Quantifying the Role of *Microcystis* Resuspension on Harmful Algal Blooms in Coastal Lake Erie Using Multidisciplinary Approaches

By

Akshata Karnik
Bangzhou Shu
Joshua Habib
Journ Galvan
Nicole Rappuhn

A Master's Project Report submitted
in partial fulfillment of the requirements
for the degree of Master of Science
(School for Environment and Sustainability)
at the University of Michigan-Ann Arbor
April 18th, 2023

Faculty Advisors
Dr. Casey M. Godwin
Dr. Dmitry Beletsky
Dr. Subba Rao Chaganti
Dr. Michael Fraker

Table of Contents

Abstract

This study evaluated the effects of sediment resuspension and *Microcystis* sediment populations on Harmful Algal Blooms (HABs) in Lake Erie. Within these HABs, *Microcystis* is a prominent species of cyanobacteria that is capable of producing the harmful hepatotoxin microcystin, when mcy genes are present. The annual occurrence of HABs can lead to negative human health effects, harm wildlife, and disrupt access to drinking water supplies. As a result, the ability to accurately predict, forecast, and characterize these blooms is extremely important. Despite previous research, there has been a knowledge gap regarding the magnitude in which resuspension impacts blooms and whether there is a significant presence of sediment population of *Microcystis* present. The differences between *Microcystis* populations in the water column and sediment have also not been well understood. Through the use of hydrodynamic modeling, sediment core collection, laboratory experiments, omics-based methods, and statistical modeling, we have assessed sediment resuspension's impact on HABs. Our first objective consisted of hydrodynamic modeling to identify the strength and frequency of total bottom stress. The second objective utilized laboratory experiments with sediment cores to determine critical bottom stress for resuspension of *Microcystis*. Samples taken during the sediment core experiments were analyzed in objective 3 using genomics to determine whether *Microcystis* genotypes in the blooms are the same as those found on sediments. Finally, our forth objective combined hindcasts and field observations to examine resuspension and phenology. We expect that our results will be able to improve current models, contribute information to public databases, and provide a basis for future research.

Acknowledgements

We would like to thank the Cooperative Institute for Great Lakes Research (CIGLR), the US Coastal Research Program (USCRP) as administered by the US Army Corps of Engineers (USACE) Department of Defense, and the University of Michigan for providing funding for this project.

We also extend thanks to the staff at NOAA Great Lakes Environmental Research Laboratory (GLERL) and CIGLR for their support and assistance. In particular, we thank Raisa Beletsky, Glenn Carter, Christine Kitchens, Russ Miller, Jasmine Mancuso, Lucas Vanderbilt, and Reagan Errera. Additional thanks to Kent Baker, Captain of NOAA R4108, for making our sediment core field sampling possible.

Introduction

Harmful algal blooms (HABs) formed by planktonic cyanobacteria impact both freshwater and marine systems worldwide (Calomeni et al., 2022; Anderson et al., 2021; Paerl et al., 2016). Out of the 5 Great Lakes, Lake Erie is the shallowest and most productive. Within Lake Erie, the western basin has been plagued with annually occurring algal blooms that have threatened both ecosystem and public health for the past twenty years (Steffen et al., 2014; Steffen et al., 2017). Two principal factors that aid HAB formation are natural processes (river flow, circulation, and upwelling relaxation) and nutrient loading (Sellner et al., 2003). Freshwater ecosystems are vital in regulating Earth's climate; however, they are threatened by several anthropogenic activities such as release of sewage, industrial waste, agricultural runoff, and nutrient-rich effluent deposition. Accelerated eutrophication, due to excessive inputs of the limiting nutrients phosphorus and nitrogen, has resulted in global proliferation of HABs caused by cyanobacteria (Steffen et al., 2014; Watson et al., 2016; Benayache et al., 2019). Lake Erie experiences intense and recurrent blooms in coastal areas, such as areas near the Sandusky and Maumee Rivers, due high nutrient availability and shallow bathymetry (Carmichael & Boyer, 2016).

Within these blooms, *Microcystis* is a prominent species of cyanobacteria that is capable of accumulating in high biomass concentrations towards the top of the water column, due to its ability to regulate its buoyancy (Verspagen et al., 2004). This accumulation of *Microcystis* is largely concerning due to its ability to produce the harmful hepatotoxin microcystin, when mcy genes are present (Carmichael 1992; Carmichael, 1994; Carmichael & Boyer, 2016). A large HAB was recorded in Lake Erie

in 2011, during which Microcystin concentrations in the western basin ranged between 0.1 µg/L to 8.7 µg/L (WHO guideline limit for safe drinking water is 1µg/L), and was attributed to increased precipitation, weak lake circulation, and farming practices (Michalak et al., 2013). Similarly, in 2014 the city of Toledo issued a “do not drink” notice and temporarily shut down the Toledo drinking water plant, after microcystin was detected in treated drinking water above WHO permissible limits (Steffen et al., 2017). The 2014 Toledo Water Crisis ultimately impacted access to fresh drinking water for over half a million residents (Bullerjahn et al., 2016). The presence of microcystin in Lake Erie also has harmful effects for the region's economy, as recreational use of the lake becomes dangerous when the toxin is present. Given that Lake Erie supports more than 11 million consumers and the industries dependent on the lake generate more than $50 billion yearly, recurring HABs are a serious issue that need to be addressed (Watson et al., 2016).

As a result, it is extremely important to be able to better understand and predict both the short-term dynamics of HABs and the long-term response of the lake. The Cooperative Institute for Great Lakes Research (CIGLR) and NOAA Great Lakes Environmental Research Laboratory (GLERL) are responsible for monitoring Lake Erie’s HABs and report out useful data, including toxin concentrations in the lake. Past monitoring by CIGLR and GLERL, as well as other studies, has shown that there are large populations of viable *Microcystis* on the lake sediment in Fall, Winter, and Spring (Kitchens et al., 2018; Calomeni et al., 2022). In temperate regions the annual cycle of *Microcystis* follows the pattern of overwintering in upper layers of the lake sediment, resuspension in the water column in spring, bloom formation in summer, and settling

back in the sediment in autumn (Kitchens et al., 2018). Although *Microcystis* abundance decreases over winter, lab experiments found that these cells are viable at the end of the overwintering period (Kitchens et al., 2018)

Further, past work has shown that the genotypes comprising the bloom in the lake are different from those in the main tributaries (Kutovaya et al., 2012; Rinta-Kanto 2009; Davis et al., 2014; Kitchens et al., 2018). The Maumee river is considered a source of essential nutrients that support blooms and potentially a source of the cyanobacteria (Stow et al., 2015; Stumpf et al., 2012). A study by (Bridgeman et al., 2012) found that *Microcystis* populations were observed first in the river and then in the lake suggesting that the river is a probable source of lake *Microcystis*. However, phylogenetic analysis of cyanobacterial samples from Maumee River and Lake Erie revealed that Maumee River samples belonged to Planktothrix spp, while Lake Erie samples predominantly belonged to *Microcystis spp.* (Kutovaya et al., 2012). Similarly, samples from Sandusky Bay also belong to *Planktothrix spp.*, which suggests that toxic *Microcystis spp.* found within HABs may originate from Lake Erie sediments (Kutovaya et al., 2012).

Past research has helped with better understanding the nature of *Microcystis* in Lake Erie HABs, but there remains a knowledge gap in predicting when the bloom will start each year, and in determining what causes *Microcystis* populations to move up into the water column. Additionally, potential differences in the genotypes between *Microcystis* populations in the water column and in the lake sediment are still somewhat unclear. According to past monitoring data from CIGLR and NOAA-GLERL, it has been observed that the western basin will go from exhibiting low biomass concentrations to a

rapidly accelerating bloom sometime between the end of June and late July. *Microcystis* possesses the ability to alter its buoyancy, so it has been thought that its movement up into the water column is triggered by temperature and light intensity, thus enabling the overwintering populations to reenter the water column and contribute to bloom formation (Harke et al., 2016). However, a study by Mission and Latour (2011) suggests that the introduction of overwintering *Microcystis* population in the water column is primarily due to sediment resuspension.

Although Lake Erie experiences sediment resuspension on a regular basis, the full impact of resuspension on the dynamics of HABs is also not well understood. The western basin is the shallowest of the three Lake Erie basins, with an average depth of less than eleven meters (Dusini et al., 2009). Reduction in water levels in Lake Erie make it more susceptible to sediment resuspension because of small sediment size and an increased fluid force (induced by wind waves and current) applied to the lakebed also known as total bottom stress (Kang et al., 1982; Hunter et al., 1993; Dusini et al., 2009). Through the use of hydrodynamic modeling, laboratory experiments, ‘omics enabled methods, and statistical modeling, this study set out to evaluate the effects of *Microcystis* resuspension on Lake Erie HABs.

Objective 1: Hydrodynamic Modeling

Methods

FVCOM (Chen et al, 2003) and WAVEWATCH III (WW3DG 2019) unstructured grid models were independently run for Lake Erie (no coupling) to produce hourly simulated data over the span of three years (2017-19). WAVEWATCH III model was run on the FVCOM grid using the same forcing as FVCOM. Interpolated meteorological forcing was used for 2017-2018 runs, HRRR atmospheric model forcing was used for 2019 runs. Model outputs were validated against real time observations collected at NOAA GLERL's Reno Beach Station buoy. While FVCOM output already contains current-induced bottom stress, wave-induced bottom stress has to be calculated from basic WAVEWATCH III output variables before both stresses can be combined. In this study, we focus on the May-October period when Microcystis blooms typically occur in the western basin (sampling sites are shown on the map).

WAVEWATCH III (WW3)

Variables of interest located at nodes in the WAVEWATCH III (WW3) model output were bottom wave orbital velocity $(ubot)$, root mean squared of bottom displacement amplitude zonal $(uabr)$, root mean squared of bottom displacement amplitude meridional $(vabr)$, significant height (meters), max height (meters), period (seconds), and direction (degrees). Modeled outputs of significant wave height, maximum height, and period associated at the node number geographically closest to the Reno Beach station buoy (Table 1.1), were compared to observations. Reno Beach

station buoy wave observations were collected by NORTEC AWAC. Modeled bottom displacement amplitude (A) was derived from $uabr$ and $vabr$ using the below formula:

$$A = \sqrt{uabr^2 + vabr^2} \quad (1)$$

Wave Reynolds number (Re) characterizes a regime of turbulence in the near-bed environment that is important for estimation of wave-induced bottom stress. Re was calculated using the below formula where the viscosity of water is 10^{-6}

$$Re = A * ubot/viscosity \quad (2)$$

A friction factor (Fw) was calculated for both laminar $(Re < 10^4)$, smooth turbulent $(10^4 < Re < 10^5)$, and turbulent flow $(Re > 10^5)$ using the below formulas from Lin et al. 2021

$$Fw_{laminar} = 2 * Re^{-0.5} \quad (3)$$

$$Fw_{smooth\ turbulent} = 0.09 * Re^{-0.2} \quad (4)$$

$$Fw_{turbulent} = exp(-6 + 5.2 * \left(\frac{A}{Ks}\right)^{-0.19}) \quad (5)$$

where $Ks = 440\ microns$. Wave induced bottom stress (Tau_w), was given by

$$Tau_w = 0.5 * density * Fw * ubot^2 \quad (6)$$

where the density of water is $1000\ \mathrm{kg/m^3}$.

Analysis of Re in western basin during periods of study revealed that the near-bed environment was either in the laminar or smooth turbulent regime, making the friction factor (and wave-induced bottom stress) independent of grain size (Ks).

Table 1.1 *Coordinates for study sites in the western basin of Lake Erie with associated node in the FVCOM unstructured grid.*

Site	**Longitude**	**Latitude**	**Node**
Reno Beach	-83.25	41.68	1136
WE2	-83.33	41.76	758
WE4	-83.2	41.83	374
WE6	-83.39	41.71	1209
WE8	-83.36	41.83	458
WE12	-83.26	41.7	976

Finite Volume Community Ocean Model (FVCOM)

FVCOM was used in this study to model bed stress magnitude from currents also known as current-induced bottom stress (Tau_c). It is proportional to the bottom current speed squared determined by

$$Bottom\ current\ speed = \sqrt{u^2 + v^2} \tag{7}$$

where (u) is Eastward water velocity, and (v) is Northward water velocity.

Interpolation

Although the unstructured (triangular) grid for WW3 and FVCOM was the same, modeled variables were defined at different locations within triangles. Bottom stresses

in FVCOM were associated with elements (triangle centroids) while WW3 variables were associated with nodes (triangle vertices). Therefore, current bottom stress and bottom current speed from elements were spatially interpolated to nodes using the Inverse Distance Weighting method (IDW)

$$\hat{z}(x) = \frac{\sum_{i}^{n} w_i z_i}{\sum_{i}^{n} (w_i)} \quad (8)$$

Once bed stress magnitude from currents was interpolated to nodes with corresponding wave induced bottom stress, total bottom stress $\left(Tau_b\right)$ could be calculated by adding wave and current bottom stress together.

$$Tau_b = Tau_w + Tau_c \quad (9)$$

Turbidity observations taken from a YSI EXO2 Multiparameter Water Quality Sonde were compared to total bottom stress values at each site for months May-October as it is assumed that turbidity measurements would be correlated with bottom stress (citation?). All analysis was performed in R version 4.2.2

Results

Wave model validation

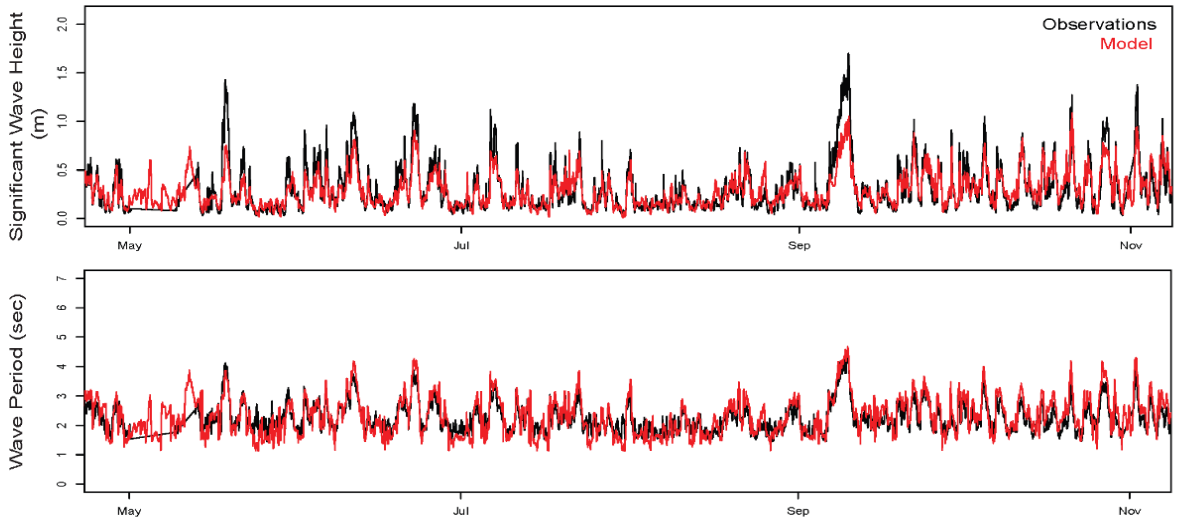


Figure 1.1. Time series of observed significant wave height (m) and wave period (sec) taken from NORTEK AWAC deployed at Reno Beach station buoy, and WW3 modeled significant wave height and wave period (red lines) in the year 2018, months May-October.

To test WW3 model robustness, simulated significant wave height (m) and wave period (sec) were validated with observations in Lake Erie (Fig. 1.1). WW3 reproduced significant wave height measurements with a root-mean-square-error (RMSE) of 0.15 m and wave period measurements with an RMSE of 0.43 sec which was in agreement with observations, but the model underestimated wave height during peak events.

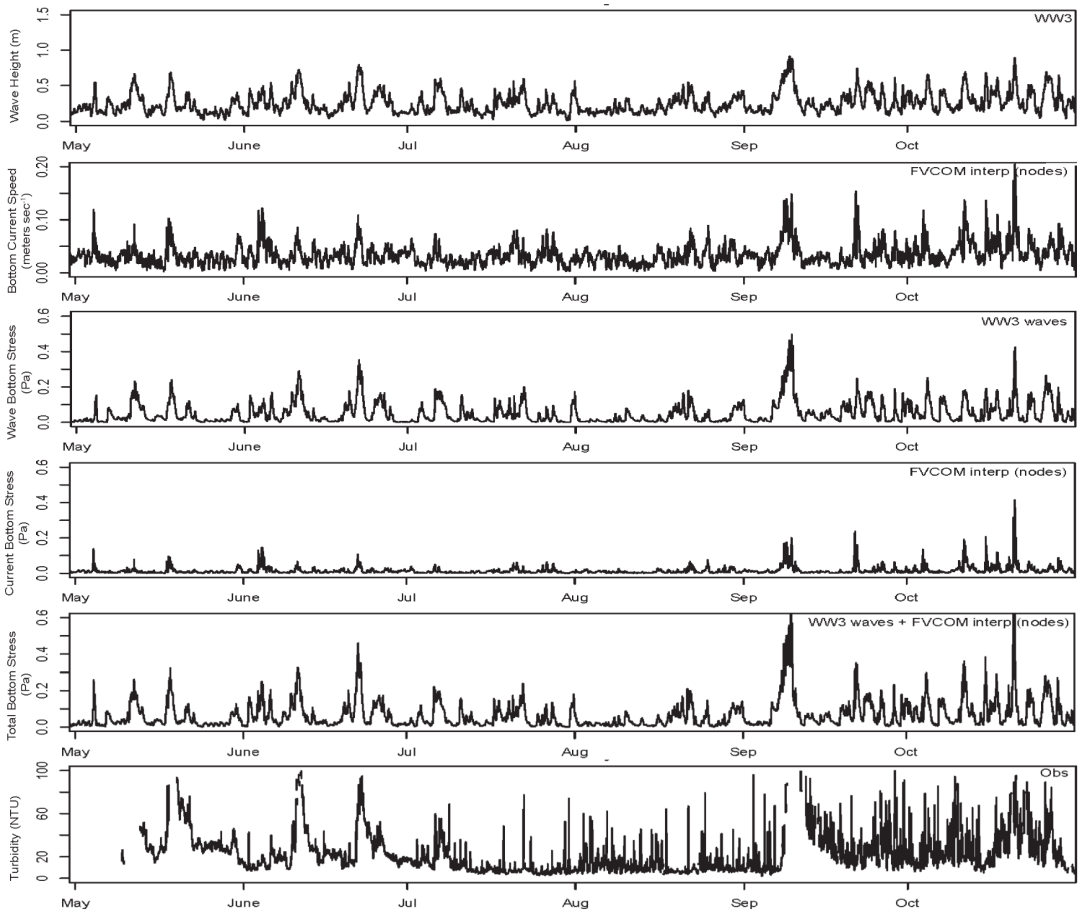


Figure 1.2. Time-series of calculations outlined in methods. Wave bottom stress (Pa) was derived from WW3 simulated wave height (m). Bottom current speed (meters sec-1) and current bottom stress (Pa) were simulated by FVCOM. Total bottom stress (pa) was calculated by adding wave bottom stress (pa) derived from modeled WW3 measurements, and FVCOM simulated current bottom stress. Turbidity (NTU) observations from a YSI EXO2 Multiparameter Water Quality Sonde deployed 0.5 mab were plotted to validate the calculated total bottom stress.

