https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Evaluating the performance of CMIP6 models in simulating Southern Ocean biogeochemistry

Ming Cheng[1,2], Nicola Maher[3], Michael J. Ellwood[1,2]

[1]Research School of Earth Sciences, Australian National University, Canberra, ACT2601, Australia
[2]Australian Centre for Excellence in Antarctic Science, Research School of Earth Sciences, Australian National University, Canberra, ACT2601, Australia
[3]ARC Centre of Excellence for Weather of the 21st Century, Australian National University, Canberra, ACT2601, Australia

Correspondence to: Ming Cheng (ming.cheng@anu.edu.au)

Abstract. The Southern Ocean plays a vital role in global biogeochemical cycles, yet the quality of its representation in Earth System Models (ESMs) remains unquantified. This study evaluates the performance of 14 Coupled Model Intercomparison Project Phase 6 (CMIP6) models in simulating key biogeochemical variables south of 30°S, including austral-summer surface chlorophyll, deep chlorophyll maxima (DCMs), nitrate, silicate, dissolved iron, and yearly particulate organic carbon (POC). Model output for the period 2000–2014 is compared to multiple observational datasets, such as Copernicus for chlorophyll and POC, the World Ocean Atlas (WOA) for nitrate and silicate, and GEOTRACES for dissolved iron. Model performance is assessed using statistical metrics including mean bias error (MBE), standardised standard deviation (SSD), root mean squared deviation (RMSD), and correlation coefficient (CC). The results reveal substantial inter-model variability, with individual models exhibiting strengths in simulating different variables. GFDL-ESM4 best reproduces chlorophyll and DCM patterns, IPSL-CM6A-LR performs well for nutrients, MIROC-ES2L for dissolved iron, and CMCC-ESM2 for POC. Based on composite rankings, the top-performing models are IPSL-CM6A-LR, GFDL-ESM4, CMCC-ESM2, UKESM1-0-LL, and CNRM-ESM2-1. This work underscores the importance of multi-model evaluation for identifying model strengths and guiding future improvements in biogeochemical (BGC) model development, particularly in the context of understanding and projecting Southern Ocean biogeochemistry under climate change.

1 Introduction

Climate change is a critical global challenge, driving major shifts in marine conditions and ecosystems. The Southern Ocean, covering 30% of the global ocean, plays a crucial role in the oceanic carbon and nutrient cycles, absorbing over 40% of anthropogenic CO_2 and 70% of human-induced warming (Gruber et al., 2019; Petrou et al., 2016; Xue et al., 2024). The Southern Ocean is characterised by complex interactions among physical circulation, biogeochemistry, and biological productivity, making it a challenge to model (Henley et al., 2020; Morley et al., 2020). The powerful eastward-moving Antarctic Circumpolar Current (ACC), one of the Earth's strongest currents, connects ocean basins and regulates global

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

climate and ocean circulation, supports diverse marine ecosystems, and distributes nutrients (Böning et al., 2008; Rintoul et al., 2001; Lopes et al., 2011; Song, 2020). The upwelling of deep, nutrient-rich waters, driven by ACC, supports phytoplankton growth, influencing global carbon sequestration and ecosystem dynamics (Venables and Moore, 2010; Morrison et al., 2015; Hunt et al., 2021; Pollard et al., 2006). This complex region of both physical and biological processes is important due to its significant impact on global climate regulation, carbon sequestration, and the health of marine ecosystems.

Phytoplankton, particularly silicifying diatoms, are a key component of the Southern Ocean food web and the global carbon cycle, playing a crucial role in carbon sequestration and nutrient cycling (Deppeler and Davidson, 2017; Baldry et al., 2020; Petrou et al., 2016; Nissen and Vogt, 2021; Timmermans et al., 2004; Hoffmann et al., 2008). Their biomass and primary production are often assessed through chlorophyll concentrations, which serve as an essential indicator in oceanic carbon fixation and ecosystem productivity (Carranza and Gille, 2015; Johnson et al., 2013). However, despite the abundance of macronutrients such as nitrate and silicate, phytoplankton growth is frequently constrained by light limitation and iron deficiency, both of which regulate their distribution and productivity (Boyd and Ellwood, 2010; Boyd, 2002). In response to these physiochemical conditions, deep chlorophyll maxima (DCMs) have been observed in nutrient-stratified waters during austral summer in the Southern Ocean, indicating robust phytoplankton production in the subsurface layer (Boyd et al., 2024; Cornec et al., 2021; Cullen, 1982; Cullen, 2015; Hopkinson and Barbeau, 2008; Li et al., 2012). These DCMs contribute significantly to the regional carbon cycle, for example, approximately 40% of primary production in the Southern Ocean occurs below the mixed layer (Vives et al., 2024), and support marine food webs by sustaining primary production below the surface, where light and nutrient conditions are more favourable for certain phytoplankton communities (Signorini et al., 2015; Cornec et al., 2021; Sauzède et al., 2018).

Ocean biogeochemical (BGC) modules, are an important component of coupled Earth system models (ESMs), and are indispensable for understanding the complicated physical and biogeochemical processes in the ocean (Follows and Dutkiewicz, 2011; Séférian et al., 2020). Depending on their complexity, these models simulate the cycles of key elements such as carbon, oxygen, nitrogen, phosphorus, silicate, and iron, and organisms including phytoplankton, zooplankton and bacteria, which are vital for marine ecosystems and global climate regulation (Dunne et al., 2020; Aumont et al., 2015; Pak et al., 2021; Ilyina et al., 2013). BGC models enable researchers to investigate how changes in environmental conditions, such as temperature, light, and nutrient availability, impact marine biogeochemistry and ecosystem dynamics (Kwiatkowski et al., 2020). They are particularly valuable for studying regions like the Southern Ocean, where observational data are limited, and the interactions between physical and biogeochemical processes are highly complex (Tagliabue et al., 2017; Lauderdale et al., 2017). Despite their significance, BGC models face considerable challenges, including the need for precise parameterisation of key biological processes, accurate representation of small-scale processes, and effective integration of diverse data sources (Ackermann et al., 2024; Beadling et al., 2019).

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

The Coupled Model Intercomparison Project Phase 6 (CMIP6) represents the latest advancement in climate modelling, providing a standardised framework for evaluating ESMs across various simulations under different climate scenarios (Eyring et al., 2016; O'neill et al., 2016; Meehl et al., 2020). Compared to previous phases, CMIP6 models feature higher spatial resolution, improved physical processes, and enhanced biogeochemical components, including expanded phytoplankton functional types, refined biogeochemical cycle representations and optimised parameterisation (Séférian et al., 2020; Kwiatkowski et al., 2020). However, significant discrepancies persist in biogeochemical performance due to variations in BGC model structures, parameterisation, and ocean physics (Séférian et al., 2020). Evaluating CMIP6 models highlights these differences, offering insights for future model development and refinement (Kwiatkowski et al., 2020; Séférian et al., 2020; Hauck et al., 2015).

While some studies have assessed the performance of CMIP6 models in simulating biogeochemical variables globally and regionally, a comprehensive analysis of chlorophyll, nutrient distribution, and DCM characteristics in the Southern Ocean remains unexplored. Marshal et al. (2024) evaluated chlorophyll, phytoplankton, nitrate and dissolved oxygen across 13 CMIP6 models in the South China Sea, ranking them using statistical metrics to identify the five best-performing models. Fisher et al. (2025) synthesised CMIP6 outputs to examine climate-driven shifts in Southern Ocean primary production, projecting a 30% increase in Antarctic zone productivity under a high-emission (SSP5-8.5) scenario, albeit with regional variations. Séférian et al. (2020) compared CMIP5 and CMIP6 models, demonstrating improved CMIP6 biogeochemical representations, including chlorophyll, dissolved oxygen, silicate and nitrate, due to more comprehensive biogeochemical cycles and Earth system interactions. Rohr et al. (2023) analysed 11 CMIP6 models and found that zooplankton grazing parameterisation introduced uncertainty in marine carbon cycle projections. These studies underscore the need for further evaluation of the CMIP6 models to assess the impact of biogeochemical processes and parameterisation on model performance.

In this paper, we evaluate biogeochemical variables-including chlorophyll, silicate, nitrate and dissolved iron, across 12 CMIP6 models and assess their performance in representing DCMs in the Southern Ocean. Sect. 2 details the observed and simulated and the statistical analysis methods. Sect. 3 presents an inter-model evaluation of each biogeochemical variable. Sect. 4 discusses the ocean vertical carbon structure, model performance, as well as avenues for improvement. Sect. 5 provides a summary of our findings.

2 Data and methods

2.1 Study region

This study focuses on the open waters of the Southern Ocean (south of 30°S). We divide the Southern Ocean into four zones: the subtropical zone (STZ), subantarctic zone (SAZ), polar front zone (PFZ) and Antarctic zone (AZ) (Fig. 1). These zones are separated by three key fronts: the subtropical front, subantarctic front and polar front, which are defined by distinct

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

physical and biogeochemical properties (Orsi et al., 1995). We compare the CMIP6 model outputs of chlorophyll, nitrate, silicate and dissolved iron across these zones and across the entire Southern Ocean.

2.2 CMIP6 datasets and availability

We obtained outputs from 14 CMIP6 models from the Earth System Grid Federation (ESGF) Nodes (Cinquini et al., 2014). Specifically, we collected data from the historical experiment for model evaluation, using the ensemble member r1i1p1f1 for most models, while r1i1p1f2 was used for CNRM-ESM2-1, MIROC-ES2L, and UKESM1-0-LL. Dissolved iron and carbon data of ACCESS-ESM1-5 were collected from National Computational Infrastructure (NCI). This includes monthly data for chlorophyll, nitrate, silicate, and dissolved iron, as well as yearly data for particulate organic carbon (POC), which comprises phytoplankton, zooplankton, detritus, and bacteria (see Sect. 2.4 for details). These data were also used to compare chlorophyll and DCM distribution. The selected CMIP6 models, their properties and available variables are detailed in Table 1. To ensure consistency, we regridded all outputs to a 1°×1° common horizontal resolution using bilinear interpolation in Climate Data Operators (CDO) software (Schulzweida, 2023), covering the time range from January 2000 to December 2014.

Table 1. List of 14 CMIP6 models utilised, detailing the ESM name, coupled ocean biogeochemical model (OBGCM) name, averaged horizontal resolution and variables with available data. All variable abbreviations and their long names: chl (mass concentration of phytoplankton expressed as chlorophyll in sea water), no3 (dissolved nitrate concentration), si (total dissolved inorganic silicon concentration), dfe (dissolved iron concentration), phyc (phytoplankton carbon concentration), zooc (zooplankton carbon concentration), detoc (mole concentration of organic detritus expressed as carbon in seawater), bacc (bacterial carbon concentration).

ESM	OBGCM	Variable	ESM and OBGCM Reference
ACCESS-ESM1-5	WOMBAT	chl, no3, dfe, phyc, zooc, detoc	Ziehn et al. (2020); Oke et al. (2013)
CanESM5	CMOC	chl, no3, phyc, zooc, detoc	Swart et al. (2019); Zahariev et al. (2007)
CESM2	MARBL	chl, no3, si, dfe, phyc, zooc	Danabasoglu et al. (2020); Long et al. (2021)
CMCC-ESM2	BFM v5.2	chl, no3, si, dfe, phyc, zooc, detoc, bacc	Lovato et al. (2022); Vichi et al. (2015)
CNRM-ESM2-1	PISCES-v2-gas	chl, no3, si, dfe, phyc, zooc, detoc	Séférian et al. (2019); (Skyllas, 2018)
GFDL-ESM4	COLBALTv2	chl, no3, si, dfe, phyc, zooc, detoc, bacc	Dunne et al. (2020); Stock et al. (2020)
IPSL-CM6A-LR	PISCES-v2	chl, no3, si, dfe, phyc, zooc, detoc	Boucher et al. (2020); Aumont et al. (2015)
MIROC-ES2L	OECO-v2	chl, no3, dfe, phyc, zooc	Hajima et al. (2020)
MPI-ESM-1-2-HAM	HAMOCC6	chl, no3, si, dfe, phyc, zooc, detoc	Neubauer et al. (2019); Ilyina et al. (2013)
MPI-ESM1-2-HR	HAMOCC6	chl, no3, si, dfe, phyc, zooc, detoc	Müller et al. (2018); Ilyina et al. (2013)
MPI-ESM1-2-LR	HAMOCC6	chl, no3, si, dfe, phyc, zooc, detoc	Mauritsen et al. (2019); Ilyina et al. (2013)
NorESM2-LM	HAMOCC	chl, no3, si, dfe, phyc, zooc, detoc	Tjiputra et al. (2020)
NorESM2-MM	HAMOCC	chl, no3, si, dfe, phyc, zooc, detoc	Tjiputra et al. (2020)
UKESM1-0-LL	MEDUSA-2.0	chl, no3, si, dfe, phyc, zooc, detoc	Sellar et al. (2019); (Yool et al., 2013)

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

2.3 Observed datasets and availability

Observed chlorophyll data was obtained from the Copernicus Global Ocean 3D Chlorophyll-a Concentration, Particulate Backscattering coefficient and Particulate Organic Carbon (Sauzède et al., 2016), which estimates chlorophyll and POC using a neural network method. This reprocessed dataset has a 0.25°×0.25° horizontal resolution, covering 36 vertical levels from the surface to 1000 m depth. Observed nitrate and silicate data were sourced from the World Ocean Atlas (WOA) 2018 (Garcia et al., 2019), representing climatological averages from 1955 to 2017. Observed dissolved iron data were obtained from GEOTRACES (Tagliabue et al., 2012), which compiles bottle-sampled dissolved iron measurements from 2001 to 2014.

2.4 Data analysis

To evaluate the performance of CMIP6 models in simulating biogeochemical variables, we compared observations with model outputs for chlorophyll, nitrate, silicate and dissolved iron (Sect. 3.1), assessed DCM (peak of chlorophyll concentration in the subsurface) representation and characteristics (Sect. 3.2), analysed particulate organic carbon (Sect. 3.3), and presented model rankings by variable (Sect. 3.4).

