Geoscientific
Model Development

The utility of simulated ocean chlorophyll observations: a case study with the Chlorophyll Observation Simulator Package (version 1) in CESMv2.2

Genevieve L. Clow[1,2,3], **Nicole S. Lovenduski**[1,2], **Michael N. Levy**[4], **Keith Lindsay**[4], **and Jennifer E. Kay**[1,3]

[1]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Boulder, Colorado, USA
[2]Institute of Arctic and Alpine Research, University of Colorado, Boulder, Boulder, Colorado, USA
[3]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Boulder, Colorado, USA
[4]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, Colorado, USA

Correspondence: Genevieve L. Clow (genevieve.clow@colorado.edu)

Abstract. For several decades, a suite of satellite sensors has enabled us to study the global spatiotemporal distribution of phytoplankton through remote sensing of chlorophyll. However, the satellite record has extensive missing data, partially due to cloud cover; regions characterized by the highest phytoplankton abundance are also some of the cloudiest. To quantify potential sampling biases due to missing data, we developed a satellite simulator for ocean chlorophyll in the Community Earth System Model (CESM) that mimics what a satellite would detect if it were present in the model-generated world. Our Chlorophyll Observation Simulator Package (ChlOSP) generates synthetic chlorophyll observations at model runtime. ChlOSP accounts for missing data – due to low light, sea ice, and cloud cover – and it can implement swath sampling. Here, we introduce this new tool and present a preliminary study focusing on long timescales. Results from a 50-year pre-industrial control simulation of CESM–ChlOSP suggest that missing data impact the apparent mean state and variability of chlorophyll. The simulated observations exhibit a nearly −20 % difference in global mean chlorophyll compared with the standard model output, which is the same order of magnitude as the projected change in chlorophyll by the end of the century. Additionally, missing data impact the apparent seasonal cycle of chlorophyll in subpolar regions. We highlight four potential future applications of ChlOSP: (1) refined model tuning; (2) evaluating chlorophyll-based net primary productivity (NPP) algorithms; (3) revised time to emergence of anthropogenic chlorophyll trends; and (4) a test bed for the assessment of gap-filling approaches for missing satellite chlorophyll data.

1 Introduction

The spatiotemporal distribution of marine phytoplankton, unicellular algae responsible for ∼ 50 % of global net primary production, greatly impacts fisheries, ecosystems, and the marine carbon cycle (Chassot et al., 2010; Fay and McKinley, 2017). Phytoplankton growth is dependent on temperature, light, and nutrient availability. Regions characterized by upwelling of nutrient-rich water, such as the equatorial, subpolar, and eastern boundary current regions, are some the most biologically productive (Siegel et al., 2013). In subpolar regions, wintertime mixing brings nutrients to the surface, but the lack of sunlight prohibits growth until the spring. This results in a pronounced seasonal cycle in phytoplankton abundance at high latitudes. In contrast, subtropical regions have abundant light but lack nutrients due to density stratification, which reduces vertical mixing. This nutrient limitation is relieved at the Equator due to ocean dynamics, leading to elevated productivity throughout the year.

For over 20 years, a suite of satellite sensors has enabled us to study the global spatiotemporal distribution of phytoplankton through remote sensing of chlorophyll *a* (McClain, 2009; Siegel et al., 2013). Chlorophyll *a* (hereafter referred to as chlorophyll) is the primary photosynthetic molecule in

plant cells, and it affects ocean spectral properties in identifiable wavelengths that can be remotely observed by passive satellite spectroradiometers. Since remote sensing of chlorophyll relies on visible light, detection is not possible at night or beneath cloud cover and sea ice. The most commonly used algorithms to derive chlorophyll concentration from remote sensing reflectance rely primarily on the ratio of blue to green wavelengths. The Hu (Hu et al., 2012) and O'Reilly (O'Reilly et al., 1998; O'Reilly and Werdell, 2019) chlorophyll algorithms are based on empirical relationships between remote sensing reflectance and in situ measurements. These satellite measurements have provided a global dataset with which we can study phytoplankton abundance and variability.

Earth system models (ESMs) can generate projections of future phytoplankton abundance and productivity in a changing climate and are thus a key tool for quantifying the impacts of changing climate on the carbon cycle (Wilson et al., 2022) and fisheries productivity (Tittensor et al., 2018), as well as the avoided impacts under climate change mitigation (Krumhardt et al., 2017). ESMs produce century-scale projections of climate, using mathematical equations to describe atmospheric and oceanic processes, including a full terrestrial and ocean carbon cycle. ESMs simulate nutrient cycling in the ocean by accounting for the role of phytoplankton and their zooplankton predators. In simulating the abundance of phytoplankton, models include processes such as photosynthesis, respiration, grazing, and sinking. These biological terms depend on physical and chemical oceanography simulated by the model. ESM projections suggest that phytoplankton abundance is affected by anthropogenic climate change due to changes in stratification and the consequent nutrient and light availability brought on by upper-ocean warming (Kwiatkowski et al., 2020). As the climate changes, oligotrophic regions are expected to see a decline in phytoplankton abundance, while regions with light-limited production are likely to see increases in abundance (Kwiatkowski et al., 2020; Marinov et al., 2010). These regional changes in primary productivity have critical implications for the coupled carbon–climate system, as well as for marine ecosystems and fisheries.

In order to use ESMs to project the future, we first need to validate them using present-day observations. During the model development phase, the biogeochemical components of ESMs are often tuned to the satellite record of chlorophyll. Based on simulated phytoplankton biomass, ESMs calculate the chlorophyll concentration at each time step and grid cell. The process of tuning involves making minor adjustments to various parameters so that the model outputs more closely align with observations. Some examples of ocean biogeochemical parameters are nutrient uptake rates, maximum grazing rates, and growth efficiency coefficients for each phytoplankton functional group (Long et al., 2021). These parameter values are based on information provided from laboratory studies, but the exact values are not known.
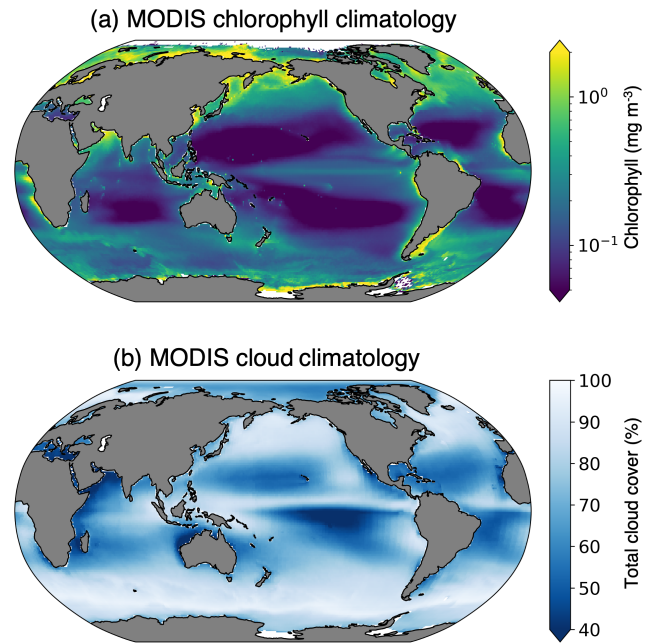

Figure 1. Multi-year means, with the **(a)** Aqua MODIS chlorophyll concentration entire mission composite image (2002–2023) (NASA Ocean Biology Processing Group, 2022) and **(b)** Aqua/Terra MODIS total cloud cover (2000–2011) mean (NASA Goddard Space Flight Center, 2019).

A typical tuning target involves minimizing error in the broad global patterns of the modeled climatology (Danabasoglu et al., 2020).

Due to the availability of surface chlorophyll data from satellite measurements, this dataset is a convenient tuning target. However, ESM-produced chlorophyll is not identical to that estimated via satellite. Satellite observations represent a vertically, optically weighted chlorophyll signal, which is generally limited to the near-surface ocean due to light attenuation at depth. Therefore, the comparison with the vertically resolved model output is limited to the surface layer. Additionally, ESMs provide a complete record of chlorophyll across the global ocean, whereas there are significant data gaps in the satellite record. Gregg and Casey (2007) estimated that satellite sampling bias leads to an 8 % overestimate of global mean chlorophyll. Additionally, ESMs calculate chlorophyll directly, while satellite estimates of chlorophyll are derived from optical properties of ocean water, leading to further uncertainty. As such, we may be tuning our ESMs to biased observations, which could inflate inaccuracies in ESM projections. Previous validation efforts have tried to make models more satellite-like by generating remote sensing reflectances within the model (Dutkiewicz et al., 2018). However, one of the most important causes of potential bias and model mismatch has not yet been addressed: the role of missing data.

Globally, the largest impediment to satellite chlorophyll detection arises from solar zenith angle limits (Gregg and Casey, 2007), which prevent detection at nighttime. This is especially important in the high-latitude regions during wintertime, when sea ice (a further detection challenge) is present. Another significant barrier to chlorophyll retrieval via satellite is cloud cover (Gregg and Casey, 2007; Mikelsons and Wang, 2019). On an average day, Moderate Resolution Imaging Spectroradiometer (MODIS) sensors aboard the Terra and Aqua satellites are unable to detect chlorophyll in approximately 72 % of the ocean's surface due to clouds (King et al., 2013). Some of the cloudiest ocean regions, such as the subpolar North Atlantic, North Pacific, and Southern Ocean, also have some of the highest rates of primary productivity (Fig. 1). The co-location of high chlorophyll and cloud coverage results from atmospheric and oceanic dynamics; global wind patterns control the climatological distribution of clouds and ocean upwelling. Ocean upwelling tends to cool the overlying atmosphere, which raises the humidity and leads to enhanced cloud cover. Upwelling also leads to increased nutrient concentrations, allowing more phytoplankton growth. Therefore, we are unable to reliably detect phytoplankton in the regions where they are most abundant.

To help bridge the gap between modeled and observed chlorophyll, we developed a satellite observation system simulator for ocean chlorophyll in the Community Earth System Model (CESM): the Chlorophyll Observation Simulator Package, ChlOSP (Fig. 2). Using ChlOSP, CESM simultaneously generates an estimate of standard modeled (full-field) chlorophyll and synthetic observations (obscured by simulated solar zenith angle, sea ice, and clouds). This enables us to (1) estimate sampling biases due to cloud cover and other sources of missing data and (2) make a more direct comparison between the model outputs and real-world data to improve model tuning and validation. Here, we present initial results from a 50-year pre-industrial simulation using ChlOSP and briefly explore future applications of this new tool. As we will show, clouds can alter the apparent mean state, seasonality, and variability in the chlorophyll. In addition to improving model tuning exercises, applications of ChlOSP include estimating the time of emergence of anthropogenic trends in the chlorophyll record, evaluating methods for calculating net primary productivity from satellite-observed chlorophyll, and creating a self-consistent gap-filling test bed.

2   Methods

2.1   Community Earth System Model version 2

The Community Earth System Model version 2 (hereafter referred to as CESM) is a fully coupled global climate model developed at the National Center for Atmospheric Research (Danabasoglu et al., 2020). The model includes components for the ocean (POP2), atmosphere (CAM6), sea ice (CICE5), land (CLM5), land ice (CISM2), waves (WW3), and rivers (MOSART), which exchange information through the coupler (CPL7). The carbon cycle is represented through land and ocean biogeochemistry subcomponents, which exchange carbon fluxes through the atmosphere. Here, we use version 2.2 of CESM, which was tuned via parameter adjustment and expert evaluation to correct ocean biogeochemical biases (Yeager et al., 2022).

2.1.1   Ocean model

The ocean component in CESM is the Parallel Ocean Program version 2 (POP2) (Danabasoglu et al., 2020; Smith et al., 2010). The coupler passes states and fluxes between CAM6 and POP2 at 30 min and hourly intervals, respectively. The standard grid for POP2 has an approximately 1° horizontal resolution, with 60 vertical levels ranging in thickness from 10 m at the surface to 250 m at depth. The model includes parameterizations of subgrid-scale processes, which are important for modeling ocean biogeochemistry. For example, biogeochemical tracers are impacted by parameterizations for eddy diffusivity, along with estuary, wave-driven, and vertical mixing (Danabasoglu et al., 2020). CESM also includes subgrid-scale light availability, which impacts photosynthesis rates and improves the representation of phytoplankton in regions with sea ice (Long et al., 2015).

