



## **Incorporating Recalcitrant Dissolved Organic Carbon and Microbial Carbon Pump Processes into the cGENIE Earth System Model (cGENIEv0.9.35-MCP)**

Yuxian Lai<sup>1</sup>, Wentao Ma<sup>2\*</sup>, Zihui Zhao<sup>3</sup>, Peng Xiu<sup>1</sup>, Nianzhi Jiao<sup>1\*</sup>

<sup>1</sup>State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Fujian

5 Key Laboratory of Marine Carbon Sequestration, Xiamen University, Xiamen, China

<sup>2</sup>State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, MNR, China

<sup>3</sup>Changshan Meteorological Bureau, Quzhou, China

Correspondence to: Wentao Ma ([wtma@sio.org.cn](mailto:wtma@sio.org.cn)), Nianzhi Jiao ([jiao@xmu.edu.cn](mailto:jiao@xmu.edu.cn))

10

**Abstract.** Recalcitrant dissolved organic carbon (RDOC) is a significant component of dissolved organic carbon (DOC), produced through the microbial carbon pump (MCP), and plays a crucial role in long-term carbon sequestration. In this study, we extend the cGENIE Earth System Model by integrating the RDOC fraction and embedding MCP-driven transformations, resulting in the enhanced cGENIE-MCP model.

15 We implement temperature-dependent limitations on nutrient uptake and organic matter remineralization to simulate MCP processes. Model outputs are compared with contemporary observations and previous cGENIE versions. The model effectively simulates the spatial distribution of concentrations and production rates of labile (LDOC), semi-labile (SLDOC), and RDOC. The cGENIE-MCP model demonstrates improved accuracy over previous versions, capturing spatial variability in DOC pools and  
20 quantifying MCP contributions to long-term carbon sequestration. For instance, sea surface DOC concentrations exhibit a latitudinal gradient, with values ranging from 65-80  $\mu\text{mol kg}^{-1}$  in tropical-subtropical zones to 40-50  $\mu\text{mol kg}^{-1}$  in subpolar regions. RDOC concentrations remain relatively stable at 40-50  $\mu\text{mol kg}^{-1}$  throughout the water column, while LDOC and SLDOC concentrations are typically below 10  $\mu\text{mol kg}^{-1}$  and 34  $\mu\text{mol kg}^{-1}$ , respectively, in high-production areas. The model reveals a strong  
25 spatial correlation between primary production and LDOC production in upwelling zones, while RDOC production exhibits long-term carbon sequestration. These results emphasize the importance of incorporating MCP processes into Earth system models to better predict ocean carbon sink efficiency and biogeochemical responses to climate change. The cGENIE-MCP model provides a tool for studying the dynamics of ocean DOC and carbon cycle across timescales from paleo to future projections.

30



## 1 Introduction

The ocean serves as the largest active carbon reservoir on Earth, sequestering nearly 25% of anthropogenic CO<sub>2</sub> emissions (Gruber et al., 2023). With rising anthropogenic CO<sub>2</sub> emissions, understanding the response of the ocean carbon cycle to future CO<sub>2</sub> levels is vital for predicting biogeochemical and physical changes in marine systems (Devries, 2022). Dissolved organic carbon (DOC), generally defined as organic matter passing through 0.2-0.7 µm filters, is the largest ocean reservoir of reduced carbon and plays a key role in ocean biogeochemical cycles (Hansell et al., 2009; Hansell and Carlson, 2014). A significant portion of DOC originates from photosynthesis by phytoplankton in the surface ocean (Hansell, 2013). DOC mainly supports heterotrophic prokaryotes, which respire it to CO<sub>2</sub> or transform and diversify it through various metabolic processes (Keller et al., 2014).

DOC can be further categorized based on its degradation rates: labile DOC (LDOC), which degrades within hours to days; semi-labile DOC (SLDOC), with turnover times of weeks to months; and refractory DOC (RDOC), which can persist for tens of thousands of years. LDOC and SLDOC fractions undergo microbial transformations, leading to RDOC production through processes known as the microbial carbon pump (MCP) (Jiao et al., 2010; Jiao et al., 2024b). The MCP is influenced by environmental parameters such as temperature, nutrient availability, and dissolved oxygen concentration (Jiao et al., 2024b; Legendre et al., 2015). It plays a regulatory role in climate change due to RDOC's recalcitrance and large-scale accumulation (Legendre et al., 2015).

The evaluation of the ocean's long-term carbon sequestration potential under climate change relies on numerical models. Despite the advances of the Coupled Model Intercomparison Project Phase 6 (CMIP6), state-of-the-art Earth System Models (ESMs) still exhibit insufficient representation of the MCP and RDOC reservoir (Polimene et al., 2018; Ma et al., 2022). Many ESMs simplify DOC into a single dynamic pool, neglecting the explicit representation of RDOC or the corresponding degradation timescales (Anderson et al., 2015). These simplifications lead to systematic biases: observed RDOC concentrations in deep basins exceed model outputs by up to 40%, while  $\Delta^{14}\text{C}$  signatures cannot adequately replicate RDOC's millennial-scale lifetime (Yamashita and Tanoue, 2008; Hansell et al., 2012). These discrepancies may arise from the lack of explicit MCP parameterizations that link the transformation of LDOC into RDOC (Hansell et al., 2012; Séférian et al., 2020). Consequently, ESMs underestimate the ocean's carbon sequestration capacity, particularly under global warming scenarios,



where changes in temperature and nutrient availability may alter RDOC accumulation (Séférian et al., 2020; Jokulsdottir, 2011; Sarmiento, 2013; Jiao et al., 2024b; Li et al., 2019).

To address these limitations, integrating MCP processes into ESMs is essential (Ridgwell and Arndt, 2015). While traditional ESMs prioritize high-resolution, near-term climate projections (Séférian et al., 2020), the cGENIE Earth system model emphasizes computational efficiency and flexibility, enabling the exploration of earth system feedbacks over multi-millennial scales (Ridgwell et al., 2007). The GENIE-1 model was initially used by Lenton et al. (2006) to explore the carbon cycle and climate change over millennial timescales. Ridgwell et al. (2007) extended this work to examine the influence of ocean biogeochemistry and marine sediments on atmospheric CO<sub>2</sub>. The cGENIE model couples ocean biogeochemistry with physical ocean and sea ice dynamics, facilitating studies on ocean carbon sequestration, DOC cycling, and paleoclimate feedbacks (Crichton et al., 2021; Van De Velde et al., 2021; Crichton et al., 2023; Ridgwell and Zeebe, 2005; Anagnostou et al., 2016; Boscolo-Galazzo et al., 2021; Vervoort et al., 2024; Ma and Tian, 2014; Wang et al., 2023; Ridgwell et al., 2007).

Despite recent advances, the explicit representation of MCP processes and RDOC cycling within cGENIE has remained limited. The Minnesota Earth System Model for Ocean biogeochemistry (MESMO 3), an Earth system model of intermediate complexity derived from cGENIE, includes an explicit treatment of semi-labile and refractory DOM pools (Matsumoto et al., 2021). However, MESMO 3 lacks a mechanistic representation of DOM production pathways associated with MCP processes—specifically, the transformation of semi-labile DOC (SLDOC) into RDOC—and has not been calibrated against global DOM observations. In this study, we propose an extension to the cGENIE model that integrates RDOC and MCP processes to investigate their long-term response to climate change. We introduce a new framework, cGENIE-MCP, which partitions total DOC into three distinct fractions—labile (LDOC), semi-labile (SLDOC), and refractory (RDOC)—enabling improved simulation of DOC production and remineralization based on prior formulations from cGENIE and MESMO 3. We evaluate the performance of the cGENIE-MCP model against global observational datasets and compare its outputs with those from the standard cGENIE model. Finally, we analyze the spatial distribution and production of LDOC, SLDOC, and RDOC in relation to primary production to assess the model’s capability in capturing essential features of the MCP.

## 2 Model description

### 2.1 The cGENIE model framework



The cGENIE model, derived from the Grid-Enabled Integrated Earth System Model (GENIE), is an Earth system model of intermediate complexity (EMIC) designed for simulating long-term carbon cycle dynamics. It is structured based on a 3-D frictional geostrophic ocean model, a 2-D energy-moisture balance model (EMBM), and a dynamic-thermodynamic sea-ice model, all within a modular framework that integrates essential Earth system components. These include ocean circulation, marine biogeochemistry, ocean-atmosphere and ocean-sediment exchanges, and the long-term geological carbon cycle on land (Lenton et al., 2006; Van De Velde et al., 2021; Marsh et al., 2011; Williamson et al., 2006; Ridgwell). Compared to high-resolution Earth System Models (ESMs), cGENIE operates at coarser spatial and temporal resolutions, offering enhanced computational efficiency and flexibility. This makes it particularly suitable for exploring Earth system feedbacks over multi-millennial timescales.

In this study, we adopt the global ocean circulation configuration of Cao et al. (2009), which forms the basis for numerous subsequent biogeochemical developments within cGENIE (Crichton et al., 2021; Reinhard et al., 2020; Reinhard et al.). Physical parameters and boundary conditions governing the climate system are retained from Cao et al. (2009). The model includes an explicit representation of iron cycling and co-limitation of primary productivity, following the formulation by Ward et al. (2018). Simulations are initialized from a pre-industrial climate state with atmospheric CO<sub>2</sub> set at 278 ppm. The model is configured on a 36×36 horizontal grid with equal area (i.e. equal divisions in longitude and the sine of latitude) and 16 vertical ocean layers of increasing thickness (Ridgwell et al., 2007; Anagnostou et al., 2016).

## 2.2 Ocean biogeochemical module

The ocean biogeochemistry in cGENIE is simulated using the BIOGEM module, which models nutrient-driven biological productivity and the cycling of carbon and associated elements (Van De Velde et al., 2021). BIOGEM includes representations of air-sea gas exchange, nutrient uptake by primary producers, and the remineralization of organic matter in the water column. Phytoplankton are not explicitly represented; instead, primary productivity is calculated diagnostically based on the availability of limiting nutrients (phosphate and dissolved iron), solar radiation, and temperature, following previous studies (Ward et al., 2018; Matsumoto et al., 2008b; Tanioka et al., 2021; Matsumoto et al., 2008a; Matsumoto et al., 2021). Phosphate (PO<sub>4</sub>) and other nutrients are converted to particulate organic matters (POMs) in the euphotic zone according to the Redfield stoichiometry. POMs are exported from the surface layer and undergo depth-dependent remineralization, releasing nutrients and carbon back into the water column. In this study, we extend the BIOGEM module by partitioning DOC into three fractions—LDOC,



SLDOC, and RDOC—to better represent MCP processes. A fraction of primary production is allocated  
 125 to each DOC pool. LDOC degradation is regulated by temperature, SLDOC degradation is driven  
 primarily by photodegradation, and a portion of SLDOC is converted into RDOC. The transformation  
 from SLDOC to RDOC is governed by the characteristic lifetime of SLDOC, reflecting microbial  
 reprocessing pathways consistent with the MCP framework. The updated DOC cycling scheme enables  
 us to track the production, transformation, and persistence of different DOC fractions under varying  
 130 environmental conditions. In the following sections, we provide a detailed description of these carbon  
 cycling processes and the associated parameterizations.

