

# ASSESSMENT OF SEAGRASS PERCENT COVER AND WATER QUALITY USING UAV IMAGES AND FIELD MEASUREMENTS IN BOLINAO, PANGASINAN

M.K.M.R. Guerrero<sup>1</sup>, J.A.M Vivar<sup>1</sup>, R.V. Ramos<sup>1</sup>, A.M. Tamondong<sup>1</sup>

<sup>1</sup> Department of Geodetic Engineering, University of the Philippines- Diliman, Quezon City, Philippines - (mrguerrero1, jmvivar2, rvramos, amtamondong) @up.edu.ph

## Commission IV

**KEY WORDS:** Chlorophyll-a, Turbidity, Seagrass Cover, Support Vector Machine, UAV images, Geostatistical techniques, Morans I, Ordinary Least Squares

## ABSTRACT:

The sensitivity to changes in water quality inherent to seagrass communities makes them vital for determining the overall health of the coastal ecosystem. Numerous efforts including community-based coastal resource management, conservation and rehabilitation plans are currently undertaken to protect these marine species. In this study, the relationship of water quality parameters, specifically chlorophyll-a (chl-a) and turbidity, with seagrass percent cover is assessed quantitatively. Support Vector Machine, a pixel-based image classification method, is applied to determine seagrass and non-seagrass areas from the orthomosaic which yielded a 91.0369% accuracy. In-situ measurements of chl-a and turbidity are acquired using an infinity-CLW water quality sensor. Geostatistical techniques are utilized in this study to determine accurate surfaces for chl-a and turbidity. In two hundred interpolation tests for both chl-a and turbidity, Simple Kriging (Gaussian-model type and Smooth- neighborhood type) performs best with Mean Prediction equal to -0.1371 FTU and 0.0061 µg/L, Root Mean Square Standardized error equal to -0.0688 FTU and -0.0048 µg/L, RMS error of 8.7699 FTU and 1.8006 µg/L and Average Standard Error equal to 10.8360 FTU and 1.6726 µg/L. Zones are determined using fishnet tool and Moran's I to calculate for the seagrass percent cover. Ordinary Least Squares (OLS) is used as a regression analysis to quantify the relationship of seagrass percent cover and water quality parameters. The regression analysis result indicates that turbidity has an inverse relationship while chlorophyll-a has a direct relationship with seagrass percent cover.

## 1. INTRODUCTION

### 1.1 Background of the Study

The Philippines is one of the countries that has an abundance in seagrasses. Currently, the seagrasses found here are recorded to be of eighteen (18) species from three (3) families found in five hundred and twenty-nine (529) sites (Fortes, 2013). Seagrasses play ecological and economic roles in Bolinao in a way that it contributes to the uplifting of the residents (Montano, 2005). In line with this, there are various studies in Bolinao, Pangasinan that investigate the extent and cover of seagrasses as well as the water quality conditions in the coastal areas. In 2015, water quality parameters Chlorophyll-a (Chl-a) and turbidity were estimated in Santiago Island using remote sensing techniques applied on Landsat-8 OLI image (Dumalag, Villafior, 2015). In 2017, seagrass percent cover is determined in the same location based on aerial image interpretation and analysis (Dalagan, Manasan, 2017). The gap between these two studies gave way to studying the correlation between the seagrass percent cover and water quality.

A declaration called the "Bolinao Declaration- a seagrass charter" by Professor Miguel Fortes of the University of the Philippines Marine Science Institute (UP MSI) was created in order to encourage countries to consider the value of seagrasses. In this declaration, he mentioned that "environmentally sensitive seagrass areas should be given priority in technical cooperation and financial aid for sustainable coastal development (Fortes, 1998)." A community-based coastal resource management project that is hosted by UP MSI, UP College of Social Work and Community Development (CSWCD), and the Haribon Foundation aims to address the said issue.

In this study, the researchers will focus on determining the gap between the previous studies and provide an assessment on the relationship of seagrass percent cover and water quality parameters, specifically Chlorophyll-a and turbidity, in Bolinao, Pangasinan. Field techniques using an Unmanned Aerial Vehicle (UAV) and a water quality sensor called the infinity-CLW, shall be conducted on selected seagrass areas. It will also be helpful to the community-based coastal resource management project which has an objective of determining and evaluating appropriate coastal resource and environmental management strategies which will ensure a sustainable base of living resources in the coastal area (McManus, L., Ferrer, E., dela Cruz, L., & Cadavos, A., n.d.).

### 1.2 Research Objectives

The primary objective of this study is to determine the relationship of seagrass percent cover and specific water quality parameters (Chl-a and turbidity) in Bolinao, Pangasinan using Aerial Photogrammetry, Remote Sensing techniques and Geographic Information Systems (GIS) techniques. The following are the specific objectives of the study: (1) to extract seagrass percent cover from images acquired by an Unmanned Aerial Vehicle (UAV) using pixel-based classification and remote sensing (RS) techniques; (2) to estimate water quality parameters, specifically Chl-a and turbidity, using field measurements and geostatistical methods; and (3) to provide a quantitative assessment using regression analysis in determining the relationship of seagrass cover and water quality parameters- Chl-a and turbidity.

