Inadequacies in the representation of sub-seasonal phytoplankton dynamics in Earth system models

Madhavan Girijakumari Keerthi[1,2], Olivier Aumont[1], Lester Kwiatkowski[1], Marina Levy[1]

[1]LOCEAN-IPSL, Sorbonne Université, CNRS, IRD, MNHN, Paris, France

[2]LMD-IPSL, École Normale Supérieure, Université PSL, CNRS, École Polytechnique, Paris, France

Correspondence to: Madhavan Girijakumari Keerthi (keerthi.madhavan-girijakumari@locean.ipsl.fr)

Abstract. Sub-seasonal phytoplankton dynamics on timescales between 8 days and 3 months significantly contribute to annual fluctuations, making it essential to accurately represent this variability in ocean models to avoid distorting long-term trends. This study assesses the capability of Earth System Models (ESMs) participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6) to reproduce sub-seasonal surface ocean phytoplankton variations observed in ocean color satellite data. Our findings reveal that, unlike sea surface temperature, all models struggle to accurately reproduce the total surface ocean phytoplankton variance and its decomposition across sub-seasonal, seasonal, and multi-annual timescales. Over the historical period, some models strongly overestimate sub-seasonal variance and exaggerate its role in annual fluctuations, while others underestimate it. Our analysis suggest that underestimation of sub-seasonal variance is likely a consequence of the coarse horizontal resolution of CMIP6 models, which is insufficient to resolve mesoscale processes—a limitation potentially alleviated with higher-resolution models. Conversely, we suggest that the overestimation of sub-seasonal variance is potentially the consequence of intrinsic oscillations such as predator-prey oscillations in certain biogeochemical models. ESMs consistently show a reduction in variance at sub-seasonal and seasonal timescales during the 21st century under high-emission scenarios. The poor capability of CMIP6 models at simulating sub-seasonal chlorophyll dynamics casts doubt on their projections at these temporal scales and multi-annual timescales. This study underscores the need to enhance spatial resolution and constrain intrinsic biogeochemical oscillations to improve projections of ocean phytoplankton dynamics.

1 Introduction

Phytoplankton, the photoautotrophic microscopic organisms populating the upper layers of the ocean, form the base of marine food webs and play a crucial role in driving ocean biogeochemical cycles. Over recent decades, climate change due to anthropogenic activities has emerged as a significant threat to ocean phytoplankton, altering the key environmental factors essential for their growth and survival (Behrenfeld et al., 2006; Bindoff et al., 2019). The repercussions extend beyond the

marine environment, impacting the global carbon cycle and the future absorption of atmospheric carbon dioxide by the ocean (Bopp et al., 2005; Gregg et al., 2005).

30   Earth system models (ESMs) are indispensable tools for forecasting the impacts of climate change on ocean primary productivity, and comprehending the intricate interplay between oceanic physical and biological processes. ESMs consistently project increased stratification across various climate change scenarios, enhancing phytoplankton nutrient limitation in low-latitude oceans (Steinacher et al., 2010; Bopp et al., 2013; Krumhardt et al., 2017; Kwiatkowski et al., 2017; Moore et al., 2018). As a consequence, marine primary production is globally projected to decrease (Sarmiento et al., 2004; Cabré et al.,

35   2014; Fu et al., 2016; Kwiatkowski et al., 2020). However, the extent of this decline remains highly uncertain across model ensembles, including uncertainty in even the direction of change (Bopp et al., 2013; Krumhardt et al., 2017; Kwiatkowski et al., 2020).

A comprehensive comparison of the ocean biogeochemistry simulated by ESMs with observations can shed light on model
40   deficiencies and associated driving factors (Séférian et al., 2020; Kwiatkowski et al., 2018; Kessler & Tjiputra, 2016; Planchat et al., 2023). The availability of two decades of daily satellite ocean color measurements of surface chlorophyll (SChl, a proxy for phytoplankton biomass) at global scale represents a unique means to evaluate the skill of ESMs to simulate phytoplankton. However, assessing multi-model uncertainty in climate projections has to go beyond evaluating solely the model mean state performance. It is crucial to assess models against observed variations across all timescales to bolster confidence in their
45   projections (Séférian et al., 2020). This is particularly critical for phytoplankton as it is characterized by large natural variability at diverse timescales, which often masks the long-term trends (Henson et al., 2010, 2016; Doney et al., 2014; Keerthi et al., 2022).

The seasonal cycle represents the primary mode of SChl variability (Demarcq et al., 2012). However, in many oceanic regions,
50   sub-seasonal variability is equally significant and occasionally surpasses seasonal fluctuations (Keerthi et al., 2022; Prend et al., 2022; Levy et al., 2024). Sub-seasonal variability comprises high-frequency fluctuations associated with sub-seasonal atmospheric variability including storms and tropical cyclones (Carranza et al., 2015), sub-seasonal climate modes (Resplandy et al., 2009), mesoscale and submesoscale eddies (Gaube et al., 2014), and intrinsic biological processes (Mayersohn et al., 2021). In various locations, phytoplankton variations at sub-seasonal frequencies can be more than two times as large as the
55   climatological mean (Resplandy et al., 2009; Thomalla et al., 2011; Keerthi et al., 2021). In contrast, low-frequency (multi-annual) variations with distinct regional characteristics are evident, and correlated with large-scale climate modes (Wilson and Adamec, 2001; Racault et al., 2017; Park et al., 2018; Resplandy et al., 2009; Lovenduski and Gruber, 2005; Martinez et al., 2016). But with the exception of specific tropical regions, their contribution to total variability remains relatively modest (Keerthi et al., 2022).

60

Previous studies on simulated ocean primary production have predominantly focused on evaluating the mean state performance of models (Séférian et al., 2020; Bopp et al., 2013; Kwiatkowski et al., 2020), neglecting a comprehensive exploration of different temporal scales in model assessments. Capitalizing on high frequency global measurements of satellite ocean color SChl, we evaluated the performance of historical simulations produced by a subset of ESMs participating in the Coupled Model Intercomparison Project Phase 6 (CMIP6) to simulate global surface ocean phytoplankton dynamics across diverse temporal scales (sub-seasonal, seasonal, and multi-annual), with a specific focus on high frequency sub-seasonal variability. To do so, we applied the temporal decomposition methodology developed for SChl satellite data in Keerthi et al. (2022) to CMIP6 historical simulations. Our analysis of SChl is additionally contrasted with that of sea surface temperature (SST), a typically well-simulated physical ocean parameter, particularly in comparison to SChl.

Satellite ocean color measurements of SChl reveal that the cumulative effect of high-frequency sub-seasonal fluctuations can modulate year-to-year variations of SChl, a factor that has historically been overlooked (Keerthi et al., 2022; Prend et al., 2022). The changing frequency of extreme atmospheric events, such as marine heatwaves (Frolicher et al., 2018) and tropical cyclones (Knutson et al., 2020; Walsh et al., 2016), coupled with mesoscale and submesoscale variability linked to global warming scenarios (Martínez-Moreno et al., 2021), may actively contribute to alter the sub-seasonal variability of SChl. The intricate interplay between the different timescales has therefore the potential to shape overarching long-term trends in surface ocean phytoplankton and thus deserves a specific focus. We therefore extend our analysis to future model projections using simulations of the high-emission scenario SSP5-8.5.

The enhancement of resolution in coupled climate models improves atmospheric and oceanic dynamics, thereby reducing biases in the mean state and variability of various quantities (Müller et al., 2018). In our analysis of CMIP6 simulations, we also used the opportunity to compare the performance of a higher-resolution model version (MPI-ESM1.2-HR) and its lower-resolution counterpart (MPI-ESM1.2-LR) in simulating SChl variability across different timescales. MPI-ESM1.2-HR has a horizontal resolution twice as high for the atmospheric component (100 km) and more than twice as high for the oceanic component (~40 km) compared to MPI-ESM1.2-LR (200 km and 150 km for the atmospheric and oceanic components, respectively).

2 Data and Methods

Observation data: We utilised the datasets outlined in Keerthi et al. (2022) for observed SChl and SST. The SChl data is the Level 3 Mapped 9x9 km resolution 8-day averaged product (release 4.1), covering the period from January 1998 to December 2014. This dataset was obtained from the European Space Agency Ocean Color Climate Change Initiative (ESA OC-CCI; Sathyendranath and Krasemann, 2014) and can be accessed at http://www.oceancolour.org/. The product is a merged compilation from various ocean color satellite missions, including the Moderate Resolution Imaging Spectroradiometer

(MODIS)-Aqua, the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) and the Medium Resolution Imaging Spectrometer (MERIS). Given the limited coverage of the satellite-derived SChl data in polar regions, our analysis is concentrated on the region between 60°S and 60°N.

95

For SST, we used the daily 25x25 km resolution Optimum Interpolation Sea Surface Temperature (OISST) data, spanning from January 1998 to December 2014. This dataset is accessible through the National Oceanic and Atmospheric Administration (NOAA) at https://www.ncdc.noaa.gov/oisst/optimum-interpolation-sea-surface-temperature-oisst-v20). The OISST data integrates observations from satellites, ships, buoys, and Argo floats.

100

CMIP6 Simulations	Physical Ocean Model	Ocean BGC Model	Horizontal resolution (Physical & BGC Model)	Model Simulations	References
IPSL-CM6A-LR	NEMO-OPA	PISCES	100 km	Historical	Boucher et al. 2018, 2021; Séférian, 2018
IPSL-CM6A-LR-INCA					
CNRM-ESM2-1					
CESM2	POP2	MARBL	100 km	Historical	Danabasoglu, 2019a, b, c;
CESM2-FV2					
CESM2-WACCM-FV2					
MPI-ESM1.2-HAM	MPIOM	HAMOCC6	150 km	Historical	Neubauer et al., 2019 ; Wieners et al., 2019a, b, c; Jungclaus et al., 2019a, b; Schupfner et al., 2019
MPI-ESM1.2-LR				Historical, SSP5-8.5, piControl	
MPI-ESM1.2-HR			40 km		
NorESM2-LM	MICOM	HAMOCC	100 km	Historical, SSP5-8.5 piControl	Seland et al., 2019a, b, c ; Bentsen et al., 2019a, b
NorESM2-MM				Historical, SSP5-8.5	

Table 1. The CMIP6 Earth system models used in this study, their physical and biogeochemical ocean components, nominal horizontal ocean resolution and the simulations assessed.

105

CMIP6 Historical Simulations: We obtained SChl and SST data for the period 1981-2014 from https://esgf-node.ipsl.upmc.fr/search/cmip6-ipsl/. Our analysis focused on the 11 CMIP6 historical simulations that had daily SChl outputs. The comparison between satellite observations and CMIP6 historical simulations is provided for the common period, 1998–2014. The horizontal nominal resolution for ocean dynamics in most models is 100 km, except for MPI-ESM1-2-LR, MPI-ESM1.2-HAM and MPI-ESM1-2-HR, which have a resolution of 150, 150 and 40 km, respectively. For SST analysis, NorESM2-LM and NorESM2-MM are excluded due to the absence of daily outputs. The ensemble member "r1i1p1f1" is utilized for all models, except for CNRM-ESM2-1, where "r1i1p1f2" is used. Details relating to the eleven models utilized in our study are provided in Table 1.

