Combining Remote and In-situ Sensing for Persistent Monitoring of Water Quality

Cesar A. Rojas[1], Gregory M. Reis[1], Arif R. Albayrak[2,3],
Batuhan Osmanoglu[2], Leonardo Bobadilla[1], and Ryan N.Smith[4]

***Abstract*—Many studies suggest that water quality parameters can be estimated by applying statistical and machine learning methods using remote sensing or in-situ data. However, identifying best practices for implementing solutions appears to be done on a case-by-case basis. In our case, we have in-situ data that covers a large period, but only small areas of Biscayne Bay, Florida. In this paper, we combine available in-situ data with remote sensing data captured by Landsat 8 OLI-TIRS Collection 2 Level 2(L8), Sentinel-2 L2A(S2), and Sentinel-3 OLCI L1B(S3). The combined data set is for use in a water quality parameter estimation application. Our contributions are two-fold. First, we present a pipeline for data collection, processing, and co-location that results in a usable data set of combined remote sensing and in-situ data. Second, we propose a classification model using the combined data set to identify areas of interest for future data collection missions based on chlorophyll-a in-situ measurements. To further prove our methodology, we conduct a data collection mission using one of the predicted paths from our model.**

***Index Terms*—remote sensing, machine learning, water quality, logistic regression, estimation, robots, chlorophyll-a**

I. Introduction

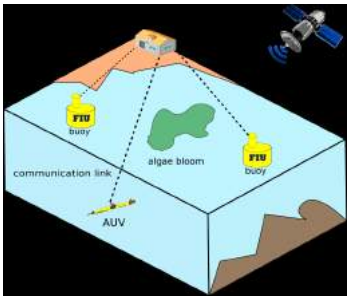

Fig. 1. A shore station using data gathered from water buoys and a satellite to deploy an AUV on data collection missions.

In recent years, Florida's coastal waters have endured several algal blooms, and the Biscayne Bay in Miami experienced significant fish kill events in August 2020 [1] and September 2021 [2]. Investigations on the Biscayne Bay fish kills conclude that the cause is from a culmination of several years of nutrient-rich run-off seeping into the bay. The source of the nutrients is a combination of physical and environmental factors including, but not limited to leaking septic tanks, grass fertilizer, and trash buildup in the sewage and canal system. The Biscayne Bay has responded with a decrease in seagrass and drops in oxygen levels. Conditions get worse in the summer months due to rising temperatures causing marine life to suffocate. To minimize losses during these events, Miami city officials respond by deploying vessels that can oxygenate water onto the bay [1], [3].

As a result, water quality measuring, monitoring, and estimation are critical for threat detection and mitigation. To further understand the nature of these harmful water events, Florida International University [4] continuously collects in-situ data in coastal waters by deploying water quality sensing buoys. Deployments cover over a limited amount of locations and also using in-house built robotic platforms with fine spatial and temporal resolutions. There are two caveats when using the in-situ data. First, data collection deployments are resource-intensive. Second, collected data may only report on small pockets of areas due to constraints such as battery life. To overcome these problems, we consider complementing our in-situ data with remote sensing images captured in the same area. To expand our coverage, we plan to implement a supervised machine learning-based approach in which in-situ data, buoys, and robot data are for use as the labels to the collocated and plentiful satellite training data set. [5].

Water buoy placement and data collection mission deployments are a reaction to a disaster event. This results in the loss of data leading up to the disaster event. Our goal is to identify locations we should target for future water buoy placements and data collection missions. An overview of our approach is in Figure 1 and depicts a shore station deploying in-situ data collection robots in key areas identified by a trained classification model.

In our work, we aim to combine the strengths of both sources, in-situ and remote, to obtain synergies. Our approach has the following caveat; we test our methodology using remote sensing data points across multiple scenes captured on different dates downloaded from Sentinel-Hub. Scene contamination occur by changes in the atmosphere, clouds, rain, sun glint, vehicles, and other disturbances. Atmospheric correction processes attempt to correct some of these issues, however

[1]Cesar A. Rojas, Gregory M. Reis, and Leonardo Bobadilla are with the Florida International University Knight Foundation School of Computing and Information Sciences, Miami, FL 33199, USA (e-mail: croja022@fiu.edu, greis003@fiu.edu, bobadilla@cs.fiu.edu)

[2]Arif R. Albayrak and Batuhan Osmanoglu are with the NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA(e-mail:rustem.a.albayrak@nasa.gov, batuhan.osmanoglu@nasa.gov)