### 1.3 Significance of the Study

Seagrasses are important to the ecosystem as they are considered a "carbon sink". If possible losses of seagrass are monitored through the relationship of these variables, water quality

parameters and seagrass percent cover, this may serve as a basis for any future conservation plans. The study aims to relate water quality and seagrass percent cover in an area using RS, GIS and field techniques in order to understand how these water quality parameters affect the seagrass percent cover. The outputs of the study may be used further for the rehabilitation and coastal management plans in Bolinao, Pangasinan.

#### 1.4 Review of Related Literature

Seagrasses are being threatened by several factors such as physical disturbances from wind driven waves and storms, marine animals like skates and rays that destroy the seagrasses because of their foraging method, and various human activities that affect the seagrasses the most. It was estimated that about twenty percent of seagrass meadows have perished in the past century. (Reynolds, 2018) There are various ways to measure the percent cover of seagrasses in the study area. The methods range from diver observations to remotely sensed data such as satellite or airborne images. (Krause-Jensen, Quaresma, n.d.)

Aerial photogrammetry is the most common method used for determining seagrass percent cover. A well-achieved data for mapping seagrasses is when the image is acquired in clear, shallow water. Turbidity or deep water can affect the interpretation of the data. Aerial photography can provide higher-resolution of images, but sun-glint reflection can occur and may cause problems for data interpretation. (Krause-Jensen, Quaresma, n.d.) Mapping of seagrasses by remote sensing relies on the fact that information regarding bottom features shows up as variations in the radiance directed towards the sensor (Krause-Jensen, Quaresma, n.d.). Pixel-based classification, most common method used in classifying images in remote sensing, uses spectral response pattern. (Matinfar et al., 2007). One of the two types of pixel-based classification is the supervised classification. It is a human guided classification which lets the user select pixels and assign it to different features (Geospatial Technology, 2015). An example of algorithm for supervised classification is the Support Vector Machine. It finds the ideal or most suitable line or hyperplane that divides the pixels into two classes. The goal is to maximize the clearance distance from the points that lies nearest to the dividing line, also known as support vectors (Avjan H., 2018) A study conducted by Huang Davis and Townshend entitled “An assessment of support vector machine for land cover classification”, concluded that SVM provides more accurate results than Maximum Likelihood Classifier, Neural Network Classifiers and Decision Tree Classifiers.

There are general variables that can be used to assess water quality, however, the only parameters that affect the growth of seagrasses are temperature, dissolved oxygen, salinity, turbidity and chlorophyll-a (Fitzhugh L., 2012) Since one of the objective of this study is to determine water quality parameters which can be directly derived from in-situ measurements, only chl-a and turbidity are measured. Geostatistics analyse and predict value using statistics linked with spatial or spatiotemporal phenomena (ESRI, n.d.). It is used for accurately interpolating values where there is no data and have greater confidence. Geostatistics provides basic parametric statistics and fast interpolation. Types of kriging provides an interpolation that is optimal to the area around a sample point location. (Robertson, 2008) A study made by John Zirschky entitled Geostatistics for environmental monitoring and survey design uses kriging, to produce minimum and unbiased variance estimates of the spatial distribution of a pollutant. Another study by Khalil, Orarda and St. Hilaire uses

geostatistical techniques for optimizing water quality monitoring in large lakes.

Regression analysis is a technique that investigates the relationship between a dependent and an independent variable and usually used for forecasting, analysis of causal effect relationship between variables and time series modelling (Ray, 2015). This statistical technique obtains the best fit line or curve given the set of data points (Daniel, 1999). Equation 1 shows the general formula for a regression analysis

$$y = a + bx \quad (1)$$

where  $y$  = dependent variable  
 $x$  = independent variable  
 $a$  = intercept of the line  
 $b$  = slope of the line

In this study, Ordinary Least Squares (OLS) regression was used. It is a statistical method that estimates the relationship of the independent and dependent variable by minimizing the sum of the squares of the difference between the observed and predicted values of the dependent variable configured as a straight line (Kotz S., 2004). It is also the proper starting point of spatial regression analysis and creates an equation that represents the model being analysed (ESRI, n.d.)

## 2. METHODOLOGY

### 2.1 Data Gathering

Data gathering involves these activities: (a) measurement of Chlorophyll-a and turbidity at sample points on the study area using the infinity-CLW water quality sensor and (b) UAV surveys, simultaneous with water quality measurements, using a DJI Phantom 4 Pro quadcopter drone attached with an optical camera that has RGB bands. Water quality parameters are usually measured in the field using sensors, one of which is the infinity-CLW. It is a deployable instrument that measure chlorophyll-a and turbidity. It has sensors that measure the amount of Chl-a, turbidity and temperature. The instrument provides high accuracy and stable measurements of the parameters in oceans, rivers, and freshwater. The DJI Phantom 4 Pro quadcopter which was used to capture images has a  $\pm 0.5$  m vertical accuracy and  $\pm 1.5$  m horizontal accuracy (Aerial Guide, 2018). The drone was operated using a remote controller that is connected to a smartphone or a tablet via a USB cable. The images are stored in a micro SD card. The coordinates were measured using the GNSS Status Phone Application. It provides the user with GNSS Status information through its configuration of Trimble R1, R2 and PG200 GNSS receivers (App Advice, 2019).

On the day of the fieldwork, there were no Ground Control Points (GCPs) established for georeferencing the images because the drone has a built-in GPS that determines the location of the photos. Simultaneous with the UAV survey, the coordinates of points where the measurements of the water quality parameters were noted. Twenty-five (25) random samples from the study area were taken within the extent of the UAV survey.

