A machine learning approach to produce a continuous solar-induced chlorophyll fluorescence dataset for understanding Ocean productivity

Nima Madani[1,2], Nicholas C. Parazoo[2], Manfredi Manizza[3], Abhishek Chatterjee[2], Dustin Carroll[4], Dimitris Menemenlis[2], Vincent Le Fouest[5], Atsushi Matsuoka[6], Kelly M. Luis[2], Camila Serra Pompei[7], Charles E. Miller[2]

[1]UCLA Joint Institute for Regional Earth System Science and Engineering. 4242 Young Hall, 607 Charles E. Young Drive East, Los Angeles, CA 90095, USA.
[2]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA.
[3]University of California San Diego, Geosciences Research Division, Scripps Institution of Oceanography 500 Gilman Drive La Jolla, CA 92093, USA.
[4]Moss Landing Marine Laboratories, San José State University, Moss Landing, California, USA
[5]Littoral Environnement et Sociétés, UMR 7266, Université de La Rochelle, La Rochelle, France
[6]UNH Marine Science & Ocean Engineering, Morse Hall, Durham, NH 03824. USA
[7]Center for Climate Change Science, Massachusetts Institute of Technology, Cambridge, MA, USA

Corresponding author: Nima Madani, Email:nima.madani@jpl.nasa.gov

Abstract

Phytoplankton primary production is a crucial component of Arctic Ocean (AO) biogeochemistry, playing a pivotal role in the carbon cycling by supporting higher trophic levels and removing atmospheric carbon dioxide. The advent of satellite observations measuring chlorophyll a concentration (Chl_ a) has yielded unprecedented insights into the distribution of AO phytoplankton, enhancing our ability to assess oceanic productivity. However, the optical properties of AO waters differ significantly from those of lower-latitude waters, and standard Chl_a algorithms perform poorly in the AO. In particular, Chl_a retrievals are challenged by interferences from other marine constituents including higher pigment packaging and higher proportion of light absorption by colored dissolved organic matter. To derive phytoplankton-originating signature as well as mitigate those effects, solar-induced chlorophyll fluorescence (SIF) emerges as a valuable tool for acquiring physiological insights into the direct photosynthetic processes in the AO. In this study, we leverage satellite-based SIF measurements to assess their correlation with a set of predictive factors influencing phytoplankton photosynthesis. We extend the temporal coverage of AO SIF data to cover the period 2004 - 2020. This novel dataset offers a pathway to monitor the physiological interactions of phytoplankton with changes in climate, promising to significantly improve our understanding of the Arctic water's productivity. The application of this data is expected to provide insights into how phytoplankton respond to shifts in environmental changes, contributing to a more nuanced understanding of their role in High-Latitude Northern Oceans ecosystems.

Key Points

- We extrapolated red SIF over the period of 2004-2020 using a set of predictive variables influencing photosynthesis over the Arctic Ocean.
- The reconstructed SIF data demonstrates a strong correlation with independent data records.

- Data produced is expected to provide a new insight into assessment of Arctic Ocean productivity.

Plain Language Summary

Phytoplankton communities, via means of photosynthesis, play a crucial role in the global carbon cycle by transforming carbon dioxide into organic matter. Recognizing the importance of ocean productivity is essential for effectively managing and conserving marine ecosystems, promoting sustainable fisheries, and comprehending the broader ramifications of climate change on the world's oceans. Alterations in ocean productivity, especially shifts in the abundance and composition of phytoplankton, can serve as early indicators of the health of aquatic ecosystems. While satellite observations have provided an unprecedented overview of phytoplankton distribution by estimating chlorophyll concentrations over oceans, uncertainties persist regarding the accurate estimation of the total photosynthetic activity of organisms in the ocean. Recently, the TROPOMI satellite instrument has made solar-induced chlorophyll fluorescence (SIF) data available, offering another metric for understanding photosynthetic activity. However, the short latency of the data record makes it challenging to assess the impact of rapid climate change in the Arctic domain. In this paper, we employ a modeling framework to extend SIF data over a more extended period, facilitating a more comprehensive assessment of ocean productivity.

1- Introduction

Increasing atmospheric growth rate of CO_2 represents a higher proportion of fossil fuel emission relative to natural compensation and negative feedback of terrestrial and aquatic ecosystems through photosynthesis. While global ocean productivity is estimated to be ~50–60 Pg yr^{-1} (Johnson and Bif 2021, Buitenhuis *et al* 2013), oceans are responsible for an annual carbon sink of 2.9±0.4 Pg (Friedlingstein *et al* 2023). Oceanic primary productivity model estimates rely on satellite estimation of phytoplankton biomass (mg C m^{-3}) and the phytoplankton growth rate (μ, d^{-1}), integrated over the euphotic depth (Westberry *et al* 2008, Silsbe *et al* 2016, Behrenfeld and Falkowski 1997). Apart from limitation of the models, there are uncertainties arise from satellite observed reflectance band employed to retrieve chlorophyll-a (Chl_a) concentration estimates that rely on empirical to semi-analytical relationships derived from the correlations between in-situ measurements and satellite reflectance bands in the blue-to-green region of the visible spectrum (Hu *et al* 2019, 2012, Li *et al* 2023). These algorithms have been implemented by multiple satellite instruments (e.g. CZCS; SeaWiFS and MODIS) since1978 and have been used to estimate photosynthetic rates from estimated chlorophyll concentration estimates (Behrenfeld and Falkowski 1997).

Rapid climate change is actively influencing the Arctic Ocean's unique ecosystem, resulting in rapid alterations in Chl_a concentration and spatial distribution (Bouman *et al* 2020, Lewis *et al* 2020). The retreat of sea- ice cover and subsequent changes in light availability and nutrient cycling have complex and sometimes opposing effects on chlorophyll levels (Dvoretsky *et al* 2023, Castagno *et al* 2023). Warmer temperatures and melting sea ice have increased light availability and lengthened phytoplankton growing season, even promoting a second, late season bloom in some locations (Zhao *et al* 2022, Ardyna *et al* 2014, Manizza *et al* 2023). Changes in surface temperature, nutrient and light availability, salinity, wind stress, and increased freshwater input from melting ice and increased river runoff are among the factors that could affect phytoplankton photosynthesis and thus oceanic NPP (Ko *et al* 2022, Singh *et al* 2023).

Satellite measurements of Chl_a provide a valuable and efficient tool for monitoring ocean chlorophyll content, spatial and temporal distribution patterns, and marine primary productivity. However, accuracy of ocean color algorithms is also impacted by high concentration of colored dissolved organic matter and pigment packaging effect in the AO which interferes with chlorophyll retrievals (Matsuoka *et al* 2017, Cota *et al* 2003). Additionally, the frequent presence of surface chlorophyll in the AO as well as a subsurface chlorophyll maximum (SCM) caused by sea ice melt can be a source of pixel contamination in the retrieval of satellite chlorophyll measurements, which potentially increases the uncertainty in primary productivity estimates (Bélanger *et al* 2007, Arrigo and Van Dijken 2011, Bouman *et al* 2020, Lee *et al* 2015).