CMIP6 Future projections: To project future variability in SChl at various timescales, we utilized a subset of ESMs (MPI-ESM1.2-LR, MPI-ESM1.2-HR, NorESM2-LM and NorESM2-MM) that performed the SSP5-8.5 scenario, providing daily resolution data for the period 2084-2100. Pre-industrial control simulations (piControl) were used to ensure that observed climate change signals were not influenced by model drift. Analysis of piControl simulations is presented for MPI-ESM1.2-LR, MPI-ESM1.2-HR and NorESM2-LM only, as NorESM2-MM does not provide daily SChl outputs in piControl simulations.

All analyses were performed on satellite observations and CMIP6 simulations regridded on a common 1°x1° spatial grid and a temporal resolution of 8 days. Satellite observations were regridded using area-weighted averaging. CMIP6 simulations were transformed using the CDO remapping tool remapdis (Schulzweida, 2023).

Temporal decomposition and variance explained: We applied a decomposition methodology akin to that in Keerthi et al., (2020, 2022) and Vantrepotte & Mélin (2009, 2011), to decompose the SChl and SST timeseries at each grid point to seasonal (S_t), multi-annual (MA_t) and sub-seasonal (SS_t) components. A comprehensive description of this methodology is available in Keerthi et al., (2020, 2022). This decomposition ensures that at every geographical location, the total time series (X_t) can be expressed as the sum of its sub-components: $X_t = SS_t + S_t + MA_t$. The seasonal component (S_t) encapsulates variability within a period of 3 months to 1 year as well as year-to-year variations in the seasonal cycle. The multi-annual component (MA_t) represents variability with a timescale longer than 8 months, while the sub-seasonal component (SS_t) captures variability with a period shorter than 88 days along with any irregular variability outside of that specified range. This method allows for small overlaps in the frequency ranges associated with each component.

The total variance of the SChl and SST timeseries can be decomposed into the cumulative variance explained by its different components along with the covariance amongst these components. In practice, the covariance terms are generally negligible. The proportional contribution of each component to the total variance is expressed as a percentage.

Spatial scale of coherence: The spatial scale of coherence associated with each time component (seasonal, multi-annual, and sub-seasonal) is defined as the extent over which the temporal signal remains self-coherent. We conducted cross-correlation analyses by comparing the decomposed time series of all grid cells included in a disk with a diameter of 2400 km. This sets then an upper limit of 2400 km to the scale we can infer with this method. We then counted the number of grid cells where the cross-correlation exceeded 0.8 and converted this count into a distance measurement. The chosen threshold value of 0.8 aligns with that in Keerthi et al., (2020, 2022).

Spatial decomposition: To assess the relative contribution of spatial scales at intervals of 100 km, 200 km, and so forth, to the sub-seasonal signal, we executed a spatial decomposition at every 8-day time step. This decomposition methodology is based on an iterative application of the heat diffusion equation, as presented in Weaver and Courtier (1990), that has been previously implemented in the work of Keerthi et al., (2013, 2016).

3 Results and Discussions

3.1 Evaluation of the mean state

Before turning to the analysis of temporal variability simulated by the models, we initially compare the mean state of the models with satellite-derived SChl. It is important to note that, due to our specific model selection process, the ensemble mean state presented here may differ from standard CMIP6 analyses, which typically include a wider selection of models. Our study focuses solely on simulations providing daily SChl outputs, discarding models that do not meet this criterion.

The ensemble mean of SChl from 11 CMIP6 simulations, as detailed in Table 1, is compared with satellite-based estimates derived from ESA OC-CCI Ocean color data (Figure 1a). Key features include elevated SChl levels in temperate, subpolar, and upwelling regions, contrasting with notably lower levels in the subtropical gyres. The latter areas are characterized by consistently low-nutrient conditions, while the former receive intermittent nutrient influxes through upwelling or deep mixing. Although the CMIP6 ensemble mean generally aligns with observations, there is a notable overestimation across the entire ocean (Figure 1a, b).

Séférian et al., (2020) undertook a comparison between the mean state of CMIP6 simulations and satellite SChl measurements (ESA OC-CCI) also spanning 1998-2014. Their results indicate significant discrepancies between models and observational data in reproducing the SChl mean state. The models assessed in both studies are CESM2, CNRM-ESM2-1, IPSL-CM6A-LR, MPI-ESM1.2-LR, and NorESM2-LR. Their findings suggest MPI-ESM1.2-LR persistently and globally overestimates SChl. NorESM2-LR slightly overestimates SChl in the tropics and subtropics but underestimates it in polar regions. CESM2, CNRM-ESM2-1, and IPSL-CM6A-LR display varying biases relative to satellite SChl across regions.

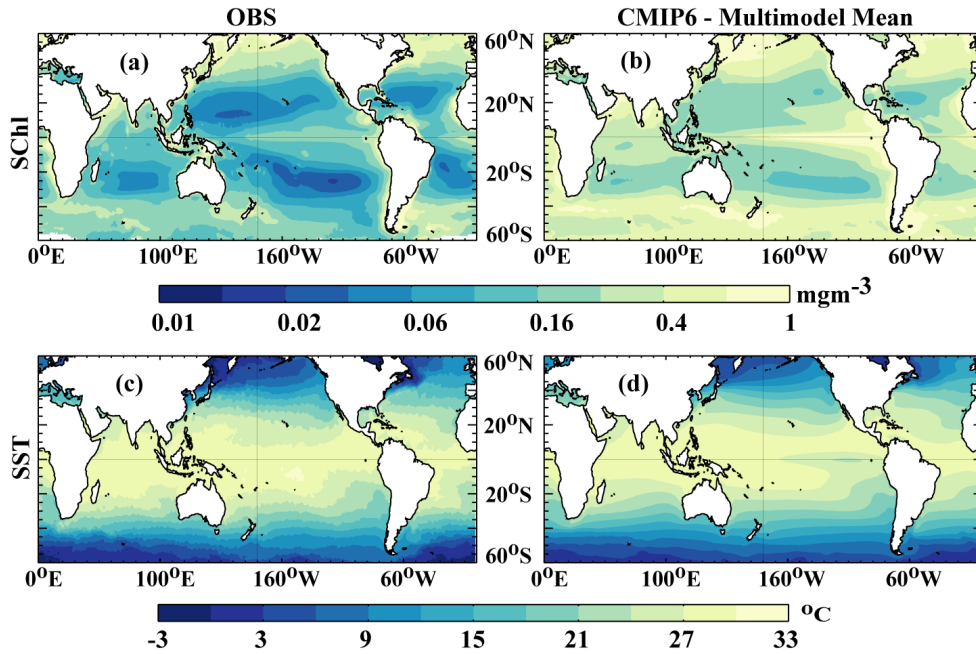



Figure 1: Mean state evaluation: Annual mean SChl (a) Observed (ESA OC-CCI product) and (b) CMIP6 multi-model mean for the years 1998-2014 and domain 60°N-60°S. (c & d) Similarly for SST.

175

The ability of the different model configurations to represent spatial variability is quantified in the Taylor diagram (Fig. 2a). This analysis reveals important differences between models. Most models analyzed here systematically underestimate the observed SChl spatial variance. All models, with the exception of the MPI models, exhibit weak spatial variability ranging from 0.15 to 0.3 mg Chl/m³, compared to 0.60 mg Chl/m³ in satellite observations. MPI models show a similar spatial

180 variability to the satellite observations, ranging from 0.6 to 0.7 mg Chl/m³. The spatial correlation (Pearson correlation) between CMIP6 models and observations remains below 0.6 (Fig. 2a), with MPI models showing particularly low correlations, below 0.2.

In agreement with Séférian et al., (2020), models sharing a common physical ocean model generally have similar skill, though

185 exceptions are noted for CNRM-ESM2-1 and MPI-ESM1.2-HR. CNRM-ESM2-1, which, like IPSL-CM6A-LR and IPSL-CM6A-LR-INCA, includes the coupled physical biogeochemical model NEMO-PISCES, shows a slightly higher spatial standard deviation than the IPSL models. MPI-ESM1.2-HR, which shares the same physical and biogeochemical model MPIOM-HAMOCC as MPI-ESM1.2-HAM and MPI-ESM1.2-LR, exhibits a higher spatial standard deviation. CESM and NorESM2 configurations, which respectively use the coupled physical biogeochemical models POP2-MARBL and MICOM-

190 HAMOCC, simulate similar spatial correlations. Despite MPI and NorESM2 models using the same ocean biogeochemical

model HAMOCC, there are notable differences in their simulated spatial standard deviations. However, the spatial correlation with observations remains consistent among these models.

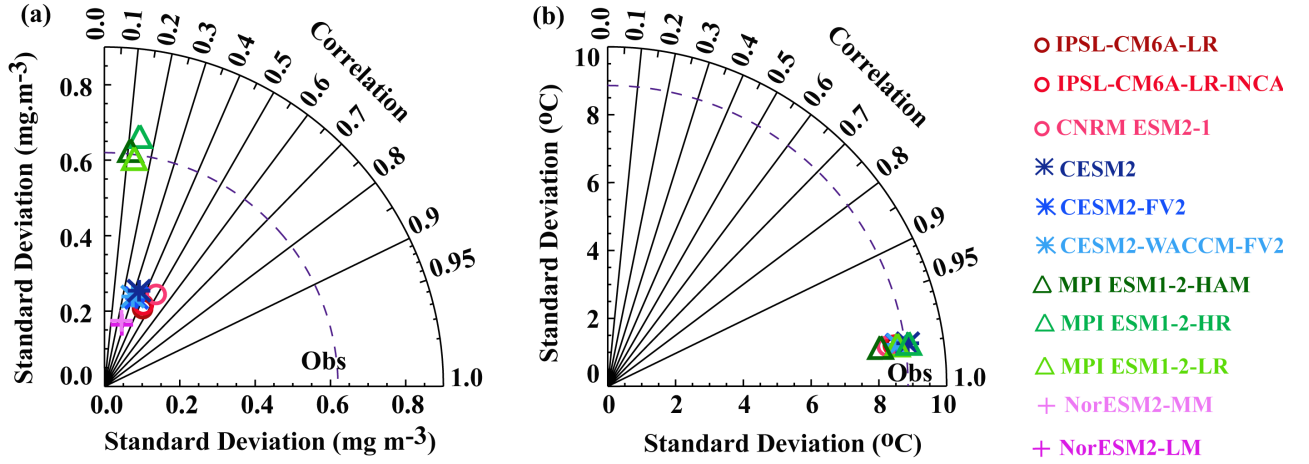



Figure 2: Evaluation of the mean spatial distribution: Taylor diagram for the annual mean (a) SChl and (b) SST over the years 1998-2014 and domain 60ºN-60ºS. The dashed curve represents the standard deviation of the observational data.

A comparison of model ensemble mean SST with observed SST reveals that model ensemble mean and observations exhibit a similar spatial variability, both in terms of amplitude and patterns (Figure 1c, d). All models display a particularly high correlation (>0.95) as well as comparable standard deviations (8-9°C) to the satellite observations (9°C) (Figure 2b). Among the models, MPI-ESM1.2-HAM is positioned at the outer boundary with a spatial standard deviation of approximately 8°C. In conclusion, all CMIP6 models examined here achieve a much better agreement with the observed spatial patterns of SST than SChl.

3.2 Exploring differences in model performance across temporal scales

Here, we assess the capability of each CMIP6 model to capture the variability of SChl across different timescales (seasonal, sub-seasonal, and multi-annual), as defined in Section 2. Figures 3a and 3b illustrate the SChl variance across these timescales and their respective contributions to the overall SChl variance, averaged globally, in comparison with satellite-derived SChl observations. Our analysis reveals that, for satellite-derived SChl, seasonal variability demonstrates the highest normalised standard deviation (~0.3), followed by sub-seasonal variability (~0.2), and then multi-annual variability (~0.15). The relative contribution of these timescales to the overall SChl variance mirrors this pattern, with seasonal variability accounting for approximately half of the total variance, while sub-seasonal and multi-annual variability contribute 30% and 20% respectively. Despite being based on the same satellite SChl dataset as in Keerthi et al., (2022), we observe a notable reduction in the relative

contribution of sub-seasonal variability to the total SChl variance in the current satellite SChl product. This difference is discussed in section 3.4.

