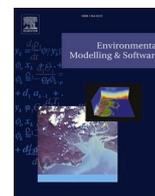




Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: <http://www.elsevier.com/locate/envsoft>

Probabilistic forecast of microcystin toxin using satellite remote sensing, *in situ* observations and numerical modeling

Qianqian Liu^{a,b,*}, Mark D. Rowe^c, Eric J. Anderson^c, Craig A. Stow^c, Richard P. Stumpf^d, Thomas H. Johengen^b

^a Department of Physics and Physical Oceanography, University of North Carolina Wilmington, Wilmington, NC, 28403, USA

^b Cooperative Institute for Great Lakes Research (CIGLR), University of Michigan, Ann Arbor, MI, 48018, USA

^c Great Lakes Environmental Research Laboratory, National Oceanic and Atmospheric Administration, Ann Arbor, MI, 48018, USA

^d National Centers for Coastal Ocean Science, National Oceanic and Atmospheric Administration, Silver Spring, MD, 20910, USA

ARTICLE INFO

Keywords:

Lake Erie
Cyanobacterial harmful algal bloom
Microcystin
Short term forecast
Probabilistic forecast

ABSTRACT

Lake Erie has experienced a resurgence of cyanobacterial harmful algal blooms (CHABs) since the early 2000's dominated by *Microcystis aeruginosa*, which produce toxins known as microcystins. We develop an approach to predict the spatially- and temporally-resolved probability of exceeding a public health advisory (PHA) level (6 µg/L) of microcystins in the western basin of Lake Erie that would be suitable for use in a forecast system, consisting of 1) an existing HAB chlorophyll forecast system, 2) a toxin-chlorophyll-a relationship that is updated weekly from observations, and 3) a statistical model relating observed relative frequency of exceeding the PHA to model predictions over a hindcast period. We evaluate the system's performance and the system's useful level of skill. This novel approach to a CHAB toxin forecast system could provide a decision support tool to Lake Erie stakeholders, and the approach may be adapted to other systems.

1. Introduction

Harmful algal blooms (HABs) occur when phytoplankton grow rapidly and produce toxins that can have adverse effects on people and aquatic ecosystems. Lake Erie, the most productive, warm and shallow of the Laurentian Great Lakes of North America (Fig. 1a), has experienced a re-emergence of cyanobacterial harmful algal blooms (CHABs) since the early 2000's. CHABs are linked to excess nutrient input from agricultural land and urbanized areas, and may intensify with climate change due to their competitive ability to grow at warm temperatures (O'Neil et al., 2012). The most severe blooms in Lake Erie originate in the shallow western basin (Fig. 2), which receives runoff from a large agricultural watershed through the Maumee River (Kane et al., 2014). In 2014, the city of Toledo issued a "do not use" order for their drinking water system due to contamination with microcystins that affected 0.5 million people for 2.5 days (Brittain et al., 2000; Michalak et al., 2013; Wynne and Stumpf, 2015; Levy, 2017; Steffen et al., 2017). Furthermore, in 2015, Lake Erie experienced the most extensive algal bloom on record (Stumpf et al., 2016a). These events motivate continued research into strategies for reducing the extent and toxicity of CHABs and the

development of decision support systems that can help mitigate their negative impacts.

CHABs in Lake Erie are dominated by *Microcystis aeruginosa*, which produces microcystins, a class of cyclic peptides (Rinta-Kanto et al., 2009). Microcystins are hepatotoxins and can cause liver and kidney disease in humans who have been exposed through drinking water or involuntarily ingesting the toxins when swimming in contaminated water. To protect public health, Ohio EPA established microcystin drinking water threshold concentrations for children under 6 years of age and sensitive population as 0.3 µg/L, and 1.6 µg/L for children 6 and older including adults. Ohio also implemented guidelines for recreational contact. A Public Health Advisory (PHA) Microcystin-LR concentration of 6 µg/L was established, above which swimming and wading are not recommended and water should not be swallowed and surface scum should be avoided. A No Contact Advisory (NCA) Microcystin-LR concentration of 20 µg/L was established, above which the public should avoid all contact with the water. Based on peer-reviewed and published science, in May 2019, USEPA also recommended new microcystin public health threshold concentrations as 8 µg/L (U.S. EPA, 2019). Therefore, predictions of when and where the

* Corresponding author. Department of Physics and Physical Oceanography, University of North Carolina Wilmington, Wilmington, NC, 28403, USA.
E-mail address: liuq@uncw.edu (Q. Liu).

<https://doi.org/10.1016/j.envsoft.2020.104705>

Received 25 November 2019; Received in revised form 19 March 2020; Accepted 23 March 2020

Available online 29 March 2020

1364-8152/© 2020 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

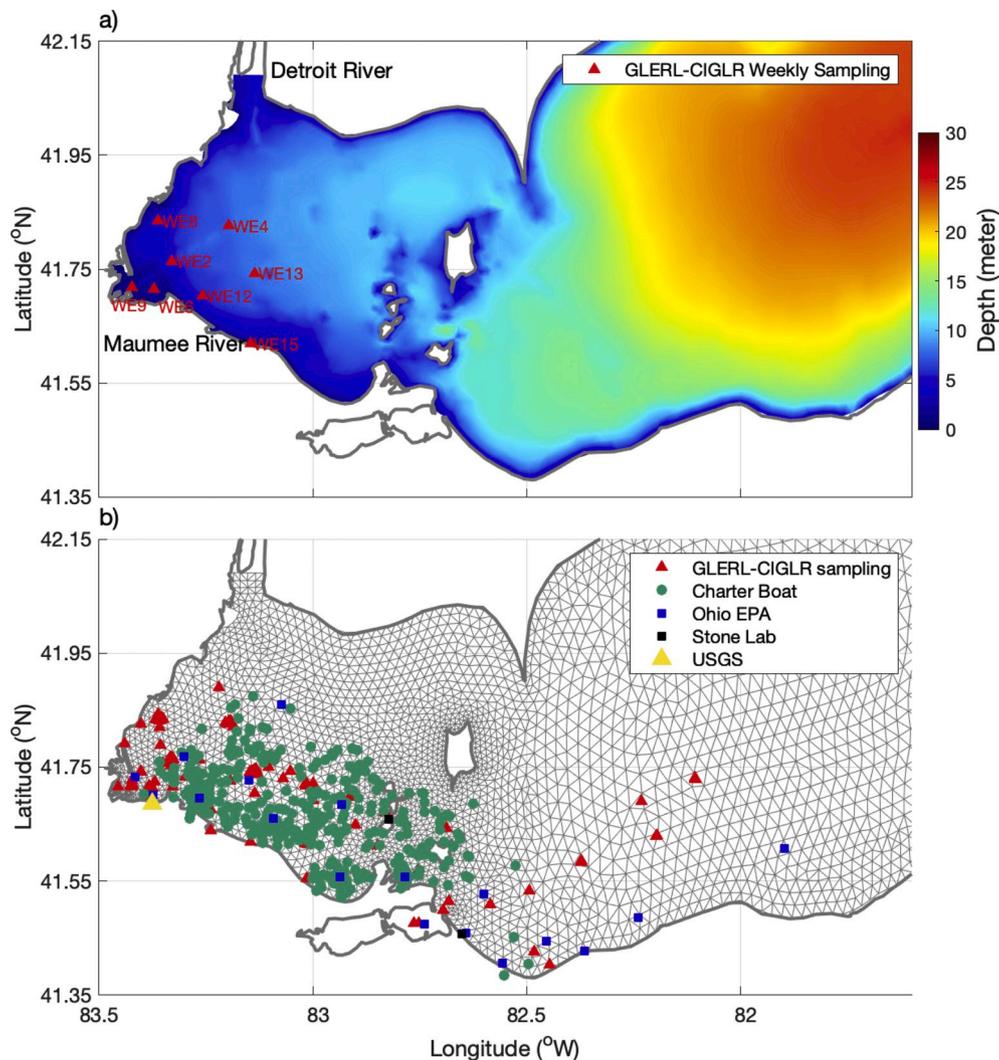


Fig. 1. a) Bathymetry of western Lake Erie, and locations of stations sampled weekly by GLERL and CIGLR. b) FVCOM model configuration of western portion of Lake Erie, and locations of all microcystin data used in the forecast system. The spatial resolution of FVCOM grid is ~ 2 km in the central basin, 1.5 km in the western basin, and 0.5 km in Maumee Bay and the islands.

microcystin concentration will exceed advisory concentrations are potentially useful to support decision making by stakeholders including public drinking water systems, and recreational users affected by CHABs.

Even though cyanobacteria are the source of microcystins, the association between cyanobacteria densities and microcystins is not consistent (Stumpf et al., 2016b). Toxic and non-toxic strains of *Microcystis* exist within the same CHAB, and toxin production varies over time (Hollister and Kreakie, 2016; Davis et al., 2009). In addition, other non-toxic forms of phytoplankton exist in Lake Erie; thus, chlorophyll-a may be present in the absence of either cyanobacteria or associated microcystin toxin. However, the highest toxin concentrations in Lake Erie occur in association with increased *Microcystis* biomass in blooms and in surface scums, conditions also associated with high concentrations of chlorophyll-a (Rinta-kanto et al., 2005; 2009). Therefore, when toxin concentration is high, the majority of the phytoplankton biomass is due to cyanobacteria (with detailed analysis shown in Section 3.2), and we can potentially use chlorophyll-a concentration to predict the probability of microcystin concentration exceeding advisory levels under these conditions (Hollister and Kreakie, 2016; Stumpf et al., 2016b).

