Remote Sensing of Environment 191 (2017) 273–285

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Using Landsat to extend the historical record of lacustrine phytoplankton blooms: A Lake Erie case study

Jeff C. Ho [a,b,*], Richard P. Stumpf [c], Thomas B. Bridgeman [d], Anna M. Michalak [b]

[a] Department of Civil & Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, CA 94305, USA
[b] Department of Global Ecology, Carnegie Institution for Science, 260 Panama St., Stanford, CA 94305, USA
[c] NOAA National Centers for Coastal Ocean Science, 1305 East-West Highway code N/SCI1, Silver Spring, MD 20910, USA
[d] Department of Environmental Sciences and the Lake Erie Center, University of Toledo, 6200 Bayshore Rd., Oregon, OH 43616, USA

ARTICLE INFO

Article history:
Received 15 July 2016
Received in revised form 9 December 2016
Accepted 20 December 2016
Available online xxxx

Keywords:
Landsat
Lake Erie
Harmful algal bloom
Phytoplankton
Cyanobacteria
Freshwater
Long-term data
Algorithm evaluation
Hindcasting

ABSTRACT

Long-term records of phytoplankton blooms in freshwater lakes are necessary both for understanding basin-scale changes to watersheds and for providing a key constraint for assessing processes driving blooms. However, due to the inherent constraints of in situ sampling and the short time period covered by modern space borne sensors, few long-term records exist. Historical data from sensors such as Landsat offer strong potential for creating new records of past blooms. Here, we use a novel evaluation procedure based on multiple metrics to assess algorithm suitability and robustness for generating long-term bloom records using Landsat 5 imagery. Evaluation metrics are based on bloom presence, spatial distribution, magnitude and timing, using both in situ *Microcystis* biovolume and remotely-sensed Cyanobacterial Index (CI) data from MERIS for 2002–2011. Applying this procedure for a test case focusing on Lake Erie's western basin, an algorithm based on a near infrared threshold with simple atmospheric correction through subtraction of the shortwave infrared band, combined with an additional "greenness" filter based on a hue threshold, performs best. Implementing this algorithm for 1984–2001 reveals the long-term trends in peak bloom magnitude prior to the start of the MERIS and MODIS record (2002–2015), and more than doubles the period of record that can be used to understand bloom occurrence and growth for this system. More broadly, we demonstrate that Landsat observations can be used to identify macro-scale features of blooms. For Lake Erie, the performance of the final Landsat algorithm is comparable to that of the MERIS CI algorithm, despite Landsat's broad spectral bands and long revisit time.

1. Introduction

Excessive and harmful phytoplankton blooms are a growing threat in freshwater systems. Blooms have become larger and more frequent due to anthropogenic phosphorus input (Schindler et al., 2016), and have been exacerbated by global climate change (Paerl and Huisman, 2009). Shallow, eutrophic systems are especially susceptible, due to the elevated risk from high nutrient inputs and high light availability providing favorable conditions for bloom growth. One well-studied system is Lake Erie, the most productive of the Laurentian Great Lakes. Lake Erie is an important drinking water source for the region and has experienced intensifying cyanobacterial blooms in the past decade (Allinger and Reavie, 2013). The lake's western basin is emblematic of the growing challenge of managing blooms in freshwater systems, as exemplified by a toxic bloom in 2014 that led to a three-day "do-not-drink" tap water advisory for local residents (Wilson, 2014). Improvements to existing approaches for bloom monitoring and prediction are necessary in order to better manage the problem of harmful phytoplankton blooms.

The information used to guide phosphorus reduction targets for managing blooms in Lake Erie comes from a relatively small number of modern bloom events, in large part because less is known about the severity of historical blooms. For instance, statistical models providing a quantitative link between nutrient load and bloom severity have only utilized data on bloom events from 2002 onward (e.g., Bertani et al., 2016; Obenour et al., 2014; Stumpf et al., 2016; Stumpf et al., 2012). This is because data on bloom magnitude comes mainly from remotely-sensed measurements by the MEdium Resolution Imaging Spectrometer (MERIS), which collected data from 2002 to 2011, and from the Moderate Resolution Image Spectrometer (MODIS) thereafter (ESA, 2016; NASA, 2016a). Most information on blooms prior to 2002 is based on types and/or locations of observations that are not entirely

* Corresponding author at: Department of Civil & Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, CA 94305, USA.
E-mail address: jeffho@stanford.edu (J.C. Ho).

http://dx.doi.org/10.1016/j.rse.2016.12.013

intercomparable, further exacerbating the challenge (Ho and Michalak, 2015). More observations of historical phytoplankton blooms would improve current management efforts by adding to the data available for modeling phosphorus reduction targets. On a more basic level, additional information on historical blooms would also provide new insight into processes driving historical variability.

Landsat observations have the potential to provide a longer-term bloom record for western Lake Erie and other eutrophic systems, as Landsat has been collecting data since the early 1980s through its Thematic Mapper sensors. Landsat has been used to identify freshwater blooms in many other studies (e.g., through measurement of chlorophyll-*a* concentrations: Duan, et al., 2007; Tebbs et al., 2013; Tyler et al., 2006; or through detection of bloom presence/absence: Duan et al., 2009). Its spatial resolution (30 m) and long time span have proved especially useful for inland water investigations (Mouw et al., 2015). However, Landsat's broad spectral bands (630-690 nm red band and 760–900 nm near-infrared band) make it difficult to identify signals unique to phytoplankton, because narrower bands are needed to accurately measure shifts in phytoplankton pigment concentrations (e.g., at 670 nm and 700 nm, or around 685 nm, for the best proxies for chlorophyll-*a*, or at 620 nm for phycocyanin; Matthews, 2011). Landsat also has a long revisit time (16 days in most locations), which makes it difficult both to track changes in phytoplankton abundance and also to reliably obtain cloud-free images (Matthews, 2011; Sayers et al., 2015).

Assessing the potential benefits of using Landsat for characterizing blooms would substantially advance the state of the science for remote sensing of inland waters. For Lake Erie, understanding the degree to which robust estimates of bloom severity can be obtained using Landsat, especially relative to estimates from MERIS and MODIS, would be particularly useful for augmenting the record on historical blooms and thereby informing models used to guide nutrient reduction strategies (Bertani et al., 2016; Stumpf et al., 2016). As the use of MERIS cyanobacteria detection algorithms for monitoring freshwater lakes has expanded to more lakes across the eastern United States (Lunetta et al., 2015), achieving comparable bloom detection with Landsat would also open up new possibilities for monitoring long term trends across a large number of inland systems. The additional potential for ongoing monitoring of bloom events using Landsat 8 (NASA, 2016b), which was launched in 2013, further motivates the evaluation of Landsat's accuracy in characterizing blooms.

In this study, we investigate the accuracy with which Landsat can be used to identify historical blooms in Lake Erie by comparing the outputs of different Landsat detection algorithms with ten years of data from in situ sampling and remote sensing. Various Landsat algorithms are evaluated based on their ability to identify blooms during a period for which the occurrence of blooms in Lake Erie is relatively well understood (Ho and Michalak, 2015), and assessed based on the accuracy of their estimates of bloom occurrence, spatial extent, and timing. The aims of this study are to (i) identify the best algorithm for classifying blooms during this evaluation period; (ii) characterize the potential uncertainties associated with that algorithm in hindcasting historical blooms; and (iii) hindcast bloom extent for western Lake Erie during the period 1984–2001.

2. Materials & methods

2.1. Site description

Lake Erie is the shallowest of the Laurentian Great Lakes in North America, with an average depth of 19 m (Lake Erie LaMP, 2011). Separated into three distinct basins by bathymetry, the western basin is the shallowest, with an average depth of 7.4 m (Lake Erie LaMP, 2011). Western basin waters receive the majority of the nutrient inputs to the lake (Maccoux et al., 2016), and have been classified as mesotrophic in the 1980s and early 1990s to eutrophic and hypereutrophic starting from the 2000s (Allinger and Reavie, 2013; Kane et al., 2015). The majority of the annual nutrient load occurs in the spring, with the highest productivity occurring over several months in the summer. Since the mid-1990s, limited phytoplankton surveys have suggested that summer blooms have increased in biomass and have been populated primarily by the cyanobacterium *Microcystis aeruginosa* (Allinger and Reavie, 2013; Lake Erie LaMP, 2011). From 2002 to 2015, the median annual peak bloom size in the western basin was 912 km^2, based on data from Stumpf et al. (2012) and Stumpf et al. (2016), with the largest blooms occurring in 2011 and 2015 (Michalak et al., 2013; Stumpf et al., 2016). Based on this increase in severity, nutrient reduction targets were set by Canada and the U.S. to restore Lake Erie's ecosystem health and provide protection for the 11 million people for whom Lake Erie provides drinking water (Lake Erie LaMP, 2011; Scavia et al., 2016).

2.2. Data

We focus on Landsat images covering the western basin of Lake Erie in July to October for 1984–2011 (i.e., all the years with Landsat 5 data). Though blooms occasionally extend into the central basin, the western basin is where bloom biovolume is most highly concentrated (Chaffin et al., 2011), where in situ samples are available most frequently (Bridgeman et al., 2013), and where overlapping Landsat scenes offer a higher frequency of image collection (8 days rather than the 16 day return time). A total of 303 images are available during the selected months, with 109 scenes during the evaluation period (2002 − 2011) and 194 during the historical period (1984–2001). The analysis is based on images from both the Landsat 5 raw digital number (DN) and top-of-atmosphere (TOA) reflectance image collections in Google Earth Engine (Google, 2016), collected originally from the U.S. Geological Survey (USGS, 2016). The images in these collections represent Level 1 quantized calibrated values (Q_{cal}) and planetary TOA reflectance (ρ_λ), respectively, as defined in Chander et al. (2009):

$$Q_{cal} = \frac{L_\lambda - B_{rescale}}{G_{rescale}} \quad (1)$$

$$\rho_\lambda = \frac{\pi L_\lambda d^2}{ESUN_\lambda \cos(\theta_s)} \quad (2)$$

where

L_λ is the spectral radiance at the sensor's aperture [$W/(m^2 \cdot sr \cdot \mu m)$];

$B_{rescale}$ is the band-specific rescaling bias factor from Chander et al. (2009) [$W/(m^2 \cdot sr \cdot \mu m)$];

$G_{rescale}$ is the band-specific rescaling gain factor from Chander et al. (2009) [DN $W/(m^2 \cdot sr \cdot \mu m)$];

d is the Earth-sun distance [astronomical units];

$ESUN_\lambda$ is the mean exoatmospheric solar irradiance [$W/(m^2 \cdot \mu m)$]; and

θ_s is the solar zenith angle [degrees].

