2019

A Satellite Assessment of Environmental Controls of Phytoplankton Community Size Structure

Colleen B. Mouw
University of Rhode Island, cmouw@uri.edu

Audrey B. Ciochetto
University of Rhode Island

See next page for additional authors

This is a pre-publication author manuscript of the final, published article.

Citation/Publisher Attribution

Mouw, C. B., Ciochetto, A. B., & Yoder, J. A. (2019). A satellite assessment of environmental controls of phytoplankton community size structure. *Global Biogeochemical Cycles*, 33. https://doi.org/10.1029/2018GB006118 Available at: https://doi.org/10.1029/2018GB006118

Authors
Colleen B. Mouw, Audrey B. Ciochetto, and James A. Yoder

A Satellite Assessment of Environmental Controls of Phytoplankton Community Size Structure

Colleen B. Mouw[1], Audrey B. Ciochetto[1], and James A. Yoder[1]

[1]University of Rhode Island, Graduate School of Oceanography
215 South Ferry Road, Narragansett, RI, 02882, USA

Corresponding author: Colleen Mouw (cmouw@uri.edu), +1-401-874-6506

Key Points:

- Globally, light availability in the water column is the most important parameter for phytoplankton size distribution
- Regionally, phytoplankton size distributions vary, responding to variable light and modes of nutrient delivery
- Cell size is increasing in the cold ocean and the dynamic regions in the warm ocean and declining in the warm ocean

Abstract

Phytoplankton play a key role as the base of the marine food web and a crucial component in the Earth's carbon cycle. There have been a few regional studies that have utilized satellite-estimated phytoplankton functional type products in conjunction with other environmental metrics. Here we expand to a global perspective and ask, what are the physical drivers of phytoplankton composition variability? Using a variety of satellite observed ocean color products and physical properties spanning 1997-2015, we characterize spatial and temporal variability in phytoplankton community size structure in relation to satellite-based physical drivers. We consider the relationships globally and by major thermal regimes (cold and warm), dominant size distribution, and chlorophyll concentration variability. Globally, euphotic depth is the most important parameter driving phytoplankton size variability and also over the majority of the high latitude ocean and the central gyres. In all other regions, size variability is driven by a balance of light and mode of nutrient delivery. We investigated the relationship between size composition and chlorophyll concentration and the physical drivers through correlation analysis. Changes in size composition over time are regionally varying and explained by temporal shifts in the varying physical conditions. These changes in phytoplankton size composition and the varying underlaying physical drivers will ultimately impact carbon export and food web processes in our changing ocean.

1 Introduction

Phytoplankton play a key role as the base of the marine food web and are a crucial component in the earth's carbon cycle. Given this importance, there have been many studies that have characterized phytoplankton distributions from field observations (Boyd et al., 2010; Brun et al., 2015), satellite estimates (McClain et al., 2009; Siegel et al., 2013; Mouw et al., 2017) and modeling (Dutkiewicz et al., 2013; Henson et al., 2017). Previous investigations have suggested alterations in biomass, productivity, and community composition as a result of the changing environment (Bopp et al., 2005; Behrenfeld et al., 2006; Vantrepotte and Melin, 2009; Siegel et al., 2013; Gregg & Rousseaux, 2014; Rousseaux & Gregg, 2015; Dutkiewicz et al., 2013). These changes are regionally variable and complex with multiple drivers at play at once, yet signatures of trends are being detected over the timeframes that satellites have observed (e.g. Henson et al., 2017).

Considering more than just abundance, phytoplankton community composition plays an important role in the intricacies of food web dynamics and their influence on carbon export flux (Guidi et al., 2015; Mouw et al., 2016). It is well understood that small cells are more commonly dominant in stratified, high light conditions, while large cells dominate in well mixed, lower light conditions (Chisholm, 1992). Now with a variety of approaches to estimate phytoplankton community structure from satellite (Mouw et al., 2017; IOCCG, 2014), we have the ability to take a broader look beyond abundance. A variety of approaches have emerged that attempt to discriminate phytoplankton functional types (PFT), which include algorithms that retrieve phytoplankton size classes (PSC), phytoplankton taxonomic composition (PTC), or particle size distribution (PSD). A PFT is an aggregation of phytoplankton, where irrespective of their phylogeny, they share similar biogeochemical or ecological roles. The existing approaches vary in what phytoplankton groupings they retrieve and the underlying mechanisms in which they derive group information (Mouw et al., 2017).

There have been a few studies that have utilized satellite PFT products in conjunction with other environmental metrics. Thus far, these have been regional in scope. In high latitudes, an increase in diatoms were observed during positive phases of local climate indices, suggesting the increases were driven by nutrient supply (Alvian et al., 2013). Coccolithophore abundance was found to correspond to shallow mixed layer with, low wind speed, and increasing sea surface temperature (Sadeghi et al., 2012). In the Indian Ocean, the variance in phytoplankton structure was explained by sea surface height, stratification and sea surface temperature (Brewin et al., 2012). Southern Ocean diatom phenology was found to be driven by the polar front, ice extent and oppositely correlated with El Niño Southern Oscillation and the Southern Annual Mode (Soppa et al., 2016).

Here we expand to a global view to assess the physical drivers of phytoplankton composition variability from satellite products. How have phytoplankton (i.e. chlorophyll concentration and composition) distributions changed over the satellite record? What are the physical drivers of this variability? Over the satellite record, we characterize the relationship between chlorophyll *a* concentration and phytoplankton composition and the variability of phytoplankton distributions to define regions based on persistent patterns. We then determine the dominant physical processes that are responsible for the observed variability and change over time.

2 Materials and Methods

2.1 Imagery and Reanalysis Products

A summary of data products, descriptions and sources, including website links, can be found in Table 1.

2.1.1 Ocean Color Imagery

To allow the greatest length of time of continuous ocean color imagery, we utilized merged imagery products obtained from the Ocean Color Climate Change Initiative (OC-CCI, version 3.0, the latest version at the time of analysis, Grant et al., 2017). OC-CCI has globally merged Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), Medium Resolution Imaging Spectrometer (MERIS), Moderate Resolution Imaging Spectroradiometer (MODIS-Aqua), and Visible Infrared Imaging Radiometer Suite (VIIRS) imagery for a continuous record from 1997 through 2015 (Sathyendranath et al., 2018). We utilized monthly resolution, 4 km products. The OC-CCI products used here include: chlorophyll *a* concentration ([Chl], mg m^{-3}), spectral remote sensing reflectance ($R_{rs}(\lambda)$, sr^{-1}), spectral dissolved and detrital absorption ($a_{dg}(\lambda)$, m^{-1}), and the diffuse attenuation coefficient at 490 nm ($K_d(490)$, m^{-1}). OC-CCI derives $K_d(490)$ from the Lee et al. (2005) algorithm, which is independent of [Chl]. Here, $K_d(490)$ was used to calculate euphotic depth ($z_{eu} = 4.6/K_d(490)$, Morel & Berthon 1989). OC-CCI provides the spectral products at the SeaWiFS bands, by band-shifting the $R_{rs}(\lambda)$ values from MERIS, MODIS and VIIRS to match those of SeaWiFS. The OC-CCI [Chl] product is estimated by a blended combination of the empirical band ratio algorithm (OCx, O'Reilly et al., 1998), and the ocean color index algorithm (OCI) which itself blends the band ratio algorithm and color index (CI, Hu et al., 2012) (Jackson & Grant, 2016). The OC-CCI $a_{dg}(\lambda)$ product is calculated from the quasi semi-analytical algorithm (Lee et al., 2002; 2007).

Satellite PFT algorithms have a variety of phytoplankton product outputs and units (Mouw et al., 2017; IOCCG, 2014). This presents an additional layer of challenge, precluding direct comparison of algorithm performance. Instead, metrics such as phenological cycle, have been used as a way to inter-compare PFT algorithms (Kostadinov et al., 2017). This intercomparison revealed that while PFT algorithms agree across broad scales, they do not all agree under all circumstances. Here we sought to utilize a PFT product that performed near the mean of the phenological metrics (phenological shape, magnitude and month of maximum) that Kostadinov et al. (2017) assessed, as well as that with high validation metrics reported from the original publication (compiled by Mouw et al., 2017). Further, phytoplankton size is one of the best characterized traits structuring food webs due to many ecosystem and physiological processes that are mediated by size such as: nutrient acquisition and utilization, light acquisition, sinking, and grazer interactions (Finkel, 2007; Litchman & Klausmeier, 2008; Finkel et al., 2009; Litchman et al., 2010; Wirtz, 2012). We selected the satellite output from Mouw and Yoder (2010) that estimates phytoplankton size class as percent microplankton (S_{fm}, > 20 μm). The uncertainty metrics of this product are one of the best performing of the PFT algorithms reviewed by Mouw et al. (2017) with r^2=0.6 and RMSE of 12.64. The calculation of S_{fm} requires satellite $R_{rs}(\lambda)$, [Chl], and $a_{dg}(\lambda)$. These are taken from the OC-CCI products described above. This is an absorption-based approach where the chlorophyll-specific absorption spectra for phytoplankton size class extremes, pico- (0.2–2 μm) and microplankton (> 20 μm), are weighted by S_{fm} (Ciotti et al., 2002; Ciotti & Bricaud, 2006). S_{fm} is estimated from a look-up table containing simulated [Chl], $a_{dg}(443)$, $R_{rs}(\lambda)$, and S_{fm}. For a given pixel, satellite-estimated [Chl] and $a_{dg}(443)$, are used to narrow the search space within the look-up table. Of the remaining options, the closest simulated

$R_{rs}(\lambda)$ to the satellite-observed $R_{rs}(\lambda)$ is selected and the associated S_{fm} is assigned. The S_{fm} product processed from OC-CCI imagery is available on PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.892211.

