Temporal and spatial scales of correlation in marine phytoplankton communities

A. M. Kuhn[1], S. Dutkiewicz[2], O. Jahn[2], S. Clayton[3], T. A. Rynearson[4], M. R. Mazloff[1] and A. D. Barton[1]

[1] Scripps Institution of Oceanography, University of California, San Diego, CA, USA

[2] Massachusetts Institute of Technology, Cambridge, MA, USA

[3] Department of Ocean, Earth and Atmospheric Sciences, Old Dominion University, Norfolk, VA, USA

[4] Graduate School of Oceanography, University of Rhode Island, Kingston, RI, USA

Key Points

- Correlation timescales in phytoplankton communities are longer in the subtropical gyres and shorter in regions of strong circulation
- Spatial correlations in phytoplankton communities are strongly anisotropic along frontal zones and boundary currents
- Ocean currents shape global patterns of temporal and spatial correlation scales in phytoplankton communities

Abstract

Ocean circulation shapes marine phytoplankton communities by setting environmental conditions and dispersing organisms. In addition, processes acting on the water column (e.g., heat fluxes and mixing) affect the community structure by modulating environmental variables that determine *in situ* growth and loss rates. Understanding the scales over which phytoplankton communities vary in time and space is key to elucidate the relative contributions of local processes and ocean circulation on phytoplankton distributions. Using a global ocean ecosystem model, we quantify temporal and spatial correlation scales for phytoplankton phenotypes with diverse functional traits and cell sizes. Through this analysis, we address these questions: 1) Over what timescales do perturbations in phytoplankton populations persist?; and 2) Over what distances are variations in phytoplankton populations synchronous? We find that correlation timescales are short in regions of strong currents, such as the Gulf Stream and Antarctic Circumpolar Current. Conversely, in the subtropical gyres, phytoplankton population anomalies persist for relatively long periods. Spatial correlation length scales are elongated near ocean fronts and narrow boundary currents, reflecting flow paths and frontal patterns. In contrast, we find nearly isotropic spatial correlation fields where current speeds are small, or where mixing acts roughly equally in all directions. Phytoplankton timescales and length scales also vary coherently with phytoplankton body size. In addition to aiding understanding of phytoplankton population dynamics, our results provide global insights to guide the design of biological ocean observing networks, and to better interpret data collected at long-term monitoring stations.

Plain Language Summary:

Using a global model of the marine planktonic ecosystem, we quantify the temporal and spatial correlation scales of diverse types of phytoplankton. The timescales reflect the persistence of anomalies in time and the stability of the planktonic system. The spatial scales measure over what distances variations in phytoplankton populations are synchronous. We find that timescales and length scales vary with cell size, and that global patterns of correlation are shaped by ocean currents. These results provide valuable insights for the design of ocean observing systems with a unique ecological perspective. We also discuss how regional differences in phytoplankton community correlation scales are relevant to interpret data collected at long-term monitoring stations.

1 Introduction

Marine phytoplankton generate roughly half of the global net primary production [*Field et al.*, 1998]. Their community structure, including both size and taxonomic diversity, determines the local transfer of energy to higher trophic levels, as well as carbon export to the deep ocean [*Legendre & Le Fèvre*, 1995; *Richardson & Jackson*, 2007; *Smetacek*, 1999]. The mechanisms that shape the diversity of phytoplankton communities are complex and vary among ocean regions and temporal scales [*Acevedo-Trejos et al.*, 2015; *Barton et al.*, 2010; *Fuhrman et al.*, 2008; *de Vargas et al.*, 2015]. Quantifying the temporal and spatial persistence of individual phytoplankton phenotypes and groups is essential for disentangling the underlying controls of phytoplankton diversity globally. This task, however, is difficult due to the continuous dispersal of phytoplankton communities and the sparseness of phytoplankton observations even in the most sampled regions of the ocean. Here, we quantify global temporal and spatial scales of correlation for a diverse set of modeled phytoplankton, spanning a broad range of cell sizes and traits. Specifically, we analyze 24 years of output from the Massachusetts Institute of Technology general circulation model (MITgcm), where a range of phytoplankton and their predators, as well as ocean biogeochemical cycles of N, P, Si, and Fe, are embedded in a high-resolution (1/5º) ocean state estimate. We address how temporal and spatial correlation structures vary across phytoplankton phenotypes, and how these properties are affected by circulation and mixing in the ocean surface. Additionally, we compare phytoplankton correlation scales to sea surface temperature and nitrate correlation scales.

Several previous studies have analyzed the temporal and spatial scales of correlation of physical and chemical ocean properties, such as sea surface temperature (SST), sea surface height (SSH) and dissolved inorganic carbon (DIC), often with the purpose of informing the design of sampling strategies and observing systems [e.g., *Kessler et al.*, 1996; *Kuragano & Kamachi,* 2000; *Mazloff et al.,* 2018]. For example, *Kessler et al.* [1996] analyzed the scales of thermal variability in the equatorial Pacific using SST and thermocline depth data from the Tropical Atmosphere-Ocean moored buoy array, concluding that the distance between buoys needed to be reduced to adequately sample signals with periods between 1 – 2 months. *Kuranago and Kamachi* [2000] used global spatial correlation scales obtained from altimeter data to design an optimal interpolation method that improved the correlation between interpolated satellite altimeter data and *in situ* sea levels from tide gauges. Most recently, *Mazloff et al.* [2018] addressed the needs for the biogeochemical-Argo floats observational networks in the Southern Ocean by estimating spatial correlation scales for oceanic dissolved inorganic carbon, heat content, and

carbon and heat exchanges. Their results suggested that a minimum of 100 Argo floats are required to monitor biogeochemical properties in the Southern Ocean [*Mazloff et al.,* 2018]. Correlation scales are also necessary to improve optimal interpolation and data assimilation methods for operational forecasts and evaluation of high resolution ocean models [*Gaillard et al.*, 2009; *Glover et al.*, 2018; *Ninove et al.*, 2016].

Ocean currents are likely to play a key role in setting spatial and temporal correlation scales in biological and physical properties. For example, spatial correlations in surface chlorophyll have been found to decrease faster along currents than in more quiescent regions [*Denman & Abbott,* 1994]. In highly dynamic locations, power spectra of SST and chlorophyll are very similar, suggesting that dispersal by strong ocean currents is more important than ecological interactions in determining phytoplankton spatial patterns in these regions [*Denman & Abbot,* 1994]. At global scales, satellite chlorophyll length scales and mixing scales have been shown to vary seasonally and temporally in coherence with biological and physical factors such as upwelling and western boundary currents [*Doney et al.,* 2003; *Glover et al.,* 2018].

Modelling studies also increasingly emphasize the role of physical dispersal in shaping the diversity of planktonic ecosystems [*Adjou et al.,* 2012; *Barton et al.,* 2010; *Bracco et al.,* 2009; *Clayton et al.,* 2013; *Lévy et al.,* 2014]. For example, systematically introducing dispersal by vertical mixing, horizontal currents, and eddies increased phytoplankton diversity in an idealized model representing the North Atlantic and the Gulf Stream [*Lévy et al.,* 2014]. At global scales, hot-spots of phytoplankton diversity appear related to ocean areas with high eddy-kinetic energy [*Barton et al.,* 2010; *Clayton et al.,* 2013]. We follow-up on these modelling studies and investigate how the differences in the temporal and spatial scales of correlation for different phenotypes relate to current speed and cell size. The results of this study thus provide a framework to improve our understanding of the distribution of marine phytoplankton phenotypes and traits in the ocean.

The goals of this study are: 1. to quantify and contrast patterns of temporal and spatial correlation scales for a broad range of model phytoplankton phenotypes, functional groups, and cell sizes, and 2. to discuss how these patterns vary across regions, and how dispersal by ocean currents and eddy activity affects them. We investigate variations across cell size because it is often considered a master trait constraining many physiological rates (e.g., growth rate and nutrient uptake kinetics), as well as predator-prey interactions [e.g., *B. Hansen et al.*, 1994; *P. J. Hansen et al.*, 1997; *Marañón et al.*, 2013]. After a brief description of the model and methodology, we organize this paper into two main sections. In the first part of our study we estimate correlation timescales. In essence, we ask: how long do perturbations in

phytoplankton surface biomass persist? We hypothesize that phytoplankton populations that exhibit rapid temporal changes will have short correlation timescales that may be linked to strong dispersal and mixing of organisms or to high variability in environmental conditions imposed by either advection, mixing or air-sea interactions. In contrast, regions with relatively sluggish circulation and relatively invariant environmental conditions over time may have relatively long correlation timescales.

In the second part of our study we quantify the spatial patterns of correlation between each grid point and the surrounding waters. We ask: over what distances are variations in surface phytoplankton populations synchronous? We address this question in two-dimensions and discuss how the shape of the spatial correlation field contains information about the processes and physical features driving the correlation patterns. When the dispersal of phytoplankton is greater in one direction (i.e., anisotropic), such as along ocean currents, we hypothesize that elongated phytoplankton spatial correlation fields will resemble the physical features affecting it. In contrast, round correlation fields would develop when mixing or environmental forcing is approximately equal in all directions (i.e., isotropic). We discuss our results in the context of ocean dynamics and trait differences in phytoplankton groups, taking into consideration the scales of correlation found in previous studies for physical variables known to affect phytoplankton distributions.

2 Model description

We use output from a coupled physical-ecosystem model using the MITgcm [https://doi.org/10.6075/J0BR8QJ1; *Jahn et al., 2019*]. The physical component of the model uses the ECCO2 physical configuration with a resolution of 1/5º, permitting the formation of eddies and narrow currents [*Menemenlis et al.,* 2008]. The ecosystem and biogeochemical components include a complex plankton community model, and resolve the cycling of carbon, phosphorus, nitrogen, silica, iron, and oxygen. The ecosystem model is based on *Dutkiewicz et al.,* [2015], and incorporates both functional and size diversity of plankton (with parameterizations based on *Ward et al.,* [2012]), resolving a total of 51 plankton types (35 phytoplankton and 16 zooplankton). Phytoplankton are sub-classified into 6 functional groups: prokaryotes, picoeukaryotes, coccolithophores, diazotrophs, diatoms, and mixotrophic dinoflagellates (see supplemental Fig. S1). All groups are modeled with Monod kinetics with constant C:N:P:Fe stoichiometry over time. Phytoplankton functional groups differ in nutrient requirements, maximum growth rates, pigment composition, and palatability to predators. Phytoplankton cell sizes increase logarithmically from 0.6 to 228 μm in diameter, with each functional group having a characteristic range of sizes (Fig. S1). Cell size determines differences in maximum growth rates, grazing, and sinking, as described in *Dutkiewicz et al.*, [2019]. Following observations, the smallest phytoplankton

(the prokaryotes and picoeukaryotes) have the lowest nutrient affinity, and the fastest growing are in the 3 μm cell size range [*Marañon et al.,* 2013]. The zooplankton graze, using a Holling III function, on plankton 5 to 15 times smaller than themselves, but with a preference for 10 times smaller than themselves.

The model represents with fidelity the regional and seasonal patterns of total chlorophyll and the distribution of chlorophyll concentrations between key size classes in the ocean surface (Fig. S2, S3).

Output from this model has been used in previous studies on community structure [*Benoiston et al.,* 2017; *McParland & Levine,* 2019; *Tréguer et al.,* 2018]. Here we analyze 3-day means of biomass output for each of the 35 phytoplankton phenotypes over 24 years (1992-2016). We analyze surface (upper 10 m) averaged output from the physical and marine ecosystem model components. The 10 m output are representative of the mixed layer. Our study is focused on the mixed layer community and do not take into account the deeper "shade" communities such as in the deep Chl-a maximum. Additionally, we use surface current speed and sea surface temperature (SST) model output to provide context and discuss the correlation patterns obtained.

3 Analysis of correlation scales

We estimate temporal and spatial scales of correlation for each of the 35 phytoplankton types in the model simulation, P_i (mmol N m^{-3}), (where $i = 1, \ldots, 35$), as well as for total phytoplankton biomass $\sum_{i=1}^{35} P_i$ (mmol N m^{-3}). The temporal correlation analysis indicates how fast the phytoplankton community at a particular location becomes uncorrelated with itself, while the spatial analysis shows how far the phytoplankton community at a given point varies in synchrony with the communities in surrounding locations.

In order to identify the effects of dispersal, we focus our analyses on the intra-annual variability of phytoplankton biomass. For this reason, we first remove the seasonal and long-term fluctuations in the series. We remove seasonality by subtracting daily climatological mean fields from the surface biomass values. Additionally, we remove long-term fluctuations by subtracting a 90-day running average filter from the series. The method effectively dampens interannual variability and other long-term oscillations (Fig. S4). We refer to the resulting, derived data as biomass anomalies (P_i') and use these data to estimate both the temporal and spatial correlation scales.

While we perform the analyses for all phenotypes, we highlight the differences between phytoplankton functional groups by comparing organisms classified as gleaners and opportunists

throughout the results and discussion (Fig. S1). Gleaners and opportunists, also known as K- versus r-strategists [*Kilham & Hecky,* 1988; *Mac Arthur & Wilson,* 1967], have strongly contrasting ecological dynamics and impacts on food webs and biogeochemical cycles [*Dutkiewicz et al.,* 2009]. The gleaners are phytoplankton types with relatively high nutrient affinity that typically dominate in oligotrophic regions [*Edwards et al.,* 2012]. The gleaners here include the four smallest size classes simulated by the model (i.e., prokaryotes and picoeukaryotes with ≤2μm spherical diameter). In order to compare the same number of phenotypes, the opportunists are represented by the four smallest and most abundant size classes of diatoms (3 -10 μm spherical diameter). The small diatoms can be considered opportunists because of their high maximum specific growth rates relative to other phytoplankton [*Marañón et al.,* 2013]. Diatoms such as these typically dominate during seasonal blooms [e.g., *Bruland et al.,* 2001; *Klais et al*., 2011; *Leblanc et al.*, 2018]. In order to analyze the relationships between cell size and timescales and spatial scales, we use a second classification to differentiate between small cells (≤3μm; 7 phenotypes) and large cells (>3μm; 28 phenotypes). Hence, the small phytoplankton group includes all gleaners, the smallest of the opportunist diatoms, the smallest coccolithophore and the smallest diazotroph. All other phenotypes are included in the large phytoplankton group (Fig. S1). This classification is based both on typical allometric scaling for various phytoplankton rates [*e.g. Marañón et al.,* 2013], and on the emergent properties of the present analysis (see sections 4.2.1 and 4.2.3).

Additionally, we estimated correlation scales for sea surface temperature (SST) and nitrate (NO_3) in order to provide context about the physical and chemical drivers of phytoplankton growth. In the next sections (sections 3.1 and 3.2), we describe the methods in terms of phytoplankton biomass only.

3.1 Temporal correlation

Temporal correlation scales for phytoplankton are estimated through the autocorrelation function of the biomass anomalies. For each model grid point, we calculate the temporal autocorrelation function ($r_i(\tau)$), where τ is the time lag in days. The autocorrelation function ($r_i(\tau)$) is:

$$r_i(\tau) = \frac{1}{(n-1)\sigma_i^2}\sum(P'_{i,t} - \overline{P'_i})(P'_{i,t+\tau} - \overline{P'_i}) \quad (1)$$

where *t* refers to the model timestep in days, $\overline{P'_i}$ is the time mean of the biomass anomalies for each phenotype *i*, σ_i^2 is the temporal variance of the biomass anomalies for each phenotype, and *n* is the number of data points. In general, autocorrelation is high at shorter time lags and decreases with increasing time. We define an autocorrelation decay timescale as the e-folding correlation timescale, τ_e, which is the first lag (day) when the autocorrelation drops below $r_i(\tau_e) = e^{-1} \approx 0.37$. We call τ_e the

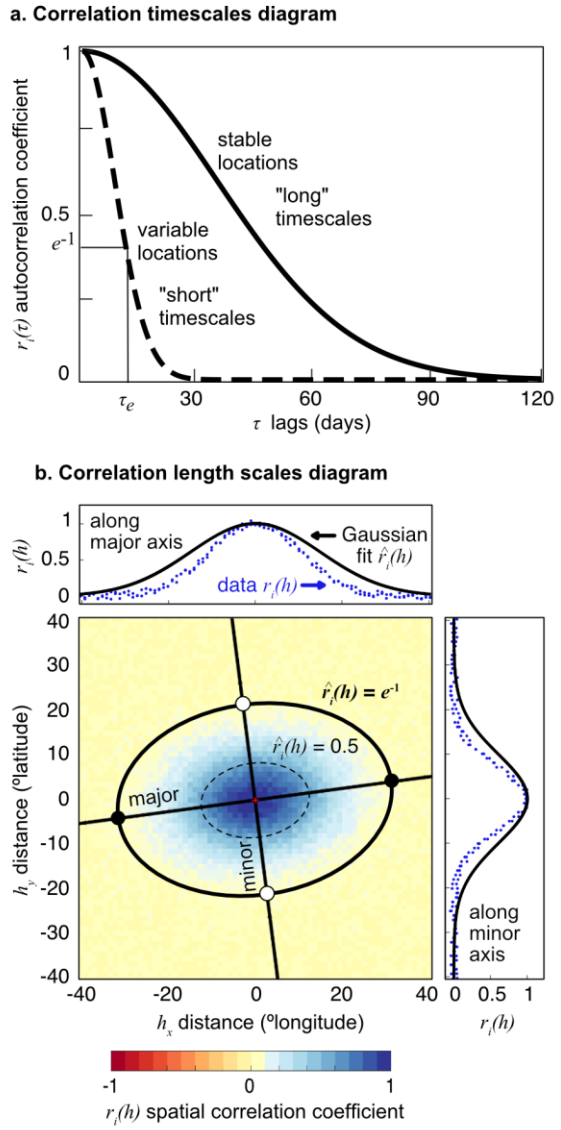
"correlation timescale" hereafter. E-folding timescales are a commonly used measure of decorrelation for intra-seasonal or "short-term" timescales, such those considered in this study. This method differs from integral timescales, which are the time lag at which the area under the autocorrelation function is maximum [*Talley,* 2011]. In general, integral timescales tend to be longer than timescales estimated using a fixed correlation threshold. We anticipate that locations where biomass anomalies are highly variable in time will have short correlation timescales, whereas locations that are stable through time will have long correlation timescales (Fig. 1a).

3.2 Spatial correlation

We next calculate the phytoplankton spatial scales of correlation from the biomass anomaly time series. This analysis correlates time series of biomass anomaly at each grid point to that of surrounding grid points to assess the similarity of population dynamics in space (e.g., Fig. 1b). We calculate the spatial correlation field $r_i(h_x, h_y)$, where h_x and h_y are horizontal distances away from a grid point of interest *k*, which has coordinates x_k, y_k. To simplify the notation, we drop the coordinates subscripts and define $r_i(h_x, h_y) = r_i(h)$ as:

$$r_i(h) = \frac{\sum(P'_{i,k} - \overline{P'_{i,k}})(P'_{i,k+h} - \overline{P'_{i,k+h}})}{\sigma_{i,k}\sigma_{i,k+h}}, \quad (2)$$

here $\sigma_{i,k}$ and $\sigma_{i,k+h}$ are the standard deviations of the biomass anomalies series of each phenotype *i* at the grid point *k,* and at *k* plus a given distance (*k*+*h*), respectively. We evaluate the correlation between the central grid point and all grid points within 15º of latitude and longitude in order to avoid comparing phytoplankton dynamics at distant locations, for example across ocean basins. For computational efficiency, we do not calculate the length scales for every grid point, but central grid points are selected once every 10 grid points (2˚ of latitude and longitude) globally. In regions with horizontal current speeds higher than 20 cm s^{-1}, we calculate the length scales every 2 grid points (~0.4˚).



226

227 **Figure 1.** Diagrams of the methods used to estimate phytoplankton temporal and spatial scales of
228 correlation. **a.** Idealized examples of the autocorrelation function for locations with long and short
229 timescales. The timescale value, τ_e, is defined as the time lag when the autocorrelation function falls
230 below e^{-1} and is marked for the short timescale example. **b.** Idealized example of spatial correlation
231 field and its corresponding 2D Gaussian fit. The central grid point *k* is marked with red + marker. The
232 major and minor axes length scales are defined using the e-folding contour ($\hat{r}_i(h) = e^{-1}$) of the fitted 2D-
233 Gaussian (outer ellipse, bold black line). In this example, the major axis length is the distance between the
234 two black dots located in the east – west direction, and the minor axis length is the distance between the
235 two white dots located in the north – south direction. The inner ellipse (dashed black line) represents
236 $\hat{r}_i(h) = 0.5$ for reference.