We use two types of observations for algorithm evaluation. The first data set includes *Microcystis* biovolume measurements collected in situ every 10–20 days at six locations in Maumee Bay at the southwest corner of the western basin (Bridgeman et al., 2013). *Microcystis* biovolume is measured by separating the floating *Microcystis* from a vertical plankton tow taken from lake bottom to surface at each location (more details in Bridgeman et al., 2013). The second data set consists of daily and ten-day composite Cyanobacterial Index (CI) values of the western basin calculated from MERIS imagery (Stumpf et al., 2012). The CI algorithm is a spectral curvature method around the 681 nm band applied to daily MERIS images acquired from the ESA's Envisat Mission and to daily MODIS images from 2012 onward (ESA, 2016; NASA, 2016a; more details in Wynne et al., 2010; Wynne et al., 2008). These two data sets were selected because they span the 2002–2011 evaluation period and because they are largely consistent in terms of their assessment of relative bloom magnitudes over this time (Ho and Michalak, 2015).

2.3. Selection of algorithms to be evaluated

We evaluate eleven Landsat algorithms (Table 1). Although more complex algorithms are possible, for example ones that utilize machine learning techniques (e.g., Pozdnyakov et al., 2005), we consider only algorithms that can be implemented on a per-pixel basis. We do this in order to take advantage of cloud-based parallel computing tools available via Google Earth Engine that have substantially reduced the time and resources necessary to perform analyses on remote-sensing imagery (Google, 2016). Among the examined algorithms, six have previously been recommended for use in freshwater systems (#1–6), while five are based on broader best practices in the remote sensing literature (#7–11). Some are intended to measure pigment concentrations indicative of phytoplankton abundance (either chlorophyll-*a* or phycocyanin), while others are more closely tied to binary measures of bloom presence or absence. To allow a direct comparison between the algorithms, and also because measurement of pigment concentrations typically requires sufficient in situ pigment measurements for robust parameterization, we use each algorithm to characterize bloom presence/absence with a threshold for bloom classification.

The first four algorithms were gleaned from studies describing empirical algorithms for remote sensing of inland water quality. Algorithm #1 is based on a ratio of the red and blue bands, two bands frequently identified in a review of Landsat algorithms for Canadian lakes (Sass et al., 2007) and also from bio-optical theory indicative of high chlorophyll-*a* (>20 mg/m^3; Matthews, 2011). Algorithm #2, a ratio of the red and NIR bands, is also recommended for use in high chlorophyll-*a* situations (Matthews, 2011), and is based on the assumption that reflectance in the NIR channel is low or close to zero due to absorption by water. Algorithms #3 and #4 are both recommended for use in low chlorophyll-*a* situations (<20 mg/m^3; Matthews, 2011). These two attempts to account for the shift in reflectance to lower wavelengths and the increased effect of inorganic suspended solids, respectively, that occur with decreased chlorophyll-*a* (Matthews, 2011).

The next two empirical algorithms (#5 and 6) were developed specifically for individual lakes. Algorithm #5 was developed using Lake Erie phycocyanin data to assess phycocyanin concentrations (Vincent et al., 2004). Phycocyanin is an indicator pigment in cyanobacteria, and has been used to distinguish cyanobacteria from other types of phytoplankton (Dekker, 1993). Algorithm #6, which is based on Landsat's near-infrared band (Band 4), identifies chlorophyll-*a* and dense floating algae under high bloom conditions (Mouw et al., 2015) and has been used to observe blooms in Lake Tai (Duan et al., 2009).

The remaining algorithms (#7 through 11) apply well-known approaches from the ocean color, terrestrial vegetation, and atmospheric correction literatures. Algorithm #7 uses Landsat's shortwave infrared band as a simple atmospheric correction, in a similar approach to ocean color processing for turbid waters (Wang and Shi, 2007). Algorithm #8 looks for a difference between Landsat's NIR and red bands, similar to the NDVI (the Normalized Difference Vegetation Index) used for observing terrestrial vegetation, providing an implicit atmospheric correction for the NIR band using the red band. Algorithm #9 combines these two variants by adding an atmospheric correction, following others (Stumpf and Tyler, 1988), and algorithm #10 is one

Table 1
Algorithms evaluated for tracking bloom occurrence, spatial extent, and timing in Lake Erie.

#	Algorithm	Algorithm reference(s)	Algorithm short name	Formula(s)[a]	Threshold for bloom classification
1	Red to blue ratio	Lathrop, 1992	Red to blue	$\frac{B3}{B1}$	>0.462
2	Red to near-infrared ratio	Yacobi et al., 1995	Red to NIR	$\frac{B3}{B4}$	N/A[b]
3	Green to blue ratio	Gitelson et al., 1993	Green to blue	$\frac{B2}{B1}$	>0.841
4	(Blue minus red) over green	Mayo et al., 1995	Blue minus red over green	$\frac{(B1-B3)}{B2}$	N/A[b]
5	Phycocyanin detection, validated using Lake Erie data[c]	Vincent et al., 2004	Phycocyanin detection	$47.7-9.21(B3/B1)+29.7(B4/B1)-118(B4/B3)$ $-6.81(B5/B3)+41.9(B7/B3)-14.7(B7/B4)$	>5.93
6	Near-infrared (NIR), validated using Lake Tai data	Duan et al., 2009	NIR	$B4$	>0.0327
7	NIR with simple atmospheric correction (SAC)	Wang and Shi, 2007	NIR with SAC	$B4-B5$	>0.0277
7b	Further refinement of NIR with SAC (#7), as discussed in Section 3.1	Gordon, 1978 Haydn et al., 1982	Improved NIR with SAC	$B4-1.03\times B5$ $Hue=\begin{cases}\frac{B2-B1}{B3+B2-2\times B1} & \text{if } B1=\min(B1,B2,B3)\\ \frac{B1-B3}{B2+B1-2\times B3}+1 & \text{if } B3=\min(B1,B2,B3)\\ \frac{B3-B2}{B3+B1-2\times B2}+2 & \text{if } B2=\min(B1,B2,B3)\end{cases}$	>0.0235 <1.6
8	NIR minus red	Tucker, 1979	NIR minus red	$B4-B3$	N/A[b]
9	NIR over red, with SAC	Stumpf and Tyler, 1988	NIR over red with SAC	$\frac{(B4-B5)}{(B3-B5)}$	>0.495
10	NIR over red, with baseline atmospheric correction (BAC)[d]	ESA, 2015	NIR over red with BAC	$\frac{(B4-L4)}{(B3-L3)}$	N/A[b]
11	Curvature method around red band	Wynne et al., 2008	Curvature around red	$(B4-B3)+0.5(B3-B5)$	N/A[b]

[a] B# stands for the band # pixel value in the input data set (i.e., level 1 quantized calibrated values Q_{cal} for algorithm #5 and planetary TOA reflectance ρ_λ for all other algorithms).
[b] No threshold performed better than classifying all pixels as "non-bloom."
[c] This algorithm was derived from Landsat 7 data, so Landsat 5 bands are calibrated beforehand according to Vincent et al. (2011).
[d] Where: $L4=(B4-B1)+(B1-B5)\left(\frac{850-490}{1650-490}\right)$ and $L3=(B3-B1)+(B1-B5)\left(\frac{660-490}{1650-490}\right)$.

final variant that includes a baseline atmospheric correction using a procedure applied by the Sentinel-3 mission for accommodating turbid waters (ESA, 2015). The final algorithm (#11) uses the curvature method around the red band, and is similar to the spectral shape algorithm for identifying cyanobacteria (Wynne et al., 2008) by attempting to identify chlorophyll-*a* absorption relative to adjacent scattering.

All algorithms except for algorithm #5 use the TOA reflectance images as input. We do not implement formal atmospheric corrections (e.g. Hadjimitsis et al., 2004) beyond the simple approaches described for specific algorithms above. Algorithm #5 uses a different input because its coefficients are parameterized to dark object (DO)-subtracted spectral radiance images. However, raw DN images are instead used as input here because, during initial implementation, we found that the digital number images resulted in phycocyanin values closer to those in the original study (Vincent et al., 2004). This is likely because revised calibration coefficients in newer downloaded Landsat data changed the values in spectral radiance images (Chander et al., 2007). Because raw DN images are not impacted by such updates, they would be closer to the original study's source data. Although this may impact the applicability of the algorithm across different time periods, because the same DN values on different dates may correspond to different radiances, we select this approach in the interest of obtaining the most realistic phycocyanin output concentrations possible.

2.4. Algorithm implementation

We use a classification approach for identifying blooms, fitting a threshold to the output of each algorithm to classify pixels as "bloom" or "non-bloom" (Table 1). The thresholds used in classification are optimized by maximizing the number of correctly classified pixels for spatially and temporally collocated images based on MERIS CI (Stumpf et al., 2012) using a nearest-neighbor sampling technique (Google, 2016: https://developers.google.com/earth-engine/resample). The fitted thresholds are therefore as close as possible to the MERIS CI threshold for a bloom, 0.001 CI, intended to be equivalent to 10^5 cells/mL.

We use this approach for three reasons. First, there are not enough measurements of summer pigment concentration in western Lake Erie (either chlorophyll-*a* or phycocyanin) to robustly assess correlations with algorithm estimates. Combining all chlorophyll-*a* measurements from the U.S. Environmental Protection Agency and Environment Canada for 2002–2011, for example, yields only 111 measurements over 10 years (Environment Canada, 2012; EPA GLNPO, 2012), which is small relative to studies that have used in situ chlorophyll-*a* to validate remote sensing algorithms for similar lakes (e.g., Palmer et al., 2015, with 692 samples over six years). Moreover, the 111 measurements were all taken during single cruises in late July or early August each year (when blooms are just forming) rather than spanning the bloom season (July through September), limiting the number of concurrent Landsat overpasses and potentially biasing the parameterization toward lower chlorophyll-*a* values. Second, Landsat's broad bands are known to yield a low signal-to-noise ratio that makes discerning between fine gradations in water quality challenging (Matthews, 2011). Third, binary classification of pixels as "bloom" or "no bloom" has previously been shown to serve as a good proxy for bloom magnitude (e.g. for bloom biomass in Lake Erie: Stumpf et al., 2012; and for bloom severity in the Baltic Sea: Kahru and Elmgren, 2014). The classification approach is thus appropriate given the constraints of in situ data and the low signal-to-noise ratio of Landsat.