Previous forecast research has focused on CHAB abundance and spatial distributions, and did not attempt to predict toxin

concentrations. Wynne et al. (2013) developed a short-term forecast of HAB location and transport that combined satellite remote sensing data with a hydrodynamic model. Stumpf et al. (2016a) developed a seasonal forecast of maximum HAB extent, based on statistical association between spring phosphorus loads from the Maumee River and maximum HAB extent. Rowe et al. (2016) developed a short-term (~ 5 -day) HAB forecast model, known as the Lake Erie HAB Tracker, which included the vertical distribution of buoyant *Microcystis* colonies in the water column. The HAB location in the short-term forecast is initiated from satellite-derived cyanobacterial index data (Wynne et al., 2008), and movement is predicted using modeled currents from the NOAA Lake Erie Operational Forecast System (LEOFS; Kelley et al., 2018), an application of the Finite Volume Community Ocean Model (FVCOM; Chen et al., 2003), and a Lagrangian particle tracking model (Churchill et al., 2011; Huret et al., 2007). The Experimental Lake Erie HAB Tracker model has provided daily nowcast and forecast information in July–October of 2016–2019 (www.glerl.noaa.gov/res/HABs_and_Hypoxia/habTracker.html).

We present and assess an approach to forecast the spatial and temporal probability of microcystins exceeding various advisory guidelines in Lake Erie. The toxicity probability forecast was achieved through a combination of a 1) HAB Tracker model (Rowe et al., 2016), 2) linear regression model to correlate extracted chlorophyll-a and microcystin



Fig. 2. An aerial view of a boat in a western Lake Erie harmful algal bloom on August 3, 2019, showing variable spatial distribution of algal biomass, including variation caused by presence of the boat. Photo credit: Zachary Haslick, Aerial Associates Photography Inc.

concentrations derived from weekly sampling in western Lake Erie (section 2), and 3) a functional relationship giving the probability of *in situ* microcystin measurements exceeding advisory levels, conditional upon model-predicted microcystin levels, based on data compiled over a hindcast period (section 3). Because in Lake Erie the microcystin: chlorophyll-a ratio varies temporally over the bloom season (July to October) more so than spatially on a given day, the forecast system assumes that the relationship between microcystins and chlorophyll-a is spatially uniform, but varies over time according to the most recent weekly sampling data. We trained the model over 2014 to 2016 and assessed the model in 2017, a year that was outside the calibration period.

2. Materials and methods

2.1. In situ observations

Starting in 2008, water samples were collected between weekly to bi-

weekly at eight stations by the Cooperative Institute of Great Lakes Research (CIGLR) and the National Oceanic and Atmospheric Administration (NOAA) Great Lakes Environmental Research Laboratory (GLERL) in western Lake Erie (Fig. 1; <https://accession.nodc.noaa.gov/0187718>). Many of the stations were sampled consistently from 2008 to 2017. Sampling usually occurred from June to September, covering the peak growing season for cyanobacteria. For this study, we used paired chlorophyll-a and microcystin data from a given sample. To obtain a representative sample in the presence of surface scums, a Niskin bottle was inserted vertically into the water until it was submerged. To avoid surface scum accumulations, the top of the Niskin bottle was submerged about 0.5 m below the surface and collected an integrated surface sample between 0.5 and 1.25 m in depth. Samples were kept cool and transported to NOAA-GLERL for analysis of, including, chlorophyll-a, particulate (intracellular) and dissolved (extracellular) microcystins. The weekly samples were collected on Monday, with results for chlorophyll-a and microcystins typically available on Wednesday. Particulate microcystins were measured by filtering water samples onto a 3 µm polycarbonate membrane, which were stored at -20 °C until analysis. Samples were lysed using three freeze/thaw cycles followed by addition of the QuikLyse reagents (Abraxis LLC, Warminster, PA). The concentration of microcystins were measured using a microcystin enzyme-linked immunosorbent assay (ELISA; Abraxis LLC) and reported as microcystin-LR equivalents. The detailed method for microcystins was described by Steffen et al. (2017).

In addition, we used western Lake Erie toxin data from multiple organizations to assess the toxin forecast system hindcast, which we refer to as the toxin database. The toxin database included all GLERL-CIGLR weekly sampling (red stars in Fig. 1b), Ohio EPA data (blue squares) including Ohio EPA Lake Erie Ambient Stations and Lake Erie Public Water System Intakes data (http://wwwapp.epa.ohio.gov/gis/mapportal/HAB_Monitoring.html), Ohio State University Stone Lab sampling (black squares) and Charter boat captains sampling data (<https://ohioseagrant.osu.edu/research/live/water>; green circles), and USGS data at Maumee Bay state Park Cove 3 in Ohio (yellow star). The Charter boat captains' data were sampled using a 2-m long tube lowered from the surface, which is referred to as an "integrated" sample. Model results were compared with observations over a comparable depth for surface or integrated samples.

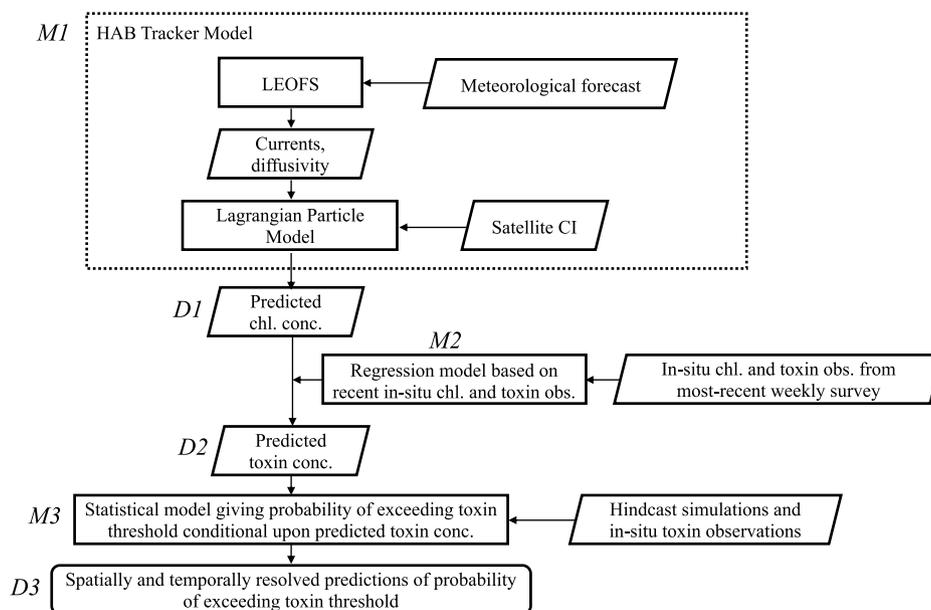


Fig. 3. Flow chart of experimental Lake Erie HAB toxin forecast system. Model processes are denoted by rectangles, while data are denoted by parallelograms. M1 to M3, and D1 to D3 are explained in the text.

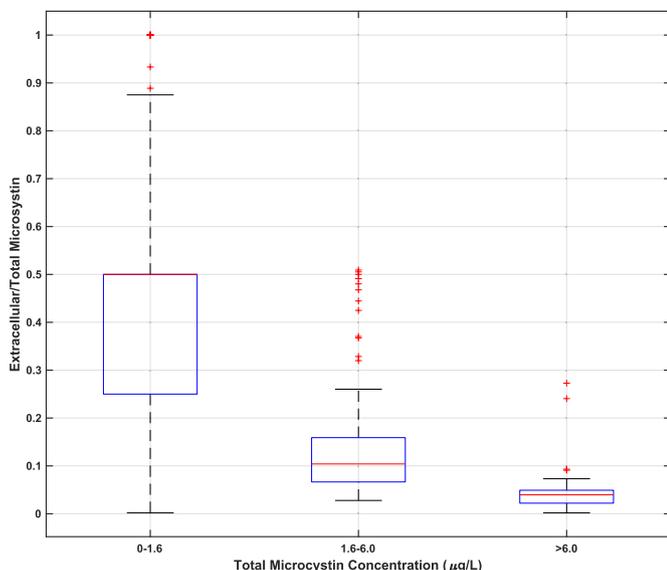


Fig. 4. Box plot of ratio of extracellular to total microcystin toxin.

2.2. Satellite remote sensing data

We used a series of images of cyanobacterial blooms in Lake Erie from July to October 2014–2017 obtained from the Moderate Resolution Imaging Spectrometer (MODIS) with ~1 km spatial resolution (Wynne and Stumpf, 2015). The cyanobacterial index (CI) was obtained using a spectral shape algorithm based around 680 nm (Wynne et al., 2008; Tomlinson et al., 2016). The CI varies linearly with biomass, with a value of 10^{-3} sr^{-1} corresponding to approximately $10^5 \text{ cells mL}^{-1}$ (Stumpf et al., 2012). MODIS provides two overpasses daily (Aqua and Terra), which result in 60–70 cloud free images per year during the summer when HABs occur (Vander Woude et al., 2019). In this study, we used images with greater than ~50% cloud-free coverage in western Lake Erie. For initialization of the HAB Tracker model, the cloud-covered areas of satellite images were filled in using model data from the previous day’s run (Rowe et al., 2016).

