University of Cincinnati

Date: 3/7/2024

I, Catherine Gottsacker, hereby submit this original work as part of the requirements for the degree of Master of Science in Environmental Engineering.

It is entitled:

Integrating UAV with sensors to monitor harmful algal blooms in surface waters

Student's name: **Catherine Gottsacker**

This work and its defense approved by:

Committee chair: Dongmei Feng, Ph.D.

Committee member: Drew McAvoy, Ph.D.

Committee member: Richard Beck, Ph.D.

47974

Integrating UAV with sensors to monitor harmful algal blooms in surface waters

A thesis submitted to the

Graduate College of the University of Cincinnati

in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE in ENVIRONMENTAL ENGINEERING

in the Department of Chemical and Environmental Engineering

of the College of Engineering and Applied Science

on

19 March 2024

Submitted by:

Catherine Gottsacker

Committee Chair:	Dongmei Feng, Ph.D.
Committee Members:	Drew McAvoy, Ph.D., P.E.
	Richard Beck, Ph.D

Abstract

Harmful algae blooms in surface waters are a global environmental concern and threaten both human and environmental health. By outcompeting aquatic diversity, causing dissolved oxygen levels in surface waters to fall, and secreting toxins, algae blooms stress water treatment infrastructure and result in large economic losses. To control and manage the impact of harmful algae blooms, timely detection and monitoring is critical. However, current monitoring methods, such as permanent monitoring stations or water sampling, can be very costly or time-intensive, and require direct water access. The methods become dangerous or impractical in areas surrounded by cliffs or wetlands.

In this study, a flexible, efficient, and cost-effective approach for monitoring surface water quality was developed by integrating water quality sensors and unmanned aerial vehicles (UAV). The integration platform was designed, constructed, and deployed through the summer of 2023 to monitor chlorophyll, phycocyanin, and turbidity in William H. Harsha Lake of Clermont County, Ohio. The water quality parameters, used as an indicator of algae blooms, were then correlated to reflectance from Landsat 8 and 9 and Sentinel 2 satellites through single and multiple linear regressions. Multiple linear regressions using reflectance from Sentinel 2 satellites yielded the highest correlations between reflectance and water quality, with R^2 values of 0.70, 0.86, and 0.97 for chlorophyll, phycocyanin and turbidity, respectively. From the regressions, visible, near infrared, and red-edge bands were identified as useful for algae detection, and commercially available multispectral cameras capable of integration with UAVs were identified for future improvement of the UAV monitoring platform.

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Dongmei Feng, for the opportunity to work on this research, and for her unwavering support and insight through all stages of this project. I would also like to thank my committee members Dr. Drew McAvoy and Dr. Richard Beck for their feedback and encouragement.

Additionally, I would like to thank the UAV MASTER Lab at the University of Cincinnati for building and supplying the UAV used in this project. I am especially grateful for Brian Kowalczyk, Justin Ouwerkerk, and Austin Wessels for providing advice and feedback on the design of the integration platform, skilfully piloting the UAV, and supporting fieldwork.

Further, I would like to acknowledge my friends and colleagues Pradeep Ramtel, Aaditya Bhattarai, Aakriti Regmi, Praval Devkota, Sumit Mahato, Emily Weidner, and Alex Suer for their assistance with data collection and visualization for my study.

Finally, I would like to express my heartfelt thanks to my family, who were always willing to lend a listening ear or jump in a kayak, for their constant support.

This project was supported by University of Cincinnati OoR Collaborative Research Advancement Program: Pilot Grants Award, and by the Ohio Space Grant Consortium.

Table of Contents

List of Figures

List of Tables

1. Introduction

Harmful algae blooms (HABs) in surface waters are a global concern, threatening both human and environmental health. Often caused by overfertilization of waters with nutrients from human activities, algae blooms and cyanobacteria (also known as blue-green algae) can outcompete other aquatic diversity and cause oxygen levels in water to fall (Vincent 2018). Freshwater cyanobacteria can also secrete hepatoxins, lipopolysaccharide, and neurotoxins, all of which are dangerous to human health (Dittmann et al. 2013). People can be exposed to these toxins through ingestion of contaminated water, skin contact, consumption of contaminated seafood, or even inhalation of aerosolized toxins while swimming or boating. Consequences can be minor, such as skin or eye irritation, but exposure can also cause serious neurological or muscular conditions and even death (Igwaran et al. 2024). Beyond causing human health concerns, these toxins can also result in fish kills, and death or illness in other animals or livestock (Carmichael and Boyer 2016).

Because of these health risks, harmful algae blooms can have detrimental impacts on the safety and availability of drinking water, particularly from freshwater reservoirs. For example, 2014 algae blooms in Lake Erie caused tap water to be unusable through Toledo, Ohio for two days (Watson et al. 2016). Disruptions to drinking water supplies and other activities have huge economic penalties, with annual costs ranging from $150 million to $6.5 billion in freshwater systems (Kudela et al. 2015). Furthermore, the incidence of HABs is increasing in frequency, magnitude, and duration globally, due to increases in nutrient pollution, rising atmospheric CO_2 concentrations and temperatures, and changing patterns of precipitation (Huisman et al. 2018; Smucker et al. 2021).

In order to control the impact of harmful algae blooms, prevention, control, and mitigation strategies are often employed (Corcoran and Hunt 2021). Prevention strategies generally seek to reduce the incidence and severity of blooms, though actions such as nutrient reduction and management. Control techniques attempt to suppress the severity, extent, and duration of blooms, such as through harvest or algicidal bacteria. Finally, mitigation strategies seek to reduce the impact of blooms, and include efforts such as drinking water treatment or lake closures. Regardless of which control strategy is employed, timely detection of the HABs is always important, whether to identify action thresholds or evaluate the strategy effectiveness.

To detect HABs in surface water, permanent monitoring stations equipped to measure chlorophyll or phycocyanin concentration as proxies for algae biomass are often used. However, establishing and maintaining the stations is a costly and labor-intensive endeavor. Chlorophyll data can also be gathered in the field using sensors or by collecting water samples, but these methods are time consuming and require direct access to the water. Both methods become dangerous and impractical in areas surrounded by cliffs or wetlands.

To decrease the burden of collecting water quality data while still understanding algae dynamics over large areas, remote sensing is often employed. Remote sensing models can be used to derive chlorophyll concentrations, as the pigment has a reflectance peak in the near-infrared region of the visible spectrum. Satellites with higher spectral resolution could also be used to identify the narrow 620 nm peak associated with phycocyanin, the pigment in cyanobacteria (Keith et al. 2018). However, the convenience of remote sensing models comes at the expense of high spatial and temporal resolutions, which are important to accurately map algae blooms that can change rapidly in response to varying environmental conditions (Egerton et al. 2014). In addition,

building remote sensing models still requires direct measurement of water quality, which can involve colossal efforts for data collection (Chaffin et al. 2021).

These challenges in data collection and timely monitoring of algae blooms can be addressed by using Unmanned Aerial Vehicles (UAV). Here, a flexible, efficient, and cost-effective approach for monitoring surface water quality is developed using UAV and a water quality sensor. The platform is deployed to collect chlorophyll, phycocyanin, and turbidity data during six flights at Lake Harsha, Ohio. The collected data, supplemented with additional boat-collected data, is used in combination with Landsat 8 and 9 and Sentinel 2 satellite imagery to identify algorithms capable of expanding HAB monitoring over larger areas and time periods within the same body of water. Finally, the spectral bands identified as useful for monitoring algae blooms at Lake Harsha are used to evaluate multispectral cameras with the potential for UAV integration that could further improve accessibility, accuracy, and convenience of algae bloom monitoring in surface waters.

2. Methods

2.1 Drone Sonde Integration

To begin the design of the drone-sonde integration platform, the water quality sensor payload was selected. The water quality sensor needed to have the capability to monitor algae concentrations in real time, locally store data, contain a power source, and have a minimal weight. To meet these criteria, a YSI EXO 3 Sonde was selected and used for all water quality data collection in the study. The YSI EXO 3 Sonde (hereby referred to as sonde) was customized to include a total algae sensor, turbidity sensor, and combined temperature and conductivity sensor. The sonde contains an internal power supply, and was configured to record water quality

data every 500 milliseconds during sampling. The sonde was also put into a rapid system-wide averaging mode during sampling, which is the manufacturer's setting to be used in applications where the sonde is moving.

