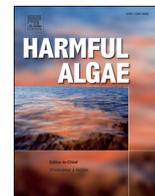


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Harmful Algae

journal homepage: www.elsevier.com/locate/hal

Phytoplankton assemblage structure, drivers, and thresholds with a focus on harmful algal bloom ecology in the Lake Okeechobee system, Florida, USA

Viviana Mazzei ^{a,*} , Kristy Lee Sullivan ^{b,1} , Keith Loftin ^c ^a U.S. Geological Survey, Orlando, FL, 32826, USA^b Contractor to the U.S. Geological Survey, Orlando, FL, 32826, USA^c U.S. Geological Survey, Lawrence, KS, 66049, USA

ARTICLE INFO

Keywords:

Phytoplankton
Harmful algal blooms
Cyanobacteria
Community assembly
Indicator species
Ecological thresholds
Lake Okeechobee

ABSTRACT

Untangling the complexities of harmful algal bloom (HAB) dynamics is an ongoing effort that requires a fundamental understanding of spatiotemporal phytoplankton patterns and the environmental filters through which assemblages are structured. To this aim, monthly field surveys were conducted from 2019 to 2021 at 21 sites in Lake Okeechobee, Florida – a large, shallow, eutrophic, and heavily managed lake with coastal connectivity that experiences intense and recurrent HABs. Phytoplankton assemblages were strongly spatially structured forming 7 distinct lake zones with significant dissimilarity in composition and total abundance. While successional patterns were not apparent across seasons or wet/dry periods, total phytoplankton abundance was significantly greater towards the end of the wet season. Distance-based linear models using 16 abiotic variables were used to identify significant explanatory variables of spatial and temporal patterns. The spatial model explained 93 % of the variability suggesting deterministic processes largely control spatial patterns. The temporal model explained only 48 % of the temporal variability suggesting stochasticity in lake-wide shifts in assemblages over time. However, the strong spatial structuring of assemblages may preclude lake-wide succession patterns. Total algal abundance metrics were inversely related to nitrate, orthophosphate, and total alkalinity, the strongest explanatory variables of assemblage patterns, suggesting a lag between peak resources and peak abundance as phytoplankton cycle “boom-to-bust” phases. Consistent with this inverse relationship, Threshold Indicator Taxa Analysis returned almost exclusively negative responder indicator taxa for all three explanatory variable gradients. The assemblage-level threshold defined the gradient boundary between boom- and bust-associated indicator taxa. These data contribute novel information about HABs ecology pertinent to management strategies.

1. Introduction

Understanding the processes that structure algal assemblages in space and time is foundational to unraveling the complexities of harmful algal bloom (HAB) dynamics. Community ecology is largely concerned with describing the spatial and temporal patterns of species distributions and identifying the drivers that organize assemblages into observed patterns (Leibold et al., 2004; Austin, 2007; Vellend, 2010; Weiher et al., 2011; Rojo, 2021). Although generalized patterns and drivers of algal community assembly exist, such as those described by Margalef’s mandala and the C-S-R model (Smayda and Reynolds, 2001; Reynolds

2006; Wyatt, 2014; Glibert, 2016), system-specific or local processes, especially in highly managed aquatic habitats, must be identified to truly understand HAB dynamics in the system of interest.

Ecological community assembly theory includes both deterministic and stochastic components encompassed by four main mechanisms: selection, drift, dispersal, and speciation (Vellend, 2010). Environmental selection by abiotic and biotic factors, commonly called environmental filtering or species sorting, is a deterministic mechanism that involves the sorting of species from the regional species pool through a series of environmental filters that determine the structure of the local community. Ecological drift, dispersal, and speciation are considered

* Corresponding author.

E-mail address: vmazzei@usgs.gov (V. Mazzei).

¹ Present Address: Colorado Springs Utilities, Laboratory Services Section, Colorado Springs, CO 80903, USA

<https://doi.org/10.1016/j.hal.2024.102744>

Received 8 July 2024; Received in revised form 28 October 2024; Accepted 4 November 2024

Available online 9 November 2024

1568-9883/Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

stochastic mechanisms that occur because of the inherent randomness or probability in biological and environmental processes. The relative strength of these mechanisms in structuring ecological communities depends on the spatial and temporal scale under investigation as well as on the organism being studied (Chase and Myers, 2011).

Most HAB-forming species in freshwaters are microalgae (Paerl et al., 2001) and can therefore be expected to follow microbial community assembly rules (Nemergut et al., 2013). Lourens G. M. Baas Becking (1934; Canfield, 2015) famously said of microbial biogeography, “everything is everywhere, but the environment selects,” asserting that because microbes are not limited by dispersal barriers like many macroorganisms, they can hypothetically disperse globally, but the environment they randomly colonize determines survival (Van der Gucht et al., 2007; Nemergut et al., 2013). Although this paradigm has been challenged by modern microbial biogeography studies that have found microbial endemism (O’Malley, 2008; Ribeiro et al., 2018; Spatharis et al., 2019), local deterministic processes are thought to be a stronger driver of microbial community assembly than stochastic processes, particularly at smaller spatial scales (Van der Gucht et al., 2007; Langenheder and Székely, 2011; Nemergut et al., 2013; Zhou and Ning, 2017). Species sorting and environmental filtering are useful concepts for understanding how phytoplankton assemblages are structured along environmental gradients at the landscape scale, and identifying ecological thresholds or change points where individual species or whole assemblages undergo significant shifts in composition and abundance (Baker and King, 2010; King and Baker, 2010). Knowledge of species-level and community-level thresholds to a particular environmental driver is valuable for management and prediction of ecosystem health and can be useful for establishing total maximum daily loads and numerical nutrient criteria aimed at mitigating HABs (Guntenspergen and Gross, 2014; King and Baker, 2014; Nichols et al., 2014). Likewise, indicator species, organisms having strong sensitivity to specific environmental drivers that provide early warning signals of impending ecosystem changes, can also be a powerful tool in HABs management (Dufrene and Legendre, 1997; Siddig et al., 2016; McQuatters-Gollop et al., 2017). Algae are especially reliable bioindicators and are routinely used in ecosystem assessments (Stevenson, 2014; McQuatters-Gollop et al., 2017).

The Lake Okeechobee system, located within the Greater Everglades Watershed in south-central Florida, presents a model ecosystem to study the patterns and processes of phytoplankton assembly in large, shallow subtropical lakes with coastal connectivity. Novel anthropogenic environmental gradients – environmental gradients that have developed as a result of human activities such as watershed development, managed hydrologic processes, and climate change – create new spatial and temporal heterogeneity that likely alter natural phytoplankton assemblage structure and play an important role in HAB dynamics. The surrounding agricultural and pasture lands along with urban development have contributed to the eutrophication of the Lake Okeechobee system which is in part responsible for the increasingly common and more intense HABs in the system (Zhang and Welch, 2018). Despite considerable research into HABs in the Lake Okeechobee system, there remains large uncertainty about the immediate processes that cause steady-state phytoplankton assemblages to shift towards dominance by one or a few cyanobacteria species that proliferate into a HAB (Brooks et al., 2016; Haakonsson et al., 2017; Ho et al., 2019). Nutrient enrichment is undeniably linked to HABs, not only in this system, but across the globe (Glibert et al., 2005; Glibert and Buford, 2017; Glibert, 2020). In large, shallow lakes like Lake Okeechobee, physical processes such as wind and hydrology can also play an important role in HAB dynamics (Havens et al., 1994). However, a complex interplay between nutrients, lake physics, and other abiotic and biotic factors determine the occurrence, extent, duration, and composition of a HAB (Paerl et al., 2001), and has yet to be understood in a manner supporting actionable management strategies.

In this paper we (1) describe the spatial and temporal patterns of

phytoplankton assemblages based on genus-level (or lowest possible) taxonomic classifications in Lake Okeechobee and the Caloosahatchee and St. Lucie Rivers (henceforth collectively referred to as the Lake Okeechobee system); (2) identify abiotic variables explaining the greatest proportion of the total spatial and temporal phytoplankton assemblage variability using multidimensional, multivariate regression models; and (3) describe the distribution of indicator taxa along gradients of the variables shown to be highly explanatory of assemblage patterns and define thresholds for those variables using Threshold Indicator Taxa Analysis. We hypothesized that (1) the phytoplankton assemblages of the Lake Okeechobee system are spatially and temporally structured exhibiting geographically distinct assemblages and predictable seasonal succession; (2) deterministic processes such as spatial differences in water quality and seasonal fluctuations in meteorological conditions significantly structure phytoplankton assemblages; and (3) phytoplankton assemblages exhibit clear thresholds along the prevailing environmental gradients influencing phytoplankton assembly and HABs with indicator species on either side of the assemblage-level thresholds representing the different ecosystem states.

2. Methods

2.1. Study area

Lake Okeechobee (27° N latitude and 81° W longitude), meaning “big water” in the Seminole language, is the largest lake in the southeastern United States occupying 1800 km² (Zhang and Welch, 2018). It is located within the Greater Everglades Watershed in south-central Florida, USA and is connected to the eastern and western coasts by the St. Lucie River and the Caloosahatchee River which flow into the Atlantic Ocean and the Gulf of Mexico, respectively (Fig. 1). The lake is characterized as eutrophic, turbid, and shallow with an average depth of 2.7 m (Zhang and Welch, 2018). The Lake Okeechobee watershed (LOW) covers the northern section of the larger Greater Everglades Watershed which extends south through Everglades National Park to Florida Bay. Water inflows to Lake Okeechobee are primarily from the Upper Kissimmee, Lower Kissimmee, Taylor Creek/Nubbin Slough, Lake Istokpoga, Indian Prairie, and Fisheating Creek subwatersheds of the LOW that drain into the northern portion of the Lake. The primary land use types in these subwatersheds north of the lake are pastures for beef cattle grazing and citrus groves, while sugarcane production is the principal land use type south of the lake (Zhang and Welch 2018; Fig. 1A). Primary water outflows from Lake Okeechobee are to the St. Lucie River, the Caloosahatchee River, and the Everglades Agricultural Area in the South Lake Okeechobee subwatershed.

2.2. Field and laboratory methods

Surface water samples were collected monthly from March, 2019 through October, 2021 (32 months) from 17 sites in Lake Okeechobee, a site at the upper (S-77) and lower (S-79) Caloosahatchee River, and a site at the upper (S-308) and lower (S-80) St. Lucie River for a total of 21 field sites (Fig. 1). Quantitative taxonomic analysis of phytoplankton assemblages was performed using an Imaging FlowCytobot East Falmouth, MA, “<https://mclanelabs.com/imaging-flowcytobot/>” Imaging FlowCytobot - McLane Labs. Live and glutaraldehyde-preserved samples were sent to Phycotech, Inc. for algal identification and enumeration. The FlowCytobot captures images of particulates between 2 and 250 µm as they pass through a flow cell. Laser-induced fluorescence and light scattering from individual particles are measured and used to trigger targeted image acquisition allowing for the discrimination between algal cells, detritus, zooplankton, and other particulates. Algae are identified by an automated classifier which identifies algal units above 9 µm to genus level. When image resolution did not allow for genus-level identification, phytoplankton were classified at higher taxonomic or morphological groups (e.g., chlorophytes other; centric diatoms other;

