

Phytoplankton traits predict ecosystem function in a global set of lakes

JACOB A. ZWART,^{1,3} CHRISTOPHER T. SOLOMON,² AND STUART E. JONES¹

¹*Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556 USA*

²*Department of Natural Resource Sciences, McGill University, Ste. Anne de Bellevue, Quebec H9X 3V9 Canada*

Abstract. Predicting ecosystem function from environmental conditions is a central goal of ecosystem ecology. However, many traditional ecosystem models are tailored for specific regions or ecosystem types, requiring several regional models to predict the same function. Alternatively, trait-based approaches have been effectively used to predict community structure in both terrestrial and aquatic environments and ecosystem function in a limited number of terrestrial examples. Here, we test the efficacy of a trait-based model in predicting gross primary production (GPP) in lake ecosystems. We incorporated data from >1000 United States lakes along with laboratory-generated phytoplankton trait data to build a trait-based model of GPP and then validated the model with GPP observations from a separate set of globally distributed lakes. The trait-based model performed as well as or outperformed two ecosystem models both spatially and temporally, demonstrating the efficacy of trait-based models for predicting ecosystem function over a range of environmental conditions.

Key words: ecosystem function; GLEON; gross primary productivity; lake primary production; phytoplankton traits; trait-based ecology; U.S. EPA National Lake Assessment.

INTRODUCTION

The ecosystem concept has enabled tremendous gains in understanding of biogeochemical fluxes at local, regional, and even global scales. However, this concept is often applied in a static manner, relying on empirical, correlative relationships that fail to consider the ecological interactions that underpin ecosystem processes (Suding et al. 2008, Loreau 2010). Unfortunately, the inflexibility of traditional ecosystem approaches limits their utility when attempting to predict ecosystem function for systems outside of their training data set, such as in dissimilar regions or under environmental change scenarios. The resurgence in trait-based ecology has provided a useful alternative for predicting emergent ecosystem processes across a wide range of ecosystems and under future climate and land-use scenarios by considering additional ecological detail within ecosystems (Díaz et al. 2007, Suding et al. 2008).

Trait-based approaches are now frequently applied to predict species or trait composition of communities (McGill et al. 2006) and species' spatiotemporal distribution through mechanistic niche modeling (Kearney and Porter 2009). These studies exploit knowledge of response traits (traits that predict species' fitness as a function of environmental conditions) to explain species' niches and consequently, community assembly dynamics. For example, habitat filtering mediated by response

traits explained divergent strategies of tree species in an Amazonian forest (Kraft et al. 2008). In aquatic environments, Edwards et al. (2013a) used phytoplankton response traits, quantified in laboratory cultures, to predict the abundance of resource utilization traits in response to a fluctuating oceanic environment in the English Channel, as well as predicting freshwater phytoplankton abundances across a gradient of light and phosphorus availability in United States lakes (Edwards et al. 2013b).

In addition to response traits, effect traits (those that govern species' contributions to ecosystem function) can be used to scale from environmental variables to community dynamics and ecosystem processes within a single trait-based framework (Suding et al. 2008). Therefore, trait-based approaches provide ecosystem scientists with a tool to incorporate additional ecological knowledge (species' traits) when attempting to predict changes in ecosystem function. A handful of terrestrial studies have led the way in the use of traits to predict ecosystem function (e.g., Lavorel and Garnier 2002, Díaz et al. 2007, de Bello et al. 2010), but we are unaware of any aquatic studies that have achieved this goal.

Here, we use models of freshwater phytoplankton response and effect traits to predict aquatic ecosystem primary productivity for a set of globally distributed lakes. We leverage division-specific phytoplankton biovolume data from >1000 United States lakes to generate niche models, and combine these models with literature-derived estimates of division-specific light use efficiency to generate lake-specific estimates of gross

Manuscript received 4 November 2014; revised 20 February 2015; accepted 23 February 2015. Corresponding Editor: C. E. Cáceres.

³ Email: jzwart@nd.edu

primary production (GPP). We evaluate the performance of our trait-based model of GPP relative to two ecosystem models of GPP using a set of globally distributed lakes for which GPP observations are available. Additionally, we examine the temporal performance of both the trait-based and ecosystem models using one lake for which temporally resolved nutrient data were available. We hypothesize that our trait-based model, by capturing key ecological mechanisms, will provide accurate estimates of lake GPP and will outperform widely used traditional ecosystem models.

METHODS

Trait-based model of lake GPP

To calibrate our trait-based model of GPP, we used the 2007 U.S. Environmental Protection Agency National Lake Assessment data set (NLA), a spatially robust sampling effort designed to provide an estimate of the condition of lakes in the contiguous United States (USEPA 2009). We created niche models for seven phytoplankton divisions (cyanobacteria, diatoms, cryptomonads, chlorophyta, euglenoids, chrysophyta, and dinoflagellates) using lake physiochemical and phytoplankton data gathered from the NLA data set. The coefficients of the lake physiochemical parameters in the niche models represent the response traits of each phytoplankton division. Because our ultimate goal was to evaluate the performance of a trait-based model of GPP against sensor-network-derived observations from the validation data set, we selected only lake physiochemical parameters from the NLA data set that were also readily available in the validation data set. These included total phosphorus (TP), total nitrogen (TN), dissolved organic carbon (DOC), surface temperature, and average light climate experienced by phytoplankton in the upper mixed lake layer.