[3]Arif R. Albayrak is with the University of Maryland, Baltimore County, Baltimore MD 21228, USA(e-mail:albayrak@umbc.edu)

[4]Ryan N. Smith is with the Florida International University Institute of Environment, Miami, FL 33199, USA(e-mail:rysmith@fiu.edu)

not all processes may be appropriate for shallow estuaries such as the Biscayne Bay. This is because the Biscayne Bay is a Class-II water. Class-I waters are open ocean waters and Class-II are estuaries and coastal waters. [6] Class-I and Class-II waters differ in their inherent optical properties(IOPs). For our experiments, we use atmospherically corrected data sets from L8 and S2 satellites. L8 and S2 are a result from using LaSRC [7] and Sen2Cor [8] algorithms respectively. Algorithms designed for Case-I waters may not function correctly on Case-II waters [6]. Selecting a proper atmospheric correction algorithm for our ocean work environment needs to be studied further. S3 is not atmospherically corrected, so it is used purely to test the methodology.

II. Related Work

Studies by Gholizadeh et al. [9], Nazeer et al. [10], Hafeez et al. [11], Cruz et al. [12] examine several approaches to applying statistical methods and machine learning algorithms with in-situ and remote sensing data and report metrics such as cross-validation accuracy scores and root-mean-square error (RMSE). These findings provide some context for our experiment design as we highlight relevant studies below.

A. Approaches using only in-situ data

Chen et al. [13] implemented an auto-regressive integrated moving average (ARIMA) model to predict chlorophyll-a to detect algal blooms in the Taihu Lake of Jiangsu, China. For 11 months, water quality samples were collected every 1-3 days. The ARIMA model was implemented using chlorophyll-a measurements and compared against a multi-variate linear regression (MVLR) model that uses water temperature, water transparency, total inorganic nitrogen concentration, and phosphate concentration. The results of this study showed that the ARIMA model outperformed the MVLR model despite being a uni-variate method. Logistic regression has been used to effectively predict and classify algal blooms by Yun [14] using in-situ data collected by the National Institute of Fisheries Science in South Korea.

B. Approaches using only Remote Sensing data

There are cases where remote-sensing is the only data available for a given area of interest and we must rely on deriving measurements using band ratios and spectral indices. Mishra et al. [15] propose the Normalized Difference Chlorophyll Index (NDCI) for use in chlorophyll-a estimation on Case-II waters. Caballero et al. [16] use NDCI to detect algal blooms in coastal waters, but use in-situ data to be able to detect harmful algal blooms caused by toxic algae species. Furthermore, Tang et al. [17] built logistic regression models using remote sensing data from 2013 to predict the distribution area of the Enteromorpha prolifera in the Yellow Sea, China.

C. Combining in-situ and remote sensing using only one scene

NDCI is possible to derive using S2 and S3. [18] L8 is missing bands for the red-edge spectral range between $700nm$ and $720nm$ for this particular index. [16] However, this does not mean L8 images cannot be used to derive a chlorophyll-a measurement. Yang et al. [19] conducted a study comparing different band ratios and their correlation with chlorophyll-a in-situ measurements. They co-located in-situ data with remote sensing data of two L8 images, one for Summer and the other for Fall. In-situ data was collected by 20 different sampling stations in Jordan Lake, North Carolina, within two days of the date of the closest L8 image. Their results showed that the correlations found between the two co-located remote sensing data differed in the ratio band with the highest correlation and could mean that different models should be implemented for different seasons. Han and Jordan [20] identify $\frac{Band1(443nm)}{Band3(561.5nm)}$ as the L8 band ratio with the highest correlation with chlorophyll-a measurements in Pensacola Bay, Florida.

III. Problem Formulation

We consider the area of interest as a 2-D environment where the ocean surface is denoted as $\mathcal{W} \subset \mathbb{R}^2$. We discretize the workspace into a 2-D grid. Let $\mathcal{C}$ be the collection of all 2-D cells within a bounding convex polygon $\mathcal{P}$ where $\mathcal{P} \subseteq \mathcal{W}$. Each cell has a center in the form of WGS84 coordinates denoted as (x, y), where $x, y \in \mathbb{R}$. The geographic coordinate (x, y) represents the center of an equal-sized grid tile. This grid size represents the resolution of the *remote sensing* data. Let $\mathcal{Q} \leftarrow (d, b, i, m)$ be a satellite query configuration where d is a satellite identifier and b, i, m are lists of requested band measurements, indices, and metadata respectively. We will represent the remote sensing data as $\rho : \mathcal{C} \times \mathcal{Q} \times T_1 \rightarrow \mathbb{R}^k$, where $T_1 = [0, t_1)$ is the time interval data and $k \leftarrow |b| + |i| + |m|$ is the total number of variables. The unit of frequency of the remote sensing data examined is days (e.g., five days in the satellite data we use, S2).