215

It should be noted that biases introduced by gap-filling in satellite-derived data can lead to an inaccurate representation of SChl variability, as missing or interpolated data points may not capture the true temporal or spatial patterns of SChl concentrations. Additionally, uncertainties in the retrieval process, such as atmospheric corrections and sensor calibration, can further distort the observed variability, affecting the reliability of satellite-derived estimates of surface chlorophyll. However, satellite ocean

220 color measurements remain the only available source of high-frequency observations of SChl over extended periods at a global scale. Furthermore, a comparison of SChl at the mooring station BOUSSOLE in the Gulf of Lion showed that satellites can capture SChl variability at high temporal resolution reasonably well (Keerthi et al., 2020). Nonetheless, cloud cover remains a limitation that can affect the accuracy of our analysis.

225 The variability of SChl across different timescales varies significantly among the CMIP6 simulations (Figure 3). With the exception of IPSL-CM6A-LR, IPSL-CM6A-LR-INCA, and CNRM-ESM2-1, most models overestimate the observed standard deviation at both seasonal and sub-seasonal timescales, often exceeding it by a factor of 3. With the exception of MPI-ESM1.2-HAM and MPI-ESM1.2-LR, the distribution of standard deviation among the three defined timescales resembles that observed in satellite SChl: variability at seasonal timescales is the largest followed by sub-seasonal variability and then multi-annual. In

230 contrast, MPI-ESM1.2-HAM and MPI-ESM1.2-LR exhibit the largest variance at sub-seasonal timescales. In the case of IPSL-CM6A-LR and IPSL-CM6A-LR-INCA, there is a slight overestimation of the standard deviation at seasonal timescales and conversely a slight underestimation at sub-seasonal timescales. Multi-annual variance, compared to other timescales, is relatively similar between models. CESM2, CESM2-FV2, CESM2-WACCM-FV2, MPI-ESM1.2-HAM, and MPI-ESM1.2-LR slightly overestimate the observed variance at the multi-annual timescale, whereas CNRM-ESM2-1, NorESM2-LM, and

235 NorESM2-MM models show a slight underestimation.

To compare the relative importance of each component, we calculated the normalized standard deviation for each time series component (SS_t, S_t, MA_t). This normalization allows for a standardized comparison across different locations and variables, providing insight into the dominant modes of variability in the SChl and SST time series.

240

When examining the relative contribution of each component to the total variance, differences between models are more apparent. With the potential exception of MPI-ESM1.2-HR, none of the models accurately replicate the observed decomposition. IPSL-CM6A-LR, IPSL-CM6A-LR-INCA, and CNRM-ESM2-1 overestimate the variance attributed to the seasonal timescale (60-70%), while the remaining 30-40% is evenly distributed between the sub-seasonal and multi-annual

245 components. Conversely, CESM2, CESM2-FV2, CESM2-WACCM-FV2, MPI-ESM1.2-HAM, MPI-ESM1.2-LR, NorESM2-LM, and NorESM2-MM overestimate the relative contribution of sub-seasonal variability (40-50%) and consistently

underestimate the contribution of the multi-annual timescale (5-15%). In these simulations, both seasonal and sub-seasonal variations contribute approximately equally to the total variance, deviating from the observed patterns.

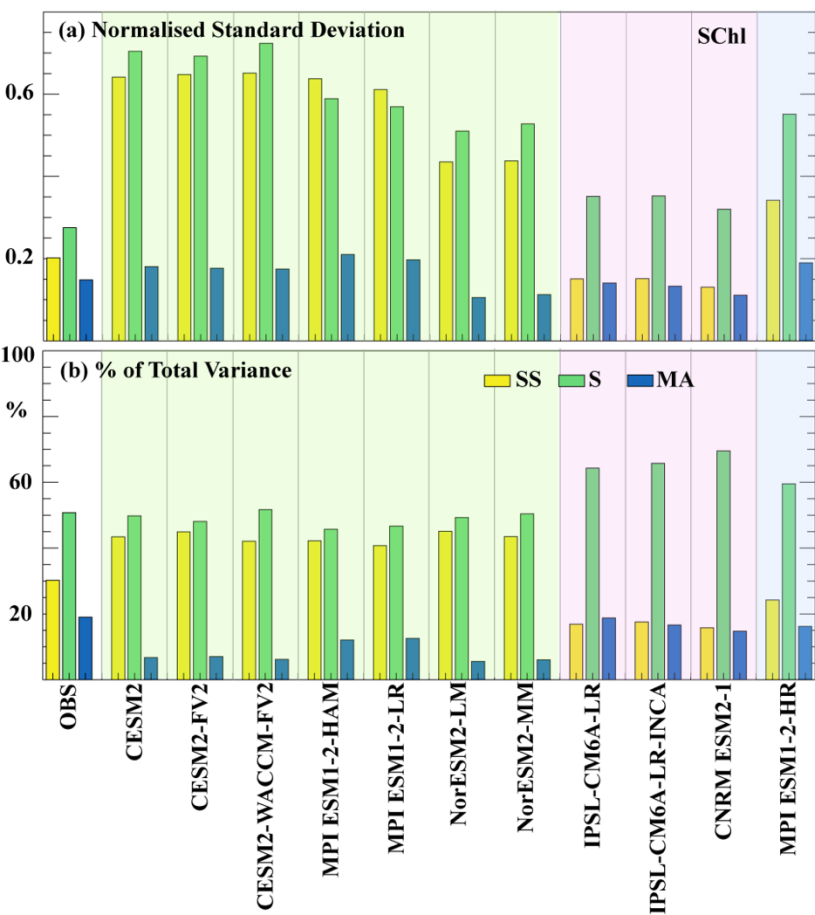



Figure 3: Variability across timescales for SChl : (a) Normalised standard deviation of SChl from observation and CMIP6 historical models. Standard deviation at each grid point is normalised by the mean over each grid. (b) Percentage of SChl variance explained by each component (sub-seasonal, seasonal and multi-annual) for observations and CMIP6 historical models. Shading represents the different model groups described in Section 3.2, with green for Group 1, pink for Group 2, and
255 blue for Group 3.

The CMIP6 SChl simulations can be broadly categorized into three distinct groups based on their performance in capturing SChl temporal variability:

260 **1. Overestimation of sub-seasonal variability**: Models falling into this category, including CESM2, CESM2-FV2, CESM2-WACCM-FV2, MPI-ESM1.2-LR, MPI-ESM1.2-HAM, NorESM-LM, and NorESM-MM, predominantly overestimate the

relative contribution of sub-seasonal variance to the total variance. Consequently, the relative contribution of seasonal and multi-annual timescales is strongly underestimated. They strongly overestimate the observed SChl standard deviation by approximately threefold at both seasonal and sub-seasonal timescales.

265

2. Underestimation of sub-seasonal variability: The models in this category, which are IPSL-CM6A-LR, IPSL-CM6A-LR-INCA, and CNRM-ESM2-1, underestimate both the variance at sub-seasonal timescales and its relative contribution to the total variance. Nevertheless, they approximately reproduce the observed total variance because the variance at seasonal timescales and its relative contribution to total variance are both overestimated.

270

3. Overestimation of total variance but consistent temporal decomposition: The only model in this category is MPI-ESM1.2-HR. This model correctly simulates the relative contribution of the three considered timescales to the total variance. However, the variances are strongly overestimated.

275 The normalised standard deviation across different timescales and the relative contribution of these timescales to the total SST variance (Figure 4) show distinct patterns compared to SChl (Figure 3). As for SChl, the primary driver of natural variability in SST is the seasonal cycle (Keerthi et al., 2022). In observations, the seasonal cycle exhibits the largest standard deviation (0.19) and accounts for approximately 80% of the total SST variance. Multi-annual variability has a standard deviation of 0.05 and explains around 10-12% of the total variance, while sub-seasonal variability is characterized by the lowest standard 280 deviation (0.04) and contributes the least to the total variance (<10%). The multi-annual component makes a relatively minor contribution everywhere, except in the tropics where it is largely related to El Niño–Southern Oscillation (Keerthi et al., 2022). Sub-seasonal variability has a minimal impact on the total SST variance everywhere in the ocean (Keerthi et al., 2022).

Consistent with observations, all simulations exhibit the highest SST standard deviation at the seasonal timescale, followed by 285 the multi-annual and then the sub-seasonal timescale. Across all simulations, seasonal variability accounts for approximately 80% of the total variance, followed by the multi-annual component (~10%). In contrast to SChl, the standard deviation of the sub-seasonal and multi-annual timescales and their relative contribution to the total variance show minimal differences between models and closely resemble observations. All simulations slightly overestimate the standard deviation at the seasonal timescale and its relative contribution (by approximately 5%) whereas both are slightly underestimated at sub-seasonal 290 timescales (by about 5%). Among the models, the IPSL-CM6A-LR and MPI-ESM1.2-LR show a larger overestimation of the observed standard deviation at both seasonal and multi-annual timescales.