The total algae sensor provides the capability to simultaneously measure chlorophyll, present in almost all algae, and phycocyanin, which is present in blue-green algae. Both parameters are measured using fluorescence, by emitting a blue light at 470 nm to excite chlorophyll, and a second orange bean at 590 nm to excite phycocyanin. The excited compounds then emit light at a higher wavelength than the initial excitation beam, which is filtered and detected by the probe. Separate datasets for chlorophyll and phycocyanin in relative fluorescence units (RFU) are reported, which can be correlated to the concentration of the respective pigments if necessary. Fluorescence signals from the measurement of chlorophyll and phycocyanin have also been shown to significantly correlate to algae density and biovolume and are often used alongside turbidity for the development of remote sensing models of algae blooms (Canfield Jr. et al. 2019; Johansen et al. 2019; Izydorczyk et al. 2005). However, in this study RFU measurements are used directly, which is strongly recommended by the manufacturer. The temperature and conductivity sensor is used to temperature-correct RFU readings by the algae sensor, and also provides a useful indicator of if the sensor is submerged in water. The turbidity sensor used also relies on light, emitting an 860nm beam with an LED. This light source is then scattered by particles, including algae, in the water, and a photodetector measures the light scattering at a 90˚ angle in accordance with ISO7027. From this measurement, turbidity is recorded in Formazin Nephelometric Unit (FNU).

The second critical component of the drone-sonde integration is the drone itself. For this study, an existing drone created by the UAV MASTER Lab at the University of Cincinnati was utilized.

The drone features an integrated geographic positioning system, 1.4-meter diameter, 6 motors with 24-inch propellers, a 20-minute maximum flight time, and a 44-pound construction weight. This drone was selected because of its large size, which allows the drone to safely manage the sensor as a slug load while remaining stable in the air. The entire system, including the platform, sensor, and drone, was constrained to a maximum weight of 55 lbs. to be classified as a small UAV by the Federal Aviation Administration.

The third component of the sonde-drone integration is a custom designed and built platform to connect the two. The solution needed to provide consistent underwater submergence depth, as water quality parameters can vary greatly with water column depth, but drones generally fluctuate in altitude by a meter or more even when attempting altitude locks. The solution also needed to have the ability to be flown though air and water, provide protection against in-water obstacles invisible to the pilot on land, and easy access to the sonde for cleaning a data collection. In addition, the sonde has sensitive internal electronics and weight-bearing connections or clams can warp the casing and risk water exposure, so the manufacturer suggests that no direct connections are made to the sonde.

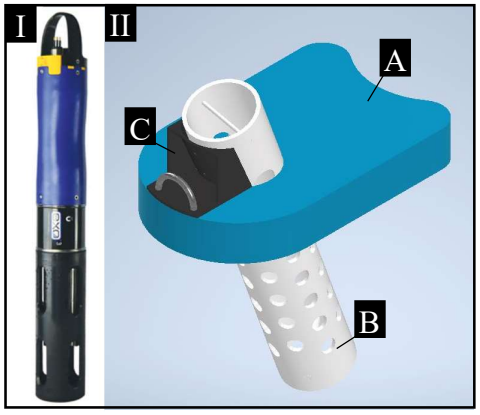


Figure 1: I: YSI Exo 3 Sonde. II: Sonde-Drone Integration Platform, including (A) foam boards, (B) sonde containment, and (C) stabilizer.

The initial platform design is shown in Figure 1, and is created using foam boards, PVC pipe, stainless steel bolts, and 3D-printed PLA, all low-cost and readily available materials. The closed-cell ethylene vinyl acetate (EVA) foam boards provide flotation, and completely support the weight of the sonde in the water. This decreases load for the drone, while ensuring that the sonde maintains a consistent submergence depth. The sonde itself is contained within the 4-inch diameter PVC pipe, which has a series of holes drilled to allow flow to the sensor. Stop bolts are placed at the top and bottom of the pipe, keeping the sonde secure and allowing enhanced flow at the bottom, without containing any direct attachment to the sonde. In addition, the top stop bolt can be easily removed to access the sonde for calibration and maintenance. Most of the weight of the sensor is carried below the board to increase the stability of the platform in the water, and the PVC pipe is able to protect the sensor against any invisible underwater obstacles. The foam boards and PVC pipe are secured using a custom printed black stabilizer piece, which also holds the containment at an angle to minimize drag and any spinning of the platform. The entire platform can be tied to the drone using paracord, connecting to the top stop bolt or U-bolt. The platform was modeled in Autodesk Inventor and evaluated under simulated stress-testing, fabricated at the University of Cincinnati 1819 Makerspace, tested in pool float tests, tested with a weighted sandbag in place of the sensor, and flight tested in an airfield before deployment for fieldwork above water.

2.2 Satellite Data Collection

To allow for the creation of remote sensing models based off water quality data collected with the drone-sonde integration, Landsat and Sentinel satellites were selected as image sources. Both Landsat and Sentinel satellites are currently in orbit and collecting images, with Sentinel 2A and 2B satellites providing a combined revisit time of just 5 days, and Landsat 8 and 9 offset in orbit

Table 1: Spectral bands from Landsat 8 and 9, using the Operational Land Imager 2 (OLI-2) and Thermal Infrared Sensor 2 (TIRS-2) instruments.

Landsat 8 & 9 Spectral Specifications				
Band #	**Name**	**Center Wavelength (nm)**	**Range (nm)**	**Resolution (m)**
B1	Coastal/ Aerosol	443	433-453	30
B2	Blue	482	450-515	30
B3	Green	562	525-600	30
B4	Red	655	630-680	30
B5	NIR	865	845-885	30
B6	SWIR 1	1610	1560-1660	30
B7	SWRI 2	2200	2100-2300	30
B8	Panchromatic	590	500-680	15
B9	Cirrus	1375	1360-1390	30
B10	Thermal	10800	10300-11300	100
B11	Thermal	12000	11500-12500	100

Table 2: Spectral bands from Sentinel 2, using the MulitSpectral Instrument (MSI) instruments.

Sentinel 2 A & B Spectral Specifications				
Band #	**Name**	**Center Wavelength (nm)**	**Range (nm)**	**Resolution (m)**
B1	Aerosol	443	423-463	60
B2	Blue	493	428-558	10
B3	Green	560	525-595	10
B4	Red	665	635-695	10
B5	Red Edge 1	704	690-718	20
B6	Red Edge 2	740	726-754	20
B7	Red Edge 3	783	764-802	20
B8	NIR	833	728-938	10
B8A	NIR narrow	864	843-885	20
B9	Water Vapour	945	926-964	60
B10	Cirrus	1375	1346-1404	60
B11	SWIR 1	1610	1520-1700	20
B12	SWIR 2	2190	2011-2369	20

to provide a combined revisit time of 8 days. As seen in the specifications in Tables 1 and 2, Landsat generally has a lower spatial resolution of 30 meters when compared to Sentinel 2's 10-

meter resolution. While Landsat satellites include two thermal bands, Sentinel Satellites include a narrow near infrared band and three red edge bands with centers ranging from 704nm to 783 nm.

Satellite pass times for the field collection site were predicted using the European space agency (ESA) acquisition calendar for Sentinel 2A and B, and United States Geological Survey (USGS) Landsat acquisition tool for Landsat 8 and 9. Once acquired, satellite imagery from all sources was accessed through the Google Earth Engine Data Catalog, and also visualized directly in Google Earth Engine (GEE). Level 2A orthorectified and atmospherically corrected surface reflectance from the ESA were used for all Sentinel 2 images, and comparably processed Level 2, collection 2, tier 1 atmospherically corrected surface reflectance from the USGS was used for Landsat imagery.

2.3 Water Quality Data Collection

To test the drone-sonde integration platform for collection of water quality in surface waters, 5 field tests were conducted from May to October of 2023. Lake Harsha, of Clermont County, Ohio, was selected for field studies. Lake Harsha is local to Cincinnati, and available for research activities through the Army Corp of Engineers. Lake Harsha is also known to be plagued by recurring harmful algae blooms. Toxic algae blooms caused recreational public health advisories at the main beach on the western half of the lake twice in 2016, once in 2020, and even caused a 41-day long advisory in 2021. At the campground beach on the eastern half of the lake, there were three advisories in 2016, four advisories in 2021, and one in 2022. Lake Harsha is also a popular recreation site and supplies six million gallons of drinking water per day to citizens of Clermont County.

Sites were selected around the lake at areas with clear, flat surfaces conducive to safe drone take-off and landing. For each data collection instance using a drone, personnel included the pilot, ground control, sensor manager, and water safety lookout. When collecting data by drone, the drone took off, flew over the water, dipped the platform for 10-15 seconds while remaining stationary, and then returned to the shore. To supplement data collected by drone, data was also collected by kayaking in two-person teams, one managing the sensor and a second acting as a water safety lookout. The integration system was still used for data collection by kayak, with the platform towed behind the boat. All sampling dates were organized to fall within 48 hours of a Landsat or Sentinel satellite passing over the area, with priority given to dates where both satellites passed over the lake.