2.3. Predictive model (hindcast simulation)

A conceptual diagram depicts the data flow through our forecast system (Fig. 3). The system builds on the output of the HAB Tracker model (process *M1*), described by Rowe et al. (2016), to predict the spatially resolved CHAB chlorophyll-a concentration and movement (5-day forecast) using a Lagrangian particle tracking model initialized from satellite-derived CHAB concentration and forced by currents and turbulent diffusivity from a numerical hydrodynamic model (LEOFS). The HAB Tracker model produces concentration output with a vertical resolution of 1 m, with the number of layers depending on the water depth. The modeled CHAB chlorophyll-a concentration (data *D1*) is transformed to toxin concentration (data *D2*), using a linear regression model of chlorophyll-a concentration versus microcystin concentration

Probability of Exceeding Various Microcystin-LR Health Advisory Levels

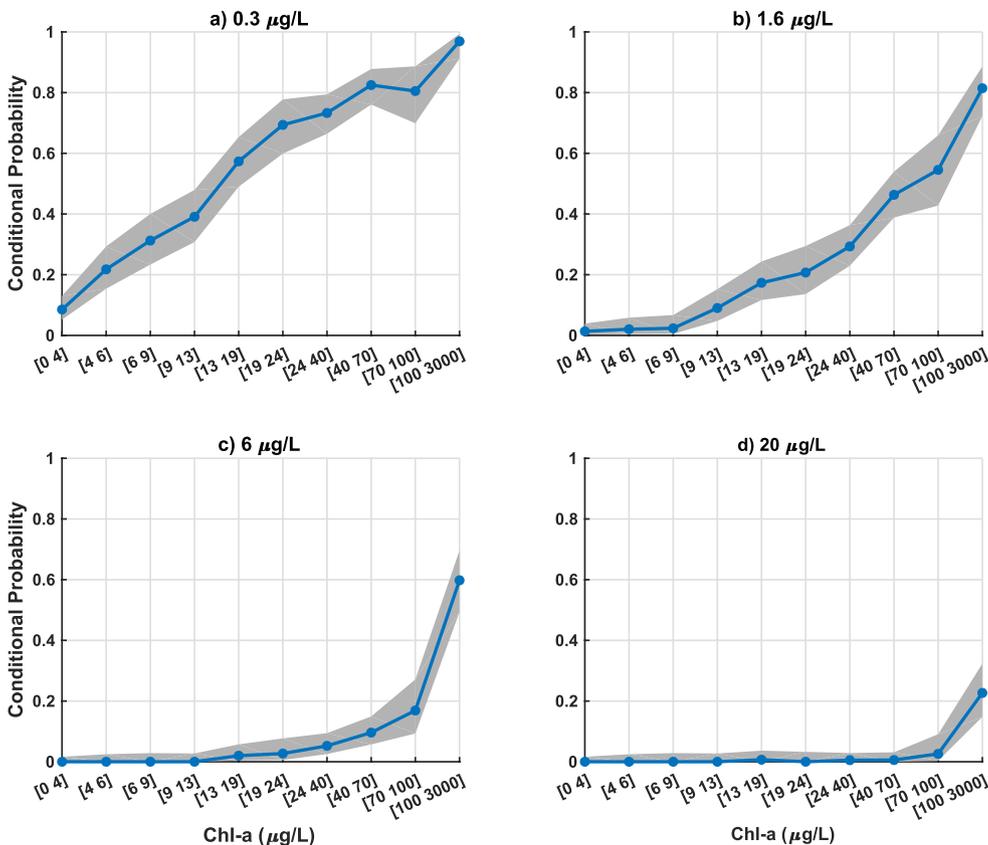


Fig. 5. Probability of microcystins exceeding microcystin advisory levels as a function of chlorophyll-a concentration. Microcystins and chlorophyll-a were measured in the same water sample, from the GLERL-CIGLR data (2008–2017). Ohio EPA threshold levels include a) children under 6 and sensitive population drinking water threshold - 0.3 µg/L, b) children 6 and older and adults drinking - 1.6 µg/L, c) Ohio PHA Microcystin-LR concentration - 6 µg/L, d) Ohio NCA Microcystin-LR concentration - 20 µg/L, as well as the 95% confidence intervals (gray shading).

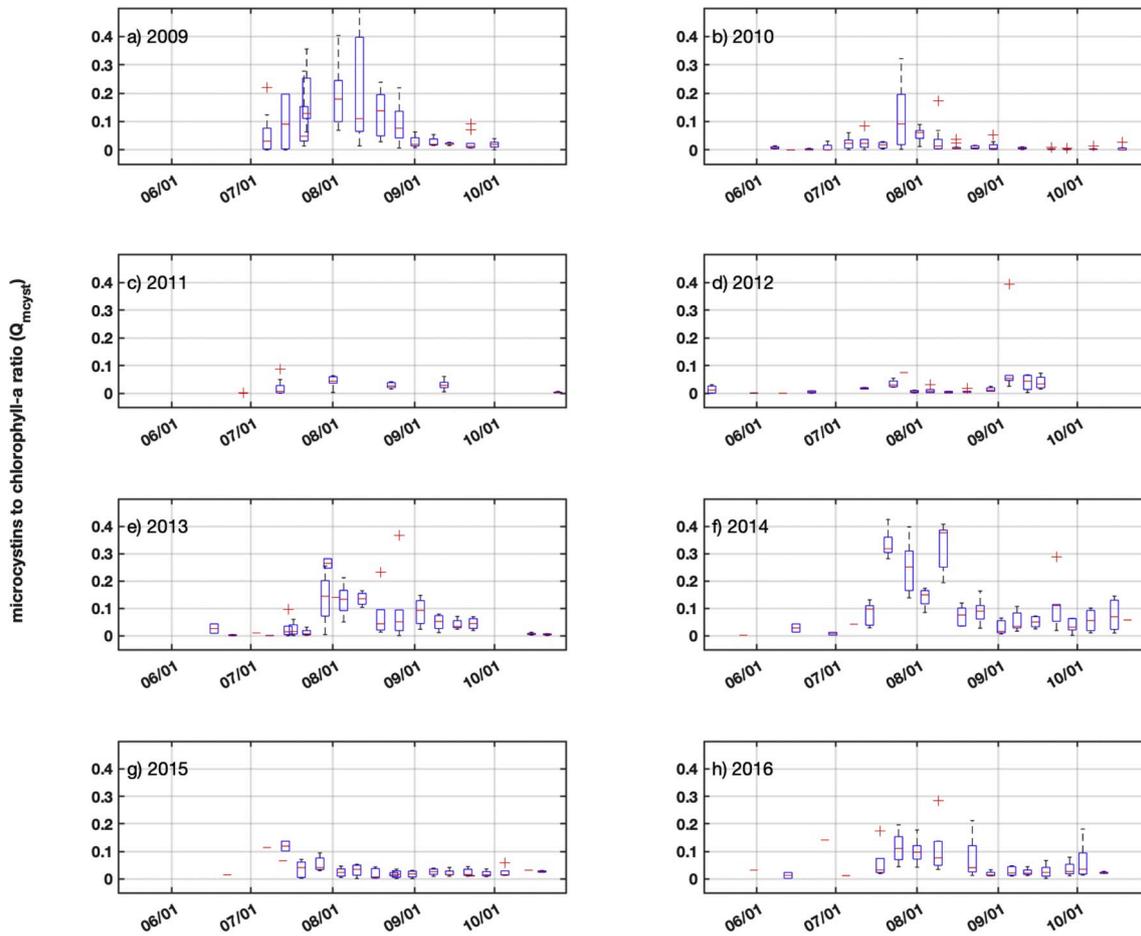


Fig. 6. Boxplots of the ratio of microcystins to chlorophyll-a (Q_{myst}) from the GLERL-CIGLR sampling program, where the boxplots represent the range of values observed across the eight sampling stations (Fig. 1) on a given date. On each box, the central mark indicates the median, and box represents the interquartile range. Whiskers extend to three times the interquartile range, with additional points plotted individually.

(process M2),

$$C_t^{pred} = \beta_0 + \beta_1 C_{chl} + \varepsilon \quad (1)$$

where C_t^{pred} is a predicted microcystin toxin concentration and C_{chl} is the chlorophyll-a concentration from the HAB Tracker model. The parameters β_0 and β_1 are estimated by linear regression and ε is uncertainty component reflecting the difference between observations and the fitted linear relationship. The observations used to estimate the parameters in Eq. (1) were from GLERL-CIGLR *in situ* grab samples in which toxin and chlorophyll-a were measured in the same samples, at a time before and closest to the model initiation time. The average time gap between *in situ* observations and model initialization from 2014 to 2017 is 1.2 days. The parameters in Eq. (1) were re-calculated for every forecast, using observations from the most recent weekly sampling date. On sample dates when fewer than three observations were available, we also included observations from the preceding sample date. Finally, after matching predicted toxin concentration (C_t^{pred}) to *in situ* observations (C_t^{obs}), we transformed the predicted toxin concentration into a probability of exceeding specified toxin concentration thresholds, C_{ii}^{thresh} , using a statistical model (process M3),

$$P(C_t^{obs} \geq C_{ii}^{thresh}) = f(C_t^{pred}) \quad (2)$$

where $P(C_t^{obs} \geq C_{ii}^{thresh})$ represents the probability of *in situ* observed microcystin concentration exceeding the threshold value, which was estimated by binning C_t^{obs} into intervals of C_t^{pred} over the hindcast period,

and calculating the fraction of $C_t^{obs} \geq C_{ii}^{thresh}$ for each interval of C_t^{pred} . The function $f(C_t^{pred})$ was then fit to the calculated values of $P(C_t^{obs} \geq C_{ii}^{thresh})$ by nonlinear least-squares regression.