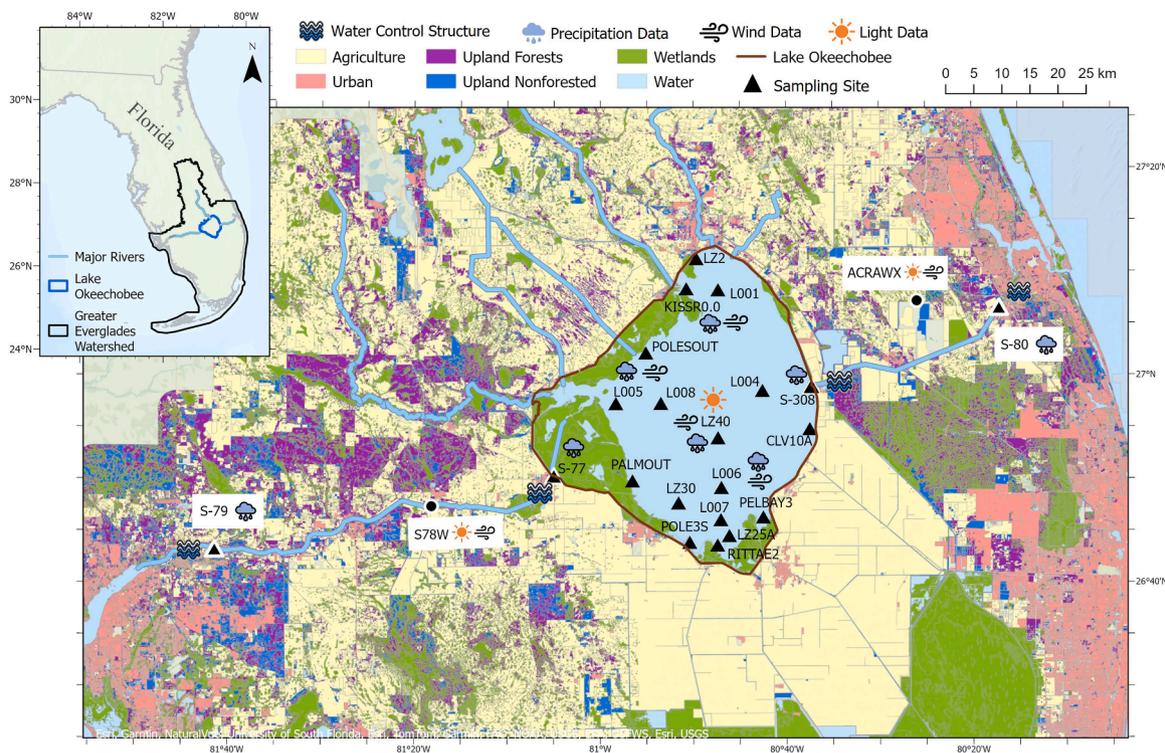


Fig. 1. Map of the Lake Okeechobee, Florida (USA) study area and surrounding land use. The locations of the twenty-one stations where phytoplankton and water quality samples were taken are shown along with the location of weather stations with icons indicating what meteorological data were available at each. The three major river connections to Lake Okeechobee are the Caloosahatchee River flowing west into the Gulf of Mexico, the St. Lucie River flowing east into the Atlantic Ocean, and the Kissimmee River flowing south into northern Lake Okeechobee near station KISSR0.0. The land use feature layer was obtained from the South Florida Water Management District's Geospatial Open Data Portal.

etc) or placed into the “unclassified” category used for all algae that were counted but could not be identified. Algal taxa below $9\ \mu\text{m}$ were counted but not identified and placed in a separate category, “taxa $< 9\ \mu\text{m}$.” The automated algal classifications were then validated by a two-person verification system. Abundance measurements for each taxon were calculated as natural units per milliliter (NU mL^{-1}), cells per milliliter (cells mL^{-1}), and biovolumes per milliliter ($\mu\text{m}^3\ \text{mL}^{-1}$). Biovolumes were calculated manually using cell measurements taken from the images and applying the appropriate shape equation for each taxon.

Abiotic data for each site collected concomitantly with our samples were downloaded from the South Florida Water Management District (SFWMD) environmental database (DBHYDRO Browser (sfwmd.gov)) and the U.S. Geological Survey (USGS) National Water Information System (NWIS) database (USGS Water Data for the Nation). Abiotic variables selected for statistical analyses included total ammonia as nitrogen ($\text{NH}_3 + \text{NH}_4 - \text{N}$), nitrate + nitrite as nitrogen ($\text{NO}_3 + \text{NO}_2 - \text{N}$), total nitrogen (TN), orthophosphate (ortho-P), total phosphorus (TP), total alkalinity (TA), chlorophyll-*a* (Chl-*a*), depth, dissolved oxygen (DO), pH, turbidity, specific conductivity (SpCond), water temperature (temp), precipitation (precip), total solar radiation (TSR), wind direction (WNDD), and wind speed (WNDS). Precipitation data were available at only 8 of our sampling sites (Fig. 1) and were extrapolated to the nearest sites where data were not available. TSR data within the lake were only available at one of our sampling sites, the LZ40 weather station in the center of Lake Okeechobee, and was applied to all sites within the lake and the upper Caloosahatchee River (S-77) and upper St. Lucie (S-308) sites (Fig. 1). TSR data for the lower Caloosahatchee site (S-79) were obtained from the S78 weather station on the Caloosahatchee River at Ortona ($\sim 39\ \text{km}$ east 78.9° of S-79) and TSR data for the lower St. Lucie site (S-80) were obtained from the Allapattah Tract weather station (ACRAWX, $\sim 14.5\ \text{km}$ west 93.9° of S-80). Wind data were available from S78 W and ACRAWX which was applied to sites

S79 and S80, respectively, as well as from four sites within the lake (Fig. 1) which were extrapolated to the nearest site where data were not available. Satellite imagery of Lake Okeechobee from the Harmful Algal Bloom Forecasting Branch of the National Centers for Coastal Ocean Science (NCCOS; Harmful Algal Bloom Monitoring System - NCCOS Coastal Science Website (noaa.gov)) were downloaded for each of the 32 months sampled for comparison to spatial and temporal distribution of algal biomass estimated through remote sensing (Fig. S1). The images were derived from Copernicus Sentinel-3 satellite data from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) and were processed by NCCOS.

2.3. Statistical analysis

Multidimensional, nonparametric statistical tools in PRIMER-e v7 software (Clarke and Gorley, 2015; Anderson et al., 2008) were used to investigate spatial and temporal patterns of phytoplankton assemblage structure and construct models describing the abiotic drivers of those patterns. Phytoplankton abundance (cells mL^{-1}) data from all 639 samples collected over 32 months from 21 sampling locations were averaged across months for each site (integrating temporal variability) and across sites for each month (MMYY; integrating spatial variability) to separately investigate spatial and temporal patterns. Abundance data were fourth-root transformed and separate Bray-Curtis resemblance matrices were constructed from the spatial and temporal datasets.

Hierarchical unconstrained binary divisive clustering (UNCLUSTER) and a similarity profile test (SIMPROF) were run on the spatial resemblance matrix to test for multivariate structure among sampling sites and identify clusters of sites with significantly ($P < 0.05$) distinct phytoplankton assemblages. A three-way crossed analysis of similarity (ANOSIM) was used to test for significant ($P < 0.05$) assemblage dissimilarity within three temporal factors: the two precipitation

seasons (wet season: May – October and dry season: November – April), the four meteorological seasons (spring: March, April, May; summer: June, July, August; fall: September, October, November; winter: December, January, February), and the years sampled (2019, 2020, and 2021). Additionally, a one-way and a three-way analysis of variance (ANOVA, RStudio 2023.09.0, `anova_test` {rstatix}) was used to test for significant differences ($p < 0.05$) in three univariate algal abundance metrics - total phytoplankton cell concentration, total phytoplankton biovolume, and Chl-*a* among spatial (lake zones) and temporal (precipitation season, meteorological season, year) factors, respectively. Bonferroni adjusted post hoc pairwise tests were performed on all ANOVAs having more than three factors to identify differences between each pair of factors.

Distance-based linear models (DISTLM) and redundancy analysis (dbrDA) were used to explore the spatial and temporal relationship between the multivariate taxa abundance data and 16 abiotic explanatory variables. Spatial tests were performed on time-averaged sample data for each site ($n = 21$) while temporal tests were performed on site-averaged sample data for each sampling month ($n = 32$). Marginal tests identified variables individually explaining a significant ($P < 0.05$) proportion of the total phytoplankton assemblage variation while sequential conditional tests selected the combination of variables that cumulatively explained the greatest proportion of the total variation. In addition to the multivariate assemblage models, DISTLMs were also run on Chl-*a* to identify the best explanatory variables describing the spatial and temporal patterns of this commonly used algal abundance metric. The most parsimonious models were determined using a forward-selection procedure based on adjusted R^2 section criterion. Histograms were examined to determine appropriate statistical transformations for individual explanatory variables. Draftsman plots and Pearson correlation matrices were examined to search for spatial and temporal collinearity between explanatory variables. The strength of the Pearson correlation coefficient (r) was interpreted as follows: $r < 0.25 =$ no relationship, $0.25 < r < 0.5 =$ weak relationship, $0.5 < r < 0.75 =$ moderate relationship, and $r > 0.75 =$ strong relationship. Draftsman plots and Pearson correlation matrices of spatial and temporal relationships between Chl-*a*, total cell concentration, and biovolume abundance metrics and the 16 abiotic variables were also examined to identify relationships between the three total abundance metrics and the explanatory variables. Plots of all untransformed abiotic variables averaged across sites and sampling dates were used to study site-specific spatial trends and monthly temporal trends of the chemical and physical conditions during the study.

Threshold Indicator Taxa Analysis (TITAN, Baker et al. 2020; Baker et al., 2023) was performed in R (TITAN2 package, R Core Team, 2023) along the gradients of the abiotic variables identified by the DISTLM as significant explanatory variables of phytoplankton assemblage structure to detect species-level and assemblage-level thresholds to each explanatory variable. TITAN calculates the environmental change point (CP), indicator value (IndVal) and its z-score (z), and the purity and reliability of each taxon classifying them as either negative (z^-) or positive (z^+) taxa. Negative indicator taxa (z^-) are those whose frequency and abundance decline moving up the environmental gradient while positive indicator taxa (z^+) increase moving up the gradient. Because raw IndVal scores will favor the most frequent and abundant taxa, IndVal z-scores are used to standardize the distribution thereby emphasizing the contributions of taxa with low occurrence frequencies but high sensitivity to the gradient. Purity is defined as the proportion of bootstrap replicates with CP directions (decreasing z^- or increasing z^+) that agree with the observed response direction while reliability is the proportion of bootstrap CPs with significant IndVal scores ($p < 0.05$) so that taxa with repeatable and consistently large IndVal scores are deemed reliable. Taxa with frequencies ≤ 3 or a frequency of 100 must be excluded from the analysis.

3. Results

3.1. Objective 1 – Identify the spatial and temporal patterns of phytoplankton assemblage structure and univariate abundance metrics

The phytoplankton assemblages of the Lake Okeechobee system were strongly spatially structured. Seven lake zones with significant phytoplankton assemblage dissimilarity were identified by the UNCTREE and SIMPROF tests (Fig. 2A, B). The two lower riverine sites S-79 and S-80, respectively named West Coast and East Coast, had distinct phytoplankton assemblages from each other and all other sites. A cluster named “Northwest (NW) Lake” zone containing the 5 sites closest to the main water inflow areas for Lake Okeechobee was identified. The sites included LZ2, KISSR0.0, and L001 (nearest to inflows from the Kissimmee River, Taylor Creek, Mosquito Creek, and Nubbin Slough), plus POLESOUT located downstream of the C-40 canal inflow and L005 located downstream of water inflows from the C-41 canal and Fisheating Creek. The central portion of Lake Okeechobee was split between a “MidWest Lake” zone containing L008, PALMOUT, and the S-77 Caloosahatchee River outflow site and a “MidEast Lake” zone containing S-308, the St. Lucie outflow site, and CLV10A, L004, and LZ40. The southern portion of Lake Okeechobee is dominated by water outflows directed south through the S354 and S351 spillways into the Everglades Agricultural Area, Water Conservation Areas, and eventually Everglades National Park. Two clusters were identified in this area – “South (S) Lake” containing L006, LZ30, L007, LZ25A, and PELBAY3 and “S Lake 2” comprised of POLES3 and RITTAE2. A gradient of increasing phytoplankton abundance from south to north was observed (Fig. 2C). The NW Lake zone had significantly greater total cell concentrations, total biovolume, and Chl-*a* concentrations compared to all other lake zones (post-hoc pairwise Bonferroni adjusted $p < 0.05$). The MidWest and MidEast Lake zones had significantly higher cell concentrations, biovolumes, and Chl-*a* compared to the South Lake and South lake 2 zones. The East and West Coast zones did not differ significantly from each other in any of the algal abundance metrics and were only significantly different from the NW Lake zone.

