

OPEN

Measurement of Cyanobacterial Bloom Magnitude using Satellite Remote Sensing

Sachidananda Mishra^{1,2*}, Richard P. Stumpf², Blake A. Schaeffer³, P. Jeremy Werdell⁴, Keith A. Loftin⁵ & Andrew Meredith^{1,2}

Cyanobacterial harmful algal blooms (cyanoHABs) are a serious environmental, water quality and public health issue worldwide because of their ability to form dense biomass and produce toxins. Models and algorithms have been developed to detect and quantify cyanoHABs biomass using remotely sensed data but not for quantifying bloom magnitude, information that would guide water quality management decisions. We propose a method to quantify seasonal and annual cyanoHAB magnitude in lakes and reservoirs. The magnitude is the spatiotemporal mean of weekly or biweekly maximum cyanobacteria biomass for the season or year. CyanoHAB biomass is quantified using a standard reflectance spectral shape-based algorithm that uses data from Medium Resolution Imaging Spectrometer (MERIS). We demonstrate the method to quantify annual and seasonal cyanoHAB magnitude in Florida and Ohio (USA) respectively during 2003–2011 and rank the lakes based on median magnitude over the study period. The new method can be applied to Sentinel-3 Ocean Land Color Imager (OLCI) data for assessment of cyanoHABs and the change over time, even with issues such as variable data acquisition frequency or sensor calibration uncertainties between satellites. CyanoHAB magnitude can support monitoring and management decision-making for recreational and drinking water sources.

Cyanobacterial harmful algal blooms (cyanoHABs) are a serious environmental, water quality and public health issue worldwide because of their ability to form dense biomass and scum and to produce toxins such as neurotoxins (anatoxin-a), hepatotoxins (microcystins), and cytotoxins (cylindrospermospin)¹. Cyanotoxins are capable of causing a wide variety of adverse human health issues including gastrointestinal distress, dermatitis, liver failure, or even death of domestic and livestock animals when they are exposed to water with toxins from intense cyanoHABs^{2,3}. CyanoHABs are considered to be increasing globally over the past few decades although observations are quite limited^{4,5}. The frequency and magnitude of these blooms are expected to worsen in the future with increased surface water temperatures and vertical stratification⁶. In addition, cyanoHAB growth and intensity are known to be affected by weather-driven environmental and anthropogenic factors, such as shifts in rainfall patterns with climatology and changing agricultural practices^{7–9}.

In order to reduce the risk of exposure to cyanotoxins, more frequent water quality assessments are needed to monitor the status and historical trends of cyanoHABs in inland lakes. This information is needed at regional as well as national scales, including lakes and reservoirs designated as drinking water sources and recreational water bodies. In the United States, the National Lakes Assessment (NLA)¹⁰ was designed to provide national estimates of lake conditions with biological, chemical, physical, and recreational/human health indicators. Through this program, assessments of ecologically representative samples of U.S. lakes >1 ha in size are conducted every five years and were last completed in 2012 (2017 is pending). However, a yearly assessment of individual lakes will aid in developing management strategies.

Lake assessments using traditional field sampling methods (routine laboratory analysis for measurements of phytoplankton pigment and cell concentration, biovolume, and biomass) are expensive, time-consuming, and often not feasible to carry out in multiple waterbodies or across multiple states. However, satellite-based remote sensing methods

¹Consolidated Safety Services Inc., Fairfax, VA, 22030, USA. ²National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science, Silver Spring, MD, 20910, USA. ³Office of Research and Development, U.S. Environmental Protection Agency, Durham, NC, 27709, USA. ⁴Ocean Ecology Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA. ⁵U.S. Geological Survey, Organic Chemistry Research Laboratory, Kansas Water Science Center, Lawrence, KS, 66049, USA. *email: sachi.mishra@noaa.gov

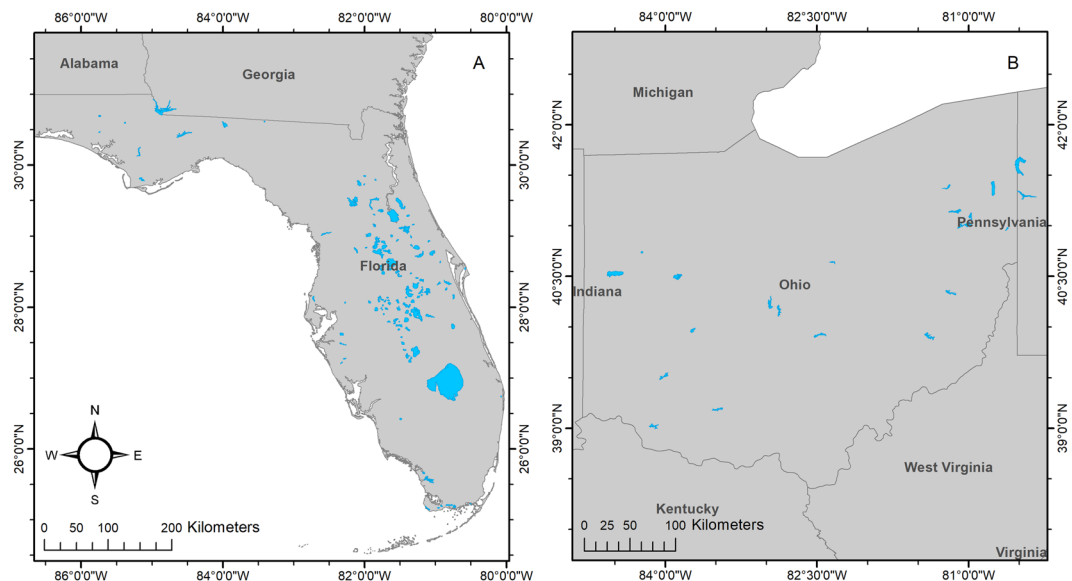


Figure 1. Map of the study region showing the location of lakes in (A) Florida and (B) Ohio. In total, 135 lakes in Florida and 21 lakes in Ohio, were resolvable with the full resolution MERIS data and are used in this study. Land and lakes are shown in gray and blue colors respectively.

can be used to monitor the current status of cyanoHABs in numerous larger water bodies on a routine basis^{11,12} and to retrospectively assess the historical status of these water bodies¹³. Among cyanoHAB assessment studies using remote sensing methods, a few have looked at several lakes at a time^{14–16}, but a majority of the efforts have focused on one body of water such as the Baltic Sea¹⁷, Lake Balaton, Hungary¹⁸, the Caspian Sea¹⁹, Lake Taihu, China^{20,21}, and Lake Erie^{11,12,22}.

Researchers often quantify cyanobacteria as biomass measured directly or by surrogate approaches such as with concentrations of chlorophyll-*a* (Chl-*a*) (mg m^{-3})^{17,22–24}, phycocyanin (PC) (mg m^{-3})^{25–27}, or as cell concentrations^{12,28}. Recent studies on multiple lakes have focused on quantifying cyanoHAB spatial extent (km^2)²⁹ and temporal frequency (% of observations)¹⁵. These studies have started to address management questions of change in spatial and temporal cyanobacteria bloom dynamics over time. Resource managers have limited resources for assessment and monitoring of lakes for public and environmental health. Knowing which lakes have severe or worsening blooms, and which do not, allows the managers to determine viable lake management strategies. An indicator of bloom magnitude would provide a key addition to the previous metrics by characterizing algal bloom biomass for an observational time (season or year), thereby highlighting the annual scale of the blooms. In this study, we introduce a metric that focuses on the magnitude of cyanoHABs in lakes and other inland water bodies. We define bloom magnitude as the spatiotemporal mean of the sequence of 7 or 14-day composites of maximum bloom biomass collected over the bloom season. In order to make the bloom magnitude comparable across lakes of different size, magnitude is further normalized to lake surface area. We also use rank as a key metric to compensate for significant changes in data frequency, and ultimately other factors like differences between satellites or calibration drift.

Nine years of European Space Agency (ESA) Medium Resolution Imaging Spectrometer (MERIS) data (2003–2011) with a nominal pixel resolution of 300×300 m was used to estimate the annual magnitude of cyanoHABs in Florida and Ohio lakes. The primary objectives of this study were: (1) to develop a method for estimating annual cyanoHAB magnitude in inland lakes and reservoirs using MERIS observations; and (2) to generate a baseline cyanoHAB magnitude dataset during the 2003–2011 time-period in Florida and Ohio as a case study. The methods are applicable to the Sentinel-3 Ocean Land Color Imager (OLCI), the replacement for MERIS, which was first launched on Sentinel-3A in 2016.

Data and Methods

Study area. We selected the states of Florida and Ohio as our study areas for three primary reasons (Fig. 1):

1. **To examine cyanoHAB magnitudes in lakes that are known to have cyanoHAB related water quality issues.** These two states also have a significant number of lakes that are resolvable in MERIS/OLCI data. Many lakes in the Coastal Plains ecoregion, which includes Florida, are known to have cyanoHAB issues, with 34% of lakes known to be hypereutrophic by the NLA in 2007³⁰. The 2007 NLA also reported that 43% of lakes in Florida had microcystin present³¹. Similarly, cyanoHABs are a common water quality issue in the Temperate Plains ecoregion, which includes western Ohio, where 45% lakes are considered hypereutrophic³⁰. Approximately 32% of lakes in Ohio had microcystin present in 2007³¹.
2. **To consider results from lakes located in different geographic and climatic regimes.** The climate in Florida is subtropical, with hot, humid, high precipitation summers and mild, dry winters. In contrast, Ohio has a temperate climate with cold winters, hot and humid summers, and year-round moderate precipitation³².

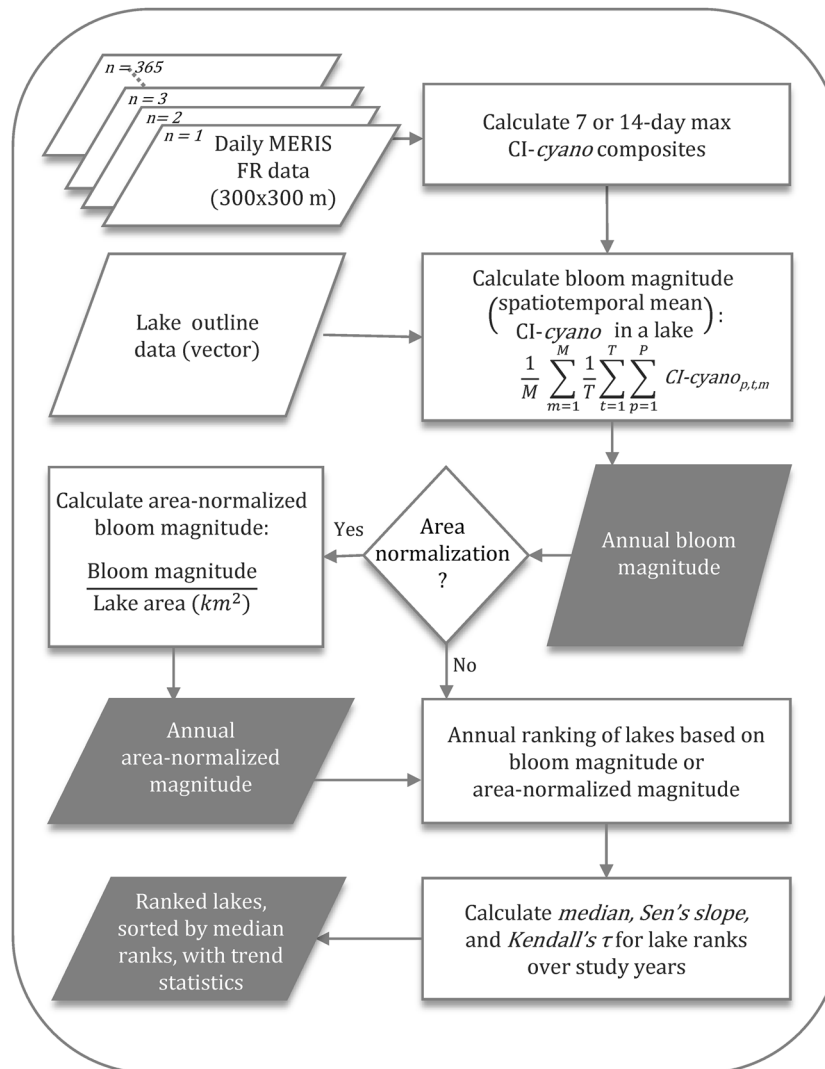


Figure 2. Schematic diagram of data processing and workflow for calculating bloom magnitude and area-normalized magnitude. Output stages in the workflow are shaded gray.