Through the use of a probabilistic forecast, the level of uncertainty in the forecast system can be properly conveyed (Jolliffe and Stephenson, 2003), including uncertainties in satellite-derived estimates of CHAB abundance, *in situ* toxin measurements, a transport model, and the regression model (Eq. (1)). The transition from microcystin concentration to the probability of exceeding a threshold is achieved by a statistical model (Eq. (2)). Equation (2) gives the probability of observations exceeding the health advisory concentrations given that the model-predicted microcystins are within a range, based on observed and modeled microcystin concentration matchups over the 2014 to 2016 hindcast.

2.4. Model assessment

We used skill metrics for binary event forecasts to assess the forecast system (Hogan and Mason, 2012; Rowe et al., 2016; Anderson et al., 2016), including the probability of detection (POD), the probability of false detection (POFD), frequency bias (B), and the Pierce skill score (PSS). The skill scores were defined as,

$$POD = \frac{a}{a+c} \quad (3)$$

$$POFD = \frac{b}{b+d} \quad (4)$$

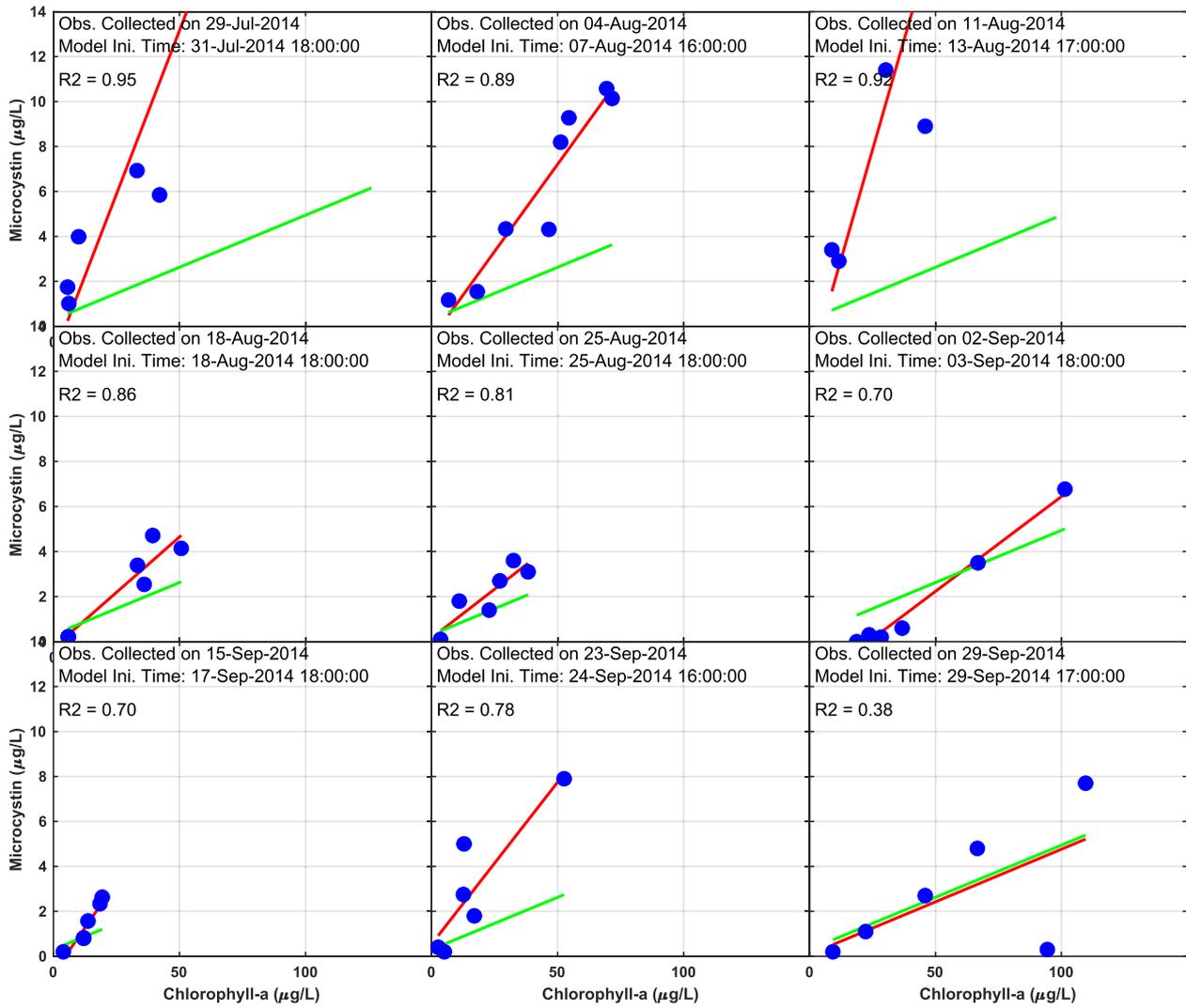


Fig. 7. Linear regression between chlorophyll-a and microcystins for all CIGLR/GLERL stations over the western basin of Lake Erie at a day closest to the initialization of the toxin forecast (red line represents the linear regression by Eq. (1), and the blue dots are *in situ* observations used for the linear regression). The green line represents the linear regression based on all CIGLR/GLERL microcystin and chlorophyll-a sampling from 2008 to 2017. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

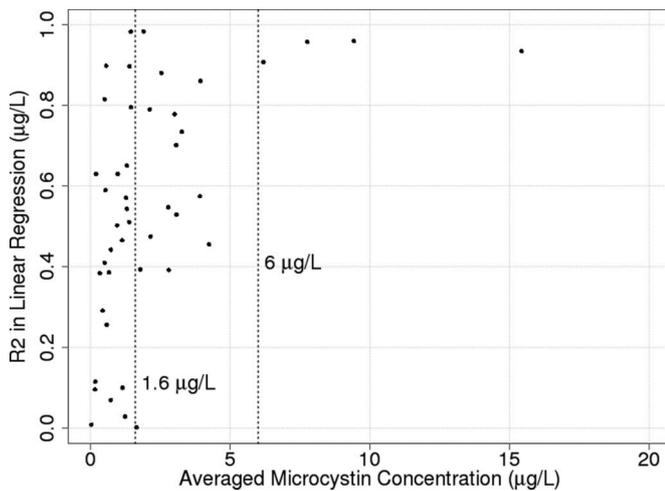


Fig. 8. Coefficient of determination (R^2) for all linear regressions between chlorophyll-a and microcystins for the hindcast simulations from 2014 to 2017. X-axis represents the averaged microcystin concentration for a single day CIGLR/GLERL sampling over the western basin of Lake Erie.

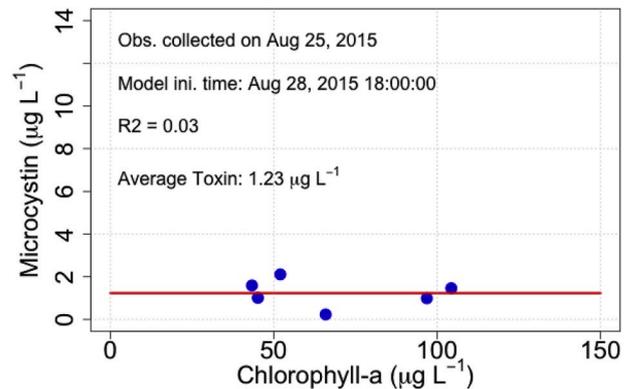


Fig. 9. Linear regression between chlorophyll-a and microcystins based on CIGLR/GLERL sampling on August 25, 2015. On this date, microcystins-producing cyanobacteria were not dominant, chlorophyll-a was not a predictor of microcystins ($R^2 = 0.03$), and therefore low concentration of microcystins was predicted regardless of chlorophyll-a concentration.

