Appendix A. Supplemental data set description, methods, and coefficient estimates for phytoplankton presence / absence and biovolume models.

Data sets

We used the United States Environmental Protection Agency's 2007 National Lake Assessment (NLA) to create response trait models of seven distinct phytoplankton divisions (Cyanobacteria, Diatoms, Cryptomonads, Chlorophyta, Euglenoids, Chrysophyta, and Dinoflagellates). The NLA is a robust sampling effort designed to provide an estimate of the condition of lakes in the contiguous United States, and in the 2007 sampling year, 1157 lakes were sampled. A vast majority of lakes were sampled only once (>90%), and for the lakes that were sampled more than once, we chose to keep only the first sampling time point. Sampling protocols were standard across all lakes and a summary of the data and protocols can be found on their website (http://water.epa.gov/type/lakes/lakessurvey_index.cfm). For the NLA data set, we used nutrient and dissolved organic carbon (DOC) concentration data obtained from the 'Lake Water Quality Data' spreadsheet, phytoplankton divisions and biovolume from the 'Lake Phytoplankton Soft Algae Count Data' spreadsheet, water temperature profiles from the 'Lake Profile Data' spreadsheet, maximum depth from the 'Information for Lakes that were Sampled' spreadsheet, and secchi depth from the 'Lake Secchi Disk Data' spreadsheet.

For the validation data set, we used data from automated sensors that were collected and compiled from 20 globally distributed lakes, although a vast majority of lakes (18) were in the Northern Hemisphere. Only one validation lake, Kentucky Lake, was a part of

the NLA set of lakes, however, it was sampled in separate years for the two data sets. All validation lakes' high- and low-frequency data and results used in this analysis were from Solomon et al. (2013).

Estimation of Average Light Climate

Estimation of average light climate requires a light extinction coefficient (K_d), incoming light intensities (I_0), and knowledge of the depth of the upper mixed layer of the lake (z_{mix}). The light concentration (I_z) at a given depth (z) was estimated with equation A.1, where I_0 was the average incident PAR for a given day. I_z was integrated over the z_{mix} and then divided by z_{mix} to obtain an average light climate experienced by phytoplankton in units of mmol PAR m^{-2} s^{-1}.

$$I_z = I_0 * e^{-Kd * z} \qquad (A.1)$$

z_{mix} was calculated using the temperature profile data from the NLA dataset and the R package rLakeAnalyzer (Winslow et al. 2014). Since light was not measured in the NLA data set, we set I_0 equal to the average incident PAR from all validation lakes. Because the light extinction coefficient (K_d) was not available for the NLA lakes, we estimated K_d using secchi depth, another proxy for lake light environment. We estimated K_d through basic rearrangement of equation A.1 and knowledge of the 1% light level depth which we calculated from secchi depth (1% light = 2 * secchi depth). If secchi depth was marked as 'clear to bottom', secchi depth was set to the maximum observed depth.

For the validation lakes, if K_d or secchi depth were not available, we estimated K_d with the following model (Morris et al. 1995):

$$K_{d\,PAR} = 0.22 * DOC + 0.07 * chl\text{-}a - 0.05 \qquad (A.2)$$

where DOC and chl-a (chlorophyll-a) are in units of g m^{-3}. Average light climate for each validation lake metabolism day was estimated using equation A.1. Daily average light climate values were averaged over the GPP time series to obtain a single average light climate for each lake.

Phytoplankton Biovolume to Carbon Conversion

The phytoplankton response trait model was used to estimate division-specific biovolume for each validation lake using the validation lakes' environmental data. Division biovolumes were converted to carbon using two separate biovolume to carbon relationships (Menden-Deuer and Lessard 2000), one for diatoms (A.3) and one for all other phytoplankton groups (A.4):

$$\log_{10} C = -0.541 + 0.811 * \log_{10} V \qquad (A.3)$$

$$\log_{10} C = -0.665 + 0.939 * \log_{10} V \qquad (A.4)$$

where C is phytoplankton cellular carbon content in pg C $cell^{-1}$ and V is phytoplankton cellular volume in μm^3. We convert from division-specific mean cell size to mean cellular carbon content using equations A.3 and A.4, and convert from division-specific biovolume to division-specific carbon by multiplying biovolume by mean carbon content per mean cell volume. Division-specific mean cell size, mean cellular carbon content, and mean carbon content per mean cell volume are reported in Table A.5.