Figure 1.2 illustrates the relationship between all modeled variables and observations related to sediment and *Microcystis* resuspension in the water column. Reno Beach was chosen to validate simulated wave (WW3) and current (FVCOM) measurements, as this site had the most reliably accurate buoy observations for the western basin. Time series composite plots also assisted in validating spatial interpolation methods (IDW) for current-induced bottom stress from elements to nodes in the FVCOM unstructured grid. Current bottom stress is also a function of bottom

current speed, which was plotted in the second panel (Fig. 1.2). Variables agree throughout the summer months (May-October), with wave-induced bottom stress having a larger contribution to total bottom stress for the western basin.

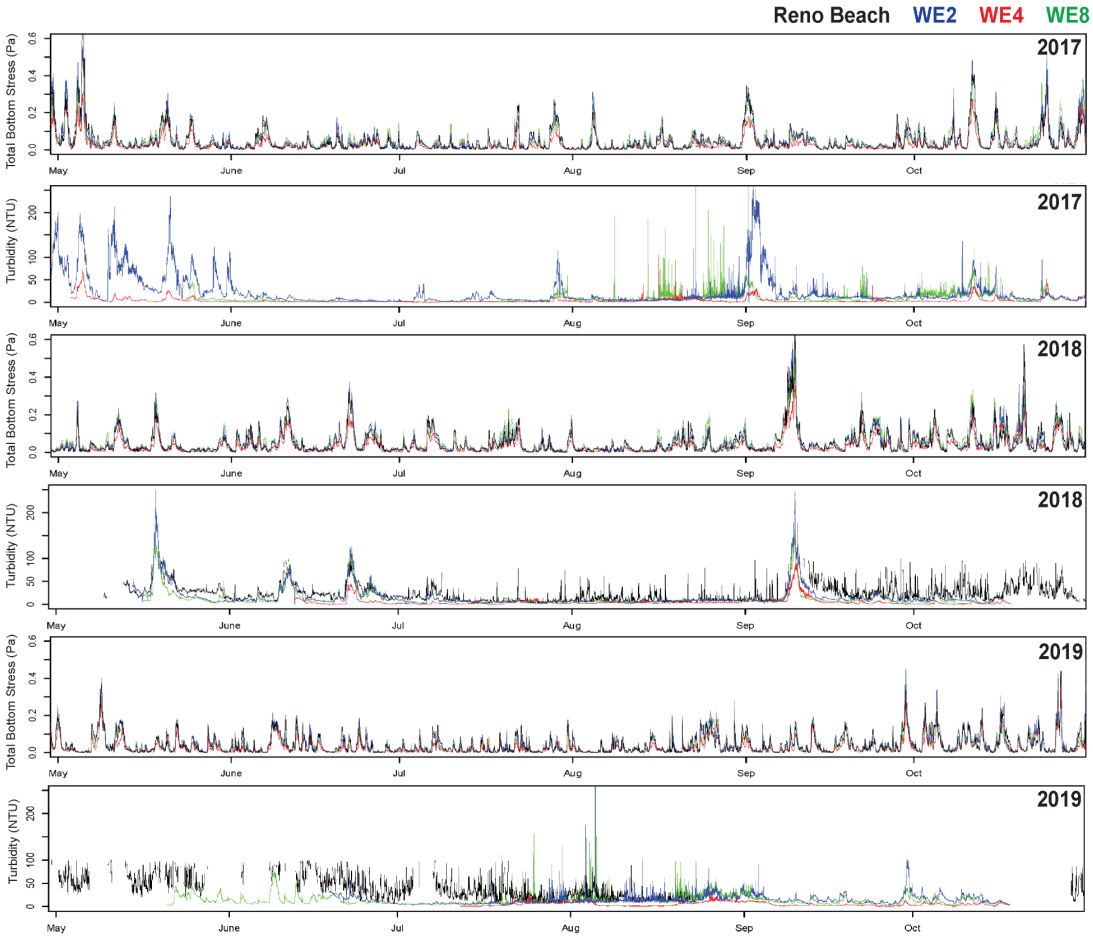


Figure 1.3. Total bottom stress (Pa) calculated from hydrodynamic model output and turbidity observations (NTU) hourly time series for Reno Beach, WE2, WE4, and WE8 sites for months May-October and across years 2017-19. Sites WE6 and WE12 were lacking turbidity data and were not shown.

Hourly time series for calculated total bottom stress (Pa) and observed turbidity (NTU) generally agree for years 2017 and 2018, although there were more peaks in the bottom stress time series than in turbidity observations. For the year 2017, both bottom stress and turbidity varied seasonally, with pronounced peaks throughout May,

September, and October. The end of July and beginning of August also showed a minor peak for bottom stress and turbidity, although it was not as pronounced as in the Spring and Fall months. In 2018, minor peaks for total bottom stress and turbidity were observed throughout the Spring and Summer months, with a pronounced peak in September and moderate peaks that continued through October. For the observed period in 2019, total bottom stress and turbidity were less pronounced, with energy being the highest in October.

Across sites, WE2 appeared to have significantly higher turbidity observations for the year 2017, which could have been due to its closer proximity to the Maumee River mouth but was in agreement with other sites in 2018. WE4, which was the furthest site from the Maumee River mouth, had the lowest turbidity observations even during peak months when other sites saw dramatic spikes, but was in fair agreement with all other sites in 2018. Increased fluctuations in turbidity and gaps in hourly sensor observations for 2019, made it difficult to compare turbidity across sites. Modeled total bottom stress was in agreement for sites at Reno Beach, WE2, WE4, and WE8 during the observed time period May-October and across all years (2017-19) (Fig. 1.3).

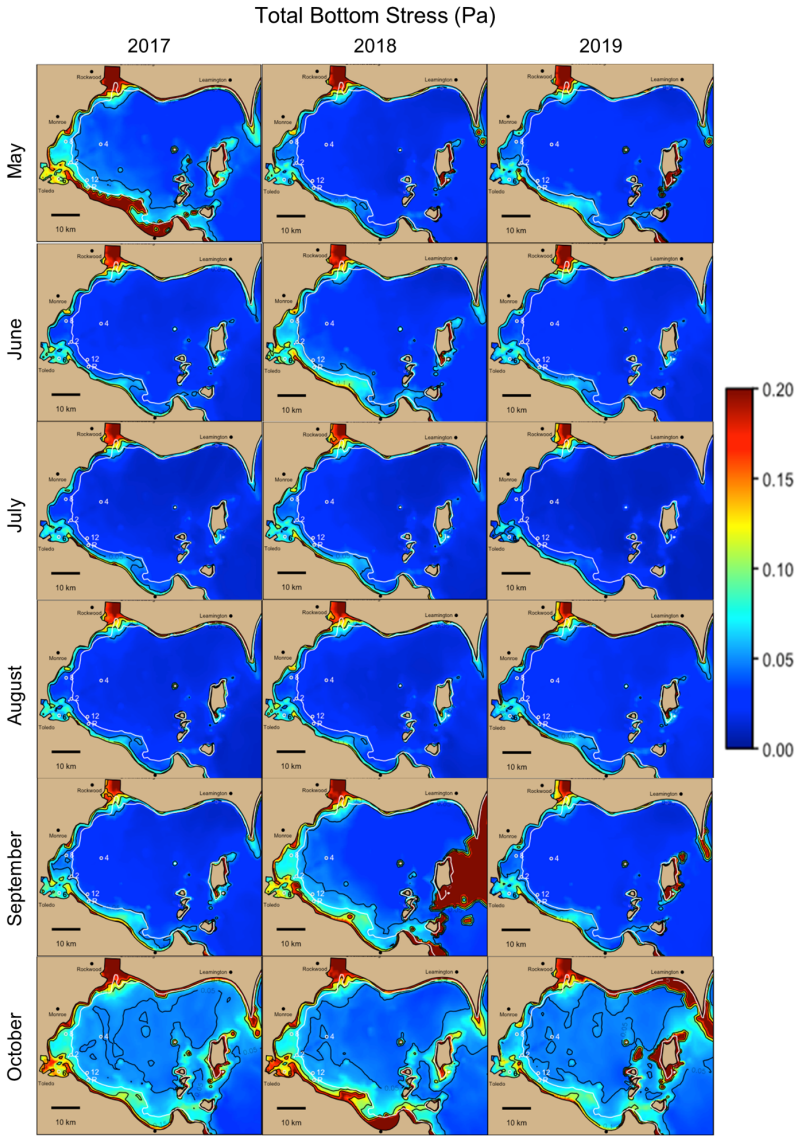


Figure 1.4. Monthly averages of total bottom stress (pa) for the entire western basin of Lake Erie. Hourly model outputs of calculated total bottom stress at each node were averaged for months May-October across years 2017-19 with a white line representing 5m isobath. Color scale saturates at 0.2 Pa and larger.

Monthly averages of total bottom stress at each node in the FVCOM grid were calculated and represented spatially for the entire western basin of Lake Erie (Fig. 1.4). Total bottom stress was highest in shallow water (<5 m), which included Maumee Bay and the Detroit River mouth. Other shallow areas of the western basin with high bottom stress averages were the southwestern shore and east of Pelee Island. Total bottom

stress varied seasonally throughout the entire western basin, with more significant values in May and September-October and annual variations (2019 was less energetic than 2017 and 2018).

Critical shear factor

Upon receiving critical shear values needed for *Microcystis* (0.06 Pa) and bulk sediment (0.07 Pa) resuspension, hourly animations for the western basin of Lake Erie from May to October were generated to show only the areas that exceed these shear values. As shown in Figure 1.5, although the critical shears needed for resuspension are close in magnitude, there is still great variation in the spatial distribution of areas exceeding these values. Additionally, although wave bottom stress dominates in the western basin, current bottom stress can still cause resuspension during peak events; which is demonstrated by the additional areas in red when comparing the top and bottom panels.

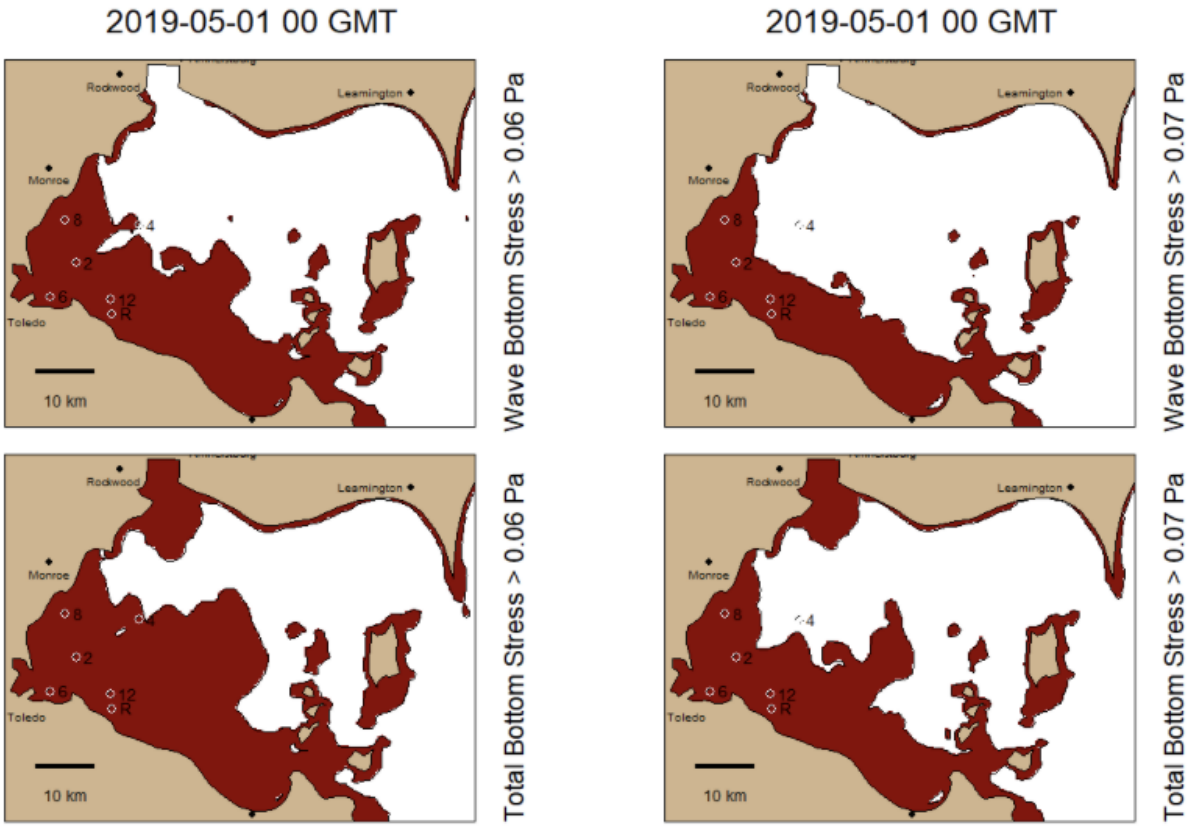


Figure 1.5. Calculated wave bottom stress (top panel) and total bottom stress (bottom panel) plotted for the entire western basin of Lake Erie for May 1st, 2019, at 00 GMT. The left image displays all areas where the bottom stress exceeds 0.06 Pa, while the right image displays stresses larger than 0.07 Pa.

Objective 2: Laboratory Experiments

In this study, we perform experiments in the laboratory using sediment core samples from western basin of Lake Erie to estimate the critical total bottom stress (TBS, referred to as bottom stress henceforth) at which resuspension of cyanobacteria (Microcystis) and sediment takes place in the water column. We hypothesized three potential outcomes: 1) cyanobacteria would resuspend at a lower critical TBS compared to the sediment, 2) cyanobacteria and sediment would resuspend at the same critical TBS, and 3) sediment would resuspend at lower critical TBS as compared to cyanobacteria.

Methods

Sediment Core Collection

We sampled sediment cores from five sites in Lake Erie (Figure 2.1). Samples were collected on May 30, 2022 (WE2), July 7, 2022 (WE2, WE4, WE12), and July 12, 2022 (WE6, WE8) using an Ekman box corer with a 30 cm by 30 cm cross-sectional area. We would then hand-core these experimental cores from within the Ekman box corer and transfer to individual sediment core barrels (Anderson et al., 2021b).

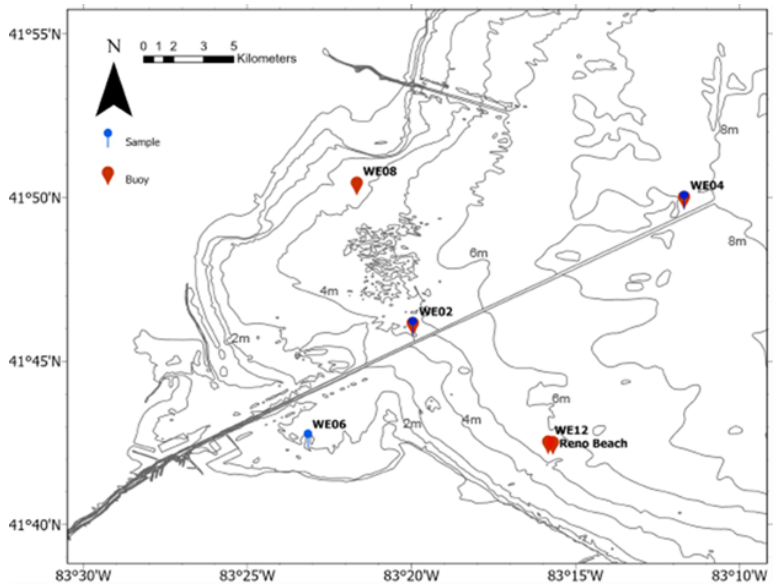


Figure 2.1. Sites where sediment cores were collected in Lake Erie.

Sediment Core Experiments

Sediment samples were allowed to settle overnight in sediment core barrels (Figure 2.2). The design of the erosion measurement system (EROMES) used at CIGLR follows that of Kalnejais et al (2007) and Tolhurst et al (2000). Our EROMES was designed to work inside the sediment core barrels used at CIGLR (Anderson et al., 2021b), which have an inside diameter of 14.6 cm. An acrylic baffle was placed inside the core barrel and on top of the sediment. That baffle has an inside diameter of 132 mm and height of 100 mm. The inside of the baffle had strips of size 15.5mm by 3.35 mm affixed in a radial pattern at 60º increments with respect to the center.

Turbulence was applied to the water inside the baffle using an impeller coupled to a stepper motor. The impeller was 76.2 mm diameter and made of stainless steel. The impeller was mounted at the end of a stainless-steel shaft coupled to the stepper motor. The stepper motor was attached to a vertical track such that the impeller can be

lowered into position. The experiments and calibration were performed with the bottom of the impeller 5 cm above the sediment surface.