Temporal patterns of assemblage structure were more variable and less defined than spatial patterns (Fig. 3). Significant compositional dissimilarity was present across years 2019, 2020, and 2021 (ANOSIM $R = 0.632$, $P = 0.001$), but there were no significant differences between wet and dry seasons (ANOSIM $R = -0.4$, $P = 0.21$) or between any of the meteorological seasons (ANOSIM $R = 0.255$, $P = 0.06$). The univariate metrics of phytoplankton abundance examined – total cell concentration, biovolume, and Chl-*a* – exhibited stronger temporal trends than the multivariate assemblages. Significant differences in all three metrics occurred between years and precipitation season (ANOVA $p < 0.05$) with all metrics increasing from 2019 to 2021 and higher values recorded during the wet season. Total cell concentration exhibited significant differences between the meteorological seasons, particularly between fall when cell density was highest and spring when it was lowest (ANOVA $p < 0.05$). Mean Chl-*a* and biovolume were highest in the summer despite the non-significant p-values returned by the 3-way ANOVAs.

Chl-*a* $> 20 \mu\text{g L}^{-1}$, cell concentration $> 100,000 \text{ cells mL}^{-1}$, and supporting satellite imagery were used to infer the presence of a bloom. Monthly NOAA satellite imagery (Harmful Algal Bloom Monitoring System - NCCOS Coastal Science Website (noaa.gov)) showed mild to moderate blooms occurring in June – August 2019, June – September 2020, and May – October 2021, and a major HAB event in May 2021. The May 2021 HAB covered the entire lake, but obvious hot spots occurred in the NW Lake zone shoreline and along the eastern shoreline of Lake Okeechobee (Fig. S1). Cell density and Chl-*a* data reflected the spatial patchiness of the bloom with the greatest cell concentrations ($\sim 300,000 \text{ cells mL}^{-1}$) measured at CLV10A, located in the MidEast zone near the eastern shoreline, and LZ2, located in the NW Lake zone near the mouth of the Kissimmee River. Cell concentrations between 100,000

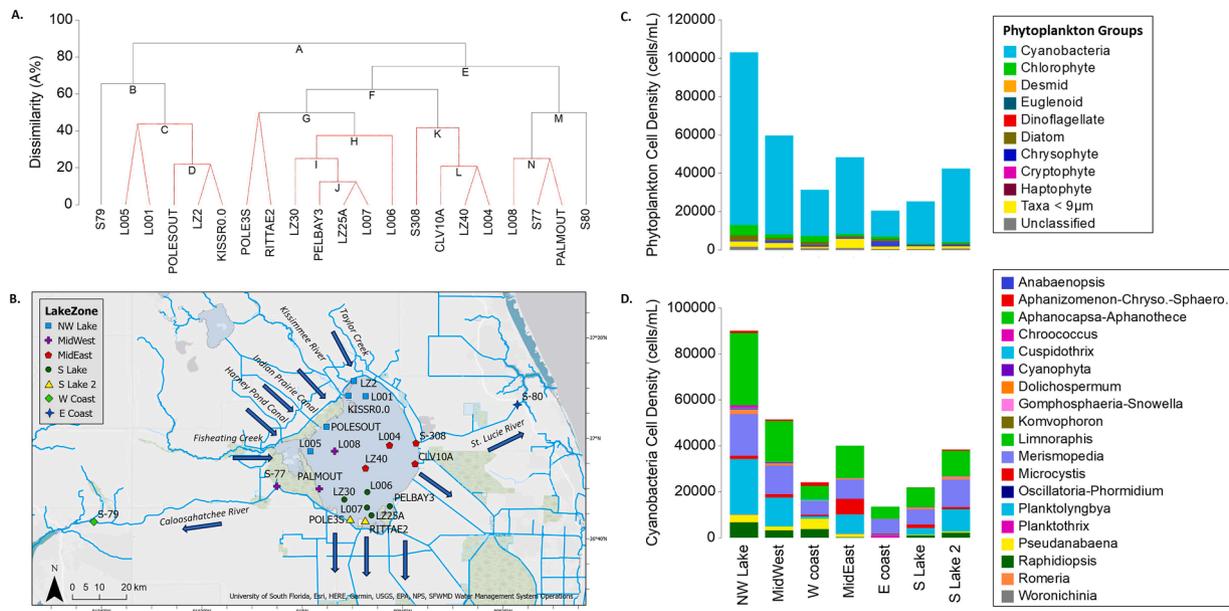


Fig. 2. Spatial patterns of phytoplankton assemblage structure. A) Dendrogram of the hierarchical unconstrained binary divisive clustering (UNCLUST) with significant groupings identified by SIMPROF displayed with red dotted lines. Lake zones identified by SIMPROF are defined as Northwest Lake (NW Lake), Mid-West Lake (MidWest), Mid-East Lake (MidEast), South Lake (S Lake), South Lake 2 (S Lake 2), West Coast (W coast), and East Coast. B) Map showing the geographic location of SIMPROF lake zones and the flow directions of primary waterways in the Lake Okeechobee watershed. C) Bar plot showing the relative contribution of different phytoplankton groups to the total phytoplankton cell concentration at each lake zone. D) Bar plot showing the relative contribution of individual cyanobacteria taxa to the total cyanobacterial cell concentration at each lake zone. Aphanizomenon-Chryso.-Sphaero. = Aphanizomenon-Chryso.-Sphaerospormopsis. The color scheme of the legend corresponds to order in the bars. Some phytoplankton groups and cyanobacteria taxa listed in legends were present in very low abundance and may not be discernable in the plots.

and 150,000 cells mL⁻¹ were recorded at three other NW Lake zone sites, L001, KISSR0.0, and POLESOUT, during the May 2021 HAB event and at the MidEast zone site S-308 adjacent to the Port Mayaca Lock and Spillway controlling water releases from Lake Okeechobee to the St. Lucie River.

A total of 74 unique taxonomic identifications were made across the 21 sites and 32-month sampling period, 19 of which were cyanobacteria. Cyanobacteria abundance dominated the phytoplankton assemblages in all lake zones and over the entire study period (Figs. 2C and 3A). *Merismopedia*, *Aphanocapsa*, and *Planktolyngbya* were the most abundant taxa across time and space, although *Planktolyngbya* abundance was noticeably lower at both lower riverine sites compared to zones within the lake (Fig. 2D and 3B). The relative abundances of *Merismopedia* and *Aphanocapsa* were evenly distributed across all spatial zones, but there was a shift in dominance from *Merismopedia* in 2019 through May 2020 to *Aphanocapsa* from June 2020 through October 2021 (Fig. 3B). *Merismopedia* and *Aphanocapsa* are both colonial genera in the order Chroococcales, family Microcystaceae and both are potential microcystin producers (Bernard et al., 2016; de J Magalhães et al., 2019), while *Planktolyngbya* is a filamentous genus within the order *Lepidodermatales* for which toxin production has not been reported. The toxin-producing taxa *Microcystis*, *Dolichospermum*, *Raphidiopsis*, and *Aphanizomenon*, and the taste and odor-producing *Pseudanabaena* were also conspicuously present but at lower concentrations and frequency (Figs. 2D and 3B). Green algae, diatoms, and taxa <9 µm were the next most abundant phytoplankton groups after cyanobacteria.

3.2. Objective 2 – Identify explanatory variables of spatiotemporal patterns of phytoplankton assemblage structure and chlorophyll-a

The DISTLM performed on time-averaged samples for each site ($n = 21$) determined that NO₃+NO₂-N individually explained the greatest proportion (44.9 %, $P = 0.001$) of the total spatial variation in the phytoplankton assemblage (Table 1). Other variables that significantly explained the spatial structure of the phytoplankton assemblage

when considered alone where TA (36.7 %, $P = 0.001$), ortho-P (33.4 %, $P = 0.001$), SpCond (23.9 %, $P = 0.003$), turbidity (18.2 %, $P = 0.03$), and temp (18.7 %, $P = 0.016$). Conditional tests revealed that after accounting for the variation explained by NO₃+NO₂-N, the addition of TP (23.7 %, $P = 0.001$), +WNDS (11.5 %, $P = 0.001$), +SpCond (4.1 %, $P = 0.003$), +turbidity (3.5 %, $P = 0.001$), +ortho-P (3.2 %, $P = 0.001$), +TA (1.3 %, 0.026), +DO (1.1 %, $P = 0.045$) cumulatively explained 93.3 % of the total spatial variation (Table 1, Table S4). The model was visualized in ordination space by distance-based redundancy analysis (dbRDA) and overlaid with vectors for the 8 significant explanatory variables selected by the conditional tests (Fig. 4). Together dbRDA axes 1 and 2 explained 83.9 % of the fitted variation and 81.4 % of the total variation.

The DISTLM performed on the site-averaged samples for each month ($n = 32$) revealed that only a small proportion of the variation in temporal assemblage structure was explained by any one variable individually with ortho-P explaining the greatest (13.8 %, $P = 0.001$), followed by depth (13.5 %, $P = 0.001$), turbidity (12.4 %, $P = 0.002$), and TA (12.1 %, $P = 0.001$, Table S1). The most parsimonious conditional model which included ortho-P, +depth (12.5 %, $P = 0.001$), +WNDD (8.7 %, $P = 0.001$), +precip (8.8 %, $P = 0.001$), and TN (4 %, $P = 0.043$) cumulatively explained 48 % of the total temporal variation in phytoplankton assemblages (Table S1, Table S4).

In addition to the models describing multivariate phytoplankton assemblage structure, DISTLM was also used to identify the best explanatory variables of the univariate phytoplankton biomass metric Chl-a. NO₃+NO₂-N (62 %, $P = 0.001$), ortho-P (51 %, $P = 0.002$), and temp (31 %, $p = 0.008$) individually explained the largest proportion of Chl-a spatial variability (Table S2). The most parsimonious model, which included NO₃+NO₂-N, +TP (17.5 %, $P = 0.002$), +temp (0.07 %, $P = 0.008$), +depth (3.2 %, $P = 0.036$), and +DO (4 %, $P = 0.007$), cumulatively accounted for 93.7 % of the total Chl-a spatial variation (Table S2, Table S4).

The temporal model for Chl-a performed slightly better than the temporal model of assemblage structure (Table S3). In marginal tests,