Biovolume data was not available for every phytoplankton observation within the NLA data set, thus, we supplemented missing biovolume data where possible by multiplying observed phytoplankton abundance by mean genera cell size from Kremer et al. (2014). Mean genera cell size data that were available for both the NLA data set and Kremer et al. (2014) corresponded well ($R^2 = 0.51$, linear regression $P < 0.001$, $n = 108$). After supplementing with cell size data from Kremer et al. (2014), biovolume attributed to our seven divisions accounted for on average 92% of the phytoplankton observed across all NLA lakes.

The absence of phytoplankton divisions from several NLA lakes created biovolume data containing many zeros. As a result, we modeled phytoplankton biovolume conditional on presence, as previously done with similar phytoplankton biovolume data sets (two-stage conditional approach; Cunningham and Lindenmayer 2005, Edwards et al. 2013a, b). Phytoplankton presence/absence was modeled as a binary response to our selected physiochemical parameters and log phyto-

plankton biovolume of the nonzero biovolumes was modeled using the same candidate physiochemical parameters. The binary presence/absence model was implemented using logistic regression (binomial error distribution with a logit link function), and linear models with normally distributed errors were used for response trait models with log-transformed biovolume of each phytoplankton division as the dependent variable. Akaike's information criterion (AIC) was used to identify the top model for each phytoplankton division's presence/absence and biovolume. In a number of cases, multiple models were within two Δ AIC units of the best model, but we chose to use only the model with the lowest AIC. Total N, TP, and DOC were transformed using natural log, and light climate was transformed using \log_{10} , while water temperature was not transformed. We used these niche models and seasonal averages of observed physiochemical characteristics to estimate phytoplankton community composition in the validation lakes. More information on the data sets and trait-based model can be found in the Appendix.

We linked phytoplankton community composition to GPP via the phytoplankton trait of light use efficiency (α , GPP per unit incident light). We assume α can be expressed as a community-aggregated trait, meaning that the community trait is equal to the summation of each individual division's trait weighted by its contribution to community biomass, or in our case biovolume (biomass ratio hypothesis; Violle et al. 2007). Laboratory-measured α 's, estimated from growth-irradiance curves, were obtained for 67 phytoplankton species representing the seven phytoplankton divisions described by our niche models (Schwaderer et al. 2011, Edwards et al. 2013b). When more than one estimate of α was available, a mean α was calculated from the shape parameters of a gamma distribution fit to all available α for that division (Appendix: Table A4 and Fig. A2). To convert laboratory-measured α 's to comparable units used in the validation data set, we combined division-specific phytoplankton biovolume-to-carbon conversions, division-specific laboratory-derived α 's, and a respiratory quotient of one to generate a division-specific effect trait α (with units of $\text{mg O}_2 \cdot (\text{L lake water})^{-1} \cdot \text{d}^{-1} \cdot (\mu\text{m phytoplankton})^{-3} \cdot (\text{mmol photosynthetically active radiation [PAR]})^{-1} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$).

We generate division-specific estimates of GPP for each validation lake by multiplying niche model predictions of biovolumes for each division, division-specific effect trait α , and each validation lake's mean daily incident PAR; assuming the biomass ratio hypothesis (Violle et al. 2007), we summed all division-specific GPPs to produce an average daily estimate of ecosystem GPP.

Ecosystem models of lake GPP

We compare our trait-based model to a traditional ecosystem model of lake GPP that uses TP as the sole

predictor variable (Hanson et al. 2003, 2004). This model was calibrated using summertime TP and GPP data collected from 25 lakes located in the Northern Highlands Lake District of Wisconsin, USA and the Upper Peninsula of Michigan, USA. It focuses on a phenomenological description of the relationship between nutrient concentration and primary production, rather than a mechanistic description of response and effect traits. Total phosphorus is often used to predict lake primary production because it represents a relatively conservative indicator of the availability of the often limiting nutrient for phytoplankton productivity (Schindler 1977), and it builds upon the relationship between TP and chlorophyll *a* (Dillon and Rigler 1974), an indicator of both phytoplankton biomass (Vörös and Padisák 1991) and primary productivity (Cole et al. 2000). Furthermore, the GPP–TP relationship may be stronger than the GPP–chlorophyll *a* relationship due to differences in phytoplankton community composition or strong zooplankton grazing on phytoplankton (Smith 1979).