Since Southern Ocean DCMs predominantly occur during austral summer (Cornec et al., 2021; Prakash and Bhaskar, 2024), all datasets (except dissolved iron and POC) were restricted to December, January and February (DJF). We calculated temporal averages for CMIP6-simulated variables and observed chlorophyll over DJF from 2000 to 2014. Similarly, we computed DJF-averaged nitrate and silicate from observations. Given that dissolved iron observations are derived from bottle-sampled data rather than gridded products, we selected observations from depths less than 10 m to represent surface iron concentrations and interpolated CMIP6 outputs to these observation sites.

In cases where CMIP6 models do not provide a specific variable representing total particulate organic carbon (POC), we manually derive it by summing different species of POC. The simulated POC concentration in this paper is calculated as the sum of phytoplankton carbon, zooplankton carbon, detrital organic carbon (absent in CESM2 and unavailable in MIROC-ES2L), and bacterial carbon (optional; available only in CMCC-ESM2 and GFDL-ESM4). Because many CMIP6 models lack monthly POC-related data, we utilise yearly data instead, as carbon export predominantly occurs during summer months (Boyd et al., 2019; Buesseler et al., 2007; Blain et al., 2007).

To quantify model performance, we calculated spatial variations, mean bias error (MBE), standardised standard deviation (SSD), correlation coefficient (CC), and root mean squared deviation (RMSD) for chlorophyll, nitrate, silicate and dissolved iron. We visualised spatial variations using Southern Ocean maps, MBE in bar charts, SSD, CC and RMSD using Taylor Diagram (TD) to illustrate the agreement between models and observations (Taylor, 2001). The TDs and their related

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

statistics-SSD, CC, and RMSD-are provided in Supplementary Materials. The equations for MBE, SSD, CC, and RMSD are presented below:

$$MBE = \frac{1}{n}\sum(y_i - x_i)\,, \tag{1}$$

$$SSD = \frac{\sqrt{\frac{1}{n}\sum(y_i-\bar{y})^2}}{\sqrt{\frac{1}{n}\sum(x_i-\bar{x})^2}}, \tag{2}$$

$$CC = \frac{\sum(y_i-\bar{y})(x_i-\bar{x})}{\sqrt{\sum(y_i-\bar{y})^2\sum(x_i-\bar{x})^2}}\,, \tag{3}$$

$$RMSD = \sqrt{\frac{1}{n}\sum(y_i - x_i)^2}\,, \tag{4}$$

where x_i and y_i represent the observed and simulated values, respectively. $\bar{x}$ and $\bar{y}$ denote the mean values of observations and simulations. n is the number of grid points in the datasets.

DCMs are identified as the vertical peak of chlorophyll concentration, where the chlorophyll value exceeds 1.1 times the surface chlorophyll concentration. The 1.1 threshold is applied to account for potential measurement errors in the observation. To evaluate DCM characteristics, we calculated peak chlorophyll concentration at the identified DCM depth and frequency of DCM occurrence, which is defined as the number of grid points where DCMs are detected.

The assessment of CMIP6 model performance relies on the ranking of four statistical metrics, containing MBE, SSD, RMSD, and CC for chlorophyll, nitrate, silicate, dissolved iron, and POC. For the evaluation of DCMs, both the chlorophyll rank and DCM occurrence frequency are considered. The overall performance of each CMIP6 model is represented by the average rank across these six variables (Sect. 3.4).

All data processing and analysis were performed using MATLAB R2024a and its numerical toolboxes. Maps were generated using the M_Map toolbox (Pawlowicz, 2020). Taylor diagrams in supplementary materials were generated using MATLAB functions from Haroon Haider (https://www.youtube.com/@EngrHaroonHaider, last accessed: 22 April 2025).

3 Results

3.1 Southern Ocean biogeochemistry

We evaluate the performance of 14 CMIP6 models in simulating Southern Ocean biogeochemistry by comparing their outputs for chlorophyll, nitrate, silicate, and dissolved iron with observational data. The surface chlorophyll concentration in the Southern Ocean exhibits a general increase from north to south, reaching its highest concentrations in the coastal regions of Antarctica (Fig. 1), with some exceptions associated with island wake effects related to continental iron input (Blain et al., 2007). In contrast, the chlorophyll simulations exhibit significant discrepancies across models. The three MPI-ESM models,

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

MPI-ESM-1-2-HAM, MPI-ESM1-2-HR, and MPI-ESM1-2-LR, tend to overestimate chlorophyll concentrations throughout the Southern Ocean, with MBEs of 1.0, 1.8 and 0.8 mg/m^3 (Fig. 2), respectively, compared to a mean chlorophyll concentration of only 0.6 mg/m^3 in observations. Conversely, the CanESM5, CMCC-ESM2, CNRM-ESM2-1, and IPSL-CM6A-LR models underestimate chlorophyll concentrations (Figs. 1 and 2). The ACCESS-ESM1-5, CESM2, MIROC-ES2L, NorESM2-LM, NorESM2-MM, and UKESM1-0-LL models exhibit small and negative MBEs for the entire Southern Ocean but showed opposing biases across regions. For instance, they overestimate chlorophyll concentrations north of the subtropical front and underestimated concentrations to the south (Fig. 2). The GFDL-ESM4 model provides the most accurate simulation of chlorophyll concentration north of the polar front but underestimates concentrations south of the polar front (Fig. 1).

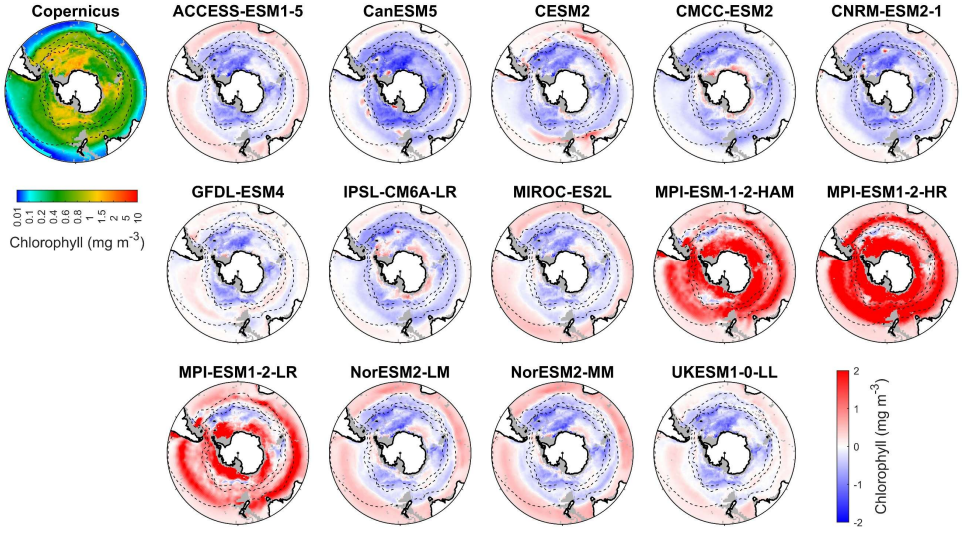


Figure 1: Observed surface chlorophyll concentrations from Copernicus in DJF and spatial biases of surface chlorophyll concentrations for 14 CMIP6 models (model chlorophyll – Copernicus chlorophyll) in DJF for the Southern Ocean (>30°S). Black dashed lines in the maps denote the subtropical front, the subantarctic front, and the polar front, from north to south. Grey areas denote regions where no data are available.

When considering other metrics such as standardised standard deviation (SSD), root mean-squared deviation (RMSD), and correlation coefficient (CC), we find that among the models, GFDL-ESM4, IPSL-CM6A-LR, and CMCC-ESM2 have the lowest RMSD, small bias errors, and CC values above 0.6, indicating that they were the best-performing models for simulating the distribution of chlorophyll across the Southern Ocean (Fig. S1 and Table S1). In contrast, the three MPI-ESMs are less reliable due to their overestimation of chlorophyll concentration. Additionally, the ACCESS-ESM1-5, CanESM5, and NorESMs models exhibit poor performance, such as their low CC (<0.2), despite moderate bias errors (Fig.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

S1 and Table S1). The remaining models, including CESM2, CNRM-ESM2-1, MIROC-ES2L, and UKESM1-0-LL, exhibit moderate performance in simulating chlorophyll.

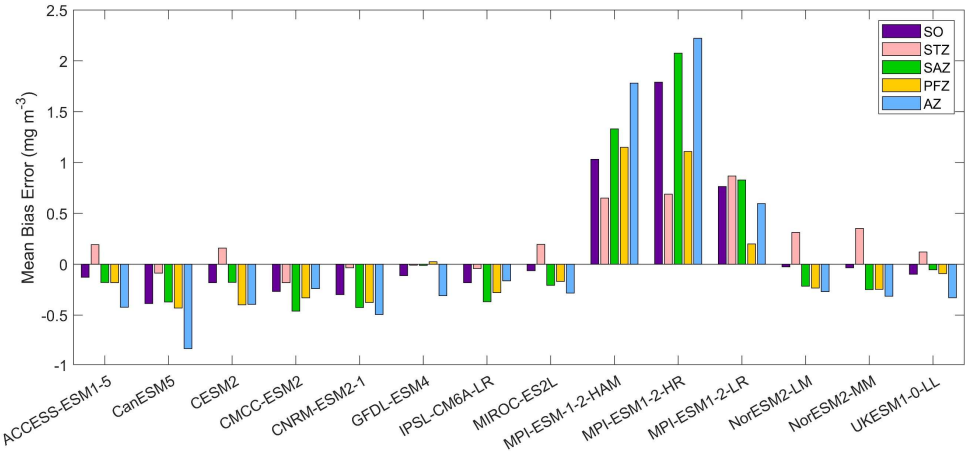


Figure 2: The mean bias errors in surface chlorophyll concentrations for the Southern Ocean (SO), the subtropical zone (STZ), the subantarctic zone (SAZ), the polar front zone (PFZ), and the Antarctic zone (AZ) in DJF. All MBEs and means in each region are calculated using area-weighted averages.

Nitrate, a key macronutrient that regulates phytoplankton growth and primary production, is abundant in the Southern Ocean, particularly south of 50°S (Fig. 3). Three MPI-ESMs (MPI-ESM-1-2-HAM, MPI-ESM1-2-HR, and MPI-ESM1-2-LR) underestimate nitrate concentrations, with MBEs of -4.7, -5.8, and -3.5 mmol/m^3 (Fig. 4), respectively, compared to the observed mean surface nitrate concentration of 11.9 mmol/m^3 from WOA. This underestimation may be linked to the high simulated chlorophyll levels, which could lead to excessive nutrient consumption. In addition, the CESM2, CMCC-ESM2, and GFDL-ESM4 models also underestimate nitrate concentrations (Fig. 3). In contrast, the ACCESS-ESM1-5, CanESM5, CNRM-ESM2-1, MIROC-ES2L, NorESM2-LM, NorESM2-MM, and UKESM1-0-LL models overestimate nitrate concentration, although the two NorESMs underestimate it in the Antarctic zone. Among all models, IPSL-CM6A-LR has the best performance, with the lowest MBE of 0.29 mmol/m^3 and a relative error of just 2.43% (Fig. 4).

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

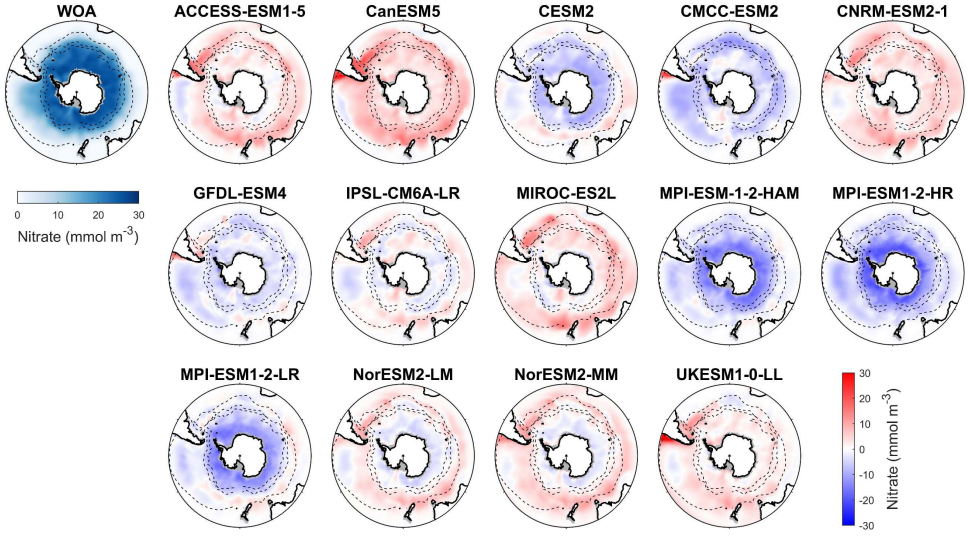


Figure 3: Observed surface nitrate concentrations from WOA in DJF and spatial biases of surface nitrate concentrations for 14 CMIP6 models (model nitrate – WOA nitrate) in DJF for the Southern Ocean (>30°S).