We utilized the vertically generalized primary production model (VGPM) as the net primary productivity (NPP) product (Behrenfeld & Falkowski, 1997). NPP is a function of chlorophyll, available light, and temperature-dependent photosynthetic efficiency. We accessed the monthly, 9 km VGPM NPP imagery for SeaWiFS and MODIS (derived from R2014 processing). Following the recommendations of Mélin (2016) to prevent the introduction of long-term anomalous trends from cross-mission differences, SeaWiFS and MODIS data were merged from a bias-corrected signal. Briefly, monthly climatologies for each pixel were created from years with mission overlap (2003-2007) for SeaWiFS and MODIS records. Gaps of less than two months in the climatology were filled using a spline interpolation. The entire MODIS record was then adjusted by the difference between SeaWiFS and MODIS climatologies:

$$x_{a,corr}(p,m) = x_a(p,m) + x_{s,clim}(p,m) - x_{a,clim}(p,m)$$

where $x_{a,corr}$ is the corrected MODIS signal for a given pixel and month, x_a is the original MODIS signal, and $x_{s,clim}$ and $x_{a,clim}$ are the SeaWiFS and MODIS climatologies, respectively. The final combined record was created by averaging SeaWiFS with the bias-corrected MODIS signal. The pre-MODIS time period includes data from SeaWiFS alone and the post-SeaWiFS time period comprises $x_{a,corr}$ alone.

2.1.2 Physical Data Sets:

Several physical products were utilized to diagnose the drivers of phytoplankton community variability. Satellite and blended products were used to characterize photosynthetically active radiation (PAR, μmol quanta m^{-2} s^{-1}), sea surface temperature (SST, °C) and sea-level anomaly (SLA, m). PAR is the quantum energy flux from the sun between 400 and 700 nm and is a standard product hosted on the NASA Ocean Color web (https://oceancolor.gsfc.nasa.gov/) at monthly, 9 km resolution. PAR from SeaWiFS and MODIS missions (R2014 processing) were merged following Mélin (2016) as described above. The Group for High Resolution Sea Surface Temperature (GHRSST) retrieves SST products that are hosted by the National Oceanographic Data Center (NODC). We utilized the Level 4 global product, which provides gap-free data at quarter-degree, daily resolution by combining *in situ* sensors with satellite products from the Advanced Very High-Resolution Radiometer (AVHRR) Pathfinder missions (Version 5 processing) (www.ghrsst.org). SLA represents the difference in sea-level height from a reference period. The SSALTO/Data Unification and Altimeter Combination System (DUACS) hosted by Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO) is a multi-sensor satellite product derived from seven international satellite missions (Saral/AltiKa, Jason-1 and -2, Cryosat-2, Envisat, ERS-1 and -2, GFO and HY-2A). SLA is retrieved in quarter-degree, monthly resolution relative to the twenty-year mean profile from 1993-2012. The seasonal cycle is not removed from the dataset.

Reanalysis data products were used to retrieve net total heat flux (Q_{net}, W m^{-2}), mixed layer depth (MLD, m) and stratification index ($\Delta\rho_{200}$, kg m^{-3}). Net total heat flux is used to determine if the ocean is a source or sink for heat energy. The National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) provide reanalysis products generated from a variety of satellite, airborne and *in situ* platforms, for latent and sensible heat fluxes on a T62 Gaussian grid with monthly resolution (Kalnay et al., 1996). Surface fluxes for net latent heat flux, net longwave radiation, net shortwave radiation and sensible heat flux were summed to retrieve Q_{net}. We utilized MLD and potential density (σ_θ, kg m^{-3}) from the Simple Ocean Data Assimilation (SODA, version 3.3.1, Carton, et al., 2018), which is forced by the Modern-Era Retrospective analysis for Research and Applications (MERRA-2) dataset. SODA assimilates a variety of *in situ* and satellite observations with a model framework to reconstruct the 3-D physical history of the ocean on a half-degree grid with 50 depth levels ranging from 5 m to 5000 m. For the density-based MLD product, the mixed layer is defined as the depth were density exceeds surface density by 0.03 kg m^{-3}. We retrieved the stratification index from σ_θ as the difference in density between the surface and 200 m (Behrenfeld et al., 2006; Brewin et al., 2012). We also considered bathymetry and wind speed, but did not find compelling relationships, thus they were left out of further analysis and not reported here.

2.2 Data Processing Procedures:

The original downloaded satellite and modeled data products described above have a variety of gridding, spatial resolution and time scales. In order to directly relate one product to another, we created a uniform 1 x 1 degree, gap-free time series for each data product based on the processing steps of Yoder and Kennelly (2003). First, derived products were retrieved from the original dataset (Step 1). Next, data were spatially smoothed or re-gridded to a 1 x 1 degree product (Steps 2 and 3) followed by log transformation where appropriate (Step 4). Finally, data were temporally smoothed and filled in an attempt to produce a gap-free time series (Steps 5 and 6) and quality controlled (Step 7). Details of these processing steps are as follows:

1) Derived products (S_{fm}, z_{eu}, and $\Delta\rho_{200}$) were retrieved, or SeaWiFS and MODIS signals were combined for NPP and PAR products following Mélin (2016).
2) Data were spatially smoothed to 1/4 degree via geometric mean for 4 km and 9 km products, or daily images for SST were combined to create a monthly mean product.
3) Data were spatially smoothed to 1 degree via median filter, or data were re-gridded to 1 degree for Q_{net} product via bi-linear interpolation, or data were re-gridded to 1 degree for MLD and $\Delta\rho_{200}$ products via geometric mean.
4) Products with non-normal distributions were base-10 log transformed (S_{fm}, [Chl], NPP, MLD and $\Delta\rho_{200}$). Normality was assessed by comparing the skewness of the original dataset to log transformed values.
5) Data were temporally smoothed via a 3-month moving average.
6) Gaps of 6 months or less were filled via spline interpolation. Gaps ranging from 5 to 6 months existed in, at most, 5% of global pixels for any given variable and were concentrated at the very northern and southern most edges of the dataset.
7) Outliers greater than 5 standard deviations from the mean were removed.

The final data are 1º by 1º latitude/longitude bins with monthly resolution from January 1998 to March 2015. Only those pixels with 100% data coverage for all data products were used in further analysis; this includes almost all pixels between 50ºN and 50ºS.

2.3 Analysis

Long-term trends and correlation are used to understand temporal and spatial variability of the dataset. To determine the long-term trend, the monthly climatological cycle for each pixel is first removed from the dataset. The remaining linear trend was calculated using the Theil-Sen approach, which is a non-parametric method insensitive to outliers where slope is retrieved as the median of the distribution of slopes between every pair of points in the data set (Barton, Lozier &

William, 2015). Bayes factors (BF_{10}, unitless) were calculated to assess fit significance (Wetzels & Wagenmakers, 2012). Bayes factors represent the likelihood that a slope should be included in the model (slope is non-zero) versus that it should not (slope is zero). For example, BF_{10}=10 means the retrieved slope is ten times more likely to exist than a slope of zero. Here, we only present results with a $BF_{10}>3$, which is considered the cutoff for "substantial evidence" that a slope exists (Wetzels & Wagenmakers, 2012). Correlation between S_{fm} and each of the parameters was determined with Kendall's rank correlation coefficient. Prior to retrieving correlation, products were standardized by subtracting the mean and dividing by the standard deviation to express them on the same scale. Again, only significant correlations, $BF_{10}>3$, are reported, with BF_{10} representing the likelihood that a correlation exists versus that it does not.

Partial least squares regression (PLSR) with 10-fold cross validation is used to determine the relative importance of each parameter to S_{fm} (Wold et al., 2001). Again, data are standardized prior to analysis to express them on the same scale. PLSR combines predictor variables into principle components that are then regressed with S_{fm}. The method allows co-linearity between predictors since they all contribute to forming principle components. VIP (variable influence on projection) scores represent the relative importance of each predictor to S_{fm} variability, while regression coefficients indicate the magnitude and direction of the relationship with S_{fm}. For a given predictor, the VIP score quantifies the cumulative contribution of that predictor to each principle component weighted by the proportion of variance in S_{fm} explained by that component (Mehmood et al., 2012). Here, we consider results with VIP>0.5 to be significant (Wold et al., 2001). Since data were standardized, the relative magnitude of regression coefficients also reflects the importance of each predictor to S_{fm}.

With PLSR, there is the possibility of finding significant correlation by chance, although this likelihood decreases as the dataset gets larger (Clark & Cramer, 1993). We performed a bootstrap test with the global dataset, where the rows of each predictor variable (i.e. latitude/longitude locations and times) were randomly paired with S_{fm} estimates prior to performing a PLSR. We repeated this process 1000 times and none of these cases explained a significant portion of the variance in S_{fm} or had significant VIP scores for any of the randomly ordered predictors. Thus, we are confident that our results with the ordered dataset are more than chance.