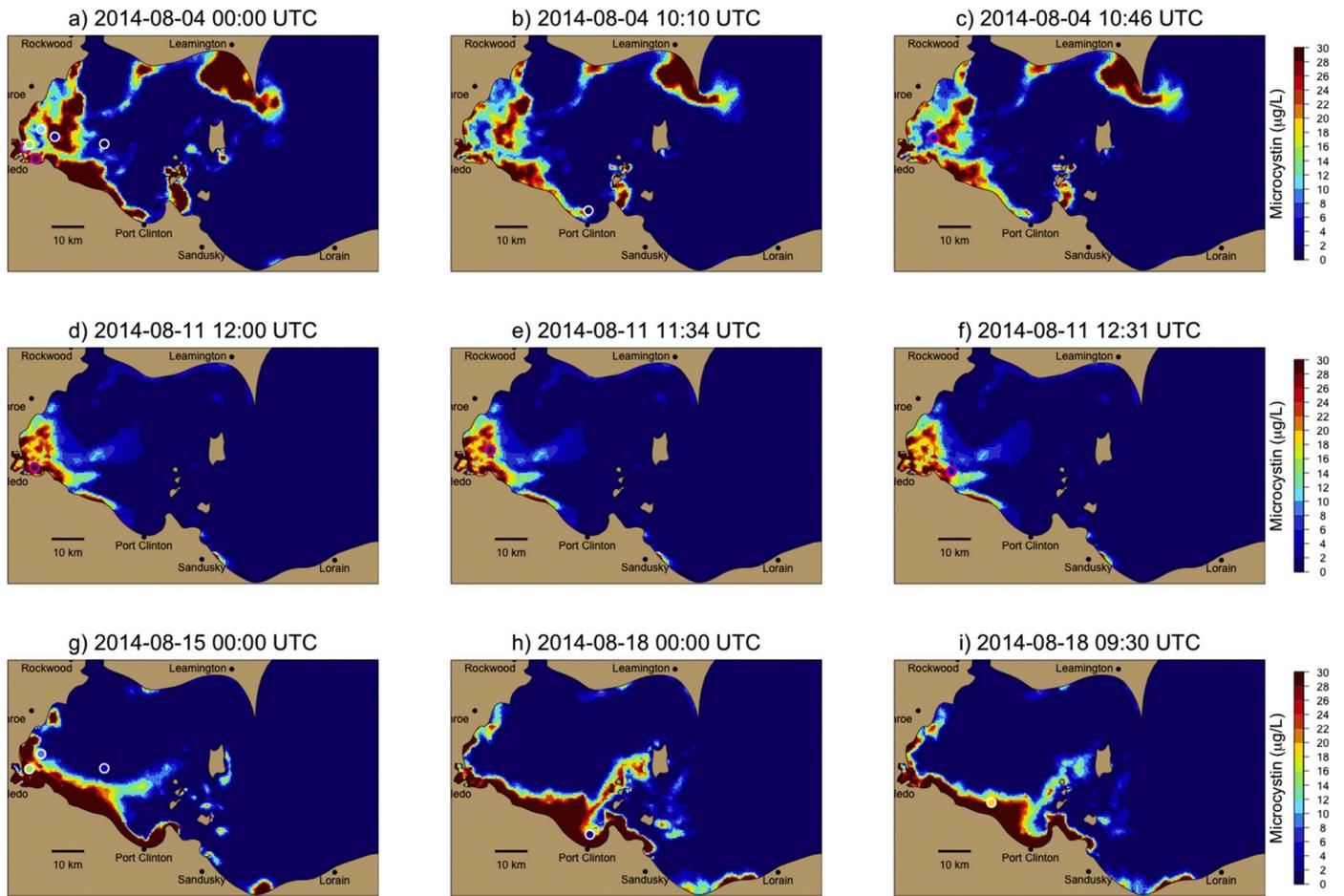


Fig. 10. Modeled surface microcystin concentration from hindcast runs initialized on a-c) August 1, 2014, d-f) August 10, 2014, and g-i) August 14, 2014 based on the five-day chlorophyll-a forecast and linear regression of toxin vs chlorophyll-a. The dots represent observations, with the face color representing microcystin concentration and edge color representing data type. Magenta-edged dots represent surface observations, and white-edged dots represent integrated observations.

$$B = \frac{a + b}{a + c} \tag{5}$$

$$PSS = POD - POFD = \frac{ad - bc}{(b + d)(a + c)} \tag{6}$$

where a represents the number of correctly predicted events (hits); b , incorrectly predicted events (false alarms); c , false negatives (misses); and d , correct nonevents, with an event defined as microcystin concentration exceeding a specific health advisory level. POD is a positively oriented score (higher values represent better performance) with values

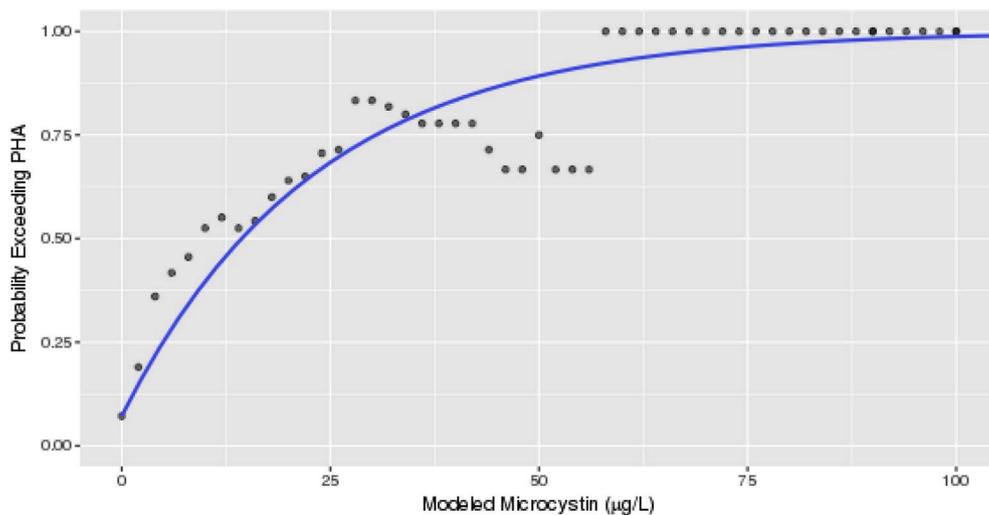


Fig. 11. Observed probability of microcystin concentration exceeding Ohio PHA level (6 µg/L) for different model-predicted microcystin concentrations over the 2014–2016 hindcast (symbols). The line is a nonlinear fit of Eq. (7).

Table 1

Skill assessment for prediction of exceeding microcystin concentration thresholds in the western basin of Lake Erie using 548 pairs of surface predictions and observations from 2014 to 2016. PHA represents Ohio public health advisory level (PHA, 6 $\mu\text{g/L}$), and NCA represents Ohio No Contact Advisory level (NCA, 20 $\mu\text{g/L}$). Skill metrics for binary predictions include the bias (B) and Pierce skill score (PSS). Exceedance number represents number of observations greater than the specified threshold.

Health Advisory Level (conc.)	B	PSS	Exceedance Number
Ohio EPA Child (0.3)	0.84	0.19	483
Ohio EPA Adult (1.6)	0.87	0.37	260
PHA (6)	1.14	0.41	57
NCA (20)	1.0	0.35	19

in the range [0,1]; *POFD* is a negatively oriented score (higher values represent worse performance) in the range of [0,1]. *PSS* is a positively oriented score in the range of [-1,1], with positive values indicating larger *POD* than *POFD*. *B* is the ratio of predicted events to observed events, in the range of [0, ∞], and equals 1 for an unbiased forecast.

Numerous studies describe methods to assess probabilistic forecasts in meteorology (e.g., Bettge and Baumhefner, 1981; Doswell and Flueck, 1989; Wilks, 1995). We applied two methods, the reliability diagram and the receiver operating characteristic (ROC) curve to evaluate the probabilistic forecast. Reliability represents the degree of agreement between forecast frequency, or probability, and observed frequency. A reliability diagram (Hartmann et al., 2002) with the observed frequency of an event plotted against the forecast probability of an event can effectively tell the user how closely the forecast probability corresponds to the actual chance of observing the event. In the case of perfect reliability, the forecast probability and the observed frequency should be equal, and the plotted points should be on the diagonal.

The ROC curve is based on signal detection theory and attempts to measure the relative “signal” and “noise” ratios contained in forecast information in the form of *POD* to *POFD* ratios (Egan, 1975; Mason, 1982). In a ROC curve, the forecast probability is converted into a continuous set of yes/no binary events based on the given probability thresholds, and *POD* and *POFD* are compared within the probability thresholds (Swets, 1973; Egan, 1975; Mason, 1982; Mason and Graham, 1999). When the ROC curve lies above and to the left of the (0,0) to (1,1) diagonal, the model has good prediction skill; when the curve lies below and to the right of the diagonal, the model lacks skill. The ROC score *AUC* is defined as the area under the ROC curve. Correspondingly, when *AUC* is larger than 0.5, the model performs better than random chance, and worse than by chance when *AUC* is smaller than 0.5.

3. Results

3.1. Association between chlorophyll-a and microcystin concentrations

Chlorophyll-a may be a useful indicator of microcystin concentration, if microcystins are mainly contained within cyanobacterial cells. In our data, extracellular microcystins were a relatively small fraction of total microcystins (Fig. 4) when total microcystin concentrations (particulate plus dissolved) were greater than 1.6 $\mu\text{g/L}$ (adult drinking water advisory level). Given that microcystins were primarily intracellular (particulate) when advisory levels were exceeded, it may be anticipated that microcystins are unlikely to exceed advisory levels in the absence of high chlorophyll concentrations. To investigate this concept, we calculated the proportion of microcystin concentrations exceeding threshold values within intervals of chlorophyll concentration, and calculated a 95% confidence interval on the proportion. The probability of exceeding microcystin advisory levels increased monotonically with increasing chlorophyll-a concentration (Fig. 5), which supports the concept of using cyanobacterial chlorophyll-a concentrations as a means to predict the probability of exceeding microcystin advisory levels. These exceedance probabilities only apply to western Lake Erie, and may change

over time with changes in the phytoplankton community. Such exceedance probability analysis can be used to predict toxin concentration in the absence of microcystin measurements (Hollister and Kreackie, 2016).

We investigated temporal variation in the microcystin:chlorophyll-a ratio ($Q_{mcyst} = mc : chl$) over the bloom season (July to October) and across years. Here, Q_{mcyst} is taken to be a surrogate for the toxin to biomass ratio of the phytoplankton community. If we assume that Q_{mcyst} varies depending on environmental and ecological conditions, and that the transport time of water masses among the limited spatial domain of our sampling program is shorter than the time scale over which the drivers of Q_{mcyst} vary, we may expect the microcystin:chlorophyll-a ratio ($Q_{mcyst} = mc : chl$) to vary temporally more so than spatially on a given day. Values of Q_{mcyst} show strong seasonal variability (Fig. 6), which usually peaks in July or August. Inter-annual variability is also strong, with a much greater maximum value of Q_{mcyst} in 2009 and 2014 than in other years. In contrast, the range of Q_{mcyst} on a given date, representing spatial and sampling uncertainty combined, is smaller than the range of Q_{mcyst} values in the whole data set since 2008 or in one year. Values of Q_{mcyst} also exhibit temporal autocorrelation, with low values typically occurring early in the season, high values mid-season, and low values again late in the bloom season.