Among the 14 CMIP6 models, IPSL-CM6A-LR, GFDL-ESM4, and CNRM-ESM2-1 produce the most accurate simulations of surface nitrate concentration for the Southern Ocean. They exhibit the lowest RMSD (<0.3), minimal MBE (absolute MBE < 4 mmol/m^3), high CC (>0.95), and SSDs close to 1, indicating strong agreement with observations (Fig. S2 and Table S2). Conversely, the three MPI-ESMs models produce less accurate simulations of surface nitrate concentration for the Southern Ocean due to their large bias errors and significant deviations (represented by SSD, RMSD, and CC on a Taylor diagram) (Fig. S2 and Table S2). The remaining models including ACCESS-ESM1-5, CanESM5, CESM2, CMCC-ESM2, MIROC-ES2L, NorESM2-LM, NorESM2-MM and UKESM1-0-LL demonstrate moderate performance.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

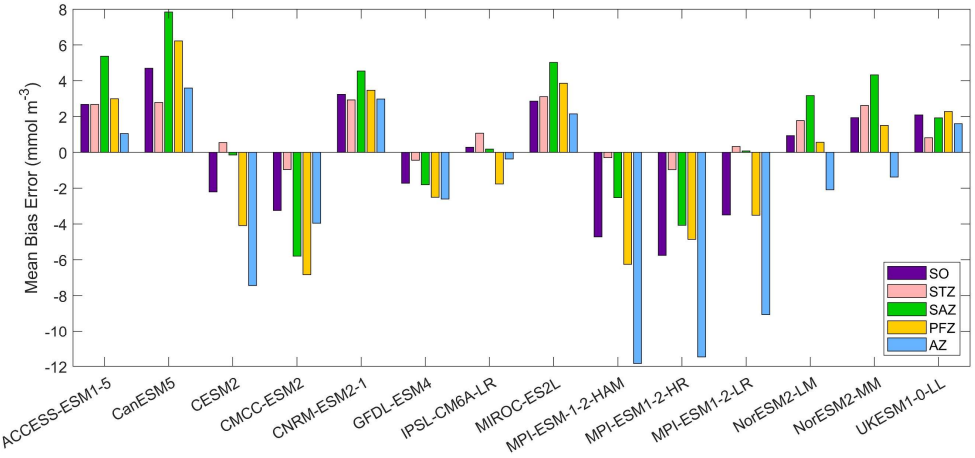


Figure 4: The mean bias errors in surface nitrate concentrations for the Southern Ocean (SO), the subtropical zone (STZ), the subantarctic zone (SAZ), the polar front zone (PFZ), and the Antarctic zone (AZ) in DJF.

Among the CMIP6 models analysed, silicate concentrations are generally overestimated across the Southern Ocean (Fig. 5). The three MPI-ESMs exhibit the most significant overestimation, with MBEs exceeding 30 mmol/m^3 (Fig. 6), over twice the observed surface silicate concentration of 12.7 mmol/m^3 from WOA. The CMCC-ESM2, NorESM2-LM, NorESM2-MM, and UKESM1-0-LL models also show large positive biases, with their mean silicate concentrations roughly double that of observed values (Fig. 6). The CMCC-ESM2 and UKESM1-0-LL models underestimate silicate concentrations in the subtropical zone (STZ), while the two NorESMs models underestimate silicate concentrations in the Ross Sea, Weddell Sea, and adjacent waters (Fig. 5). CESM2, CNRM-ESM2-1, GFDL-ESM4, and IPSL-CM6A-LR exhibit the lowest positive MBEs among the models (Fig. 6). and underestimate silicate concentrations in the STZ. Interestingly, in some regions around Antarctica, simulated silicate concentrations are lower than observations, particularly in areas where the GFDL-ESM4 and IPSL-CM6A-LR models overestimate chlorophyll (Fig. 5), suggesting a possible link between silicate availability and diatom growth. Three models, including ACCESS-ESM1-5, CanESM5, and MIROC-ES2L are excluded from the silicate comparison because they do not include diatoms as one of their phytoplankton species or silicate as a nutrient variable.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

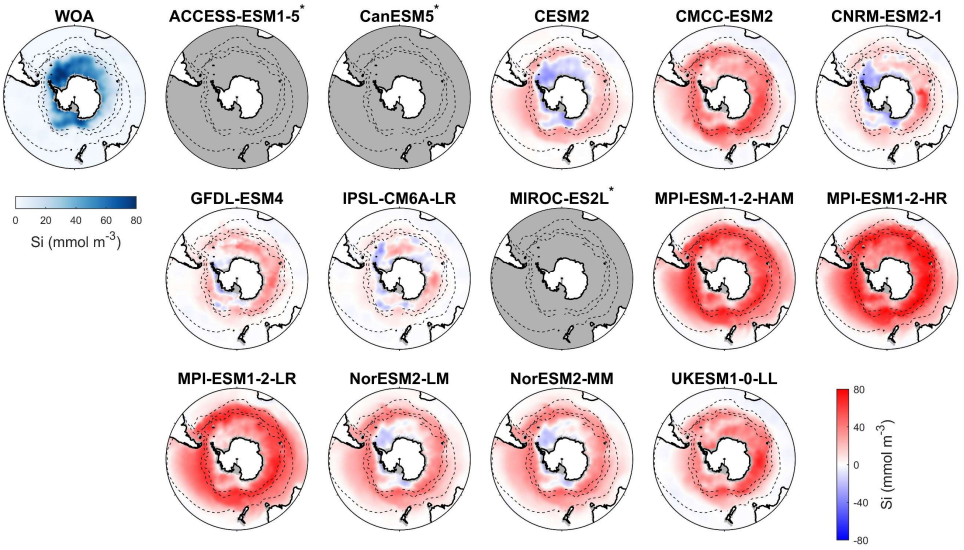


Figure 5: Observed surface silicate concentrations from WOA in DJF and spatial biases of surface silicate concentrations for 14 CMIP6 models (model silicate – WOA silicate) in DJF for the Southern Ocean (>30°S). Models with unavailable silicate are labelled with *.

Among the 11 CMIP6 models with available silicate data, IPSL-CM6A-LR is the best-performing model for representing silicate distribution across the Southern Ocean. It has the lowest MBE (1.50 mmol/m^3, compared to the observation of 12.65 mmol/m^3), an SSD closest to 1 (1.04), the lowest RMSD (0.37), and the highest CC (0.94) (Fig. S3 and Table S3), making it the most reliable model for simulating silicate concentrations. Following IPSL-CM6A-LR, the CNRM-ESM2-1, GFDL-ESM4, and CESM2 models also show relatively good performance, although their statistical metrics are not as strong as IPSL-CM6A-LR. The remaining models, CMCC-ESM2, MPI-ESMs, NorESMs, and UKESM1-0-LL, are less reliable due to their large bias errors, which suggests significant discrepancies in their silicate simulations.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

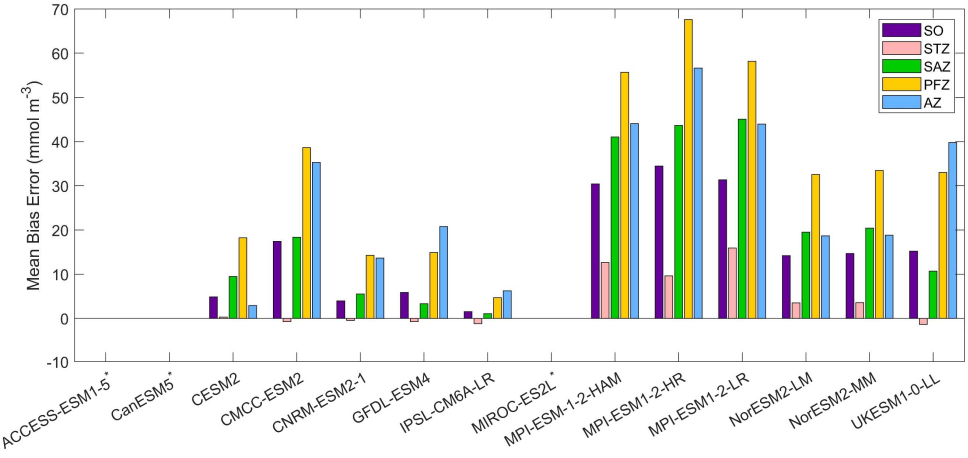


Figure 6: The mean bias errors in surface silicate concentrations for the Southern Ocean (SO), the subtropical zone (STZ), the subantarctic zone (SAZ), the polar front zone (PFZ), and the Antarctic zone (AZ) in DJF. Models with unavailable silicate are labelled with *.

Dissolved iron concentrations in the Southern Ocean are generally low in open waters and higher in coastal regions, as observed from GEOTRACES data (Fig. 7). Among the CMIP6 models analysed, most tend to underestimate dissolved iron concentrations, with MBEs ranging from -0.06 to -0.28 μmol/m^3 compared to the observed mean of 0.57 μmol/m^3 (Fig. 8). The only exceptions are ACCESS-ESM1-5, NorESMs and UKESM1-0-LL, which overestimate the dissolved iron in Southern Ocean surface waters, except in some coastal regions around Antarctica (Fig. 7). No strong correlation is found between the spatial deviation of chlorophyll and dissolved iron concentrations across the models, despite iron limitation being a key factor controlling phytoplankton growth (Tagliabue et al., 2017). For example, the three MPI-ESM models simulate low dissolved iron and have a high half-saturation coefficient for iron (3.6 μmol/m^3), yet they significantly overestimate chlorophyll concentrations (Fig. 1). Conversely, NorESM2-LM and NorESM2-MM models simulate higher dissolved iron concentrations in the polar front zone and subantarctic zone, but their chlorophyll levels remain low in these regions. CanESM5 is excluded from dissolved iron comparison because it does not explicitly simulate dissolved iron; instead, iron limitation on phytoplankton growth is parameterised through a functional relationship with nitrate.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

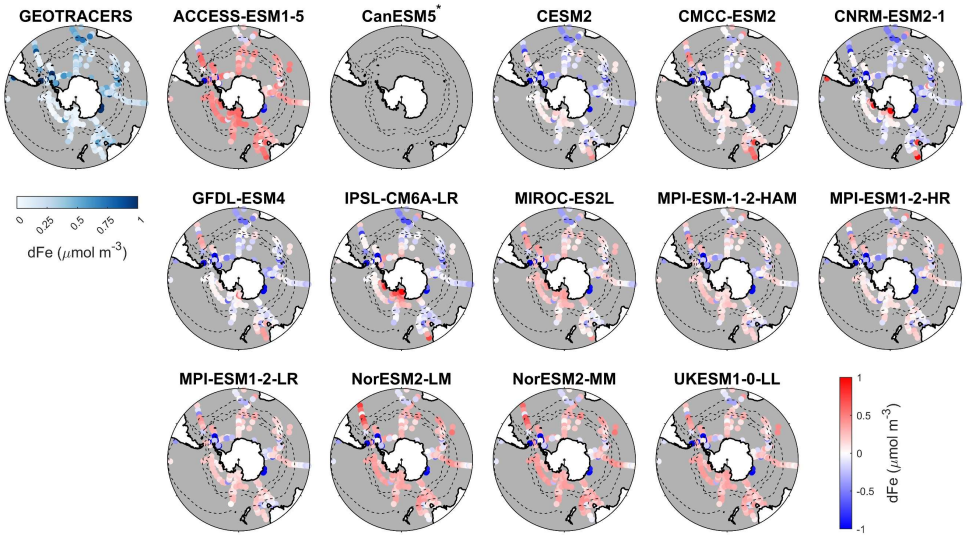


Figure 7: Observed surface dissolved iron concentrations from GEOTRACES and spatial biases of surface dissolved iron concentrations for 14 CMIP6 models (model dissolved iron – GEOTRACES dissolved iron) for the Southern Ocean (>30°S). Due to the limited availability of observed data, all CMIP6 model outputs were regridded to match the spatial resolution of the observational dataset, ensuring a consistent grid for comparison. Models with unavailable dissolved iron are labelled with *.

All models exhibit poor statistics for dissolved iron, with SSD less than 0.5, RMSD larger than 0.9, and CC values lower than 0.4 (Fig. S4 & Table S4). Most models, except the CMCC-ESM2, CNRM-ESM2-1, IPSL-CM6A-LR, MIROC-ES2L, and UKESM1-0-LL models, have negative CC values, indicating a distribution trend opposite to observations. Among them, the MIROC-ES2L model performs relatively better, with an MBE of -0.06 µmol/m^3 (the fourth lowest among models), an SSD closest to 1 (0.19), the lowest RMSD (0.91), and the largest positive CC (0.40) (Fig. S4 and Table S4). Despite these findings, the evaluation of dissolved iron simulation remains uncertain due to the limited availability of observational data, making it difficult to draw definitive conclusions about model performance in this regard.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

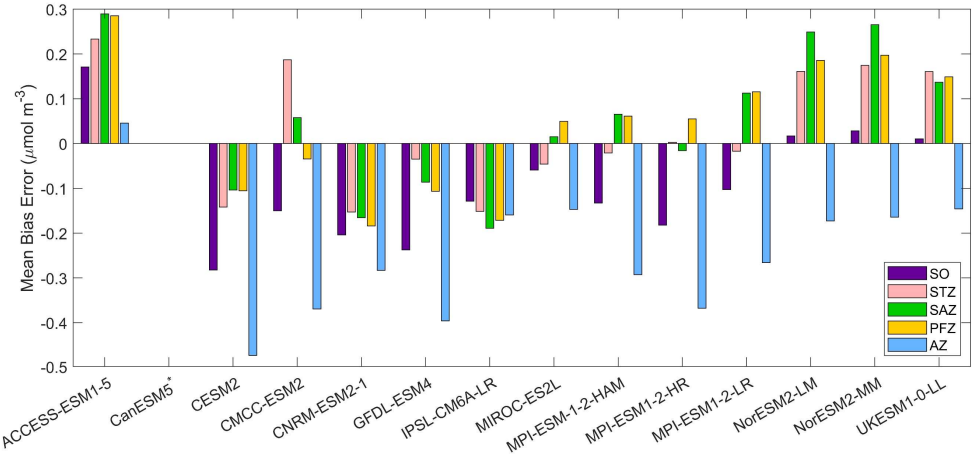


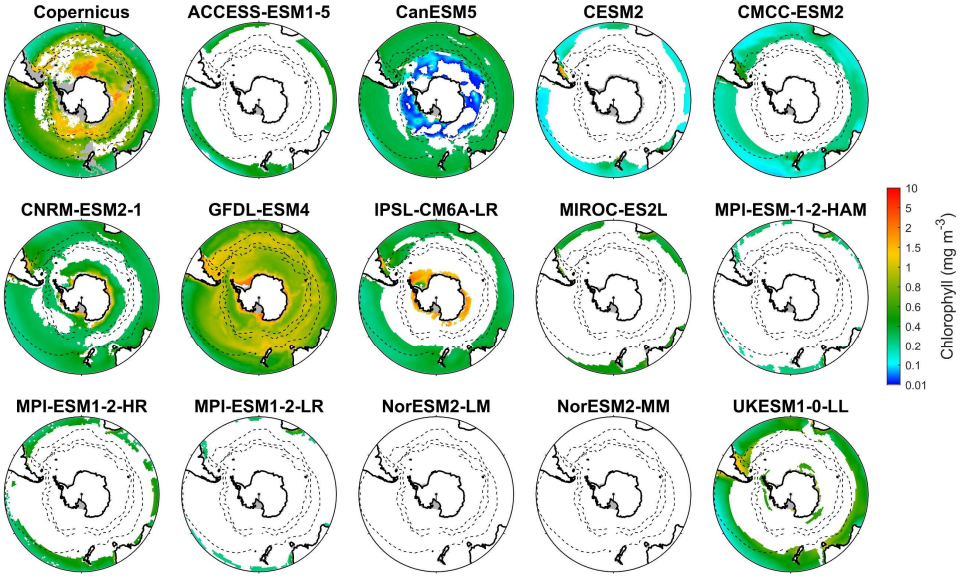
Figure 8: The mean bias errors in surface dissolved iron concentrations for the Southern Ocean (SO), the subtropical zone (STZ), the subantarctic zone (SAZ), the polar front zone (PFZ), and the Antarctic zone (AZ) in DJF. Models with unavailable dissolved iron are labelled with *.

