Electronic Supplementary Material

Phytoplankton blooms in Lake Erie impacted by both long-term and springtime phosphorus loading

Jeff C. Ho[a, b, *] and Anna M. Michalak[b]

[a]Department of Civil & Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, California, USA 94305

[b]Department of Global Ecology, Carnegie Institution for Science, 260 Panama St., Stanford, California, USA 94305

[*]Corresponding author (jeffho@stanford.edu)

Appendix S1.

Additional ancillary data

To interpret the historical record (Fig. 1), bloom extent in certain years is compared to concurrent *in situ* chlorophyll-a (chl-a) measurements and local wind stress and wind direction information. Chl-a is a proxy for phytoplankton abundance, and strong winds influence bloom extent through mixing (Wynne et al., 2010). We use surface measurements of chl-a from the U.S. EPA Great Lakes National Program Office (EPA GLNPO, 2012); wind direction from National Data Buoy Center Station 45005 (NOAA NDBC, 2015); and buoy measurements of wind speed (NOAA NDBC, 2015), from which wind stress τ [Pa] is derived:

$$\tau = \rho w^2[0.001(0.69 + 0.081w)] \tag{S1}$$

where ρ is the air density, estimated as 1.25 $\mathrm{kg/m^3}$, w is the mean hourly wind speed [m/s], and the term in square brackets [unitless] is an empirical drag coefficient estimated by Hsu (1974), following the approach of Wynne et al. (2010).

Impact of dreissenids and other potential explanatory factors

In addition to exploring the impact of loading, we also examined whether the impact from dreissenid mussels could aid in explaining the historical bloom record. The invasion by dreissenid mussels was complete by the early-1990s (Allinger and Reavie, 2013), and was well-documented to have changed water clarity and nutrient cycling in Lake Erie (Vanderploeg et al., 2001).

We performed two analyses to test whether the pre- and post-dreissenid distinction explained any additional variability beyond loading. First, we identified variables that explained bloom size during only the pre- (1984-1992) or post- (1993-2015) dreissenid periods. In both periods, a two-variable model including one spring loading variable and one decadal-scale cumulative loading variable explained a substantially greater proportion of variability than either a spring or cumulative variable alone. Second, we explored whether a binary variable stratifying pre- and post-dreissenid years was useful in explaining variation over the whole period. We found that the best model that included this variable (a two-variable model with the binary variable and March-July DRP) explained only 55% of the variability across the whole period, and a standard F-test showed that adding this binary variable to the spring loading term did not improve model fit significantly (p=0.20). We therefore observed no evidence that the presence of dreissenids provided significant additional explanatory power, either beyond or as a surrogate for springtime or long-term loading.

We also found no evidence of systematic differences in how the Landsat algorithm captures bloom size before vs. after the appearance of dreissenids, despite the changes in turbidity associated with dreissenid mussels (Budd et al., 2001) and shifts in the phytoplankton community during this time (Table 1). The consistency of the model in Eqn. 2 between the historical (1984-2000) and current (2001-2015) periods (Table 2) is also indicative of this, as are consistencies between the Landsat-observed decline in bloom area and documented decreases in phytoplankton biomass in the late 1980s (Ho et al., 2017). The inclusion of a “greenness” filter in the algorithm for improving accuracy in areas with higher turbidity (Ho et al., 2017) further mitigates any impact due to changes in water clarity.

Table. S1. Posterior probabilities of individual variables (Yadav et al., 2010), based on Bayesian model averaging (Raftery et al., 1997) of all possible two-variable models. Of the variables in Section 2.3, only those with a posterior probability above 1% are listed here. The posterior probability P is a measure of the importance of a variable in explaining the observed variability in bloom extent (with higher probabilities corresponding to greater importance), when considering all possible sets of two candidate variables. Note that these probabilities should not be confused with *p*-values, which instead denote a level of statistical significance. Shading indicates the relative probabilities, with dark grey for $P>0.15$, light grey for $0.15>P>0.05$, and white for $0.05>P>0.01$. Note that TBP ($\theta=0.63$) and TP do not appear in the table because none of the variables based on these quantities had a probability exceeding 1%.