In general, points closer to the central grid point are highly correlated, whereas points further away are less correlated (Fig. 1b). Correlation fields can be either isotropic or anisotropic. For example, population dynamics along a zonal front could be correlated over a great distance of longitude, but decorrelate rapidly with latitude. Along Western Boundary Currents, such as the Gulf Stream, spatial correlation could be high in the direction of the current but low across the orthogonal axis of the current. Alternatively, the correlation could be equal in all directions (isotropic). The shape of the correlation field therefore contains information about the underlying processes that link adjacent ocean regions. Thus, to approximate correlation length scales, we fit an ellipsoidal or two-dimensional Gaussian function $\hat{r}_i(h)$ to the correlation field:

$$\ln\left(\hat{r}_i(h)\right) = -\left[\frac{[(x_{k+h}-x_k)\,cos(\theta)+(y_{k+h}-y_k)\,sin(\theta)]^2}{a^2} + \frac{[(x_{k+h}-x_k)\,sin(\theta)+(y_{k+h}-y_k)\,cos(\theta)]^2}{b^2}\right], \quad (3)$$

where, x and y refer to the coordinates of the central grid point k (x_k, y_k) and its neighbors (x_{k+h}, y_{k+h}). The parameters a and b are the two axes of the Gaussian fit to the correlation field, and θ is the angle. If a is equal to b, the correlation field is isotropic or equal in all directions (i.e., a circle). If a and b are different, the correlation field is anisotropic or elongated (i.e., an ellipse). The longest of the two axes is referred as the major axis and the shortest as the minor axis. We use the major-to-minor axes aspect ratio ($AR = \max(a,b)/\min(a,b)$) to determine whether the shape of the fitted ellipse is anisotropic or isotropic. The angle θ is given as absolute values ranging from 0º to 180º, such that if θ is 0º or 180º it means that the correlation field is elongated zonally or along lines of constant latitude. If θ is 90º the correlation field is elongated along lines of constant longitude (i.e., meridionally elongated).

Parameters a, b and θ are estimated using a weighted least-squares fit to the Gaussian function, where the weight $\omega = h^{-1}$ decreases the importance of fitting locations further away from the grid point k. *A priori* parameter guesses (a^*, b^*, θ^*) are defined by finding the maximum zonal and meridional distances from the center k to the contour $r_i(h) = 0.5$, and the corresponding angle. *A priori* values are assigned a 90% uncertainty. This 2D Gaussian fit methodology is based on the approach used for spatial scales of heat and carbon content in the Southern Ocean by *Mazloff et al.* [2018], where the importance of *a priori* parameter guesses and uncertainty considerations is stressed. As spurious correlations may develop at long distances, the a *priori* guess and weighting factors used in our application ensure that the structure closest to the point of interest is influential in determining the shape of the ellipse. Various other functions and models have been used to fit spatial correlation coefficients, however the 2D Gaussian approach has been shown effective in determining physical length scales in the ocean [*Kuranago and Kamachi,* 2000; *Mazloff et al.,* 2018]. The strength of our approach fitting a 2D Gaussian function to

spatial correlation fields is that it allows us to test in a systematic way across the oceans whether correlation fields tend to be elongated or isotropic.

4 Results

4.1 Model environment

To contextualize the temporal and spatial correlation analysis, we first describe the simulated global patterns of climatological mean phytoplankton biomass (mmol N m^{-3}), SST (ºC), ocean sea surface current speed (cm s^{-1}), and eddy kinetic energy (cm^2 s^{-2}). Eddy kinetic energy, or the energy associated with turbulent motions in the ocean, is calculated as $EKE = \frac{1}{2}\left(\overline{u'^2} + \overline{v'^2}\right)$, where (u', v') are horizontal velocities anomalies from the climatological mean [e.g., *Richardson,* 1983]. While current speed is a measure of the magnitude of ocean currents, EKE is a measure of the magnitude of variability in currents through time.

The model captures global patterns of phytoplankton biomass: subtropical gyres exhibit low annual average phytoplankton biomass, while high biomass exceeding 3 mmol N m^{-3} is found in the high latitudes, equatorial regions, and major coastal upwelling areas (Fig. 2a). Strong gradients in annual average sea surface temperature are apparent, for example crossing from the South Pacific subtropical gyre into the Southern Ocean or from the subtropical to subpolar North Atlantic (Fig. 2b).

Simulated mean ocean current speeds and eddy kinetic energy (EKE) values in Fig. 2c,d agree well with those estimated from global drifter data [*Lumpkin & Johnson,* 2013]). In the Southern Ocean, the ACC shows both strong jets and regions of weaker current speed (Fig. 2c), and its position is bounded by ocean fronts [*Lumpkin & Johnson,* 2013]. The Sub-Antarctic Front at ~55ºS corresponds to the northern boundary, while the Southern Boundary Front (typically with SST < 0º) limits the ACC to the south [*Carter et al.,* 2008]. The equatorial current systems are prominent in the Atlantic, Indian and Pacific Oceans with mean speeds higher than 40 cm s^{-1} (Fig. 2c). The model resolution is eddy-permitting and resolves regions of high current speeds and enhanced EKE associated with narrow western boundary currents and their corresponding extensions (Fig. 2c,d). All major western boundary currents are prominent in the climatological mean current speed: the Gulf Stream (GS), the Kuroshio Current (KC), the Agulhas Current (AC), the Somali Current (SC) and the Brazil Current (BC). Correspondingly, EKE is high in eddy dominated regions, as well as in the path of the North Brazil Current (NBC) and the Brazil-Malvinas Confluence (BMC). EKE is also high in the equatorial currents due to the energy of tropical instability waves [*Chelton et al.,* 2000; *Lumpkin & Johnson,* 2013]. The eastern North Pacific,

eastern South Pacific, eastern South Atlantic, and western South Atlantic have low EKE and correspond to regions classified as "eddy deserts" [*Lumpkin & Johnson,* 2013].

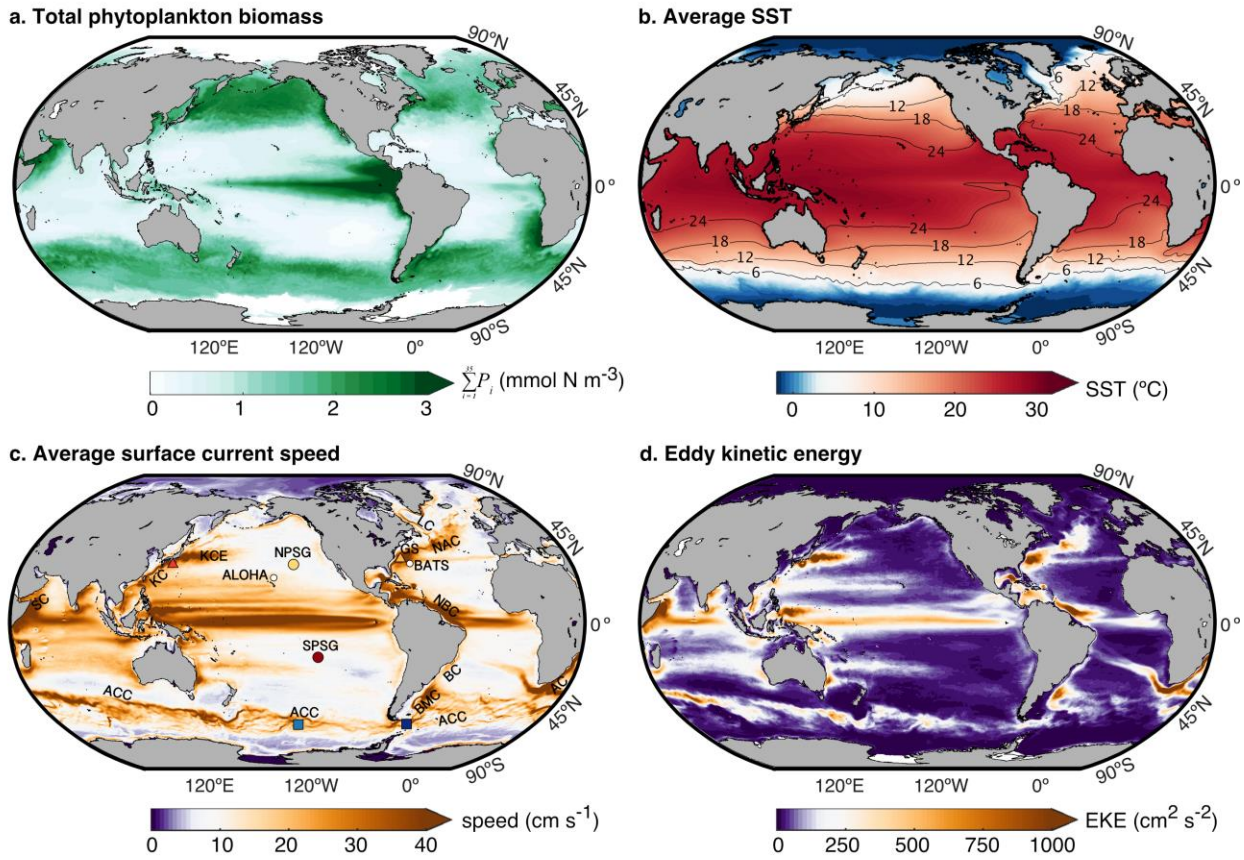


Figure 2. Model annual average: **a.** total surface phytoplankton biomass, **b.** sea surface temperature (SST), **c.** surface current speed, and **d.** eddy kinetic energy (EKE). Ocean currents and gyres discussed throughout the text are indicated in c.: *AC* – Agulhas Current, *ACC* – Antarctic Circumpolar Current, *BC* – Brazil Current, *BMC* – Brazil – Malvinas Confluence, *GS* – Gulf Stream, *KC* – Kuroshio Current, *KCE* – Kuroshio Current Extension, *LC* – Labrador Current, *NAC* - North Atlantic Current, *NBC* – North Brazil Current, *NPSG* – North Pacific Subtropical Gyre, *SC* – Somali Current, and *SPSG* – South Pacific Subtropical Gyre. Markers in c. indicate the location of illustrative regions also used in Fig. 6e. Long-term monitoring stations ALOHA (A Long-term Ocean Habitat Assessment) and BATS (Bermuda Atlantic Time Series), also referred to in Fig. 8f-g, are marked for reference.

In our analysis, we consider separately the four smallest phytoplankton types with the highest nutrient efficiencies ("gleaners") and the four fastest growing types (the four smallest diatoms, the "opportunists"). Gleaners dominate in the most oligotrophic parts of the ocean, while the opportunists

dominate in the highly productive regions (Fig. 3a, b, S1). We also examine how cell size impacts our results and differentiate "small" (≥ 3μm) and "large" (> 3μm) phytoplankton (Fig. 3c, d, S1). We separate phytoplankton by size in this manner because the allometric scaling for specific growth rates has a unimodal shape: it peaks for cells approximately 3μm in equivalent spherical diameter but decreases for smaller and larger cells [*Dutkiewicz et al*., 2019; *Marañón et al.*, 2013]. Much like for gleaners and opportunists, small phytoplankton dominate in oligotrophic regions while larger phytoplankton dominate in more seasonal and eutrophic regions (Fig. 3c, d).

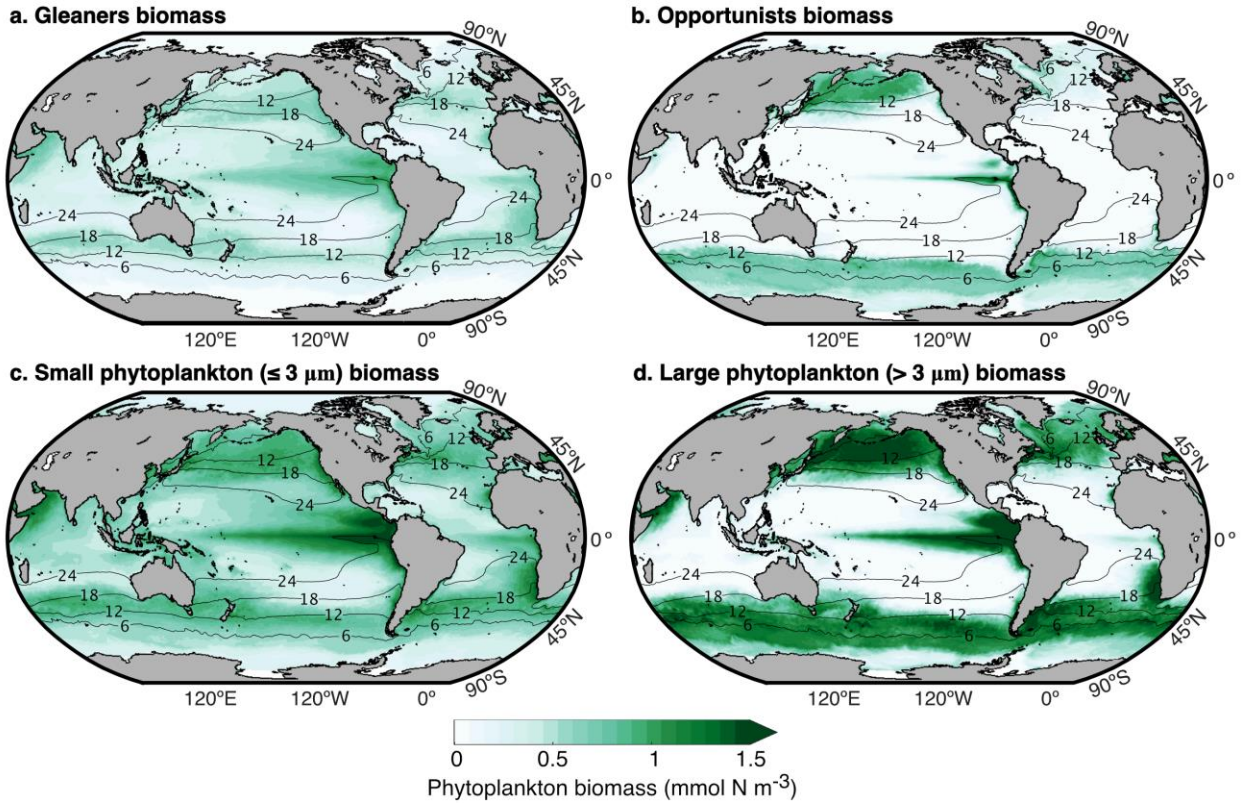


Figure 3. Model annual average: **a.** total gleaners phytoplankton biomass (i.e., the 2 prokaryotes and 2 picoeukaryotes with ≤2μm spherical diameter); **b.** total opportunists phytoplankton biomass (i.e., the 4 smallest diatoms with spherical diameter raging from 3 to 10 μm); **c.** small phytoplankton biomass (including 2 prokaryotes, 2 picoeukaryotes, 1 coccolithophore, 1 diazotroph, and 1 diatom); and **d.** large phytoplankton biomass (including 4 coccolithophore, 4 diazotrophs, 10 diatoms, and 10 mixotrophic dinoflagellates). Contours indicate lines of constant SST (ºC; contours are the same as Fig. 2b).

4.2 Phytoplankton correlation timescales (τ_e)

4.2.1 Correlation timescales and ocean dynamics

In general, the phytoplankton correlation timescales, or τ_e, estimated from biomass anomaly time series are shorter than 30 days. The median phytoplankton τ_e is 18.7, 19.5, and 18 days for gleaners, opportunists, and total biomass, respectively (Fig. 4). Correlation timescales of 30 days or shorter for gleaners, opportunists, and total biomass correspond to 85, 68.5, and 95.9% of ocean area, respectively. Correlation timescales of 15 days or shorter for gleaners, opportunists, and total biomass correspond to 37, 37, and 48.5% of the ocean area, respectively.

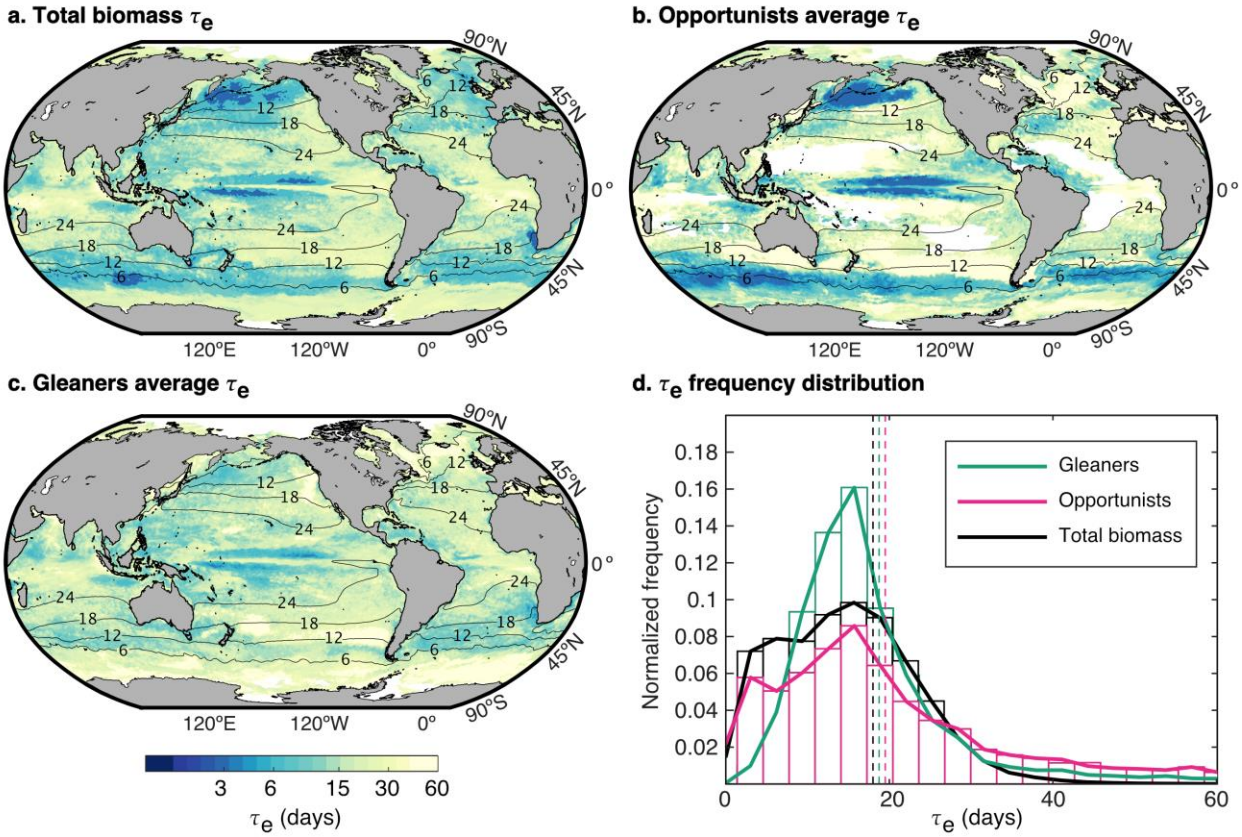


Figure 4. Correlation timescales, τ_e, estimated for **a.** total biomass **b.** opportunists, and **c.** gleaners. Gleaners include the four smallest size classes simulated by the model (i.e., prokaryotes and picoeukaryotes with ≤2μm spherical diameter), and opportunists include the four smallest size classes of diatoms (3 -10 μm spherical diameter). In the case of total biomass, the correlation timescale is calculated from total biomass anomaly time series. In the case of gleaners and opportunists, τ_e is estimated individually for each phenotype in the group and then

averaged. White regions indicate where total biomass is below 10^{-6} mmol N m^{-3} or group biomass is below 1% of total biomass. Contours indicate lines of constant SST (ºC; contours are the same as Fig. 2b). The color scale is logarithmic to improve visualization and comparison between maps. **d.** The bars represent the area-weighted, normalized frequency distribution of global correlation timescales for total biomass anomalies, gleaners and opportunists. Solid lines represent the corresponding probability density functions. Vertical dashed lines represent the median value for each group.

Here we differentiate and discuss regions of the ocean with short ($\tau_e \leq 15$ days) and long ($\tau_e > 15$ days) correlation scales. The overall spatial patterns in τ_e between gleaners, opportunists, and total biomass are qualitatively similar. Short correlation timescales (<15 days) occur in the Southern Ocean (Fig. 5b, d), Equatorial zones, and subpolar North Pacific (Fig. 4a-c), as well as in in Western Boundary Currents such as the Gulf Stream (Fig. 5a, c). Short correlation timescales along the westward travelling Equatorial Currents (located to the north and south of the equator) in the Pacific and Indian Oceans, and to a lesser extend in the Atlantic, may be tied to variable currents [*Masumoto et al.,* 2005] and tropical instability waves [*Han et al.,* 2008; *Kessler et al.,* 1996], which generate high-frequency variations in the environmental conditions in this region [*Han et al.,* 2008; *Kessler et al.,* 1996; *Lyman et al.,* 2005]. Correlation timescales are short in the Subpolar North Pacific but do not correspond closely to enhanced currents or EKE (Fig. 2c, d). However, this short τ_e area coincides with the location of the Aleutian Low, a zone of high storm frequency comprising the Subpolar North Pacific and Bering Sea [*Pickart et al.,* 2009]. Correlation timescales are also short in the Antarctic Circumpolar Current (ACC), but increase both north and south of the ACC (Figs. 4a-c; 5b, d).