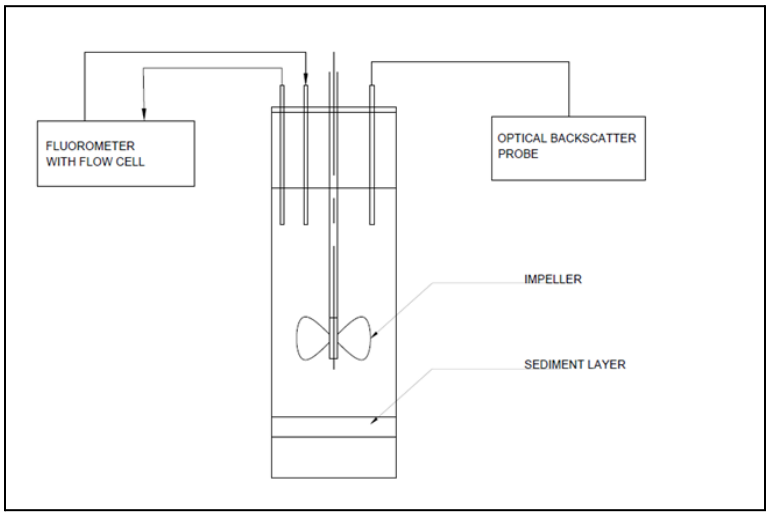


Figure 2.2. Experimental setup of sediment core used to determine critical shear required for resuspension.

The velocity of the impeller was increased stepwise every ten minutes starting from 1 RPM to 125 RPM. We followed the EROMES calibration approach of Kalnejais et al. (2006) whereby the RPM required to resuspend sand of different sizes is recorded and converted to bottom stress using the Shields curve or parameterization (Soulsby & Whitehouse, 1997). An optical backscatter probe (Ocean Insight) and a fluorometer (BBE Moldaenke Fluoroprobe) were inserted in the water column to detect bulk sediment and cyanobacterial resuspension. These instruments recorded continuous observations. The experiment was repeated for all cores collected from sites mentioned in Table 2.1.

Table 2.1. Dates and sites of sediment core collection in western basin of Lake Erie.

Date of sample collection	***Site***	***Replicates***
May 30 2022	WE2	4
July 7 2022	WE2	4
July 7 2022	WE4	2
July 7 2022	WE 12	2
July 12 2022	WE 6	3
July 12 2022	WE 8	3

The optical backscatter probe recorded observations for 2048 backscatter intensities between wavelengths 190 and 1020 nm. To determine resuspension, we used an average of observations between wavelengths 540-550 nm. The fluorometer measured continuous observations of concentration of cyanobacteria (µg/L), total chlorophyll-a (µg/L), and beam transmission (%). We constructed time series plots of backscatter intensity, transmission, and algal concentration to determine the trend at the increasing bottom stress.

In order to determine the changepoint in backscatter intensity, transmission percentage, and algal (cyanobacteria and total chlorophyll-a) concentration in the cylinder we fit segmented regressions using a bayesian regression. We used the changepoint in cyanobacterial concentrations to identify critical bottom stress for *Microcystis* resuspension and the changepoint in beam transmission percentage to identify critical shear stress for bulk sediment resuspension. The time at which the changepoint occurred was mapped to the RPM of the impeller which was then equated to bottom stress using results from the EROMES calibration experiment. We performed the changepoint detection analyses in RStudio (4.2.2) using the brms package (Bürkner, 2018) and compiled segmented Bayesian regression models (Anderson et al., 2021)

that were fit using a Hamiltonian Markov Chain Monte Carlo simulation in ‘stan’ (Stan Development Team, 2020). Figure 2.3 represents a plot of regression with the fitted terms.

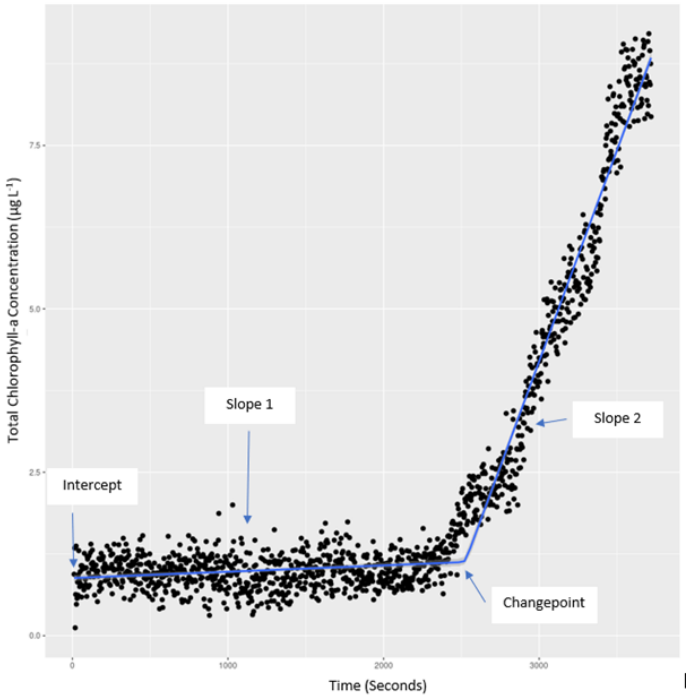


Figure 2.3. Segmented Bayesian regression with fitted terms.

We used the posterior distributions to estimate the mean, standard errors, and 95% credible intervals. Four Markov chains and informative priors (Table 2.2) informed each regression. After an initial period of 1000 warmup iterations, we sampled each chain for 1000 iterations. We assessed model convergence using the Gelman-Rubin statistic, R-hat <1.01.

Table 2.2. Priors used to inform regression for the four response variables used to determine changepoint

Response Variable	***Intercept Mean & SD***	***Slope1 (pre-resuspension)***	***Slope 2 (resuspension)***	***Changepoint***
Backscatter Intensity	9000, 3200	0,10	0,10	3000,1500
Cyanobacteria	1,1	0,10	0,10	3000,1500
Total Chlorophyll-a	7,5	0,10	0,10	3000,1500
Beam Transmission	85, 20	0, 1	0,1	3000,1500

SD: Standard Deviation

Often the backscatter probe showed excessive drift, and, in those cases, we were not able to use the data to produce estimates of bulk sediment erosion. For some cores with coarse sediment, we did not observe bulk sediment erosion across the range of bottom stress used for the experiment. In those cases, we did not estimate the point at which bulk sediment erosion began.

Results

We analyzed data for sediment cores collected across five sites over three sampling events between May 30, 2022, and July 12, 2022 (Table 2.1). The aim of the experiment was to identify critical bottom stress at which bulk sediment and cyanobacterial resuspension takes place in the water column. We achieved this by assessing the changepoint in beam transmission and cyanobacterial concentration in the water column when subjected to increasing impeller speed. These changepoint timing estimates (Table 2.3) were used to identify the corresponding impeller speed at which it occurred, which was converted to bottom stress based on the EROMES

calibration (Table 2.4). Figure 2.4 represents the trend of algal concentration and beam transmission percentage at successive bottom stresses.

Table 2.3 Changepoint estimates identified for the three response variables using Bayesian regression

Date	***Site***	***Replicate***	***Cyanobacteria Changepoint (seconds)***	***Total Chlorophyll-a Changepoint (seconds)***	***Beam Transmission Changepoint (seconds)***
May 30	WE2	C	2533 (2469 – 2600)[#]	2511 (2459 – 2563)	2428 (2396 – 2460)
	WE2	D	688 (36 – 1746)[*]	1745 (1687 – 1797)	1547 (1513 – 1584)
	WE2	E	1667 (1627 – 1711)	1653 (1620 – 1690)	1181 (1160 – 1202)
July 7	WE2	E	2406 (2361 – 2449)	2404 (2372 – 2433)	2378 (2368 – 2388)
	WE2	F	584 (1 – 2114)[*]	1760 (479 – 2271)[*]	2460 (2441 – 2479)[#]
	WE2	H	1568 (1470 – 1666)	2653 (2622 - 2682)	2522 (2504 – 2540)
	WE4	B	2412 (2394 – 2433)	2524 (2502 – 2545)	2467 (2454 – 2479)
	WE4	D	2297 (2234 – 2358)[#]	2388 (2325 – 2477)[#]	2583 (2556 – 2612)
	WE12	G	2729 (2669 – 2779)[#]	2663 (2626 – 2694)	2488 (2472 – 2506)
	WE12	H	3094 (3043 – 3149)[#]	2284 (3 – 3069)[*#]	2957 (2941 – 2974)
July 12	WE6	A[ł]	678 (645 – 711)[*]	478 (17 – 701)[*]	747 (37 – 5121)[*]
	WE6	C[ł]	2881 (2763 – 2986)	2879 (2772 – 2972)	1140 (11 – 3170)[*]
	WE6	D[ł]	2422 (628 – 6012)[*]	1891 (1409 – 2479)	1553 (76 – 3126)[*]
	WE8	A	1499 (2 – 2355)[*]	603 (1 – 2359)[*]	66 (2 – 373)[*^]
	WE8	B[ł]	155 (115 – 204)	124 (79 – 183)	155 (76 – 283)[^]
	WE8	D	2624 (2528 – 2739)	2600 (2504 – 2694)	543 (73 – 1062)[*^]

*Values in brackets represent 95% confidence intervals. * - Large confidence intervals suggesting the model did not converge. # - Sigma r-hat greater than 1.01. ^ - Bulk sediment resuspension not observed, changepoint estimate not valid. ł - Changepoint estimates not used in calculation of critical shear stress*

The changepoint for total chlorophyll-a concentration occurred after beam transmission (BT) changepoint seven times. For all May 30 samples for site WE2 total chlorophyll-a (TC-A) changepoint occurred after beam transmission. However, for samples from July 7 half the sample results suggested that changepoint for total chlorophyll-a occurred before beam transmission while for the rest beam transmission

changepoint took place before TC-A and there was no apparent trend across the different sites.

We observed that seven times cyanobacterial resuspension occurred before bulk sediment resuspension. Out of the three samples collected in May, two samples had cyanobacterial resuspension before bulk sediment resuspension. While for samples collected on July 7 out of seven samples, bulk sediment resuspension occurred before cyanobacterial resuspension four times. Cyanobacterial resuspension occurred before bulk sediment resuspension for site WE2 and WE12. For WE4 samples bulk sediment resuspension was detected first. We did not observe bulk sediment erosion across the range of shear stress for sample for July 12 and hence not use these samples in final analysis. Across all the samples, the changepoint estimates for TC-A and BTP occurred at similar time points while those for cyanobacterial concentration varied between samples.

Table 2.4. Bottom Stress values established at different impeller RPM using EROMES calibration.

RPM	***Bottom Stress (Nm^{-2})***
1	*0.000697*
10	*0.006966*
25	*0.017415*
50	*0.034829*
75	*0.052244*
100	*0.069658*
125	*0.087073*
150	*0.104488*

We identified the impeller speed at which these changepoints occurred and translated the impeller speed to bottom stress using the EROMES calibration experiment for all samples. Table 2.4 provides values for impeller RPM and the corresponding bottom stress. The median shear stress value for cyanobacterial

concentration was 0.060951 N m^{-2} and for TC-A and BT was found to be 0.069658 N m^{-2} (Table 2.5).

Table 2.5. Critical bottom stress values at which resuspension occurred for the response variables.

Date	***Site***	***Replicate***	***Cyanobacteria (Nm^{-2})***	***Total Chlorophyll-a (Nm^{-2})***	***Beam Transmission (Nm^{-2})***
May 30	WE2	C	0.069658	0.069658	0.069658
	WE2	D	0.017415	0.069658	0.052244
	WE2	E	0.069658	0.069658	0.034829
July 7	WE2	E	0.069658	0.069658	0.069658
	WE2	F	0.006966	0.034829	0.069658
	WE2	H	0.034829	0.069658	0.069658
	WE4	B	0.069658	0.069658	0.069658
	WE4	D	0.052244	0.069658	0.069658
	WE12	G	0.069658	0.069658	0.069658
	WE12	H	0.087073	0.052244	0.069658
July 12	WE8	A	0.017415	0.006966	-
	WE8	D	0.052244	0.052244	-
Median			**0.060951**	**0.069658**	**0.069658**

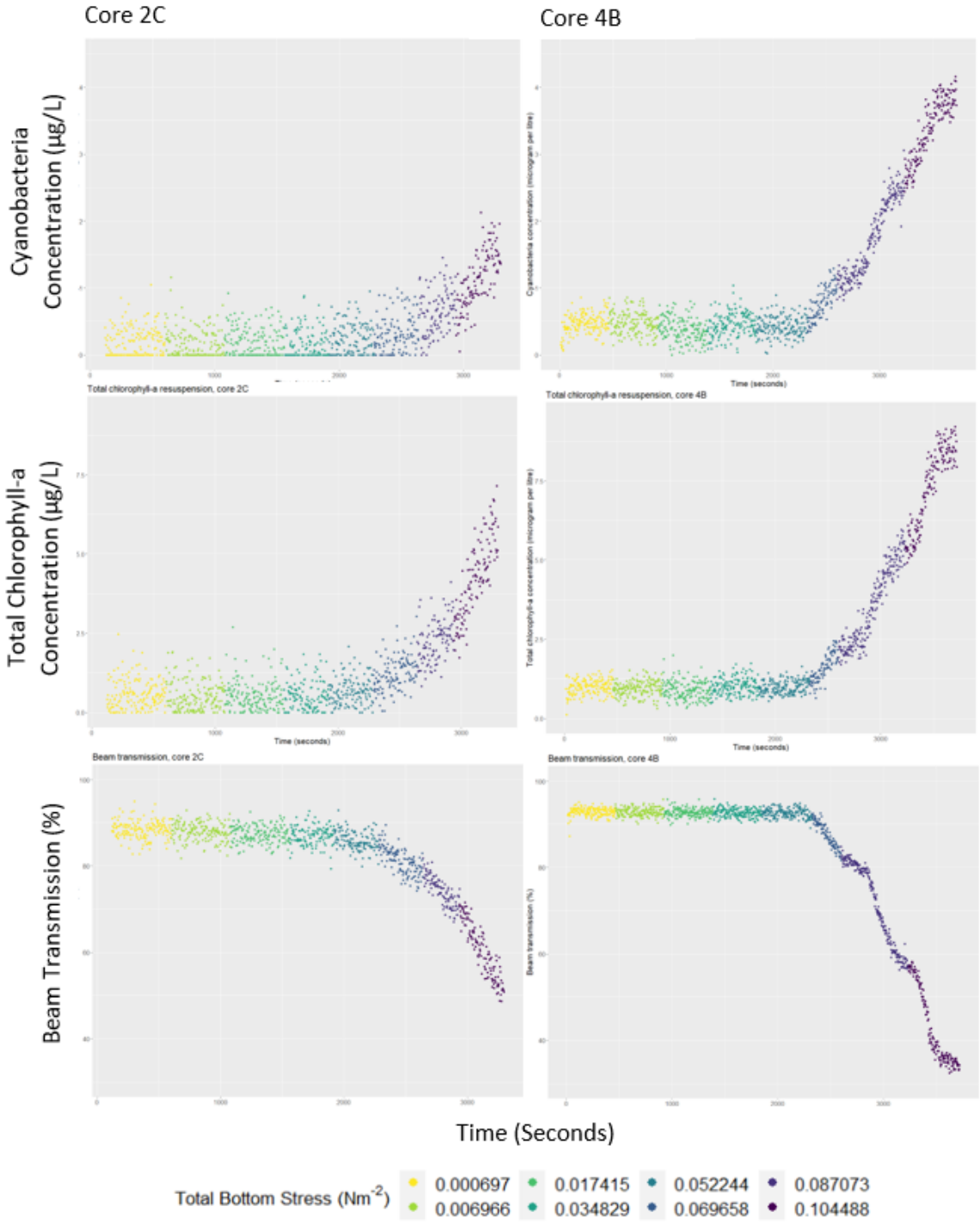


*Figure 2.4.*Cyanobacteria and total chlorophyll-a concentrations, and beam transmission percentage observed in water column over a series of increasing shear stress.

Objective 3: Genomics

Through the use of genomics, objective 3 set out to determine whether the *Microcystis* genotypes found in blooms are the same as those found on sediments.

Methods

Sediment Core Experiments

During the summer of 2022, sediment cores were collected from a variety of sites within the western basin of Lake Erie on May 31st, 2022 and July 7th, 2022. These sediment cores were assessed using in-lab experiments, described in detail within Objective 2, where different rates of shear stress were used to observe how the sediment cores with overlying lake water would show sediment and *Microcystis* resuspension. The shear speeds tested included 0.000697 N m^{-2}, 0.006966 N m^{-2}, 0.017415 N m^{-2}, 0.034829 N m^{-2}, 0.052244 N m^{-2}, 0.069658N m^{-2}, 0.087073 N m^{-2}, and 0.104488 N m^{-2}. During the laboratory experiments, overlying water samples were collected from each sediment core at each shear speed interval. The water was pulled through clean tubing directly from the top of the sediment core, into a sterile syringe, and measured out to 50 mL per sample. Using a vacuum manifold, the water was then filtered through sterile 0.22 µm flat filters that were collected into cryotubes and stored at -20°C for future DNA extraction.

DNA Extraction and Quantification

The DNA extractions were completed utilizing QIAgen DNeasy kits with QIAshredder columns added. The standard protocol was modified from the QIAgen

DNeasy Standard Operating Procedure #001 that was summarized from DNeasy Handbook 07/2006. Three different phases were completed that included DNA preparation disruption, capture and purification, and elution. During the preparation disruption stage, 130 µL of master mix (100 µL Buffer ATL + 30 µL proteinase K) was added to each sample tube containing a flat filter. Then samples were mixed by vortexing for 10 seconds before adding 300 µL Buffer AL, then mixed by vortexing for 10 seconds again. Once the samples were well mixed, they were incubated in a shaking dry bath for 60 minutes at 56°C, with shaking set to 300 RPM. The samples were mixed every 15 minutes of incubation by quickly vortexing. After, all of the samples were vortexed at maximum speed for 10 minutes then centrifuged for 30 seconds at 13.3K x g. The lysate (including filter) was then transferred to a QIAshredder column using a sterile pipet tip, flipping the filter so that biomass is most accessible. Finally all of the samples were centrifuged for 30 seconds at 13.3K x g.

