

Supplementary Material (Model)

Equations governing heterotrophic bacterial metabolism

Overall growth of bacteria, G (mol C kg⁻¹ s⁻¹), is defined by

$$G = \mu B \quad \text{Eq. S1}$$

Where μ is the per capita growth rate of bacteria (s⁻¹) and B is the biomass of bacteria in concentration units (mol C kg⁻¹).

$$\mu = \min(\mu(S), \mu(dFe), \mu(N), \mu(O_2)) \beta^T \quad \text{Eq. S2}$$

Where:

- $\mu(S)$ is the growth rate associated with organic carbon substrate assimilation (s⁻¹).
- $\mu(dFe)$ is the growth rate associated with dissolved iron assimilation (s⁻¹).
- $\mu(N)$ is the growth rate associated with ammonium assimilation (s⁻¹), if required.
- $\mu(O_2)$ is the growth rate associated with dissolved oxygen assimilation (s⁻¹).
- β^T is the temperature-dependent scaling on metabolism (dimensionless)

These resource-specific growth rates (s⁻¹) are defined by the uptake rate into the cell, V (mol resource (mol C biomass)⁻¹ s⁻¹), multiplied by the growth yield, y (mol C biomass (mol resource)⁻¹), such that:

$$\begin{aligned} \mu(S) &= V_S y_S \\ \mu(dFe) &= V_{dFe} R^{C:Fe} \\ \mu(N) &= \frac{V_{NH_4}}{\max(0.0, -p_{NH_4})} \\ \mu(O_2) &= V_{O_2} y_{O_2} \end{aligned} \quad \text{Eq. S3}$$

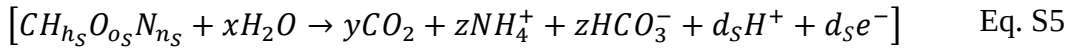
For nitrogen, bacteria only take up ammonium from the environment when their rate of production of NH₄ is negative ($p_{NH_4} < 0$), which only occurs when the dissolved organic matter substrate that they consume is highly deplete in organic nitrogen and thus enriched in C:N ratios. Uptake rates of these resources into the cell are dependent on a prescribed maximum uptake rate and varying resource availabilities in the water column.

$$\begin{aligned} V_S &= \frac{V_S^{max} S}{S + K_S} \\ V_{dFe} &= \frac{V_{dFe}^{max} dFe}{dFe + K_{dFe}} \\ V_N &= \frac{V_{NH_4}^{max} NH_4}{NH_4 + K_{NH_4}} \\ V_{O_2} &= \rho_{O_2} O_2 \end{aligned} \quad \text{Eq. S4}$$

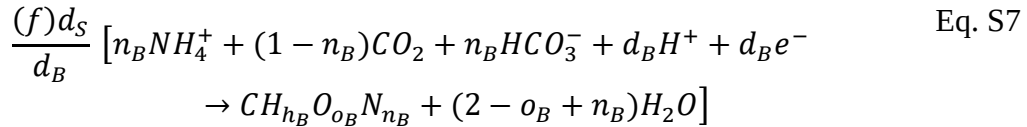
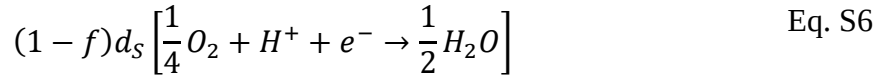
Where

- V_S^{max} is the maximum uptake rate of organic substrate into the cell (mol C (mol C biomass)⁻¹ s⁻¹)
- V_{dFe}^{max} is the maximum uptake rate of dissolved iron into the cell (mol Fe (mol C biomass)⁻¹ s⁻¹)
- $V_{NH_4}^{max}$ is the maximum uptake rate of ammonium into the cell (mol N (mol C biomass)⁻¹ s⁻¹)
- ρ_{O_2} is the diffusive uptake limit for O₂ into the cell ((mol C biomass)⁻¹ s⁻¹)
- K_S is the half-saturation coefficient for uptake of organic substrate
- K_{dFe} is the half-saturation coefficient for uptake of dissolved iron
- K_{NH_4} is the half-saturation coefficient for uptake of ammonium

Resource yields are dependent on the energy of the substrate - the number of electrons released per C atom – and this determines the overall stoichiometry of the heterotrophic metabolism. We compute the resource requirements and products of heterotrophic bacteria building from the work of Zakem et al. (2020) and Rittman and McCarthy (2001). For a heterotrophic bacteria performing aerobic metabolism, that is consuming an organic matter substrate (S) and oxygen (O₂) to produce biomass (B), CO₂ and inorganic nutrient, we split their metabolism into three half reactions and normalize by carbon:



$$\begin{aligned} x &= (2 - o_S + n_S) \\ y &= (1 - n_S) \\ z &= n_S \end{aligned}$$



Where:

- $CH_{h_S}O_{o_S}N_{n_S}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the organic substrate (S), normalized by carbon.
- $CH_{h_B}O_{o_B}N_{n_B}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the bacterial biomass (B), normalized by carbon.
- f is the fraction of electrons that are routed to biomass synthesis, rather than oxygen reduction via the respiratory electron chain.
- d_S is the number of electrons per carbon atom within the substrate (S)
- d_B is the number of electrons per carbon atom within the bacterial biomass (B)

The first half reaction is the electron donor oxidation, where organic matter is fully oxidized to its inorganic constituents and some electrons are released. The second and third constitute the electron acceptor reduction, which generates ATP, and biomass synthesis, which uses ATP. Again, f represents the fraction of electrons that are routed from oxidation of organic matter to biomass synthesis and is an overall cellular efficiency for growth.

If we sum these equations 5-7 (ignoring H_2O), treating organic matter substrate S as equal to $CH_{h_S}O_{o_S}N_{n_S}$, biomass B as equal to $CH_{h_B}O_{o_B}N_{n_B}$ and we define:

$$\begin{aligned} d_S &= 4 + h_S - 2o_S - 3n_S \\ d_B &= 4 + h_B - 2o_B - 3n_B \end{aligned} \quad \text{Eq. S8}$$

$$y_S = \frac{f d_S}{d_B} \quad \text{Eq. S9}$$

we retrieve:

$$S + \frac{(1-f)d_S}{4} O_2 \rightarrow y_S B + (1-y_S)CO_2 + (n_S - n_B y_S)NH_4^+ \quad \text{Eq. S10}$$

A higher d_S increases the potential biomass yield, as does a higher f . A high d value indicates an organic molecule that is highly reduced with a high energy content (i.e., many electrons per carbon atom), whereas a low d value means the molecule is more oxidized with less electrons and less potential energy (Rittman and McCarthy, 2001; Stephens et al., 2025). For the biomass of heterotrophic marine bacteria we assume a stoichiometry of $CH_{1.4}O_{0.4}N_{0.2}$ (White et al., 2019; Zimmerman et al., 2014), which returns $d_B = 4$. Note that in Eq. 10 the coefficient for NH_4 production, $(n_S - n_B y_S)$, is equivalent to p_{NH_4} .

We can expand this framework to include partial oxidation by a heterotrophic bacteria to an organic matter product or products. For one product:

$$\begin{aligned} &[CH_{h_S}O_{o_S}N_{n_S} + xH_2O \\ &\rightarrow \alpha CH_{h_P}O_{o_P}N_{n_P} + yCO_2 + zNH_4^+ + zHCO_3^- \\ &+ (d_S - \alpha d_P)H^+ + (d_S - \alpha d_P)e^-] \end{aligned} \quad \text{Eq. S11}$$

$$\begin{aligned} x &= 2 - o_S + n_S + \alpha(o_P - 2 - n_P) \\ y &= (1 - \alpha - n_S + \alpha n_P) \\ z &= (n_S - \alpha n_P) \end{aligned}$$

$$(1-f)(d_S - \alpha d_P) \left[\frac{1}{4} O_2 + H^+ + e^- \rightarrow \frac{1}{2} H_2O \right] \quad \text{Eq. S12}$$

$$\begin{aligned} &\frac{(f)(d_S - \alpha d_P)}{d_B} [n_B NH_4^+ + (1 - n_B)CO_2 + n_B HCO_3^- + d_B H^+ + d_B e^- \\ &\rightarrow CH_{h_B}O_{o_B}N_{n_B} + (2 - o_B + n_B)H_2O] \end{aligned} \quad \text{Eq. S13}$$

Where:

- α is the fraction of organic substrate that undergoes partial oxidation.
- $CH_{h_S}O_{o_S}N_{n_S}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the organic substrate (S), normalized by carbon.
- $CH_{h_B}O_{o_B}N_{n_B}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the bacterial biomass (B), normalized by carbon.
- $CH_{h_P}O_{o_P}N_{n_P}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the organic product (P), normalized by carbon.

- f is the fraction of electrons that are routed to biomass synthesis, rather than oxygen reduction via the respiratory electron chain.
- d_S is the number of electrons per carbon atom within the substrate (S)
- d_B is the number of electrons per carbon atom within the bacterial biomass (B)
- d_P is the number of electrons per carbon atom within the product (P)

And again summing Eqs. 11 to 13 (ignoring H_2O), treating organic matter substrate S as equal to $CH_{h_S}O_{o_S}N_{n_S}$, the product P as $CH_{h_P}O_{o_P}N_{n_P}$, biomass B as equal to $CH_{h_B}O_{o_B}N_{n_B}$ and we define:

$$d_S = 4 + h_S - 2o_S - 3n_S \quad \text{Eq. S14}$$

$$d_B = 4 + h_B - 2o_B - 3n_B$$

$$d_P = 4 + h_P - 2o_P - 3n_P$$

$$D = d_S - \alpha d_P \quad \text{Eq. S15}$$

$$y_S = \frac{fD}{d_B} \quad \text{Eq. S16}$$

we retrieve:

$$S + \frac{(1-f)D}{4}O_2 \rightarrow y_SB + \alpha P + (1-\alpha-y_S)CO_2 + (n_S - \alpha n_P - n_By_S)NH_4^+ \quad \text{Eq. S17}$$

Now, a higher D increases the potential biomass yield. Increasing α lowers D because more substrate carbon is retained in the partially oxidized organic product rather than oxidized to release electrons.

We can also see that the growth yield of bacteria on organic carbon substrate, y_S , increases as a function of D and f . This is equivalent to Zakem et al. (2020) except that d_S has been extended to D , which captures the dependency of bacterial growth on the degree of partial oxidation, α , and the degree to which the product is oxidized (d_P).