In contrast, longer phytoplankton τ_e (>15 days) occurs in subtropical regions and large parts of the Southern Ocean, specifically north of the Subtropical Front at ~45ºS and south of the Sub-Antarctic Front at ~55ºS, approximately denoted by the 6 and 12ºC contours, respectively [*Carter et al.,* 2008]. These areas of very long correlation timescale broadly, though not exclusively, coincide with regions of low current speed and low EKE (Fig. 2c, d).

Compared with gleaners, opportunist phytoplankton exhibit stronger spatial gradients in correlation timescales (compare Fig. 4b-c). Opportunists have both shorter and longer correlation timescales than gleaners depending on the region (Fig. 4d). Regions of long correlation timescales for opportunists are primarily found in the subtropical gyres and in polar zones, with portions of the subtropical and subpolar regions having long τ_e of up to several months. The longest correlation

timescales for opportunists (exceeding 100 days) are found in the Subtropical Front at ~45ºS extending from the coast of Chile to the south of Australia (Fig. 4b, Fig. 5d), and the shortest timescales are found in the central Equatorial Pacific and Subpolar Pacific (Fig. 4b). In contrast, correlation timescales for gleaners are more uniform in space (Fig. 4c) and exhibit a slightly narrower distribution of timescales globally (Fig. 4d). The longest correlation timescales are found in the Subtropical Front at ~45ºS off the coast of Chile and east of the Weddell Sea area of the Southern Ocean up to 60ºE (Fig. 4c, Fig. 5e). The shortest timescales are found north of the equator in the central Pacific (Fig. 4c).

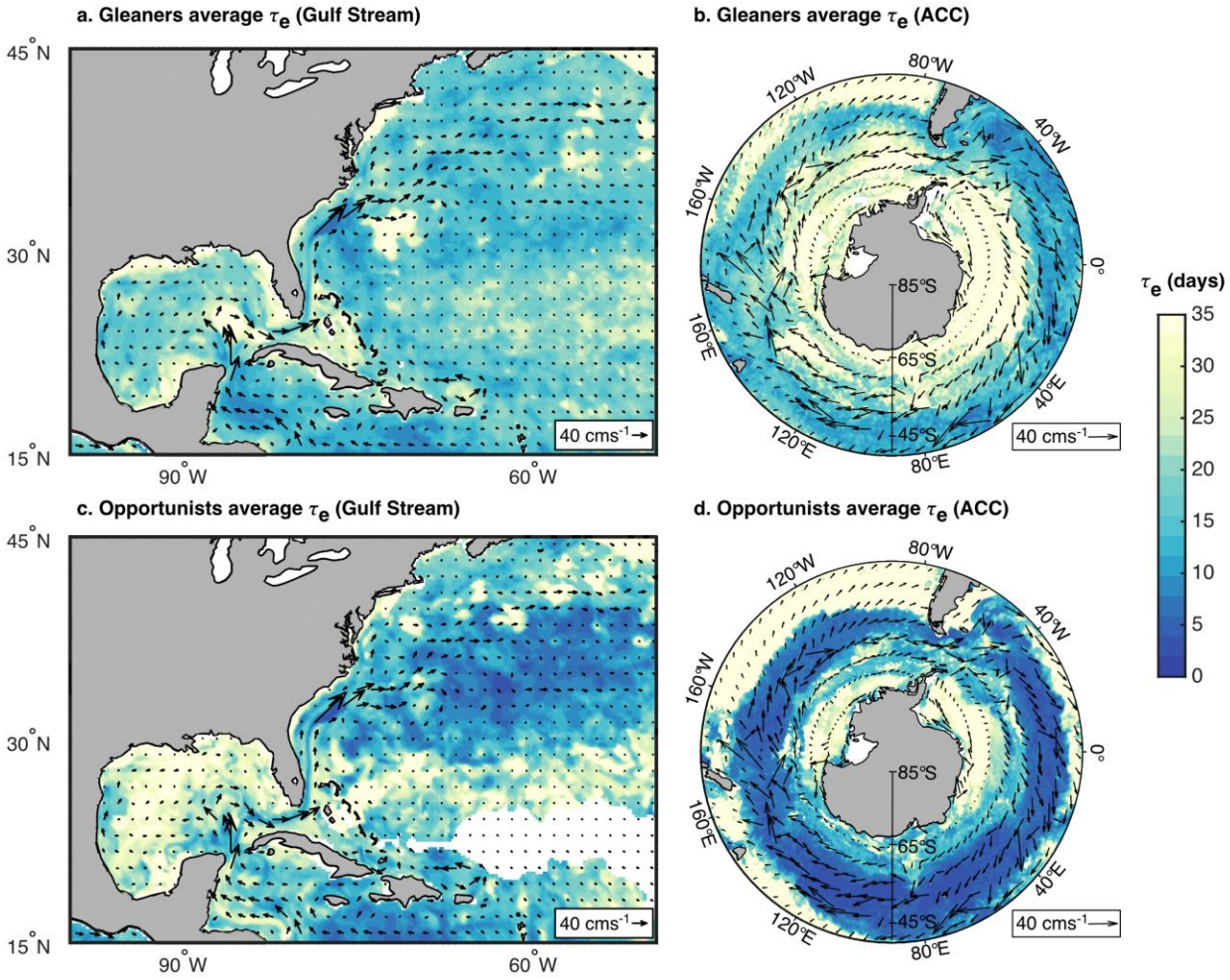


Figure 5. Regional examples of correlation timescales: **a.** Average τ_e estimated from biomass anomalies of gleaners in the Gulf Stream, and **b.** in the Antarctic Circumpolar Current (ACC) region. **c.** Average τ_e estimated from biomass anomalies of opportunists in the Gulf Stream, and **d.** in the ACC region. Current velocity vectors are indicated by black arrows and the color scale is truncated at 35 days to improve visualization. Low biomass areas (group biomass below 1% of total phytoplankton biomass) are masked in white.

Using the same methodology as we did for estimating phytoplankton correlation timescale (see section 3), we estimated correlation timescales for SST and surface NO_3 using the MITgcm model output. SST exhibits long timescales exceeding 45 days in subtropical and certain temperate latitudes, while polar and tropical latitudes and areas of high kinetic energy show correlation timescales under 45 days and down to 15 days (Fig. S5a). Long persistence of SST anomalies, typically 3-5 months, is known to occur in the North Atlantic and North Pacific Oceans, as a consequence of the high heat capacity of the ocean [*Deser et al.,* 2003; *Frankignoul,* 1985; *Frankignoul & Hasselmann,* 1977; *Kushnir et al.,* 2002]. Surface NO_3 correlation timescales exhibit long timescales in subpolar latitudes and short timescales in the tropical and subtropical oceans, as well as polar regions (Fig. S5b). In tropical and sub-tropical oceans, phytoplankton growth quickly consumes any nitrate available (e.g., nutrients supplied by the passing of eddies), yielding short NO_3 correlation timescales. In regions with deeper winter mixed layers and seasonal limitation of phytoplankton growth by temperature or light, nutrients may accumulate, leading to longer NO_3 correlation timescales. Our analysis also shows that, in most of the ocean, phytoplankton timescales are shorter than the timescales of both SST and NO_3 (Fig. S5c, d). This contrast is consistent with the different dominant timescales of the processes influencing each variable: days to weeks in the case of phytoplankton growth and weeks to months in the case of SST, for example. In regions of high EKE, such as the Gulf Stream, the differences between SST, NO_3 and phytoplankton correlation timescales are relatively small, suggesting that in these regions the dynamics of biotic and environmental conditions are tied to ocean mixing and currents (Fig. S5c, d).

In agreement with our estimates, chlorophyll correlation timescales in the North Atlantic Ocean show shorter timescales than measured for the physical variables [*Boss et al.,* 2008]; while *Denman and Abbot* [1994] found equal timescales for SST and chlorophyll in dynamic areas of the California Current System region. As expected, we find a tighter relationship between nitrate timescales and phytoplankton timescales, particularly within nutrient limited areas in the tropics and subtropics, where light is sufficient to sustain phytoplankton growth year-round.

Overall, phytoplankton populations and total biomass in most of the ocean have relatively fast decorrelation timescales, and opportunists and gleaners exhibit similar spatial patterns in correlation timescales that differ in magnitude. Next, we examine more closely how these spatial patterns and differences across gleaners and opportunists in correlation timescale may be linked to variations in current speed in the ocean and cell sizes across phytoplankton types.

4.2.2 Correlation timescales and current speed

For both gleaners and opportunists, short correlation timescales (≤ 15 days) occur more frequently where mean current speed is relatively high (Fig. 6a-d). Conversely, long correlation timescales (>15 days) occur more frequently where mean current speed is relatively low (Fig. 6a-d). This is more clearly represented in the marginal frequency distributions (Fig. 6c, d) by the larger area of blue bars (short timescales) at current speeds >10 cm s^{-1}, in comparison to the area of red bars (long timescales) for the same speeds. In areas of strong advection by currents, such as the Gulf Stream, Kuroshio Current, or Antarctic Circumpolar Current, short correlation timescales of the phytoplankton biomass anomalies may be due to dispersal of phytoplankton driven by advection, assuming upstream phytoplankton communities are different from the location of interest. Similarly, short correlation timescales of the phytoplankton community also occur where horizontal mixing due to mesoscale and submesoscale processes facilitates dispersal, such as in Equatorial zones and Western Boundary Currents [*Abernathey & Marshall,* 2013; *Cole et al.,* 2012].

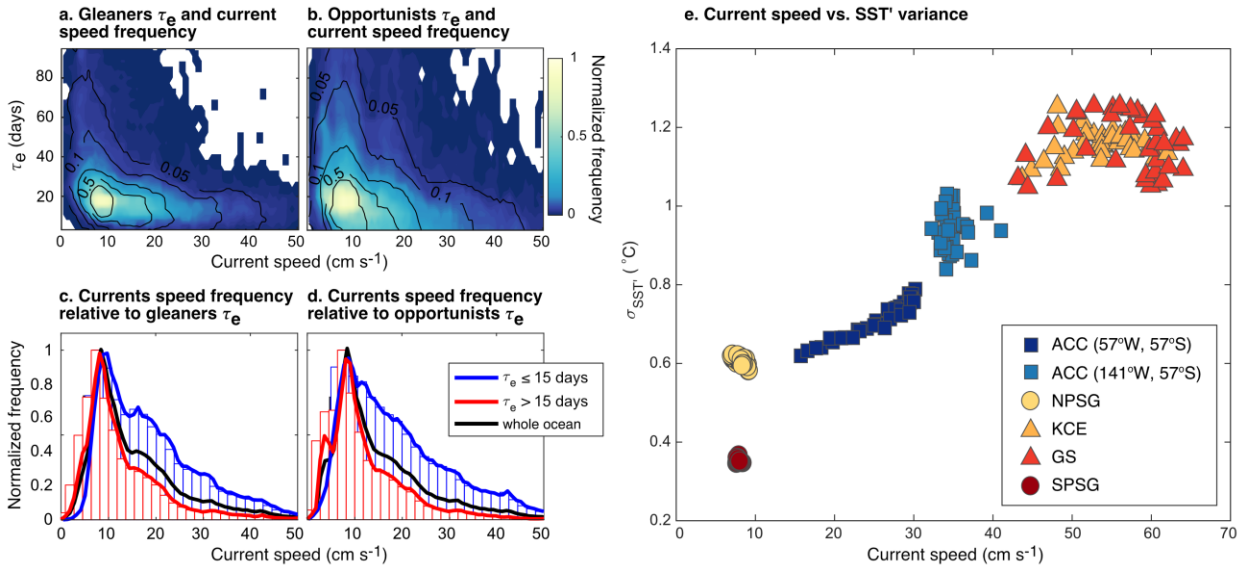


Figure 6. Relationship between current speed, timescales and environmental variability. Contour plots show the bi-variate normalized frequency distribution of correlation timescales with respect to current speed for **a.** gleaners and **b.** opportunists. Histograms show normalized frequency distributions of current speed and their corresponding probability density functions summarized for areas with short timescales (≤15 days) and long timescales (>15 days) for **c.** gleaners and **d.** opportunists. The probability density functions for the whole ocean (black line) are shown for reference. The distribution of current speeds for areas with short and long timescales are significantly different (Smirnoff-Kolmogorov test, $p<0.01$). In plots a. to d., the frequency of grid points is area-weighted, and normalized to the maximum frequency in the dataset. **e.** Relationship between current speed and variance in sea surface temperature anomalies at

six 1º by 1º illustrative regions: the Antarctic Circumpolar Current at the South Pacific and South Atlantic portions (ACC at 141ºW, 57ºS and 57ºW, 57ºS, respectively), the North Pacific Subtropical Gyre (NPSG), the South Pacific Subtropical Gyre (SPSG), the Kuroshio Current (KC) and the Gulf Stream (GS). Regions are color coded and ordered from colder (dark blue) to warmest (dark red), according to mean SST. The corresponding locations are indicated in Fig. 2c using the same markers.

In addition to the dispersal of organisms, both advection and mixing introduce environmental variability, including temperature, nutrients, and light, which can influence phytoplankton growth. Regions of stronger currents and EKE tend to have greater variability in SST anomalies [*Deser et al.,* 2010], for example. We find a similar pattern linking current speed to SST variance in our model (Fig. 6e). For 6 selected, illustrative regions, SST variance through time increases with mean current speed. Thus, the increased variability in the environment through time is likely to promote shorter correlation timescales within the phytoplankton populations. The present analysis in a Eulerian framework is, however, unable to differentiate whether the correlation timescales are more closely tied to rapid transport of organisms or intermittent favorable conditions for phytoplankton growth. We suggest that further study using a Lagrangian perspective will be useful to separate the different effects.

4.2.3 Correlation timescales and cell size

In order to evaluate why gleaners and opportunists have different correlation timescales, we consider the ecological importance of cell size. Cell size is a critical trait constraining important physiological rates, such as the growth rate [e.g., *Marañón et al.,* 2013], and predator – prey interactions [*B. Hansen et al.,* 1994; *P. J. Hansen et al.,* 1997]. We find that in vast ocean areas, large phytoplankton tend to have longer correlation timescales, while small phytoplankton tend to have shorter correlation timescales (Fig. 7b). In Fig. 7, red means that correlation timescale becomes shorter with increasing cell size (negative relationship), while blue means that correlation timescale becomes longer with increasing cell size (positive relationship). Because the number of species present in the community (here defined as phytoplankton phenotypes with abundance larger than 1% the total biomass) varies in space (Fig. 7a), this pattern linking cell size to correlation timescale spans a different number of cell sizes or phenotypes by location.

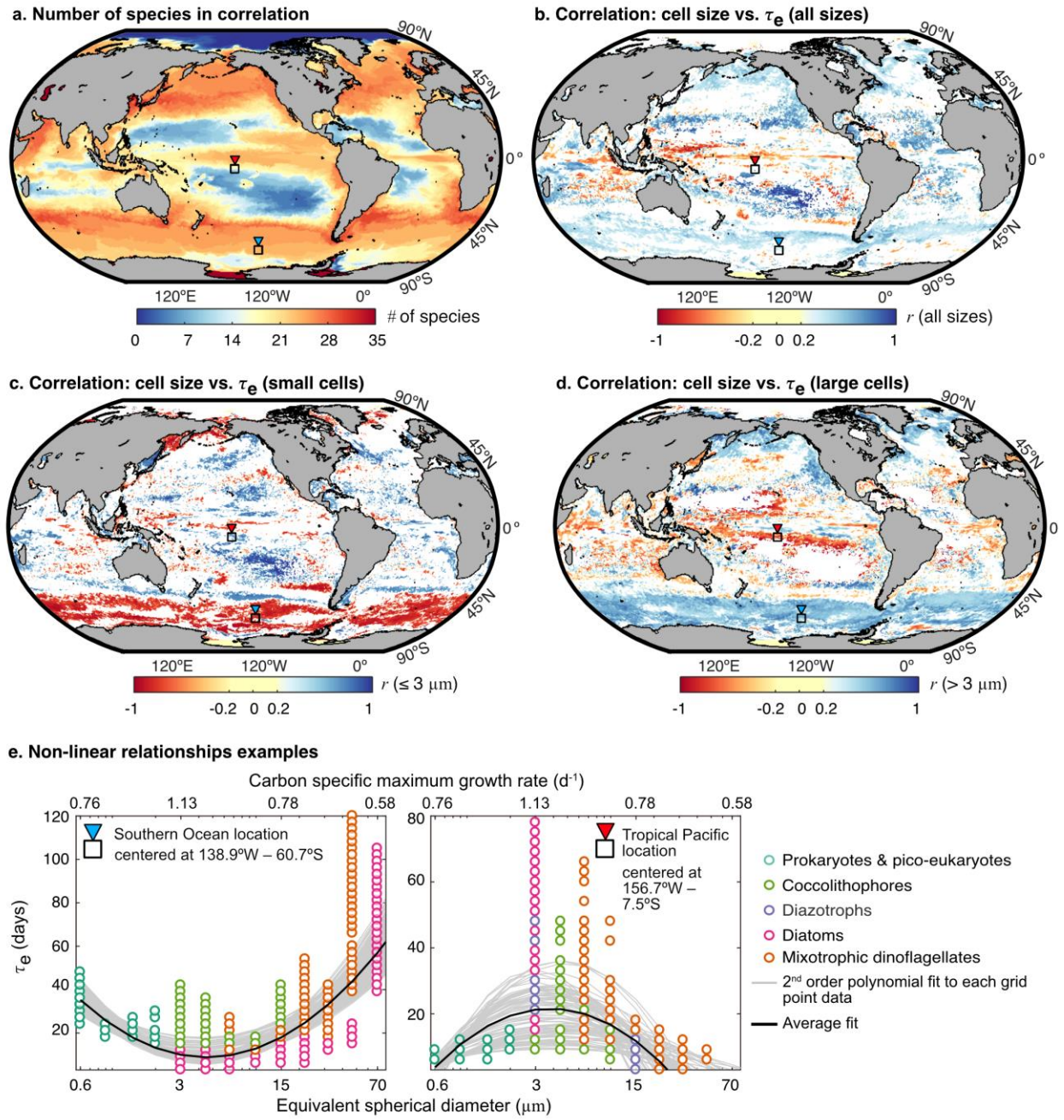


Figure 7. Analysis of the relationships between timescales and cell size. **a.** Total number of phytoplankton phenotypes ("richness"), defined as the number of phytoplankton phenotypes with abundance larger than 1% the total biomass in each grid point. Relationship (Pearson correlation coefficient, *r*) between cell size and correlation timescales (τ_e) for **b.** all cell sizes, **c.** only phytoplankton smaller ≤ 3μm, and **d.** only larger phytoplankton >3μm. Note that cell size increases logarithmically (Fig. S1). White patches in **b-d** mask areas where linear correlations are non-significant (p>0.1). Areas marked with a square and an inverted triangle mark the locations used for **e.** two scatter plots with contrasting examples of relationships between cell size (bottom x-axes) and correlation timescales. For reference, the

478 top x-axes show the corresponding carbon specific maximum growth rates (d^{-1}), which peak at a cell size
479 of 3 μm. The scatter plots show the timescales of all phenotypes at grid points in an area of 1º by 1º
480 centered at 138.9ºW 60.7ºS and 156.7ºW 7.5ºS (Southern Ocean and Central Equatorial Pacific,
481 respectively. Different phenotypes are identified with different marker colors. We fit a second-order
482 polynomial function to the relationship between cell size and τ_e relationships at each grid point (gray
483 lines). The thick black line shows the average fit for all grid points.

484 We separated the connections between cell size and correlation timescale into two size classes:
485 the 7 phenotypes that are 3μm or smaller, and the 28 phenotypes larger than 3μm (Fig. 7c, d). Though
486 there are many regions with no significant relationship between cell size and correlation timescale (white
487 areas in Figs. 7b-d), small and large phytoplankton phenotypes exhibit contrasting relationships between
488 correlation timescale and cell size in some regions. In the Southern Ocean, correlation timescales are
489 shorter with increasing cell size for small phytoplankton (≤3μm; Fig. 7c), but longer with increasing cell
490 size for large phytoplankton (>3μm; Fig. 7d). In the Tropical Pacific, correlation timescale is longer with
491 increasing cell size for small phytoplankton (≤3μm; Fig. 7c), but is shorter with increasing cell size for
492 larger phytoplankton (>3μm; Fig. 7d). Upon further inspection, we find that the relationship between cell
493 size and correlation timescale for locations in the tropics is characterized by a "humpback" shape or
494 convex curve, whereas the relationship for locations in the Southern Ocean is characterized by a "U"
495 shape or concave curve (Fig. 7e).