3.2 Performance of DCMs

The observational data from Copernicus indicates that DCMs are widespread across approximately 85% of the Southern Ocean in austral summer (Fig. 9). Their occurrence frequency is lower in the SAZ (below 70%) but exceeds 90% in other regions. Areas without DCMs are primarily located south of Australia, southwest of Chile, and in the Weddell and Ross Seas and surrounding waters. CMIP6 models exhibit varying performance in simulating DCMs. GFDL-ESM4 has DCM occurrence frequency close to 100% across the Southern Ocean (Fig. 10), while the CanESM5 model simulates a DCM frequency similar to observations, but its spatial distribution deviates from observations where we find no DCMs in the Antarctic waters. CNRM-ESM2-1 simulates a high occurrence of DCMs in the STZ and AZ, but a low occurrence in the SAZ and PFZ (Fig. 9). CMCC-ESM2, IPSL-CM6A-LR, and UKESM1-0-LL models simulate DCMs in the STZ but fail to capture them south of the subtropical front (Fig. 9). The ACCESS-ESM1-5, CESM2, MIROC-ES2L, and the three MPI-ESMs models sporadically simulate DCMs in the STZ, resulting in a low overall DCM frequency (<20% for the Southern Ocean). The NorESM2-LM and NorESM2-MM models fail to simulate any DCMs. Among the remaining models, CanESM5, CNRM-ESM2-1, and GFDL-ESM4 exhibit DCM frequencies closest to observations. However, the CanESM5 and CNRM-ESM2-1 models are not considered reliable for representing DCMs due to their poor chlorophyll performance (Figs. 1 and 2), which fails to reflect the actual distribution of the phytoplankton biomass in the water column, as chlorophyll serves as a key indicator of phytoplankton abundance, despite their accurate DCM frequencies. Consequently, GFDL-ESM4

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

model is identified as the best-performing model for DCM simulation, given its strong agreement with both DCM frequency and its chlorophyll distribution.



290

Figure 9: Observed chlorophyll concentration at deep chlorophyll maximum (DCM) depth during DJF for Copernicus (observation) and 14 CMIP6 models in the Southern Ocean (>30°S). The colours in the maps indicate the chlorophyll concentration at DCM depth, while white areas represent regions where no DCM occurred.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

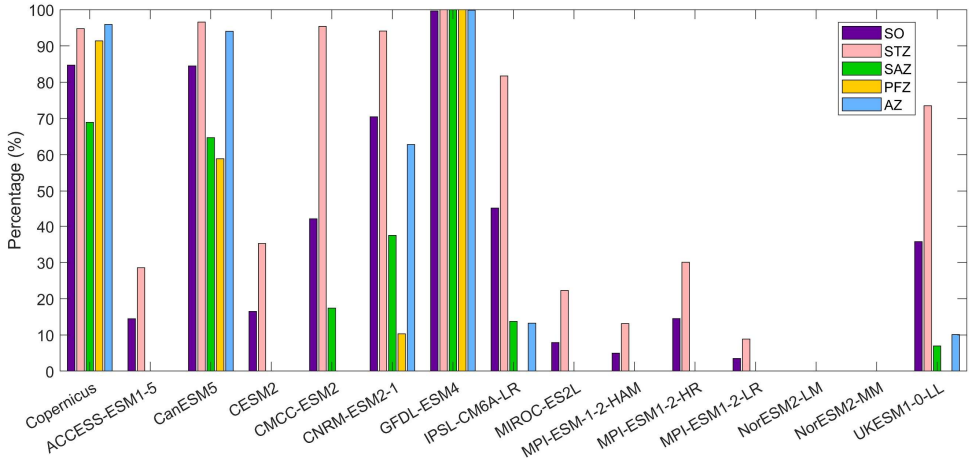


Figure 10: The percentage of DCM occurrence in the Southern Ocean (SO), the subtropical zone (STZ), the subantarctic zone (SAZ), the polar front zone (PFZ), and the Antarctic zone (AZ) in DJF. All percentages in each region are calculated using area-weighted averages.

3.3 Particulate organic carbon (POC)

Observed particulate organic carbon (POC) concentrations in the Southern Ocean are higher in Antarctic coastal waters and lower at low latitudes (Fig. 11). Model simulations diverge markedly from this pattern. CMCC-ESM2, the three MPI-ESM models, and UKESM1-0-LL generally overestimate POC across the basin, apart from CMCC-ESM2's underestimation in the subtropical zone, MPI-ESM-1-2-LR's underestimation south of the subantarctic front, and UKESM1-0-LL's underestimation in the AZ, yield MBEs of 12.1, 10.6, 28.8, 2.4, and 6.7 mg/m^3 (Fig. 12), respectively, versus the observed mean of 70.4 mg/m^3. The overestimated POC concentrations in three MPI-ESMs align with their significantly high simulated chlorophyll concentrations (Fig. 1). In contrast, the remaining models, including ACCESS-ESM1-5, CanESM5, CESM2, CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, MIROC-ES2L, and two NorESM models underestimate the surface POC. CNRM-ESM2-1, GFDL-ESM4, and IPSL-CM6A-LR simulate nearly uniform values (regional means of 50 to 65 mg/m^3), leading to large bias errors near Antarctica but small errors at lower latitudes (Fig. 12). NorESM2-LM and NorESM2-MM overestimate POC in the subtropical zone while strongly underestimating it south of the subtropical front (Fig. 12). ACCESS-ESM1-5, CanESM5, CESM2, and MIROC-ES2L show the largest negative MBEs (40-50 mg/m^3), severely underrepresenting POC, especially at high southern latitudes.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

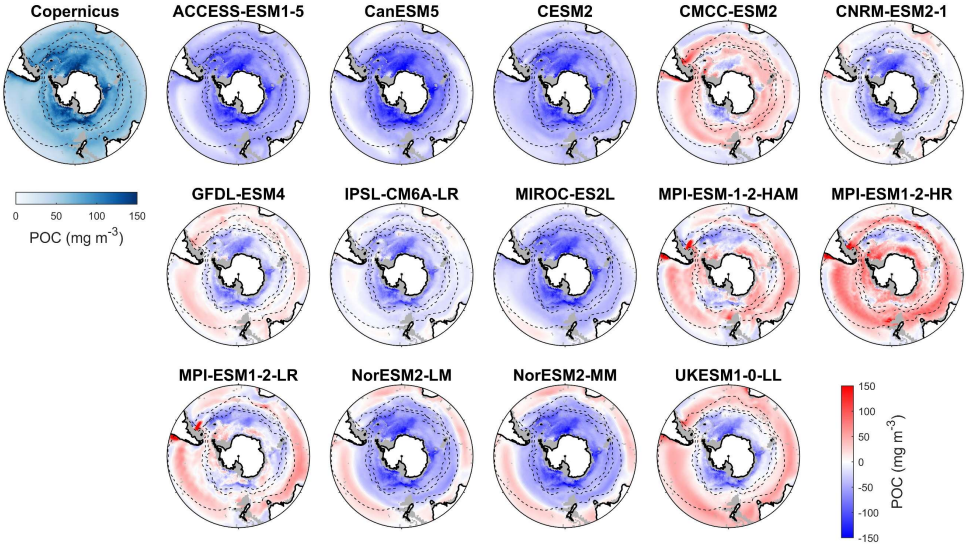


Figure 11: Observed yearly surface POC concentrations from Copernicus and spatial biases of yearly surface POC concentrations for 14 CMIP6 models (model POC – Copernicus POC) for the Southern Ocean (>30°S). POC data in CMIP6 models contains phytoplankton carbon, zooplankton carbon, and detrital organic carbon.

Among 14 CMIP6 models, CMCC-ESM2 and MPI-ESM-1-2-LR have the most realistic simulations. Both have small MBEs (<13 mg/m^3) and SSD closest to 1 (1.22 and 1.16; Fig. S5 and Table S5). The correlation coefficient of CMCC-ESM2 is the highest (0.60), making it the best-performing model for representing POC. CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM-1-2-HAM, and MPI-ESM-1-2-HR show intermediate skill with weaker statistics, whereas ACCESS-ESM1-5, CanESM5, MIROC-ES2L, NorESM2-LM, and NorESM2-MM are unreliable owing to large negative biases, high RMSDs, and negative correlations.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

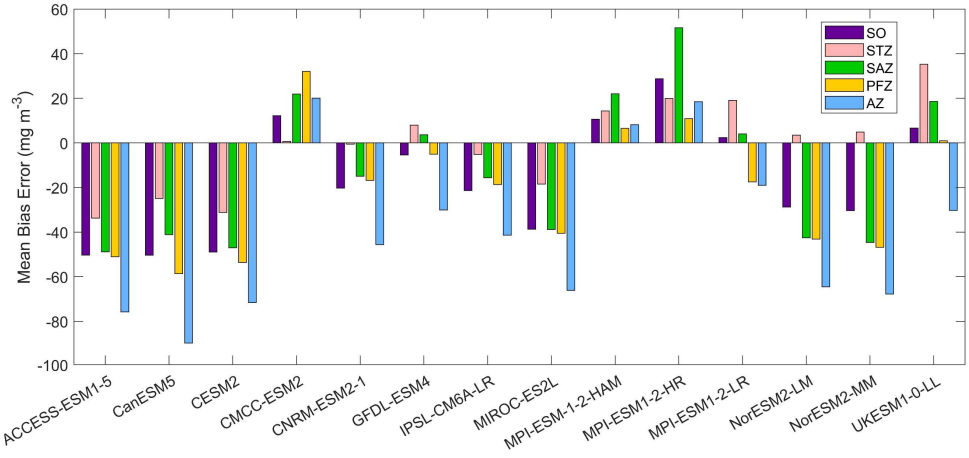


Figure 12: The mean bias errors in yearly surface POC concentrations for the Southern Ocean (SO), the subtropical zone (STZ), the subantarctic zone (SAZ), the polar front zone (PFZ), and the Antarctic zone (AZ). POC data in CMIP6 models contains phytoplankton carbon, zooplankton carbon, and detrital organic carbon.

3.4 Model ranking

Based on the statistical evaluation of surface chlorophyll, nitrate, silicate, dissolved iron, and POC using MBE, SSD, RMSD, and CC (Sect. 3.1 and 3.3), along with DCM occurrence frequency (Sect. 3.2), we computed a ranking of each variable and an overall ranking for each model following the methodology described in Sect. 2.4. The results are presented in Fig. 13 as a heat map. IPSL-CM6A-LR ranks the highest overall, placing within the top two models for all variable rankings except for POC, for which it ranks fifth. GFDL-ESM4 follows closely, achieving top three rankings in all variables except dissolved iron, where it ranks eighth. CMCC-ESM2 demonstrates strong performance in chlorophyll, DCM and POC (all rank in the top three), but its lower scores for nutrient variables reduce its overall ranking to third. UKESM1-0-LL ranks fourth, supported by its relatively balanced performance across all metrics. CNRM-ESM2-1, which also incorporates the PISCES-v2 biogeochemical model (as in IPSL-CM6A-LR) also ranks fourth, with performance slightly below that of IPSL-CM6A-LR across most variables. MIROC-ES2L, despite having the highest ranking for dissolved iron, ranks sixth due to weak performance in other variables. The three MPI-ESM models, all coupled with HAMOCC6, occupy the lowest three positions, despite showing reasonable POC estimates. Models such as NorESM2-LM, CESM2, NorESM2-MM, ACCESS-ESM1-5, and CanESM5 fall into the middle tier, ranking from seventh to eleventh. In summary, IPSL-CM6A-LR and GFDL-ESM4 emerge as the most robust models for simulating biogeochemical processes in the Southern Ocean, with consistently performance across a range of biogeochemical parameters.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

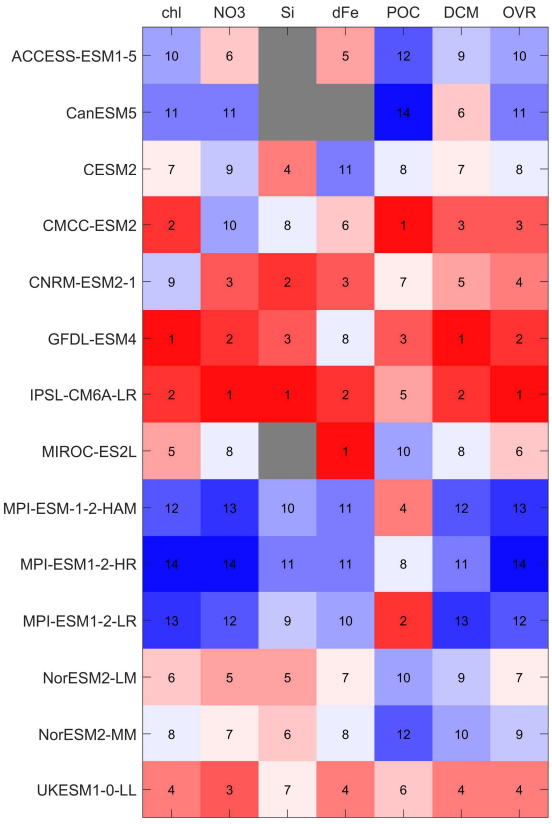


Figure 13: Heat-map of performance ranks for 12 CMIP6 models. Columns list the evaluated variables—surface chlorophyll (Chl), nitrate (NO3), silicate (Si), dissolved iron (dFe), particulate organic carbon (POC), deep-chlorophyll-maximum metrics (DCM)—and an overall score (OVR, the mean of the six individual ranks). Rows list the models. Box colours and overlaid numbers give the rank for each model–variable pair (1 = best, higher numbers = poorer performance): reds indicate higher ranks, blues lower ranks, and grey boxes indicate variables not available for that model.