- To assess the impacts of differences in data coverage in each location.** MERIS full resolution (FR) data collection frequency prior to 2008 was inconsistent. The temporal frequency of MERIS FR data over Ohio is higher than that over Florida during this time period. The consideration of two states with different temporal data coverage will illustrate the effect of reduced data frequency on the bloom magnitude metrics.

Figure 2 shows the steps of the data analysis and workflow carried out in this study. Individual components of the data and methods are presented below.

Lake outline data. The lakes were screened for size using polygons of lakes and water bodies from the National Hydrography Dataset Plus version 2.0 (NHDPlusV2) lake polygons dataset³³, with the condition that each selected water body should be resolved by a satellite image with 300 × 300 m pixel resolution. Lakes and other water bodies were considered resolvable if they had the equivalent of three connected non-mixed water pixels (i.e. three pixels without any land) within the NHDPlusV2 dataset. Further, all selected water bodies were screened and filtered using U.S. EPA's 2012 NLA³⁴ site evaluation guidelines (U.S. EPA, 2011). Waterbodies classified as intermittent or estuarine were excluded from further analysis based on NLA criteria (although some estuarine lakes in Florida were not identified and excluded, as discussed later). The final lake polygon layer included 135 lakes in Florida and 21 lakes in Ohio that would be resolved in FR MERIS/OLCI imagery. The surface area of resolvable lakes in Florida varied from 1.26 km² to 1427 km² with a median surface area of 5.31 km², whereas, surface areas of Ohio lakes varied from 1.98 to 53 km² with a median of 8.9 km². In Florida, Center Lake and Lake Okechobee were the smallest and largest lakes considered in this study, respectively. In Ohio, Evans Lake and Pymatuning Reservoir were the smallest and largest, respectively.

Satellite data. MERIS Level-2 (L2) datasets were processed using the satellite automated processing system (SAPS) administered by the National Oceanic and Atmospheric Administration (NOAA). SAPS incorporates l2gen, the NASA standard software for processing L2 ocean color data that is included in NASA's SeaWiFS Data Analysis Software (SeaDAS) package³⁵, and the Shuttle Radar Topography Mission (SRTM) global land mask³⁶. The l2gen-derived “rho_s” surface reflectance (ρ_s , unitless) product was projected to Universal Transverse Mercator (UTM) projection using nearest-neighbor interpolation. The rho_s process generates a reflectance that corrects for top-of-atmosphere solar irradiance and removes Rayleigh radiance and molecular absorption corrected for elevation. Clouds are masked using an albedo threshold algorithm corrected for turbid water, with the latter necessary to retain those pixels with bright reflectance from intense blooms that would otherwise be incorrectly masked as clouds. Land adjacency issues, including mixed land/water pixels, were detected using near-infrared and red-edge thresholds³⁷, thereby ensuring that the signals originating from land vegetation were flagged and excluded from further analysis.

Temporal coverage and data density for the MERIS FR data has changed over the mission. Prior to 2008, MERIS FR data sets for North America were obtained by onboard recording, which limited data acquisition. From 2002–2004, there was also competition for band width with other ENVISAT instruments, which may have also reduced acquisition frequency. In 2008, the Canadian Centre for Remote Sensing began reception of direct broadcast, which assured acquisition of most North American data³⁸. In the contiguous United States (CONUS), prior to 2008, FR data frequency was lower over the southern states, including Florida³⁹. Our analysis showed that the number of composites with near-100% missing data (inclusive of clouds and no satellite data collected) varied significantly in Florida and Ohio. The mean percentage of composites with near-100% missing data varied from 13% to 64% in Florida from 2003–2007. During the same time period, the mean percentage in Ohio varied from 12–28%. In contrast, from 2008–2011, the percentage of composites with missing data decreased to 6–12% (in Florida) and stayed about the same (14–29%) in Ohio. When the study period was restricted to 2008–2011 recreational months (May 1st to Oct 31st as per Ohio Environmental Protection Agency (Ohio EPA) recreational criteria⁴⁰), Florida lakes received slightly more data coverage as compared to Ohio lakes as shown by the means of missing data percentage in Florida (range = 1–9%) and Ohio (0–16%).

Cyanobacteria estimation algorithm. Cyanobacterial quantity was found through a combination of biomass estimation and cyanobacterial presence algorithms^{11,28}. The Cyanobacteria Index (CI) measures a proxy of Chl-*a* absorption and provides the cyanobacterial biomass^{11,18,19,29,41}. It is calculated with a spectral shape (SS) algorithm^{28,42} and is presented as:

$$SS(\lambda) = \rho_s(\lambda) - \rho_s(\lambda_-) + \{\rho_s(\lambda_-) - \rho_s(\lambda_+)\} \frac{(\lambda - \lambda_-)}{(\lambda_+ - \lambda_-)} \quad (1)$$

where ρ_s is the top of atmosphere reflectance corrected for Rayleigh radiance, λ is the central band, and λ_+ and λ_- are the adjacent reference bands. Cyanobacteria Index (CI) is calculated by centering the spectral shape at 681 nm and changing the sign of SS, or $CI = -SS$ (681).

The CI evaluates Chl-*a* absorption at 681 nm. At 681 nm, chlorophyll in eukaryotes typically fluoresces strongly, leading to increased apparent reflectance that obscures chlorophyll absorption. In cyanobacteria, however, chlorophyll fluorescence is much weaker⁴³, such that Chl-*a* absorption dominates the radiance signal from the water at 681 nm, causing the reflectance at 681 nm to decrease relative to 665 and 709 nm.

For more specific identification of cyanobacteria, a SS using 620, 665, and 681 nm was used to identify the presence of PC, a characteristic pigment in this taxonomic division with identifying features in this spectral region^{25,41,44}. (Estimated PC concentration is not used as it has several issues, in particular, it is not a consistent estimator of cyanobacterial biomass, and the more usable PC concentration algorithms require robust atmospheric corrections⁴⁵ greatly limiting data availability.) In this case, a conditional negative SS (665) value indicates the presence of PC. Inclusion of 620 nm, which is the absorption peak of PC, a characteristic photopigment in cyanobacteria, reduces the false detection issue. In the case of cyanobacteria, SS (665) turns negative due to lower reflectance at 620 nm band and is used as an exclusion criterion to select only cyanobacteria. This spectral shape condition has also been applied by⁴⁶ (their Eq. 3–4) to separate cyanobacteria from other algal blooms in African lakes. The CI product, when SS (665) is negative, is termed as CI-*ciano* and was used to estimate cyanobacteria biomass in this research.

For purposes of setting risk thresholds, we applied a relationship between CI and cyanobacterial cell concentration of 10^8 cells mL⁻¹ per 1 unit of CI-*ciano*¹¹. While the relationship¹¹ was developed for *Microcystis* (so we term the value as “*Microcystis*-equivalent cells”), it was validated by^{15,41} for unspecified total cell concentrations of cyanobacteria in eight U.S. eastern states across New England (Connecticut, Massachusetts, Maine, New Hampshire, Rhode Island, and Vermont), Ohio, and Florida. Mean absolute percent error (MAPE) of 28.6% was reported between field-measured cyanobacteria biomass data (cells mL⁻¹) and satellite-derived cell biomass¹⁵. This CI algorithm has also been confirmed for detecting cyanobacterial blooms and estimating biomass (cells mL⁻¹) in other areas^{18,19}.

Fourteen-day and seven-day maximum temporal CI-*ciano* composites were computed for 2003–2007 and 2008–2011 MERIS FR time series data. Fourteen-day intervals were chosen for compositing the 2003–2007 time series in order to address the FR data gaps as discussed earlier. Maximum temporal composites, that is, reporting of the maximum value retrieved during the 14-day or 7-day window, serve two purposes. First, many cyanoHAB species such as *Microcystis*, *Aphanizomenon*, and *Dolichospermum*, have buoyancy control mechanisms and will typically float to the surface in the day when vertical water column mixing is weak⁴⁷. It is expected that over a 14-day or 7-day window, cyanobacteria would be near the surface on one or more days to be captured in the satellite data^{12,47}. In addition, compositing reduces the amount of missing data, particularly due to clouds and

sun-glint. The set of 14-day composites helped to reduce the under-sampling bias before 2008 (that was caused by reduced frequency in FR data acquisition discussed earlier).

Annual or seasonal bloom magnitude metric. CyanoHAB magnitude is intended to capture the combination of two key aspects of an algal bloom: the amount and duration of the cyanoHAB biomass. The annual/seasonal bloom magnitude is the mean biomass of the 14-day or 7-day maxima found in the lake over a year/season and mathematically expressed as:

$$\text{Bloom magnitude} = \frac{1}{M} \sum_{m=1}^M \frac{1}{T} \sum_{t=1}^T \sum_{p=1}^P \text{CI-cyano}_{p,t,m} \quad (2)$$

where, the indices P and T represent, respectively, the number of valid pixels with detectable CI-cyano in a lake, and the number of composite (time) sequences in each month (e.g. two composites in 2003 and four in 2011). Index M represents the number of months in a season/annual study period. Bloom magnitude was expressed in CI units, which is dimensionless. As noted above, CI-cyano can be converted to *Microcystis-equivalent* cells by multiplying by the factor 10^8 (cells mL⁻¹)^{11,15,41}, to provide a more intuitive biomass metric. In order to be able to compare bloom magnitude across lakes with different surface area, we normalize bloom magnitude by lake surface area as below:

$$\text{Area-normalized bloom magnitude} = \frac{\text{Bloom magnitude (dl)}}{\text{Lake Surface Area (km}^2\text{)}} \quad (3)$$

Lake surface area in Eq. (3), as detectable by satellite images, was estimated by identifying all pixels inside a lake polygon vector layer that were classified as water during 2008–2011. The number of water pixels was converted to area by multiplying it by area of one MERIS FR pixel (0.09 km²). Note that the estimated surface area may change over time due to seasonal precipitation and evapotranspiration. However, this satellite-adjusted surface area is a better representation of surface area than that available in lake databases, which often include dry lake beds and/or areas not covered by standing water. Henceforth, bloom metrics in Eqs. 2 and 3 are referred to as magnitude and area-normalized magnitude for brevity.