To assess confidence in parameter importance, we applied leave-one-predictor-out validation (Martens & Martens, 2000). This method repeats the PLSR analysis n+1 times, where n is the number of predictor variables. The first run includes all predictor variables and subsequent runs remove a single predictor from the dataset each time. Results are presented as the mean coefficients and VIP scores from the resulting distribution with error bars representing minimum and maximum values in the leave-one-predictor-out analysis. This is more appropriate for our large dataset than a jack-knife leave-one-replicate-out approach, where each data point is successively removed from the repeated analysis, as single measurements are not likely to alter final relationships in large data sets (Wold et al., 2001).

3 Results

3.1 Global Analysis:

The great advantage of using satellite products and merging them over time is the ability to explore temporal and spatial variability and the interrelation of these trends. Which parameters display the greatest change over the satellite record and which show significant correlation with S_{fm}? We first explore these relationships at the global scale. The long-term linear trend of the parameters considered is variable across the globe (Figure 1 a-j2). Only S_{fm} and the parameters that have a significant relationship with S_{fm} are presented in Figure 1. The long-term trend of S_{fm}, [Chl], and NPP are nuanced. These parameters are increasing at high latitudes and portions of the subtropics; Z_{eu} is broadly the inverse of these parameters. PAR is decreasing at high latitudes and equatorial regions and increasing in the subtropics. SST is predominately increasing over the majority of the ocean with the exception of some regions of the central gyres and the southern tip of South America. $\Delta\rho_{200}$ generally follows the same spatial patters of z_{eu} with the inverse relationships found for MLD. SLA is increasing over the majority of the ocean. Q_{net} is primarily variable in the equatorial region. The variability of the long-term trends of these parameters will be explored in greater detail in the regional analysis. The correlation between these parameters and S_{fm} is less variable (Figure 1 b-j3). Overwhelmingly, [Chl] and NPP are positively correlated with S_{fm}. However, there are regions where [Chl] and S_{fm} are non- and anti-correlated that will be explored in more detail in subsequent sections. Likewise, Z_{eu} is predominately negatively correlated with S_{fm}, with the non- and anti- correlated regions inverse those of [Chl]. PAR is generally positively correlated at high latitude and equatorial regions and anti-correlated in gyre regions. SST is generally anti-correlated in the warm regions of the ocean and correlated at cold, high latitude regions. $\Delta\rho_{200}$ generally follows the same patters as SST, while MLD and Q_{net} display an inverse relationship to SST. SLA also follows a similar correlation pattern to SST but with weaker correlative relationships.

Globally, which parameters are most important to describing the variability in S_{fm}? We applied PLSR to the global ocean to explore this question. Light availability in the water column, indicated as euphotic depth, is most important to the size distribution of phytoplankton, followed by [Chl], NPP, SST, and PAR (Figure 2). Probability density plots reveal, larger cells are associated with higher [Chl] and NPP, shallower $z_{eu,}$ colder SST and lower PAR. Conversely, smaller cells are associated with deeper z_{eu}, warmer surface waters and higher PAR (Figure S1).

3.2 Regional Analysis:

3.2.1 Size Relationship with Chlorophyll:

Are S_{fm} and [Chl] changing in synchrony? This is an important question to understand as a subset of the satellite PFT algorithms are abundance-based, meaning they estimate PFTs directly from empirical relationships with [Chl] (Mouw et al., 2017). Thus, the relationships used by these approaches should only hold up where [Chl] and phytoplankton composition are strongly correlated. It is therefore of interest to further explore regions of the ocean where strong positive correlation between S_{fm} and [Chl] are not found. From the global analysis above, we identified that S_{fm} and [Chl] are correlated over the majority of the ocean, but there are regions of the ocean where there is little or no correlation between these parameters. How do the physical drivers of S_{fm} variability vary between correlated, non-correlated, and anti-correlated cases? To explore these relationships further, we partition the ocean into regions where S_{fm} and [Chl] are correlated, non-correlated, and anti-correlated. To isolate the impact of temperature, we further differentiate the ocean by warm (≥18ºC) and cold regions (<18ºC), resulting in a total of six regions (Figure 3a). We refer to these as the correlation regions.

In the correlated regions, only z_{eu}, [Chl] and NPP are significant in explaining the variance in S_{fm} (Figure 3 b, c). In these regions, [Chl], S_{fm} and NPP vary together and inversely to z_{eu} (Figure S2). In the anti-correlated regions, z_{eu}, [Chl] and NPP, PAR and Q_{net} are important.

Additionally, in the warm, anti-correlated region, MLD is also important (Figure 3 d, e). In the cold, anti-correlated region, S_{fm}, NPP, PAR and z_{eu} vary together and inversely to [Chl], Q_{net} and MLD, while in the warm anti-correlated region, S_{fm}, varies together with, PAR and z_{eu}, but [Chl], Q_{net}, and NPP vary together with a slight time lag (Figure S2). It should be noted that the cold ocean anti-correlated region is very small and immediately adjacent to the anti- and un-correlated regions found in the southern portion of the South Pacific Ocean. In the uncorrelated regions, z_{eu}, [Chl] and NPP are still significant with the addition of SST and $\Delta\rho_{200}$ in the cold ocean and Q_{net} in the warm ocean. In the uncorrelated cold ocean, S_{fm}, [Chl] and NPP are varying in opposition with each other, while z_{eu} is the inverse of [Chl], and SST and MLD track each other identically. In the warm uncorrelated region, S_{fm}, [Chl], NPP and Q_{net} display a similar temporal pattern offset from each other in time and inverse to z_{eu} (Figure S2). To sum up these relationships, in all regions, z_{eu}, [Chl] and NPP are important and in addition, A) in the correlated regions [Chl], S_{fm} and NPP vary together and inversely to z_e; B) in the anti-correlated regions, Q_{net} and PAR are important factors with addition of MLD in the cold ocean, and C) in the non-correlated regions, Q_{net} is important in the warm ocean and SST and $\Delta\rho_{200}$ in the cold ocean.

3.2.2 Physical Drivers of Phytoplankton Size:

For the majority of our analysis we focus on regions that were determined from a combination of SST, S_{fm} dominance and [Chl] variability (Figure 4). We refer to these as the S_{fm} and [Chl] regions. Within these regions we ask, what are the important physical drivers of S_{fm} variability? To isolate the impact of temperature, mean SST over the time series was used to partition the ocean into warm (≥18ºC) and cold regions (<18ºC). Other studies have used the 15ºC isotherm to delineate warm and cold regions (Behrenfeld et al., 2006; Siegel et al., 2013). However, the 18ºC isotherm corresponded better to the boundaries of S_{fm} dominance (Figure 1f1 and 4a). S_{fm} dominance was determined by assessing the percentage of time spent above the global mean for a given pixel; microplankton were considered dominant if S_{fm} was greater than the global mean for at least 50% of the record (Figure 1a3). The standard deviation of [Chl] ($\sigma_{[Chl]}$, $\mu g\ L^{-1}$) was used to characterize [Chl] variability. Regions were partitioned from the distribution of $\sigma_{[Chl]}$ as greater than the 75th percentile (high variability), between the 25th and 75th percentiles (moderate variability), and less than the 25th percentile (low variability). This results in the possibility of twelve regions. However, not all combinations contained enough pixels to proceed with analysis. We proceeded with nine regions (Figure 4a). These included the low and moderate $\sigma_{[Chl]}$ percentiles (<25th and 25th – 75th) when the phytoplankton community was dominated by small cells for both the warm and cold ocean, the moderate and high $\sigma_{[Chl]}$ percentiles (25th – 75th and >75th) when the phytoplankton community was dominated by large cells for both the warm and cold ocean, and the warm, large-dominated low $\sigma_{[Chl]}$ percentiles (<25th).

PLSR was run on all nine of the S_{fm} and [Chl] regions (Figure 4). To help simplify the variability of the primary drivers of S_{fm} across these regions as determined from the PLSR, we have color coded a map of the regions by the dominant physical drivers (Figure 5), which are referred to as environmental regions, and to view the importance of the parameters driving S_{fm} variability in each of these regions, we have mapped the VIP scores for each parameter considered (Figure 6*)*. The six environmental regions represented in figure 5 correspond to the nine S_{fm} and [Chl] regions in Figure 4 as +SST, MLD, $\Delta\rho_{200}$ – NPP = small, cold, moderate $\sigma_{[Chl]}$; +SST and MLD = large, warm, low $\sigma_{[Chl]}$; +SLA = large, warm, moderate $\sigma_{[Chl]}$; +Q_{net} and MLD = small, warm, moderate $\sigma_{[Chl]}$; +Q_{net}, MLD, PAR – z_{eu} = small, cold, low $\sigma_{[Chl]}$; and z_{eu}, [Chl] and NPP =

4 S_{fm} and [Chl] regions including 1) large, cold, high $\sigma_{[Chl]}$, 2) large, cold, moderate $\sigma_{[Chl]}$, 3) large, warm, high $\sigma_{[Chl]}$, and 4) small, warm, low $\sigma_{[Chl]}$.