4 Discussion

4.1 Vertical structure of carbon

Most CMIP6 models perform relatively well in simulating surface chlorophyll in the Southern Ocean, but they exhibit only moderate skill in representing surface particulate organic carbon (POC). In contrast, the majority of models struggle to accurately simulate the deep chlorophyll maxima (DCMs), which is crucial for capturing the vertical structure of chlorophyll

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

distributions. As discussed in Sect 3.2, models such as CanESM5, CNRM-ESM2-1, and GFDL-ESM4 reproduce the horizontal frequency patterns of DCMs reasonably well. However, when surface chlorophyll performance is also considered, GFDL-ESM4 emerges as the only model that satisfactorily represents both surface chlorophyll concentrations and DCM frequency. This finding suggests that most CMIP6 models face challenges in simulating the vertical structure of chlorophyll, as well as POC distributions.

To compare the vertical structure of chlorophyll and POC between models and observations, we integrated their concentrations over the top 100m of the water column, where the majority of primary production occurs (Henley et al., 2020; Arrigo et al., 2008). Unlike the surface chlorophyll and POC, which are generally close to observations, the vertically integrated chlorophyll and POC in the upper 100m are significantly underestimated by most CMIP6 models, except chlorophyll in MPI-ESM-1-2-HR and POC in CMCC-ESM2, both of which are overestimated (Fig. 14).

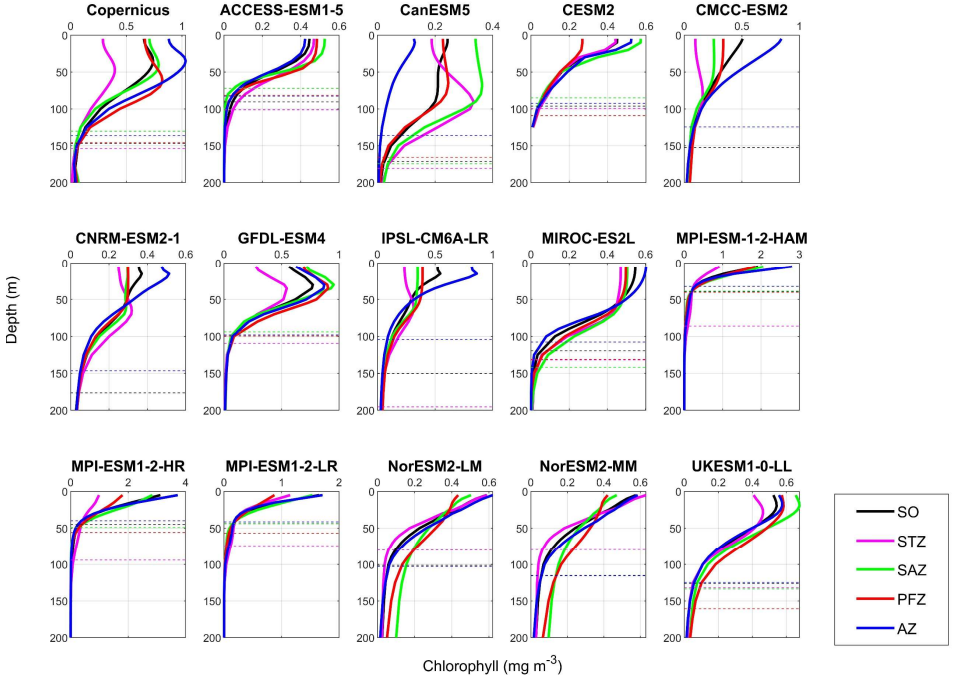


Figure 14: Mean vertical profiles of chlorophyll during DJF (December–January–February) across the Southern Ocean (SO) and its subregions: the Subtropical Zone (STZ), Subantarctic Zone (SAZ), Polar Frontal Zone (PFZ), and Antarctic Zone (AZ), based on observations (Copernicus) and 14 CMIP6 models. Solid lines represent chlorophyll profiles in different regions, while dashed lines indicate the threshold depth of chlorophyll, defined as the depth at which chlorophyll concentration reaches 10% of the maximum value.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

The underestimation of vertically integrated chlorophyll in the top 100 m ranges from −63% for CESM2 to −16% for GFDL-ESM4 (Fig. 14) and is influenced by both surface chlorophyll concentrations and the vertical structure of the water column. For example, ACCESS-ESM1-5, CESM2, NorESM2-LM, and NorESM2-MM exhibit similar vertical chlorophyll profiles, characterised by low surface concentrations, almost no deep chlorophyll maxima (DCMs), and shallow chlorophyll threshold depth (CTD; defined as the depth where chlorophyll falls to 10% of the maximum), resulting in underestimations exceeding 50% (Fig. 14). In contrast, MPI-ESM-1-2-HAM and MPI-ESM1-2-LR show high surface chlorophyll levels but extremely shallow CTD (<50 m), leading to low vertically integrated chlorophyll. A third pattern is found in CanESM5, CMCC-ESM2, CNRM-ESM2-1, IPSL-CM6A-LR, MIROC-ES2L, and UKESM1-0-LL, which simulate appropriate threshold depths (~150 m) and some occurrence of DCMs, but their low surface chlorophyll leads to insufficient primary production in the water column. GFDL-ESM4 demonstrates a vertical structure most similar to observations, with a slightly shallower threshold depth, resulting in only an 18% underestimation of integrated chlorophyll. While CMIP6 models vary widely in their simulation of surface chlorophyll concentrations and generally manage to control these levels they largely lack the capability to accurately simulate the vertical structure of chlorophyll, including both DCMs and CTD.

The vertical structure of chlorophyll and the formation of DCMs are influenced by various environmental and biological factors. DCMs are primarily driven by photoacclimation, as the carbon to chlorophyll (C:Chl) ratio decreases from values exceeding 100 g:g at the surface to below 50 at the base of euphotic layer (Marañón et al., 2021; Boyd et al., 2024). Consequently, the poor representation of DCMs in ACCESS-ESM1-5 (with its coupled biogeochemical component WOMBAT), the MPI-ESM models (coupled with HAMOCC6), and the NorESM models (coupled with HAMOCC) is likely due to their use of fixed C:Chl ratio (Oke et al., 2013; Ilyina et al., 2013; Tjiputra et al., 2020). This simplification prevents the models from capturing photoacclimation processes, thereby limiting their ability to simulate realistic DCM structures. Additionally, phytoplankton functional types (PFTs) significantly influence the vertical distribution of chlorophyll. For instance, siliceous diatoms, which account for approximately 75% of primary production in the Southern Ocean (Crosta et al., 2005), are not represented in ACCESS-ESM1-5 and CanESM5. This omission leads to the underestimation of chlorophyll, particularly in the Antarctic zone (Fig. 14). CMIP6 models represent no more than three PFTs, typically small phytoplankton, diatoms, and diazotrophs. In contrast, observational studies, such as Yingling et al. (2025), identify at least five ecologically significant PFTs in the Southern Ocean, including *Synechococcus*, *Picoeukaryotes*, nanoplankton, diatoms, and microplankton. This simplification of PFT diversity in CMIP6 models likely contributes to inaccurate chlorophyll estimates and unrealistic vertical chlorophyll structures. Moreover, the vertical structure of chlorophyll is linked to the mixed layer depth (MLD), which modulates nutrient supply (Durán-Campos et al., 2019; Zampollo et al., 2023). Our analysis indicates a positive correlation between the CTD and MLD (Fig. S7a; R^2=0.24, p=0.075), suggesting that deep mixing enables phytoplankton to extend further into the water column while maintaining detectable concentrations (Mignot et al., 2014). Conversely, the integrated chlorophyll within the upper 100m shows a negative correlation with MLD (Fig. S7b;

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

R^2=0.23, p=0.082), likely due to reduced light availability and dilution effects associated with deeper mixed layers (Behrenfeld and Boss, 2006).

Furthermore, the occurrence frequency of DCMs exhibits a Gaussian-like relationship with MLD (Fig. S7c; R^2=0.42), peaking at MLD of 31 m. When the MLD is excessively shallow, nutrient replenishment to the euphotic zone is limited, inhibiting phytoplankton growth below the surface, thereby reducing the likelihood of DCM formation (Letelier et al., 2004). Conversely, when the MLD becomes too deep, light availability at depth decreases to levels insufficient for sustaining phytoplankton biomass accumulation, which similarly suppresses DCM development (Mignot et al., 2014). Thus, the observed distribution reflects a balance between light limitation from above and nutrient supply from below, a mechanism well-documented in earlier studies (Cullen, 1982; Fennel and Boss, 2003).

Similar to chlorophyll, the vertical distribution of POC is significantly underestimated by most CMIP6 models (Fig. 14). In this study, POC consists of four carbon pools: phytoplankton carbon, zooplankton carbon, detrital organic matter carbon, and heterotrophic bacteria carbon. Observational estimates suggest an approximate partitioning of these pools in the Southern Ocean at 20% phytoplankton, 37% zooplankton, 33% detritus, and 10% heterotrophic bacteria (Yingling et al., 2025; Liu et al., 2025; Yang et al., 2022). However, the allocation among POC components varies across CMIP6 models.

Most models simulate integrated phytoplankton carbon reasonably well, with values comparable to observations, except for MPI-ESM1-2-HR and UKESM1-0-LL, which show significant overestimation (Table S6). The general agreement in phytoplankton carbon across the models contrasts sharply with the widespread underestimation of integrated chlorophyll, suggesting that models may be applying high C:Chl ratio below the surface. Integrated zooplankton carbon is substantially underestimated (Table S6), likely due to oversimplified zooplankton physiology and trophic structure. Only a few models, such as CMCC-ESM2, CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, and UKESM1-0-LL, include more than two zooplankton types, and many may apply low growth and grazing efficiency (Rohr et al., 2023), contributing to low biomass estimates.

Detrital organic carbon shows the widest range of discrepancies. For example, CMCC-ESM2 overestimates detritus by more than threefold compared to observations, while GFDL-ESM4 and three MPI-ESMs simulate less than 2% of the observed values (Table S7). In contrast, CNRM-ESM2-1, IPSL-CM6A-LR, and UKESM1-0-LL provide detritus concentrations that align well with observations (Table S7). The success of CNRM-ESM2-1 and IPSL-CM6A-LR is attributed to their use of the PISCES-v2 model, which offers a detailed carbon pool structure, including small and large size particulate organic detritus with size-dependent sinking rates and complex exchanges with dissolved organic carbon (DOC) (Aumont et al., 2015). UKESM1-0-LL's high detritus levels may result from elevated phytoplankton and zooplankton concentrations, potentially driven by a high C:Chl ratio and the absence of a DOC pool (Sellar et al., 2019). In contrast, the low detritus levels in GFDL-ESM4 and MPI-ESMs may result from a lack of exchange between DOC and particulate detritus (Stock et al., 2020; Ilyina et al., 2013). This structural limitation can result in unrealistically low detritus levels, especially under strong remineralisation conditions, and when the exudation and residual matter from phytoplankton and zooplankton are directed

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

primarily into the DOC pool rather than contributing to particulate detritus pool. While NorESMs share the same biogeochemical framework as the MPI-ESMs, their relatively higher detritus levels may stem from parameter tuning specific to HAMOCC (Tjiputra et al., 2020). CESM2 adopts a more simplified approach, lacking an explicit detritus tracer. This means there is no time lag between surface production and deep remineralisation, leading to unrealistic vertical carbon fluxes. Other models, such as ACCESS-ESM1-5, CanESM5, and MIROC-ES2L, employ a basic NPZD (nutrient-phytoplankton-zooplankton-detritus) framework, which simplifies the marine food web and organic carbon cycling (Oke et al., 2013; Zahariev et al., 2007; Hajima et al., 2020).

Only CMCC-ESM2 and GFDL-ESM4 simulate an explicit bacteria pool, and their integrated bacterial carbon concentrations are reasonably consistent with observational estimates (Table S7). Including bacteria is important in biogeochemical models, as it allows dynamic regulation of remineralisation and other microbial processes based on bacteria biomass. Furthermore, bacteria contribute significantly to carbon export, highlighting their importance as a key component for future model development and improvement.