Our domain will have a fixed *buoy* $\mathcal{B}$ at position (x_b, y_b) in cell c_b [4]. This buoy collects in-situ data modeled as $\beta : T_2 \rightarrow \mathbb{R}^n$, where $T_2 = [0, t_2)$ represent the time collection interval and n is the number of in-situ sensors [4]. In the buoys we are using, the sensors include dissolved oxygen, temperature, salinity, and turbidity. The measurement frequency of this buoy is higher, in our example observations are made every 15 minutes [4].

We will also deploy an Autonomous Underwater Vehicle (AUV) $\mathcal{A}$. The state-space for the vehicle is defined as $X = \mathcal{W} \times [0, 2\pi)$ in which $[0, 2\pi)$ is the set of angles that represents the vehicle's orientations. We use the unicycle vehicle model as described in [21]. A path for the vehicle is denoted as $\pi : T_3 \rightarrow X$, where $T_3 = [0, t_3)$ is the deployment mission time.

Problem 1. Computing a critical data collection plan for the AUV based on remote and in-situ information: *Given an environment $\mathcal{W}$ of interest, a bounding convex polygon $\mathcal{P}$, a satellite configuration $\mathcal{Q}$, a collection of remote sensing data ρ, a collection of buoy data β, produce a path π that visit locations of importance.*

IV. Methods

A. Data Collection

TABLE I
RESEARCH BUOYS DEPLOYMENTS

Buoy Location	Deployment Details[a]		
	Period	*Longitude*	*Latitude*
Miami Shores	08/27/2020-09/19/2021	-80.16712	25.86171
Haulover Inlet	10/05/2018-07/30/2020	-80.13311	25.90273
North Bay Village	08/13/2020-01/27/2021	-80.14203	25.83991

[a]May have small gaps for servicing and maintenance.

TABLE II
DATA COLLECTED

Buoy Location	Source[a]			
	In-Situ	*Landsat 8*	*Sentinel-2*	*Sentinel-3*
Miami Shores	37,042	24	78	702
Haulover Inlet	47,532	43	138	818
North Bay Village	15,345	10	33	315

[a]Collected every 15 minutes, 16 days, 5 days, and 1-2 days respectively

Fig. 2. Research buoy deployment locations.

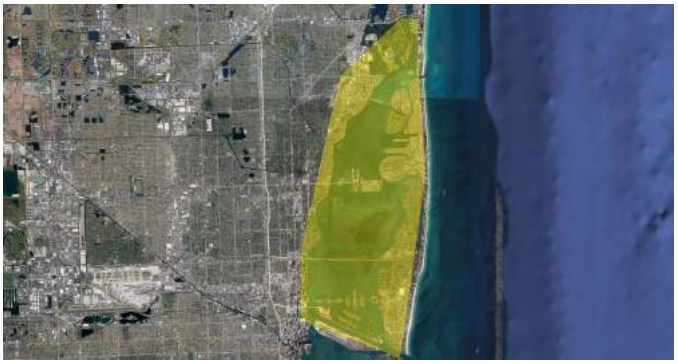

Fig. 3. The ocean work environment $\mathcal{W}$ is colored in yellow.

First, we collect data from the water buoys which are stored as comma-separated value files(CSV) with 26 columns to create β [4]. However, we only examine four metadata columns (date, time, latitude, and longitude) and $n = 6$ measurement columns (pH, Temperature, Salinity, Turbidity, Oxygen, and Chlorophyll-a). Table I shows the locations and time span of the water buoy deployments. Figure 2 illustrates the location of the water buoy deployments.

We create $\mathcal{P}$ by collecting remote sensing data using several satellite query configurations for L8, S2, and S3 satellites and a bounding polygon at the location of the water buoy and querying a public satellite data store for the time interval covered by the water buoys [4]. Each water buoy placement is associated with a pixel on a satellite image capture. Pixels represent one or more longitude and latitude coordinates. If the pixel is contaminated by land, such as a nearby island, we select a bounding box with the nearest pixel associated only with coordinates to water. Figure 3 depicts our ocean work environment $\mathcal{W}$. Table II shows a count of in-situ measurements and satellite images gathered for each deployment location.