#### 2.1.1 Study Area

The study area is located on the north-eastern part of Siapar Island in Bolinao, Pangasinan as shown in Figure 1. This was chosen due to the varying seagrass distribution and the quality of water present in the area.

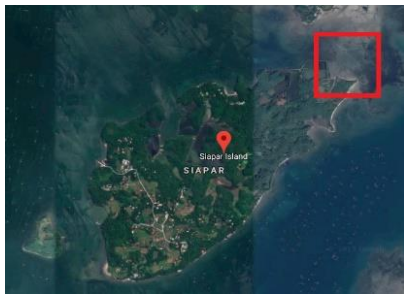


Figure 1: Area of study in Siapar Island as highlighted by the red box (Google Maps, 2019)

### 2.1.2 Fieldwork Preparation

The data gathering was conducted on February 10, 2019, 8:00 am, in Bolinao, Pangasinan particularly on the north-eastern part of Siapar Island. Fieldwork preparations were made prior the date of data gathering. These preparations include the creation of flight plan using the Pix4D Mapper phone application, setting-up of the infinity-CLW water sensor for delayed launch, and determination of tide levels & weather forecast on the set fieldwork date. Since seagrasses are submerged in water, the UAV survey was conducted during low tide in order to capture maximum surface objects during flights. The flight plan prepared can be seen in Figure 2 below.

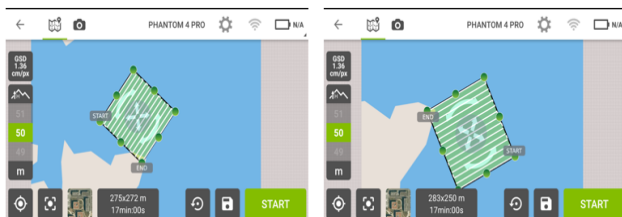


Figure 2: Flight plans 1 and 2 for the north eastern part of Santiago Island

## 2.2 Pre-processing

### 2.2.1 UAV Images

The UAV derived images were pre-processed using Agisoft Photoscan Professional. The photos were processed according to the workflow. It starts with Align Photos which includes matching points between overlapping images. The next step is Build Dense Cloud that calculates depth information through estimated camera positions. A polygonal mesh model which was based on the dense point cloud was then built. Building texture is not necessary in orthomosaic photo but it helps in precise marker placement. Next is Build Tiled Model that bases its process on the dense cloud or mesh model produced. The last step is Build Orthomosaic that produces an orthoimage which can be exported to JPEG/TIFF/ PNG. (Agisoft, n.d.) All steps were set to high accuracy.

### 2.2.2 Water Quality Parameters

The data acquired using the sensor were sorted out and data outliers were removed. Data outliers were identified by looking at the values within the 2 minutes that the sensor was submerged in water. It can be seen during that time that measurements have similar values and extreme values were evident.

## 2.3 Data Extraction

Data extraction involves a) determination of seagrass percent cover from UAV derived images through ENVI software using SVM and b) interpolation of water quality parameters, acquired from fieldwork using the Geostatistical Wizard in ArcGIS.

### 2.3.1 Determination of Seagrass Percent Cover

The pre-processed drone images were stitched together to form the whole image of the study area. SVM was used in classifying the image. Two types of Regions of Interest (ROIs) were made: the train ROI and the check ROI. Both contains two land cover types: seagrass and water with no seagrass. The ROIs created should achieve a separability values greater than 1.9. This was checked using the compute ROI Separability tool. The SVM has four different types of kernel: linear, polynomial, Radial Basis Function (RBF) and sigmoid. However, the rule of thumb is that linear SVMs can only be used in solving linear problems, and for non-linear problems, RBF will be used. RBF is the most commonly used kernel type in SVM. There are also parameters for the SVM to work successfully: gamma and penalty. Gamma controls different transformations and its default value is the inverse of the band numbers. Higher gamma value will be dependent on the points that is near the line. Lower value considers the points that is far from the line. Penalty parameter is the trade-off between classifying the data correctly and smooth decision boundary. Higher penalty parameters mean that the training points will be classified correctly but the test set will not be generalized well. Gamma and penalty affect the accuracy of the classification. The numbers that were assigned in gamma and penalty were chosen after several iterations. An accuracy assessment was then checked using the Confusion Matrix using Ground Truth ROIs tool. The train and check datasets were matched and assessed through Confusion Matrix using Ground Truth ROIs tool. A confusion matrix report, which included both the user's and producer's accuracies were checked and made sure that it has et the 85% threshold.

The seagrass percent cover was extracted by creating a fishnet covering the boundaries of the study area where water quality parameter samples were also taken. The proper zonation was based on the results of Moran's I which indicated the spatial autocorrelation of the data being analysed. There were 8 different number of zones tested in order to see which will be more appropriate for the data acquired. In order to determine the zone that will be valid for OLS, Moran's I should result to a random data. Clustered and dispersed data mean that the data being measured is spatially autocorrelated, thus violating the assumption in statistics that the data should be independent from one another. The tabulate intersection tool was used to intersect the input zone feature (fishnet) and the input class feature (classified drone image of the study area) to generate the area and seagrass percent cover of each zone.

