis in shallow water is true color composite, which is a blue-red-green band (Prayudha, 2014).

2. Cropping and Masking Data. Cropping is an image cropping process carried out to limit the study area to be mapped to match the research location. Image cropping is done using the Data Masking process. This process is an analysis to eliminate land areas, deep waters, and other objects that are not needed in the process of processing image data. This process utilizes channel 4 (NIR) because of the considerable spectral response between land and sea areas. Cropping and masking aim to limit the area of image analysis. The process is well done using delimiter values on certain objects to be analyzed digitically.

Data Analysis
Benthic Habitat Cover Categories

Each photograph using a 1m x 1m squared transect was analysed using Coral Point Count with Excel extensions (CPCe) software to obtain a benthic habitat cover category [8]. CPCe is software developed through visual basic that automatically analyzes random point counts. In addition, the software is able to perform statistical analyses (e.g. relative abundance, average, standard division, standard error) for each spreadsheet type or substrate in Microsoft Excel and calculate Shannon-Weaver for each species diversity index. CPCe has four random point calculation models, namely simple random, stratified random, uniform grid, and equally spaced grid. The type of stratifield point specification used in research is a uniform grid, namely points identified in the form of rows and columns, consisting of 30 points superimposed on each quadrant photo. (Fig. 3)

Figure 3. Percentage Analysis Of Benthic Habitat Using Cpce

AHC (Agglomerative Hierarchical Clustering) Analysis

AHC (Agglomerative Hierarchical Clustering) is a clauster analysis method designed to group objects based on their characteristics. Clauster analysis (AHC) was performed using XLSTAT software with bray-curtis dissimilarity with unweighted pair-group average algorithm, resulting from a dendogram classification scheme based on field survey data by eliminating class cover that has a frequency of occurrence of less than 40% (Green et al., 2000). Bray-curtis is an index used to see the level of similarity or difference between the two components. This study used a dissimilarity index of 0.4 and produced 4-7 clauster or

grades. According to Sneath and Sokal (1973), clauster analysis can be calculated using the following formula:

$$\frac{1}{|A|.|B|}\sum_{x\in A}\sum_{y\in B} d(x,y) \dots\dots\dots\dots\dots\dots(1)$$

Where:
A ,B   : Set or clauster
d   : Average distance
x,y   : Pair in clauster

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a guided classification algorithm that is able to separate two classes by optimizing the margins between them and is able to work effectively on data types that have low texture resolution (Wahidin et al., 2015). This classification algorithm is classified as machine learning classification based on the linear classifer principle r (Nugroho et al., 2003). According to Tzotsos (2006), the Support Vector Machine (SVM) classification equation is as follows:

$$f(x): \quad \sum_{i\in s}\lambda_i\, y_i\, K(x_i x) + w_0 \dots\dots\dots\dots\dots\dots(2)$$

Where:
K   : Kernel functions
x_i : Training samples
Λ_i : Lagrange
S   : Part of the training sample corresponding to the non-zero Lagrange multiplier
w_0 : Parameter hyperplane

Minimum Distance Classification (MDC)

Another classification method used in this study is the Minimum Distance Classification (MDC) method. This method is guided classification, where the classification method is based on the shortest distance between the pixel brightness value and the average of the classified class. The brightness value is calculated by the average spectrum for each category and the pixel viewing distance you want to classify. In other words, the closest distance will be in the same category as the average value. Conversely, if the pixel distance is greater than the average, then the pixel is not classified (Mukhaiyar, 2010). Mathematically the characteristics of the categories can be seen by the following equation:

$$m_j = 1/N \sum_{x\in mj} X \dots\dots\dots\dots\dots\dots\dots\dots\dots(3)$$

Remarks: j = 1, 2, 3, ...M

If there is a new object with characteristic x, the calculation of the closest distance to the eucledian distance is:

$$D_j(x) = \|x - m_i\| \dots\dots\dots\dots\dots\dots\dots\dots\dots(4)$$

Remarks: j = 1, 2, 3, ...M

Accuracy Test

Accuracy tests are carried out on the classification results map with the aim of determining the level of accuracy of the classification technique applied. A commonly used accuracy test on classification result data is the confusion matrix (Congalton & Green 2009) and compares the image classification results as the basis of the actual category with some field data representing the research area (Siregar 2010). Accuracy categories include user accuracy (UA), producer accuracy (PA), kappa accuracy, and overall accuracy (OA).