Biogeochemical ocean processes are modeled in CESM by the Marine Biogeochemistry Library (MARBL) (Long et al., 2021). We use a configuration of MARBL that simulates three phytoplankton functional groups: diatoms, diazotrophs, and small (pico and nano) phytoplankton. The growth term depends on light, nutrients, and temperature. The loss terms include sinking and grazing, which is controlled by one zooplankton functional group. For each phytoplankton type, nutrient limitation is calculated based on phosphorus, nitrogen, iron, and silicon tracers. Nutrient concentrations evolve through biological processes, as well as nutrient fluxes from dust deposition and river inputs. The light limitation term is a function of photosynthetically active radiation (PAR), which is calculated as 45 % of incoming shortwave radiation at the surface. PAR varies with time of day, depth in the water column, cloud coverage, and sea ice. MARBL includes a dynamic chlorophyll-to-carbon ratio (θ) within each phytoplankton group. The optimal θ depends on temperature, light, and nutrient availability, allowing phytoplankton to adapt to changing environments through the process of photo-acclimation (Geider et al., 1998). Since satellite observations of chlorophyll are used as a proxy for phytoplankton biomass, including the photo-acclimation term is key for estimating biases in the satellite record.
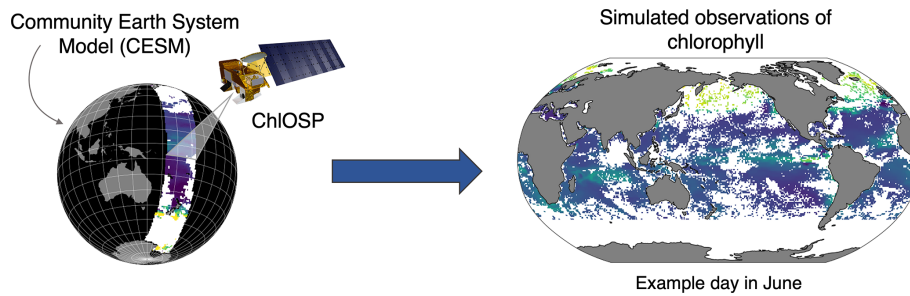
Figure 2. Conceptual diagram of ChlOSP.

2.1.2 Atmosphere model

The atmosphere model used in CESM is the Community Atmosphere Model version 6 (CAM6) (Danabasoglu et al., 2020). The default configuration for CAM6 is a finite-volume dynamical core with a horizontal resolution of 1.25° in longitude and 0.9° in latitude. It has 32 vertical levels up to about 40 km in height. Clouds are simulated through parameterizations of the planetary boundary layer and shallow convection following the Cloud Layers Unified By Binormals method (Golaz et al., 2002; Bogenschutz et al., 2013). This allows for subgrid-scale variations in temperature, humidity, and vertical velocity, leading to partial cloud cover within a grid cell. The cloud microphysics scheme is based on Gettelman and Morrison (2015), with ice nucleation depending on both temperature and aerosol concentration (Wang et al., 2014).

This work builds on existing satellite simulator software designed for clouds – the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) (Bodas-Salcedo et al., 2011; Webb et al., 2017). Within CAM6, COSP provides model outputs that are directly comparable to real-world satellite observations (Pincus et al., 2012). This software package has been incorporated into many climate models (Klein et al., 2013), including CESM (Kay et al., 2012). The latest version of COSP, COSP2, is functional in CESM version 2 (Swales et al., 2018). COSP simulates the observations of several satellite sensors, including the Multi-angle Imaging SpectroRadiometer (MISR), Moderate Resolution Imaging Spectroradiometer (MODIS), CloudSat, Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO), and the International Satellite Cloud Climatology Project (ISCCP). In COSP, the atmospheric model grid cells are divided into internally homogeneous subcolumns that roughly correspond to the spatial resolution of satellite data, and then forward modeling is applied to each subcolumn to generate satellite-like measurements. The results from each subcolumn are then aggregated back to the model resolution. We developed ChlOSP using both the ISCCP and MODIS cloud simulators. Cloud properties are calculated at hourly radiation time step intervals using 250 subcolumns.

2.2 ChlOSP description

The goal of ChlOSP is to generate model output that is comparable to the NASA Ocean Color Level 3 chlorophyll concentration data product. Level 3 data are an imperfect estimate of the actual surface chlorophyll concentration due to atmospheric correction, instrumentation, and ocean color algorithm uncertainties. Here, we focus on sampling biases that arise due to missing data by assuming that the satellite can detect the true ocean surface chlorophyll with 100 % accuracy in clear-sky conditions. We discuss the implications of this assumption in the Discussion and Conclusions section below.

Simulated observations of surface chlorophyll are generated within the ocean model component, POP2. We assume that the satellite can only see the surface layer of the POP2 grid, which represents depths from 0 to 10 m. Although the depth seen by a satellite depends on the optical constituents present in the water column, the surface layer of the model roughly aligns with depths of in situ measurements used to validate the SeaWiFS (Sea-viewing Wide-Field-of-view Sensor) chlorophyll retrievals (Gregg and Casey, 2004). The chlorophyll concentration in each surface model grid cell is calculated as the sum of chlorophyll from each phytoplankton functional group represented in MARBL. At each model time step, POP2 uses multiple variables to calculate the chlorophyll weights, which represent the fraction of each model grid cell that would be viewable by a satellite (i.e., the fraction of the surface area that is not obscured). The weighted chlorophyll field can then be compared with the total surface chlorophyll to assess biases due to missing data.

2.2.1 Calculation of the weights

ChlOSP accounts for clouds, sea ice, and low sunlight (high solar zenith angle), all of which prevent satellite detection using passive instruments. In the default CESM configuration, the sea ice fraction is calculated in CICE5 and then passed to POP2. For clouds and solar zenith angle, the CESM coupler is modified to pass these additional variables from CAM6 into POP2. Specifically, we use the COSP-generated MODIS and ISCCP total cloud cover and cosine of the solar

zenith angle. Since the quality of chlorophyll retrievals starts to decline at a solar zenith angle of about 70° (Mikelsons and Wang, 2019), we apply a masking threshold of 0.342 for the cosine of the solar zenith angle.

Satellite-derived level-3 ocean color products have a high spatial resolution (4.6 km for MODIS) compared to the coarse model grid (1°). To account for the discrepancy in resolution, we apply a weighting method for sea ice and cloud cover. The weights range from 0 to 1, where 1 indicates that 100 % of the cell was viewable by the satellite, and 0 indicates that no detection was possible. To calculate the weight from modeled sea ice and cloud cover fields, which are both expressed in terms of the fraction of a grid cell that is covered, these values are subtracted from 1. All weights are assumed to be independent from one another, so the final weight is the product of the weights calculated from each input parameter. At every model time step, the surface chlorophyll is multiplied by the weights. Then, the weighted chlorophyll and the weights are both output by the model at the frequency specified by the user when running CESM (hourly, monthly, etc.). Both outputs are needed to calculate the weighted mean over space and/or time. The weighted chlorophyll is not a physical value that should be analyzed independently from the weights. When calculating the weighted mean of chlorophyll, the weighted chlorophyll output corresponds to the numerator, and the weights output corresponds to the denominator in Eq. (1):

$$\text{weighted mean} = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}. \tag{1}$$

2.2.2 Simulator configurations

In order to test the sensitivity of ChlOSP to the modeled representation of cloudiness, we generate outputs using two different simulated cloud observations. We also test the impact of sampling frequency by comparing results from sampling chlorophyll once per day vs. all-sunlit time steps. Here, we present the results from three configurations of ChlOSP: (1) all-daylight sampling with simulated ISCCP cloud observations, (2) all-daylight sampling with simulated MODIS cloud observations, and (3) 13:30 LT swath sampling with simulated MODIS cloud observations. The ISCCP configuration is comparable to a chlorophyll observing system that combines multiple satellites, while the 13:30 LT (local time) sampling of MODIS is more similar to observations from an individual satellite.

MODIS and ISCCP cloud cover are simulated observation fields generated in COSP. The simulated observations are generated using the same model information, but different cloud-detection algorithms result in different observed total cloud fraction (Bodas-Salcedo et al., 2011; Pincus et al., 2012). In the real world, the ISCCP cloud cover product combines data from multiple passive sensors, including geostationary weather satellites (Rossow and Schiffer, 1991), and the MODIS instrument is aboard two polar-orbiting satellites

(Aqua and Terra). COSP in CESM samples each location at all time steps and does not include varying orbits for the different satellite simulators. However, since both MODIS and ISCCP rely on visible wavelengths, only sunlit time steps are included (Kay et al., 2012). For the all-daylight ISCCP and MODIS configurations in ChlOSP, we similarly sample chlorophyll at all sunlit locations at each model time step.

The two MODIS configurations can be compared to assess the impact of satellite-like sampling vs. sampling at all daylight time steps. Since phytoplankton and chlorophyll concentrations exhibit a diurnal cycle, the time of detection may impact the results (Salisbury et al., 2021; O'Malley et al., 2014). We simulate a simplified version of NASA's Aqua orbit. Aqua is a polar-orbiting satellite that collects data at approximately 13:30 LT, with a swath width of 2300 km. On a given day, Aqua samples the poles several times but has data gaps at low latitudes because successive orbits are not aligned longitudinally. These low-latitude gaps are then filled during an orbit on the subsequent day, and the orbital pattern repeats every 16 d. To simplify the complex orbital geometry, the simulated Aqua satellite in ChlOSP flies exactly the same orbit every day. The swaths are vertically centered on 13:30 LT and have a swath width of 1668 km. Since successive orbits are aligned, there are no inter-orbit gaps near the Equator. While simplified, this method simulates the general sampling pattern of Aqua, which is approximately once per day at low latitudes, with increasing frequency at higher latitudes (Fig. S1).

2.3 Initial simulation and analysis

2.3.1 CESM simulation setup

We tested ChlOSP in a pre-industrial control simulation. In this type of simulation, forcing fields (greenhouse gases, aerosols, etc.) are fixed at values for the year 1850. Therefore, fluctuations in the system are a result of internal climate variability, rather than a response to external forcing. After initializing the model, equilibration of the deep ocean can take thousands of years. However, we are interested primarily in surface ocean variables, which reach equilibrium relatively quickly. In our model simulation, equilibrium of surface chlorophyll is reached after approximately 15 years (Fig. S2). We ran the model for 50 years but only analyzed the last 30 years of data.

2.3.2 Model outputs

Each configuration of ChlOSP generates a chlorophyll output, along with the corresponding weights (see Table A1 for a complete list of new model outputs). ChlOSP outputs were added to a new hourly POP2 output file stream. We use these hourly model outputs to assess the impact of photo-acclimation and the diurnal cycle of chlorophyll. We expected that the diurnal cycle may impact our MODIS

13:30 LT swath results, since this version samples each grid cell fewer times per day (once a day at low latitudes).

In post-processing, the weights are used along with their corresponding variable to calculate means over space and time (see Appendix B for equations). To investigate sampling biases in synthetic observations, we calculate the chlorophyll climatology using three ChlOSP outputs: "standard", "clear sky", and "cloudy". The standard climatology is the unweighted, standard model output (i.e., total surface chlorophyll). The cloudy output includes solar angle, sea ice, and cloud cover weights, and the clear-sky version includes only the daylight and sea ice weights, allowing us to isolate the impact of cloud cover. The cloudy and clear-sky outputs vary, depending on which configuration of ChlOSP is used (i.e., ISCCP, MODIS, and MODIS swath).

When calculating the global mean of chlorophyll, we weight each grid cell by how frequently it was sampled (Eq. B3). To do this, we calculate the time mean of the weights and then multiply this by the area of each grid cell, which effectively represents the sample size for each grid point. Figure S3 shows the chlorophyll climatologies along with the corresponding time mean of the weights for each cloudy configuration. The normalized weights represent the mean area seen by the satellite relative to other points on the globe.

2.3.3 Regions of interest

For our evaluation of ChlOSP and subsequent analysis, we focus on highly biologically productive and cloudy open-ocean regions, particularly those with large seasonal cycles. We use ocean biomes defined by Fay and McKinley (2014) (Fig. S4). These regions were defined using observations of chlorophyll, sea surface temperature, ice fraction, and mixed layer depth. The biomes that we highlight are the North Pacific (biome 2 is North Pacific subpolar seasonally stratified), North Atlantic (biome 9 is North Atlantic subpolar seasonally stratified), Arctic (biome 1 is North Pacific ice; biome 8 is North Atlantic ice), and Southern Ocean (biome 16 is Southern Ocean subpolar seasonally stratified). Note that the Arctic biome used in our analysis includes only regions that are seasonally ice-free and does not correspond to the entire Arctic ocean.