Based on the World Health Organization's (WHO) cell abundance threshold⁴⁸, three magnitude classes were considered for categorizing lakes as Low ($\leq 20,000$ cells mL⁻¹), Moderate ($20,000 \leq \text{cells mL}^{-1} \leq 100,000$), and High ($> 100,000$ cells mL⁻¹) exposure health risk. We estimated the area-normalized magnitudes that are equivalent to the WHO thresholds using the CI-cell abundance relation^{11,15,41} and normalizing this cell-equivalent threshold by the pixel unit area (CI threshold/0.09 km²). These area-normalized CI thresholds are ≤ 0.0022 for WHO-low, 0.0022 to 0.0111 for moderate, and > 0.0111 for high. We have also added 'Very High' (V.High) category when estimated cyanobacteria concentration and area-normalized magnitude exceeded 1,000,000 cell mL⁻¹ and 0.111, respectively.

Ranking of lakes based on the area-normalized magnitude. Satellite data frequency from 2003–2007 was not spatially homogeneous across CONUS as discussed earlier. Varying MERIS FR data frequency in the CONUS adds a bias in the MERIS time series data which, in turn, can bias comparisons of lakes across space and time. As a result, the analysis of trends in the area-normalized magnitude data across 2003–2011 could produce misleading results^{15,29}. To address this issue, the lakes were ranked based on their seasonal or annual area-normalized magnitude with rank 1 assigned to the lake with the greatest area-normalized magnitude in a specific year. When more than one lake had the same area-normalized magnitude level, the minimum possible rank was assigned to all lakes in the group. To summarize across years, each lake's median rank for the observational period was used. While ranks reduce the information on absolute bloom impact that is found in magnitude, they offer a key advantage over magnitude by allowing us to compare lakes between years, even when differences in data frequency biases the magnitudes. Ranks might still be region-specific, depending on differences in data coverage. For example, magnitude scores cannot be compared (ranked) between regions (states) having widely varying data sampling, because the lakes with sufficiently higher data frequency might appear to have more severe blooms than the lakes with reduced frequency. Lakes in regions (states) with similar data coverage can be compared with minimal bias.

We used a non-parametric statistic, Theil-Sen's slope estimator⁴⁹ for assessing trends in the ranks of cyanoHAB magnitude in a lake over time, with Kendall's τ ⁵⁰ for the strength of the trend. Theil-Sen's slope is estimated as the median of the set of slopes in the ranked and paired data. Theil-Sen's estimator for slope makes no assumptions about error distribution and provides an unbiased estimator of trend⁵¹. Theil-Sen's slope was expressed in the units of ranks yr⁻¹ and interpreted as the number of ranks increased or decreased over time for a lake in question. A negative trend (toward the rank of 1) indicates that the lake is getting relatively worse. Kendall's τ , by determining the concordance between all pairs of two ranked variables indicates the strength of a monotonic trend⁵¹. Kendall's τ varies between -1 and $+1$, where a positive τ indicates that the ranks of both variables increase together, and a negative τ indicates that as the rank of one variable increases the other decreases. As we have a slope for direction, we report absolute value of τ . A τ value of $[0.2]$ to $[0.5]$ indicates a moderate trend and $> [0.5]$ indicates a strong trend.

Results

Annual area-normalized magnitude in Florida lakes. CyanoHAB magnitude is calculated as an annual mean of 7- or 14-day maximum accumulation of CI in the lakes in Florida over the observation period. Therefore, large lakes are more likely to have a higher accumulation of biomass as compared to lakes with smaller surface area (Fig. 3). Normalization of magnitude by lake surface area removes the influence of lake size from the metric

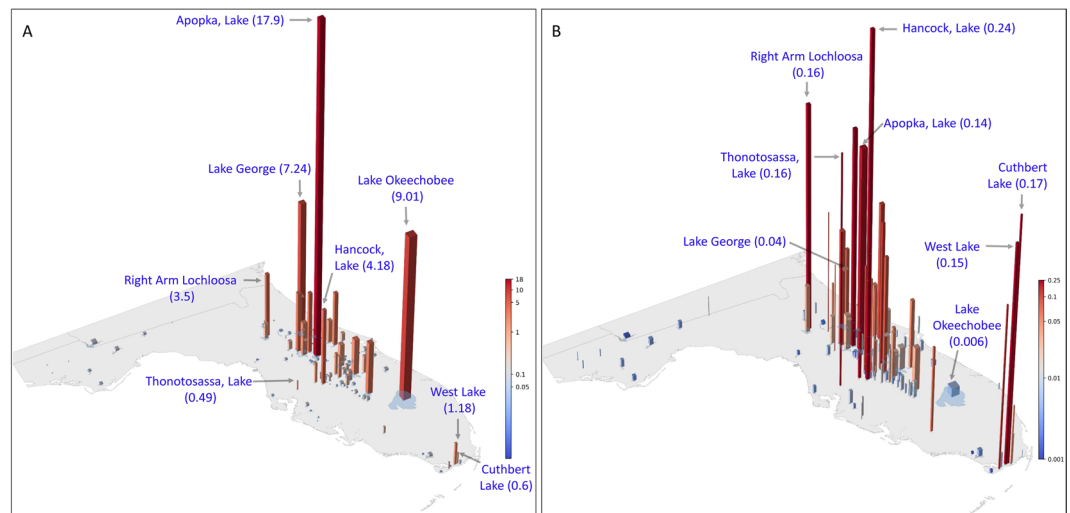


Figure 3. (A). Algal bloom magnitude in Florida lakes in 2011 before normalization and (B) after normalization by lake surface area. Area-normalized magnitude (km^{-2}) of selected lakes provided as part of bar labels in parenthesis. Bar height and color are proportional to annual bloom magnitude (in A) and annual area-normalized magnitude (in B). The width of the bars is proportional to the lake surface area. Note that the color bars are log-scaled.

and allows comparison of cyanobacteria area-normalized magnitude among lakes of different sizes. Results from 2011 are shown in Fig. 3, which highlights the impact of normalization of annual bloom magnitude using the lake surface area. Before normalization, Lake Apopka, Lake Okeechobee, Lake George, Hancock Lake, and Right Arm Lochloosa Lake ranked 1st to 5th respectively in 2011. Lake Apopka, which has the fourth largest surface area, but year-round blooms, ranked first, whereas Lake Okeechobee, with the largest surface area but less frequent blooms, ranked second in terms of cyanoHAB magnitude in 2011 (Fig. 3A). When the mean annual biomass estimates were normalized by lake surface area, Hancock Lake, Cuthbert Lake, Thonotosassa Lake, Right Arm Lochloosa, and West Lake ranked in the top five positions respectively (Fig. 3B). These results highlight that: 1) in 2011, while Lake Okeechobee had an intense bloom that covered only a portion of the lake, most of Hancock Lake was impacted; and 2) the average area of Hancock Lake was more severely affected by cyanobacteria than the average area in Lake Okeechobee.

Area-normalized magnitude rankings for all lakes during the study period were analyzed to identify those lakes with the most severe cyanobacteria blooms needing attention (Fig. 4). Lakes are ordered by their median rank from 2003–2011 in ascending order. Hancock Lake, Lake Apopka, Lake Dora/Beauclair/Carlton, Cuthbert Lake, and West Lake are the top five lakes for annual area-normalized bloom magnitude in Florida (Fig. 4, Table 1). The three top-ranked Florida lakes exhibited little variation over time (Hancock Lake, Lake Apopka, and Lake Dora/Beauclair/Carlton) (Fig. 4). Obviously, for a change in rank to occur one lake changes to a lower rank, another lake must move to a higher rank. The changes are not evenly distributed. Six lakes showed large declines: Right Arm Lochloosa and Lake George declined at ~ 6 ranks yr^{-1} , and Clinch, Hamilton, Panasoffkee, and Buffum lakes declined at 3 to 5 ranks yr^{-1} . Of these lakes, the decline was highly consistent ($\tau > 0.5$) for Lochloosa, Panasoffkee, George, and Buffum, and moderately consistent ($\tau = 0.3$ – 0.5) for the others. In contrast, only three lakes (Seven Palms, Leonore, Konomac) had consistent ($\tau > 0.3$) and large increases in rank (better) changing at $+6$ – 7 ranks yr^{-1} . Overall, area-normalized magnitude improved in Dias Lake, Monroe Lake, Deaton Lake, and Lake Griffin, which resulted in their lake ranks increasing at the rate of $\sim +2$ – 3 ranks yr^{-1} . Several of the lakes at the southern tip of Florida (e.g., Cuthbert, West), while in the Everglades, are actually estuarine with salinity influenced by Florida Bay. These are noted by asterisks (*) in the table.

In order to infer potential exposure risk of cyanoHABs in Florida lakes, we determined the number of lakes where cyanoHAB abundance exceeded the specific WHO risk thresholds of Low, Moderate, High, and V.High levels. A recreational Low WHO limit indicates lakes that are unlikely to have a management concern⁴⁸. Out of 135 lakes, 34–58 (range represents the number of lakes in a specific year) lakes were assigned to the High category with area-normalized magnitude over the study period (Fig. 5A). 2011 witnessed the maximum number of lakes in the high category ($n = 58$, $\sim 43\%$ of all lakes in Florida). Similarly, the area-normalized magnitude for 66–90 lakes were in the Moderate category and 1–11 lakes were in the Low category. In 2010, all lakes were in Moderate and High categories. Area-normalized magnitude fell into the V.High range in 10 lakes (Right Arm Lochloosa Lake, Bear Lake, Parker Lake, Apopka Lake, Thonotosassa Lake, West Lake, Hancock Lake, Cuthbert Lake, Lake Griffin, and Lake Dora/Beauclair/Carlton, not in order) over the study period. In 2008 and 2011, eight (excluding Lake Griffin and Thonotosassa Lake) and nine (excluding Lake Griffin) out of those 10 lakes were categorized as V.High.