496 The distinct and spatially coherent regional patterns in the relationship between cell size and
497 correlation timescales, particularly the concave and convex patterns with an inflection point occurring at
498 around 3-5μm (Fig. 7e), suggest an underlying ecological origin. Phytoplankton in the 3-5μm size range
499 have higher growth rates than both smaller and larger phytoplankton [e.g., *Marañón et al.,* 2013], as
500 parameterized in our model. They also have relatively high nutrient and light affinities compared with
501 larger phytoplankton [e.g., *Edwards et al.,* 2012]. These opportunist phytoplankton are therefore well-
502 positioned to take advantage of pulses of nutrients, however scarce. In nutrient limited regions of the
503 ocean (such as the Tropical Pacific, Fig. 7e), we speculate that the relatively long correlation timescale of
504 opportunist phytoplankton may be caused by episodic blooms due to nutrient anomalies, for example
505 driven by the passage of eddies [*Chelton et al.,* 2011]. Microzooplankton predators ultimately graze down
506 the bloom, but returning to the pre-bloom biomass takes longer because the peak of the bloom of these
507 phenotypes is higher in magnitude. Phytoplankton larger and smaller than these fast-growing
508 opportunistic phytoplankton would not bloom to the same extent, either due to their higher nutrient

requirements or because they are grazed down more quickly than the blooming opportunist, thus yielding faster decorrelation timescales.

In regions with generally high nutrient supply (such as the Southern Ocean, Fig. 7e), the underlying dynamics may be different. The U-shaped pattern here may reflect three contrasting regimes for very small (0.6 - 2 μm), small (3 - 5 μm), and large phytoplankton (>5 μm). The smallest phytoplankton in this region have very low and relatively constant biomass (Fig. 3a). Small phytoplankton (3 - 5 μm) have rapid growth timescales, but they are also grazed readily by microzooplankton grazers, such that blooms do not persist for long. In contrast, large phytoplankton (>5 μm) have large microzooplankton grazers [e.g., *B. Hansen et al.,* 1994], which have lower specific ingestion and growth rates than smaller microzooplankton [e.g., *P. J. Hansen et al.,* 1997], as captured in the model parameterization [*Ward et al.,* 2012]. In habitats such as this, the growth of the largest phytoplankton would be periodically decoupled from their grazers, allowing a positive biomass anomaly to persist for a longer duration than for smaller phytoplankton. This may explain the very long correlation timescales for the largest phytoplankton types. Thus, the concave and convex shaped patterns in Fig. 7e may reflect the relative importance of distribution patterns (Fig. 3c, d), resource acquisition, growth traits, and predator-prey imbalances in shaping the persistence of phytoplankton anomalies.

4.3 Spatial scales of correlation

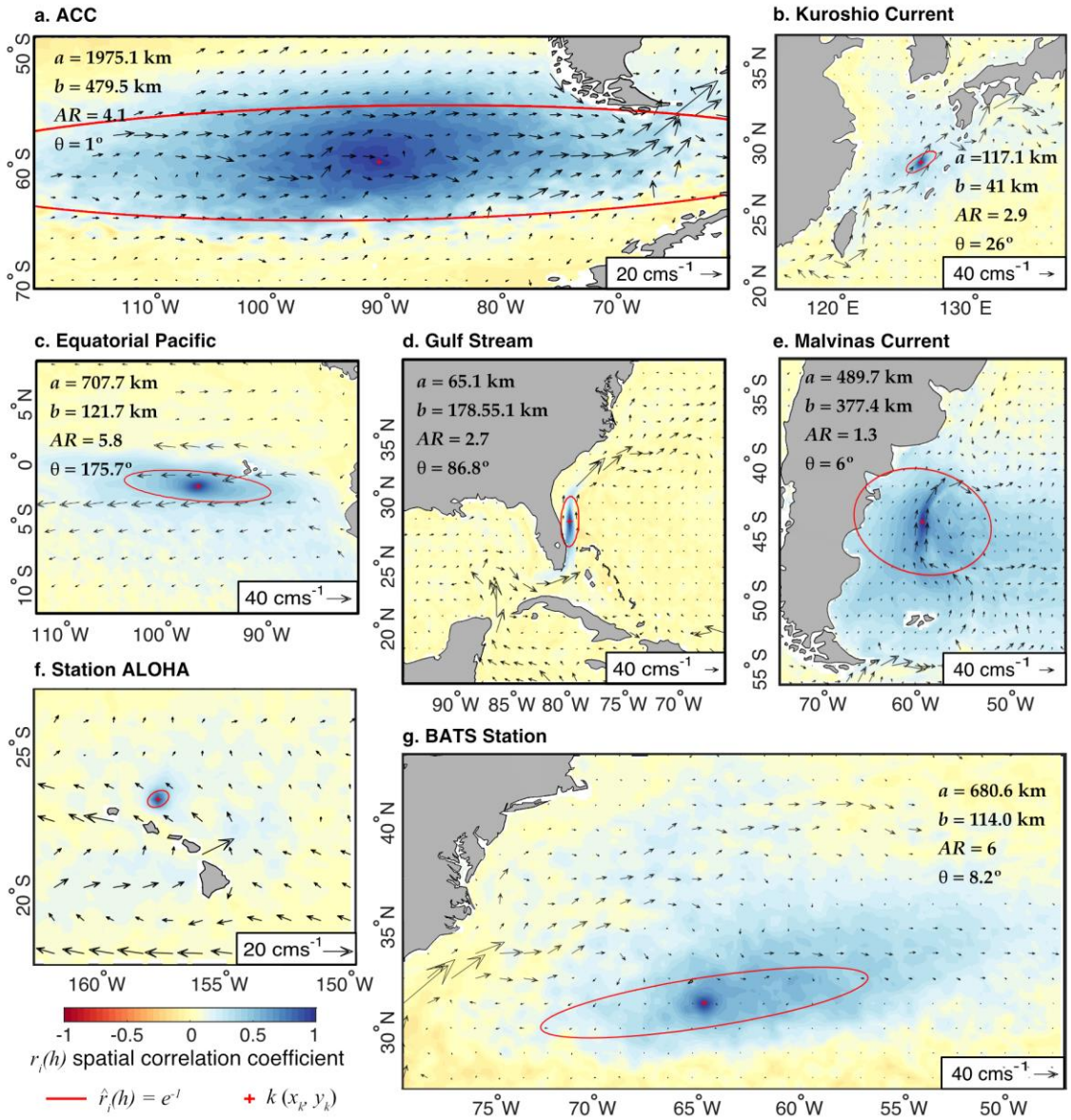
4.3.1 Length scales and ocean dynamics

Over what distance do phytoplankton populations vary synchronously? To answer this, we calculated the spatial correlation structure of phytoplankton populations on a point by point and phenotype by phenotype basis with no time lags. A feature of our analysis is that we are able to quantify the shape of the correlation fields, not just the correlation length scales. A key shape property is the aspect ratio of spatial correlation structure. If the aspect ratio is equal to one, (AR=1), the spatial correlation is roughly equal in all directions; we term this isotropic. Isotropic correlation structures may be due to horizontal mixing homogenizing local properties, such as nutrient concentrations and temperature, or equal dispersal of organisms in all directions. An elongated spatial correlation field has AR>1, possibly indicating the presence of strong advection or persistent frontal zones (Fig. 8). We call this type of correlation structure anisotropic.

In general, spatial correlation patterns are similar for gleaners and opportunists (Fig. 9a-d, S8). The longest correlation scales (major axis) occur in the Subpolar oceans exceeding 2000 km (Fig. 8a, Fig. 9a, b), and coincide roughly with the position of strong frontal zones in both hemispheres, and with the

location of the ACC in the Southern Ocean (Fig. 2b-d). Major axis length scales range between 150 km and 500 km in the Equatorial regions (Fig. 8c) and between 150 km and 250 km along boundary currents (Fig. 8b, d). The shortest major axis scales, up to 100 km, occur in Subtropical regions (Fig. 8f) and in some areas near Western Boundary Currents and their corresponding extensions (Fig. 9a, b). Previous studies have suggested long correlation length scales in phytoplankton can be achieved via rapid advection and turbulent mixing [*Lévy et al.,* 2014]; where as long correlation length scales in SST have been also attributed to synoptic forcing over large distances leading to uniform conditions in the upper ocean [*Hosoda and Kawamura*, 2005]. The minor axes are, by definition, shorter than the major axes (Fig. 9c-d), but exhibit similar spatial patterns as for major axes correlation length. In general, the orientation of the correlation structures aligns well with the direction of the mean flow (Fig. 8).

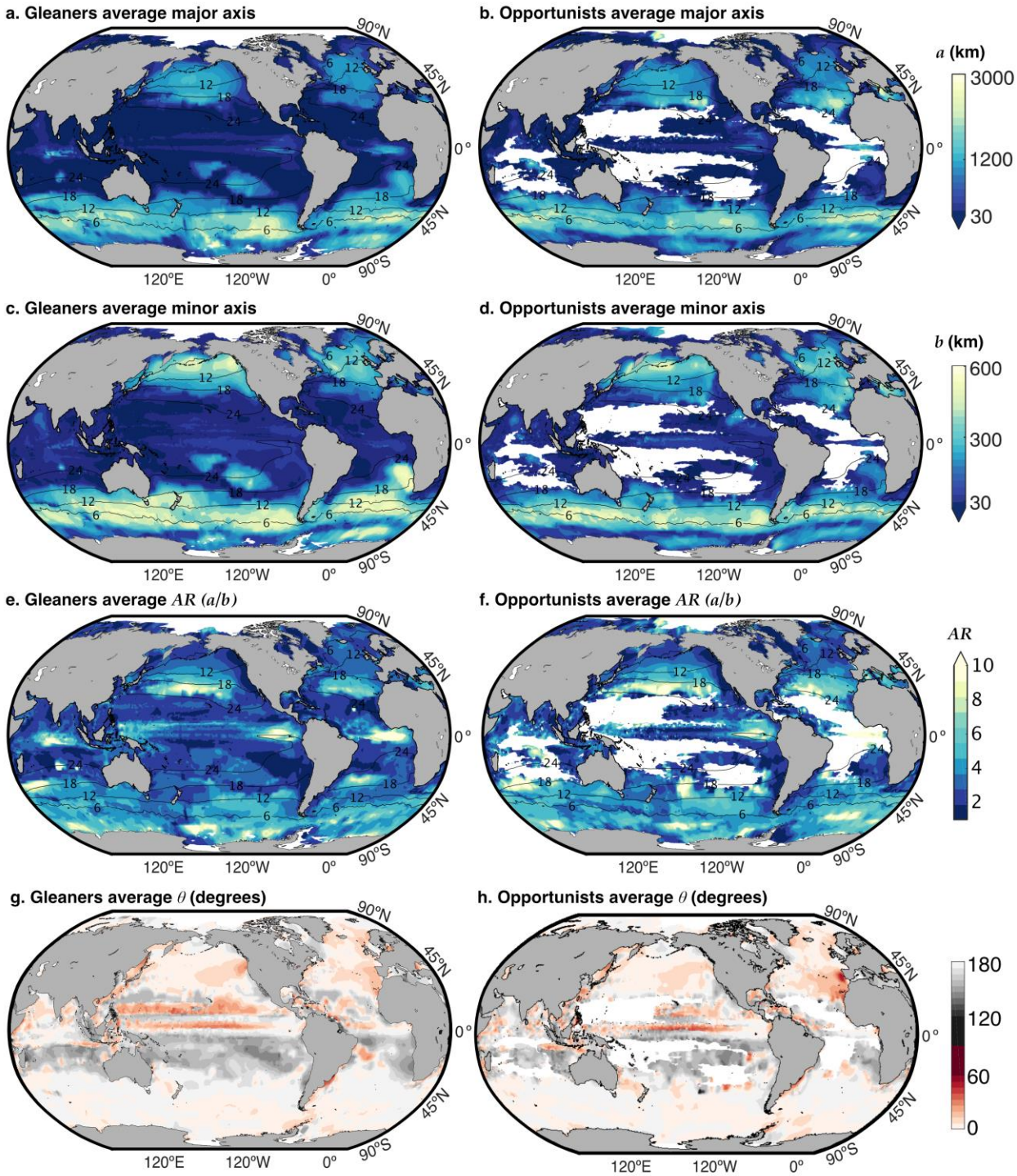
We find that, regardless of regional differences in correlation length scales, the spatial correlations of phytoplankton biomass anomaly in the ocean is strongly anisotropic (Fig. 9e, f). The total ocean area with isotropic correlation fields (AR=1) is very small and aspect ratios below 2:1 (AR<2) represent only 8.5% and 12.3% of the global ocean for gleaners and opportunists, respectively. The median AR for gleaners and opportunists is 2.9 and 3.1, respectively (Fig. S8). High aspect ratios ranging between 2.5 to 5 occur along the Antarctic Circumpolar Current and Western Boundary Currents (Fig. 8a, c, d). In some cases, the elongation of the correlation contours due to the presence of an ocean current is most obvious near the core of the current, but decreases rapidly away from the center of the current (Fig. 8e). The effect of some of the narrow boundary currents is apparent in the Alguhas Current flowing southward along the east coast of South Africa, the Kuroshio Current to the southwest of Japan, and the Somali Current along the coast of Somalia and Oman in the western Indian Ocean (Fig. 9e, f). The most elongated shapes (AR>6) are found within the Eastern Equatorial regions (Fig. 8c), and roughly coinciding with major extratropical ocean fronts (Fig. 9e, f), including the Subpolar Front in the North Pacific (approximately located between 40ºN and 45ºN, at isotherms ranging from 9ºC to 18ºC [*Yuan & Talley,* 1996]), the Subpolar Front in the North Atlantic (which typically follows the NAC, and is approximately located south of the 18ºC SST contour in Figs. 9e, f), and the region between the Southern and Subantarctic Fronts in the Southern Ocean (typically defined by the 6 ºC and 12ºC SST contours), including the Polar Front [*Carter et al.,* 2008]). The meridional length scales (i.e., minor axis, orthogonal to the front) become shortened in the vicinity of the front, while zonal length scales are long. For example, in the South Pacific sector of the ACC (Fig. 8a, Fig. 9), the average flow of the ACC is roughly west to east and AR=4.



572

573 **Figure 8.** Spatial correlation structures (color background) and corresponding 2D Gaussian fit (red ellipse
574 marking the $\hat{r} = e^{-1}$ contour) obtained from the biomass anomalies of the smallest gleaner in key regions
575 of interest. The central grid point *k* for each spatial correlation structure is marked with a red + symbol.
576 Mean velocity vectors are shown in black, emphasizing the elongation of correlation length scales along
577 currents at **a.** the Antarctic Circumpolar Current, **b.** the Kuroshio Current, **c.** the eastern Equatorial
578 Pacific, **d.** the Gulf Stream, and **e.** the Malvinas Current. Spatial correlation for the long-term
579 observational stations ALOHA (A Long-term Oligotrophic Habitat Assessment) in Hawaii, and BATS
580 (Bermuda Atlantic Time-series Study) are shown in **f.** and **g.**, respectively. Each spatial correlation
581 structure details its corresponding axes length scales (a and b), the major-to-minor aspect ratio ($AR = max(a,b)/min(a,b)$), and the angle of orientation (θ).
582

583



584

585 **Figure 9.** Spatial correlation scales of gleaners and opportunists: **a., b.** major axis, **c., d.** minor axis, **e., f.**
586 major-to-minor aspect ratio, and **g. h.** angles. Angles are reported from 0º to 180º, such that meridional

patterns correspond to 90º and zonal patterns correspond to either 0º or 180º. White areas are where opportunist biomass is very low (group biomass below 1% of total phytoplankton biomass).

Our estimates of phytoplankton population correlation length scales broadly agree with regional length scales estimated from physical and chemical variables from the same model simulation (Fig. S6, S7) and from previous studies [*Hosoda & Kawamura,* 2005; *Kessler et al.,* 1996; *Kuragano & Kamachi,* 2000; *Mazloff et al.,* 2018]. In general, phytoplankton and environmental correlation length scales are long in the Southern Ocean and Equatorial Pacific, but shorter in western boundary currents and subtropical regions. A recent analysis of satellite derived and simulated values of low-passed filtered SSH, SST, heat and carbon content in the Southern Ocean suggests large correlation length scales on the order of 500 km to 4000 km zonally and 500 and 1000 km meridionally [*Mazloff et al.,* 2018]. Zonal length scales from *in situ* SST and thermocline depth in the Equatorial Pacific are about 10.2º (~1110 km) and 11.6º (~1280 km), respectively [*Kessler et al.,* 1996]. In the Kuroshio region, the spatial correlation analysis of satellite SST revealed correlation length scales between 78 and 230 km, with the smallest length scales observed in the most dynamical regions [*Hosoda and Kawamura,* 2005]. In a global analysis using TOPEX-POSEIDON SSH anomaly data from 1993 to 1996, *Kuragano and Kamachi* [2000] found large zonal length scales in the Equatorial region (1300 – 2600 km), as well as in the Subpolar gyres (470 – 960 km) and small length scales in the Subtropical gyres and boundary currents (100 – 260 km). Though we do not directly correlate the length scales of physical variables to the modeled phytoplankton variables in this study, the available evidence suggests that the gradients in correlation length scales are qualitatively similar.

4.3.1 Length scales and current speed

As with the earlier discussion of correlation timescales, we now consider how correlation length scales may be tied to advection. In this case, the relationship between current speed and length scales or aspect ratio is complicated due to the many possible drivers of spatial correlation structure.

Very large length scales and aspect ratio can occur in areas of the ocean with relatively low average current speed, due to large-scale uniform synoptic forcing [e.g., *Hosoda and Kawamura,* 2005]. Turbulent mixing may also disperse and homogenize phytoplankton further away from the region of direct influence of an ocean current. Western Boundary currents have elongated spatial correlation fields, but their length scales are shorter and aspect ratios smaller. For example, the correlation structure in the poleward flowing Gulf Stream (Fig. 8d) has a smaller aspect ratio than the zonal ACC (Figs 8a). This

suggests that when the direction of the correlation structure is across strong environmental and biomass gradients, such as the Gulf Stream flowing northward from tropical to temperate waters (Fig. 8d), synoptic atmospheric forcing and eddy activity along the current edges modify the major and minor axes. This creates overall smaller and less elongated correlation structures than in zonal currents. The short minor axis length scales in Western Boundary Currents also reflect the limited cross-jet exchange with the adjacent water masses.

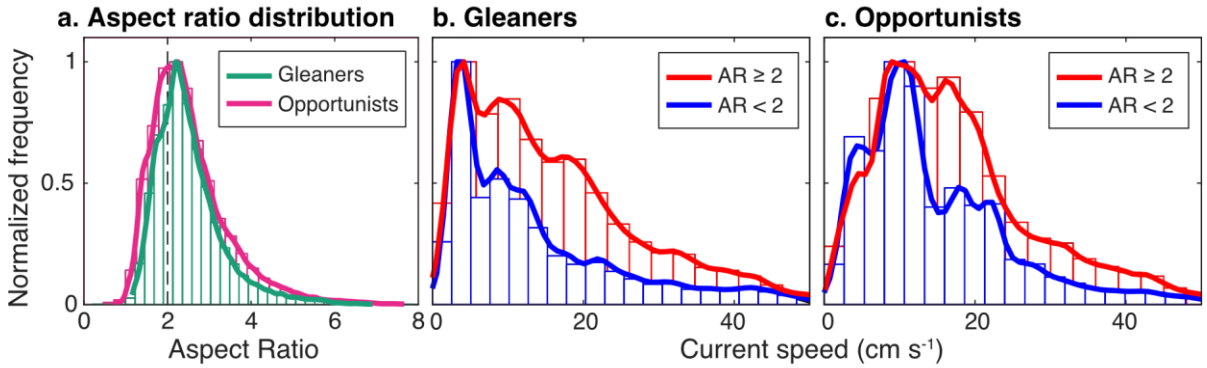


Figure 10 a. Frequency distribution of global aspect ratios from biomass anomalies of gleaners and opportunists for areas with length scales below 300 km. The 25th percentile of both distributions is found at approximately AR = 2 (black dashed line). This value is used as a threshold for the comparisons of frequency distributions of current speed in areas with very elongated correlation structures (AR ≥ 2) and more isotropic structures (AR < 2) for gleaners (**b.**) and opportunists (**c.**). Distributions of current speed with AR ≥ 2 and AR < 2 are significantly different (Smirnoff-Kolmogorof test $p < 0.01$).