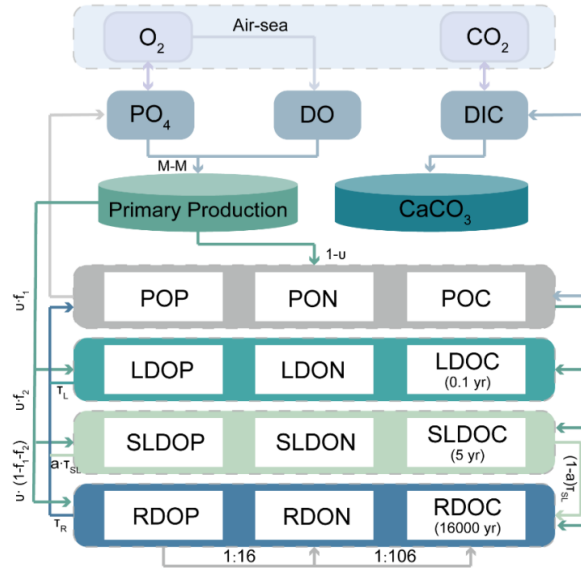


Figure1. Schematic representation of phosphorus, carbon, and oxygen cycling in the BIOGEM module  
 135 of the cGENIE-MCP model.

### 2.2.1 Ocean-atmosphere gas exchange

The flux of  $CO_2$  across the air-sea interface is calculated following Ridgwell et al. (2007):

$$f_{atm \rightarrow ocn}^{CO_2} = k \cdot Sol \cdot (pCO_{2atm} - pCO_{2sea}),$$

140 where  $pCO_{2atm}$  and  $pCO_{2sea}$  are the partial pressure of  $CO_2$  ( $\mu atm$ ) in the atmosphere and ocean surface, respectively.  $k$  is the gas transfer velocity ( $m s^{-1}$ ), and  $Sol$  ( $mol m^{-3} \mu atm^{-1}$ ) is solubility of  $CO_2$  in sea water. The gas transfer velocity is calculated as:



$$k = l \cdot [2.5 \cdot (a + b \cdot T + c \cdot T^2) + 0.38 \cdot \mu_{av}^2] \cdot \left(\frac{Sc}{660}\right)^{-0.5},$$

where  $a$ ,  $b$ , and  $c$  are constants, with values of 0.5246, 0.01625, and  $4.9946 \times 10^{-4}$ , respectively;  $T$  is the  
 145 ocean temperature in Kelvin (K).  $l$  is the scaling factor;  $\mu_{av}$  is the average wind speed ( $\text{m s}^{-1}$ ); and  $Sc$   
 is the Schmidt number, calculated as:

$$Sc = A - B \cdot T + C \cdot T^2 - D \cdot T^3,$$

where  $A$ ,  $B$ ,  $C$ , and  $D$  are empirical constants with constants:  $A=2073.1$ ,  $B=125.62$ ,  $C=3.6267$ , and  
 $D=0.0432$ .

150 The total  $\text{CO}_2$  flux is area-weighted by the ice-free ocean surface area:

$$F_{ocn \rightarrow atm}^{CO_2} = \sum_c (f_{atm \rightarrow ocn}^{CO_2} \cdot (A_{(c)} - A_{ice(c)})),$$

where  $A_{(c)}$  is the grid cell area and  $A_{ice(c)}$  is the ice-covered portion of the cell.

### 2.2.2 Biological production

155 Nutrient uptake in each surface grid cell is given by:

$$\Gamma = (1 - A) \cdot \gamma^T \cdot \min\left(\frac{PO_4^{3-}}{PO_4^{3-} + K^{PO_4^{3-}}}, \frac{FeT}{FeT + K^{FeT}}\right) \cdot \frac{I}{I_0} \cdot \frac{\min(PO_4^{3-}, red_{P/Fe} \cdot FeT)}{\tau_{bio}},$$

$$\gamma^T = a e^{\left(\frac{T}{b}\right)},$$

Where  $A$  represents the ice-covered fraction of the grid cell, the temperature factor  $\gamma^T$  is derived from  
 Eppley's initial values, with  $a = 0.59$  and  $b = 75.80$ , which is associated with the  $Q_{10}$  factor.  $K^{PO_4^{3-}}$  and  
 160  $K^{FeT}$  are half-saturation constants for phosphate and iron ( $\text{mol kg}^{-1}$ ), respectively.  $PO_4^{3-}$  is the phosphate  
 concentration ( $\text{mol kg}^{-1}$ ), and  $FeT$  is the iron concentration ( $\text{mol kg}^{-1}$ ).  $I$  is the local solar radiation, and  
 $I_0$  is the solar constant,  $red_{P/Fe}$  is the Fe:P Redfield ratio,  $\tau_{bio}$  is an optimal uptake timescale (h).

The primary production ( $F_{z=h_e}^{POP}$ ) is derived from nutrient uptake ( $\text{mol m}^{-2} \text{yr}^{-1}$ ). Biologically  
 assimilated  $\text{PO}_4$  is converted into particulate organic phosphorus (POP), with a fraction partitioned into  
 165 labile dissolved organic phosphorus (LDOP), semi-labile dissolved organic phosphorus (SLDOP), and  
 refractory dissolved organic phosphorus (RDOP). The total phosphorus uptake is partitioned as:

$$\int_{h_e}^0 \rho \cdot \Gamma \, dz = F_{z=h_e}^{POP} + F_{z=h_e}^{LDOP} + F_{z=h_e}^{SLDOP} + F_{z=h_e}^{RDOP},$$

with each term defined as:

$$F_{z=h_e}^{POP} = \int_{h_e}^0 \rho \cdot (1 - v) \cdot \Gamma \, dz,$$



170

$$F_{z=h_e}^{LDOP} = \int_{h_e}^0 \rho \cdot \nu \cdot f_1 \cdot \Gamma \, dz,$$

$$F_{z=h_e}^{SLDOP} = \int_{h_e}^0 \rho \cdot \nu \cdot f_2 \cdot \Gamma \, dz,$$

$$F_{z=h_e}^{RDOP} = \int_{h_e}^0 \rho \cdot \nu \cdot (1 - f_1 - f_2) \cdot \Gamma \, dz,$$

where  $\nu = 0.66$  is the fraction of nutrient uptake allocated to dissolved organic phosphorus following the OCMIP-2 protocol (Ridgwell et al., 2007),  $f_1=0.9599$  and  $f_2=0.04$  are the export partitioning coefficients for LDOP and SLDOP, respectively, based on Wang et al. (2023),  $\rho$  is the seawater density (kg m<sup>-3</sup>),  $\Gamma$  is the nutrient uptake rate, and  $h_e$  is the depth of the euphotic layer (m). Organic phosphorus fluxes are converted to organic carbon fluxes using the Redfield ratio (C:N:P=106:16:1):

175

$$F_{z=h_e}^{POC} = 106 \cdot F_{z=h_e}^{POP},$$

$$F_{z=h_e}^{LDOC} = 106 \cdot F_{z=h_e}^{LDOP},$$

180

$$F_{z=h_e}^{SLDOC} = 106 \cdot F_{z=h_e}^{SLDOP},$$

$$F_{z=h_e}^{RDOC} = 106 \cdot F_{z=h_e}^{RDOP}.$$

POC is categorized to labile POC (POC1) and recalcitrant POC (POC2) based on their relative degradability, following Ridgwell et al. (2007):

$$F_{z=h_e}^{POC1} = 0.992 \cdot F_{z=h_e}^{POC},$$

185

$$F_{z=h_e}^{POC2} = 0.008 \cdot F_{z=h_e}^{POC}.$$

This implies that 0.8% of the total POC flux is assigned to the recalcitrant pool, with the remaining 99.2% attributed to the labile pool.

### 2.2.3 Remineralization in the water column

190

Remineralization in the water column converts portions of POC and DOC into inorganic nutrients, supporting nutrient cycling. While some POC sinks to the bottom waters and is eventually buried in sediments, a substantial fraction is remineralized within the water column.

#### POC remineralization

195

The remineralization of POC is modeled as a temperature-dependent process, with separate treatment for labile (POC1) and recalcitrant (POC2) components. The remineralization rate is given by:

$$k_{zPOCn} = \beta_{POCn} \cdot e^{\left(\frac{-E_{an}}{RT_z}\right)},$$



$$F_{rz}^{POCn} = k_{zPOCn} \cdot POCn_z,$$

where  $POCn_z$  is the concentration of POC component  $n$  (either POC1 or POC2) at depth  $z$ ;  $k_{zPOCn}$  represents the remineralization rate of  $POCn$ ;  $E_{an}$  is the activation energy of POC component  $n$  ( $J\ mol^{-1}$ );  $R$  is the gas constant ( $J\ K^{-1}\ mol^{-1}$ );  $T_z$  is the temperature at  $z$  layer (K);  $\beta_{POCn}$  is the  
 200 remineralization rate constant for  $POCn$  ( $yr^{-1}$ ).  $F_{rz}^{POCn}$  denotes the remineralization flux of  $POCn$  at  $z$  layer.

### LDOC remineralization

LDOC is rapidly remineralized in the water column, releasing inorganic carbon and nutrients. The  
 205 remineralization follows a similar temperature-dependent formulation:

$$k_{zLDOC} = \beta_{LDOC} \cdot e^{\left(\frac{-E_a}{RT_z}\right)},$$

$$F_{rz}^{LDOC} = k_{zLDOC} \cdot LDOC_z,$$

where  $k_{zLDOC}$  is the temperature-dependent remineralization rate of LDOC at  $z$  layer;  $E_a$  is the activation energy ( $J\ mol^{-1}$ );  $R$  is the gas constant ( $J\ K^{-1}\ mol^{-1}$ );  $T_z$  is the temperature at  $z$  (K);  $\beta_{LDOC}$  is  
 210 the remineralization rate constant ( $yr^{-1}$ );  $LDOC_z$  is the LDOC concentration at  $z$  layer ( $mol\ kg^{-1}$ ).  $F_{rz}^{LDOC}$  denotes the remineralization flux of LDOC at  $z$  layer.

### SLDOC remineralization

SLDOC remineralizes more slowly than LDOC and contributes to nutrient cycling on intermediate timescales. The remineralization rate is described by:

$$k_{zSLDOC} = \frac{1}{\tau_{SL}},$$

$$F_{rz}^{SLDOC} = k_{zSLDOC} \cdot SLDOC_z,$$

where  $k_{zSLDOC}$  is the temperature-dependent remineralization rate of SLDOC at  $z$  layer ( $yr^{-1}$ );  $\tau_{SL}$  is the SLDOC lifetime (yr);  $SLDOC_z$  is the SLDOC concentration at  $z$  layer ( $mol\ kg^{-1}$ ).  $F_{rz}^{SLDOC}$  denotes the remineralization flux of SLDOC at  $z$  layer.

### RDOC remineralization

RDOC remineralizes on much longer timescales and is partly derived from the transformation of SLDOC. The RDOC remineralization and conversion fluxes are:

$$k_{zRDOC} = \frac{1}{\tau_R},$$

$$F_{conv}^{RDOC} = F_{rz}^{SLDOC} \cdot a,$$

$$F_{rz}^{RDOC} = k_{zRDOC} \cdot RDOC_z,$$

225





where  $k_{zRDOC}$  is the temperature-dependent remineralization rate of RDOC at  $z$  layer ( $\text{yr}^{-1}$ );  $\tau_R$  is the RDOC lifetime (yr);  $RDOC_z$  is the RDOC concentration at  $z$  layer.  $F_{conv}^{RDOC}$  denotes the remineralization flux of converted from SLDOC to RDOC at  $z$  layer.  $a$  is the conversion rate from SLDOC to RDOC, which is based on the Wang et al. (2023);  $F_z^{RDOC}$  denotes the remineralization flux of RDOC at  $z$  layer.