The next stage of extraction was for DNA capture and purification. After being centrifuged, 300 µL of 100% Ethanol was added to the QIAshredder column's flowthrough then the sample was mixed by vortexing for 10 seconds. 500 µL of the flowthrough lysate was then transferred into DNA minispin columns and centrifuged for 30 seconds at 13.3 K x g. Afterwards, the flowthrough was discarded, and then the prior steps were repeated, until all lysate had passed through the column. The final flowthrough was discarded along with the collection tube, and the DNA column was put into a new collection tube. Next, 500 µL of Buffer AW1 was added to the DNA column, then it centrifuged for 30 seconds at 13.3 K x g, with the flowthrough getting discarded afterwards. Following that step, 500 µL of Buffer AW2 was added to the DNA spin

column and centrifuged for 2 minutes at 13.3 K x g. The last part of this stage was discarding the flowthrough and centrifuging the sample for 1 minute 13.3 K x g.

The final part of the DNA extractions was for DNA elution. First, the DNA spin column was placed into a sterile 1.5mL centrifuge tube. Then 30 µL of Buffer AE was added, and the sample was incubated at room temperature for 3 minutes, then centrifuged for 1 minute at 8 K x g. The prior steps were repeated for a second 30 µL elution. Finally, the DNA spin column was discarded and the 60 µL of extracted DNA is kept in the centrifuge tube. After the extractions were completed, the processed samples were carefully labeled and stored at -20°C. Quality checks were later performed on the samples to ensure that the DNA extraction had produced enough DNA concentration to be able to proceed with quantitative polymerase chain reaction. The DNA concentrations were measured using standard protocol for both NanoVue and Qubit Assay. The plots for analyzing the DNA concentrations were made in Microsoft Excel.

Quantitative Polymerase Chain Reaction

Quantitative polymerase chain reaction (qPCR) was completed for the DNA extractions from each sediment core collection site. The qPCR was completed using a Phytoxigene™ CyanoDTec assay that consists of the 205-0050 Phytoxigene™ CyanoDTec Total Cyanobacteria; (16S rRNA) and Internal Amplification Control (IAC) target kit, followed by a 205-0051 Phytoxigene™ CyanoDTec Toxin Gene; microcystin/nodularin, cylindrospermopsin and saxitoxin kit, as well as a QuantStudio™ 6 Flex Real-Time PCR System. The qPCR was able to detect and quantify the presence of Total Cyanobacteria, as well as the presence of toxin producing genes. The Total

Cyanobacteria 16s rRNA and IAC target assay was run first to ensure cyanobacteria was detected and to test whether the DNA sample contained any inhibitors that would have invalidated the qPCR going forward. After the Total Cyanobacteria was run, the IAC was referenced in order to check for inhibition, with the ideal result being no inhibition. Any samples that had results suggesting quality issues were rerun or discarded. At this point, the 16S amplified all cyanobacteria, not just *Microcystis*, so the toxin genes kit was used next to confirm whether microcystin was present in each of the DNA samples.

In order to analyze the toxin genes, the Toxin Gene assay was conducted with the second kit, which allows for *Microcystis* to be detected through the mcyE gene that is associated with the microcystin synthetase gene operon that is responsible for the potential to produce microcystin (Tillett et al., 2000). Although the Toxin Gene assay is capable of detecting multiple toxins, we only analyzed the microcystin results because the focus of our study is specifically on *Microcystis*. When analyzing the results from the qPCR data, the total copies of 16S rRNA gene showing either Total Cyanobacteria or Microcystin (mcyE gene) per mL of water sample were calculated by first taking the copies per qPCR reaction (the volume from each reaction was 2.5µL) then calculating how many copies were present per 1 µL. The total copies per 1 µL were then multiplied by the DNA Extraction sample volume of 60µL. Since the samples came from 50 mL of filtered lake water that is next taken into account by dividing by 50 mL to provide total copies per mL of water sample. If a sample had a dilution factor associated with it from the qPCR for quality control purposes that would next be multiplied to the total copies per mL. The plots for qPCR data were made using RStudio. After the qPCR process

was complete, the samples were submitted for metabarcode high throughput sequencing.

High Throughput 16S rRNA Amplicon Sequencing

Our samples were sequenced, utilizing 16S rRNA so that we could further confirm that *Microcystis* was present in the samples and learn more about potential genetic variation in the samples. The amplicon libraries primers (forward 5'-GTGCCAGCMGCCGCGGTAA-3' reverse 5'-GGACTACHVGGGTWTCTAAT-3') targeted the bacterial 16S V4 hypervariable region and were prepared using the dual indexed NEXTflex 16S V4 2.0 Illumina compatible barcoded kit. Metabarcoded samples were sequenced on Illumina's Miseq using V2 version 500 cycle, 2 X 250 base pair (bp), paired end reads. High throughput sequencing data obtained from Miseq was processed using Quantitative Insights Into Microbial Ecology 2 (QIIME2) pipeline commands, then forward and reverse primers were trimmed from all sequences. Next, all low quality reads, with a quality score of less than 30 were removed. Forward and reverse reads were ultimately truncated at 200 bp length because the quality scores dropped below 30 after 200 bp.

After the trimming was completed, truncated forward and reverse reads were merged. Merged reads were clustered into amplicon sequence variants (ASVs) based on their similarity. The quality control, denoising, and merging of the reads was performed using daad2 denoise paired command. A future table was also created which included the number of reads per sample and number of unique ASVs. Taxonomy was assigned to ASVs using a 16S V4 classifier that is generated using the SILVA database SILVA_138.1_SSURef_NR99. Afterwards, Mitochondria, Chloroplast, Archaea,

Eukaryota, and ambiguous_taxa were filtered out to keep only the bacteria. From the future table, *Microcystis* ASVs and their relative abundance across all samples were pooled separately. Percent relative abundance of the *Microcystis* ASVs per individual sample was calculated using the total bacterial relative abundance of the same sample and used for comparison between the samples. Additionally, *Microcystis* ASVs that were identified in the sediment cores were compared with the *Microcystis* ASVs detected within samples from the water column. The sequencing data for water column samples came from NOAA-GLERL's 2022 Lake Erie Weekly Field Sampling Data. The plots for analyzing the sequencing data were made in Microsoft Excel.

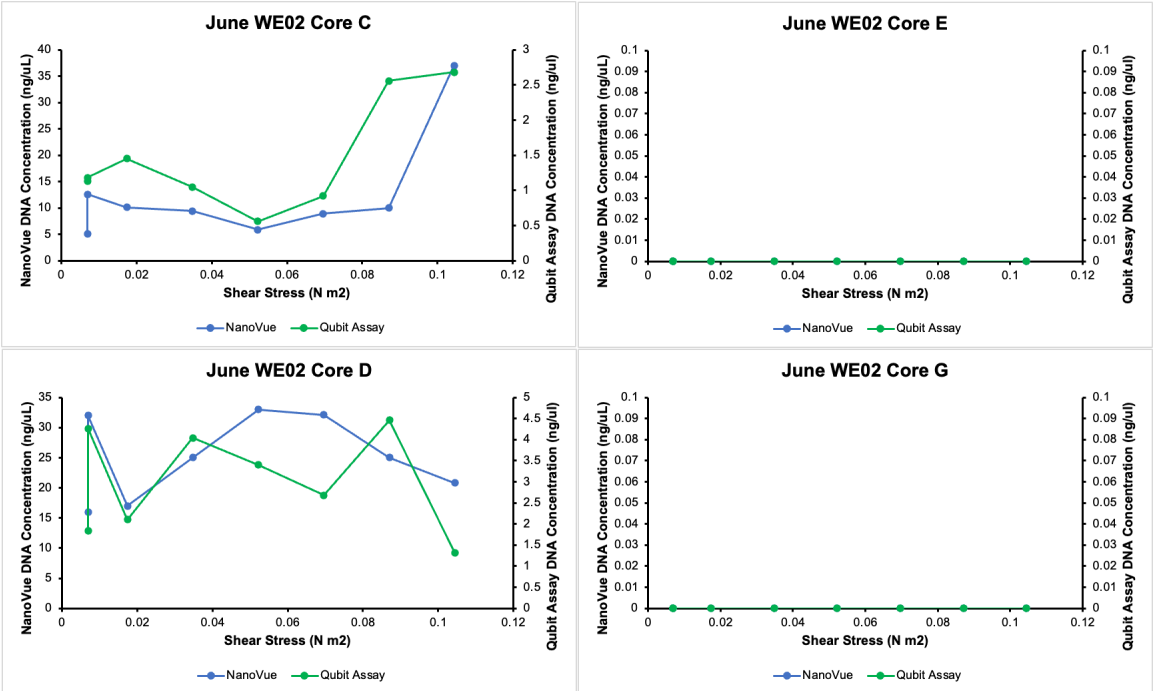
Results

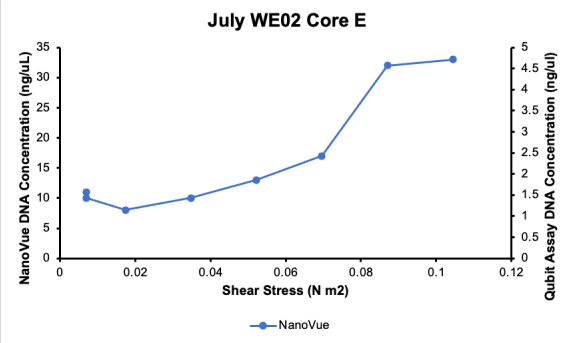
DNA Extraction and Quantification

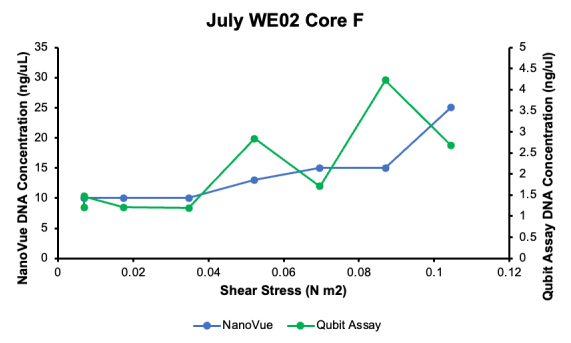
The concentration of DNA from each DNA extraction was measured and compared. The samples that were analyzed in June 2022 show variation within the measured DNA concentration between the different cores from site WE02, but the DNA concentration was relatively low (Figure 3.1). Overall, the NanoVue DNA Concentration measurements ranged from 0-37 ng/μL, and the Qubit Assay DNA Concentration measurements ranged from 0-4.46 ng/μL. Only Cores C and D showed reliable detection of DNA, with Core D having a higher concentration (peaked at 33 ng/μL for NanoVue during the 5th time interval at 0.052244 N m^{-2} shear speed and at 4.46 ng/μL for Qubit Assay during the 8th time interval at 0.104488 N m^{-2} shear speed), until the final two shear speed intervals, where Core C showed a spike in concentration (37 ng/μL NanoVue and 2.68 ng/μL Qubit Assay at 0.104488 N m^{-2} shear speed). Cores E

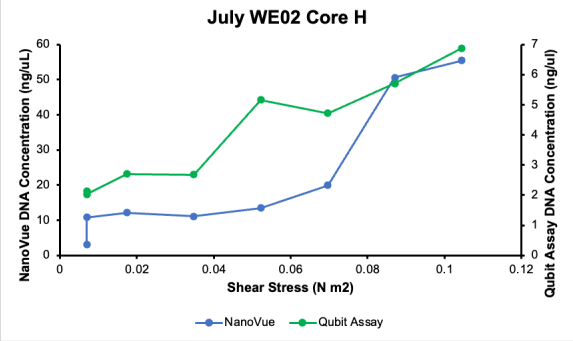
and G did not show clear DNA concentration. Since there were only samples collected from site WE02 during the early stage of the bloom, we were unable to assess the DNA concentration at the other study sites for that time point.

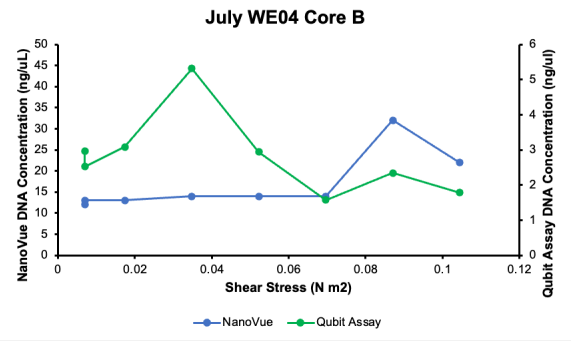
The samples that were later collected in July 2022 also show variation in DNA concentration between the different sites and cores (Figure 3.1). Compared to the earlier samples, there was a greater concentration of DNA consistently found across the July samples that spanned all 5 study sites. Overall, the NanoVue DNA Concentration measurements ranged from 3-55.4 ng/μL, and the Qubit Assay DNA Concentration measurements ranged from 1.19-8.92 ng/μL. While there was only the NanoVue DNA concentration measured for WE02 Core E and WE06 Core C, the other July cores show interesting trends between the NanoVue and Qubit Assay measurements of DNA concentration. While site WE08 showed a relatively similar looking pattern between the NanoVue and Qubit Assay DNA concentrations, the other cores showed different patterns (Figure 3.1). Overall, there was enough DNA in most samples across all of the cores and sites to justify proceeding with qPCR.

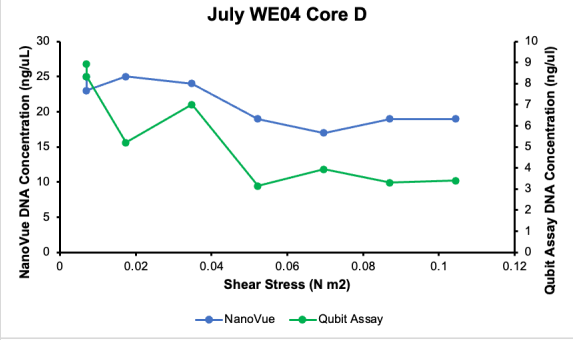


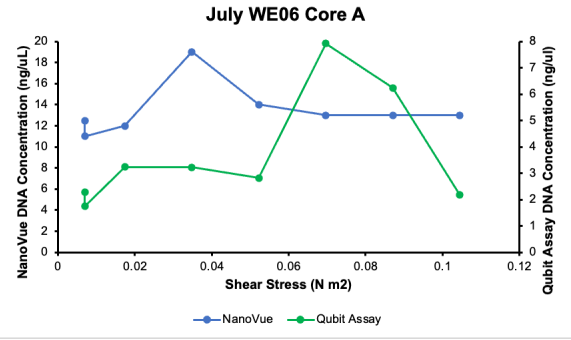

July WE02 Core E
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue

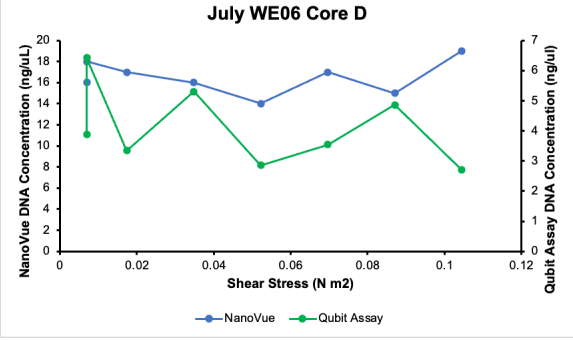

July WE02 Core F
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue
Qubit Assay

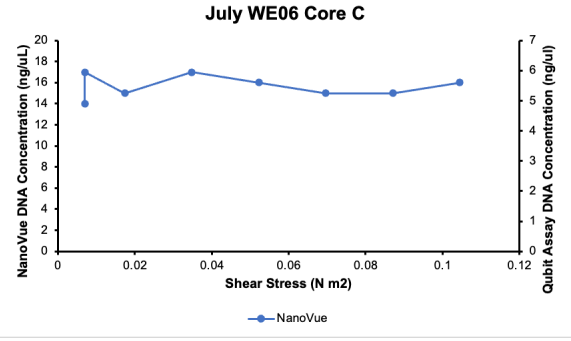

July WE02 Core H
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue
Qubit Assay


July WE04 Core B
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue
Qubit Assay


July WE04 Core D
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue
Qubit Assay


July WE06 Core A
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue
Qubit Assay


July WE06 Core D
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue
Qubit Assay


July WE06 Core C
NanoVue DNA Concentration (ng/uL)
Qubit Assay DNA Concentration (ng/ul)
Shear Stress (N m2)
NanoVue