4.2 Model components and their performance

The performance of CMIP6 models in simulating key biogeochemical variables such as chlorophyll, nitrate, silicate, dissolved iron, POC and DCMs is jointly determined by the complexity of the biogeochemical (BGC) module, the adopted parameterisations of key biogeochemical processes, and the resolution of their coupled ocean and atmosphere model.

Among these, the complexity of the BGC module is the most crucial factor. Key aspects include the representation of phytoplankton functional types (PFTs), stoichiometry flexibility, and nutrient uptake and regeneration schemes. Models that incorporate multiple PFTs, particularly those distinguishing between diatoms and non-diatom phytoplankton, tend to outperform models with a single phytoplankton type in simulating chlorophyll and overall biogeochemical patterns (Fig. 15a; $p<0.01$). In contrast, the inclusion of diazotrophs has a limited impact on chlorophyll performance, as nitrate is rarely limiting in the Southern Ocean (Fig. 15b; $p=0.17$).

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

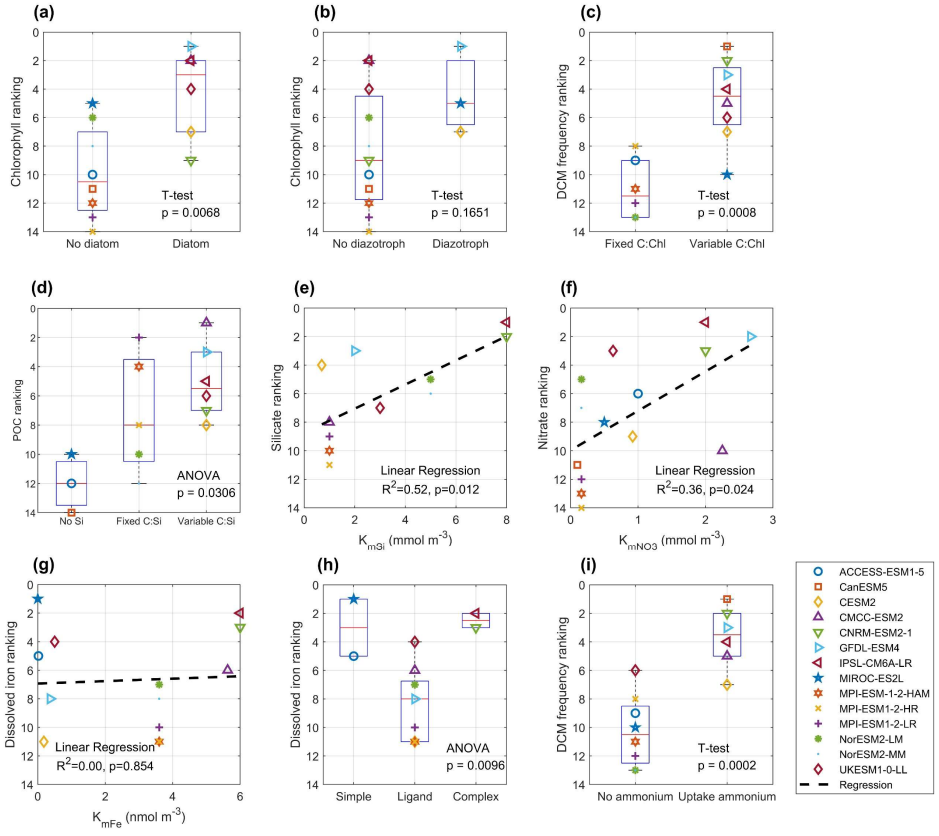


Figure 15: Panels show statistical relationships between model rankings and key biogeochemical descriptors: (a) surface chlorophyll ranking vs. inclusion of diatom; (b) surface chlorophyll ranking vs. inclusion of diazotroph; (c) DCM frequency ranking vs. use of a variable C:Chl ratio; (d) POC ranking vs. presentation of silica cycling (presence of an explicit Si pool or variable C:Si ratio); (e) silicate ranking vs silicate half-saturation coefficient (K_{mSi}); (f) nitrate ranking vs. nitrate half-saturation coefficient (K_{mNO3}); (g) dissolved iron ranking vs. iron half-saturation coefficient (K_{mFe}); (h) dissolved iron ranking vs. iron chemistry complexity (simple-no ligand, simple ligand, or complex ligand scheme); (i) DCM frequency ranking vs model ability to assimilate ammonium for photosynthesis. (a), (b), (c), (i) are performed using T-test, (d) and (h) are performed using ANOVA (Analysis of Variance), (e), (f), (g) are performed using linear regression. Tests applied: two-sample t-tests for (a), (b), (c), (i); one-way ANOVA for (d), (h); linear regression for (e)–(g). Each point (colour/shape) represents a CMIP6 model, and dashed lines indicate regression fits where relevant. Corresponding P-values and R^2 statistics (for regressions) are displayed on each panel.

Cellular plasticity (stoichiometry) plays a vital role in regulating nutrient uptake and the cellular elemental composition under variable environmental conditions. Most models employ fixed carbon:nitrogen:phosphorus (C:N:P) ratios consistent with the Redfield Ratio, while carbon:iron ratios are generally dynamic. However, carbon:chlorophyll and carbon:silicate

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

ratios vary across models. A dynamic carbon:chlorophyll ratio significantly improves the simulation of DCM (Fig. 15c; $p<0.01$), as mentioned in Sect. 4.1, while a variable carbon:silicate ratio enhances POC representation (Fig. 15d; $p=0.03$), especially given the dominance of diatoms Southern Ocean primary production (Crosta et al., 2005).

Phytoplankton growth in models is typically limited by light and nutrient availability, often represented using Michaelis-Menten kinetics (Michaelis and Menten, 1913). However, our analysis did not reveal a clear relationship between model performance in simulating surface chlorophyll or DCMs and specific light or nutrient uptake parameters, such as initial PI (production-irradiance) slope or half-saturation coefficients for nitrate, silicate, and dissolved iron. This suggests that chlorophyll distribution is governed by a complex interplay of environmental drivers rather than any single parameter. In contrast, nutrient concentrations are more directly influenced by process parameterisation. For example, higher silicate half-saturation coefficients (e.g. 8 mmol/m^3 in PISCES-v2, as used in CNRM-ESM2-1 and IPS-CM6A-LR) spear to improve silicate simulations (Fig. 15e; $R^2=0.52$, $p=0.01$) (Nelson et al., 2001). Similarly, nitrate half-saturation coefficients in the range of 1–3 mmol m^{-3} tend to yield better agreement with observations (Fig. 15f; $R^2=0.36$, $p=0.02$) (Eppley et al., 1969). For dissolved iron, no clear correlation was found between model performance and the half-saturation coefficient (Fig. 15g; $R^2=0.00$, $p=0.86$). The complexity of the iron cycle contributes to variability in simulated dissolved iron performance (Fig. 15h; $p<0.01$). Models with more advanced iron chemistry, such as PISCES-v2 (BGC model coupled in CNRM-ESM2-1 and IPSL-CM6A-LR), which includes strong and weak ligands, and five iron forms (free Fe(II), Fe(III), Fe(III) bounded to strong and weak ligands, and particulate iron) tend to simulate dissolved iron more accurately than those with simple iron complexation (Tagliabue et al., 2023). In contrast, models with simple iron complexation schemes do not show strong ability to simulate better iron concentrations than a simple iron model, which only contains basic iron processes such as scavenging. These inconsistencies are likely due to the limited spatial and temporal coverage of iron observations, which hinders robust evaluation and may mask the benefits of advanced iron cycling mechanisms. Additionally, the utilisation of ammonium appears to promote the formation of DCMs (Fig. 15i; $p<0.01$), as ammonium-primarily produced through remineralisation-is more readily and rapidly assimilated by phytoplankton than nitrate. This is due to its lower energy and electron requirements for incorporation into cellular biomass. Consequently, substantial ammonium production by heterotrophic bacteria in the subsurface can enhance phytoplankton growth and contribute to the development of DCMs (Boyd et al., 2024).

We also found that the resolution of the ocean component in ESMs can influence the performance of simulated biogeochemical variables. For example, MPI-ESM1-2-HR and MPI-ESM1-2-LR, both coupled with the same biogeochemical model (HAMOCC6), differ significantly in ocean resolution 0.4° vs 1.5°, respectively, and show notable differences in biogeochemical performance. The mean surface chlorophyll concentration in austral summer is 2.37 mg/m^3 in MPI-ESM1-2-HR, compared to 1.35 mg/m^3 in MPI-ESM1-2-LR which is closer to the Copernicus chlorophyll dataset. These discrepancies may arise from resolution-induced differences in ocean circulation and physical conditions, which influence nutrient availability, light penetration, and phytoplankton dynamics. In contrast, variations in atmospheric model resolution appear to have a limited impact on ocean biogeochemistry. For instance, NorESM2-MM and NorESM2-LM,

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

which use the same ocean biogeochemical model (HAMOCC) but differ in atmospheric resolution (2° vs 1°), exhibit nearly identical biogeochemical outcomes such as mean austral summer surface chlorophyll concentrations of 0.56 and 0.55 mg/m^3, respectively. These findings suggest that while higher ocean resolution can improve the realism of physical processes affecting biogeochemical simulations, it does not necessarily guarantee better biogeochemical performance.

4.3 Avenues for improvement in biogeochemical representation

This study provides a comparative assessment of several ocean biogeochemical indicators for 14 CMIP6 ESMs over the Southern Ocean. Although some models performed adequately, there remain several key directions for future improvements:

- The representation of key biogeochemical processes in most BGC models remains simplified or parameterised based on limited observations. For instance, differences in the phytoplankton functional types (PFTs), elemental composition (fixed or variable stoichiometry), and nutrient uptake parameterisation contribute to model divergence. Future models should incorporate a more complex marine food web, and more dynamic parameterisations informed by field and laboratory experiments, especially under Southern Ocean specific conditions.
- As the key factor controlling the Southern Ocean primary production, iron cycles and their representations remain poor in most models, compared to limited iron sampled data. Improvements in the simulation of iron sources (e.g., dust deposition, sediment resuspension), bioavailability (i.e., more complex iron chemistry module such as including iron-binding ligands), and biological recycling are essential to help reduce the bias in simulated chlorophyll.
- Most models lack a good representation of the vertical structure of chlorophyll and biomass. For example, come models have discrepancies in mixed layer depth and other physical properties simulation, which influences nutrient supply. There is also an oversimplified remineralisation by heterotrophic bacteria, and lack of diversity of PFTs. Future efforts could expand the model structure to capture these ecological dynamics, which are particularly important in determining vertical profiles and export efficiency for biomass.
- Observational constraints remain limited, especially for subsurface variables such as DCMs, dissolved iron, and POC and its classification. Future work should prioritise the integration of additional in situ datasets to validate and improve model parameterisations. Ensemble data assimilation or machine learning approaches could also be explored for model tuning.

5 Conclusion

This study evaluated the performance of key biogeochemical variables, including austral summer surface chlorophyll and deep chlorophyll maxima (DCMs), nitrate, silicate, dissolved iron, and annual particulate organic carbon (POC) across 14 CMIP6 models in the Southern Ocean (south of 30°S). The results reveal substantial variability in model skill. While some

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

models demonstrated strong performance, others showed significant over- or underestimations. Among them, GFDL-ESM4 was the most effective in reproducing surface chlorophyll and DCM features, while IPSL-CM6A-LR excelled in simulating nutrient distribution, particularly nitrate and silicate. MIROC-ES2L performed best for dissolved iron, and CMCC-ESM2 provided the most accurate representation of POC. Based on aggregated performance across all variables, the top five models for simulating Southern Ocean biogeochemistry were IPSL-CM6A-LR, GFDL-ESM4, CMCC-ESM2, UKESM1-0-LL, and CNRM-ESM2-1. Our analysis highlights a common limitation across CMIP6 models: the underrepresentation of vertical biogeochemical structures, including DCMs and subsurface POC distributions. Additionally, spatial mismatches and persistent biases, particularly for dissolved iron and POC, underscore the need for targeted model improvements. Overall, this study not only provides a comprehensive evaluation of model performance for key biogeochemical variables but also offers insights into areas requiring refinement. These insights can guide future model development and support more informed model selection. Enhancing the representation of biogeochemical processes in Earth system models is essential for improving projections of the Southern Ocean's role in the global carbon and nutrient cycles under ongoing climate change.

Code availability

All codes for regridding datasets and data analysis are available at https://github.com/mingcheng7/Evaluation-CMIP6-historical.

Data availability

Raw CMIP6 used in this study are available on the Earth System Grid Federation (ESGF) Nodes for the CMIP6 Archive at https://esgf.github.io/nodes.html (Cinquini et al., 2014). Copernicus Global Ocean 3D Chlorophyll-a Concentration, Particulate Backscattering coefficient and Particulate Organic Carbon Product can be accessed at https://doi.org/10.48670/moi-00046 (Sauzède et al., 2016). The World Ocean Atlas (WOA) 2018 data can be accessed at https://www.ncei.noaa.gov/access/world-ocean-atlas-2018/ (Garcia et al., 2019). The GEOTRACES dissolved iron data can be accessed at https://www.bodc.ac.uk/geotraces/data/historical/ (Tagliabue et al., 2012).

Author contribution

MC, NM, and MJE contributed to the conceptualisation of the study and participated in writing and revising the manuscript. MC was responsible for data collection, analysis and figure preparation. NM and MJE provided supervision and guidance throughout the project. All authors reviewed and approved the final version of the manuscript.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF.

Financial support

MC was supported by an Australian National University PhD scholarship. MC and MJE were supported by the Australian Research Council Special Research Initiative, Australian Centre for Excellence in Antarctic Science (Project Number SR200100008). NM was supported by the Australian Research Council Centre of Excellence for the Weather of the 21st Century (CE230100012).

References

Ackermann, L., Rackow, T., Himstedt, K., Gierz, P., Knorr, G., and Lohmann, G.: A comprehensive Earth system model (AWI-ESM2.1) with interactive icebergs: effects on surface and deep-ocean characteristics, Geoscientific Model Development, 17, 3279-3301, 10.5194/gmd-17-3279-2024, 2024.

