Empirical Models to Estimate Seawater Parameters in Mayagüez Bay

Nazario D. Ramirez-Beltrán[1], Fernando Gilbes[2] and Joan Manuel Castro[3]

[1] Department of industrial Engineering, University of Puerto Rico, Mayagüez P.R. 00680
[2] Department of Geology University of Puerto Rico, Mayagüez P.R. 00680
[3] Department of Electrical and Computer Engineering, University of Puerto Rico, Mayagüez P.R. 00680

Abstract - The application of remote sensing in coastal environment has been very difficult because of the optical complexity of these waters. Seasonal river discharge and land run-off increase the concentration of phytoplankton biomass along with suspended sediments, and colored chromophoric dissolved organic matter (CDOM). The occurrence of these events in coastal waters affects the penetration of visible radiation into the water column and can confound remote sensing estimates of chlorophyll *a* concentrations. The Mayaguez Bay, located in the west coast of Puerto Rico is not exception. There are three rivers whose discharges occur on the Mayaguez Bay and consequently the SeaWiFS sensor is not able to estimate concentration of chlorophyll *a*.

Keywords: chlorophyll a, spectral slope, organic matter (CDOM), model identification

1. INTRODUCTION

1.1 Ocean color.

The color of seawater is determined by the interactions of light with the ocean. Colors can be observed when the light is reflected by the objects. A white light is made up of a spectrum or combination of colors, which are broken apart by water droplets in a rainbow. When the light hits the surface of an object the different colors can be absorbed, transmitted, scattered, or reflected in different kind of intensities. The substances in the ocean water that mostly affect the reflected light are: phytoplankton, dissolved organic matter, and suspended sediments, and the water itself. When the ocean is observe from the space, it reveals a blue color because water absorbs red and reflects blue light. If an instrument that is more sensitive than the human eye is used then it can measure a wide array of blue shades, which reveal the presence of varying amounts of phytoplankton, dissolved organic matter, and suspended sediments.

1.2 Ocean Color and Remote Sensing

The application of remote sensing to study the coastal environment has been very hard because there are many obstacles that affect the penetration of the light and consequently the optical parameters cannot properly be measured. For example the seasonal river discharges, land run-off increases, and the internal sea cycles affect the phytoplankton biomass concentration along with suspended sediments and colored dissolved organic matter. The influence of phytoplankton on the color of seawater has been studied for many years. It is understood that the primary photosynthetic pigments in phytoplankton called chlorophyll *a*, absorbs blue and red light in the visible band and the color of the ocean water progressively shifts from deep blue to green as the concentration of phytoplankton increases.

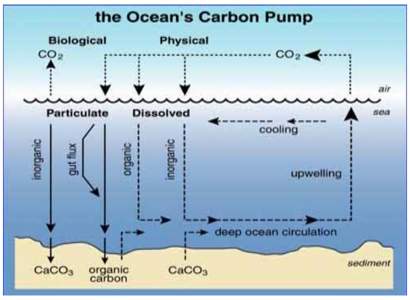

Figure 1: The synthesis of organic matter by plants (primary production) is a basic source of food for all living organisms.

In summary the conventional remote sensing techniques to measure the water optical properties do not work property in coastal areas, especially because of large concentrations of phytoplankton, dissolved organic matter, and suspended sediments. Thus, the main purpose of this research effort is to develop a set of empirical equations to improve the estimation of chlorophyll *a*, solid suspended and spectral slope.

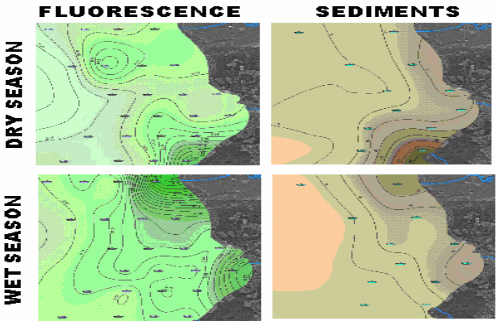

Figure 2: The fluorescence and the suspended sediments are highly correlated

The optical organic matter has been problematic to calibrate the sensor to measure the correct phytoplankton concentration in seawater, especially in the river mouth near the coastal area. The absorption of the CDOM and phytoplankton occurs in the same bands in the visible spectrum. Thus, it is needed to filter the CDOM and phytoplankton to determine the water properties. A second problem is the fluorescence of the chlorophyll *a*, which is highly correlated with suspended sediments, as shown in Figure 2. The study of concentration of the suspended sediments is important to trace rivers plumes, and to identify areas with large turbulence associated to the wave action and storm events.