Solar induced chlorophyll fluorescence (SIF) represents a promising alternative to quantify AO productivity without the spectroscopic challenges that hamper chlorophyll retrievals. The SIF signal represents 1-2% of photosynthetically active radiation re-emitted in red to near-infrared spectral range (Köhler *et al* 2020a, Parazoo *et al* 2019, Köhler *et al* 2018). SIF is closely related to terrestrial gross primary productivity (Frankenberg *et al* 2011, Guanter *et al* 2014, Parazoo *et al* 2014). The recent development of red SIF (Köhler et al 2020a) provides an unprecedented opportunity to quantify primary productivity from aquatic ecosystems. Red SIF is derived from Bands 5 and 6 of the TROPOspheric Monitoring Instrument (TROPOMI), encompassing wavelengths of 661–725 nm and 725–775 nm, respectively. Red SIF generally correlates with

MODIS normalized fluorescence line height (nFLH, Behrenfeld et al 2009); however, notable differences are reported for specific regions (Köhler *et al* 2020a). In particular, nFLH retrievals face challenges in regions characterized by high chlorophyll concentrations (Gupana et al 2021); red SIF does not have this limitation.

Despite its advantages, the red SIF record only extends back to 2018, limiting its use in studying long-term change. We overcome this limitation by employing a randomForest (RF) machine learning model to provide continuous red SIF data spanning the period from 2004 to 2020. We compare our AO red SIF product to AO chlorophyll and NFLH products and discuss their similarities and differences. We conclude with observations on the potential for using red SIF to study marine PP trends in the rapidly evolving Arctic ecosystems.

2- Methodology

2.1 Datasets

The TROPOMI sensor onboard the Sentinel-5 satellite provides wavelengths to capture SIF spectra for monitoring terrestrial and aquatic photosynthetic activity. Global retrievals of red wavelength SIF data have the advantage of retrieving photosynthetic information in variable atmospheric conditions with ~5km spatial resolution (Köhler *et al* 2020b). SIF signal is capable of penetrating through cloud and aerosol layer and unlike traditional usage of visible spectral bands, SIF is a photosynthesis by-product and insensitive to ocean color, which provides an unprecedented opportunity to monitor oceanic photosynthetic activity. Despite advantages of oceanic red SIF from TROPOMI, the recent availability of data (from Apr-2018) makes quantifying long-terms trends and anomalies in the aquatic systems challenging.

Recently, machine learning methods have been used to extrapolate and upscale SIF in terrestrial ecosystems using a combination of MODIS reflectance (e.g. CSIF; (Zhang *et al* 2018)) or reflectance and meteorological data (GOSIF; (Li and Xiao 2019). These methods that are very common in remote sensing and ecosystem process analyses and predictions (Jung *et al* 2011, Madani *et al* 2018, Natali *et al* 2019), operate by predicting unavailable data using quantified relationships between observable and explanatory variables.

Here, we generate spatial and temporally coherent red-SIF products beyond the original retrieved data that covers Apr 2018 to Apr 2021. In this approach, we first train monthly RF models using the TROPOMI SIF climatological records over selected predictive variables that are believed to impact quantic vegetation’s photosynthetic activity (Table 1). Subsequently, we predict SIF over the 2004-2020 period using temporal information provided by each of the explanatory variables.

Table 1. List of candidate explanatory variables for prediction of TROPOMI red-SIF over the period of 2004-2020.

Variable	Abbreviation	Spatial resolution	Source
Chlorophyll-a Concentration	Chl_a	Retrieved at 0.05^{o}	(NASA 2014)
Normalized Fluorescence Line Height	nFLH*	Retrieved at 0.05o	(NASA 2014)
Sea Surface Temperature	SST*	Retrieved at 0.05^{o}	(NASA 2014)
Sea Surface Salinity	SSS*	0.25^{o}	(Carroll *et al* 2020)
Meridional Wind Stress	vWind	0.25^{o}	(Carroll *et al* 2020)
Zonal Wind Stress	uWind*	0.25^{o}	(Carroll *et al* 2020)
Surface-ocean U Velocity	U	0.25^{o}	(Carroll *et al* 2020)
Surface-ocean V Velocity	V*	0.25^{o}	(Carroll *et al* 2020)
Distance from Coastal Zones	Distance*	0.05^{o}	(Carroll *et al* 2020)
Aquatic Ecoregions	Ecoregions*	-	(Spalding *et al* 2007)

* Next to the abbreviated variables indicates that they were used in the final model.

Explanatory variables were selected to represent spatial distribution of phytoplankton communities as well as representing biotic, abiotic and physical characteristics of marine ecosystems across the Arctic domain. We obtained MODIS Chl_a and NFLH from Google Earth Engine (Gorelick *et al* 2017), where we calculated monthly means based on 0.05 degree spatial resolution, consistent with TROPOMI SIF data. Marine biophysical data obtained were from ECCO-Darwin data assimilation model estimates (Carroll *et al* 2020). Aquatic ecoregions (Spalding *et al* 2007) were used as a proxy to represent seasonal nutrient cycle and availability across the Arctic ecosystems. Additionally, we calculated Euclidian distance from the coastal zones to represent nutrient transport along the land-ocean continuum. All analysis were performed in R (Core Team 2017) using open source libraries.

2.2 Modeling and validation

We developed a RF machine-learning model (Liaw and Wiener 2002) to examine the factors influencing the seasonal variation in SIF cross our 50-90˚ N study area (Figure 1). Our objective was to attribute the observed SIF patterns to potential underlying biotic and abiotic factors, as outlined in Table 1. For each month, we randomly sampled 70% of the data within the study area and linked it to corresponding information extracted from predictor variables.

Figure 1. Study domain indicating ocean bathymetry (in meters) and major regions.

Decision Trees offer the advantage of capturing both linear and non-linear relationships between responses and target variables by categorizing data through a series of if-else nodes. At each terminal node, the mean value of observations within that region is calculated. We assessed the predictive power of MODIS Chlor_a and NFLH as the main predictive component of the model (Figure 2). MODIS NFLH provided higher proportion of the variance in spatio-temporal correlations with the observed SIF over the study domain compared to Chl_a.