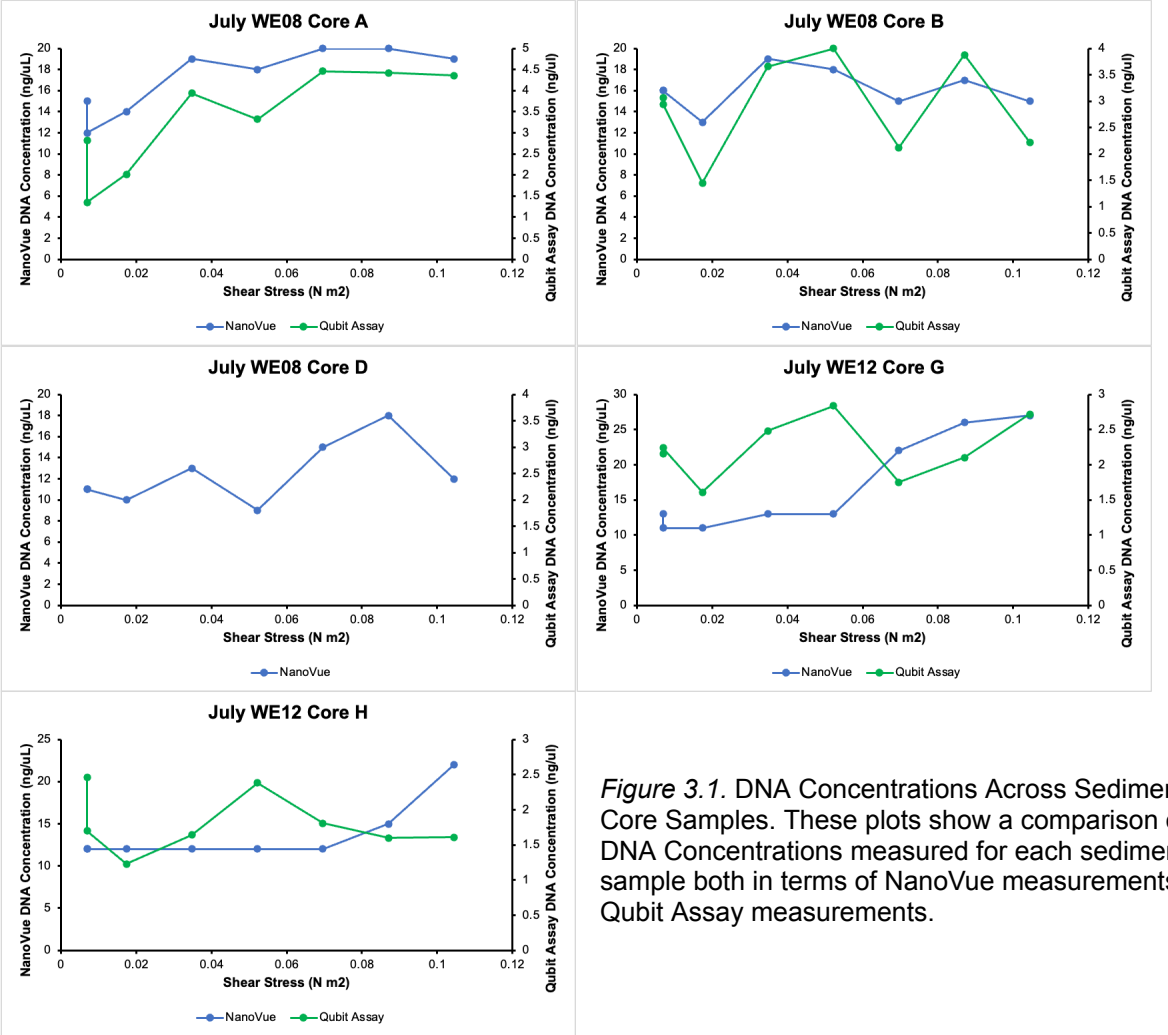


Figure 3.1. DNA Concentrations Across Sediment Core Samples. These plots show a comparison of the DNA Concentrations measured for each sediment core sample both in terms of NanoVue measurements and Qubit Assay measurements.

Quantitative Polymerase Chain Reaction

From the qPCR, it was found that all sediment core samples tested showed signals for the presence of cyanobacteria (Figure 3.2). Although there is low biomass expected towards the start of the bloom, the signals from the cores analyzed in June 2022 show that there is *Microystis* with some toxins present at that time (Figure 3.3). When looking at Figure 3.2, there was a range in Total Cyanobacteria of 1-12,502 copies/mL in June 2022 (Figure 3.2). In terms of the mcyE gene that is linked to

microcystin production, there is a range of 0-4 copies/mL in June 2022 (Figure 3.3). This means that although the water column had low biomass towards the start of the bloom, there is still detectable *Microcystis* in the sediments, but some of the *Microcystis* is likely non-toxic.

Later into the bloom, the cores analyzed in July 2022 show that there is an overall increase in the amount of Total Cyanobacteria detected, along with the amount of the mcyE gene detected. For the Total Cyanobacteria, July 2022 showed a range of 2449-471,807 copies/mL, which is much higher than the range from June 2022. The mcyE gene had a detected range of 0-10,036 copies/mL, which also portrays an increase from the June 2022 samples. However, since the low end of the mcyE is still 0 copies/mL, there are instances where no toxin strains are detected.

When comparing the overall trends between Total Cyanobacteria and the presence of the mcyE gene, there is a lot more variability shown in the detection of the mcyE gene over the course of the difference shear speeds from the sediment core experiments (Figure 3.3). While some cores, such as those collected from site WE08 in July 2022 show consistent patterns across both Total Cyanobacteria and mcyE gene, many other cores show differences.

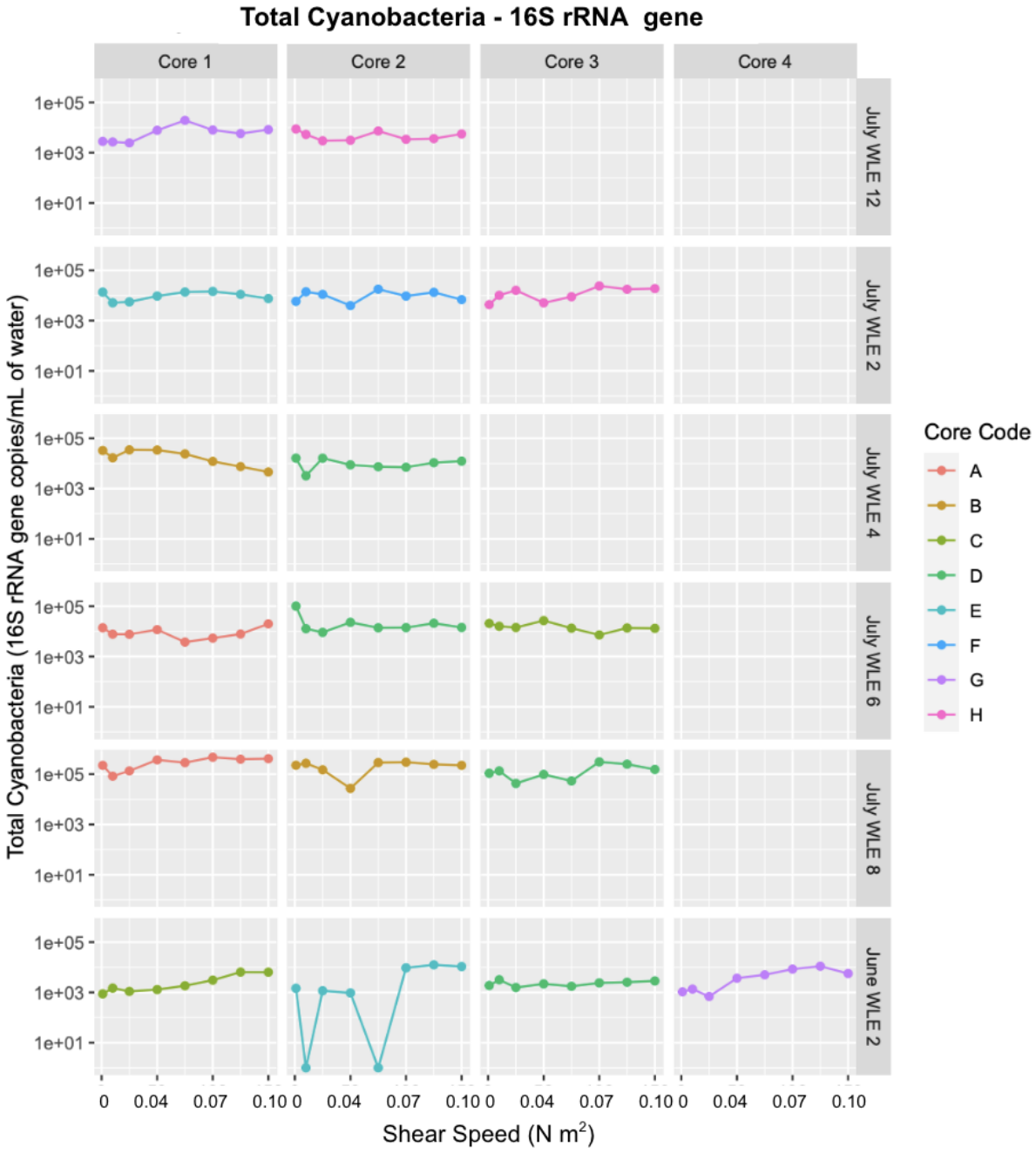


Figure 3.2. Total Cyanobacteria Measured from 16S rRNA Gene Using qPCR. These plots show Total Cyanobacteria detected in each sediment core over the different experimental time intervals. Total Cyanobacteria was measured by taking the number of 16S rRNA gene copies per mL of sample.

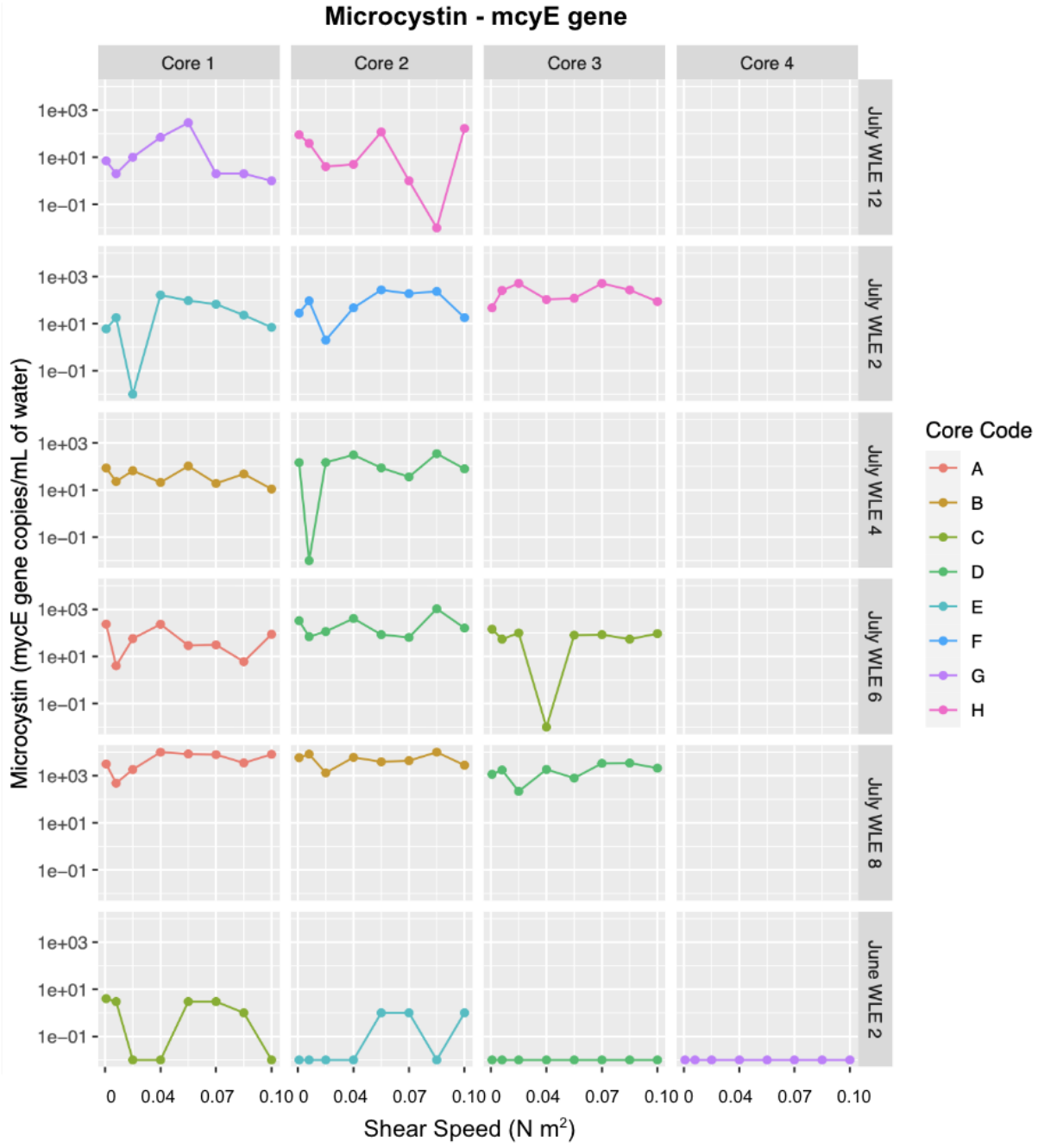


Figure 3.3. Microcystin Measured from mcyE Gene Using qPCR. These plots show the mcyE gene detected across the sediment core samples over different speeds of shear stress. Microcystin was measured by taking the amount of mcyE gene copies per mL of sample.

High Throughput 16S rRNA Amplicon Sequencing

There were 13 unique *Microcystis* ASVs initially detected from the 16S rRNA sequencing data associated with the sediment cores, but after the data was quality checked to limit noise, only 3 were kept as valid ASVs for further analysis. The 3 ASVs were named based on common nucleotide combinations identified. As seen in Figure 3.4, the 3 ASVs are labeled as ATC, CCG, and CTG+AGC. These ASVs were found to be very dominant within the sediment, as they could be tracked across the sediment cores. Most of the different sampling sites only showed 2 of the ASVs, CCG and CTG+AGC. The third ASV, ATC, could be found in some of the sediment cores, as shear speed was increased a few steps but was more rare overall.

In addition to variation in the presence of ASVs, there were also variations across different sites, in the measured relative abundance of 16s rRNA from each individual sample. The lowest relative abundance was observed during the early bloom samples at site WE02, with all of the July samples showing higher abundance. However, within the July samples, there was notable variation specifically between the cores from site WE08 and the rest of the samples. The cores from site WE08 show much greater 16s rRNA relative abundance.

After analyzing the sequencing data that came from the sediment cores, the sequencing data from the water column was also analyzed to check for the same ASVs. Based on the results from the weekly monitoring shown in Figure 3.5, the water column shows clear evidence of the same ASVs from the sediment cores also being present in the water column. Similar to the sediment core data, CCG and CTG+AGC were the more common ASVs, with ATC occasionally appearing at the different study sites.

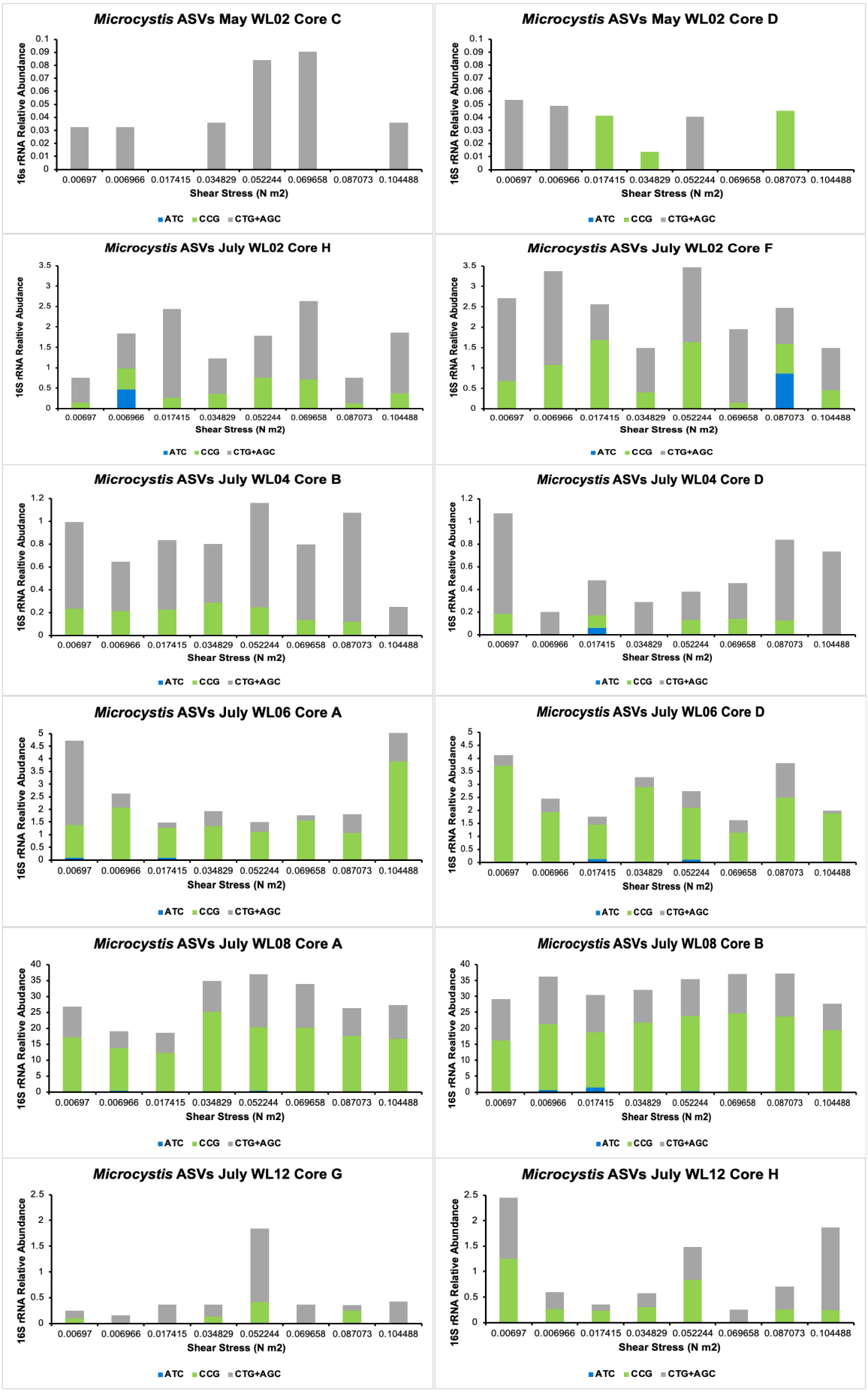

Microcystis ASVs May WL02 Core C
Microcystis ASVs May WL02 Core D
Microcystis ASVs July WL02 Core H
Microcystis ASVs July WL02 Core F
Microcystis ASVs July WL04 Core B
Microcystis ASVs July WL04 Core D
Microcystis ASVs July WL06 Core A
Microcystis ASVs July WL06 Core D
Microcystis ASVs July WL08 Core A
Microcystis ASVs July WL08 Core B
Microcystis ASVs July WL12 Core G
Microcystis ASVs July WL12 Core H
16S rRNA Relative Abundance
Shear Stress (N m2)
0.00697 0.006966 0.017415 0.034829 0.052244 0.069658 0.087073 0.104488
ATC CCG CTG+AGC

Figure 3.4. Microcystis ASVs Across All Sediment Core Samples. These plots show the distribution of the 3 *Microcystis* ASVs for each core at each given time interval.