Phytoplankton Effect Trait

Laboratory-measured α's, estimated from growth-irradiance curves, were obtained for 67 phytoplankton species representing the seven phytoplankton divisions (Schwaderer et al. 2011; Edwards et al. 2013*b*). Three phytoplankton divisions (Chrysophyta, Dinoflagellates, and Euglenoids) were represented by only one species each, and subsequently this single α was used as the respective

phytoplankton division's α. For the remaining divisions, a gamma distribution of α was estimated by maximum likelihood to estimate the mean α of each division (Fig. A2).

TABLE A1. Summary of lake parameters for the validation lakes, where Long and Lat indicate the longitude and latitude in decimal degrees; lake area is the surface area of the lake; TP, TN, Chl-a, DOC, and color are epilimnetic summer-time means of total phosphorus, total nitrogen, chlorophyll-a, dissolved organic carbon, and water color absorbance measured at 440nm, respectively; K_d is the light attenuation coefficient calculated from light profiles; water temp is the average summer-time water temperature measured at the depth of the dissolved oxygen sensor; and light climate is the summer-time average light climate experienced by the phytoplankton as described in the supplementary methods.

Lake	Long	Lat	Lake Area (ha)	TP (μg L^{-1})	TN (μg L^{-1})	Chl-a (μg L^{-1})	DOC (mg L^{-1})	Color (m^{-1})	Kd (m^{-1})	Water Temp (°C)	Light Climate (mmol PAR m^{-2} s^{-1})
Acton	-84.744	39.575	253.0	114.0	5836.0	55.7	3.6	1.5	1.6	27.4	0.203
Annie	-81.351	27.207	36.5	4.3	236.0	2.3	7.7	1.1	1.8	29.4	0.108
Balaton	17.245	46.717	3800.0	72.0	1664.0	17.7	7.7	1.4	-	23.5	0.332
Crampton	-89.470	46.200	25.7	8.9	322.1	2.6	3.8	0.6	-	22.5	0.160
Crystal Bog	-89.606	46.008	0.5	27.0	684.0	19.2	11.5	5.1	2.3	23.3	0.244
Feeagh	-9.575	53.948	400.0	7.3	130.0	1.8	7.8	4.0	1.8	16.6	0.016
Fredriksburg Slotsø	12.000	56.000	22.3	102.1	1735.0	64.5	5.4	2.2	-	21.6	0.027
Hampensø	9.333	56.000	76.0	22.7	580.0	5.3	3.1	0.5	0.7	18.6	0.117
Kentucky	-88.109	36.739	97000.0	47.0	792.0	16.9	3.0	-	-	28.3	0.025
Mendota	-89.652	43.099	3937.7	85.0	957.0	3.8	5.2	0.8	1.4	24.3	0.106

Muggelsee	13.650	52.440	746.1	105.0	910.0	33.5	8.0	-	4.1	20.9	0.034
Rotoiti	176.428	-38.039	3460.0	30.3	295.9	4.2	1.4	0.2	-	21.0	0.157
Rotorua	176.266	-38.066	7980.0	32.7	476.5	14.3	2.3	0.2	0.6	21.1	0.058
St Gribsø	12.300	55.983	10.0	69.0	699.0	30.3	12.8	6.5	2.5	21.4	0.092
Sunapee	-72.033	43.400	1667.0	5.3	171.1	1.9	2.4	-	-	21.9	0.206
Taihu	120.317	31.300	233800.0	186.0	3600.0	46.0	5.6	1.1	4.4	28.6	0.079
Trout	-89.665	46.029	1608.0	13.0	203.0	1.7	2.8	1.0	0.3	20.8	0.166
Trout Bog	-89.686	46.041	1.1	29.0	634.0	15.0	17.3	13.1	2.9	22.1	0.215
Vedstedsø	9.333	55.167	9.0	19.5	547.6	41.2	4.8	1.1	0.8	20.3	0.227
Yuan Yang	121.400	24.583	3.6	6.4	1882.3	11.6	8.4	5.3	1.3	19.5	0.102

TABLE A2. Coefficients used in the best models of phytoplankton division presence / absence across the NLA lakes. The best model for each division was the model with the lowest AIC.