2.4 Data Preprocessing

From the fieldwork, two distinct datasets were collected. The first is water quality, collected from the sonde. The collected data includes measurements of chlorophyll and phycocyanin concentrations in RFU, turbidity in the Formazin Nephelometric Unit (FNU), temperature, and conductivity. Each measurement is also correlated with a timestamp, and a general location note for the site of sampling. This data was accessed via YSI Kor software and exported as a csv file. Specific locations of the sampling are stored separately, as the sonde itself does not contain an internal GPS unit. When using a drone, time, latitude, and longitude are recorded by the drone's GPS unit and accessed through the drone autopilot log as a .gpx file. When data was collected by kayak, cell phones were used to track location, which were be accessed by exporting .gpx files from the phones.

In order to analyze the satellite reflectance at the points with collected water quality, the water quality data and position data first had to be correlated. This was done for each date of sampling,

to facilitate later reflectance collection from satellite imagery. Python was chosen as the core programming language to achieve this correlation, because of its capability to manage and process large datasets using the pandas library. To begin, water quality data was transferred into python from the sonde .csv file. While the sonde constantly records readings, only those recorded when the sonde was submerged in water are of interest. The conductivity parameter is useful to identify this, as there is a sharp cutoff when the sonde is in air versus in water. Therefore, the water quality datasets were first trimmed to include only points with a conductivity above 100 µS/cm. In addition, water quality readings are recorded every 500 ms, yielding two readings per second. In contrast, location datapoints are recorded to a precision of whole seconds. Therefore, to avoid duplicated records water quality readings were averaged within the same second. At this point, for each date of sampling there was a pandas dataframe including water quality parameters and associated time.

To correlate to positions, .gpx files of timestamped locations were read into python dataframes using the convenient gpx_converter package for python. For location data from both the drone and kayak, the timestamps were adjusted to be in Eastern Daylight Time. In the case of the drone position data collected on September 11th, conductivity and stationary positions did not initially align, as seen in Figure 2. This is likely because the sonde adjusts its time to the laptop it was deployed with, while the drone time stamp is generated based off the last system reboot and connected to a different computer. In this case, the data was manually analyzed and aligned, using a 2.75-minute time offset on the drone timestamp.

Using the time-corrected position datasets and timestamped water quality datasets, a merge function with an inner join was performed to find save only the rows with records from each dataset. At this point, each sampling date had a position-quality dataframe containing datetime,

latitude, longitude, general site name, chlorophyll, conductivity, phycocyanin, turbidity, and temperature. Latitude and longitude points were exported into a shapefile for analysis of satellite imagery.

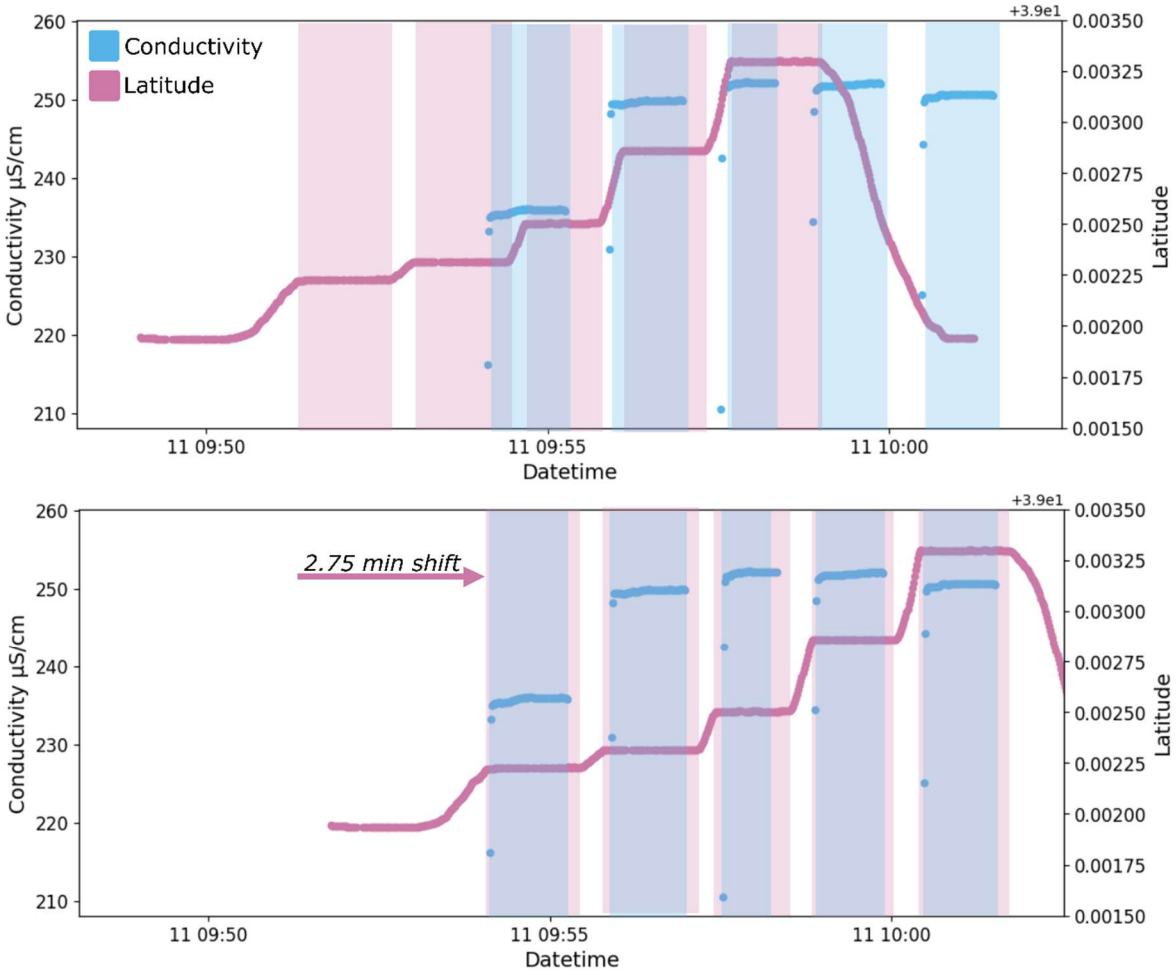


Figure 2: Quality and Position data alignment. Blue points correspond to water conductivity readings and indicate time frames when the sonde is submerged. Horizontal segments of the pink latitude points indicate that the drone was hovering at a stationary point while dipping the sonde. The top graph shows mismatched data before manual alignment, while the bottom shows time-offset position data used for continued analysis.

At this point, reflectance was gathered using google earth engine, which is better suited to spatial image analysis. Shapefiles of collected data for each date were imported as assets into the project, and then used as the input for both Landsat and Sentinel analysis. Satellite images from the google earth engine catalog were searched for images of Lake Harsha withing 48 hours of the sample date, loaded, and masked for cloud cover. From these images, reflectance values for all available bands at each water sampling point were extracted and exported as a .csv file back into python.

The reflectance and position-quality datasets were then merged using both the latitude and longitude columns as keys, including only the rows contained in both data frames. Some data is lost at this point, due to cloud cover. After the merge, there are cases where identical reflectance values for all bands are associated with multiple water quality values, which occurs when water quality readings were collected within a pixel of the satellite imagery. To achieve a one-to-one correlation of the reflectance matrix to water quality, any rows with identical values for every reflectance band were combined by averaging the water quality readings for those rows.

At the end of pre-processing there was a quality-reflectance dataset for the May 4th sampling using Sentinel reflectance, a dataset from July 27th using Landsat reflectance, and two datasets each for September 2nd, September 11th, and October 10th using either Sentinel and Landsat reflectance. The coastal/ aerosol, cirrus, water vapor, and thermal targeted bands were not considered in analysis for either data set.

2.5 Single Linear Regressions

To begin regressions correlating reflectance to water quality, data across all dates was combined for the Sentinel derived reflectance sets, and separately for Landsat derived reflectance sets.

Then, indexes, or algebraic combinations of reflectance bands, were calculated for each row of data with associated water quality. Indexes were calculated with three general formulae for independent combination of bands: band ratios (equation 1), normalized differences (equation 2) and additive combinations (equation 3).

$$\frac{band\ A}{band\ B} \tag{1}$$

$$\frac{band\ A - band\ B}{band\ A + band\ B} \tag{2}$$

$$\frac{band\ A}{band\ B + band\ C} \tag{3}$$

For the Landsat datasets, this resulted in 6 reflectance bands, 15 reflectance band ratios, 15 normalized differences and 47 additive combinations. For the Sentinel datasets, this resulted in 10 reflectance bands, 45 reflectance band ratios, 45 normalized differences, and 362 additive combinations.

For each water quality parameter of chlorophyll, phycocyanin, and turbidity, a linear regression was run using every band using the Scipy.stats.linregress package. For each regression, a slope, intercept, R^2 value, root mean squared error (RSME), and p-value were recorded, which are all direct outputs from the regression package. The p-value is that of a hypothesis test whose null hypothesis is that the slope is 0, using a Wald Test with a t-distribution of the test statistic.

