Nutrient and environmental factors regulating western Lake Erie cyanobacterial blooms

Hounshell, A.G.[1], Johnson, L.T.[2], Stumpf, R.P.[3]

[1] National Centers for Coastal Ocean Science, National Oceanic and Atmospheric Administration, 101 Pivers Island Rd., Beaufort, NC, 28516; alexandria.hounshell@noaa.gov

[2] National Center for Water Quality Research, Heidelberg University, 310 E. Market St., Tiffin, OH, 44883; ljohnson@heidelberg.edu

[3] National Centers for Coastal Ocean Science, National Oceanic and Atmospheric Administration, 1305 East-West Highway, Silver Spring, MD, 20910; richard.stumpf@noaa.gov

Abstract:

Over the past two decades, western Lake Erie has experienced recurring summer cyanobacterial blooms that pose severe negative impacts on human, animal, and ecological health. Previous research has identified a strong correlation between annual cyanobacterial bloom intensity and preceding spring (March-July) phosphorus loading from the Maumee river, the largest tributary to western Lake Erie, which is used to predict upcoming summer bloom severity. Maumee river spring phosphorus load, however, does not explain all the variation of bloom severity between years. Considering additional environmental parameters may help to better capture the physical and biogeochemical processes that regulate bloom severity, eventually leading to improved cyanobacterial forecasts which serve as an early warning for Lake Erie stakeholders. We aggregated various environmental parameters that may influence western Lake Erie cyanobacterial blooms to examine these factors as potential predictors for annual bloom severity. These included nitrogen and phosphorus loading from the Maumee river, freshwater discharge from the primary rivers and tributaries (Detroit, Huron, Raisin, Maumee, and Portage rivers), seasonal lake surface water temperature (mean winter, spring, and summer temperature), and Lake Erie winter ice extent and duration from 2002-2022. Empirical model results show that spring phosphorus loading, as total bioavailable phosphorus, from the Maumee river remains the dominant environmental factor controlling cyanobacterial blooms. However, additional environmental factors, such as Maumee river winter phosphorus loads and Lake Erie winter ice extent and timing, are likely important in modulating bloom severity, particularly in years with moderate phosphorus loads. Finally, we suggest incorporating mechanistic or rule-based models, in addition to empirical models, to better understand and predict annual cyanobacterial bloom severity. The updated models not only improve seasonal forecast accuracy which provides advanced warning of bloom severity to Lake Erie stakeholders, but also helps identify which factors we can better manage to reduce the frequency of severe blooms.

Keywords: phosphorus, winter ice-cover, Great Lakes, harmful algal blooms, empirical modeling

Introduction:

Western Lake Erie began experiencing intense cyanobacterial blooms in the 1960s and 1970s, which led, in part, to the passage of the Clean Water Act in 1972 (e.g., Kane et al. 2014). The point-source pollution reductions mandated by the act led to the recovery of the lake in the 1980s (Bridgeman et al. 2013). Since the early 2000s, however, annual cyanobacterial blooms have reappeared in the western basin, with up to a 1000 km^2 of bloom coverage during the most intense bloom years (Stumpf et al. 2012, 2016). One key factor in the resurgence of cyanobacterial blooms appeared to be the doubling of dissolved reactive phosphorus (DRP) concentrations, particularly in the Maumee river, starting in the mid-1990s through early-2000s (Baker et al. 2014). Almost all DRP is considered bioavailable and constitutes the largest fraction of total bioavailable phosphorus (TBP), the fraction of river phosphorus (P) loading that is directly available to phytoplankton (Baker et al. 2014). In addition, the Maumee river is the largest tributary to any Great Lake, dominating the river load in the southwestern part of Lake Erie's western basin (Figure S1; Bocaniov et al. 2023; Maccoux et al. 2016). Thus, there is a strong correlation between TBP loading from the Maumee river and annual cyanobacterial bloom severity (e.g., Kane et al.

50   2014; Stumpf et al. 2016).
51   Due to this relationship, multiple empirical models have been developed which demonstrate a
52   strong, positive relationship between TBP loading from the Maumee river and annual summer
53   cyanobacterial bloom intensity (Bertani et al. 2016; Ho and Michalak, 2017; Obenour et al. 2014; Scavia
54   et al. 2023; Stumpf et al. 2012, 2016). This suite of empirical models has been used in the annual western
55   Lake Erie Seasonal forecast, which provides an advanced warning of potential summer bloom severity to
56   Lake Erie water treatment plant operators, coastal businesses reliant on tourism, and the broader
57   management community. Overall, these empirical models have worked well in predicting upcoming
58   annual bloom severity; however, not all the variation between years is completely explained by TBP
59   loading (e.g., Ho and Michalak, 2017; Scavia et al. 2021, 2023; Stumpf et al. 2016). For example, Stumpf
60   et al. (2016) hypothesized that early summer water temperatures may influence the overall bloom uptake
61   of TBP, with cooler temperatures in June leading to reduced TBP uptake and bloom growth in subsequent
62   months. Other studies have proposed that TBP loading from multiple previous years may be an additional
63   factor in annual bloom severity (Ho and Michalak, 2017; Scavia et al. 2023), however, this would likely
64   alter longer term trends in the base level of the blooms, rather than specific interannual bloom variability.
65   In addition, with climate change, western Lake Erie is also seeing other changes in physical properties
66   such as decreased winter ice cover (Ozersky et al. 2021; Wang et al. 2012).
67   In this study, we examined several environmental parameters that likely influence bloom severity,
68   including nutrient loading from the Maumee river and various physical characteristics of the western Lake
69   Erie basin, coupled with empirical modeling, to identify potential environmental factors that influence
70   interannual variability of cyanobacterial bloom severity. We hypothesized that spring TBP loading from
71   the Maumee river would be the dominant driver of cyanobacterial bloom intensity, but that additional
72   environmental variables would further modulate annual bloom severity. Ultimately, we seek to use these
73   updated empirical models to better predict upcoming summer bloom severity (NOAA, 2022) and will be
74   directly incorporated into the annual western Lake Erie Seasonal forecasts provided to lake managers.
75
76   **Methodology:**
77   ***Annual cyanobacterial bloom intensity***: Western Lake Erie cyanobacterial bloom intensity was obtained
78   from satellite imagery as described in Stumpf et al. (2016). The data sets were obtained from the Medium
79   Resolution Imaging Spectrometer (MERIS, 2002-2011), the Moderate Resolution Imaging
80   Spectroradiometer (MODIS, 2012-2018), and the Sentinel-3 Ocean and Land Colour Instrument (OLCI,
81   2019-2022), as the replacement for MERIS. Wynne et al. (2021) applied a multiplier to the MODIS data
82   which allows for the direct comparison of MODIS to MERIS/OLCI images. Sequential 10-day composite
83   images were generated following the methods of Stumpf et al. (2012) and Wynne and Stumpf (2015). The
84   composites were created from the maximum cyanobacterial chlorophyll-related index (CI) value for each
85   map pixel from individual scenes (Stumpf et al. 2012; Wynne and Stumpf, 2015; Figure S2). Generating
86   10-day composites provides several advantages over single-day images by removing interference from
87   clouds and capturing maximum bloom extent during calm winds when cyanobacteria can be more easily
88   seen by satellite. Finally, for each 10-day period, all the pixels in the western basin were summed to
89   obtain the total biomass in CI units (Stumpf et al. 2016). The resulting annual CI max was used to
90   estimate annual bloom severity. CI has units of dimensionless reflectance but can be equated to
91   cyanobacterial concentration by: 1 CI$=10^8$ *Microcystis* cells mL^{-1} (Lunetta et al. 2015). In addition, a
92   severity index (SI), which is used in annual bloom severity reporting (NOAA, 2022), was used to identify
93   annual blooms with low and high biomass. The SI uses the average of the biomass in the three 10-day
94   composites with the greatest biomass, normalized by log-scale with a value of 1 corresponding to the SI
95   for 2005, and 10 corresponding to the SI in 2011 (Table S3).
96
97   ***Environmental and nutrient data***: We aggregated a variety of environmental and nutrient loading data
98   from openly available sources, which could be used as potential predictors of cyanobacterial bloom

