LIMNOLOGY and OCEANOGRAPHY LETTERS

ASLO

Open Access

Limnology and Oceanography Letters 2025

doi: 10.1002/lol2.70062

LETTER

Predicting photosynthesis–irradiance relationships from satellite remote-sensing observations

Gregory L. Britten,[1,2*] **Bror Jönsson**,[3] **Gemma Kulk**,[3,4] **Heather A. Bouman**,[5] **Michael J. Follows**,[2] **Shubha Sathyendranath**[3,4]

[1]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA; [2]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; [3]Earth Observation Science and Applications, Plymouth Marine Laboratory, Plymouth, UK; [4]National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, UK; [5]Department of Earth Sciences, University of Oxford, Oxford, UK

Scientific Significance Statement

The relationship between photosynthesis–irradiance (PI) is important in marine ecosystems and the global carbon cycle but is difficult to predict due to multiple relevant processes interacting on different timescales. This study leverages machine learning algorithms and an in situ PI database to predict these relationships using oceanographic data observed from satellites. We develop a method to identify the optimal environmental integration timescale to predict PI relationships and then use the method to estimate how PI parameters vary seasonally. We link these timescales to patterns of variability in light, temperature, and surface chlorophyll. This study helps us better understand how photosynthesis responds to environmental variation and will improve our ability to model marine primary production from satellite remote-sensing observations.

Abstract

Photosynthesis–irradiance (PI) relationships are important for phytoplankton ecology and quantifying carbon fixation rates in the environment. However, the parameters of PI relationships are typically unknown across space and time. Here we use machine learning, satellite remote-sensing, and a database of in situ PI relationships to build models that predict the seasonal cycle of PI parameters as a function of satellite-observed variables. Using only surface light, temperature, and chlorophyll, we achieve an R^2 of 58% for predicting photosynthesis rates at saturating light ($P^B_{\max}$) and an R^2 of 78% for predicting the light saturation parameter (E_k). Predictability is maximized when averaging environmental covariates over 30-d ($P^B_{\max}$) and 25-d (E_k) timescales, indicating that environmental history and community turnover timescales are important for predicting in situ PI relationships. These results will help improve the parameterization of satellite-based primary production models and quantify emergent environmental integration timescales in photosynthetic communities.

*Correspondence: gregory.britten@whoi.edu

Associate editor: Raphael Kudela

The relationship between the rate of photosynthesis and environmental light availability (termed the photosynthesis–irradiance relationship or PI relationship) is central to quantifying rates of carbon fixation and forms the basis for many models of marine primary productivity. As such, an accurate description of the PI relationship and its response to environmental variation is important for understanding marine ecosystem productivity.

The most common form of the PI relationship is given by

$$P^B = P^B_{\max}\left(1 - e^{-\frac{\alpha^B I}{P^B_{\max}}}\right) \quad (1)$$

where P^B (mg C mg Chl $a^{-1}\,h^{-1}$) is the chlorophyll-normalized rate of photosynthesis, $P^B_{\max}$ (mg C mg Chl $a^{-1}\,h^{-1}$) is the maximum rate of normalized photosynthesis at saturating light, α^B (mg C mg Chl $a^{-1}\,s^{-1}$ [μmol photons $m^{-2}\,s^{-1}$]$^{-1}$) is the slope at the origin of the PI curve, and I is the in situ irradiance (μmol photons $m^{-2}\,s^{-1}$). The ratio of the two parameters, $E_k = \frac{P^B_{\max}}{\alpha^B}$ (μmol photons $m^{-2}\,s^{-1}$), governs the shape of the saturating function and is referred to as the light saturation parameter (Bouman et al. 2018). Since E_k is a function of α^B and $P^B_{\max}$, the three parameters are redundant as a set. We focus on E_k and $P^B_{\max}$ for our analysis due to their broad interpretability. Parameters of the PI relationship are estimated by incubating seawater samples to a range of light levels and measuring photosynthesis (typically by ^{14}C uptake), yielding PI data to which the parameters are fitted by regression (Platt and Jassby 1976; Lewis and Smith 1983; Kulk et al. 2023).

In natural phytoplankton communities, PI relationships can be difficult to predict due to interacting factors related to species physiology, community turnover, and their dependence on environmental conditions. Photosynthesis is performed by multispecies assemblages due to species coexistence (Hutchinson 1961; Dutkiewicz et al. 2020), with individual phytoplankton species exhibiting distinct PI relationships (Kulk et al. 2011). The physiological state of individual phytoplankton species is also time-dependent with physiological processes regulated on timescales of microseconds to days (Geider et al. 1998; Liefer et al. 2018). Despite the complexity, predictive relationships have emerged between environmental variables and community-level PI relationships. For example, Bouman et al. (2005) found strong but contrasting relationships across ocean regions, while Saux Picart et al. (2014) found that $P^B_{\max}$ varied positively with temperature. The satellite-based vertically generalized productivity model parameterizes the $P^B_{\max}$–temperature relationship with a dome-shaped 5$^{\text{th}}$-order polynomial (Behrenfeld and Falkowski 1997). An alternative power–law relationship has also been used in satellite-based models, sensu Eppley (Eppley 1972; Taboada et al. 2019). Regarding timescales, Platt and Sathyendranath (1993) found that community-level PI relationships were better predicted using data averaged over 3 d relative to using environmental data from the day of sampling, suggesting emergent environmental timescales that integrate beyond the sampling day. Here, we leverage a database of experimentally determined in situ PI relationships to further develop predictive models of PI parameters and apply them across regions of the global ocean.

Methods

We analyzed the database curated by Bouman et al. (2018) and updated by Kulk et al. (2020, 2021). All measurements analyzed here were taken using ^{14}C uptake as a measure of photosynthetic rate using incubations between 1.5 and 4 h in duration (Bouman et al. 2018). The database reports parameters for fitted PI relationships following either Eq. 1 when photoinhibition was absent. When photoinhibition is present, the PI curve does not contain an explicit asymptotic $P^B_{\max}$ so $P^B_{\max}$ was derived via analytical expressions as the maximum of the PI curve using formulae from Bouman et al. (2018). We chose subsets of PI relationships originating from established regional sampling programs wherein the same locations are sampled across time and over a range of environmental conditions using consistent sampling and experimental methodology. Analyzing these subsets acts to minimize bias when collating datapoints that differ in collection methodology (e.g., Britten et al. 2021). Samples were taken from the mixed layer based on an established monthly mixed layer depth climatology (Montegut et al. 2004) to best correspond with satellite observations. Figure 1 gives the sampling locations of data analyzed in this study. Metadata for the chosen subsets are given in Supporting Information Table S1.