As the calibration data set for the Hanson et al. (2004) ecosystem model is geographically constrained, we also compare our trait-based model to an ecosystem model of GPP produced using the NLA data set. Our ecosystem model structure is similar to our trait model; however, we do not distinguish between phytoplankton divisions for both the response and effect traits. As such, we model total phytoplankton biovolume using the same physiochemical parameters in the division-specific niche models (for coefficients of parameters, see Appendix: Table A3), and use these response coefficients along with average observed physiochemical parameters to estimate total phytoplankton biovolume for the validation lakes. We generate a division-ambiguous phytoplankton effect trait α by combining the mean laboratory-measured α for all divisions (Schwaderer et al. 2011, Edwards et al. 2013b), a single phytoplankton biovolume-to-carbon conversion for protist plankton (Menden-Deuer and Lessard 2000), and a respiratory quotient of one (with units of $\text{mg O}_2 \cdot (\text{L lake water})^{-1} \cdot \text{d}^{-1} \cdot (\mu\text{m phytoplankton})^{-3} \cdot (\text{mmol PAR})^{-1} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$). We generate mean daily estimates of ecosystem GPP for each validation lake by multiplying predicted total phytoplankton biovolume, the division-ambiguous phytoplankton effect trait α , and each validation lake's mean daily incident PAR.

We use the TP ecosystem model described in Hanson et al. (2004; referred to as the Hanson-ecosystem model) and our own ecosystem model (referred to as the NLA-ecosystem model) as benchmarks to evaluate the performance gain afforded by a more ecologically detailed trait-based model when predicting ecosystem function.

Model validation

We utilized a set of 20 lakes from the Global Lake Ecological Observatory Network (GLEON), a grass-

roots organization of researchers and autonomous sensor platforms designed to provide high-resolution sensor data for a global set of lakes. For each of the validation lakes, low-frequency lake parameters (e.g., TP, TN, DOC) were sampled at the same location as the high-frequency measurements. Sampling date ranges were from 1 June to 31 August for lakes in the northern hemisphere and from 1 December to 2 March for lakes in the southern hemisphere, and we averaged all sampling time points of the low-frequency lake parameters to generate seasonal means for use in the trait-based and ecosystem models. Additionally, observed phytoplankton biovolume was available for one lake in our data set (Mendota, Wisconsin, USA; North Temperate Lakes LTER 2011), and we use this data to evaluate the performance of our niche models. A summary of the validation lakes' environmental parameters is in Appendix: Table A1.

Observed rates of GPP for the validation lakes were calculated using a free-water metabolism model of dissolved oxygen (DO) measurements in the deepest part of the lakes to capture the pelagic primary production signal, or phytoplankton contribution to whole-lake primary productivity (Van de Bogert et al. 2007). Diel DO cycles were inverse modeled by Solomon et al. (2013) to estimate daily rates of GPP for each of the validation lakes and these rates were used in our analysis. Only summertime metabolism data were used, as the niche models were generated using summertime sampling in the NLA data set and environmental data from the validation data set was collected over this same time span. Days for which metabolism fits had a coefficient of variation, based upon bootstrapped model fits, of greater than 50% were not used, as the model for the free-water method can sometimes produce low signal to noise ratios to which a simple metabolism model fit performs poorly (Solomon et al. 2013, Rose et al. 2014). This occurred in only 18% of the metabolism days. The mean of daily GPP observations represents each lake's seasonal GPP observation, and these means are compared to output from the trait-based and ecosystem models to evaluate model performance.

Daily metabolism estimates provided by the autonomous sensor technology allow for evaluation of temporal model performance in addition to spatial performance. Using a single lake, St. Gribso, Denmark, which had temporally resolved nutrient data available, we also examined the temporal performance of both the trait-based and ecosystem models when predicting daily GPP. This lake had at least weekly measurements of all physiochemical parameters and we linearly interpolated between sampling time points to generate daily estimates of physiochemical parameters. Both the trait-based and ecosystem models were simulated with the daily physiochemical parameters to generate daily estimates of GPP.