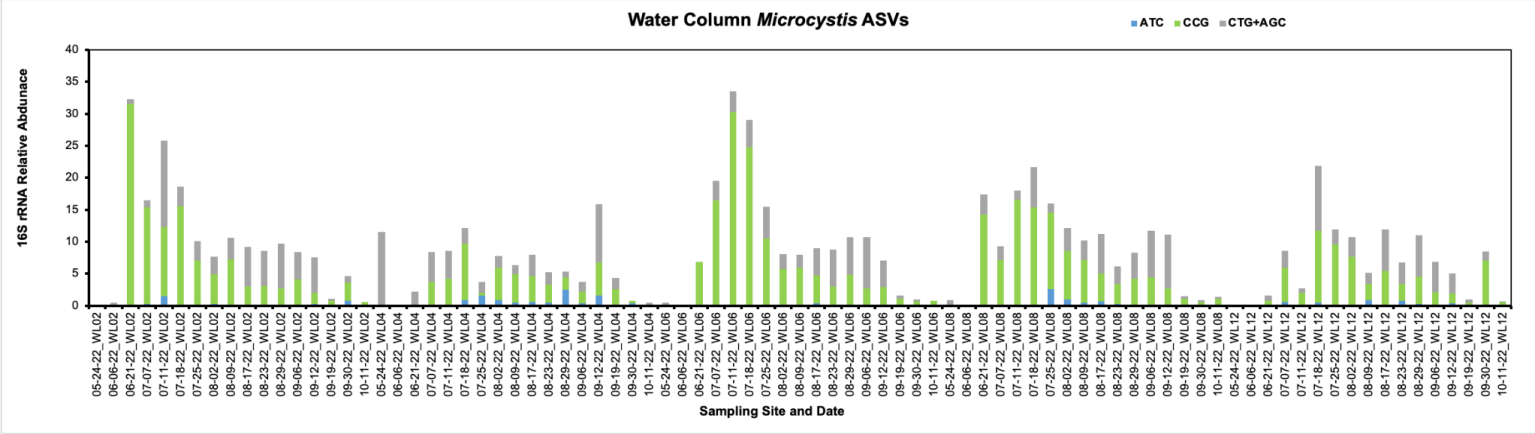


Figure 3.5. Microcystis ASVs from the Water Column over Time. This plot shows the presence of the three identified ASVs across water column samples that were collected as a part of NOAA-GLERL's weekly Lake Erie HABS monitoring between May 2022 - October 2022.

Objective 4: Statistical Modeling

In objective 4, we utilized statistical models to make predictions about concentrations of cyanobacterial and algal biomass at various monitoring sites located in western Lake Erie. Initially, we performed change point analysis on extracted chlorophyll and dissolved microcystin to identify potential relationships between them and environmental variables. Next, we implemented generalized linear models with stepwise model selection to determine relevant environmental variables by analyzing observational data. These variables included nutrient concentrations, river discharge, surface water temperature, and wind speed. In the future, we plan to evaluate the impact of sediment resuspension by comparing the performance of models that do and do not include an index of resuspension (calculated from output from Objectives 1 and 2).

Methods

Data

For this objective, we used generalized linear models to predict cyanobacterial and algal biomass concentrations by environmental variables. To accomplish this, we collected several datasets, including physical, chemical, and biological water quality monitoring data (NCEI, 2019), nutrient concentration data from the Maumee river (National Center for Water Quality Research), Thro1h station buoy data (National Data Buoy Center, 2021), Maumee river discharge data (USGS, 2022), and sediment resuspension data from objective 1.

The physical, chemical, and biological water quality monitoring data were collected each week or every other week from May to October since 2017 from nine stations in West Lake Erie: WE2, WE4, WE6, WE8, WE12, WE13, WE15, and WE16. This dataset includes important variables that measure the biomass of algae, such as extracted chlorophyll, dissolved microcystin, and particulate microcystin.

The nutrient concentration data contains nutrient data of Maumee river such as total phosphorus, soluble reactive phosphorus, nitrate, and nitrogen, and is collected daily. The Thro1h station buoy data contains physical variables like wind speed, wind direction, air temperature, and air pressure data of station Thro1h, and is collected every six minutes. Finally, the Maumee river discharge data is collected every 30 minutes.

Change point analysis

Change point analysis is a statistical technique employed to detect changes in the underlying characteristics of a given dataset. In this study, we conducted change point analyses with varying time lags to determine the timing of bloom events and their relationship with recent environmental conditions. We employed extracted chlorophyll and dissolved microcystin as observational variables and identified the change points of the concentration of these two variables from the sampling data. Subsequently, we visualized the change points alongside the environmental variables. This allowed us to investigate potential relationships between sudden changes in bloom events and other environmental variables.

Statistical modeling

Harmful algal blooms (HABs) are influenced by various environmental factors such as sunlight, temperature, nutrients, salinity, wind and water currents (NOAA, 2019). High concentrations of nutrients, including nitrogen and phosphorus, can lead to the overgrowth of algae and contribute to the development of HABs. Sources of these nutrients can include agricultural runoff, sewage treatment plants, and stormwater runoff (Paerl & Otten, 2013). Algae thrive in warm water, so warmer temperatures can also promote the growth of HABs (Stumpf et al., 2012). High levels of light are required for photosynthesis, which is why sunlight can also promote the growth of HABs. However, strong water currents can help disperse algae and prevent the development of HABs (Anderson et al., 2002).

To evaluate the role of sediment resuspension, we constructed generalized linear models using stepwise model selection to identify relevant environmental variables from available observational data. We included extracted chlorophyll and dissolved microcystin as observational variables and tried various combinations of environmental variables to minimize the AIC value of our models and determine statistically significant variables. The environmental variables include total phosphorus, soluble reactive phosphorus, Maumee river discharge, air temperature, air pressure, wind speed, etc.

It is worth noting that there can be time lags between the increase in nutrient concentration and the occurrence of HABs (Urban et al., 2006). To account for these potential time lags in the impact of environmental variables on our observational variables, we introduced time lags ranging from day 0 to day 8 and applied smoothing techniques to our model.

Results

Change point analysis

We conducted change point analysis on the extracted chlorophyll and dissolved microcystin data from 2017 to 2021 using binary segmentation. The observational variables were averaged using the sampling data of WE2, WE4, WE6, WE8, and WE12. Our analysis showed that the time of change points had a strong seasonal pattern. Specifically, the concentration of extracted chlorophyll increased around Julian date 220 (early August), while the decrease occurred around Julian date 250 (early September). This pattern was similar to the activity of the Maumee river, which also had similar increases and decreases in the summer season.

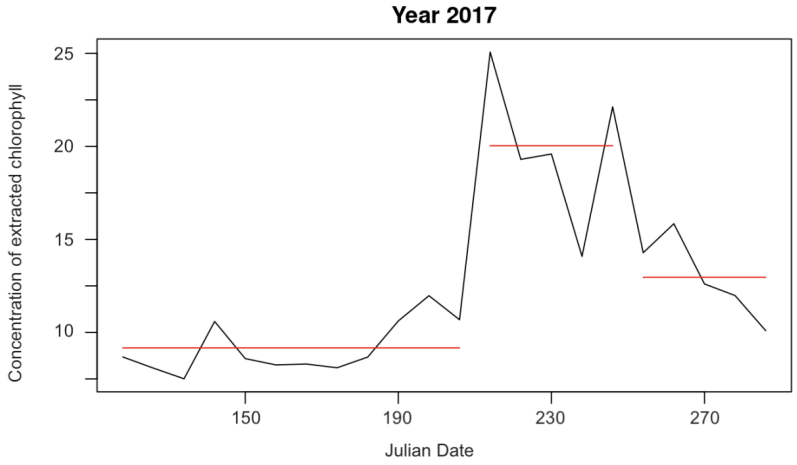


Figure 4.1. Change point analysis of year 2017.

To identify potential relationships between the change points and environmental variables, we plotted changepoints and environmental variables. Figure 4.2 is an example of our visualizations. In the figure, the green lines represent the increase in extracted chlorophyll, while the red lines represent the decrease. The y-axis represents soluble reactive phosphorus, and the x-axis represents Julian date. The plot shows that

the change points occur around Julian date 230 to 250, which coincides with a peak in the concentration of soluble reactive phosphorus, albeit with time lags. This plot suggests a potential relationship between soluble reactive phosphorus and the change points of extracted chlorophyll.

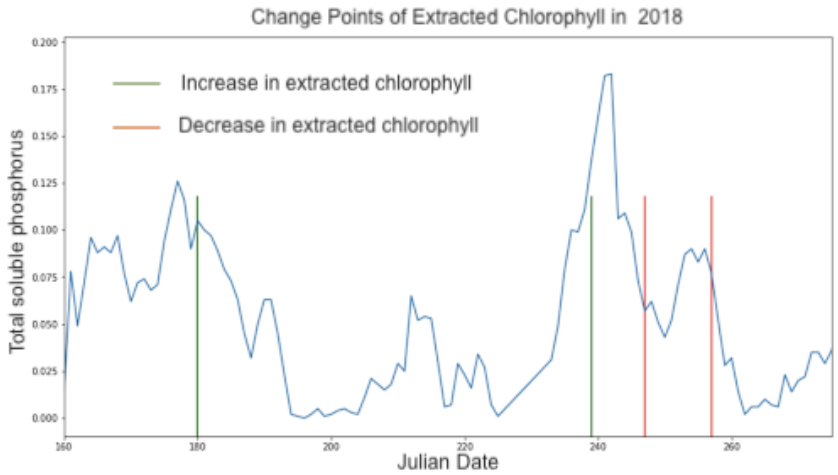


Figure 4.2. Change points and soluble reactive phosphorus of year 2018.

Statistical modeling

Models for extracted chlorophyll

Aftering identifying the potential relationships between the change points and environmental variables, we constructed generalized linear models to predict extracted chlorophyll and dissolved microcystin, using stepwise model selection. The potential variables we included in our models were suspended solids, Maumee river discharge, Maumee river total phosphorus, soluble reactive phosphorus, total nitrogen, wind direction, wind speed, gust speed, air temperature, and dew point temperature. To improve the models' predictive capabilities, we created new smoothed variables and implemented time lags of up to eight days to capture the time lags between the variables and algae biomass.

Our main objective was to find models that can accurately predict our observational variables while also minimizing the Akaike Information Criterion (AIC). We found six models that outperformed the others in our study. These models are detailed in the appendix.

Model 1 (Figure A1) consisted of four variables: total phosphorus with a 1-day time lag, smoothed air temperature with an 8-day time lag, gust speed with a 4-day time lag, and wind direction with a 1-day time lag. Wind direction is a categorical variable indicating whether the wind is west wind. The coefficients of these variables varied, with total phosphorus and gust speed having positive coefficients, while wind direction and air temperature had negative coefficients. Most of the variables were statistically significant except gust speed, and the model had an AIC value of 2149.3.

Model 2 (Figure A2) used the same variables as Model 1 but replaced gust speed with wind speed with a 4-day time lag, resulting in a slightly lower AIC value of 2149.2. Model 3 (Figure A3) introduced air pressure with a 4-day time lag based on Model 1, with a positive coefficient and a slightly reduced AIC value of 2142.2. Model 4 (Figure A4) replaced total phosphorus of Model 1 with Maumee river discharge with a 1-day time lag, achieving the lowest AIC value of 2128.4.

Models for dissolved microcystin

Similar to models for extracted chlorophyll, after testing various models, we selected five that showed superior performance compared to others. In Model 1 (Figure A7), we achieved an impressive AIC value of -1291.2 by incorporating wind speed with 5-day time lags, air temperature with 7-day time lags, total kjeldahl nitrogen with 1-day

time lag, nitrate concentration with 7-day time lags, sulfate concentration with 7-day time lags, total phosphorus with 6-day time lags, and air pressure with 1-day time lag.

Model 2 (Figure A8) includes gust speed with 5-day time lags, dew point temperature with 7-day time lags, nitrate concentration with 7-day time lags, sulfate concentration with 7-day time lags, total kjeldahl nitrogen with 3-day time lags, and total phosphorus with 6-day time lags. Similarly, Model 3, 4, and 5 also exhibit comparable AIC values by substituting some environmental variables of Model 1 and 2.

Discussion

The work conducted in objective 1 provided a total bottom stress output that can be used as a predictor variable in future Bayesian hierarchical modeling to better forecast HAB events. Additionally, the temporal and spatial variation of bottom stress was further explained in the western basin of Lake Erie. These high bottom stress values, and therefore resuspension events, tend to occur in areas of 5 meters or less in depth. These areas are not only highly recreated, but are also near public water intake sites that provide potable water to hundreds of thousands of residents.

Dredging takes place annually in the western basin to maintain access for deep draft commercial navigation in designating shipping lanes. This requires the removal of sediment to provide adequate depth (~28 feet) for the passage of container ships which may affect sediment resuspension and turbidity observations. Although this objective only calculated values of total bottom stress where sediment resuspension would most likely occur, future work should take into consideration the effect of dredging on sediment resuspension observations.

Based on the EROMES calibration approach (Kalnejais et al. (2006) we identified critical total bottom stress values (hereon referred to as bottom stress) required for resuspension of cyanobacteria (*Microcystis*) in sediment core samples collected from five sites in Lake Erie in May and July 2022. As mentioned earlier we proposed three possible scenarios: 1) Cyanobacteria (*Microcystis*) in the sediment core would resuspend at a lower TBS than the sediment. This would imply that cyanobacteria are present in the upper layers of sediment and would readily resuspend when bottom stress on the lakebed crosses the threshold value. 2) cyanobacteria and sediment at the same critical total bottom stress, suggesting that cyanobacteria and sediment are homogenized on the lakebed. 3) Sediment from the lakebed would resuspend at a lower bottom stress compared to the cyanobacteria, indicating the cyanobacteria are present beneath the sediment and can only resuspend after the initial erosion of the top sediment layer.

Even though the five sampling sites represent a range of depths, different substrate types, and depositional environments we did not observe a clear trend in the changepoint estimates for both cyanobacteria and sediment across the sediment cores. We collected sediment cores in late May and early July and did not see a clear trend in change point estimates for samples collected at different sampling occasions. It is likely that lakebed conditions during these periods are similar and sediment cores representing different time points could give a better understanding of resuspension trends.

We found that the median bottom stress required for resuspension of cyanobacteria was 0.060951 N m^{-2} while resuspension of sediment was 0.069658 N m^{-2}.

This indicates that cyanobacteria possibly resuspend into the water column at a slightly lower bottom stress as compared to the sediment. However, we did not quantify if this difference in critical bottom stress value required for resuspension was significant. Furthermore, we mapped changepoint estimates which were continuous onto discrete intervals when the corresponding bottom stress was exerted, which would have resulted in potential loss of information.

Using these critical bottom stress values and outputs of total bottom stress from WW3 and FVCOM, it would be informative to identify how many times these thresholds are crossed and further investigate if blooms of toxic *Microcystis* align with the total bottom stress exceeding thresholds in Lake Erie.

The genomics work in this study allowed for ASVs unique to *Microcystis* to be identified in both the sediments and the water column of Lake Erie, early-on and midway through the algal bloom. Due to the variation in the presence of the ASVs, it seems that different strains of *Microcystis* may appear in the water column as the shear speed is increased. This is especially supported by the ATC ASV only appearing after certain thresholds of shear speed were met (Figure 3.4). Since the shear speed seems capable of potentially controlling which species of *Microcystis* would be resuspended, sediment resuspension across the lake likely has an important impact on which *Microcystis* genotypes are present.

Although our sediment core sampling was limited, the water column data from NOAA-GLERL was collected for a majority of the bloom's duration, which allowed for the ASVs to be tracked over the rise, peak, and decline of the 2022 HAB. The water column data from May and early June barely picked up any detectable *Microcystis*

ASVs, but the results from our sediment core ASVs from that time suggest that *Microcystis* found within the sediments was likely coming up into the water column later into the summer season (Figure 3.2 and 3.3).

When looking at the cores from each site, the samples from site WE08 are particularly interesting, as they seem to provide the most clear trends, compared to the other sampling sites. Since the cores from WE08 were much higher in measured relative abundance of the 16S rRNA gene for Total Cyanobacteria and the mcyE gene, shown within the qPCR data, it seems to suggest that at that site in particular, the *Microcystis* from the sediments is likely coming up in the water column. There seems to still be similar trends at the other sites, but they are more challenging to confirm due to lower concentrations of biomass detected.

In terms of taking toxicity of the bloom into account, the results from the qPCR analysis allowed for the presence of microcystin to be detected across multiple samples, which helped to confirm that there are some toxic strains of *Microcystis* in both the sediments and the water column. Due to some qPCR results showing 0 copies/mL even when Total Cyanobacteria and *Microcystis* ASVs were detected, there also is potential evidence of non-toxic strains also being present (Figure 3.3). However, we have not fully isolated the ASVs, so we are unable to pinpoint which are linked to toxic strains compared to non-toxic strains, since the 16S rRNA data alone does not provide that level of detail. Future work may help to further investigate this.

Overall there may be multiple reasons that there is variation between the genomics data collected from different sites and cores. Irregular patterns could be related to the fact that the different sites had sand grains of different sizes. There is also

the potential for matchiness in the data. This is because if there were *Microcystis* colonies significantly smaller or larger compared to others in the sample, the data could be skewed.