We matched PI relationships with satellite-derived photosynthetically available radiation (PAR), sea surface temperature (SST), and chlorophyll *a* (Chl *a*). We used 4 km daily-averaged surface PAR from merged MODIS and SeaWIFS version R2018.0 (Frouin et al. 2012). Sea surface temperature at 6 km resolution was from version 2.0 of the multi-sensor Operational Sea Surface Temperature and Ice Analysis foundation temperature product (Donlon et al. 2012). We used Chl *a* from version 5.0 of the multi-sensor Ocean Color Climate Change Initiative 4 km product (Sathyendranath et al. 2019, 2021). We also used SST and Chl *a* to compute an estimate of the picophytoplankton contribution to Chl *a*, denoted $\frac{C_{\text{pico}}}{C_{\text{total}}}$ according to equations given in Brewin et al. (2010). We constructed a series of temporal averages of the variables, with each average extending 1 d further back in time. We do this over 100 d for individual in situ sampling locations, generating 100 covariate datasets to be used in the modeling of the PI parameters as a function of environmental integration timescale. Temporally averaged time series for PAR, SST, and Chl *a* are shown in Supporting Information Fig. S1. *e*-Folding timescales for the time-averaged covariates were estimated for each region by finding the time lag when the autocorrelation

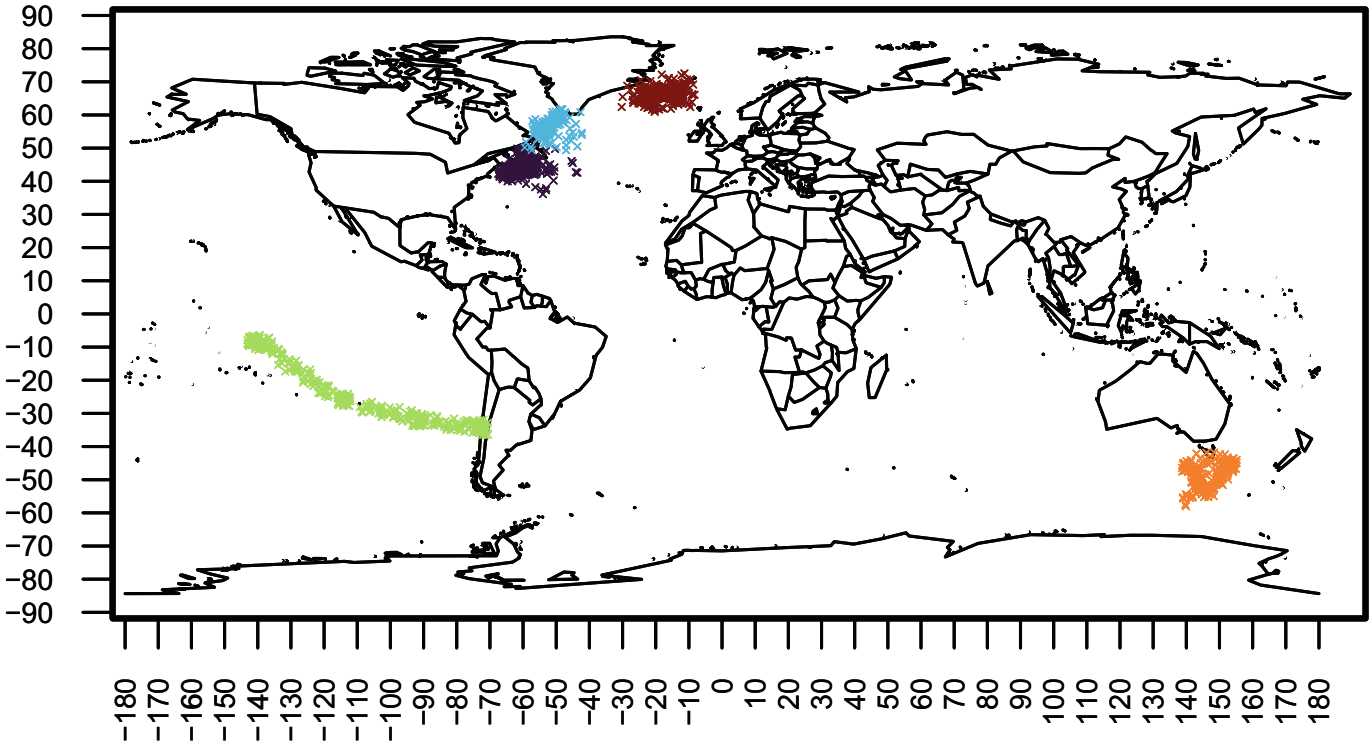

Fig. 1. Locations of the satellite-matched in situ photosynthesis–irradiance experiments contained within the MArine primary Production: model Parameters from Space (MAPPS) database (Bouman et al. 2018; Kulk et al. 2020, 2021). Regions are given in color (purple: Scotian Shelf, light blue: Labrador Sea, red: Iceland Shelf, green: South Pacific, orange: Southern Ocean). The full satellite-matched dataset is available on Dryad (10.5061/dryad.w6m905r1j).

of the time series drops by a factor of $1/e$, yielding 7–12 d, 11–15 d, and greater than 60 d for PAR, SST, and Chl *a*, respectively (Supporting Information Fig. S2). We note that time series of daily satellite-based Chl *a* (and therefore $\frac{\overline{c_{\text{pico}}}}{c_{\text{total}}}$) suffer from a relatively large number of missing values due to cloud cover on daily timescales, causing the stepwise character of many Chl *a* histories (Supporting Information Fig. S1k–o).

We applied random forest regression algorithms as a robust and easily implemented machine learning model to generate predictions across environmental integration timescales. Random forests were chosen over other predictive machine learning algorithms (e.g., neural networks) for ease of use, requiring little hyperparameter tuning while rivaling neural networks for predictive accuracy for smaller datasets (Cutler et al. 2007; Liu et al. 2013; Grinsztajn et al. 2022). Random forests were fitted by randomly generating bootstrap subsets of the samples and covariates, then computing decision trees for each subset. Predictions were then averaged across individual trees. We present results of random forests of the form

$$P^B_{\max}(t_i) = \text{RF}\left(\overline{\text{PAR}}(t_i), \overline{\text{SST}}(t_i), \frac{\overline{c_{\text{pico}}}}{c_{\text{total}}}(t_i), \text{depth}, \text{region}\right)$$

$$E_k(t_i) = \text{RF}\left(\overline{\text{PAR}}(t_i), \overline{\text{SST}}(t_i), \frac{\overline{c_{\text{pico}}}}{c_{\text{total}}}(t_i), \text{depth}, \text{region}\right)$$