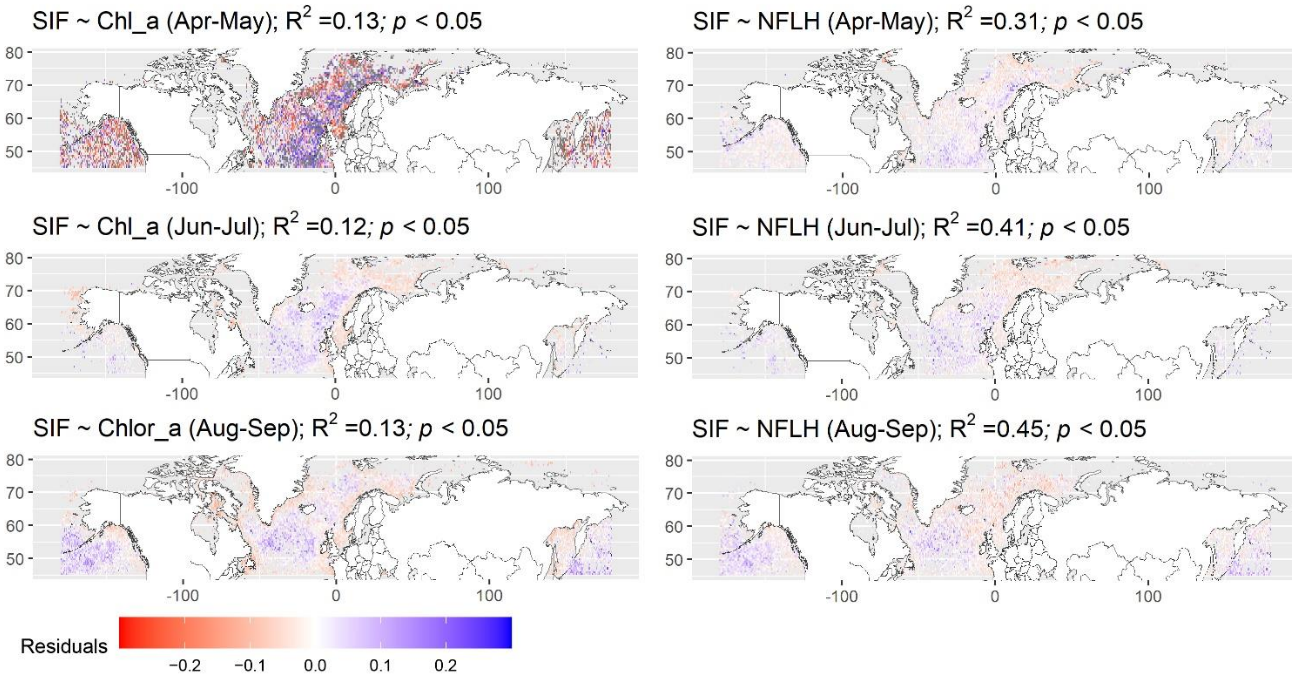

Figure 2. Spatial correlation between observed TROPOMI SIF (Solar-Induced Fluorescence) with MODIS Chlorophyll-a (Chl_a) and Normalized Fluorescence Line Height (NFLH). Color palette indicates residuals of SIF values when linearly associated with Chl_a and NFLH. Warmer and cooler colors indicate regions where SIF may be underestimated and overestimated when using Chl_a and NFLH as linear predictors. A higher proportion of variance (R^2) in SIF is explained by NFHL compared to Chl_a.

A stronger spatial correlation between the observed seasonality SIF relative to MODIS-derived NFLH and Chl_a is also evident over temperate and polar ecoregions as well as selected regions (Figure S1, S2; supplementary materials). To prevent overfitting, we constrained the number of trees to reduce the RMSE of the prediction and maximize the performance of the model. Variable selection involved using a stepwise technique to identify a minimal set of variables adequate for robustly predicting the response variable (Genuer *et al* 2010).

We constructed RF models to evaluate the importance of explanatory variables in elucidating the heterogeneity in SIF trends throughout the seasons. These models quantify an increase in mean squared error (IncMSE) upon permuting each variable, indicating the significance of individual factors in explaining pixel-level heterogeneity relative to monthly SIF observations. Model validation and testing were conducted on the remaining 30% of independent pixels. Additional analyses were performed at regional scales, and the predicted SIF data were compared with in-situ chlorophyll data obtained from the Tara Oceans Polar Circle 2013 cruises (Guidi *et al* 2017).

3- Results and Discussion

The RF model was used to assess the importance of each selected explanatory variable to predict SIF over the Arctic domain. Our model demonstrated an accuracy of 86% in elucidating spatial variability in SIF across independent testing data during the peak of the chlorophyll bloom. Variables such as NFLH, SST, and SSS emerge as among the most important factors in explaining SIF spatial variability (Fig S3). Both SST and SSS served as indicators for spatial variations in optimal environmental conditions for phytoplankton photochemical processes. When used in conjunction with NFLH as a proxy for fluorescence reflectance, they improved SIF prediction significantly.