In this study, we also utilized statistical models to predict the concentrations of cyanobacterial and algal biomass at various monitoring sites in western Lake Erie. Firstly, we conducted a change point analysis on chlorophyll and dissolved microcystin data to identify potential relationships between these variables and environmental factors. We found that there was a strong seasonal pattern in the change points, with increases in extracted chlorophyll occurring around early August and decreases occurring in early September.

To determine the relevant environmental variables, we used generalized linear models with stepwise model selection. We built two series of generalized linear models to predict extracted chlorophyll and dissolved microcystin, which had high model performance. In the models, we identified environmental variables that were significant predictors of cyanobacterial and algal biomass concentrations in our study sites.

In the future, to assess the impact of sediment resuspension and *Microcystis* sediment populations on harmful algal blooms (HABs) in Lake Erie, we will incorporate the resuspension index obtained from objectives 1 and 2 into our models, and compare the performance of models with and without an index of resuspension.

Conclusion

Overall, this study provides further knowledge towards the effects that *Microcystis* resuspension has on Lake Erie HABs. The hydrodynamic modeling completed in Objective 1 provided spatial and temporal measurements of total bottom stress that relates to sediment resuspension. This knowledge is especially important when *Microcystis* has been shown to overwinter in the sediment layer and therefore, may better predict future HAB events. From Objective 2's laboratory experiments, we determined critical bottom stress values at which cyanobacteria and sediment resuspended in the water column and found that cyanobacteria resuspend at a slightly lower bottom stress compared to the sediment. The genomics work completed in Objective 3 supports that *Microcystis* found in the sediments is consistent with the *Microcystis* later found within the water column during blooms. Statistical modeling from objective 4 identifies the environmental variables that affect the concentrations of cyanobacterial and algal biomass in Lake Erie. Additionally, the models will include the resuspension index to investigate the impact of *Microcystis* resuspension on HABs in the lake in the future.

Appendix

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.84744  -0.46185  -0.09599   0.29784   1.73821

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)              0.1277742  0.0348071   3.671 0.000289 ***
total_phosphorus_1day    0.0904330  0.0242064   3.736 0.000226 ***
atmp_smooth_8day        -0.0022618  0.0004770  -4.742 3.35e-06 ***
gst_m.s_4day             0.0007142  0.0011830   0.604 0.546529
wdir_if_west_1dayWest -0.0180745  0.0037774  -4.785 2.76e-06 ***
siteWE1                 -0.0304416  0.0342814  -0.888 0.375299
siteWE12                -0.0236637  0.0328767  -0.720 0.472260
siteWE13                 0.0020727  0.0336860   0.062 0.950981
siteWE14                -0.0159458  0.0355532  -0.449 0.654132
siteWE16                -0.0105879  0.0333337  -0.318 0.750998
siteWE17                 0.0194751  0.0510998   0.381 0.703401
siteWE2                 -0.0236157  0.0327653  -0.721 0.471656
siteWE4                 -0.0024434  0.0332187  -0.074 0.941417
siteWE5                 -0.0113314  0.0360074  -0.315 0.753223
siteWE6                 -0.0506877  0.0324400  -1.563 0.119286
siteWE8                 -0.0362015  0.0326465  -1.109 0.268417
siteWE9                 -0.0543440  0.0324139  -1.677 0.094732 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3251383)

    Null deviance: 161.020  on 299  degrees of freedom
Residual deviance:  99.589  on 283  degrees of freedom
  (69 observations deleted due to missingness)
AIC: 2149.3
```

Figure A1 Model 1 for extracted chlorophyll

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.84814  -0.46172  -0.09446   0.29622   1.73552

Coefficients:
                           Estimate Std. Error t value Pr(>|t|)
(Intercept)              0.1274499  0.0347957   3.663 0.000298 ***
total_phosphorus_1day    0.0894935  0.0243342   3.678 0.000282 ***
atmp_smooth_8day        -0.0022649  0.0004777  -4.741 3.37e-06 ***
wspd_m.s_4day            0.0012473  0.0017820   0.700 0.484552
wdir_if_west_1dayWest -0.0180360  0.0037542  -4.804 2.52e-06 ***
siteWE1                 -0.0303757  0.0342849  -0.886 0.376382
siteWE12                -0.0235222  0.0328821  -0.715 0.474983
siteWE13                 0.0022370  0.0336928   0.066 0.947110
siteWE14                -0.0158899  0.0355575  -0.447 0.655305
siteWE16                -0.0104282  0.0333398  -0.313 0.754675
siteWE17                 0.0197563  0.0511253   0.386 0.699469
siteWE2                 -0.0234844  0.0327705  -0.717 0.474192
siteWE4                 -0.0023206  0.0332235  -0.070 0.944363
siteWE5                 -0.0112777  0.0360120  -0.313 0.754387
siteWE6                 -0.0505414  0.0324455  -1.558 0.120414
siteWE8                 -0.0360536  0.0326521  -1.104 0.270456
siteWE9                 -0.0542182  0.0324188  -1.672 0.095544 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3253835)

    Null deviance: 161.020  on 299  degrees of freedom
Residual deviance:  99.547  on 283  degrees of freedom
  (69 observations deleted due to missingness)
AIC: 2149.2
```

Figure A2 Model 2 for extracted chlorophyll

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.81139  -0.44876  -0.08823   0.28335   1.63856

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)            9.150e-02  3.679e-02   2.487  0.01345 *
total_phosphorus_1day  1.119e-01  2.612e-02   4.285 2.51e-05 ***
atmp_smooth_8day      -2.241e-03  4.590e-04  -4.883 1.76e-06 ***
gst_m.s_4day           8.379e-04  1.166e-03   0.719  0.47279
pres_hPa_4day          2.967e-05  1.069e-05   2.775  0.00588 **
wdir_if_west_1dayWest -1.721e-02  3.717e-03  -4.629 5.60e-06 ***
siteWE1               -3.054e-02  3.403e-02  -0.898  0.37016
siteWE12              -2.646e-02  3.266e-02  -0.810  0.41857
siteWE13              -1.939e-04  3.347e-02  -0.006  0.99538
siteWE14              -1.617e-02  3.529e-02  -0.458  0.64710
siteWE16              -1.293e-02  3.312e-02  -0.391  0.69645
siteWE17               2.049e-02  5.074e-02   0.404  0.68660
siteWE2               -2.597e-02  3.255e-02  -0.798  0.42568
siteWE4               -4.292e-03  3.300e-02  -0.130  0.89660
siteWE5               -1.159e-02  3.574e-02  -0.324  0.74597
siteWE6               -5.325e-02  3.222e-02  -1.653  0.09952 .
siteWE8               -3.898e-02  3.243e-02  -1.202  0.23045
siteWE9               -5.590e-02  3.219e-02  -1.737  0.08357 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3208893)

    Null deviance: 161.020  on 299  degrees of freedom
Residual deviance:  96.742  on 282  degrees of freedom
  (69 observations deleted due to missingness)
AIC: 2142.2
```

Figure A3 Model 3 for extracted chlorophyll

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.81045  -0.46796  -0.09702   0.24246   1.29109

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)            9.894e-02  3.503e-02   2.825  0.00507 **
flow_cfs_1day          1.987e-06  3.868e-07   5.137 5.21e-07 ***
atmp_degC_8day        -1.600e-03  3.685e-04  -4.341 1.98e-05 ***
gst_m.s_4day           1.286e-03  1.130e-03   1.138  0.25605
pres_hPa_4day          2.494e-05  1.027e-05   2.428  0.01579 *
wdir_if_west_1dayWest -1.753e-02  3.616e-03  -4.848 2.06e-06 ***
siteWE1               -3.298e-02  3.322e-02  -0.993  0.32177
siteWE12              -2.840e-02  3.191e-02  -0.890  0.37429
siteWE13              -2.254e-03  3.270e-02  -0.069  0.94510
siteWE14              -1.884e-02  3.444e-02  -0.547  0.58489
siteWE16              -1.484e-02  3.236e-02  -0.459  0.64694
siteWE17               1.884e-02  4.957e-02   0.380  0.70421
siteWE2               -2.832e-02  3.180e-02  -0.891  0.37393
siteWE4               -6.749e-03  3.224e-02  -0.209  0.83433
siteWE5               -1.432e-02  3.488e-02  -0.410  0.68178
siteWE6               -5.528e-02  3.148e-02  -1.756  0.08020 .
siteWE8               -4.113e-02  3.169e-02  -1.298  0.19536
siteWE9               -5.769e-02  3.146e-02  -1.834  0.06770 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3056254)

    Null deviance: 161.02  on 299  degrees of freedom
Residual deviance:  92.62  on 282  degrees of freedom
  (69 observations deleted due to missingness)
AIC: 2128.4
```

Figure A4 Model 4 for extracted chlorophyll

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8184  -0.4605  -0.1060   0.2770   1.5576

Coefficients:
                         Estimate Std. Error t value Pr(>|t|)
(Intercept)             9.088e-02  3.693e-02   2.461  0.01447 *
total_phosphorus_1day   1.149e-01  2.620e-02   4.386 1.63e-05 ***
atmp_smooth_8day       -2.209e-03  4.507e-04  -4.901 1.61e-06 ***
gst_m.s_4day            1.243e-03  1.162e-03   1.069  0.28578
pres_hPa_4day           3.171e-05  1.084e-05   2.926  0.00371 **
wind_dir_1day          -8.051e-05  1.881e-05  -4.279 2.57e-05 ***
siteWE1                -2.939e-02  3.409e-02  -0.862  0.38924
siteWE12               -2.621e-02  3.273e-02  -0.801  0.42390
siteWE13               -2.121e-04  3.354e-02  -0.006  0.99496
siteWE14               -1.505e-02  3.535e-02  -0.426  0.67066
siteWE16               -1.309e-02  3.319e-02  -0.394  0.69361
siteWE17                2.234e-02  5.081e-02   0.440  0.66043
siteWE2                -2.551e-02  3.261e-02  -0.782  0.43486
siteWE4                -3.968e-03  3.307e-02  -0.120  0.90458
siteWE5                -1.047e-02  3.580e-02  -0.292  0.77016
siteWE6                -5.303e-02  3.229e-02  -1.642  0.10161
siteWE8                -3.875e-02  3.250e-02  -1.192  0.23422
siteWE9                -5.550e-02  3.225e-02  -1.721  0.08643 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3217463)

    Null deviance: 161.020  on 299  degrees of freedom
Residual deviance:  98.007  on 282  degrees of freedom
  (69 observations deleted due to missingness)
AIC: 2146.3
```

Figure A5 Model 5 for extracted chlorophyll

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8088  -0.4769  -0.0725   0.2395   1.2268

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              1.047e-01  3.546e-02   2.952  0.00342 **
flow_cfs                 2.667e-06  4.998e-07   5.337 1.95e-07 ***
nitrate_nitrate         -1.283e-03  6.231e-04  -2.059  0.04043 *
atmp_degC_8day          -1.879e-03  3.811e-04  -4.932 1.39e-06 ***
wspd_m.s_4day            2.192e-03  1.677e-03   1.307  0.19231
pres_hPa_4day            2.058e-05  1.047e-05   1.966  0.05028 .
wdir_if_west_1dayWest -1.766e-02  3.524e-03  -5.010 9.63e-07 ***
siteWE1                 -2.708e-02  3.290e-02  -0.823  0.41112
siteWE12                -2.179e-02  3.167e-02  -0.688  0.49208
siteWE13                 4.495e-03  3.245e-02   0.139  0.88994
siteWE14                -1.329e-02  3.409e-02  -0.390  0.69706
siteWE16                -8.019e-03  3.212e-02  -0.250  0.80301
siteWE17                 2.852e-02  4.922e-02   0.579  0.56278
siteWE2                 -2.197e-02  3.155e-02  -0.696  0.48678
siteWE4                 -3.679e-04  3.198e-02  -0.012  0.99083
siteWE5                 -8.856e-03  3.452e-02  -0.257  0.79771
siteWE6                 -4.893e-02  3.125e-02  -1.566  0.11854
siteWE8                 -3.464e-02  3.145e-02  -1.101  0.27164
siteWE9                 -5.138e-02  3.123e-02  -1.645  0.10101
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.3004845)

    Null deviance: 161.020  on 299  degrees of freedom
Residual deviance:  91.499  on 281  degrees of freedom
  (69 observations deleted due to missingness)
AIC: 2126.6
```

Figure A6 Model 6 for extracted chlorophyll

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5636  -0.6078  -0.2415   0.2564   2.3295

Coefficients:
                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                  -144.59745   41.49471  -3.485 0.000535 ***
SiteWE13                       2.19319    1.12310   1.953 0.051384 .
SiteWE15                       2.44130    2.03524   1.200 0.230878
SiteWE16                       0.67398    1.08511   0.621 0.534798
SiteWE2                       -1.65091    0.76769  -2.150 0.031981 *
SiteWE4                        1.52092    1.04248   1.459 0.145186
SiteWE6                       -1.93908    0.74696  -2.596 0.009701 **
SiteWE8                       -1.22840    0.79947  -1.537 0.125026
SiteWE9                       -1.85297    0.75444  -2.456 0.014375 *
wspd_m.s_lag5                  0.79086    0.23246   3.402 0.000720 ***
atmp_degC_lag7                -0.25943    0.06224  -4.168 3.60e-05 ***
total_kjeldahl_nitrogen_lag1   3.67583    0.66035   5.567 4.18e-08 ***
nitrate_nitrate_lag7           0.66924    0.11769   5.687 2.17e-08 ***
sulfate_lag7                   0.10164    0.01461   6.959 1.05e-11 ***
total_phosphorus_lag6         19.96850    4.11484   4.853 1.62e-06 ***
pres_hPa_lag1                  0.13851    0.04035   3.433 0.000645 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.5289619)

    Null deviance: 358.72  on 531  degrees of freedom
Residual deviance: 225.86  on 516  degrees of freedom
AIC: -1291.2

Number of Fisher Scoring iterations: 6
```

Figure A7 Model 1 for dissolved microcystin

```
Call:
glm(formula = dissolve_microcystin ~ Site + gst_m.s_lag5 + dewp_degC_lag7 +
    nitrate_nitrate_lag7 + sulfate_lag7 + total_kjeldahl_nitrogen_lag3 +
    total_phosphorus_lag6, family = "Gamma", data = data_nona)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4045  -0.6432  -0.2659   0.2772   2.2673

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  -0.38364    2.44181  -0.157  0.87522
SiteWE13                      2.27669    1.16647   1.952  0.05150 .
SiteWE15                      2.76149    2.08727   1.323  0.18641
SiteWE16                      0.65955    1.13628   0.580  0.56187
SiteWE2                      -1.72717    0.80817  -2.137  0.03305 *
SiteWE4                       1.61112    1.08467   1.485  0.13806
SiteWE6                      -2.02802    0.78727  -2.576  0.01027 *
SiteWE8                      -1.27463    0.84050  -1.517  0.13000
SiteWE9                      -1.93854    0.79500  -2.438  0.01509 *
gst_m.s_lag5                  0.42040    0.16034   2.622  0.00900 **
dewp_degC_lag7               -0.37281    0.06968  -5.350 1.32e-07 ***
nitrate_nitrate_lag7          0.54839    0.13136   4.175 3.50e-05 ***
sulfate_lag7                  0.08903    0.01571   5.666 2.42e-08 ***
total_kjeldahl_nitrogen_lag3  2.10473    0.66505   3.165  0.00164 **
total_phosphorus_lag6        19.76414    4.64426   4.256 2.48e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.5612257)

    Null deviance: 358.72  on 531  degrees of freedom
Residual deviance: 243.65  on 517  degrees of freedom
AIC: -1250

Number of Fisher Scoring iterations: 6
```

Figure A8 Model 2 for dissolved microcystin

```
Call:
glm(formula = dissolve_microcystin ~ Site + wspd_m.s_lag5 + atmp_degC_lag7 +
    discharge_lag0 + pres_hPa_lag1, family = "Gamma", data = data_nona)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4960  -0.6726  -0.2650   0.2328   2.2701

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.019e+02  4.056e+01  -2.513  0.01226 *
SiteWE13        2.410e+00  1.234e+00   1.954  0.05128 .
SiteWE15        2.068e+00  2.222e+00   0.931  0.35236
SiteWE16        9.986e-01  1.203e+00   0.830  0.40703
SiteWE2        -1.831e+00  8.602e-01  -2.128  0.03377 *
SiteWE4         1.701e+00  1.149e+00   1.481  0.13917
SiteWE6        -2.172e+00  8.378e-01  -2.592  0.00981 **
SiteWE8        -1.353e+00  8.937e-01  -1.514  0.13075
SiteWE9        -2.103e+00  8.465e-01  -2.484  0.01331 *
wspd_m.s_lag5   7.019e-01  2.656e-01   2.643  0.00847 **
atmp_degC_lag7 -3.444e-01  6.125e-02  -5.623 3.08e-08 ***
discharge_lag0  2.930e-04  5.304e-05   5.523 5.27e-08 ***
pres_hPa_lag1   1.126e-01  3.982e-02   2.827  0.00487 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.610788)

    Null deviance: 358.72  on 531  degrees of freedom
Residual deviance: 262.79  on 519  degrees of freedom
AIC: -1210.6

Number of Fisher Scoring iterations: 6
```

Figure A9 Model 3 for dissolved microcystin

```
glm(formula = dissolve_microcystin ~ Site + wspd_m.s_lag5 + atmp_degC_lag7 +
    total_kjeldahl_nitrogen_lag1 + nitrate_nitrate_lag7 + sulfate_lag7 +
    total_phosphorus_lag6, family = "Gamma", data = data_nona)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4816  -0.6171  -0.2455   0.2481   2.2853