The results were then ranked by R^2 value for each water quality parameter to identify the index best correlated to that water quality parameter, and the regression using the best index visualized.

2.6 Multiple Linear Regressions

To test more complex models, multiple linear regressions using more than one index concurrently were also evaluated. In this case, overfitting was a concern, so data was split into testing and training sets, with 30% used for the testing of all data. In addition, highly correlated indexes used within the same model are undesirable, so multilinear regressions were evaluated for each water quality parameter with four separate groups of feature inputs: reflectance bands, band ratios, normalized differences, and additive combinations.

To minimize overfitting and the number of features used within each model, an ElasticNet regression strategy from the sklearn package was used. ElasticNet regression is a combination of ridge and lasso regressions, which are methods of multiple linear regressions that act to regularize models and decrease variance. Ridge regression seeks to minimize the sum of squared residuals and some lambda multiplied by the slope squared. Lasso regression is extremely similar, but uses an added penalty of a lambda multiplied by the absolute value of the slope instead of the slope squared. ElasticNet regression is able to identify and reduce the number of parameters by including both penalties and is useful for situations where there may be correlated parameters. For each model, the parameters for the ElasticNet regression were tuned using a grid search with l1 ratios of 0.1, 0.5, 0.8, 0.9, 0.95, 0.99, and 1, and alphas of 0.05, 0.1, 0.15, 0.3, 0.5, and 0.7.

Once trained, a R^2 value and root mean squared error (RSME) value from the predicted testing and training data and compared. The testing and training data was then recombined to calculate overall metrics and fit a regression line through the predicted and actual water quality parameters before visualizing. This yielded 4 multilinear models for each water quality parameter for each satellite reflectance source, for 24 models total.

3. Results and Discussion

3.1 Drone Sonde Integration

The drone platform designed and built in this study was able to successfully collect water quality data on six different flights and three locations at Lake Harsha, seen in Figure 3. On each flight, the drone took off from the shore, flew over obstacles to the water, dipped the sensor into the water while hovering for 10-15 seconds, repeated dips at 5-8 locations, and returned to the takeoff location. The platform supported the weight of the sonde while in the water, ensuring a consistent submergence at all sampling points. The platform also has the added benefit of increasing visibility for the pilot, as the paracord and sensor are difficult to see.

Figure 3: Drone-sonde integration in use at Lake Harsha

When compared to sampling by boat, sampling by drone does require a larger team and trained pilots to fly. However, the sampling takes less time than by boat, is less strenuous, and allows the sampling to be more flexible in location. The integration platform itself is also extremely affordable, costing less than $75 in materials. It is also easy to construct with low skill and common tools, the most advanced of which being a 3D printer which are found at many universities and public libraries. Using this integration platform could allow water quality sampling in hard-to-reach areas such as water bodies surrounded by cliffs or wetlands and allow a greater understanding of algae blooms or other water quality parameters to be developed in these areas. In addition, the YSI sonde central to the

integration has a flexible system with replaceable sensors, which could allow monitoring of water quality parameters beyond algae and turbidity.

However, the platform could still be improved. This integration lacked a GPS sensor integrated into the water quality sensor itself, which resulted in extensive preprocessing required to map water quality to specific locations. Matching timestamps from separate water quality and location datasets was used to do so, but undoubtedly introduced error into analysis, as the data sources use different time systems and sources of "truth". This was exemplified in the data collected by drone on September 11th, where the water quality data and location data were clearly mismatched and had to be offset by over 2 minutes. In addition, there is some position offset between the drone and the platform, as they are not always perfectly vertically aligned. When collecting by kayak, this error is likely even greater, as the kayak and associated GPS is consistently ahead of the sensor platform. The location accuracy could be improved by adding a GPS directly onto the platform, and by recording the timestamp from both data sources instantaneously at the start of sampling. Alternatively, if a multispectral camera was attached to the drone and imaged at the same time as the water quality, the position of the platform could be derived from the image knowing the exact location and altitude of the drone.

While the drone was able to successfully dip the platform to collect data, it was not able to tow the platform through the water. Although the platform sits level in the water while still, the nose begins to dip when in motion, greatly increasing drag and stress on the drone. This severely limits the amount of data that the drone was able to collect, and prevents the use of the platform in rivers with strong currents. To attempt to address this problem, a second iteration of the design was completed and constructed with a center of mass further back on the platform, and the connection point was adjusted to attach to the front bottom edge of the platform. While this

improved the performance, the drag was still too large for the drone to safely manage. A design that allows successful towing of the platform would need to be extremely smooth, as the risk is very high for testing the drone above water. In future designs, the platform could be made longer, with a sloped bottom to prevent the nose from submerging. In addition, drag could also be reduced by optimizing the angle, hole pattern, and diameter of the pipe used to contain the sensor. While the pipe design and hole pattern were recommended by YSI for dock deployments, a sensitivity analysis of the pipe color and flow patterns could also be conducted experimentally to increase confidence in the accuracy of collected water quality.

3.2 Water Quality Data

From the combined drone and kayak sampling, chlorophyll, phycocyanin, and turbidity data was collected across Lake Harsha over five dates. Sampling sites were distributed across the lake and included areas near the dam on the west side of the lake as well as the eastern half. Data was preprocessed using both Sentinel and Landsat imagery for each date to acquire a correlated water quality dataset, resulting in slightly different datasets each satellite source due to cloud cover and image availability.

The sample spatial distributions are visualized in Figure 4. For the May 4th sampling, shown in dark orange in Figure 4, only Sentinel imagery was available. For the second sampling on July 27th, both satellites took images, but those from Sentinel were completely obscured by clouds for all sampling locations and yielded no usable data. Sampling completed on September 2nd had both Sentinel and Landsat imagery available, although cloud cover in the Landsat imagery obscured some sampling points at the swim beach and Reisinger. For the final two sampling dates by drone, both satellite sources were available, but much smaller quantities of data were collected due to the dipping technique used by the drone.

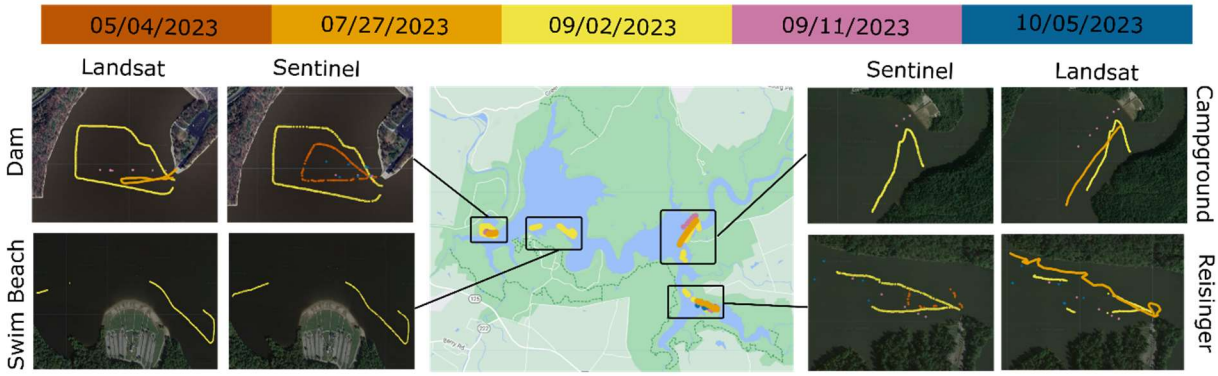


Figure 4: Water quality-reflectance dataset locations by date, location, and satellite source.

As algae blooms can develop rapidly over time, the time difference between sampling and the satellite image capture was examined and is presented in Table 3. Images were only considered if they were within 2 days of the water quality sampling, and both satellites fly over Lake Harsha in the morning or just after noon. As sampling was also conducted in the morning, many datapoints were collected within four hours of sampling if on the same day. However, other data was collected as far as 26 hours after the water quality sampling. Not including the July 27th sampling with no usable data, the time lag between the Sentinel sampling and satellite image capture was an average of 7.68 hours, while that of Landsat was 12.54 hours, much larger.

Table 3: Time difference between water quality sampling and satellite image capture.