Without normalization by lake surface area, median magnitude data in Florida looked completely skewed (min = 0.007 CI, max = 15.49 CI, median = 0.05 CI), where about 90% of lakes were below annual bloom magnitude of one CI. The top 10% (13) having the highest median bloom magnitude had bloom magnitudes

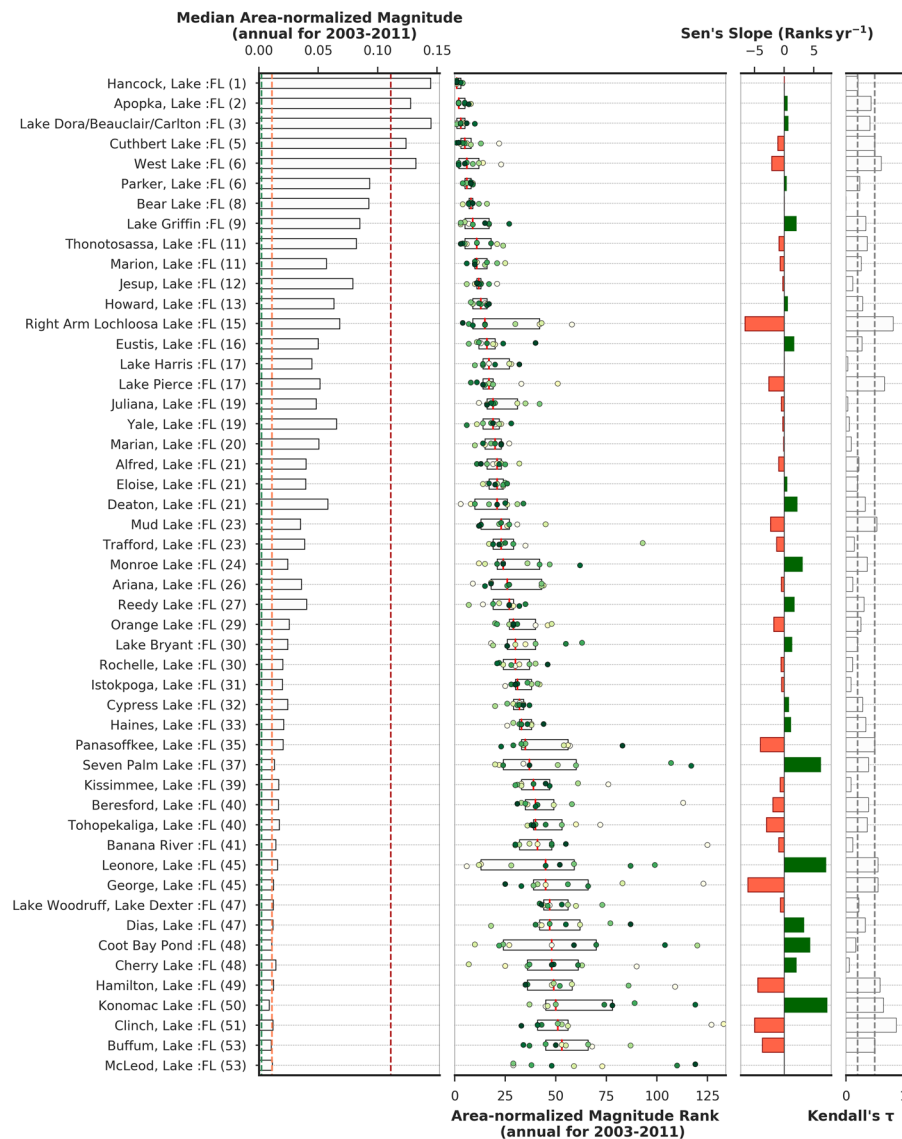


Figure 4. First panel: Annual area-normalized magnitude (km^{-2}) in the top 50 Florida lakes. Green, orange, and red dotted lines represent equivalent WHO thresholds of 20,000, 100,000, and 1,000,000 cells mL^{-1} limits; second panel: the inter-quartile range of area-normalized magnitude ranks in top-ranked Florida lakes over 2003–2011 ordered by their median rank over the 9 year period. Median values or ranks are highlighted in red vertical lines inside the box. Annual area-normalized magnitude rank data points are overlaid on inter-quartile boxes to highlight the variation, where the color of the scatter points indicate data year (tan: start year and deep green: end year). Third panel: green/red bar plot shows Sen's slope (trends in rank change) during 2003–2011. Green/red color represents positive/negative trend meaning area-normalized magnitude for a lake is decreasing/increasing over time. Fourth panel: bars show Kendall's τ (absolute values) representing consistency in rank-change trend over time. Dotted lines in Kendall's τ plot mark the τ values at 0.2 and 0.5.

between 1.11–15.5 CI. Of these, six are also in the top 10% of the area-normalized magnitude (Apopka, Griffin, Dora/Beaulair/Carlton, Hancock, and Jesup). Some of these lakes with large bloom magnitude, e.g., Istokpoga, Kissimmee, and Tohopekaliga, ranked in the 2nd quartile (27 and below) when area-averaged.

Seasonal area-normalized magnitude in Ohio lakes. CyanoHAB magnitude analyses in Ohio lakes during the recreational season (May 1st to Oct 31st) showed that Grand Lake St. Marys (Rank 1, IQR = 0), Buckeye Lake (Rank 2, IQR = 0.25), and Indian Lake (Rank 3, IQR = 0.25) are the top three lakes by median seasonal area-normalized magnitude rank in Ohio from 2003–2011 (Figs. 6–7, Table 2). These top three lakes maintained the ranks consistently and were in WHO High category with area-normalized magnitude > 0.011 confirming severe CyanoHAB issue over time. Unlike the case for Florida, the variance in rank change for Ohio lakes during the study period is negligible, indicating that lakes maintained their area-normalized magnitude ranks consistently every year. However, there were substantial differences in cyanoHAB magnitude among different Ohio lakes. For instance, the median area-normalized magnitude in Grand Lake St. Marys (0.27 km^{-2}) was ~12 times

Lake Name	Surface Area ⁺ (km ²)	State	Median Bloom Magnitude (dl)	Median Area-normalized Magnitude (km ⁻²)	Median Rank	Sen's Slope (Ranks yr ⁻¹)	Kendall's τ
Hancock, Lake	17.01	FL	2.457	0.144	1	0.00	-0.20
Apopka, Lake	121.5	FL	15.497	0.128	2	0.46	0.43
Lake Dora/Beauclair/Carlton	21.24	FL	3.077	0.145	3	0.65	0.41
Cuthbert Lake*	3.42	FL	0.423	0.124	5	-1.10	-0.50
West Lake*	7.47	FL	0.987	0.132	6	-2.13	-0.61
Parker, Lake	7.65	FL	0.712	0.093	6	0.35	0.24
Bear Lake*	3.15	FL	0.291	0.092	8	0.00	0.00
Lake Griffin	38.88	FL	3.297	0.085	9	2.00	0.34
Thonotosassa, Lake	2.97	FL	0.243	0.082	11	-0.88	-0.37
Marion, Lake	10.71	FL	0.607	0.057	11	-0.70	-0.26
Jesup, Lake	29.61	FL	2.338	0.079	12	-0.27	-0.11
Howard, Lake	2.16	FL	0.137	0.063	13	0.55	0.29
Right Arm Lochloosa Lake	21.51	FL	1.461	0.068	15	-6.63	-0.82
Eustis, Lake	29.79	FL	1.485	0.050	16	1.62	0.28
Lake Harris	71.1	FL	3.166	0.045	17	0.00	0.03
Lake Pierce	13.77	FL	0.707	0.051	17	-2.58	-0.67
Juliana, Lake	3.51	FL	0.169	0.048	19	-0.50	-0.03
Yale, Lake	14.94	FL	0.974	0.065	19	-0.26	-0.06
Marian, Lake	18.27	FL	0.920	0.050	20	-0.08	-0.09
Alfred, Lake	2.43	FL	0.096	0.040	21	-0.95	-0.22
Eloise, Lake	4.14	FL	0.163	0.039	21	0.45	0.20
Deaton, Lake	1.8	FL	0.104	0.058	21	2.13	0.33
Mud Lake	1.71	FL	0.060	0.035	23	-2.31	-0.54
Trafford, Lake	5.49	FL	0.211	0.038	23	-1.27	-0.14
Monroe Lake*	2.7	FL	0.065	0.024	24	3.00	0.37
Ariana, Lake	3.69	FL	0.132	0.036	26	-0.50	-0.11
Reedy Lake	13.41	FL	0.538	0.040	27	1.65	0.31
Orange Lake	23.58	FL	0.602	0.026	29	-1.79	-0.25
Lake Bryant	3.78	FL	0.091	0.024	30	1.27	0.20
Rochelle, Lake	1.98	FL	0.040	0.020	30	-0.55	-0.11
Istokpoga, Lake	90.99	FL	1.797	0.020	31	-0.45	-0.08
Cypress Lake	11.61	FL	0.282	0.024	32	0.73	0.29
Haines, Lake	2.52	FL	0.052	0.021	33	1.08	0.34
Panasoffkee, Lake	9.99	FL	0.204	0.020	35	-4.00	-0.50
Seven Palm Lake*	5.04	FL	0.065	0.013	37	6.08	0.39
Kissimmee, Lake	118.8	FL	1.969	0.017	39	-0.71	-0.08
Beresford, Lake	2.7	FL	0.044	0.016	40	-1.90	-0.39
Tohopekaliga, Lake	64.71	FL	1.113	0.017	40	-3.00	-0.37
Banana River	3.24	FL	0.046	0.014	41	-0.95	-0.11
Leonore, Lake	1.35	FL	0.021	0.016	45	7.00	0.56
George, Lake	172.71	FL	2.092	0.012	45	-6.13	-0.56
Lake Woodruff, Lake Dexter	17.37	FL	0.208	0.012	47	-0.63	-0.22
Dias, Lake	2.52	FL	0.029	0.012	47	3.25	0.33
Coot Bay Pond*	4.14	FL	0.044	0.011	48	4.29	0.17
Cherry Lake	1.62	FL	0.023	0.014	48	2.00	0.06
Hamilton, Lake	8.1	FL	0.100	0.012	49	-4.47	-0.59
Konomac Lake	4.05	FL	0.035	0.009	50	7.20	0.65
Clinch, Lake	4.14	FL	0.049	0.012	51	-5.00	-0.87
Buffum, Lake	4.95	FL	0.051	0.010	53	-3.67	-0.50
McLeod, Lake	1.53	FL	0.017	0.011	53.5	6.45	0.40

Table 1. Summary of median annual bloom magnitude, median annual area-normalized magnitude and their ranks of top 50 lakes in Florida (2003–2011). +Lake area represented by MERIS pixels. * Indicates estuarine lake at the southern tip of Florida.

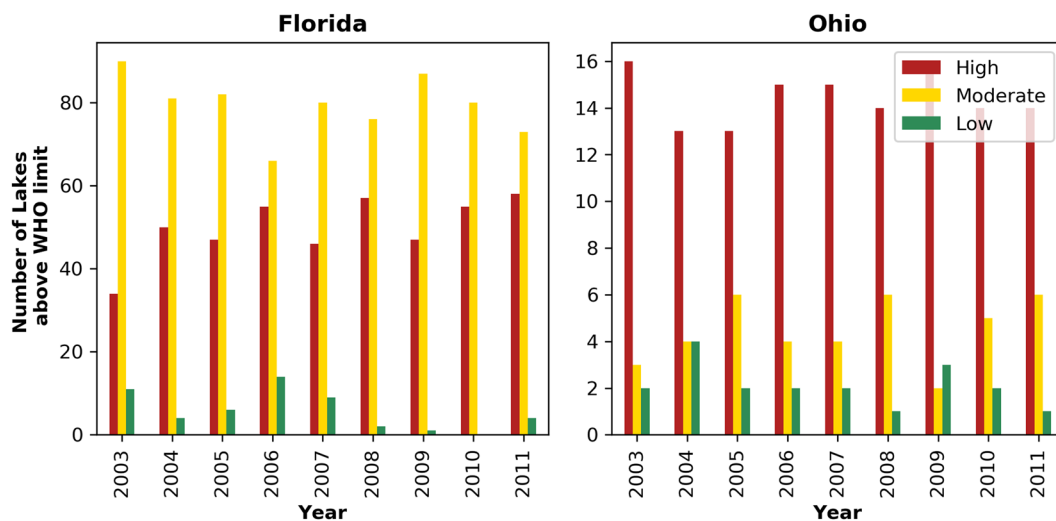


Figure 5. The number of lakes in Florida and Ohio classified as high, moderate, and low bloom categories based on recreational WHO cyanobacterial cell density limits.