	Intercept	Temp	ln DOC	ln TN	ln TP	log10 Light Climate
Chlorophyta	-0.44	-	-	1.21	-0.95	-
Chrysophyta	4.97	-0.06	0.39	-0.32	-0.28	-
Cryptomonads	4.62	-0.05	-	-	-0.28	-
Cyanobacteria	-2.19	0.18	-	0.25	-	-
Diatoms	4.95	-	-	-0.30	-	-
Dinoflagellates	0.26	0.12	-	-0.27	-0.33	-0.46
Euglenoids	-4.04	0.14	-	-	0.24	-

TABLE A3. Coefficients used in the best models of phytoplankton division biovolume and total phytoplankton biovolume across the NLA lakes. The best model for each division and total phytoplankton biovolume was the model with the lowest AIC.

	Intercept	Temp	ln DOC	ln TN	ln TP	log10 Light Climate
Chlorophyta	5.98	0.10	0.43	0.47	-	-
Chrysophyta	6.26	-	0.29	0.31	-	-
Cryptomonads	5.72	-0.02	0.41	0.28	-	-0.53
Cyanobacteria	5.51	0.08	-	0.57	0.21	-0.30
Diatoms	7.64	-0.03	-	0.37	0.22	-0.57
Dinoflagellates	7.94	0.10	-	0.51	-	-
Euglenoids	6.90	0.05	0.51	-	0.46	-0.43
Total Biovolume	9.27	0.07	-	0.50	0.08	-

TABLE A4. Shape and scale parameters used to generate gamma distributions of light use efficiency for the divisions with more than one laboratory-measured light use efficiency value available. Divisions with only one light use efficiency value included Chrysophyta (0.063), dinoflagellates (0.004), and Euglenoids (0.203).

	Shape	Scale
Chlorophyta	0.9773	0.0355
Cryptomonads	3.9390	0.0027
Cyanobacteria	2.8004	0.0171
Diatoms	1.7065	0.0190

TABLE A5. Division-specific mean cell size (V), mean cellular carbon content (C), and mean carbon content per mean cell volume (C / V) used to convert from phytoplankton biovolume to carbon. Mean values are based on NLA lakes after supplementing with phytoplankton data from Kremer et al. 2014.

	V (μm^3 $cell^{-1}$)	C (pg C $cell^{-1}$)	C / V (pg C μm^{-3})
Chlorophyta	951.4	135.1	0.142
Chrysophyta	198.0	31.1	0.157
Cryptomonads	109.6	17.8	0.162
Cyanobacteria	45.2	7.7	0.171
Diatoms	785.3	64.4	0.082
Dinoflagellates	28169.9	3267.7	0.116
Euglenoids	4626.2	596.8	0.129
Total Biovolume	200.3	31.4	0.157