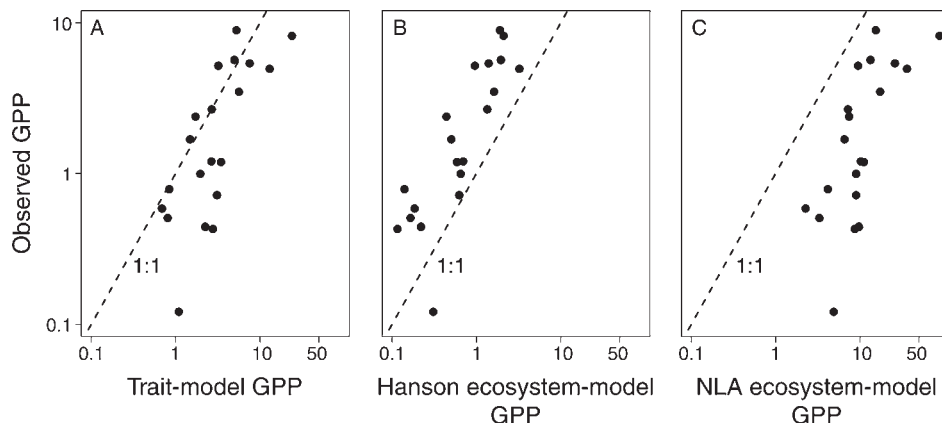


FIG. 1. Observed mean summer gross primary production (GPP) plotted against estimates from (A) the trait-based, (B) Hanson-ecosystem, and (C) NLA-ecosystem models. Each point represents a single lake ($n = 20$), and the dashed line represents the 1:1 line. All GPP observations and estimates were measured in $\text{mg O}_2\text{-L}^{-1}\text{-d}^{-1}$. Note that values are plotted on a log scale.

Model performance

We evaluate the performance of the trait-based and ecosystem models by comparing their predicted GPP values to observed GPP using several summary statistics. As an indicator of overall model predictive power we use root mean square error (RMSE). RMSE can be broken down into two components, bias and variance. We use mean error squared (MES) to describe model prediction bias, either systematically over- or underestimating observations, and we use R^2 to describe model prediction variance. As applied here, R^2 does not represent the variance in observations explained by our models. Rather, we use one minus the R^2 of model predictions regressed on observations, reflecting variance in model prediction after removing any effect of model bias. When reporting these statistics, we refer to one minus R^2 as model variance to avoid confusion with the traditional interpretation of R^2 , and we refer to MES as model bias. Model predictions with high bias have a high MES and model predictions with high variance have a high one minus R^2 . We used the R statistical program for all model fitting and subsequent analyses (R Development Core Team 2013).

RESULTS

Despite the coarse taxonomic resolution, our niche models explained a modest but significant ($P < 0.05$) amount of variation in division-specific phytoplankton biovolume across the NLA lakes ($R^2 = 0.07\text{--}0.30$; all likelihood-ratio tests $P < 0.001$). Our niche model parameters, representative of response traits, were consistent with previous observations of phytoplankton traits and community dynamics. All coefficients describing response to nutrients were positive, but this was not the case for temperature. Diatoms and cryptomonads had negative responses to water temperature, while the rest of the divisions responded positively (Appendix: Table A3).

Although phytoplankton community data were only available for one validation lake (Mendota), our niche models predicted phytoplankton biovolumes fairly well in this lake. Modeled and observed mean summer total phytoplankton biovolume were within an order of magnitude (off by a factor of 7.5), whereas observed total phytoplankton biovolume varied by well over two orders of magnitude across the sampling time points. Additionally, rank order of mean summer phytoplankton division abundance is consistent between our modeled and observed data for Mendota.

The trait-based model, which combined niche models with division-specific light use efficiency, made reasonably accurate predictions of mean summer GPP in the validation lakes (Fig. 1A). Bias in model output was low, although the trait-based model tended to overestimate observed rates of GPP. Both the trait-based model and Hanson-ecosystem model (Fig. 1B) performed much better than the NLA-ecosystem model (Fig. 1C), as the NLA-ecosystem model had poor predictive power and very high bias (Table 1). The Hanson-ecosystem model had better overall predictive power than our trait-based model, but higher bias. Despite higher bias, the Hanson-ecosystem model had the lowest variance in model output, highlighting a trade-off between bias and variance for the trait-based and Hanson-ecosystem models.

There was one highly influential, overestimated data point (Acton Lake, Ohio, USA) in the validation lakes for both the trait-based and NLA-ecosystem models. Removing this lake from the spatial analysis significantly improved model performance metrics for both the trait-based (RMSE = 2.47, bias = 0.88, variance = 0.59) and NLA-ecosystem model (RMSE = 10.92, bias = 74.00, variance = 0.63), while minimally affecting the Hanson-ecosystem model performance (RMSE = 2.42, bias = 2.53, variance = 0.39). This lake's physiochemical parameters were very high for a number of environmental variables in the NLA data set, including water

TABLE 1. Spatial and temporal model performance metrics for the trait-based and ecosystem models.

Model and metric	Trait-based	Hanson-ecosystem	NLA-ecosystem
Spatial			
RMSE	4.26	2.73	20.73
Bias	2.81	3.30	147.56
Variance	0.51	0.36	0.57
Temporal			
RMSE	1.37	1.58	5.76
Bias	0.31	1.66	22.35
Variance	0.72	0.995	0.87

Notes: Spatial model performance summarizes predicted gross primary production (GPP) values compared to observed GPP for the validation lakes ($n = 20$) and temporal model performance summarizes daily predicted GPP values compared to daily observed GPP for St. Gribso, Denmark, the lake at which the nutrient data required for the trait-based and ecosystem models were available at high temporal frequency ($n = 90$). Performance metrics are root mean square error (RMSE), mean error square (bias), and one minus the R^2 of observations regressed on model predictions (variance). The best-performing model for each metric is shown in boldface type.

temperature, TP, and TN (76th, 80th, and 97th percentile in the NLA data set respectively). Based on z -scored predictor variables, TN had the largest effect size (up to nine times the next most influential predictor) for phytoplankton biovolume in five of the seven division-specific niche models and the total biovolume niche model, while it had the second-strongest effect size for the two remaining division-specific niche models.