2.4 ChlOSP evaluation

Before using ChlOSP to quantify the impact of missing data on chlorophyll, we demonstrate that ChlOSP is able to realistically simulate satellite observations. Here, we focus on evaluating how well the simulator mimics real-world satellite data by calculating the percentage of missing chlorophyll data during the sunlit period of the day. This metric was selected for evaluation because it captures the efficacy of the simulator in an imperfect representation of the Earth system; CESM exhibits known biases in both chlorophyll (Long et al., 2021) and clouds (Danabasoglu et al., 2020).

Total cloud coverage greatly impacts the amount of missing chlorophyll data. Due to differing cloud detection algorithms, ISCCP and MODIS produce different estimates of total cloud cover (Fig. 3). The global mean cloud coverage from the ISCCP observations is approximately 13 % higher than the MODIS observations. The primary difference is in the treatment of partially cloudy pixels, which are treated as fully cloudy in the ISCCP simulator and fully clear in the MODIS simulator (Pincus et al., 2012). Since partial cloud cover within a pixel would also prevent accurate satellite measurements of chlorophyll, it is more appropriate here to use the higher estimate of total cloud cover. Therefore, our results focus on the ISCCP cloud configuration. The ISCCP-simulated chlorophyll sampling strategy would be most comparable to a global network of geostationary satellites with passive ocean color instruments. Since this does not exist in the real world, we use a merged chlorophyll product that combines several polar-orbiting sensors to increase daily data coverage: the Ocean Colour Climate Change Initiative (OC-CCI, https://www.oceancolour.org/, last access: 22 January 2024) dataset, version 6.0 (Sathyendranath et al., 2019). This product combines chlorophyll data from SeaWiFS, MERIS (Medium Resolution Imaging Spectrometer), MODIS (Moderate Resolution Imaging Spectroradiometer, aboard the Aqua satellite), and VIIRS (Visible Infrared Imaging Radiometer Suite). The data are available daily at 4 km spatial resolution.

The real world has more missing data than the simulated observations (Fig. 4). The OC-CCI chlorophyll product has a median global daily coverage of 21 %, whereas the median daily coverage for the ChlOSP ISCCP configuration is 40 %. The ISCCP configuration of ChlOSP samples more frequently than real-world sensors because it samples at every sunlit time step rather than once per day. Additionally, ChlOSP does not account for all conditions that prevent chlorophyll detection. Many of these factors are difficult to predict and model; for example, observations may be excluded from level-3 data due to atmospheric correction failure, saturated observed radiance, stray light contamination, algorithm failures, or satellite navigation failure (Scott and Werdell, 2019). However, there are several factors that are candidates for future versions of the simulator; some atmospheric and oceanic constituents that prevent chlorophyll retrieval, such as white caps, coccolithophores, and aerosols, are already simulated in some capacity in CESM. Instrument-related challenges, such as sun glint and high sensor zenith angle, would also be valuable additions to future versions of ChlOSP. Sun glint and inter-orbit gaps mainly impact low- to mid-latitude regions. However, Gregg and Casey (2007) found that chlorophyll sampling biases in these regions are small, so addressing these issues was not the primary goal of ChlOSP. Since we are not accounting for all of these factors currently, we expect that the total percent
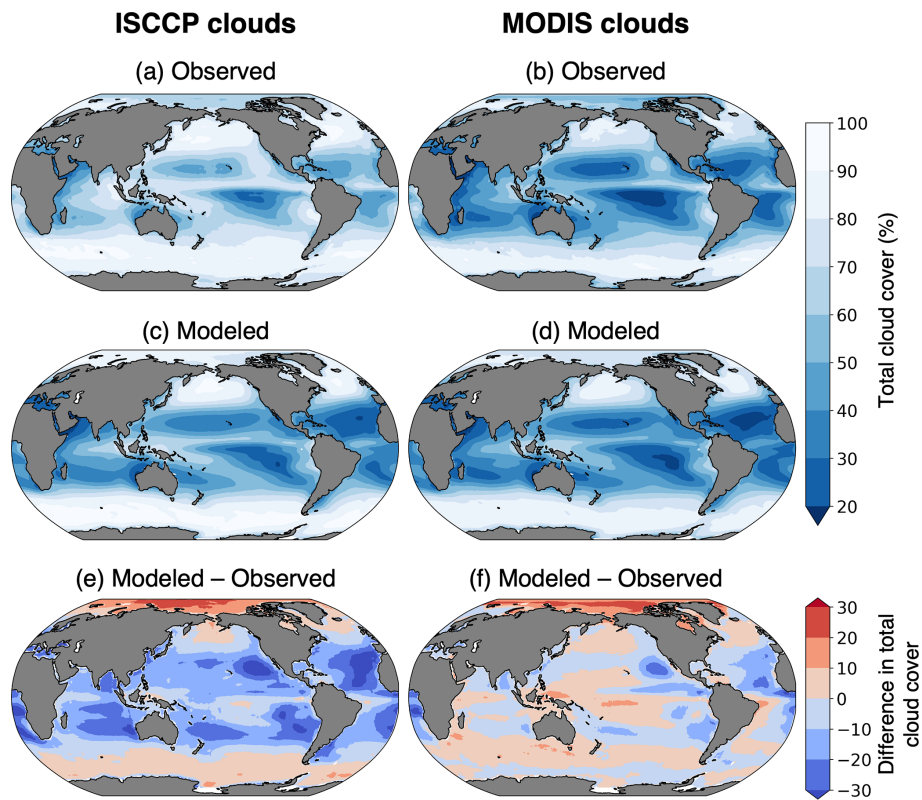
Figure 3. Cloud fraction climatology from CESM and observations. **(a)** ISCCP cloud cover observations from 2002 to 2016 (Rossow et al., 2017). **(b)** MODIS cloud cover (percent of pixels within 1° cell that had successful cloud property retrievals) from 2002 to 2016 (NASA Goddard Space Flight Center, 2019). **(c)** ISCCP and **(d)** MODIS modeled clouds from 30 years of the CESM pre-industrial simulation.

missing data will be lower in ChlOSP than in the real world. Therefore, our results represent a conservative estimate of biases due to missing data on a global scale.

In addition to the viewing conditions built into the simulator, differences in missing data arise due to the modeled representation of Earth. Comparing the temporal coverage across the globe provides more insight into the distribution of missing data (Fig. 4). ChlOSP captures missing data in the subpolar and polar regions well. However, ChlOSP collects more data in the tropics and subtropics compared to the real-world observations. This is due to cloud biases in the model, as shown in Fig. 3. We focus our study on the highly productive and very cloudy subpolar North Atlantic, Pacific, Arctic, and Southern Ocean regions, where the modeled cloud bias is minimal; as such, ChlOSP is an appropriate modeling tool for our purposes.

Since we are interested in the seasonal cycle of chlorophyll in the productive and cloudy subpolar regions, accurately representing the seasonality of missing data is also important. The timing of missing data relative to the seasonal cycle in chlorophyll effects how the climatology is weighted in time. Figure 5 shows the mean percent area missing from our four biomes of interest for each day of the year. To provide seasonal context for each biome, the mean seasonal cycle

of modeled chlorophyll is plotted as a heatmap above each panel, along with the mean sea ice fraction in the Arctic. For the model output, the weights were used to calculate the total area that was observed during the sunlit portion of the day. Overall, we find that the seasonality of missing data is appropriately captured in the model. Differences in the percent missing data between the model and real world may arise due to biases in modeled clouds and sea ice, as well as the mean state of the climate. For example, the modern-day satellite observations have lower sea ice coverage in the Arctic compared with the pre-industrial climate simulated in the model (Kay et al., 2022), leading to fewer instances of missing data in the real world.

The amount of missing data in ChlOSP depends on both the simulator setup (conditions used for masking) and the representation of physical variables in the model (concentration of clouds and sea ice). Despite some biases, our model evaluation demonstrates that, overall, ChlOSP realistically simulates the number of missing observations in a merged chlorophyll data product, particularly in regions with high biological productivity. Therefore, we can now use ChlOSP to investigate how these missing data impact our interpretation of chlorophyll climatology and seasonal cycles.
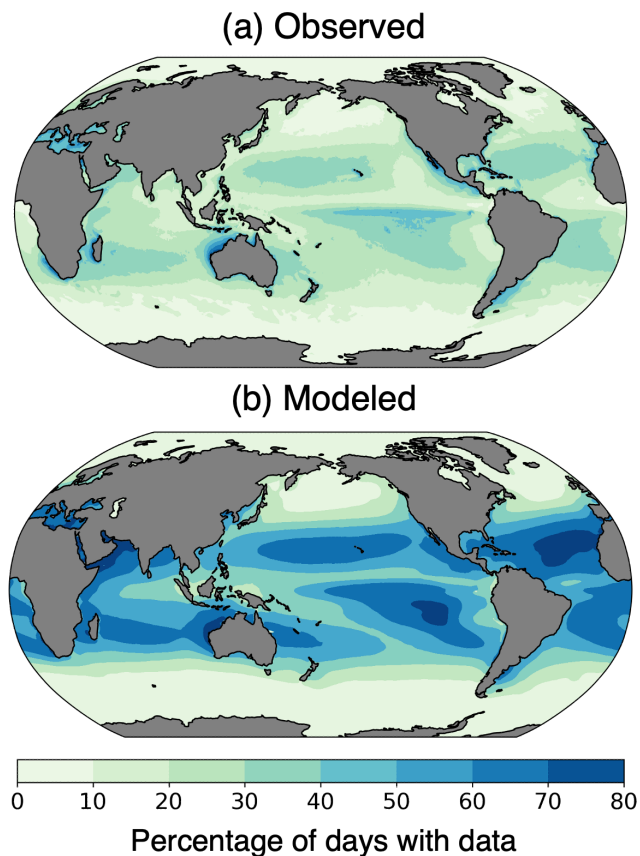
(a) Observed

(b) Modeled

Figure 4. Percentage of days with chlorophyll data coverage from **(a)** the Ocean Colour Climate Change Initiative (OC-CCI) (European Space Agency, 2022), 2006–2016, and **(b)** the pre-industrial simulation with the ISCCP configuration of ChlOSP.

3 Results

3.1 Climatology and global mean

A comparison of the standard, clear-sky, and cloudy (ISCCP) chlorophyll climatologies reveals that the temporal mean is impacted by missing data (Fig. 6). We highlight the differences between the various model outputs and configurations by calculating the percent differences in the climatology outputs (Fig. 7). To estimate the total sampling bias in simulated observations relative to the standard model output, we subtract the standard climatology from the cloudy climatology. We further estimate the contributions of sunlight and sea ice (clear sky − standard) and cloud cover (cloudy − clear sky) to the simulated observations of chlorophyll. The cloudy minus standard (cloudy − standard) maps indicate that sampling biases lead to > 100 % overestimates of chlorophyll in the high latitudes. This pattern emerges largely from the clear-sky minus standard (clear sky − standard) maps, indicating that daylight-only sampling has the largest influence. Figure 7 reveals that the different configurations of ChlOSP show similar overall spatial patterns.

The greatest differences in the clear-sky minus standard climatologies are located in the high latitudes, where there is insufficient light for satellite detection during the winter months. The winter months also correspond with low chlorophyll concentrations because the lack of sunlight limits phytoplankton growth. Additionally, in the polar regions, sea ice further prevents the satellite detection of chlorophyll during the start of the bloom season. As such, undersampling in the winter leads to an overestimate of the mean chlorophyll concentration. In the low latitudes, a small underestimation of chlorophyll arises due to the phasing of the diurnal cycle relative to sampling time (Fig. 8). In Fig. 8, we select grid cells near the Equator to illustrate how these sampling biases arise. This region exhibits the largest diurnal range, as shown in Fig. S5. Both the swath and daylight configurations of ChlOSP have a negative anomaly compared to the true mean (cf. dashed yellow, blue, and gray lines in Fig. 8).