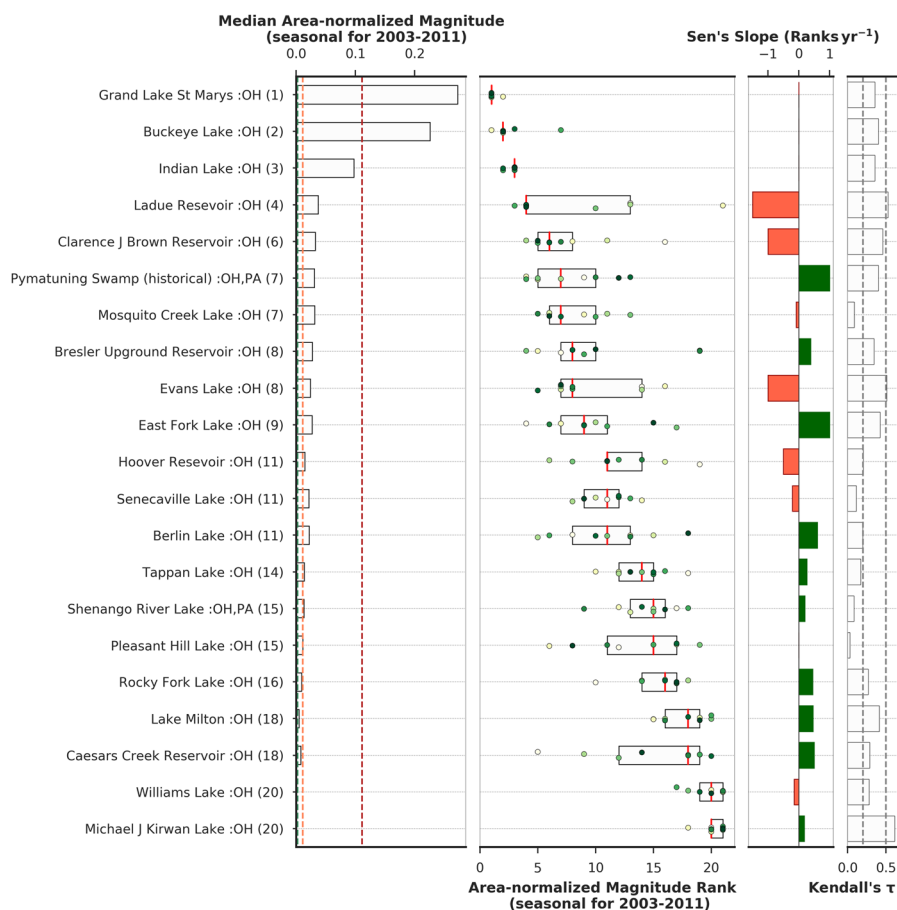


Figure 6. First panel: Seasonal area-normalized magnitude (km^{-2}) in Ohio lakes. Green, orange, and red dotted lines represent equivalent WHO thresholds of 20,000, 100,000, and 1,000,000 cells mL^{-1} limits; second panel: the inter-quartile range of area-normalized magnitude ranks in top-ranked Ohio lakes over 2003–2011 ordered by their median rank over the 9 year period. Median values or ranks are highlighted in red vertical lines inside the box. Annual area-normalized magnitude rank data points are overlaid on inter-quartile boxes to highlight the variation, where the color of the scatter points indicate data year (tan: start year and deep green: end year). Third panel: green/red bar plot shows Sen's slope (trends in rank change) during 2003–2011. Green/red color represents positive/negative trend meaning area-normalized magnitude for a lake is decreasing/increasing over time. Fourth panel: bars show Kendall's τ (absolute values) representing consistency in rank-change trend over time. Dotted lines in Kendall's τ plot mark the τ values at 0.2 and 0.5.

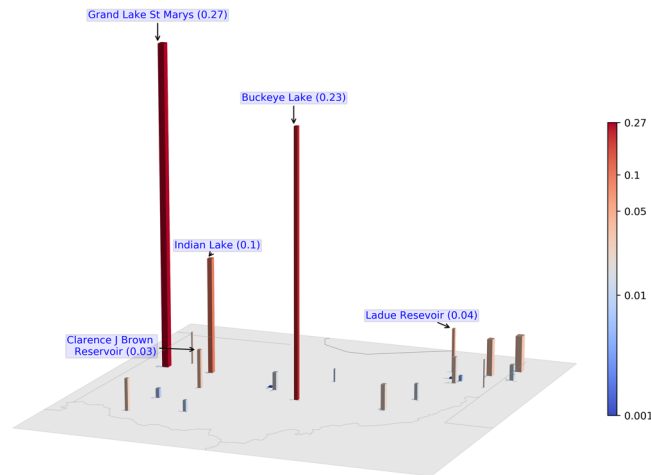


Figure 7. Bar plot showing median area-normalized magnitude (km^{-2}) in Ohio lakes during the recreational season over the study period (2003–2011). Width of bars is proportional to lake surface area, height and color of the bars are proportional to the median annual area-normalized magnitude (km^{-2}). Median values of the top five lakes are provided as part of bar labels inside the parenthesis.

Lake Name	Surface Area ⁺ (km^2)	State	Median Bloom Magnitude (dl)	Median Area-normalized Magnitude (km^{-2})	Median Rank	Sen's Slope (Ranks yr^{-1})	Kendall's τ
Grand Lake St. Marys	47.97	OH	13.07	0.272	1	0	−0.35
Buckeye Lake	9.27	OH	2.091	0.226	2	0	0.4
Indian Lake	16.29	OH	1.591	0.098	3	0	−0.36
Ladue Reservoir	4.77	OH	0.177	0.037	4	−1.5	−0.53
Clarence J Brown Reservoir	7.38	OH	0.24	0.032	6	−1	−0.46
Pymatuning Reservoir	53.1	OH, PA	1.636	0.031	7	1	0.4
Mosquito Creek Lake	27	OH	0.846	0.031	7	−0.08	−0.09
Bresler Upground Reservoir	2.25	OH	0.061	0.027	8	0.39	0.34
Evans Lake	1.98	OH	0.047	0.024	8	−1	−0.51
East Fork Lake	6.57	OH	0.176	0.027	9	1	0.42
Hoover Reservoir	9.72	OH	0.145	0.015	11	−0.5	−0.2
Senecaville Lake	11.61	OH	0.248	0.021	11	−0.21	−0.11
Berlin Lake	8.91	OH	0.197	0.022	11	0.61	0.2
Tappan Lake	6.75	OH	0.092	0.014	14	0.27	0.17
Shenango River Lake	11.07	OH, PA	0.146	0.013	15	0.2	0.08
Pleasant Hill Lake	2.43	OH	0.027	0.011	15	0	−0.03
Rocky Fork Lake	6.12	OH	0.057	0.009	16	0.45	0.27
Lake Milton	5.94	OH	0.028	0.005	18	0.46	0.41
Caesars Creek Reservoir	9.09	OH	0.071	0.008	18	0.5	0.29
Williams Lake	9.72	OH	0.008	0.001	20	−0.15	−0.28
Michael J Kirwan Lake	7.92	OH	0.007	0.001	20	0.18	0.61

Table 2. Summary of median seasonal bloom magnitude, median seasonal area-normalized magnitude and their ranks and statistics (Sen's Slope and Kendall's τ) of lakes in Ohio (2003–2011). Seasonal data (May 1st–Oct 31st). ⁺Lake area represented by MERIS FR pixels.

higher than 50% of the lakes in Ohio (Table 2). Out of 21 lakes in Ohio, there are three lakes – Ladue Reservoir, Clarence J. Brown Reservoir, and Evans Lake, where the relative rank decreased, or the area-normalized magnitude deteriorated over time (slope = ~ 1 – 1.5 ranks yr^{-1}). These lakes showed relative decline against other lakes that showed relative improvement, especially East Fork Lake, Pymatuning reservoir, Lake Milton, Bresler Upground Reservoir, and Michael J Kirwan Lake.

Based on the WHO recreational thresholds, 13 to 16 lakes (~ 62 – 76%) had an area-normalized magnitude in the High category (Fig. 5B). Years 2003 and 2009 had the maximum number of lakes ($n = 16$) in the High magnitude category and years 2004 and 2005 had the minimum number of lakes ($n = 13$) in the High category. Similarly, 2–6 (~ 10 – 29%) and 1–4 (~ 5 – 14%) lakes were under Moderate and Low categories, respectively.

Area-normalized magnitude in Grand Lake St. Marys stayed in the V.High category throughout the study period. Similarly, the area-normalized magnitude in Buckeye Lake fell in the V.High category during 2003–2007 and in 2011. Bloom magnitude in Indian Lake reached the V.High threshold only in 2007. Median bloom magnitude data before normalization clearly highlighted four lakes: Grand Lake St. Marys (13.07 CI), Buckeye Lake (2.09 CI), Pymatuning Reservoir (1.64 CI), and Indian Lake (1.59 CI), where bloom magnitudes were one order of magnitude higher than the rest of the lakes.

Comparison of area-normalized magnitude in Florida and Ohio lakes. In order to compare Florida and Ohio, the magnitudes from the two states were combined and lakes were then ranked. In the combined dataset, the seasonal area-normalized magnitude estimation was limited to a time period of May 1st–Oct 31st for each year to exclude the snow/ice-covered winter months and to include area-normalized magnitude status during the typical cyanoHAB season. In addition, a comparison between the two states was only performed from 2008–2011 to avoid any positive bias in the MERIS FR data in Ohio relative to Florida prior to the year 2008.

Among all lakes from the two states, Grand Lake St. Marys (OH), Hancock Lake (FL), Apopka Lake (FL), Cuthbert Lake (FL), Lake Dora/Beauclair/Carlton (FL), West Lake (FL), Right Arm Lochloosa Lake (FL), Parker Lake (FL), Buckeye Lake (OH), and Thonotosassa Lake (FL) are among the top ten lakes based on median area-normalized magnitude rank as observed from 2008 to 2011 (Fig. 8, Table 3). Based on seasonal magnitude (not normalized), Apopka Lake in Florida (15.75 CI), Grand Lake St. Marys in Ohio (12.78 CI), and Lake Okeechobee in Florida (12.71 CI) are top three lakes in descending order (Fig. 9), reflecting the large magnitude blooms that occurred in these large lakes.

Discussion

Evaluating severity each year and comparing lakes provides a potentially important resource for managers, including under such regulatory frameworks as the European Union Water Framework Directive⁵² and the U.S. Clean Water Act^{53,54}. In Lake Erie, an assessment of magnitude (over a 30-day period) from satellite proved essential to the development of nutrient target strategies⁵⁵, and has also led to an annual forecast of bloom severity⁴⁵. In comparison to satellite, traditional routine monitoring is difficult and expensive. Exceptionally strong programs such as Ohio EPA's drinking water program may collect a sample each week in a water body (although for toxin only)⁵⁶. More common for a water quality monitoring program is monthly or quarterly sampling of water quality (including Chl-*a*), such as seen in Florida's several monitoring programs^{57,58}.