where separate random forests are fit to $P^B_{\max}$ and E_k as a function of satellite-estimated covariates averaged over t_i preceding days. RF represents a function constructed by averaging decision trees in the random forest, $\overline{\text{PAR}}(t_i)$, $\overline{\text{SST}}(t_i)$, and $\frac{\overline{c_{\text{pico}}}}{c_{\text{total}}}(t_i)$ are PAR, SST, and the fraction of Chl *a* that is estimated as picophytoplankton, respectively. We tested both Chl *a* and $\frac{c_{\text{pico}}}{c_{\text{total}}}$ as covariates (but not both due to redundancy) and found $\frac{c_{\text{pico}}}{c_{\text{total}}}$ to give a slightly higher R^2. We also tested mixed layer-averaged PAR using additional estimates of attenuation depth but achieved lower R^2, possibly due to the uncertainty introduced using satellite-estimates of attenuation and climatological mixed layer depth. Depth is the water column depth of the PI sample within the mixed layer and region is a categorical variable representing the geographic location of the collected sample (colors in Fig. 1). depth and region do not change with timescale so are not a function of t_i. We fit a separate random forest to each temporal average, constituting 100 random forests for each parameter. All model fits and diagnostics were performed using the *randomForest* package in the R programming language (Liaw and Wiener 2002). We computed predictability as the out-of-bag coefficient of determination (R^2), where out-of-bag refers to predictions made using trees that did not contain the given datapoint in the training set. Fitted relationships to individual covariates were

visualized using partial effects where random forest predictions are plotted against a single variable after averaging (marginalizing) across all other covariates in the model (Liaw and Wiener 2002). We also compare the random forest performance to a linear mixed effects model using the same covariate sets, including region as a categorical random effect. We used the *lme4* package in R to fit the mixed effects models (Bates et al. 2013). To aid in interpreting results, we also conducted idealized numerical simulations of a multispecies plankton community to estimate community turnover timescales and help interpret statistical results. Finally, we used the fitted models to predict maps of surface (depth = 0) $P^B_{\max}$ and E_k at the global scale, using global maps of input parameters $\overline{\mathrm{PAR}}(t_i), \overline{\mathrm{SST}}(t_i)$, and $\overline{\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}}}(t_i)$ and the identified optimal averaging timescale.

Results

We found that predictive performance of the fitted random forests varied as a function of environmental averaging timescale, with relationships differing for parameters $P^B_{\max}$ and E_k (solid lines in Fig. 2). Predictive capacity was highest for E_k,, achieving a maximum out-of-bag R^2 of 78% when averaging environmental covariates over 25 d (Fig. 2b) and a maximum R^2 of 58% for $P^B_{\max}$ when averaging over 30 d (Fig. 2a). Predictive capacity was lower for the mixed effects regression models (dashed lines in Fig. 2) and was achieved on somewhat longer averaging timescales, although the dome-shaped relationship was consistent with predictive maxima at intermediate timescales: $P^B_{\max}$ R^2 maximized at 38% over 39 d averaging and E_k R^2 maximized at 65% over 48 d of averaging.

Partial effects of individual covariates showed consistent patterns across regions with all covariates driving variability in predicted PI parameters (Fig. 3). The mean effect sizes are given in each panel of Fig. 3 in standardized units of $\Delta P^B_{\max}$ or ΔE_k per unit standard deviation of the covariate. PAR showed the weakest partial effects across regions with minimal variability in both $P^B_{\max}$ and E_k within a region (Fig. 3a,e), while SST, $\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}}$, and depth showed stronger effect sizes (Fig. 3b–d,f–h). Partial effects of SST on $P^B_{\max}$ were regionally consistent, showing a nearly linear effect across the range of observed SST (independent of whether $\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}}$ or Chl *a* was included in the model). Effects of $\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}}$ appeared nonlinear with notable threshold behavior, where a nonlinear increase in $P^B_{\max}$ occurs at $\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}} = 0.25$ and at $\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}} = 0.4$ for E_k. In Supporting Information Figs. S3 and S4, we show that PAR is often invoked in individual decision trees but occurs at lower nodes, meaning that PAR became an important effect, conditional on a particular SST and $\frac{c_{\mathrm{pico}}}{c_{\mathrm{total}}}$ regime. Effects estimated in the random forest are broadly consistent with regression coefficients estimated by the mixed effects regression model (Supporting Information Fig. S5) but showed some differences due to the more limited linear structure of the model. The strongest effects estimated within the mixed effects regression was a positive effect of SST on $P^B_{\max}$ (Supporting Information Fig. S5a) and a positive

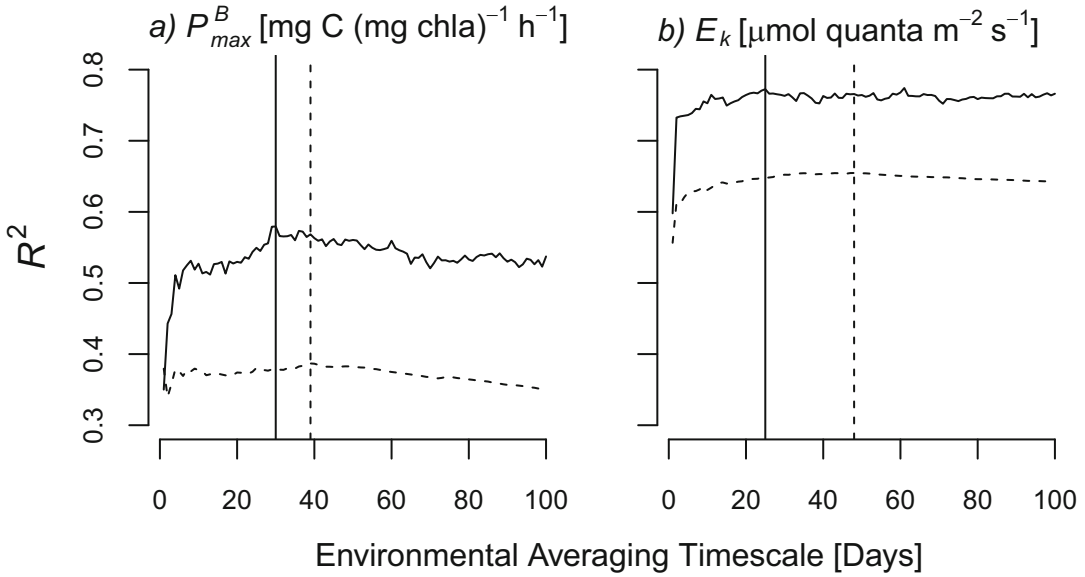

Fig. 2. Predictability of PI parameters as a function of environmental integration timescale. Left figure gives model R^2 values for $P^B_{\max}$; right figure gives R^2 values for E_k. Solid curves give R^2 for random forests; dashed curves give R^2 for mixed effects regression; vertical solid and dashed lines give the timescale that maximizes R^2 for random forests and mixed effects regression, respectively.