		DRP	**Discharge**	**TBP (θ=0.138)**
Spring loading	March-July	0.01		
	April-June	0.11		
	April-July	0.15	0.34	0.05
	April-August	0.01	0.04	
	April-September	0.02	0.10	0.01
	May-June	0.01		
	May-July	0.02	0.03	
	All monthly aggregations	**0.36**	**0.54**	**0.09**
Cumulative loading	8-year	0.01		
	9-year	0.27		0.07
	10-year	0.17		0.03
	11-year	0.12		0.01
	12-year	0.05		
	13-year	0.04		
	14-year	0.03		
	15-year	0.02		
	All yearly aggregations	**0.78**	**0.06**	**0.16**

Fig. S1. Comparisons between observed bloom area [km^2] and linear models based only on (A) the April-July DRP loading term and (B) the truncated 9-year cumulative DRP loading term in Eqn. 2. Dashed line represents bloom areas used in sensitivity test.

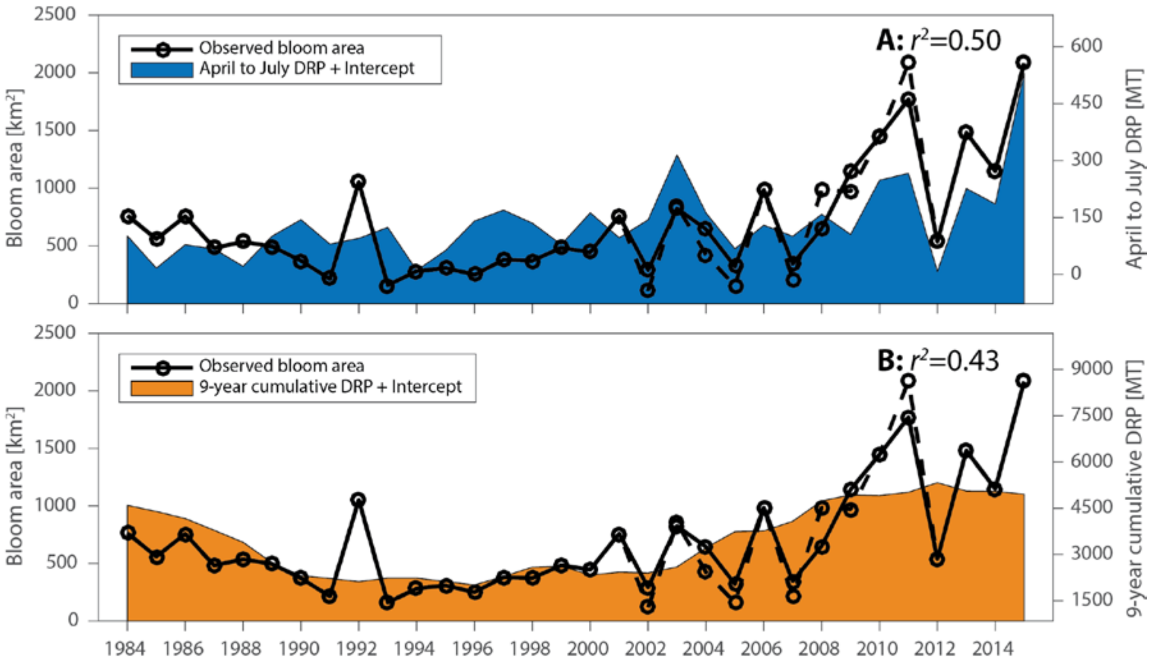


Fig. S2. Contributions to predicted bloom area from April-July discharge, 9-year cumulative DRP loading and intercept term fit to observed bloom areas. Dashed line represents bloom areas used in sensitivity test.