Once classification thresholds are identified for each algorithm, the thresholds are applied to each of the 109 evaluation period Landsat images to create binary bloom maps. Algorithms for which no threshold performs better than simply classifying all pixels as "bloom" or "non-bloom" are not included in this subsequent evaluation.

2.5. Algorithm evaluation

We use a multi-faceted approach to evaluate the algorithms based on the binary bloom maps (Table 2), comparing algorithm output with synoptic estimates based on MERIS CI (Stumpf et al., 2012) and individual observations collected in situ (Bridgeman et al., 2013). Previous studies introducing new algorithms have often relied on simple correlation statistics based on single data sets, and have been shown to yield results that are poorly comparable to other measurements for Lake Erie (Ho and Michalak, 2015). This proposed approach is expected to provide a more robust evaluation of algorithm performance.

We use the Bridgeman et al. (2013) data to evaluate algorithm accuracy in identifying bloom occurrence at specific locations (Table 2). In situ *Microcystis* biovolume measurements are compared to individual pixel values at the same coordinates from the Landsat overpass taken within three days of in situ sampling. We ran a sensitivity analysis to confirm that using observations taken within three days did not decrease algorithm accuracy relative to using observations from within 12 hours or one day. The biovolume data are transformed into binary data using the threshold for severe blooms from the Ohio Environmental Protection Agency (10 mm^3/L; Kasich et al., 2014).

We use the Stumpf et al. (2012) MERIS CI data to evaluate algorithm accuracy over the entire basin. First, we assess each algorithm's ability to correctly categorize blooms by their relative magnitude. The relative magnitude of blooms is a natural metric, especially during 2002–2011 when blooms fall into three relatively distinct categories of small, medium, and large blooms (Bridgeman et al., 2013; Stumpf et al., 2012). We first assess whether each algorithm correctly identifies 2011 as the largest bloom, and also what fraction of the set of four other large blooms (2003, 2008, 2009, 2010) based on the evaluation data sets are identified as such by each algorithm (Table 2). We then do the same for the three smallest blooms (2002, 2005, and 2007) based on the two evaluation data sets (Table 2). This approach assesses algorithm skill in making comparative statements about bloom magnitude between years.

Second, we evaluate algorithm skill in characterizing bloom magnitude, using individual MERIS CI images. CI-based bloom images are matched to the dates of Landsat overpasses (within one day), yielding 61 matched pairs. The pairs are evenly distributed across all years except for 2006, when there were large gaps in the MERIS imagery (Stumpf et al., 2012). Coefficients of determination and root mean squared errors from weighted least squares regression between the Landsat- and MERIS-estimated bloom areas are used to assess algorithm skill in characterizing bloom size variability (Table 2). Weighted least squares regression is used in situations where the measurements (in this case, the areas from each Landsat image) have uncorrelated but variable uncertainties (Carroll and Ruppert, 1988). The weights in this regression are based on the cloudy areas in the Landsat images (Eq. 3), to account for the higher uncertainties in cloud-obscured images. Clouds are identified using the Landsat Automatic Cloud Cover Assessment procedure (Irish, 2000).

$$w_i = 1 - \left(\frac{c}{c_{max}}\right) \tag{3}$$

where

w_i is the weight applied for the ith Landsat-MERIS bloom area pair in weighted linear regression [unitless];

c is the cloudy area in image [km^2]; and

c_{max} is the maximum possible cloudy area, set equal to 4000 km^2, approximately the size of the western basin.

We repeat this regression procedure using a subset of the data with only the largest bloom areas for each year (i.e., only 10 pairs of images versus 61, Table 2). This approach evaluates an algorithm's ability to provide useful data for water quality management, because estimates of annual maximum bloom magnitude are used to set nutrient input targets (Scavia et al., 2016). To more accurately measure bloom peak, we constrain the selection of the largest bloom area each year to dates

Table 2
Description of metrics used in comprehensive evaluation approach based on bloom occurrence, spatial extent, and timing. The color shading is intended to enable easy mapping to the results presented in Fig. 2.

	Evaluation data set	Extracted variable	Evaluation metric	Metric description
Bloom occurrence	Bridgeman et al., 2013	*In situ Microcystis* biovolume	% samples correct	Percentage of biovolume samples correctly identified (either as bloom or non-bloom)
	Stumpf et al., 2012	Ranking of years by bloom size	Is 2011 bloom largest?	Whether the 2011 bloom ranks largest out of all years
			% large blooms correct	Percentage of overlap among the years with the next four largest blooms
			% small blooms correct	Percentage of overlap among the three years with the smallest blooms
Bloom spatial extent		Bloom size in individual images (n=61)	wr^2_{all}	Coefficient of determination from weighted least squares regression of Landsat vs. MERIS bloom size based on individual images
			$RWMSE_{all}$ [km^2]	Root weighted mean squared error of bloom size from individual Landsat images relative to MERIS estimates
		Maximum bloom size each year (n=10)	wr^2_{max}	Coefficient of determination from weighted least squares regression of Landsat vs. MERIS bloom size based on the maximum bloom size observed using each instrument
			$RWMSE_{max}$ [km^2]	Root weighted mean squared error of maximum observed bloom size from Landsat relative to MERIS estimates
Bloom timing		Date of peak bloom selecting only overlapping dates	$RMSE_{overlap}$ [days]	Root mean squared error of date of bloom peak for Landsat vs. MERIS, only using images from overlapping dates
		Date of peak bloom selecting all possible dates	$RMSE_{all}$ [days]	Root mean squared error of date of bloom peak for Landsat vs. MERIS, using all available images

where the water temperature is above 20 °C, using water temperature data from the National Data Buoy Center (NDBC) Station 45005 (NOAA NDBC, 2014). This threshold is the midpoint of the temperature range estimated for cyanobacterial growth (Robarts and Zohary, 1987).

Third, we evaluate algorithm skill in estimating peak bloom timing by comparing the dates of peak bloom area for the Landsat algorithms vs. MERIS CI. This comparison is performed using two subsets of data: the first, only selecting images where Landsat and MERIS have overlapping dates (within one day), and the second, allowing all data from both sensors (Table 2). The first comparison highlights algorithmic limitations (i.e., differences in timing resulting from the fact that the Landsat and MERIS CI algorithms interpret information from concurrent dates differently). The second comparison measures accuracy when considering both algorithmic and temporal sampling limitations (i.e. differences in timing resulting both from algorithmic differences and the fact that the two sensors may not be observing the lake on the same dates).

Overall, this multi-metric approach for evaluating bloom detection provides a comprehensive picture of algorithm strengths and weaknesses, beyond evaluation and validation approaches that have been used traditionally. The results of this approach, and conclusions on algorithm suitability for generating historical estimates of bloom conditions, are presented in Sections 3.1 and 3.2 followed by an exploration of historical results using the best algorithm in Section 3.3.

3. Results & discussion

3.1. Algorithm intercomparison and performance

For five of the algorithms that were evaluated (#2, 4, 8, 10 and 11), simply classifying all pixels as "non-bloom" outperforms all thresholds, indicating that these algorithms provide no useful information on bloom presence for the western basin of Lake Erie based on the implemented classification approach (Table 1). For the remaining algorithms (#1, 3, 5, 6, 7 and 9), the optimized thresholds result in varying levels of agreement with MERIS CI images of blooms (e.g., Fig. 1) and in situ observations across the examined metrics.

Results from systematic evaluation of Landsat algorithms suggest that only a few of the examined algorithms are able to identify the relative magnitude, size, or timing of blooms using our approach (Fig. 2). The best of the original algorithms across most metrics is algorithm #7 (NIR with SAC), performing best across seven of the ten metrics (and tied with other algorithms for three of these). Algorithm #3 (green to blue) also performs well across many metrics, performing best for five of the ten metrics (tied with other algorithms for two of these).

We further optimize the NIR with SAC algorithm (#7) by adding a weight to the shortwave-infrared (SWIR) band and adjusting this