A comparison of Fig. 7e and h reveals that implementing a satellite swath in ChlOSP impacts the clear-sky chlorophyll climatology. Overall, the biases are slightly less extreme when the swath is implemented; the differences decrease by about 9 % when averaged globally. In low latitudes, this is due to the sampling time. Figure 8 shows that, at the Equator, the swath sampling has a smaller bias compared to the daylight sampling. In subpolar regions, the overestimate is also smaller when the swath is implemented (cf. Fig. 7e and h). The swath version samples less frequently, relative to the daylight configuration during the summer months (cf. Fig. S1a and b), leading to weights that are more evenly distributed throughout the year. Therefore, the bias towards the summer chlorophyll peak is less extreme in the clear-sky swath compared with the all-daylight version.

The cloudy minus clear-sky (cloudy − clear-sky) column in Fig. 7 isolates the impact of cloud cover on the chlorophyll climatology from simulated observations. Clouds from the ISCCP simulator were used for Fig. 7c, and clouds from the MODIS simulator were used for both Fig. 7f and i. The impact of clouds is slightly sensitive to the sampling strategy (cf. Fig. 7f and i), but the total simulated cloud cover has a larger effect on chlorophyll climatology (cf. Fig. 7c and f). The difference in magnitude between Fig. 7c and f arises from the difference in total cloud coverage simulated by the two configurations (Fig. 3). Since the ISCCP simulator has more cloud coverage on average, there is a greater impact on the apparent chlorophyll climatology. The spatial pattern in all three panels can be largely explained by the correlation between the seasonal cycle of cloud cover and clear-sky chlorophyll (Fig. 9). The similarity in the spatial patterns seen between Figs. 9 and 7c suggests that correlations on monthly timescales play an important role in the resulting cloudy chlorophyll climatology. In regions where cloudy seasons correspond with lower chlorophyll (negative correlation), the chlorophyll climatology is overestimated relative to the clear-sky chlorophyll. Similarly, in regions where cloudy seasons correspond with higher chlorophyll (positive corre-
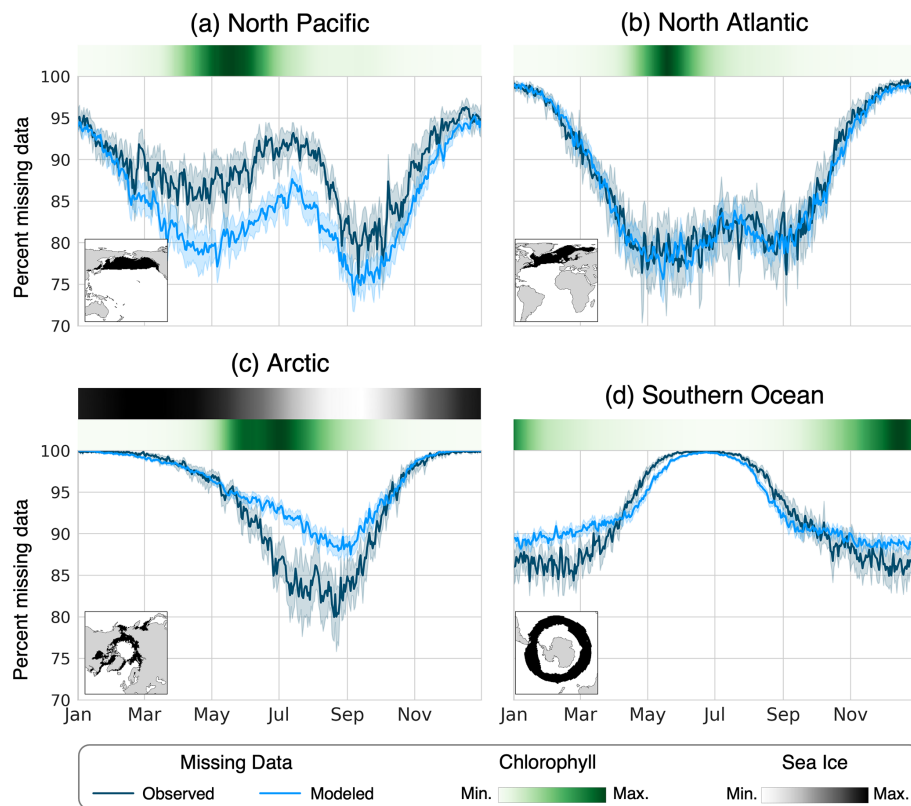
Figure 5. The mean percentage of area missing from the chlorophyll data for each day of the year in the **(a)** North Pacific, **(b)** North Atlantic, **(c)** Arctic, and **(d)** Southern Ocean. The observed data (dark blue line) were calculated using the OC-CCI data product from 2006 to 2016 (European Space Agency, 2022), and the modeled data (light blue line) were calculated from the 30 years of the pre-industrial simulation using the ISCCP configuration of ChlOSP. The error bars show the 95 % confidence interval on the daily mean. The green heatmaps show the seasonal cycle of chlorophyll derived from the standard model output, and the gray heatmap above the Arctic panel represents the sea ice seasonality.

lation), the chlorophyll climatology is underestimated. The impact of cloud cover is particularly important in the Arctic, where clouds offset some of the large overestimates of chlorophyll due to daylight-only sampling.

Global mean chlorophyll concentration estimated by the cloudy ChlOSP output is ∼ 20 % different from that estimated by the standard configuration (Table 1; differences for individual biomes are included in Table S1). From the maps in Fig. 7, the simulated observations appear to strongly overestimate global chlorophyll. However, this does not account for how often each location is sampled. The regions that show some of the highest sampling biases, such as subpolar biomes, are sampled very infrequently due to cloud cover or lack of sunlight (Fig. S3). If satellite sensors could see through clouds (as in the clear-sky configuration), then the global chlorophyll mean would be overestimated by 14 % to 22 %. This is because the clear-sky mean is heavily biased towards the productive summer months in high-latitude regions due to solar zenith angle limits in the wintertime. However, including cloud coverage leads to an underestimation of chlorophyll, which ranges from −7 % to −17 %.

This is because locations with high chlorophyll values tend to also be cloudy, and therefore, they are sampled less frequently than other regions. The sampling frequency of the swath vs. daylight-only configurations also plays an important role. The cloudy MODIS swath configuration has a lower global bias than the daylight-only version (−7 % vs. −14 %). While the MODIS swath configuration samples less frequently overall, it samples the high latitudes more frequently relative to other parts of the globe (Fig. S3). This is especially true in the summer months in high latitudes, since the orbit passes over the poles many times per day. Therefore, the biologically productive subpolar regions are weighted more strongly relative to other locations, leading to a smaller underestimation in global chlorophyll.

3.2 Seasonal cycles

In addition to impacting the chlorophyll climatology, satellite-like sampling also impacts the spatially averaged chlorophyll concentration. To investigate spatial means, we calculate the daily area-weighted chlorophyll mean within our biomes of interest using the standard, clear-sky, and

Table 1. Weighted global mean chlorophyll concentrations and weighted standard deviation for various configurations of ChlOSP, calculated using Eq. (B3). The weights are the product of the grid cell area and the ChlOSP weights (mean fraction of the grid cell that was "seen" by the satellite simulator). Note that chlorophyll is approximately log-normally distributed.

Model output	Global mean chlorophyll (mg m^{-3})	Global mean difference from standard	Standard deviation (mg m^{-3})
Standard	0.215	–	0.553
Clear sky	0.245	14.1 %	0.609
Clear-sky swath	0.263	22.2 %	0.642
Cloudy: ISCCP	0.179	−16.7 %	0.463
Cloudy: MODIS	0.184	−14.3 %	0.477
Cloudy: MODIS swath	0.200	−7.4 %	0.515

cloudy (ISCCP) outputs. The mean seasonal cycle was then calculated over the 30-year analysis period (Fig. 10).

The largest differences between the clear-sky and standard seasonal cycles occur during winter months when satellite sensors cannot detect higher-latitude grid cells due to low light. These regions also correspond to low wintertime chlorophyll values because there is limited light for photosynthesis. This difference between clear sky and standard is more pronounced in biomes that span a larger latitudinal range, such as the North Atlantic. In the Arctic, sea ice also plays a major role in the apparent seasonal cycle of the clear-sky chlorophyll. During most of the year, the standard chlorophyll is lower than the clear sky in this region. However, in July and August, the standard chlorophyll is higher than the clear sky, indicating that there are phytoplankton blooms beneath the sea ice that cannot be seen by ChlOSP.

Cloud cover also influences the apparent magnitude and timing of the seasonal phytoplankton bloom in the North Pacific. The differences between cloudy and clear-sky chlorophyll arise due to the spatial correlation between clouds and chlorophyll within each biome, which varies throughout the year. These correlations are shown in the boxes beneath each time series (Fig. 10; the corresponding correlations for all biomes can be found in Fig. S6). Many biomes exhibit a positive spatial correlation between cloud cover and chlorophyll concentration during the bloom months (e.g., Fig. 10a, b, and c). Within these biomes, model grid cells with high cloud cover tend to have higher chlorophyll, leading to lower biome mean cloudy chlorophyll concentrations relative to the clear-sky configuration (Fig. 10a, b, and c). Conversely, in the Southern Ocean, there is a negative spatial correlation between cloud cover and chlorophyll concentration through most of the year, leading to higher cloudy chlorophyll concentrations than the clear-sky configuration (Fig. 10d).

4 Applications

Our analysis so far has focused on using ChlOSP to assess how clouds, daylight vs. swath sampling, and the presence of sea ice may bias satellite observations of chlorophyll. How-

ever, there are many other potential applications that we envision for this tool.

4.1 Model tuning

Previously, tuning of the biogeochemical component of CESM has been accomplished by comparing standard chlorophyll to real-world satellite observations. The goal is to replicate the spatial pattern of the global chlorophyll climatology. However, as we have demonstrated, satellite observations of chlorophyll are biased due to missing data, whereas the standard model output is not. Figure 11 compares the real-world, observed chlorophyll climatology from Aqua MODIS to the modeled 30-year pre-industrial climatology. Calculating the model bias using standard chlorophyll vs. cloudy outputs leads to different results (Fig. 11). Given that they are both impacted similarly by missing data, the cloudy model output is more suitable for comparing model output with the real-world observations. These results indicate that the actual bias of CESM in the subpolar regions may be greater than previously thought, demonstrating the importance of taking sampling bias into account during the tuning process.

4.2 Net primary productivity

Another metric used for model tuning is the rate of globally integrated net primary productivity (NPP). NPP, the rate at which dissolved inorganic carbon is converted into organic matter, is particularly relevant for quantifying the global carbon cycle. The true modeled NPP can be determined directly by calculating the sum of the total carbon fixation vertical integral for all phytoplankton groups. This value is then compared to estimates of real-world marine NPP. There are many methods for estimating real-world NPP, many of which rely on satellite-observed chlorophyll. For example, the Vertically Generalized Production Model (VGPM) uses chlorophyll, along with sea surface temperature (SST) and photosynthetically active radiation (PAR), which are also derived from satellite products (Behrenfeld and Falkowski, 1997).
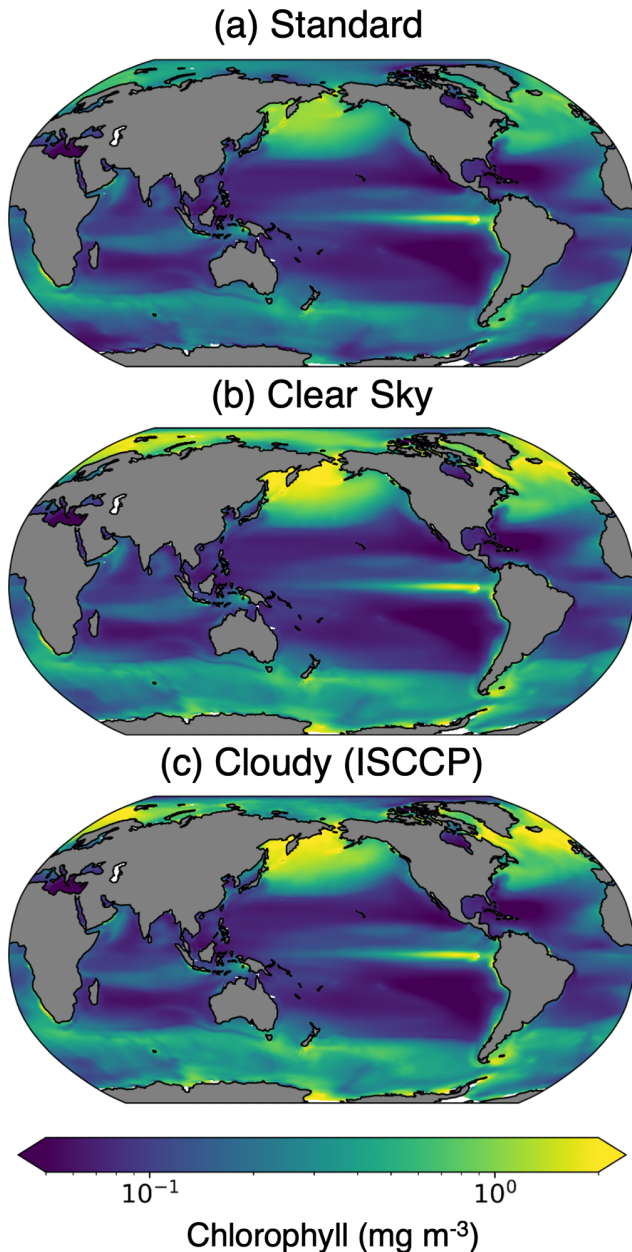
(a) Standard

(b) Clear Sky

(c) Cloudy (ISCCP)

Chlorophyll (mg m⁻³)

Figure 6. The 30-year chlorophyll climatology from the pre-industrial simulation calculated with (**a**) standard, (**b**) clear-sky, and (**c**) cloudy (ISCCP configuration) model outputs.