Based on the above analysis, we assumed that toxin concentration was proportional to cyanobacterial biomass, as represented by chlorophyll-a concentration. We further assumed that Q_{mcyst} was spatially uniform, but varied by sampling week; thus, we used linear regression to determine parameters in Eq. (1) for each sampling date. Uncertainties introduced by these assumptions will ultimately be conveyed by a probabilistic forecast (Fig. 3, M3). Fig. 7 shows all the linear regressions for the hindcast simulations in 2014 with data collection time and model initialization time on each plot, most of which are significantly different from the climatological regression based on all CIGLR/GLERL samplings from 2008 to 2017 (green lines in Fig. 7). The Chow test (Chow, 1960) demonstrated that the coefficients of the regression models change with time. The R^2 values for the hindcast simulations from 2014 to 2017 generally increased with the increase of averaged microcystin concentration on a single sampling day (Fig. 8). When the averaged microcystin concentration is larger than 1.6 $\mu\text{g/L}$, most of R^2 values are larger than 0.39; when the averaged microcystin concentration is larger than 6 $\mu\text{g/L}$ (PHA threshold), R^2 is larger than 0.9.

When average microcystin concentration is low, a situation when the microcystins-producing cyanobacteria are not dominant, the R^2 for linear regression between chlorophyll-a and microcystins can be small, as shown in Fig. 8. Under this condition, the model predicts low microcystin concentration, resulting from a low microcystin:chlorophyll ratio. For example, the point on the lower left corner of Fig. 8, with R^2 of 0.03, and average microcystin of 1.23 $\mu\text{g L}^{-1}$, from August 25, 2015 (Fig. 9), represents the lowest R^2 of our hindcast period. Using the linear regression model based on this day’s sampling, the model appropriately predicts low microcystin concentration (Fig. 9), and a low probability (12%) of exceeding the public health advisory level (Fig. 11).

3.2. Spatial and temporal prediction of microcystin concentration

We predicted microcystin concentration during the five-day forecast period by combining Eq. (1) with chlorophyll-a predicted by the HAB Tracker model, (Fig. 3, D2). We then compared the modeled microcystin concentrations with the toxin database, using matchup criteria described in Section 2.1. For example, snapshots of modeled surface toxin concentrations in 2014 are shown in Fig. 10, with the corresponding surface and integrated observations plotted using symbols on the same color scale.

We assessed model skill in predicting exceedance of advisory levels by comparison with *in situ* observations over 2014 to 2016. With

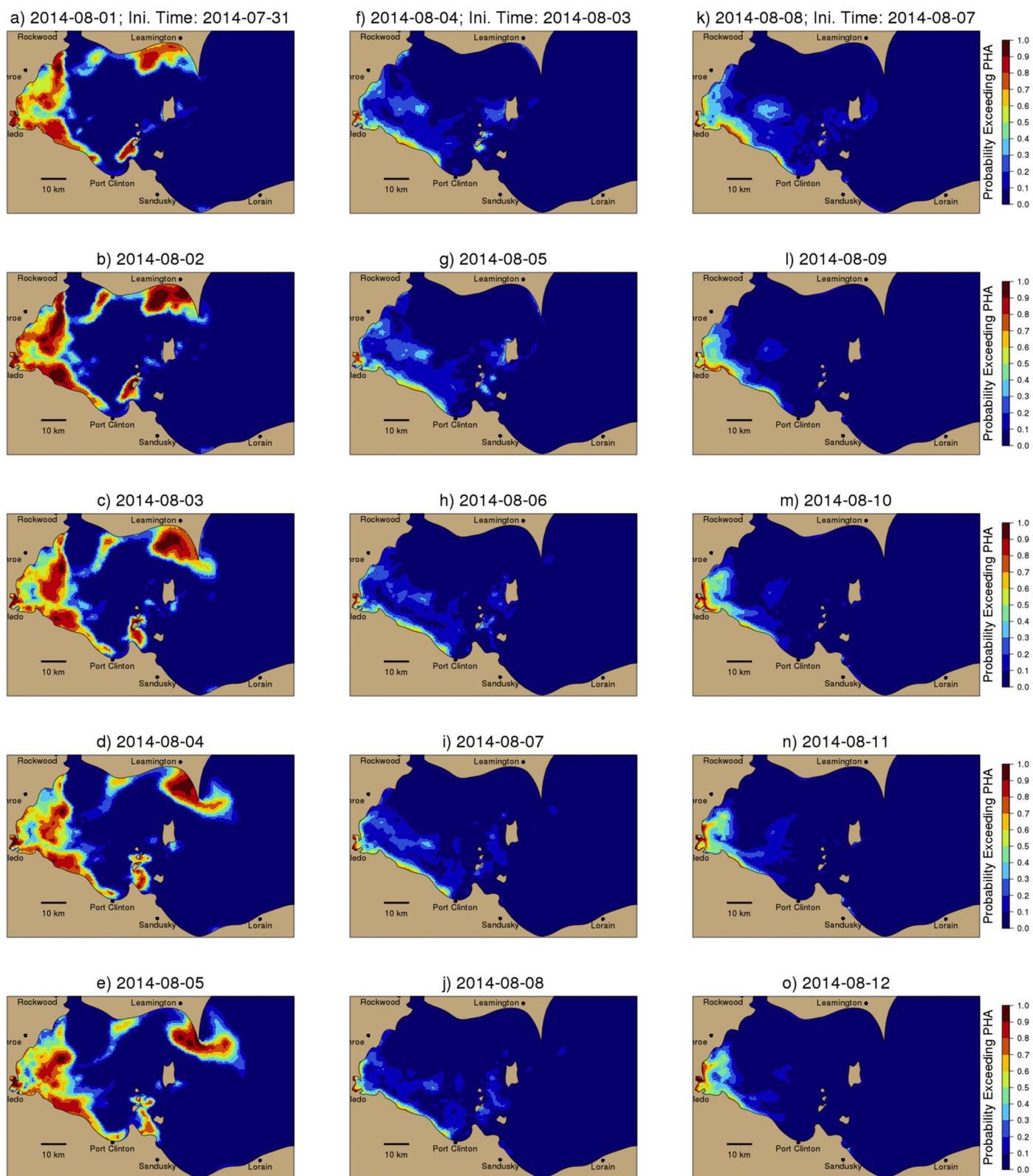


Fig. 12. The short-term (5-day) probabilistic forecasts of microcystin concentration exceeding Ohio PHA level ($6 \mu\text{g/L}$) in 2014. The forecasts initialized from July 31st, August 3rd and August 7th are shown in each column.

comparable bias for different advisory levels, the model performed best at mid-range concentrations associated with the PHA and Ohio EPA adult drinking thresholds, having PSS values of 0.41 and 0.37, respectively (Table 1). Because PSS is largest for the PHA threshold, we focused on prediction of exceeding the PHA threshold as a goal for the forecast

system.