If we focus on only those regions with a major axis length scale less than 300 km (i.e., the maximum length scales found at boundary currents), a clearer picture of the relationship between current speed and the elongation of correlation structures emerges. This subset of the global ocean represents 51.2% and 59.7% of the ocean area for gleaners and opportunists, respectively. The subset includes western boundary currents, coastal upwelling areas and subtropical gyres, but excludes the Southern Ocean, open waters in the North Pacific, the central and eastern North Atlantic, and certain equatorial areas. Differences in the distribution of AR for gleaners and opportunists are negligible for this subset (Fig. 10a), as for the entire global results (Fig. S8). The median AR of the subset is 2.36 for gleaners and 2.34 for opportunists, and the first quartile is approximately AR = 2 for both types of phytoplankton (Fig. 10a). We use this threshold to distinguish regions with more or less anisotropic correlation structures. Elongated correlation structures (AR > 2) are more common in regions of relatively high current speed, and more isotropic correlations structures (AR < 2) are more common in regions of relatively low

currents speeds (Fig. 10b, c). Thus, we conclude that current speed is an important factor affecting the spatial correlation of phytoplankton communities in this subset of the global ocean. Nevertheless, some regions of strong advection have low aspect ratios. This effect occurs mainly in the eastward extensions of Western Boundary Currents, which are ocean regions characterized by large numbers of eddies [*Chelton et al.,* 2011]. In these eddy dominated areas, the dispersing effect of turbulent mixing in all directions is a relatively important driver of tracer distributions compared with other regions of the ocean.

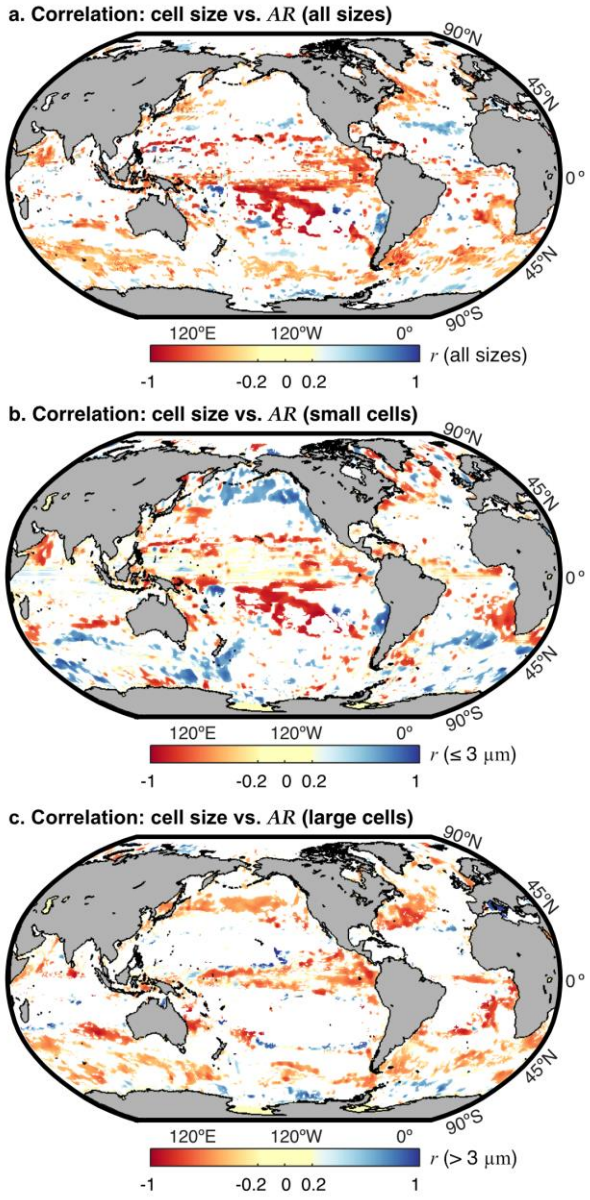


Figure 11 Analysis of the relationship between cell size and aspect ratio (*AR*) derived from length scales of correlation: **a.** Pearson correlation coefficient (*r*) of the relationship between cell size and *AR*

652 considering all significantly abundant phytoplankton phenotypes. The number of the significantly
653 abundant phytoplankton phenotypes used in the correlation is the same as Figure 7a.; **b.** Same analysis as
654 panel a., but considering only phenotypes ≤ 3μm; **c.** Same analysis as panel a., but considering only
655 phenotype > 3μm.

656 **4.3.2 Length scales and cell size**

657 Smaller phytoplankton have, on average, larger aspect ratios than larger phytoplankton cells (Fig.
658 11a; indicated by areas of negative correlation, in red, between aspect and cell size). The negative
659 relationship between cell size and aspect ratio is most apparent in the Southern Ocean, Western Boundary
660 Currents, and the Subtropics (Fig.11a). Similar spatial patterns are observed when we analyze large
661 phytoplankton (≥ 3μm; Fig. 11c), with mostly negative correlations between cell size and aspect ratio.
662 For small cells (< 3μm; Fig. 11b), correlations between aspect ratio and cell size are negative in the
663 tropics but more varied and equivocal elsewhere, likely due to the fewer number of taxa included in this
664 category. We speculate that the overall negative relationship between aspect ratio and cell size may be
665 related to the abundance of each phytoplankton phenotype, which generally decreases with size. Smaller
666 phytoplankton have larger populations and disperse greater distances than do larger phytoplankton
667 [*Villarino et al.,* 2018].

668 **5 Discussion and conclusions**

669 In this study, we estimated global timescales and spatial scales of correlation in a state-of-the-art
670 physical-ecosystem ocean model, with 35 simulated phytoplankton phenotypes covering a broad range of
671 phytoplankton functions and sizes. By calculating and using biomass anomalies, we focus on the intra-
672 annual variability of phytoplankton biomass, which is more likely to reflect the effects of dispersal. The
673 correlation timescale analysis provides information about the persistence of anomalies and the stability of
674 the planktonic system. The results from the spatial correlation analysis indicate the extent of regions in
675 the ocean which act in concert, driven by advection, mixing, synoptic events, or a mixture of all these
676 drivers.

677 Although the model we use here is unique in its ecosystem complexity, it still only represents a
678 few tens of phytoplankton phenotypes compared to the many thousands of species in the real ocean.
679 Additionally, the model resolution is about 18 km, which permits the development of eddies and narrow
680 currents, but does not capture sub-mesoscale processes. Sub-mesoscale processes are also likely to affect
681 both spatial and temporal timescales of phytoplankton communities, and understanding their impacts
682 would require investigating model output at much higher spatial resolution and temporal frequency.

683 Even though the model resolution is nominally 18 km, coastal regions and inland seas may not be
684 simulated with fidelity. Therefore we focus our interpretation of the results on the pelagic ocean.

685 The results presented in this study should be considered only as a first step in defining the
686 phytoplankton community correlation scales globally. Previous studies of biogeochemically important
687 scales in the oceans have been mostly limited to satellite chlorophyll [*Doney et al.,* 2003; *Fuentes et al.,*
688 2000; *Glover et al.,* 2018]. These previous studies have largely focused on quantifying the impact of
689 mesoscale features, and using different methodologies to overcome data gaps due to the presence of
690 clouds. In a different type of study, *Henson et al.,* [2016] used model output from Earth System Models to
691 explore the role of temporal and spatial scales on trend detection in several biogeochemical variables (e.g.
692 chlorophyll, primary production, pH), and their implications for Earth observing systems. In that study,
693 the footprint of existing and planned timeseries stations were obtained based on statistical similarity in
694 terms of means and variability for surrounding grid cells.

695 Here, for the first time, we instead look at timescales and spatial scales of correlation of
696 ecological variables, and we provide robust point by point evaluations of these scales. Our results can thus
697 provide information for any future observing system design from an ecologically relevant perspective. For
698 instance, we find that regions with fast currents are more likely to exhibit short correlation timescales
699 (Figs. 4-6). Conversely, in more quiescent regions such as the cores of the subtropical ocean gyres,
700 phytoplankton anomalies persist for long periods. As such, observing these different systems would
701 require different sampling strategies: much more frequent in the former than the latter. In the northern
702 hemisphere high latitudes, mesoscale currents and eddy activity imprint a signal of short timescales (Fig.
703 5 a, c), again suggesting that monitoring systems in these regions would require more frequent sampling.
704 Phytoplankton timescales are shorter than the timescales of physical and chemical variables that control
705 phytoplankton growth, such as temperature and nitrate. In the case of temperature, the large heat capacity
706 of the ocean imprints a general pattern of long timescales on the order of several months. In the case of
707 nitrate, the decoupling of timescales is prevalent at high latitudes where light, rather than nutrients, is a
708 seasonal limiting factor for phytoplankton growth.

709 Monitoring of different types of phytoplankton is also likely to be complicated. The relationships
710 between correlation timescales and cell size are noisy and complex, but exhibit spatial coherence globally
711 suggesting underlying dynamical or ecological origins. Overall, we find that larger phytoplankton tend to
712 have longer correlation timescales than smaller phytoplankton (Fig. 7b), but there are contrasting patterns
713 for the smallest ($\leq$3μm) and largest (>3μm) phytoplankton, as well as by region (Fig. 7c-e). We separate
714 phytoplankton into these two size categories as observations suggest that there are contrasting patterns of

correlation between growth rates and size: growth rates increase with cell size in the smaller category and decrease with size in the larger category [e.g. *Marañón et al.,* 2013]. This characteristic is also captured in the model configuration and emergent in the correlation timescales results (Fig 7e). In the Southern Ocean, correlation timescales decrease with increasing body size for small phytoplankton (negative relationship), but increase with body size for larger phytoplankton (positive relationship). Conversely, in the subtropics correlation timescales increase with body size for small phytoplankton (positive relationship), but decrease with increasing body size for larger phytoplankton (negative relationship). The relationships can be thus characterized by a "humpback" shape or convex curve in the subtropics, and by a "U" shape or concave curve in the Southern Ocean (Fig. 7e). These patterns may reflect trade-offs between resource acquisition and predation. In the model, the smaller category has nutrient affinity decreasing with cell size [e.g., *Edwards et al.,* 2012], the growth rate increasing [*Marañón et al.,* 2013], and the predators grazing rates decreasing [*Hansen et al.,* 1997]. In the larger size class, nutrient affinity, grazing, and growth rate all decline with cell size [e.g., *Edwards et al.,* 2012; *B. Hansen et al.,* 1994; *Marañón et al.,* 2013]. These factors lead to different regional distributions, with the smaller category having a much more regionally uniform distribution (Fig. 3). Monitoring the differences in correlation scales between phytoplankton phenotypes could therefore help us understand some of the major controlling mechanisms across sizes. These differential relationships between cell size and correlation timescales also imply that sampling frequency may need to be different for different phytoplankton phenotypes.

Our study of correlation length scales provides a mechanism to understand how far (in distance) a single station observation can provide insight into phytoplankton community dynamics. We find that the global ocean is predominantly anisotropic (Fig. 9, S8). The strongest anisotropic features are zonal, along the equatorial region, in the ACC, and along major ocean fronts. In such regions an observational site will provide context and insight for extensive regions in the east-west direction, but less insight to the north and south. Elongated spatial correlation fields also occur in Western Boundary Currents, but their length scales are shorter because of strong eddy mixing, limited cross-jet exchange, and meridional variations in local forcing, such as heat fluxes. Thus, biological measurements taken within a current jet provide insight about processes along-flow, but very little information of across-current processes. More isotropic correlation structures are present where current speeds are low and where eddies or recirculation disperse phytoplankton equally in all directions. In contrast with correlation timescales (Fig. 7b-e), the correlation length scales have a predominantly negative relationship with cell size across most areas of the ocean (Fig. 11). This implies that resolving spatial dynamics of large cells requires denser spatial observations than for small cells, and that this relationship does not vary strongly in space.

Our results could also be a starting place to explore implications for our existing observing systems. For instance, the correlation timescales for total biomass at ALOHA and BATS are 13 and 6 days. Given that sampling at these stations is monthly, our results suggest that transient non-seasonal changes to their phytoplankton communities are not adequately captured by the current sampling strategy (though clearly seasonal and interannual variability are captured by these sites). The spatial correlation structures and fitted ellipses we calculated for these locations (Fig. 8f, g) indicate the regions that will have similar responses to what is seen at these timeseries sites. For a small prokaryote, the major axis length scales we estimate for these two locations are 41.1 and 680.6 km. Thus, suggesting that any variability seen at the timeseries may be relevant only over these spatial scales, and the stations do not provide insight into the full subtropical gyres as is sometimes assumed.

This study thus offers a quantitative, global-scale estimation of the temporal and spatial scales of correlation in phytoplankton communities, with a unique ecological perspective that cannot be obtained with current observations alone. Its results provide unprecedented background information to explore regional differences, as well as differences between sub-populations of the planktonic system. Understanding the scales of correlation of different phytoplankton phenotypes can also influence our understanding of global patterns of distribution of their predators, including zooplankton and larval fish. As in the case of correlation scales from physical ocean variables, this information can potentially aid in the design of biological ocean observing networks and monitoring campaigns by guiding decisions about optimal sampling frequency and distance between monitoring stations in different regions. This is particularly important as new biological sampling methods capable of acquiring massive amounts of data, such as genomics and biogeochemical-Argo, become more widely used in the study of marine microbial ecology.

Acknowledgements

Model data used in this study is publicly available at the *UC San Diego Library Digital Collections:* https://doi.org/10.6075/J0BR8QJ1. Additional model data and visualization tools are available through the following OPeNDAP server: http://engaging-opendap.mit.edu:8080/las/UI.vm, data set Darwin v0.2 cs510. We acknowledge funding from NASA grant 80NSSC17K0561 (to SD and OJ); the Moore/Sloan Data Science and Washington Research Foundation Innovation in Data Science Fellowship (to SC); NSF grant OCE-1638834 and NSF grant OPP-1543245 (to TAR); NASA grant NNX16AH67G, NSF grant OPP-1750035 and NSF PLR-1425989 (to MRM); and from the Simons Foundation (to ADB). We thank our editor M. Friedrichs and three anonymous reviewers whose comments helped improving this manuscript.

References

782 Abernathey, R. P., & Marshall, J. (2013). Global surface eddy diffusivities derived from satellite
783 altimetry. *Journal of Geophysical Research: Oceans*, *118*(2), 901–916.
784 https://doi.org/10.1002/jgrc.20066

785 Acevedo-Trejos, E., Brandt, G., Bruggeman, J., & Merico, A. (2015). Mechanisms shaping size structure
786 and functional diversity of phytoplankton communities in the ocean. *Scientific Reports*, *5*(1).
787 https://doi.org/10.1038/srep08918

788 Adjou, M., Bendtsen, J., & Richardson, K. (2012). Modeling the influence from ocean transport, mixing
789 and grazing on phytoplankton diversity. *Ecological Modelling*, *225*, 19–27.
790 https://doi.org/10.1016/j.ecolmodel.2011.11.005

791 Barton, A. D., Dutkiewicz, S., Flierl, G., Bragg, J., & Follows, M. J. (2010). Patterns of Diversity in
792 Marine Phytoplankton. *Science*, *327*(5972), 1509–1511. https://doi.org/10.1126/science.1184961

793 Benoiston, A.S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., & Bowler, C. (2017).
794 The evolution of diatoms and their biogeochemical functions. *Philosophical Transactions of the*
795 *Royal Society B: Biological Sciences*, *372*(1728), 20160397.
796 https://doi.org/10.1098/rstb.2016.0397

797 Boss, E., Swift, D., Taylor, L., Brickley, P., Zaneveld, R., Riser, S., Perry, M. J., & Strutton, P.G., (2008).
798 Observations of pigment and particle distributions in the western North Atlantic from an
799 autonomous float and ocean color satellite. *Limnology and Oceanography, 53*(5, part 2), 2112 –
800 2122.

801 Bracco, A., Clayton, S., & Pasquero, C. (2009). Horizontal advection, diffusion, and plankton spectra at
802 the sea surface. *Journal of Geophysical Research*, *114*(C2).
803 https://doi.org/10.1029/2007JC004671

804 Bruland, K. W., Rue, E. L., & Smith, G. J. (2001). Iron and macronutrients in California coastal
805 upwelling regimes: Implications for diatom blooms. *Limnology and Oceanography*, *46*(7), 1661–
806 1674. https://doi.org/10.4319/lo.2001.46.7.1661

807 Carter, L., McCave, I. N., & Williams, M. J. M. (2008). Chapter 4 Circulation and Water Masses of the
808 Southern Ocean: A Review. In *Developments in Earth and Environmental Sciences* (Vol. 8, pp.
809 85–114). Elsevier. https://doi.org/10.1016/S1571-9197(08)00004-9

810 Chelton, D. B., Wentz, F. J., Gentemann, C. L., de Szoeke, R. A., & Schlax, M. G. (2000). Satellite
811 microwave SST observations of transequatorial tropical instability waves. *Geophysical Research*
812 *Letters*, *27*(9), 1239–1242. https://doi.org/10.1029/1999GL011047

Chelton, D. B., Gaube, P., Schlax, M. G., Early, J. J., & Samelson, R. M. (2011). The Influence of Nonlinear Mesoscale Eddies on Near-Surface Oceanic Chlorophyll. *Science*, *334*(6054), 328–332. https://doi.org/10.1126/science.1208897

Clayton, S., Dutkiewicz, S., Jahn, O., & Follows, M. J. (2013). Dispersal, eddies, and the diversity of marine phytoplankton. *Limnology and Oceanography: Fluids and Environments*, *3*(1), 182–197. https://doi.org/10.1215/21573689-2373515

Cole, H., Henson, S., Martin, A., & Yool, A. (2012). Mind the gap: The impact of missing data on the calculation of phytoplankton phenology metrics. *Journal of Geophysical Research*, *117*(C08030).

Denman, K. L., & Abbott, M. R. (1994). Time scales of pattern evolution from cross-spectrum analysis of advanced very high-resolution radiometer and coastal zone color scanner imagery. *Journal of Geophysical Research*, *99*(C4), 7433. https://doi.org/10.1029/93JC02149

Deser, C., Alexander, M. A., Xie, S.-P., & Phillips, A. S. (2010). Sea Surface Temperature Variability: Patterns and Mechanisms. *Annual Review of Marine Science*, *2*(1), 115–143. https://doi.org/10.1146/annurev-marine-120408-151453

Doney, S. C., Glover, D. M., McCue, S. J., & Fuentes, M. (2003). Mesoscale variability of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite ocean color: Global patterns and spatial scales. *Journal of Geophysical Research: Oceans*, *108*(C2). https://doi.org/10.1029/2001JC000843

Dutkiewicz, S., Follows, M., & Bragg, J. G. (2009). Modelling the coupling of ocean ecology and biogeochemistry. *Global Biogeochemical Cycles*, *23*(GB4017). https://doi.org/10.1029/2008GB003405

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., & Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences*, *12*(14), 4447–4481. https://doi.org/10.5194/bg-12-4447-2015

Dutkiewicz, S., Cermeno, P., Jahn, O., Follows, M. J., Hickman, A. E., Taniguchi, D. A. A., & Ward, B. A. (2019). Dimensions of Marine Phytoplankton Diversity. *Biogeosciences Discussions*, 1–46. https://doi.org/10.5194/bg-2019-311

Edwards, K. F., Thomas, M. K., Klausmeier, C. A., & Litchman, E. (2012). Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton. *Limnology and Oceanography*, *57*(2), 554–566. https://doi.org/10.4319/lo.2012.57.2.0554

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. G. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, *281*, 237–240.

844 Frankignoul, C. (1985). Sea surface temperature anomalies, planetary waves, and air-sea feedback in the
845 middle latitudes. *Reviews of Geophysics*, *23*(4), 357. https://doi.org/10.1029/RG023i004p00357

846 Frankignoul, C., & Hasselmann, K. (1977). Stochastic climate models, Part II Application to sea-surface
847 temperature anomalies and thermocline variability. *Tellus*, *29*(4), 289–305.
848 https://doi.org/10.1111/j.2153-3490.1977.tb00740.x

849 Fuentes, M., Doney, S. C., Glover, D. M., & McCue, S. J. (2000). Spatial structure of the SeaWiFS ocean
850 color data for the North Atlantic Ocean. In *Studies in the Atmospheric Sciences* (pp. 153–171).
851 Springer.