B. Classification Model

We will formalize the problem of finding regions with high chlorophyll-a as a classification or supervised learning problem. More concretely, a two-class classification where y is the class where 0 represents normal chlorophyll-a levels and 1 represents high levels. The features used for the classification x represent the spectral bands and indices data from an image captured by L8, S2, or S3.

$$h_w(x) = \frac{1}{1 + exp(-w^T x)} \tag{1}$$

$$p(y = 1|x; w) = h_w(x) \tag{2}$$

$$p(y = 0|x; w) = 1 - h_w(x) \tag{3}$$

We would like a score that represents how likely a given pixel has high chlorophyll-a. We will use *logistic regression* where the classification hypothesis is represented by a logistic function in equation 1 [22]. w represents the parameters that need to be learned. We will interpret the output of the logistic regression as equation 2 and equation 3.

TABLE III
TOTAL NUMBER OF CO-LOCATED DATA POINTS

Satellite	Chlorophyll-a
Landsat 8	15
Sentinel-2	118
Sentinel-3	1071

The model will be fitted using multi-variate samples and the indices will be included with the raw band values from remote sensing satellite images co-located with the in-situ measurements. Any quality assurance bands, data masks, and probabilities are not included in the samples. To start the co-location process, we first eliminate data points from satellite images that have errors or are contaminated by clouds using data masks. L8 data can derive a water mask using the quality assurance band and S2 data has a cloud mask available. We attempt to remove in-situ data with errors by removing values

that are all zero values or "NaN" values, but a quality assurance column is not available at the moment. Any appropriate indices such as NDCI for S2 and S3 images are now calculated and concatenated to the in-situ data. For L8 we use the indices identified by Yang et al. [19]. Next, the in-situ data are re-sampled to daily maximums. Each satellite image data point is paired with a measurement column in the re-sampled in-situ data. If a satellite image data point is unable to be paired with a measurement recorded on the same date of capture, it is paired with the last known value. The total number of co-located data points for each satellite is in Table III. Then each co-located data point to classified as a "1" if it meets or exceeds a selected threshold, otherwise it is classified as a "0". For chlorophyll-a, the threshold we select is $72\mu g/L$ as it is the minimum value chosen for one of the highest concentration groups in [19].

C. Robot Path Construction

Once the classification model is fitted, we make a request to Sentinel Hub for an image with the most recent cloud-free pixels within a bounding box representing our area of interest in the Biscayne Bay. Sentinel Hub returns an image with multiple pixels. Each pixel covers multiple longitude and latitude coordinates, this means multiple entries for the same pixel are needed to render it as an RGB image. Since multiple longitude and latitude coordinates are associated with each pixel, we drop duplicate pixels from each satellite image and keep the coordinates of the first pixel encountered. The remaining pixels is input for the classification model's prediction step. The model assigns a probability percentage for each class to each pixel and the top ten pixels is input for a Christofides algorithm implementation; this results in an ordered path that is within a factor of 3/2 of the optimal path [23]. The resulting path is π and is translated to a valid navigation path for a data collection robot. The translation is necessary since not all robots may be able to make sharp turns or run into obstacles.

V. Results

TABLE IV
5-Folds Accuracy Scores

Satellite	Score
Landsat 8	66%
Sentinel-2	87%
Sentinel-3	88%[a]

[a]based on atmospherically uncorrected satellite images.

The performance of the classification model is measured using a K-Folds cross validator set to use 5 folds with shuffling and random state disabled. Table IV shows accuracy scores of 66% fitted with co-located L8 images, 87% fitted with co-located S2 images, 88% fitted with co-located S3 images. The model predicts which coordinates have high concentrations of chlorophyll-a in several areas of interest of varying sizes.

TABLE V
Classification Input Details

Satellite	Input Size(Pixels)	Resolution[a]
Landsat 8	690×1523	$30m$
Sentinel-2	132×174	$10m$
Sentinel-3	2245×2311	$300m$

[a]Approximate resolution of RGB bands, others bands may be up-sampled.

Please note S3 performance is based on atmospherically uncorrected satellite images and should hold no meaning.