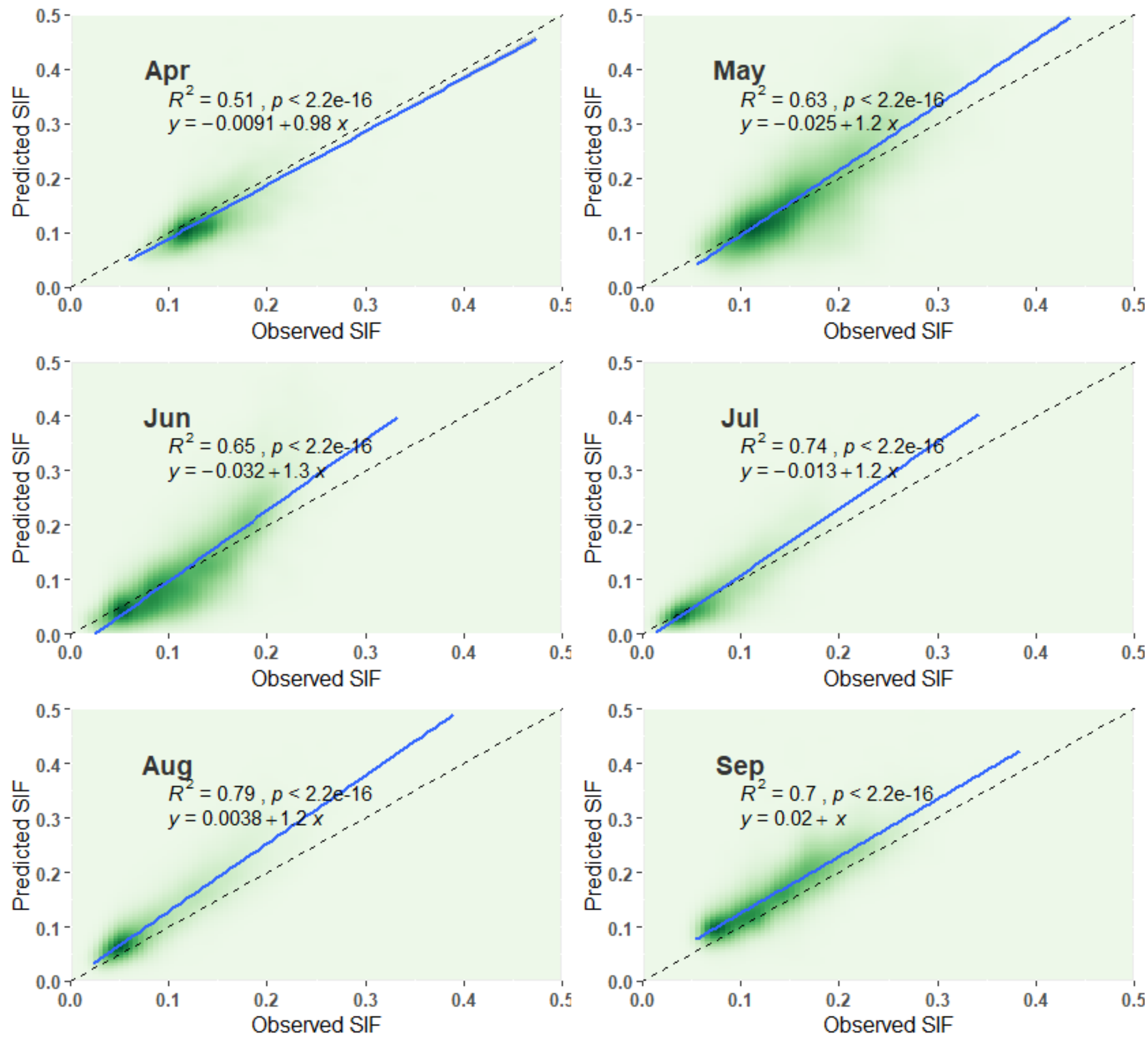

Figure 3. Comparison between spatially explicit monthly climatology in Random Forest predicted SIF relative to the observed SIF over the months of Apr to Sep.

The model results indicate remarkable performance in predicting SIF monthly spatial variability relative to observed SIF data (Figure 3). When comparing the seasonality of observed and predicted SIF over temperature and polar regions, results demonstrate strong and close relationships at regional scales (Figure S4). The performance of the predicted SIF became evident when compared with MODIS Chl_a and NFLH, in conjunction with in-situ observations of Chl_a concentrations from the Tara Oceans Polar Circle 2013 cruises. It is important to highlight that the robust correlations observed between SIF and measured Chl_a are particularly compelling, given that we utilized predicted SIF data for the year 2013, while relying on 2013 observed MODIS chlorophyll-a and NFLH data (Figure 4). The spatial correlation between predicted SIF and in situ observations were 28% stronger than was indicated by MODIS Chl_a data.

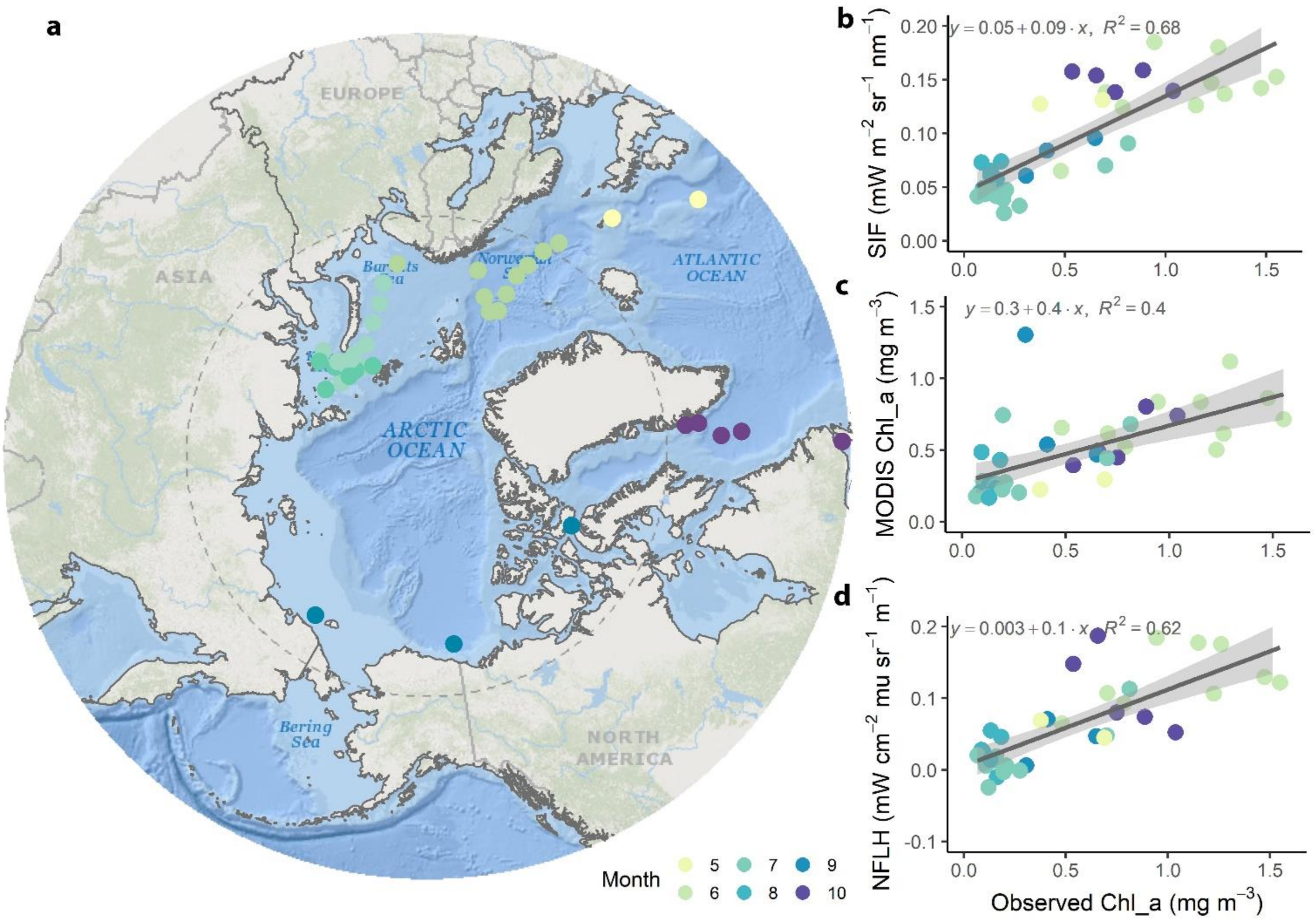

Figure 4. Comparison between predicted SIF (2013), Chl_a and NFLH with HPLC Chl_a observations obtained from Tara Oceans Polar Circle 2013 cruises. **a** location of the retrieved Chl_a data from Ocean Circle cruises. **b** *Predicted SIF over 2013 compared to observed Chl_a.* ***c*** *MODIS derived Chl_a over 2013 compared with observed Chl_a.* **d** *MODIS NFLH compared to observed Chl_a. Colored dots represent the month of retrieved Chl_a observations.*

We conducted a comparative analysis of trends in MODIS Chl_a, NFLH, and reconstructed SIF across temperate and polar ecoregions (Figure 5). In temperate zones, MODIS Chlor_a does not exhibit a significant trend, but there is a noteworthy decline in NFLH. Conversely, SIF indicates a slight decline, particularly evident after 2015, aligning with the observed NFLH trends. In the polar regions, Chl_a exhibits a significant increasing trend, while the NFLH trend is not statistically significant. Notably, SIF shows increasing trends, although they are not as pronounced as those observed in Chl_a. The alignment becomes more apparent when examining regional scales and analyzing anomalies in the average SIF over four distinct periods. In this context, anomalies were defined as departure from annual means within specified years of data records in predicted SIF and NFLH. Notably, differences emerged between SIF and NFLH anomalies, particularly from 2008–2011. During this period, predicted SIF exhibited positive anomalies over the Kara and Barents seas, in contrast to NFLH (Figure 6). It should be noted that unlike NFLH, the RF-predicted SIF is trained on a set of predictive variables that influence ocean processes, which can indirectly impact phytoplankton concentration and photochemical processes.