For the majority of the high latitude ocean and the central gyres, S_{fm} variability is well explained by only variability in z_{eu}, [Chl], and NPP (Figures 4d, 4e, 4g, 4h, and 5). The importance of z_{eu} points to light availability in the water column being the most important factor in describing the variability of phytoplankton community size composition in these regions. These three parameters have importance across the majority of the ocean, with z_{eu} and [Chl] having the greatest and NPP the least importance at high latitudes (Figure 6). The four regions where only z_{eu}, [Chl] and NPP are significant in explaining S_{fm} variance include, 1-2) cold, large-dominated, high (Figure 4d) and moderate $\sigma_{[Chl]}$ (Figure 4e); 3) warm, small-dominated, low $\sigma_{[Chl]}$ (Figure 4g) and 4) warm, large dominated, high $\sigma_{[Chl]}$ (Figure 4h). In these regions, S_{fm}, [Chl], and NPP varied together and inversely with z_{eu} (Figure S3c, d, f,g). These regions represent the extremes of size and chlorophyll variability: large-dominated with the greatest [Chl] variability in the warm and cold ocean and small-dominated with the lowest [Chl] variability in the warm ocean. (The other region that small-dominated with the lowest [Chl] variability found in the cold ocean, is more complex and discussed later.) The only regions where one of these parameters is not significant in describing S_{fm} variability are found in a small section of the Southern Ocean (Figure 5). NPP is excluded from the cold, small dominated, moderate $\sigma_{[Chl]}$ region where instead just z_{eu} and [Chl], in addition to $\Delta\rho_{200}$, SST and MLD are important (Figure 4b). In this region, SST and $\Delta\rho_{200}$ vary together but inversely from MLD, while [Chl] and z_{eu} vary inversely and S_{fm} and [Chl] are uncorrelted (Figure S3a). Euphotic depth is non-significant in the cold, small dominated, low $\sigma_{[Chl]}$ region with [Chl] and NPP, in addition to PAR, Q_{net} and MLD remaining important (Figure 4c). Here, S_{fm}, NPP and PAR vary together and inversely to [Chl], MLD and Q_{net} (Figure S3b). Overall, PAR and $\Delta\rho_{200}$ are only important in these Southern Ocean regions as well (Figure 6).

All other regions have a balance of the importance of light (indicated by z_{eu} and/or PAR) and a mode of nutrient delivery to the surface ocean (MLD, Q_{net}, SLA and $\Delta\rho_{200}$) beyond z_{eu}, [Chl] and NPP alone (Figure 5). The VIP scores of all other variables are much lower than the ones for z_{eu}, [Chl] and NPP (Figure 6). These include the upwelling and transition regions (adjacent to the gyres or subpolar regions). The only region where one physical parameter is significant, in addition to z_{eu}, [Chl] and NPP, is the warm, large-dominated, moderate $\sigma_{[Chl]}$ region, where SLA is significant (Figure 4i). SLA varies with S_{fm}, [Chl] and NPP and inversely to z_{eu} (Figure S3h). This region is found across the equatorial Pacific and Atlantic indicating a connection to El Niño dynamics, regions of western boundary currents, and fringing some sub-polar regions (Figure 4, 5, and 6). SLA has some of the lowest VIP scores of all parameters (Figure 6). The warm, large-dominated, low $\sigma_{[Chl]}$ region is also found in the equatorial Pacific (Figure 4j and 5). Here MLD and SST are significant in addition to z_{eu}, [Chl] and NPP. S_{fm} variability here is driven by the deepening of MLD, leading to cooling SST, associated with equatorial counter-current dynamics (Figure 4j) that drives the seasonal timing of S_{fm}, [Chl] and NPP peaks (Figure S3i). The mixed layer depth remains an important driver in the remaining region (warm, small-dominated, moderate $\sigma_{[Chl]}$ region) in addition to Q_{net} (Figure 4f), which is found around the outer edges of the gyres (Figures 4 and 5). Here Q_{net} leads the seasonal succession of MLD, S_{fm}, [Chl], NPP, with z_{eu} varying inversely (Figure S3f).

3.3 Temporal Variability:

At the regional scale, which parameters show the greatest change over the satellite record? To answer this, we considered how parameters change over time (Figure 1) within the regions defined in Figure 4a (Figure 7). On average, [Chl] and NPP show similar trends across the various region, increasing nearly everywhere, except the warm, small-dominated, moderate $\sigma_{[Chl]}$ region that captures the equatorial counter current dynamics and additionally for NPP the warm, small-dominated, low $\sigma_{[Chl]}$ region which covers the central gyres. Euphotic depth shows the inverse relationship with NPP. Q_{net} and PAR are mostly neutral across the globe. For PAR the only exceptions are found in the Southern Ocean were the cold, small-dominated, low $\sigma_{[Chl]}$ region show increases and the cold, small-dominated, moderate $\sigma_{[Chl]}$ region showing decreases. These same two regions are slightly decreasing for Q_{net} while; warm, small-dominated low and moderate $\sigma_{[Chl]}$ region is slightly increasing. SST and SLA are increasing across the globe. MLD is increasing across all regions in the cold ocean and slightly decreasing in the warm ocean with the exception of the equatorial counter current region. Likewise, $\Delta\rho_{200}$ is decreasing in the cold ocean and only slightly decreasing in the warm ocean, with the exception of the equatorial counter current region, which is slightly increasing. Changes in S_{fm} are more nuanced. S_{fm} is increasing in the cold ocean, and the dynamic regions in the warm ocean (large-dominated, high and moderate $\sigma_{[Chl]}$). However, S_{fm} is declining in the warm ocean where small cells dominate and $\sigma_{[Chl]}$ is low. Merging these aspects together, in the warm small-dominated ocean, MLD is decreasing, while S_{fm} is decreasing but [Chl] is increasing. This suggests a shift toward greater prevalence of small cells, which are less dependent on nutrients introduced from mixing. In the equatorial counter current region, S_{fm} and [Chl] are decreasing while MLD is increasing suggesting a possible dilution effect.

To provide an example of temporal changes, we selected a transect in the Pacific Ocean (from 8°S to 22°S, along 100°W) (location displayed in Figures 1,3 and 4) that transverses regions that are declining in S_{fm} and [Chl] in the north and increasing in both of these parameters in the south (Figures 1a2, b2). Over the length of the transect from north to south, z_{eu} and MLD deepen and cooling occurs (Figures 1e2, 1i2). The transect transverses three small-dominated, warm ocean regions across all three $\sigma_{[Chl]}$ percentiles (Figure 4). We use Hovmöller plots of the transect (Figure 8) to show the transition over time of these and the other parameters with statistically significant relationships from the PLSR analysis (Figure 4). Interannual variability is evident with an increase in both S_{fm} and [Chl] in the northern reaches of the transect in 2004 and the boundary of smallest percent microplankton contribution and low [Chl] to the south of the transect (Figure 8a, b). To allow comparison between percent microplankton and [Chl] simultaneously, we have coded them on a pixel-by-pixel basis, where if the given value was above the mean it was coded "high" and conversely for "low." In this way, we are able to visualize where size and [Chl] are changing concurrently or oppositely. The period of time that S_{fm} and [Chl] are changing in the same direction (either both increasing or both decreasing) declines over time. Over the timeseries, predominately, [Chl] is remaining above the mean, but the phytoplankton community is shifting toward smaller cells (Figure 8c). NPP, which is a function of [Chl], available light (PAR) and photosynthetic efficiency (which in turn is temperature-dependent) increases over the timeseries (Figures 1c2, 8d), even though the change in [Chl] is variable between north and south (Figures 1b2, 8b), PAR is neutral (Figures 1d2) and SST is cooling (Figures 1f2, 8e). Euphotic depth is deeper to the south and is shallowing over the timeseries (Figures 1e2, 8f), particularly to the south, which is expected with the noted increase in [Chl] in the south. The overall change in heat flux is neutral over the transect with minor interannual variability (Figures 1j2, 8g). MLD is slightly deepening over the timeseries (Figures 1i2, 8h) consistent with the noted SST cooling (Figures 1f2, 8e). Merging these together, at the beginning of the transect we observe conditions where S_{fm} and [Chl] are changing together (either both high or both low, relative to the mean). Over time, there are greater instances of S_{fm} and [Chl] changing on the opposite directions (one is increasing while the other is decreasing and vice versa).

4 Discussion

4.1 Importance of Light Availability

Satellite radiometers sample from roughly the first attenuation length of the water column ($1/K_d$) (Kirk, 1994), which often is much shallower than the mixed layer depth, consequently they are not able to fully capture water column processes associated with mixing. Thus, observing the ocean from satellite biases to a portion of the water column that is most responsive to variable light availability. To some extent, the finding of euphotic depth being the most important parameter is not surprising when considering the sampling method. That being said, it should be noted that light penetration in the water column (euphotic depth) rather than absolute incident light level (PAR) is consistently the more important parameter, with PAR only playing a significant role in a small part of the Southern Ocean.

Cell size is also highly influenced by how pigments are packaged within the cell, known as the packaging effect (Morel & Bricaud, 1981). Small cells have little cellular material between the chloroplast and cell wall making them highly efficient absorbers, resulting in higher magnitude and more peaked absorption. With large cells, light has to penetrate more cellular material to reach the chloroplast after passing through the cell wall, resulting in muted absorption affinity and in some cases shelf-shading (see Figure 7E in Ciotti et al., 2002). The results of the primary importance of euphotic depth in predicting changes in S_{fm} is also not surprising considering these direct relationships between light and pigment packaging within various sized phytoplankton cells.