### 2.3.2 Interpolation of Water Quality Parameters

On the other hand, the pre-processed water quality parameters, as seen in Table 1, measured were interpolated using the Geostatistical Wizard. To determine the value of chlorophyll-a and turbidity on different points within the study area an interpolation is needed. Interpolation is a process wherein limited number of sample points are used in order to predict the values for cells in a raster of unknown values for any geographic data such as elevation, noise levels, chemical concentrations and a lot

more (ESRI, n.d.). Interpolation follows the assumption that things that are near each other tend to have similar characteristics (ESRI, n.d.). There are various interpolation methods that can be used from the ArcGIS software. In this study, the Geostatistical Wizard was used to find the best interpolation method for the water quality parameters-chlorophyll-a and turbidity.

Point	Latitude	Longitude	Chlorophyll-a	Turbidity
1	16.36553359	119.964899	2.496725664	40.53137255
2	16.36568142	119.9649349	1.911525424	8.448222222
3	16.36576836	119.9649528	1.480756303	5.682666667
4	16.36573954	119.9649315	2.597244898	7.708023256
5	16.365843	119.9649268	1.806226415	7.65752381
6	16.36587507	119.9649703	2.599313725	4.07688172
7	16.36597698	119.9649776	2.3206	8.142989691
8	16.36603382	119.9650105	4.759746835	9.331538462
9	16.366134	119.9649911	5.431774194	6.743703704
10	16.36621361	119.9650174	3.547181818	8.242134831
11	16.36631843	119.9650535	8.875	6.606363636
12	16.36644563	119.9650951	1.759038462	2.761176471
13	16.36652608	119.965121	6.445365854	10.66134328
14	16.36661593	119.9651177	2.2183	2.242688172
15	16.36659734	119.9650042	5.215617978	7.879090909
16	16.36661836	119.9649183	3.173303571	5.17525641
17	16.36647749	119.9649741	2.614489796	5.470243902
18	16.36644473	119.9648988	3.564606742	5.049
19	16.36634104	119.964819	7.927763158	12.33977273
20	16.36618214	119.9647274	2.841894737	2.837303371
21	16.36608516	119.9646858	3.873240741	7.701375
22	16.36583449	119.9646755	2.215833333	13.8152381
23	16.36573276	119.9644193	2.349210526	20.97864198
24	16.3655901	119.9645614	1.752019231	21.33084746
25	16.36550766	119.9647495	2.761083333	56.59196078

Table 1: Turbidity and Chlorophyll-A values and their corresponding coordinates

The data was examined using the explore data in Geostatistical Analyst. Before proceeding to interpolation using Simple and Ordinary Kriging, the data should be normally distributed. Since the data is not normally distributed, log transformation was applied. After examining the data, the next step is calculating the empirical semi variogram. Ordinary Kriging and Simple Kriging were tested, and several iterations were made to determine what combination of parameters were suitable for producing the best interpolation method with little errors possible. Ordinary Kriging assumes that the unknown mean is constant along the neighborhood in each point. On the other hand, Simple Kriging has an assumption that the known mean is constant (ESRI, n.d.). In choosing the model that is best-fit for the semi-variance calculated by the software, different models were considered; Stable, Circular, Exponential, Gaussian and Linear. These models influence the prediction of the unknown values. Another parameter that influences the prediction of the unknown value is the shape of the curve near the origin. A steeper curve entails a stronger influence of the closest neighbor with the predictions. Other parameters which are used also to describe the models are range, sill, and nugget. These parameters are found on the semi variogram graph that measures the strength of statistical correlation as a function of distance. Iterations of different combinations of these parameters were tested. To help create matrices in determining the kriging weights, the search neighborhood types- Standard and Smooth- were also tested. Different parameters for each type were also iterated together with the combinations of the parameters in the semi variogram. For Standard Neighborhood type, the parameters iterated are Sector type and Maximum & Minimum Neighborhood. On the other hand, only the Smoothing Factor was iterated for Smooth Neighborhood type.

Among the iterations done, the interpolation method with the least possible errors and errors that satisfy the given criteria were used for predicting water quality parameter values for both chlorophyll-a and turbidity of points which were not measured using the infinity-CLW water sensor. The mean prediction error should be close to 0 so that the predicted values are unbiased. Root-mean-square standardized error should be close to 1 for standard errors to be accurate. RMSE and ASE should be small so not to deviate from the measured values. These errors could be interpreted using the output- reports and/or maps- of the method used. Through these outputs, it can be checked if the model has produced reasonable results for the predictions and uncertainties (ArcGIS, n.d.).

The results of interpolation for chlorophyll-a and turbidity were used as basis for the value layer for further analysis using the Zonal Statistics tool in the ArcGIS software. Zonal Statistics calculates statistic for each zone delineated by the zone dataset from a value raster dataset (ArcGIS Pro, n.d.). The Zone Statistics as Table tool does the same work as the Zonal Statistics tool but provides the output in a table format (ArcGIS Pro, n.d.). This tool produces a different output type such as the mean, median, mode, maximum value, minimum value, minority, majority, range, standard deviation, sum and variety. However, only the mean value was used for analysis, as this represents the mean value of the water quality parameter being measured per zone. The zone layer used for this tool was the same fishnet used for extracting the seagrass percent cover.