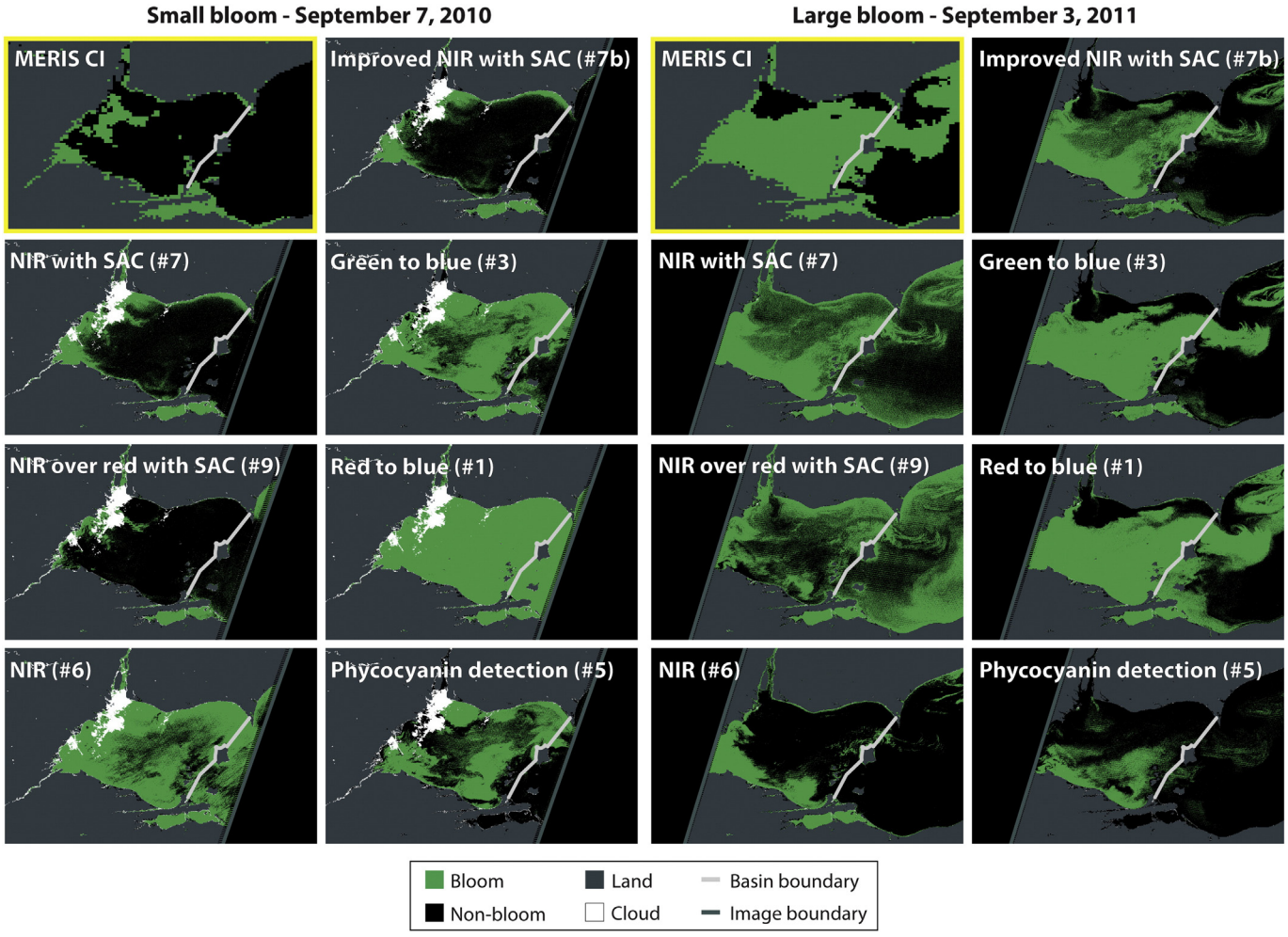

Fig. 1. Bloom classification images showing algorithm output for prototypical examples of small and large blooms, on September 7, 2010 and September 3, 2011, respectively. Image highlighted in yellow is the MERIS CI image to which algorithm output images are compared. Classification thresholds for each algorithm are reported in Table 1. Algorithms for which no threshold performed better than classifying all pixels as "non-bloom" are not shown.

weight to maximize coherence with the MERIS CI bloom images (i.e., adjusting the amount of atmospheric correction provided by Band 5, similar to Neukermans et al. (2009) and Stumpf and Pennock (1989) for other sensors). For this "Improved NIR with SAC" algorithm (#7b), both the weight and the threshold are optimized together to achieve a new threshold (Table 1). Because of misclassification in some images due to suspended sediment, we add a secondary filter based on the "greenness" of the pixel, implemented as hue values below a threshold fitted from MERIS data (fitted hue threshold = 1.6, Table 1; Haydn et al., 1982). The improvement in matching the MERIS CI bloom images after adjusting the weight of the SWIR band and adding the secondary filter step can be seen through the reduction in nearshore pixels erroneously identified as "bloom" in Fig. 1, and should minimize potential misclassification errors from events such as storm-induced sediment resuspension.

The refinements made in algorithm #7b result in improved performance across the evaluation metrics (Fig. 2), producing the best performance for nine of the ten examined metrics (tied with the original algorithm #7 for two of these, and also with algorithm #3 (Green to Blue) for one of these two). This algorithm is therefore selected as the best algorithm for evaluating bloom occurrence, spatial extent, and timing in the western basin of Lake Erie.

Starting with assessing bloom occurrence, this final algorithm identifies 65% of measurements correctly for the point comparisons to in situ *Microcystis* biovolume, compared to 63% for the original algorithm #7 and 60% for the next best algorithm. For identifying individual years with blooms, algorithms #7 and #7b are two of only three algorithms to correctly identify 2011 as the year with the largest bloom (Fig. 2). The final algorithm also correctly ranks 6/7 (85%, highest among all algorithms) of the four next largest and the three smallest blooms correctly, and is the only algorithm to rank all of the small blooms correctly (3/3). For assessing bloom spatial extent, the final algorithm achieves the highest weighted r^2 with MERIS CI images out of all algorithms, ($wr^2_{all} = 0.64$ versus $wr^2_{all} = 0.54$ for the next best algorithm (#7)) and also the lowest root weighted mean squared errors ($RWMSE_{all} = 262$ km^2 versus 297 km^2 for the next best algorithm (#7)). For the maximum bloom area subset ($n = 10$), the final algorithm also performs best, with the highest weighted r^2 ($wr^2_{max} = 0.71$ versus 0.68 for the next best algorithm (#3), Fig. S1) and lowest mean error ($RWMSE_{max} = 352$ km^2 versus 543 km^2 for algorithm #7 and 988 km^2 for the next best). When measuring bloom timing based on images for overlapping dates, the final algorithm again performs best, with an $RMSE_{overlap}$ of 15 days. Notably, algorithm #3 (Green to blue) performs marginally better than the final algorithm at identifying peak bloom timing when considering all available images.

All other algorithms perform relatively well for one or a few metrics, but not consistently well overall (Fig. 2). For example, the NIR over red with SAC algorithm (#9) is the only algorithm other than NIR with SAC (#7) and Improved NIR with SAC (#7b) to identify 2011 as the largest, but performs poorly at estimating bloom timing. Taken all together, these findings demonstrate the importance of using a multi-faceted approach for algorithm evaluation. Our results show that no algorithm

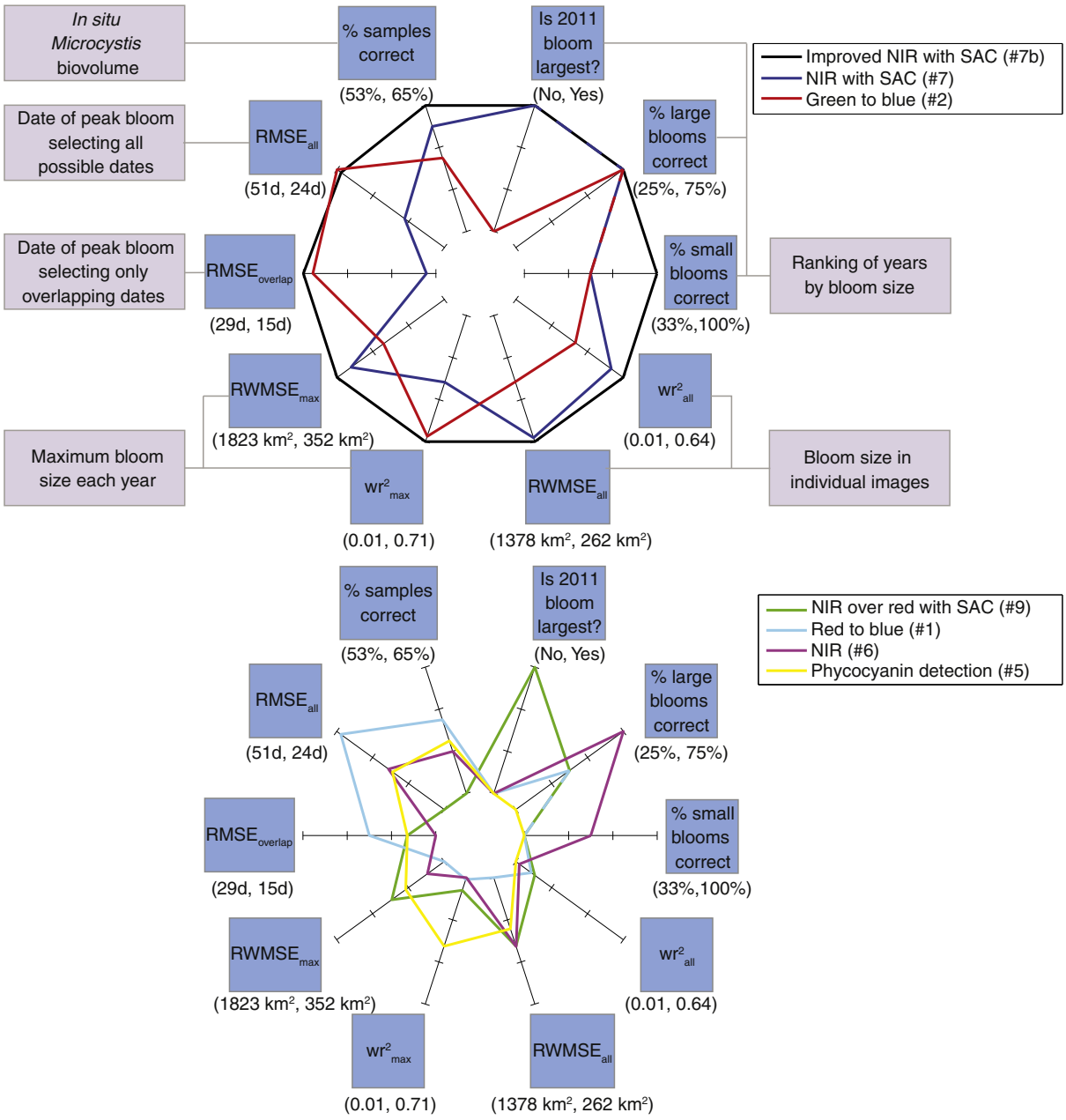

Fig. 2. Performance of algorithms across evaluation metrics, where the outer edge of the circle represents the value of a particular metric for the best performing algorithm, and the innermost portion of the circle represents the value for the worst performing algorithm. The observed ranges for individual metrics across algorithms are noted below each evaluation metric (abbreviations explained in Table 2). The bottom plot shows the same metrics and ranges as the top plot, but with different algorithms. Algorithms for which no classification threshold performed better than classifying all pixels as "non-bloom" (see Table 1) are not shown.

outperforms all others across all the examined metrics, indicating that the various algorithms capture different specific features of blooms. This finding supports the conclusion that improved evaluation approaches based on multiple metrics provide more accurate overall assessment of algorithm strengths and limitations.

3.2. Potential for classifying historical blooms

Overall, our results show that the Improved NIR with SAC algorithm (#7b) performs best in measuring bloom occurrence, spatial extent, and timing (Fig. 2). Visual inspection of maps of peak bloom extent by year moreover indicates that the spatial distribution of the observed blooms agrees well with other remote sensing studies of blooms in Lake Erie (Fig. 3; Fig. S2), and visual comparison to true color images (not shown) does not suggest any conspicuous sources of error (e.g., sun glint or haze). In order to further gauge the potential for hindcasting estimates using the final selected algorithm, we next explore more deeply the years for which estimates are less consistent with the size and timing of blooms observed from MERIS CI.