Arrigo, K. R., Van Dijken, G. L., and Bushinsky, S.: Primary production in the Southern Ocean, 1997–2006, Journal of Geophysical Research: Oceans, 113, 10.1029/2007jc004551, 2008.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies, Geoscientific Model Development, 8, 2465-2513, 10.5194/gmd-8-2465-2015, 2015.

Baldry, K., Strutton, P. G., Hill, N. A., and Boyd, P. W.: Subsurface chlorophyll-a maxima in the Southern Ocean, Frontiers in Marine Science, 7, 671, 10.3389/fmars.2020.00671, 2020.

Beadling, R. L., Russell, J. L., Stouffer, R. J., Goodman, P. J., and Mazloff, M.: Assessing the Quality of Southern Ocean Circulation in CMIP5 AOGCM and Earth System Model Simulations, Journal of Climate, 32, 5915-5940, 10.1175/jcli-d-19-0263.1, 2019.

Behrenfeld, M. J. and Boss, E.: Beam attenuation and chlorophyll concentration as alternative optical indices of phytoplankton biomass, Journal of Marine Research, 64, 2006.

Blain, S., Quéguiner, B., Armand, L., Belviso, S., Bombled, B., Bopp, L., Bowie, A., Brunet, C., Brussaard, C., Carlotti, F., Christaki, U., Corbière, A., Durand, I., Ebersbach, F., Fuda, J.-L., Garcia, N., Gerringa, L., Griffiths, B., Guigue, C., Guillerm, C., Jacquet, S., Jeandel, C., Laan, P., Lefèvre, D., Lo Monaco, C., Malits, A., Mosseri, J., Obernosterer, I., Park, Y.-H., Picheral, M., Pondaven, P., Remenyi, T., Sandroni, V., Sarthou, G., Savoye, N., Scouarnec, L., Souhaut, M., Thuiller, D., Timmermans, K., Trull, T., Uitz, J., Van Beek, P., Veldhuis, M., Vincent, D., Viollier, E., Vong, L., and Wagener, T.:

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Effect of natural iron fertilization on carbon sequestration in the Southern Ocean, Nature, 446, 1070-1074, 10.1038/nature05700, 2007.
Böning, C. W., Dispert, A., Visbeck, M., Rintoul, S. R., and Schwarzkopf, F. U.: The response of the Antarctic Circumpolar Current to recent climate change, Nature Geoscience, 1, 864-869, 10.1038/ngeo362, 2008.
Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., De Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J. L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M. A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J. Y., Guenet, B., Guez, L., E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J. B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, Journal of Advances in Modeling Earth Systems, 12, 10.1029/2019ms002010, 2020.
Boyd, P. W.: The role of iron in the biogeochemistry of the Southern Ocean and equatorial Pacific: a comparison of in situ iron enrichments, Deep Sea Research Part II: Topical Studies in Oceanography, 49, 1803-1821, 10.1016/S0967-0645(02)00013-9, 2002.
Boyd, P. W. and Ellwood, M. J.: The biogeochemical cycle of iron in the ocean, Nature Geoscience, 3, 675-682, 10.1038/ngeo964, 2010.
Boyd, P. W., Claustre, H., Levy, M., Siegel, D. A., and Weber, T.: Multi-faceted particle pumps drive carbon sequestration in the ocean, Nature, 568, 327-335, 10.1038/s41586-019-1098-2, 2019.
Boyd, P. W., Antoine, D., Baldry, K., Cornec, M., Ellwood, M., Halfter, S., Lacour, L., Latour, P., Strzepek, R. F., Trull, T. W., and Rohr, T.: Controls on Polar Southern Ocean Deep Chlorophyll Maxima: Viewpoints From Multiple Observational Platforms, Global Biogeochemical Cycles, 38, 10.1029/2023gb008033, 2024.
Buesseler, K. O., Lamborg, C. H., Boyd, P. W., Lam, P. J., Trull, T. W., Bidigare, R. R., Bishop, J. K. B., Casciotti, K. L., Dehairs, F., Elskens, M., Honda, M., Karl, D. M., Siegel, D. A., Silver, M. W., Steinberg, D. K., Valdes, J., Van Mooy, B., and Wilson, S.: Revisiting Carbon Flux Through the Ocean's Twilight Zone, Science, 316, 567-570, 10.1126/science.1137959, 2007.
Carranza, M. M. and Gille, S. T.: Southern Ocean wind-driven entrainment enhances satellite chlorophyll-a through the summer, Journal of Geophysical Research: Oceans, 120, 304-323, 10.1002/2014jc010203, 2015.
Cinquini, L., Crichton, D., Mattmann, C., Harney, J., Shipman, G., Wang, F., Ananthakrishnan, R., Miller, N., Denvil, S., and Morgan, M.: The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data, Future Generation Computer Systems, 36, 400-417, doi.org/10.1016/j.future.2013.07.002, 2014.
Cornec, M., Claustre, H., Mignot, A., Guidi, L., Lacour, L., Poteau, A., D'Ortenzio, F., Gentili, B., and Schmechtig, C.: Deep Chlorophyll Maxima in the Global Ocean: Occurrences, Drivers and Characteristics, Global Biogeochemical Cycles, 35, 10.1029/2020gb006759, 2021.
Crosta, X., Romero, O., Armand, L. K., and Pichon, J.-J.: The biogeography of major diatom taxa in Southern Ocean sediments: 2. Open ocean related species, Palaeogeography, Palaeoclimatology, Palaeoecology, 223, 66-92, 10.1016/j.palaeo.2005.03.028, 2005.
Cullen, J. J.: The deep chlorophyll maximum: comparing vertical profiles of chlorophyll a, Canadian Journal of Fisheries and Aquatic Sciences, 39, 791-803, 10.1139/f82-108, 1982.
Cullen, J. J.: Subsurface Chlorophyll Maximum Layers: Enduring Enigma or Mystery Solved?, Annual Review of Marine Science, 7, 207-239, 10.1146/annurev-marine-010213-135111, 2015.
Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D., DuVivier, A., Edwards, J., Emmons, L., Fasullo, J., Garcia, R., and Gettelman, A.: The community earth system model version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001916, 2020.
Deppeler, S. L. and Davidson, A. T.: Southern Ocean Phytoplankton in a Changing Climate, Frontiers in Marine Science, 4, 10.3389/fmars.2017.00040, 2017.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Dunne, J. P., Horowitz, L., Adcroft, A., Ginoux, P., Held, I., John, J., Krasting, J. P., Malyshev, S., Naik, V., and Paulot, F.: The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics, Journal of Advances in Modeling Earth Systems, 12, e2019MS002015, 10.1029/2019MS002015, 2020.

Durán-Campos, E., Monreal-Gómez, M. A., de León, D. A. S., and Coria-Monter, E.: Chlorophyll-a vertical distribution patterns during summer in the Bay of La Paz, Gulf of California, Mexico, Egyptian Journal of Aquatic Research, 45, 109-115, 10.1016/j.ejar.2019.04.003, 2019.

Eppley, R. W., Rogers, J. N., and Mccarthy, J. J.: HALF-SATURATION CONSTANTS FOR UPTAKE OF NITRATE AND AMMONIUM BY MARINE PHYTOPLANKTON1, Limnology and Oceanography, 14, 912-920, 10.4319/lo.1969.14.6.0912, 1969.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937-1958, 10.5194/gmd-9-1937-2016, 2016.

Fennel, K. and Boss, E.: Subsurface maxima of phytoplankton and chlorophyll: Steady-state solutions from a simple model, Limnology and Oceanography, 48, 1521-1534, 10.4319/lo.2003.48.4.1521, 2003.

Fisher, B. J., Poulton, A. J., Meredith, M. P., Baldry, K., Schofield, O., and Henley, S. F.: Climate-driven shifts in Southern Ocean primary producers and biogeochemistry in CMIP6 models, Biogeosciences, 22, 975-994, 10.5194/bg-22-975-2025, 2025.

Follows, M. J. and Dutkiewicz, S.: Modeling diverse communities of marine microbes, Annual review of marine science, 3, 427-451, 10.1146/annurev-marine-120709-142848, 2011.

Garcia, H. E., Boyer, T. P., Baranova, O. K., Locarnini, R. A., Mishonov, A. V., Grodsky, A., Paver, C. R., Weathers, K. W., Smolyar, I. V., Reagan, J. R., Seidov, D., and Zweng, M. M.: World Ocean Atlas 2018: Product Documentation, 2019.

Gruber, N., Landschützer, P., and Lovenduski, N. S.: The Variable Southern Ocean Carbon Sink, Annual Review of Marine Science, 11, 159-186, 10.1146/annurev-marine-121916-063407, 2019.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M.: Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks, Geoscientific Model Development, 13, 2197-2244, 10.5194/gmd-13-2197-2020, 2020.

Hauck, J., Völker, C., Wolf-Gladrow, D. A., Laufkötter, C., Vogt, M., Aumont, O., Bopp, L., Buitenhuis, E. T., Doney, S. C., Dunne, J., Gruber, N., Hashioka, T., John, J., Quéré, C. L., Lima, I. D., Nakano, H., Séférian, R., and Totterdell, I.: On the Southern Ocean CO₂ uptake and the role of the biological carbon pump in the 21st century, Global Biogeochemical Cycles, 29, 1451-1470, 10.1002/2015gb005140, 2015.

Henley, S. F., Cavan, E. L., Fawcett, S. E., Kerr, R., Monteiro, T., Sherrell, R. M., Bowie, A. R., Boyd, P. W., Barnes, D. K. A., Schloss, I. R., Marshall, T., Flynn, R., and Smith, S.: Changing Biogeochemistry of the Southern Ocean and Its Ecosystem Implications, Frontiers in Marine Science, 7, 10.3389/fmars.2020.00581, 2020.

Hoffmann, L. J., Peeken, I., and Lochte, K.: Iron, silicate, and light co-limitation of three Southern Ocean diatom species, Polar Biology, 31, 1067-1080, 10.1007/s00300-008-0448-6, 2008.

Hopkinson, B. M. and Barbeau, K. A.: Interactive influences of iron and light limitation on phytoplankton at subsurface chlorophyll maxima in the eastern North Pacific, Limnology and Oceanography, 53, 1303-1318, 10.4319/lo.2008.53.4.1303, 2008.

Hunt, B. P. V., Espinasse, B., Pakhomov, E. A., Cherel, Y., Cotté, C., Delegrange, A., and Henschke, N.: Pelagic food web structure in high nutrient low chlorophyll (HNLC) and naturally iron fertilized waters in the Kerguelen Islands region, Southern Ocean, Journal of Marine Systems, 224, 103625, 10.1016/j.jmarsys.2021.103625, 2021.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., and Núñez-Riboni, I.: Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations, Journal of Advances in Modeling Earth Systems, 5, 287-315, 10.1029/2012ms000178, 2013.

Johnson, R., Strutton, P. G., Wright, S. W., Mcminn, A., and Meiners, K. M.: Three improved satellite chlorophyll algorithms for the Southern Ocean, Journal of Geophysical Research: Oceans, 118, 3694-3703, 10.1002/jgrc.20270, 2013.