99   severity. Freshwater and nutrient loading from the Maumee river were obtained from Heidelberg
100   University's National Center for Water Quality Research (NCWQR, 2022), including TBP and dissolved
101   inorganic nitrogen (DIN) from 2002-2022. Spring monthly (March, April, May, June, July), spring
102   (March-July), and winter (November-February) loads were summed from daily freshwater flow and
103   nutrient concentrations (Richards et al. 2009). TBP was calculated from total P and DRP following Baker
104   et al. (2014), where total bioavailable particulate P was calculated as the proportion (β=0.26) of total
105   particulate P (total particulate P=total P - DRP). A settling term was also applied to the total bioavailable
106   particulate P to account for settling of particulate P from the location of monitoring (Waterville, OH) to
107   western Lake Erie (S=0.70). Finally, TBP is the sum of total bioavailable particulate P, scaled for settling,
108   and DRP (Baker et al. 2014). DIN loads were calculated by summing nitrate (NO_3^-) and ammonium
109   (NH_4^+) loads. NO_3^- loads were calculated as noted above, whereas monthly NH_4^+ loads were calculated by
110   extrapolating weekly NH_4^+ concentrations using the monthly NH_4^+:NO_3^- ratio. TBP and DIN loads were
111   used to calculate molar DIN:TBP (Figure S3).
112   In addition, we estimated the proportion of freshwater flow from the Maumee river as compared
113   to other major tributaries entering the western basin including: the Huron, Raisin, and Portage rivers
114   (Figure S1). Freshwater flows were obtained from the most downstream USGS gauging station (Table
115   S1) and scaled to the ungauged watershed area. The proportion of Maumee river flow was calculated by
116   dividing the Maumee river flow by the total tributary flow for the spring period. We estimated spring
117   (March-July) Detroit river flow using monthly connecting channel flows provided by the U.S. Army
118   Corps of Engineers-Detroit District (Figure S1, S4; D.C. Fielder, USACE-Detroit District, Detroit, MI,
119   USA, pers. comm.). Flow from the Detroit river was not included in the proportion of Maumee river flow.
120   Due to its low TBP concentration (~0.01 μg L^{-1} vs. 0.1 μg L^{-1} for the Maumee River), the Detroit river
121   acts as a diluent in the western basin, rather than a source of TBP for summer cyanobacteria blooms (e.g.,
122   Bocaniov et al. 2023; Burniston et al. 2018; Downing et al. 2001).
123   Finally, we included several environmental parameters such as lake surface water temperature
124   (SST), winter ice cover extent and duration, and the number of low wind speed days during the early
125   bloom season. Lake SST was estimated via MODIS thermal data obtained from NASA Giovanni (NASA,
126   2023a, 2023b; Wynne et al. 2022). Night and day SST were averaged seasonally for winter (January-
127   March), spring (April-June), and summer (July-September) from 2003-2022. Winter ice cover extent and
128   duration from 2002-2022 for Lake Erie was obtained from NOAA's Great Lakes Environmental Research
129   Laboratory and included the percent of maximum ice cover and day of maximum ice extent (GLERL,
130   2022). The day of maximum ice extent was then converted to day of year (DOY). While we acknowledge
131   that the central and eastern basins likely play a role in controlling Lake Erie ice cover, we assume that
132   patterns observed in all Lake Erie basins are representative of winter conditions in western Lake Erie.
133   Finally, we used the number of low wind speed days during the early bloom season (July-August) as an
134   estimate of internal nutrient loading due to potential periods of hypoxia/anoxia (Bridgeman et al. 2006).
135   Briefly, hourly wind data were obtained from the Toledo Light weather station (NDBC, 2023) and
136   averaged to daily values. Daily wind values <3 m s^{-1} were summed for the two-month period (July-
137   August) for each year from 2005-2022 (Figure S5). We note that our identification of hypoxic/anoxic
138   events is an over-simplification and does not explicitly account for other environmental conditions that
139   might lead to these events, including entrainment of hypoxic bottom water from the central basin (Jabbari
140   et al. 2019) and increased respiration due to elevated temperature (Bridgeman et al. 2006).
141
142   ***Empirical modeling:*** We used several empirical models to identify important predictors of western Lake
143   Erie annual bloom severity. First, we applied the standard least squares linear regression model developed
144   by Stumpf et al. (2016) which identified a strong, positive relationship between $\log_{10}$(CI max) and spring
145   Maumee river TBP loading. Specifically, we used the unweighted, March-July TBP model following:
146
147          $$CI\ Max = b * 10^{(a*\text{Mar-Jul TBP Loading})}$$ (Eq. 1)

148   where b=0.37 and a=3.26×10^3 (Stumpf et al. 2016). We then updated the standard least squares linear
149   regression between $\log_{10}$(CI max) and spring TBP loading using available data from 2002-2022. We
150   excluded 2003 and 2013 as anomalous years during model fitting (n=19; Stumpf et al. 2016).
151   Next, we categorized historical blooms based on their annual SI where low blooms were
152   classified as SI<5 and high blooms were categorized as SI>5 (Table S3). We then applied a single logistic
153   regression to the categorized data using Maumee river spring TBP loading as the predictor variable. We
154   excluded 2003 as a year with an anomalously high spring TBP loading, but low overall bloom severity
155   (Stumpf et al. 2016).
156   We also identified other environmental variables that could explain some of the additional
157   variability in annual cyanobacterial bloom severity. To accomplish this, we conducted a series of multiple
158   linear regressions. We first removed any collinear (Pearsons correlation coefficient, ρ>0.7) environmental
159   variables (Table S2). We then developed an algorithm to randomly select up to two environmental
160   variables to be used in forward, stepwise multiple linear regression. The developed models were then
161   ranked by the Akaike Information Criteria (AIC) and the models within 2 AIC of the best fit model were
162   selected for further analysis. For multiple linear regression, we used all years with available data for
163   model fitting (2005-2007, 2009-2014, 2016-2019, 2021; n=14; Table S3). All data visualizations and
164   empirical modeling were conducted in R v4.1.0 (R Core Team, 2021).
165
166 **Results:**
167   Overall, the composited 10-day accumulated CI was highly variable (Figure 1). 2011 had the
168   highest median 10-day CI (7.72), followed by 2015 (7.45), and 2013 (5.73). 2011 also had the highest CI
169   max (29.9), followed by 2015 (29.2), then 2017 (15.5). 2019 had the highest CI max in the past five years
170   (11.5) while 2022 had the highest median 10-day CI (3.3). The aggregated environmental predictors were
171   also highly variable among years and between months and seasons (Figure 2, S3, S4, S5; Table S4). Total
172   spring TBP loading ranged from 89 metric tons (mton) in 2012 to 722 mton in 2015 with a mean of
173   344±155 mton (Figure 2). For July, TBP loading ranged from 1.33-216 mton in 2012 and 2015
174   respectively, with a mean of 40.7±58.3 mton (Figure 2). Throughout the study period, the proportion of
175   Maumee river flow to total tributary flow (i.e., from the Maumee, Huron, Raisin, and Portage rivers) into
176   western Lake Erie remained >65% with a mean of 77±4.2% (Figure S4). Seasonal lake SST was
177   relatively consistent among years (2004-2022; Figure 2; Table S4). Percent of maximum winter ice cover
178   for Lake Erie was generally high with a few years of <50% ice cover (2002, 2006, 2012, 2017, 2020) and
179   mean percent maximum winter ice cover of 75±32% (Figure 3). Finally, the day of year (DOY) of ice off
180   ranged from 3 (2002) to 70 (2008) with a mean of 43±16 days (Figure 2).
181   Prior to empirical modeling, we removed environmental and nutrient parameters that were highly
182   linearly correlated (ρ>0.70; Table S2). Monthly and seasonal TBP loading were highly correlated with
183   freshwater discharge (Q) and DIN loading. Ultimately, we removed all monthly and seasonal Q and DIN
184   loading. In addition, we removed all monthly TBP loading for March, April, May, and June; only
185   retaining winter (November-February), spring (March-July), and July TBP loading for modeling.
186
187 *Empirical model results:* First, we updated the logarithmic relationship between CI max and Maumee
188   river spring TBP loading following methods in Stumpf et al. (2016) and using the full dataset (2002-2022,
189   excluding 2003, 2013; Figure 3). Similar to previous studies, we found a strong, positive correlation
190   between $\log_{10}$(CI max) and Maumee river spring TBP loading (adjusted r^2=0.76; Table 1). Model
191   residuals (modeled-observed) suggest the 2023 model (this study) was marginally better at predicting
192   annual CI max, especially for high bloom years such as in 2015 (Figure 2). In both cases, 2022 had a
193   relatively large error (Figure 3b). The two fits matched closely for TBP loading<400 mtons, and deviated
194   as the load increased.
195   Logistic regression results suggest Maumee river spring TBP loading can be used to distinguish
196   overall bloom severity (low vs. high blooms), for loads <273 mtons or >332 mtons. Between 273-332