The size of the input images needs to be appropriate to the satellite used to train the model or there may not be enough pixels to classify. For example, attempting to classify an L8 image with the same bounding box used for S2 will result in classifying less than ten pixels. Table V shows the resolution of each RGB pixel, however not all band information downloaded for each satellite is of the same resolution and may have been up-sampled. Sentinel-Hub API use nearest-neighbor interpolation by default when requesting pixels with resolution greater than the source for L8, S2, and S3 [24]. All areas of interest are within the ocean work environment $\mathcal{W}$ so RGB pixels are expected to be different shades of blue due to different water quality, and the amount of sunlight or cloud shadow over the water. White pixels may represent clouds or objects such as boats and piers. Green pixels could indicate an increase in chlorophyll-a.

A. Classification Results - BBC Campus Pier

Our first area of interest is a pier located at the Florida International University Biscayne Bay Campus. Due to how small this area is, images from S2 are ideal because of the $10m$ resolution bands. This area is also where we conducted an actual data collection mission planned using our model. Figure 4(a) shows a YSI EcoMapper IVER2 AUV [25] deployed at the area of interest conducting the calculated data collection mission escorted by a manned vessel. The bounding box representing the area of interest is in Figure 4(b) and the RGB bands for the S2 image used by the model to make its classification is in Figure 4(c). And Figure 4(d) shows the Christofides path used in the motion planning of the deployed IVER2 AUV.

B. Classification Results - Haulover

The next area of interest expands the previous one more into the Haulover Inlet. For this area, we examine the use of L8 imagery. Unlike the previous area, the resolution of the $30m$ resolution bands of L8 gives us many pixels for the model to classify. We also have easy access to a water mask using the quality assurance band in the Level 2 collection to help us filter out contaminated pixels. The bounding box representing this area of interest is in Figure 5(a) and the RGB bands for the L8 image used by the model to make its classification is in Figure 5(b). Figure 5(c) shows the Christofides path to be used in the motion planning of a future robot data collection mission.

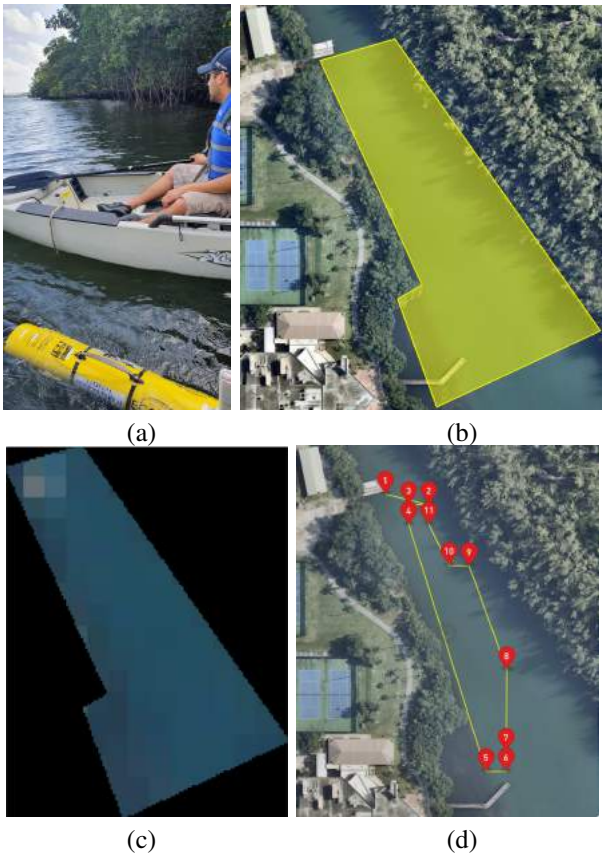

Fig. 4. (a) The AUV deployed to the area of interest in the ocean work environment; (b) The bounding convex polygon; (c) A recent S2 RGB image capture downloaded from Sentinel Hub within the bounding convex polygon. (d) Resulting path using our classification model to identify areas with highest chlorophyll-a concentrations on the Biscayne Bay pier at Florida International University Biscayne Bay Campus.

Fig. 5. (a) The bounding convex polygon; (b) A recent L8 RGB image capture downloaded from Sentinel Hub within the bounding convex polygon. (c) Resulting path using our classification model to identify areas with the highest chlorophyll-a concentrations on the Biscayne Bay Haulover Inlet near Florida International University Biscayne Bay Campus.