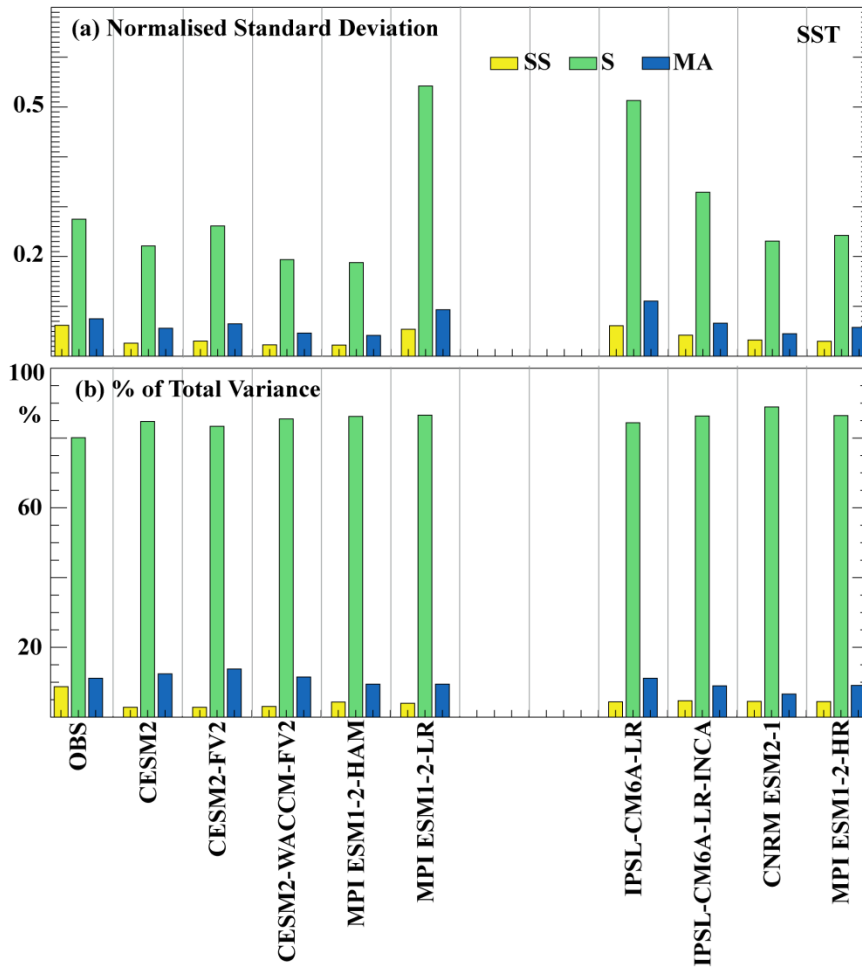



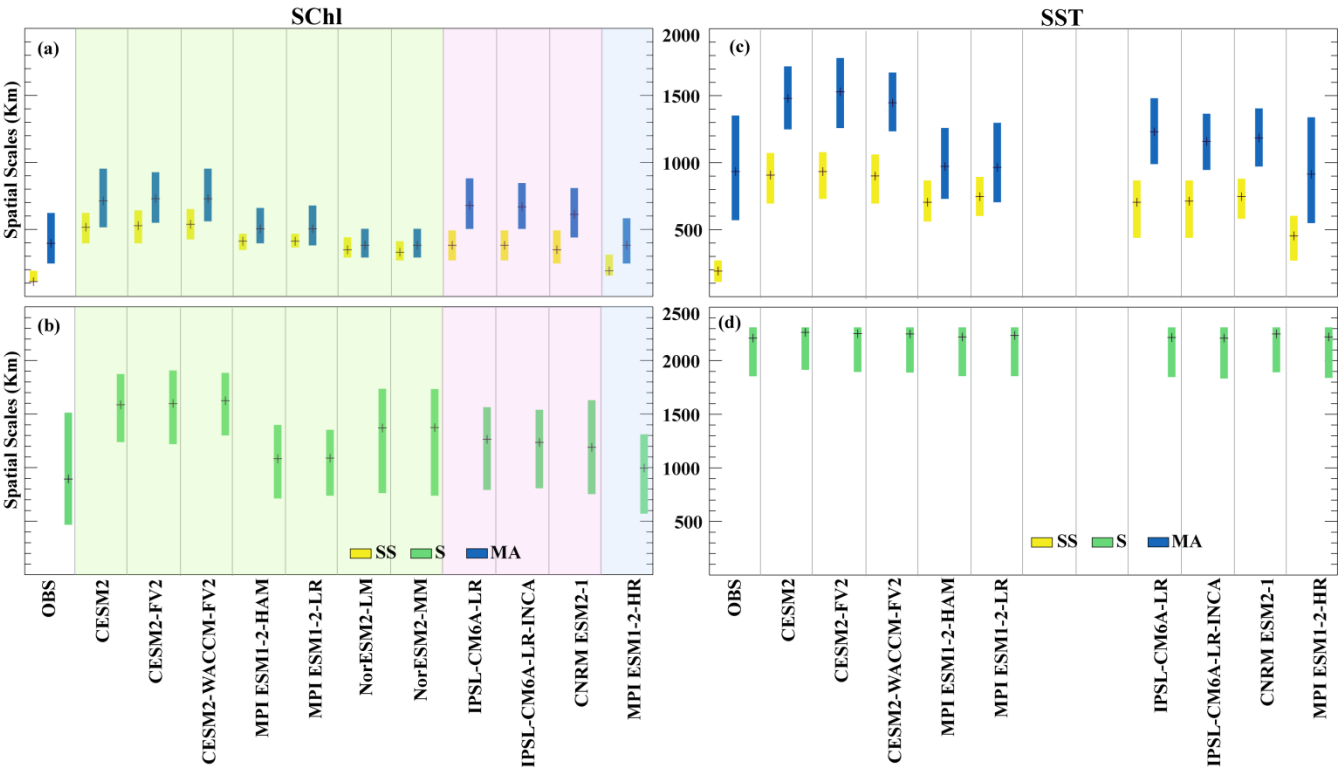
Figure 4: Variability across timescales for SST: Similar to Figure 3, but for SST. (a) Normalised standard deviation of SST from observations and CMIP6 historical simulations. Standard deviation at each grid point is normalised by the mean over each grid. (b) Percentage of SST variance explained by each component (sub-seasonal, seasonal and multiannual) for observations and CMIP6 historical simulations. Note that NorESM2-LM and NorESM2-MM are excluded, as daily SST data for these models are not available on the CMIP6 data portal.

3.3 Spatial scales corresponding to timescales of variability

The evaluation of spatial scales provides insights into the distance over which a signal remains coherent. This helps identify the relevant driving processes at each timescale, aiding understanding of the differences between models and observations.

In the satellite-derived observations of SChl, the seasonal component displays the largest spatial scales, between 500 and 1500 km (Figure 5a, b). These scales are coherent with factors driving seasonality, such as variations in surface stratification and

305 solar irradiance which operate at basin scales. Likewise, in all CMIP6 simulations, the largest spatial scales (>~800 km) correspond to the seasonal cycle. Two groups can be broadly identified among the models (Figure 5a,b). The first group, that includes CESM2, CESM2-FV2, CESM2-WACCM-FV2, NorESM2-LM, and NorESM2-MM, is characterized by the largest spatial scales, approximately ~1500 km. In contrast, in the second group that comprises IPSL-CM6A-LR, IPSL-CM6A-LR-INCA, CNRM-ESM2-1, MPI-ESM1.2-HAM, MPI-ESM1.2-LR and MPI-ESM1.2-HR, spatial scales associated to seasonal

310 variability are smaller, ~1000 km. Similarly, the seasonal cycle of SST is characterized by very large spatial scales exceeding 2000 km which are even greater than those of SChl (Figure 5c, d). Indeed, both in the models and the observations, the computed scales reach the upper limit of 2400 km we set in our methodology which is not the case for SChl.

315 **Figure 5: Spatial scales corresponding to each timescales: The** spatial scales associated with sub-seasonal (yellow), seasonal (green) and multi-annual (blue) variations of SChl (a,b) and SST (c,d). The black line within each box denotes the median. Shading in panel a and b represents the different model groups described in Section 3.2, with green for Group 1, pink for Group 2, and blue for Group 3.

320 For both SChl and SST, the spatial scales corresponding to multi-annual variability are the second largest (Figure 5 a,c). Satellite SChl measurements associate multi-annual variability with spatial scales ranging from about 300 to about 600 km with a median close to 400 km, aligning with climate mode scales on average. However, in the tropics, where this temporal

component dominates, spatial scales can extend beyond these averages (Keerthi et al., 2022). In CMIP6 simulations, the spatial scales simulated for this component vary among models. IPSL-CM6A-LR, IPSL-CM6A-LR-INCA, CNRM-ESM2-1, CESM2, CESM2-FV2, and CESM2-WACCM-FV2 simulate the largest values with a median close to about 700 km, whereas MPI-ESM1.2-HAM, MPI-ESM1.2-LR, MPI-ESM1.2-HR, NorESM2-LM, and NorESM2-MM have comparatively smaller spatial scales, with a median close to that of the observations (~400-500 km). For SST, observational data show spatial scales of about 900 km, while CMIP6 simulations display a broader range from ~500-1800 km. This variability in models may stem from averaging across regions where multi-annual variability is not dominant, given that for both SST and SChl, multi-annual variability is primarily significant only over the equatorial Pacific and Indian Ocean where ENSO is dominant (Keerthi et al., 2022). Additionally, models generally display broader and more intense ENSO patterns that extend too far west compared to observations (Vaittinada Ayar et al., 2023).

Sub-seasonal variability exhibits the smallest spatial scales for both SChl and SST (Figure 5 a,c). In satellite-derived SChl, sub-seasonal variability has spatial scales of around 100-150 km, at the resolution limit of the grid to which we regridded the satellite data (100 km). Using the same satellite product at 25 km resolution, Keerthi et al (2022) identified sub-seasonal spatial scales of around 50 km. Most simulated sub-seasonal variability of SChl, except in MPI-ESM1.2-HR, is characterized by spatial scales exceeding 350 km. Among the simulations, the CESM2 models exhibit the largest scales close to 500 km, followed by the MPI models (excluding MPI-ESM1.2-HR), then IPSL, CNRM, and NorESM2 models. MPI-ESM1.2-HR, an eddy-permitting model, exhibits the smallest scales (~200 km) for sub-seasonal timescales, although still larger than in observations. The mean spatial scale corresponding to SST sub-seasonal variability is ~200 km in observations. In contrast, simulations, with the exception of MPI-ESM1.2-HR, display scales exceeding 600 km.

3.4 Drivers of sub-seasonal variability

Sub-seasonal SChl fluctuations result from various drivers across different spatial scales: submesoscale/mesoscale processes (1–100 km), cyclones and tropical storms (100–1,000 km), large-scale climate modes (>1,000 km) and internal variability. A prior study (Keerthi et al., 2022), based on satellite ocean color SChl measurements, suggests that sub-seasonal variability is strongly associated with mesoscale and submesoscale variations, as evidenced by their mean spatial scales of about 50 km. However, in the high latitudes and tropics where sub-seasonal variability has a large contribution to the SChl total variability, >50% of the sub-seasonal variability has spatial scales greater than 100 km. At high latitudes, these large spatial scales reflect synoptic storms (Prend et al., 2022; Keerthi et al., 2022; Thomalla et al., 2011), while in the tropics, they reflect intraseasonal climate modes such as Madden Julian Oscillations, Kelvin waves, and tropical instability waves (Keerthi et al., 2022; Resplandy et al., 2009; Strutton et al., 2001; Xu et al., 2018; Jin et al., 2013).

We observe a significant reduction in the relative contribution of sub-seasonal variability to the total SChl variance in the satellite SChl product compared to that in Keerthi et al., (2022). This disparity can be attributed to the spatial regridding process used to analyse the satellite SChl data at a mean horizontal resolution comparable to that of the CMIP6 models. In the present study, the horizontal resolution of the satellite product is approximately 100 km, while in Keerthi et al., (2022), it is 25 km. Since the mean spatial scale of sub-seasonal variability has been shown to be of the order of 50 km (Keerthi et al., 2022), regridding to coarser resolution removes substantial variability. Due to effective coarse-resolution, which is several times lower than grid resolution (Levy et al., 2012), CMIP6 simulations should structurally underestimate sub-seasonal variability. However, contrary to this expectation, most models overestimate the relative contribution of sub-seasonal variability to the total SChl variance. Furthermore, the mean spatial scale is about 4 to 6 times larger than in observations.