197 mtons, the model was less able to predict bloom severity, which included three years with low blooms:
198 2007, 2018, 2020, and three with high blooms: 2009, 2021, and 2022. Finally, multiple linear regression
199 identified three models that were within 2 AIC of the best-fit model (Table 1) and included a combination
200 of: a positive relationship with Maumee river spring TBP loading and DOY of maximum ice extent, but a
201 negative relationship with Maumee river winter TBP loading. Adjusted r^2 for the three identified models
202 ranged from 0.65-0.66.
203
204 **Discussion:**
205 ***Temporal variability of western Lake Erie cyanobacterial blooms:*** Overall, there was a high degree of
206 inter- and intra-annual variability in the observed 10-day accumulated CI and CI max from 2002-2022,
207 which likely reflects a range of competing external (i.e., nutrient loading, freshwater delivery) and
208 internal (i.e., internal nutrient loading, light availability, water column mixing) environmental factors that
209 modulate bloom severity. Several years had a wide range in 10-day accumulated CI values, including
210 years with anonymously high 10-day periods (2003, 2008, 2009, 2010, 2011, 2017; Figure 1), that
211 skewed the overall CI max. We note that the current method of assessing annual bloom severity (i.e., CI
212 max) does not explicitly take into account temporal variability during the bloom. Thus, this may
213 overestimate annual bloom severity by relying on a single 10-day period with abnormally high
214 accumulated CI or it might underestimate bloom severity for years when there were multiple 10-day
215 periods with similarly high accumulated CI. However, the CI max and CI average for the 30-day period
216 are closely correlated (Stumpf et al., 2016). While updates to the CI max are likely needed to take into
217 consideration both the maximum and the temporal intensity of the bloom over multiple 10-periods, the
218 current CI max does reflect the maximum bloom severity, such that models of CI max can be used to
219 reliably predict maximum annual bloom severity (Stumpf et al. 2012, 2016).
220
221 ***Spring phosphorus loading from the Maumee river is the dominant driver of annual bloom severity***:
222 Following Stumpf et al. (2016), Maumee river spring TBP loading was still a reliable predictor of annual
223 maximum bloom severity (Table 1; Figure 3). The most apparent difference between the Stumpf et al.
224 (2016) model and the model presented here (2023) was for high bloom years, primarily driven by 2015,
225 but also due to intense blooms in 2017 and 2019. We do note that this regression results in a non-linear
226 relationship between CI max and TBP loading (Figure 3), such that higher spring TBP loads lead to a
227 proportionally larger annual bloom severity as compared to smaller TBP spring loads. A similar non-
228 linearity has appeared in other models (e.g., Obenour et al. 2014; Scavia et al., 2023). This change with
229 load suggests that either small blooms are less effective at using the available TBP, or less of the TBP is
230 available to blooms during low load years.
231   Most relationships developed using TBP loading as a predictor of bloom severity, rely on TBP
232 loads from the Maumee river, thus making the assumption that the majority of P originates from the
233 Maumee watershed (e.g., Ho and Michalak, 2017; Scavia et al. 2023; Stumpf et al. 2016). To confirm
234 this, we used the annual proportion of Maumee river flow to scale Maumee river TBP loading to total
235 potential TBP loading for the entire western basin, as a predictor for annual bloom severity (Figure S6).
236 While this resulted in a slight reduction in total model RMSE (4.66 to 3.91 CI max for un-scaled and
237 scaled TBP loads, respectively), overall, the total TBP loading did not substantially improve model
238 results. While not an exact representation of overall TBP loading to western Lake Erie, this supports the
239 assumption that Maumee river loading is the dominant source of external P for summer cyanobacterial
240 blooms, and is suitable for empirical models used to estimate overall bloom severity. Future studies and
241 mechanistic modeling efforts should, however, consider loads from additional river sources.
242   Following studies which have found the Detroit river to be a substantial source of river P load to
243 Lake Erie, we also included spring Detroit river load as a potential environmental predictor. While the
244 Detroit river is a substantial freshwater and nutrient source to Lake Erie (Bocaniov et al. 2023; Burniston
245 et al. 2018; Scavia and Calappi, 2023), our results show it was not a factor in determining summer