The results presented here capture the relative severity of cyanobacterial blooms observed in state monitoring programs. In Ohio, the concentration of microcystin toxins is the most common water quality measurement. Three of the top four lakes (Grand Lake St. Marys, Buckeye, and LaDue) consistently reported the highest microcystin concentrations of the observed lakes when Ohio EPA started sampling in 2010, and these were well above the WHO recreational risk levels ($10 \mu\text{g L}^{-1}$)¹⁵. These lakes are also currently listed as impaired due to algae and associated microcystin⁵⁶. The fourth, Indian Lake, does not have routine sampling.

In Florida, we verified the ranking using field Chl-*a* data for lakes found in both the Florida Water Atlas⁵⁸ and in the top 50 lakes from our satellite-based observations (as in Table 1). Similar to annual area-normalized magnitude estimation, we calculated annual mean Chl-*a* concentration for the study years 2003–2011 by taking the mean of monthly mean Chl-*a* concentrations for all samples from a lake available in the database. To match area-normalized magnitude, we calculated the median of annual mean Chl-*a* concentrations over the study period. We further ranked the ten lakes based on median area-normalized magnitude and median of annual mean Chl-*a* concentration over nine-year study period (Fig. 10), and found similar results across these lakes from Lake Hancock (1 in both) to Hamilton Lake (rank 10 among this set of lakes, and 49 in the larger satellite dataset).

Two studies with satellite-based methods examined Ohio and Florida lakes, and had similar conclusions about the most impacted lakes as found here, although they use different methods. Gorham *et al.*²⁷ used 10 years of MERIS data (2002–2011) to estimate PC with a semi-analytical model²⁵. Lakes were evaluated using the maximum PC concentration for the year at each pixel in each lake. This approach would rank a lake having a single day of high concentration as more severe than a lake with a slightly less severe long-duration bloom. Regardless, this approach also put Grand Lake St. Marys, Buckeye Lake, Indian Lake, and Seneca Lake as the top four, matching our result.

Clark *et al.*¹⁵ focused on bloom frequency. They calculated cyanoHAB frequency for Ohio and Florida lakes as the fraction of total pixel observations where cyanoHAB abundance exceeded the WHO's threshold of 100,000 cells mL^{-1} , and then ranked those lakes by the overall frequency during 2008–2011 study period. They concluded that Lake Apopka in Florida (cyanoHAB frequency = 99.1%) and Grand Lake St. Marys in Ohio (cyanoHAB frequency = 83.1%) had the highest cyanoHAB frequency during 2008–2011. Their top ten lakes from Florida based on cyanoHAB bloom frequency¹⁵ (in order: Apopka (1), Pierce, Dora, Marion, Howard, Parker, Hancock, Harris, Jesup, and Juliana (10)) are in the top 18 lakes in our study (Fig. 4). As expected, a comparison of Florida lake ranks based on the frequency and area-normalized magnitude highlights differences in information between the methods. A lake with a persistent moderate bloom would rank higher in frequency than in magnitude; an example is Lake Pierce (rank 2 in frequency, 9 in magnitude). A lake with short intense blooms would rank higher in magnitude than in frequency. An example is Hancock Lake, which has intense annual blooms (Chl-*a* of 300–500 mg m^{-3}) that only last for a few months⁵⁸, and ranks 1 in area-normalized magnitude and 7 in frequency. The difference between metrics can be more acute in large lakes, like Lake Okeechobee, which has blooms that are large in magnitude but do not cover most of the lake (Fig. 9). A single metric cannot highlight all aspects of cyanoHABs; complementary cyanoHAB metrics representing factors like bloom frequency, area, and area-normalized magnitude are needed.

Wind-driven mixing in the water column can add uncertainty to the cyanoHAB magnitude estimates for species with buoyancy regulation, such as *Microcystis aeruginosa* and *Aphanizomenon flos-aquae*. Satellite

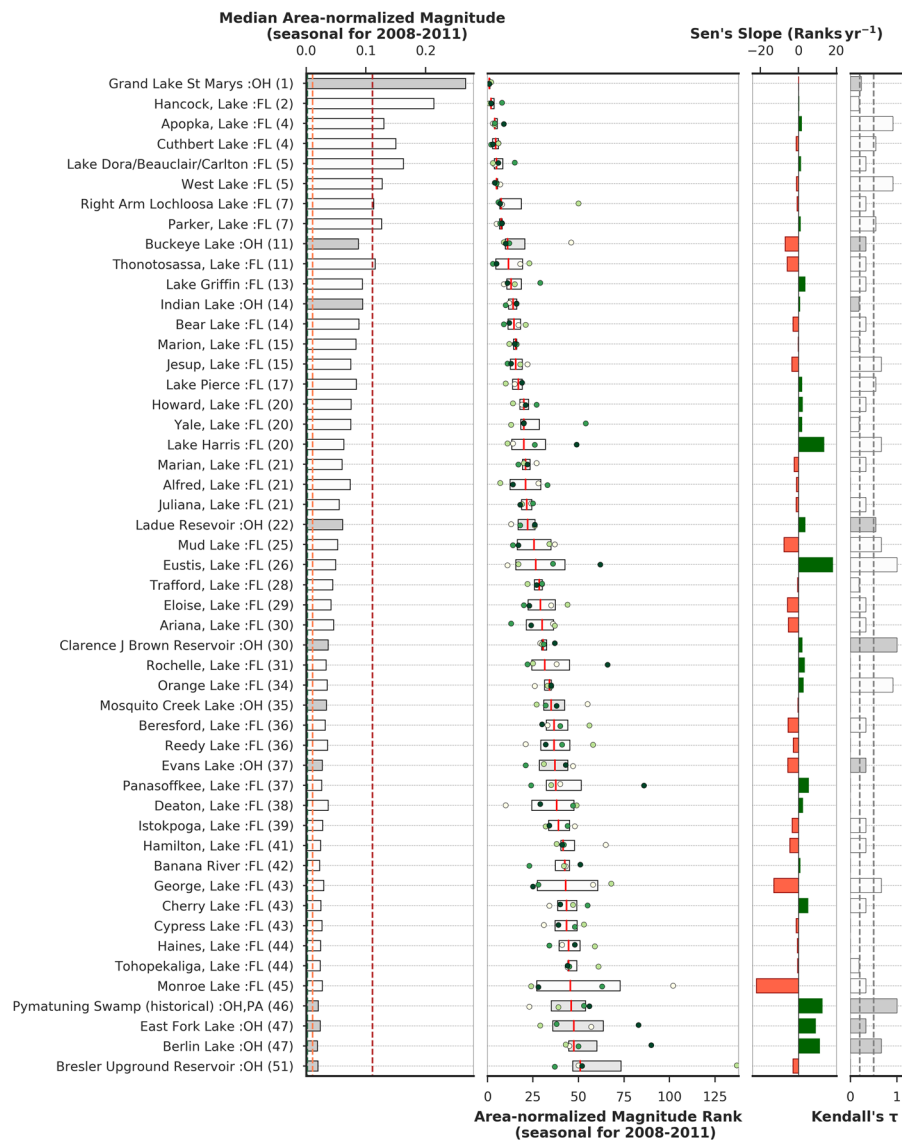


Figure 8. First panel: Seasonal area-normalized magnitude (km^{-2}) in the top 50 Florida and Ohio lakes. Green, orange, and red dotted lines represent the equivalent WHO thresholds of 20,000, 100,000, and 1,000,000 cells mL^{-1} limits. Lakes with white and shaded bars are located in Florida and Ohio, respectively; second panel: the inter-quartile range of area-normalized magnitude ranks in top-ranked Florida lakes over 2008–2011 ordered by their median rank over the 4 year period. Median values or ranks are highlighted in red vertical lines inside the box. Annual area-normalized magnitude rank data points are overlaid on inter-quartile boxes to highlight the variation, where the color of the scatter points indicate data year (tan: start year and deep green: end year). Third panel: green/red bar plot shows trends in rank change during 2008–2011. Green/red color represents positive/negative trend meaning area-normalized magnitude for a lake is decreasing/increasing over time. Fourth panel: bars show Kendall's τ (absolute values) representing consistency in rank-change trend over time. Dotted lines in Kendall's τ plot mark the τ values at 0.2 and 0.5.

algorithms, in bloom conditions, detect cyanobacteria concentration only near the surface. Previous studies have reported that wind stress can increase vertical mixing of cyanobacteria cells and thereby reduce the ability of the ocean or water color sensor to detect the majority of the biomass¹². Therefore, persistent high wind ($>7.7 \text{ m s}^{-1}$ as observed in Lake Erie)¹¹, when combined with frequent cloud cover (such that only windy days during the bloom season are imaged) may occasionally lead to underestimation of cyanoHAB biomass or area-normalized magnitude. Cloud cover, sun glint, and the effectiveness of masking algorithms for other invalid pixels (e.g. mixed land and water at the shore, dry lake bed, algal mats, and vegetative areas) may add uncertainty to the satellite-based measurements. Cloud and glint impacts should be uniform through a region, but the other factors might bias specific lakes. Another source of uncertainty could come from the use of the CI-cyanobacteria cell count relationship when the analysis would be scaled up to the CONUS lakes. In Lunetta *et al.*⁴¹, a CI-cyanobacteria biomass relationship was demonstrated using field data collected from lakes in the US states across Ohio, Florida, and throughout New England. The same CI-cyano and cyanobacteria biomass relationship was revised and presented