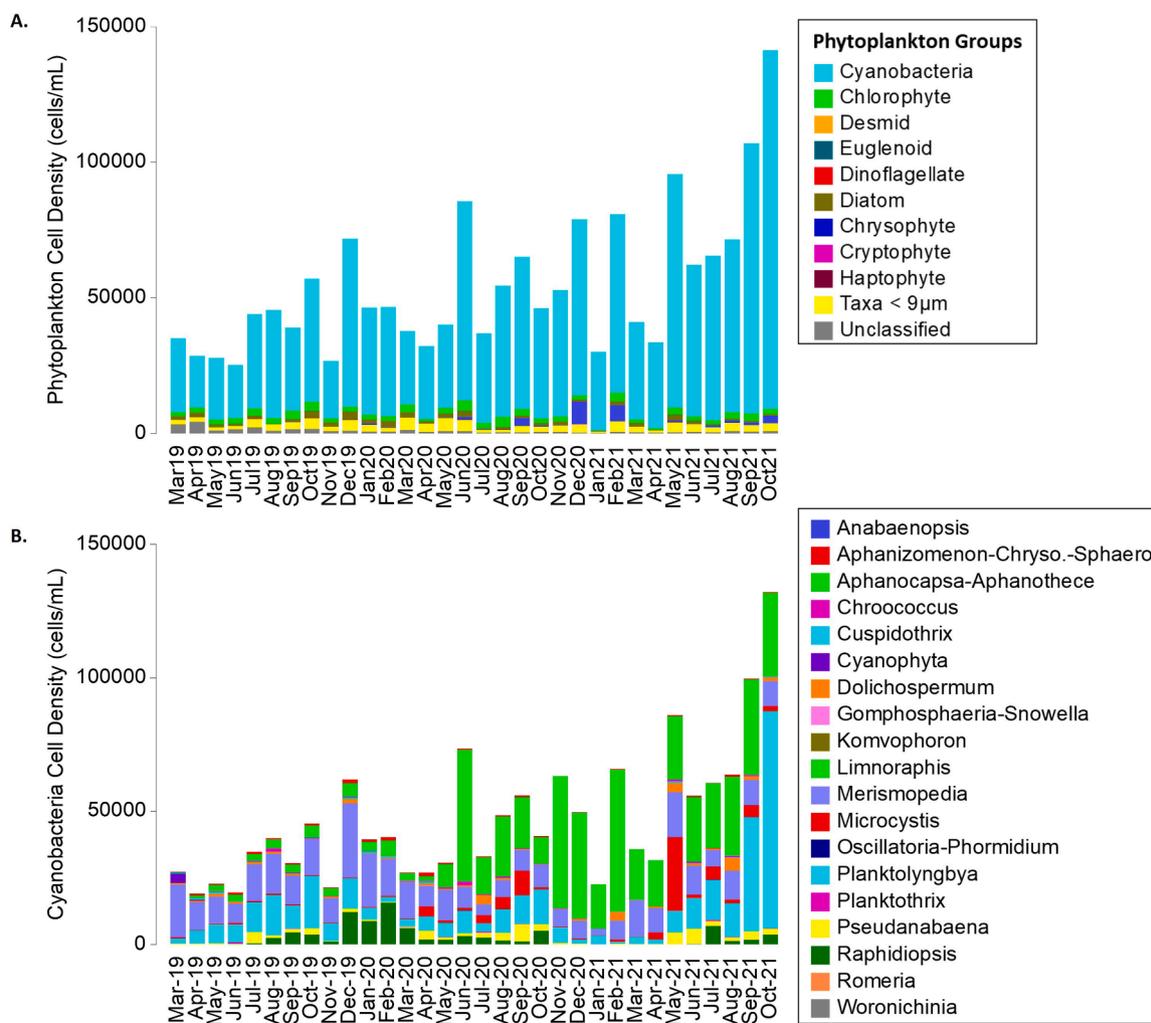


Fig. 3. Temporal patterns of phytoplankton assemblage structure. A) Bar plot showing the relative contribution of different algal functional groups to the total phytoplankton cell concentration during each of the 32 months samples (month-year). B) Bar plot showing the relative contribution of individual cyanobacteria taxa to the total cyanobacterial cell concentration during each of the 32 months samples. Aphanizomenon-Chryso.-Sphaero = Aphanizomenon-Chryso.-Sphaero. The color scheme of the legend corresponds to order in the bars. Some phytoplankton groups and cyanobacteria taxa listed in legends were present in very low abundance and may not be discernable in the plots.

$\text{NO}_3+\text{NO}_2-\text{N}$ explained the largest proportion of temporal Chl- α variation on its own (48.8 %, $P = 0.001$), followed by temp (48 %, $P = 0.001$), NH_3-N (39.3 %, $P = 0.001$), DO (37.3 %, $P = 0.001$), and precip (31 %, $P = 0.002$). The most parsimonious conditional model included $\text{NO}_3+\text{NO}_2-\text{N}$ (49 %, $P = 0.001$), $+\text{NH}_3-\text{N}$ (8.2 %, $P = 0.022$), $+\text{TSR}$ (6.6 %, $P = 0.31$) which together explained 64 % of the total temporal variation (Table S3, Table S4).

Chl- α , total cell concentration, and total biovolume abundance metrics exhibited moderate negative spatial correlations with $\text{NO}_3+\text{NO}_2-\text{N}$, ortho-P, and TA (Fig S2A). $\text{NO}_3+\text{NO}_2-\text{N}$, ortho-P, TA, and SpCond were lowest in the NW Lake zone where all three phytoplankton abundance metrics were highest (Fig. S2A and S3A). $\text{NO}_3+\text{NO}_2-\text{N}$ displayed a conspicuous seasonal trend of decreasing concentrations during the summer wet season months and increasing into the fall reaching peak concentrations in the dry season winter and spring months (Fig. S2A and S3B). This seasonal $\text{NO}_3+\text{NO}_2-\text{N}$ pattern was inversely related to phytoplankton abundance seasonal patterns which typically peaked during months when $\text{NO}_3+\text{NO}_2-\text{N}$ was lowest. In fact, the major HAB event that occurred in May 2021 was preceded by the highest mean $\text{NO}_3+\text{NO}_2-\text{N}$ concentrations measured over the study period, which then crashed during the bloom (Fig. S3B). Ortho-P fluctuated from month to month without obvious seasonal cycles, but the highest concentration outliers were observed in summer and fall (Fig

S3B). Conspicuous peaks in NH_3-N were observed in the summer particularly at the two lake outlet sites (S-77 and S-308) and at the site closest to the mouth of the Kissimmee River (KISSR0.0). TSR was highest in spring and summer while temp and precip were highest in summer and fall coincident with lowest DO, turbidity, pH, and TA measurements. The average depth of the Lake Okeechobee system tended to increase in the fall and winter accompanied by declines in SpCond following the summer rainy season. Temporally, WNDD and WND5 had slight negative correlations to all abundance metrics though these relationships were not strong (Fig. S2B).

3.3. Objective 3 - Identify Indicator Species and Thresholds using

TITAN was performed on the gradients of abiotic variables identified by the DISTLM marginal test as explaining the greatest proportion of spatial variation in phytoplankton assemblage structure when examined individually: $\text{NO}_3+\text{NO}_2-\text{N}$, ortho-P, and TA (Table 1). The gradient of each variable is the range of values observed across the 21 sites over the 32-month sampling period. The $\text{NO}_3+\text{NO}_2-\text{N}$ gradient ranged from 0.5 – 585 $\mu\text{g L}^{-1}$, the ortho-P gradient ranged from 0.5 – 244 $\mu\text{g L}^{-1}$, and the TA gradient ranged from 27.5 – 161 mg L^{-1} . The analysis identified 39 negative (z-) pure and reliable indicator taxa for $\text{NO}_3+\text{NO}_2-\text{N}$ and a community threshold (i.e. change point = CP) at 179 $\mu\text{g L}^{-1}$ (Fig. 5,

Table 1

Distance-based linear model (DISTLM) results for time-averaged spatial patterns in phytoplankton assemblage structure. Marginal tests report statistics for each of the 16 explanatory variables individually and conditional tests report statistics for variables selected by the forward selection model that cumulatively explain the greatest proportion of total spatial variation. SS(regression) = sum of squares for the regression; Pseudo-F = pseudo-F statistic, Prop. (R²) = proportion of explained variation, Cumul.Prop = cumulative proportion of explained variation; res.df = residual degrees of freedom. Sqr() denoted a square-root transformation and Log() denotes a logarithmic transformation.

Marginal tests							
Variable	SS(regression)	Pseudo-F	P-value	Prop. (R ²)			
Sqr(NH3-N)	320.8	1.3	0.262	0.065			
Sqr(NO3+NO2-N)	2219.4	15.5	0.001	0.449			
Log(TN)	832.3	3.8	0.028	0.168			
Log(TP)	622.0	2.7	0.074	0.126			
Log(Turbidity)	901.7	4.2	0.022	0.182			
Sqr(Ortho-P)	1653.0	9.5	0.001	0.334			
TA	1814.4	11.0	0.001	0.367			
Depth	347.7	1.4	0.232	0.070			
DO	247.7	1.0	0.357	0.050			
pH	287.6	1.2	0.312	0.058			
SpCond	1179.2	6.0	0.005	0.239			
Temp	922.2	4.4	0.022	0.187			
Sqr(Precip)	316.3	1.3	0.261	0.064			
TSR	516.9	2.2	0.107	0.105			
WNDD	243.2	1.0	0.362	0.049			
WNDS	506.5	2.2	0.087	0.102			
Total SS = 4942.5							
Conditional tests							
Variable	Adj R ²	SS(regression)	Pseudo-F	P-value	Prop. (R ²)	Cumul.Prop.	res.df
+Sqr(NO3+NO2-N)	0.420	2219.4	15.5	0.001	0.449	0.449	19
+LOG(TP)	0.651	1170.5	13.6	0.001	0.237	0.686	18
+WNDS	0.766	568.6	9.8	0.001	0.115	0.801	17
+SpCond	0.802	200.6	4.1	0.003	0.041	0.841	16
+LOG(Turbidity)	0.836	173.7	4.3	0.001	0.035	0.877	15
+Sqr(Ortho-P)	0.869	156.9	4.9	0.001	0.032	0.908	14
+TA	0.879	63.9	2.1	0.026	0.013	0.921	13
+DO	0.888	56.4	2.0	0.045	0.011	0.933	12
Residual SS = 225.8							

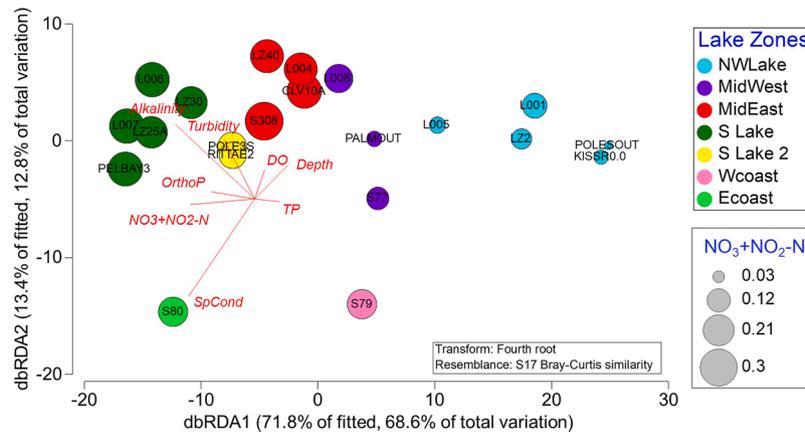


Fig. 4. Distance-based redundancy analysis (dbRDA) ordination for the most parsimonious model explaining the spatial patterns in phytoplankton assemblage structure. The significant environmental explanatory variables (transformed) identified by the model are plotted as vectors in red. Bubble size represent nitrate concentrations ($\mu\text{g L}^{-1}$, untransformed,) at each site as this variable explained the greatest amount of the spatial variation in phytoplankton assemblage structure.

Table 2. There were no positive (z+) pure and reliable indicator taxa for $\text{NO}_3+\text{NO}_2\text{-N}$. For ortho-P, TITAN returned 34 z- indicator taxa with a community threshold of $50 \mu\text{g L}^{-1}$ and 5 z+ indicator taxa with a community threshold of $94 \mu\text{g L}^{-1}$ (Fig. 5, Table 2). However, z+ indicators had relatively low IndVal z-scores and wide confidence intervals indicating large uncertainty around the CP. For the TA gradient, TITAN returned 41 z- indicator taxa with a community change point of 92mg L^{-1} and zero z+ indicator taxa (Fig. 5, Table 2). The prevalence of negative responder indicator taxa is supported by the negative correlations measured between total abundance metrics and $\text{NO}_3+\text{NO}_2\text{-N}$, ortho-P, and TA.