Despite being trained on spatial data, the trait-based model also performed well in predicting daily fluctuations in GPP within a single lake through time, and performed much better than both ecosystem models in this context (Fig. 2). The trait-based model had reasonable predictive power with relatively low bias and variance (Fig. 2A), whereas the Hanson-ecosystem model had a negative correlation with observed GPP (Pearson's $r = -0.07$) and both ecosystem models produced higher bias and variance and worse predictive power (Fig. 2B, C, Table 1).

DISCUSSION

The renewed interest in trait-based ecology to model both species distributions and ecosystem function is timely, given the pressing issues of global climate and land use change (McGill et al. 2006, Litchman and Klausmeier 2008, Suding et al. 2008). The promise of this approach is that it can explain the organization of ecological communities, and consequently ecosystem function. Our results provide compelling evidence for this trait linkage from environmental conditions to ecosystem function for aquatic systems, following previous work in terrestrial systems (Lavorel and Garnier 2002, Diaz et al. 2007, Violle et al. 2007). By integrating autonomous sensor-based estimates of ecosystem function and trait-based models, we have improved our ability to predict how environmental conditions dictate aquatic primary productivity via phytoplankton sorting.

Trait-based vs. ecosystem models

Without being calibrated by any ecosystem-scale measurements of GPP, our trait-based model predicted ecosystem GPP remarkably well in a global set of lakes, and displayed little bias. This emphasizes the portability afforded by consideration of traits and the efficacy of models incorporating some level of ecological detail to predict ecosystem function. Because traditional ecosystem models often use only a single predictor variable (TP in the case of the Hanson-ecosystem model), their output can be highly biased depending on the level of stationarity between environmental predictors in the training and prediction data sets. When applying the Hanson-ecosystem model to the validation lakes, we observed significant bias in our predictions. This likely reflects the geographically constrained nature of the training data set relative to the diversity of the global set of lakes used for validation.

The NLA-ecosystem model, although much less geographically constrained, also produced significant bias and poor model performance, highlighting the importance of including at least some taxonomic resolution when using response and/or effect traits to predict ecosystem function. Our trait-based model was constrained to the taxonomic resolution of phytoplankton divisions due to the limited availability of effect trait data. We think including more taxonomic resolution would improve our trait-based estimates of ecosystem function, echoing the need for a global database of phytoplankton traits (Litchman and Klausmeier 2008).

Although the trait-based model had lower bias than the Hanson-ecosystem model, the model complexity may promote variation in the model output, resulting in higher model variance. Mathematically, there is a trade-off in variance and bias for a given model, and it is often difficult to reduce both simultaneously (Hastie et al. 2009). More complex models tend to have higher variance in model output as they are typically more sensitive to small changes in driver data. Conversely,