Table 2. Global net primary productivity calculated with VGPM model using ChlOSP (ISCCP) outputs.

Chlorophyll input used in VGPM	Global NPP ($\mathrm{Pg\,C\,yr}^{-1}$)
Standard	49.26
Clear sky	54.54
Cloudy	50.10

our NPP estimate to the full area of the ocean. The resulting NPP values are impacted by the version of chlorophyll from ChlOSP used as the input (Table 2). Since the cloudy output is most similar to real-world satellite data, the $50.10\,\mathrm{Pg\,C\,yr}^{-1}$ value should be used when tuning the model.

In addition to using simulated satellite-derived NPP to improve model tuning, we also demonstrate how ESMs can be used as a test bed for NPP algorithms. In our CESM simulation, the true globally integrated NPP is $48.43\,\mathrm{Pg\,C\,yr}^{-1}$. This value is remarkably similar to the VGPM-derived values, increasing our confidence in the accuracy of the real-world globally integrated NPP from VGPM.

4.3 Time of emergence

ChlOSP can also be used to calculate the time of emergence for chlorophyll trends in simulated observations. The time of emergence is the length of the observational record required to identify a statistically significant trend within the context of internal variability. The impact of anthropogenic climate change on phytoplankton abundance is critically important to marine ecosystems and fisheries around the world. However, there is great uncertainty in global chlorophyll trends in the current satellite record (Beaulieu et al., 2013; Boyce et al., 2014; Gregg and Rousseaux, 2014; Hammond et al., 2017; van Oostende et al., 2023), due in part to the limitations of satellite data, i.e., the shortness of the record and the prevalence of missing data. Additionally, there is further uncertainty in how surface chlorophyll trends translate to changes in total phytoplankton biomass (Siegel et al., 2013; Behrenfeld et al., 2016).

Using Earth system models, we can project global phytoplankton biomass into the future, using various forcing scenarios, where we know the true trend in surface chlorophyll and primary productivity. With ChlOSP enabled, we can also calculate the apparent trend from the simulated observations. Because we have a fully coupled model, we can account for any changes in cloud cover and sea ice due to warming. The time of emergence from the simulated observations gives us greater insight into when we might detect significant trends in the real world.

Since we have not yet generated a future projection with ChlOSP, here we compare the variability in monthly chlorophyll anomalies using the cloudy and standard outputs from our pre-industrial simulation. Throughout the majority of the
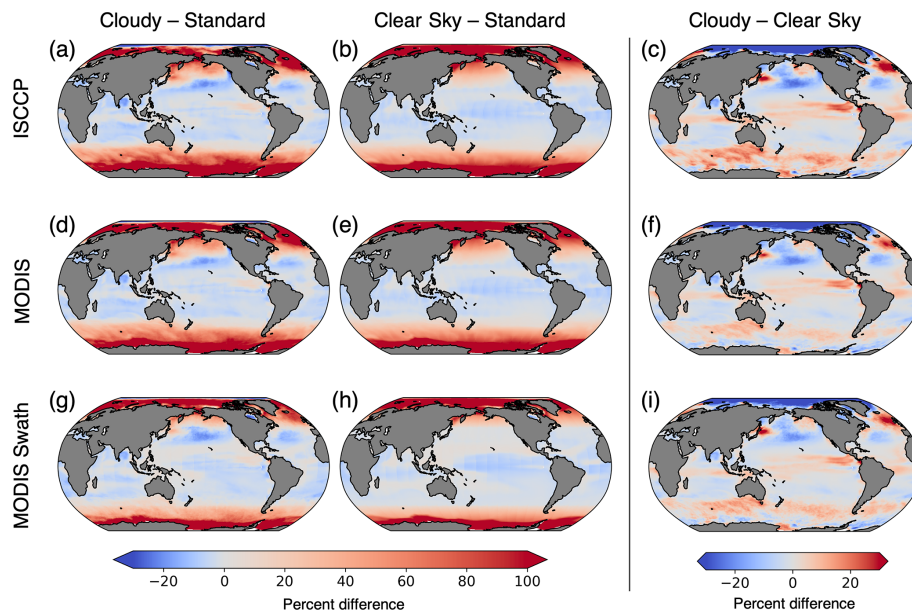
We can make a more direct comparison between the model and the real world by calculating ChlOSP-estimated NPP (using VGPM) rather than the true model output (i.e., the vertically integrated phytoplankton carbon fixation). To calculate satellite-like NPP from ChlOSP, we use the chlorophyll, SST, and PAR climatologies in the VGPM algorithm. When integrating the resulting NPP over the global oceans, we weight each grid cell area by the time-averaged satellite weights. We then calculate the total fraction of the ocean that was seen by the satellite and use this fraction to scale

Figure 7. Percent difference in chlorophyll climatologies calculated with the three configurations of ChlOSP. The rows show the outputs from ISCCP, MODIS, and MODIS swath, respectively. The first column is the difference between cloudy chlorophyll and standard. The second column is the difference between clear-sky chlorophyll and standard, which shows the impact of daylight-only sampling and sea ice on observations. The third column is the difference between cloudy chlorophyll and clear-sky chlorophyll, which isolates the impact of cloud cover on observations. Note that panels (**b**) and (**e**) are equivalent.

globe, the temporal variability is higher in the cloudy output relative to the standard output (Fig. 12). This is particularly apparent in high-latitude regions, where the low variance in wintertime is not seen by satellites. Since the cloudy dataset has more noise, a longer time series is necessary to identify a statistically significant trend. To estimate the length of the time series needed, we assumed a global trend in the surface chlorophyll of $-5 \times 10^{-4}\,\mathrm{mg\,m^{-3}\,yr^{-1}}$ and applied the method described in Weatherhead et al. (1998). Our calculations (not shown) indicate that the time to emergence may be delayed by more than 10 years in the subpolar regions due to missing data. This preliminary analysis indicates that typical chlorophyll model outputs may underestimate the time of emergence because they do not account for the enhanced variability in real-world observations. Therefore, ChlOSP will be a valuable addition to future simulations by providing more realistic estimates of the time of emergence.

4.4 Gap-filling

Satellite-derived chlorophyll data are often gap-filled to generate a more complete dataset at high spatial and temporal resolution. A wide variety of methods have been applied to this problem, from simple linear interpolation to more complex methods such as EOFs (empirical orthogonal functions) (Liu and Wang, 2018), neural networks (Krasnopolsky et al., 2016), and self-organizing maps (Jouini et al., 2013). Stock et al. (2020) compared many of these methods within four study areas and found that ordinary kriging, spatiotempo-

ral kriging, DINEOF (data interpolating empirical orthogonal functions), and random forests were the most successful methods, although results varied by region. Validating gap-filling methods is a challenge in the real world because we do not know what the truth is in places where we have missing data. This is often solved by transplanting artificial cloud masks onto clear-sky images or using data at a later time step. As we have shown, clouds and chlorophyll exhibit correlations, and there is a diurnal cycle in surface chlorophyll. Therefore, these methods introduce additional errors, making it difficult to quantify the error from the gap-filling method alone.

Here, we propose using an Earth system model test bed to apply various gap-filling techniques to the simulated observations from ChlOSP (Fig. 13). In the model world, we know the true chlorophyll values at every location, thus improving our ability to validate the gap-filled results. To generate gaps in the ChlOSP output, we use the weights as the probability of a grid cell being masked out. As an example, we have done a linear interpolation in Fig. 13. While the overall spatial pattern of chlorophyll is a close match between the gap-filled estimate and the model truth, smaller-scale features are not captured well by the linear interpolation, particularly in regions with large amounts of missing data.

Future work will involve testing a wide variety of methods, with the goal of identifying the best method for gap-filling chlorophyll on a global scale. In the model, we have full knowledge of variables that impact phytoplankton growth –
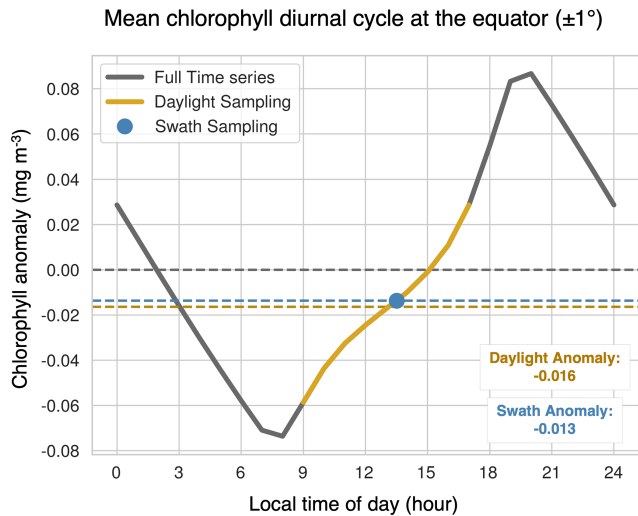
Mean chlorophyll diurnal cycle at the equator (±1°)

Figure 8. The mean diurnal cycle of chlorophyll at the Equator over all months of the year, represented as chlorophyll anomalies. The gray line represents the entire cycle, the yellow line indicates the sunlit period where the solar zenith angle is less than 70°, and the blue dot is the swath sampling time. The horizontal dashed yellow and blue lines correspond to the mean anomaly observed with daylight sampling and swath sampling, respectively, and the dashed gray line highlights the mean of the full period. To analyze the mean diurnal cycle, chlorophyll concentrations from the standard output were grouped by local time of day (binned hourly) and then averaged over all months and years. The time stamp of each data point is the end time of the averaging interval.

Cloud cover and chlorophyll seasonal cycle correlation
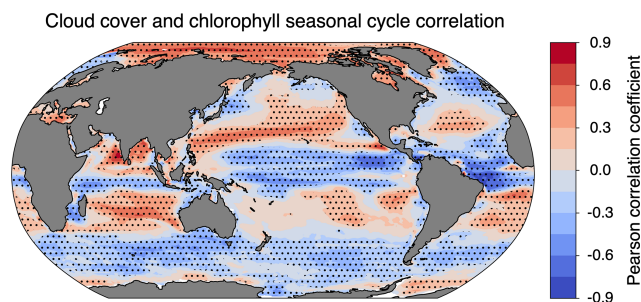
Figure 9. Pearson coefficient of correlation between monthly means of ISCCP cloud cover and clear-sky chlorophyll. Dotted cells indicate that the correlation was significantly different from zero at the 95 % confidence level. The effective sample size was determined by calculating the lag 1 autocorrelation for clouds and chlorophyll in each grid cell.

such as temperature, salinity, and wind – and can use this additional information in random forests or neural networks to predict chlorophyll. Many of these variables can be detected beneath cloud cover using microwave remote sensing (Gentemann et al., 2010), so these methods could be applied in the real world. One disadvantage of using the model as a test bed is that the resolution is much coarser than real-world satellite data, so it would likely not be suitable for gap-filling small-scale features. However, this method allows for the quantification of the error due to the gap-filling method.