Considering multiple products is a simple extension of this framework:

$$\begin{aligned} & [CH_{h_S}O_{o_S}N_{n_S} + xH_2O \\ & \rightarrow (\alpha\beta)CH_{h_{P1}}O_{o_{P1}}N_{n_{P1}} + (\alpha(1-\beta))CH_{h_{P2}}O_{o_{P2}}N_{n_{P2}} \\ & + yCO_2 + zNH_4^+ + zHCO_3^- \\ & + (d_S - (\alpha\beta)d_{P1} - (\alpha(1-\beta))d_{P2})H^+ \\ & + (d_S - (\alpha\beta)d_{P1} - (\alpha(1-\beta))d_{P2})e^-] \end{aligned} \quad \text{Eq. S18}$$

$$\begin{aligned} x &= 2 - o_S + n_S + \alpha\beta(o_{P1} - 2 - n_{P1}) + \alpha(1-\beta)(o_{P2} - 2 - n_{P2}) \\ y &= (1 - \alpha - n_S + \alpha\beta n_{P1} + \alpha(1-\beta)n_{P2}) \\ z &= (n_S - \alpha\beta n_{P1} - \alpha(1-\beta)n_{P2}) \end{aligned}$$

$$(1-f)(d_S - \alpha\beta d_{P1} - \alpha(1-\beta)d_{P2}) \left[\frac{1}{4}O_2 + H^+ + e^- \rightarrow \frac{1}{2}H_2O \right] \quad \text{Eq. S19}$$

$$\frac{(f)(d_S - \alpha\beta d_{P1} - \alpha(1 - \beta)d_{P2})}{d_B} [n_B NH_4^+ + (1 - n_B)CO_2 + n_B HCO_3^- + d_B H^+ + d_B e^- \rightarrow CH_{h_B} O_{o_B} N_{n_B} + (2 - o_B + n_B)H_2O] \quad \text{Eq. S20}$$

Where:

- α is the fraction of organic substrate that undergoes partial oxidation.
- β is the fraction of organic partially oxidized product that goes to product 1.
- $CH_{h_S} O_{o_S} N_{n_S}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the organic substrate (S), normalized by carbon.
- $CH_{h_B} O_{o_B} N_{n_B}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the bacterial biomass (B), normalized by carbon.
- $CH_{h_{P1}} O_{o_{P1}} N_{n_{P1}}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the first organic product (P1), normalized by carbon.
- $CH_{h_{P2}} O_{o_{P2}} N_{n_{P2}}$ is the carbon, hydrogen, oxygen and nitrogen stoichiometry of the second organic product (P2), normalized by carbon.
- f is the fraction of electrons that are routed to biomass synthesis, rather than oxygen reduction via the respiratory electron chain.
- d_S is the number of electrons per carbon atom within the substrate (S)
- d_B is the number of electrons per carbon atom within the bacterial biomass (B)
- d_{P1} is the number of electrons per carbon atom within the first product (P1)
- d_{P2} is the number of electrons per carbon atom within the second product (P2)

And again summing Eqs. 18 to 20 (ignoring H_2O), treating organic matter substrate S as equal to $CH_{h_S} O_{o_S} N_{n_S}$, the products P1 and P2 as $CH_{h_{P1}} O_{o_{P1}} N_{n_{P1}}$ and $CH_{h_{P2}} O_{o_{P2}} N_{n_{P2}}$, biomass B as equal to $CH_{h_B} O_{o_B} N_{n_B}$ and we define:

$$\begin{aligned} d_S &= 4 + h_S - 2o_S - 3n_S \\ d_B &= 4 + h_B - 2o_B - 3n_B \\ d_{P1} &= 4 + h_{P1} - 2o_{P1} - 3n_{P1} \\ d_{P2} &= 4 + h_{P2} - 2o_{P2} - 3n_{P2} \end{aligned} \quad \text{Eq. S21}$$

$$D = d_S - \alpha\beta d_{P1} - \alpha(1 - \beta)d_{P2} \quad \text{Eq. S22}$$

$$y_S = \frac{fD}{d_B} \quad \text{Eq. S23}$$

we retrieve:

$$\begin{aligned} S + \frac{(1-f)D}{4} O_2 \\ \rightarrow y_S B + \alpha\beta P1 + \alpha(1 - \beta)P2 + (1 - \alpha - y_S)CO_2 \\ + (n_S - \alpha\beta n_{P1} - \alpha(1 - \beta)n_{P2} - n_B y_S)NH_4^+ \end{aligned} \quad \text{Eq. S24}$$

Model microbial loop and dissolved organic matter cycling

Three types of dissolved organic matter (DOM) are each consumed by three bacterial heterotrophs: sharing, selfish and scavenging types (Reintjes et al., 2019). The sharing bacteria consume long-chain DOM (l-DOM) and may partially oxidize this DOM to produce

downstream pools of oxidized DOM (o-DOM) and/or short-chain DOM (s-DOM). These sharing bacteria are copiotrophic generalists with fast maximum growth rates and low substrate affinity (Reynolds et al., 2026; Zakem et al., 2025). Selfish bacteria consume and fully oxidize the oxidized DOM pool to CO_2 , but due to the oxidized nature of the substrate they consume and the need for specialised enzymes to consume it these bacteria have the lowest substrate affinity and low growth yields. This allows the oxidized DOM pool to accumulate, consistent with the "island of stability" of slightly oxidized, CRAM-like DOM material that persists in the ocean (Cai et al., 2026; Lechtenfeld et al., 2014). Scavenging bacteria consume the short-chain DOM pool, which we consider to be a pool of highly labile compounds like simple carbohydrates, carboxylic acids and amino acids fed largely by phytoplankton overflow production. Scavenging bacteria, like SAR11, thus have high resource affinities but slow maximum growth rates (Norris et al., 2021; Reynolds et al., 2026; Zakem et al., 2025).

We consider long-chain DOM to inherit a typical stoichiometry of fresh marine organic matter of $\text{CH}_{1.65}\text{O}_{0.3}\text{N}_{0.131}$ (Anderson, 1995), which gives a $d_S = 4.65$. A similar d_S is achieved for the short-chain DOM with a different stoichiometry of $\text{CH}_{2.0}\text{O}_{0.6}\text{N}_{0.05}$. The oxidized DOM, however, contains less energy than these other DOM pools and has a stoichiometry of $\text{CH}_{1.0}\text{O}_{0.6}\text{N}_{0.05}$ with a $d_S = 3.65$ to be consistent with CRAM-like material (Cai et al., 2026; Lechtenfeld et al., 2014).

We note, however, that because we track variations in the N:C content of each DOM pool, the exact values of d_S for each substrate also vary dynamically since n_S ratios change in space and time. The best example of this is for the scavenging bacterial type. In scenarios of strong overflow production of s-DOC by phytoplankton, which has a unique but constant stoichiometry of $\text{CH}_{2.0}\text{O}_{0.6}$ to represent exudation of carbohydrates, carboxylic acids and lipids (Fogg, 1983; Hansell and Carlson, 2014) with no nitrogen content, the N:C ratios of s-DOM can become strongly depleted and consequently the energy of the substrate increases ($d_S > 4.65$), which increases growth yields. However, high growth rates in this scenario can only be achieved if scavenging bacteria can access NH_4 from the environment ($p_{\text{NH}_4} < 0$) and must therefore compete directly with phytoplankton and/or ammonia oxidizing archaea.

1.2 Model configuration

We use the World Ocean Model of Biogeochemistry and Trophic dynamics (WOMBAT) version "full" coupled to the Modular Ocean Model version 6 (MOM6) in a one-dimensional configuration located at SOTS (47°S, 142°E). The coupling framework comes from the Australian Community Climate and Earth System Simulator (ACCESS-OM3) (<https://github.com/ACCESS-NRI/ACCESS-OM3>). The ocean model, MOM6, is forced by the atmospheric conditions of the Japanese Reanalysis (Tsujino et al., 2018) in a repeat year forcing spanning 1990-1991, which was chosen to avoid extremes in climate modes (Stewart et al., 2020). The ocean model was initialised with mean January temperature and salinity taken from the World Ocean Atlas (Locarnini et al., 2024; Reagan et al., 2024). Near-surface mixing was parameterized according to the ePBL (energetic Planetary Boundary Layer) scheme (Reichl and Hallberg, 2018) while interior mixing occurred at a background diffusivity of 2.022×10^{-5} . Atmospheric CO_2 was set at 395.3 ppm, which represents the mean concentration in calendar year 2013 and is roughly the central year of SOTS observations. Aeolian deposition of iron included mineral, fire and anthropogenic sources at the SOTS site (Hamilton et al., 2020). Biogeochemical fields were initialized with available data products when available (Table S1) or otherwise set to constant values.

Table S6. All active tracers in WOMBAT-full, their units and the initial values used in this study.