## 2.4 Data Analysis

### 2.4.1 Data Analysis using OLS

Data analysis uses the extracted data from the previous step and involves regression analysis method for the correlation of seagrass percent cover and water quality parameters (chlorophyll-a and turbidity). A feature class containing the dependent and explanatory variables for analysis was determined. The tables produced from Zonal Statistics of both water quality parameters were joined with the layer of seagrass percent cover having the Object ID and/or FID as the unique ID field for basis of intersection. The fishnet grid size, which was determined based on the spatial autocorrelation of the data using Moran's I, was used to define the zones within the study area. The output of the OLS regression tool is a map of the residuals of the data being analysed and a report containing the statistical reports pertaining to model performance, model significance, and explanatory variables. Moreover, the validity of the OLS depends on the spatial autocorrelation of the area and the criteria discussed in the next sub-chapter.

### 2.4.2 Validation of OLS result using Set Criteria and Spatial Autocorrelation Tool

One way to know the spatial autocorrelation is by using the Moran's I. The Moran's I tool produces five values; Moran's Index, Expected Index, Variance, p-value and z-value. The null hypothesis of this spatial autocorrelation is that the data is randomly distributed. In order to not reject the null hypothesis, the p-value must be greater than the assigned level of significance which is dependent on the z-value. If the null hypothesis is not rejected, OLS is valid. The zone which satisfies the criteria would be used in the data analysis. The criteria are as follows: a. Maximum Variance Inflation Factor (VIF) Value < 7.50, b. Minimum Jarque Bera p-value > 0.10, c. Minimum Spatial Autocorrelation p-value > 0.10 (ArcGIS Pro, n.d.). The VIF

value represents multicollinearity or the redundancy willing to tolerate among the explanatory variables. Higher VIF value can make the model unstable. For Jarque-Bera p value, it tells if the residual of the model is normally distributed. Smaller p-value means that the model is not normally distributed thus making the model bias. Moreover, spatial autocorrelation pvalue checks the model residuals for spatial clustering using Moran's I. Smaller p-value means misspecification or there are missing explanatory variables. (ESRI, n.d.)

### 3. RESULTS AND DISCUSSION

#### 3.1 Data Acquisition

Upon completion of the fieldwork, the acquired UAV images were then processed using Agisoft Photoscan. Two UAV flights on Siapar island were processed and Table 2 shows information regarding the processing. There were no corrections-atmospheric, radiometric and sun glint applied in the image because the drone was only set to obtain image at a low altitude and the only available band is RGB. The spatial resolution of the orthomosaic image is 0.08 m x 0.08 m. The built-up found in the images were masked out since it was not needed in the study.

No. of Images	284
Flying Altitude	209 m
Coverage Area	0.0562 km <sup>2</sup>
Focal Length	8.8 mm
Spatial Resolution of the Orthomosaic	0.079324 x 0.079324

Table 2: Details of the first UAV flight

#### 3.2 Data Extraction

##### 3.2.1 Determination of Seagrass Percent Cover

The stitched image was then classified using ENVI software to differentiate the seagrass from non-seagrass elements. Below is the result of the classified image using SVM. The type of Kernel used was Radial Basis Function. Two parameters were also set-Gamma in Kernel Function and Penalty Parameter. After several iterations, the values used for the parameters were determined as follows: Gamma= 0.33 and Penalty Parameter= 100.

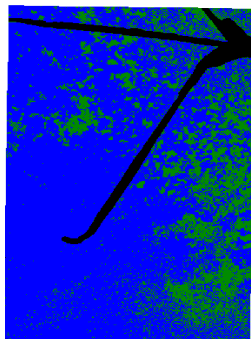


Figure 3: Classified Stitched Image

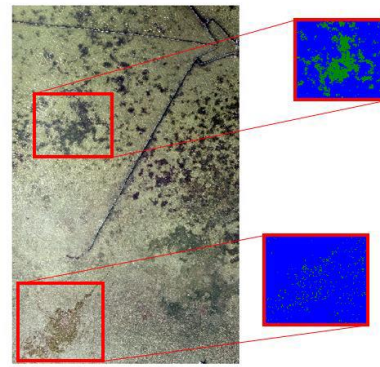


Figure 4: Comparison of Stitched and Classified Stitched Image

The Regions of Interest (ROIs) were created in order to classify the image accordingly. Four sets of ROIs were created, half of which are train datasets and the other half are check datasets. Separability of the ROIs were also constantly monitored to make sure that the classification process was still within the allowed value of separability which is more than 1.9. The overall accuracy of the classification was also checked through the confusion matrix tool. The figure below shows the overall accuracy, kappa coefficient, user and producer's accuracy of the classification which are all within the allowable value which is 85 percent.

Accuracy Assessment		
Overall Accuracy	91.0369%	
Kappa Coefficient	0.8210	
Commission and Omission		
Class	Commission (%)	Omission (%)
Seagrass	15.32	0.00
Non-Seagrass	0.00	17.77
Producer and User Accuracy		
Class	Producer Accuracy (%)	User Accuracy (%)
Seagrass	100.00	84.68
Non-Seagrass	82.23	100.00

Table 3: Accuracy Assessment

The classified image was then processed in ArcMap to determine the seagrass percent cover. The fishnet tool was used in order to create a grid for determining the percent cover of seagrass for each area containing the sampled point. The fishnet was used to determine the zones for seagrass percent cover and water quality parameters. There were eight (8) different fishnets- 9, 16, 25, 36, 49, 64, 81, 100 tried in determining what number of zones would be applicable in this study. Moran's I was run to help determine the number of zones. Only 9 and 16 zones appeared to have random data, which means that the seagrass percent cover in each grid are independent from each other.