1.3 Mayagüez Bay Seawater properties

In Puerto Rico (PR), several studies have been conducted to understand the bio-optical properties of Mayagüez Bay. This semi-enclosed bay localized in the western coast of the PR has been suffering temporal variations in the phytoplankton pigments and suspended sediments due to seasonal discharges of local rivers (Giles et al., 1996). About 10 years ago a joint project with researchers from NASA-Stennis Space Center and the University of Puerto Rico – Mayagüez Campus experimented with remote sensing to have a better knowledge of the land-sea interface in the Mayagüez Bay. However the complexity of the optical properties in this bay and some limitations of the technology at that time made this study very difficult. New efforts using improved methods and betters instruments are accomplishing the original objectives.

The major contribution of this work is the developing of an algorithm to identify the bands and the mathematical transformation that best explain the concentration of chlorophyll a, sediments, organic matter.

2. Method:

Two cruises per year have been conducted in the Mayaguez Bay and the data used in this paper cover the period from 2001 to 2004. The samplings were planned considering the dry and wet seasons of the region and from inshore to offshore waters. The sampling layout includes 24 stations as shown in Figure 3. Mayagüez Bay coast includes three towns in the west area of Puerto Rico (Añasco, Mayagüez and Cabo Rojo). Each town has a river that discharges water and other components to the bay. Añasco city has the Añasco River located in the North of the bay. In the center of the bay is located the Mayagüez city and has the Yagüez River. Cabo Rojo city has the Guanajibo River located in the South of the bay. The region also is affect by the dumping of the tuna factory and sewage pipe located between Añasco and Yagüez rivers, as shown in Figure 3. Twenty-four stations were sampled with an optical package, in which the half of the stations (12) had ancillary data.

These stations were sampled with the optical rosette, which is an optical package that includes several instruments to measure profiles of different water properties: temperature and salinity, chlorophyll fluorescence, spectral transmittance, spectral adsorption, the backscattering coefficient, etc. Water living radiance, and the above-surface downwelling irradiance were measure with GER 1500 portable spectroradiomenter. Results obtained by the optical rosette, and laboratory sampling are studied. Reflectance measure by GER was used to evaluate the SeaWiFS algorithm to estimate chlorophyll *a* and the estimated were compared with laboratory measurements. It was found that SeaWiFS algorithm in costal areas overestimate the chlorophyll *a* concentrations.

Collected data includes four samplings conducted during drought and lower river discharge season and three sampling conducted during rainy and greater river discharge season.

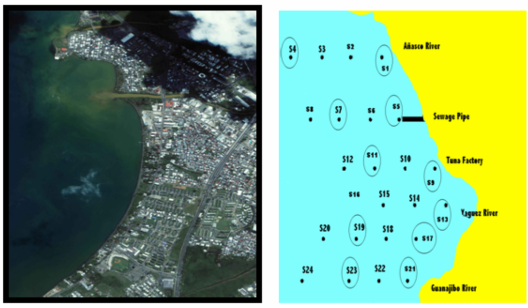

Figure 3: Hyperspectral Satellite Image of Mayagüez Bay (left) and, 24 Sample Stations located across Mayagüez Bay, inshore and offshore stations (right)

The study of chlorophyll *a* concentration was conducted with five samplings, the sediments studies includes six samplings and CDOM was conducted with only fourth samplings. The study was conducted in this way because of data limitation. The sampling dates are listed in Table 1.

Table 1. Sampling for model fitting and validation

No	Date	Sampling Used on					
		Chlorophyll *a*		Sediments		CDOM	
		Fit	Val.	Fit	Val	Fit	Val
1	04-01			x		x	
2	10-01	x		x		x	
3	02-02	x		x			x
4	08-02				x		x
5	02-03	x			x		
6	10-03		x				
7	02-04		x		x		

A preliminary analysis of the Mayaguez Bay shows that the bio-optical variability is very large, even though this is a relative small geographical area. Reflectance spectrum show different behavior depending of the water conditions. Figure 4 shows the oceanic sea water spectrum (green) provides the lowest reflectance in the visible range. In addition, in the blue area (400 and 500 nm) maintain about the same reflectance intensity and start decreasing in the green area (500 and 600 nm) and after that remains very stable and close to zero. On the other hand, if reflectance were obtained from the Añasco River discharge or sewage pipe water discharge system a different spectrum is obtained. In both cases, the reflectance intensity of the water was greater than the offshore ocean water. In addition, in the blue area of the visible band, the reflectance increased (positive slope). The Anasco River discharge shows an increment larger than sewage pipe discharge system. This event occurs fundamentally because of the suspended sediments and organic matter