246 cyanobacterial bloom severity. As compared to the Maumee, the Detroit river has about an order of
247 magnitude lower bioavailable P concentrations (~0.01 µg L^{-1} total P, with about half being DRP vs. ~0.1
248 µg L^{-1} for Maumee TBP flow weighted mean concentration; Bocaniov et al. 2023; Burniston et al. 2018;
249 Scavia and Calappi, 2023). The Detroit river also tends to have relatively little interannual variability in P
250 load, with more variation occurring in long term trends associated with Great Lakes water levels (Scavia
251 and Calappi, 2023). In addition, blooms were shown to develop in Maumee Bay during the early bloom
252 season (July-August) where TBP concentrations from the Maumee river are high, and are rarely observed
253 in the Detroit river plume (Wynne and Stump, 2015). While we cannot discount the Detroit river flow as
254 an important contributor to background TP concentrations in the western and central basins of Lake Erie,
255 our results suggest it is not the driving factor of western basin cyanobacterial bloom intensity.
256   Finally, we identified a relationship with spring TBP loading to separate low and high blooms,
257 based on bloom SI, where low blooms were identified with >75% certainty at TBP loads <273 mton and
258 high blooms were identified with >75% certainty at TBP loads >332 mton (Figure 4). This suggests there
259 are TBP thresholds at which we can reliably predict a low versus high bloom which may help to guide
260 early season predictions for upcoming bloom severity. Conversely, this also implies there is a range of
261 TBP loads (273-332 mton) where we are unable to use spring TBP loading to reliably predict bloom size,
262 suggesting additional environmental parameters may be necessary to fully separate blooms by size.
263   While Maumee river spring TBP loading was identified as the primary environmental variable
264 constraining annual CI max in all regression models, one multiple linear regression model also identified
265 a negative relationship with Maumee river winter TBP loading. The negative relationship was surprising,
266 given the strong correlation between spring TBP loading and annual bloom severity, as well as studies
267 that have identified western Lake Erie as primarily P-limited (Chaffin et al. 2014; Steffen et al. 2014). We
268 hypothesize the negative relationship with winter TBP loading is a proxy for winter temperature and
269 precipitation conditions that influence both watershed and in-lake processes. First, higher winter TBP
270 loading implies a warmer winter, which could decrease spring runoff due to drier soil antecedent
271 conditions. Second, differences in winter TBP loading may influence winter under-ice (Beall et al. 2015)
272 and/or spring diatom blooms that occur nearly annually in western Lake Erie (Reavie et al. 2014).
273 Specifically, higher winter TBP loading may fuel larger winter or spring phytoplankton growth, altering
274 the timing, duration, and intensity of these blooms which may influence the availability of P into summer.
275   While external TBP loading was shown to be a dominant driver of annual bloom severity, we
276 note that we did not explicitly include a term for internal P loading, which may be a potentially important
277 source of P for western Lake Erie cyanobacterial blooms (Bocaniov et al. 2023; Matisoff et al. 2016). The
278 overall magnitude of internal P-load would be influenced by both sediment disturbance and by the
279 development of near-bottom hypoxia, which varies annually (Matisoff et al. 2016). However, determining
280 the exact role of hypoxia-driven internal P-loading, requires additional modeling efforts to account for
281 temporary water column stratification and subsequent hypoxic events. Additional internal loading during
282 high biomass years, when hypoxia might be more likely, may be an additional source of P during high
283 bloom years, contributing to the non-linear relationship observed between CI max and spring TBP.
284
285 ***Other environmental variables modulate annual bloom severity:*** While Maumee river spring TBP
286 loading was identified as a reliable predictor of cyanobacterial bloom intensity, there was still variability
287 in our ability to accurately estimate annual blooms. The TBP model assumes that all TBP loading from
288 the Maumee river is used by the cyanobacterial bloom; however, in some years, this relationship led to an
289 overestimation of bloom severity (i.e., 2003, 2007) while in others, it led to an under-estimation (i.e.,
290 2013, 2022; Figure 3). In addition, the logistic relationship used to identify low versus high blooms
291 under-estimated blooms in 2009, 2021, and 2022, but over-estimated blooms in 2007, 2018, and 2020,
292 which corresponded to years with TBP loads between 273-332 mton. Below, we discuss other
293 environmental variables that may further constrain annual bloom severity.
294   For example, Stumpf et al. (2016) identified June water temperature as a modulator for the spring

295 TBP versus CI max relationship. Specifically, they hypothesized that a cool June (water temperature
296 <20°C) resulted in lower uptake of TBP during the early bloom, negating the influence of the July TBP
297 load. Instead, they used March-June TBP loading to estimate bloom severity for years with June water
298 temperatures <20°C, which helped to adjust the predicted CI max (Figure 3). In 2007, an anomalously
299 high flow event occurred in late August (Figure S7; Baker et al. 2014; Bridgeman et al., 2013), which
300 likely disrupted the bloom by either introducing turbidity or flushing the bloom into the lake. Wynne and
301 Stumpf (2015) found that blooms begin rapid expansion during the last 10 days of August; thus, it is
302 possible that this elevated discharge event in August 2007 led to a smaller bloom than was predicted by
303 spring TBP loading.
304   We also hypothesize that July TBP loading may further modulate bloom severity. For example,
305 low bloom years in the mid-TBP loading range (273-332 mton; i.e., 2007, 2018, 2010; Figure 4) had July
306 TBP loads <5.75 mtons which is substantially lower than both the mean (40.7 mton) and median (9.7
307 mton; Table S4) loads. In contrast, both 2021 and 2022, which were high blooms under this same TBP
308 range, had substantially higher July loads (90.3 and 42.5 mton, respectively). The 2009 bloom had a low
309 July TBP load (2.2 mton), but had no load feature that distinguished it from the low bloom years within
310 this TBP range. The complexity of bloom variability is evident, but the result for five out of six years
311 suggests that the timing of TBP loads, as lower or higher July loads may modulate the overall spring TBP
312 loading relationship. The relative influence of July loads will likely become clearer over the next few
313 years, given that 7 of the last 10 years had July loads exceeding the 20-year median (Table S3, S4).
314   Finally, we speculate that winter ice cover might also influence annual bloom severity. We found
315 that the DOY of maximum ice extent was identified as having a weak, positive relationship with bloom
316 severity (Table 2). Similarly, extensive Lake Erie ice cover was more common in winters prior to high
317 bloom years, where 9 out of 10 years had ≥84% maximum ice cover, while in low bloom years, 4 out of
318 10 years had ≤22% cover. As noted earlier, we hypothesize that cooler winter and spring temperatures
319 alter the delivery of spring nutrient loading as well as the duration and intensity of under-ice winter (Beall
320 et al. 2016) and spring diatom blooms (Reavie et al. 2016), which could influence summer blooms.
321
322 ***Study limitations and future directions:*** While this study represents an increase in the amount of data
323 used to identify environmental predictors of western Lake Erie cyanobacterial bloom intensity, we still
324 acknowledge some study limitations. First, we focused on TBP due to the extensive literature which has
325 identified Maumee river TBP loading as the dominant factor driving cyanobacterial blooms (Bertani et al.
326 2016; Ho and Michalak, 2017; Obenour et al. 2014; Scavia et al. 2023; Stumpf et al. 2012, 2016). While
327 there is an inherent autocorrelation among Q, DIN, and TBP loading (Table S2), spring DIN:TBP in the
328 Maumee river suggests substantially more N than P in incoming loads (mean molar DIN:TBP=118±33;
329 Figure S3; Table S4). Similarly, Stow et al. (2022) found that TBP loads from the Maumee river have
330 almost doubled over the past 30+ years while DIN loads have remained relatively constant, all while
331 summer cyanobacterial bloom intensity has increased. Finally, the western basin is largely P-limited,
332 especially during the early and peak bloom season, and does not become N-limited until after the bloom is
333 established (Chaffin et al. 2014; Steffen et al. 2014). These lines of evidence suggest that TBP is still the
334 dominant driver of annual cyanobacterial bloom intensity, though N-limitation is likely a factor during the
335 late bloom season (Chaffin et al. 2013; Jankowiak et al. 2019).
336   Second, the majority of annual bloom prediction models rely on statistical relationships between
337 Maumee river TBP loading and cyanobacterial bloom severity. While TBP loads are the dominant driving
338 factor of bloom intensity, this study suggests other environmental factors modulate this relationship. With
339 the relatively small number of study years (~20 years) and the large number of environmental variables
340 (>20) identified, statistical models cannot accurately represent how additional environmental variables
341 may work to modulate bloom severity. While they can guide hypotheses, additional years of data are
342 needed to more clearly elucidate these relationships.
343   These limitations suggest additional modeling strategies are likely needed. The information