Method	Location	Sample Date Time	Sentinel 2A or 2B		Landsat 8 or 9	
			Image Capture	Difference (hours)	Image Capture	Difference (hours)
Kayak	Reisinger	5/4/23 11:30	5/4/23 9:27	2	N/A	N/A
	Dam	5/4/23 13:30		4	N/A	N/A
	Reisinger	7/27/23 9:00	7/28/23 9:20 (full clouds)	24.25	7/26/23 12:10	20.75
	Campground	7/27/23 10:00		23.25		21.75
	Dam	7/27/23 12:15		21		24
	Reisinger	9/2/23 10:00	9/1/23 9:24	0.5	9/3/23 12:17	26.25
	Campground	9/2/23 10:45		1.25		25.5
	Dam	9/2/23 13:30		4		22.75
	Swim Beach	9/2/23 14:00		4.5		22.25
Drone	Reisinger	9/11/23 10:00	9/11/23 9:31	0.5	9/11/23 12:17	2.25
	Campground	9/11/23 11:15		1.75		1
	Dam	9/11/23 12:30		3		0.25
	Reisinger	10/5/23 11:00	10/6/23 9:25	22.5	10/5/23 12:17	1.25
	Campground	10/5/23 10:00		23.5		2.25
	Dam	10/5/23 8:45		24.75		3.5

Overall, 848 datapoints with water quality and reflectance across the sampling dates were collected for the Sentinel dataset, and 3108 datapoints were collected using Landsat reflectance values. The distribution of each water quality parameter is visualized in stacked histograms in

Figure 5. Measured chlorophyll RFU had the most normal distribution, and ranged from 0.75 RFU to 11.33 RFU for Sentinel and from 1.23 RFU to 13.12 RFU for Landsat. Phycocyanin has multiple distinct peaks for both datasets, and ranged from 0.01 RFU to 4.375 for Sentinel and from 1.61 to 7.26 for Landsat. The distinct points observed in the phycocyanin and turbidity datasets are largely due to the different locations sampled even on the same day, which have different water quality. Turbidity data ranged from 3.74 FNU to 28.7 FNU for the Sentinel dataset, and from 3.99 FNU to 27.05 FNU for the Landsat dataset. Across all the distributions, data collected on the first three sampling dates makes up the majority of the points, as a kayak was used with continuous collection.

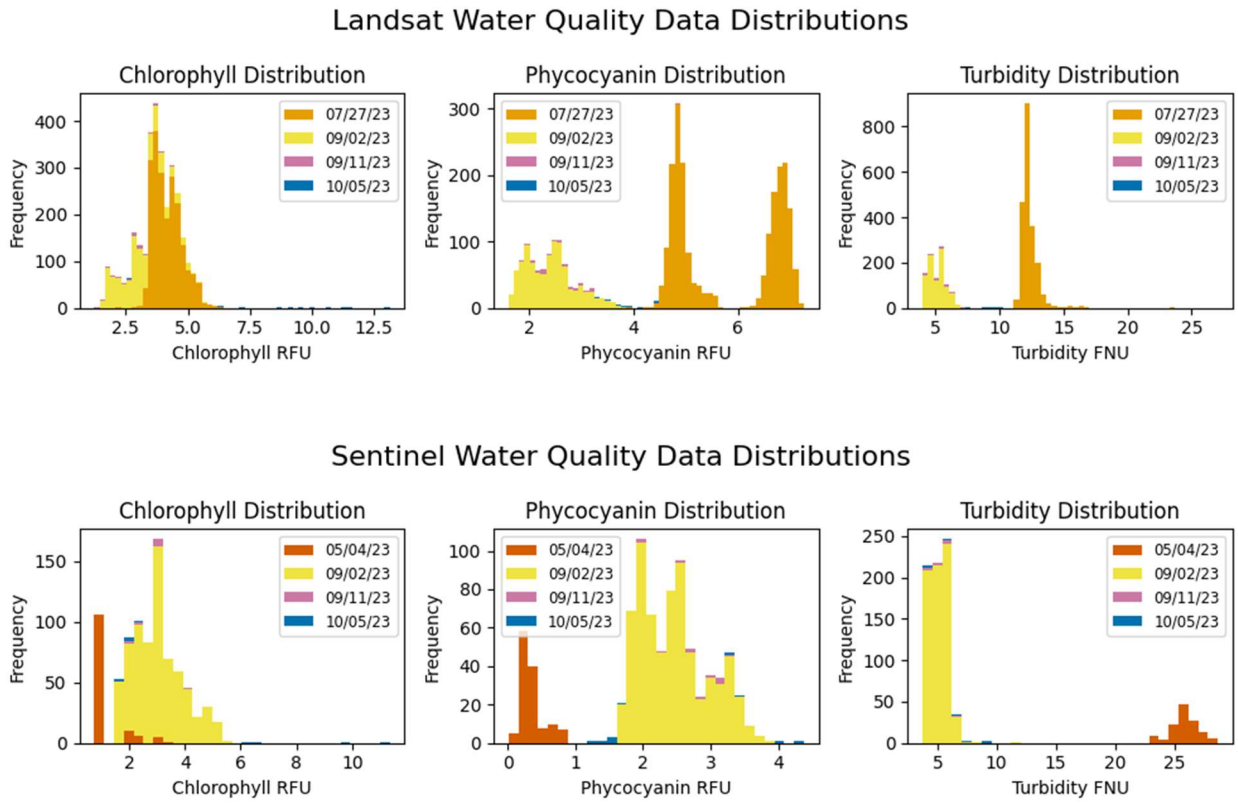


Figure 5: Distribution of collected chlorophyll, phycocyanin, and turbidity data.

3.3 Remote Sensing of Water Quality

To examine the potential indexes or bands most useful for algae bloom detection, single linear regressions were first utilized on all the bands and band combinations. The top five best performing indexes, ranked by their R^2 and RSME vales, are tabulated for each model in Figure 6. Every model identified in this top group had p-values much smaller than 0.0001, indicating that relationships between the indexes and water quality parameters are significant.

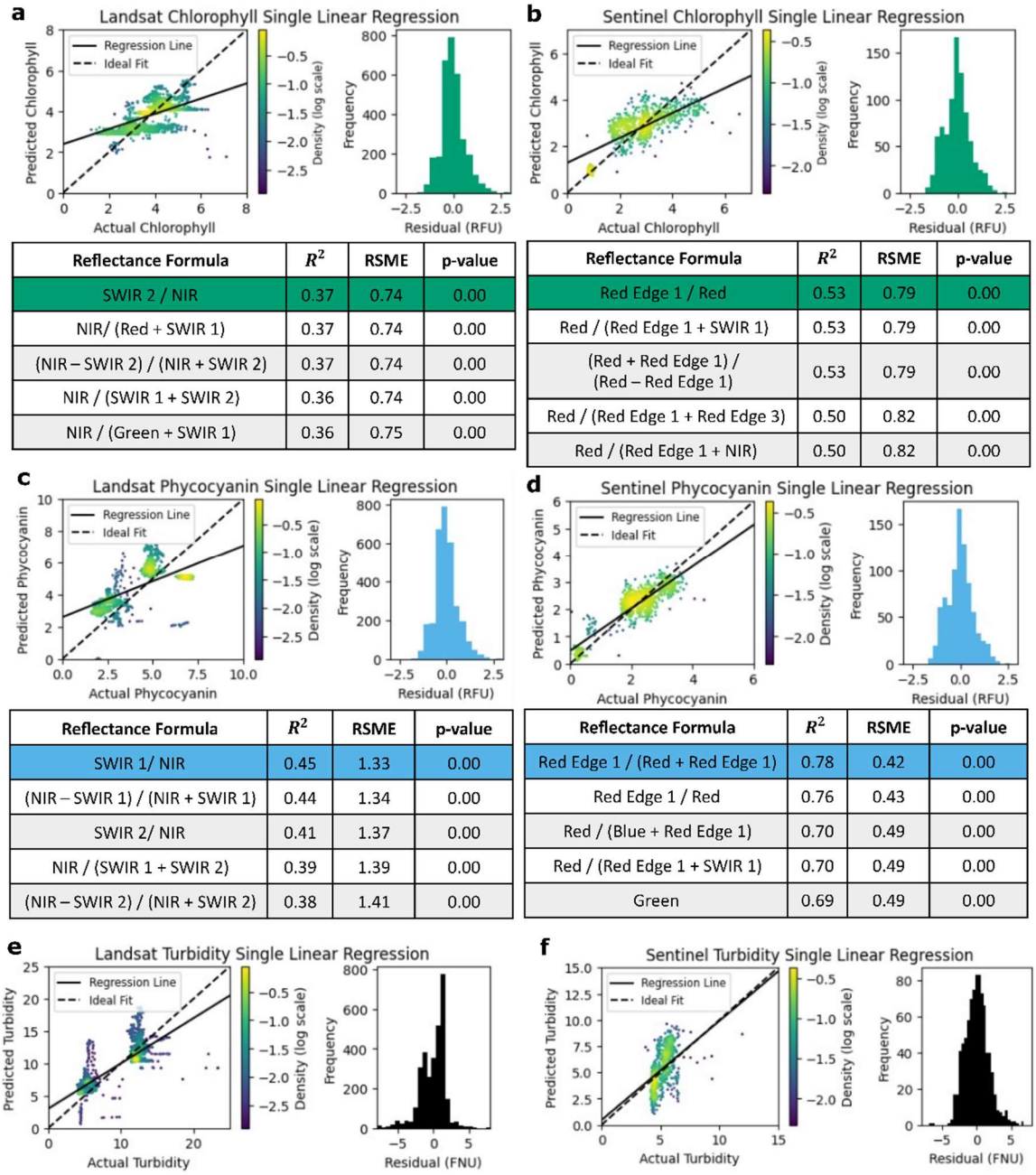
When using the same satellite source, very similar bands or band combinations were most highly correlated to all three water quality parameters, as seen when examining the tabulated indexes in each column of Figure 6. This trend is supported logically, as cyanobacteria contain both phycocyanin and chlorophyll, so these two parameters should be highly correlated. Although turbidity can be from several factors, it too can be impacted by algae growth. In addition, highly turbid water due to erosion can indicate higher levels of nitrogen or phosphorus in the water, which fuel algae growth.