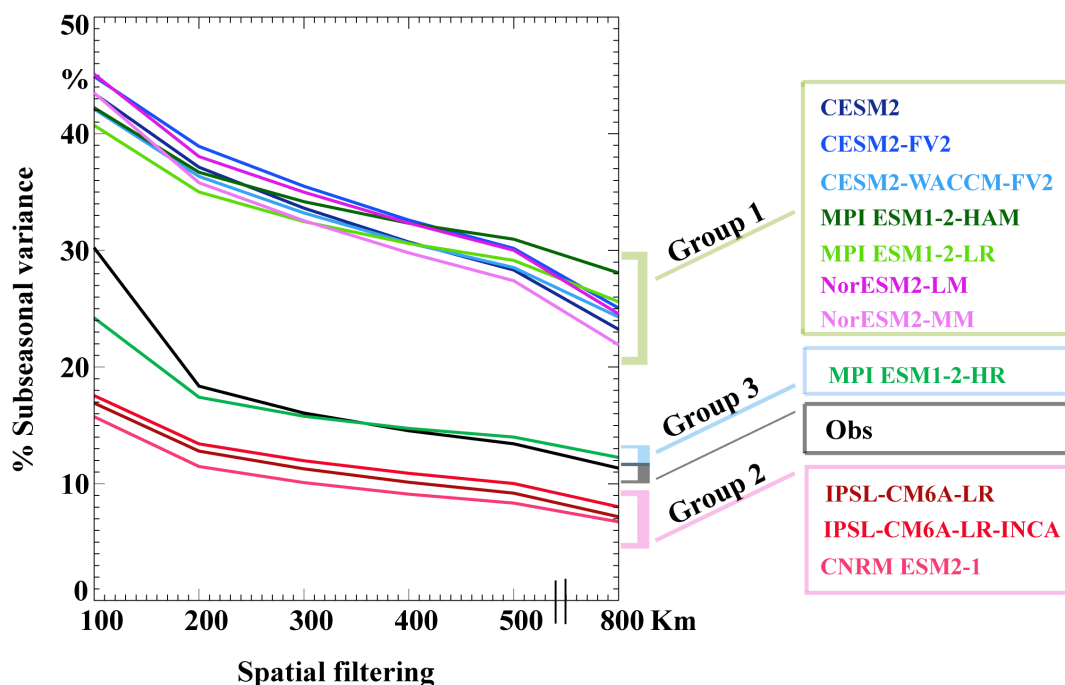



Figure 6: Sub-seasonal variance across spatial scales: Percentage of total SChl variance explained by sub-seasonal variability after applying a spatial smoothing of varying scale from 100 km to 800 km to the SChl data. Coloured boxes for the model names represent the different model groups described in Section 3.2, with green for Group 1, pink for Group 2, and blue for Group 3.

When the sub-seasonal variance is assessed at spatial scales from 100 to 800 km, a clear pattern emerges in the SChl data. The relative contribution of sub-seasonal variations to total SChl variance in satellite observations drops from 30% at 100 km to 18% at 200 km, followed by a gradual decline up to 800 km (Figure 6). Using the categories of models defined in section 3.2, the model in category 3 (MPI-ESM1.2-HR) shows a decline of 7% from 100 km to 200 km to values relatively close to the

observations. Above 200 km, the simulated sub-seasonal relative contribution then decreases similar to observations. Models in category 2 also exhibit a pattern similar to that of observations with a modest decrease of (~5%) from 100-200km and then a gradual decrease up to 800km. But these models underestimate the contribution of sub-seasonal variability by about 4-5% at all spatial scales from 200-800km. This suggests that these models correctly simulate the large-scale component of sub-seasonal variability but are unable to capture its small-scale component as expected due to their limited resolution. Category 1 models systematically overestimate the sub-seasonal contribution by about 20%. In addition, the downward slope tends to be steeper than observed, particularly in CESM2 configurations. We hypothesize that the large sub-seasonal variability simulated by models in this category is generated by driving mechanisms that differ from observations.

The discrepancies between the three model categories in capturing SChl sub-seasonal variability are also evident in an analysis of temporal subscales (Figure 7). Satellite observations show that 40-50% of sub-seasonal variability is present in the 16-32day temporal window, followed by 20% in the 33-48day window, and 0-15% in days 49-64. Category 2 and 3 models show a similar pattern, but tend to underestimate the contribution of the 16-32day period, probably due to an underestimation of mesoscale variability. In contrast, all other simulations predominantly overestimate sub-seasonal variability in the 16-32day period.

Much of the phytoplankton variability is often attributed to fluctuations in their physical environment. However, phytoplankton time series also exhibit variability that is not strongly correlated with key physical variables and is distinctly nonlinear (Mayersohn et al., 2021, 2022). These intrinsic oscillations are primarily associated with two mechanisms: one related to species succession and resource changes (Tilman, 1977; Huisman and Weissing, 1999, 2001), and the other to changes in total phytoplankton biomass due to predator-prey interactions (Gilpin, 1979; Edwards and Brindley, 1996). Predator-prey oscillations typically occur on shorter subseasonal time scales of up to 60 days while resource-related oscillations can extend to longer time scales (Mayersohn et al., 2021, 2022). This intrinsic variability can complicate the accurate simulation of high-frequency, subseasonal fluctuations in phytoplankton populations by models.

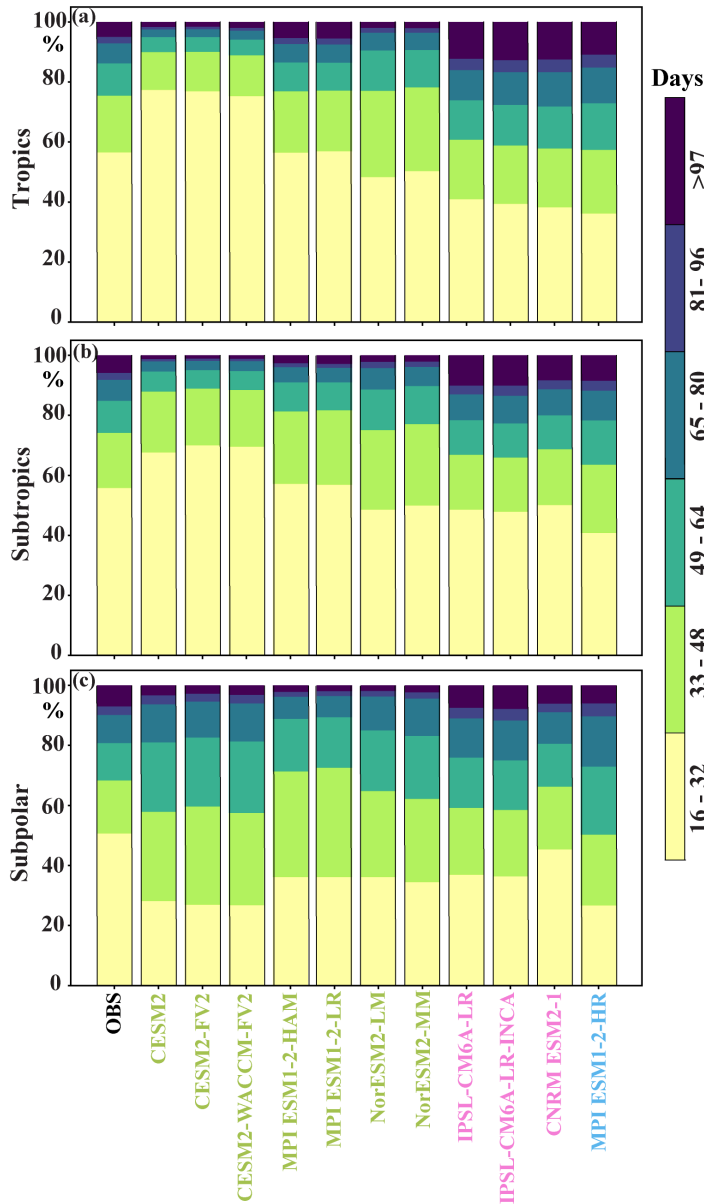



Figure 7: **Sub-seasonal SChl variability across temporal subperiods:** Bar plot showing the relative contribution of each temporal subperiod to the total SChl sub-seasonal variance in the observations and different CMIP6 historical simulations for (a) Tropics (20°N-20°S), (b) Subtropics (20°N-20°N and 20°S-40°S) and (c) Subpolar (40°N-60°N and 40°S-60°S) regions. Each temporal subperiod corresponds to a 16-day window within the sub-seasonal band. The colors assigned to the names of the CMIP6 simulations indicate the model groups discussed in Section 3.2, with green representing Group 1, pink representing Group 2, and blue representing Group 3.

A recent study (Rohr et al., 2023) emphasizes significant differences in CMIP6 simulations regarding the representation of zooplankton-specific grazing, which could have a profound impact on the temporal variability of phytoplankton (Mayersohn et al., 2021, 2022; Talmy et al., 2024). Despite the diversity of functional roles and distributions of zooplankton species (Kiorboe, 2011; Benedetti et al., 2023; Pata and Hunt, 2023), biogeochemical (BGC) models must represent aggregated behaviour using a limited number of zooplankton groups. This limitation introduces considerable uncertainty into the modeling of complex zooplankton communities and their role in the marine carbon cycle. The representation of grazing in CMIP-class BGC models varies considerably, from models with a single zooplankton functional type grazing on a specific phytoplankton type to those incorporating multiple zooplankton groups and potential preys (Sailley et al., 2013; Petrik et al., 2022; Rohr et al., 2023). Beyond differences in grazing formulations, there are significant variations between models in terms of phytoplankton groups/size classes, temperature-dependent phytoplankton growth, biogeochemical factors influencing phytoplankton growth, and resource competition (Séférian et al., 2020), further contributing to the overall uncertainty in model simulations. For instance, many studies have shown that the mathematical formulation and the parameter values of the closure term representing predation by unresolved higher trophic levels have profound impacts on the temporal stability of the biogeochemical model, especially at high resource levels (e.g., Edwards and Brindley, 1999; Edwards and Yool, 2000; Omta et al., 2023).

Analysing the relative significance of each factor—biogeochemical models, ocean dynamical models, and horizontal resolution—influencing the observed differences between simulations in capturing the temporal variability of SChl proves challenging. For example, even though MPI-ESM1.2-HR and MPI-ESM1.2-LR use similar ocean, atmosphere, and biogeochemical models, they show notable differences in the simulation of SChl temporal variability, with improved horizontal resolution playing a crucial role. Interestingly, models like IPSL and CNRM, classified as type 2, which differ from MPI-ESM1.2-HR in ocean, atmosphere, and biogeochemical components and utilize a coarse horizontal resolution (100 km), show comparable, if not superior, performance in simulating SChl temporal variability. This is despite a minor underestimation in sub-seasonal variability contribution to the total SChl variance. Furthermore, the MPI and NorESM models, which share the same biogeochemical model but not the same physical ocean model, simulate contrasting SChl variability.

3.5 Role of sub-seasonal variability in year-to-year variations of SChl

Recent studies have highlighted the significant role of sub-seasonal variability in modulating annual variations in SChl (Keerthi et al., 2022; Prend et al., 2022; Levy et al., 2024). In the Southern Ocean and western and eastern boundary upwelling systems, high frequency sub-seasonal SChl variations can accumulate and modulate the year-to-year variations in SChl. We quantify the impact of sub-seasonal variability on year-to-year SChl variations using the annual mean low-frequency index, following

the methodology of Keerthi et al., (2022). This index is computed as the squared correlation between the annual mean SChl and the mean multi-annual component of SChl for each year. An index approaching 1 implies that year-to-year variations are primarily associated with multi-annual variations, while an index less than 1 indicates a contribution of sub-seasonal and seasonal variability to year-to-year variations in SChl. A smaller index thus signifies a greater contribution of sub-seasonal variability to year-to-year variations.