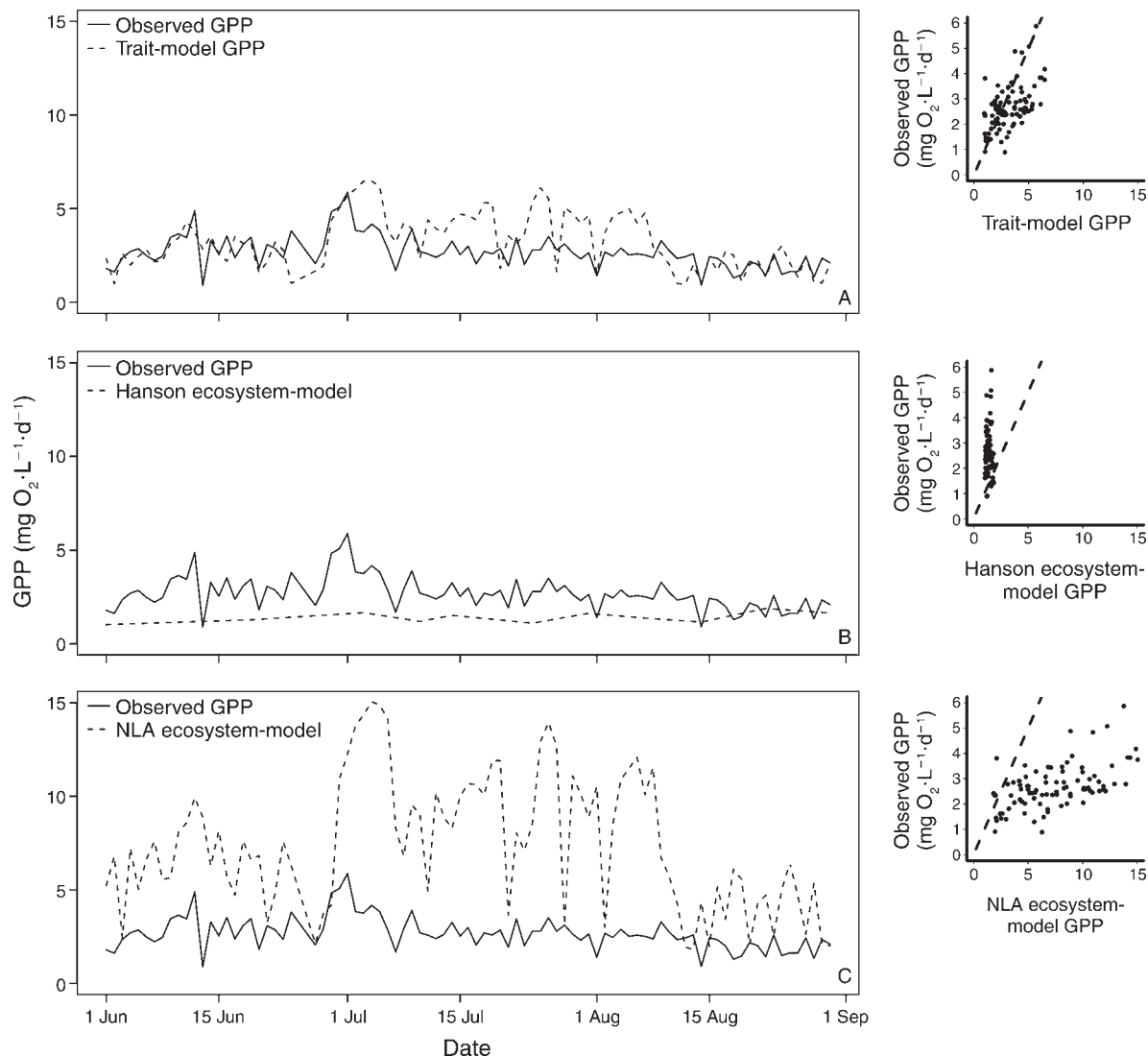


FIG. 2. Daily observations of GPP ($n = 90$) from a single lake compared with daily predictions from (A) the trait-based, (B) Hanson-ecosystem, and (C) NLA-ecosystem models. Each panel includes time series of the daily observations and model predictions and smaller scatter plots (right) of daily observations and predictions where each point is a day and the dashed line is the 1:1 line. Data are from St. Gribsoe, Denmark, the lake at which the nutrient data required for the trait-based and ecosystem models were available at high temporal frequency.

simple models often have higher bias in model output, as their simplicity may not capture important additional drivers that vary across systems and regions (e.g., light, in addition to TP, regulating phytoplankton productivity; Karlsson et al. 2009). Our comparison of trait-based and Hanson-ecosystem GPP models provides further support for the trade-off between model bias and variance along a gradient of model complexity.

Predicting ecosystem function through time in a single location with models that are trained using spatial data may fail to accurately capture more subtle temporal changes in environmental drivers. All models used in our analysis were trained using only spatial data, however, the trait-based model outperformed both ecosystem

models when applied temporally (Fig. 2, Table 1). Previous efforts to substitute space for time in an attempt to predict ecological processes under changing environmental conditions have produced mixed results (Pickett 1989, Johnson and Miyanishi 2008, Blois et al. 2013), and can perform poorly when spatial variation in driver data is much larger than temporal variation (Alder and Levine 2007, Blois et al. 2013). Even though every physiochemical driver used in the trait-based model had more spatial variation than temporal variation (indicated by a higher coefficient of variation), the temporal performance is still strong, underlining the flexibility of trait-based approaches to predict ecosystem function both spatially and temporally.

We show that laboratory-measured phytoplankton effect traits can be used to predict ecosystem function across both space and time. Although our trait-based predictions were accurate, we did observe a tendency for our trait-based model to overestimate GPP. One explanation for this overestimation is that α is, on average, reduced in the natural environment relative to optimal conditions provided in a laboratory setting. For instance, photosynthetic rates for three species of phytoplankton were reduced under phosphorus limitation (Litchman et al. 2003), and temperature changes can effect growth rate more than fourfold in some species of phytoplankton (Maddux and Jones 1964). Additionally, light saturation has been shown to reduce photosynthetic efficiency in some natural environments, and might explain some of the observed overestimation (Platt and Jassby 1976). Assuming that phytoplankton performance is reduced in the suboptimal natural environmental, we can ask by how much we would have to discount lab-derived α estimates to most accurately predict the observed GPP. Discounting all division-specific α values by 54% provided the best correspondence between trait-based and observed GPP. It is reasonable to expect a slight overestimate, perhaps as high as 54%, of light use efficiency under optimal laboratory conditions. Just as adjusting for optimal light use efficiency in terrestrial plants improved terrestrial ecosystem GPP estimates (Madani et al. 2014), our findings suggest that estimation of phytoplankton effect traits under natural conditions or the inclusion of more sophisticated models of photosynthesis that include the potential for photoinhibition (Jassby and Platt 1976) would further improve aquatic trait-based model performance.