No. of Zones	Dispersed/Random/Clustered	Moran's Index, Z-Value and P-Value
9	Random	Moran's Index: -0.127487
		Z-Value: -0.006407
		P-Value: 0.994888
16	Random	Moran's Index: 0.238203
		Z-Value: 1.099304
		P-Value: 0.271636
25	Clustered	Moran's Index: 0.442581
		Z-Value: 2.211897
		P-Value: 0.026974
36	Clustered	Moran's Index: 0.558320
		Z-Value: 3.262342
		P-Value: 0.001105

49	Clustered	Moran's Index: 0.654057 Z-Value: 4.420846 P-Value: 0.000010
64	Clustered	Moran's Index: 0.678997 Z-Value: 5.245772 P-Value: 0.000000
81	Clustered	Moran's Index: 0.747318 Z-Value: 6.50118 P-Value: 0.000000
100	Clustered	Moran's Index: 0.770333 Z-Value: 7.454159 P-Value: 0.000000

Table 4: Zonation and their corresponding Moran's I result

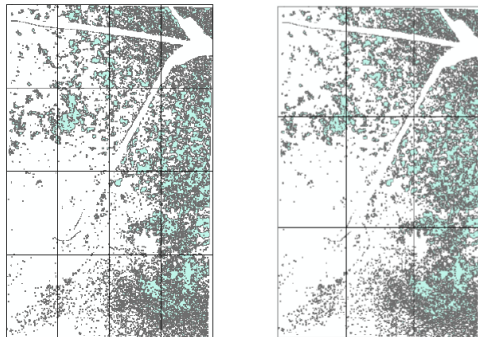
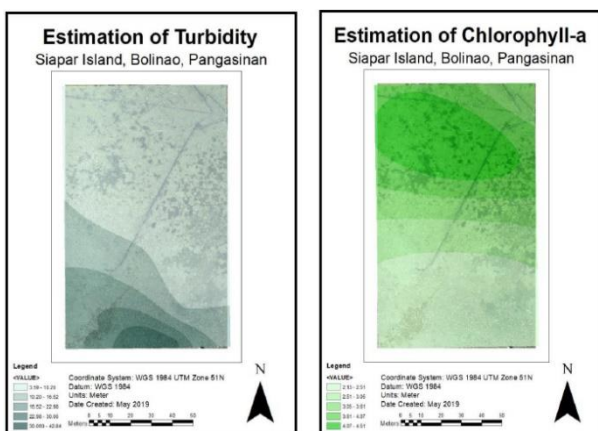


Figure 5: Classified Image after using the Fishnet Tool in ArcMap

### 3.2.2 Interpolation of Water Quality Parameters

Turbidity and chlorophyll-a were interpolated using the Geostatistical Wizard in ArcMap. The criteria for choosing the best model for interpolating turbidity and chlorophyll-a are as follows: a) Mean Prediction error is close to zero, b) Root Mean Square Standardized prediction error is close to 1, and c) Root Mean Square error and Average Standard Error are small (Sharma, P., Vijay, R., & MP, P., 2015). From two hundred iterations and evaluations based on the criteria set, below are the best interpolation methods for chlorophyll-a and turbidity.



Figures 6 and 7: Maps of the Estimated Turbidity and Chl-a in Siapar Island

It can be observed on Figure 6 that higher turbidity values can be found near the shore ranging from 30.05 to 42.84. Based on the seagrass percent cover shown on Figure 5, the area with high turbidity values only contain patches of seagrass. During

fieldwork, it was also observed that seagrasses along this area were withered. Moreover, numerous dead clams and trash were found in this area which might have contributed to the higher values of turbidity.

Prediction Errors	
Mean	-0.137104079
Root-Mean Square	8.76989783
Mean Standardized	-0.068814929
Root-Mean Square Standardized	0.977045747
Ave. Standard Error	10.83597938

Table 5: Values of the Prediction Errors of the Best Surface for Turbidity

From the iterations made, the best interpolation for Turbidity was Simple Kriging with Gaussian as the model type and Smooth as the neighborhood type. It yielded the best set of prediction errors based on the criteria that was set as seen in Table 5. The high values of Root Mean Square error and Average Standard Error were observed on almost all the iterations done and this was due to the extreme values of turbidity measured near the shore of the island. As discussed earlier, there were withered seagrasses and trash in this area which may have contributed to the turbidity level of the water.

Lower values of chlorophyll-a were observed in the area where the higher values of turbidity were located. The same reason, as stated on the discussion of higher values of turbidity on the same area, can be inferred regarding this. The withered seagrasses in the area may have produced less chlorophyll-a than those found on areas with less turbid waters. Since chlorophyll-a is an indication of light stress as stated in the RRL section of this study, the seagrasses in this area might have received less light hence becoming withered.

Prediction Errors	
Mean	-0.006088516
Root-Mean Square	1.800581172
Mean Standardized	-0.004764967
Root-Mean Square Standardized	0.964018174
Ave. Standard Error	1.672616305

Table 6: Values of the Prediction Errors of the Best Surface for Chlorophyll-a

From the iterations made, the best interpolation model for Chlorophyll-a was Simple Kriging with Gaussian as the model type and Smooth as the neighborhood type. This was the same type as the turbidity model but the values are different. The set of prediction errors from this model was the best fit according to the criteria that was set, as seen in Table 6. Compared to the error produced in turbidity, chlorophyll-a has smaller root mean square and average standard error because all the values range from 1-8 there were no higher values like on turbidity.