Based on the collected information (such as: chlorophyll *a*, suspended sediments, spectral slope, reflectance and absorption) a set of empirical equations were developed to estimate chlorophyll a, suspended sediments and spectral slope. The main advantage of these equations is that they can provide reasonable estimates under coastal waters that contain large concentration of suspended sediments and CDOM. The proposed methodology to identify the empirical equations follows three major steps: 1) A band was randomly selected in a given specific range and a mathematical transformation was implemented. 2) The large amount of selected bands was organized in small groups to have the appropriate number of degree of freedom and a variable selection algorithm was implemented. 3) A multicolinearity routine was used to avoid the multicollinearity problem and derive a robust empirical model.

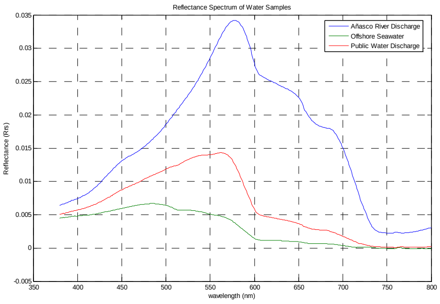

Figure 4: Reflectance Spectrum of Water Samples obtained on Anasco River Discharge (blue), Oceanic Seawater (green), and Sewage Pipe Discharge (red)

3. RESULTS

The rainy versus the drought seasons, and the river discharge are the principal factors to regulate the bio-optical properties in Mayagüez Bay, including phytoplankton populations. Anthropogenic activities in the river basins affect the composition of the rivers input in the bay and the characteristics of the water masses entering in the bay. The western basin of PR is highly developed and deforested, which favors erosion and transferences of soil particle into the river waters. The suspended particles increases scattering and absorption, effectively attenuating light, but also increase nutrient concentrations.

The main strategy to derive the empirical equation that correlates chlorophyll *a* with reflectance is as follows:

$$C_{ij} = f(R_{rs}(x_{ij})), \quad 500nm \le R_{rs}(x_{ij}) \le 600nm \quad (1)$$

Where C_{ij} is the i^{th} concentration of chlorophyll *a* at the j^{th} station, and $R_{rs}(x_{ij})$ is the reflectance of the i^{th} observation at the j^{th} station. The identified equation is given as follows:

$$C = 12.4 - 8.61\left(\sin\left(\frac{R_{rs}(570nm)}{R_{rs}(542nm)}\right) + \cos\left(\frac{R_{rs}(570nm)}{R_{rs}(542nm)}\right)\right) \quad (2)$$

This model explains 79% of the variance of chlorophyll *a*. The model fitting and validation is given in Table 2. The model was developed using 3 sampling results and validated with two samplings. Figure 5 shows the model fitting performance and Figure 6 shows the model validation.

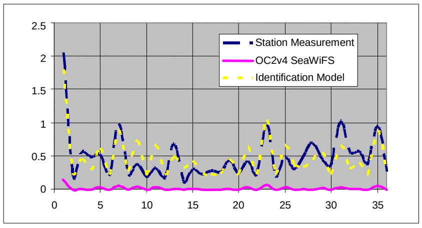

Figure 5: Model Fitting Performances of Chlorophyll using three samplings (August 2001, February 2002, and February 2003)

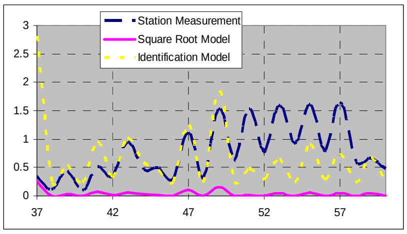

Figure 6: Validate Chlorophyll Estimations using two stations (October 2003, and February 2004)

The strategy to derive the empirical equation for suspended solids is as follows:

$$S_{ij} = f(R_{rs}(x_{ij})), \quad 600nm \le R_{rs}(x_{ij}) \le 900nm \quad (3)$$

Where S_{ij} is the i^{th} suspended solids at the j^{th} station, and $R_{rs}(x_{ij})$ is the reflectance of the i^{th} observation at the j^{th} station. The identified equation is given as follows:

$$S = 0.19 + 3.31 e^{-\frac{R_{rs}(550\ nm)}{R_{rs}(600\ nm)}} \quad (4)$$

Model performances and validation are given in Figures 7 and 8 as well in Table 2. This model explains only 55% of the total variance of the suspended sediments.