According to SNI 7716:2011, states that the limit of the accuracy value is ≥60% for mapping various benthic habitats in shallow waters (LIPI, 2014).

Overall Accuracy (OA) = $\frac{\sum_{i=1}^{n} X_{ii}}{N} \times 100\%$ ....(5)

User Accuracy (UA) = $\frac{X_{ii}}{X_{+I}} \times 100\%$...(6)

Producer Accuracy (PA) =$\frac{X_{ii}}{X_{i+}} \times 100\%$...... ....(7)

Kappa accuracy = $\frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=}^{r} x_i + x_i}{N^2 - \sum_{i=}^{r} x_i + x_i} \times 100\%$ ....(8)

Where:
n   : Number of rows in the User Accuracy (UA) matrix
K   : Number of observations
Xii   : The number of observations in the i-th column and i-th row
X+i   : Number of observation in the i-th column
Xi+   : Number of i-row observation

RESEARCH RESULTS AND DISCUSSION
Classification Scheme

The classification scheme using clauster analysis produces data in the form of a dendogram and the central object value is generated from field data. Based on the clauster analysis, the results of the classification of benthic habitats in the waters of Pengudang Village were obtained as many as 6 classes, namely, class 1; seagrass mixed sand and rubble (coral faults), grade 2; seagrass mixed with sand, grade 3; dead corals mixed with sand, algae and live coral, grade 4; mixed sand of live coral and dead coral, grade 5; seagrass mixed with sand and live coral, grade 6; sand mixed live coral. The following is presented a dendogram image of the results of the Bray-curtis clauster analysis. (Fig. 3)
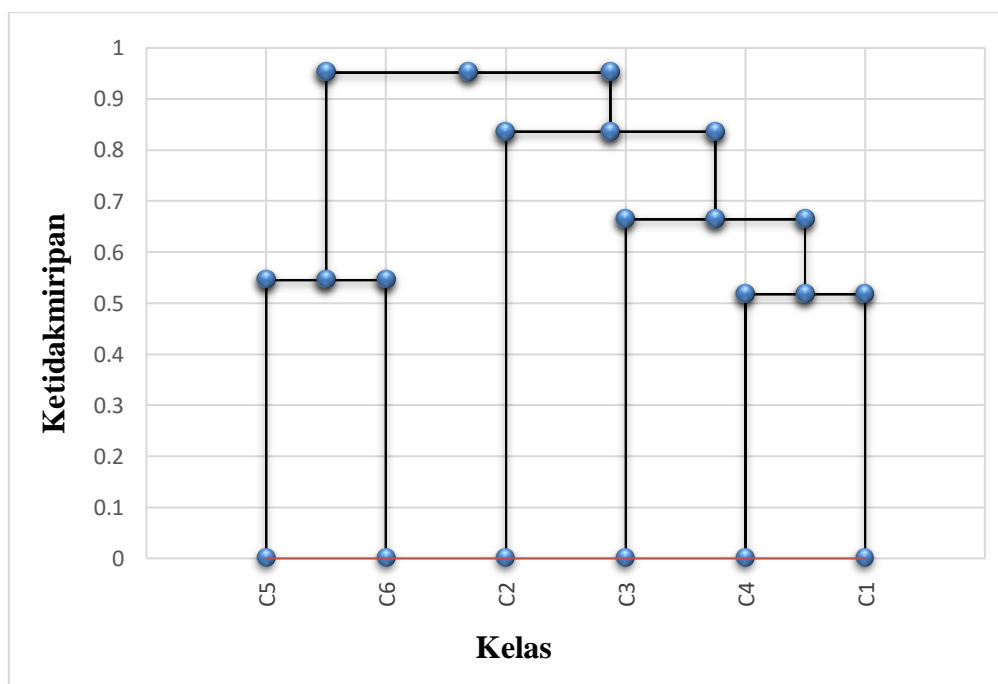