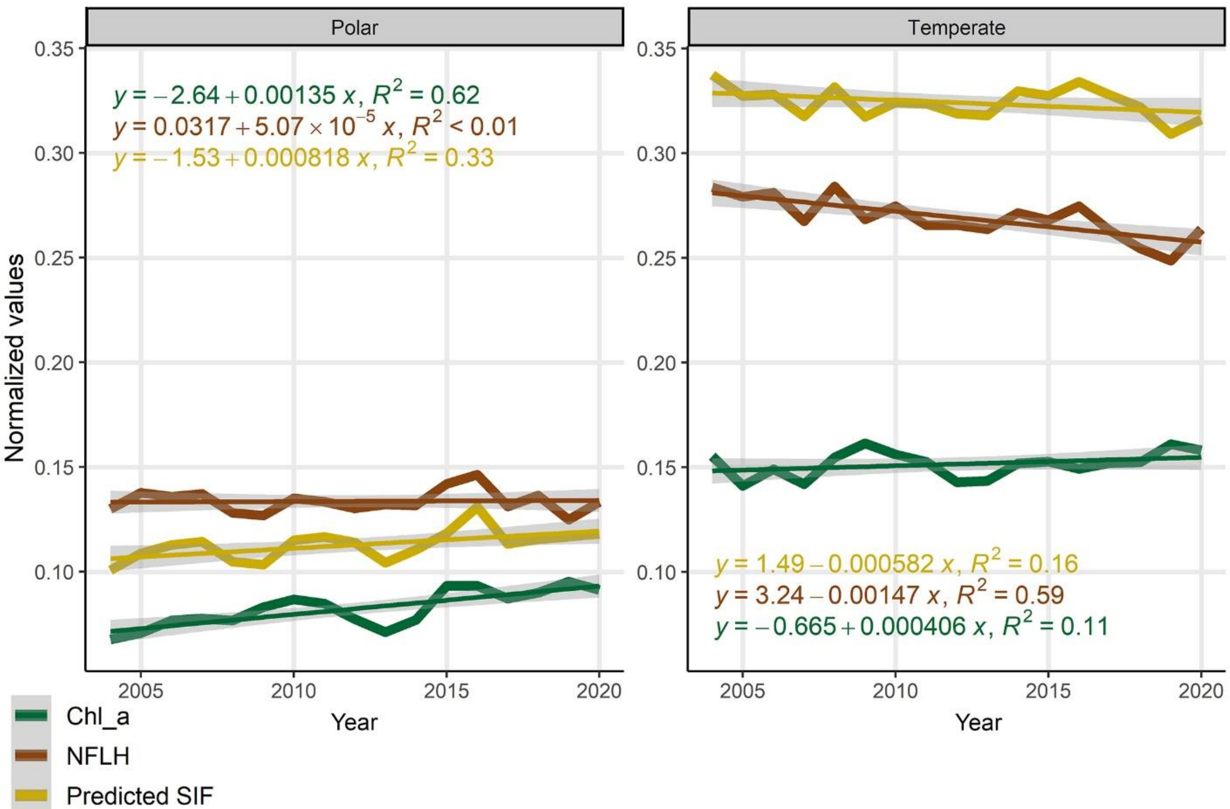

Figure 5. Comparison of observed trends in annual Chl_a, NFLH, and predicted SIF from 2004–2020. Analysis of the observed trends indicates that the overall trajectory of SIF follows NFLH. This similarity is especially noticeable in temperate zones, although there are some variations in specific periods.

However, these disparities in trends and anomalies in SIF concerning Chl_a and NFLH observations underscore the need for cautious utilization of each of these datasets. The observed trends and annual variability may not accurately depict the intricacies of ocean photosynthesis process. The representativeness of SIF in capturing the vertical distribution of phytoplankton remains uncertain at this stage (Köhler *et al* 2020a). Nevertheless, the robust correlations identified between in situ Chl_a observations and predicted SIF are highly encouraging. These findings establish a promising foundation for subsequent analyses and the practical application of the long-term SIF data generated by this research.