#	Tracer	Code name	Description	Units	Initial values
1	NH ₄	f_nh4	Ammonium	mol N kg ⁻¹	0.0
2	NO ₃	f_no3	Nitrate	mol N kg ⁻¹	(Garcia et al., 2024b)
3	dFe	f_fe	Dissolved iron	mol Fe kg ⁻¹	(Huang et al., 2022)
4	Si(OH) ₄	f_sil	Silicic acid	mol Si kg ⁻¹	(Garcia et al., 2024b)
5	O ₂	f_o2	Dissolved oxygen	mol O ₂ kg ⁻¹	(Garcia et al., 2024a)
6	Alk	f_alk	Alkalinity	mol Eq kg ⁻¹	(Lauvset et al., 2016)
7	DIC	f_dic	Dissolved inorganic carbon	mol C kg ⁻¹	(Lauvset et al., 2016)
8	B _{np} ^C	f_phy	Nano-phytoplankton carbon biomass	mol C kg ⁻¹	0.1 × 10 ⁻⁶
9	B _{np} ^{Chl}	f_pchl	Nano-phytoplankton chlorophyll biomass	mol C kg ⁻¹	(B _{np} ^C) 0.004
10	B _{np} ^{Fe}	f_phyfe	Nano-phytoplankton iron biomass	mol Fe kg ⁻¹	(B _{np} ^C) 7 × 10 ⁻⁶
11	B _{mp} ^C	f_dia	Micro-phytoplankton carbon biomass	mol C kg ⁻¹	0.1 × 10 ⁻⁶
12	B _{mp} ^{Chl}	f_dchl	Micro-phytoplankton chlorophyll biomass	mol C kg ⁻¹	(B _{mp} ^C) 0.004
13	B _{mp} ^{Fe}	f_diafe	Micro-phytoplankton iron biomass	mol Fe kg ⁻¹	(B _{mp} ^C) 7 × 10 ⁻⁶
14	B _{mp} ^{Si}	f_diasi	Micro-phytoplankton silicon biomass	mol Si kg ⁻¹	(B _{mp} ^C) $\frac{16}{122}$
15	B _{mz} ^C	f_zoo	Micro-zooplankton carbon biomass	mol C kg ⁻¹	0.1 × 10 ⁻⁶
16	B _{mz} ^{Fe}	f_zoofe	Micro-zooplankton iron biomass	mol Fe kg ⁻¹	(B _{mz} ^C) 7 × 10 ⁻⁶
17	B _{Mz} ^C	f_mes	Meso-zooplankton carbon biomass	mol C kg ⁻¹	0.1 × 10 ⁻⁶
18	B _{Mz} ^{Fe}	f_mesfe	Meso-zooplankton iron biomass	mol Fe kg ⁻¹	(B _{Mz} ^C) 7 × 10 ⁻⁶
19	B _{sd} ^C	f_sdet	Small particulate detritus carbon biomass	mol C kg ⁻¹	0.1 × 10 ⁻⁶
20	B _{sd} ^{Fe}	f_sdetfe	Small particulate detritus iron biomass	mol Fe kg ⁻¹	(B _{sd} ^C) 7 × 10 ⁻⁶
21	B _{sd} ^{CaCO3}	f_caco3	Small particulate detrital calcium carbonate	mol C kg ⁻¹	(B _{sd} ^C + B _{ld} ^C) 0.08
22	B _{ld} ^C	f_ldet	Large particulate detritus carbon biomass	mol C kg ⁻¹	0.1 × 10 ⁻⁶
23	B _{ld} ^{Fe}	f_ldetfe	Large particulate detritus iron biomass	mol Fe kg ⁻¹	(B _{ld} ^C) 7 × 10 ⁻⁶
24	B _{ld} ^{Si}	f_ldetsi	Large particulate detritus silicon biomass	mol Si kg ⁻¹	(B _{ld} ^C) $\frac{16}{122}$
25	B _{l-DOM} ^C	f_ldoc	Long-chain dissolved organic carbon (DOC)	mol C kg ⁻¹	0.0
26	B _{l-DOM} ^N	f_ldon	Long-chain dissolved organic nitrogen (DON)	mol N kg ⁻¹	0.0
27	B _{o-DOM} ^C	f_odoc	Oxidized dissolved organic carbon (DOC)	mol C kg ⁻¹	0.0
28	B _{o-DOM} ^N	f_odon	Oxidized dissolved organic nitrogen (DON)	mol N kg ⁻¹	0.0
29	B _{s-DOM} ^C	f_sdoc	Short-chain dissolved organic carbon (DOC)	mol C kg ⁻¹	0.0
30	B _{s-DOM} ^N	f_sdon	Short-chain dissolved organic nitrogen (DON)	mol N kg ⁻¹	0.0
31	B _{lbac} ^C	f_lbac	Long-chain DOM-consuming heterotrophic bacteria	mol C kg ⁻¹	0.1 × 10 ⁻⁶
32	B _{obac} ^C	f_obac	Oxidized DOM-consuming heterotrophic bacteria	mol C kg ⁻¹	0.1 × 10 ⁻⁶
33	B _{sbac} ^C	f_sbac	Short-chain DOM-consuming heterotrophic bacteria	mol C kg ⁻¹	0.1 × 10 ⁻⁶
34	B _{aoa} ^C	f_aoa	Ammonia oxidizing archaea	mol C kg ⁻¹	0.1 × 10 ⁻⁶
35	aFe _s	f_safe	Small authigenic iron particles	mol Fe kg ⁻¹	(dFe) 0.1
36	aFe _l	f_lafe	Large authigenic iron particles	mol Fe kg ⁻¹	(dFe) 0.1

1.3 Observational target datasets

The SOTS is the longest running observatory in the Southern Ocean and provides hydrographic and biogeochemical data over decades (Shadwick et al., 2025). We collected measurements of chlorophyll (mooring and bottle), nutrients including ammonium, nitrate, silicate and dissolved iron (Traill et al., 2025), particulate organic carbon, $p\text{CO}_2$ and estimates of air-sea flux of CO_2 at SOTS. These data were used to create monthly climatologies (Fig. SM1). For NO_3 , POC and bottle chlorophyll, we chose to use Argo profiles from the region (Fig. SM2) to replace the raw measurements from SOTS. Estimates of zooplankton and bacteria were also gathered to supplement the SOTS observations. Micro- and meso-zooplankton from the MAREDAT database were averaged between 40°S and 60°S (Buitenhuis et al., 2013). Bacterial abundances in cells per mL from the nearby Chatham Rise east of New Zealand in spring 2018 (Deans et al., 2025) were converted through to mmol C m^{-3} assuming $20 \text{ fg C cell}^{-1}$. Estimates of net primary production from the Carbon-based Production Model (CbPM) (Behrenfeld et al., 2006) and the Vertically Generalized Production Model (VGPM) (Behrenfeld and Falkowski, 1997) averaged between $140\text{--}144^\circ\text{E}$, $45\text{--}49^\circ\text{S}$ and between calendar years 2012 to 2020 were included as post-optimisation constraints (see main Figure 3b).

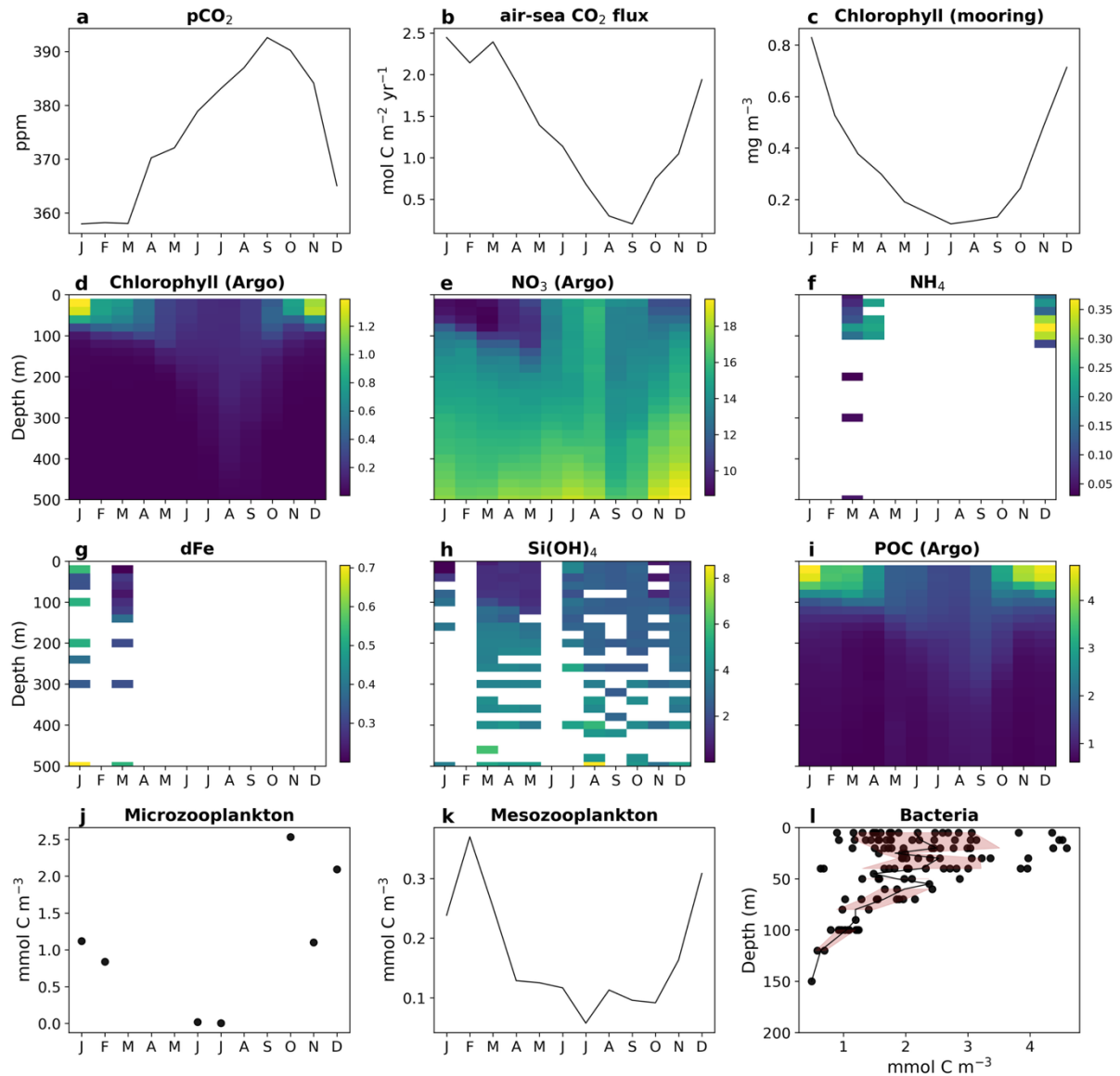


Figure S7. Observational targets for guiding the sensitivity analysis and model optimization at SOTS. Spring bacteria from the Chatham Rise east of New Zealand in spring 2018 (Deans et al., 2025). Mean micro-zooplankton and meso-zooplankton between 0 to 200 m from the MAREDAT dataset averaged between 40°S and 60°S (Buitenhuis et al., 2013).

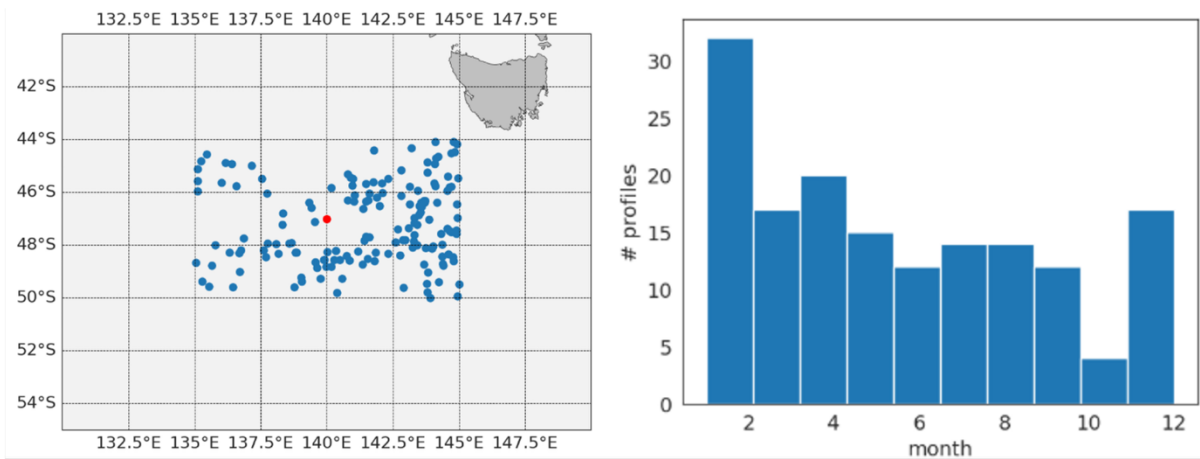


Figure S8. Pre-processing of the Argo data showing the locations of Argo profiles (blue) in reference to the SOTS site (red) and the number of profiles taken in each calendar month.

1.4 Sensitivity analysis

To calibrate the ocean biogeochemical model against the observed datasets shown in Figure SM1 we performed a comprehensive two-step sensitivity analysis to identify the most influential parameters on model performance. Due to the large number of experiments required, we used surrogate machine learning methods (Buchanan et al., 2025; Reddy et al., 2024b).