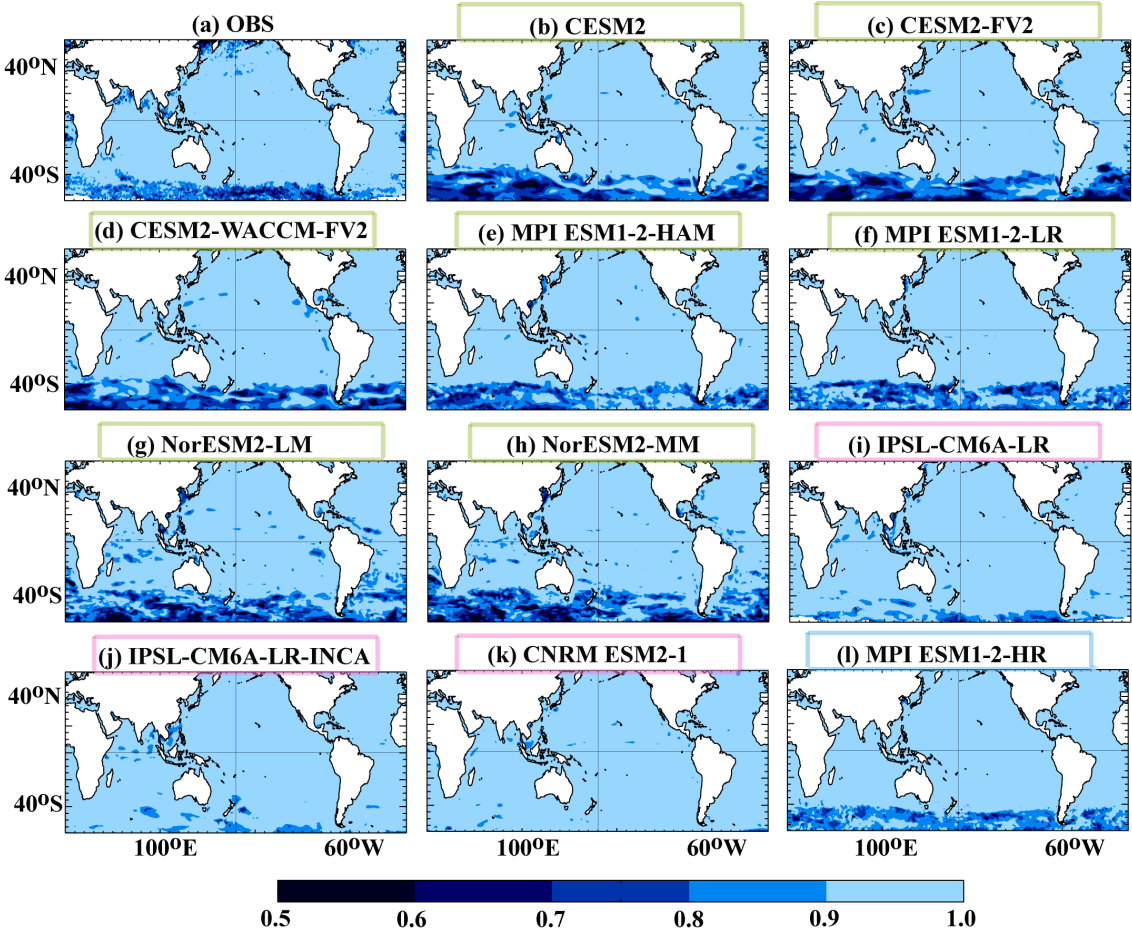



Figure 8: Annual mean low-frequency index for SChl, which is defined as the correlation square between annual mean and annual mean of the multi-annual component. When the index is close to one, year-to-year fluctuations in the annual mean reflect low-frequency variability. The value of the index decreases as high-frequency variability contributes more to year-to-year variations. Coloured boxes for the model names represent the different model groups described in Section 3.2, with green for Group 1, pink for Group 2, and blue for Group 3.

For satellite-derived SChl data with a horizontal resolution of 25km, consistent with Keerthi et al. (2022), the index is very close to 1 in parts of the tropics and subtropics. However, it decreases, reaching values as low as 0.5, in regions such as the equatorial Atlantic, the Southern Ocean, western boundary current regions and eastern boundary upwelling systems, where sub-seasonal variability and irregular seasonal cycles intensify. These regions are characterized by substantial eddy and frontal activity. When regridded to a 100km horizontal resolution, the index is close to 1 over most of the ocean, except in the southern subpolar regions and in the northern Indian Ocean (Figure 8a). This implies that much of the sub-seasonal variability that imprints on year-to-year variations is due to mesoscale processes.

The distribution of the index computed for the CMIP6 models is presented in Figure 8 b-l. Models simulate contrasting distributions which can be structured according to the three categories defined in section 3.2. Category 2 models, with the exception of CNRM ESM2-1, show sporadic contributions in the Southern Ocean. All models in category 1 strongly overestimate the impact of sub-seasonal variability on year-to-year variations in the Southern Ocean. The category 3 model shows a similar pattern to observations with a slight overestimation. This is further supported by regression of annual mean SChl with the annually-averaged sub-seasonal component of SChl (Supplementary Figure 1). In the observations, sub-seasonal variability contributes 10 to 30% of the amplitude of year-to-year variations in the Southern Ocean. In models, with the exception of category 2 models, this contribution is significantly larger, exceeding 30% in large parts of the Southern Ocean.

3.6 Projected changes to SChl temporal variability

Under climate change scenarios, ESMs consistently project increased stratification and a reduction in nutrient concentrations in the euphotic zone (Bopp et al., 2001; Sarmiento et al., 2004; Cabré et al., 2014; Fu et al., 2016). These changes generally lead to an overall reduction in net primary production due to increased limitation of phytoplankton growth by nutrients (Bopp et al., 2013; Krumhardt et al., 2017; Moore et al., 2018), although results from CMIP6 models show a more modest reduction associated with greater uncertainty than CMIP5 models (Kwiatkowski et al., 2020; Tagliabue et al., 2021). In addition to these simulated mean changes, global warming has been also shown to alter seasonal cycles (Henson et al., 2013; Thomalla et al., 2023), to modify interannual and decadal climate modes (Cai et al., 2014, 2021), and to increase the frequency of extreme events like heatwaves and tropical cyclones (Frolicher et al., 2018; Knutson et al., 2020; Walsh et al., 2016; Jo et al., 2022). The multi-model mean seasonal amplitude of global SST is projected to increase by +0.59±0.21°C under SSP5-8.5 (Kwiatkowski et al., 2020), mainly resulting from an overall shoaling and increasing seasonal amplitude of the mixed layer (Alexander et al., 2018; Jo et al., 2022). By the end of the 21st century, most models forecast an increase in frequency and amplitude of central Pacific El Niño events and a rise in the frequency of eastern Pacific El Niño events (Vaittinada Ayar et al., 2023). The increased frequency of extreme events such as marine heatwaves (Frolicher et al., 2018) and tropical cyclones (Knutson et al., 2020; Walsh et al., 2016), coupled with mesoscale and submesoscale variability linked to global warming scenarios (Martínez-Moreno et al., 2021, 2022), contributes to an increase in sub-seasonal variability. With respect to the SChl

temporal variability, our knowledge of its potential changes is more limited. Specifically, the simulated impact of climate change on the sub-seasonal variability of SChl has, to the best of our knowledge, not been previously assessed in CMIP-type models.

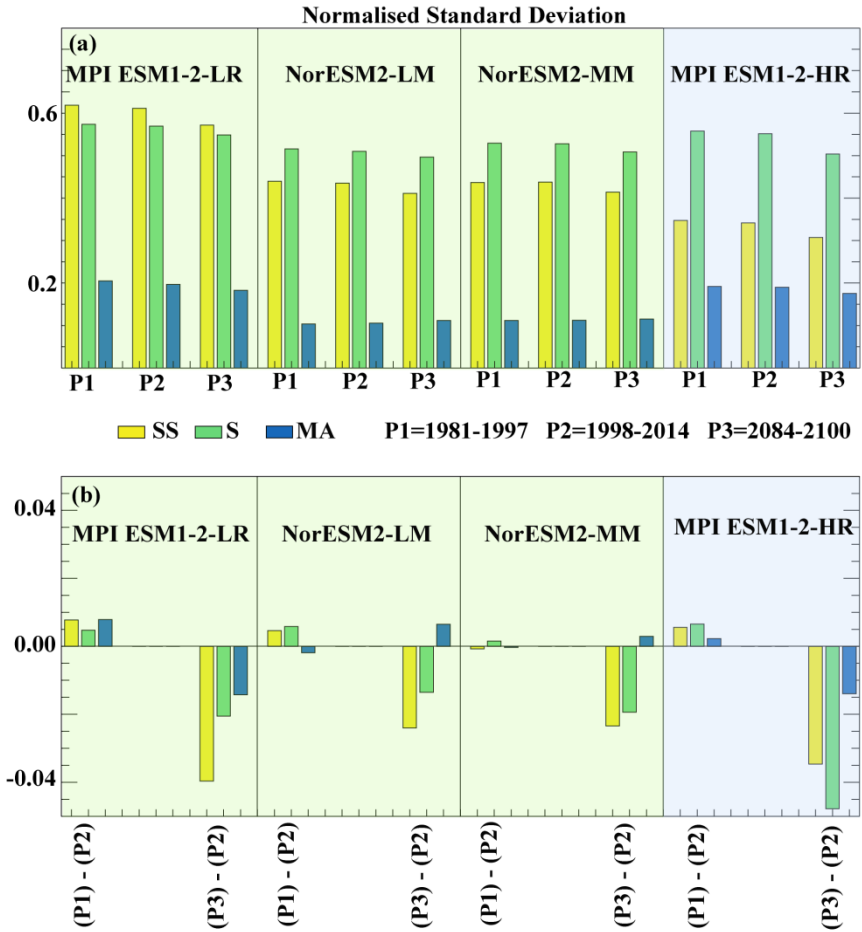



Figure 9: Future projections (a) Normalised standard deviation of SChl from two different periods of CMIP6 historical simulations (1981-1997 & 1998-2014) and CMIP6 SSP585 simulation for the period (2084-2100). (b) Difference between different periods are shown. Shading in panels a and b represents the different model groups described in Section 3.2, with green for Group 1, and blue for Group 3.

Here, we analyzed the changes in SChl temporal variability under the high-warming scenario SSP5-8.5, using simulations that provide daily SChl outputs to be able to include sub-seasonal variability in our analysis. SChl standard deviations at sub-seasonal, seasonal and multi-annual timescales for the end of the century (2084–2100) and for the end of historical simulations (1998-2014) are compared in Figure 9. To determine whether the changes noted between the two periods can be explained by

decadal variability, we also applied our analysis to the 1981-1997 period. All models simulate a consistent decrease in the normalized standard deviation of both seasonal and sub-seasonal timescales from over the 21st century. The MPI models tend to simulate a stronger decrease at both timescales (7-10% sub-seasonally and about 3-8% seasonally), with larger changes in the high resolution configuration. In NorESM models the simulated decrease is smaller, 4-5% and 2-3% at sub-seasonal and seasonal timescales respectively. At multi-annual timescales, changes are of a similar relative magnitude, about 3-8% but show opposite signs between MPI and NorESM models. Comparison with the period 1981-1997 and a similar analysis carried out with the piControl experiments (Figure S2) shows that these changes cannot be explained by decadal natural variability.

4 Summary

In this study, we assessed how ESMs participating in CMIP6 reproduce surface phytoplankton variations across various temporal scales, with a particular focus on the often-overlooked sub-seasonal timescales. We compared 13 ESMs that have daily SChl outputs for the historical period (1998-2014) with the ESA OC-CCI merged ocean color satellite SChl product. Unlike SST, where ESMs generally exhibit consistent behaviour, we find significant intermodel variability and discrepancies between SChl simulations and observational data, both in terms of the amplitude of variability and likely driving mechanisms. Our findings indicate that none of the analyzed models accurately replicate both the observed variability across timescales and their relative contributions to the total temporal variance in SChl.