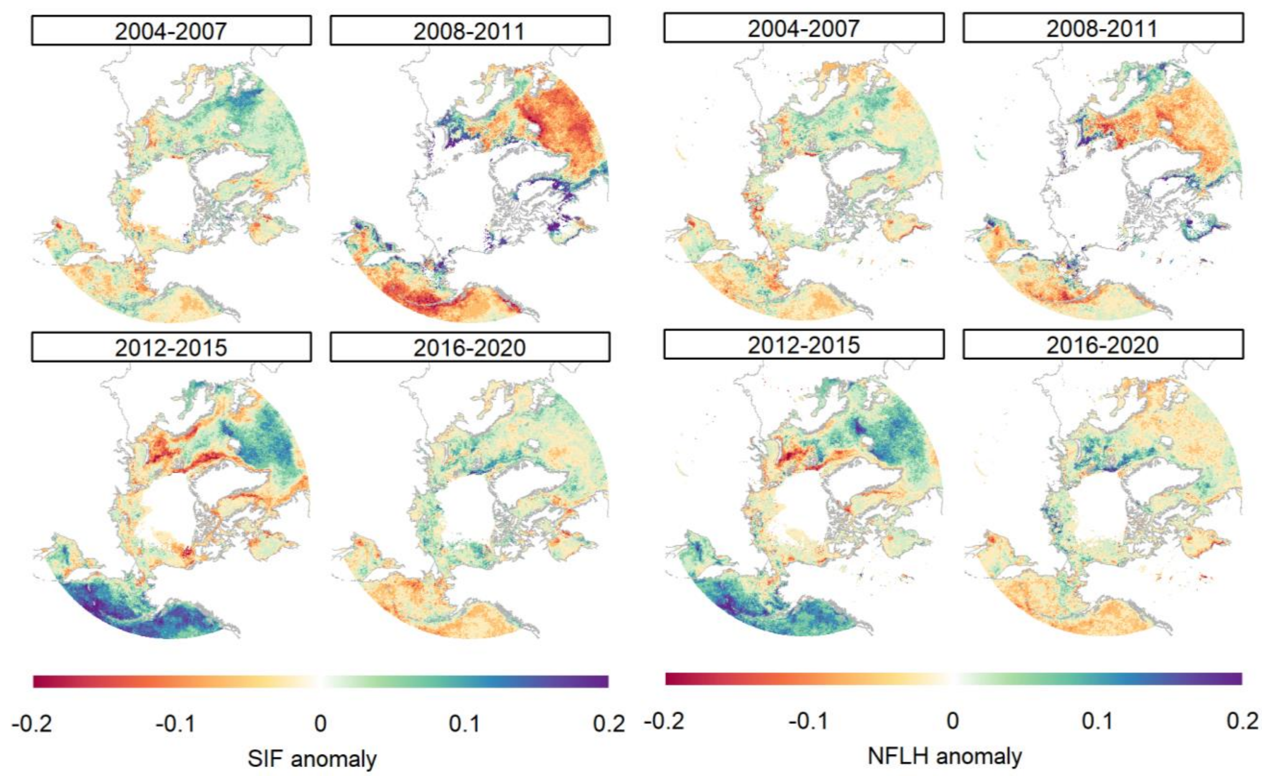

Figure 6. Comparison of annual-mean anomalies in predicted SIF and observed MODIS NFLH over specific time periods. Anomaly in each period is calculated based on the departure from the 2004-2020 means. The analysis of the anomalous data reveals slight regional variations between SIF and NFLH.

4- Conclusions

We presented a pioneering spatially explicit, long-term SIF dataset over an Arctic Ocean domain. Our methodology involved utilizing TROPOMI-observed SIF over the ocean, despite its inherent limitation of a short temporal coverage. To overcome this constraint, we extrapolated the data to a more extended timeframe by leveraging the relationships among explanatory variables that govern the distribution of phytoplankton. Notably, the TROPOMI SIF data exhibited a high degree of consistency with MODIS NFLH observations. However, disparities in long-term trends and anomalies occur, which warrants focused attention in future studies. Although TROPOMI observations provide the advantage of measuring SIF through optically thin cloud and aerosol layers, the presence of optically-thick clouds introduces measurement artifacts (Köhler *et al* 2020a). Furthermore, our predictive modeling approach entails uncertainties due to the utilization of some predictive variables that are themselves modeled (ocean state estimates). Nevertheless, this represents the initial step in aiding our comprehension of long-term changes in Arctic Ocean ecosystems and the influence of ongoing climate change on ocean productivity and ecosystem dynamics.

Acknowledgments

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Government funding acknowledged. MM acknowledges financial support from NASA via the Inter Disciplinary Science Program (IDS19-0113).

Open Research

All data used in this research are publicly available from the cited literature and the links below:
MODIS data including Chl_a and NFLH are available to download from LP DACC data pool:
https://lpdaac.usgs.gov/tools/data-pool/
TROPOMI SIF data are available on Caltech Data repository:
https://data.caltech.edu/records/8hm1f-w5492
ECCO-Darwin model fields are available at:
https://data.nas.nasa.gov/ecco
All analysis were performed in R using open-source software packages:
https://www.R-project.org
R software codes are available on:
https://github.com/MadaniN/Ocean-SIF
Dataset produced by this research are available on ORNL DAAC:
Temporary link: [https://shorturl.at/fqEP1]

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

References

Ardyna M, Babin M, Gosselin M, Devred E, Rainville L and Tremblay J-É 2014 Recent Arctic Ocean sea ice loss triggers novel fall phytoplankton blooms *Geophys. Res. Lett.* **41** 6207–12

Arrigo K R and Van Dijken G L 2011 Secular trends in Arctic Ocean net primary production *J. Geophys. Res. Ocean.* **116** 1–15

Arrigo K R, Matrai P A and Dijken G L Van 2011 Primary productivity in the Arctic Ocean: Impacts of complex optical properties and subsurface chlorophyll maxima on large scale estimates *J. Geophys. Res.* **116** 1–15

Behrenfeld M J and Falkowski P G 1997 Photosynthetic rates derived from satellitebased chlorophyll concentration *Limnol. Oceanogr.* **41** 1–20

Behrenfeld M J, Westberry T K, Boss E S, O'Malley R T, Siegel D A, Wiggert J D, Franz B A, McClain C R, Feldman G C, Doney S C, Moore J K, Dall'Olmo G, Milligan A J, Lima I and Mahowald N 2009 Satellite-detected fluorescence reveals global physiology of ocean phytoplankton *Biogeosciences* **6** 779–94

Bélanger S, Ehn J K and Babin M 2007 Impact of sea ice on the retrieval of water-leaving reflectance, chlorophyll a concentration and inherent optical properties from satellite ocean color data *Remote Sens. Environ.* **111** 51–68

Bouman H A, Jackson T, Sathyendranath S and Platt T 2020 Vertical structure in chlorophyll profiles: Influence on primary production in the Arctic Ocean: Vertical Structure in Arctic Chlorophyll *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **378**

Buitenhuis E T, Hashioka T and Quéré C Le 2013 Combined constraints on global ocean primary production using observations and models *Global Biogeochem. Cycles* **27** 847–58

Carroll D, Menemenlis D, Adkins J F, Bowman K W, Brix H, Dutkiewicz S, Fenty I, Gierach M M, Hill C, Jahn O, Landschützer P, Lauderdale J M, Liu J, Manizza M, Naviaux J D, Rödenbeck C, Schimel D S, Van der Stocken T and Zhang H 2020 The ECCO-Darwin Data-Assimilative Global Ocean Biogeochemistry Model: Estimates of Seasonal to Multidecadal Surface Ocean pCO2 and Air-Sea CO2 Flux *J. Adv. Model. Earth Syst.* **12**

Castagno A P, Wagner T J W, Cape M R, Lester C W, Bailey E, Alves-de-Souza C, York R A and Fleming A H 2023 Increased sea ice melt as a driver of enhanced Arctic phytoplankton blooming *Glob. Chang. Biol.* **29** 5087–98

Core Team R 2017 R: a Language and Environment for Statistical computing

Cota G F, Harrison W G, Platt T, Sathyendranath S and Stuart V 2003 Bio-optical properties of the Labrador Sea *J. Geophys. Res. Ocean.* **108**

Dvoretsky V G, Vodopianova V V. and Bulavina A S 2023 Effects of Climate Change on Chlorophyll a in the Barents Sea: A Long-Term Assessment *Biology (Basel).* **12**

Frankenberg C, Fisher J B, Worden J, Badgley G, Saatchi S S, Lee J E, Toon G C, Butz A, Jung M, Kuze A and Yokota T 2011 New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity *Geophys. Res. Lett.* **38** 1–6