3.3. Probability of exceeding toxin threshold

Using the model and *in situ* microcystins matchups between 2014 and

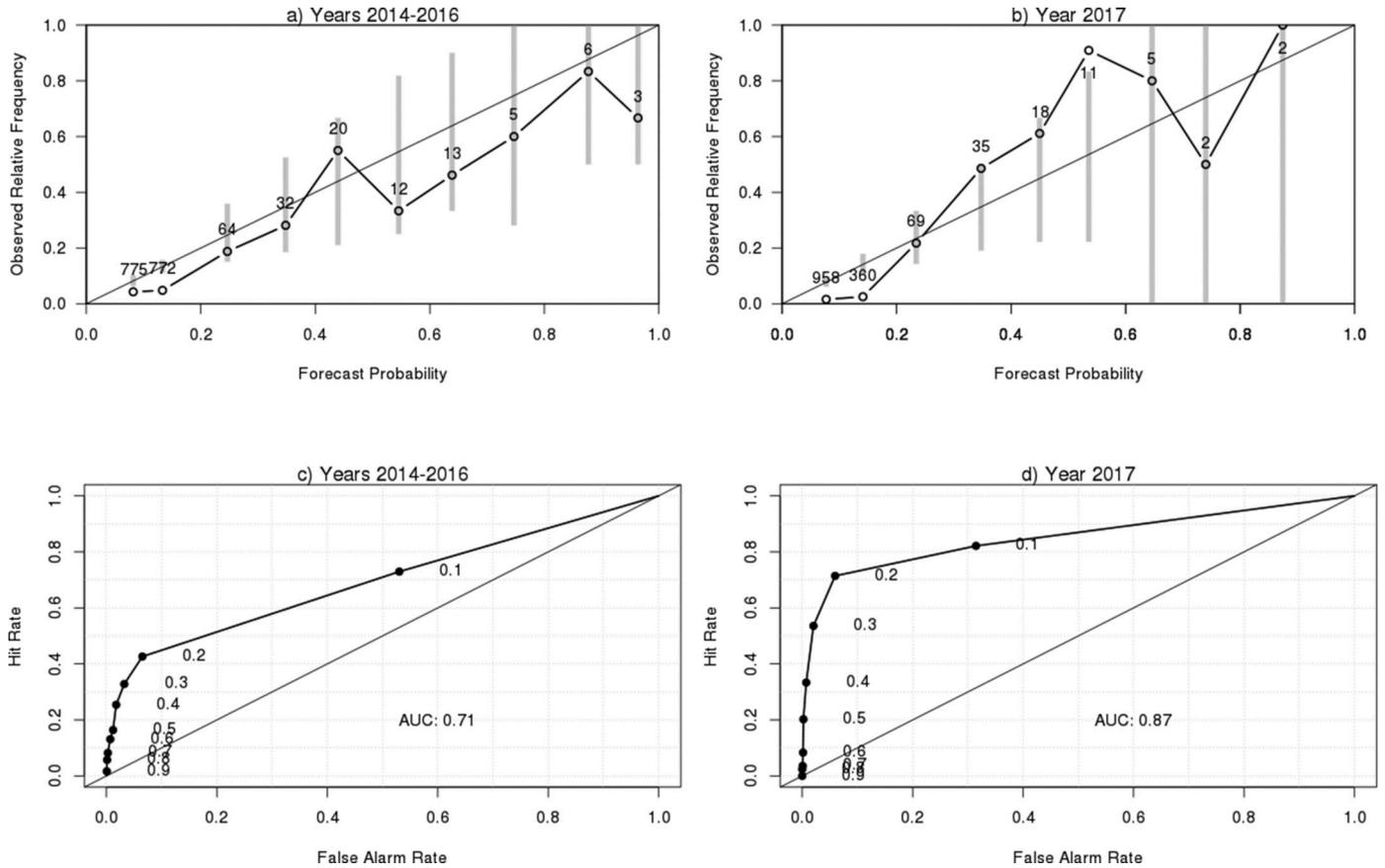


Fig. 13. Reliability diagrams for the probabilistic forecasts of microcystin concentration exceeding Ohio PHA advisory concentration in a) years 2014–2016, and b) year 2017, and ROC curves for forecasts in c) years 2014–2016, and d) year 2017. The forecast probabilities are divided into bins with 10% increment. The numbers in the reliability diagrams represent the numbers of samples used in each probability bin, and the vertical gray bars represent the 5th and 95th percentiles confidence intervals determined by 1000 bootstrap samples.

2016, we calculated the probability of exceeding a health advisory level given that the model-predicted microcystins were within a specified range (Fig. 11). For example, when the modeled microcystin concentration is larger than 10 $\mu\text{g/L}$, the probability of in-situ observations exceeding the PHA level is $\sim 50\%$. The relationship between modeled microcystin concentration (C_t^{pred}) and exceedance probability ($f(C_t^{\text{pred}})$) as in equation (2) was fit through nonlinear least-squares estimates of a nonlinear model (Fig. 11)

$$f(C_t^{\text{pred}}) = 1 + P_0 \left(-\exp\left(-\frac{C_t^{\text{pred}}}{\gamma}\right) \right) + \varepsilon \quad (7)$$

where γ is an unknown parameter, and $1 - P_0$ is the vertical intercept, approximated as the probability calculated by model-*in situ* matchups when C_t^{pred} is 0. ε is uncertainty component reflecting the difference between calculated probability (gray dots in Fig. 11) and the fitted nonlinear relationship. For the modeled microcystin concentration larger than 65 $\mu\text{g/L}$, the probability of event occurrence was $\sim 90\%$. In process M3 (Fig. 3), Eq. (7) was used to convert modeled microcystin concentrations to a probability of exceeding the PHA level (Fig. 12).

3.4. Skill assessment of the probabilistic forecast

The reliability diagram compares the observed relative frequency to the forecast probability. Microcystin concentrations in the toxin database were used to assess the toxin forecast from 2014 to 2017, with 2017 representing an independent time period not used in calibration of Eq. (7). Fig. 13a shows the reliability diagram for the probabilistic forecast of microcystin concentration exceeding PHA advisory concentration (6

$\mu\text{g/L}$) based on 85 forecast runs for years 2014–2016, and Fig. 13b based on 56 forecast runs for year 2017. The 1:1 line represents the situation with perfect reliability, where the predicted probability exactly matches the observed. Fig. 13a shows that, from 2014 to 2016, the confidence interval of the model reliability curve overlaps the 1:1 line, although the model tended to over-predict the frequency of exceedance on average in 2014–2016. The area under the ROC curve (AUC) in years 2014–2016 is 0.71, and 0.87 for year 2017, which means the forecast is better than by chance (AUC of 0.5).

4. Discussion

Drawing on a large data set of microcystins and chlorophyll-a observations in western Lake Erie, we found that microcystin PHA concentrations were unlikely to be exceeded when chlorophyll-a concentrations were low (Fig. 5), and that chlorophyll-a was a useful predictor of microcystin concentrations when the mean microcystin concentration was near PHA levels (Fig. 8). At PHA levels, microcystins are primarily intracellular (Fig. 4), and thus it is reasonable to assume that microcystins scale with algal biomass. However, these conditions do not always prevail, and at times non-toxic phytoplankton are dominant. In which case, chlorophyll-a is a weak predictor of microcystins, but regression of microcystins on chlorophyll-a appropriately predicts toxin concentrations near the mean at all chlorophyll levels (Fig. 7a). This approach could fail if chlorophyll-a concentrations at the monitoring stations do not cover the range of values over the full prediction spatial domain, requiring extrapolation of a poorly-predictive relationship. Our monitoring stations are located in the area most affected by HABs (Wynne and Stumpf, 2015), and are likely to capture the highest

chlorophyll-a concentrations, but extrapolation should be avoided.

Given that predictive models relating microcystin toxin production to environmental drivers are lacking (Bullerjahn et al., 2016), we resorted to an approach driven by weekly updated observations. The microcystin:chlorophyll-a ratio (Q_{mcyst}) varied in a seasonal pattern, and changed over time with changes in the phytoplankton community, reaching maximum values in July or August, with the maximum value varying by year. The variance in Q_{mcyst} ratio on a given date represents a combination of sampling uncertainty and any spatial pattern that may exist. Our sampling program lacked the level of replication and spatial extent that would be required to distinguish a spatial pattern from sampling uncertainty, thus we resorted to the assumption of spatially uniform Q_{mcyst} determined by linear regression. We anticipate that over a sufficiently broad spatial extent, the assumption of spatial uniformity of Q_{mcyst} would fail. For example, toxic HABs rarely occur in the central basin of Lake Erie (Ohio EPA data). However, the database of microcystin observations that we used for model assessment fully covered the southern half of the western basin (Fig. 1b) where HABs occur, and covered a broader spatial extent than the monitoring stations that we used to determine Q_{mcyst} (Fig. 1a); thus, our skill assessment is largely representative of the region over which the forecast would be applied. However, it may be appropriate to limit model predictions to the region over which the model has been assessed. If improved predictive models of Q_{mcyst} as a function of environmental drivers become available, the spatially-uniform linear regression model used here could be replaced within the forecast system framework we describe.

The spatial distribution of cyanobacterial biomass, and associated toxin concentration, is quite variable at small spatial scales (Fig. 2), which introduces challenges in obtaining in-situ observations that are representative over spatial scales of satellite or model data (~km). Our predictions of microcystin concentrations showed some skill, but large mis-matches also occurred (Fig. 10). Therefore, it is most appropriate to represent our toxin predictions in terms of probability.

Our forecast system was most skillful at predicting exceedance of toxin thresholds at intermediate toxin concentrations (Table 1). At the lowest threshold level (0.3 µg/L), dissolved microcystins may be equal or greater than particulate microcystins, conditions under which chlorophyll-a (algal biomass) was not predictive of total microcystins. At the highest threshold level (20 µg/L), relatively few observations existed for calibration of the conditional probability function (Eq. (7)). The conditional probability function can be refined as additional observations accumulate over years.

Both measures of performance of the probabilistic forecast indicated a potentially useful level of skill. We characterized the performance of the probabilistic forecast of exceeding the microcystin PHA level over the 2014 to 2016 hindcast period, and for 2017, a year that was not included in calibration of the model, using the reliability diagram and the ROC curve. In the reliability diagram, the forecast probability was within uncertainty bounds of observed relative frequency, given limited observations. The ROC curve provides an analysis of the forecast performance from the perspective of a decision maker (Fig. 13c and d). A user whose decision is sensitive to the POFD of the forecast system may use this plot to identify a forecast probability threshold that maximizes POD while keeping an acceptable POFD. For example, such a user would be able to achieve a 70% POD while maintaining a 6% POFD using a probability threshold of 20% for year 2017. Using the conditional probability function (Fig. 7), obtained from comparison of model predictions to observations over the hindcast period, ensures that the various sources of uncertainty are reflected in the predicted probabilities. Even though multiple sources of uncertainty in our complex forecast system contribute to limited skill, the final product reasonably discriminated between high and low probabilities of exceeding the PHA.

Sources of uncertainty in the probabilistic forecast point to opportunities to improve the forecast system. Uncertainty arises from satellite-derived initial conditions for bloom location and biomass, the

hydrodynamic forecast, and the Lagrangian particle transport model. Improved accuracy in each of these components will improve skill of the forecast system. As additional toxin observations accumulate, the conditional probability function can be refined. Because the linear regression model ignores spatial variability of the association between chlorophyll-a and microcystins, and the temporal variability is updated weekly at the most frequent, uncaptured variability in the microcystin:chlorophyll-a relationship contribute to model error. Increased frequency or spatial extent of microcystin and chlorophyll-a paired observations would likely improve forecast skill. If predictive models of the microcystin:chlorophyll-a as a function of environmental variables become available, they may be substituted for the simple linear regression in the forecast system. This effort provides a path toward an operational short-term forecast system for the probability of exceeding HAB toxin advisory levels in Lake Erie, provides a benchmark against which to assess future efforts, and provides an approach that may be applied to other systems affected by toxic cyanobacterial blooms.