230

**Table 1. Parameters of biogeochemical processes in the cGENIE-MCP model**

Parameters	Symbol	Value	Unit	Reference
The scaling factor	$l$	$2.778 \times 10^{-6}$	-	(Crichton et al., 2021)
The half-saturation constant	$K^{PO_4^{3-}}$	$0.05 \times 10^{-6}$	$\text{mol kg}^{-1}$	(Crichton et al., 2021)
	$K^{FeT}$	$0.03 \times 10^{-9}$	$\text{mol kg}^{-1}$	(Matsumoto et al., 2021)
A proportion of $PO_4$ take up by the biota is partitioned into DOP	$v$	0.66	-	(Najjar et al., 1992; Orr et al., 2017)
Optimal uptake timescale	$\tau_{bio}$	95.63	h	(Van De Velde et al., 2021)
A proportion of $PO_4$ take up by the biota is partitioned into LDOP	$f_1$	0.9599	-	(Wang et al., 2023)
A proportion of $PO_4$ take up by the biota is partitioned into SLDOP	$f_2$	0.04	-	(Wang et al., 2023)
The activation energy of $POC_1$	$E_{a1}$	54000	$\text{J mol}^{-1}$	(Crichton et al., 2021)
The activation energy of $POC_2$	$E_{a2}$	80000	$\text{J mol}^{-1}$	(Crichton et al., 2021)
The activation energy of LDOC	$E_{aL}$	54000	$\text{J mol}^{-1}$	(Crichton et al., 2021)
The SLDOC lifetime	$\tau_{SL}$	5	yr	(Letscher et al., 2015)
The RDOC lifetime	$\tau_R$	16000	yr	(Letscher et al., 2015)
A proportion of SLDOC take up by the biota is partitioned into RDOC	$a$	0.015	-	(Wang et al., 2023)

## 2.3 Model validation method

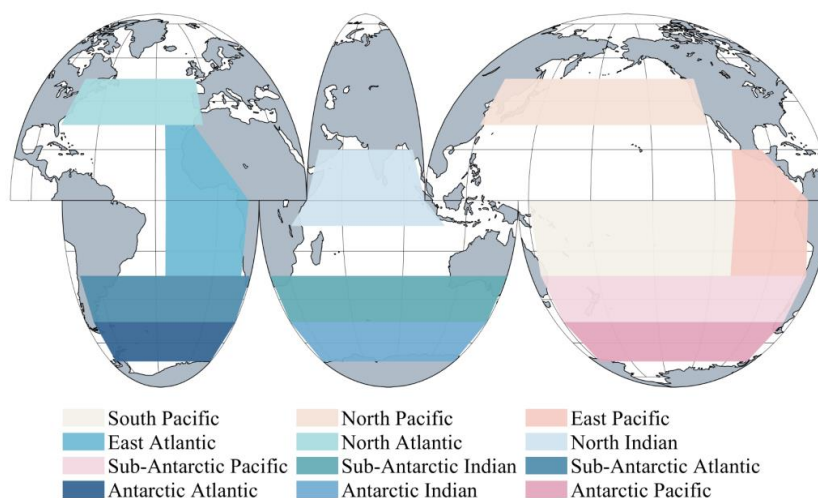


235 To evaluate the performance of the cGENIE-MCP model, we compare simulated outputs with climatological and observational datasets. Modeled ocean temperature, salinity, PO<sub>4</sub>, and DO are evaluated against the World Ocean Atlas 2023 (WOA23) climatology. DIC and alkalinity are compared with the Global Ocean Data Analysis Project (GLODAP) dataset (Key et al., 2004; Lauvset et al., 2024; Lauvset et al., 2021; Lauvset et al., 2023). Model skill is quantified using the root mean square deviation (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [(X_i - \bar{X}) - (Y_i - \bar{Y})]^2},$$

where  $X_i$  and  $\bar{X}$  represent the simulated tracer concentration at position  $i$  and its mean across time,  $Y_i$  and  $\bar{Y}$  denote the observed tracer concentration and its mean;  $N$  is the number of data points used in the calculation.

245 For consistency, WOA23 and GLODAP datasets are interpolated onto the cGENIE-MCP model grid (36×36×16 grids). Ocean regions used for the validation of vertical profiles are based on regional divisions from Crichton et al. (2021) with minor modifications to ensure each region maintains consistent oceanographic characteristics and particle flux regimes (Figure 2).



250 Figure 2. Ocean regions used for model validation in the cGENIE-MCP simulations. Colored regions indicate the spatial domains included in the validation of water column biogeochemical properties. Unshaded areas are excluded from the analysis due to limited relevance or observational data availability.

## 255 3 Model Results

### 3.1 Temperature



Ocean temperature strongly influences biogeochemical cycling and nutrient uptake, as demonstrated in prior studies (Yan et al., 2024). To assess the performance of the cGENIE-MCP model, simulated temperature fields were evaluated against observations from the WOA23 at three representative depth levels: surface, intermediate (400 m), and deep (3000 m) layers. As shown in Fig. 3, the cGENIE-MCP model reproduces the large-scale spatial distribution of ocean temperature with good accuracy, especially in the surface layer and in low-latitude regions. Temperature exhibits a clear latitudinal gradient, decreasing from the tropics toward the poles in the upper and intermediate ocean, with relatively small variability in the deep ocean (Fig. 4). The RMSE in the tropical regions is approximately 1.2 °C, while RMSE values in the Atlantic Ocean range from 0.6 °C to 1.0 °C, indicating strong agreement with observations (Table 2).

Despite the model performing well overall, it underestimates temperature in high-latitude regions, deep layers, and equatorial upwelling zones. These biases likely arise from the limitations of the simplified two-dimensional energy-moisture balance atmospheric model used in cGENIE, which affects surface forcing and ocean circulation. Additionally, insufficient stratification in the upper ocean and uncertainties associated with the interpolation of WOA23 data onto the model grid may also contribute to discrepancies in the water column.

Comparisons between the standard cGENIE and cGENIE-MCP versions reveal negligible differences in temperature fields, with RMSE values of 0.01 °C (Fig. 3-4 and Table 2). These results confirm that incorporating the refractory dissolved organic matter (RDOM) component into the cGENIE-MCP model does not degrade the model’s performance in simulating ocean temperature, thereby preserving the core physical integrity of the system.

**Table 2. RMSE of modeled tracers for cGENIE-MCP and cGENIE compared to observational data**

Tracers	cGENIE-MCP	cGENIE
T (°C)	1.25	1.26
Salinity	0.34	0.34
PO <sub>4</sub> (μmol kg <sup>-1</sup> )	0.10	0.12
DO (μmol kg <sup>-1</sup> )	26.89	28.50
DIC (μmol kg <sup>-1</sup> )	51.33	50.92

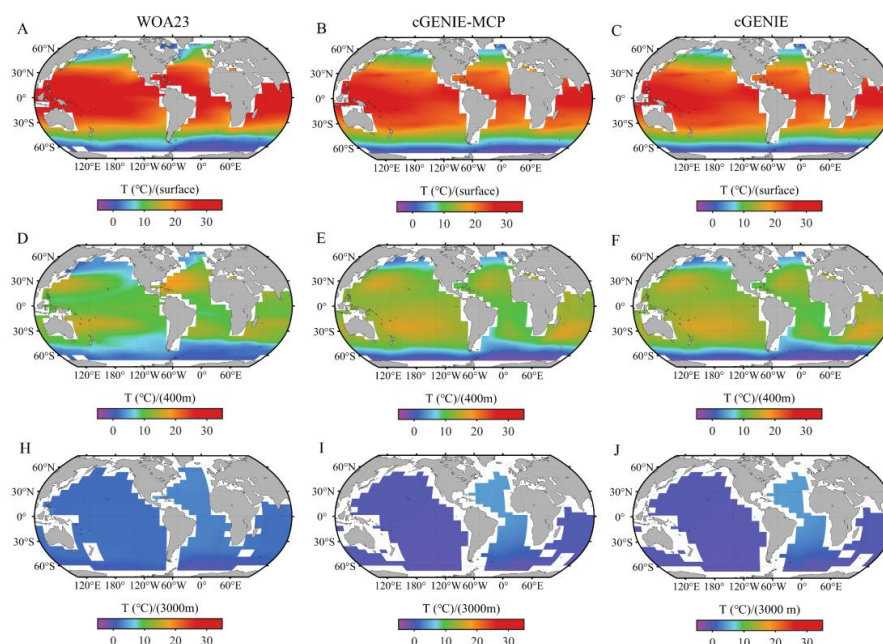


Figure 3. Ocean temperature distributions (°C) at three depth levels: (A-C) the surface, (D-F) 400 m, and (H-J) 3000 m. Panels show: (A, D, H) observational data from WOA23; (B, E, I) simulated results from the cGENIE-MCP model; and (C, F, J) outputs from the standard cGENIE.

285

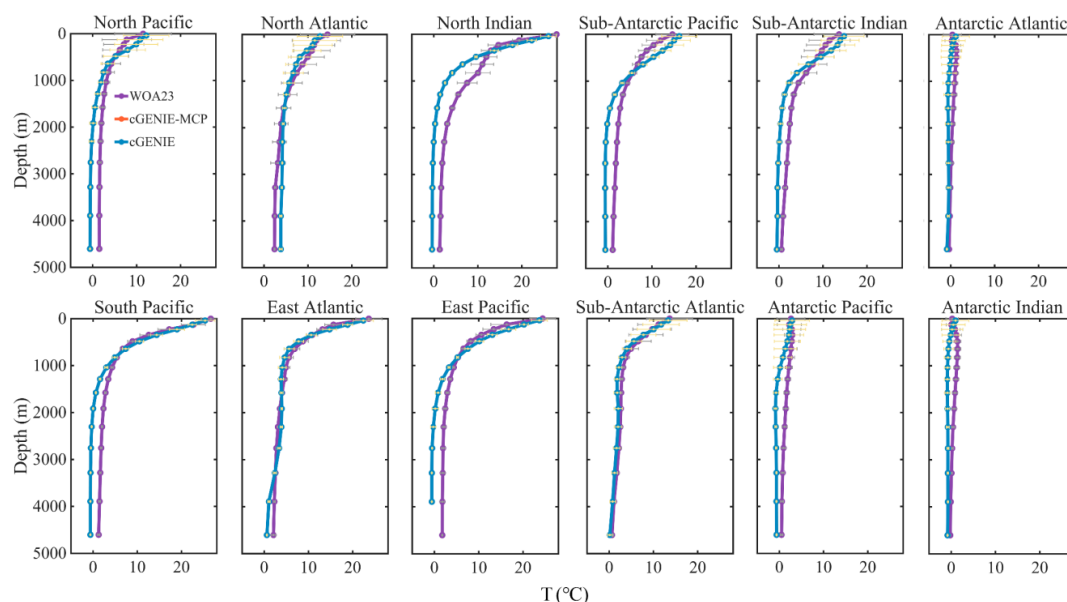


Figure 4. Vertical profiles of temperature(°C) in selected ocean regions, comparing among WOA23  
 290 data, cGENIE-MCP and standard cGENIE model outputs. Regions include the Atlantic Ocean, low-  
 latitude tropics, South Pacific, and Northern Indian Ocean, as defined in the model evaluation.