Friedlingstein P, O'Sullivan M, Jones M W, Andrew R M, Bakker D C E, Hauck J, Landschützer P, Le Quéré C, Luijkx I T, Peters G P, Peters W, Pongratz J, Schwingshackl C, Sitch S, Canadell J G, Ciais P, Jackson R B, Alin S R, Anthoni P, Barbero L, Bates N R, Becker M, Bellouin N, Decharme B, Bopp L, Brasika I B M, Cadule P, Chamberlain M A, Chandra N, Chau T-T-T, Chevallier F, Chini L P, Cronin M, Dou X, Enyo K, Evans W, Falk S, Feely R A, Feng L, Ford D J, Gasser T, Ghattas J, Gkritzalis T, Grassi G, Gregor L, Gruber N, Gürses Ö, Harris I, Hefner M, Heinke J, Houghton R A, Hurtt G C, Iida Y, Ilyina T, Jacobson A R, Jain A, Jarníková T, Jersild A, Jiang F, Jin Z, Joos F, Kato E, Keeling R F, Kennedy D, Klein Goldewijk K, Knauer J, Korsbakken J I, Körtzinger A, Lan X, Lefèvre N, Li H, Liu J, Liu Z, Ma L, Marland G, Mayot N, McGuire P C, McKinley G A, Meyer G, Morgan E J, Munro D R, Nakaoka S-I, Niwa Y, O'Brien K M, Olsen A, Omar A M, Ono T, Paulsen M, Pierrot D, Pocock K, Poulter B, Powis C M, Rehder G, Resplandy L, Robertson E, Rödenbeck C, Rosan T M, Schwinger J, et al 2023 Global Carbon Budget 2023 *Earth Syst. Sci. Data* **15** 5301–69

Genuer R, Poggi J M and Tuleau-Malot C 2010 Variable selection using random forests *Pattern Recognit. Lett.* **31** 2225–36 Online: http://dx.doi.org/10.1016/j.patrec.2010.03.014

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D and Moore R 2017 Google Earth Engine: Planetary-scale geospatial analysis for everyone *Remote Sens. Environ.* **202** 18–27

Guanter L, Zhang Y, Jung M, Joiner J, Voigt M, Berry J a., Frankenberg C, Huete A R, Zarco-Tejada P, Lee J-E J-E, Moran M S, Ponce-Campos G, Beer C, Camps-Valls G, Buchmann N, Gianelle D, Klumpp K, Cescatti A, Baker J M and Griffis T J 2014 Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence *Proc. Natl. Acad. Sci.* **111** E1327-33 Online: http://www.ncbi.nlm.nih.gov/pubmed/24706867

Guidi L, Ras J, Claustre H, Pesant S, Tara Oceans Consortium C and Tara Oceans Expedition P 2017 Environmental context of all samples from the Tara Oceans Expedition (2009-2013), about pigment concentrations (HPLC) in the targeted environmental feature *Tara Ocean. Consortium, Coord. Tara Ocean. Exped. Particip. Regist. all samples from Tara Ocean. Exped. (2009-2013). PANGAEA, https//doi.org/10.1594/PANGAEA.875582* Online: https://doi.org/10.1594/PANGAEA.875569

Gupana R S, Odermatt D, Cesana I, Giardino C, Nedbal L and Damm A 2021 Remote sensing of sun-induced chlorophyll-a fluorescence in inland and coastal waters: Current state and future prospects *Remote Sens. Environ.* **262** 112482 Online: https://doi.org/10.1016/j.rse.2021.112482

Hu C, Feng L, Lee Z, Franz B A, Bailey S W, Werdell P J and Proctor C W 2019 Improving Satellite Global Chlorophyll a Data Products Through Algorithm Refinement and Data Recovery *J. Geophys. Res. Ocean.* **124** 1524–43

Hu C, Lee Z and Franz B 2012 Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference *J. Geophys. Res. Ocean.* **117** 1–25

Jiao W, Chang Q and Wang L 2019 The Sensitivity of Satellite Solar-Induced Chlorophyll Fluorescence to Meteorological Drought *Earth's Futur.* **7** 558–73

Johnson K S and Bif M B 2021 Constraint on net primary productivity of the global ocean by Argo oxygen measurements *Nat. Geosci.* **14** 769–74 Online: http://dx.doi.org/10.1038/s41561-021-00807-z

Jung M, Reichstein M, Margolis H A, Cescatti A, Richardson A D, Arain M A, Arneth A, Bernhofer C, Bonal D, Chen J, Gianelle D, Gobron N, Kiely G, Kutsch W, Lasslop G, Law B E, Lindroth A, Merbold L, Montagnani L, Moors E J, Papale D, Sottocornola M, Vaccari F and Williams C 2011 Global patterns of land - atmosphere fluxes of carbon dioxide , latent heat , and sensible heat derived from eddy covariance , satellite , and meteorological observations **116** 1–16

Kim J E, Wang J A, Li Y, Czimczik C I and Randerson J T 2024 Wildfire-induced increases in photosynthesis in boreal forest ecosystems of North America *Glob. Chang. Biol.* **30**

Ko E, Gorbunov M Y, Jung J, Lee Y, Cho K H, Yang E J and Park J 2022 Phytoplankton photophysiology varies depending on nitrogen and light availability at the subsurface chlorophyll maximum in the northern Chukchi Sea *Front. Mar. Sci.* **9** 1–18

Köhler P, , Michael J. Behrenfeld , Jochen Landgraf, Troy S. Magney C F, Behrenfeld M J, Landgraf J, Joiner J, Magney T S, Frankenberg C, Köehler P, Frankenberg C, Magney T S, Guanter L, Joiner J, Landgraf J and Köhler, P., Behrenfeld, M. J., Landgraf, J., Joiner, J., Magney, T. S., & Frankenberg C 2020a Global Retrievals of Solar-Induced Chlorophyll Fluorescence at Red Wavelengths With TROPOMI *Geophys. Res. Lett.* **47** 456–63 Online: http://doi.wiley.com/10.1029/2018GL079031

Köhler P, Behrenfeld M J, Landgraf J, Joiner J, Magney T S and Frankenberg C 2020b Global Retrievals of Solar-Induced Chlorophyll Fluorescence With TROPOMI

Köhler P, Frankenberg C, Magney T S, Guanter L, Joiner J and Landgraf J 2018 Global Retrievals of Solar-Induced Chlorophyll Fluorescence With TROPOMI: First Results and Intersensor Comparison to OCO-2 *Geophys. Res. Lett.* **45** 10,456-10,463

Lee Y J, Matrai P A, Friedrichs M A M, Saba V S, Antoine D, Ardyna M, Asanuma I, Babin M, Bélanger S, Benoît-Gagné M, Devred E, Fernández-Méndez M, Gentili B, Hirawake T, Kang S H, Kameda T, Katlein C, Lee S H, Lee Z, Mélin F, Scardi M, Smyth T J, Tang S, Turpie K R, Waters K J and Westberry T K 2015 An assessment of phytoplankton primary productivity in the Arctic Ocean from satellite ocean color/in situ chlorophyll-a based models *J. Geophys. Res. Ocean.* **120** 6508–41