The final algorithm estimates of peak bloom size perform well at reproducing estimates from MERIS CI, except for notable differences in 2006 and in 2010 (Fig. 4). We find that the large differences in peak bloom size during those two years are primarily the result of missing MERIS data in 2006 and high cloud cover in the Landsat images in 2010. In 2006, Landsat captures a peak during the first week of September, coinciding with a gap in the MERIS data (Fig. 4). This suggests that the peak in the bloom occurred during this gap, and that the Landsat estimate is likely be to more accurate than the MERIS estimate of peak bloom size for that year. The in situ *Microcystis* biovolume observations confirm this, indicating a larger bloom in 2006 than 2005 or 2007. In

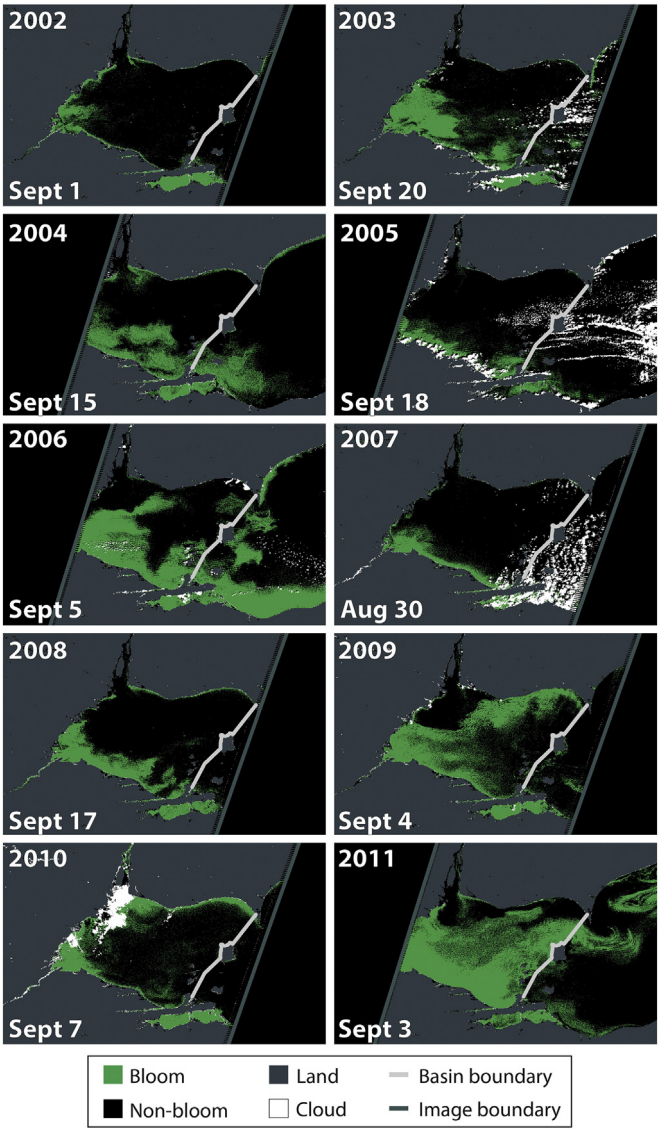

Fig. 3. Bloom image from scene with maximum observed bloom size using the Improved NIR with SAC algorithm (#7b in Table 1) for 2002 to 2011. See Fig. S2 for equivalent MERIS CI images.

2010, on the other hand, clouds obscure the Landsat image coinciding with the MERIS-estimated bloom peak on September 23 (Fig. 4), and there are no other clear images of that bloom. Viewed within the context of estimates from other years, these results suggest that Landsat's long revisit time does not hinder bloom size estimates in nine of the ten cases examined here, and can even improve upon estimates from other sensors when those other data sets have gaps.

Algorithm estimates of the peak bloom date are similarly consistent with those from MERIS CI for most years, but the differences in 2002 and 2005 are much larger (57 and 44 days, respectively; Fig. 4). The large differences for peak bloom timing in 2002 and 2005, however, are mostly due to the fact that those years have little bloom activity overall. The blooms are small, and there is a lack of clear seasonality in the blooms for those years (Fig. 4). As a result, a large difference in matching "bloom peak" timing for those years does not mean that the Landsat algorithm is performing poorly. For years with medium or large blooms (i.e., 2003, 2004, 2006, 2008, 2009, 2010, and 2011), the average difference between Landsat and MERIS peak bloom estimates is just 10 days.

We find that these two sets of years (2006 & 2010, 2002 & 2005) skew the mean error statistics reported for size and timing in the previous section, respectively, and that indeed the final algorithm performs even better than suggested by the overall metrics for the majority of years. Peak bloom size is estimated within 250 km^2 for eight out of ten years, which is a small error relative to the 4110 km^2 area of the western basin as outlined in Fig. 3. Peak bloom date is captured within 8 days and 16 days in five and eight out of the ten years, respectively, a promising result given the 8-day revisit time for Landsat images

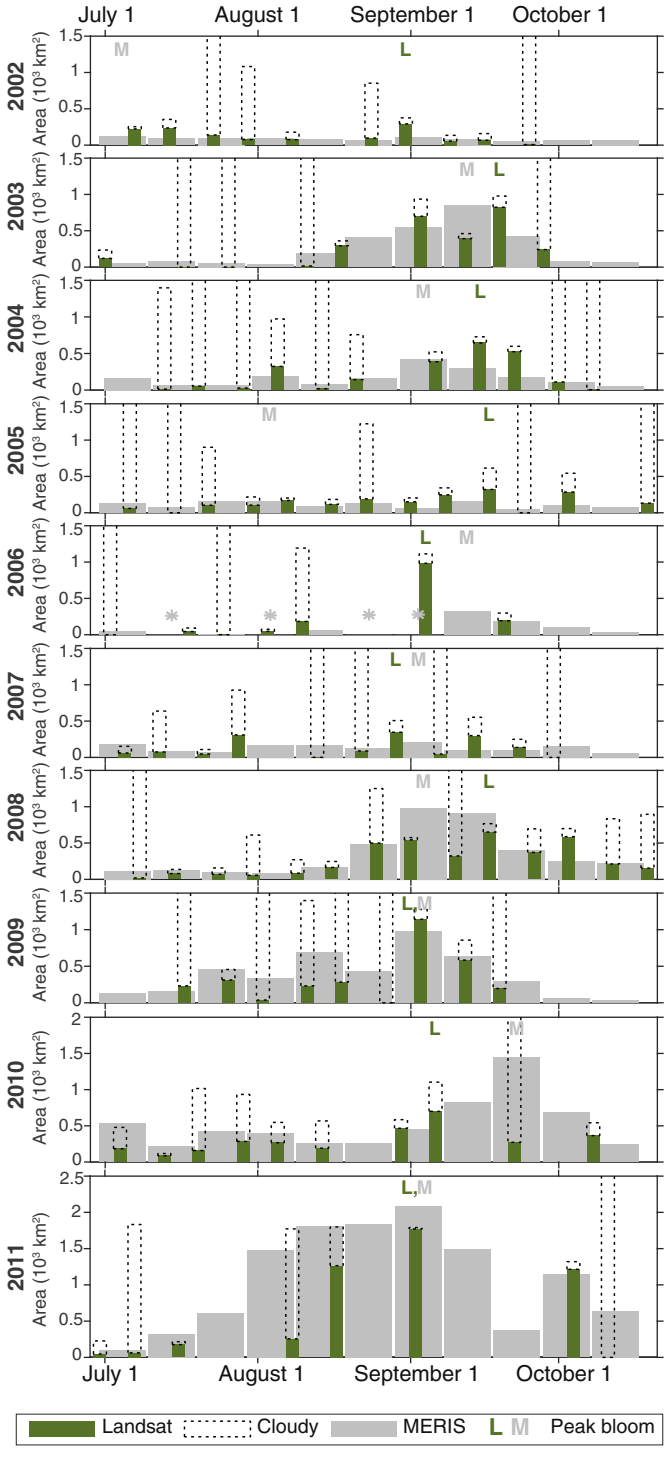

Fig. 4. Comparison of Landsat Improved NIR with SAC algorithm and MERIS CI bloom areas for 2002–2011. The grey bars for MERIS represent estimates of 10-day average areas, while the green bars for Landsat represent estimates based on individual scenes. The empty rectangles represent cloud cover over the western basin in Landsat scenes, where the presence or absence of a bloom could not be directly assessed. Letters 'L' and 'M' at the top of each panel represent the peak observed bloom date for a given year for Landsat and MERIS, respectively. Gaps in MERIS data in 2006 are denoted with an asterisk.

over Lake Erie. Taken together, these findings suggest that differences between Landsat and MERIS are relatively minor and do not preclude the use of Landsat to hindcast blooms.

One potential issue is that Landsat may underestimate bloom size in scenes with high cloud cover, but even this only led to substantial underestimation of the peak bloom area in one out of ten years. For characterization of historical blooms, tagging cloud cover in areas of the western basin adjacent to observed bloom areas could be used as an empirical approach to assess cloud-cover-induced uncertainty in bloom area estimates. This approach is implemented when exploring historical blooms in Section 3.3.

Overall, we find that the accuracy of Landsat-derived bloom estimates is not significantly hindered by its spectral and revisit cycle limitations. We note, however, that the long duration of phytoplankton blooms in this system (up to several months) serves to limit the impact of Landsat's long revisit time, and applicability to systems where blooms are more episodic or where Landsat has a longer revisit time would need to be further evaluated.

3.3. Hindcasted bloom results

Applying the algorithm to the period 1984–2001 reveals the long-term trends in peak bloom magnitude prior to the start of the MERIS and MODIS record (Fig. 5). This expands the historical record of blooms in Lake Erie by nearly two decades, more than doubling the period of record that can be used to understand bloom occurrence and growth. Peak bloom areas decrease gradually from 1984 (760–910 km^2) to 1991 (210–320 km^2), where the ranges represent the observed peak bloom extent (lower bound) and the extent with additional cloudy area tagged manually as potentially obscuring bloom (upper bound) (Fig. 5). The observed decline in the late 1980s is consistent with significant decreases in phytoplankton biomass and chlorophyll-*a* following reductions in phosphorus loading reported in the literature (e.g., Allinger and Reavie, 2013; Makarewicz, 1993; Munawar and Munawar, 1999). The period of declining bloom areas is followed by a large bloom in 1992 (1050–1360 km^2), and then a subsequent period with low bloom activity from 1993 to 2000 (Fig. 5). Starting in 2001, larger blooms become more frequent, and bloom size shows considerably more interannual variability. This progression from low bloom activity to larger and more frequent blooms is consistent with known increases in phytoplankton biomass (e.g., Conroy et al., 2005; Lake Erie LaMP, 2011).