### 3.2 Phosphate

Fig. 5 illustrates the spatial distribution of  $\text{PO}_4$  concentration in both surface and deep water. Both  
 295 the cGENIE-MCP and standard cGENIE models show good agreement with WOA23 dataset. Modeled  
 $\text{PO}_4$  concentrations range from near 0 to a maximum of  $4.5 \mu\text{mol kg}^{-1}$ , with lower values observed in the  
 surface layer and upper thermocline. In surface waters of low- and mid-latitude regions, concentrations  
 vary from 0 to  $0.5 \mu\text{mol kg}^{-1}$ , exhibiting an increasing trend toward higher latitudes, consistent with prior  
 studies (Fig. 5).

300 Fig. 6 shows regional vertical profiles of  $\text{PO}_4$  concentrations, comparing WOA23 data with outputs  
 from both the cGENIE-MCP and standard cGENIE models. The cGENIE-MCP model reproduces  
 vertical structures more accurately in middle- and high-latitude regions, notably in the sub-Antarctic zone,  
 Northwest Pacific, and Eastern Pacific, where observed  $\text{PO}_4$  concentrations peak within the thermocline.  
 In these regions, the cGENIE-MCP achieves a RMSE of approximately  $0.1 \mu\text{mol kg}^{-1}$ . In the deep ocean,  
 305  $\text{PO}_4$  concentrations are lowest in the North Atlantic and highest in the North Pacific, with intermediate  
 levels in the South Indian and Southern Oceans.

Some discrepancies are observed in the intermediate and deep layers of the Indian, Atlantic, and



North Pacific Oceans, where both models tend to overestimate  $\text{PO}_4$  concentrations compared to observations. These discrepancies are consistent with those reported in previous studies (Stappard et al., 2025). The overall RMSE between cGENIE-MCP and WOA23 is  $0.1 \mu\text{mol kg}^{-1}$ , indicating that the model is capable of reproducing large-scale phosphate distributions with a high degree of accuracy.

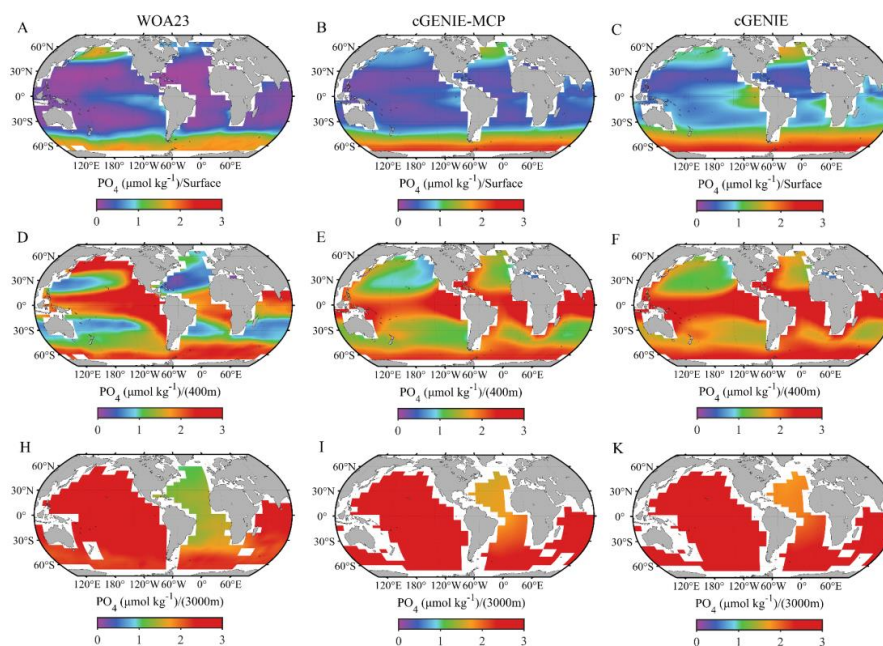


Figure 5. Spatial distributions of  $\text{PO}_4$  concentrations ( $\mu\text{mol kg}^{-1}$ ) at three depth levels: (A-C) surface, and (D-F) 400 m, and (H-J) 3000 m. Panels show: (A, D, H) observational data from WOA23; (B, E, I) simulated results from the cGENIE-MCP model; and (C, F, J) outputs from the standard cGENIE ( $\mu\text{mol kg}^{-1}$ ).

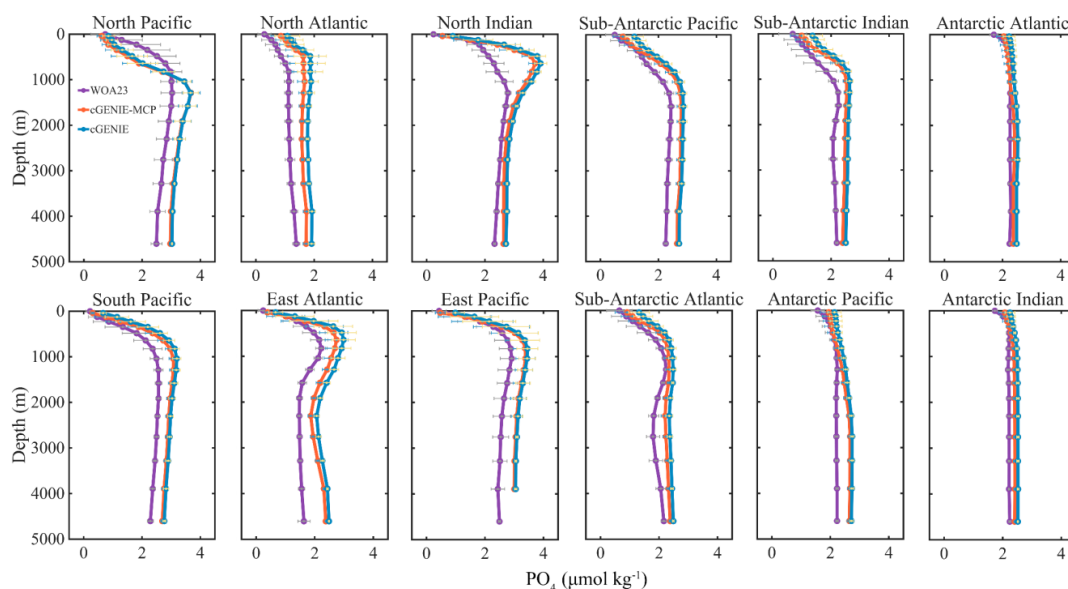


Figure 6. Vertical profiles of  $\text{PO}_4$  ( $\mu\text{mol kg}^{-1}$ ) in selected ocean regions, comparing among WOA23 data, cGENIE-MCP and standard cGENIE model outputs.

### 3.3 Dissolved oxygen

DO is a critical determinant of aerobic respiration and species distribution in marine ecosystems. Figure 7 shows that the simulated spatial distribution of DO aligns well with the WOA23 dataset, with an RMSE of 50-52  $\mu\text{mol kg}^{-1}$ . The spatial pattern of DO varies significantly in both horizontal and vertical dimensions (Alhassan et al., 2024). DO concentrations decline with depth, reaching minima within the oxygen minimum zones (OMZs) typically located between 200 and 1000 m. The cGENIE-MCP model successfully reproduces these vertical and horizontal DO gradients, particularly in the Atlantic Ocean (Fig. 8). However, moderate discrepancies are observed in the Antarctic region, potentially reflecting limitations in the simulation of regional ocean circulation.



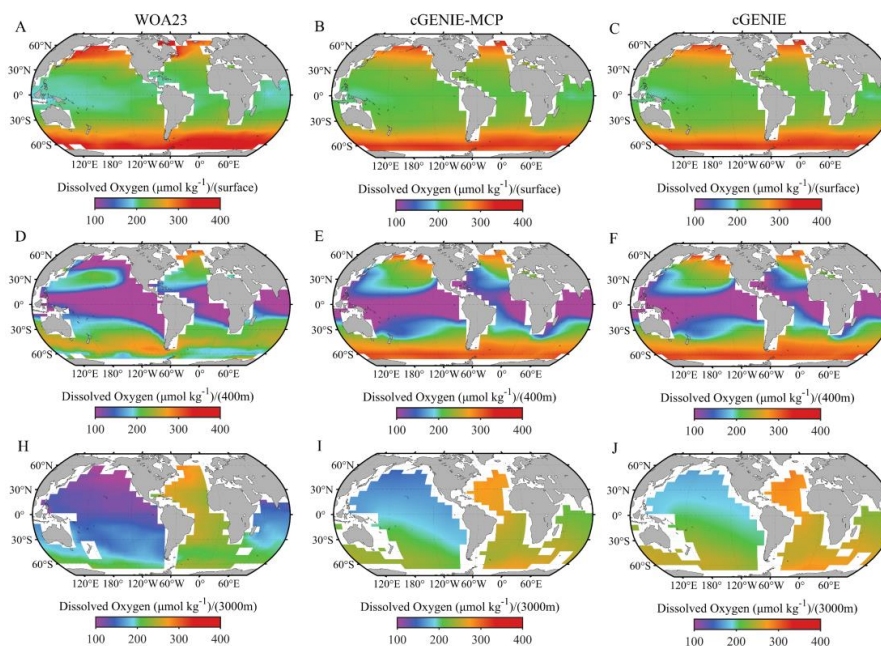
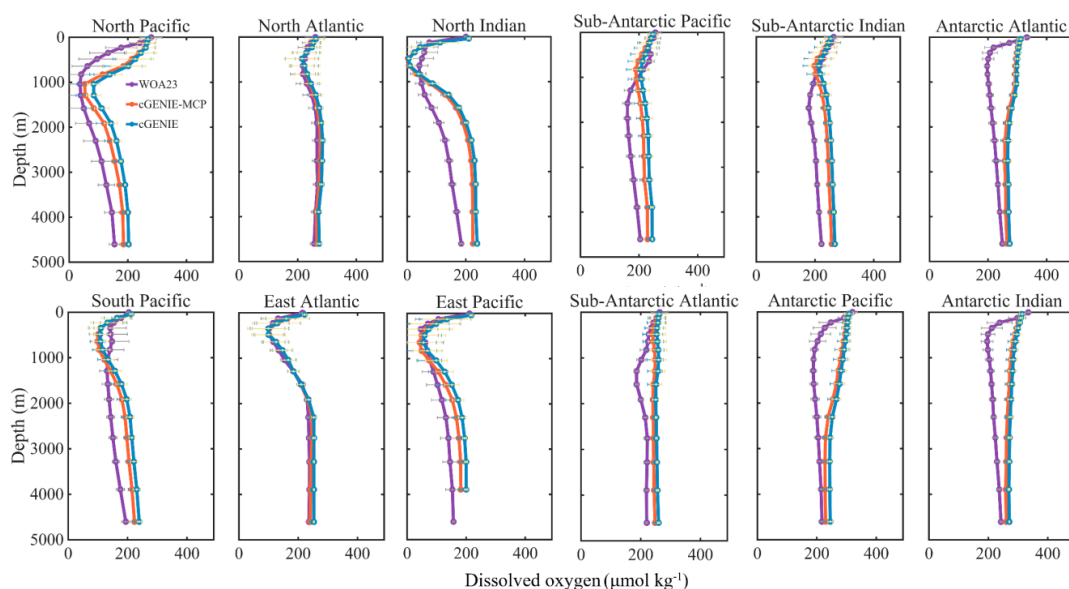


Figure 7. Spatial distributions of DO concentrations ( $\mu\text{mol kg}^{-1}$ ) at three depth levels: (A-C) surface,  
 335 and (D-F) 400 m, and (H-J) 3000 m. Panels show: (A, D, H) observational data from WOA23; (B, E, I)  
 simulated results from the cGENIE-MCP model; and (C, F, J) outputs from the standard cGENIE ( $\mu\text{mol kg}^{-1}$ ).