Lewis K M, Van Dijken G L and Arrigo K R 2020 Changes in phytoplankton concentration now drive increased Arctic Ocean primary production *Science (80-.).* **369** 198–202

Li J, Matsuoka A, Hooker S B, Maritorena S, Pang X and Babin M 2023 A tuned ocean color algorithm for the Arctic Ocean: a solution for waters with high CDM content *Opt. Express* **31** 38494

Li X and Xiao J 2019 A global, 0.05-degree product of solar-induced chlorophyll fluorescence derived from OCO-2, MODIS, and reanalysis data *Remote Sens.* **11**

Liaw A and Wiener M 2002 Classification and Regression by randomForest *R News* **2** 18–22

Liu J, Wennberg P O, Parazoo N C, Yin Y and Frankenberg C 2020 Observational Constraints on the Response of High-Latitude Northern Forests to Warming *AGU Adv.* **1**

Madani N, Kimball J S, Ballantyne A P, Affleck D L R, Van Bodegom P M, Reich P B, Kattge J, Sala A, Nazeri M, Jones M O, Zhao M and Running S W 2018 Future global productivity will be affected by plant trait response to climate *Sci. Rep.* **8**

Madani N, Parazoo N C, Kimball J S, Reichle R H, Chatterjee A, Watts J D, Saatchi S, Liu Z, Endsley A, Tagesson T, Rogers B M, Xu L, Wang J A, Magney T and Miller C E 2021 The Impacts of Climate and Wildfire on Ecosystem Gross Primary Productivity in Alaska *J. Geophys. Res. Biogeosciences* **126** 1–14

Madani N, Parazoo N C and Miller C E 2023 Climate change is enforcing physiological changes in Arctic Ecosystems *Environ. Res. Lett.* **18** 074027

Magney T S, Bowling D R, Logan B A, Grossmann K, Stutz J, Blanken P D, Burns S P, Cheng R, Garcia M A, Köhler P, Lopez S, Parazoo N C, Raczka B, Schimel D and Frankenberg C 2019 Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence *Proc. Natl. Acad. Sci. U. S. A.* **116** 11640–5

Manizza M, Carroll D, Menemenlis D, Zhang H and Miller C E 2023 Modeling the Recent Changes of Phytoplankton Blooms Dynamics in the Arctic Ocean *J. Geophys. Res. Ocean.* **128**

Matsuoka A, Boss E, Babin M, Karp-Boss L, Hafez M, Chekalyuk A, Proctor C W, Werdell P J and Bricaud A 2017 Pan-Arctic optical characteristics of colored dissolved organic matter: Tracing dissolved organic carbon in changing Arctic waters using satellite ocean color data *Remote Sens. Environ.* **200** 89–101 Online: http://dx.doi.org/10.1016/j.rse.2017.08.009

Mohammadi K, Jiang Y and Wang G 2022 Flash drought early warning based on the trajectory of solar-induced chlorophyll fluorescence *Proc. Natl. Acad. Sci. U. S. A.* **119** 1–9

NASA 2014 MODIS-Aqua Ocean Color Data *NASA Goddard Sp. Flight Centre, Ocean Ecol. Lab. Ocean Biol. Process. Gr.*

Natali S M, Watts J D, Rogers B M, Potter S, Ludwig S M, Selbmann A K, Sullivan P F, Abbott B W, Arndt K A, Birch L, Björkman M P, Bloom A A, Celis G, Christensen T R, Christiansen C T, Commane R, Cooper E J, Crill P, Czimczik C, Davydov S, Du J, Egan J E, Elberling B, Euskirchen E S, Friborg T, Genet H, Göckede M, Goodrich J P, Grogan P, Helbig M, Jafarov E E, Jastrow J D, Kalhori A A M, Kim Y, Kimball J S, Kutzbach L, Lara M J, Larsen K S, Lee B Y, Liu Z, Loranty M M, Lund M, Lupascu M, Madani N, Malhotra A, Matamala R, McFarland J, McGuire A D, Michelsen A, Minions C, Oechel W C, Olefeldt D, Parmentier F J W, Pirk N, Poulter B, Quinton W, Rezanezhad F, Risk D, Sachs T, Schaefer K, Schmidt N M, Schuur E A G, Semenchuk P R, Shaver G, Sonnentag O, Starr G, Treat C C, Waldrop M P, Wang Y, Welker J, Wille C, Xu X, Zhang Z, Zhuang Q and Zona D 2019 Large loss of CO2 in winter observed across the northern permafrost region *Nat. Clim. Chang.* **9** 852–7

Parazoo N C, Bowman K, Fisher J B, Frankenberg C, Jones D B A, Cescatti A, Pérez-Priego Ó, Wohlfahrt G and Montagnani L 2014 Terrestrial gross primary production inferred from satellite fluorescence and vegetation models *Glob. Chang. Biol.* **20** 3103–21

Parazoo N C, Frankenberg C, Köhler P, Joiner J, Yoshida Y, Magney T, Sun Y and Yadav V 2019 Towards a Harmonized Long-Term Spaceborne Record of Far-Red Solar-Induced Fluorescence *J. Geophys. Res. Biogeosciences*

Silsbe G M, Behrenfeld M J, Halsey K H, Milligan A J and Westberry T K 2016 The CAFE model: A net production model for global ocean phytoplankton *Global Biogeochem. Cycles*

30 1756–77

Singh A, Fietz S, Thomalla S J, Sanchez N, Ardelan M V., Moreau S, Kauko H M, Fransson A, Chierici M, Samanta S, Mtshali T N, Roychoudhury A N and Ryan-Keogh T J 2023 Absence of photophysiological response to iron addition in autumn phytoplankton in the Antarctic sea-ice zone *Biogeosciences* **20** 3073–91

Spalding M D, Fox H E, Allen G R, Davidson N, Ferdaña Z A, Finlayson M, Halpern B S, Jorge M A, Lombana A, Lourie S A, Martin K D, McManus E, Molnar J, Recchia C A and Robertson J 2007 Marine ecoregions of the world: A bioregionalization of coastal and shelf areas *Bioscience* **57** 573–83

Westberry T, Behrenfeld M J, Siegel D A and Boss E 2008 Carbon-based primary productivity modeling with vertically resolved photoacclimation *Global Biogeochem. Cycles* **22** 1–18

Zhang Y, Joiner J, Alemohammad S H, Zhou S and Gentine P 2018 A global spatially contiguous solar-induced fluorescence (CSIF) dataset using neural networks *Biogeosciences* **15** 5779–800 Online: https://www.biogeosciences.net/15/5779/2018/

Zhao H, Matsuoka A, Manizza M and Winter A 2022 Recent Changes of Phytoplankton Bloom Phenology in the Northern High-Latitude Oceans (2003–2020) *J. Geophys. Res. Ocean.* **127**