The sensitivity analysis was a two-stage procedure using a sobol sensitivity analysis (Sobol', 2001) – one of the most effective global sensitivity analysis methods. First, we identified 87 uncertain input parameters to WOMBAT-full (Supplementary Data Table 1) and ran 2048 sensitivity experiments for 10 years at the SOTS site, starting from initial conditions (Table S1). Output from the final year of the model was compared against a subset of the observational targets via the normalized root mean square error (NRMSE). We trained a Gaussian Process Regression (GPR) model on this data to predict the NRMSE for each observational target with variations in the 87 input parameters. This surrogate GPR model generated 360,448 synthetic NRMSE outcomes, varying the 87 input parameters through their *a priori* ranges. This initial sensitivity analysis was a wide by coarse net designed to identify which parameters were influential to the model performance against the 12 observational targets, with the idea being to reduce the degrees of freedom down from 87 to make the eventual optimization more manageable. It did so successfully by identifying 40 influential parameters out of the original 87 that were key to delivering good outcomes.

Subsequently we performed another 2048 10-year experiments varying the 40 influential parameters through their *a priori* ranges (Supplementary Table S2), trained GPR models on the resulting NRMSE (Fig. SM3, SM4), and repeated the sensitivity analysis. Our GPR models were able to perform satisfactorily for most observational targets, except for NH_4 (Fig. SM4). This identified that of the 40 parameters, 30 were most influential, with 10 having very small or negligible influence (Fig. SM5). The 30 parameters of influence included controls on phytoplankton growth, zooplankton growth, sinking and remineralisation of particles, iron cycling, and the microbial loop. A brief description of these influential parameters is in Supplementary Table 2 but is also showcased in the WOMBAT-full documentation (<https://wombat-docs.readthedocs.io/latest/>).

Table S7. The influential parameters and their *a priori* ranges that were the subject of the sensitivity analysis and were optimized to produce the best agreement with the observational targets. Greyed rows are those parameters that were selected as sensitive following the large but less accurate sensitivity analysis using 87 parameters but were subsequently identified as insensitive following the second constrained sensitivity analysis. These 10 parameters were not carried forward during Markov-chain Monte-Carlo optimisation. Optimal values represent the best performing simulation of 100 final simulations, whose parameter values were selected from the posterior distributions generated by the Markov-chain Monte-Carlo sampling.

#	Parameter	Description	Units	Range	Optimal
1	alphabio_phy	Initial slope of P-I curve (nano-phytoplankton)	$\text{mol C (mol Chl)}^{-1} (\text{W m}^{-2})^{-1}$	1 – 4	3.2
2	abioa_phy	Max growth rate parameter a (nano-phytoplankton)	s^{-1}	2.89E-06 – 1.74E-05	1.20E-05
3	bbioa_phy	Max growth rate parameter b (nano-phytoplankton) ($Q_{10} = b^{(10)}$)	dimensionless	1.04 – 1.08	1.070
4	abioa_dia	Max growth rate parameter a (micro-phytoplankton)	s^{-1}	5.78E-06 – 1.74E-05	7.65E-06
5	bbioh	Max growth rate parameter b (heterotrophy) ($Q_{10} = b^{(10)}$)	dimensionless	1.06 – 1.08	1.070

6	phykf	Half-saturation coefficient Fe uptake (nano-phytoplankton)	nM Fe	0.1 – 3.0	0.43
7	phymaxqc	Max Chl:C (nano-phytoplankton)	mg Chl (mg C) ⁻¹	0.01 – 0.08	0.016
8	phylmor	Linear mortality rate (nano-phytoplankton)	s ⁻¹	1.16E-09 – 1.16E-06	6.21E-08
9	diamaxqs	Max Si:C (micro-phytoplankton)	mol Si (mol C) ⁻¹	0.3 – 1.0	0.77
10	diaVmaxs	Max Si uptake (micro-phytoplankton)	mol Si (mol C) ⁻¹ s ⁻¹	3.47E-07 – 6.94E-06	2.26E-06
11	dialmor	Linear mortality rate (micro-phytoplankton)	s ⁻¹	1.16E-09 – 1.16E-06	7.00E-07
12	overflow	Max DOC exudation fraction by phytoplankton	mol C (mol C) ⁻¹	0.1 – 0.9	0.20
13	zooCingest	Micro-zooplankton C ingestion efficiency	mol C (mol C) ⁻¹	0.50 – 0.95	0.80
14	zooCassim	Micro-zooplankton C assimilation efficiency	mol C (mol C) ⁻¹	0.15 – 0.60	0.55
15	zooepps	Micro-zooplankton prey capture efficiency	(μM) ⁻² s ⁻¹	5.78E-07 – 5.79E-06	3.80E-06
16	zprefdia	Micro-zooplankton preference (micro-phytoplankton)	dimensionless	0.01 – 1.0	0.50
17	zoolmor	Micro-zooplankton linear mortality rate	s ⁻¹	1.16E-09 – 1.16E-06	2.50E-08
18	mesCassim	Meso-zooplankton C assimilation efficiency	mol C (mol C) ⁻¹	0.15 – 0.60	0.55
19	meseps	Meso-zooplankton prey capture efficiency	(μM) ⁻² s ⁻¹	1.16E-07 – 2.31E-06	2.00E-06
20	mprefphy	Meso-zooplankton preference (micro-phytoplankton)	dimensionless	0.0 – 1.0	0.80
21	meslmor	Meso-zooplankton linear mortality rate	s ⁻¹	1.16E-09 – 1.16E-06	1.80E-09
22	mesqmor	Meso-zooplankton quadratic density-dependent mortality rate	(mmol C m ⁻³) ⁻¹ s ⁻¹	5.79E-07 – 1.16E-05	3.50E-06
23	mespreyswitch	Meso-zooplankton prey switching exponent	dimensionless	1.0 – 3.0	2.5
24	detqrem	Detritus hydrolysis rate	(mmol C m ⁻³) ⁻¹ s ⁻¹	1.16E-06 – 1.74E-05	8.10E-06
25	ldetphi	Porosity (large detritus)	dimensionless	0.5 – 0.99	0.88
26	phyrad0	Nano-phytoplankton mean radius	μm	2 – 20	11.0
27	diarad0	Micro-phytoplankton mean radius	μm	20 – 200	110.0
28	zoorad0	Micro-zooplankton mean radius	μm	5 – 50	27.5
29	phybiot	Biomass threshold (nano-phytoplankton)	μM C	0.1 – 2.0	1.6
30	diabiot	Biomass threshold (micro-phytoplankton)	μM C	0.1 – 2.0	0.9
31	dfefloor	Minimum open water concentration of dissolved iron (detection limit)	nM Fe	0.001 – 0.05	0.025
32	kscav_dfe	Free dissolved iron scavenging rate	(μM mass) ⁻¹ s ⁻¹	1.16E-09 – 1.16E-05	2.80E-08
33	kcoag_dfe	Colloidal dissolved iron coagulation rate	(μM C) ⁻¹ s ⁻¹	1.16E-14 – 1.16E-09	1.40E-13
34	bsi_alpha	Natural-log intercept of temperature-dependent biogenic silica dissolution	ln(per hour)	(-10) – (-7.5)	-8.60
35	aoalmor	AOA linear mortality rate	s ⁻¹	1.16E-09 – 1.16E-06	2.45E-07
36	pbac_alpha	Degree of implicit partial oxidation of particulate detritus	mol C (mol C) ⁻¹	0.0 – 1.0	1.0
37	lbac_alpha	Degree of explicit partial oxidation by long-chain DOM-consuming sharing bacteria	mol C (mol C) ⁻¹	0.0 – 0.95	0.30
38	lbac_beta	Fraction of partially oxidized product routed to short-chain DOM by sharing bacteria	mol C (mol C) ⁻¹	0.0 – 1.0	0.10

39	baclmor	Bacteria (all types) linear mortality rate	s^{-1}	1.16E-09 1.16E-06	–	5.50E-08
40	bacqmor	Bacteria (all types) quadratic density-dependent mortality rate	$(mmol\ C\ m^{-3})^{-1}\ s^{-1}$	1.16E-07 2.89E-06	–	2.23E-07

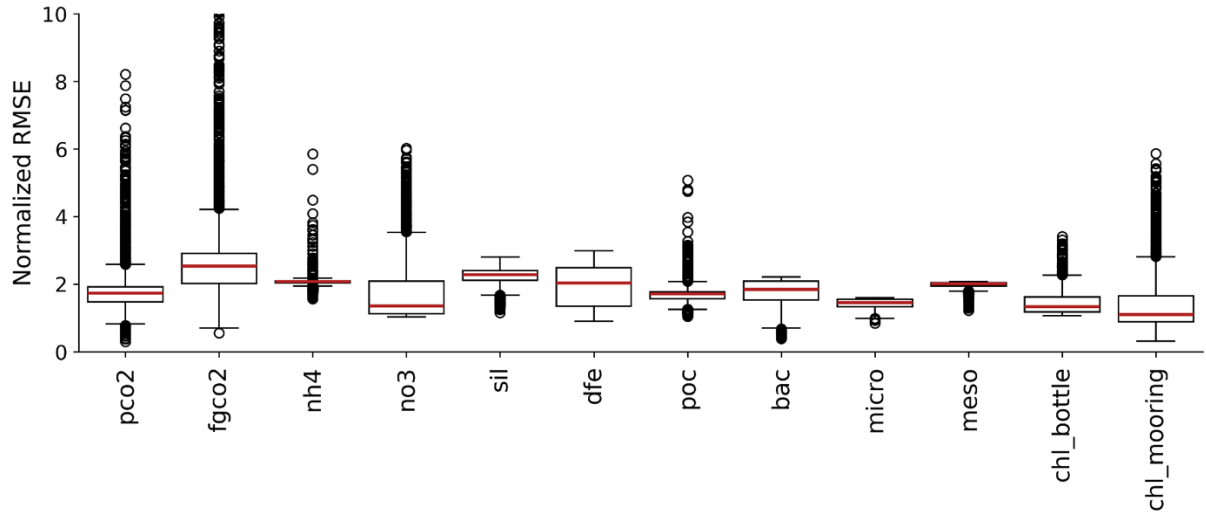


Figure S9. Normalized root mean square error (NRMSE) of the 2048 experiments run with the 35 sensitive parameters chosen following the first stage of the sensitivity analysis. The NRMSE is shown for each of the 12 observational targets. “pco2” = surface pCO₂. “fgco2” = air-sea flux of CO₂. “nh4” = bottle NH₄. “no3”, “poc” and “chl_bottle” are in fact Argo NO₃, Argo POC and Argo chlorophyll. “sil” = bottle Si(OH)₄. “dfe” = bottle dFe. “bac” = Spring bacteria from the Chatham Rise east of New Zealand in spring 2018 (Deans et al., 2025). “micro” and “meso” are the mean micro-zooplankton and meso-zooplankton between 0 to 200 m from the MAREDAT dataset averaged between 40°S and 60°S (Buitenhuis et al., 2013).