340 Figure 8. Vertical profiles of DO ( $\mu\text{mol kg}^{-1}$ ) in selected ocean regions, comparing among WOA23 data, cGENIE-MCP and standard cGENIE model outputs.

### 3.4 Dissolved inorganic carbon

The global distribution of surface dissolved inorganic carbon (DIC) serves as a key indicator of  
345 oceanic carbon cycling, shaped by atmospheric CO<sub>2</sub> levels, alkalinity, and biological activity. Figures 9-10 compare the simulated surface DIC concentrations from the cGENIE-MCP model with observations from the GLODAP dataset. The model generally underestimates DIC concentrations relative to observations, primarily due to the use of pre-industrial atmospheric CO<sub>2</sub> levels (278 ppm), whereas present-day CO<sub>2</sub> concentrations are substantially higher. When the model is forced with modern CO<sub>2</sub>  
350 levels (400 ppm), the agreement with the GLODAP dataset improves significantly (Fig. S1). Additional factors contributing to the discrepancy include uncertainties in surface alkalinity, biological carbon uptake, and remineralization-driven carbon release (Wu et al., 2019).

Surface and subsurface alkalinity distributions are also evaluated (Figs. S2-S3). Similar to DIC, simulated alkalinity concentrations are lower than observed values, reflecting consistent biases across  
355 carbon system variables. The global pattern of surface DIC (Fig. 9) closely mirrors that of major nutrients such as phosphate (PO<sub>4</sub><sup>3-</sup>), with elevated concentrations in high-latitude regions and reduced levels in low-latitude zones. This distribution highlights the role of temperature and physical dynamics—particularly vertical mixing—in shaping surface DIC concentrations. Lower temperature and well-mixed high-latitude surface waters typically support higher DIC levels, whereas warmer, more stratified low-  
360 latitude waters exhibit lower concentrations.

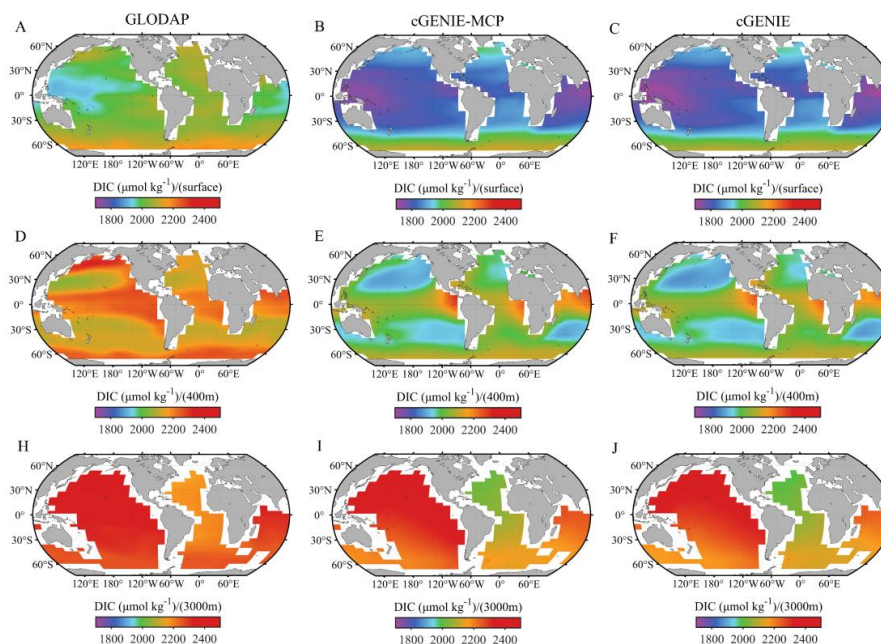


Figure 9. Spatial distributions of DIC concentrations ( $\mu\text{mol kg}^{-1}$ ) at three depth levels: (A-C) surface, and (D-F) 400 m, and (H-J) 3000 m. Panels show: (A, D, H) observational data from GLODAP; (B, E, I) simulated results from the cGENIE-MCP model; and (C, F, J) outputs from the standard cGENIE ( $\mu\text{mol kg}^{-1}$ ).

365

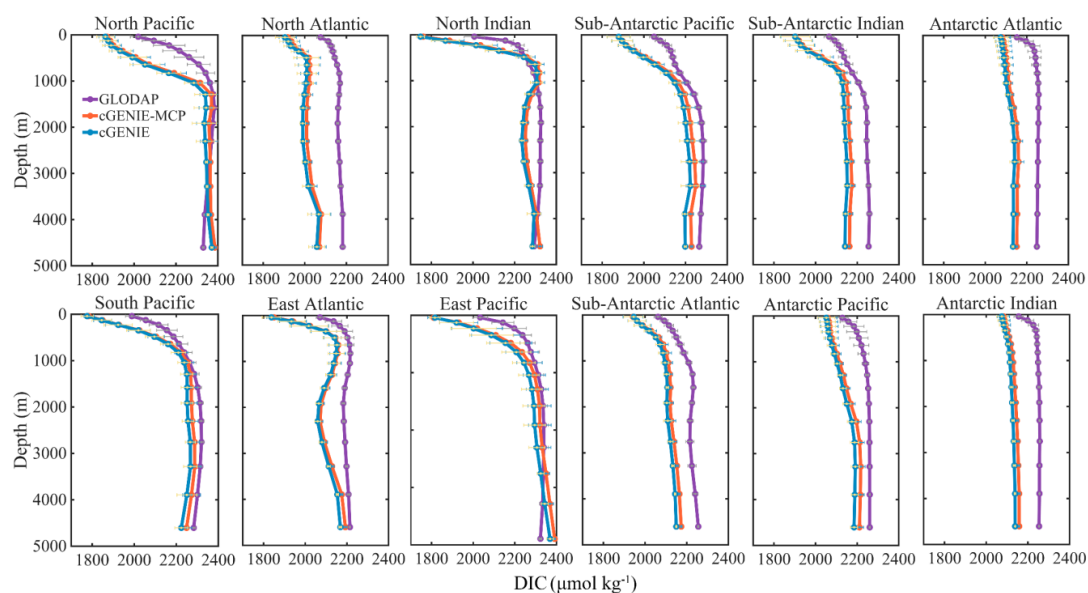




Figure 10. Vertical profiles of DIC ( $\mu\text{mol kg}^{-1}$ ) in selected ocean regions, comparing among GLODAP data, cGENIE-MCP and standard cGENIE model outputs.

### 370 3.5 Dissolved organic carbon

DOC is a key component of ocean carbon cycling, shaped by biological production, ventilation, and circulation dynamics. Figure 11 illustrates the global distribution of surface DOC concentrations simulated by the cGENIE-MCP model and compares them with observational data. Simulated surface DOC concentrations range from 40 to 80  $\mu\text{mol kg}^{-1}$ , with elevated values (70–80  $\mu\text{mol kg}^{-1}$ ) occurring primarily in tropical and subtropical regions between  $\sim 40^\circ\text{N}$  and  $40^\circ\text{S}$ . These elevated concentrations reflect the combined effects of high biological production and strong vertical stratification, which limits vertical mixing and facilitates DOC accumulation in surface waters. Compared to MESMO 3—which substantially overestimates DOC concentrations, with surface values reaching 70–150  $\mu\text{mol kg}^{-1}$  and deep values of 60–70  $\mu\text{mol kg}^{-1}$ —the cGENIE-MCP model exhibits improved agreement with observed DOC distributions in both surface and deep layers. However, the model underrepresents equatorial upwelling, consistent with similar limitations in simulating temperature fields, leading to discrepancies in equatorial DOC concentrations. In contrast, lower surface DOC concentrations (40–50  $\mu\text{mol kg}^{-1}$ ) are simulated in subpolar regions and the Southern Ocean (Fig. 12), consistent with enhanced vertical mixing and the upwelling of deep DOC-poor waters (Carlson et al., 2010). In additionally, Hansell (2013) studies have indicated that in the western Sargasso Sea region, the DOC concentration at 150–200 meters is approximately 56  $\mu\text{mol kg}^{-1}$ , the SLDOC usually reaches approximately 11  $\mu\text{mol kg}^{-1}$ , and the RDOC concentration is approximately 46  $\mu\text{mol kg}^{-1}$ , which is consistent with our simulation results (Fig. S6A). The large-scale patterns of simulated DOC distributions, such as elevated values in stratified subtropical gyres and lower concentrations in high-latitude and upwelling regions, are qualitatively consistent with observations (Hansell et al., 2012; Carlson et al., 2010; Hansell, 2013), which suggest that the patterns of DOC cycling are approximately reasonable.

LDOC is rapidly remineralized and does not accumulate significantly in surface waters, with concentrations typically below 10  $\mu\text{mol kg}^{-1}$  in regions of high productivity. SLDOC concentrations range from 10 to 30  $\mu\text{mol kg}^{-1}$  in tropical and subtropical regions, peaking at approximately 34  $\mu\text{mol kg}^{-1}$ , indicative of high production efficiency in these zones (Carlson et al., 2010; Hansell and Carlson, 2014; Hansell, 2013). RDOC concentrations range from 40 to 50  $\mu\text{mol kg}^{-1}$  in the upper ocean, with maximum values found in the centers of subtropical gyres. These high concentrations are associated with convergent circulation and weak vertical mixing, which facilitate RDOC accumulation. The RDOC concentration



and its vertical variation simulated by the model are in agreement with the results of Wang et al. (2023)  
 (Figure S6B). The simulation of LDOC, SLDOC, and RDOC captures the essential patterns consistently  
 reported in observational and previous studies (Table S2) (Letscher et al., 2015; Legendre et al., 2015;  
 Hansell and Carlson, 2014; Hansell, 2013; Hansell et al., 2012; Hansell et al., 2009; Ge et al., 2022).  
 Specifically, the model reproduces the rapid turnover and low accumulation of labile fractions in  
 productive regions.

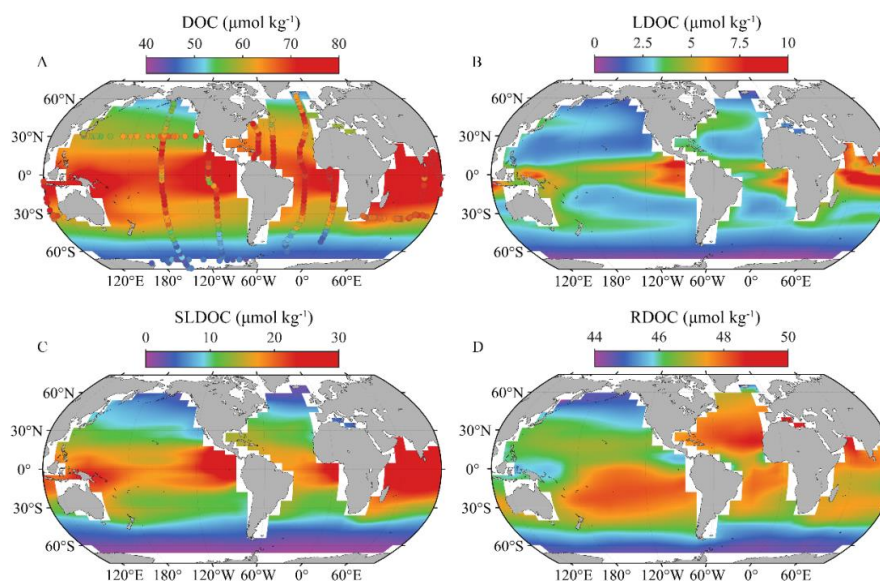


Figure 11. Global distributions of surface (A) DOC, (B) LDOC, (C) SLDOC, (D) RDOC concentrations ( $\mu\text{mol kg}^{-1}$ ). Shaded represents cGENIE-MCP model outputs, overlaid with scatter points of observed values from Hansell's laboratory (Hansell, 2013).