Given that the uncertainty associated with these Landsat-derived peak bloom areas is approximately 350 km^2, and for the majority of years is below 250 km^2 (Section 3.2), these broad temporal trends are expected to be robust. There is substantial cloud cover in the vicinity of the bloom in images of the peak bloom in 1989, 1994, and 1997, but this does not impact the overall trends in bloom size and variability as a whole (Fig. 5; Fig. 6). The overall consistency between the historical trends observed here and existing understanding of long-term trends in Lake Erie phytoplankton blooms (e.g., Conroy et al., 2005; Makarewicz, 1993) also suggests that the algorithm performs well even prior to *Microcystis* sp. dominance starting in the late 1990s.

We note that an initial analysis of the Landsat record also suggested a large outlier "bloom" in 1991 (Fig. 5), which was later determined to represent a false positive. This outlier "bloom" was only observed by one Landsat image out of the twelve that year, and lacked the characteristic spatial gradients observed in other blooms (Fig. 7A, C). Moreover, the difference between the mode in the near-infrared (Band 4) and the shortwave-infrared (Band 5) for "bloom" pixels in this image was much greater than in images in other years (Fig. 7B, D). This behavior is consistent with high concentrations of fine aerosol particles in the atmosphere, which would have scattered light disproportionately in the shortwave-infrared region relative to the near-infrared (Alexandrov et al., 2008). This differential scattering behavior is also more common in continental aerosols (as would be expected over Lake Erie) versus maritime or oceanic aerosols (Wang, 2007), further lending credence to this hypothesis. For completeness, we have shown the bloom area implied by the outlier image for 1991 in Fig. 5, in addition to what we believe to be a better estimate of the true bloom size for that year based on the remaining Landsat images.

The outlier "bloom" in 1991 points to the potential impact of atmospheric conditions on algorithm performance in some circumstances. Given our use of TOA reflectance images as input and implementation of a relatively simple atmospheric correction approach, meteorological conditions that affect factors such as water vapor concentrations and aerosol optical thickness could add uncertainty to bloom areal estimates. In the case of Lake Erie, visual observation of historical bloom images did not indicate substantial atmospheric effects (except for the 1991 outlier), and comparison with historical literature suggests strong coherence in bloom estimates despite the implementation of only a simple atmospheric correction. The need for a more rigorous approach for quantifying atmospheric effects (e.g., Barnes et al., 2014) may need to be re-examined for application to other lakes.

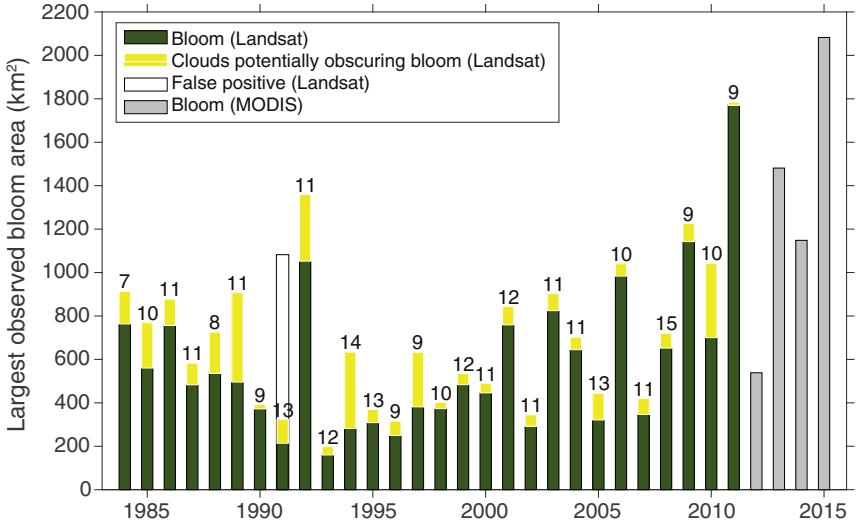

Fig. 5. Largest observed bloom areas estimated via Landsat Improved NIR with SAC (1984–2011) and MODIS CI (2012–2015) algorithms. Bars indicate observed bloom area (green) and additional cloud-obscured potential bloom area (yellow). The number on top of each bar indicates the number of Landsat images available in July to October each year.

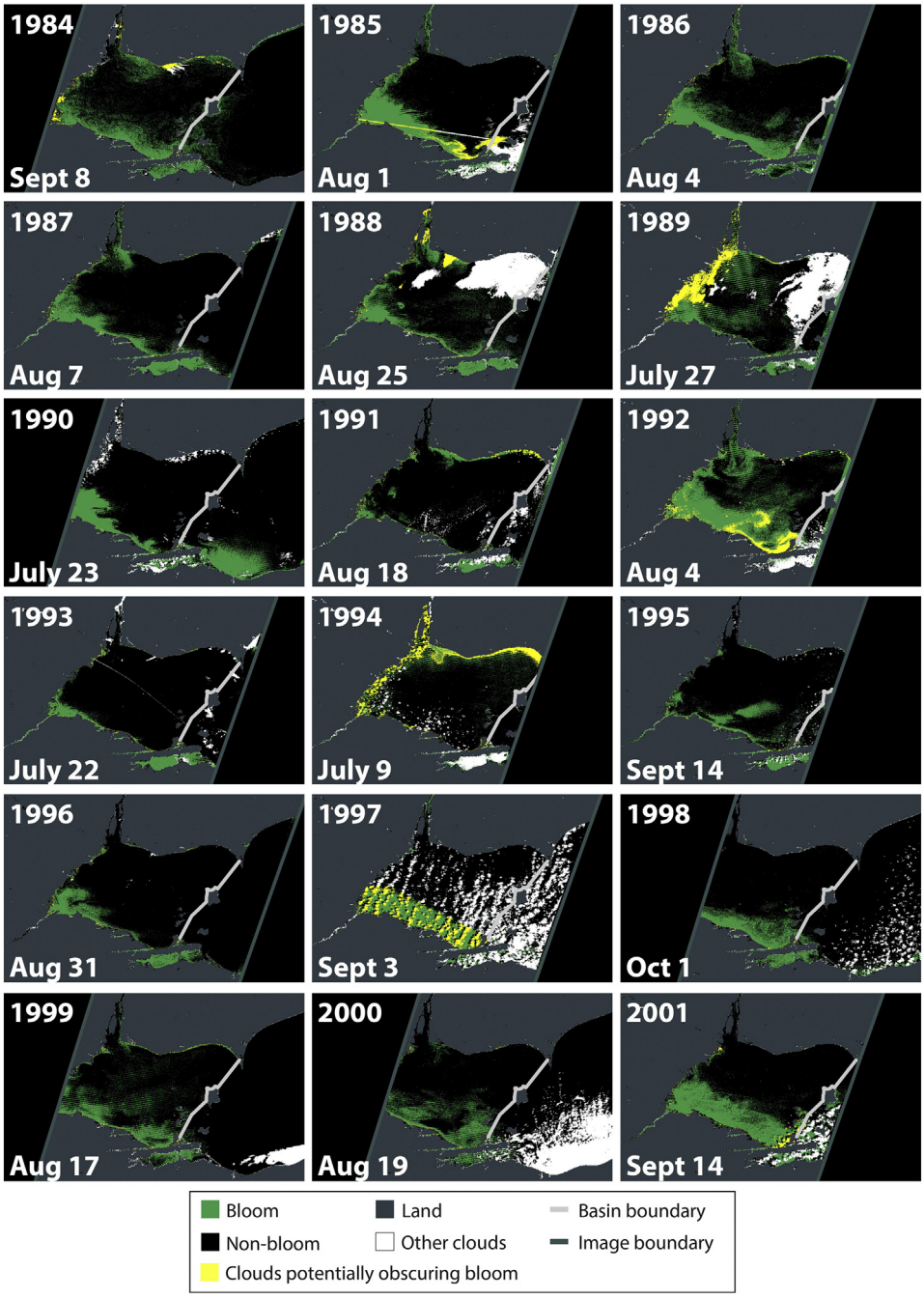

Fig. 6. Images of largest observed blooms for 1984–2001 displaying bloom-classified values in green (bloom) and black (non-bloom). Cloud pixels are shown in yellow (clouds in the vicinity of bloom) and white (other clouds).

4. Conclusions

In this study, we assess the effectiveness of Landsat observations for identifying historical phytoplankton blooms in a shallow eutrophic lake, by comparing algorithms based on Landsat imagery to in situ *Microcystis* biovolume and remotely-sensed MERIS CI observations. We evaluate an initial set of eleven Landsat algorithms for 2002–2011 using a novel evaluation procedure based on accurate identification of bloom occurrence, spatial extent, and timing. We find that a near-infrared threshold algorithm with simple atmospheric correction (Algorithm #7) performs best, and further optimize this algorithm with a secondary greenness filter based on a hue threshold (Algorithm #7b). The final algorithm outperforms the other 11 evaluated algorithms across nine of the ten evaluation metrics.

The novel evaluation procedure presented here also demonstrates the utility of a multi-metric approach for assessing algorithm performance. The procedure intercompares algorithms based on their accuracy at identifying different aspects of the annual blooms (e.g., bloom occurrence versus bloom timing), yielding a more accurate assessment of algorithm strengths and limitations relative to previous approaches based on individual correlation statistics and single validation data sets.