C. Classification Results - North Biscayne Bay

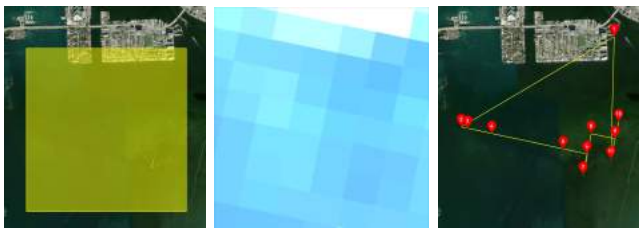

Fig. 6. (a) The bounding convex polygon; (b) A recent S3 OLCI RGB image capture downloaded from Sentinel Hub within the bounding convex polygon. (c) Resulting path using our classification model to identify areas with the highest chlorophyll-a concentrations near the Biscayne Bay North Bay Village.

The last area of interest we examine is in North Bay Village and it is located near the center in between all the water buoy deployment locations. The bounding box depicted in Figure 6(a) was primarily chosen due to size limitations enforced by the Sentinel Hub API. Figure 6(b) shows a S3 RGB image downloaded from Sentinel Hub. The $300m$ resolution of S3 requires a large bounding box to get a good amount of pixels for our model to classify. Figure 6(c) shows the Christofides path to be used in the motion planning of a future robot data collection mission. For this area of interest, the starting point of the closest pier happens to be outside the bounding box.

VI. Conclusions and Future Work

The results reported are based on actual ocean model prediction data that demonstrate the applicability of our method. We implemented and applied a logistic regression model to a chlorophyll-a classification problem using three different satellites and achieved acceptable accuracy scores. The model can be further improved through the selection of an atmospheric correction algorithm, different indices, increasing the number of in-situ measurements collected, different data masks, changing the classification model algorithm, and other avenues.

This work primarily focused on implementing models based on chlorophyll-a, but the water buoy in-situ data has five other measurements to explore. Measurements such as temperature and salinity can be used to identify the conditions in which a toxic algae species can bloom [26]. Furthermore, Kim et al. [27] show remote sensing can be used to monitor dissolved oxygen of coastal waters. A classification model based on oxygen can help city officials determine where to send oxygenation vessels on the Biscayne Bay in the summer months.

A promising atmospheric correction processor is the ACOLITE processor. Vanhellemont and Ruddick [28] show ACOLITE corrected L8 and S2 images can be used to effectively detect and map phytoplankton blooms in coastal and inland waters. Caballero et al. [16] also use the ACOLITE process for their work due to its effectiveness to correct sun glint. We have already begun investigating the applicability of the

ACOLITE processor to remote sensing data of the Biscayne Bay.

We are also interested in using other remote sensing data sets. Claverie et al. [29] present the Harmonized L8 and S2 surface reflectance data set. The rate of revisit is higher by combining L8 and S2 data sets. Another promising data set is from the recently launched Landsat 9 satellite which complements L8 remote sensing data and offers improved observations of coastal waters. [30]

Additional work includes accounting for bathymetric information to correct remote sensing data, introduction of other variables of interest, and conducting more field experiments with AUVs as well as unmanned surface vehicles (USVs) and unmanned aerial vehicles (UAVs).

VII. Acknowledgment

This work is supported in part by the NSF grants IIS-2034123, IIS-2024733, by the U.S. Dept. of Homeland Security grant 2017-ST-062000002, and by the ESA Network of Resources Initiative. We also acknowledge the equipment loan from Fort Lewis College, Durango, CO.