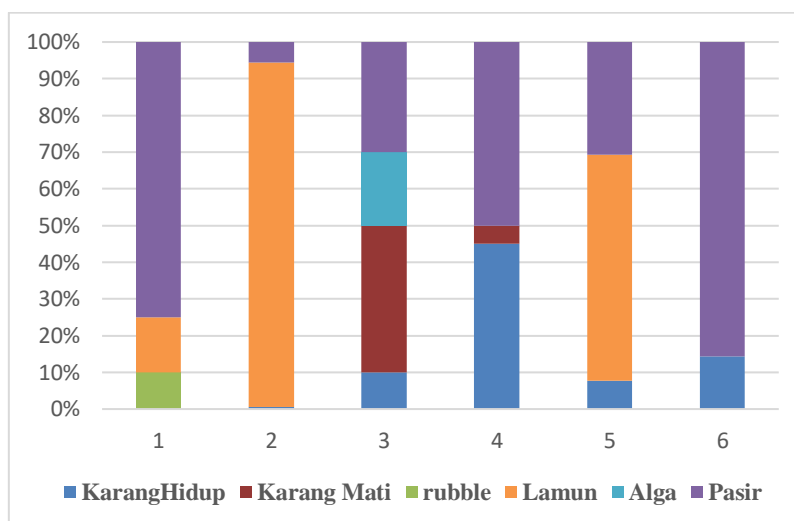Figure 3. Dendogram Of Bray-Curtis Clauster Analysis Results

Figure 4. Percentage Of Central Object Values That Make Up The Benthic Habitat Classification Scheme

Benthic Habitat Classification

Classification of benthic habitat in the waters Pengudang Village, Bintan Regency using Sentinel-2A with pixel-based image classification with two algorithms, namely SVM and MDC with thematic layer input from field data. The number of points used for classification is 70 points representing each class of benthic habitat cover. The following are the classification results using the support vector machine (SVM) algorithm model and minimum distance classification (MDC). (Fig. 5 and Fig. 6)
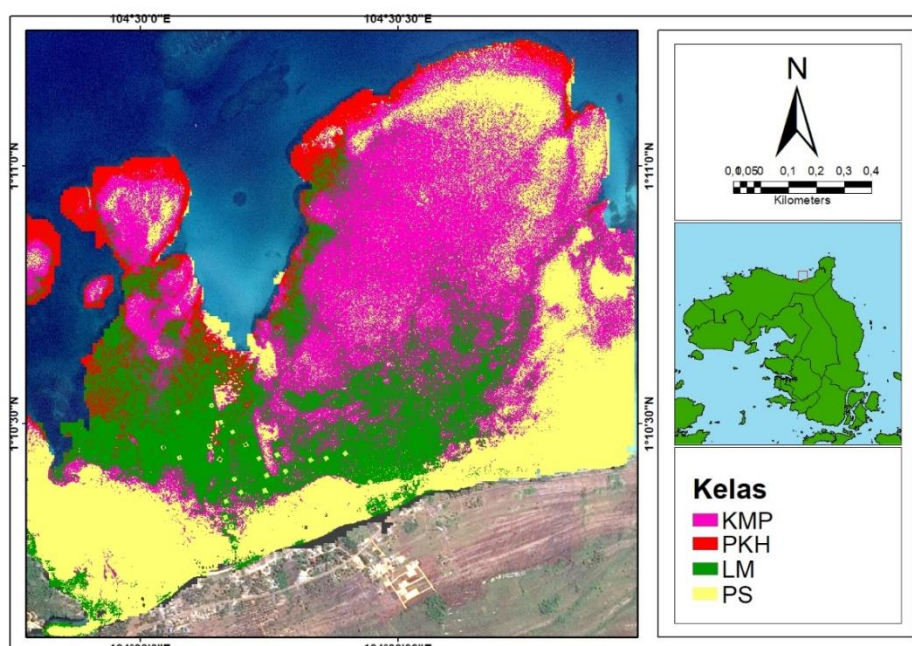