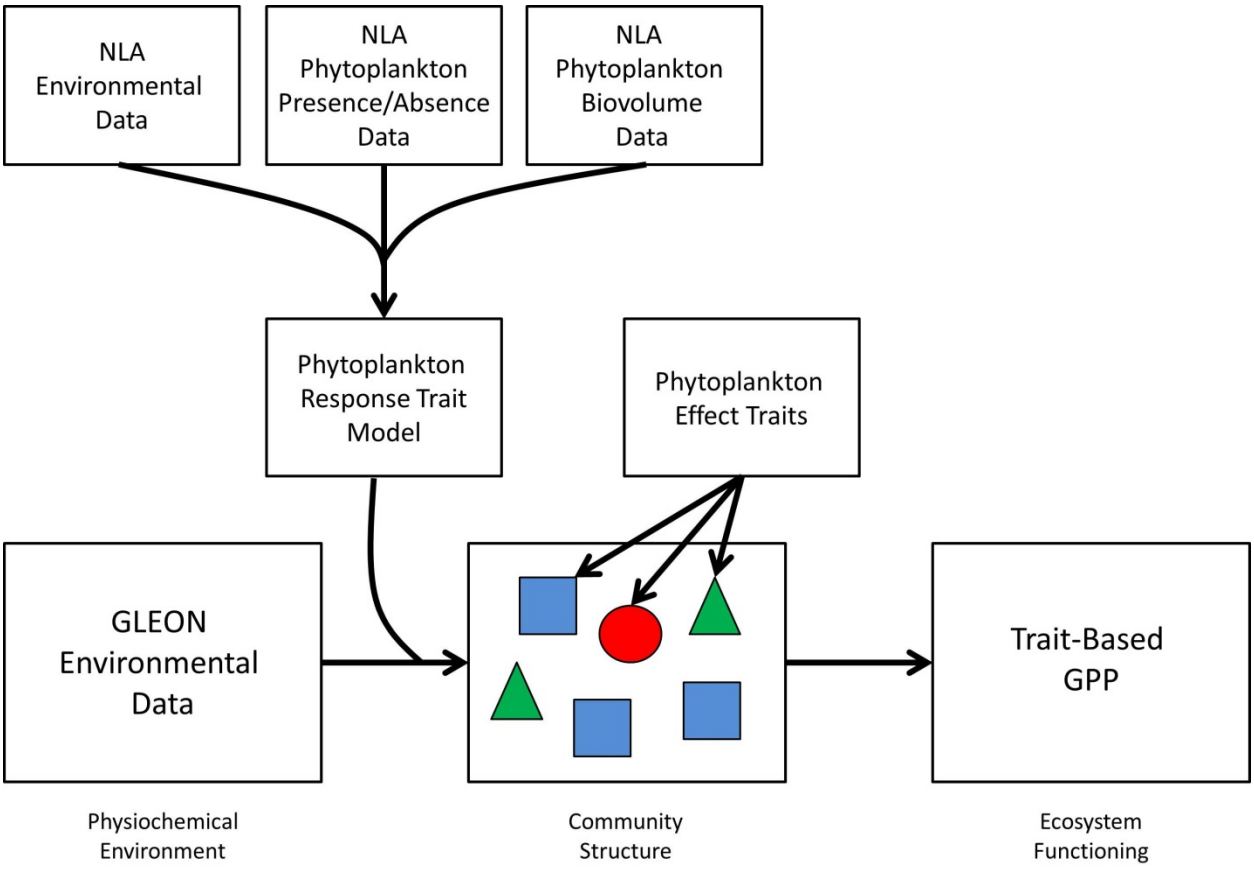

FIG. A1. The NLA Environmental Data was used to explain the NLA Phytoplankton Presence / Absence and Biovolume, ultimately creating a two-stage conditional Phytoplankton Response Trait Model. The response model is applied to the GLEON Environmental Data to estimate phytoplankton community structure, indicated by the different shapes, for each GLEON lake. Phytoplankton Effect Trait of light use efficiency is applied to the modeled phytoplankton communities, indicated by the color of the shapes, resulting in a Trait-Based estimate of GPP. This conceptual figure is redrawn from Suding et al. 2008.