Future applications

We use GPP as the ecosystem function of choice, but there are many other aquatic ecosystem processes that could be modeled under the trait-based framework. For example, cyanobacteria toxin production may be accurately modeled if taxon-specific toxin production rates were known. Environmental change will favor cyanobacteria over other phytoplankton through their response traits, and predicting community toxin production would benefit resource managers who need to predict toxin levels that may be dangerous to livestock, pets, and humans (Carey et al. 2012). Additionally, the trait-based approach could be applied to phytoplankton production of dimethylsulfide (DMS), which can influence global climate. Dimethylsulfide production rates can vary over three orders of magnitude across phytoplankton species, which is 10 times more than in the set of α 's that we used in the current study; thus it is likely that estimates of community DMS production would benefit from taxon-specific modeling under a trait-based framework (Charlson et al. 1987).

Aquatic environments are extremely dynamic and there is a quick turnover time in phytoplankton communities.

Coupled with the ability to estimate temporally resolved ecosystem function using autonomous sensor technology, we suggest that these ecosystems are ideal for examining emergent properties of complex systems and, ultimately, can help bridge the gap between population, community, and ecosystem ecology.

ACKNOWLEDGMENTS

We are grateful to the Global Lake Ecological Observatory Network for facilitating lake data sharing and thank all the data providers to Solomon et al. (2013). Data providers: Michael Vanni, Miami University (Acton); Hilary Swain and Evelyn Gaiser, Archbold Biological Station (Annie); Peter Staehr, Aarhus University (Fredriksburg Slotsø, Hampensø, St. Gribso, and Vedstedso); Paul Hanson, University of Wisconsin (Crystal Bog, Mendota, Trout, and Trout Bog); David Hamilton, University of Waikato (Rotorua and Rotorua); Charles Chiu, Academia Sinica (Yuan Yang); Guangwei Zhu, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences (Taihu); Vera Istvanovics, Water Quality Management Group, Hungarian Academy of Sciences (Balaton); Michael Pace, University of Virginia (Crampton); Elvira de Eyto and Elizabeth Ryder, Marine Institute (Feeagh); Susan Hendricks, Hancock Biological Station (Kentucky); Rita Adrian, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (Muggelsee); and Kathleen Weathers, Cary Institute of Ecosystem Studies (Sunapee). We thank the Jones lab members and M. F. Ebenezer for helpful comments on previous versions of the manuscript and two anonymous referees for their tremendously useful comments. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under NSF DGE-1313583 to J. A. Zwart, an NSERC Discovery grant to C. T. Solomon, and NSF DEB-1442230 to S. E. Jones. J. A. Zwart, C. T. Solomon, and S. E. Jones conceived the study; C. T. Solomon produced estimates of GPP for the validation lakes, J. A. Zwart carried out the rest of the analyses and wrote the manuscript with input from C. T. Solomon and S. E. Jones.

LITERATURE CITED

- Alder, P. B., and J. M. Levine. 2007. Contrasting relationships between precipitation and species richness in space and time. *Oikos* 116:221–232.
- Blois, J. L., J. W. Williams, M. C. Fitzpatrick, S. T. Jackson, and S. Ferrier. 2013. Space can substitute for time in predicting climate-change effects on biodiversity. *Proceedings of the National Academy of Sciences USA* 110:9374–9379.
- Carey, C. C., B. W. Ibelings, E. P. Hoffmann, D. P. Hamilton, and J. D. Brookes. 2012. Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. *Water Research* 46:1394–1407.
- Charlson, R. J., J. E. Lovelock, M. O. Andreae, and S. G. Warren. 1987. Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* 326:655–661.
- Cole, J. J., M. L. Pace, S. R. Carpenter, and J. F. Kitchell. 2000. Persistence of net heterotrophy in lakes during nutrient addition and food web manipulations. *Limnology and Oceanography* 45:1718–1730.
- Cunningham, R. B., and D. B. Lindenmayer. 2005. Modeling count data of rare species: some statistical issues. *Ecology* 86: 1135–1142.
- de Bello, F., et al. 2010. Towards an assessment of multiple ecosystem processes and services via functional traits. *Biodiversity and Conservation* 19:2873–2893.
- Diaz, S., S. Lavorel, F. de Bello, F. Quétiér, K. Grigulis, and T. M. Robson. 2007. Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences USA* 104:20684–20689.