4.2 Modes of Nutrient Delivery

While the physics of the ocean are highly dynamic and cannot be characterized by one parameter alone, our analysis treats each parameter independently but allows co-variability. Here we aim to describe the possible modes of nutrient delivery from each parameter and also where multiple parameters are coincidently important. First, we revisit the relationship between nutrient delivery and each physical parameter. Relatively cold expressions of SST are indicative of upwelling or wind mixing strong enough to bring deep, cold, nutrient rich water to the surface. When the MLD is deeper than the nutricline, nutrient rich water is mixed into the euphotic zone, impacting phytoplankton growth and composition. SLA is indicative of major upwelling and eddy features. There is also an inverse relationship between sea surface height and nutricline depth where a negative SLA is indicative of isopycnal uplift, and a positive SLA is indicative of deepening of the thermocline and nutricline. The $\Delta\rho_{200}$ is the difference in density between the surface and 200 m and indicative of the stability of the water column. When the ocean is highly stratified, there is reduced vertical mixing, thus a lower likelihood of the entrainment of deep, nutrient rich water into the euphotic zone. There are many fluxes that are summed to total Q_{net} including net latent heat flux, net longwave radiation, net shortwave radiation and sensible heat flux. Thus Q_{net}, is indicative of the magnitude of the temperature differential between the surface ocean and the atmosphere. A higher Q_{net} represents a greater temperature differential indicative of a source of deep, cool, nutrient-rich water.

Many studies have pointed to the importance of stratification in controlling phytoplankton communities (Behrenfeld et al., 2006; Polovina et al., 2008). These studies suggest that with greater warming, stratification will increase, resulting in a decrease in overall production and a shift toward smaller cells. However, this simple explanation cannot be expected to work across the whole ocean and other studies have concluded interannual stratification variability is not large enough to drive a coherent phytoplankton response (Dave & Lozier, 2010). The relationships between phytoplankton response and physical drivers are complicated and depend on a host of variables (Lozier et al., 2011; Barton et al., 2014).

Studies that have considered a host of environmental variables have concluded differing mechanisms of phytoplankton control. Using a global-scale dataset including [Chl], PAR, nutrients, MLD, SST, latitude, longitude and month of the year, Irwin and Finkel (2008) found that SST was the best parameter for explaining [Chl] variability (51%) and that light and nitrate concentration explain 47% of the variation in [Chl]. Further, MLD, surface nitrate, SST, latitude and longitude explain 83% of this variation, thus concluding that either light or macronutrients are often limiting. Thus, most of the variation in [Chl] is explained by bottom-up mechanisms. Conversely, when examining a much longer data record for just the North Atlantic, Barton et al. (2014) found strong seasonal relationships in the physical drivers of various phytoplankton communities, but on interannual to multidecadal timescales, the links to physical changes were much weaker. They attribute this to the year-to-year variability in phytoplankton assemblages being greater than that of the physical drivers, suggesting that top-down controls and/or perhaps changes in ocean circulation may be more important than the physical parameters that they considered (wind speed, heat flux, turbulent kinetic energy generation, SST, stratification, and MLD). Wilson & Coles (2005) conducted a global analysis of [Chl], SST, MLD, thermocline depth, and nutricline depth and identified three regional mechanisms of nutrient delivery to the euphotic zone that were controlling phytoplankton distributions. These include, 1) dynamic thermocline uplift in the tropics, 2) nutrient entrainment at mid-latitudes, and 3) seasonal light limitation at high latitudes. Here we expand this debate to include consideration of satellite-estimated phytoplankton size structure at global scales. We find that our results generally follow the Wilson and Coles (2005) framework, but our regions are less latitudinally bound.

Beginning with the equatorial regions, we find these regions are driven by SST and MLD or SLA. SLA was only found to be significant in the large-dominated, moderate $\sigma_{[Chl]}$, warm ocean which is found in regions primarily impacted by upwelling and El Niño dynamics. Previous studies have documented [Chl] variability in the eastern subtropical Atlantic to be highly correlated to sea surface height variability caused by divergent surface currents leading to vertical upwelling and a shallower thermocline and nutricline (Pastor et al., 2013). Negative SLA (found during La Niña) results in higher [Chl] due to isopycnal uplift, and positive SLA (found during El Niño) results in lower [Chl] due to deepening of the thermocline and nutricline. These dynamics resulting in off-equator [Chl] variability was also described by Wilson and Adamec (2001). [Chl] and S_{fm} are strongly positively corelated in this region, thus driven by the same mechanisms. SST and MLD are the driving parameters in the equatorial counter current region that is large-dominated with low [Chl] variability. Here SST and MLD are proxies to the vertical advection of nutrients to the surface ocean with the dominant nutrient delivery mode in this region (Pastor et al., 2013). SST is indicative of cool, nutrient rich water being brought to the surface and MLD is important due to the variability imposed by El Niño.

Moving to the subtropical ocean, Signorini et al. (2015) investigated the physical drivers of the change in [Chl] and NPP in the subtropical gyres from a satellite perspective. They found downward trends in NPP for all gyres and a downward trend in [Chl] for all gyres except the South Pacific which had a non-significant weak upward trend. They found seasonality in [Chl] was tightly coupled with variability in the MLD confirming vertical mixing is the major driver of phytoplankton photosynthesis in the gyres. In our study, the subtropical gyres correspond to the warm, small-dominated, low $\sigma_{[Chl]}$ region where our results indicate that S_{fm} is driven by z_{eu}, [Chl] and NPP. Our results did not indicate the importance of MLD as in the Signorini et al. (2015) study. However, Signorini et al. (2015) regressed each of the parameters that they considered only against time. Their results suggesting [Chl] and MLD are tightly coupled were deduced from comparing the trend of [Chl] over time with that of the MLD over time. We also find that MLD is generally decreasing over time in the gyres, with the exception of the North Atlantic sub-tropical gyre, but the [Chl] trends over time in these regions are more nuanced. Our analysis considers a longer period of time and simultaneously compares each parameter considered to S_{fm}, therefore reporting the statistical significance of each parameter to predicting S_{fm} rather than only considering which parameters had statistically significant trends over time.

In the mid-latitudes outside of the gyres, we find a variety of different drivers of S_{fm}. There are regions dominated by MLD, SLA, and Q_{net}. We isolated the impact of SST by dividing the ocean into warm and cold regions. Thus, the impact of changing thermal regimes within the warm ocean are now evidenced by heat flux which is indicative of the changing seasonal heating/cooling patterns. In agreement with Wilson and Coles (2005), we would expect MLD to impact the phytoplankton response due to seasonally variable mixing. SLA is found to impact the mid-latitudes in upwelling and western boundary current regions and the dynamical impact of these processes on phytoplankton response have been well documented (Schollart et al., 2004; Clayton et al., 2014).

In agreement with other studies of high latitudes and Wilson and Coles (2005), we find S_{fm} to be driven by light availability in the water column. However, we also find SST, MLD, $\Delta\rho_{200}$, Q_{net}, and PAR to be important in sub-regions of the Southern Ocean indicating phytoplankton composition in this region is not simply light driven. Ardyna et al. (2017) also found the Southern Ocean to be latitudinally and regionally divided. At the coarse latitudinal scale, variability is driven by seasonally varying light availability. At the regional scale, phytoplankton variability is driven by iron supply and local advection processes.

4.3 Relationship between S_{fm} and [Chl]

Overwhelmingly, S_{fm} and [Chl] are changing together driven by z_{eu} and SST in the warm ocean and z_{eu}, SST, $\Delta\rho_{200}$, and MLD in the cold ocean. However, for the regions where S_{fm} and [Chl] are not positively correlated, heat flux is an important variable. When S_{fm} and [Chl] are anti-correlated, PAR is additionally important. Heat flux indicates a temperature differential between the ocean and atmosphere, leading to greater cloud formation and less PAR. Heat flux is a potential indicator of nutrient delivery to the surface ocean, as deep nutrient-rich water would be cold when brought to the surface ocean. In the anticorrelated regions, we see [Chl] increasing when S_{fm} is small. In the uncorrelated regions, we see a timing offset with [Chl] tending to increase prior to an increase in cell size. Physiological compensation is a likely cause. When PAR decreases due to a thermal gradient, leading to cloud formation and when nutrients remain replete, phytoplankton increase their chlorophyll content to more efficiently capture light. Siegel et al. (2013) found biomass changes dominate [Chl] at high latitudes, while physiological processes dominate [Chl] variability in the tropical and sub-tropical regions. However, they also note [Chl] changes in coastal and equatorial upwelling areas within the tropical and sub-tropical regions were dominated by biomass. The majority of the anticorrelated and uncorrelated regions fell in these upwelling dominated areas not identified by Siegel et al. (2013) to be dominated by physiological compensation. However, their study considered annual climatology of the whole SeaWiFS mission, while we investigate at monthly timescales. Thus, we are capturing physiological compensation occurring at seasonal timescales.

4.4 Relation to Other Observational Methods

Field studies have reported mixed results as to the most important drivers of phytoplankton composition. Acevedo-Trejos et al. (2013), using the Atlantic Meridional dataset found temperature and nitrite+nitrate to be the most important variables, with light not playing a significant role in structuring the community. Conversely, Brun et al. (2015), using a globally distributed dataset, found MLD to be the most important environmental parameter followed by temperature and PAR within the mixed layer. In a review of many phytoplankton groups considered separately, Boyd et al. (2010) found that nitrogen was most important for diatoms, PAR was most important for *Phaeocystis antarctica* and picocyanobacteria, and temperature was most important for coccolithophores, nitrogen-fixers and *Prochlorococcus*. Across all of these studies, temperature is a common driving parameter. When running our analysis on the whole globe (i.e. Figure 2), we found SST to be the fourth most important parameter behind, z_{eu}, [Chl] and NPP. The field studies were not using [Chl] and NPP as environmental variables, thus our results are consistent after taking this into account. Yet, light in the water column still remains the most important, most likely due to the sampling method and packaging effects discussed above. We would have liked to include varying nutrients in our study, however, such a product does not exist that is global and varying at monthly timescales. We opted not to use monthly climatology of nutrients as this would have biased the interannual relationships. Instead, we have focused on physical drivers that are indicative of nutrient delivery.