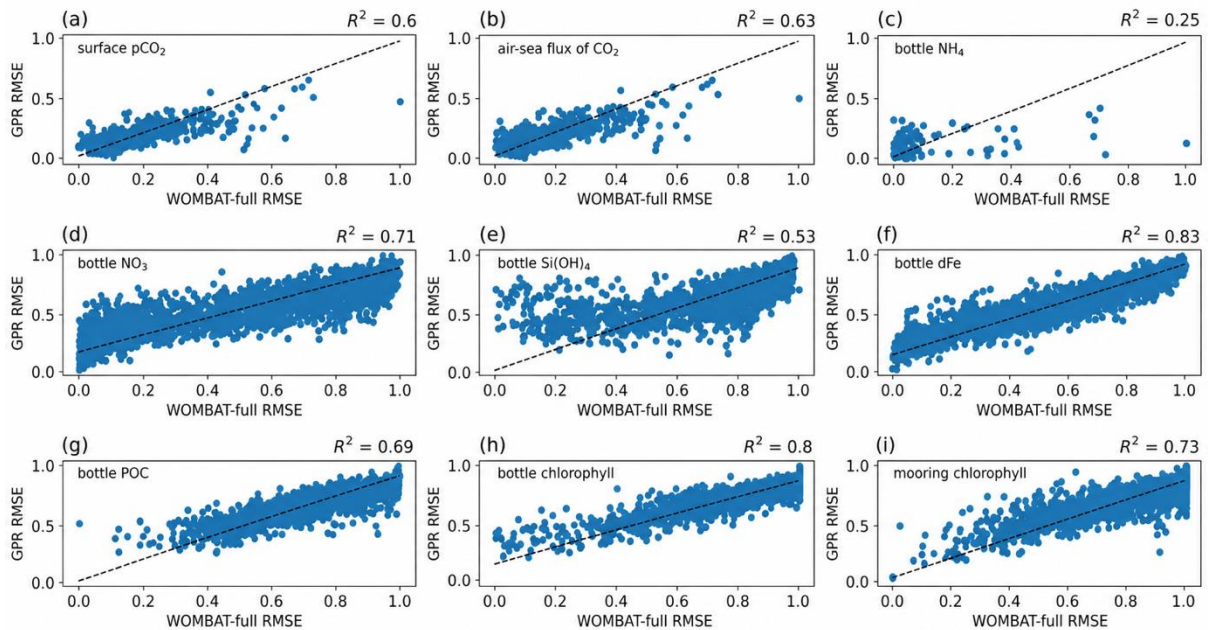


Figure S10. Skill of the Gaussian Process Regression (GPR) model to reproduce the WOMBAT-full model NRMSE outcomes for the subset of observational targets at the SOTS site. We consider that R^2 greater than 0.6 to be sufficient

for accurate sensitivity analysis, with values lower than this producing questionable or inaccurate results of the sensitivity analysis.

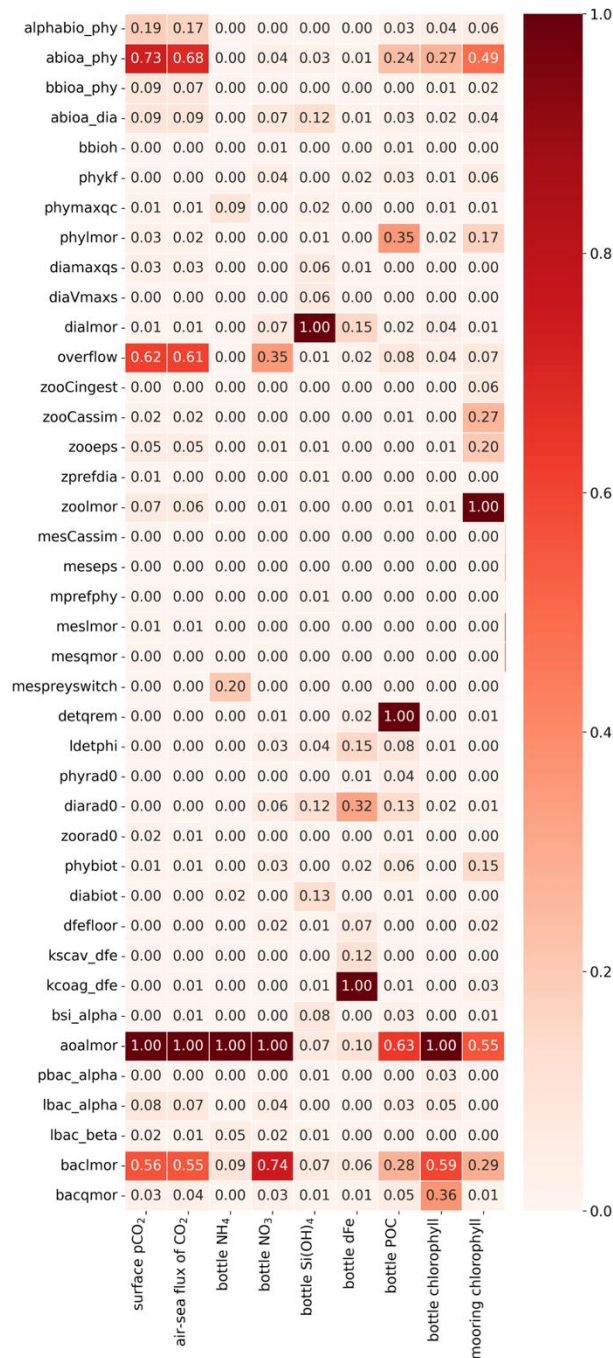


Figure S11. Results of the second global sensitivity analysis of WOMBAT-full against subset of observational datasets. Performance is measured by the NRMSE. Darker colours and higher numbers indicate a greater sensitivity of the observational target to the parameter in question. Parameter names and their descriptions can be found directly within the model documentation (<https://wombat-docs.readthedocs.io/stable/>).

1.5 Optimization

A Bayesian optimization procedure followed the identification of the most sensitive parameters (Reddy et al., 2024). Gaussian Process Regression (GPR) surrogate models were trained on the

either the NRMSE (Eq. S25) or a cost function composed of the Pearson's correlation and NRMSE (Eq. S26) from the 2048 sensitivity experiments (Table S3). Rather than optimize to each target observation in isolation, we performed a simultaneous optimization of the 30 sensitive parameters (Table S2) to all targets (Eq. S27). In essence, we attempted to minimize the global cost function, J .

$$J_n = \text{NRMSE}_n^{z,t} \quad \text{Eq. S25}$$

$$J_n = (1 - r_n^{z,t}) \cdot \text{NRMSE}_n^{z,t} \quad \text{Eq. S26}$$

$$J = \sum_{n=1}^O J_n \quad \text{Eq. S27}$$

Table S8. Individual cost functions, J_n , contributing to the global cost function, J . Skill of the GPR in reproducing the full model skill statistic is represented by the R^2 . Greyed rows indicate those targets that were not included in the optimisation due to poor surrogate performance.

#	Target	Statistic	GPR R^2
1	Surface pCO ₂ (ppm)	NRMSE	0.60
2	Flux of CO ₂ between ocean and atmosphere (mol C m ⁻² year ⁻¹)	NRMSE	0.63
3	Bottle ammonium data (μM)	Correlation	0.31
4	Argo-based nitrate data (μM)	(1-r)*NRMSE	0.71
5	Bottle silicic acid data (μM)	(1-r)*NRMSE	0.59
6	Bottle dissolved iron data (nM)	(1-r)*NRMSE	0.67
7	Argo-based particulate organic carbon data (μM)	NRMSE	0.70
8	Spring vertical profile of heterotrophic bacterial biomass	NRMSE	0.80
9	MAREDAT microzooplankton biomass (mean 0-200 metres)	NRMSE	0.74
10	MAREDAT mesozooplankton biomass (mean 0-200 metres)	NRMSE	0.52
11	Argo-based chlorophyll data (mg m ⁻³)	(1-r)*NRMSE	0.64
12	Mooring chlorophyll data at 30 metres (mg m ⁻³)	NRMSE	0.49

Where subscript n in the n^{th} target observation (Fig. SM1; Table S3) and superscript z,t are the cost functions, J_n , evaluated over each depth and time over an annual cycle where there are observations. We used a composite covariance kernel composed of a constant amplitude term multiplied by a Matérn kernel, plus a white-noise kernels. The initial hyperparameter values were guided by previous work (Reddy et al., 2024a). GPR models trained to predict model performance via K-fold cross validation (K=8) showed generally good skill except for NH₄ which achieved poor skill ($R^2 \sim 0.3$) (Table S3; Fig. SM6).

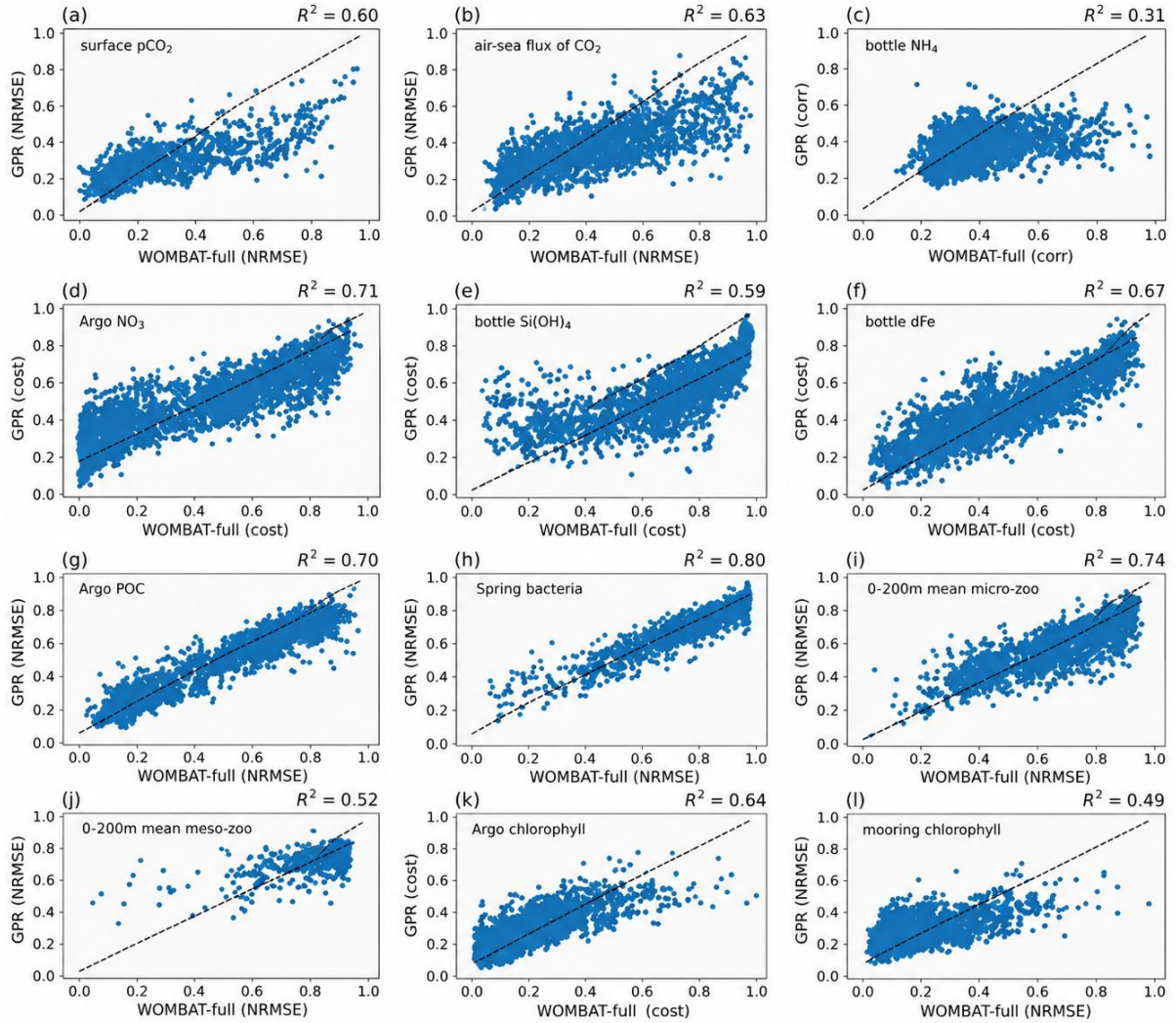


Figure S12. Accuracy of the Gaussian Process Regression (GPR) model to true model output (2048 samples) during the optimisation procedure. Performance of the GPR is measured as the R^2 when comparing predicted versus definite normalized root mean square errors (NRMSE), the Pearson's correlation (corr) or a cost function ($(1-r)*\text{NRMSE}$) and is shown above each panel for each target variable.