### 3.3 Data Analysis

#### 3.3.1 Validation of OLS result using Set Criteria and Spatial Autocorrelation Tool

Out of the eight number of zones, only two were classified as randomly distributed hence only two were analysed using OLS. In order to choose which of the two is more appropriate, the criteria mentioned in the methodology was followed. Based on the report generated from OLS, the Variance Inflation Factor (VIF) indicates the redundancy among explanatory variables, the

Jarque- Bera p-value tells whether the model prediction was biased, and the Spatial Autocorrelation p-value determines whether the model has an overcounting type of bias.

Criteria	9 zones	16 zones
Max VIF Value < 7.50	1.752804 (Passing Model)	2.038740 (Passing Model)
Min Jarque-Bera p-value > 0.10	0.997614 (Passing Model)	1.529898 (Passing Model)
Min Spatial Autocorrelation p-value > 0.10	0.321118 (Passing Model)	0.002572 (Model Failed)

Table 7: Comparison of Nine and Sixteen Zones

No. of Zones	Equation of Regression
9	$y = -0.106502 * Turbidity + 3.813341 * Chlorophyll - a + 19.123618$
16	$y = -0.008097 * Turbidity + 6.378128 * Chlorophyll - a + 16.610613$

Table 8: Equations of Regression of the Two Zonations

Based on the Table 7, only nine zones have passed all of the criteria. Therefore, this zonation will be used to further analyze the relationship between the seagrass percent cover and the water quality parameters- Turbidity and Chlorophyll-a. The regression formula derived from this zonation is:

$$Seagrass\ Percentage = -0.106502 * Turbidity + 3.813341 * Chlorophyll - a + 19.123618 \quad (2)$$

The coefficients of regression as shown in Table 8 explains the strength and type of relationship that the explanatory variable has with the dependent variable. Turbidity has a negative relationship with seagrass percent cover, therefore higher value of turbidity would lessen the percentage of seagrasses in an area. It will continue to lessen if there is a lower value of chlorophyll-a since it has a positive relationship with the seagrass percent cover. This type of relationship between the dependent and independent variables can also be seen on the maps of the estimation of turbidity and chlorophyll-a on Figures 6 and 7.

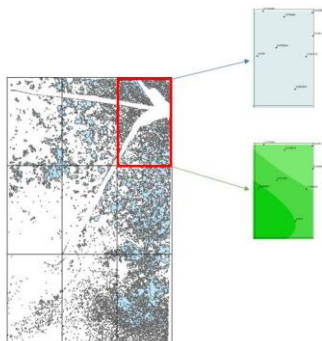


Figure 8: Application of the equation produced in OLS

The values gathered in the field within the chosen grid were averaged and substituted to the equation produced in OLS. There are 8 sample points within the chosen grid, the average value for turbidity and chlorophyll-a are 5.73064525 and 4.233215375 respectively. The computed seagrass cover is 34.66% while the seagrass percent cover from the tabulate intersection tool is 31.79%. There is a deviation of 2.87%, this can be caused by the error produced in the interpolation especially in the interpolation of turbidity because of the large difference of values. The y-intercept in the equation does not necessarily mean that if the

explanatory variables are 0, the seagrass percent cover is 19.12%. In reality, it is impossible to have 0 value of turbidity and chlorophyll-a, thus the y-intercept has no intrinsic meaning.

The validity of the OLS can be based on the result of Moran's I, a spatial autocorrelation tool. It always returns five statistical values. Z-value and P-value measures the statistical level of significance of the model. Positive value of Moran's I indicates the tendency towards clustering while negative value towards dispersing. The other values are Variance and Expected Index.

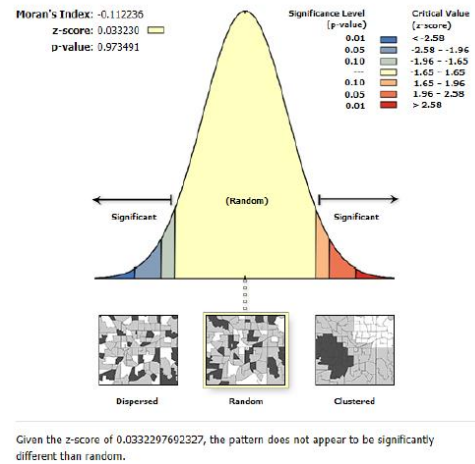


Figure 9: Spatial Autocorrelation

The z-value of the model is 0.033230 which falls under the mid-range of the level of significance, implying that the model used in OLS is valid.