Modeling studies that have the luxury of full depth resolution predict globally integrated primary production will decrease (Bopp et al. 2013) as a result of reduced supply of macro-nutrients to the euphotic zone. However, the response is not uniform across the globe, rather some regions have an increase in productivity due to reduction in light limitation due to increased stratification, and higher growth rates due to increased temperatures (Taucher & Oschlies, 2011; Dutkiewicz et al., 2013). Models also suggest geographical shifts in temperature structure will dramatically change local community composition with a shift toward greater abundance of small cells (Bopp et al., 2005; Marinov et al., 2013; Dutkiewicz et al., 2013), since they require less nutrients than larger cells. In modeling results, reduced nutrient supply was most pronounced on biomass and primary productivity at lower latitudes, with increased growth rates playing a stronger role in nutrient-rich higher latitudes (Dutkiewicz et al., 2013). Our results suggest similar latitudinal variations in productivity and community size shifts, with each increasing at high latitudes and decreasing at low latitudes.

4.5 Temporal Trends

Many studies have been aimed at predicting how phytoplankton has changed over our observational record and how it will change in the future. What is clear is that the ocean has changed and will continue to change (Barton et al., 2016) and this change may be more rapid than estimated just a few years ago (Henson et al., 2017). There will be winners and losers with shifts in geographical temperature structure dramatically changing local phytoplankton community composition (Dutkiewicz et al., 2013). Our analysis shows regional increase and decline in the size of phytoplankton over our observational record. S_{fm} is increasing in the cold ocean, and the dynamic regions of the warm ocean where MLD is increasing. However, S_{fm} is declining in the warm ocean where small cells dominate, [Chl] is increasing but has low variability, and MLD is decreasing. These changes will ultimately impact food web processes (Litchman et al., 2008; 2010) and carbon export (Mouw et al., 2016).

5 Conclusions

Light availability in the water column is the most important parameter for the size distribution of phytoplankton as sampled from a satellite platform. As expected, larger cells are associated with higher [Chl] and NPP, shallower z_{eu}, colder SST and lower PAR. When considering the ocean by major thermal regimes (cold and warm), dominant size distribution, and [Chl] variability, for the majority of the high latitude ocean and the central gyres, S_{fm} variability is well explained by only variability in z_{eu}, [Chl], and NPP indicating light availability drives the phytoplankton community. In all other regions of the ocean there is a balance of the importance of light (indicated by z_{eu} and/or PAR) and mode of nutrient delivery to the surface ocean (MLD, Q_{net}, SLA and $\Delta\rho_{200}$). These results point to regionally varying phytoplankton distributions, responding to variable light and mixing regimes. For the majority of the ocean, S_{fm} and [Chl] are correlated and vary together with NPP and inversely to SST and z_{eu}. There are regions of the ocean where phytoplankton size distribution and [Chl] are not positively correlated. In these regions, Q_{net} becomes important, in addition to z_{eu}, [Chl], NPP. PAR is also important in anti-correlated regions and various modes of mixing (as indicated by MLD, SST and $\Delta\rho_{200}$) are important in the cold ocean.

S_{fm} is increasing in the cold ocean, and the dynamic regions in the warm ocean (large-dominated, high/moderate $\sigma_{[Chl]}$) where MLD is increasing. However, S_{fm} is declining in the warm ocean where small cells dominate, $\sigma_{[Chl]}$ is low, [Chl] is increasing, and MLD is decreasing; suggesting a shift toward greater prevalence of small cells, which are less dependent on nutrients introduced from mixing. In the equatorial counter current region, S_{fm} and [Chl] are decreasing while MLD is increasing, suggesting a possible dilution effect. Temporal change suggests the vulnerability of phytoplankton size distributions in a changing ocean will be regionally varying, ultimately impacting carbon export and food web processes.

Acknowledgments and Data

Funding for this effort was provided by NASA grants NNX13AC34G and NNX16AR47G. This manuscript was improved from the comments of two anonymous reviewers. We gratefully thank all of the data provides including: European Space Agency for the Ocean Colour Climate Change Initiative dataset (http://www.esa-oceancolour-cci.org/), NASA for providing the ocean color products that were utilized by the Ocean Colour Climate Change Initiative dataset and the photosynthetically active radiation imagery product (https://oceancolor.gsfc.nasa.gov/), NCEP/NCAR for hosting the reanalysis products (https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.derived.surfaceflux.html), AVISO for the sea level anomaly product (https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/msla-mean-climatology.html#c10358), Group for High Resolution SST for hosting the SST product (https://data.nodc.noaa.gov/ghrsst/), the University of Maryland for hosting the Simple Ocean Data Assimilation products (http://www.atmos.umd.edu/~ocean/), and Oregon State University for hosting the net productivity imagery

(http://www.science.oregonstate.edu/ocean.productivity/index.php). The percent microplankton imagery can be accessed on PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.892211.

References

Acevedo-Trejos, E., Brandt, G., Merico, A., & Smith, S. L. (2013). Biogeographical patterns of phytoplankton community size structure in the oceans. *Global Ecology and Biogeography*, *22*(9), 1060–1070. http://doi.org/10.1111/geb.12071

Alvain, S., Le Quéré, C., Bopp, L., Racault, M.-F., Beaugrand, G., Dessailly, D., & Buitenhuis, E. T. (2013). Rapid climatic driven shifts of diatoms at high latitudes. *Remote Sensing of Environment*, 132, 195–201. http://doi.org/10.1016/j.rse.2013.01.014

Ardyna, M., Claustre, H., Sallée, J.-B., D'Ovidio, F., Gentili, B., van Dijken, G., et al. (2017). Delineating environmental control of phytoplankton biomass and phenology in the Southern Ocean. *Geophysical Research Letters*, *44*(10), 5016–5024. http://doi.org/10.1002/2016GL072428

Barton, A. D., Irwin, A. J., Finkel, Z. V., & Stock, C. A. (2016). Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. *Proceedings of the National Academy of Sciences*, *113*(11), 2964–2969. http://doi.org/10.1073/pnas.1519080113

Barton, A. D., Lozier, M.S., & Williams, R.G. (2014). Physical controls of variability in North Atlantic phytoplankton communities. *Limnology and Oceanography*, *60*(1), 181–197. http://doi.org/10.1002/lno.10011

Behrenfeld, M., & Falkowski, P. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, 42(1), 1–20

Behrenfeld, M., O'Malley, R., Siegel, D., McClain, C., Sarmiento, J., Feldman, G., et al. (2006). Climate-driven trends in contemporary ocean productivity. *Nature*, *444*, doi:10.1038/nature05317

Bopp, L., Aumont, O., Cadule, P., Alvain, S., & Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: A global model study. *Geophysical Research Letters*, 32(19), L19606. http://doi.org/10.1029/2005GL023653

Bopp, L., Resplandy, L., Orr, J., Doney, S., Dunne, J., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences*, 10, 6225 -6245. doi:10.5194/bg-10-6225-2013.

Boyd, P.W., Strzepek, R., Fu, F., & Hutchins, D.A. (2010). Environmental control of open-ocean phytoplankton groups: Now and in the future. *Limnology and Oceanography*, *55*(3), 1353–1376. http://doi.org/10.4319/lo.2010.55.3.1353

Brewin, R.J.W., Hirata, T., Hardman-Mountford, N. J., Lavender, S. J., Sathyendranath, S., & Barlow, R. (2012). The influence of the Indian Ocean Dipole on interannual variations in phytoplankton size structure as revealed by Earth Observation. *Deep Sea Research Part II: Topical Studies in Oceanography*, *77-80*(C), 117–127. http://doi.org/10.1016/j.dsr2.2012.04.009

Brun, P., Vogt, M., Payne, M. R., Gruber, N., O'Brien, C. J., Buitenhuis, E.T., et al. (2015). Ecological niches of open ocean phytoplankton taxa. *Limnology and Oceanography*, *60*(3), 1020–1038. http://doi.org/10.1002/lno.10074

Carton, J.A., Chepurin, G.A. & Chen, L. (2018). SODA3: a new ocean climate reanalysis, *Journal of Climate.*, doi:10.1175/JCLI-D-18-0149.1

Chisholm, S.W. (1992). Phytoplankton size. *Primary productivity and biogeochemical cycles in the sea* (ed. by P.G. Falkowski and A.D. Woodhead), pp. 213–237. Plenum Press, New York.

Ciotti, A., & Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnology and Oceanography: Methods*, 4, 237–253.