We performed surrogate-assisted Bayesian calibration of the model using the trained GPRs. For our priors, we used a bounded gaussian distribution centred on the middle of the prior distributions (Table S2) with a standard deviation of 0.35 with hard bounds 0 to 1. This mildly penalized the optimization when selecting values from the edge of the prior distributions. The assumption underlying this decision is that the priors were all chosen to reflect broad ranges and their central values were informed by the literature (Supplementary Data Table X). Parameters not used in the calibration procedure were centred at 0.5.

The likelihood was calculated from the GPR surrogate models trained to predict the normalized model-data error (J) for each observational target. For a given parameter vector, each GPR emulator returned a predictive mean and standard deviation. The predicted error was clipped at zero to avoid penalizing negative predicted costs, and each target contributed a Gaussian log-likelihood term with uncertainty. The total log-likelihood was calculated as the sum of the individual target likelihoods, and the log-posterior was then obtained by adding this likelihood to the log-prior (Reddy et al., 2024a). A parameter set was considered more optimal when the GPR surrogates predicted low model-data errors across all targets, while also not being too far towards the edges of the prior ranges.

We used Markov chain Monte Carlo (MCMC) sampling to estimate the posterior distribution of each of the 30 sensitive parameters (Table S2) without requiring an explicit calculation of these values (Issan et al., 2023). First, a differential-evolution search as used to identify a high-posterior starting point for the ensemble. We normalized the priors via min-max scaling between 0 and 1 to ensure more efficient exploration of the high-dimensional parameter space (i.e., better “mixing”) and therefore more reliable convergence, while normalization of the cost function, J , ensured equal weighting to each of the 11 target observations (NH₄ data was excluded). We used affine invariant ensemble sampling from the “emcee” python package (Foreman-Mackey et al., 2013). This approach deploys an ensemble of chains to simplify sampling from anisotropic distributions. 140 walkers were applied with stretch moves of 1.3 occurring 70% of the time and with differential evolution moves occurring 30% of the time. The chains were run for 300,000 steps to achieve stable posterior distributions (Foreman-Mackey et al., 2013; Goodman and Weare, 2010), with a burn-in phase defined by 5-times the maximum autocorrelation time. Samples were thinned by half of the minimum estimated autocorrelation time and the burn-in phase was cut before constructing the final posterior distributions shown in Fig. SM7 (Foreman-Mackey et al., 2013). Only surrogate machine learning models like the GPR employed here could provide this degree of convergence of posteriors across 30 input parameters attempting to minimize 11 targets. Convergence was also evident by a mean acceptance fraction of 0.163 (Foreman-Mackey et al., 2013), a Gelman–Rubin statistic of between 1.005 and 1.009 for each parameter and visual assessment of the chains.

The final posterior distributions of the 30 parameters were sampled randomly from within their 90% confidence intervals, returning 100 unique and “optimal” parameter sets (Fig. SM7). These 100 parameter sets were used to run the model again for 10 years, assessed against the targets, and the best parameter set of these 100 was chosen as our calibrated model for further experimentation.

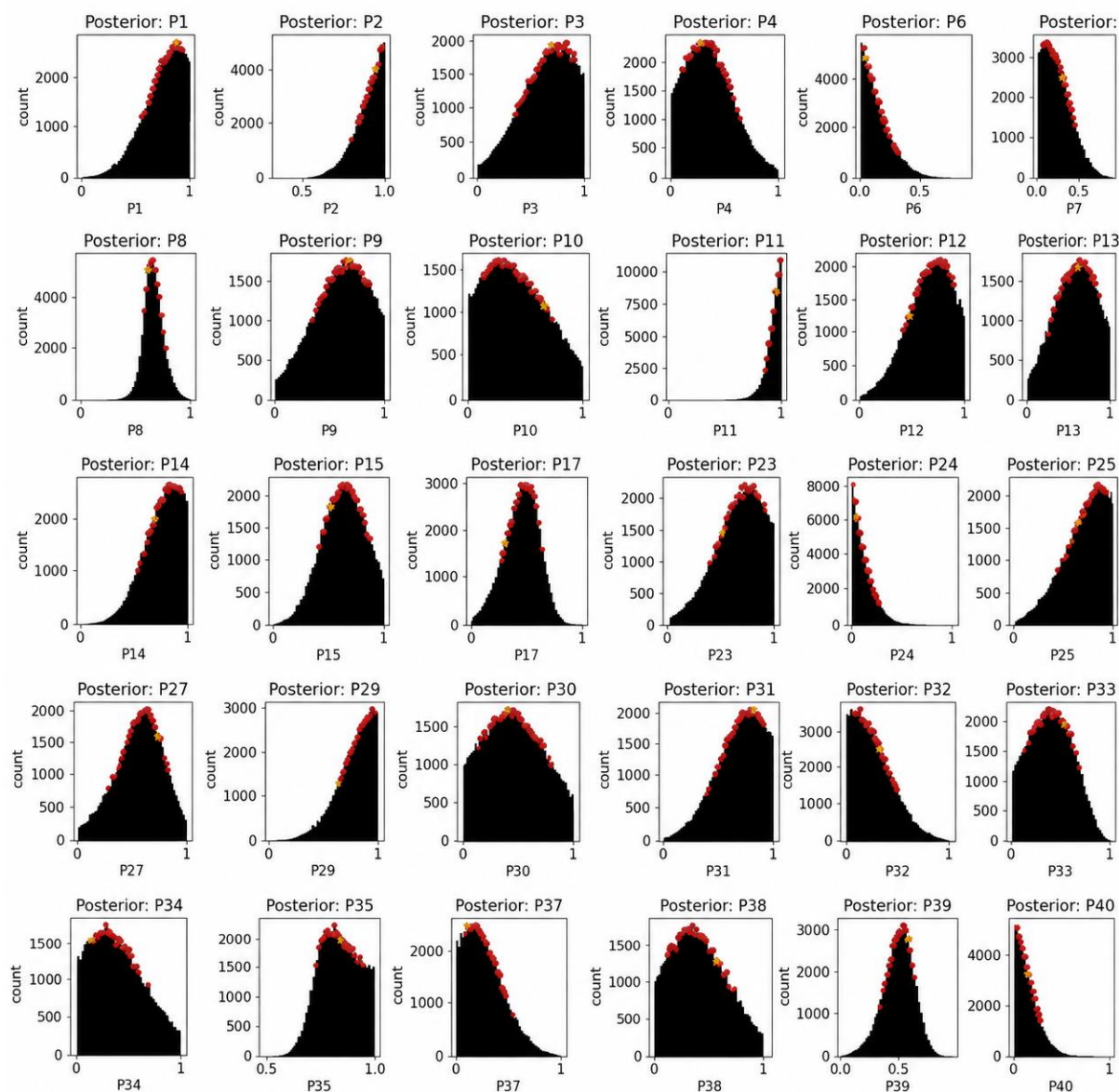


Figure S13. Optimal probability density functions of 30 important parameters (Table S2). Optimization involved minimizing the global (summed) cost-function (Eq. S27) of model performance across the twelve target variables shown in Fig. SM1. We show normalized distributions of each parameter here and refer the reader to the ranges (a priori and optimal) shown in Table S2 for their actual values. Red dots are the 100 random samples from within the 90% confidence interval and the gold star is the optimal parameters set taken forward for experimentation at the SOTS site.

1.6 Experiments and model performance

We performed 21 experiments with the calibrated model. For our first experiment we set ``pbac_alpha`` = 0.0, meaning that all particulate organic carbon was immediately remineralised back to CO_2 and did not enter the DOM cycle. For the remaining 20 experiments we set ``pbac_alpha`` = 1.0 and then varied the degree of partial oxidation of the sharing bacteria from 0% to 95% over 5% intervals. All other input parameters remained fixed at their default, optimized values shown in Table S2. Experiments were run for 10 years from initial conditions (Table S1) and evaluated in the final year.

We present in Fig. SM8 the annual cycle of nutrient concentrations (NO_3 , NH_4 , Si(OH)_4 , and dFe), nano-phytoplankton biomass, micro-phytoplankton biomass, chlorophyll concentration,

total zooplankton biomass, small and large particulate organic detritus, export rates of small and large particulate organic detritus, total dissolved organic carbon concentration, its C:N ratio, and the total biomass of sharing bacteria and total bacterial types in the optimal model configuration. We also show the model performance against the SOTS target observations following optimization on Taylor Diagrams in Fig. SM9 and Fig. SM10. As seen in these Taylor diagrams, the model consistently performs well compared to the observations, with positive correlations > 0.6 for most variables. The exception was for $p\text{CO}_2$ and air-sea CO_2 flux, which only performed with moderate degrees of partial oxidation of DOM by the free-living sharing heterotrophic bacteria.