- Dillon, P. J., and F. H. Rigler. 1974. The phosphorus-chlorophyll relationships in lakes. *Limnology and Oceanography* 19:767–773.
- Edwards, K. F., E. Litchman, and C. A. Klausmeier. 2013a. Functional traits explain phytoplankton community structure and seasonal dynamics in a marine ecosystem. *Ecology Letters* 16:56–63.
- Edwards, K. F., E. Litchman, and C. A. Klausmeier. 2013b. Functional traits explain phytoplankton responses to environmental gradients across lakes of the United States. *Ecology* 94:1626–1635.
- Hanson, P. C., D. L. Bade, S. R. Carpenter, and T. K. Kratz. 2003. Lake metabolism: relationships with dissolved organic carbon and phosphorus. *Limnology and Oceanography* 48:1112–1119.
- Hanson, P. C., A. I. Pollard, D. L. Bade, K. Predick, S. R. Carpenter, and F. A. Foley. 2004. A model of carbon evasion and sedimentation in temperate lakes. *Global Change Biology* 10:1285–1298.
- Hastie, T., R. Tibshirani, and J. Friedman. 2009. The elements of statistical learning: data mining, inference, and prediction. Second edition. Springer, New York, New York, USA.
- Jassby, A. D., and T. Platt. 1976. Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. *Limnology and Oceanography* 21:540–547.
- Johnson, E. A., and K. Miyanishi. 2008. Testing the assumptions of chronosequences in succession. *Ecology Letters* 11:419–431.
- Karlsson, J., P. Byström, J. Ask, P. Ask, L. Persson, and M. Jansson. 2009. Light limitation of nutrient-poor lake ecosystems. *Nature* 460:506–510.
- Kearney, M., and W. Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters* 12:1–17.
- Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* 332:580–582.
- 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.
- Lavorel, S., and E. Garnier. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology* 16:545–556.
- Litchman, E., and C. A. Klausmeier. 2008. Trait-based community ecology of phytoplankton. *Annual Review of Ecology, Evolution, and Systematics* 39:615–639.
- Litchman, E., D. Steiner, and P. Bossard. 2003. Photosynthetic and growth responses of three freshwater algae to phosphorus limitation and daylength. *Freshwater Biology* 48:2141–2148.
- Loreau, M. 2010. Linking biodiversity and ecosystems: towards a unifying ecological theory. *Philosophical Transactions of the Royal Society B* 365:49–60.
- Madani, N., J. S. Kimbell, D. L. R. Affleck, J. Kattge, J. Graham, P. M. van Bodegom, P. B. Reich, and S. W. Running. 2014. Improving ecosystem productivity modeling through spatially explicit estimation of optimal light use efficiency. *Journal of Geophysical Research—Biogeosciences* 119:1755–1769.
- Maddux, W. S., and R. F. Jones. 1964. Some interactions of temperature, light intensity, and nutrient concentration during the continuous culture of *Nitzschia closterium* and *Tetraselmis* sp. *Limnology and Oceanography* 9:79–86.
- McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution* 21:178–185.
- 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.
- North Temperate Lakes LTER. 2011. North Temperate Lakes LTER: phytoplankton. Madison Lakes Area, 1995–current. Center for Limnology, University of Wisconsin, Madison, Wisconsin, USA.
- Pickett, S. T. A. 1989. Space-for-time substitution as an alternative to long-term studies. Pages 110–135 in G. E. Likens, editor. *Long-term studies in ecology*. Springer-Verlag, New York, New York, USA.
- Platt, T., and A. D. Jassby. 1976. The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. *Journal of Phycology* 12:421–430.
- R Development Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.r-project.org
- Rose, K. C., L. A. Winslow, J. S. Read, E. K. Read, C. T. Solomon, R. Adrian, and P. C. Hanson. 2014. Improving the precision of lake ecosystem metabolism estimates by identifying predictors of model uncertainty. *Limnology and Oceanography: Methods* 12:303–312.
- Schindler, D. W. 1977. Evolution of phosphorus limitation in lakes. *Science* 195:260–262.
- 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.
- Smith, V. H. 1979. Nutrient dependence of primary productivity in lakes. *Limnology and Oceanography* 24:1051–1064.
- Solomon, C. T., 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. Diaz, E. Garnier, D. Goldberg, D. U. Hooper, S. T. Jackson, and M. Navas. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Global Change Biology* 14:1125–1140.
- USEPA. 2009. National lakes assessment: a collaborative survey of the nation's lakes. United States Environmental Protection Agency, Washington, D.C., USA.
- Van de Bogert, M. C., S. R. Carpenter, J. J. Cole, and M. L. Pace. 2007. Assessing pelagic and benthic metabolism using free water measurements. *Limnology and Oceanography: Methods* 5:145–155.
- Violle, C., M. Navas, D. Vile, E. Kazakou, C. Fortunel, and I. Hummel. 2007. Let the concept of trait be functional! *Oikos* 116:882–892.
- Vörös, L., and J. Padisák. 1991. Phytoplankton biomass and chlorophyll-*a* in some shallow lakes in central Europe. *Hydrobiologia* 215:111–119.

SUPPLEMENTAL MATERIAL

Ecological Archives

The Appendix is available online: <http://dx.doi.org/10.1890/14-2102.1.sm>