Based on globally averaged metrics of sub-seasonal timescales we categorized the models into three distinct groups. Group 1 models strongly overestimate sub-seasonal SChl standard deviation and its relative contribution to total variance, despite their coarser horizontal resolution compared to observations. Group 2 models better represent the observed SChl variability across timescales but underestimate the sub-seasonal variance and its relative contribution to total variance. These models capture large-scale sub-seasonal variability but fail to resolve small-scale components, partly due to their resolution, which does not permit mesoscale processes—a bias that can be reduced with higher-resolution models. Group 3 models correctly simulate the relative contribution of different timescales to total variance but significantly overestimate SChl variances. This overestimation of sub-seasonal variance in Group 1 and Group 3 models is possibly due to intrinsic oscillations (e.g., predator-prey oscillations) inconsistent with observations and potentially stemming from the structure of the biogeochemical models. Models that overestimate sub-seasonal variability exaggerate its influence on annual variations, potentially impacting long-term trends. In contrast, Group 2 models exhibit a diminished impact of sub-seasonal variations on annual variations, which could also influence long-term projections.

Overall, our findings highlight the challenges and discrepancies in ESM representation of surface phytoplankton dynamics, emphasizing the crucial role of spatial resolution and the accurate representation of biogeochemical processes in determining model accuracy. A direct relationship between model performance and horizontal resolution might not always exist. Group 2

models, despite having a lower resolution comparable to Group 1 models, exhibit comparatively better performance among the models analyzed. However, increasing the resolution of both the atmospheric and ocean components of a model significantly improved its performance.

By the end of the 21st century, models project a modest global decrease in both seasonal and sub-seasonal variability. However, projected changes on multi-annual timescales diverge. This analysis is however limited by the number of models that provide daily output for the historical and future periods (only four ESMs). Furthermore, we hypothesize that substantial differences in zooplankton grazing dynamics among CMIP6 models (Rohr et al., 2023) lead to notable variations in higher-frequency timescales. However, this hypothesis remains challenging to confirm without daily zooplankton data. We therefore advocate that, in future exercises, more modeling groups submit daily surface outputs of biogeochemical variables, particularly SChl, along with zooplankton concentrations and grazing rates. The poor model performance in simulating sub-seasonal SChl dynamics raises concerns about the reliability of projections at these timescales and potentially undermines long-term projections, given the influence of sub-seasonal dynamics on year-to-year variability.

Acknowledgments

The authors acknowledge the support from the ENS-PSL CHANEL chair and the Centre National d'Etudes Spatiales (TOSCA).

Author Contributions

MGK, OA, ML and LK conceived and developed the study. MGK performed the data analysis and made the plots. MGK, OA, LK and ML made the interpretation of the results and wrote the manuscript.

Data Availability

All data analyzed in this study are freely available from the respective websites mentioned in the Methods Section.

Competing interests' statement:

The authors declare no competing interests

References

Alexander, M. A., Scott, J. D., Friedland, K. D., Mills, K. E., Nye, J. A., Pershing, A. J., and Thomas, A. C.: Projected sea
surface temperatures over the 21st century: Changes in the mean, variability and extremes for large marine ecosystem regions
of Northern Oceans, Elem. Sci. Anth., 6, 9, https://doi.org/10.1525/elementa.191, 2018.

Behrenfeld, M. J. et al. : Climate-driven trends in contemporary ocean productivity. Nature, 444(7120), 752-755, 200.

Benedetti, F., Gruber, N., & Vogt, M. : Global gradients in species richness of marine plankton functional groups. Journal of
Plankton Research, 45(6), 832-852, 2023.

Bentsen, M., Oliviè, D.J.L., Seland, O., Toniazzo, T., Gjermundsen, A. et al.: NCC NorESM2-MM model output prepared
for CMIP6 CMIP historical. Version 20191108, Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8040,
2019a.

Bentsen, M., Oliviè, D.J.L., Seland, O., Toniazzo, T., Gjermundsen, A. et al.: NCC NorESM2-MM model output prepared
for CMIP6 ScenarioMIP ssp585. Version 20230616, Earth System Grid Federation.
https://doi.org/10.22033/ESGF/CMIP6.8321, 2019b.

Bindoff, N.L., W.W.L. Cheung, J.G. Kairo, J. Arístegui, V.A. Guinder, R. Hallberg, et al.: Changing ocean, marine ecosystems,
and dependent communities. IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.!O. Pörtner, D.C.
Roberts, V. Masson!Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J.
Petzold, B. Rama, N.M. Weyer(eds.)], https://www.ipcc.ch/srocc/chapter/chapter!5/, 2019.

Bopp, L., Monfray, P., Aumont, O., Dufresne, J.-L., Treut, H. L., Madec, G., Terray, L., and Orr, J. C.: Potential impact of
climate change on marine export production, Global Biogeochem. Cy., 15, 81–99, https://doi.org/10.1029/1999GB001256,
2001.

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

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R.,
Tjiputra, J., and Vichi, M.: Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models,
Biogeosciences, 10, 6225–6245, https://doi.org/10.5194/bg-10- 6225-2013, 2013.

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A. et al.: IPSL IPSL-CM6A-LR model output prepared for CMIP6
CMIP historical. Version 2021, Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5195, 2018.

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A. et al.: IPSL IPSL-CM6A-LR-INCA model output prepared
for CMIP6 CMIP historical. Version 2021, Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.13601,
2021.

Cabré, A., Marinov, I., and Leung, S.: Consistent global responses of marine ecosystems to future climate change across the
IPCC AR5 earth system models, Clim. Dynam., 45, 1253–1280, https://doi.org/10.1007/s00382-014-2374-3, 2014.

580 Cai, W., Santoso, A., Wang, G., Yeh, S.W., An, S.I., Cobb, K.M., Collins, M., Guilyardi, E., Jin, F.F., Kug, J.S. and Lengaigne, M.: ENSO and greenhouse warming. Nature Climate Change, 5(9), pp.849-859, 2015.

Cai, W., Santoso, A., Collins, M., Dewitte, B., Karamperidou, C., Kug, J.S., Lengaigne, M., McPhaden, M.J., Stuecker, M.F., Taschetto, A.S. and Timmermann, A.: Changing El Niño–Southern oscillation in a warming climate. Nature Reviews Earth & Environment, 2(9), pp.628-644, 2021.

585 Carranza, M. M. & Gille, S. T: Southern Ocean wind-driven entrainment enhances satellite chlorophyll-a through the summer. Journal of Geophysical Research: Oceans, 120, 304-323, 2015.

Danabasoglu, G.: NCAR CESM2 model output prepared for CMIP6 CMIP historical, Version 20190308, Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.7627, 2019a.

Danabasoglu, G.: NCAR CESM2-FV2 model output prepared for CMIP6 CMIP historical, Version 20191120 , Earth System
590 Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.11297, 2019b.

Danabasoglu, G.: NCAR CESM2-WACCM-FV2 model output prepared for CMIP6 CMIP historical, Version 20191120, Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.11298, 2019c.

Demarcq, H., Reygondeau, G., Alvain, S. & Vantrepotte, V: Monitoring marine phytoplankton seasonality from space. Remote sensing of environment, 117, 211-222, 2012.

595 Doney, S.C., L. Bopp, and M.C. Long: Historical and future trends in ocean climate and biogeochemistry. Oceanography 27(1):108–119, http://dx.doi.org/10.5670/oceanog.2014.14, 2014.

Edwards, A.M., Brindley, J.: Oscillatory behaviour in a three-component plankton population model. Dyn. Syst. 11 (4), 347–370. http://dx.doi.org/10.1080/ 02681119608806231, 1996.

Edwards, A.M., and Brindley, J.: Zooplankton mortality and the dynamical behaviour of plankton population models, Bulletin
600 of mathematical biology 61, no. 2, 303-339, 1999.

Edwards, A.M., and Yool, A.: The role of higher predation in plankton population models, Journal of Plankton Research 22, no. 6, 1085-1112, 2000.

Frölicher, T. L., Fischer, E. M., & Gruber, N.: Marine heatwaves under global warming. Nature, 560(7718), 360-364, 2018.

Fu, W., Randerson, J. T., and Moore, J. K.: Climate change impacts on net primary production (NPP) and export production
605 (EP) regulated by increasing stratification and phytoplankton community structure in the CMIP5 models, Biogeosciences, 13, 5151–5170, https://doi.org/10.5194/bg-13-5151-2016, 2016.

Gaube, P., McGillicuddy Jr, D. J., Chelton, D. B., Behrenfeld, M. J., & Strutton, P. G. Regional variations in the influence of mesoscale eddies on near-surface chlorophyll. Journal of Geophysical Research: Oceans, 119(12), 8195-8220, 2014.

Gilpin, M.E.: Spiral chaos in a predator-prey model. Amer. Nat. 113 (2), 306–308, 1979.

610 Gregg, W.W., Casey, N.W. and McClain, C.R.: Recent trends in global ocean chlorophyll. Geophysical research letters, 32(3), 2005

Henson, S. A. et al.: Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. Biogeosciences, 7(2), 621-640, 2010.

Henson, S., Cole, H., Beaulieu, C., & Yool, A.: The impact of global warming on seasonality of ocean primary production. Biogeosciences, 10(6), 4357–4369, 2013.

Henson, S. A., Beaulieu, C. & Lampitt, R. Observing climate change trends in ocean biogeochemistry: when and where. Global change biology, 22(4), 1561-1571, 2016.

Huisman, J. & Weissing, F. J.: Biodiversity of plankton by species oscillations and chaos. Nature 402, 407–410, 1999.

Huisman, J., & Weissing, F. J.: Biological conditions for oscillations and chaos generated by multispecies competition. Ecology, 82(10), 2682–2695, 2001.

Jin, D., Waliser, D. E., Jones, C. & Murtugudde, R. Modulation of tropical ocean surface chlorophyll by the Madden–Julian Oscillation. Climate dynamics, 40(1-2), 39-58, 2013.

Jo, A.R., Lee, J.Y., Timmermann, A., Jin, F.F., Yamaguchi, R. and Gallego, A.: Future amplification of sea surface temperature seasonality due to enhanced ocean stratification. Geophysical Research Letters, 49(9), p.e2022GL098607, 2022.

Jungclaus, J., Bittner, M., Wieners, k., Wachsmann, F., Schupfner, M., et al.: MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP piControl. Version 20190710 ,Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6674, 2019a.

Jungclaus, J., Bittner, M., Wieners, k., Wachsmann, F., Schupfner, M., et al.: MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP historical. Version 2019].Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6594, 2019b.

Keerthi, M. G., Lengaigne, M., Vialard, J., de Boyer Montégut, C. & Muraleedharan, P. M.: Interannual variability of the Tropical Indian Ocean mixed layer depth. Climate dynamics, 40(3), 743-759, 2013.

Keerthi, M.G., Lengaigne, M., Drushka, K., Vialard, J., de Boyer Montégut, C., Pous, S., Levy, M. and Muraleedharan, P.M., 2016.: Intraseasonal variability of mixed layer depth in the tropical Indian Ocean. Climate dynamics, 46, pp.2633-2655, 2016.

Keerthi, M. G., Levy, M., Aumont, O., Lengaigne, M. & Antoine, D.: Contrasted contribution of intraseasonal time scales to surface chlorophyll variations in a bloom and an oligotrophic regime. Journal of Geophysical Research: Oceans, 125(5), e2019JC015701, 2020.

Keerthi, M. G., Lévy, M. & Aumont, O.: Intermittency in phytoplankton bloom triggered by modulations in vertical stability. Scientific reports, 11(1), 1-9, 2021.

Keerthi MG, Prend CJ, Aumont O, Levy M.: Annual variations in phytoplankton biomass driven by small-scale physical processes. Nature Geoscience :1–14, 2022.