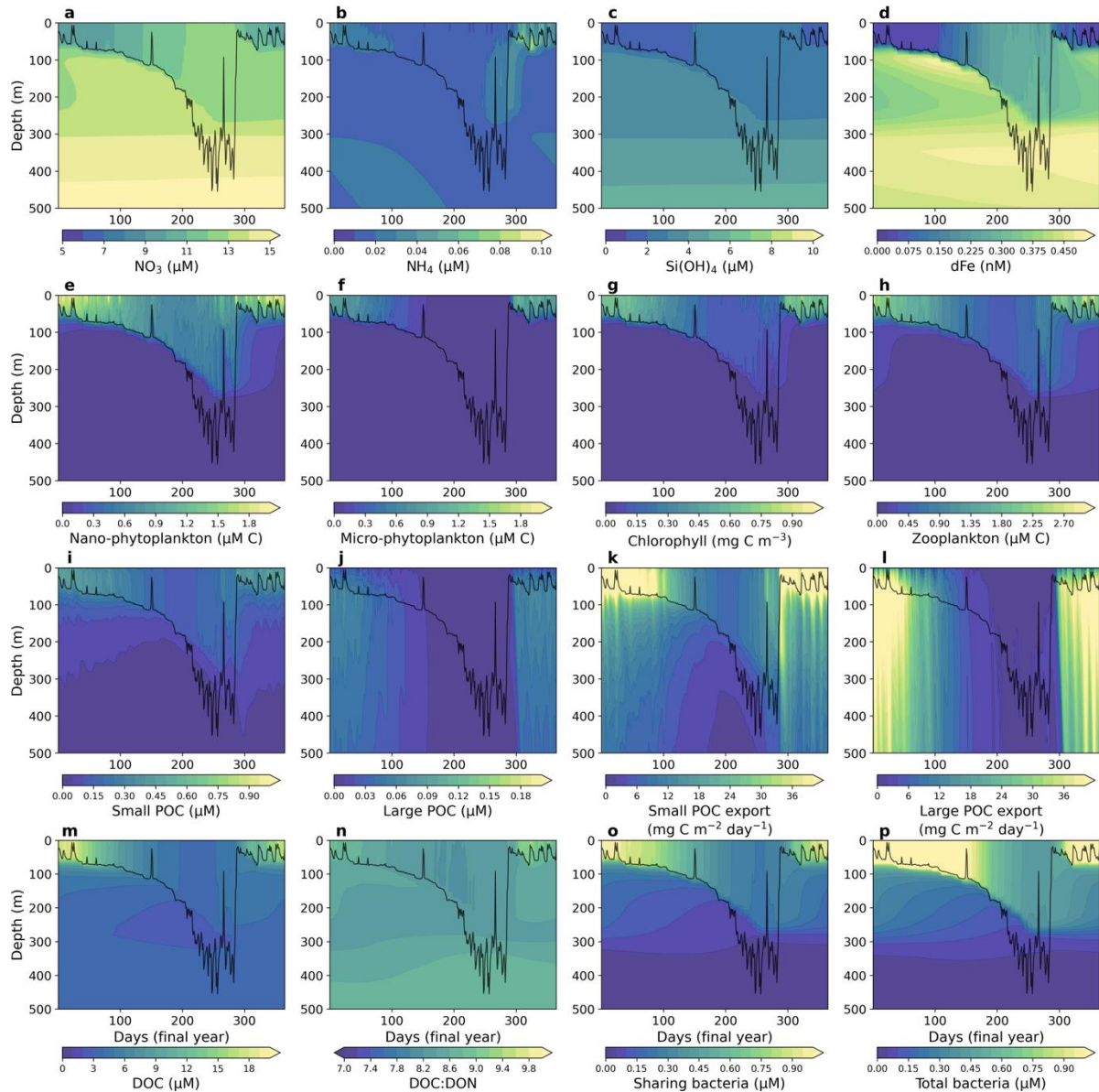


Figure S14. Output from the final year of the biogeochemical model calibrated to conditions at the Southern Ocean Timeseries site (SOTS). Top row (a-d) shows inorganic nutrients. Second row (e-h) shows phytoplankton, chlorophyll and zooplankton. Third row (i-l) shows sinking detritus and export rates. Bottom row (m-p) shows dissolved organic carbon, its C:N ratio, and biomass of sharing bacteria (partially oxidizing type) and total free-living bacteria.

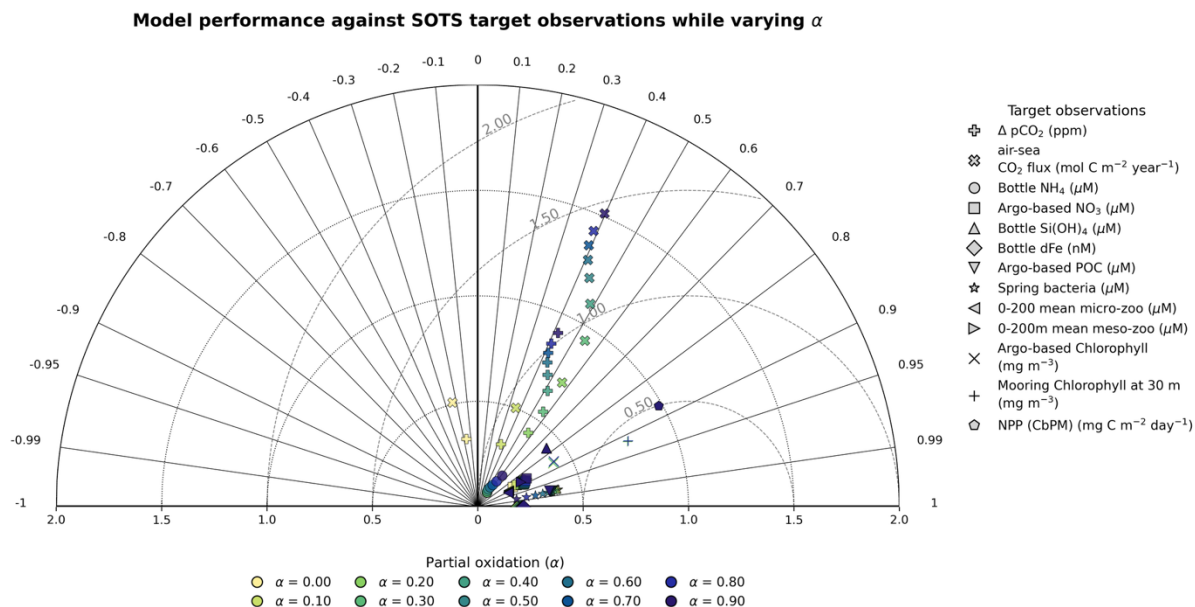


Figure S15. Taylor Diagram of WOMBAT-full performance against the observational target datasets. Model parameters used default values listed in Table S2 and with variations in the degree of partial oxidation of the sharingheterotrophic bacteria. As discussed, variations in partial oxidation (α) have a strong influence of both $p\text{CO}_2$ and the air-sea flux of CO_2 into the ocean, while having small effects on other biogeochemical variables that tend to perform well (Pearson's correlation coefficients tend to be greater than 0.6).

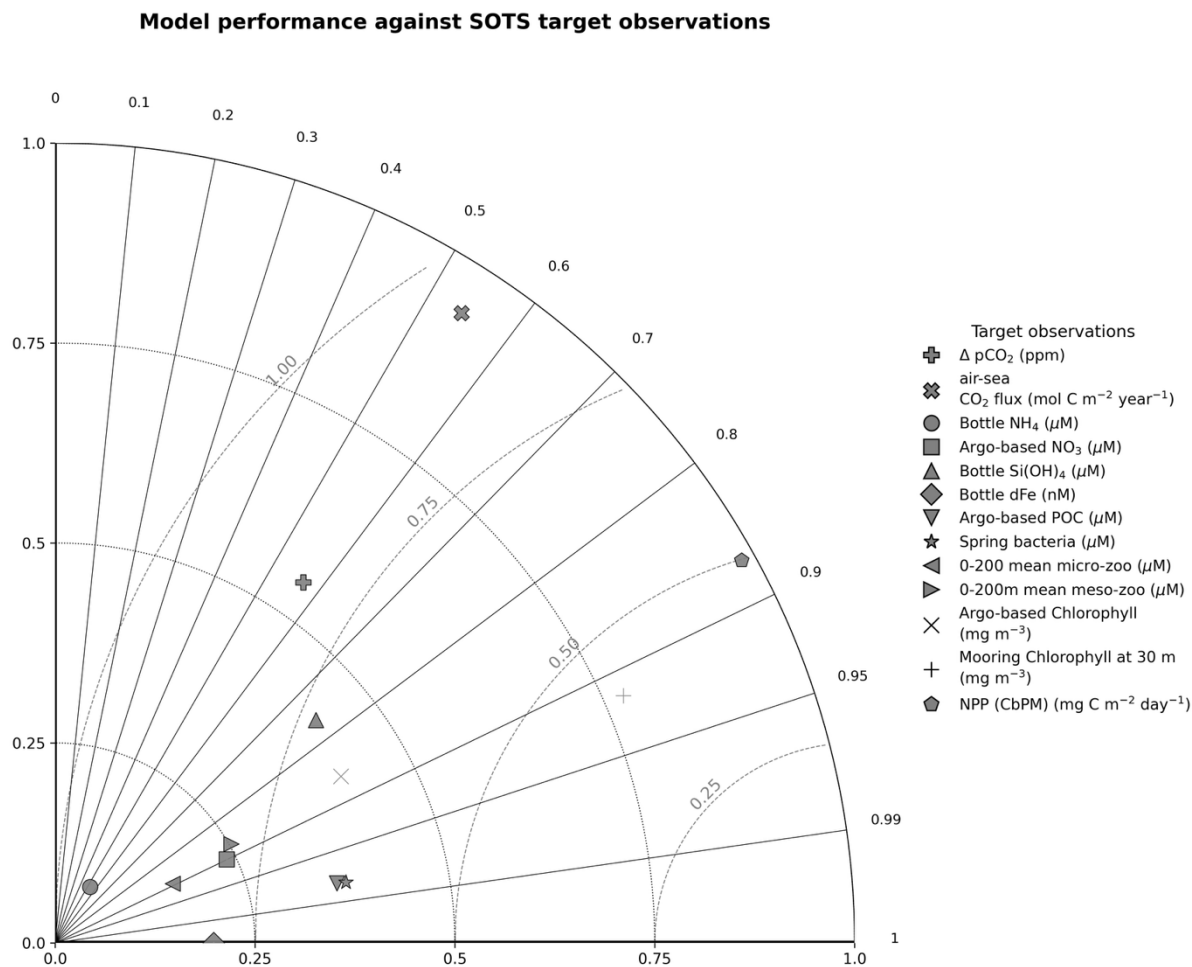


Figure S16. Taylor Diagram of WOMBAT-full performance against the observational target datasets for the optimal degree of partial oxidation ($\alpha = 0.30$). Model parameters used default values listed in Table S2.