5   Discussion and conclusions

We developed the Chlorophyll Observation Simulator Package (ChlOSP) for CESM, a fully coupled Earth system model. This new tool generates synthetic observations of surface ocean chlorophyll that are obscured by simulated cloud cover, sea ice, and high solar zenith angle. As a proof of concept, we ran ChlOSP in a 50-year pre-industrial simulation of CESM and analyzed the last 30 years. We tested three configurations of ChlOSP using different simulated cloud observations from COSP: ISCCP clouds, MODIS clouds, and MODIS clouds with an Aqua-like swath (13:30 LT sampling time). For each configuration, we compared the cloudy (obscured by sea ice, high zenith angle, and cloud cover), clear-sky (obscured by sea ice and high zenith angle only), and standard (not obscured) chlorophyll outputs to assess the sampling bias that arises due to missing data. We found that missing data impact the apparent climatology, overall global mean, and seasonal cycle in subpolar regions. We further demonstrated that ChlOSP can be used in future simulations to improve model tuning, calculate the time of emergence of a trend, and test gap-filling methods.

We found that the simulated observations from the ISCCP configuration underestimate the true global mean chlorophyll by $\sim 0.03\,\mathrm{mg\,m^{-3}}$, which is the same order of magnitude as the expected change by the end of the century ($\sim 0.05\,\mathrm{mg\,m^{-3}}$; Schlunegger et al., 2020). The largest differences between the simulated observations and the standard chlorophyll output are due to daylight-only sampling. These differences are particularly pronounced in the high-latitude regions, where the detection of low chlorophyll values during winter is not possible. This leads to an overestimation of mean chlorophyll over time, which is consistently over 100 % throughout the polar regions. In the real world, we have sparse in situ observations to compare with satellite data; therefore, our ability to estimate sampling bias is limited. However, results in Fig. 7 agree well with estimated real-world satellite sampling bias from Gregg and Casey (2007), where they applied satellite sampling to a global biogeochemical model with data-assimilated chlorophyll. Their results similarly showed that the largest sampling biases were due to the solar zenith angle threshold in high-latitude regions, with clouds being the second most important factor. They found a positive bias in the annual global mean ($+8\,\%$), which differs from the negative bias that we report ($-16.7\,\%$). However, this is because they report the standard mean rather than weighting by sampling frequency as we have done here.

Cloud cover plays an important role in the apparent mean of chlorophyll. Depending on the region, cloud cover can cause positive or negative sampling bias. In many ocean re-
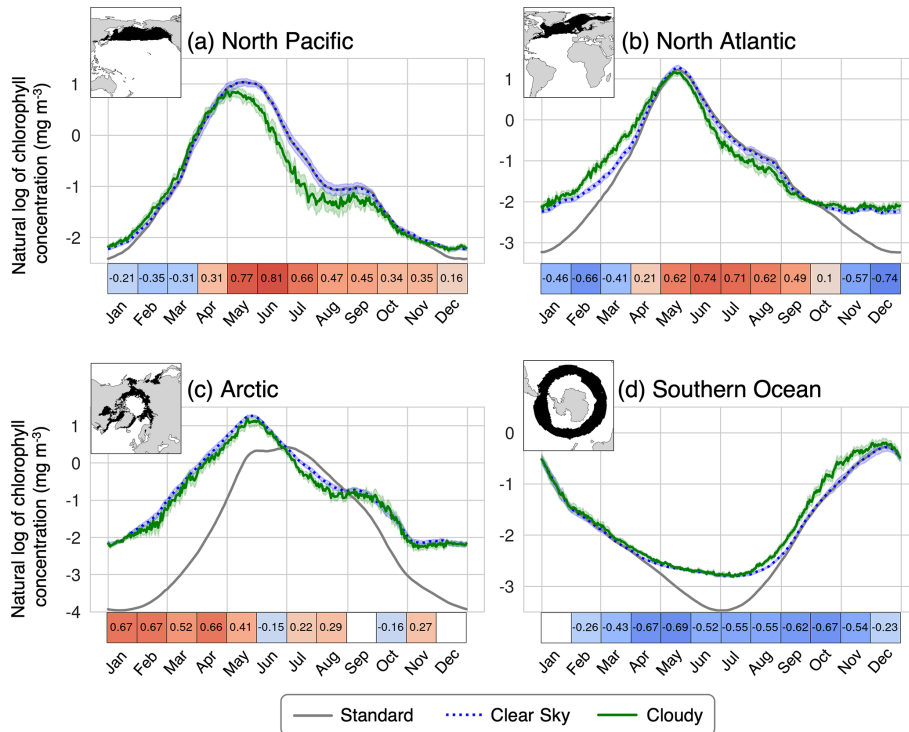
Figure 10. Seasonal cycle of standard, clear-sky, and cloudy chlorophyll (from the ISCCP configuration) in the **(a)** North Pacific (biome 2), **(b)** North Atlantic (biome 9), **(c)** Arctic (biomes 1 and 8), and **(d)** Southern Ocean (biome 16). The error bars on the clear-sky and cloudy lines indicate the 95 % confidence interval on the mean. The boxes below the time series indicate the spatial correlation (Pearson's coefficient) between the mean cloud coverage and clear-sky chlorophyll within the biome for each month. White cells indicate that the correlation is not significantly different from zero at the 95 % confidence level. The effective sample size was calculated using Moran's I spatial autocorrelation index.
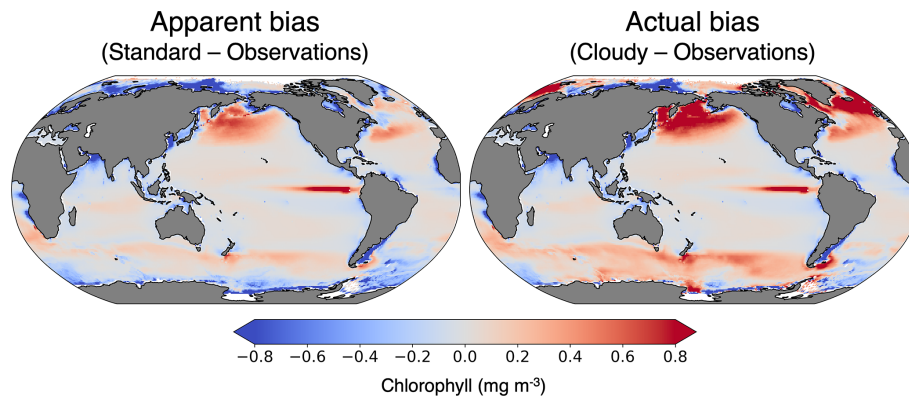
Figure 11. Model biases in chlorophyll climatology calculated with standard and cloudy (ISCCP) chlorophyll outputs. The observations are from Aqua MODIS from 2002 to 2023 (NASA Ocean Biology Processing Group, 2022) and are re-gridded to the model resolution.

gions, cloud cover and chlorophyll exhibit statistically significant correlations in both space and time (Figs. 9 and S6). Spatial correlations impact the mean chlorophyll over ocean biomes, while temporal correlations impact the climatology in a given location. The mechanisms driving these correlations are not explored here, but we expect that this is a result of large-scale dynamics rather than a direct interac-

tion between clouds and phytoplankton. In CESM, the PAR field at the ocean surface includes the influence of clouds, but the extent to which this impacts phytoplankton growth and/or photo-acclimation is not investigated here. There is some evidence that biogenic aerosols (dimethyl sulfide) produced by phytoplankton can increase cloud cover by acting as cloud condensation nuclei (Andreae and Crutzen, 1997),

Difference in variance of monthly anomalies
(Cloudy – Standard)
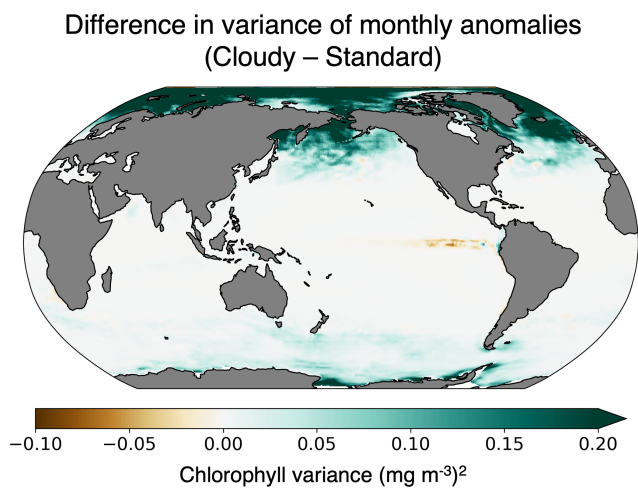
Figure 12. Difference in the variance of the monthly anomalies in surface chlorophyll concentration. The cloudy output is from the ISCCP ChlOSP configuration.
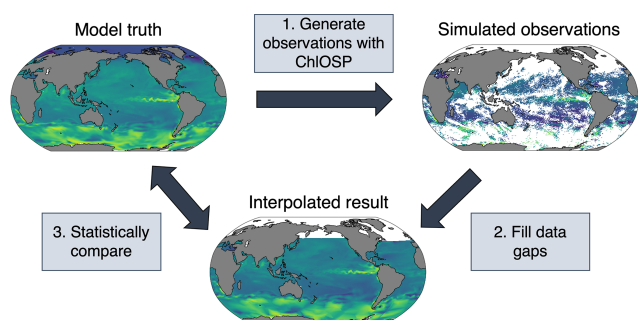

Figure 13. Diagram of the gap-filling test bed using ChlOSP-CESM.

but this process is not represented in CESM. Further work is needed to investigate the direct effects of cloud cover on surface chlorophyll concentration and to validate the seasonal phasing of clouds and chlorophyll in the real world.

Through testing various configurations of ChlOSP, we found that the results are sensitive to both the definition of cloud cover and the sampling pattern. The ISCCP configuration had higher cloud cover than MODIS, which amplified the differences between the standard and cloudy chlorophyll outputs (Fig. 7; Table 1). Since the ISCCP cloud simulator can detect partial cloud cover, it provides a more realistic representation of missing chlorophyll data. Therefore, we chose to focus the majority of our results on this configuration. Despite some differences in magnitude, we found that the overall patterns were consistent in all configurations.

Interestingly, when compared with the all-daylight versions, the swath configuration led to greater chlorophyll biases in the clear-sky global mean yet smaller biases in the cloudy global mean (Table 1). Figure S1 highlights that in the swath configuration, the subpolar regions (generally more productive) are sampled more frequently than lower-

latitude regions (generally less productive) due to the polar orbit. Therefore, the swath configuration further enhances the overestimation of the global mean that arises in clear-sky chlorophyll due to summertime-only sampling. However, when we add clouds, the swath sampling more accurately captures the global mean chlorophyll (Table 1). Figure S3 reveals that when clouds are included, the resulting weights are more evenly distributed throughout the globe because the productive subpolar regions tend to be cloudy. Additionally, we showed that at low latitudes, the 13:30 LT sampling time provides a more representative sample of the diurnal cycle than the all-daylight version (Fig. 8), which impacts the climatological mean of the clear-sky output (Fig. 7). This demonstrates the importance of simulating a realistic sampling pattern when assessing sampling biases in chlorophyll. Future work will involve implementing various swath widths, times, and orbital geometries to simulate a variety of sensors, including the upcoming NASA PACE (Plankton, Aerosol, Cloud, ocean Ecosystem) mission.

Our work focuses on sampling biases that arise due to missing data, but there are many other differences between modeled and observed chlorophyll. As shown in Dutkiewicz et al. (2018), large errors in observed chlorophyll – comparable in magnitude to what we found here – arise from the choice of algorithm used for estimating chlorophyll from remote sensing reflectance. We are unable to estimate these errors in CESM, as it currently lacks an optical model. While it is useful to isolate the biases that arise due to certain factors, it would also be beneficial to understand the cumulative effect. We hope that ChlOSP will be implemented in an Earth system model capable of combining these various components.

While ChlOSP provides an improved model output for real-world comparison, it is not a perfect representation of satellite data. We have demonstrated that ChlOSP reasonably represents the amount of missing chlorophyll data (Figs. 4 and 5). However, ChlOSP does not include all factors that prevent satellite detection of ocean chlorophyll. Future improvements to ChlOSP could involve adding more of these factors – such as white caps, coccolithophores, and aerosols – along with more realistic satellite orbits and associated sensor challenges, including sun glint and high sensor zenith angle. Additional discrepancies between missing data in the real world and ChlOSP arise due to the model's representation of the Earth system (i.e., cloud cover and sea ice in a pre-industrial vs. present-day climate).