Figure 5. Results Of Sentinel-2A Image Classification With SVM Algorithm
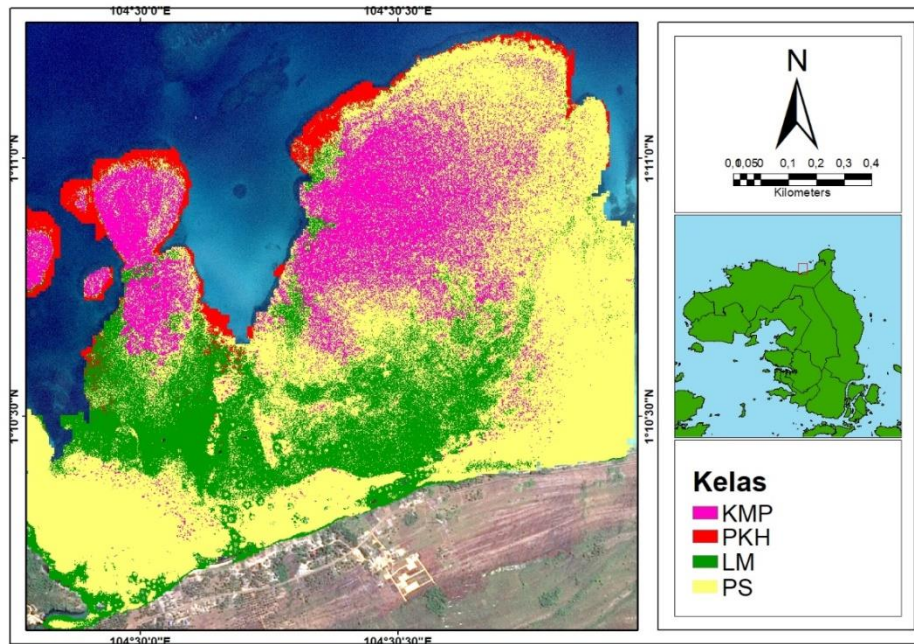
Figure 6. Results Of Sentinel-2A Image Classification With MDC Algorithm

Accuracy Test

The accuracy test in this research uses the confusion matrix method which compares the results of image classification as the basis for the actual class with the number of survey points used. Following are the accuracy of classification results using the SVM algorithm model (Table 3) and MDC (Table 4) presented in the table below.

Table 3. Accuracy Test Results Using SVM Algorithm

Habitat Class	Field					
	KMP	PS	PKH	LM	TOTAL	UA
KMP	10	1	1	0	12	83,33
PS	2	12	1	2	17	70,59
PKH	1	1	5	0	7	71,43
LM	2	1	0	11	14	78,57
TOTAL	15	15	7	13	50	
PA	66,67	80,00	71,43	84,62	OA=	76,00

Table 4. Accuracy Test Results Using MDC Algorithm

Habitat Class	Field					
	KMP	PS	PKH	LM	TOTAL	UA
KMP	7	1	1	1	10	70,00
PS	0	13	1	2	16	81,25
PKH	1	2	5	0	8	62,50
LM	4	3	1	8	16	50,00
TOTAL	12	19	8	11	50	
PA	58,33	68,42	62,50	72,73	OA=	66,00

Discussion
Classification Scheme

The classification scheme based on the results of field observations obtained using Bray-Curtis clauster analysis produces 6 classes, namely class 1; sand (75%) mixed with seagrass (15%) and rubble (15%), class 2; seagrass (93.81%) mixed with sand (6.19%), class 3; dead coral (40%) mixed with sand (30%) algae (20%) and live coral (10%), class 4; sand (50%) mixed with live coral (45%) and dead coral (5%), class 5; seagrass (60%) mixed with sand (30%) and live coral (10%), class 6; sand (90%) mixed with live coral (10%). Then simplification was carried out by naming the dominant classes from the 6 classes into 4 classification classes based on the central object value, namely class 1; sand (75%), class 2; seagrass (93.81%), class 3; dead coral (40%) mixed with sand (30%), class 4; sand (50%) mixed with live coral (45%). This simplification was carried out because there are several classes that have the same components of the benthic habitat classification scheme. The Bray-Curtis clauster analysis uses a dissimilarity of 0.4 or 40%, where each benthic habitat class has a similarity of 0.6 or 60% based on field data results. This use is adjusted to the location of the satellite imagery used and is caused by conditions and variations in different observation locations (Green et al., 2000).

Benthic Habitat Classification

Benthic habitat classification in this study uses pixel-based image classification with the application of two algorithms, namely SVM and MDC, which include supervised classification. In this study, 4 classes of benthic habitat were produced, namely sand (PS), seagrass (LM), dead coral mixed with sand (KMP), and sand mixed with live coral (PKH) using the SVM and MLH classifications, each of which has a different color. (Figure 6 and Figure 7). The classification results using the SVM and MLH algorithm models show that benthic habitat is distributed in the shallow waters of Pengudang Village. Based on the classification results using the SVM and MDC algorithms, the sand (PS) class is yellow, seagrass (LM) is green, dead coral mixed with sand (KMP) is pink which is spread throughout shallow waters, while the sand mixed with live coral (PKH) class is red can be found near the waters.