Negative indicator taxa with CPs at the lower end of the driver gradients are ones whose abundance declines at low nutrients and TA, while those with CPs at the upper end of the gradient decline in abundance when nutrients and TA are high. Given that high nutrients and TA were correlated to low algal abundance, conditions associated with the absence of bloom or the “bust” phase of the bloom cycle, taxa that declined in abundance during the bust phase represent bloom-associated taxa (Figs. 5 and 6). Therefore, z- taxa having CPs at the upper end of the environmental gradients are associated with blooms whether they be the dominant bloom formers or simply form an association with the bloom forming taxa. Z- taxa with CPs at the lower end of the gradient are ones

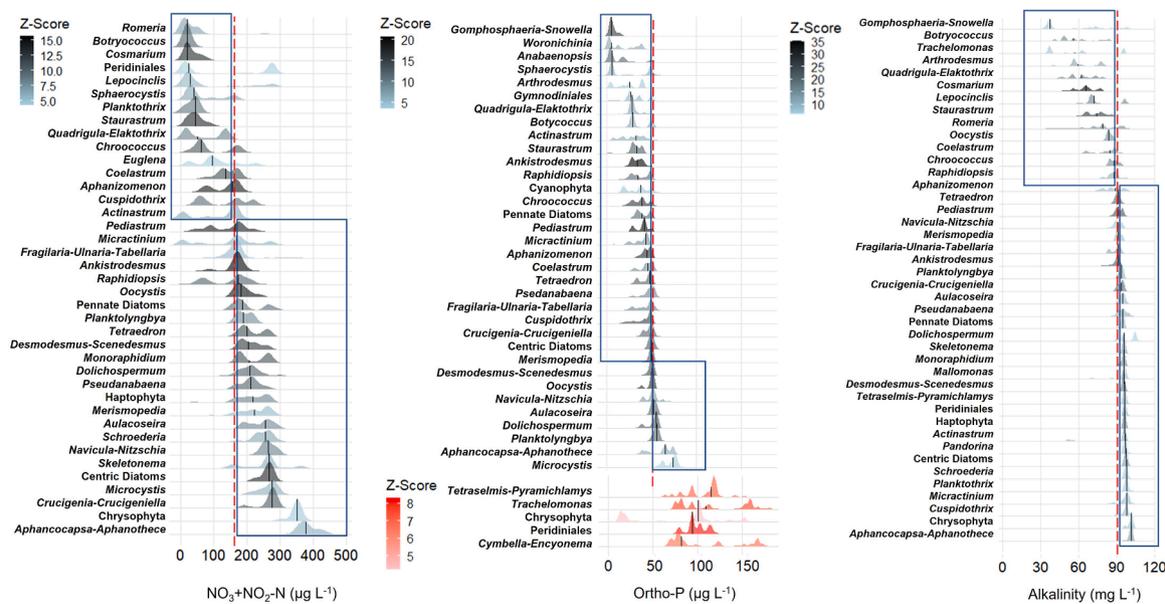


Fig. 5. Change point (CP) values for pure and reliable indicator taxa of nitrate ($\text{NO}_3 + \text{NO}_2\text{-N}$), ortho-Phosphate (ortho-P), and total alkalinity plotted as probability density functions with central tendency signaling the value of x resulting in the greatest indicator z value. Taxa with higher z -score values have a higher indicator value (Table 2). The red dotted lines represent community thresholds (i.e., CPs) where there is a synchronous response by indicator taxa. Negative indicator taxa with CPs to the left of the community threshold (red dotted line) are considered associated with non-bloom conditions while taxa to the right of the community thresholds are considered bloom-associated taxa.

that decline under “boom” conditions, when nutrients and alkalinity are lowest, and represent non-bloom-associated taxa (Figs. 5 and 6). The assemblage-level threshold was used as the cutoff between non-bloom and bloom associated taxa (Table 2).

The top ten pure and reliable indicator taxa (highest IndVal z -scores) of $\text{NO}_3 + \text{NO}_2\text{-N}$ were the green algae *Ankistrodesmus*, *Oocystis*, *Pediastrum*, *Desmodesmus-Scenedesmus*, *Tetraedron*, and *Coelastrum* and the desmids *Cosmarium* and *Staurastrum* along with centric diatoms and the cyanobacteria *Aphanizomenon* (Fig. 5, Table 2). The chlorophytes and centric diatoms were associated with bloom conditions, having CPs at concentrations above the assemblage-level threshold ($179 \mu\text{g L}^{-1}$), while the two desmids, *Cosmarium* and *Staurastrum*, were associated with non-bloom conditions having change points of declining abundance at $20 \mu\text{g L}^{-1}$. The top 10 ortho-P indicator taxa also included *Ankistrodesmus* (CP = $21 \mu\text{g L}^{-1}$), *Pediastrum* (CP = $35 \mu\text{g L}^{-1}$), *Oocystis* (CP = $51 \mu\text{g L}^{-1}$), *Desmodesmus-Scenedesmus* (CP = $54 \mu\text{g L}^{-1}$), *Staurastrum* (CP = $33 \mu\text{g L}^{-1}$) and *Aphanizomenon* (CP = $49 \mu\text{g L}^{-1}$), but also the cyanobacteria *Chroococcus* (CP = $39 \mu\text{g L}^{-1}$), *Gomphosphaeria-Snowella* (CP = $3 \mu\text{g L}^{-1}$), *Cuspidothrix* (CP = $49 \mu\text{g L}^{-1}$), and *Dolichospermum* (CP = $56 \mu\text{g L}^{-1}$, Fig. 5, Table 2). Green algae and desmids were the top indicator taxa for TA including *Cosmarium* (CP = 53mg L^{-1}) and *Staurastrum* (CP = 74mg L^{-1}), and *Pediastrum* (CP = 90mg L^{-1}), *Ankistrodesmus* (CP = 90mg L^{-1}), *Tetraedron* (CP = 92mg L^{-1}), *Crucigenia-Crucigeniella* (CP = 95mg L^{-1}), *Desmodesmus-Scenedesmus* (CP = 96mg L^{-1}), *Quadrigula-Elakatothrix* (CP = 53mg L^{-1}), and *Monoraphidium* (CP = 96mg L^{-1}) as well as the euglenid *Lepocinclis* (CP = 71mg L^{-1} , Fig. 5, Table 2).

4. Discussion

This study (1) examined the spatial and temporal structure of phytoplankton assemblages in Lake Okeechobee and the two outflowing rivers connecting it to the western and eastern coasts of Florida, (2) identified explanatory abiotic variables shaping phytoplankton assemblage and HAB patterns, and (3) defined indicator species and their thresholds along the gradients of the strongest explanatory variables. As hypothesized, strong spatial structuring of phytoplankton assemblages by species-sorting and environmental filtering was observed. $\text{NO}_3 + \text{NO}_2\text{-N}$, ortho-P, and TA were identified as the strongest abiotic

variables driving spatial assemblage dissimilarity. Temporal phytoplankton variability was less deterministic at the whole-system scale perhaps due to the strong geographic zonation of assemblages, unexamined variables excluded from the models, the relatively short duration of this study (32 months), or stochastic processes. Assemblage-level and taxon-specific (i.e. indicator species) thresholds along the gradients of the top three explanatory variables of spatial assemblage patterns were identified allowing inferences regarding genera associated with “boom” and “bust” phases of HAB cycles, characterized by low and high $\text{NO}_3 + \text{NO}_2\text{-N}$, ortho-P, and TA, respectively.

Seven distinct lake zones with significantly dissimilar phytoplankton assemblages and a trend of decreasing biomass from north to south were identified. The deterministic process of species sorting through abiotic environmental filters was largely responsible for the spatial structuring of assemblages as evidenced by the DISTLM model which explained 93 % of the total spatial assemblage variability. The significant explanatory variables selected by the most parsimonious conditional model represent the set of abiotic filters through which the regional phytoplankton pool passes before being sorted into these 7 geographic zones. The classification of lake zones based on distinct phytoplankton assemblages described here generally aligns with previously classified limnological zones in Lake Okeechobee based on Chl- a and other water quality metrics (Phillips et al., 1993). Chl- a in the northwestern region of the lake remains significantly higher than other regions, supporting recent findings (Wachnicka et al., 2023; Krausfeldt et al., 2024) as well as studies going back three decades (Phillips et al., 1993; Havens et al., 1994). Total phytoplankton cell concentration and biovolume were also elevated in this region; these metrics of algal abundance provide additional information about the quantity of phytoplankton. Chl- a is commonly used as a proxy for algal biomass but its content varies intra- and interspecifically preventing a direct correlation to the number of cells per unit volume, a more accurate measure of abundance, while biovolume estimates the 3D space occupied by those cells given different cells sizes and shapes. The NW lake was also characterized by significantly dissimilar phytoplankton assemblages from the rest of the lake, a finding supported by Krausfeldt et al. (2024) who reported significantly distinct cyanobacterial assemblages based on 16S rRNA gene sequencing in the nearshore, northwest area of the lake correlated to distinct

Table 2

TITAN results of taxon and community-level change points (CP) for pure and reliable indicator species (≥ 0.95) along the $\text{NO}_3+\text{NO}_2\text{-N}$, ortho-P, and total alkalinity gradients observed across the 21 sites and over the 32-month sampling period. Taxon-specific indicator value z-scores (zscore), taxon frequency (Freq), and classification as positive (z+) or negative (z-) responder are also provided. Differences in taxa frequencies for each gradient are due to lower sample sizes for the $\text{NO}_3+\text{NO}_2\text{-N}$ and ortho-P gradients where samples with zero or negative values were removed ($n = 526$ for $\text{NO}_3+\text{NO}_2\text{-N}$; $n = 600$ for ortho-P; $n = 639$ for total alkalinity which had no zero values).