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

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

690 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, 10.5194/bg-17-3439-2020, 2020.
Lauderdale, J. M., Williams, R. G., Munday, D. R., and Marshall, D. P.: The impact of Southern Ocean residual upwelling on atmospheric CO2 on centennial and millennial timescales, Climate Dynamics, 48, 1611-1631, 10.1007/s00382-016-3163-y, 2017.
695 Letelier, R. M., Karl, D. M., Abbott, M. R., and Bidigare, R. R.: Light driven seasonal patterns of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical Gyre, Limnology and Oceanography, 49, 508-519, 10.4319/lo.2004.49.2.0508, 2004.
Li, G., Lin, Q., Ni, G., Shen, P., Fan, Y., Huang, L., and Tan, Y.: Vertical Patterns of Early Summer Chlorophyll a Concentration in the Indian Ocean with Special Reference to the Variation of Deep Chlorophyll Maximum, Journal of
700 Marine Biology, 2012, 1-6, 10.1155/2012/801248, 2012.
Liu, Q., Wang, J., Li, X.-J., Meng, N., Yang, G.-P., Zhang, G., and Zhuang, G.-C.: Regulation and response of heterotrophic bacterial production to environmental changes in marginal seas of the Western Pacific Ocean, Global and Planetary Change, 245, 104678, 10.1016/j.gloplacha.2024.104678, 2025.
Long, M. C., Moore, J. K., Lindsay, K., Levy, M., Doney, S. C., Luo, J. Y., Krumhardt, K. M., Letscher, R. T., Grover, M.,
705 and Sylvester, Z. T.: Simulations With the Marine Biogeochemistry Library (MARBL), Journal of Advances in Modeling Earth Systems, 13, 10.1029/2021ms002647, 2021.
Lopes, F., Viollier, E., Thiam, A., Michard, G., Abril, G., Groleau, A., Prévot, F., Carrias, J. F., Albéric, P., and Jézéquel, D.: Biogeochemical modelling of anaerobic vs. aerobic methane oxidation in a meromictic crater lake (Lake Pavin, France), Applied Geochemistry, 26, 1919-1932, 10.1016/j.apgeochem.2011.06.021, 2011.
710 Lovato, T., Peano, D., Butenschön, M., Materia, S., Iovino, D., Scoccimarro, E., Fogli, P. G., Cherchi, A., Bellucci, A., Gualdi, S., Masina, S., and Navarra, A.: CMIP6 Simulations With the CMCC Earth System Model (CMCC-ESM2), Journal of Advances in Modeling Earth Systems, 14, 10.1029/2021ms002814, 2022.
Marañón, E., Van Wambeke, F., Uitz, J., Boss, E. S., Dimier, C., Dinasquet, J., Engel, A., Haëntjens, N., Pérez-Lorenzo, M., Taillandier, V., and Zäncker, B.: Deep maxima of phytoplankton biomass, primary production and bacterial production in
715 the Mediterranean Sea, Biogeosciences, 18, 1749-1767, 10.5194/bg-18-1749-2021, 2021.
Marshal, W., Xiang Chung, J., Roseli, N. H., Md Amin, R., and Mohd Akhir, M. F. B.: Evaluation of CMIP6 model performance in simulating historical biogeochemistry across the southern South China Sea, Biogeosciences, 21, 4007-4035, 10.5194/bg-21-4007-2024, 2024.
Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M.,
720 Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-De-La-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S. S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H.,
725 Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., Von Storch, J. S., Tian, F., Voigt, A., Vrese, P., Wieners, K. H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO₂, Journal of Advances in Modeling Earth Systems, 11, 998-1038, 10.1029/2018ms001400, 2019.
Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., Taylor, K. E., and Schlund, M.: Context
730 for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models, Science Advances, 6, eaba1981, 10.1126/sciadv.aba1981, 2020.
Michaelis, L. and Menten, M. L.: Die kinetik der invertinwirkung, Biochem. z, 49, 352, 1913.
Mignot, A., Claustre, H., Uitz, J., Poteau, A., D'Ortenzio, F., and Xing, X.: Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: A Bio-Argo float investigation,
735 Global Biogeochemical Cycles, 28, 856-876, 10.1002/2013gb004781, 2014.
Morley, S. A., Abele, D., Barnes, D. K. A., Cárdenas, C. A., Cotté, C., Gutt, J., Henley, S. F., Höfer, J., Hughes, K. A., Martin, S. M., Moffat, C., Raphael, M., Stammerjohn, S. E., Suckling, C. C., Tulloch, V. J. D., Waller, C. L., and Constable, A. J.: Global Drivers on Southern Ocean Ecosystems: Changing Physical Environments and Anthropogenic Pressures in an Earth System, Frontiers in Marine Science, 7, 10.3389/fmars.2020.547188, 2020.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Morrison, A. K., Frölicher, T. L., and Sarmiento, J. L.: Upwelling in the Southern Ocean, Physics Today, 68, 27-32, 10.1063/pt.3.2654, 2015.
Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), Journal of Advances in Modeling Earth Systems, 10, 1383-1413, 10.1029/2017ms001217, 2018.
Nelson, D. M., Brzezinski, M. A., Sigmon, D. E., and Franck, V. M.: A seasonal progression of Si limitation in the Pacific sector of the Southern Ocean, Deep Sea Research Part II: Topical Studies in Oceanography, 48, 3973-3995, 10.1016/S0967-0645(01)00076-5, 2001.
Neubauer, D., Ferrachat, S., Siegenthaler-Le Drian, C., Stier, P., Partridge, D. G., Tegen, I., Bey, I., Stanelle, T., Kokkola, H., and Lohmann, U.: The global aerosol–climate model ECHAM6.3–HAM2.3 – Part 2: Cloud evaluation, aerosol radiative forcing, and climate sensitivity, Geoscientific Model Development, 12, 3609-3639, 10.5194/gmd-12-3609-2019, 2019.
Nissen, C. and Vogt, M.: Factors controlling the competition between Phaeocystis and diatoms in the Southern Ocean and implications for carbon export fluxes, Biogeosciences, 18, 251-283, 10.5194/bg-18-251-2021, 2021.
O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geoscientific Model Development, 9, 3461-3482, 10.5194/gmd-9-3461-2016, 2016.
Oke, P. R., Griffin, D. A., Schiller, A., Matear, R. J., Fiedler, R., Mansbridge, J., Lenton, A., Cahill, M., Chamberlain, M. A., and Ridgway, K.: Evaluation of a near-global eddy-resolving ocean model, Geoscientific Model Development, 6, 591-615, 10.5194/gmd-6-591-2013, 2013.
Orsi, A. H., Whitworth III, T., and Nowlin Jr, W. D.: On the meridional extent and fronts of the Antarctic Circumpolar Current, Deep Sea Research Part I: Oceanographic Research Papers, 42, 641-673, 10.1016/0967-0637(95)00021-W, 1995.
Pak, G., Noh, Y., Lee, M.-I., Yeh, S.-W., Kim, D., Kim, S.-Y., Lee, J.-L., Lee, H. J., Hyun, S.-H., Lee, K.-Y., Lee, J.-H., Park, Y.-G., Jin, H., Park, H., and Kim, Y. H.: Korea Institute of Ocean Science and Technology Earth System Model and Its Simulation Characteristics, Ocean Science Journal, 56, 18-45, 10.1007/s12601-021-00001-7, 2021.
Pawlowicz, R.: M_Map: A mapping package for MATLAB, Computer Software, 2020.
Petrou, K., Kranz, S. A., Trimborn, S., Hassler, C. S., Ameijeiras, S. B., Sackett, O., Ralph, P. J., and Davidson, A. T.: Southern Ocean phytoplankton physiology in a changing climate, Journal of Plant Physiology, 203, 135-150, 10.1016/j.jplph.2016.05.004, 2016.
Pollard, R., Tréguer, P., and Read, J.: Quantifying nutrient supply to the Southern Ocean, Journal of Geophysical Research: Oceans, 111, 10.1029/2005jc003076, 2006.
Prakash, P. and Bhaskar, T. V. S. U.: Temporal insights into deep chlorophyll maxima dynamics in the Indian sector of the Southern Ocean: a Bio-Argo float study, Journal of Oceanography, 80, 407-419, 10.1007/s10872-024-00736-6, 2024.
Rintoul, S. R., Hughes, C. W., and Olbers, D.: The Antarctic circumpolar current system, in: International geophysics, Elsevier, 271-XXXVI, 10.1016/S0074-6142(01)80124-8, 2001.
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, 10.1038/s43247-023-00871-w, 2023.
Sauzède, R., Claustre, H., Uitz, J., Jamet, C., Dall'Olmo, G., D'Ortenzio, F., Gentili, B., Poteau, A., and Schmechtig, C.: A neural network-based method for merging ocean color and Argo data to extend surface bio-optical properties to depth: Retrieval of the particulate backscattering coefficient, Journal of Geophysical Research: Oceans, 121, 2552-2571, 10.1002/2015jc011408
10.48670/moi-00046 (data), 2016.
Sauzède, R., Martinez, E., Pasqueron De Fommervault, O., Poteau, A., Mignot, A., Maes, C., Claustre, H., Uitz, J., Maamaatuaiahutapu, K., Rodier, M., Schmechtig, C., and Laurent, V.: Seasonal dynamics and disturbance of phytoplankton biomass in the wake of Tahiti as observed by Biogeochemical-Argo floats, 10.5194/bg-2017-541, 2018.
Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J., Dunne, J., Gehlen, M., Ilyina, T., John, J. G., Li, H., Long, M. C., Luo, J. Y., Nakano, H., Romanou, A., Schwinger, J., Stock, C., Santana-Falcón, Y., Takano, Y., Tjiputra, J., Tsujino, H., Watanabe, M., Wu, T., Wu, F., and Yamamoto, A.: Tracking

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6, Current Climate Change Reports, 6, 95-119, 10.1007/s40641-020-00160-0, 2020.
Séférian, R., Nabat, P., Michou, M., Saint-Martin, D., Voldoire, A., Colin, J., Decharme, B., Delire, C., Berthet, S., Chevallier, M., Sénési, S., Franchisteguy, L., Vial, J., Mallet, M., Joetzjer, E., Geoffroy, O., Guérémy, J. F., Moine, M. P., Msadek, R., Ribes, A., Rocher, M., Roehrig, R., Salas-Y-Mélia, D., Sanchez, E., Terray, L., Valcke, S., Waldman, R., Aumont, O., Bopp, L., Deshayes, J., Éthé, C., and Madec, G.: Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate, Journal of Advances in Modeling Earth Systems, 11, 4182-4227, 10.1029/2019ms001791, 2019.
Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'connor, F. M., Stringer, M., Hill, R., and Palmieri, J.: UKESM1: Description and evaluation of the UK Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513-4558, 10.1029/2019MS001739, 2019.
Signorini, S. R., Franz, B. A., and Mcclain, C. R.: Chlorophyll variability in the oligotrophic gyres: mechanisms, seasonality and trends, Frontiers in Marine Science, 2, 10.3389/fmars.2015.00001, 2015.
Skyllas, N.: Evaluation of NEMO–PISCES v2 biogeochemical model with field data from a north-south gradient in the Northeast Atlantic Ocean & Growth of Emiliania huxleyi and Thalassiosira oceanica under 3 irradiance regimes, 2018.
Song, X.: Explaining the Zonal Asymmetry in the Air-Sea Net Heat Flux Climatology Over the Antarctic Circumpolar Current, Journal of Geophysical Research: Oceans, 125, 10.1029/2020jc016215, 2020.
Stock, C. A., Dunne, J. P., Fan, S., Ginoux, P., John, J., Krasting, J. P., Laufkötter, C., Paulot, F., and Zadeh, N.: Ocean Biogeochemistry in GFDL's Earth System Model 4.1 and Its Response to Increasing Atmospheric CO₂, Journal of Advances in Modeling Earth Systems, 12, 10.1029/2019ms002043, 2020.
Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., Von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geoscientific Model Development, 12, 4823-4873, 10.5194/gmd-12-4823-2019, 2019.
Tagliabue, A., Bowie, A. R., Boyd, P. W., Buck, K. N., Johnson, K. S., and Saito, M. A.: The integral role of iron in ocean biogeochemistry, Nature, 543, 51-59, 10.1038/nature21058, 2017.
Tagliabue, A., Mtshali, T., Aumont, O., Bowie, A. R., Klunder, M. B., Roychoudhury, A. N., and Swart, S.: A global compilation of dissolved iron measurements: focus on distributions and processes in the Southern Ocean, Biogeosciences, 9, 2333-2349, 10.5194/bg-9-2333-2012, 2012.
Tagliabue, A., Buck, K. N., Sofen, L. E., Twining, B. S., Aumont, O., Boyd, P. W., Caprara, S., Homoky, W. B., Johnson, R., König, D., Ohnemus, D. C., Sohst, B., and Sedwick, P.: Authigenic mineral phases as a driver of the upper-ocean iron cycle, Nature, 620, 104-109, 10.1038/s41586-023-06210-5, 2023.
Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106, 7183-7192, 10.1029/2000jd900719, 2001.
Timmermans, K. R., Van Der Wagt, B., and De Baar, H. J. W.: Growth rates, half-saturation constants, and silicate, nitrate, and phosphate depletion in relation to iron availability of four large, open-ocean diatoms from the Southern Ocean, Limnology and Oceanography, 49, 2141-2151, 10.4319/lo.2004.49.6.2141, 2004.
Tjiputra, J. F., Schwinger, J., Bentsen, M., Morée, A. L., Gao, S., Bethke, I., Heinze, C., Goris, N., Gupta, A., He, Y.-C., Olivié, D., Seland, Ø., and Schulz, M.: Ocean biogeochemistry in the Norwegian Earth System Model version 2 (NorESM2), Geoscientific Model Development, 13, 2393-2431, 10.5194/gmd-13-2393-2020, 2020.
Venables, H. and Moore, C. M.: Phytoplankton and light limitation in the Southern Ocean: Learning from high-nutrient, high-chlorophyll areas, Journal of Geophysical Research: Oceans, 115, 10.1029/2009jc005361, 2010.
Vichi, M., Lovato, T., Lazzari, P., Cossarini, G., Gutierrez Mlot, E., Mattia, G., Masina, S., McKiver, W., Pinardi, N., and Solidoro, C.: The biogeochemical flux model (BFM): Equation description and user manual, BFM version, 5, 104, 2015.
Vives, C. R., Schallenberg, C., Strutton, P. G., Bendtsen, J., Richardson, K., and Boyd, P. W.: An inter-comparison of Deep Chlorophyll Maxima characteristics from 30S to 74S and their contribution to Net Primary Production, 10.22541/au.172417549.92336092/v1, 2024.
Xue, T., Terhaar, J., Prowe, A. E. F., Frölicher, T. L., Oschlies, A., and Frenger, I.: Southern Ocean phytoplankton under climate change: a shifting balance of bottom-up and top-down control, Biogeosciences, 21, 2473-2491, 10.5194/bg-21-2473-2024, 2024.

https://doi.org/10.5194/egusphere-2025-2633
Preprint. Discussion started: 25 July 2025

Yang, G., Atkinson, A., Pakhomov, E. A., Hill, S. L., and Racault, M. F.: Massive circumpolar biomass of Southern Ocean zooplankton: Implications for food web structure, carbon export, and marine spatial planning, Limnology and Oceanography, 67, 2516-2530, 10.1002/lno.12219, 2022.

Yingling, N., Selph, K. E., Décima, M., Safi, K. A., Gutiérrez-Rodríguez, A., Fender, C. K., and Stukel, M. R.: Investigating plankton size spectra, biomass, abundance, and community composition in the Subtropical Convergence Front in the Southern Ocean, Frontiers in Marine Science, 12, 10.3389/fmars.2025.1465125, 2025.

Yool, A., Popova, E. E., and Anderson, T. R.: MEDUSA-2.0: an intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies, Geoscientific Model Development, 6, 1767-1811, 10.5194/gmd-6-1767-2013, 2013.

Zahariev, K., Christian, J. R., and Denman, K. L.: The Canadian Model of Ocean Carbon (CMOC) v1. 0, 2007.

Zampollo, A., Cornulier, T., O'Hara Murray, R., Tweddle, J. F., Dunning, J., and Scott, B. E.: The bottom mixed layer depth as an indicator of subsurface Chlorophyll a distribution, Biogeosciences, 20, 3593-3611, 10.5194/bg-20-3593-2023, 2023.

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, Journal of Southern Hemisphere Earth Systems Science, 70, 193-214, 10.1071/es19035, 2020.