344 gleaned from the broad range of statistical models (e.g., Ho and Michalak, 2017; Scavia et al. 2023;
345 Stumpf et al. 2016), along with experimental and modeling studies, can be used to develop mechanistic
346 models which may help to fill some of the gaps left by empirical models, including dynamic, process-
347 based, numerical models like the Western Lake Erie Ecosystem Model (WLEEM; e.g., Verhamme et al.
348 2016). Empirical models are also unable to account for any fundamental ecosystem changes that may
349 have occurred which influence annual bloom severity, limiting the ability of these relationships to predict
350 future ecosystem changes. Finally, in a predictive framework, we are limited to environmental parameters
351 that can be appropriately forecasted, such as river discharge forecasts produced for the Maumee river
352 (OHRFC, 2023) or previously observed winter or spring temperature. Thus, while empirical models are
353 useful in predicting upcoming bloom severity, we argue that additional, mechanistic or rule-based models,
354 will help to better identify other environmental processes modulating bloom severity and enhance our
355 ability to predict upcoming bloom severity.
356
357 ***Providing advanced warning for the Lake Erie management community:*** Recurring summer
358 cyanobacterial HABs pose severe negative impacts on human, animal, and ecological health, which
359 impact drinking water resources, lead to beach closures, and may discourage or prevent people from
360 fishing, swimming, boating, or visiting the shoreline. Since 2016, the annual western Lake Erie seasonal
361 forecast has provided advanced warning to lake managers, drinking water treatment plant operators, and
362 the general public on the potential bloom severity for the upcoming summer. The updated and new
363 empirical models developed in this study will be directly incorporated into these annual forecasts,
364 improving our ability to provide advanced warning of potential bloom severity to the Lake Erie
365 management community. In addition, previous relationships between Maumee river TBP loading and
366 annual bloom severity were used to identify the necessary 40% reduction in watershed TBP loading
367 needed to reduce overall cyanobacterial bloom severity (Annex 4 Objectives and Targets Task Team,
368 2015). Additional analysis from this study could help to identify additional environmental factors to better
369 manage the frequency of severe blooms in the future.
370
371 **Conclusions:**
372   Overall, results from this study show that Maumee river spring TBP loading is still the dominant
373 factor controlling annual western Lake Erie cyanobacterial bloom severity and can be reliably used to
374 predict upcoming bloom severity (Bertani et al. 2016; Ho and Michalak, 2017; Obenour et al. 2014;
375 Scavia et al. 2023; Stumpf et al. 2012, 2016). Thus, we recommend continuing to use established
376 relationships between Maumee river spring TBP loading and annual bloom severity for seasonal
377 cyanobacterial bloom predictions. However, our results suggest additional models should be considered to
378 distinguish the threshold between a low or high blooms, especially during the early spring season when
379 predictions of annual spring TBP loading may be more variable. Results also suggest we are likely
380 missing key environmental factors that further modulate bloom severity, particularly in the mid-TBP
381 loading range (273-332 mton). Under these mid-range conditions, additional environmental parameters
382 may be needed to better predict bloom severity. Due to the limited years of data, these additional
383 relationships are still under consideration, though winter TBP loading, as well as winter ice dynamics,
384 along with July TBP loading, may be important. We suggest that alternative, mechanistic or rule-based
385 models which can explicitly incorporate known mechanisms, are an important future direction for annual
386 bloom severity predictions, which would also allow for future climate predictions to be incorporated into
387 models. Importantly, these and future model updates will be incorporated into the NOAA western Lake
388 Erie Seasonal Forecasts to provide advanced warning of upcoming summer bloom severity to the Lake
389 Erie management community and can help identify additional environmental factors to better manage the
390 frequency of severe blooms.
391
392

403 **References:**
404 Annex 4 Objectives and Targets Task Team, 2015. Recommended phosphorus loading targets for Lake
405   Erie.
406 Baker, D.B., Confesor, R., Ewing, D.E., Johnson, L.T., Kramer, J.W., Merryfield, B.J. 2014. Phosphorus
407   loading to Lake Erie from the Maumee, Sandusky and Cuyahoga rivers: The importance of
408   bioavailability. Journal of Great Lakes Research, 40(3), 502-517.
409 Beall, B.F., Twiss, M.R., Smith, D.E., Oyserman, B.O., Rozmarynowycz, M.J., Binding, C.E.,
410   Bourbonniere, R.A., Bullerjahn, G.S., Palmer, M.E., Reavie, E.D., Waters, L.M., Woityra, L.W.,
411   McKay, R.M. 2016. Ice cover extent drives phytoplankton and bacterial community structure in a
412   large north-temperate lake: implications for a warming climate. Environ Microbiol, 18(6), 1704-
413   1719.
414 Bertani, I., Obenour, D.R., Steger, C. E., Stow, C.A., Gronewold, A.D., Scavia, D. 2016. Probabilistically
415   assessing the role of nutrient loading in harmful algal bloom formation in western Lake Erie.
416   Journal of Great Lakes Research, 42(6), 1184-1192.
417 Bocaniov, S.A., Scavia, D., Van Cappellen, P. 2023. Long-term phosphorus mass-balance of Lake Erie
418   (Canada-USA) reveals a major contribution of in-lake phosphorus loading. Ecological
419   Informatics, 77: 102131. https://doi.org/10.1016/j.ecoinf.2023.102131
420 Bridgeman, T.B., Chaffin, J.D., Filbrun, J.E. 2013. A novel method for tracking western Lake Erie
421   Microcystis blooms, 2002–2011. Journal of Great Lakes Research, 39(1), 83-89.
422 Bridgeman, T.B., Schloesser, D.W., Krause, A.E. 2006. Recruitment of *Hexagenia* mayfly nymphis in
423   western Lake Erie linked to environmental variability. Ecological Applications, 16(2), 601-611.
424 Burniston, D., Dove, A., Backus, S., Tomphson, A. 2018. Nutrient concentrations and loading in the St.
425   Clair River-Detroit River Great Lakes interconnecting channel. Journal of Great Lakes Research,
426   44(3): 398-411. https://doi.org/10.1016/j.jglr.2018.02.005
427 Chaffin, J.D., Bridgeman, T.B., Bade, D.L. 2013. Nitrogen Constrains the Growth of Late Summer
428   Cyanobacterial Blooms in Lake Erie. Advances in Microbiology, 03(06), 16-26.
429 Chaffin, J.D., Bridgeman, T.B., Bade, D.L., Mobilian, C.N. 2014. Summer phytoplankton nutrient
430   limitation in Maumee Bay of Lake Erie during high-flow and low-flow years. Journal of Great
431   Lakes Research, 40(3), 524-531.
432 Downing, J.A., Watson, S.B., McCauley, E. 2001. Predicting cyanobacteria dominance in lakes.
433   Canadian Journal of Fisheries and Aquatic Sciences, 58(10): 1905-1908.
434   https://doi.org/10.1139/f01-143
435 GLERL. 2022. Historical Ice Cover. https://www.glerl.noaa.gov/data/ice/#historical [Date accessed: 21
436   Aug. 2023]
437 Ho, J.C., Michalak, A.M. 2017. Phytoplankton blooms in Lake Erie impacted by both long-term and
438   springtime phosphorus loading. Journal of Great Lakes Research, 43(3), 221-228.
439 Jabbari, A., Ackerman, J.D., Boegman, L., Zhao, Y. 2019. Episodic hypoxia in the western basin of Lake
440   Erie. Limnology and Oceanography, 64: 2220-2236. https://doi.org/10.1002/lno.11180
441 Jankowiak, J., Hattenrath-Lehmann, T., Kramer, B.J., Ladds, M., Gobler, C.J. 2019. Deciphering the