Algal group	Indicator taxa	$\text{NO}_3+\text{NO}_2\text{-N}$				ortho-P				Total alkalinity			
		CP ($\mu\text{g}/\text{L}$)	Freq	zscore	-/+	CP ($\mu\text{g}/\text{L}$)	Freq	zscore	-/+	CP (mg/L)	Freq	zscore	-/+
Cyanobacteria	<i>Anabaenopsis</i>	–	–	–	–	2.0	17	12.1	z-	–	–	–	–
Cyanobacteria	<i>Aphanizomenon</i>	175.5	214	14.2	z-	49.0	275	17.2	z-	90.0	311	9.1	z-
Cyanobacteria	<i>Aphanocapsa- Aphanothece</i>	378.5	520	6.3	z-	64.0	594	6.6	z-	101.8	633	12.9	z-
Cyanobacteria	<i>Chroococcus</i>	58.8	187	10.9	z-	39.0	242	16.0	z-	86.3	276	15.1	z-
Cyanobacteria	<i>Cuspidothrix</i>	53.5	127	9.3	z-	49.0	170	15.0	z-	98.5	196	6.9	z-
Cyanobacteria	<i>Cyanophyta</i>	–	–	–	–	38.5	65	4.4	z-	–	–	–	–
Cyanobacteria	<i>Dolichospermum</i>	207.3	332	9.5	z-	55.8	396	13.1	z-	105.0	434	5.9	z-
Cyanobacteria	<i>Gomphosphaeria- Snowella</i>	–	–	–	–	3.3	23	16.0	z-	38.0	27	6.7	z-
Cyanobacteria	<i>Merismopedia</i>	264.0	525	5.4	z-	48.9	598	12.4	z-	90.0	637	7.8	z-
Cyanobacteria	<i>Microcystis</i>	278.5	411	5.9	z-	78.0	480	3.6	z-	–	–	–	–
Cyanobacteria	<i>Planktolyngbya</i>	209.0	502	7.9	z-	55.0	572	12.3	z-	96.0	608	9.0	z-
Cyanobacteria	<i>Planktothrix</i>	41.6	44	9.2	z-	–	–	–	–	98.0	64	6.6	z-
Cyanobacteria	<i>Pseudanabaena</i>	220.5	389	10.7	z-	46.6	450	9.4	z-	92.3	486	9.2	z-
Cyanobacteria	<i>Raphidiopsis</i>	178.8	160	9.9	z-	28.0	197	11.5	z-	90.0	229	10.5	z-
Cyanobacteria	<i>Romeria</i>	50.0	60	6.2	z-	–	–	–	–	78.8	71	14.6	z-
Cyanobacteria	<i>Woronichinia</i>	–	–	–	–	3.5	6	6.9	z-	–	–	–	–
Diatom	<i>Atuloseira</i>	260.5	457	8.2	z-	52.3	521	11.8	z-	93.3	554	8.9	z-
Diatom	<i>Cymbella-Encyonema</i>	–	–	–	–	51.3	20	12.6	z-	–	–	–	–
Diatom	<i>Fragilaria-Ulnaria- Tabellaria</i>	162.9	386	5.5	z-	78.0	447	7.0	z+	92.0	481	11.0	z-
Diatom	<i>Navicula-Nitzschia</i>	291.5	478	8.7	z-	48.8	550	9.4	z-	83.3	588	7.1	z-
Diatom	<i>Skeletonema</i>	267.5	171	5.8	z-	–	–	–	–	96.0	194	12.1	z-
Diatom	Centric Diatoms	287.0	480	13.2	z-	51.5	547	7.6	z-	96.3	581	12.4	z-
Diatoms	Pennate Diatoms	170.0	412	9.5	z-	46.0	481	11.7	z-	95.0	516	10.2	z-
Dinoflagellate	Gymnodinales	–	–	–	–	36.5	51	5.4	z-	–	–	–	–
Dinoflagellate	Peridinales	7.8	150	4.4	z-	93.9	172	8.3	z+	96.0	182	9.5	z-
Euglenid	<i>Euglena</i>	116.3	56	4.6	z-	–	–	–	–	–	–	–	–
Euglenid	<i>Lepocinclis</i>	29.5	91	5.2	z-	–	–	–	–	70.8	107	15.9	z-
Euglenid	<i>Trachelomonas</i>	–	–	–	–	159.5	22	7.9	z+	38.5	24	7.3	z-
Green algae	<i>Actinastrum</i>	169.0	57	6.3	z-	32.3	79	6.5	z-	97.0	83	10.3	z-
Green algae	<i>Ankistrodesmus</i>	175.0	95	15.7	z-	21.0	133	20.7	z-	90.0	153	23.0	z-
Green algae	<i>Botryococcus</i>	16.8	34	10.4	z-	29.0	40	6.6	z-	49.3	44	15.9	z-
Green algae	<i>Coelastrum</i>	170.0	83	11.4	z-	50.8	105	8.7	z-	64.0	112	14.5	z-
Green algae	<i>Crucigenia- Crucigeniella</i>	276.5	350	10.6	z-	50.5	410	11.9	z-	95.0	438	17.7	z-
Green algae	<i>Desmodesmus- Scenedesmus</i>	230.0	489	12.3	z-	53.8	562	14.3	z-	96.0	601	17.3	z-
Green algae	<i>Micractinium</i>	184.3	82	6.0	z-	46.0	97	8.4	z-	99.0	104	8.9	z-
Green algae	<i>Monoraphidium</i>	175.0	442	11.1	z-	–	–	–	–	96.3	531	16.6	z-
Green algae	<i>Oocystis</i>	222.5	251	15.6	z-	51.0	309	17.2	z-	83.5	342	13.9	z-
Green algae	<i>Pandorina</i>	–	–	–	–	–	–	–	–	96.0	41	8.2	z-
Green algae	<i>Pediastrum</i>	175.0	210	13.6	z-	34.5	263	20.3	z-	90.0	290	23.2	z-
Green algae	<i>Quadrigula- Elakatothrix</i>	135.5	25	8.1	z-	28.0	35	6.5	z-	53.3	40	17.0	z-
Green algae	<i>Schroederia</i>	230.0	350	6.5	z-	–	–	–	–	98.0	419	7.0	z-
Green algae	<i>Sphaerocystis</i>	26.0	10	6.1	z-	6.0	19	6.6	z-	–	–	–	–
Green algae	<i>Tetraedron</i>	186.3	325	11.5	z-	34.0	382	12.3	z-	92.0	413	19.4	z-
Green algae	<i>Tetraselmis- Pyramichlamys</i>	–	–	–	–	116.5	85	7.3	z+	96.0	89	8.8	z-
Desmid	<i>Arthrodesmus</i>	–	–	–	–	39.0	12	4.0	z-	–	–	–	–
Desmid	<i>Cosmarium</i>	20.0	52	14.1	z-	–	–	–	–	53.3	81	33.5	z-
Desmid	<i>Staurastrum</i>	19.8	82	13.1	z-	32.8	107	12.7	z-	74.0	124	27.0	z-
Chrysophyte	Chrysophyta	350.0	306	4.7	z-	101.7	336	4.2	z+	102.0	350	8.3	z-
Chrysophyte	<i>Mallomonas</i>	–	–	–	–	–	–	–	–	97.0	64	8.0	z-
Haptophyte	Haptophyta	250.3	267	8.0	z-	–	–	–	–	97.0	321	14.6	z-
	Community-level threshold	CP	0.05 quantile	0.95 quantile		CP	0.05 quantile	0.95 quantile		CP	0.05 quantile	0.95 quantile	
	fsumz- (negative responder community)	178.8	162.9	226.6		50.3	30.0	51.5		92.0	84.5	96.8	
	fsumz+ (positive responder community)	NA	NA	NA		94.0	80.0	153.7		NA	NA	NA	

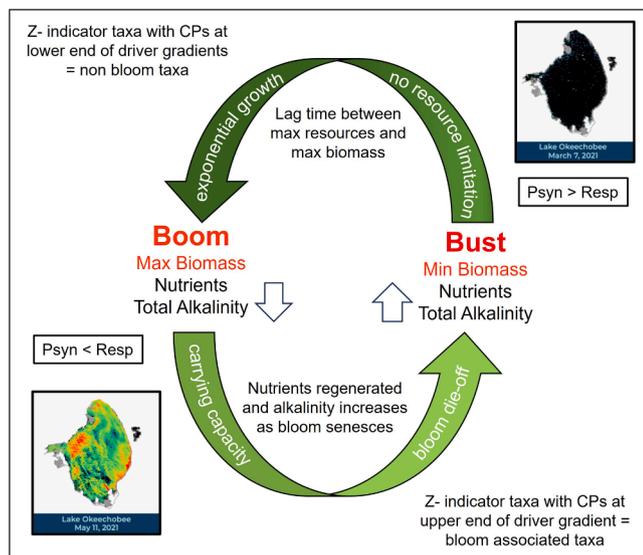


Fig. 6. A conceptual diagram depicting the hypothetical phytoplankton boom-to-bust feedback loop in Lake Okeechobee. The “boom” phase occurs when algal abundance peaks resulting in depletion of nutrients and decreased total alkalinity. Once carrying capacity is reached, high concentrations of algae can no longer be supported, the bloom begins to senesce, and the system becomes net heterotrophic as decomposition rates increase and respiration rates exceed photosynthetic rates. These changes lead into the “bust” phase where algal biomass declines, nutrients are regenerated, light availability improves, and phytoplankton are released from resource limitation, setting the stage for the next “boom” phase. Taxa declining during the boom to bust part of the cycle are bloom-associated, while those that decline during the bust to boom period are considered non-bloom associated taxa. Different shades of green arrows represent increasing (darker) or decreasing (lighter) algal biomass. Nutrients refer to nitrate and orthophosphate. Z- = negative responder indicator taxa identified by the Threshold Indicator Taxa Analysis (TITAN). Psynth = photosynthesis. Resp = respiration.

chemical and physical limnological characteristics compared to the pelagic region. The finding of distinct zones within this highly connected and relatively small-scale aquatic system, extending ~145 km east-west and ~50 km north-south, suggests environmental filtering and species sorting may override the especially high dispersal and homogenization potential for microorganisms at this scale, where individuals are constantly shuffled by natural process as well as actively transported by hydrological management schedules (Van der Gucht et al., 2007; Soinenen, 2014; Mazzei and Gaiser, 2018).

Temporal patterns of assemblage structure were also examined but no evidence of seasonal succession was found across meteorological seasons or subtropical wet/dry seasons. The most parsimonious DISTLM model only accounted for 48 % of the total temporal variation and cannot be expected to reliably predict assemblage patterns; nonetheless it offers insights into which variables may play a role in seasonal variability. We theorize that the strong geographic zonation of phytoplankton assemblages discovered in this study obscures whole lake succession patterns and that temporal turnover of assemblages may be occurring separately within each lake zone, but this remains to be tested. However, temporal stochasticity of factors such as major weather events or hydrological management decisions that affect many of the environmental variables shaping phytoplankton assemblages may also be responsible for seasonal unpredictability. A study by Deng et al. (2014) examined annual spring succession of phytoplankton in Lake Taihu, China, a large, shallow, subtropical lake similar to Lake Okeechobee. They found that nutrients and temperature were largely responsible for year-to-year differences in spring assemblages thus demonstrating the unpredictability of assemblages within a given season across years. In our study, significant phytoplankton assemblage dissimilarity was

present between 2019 and 2021, however there was no clear link to any of the variables examined. Unaccounted for abiotic, physical, or biological variables like concentrations of different dissolved inorganic carbon species or trace metals, underwater photosynthetically active radiation (PAR), and grazing may also be responsible for the absence of significant seasonal succession patterns. Although previous studies have demonstrated that grazing pressure does not play a significant role in phytoplankton dynamics in Lake Okeechobee (Havens et al., 1996), the influence of top-down control by grazing should be reexamined. Furthermore, higher frequency (e.g. daily), long-term phytoplankton assemblage and environmental data collection would greatly aid in the accurate detection of successional patterns and the identification of triggers leading to the development of HABs, but there are many challenges to implementing this type of monitoring (Johnston et al., 2024). Although no evidence of predictable seasonal succession was found, total abundance metrics did exhibit significant temporal trends consistent with typical phytoplankton growth patterns for subtropical systems where algal abundance is highest during the summer and fall months of the wet season (Havens et al., 1994; Philips et al., 1997, 2012; Wachnicka et al., 2022, 2023). The results imply that significant seasonal trends in total abundance (i.e. Chl-*a*, cell density, and biovolume) do not necessarily correspond to significant turnover of phytoplankton assemblages but rather that dominant taxa remain consistently dominant across seasons and changes in their abundance drive the significant increases in total phytoplankton abundance during summer/fall wet seasons.

Inorganic nutrients were important explanatory variables of both spatial and temporal assemblage and Chl-*a* patterns. The inverse relationship between algal abundance and $\text{NO}_3+\text{NO}_2-\text{N}$ and ortho-P may seem counterintuitive but is well documented in the Lake Okeechobee system (McPherson and Rose, 1981; Havens et al. 1994; James et al., 2011; Wachnicka et al., 2022; Krausfeldt et al., 2024) and other eutrophic lakes (Filstrup and Downing, 2017). This phenomenon can be attributed to a lag between peak inorganic nutrient availability and peak abundance during phytoplankton “boom-and-bust” periods in systems that are chronically enriched with both N and P. As phytoplankton reach carrying capacity during the “boom” phase, they deplete nutrients in the water column and reduce light availability by self-shading leading to declines in biomass as the bloom senesces and eventually dies off or “busts.” This in turn replenishes the nutrient pool as the organic matter decomposes and light availability increases, setting up conditions for the next “boom” phase. Havens et al. (1994) proposed an alternative mechanism for the inverse relationship between P and Chl-*a* and blooms in Lake Okeechobee whereby periods of high wind velocity resuspend P-rich sediments in the shallow lake simultaneously reducing light availability which limits algal growth and prevent bloom development despite elevated P. High concentration of legacy P stored in Lake Okeechobee sediments is a well-studied characteristic of the system (Missimer et al., 2021). Likely due to the much shorter duration of our study, wind and light were not strong explanatory variables of phytoplankton dynamics, though wind speed was positively correlated to ortho-P and turbidity but negatively correlated to algal abundance which supports, or at least does not refute, the Havens et al. theory. Nevertheless, our results point to a greater role of external nutrient loads and rapid nutrient recycling during boom-to-bust phases in determining HAB dynamics, particularly in the nearshore, northwestern section of the lake. Higher algal abundance and prevalence of HABs in the NW lake zone are generally thought to be driven by the high volumes of nutrient-rich water delivered via watershed runoff and point source releases from the Kissimmee River and several canals and creeks draining the surrounding agricultural areas (Ma et al., 2020; Wachnicka et al., 2022; Zhang et al., 2022). Low $\text{NO}_3+\text{NO}_2-\text{N}$ and ortho-P concentrations in the NW lake zone despite large inputs of nutrient-rich water supports the idea that nutrients are rapidly assimilated into algal and plant biomass. The lag between inorganic nutrient highs and algal abundance highs was especially apparent during the May 2021 HAB event which occurred following a peak in $\text{NO}_3+\text{NO}_2-\text{N}$

concentrations in the prior dry season and a nutrient crash during the bloom.

Low total alkalinity, one of the strongest explanatory variables of phytoplankton spatial patterns along with $\text{NO}_3+\text{NO}_2-\text{N}$ and ortho-P, in the NW lake zone, could be due to natural hydrogeomorphology but could also be the result of the frequent high biomass blooms that occur there. In water where pH is near neutral, bicarbonate (HCO_3) is the dominant inorganic carbon species and algae must convert HCO_3 to carbon dioxide (CO_2) intracellularly for use in photosynthesis (Reynolds, 2006; Falkowski and Raven, 2013). Naturally lower TA may aid in bloom formation by making the preferred carbon species (CO_2) more abundant (Raven et al., 2020). On the other hand, the high abundance blooms that occur in this zone might lower TA through the removal of large quantities of HCO_3 for photosynthesis and carbonate (CO_3) by calcite precipitation (Wetzel, 2001). Cyanobacteria are well known biocalcifiers as are the benthic green macroalgae *Chara* and *Nitella* common in the nearshore areas of the lake (Raven, 2012). High respiration rates and CO_2 production, especially during bloom senescence, could also be contributing to lower TA. Nutrient assimilation and remineralization during photosynthesis and respiration, respectively, also alter TA. The relationship between TA and dissolved inorganic carbon composition with HABs needs to be studied in more detail, especially in the context of climate change scenarios.