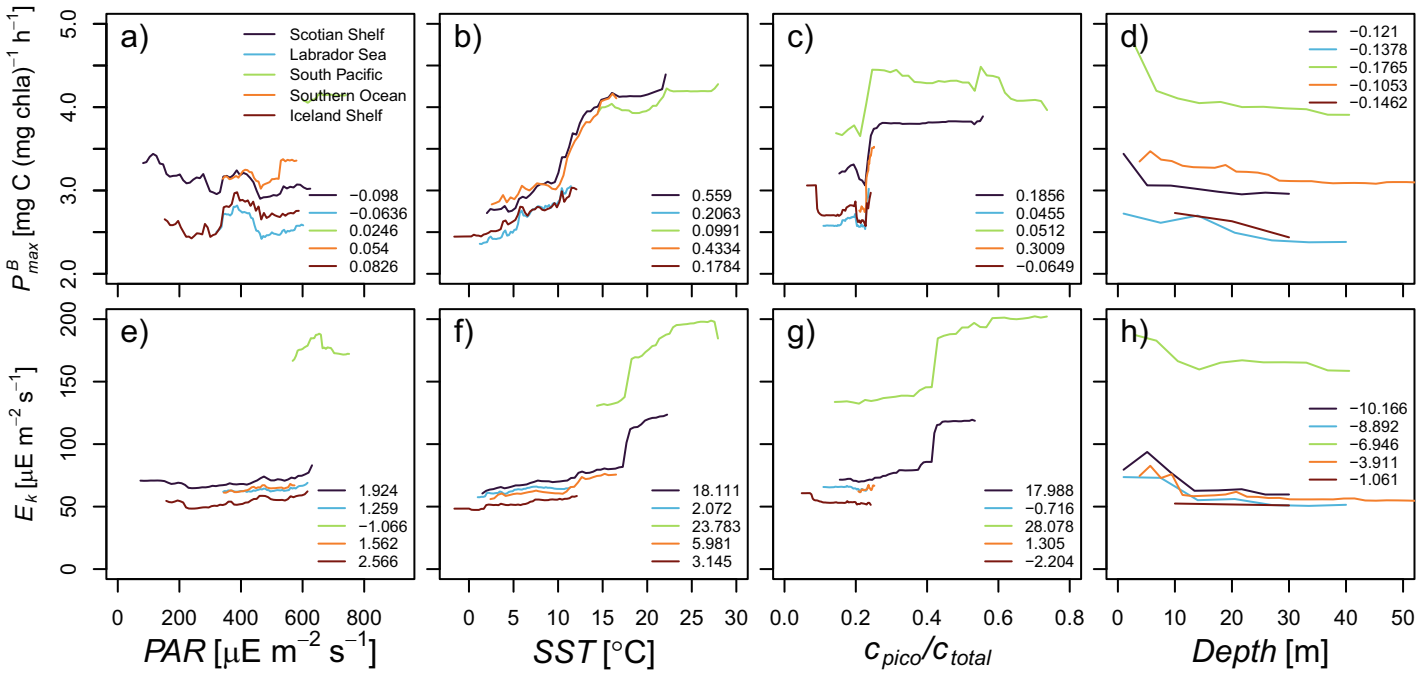

Fig. 3. Partial effects of predictor variables for the random forest with the identified best-fitting timescale. Model for $P^B_{\max}$ averages environmental covariates over preceding 25 d (**a**–**d**); model for E_k averages environmental covariates over 30 d (**e**–**h**). Partial effects are computed by averaging (marginalizing) model predictions across all other covariates in the model. The mean partial effect size for each region is given in each panel in standardized units of $\Delta P^B_{\max}$ or ΔE_k per unit standard deviation of the covariate.

effect of $\frac{c_{\text{pico}}}{c_{\text{total}}}$ on E_k (Supporting Information Fig. S5b). In Supporting Information Fig. S6, we show partial effects from the model fitted using an environmental integration timescale of 10 d (where R^2 begins to plateau but does not yet achieve a maximum in Fig. 2), demonstrating robust patterns in covariate relationships across timescales.

Seasonal analysis at the regional-scale shows variability in environmental covariates both within and between regions (Fig. 4a–e). The shape of the time-integrated Chl *a* time series was most variable within regions, with a temporally shifted spring bloom present in all Chl *a* time series and a fall bloom present in the Scotian Shelf, Iceland Shelf, and to a lesser extent the Labrador Sea. Integrated SST and PAR time series are less variable, following the latitude of individual samples. The largest seasonal cycle predicted for $P^B_{\max}$ occurred on the Scotian Shelf with $\Delta P^B_{\max}$ over the seasonal predicted to be approximately 1.8 (mg C mg Chl a^{-1} h^{-1}). Less seasonal variation was observed in other regions, particularly the South Pacific which showed no discernable $P^B_{\max}$ cycle ($\Delta P^B_{\max} < 0.2$ mgC mg Chl a^{-1} h^{-1}). Seasonal cycles in E_k were relatively large on the Scotian Shelf with greater than two-fold variation, followed by the South Pacific. Seasonal cycles for E_k were approximately flat for the Labrador Sea and Iceland Shelf.

Using the fitted models, we also extrapolate the seasonal cycle globally to assess whether interpretable patterns emerge from a regional analysis at the global scale (Fig. 5). Contrasting PI parameter predictions in winter and summer, we see that $P^B_{\max}$ follows a distinct latitudinal structure with maxima occurring in mid-latitude winter in both northern and southern hemispheres, along with equatorial maxima occurring year-round (Fig. 5a,b). Lowest $P^B_{\max}$ is predicted at high latitudes in hemispheric summers. Predictions for E_k follow a different latitudinal pattern where maxima are predicted in subtropical summer within the oligotrophic gyres (Fig. 4c, d). Predicted values for E_k fall by roughly 20–30% when moving from subtropical to equatorial environments and by more than half when moving from subtropical to high latitude environments. We note that global extrapolations are based on limited regional coverage and therefore should be interpreted with inherent uncertainty.