Ciotti, A., Lewis, M., & Cullen, J. (2002). Assessment of the relationship between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnology and Oceanography*, 47, 404–417. doi: 10.4319/lo.2002.47.2.0404

Clark, M., & Cramer III, R.D. (1993). The probability of chance correlation using partial least squares. *Quantitative Structure-Activity Relationships*, 12, 137-145

Clayton, S., Nagai, T., & Follows, M.J. (2014). Fine scale phytoplankton community structure across the Kuroshio Front. *Journal of Plankton Research*, 36(4), 1017–1030. http://doi.org/10.1093/plankt/fbu020

Dave, A.C., & Lozier, M.S. (2010). Local stratification control of marine productivity in the subtropical North Pacific. *Journal of Geophysical Research*, 115(C12), C05001–16. http://doi.org/10.1029/2010JC006507

Dutkiewicz, S., Scott, J.R., & Follows, M.J. (2013). Winners and losers: Ecological and biogeochemical changes in a warming ocean. *Global Biogeochemical Cycles*, 27(2), 463–477. http://doi.org/10.1002/gbc.20042

Finkel, Z.V. (2007). Does phytoplankton cell size matter? The evolution of modern marine food webs. *Evolution of primary producers in the Sea* (ed. by P.G. Falkowski and A.H. Knoll), pp. 333–350. Academic Press, San Diego, CA.

Finkel, Z.V., Beardall, J., Flynn, K.J., Quigg, A., Rees, T.A.V. & Raven, J.A. (2009). Phytoplankton in a changing world: cell size and elemental stoichiometry. *Journal of Plankton Research*, 32, 119–137.

Grant, M., Jackson, T., Chuprin, A., Sathyendranath, S., Zühlke, M., Dingle, J., Storm, T., et al. (2017). Ocean Color Climate Change Initiative Product User's Guide Version 3.1, Plymouth Marine Laboratory, Plymouth, UK (http://www.esa-oceancolour-cci.org)

Gregg, W.W., & Rousseaux, C.S. (2014). Decadal trends in global pelagic ocean chlorophyll: A new assessment integrating multiple satellites, in situ data, and models. *Journal of Geophysical Research-Oceans*, *119*(9), 5921–5933. http://doi.org/10.1002/2014JC010158

Guidi, L., Legendre, L., Reygondeau, G., Uitz, J., Stemmann, L., & Henson, S.A. (2015). A new look at ocean carbon remineralization for estimating deepwater sequestration. *Global Biogeochemical Cycles*, 1–16. http://doi.org/10.1002/2014GB005063

Henson, S.A., Beaulieu, C., Ilyina, T., John, J. G., Long, M., Séférian, R., et al. (2017). Rapid emergence of climate change in environmental drivers of marine ecosystems. *Nature Communications*, 8, 14682. http://doi.org/10.1038/ncomms14682

Hu, C., Lee, Z. & Franz, B. (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *Journal of Geophysical Research*, 117, C01011. doi:01010.01029/02011JC007395

IOCCG (2014). *Phytoplankton Functional Types from Space*. Reports of the International Ocean-Colour Coordinating Group, No. 15, IOCCG, Dartmouth, NS

Irwin, A. J., & Finkel, Z. V. (2008). Mining a sea of data: Deducing the environmental controls of ocean chlorophyll. *PLoS ONE*, *3*(11), e3836. http://doi.org/10.1371/journal.pone.0003836

Jackson, T., & Grant, M. (2016). Ocean Colour Algorithm Blending Utilizing Fuzzy Optical Water Classification. Algorithm Theoretical Baseline Document (Ocean Colour Algorithm Blending), http://www.esa-oceancolour-cci.org/?q=webfm_send/587

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., et al. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77, 437-470

Kostadinov, T.S., Cabré, A., & Vedantham, H. (2017). Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and Earth System Models. *Remote Sensing of Environment*, *190*, 162–177. doi:http://doi.org/10.1016/j.rse.2016.11.014

Lee, Z.P., Carder, K.L. & Arnone, R. (2002). Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, *Applied Optics*, 41, 5755–5772

Lee, Z.P., Du, K.P. & Arnone, R. (2005). A model for the diffuse attenuation coefficient of downwelling irradiance. *Journal of Geophysical Research*, 110(C02016). Doi: 10.1029/2004JC002275.

Lee, Z.P., Weidemann, A., Kindle, J., Arnone, R., Carder, K.L. & Davis, C. (2007). Euphotic zone depth: Its derivation and implication to ocean-color remote sensing, *Journal of Geophysical Research*, 112, C03009. doi:10.1029/2006JC003802

Litchman, E. & Klausmeier, C.A. (2008). Trait-based community ecology of phytoplankton. *Annual Review of Ecology, Evolution, and Systematics*, 39, 615–639.

Litchman, E., Pinto, P., Klausmeier, C.A., Thomas, M.K. & Yoshiyama, K. (2010). Linking traits to species diversity and community structure in phytoplankton. *Hydrobiologia*, 653, 15–28.

Lozier, M. S., Dave, A. C., Palter, J. B., Gerber, L. M., & Barber, R. T. (2011). On the relationship between stratification and primary productivity in the North Atlantic. *Geophysical Research Letters*, 38(18), L18609. http://doi.org/10.1029/2011GL049414

Marinov, I., Doney, S. C., Lima, I. D., Lindsay, K., Moore, J. K., & Mahowald, N. (2013). North-South asymmetry in the modeled phytoplankton community response to climate change over the 21st century. *Global Biogeochemical Cycles*, *27*(4), 1274–1290. http://doi.org/10.1002/2013GB004599

Martens, H., & Martens, M. (2000). Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference*, 11, 5-16. doi: https://doi.org/10.1016/S0950-3293(99)00039-7

McClain, C. R. (2009). A Decade of Satellite Ocean Color Observations. *Annual Review of Marine Science*, *1*(1), 19–42. http://doi.org/10.1146/annurev.marine.010908.163650

Mehmood, T., Liland, K.H. Snipe, L. & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62-69. doi:10.1016/j.chemolab.2012.07.010

Mélin, F., Chuprin, A., Grant, M., Jackson, T., & Sathyendranath, S. (2016). Ocean Colour Data Bias Correction and Merging, Algorithm Theoretical Baseline Document, doi:http://esa-oceancolour-cci.org/?q=webfm_send/619

Morel, A., & Berthon, J.-F. (1989). "Surface pigments, algal biomass, and potential production of the euphotic layer: relationships reinvestigation in view of remote-sensing applications." *Limnology and Oceanography*, (8), 1545–1562.

Morel, A., & Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep Sea Research*, 11, 1375–1393. doi: 10.1016/0198-0149(81)90039-X

Mouw, C. B., Barnett, A., & McKinley, G. A. (2016). Phytoplankton size impact on export flux in the global ocean. *Global Biogeochemical Cycles*, 30(10), 1542–1562. http://doi.org/10.1002/2015GB005355

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. *Frontiers in Marine Science*, 4, 497–19. doi:http://doi.org/10.3389/fmars.2017.00041

Mouw, C. B., & Yoder, J. A. (2010). Optical determination of phytoplankton size composition from global SeaWiFS imagery. *Journal of Geophysical Research: Oceans*, *115*(C12), C12018. http://doi.org/10.1029/2010JC006337

O'Reilly, J., Maritorena, S., Mitchell, B., Siegel, D., Carder, K., Garver, S., Kahru, M., & McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, 103, 24937-24953

Pastor, M. V., Palter, J. B., Pelegrí, J. L., & Dunne, J. P. (2013). Physical drivers of interannual chlorophyll variability in the eastern subtropical North Atlantic. *Journal of Geophysical Research-Oceans*, 118(8), 3871–3886. http://doi.org/10.1002/jgrc.20254

Polovina, J. J., Howell, E. A., & Abecassis, M. (2008). Ocean's least productive waters are expanding. *Geophysical Research Letters*, 35(3), L03618. http://doi.org/10.1029/2007GL031745

Rousseaux, C. S., & Gregg, W. W. (2015). Recent decadal trends in global phytoplankton composition. *Global Biogeochemical Cycles*, 29(10), 1674–1688. http://doi.org/10.1002/2015GB005139

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Altenburg-Soppa, M., & Bracher, A. (2012). Remote sensing of coccolithophore blooms in selected oceanic regions using the PhytoDOAS method applied to hyper-spectral satellite data. *Biogeosciences*, *9*(6), 2127–2143. http://doi.org/10.5194/bg-9-2127-2012

Sathyendranath, S., Grant, M., Brewin, R.J.W., Brockmann, C., Brotas, V., Chuprin, A., Doerffer, R., et al. (2018). ESA Ocean Colour Climate Change Initiative (Ocean_Colour_cci): Version 3.1 Data. Centre for Environmental Data Analysis, *04 July 2018*. doi:10.5285/9c334fbe6d424a708cf3c4cf0c6a53f5. http://dx.doi.org/10.5285/9c334fbe6d424a708cf3c4cf0c6a53f5.

Schollaert, S., Rossby, T., & Yoder, J. (2004). Gulf Stream cross-frontal exchange: possible mechanisms to explain interannual variations in phytoplankton chlorophyll in the Slope Sea during the SeaWiFS years. *Deep-Sea Research II*, 51, 173–188.