Assemblage-level thresholds, indicator taxa, and taxon-specific thresholds were identified for the $\text{NO}_3+\text{NO}_2-\text{N}$, ortho-P, and TA gradients. Identification of indicator taxa and their environmental thresholds to prevailing environmental gradients is a valuable tool for the management of specific taxa of interest such as those associated with bloom and non-bloom conditions (Groffman et al., 2006; Foley et al., 2015). TITAN provides this information in addition to assemblage-level thresholds that mark the critical point along a gradient at which the assemblage synchronistically shifts from one state to another, an important criterion for ecosystem management that promotes environmental conditions for favorable phytoplankton assemblages (e.g. those associated with non-bloom conditions). Here, taxa that declined in abundance towards the upper end of the nutrient and alkalinity gradients were associated with “boom” phase conditions and decline as the bloom enters the “bust” phase during which high nutrients and alkalinity are restored. These included the two most consistently abundant taxa across space and time, the cyanobacteria *Aphanocapsa* and *Merismopedia*. Both genera belong to the family Microcystaceae along with *Microcystis* (Strunecý et al., 2023), also identified as a bloom associated taxon by TITAN and a ubiquitous HAB-former worldwide (Paerl et al., 2001; LaPointe et al., 2024). Despite their phylogenetic proximity, *Aphanocapsa* is not known to contain aerotopes like *Merismopedia* and *Microcystis*, a trait which is thought to give species an advantage in forming high-biomass planktonic blooms. Several toxic cyanobacteria of concern, including *Aphanizomenon* and *Raphidiopsis*, were associated with the “bust” phase or were intermediate between phases, but as toxin producers these taxa can be hazardous even at low concentrations (Metcalf and Codd, 2012). Contrasting nutrient requirements and niches between *Microcystis* and *Raphidiopsis* in Lake Okeechobee were reported by Lefler et al. (2023). In addition to cyanobacteria, chlorophytes, centric diatoms, and *Navicula-Nitzschia* species were also significant indicators of “boom” conditions. Although chlorophytes are not generally known to produce toxins, they are common causes of high-biomass nuisance blooms with negative ecological consequences (Watson et al. 2015). Toxin-producing diatoms, while a concern in coastal and marine systems, are not common in freshwaters, but nontoxic diatoms can also create high-biomass nuisance blooms in freshwaters (Watson et al. 2015).

This study highlights the complexity of phytoplankton and HAB dynamics, particularly in heavily managed and impaired lakes where novel anthropogenic gradients have developed from decades of human intervention and potential long-term climate changes. Continued efforts that build upon and expand existing knowledge are necessary to develop

the predictive models for HABs that resource managers need to adequately adjust strategies for HAB mitigation, and ultimately, prevention. The implementation of high-frequency, high-resolution monitoring systems could also greatly improve model-based management decisions. These goals will likely take a long time to achieve, even after preventative measures are taken due to legacy effects of nutrient loading and will require a continuous adaptive management approach as new information becomes available.

Data availability

Phytoplankton abundance data are available in Mazzei and Sullivan 2022. All water quality data for the 21 sampling locations are available in the South Florida Water Management District’s DBHYDRO Browser (sfwmd.gov) or the USGS National Water Information System (NWIS, USGS Water Data for the Nation).

CRedit authorship contribution statement

Viviana Mazzei: Writing – original draft, Visualization, Project administration, Investigation, Formal analysis, Data curation, Conceptualization. **Kristy Lee Sullivan:** Writing – review & editing, Investigation, Data curation. **Keith Loftin:** Writing – review & editing, Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the U.S. Army Corps of Engineers Aquatic Nuisance Species Research Program (IAA # 7600A-2019-El-325) and the USGS Toxic Substances Hydrology Program for funding this project. We thank the South Florida Water Management District for collecting and shipping water samples from field sites and Emily Karwacki for providing sample shipping and receiving services. We also thank Dr. Barry Rosen for establishing this project in cooperation with Dr. Jose Lopez and Dr. Hidetoshi Urakawa. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.hal.2024.102744.

References

- Anderson, M.J., Clarke, K.R., Gorley, R.N., 2008. PERMANOVA+ for PRIMER: guide to software and statistical methods. PRIMER-E: Plymouth, UK. PRIMER-e.
- Austin, M., 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecol. Model.* 200 (1–2), 1–19. <https://doi.org/10.1016/j.ecolmodel.2006.07.005>.
- Baas Becking L.G.M. (1934). *Geobiologie of Inleiding Tot De milieukunde*. The Hague, the Netherlands: W.P. Van Stockum & Zoon, <https://lib.ugent.be/catalog/rug01:001271403>.
- Baker, M.E., King, R.S., 2010. A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods Ecol. Evol.* 1 (1), 25–37. <https://doi.org/10.1111/j.2041-210X.2009.00007.x>.
- Baker, M., King, R., Kahle, D., 2020. An Introduction to TITAN2, Version 2.4.3. GitHub - dkahle/TITAN2: Threshold Indicator Taxa Analysis.
- Baker M., King R., Kahle D. (2023). TITAN2: threshold Indicator Taxa Analysis. R package version 2.4.3, <https://CRAN.R-project.org/package=TITAN2>.
- Bernard, C., Ballot, A., Thomazeau, S., Maloufi, S., Furey, A., Mankiewicz-Boczek, J., Pawlik-Skowrońska, B., Capelli, C. and Salmaso, N. (2016). Appendix 2: cyanobacteria associated with the production of cyanotoxins. *Handbook Of*