In Landsat models, short wave infrared bands (SWIR) 1 and 2, as well as the near infrared band (NIR) were most highly correlated with chlorophyll, phycocyanin, and turbidity. The identification of the SWIR / NIR ratio align with the results of Rivani et al, who also found that a ratio of SWIR 1 to NIR bands were the most correlated to chlorophyll in Maninjau lake (A. Rivani and P. Wicaksono 2018). However, the coefficient of determination values for the chlorophyll and phycocyanin regressions were low for the Landsat models, at only 0.37 and 0.45, indicating that they are not highly reliable models. In addition, the visualization of the chlorophyll regression, in Figure 6a, shows a strong pattern of horizontal lines, which may indicate that other unidentified features could be important in identification of chlorophyll.

Sentinel models for chlorophyll and phycocyanin, visualized in Figure 6b and 6d, had improved coefficients of determination, at 0.53 and 0.78, respectively. In both cases, band combinations including the red and red edge 1 bands were most highly correlated, aligning with the work of Meng et al. monitoring chlorophyll in another inland reservoir (Meng et al. 2022). This red to red edge band ratio is also commonly used to monitor coastal algae blooms, where it is also among the most effective remote monitoring techniques (Rodríguez-Benito et al. 2020; Caballero et al. 2020; Jordan et al. 2021).

Turbidity correlations from Landsat models, visualized in Figure 6e, relied on the same NIR, SWIR 1, and SWIR 2 bands used for the phycocyanin and chlorophyll regressions. This is unusual and could be a result of strong correlation between algae and turbidity and in the dataset. In contrast, the Sentinel regressions shown in Figure 6f utilized red, blue, and green bands found as common predictors of turbidity in literature (Rodríguez-López et al. 2021).

By using multiple linear regression and including more than one feature at a time, the model accuracy for all water quality parameters were able to be improved, as shown in figure 7. Linear regression combinations using either reflectance, band ratios, normalized differences, or additive combinations were considered for all water quality parameters, using an ElasticNet regression to minimize features. Data was also split into testing and training sets to evaluate potential overfitting. Although some models still use up to 100 features, the RSME and R^2 values were still very similar between testing and training datasets, with differences of less than 0.09 RFU for RSME and less than 0.04 for R^2 in for the largest model. This indicates that the models are not overfit, as overfit models would yield much better coefficients of determination and lower errors in training data than in test data.



Reflectance Formula	R^2	RSME	p-value
SWIR 2 / NIR	0.37	0.74	0.00
NIR/ (Red + SWIR 1)	0.37	0.74	0.00
(NIR – SWIR 2) / (NIR + SWIR 2)	0.37	0.74	0.00
NIR / (SWIR 1 + SWIR 2)	0.36	0.74	0.00
NIR / (Green + SWIR 1)	0.36	0.75	0.00

Reflectance Formula	R^2	RSME	p-value
Red Edge 1 / Red	0.53	0.79	0.00
Red / (Red Edge 1 + SWIR 1)	0.53	0.79	0.00
(Red + Red Edge 1) / (Red – Red Edge 1)	0.53	0.79	0.00
Red / (Red Edge 1 + Red Edge 3)	0.50	0.82	0.00
Red / (Red Edge 1 + NIR)	0.50	0.82	0.00

Reflectance Formula	R^2	RSME	p-value
SWIR 1/ NIR	0.45	1.33	0.00
(NIR – SWIR 1) / (NIR + SWIR 1)	0.44	1.34	0.00
SWIR 2/ NIR	0.41	1.37	0.00
NIR / (SWIR 1 + SWIR 2)	0.39	1.39	0.00
(NIR – SWIR 2) / (NIR + SWIR 2)	0.38	1.41	0.00

Reflectance Formula	R^2	RSME	p-value
Red Edge 1 / (Red + Red Edge 1)	0.78	0.42	0.00
Red Edge 1 / Red	0.76	0.43	0.00
Red / (Blue + Red Edge 1)	0.70	0.49	0.00
Red / (Red Edge 1 + SWIR 1)	0.70	0.49	0.00
Green	0.69	0.49	0.00

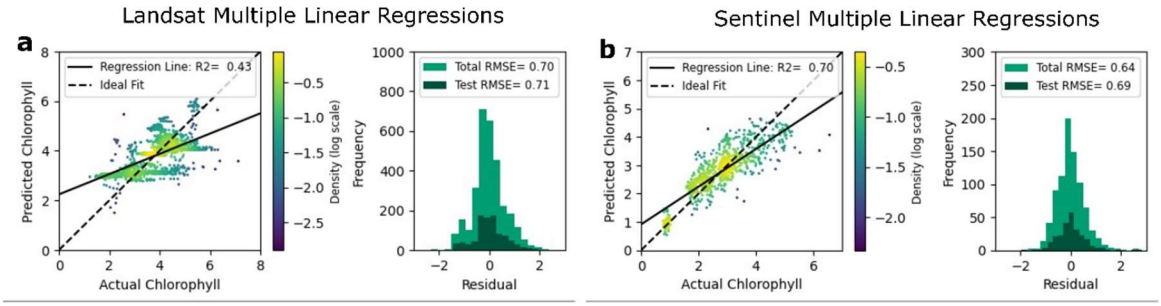
Reflectance Formula	R^2	RSME	p-value
SWIR 2 / NIR	0.70	1.89	0.00
(NIR – SWIR 1) / (NIR + SWIR 1)	0.69	1.92	0.00
SWIR 1/ NIR	0.68	1.94	0.00
NIR / (SWIR 1 + SWIR 2)	0.68	1.95	0.00
(NIR – SWIR 2) / (NIR + SWIR 2)	0.67	1.98	0.00

Reflectance Formula	R^2	RSME	p-value
Red / Green	0.94	1.76	0.00
Red / (Green + Red Edge 1)	0.94	1.83	0.00
Red	0.94	1.89	0.00
Red/ (Blue + Green)	0.93	1.97	0.00
Red / (Green + SWIR 2)	0.92	2.07	0.00

Figure 6: Best performing indexes for prediction of water quality parameters with single linear regression.

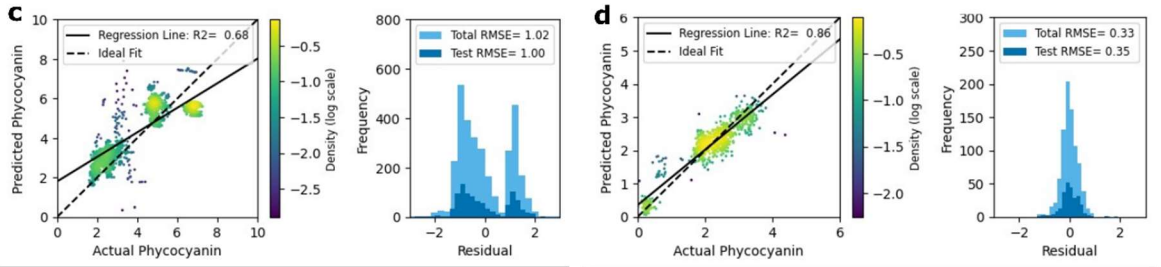
For each multiple linear regression model, the features with the largest coefficients, which are the most important, were ranked. For the best Landsat chlorophyll model, visualized in Figure 7a, which used additive band combinations to achieve an in R^2 value of 0.43, the [red / (green + blue)] index had the highest coefficient, with [red / (blue + SWIR 2)] and [red / (green + NIR)] following as the next most important indexes in the model. This is interesting, because the [red / (green + blue)] index was not identified as a highly correlated ratio in the single linear regressions. However, ratios using red, green, and blue bands to identify chlorophyll do appear in literature (Boucher et al. 2018). The most indexes with the highest coefficients for the multiple linear regression of phycocyanin, shown in Figure 7c, also included red, blue, and green bands, but in relation to the SWIR 1 and SWIR 2 bands identified in the linear regressions. Importantly, the indexes were different than those for the chlorophyll regressions, showing that more complex models with multiple features may be able to better differentiate between chlorophyll and phycocyanin.