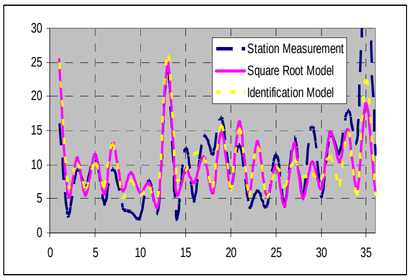

Figure 7: Sediments Model Fitting using three samplings (April 2001, August 2001, and February 2002)

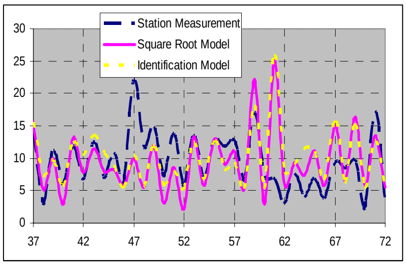

Figure 8: Sediments Model Validation using three samplings (August 2002, February 2003, and February 2004)

The strategy to derive the empirical equation for spectral slope is as follows:

$$SS_{ij} = f(R_{rs}(x_{ij})), \qquad 400nm \le R_{rs}(x_{ij}) \le 500nm \qquad (5)$$

Where SS_{ij} is the i^{th} spectral slope at the j^{th} station, and $R_{rs}(x_{ij})$ is the reflectance of the i^{th} observation at the j^{th} station. The identified equation is given as follows:

$$SS = 0.20 + 0.067 \tan\left(\frac{R_{rs}(450nm)}{R_{rs}(500nm)}\right) \qquad (6)$$

Again the model fitting performances and validation is given in Figures 9 and 10, respectively. Table 2 also presents some assessment for model fitting and validation. This model explains 82% of the total variance of the spectral slope.

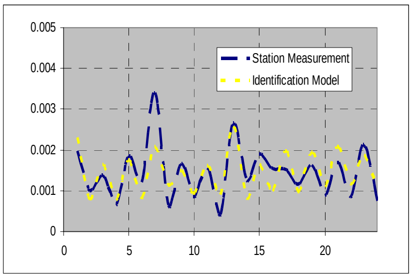

Figure 9: Model Fitting Performance for Absorption of Spectral Slope using two samplings (April 2001, August 2001)

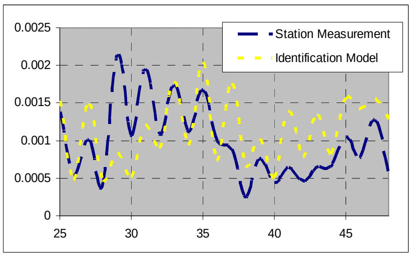

Figure 10: Model Validation for Absorption Spectral Slope using two samplings (February 2002, and August 2002)

Table 2 summarizes the model fitting performance and the assessment of model validation.

Table 2: Mean Square Error results using models to estimate seawater properties

Chlorophyll	**MSE estimation**	**MSE validation**	Band 1	Band 2
Equation (2)	0.03	0.46	570	542
OC2V4	0.33	0.73	490	555
Sediments	**MSE estimation**	**MSE validation**	Band 1	Band 2
Equation (4)	22.9	25.1	550	600
Square Root	22.7	23.1	600	
Spectral Slope	**MSE estimation**	**MSE validation**	Band 1	Band 2
Equation (6)	1.69E-07	2.76E-07	500	475

4. Conclusion:

The choice of the sampling for model building and calibration affects the performance of the models. The largest effect was observed at the station located at the mouth the Añasco River, because there is a large discharge of sediments and organic matter, the proposed models can not estimate correctly the chlorophyll concentration at that station. However, in the remaining stations, the new algorithms provide appropriate estimation of water properties. The chlorophyll proposed model autoperforms the SeaWiFS algorithms.

The algorithms designed to estimate chlorophyll *a* at global scales are less accurate than the algorithms designed fro local and regional purposes. Combination of global, regional and local algorithm are required to properly estimate the optical properties in coastal waters.

5. Acknowledgments

This research was partially support by the NOAA-CREST grant NA17AE1625 and the Institute of Water Resources of the University of Puerto Rico at Mayagüez.

6. References

Gilbes F., Lopez J.M., and Yoshioka P.M., 1996. Spatial and temporal variations of phytoplankton chlorophyll-a suspended particular manner in Mayagüez Bay, Puerto Rico. Journal of Plankton Research, 18(1): 29-43