Figure 12 shows the vertical distribution of modeled DOC concentrations compared with  
 observational data. In the euphotic zone, mean DOC concentrations exceed  $60 \mu\text{mol kg}^{-1}$ , with maxima  
 over  $70 \mu\text{mol kg}^{-1}$  observed in the Atlantic, Pacific, and Indian Oceans. Within the mesopelagic and  
 bathypelagic zones, DOC concentrations range from  $40$  to  $60 \mu\text{mol kg}^{-1}$ , with smoother vertical gradients  
 at high latitudes due to reduced remineralization rates, strong convective mixing, and enhanced isopycnal  
 ventilation.

In the North Atlantic, the formation of North Atlantic Deep Water (NADW) facilitates the transport  
 of DOC to depth, with concentrations reaching up to  $47 \mu\text{mol kg}^{-1}$  below  $2000 \text{ m}$ . These concentrations  
 decline progressively to approximately  $42 \mu\text{mol kg}^{-1}$  toward the southern Atlantic, in agreement with the



classical pattern of Atlantic meridional overturning circulation. However, the modeled deep DOC  
 concentrations in both the Atlantic and Indian Oceans tend to be higher than observed values, potentially  
 reflecting limitations in the representation of DOC degradation or ventilation processes. In contrast, the  
 modeled DOC concentrations in the deep Pacific ( $\sim 40 \mu\text{mol kg}^{-1}$ ) are consistent with observations,  
 reflecting the progressive degradation of DOC during the northward flow of deep waters from the  
 Southern Ocean and their subsequent return southward through the mesopelagic layer (Hansell, 2013).

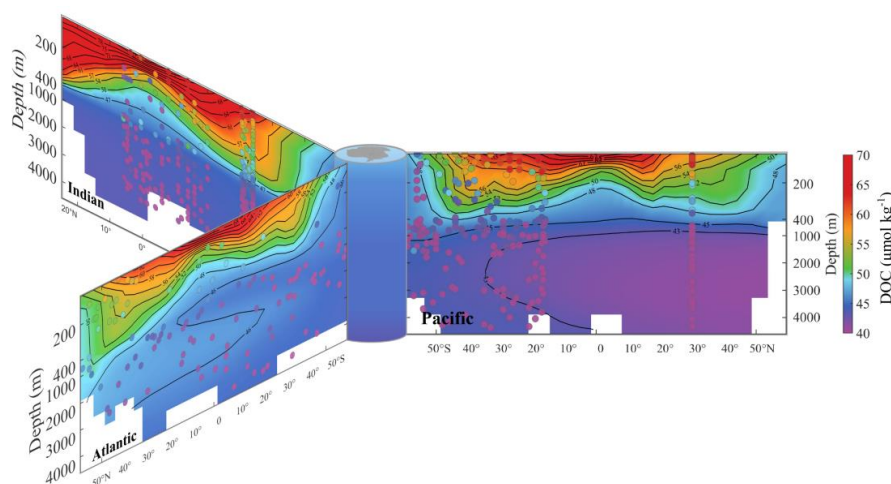


Figure 12. Distributions of DOC ( $\mu\text{mol kg}^{-1}$ ) along meridional sections in the central Atlantic, central  
 Pacific, and western Indian oceans, respectively. The shaded represents cGENIE-MCP model outputs,  
 overlaid with scatter points of observed values from Hansell's laboratory (Hansell, 2013).

## 4 Discussion

### 4.1 Effect of DOC partitioning

The distribution of DOC and its constituent fractions provides essential insights into the production,  
 degradation, and transport processes that govern the oceanic carbon cycle. The cGENIE-MCP model  
 effectively simulates the vertical transformation of labile DOC (LDOC), which exhibits a global mean  
 surface concentration of approximately  $10 \mu\text{mol} \cdot \text{kg}^{-1}$ . The rapid microbial consumption makes LDOC  
 concentrations approach zero below the euphotic zone (Fig. 13). This rapid cycling underscores LDOC's  
 role in supporting microbial loop dynamics in surface waters. LDOC is also primarily exits in equatorial  
 regions, further reflecting its rapid turnover driven by the microbial loop process. SLDOC is abundant in  
 the euphotic zone, particularly in tropical and subtropical regions, where it supports heterotrophic  
 bacterial carbon demand. It is laterally redistributed by subtropical gyres and is typically exported to the





upper mesopelagic zone (100-1000 m) due to its moderate lability. With a lifetime of months to years, SLDOC is either remineralized or converted into RDOC during its transit. The simulated spatial extent of SLDOC extends slightly deeper and spans a broader latitudinal range, particularly in the subtropics, suggesting its role as a transitional carbon reservoir that can be exported through physical processes such as subduction and mesoscale eddy transport.

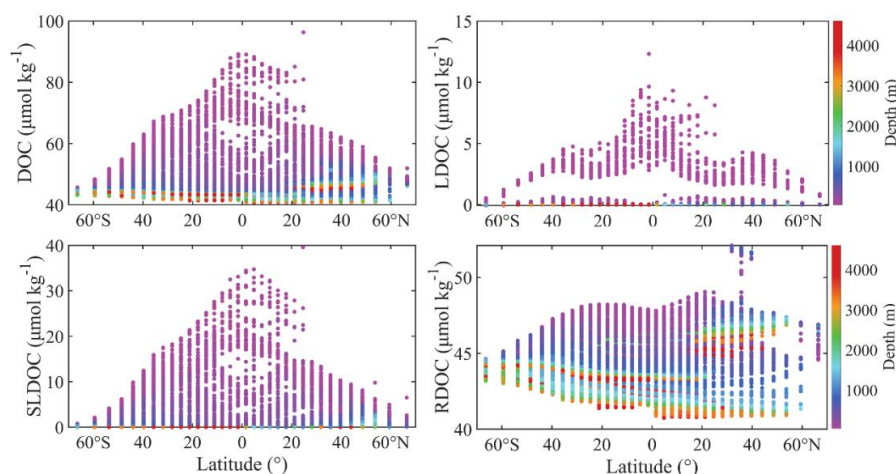


Figure 13. Meridional distributions of mean DOC concentrations (A), mean LDOC concentrations (B), mean SLDOC concentrations (C), mean RDOC concentrations (D) ( $\mu\text{mol kg}^{-1}$ ).

In contrast, RDOC is ubiquitous throughout the ocean interior, with relatively uniform concentrations ranging from 40 to 50  $\mu\text{mol kg}^{-1}$  in deep waters. This homogeneity reflects its recalcitrance to microbial degradation and highlights its significance as a long-term carbon reservoir. Slightly elevated RDOC concentrations in high-latitude deep waters may reflect entrainment of surface-derived labile and semi-labile DOC during deep water formation and ventilation. These results underscore the dynamic role of LDOC and SLDOC as intermediaries in the microbial carbon pump (MCP), mediating the transformation of labile carbon into more persistent forms that contribute to long-term carbon sequestration.

DOC concentrations are highest in the tropical and subtropical surface waters and decrease with depth and latitude. This pattern reflects enhanced primary production in low-latitude regions and limited vertical mixing due to strong stratification, allowing DOC to accumulate in surface layers. In contrast, high-latitude regions exhibit lower DOC concentrations due to weaker stratification and lower biological production. The distribution of DOC reflects the functional partitioning of the oceanic carbon reservoir-



from rapidly cycling, bioavailable fractions to long-lived, deeply sequestered refractory pools-  
 highlighting the essential roles of biological production, microbial transformation, and physical transport  
 465 in shaping carbon dynamics and supporting long-term ocean carbon sequestration.

In this study, the cGENIE-MCP model has been extended to incorporate temperature-dependent  
 nutrient uptake and remineralization of sinking organic matter. These enhancements allow the model to  
 more accurately capture observed DOC distributions and improve the representation of spatial and  
 vertical gradients. Furthermore, integrating temperature-sensitive DOC cycling processes enables the  
 470 cGENIE-MCP to simulate the dynamic feedbacks between biological activity, ocean carbon sequestration,  
 and climate warming (Crichton et al., 2021). Partitioning DOC into LDOC, SLDOC, and RDOC provides  
 a mechanistic framework for evaluating the efficiency of the microbial carbon pump under future climate  
 scenarios and contributes to a deeper understanding of the ocean's role in long-term carbon storage.

## 475 **4.2 Distribution of DOC production**

Primary production (PP) is driven by the availability of nutrients and light within the euphotic zone,  
 where phytoplankton convert inorganic carbon into organic matters. The spatial distribution  
 characteristics of PP arise from abiotic controls (e.g., light, nutrient upwelling, stratification) and biotic  
 feedbacks (e.g., grazing pressure, microbial loop), shaping global productivity gradients from nutrient-  
 480 rich coastal upwelling zones to oligotrophic gyres. The DOC originates from biologically mediated  
 processes in the euphotic zone including phytoplankton extracellular release, top-down pressures (grazers,  
 bacterial and viral lysis), and bottom-up recycling (particle solubilization, prokaryotic excretion). Spatial-  
 temporal variations in DOC emerge from the interplay of these biological and abiotic drivers.

Figure 14 illustrates the spatial distributions of PP and DOC production (LDOC, SLDOC, RDOC)  
 485 simulated by the cGENIE-MCP model, revealing region-specific variability driven by nutrient supply,  
 light availability, hydrodynamic processes, and biological community structure. In equatorial-tropical  
 ocean regions, PP and LDOC production ( $P_{LDOC}$ ) remain relatively high, reaching values of 400-1000 mg  
 $C\ m^{-2}\ d^{-1}$  and 300-800 mg  $C\ m^{-2}\ d^{-1}$ , respectively (Fig. 14). The spatial coherence between high  $P_{LDOC}$  and  
 high-PP zones indicates that LDOC is primarily derived from PP. SLDOC production ( $P_{SLDOC}$ ) ranges  
 490 from 20 to 50 mg  $C\ m^{-2}\ d^{-1}$ , closely mirroring PP and  $P_{LDOC}$  patterns, reflecting a progressive attenuation  
 of DOC bioavailability from labile to semi-labile fractions. RDOC production ( $P_{RDOC}$ ), ranging from 0.4  
 to 0.8 mg  $C\ m^{-2}\ d^{-1}$  and constituting approximately 0.1-0.2% of PP, exhibits a relatively uniform spatial  
 distribution, consistent with Hansell (2013) and Legendre et al. (2015) studies (Table S3). This uniformity



underscores RDOC's role in long-term biogeochemical cycling and carbon sequestration.