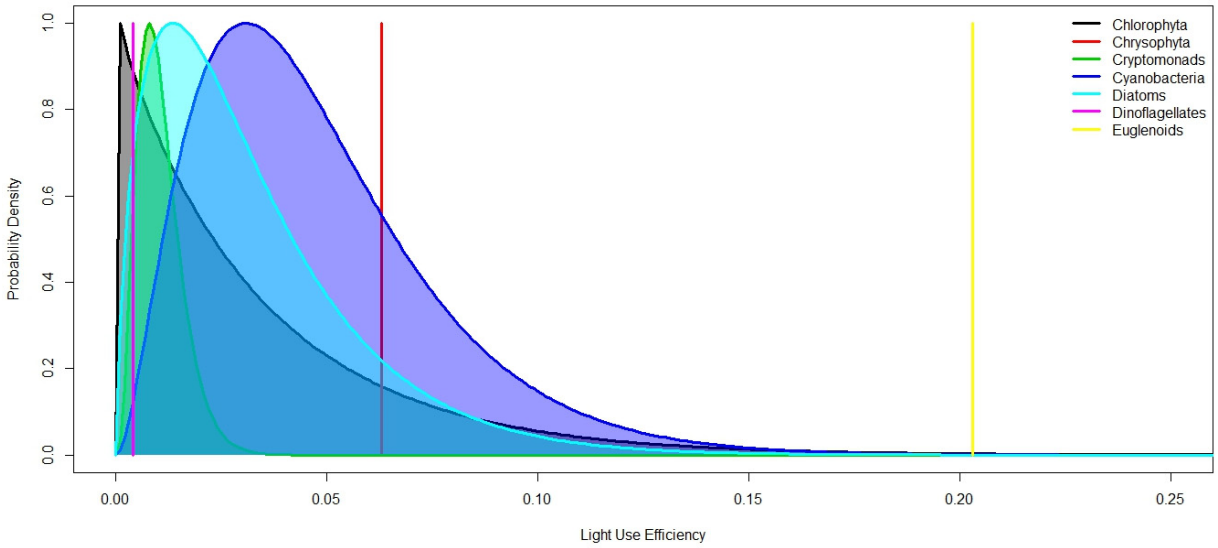

FIG. A2. For each phytoplankton division, we fit a gamma distribution to the available estimates of light use efficiency for each of the seven phytoplankton divisions. Light use efficiency values were obtained from Schwarderer et al. 2011 and Edwards et al. 2013b (n = 67).

LITERATURE CITED

Edwards, K. F., E. Litchman, and C. A. Klausmeier. 2013*a*. Functional traits explain phytoplantkon community structure and seasonal dynamics in a marine ecosystem. Ecology Letters 16:56–63.

Edwards, K. F., E. Litchman, and C. A. Klasumeier. 2013*b*. Functional traits explain phytoplankton responses to environmental gradients across lakes of the United States. Ecology 94:1626–1635.

Kremer, C. T., J. P. Gillette, L. G. Rudstam, P. Brettum, and R. Ptacnik. 2014. A compendium of cell and natural unit biovolumes for >1200 freshwater phytoplankton species. Ecology 95:2984.

Menden-Deur, S., and E. J. Lessard. 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. Limnology and Oceanography 45:569–579.

Morris, D. P., H. Zagarese, C. E. Williamson, E. G. Balseiro, B. R. Hargreaves, B. Modenutti, et al. 1995. The attenuation of solar UV radiation in lakes and the role of dissolved organic carbon. Limnology and Oceanography 40:1381–1391.

Schwaderer, A. S., K. Yoshiyama, D. T. Pinto, N. G. Swenson, C. A. Klausmeier, and E. Litchman. 2011. Eco-evolutionary differences in light utilization traits and distributions of freshwater phytoplankton. Limnology and Oceanography 56:589–598.

Solomon, C. T., D. A. Bruesewitz, D. C. Richardson, K. C. Rose, M. C. Van de Bogert, P. C. Hanson, et al. 2013. Ecosystem respiration: drivers of daily variability and background respiration in lakes around the globe. Limnology and Oceanography 58:849–866.

Suding, K. N., S. Lavorel, F. S. Chapin III, J. H. C. Cornelissen, S. Díaz, E. Garnier, et al. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. Global Change Biology 14:1125–1140.

Winslow, L., J. Read, R. Woolway, J. Bentrup, T. Leach, and J. Zwart. 2014. rLakeAnalyzer: package for the analysis of lake physics. R package version 1.4. http://CRAN.R-project.org/package=rLakeAnalyzer