Lake Name	Surface Area (km ²) ⁺	State	Median Bloom Magnitude (dl)	Median Area-normalized Magnitude (km ⁻²)	Median Rank	Sen's Slope (Ranks yr ⁻¹)	Kendall's τ
Grand Lake St. Marys	47.97	OH	12.788	0.267	1.00	0.00	-0.24
Hancock, Lake	17.01	FL	3.630	0.213	2.00	0.25	0.18
Apopka, Lake	121.50	FL	15.754	0.130	4.00	1.50	0.91
Cuthbert Lake*	3.42	FL	0.513	0.150	4.50	-1.25	-0.55
Lake Dora/Beauclair/Carlton	21.24	FL	3.452	0.163	5.00	1.08	0.33
West Lake*	7.47	FL	0.950	0.127	5.00	-1.00	-0.91
Right Arm Lochloosa Lake	21.51	FL	2.424	0.113	7.50	-0.67	-0.33
Parker, Lake	7.65	FL	0.963	0.126	7.50	1.00	0.55
Buckeye Lake	9.27	OH	0.809	0.087	11.00	-7.00	-0.33
Thonotosassa, Lake	2.97	FL	0.342	0.115	11.50	-5.92	-0.33
Lake Griffin	38.88	FL	3.649	0.094	13.00	3.33	0.33
Indian Lake	16.29	OH	1.535	0.094	14.00	0.67	0.18
Bear Lake*	3.15	FL	0.278	0.088	14.50	-2.83	-0.33
Jesup, Lake	29.61	FL	2.212	0.075	15.50	-3.50	-0.67
Marion, Lake	10.71	FL	0.893	0.083	15.50	-0.17	-0.18
Lake Pierce	13.77	FL	1.155	0.084	17.00	1.67	0.55
Lake Harris	71.10	FL	4.481	0.063	20.00	13.33	0.67
Yale, Lake	14.94	FL	1.114	0.075	20.00	1.75	0.18
Howard, Lake	2.16	FL	0.162	0.075	20.00	2.08	0.33
Marian, Lake	18.27	FL	1.093	0.060	21.00	-2.33	-0.33
Alfred, Lake	2.43	FL	0.179	0.073	21.00	-1.08	0.00
Juliana, Lake	3.51	FL	0.194	0.055	21.50	-1.25	-0.33
Ladue Reservoir	4.77	OH	0.291	0.061	22.00	3.42	0.55
Mud Lake	1.71	FL	0.090	0.053	25.50	-7.58	-0.67
Eustis, Lake	29.79	FL	1.475	0.050	26.50	18.00	1.00
Trafford, Lake	5.49	FL	0.244	0.044	28.50	-0.50	-0.18
Eloise, Lake	4.14	FL	0.172	0.042	29.00	-5.75	-0.33
Ariana, Lake	3.69	FL	0.169	0.046	30.00	-5.25	-0.33
Clarence J Brown Reservoir	7.38	OH	0.270	0.037	30.50	1.83	1.00
Rochelle, Lake	1.98	FL	0.066	0.033	31.50	3.17	0.00
Orange Lake	23.58	FL	0.834	0.035	34.00	2.50	0.91
Mosquito Creek Lake	27.00	OH	0.912	0.034	35.00	-0.33	0.00
Beresford, Lake	2.70	FL	0.086	0.032	36.50	-5.50	-0.33
Reedy Lake	13.41	FL	0.479	0.036	36.50	-2.67	0.00
Evans Lake	1.98	OH	0.053	0.027	37.00	-5.67	-0.33
Panasoffkee, Lake	9.99	FL	0.259	0.026	37.50	5.17	0.00
Deaton, Lake	1.80	FL	0.066	0.036	38.00	2.17	0.00
Istokpoga, Lake	90.99	FL	2.493	0.027	39.00	-3.33	-0.33
Hamilton, Lake	8.10	FL	0.193	0.024	41.50	-4.50	-0.33
Banana River	3.24	FL	0.074	0.023	42.50	0.83	0.00
George, Lake	172.71	FL	5.049	0.029	43.00	-13.00	-0.67
Cypress Lake	11.61	FL	0.306	0.026	43.50	-1.17	0.00
Cherry Lake	1.62	FL	0.040	0.025	43.50	5.00	0.33
Tohopekaliga, Lake	64.71	FL	1.536	0.024	44.50	-0.50	-0.18
Haines, Lake	2.52	FL	0.060	0.024	44.50	-0.58	0.00
Monroe Lake*	2.70	FL	0.073	0.027	45.50	-22.08	-0.33
Pymatuning Reservoir	53.10	OH, PA	1.056	0.020	46.00	12.50	1.00
Berlin Lake	8.91	OH	0.168	0.019	47.50	11.00	0.67
East Fork Lake	6.57	OH	0.154	0.023	47.50	8.83	0.33
Bresler Upground Reservoir	2.25	OH	0.044	0.019	51.00	-2.92	0.00

Table 3. Summary of median seasonal bloom magnitude, median seasonal area-normalized magnitude and their ranks and statistics (Sen's Slope and Kendall's τ) of lakes in Florida and Ohio (2008–2011). ⁺Surface area determined by MERIS FR pixels within the lake polygon. * Indicates estuarine lake at the southern tip of Florida.

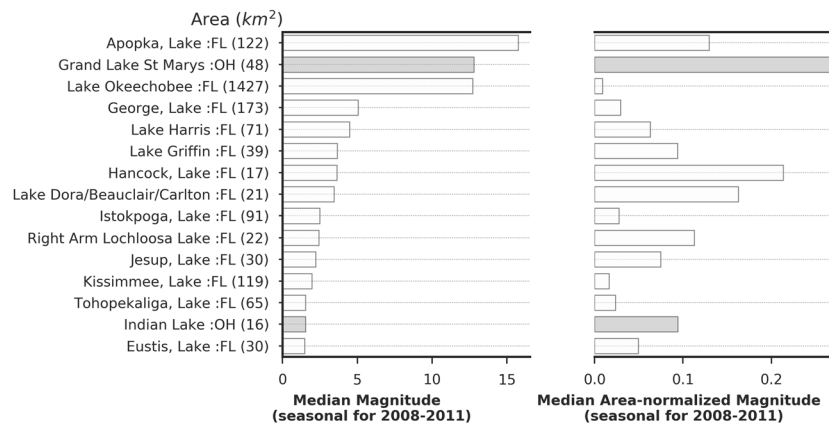


Figure 9. Left panel: Median seasonal bloom magnitude in 15 Florida and Ohio lakes over the study period (2008–2011) ordered by their values. Lake labels include the state name the lake is associated with and the number inside brackets represents the median area-normalized magnitude rank over the same study period. Right panel: median area-normalized magnitude for the same lakes during the same study period provided for comparison. Gray-colored bars represent lakes from Ohio.

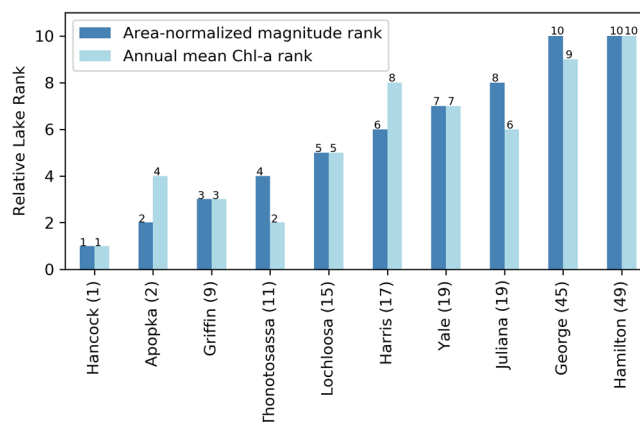


Figure 10. Relative comparison of lake ranks calculated from the annual area-normalized magnitude and measured annual mean Chl-*a* concentration. Numbers associated with the lake names in the x-axis tick label represent the median lake rank as in Table 1.

in Clark *et al.*¹⁵ with more meaningful error estimates (MAPE = 28.6%). Error in manual cell enumeration is inversely proportional to the number of colonies counted⁴⁸. 20–30% error is expected and considered acceptable when at least 400 units or colonies are counted although Chorus and Bartram⁴⁸ report that normally 20–40 colonies may be present in 100 mL of sampled water from the field. Therefore, even higher than 20–30% error, only from field cell density, cannot be ruled out. Similarly, variability in the spatial distribution of biomass in a bloom can add up to two orders of magnitude difference in biomass, as observed by⁵⁹ in a cyanobacteria bloom in the Gulf of Finland. Spatiotemporal variability in biomass in a diurnal scale can add significant uncertainty as well⁶⁰. Therefore, after considering errors from multiple sources, it is expected to have greater than 30% error or difference in satellite estimates, when compared with field measured cell density from a point source in a bloom event. Additionally, a CI-*cyano* to Chl-*a* relationship, established by Tomlinson *et al.*¹⁶ for Florida lakes, could be used in future studies, although the CI-*cyano* to Chl-*a* relationship may require additional examination before applying to all CONUS lakes.

Lake size presents a potential limitation on decisions based on area-normalized magnitude. One such example is Lake Okeechobee in Florida, which is the largest freshwater lake in Florida and the 9th largest freshwater lake (by area) in the United States. Due to its size and societal importance (for water supply, tourism, and ecological impacts), cyanoHAB issues in this lake have been widely covered by the press and media, thereby creating cyanoHAB awareness at state and national levels. While Lake Okeechobee was ranked second in bloom magnitude (behind Apopka, a moderately large lake), based on the area-normalized magnitude, Lake Okeechobee was ranked 95th among the Florida lakes. This is because the bloom area is simply a small percentage of surface area in such a large lake. Therefore, for larger lakes, annual or seasonal bloom magnitude numbers should be used for lake water management and decision making related to water quality. In contrast, area normalization highlights cyanoHABs in most of the smaller lakes such as Hancock Lake, Lake Dora/Beauclair/Carlton, and many others,

which may not get enough attention due to their size, although they equally bear the potential of causing an adverse effect on health and the environment.

The proposed bloom magnitude metric provided a synoptic view in lakes by capturing spatiotemporal mean areal cyanobacteria biomass. Further normalization of bloom magnitude by lake surface area provided comparable and actionable information for water quality managers inside states or other jurisdictional boundaries. The relative ranking of lakes allows the MERIS record (2003–2011) to be utilized within a state or a region, as the MERIS FR temporal frequency is expected to be similar. Ranks and nonparametric statistics provide robust parameters that do not depend on the calibration accuracy and precise thresholds, as compared to the direct metrics like a bloom area, bloom frequency, or area-normalized magnitude. The rank-based metric has additional power of allowing the use of multiple satellites without introducing biases between the different satellite data sets. For example, OLCI on Sentinel 3A and 3B may not currently match the MERIS calibration as the OLCI calibration is still on-going as of this writing, but the ranking of the lakes would eliminate the systematic bias in the data due to differences in calibration coefficients. Therefore, area-normalized magnitude ranks estimated from OLCI should be consistent with those from MERIS in each season, allowing identification of those lakes that are changing in bloom magnitude. The area-normalized magnitude should be inspected (together with sample frequency) to confirm that there is not a systemic change in all the lakes in a region over the time period of interest. This problem is less likely to occur if many lakes are considered in the analysis, or if they have fundamentally different environmental characteristics. These sensors cannot resolve all lakes of interest in a state. Narrow lakes and some rivers are a particular problem. For those small or narrow water bodies, Sentinel-2 may provide a solution, but that requires more research, as some key bands (620 and 681 nm) are not on that sensor.

The method presented in this study captures an assessment of cyanoHAB magnitude, which is the cyanobacteria biomass for the year or season. Normalization of bloom magnitude by lake surface area let us compare the cyanoHAB magnitude across lakes with varying size. Our approach to rank the lakes by median area-normalized magnitude helped us to highlight the top lakes, which need immediate attention from water quality managers. Provided below are three advantages of the ranking:

1. This approach uses the power of ranked and non-parametric statistics in order to be able to use the MERIS time series (2003–2011) in a state or a localized region irrespective of bias in temporal coverage. However, if contemporaneous satellite data collection frequency is different between two areas, they cannot be compared side-by-side and lakes in those areas should be analyzed separately.
2. This approach would also allow the use of Sentinel-3 OLCI data along with MERIS time series even though the sensor calibration coefficients of OLCI are still being refined and may not match with MERIS. This would enable a comparison of cyanoHAB bloom magnitude derived from OLCI with the historic cyanoHAB magnitude derived from MERIS.
3. As our method also included WHO recreational thresholds, the same information may also be used for categorizing which lakes need pressing attention for cyanoHAB management. A specific threshold representing exposure risk can be set and lakes above that threshold may be identified as a priority during the observational period.

No one metric can completely represent the attributes of cyanoHAB severity of interest to water quality managers. Area-normalized magnitude can provide awareness of smaller lakes that have significant blooms. However, the area-normalized magnitude can equate a short-lived, large, intense bloom with a long-lived, moderate bloom in any sized lake. Therefore, the reader is encouraged to compare other metrics such as temporal frequency¹⁵ and bloom spatial extent²⁹ to address related questions. These methods complement each other and can provide a more complete picture of cyanoHABs on a regional or national scale. The Cyanobacteria Assessment Network (CyAN) project¹³ provides the capability to scale this effort to CONUS fresh water lakes and water bodies. Our future work could focus on providing a comprehensive analysis of cyanoHAB magnitude in CONUS lakes and identifying lakes of concern.