Discussion

This study quantified the predictability of PI parameters across environmental integration timescales as a function of satellite-estimated covariates. We argue that environmental integration timescales (t_i^{opt} ~25–30 d) of PI parameter predictability found here are consistent with what is expected from community turnover timescales. From basic theoretical calculations, we expect community turnover timescales to range from 10 d to 100+ d (Supporting Information Fig. S7), as consistent with other theoretical and empirical studies

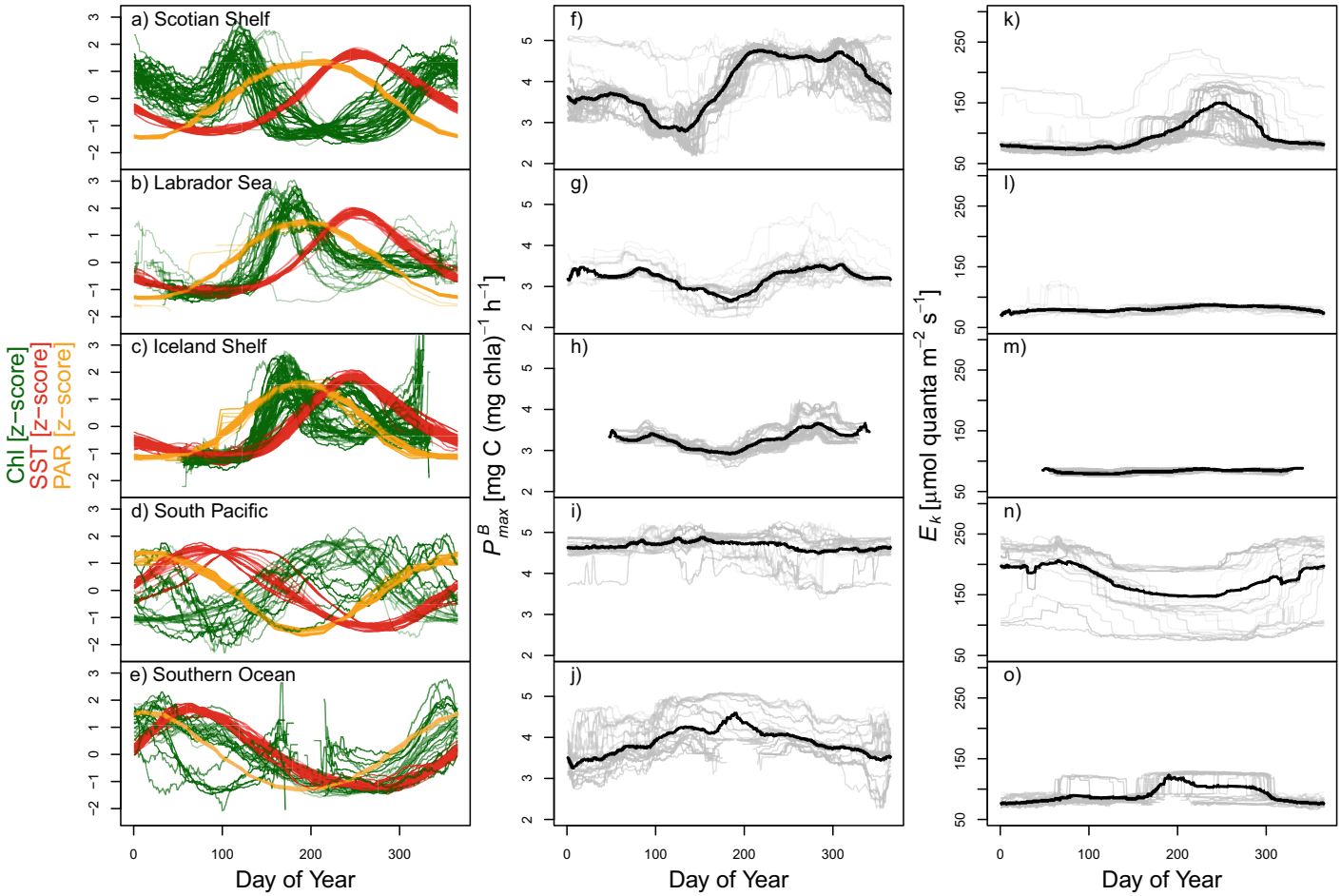

Fig. 4. Predicted seasonal dynamics of PI parameters. (**a**–**e**) 25 day–integrated satellite covariates chlorophyll (Chl *a*), sea surface temperature (SST), and photosynthetically active radiation (PAR) for all sampling sites per region. Time series are z-score transformed to compare trajectories. (**f**–**j**) give seasonal predictions for $P^B_{\max}$ across regions. Thin gray lines give site-specific predictions and solid black line gives the mean. (**k**–**o**) Same but for E_k.

(Hutchinson 1961; Anderson et al. 2022; Gallego and Narwani 2022). This result supports the broader notion that community turnover is important for predicting community-integrated rates (Leibold et al. 1997). We find that predictability sharply declines on timescales less than 5 d for $P^B_{\max}$ and less than 10 d for E_k, indicating a degree of environmental 'noise' on these timescales. Sources of noise may include both satellite-based measurement error and variability in the community response to environmental processes. Five to 10 d roughly matches the synoptic meteorological timescale driving clouds, suggesting that phytoplankton communities may buffer meteorological PAR variability with respect to their PI parameters. The balance of environmental and community turnover timescales relates to the Hutchinsonian solution to the paradox of the plankton (Hutchinson 1961), where communities experience perpetual turnover when environmental and turnover timescales match. We suggest a similar balance of timescales may be operating here.

Lower predictability of $P^B_{\max}$ relative to E_k may be due to several factors. For one, we note that raw correlations between $P^B_{\max}$ and E_k are regionally variable and generally weak, achieving R^2 values across regions of 2–37% (Supporting Information Fig. S8) indicating that distinct processes are controlling variation in the two parameters, consistent with notions of E_k-dependent physiological processes (Behrenfeld et al. 2008). Second, previous studies have identified a close association between E_k and in situ irradiance (Sathyendranath et al. 2020), noting that community level E_k is often measured close, but slightly lower, than in situ irradiance. $P^B_{\max}$, on the other hand, is often a large extrapolation at light levels much higher than those measured in situ. We also tried fitting models to predict α^B and constructing E_k predictions from the ratio of predicted $P^B_{\max}$ and α^B (results not shown), but generally found poorer results than those presented here.