The spatial resolution of the simulated observations from ChlOSP match the spatial resolution of the CESM configuration used, which is ∼ 1° in this case. This coarse spatial resolution is needed to run long-term climate simulations, but it is much lower than the resolution of satellite data. As such, ChlOSP does not capture small-scale heterogeneity that exists in the real-world, including coastal variability. Figure 11 demonstrates how CESM does not resolve coastal regions, which tend to have high chlorophyll concentrations. There-

fore, this tool is best suited for large-scale, open-ocean analyses. To address partial cloud cover and sea ice within a model grid cell, we implemented a weighting method, which differs from the subcolumn strategy utilized in COSP.

Currently, our ability to compare modeled and real-world chlorophyll is limited because ChlOSP was developed for a free-running, fully coupled climate model simulation. While this configuration permits future projections, internal variability complicates our ability to compare the model to the real world. We anticipate that the next version of ChlOSP will be implemented in a hindcast configuration, i.e., an ocean-only version of the model forced with momentum, heat, and freshwater fluxes from historical observations spanning the duration of the satellite chlorophyll record. Using this version of the model along with in situ data, we plan to assess how clouds and missing data may have impacted our understanding of historical chlorophyll evolution.

Despite these uncertainties, our proof-of-concept simulation has demonstrated the utility of ChlOSP. The new model outputs allow for more robust comparisons between modeled and real-world chlorophyll, leading to improved model tuning and data assimilation capabilities. Initial results indicate that there are differences in the chlorophyll concentration between the typical model output and the satellite-like version. While we do not address all errors associated with ocean color remote sensing, we focus on one of the largest sources of uncertainty, namely missing data due to solar zenith angle and clouds. In the real world, we rarely know the true chlorophyll value where data are missing. However, in the model world, we know the exact values of all variables at every location and every time step, making it a powerful tool for estimating sampling bias. We anticipate that this tool will open the door to a wide body of future work.

Appendix A: Model outputs

We added the cloudfrac_modis and cloudfrac_isccp to the POP2 outputs for easier comparison to the other POP2 variables. These cloud fraction outputs include the daylight mask, which is written out as the cloudfrac_wgt variable. This output differs from the clear-sky weights because it does not include the weights from sea ice.

Table A1. New model outputs from CESM–ChlOSP. These outputs are written out in a new file stream within POP2.

	Name	Description
1	totChl	Sum of chlorophyll in the surface layer (10 m) over all phytoplankton functional types.
2	totChl_sat_nocld	Clear-sky chlorophyll (daylight-only sampling and sea ice mask).
3	totChl_sat_nocld_wgt	Weights used for clear-sky chlorophyll.
4	totChl_modis	Clear-sky chlorophyll weighted by MODIS-simulated cloud cover.
5	totChl_modis_wgt	Weights used for totChl_modis.
6	totChl_isccp	Clear-sky chlorophyll weighted by ISCCP-simulated cloud cover.
7	totChl_isccp_wgt	Weights used for totChl_isccp.
8	cloudfrac_modis	Total cloud fraction from MODIS simulator.
9	cloudfrac_isccp	Total cloud fraction from ISCCP simulator.
10	cloudfrac_wgt	Daylight weights for cloud fraction outputs.
11	totChl_sat_nocld_swath	Clear-sky chlorophyll with 13:30 LT swath sampling.
12	totChl_sat_nocld_wgt_swath	Weights used for clear-sky chlorophyll with 13:30 LT swath sampling.
13	totChl_modis_swath	MODIS chlorophyll with 13:30 LT swath sampling.
14	totChl_modis_wgt_swath	Weights used for MODIS chlorophyll with 13:30 LT swath sampling.
15	cloudfrac_modis_swath	MODIS cloud fraction with 13:30 LT swath sampling.
16	cloudfrac_modis_wgt_swath	Weights used for MODIS cloud fraction with 13:30 LT swath sampling.

Appendix B: Equations

For ChlOSP outputs, the climatology was calculated by applying the weighted mean. As an example, to calculate the mean of satellite-observed ISCCP chlorophyll over time, we use the model outputs totChl_isccp and totChl_isccp_wgt in Eq. (B1):

$$\text{weighted time mean} = \frac{\sum_t (\text{totChl\_isccp}(x, y, t))}{\sum_t (\text{totChl\_isccp\_wgt}(x, y, t))}, \quad \text{(B1)}$$

where totChl_isccp = chlorophyll × weight (calculated within POP2) and totChl_isccp_wgt = weight. The chlorophyll and weight variables correspond to the surface chlorophyll concentration in a grid cell and the fraction of the grid cell observed during a given time step, respectively. Seasonal cycles were evaluated in ocean biomes (Fig. S4) using Eq. (B2):

$$\text{weighted spatial mean} =$$

$$\frac{\sum_{x,y} (\text{totChl\_isccp}(x, y, t) \times \text{TAREA}(x, y))}{\sum_{x,y} (\text{totChl\_isccp\_wgt}(x, y, t) \times \text{TAREA}(x, y))}. \quad \text{(B2)}$$

This is equivalent to taking the weighted average in space, where the weights are equal to the total area observed (in square centimeters) within the biome at each time step. TAREA is the area of each model grid cell and is included in all POP2 output files. TAREA was subset for certain biomes of interest in these calculations. The total weighted mean over time and space is calculated with Eq. (B3):

$$\text{weighted total mean} =$$

$$\frac{\sum_{x,y,t} (\text{totChl\_isccp}(x, y, t) \times \text{TAREA}(x, y))}{\sum_{x,y,t} (\text{totChl\_isccp\_wgt}(x, y, t) \times \text{TAREA}(x, y))}. \quad \text{(B3)}$$

References

Andreae, M. O. and Crutzen, P. J.: Atmospheric Aerosols: Biogeochemical Sources and Role in Atmospheric Chemistry, Science, 276, 1052–1058, https://doi.org/10.1126/science.276.5315.1052, 1997.

Beaulieu, C., Henson, S. A., Sarmiento, J. L., Dunne, J. P., Doney, S. C., Rykaczewski, R. R., and Bopp, L.: Factors challenging our ability to detect long-term trends in ocean chlorophyll, Biogeosciences, 10, 2711–2724, https://doi.org/10.5194/bg-10-2711-2013, 2013.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnol. Oceanogr., 42, 1–20, https://doi.org/10.4319/lo.1997.42.1.0001, 1997.

Behrenfeld, M. J., O'Malley, R. T., Boss, E. S., Westberry, T. K., Graff, J. R., Halsey, K. H., Milligan, A. J., Siegel, D. A., and Brown, M. B.: Revaluating ocean warming impacts on global phytoplankton, Nat. Clim. Change, 6, 323–330, https://doi.org/10.1038/nclimate2838, 2016.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., and John, V. O.: COSP: Satellite simulation software for model assessment, B. Am. Meteorol. Soc., 92, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1, 2011.

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., and Schanen, D. P.: Higher-Order Turbulence Closure and Its Impact on Climate Simulations in the Community Atmosphere Model, J. Climate, 26, 9655–9676, https://doi.org/10.1175/JCLI-D-13-00075.1, 2013.

Boyce, D. G., Dowd, M., Lewis, M. R., and Worm, B.: Estimating global chlorophyll changes over the past century, Prog. Oceanogr., 122, 163–173, https://doi.org/10.1016/j.pocean.2014.01.004, 2014.

Chassot, E., Bonhommeau, S., Dulvy, N. K., Mélin, F., Watson, R., Gascuel, D., and Le Pape, O.: Global marine primary production constrains fisheries catches, Ecol. Lett., 13, 495–505, https://doi.org/10.1111/j.1461-0248.2010.01443.x, 2010.

Clow, G. L., Lovenduski, N. S., Levy, M. N., Lindsay, K., and Kay, J. E.: ChlOSPv1.0 output from 30 year pre-industrial simulation of CESMv2.2, Zenodo [data set], https://doi.org/10.5281/zenodo.8097543, 2023.

Clow, G. and CESM Team: CESM: ChlOSP Initial Release (v1.0), Zenodo [code], https://doi.org/10.5281/zenodo.8071063, 2023.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), J. Adv. Model. Earth Sy., 12, e2019MS001916, https://doi.org/10.1029/2019MS001916, 2020.

Dutkiewicz, S., Hickman, A. E., and Jahn, O.: Modelling ocean-colour-derived chlorophyll a, Biogeosciences, 15, 613–630, https://doi.org/10.5194/bg-15-613-2018, 2018.

European Space Agency: Ocean Colour Climate Change Initiative dataset, Version 6.0, ESA [data set], http://www.esa-oceancolour-cci.org (last access: 22 January 2024), 2022.

Fay, A. R. and McKinley, G. A.: Global open-ocean biomes: mean and temporal variability, Earth Syst. Sci. Data, 6, 273–284, https://doi.org/10.5194/essd-6-273-2014, 2014.

Fay, A. R. and McKinley, G. A.: Correlations of surface ocean pCO$_2$ to satellite chlorophyll on monthly to interannual timescales, Global Biogeochem. Cy., 31, 436–455, https://doi.org/10.1002/2016GB005563, 2017.

Geider, R. J., MacIntyre, H. L., and Kana, T. M.: A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature, Limnol. Oceanogr., 43, 679–694, https://doi.org/10.4319/lo.1998.43.4.0679, 1998.

Gentemann, C. L., Wentz, F. J., Brewer, M., Hilburn, K., and Smith, D.: Passive Microwave Remote Sensing of the Ocean: An Overview, Springer Netherlands, Dordrecht, 13–33, ISBN 978-90-481-8681-5, https://doi.org/10.1007/978-90-481-8681-5_2, 2010.

Gettelman, A. and Morrison, H.: Advanced Two-Moment Bulk Microphysics for Global Models. Part I: Off-Line Tests and Comparison with Other Schemes, J. Climate, 28, 1268–1287, https://doi.org/10.1175/JCLI-D-14-00102.1, 2015.

Golaz, J.-C., Larson, V. E., and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part I: Method and Model Description, J. Atmos. Sci., 59, 3540–3551, https://doi.org/10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2, 2002.

Gregg, W. W. and Casey, N. W.: Global and regional evaluation of the SeaWiFS chlorophyll data set, Remote Sens. Environ., 93, 463–479, https://doi.org/10.1016/j.rse.2003.12.012, 2004.

Gregg, W. W. and Casey, N. W.: Sampling biases in MODIS and SeaWiFS ocean chlorophyll data, Remote Sens. Environ., 111, 25–35, https://doi.org/10.1016/j.rse.2007.03.008, 2007.

Gregg, W. W. and Rousseaux, C. S.: Decadal trends in global pelagic ocean chlorophyll: A new assessment integrating multiple satellites, in situ data, and models, J. Geophys. Res.-Oceans, 119, 5921–5933, https://doi.org/10.1002/2014JC010158, 2014.

Hammond, M. L., Beaulieu, C., Sahu, S. K., and Henson, S. A.: Assessing trends and uncertainties in satellite-era ocean chlorophyll using space-time modeling, Global Biogeochem. Cy., 31, 1103–1117, https://doi.org/10.1002/2016GB005600, 2017.

Hu, C., Lee, Z., and Franz, B.: Chlorophyll-a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference, Journal of Geophysical Research: Oceans, 117, https://doi.org/10.1029/2011JC007395, 2012.

Jouini, M., Levy, M., Crépon, M., and Thiria, S.: Reconstruction of satellite chlorophyll images under heavy cloud coverage using a neural classification method, Remote Sens. Environ., 131, 232–246, https://doi.org/10.1016/j.rse.2012.11.025, 2013.

Kay, J. E., Hillman, B. R., Klein, S. A., Zhang, Y., Medeiros, B., Pincus, R., Gettelman, A., Eaton, B., Boyle, J., Marchand, R., and Ackerman, T. P.: Exposing Global Cloud Biases in the Community Atmosphere Model (CAM) Using Satellite Observations and Their Corresponding Instrument Simulators, J. Climate, 25, 5190–5207, https://doi.org/10.1175/JCLI-D-11-00469.1, 2012.