Data availability

The satellite dataset used in the current study can be downloaded from <https://oceancolor.gsfc.nasa.gov/> and datasets generated during this study are available from the corresponding author on request.

Received: 29 May 2019; Accepted: 12 November 2019;

Published online: 04 December 2019

References

1. Boyer, G.L. In *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs* 153–165 (Springer, 2008).
2. Carmichael, W. W. Health effects of toxin-producing cyanobacteria: “The CyanoHABs”. *Human and ecological risk assessment: An International Journal* 7, 1393–1407 (2001).
3. Carmichael, W. W. & Boyer, G. L. Health impacts from cyanobacteria harmful algae blooms: Implications for the North American Great Lakes. *Harmful Algae* 54, 194–212 (2016).
4. Paerl, H. W. & Paul, V. J. Climate change: links to global expansion of harmful cyanobacteria. *Water Res* 46, 1349–1363 (2012).
5. Taranu, Z. E. *et al.* Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecol Lett* 18, 375–384 (2015).
6. Paerl, H. W. & Huisman, J. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environ Microbiol Rep* 1, 27–37 (2009).
7. King, K. W., Williams, M. R. & Fauser, N. R. Contributions of systematic tile drainage to watershed-scale phosphorus transport. *Journal of environmental quality* 44, 486–494 (2015).

8. Michalak, A. M. *et al.* Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc Natl Acad Sci USA* **110**, 6448–6452 (2013).
9. Paerl, H. W., Hall, N. S. & Calandrino, E. S. Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. *Sci Total Environ* **409**, 1739–1745 (2011).
10. Loftin, K. A. *et al.* Cyanotoxins in inland lakes of the United States: Occurrence and potential recreational health risks in the EPA National Lakes Assessment 2007. *Harmful Algae* **56**, 77–90 (2016).
11. Wynne, T. T., Stumpf, R. P., Tomlinson, M. C. & Dyble, J. Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and meteorological data. *Limnol Oceanogr* **55**, 2025–2036 (2010).
12. Stumpf, R. P., Wynne, T. T., Baker, D. B. & Fahnenstiel, G. L. Interannual variability of cyanobacterial blooms in Lake Erie. *PLoS One* **7**, e42444 (2012).
13. Schaeffer, B. A., Loftin, K. A., Stumpf, R. P. & Werdell, P. J. Agencies collaborate, develop a cyanobacteria assessment network. *EOS-Earth & Space Science News* **96** (2015).
14. Duan, H., Ma, R. & Hu, C. Evaluation of remote sensing algorithms for cyanobacterial pigment retrievals during spring bloom formation in several lakes of East China. *Remote Sensing of Environment* **126**, 126–135 (2012).
15. Clark, J. M. *et al.* Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. *Ecological Indicators* **80**, 84–95 (2017).
16. Tomlinson, M. C. *et al.* Relating chlorophyll from cyanobacteria-dominated inland waters to a MERIS bloom index. *Remote Sensing Letters* **7**, 141–149 (2016).
17. Kahru, M. & Elmgren, R. Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. *Biogeosciences* **11**, 3619–3633 (2014).
18. Palmer, S. C. J. *et al.* Satellite remote sensing of phytoplankton phenology in Lake Balaton using 10 years of MERIS observations. *Remote Sensing of Environment* **158**, 441–452 (2015).
19. Moradi, M. Comparison of the efficacy of MODIS and MERIS data for detecting cyanobacterial blooms in the southern Caspian Sea. *Mar Pollut Bull* **87**, 311–322 (2014).
20. Hu, C. *et al.* Moderate Resolution Imaging Spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China. *Journal of Geophysical Research* **115** (2010).
21. Shi, K. *et al.* Long-term MODIS observations of cyanobacterial dynamics in Lake Taihu: Responses to nutrient enrichment and meteorological factors. *Sci Rep* **7**, 40326 (2017).
22. Binding, C. E., Greenberg, T. A. & Bukata, R. P. An analysis of MODIS-derived algal and mineral turbidity in Lake Erie. *Journal of Great Lakes Research* **38**, 107–116 (2012).
23. Kahru, M., Elmgren, R. & Savchuk, O. P. Changing seasonality of the Baltic Sea. *Biogeosciences* **13**, 1009 (2016).
24. Binding, C. E., Greenberg, T. A., Bukata, R. P., Smith, D. E. & Twiss, M. R. The MERIS MCI and its potential for satellite detection of winter diatom blooms on partially ice-covered Lake Erie. *Journal of Plankton Research* **34**, 569–573 (2012).
25. Simis, S. G. H., Peters, S. W. M. & Gons, H. J. Optical changes associated with cyanobacterial bloom termination by viral lysis. *Journal of Plankton Research* **27**, 937–949 (2005).
26. Mishra, S., Mishra, D. R., Lee, Z. & Tucker, C. S. Quantifying cyanobacterial phycocyanin concentration in turbid productive waters: A quasi-analytical approach. *Remote Sensing of Environment* **133**, 141–151 (2013).
27. Gorham, T., Jia, Y., Shum, C. & Lee, J. Ten-year survey of cyanobacterial blooms in Ohio's waterbodies using satellite remote sensing. *Harmful algae* **66**, 13–19 (2017).
28. Wynne, T. T. *et al.* Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. *Int J Remote Sens* **29**, 3665–3672 (2008).
29. Urquhart, E. A., Schaeffer, B. A., Stumpf, R. P., Loftin, K. A. & Werdell, P. J. A method for examining temporal changes in cyanobacterial harmful algal bloom spatial extent using satellite remote sensing. *Harmful algae* **67**, 144–152 (2017).
30. USEPA National Lakes Assessment: A Collaborative Survey of the Nation's Lakes. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C. (2009).
31. Loftin, K. A. *et al.* Total cylindrospermopsins, microcystins/nodularins, and saxitoxins data for the 2007 United States Environmental Protection Agency National Lake Assessment: U.S. Geological Survey Data Series 929 (2016).
32. US Climate-Data Version 2.3. Assessed March 03, 2019 from, <https://www.usclimatedata.com/climate/florida/united-states/3179>. (2019).
33. McKay, L. *et al.* NHDPlus Version 2: User Guide (2012).
34. USEPA 2012 National Lakes Assessment Site Evaluation Guidelines. Office of Water, Washington, DC EPA 841-B-11-005. (2011).
35. Baith, K., Lindsay, R., Fu, G. & McClain, C. R. Data analysis system developed for ocean color satellite sensors. *Eos, Transactions American Geophysical Union* **82**, 202–202 (2001).
36. Carroll, M., Townshend, J. R., DiMiceli, C. M., Noojipady, P. & Sohlberg, R. A. A new global raster water mask at 250 m resolution. *International Journal of Digital Earth* **2**, 291–308 (2009).
37. Wynne, T., Meredith, A., Briggs, T., Litaker, W. & Stumpf, R. Harmful Algal Bloom Forecasting Branch Ocean Color Satellite Imagery Processing Guidelines. NOAA Technical Memorandum NOS NCCOS **252**, 48 (2018).
38. Tum, M. *et al.* Global gap-free MERIS LAI time series (2002–2012). *Remote Sensing* **8**, 69 (2016).
39. MERIS FR Density-Maps Accessed June 15, 2018, from, <https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat/news/-/article/envisat-meris-frs-density-maps>
40. Ohio, E. P. A. Factsheet Draft Rule – Recreational Water Criteria, Accessed December 15, 2018, from, https://www.epa.ohio.gov/Portals/35/rules/fs_IPR_WQS_RecUse_1_mar15.pdf.
41. Lunetta, R. S. *et al.* Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. *Remote Sensing of Environment* **157**, 24–34 (2015).
42. Stumpf, R. P. & Werdell, P. J. Adjustment of ocean color sensor calibration through multi-band statistics. *Optics Express* **18**, 401–412 (2010).
43. Seppala, J. *et al.* Ship-of-opportunity based phycocyanin fluorescence monitoring of the filamentous cyanobacteria bloom dynamics in the Baltic Sea. *Estuarine, Coastal and Shelf Science* **73**, 489–500 (2007).
44. Matthews, M. W. & Odermatt, D. Improved algorithm for routine monitoring of cyanobacteria and eutrophication in inland and near-coastal waters. *Remote Sensing of Environment* **156**, 374–382 (2015).
45. Stumpf, R. P. *et al.* Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae* **54**, 160–173 (2016).
46. Matthews, M. W., Bernard, S. & Robertson, L. An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Remote Sensing of Environment* **124**, 637–652 (2012).
47. Wynne, T. T. & Stumpf, R. P. Spatial and temporal patterns in the seasonal distribution of toxic cyanobacteria in western lake erie from 2002–2014. *Toxins* **7**, 1649–1663 (2015).
48. Chorus, I. & Bartram, J. Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management. (CRC Press, 1999).
49. Sen, P. K. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American statistical association* **63**, 1379–1389 (1968).
50. Kendall, M. G. A new measure of rank correlation. *Biometrika* **30**, 81–93 (1938).

51. Hirsch, R. M. & Slack, J. R. A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research* **20**, 727–732 (1984).
52. European Union Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official journal of the European communities* **22**, 2000 (2000).
53. Alikas, K. *et al.* Satellite-based products for monitoring optically complex inland waters in support of EU Water Framework Directive. *Int J Remote Sens* **36**, 4446–4468 (2015).
54. Davis, T. W. *et al.* Science meets policy: a framework for determining impairment designation criteria for large waterbodies affected by cyanobacterial harmful algal blooms. *Harmful algae* **81**, 59–64 (2019).
55. Scavia, D., DePinto, J. V. & Bertani, I. A multi-model approach to evaluating target phosphorus loads for Lake Erie. *Journal of Great Lakes Research* **42**, 1139–1150 (2016).
56. OEPA Ohio Environmental Protection Agency, Accessed August 15, 2019, from, <https://www.epa.state.oh.us/>.
57. Hoyer, M. V., Bigham, D. L., Bachmann, R. W. & Canfield, D. E. Jr. Florida LAKEWATCH: Citizen Scientists protecting Florida's aquatic systems. *Florida Scientist* **77**, 184 (2014).
58. The Water Atlas The Water Atlas Program, Accessed August 15, 2019, from, <http://www.wateratlas.usf.edu/>.
59. Kutser, T. Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnol Oceanogr* **49**, 2179–2189 (2004).
60. Hunter, P., Tyler, A., Willby, N. & Gilvear, D. The spatial dynamics of vertical migration by *Microcystis aeruginosa* in a eutrophic shallow lake: A case study using high spatial resolution time-series airborne remote sensing. *Limnol Oceanogr* **53**, 2391–2406 (2008).

Acknowledgements

The work was supported by NASA Ocean Biology and Biogeochemistry Program/Applied Sciences Program proposal 14-SMDUNSOL14-0001, NOAA, USGS, and EPA. We would like to thank three anonymous reviewers for providing constructive comments and suggestions, which helped to significantly improve the quality of the manuscript. This article has been reviewed by the National Exposure Research Laboratory and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. Government. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA.

Author contributions

S.M. and R.P.S. designed the research; S.M. and A.M. conducted data preprocessing. S.M. carried out data processing and statistical analysis; P.J.W. and R.P.S. provided datasets; R.P.S., B.S., P.J.W., K.L., and A.M. contributed to the interpretation and discussion of the results; and S.M. wrote the paper with contributions from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019