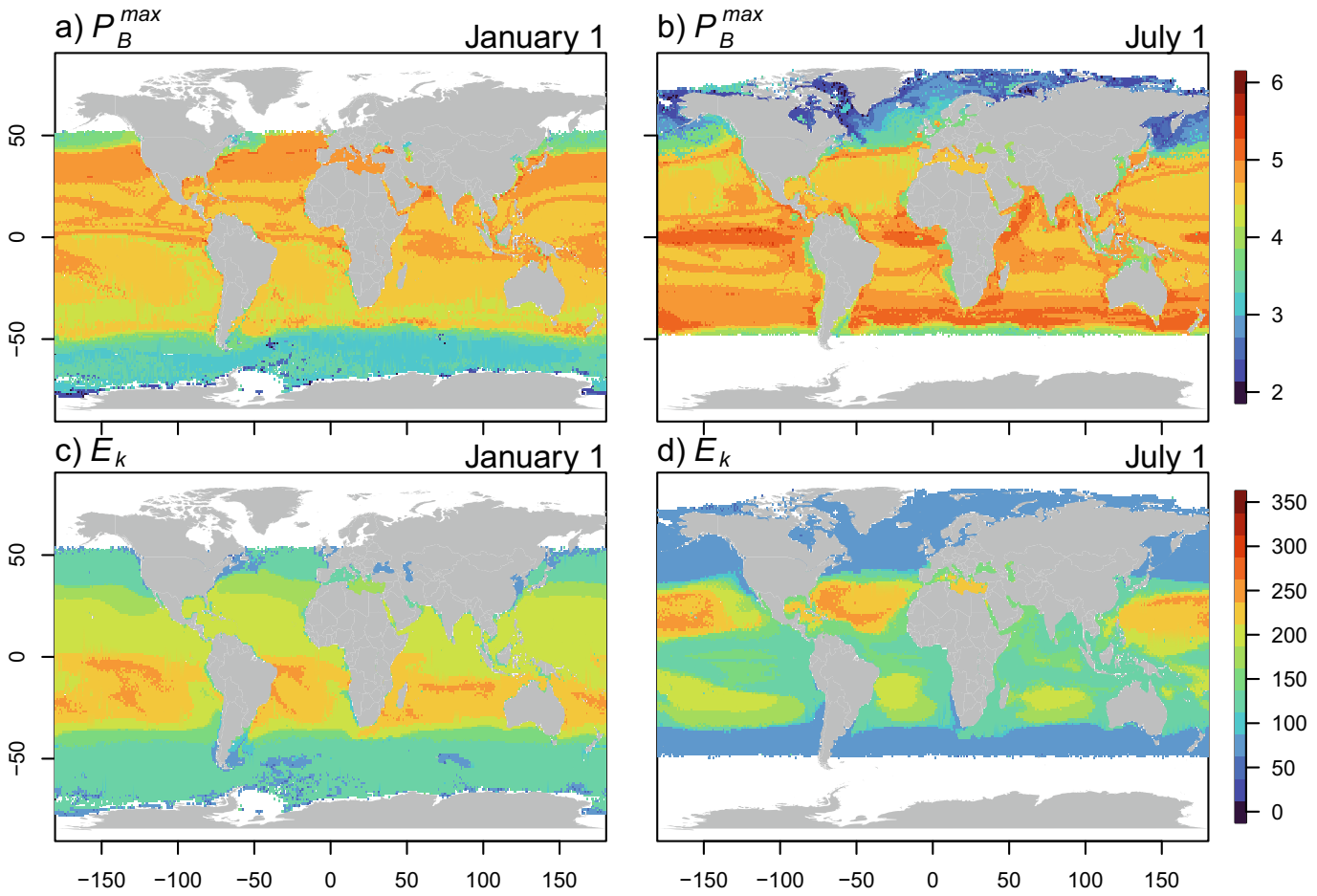

Fig. 5. Global prediction of PI parameters. Predictions for P^B_{max} (mg C (mg Chl *a*)$^{-1}$ h^{-1}) are given in (**a**) and (**b**) for 01 January and 01 July, respectively. Predictions for E_k (μmol photons m^{-2} s^{-1}) are given in (**c**) and (**d**) for 01 January and 01 July, respectively. Environmental covariates are averaged over the preceding 25 d for P^B_{max} and over the preceding 30 d for E_k following results presented in Fig. 2.

While global predictions of PI parameters are extrapolations, we suggest that largescale patterns of predicted PI parameters may indicate effects of regional light limitation. Surface nutrients are highest in winter but insufficient light limits growth. PI experiments under nutrient replete wintertime conditions could release communities from light limitation at high experimental irradiance which would drive high measured rates of photosynthesis and therefore P^B_{max}. These extrapolated patterns warrant further examination and validation to better understand the role of nutrient limitation in driving rates of photosynthesis observed in PI experiments. We find that E_k, on the other hand, is consistently maximized within the oligotrophic gyres with high surface PAR and weak light attenuation in the water column. Predicted patterns are broadly consistent with previous empirically-derived seasonal maps of P^B_{max} and E_k (Kulk et al. 2020, 2021) and regional analyses demonstrating relationships with nutrient availability (Platt et al. 1992). We also note that covariate relationships may differ from what is expected based on single species PI curves (e.g., Cullen et al. 1992) since community-level PI curves include seasonal changes in community composition.

Our findings have implications for the development of satellite-based primary productivity models. Many current algorithms implicitly assume instantaneous acclimation of community-level PI relationships. For example, the vertically generalized productivity model (Behrenfeld and Falkowski 1997) assumes an instantaneous relationship between temperature and a parameter related to P^B_{max} (P^B_{opt}), while non-chlorophyll–based algorithms assume analogous relationships between growth rate, light, and temperature (Behrenfeld et al. 2005; Silsbe et al. 2016). Our findings demonstrate that predictability emerges on longer timescales than considered in these models. Theoretically, these timescales may induce error from nonlinear temporal averaging of the PI curve (Bernhardt et al. 2018). We suggest that timescales of weeks to months be considered when

constructing satellite-based models that take PAR, SST, and Chl *a* as inputs. We also note however that photosynthesis rates inferred from short-term PI experiments may have a complex relationship to time-integrated rates of net primary productivity due to turnover timescales of internal carbon pools of phytoplankton (Halsey and Jones 2015). Photosynthate produced on hourly timescales in nutrient replete conditions may be allocated to high turnover polysaccharide pools rather than protein or structural biomass which impacts the time-integrated balance of gross production and respiration (Halsey and Jones 2015; but see Platt and Sathyendranath 1988). We further note that the Lagrangian transport of phytoplankton means that satellite-based Eulerian characterizations of environmental histories only approximate the environment experienced by individual organisms. More advanced Lagrangian backtracking of environmental histories is suggested for future work (Sebille et al. 2018). Finally, satellite data unavoidably contained missing values. Multi-sensor composite products used here limited the number of missing values for SST, PAR, and Chl *a* to 0%, 15% and 66%, respectively, on daily timescales. We used daily observations best match the timescale of phytoplankton growth.

In closing, we have shown predictability of in situ PI parameters as a function satellite-observed environmental covariates. Predictability was maximized at roughly monthly timescales, suggesting that community turnover timescales are important for capturing PI parameter variability. These results help us understand how phytoplankton communities respond to environmental variability and will improve satellite models of marine primary production.