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  -2.42894    2.58039  -0.941 0.346988
SiteWE13                      2.25764    1.14000   1.980 0.048190 *
SiteWE15                      2.76356    2.05504   1.345 0.179289
SiteWE16                      0.67028    1.10554   0.606 0.544588
SiteWE2                      -1.68852    0.78446  -2.152 0.031822 *
SiteWE4                       1.58720    1.05829   1.500 0.134284
SiteWE6                      -1.98439    0.76354  -2.599 0.009619 **
SiteWE8                      -1.25536    0.81639  -1.538 0.124731
SiteWE9                      -1.89986    0.77110  -2.464 0.014071 *
wspd_m.s_lag5                 0.91558    0.23419   3.910 0.000105 ***
atmp_degC_lag7               -0.29246    0.06105  -4.791 2.18e-06 ***
total_kjeldahl_nitrogen_lag1  3.17638    0.63819   4.977 8.80e-07 ***
nitrate_nitrate_lag7          0.64089    0.11844   5.411 9.60e-08 ***
sulfate_lag7                  0.09833    0.01495   6.578 1.17e-10 ***
total_phosphorus_lag6        18.94035    4.21166   4.497 8.51e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.541488)

    Null deviance: 358.72  on 531  degrees of freedom
Residual deviance: 232.31  on 517  degrees of freedom
AIC: -1277.2

Number of Fisher Scoring iterations: 6
```

Figure A10 Model 4 for dissolved microcystin

```
Call:
glm(formula = dissolve_microcystin ~ Site + gst_m.s_lag5 + atmp_degC_lag7 +
    nitrate_nitrate_lag7 + sulfate_lag7 + total_kjeldahl_nitrogen_lag1 +
    total_phosphorus_lag6, family = "Gamma", data = data_nona)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4502  -0.6170  -0.2660   0.2644   2.2668

Coefficients:
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                 -0.81059    2.60192  -0.312   0.7555
SiteWE13                     2.23353    1.16129   1.923   0.0550 .
SiteWE15                     2.81144    2.08066   1.351   0.1772
SiteWE16                     0.59009    1.12994   0.522   0.6017
SiteWE2                     -1.69795    0.79974  -2.123   0.0342 *
SiteWE4                      1.57100    1.07817   1.457   0.1457
SiteWE6                     -1.99365    0.77871  -2.560   0.0107 *
SiteWE8                     -1.25840    0.83223  -1.512   0.1311
SiteWE9                     -1.90749    0.78640  -2.426   0.0156 *
gst_m.s_lag5                 0.41703    0.16043   2.599   0.0096 **
atmp_degC_lag7              -0.29678    0.06279  -4.726 2.95e-06 ***
nitrate_nitrate_lag7         0.63530    0.12039   5.277 1.93e-07 ***
sulfate_lag7                 0.09459    0.01469   6.439 2.76e-10 ***
total_kjeldahl_nitrogen_lag1 3.08864    0.64943   4.756 2.57e-06 ***
total_phosphorus_lag6       16.58577    4.16523   3.982 7.81e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Gamma family taken to be 0.5594099)

    Null deviance: 358.72  on 531  degrees of freedom
Residual deviance: 237.19  on 517  degrees of freedom
AIC: -1265.3
```

Figure A11 Model 5 for dissolved microcystin

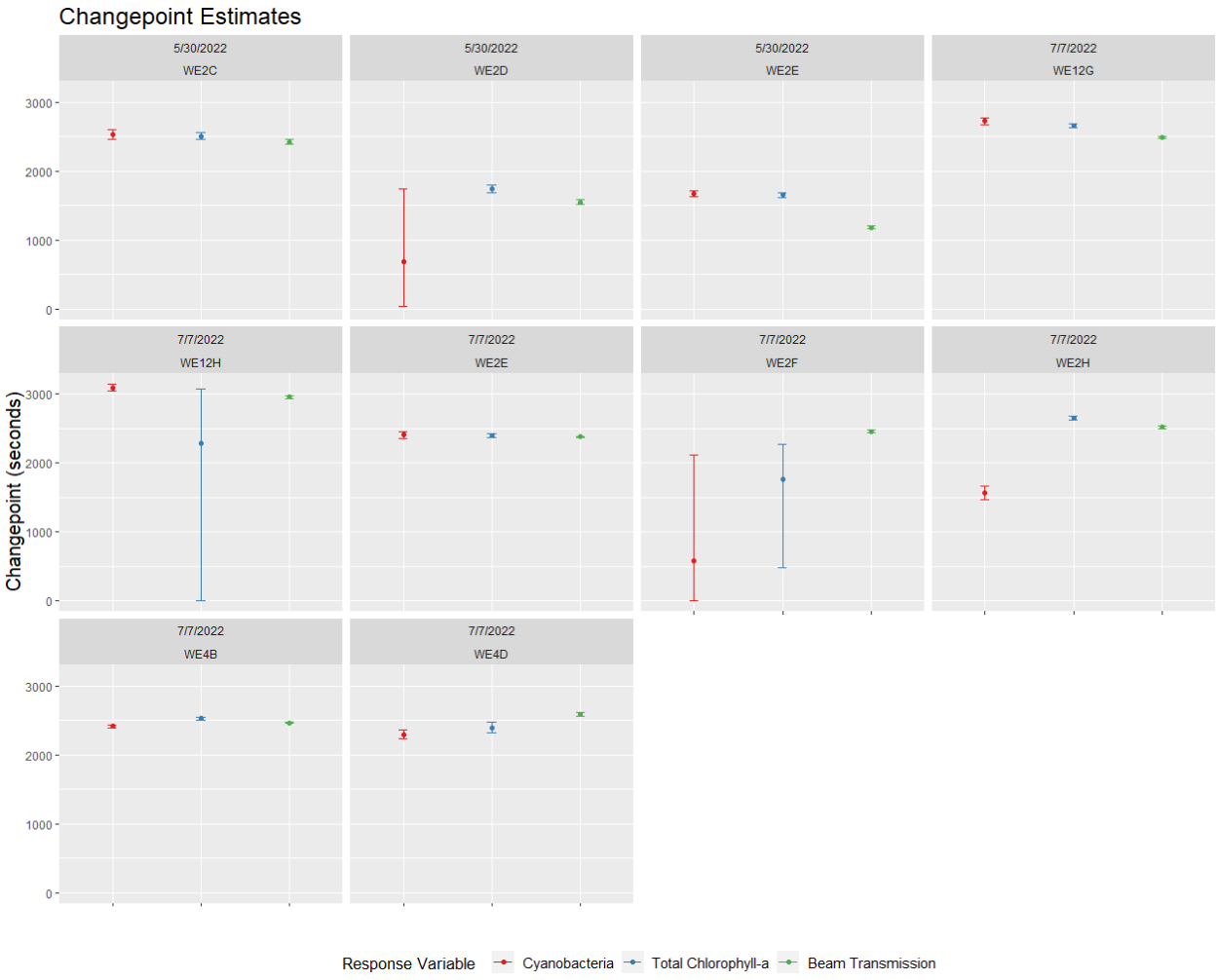


Figure A12. Changepoint estimates for cyanobacteria, total chlorophyll-a, and beam transmission in sediment cores across the five sampling sites collected in May and July 2022

References

Anderson, D. M., Fensin, E., Gobler, C. J., Hoeglund, A. E., Hubbard, K. A., Kulis, D. M., Landsberg, J. H., Lefebvre, K. A., Provoost, P., Richlen, M. L., Smith, J. L., Solow, A. R., & Trainer, V. L. (2021a). Marine harmful algal blooms (HABs) in the United States: History, current status and future trends. *Harmful Algae*, *102*, 101975. https://doi.org/10.1016/J.HAL.2021.101975

Anderson, H. S., Johengen, T. H., Miller, R., & Godwin, C. M. (2021b). Accelerated sediment phosphorus release in Lake Erie's central basin during seasonal anoxia. *Limnology and Oceanography*, *66*(9), 3582–3595. https://doi.org/10.1002/lno.11900

Benayache, N. Y., Nguyen-Quang, T., Hushchyna, K., McLellan, K., Afri-Mehennaoui, F. Z., & Bouaïcha, N. (2019). An overview of cyanobacteria harmful algal bloom (CyanoHAB) issues in freshwater ecosystems. Limnology-Some New Aspects of Inland Water Ecology, 1-25.

Bridgeman, T. B., Chaffin, J. D., Kane, D. D., Conroy, J. D., Panek, S. E., & Armenio, P. M. (2012). From River to Lake: Phosphorus partitioning and algal community compositional changes in Western Lake Erie. *Journal of Great Lakes Research*, *38*(1), 90–97. https://doi.org/10.1016/J.JGLR.2011.09.010

Bullerjahn, G. S., McKay, R. M., Davis, T. W., Baker, D. B., Boyer, G. L., D'Anglada, L. V., Doucette, G. J., Ho, J. C., Irwin, E. G., Kling, C. L., Kudela, R. M., Kurmayer, R., Michalak, A. M., Ortiz, J. D., Otten, T. G., Paerl, H. W., Qin, B., Sohngen, B. L., Stumpf, R. P., … Wilhelm, S. W. (2016). Global Solutions to regional problems: Collecting global expertise to address the problem of harmful cyanobacterial blooms. A lake erie case study. *Harmful Algae*, *54*, 223–238. https://doi.org/10.1016/j.hal.2016.01.003

Calomeni, A., McQueen, A., Kinley-Baird, C., & Clyde, G. (2022). Identification and preventative treatment of overwintering cyanobacteria in sediments : A literature review. https://doi.org/10.21079/11681/45063

Carmichael, W. W. (1992). Cyanobacteria secondary metabolites-the cyanotoxins. *Journal of Applied Bacteriology*, *72*(6), 445–459. https://doi.org/10.1111/j.1365-2672.1992.tb01858.x

Carmichael, W. W. (1994). The toxins of cyanobacteria. *Scientific American*, *270*(1), 78–86. https://doi.org/10.1038/scientificamerican0194-78

Carmichael, W. W., & Boyer, G. L. (2016). Health impacts from cyanobacteria harmful algae blooms: Implications for the north american great lakes. *Harmful Algae*, *54*, 194–212. https://doi.org/10.1016/j.hal.2016.02.002

Chen, C., H. Liu, and R. C. Beardsley. (2003). An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries, *Journal of Atmospheric and Oceanic Technology*, 20: 159-86.

Cooperative Institute for Great Lakes Research, University of Michigan; NOAA Great Lakes Environmental Research Laboratory (2019). Physical, chemical, and biological water quality monitoring data to support detection of Harmful Algal Blooms (HABs) in western Lake Erie, collected by the Great Lakes Environmental Research Laboratory and the Cooperative Institute for Great Lakes Research since 2012. NOAA National Centers for Environmental Information. Dataset. https://doi.org/10.25921/11da-3x54

Davis, T. W., Watson, S. B., Rozmarynowycz, M. J., Ciborowski, J. J., McKay, R. M., & Bullerjahn, G. S. (2014). Phylogenies of microcystin-producing cyanobacteria in the lower Laurentian Great Lakes suggest extensive genetic connectivity. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0106093

Dusini, D. S., Foster, D. L., Shore, J. A., & Merry, C. (2009). The effect of Lake Erie water level variations on sediment resuspension. *Journal of Great Lakes Research*, *35*(1), 1–12. https://doi.org/10.1016/j.jglr.2008.04.001

Harke, M. J., Steffen, M. M., Gobler, C. J., Otten, T. G., Wilhelm, S. W., Wood, S. A., & Paerl, H. W. (2016). A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, Microcystis spp. In *Harmful Algae* (Vol. 54, pp. 4–20). Elsevier B.V. https://doi.org/10.1016/j.hal.2015.12.007

Hunter, T. S., & Croley, T. E. (1993). *Great lakes monthly hydrologic data, NOAA Data report GLERL*. Technical report, National Technical Information Service, Springfield, Virginia.

Kalnejais, L. H., Martin, W. R., Signell, R. P., & Bothner, M. H. (2006). *The Role of Sediment Resuspension in the Remobilization of Particulate-Phase Metals from Coastal Sediments Supporting Information*.

Kang, S. W., Sheng, Y. P., & Lick, W. (1982). Wave action and bottom shear stresses in Lake Erie. *Journal of Great Lakes Research*, *8*(3), 482-494.

Kitchens, C. M., Johengen, T. H., & Davis, T. W. (2018). Establishing spatial and temporal patterns in Microcystis sediment seed stock viability and their relationship to subsequent bloom development in western Lake Erie. *PLOS ONE*, *13*(11). https://doi.org/10.1371/journal.pone.0206821

Kutovaya, O. A., McKay, R. M., Beall, B. F. N., Wilhelm, S. W., Kane, D. D., Chaffin, J. D., Bridgeman, T. B., & Bullerjahn, G. S. (2012). Evidence against fluvial seeding of recurrent toxic blooms of Microcystis spp. in Lake Erie's Western Basin. *Harmful Algae*, *15*, 71–77. https://doi.org/10.1016/j.hal.2011.11.007

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgeman, T. B., Chaffin, J. D., Cho, K., Confesor, R., Daloglu, I., DePinto, J. V., Evans, M. A., Fahnenstiel, G. L., He, L., Ho, J. C., Jenkins, L., Johengen, T. H., Kuo, K. C., LaPorte, E., … Zagorski, M. A. (2013). Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6448–6452. https://doi.org/10.1073/pnas.1216006110

Misson, B., & Latour, D. (2012). Influence of light, sediment mixing, temperature and duration of the benthic life phase on the benthic recruitment of Microcystis. Journal of Plankton Research, 34(2), 113-119.

National Data Buoy Center (2022) Station THRO1 - Toledo, OH, NDBC. https://www.ndbc.noaa.gov/station_history.php?station=thro1

Paerl, H. W., Gardner, W. S., Havens, K. E., Joyner, A. R., McCarthy, M. J., Newell, S. E., Qin, B., & Scott, J. T. (2016). Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients. *Harmful Algae*, *54*, 213–222. https://doi.org/10.1016/J.HAL.2015.09.009

Rinta-Kanto, J. M., Saxton, M. A., DeBruyn, J. M., Smith, J. L., Marvin, C. H., Krieger, K. A., Sayler, G. S., Boyer, G. L., & Wilhelm, S. W. (2009). The diversity and distribution of toxigenic Microcystis spp. in present day and archived pelagic and sediment samples from Lake Erie. *Harmful Algae*, *8*(3), 385–394. https://doi.org/10.1016/j.hal.2008.08.026

Sellner, K. G., Doucette, G. J., & Kirkpatrick, G. J. (2003). Harmful algal blooms: Causes, impacts and detection. In *Journal of Industrial Microbiology and Biotechnology* (Vol. 30, Issue 7, pp. 383–406). https://doi.org/10.1007/s10295-003-0074-9

Soulsby, R. L., & Whitehouse, R. J. S. (n.d.). *Threshold of Sediment Motion in Coastal Environments*.

Steffen, M. M., Belisle, B. S., Watson, S. B., Boyer, G. L., & Wilhelm, S. W. (2014). Status, causes and controls of cyanobacterial blooms in Lake Erie. *Journal of Great Lakes Research*, *40*(2), 215–225. https://doi.org/10.1016/j.jglr.2013.12.012

Steffen, M. M., Davis, T. W., McKay, R. M. L., Bullerjahn, G. S., Krausfeldt, L. E., Stough, J. M. A., Neitzey, M. L., Gilbert, N. E., Boyer, G. L., Johengen, T. H., Gossiaux, D. C., Burtner, A. M., Palladino, D., Rowe, M. D., Dick, G. J., Meyer, K. A., Levy, S., Boone, B. E., Stumpf, R. P., … Wilhelm, S. W. (2017). Ecophysiological Examination of the Lake Erie Microcystis Bloom in 2014: Linkages between Biology and the Water Supply Shutdown of Toledo, OH. *Environmental Science and Technology*, *51*(12), 6745–6755. https://doi.org/10.1021/acs.est.7b00856

Stow, C. A., Cha, Y., Johnson, L. T., Confesor, R., & Richards, R. P. (2015). Long-term and seasonal trend decomposition of maumee river nutrient inputs to western lake erie. *Environmental Science and Technology*, *49*(6), 3392–3400. https://doi.org/10.1021/es5062648

Stumpf, R. P., Wynne, T. T., Baker, D. B., & Fahnenstiel, G. L. (2012). Interannual variability of cyanobacterial blooms in Lake Erie. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0042444

The WAVEWATCH III® Development Group (WW3DG), 2019: User manual and system documentation of WAVEWATCH III® version 6.07. Tech. Note 333, NOAA/NWS/NCEP/MMAB, College Park, MD, USA, 326 pp. + Appendices.

Tillett, D., Dittmann, E., Erhard, M., von Döhren, H., Börner, T., & Neilan, B. A. (2000). Structural organization of microcystin biosynthesis in Microcystis aeruginosa PCC7806:

An integrated peptide–polyketide synthetase system. *Chemistry & Biology*, *7*(10), 753–764. https://doi.org/10.1016/s1074-5521(00)00021-1

Tolhurst, T. J., Black, K. S., Paterson, D. M., Mitchener, H. J., Termaat, G. R., & Shayler, S. A. (2000). A comparison and measurement standardisation of four in situ devices for determining the erosion shear stress of intertidal sediments. Continental Shelf Research, 20(10-11), 1397-1418.

United States Geological Survey. Maumee River at Waterville Oh. USGS Water Data for the Nation. https://waterdata.usgs.gov/monitoring-location/04193500/

US Department of Commerce, NOAA. *NOAA-GLERL Western Lake Erie Sampling Data*. Home: NOAA Great Lakes Environmental Research Laboratory - Ann Arbor, MI, USA. Retrieved April 2023, from https://www.glerl.noaa.gov/res/HABs_and_Hypoxia/wle-weekly-current/

US Department of Commerce, N.O.A.A. (2019) Harmful algal blooms (red tide), (Red Tide). Available at: https://oceanservice.noaa.gov/hazards/hab/

Verspagen, J. M. H., Snelder, E. O., Visser, P. M., Huisman, J., Mur, L. R., & Ibelings, B. W. (2004). Recruitment of benthic microcystis (Cyanophyceae) to the water column: Internal buoyancy changes or resuspension?1. *Journal of Phycology*, *40*(2), 260–270. https://doi.org/10.1111/j.1529-8817.2004.03174.x

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., & Wilhelm, S. W. (2016). The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia. *Harmful Algae*, *56*, 44–66. https://doi.org/10.1016/j.hal.2016.04.010