Accuracy Comparison

The estimation of classification accuracy is calculated using a confusion matrix, where the overall accuracy (OA) of the classification results using SVM and MDC algorithms results in an accuracy rate of 76% (Table 3) and 66% (Table 4) respectively. The accuracy of benthic habitat maps produced by pixel-based classification methods using the SVM algorithm is higher than the use of the MDC algorithm with a difference in accuracy values of 10%. The accuracy procedure (PA) of the SVM algorithm states that the seagrass class (LM) has the hishest accuracy rate of 84.62% and the dead coral mixed with sand (KMP) class has the lowest accuracy rate of 66.67%, while the accuracy procedure (PA) of the MDC algorithm states that the one with the highest accuracy level is the seagrass class (LM) which is 72.73% and the lowest is the dead coral mixed with sand (KMP) class which is 58.33%. User accuracy (UA) SVM algorithm can be seen that the dead coral mixed sand class (KMP) has the highest accuracy level of 83,33% and the lowest accuracy level is found in the sand class which is 70,59%, while the user accuracy (UA) of the MDC algorithm can be seen that the highest accuracy level is found in the sand class of 81.25% and the lowest accuracy level is found in the seagrass class (LM) of 50%. The SVM algorithm model used produced a kappa accuracy of 67.28% while the MDC algorithm model produced a kappa accuracy of 53.70%. This shows that the accuracy between field data and satellite imagery is good, based on research by

Landis and Koch (1997) which produces a kappa value of 0.762%. The accuracy results show that the use of Sentinel-2A imagery with the application of SVM algorithm for benthic habitat mapping has better results, and the accuracy of the resulting image can be used as spatial information on the distribution of benthic habitat in shallow waters of Pengudang Village. The test results meet SNI 7716:2011 standards of 60% of shallow water benthic habitat mapping and support previous studies that show the effectiveness of Sentinel-2A imagery in mapping benthic habitats through comparison with other satellite image such as the ability of Landsat 8 which has an accuracy of 47,57% (Muhammad Hafizt, 2017).

The results of research using the SVM algorithm are also evidenced by previous research conducted by (Rahmani, 2022), using object-based classification methods using Sentinel-2A imagery and UAV imagery, and resulted in accuracy of 73.95% and 81.51%. Mukrimin et al., (2021), also conducted research on object-based benthic habitat mapping using Sentinel-2A imagery with the application of SVM algorithm which resulted in an accuracy of 88%. In addition, other studies were also conducted using Sentinel-2A images with the application of MLC, MDC and MaDC algorithms which resulted in accuracy rates of 74.33%, 73.45%, and 78.35% respectively (La Ode Alifatri, 2022). The difference in the accuracy of benthic habitat mapping is caused by several factors such as differences in classification methods, the number of field observation points, the number of benthic habitat classes and the imagery used (Prabowo, 2018). In addition, the use of classification algorithms can also affect the results of accuracy tests. According to (Prabowo, 2018), other errors in accuracy such as errors in the classification scheme, data input errors, and reference data labeling errors.

CONCLUSION

Based on the results of research on mapping benthic habitat using Sentinel-2A imagery with the application of SVM and MDC algorithms in the waters of Pengudang Village with 4 classes of benthic habitat, namely seagrass class (LM), sand (PS), dead coral mixed with sand (KMP), and live coral mixed sand (PKH) resulted in overlay accuracy of 76% for the application of the SVM algorithm and 66% for the application of the MDC algorithm. Where the application of the SVM algorithm has the highest overlay accuracy of 10%, so it can be used as a reference in mapping benthic habitats in shallow waters.

BIBLIOGRAPHY

Alifatri, L. O., Bayu, P., & Kasih, A. 2022. Klasifikasi Habitat Bentik Berdasarkan Citra Sentinel-2 di Kepulauan Kei,. Jurnal Ilmu Pertanian Indonesia, 27(3), 372-384.

Congalton, R. G & Green, K. 2009. Menilai Keakuratan Prinsip Dan Praktik Data Penginderaan Jauh 2 Nd Ed. CRC Press Indiana. 201 hal.