In Sentinel models, chlorophyll was also predicted with blue, green, and red bands, but also incorporated the important red edge band seen in single linear regressions, as well as short wave infrared bands. The best model in this case achieved an R^2 of 0.70, shown in Figure 7b. Again, phycocyanin models used similar bands, also including red, green, blue, and red edge bands, but in different combinations. The best model for phycocyanin utilized additive combinations of bands and achieved an R^2 of 0.86, visualized in Figure 7d. In this case, the improvement over the Landsat model is likely partially due to the difference in phycocyanin distributions input into the models. The Landsat model had a wider distribution of phycocyanin, but the model fails to correctly predict phycocyanin at higher RFU values, and underpredicts these high values. In



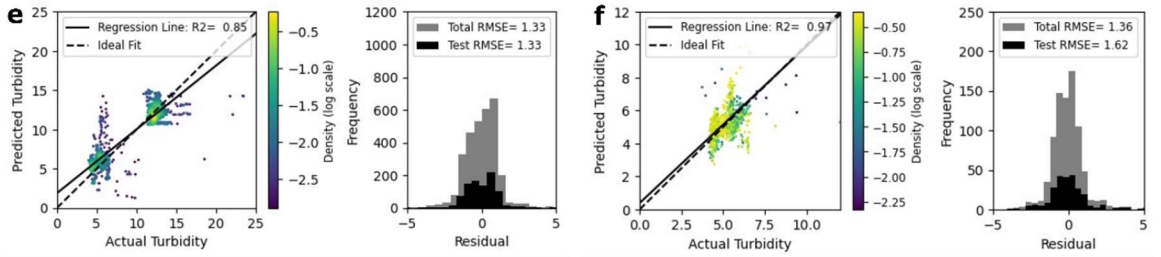
Landsat Chlorophyll Models	# Features	R^2	RSME
Reflectance Bands	6	0.34	0.76
Band Ratios	13	0.43	0.70
Normalized Differences	11	0.41	0.72
Additive Combinations	26	0.43	0.71

Sentinel Chlorophyll Models	# Features	R^2	RSME
Reflectance Bands	10	0.53	0.79
Band Ratios	29	0.64	0.69
Normalized Differences	28	0.63	0.71
Additive Combinations	100	0.70	0.64



Landsat Phycocyanin Models	# Features	R^2	RSME
Reflectance Bands	5	0.51	1.25
Band Ratios	10	0.56	1.19
Normalized Differences	14	0.56	1.18
Additive Combinations	39	0.68	1.02

Sentinel Phycocyanin Models	# Features	R^2	RSME
Reflectance Bands	9	0.80	0.40
Band Ratios	30	0.83	0.37
Normalized Differences	22	0.80	0.40
Additive Combinations	75	0.86	0.33



Landsat Turbidity Models	# Features	R^2	RSME
Reflectance Bands	5	0.76	1.70
Band Ratios	7	0.77	1.65
Normalized Differences	14	0.75	1.72
Additive Combinations	43	0.85	1.33

Sentinel Turbidity Models	# Features	R^2	RSME
Reflectance Bands	6	0.97	1.36
Band Ratios	10	0.95	1.26
Normalized Differences	8	0.94	1.86
Additive Combinations	218	0.97	1.26

Figure 7: Best performing indexes for prediction of water quality parameters with multiple linear regression. Statistics are calculated using combined testing and training data after training.

contrast, the Sentinel dataset lacked these higher values above 6 RFU and still performs well in the low RFU range.

In the Landsat turbidity model, the most effective multilinear regression also used additive combinations of bands, with an R^2 of 0.85 visualized in Figure 7e. With multiple features, green, red, and NIR bands were among the most common in indexes with high coefficients, which aligns more with literature than the single linear regressions. For the Sentinel turbidity model, shown in Figure 7f, a model with only the reflectance bands was able to predict turbidity as well as models with additive combinations. Although the reflectance band model has a slightly higher error, it uses many times less features, making it more attractive. With single reflectance bands, red had by far the highest coefficient, and 6 of the 10 available features were used in the final model.

In nearly every model tested, reflectance values derived from Sentinel outperform those from Landsat. There are several contributing factors that could be the case in these two datasets. The first is simply the spatial resolution of the two satellites. Most Landsat bands have a spatial resolution of 30 meters, while Sentinel bands have spatial resolutions of 10 and 20 meters. This means that water quality within the grids can be more closely matched to reflectance values, and better models can be created. Secondly, in this dataset there was a larger time lag between the water quality sampling and images captured by Landsat, as established in Table 3. As algae blooms can develop rapidly over days, and move due to wind or precipitation patterns, these time lags could be causing substantial differences in the water quality over time. However, the difference in model performance could also be due to the different bands available for use. Landsat does not have any red edge bands, which were among the most important in the Sentinel

models. This indicates that expanding the available spectral bands can lead to more complex models that are able to more accurately predict algae concentrations.

In this study, only simple single and multiple linear regressions were used to predict algae blooms, although more advanced machine learning models are used elsewhere (C. Yang et al. 2023). While machine learning models shine in scenarios with large amounts of data balanced over a wide range, the data collected in this study is relatively small in scale and range, and therefore not well suited to advance machine learning strategies. However, with larger amounts of data collection over longer time periods, machine learning techniques could be applied.

It is worth noting that chlorophyll and phycocyanin are proxy pigments, and although they can be correlated to algae biomass or cell counts with field measurements at specific sites, they cannot be used to quantify toxicity (Bastien et al. 2011). Even among cyanobacteria, it is estimated that only 25% to 75% of cyanobacteria blooms are toxic (Bláha et al. 2009). There are also many different cyanobacterias, such as *Microcystis*, *Schizotrix*, and *Cylindrospermopis*, which produce different toxins that are dangerous at different levels of cyanobacterial biomass. So, while remote sensing models can be useful as early-warning systems to determine the location and extent of harmful algae blooms, water samples and more advanced analytical techniques are still needed to quantify toxicity (Stumpf et al. 2016). With this purpose in mind, simple, affordable models still retain value as useful tools for early-warning systems and can be more accessible to analysts and decision makers. The models presented here, especially using the Sentinel satellite, are accurate enough to be used as an early warning system to prompt more rigorous analytical testing as needed to make decisions to control the impact of algae blooms.

3.4 Application to Camera-based Drone Monitoring Systems

One of the major difficulties with satellite-based monitoring is the low temporal and spatial resolutions. These limitations make it difficult to track the development of algae blooms over time, or collect timely data if needed urgently, such as by a water municipality managing a water intake. While the drone-sonde integration can increase accessibility to gather water quality data in hard-to-reach areas and then be used to develop satellite-based remote sensing models, it also has extremely high potential to be used with imagers carried by drones. If the drone carrying the sonde is also equipped with a multispectral or hyperspectral imager, images can be taken at the same time as water quality data collection, minimizing error due to time lags and low spatial resolution and eliminate lost data from cloud cover, without requiring larger teams of personnel than water quality data collection alone.

Today, many commercially available multispectral cameras are available that can either be easily mounted to or are already integrated with UAVs, and are presented in Table 4. Many of these drones are intended for agricultural purposes, but include the red, red edge, and NIR bands that can be useful for monitoring algae blooms with the same indexes as those used by Sentinel 2 satellites. However, these cameras have the added benefit of much higher spatial resolutions. For example, the Sentera 6X Multispectral Sensor has a 1-foot ground sampling distance when flown at a height of 200 ft. With this resolution, chlorophyll and phycocyanin concentrations could be mapped much more closely to the image reflectance.

In addition, more advanced multispectral and hyperspectral cameras with a wider range of available reflectance bands are becoming available. Examples include the MapIR Kernal 2, currently under development, which will allow purchasers to select 6 specific

Table 4: Commercially available multispectral cameras with potential for UAV-based HAB monitoring. Spectral bands as listed as the band center x the full width at half maximum (FWHM).

		Spectral Band Centers				
Camera	**# of Bands**	**Blue**	**Green**	**Red**	**Red Edge**	**NIR**
Sentera 6X Multispectral Sensor	5	475 x 30	550 x 20	670 x 30	715 x 10	840 x 20
Parrot Sequoia+ Multispectral Sensor	4		550 x 40	660 x 40	735 x 10	790 x 40
DJI Mavic 3M	4		560 x 16	650 x 16	730 x 16	860 x 26
DJI P4 Multispectral	5	450 x 16	560 x 16	650 x 16	730 x 16	840 x 26
MapIR Kernal 2	6	342-915 (x 15)				
Cubert Ultris 5 Hyperspectral	51	540-850 (x 26)				

wavelengths. This could be extremely beneficial for algae monitoring, as the users could reap the benefits of hyperspectral monitoring with targeted features such as 620 nm for detection of phycocyanin, without paying the cost of a hyperspectral camera. Another option is the Cubert Ultris 5 hyperspectral camera, which is the smallest of those listed, and uses a snapshot image capture to simultaneously capture 51 spectral bands ranging from 540-850nm. The large amount of data collected by this camera could make it an excellent candidate for the development of advanced machine learning algorithms.