References

1. Anderson, L. A.: On the hydrogen and oxygen content of marine phytoplankton, *Deep Sea Research Part I: Oceanographic Research Papers*, 42, 1675–1680, [https://doi.org/10.1016/0967-0637\(95\)00072-E](https://doi.org/10.1016/0967-0637(95)00072-E), 1995.
2. Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, *Limnol. Oceanogr.*, 42, 1–20, <https://doi.org/10.4319/lo.1997.42.1.0001>, 1997.
3. Behrenfeld, M. J., O'Malley, R. T., Siegel, D. a, McClain, C. R., Sarmiento, J. L., Feldman, G. C., Milligan, A. J., Falkowski, P. G., Letelier, R. M., and Boss, E. S.: Climate-driven trends in contemporary ocean productivity., *Nature*, 444, 752–755, <https://doi.org/10.1038/nature05317>, 2006.
4. Buchanan, P. J., Reddy, P. J., Matear, R. J., Chamberlain, M. A., Rohr, T., Squire, D., and Shadwick, E. H.: Optimization of the World Ocean Model of Biogeochemistry and Trophic dynamics (WOMBAT) using surrogate machine learning methods, *Biogeosciences*, 22, 5349–5385, <https://doi.org/10.5194/bg-22-5349-2025>, 2025.
5. Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S. C., Leblanc, K., Le Quéré, C., Luo, Y.-W., O'Brien, C., O'Brien, T., Peloquin, J., Schiebel, R., and Swan, C.: MAREDAT: towards a world atlas of MARine Ecosystem DATA, *Earth Syst. Sci. Data*, 5, 227–239, <https://doi.org/10.5194/essd-5-227-2013>, 2013.
6. Cai, R., Lechtenfeld, O. J., Tanentzap, A. J., Sachs, J. P., Kao, S., Osterholz, H., Yan, Z., Zhao, C., Yi, Y., Li, P., Zhang, C., Jiao, N., Koch, B. P., Herndl, G. J., and He, D.: Rapid microbial production of long-lived dissolved organic carbon in the global ocean, *Proceedings of the National Academy of Sciences*, 123, <https://doi.org/10.1073/pnas.2601044123>, 2026.
7. Deans, F., Décima, M., Gutiérrez-Rodríguez, A., Morales, S. E., Baltar, F., and Law, C. S.: Bacterioplankton abundance and heterotrophic production from the Chatham Rise, east of New Zealand, measured during the SalpPOOP (TAN1810) campaign onboard the RV Tangaroa, 2025.
8. Fogg, G. E.: The Ecological Significance of Extracellular Products of Phytoplankton Photosynthesis, *Botanica Marina*, 26, <https://doi.org/10.1515/botm.1983.26.1.3>, 1983.
9. Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J.: `emcee`: The MCMC Hammer, *Publications of the Astronomical Society of the Pacific*, 125, 306–312, <https://doi.org/10.1086/670067>, 2013.
10. Garcia, H. E., Wang, Z., Bouchard, C., Cross, S. L., Paver, C. R., Reagan, J. R., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Seidov, D., and Dukhovskoy, D.: World Ocean Atlas 2023, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, Dissolved Oxygen Saturation and 30-year Climate Normal, <https://doi.org/10.25923/rb67-ns53>, 2024a.
11. Garcia, H. E., Bouchard, C., Cross, S. L., Paver, C. R., Reagan, J. R., Boyer, T. P., Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Seidov, D., Wang, Z., and Dukhovskoy, D.: World Ocean Atlas 2023, Volume 4: Dissolved Inorganic Nutrients (Phosphate, Nitrate, and Silicate), <https://doi.org/10.25923/39qw-7j08>, 2024b.
12. Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, *Comm. App. Math. Comp. Sci.*, 5, 65–80, <https://doi.org/10.2140/camcos.2010.5.65>, 2010.
13. Hamilton, D. S., Scanza, R. A., Rathod, S. D., Bond, T. C., Kok, J. F., Li, L., Matsui, H., and Mahowald, N. M.: Recent (1980 to 2015) Trends and Variability in Daily-to-Interannual Soluble Iron Deposition from Dust, Fire, and Anthropogenic Sources, *Geophys. Res. Lett.*, 47, <https://doi.org/10.1029/2020GL089688>, 2020.

14. Hansell, D. A. and Carlson, C. A.: Biogeochemistry of marine dissolved organic matter, Academic Press, 2014.
15. Huang, Y., Tagliabue, A., and Cassar, N.: Data-Driven Modeling of Dissolved Iron in the Global Ocean, *Front. Mar. Sci.*, 9, <https://doi.org/10.3389/fmars.2022.837183>, 2022.
16. Issan, O., Riley, P., Camporeale, E., and Kramer, B.: Bayesian Inference and Global Sensitivity Analysis for Ambient Solar Wind Prediction, *Space Weather*, 21, <https://doi.org/10.1029/2023SW003555>, 2023.
17. Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnack, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the $1^\circ \times 1^\circ$ GLODAP version 2, *Earth Syst. Sci. Data*, 8, 325–340, <https://doi.org/10.5194/essd-8-325-2016>, 2016.
18. Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmitt-Kopplin, P., and Koch, B. P.: Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean, *Geochim. Cosmochim. Acta*, 126, 321–337, <https://doi.org/10.1016/j.gca.2013.11.009>, 2014.
19. Locarnini, R. A., Mishonov, A. V., Baranova, O. K., Reagan, J. R., Boyer, T. P., Seidov, D., Wang, Z., Garcia, H. E., Bouchard, C., Cross, S. L., Paver, C. R., and Dukhovskoy, D.: World Ocean Atlas 2023, Volume 1: Temperature, in: NOAA Atlas NESDIS 89, edited by: Mishonov, A., 52, 2024.
20. Norris, N., Levine, N. M., Fernandez, V. I., and Stocker, R.: Mechanistic model of nutrient uptake explains dichotomy between marine oligotrophic and copiotrophic bacteria, *PLoS Comput. Biol.*, 17, e1009023, <https://doi.org/10.1371/journal.pcbi.1009023>, 2021.
21. Reagan, J. R., Seidov, D., Wang, Z., Dukhovskoy, D., Boyer, T. P., Locarnini, R. A., Baranova, O. K., Mishonov, A. V., Garcia, H. E., Bouchard, C., Cross, S. L., and Paver, C. R.: World Ocean Atlas 2023, Volume 2: Salinity, in: NOAA Atlas NESDIS 90, edited by: Mishonov, A., 52, 2024.
22. Reddy, P. J., Chinta, S., Baki, H., Matear, R., and Taylor, J.: Gaussian Process Regression-Based Bayesian Optimisation (G-BO) of Model Parameters—A WRF Model Case Study of Southeast Australia Heat Extremes, *Geophys. Res. Lett.*, 51, <https://doi.org/10.1029/2024GL111074>, 2024a.
23. Reddy, P. J., Chinta, S., Matear, R., Taylor, J., Baki, H., Thatcher, M., Kala, J., and Sharples, J.: Machine learning based parameter sensitivity of regional climate models—a case study of the WRF model for heat extremes over Southeast Australia, *Environmental Research Letters*, 19, <https://doi.org/10.1088/1748-9326/ad0eb0>, 2024b.
24. Reichl, B. G. and Hallberg, R.: A simplified energetics based planetary boundary layer (ePBL) approach for ocean climate simulations., *Ocean Model. (Oxf.)*, 132, 112–129, <https://doi.org/10.1016/j.ocemod.2018.10.004>, 2018.
25. Reintjes, G., Arnosti, C., Fuchs, B., and Amann, R.: Selfish, sharing and scavenging bacteria in the Atlantic Ocean: a biogeographical study of bacterial substrate utilisation, *ISME J.*, 13, 1119–1132, <https://doi.org/10.1038/s41396-018-0326-3>, 2019.
26. Reynolds, R. C., Weiss, A. C. B., James, C. C., Kojima, C. Y., Weissman, J. L., Thrash, J. C., and Levine, N. M.: Defining metabolic niches for marine microbial heterotrophs, *Sci. Adv.*, 12, <https://doi.org/10.1126/sciadv.adz0537>, 2026.
27. Rittman, B. E. and McCarthy, P. L.: Environmental Biotechnology: Principles and Applications, McGraw-Hill Education, 2001.

28. Shadwick, E. H., Wynn-Edwards, C. A., Eriksen, R. S., Jansen, P., Yang, X., Woodward, G., and Davies, D.: The Southern Ocean Time Series: a climatological view of hydrography, biogeochemistry, phytoplankton community composition, and carbon export in the Subantarctic Zone, *Ocean Science*, 21, 1549–1573, <https://doi.org/10.5194/os-21-1549-2025>, 2025.
29. Sobol', I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Math. Comput. Simul.*, 55, 271–280, [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6), 2001.
30. Stephens, B. M., Stincone, P., Petras, D., English, C. J., Opalk, K., Giovannoni, S., and Carlson, C. A.: Oxidation state of bioavailable dissolved organic matter influences bacterioplankton respiration and growth efficiency, *Commun. Biol.*, 8, 145, <https://doi.org/10.1038/s42003-025-07574-2>, 2025.
31. Stewart, K. D., Kim, W. M., Urakawa, S., Hogg, A. McC., Yeager, S., Tsujino, H., Nakano, H., Kiss, A. E., and Danabasoglu, G.: JRA55-do-based repeat year forcing datasets for driving ocean–sea-ice models, *Ocean Model. (Oxf.)*, 147, 101557, <https://doi.org/10.1016/j.ocemod.2019.101557>, 2020.
32. Traill, C. D., Rohr, T., Shadwick, E., Schallenberg, C., Ellwood, M., and Bowie, A.: Coupling Between the Subantarctic Seasonal Iron Cycle and Productivity at the Southern Ocean Time Series (SOTS), *AGU Advances*, 6, <https://doi.org/10.1029/2024AV001599>, 2025.
33. Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber, J. L., Bentsen, M., Böning, C. W., Bozec, A., Chassignet, E. P., Curchitser, E., Boeira Dias, F., Durack, P. J., Griffies, S. M., Harada, Y., Ilicak, M., Josey, S. A., Kobayashi, C., Kobayashi, S., Komuro, Y., Large, W. G., Le Sommer, J., Marsland, S. J., Masina, S., Scheinert, M., Tomita, H., Valdivieso, M., and Yamazaki, D.: JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do), *Ocean Model. (Oxf.)*, 130, 79–139, <https://doi.org/10.1016/j.ocemod.2018.07.002>, 2018.
34. White, A. E., Giovannoni, S. J., Zhao, Y., Vergin, K., and Carlson, C. A.: Elemental content and stoichiometry of SAR11 chemoheterotrophic marine bacteria, *Limnol. Oceanogr. Lett.*, 4, 44–51, <https://doi.org/10.1002/lol2.10103>, 2019.
35. Zakem, E. J., Mahadevan, A., Lauderdale, J. M., and Follows, M. J.: Stable aerobic and anaerobic coexistence in anoxic marine zones, *ISME J.*, 14, 288–301, <https://doi.org/10.1038/s41396-019-0523-8>, 2020.
36. Zakem, E. J., McNichol, J., Weissman, J. L., Raut, Y., Xu, L., Halewood, E. R., Carlson, C. A., Dutkiewicz, S., Fuhrman, J. A., and Levine, N. M.: Functional biogeography of marine microbial heterotrophs, *Science* (1979)., 388, <https://doi.org/10.1126/science.ado5323>, 2025.
37. Zimmerman, A. E., Allison, S. D., and Martiny, A. C.: Phylogenetic constraints on elemental stoichiometry and resource allocation in heterotrophic marine bacteria, *Environ. Microbiol.*, 16, 1398–1410, <https://doi.org/10.1111/1462-2920.12329>, 2014.