- Cyanobacterial Monitoring And Cyanotoxin Analysis, pp.501–525. <https://doi.org/10.1002/9781119068761>.
- Brooks, B.W., Lazorchak, J.M., Howard, M.D., Johnson, M.V.V., Morton, S.L., Perkins, D.A., Reavie, E.D., Scott, G.I., Smith, S.A., Stevens, J.A., 2016. Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic ecosystems? *Environ. Toxicol. Chem.* 35 (1), 6–13. <https://doi.org/10.1002/etc.3220>.
- Canfield, D.E., 2015. *Baas Becking's geobiology: Or Introduction to Environmental Science*. English translation John Wiley & Sons. <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118295472>.
- Chase, J.M., Myers, J.A., 2011. Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. Royal Soc. B: Biol. Sci.* 366, 2351–2363. <https://doi.org/10.1098/rstb.2011.0063>, 1576.
- Clarke, K.R., Gorley, R.N., 2015. *PRIMER v7: user Manual/Tutorial*. PRIMER-E Ltd. Plymouth, UK. PRIMER-e.
- Deng, J., Qin, B., Pearl, H.W., Zhang, Y., Wu, P., Ma, J., Chen, Y., 2014. Effects of nutrients, temperature and their interactions on spring phytoplankton community succession in Lake Taihu, China. *PLoS One* 9 (12), e113960. <https://doi.org/10.1371/journal.pone.0113960>.
- de J Magalhães, A.A., da Luz, L.D., de Aguiar Jr, T.R., 2019. Environmental factors driving the dominance of the harmful bloom-forming cyanobacteria *Microcystis* and *Aphanocapsa* in a tropical water supply reservoir. *Water Environ. Res.* 91, 1466–1478. <http://doi.org/10.1002/wer.1141>.
- Falkowski, P.G., Raven, J.A., 2013. *Carbon Acquisition and Assimilation In Aquatic Photosynthesis*, Chapter 5. Princeton University Press. *Aquatic Photosynthesis | Princeton University Press*.
- Filstrup, C.T., Downing, J.A., 2017. Relationship of chlorophyll to phosphorus and nitrogen in nutrient-rich lakes. *Inland. Waters.* 7 (4), 385–400. <https://doi.org/10.1080/20442041.2017.1375176>.
- Foley, M.M., Martone, R.G., Fox, M.D., Kappel, C.V., Mease, L.A., Erickson, A.L., Halpern, B.S., Selkoe, K.A., Taylor, P., Scarborough, C., 2015. Using ecological thresholds to inform resource management: current options and future possibilities. *Front. Mar. Sci.* 2, 95. <https://doi.org/10.3389/fmars.2015.00095>.
- Glibert, P.M., 2020. Harmful algae at the complex nexus of eutrophication and climate change. *Harmful. Algae* 91, 101583. <https://doi.org/10.1016/j.hal.2019.03.001>.
- Glibert, P.M., Burford, M.A., 2017. Globally changing nutrient loads and harmful algal blooms: recent advances, new paradigms, and continuing challenges. *Oceanography* 30 (1), 58–69. <https://doi.org/10.5670/oceanog.2017.110>. Recent advances, new paradigms, and continuing challenges. *Oceanography* 30(1):58–69.
- Glibert, P.M., 2016. Margalef revisited: a new phytoplankton mandala incorporating twelve dimensions, including nutritional physiology. *Harmful. Algae* 55, 25–30. <https://doi.org/10.1016/j.hal.2016.01.008>.
- Glibert, P.M., Seitzinger, S., Heil, C.A., Burkholder, J.M., Parrow, M.W., Codispoti, L.A., Kelly, V., 2005. The role of eutrophication in the global proliferation of harmful algal blooms. *Oceanography* 18 (2), 198–209. <https://doi.org/10.5670/oceanog.2005.54>.
- Groffman, P.M., Baron, J.S., Blett, T., Gold, A.J., Goodman, I., Gunderson, L.H., Levinson, B.M., Palmer, M.A., Paerl, H.W., Peterson, G.D., Poff, N.L., 2006. Ecological thresholds: the key to successful environmental management or an important concept with no practical application? *Ecosystems* 9, 1–13. <https://doi.org/10.1007/s10021-003-0142-z>.
- Guntenspergen, G.R., Gross, J., 2014. Threshold concepts: implications for the management of natural resources. Editor. In: Guntenspergen, G.R. (Ed.), *Application of Threshold Concepts in Natural Resource Decision Making*. Springer, New York, pp. 1–7. https://doi.org/10.1007/978-1-4899-8041-0_1.
- Haakonsson, S., Rodríguez-Gallego, L., Somma, A., Bonilla, S., 2017. Temperature and precipitation shape the distribution of harmful cyanobacteria in subtropical lotic and lentic ecosystems. *Sci. Total Environ.* 609, 1132–1139. <https://doi.org/10.1016/j.scitotenv.2017.07.067>.
- Havens, K., East, T., Beaver, J., 1996. Experimental studies of zooplankton–phytoplankton–nutrient interactions in a large subtropical lake (Lake Okeechobee, Florida, USA). *Freshw. Biol.* 36 (3), 579–597. <https://doi.org/10.1046/j.1365-2427.1996.00122.x>.
- Havens, K., Hanlon, C., James, R.T., 1994. Seasonal and spatial variation in algal bloom frequencies in Lake Okeechobee, Florida, USA. *Lake Reserv. Manag.* 10, 139–148. <https://doi.org/10.1080/07438149409354185>.
- Ho, J.C., Michalak, A.M., Pahlevan, N., 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature* 574, 667–670. <https://doi.org/10.1038/s41586-019-1648-7>.
- James, R.T., Gardner, W.S., McCarthy, M.J., Carini, S.A., 2011. Nitrogen dynamics in Lake Okeechobee: forms, functions, and changes. *Hydrobiologia* 669, 199–212. <https://doi.org/10.1007/s10750-011-0683-7>.
- Johnston, B.D., Finkelstein, K.M., Gifford, S.R., Stouder, M.D., Nystrom, E.A., Savoy, P., Rosen, J.J., Jennings, M.B., 2024. Evaluation of Sensors For Continuous Monitoring of Harmful Algal Blooms in the Finger Lakes Region. 2019 and 2020: U.S. Geological Survey Scientific Investigations Report 2024-5010, New York, p. 54. <https://doi.org/10.3133/sir20245010>.
- King, R.S., Baker, M.E., 2010. Considerations for analyzing ecological community thresholds in response to anthropogenic environmental gradients. *J. North Am. Benthol. Soc.* 29 (3), 998–1008. <https://doi.org/10.1899/09-144.1>.
- King, R.S., Baker, M.E., 2014. Use, misuse, and limitations of Threshold Indicator Taxa Analysis (TITAN) for natural resource management. Editor. In: Guntenspergen, G.R. (Ed.), *Application of Threshold Concepts in Natural Resource Decision Making*. Springer, New York, pp. 231–254. https://doi.org/10.1007/978-1-4899-8041-0_11.
- Krausfeldt, L.E., Shmakova, E., Lee, H.W., Mazzei, V., Loftin, K.A., Smith, R.P., Karwacki, E., Fortman, P.E., Rosen, B.H., Urakawa, H., Dadlani, M., Colwell, R.R., Lopez, J.V., 2024. Microbial diversity, genomics, and phage-host interactions of cyanobacterial harmful algal blooms. *Environ. Microbiol.* 7 (9). <https://doi.org/10.1128/mSystems.00709-23>.
- Langenheder, S., Székely, A.J., 2011. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISME J.* 5 (7), 1086–1094. <https://doi.org/10.1038/ismej.2010.207>.
- Lapointe, B.E., Brewton, R.A., McFarland, M.N., Stockley, N., 2024. Nutrient availability in a freshwater-to-marine continuum: cyanobacterial blooms along the Lake Okeechobee Waterway. *Harmful. Algae* 139. <https://doi.org/10.1016/j.hal.2024.102710>.
- Lefler, F.W., Barbosa, M., Zimba, P.V., Smyth, A.R., Berthold, D.E., Laughinghouse, H.D., 2023. Spatiotemporal diversity and community structure of cyanobacteria and associated bacteria in the large shallow subtropical Lake Okeechobee (Florida, United States). *Front. Microbiol.* 14, 1219261. <https://doi.org/10.3389/fmicb.2023.1219261>.
- Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M., Gonzalez, A., 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.* 7 (7), 601–613. <https://doi.org/10.1111/j.1461-0248.2004.00608.x>.
- Ma, P., Zhang, L., Mitsch, W.J., 2020. Investigating sources and transformations of nitrogen using dual stable isotopes for Lake Okeechobee restoration in Florida. *Ecol. Eng.* 155, 105947. <https://doi.org/10.1016/j.ecoleng.2020.105947>.
- Mazzei, V., Gaiser, E., 2018. Diatoms as tools for inferring ecotone boundaries in a coastal freshwater wetland threatened by saltwater intrusion. *Ecol. Indic.* 88, 190–204. <https://doi.org/10.1016/j.ecolind.2018.01.003>.
- Mazzei, V., Sullivan, K.L. (2022). Monthly phytoplankton taxonomic quantification in Lake Okeechobee and the Okeechobee Waterway, Florida, USA, 2019–2021: U.S. Geological Survey Data Release, <https://doi.org/10.5066/P9QOJ4JZ>. Dataset.
- McPherson, B.F., La Rose, H.R., 1981. Algal conditions in the Caloosahatchee river (1975–79). In: *Lake Okeechobee to Franklin Lock, Florida*, 81. US Geological Survey, Water Resources Division. <https://doi.org/10.3133/wri8181>.
- McQuatters-Gollop, A., Johns, D.G., Bresnan, E., Skinner, J., Rombouts, I., Stern, R., Aubert, A., Johansen, M., Bedford, J., Knights, A., 2017. From microscope to management: the critical value of plankton taxonomy to marine policy and biodiversity conservation. *Mar. Policy* 83, 1–10. <https://doi.org/10.1016/j.marpol.2017.05.022>.
- Metcalf, J.S., Codd, G.A., 2012. Cyanotoxins. In *Ecology of cyanobacteria II: their diversity in space and time*. Springer Netherlands, Dordrecht, pp. 651–675. https://doi.org/10.1007/978-94-007-3855-3_24.
- Missimer, T.M., Thomas, S., Rosen, B.H., 2021. Legacy phosphorus in Lake Okeechobee (Florida, USA) sediments: a review and new perspective. *Water (Basel)* 13 (1), 39. <https://doi.org/10.3390/w13010039>.
- Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P., Ferrenberg, S., 2013. Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77 (3), 342–356. <https://doi.org/10.1128/mbr.00051-12>.
- Nichols, J.D., Eaton, M.J., Martin, J., 2014. Thresholds for conservation and management: structured decision making as a conceptual framework. Editor. In: Guntenspergen, G.R. (Ed.), *Application of Threshold Concepts in Natural Resource Decision Making*. Springer, New York, pp. 9–28. https://doi.org/10.1007/978-1-4899-8041-0_2.
- O'Malley, M.A., 2008. Everything is everywhere: but the environment selects: ubiquitous distribution and ecological determinism in microbial biogeography. *Stud. Hist. Philos. Biol. Biomed. Sci.* 39 (3), 314–325. <https://doi.org/10.1016/j.shps.2008.06.005>.
- Paerl, H.W., Fulton, R.S., Moisaner, P.H., Dyble, J., 2001. Harmful freshwater algal blooms, with an emphasis on cyanobacteria. *Sci. World J.* 1, 76–113. <https://doi.org/10.1100/tsw.2001.16>.
- Phlips, E.J., Cichra, M., Havens, K., Hanton, C., Badylak, S., Rueter, B., Randall, M., Hansen, P., 1997. Relationships between phytoplankton dynamics and the availability of light and nutrients in a shallow sub-tropical lake. *J. Plankton Res.* 19 (3), 319–342. <https://doi.org/10.1093/plankt/19.3.319>.
- Phlips, E.J., Aldridge, F.J., Hansen, P., Zimba, P.V., Inhat, J., Conroy, M., 1993. Spatial and temporal variability of trophic state parameters in a shallow subtropical lake (Lake Okeechobee, Florida, USA). *Archiv. für Hydrobiologie* 128 (4), 437–458. <http://doi.org/10.1127/archiv-hydrobiol/128/1993/437>.
- Phlips, E.J., Badylak, S., Hart, J., Haunert, D., Lockwood, J., O'Donnell, K., Sun, D., Viveros, P., Yilmaz, M., 2012. Climatic influences on autochthonous and allochthonous phytoplankton blooms in a subtropical estuary. *Estuar. Coasts* 35, 335–352. <https://doi.org/10.1007/s12237-011-9442-2>.
- R. Core Team. (2023). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Raven, J.A., 2012. Carbon. In: Whitton, B. (Ed.), *Ecology of Cyanobacteria II*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-3855-3_17.
- Raven, J.A., Gobler, C.J., Hansen, P.J., 2020. Dynamic CO₂ and pH levels in coastal, estuarine, and inland waters: theoretical and observed effects on harmful algal blooms. *Harmful. Algae* 91, 101594. <https://doi.org/10.1016/j.hal.2019.03.012>.
- Reynolds, C., 2006. *The Ecology of Phytoplankton (Ecology, Biodiversity and Conservation)*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511542145>.
- Ribeiro, K.F., Duarte, L., Crossetti, L.O., 2018. Everything is not everywhere: a tale on the biogeography of cyanobacteria. *Hydrobiologia* 820, 23–48. <https://doi.org/10.1007/s10750-018-3669-x>.
- Rojo, C., 2021. Community assembly: perspectives from phytoplankton's studies. *Hydrobiologia* 848, 31–52. <https://doi.org/10.1007/s10750-020-04249-3>.
- Siddig, A.A., Ellison, A.M., Ochs, A., Villar-Leeman, C., Lau, M.K., 2016. How do ecologists select and use indicator species to monitor ecological change? Insights

- from 14 years of publication in Ecological Indicators. *Ecol. Indic.* 60, 223–230. <https://doi.org/10.1016/j.ecolind.2015.06.036>.
- Smayda, T.J., Reynolds, C.S., 2001. Community assembly in marine phytoplankton: application of recent models to harmful dinoflagellate blooms. *J. Plankton Res.* 23 (5), 447–461. <https://doi.org/10.1093/plankt/23.5.447>.
- Soininen, J., 2014. A quantitative analysis of species sorting across organisms and ecosystems. *Ecology*. 95 (12), 3284–3292. <https://doi.org/10.1890/13-2228.1>.
- Spatharis, S., Lamprinou, V., Meziti, A., Kormas, K.A., Danielidis, D.D., Smeti, E., Roelke, D.L., Mancy, R., Tsirtis, G., 2019. Everything is not everywhere: can marine compartments shape phytoplankton assemblage? *Proc. Royal Soc. B* 286. <https://doi.org/10.1098/rspb.2019.1890>, 1914.
- Stevenson, J., 2014. Ecological assessments with algae: a review and synthesis. *J. Phycol.* 50 (3), 437–461. <https://doi.org/10.1111/jpy.12189>.
- Strunecký, O., Ivanova, A.P., Mareš, J., 2023. An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. *J. Phycol.* 59 (1), 12–51. <https://doi.org/10.1111/jpy.13304>.
- Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S., Jeppesen, E., Conde-Porcuna, J.M., Schwenk, K., Zwart, G., Degans, H., 2007. The power of species sorting: local factors drive bacterial community composition over a wide range of spatial scales. *Proc. Natl Acad. Sci.* 104 (51), 20404–20409. <https://doi.org/10.1038/ismej.2010.207>.
- Vellend, M., 2010. Conceptual synthesis in community ecology. *Q. Rev. Biol.* 85 (2), 183–206. <https://doi.org/10.1086/652373>.
- Wachnicka A., Jones P., Marchio D., Marley D. (2022). Appendix 8B-3: results from water year 2021 Expanded Lake Okeechobee phytoplankton and water quality monitoring program. *In* 2022 South Florida Environmental Report – Volume 1, South Florida Water Management District, West Palm Beach, FL. South Florida Environmental Report and Other Publications | South Florida Water Management District.
- Wachnicka A., Jones P., Marchio D., Marley D. (2023). Appendix 8B-2: results from water year 2022 expanded Lake Okeechobee phytoplankton and water quality monitoring program. *In* 2023 South Florida Environmental Report – Volume 1, South Florida Water Management District, West Palm Beach, FL. South Florida Environmental Report and Other Publications | South Florida Water Management District.
- Watson, S.B., Whitton, B.A., Higgins, S.N., Paerl, H.W., Brooks, B.W., Wehr, J.D., 2015. Harmful algal blooms. *In*: *Freshwater Algae of North America*. Academic Press, pp. 873–920. <https://doi.org/10.1016/B978-0-12-385876-4.00020-7>.
- Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T., Bentivenga, S., 2011. Advances, challenges and a developing synthesis of ecological community assembly theory. *Philos. Trans. Royal Soc. B: Biol. Sci.* 366, 2403–2413. <https://doi.org/10.1098/rstb.2011.0056>, 1576.
- Wetzel, R.G., 2001. *The Inorganic Carbon Complex In Limnology: Lake and River Ecosystems*, 3rd Ed, Chapter 11. Academic Press.
- Wyatt, T., 2014. Margalef's mandala and phytoplankton bloom strategies. *Deep Sea Res. Part II: Top. Stud. Oceanogr.* 101, 32–49. <https://doi.org/10.1016/j.dsr2.2012.12.006>.
- Zhang, J., Welch, Z., 2018. Chapter 8B: lake Okeechobee Watershed Research and Water Quality Monitoring Results and Activities. 2018 South Florida Environmental Report – Volume I, South Florida Water Management District. South Florida Environmental Report and Other Publications | South Florida Water Management District, West Palm Beach, FL.
- Zhang, J., Jone, P., Betts, A., 2022. Chapter 8B: lake Okeechobee Watershed Protection Plan Annual Progress Report. 2022 South Florida Environmental Report – Volume I, South Florida Water Management District. South Florida Environmental Report and Other Publications | South Florida Water Management District, West Palm Beach, FL.
- Zhou, J., Ning, D., 2017. Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81 (4). <https://doi.org/10.1128/mmlbr.00002-17>.