Kay, J. E., DeRepentigny, P., Holland, M. M., Bailey, D. A., DuVivier, A. K., Blanchard-Wrigglesworth, E., Deser, C., Jahn, A., Singh, H., Smith, M. M., Webster, M. A., Edwards, J., Lee, S.-S., Rodgers, K. B., and Rosenbloom, N.: Less Surface Sea Ice Melt in the CESM2 Improves Arctic Sea Ice Simulation With Minimal Non-Polar Climate Impacts, J. Adv. Model. Earth Sy., 14, e2021MS002679, https://doi.org/10.1029/2021MS002679, 2022.

King, M. D., Platnick, S., Menzel, W. P., Ackerman, S. A., and Hubanks, P. A.: Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites, IEEE T. Geosci. Remote, 51, 3826–3852, https://doi.org/10.1109/TGRS.2012.2227333, 2013.

Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., and Gleckler, P. J.: Are climate model simulations of clouds improving? An evaluation using the ISCCP simulator, J. Geophys. Res.-

Atmos., 118, 1329–1342, https://doi.org/10.1002/jgrd.50141, 2013.

Krasnopolsky, V., Nadiga, S., Mehra, A., Bayler, E., and Behringer, D.: Neural Networks Technique for Filling Gaps in Satellite Measurements: Application to Ocean Color Observations, Comput. Intel. Neurosc., 2016, 6156513, https://doi.org/10.1155/2016/6156513, 2016.

Krumhardt, K., Lovenduski, N., Long, M., 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.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Lenton, A., Li, H., Lovenduski, N. S., Orr, J. C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C. A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, Biogeosciences, 17, 3439–3470, https://doi.org/10.5194/bg-17-3439-2020, 2020.

Liu, X. and Wang, M.: Gap Filling of Missing Data for VIIRS Global Ocean Color Products Using the DINEOF Method, IEEE T. Geosci. Remote, 56, 4464–4476, https://doi.org/10.1109/TGRS.2018.2820423, 2018.

Long, M. C., Lindsay, K., and Holland, M. M.: Modeling photosynthesis in sea ice-covered waters, J. Adv. Model. Earth Sy., 7, 1189–1206, https://doi.org/10.1002/2015MS000436, 2015.

Long, M. C., Moore, J. K., Lindsay, K., Levy, M., Doney, S. C., Luo, J. Y., Krumhardt, K. M., Letscher, R. T., Grover, M., and Sylvester, Z. T.: Simulations With the Marine Biogeochemistry Library (MARBL), J. Adv. Model. Earth Sy., 13, e2021MS002647, https://doi.org/10.1029/2021MS002647, 2021.

Marinov, I., Doney, S. C., and Lima, I. D.: Response of ocean phytoplankton community structure to climate change over the 21st century: partitioning the effects of nutrients, temperature and light, Biogeosciences, 7, 3941–3959, https://doi.org/10.5194/bg-7-3941-2010, 2010.

McClain, C. R.: A Decade of Satellite Ocean Color Observations, Annu. Rev. Mar. Sci., 1, 19–42, https://doi.org/10.1146/annurev.marine.010908.163650, 2009.

Mikelsons, K. and Wang, M.: Optimal satellite orbit configuration for global ocean color product coverage, Opt. Express, 27, A445–A457, https://doi.org/10.1364/OE.27.00A445, 2019.

NASA Goddard Space Flight Center: MODIS MOD08_M3 Cloud Fraction, obs4MIPS, NASA [data set], https://aims3.llnl.gov/thredds/dodsC/user_pub_work/obs4MIPs/NASA-GSFC/MODIS-1-0/mon/clt/gn/v20180305/clt_mon_MODIS-1-0_BE_gn_200003-201109.nc, (last access: 30 October 2023), 2019.

NASA Ocean Biology Processing Group: Aqua MODIS Level 3 Mapped Chlorophyll Data, Version R2022.0 [data set], https://doi.org/10.5067/AQUA/MODIS/L3M/CHL/2022, 2022.

O'Malley, R. T., Behrenfeld, M. J., Westberry, T. K., Milligan, A. J., Shang, S., and Yan, J.: Geostationary satellite observations of dynamic phytoplankton photophysiology, Geophys. Res. Lett., 41, 5052–5059, https://doi.org/10.1002/2014GL060246, 2014.

O'Reilly, J. E. and Werdell, P. J.: Chlorophyll algorithms for ocean color sensors – OC4, OC5 & OC6, Remote Sens. Environ., 229, 32–47, https://doi.org/10.1016/j.rse.2019.04.021, 2019.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., and McClain, C.: Ocean color chlorophyll algorithms for SeaWiFS, J. Geophys. Res.-Oceans, 103, 24937–24953, https://doi.org/10.1029/98JC02160, 1998.

Pincus, R., Platnick, S., Ackerman, S. A., Hemler, R. S., and Hofmann, R. J. P.: Reconciling Simulated and Observed Views of Clouds: MODIS, ISCCP, and the Limits of Instrument Simulators, J. Climate, 25, 4699–4720, https://doi.org/10.1175/JCLI-D-11-00267.1, 2012.

Rossow, W., Golea, V., Walker, A., Knapp, K., Young, A., Hankins, B., and Inamdar, A.: International Satellite Cloud Climatology Project (ISCCP) Climate Data Record, H-Series, NOAA National Centers for Environmental Information [data set], https://doi.org/10.7289/V5QZ281S, 2017.

Rossow, W. B. and Schiffer, R. A.: ISCCP Cloud Data Products, B. Am. Meteorol. Soc., 72, 2–20, https://doi.org/10.1175/1520-0477(1991)072<0002:ICDP>2.0.CO;2, 1991.

Salisbury, J. E., Jönsson, B. F., Mannino, A., Kim, W., Goes, J. I., Choi, J.-Y., and Concha, J. A.: Assessing Net Growth of Phytoplankton Biomass on Hourly to Annual Time Scales Using the Geostationary Ocean Color Instrument, Geophys. Res. Lett., 48, e2021GL095528, https://doi.org/10.1029/2021GL095528, 2021.

Sathyendranath, S., Brewin, R. J., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Cipollini, P., Couto, A. B., Dingle, J., Doerffer, R., Donlon, C., Dowell, M., Farman, A., Grant, M., Groom, S., Horseman, A., Jackson, T., Krasemann, H., Lavender, S., Martinez-Vicente, V., Mazeran, C., Mélin, F., Moore, T. S., Müller, D., Regner, P., Roy, S., Steele, C. J., Steinmetz, F., Swinton, J., Taberner, M., Thompson, A., Valente, A., Zühlke, M., Brando, V. E., Feng, H., Feldman, G., Franz, B. A., Frouin, R., Gould, R. W., Hooker, S. B., Kahru, M., Kratzer, S., Mitchell, B. G., Muller-Karger, F. E., Sosik, H. M., Voss, K. J., Werdell, J., and Platt, T.: An Ocean-Colour Time Series for Use in Climate Studies: The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI), Sensors, 19, 4285, https://doi.org/10.3390/s19194285, 2019.

Schlunegger, S., Rodgers, K. B., Sarmiento, J. L., Ilyina, T., Dunne, J. P., Takano, Y., Christian, J. R., Long, M. C., Frölicher, T. L., Slater, R., and Lehner, F.: Time of Emergence and Large Ensemble Intercomparison for Ocean Biogeochemical Trends, Global Biogeochem. Cy., 34, e2019GB006453, https://doi.org/10.1029/2019GB006453, 2020.

Scott, J. P. and Werdell, P. J.: Comparing level-2 and level-3 satellite ocean color retrieval validation methodologies, Opt. Express, 27, 30140–30157, https://doi.org/10.1364/OE.27.030140, 2019.

Siegel, D., Behrenfeld, M., Maritorena, S., McClain, C., Antoine, D., Bailey, S., Bontempi, P., Boss, E., Dierssen, H., Doney, S., Eplee, R., Evans, R., Feldman, G., Fields, E., Franz, B., Kuring, N., Mengelt, C., Nelson, N., Patt, F., Robinson, W., Sarmiento, J., Swan, C., Werdell, P., Westberry, T., Wilding, J., and Yoder, J.: Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission, Remote Sens. Environ., 135, 77–91, https://doi.org/10.1016/j.rse.2013.03.025, 2013.

Smith, R., Jones, P., Briegleb, B. P., Bryan, F. O., Danabasoglu, G., Dennis, J. M., Dukowicz, J., Eden, C., Fox-Kemper, B., Gent,

P. R., Hecht, M., Jayne, S., Jochum, M., Large, W. G., Lindsay, K., Maltrud, M., Norton, N. J., Peacock, S. L., Vertenstein, M., and Yeager, S.: The Parallel Ocean Program (POP) reference manual: Ocean component of the Community Climate System Model (CCSM), Tech. Rep. LAUR-10-01853, Los Alamos National Laboratory, 2010.

Stock, A., Subramaniam, A., Van Dijken, G. L., Wedding, L. M., Arrigo, K. R., Mills, M. M., Cameron, M. A., and Micheli, F.: Comparison of Cloud-Filling Algorithms for Marine Satellite Data, Remote Sensing, 12, 3313, https://doi.org/10.3390/rs12203313, 2020.

Swales, D. J., Pincus, R., and Bodas-Salcedo, A.: The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2, Geosci. Model Dev., 11, 77–81, https://doi.org/10.5194/gmd-11-77-2018, 2018.

Tittensor, D. P., Eddy, T. D., Lotze, H. K., Galbraith, E. D., Cheung, W., Barange, M., Blanchard, J. L., Bopp, L., Bryndum-Buchholz, A., Büchner, M., Bulman, C., Carozza, D. A., Christensen, V., Coll, M., Dunne, J. P., Fernandes, J. A., Fulton, E. A., Hobday, A. J., Huber, V., Jennings, S., Jones, M., Lehodey, P., Link, J. S., Mackinson, S., Maury, O., Niiranen, S., Oliveros-Ramos, R., Roy, T., Schewe, J., Shin, Y.-J., Silva, T., Stock, C. A., Steenbeek, J., Underwood, P. J., Volkholz, J., Watson, J. R., and Walker, N. D.: A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1.0, Geosci. Model Dev., 11, 1421–1442, https://doi.org/10.5194/gmd-11-1421-2018, 2018.

van Oostende, M., Hieronymi, M., Krasemann, H., and Baschek, B.: Global ocean colour trends in biogeochemical provinces, Frontiers in Marine Science, 10, 1–13, https://doi.org/10.3389/fmars.2023.1052166, 2023.

Wang, Y., Liu, X., Hoose, C., and Wang, B.: Different contact angle distributions for heterogeneous ice nucleation in the Community Atmospheric Model version 5, Atmos. Chem. Phys., 14, 10411–10430, https://doi.org/10.5194/acp-14-10411-2014, 2014.

Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., Keller, T., DeLuisi, J., Wuebbles, D. J., Kerr, J. B., Miller, A. J., Oltmans, S. J., and Frederick, J. E.: Factors affecting the detection of trends: Statistical considerations and applications to environmental data, J. Geophys. Res.-Atmos., 103, 17149–17161, https://doi.org/10.1029/98JD00995, 1998.

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., Klein, S. A., Marchand, R., Medeiros, B., Siebesma, A. P., Skinner, C. B., Stevens, B., Tselioudis, G., Tsushima, Y., and Watanabe, M.: The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, Geosci. Model Dev., 10, 359–384, https://doi.org/10.5194/gmd-10-359-2017, 2017.

Wilson, J. D., Andrews, O., Katavouta, A., de Melo Viríssimo, F., Death, R. M., Adloff, M., Baker, C. A., Blackledge, B., Goldsworth, F. W., Kennedy-Asser, A. T., Liu, Q., Sieradzan, K. R., Vosper, E., and Ying, R.: The biological carbon pump in CMIP6 models: 21st century trends and uncertainties, P. Natl. Acad. Sci. USA, 119, e2204369119, https://doi.org/10.1073/pnas.2204369119, 2022.

Yeager, S. G., Rosenbloom, N., Glanville, A. A., Wu, X., Simpson, I., Li, H., Molina, M. J., Krumhardt, K., Mogen, S., Lindsay, K., Lombardozzi, D., Wieder, W., Kim, W. M., Richter, J. H., Long, M., Danabasoglu, G., Bailey, D., Holland, M., Lovenduski, N., Strand, W. G., and King, T.: The Seasonal-to-Multiyear Large Ensemble (SMYLE) prediction system using the Community Earth System Model version 2, Geosci. Model Dev., 15, 6451–6493, https://doi.org/10.5194/gmd-15-6451-2022, 2022.