4. Conclusion

In this thesis, a new technique to monitor algae blooms using integrated unmanned aerial vehicles and water quality sensors was designed, tested, and deployed. The platform was successfully utilized to collect chlorophyll, phycocyanin, and turbidity data in Lake Harsha, Ohio through the summer of 2023. Data was correlated with Landsat 8 and 9 and Sentinel 2 imagery, and the improved capabilities of Sentinel to monitor harmful algae blooms due to higher spatial

resolution and the availability of narrow red-edge bands from 704nm to 783nm was demonstrated. Useful bands were identified, and commercially available multispectral cameras capable of integration with UAV were evaluated. By applying improved monitoring techniques for HABs that utilize unmanned aerial vehicles, researchers and water quality managers can build an understanding of the larger patterns of algae blooms in surface waters with less resources. This can enable more effective and timely control of algae bloom impacts, and protect surface water as safe sites of outdoor recreation and critical drinking water supply.

References

A. Rivani, and P. Wicaksono. (2018). "Water Trophic Status Mapping of Tecto-Volcanic Maninjau Lake during Algae Bloom using Landsat 8 OLI Satellite Imagery." *2018 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES),* 1-7.

Bastien, C., Cardin, R., Veilleux, E., Deblois, C., Warren, A., and Laurion, I. (2011). "Performance evaluation of phycocyanin probes for the monitoring of cyanobacteria." *Journal of Environmental Monitoring,* 13(1), 110-118.

Bláha, L., Babica, P., and Maršálek, B. (2009). "Toxins produced in cyanobacterial water blooms - toxicity and risks." *Interdiscip.Toxicol.,* 2(2), 36-41.

Boucher, J., Weathers, K. C., Norouzi, H., and Steele, B. (2018). "Assessing the effectiveness of Landsat 8 chlorophyll a retrieval algorithms for regional freshwater monitoring." *Ecol.Appl.,* 28(4), 1044-1054.

C. Yang, Z. Tan, Y. Li, M. Shen, and H. Duan. (2023). "A Comparative Analysis of Machine Learning Methods for Algal Bloom Detection Using Remote Sensing Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 16 7953-7967.

Caballero, I., Fernández, R., Escalante, O. M., Mamán, L., and Navarro, G. (2020). "New capabilities of Sentinel-2A/B satellites combined with in situ data for monitoring small harmful algal blooms in complex coastal waters." *Scientific Reports,* 10(1), 8743.

Canfield Jr., D. E., Bachmann, R. W., Hoyer, M. V., Johansson, L. S., Søndergaard, M., and Jeppesen, E. (2019). "To measure chlorophyll or phytoplankton biovolume: an aquatic conundrum with implications for the management of lakes." *Lake Reserv.Manage.,* 35(2), 181-192.

Carmichael, W. W., and Boyer, G. L. (2016). "Health impacts from cyanobacteria harmful algae blooms: Implications for the North American Great Lakes." *Harmful Algae,* 54 194-212.

Chaffin, J. D., Bratton, J. F., Verhamme, E. M., Bair, H. B., Beecher, A. A., Binding, C. E., Birbeck, J. A., Bridgeman, T. B., Chang, X., Crossman, J., Currie, W. J. S., Davis, T. W., Dick, G. J., Drouillard, K. G., Errera, R. M., Frenken, T., MacIsaac, H. J., McClure, A., McKay, R. M., Reitz, L. A., Domingo, J. W. S., Stanislawczyk, K., Stumpf, R. P., Swan, Z. D., Snyder, B. K., Westrick, J. A., Xue, P., Yancey, C. E., Zastepa, A., and Zhou, X. (2021). "The Lake Erie HABs Grab: A binational collaboration to characterize the western basin cyanobacterial harmful algal blooms at an unprecedented high-resolution spatial scale." *Harmful Algae,* 108 102080.

Corcoran, A. A., and Hunt, R. W. (2021). "Capitalizing on harmful algal blooms: From problems to products." *Algal Research,* 55 102265.

Dittmann, E., Fewer, D. P., and Neilan, B. A. (2013). "Cyanobacterial toxins: biosynthetic routes and evolutionary roots." *FEMS Microbiol.Rev.,* 37(1), 23-43.

Egerton, T. A., Morse, R. E., Marshall, H. G., and Mulholland, M. R. (2014). "Emergence of Algal Blooms: The Effects of Short-Term Variability in Water Quality on Phytoplankton Abundance, Diversity, and Community Composition in a Tidal Estuary." *Microorganisms,* 2(1), 33-57.

Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M. H., and Visser, P. M. (2018). "Cyanobacterial blooms." *Nature Reviews Microbiology,* 16(8), 471-483.

Igwaran, A., Kayode, A. J., Moloantoa, K. M., Khetsha, Z. P., and Unuofin, J. O. (2024). "Cyanobacteria Harmful Algae Blooms: Causes, Impacts, and Risk Management." *Water, Air, & Soil Pollution,* 235(1), 71.

Izydorczyk, K., Tarczynska, M., Jurczak, T., Mrowczynski, J., and Zalewski, M. (2005). "Measurement of phycocyanin fluorescenceas an online early warning system for cyanobacteria in reservoir intake water." *Environ.Toxicol.,* 20(4), 425-430.

Johansen, R. A., Reif, M. K., Emery, E. B., Nowosad, J., Beck, R. A., Xu, M., and Liu, H. (2019). " waterquality : an open-source R package for the detection and quantification of cyanobacterial harmful algal blooms and water quality." .

Jordan, C., Cusack, C., Tomlinson, M. C., Meredith, A., McGeady, R., Salas, R., Gregory, C., and Croot, P. L. (2021). "Using the Red Band Difference Algorithm to Detect and Monitor a Karenia spp. Bloom Off the South Coast of Ireland, June 2019." *Frontiers in Marine Science,* 8.

Keith, D., Rover, J., Green, J., Zalewsky, B., Charpentier, M., Thursby, G., and Bishop, J. (2018). "Monitoring Algal Blooms in drinking water reservoirs using the Landsat 8 Operational Land Imager." *Int.J.Remote Sens.,* 39(9), 2818-2846.

Kudela, R. M., Berdalet, E., Bernard, S., Burford, M., Fernand, L., Lu, S., Roy, S., Tester, P., Usup, G., Magnien, R., Anderson, D. M., Cembella, A., Chinain, M., Hallegraeff, G., Rguera, B., Zingone, A., Enevoldsen, H., and Urban, E. (2015). "Harmful Algal Blooms. A scientific Summary for Policy Makers." .

Meng, H., Zhang, J., and Zheng, Z. (2022). "Retrieving Inland Reservoir Water Quality Parameters Using Landsat 8-9 OLI and Sentinel-2 MSI Sensors with Empirical Multivariate Regression." *Int.J.Environ.Res.Public.Health.,* 19(13), 7725. doi: 10.3390/ijerph19137725.

Rodríguez-Benito, C. V., Navarro, G., and Caballero, I. (2020). "Using Copernicus Sentinel-2 and Sentinel-3 data to monitor harmful algal blooms in Southern Chile during the COVID-19 lockdown." *Mar.Pollut.Bull.,* 161(Pt A), 111722.

Rodríguez-López, L., Duran-Llacer, I., González-Rodríguez, L., Cardenas, R., and Urrutia, R. (2021). "Retrieving Water Turbidity in Araucanian Lakes (South-Central Chile) Based on Multispectral Landsat Imagery." *Remote Sensing,* 13(16),.

Smucker, N. J., Beaulieu, J. J., Nietch, C. T., and Young, J. L. (2021). "Increasingly severe cyanobacterial blooms and deep water hypoxia coincide with warming water temperatures in reservoirs." *Glob.Chang Biol.,* 27(11), 2507-2519.

Stumpf, R. P., Davis, T. W., Wynne, T. T., Graham, J. L., Loftin, K. A., Johengen, T. H., Gossiaux, D., Palladino, D., and Burtner, A. (2016). "Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria." *Harmful Algae,* 54 160-173.

Vincent, W. F. (2018). *Lakes: A Very Short Introduction.* Oxford University Press, .

Watson, S. B., Miller, C., Arhonditsis, G., Boyer, G. L., Carmichael, W., Charlton, M. N., Confesor, R., Depew, D. C., Höök, T. O., Ludsin, S. A., Matisoff, G., McElmurry, S. P., Murray, M. W., Peter Richards, R., Rao, Y. R., Steffen, M. M., and Wilhelm, S. W. (2016). "The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia." *Harmful Algae,* 56 44-66.