For the final selected algorithm (#7b), exploration of potential uncertainties for hindcasting historical blooms reveals that differences between Landsat and MERIS are encouragingly minor, with

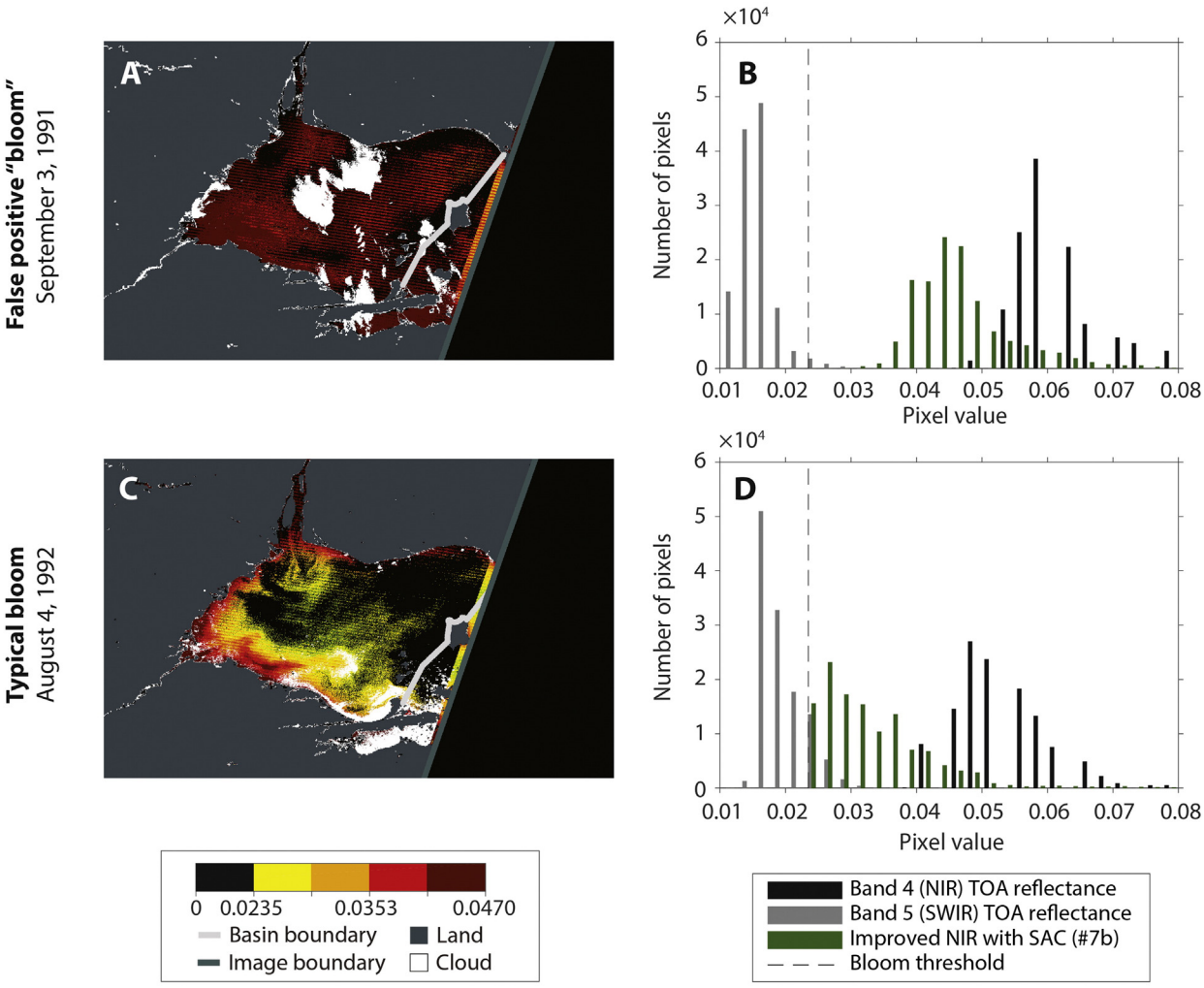

Fig. 7. Comparison between false positive "bloom" detected in 1991 (A, B) and a typical bloom from 1992 (C, D). Raw (i.e., before thresholding) Improved NIR with SAC (#7b) values (A, C) show that the 1991 image lacks the characteristic spatial gradient of blooms. This behavior is due to high pixel values in the near-infrared (NIR) band compared to the shortwave-infrared (SWIR) band relative to images of actual blooms (B, D). Only "bloom" pixels (i.e., Improved NIR with SAC (#7b) value above the bloom threshold of 0.0235) are shown in panels B and D.

underestimation of bloom size in cloudy scenes being the only potential issue of concern. Overall, we find that the final algorithm can provide accurate estimates of bloom extent and timing during 2002–2011, and that the algorithm can be expected to provide reliable information for the period preceding MERIS CI observations.

Applying the best algorithm to scenes for 1984–2001, we present new quantitative information on historical Lake Erie phytoplankton blooms, expanding remotely-sensed information on blooms in Lake Erie by nearly two decades. Bloom size declined in the late 1980s, stayed relatively low in the 1990s, and increased thereafter. These new data make it possible to assess long-term trends in phytoplankton blooms and provide new insights about historical bloom spatial distributions and timing.

Overall, our findings demonstrate the potential for using Landsat to generate long-term data on phytoplankton blooms in similar systems. For example, the novel evaluation procedure presented here could be used to identify appropriate Landsat algorithms in other systems where the MERIS CI algorithm has been successfully applied for water quality monitoring (e.g., Lunetta et al., 2015), especially given evidence suggesting that the final algorithm selected here performed well during periods when different phytoplankton species were dominant. Future work should also examine factors affecting Landsat's ability to identify bloom features for other lakes, including examining the impact of bloom duration on accuracy and the degree to which estimates at other locations may be more strongly impacted by atmospheric effects. Application to other systems would also reveal the extent to which this approach can provide new insights about historical trends in phytoplankton blooms across locations.

Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF) under Grants 1039043 and 1313897. Additional support for J.C. Ho was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) under a Postgraduate Scholarship-Doctoral award (PGSD3-438855-2013). Funding sources had no involvement in the conduct of the research or the preparation of the article. Landsat 5 imagery is available from the U.S. Geological Survey and through Google Earth Engine. Chlorophyll-*a* measurements are available from the U.S. Environmental Protection Agency and Environment Canada. Buoy water temperature information are available from the NOAA National Data Buoy Center. The authors thank Eva Sinha and Yoichi Shiga for comments on early manuscripts and figures, and thank Tyler Erickson, other Google Earth Engine staff, and users on the Earth Engine developer's list for support on Google Earth Engine. The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2016.12.013.

References

Alexandrov, M.D., Lacis, A.A., Carlson, B.E., Cairns, B., 2008. Characterization of atmospheric aerosols using MFRSR measurements. J. Geophys. Res. Atmos. 113 (8):1–23. http://dx.doi.org/10.1029/2007JD009388.

Allinger, L., Reavie, E., 2013. The ecological history of Lake Erie as recorded by the phytoplankton community. J. Great Lakes Res. 39 (3):365–382. http://dx.doi.org/10.1016/j.jglr.2013.06.014.

Barnes, B.B., Hu, C., Holekamp, K.L., Blonski, S., Spiering, B.A., Palandro, D., Lapointe, B., 2014. Use of Landsat data to track historical water quality changes in Florida Keys marine environments. Remote Sens. Environ. 140:485–496. http://dx.doi.org/10.1016/j.rse.2013.09.020.

Bertani, I., Obenour, D.R., Steger, C.E., Stow, C.A., Gronewold, A.D., Scavia, D., 2016. Probabilistically assessing the role of nutrient loading in harmful algal bloom formation in western Lake Erie. J. Great Lakes Res. 42 (6):1184–1192. http://dx.doi.org/10.1016/j.jglr.2016.04.002.

Bridgeman, T.B., Chaffin, J.D., Filbrun, J.E., 2013. A novel method for tracking western Lake Erie Microcystis blooms, 2002–2011. J. Great Lakes Res. 39 (1):83–89. http://dx.doi.org/10.1016/j.jglr.2012.11.004.

Carroll, R.J., Ruppert, D.J., 1988. Transformation and Weighting in Regression. Chapman and Hall, New York.

Chaffin, J.D., Bridgeman, T.B., Heckathorn, S.A., Mishra, S., 2011. Assessment of Microcystis growth rate potential and nutrient status across a trophic gradient in western Lake Erie. J. Great Lakes Res. 37 (1):92–100. http://dx.doi.org/10.1016/j.jglr.2010.11.016.

Chander, G., Markham, B.L., Barsi, J.A., 2007. Revised landsat-5 thematic mapper radiometric calibration. IEEE Geosci. Remote Sens. Lett. 4 (3):490–494. http://dx.doi.org/10.1109/LGRS.2007.898285.

Chander, G., Markham, B.L., Helder, D.L., 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM +, and EO-1 ALI sensors. Remote Sens. Environ. 113 (5):893–903. http://dx.doi.org/10.1016/j.rse.2009.01.0s07.

Conroy, J.D., Kane, D.D., Dolan, D.M., Edwards, W.J., Charlton, M.N., Culver, D.A., 2005. Temporal trends in Lake Erie plankton biomass: roles of external phosphorus loading and Dreissenid mussels. J. Great Lakes Res. 31:89–110. http://dx.doi.org/10.1016/S0380-1330(05)70307-5.

Dekker, A.G., 1993. Detection of Optical Water Quality Parameters for Eutrophic Waters by High Resolution Remote Sensing. VU University Amsterdam.

Duan, H., Ma, R., Xu, X., Kong, F., Zhang, S., Kong, W., ... Shang, L., 2009. Two-decade reconstruction of algal blooms in China's Lake Taihu. Environ. Sci. Technol. 43 (10), 3522–3528.

Duan, H., Zhang, Y., Zhang, B., Song, K., Wang, Z., 2007. Assessment of chlorophyll-*a* concentration and trophic state for Lake Chagan using Landsat TM and field spectral data. Environ. Monit. Assess. 129 (1–3):295–308. http://dx.doi.org/10.1007/s10661-006-9362-y.

Environment Canada, 2012. Great Lakes Surveillance Program. Retrieved November 16, 2012, from. http://www.ec.gc.ca/scitech/default.asp?lang=en&n=3F61CB56-1.

EPA GLNPO, 2012. Great Lakes Environmental Database (GLENDA). Retrieved February 20, 2010, from. http://www.epa.gov/greatlakes/monitoring/data_proj/glenda/.

ESA, 2015. Baseline Atmospheric Correction. Retrieved October 27, 2015, from. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-olci/level-2/baseline-atmospheric-correction.

ESA, 2016. Envisat Instruments - MERIS. Retrieved December 23, 2016, from. https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat/instruments/meris.

Gitelson, A., Garbuzov, G., Szilagyi, F., Mittenzwey, K.-H., Karnieli, A., Kaiser, A., 1993. Quantitative remote sensing methods for real-time monitoring of inland waters quality. Int. J. Remote Sens. 14 (7):1269–1295. http://dx.doi.org/10.1080/01431169308953956.

Google, 2016. Earth Engine. Retrieved November 24, 2015 from https://earthengine.google.com/.

Gordon, H.R., 1978. Removal of atmospheric effects from satellite imagery of the oceans. Appl. Opt. 17 (10):1631. http://dx.doi.org/10.1364/AO.17.001631.

Hadjimitsis, D.G., Clayton, C.R.I., Hope, V.S., 2004. An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs. Int. J. Remote Sens. 25 (18):3651–3674. http://dx.doi.org/10.1080/01431160310001647993.

Haydn, R., Dalke, G.W., Henkel, J., Bare, J.E., 1982. Application of the IHS color transform to the processing of multisensor data and image enhancement. Proceedings of the International Symposium on Remote Sensing of Arid and Semiarid Lands, Cairo, Egypt 599–616.