References

[1] T. Florin, "Miami-dade county continues to respond to recent fish kill in biscayne bay with multi-pronged approach," Aug 2020.
[2] T. Florin, "Update on fish kill from miami-dade county," Sep 2021.
[3] L. M. Gomez, "Fish kill," 2021.
[4] M. Presa-Reyes, B. Bogosian, B. Schonhoff, C. Jerauld, C. Moreyra, P. Gardinali, and S. C. Chen, "A Water Quality Research Platform for the Near-real-Time Buoy Sensor Data," in *Proceedings - 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science, IRI 2020*, 2020.
[5] A. Albayrak, J. Wei, M. Petrenko, C. Lynnes, and R. C. Levy, "Global bias adjustment for MODIS aerosol optical thickness using neural network," *Journal of Applied Remote Sensing*, vol. 7, no. 1, 2013.
[6] B. Matsushita, W. Yang, P. Chang, F. Yang, and T. Fukushima, "A simple method for distinguishing global case-1 and case-2 waters using seawifs measurements," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 69, 2012.
[7] USGS, "Landsat collection 2 surface reflectance u.s. geological survey," 2021.
[8] Sentinel-Hub, "Atmospheric correction," 2021.
[9] M. H. Gholizadeh, A. M. Melesse, and L. Reddi, "A comprehensive review on water quality parameters estimation using remote sensing techniques," 2016.
[10] M. Nazeer, M. Bilal, M. M. Alsahli, M. I. Shahzad, and A. Waqas, "Evaluation of empirical and machine learning algorithms for estimation of coastal water quality parameters," *ISPRS International Journal of Geo-Information*, vol. 6, no. 11, 2017.
[11] S. Hafeez, M. Wong, H. Ho, M. Nazeer, J. Nichol, S. Abbas, D. Tang, K. Lee, and L. Pun, "Comparison of Machine Learning Algorithms for Retrieval of Water Quality Indicators in Case-II Waters: A Case Study of Hong Kong," *Remote Sensing*, vol. 11, no. 6, 2019.
[12] R. C. Cruz, P. R. Costa, S. Vinga, L. Krippahl, and M. B. Lopes, "A review of recent machine learning advances for forecasting harmful algal blooms and shellfish contamination," 2021.
[13] Q. Chen, T. Guan, L. Yun, R. Li, and F. Recknagel, "Online forecasting chlorophyll a concentrations by an auto-regressive integrated moving average model: Feasibilities and potentials," *Harmful Algae*, vol. 43, 2015.
[14] H. Yun, "Prediction model of algal blooms using logistic regression and confusion matrix," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, p. 2407, 06 2021.
[15] S. Mishra and D. R. Mishra, "Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters," *Remote Sensing of Environment*, vol. 117, 2012.
[16] I. Caballero, R. Fernández, O. M. Escalante, L. Mamán, and G. Navarro, "New capabilities of Sentinel-2A/B satellites combined with in situ data for monitoring small harmful algal blooms in complex coastal waters," *Scientific Reports*, vol. 10, no. 1, 2020.
[17] Y. Tang, S. He, H. He, P. He, and H. Qi, "A study on the growth characteristics of green algae blooms in the yellow sea using categorical analysis," 2016.
[18] C. V. RodrÃguez-Benito, G. Navarro, and I. Caballero, "Using copernicus sentinel-2 and sentinel-3 data to monitor harmful algal blooms in southern chile during the covid-19 lockdown," *Marine Pollution Bulletin*, vol. 161, p. 111722, 2020.
[19] Z. Yang and Y. Anderson, "Estimating Chlorophyll-A Concentration in a Freshwater Lake Using Landsat 8 Imagery," *Journal of Environment and Earth Science*, vol. 6, no. 4, 2016.
[20] L. Han and K. J. Jordan, "Estimating and mapping chlorophyll-a concentration in pensacola bay, florida using landsat etm + data," *International Journal of Remote Sensing*, vol. 26, 2005.
[21] S. M. LaValle, *Planning algorithms*, vol. 9780521862059. 2006.
[22] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression: Third Edition*. 2013.
[23] N. Christofides, "Worst-case analysis of a new heuristic for the travelling salesman problem," 1976.
[24] "Sentinel-hub api documentation," 2021.
[25] AUVAC, "Ysi ecomapper," 2021.
[26] L. E. Brand, L. Campbell, and E. Bresnan, "Karenia: The biology and ecology of a toxic genus," *Harmful Algae*, vol. 14, 2012.
[27] Y. H. Kim, S. Son, H. C. Kim, B. Kim, Y. G. Park, J. Nam, and J. Ryu, "Application of satellite remote sensing in monitoring dissolved oxygen variabilities: A case study for coastal waters in Korea," *Environment International*, vol. 134, 2020.
[28] Q. Vanhellemont and K. G. Ruddick, "Acolite processing for sentinel-2 and landsat-8 : atmospheric correction and aquatic applications," 2016.
[29] M. Claverie, J. Ju, J. G. Masek, J. L. Dungan, E. F. Vermote, J. C. Roger, S. V. Skakun, and C. Justice, "The Harmonized Landsat and Sentinel-2 surface reflectance data set," *Remote Sensing of Environment*, vol. 219, 2018.
[30] NASA, "Landsat 9," 2019.