#### 4. CONCLUSION AND RECOMMENDATIONS

##### 4.1 Conclusion

There are three main objectives to this study: a.) To extract seagrass percent cover from images acquired by an Unmanned Aerial Vehicle (UAV) using pixel-based classification and remote sensing (RS) techniques, b.) To estimate water quality parameters, specifically Chlorophyll-a (Chl-a) and turbidity, using field measurements, c.) To provide a quantitative assessment using statistical methods in determining the relationship of seagrass cover and water quality parameters-chlorophyll-a and turbidity. Based on the results on chapter 4, all objectives were met. The seagrass percent cover was extracted using Support Vector Machine which yielded an accuracy of 91.0369% after several iterations. The water quality parameters-chlorophyll-a and turbidity were also obtained using the interpolation models that satisfy the criteria for choosing the most appropriate interpolation model for the data. Lastly, the relationship of the dependent variable (seagrass percent cover) and independent variables (chlorophyll-a and turbidity) were assessed using OLS and Moran's I. The relationship was analysed using the equation:

$$Seagrass\ Percentage = -0.106502 * Turbidity + 3.813341 * Chlorophyll - a + 19.123618 \quad (3)$$

It can be concluded from the equation that turbidity has an inverse relationship with the seagrass percent cover. It means that every unit increase in turbidity could lessen the seagrass percent cover. On the contrary, chlorophyll-a has a direct relationship with seagrass percent cover. A unit increase in chlorophyll-a, could increase seagrass percent cover.

## 4.2 Recommendations

The study can further be enhanced if the results of the other study area, which is the sandbar near Santiago Island, can be analysed because the quality of water surrounding the sandbar is pristine compared to the one in Siapar island. The water quality condition on the said areas can influence the water quality parameters measured by the infinity-CLW.

Also, the use of an Unmanned Surface Vehicle (USV) can increase the number of water quality parameters measured by the infinity-CLW. Lastly, the iteration for choosing the best interpolation for the water quality parameters, can also be improved if more exploration was done to exhaust all the possible combinations of the parameters in the interpolation method used.

## REFERENCES

- Aerial Guide, 2018. Phantom 4 Pro Review, Specs & Price. Retrieved from <https://aerial-guide.com/article/phantom-4-pro-review-specs-price>
- Agisoft, 2016. Agisoft PhotoScan User Manual Professional Edition, Version 1.2. Retrieved from [https://www.agisoft.com/pdf/photoscan-pro\\_1\\_2\\_en.pdf](https://www.agisoft.com/pdf/photoscan-pro_1_2_en.pdf)
- App Advice, 2019. GNSS Status by Trimble Inc. Retrieved from <https://appadvice.com/app/gnss-status/932189198>
- ArcGIS Pro, n.d. How Zonal Statistics works. Retrieved from Desktop: <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/h-how-zonal-statistics-works.htm>
- Avjyan, H., 2018. Overview of Supervised Learning model SVM (support vector machines). Retrieved from Towards Data Science: <https://towardsdatascience.com/overview-of-supervised-learning-model-svm-support-vector-machines-20b683a4eaf>
- Daniel, W. W., 1999. Regression and Correlation. Kean University. Retrieved from <https://www.kean.edu/~fosborne/bstat/09rc.html>
- Dalagan, G., Manasan, L., 2017. Extraction of Seagrass Biophysical Parameters using Unmanned Aerial Systems. University of the Philippines Diliman.
- Dumalag, J.B., Villaflor, D.A., 2015. Estimation of Chlorophyll-a and Turbidity in Bolinao, Pangasinan Using Landsat 8 OLI Image. University of the Philippines Diliman.
- ESRI, n.d.. Get started with Geostatistical Analyst in ArcGIS Pro. Retrieved from ArcGIS Pro: <https://pro.arcgis.com/en/pro-app/help/analysis/geostatistical-analyst/get-started-with-geostatistical-analyst-in-arcgis-pro.htm>
- Fitzhugh, L., 2012. Water Quality and its Effect on Seagrasses: Depth Distribution, Biomass, Blade Width, and Carbon Content. Retrieved from [sabrma.org: http://www.sabrma.org/images/Water\\_quality\\_and\\_its\\_effect\\_on\\_seagrasses\\_fall\\_2012.pdf](http://www.sabrma.org/images/Water_quality_and_its_effect_on_seagrasses_fall_2012.pdf)
- Fortes, P. M., 1998. Bolinao Declaration. Retrieved from World Seagrass Association: <http://wsa.seagrassonline.org>
- Geospatial Technology, 2013. What's the difference between a supervised and unsupervised image classification. From extension.org: <https://articles.extension.org/pages/40214/whats-the-difference-between-a-supervised-and-unsupervised-image-classification>
- Krause-Jensen, D., Quaresma, A., Cunha, A. H., & Greve, T. M., n.d. How are seagrass distribution and abundance monitored. Retrieved from Vlaams Instituut De Zee: <http://www.vliz.be/imisdocs/publications/67189.pdf>
- Kotz, S., 2004. Least Squares Regression, Convex. Encyclopedia of Statistical Sciences. Retrieved from <https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/ordinary-least-squares-regression>.
- Matinfar, H., Sarmadian, F., 2007. Comparisons of Object-Oriented and Pixel-Based Classification of Land Use/Land Cover Types Based on Landsat7, Etm Spectral Bands (Case Study: Arid Region of Iran). American-Eurasian J. Agricultural & Environmental Science, 448-456.
- Montano, M. N. 2005. *BOLINAO, PHILIPPINES - SEAGRASS DEMONSTRATION SITE*. Bolinao, Pangasinan, Philippines.
- Ray, S., 2015. 7 Types of Regression Techniques you should know! Retrieved from Analytics Vidhya: <https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/>
- Reynolds, P. L., 2018. Seagrass and Seagrass Beds. Retrieved from Smithsonian: <https://ocean.si.edu/ocean-life/plants-algae/seagrass-and-seagrass-beds>
- Robertson, G.P., 2008. GS+ : Geostatistics for the Environmental Sciences. Gamma Design Software, Plainwell, Michigan USA.