Author Contributions

All authors conceived of the work, assessed results, and wrote the paper. Gregory L. Britten and Bror Jönsson performed data processing. Gregory L. Britten performed analyses and produced figures.

Acknowledgments

This work was carried out under the auspices of the Simons Collaboration on Computational Biogeochemical Modeling of Marine Ecosystems (CBIOMES). Gregory L. Britten acknowledges funding from the Simons Foundation (645921) and from the Woods Hole Oceanographic Institution (27100177). Shubha Sathyendranath, Gemma Kulk, and Heather A. Bouman acknowledge funding from the European Space Agency. Bror Jönsson, Gemma Kulk, and Shubha Sathyendranath acknowledge fundings from the Simons Foundation (549947 to Shubha Sathyendranath). This work is a contribution to the activities of the National Centre of Earth Observation of the United Kingdom. Shubha Sathyendranath and Gemma Kulk acknowledge funding support from National Centre of Earth Observation.

Data Availability Statement

All data used in this study are cited in the main text and publicly available with the following DOIs: Ocean Color Climate Change Initiative chlorophyll (https://doi.org/doi:10.5285/5011d22aae5a4671b0cbc7d05c56c4f0); Operational Sea Surface Temperature and Ice Analysis SST (https://doi.org/doi:10.48670/moi-00168); MODIS and SeaWifs PAR (https://doi.org/doi:10.5067/AQUA/MODIS/L3B/PAR/2022.0, https://doi.org/doi:10.5067/ORBVIEW-2/SeaWiFS/L3B/PAR/2022.0). All processed data, including in situ PI parameters and satellite covariates are available in a Dryad repository (https://doi.org/10.5061/dryad.w6m905r1j). All code used to perform analyses and produce figures will be publicly available at: https://github.com/gregbritten/pi_parameters_public.

References

Anderson, S. I., G. Franzè, J. D. Kling, et al. 2022. "The Interactive Effects of Temperature and Nutrients on a Spring Phytoplankton Community." *Limnology and Oceanography* 67: 634–645. https://doi.org/10.1002/lno.12023.

Bates, D. M., M. Maechler, B. Bolker, and S. Walker. 2013. Lme4: Linear Mixed-Effects Models Using Eigen and S4 R Package Version 1.0-5. The R Foundation. https://doi.org/10.32614/cran.package.lme4.

Behrenfeld, M. J., and P. G. Falkowski. 1997. "Photosynthetic Rates Derived From Satellite-Based Chlorophyll Concentration." *Limnology and Oceanography* 42: 1–20. https://doi.org/10.4319/lo.1997.42.1.0001.

Behrenfeld, M. J., E. Boss, D. A. Siegel, and D. M. Shea. 2005. "Carbon-Based Ocean Productivity and Phytoplankton Physiology From Space." *Global Biogeochemical Cycles* 19: 1–14. https://doi.org/10.1029/2004GB002299.

Behrenfeld, M. J., K. H. Halsey, and A. J. Milligan. 2008. "Evolved Physiological Responses of Phytoplankton to Their Integrated Growth Environment." *Philosophical Transactions of the Royal Society B* 363: 2687–2703. https://doi.org/10.1098/rstb.2008.0019.

Bernhardt, J. R., J. M. Sunday, P. L. Thompson, and M. I. O'Connor. 2018. "Nonlinear Averaging of Thermal Experience Predicts Population Growth Rates in a Thermally Variable Environment." *Proceedings of the Royal Society B: Biological Sciences* 285: 20181076. https://doi.org/10.1098/rspb.2018.1076.

Bouman, H. A., T. Platt, M. Doblin, et al. 2018. "Photosynthesis–Irradiance Parameters of Phytoplankton: Synthesis of a Global Data Set." *Earth System Science Data* 10: 251–266. https://doi.org/10.5194/essd-10-251-2018.

Bouman, H., T. Platt, S. Sathyendranath, and V. Stuart. 2005. "Dependence of Light-Saturated Photosynthesis on Temperature and Community Structure." *Deep Sea Research Part*

I: Oceanographic Research Papers 52: 1284–1299. https://doi.org/10.1016/j.dsr.2005.01.008.

Brewin, R. J. W., S. Sathyendranath, T. Hirata, S. J. Lavender, R. M. Barciela, and N. J. Hardman-Mountford. 2010. "A Three-Component Model of Phytoplankton Size Class for the Atlantic Ocean." *Ecological Modelling* 221: 1472–1483. https://doi.org/10.1016/j.ecolmodel.2010.02.014.

Britten, G. L., Y. Mohajerani, L. Primeau, et al. 2021. "Evaluating the Benefits of Bayesian Hierarchical Methods for Analyzing Heterogeneous Environmental Datasets: A Case Study of Marine Organic Carbon Fluxes." *Frontiers in Environmental Science* 9: 1–12. https://doi.org/10.3389/fenvs.2021.491636.

Cullen, J. J., X. Yang, and H. L. MacIntyre. 1992. "Nutrient Limitation of Marine Photosynthesis." In *Primary Productivity and Biogeochemical Cycles in the Sea*. Plenum Press.

Cutler, D. R., T. C. Edwards Jr., K. H. Beard, et al. 2007. "Random Forests for Classification in Ecology." *Ecology* 88: 2783–2792. https://doi.org/10.1890/07-0539.1.

Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer. 2012. "The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) System." *Remote Sensing of Environment* 116: 140–158. https://doi.org/10.1016/j.rse.2010.10.017.

Dutkiewicz, S., P. Cermeno, O. Jahn, et al. 2020. "Dimensions of Marine Phytoplankton Diversity." *Biogeosciences* 17: 609–634. https://doi.org/10.5194/bg-17-609-2020.

Eppley, R. W. 1972. "Temperature and Phytoplankton Growth in the Sea." *Fishery Bulletin* 70: 1063–1085. https://spo.nmfs.noaa.gov/content/temperature-and-phytoplankton-growth-sea.

Frouin, R., J. McPherson, K. Ueyoshi, and B. A. Franz. 2012. "A Time Series of Photosynthetically Available Radiation at the Ocean Surface from SeaWiFS and MODIS Data." *Remote Sensing of the Marine Environment II* 8525: 852519. https://doi.org/10.1117/12.981264.

Gallego, I., and A. Narwani. 2022. "Ecology and Evolution of Competitive Trait Variation in Natural Phytoplankton Communities under Selection." *Ecology Letters* 25: 2397–2409. https://doi.org/10.1111/ele.14103.

Geider, R. J., H. L. MacIntyre, and T. M. Kana. 1998. "A Dynamic Regulatory Model of Phytoplanktonic Acclimation to Light, Nutrients, and Temperature." *Limnology and Oceanography* 43: 679–694. https://doi.org/10.4319/lo.1998.43.4.0679.