495

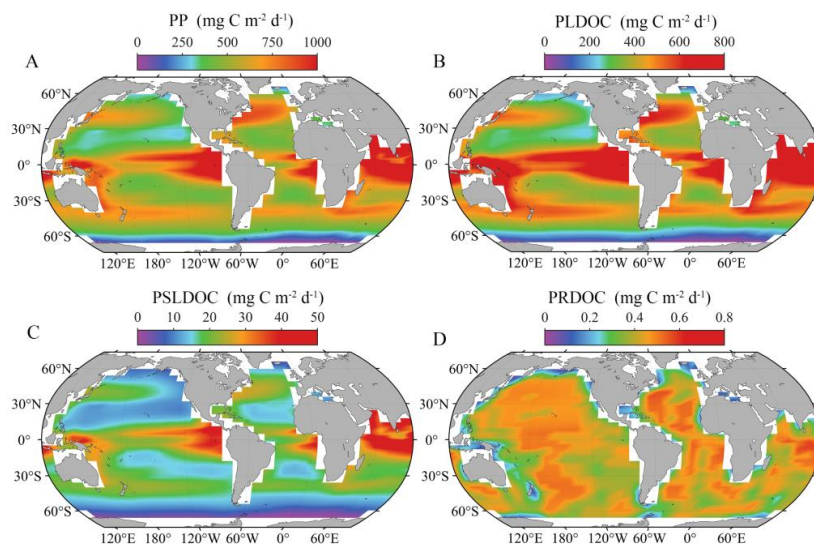


Figure 14. Global distributions of the production rate for (A) DOC, (B) LDOC, (C) SLDOC, (D) RDOC ( $\text{mg C m}^{-2} \text{d}^{-1}$ ), respectively.

## 500 5 Perspectives

Several "Snowball Earth" events occurred throughout geological history. According to the Snowball Earth hypothesis, the biogeochemical cycle and the PP have severely slowed down or even stagnated under global freezing conditions. However, previous studies have found PP and DOC reservoirs still exist during glaciations (Jiao et al., 2024a; Man et al., 2024). These results suggest that organic matter produced in the surface ocean may have been degraded into DOC or RDOC in the water column. Therefore, understanding the changes of the DOC pools during Snowball Earth periods is of great significance (Hoffman and Schrag, 2002; Hoffman et al., 2017; Sansjofre et al., 2011). Nevertheless, the storage, spatial distribution, and source-sink processes of DOC during "Snowball Earth" remain largely unexplored. The cGENIE-MCP model proposed in this study can simulate the DOC dynamics over geological historical spans and explore the relationship between DOC dynamics and long-term climate change. It is possible to analyze the relationship between  $\delta^{13}\text{C}$  negative excursion and carbon pumps in geological records on a global scale, quantify the efficiency of MCP, and evaluate the impact of ocean environmental changes on the distribution of DOC.





Furthermore, reducing emissions and enhancing carbon sinks have become a global consensus in response to global warming, with ocean carbon sinks playing a vital role in achieving this goal. Previous studies have pointed out that both "Snowball Earth" events and glacial-interglacial cycles are not only driven by orbital forcing but are also influenced by the ocean carbon cycle (Hoffman et al., 2017). The primary reason is due to the huge RDOC, which can remain in the ocean for tens of thousands of years and also release carbon into the atmosphere, thereby regulating climate change. Therefore, evaluating the efficiency of the MCP is crucial for understanding of long-term climate regulation. The cGENIE-MCP model offers selectable modular combination methods and is suitable for long-term climate research. For example, by combining SSP scenarios, the response of MCP to CO<sub>2</sub> concentration will be investigated and reveal the feedback mechanism between climate change and the ocean carbon cycle. Evaluating the application potential of ocean negative carbon emission technologies (e.g., ocean alkalization enhancement technology) under different climate scenarios. Analyzing the long-term effects of ocean negative carbon emission technologies on the environment by simulating different implementation pathways, and comprehensively evaluating the optimal implementation scenarios and strategies for the coupled of ocean carbon pump.

## 6 Conclusions

This paper developed an RDOC parameterized-based method of the MCP process and embedded a temperature-dependent process in cGENIE model and successfully incorporated RDOC as a tracer. The model outputs towards cGENIE-MCP and cGENIE were evaluated and verified the reasonable simulation performance of the models via WOA 2023 products and GLODAP datasets. We derive the spatial distribution of concentrations and production rates of LDOC, SLDOC, and RDOC. The LDOC and SLDOC were hardly accumulated in the surface layer due to their rapid degradation and exhibited a lower concentration than RDOC. The production rates reveal the spatial distribution of LDOC, SLDOC, and RDOC strongly associated with primary production and indicate the critical role of bottom-up control driven by the nutrient cycle in pelagic ecosystems and the primary production process. This study demonstrated that RDOC is the most vital component contributing to carbon sequestration. With this new extension, there is a potential for further study regarding the role of RDOC during future and past climatological perturbations and the role in the biological pump.

## Code availability



545 The code of cGENIE-MCP model used in this paper is available at  
<https://doi.org/10.5281/zenodo.15734206> (Ma and Lai, 2025).

Configurations for the specific experiments presented in the paper can be found in the file  
 README.md at <https://doi.org/10.5281/zenodo.15734206> (Ma and Lai, 2025).

## 550 **Author contributions**

WM and NJ designed the MCP module. YL and WM developed the code, ran, and analysed the model  
 ensemble and data, and all authors co-wrote the paper.

## **Acknowledgements**

555 We are grateful to Prof. Weilei Wang who provided parameters on DOC transformation.

## **Financial support**

This study was supported by the Ocean Negative Carbon Emissions (ONCE) Program. National. Key  
 Research and Development Program of China (2023YFF0805004), National Natural Science Foundation  
 560 of China (No. 42276063).

## **References**

- Alhassan, Y., Siekmann, I., and Petrovskii, S.: Mathematical model of oxygen minimum zones in the vertical distribution  
 of oxygen in the ocean, *Scientific Reports*, 14, 22248, <https://doi.org/10.1038/s41598-024-72207-3>, 2024.
- 565 Anagnostou, E., John, E. H., Edgar, K. M., Foster, G. L., Ridgwell, A., Inglis, G. N., Pancost, R. D., Lunt, D. J., and Pearson,  
 P. N.: Changing atmospheric CO<sub>2</sub> concentration was the primary driver of early Cenozoic climate, *Nature*, 533, 380-  
 384, <https://doi.org/10.1038/nature17423>, 2016.
- Anderson, T. R., Christian, J. R., and Flynn, K. J.: Modeling DOM biogeochemistry, in: *Biogeochemistry of marine  
 dissolved organic matter*, Elsevier, 635-667, 2015.
- 570 Boscolo-Galazzo, F., Crichton, K. A., Ridgwell, A., Mawbey, E. M., Wade, B. S., and Pearson, P. N.: Temperature controls  
 carbon cycling and biological evolution in the ocean twilight zone, *Science*, 371, 1148-1152,  
<https://doi.org/10.1126/science.abb6643>, 2021.
- Cao, L., Eby, M., Ridgwell, A., Caldeira, K., Archer, D., Ishida, A., Joos, F., Matsumoto, K., Mikolajewicz, U., and Mouchet,  
 A.: The role of ocean transport in the uptake of anthropogenic CO<sub>2</sub>, *Biogeosciences*, 6, 375-390,  
 575 <https://doi.org/10.5194/bg-6-375-2009>, 2009.
- Carlson, C. A., Hansell, D. A., Nelson, N. B., Siegel, D. A., Smethie, W. M., Khaliwala, S., Meyers, M. M., and Halewood,  
 E.: Dissolved organic carbon export and subsequent remineralization in the mesopelagic and bathypelagic realms of  
 the North Atlantic basin, *Deep Sea Research Part II: Topical Studies in Oceanography*, 57, 1433-1445,  
<https://doi.org/10.1016/j.dsr2.2010.02.013>, 2010.
- 580 Crichton, K. A., Wilson, J. D., Ridgwell, A., and Pearson, P. N.: Calibration of temperature-dependent ocean microbial  
 processes in the cGENIE. muffin (v0. 9.13) Earth system model, *Geoscientific Model Development*, 14, 125-149,



- <https://doi.org/10.5194/gmd-14-125-2021>, 2021.
- Crichton, K. A., Wilson, J. D., Ridgwell, A., Boscolo-Galazzo, F., John, E. H., Wade, B. S., and Pearson, P. N.: What the geological past can tell us about the future of the ocean's twilight zone, *Nature Communications*, 14, 2376, <https://doi.org/10.1038/s41467-023-37781-6>, 2023.
- DeVries, T.: The ocean carbon cycle, *Annual Review of Environment and Resources*, 47, 317-341, <https://doi.org/10.1146/annurev-environ-120920-111307>, 2022.
- Ge, T., Luo, C., Ren, P., Zhang, H., Chen, H., Chen, Z., Zhang, J., and Wang, X.: Dissolved organic carbon along a meridional transect in the western north pacific ocean: Distribution, variation and controlling processes, *Frontiers in Marine Science*, 9, 909148, <https://doi.org/10.3389/fmars.2022.909148>, 2022.
- Gruber, N., Bakker, D. C., DeVries, T., Gregor, L., Hauck, J., Landschützer, P., McKinley, G. A., and Müller, J. D.: Trends and variability in the ocean carbon sink, *Nature Reviews Earth & Environment*, 4, 119-134, <https://doi.org/10.1038/s43017-022-00381-x>, 2023.
- Hansell, D. A.: Recalcitrant dissolved organic carbon fractions, *Ann Rev Mar Sci*, 5, 421-445, <https://doi.org/10.1146/annurev-marine-120710-100757>, 2013.
- Hansell, D. A. and Carlson, C. A.: Biogeochemistry of marine dissolved organic matter, Academic press 2014.
- Hansell, D. A., Carlson, C. A., and Schlitzer, R.: Net removal of major marine dissolved organic carbon fractions in the subsurface ocean, *Global Biogeochemical Cycles*, 26, <https://doi.org/10.1029/2011GB004069>, 2012.
- Hansell, D. A., Carlson, C. A., Repeta, D. J., and Schlitzer, R.: Dissolved organic matter in the ocean: A controversy stimulates new insights, *Oceanography*, 22, 202-211, <https://www.jstor.org/stable/24861036>, 2009.
- Hoffman, P. F. and Schrag, D. P.: The snowball Earth hypothesis: testing the limits of global change, *Terra nova*, 14, 129-155, <https://doi.org/10.1046/j.1365-3121.2002.00408.x>, 2002.
- Hoffman, P. F., Abbot, D. S., Ashkenazy, Y., Benn, D. I., Brocks, J. J., Cohen, P. A., Cox, G. M., Creveling, J. R., Donnadieu, Y., and Erwin, D. H.: Snowball Earth climate dynamics and Cryogenian geology-geobiology, *Science Advances*, 3, e1600983, <https://doi.org/10.1126/sciadv.1600983>, 2017.
- Jiao, N., Du, J., Xiao, X., and Tian, J.: Addressing the 'Negative  $\delta^{13}\text{C}$  arms' puzzle pre-and post-Neoproterozoic glaciations, *The Innovation Geoscience*, 2, 100081-100081-100081-100082, <https://doi.org/10.59717/j.xinn-geo.2024.100081>, 2024a.
- Jiao, N., Herndl, G. J., Hansell, D. A., Benner, R., Kattner, G., Wilhelm, S. W., Kirchman, D. L., Weinbauer, M. G., Luo, T., and Chen, F.: Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean, *Nature Reviews Microbiology*, 8, 593-599, <https://doi.org/10.1038/nrmicro2386>, 2010.
- Jiao, N., Luo, T., Chen, Q., Zhao, Z., Xiao, X., Liu, J., Jian, Z., Xie, S., Thomas, H., and Herndl, G. J.: The microbial carbon pump and climate change, *Nature reviews microbiology*, 1-12, <https://doi.org/10.1038/s41579-024-01018-0>, 2024b.
- Jokulsdottir, T.: Sinking biological aggregates in the ocean-a modeling study, The University of Chicago 2011.
- Keller, D. P., Feng, E. Y., and Oschlies, A.: Potential climate engineering effectiveness and side effects during a high carbon dioxide-emission scenario, *Nature Communications*, 5, <https://doi.org/10.1038/ncomms4304>, 2014.
- Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T. H.: A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP), *Global biogeochemical cycles*, 18, <https://doi.org/10.1029/2004GB002247>, 2004.
- Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Álvarez-Rodríguez, M., Azetsu-Scott, K., Becker, S., and Brown, P. J.: Global Ocean Data Analysis Project version 2.2023 (GLODAPv2. 2023), <http://doi.org/10.25921/ZYRQ-HT66>, 2023.
- Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Álvarez, M., Azetsu-Scott, K., Brown, P. J., and Carter, B. R.: The annual update GLODAPv2. 2023: the global interior ocean biogeochemical data product, *Earth System Science Data*, 16, 2047-2072, <https://doi.org/10.5194/essd-16-2047-2024>, 2024.
- Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Álvarez, M., Becker, S., Brown, P. J., and Carter, B. R.: An updated version of the global interior ocean biogeochemical data product, GLODAPv2. 2021, *Earth System*