442   effects of nitrogen, phosphorus, and temperature on cyanobacterial bloom intensification,
443   diversity, and toxicity in western Lake Erie. Limnology and Oceanography, 64(3), 1347-1370.
444 Kane, D.D., Conroy, J.D., Richards, R.P., Baker, D.B., Culver, D.A. 2014. Re-eutrophication of Lake
445   Erie: Correlations between tributary nutrient loads and phytoplankton biomass. Journal of Great
446   Lakes Research, 40(3), 496-501.
447 Lunetta, R.S., Schaeffer, B.A., Stumpf, R.P., Keith, D., Jacobs, S.A., Murphy, M.S. 2015. Evaluation of
448   cyanobacteria cell count detection derived from MERIS imagery across the eastern USA. Remote
449   Sensing of Environment, 157, 24-34.
450 Maccoux, M.J., Dove, A., Backus, S.M., Dolan, D.M. 2016. Total and soluble reactive phosphorus
451   loadings to Lake Erie. Journal of Great Lakes Research, 42(6), 1151-1165.
452 Matisoff, G., Kaltenberg, E.M., Steely, R.L., Hummel, S.K., Seo, J., Gibbons, K.J., Bridgeman, T.B., Seo,
453   Y., Behbahani, M., James, W.F., Johnson, L.T., Doan, P., Dittrich, M., Evans, M A., Chaffin,
454   J.D. 2016. Internal loading of phosphorus in western Lake Erie. Journal of Great Lakes Research,
455   42(4), 775-788.
456 NASA. 2023a. Sea Surface Temperature at 11 microns (Day),
457   MODIS_L3m_SST_Monthly_4km_vR2019.0. NASA Giovanni.
458   https://giovanni.gsfc.nasa.gov/giovanni/#service=ArAvTs&starttime=2002-07-
459   01T00:00:00Z&endtime=2022-09-
460   30T23:59:59Z&bbox=83.55W,41.45N,82.55W,42.05N&data=MODISA_L3m_SST_Monthly_4k
461   m_R2019_0_sst%2CMODISA_L3m_NSST_Monthly_4km_R2019_0_sst&portal=GIOVANNI&
462   format=json [Accessed 21 Aug. 2023]
463 NASA. 2023b. Sea Surface Temperature at 11 microns (Night),
464   MODIS_L3m_NSST_Monthly_4km_vR2019.0. NASA Giovanni.
465   https://giovanni.gsfc.nasa.gov/giovanni/#service=ArAvTs&starttime=2002-07-
466   01T00:00:00Z&endtime=2022-09-
467   30T23:59:59Z&bbox=83.55W,41.45N,82.55W,42.05N&data=MODISA_L3m_SST_Monthly_4k
468   m_R2019_0_sst%2CMODISA_L3m_NSST_Monthly_4km_R2019_0_sst&portal=GIOVANNI&
469   format=json [Accessed: 21 Aug. 2023]
470 NCWQR. 2022. Heidelberg Tributary Loading Program (HTLP) Dataset Zenodo.
471   https://doi.org/https://doi.org/10.5281/zenodo.6606949 [Accessed: 21 Aug. 2023]
472 NDBC. 2023. Station THL01-Toledo Light No. 2 OH.
473   https://www.ndbc.noaa.gov/station_page.php?station=thlo1 [Accessed: 21 Aug. 2023]
474 NOAA. 2022. NOAA Western Lake Erie Harmful Algal Bloom Seasonal Forecast.
475 Obenour, D.R., Gronewold, A.D., Stow, C.A., Scavia, D. 2014. Using a Bayesian hierarchical model to
476   improve Lake Erie cyanobacteria bloom forecasts. Water Resources Research, 50(10), 7847-
477   7860.
478 OHRFC. 2023. NWS Ohio River Forecast Center. Retrieved March 27, 2023 from
479   https://www.weather.gov/ohrfc/#
480 Ozersky, T., Bramburger, A.J., Elgin, A.K., Vanderploeg, H.A., Wang, J., Austin, J.A., Carrick, H.J.,
481   Chavarie, L., Depew, D.C., Fisk, A.T., Hampton, S.E., Hinchey, E.K., North, R.L., Wells, M.G.,
482   Xenopoulos, M.A., Coleman, M.L., Duhaime, M.B., Fujisaki-Manome, A., McKay, R.M.,
483   Meadows, G.A., Rowe, M.D., Sharma, S., Twiss, M.R., Zastepa, A. 2021. The Changing Face of
484   Winter: Lessons and Questions From the Laurentian Great Lakes. Journal of Geophysical
485   Research: Biogeosciences, 126(6).
486 Reavie, E.D., Barbiero, R.P., Allinger, L.E., Warren, G.J. 2014. Phytoplankton trends in the Great Lakes,
487   2001–2011. Journal of Great Lakes Research, 40(3), 618-639.
488 Richards, R.P., Baker, D.B., & Crumrine, J.P. 2009. Improved water quality in Ohio tributaries to Lake
489   Erie: A consequence of conservation practices. Journal of Soil and Water Conservation, 64(3),
490   200-211.

491   Scavia, D., Calappi, T.J. 2023. Detroit River load estimation: The need for a new monitoring approach.
492   Journal of Great Lakes Research, 49(4): 941-948.https://doi.org/10.1016/j.jglr.2023.05.007
493   Scavia, D., Wang, Y.C., Obenour, D.R. 2023. Advancing freshwater ecological forecasts: Harmful algal
494   blooms in Lake Erie. Sci Total Environ, 856(Pt 1), 158959.
495   Scavia, D., Wang, Y.C., Obenour, D.R., Apostel, A., Basile, S.J., Kalcic, M.M., Kirchhoff, C.J., Miralha,
496   L., Muenich, R.L., Steiner, A.L. 2021. Quantifying uncertainty cascading from climate,
497   watershed, and lake models in harmful algal bloom predictions. Sci Total Environ, 759, 143487.
498   Steffen, M.M., Belisle, B.S., Watson, S.B., Boyer, G.L., Wilhelm, S.W. 2014. Status, causes and controls
499   of cyanobacterial blooms in Lake Erie. Journal of Great Lakes Research, 40(2), 215-225.
500   Stow, C.A., Stumpf, R.P., Rowe, M.D., Johnson, L.T., Carrick, H.J., Yerubandi, R. 2022. Model
501   assumptions limit implications for nitrogen and phosphorus management. Journal of Great Lakes
502   Research, 48(6), 1735-1737.
503   Stumpf, R.P., Johnson, L.T., Wynne, T.T., Baker, D.B. 2016. Forecasting annual cyanobacterial bloom
504   biomass to inform management decisions in Lake Erie. Journal of Great Lakes Research, 42(6),
505   1174-1183.
506   Stumpf, R.P., Wynne, T.T., Baker, D.B., Fahnenstiel, G.L. 2012. Interannual variability of cyanobacterial
507   blooms in Lake Erie. Plos One, 7(8), e42444.
508   R Core Team. 2021. R: A language and environment for statistical computing. In https://www.R-
509   project.org/
510   Verhamme, E.M., Redder, T.M., Schlea, D.A., Grush, J., Bratton, J.F., DePinto, J.V. 2016. Development
511   of the Western Lake Erie Ecosystem Model (WLEEM): Application to connect phosphorus loads
512   to cyanobacteria biomass. Journal of Great Lakes Research, 42(6), 1193-1205.
513   Wang, J., Bai, X., Hu, H., Clites, A., Colton, M., Lofgren, B. 2012. Temporal and Spatial Variability of
514   Great Lakes Ice Cover, 1973–2010. Journal of Climate, 25(4), 1318-1329.
515   Wynne, T.T., Mishra, S., Meredith, A., Litaker, R.W., Stumpf, R.P. 2021. Intercalibration of MERIS,
516   MODIS, and OLCI Satellite Imagers for Construction of Past, Present, and Future Cyanobacterial
517   Biomass Time Series. Remote Sensing, 13(12).
518   Wynne, T.T., Stumpf, R.P. 2015. Spatial and temporal patterns in the seasonal distribution of toxic
519   cyanobacteria in Western Lake Erie from 2002-2014. Toxins, 7(5), 1649-1663.
520   Wynne, T.T., Stumpf, R.P., Pokrzywinski, K.L., Litaker, R.W., De Stasio, B.T., Hood, R.R. 2022.
521   Cyanobacterial Bloom Phenology in Green Bay Using MERIS Satellite Data and Comparisons
522   with Western Lake Erie and Saginaw Bay. Water, 14(17).
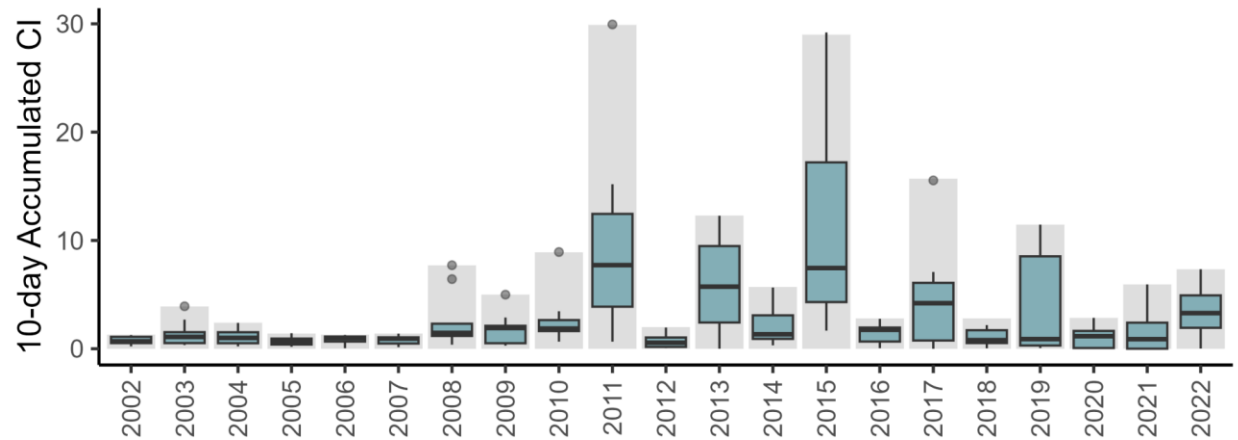
Figures:

Figure 1. Boxplots of 10-day accumulated cyanobacterial index (CI) for the bloom season (July-October) along with a bar plot (gray bars) of maximum CI (CI max). For boxplots, the boxes represent the interquartile range (25th and 75th percentile), while the thick, horizontal line shows the median weekly accumulated CI. The vertical lines represent the minimum (25th percentile-1.5×Interquartile range) and maximum (75th percentile+1.5×Interquartile range). Points represent anomalous 10-day accumulated CI for that year (>1.5×Interquartile).
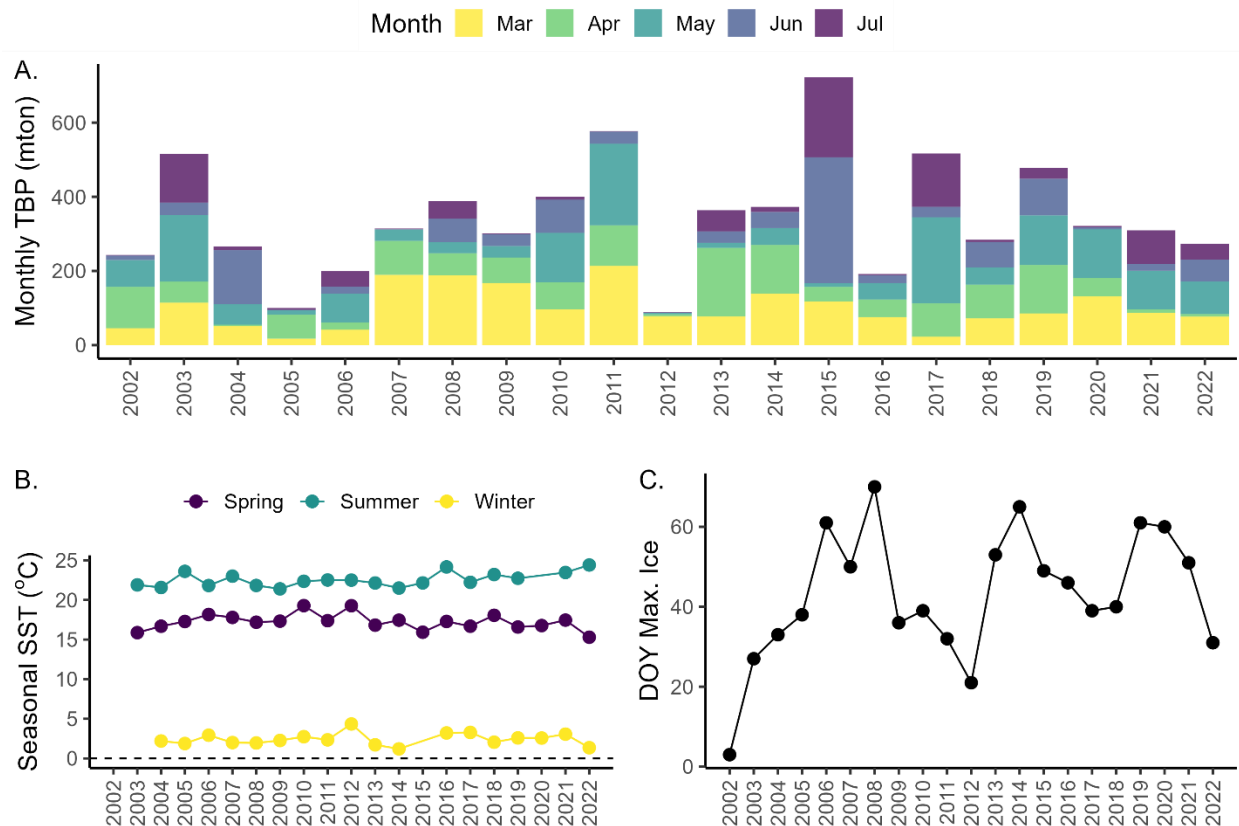
Figure 2. Selected environmental parameters plotted from 2002-2022 and used as potential predictors for annual cyanobacterial bloom intensity including: A) Monthly total bioavailable phosphorus (TBP) loading from March-July (mton); B) western Lake Erie seasonal sea surface temperature (SST, °C) as estimated using NASA Giovanni data for winter, spring, and summer; and C) Day of year (DOY) of maximum ice extent for the entirety of Lake Erie.
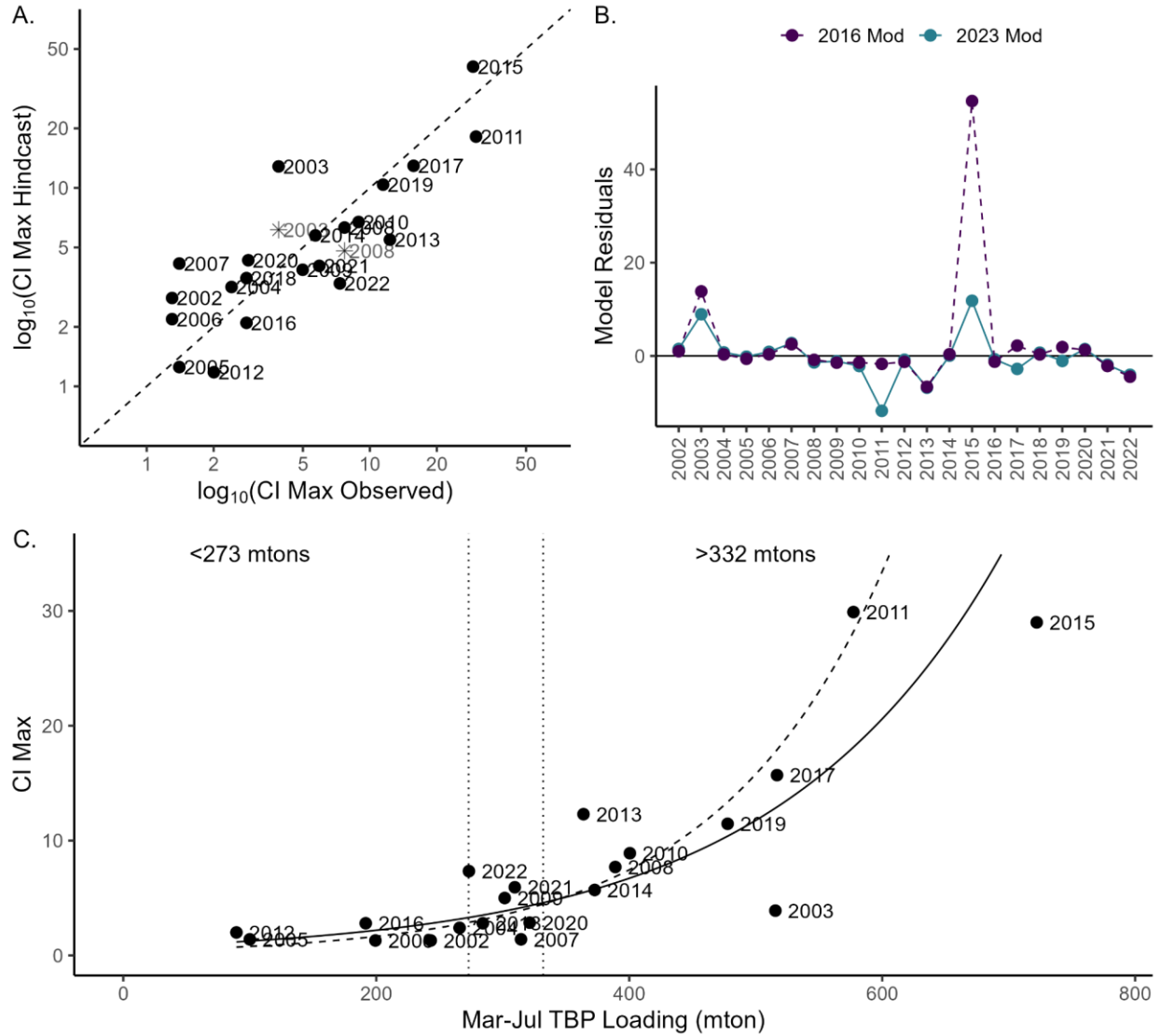
Figure 3. Regression results for the updated total bioavailable phosphorus (TBP) model. A) Comparison of observed and calculated cyanobacterial index (CI) max. Gray asterisks indicate cool June years (2003, 2008) when March-June loading was used, following Stumpf et al. (2016). The dashed line represents the 1:1 line. Results are plotted as $\log_{10}$ values for visualization. B) Model residuals (modeled-observed) plotted for the Stumpf et al. (2016) model and the updated TBP model (2023) from this study. C) Logarithmic relationship between annual CI max and spring (March-July) TBP loading. The dashed line is the relationship observed in Stumpf et al. (2016) while the solid line is the relationship identified in this study (2023). The vertical dashed lines indicate 273 mtons TBP loading and 332 mtons TBP loading, respectively, as it corresponds to the logistic regression below.
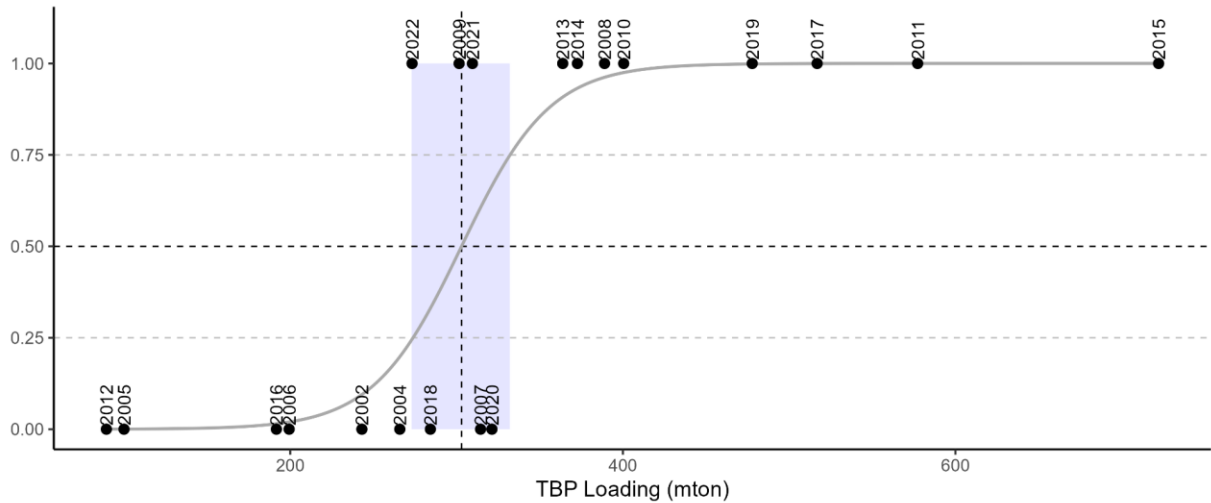
14