852 Fuhrman, J. A., Steele, J. A., Hewson, I., Schwalbach, M. S., Brown, M. V., Green, J. L., & Brown, J. H.
853 (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National*
854 *Academy of Sciences*, *105*(22), 7774–7778. https://doi.org/10.1073/pnas.0803070105

855 Gaillard, F., Autret, E., Thierry, V., Galaup, P., Coatanoan, C., & Loubrieu, T. (2009). Quality Control of
856 Large Argo Datasets. *Journal of Atmospheric and Oceanic Technology*, *26*(2), 337–351.
857 https://doi.org/10.1175/2008JTECHO552.1

858 Glover, D. M., Doney, S. C., Oestreich, W. K., & Tullo, A. W. (2018). Geostatistical Analysis of
859 Mesoscale Spatial Variability and Error in SeaWiFS and MODIS/Aqua Global Ocean Color Data.
860 *Journal of Geophysical Research: Oceans*, *123*(1), 22–39. https://doi.org/10.1002/2017JC013023

861 Han, W., Webster, P. J., Lin, J.-L., Liu, W. T., Fu, R., Yuan, D., & Hu, A. (2008). Dynamics of
862 Intraseasonal Sea Level and Thermocline Variability in the Equatorial Atlantic during 2002–03.
863 *Journal of Physical Oceanography*, *38*(5), 945–967. https://doi.org/10.1175/2008JPO3854.1

864 Hansen, B., Bjørnsen, P. K., & Hansen, P. J. (1994). The size ratio between planktonic predators and their
865 prey. *Limnology and Oceanography*, *39*(2), 395–403. https://doi.org/10.4319/lo.1994.39.2.0395

866 Hansen, P. J., Bjørnsen, P. K., & Hansen, B. (1997). Zooplankton grazing and growth: Scaling within the
867 2-2,-µm body size range. *Limnology and Oceanography*, *42*(4), 687–704.
868 https://doi.org/10.4319/lo.1997.42.4.0687

869 Henson, S. A., Beaulieu, C., & Lampitt, R. (2016). Observing climate change trends in ocean
870 biogeochemistry: when and where. *Global Change Biology*, *22*(4), 1561–1571.
871 https://doi.org/10.1111/gcb.13152

872 Hosoda, K., & Kawamura, H. (2005). Seasonal Variation of Space/Time Statistics of Short-Term Sea
873 Surface Temperature Variability in the Kuroshio Region. *Journal of Oceanography*, *61*(4), 709–
874 720. https://doi.org/10.1007/s10872-005-0078-3

875 Jahn, O., Hill, C., Dutkiewicz, S., Follows, M. (2019). MITgcm 3-daily global sea surface temperature,
876 ocean currents, nitrate and phytoplankton biomass (1992 - 2016). *UC San Diego Library Digital*
877 *Collections.* https://doi.org/10.6075/J0BR8QJ1

878 Kessler, W. S., Spillane, M. C., McPhaden, M. J., & Harrison, D. E. (1996). Scales of Variability in the
879 Equatorial Pacific Inferred from the Tropical Atmosphere-Ocan Buoy Array. *Journal of Climate*,
880 *9*, 2999–3024.

881 Kilham, P., & Hecky, R. E. (1988). Comparative ecology of marine and freshwater phytoplankton1:
882 Phytoplankton ecology. *Limnology and Oceanography*, *33*(4, part 2), 776–795.
883 https://doi.org/10.4319/lo.1988.33.4part2.0776

884 Klais, R., Tamminen, T., Kremp, A., Spilling, K., & Olli, K. (2011). Decadal-Scale Changes of
885 Dinoflagellates and Diatoms in the Anomalous Baltic Sea Spring Bloom. *PLoS ONE*, *6*(6),
886 e21567. https://doi.org/10.1371/journal.pone.0021567

887 Kuragano, T., & Kamachi, M. (2000). Global statistical space-time scales of oceanic variability estimated
888 from the TOPEX/POSEIDON altimeter data. *Journal of Geophysical Research: Oceans*,
889 *105*(C1), 955–974. https://doi.org/10.1029/1999JC900247

890 Kushnir, Y., Robinson, W. A., Bladé, I., Hall, N. M. J., Peng, S., & Sutton, R. (2002). Atmospheric GCM
891 Response to Extratropical SST Anomalies: Synthesis and Evaluation*. *Journal of Climate*,
892 *15*(16), 2233–2256. https://doi.org/10.1175/1520-0442(2002)015<2233:AGRTES>2.0.CO;2

893 Leblanc, K., Quéguiner, B., Diaz, F., Cornet, V., Michel-Rodriguez, M., Madron, X. D. de, et al. (2018).
894 Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon
895 export. *Nature Communications*, *9*(1), 1–12. https://doi.org/10.1038/s41467-018-03376-9

896 Legendre, L., & Le Fèvre, J. (1995). Microbial food webs and the export of biogenic carbon in oceans.
897 *Aquatic Microbial Ecology*, *9*, 69–77. https://doi.org/10.3354/ame009069

898 Lévy, M., Jahn, O., Dutkiewicz, S., & Follows, M. J. (2014). Phytoplankton diversity and community
899 structure affected by oceanic dispersal and mesoscale turbulence: Dispersal Impact on Plankton
900 Diversity. *Limnology and Oceanography: Fluids and Environments*, *4*(1), 67–84.
901 https://doi.org/10.1215/21573689-2768549

902 Lumpkin, R., & Johnson, G. C. (2013). Global ocean surface velocities from drifters: Mean, variance, El
903 Niño-Southern Oscillation response, and seasonal cycle: Global Ocean Surface Velocities.
904 *Journal of Geophysical Research: Oceans*, *118*(6), 2992–3006.
905 https://doi.org/10.1002/jgrc.20210

Lyman, J. M., Chelton, D. B., deSzoeke, R. A., & Samelson, R. M. (2005). Tropical Instability Waves as a Resonance between Equatorial Rossby Waves. *Journal of Physical Oceanography*, *35*(2), 232–254. https://doi.org/10.1175/JPO-2668.1

Mac Arthur, R. H., & Wilson, E. O. (1967). *The Theory of Island Biogeography* (Vol. M3). N.J.: Princeton University Press.

Marañón, E., Cermeño, P., López-Sandoval, D. C., Rodríguez-Ramos, T., Sobrino, C., Huete-Ortega, M., et al. (2013). Unimodal size scaling of phytoplankton growth and the size dependence of nutrient uptake and use. *Ecology Letters*, *16*(3), 371–379. https://doi.org/10.1111/ele.12052

Masumoto, Y., Hase, H., Kuroda, Y., Matsuura, H., & Takeuchi, K. (2005). Intraseasonal variability in the upper layer currents observed in the eastern equatorial Indian Ocean. *Geophysical Research Letters*, *32*(2). https://doi.org/10.1029/2004GL021896

Mazloff, M. R., Cornuelle, B. D., Gille, S. T., & Verdy, A. (2018). Correlation Lengths for Estimating the Large-Scale Carbon and Heat Content of the Southern Ocean. *Journal of Geophysical Research: Oceans*, *123*(2), 883–901. https://doi.org/10.1002/2017JC013408

McParland, E. L., & Levine, N. M. (2019). The role of differential DMSP production and community composition in predicting variability of global surface DMSP concentrations. *Limnology and Oceanography*, *64*(2), 757–773. https://doi.org/10.1002/lno.11076

Menemenlis, D., Campin, J.-M., Heimbach, P., Hill, C., Lee, T., Nguyen, A., et al. (2008). ECCO2: High resolution global ocean and sea ice data synthesis. *Mercator Ocean Quarterly Newslwtter*, (31), 13–21.

Ninove, F., Le Traon, P.-Y., Remy, E., & Guinehut, S. (2016). Spatial scales of temperature and salinity variability estimated from Argo observations. *Ocean Science*, *12*(1), 1–7. https://doi.org/10.5194/os-12-1-2016

Pickart, R. S., Macdonald, A. M., Moore, G. W. K., Renfrew, I. A., Walsh, J. E., & Kessler, W. S. (2009). Seasonal Evolution of Aleutian Low Pressure Systems: Implications for the North Pacific Subpolar Circulation*. *Journal of Physical Oceanography*, *39*(6), 1317–1339. https://doi.org/10.1175/2008JPO3891.1

Richardson, P. L. (1983). Eddy kinetic energy in the North Atlantic from surface drifters. *Journal of Geophysical Research*, *88*(C7), 4355. https://doi.org/10.1029/JC088iC07p04355

Richardson, T. L., & Jackson, G. A. (2007). Small Phytoplankton and Carbon Export from the Surface Ocean. *Science*, *315*(5813), 838–840. https://doi.org/10.1126/science.1133471

Smetacek, V. (1999). Diatoms and the ocean carbon cycle. *Protist*, *150*(1), 25–32.

Talley, L. D. (2011). *Descriptive Physical Oceanography: An Introduction*. Academic Press.

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., et al. (2018). Influence of diatom diversity on the ocean biological carbon pump. *Nature Geoscience*, *11*(1), 27–37. https://doi.org/10.1038/s41561-017-0028-x

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, *348*(6237), 1261605-1-1261605–11. https://doi.org/10.1126/science.1261605

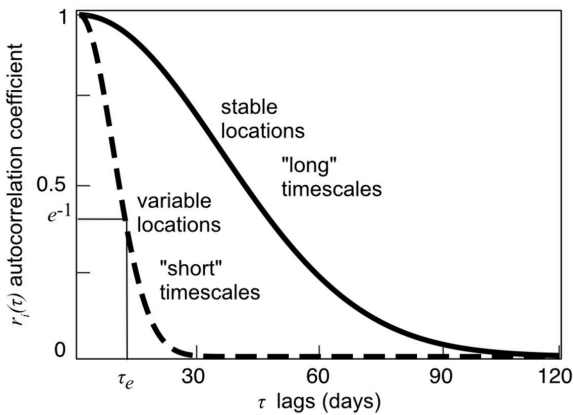
Villarino, E., Watson, J. R., Jönsson, B., Gasol, J. M., Salazar, G., Acinas, S. G., et al. (2018). Large-scale ocean connectivity and planktonic body size. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02535-8

Ward, B. A., Dutkiewicz, S., Jahn, O., & Follows, M. (2012). A size-structured food-web model for the global ocean. *Limnology and Oceanography*, *57*(6), 1877–1891.

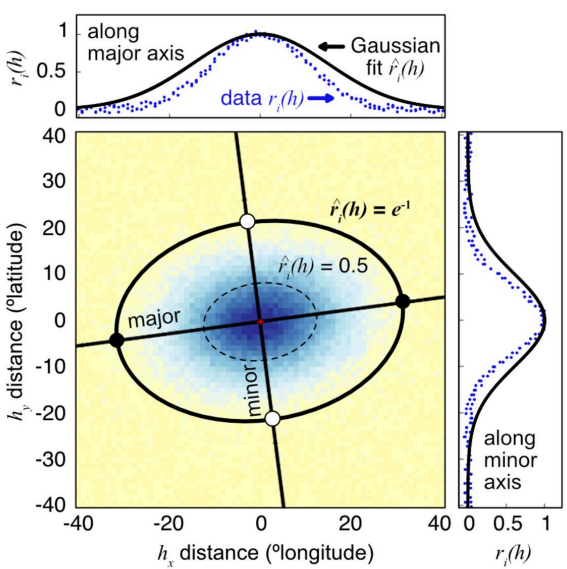
Yuan, X., & Talley, L. D. (1996). The subarctic frontal zone in the North Pacific: Characteristics of frontal structure from climatological data and synoptic surveys. *Journal of Geophysical Research: Oceans*, *101*(C7), 16491–16508. https://doi.org/10.1029/96JC01249

Figure 1.

a. Correlation timescales diagram



b. Correlation length scales diagram



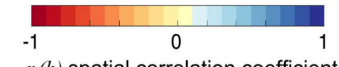


Figure 2.

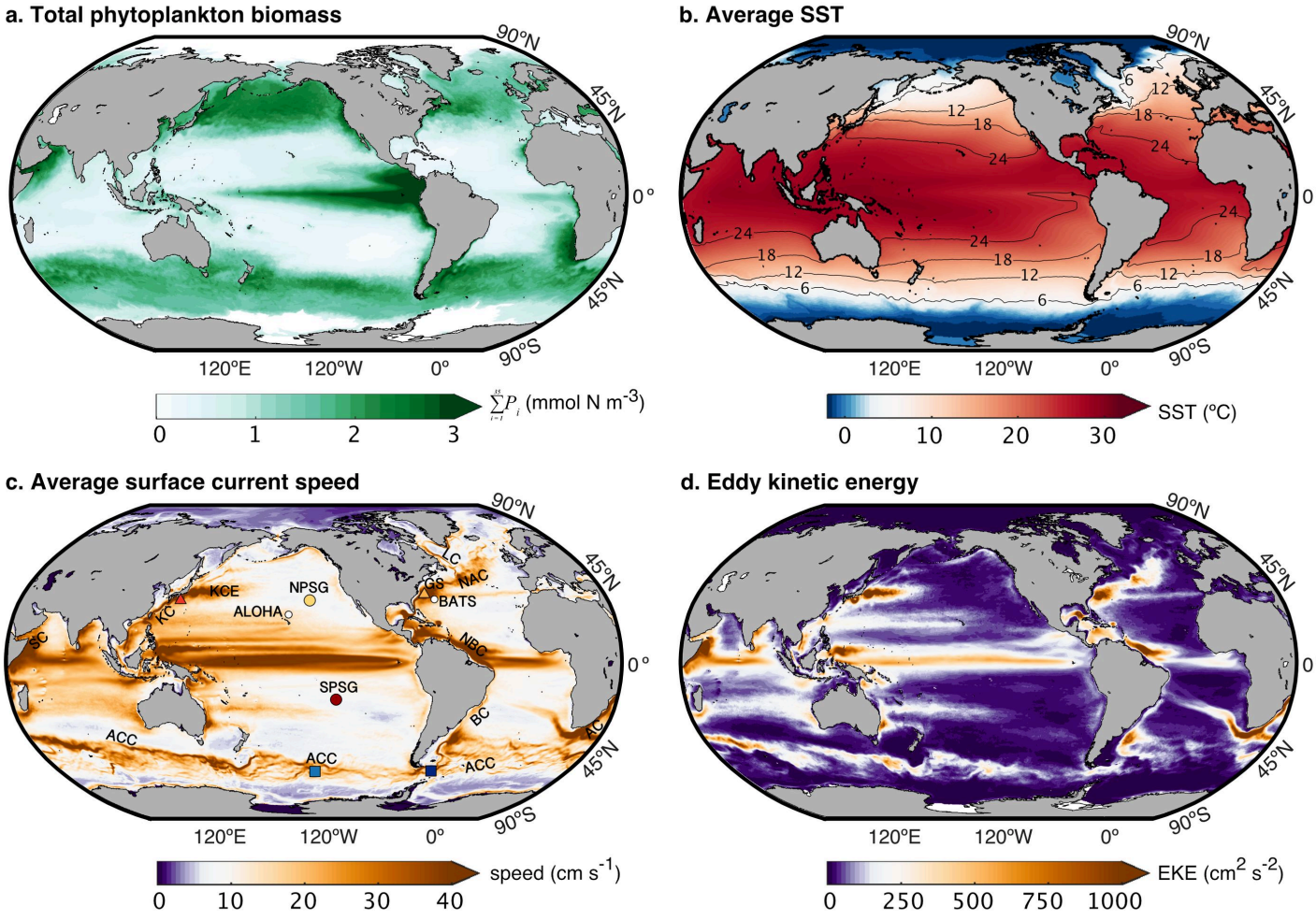

a. Total phytoplankton biomass
90°N
45°N
0°
45°S
90°S
120°E
120°W
0°
$\sum_{i=1}^{n} P_i$ (mmol N m^{-3})
0
1
2
3
b. Average SST
6
12
18
24
SST (°C)
0
10
20
30
c. Average surface current speed
KE
KCE
KC
NPSG
ALOHA
LC
GS
NAC
BATS
NBC
SC
SPSG
BC
ACC
speed (cm s^{-1})
0
10
20
30
40
d. Eddy kinetic energy
EKE (cm^2 s^{-2})
0
250
500
750
1000

Figure 3.

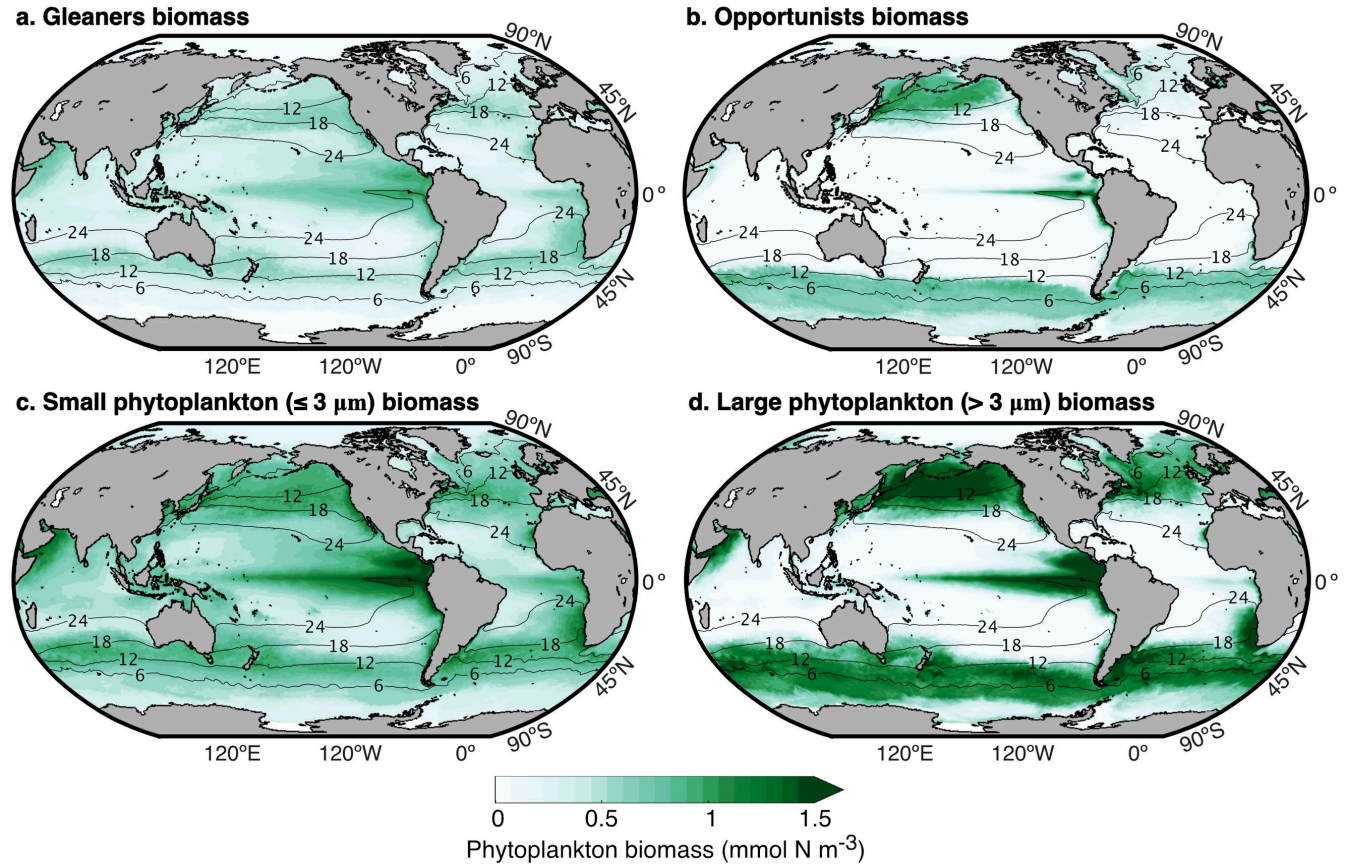

a. Gleaners biomass
b. Opportunists biomass
c. Small phytoplankton (≤ 3 μm) biomass
d. Large phytoplankton (> 3 μm) biomass
90°N
45°N
0°
45°N
90°S
120°E
120°W
0°
6
12
18
24
0
0.5
1
1.5
Phytoplankton biomass (mmol N m-3)

Figure 4.