Ho, J.C., Michalak, A.M., 2015. Challenges in tracking harmful algal blooms: A synthesis of evidence from Lake Erie. J. Great Lakes Res. 41 (2):317–325. http://dx.doi.org/10.1016/j.jglr.2015.01.001.

Irish, R.R., 2000. Landsat 7 automatic cloud cover assessment. AeroSense 2000 (4049): 348–355. http://dx.doi.org/10.1117/12.410358.

Kahru, M., Elmgren, R., 2014. Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. Biogeosciences 11 (13):3619–3633. http://dx.doi.org/10.5194/bg-11-3619-2014.

Kane, D.D., Ludsin, S.A., Briland, R.D., Culver, D.A., Munawar, M., 2015. Ten + years gone: continued degradation of offshore planktonic communities in U.S. waters of Lake Erie's western and central basins (2003–2013). J. Great Lakes Res. 41 (3):930–933. http://dx.doi.org/10.1016/j.jglr.2015.06.002.

Kasich, J.R., Taylor, M., Butler, C.W., 2014. Public Water System Harmful Algal Bloom Response Strategy. Columbus, OH.

Lake Erie LaMP, 2011. Lake Erie Binational Nutrient Management Strategy: Protecting Lake Erie by Managing Phosphorus (Prepared by the Lake Erie LaMP Work Group Nutrient Management Task Group).

Lathrop, R., 1992. Landsat thematic mapper monitoring of turbid inland water quality. Photogramm. Eng. Remote. Sens. 58.

Lunetta, R.S., Schaeffer, B.A., Stumpf, R.P., Keith, D., Jacobs, S.A., Murphy, M.S., 2015. Evaluation of cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. Remote Sens. Environ. 157:24–34. http://dx.doi.org/10.1016/j.rse.2014.06.008.

Maccoux, M.J., Dove, A., Backus, S.M., Dolan, D.M., 2016. Total and Soluble Reactive Phosphorus Loadings to Lake Erie. J. Great Lakes Res. 42 (6):1151–1165. http://dx.doi.org/10.1016/j.jglr.2016.08.005.

Makarewicz, J.C., 1993. Phytoplankton biomass and species composition in Lake Erie, 1970 to 1987. J. Great Lakes Res. 19:258–274. http://dx.doi.org/10.1016/S0380-1330(93)71216-2.

Matthews, M.W., 2011. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. Int. J. Remote Sens. 32 (21):6855–6899. http://dx.doi.org/10.1080/01431161.2010.512947.

Mayo, M., Gitelson, A., Yacobi, Y., Ben-Avraham, Z., 1995. Chlorophyll distribution in Lake Kinneret determined from Landsat Thematic Mapper data. Int. J. Remote Sens. 16 (1): 175–182. http://dx.doi.org/10.1080/01431169508954386.

Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., ... Zagorski, M.A., 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proc. Natl. Acad. Sci. U. S. A. 110 (16):6448–6452. http://dx.doi.org/10.1073/pnas.1216006110.

Mouw, C.B., Greb, S., Aurin, D., DiGiacomo, P.M., Lee, Z., Twardowski, M., ... Craig, S.E., 2015. Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions. Remote Sens. Environ. 160:15–30. http://dx.doi.org/10.1016/j.rse.2015.02.001.

Munawar, M., Munawar, I.F., 1999. The changing phytoplankton biomass and its composition in Lake Erie: a lake wide comparative analysis. In: Munawar, M., Edsall, T., Munawar, I.F. (Eds.), State of Lake Erie - Past. Present and Future. Backhuys Publisers, Leiden, The Netherlands, pp. 125–154.

NASA, 2016a. MODIS - Moderate Resolution Imaging Spectroradiometer. Retrieved March 10, 2016 from http://modis.gsfc.nasa.gov/.

NASA, 2016b. Landsat 8. Retrieved October 12, 2016 from http://landsat.usgs.gov/landsat8.php.

Neukermans, G., Ruddick, K., Bernard, E., Ramon, D., Nechad, B., Deschamps, P.Y., 2009. Mapping total suspended matter from geostationary satellites: a feasibility study with SEVIRI in the Southern North Sea. Opt. Express 17 (16):14029–14052. http://dx.doi.org/10.1364/OE.17.014029.

NOAA NDBC, 2014. Station 45005. Retrieved June 8, 2014, from. http://www.ndbc.noaa.gov/station_page.php?station=45005.

Obenour, D.R., Gronewold, A.D., Stow, C.A., Scavia, D., 2014. Using a Bayesian hierarchical model to improve Lake Erie cyanobacteria bloom forecasts. Water Resour. Res. 50. http://dx.doi.org/10.1002/2014WR015616.

Paerl, H.W., Huisman, J., 2009. Climate change: A catalyst for global expansion of harmful cyanobacterial blooms. Environ. Microbiol. Rep. 1 (1):27–37. http://dx.doi.org/10.1111/j.1758-2229.2008.00004.x.

Palmer, S.C.J., Hunter, P.D., Lankester, T., Hubbard, S., Spyrakos, E.N., Tyler, A., Tóth, V.R., 2015. Validation of envisat MERIS algorithms for chlorophyll retrieval in a large, turbid and optically-complex shallow lake. Remote Sens. Environ. 157:158–169. http://dx.doi.org/10.1016/j.rse.2014.07.024.

Pozdnyakov, D., Shuchman, R., Korosov, A., Hatt, C., 2005. Operational algorithm for the retrieval of water quality in the Great Lakes. Remote Sens. Environ. 97 (3): 352–370. http://dx.doi.org/10.1016/j.rse.2005.04.018.

Robarts, R.D., Zohary, T., 1987. Temperature effects on photosynthetic capacity, respiration, and growth rates of bloom-forming cyanobacteria. N. Z. J. Mar. Freshw. Res. 21 (3):391–399. http://dx.doi.org/10.1080/00288330.1987.9516235.

Sass, G.Z., Creed, I.F., Bayley, S.E., Devito, K.J., 2007. Understanding variation in trophic status of lakes on the boreal plain: A 20-year retrospective using Landsat TM imagery. Remote Sens. Environ. 109 (2):127–141. http://dx.doi.org/10.1016/j.rse.2006.12.010.

Sayers, M.J., Grimm, A.G., Shuchman, R.a., Deines, A.M., Bunnell, D.B., Raymer, Z.B., ... Mychek-Londer, J., 2015. A new method to generate a high-resolution global distribution map of lake chlorophyll. Int. J. Remote Sens. 36 (7):1942–1964. http://dx.doi.org/10.1080/01431161.2015.1029099.

Scavia, D., DePinto, J.V., Bertani, I., 2016. A multi-model approach to evaluationg target phosphorus loads for Lake Erie. J. Great Lakes Res. 42 (6):1139–1150. http://dx.doi.org/10.1016/j.jglr.2016.09.007.

Schindler, D.W., Carpenter, S.R., Chapra, S.C., Hecky, R.E., Orihel, D.M., 2016. Reducing phosphorus to curb Lake eutrophication is a success. Environ. Sci. Technol. 50: 8923–8929. http://dx.doi.org/10.1021/acs.est.6b02204.

Stumpf, R.P., Johnson, L.T., Wynne, T.T., Baker, D.B., 2016. Forecasting annual cyanobacterial bloom biomass to inform management decisions in Lake Erie. J. Great Lakes Res. 42 (6):1174–1183. http://dx.doi.org/10.1016/j.jglr.2016.08.006.

Stumpf, R.P., Pennock, J.R., 1989. Calibration of a general optical equation for remote sensing of suspended sediments in a moderately turbid estuary. J. Geophys. Res. 94 (C10): 14363. http://dx.doi.org/10.1029/JC094iC10p14363.

Stumpf, R.P., Tyler, M.A., 1988. Satellite detection of bloom and pigment distributions in estuaries. Remote Sens. Environ. 24 (3):385–404. http://dx.doi.org/10.1016/0034-4257(88)90014-4.

Stumpf, R.P., Wynne, T.T., Baker, D.B., Fahnenstiel, G.L., 2012. Interannual variability of cyanobacterial blooms in Lake Erie. PLoS One 7 (8), e42444. http://dx.doi.org/10.1371/journal.pone.0042444.

Tebbs, E.J., Remedios, J.J., Harper, D.M., 2013. Remote sensing of chlorophyll-*a* as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline-alkaline, flamingo lake, using Landsat ETM+. Remote Sens. Environ. 135:92–106. http://dx.doi.org/10.1016/j.rse.2013.03.024.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2):127–150. http://dx.doi.org/10.1016/0034-4257(79)90013-0.

Tyler, A.N., Svab, E., Preston, T., Présing, M., Kovács, W.A., 2006. Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment. Int. J. Remote Sens. 27 (8):1521–1537. http://dx.doi.org/10.1080/01431160500419311.

USGS, 2016. Landsat Missions. Retrieved December 23, 2016 from http://landsat.usgs.gov/.

Vincent, R.K., Bullerjahn, G., McKay, R., Wicks, J., Beall, B., Sridhar, B., ... Loew, T., 2011. Monitoring of Great Lakes Water Quality with Remote Sensing. Second Project Progress Report. Bowling Green, OH.

Vincent, R.K., Qin, X., McKay, R.M.L., Miner, J., Czajkowski, K., Savino, J., Bridgeman, T., 2004. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. Remote Sens. Environ. 89 (3):381–392. http://dx.doi.org/10.1016/j.rse.2003.10.014.

Wang, M., 2007. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: Simulations. Appl. Opt. 46 (9):1535–1547. http://dx.doi.org/10.1364/AO.46.001535.

Wang, M., Shi, W., 2007. The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing. Opt. Express 15 (24), 15722–15733.

Wilson, E.K., 2014. Danger from microcystins in Toledo water unclear. Chem. Eng. News 92 (32), 9.

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Warner, R.A., Tester, P.A., Dyble, J., Fahnenstiel, G.L., 2008. Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. Int. J. Remote Sens. 29 (12), 3665–3672.

Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Dyble, J., 2010. Characterizing a cyanobacterial bloom in western Lake Erie using satellite imagery and meteorological data. Limnol. Oceanogr. 55 (5):2025–2036. http://dx.doi.org/10.4319/lo.2010.55.5.2025.

Yacobi, Y.Z., Gitelson, A.A., Mayo, M., 1995. Remote sensing of chlorophyll in Lake Kinneret using high spectral resolution radiometer and Landsat thematic mapper: Spectral features of reflectance and algorithm development. J. Plankton Res. 17 (11), 2155–2173.