548
549 **Figure 4.** Results from the logistic regression used to identify low (0) and high (1) bloom years based on
550 total bioavailable phosphorus (TBP) loading as measured from March-July. The horizontal dashed lines
551 indicate 25% (273 mton TBP loading), 50% (303 mton TBP loading), and 75% (332 mton TBP loading)
552 certainty, respectively. The vertical dashed line corresponds to the TBP loading at 50% certainty. The
553 blue box represents TBP loading from 25-75% certainty.
554
555 **Tables:**
556
557 **Table 1.** Linear regression (TBP model) and multiple linear regression results for the top three identified
558 models (Model 1-3). Multiple linear regression models were included if they were within 2 Akaike
559 Information Criteria (AIC) of the best-fit model. Columns include the model intercept, spring (March-
560 July) total bioavailable phosphorus (TBP) loading, day of year (DOY) of maximum ice extent, and winter
561 (November-February) TBP loading. The AIC adjusted r^2 (Adj. r^2), and root mean square error (RMSE)
562 are reported. The mean and standard error (mean $\pm$ standard error) are reported for each variable.
563

	Intercept	Spring TBP	DOY	Winter TBP	AIC	Adj. r^2	RMSE
TBP Model	-0.146 $\pm$ 0.117	$2.4 \times 10^3 \pm$ 3.2×10^4				0.76	4.66
Model 1	-0.095 $\pm$ 0.134	$2.3 \times 10^3 \pm$ 3.6×10^4			-41.47	0.66	4.31
Model 2	-0.153 $\pm$ 0.187	$2.2 \times 10^3 \pm$ 3.7×10^4	$1.6 \times 10^3 \pm$ 2.5×10^3		-41.81	0.65	4.37
Model 3	-0.0720 $\pm$ 0.208	$2.4 \times 10^3 \pm$ 4.0×10^4		-4.9×10^5 $\pm 3.3 \times 10^4$	-42.64	0.65	4.32

564