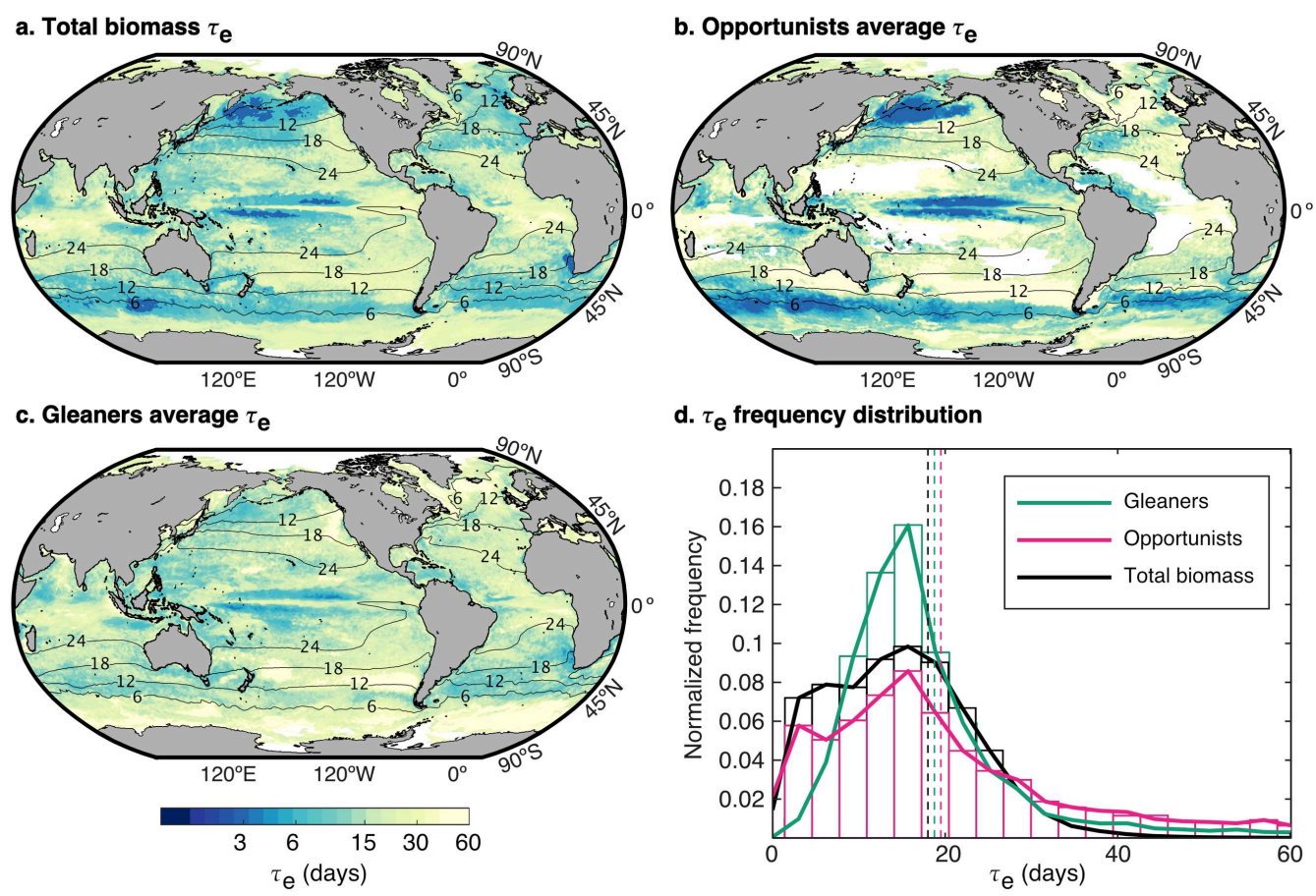

a. Total biomass τ_e
b. Opportunists average τ_e
c. Gleaners average τ_e
d. τ_e frequency distribution
90°N
45°N
0°
45°N
90°S
120°E
120°W
0°
3
6
15
30
60
τ_e (days)
Gleaners
Opportunists
Total biomass
Normalized frequency
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
20
40
60
τ_e (days)

Figure 5.

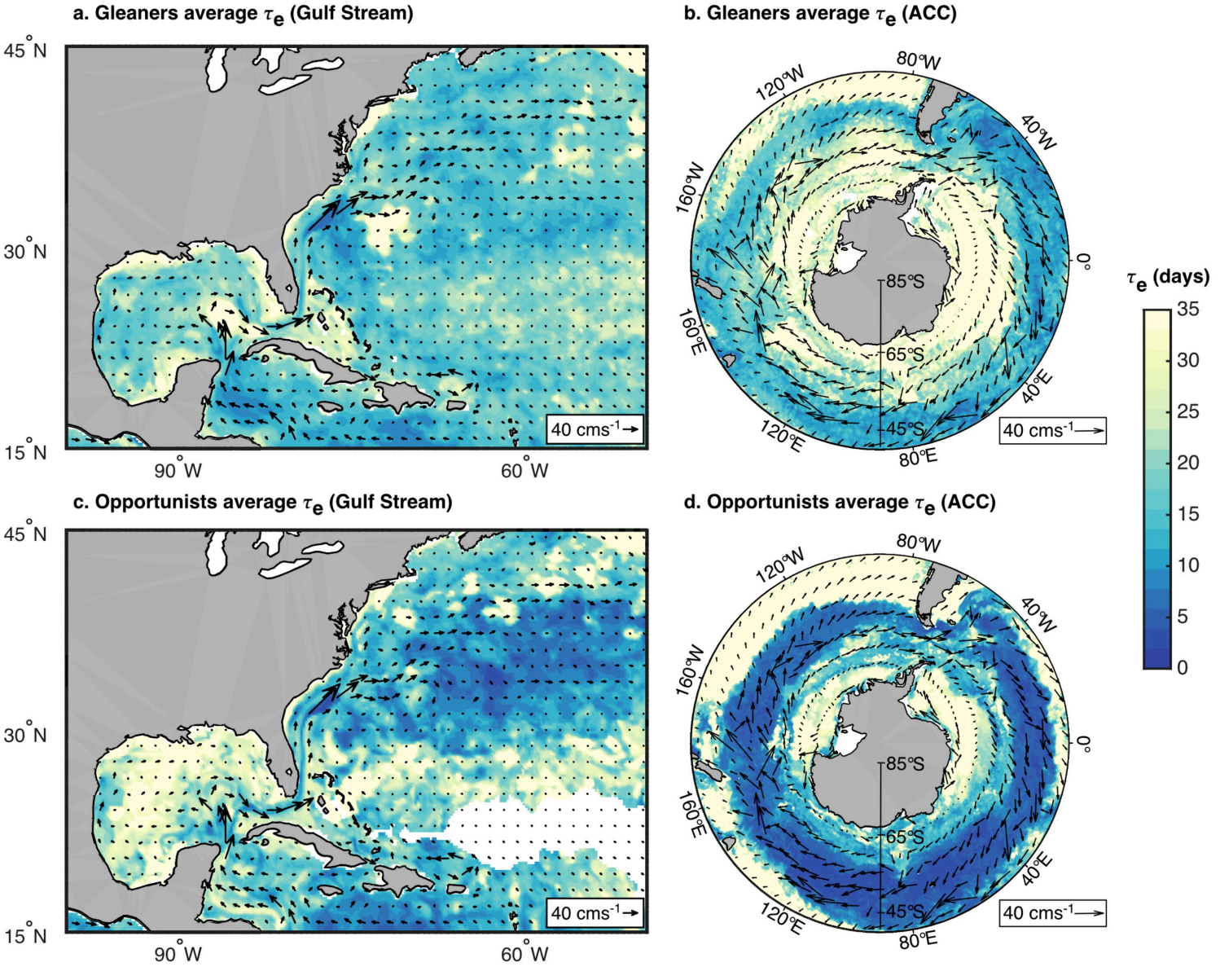

a. Gleaners average τ_e (Gulf Stream)
b. Gleaners average τ_e (ACC)
c. Opportunists average τ_e (Gulf Stream)
d. Opportunists average τ_e (ACC)
45°N
30°N
15°N
90°W
60°W
40 cms^{-1}
80°W
120°W
40°W
160°W
0°
160°E
40°E
120°E
80°E
85°S
65°S
45°S
τ_e (days)
35
30
25
20
15
10
5
0

Figure 6.

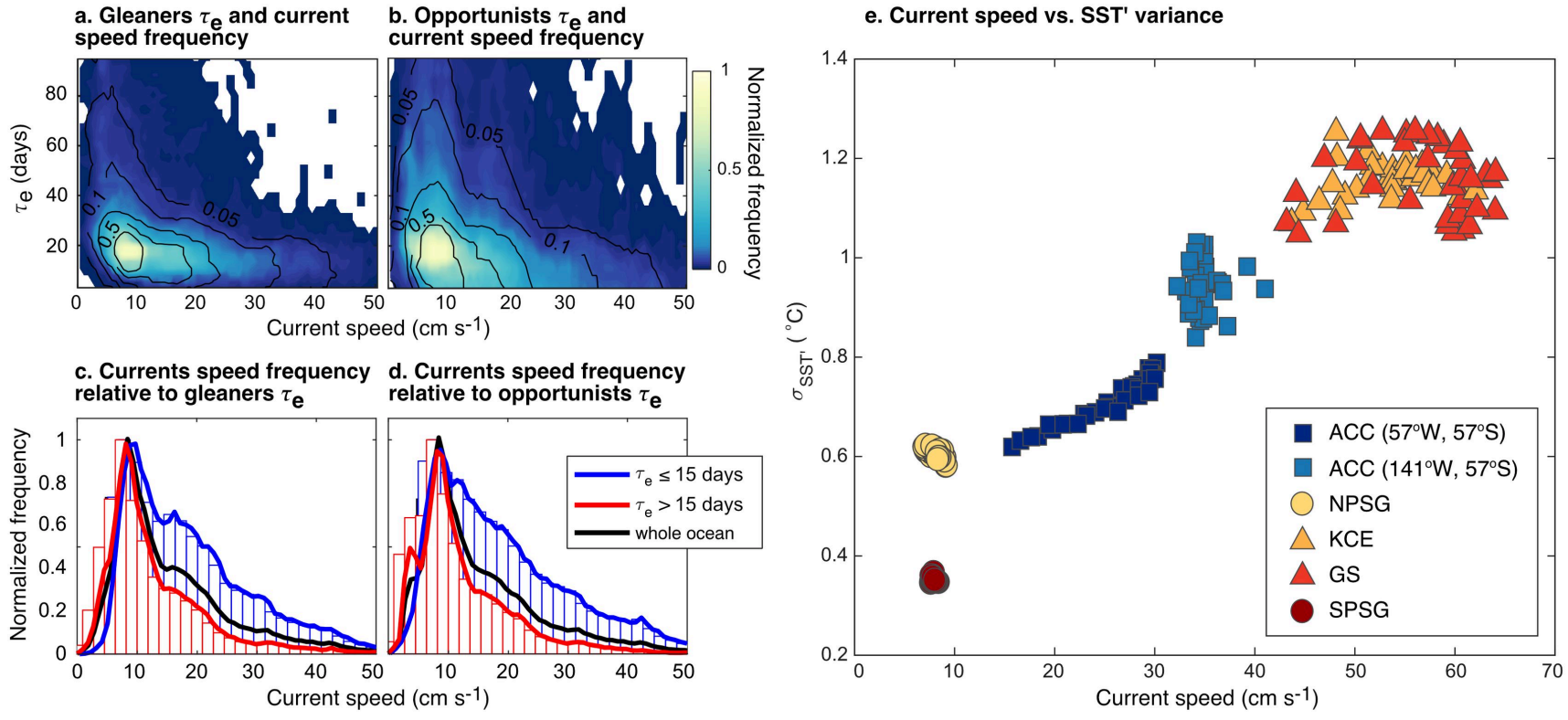
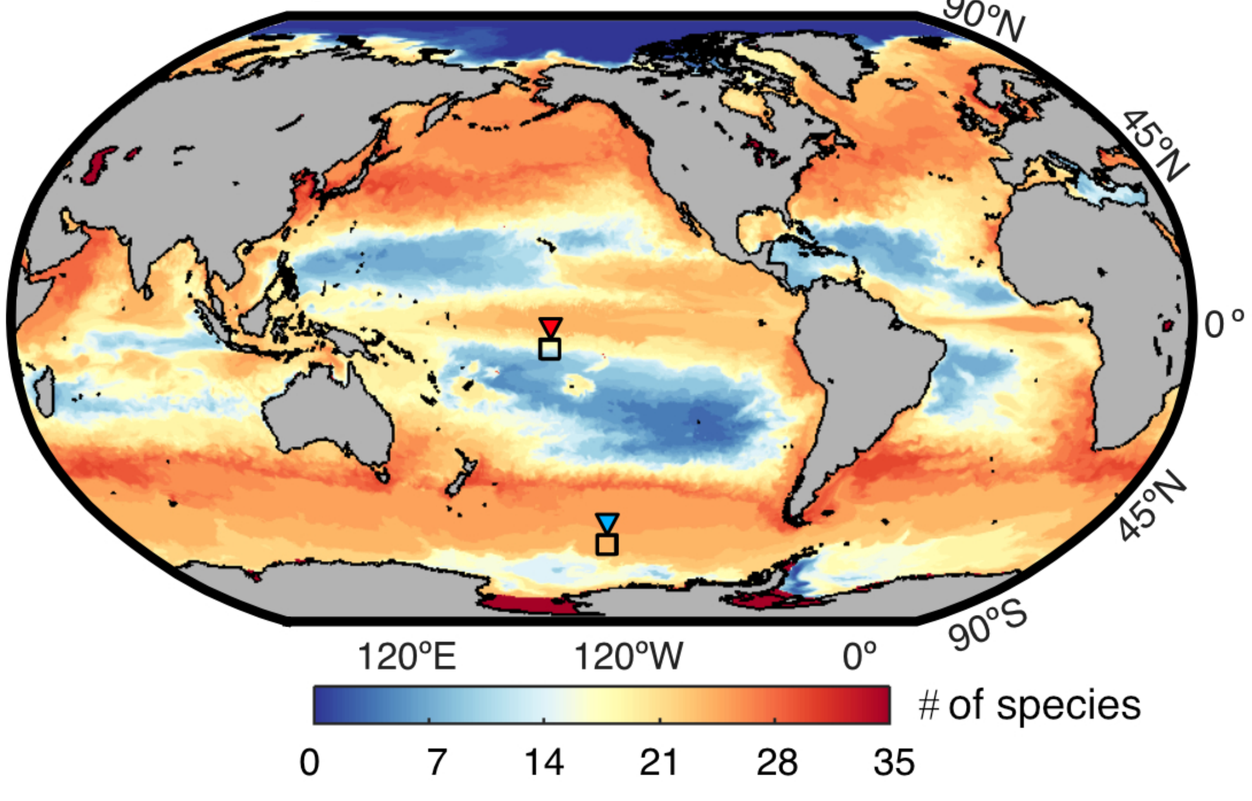
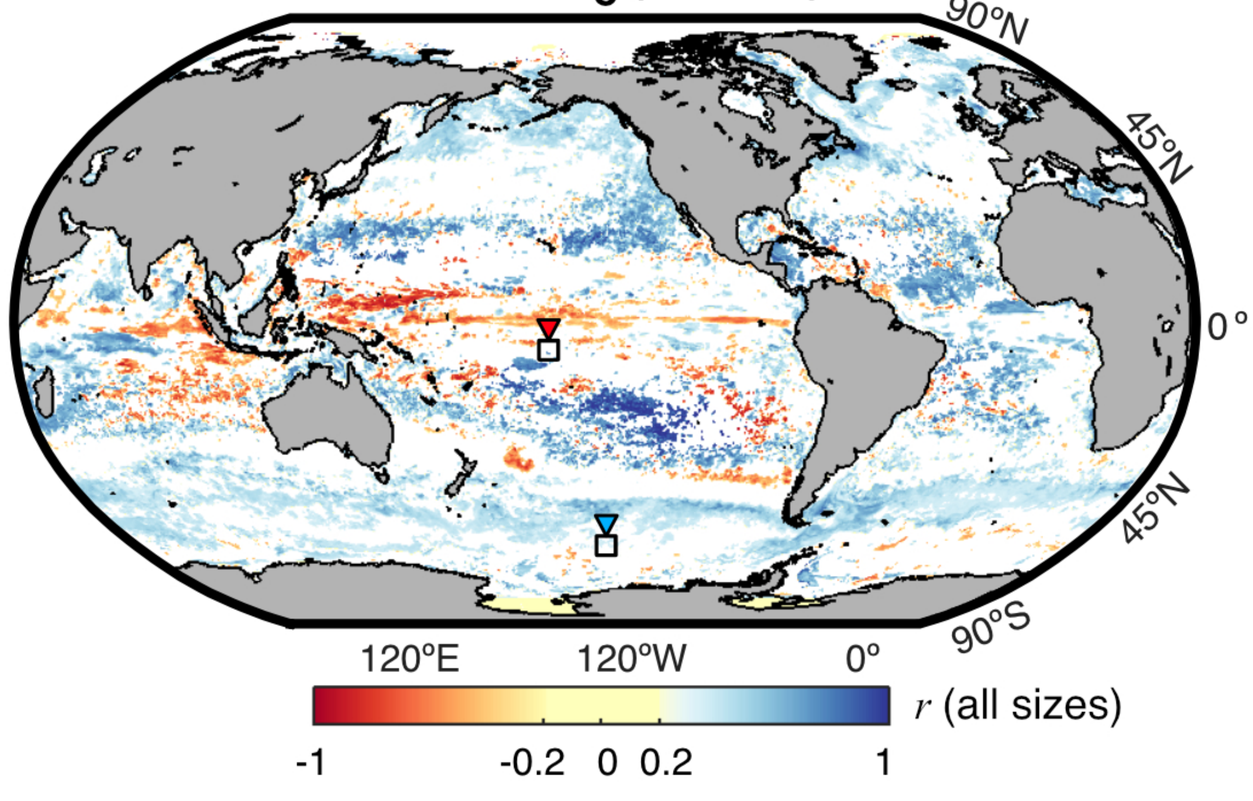

a. Gleaners τ_e and current speed frequency
b. Opportunists τ_e and current speed frequency
τ_e (days)
Current speed (cm s^{-1})
Normalized frequency
c. Currents speed frequency relative to gleaners τ_e
d. Currents speed frequency relative to opportunists τ_e
Normalized frequency
Current speed (cm s^{-1})
$\tau_e \leq 15$ days
$\tau_e > 15$ days
whole ocean
e. Current speed vs. SST' variance
$\sigma_{SST'}$ (°C)
Current speed (cm s^{-1})
ACC (57°W, 57°S)
ACC (141°W, 57°S)
NPSG
KCE
GS
SPSG

Figure 7.

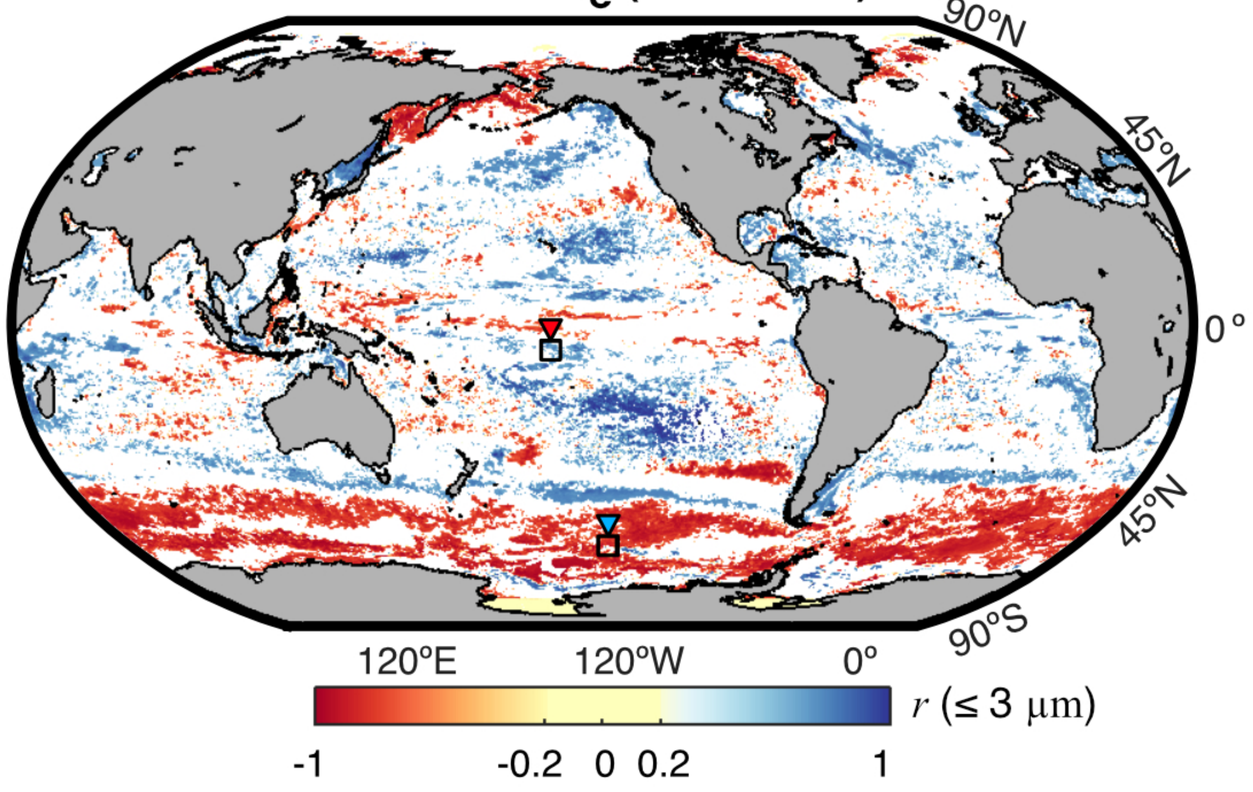
a. Number of species in correlation



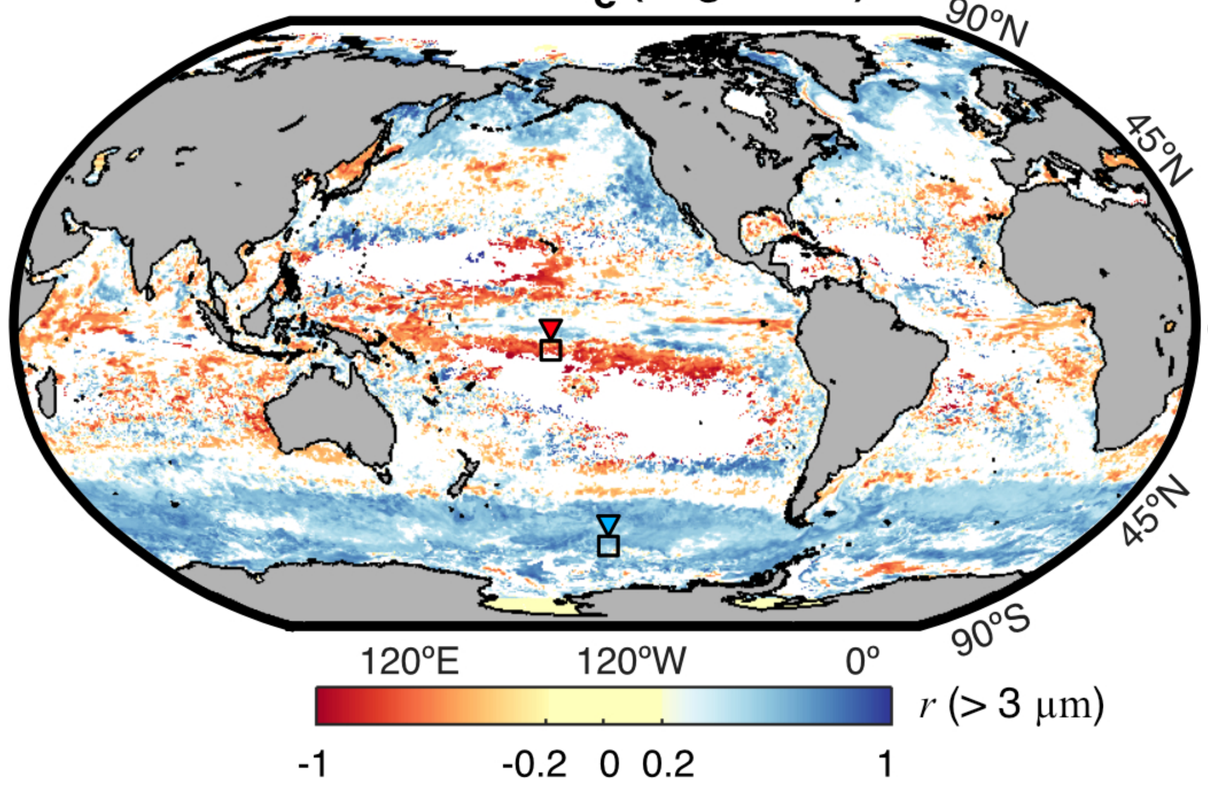
b. Correlation: cell size vs. τ_e (all sizes)