- Science Data Discussions, 2021, 1-32, <https://doi.org/10.5194/essd-13-5565-2021>, 2021.
- Legendre, L., Rivkin, R. B., Weinbauer, M. G., Guidi, L., and Uitz, J.: The microbial carbon pump concept: Potential  
 630 biogeochemical significance in the globally changing ocean, *Progress in Oceanography*, 134, 432-450,  
<https://doi.org/10.1016/j.pocean.2015.01.008>, 2015.
- Lenton, T., Williamson, M., Edwards, N., Marsh, R., Price, A., Ridgwell, A., Shepherd, J., Cox, S., and team, G.: Millennial  
 timescale carbon cycle and climate change in an efficient Earth system model, *Climate Dynamics*, 26, 687-711,  
<https://doi.org/10.1007/s00382-006-0109-9>, 2006.
- 635 Letscher, R. T., Moore, J. K., Teng, Y. C., and Primeau, F.: Variable C : N : P stoichiometry of dissolved organic matter  
 cycling in the Community Earth System Model, *Biogeosciences*, 12, 209-221, <https://doi.org/10.5194/bg-12-209-2015>,  
 2015.
- Li, H., Ilyina, T., Müller, W. A., and Landschützer, P.: Predicting the variable ocean carbon sink, *Science advances*, 5,  
 eaav6471, <https://doi.org/10.1126/sciadv.aav6471>, 2019.
- 640 Model for cGENIE-MCP: <https://doi.org/10.5281/zenodo.15734207>, last  
 Ma, W. and Tian, J.: Modeling the contribution of dissolved organic carbon to carbon sequestration during the last  
 glacial maximum, *Geo-Marine Letters*, 34, 471-482, <https://doi.org/10.1007/s00367-014-0378-y>, 2014.
- Ma, W., Xiu, P., Yu, Y., Zheng, Y., and Chai, F.: Production of dissolved organic carbon in the South China Sea: A  
 modeling study, *Science China Earth Sciences*, 1-14, <https://doi.org/10.1007/s11430-021-9817-2>, 2022.
- 645 Man, I., Zhu, S., Deng, B., Li, Y., and Lang, X.: Active Marine Carbon Cycle During the Sturtian "Snowball Earth" Glaciation,  
*Acta Sedimentologica Sinica*, 42, 29-38, <https://doi.org/10.14027/j.issn.1000-0550.2022.025>, 2024.
- Marsh, R., Müller, S., Yool, A., and Edwards, N.: Incorporation of the C-GOLDSTEIN efficient climate model into the  
 GENIE framework:" eb\_go\_gs" configurations of GENIE, *Geoscientific Model Development*, 4, 957-992,  
<https://doi.org/10.5194/gmd-4-957-2011>, 2011.
- 650 Matsumoto, K., Tanioka, T., and Zahn, J.: MESMO 3: Flexible phytoplankton stoichiometry and refractory dissolved  
 organic matter, *Geoscientific Model Development*, 14, 2265-2288, <https://doi.org/10.5194/gmd-14-2265-2021>, 2021.
- Matsumoto, K., Tokos, K., Price, A., and Cox, S.: First description of the Minnesota Earth System Model for ocean  
 biogeochemistry (MESMO 1.0), *Geoscientific Model Development*, 1, 1-15, <https://doi.org/10.5194/gmd-1-1-2008>,  
 2008a.
- 655 Matsumoto, K., Tokos, K., Price, A., and Cox, S.: GENIE-M: a new and improved GENIE-1 developed in Minnesota,  
*Geoscientific Model Development Discussions*, 1, 1-37, <https://hal.science/hal-00298260v1>, 2008b.
- Najjar, R. G., Sarmiento, J. L., and Toggweiler, J.: Downward transport and fate of organic matter in the ocean:  
 Simulations with a general circulation model, *Global Biogeochemical Cycles*, 6, 45-76,  
<https://doi.org/10.1029/91GB02718>, 1992.
- 660 Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., Doney, S. C., Dunne, J. P., Dutay, J.-C., and  
 Graven, H.: Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP),  
*Geoscientific Model Development*, 10, 2169-2199, <https://doi.org/10.5194/gmd-10-2169-2017>, 2017.
- Polimene, L., Rivkin, R. B., Luo, Y.-W., Kwon, E. Y., Gehlen, M., Pena, M. A., Wang, N., Liang, Y., Kaartokallio, H., and Jiao,  
 N.: Modelling marine DOC degradation time scales, *National Science Review*, 5, 468-474,  
 665 <https://doi.org/10.1093/nsr/nwy066>, 2018.
- Reinhard, C. T., Olson, S. L., Kirtland Turner, S., Pälike, C., Kanzaki, Y., and Ridgwell, A.: Oceanic and atmospheric  
 methane cycling in the cGENIE Earth system model-release v0. 9.14, *Geoscientific Model Development*, 13, 5687-5706,  
<https://doi.org/10.5194/gmd-13-5687-2020>, 2020.
- Reinhard, C. T., Olson, S. L., Turner, S. K., Pälike, C., Kanzaki, Y., and Ridgwell, A.: Oceanic and atmospheric methane  
 cycling in the cGENIE Earth system model 2, system (Hitchcock and Lovelock, 1967; Sagan et al., 1993; Krissansen-  
 Totton et al., 2018), 63, 64, <https://doi.org/10.5194/gmd-2020-32>,  
 670 Ridgwell, A. and Arndt, S.: Why dissolved organics matter: DOC in ancient oceans and past climate change, in:  
*Biogeochemistry of marine dissolved organic matter*, Elsevier, 1-20, <https://doi.org/10.1016/B978-0-12-405940->



- 5.00001-7, 2015.
- 675 Ridgwell, A. and Zeebe, R. E.: The role of the global carbonate cycle in the regulation and evolution of the Earth system, *Earth and Planetary Science Letters*, 234, 299–315, <https://doi.org/10.1016/j.epsl.2005.03.006>, 2005.
- Ridgwell, A., Hargreaves, J., Edwards, N. R., Annan, J., Lenton, T. M., Marsh, R., Yool, A., and Watson, A.: Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling, *Biogeosciences*, 4, 87–104, <https://doi.org/10.5194/bg-4-87-2007>, 2007.
- 680 Ridgwell, A. J.: Glacial-interglacial perturbations in the global carbon cycle, Univ. of East Anglia, Sansjofre, P., Ader, M., Trindade, R. I. F. d., Elie, M., Lyons, J., Cartigny, P., and Nogueira, A. C. R.: A carbon isotope challenge to the snowball Earth, *Nature*, 478, 93–96, <https://doi.org/10.1038/nature10499>, 2011.
- Sarmiento, J. L.: Ocean biogeochemical dynamics, in: *Ocean biogeochemical dynamics*, Princeton university press, <https://doi.org/10.1515/9781400849079>, 2013.
- 685 Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J., and Dunne, J.: Tracking improvement in simulated marine biogeochemistry between CMIP5 and CMIP6, *Current Climate Change Reports*, 6, 95–119, <https://doi.org/10.1007/s40641-020-00160-0>, 2020.
- Stappard, D. A., Wilson, J. D., Yool, A., and Tyrrell, T.: NutGENIE 1.0: nutrient cycle extensions to the cGENIE Earth system model to examine the long-term influence of nutrients on oceanic primary production, *EGUsphere*, 2025, 1–33, <https://doi.org/10.5194/egusphere-2025-436>, 2025.
- 690 Tanioka, T., Matsumoto, K., and Lomas, M. W.: Drawdown of atmospheric pCO<sub>2</sub> via variable particle flux stoichiometry in the ocean twilight zone, *Geophysical Research Letters*, 48, e2021GL094924, <https://doi.org/10.1029/2021GL094924>, 2021.
- van de Velde, S. J., Hülse, D., Reinhard, C. T., and Ridgwell, A.: Iron and sulfur cycling in the cGENIE. muffin Earth system model (v0. 9.21), *Geoscientific Model Development*, 14, 2713–2745, <https://doi.org/10.5194/gmd-14-2713-2021>, 2021.
- 695 Vervoort, P., Kirtland Turner, S., Rochholz, F., and Ridgwell, A.: Earth system model analysis of how astronomical forcing is imprinted onto the marine geological record: The role of the inorganic (carbonate) carbon cycle and feedbacks, *Paleoceanography and Paleoclimatology*, 39, e2023PA004826, <https://doi.org/10.1029/2023PA004826>, 2024.
- 700 Wang, W. L., Fu, W., Le Moigne, F. A. C., Letscher, R. T., Liu, Y., Tang, J. M., and Primeau, F. W.: Biological carbon pump estimate based on multidecadal hydrographic data, *Nature*, 624, 579–585, <https://doi.org/10.1038/s41586-023-06772-4>, 2023.
- Ward, B. A., Wilson, J. D., Death, R. M., Monteiro, F. M., Yool, A., and Ridgwell, A.: EcoGENIE 1.0: plankton ecology in the cGENIE Earth system model, *Geoscientific Model Development*, 11, 4241–4267, <https://doi.org/10.5194/gmd-11-4241-2018>, 2018.
- 705 Williamson, M. S., Lenton, T., Shepherd, J., and Edwards, N.: An efficient numerical terrestrial scheme (ENTS) for Earth system modelling, *Ecological Modelling*, 198, 362–374, <https://doi.org/10.1016/j.ecolmodel.2006.05.027>, 2006.
- Wu, Y., Hain, M. P., Humphreys, M. P., Hartman, S., and Tyrrell, T.: What drives the latitudinal gradient in open-ocean surface dissolved inorganic carbon concentration?, *Biogeosciences*, 16, 2661–2681, <https://doi.org/10.5194/bg-16-2661-2019>, 2019.
- 710 Yamashita, Y. and Tanoue, E.: Production of bio-refractory fluorescent dissolved organic matter in the ocean interior, *Nature Geoscience*, 1, 579–582, <https://doi.org/10.1038/ngeo279>, 2008.
- Yan, C., Lu, Y., Yuan, X., Lai, H., Wang, J., Fu, W., Yang, Y., and Li, F.: Simulation of Vertical Water Temperature Distribution in a Megareservoir: Study of the Xiaowan Reservoir Using a Hybrid Artificial Neural Network Modeling Approach, *Journal of Hydrologic Engineering*, 29, 04024047, <https://doi.org/10.1061/JHYEFF.HEENG-6219>, 2024.
- 715