Grinsztajn, L., E. Oyallon, and G. Varoquaux. 2022. "Why Do Tree-Based Models Still Outperform Deep Learning on Typical Tabular Data?" *Advances in Neural Information Processing Systems* 35: 1–33. https://doi.org/10.48550/arXiv.2207.08815.

Halsey, K. H., and B. M. Jones. 2015. "Phytoplankton Strategies for Photosynthetic Energy Allocation." *Annual Review of Marine Science* 7: 265–297. https://doi.org/10.1146/annurev-marine-010814-015813.

Hutchinson, G. E. 1961. "The Paradox of the Plankton." *The American Naturalist* 95: 137–145. https://doi.org/10.1086/282171.

Kulk, G., A. F. Carrera, W. M. Balch, et al. 2023. "Carbon-Based-Incubations." In *Ocean Optics & Biogeochemistry Protocols for Satellite Ocean Colour Sensor Validation*. IOCCG.

Kulk, G., T. Platt, J. Dingle, et al. 2020. "Primary Production, an Index of Climate Change in the Ocean: Satellite-Based Estimates Over Two Decades." *Remote Sensing* 12: 1–26. https://doi.org/10.3390/rs12050826.

Kulk, G., T. Platt, J. Dingle, et al. 2021. "Correction: Kulk et al. Primary Production, an Index of Climate Change in the Ocean: Satellite-Based Estimates over Two Decades. *Remote Sens*. 2020, *12*, 826." *Remote Sensing* 13: 3462. https://doi.org/10.3390/rs13173462.

Kulk, G., W. H. van de Poll, R. J. W. Visser, and A. G. J. Buma. 2011. "Distinct Differences in Photoacclimation Potential Between Prokaryotic and Eukaryotic Oceanic Phytoplankton." *Journal of Experimental Marine Biology and Ecology* 398: 63–72. https://doi.org/10.1016/j.jembe.2010.12.011.

Leibold, M. A., J. M. Chase, J. B. Shurin, and A. L. Downing. 1997. "Species Turnover and the Regulation of Trophic Structure." *Annual Review of Ecology and Systematics* 28: 467–494. https://doi.org/10.1146/annurev.ecolsys.28.1.467.

Lewis, M., and J. Smith. 1983. "A Small Volume, Short-Incubation-Time Method for Measurement of Photosynthesis as a Function of Incident Irradiance." *Marine Ecology Progress Series* 13: 99–102. https://doi.org/10.3354/meps013099.

Liaw, A., and M. Wiener. 2002. "Classification and Regression by RandomForest." *R News* 2: 18–22. https://journal.r-project.org/articles/RN-2002-022/.

Liefer, J. D., A. Garg, D. A. Campbell, A. J. Irwin, and Z. V. Finkel. 2018. "Nitrogen Starvation Induces Distinct Photosynthetic Responses and Recovery Dynamics in Diatoms and Prasinophytes." *PLoS One* 13: 1–24. https://doi.org/10.1371/journal.pone.0195705.

Liu, M., X. Liu, J. Jiang, and X. Xia. 2013. "Artificial Neural Network and Random Forest Approaches for Modeling of Sea Surface Salinity." *International Journal of Remote Sensing Applications* 3: 229. https://doi.org/10.14355/ijrsa.2013.0304.08.

Montegut, C. d. B., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone. 2004. "Mixed Layer Depth over the Global Ocean: An Examination of Profile Data and a Profile-Based Climatology." *Journal of Geophysical Research. Oceans* 109: 1–20. https://doi.org/10.1029/2004JC002378.

Platt, T., and A. Jassby. 1976. "The Relationship Between Photosynthesis and Light for Natural Assemblages of Coastal Marine Phytoplankton." *Journal of Phycology* 12: 421–430. https://doi.org/10.1111/j.0022-3646.1976.00421.x.

Platt, T., and S. Sathyendranath. 1988. "Oceanic Primary Production: Estimation by Remote Sensing at Local and Regional Scales." *Science* 241: 1613–1620. https://doi.org/10.1126/science.241.4873.1613.

Platt, T., and S. Sathyendranath. 1993. "Estimators of Primary Production for Interpretation of Remotely Sensed Data on Ocean Color." *Journal of Geophysical Research* 98: 14561–14576. https://doi.org/10.1029/93JC01001.

Platt, T., S. Sathyendranath, O. Ulloa, W. G. Harrison, N. Hoepffner, and J. Goes. 1992. "Nutrient Control of Phytoplankton Photosynthesis in the Western North Atlantic." *Nature* 356: 229–231. https://doi.org/10.1038/356229a0.

Sathyendranath, S., R. J. W. Brewin, C. Brockmann, et al. 2019. "An Ocean-Colour Time Series for Use in Climate Studies: The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI)." *Sensors (Basel)* 19: 1–31. https://doi.org/10.3390/s19194285.

Sathyendranath, S., T. Jackson, and C. Brockmann. 2021. *ESA Ocean Colour Climate Change Initiative: Version 5.0 Data*, 5. NERC EDS Centre for Environmental Data Analysis.

Sathyendranath, S., T. Platt, Ž. Kovač, et al. 2020. "Reconciling Models of Primary Production and Photoacclimation." *Applied Optics* 59: 100–114. https://doi.org/10.1364/AO.386252.

Saux Picart, S., S. Sathyendranath, M. Dowell, T. Moore, and T. Platt. 2014. "Remote Sensing of Assimilation Number for Marine Phytoplankton." *Remote Sensing of Environment* 146: 87–96. https://doi.org/10.1016/j.rse.2013.10.032.

Sebille, E. V., E. van Sebille, S. M. Griffies, et al. 2018. "Lagrangian Ocean Analysis : Fundamentals and Practices." *Ocean Modelling* 121: 49–75. https://doi.org/10.1016/j.ocemod.2017.11.008.

Silsbe, G., M. Behrenfeld, K. Halsey, A. Milligan, and T. Westberry. 2016. "The CAFE Model: A Net Production Model for Global Ocean Phytoplankton." *Global Biogeochemical Cycles* 30: 1756–1777. https://doi.org/10.1002/2016GB005521.

Taboada, F. G., A. D. Barton, C. A. Stock, J. Dunne, and J. G. John. 2019. "Seasonal to Interannual Predictability of Oceanic Net Primary Production Inferred from Satellite Observations." *Progress in Oceanography* 170: 28–39. https://doi.org/10.1016/j.pocean.2018.10.010.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Submitted 10 September 2024
Revised 14 August 2025
Accepted 24 August 2025