c. Correlation: cell size vs. τ_e (small cells)



d. Correlation: cell size vs. τ_e (large cells)



e. Non-linear relationships examples

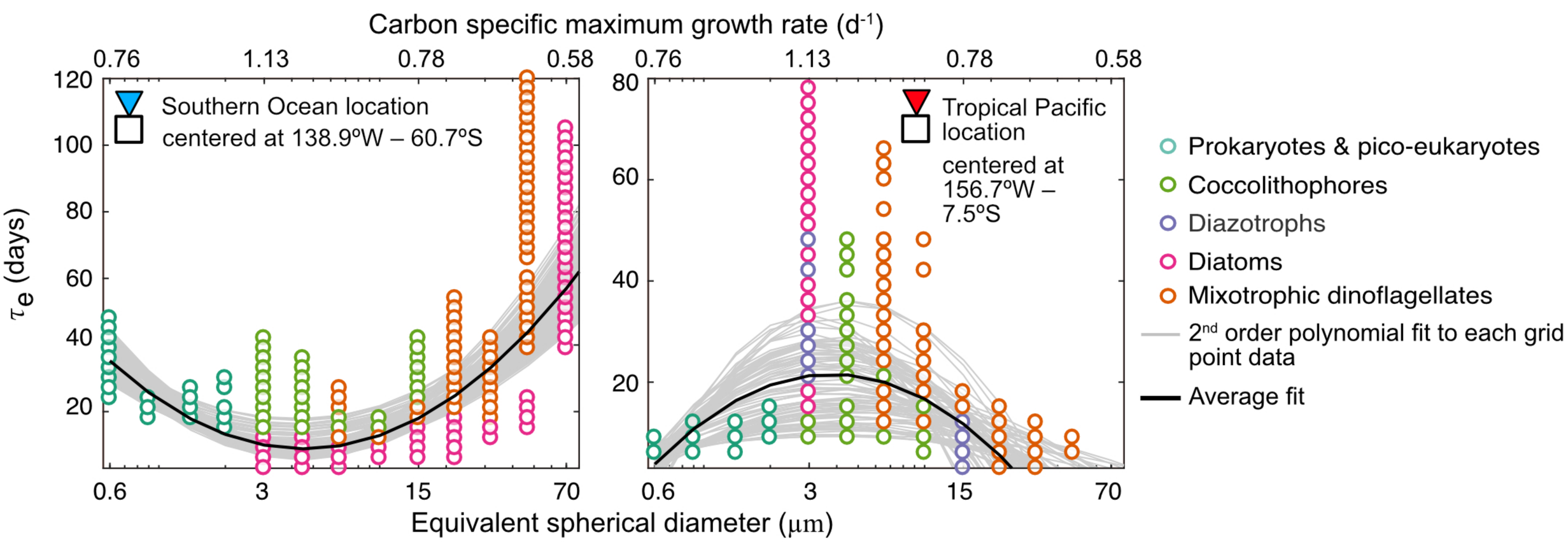


Figure 8.

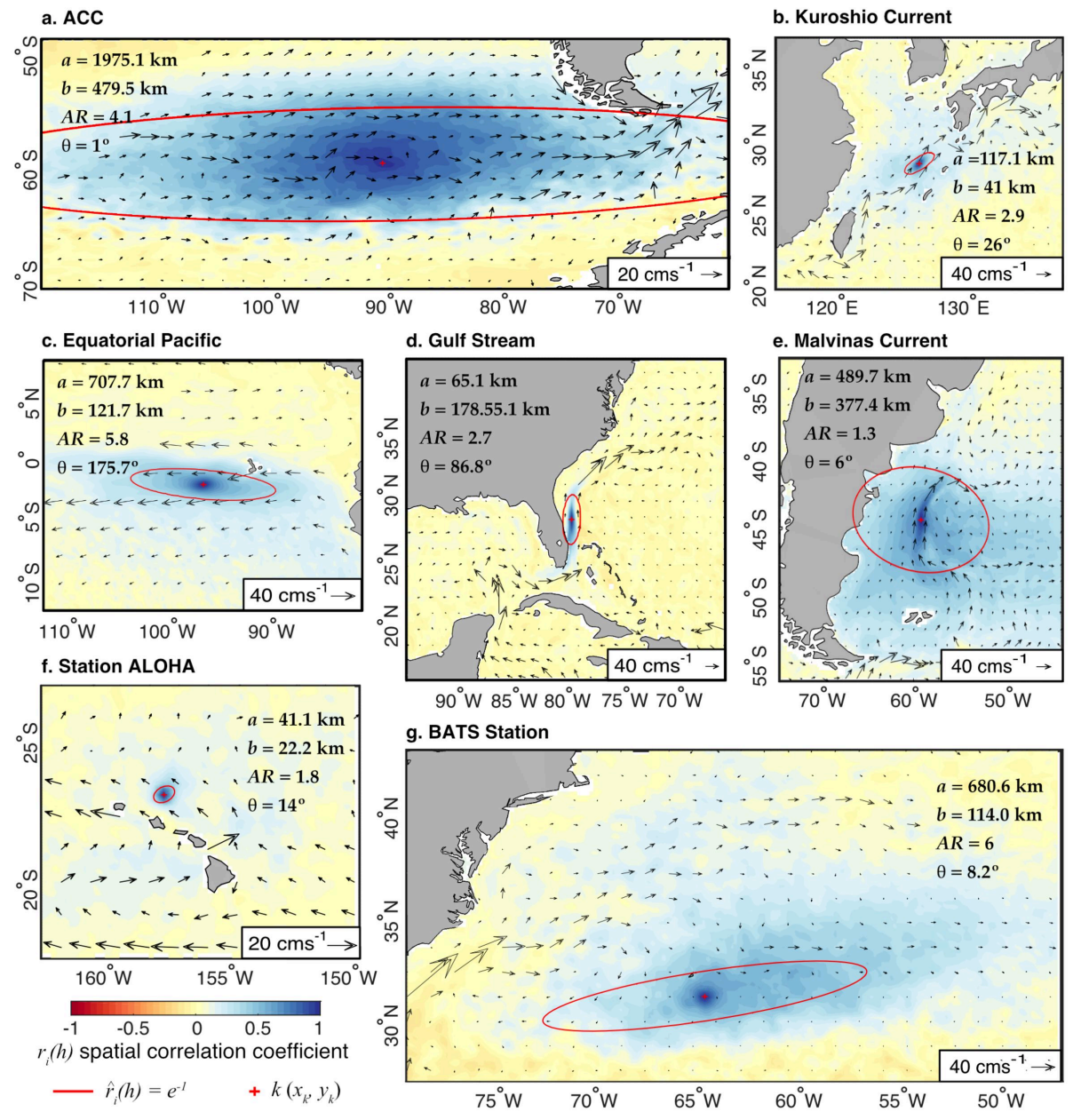

a. ACC
a = 1975.1 km
b = 479.5 km
AR = 4.1
θ = 1°
20 cms-1
50°S
60°S
70°S
110°W
100°W
90°W
80°W
70°W
b. Kuroshio Current
a =117.1 km
b = 41 km
AR = 2.9
θ = 26°
40 cms-1
35°N
30°N
25°N
20°N
120°E
130°E
c. Equatorial Pacific
a = 707.7 km
b = 121.7 km
AR = 5.8
θ = 175.7°
40 cms-1
5°N
0°
5°S
10°S
110°W
100°W
90°W
d. Gulf Stream
a = 65.1 km
b = 178.55.1 km
AR = 2.7
θ = 86.8°
40 cms-1
35°N
30°N
25°N
20°N
90°W
85°W
80°W
75°W
70°W
e. Malvinas Current
a = 489.7 km
b = 377.4 km
AR = 1.3
θ = 6°
40 cms-1
35°S
40°S
45°S
50°S
55°S
70°W
60°W
50°W
f. Station ALOHA
a = 41.1 km
b = 22.2 km
AR = 1.8
θ = 14°
20 cms-1
25°S
20°S
160°W
155°W
150°W
-1
-0.5
0
0.5
1
ri(h) spatial correlation coefficient
r̂i(h) = e-1
k (xk, yk)
g. BATS Station
a = 680.6 km
b = 114.0 km
AR = 6
θ = 8.2°
40 cms-1
40°N
35°N
30°N
75°W
70°W
65°W
60°W
55°W
50°W

Figure 9.

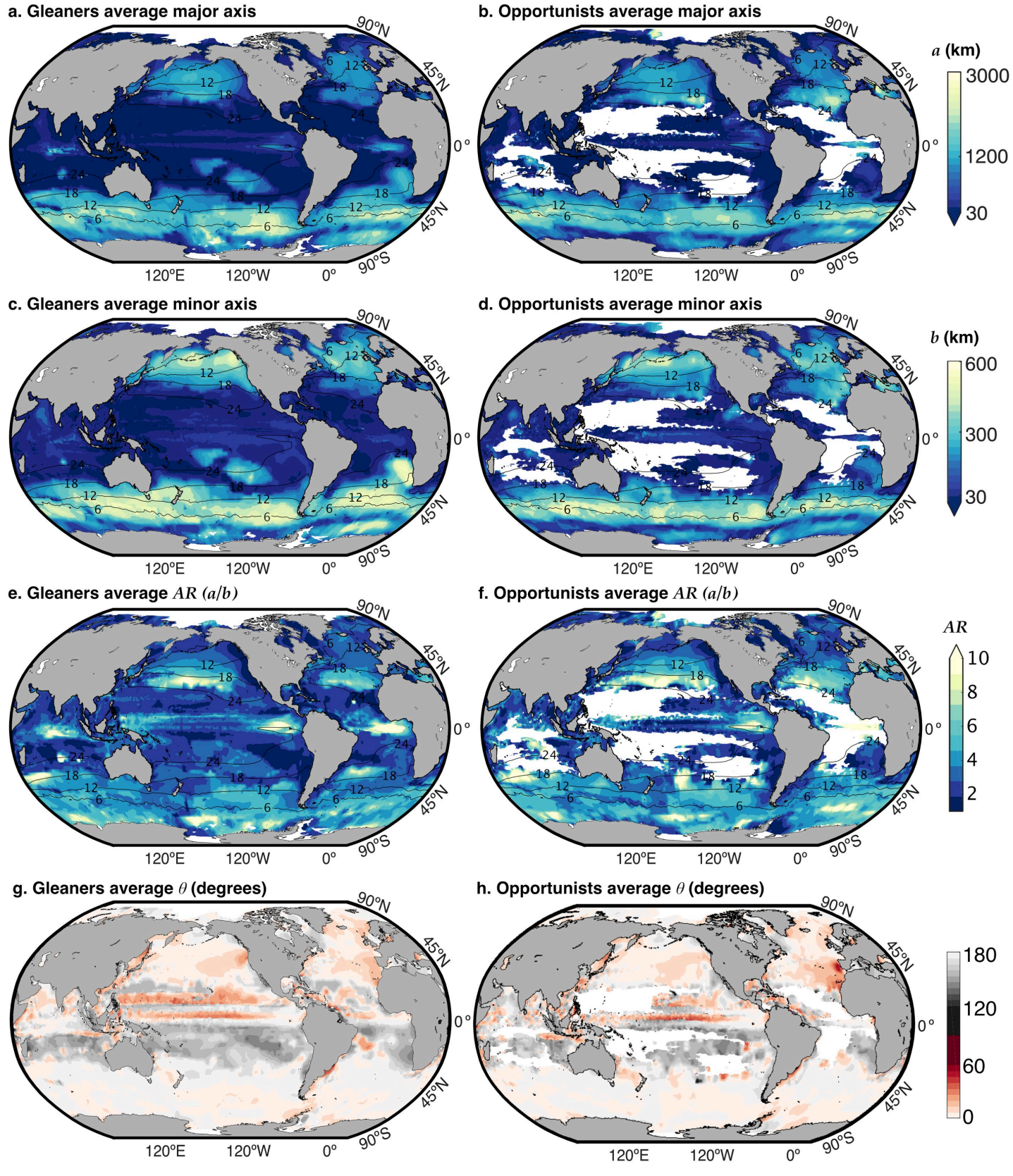

a. Gleaners average major axis
b. Opportunists average major axis
a (km)
3000
1200
30
c. Gleaners average minor axis
d. Opportunists average minor axis
b (km)
600
300
30
e. Gleaners average AR (a/b)
f. Opportunists average AR (a/b)
AR
10
8
6
4
2
g. Gleaners average θ (degrees)
h. Opportunists average θ (degrees)
180
120
60
0
90°N
45°N
0°
90°S
120°E
120°W
0°

Figure 10.

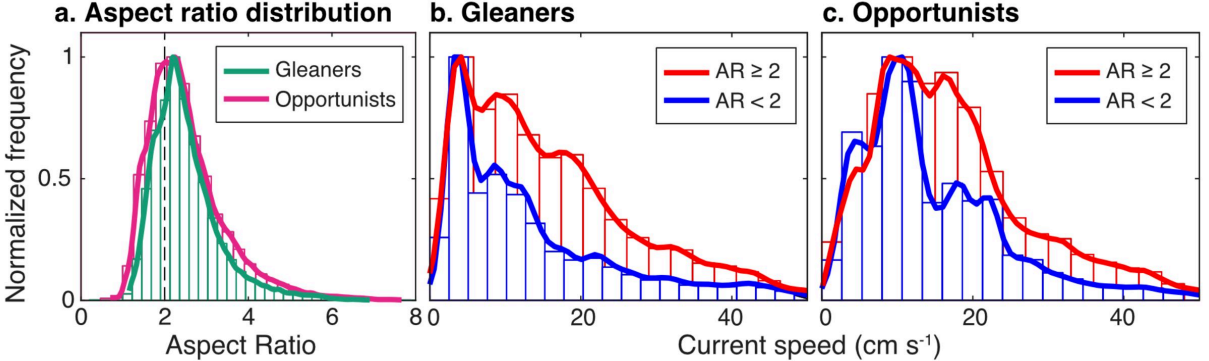
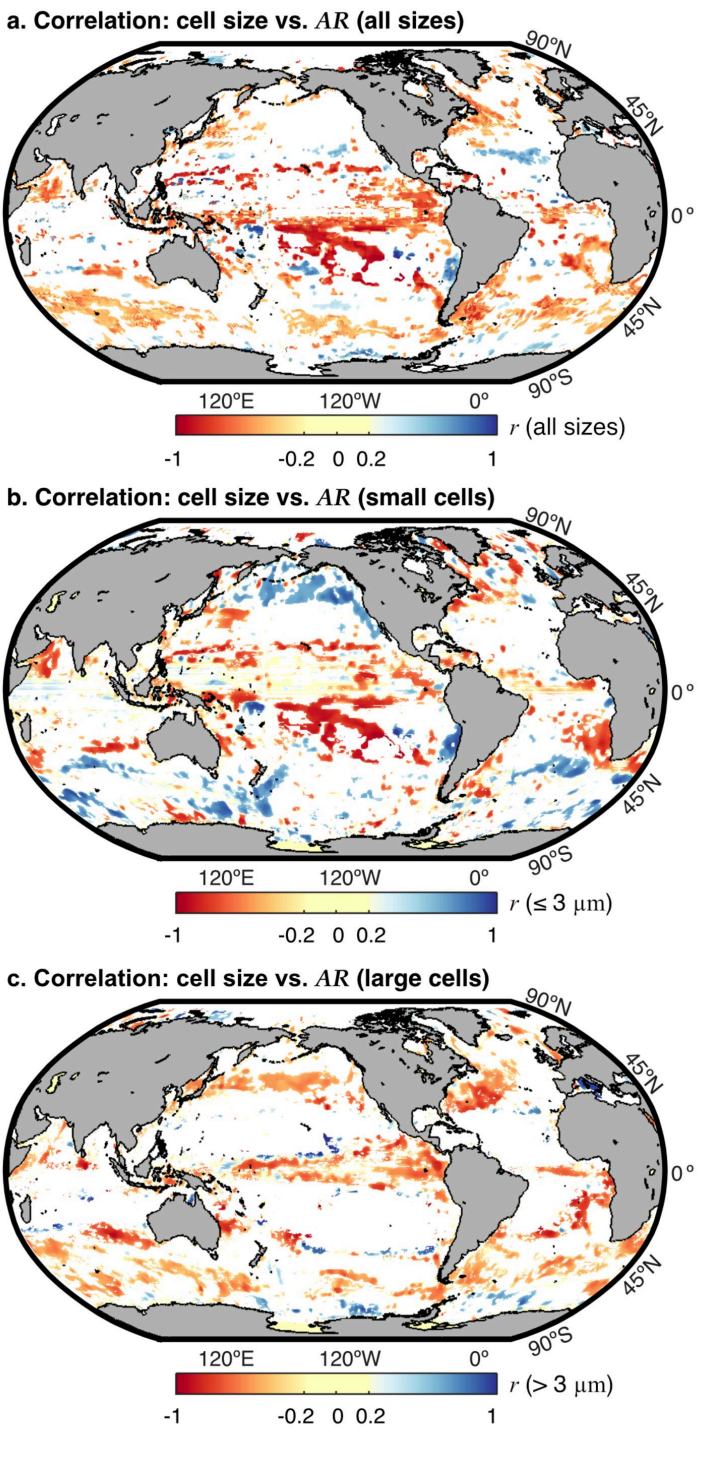

a. Aspect ratio distribution
b. Gleaners
c. Opportunists
Normalized frequency
1
0.5
0
Gleaners
Opportunists
AR ≥ 2
AR < 2
0
2
4
6
8
20
40
Aspect Ratio
Current speed (cm s-1)

Figure 11.


a. Correlation: cell size vs. AR (all sizes)
90°N
45°N
0°
45°N
90°S
120°E
120°W
0°
r (all sizes)
-1
-0.2 0 0.2
1
b. Correlation: cell size vs. AR (small cells)
90°N
45°N
0°
45°N
90°S
120°E
120°W
0°
r (≤ 3 μm)
-1
-0.2 0 0.2
1
c. Correlation: cell size vs. AR (large cells)
90°N
45°N
0°
45°N
90°S
120°E
120°W
0°
r (> 3 μm)
-1
-0.2 0 0.2
1