Siegel, D. A., Behrenfeld, M. J., Maritorena, S., McClain, C. R., Antoine, D., Bailey, S. W., et al. (2013). Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sensing of Environment*, 135, 77–91. doi:http://doi.org/10.1016/j.rse.2013.03.025

Signorini, S. R., Franz, B. A., & McClain, C. R. (2015). Chlorophyll variability in the oligotrophic gyres: mechanisms, seasonality and trends. *Frontiers in Marine Science*, 2, 1–11. http://doi.org/10.3389/fmars.2015.00001

Soppa, M., Völker, C., & Bracher, A. (2016). Diatom Phenology in the Southern Ocean: Mean Patterns, Trends and the Role of Climate Oscillations. *Remote Sensing*, 8(5), 420–17. http://doi.org/10.3390/rs8050420

Taucher, J., & Oschlies, A. (2011). Can we predict the direction of marine primary production change under global warming? *Geophysical Research Letters*, 38(2), L02603. http://doi.org/10.1029/2010GL045934

Vantrepotte, V., & Melin, F. (2009). Temporal variability of 10-year global SeaWiFS time-series of phytoplankton chlorophyll a concentration. *ICES Journal of Marine Science*, 66(7), 1547–1556. http://doi.org/10.1093/icesjms/fsp107

Wetzels, R., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for correlations and partial correlations, *Psychonomic Bulletin & Review*, 19, 1057–1064. doi:10.3758/s13423-012-0295-x

Wilson, C., & Adamec, D. (2001). Correlations between surface chlorophyll and sea surface height in the tropical Pacific during the 1997-1999 El Niño - Southern Oscillation event. *Journal of Geophysical Research*, 106(C12), 31175–31188

Wilson, C., & Coles, V. (2005). Global climatological relationships between satellite biological and physical observations and upper ocean properties. *Journal of Geophysical Research*, 110(C10), C10001–14. http://doi.org/10.1029/2004JC002724

Wirtz, K. (2012). Who is eating whom? Morphology and feeding type determine the size relation between planktonic predators and their ideal prey. *Marine Ecology Progress Series*, 445, 1–12.

Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, 109-130

Yoder, J.A. & Kennelly, M.A. (2003). Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements. Global Biogeochemical Cycles. 17(4):112. doi:10.1029/2002GB001942

Table 1. List of satellite imagery and reanalysis products and sources.

Parameter	**Description and source**	**Use**
$a_{dg}(\lambda)$ (m^{-1})	Spectral absorption of dissolved and detrital matter Ocean Colour Climate Change Initiative (OC-CCI) www.esa-oceancolour-cci.org QAA (Lee et al., 2002; 2007)	S_{fm} calculation
[Chl] (ug L^{-1})	Chlorophyll *a* concentration Ocean Colour Climate Change Initiative (OC-CCI) www.esa-oceancolour-cci.org	S_{fm} calculation & overall analysis
$K_d(490)$ (m^{-1})	Diffuse attenuation coefficient at 490 nm Ocean Colour Climate Change Initiative (OC-CCI) www.esa-oceancolour-cci.org (Lee et al., 2007)	z_{eu} calculation
MLD (m)	Mixed layer depth Simple Ocean Data Assimilation (SODA) www.atmos.umd.edu/~ocean/	Overall analysis
NPP (mg C m^{-2} d^{-1})	Net primary production Ocean Productivity www.science.oregonstate.edu/ocean.productivity/ SeaWiFS and MODIS R2014 merged following Mélin (2016)	Overall analysis
PAR (μmol quanta m^{-2} d^{-1})	Photosynthetically active radiation NASA Ocean Color Web oceancolor.gsfc.nasa.gov/ SeaWiFS and MODIS R2014 merged following Mélin (2016)	Overall analysis
Q_{net} (W m^{-2})	Net total heat flux NCEP/NCAR reanalysis www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.derived.surfaceflux.html	Overall analysis
$R_{rs}(\lambda)$ (sr^{-1})	Spectral remote sensing reflectance Ocean Colour Climate Change Initiative (OC-CCI) www.esa-oceancolour-cci.org	S_{fm} calculation
S_{fm} (%)	Percent microplankton Calculated using OC-CCI $R_{rs}(\lambda)$, [Chl], and $a_{CDM}(\lambda)$ products https://doi.pangaea.de/10.1594/PANGAEA.892211	Overall analysis
SLA (m)	Sea level anomaly AVISO www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/msla-mean-climatology.html - c10358	Overall analysis
SST (°C)	Sea Surface Temperature Group for High Resolution SST (GHRSST) data.nodc.noaa.gov/ghrsst/	Overall analysis
z_{eu} (m)	Euphotic depth Calculated from OC-CCI $K_d(490)$ as $z_{eu} = 4.6/K_d(490)$ (Morel & Berthon, 1989)	Overall analysis

Parameter	**Description and source**	**Use**
$\Delta\rho_{200}$ (kg m^{-3})	Stratification index Simple Ocean Data Assimilation (SODA) Derived σ_θ as the difference in density between the surface and 200 m www.atmos.umd.edu/~ocean/	Overall analysis
σ_θ (kg m^{-3})	Potential density Simple Ocean Data Assimilation (SODA) www.atmos.umd.edu/~ocean/	$\Delta\rho_{200}$ calculation

FIGURE CAPTIONS

Figure 1. Phytoplankton size, represented as percent microplankton (a), and relationships with environmental parameters, chlorophyll concentration (b), net primary production (c), photosynthetically available radiation (d), euphotic depth (e), sea surface temperature (f), sea level anomaly (g), stratification index (h), mixed layer depth (i), and heat flux (j). Data are 1ºx1º, monthly resolution for pixels with 100% data coverage across all variables. See Table 1 for parameter definitions and data sources. Column 1) Mean from 1998-2015. For mean SST (f1), the 18ºC isotherm is outlined. Column 2) Long term linear trend from the Theil-Sen approach (significant results $BF_{10}>3$ shown). The transect for Hovmoller plots in Figure 8 is highlighted. Column 3) Percentage of the S_{fm} time series above the global mean with 50% contour outlined (a3) or Kendall rank correlation coefficient with S_{fm} (b3-j3, significant results $BF_{10}>3$ shown).

Figure 2. Global relationship of environmental variables with percent microplankton (S_{fm}). We used partial least square regression (PLSR) which combines predictor variables into principle components that are then regressed with S_{fm}. The method allows co-linearity between predictors since they all contribute to forming principle components. a) PLSR coefficients which represent the magnitude and direction of each predictor on S_{fm} response. b) VIP scores (variable influence on projection) represent the relative importance of each predictor to S_{fm} variability. Euphotic depth, [Chl], NPP, SST, and PAR are the most important correlates with S_{fm}. Bars in grey are non-significant.

Figure 3. Relationship between percent microplankton and chlorophyll. a) Regions for PLSR analysis were defined by isolating the impact of temperature at the 18℃ isotherm (Fig. 1, f1) combined with the correlation of S_{fm} with [Chl] (Fig. 1, b3). b-g) PLSR coefficients and VIP scores for each region. Bars in grey are non-significant.

Figure 4. Relationships between percent microplankton and chlorophyll variance. a) Regions for PLSR analysis were defined by isolating the impact of temperature at the 18℃ isotherm (Fig. 1, f1) combined with areas dominated by small or large phytoplankton (Fig. 1a3) and variance in [Chl] (data not shown). Variance in [Chl] was defined as regions falling greater than the 75th percentile, between the 25th and 75th percentiles and less than the 25th percentile. The transect for Hovmoller plots in Fig. 8 is also shown. b-j) PLSR coefficients and VIP scores for each region. Bars in grey are non-significant.

Figure 5. Regions coded by significant driver of phytoplankton size variability. With few exceptions, z_{eu}, [Chl] and NPP were important in all regions. Drivers with significant importance beyond these three base parameters are indicated with a "+". A "-" indicates one of the base drivers is not statistically significant for that region. The six environmental regions represented in this figure correspond to the nine S_{fm} and [Chl] regions in Figure 4 as +SST, MLD, $\Delta\rho_{200}$ – NPP = small, cold, moderate $\sigma_{[Chl]}$; +SST and MLD = large, warm, low $\sigma_{[Chl]}$; +SLA = large, warm, moderate $\sigma_{[Chl]}$; +Q_{net} and MLD = small, warm, moderate $\sigma_{[Chl]}$; +Q_{net}, MLD, PAR – z_{eu} = small, cold, low $\sigma_{[Chl]}$; and z_{eu}, [Chl] and NPP = 4 size and [Chl] regions including 1) large, cold, high $\sigma_{[Chl]}$, 2) large, cold, moderate $\sigma_{[Chl]}$, 3) large, warm, high $\sigma_{[Chl]}$, and 4) small, warm, low $\sigma_{[Chl]}$.

Figure 6. Global maps of VIP scores for a) euphotic depth, b) chlorophyll concentration, c) net primary production, d) sea surface temperature, e) photosynthetically active radiation, f) stratification index, g) sea level anomaly, h) heat flux, and i) mixed layer depth. Areas shaded grey are below the VIP threshold of significance (VIP<0.5).

Figure 7. Frequency distribution of statistically significant (BF_{10}>3) long term linear trends from the Theil-Sen approach (Figure 1, center column) of the regions defined in Figure 4a.

Figure 8. Example of temporal change in percent microplankton and chlorophyll concentration. The transect traverses the South Equatorial Pacific Ocean (100°W) as displayed in Figure 1. The parameters depicted are those that are statistically significant with phytoplankton size. Size (a) and chlorophyll (b) is declining in the north and increasing in the south. c) To compare these parameters simultaneously, we have coded them on a pixel-by-pixel basis, where if the given value was above the mean than it was coded "high" and conversely for "low." Net primary production (d) is increasing while SST (e) is cooling. Euphotic depth is shallowing (e). Heat flux (g) varies significantly over the annual cycle, but is predominately neutral over the whole length of the record, with a decline near the northern reaches of the transect, while mixed layer depth (h) is slightly declining across the whole transect. The white horizontal lines indicate the transitions between regions depicted in Figure 4a.