Kessler A, Tjiputra J.: The Southern Ocean as a constraint to reduce uncertainty in future ocean carbon sinks. Earth Syst Dynam, 7:295–312. https://doi.org/10.5194/esd-7-295-2016, 2016.

Kiorboe, T.: How zooplankton feed: mechanisms, traits and trade-offs. Biological reviews, 86(2), pp.311-339, 2011.

Knutson, T. et al. Tropical cyclones and climate change assessment: Part II: Projected response to anthropogenic warming. Bulletin of the American Meteorological Society, 101(3), E303-E322, 2020.

Krumhardt, K. M., Lovenduski, N. S., Long, M. C., and Lindsay, K.: Avoidable impacts of ocean warming on marine primary production: Insights from the CESM ensembles, Global Biogeochem. Cy., 31, 114–133, https://doi.org/10.1002/2016GB005528, 2017.

650 Kwiatkowski, L., Bopp, L., Aumont, O., Ciais, P., Cox, P. M., Laufkötter, C., Li, Y., and Séférian, R.: Emergent constraints on projections of declining primary production in the tropical oceans, Nat. Clim. Change, 7, 355–358, https://doi.org/10.1038/nclimate3265, 2017.

Kwiatkowski, L., Aumont, O., and Bopp, L.: Consistent trophic amplification of marine biomass declines under climate change, Glob. Change Biol., 25, 218–229, https://doi.org/10.1111/gcb.14468, 2018.

655 Kwiatkowski, L, O Torres, L Bopp, O Aumont, M Chamberlain, J R. Christian, J P. Dunne et al.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections." Biogeosciences 17, no. 13, 3439-3470, 2020.

Lévy, M., Ferrari, R., Franks, P. J., Martin, A. P., & Rivière, P. Bringing physics to life at the submesoscale. Geophysical Research Letters, 39(14), 2012.

660 Lévy, M., Couespel, D., Haëck, C., Keerthi, M.G., Mangolte, I. and Prend, C.J.: The impact of fine-scale currents on biogeochemical cycles in a changing ocean. Annual Review of Marine Science, 16(1), pp.191-215, 2024.

Lovenduski NS, Gruber N.: Impact of the Southern Annular Mode on Southern Ocean circulation and biology. Geophysical Research Letters 32(11), 2005.

Martinez, E., Raitsos, D. E. & Antoine, D. Warmer, deeper, and greener mixed layers in the North Atlantic subpolar gyre over 665 the last 50 years. Global change biology, 22(2), 604-612, 2016.

Martínez-Moreno, J. et al. Global changes in oceanic mesoscale currents over the satellite altimetry record. Nature Climate change, 1–7, 2021.

Martínez-Moreno J, Hogg AM, England MH.: Climatology, Seasonality, and Trends of Spatially Coherent Ocean Eddies. J. Geophys. Res. Ocean 127(7): e2021JC017453, 2022.

670 Mayersohn, B., Smith, K. S., Mangolte, I. & Lévy, M. Intrinsic timescales of variability in a marine plankton model. Ecological Modelling, 443, 109446, 2021.

Mayersohn, B., L.vy, M., Mangolte, I., & Smith, K. S.: Emergence of broadband variability in a marine plankton model under external forcing. Journal of Geophysical Research: Biogeosciences, 127, e2022JG007011. https://doi.org/10.1029/2022JG007011, 2022.

675 Moore, J. K., Fu, W., Primeau, F., Britten, G. L., Lindsay, K., Long, M., Doney, S. C., Mahowald, N., Hoffman, F., and Randerson, J. T.: Sustained climate warming drives declining marine biological productivity, Science, 359, 1139–1143, https://doi.org/10.1126/science.aao6379, 2018.

Müller,W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., and Ilyina, T.: A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), J. Adv. Model. 680 Earth Syst., 10, 1383–1413, https://doi.org/10.1029/2017MS001217, 2018.

Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stoll, J., Folini, D.S., Tegen, I. et al.: HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP historical. Version 2019].Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.5016, 2019.

Omta, A.W., E.A. Heiny, H. Rajakaruna, D. Talmy, and M.J. Follows.: Trophic model closure influences ecosystem response to enrichment." Ecological Modelling 475, 110183, 2023.

Pata, P.R.,Hunt, B.P.: Harmonizing marine zooplankton trait data toward a mechanistic understanding of ecosystem functioning. Limnology and Oceanography, 2023.

Petrik, C.M., Jessica Y.L., R. F. Heneghan, J.D. Everett, C.S.Harrison, and A. J. Richardson.: Assessment and constraint of mesozooplankton in CMIP6 Earth system models, Global Biogeochemical Cycles 36, no. 11, e2022GB007367, 2022.

Planchat, A., Kwiatkowski, L., Bopp, L., Torres, O., Christian, J.R., Butenschön, M., Lovato, T., Séférian, R., Chamberlain, M.A., Aumont, O. and Watanabe, M.: The representation of alkalinity and the carbonate pump from CMIP5 to CMIP6 Earth system models and implications for the carbon cycle. Biogeosciences, 20(7), pp.1195-1257, 2023.

Prend,C.J., Keerthi, M.G., Levy, M., Aumont, O., Gille, S.T. & Talley, L.D. Sub-seasonal forcing drives year-to-year variations of Southern Ocean primary productivity, Global Biogeochemical Cycles, 36(7), e2022GB007329, 2022.

Racault, M. F. et al. Impact of El Niño variability on oceanic phytoplankton. Frontiers in Marine Science, 4, 133, 2017.

Resplandy, L., Vialard, J., Lévy, M., Aumont, O. & Dandonneau, Y. Seasonal and intraseasonal biogeochemical variability in the thermocline ridge of the southern tropical Indian Ocean. Journal of Geophysical Research: Oceans, 114(C7), 2009.

Rohr, T., Richardson, A.J., Lenton, A., Chamberlain, M.A. and Shadwick, E.H.: Zooplankton grazing is the largest source of uncertainty for marine carbon cycling in CMIP6 models. Communications Earth & Environment, 4(1), p.212, 2023.

Sailley, S. F., Vogt, M., Doney, S. C., Aita, M. N., Bopp, L., Buitenhuis, E. T., Hashioka, T., Lima, I., Le Quéré, C., & Yamanaka, Y.: Comparing food web structures and dynamics across a suite of global marine ecosystem models. Ecological Modelling, 261–262, 43–57, https://doi.org/10.1016/j.ecolm odel.2013.04.006, 2013.

Sarmiento, J. L., Slater, R., Barber, R., Bopp, L., Doney, S. C., Hirst, A. C., Kleypas, J., Matear, R., Mikolajewicz, U., Monfray, P., Soldatov, V., Spall, S. A., and Stouffer, R.: Response of ocean ecosystems to climate warming, Global Biogeochem. Cy., 18, GB3003, https://doi.org/10.1029/2003GB002134, 2004.

Sathyendranath, S., Ji, R. & Browman, H. I.: Revisiting Sverdrup's critical depth hypothesis. ICES Journal of Marine Science, 72(6), 1892-1896, 2015.

Schulzweida, Uwe. CDO User Guide (2.3.0). Zenodo (2022]3). https://doi.org/10.5281/zenodo.10020800

Schupfner, M., Wieners, K., Wachsmann, F., Steger, C., Bittner, M., Jungclaus, J. et al.: MPI-ESM1.2-HR model output prepared for CMIP6 ScenarioMIP ssp585. Version 2019, Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4403, 2019.

Séférian, R., CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical. Version 20180917 Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4068, 2018.

Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J. et al.:

715 Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6, Current Climate Change Reports, doi:10.1007/s40641-020-00160-0, 2020.

Seland, O., Bentsen, M., Oliviè, D.J.L., Toniazzo, T., Gjermundsen, A. et al.: NCC NorESM2-LM model output prepared for CMIP6 CMIP piControl. Version 2019.Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8217, 2019a.

Seland, O., Bentsen, M., Oliviè, D.J.L., Toniazzo, T., Gjermundsen, A. et al.: NCC NorESM2-LM model output prepared for

720 CMIP6 CMIP historical. Version 20191108, Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8036, 2019b.

Seland, O., Bentsen, M., Oliviè, D.J.L., Toniazzo, T., Gjermundsen, A. et al.: NCC NorESM2-LM model output prepared for CMIP6 ScenarioMIP ssp585. Version 2019. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.8319, 2019c.

725 Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Cocco, V., Doney, S. C., Gehlen, M., Lindsay, K., Moore, J. K., Schneider, B., and Segschneider, J.: Projected 21st century decrease in marine productivity: a multi-model analysis, Biogeosciences,7, 979–1005, https://doi.org/10.5194/bg-7-979-2010, 2010.

Strutton, P. G., Ryan, J. P., & Chavez, F. P. Enhanced chlorophyll associated with tropical instability waves in the equatorial Pacific. Geophysical Research Letters, 28(10), 2005-2008, 2001.

730 Tagliabue, A., Kwiatkowski, L., Bopp, L., Butenschön, M., Cheung, W., Lengaigne, M. and Vialard, J.: Persistent uncertainties in ocean net primary production climate change projections at regional scales raise challenges for assessing impacts on ecosystem services. Frontiers in Climate, 3, p.738224, 2021.

Talmy,D., E. Carr., H. Rajakaruna., S. Våge., and A.W. Omta.: Killing the predator: impacts of highest-predator mortality on the global-ocean ecosystem structure." Biogeosciences 21, no. 10, 2493-2507, 2024.

735 Thomalla, S. J., Fauchereau, N., Swart, S. & Monteiro, P. M. S. Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean. Biogeosciences, 8(10), 2849-2866, 2011.

Thomalla, S.J., Nicholson, S.A., Ryan-Keogh, T.J. and Smith, M.E.: Widespread changes in Southern Ocean phytoplankton blooms linked to climate drivers. Nature Climate Change, 13(9), pp.975-984, 2023.

Tilman, D.: Resource competition between plankton algae: An experimental and theoretical approach. Ecology 58 (2), 338–

740 348, 1977.

Vaittinada Ayar, P., Battisti, D.S., Li, C., King, M., Vrac, M. and Tjiputra, J.: A regime view of ENSO flavors through clustering in CMIP6 models. Earth's Future, 11(11), p.e2022EF003460, 2023.

Vantrepotte, V. & Mélin, F. Temporal variability of 10-year global SeaWiFS time-series of phytoplankton chlorophyll a concentration. ICES J. Mar. Sci. 66, 1547–1556, 2009.

745 Vantrepotte, V. & Mélin, F. Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). Deep Sea Res. Part I 58, 429–441, 2011.

Walsh, K. J. et al. Tropical cyclones and climate change. Wiley Interdisciplinary Reviews: Climate Change, 7(1), 65-89, 2016.

Weaver, A. & Courtier, P. Correlation modelling on the sphere using a generalized diffusion equation. Quarterly Journal of the Royal Meteorological Society, 127(575), 1815-1846, 2001.

750 Wieners, K., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M. et al.: MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP piControl. Version 2019.Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6675, 2019a.

Wieners, K., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M. et al.: MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical. Version 2019.Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6595, 2019b.

Wieners, K., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M. et al.: MPI-M MPI-ESM1.2-LR model output prepared for

755 CMIP6 ScenarioMIP ssp585. Version 2019,Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6705, 2019c.

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

Xu, T., Li, S., Hamzah, F., Setiawan, A., Susanto, R. D., Cao, G. & Wei, Z. Intraseasonal flow and its impact on the

760 chlorophyll-a concentration in the Sunda Strait and its vicinity. Deep-Sea Research I, 136, 84-90, 2018.

765

770