Declaration of competing interests

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

Acknowledgements

Funding was awarded to Cooperative Institute for Great Lakes Research (CIGLR) through the NOAA Cooperative Agreement with the University of Michigan (NA17OAR4320152). This is CIGLR publication # 1157 and GLERL publication # 1943. The project was supported by the U.S. EPA Great Lakes Restoration Initiative. Lake Erie algal toxin and water quality data used in this analysis were compiled by LimnoTech from public sources, including NOAA (<https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0187718>), Ohio EPA (<https://epa.ohio.gov/ddagw/HAB>), and The Ohio State University (<https://ohioseagrant.osu.edu/research/live/water>). The Lake Erie HAB Tracker model (Rowe et al., 2016) was developed using the Finite Volume Community Ocean Model (FVCOM) and associated Lagrangian particle tracking model available through University of Massachusetts Dartmouth (<http://fvcom.smast.umassd.edu/>).

References

- Anderson, C.R., Kudela, R.M., Kahru, M., et al., 2016. Initial skill assessment of the California harmful algae risk mapping (C-HARM) system. *Harmful Algae* 59, 1–8.
- Bette, T.W., Baumhoffer, D.P., 1981. On the verification of seasonal climate forecasts. *Bull. Am. Met. Soc.* 62 (12), 1654.
- Brittain, S.M., Wang, J., Babcock-Jackson, L., Carmichael, W.W., Rinehart, K.L., Culver, D.A., 2000. Isolation and characterization of microcystins, cyclic heptapeptide hepatotoxins from a Lake Erie strain of *Microcystis aeruginosa*. *J. Great Lake. Res.* 26, 241–249. [https://doi.org/10.1016/S0380-1330\(00\)70690-3](https://doi.org/10.1016/S0380-1330(00)70690-3).
- Bullerjahn, G.S., McKay, R.M., Davis, T.W., Baker, D.B., Boyer, G.L., D'Anglada, L.V., et al., 2016. Global solutions to regional problems: collecting global expertise to address the problem of harmful cyanobacterial blooms. A Lake Erie case study. *Harmful Algae* 54, 223–238.
- Chen, C., Liu, H., Beardsley, R.C., 2003. An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries. *J. Atmos. Ocean. Technol.* 20 (1), 159–186.
- Chow, Gregory C., 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28 (3), 591–605. <https://doi.org/10.2307/1910133>. JSTOR1910133.
- Churchill, J.H., Runge, J., Chen, C., 2011. Processes controlling retention of spring-spawned Atlantic cod (*Gadus morhua*) in the western Gulf of Maine and their relationship to an index of recruitment success. *Fish. Oceanogr.* 20 (1), 32–46.
- Davis, T.W., Berry, D.L., Boyer, G.L., Gobler, C.J., 2009. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. *Harmful Algae* 8, 715–725.
- Doswell III, C.A., Flueck, J.A., 1989. Forecasting and verifying in a field research project: DOPLIGHT 87. *Weather Forecast.* 4, 97–109.
- Egan, J.P., 1975. Signal Detection Theory and ROC Analysis. Series in Cognition and Perception. Academic Press, New York.

- Hartmann, H.C., Pagano, T.C., Sorooshian, S., Bales, R., 2002. Confidence builder: evaluating seasonal climate forecasts from user perspectives. *Bull. Amer. Met. Soc.* 84, 683–689.
- Hogan, R.J., Mason, I.B., 2012. Deterministic forecasts of binary events. In: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, second ed. John Wiley & Sons, Ltd., West Sussex, U. K, pp. 31–59.
- Hollister, J.W., Kreakie, B.J., 2016. Associations between chlorophyll a and various microcystin health advisory concentrations. *F1000Research* 5, 151. <https://doi.org/10.12688/f1000research.7955.2>.
- Huret, M., Runge, J., Chen, C., Cowles, G., Xu, Q., Pringle, J., 2007. Dispersal modeling of fish early life stages: sensitivity with application to Atlantic cod in the western Gulf of Maine. *Mar. Ecol. Prog. Ser.* 347, 261–274.
- Jolliffe, I.T., Stephenson, D.B., 2003. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. J. Wiley, Chichester, West Sussex, England.
- Kane, D.D., Conroy, J.D., Richards, R.P., Baker, D.B., Culver, D.A., 2014. Re-eutrophication of Lake Erie: correlations between tributary nutrient loads and phytoplankton biomass. *J. Great Lake. Res.* 40, 496–501.
- Kelley, J.G.W., Anderson, E.J., Xu, J., 2018. Upgrade of NOS Lake Erie Operational Forecast System (LEOFS) to FVCOM: Model Development and Hindcast Skill Assessment. NOAA Technical Memorandum. NOS CS 40.
- Levy, S., 2017. Microcystis rising: why phosphorus reduction isn't enough to stop cyanoHABS. *Environ. Health Perspect.* 125 (2), A34–A39.
- Mason, I., 1982. A model for assessment of weather forecasts. *Aust. Meteorol. Mag.* 30, 291–303.
- Mason, S.J., Graham, N.E., 1999. Conditional probabilities, relative operating characteristics, and relative operating levels. *Weather Forecast.* 14, 713–725.
- Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., et al., 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, 6448–6452. <https://doi.org/10.1073/pnas.1216006110>.
- O'Neil, J.M., Davis, T.W., Burford, M.A., Gobler, C.J., 2012. The rise of harmful cyanobacteria blooms: the potential roles of eutrophication and climate change. *Harmful Algae* 14, 313–334.
- Rinta-kanto, J.M., Ouellette, A.J.A., Boyer, G.L., Twiss, M.R., Bridgeman, T.B., Wilhelm, S.W., 2005. Quantification of toxic *Microcystis* spp. during the 2003 and 2004 blooms in Western Lake Erie using quantitative real-time PCR. *Environ. Sci. Technol.* 39 (11), 4198–4205.
- Rinta-kanto, J.M., Konopko, E.A., DeBruyn, J.M., Bourbonniere, R.A., Boyer, G.L., Wilhelm, S.W., 2009. Lake Erie *Microcystis*: relationship between microcystin production, dynamics of genotypes and environmental parameters in a large lake. *Harmful Algae* 8, 665–673. <https://doi.org/10.1016/j.hal.2008.12.004>.
- Rowe, M.D., Anderson, E.J., Wynne, T.T., Stumpf, R.P., Fanslow, D.L., Kijanka, K., Vanderploeg, H.A., Strickler, J.R., Davis, T.W., 2016. Vertical distribution of buoyant *Microcystis* blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. *J. Geophys. Res. Oceans* 121, 5296–5314. <https://doi.org/10.1002/2016JC017720>.
- Steffen, M.M., Davis, T.W., McKay, R.M.L., Bullerjahn, G.S., Krausfeldt, L.E., Stough, J. M., et al., 2017. Ecophysiological examination of the Lake Erie *Microcystis* bloom in 2014: linkages between biology and the water supply shutdown of Toledo, OH. *Environ. Sci. Technol.* 51 (12), 6745–6755.
- 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. <https://doi.org/10.1371/journal.pone.0042444>.
- Stumpf, R.P., Johnson, L.T., Wynne, T.T., Baker, D.B., 2016a. Forecasting annual cyanobacterial bloom biomass to inform management decisions in Lake Erie. *J. Great Lake. Res.* <https://doi.org/10.1016/j.jglr.2016.08.006>.
- Stumpf, R.P., Davis, T.W., Wynne, T.T., Graham, J.L., Loftin, K.A., Johengen, T.H., Gossiaux, D., Palladino, D., Burtner, A., 2016b. Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae* 54, 160–173.
- Swets, J.A., 1973. The relative operating characteristic in psychology. *Science* 182, 990–1000.
- Tomlinson, M.C., Stumpf, R.P., Wynne, T.T., Dupuy, D., Burks, R., Hendrickson, J., Fulton III, R.S., 2016. Relating chlorophyll from cyanobacteria-dominated inland waters to a MERIS bloom index. *Remote Sens. Lett.* 7 (2), 141–149.
- U.S. Environmental Protection Agency, 2019. Recommended human health recreational ambient water quality criteria or swimming advisories for microcystins and cylindrospermopsin. *FR* 84, 26413 (26413-26414).
- Vander Woude, A., Ruberg, S., Johengen, T., Miller, R., Stuart, D., 2019. Spatial and temporal scales of variability of cyanobacteria harmful algal blooms from NOAA GLERL airborne hyperspectral imagery. *J. Great Lake. Res.* <https://doi.org/10.1016/j.jglr.2019.02.006>.
- Wilks, D.S., 1995. *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, p. 467.
- Wynne, T.T., Stumpf, R.P., Tomlinson, M., Warner, R., Tester, P., Dyble, J., Fahnenstiel, G., 2008. Relating spectral shape to cyanobacterial blooms in the Laurentian Great Lakes. *Int. J. Rem. Sens.* 29 (12), 3665–3672.
- Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Fahnenstiel, G.L., Dyble, J., Schwab, D.J., Joshi, S.J., 2013. Evolution of a cyanobacterial bloom forecast system in western Lake Erie: development and initial evaluation. *J. Great Lake. Res.* 39, 90–99.
- Wynne, T.T., Stumpf, R.P., 2015. Spatial and temporal patterns in the seasonal distribution of toxic cyanobacteria in Western Lake Erie from 2002–2014. *Toxins* 7 (5), 1649–1663. <https://doi.org/10.3390/toxins7051649>.