Dekker, A. G, Brando, V. E., & Anstee, J. M. 2005. Deteksi Perubahan Lamun Retrospektif di Danau Australia Pasang Surut Pantai Dangkal. Penginderaan Jauh Lingkungan. 97: 415-433. https:// doi.org/10.1016/j.rse.2005.02.017.

Green, E. P., Mumby P. J., Edwards A. J., & Clark C. D. 2000. Remote Sensing Handbook for Topical Coastal Management. Paris (FR): UNESCO Pub.

Hafizt, M., Yulia, M., & Bayu, P. 2017. Kajian Metode Klasifikasi Citra Landsat-8 untuk Pemetaan Habitat Bentik di Kepulauan Padaido, Papua. Oseanologi dan Limnologi di Indonesia, 1-13.

Hedley, J., Roelfsema C., & Phinn S. R. 2009. Model Transfer Radiasi Yang Efisien dalam Versi Untuk Aplikasi Penginderaan Jauh. Penginderaan Jauh Lingkungan. 113: 2527-2532. https://doi.org/ 10.1016/j.rse.2009.07.008.

Inounu, I., Martindah, E., Saptati, R. A., & Priyanti, D. A. 2007. Potensi Ekosistem Pulau-Pulau Kecil dan Terluar untuk Pengembangan Usaha Sapi Potong. Wartazoa, 17(4), 156– 164.

Kohler K. E., & Gill S. M. 2006. Coral Point Count with Excel Extensions (Cpce): A Visual Basic Program for the Determination Of Coral And Substrate Coverage Using Random Point Count Methodology. Computers and Geosciences. 32: 1259-1269.

LIPI (Lembaga Ilmu Pengetahuan Indonesia). 2014. Panduan Teknis Pemetaan Habitat Dasar Perairan Laut Dangkal. Jakarta, Indonesia.

Mukrimin., LaOde. M. Y. H., & Amadhan, T. 2021. Pemetaan Habitat Bentik Perairan Dangkal di Pesisir Pulau Tiga (Selat Tiworo) Menggunakan Citra Satelit Sentinel-2A. Sapa Laut, 6(1), 63-74.

Nugroho, A. S., Witarto A. B., & Handoko, D. 2003. Support Vector Machine Teori dan Aplikasinya Dalam Bioinformatika. Ilmu Komputer.

Prabowo, N.W., Vicentius, P. S., & Syamsul, B. A. 2018. Klasifikasi Habitat Bentik Berbasis Objek dengan Algoritma Support Vector Machines dan Decision Tree Menggunakan Citra Multispektral Spot-7 di Pulau Harapan dan Pulau Kelapa. Jurnal Ilmu dan Teknologi Kelautan Tropis, 10, 123-134.

Prayuda, B "Panduan Teknis Pemetaan Habitat Dasar Laut Dangkal," Pusat Penelitian Oseanografi Pemetaan Dasar Laut Dangkal Lembaga Ilmu Pengetahuan Indonesia, CRITC COREMAP II LIPI, 2014

Rahmani, E., Wayan, G. A. K., & Dewa, N. N. P. 2022. Pemetaan Habitat Bentik Menggunakan Citra Sentinel-2A Dan Unmanned Aerial Vehicle (UAV) di Perairan Pemuteran, Bali. Journal Of Marine Research And Technology, 5, 29-39.

Roelfsema, C. M., & Phinn, S. R. 2008. Evaluating Eight Field And Remote Sensing Approaches For Mapping The Benthos Of Three Different Coral Reef Environments In Fiji. Pages 71500F-71500F.

Siregar, V. P. 2010. Pemetaan Substrat Dasar Perairan Dangkal Karang Congkak dan Lebar Kepulauan Seribu Menggunakan Citra Satelit Quick Bird. J. Ilmu dan Teknol. Kelaut. Trop. 2(1), 19-30.

Talitha, A. E. 2017. Pemetaan Bentik Habitat Perairan Dangkal Karang Bongkok dengan Metode OBIA Menggunakan Citra WorldView-2. [Skripsi]. Bogor (ID): Institut Pertanian Bogor.

Wahidin, N., Vicentius, P. S., Bisman, N., Indra, J., & Sam, W. 2015. Object-Based Image Analysis For Coral Reef Benthic Habitat Mapping With Several Classification Algorithms. Procedia Environmental Sciences, 222-227.