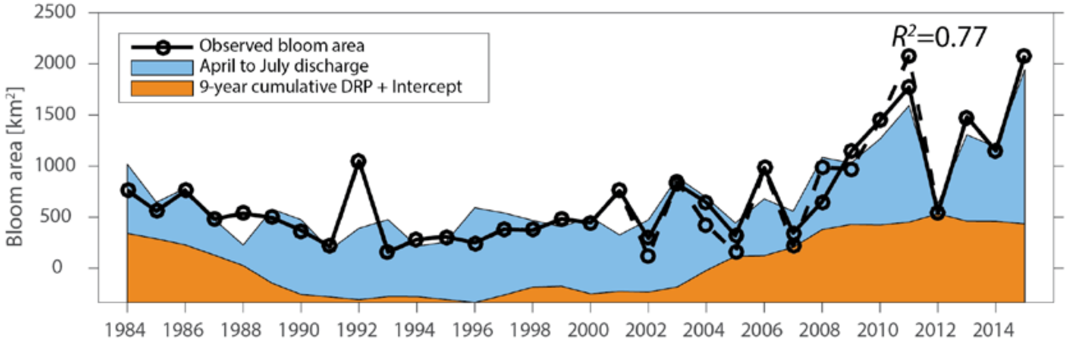
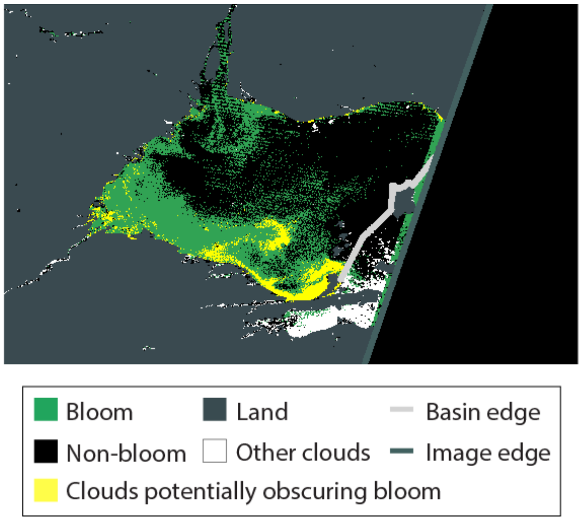


Fig. S3: Maximum bloom extent for 1992 (August 4, 1992).



References

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

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

Budd, J.W., Nalepa, T.F., Fahnenstiel, G.L., Drummer, T.D., 2001. Remote sensing of biotic effects: zebra mussels (Dreissena polymorpha) influence on water clarity in Saginaw Bay, Lake Huron. Limnol. Oceanogr. 46, 213–223.

EPA GLNPO, 2012. Great Lakes Environmental Database (GLENDA) [WWW Document]. URL http://www.epa.gov/greatlakes/monitoring/data_proj/glenda/ (accessed 2.20.10).

Ho, J.C., Stumpf, R.P., Bridgeman, T.B., Michalak, A.M., 2017. Using Landsat to extend the historical record of lacustrine phytoplankton blooms: A Lake Erie case study. Remote Sens. Environ. 191, 273–285.

Hsu, S.A., 1974. Experimental results of the drag-coefficient estimation for air-coast interfaces. Boundary-Layer Meteorol. 6, 505–507.

NOAA NDBC, 2015. Station 45005 [WWW Document]. URL http://www.ndbc.noaa.gov/station_page.php?station=45005 (accessed 3.4.16).

Raftery, A.E., Madigan, D., Hoeting, J., 1997. Bayesian model averaging for linear regression models. J. Am. Stat. Assoc. 92, 179–191. doi:10.1080/01621459.1997.10473615

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

Vanderploeg, H. a., Liebig, J.R., Carmichael, W.W., Agy, M. a., Johengen, T.H., Fahnenstiel, G.L., Nalepa, T.F., 2001. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. Can. J. Fish. Aquat. Sci. 58, 1208–1221. doi:10.1139/cjfas-58-6-1208

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

Yadav, V., Mueller, K.L., Dragoni, D., Michalak, A.M., 2010. A geostatistical synthesis study of factors affecting gross primary productivity in various ecosystems of North America. Biogeosciences 7, 2655–2671. doi:10.5194/bg-7-2655-2010